

ORIGINAL ARTICLE

Chemical diversity of labdane-type bicyclic diterpene biosynthesis in Actinomycetales microorganisms

Yuuki Yamada, Mamoru Komatsu and Haruo Ikeda

Five pairs of bacterial type-A and type-B diterpene synthases have been characterized: BAD86798/BAD86797, AHK61133/AHK61132, BAB39207/BAB39206, CldD/CldB and Rmnd/RmnB, and are involved in the formation of pimara-9(11),15-diene, terpenite-3,13,15-triene and labda-8(17),12(E),14-triene. Mining of bacterial genome data revealed an additional four pairs of type-A and type-B diterpene synthases: Sros_3191/Sros_3192 of *Streptosporangium roseum* DSM 43021, Sare_1287/Sare_1288 of *Salinispora arenicola* CNS-205, SCLAV_5671/SCLAV_5672 and SCLAV_p0491/SCLAV_p0490 of *Streptomyces clavuligerus* ATCC 27064. Since SCLAV_p0491/SCLAV_p0490 is similar to the labdane-type diterpene synthase pairs, CldD/CldB and Rmnd/RmnB based on the alignment of the deduced amino acid sequences and phylogenetic analyses of the aligned sequences, these predicted diterpene synthases were characterized by an enzymatic reaction using a pair of recombinant type-A and type-B diterpene synthases prepared in *Escherichia coli* and the heterologous expression of two genes encoding type-A and type-B diterpene synthases in an engineered *Streptomyces* host. The generation of labda-8(17),12(E),14-triene (**1**) by CldD and CldD was reconfirmed by enzymatic synthesis. Furthermore, labda-8(17),13(16),14-triene (**2**) was generated by SCLAV_p0491 and CldB, and labda-7,12(E),14-triene (**3**) by CldD and SCLAV_p0490. SCLAV_p0491 and SCLAV_p0490 catalyzed the generation of the novel diterpene hydrocarbon, labda-7,13(16),14-triene (**4**).

The Journal of Antibiotics (2016) 69, 515–523; doi:10.1038/ja.2015.147; published online 27 January 2016

INTRODUCTION

Terpenoid metabolites are an important class of natural products that act as antibiotics, hormones, flavor and odor constituents, and pigments. Some of these metabolites also possess other physiologically or commercially important properties as vitamins and antimalarial and antitumor agents. Most of these metabolites have been isolated from terrestrial and marine higher plants, liverworts and fungi. Actinomycetales microorganisms are prolific producers of a large number of natural products; however, the discovery of terpenoid metabolites from Actinomycetales microorganisms is rare. Previous studies reported the production of the volatile monoterpene and degraded sesquiterpene alcohols, 2-methylisborneol and geosmin, respectively, from soil bacteria, actinomycetes,^{1,2} cyanobacteria^{3,4} and myxobacteria.^{5,6} Although a few terpenoid metabolites have been isolated from Actinomycetales microorganisms, many genes encoding presumptive terpene synthases, which were in a cryptic state in the microorganisms, have recently been identified in public bacterial genome databases.⁷ Furthermore, many novel terpenes have been isolated by the expression of these silent genes in the heterologously engineered *Streptomyces* host, *Streptomyces avermitilis* SUKA22.⁸

Several cyclic terpene hydrocarbons and alcohols are formed by variations in the common terpene cyclization mechanism that is initiated by enzyme-catalyzed ionization to form allylic cations from the universal acyclic precursors geranyl diphosphate (for

monoterpenes), farnesyl diphosphate (for sesquiterpenes), and geranylgeranyl diphosphate (GGPP; for diterpenes). The intramolecular electrophilic attack of the intermediate allylic cations on the central or distal double bonds of the substrate, followed by well-precedented cationic transformations, including hydride shifts and carbon–carbon backbone rearrangements, and the ultimate quenching of the positive charge either by deprotonation or by the capture of a nucleophilic water molecule, may account for the formation of several hundred cyclic terpene compounds. In the formation of mono- and sesquiterpenes, these cyclization reactions are accomplished by a single polypeptide synthase, whereas the formation of diterpenes is achieved by the above single polypeptide synthase (type-A or type-B terpene synthase) or a combination of two fundamental types of terpene synthases (type-A, type-B, type-A–type-B, and type-B–type-A).⁹ The type-A diterpene synthase reaction is initiated by ionization of the diphosphate of GGPP to give an allylic cation, followed by deprotonation to generate an olefinic diterpene hydrocarbon or the capture of a water molecule to produce a diterpene alcohol. The cyclization reaction catalyzed by type-B diterpene synthases is initiated by protonation at the C-14,15-double bond of GGPP distal to the allylic diphosphate.

In bacterial terpene synthases, more than 260 presumptive synthases have been classified into type-A synthase; however, the combination of type-A and type-B diterpene synthases is extremely rare. Only three

pairs of type-A and type-B diterpene synthases, which are involved in the formation of the diterpenoid metabolites viguiepinol,¹⁰ oxaloterpin C¹¹ and terpentecin,¹² had been characterized from a bacterial origin. We recently characterized two pairs of type-A and type-B diterpene synthases from *S. cylabdanicus* K04-0144 and *S. anulatus* GM95, respectively, and they catalyzed the formation of the labdane-type bicyclic diterpene hydrocarbon, labda-8(17),12(E),14-triene.¹³ We herein determined whether other Actinomycetales microorganisms harbor the labdane-type diterpene synthases from public databases and characterized them.

