



Published in final edited form as:

*Biochemistry*. 2017 April 11; 56(14): 2010–2023. doi:10.1021/acs.biochem.7b00137.

## Exploring the Influence of Domain Architecture on the Catalytic Function of Diterpene Synthases

Travis A. Pemberton<sup>†,¶</sup>, Mengbin Chen<sup>†,¶</sup>, Golda G. Harris<sup>†</sup>, Wayne K. W. Chou<sup>‡</sup>, Lian Duan<sup>‡</sup>, Mustafa Köksal<sup>†</sup>, Alex S Genshaft<sup>†</sup>, David E. Cane<sup>‡</sup>, and David W. Christianson<sup>†,\*</sup>

<sup>†</sup>Roy and Diana Vagelos Laboratories, Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6323, United States

<sup>‡</sup>Department of Chemistry, Brown University, Box H, Providence, Rhode Island 02912, United States

### Abstract

Terpenoid synthases catalyze isoprenoid cyclization reactions underlying the generation of more than 80,000 natural products. Such dramatic chemodiversity belies the fact that these enzymes generally consist of only three domain folds designated  $\alpha$ ,  $\beta$ , and  $\gamma$ . Catalysis by class I terpenoid synthases occurs exclusively in the  $\alpha$  domain, which is found with  $\alpha$ ,  $\alpha\alpha$ ,  $\alpha\beta$ , and  $\alpha\beta\gamma$  domain architectures. Here, we explore the influence of domain architecture on catalysis by taxadiene synthase from *Taxus brevifolia* (TbTS,  $\alpha\beta\gamma$ ), fusicoccadiene synthase from *Phomopsis amygdali* (PaFS,  $(\alpha\alpha)_6$ ), and ophiobolin F synthase from *Aspergillus clavatus* (AcOS,  $\alpha\alpha$ ). We show that the cyclization fidelity and catalytic efficiency of the  $\alpha$  domain of TbTS are severely compromised by deletion of the  $\beta\gamma$  domains; however, retention of the  $\beta$  domain preserves significant cyclization fidelity. In PaFS, we previously demonstrated that one  $\alpha$  domain similarly influences catalysis by the other  $\alpha$  domain [Chen, M., Chou, W. K. W., Toyomasu, T., Cane, D. E., Christianson, D. W. (2016) *ACS Chem. Biol.* 11, 889–899]. Here, we show that the hexameric quaternary structure of PaFS enables cluster channeling. We also show that the  $\alpha$  domains of PaFS and AcOS can be swapped so as to make functional chimeric  $\alpha\alpha$  synthases. Notably, both cyclization fidelity and catalytic efficiency are altered in all chimeric synthases. Twelve newly formed and uncharacterized C<sub>20</sub> diterpene products and three C<sub>25</sub> sesterterpene products are generated by these chimeras. Thus, engineered  $\alpha\beta\gamma$  and  $\alpha\alpha$  terpenoid cyclases promise to generate chemodiversity in the greater family of terpenoid natural products.

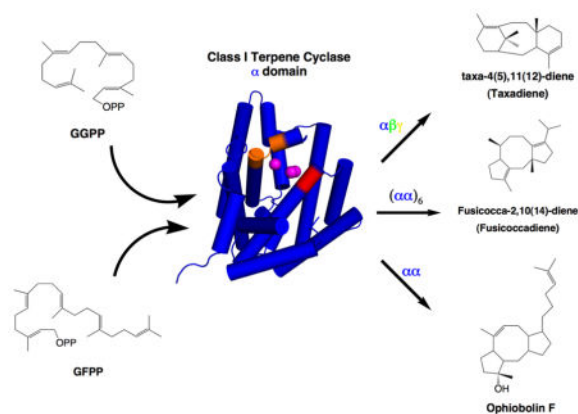
### Graphical abstract

\*To whom correspondence should be addressed: Roy and Diana Vagelos Laboratories, Department of Chemistry, University of Pennsylvania, 231 South 34<sup>th</sup> Street, Philadelphia, PA, 19104-6323. Tel: 215-898-5714; chris@sas.upenn.edu.

<sup>¶</sup>Contributed equally to this work.

#### Supporting Information

Supporting Information Available: Table S1, primers used for mutagenesis; Figure S1, detailed cyclization mechanisms of taxadiene synthase, fusicoccadiene synthase, and ophiobolin F synthase; Figure S2, mass spectra of diterpene products generated by TbTS mutants; Figure S3, mass spectra of sesquiterpene, diterpene, and sesterterpene products generated by PaFS and AcOS chimeras. This material is available free of charge via the Internet.



## Introduction

More than 80,000 terpenoids have been identified to date, thereby comprising the largest and most diverse family of natural products. Such molecular diversity fosters utility for humankind, since many terpenoids have been widely used as fragrances, insect attractants or repellents, and medicines since times of antiquity.<sup>1–5</sup> The molecular complexity of the terpenome belies simple origins in the  $C_5$  building blocks isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). Prenyl diphosphate synthases catalyze the condensation of one molecule of DMAPP and one or more molecules of IPP in head-to-tail fashion to form  $C_{10}$ ,  $C_{15}$ ,  $C_{20}$ , and  $C_{25}$  prenyl diphosphates.<sup>6</sup> Terpene cyclases then convert these linear achiral allylic diphosphates into polycyclic terpene hydrocarbons and alcohols through multi-step cyclization cascades.<sup>7–13</sup> High-energy carbocation intermediates are hallmarks of these reaction cascades, which cyclases manipulate to direct carbon-carbon bond formation along a wide range of specific pathways. The extraordinary structural diversity of terpenoid natural products arises in large part from the extraordinary functional diversity of terpene cyclases, which are capable of generating myriad carbon skeletons from a single acyclic prenyl diphosphate precursor.

The first crystal structures of terpenoid cyclases were reported twenty years ago.<sup>14–16</sup> These structures have guided the exploration of structure-function relationships for both class I cyclases, which initiate carbocation formation through the metal-dependent ionization of the substrate diphosphate group, and class II cyclases, which initiate carbocation formation through the protonation of an isoprenoid carbon-carbon double bond. Diterpene cyclases, which utilize the acyclic  $C_{20}$  substrate geranylgeranyl diphosphate (GGPP), are unique among the greater family of terpenoid cyclases in that both class I and class II enzymes are known. The first crystal structures of class I and class II plant diterpene cyclases revealed a three-domain architecture designated  $\alpha\beta\gamma$ . Class I cyclization reactions are catalyzed exclusively in the  $\alpha$  domain, while class II cyclizations are catalyzed at the interface of the  $\beta$  and  $\gamma$  domains.<sup>17–19</sup> Although class I terpene synthases are also found as single  $\alpha$ -domain enzymes, such as the bacterial diterpene cyclase cyclooctat-9-en-7-ol synthase (CotB2),<sup>20</sup> as well as  $\alpha\beta$  domain assemblies exemplified by the plant sesquiterpene cyclase 5-*epi*-aristolochene synthase<sup>15</sup> and the plant monoterpene cyclase bornyl diphosphate synthase,<sup>21</sup> the occurrence of more complex multi-domain assemblies in terpenoid cyclases raises

intriguing questions regarding the relationship between domain architecture and enzyme function. For example, consider taxadiene synthase from *Taxus brevifolia* (TbTS), a class I diterpene cyclase that catalyzes the first committed step in the biosynthesis of the blockbuster cancer chemotherapeutic drug Taxol.<sup>22–24</sup> The crystal structure of TbTS reveals an  $\alpha\beta\gamma$  domain architecture in which the catalytic function is localized solely in the  $\alpha$  domain (Figure 1, Figure S1).<sup>17</sup> How, then, does the adjacent  $\beta$  domain, or the distant  $\gamma$  domain, influence catalysis in the  $\alpha$  domain of this cyclase?

In a contrasting example, consider fusicoccadiene synthase from *Phomopsis amygdali* (PaFS). This bifunctional diterpene synthase adopts  $\alpha\alpha$  domain architecture in which the C-terminal  $\alpha$  domain generates GGPP and the N-terminal  $\alpha$  domain cyclizes GGPP to form fusicoccadiene.<sup>25,26</sup> The tricyclic ring skeleton of fusicoccadiene is critical for the anti-tumor activity of fusicoccin A, a diterpene glycoside that locks 14-3-3 proteins in place with their binding partners.<sup>27–29</sup> The recently determined crystal structures of the individual chain elongation and cyclase  $\alpha$  domains of PaFS reveal that each utilizes three  $Mg^{2+}$  ions to initiate catalysis.<sup>30</sup> Analytical ultracentrifugation and low-angle X-ray scattering experiments indicate that PaFS is a hexamer (Figure 1, Figure S1),<sup>30</sup> which raises the possibility of proximity channeling<sup>31</sup> between active sites. Proximity channeling may enhance product flux in a biosynthetic pathway even in the absence of a direct tunnel between enzyme active sites, since the product generated in one active site need not completely diffuse into solution and re-diffuse back into the second active site if the two active sites are sufficiently close. The clustering of active sites in such a hexameric assembly may enhance this effect since the product generated in one active site can re-diffuse back into one of multiple second active sites that are nearby as a result of the quaternary structure of the oligomer.<sup>32</sup> Thus, multi-domain assembly and oligomerization may enhance the efficiency of cyclic terpene biosynthesis through cluster channeling of the acyclic substrate.<sup>13</sup>

An  $\alpha\alpha$  domain architecture comparable to that of PaFS is found in ophiobolin F synthase from *Aspergillus clavatus* (AcOS), a bifunctional sesterterpene synthase that first generates  $C_{25}$  geranylarnesyl diphosphate (GFPP) in the C-terminal  $\alpha$  domain and then cyclizes GFPP in the N-terminal  $\alpha$  domain to form ophiobolin F.<sup>33</sup> This tricyclic sesterterpene is characterized by a 5-8-5 ring skeleton similar to that of fusicoccadiene (Figure 1, Figure S1). Although the crystal structure of AcOS has not yet been determined, PaFS and AcOS share 41% amino acid sequence identity and catalyze GGPP and GFPP cyclization reactions, respectively, that proceed through a common 5–11 ring system intermediate.<sup>25,26,33,34</sup> This mechanistic similarity suggests structural similarity between the active site contours of PaFS and AcOS. Pairwise comparisons of active site residues between the two cyclases reveal that while most residues are conserved or isosteric, only two residue pairs substantially differ in size: W225 and V228 in PaFS correspond to L217 and A220 in AcOS, respectively.

Here, we probe the influence of domain architecture on catalytic function in  $\alpha\beta\gamma$  and  $\alpha\alpha$  diterpene cyclases. We demonstrate that the vestigial  $\beta\gamma$  domains of TbTS influence cyclization fidelity and catalytic efficiency in the  $\alpha$  domain of this  $\alpha\beta\gamma$  cyclase. In PaFS, we demonstrate that  $\alpha\alpha$  domain architecture of the wild-type hexamer enables cluster channeling of the acyclic substrate to the cyclase domain. We also show that the  $\alpha$  domains

of PaFS and AcOS can be swapped to make functional chimeric  $\alpha\alpha$  diterpene synthases, while the cyclase domains of PaFS and AcOS can be engineered with just two amino acid substitutions so as to swap substrate specificity for  $C_{20}$  and  $C_{25}$  substrates. Cyclization fidelity and catalytic efficiency are altered in all chimeric synthases studied, however, indicating that domain-domain interactions play a significant role in stabilizing the cyclase domains in catalytically competent conformations.

## Materials and Methods

### Expression, and purification of TbTS variants

We previously described the pET22bTV plasmid encoding a catalytically-active<sup>24</sup> pseudomature TbTS truncation variant lacking the N-terminal plastidial targeting sequence, initially designated “M79-TXS” by Köksal and colleagues.<sup>17</sup> Here, this variant is referred to as TbTS. This vector was used to prepare all TbTS variants described herein as summarized in Figure 2. The  $\gamma$  domain of TbTS (S136-Y348) is an insertion domain between the first and second helix of the  $\beta$  domain. The  $\gamma$  domain was spliced out of the full-length protein using assembly PCR to yield an  $\alpha\beta$  construct designated “TbTS $\alpha\beta$ ” (Figure 2; primers are listed in Table S1). The N-terminal segment M79-I135 includes the first helix of the  $\beta$  domain and remained intact in this construct; this segment is believed to help cap the active site upon substrate binding, as observed in 5-*epi*-aristolochene synthase and bornyl diphosphate synthase.<sup>15,21</sup> Flanked by NdeI and BamHI restriction sites, the gene sequence encoding TbTS $\alpha\beta$  was ligated into plasmid pET22bTV and transformed into *Escherichia coli* DH5 $\alpha$  cells. The correct gene sequence was confirmed by DNA sequencing at the University of Pennsylvania Perelman School of Medicine Sequencing Facility. The TbTS $\alpha\beta$  protein was expressed but suffered from intractable aggregation problems, so the construct was modified so as to insert a single glycine residue at the splice site. Guided by molecular modeling, it was hypothesized that this additional residue would improve the geometry of the loop connecting the first and second helix of the  $\beta$  domain. This modified construct, designated TbTS $\alpha\beta_G$ , was prepared using site-directed insertion through PCR amplification (Figure 2; primers are listed in Table S1). The resulting PCR product was digested with DpnI for 1 h then transformed into *E. coli* DH5 $\alpha$  cells. The correct gene sequence was confirmed by sequencing.

