Biosynthesis

International Edition: DOI: 10.1002/anie.201509263
German Edition: DOI: 10.1002/ange.201509263

# An Unusual Chimeric Diterpene Synthase from *Emericella variecolor* and Its Functional Conversion into a Sesterterpene Synthase by Domain Swapping

*Bin Qin[+], Yudai Matsuda[+], Takahiro Mori, Masahiro Okada, Zhiyang Quan, Takaaki Mitsuhashi, Toshiyuki Wakimoto, and Ikuro Abe**

*Dedicated to Professor Yutaka Ebizuka on the occasion of his 70th birthday*

***Abstract:*** *Di- and sesterterpene synthases produce $C_{20}$ and $C_{25}$ isoprenoid scaffolds from geranylgeranyl pyrophosphate (GGPP) and geranylfarnesyl pyrophosphate (GFPP), respectively. By genome mining of the fungus Emericella variecolor, we identified a multitasking chimeric terpene synthase, EvVS, which has terpene cyclase (TC) and prenyltransferase (PT) domains. Heterologous gene expression in Aspergillus oryzae led to the isolation of variediene (**1**), a novel tricyclic diterpene hydrocarbon. Intriguingly, in vitro reaction with the enzyme afforded the new macrocyclic sesterterpene **2** as a minor product from dimethylallyl pyrophosphate (DMAPP) and isopentenyl pyrophosphate (IPP). The TC domain thus produces the diterpene **1** and the sesterterpene **2** from GGPP and GFPP, respectively. Notably, a domain swap of the PT domain of EvVS with that of another chimeric sesterterpene synthase, EvSS, successfully resulted in the production of **2** in vivo as well. Cyclization mechanisms for the production of these two compounds are proposed.*

Terpenoids constitute a diverse class of natural products with many important functions in plants, microbes, and humans. Cyclic terpenes are commonly biosynthesized through the cyclization of linear polyisoprenoid substrates by terpene cyclases,[1] while the linear substrates are produced through condensation of the universal $C_5$ units isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) by isoprenyl pyrophosphate synthases (also known as prenyltransferases).[2] Although the condensation and cyclization reactions usually occur independently, there exist several chimeric fungal terpene synthases that catalyze both chain elongation and cyclization reactions from DMAPP and IPP as substrates. These multifunctional enzymes include PaFS and PaPS from *Phomopsis amygdali*,[3a,b] AcOS from *Aspergillus clavatus*,[3c] EvSS from *Emericella variecolor*,[3d] and NfSS from *Neosartorya fischeri*[3e] (Figure 1). All of these chimeric enzymes consist of a C-terminal prenyltransferase (PT) and an N-terminal terpene cyclase (TC) domain.

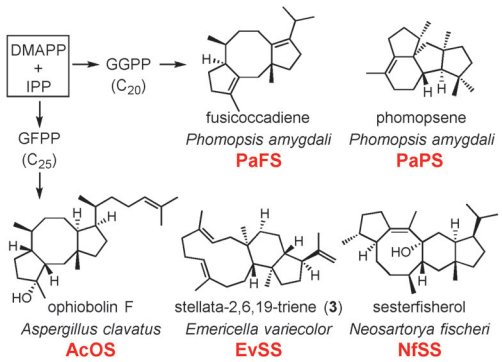


***Figure 1.*** Structures of compounds produced by chimeric fungal terpene synthases.

In the reactions catalyzed by PaFS and PaPS, the PT domains of the fusion proteins produce $C_{20}$ geranylgeranyl pyrophosphate (GGPP) as the substrate for the TC domains, while the PT domains of AcOS, EvSS, and NfSS provide $C_{25}$ geranylfarnesyl pyrophosphate (GFPP) for the following reactions. However, it still remains unknown how the PTs of these enzymes control the chain length of their products. Furthermore, there has been no report on the engineering of this class of multitasking terpene synthases to afford novel products. Herein, we report the discovery and characterization of a novel diterpene synthase that generates variediene (**1**), and the alteration of its function through a domain swapping approach to provide the $C_{25}$ sesterterpene hydrocarbon (2*E*)-α-cericerene (**2**).

