A EUROPEAN JOURNAL OF CHEMICAL BIOLOGY

CHEMBIOCHEM

SYNTHETIC BIOLOGY & BIO-NANOTECHNOLOGY

Accepted Article

Title: Biosynthesis of Sesterterpene, Head-to-tail Triterpene, and Sesquarterpene in Bacillus clausii: Identification of Multifunctional Enzymes and Analysis of Isoprenoid Metabolites

Authors: Daijiro Ueda; Hiroaki Yamaga; Mizuki Murakami; Yusuke Totsuka; Tetsuro Shinada; Tsutomu Sato

This manuscript has been accepted after peer review and the authors have elected to post their Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). This work is currently citable by using the Digital Object Identifier (DOI) given below. The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

To be cited as: ChemBioChem 10.1002/cbic.201500138

Link to VoR: http://dx.doi.org/10.1002/cbic.201500138

A Journal of ChemPubSoc Europe

www.chembiochem.org

WILEY-VCH

Biosynthesis of Sesterterpene, Head-to-tail Triterpene, and Sesquarterpene in *Bacillus clausii*: Identification of Multifunctional Enzymes and Analysis of Isoprenoid Metabolites

Daijiro Ueda,[a] Hiroaki Yamaga,[a] Mizuki Murakami, [a] Yusuke Totsuka,[b] Tetsuro Shinada,[b] and Tsutomu Sato *[a]

Abstract: We performed functional analysis of recombinant enzymes and analysis of isoprenoid metabolites in *Bacillus clausii* to gain insights into the biosynthesis of rare terpenoids of sesterterpenes, head-to-tail triterpenes, and sesquarterpenes. We have identified an (all-*E*)-isoprenyl diphosphate synthases (E-IDS) homolog as a trifunctional geranylfarnesyl diphosphate (GFPP)/hexaprenyl diphosphate (HexPP)/heptaprenyl diphosphate (HepPP) synthase. In addition, we have redefined the function of tetraprenyl-*β*-curcumene synthase homolog as a trifunctional sesterterpene/triterpene/sesquarterpene synthase. The present study revealed that GFPP, HexPP, and HepPP, intermediates of two isoprenoid pathways (acyclic terpenes and menaquinones), are biosynthesized by one trifunctional E-IDS. In addition, GFPP/HexPP and HepPP are the primary substrates for biosynthesis of acyclic terpenes and menaquinone, respectively.

Introduction

Isoprenoids (terpenoids) make up the largest group of natural products, including more than 50,000 structurally diverse compounds.[1] They are categorized by the number of C_5 isoprene units: hemi- (one C_5 unit), mono- (C_{10}; two C_5 units), sesqui- (C_{15}; three C_5 units), di- (C_{20}; four C_5 units), sester- (C_{25}; five C_5 units), tri- (C_{30}; six C_5 units), and tetra-terpenes (C_{40}; eight C_5 units). [1] Isoprenoid biosynthesis begins with the condensation of isopentenyl diphosphate (IPP) into dimethylallyl diphosphate (DMAPP) by isoprenyl diphosphate synthases (IDSs).[2] (all-*E*)- and *E*,*Z*-mixed isoprenyl diphosphates are formed by E-IDSs and Z-IDSs, respectively, which have no sequence similarity with each other.[2] Recent studies have revealed a unique biosynthetic pathway for sesquarterpenes, the new family name for C_{35} terpenes.[3] Tetraprenyl-*β*-curcumene synthase (TS), which catalyzes the conversion of a linear C_{35} isoprenoid, (all-*E*)-heptaprenyl diphosphate (HepPP), into a monocycle in *Bacillus subtilis* and *Bacillus megaterium* (Scheme 1), is a new type of terpene cyclase with no sequence homology to any known terpene synthase.[3d] HepPP is biosynthesized with IPP and farnesyl diphosphate (FPP; C_{15}) by heterodimeric E-IDS (HepS/T), and is utilized to produce menaquinone-7 (MK-7) (Scheme 1). In addition, *B. megaterium* tetraprenyl-*β*-curcumene cyclase (TC) is a bifunctional enzyme that acts as an onoceroid synthase (Scheme 1).[3e, 4]

We recently reported that the TS homolog from *Bacillus clausii* (Bcl-TS) catalyzes the conversion of (all-*E*)-geranylfarnesyl diphosphate (GFPP) and (all-*E*)-hexaprenyl diphosphate (HexPP) into *β*-geranylfarnesene (**1**) and *β*-hexaprene (**2**), respectively, *in vitro* (Scheme 2) and that both compounds are produced in *B. clausii*.[5] Thus, we proposed that Bcl-TS is a sesterterpene (C_{25})/triterpene (C_{30}) synthase, and that TS homologs in many bacteria represent a new family of terpene synthases that form sesterterpene and triterpene as well as sesquarterpene.[5] However, it remains unclear whether GFPP and HexPP, which are the substrates for Bcl-TS, are biosynthesized in *B. clausii*. Biosynthesis of sesterterpenes and head-to-tail triterpenes is not fully understood. Only one sesterterpene synthase (ophiobolin F synthase), which utilizes GFPP as a substrate, has been identified in addition to Bcl-TS.[6] Many tail-to-tail triterpene synthases, squalene, and oxidosqualene cyclases, have been found from animal, plant, fungi, and bacteria.[7] However, a head-to-tail triterpene synthase that uses HexPP as a substrate has never been found, except for Bcl-TS. We recently synthesized unnatural head-to-tail pentacyclic triterpenes by the *in vitro* enzymatic reaction of TC with acyclic triterpene **2** (Scheme 2).[8]

All reported E-IDS types of GFPP,[9] HexPP,[10] and HepPP synthases[11] have been described as mono-functional enzymes. However, they form several products *in vitro*,[9-11] and isoprenoid metabolites *in vivo* have not been reported. In this paper, the biosynthesis of rare terpenoid groups of sesterterpene, head-to-tail triterpene, and sesquarterpene are described based on the identification of multifunctional enzymes and analysis of isoprenoid metabolites.