The native forms of many class I plant terpene synthases are similarly found as  $\alpha\beta$  domain assemblies. For example, isoprene synthase from *Populus x canescens* (PcIS) has an active site in the  $\alpha$  domain and an N-terminal  $\beta$  domain with a similar N-terminal capping segment.<sup>35</sup> It is possible that removal of the  $\gamma$  domain exposes a large hydrophobic patch (the former domain-domain interface) on the  $\beta$  domain of TbTS. To probe the effects of a modified  $\beta$  domain “pre-evolved” to have a hydrophilic solvent-exposed surface in the absence of a  $\gamma$  domain, the  $\alpha$  domain (S553-V862) and 120 residues from the  $\beta$  domain (L432-Q552) of TbTS were spliced together with the complementary portion of the  $\beta$ -domain of PcIS (M1-F151) using assembly PCR to result in the 581-residue chimera designated TbTS $\alpha\beta'$  (Figure 2). PCR primers are listed in Table S1.

Finally, a minimalist truncation variant, designated TbTS $\alpha$  (Figure 2), was prepared by completely splicing out the  $\beta$  and  $\gamma$  domains and connecting the N-terminal segment M79-

I135 directly to the N-terminus of the first helix of the  $\alpha$  domain, starting at N537. This was achieved using assembly PCR (primers are listed in Table S1). Flanked by NdeI and BamHI restriction sites, the gene sequence encoding TbTS $\alpha$  was ligated into plasmid pET22bTV and transformed into *E. coli* DH5 $\alpha$  cells. The correct gene sequence was confirmed by DNA sequencing.

TbTS $\alpha$ , TbTS $\alpha\beta_G$ , and TbTS $\alpha\beta'$  were transformed into *E. coli* BL21 (DE3) RIL-CodonPlus (Novagen) cells and grown on LB-Agar plates with 50  $\mu$ g/mL ampicillin and 34  $\mu$ g/mL chloramphenicol overnight at 37 °C. For the preparation of each construct, individual colonies were transferred to two 15-mL LB broth cultures with 50  $\mu$ g/mL ampicillin and 34  $\mu$ g/mL chloramphenicol and allowed to grow at 37 °C with aeration for 12–16 h. A total of 6  $\times$  1 L LB Broth cultures containing 50  $\mu$ g/mL ampicillin and 34  $\mu$ g/mL chloramphenicol were inoculated with 5 mL of the cultures grown overnight and allowed to shake at 37 °C until cells reached OD<sub>600</sub> = 1.0. Cells were induced with 500  $\mu$ M isopropyl- $\beta$ -D-1-thiogalactopyranoside (IPTG) and the temperature was reduced to 16 °C overnight. Cells were pelleted at 5400g for 20 min. The pellet was harvested, flash frozen in liquid nitrogen and stored at –20 °C.

Cells with TbTS $\alpha\beta_G$  were thawed and resuspended in buffer A [20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) (pH 8.2), 300 mM NaCl, 1 mM tris-(2-carboxyethyl)phosphine (TCEP), 10% glycerol]. Cells were lysed by sonication and spun at 37000g for 1 h. Supernatant was loaded onto a His60 Superflow Resin (Clontech) column using buffer A. The column was washed with 30 mM imidazole and then 300 mM imidazole. The 300 mM fractions were pooled and dialyzed overnight at 4 °C in buffer B [50 mM Tris (pH 7.5), 1 mM TCEP, 10% glycerol]. Pooled fractions were loaded onto a 5-mL Hi-Trap Q HP column (GE Healthcare). The enzyme was eluted using a gradient of 0.0–1.0 M NaCl. Fractions were analyzed using SDS-PAGE, and fractions containing protein of the desired molecular weight were again pooled and dialyzed overnight in buffer C [50 mM Tris (pH 7.5), 300 mM NaCl, 1 mM TCEP, 10% glycerol]. TbTS $\alpha\beta'$  was expressed and purified in a similar fashion.

Cells with TbTS $\alpha$  were thawed and resuspended in buffer D [50 mM Tris (pH 7.5), 300 mM NaCl, 1 mM TCEP, 1% Tween20]. Cells were lysed by sonication and then spun at 37000g for 1 h. The supernatant was discarded. The pellet was resuspended in buffer D to a final volume of 50 mL. The pH of the resuspended pellet was gradually increased to 12. The pellet was then allowed to mix at 4 °C for 30 min. The pH was then gradually decreased to 7.5 and allowed to mix at 4 °C for 30 min. This mixture was then spun at 37000g for 1 h. Supernatant was loaded onto a Ni-NTA column using buffer D and purified as described for TbTS $\alpha\beta_G$ . Enzymes were concentrated using Amicon centrifuge tubes with 10-kDa pore size for TbTS $\alpha$  and 30-kDa pore size for TbTS $\alpha\beta_G$  and TbTS $\alpha\beta'$ . Enzymes were concentrated to approximately 8 mg/mL as determined by absorbance at 280 nm.

Enzyme assays were performed as described previously for TbTS and other terpenoid cyclases.<sup>23,36</sup> Briefly, enzymes were solubilized in buffer E [30 mM HEPES (pH 8.2), 1 mM MgCl<sub>2</sub>, 1 mM TCEP, 10% glycerol]. Assays were performed in triplicate using a range of 0–15  $\mu$ M [1-<sup>3</sup>H]GGPP (100 mCi/mmol, American Radiolabeled Chemicals, Inc.) for full-

length TbTS, 0–3  $\mu\text{M}$  [ $1\text{-}^3\text{H}$ ]GGPP (100 mCi/mmol) for TbTS $\alpha\beta\text{G}$ , and 0–25  $\mu\text{M}$  [ $1\text{-}^3\text{H}$ ]GGPP (100 mCi/mmol) for TbTS $\alpha$ . The final concentration of each construct used for assays was determined to be that at which the dependence of product formation on enzyme concentration was linear and the level of substrate turnover was less than 10%. A 500- $\mu\text{L}$  reaction mixture was prepared and allowed to react for 10 min at room temperature before being quenched with 20  $\mu\text{L}$  of 500 mM ethylenediaminetetraacetic acid (EDTA) at pH 8.0. The extraction of diterpene products was achieved by adding 1 mL of hexane to the reaction mixture and vortexing for 20 s. The hexane layer was removed and passed over a silica gel column (1 cm of silica gel 230–400 mesh, Grade 60, in a Pasteur pipette) into a scintillation vial containing 5 mL of scintillation solution (Ecoscint O, National Diagnostics) and the process was repeated two more times. A Beckman scintillation counter (LS6500) was used to measure total product formation. Data for the rate of product formation as a function of substrate concentration were fit by nonlinear regression using GraphPad Prism software to determine steady-state kinetic parameters  $k_{\text{cat}}$  and  $K_{\text{M}}$  based on the known total enzyme concentration.

For product analysis, reaction mixtures were set up with a total volume of 5 mL of buffer E containing 1.7  $\mu\text{M}$  enzyme and 100  $\mu\text{M}$  GGPP. The reaction mixture was overlaid with 3 mL of *n*-pentane and allowed to react for approximately 20 h at room temperature. The reaction mixture was extracted three times with *n*-pentane; pentane extracts were dried with anhydrous  $\text{MgSO}_4$ , filtered, and concentrated on an ice/water mixture under reduced pressure until the volume reached 100  $\mu\text{L}$ . Extracts were analyzed using an Agilent 5890 Series II mass spectrometer with a 30 m  $\times$  0.25 mm HP5MS capillary column in EI+ mode using a temperature program with an initial hold at 60  $^{\circ}\text{C}$  for 2 min, a 20  $^{\circ}\text{C}/\text{min}$  temperature gradient from 60 to 280  $^{\circ}\text{C}$ , a final hold at 280  $^{\circ}\text{C}$  for 2 min, and a solvent delay of 3 min. Analysis of the organic extracts resulting from the incubation of GGPP with the mutant cyclases by gas chromatography and mass spectrometry (GC-MS) revealed the formation of mixtures of diterpene hydrocarbons at  $m/z$  272. Compounds were identified by comparison of their individual mass spectra and chromatographic retention indices with those of authentic compounds using the MassFinder 4.0 Program and Database and the NIST Mass Spectral Library.<sup>37</sup>

### Expression and purification of PaFS and AcOS chimeras

The plasmid for PaFS was prepared as previously described.<sup>30</sup> Plasmids for AcOS and stellatatriene synthase (from which the GFPP synthase  $\alpha$  domain was utilized) were synthesized and codon-optimized by Genscript. The AcOS gene was on a pET28a vector, using NdeI and BamHI as restriction sites. For cloning chimeric enzymes, a pET His6 TEV LIC cloning vector (a generous gift from Dr. Scott Gradia, Addgene ID: 29653) with an N-terminal hexahistidine tag was used unless otherwise specified. Primers used for cloning and site-directed mutagenesis are listed in Table S1. DNA sequencing was performed at the DNA Sequencing Facility at the Perelman School of Medicine, University of Pennsylvania.

For expression of AcOS, *E. coli* BL21(DE3) cells (Novagen) were transformed with the AcOS plasmid and grown on an LB-Agar plate with 50  $\mu\text{g}/\text{mL}$  kanamycin overnight at 37  $^{\circ}\text{C}$ . A single colony was transferred to a 5-mL LB culture with 50  $\mu\text{g}/\text{mL}$  kanamycin,

which grew at 37 °C with aeration for 8–12 h. A total of 6 × 1 L 2XYT culture containing 50 µg/mL kanamycin was inoculated with 6 × 5 mL LB culture, and was shaken at 37 °C until OD<sub>600</sub> reached 1.0. Protein expression was induced at 16 °C by adding IPTG to a final concentration of 50 µM. After expression for 18–22 h, cells were pelleted via centrifuge (5400g for 15 min) and stored at –80 °C.

All four chimera protein constructs (Figure 2) were expressed in *E. coli* BL21(DE3) RIL-CodonPlus strain (Novagen). A total of 6 × 1 L 2XYT medium containing 50 µg/mL kanamycin and 34 µg/mL chloramphenicol was inoculated with 5 mL starting culture and grown at 37 °C with aeration. Cells were induced with 50 µM IPTG at 16 °C when OD<sub>600</sub> reached 0.8–1.0. After expression for 18 h, cells were pelleted by centrifugation (5400g for 15 min).

For each construct, the cell pellet was resuspended in Ni-NTA (QIAGEN) wash buffer [50 mM K<sub>2</sub>HPO<sub>4</sub> (pH 7.7), 300 mM NaCl, 10% (v/v) glycerol, 3 mM TCEP]. Cells were lysed by sonication for 10 min. Cell debris was discarded after centrifugation, and supernatant was loaded on a Ni-NTA column (QIAGEN). Protein was eluted with a 0–400 mM imidazole gradient in Ni-NTA wash buffer. Protein fractions were combined, concentrated, exchanged with buffer Q [10 mM Tris (pH 8.0), 10 mM NaCl] and loaded on a 5-mL HiTrap Q HP column (GE Healthcare). Elution was performed with a 0–500 mM NaCl gradient in buffer Q. Fractions were analyzed by SDS-PAGE, combined, concentrated, and loaded onto a 26/60 Superdex 200 size-exclusion column (GE Healthcare) pre-equilibrated with 10 mM Tris (pH 7.9), 200 mM NaCl, 5% (v/v) glycerol, and 3 mM TCEP. Fractions of the expected molecular weight were combined and concentrated.

Diterpene synthase assays on a 100-µL scale were conducted in triplicate, in similar fashion as described above for TbTS assays, using 0.16 nM AcOS, 0.43 nM AcOS $\alpha$ -PaFS $\alpha$ , or 42 nM AcOS $\alpha$ '-PaFS $\alpha$  in assay buffer [25 mM HEPES (pH 8.0), 200 mM NaCl, 5 mM MgCl<sub>2</sub>]. Concentration ranges of [1-<sup>3</sup>H]-GGPP were adjusted accordingly. Reactions were initiated by adding enzyme at 30 °C, and quenched with 10 µL 0.5 M EDTA. A 1-mL aliquot of hexane was used to extract hydrocarbon products twice with vigorous vortexing and subsequently loaded onto a 1-mL silica gel column (1 cm of silica gel 230–400 mesh, Grade 60, in a Pasteur pipette). The column was washed with 1 mL hexane and 1 mL ethyl ether. Eluate was collected into a scintillation vial containing 5 mL of scintillation solution (Ecoscint O, National Diagnostics), and radioactivity counts were determined using a Beckman liquid scintillation counter (LS6500). Control experiments were performed to determine background counts, and data were processed with GraphPad Prism software.