We recently identified the sesterterpene synthase EvSS, which is responsible for the biosynthesis of stellata-2,6,19-triene (**3**), from the fungus *E. variecolor*,[3d] which is a rich source of sesterterpenoids.[4a–e] The genome of *E. variecolor* NBRC 32302 encodes an unusually large number of putative chimeric terpene synthases, and therefore we searched for another possible sester- or diterpene synthase gene in the genome. A phylogenetic analysis of bifunctional terpene synthases suggested that these enzymes can be grouped into five different clades (clades A–E), and that the enzymes in different clades might have distinct first cyclization modes.[3e] We thus focused on a gene encoding a chimeric protein

[*] Dr. B. Qin,[+] Dr. Y. Matsuda,[+] Dr. T. Mori, Dr. M. Okada, Z. Quan, T. Mitsuhashi, Prof. T. Wakimoto, Prof. I. Abe
Graduate School of Pharmaceutical Sciences, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033 (Japan)
E-mail: abei@mol.f.u-tokyo.ac.jp
[+] These authors contributed equally to this work.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201509263.

  

belonging to clade C (DDBJ/EMBL/GenBank accession no. LC063849; Figure S1 in the Supporting Information), from which no terpene synthase has been characterized. This putative protein contains a partially and completely conserved aspartate-rich DDXXD/E motif[5] for $Mg^{2+}$ binding in the TC domain and PT domains, respectively. Another conserved $Mg^{2+}$-binding "NSE/DTE" triad for TC[5] was found as $^{230}$**N**DYF**S**FDI**E**$^{238}$ (Figure S2).

To analyze the metabolites of this protein, we used an *Aspergillus oryzae* heterologous expression system, which has been utilized for biosynthetic studies on fungal natural products.[6] GC–MS analysis of mycelial extracts from the transformant harboring the terpene synthase gene showed the presence of a new product with *m*/*z* 272 $[M]^+$ (**1**; Figure 2 and Figure S3 A), which was expected to be a diterpene hydrocarbon.

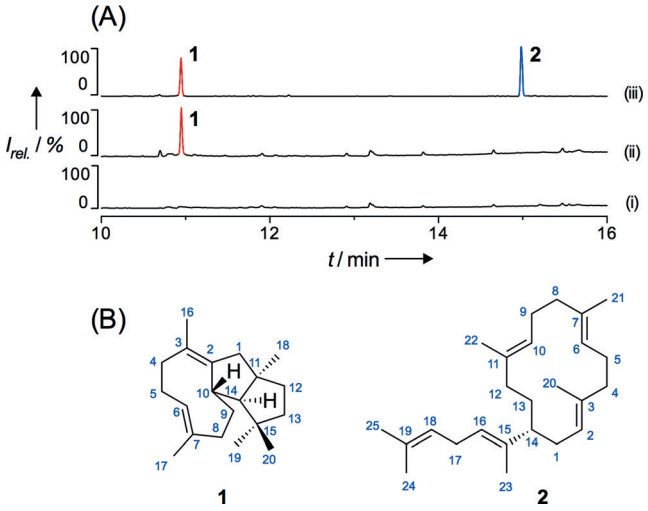


***Figure 2.*** A) GC–MS chromatogram of mycelial extracts from *A. oryzae* transformants harboring i) only empty vector, ii) EvVS, and (iii) Swap 1 (see Figure 4). B) Structures of **1** and **2** (see following paragraph). $I_{rel.}$ = relative intensity.

The structure of the diterpene was elucidated by spectroscopic analyses. The molecular formula ($C_{20}H_{32}$) of **1** was deduced from HR-EI-MS. With the aid of the NMR data (Table S2 and Figures S11–S15), the planar structure of **1** was identified as a 9-5-5 tricyclic compound (Figure 2 B). For the relative configuration, NOESY correlations indicated that H-14, Me-19, H-12α, and Me-18 are on the same side of the cyclopentane ring in an α-orientation (Figures S4C, S16), while H-10, Me-20, and H-12β are situated on the opposite face (β-oriented). The absolute stereochemistry was then established by derivatization of **1** followed by the modified Mosher's method[7] (see the Supporting Information for details). The absolute configuration of **1** was assigned as 10*S*, 11*R*, and 14*R*, which matches that of the diterpenoids phomopsene[3b] and conidiogenol[8] (Figure S6). Compound **1** was named variediene, and the chimeric terpene synthase was designated as EvVS (*E. variecolor* Variediene Synthase).