Results and Discussion

Cloning, expression, and purification of E-IDS1, E-IDS2-S, and E-IDS2-L

In order to identify GFPP and HexPP in *B. clausii*, which is essential for demonstrating that Bcl-TS is a sesterterpene/triterpene synthase, we performed functional analysis of all E-IDS homologs in *B. clausii*. The *B. clausii* genome possesses two E-IDS homologs, AB905204 and

[a] D. Ueda, H. Yamaga, M. Murakami, Prof. Dr. T. Sato,
Department of Applied Biological Chemistry, Faculty of Agriculture, and Graduate School of Science and Technology, Niigata University, Ikarashi 2–8050, Nishi-ku, Niigata 950–2181, Japan
E-mail: satot@agr.niigata-u.ac.jp

[b] Y. Totsuka, Prof. Dr. T. Shinada
Graduate School of Science, Osaka City University
Sugimoto, Sumiyoshi-ku, Osaka 558-8585, Japan

Supporting information for this article is given via a link at the end of the document.

Accepted Manuscript

AB735676. AB905204 and AB735676 are referred to as E-IDS1 and E-E-IDS2-L, respectively, in this paper. For example, multiple sequence alignments of E-IDS1 and FPP synthase from *Geobacillus stearothermophilus* [12] (identity: 52%) and E-IDS2-L and HepT from *B. subtilis* [11b] (identity: 53%) are shown in Fig. S1 and S3, respectively. The heterodimer type, which dimerizes with the E-IDS homolog (HepT, large subunit or component II) and non-homolog (HepS, small subunit or component I), was identified in *B. subtilis* (Scheme 1).[11b] *B. clausii* possesses a HepS homolog (AB735675: E-IDS2-S) in the gene cluster including E-IDS2-L. Multiple sequence alignments of E-IDS2-S and HepS from *B. subtilis* [11b] (identity: 19%) are shown in Fig. S2. Therefore, we also used E-IDS2-S to analyze the reaction. E-IDS1, E-IDS2-S, and E-IDS2-L were cloned into the pCold I vector, expressed as a soluble protein, and purified by Ni affinity chromatography (Fig. 1).

Analysis of products formed by E-IDS1, E-IDS2-S, and E-IDS2-L

To characterize the pathway from DMAPP, which is a starter allylic compound for isoprenoid biosynthesis in all organisms, to GFPP and HexPP, the reaction of DMAPP and [1-^{14}C]IPP with E-IDS1, E-IDS2-S, and E-IDS2-L was analyzed by radio-TLC (Fig. 2). DMAPP was converted into FPP (C_{15}) and geranylgeranyl diphosphate (GGPP; C_{20}) by E-IDS1, and into GFPP (C_{25}) and HexPP (C_{30}) by E-IDS2-S/L (mixture of E-IDS2-S and E-IDS2-L) (Fig. 2 and Fig. 3A).

Next, the reaction of FPP or GGPP, produced from DMAPP by E-IDS1, with E-IDS2-S/L was analyzed by radio-TLC (Fig. 2). E-IDS2-S/L produced GFPP (C_{25}) and HexPP (C_{30}) from FPP or GGPP (Fig. 2). Thus, GFPP and HexPP are also biosynthesized from FPP and GGPP by E-IDS2-S/L in *B. clausii* (Fig. 3A).

We thus demonstrated that GFPP and HexPP, the Bcl-TS substrates, are biosynthesized by E-IDS2-S/L in *B. clausii* (Fig. 3A). E-IDS2-S or E-IDS2-L alone formed no product from DMAPP, FPP, and GGPP (data not shown); thus, formation of GFPP and HexPP is mediated by heteromeric E-IDS2-S/L in *B. clausii*.

Isolation and structural analysis of novel acyclic sesquarterpene 3

Unexpectedly, C_{35} HepPP production was detected in the *in vitro* reaction of E-IDS2-S/L (Fig. 2 and Fig. S4). To search for the sesquarterpene metabolite, *n*-hexane extracts of *B. clausii* were re-analyzed by GC-MS. Injection of *n*-hexane extracts 10-fold larger than those previously reported [5] yielded minor peak **3** (Fig. 4A). We detected no sesquiterpene (C_{15}) or diterpene (C_{20}). To determine the structure of **3**, *B. clausii* was cultured in 20 flasks containing 1 L medium each, and compound **3** was isolated by silica gel column chromatography and SiO_2 HPLC to yield 1.1 mg of purified **3**.

The structure of **3** was determined by NMR (^{1}H, ^{13}C, DEPT, COSY, NOESY, HSQC, and HMBC) and EI-MS (Fig. 5 and Table 1). The molecular formula was identified by HR-EI-MS as $C_{35}H_{56}$. The partial structure (right side in Fig. 5B) of **3** (C-1–C-5 and C-29) was assigned according to data obtained from HMBC and COSY. The methylidene protons of **3** at δ_H 5.34 (d, J = 17.7 Hz) and δ_H 5.11 (d, J = 11.0 Hz) exhibited a strong HMBC correlation with C-2 (δ_C 139.4, d) and C-3 (δ_C 146.5, s), indicating that these protons are assignable to H-1. H-2 (δ_H 6.50) exhibited a dd splitting pattern (J = 17.6 and 10.7 Hz) due to the spin-spin couplings with H-1a (trans) and H-1b (cis). The methylidene protons at δ_H 5.12 (s) and δ_H 5.11 (s) were assigned to H-29 because of the clear cross-peak with C-2, C-3, and C-4 (δ_C 31.9, t) in the HMBC spectrum. Thus, the right side of **3** shown in Fig. 5B contains an isoprene moiety. We determined the structure of the isobutenyl moiety on the left side of the compound, as shown in Fig. 5B (C-26, C-27, C-28, and C-35), by HMBC spectral analysis. HMBC data showed five isopentenyl moieties (represented by squares with broken lines in Fig. 5B) linked to each other. The complete structure of **3** was elucidated by connecting the three partial structures obtained from HMBC cross peaks between H-5 (δ_H = 2.40 ppm) and C-6 (δ_C = 124.6–125.0 ppm) and those between H-25 (δ_H = 2.30–2.35 ppm) and C-27 (δ_C = 131.1 ppm). No NOE was observed for H-6 (δ_H 5.37–5.43)/H-30 (δ_H 1.69, s), H-10 (δ_H 5.37–5.43)/H-31 (δ_H 1.74, s), H-14 (δ_H 5.37–5.43)/H-32 (δ_H 1.74, s), H-18 (δ_H 5.37–5.43)/H-33 (δ_H 1.74, s), and H-22 (δ_H 5.37–5.43)/H-34 (δ_H 1.74, s), indicating that all prenyl residues in **3** are arranged in the *E* configuration. Comparison of the chemical shifts (δ_C 40.2 ppm) of C-8, C-12, C-16, C-20, and C-24 with those of ficaprenol[13] (*Z*, ca. 32.0 ppm; *E*, ca. 40.0 ppm) further supports the *E* geometry. The complete assignments of ^{1}H and ^{13}C NMR data for **3** unambiguously supported this structure (Table 1). While *β*-myrcene (C_{10}), *β*-farnesene (C_{15}), *β*-springene (C_{20}), **1**, and **2** have been identified as acyclic terpene hydrocarbons containing a similar isoprene moiety, compound **3** appears to be a novel sesquarterpene. We propose that **3** be named *β*-heptaprene.