Product assays were carried out in duplicate on a 5-mL scale in similar fashion to that described above for TbTS assays. To calibrate peak integrals, nitrobenzene (NBZ) was added to each reaction mixture to serve as internal standard. Reactions were initiated at 30 °C by adding enzyme into assay buffer containing substrate (DMAPP and IPP) overlaid with 5 mL *n*-pentane. After quenching by EDTA after 16 h of reaction, hydrocarbon products were extracted from the reaction mixture with three 5-mL portions of *n*-pentane and the extracts were run through anhydrous MgSO<sub>4</sub> to eliminate trace amounts of water. Samples were then concentrated under vacuum in an ice water bath using a rotary evaporator

and analyzed by GC-MS. Mixtures of diterpene hydrocarbons were detected at  $m/z$  272, some of which were directly identified by comparison of their individual mass spectra and chromatographic retention indices with those of authentic compounds using the MassFinder 4.0 Program and Database.<sup>37</sup> Sesterterpene hydrocarbons were detected at  $m/z$  340.

### PaFS-TbTS substrate competition assays

PaFS and TbTS were expressed and purified as described above. Monomer concentrations of 1  $\mu$ M PaFS and 1  $\mu$ M TbTS in 6 mL of assay buffer [25 mM HEPES pH 8.0, 200 mM NaCl, 5 mM MgCl<sub>2</sub>] were co-incubated with the following substrates in two separate experiments: (1) 50  $\mu$ M DMAPP + 150  $\mu$ M IPP, and (2) 50  $\mu$ M GGPP. The samples were overlaid with 5 mL HPLC grade *n*-pentane in a glass test tube and incubated at 30 °C for 16 h. The reaction mixture was extracted four times with 5 mL *n*-pentane (total 20 mL *n*-pentane extracts). The pentane extracts were dried with MgSO<sub>4</sub>, filtered, and concentrated in an ice/water bath under reduced pressure to approximately 100  $\mu$ L. Reaction mixture extracts were analyzed with an Agilent mass spectrometer with a 30 m  $\times$  0.25 mm HP5MS capillary column using a temperature program with an initial 2 min hold at 60 °C, a 20 °C/min temperature gradient from 60 °C to 280 °C, and a final 2 min hold at 280 °C. Analysis of the pentane extracts resulting from co-incubation of PaFS and TbTS with various substrates by GC-MS showed the formation of diterpene products. Compounds were identified by comparison of their mass spectra and GC retention indices with those of authentic compounds using the MassFinder 4.0 Program and Database.<sup>37</sup> Since the catalytic fidelities of PaFS and TbTS are comparable for the generation of their major products (PaFS generates 64–72% fusicoccadiene<sup>30</sup> and TbTS generates 58–82% taxadiene (Table 1)), the fusicoccadiene:taxadiene ratio measured in the GGPP competition assay provides an accurate reflection of the ratio of steady-state kinetic parameters ( $k_{cat}/K_M$ )<sub>PaFS</sub>/ $(k_{cat}/K_M)$ <sub>TbTS</sub> for GGPP utilization by individual PaFS and TbTS enzymes, even though these parameters reflect the generation of all diterpene products and not just the major product of each enzyme.

## Results

### TbTS truncation and chimera variants

Truncated domain constructs were prepared to study the effects of the  $\beta$  and  $\gamma$  domains on the class I GGPP cyclase activity that resides solely in the  $\alpha$  domain of full-length  $\alpha\beta\gamma$  TbTS. The  $\gamma$  domain was spliced out and replaced by a single glycine residue in the di-domain construct TbTS $\alpha\beta_G$ , and both  $\beta$  and  $\gamma$  domains were spliced out to yield the single-domain construct TbTS $\alpha$ . These constructs retained the N-terminal segment M79-I135 of the full-length pseudomature enzyme, since this segment is believed to help cap the active site during catalysis. In TbTS $\alpha\beta'$ , approximately half of the  $\beta$  domain of TbTS $\alpha\beta_G$  was replaced by the  $\beta$  domain of PcIS. Since deletion of the  $\gamma$  domain in TbTS $\alpha\beta_G$  exposes nonpolar surface area on the  $\beta$  domain at the former  $\beta\gamma$  domain interface, the utilization of a related  $\beta$  domain already evolved to be solvent-exposed at this region in TbTS $\alpha\beta'$  might potentially be beneficial. The primary structure of each engineered construct is schematically illustrated in Figure 2.



The double-domain construct TbTS $\alpha\beta_G$  was obtained in very poor yield after purification (0.5–1 mg/L culture, a level approximately 10-fold lower than that of wild-type TbTS), but sufficient quantities of pure protein could be prepared for kinetic and product array assays. Michaelis-Menten kinetics indicated a 2750-fold loss of catalytic efficiency ( $k_{cat}/K_M$ ) for TbTS $\alpha\beta_G$  relative to full-length TbTS (Figure 3). Thus, an intact  $\gamma$  domain is required for 99.9% of the maximum catalytic activity in the  $\alpha$  domain of TbTS, even though the  $\gamma$  domain makes no direct contact with the  $\alpha$  domain. Comparisons of the diterpene product arrays of full-length TbTS and TbTS $\alpha\beta_G$  using GC-MS are similarly informative. Full-length TbTS generates taxadiene as its major cyclization product (58–82%), along with 9.9–15.0% taxadiene isomers taxa-3(4)11(12)-diene and taxa-4(20),11(12)-diene, and 8.0–26.8% diterpene side products; interestingly, some run-to-run variations are observed (Table 1; mass spectra of diterpene products are reported in Figure S2). TbTS $\alpha\beta_G$  still generates taxadiene as its major product, although cyclization fidelity is somewhat compromised since the major cyclization products are 25.9% taxadiene and 4.3% taxadiene isomers. This di-domain construct also generates a mixture of the linear solvolysis product geranylinalool (21.0%) and the elimination product geranylmyrcene (20.3%), in addition to small amounts of other unidentified diterpenes. Comparisons of the fragmentation patterns in mass spectra of unidentified diterpenes with those of known cyclic and linear diterpenes suggest that unknown **4** is a cyclic diterpene, whereas unknowns **3** and **5** are linear diterpenes (Figure S2). With the formation of taxadiene thus compromised only 3-fold while the  $k_{cat}/K_M$  is decreased by 2750-fold, cyclization fidelity appears to be less susceptible to the deletion of the  $\gamma$  domain than is catalytic efficiency in TbTS $\alpha\beta_G$ .

To ascertain whether the substitution of part of an alternative  $\beta$  domain for the  $\beta\gamma$  domains of TbTS would support catalysis, the N-terminal  $\beta$  domain of PclS was substituted for the  $\beta$  domain of TbTS $\alpha\beta_G$  to generate the chimera TbTS $\alpha\beta'$ . This chimera was obtained in slightly better yield (approximately 1–2 mg/L culture) compared with TbTS $\alpha\beta_G$ . TbTS $\alpha\beta'$  retains the ability to generate taxadiene as its major cyclization product (37.6%) and one taxadiene isomer (4.9%). This chimera also generates the linear solvolysis product geranylinalool (16.0%) and elimination product geranylmyrcene (12.6%), as well as small amounts of additional unidentified diterpenes almost identical in composition to the unknown diterpenes generated by TbTS $\alpha\beta_G$  (Table 1). Again, comparisons of the fragmentation patterns in mass spectra of unidentified diterpenes with those of known cyclic and linear diterpenes suggest that unknown **4** is cyclic and unknowns **3** and **5** are linear diterpenes (Figure S2). The catalytic activity of TbTS $\alpha\beta'$  was sufficiently low that the steady-state kinetic parameters could not be measured.

Finally, deletion of both the  $\beta$  and  $\gamma$  domains yields the single-domain construct TbTS $\alpha$ . In full-length TbTS, a 148-residue helix (N537-F685) connects the  $\alpha$  domain to the  $\beta$  domain. This polypeptide segment was retained in TbTS $\alpha$  and connected to the N-terminal segment M79-I135 believed to help cap the active site during catalysis. The yield of TbTS $\alpha$  was comparable to that of TbTS $\alpha\beta'$  (ca. 1–2 mg/L culture), but the expressed protein was insoluble. A basic extraction protocol was performed which rescued much of the insoluble enzyme from the pellet. We were unable to measure Michaelis-Menten kinetics for TbTS $\alpha$  due to very low activity. Product array analysis revealed a complete absence of taxadiene;

instead, significant proportions of the acyclic diterpene alcohols geranylgeraniol (18.3–46.1%) and geranylgeraniol (0.0–18.7%), as well as 53.9–63.1% of new unidentified diterpenes **3** and/or **5**, were generated. Here, too, some run-to-run variations were observed in product arrays (Table 1). The mass spectra of unknown diterpenes **3** and **5** share similar fragmentation patterns with linear diterpenes (Figure S2), suggesting that these are also linear diterpene products. Thus, deletion of both  $\beta$  and  $\gamma$  domains not only compromises catalytic efficiency, but also appears to abolish the ability to cyclize GGPP since only linear diterpene products are observed.

Taken together, these results indicate that, although the  $\beta$  and  $\gamma$  domains do not directly participate in the chemistry of cyclization of GGPP by the  $\alpha$  domain, these domains influence catalysis by their direct or indirect interaction with the  $\alpha$  domain, presumably by stabilizing the  $\alpha$  domain in a catalytically competent conformation. The effects of the  $\beta$  and  $\gamma$  domains on catalytic efficiency are more pronounced than on cyclization fidelity. While retention of the  $\beta$  domain preserves substantial taxadiene formation, complete deletion of the  $\beta$  and  $\gamma$  domains abolishes taxadiene formation with production of other diterpene products.

### Cluster channeling in PaFS

We previously reported that PaFS adopts hexameric ( $\alpha\alpha$ )<sub>6</sub> quaternary structure in solution.<sup>30</sup> In addition to a modest two-fold enhancement of product flux due to covalent linkage of the GGPP synthase and GGPP cyclase domains on the same polypeptide chain,<sup>30</sup> the hexameric quaternary structure is proposed to enhance the efficiency of fusicoccadiene formation through cluster channeling.<sup>13</sup> Before preparing chimeric synthases with PaFS, we decided to evaluate the possibility of cluster channeling in the wild-type enzyme hexamer, since this potential catalytic advantage would likely be compromised in chimeric synthases that do not maintain hexameric quaternary structure.

When exogenous GGPP is utilized as a substrate, the catalytic efficiencies ( $k_{\text{cat}}/K_{\text{M}}$ ) for individual full-length PaFS<sup>30</sup> and TbTS (see below) are  $3.3 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$  and  $2.1 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$ , respectively. Theoretically, the ratio of the catalytic efficiencies for GGPP cyclization by the two enzymes,  $(k_{\text{cat}}/K_{\text{M}})_{\text{PaFS}}/(k_{\text{cat}}/K_{\text{M}})_{\text{TbTS}}$ , represents their ability to compete for a common exogenous substrate. In other words, equimolar PaFS and TbTS should yield a fusicoccadiene:taxadiene product ratio of approximately 16:1 when incubated with exogenous GGPP (i.e., in the absence of any possible channeling). Experimentally, using GC-MS we measure a fusicoccadiene:taxadiene product ratio that is slightly lower, 4.3:1, when equimolar PaFS and TbTS are incubated with exogenous GGPP (Figure 4).

Strikingly, when equimolar PaFS and TbTS are co-incubated with one equivalent of DMAPP and three equivalents of IPP, such that the only source of GGPP is that generated by the GGPP synthase domain of PaFS, the resulting fusicoccadiene:taxadiene product ratio is 46:1 (Figure 4). This excess is much larger than both the theoretical maximum of 16:1 and the experimentally-measured value of 4.3:1 when PaFS and TbTS compete for exogenously supplied GGPP. These results suggest that the GGPP synthesized by the chain elongation domain of PaFS remains mostly associated with the hexamer instead of dissociating to bulk solution and then rebinding to the cyclase domain. If GGPP were instead to dissociate into bulk solution, an equal probability of cyclization by PaFS and by TbTS would result in an

increase in taxadiene formation. These results therefore suggest that the  $(\alpha\alpha)_6$  hexameric quaternary structure of PaFS enables cluster channeling with a resultant enhancement in fusicoccadiene formation.

### Chimeric $\alpha\alpha$ domain cyclases

Four different chimeric  $\alpha\alpha$  bifunctional synthases were constructed using the cyclase  $\alpha$  domains of PaFS and AcOS, the GGPP synthase  $\alpha$  domain of PaFS, and the GFPP synthase  $\alpha$  domain of *Emericella varicolor* stellatatriene synthase (EvSS) (Figure 2). We used the GFPP synthase  $\alpha$  domain of EvSS instead of the GFPP synthase domain of AcOS because the latter generates GGPP as a side product,<sup>33</sup> whereas the former is reported to generate solely GFPP when incubated with DMAPP and IPP.<sup>38,39</sup> In accord with the  $\alpha\alpha$  architecture of the wild-type enzymes, the cyclase domain is the N-terminal  $\alpha$  domain and the chain elongation domain is the C-terminal  $\alpha$  domain in all chimeras prepared in the current study.