To further investigate the reactions of EvVS, we performed an in vitro assay with recombinant EvVS expressed in *Escherichia coli* as a fusion protein with Maltose-Binding Protein (MBP) at the N terminus. An incubation of MBP–EvVS with DMAPP and IPP in the presence of $MgCl_2$ gave the diterpene **1** (Figure 3, lane ii), thus indicating that this

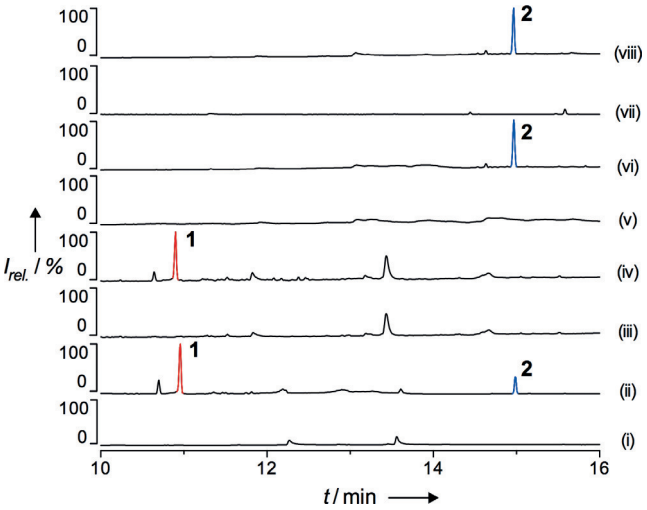


***Figure 3.*** GC–MS profiles of in vitro assays of i) boiled MBP-EvVS with DMAPP and IPP; ii) MBP-EvVS with DMAPP and IPP; iii) boiled MBP-EvVS with GGPP; iv) MBP-EvVS with GGPP; v) boiled MBP-EvVS with GFPP; vi) MBP-EvVS with GFPP; vii) boiled EvVS-N349 with GFPP; and viii) EvVS-N349 with GFPP. $I_{rel.}$ = relative intensity.

single enzyme synthesizes diterpene **1** from the five-carbon starter units. Incubation of EvVS with GGPP and $MgCl_2$ also confirmed the functions of the TC domain (Figure 3, lane iv). Surprisingly, another product (**2**) with *m*/*z* 340 $[M]^+$, expected to be a sesterterpene hydrocarbon, was detected when DMAPP and IPP were used as substrates (Figure 3, lane ii and Figure S3 B). An incubation of EvVS with GFPP further confirmed that the EvVS transforms GFPP into the sesterterpene **2** (Figure 3, lane vi). These lines of evidence suggest that the PT domain of EvVS can synthesize both GGPP and GFPP from the $C_5$ isoprene units in vitro, while the TC domain is able to cyclize both GGPP and GFPP to the corresponding diterpene **1** and sesterterpene **2**. In contrast, geranyl pyrophosphate (GPP) and farnesyl pyrophosphate (FPP) were not accepted by the TC (Figure S40). We also investigated the metal-ion dependency ($Mg^{2+}$, $Co^{2+}$ or $Mn^{2+}$) of the enzyme activities, since a previous study on an insect PT demonstrated that the chain length of its enzyme reaction products can vary depending on the metal cofactor.[9] However, EvVS was active only in the presence of $Mg^{2+}$, and no terpene-synthesizing activity was detected in the presence of $Co^{2+}$ or $Mn^{2+}$ (Figure S39).