To determine the amount of **1**–**3** produced by *B. clausii*, we performed GC analysis of *n*-hexane extracts. As shown in Fig. 3B, the amounts of **1**–**3** per 1 g dry cell weight were 0.96, 0.66, and 0.065 mg, respectively. The production ratios (%) of compounds **1**–**3** were as follows: **1**:**2**:**3** = C_{25}:C_{30}:C_{35} = 57:39:4 (Fig. 3A).

Analysis of menaquinones in *B. clausii*

(all-*E*)-Isoprenoids with chain lengths ranging from C_{25} to C_{35} are essential in cell maintenance as the hydrophobic side chains of quinones. *B. subtilis* produces MK-7, which possesses a C_{35} isoprenoid side chain; thus, we hypothesized that *B. clausii* may also produce menaquinone. To gain further insights into the biosynthetic pathways of C_{25}, C_{30}, and C_{35} isoprenoids, we performed LC-ESI-MS and HPLC of *n*-hexane extracts from *B. clausii* to assess menaquinone production. Since E-IDS2-S/L formed GFPP (C_{25}) and HexPP (C_{30}) as the major products (Fig. 2), and compounds **1** (C_{25}) and **2** (C_{30}) were mainly produced as acyclic terpenes (Fig. 3B), we expected MK-5 and MK-6, which possess C_{25} and C_{30} isoprenoids, are the primary products of *B. clausii*. However, MK-7 was the major menaquinone (Fig. 3B). The amounts of MK-5, MK-6, and MK-7 per 1 g dry cell weight were 0, 0.035, and 0.365 mg, respectively (Fig. 3B). The production ratios (%) of isoprenoids in the acyclic terpene and

Accepted Manuscript

menaquinone family were largely different: MK-5:MK-6:MK-7 = C_{25}:C_{30}:C_{35} = 0:9:91 (Fig. 3A). In addition, the amount of MK-7 (0.365 mg) was approximately 5.6-fold higher than that of **3** (0.065 mg) (Fig. 3B). These results suggest GFPP (C_{25}) and HexPP (C_{30}) produced by E-IDS2-S/L are mainly utilized for acyclic terpene (**1** and **2**) biosynthesis, while HepPP (C_{35}) is used for MK-7 synthesis (Fig. 3A).

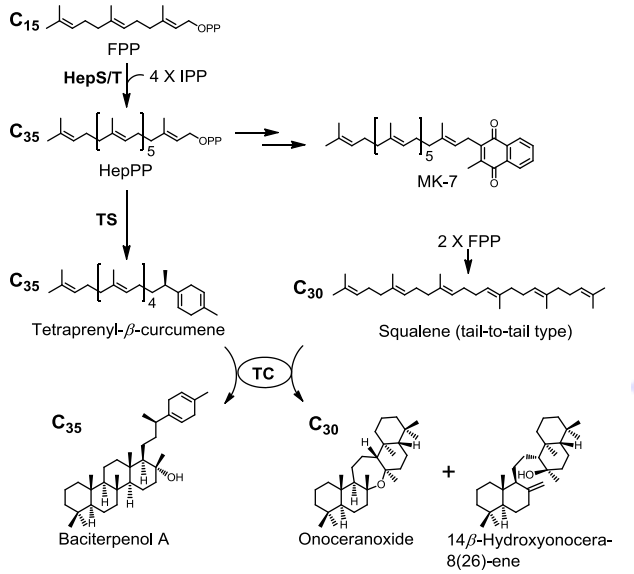
Multiple functions of Bcl-TS, E-IDS2-S/L, and E-IDS1

To demonstrate that compound **3** is formed from HepPP by Bcl-TS, chemically synthesized HepPP was incubated with purified Bcl-TS. As shown in Fig. 4B and 4C, product **3** was detected by GC-MS. Thus, Bcl-TS produces **3** as a minor product (Fig. 3B). We thus have redefined the function of Bcl-TS as a trifunctional sesterterpene/triterpene/sesquarterpene synthase. Since this study revealed that Bcl-TS can accept HepPP (C_{35}) as a substrate, mutational analysis targeting residues that differ in TS, which cyclizes HepPP (C_{35}) to a monocycle (Scheme 1),[3d] will help establish the mechanism that controls the formation of acyclic and cyclic structures.