In the first step of converting the  $C_{20}$  diterpene synthase PaFS into a  $C_{25}$  sesterterpene synthase, we swapped the C-terminal GGPP synthase  $\alpha$  domain of PaFS with the GFPP synthase  $\alpha$  domain of EvSS to ensure the generation of the  $C_{25}$  cyclization substrate GFPP. The resulting chimeric bifunctional  $\alpha\alpha$  domain enzyme was designated PaFS $\alpha$ -EvSS $\alpha$  (i.e., the N-terminal cyclase  $\alpha$  domain of fusicoccadiene synthase linked to the C-terminal GFPP synthase  $\alpha$  domain of EvSS). To complete the conversion into a sesterterpene synthase, we made the double mutant W225L/V228A in the N-terminal cyclase domain of this chimera, designated PaFS $\alpha'$ -EvSS $\alpha$ , to enlarge the active site so as to mimic that of AcOS and thus better accommodate GFPP.

To convert the bifunctional sesterterpene synthase AcOS into a  $C_{20}$  diterpene synthase, we replaced its C-terminal GFPP synthase domain with the GGPP synthase domain of PaFS. The resulting chimeric bifunctional  $\alpha\alpha$  domain enzyme was designated AcOS $\alpha$ -PaFS $\alpha$ . Additionally, we made the double mutant L217W/A220V in the N-terminal cyclase domain of this chimera, designated AcOS $\alpha'$ -PaFS $\alpha$ , in order to constrict the active site so as to mimic that of PaFS and better accommodate GGPP.

### PaFS $\alpha$ -EvSS $\alpha$ and PaFS $\alpha'$ -EvSS $\alpha$ chimeras

Upon incubation with DMAPP and IPP at a ratio of 1:10, GC-MS analysis indicates that PaFS $\alpha$ -EvSS $\alpha$  generates no cyclic sesterterpenes (Table 2, Figure S3), confirming that the active site of the wild-type PaFS cyclase domain is too small to accommodate GFPP. However, this chimera yields sesquiterpene and diterpene products resulting from the farnesyl diphosphate (FPP) and GGPP intermediates generated by the GFPP synthase domain. Although the GFPP synthase domain of EvSS generates solely GFPP in the wild-type enzyme,<sup>38</sup> its detachment and utilization in the PaFS $\alpha$ -EvSS $\alpha$  chimera apparently relaxes the chain elongation specificity and allows for the generation and release of FPP and GGPP intermediates en route to the generation of GFPP. Accordingly, the predominant product generated by the PaFS $\alpha$ -EvSS $\alpha$  chimera is fusicoccadiene (69.1%), but the  $C_{15}$  alcohols farnesol and nerolidol are also formed (total 30.9%). The farnesol:nerolidol ratio of 25.9:5.0 suggests that solvolysis is enzyme-catalyzed, since nonenzymic FPP solvolysis ordinarily yields a 1:3–4 ratio.<sup>40,41</sup> While we were unable to measure steady-state kinetics

for these chimeras due to the lack of a sesterterpene synthase assay with a C<sub>25</sub>-radiolabeled substrate, these chimeras are nonetheless interesting due to changes in their product arrays brought about by the mutations.

With its enlarged cyclase active site, the PaFS $\alpha$ '-EvSS $\alpha$  chimera exclusively generates three unknown sesterterpene products (Table 2). The mass spectra of these C<sub>25</sub> products indicate molecular weights of 340 and hence 6 degrees of unsaturation (Figure S3); GC-MS analysis further indicates that these unidentified sesterterpenes are distinct from both acyclic geranylarnesene and tricyclic ophiobolin F (mass spectra of all products generated are reported in Figure S3). These results demonstrate that just two site-specific mutations readily enlarge the active site of the PaFS cyclase domain so as to preferentially accommodate a C<sub>25</sub> substrate instead of a C<sub>20</sub> substrate.

### AcOS $\alpha$ -PaFS $\alpha$ and AcOS $\alpha$ '-PaFS $\alpha$ chimeras

Upon incubation with DMAPP and IPP in a ratio of 1:10, AcOS $\alpha$ -PaFS $\alpha$  and AcOS $\alpha$ '-PaFS $\alpha$  chimeras yield only sesquiterpene and diterpene products (Table 2). Thus, the GGPP synthase domain of PaFS is not functionally compromised by its fusion to the cyclase domain of AcOS in these chimeras.

A total of 11 terpenes are generated by AcOS $\alpha$ -PaFS $\alpha$  (Table 2). Some of these products are conclusively identified on the basis of GC-MS data (Figure S3), but the majority of these products are newly formed and uncharacterized diterpenes. A minor amount of nerolidol is detected (0.1%), indicating release of intermediate FPP in the chain elongation reaction catalyzed by the C-terminal GGPP synthase domain. This broad product array suggests that there is significant flexibility in the binding of GGPP in the active site of the AcOS cyclase domain that normally accommodates the larger substrate GFPP. A more loosely-held substrate can adopt alternative conformations leading to the generation of a wider product array upon initial carbocation formation. However, constriction of the AcOS cyclase domain through the L217W and A220V mutations in AcOS $\alpha$ '-PaFS $\alpha$  reduces the products array to 8 compounds, including newly generated cembrene A (9.8%) (Table 2). While AcOS $\alpha$ -PaFS $\alpha$  generates 43.5% of the linear alcohols geranylgeraniol and geranylinalool, AcOS $\alpha$ '-PaFS $\alpha$  generates 13.3% geranylgeraniol as the sole linear alcohol. A diterpene alcohol of unknown structure is also generated by AcOS $\alpha$ -PaFS $\alpha$  and AcOS $\alpha$ '-PaFS $\alpha$  (14.2% and 25.6%, respectively). These results suggest that while the constricted active site of AcOS $\alpha$ '-PaFS $\alpha$  cannot enforce a GGPP binding conformation leading to the formation of the core 5-8-5 ring structure of fusicoccadiene, it can enforce an alternative binding conformation leading to a 1,14-cyclization reaction that yields cembrene A, among other diterpene products.

We measured the steady-state kinetics of GGPP cyclization by these chimeric enzymes using a radiolabeled substrate (Table 3, Figure 5). Although the turnover number ( $k_{cat}$ ) of AcOS $\alpha$ -PaFS $\alpha$  is greater than that of wild-type AcOS, an increase in  $K_M$  results in equal catalytic efficiencies ( $k_{cat}/K_M$ ) for the generation of diterpenes by AcOS $\alpha$ -PaFS $\alpha$  and wild-type AcOS. For AcOS $\alpha$ '-PaFS $\alpha$ , with its cyclase active site constricted by the L217W and A220V mutations, catalytic efficiency is significantly reduced (Table 3). Interestingly, the utilization of GGPP by AcOS $\alpha$ '-PaFS $\alpha$  did not follow Michaelis-Menten kinetics; while

curve-fitting is less than ideal, we interpreted these data as reflecting cooperativity with a Hill coefficient of 1.8. This chimera exhibited a catalytic efficiency approximately 100-fold lower than that of AcOS $\alpha$ -PaFS $\alpha$ . Curiously, the fusion of the PaFS GGPP synthase domain with the AcOS cyclase domain in AcOS $\alpha$ -PaFS $\alpha$  is not sufficient to induce cooperativity – only when the L217W and A220V mutations are introduced is cooperativity observed.

## Discussion

Terpene synthases catalyze the most complex reactions in biology, in which more than half of the substrate carbon atoms typically undergo changes in chemical bonding during the course of a chain elongation or cyclization reaction. Since cyclic terpenes cannot be generated from a linear precursor such as GGPP in the absence of an enzyme, the rate enhancement over the uncatalyzed rate for even the most sluggish terpene cyclase is immeasurably large. A terpenoid cyclase is also unique among enzymes in that its catalytic prowess is represented by its product diversity. Indeed, the fidelity or promiscuity of the cyclization cascade is a defining feature of a terpenoid cyclase.

The active site of a terpenoid cyclase can be altered by mutation of single amino acid residues or the addition/deletion of exonic domains containing hundreds of amino acid residues, and these changes may or may not influence cyclization fidelity. For example, early domain-swapping experiments involving the  $\alpha\beta$  sesquiterpene cyclases *Nicotiana tabacum* 5-*epi*-aristolochene synthase and *Hyoscyamus muticus* vetispiradiene synthase demonstrated that the  $\beta$  domain of one cyclase could be partially or completely swapped with that of the other while retaining wild-type cyclization fidelity.<sup>42</sup> Specific activity measurements indicated that most chimeras retained wild-type levels of catalytic activity, which was probably a reflection of the fact that these enzymes are already related by 77% amino acid sequence identity.<sup>43</sup> Structural analysis<sup>15</sup> indicates that 75% of  $\alpha\beta$  interface residues (van der Waals contact or hydrogen bonding) are conserved and 88% are similar between these two enzymes, likely ensuring a complementary fit in domain-swapped  $\alpha\beta$  chimeras. The preparation of catalytically-active chimeric enzymes is not limited to sesquiterpene synthases, since subsequent examples of domain-swapping include monoterpene synthases from *Citrus limon* and *Salvia officinalis*.<sup>44,45</sup>

In contrast, the  $\beta$  domains of TbTS and PcIS are related by 31% sequence identity, with conservation of 42% of  $\alpha\beta$  interface residues (53% similarity) as illustrated by the crystal structure of TbTS.<sup>17</sup> Thus, to ensure an intact  $\alpha\beta$  interface in TbTS $\alpha\beta'$ , only the first half of the  $\beta$  domain was substituted with that of PcIS (Figure 2). Consequently, with a conserved  $\alpha\beta$  interface, TbTS $\alpha\beta'$  retains comparable cyclization fidelity to that of TbTS $\alpha\beta_G$  (Table 1), albeit with catalytic activity insufficiently high to permit measurement of steady-state kinetic parameters. A complementary fit between  $\alpha$  and  $\beta$  domains in an  $\alpha\beta$  cyclase is likely to be an important determinant of cyclization fidelity, since deletion of the  $\beta\gamma$  domains altogether in TbTS $\alpha$  yields a compromised catalyst that appears to generate only linear and not cyclic diterpene products.

Prior domain engineering experiments with bifunctional  $\alpha\beta\gamma$  cyclases reveal contrasting degrees of interdependence of one domain on the other. Separate  $\alpha$  and  $\beta\gamma$  domain

constructs of *ent*-kaurene synthase from *Phaeosphaeria* sp. L487 expressed as GST fusion proteins are catalytically active, but exhibit up to 30-fold reductions in catalytic activity.<sup>46</sup> In contrast, separate  $\alpha$  and  $\beta\gamma$  domain constructs of abietadiene synthase from *Abies grandis* (grand fir) are catalytically inactive,<sup>47</sup> indicating that the structure and function of the  $\alpha$  and  $\beta\gamma$  domains are much more interdependent than those of *ent*-kaurene synthase. While the  $\beta\gamma$  domains of TbTS are not catalytically active, that their deletion results in a severely compromised  $\alpha$  domain catalyst unable to generate a cyclic diterpene product is similar to the domain interdependence previously observed for abietadiene synthase.

Bifunctional  $\alpha\alpha$  diterpene cyclases make for an interesting contrast with bifunctional and monofunctional  $\alpha\beta\gamma$  diterpene cyclases in terms of the interdependence of one domain on the other. PaFS serves as a paradigm  $\alpha\alpha$  system for probing such interactions. Here, too, catalytic efficiency and cyclization fidelity are substantially altered in all chimeric synthases studied, indicating that domain-domain interactions play a significant role in stabilizing cyclase domains in catalytically competent conformations. Even so, the substitution of a GFPP synthase  $\alpha$  domain for the wild-type GGPP synthase  $\alpha$  domain in the PaFS $\alpha$ -EvSS $\alpha$  chimera does not result in the utilization of the GFPP by the cyclase domain of PaFS. Additional mutations are required to increase the volume of the PaFS active site to accommodate the larger C<sub>25</sub> substrate, as guided by structural homology between the cyclase domains of PaFS and AcOS. Such protein engineering experiments also demonstrate that the active site of a diterpene cyclase can be readily enlarged to favor the utilization of a larger isoprenoid diphosphate substrate and disfavor the utilization of the smaller native substrate to yield new products.