Despite production of the sesterterpene **2** in vitro, the *A. oryzae* transformant with the EvVS gene did not yield **2** at a detectable level. The PT domain probably does not synthesize a sufficient amount of GFPP in vivo, leading to a lack of substrate for the TC domain to form **2**. This possibility led us to hypothesize that **2** could be obtained from an in vivo system by swapping the PT domain with one that predominantly synthesizing GFPP. First, to test whether the TC domain alone retains its activity, the truncated enzyme

  

EvVS-N349 was constructed and expressed in *E. coli*, according to the truncation of PaFS.[3a] The reaction of EvVS-N349 with GFPP proceeded to yield **2** (Figure 3, lane viii). We therefore designed a new chimeric enzyme consisting of the TC domain of EvVS and the PT domain of another chimeric terpene synthase EvSS from the same fungus *E. variecolor*, which produces the sesterterpene stellatatriene (**3**).[3d] Our recent work confirmed that the PT domain of EvSS has the ability to condense DMAPP and IPP to the linear GFPP. *EvVS* and *EvSS* genes display conserved intron/exon organization, and prediction of the EvVS and EvSS secondary structures suggested that the residues 336–385 in EvVS and 333–388 in EvSS form loop regions linking the TC and PT domains. Considering this information, the chimeric gene was designed carefully to include appropriate lengths of the TC/PT domains (to retain the activities) and to avoid bulky residues in the linker region between the TC and PT domains (to prevent protein misfolding and aggregation). Finally, as shown in Figure 4, the 2.8-kb *Swap 1* fragment,

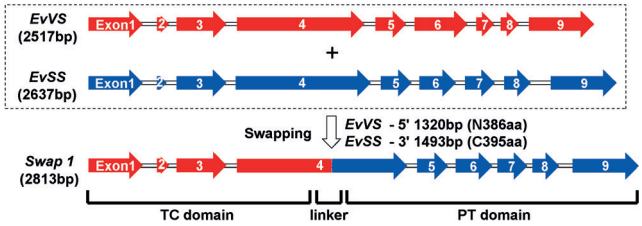


***Figure 4.*** Scheme for the domain swapping of EvVS and EvSS.

which is predicted to encode a protein consisting of the TC of EvVS (N-terminal 386 aa) and the PT of EvSS (C-terminal 395 aa), was generated. The swapped gene was expressed in *A. oryzae,* and fortunately, the transformant harboring *Swap 1* produced sesterterpene **2** as major product and diterpene **1** as minor product, although the yield of terpene compounds was decreased compared to that of the wild-type enzyme (Figure 2, lane iii). These findings indicated that the TC domain of EvVS possesses the ability to catalyze the cyclization of GGPP and GFPP, both in vivo and in vitro. The swapped enzyme still produced **1**, probably owing to the fact that the PT domain of EvSS yields GGPP as a precursor of GFPP.

The planar structure of **2** was investigated by a combination of 1D and 2D NMR measurements (see the Supporting Information, including Table S4 and Figures S17–S21, for details). Compound **2** has a simple carbon skeleton with a 14-membered ring, which is very different to that of **1** (Figure 2B). Compound **2** was named (2*E*)-α-cericerene, since **2** is the (2*E*)-isomer of α-cericerene.[10] To determine the absolute configuration at the C-14 position of **2**, (*R*)-/(*S*)-limonene and **2** were subjected to ozonolysis (Figure S9), and the keto limononaldehydes thus obtained were analyzed by GC–MS with a chiral stationary phase. The fragments obtained from **2** (both the in vitro and in vivo products) revealed that the (*R*)-enantiomer is predominant (*R*/*S* = 9:1, Figure S10).

Since the diterpene **1** and the sesterterpene **2** have different carbon skeletons, it is interesting to clarify the mechanisms underlying the cyclization reactions to form **1** and **2**. The cyclization mechanism of **2** can be easily proposed, as shown in Figure 5A. The initial C-1-C-14