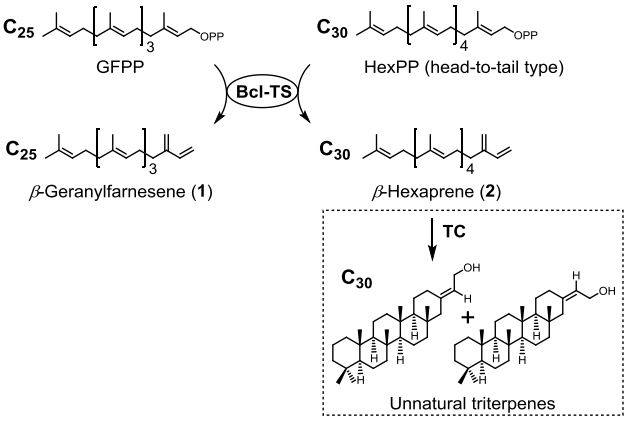
Detection of **3** and MK-7 in *B. clausii* cells and *in vitro* conversion (DMAPP→HepPP→**3**) indicated that HepPP is biosynthesized by E-IDS2-S/L in *B. clausii* (Fig. 3A). The amounts of C_{25}, C_{30}, and C_{35} metabolites (sum of acyclic terpenes and MKs) did not differ substantially (Fig. 3B). Thus, E-IDS2-S/L is the first identified trifunctional GFPP/HexPP/HepPP synthase. We have identified the first bifunctional *Z*-IDS, which forms *Z*-HepPP (C_{35}) and *Z*-decaprenyl diphosphate (C_{50}), from biosynthetic study of mycobacterial sesquarterpenes.[3h] In contrast, all reported E-IDS types of GFPP,[9] HexPP,[10] and HepPP synthases [11] have been described as mono-functional enzymes. Since they form several products *in vitro*,[9-11] they may also be multifunctional. Further study of the mono-functional E-IDSs [9-11] through analysis of isoprenoid metabolites may lead to the discovery of multifunctional enzymes and novel compounds such as sesterterpene, head-to-tail triterpene, and sesquarterpene.

Most FPP and GGPP synthases produce single products *in vitro*. In contrast, two E-IDSs from archaea and protozoans form both FPP and GGPP *in vitro* and have been identified as bifunctional FPP/GGPP synthases.[14] Since E-IDS1 also produced comparable amounts of FPP and GGPP (Fig. 2), we propose that E-IDS1 be considered a bifunctional FPP/GGPP synthase (Fig. 3A). To the best of our knowledge, this is the first report of a bifunctional FPP/GGPP synthase in bacteria.

The differences in isoprenoid (acyclic terpenes and MKs) distribution in *B. clausii* (Fig. 3B) may be determined by the substrate specificity of Bcl-TS, E-IDS2-S/L, and prenyltransferase of the new menaquinone biosynthetic pathway,[15] which has not been previously identified in any organisms (Fig. 3A). Further experiments are required to establish the regulatory mechanism of the acyclic terpene and MK pathways.



Scheme 1. Biosynthetic pathway of tail-to-tail triterpenes (C_{30}), sesquarterpenes (C_{35}), and MK-7 in *B. subtilis* and *B. megaterium*. The conversion of squalene to onoceroid by TC occurs only in *B. megaterium*. HepS/T: heterodimeric heptaprenyl diphosphate synthase; TS: tetraprenyl-β-curcumene synthase; TC: tetraprenyl-β-curcumene cyclase; MK-7: menaquinone-7.



Scheme 2. Proposed pathways for the biosynthesis of acyclic sesterterpene and head-to-tail triterpene in *B. clausii*. Dotted square shows the *in vitro* reaction. Bcl-TS: tetraprenyl-β-curcumene synthase homolog.

Accepted Manuscript

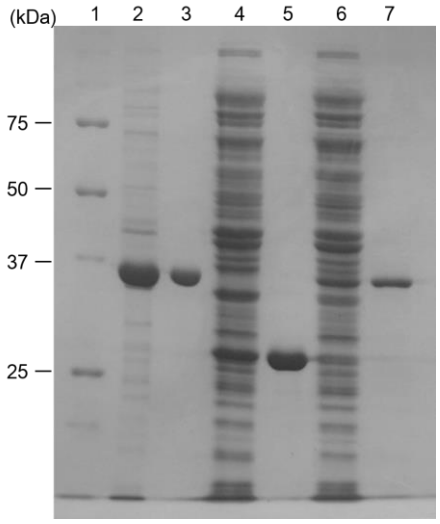


Fig. 1. SDS-PAGE of the overexpressed and purified recombinant enzymes. Lane 1: molecular weight marker; lanes 2, 4, and 6: total soluble proteins from BL21(DE3)/pColdI-E-IDS1, BL21(DE3)/pColdI-E-IDS2-S, and BL21(DE3)/pColdI-E-IDS2-L, respectively; lanes 3, 5, and 7: purified E-IDS1, E-IDS2-S, and E-IDS2-L, respectively.

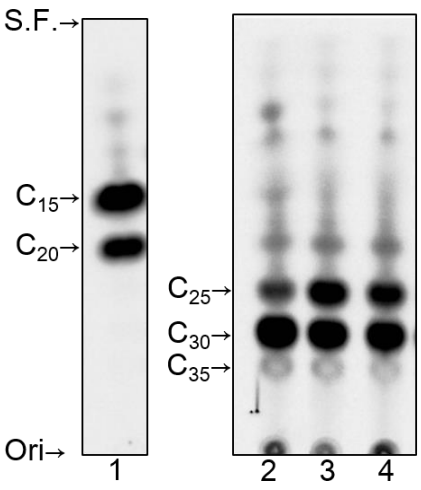


Fig. 2. TLC analysis of radiolabeled products prepared with recombinant E-IDSs. Lane 1: reaction products synthesized by incubation of DMAPP and [1-^{14}C]IPP with E-IDS1; lanes 2–4: reaction products synthesized by incubation of allylic substrate (lane 2: DMAPP; lane 3: FPP; lane 4: GGPP) and [1-^{14}C]IPP with E-IDS2-S/L. Only E-IDS2-S and E-IDS2-L formed no product from the allylic substrate (DMAPP, FPP, and GGPP) and [1-^{14}C]IPP. Ori.: origin; S.F.: solvent front. This figure represents the reaction with 133.6 µM [1–^{14}C]IPP. All reaction products of 15 µM [1-^{14}C]IPP with E-IDS1 or E-IDS2-S/L produced similar results. In Fig. S4, Fig.2 (lanes 2–4) with enhanced intensity has been provided; in this figure, the C_{35} spot can be seen more clearly.