Interestingly, Abe and colleagues have recently described the bifunctional diterpene synthase *Emericella varicolor* variediene synthase (EvVS) which, similar to PaFS, exhibits double  $\alpha\alpha$  domain architecture with a C-terminal GGPP synthase domain and an N-terminal cyclase domain.<sup>39</sup> In contrast with PaFS, however, EvVS can utilize both GGPP and GFPP as cyclization substrates. The construction of a chimeric enzyme containing the GFPP synthase domain of EvSS and the cyclase domain of EvVS resulted in the generation of a C<sub>25</sub> sesterterpene, (2*E*)- $\alpha$ -cericerene, as the major product *in vivo*.<sup>39</sup> Notably, minor amounts of variediene were also detected, and the total yield of terpenes was decreased relative to the wild-type system.<sup>39</sup> These results indicate, too, that functional  $\alpha$  domains can be swapped in bifunctional class I  $\alpha\alpha$  terpenoid synthases to engineer product size and diversity.

Geosmin synthase from *Streptomyces coelicolor* (ScGS) is an example of a bifunctional monomeric  $\alpha\alpha$  terpenoid synthase in which both domains catalyze cyclization reactions (Figure 6).<sup>48-50</sup> The N-terminal  $\alpha$  domain converts FPP into germacradienol, which undergoes further cyclization and fragmentation in the C-terminal  $\alpha$  domain to form acetone plus the earthy odorant geosmin.<sup>51,52</sup> The N-terminal  $\alpha$  domain in isolation can convert FPP to germacradienol, and the C-terminal  $\alpha$  domain in isolation can convert germacradienol to geosmin and acetone. Furthermore, co-incubation of separately expressed N-terminal and C-terminal  $\alpha$  domains with FPP results in the formation of both germacradienol and geosmin, indicating that the two halves of the protein need not be covalently linked for the enzyme to be catalytically competent. It has also been shown that the germacradienol intermediate does

not channel from one active site to the other, but instead dissociates into the medium and re-associates with the active site in the C-terminal  $\alpha$  domain to complete the reaction sequence.<sup>52</sup> Given that ScGS is an  $\alpha\alpha$  monomer in solution as determined by small-angle X-ray scattering,<sup>50</sup> there is no possibility of cluster channeling similar to that implicated in  $(\alpha\alpha)_6$  PaFS.

Product arrays observed for full-length ScGS or an equimolar mixture of individual N-terminal and C-terminal  $\alpha$  domains reveal only up to 17% geosmin formation when incubated with FPP, so the cyclization of germacradienol catalyzed by the C-terminal domain appears to be a slower reaction in comparison with the FPP cyclization reaction catalyzed by the N-terminal domain.<sup>52</sup> Steady-state kinetics for total hydrocarbon product formation by full-length ScGS and the N-terminal  $\alpha$  domain of ScGS indicate a slight loss of catalytic efficiency for the isolated N-terminal  $\alpha$  domain, with  $k_{cat}/K_M$  values of  $1.0 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$  and  $2.8 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$ , respectively.<sup>49</sup> Additionally, mutations in the C-terminal  $\alpha$  domain of full-length ScGS suppress geosmin formation.<sup>52</sup> Thus, the presence or absence of one  $\alpha$  domain modestly influences catalytic efficiency and cyclization fidelity in the other  $\alpha$  domain in this monomeric  $\alpha\alpha$  terpenoid synthase. However, the degree to which one  $\alpha$  domain influences the other in this system is smaller than that observed for PaFS.

## Conclusions

In conclusion, domain-domain interactions play a significant role in stabilizing cyclization domains in catalytically competent conformations in class I  $\alpha\beta\gamma$ ,  $\alpha\alpha$ , and  $(\alpha\alpha)_6$  diterpene cyclases, since cyclization specificity and catalytic efficiency are significantly altered in all chimeric synthases described herein. The mutant cyclases described here also generate 12 newly formed and uncharacterized  $C_{20}$  diterpene products and 3 newly formed and uncharacterized  $C_{25}$  sesterterpene products. Thus, engineered terpenoid cyclases promise to generate new molecular species in an already-diverse family of natural products.

The catalytic efficiency of the utilization of GGPP by the  $\alpha$  domain of TbTS within the context of its  $\alpha\beta\gamma$  domain architecture seems to be more sensitive to truncation or swapping of associated  $\beta\gamma$  domains than is the specificity of the cyclization itself (Table 1, Figure 3). The crystal structure of TbTS reveals that the  $\beta$  domain partially buttresses the adjacent  $\alpha$  domain (Figure 1), so the loss of such stabilizing interactions presumably causes compensatory structural changes in the active site that compromise catalysis. Based on the results with TbTS $\alpha\beta_G$  and TbTS $\alpha\beta'$ , such compensatory structural changes can be transmitted across more than one domain-domain interface, since deletion of the  $\gamma$  domain of TbTS influences cyclization fidelity and catalytic efficiency even though the  $\gamma$  domain does not directly contact the  $\alpha$  domain.

Interestingly, the catalytic efficiency of GGPP utilization by the  $\alpha$  domain of AcOS in wild-type AcOS and the AcOS $\alpha$ -PaFS $\alpha$  chimera is relatively insensitive to substitution of the C-terminal domain (Table 3, Figure 5c). Although GGPP is not the biological substrate of AcOS, its activity nonetheless provides a useful measure of catalytic efficiency. The comparable catalytic efficiencies of AcOS and AcOS $\alpha$ -PaFS $\alpha$  could imply minimal interdomain interactions in these  $\alpha\alpha$  assemblies.

Domain-domain and subunit-subunit interactions facilitate cluster channeling in PaFS (Figure 4), in which the bifunctional wild-type enzyme adopts a hexameric quaternary structure,  $(\alpha\alpha)_6$ .<sup>30</sup> Quaternary structure can also be used to regulate the catalytic activity of a class I terpenoid cyclase. A model for such functional regulation is found in *Mentha piperita* GPP synthase, in which a truncated, catalytically-inert  $\alpha'$  domain ( $\alpha'$ ) regulates the function of a catalytically-active  $\alpha$  domain in an  $\alpha_2\alpha'_2$  tetramer.<sup>53</sup> It is conceivable that such regulatory strategies could also be engineered in  $\alpha\alpha$  or  $\alpha\beta\gamma$  diterpene cyclases, such that domain-domain or subunit-subunit interactions could direct cyclization fidelity toward one product or another. There are many examples demonstrating that just single-residue substitutions,<sup>54–61</sup> or even epistatic pair substitutions,<sup>62</sup> can significantly alter product outcome for a diterpene cyclase, so interdomain interactions can be reasonably expected to influence product outcome as well. Future work from our laboratory in this regard will be reported in due course.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We are grateful to Drs. Ikuro Abe and Takaaki Mitsuhashi for helpful discussions and advice. This work was supported by NIH grant GM56838 to D.W.C. and grant GM30301 to D.E.C.

## References

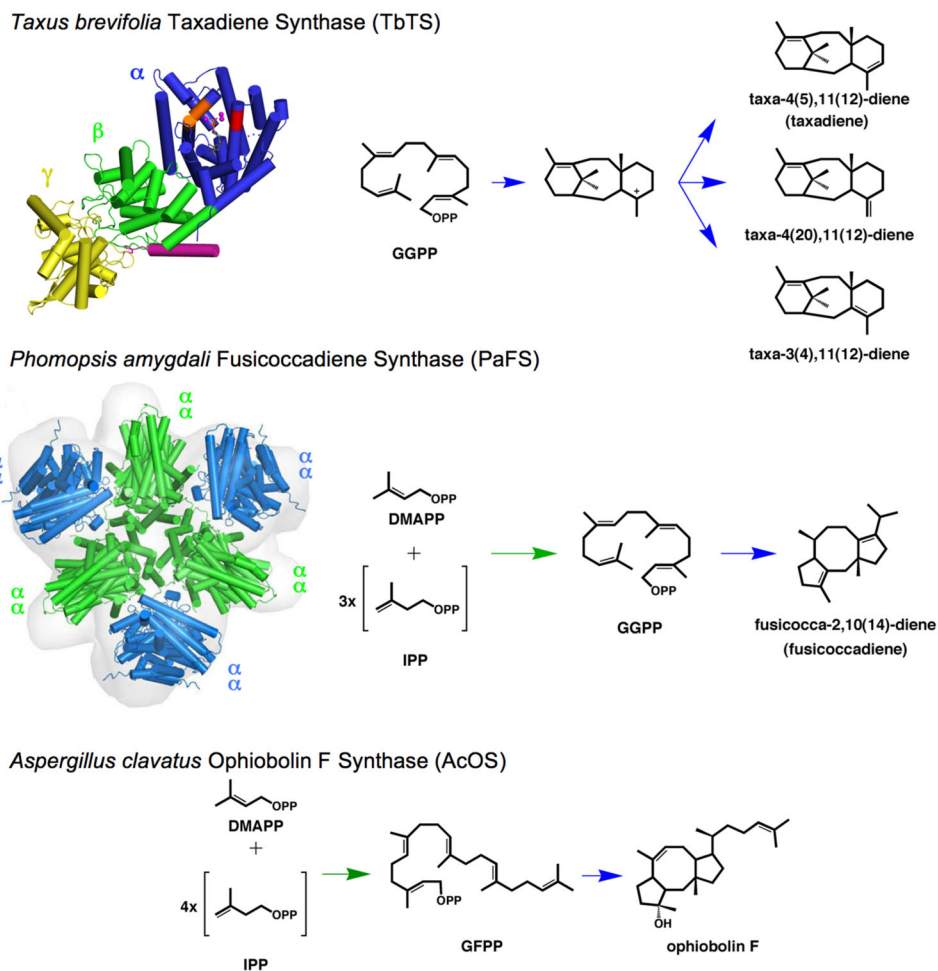
1. Dolara P, Luceri C, Ghelardini C, Monserrat C, Aiolfi S, Luceri F, Lodovici M, Menichetti S, Romanelli MN. Analgesic effects of myrrh. *Nature*. 1996; 379:29. [PubMed: 8538737]
2. Patel T, Ishiujji Y, Yosipovitch G. Menthol: a refreshing look at this ancient compound. *J Am Acad Dermatol*. 2007; 57:873–878. [PubMed: 17498839]
3. Gershenzon J, Dudareva N. The function of terpene natural products in the natural world. *Nat Chem Biol*. 2008; 3:408–414.
4. Baeten J, Deforce K, Challe S, De Vos D, Degryse P. Holy smoke in medieval funerary rites: chemical fingerprints of frankincense in southern Belgian incense burners. *PLoS One*. 2014; 9(11):e113142. [PubMed: 25391130]
5. Pichersky E, Raguso RA. Why do plants produce so many terpenoid compounds? *New Phytol*. 2016; in press. doi: 10.1111/nph.14178
6. Poulter CD, Rilling HC. The prenyl transfer reaction. Enzymic and mechanistic studies of the 1'-4 coupling reaction in the terpene biosynthetic pathway. *Acc Chem Res*. 1978; 11:307–313.
7. Cane DE. Enzymatic formation of sesquiterpenes. *Chem Rev*. 1990; 90:1089–1103.
8. Davis, EM., Croteau, R. *Biosynthesis: Aromatic polyketides, isoprenoids, alkaloids*. Vol. 209. Berlin: Springer-Verlag Berlin; 2000. Cyclization enzymes in the biosynthesis of monoterpenes, sesquiterpenes, and diterpenes; p. 53-95.
9. Tholl D. Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. *Curr Opin Plant Biol*. 2006; 9:297–304. [PubMed: 16600670]
10. Christianson DW. Structural biology and chemistry of the terpenoid cyclases. *Chem Rev*. 2006; 106:3412–3442. [PubMed: 16895335]
11. Miller DJ, Allemann RK. Sesquiterpene synthases: passive catalysts or active players? *Nat Prod Rep*. 2012; 29:60–71. [PubMed: 22068697]
12. 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]