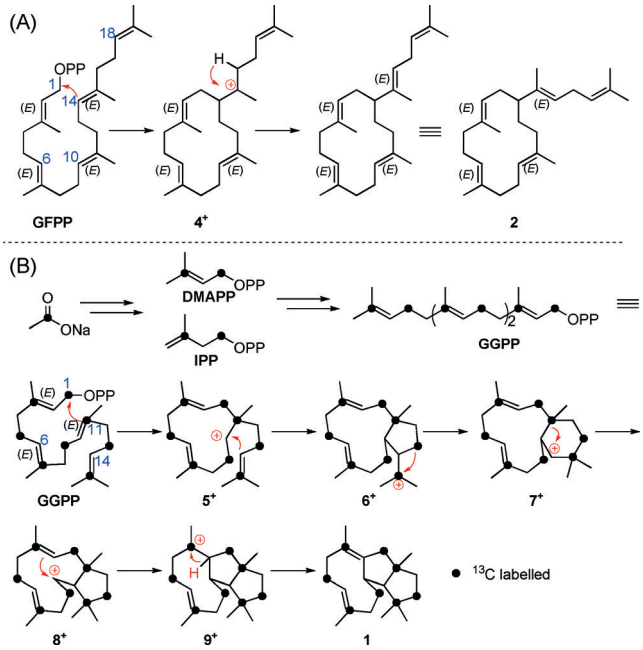


***Figure 5.*** Proposed mechanisms for the A) the cyclization of GFPP to **2** and B) the cyclization of GGPP to **1**.

cyclization, followed by deprotonation at C-16, yields **2**. To investigate the mechanism for the formation of **1**, we added (1-$^{13}$C)-labelled sodium acetate, which is incorporated at the C-1 and C-3 positions of DMAPP and IPP, to the culture broth of *A. oryzae* harboring EvVS. The $^{13}$C NMR spectrum (Figure S38) revealed efficient $^{13}$C incorporation into the C-1, C-3, C-5, C-7, C-9, C-11, C-13, and C-15 positions of **1**. On the basis of this observation, a mechanism for cyclization was proposed (Figure 5B). The initial C-1–C-11 cyclization followed by the second C-10–C-14 cyclization may proceed to produce a bicyclic tertiary cation **6**$^+$. Cation **6**$^+$ might then undergo ring expansion to yield **7**$^+$, followed by ring contraction to yield **8**$^+$. The subsequent C-2–C-10 cyclization, followed by deprotonation at C-2, finally yields **1**.

Intriguingly, the first cyclization mode for the synthesis of **1** is the same as those for fusicoccadiene, phomopsene, ophiobolin F, and stellatatriene, which are all produced by synthases that belong to clade B (Figure S1).[3e] EvVS, however, is grouped into a different clade (clade C), but it generates the same 11–5 fused bicyclic cationic intermediate. Additionally, EvVS in turn executes C-1–C-14 cyclization when GFPP is used as the substrate, thus suggesting that a single enzyme can adopt different initial cyclization patterns. Further investigations of proteins in other clades are essential for a more in-depth understanding of this class of terpene synthases.

In summary, we identified an unusual chimeric terpene synthase, EvVS, which catalyzes both condensation and

  

cyclization through its PT and TC domains. Interestingly, the TC domain of EvVS produces both a novel diterpene (**1**) and a sesterterpene (**2**) through cyclization of GGPP and GFPP, respectively. By swapping the PT domain of EvVS with a GFPP synthase, the sesterterpene **2**, which was initially only obtained in vitro, was also produced in vivo. The ophiobolin F synthase (AcOS) reportedly cyclizes GGPP in vitro, but structural determination of the resulting diterpene was not achieved owing to low productivity.[3c] Our domain-swapping strategy would bypass this type of structural characterization problem, since heterologous expression in *A. oryzae* generally results in a large amount of the product derived from the introduced gene. Furthermore, since there are numerous putative sester- or diterpene synthases with the PT and TC domains in publicly available databases, novel terpene scaffolds could most likely be obtained by applying this methodology to other terpene synthases. In addition, we proposed cyclization mechanisms for the production of **1** and **2**, which have distinct carbon skeletons. Recent studies have revealed the existence of many terpene cyclases that are different from classical ones,[11] and we believe that the unusual chimeric terpene synthase reported herein provides a remarkable addition. Further investigations of the unusual chimeric terpene synthase EvVS, including protein crystallization, would be fascinating and are now in progress in our laboratories.

## Acknowledgements

We thank Prof. K. Gomi (Tohoku University) and Prof. K. Kitamoto (The University of Tokyo) for their help in fungal transformation. This work was supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan. B.Q. is supported by The Tokyo Biochemical Research Foundation Postdoctoral Fellowship.