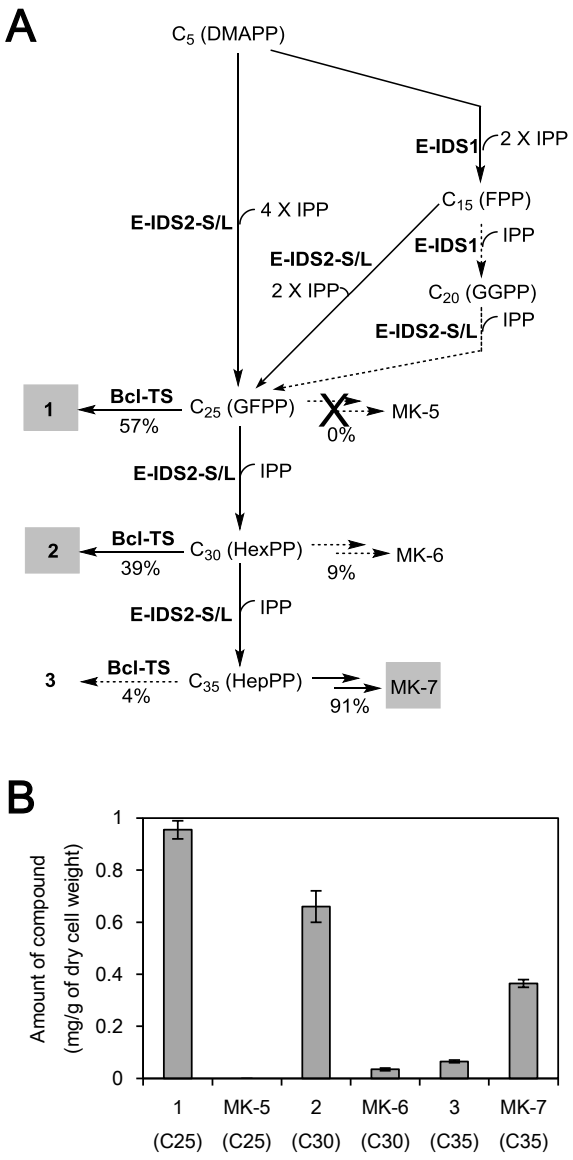


Fig. 3. Proposed pathways for the biosynthesis of acyclic terpenes and menaquinones (A) and acyclic terpenes and menaquinone production by *B. clausii* (B). Major and minor pathways are represented by solid and dotted arrows, respectively in A. Main metabolites (**1**, **2**, and MK-7) are shown by gray boxes in A. The information provided in parentheses in B indicates isoprenoid carbon numbers.

Accepted Manuscript

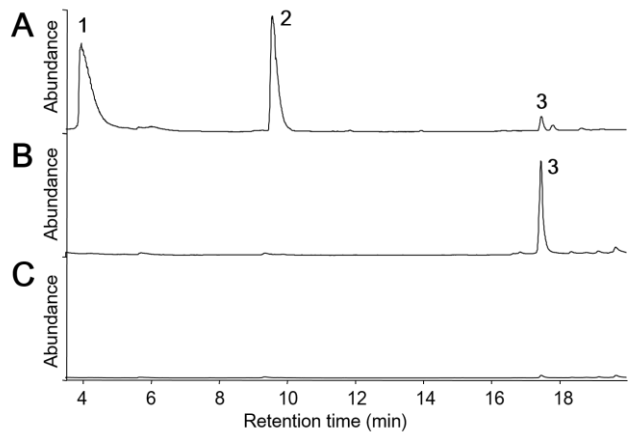


Fig. 4. GC-MS analysis of acyclic terpenes. A) *n*-Hexane extract from *B. clausii* cells. B) Reaction product formed by incubation of HepPP with purified Bcl-TS. C) Reaction product formed by incubation of HepPP with heat-denatured Bcl-TS. MK-5, MK-6, and MK-7 were not detected under these conditions.

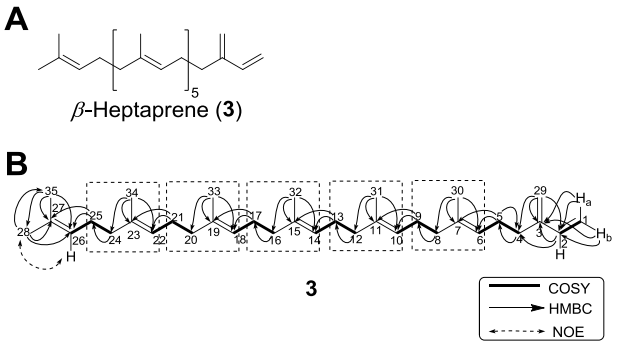


Fig. 5. Structure of **3** (A) and key correlations in the 2D-NMR data of **3** (B).

Table 1. NMR assignments for **3** in C_6D_6

No.	δ_H (mult., *J* in Hz)	δ_C (mult.)
1	H_a: 5.34 (1H, d, 17.7) H_b: 5.11 (1H, d, 11.0)	113.1 (t)
2	6.50 (dd, 17.6, 10.7)	139.4 (d)
3	–	146.5 (s)
4	2.38 (m)	31.9 (t)
5	2.40 (m)	27.1–27.3 (t)
6	5.37–5.43 (m)	124.6–125.0 (d)
7	–	135.0–135.4 (s)
8	2.21–2.25 (m)	40.2 (t)
9	2.30–2.35(m)	27.1–27.3 (t)
10	5.37–5.43 (m)	124.6–125.0 (d)
11	–	135.0–135.4 (s)
12	2.21–2.25 (m)	40.2 (t)
13	2.30–2.35(m)	27.1–27.3 (t)
14	5.37–5.43 (m)	124.6–125.0 (d)
15	–	135.0–135.4 (s)
16	2.21–2.25 (m)	40.2 (t)
17	2.30–2.35 (m)	27.1–27.3 (t)
18	5.37–5.43 (m)	124.6–125.0 (d)
19	–	135.0–135.4 (s)
20	2.21–2.25 (m)	40.2 (t)
21	2.30–2.35(m)	27.1–27.3 (t)
22	5.37–5.43 (m)	124.6–125.0 (d)
23	–	135.0–135.4 (s)
24	2.21–2.25 (m)	40.2 (t)
25	2.30–2.35 (m)	27.1–27.3 (t)
26	5.37–5.43 (m)	124.6–125.0 (d)
27	–	131.1 (s)
28	1.81 (3H, s)	25.8 (q)
29	5.11 (1H, s) 5.12 (1H, s)	115.9 (t)
30	1.69 (3H, s)	16.1 (q)
31	1.74 (3H, s)	16.1 (q)
32	1.74 (3H, s)	16.1 (q)
33	1.74 (3H, s)	16.1 (q)
34	1.74 (3H, s)	16.1 (q)
35	1.69 (3H, s)	17.7 (q)