13. Chen M, Harris GG, Pemberton TA, Christianson DW. Multi-domain terpenoid cyclase architecture and prospects for proximity in bifunctional catalysis. *Curr Op Struct Biol.* 2016; 41:27–37.
14. 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]
15. Starks CM, Back K, Chappell J, Noel JP. Structural basis for cyclic terpene biosynthesis by tobacco 5-*epi*-aristolochene synthase. *Science.* 1997; 277:1815–1820. [PubMed: 9295271]
16. Wendt KU, Poralla K, Schulz GE. Structure and function of a squalene cyclase. *Science.* 1997; 277:1811–1815. [PubMed: 9295270]
17. Köksal M, Jin Y, Coates RM, Croteau R, Christianson DW. Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis. *Nature.* 2011; 469:116–120. [PubMed: 21160477]
18. 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]
19. Zhou K, Gao Y, Hoy JA, Mann FM, Honzatko RB, Peters RJ. Insights into diterpene cyclization from structure of bifunctional abietadiene synthase from *Abies grandis*. *J Biol Chem.* 2012; 287:6840–6850. [PubMed: 22219188]
20. Janke R, Görner C, Hirte M, Brück T, Loll B. The first structure of a bacterial diterpene cyclase: CotB2. *Acta Cryst D70.* 2014:1528–1537.
21. Whittington DA, Wise ML, Urbansky M, Coates RM, Croteau R, Christianson DW. Bornyl diphosphate synthase: Structure and strategy for carbocation manipulation by a terpenoid cyclase. *Proc Natl Acad Sci U S A.* 2002; 99:15375–15380. [PubMed: 12432096]
22. Koepf AE, Hezari M, Zajicek J, Vogel BS, LaFever RE, Lewis NG, Croteau R. Cyclization of geranylgeranyl diphosphate to *taxa*-4(5),11(12)-diene is the committed step of Taxol biosynthesis in Pacific yew. *J Biol Chem.* 1995; 270:8686–8690. [PubMed: 7721772]
23. Hezari M, Lewis NG, Croteau R. Purification and characterization of *taxa*-4(5),11(12)-diene synthase from Pacific yew (*Taxus brevifolia*) that catalyzes the first committed step of Taxol biosynthesis. *Arch Biochem Biophys.* 1995; 322:437–444. [PubMed: 7574719]
24. Williams DC, Wildung MR, Jin AQ, Dalal D, Oliver JS, Coates RM, Croteau R. Heterologous expression and characterization of a “pseudomature” form of taxadiene synthase involved in Paclitaxel (Taxol) biosynthesis and evaluation of a potential intermediate and inhibitors of the multistep diterpene cyclization reaction. *Arch Biochem Biophys.* 2000; 379:137–146. [PubMed: 10864451]
25. Toyomasu T, Tsukahara M, Kaneko A, Niida R, Mitsunashi W, Dairi T, Kato N, Sassa T. Fusicoccins are biosynthesized by an unusual chimera diterpene synthase in fungi. *Proc Natl Acad Sci U S A.* 2007; 104:3084–3088. [PubMed: 17360612]
26. Toyomasu T, Tsukahara M, Kenmoku H, Anada M, Nitta H, Ohkanda J, Mitsunashi W, Sassa T, Kato N. Transannular proton transfer in the cyclization of geranylgeranyl diphosphate to fusicoccadiene, a biosynthetic intermediate of fusicoccins. *Org Lett.* 2009; 11:3044–3047. [PubMed: 19530695]
27. de Vries-van Leeuwen IJ, Kortekaas-Thijssen C, Nzigou Mandouckou JA, Kas S, Evidente A, de Boer AH. Fusicoccin-A selectively induces apoptosis in tumor cells after interferon- $\alpha$  priming. *Cancer Lett.* 2010; 293:198–206. [PubMed: 20153922]
28. Bury M, Andolfi A, Rogister B, Cimmino A, Mégalizzi V, Mathieu V, Feron O, Evidente A, Kiss R. Fusicoccin A, a phytotoxic carbocyclic diterpene glucoside of fungal origin, reduces proliferation and invasion of glioblastoma cells by targeting multiple tyrosine kinases. *Transl Oncol.* 2013; 6:112–123. [PubMed: 23544164]
29. Wurtele M, Jelich-Ottmann C, Wittinghofer A, Oecking C. Structural view of a fungal toxin acting on a 14-3-3 regulatory complex. *EMBO J.* 2003; 22:987–994. [PubMed: 12606564]
30. Chen M, Chou WKW, Toyomasu T, Cane DE, Christianson DW. Structure and function of fusicoccadiene synthase, a hexameric bifunctional diterpene synthase. *ACS Chem Biol.* 2016; 11:889–899. [PubMed: 26734760]

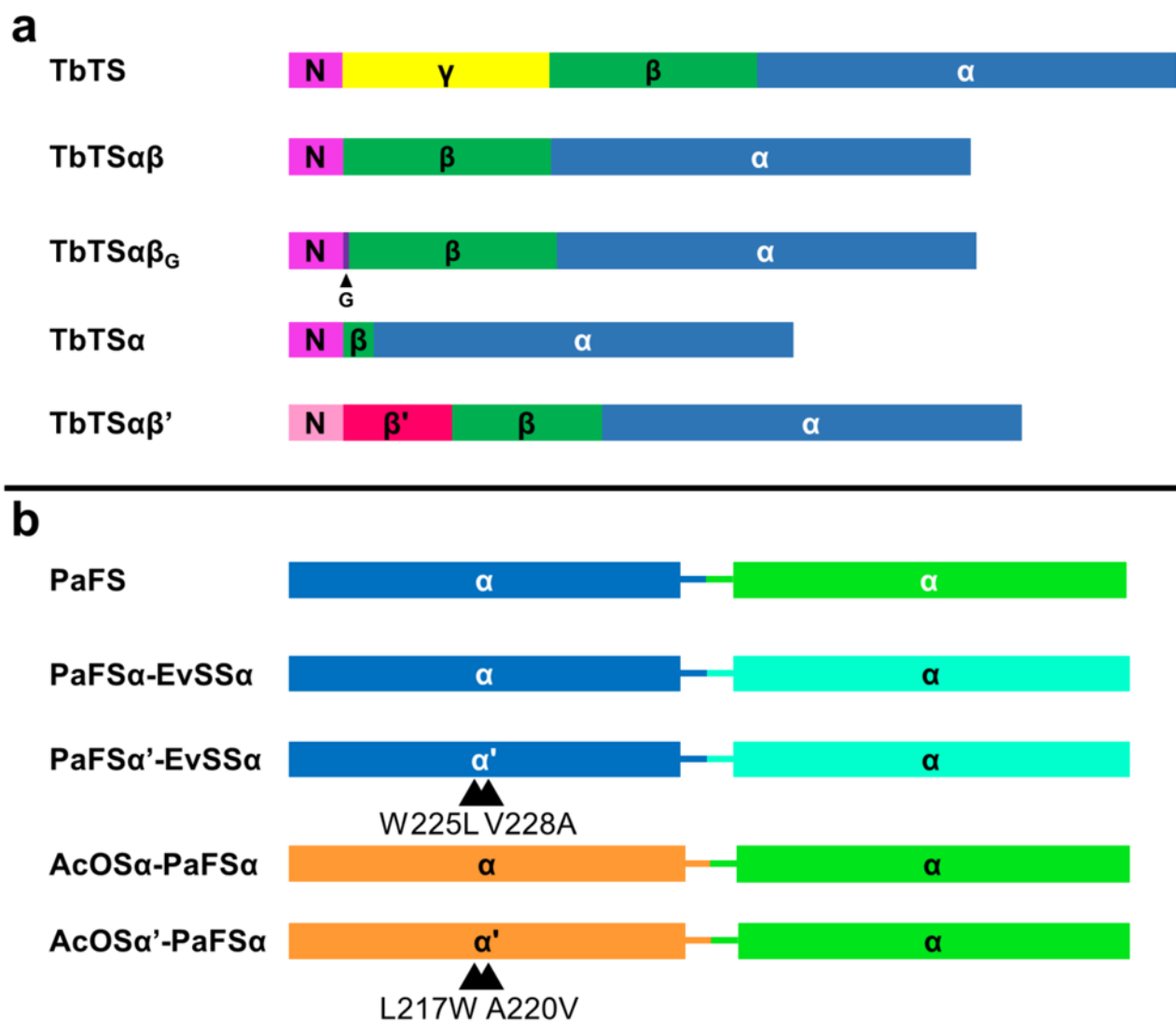
31. Bauler P, Huber G, Leyh T, McCammon JA. Channeling by proximity: the catalytic advantages of active site colocalization using Brownian dynamics. *J Phys Chem Lett.* 2010; 1:1332–1335. [PubMed: 20454551]
32. Castellana M, Wilson MZ, Xu Y, Joshi P, Cristea IM, Rabinowitz JD, Gitai Z, Wingreen NS. Enzyme clustering accelerates processing of intermediates through metabolic channeling. *Nat Biotechnol.* 2014; 32:1011–1018. [PubMed: 25262299]
33. Chiba R, Minami A, Gomi K, Oikawa H. Identification of ophiobolin F synthase by a genome mining approach: a sesterterpene synthase from *Aspergillus clavatus*. *Org Lett.* 2013; 15:594–597. [PubMed: 23324037]
34. Ye Y, Minami A, Mandi A, Liu C, Taniguchi T, Kuzuyama T, Monde K, Gomi K, Oikawa H. Genome mining for sesterterpenes using bifunctional terpene synthases reveals a unified intermediate of di/sesterterpenes. *J Am Chem Soc.* 2015; 137:11846–11853. [PubMed: 26332841]
35. Köksal M, Zimmer I, Schnitzler J, Christianson DW. Structure of isoprene synthase illuminates the chemical mechanism of teragram atmospheric carbon emission. *J Mol Biol.* 2010; 402:363–373. [PubMed: 20624401]
36. Aaron JA, Lin X, Cane DE, Christianson DW. Structure of *epi*-isozizaene synthase from *Streptomyces coelicolor* A3(2), a platform for new terpenoid cyclization templates. *Biochemistry.* 2010; 49:1787–1797. [PubMed: 20131801]
37. Harangi J. Retention index calculation without *n*-alkanes – the virtual carbon number. *J Chromatogr A.* 2003; 993:187–195. [PubMed: 12735452]
38. Matsuda Y, Mitsuhashi T, Quan Z, Abe I. Molecular basis for stellatic acid biosynthesis: a genome mining approach for discovery of sesterterpene synthases. *Org Lett.* 2015; 17:4644–4647. [PubMed: 26351860]
39. Qin B, Matsuda Y, Mori T, Okada M, Quan Z, Mitsuhashi T, Wakimoto T, Abe I. An unusual chimeric diterpene synthase from *Emericella varicolor* and its functional conversion into a sesterterpene synthase by domain swapping. *Angew Chem Int Ed.* 2016; 55:1658–1661.
40. Cramer F, Rittersdorf W. Die Hydrolyse von Phosphaten und Pyrophosphaten Einiger Monoterpenalkohole. *Tetrahedron.* 1967; 23:3015–3022.
41. George-Nascimento C, Pont-Lezica R, Cori O. Nonenzymic formation of nerolidol from farnesyl pyrophosphate in the presence of bivalent cations. *Biochem Biophys Res Commun.* 1971; 45:119–124. [PubMed: 4334520]
42. Back K, Chappell J. Identifying functional domains within terpene cyclases using a domain-swapping strategy. *Proc Natl Acad Sci U S A.* 1996; 93:6841–6845. [PubMed: 8692906]
43. Back K, Chappell J. Cloning and bacterial expression of a sesquiterpene cyclase from *Hyoscyamus muticus* and its molecular comparison to related terpene cyclases. *J Biol Chem.* 1995; 270:7375–7381. [PubMed: 7706281]
44. El Tamer MK, Lücker J, Bosch D, Verhoeven HA, Verstappen FWA, Schwab W, van Tunen AJ, Voragen AGJ, de Maagd R, Bouwmeester HJ. Domain swapping of *Citrus limon* monoterpene synthases: impact on enzymatic activity and product specificity. *Arch Biochem Biophys.* 2003; 411:196–203. [PubMed: 12623068]
45. Peters RJ, Croteau RB. Alternative termination chemistries utilized by monoterpene cyclases: chimeric analysis of bornyl diphosphate, 1,8-cineole, and sabinene synthases. *Arch Biochem Biophys.* 2003; 417:203–211. [PubMed: 12941302]
46. Kawaide H, Sassa T, Kamiya Y. Functional Analysis of the Two Interacting Cyclase Domains in *ent*-Kaurene Synthase from the Fungus *Phaeosphaeria* sp. L487 and a Comparison with Cyclases from Higher Plants. *J Biol Chem.* 2000; 275:2276–2280. [PubMed: 10644675]
47. Peters RJ, Carter OA, Zhang Y, Matthews BW, Croteau RB. Bifunctional Abietadiene Synthase: Mutual Structural Dependence of the Active Sites for Protonation-Initiated and Ionization-Initiated Cyclizations. *Biochemistry.* 2003; 42:2700–2707. [PubMed: 12614165]
48. Gust B, Challis GL, Fowler K, Kieser T, Chater KF. PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc Natl Acad Sci U S A.* 2003; 100:1541–1546. [PubMed: 12563033]