**Keywords:** biosynthesis · domain swapping · prenyltransferases · sesterterpenoids · terpene cyclases

**How to cite:** *Angew. Chem. Int. Ed.* **2016**, *55*, 1658–1661
*Angew. Chem.* **2016**, *128*, 1690–1693

[1] D. W. Christianson, *Chem. Rev.* **2006**, *106*, 3412–3442.
[2] K. Wang, S. Ohnuma, *Trends Biochem. Sci.* **1999**, *24*, 445–451.
[3] a) T. Toyomasu, M. Tsukahara, A. Kaneko, R. Niida, W. Mitsuhashi, T. Dairi, N. Kato, T. Sassa, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 3084–3088; b) T. Toyomasu, A. Kaneko, T. Tokiwano, Y. Kanno, Y. Kanno, R. Niida, S. Miura, T. Nishioka, C. Ikeda, W. Mitsuhashi, *J. Org. Chem.* **2009**, *74*, 1541–1548; c) R. Chiba, A. Minami, K. Gomi, H. Oikawa, *Org. Lett.* **2013**, *15*, 594–597; d) Y. Matsuda, T. Mitsuhashi, Z. Quan, I. Abe, *Org. Lett.* **2015**, *17*, 4644–4647; e) Y. Ye, A. Minami, A. Mandi, C. Liu, T. Taniguchi, T. Kuzuyama, K. Monde, K. Gomi, H. Oikawa, *J. Am. Chem. Soc.* **2015**, *137*, 11846–11853.
[4] a) I. H. Qureshi, S. A. Husain, R. Noorani, N. Murtaza, Y. Iitaka, S. Iwasaki, S. Okuda, *Tetrahedron Lett.* **1980**, *21*, 1961–1962; b) I. H. Sadler, T. J. Simpson, *J. Chem. Soc. Chem. Commun.* **1989**, 1602–1604; c) S. B. Singh, R. Reamer, D. Zink, D. Schmatz, A. Dombrowski, M. Goetz, *J. Org. Chem.* **1991**, *56*, 5618–5622; d) O. D. Hensens, D. Zink, J. M. Williamson, V. J. Lotti, R. S. Chang, M. A. Goetz, *J. Org. Chem.* **1991**, *56*, 3399–3403; e) H. Wei, T. Itoh, M. Kinoshita, Y. Nakai, M. Kurotaki, M. Kobayashi, *Tetrahedron* **2004**, *60*, 6015–6019.
[5] E. Y. Shishova, L. Di Costanzo, D. E. Cane, D. W. Christianson, *Biochemistry* **2007**, *46*, 1941–1951.
[6] a) Y. Matsuda, T. Iwabuchi, T. Wakimoto, T. Awakawa, I. Abe, *J. Am. Chem. Soc.* **2015**, *137*, 3393–3401; b) C. Liu, K. Tagami, A. Minami, T. Matsumoto, J. C. Frisvad, H. Suzuki, J. Ishikawa, K. Gomi, H. Oikawa, *Angew. Chem. Int. Ed.* **2015**, *54*, 5748–5752; *Angew. Chem.* **2015**, *127*, 5840–5844.
[7] T. R. Hoye, C. S. Jeffrey, F. Shao, *Nat. Protoc.* **2007**, *2*, 2451–2458.
[8] T. Roncal, S. Cordobés, U. Ugalde, Y. He, O. Sterner, *Tetrahedron Lett.* **2002**, *43*, 6799–6802.
[9] S. Frick, R. Nagel, A. Schmidt, R. R. Bodemann, P. Rahfeld, G. Pauls, W. Brandt, J. Gershenzon, W. Boland, A. Burse, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 4194–4199.
[10] J. K. Pawlak, M. S. Tempesta, T. Iwashita, K. Nakanishi, Y. Naya, *Chem. Lett.* **1983**, *12*, 1069–1072.
[11] M. Baunach, J. Franke, C. Hertweck, *Angew. Chem. Int. Ed.* **2015**, *54*, 2604–2626; *Angew. Chem.* **2015**, *127*, 2640–2664.

Received: October 4, 2015
Published online: November 5, 2015

  