Accepted Manuscript

Conclusions

Functional analysis of recombinant enzymes and analysis of isoprenoid metabolites revealed that Bcl-TS, E-IDS1, and E-IDS2-S/L are the trifunctional sesterterpene/triterpene/sesquarterpene synthase, bifunctional FPP/GGPP synthase, and trifunctional GFPP/HexPP/HepPP synthase, respectively (Fig. 3A). To our knowledge, this is the first report to identify the trifunctional sesterterpene/triterpene/sesquarterpene synthase and trifunctional GFPP/HexPP/HepPP synthase. In particular, the present study suggests that isoprenoid metabolite analysis is crucial for the identification of multifunctional E-IDSs. GFPP, HexPP, and HepPP are intermediates in two isoprenoid pathways (acyclic terpenes and menaquinones) and are biosynthesized by one trifunctional E-IDS. In addition, GFPP/HexPP and HepPP are the main substrates for biosynthesis of acyclic terpenes and menaquinone, respectively (Fig. 3A). The present study will promote the search for and creation of novel sesterterpenes, head-to-tail triterpenes, and sesquarterpenes.

Experimental Section

Materials

DMAPP, FPP, and GGPP were diphosphorylated from 3-methyl-2-buten-1-ol, farnesol, and geranylgeraniol, respectively, as described by Davisson et al. [16]. GFPP and HexPP were prepared as previously described[5]. HepPP was diphosphorylated from (all-*E*)-heptaprenol, which was chemically synthesized as described in the Supporting Information. [1-^{14}C]IPP was purchased from GE Healthcare, USA. All other chemicals were of analytical grade.

Cloning, expression, and purification of E-IDS1, E-IDS2-S, and E-IDS2-L

The following sets of primers and *B. clausii* JCM9138 genomic DNA were used to PCR amplify three genes (E-IDS1: ABC2463, E-IDS2-S: ABC1885; and E-IDS2-L: ABC1889), including approximately 200-bp regions upstream and downstream of the open reading frames: E-IDS1-S, 5′-actgaaagtcgagaagcagcttg-3′; E-IDS1-A, 5′-ggttcagactgcaggcaaacgctc-3′; E-IDS2-S-S, 5′-GGATACGGAGGAATCAGGCGC-3′; E-IDS2-S-A, 5′-TCATTGTATCTTTACGCCAGC-3′; E-IDS2-L-S, 5′-CTGCTTCATCAGCCAACAAGC-3′; E-IDS2-L-A, 5′-CGTTCCATTGTCTAACCTCGC-3′. Amplified DNA fragments were inserted into pGEM-T (Promega) and the constructs were sequenced. The DNA sequences of *e-ids1*, *e-ids2-s*, and *e-ids2-l* from *B. clausii* JCM9138 were deposited in the DNA Data Bank of Japan under accession numbers AB905204, AB735675, and AB735676. The amino acid sequences of E-IDS1, E-IDS2-S, and E-IDS2-L from *B. clausii* JCM9138 showed 96%, 95%, and 98% identities to those from *B. clausii* KSM-K16, for which complete sequence data are available. The following primer pairs and the pGEM-T derivatives of *e-ids1*, *e-ids2-s*, and *e-ids2-l* were used to add restriction sites: 2-E-IDS1-S, (KpnI), 5′-ggaggcggtaccATGAGCGCATTGTTCCAA ACG-3′; 2-E-IDS1-A, (XbaI), 5′-cctgtttctagaTTAATGGTCACGCGTAACGATG-3′; 2-E-IDS2-S-S, (EcoRI), 5′-gggacagaattcATGAGCAATACAACGTTGCCG-3′; 2-E-IDS2-S-A (XbaI), 5′-tgggcctctagaCTACCTCCTTGCTTTTTCAAG-3′; 2-E-IDS2-L-S (EcoRI), 5′-agggtggAAtTcATGAAGTTAGCAGATATTTATAAA G-3′; 2-E-IDS2-L-A (*Xba*I), 5′-CTCGTCTCTAGATTAGTATTTTCTCCGTCCTATG-3′. Amplified DNA fragments were inserted into pGEM-T. Fragments (*Kpn*I/*Xba*I or *Eco*RI/*Xba*I) were inserted into pColdI. *E. coli* BL21(DE3) harboring pColdI derivatives (pColdI-E-IDS1, pColdI-E-IDS2-S, and pColdI-E-IDS2-L) were grown at 37°C in Luria-Bertani medium containing 100 μg mL^{-1} ampicillin. Expression of the recombinant protein was induced by addition of 0.1 mM IPTG at OD_{600} 0.6, followed by culture for an additional 24 h at 15°C. *E. coli* expressing recombinant enzymes were harvested by centrifugation and resuspended in buffer containing 20 mM Tris-HCl (pH 7.9), 10 mM imidazole, and 300 mM NaCl. Cells were disrupted by sonication in a UP200s sonicator (Hielscher Ultrasonics GmbH) at 4–10°C for 15 min, and then centrifuged at 10,000 × *g* for 20 min. The pellet was discarded and the supernatant containing soluble His-tagged fusion enzyme was purified according to manufacturer protocols (Takara). Expression and purification of E-IDS1, E-IDS2-S, and E-IDS2-L were analyzed by 10% SDS-PAGE (Fig. 1).

Analysis of products formed by E-IDS1, E-IDS2-S, and E-IDS2-L

The reaction mixture for analyzing the chain length of products formed by E-IDS1, E-IDS2-S, and E-IDS2-L contained 25 mM Tris-HCl buffer (pH 8.5), 10 mM dithiothreitol, 1 mM $MgCl_2$, 25 mM NH_4Cl, 133.6 or 15 μM [1–^{14}C]IPP, 15 μM allylic substrate (DMAPP, FPP or GGPP), and 3.0 μM purified enzyme, in a total volume of 50 μL. The reaction was performed at 37°C for 1 h and terminated by boiling. Polyprenyl diphosphate products were hydrolyzed to the corresponding alcohols by using potato acid phosphatase (Sigma). A phosphatase reaction mixture (50 μL of reaction mixture, 30 μL of 2-propanol, 15 μL of 1% Triton X-100, 53 μL of 3 M acetate buffer, and 40 μg of potato acid phosphatase) was incubated at 37°C for 12 h. After addition of 150 μL MeOH, the *n*-hexane extract (100 μL) was analyzed by TLC on a reverse-phase RP-18 plate (Merck) with a solvent system of MeOH:acetone (8:2). Radioactive spots were detected using an image analyzer (FLA-9000; Fujifilm).