49. Cane DE, Watt RM. Expression and mechanistic analysis of a germacradienol synthase from *Streptomyces coelicolor* implicated in geosmin biosynthesis. *Proc Natl Acad Sci USA*. 2003; 100:1547–1551. [PubMed: 12556563]
50. Harris GG, Lombardi PM, Pemberton TA, Matsui T, Weiss TM, Cole KE, Köksal M, Murphy FV IV, Vedula LS, Chou WKW, Cane DE, Christianson DW. Structural studies of geosmin synthase, a bifunctional sesquiterpene synthase with  $\alpha$ . $\alpha$  domain architecture that catalyzes a unique cyclization-fragmentation reaction sequence. *Biochemistry*. 2015; 54:7142–7155. [PubMed: 26598179]
51. 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]
52. Jiang J, He X, Cane DE. Biosynthesis of the earthy odorant geosmin by a bifunctional *Streptomyces coelicolor* enzyme. *Nat Chem Biol*. 2007; 3:711–715. [PubMed: 17873868]
53. Chang TH, Hsieh FL, Ko TP, Teng KH, Liang PH, Wang AHJ. Structure of a heterotetrameric geranyl pyrophosphate synthase from mint (*Mentha piperita*) reveals intersubunit regulation. *Plant Cell*. 2010; 22:454–467. [PubMed: 20139160]
54. Xu M, Wilderman PR, Peters RJ. Following evolution's lead to a single residue switch for diterpene synthase product outcome. *Proc Natl Acad Sci U S A*. 2007; 104:7397–7401. [PubMed: 17456599]
55. Wilderman PR, Peters RJ. A single residue switch converts abietadiene synthase into a pimaradiene specific cyclase. *J Am Chem Soc*. 2007; 129:15736–15737. [PubMed: 18052062]
56. Morrone D, Xu M, Fulton DB, Determan MK, Peters RJ. Increasing complexity of a diterpene synthase reaction with a single residue switch. *J Am Chem Soc*. 2008; 130:5400–5401. [PubMed: 18366162]
57. Keeling CI, Weisshaar S, Lin RPC, Bohlmann J. Functional plasticity of paralogous diterpene synthases involved in conifer defense. *Proc Natl Acad Sci USA*. 2008; 105:1085–1090. [PubMed: 18198275]
58. Zerbe P, Chiang A, Bohlmann J. Mutational analysis of white spruce (*Picea glauca*) *ent*-kaurene synthase (*PgKS*) reveals common and distinct mechanisms of conifer diterpene synthases of general and specialized metabolism. *Phytochemistry*. 2012; 74:30–39. [PubMed: 22177479]
59. Kawaide H, Hayashi K-i, Kawanabe R, Sakigi Y, Matsuo A, Natsume M, Nozaki H. Identification of the single amino acid involved in quenching the *ent*-kauranyl cation by a water molecule in *ent*-kaurene synthase of *Physcomitrella patens*. *FEBS J*. 2011; 278:123–133. [PubMed: 21122070]
60. Irmisch S, Müller AT, Schmidt L, Günther J, Gershenzon J, Köllner TG. One amino acid makes the difference: the formation of *ent*-kaurene and 16 $\alpha$ -hydroxy-*ent*kaurane by diterpene synthases in poplar. *BMC Plant Biol*. 2015; 15:262.doi: 10.1186/s12870-015-0647-6 [PubMed: 26511849]
61. Jia M, Peters RJ. Extending a single residue switch for abbreviating catalysis in plant *ent*-kaurene synthases. *Front Plant Sci*. 2016; 7:1765.doi: 10.3389/fpls.2016.01765 [PubMed: 27920791]
62. Jia M, Zhou K, Tufts S, Schulte S, Peters RJ. A pair of residues that interactively affect diterpene synthase product outcome. *ACS Chem Biol*. 2017; 12:862–867. [PubMed: 28170228]
63. Jin Q, Williams DC, Hezari M, Croteau R, Coates RM. *J Org Chem*. 2005; 70:4667–4675. [PubMed: 15932303]



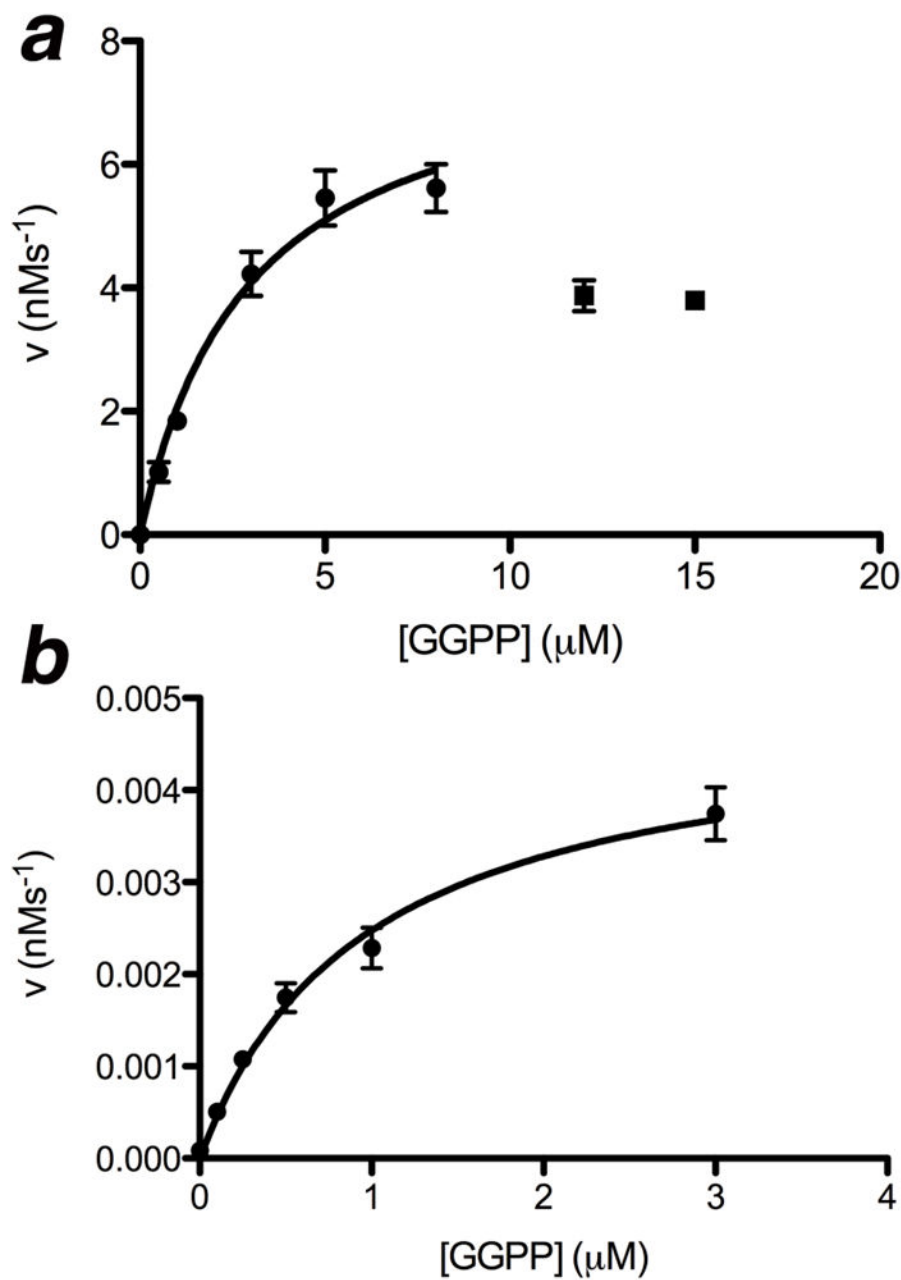
### Figure 1. Domain architecture in $\alpha\beta\gamma$ and $\alpha\alpha$ diterpene synthases

Taxadiene synthase from *Taxus brevifolia* (TbTS) catalyzes the cyclization of GGPP to form taxa-4(5),11(12)-diene as the major product (henceforth designated simply “taxadiene”). Alternative proton elimination steps yield taxa-4(20),11(12)-diene and taxa-3(4),11(12)-diene as minor products. This cyclization occurs exclusively in the  $\alpha$  domain (blue); neither the  $\beta$  domain (green) nor the  $\gamma$  domain (yellow) contain a functional active site. A polypeptide segment connected to the N-terminal helix of the  $\beta$ -domain (magenta) is believed to help cap the active site in the  $\alpha$  domain during catalysis. Fusicoccadiene synthase from *Phomopsis amygdali* (PaFS) is a hexamer of subunits with  $\alpha\alpha$  domain architecture. The chain elongation reaction of DMAPP and 3 IPP molecules is catalyzed in the C-terminal GGPP synthase  $\alpha$  domain (green), and the GGPP cyclization reaction forming fusicoccadiene is catalyzed in the N-terminal cyclase  $\alpha$  domain (blue). Although neither the tertiary structure nor the quaternary structure of ophiobolin F synthase from *Aspergillus clavatus* (AcOS) are known, this bifunctional sesterterpene synthase also adopts an  $\alpha\alpha$  domain architecture in which the chain elongation reaction of DMAPP and 4 IPP molecules is catalyzed in the C-terminal  $\alpha$  domain to form GFPP, which then undergoes cyclization in the N-terminal  $\alpha$  domain to form ophiobolin F. Detailed catalytic mechanisms for each enzyme are found in Figure S1.



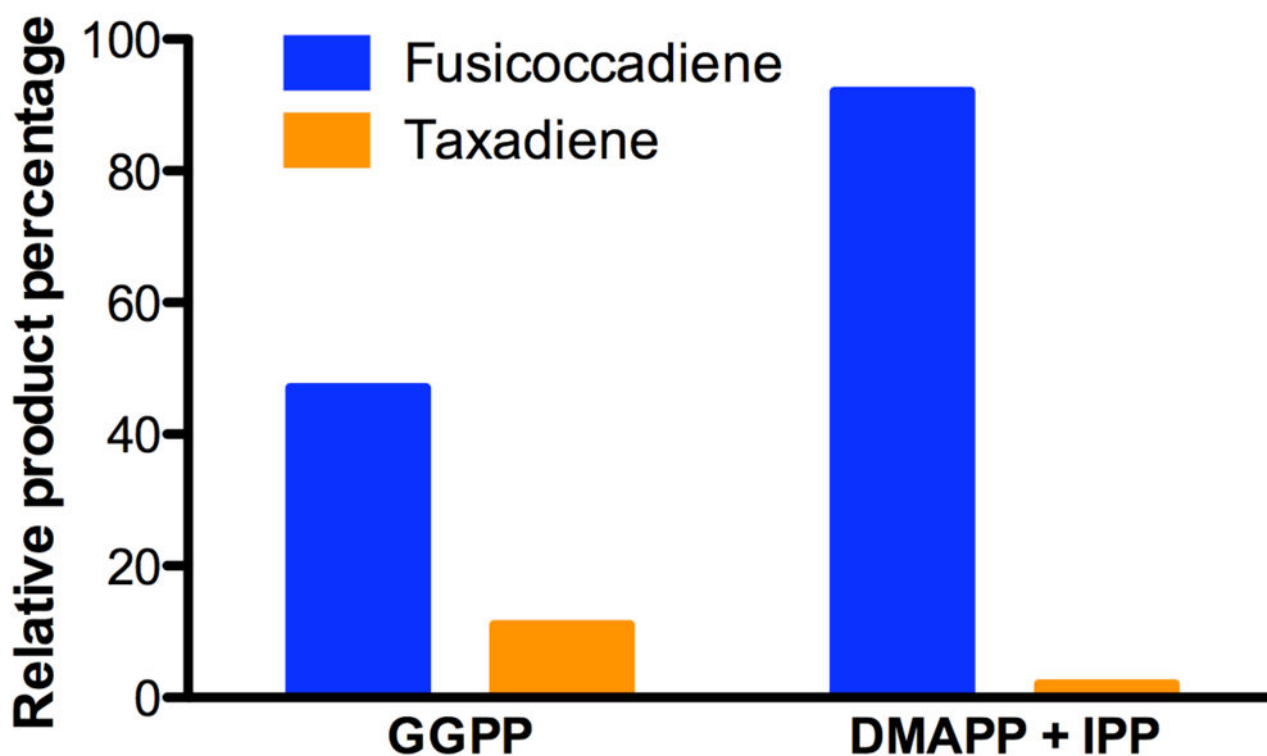
**Figure 2. Primary structure of truncated and chimeric terpenoid synthases**

(a) For the class I diterpene synthase taxadiene synthase (TbTS), four engineered domain constructs successively downsized the  $\alpha\beta\gamma$  domain architecture, color-coded as follows:  $\alpha$  domain, blue;  $\beta$  domain, green;  $\gamma$  domain, yellow; N-terminal segment, magenta. The location of the single inserted glycine residue (G) in TbTS $\alpha\beta_G$  is indicated by an arrow. The TbTS $\alpha\beta'$  is a chimera containing the N-terminal segment (pink) and part of the  $\beta$  domain ( $\beta'$ , red) of isoprene synthase. (b) Engineered bifunctional class I diterpene and sesterterpene synthases with  $\alpha\alpha$  domain architecture are color-coded as follows: dark blue, cyclization  $\alpha$  domain of fusicoccadiene synthase (PaFS); orange, cyclization  $\alpha$  domain of ophiobolin F synthase (AcOS); aquamarine, GGPP synthase  $\alpha$  domain of stellatatriene synthase (EvSS); green, GGPP synthase  $\alpha$  domain of fusicoccadiene synthase (PaFS).