Analysis of isoprenoids in *B. clausii* cells

B. clausii was cultured under shaker conditions at 37°C for 4 d in 1 L Horikoshi-I medium, as recommended by JCM (http://www.jcm.riken.go.jp/). Cells were then collected by centrifugation (6,000 × *g*, 4°C, 20 min). After lyophilization, cell extracts were made in MeOH (3× 200 mL) and the MeOH fractions were partitioned by using *n*-hexane (3× 600 mL). To analyze the acyclic terpenes **1**–**3**, GC-MS (injection temperature, 300°C; oven temperature, 220–300°C at an increment of 3°C min^{-1}) was performed on a JMS-T100GCV spectrometer (JEOL) equipped with a DB-1 capillary column (30 m × 0.25 mm × 0.25 μm; J&W Scientific, Inc.) by using the EI mode operated at 70 eV. GC analysis (injection temperature, 300°C; oven temperature, 220–300°C at an increment of 3°C min^{-1}) was performed twice to determine the amounts of **1**–**3** on a Shimadzu GC-2014 chromatograph fitted with a flame ionization detector (30 m × 0.32 mm × 0.25 μm; J&W Scientific, Inc.). To analyze the menaquinones (MK-5, MK-6, and MK-7), LC-MS was performed on a JMS-T100LP spectrometer (JEOL), using ESI mode. HPLC (MeOH/2-propanol 6:1, flow = 1.0 mL/min) was performed twice to determine the amount of menaquinones with an SPD-10A detector (270 nm) and LC-10AS pump (Shimadzu) equipped with a C_{18} reverse-phase column (SHISEIDO, CAPCELL PAK C_{18}). Authentic MK-5, MK-6, and MK-7 isolated from *B. krulwichiae* [17] were detected at 9.3, 13.6, and 20.6 min, respectively. Molecular ions $[M+Na]^+$ of MK-5, MK-6, and MK-7 were *m*/*z* 535, 603, and 671, respectively. Isolated MK-6 and MK-7 from *B. clausii* was found to have fragmentation patterns identical to those of the authentic samples on performing direct-EI-MS, including the characteristic *m*/*z* 225, which supported methylnaphtoquinone substructure.

Accepted Manuscript

Isolation and structural analysis of 3

B. clausii was cultured in 20 × 1 L medium at 37°C for 4 d. Lyophilized cells (dry weight, 37 g) were extracted using MeOH (3 × 800 mL), and the MeOH fraction was concentrated to 800 mL and partitioned using *n*-hexane (3 × 800 mL). The *n*-hexane layer (295.6 mg) was partially purified by silica gel (20 g) column chromatography with *n*-hexane and EtOAc. The fraction eluted with *n*-hexane (48.3 mg) was subjected to further silica gel (20 g) column chromatography with *n*-hexane and SiO_2 HPLC with *n*-hexane, yielding 1.1 mg of purified **3**. The structure of **3** was determined by NMR and MS. *β*-Heptaprene (**3**): Colorless oil; ^{1}H and ^{13}C NMR assignments are shown in Table 1. EI-MS *m*/*z* (%): 69 (100), 81 (65), 93 (32), 107 (20), 135 (17), 161 (15), 203 (10), 271 (1), 339 (1), 407 (1), 476 ([*M*]$^{+}$, 1). HR-EI-MS: *m*/*z* calcd for $C_{35}H_{56}$: 476.4390 [*M*]$^{+}$; found: 476.4382. Mass spectrum (EI) and NMR data of **3** are shown in Fig. S5 and S6-S10, respectively.

Enzyme assay of Bcl-TS, using HepPP as the substrate

Bcl-TS was purified as described previously.[5] The Bcl-TS assay reaction mixture contained 25 mM Tris-HCl buffer (pH 8.5), 10 mM dithiothreitol, 1 mM $MgCl_2$, 25 mM NH_4Cl, 65.5 µM HepPP, and 2.0 µM purified Bcl-TS in a total volume of 2 mL. The reaction was performed at 37°C for 16 h. A 15% KOH/MeOH solution (2.4 mL) was added to terminate the reaction, and the mixture was then heated at 80°C for 30 min. The lipophilic product was extracted with *n*-hexane (3 × 4 mL) and analyzed by GC-MS.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Numbers 25450149 and 25108709. We thank Dr. M. Fujihashi (Kyoto University) for the helpful discussions.

Keywords: isoprenoids • isoprenyl diphosphate synthase • prenyltransferase • terpene synthase • menaquinone

[1] a) L. Buckingham, "*Dictionary of Natural Products*", online web ed., Chapman & Hall/CRX Press, London **2002**. b) D.W. Christianson, *Science* **2007**, *316*, 60–61.

[2] a) K. Ogura, T. Koyama, *Chem. Rev.* **1998**, *98*, 1263–1276. b) K.C. Wang, S. Ohnuma, *Biochim. Biophys. Acta*. **2000**, *1529*, 33–48. c) P.L. Liang, T. Ko, A.H. Wang, *Eur. J. Biochem*. **2002**, *269*, 3339–3354.

[3] a) T. Bosak, R.M. Losick, A. Pearson, *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105***,** 6725–6729. b) R. Kontnik, T. Bosak, R.A. Butcher, J.J. Brocks, R. Losick, J. Clardy, A. Pearson, *Org. Lett.* **2008**, *10*, 3551–3554. c) H. Takigawa, M. Sugiyama, Y. Shibuya, *J. Nat. Prod*. **2010**, *73*, 204–207. d) T. Sato, S. Yoshida, H. Hoshino, M. Tanno, M. Nakajima, T. Hoshino, *J. Am. Chem. Soc.* **2011**, *133*, 9734–9737. e) T. Sato, H. Hoshino, S. Yoshida, M. Nakajima, T. Hoshino, *J. Am. Chem. Soc.* **2011**, *133*, 17540–17543. f) T. Sato, A. Kigawa, R. Takagi, T. Adachi, T. Hoshino, *Org. Biomol. Chem.* **2008***,* *6*, 3788–3794. g) T. Sato, R. Takagi, Y. Orito, E. Ono, T. Hoshino, *Biosci. Biotechnol. Biochem*. **2010**, *74*, 147–151. h) T. Sato, K. Takizawa, Y. Orito, H. Kudo, T. Hoshino, *Chembiochem.* **2010**, *11*, 1874–1881. i) T. Sato, E. Ono, M. Nakajima, C. Nakano, T. Hoshino, *Tetrahedron Lett*. 2012, *53*, 2522–2524 (2012). j) T. Sato, *Biosci. Biotechnol. Biochem*. **2013**, *77*, 1155–1159.

[4] D. Ueda, T. Hoshino, T. Sato, *J. Am. Chem. Soc.* **2013**, *135*, 18335–18338.

[5] T. Sato, H. Yamaga, S. Kashima, Y. Murata, T. Shinada, C. Nakano, T. Hoshino, *Chembiochem.* **2013**, *14*, 822–825.

[6] R. Chiba, A. Minami, K. Gomi, H. Oikawa, *Org. Lett.* **2013**, *15*, 594–597.

[7] a) T. Hoshino, T. Sato, *Chem. Commun.* **2002**, 291–301. b) R. Xu, G.C. Fazio, S.P.T. Matsuda, *Phytochemistry* **2004**, *65*, 261–291. c) I. Abe, *Nat. Prod. Rep*. **2007**, *24*, 1311–1331. d) G. Siedenburg, D. Jendrossek, *Appl. Environ. Microbiol*. **2011**, *77*, 3905–3915. e) S.C. Hammer, P. Syre, M. Seitz, B.M. Nestl, B. Hauer, *Curr. Opin. Chem. Biol.* **2013**, *17*, 293–300.

[8] W. Okamoto, T. Sato, *Tetrahedron Lett.* **2013**, *54*, 6747–6750.

[9] a) A. Tachibana, *FEBS Lett*. **1994**, *341*, 291–294. b) A. Tachibana, Y. Yano, S. Otani, N. Nomura, Y. Sako, M. Taniguchi, *Eur. J. Biochem*. **2000**, *267*, 321–328. c) T. Ogawa, T. Yoshimura, H. Hemmi, *Biochem. Biophys. Res. Commun*. **2010**, *393*, 16–20.

[10] a) M.N. Ashby, P.A. Edwards, *J. Biol. Chem*. **1990**, *265*, 13157–13164. b) N. Shimizu, T. Koyama, K. Ogura, *J. Bacteriol*. **1998**, *180*, 1578–1581. c) H. Hemmi, S. Ikejiri, S. Yamashita, T. Nishino, *J. Bacteriol.* **2002**, *184*, 615–620.

[11] a) A. Koike-Takeshita, T. Koyama, S. Obata, K. Ogura, *J. Biol. Chem.* **1995**, *270*, 18396–18400. b) Y. Zhang, T. Koyama, D.M. Marecak, G.D. Prestwich, Y. Maki, K. Ogura, *Biochemistry* **1998**, *37*, 13411–13420. c) C. Ohto, C. Ishida, H. Nakane, M. Muramatsu, T. Nishino, *Plant Mol. Biol.* **1999**, *40*, 307–321.

[12] T. Koyama, S. Obata, M. Osabe, A. Takahashi, K. Yokoyama, M. Uchida, T. Nishino, K. Ogura, *J. Biochem.* **1993**, *113*, 355–363.

[13] T. Tanaka, M. Takagi, *Biochem. J.* **1979**, *183*, 163–165.

[14] a) S. Fujiwara, A. Yamanaka, K. Hirooka, A. Kobayashi, T. Imanaka, E. Fukusaki, *Biochem. Biophys. Res. Commun.* **2004**, *325*, 1066–1074. b) Y. Ling, Z. Li, K. Miranda, E. Oldfield, S.N.J. Moreno, *J. Biol. Chem*. **2007**, *282*, 30804–30816.

[15] T. Dairi, *J. Antibiot.* **2009**, *62*, 347–352.

[16] V.J. Davisson, A.B. Woodside, C.D. Poulter, *Methods Enzymol.* **1985**, *110*, 130–144.

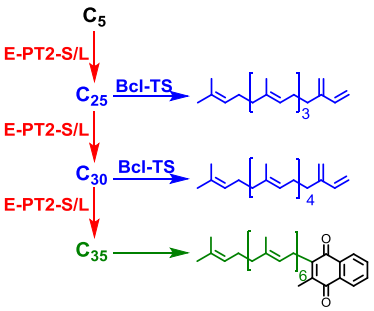
[17] I. Yumoto, S. Yamada, Y. Sogabe, Y. Nodasaka, H. Matsuyama, K. Nakajima, A. Suemori, *Int. J. Syst. Evol. Microbiol.* **2003**, *53*, 1531–1536.

Accepted Manuscript

Entry for the Table of Contents

FULL PAPER

Geranylfarnesyl diphosphate (GFPP), hexaprenyl diphosphate (HexPP), and heptaprenyl diphosphate (HepPP) are intermediates in two isoprenoid pathways (acyclic terpenes and menaquinones) and are biosynthesized by one trifunctional (all-*E*)-isoprenyl diphosphate synthase. GFPP/HexPP and HepPP are mainly utilized for the biosynthesis of acyclic terpenes and menaquinone, respectively.



Daijiro Ueda, Hiroaki Yamaga, Mizuki Murakami, Yusuke Totsuka, Tetsuro Shinada, and Tsutomu Sato *

Page No. – Page No.

Biosynthesis of Sesterterpene, Head-to-tail Triterpene, and Sesquarterpene in *Bacillus clausii*: Identification of Multifunctional Enzymes and Analysis of Isoprenoid Metabolites

Accepted Manuscript