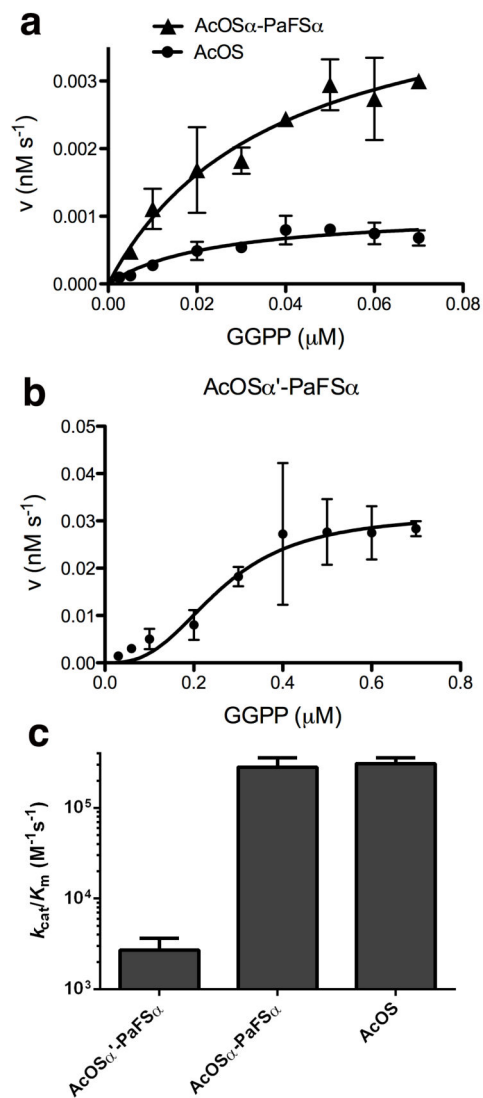
**Figure 3. Catalytic activity measurements**

(a) The cyclization of GGPP by full-length TbTS exhibits Michaelis-Menten kinetics with  $k_{\text{cat}} = 0.16 \text{ s}^{-1}$ ,  $K_{\text{M}} = 2.9 \mu\text{M}$ , and  $k_{\text{cat}}/K_{\text{M}} = 5.5 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$ . The highest concentrations of GGPP are not included in the final curve due to substrate inhibition. (b) The cyclization of GGPP by TbTS $\alpha\beta\text{G}$  exhibits Michaelis-Menten kinetics with  $k_{\text{cat}} = 0.000019 \text{ s}^{-1}$ ,  $K_{\text{M}} = 0.96 \mu\text{M}$ , and  $k_{\text{cat}}/K_{\text{M}} = 20 \text{ M}^{-1}\text{s}^{-1}$ .



**Figure 4. Substrate competition between PaFS and TbTS**

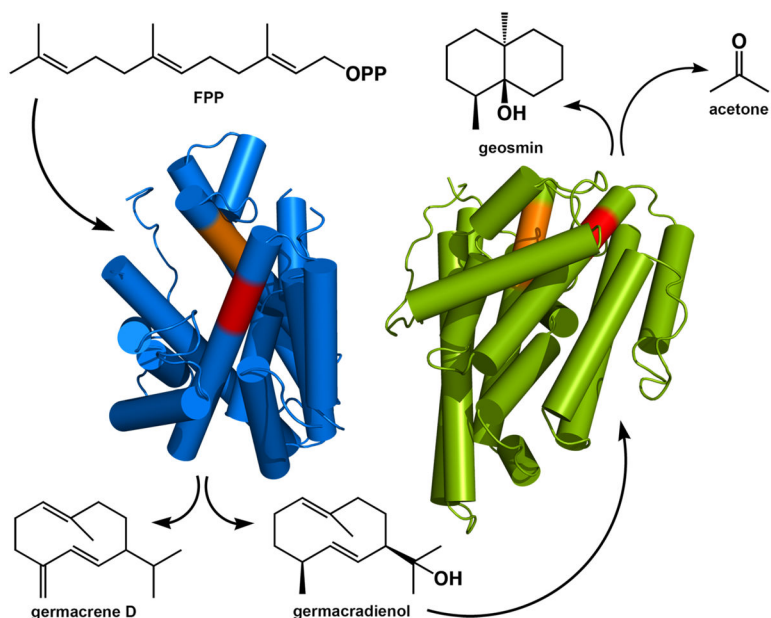
Relative product percentages for the generation of fusicoccadiene and taxadiene in reaction mixtures containing equimolar PaFS and TbTS. When the enzyme mixture is incubated with exogenous GGPP, the fusicoccadiene:taxadiene product ratio is 4.3:1. When the enzyme mixture is incubated with DMAPP and IPP, the only source of cyclization substrate GGPP is that generated by the C-terminal  $\alpha$  domain of PaFS and the resulting fusicoccadiene:taxadiene product ratio increases to 46:1. That very little taxadiene is generated in this experiment strongly suggests a cluster channeling model for bifunctional catalysis, in which most of the GGPP generated remains bound in the PaFS hexamer to be utilized for cyclization to fusicoccadiene.



**Figure 5. Steady-state kinetics**

(a) AcOS and AcOS $\alpha$ -PaFS $\alpha$  exhibit Michaelis-Menten kinetics for GGPP utilization. (b) AcOS $\alpha'$ -PaFS $\alpha$  exhibits sigmoidal kinetics for GGPP utilization. (c) Comparison of catalytic efficiencies for AcOS, AcOS $\alpha$ -PaFS $\alpha$ , and AcOS $\alpha'$ -PaFS $\alpha$ .



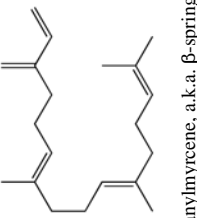
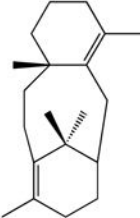
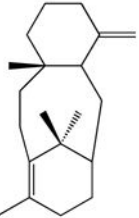
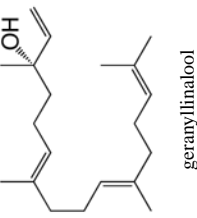


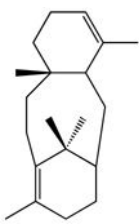
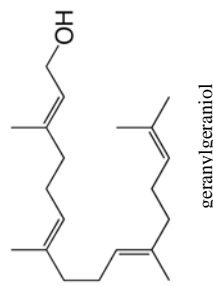
**Figure 6. Geosmin synthase**

Geosmin synthase is a bifunctional enzyme with  $\alpha\alpha$  domain architecture that catalyzes FPP cyclization in the N-terminal domain (blue) to generate germacrene D and germacradienol, the latter of which undergoes a cyclization and fragmentation reaction in the C-terminal domain (green) to yield geosmin and acetone. The structure of the N-terminal domain was determined by X-ray crystallographic methods, and the structure of the C-terminal domain was generated through homology modeling. The domain assembly shown is based on the best fit of small-angle X-ray scattering data (the 41-residue interdomain linker is not shown).<sup>50</sup>

Table 1

Product arrays generated by TbTS mutants<sup>a</sup>

Retention Index (calc)	Retention Index (ref)	Compound	TbTS run 1 (%)	TbTS run 2 (%)	TbTSc $\alpha\beta\gamma$ (%)	TbTSc $\alpha\beta'$ (%)	TbTSc run 1 (%)	TbTSc run 2 (%)
1893	1922 <sup>b</sup>	 geranylmyrcene, a.k.a. $\beta$ -springene	1.7	–	20.3	12.6	–	–
1912	–	 taxa-3(4),11(12)-diene <sup>c</sup>	2.7	–	–	–	–	–
1919	–	unknown diterpene 1	–	16.8	–	–	–	–
1954	–	unknown diterpene 2	1.5	–	–	–	–	–
1965	–	unknown diterpene 3	0.3	10.0	7.1	8.6	53.9	24.2
1970	–	unknown diterpene 4	3.0	–	5.7	1.7	–	–
1998	–	 taxa-4(20),11(12)-diene <sup>c</sup>	7.2	15.0	4.3	4.9	–	–
2008	2008 <sup>d</sup>	 geranylmyrcene geranyl linalool	1.4	–	21.0	16.0	46.1	18.3

Retention Index (calc)	Retention Index (ref)	Compound	TbTS run 1 (%)	TbTS run 2 (%)	TbTSa <sub>β</sub> (%)	TbTSa <sub>β</sub> (%)	TbTSa <sub>β</sub> (%)	TbTSa <sub>β</sub> (%)	TbTSa <sub>β</sub> (%)
2043	–	 taxa-4(5),11(12)-diene (taxadiene) <sup>c</sup>	82.0	58.2	25.9	37.6			
2066	–	unknown diterpene <b>5</b>	0.1	–	15.6	18.5	38.9	–	–
2173	2167 <sup>d</sup>	 geranylgeraniol geranylgeranitol	–	–	–	–	18.7	–	–

<sup>a</sup>Mass spectra for all compounds are recorded in Figure S2.

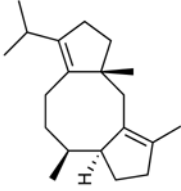
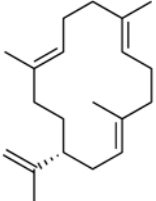
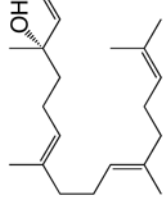
<sup>b</sup>Retention index referenced from the NIST database.

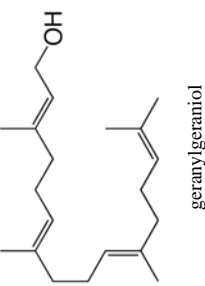
<sup>c</sup>Mass fragmentation pattern matched that reported in ref. 63.

<sup>d</sup>Retention index referenced from the MassFinder database.

**Table 2**

Product arrays generated by  $\alpha,\alpha$  synthase chimeras<sup>a</sup>

Retention Index (calc)	Retention Index (ref)	Compound	PaFSa-EvSSa (%)	PaFSa-EvSSa (%)	AcOSa-PaFSa (%)	AcOSa'-PaFSa (%)
1539	1553 <sup>b</sup>	nerolidol	5.0	-	0.1	5.0
1695	1694 <sup>b</sup>	farnesol	25.9	-	-	-
1843	-	 fusicoccadiene <sup>c</sup>	69.1	-	-	-
1888	-	unknown diterpene <b>6</b>	-	-	8.4	5.8
1928	-	unknown diterpene <b>7</b>	-	-	7.1	6.1
1937	-	unknown diterpene <b>8</b>	-	-	7.0	22.1
1964	1962 <sup>b</sup>	 cembrene A	-	-	-	9.8
2002	2008 <sup>b</sup>	 geranylinalool	-	-	18.1	13.3
2016	-	unknown diterpene <b>9</b>	-	-	4.3	-
2128	-	unknown diterpene <b>10</b>	-	-	3.0	-
2140	-	unknown diterpene <b>11</b>	-	-	5.7	-
2149	-	unknown diterpene <b>12</b>	-	-	6.7	12.2

Retention Index (calc)	Retention Index (ref)	Compound	PaFSa'-EvsSSa. (%)	PaFSa'-EvsSSa. (%)	AcOSa'-PaFSa. (%)	AcOSa'-PaFSa. (%)
2162	2167 <sup>b</sup>	 geranylgeraniol unknown diterpene alcohol	-	-	25.4	-
2214	-	unknown diterpene alcohol	-	-	14.2	25.6
2299	-	unknown sesquiterpene <b>1</b>	-	40.7	-	-
2338	-	unknown sesquiterpene <b>2</b>	-	14.5	-	-
2400	-	unknown sesquiterpene <b>3</b>	-	44.8	-	-

<sup>a</sup>Mass spectra for all compounds are recorded in Figure S3.

<sup>b</sup>Retention index referenced from the MassFinder database.

<sup>c</sup>Mass fragmentation pattern matched that reported in ref. 25.

**Table 3**Steady-state kinetics for wild-type and chimeric bifunctional  $\alpha\alpha$  terpene synthases with substrate GGPP

Enzyme	PaFS <sup>a</sup>	AcOS	AcOS $\alpha$ -PaFS $\alpha$	AcOS $\alpha'$ -PaFS $\alpha$
$k_{\text{cat}}$ (s <sup>-1</sup> )	0.021 $\pm$ 0.002	0.0069 $\pm$ 0.001	0.011 $\pm$ 0.002	0.0012 $\pm$ 0.0003
$K_M$ ( $\mu$ M)	0.63 $\pm$ 0.03	0.024 $\pm$ 0.01	0.038 $\pm$ 0.01	0.48 $\pm$ 0.2
$h$	2.2	1	1	1.8
$k_{\text{cat}}/K_M$ (M <sup>-1</sup> s <sup>-1</sup> )	3.3 $\times$ 10 <sup>4</sup>	2.9 $\times$ 10 <sup>5</sup>	2.9 $\times$ 10 <sup>5</sup>	2.5 $\times$ 10 <sup>3</sup>

<sup>a</sup>From ref. 30.