MATERIALS AND METHODS

Preparation of recombinant type-A and type-B diterpene synthases in *Escherichia coli*

Four genes, *cldB*, *cllD*, *sclav_p0491* and *sclav_p0490*, were amplified by PCR with cosmid clones containing the *clD* cluster and *sclav_p0491/sclav_p0490* as template DNA, respectively, using the following primer pairs: forward for *cldB*: 5'-GCTGGCACATATGAGCTCCATATCCCACCAGCGGC-3' (underlined and bold characters indicate the *NdeI* site and start codon of *cldB*, respectively) and reverse for *cldB*: 5'-CTGCAGAAGCTTTCAGGGCCGACGGCCTGCA GCA-3' (underlined and bold characters indicate *HindIII* and the stop codon of *cldB*, respectively), forward for *cllD*: 5'-GGGAGAGCATATGACCCGACCCGG CGATGCGGTGAC-3' (underlined and bold characters indicate the *NdeI* site and start codon of *cllD*, respectively) and reverse for *cllD*: 5'-GTCTCC AAGCTTGGTTCAACGACTGGTTGGGGCCGCG-3' (underlined and bold characters indicate *HindIII* and the stop codon of *cllD*, respectively), forward for *sclav_p0490*: 5'-CCTGTCCCATATGCCCGTCGACGTCGGCACCCCTTC-3' (underlined and bold characters indicate the *NdeI* site and start codon of *sclav_p0490*, respectively) and reverse for *sclav_p0490*: 5'-GACCAC AAGCTTTCAGTCGCGGGCCCGTCCGAG-3' (underlined and bold characters indicate *HindIII* and the stop codon of *sclav_p0490*, respectively), and forward for *sclav_p0491*: 5'-GGGAGTGCATATGAGGTCTGGGGACCGTGGC GAG-3' (underlined and bold characters indicate the *NdeI* site and start codon of *sclav_p0491*, respectively) and reverse for *sclav_p0491*: 5'-GAGCAG AAGCTTCTACTCCTCCCGCCAGCGGTGCGC-3' (underlined and bold characters indicate *HindIII* and the stop codon of *sclav_p0491*, respectively). Initial denaturation at 96 °C for 180 s was followed by 25 cycles of amplification (95 °C for 30 s, 60 °C for 30 s and 72 °C for 60 s (*cllD* and *sclav_p0491*) or 100 s (*cldB* and *sclav_p0490*)) and then a final incubation at 72 °C for 5 min. After amplification, the reaction mixture was treated with *DpnI* to remove template DNA. Amplified DNA was digested with *NdeI* and *HindIII*, and the *NdeI/HindIII* amplicon was then ligated with the large *NdeI/HindIII*-pET28a (+). After confirming the sequence of each amplicon, each recombinant plasmid was expressed as the N-terminal His₆-tagged fusion protein in *E. coli* JM109 (DE3). Transformants were grown at 37 °C in Luria broth with 50 µg of kanamycin per ml. Expression of the recombinant protein was induced by adding 0.1 mM isopropyl-β-D-thiogalactopyranoside when optical density at 600 nm reached approximately 0.8. Cultivation was continued for an additional 12 h at 18 °C. After the induced cells were harvested by centrifugation at 3000 r.p.m. at 4 °C for 10 min, the cells were disrupted by sonication 10 times for 15 s (20-s cooling intervals). Since transformants carrying pET28a(+):*cldB* and pET28a(+):*sclav_p0490* produced insoluble proteins, these proteins were sedimented by centrifugation at 5000 r.p.m. at 4 °C for 10 min. Insoluble proteins were dissolved in buffer containing 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 5 mM dithiothreitol (DTT) and 8 M urea at 37 °C for 20 h. After insoluble materials had been removed by centrifugation at 10 000 r.p.m. at 4 °C for 10 min, the supernatant was dialyzed against buffer containing 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 mM DTT and 20 v/v% glycerol. Dialyzed solution was centrifuged at 10 000 r.p.m. at 4 °C for 20 min in order to remove insoluble materials. The purification of His₆-tagged recombinant proteins was performed in accordance with the manufacturer's (Qiagen K. K., Tokyo, Japan) protocol using Ni-NTA agarose. Purified proteins were analyzed by sodium dodecyl sulfate 12.5 w/v% polyacrylamide gel electrophoresis.

Enzymatic synthesis of diterpenes by type-A and type-B diterpene synthases

The reaction mixture (0.5 ml) contained 50 mM Tris-HCl (pH 7.5), 10 mM MgCl₂, 1 mM DTT, 0.1 w/v% Tween 80, 20 w/v% glycerol, 30 µM IPP, 10 µM DMAPP, 100 µg of purified recombinant GGPP synthase (N-terminus His₆-tagged CldA prepared from *E. coli* JM109 (DE3) harboring pET28a(+):*cllD*) and 150 µg each of the purified type-A and type-B diterpene synthases per ml. The reaction was carried out at 30 °C for 16 h, was stopped by the addition of 50 µl of 0.5 M EDTA (pH 7.5), and the product was extracted two times with 0.5 ml of *n*-pentane. The upper organic layers were combined and concentrated under reduced pressure to remove *n*-pentane. The residue was dissolved in 100 µl of methanol and a 2-µl portion of the extract was subjected to GC-MS.

Expression of genes encoding type-A and type-B diterpene synthases in the engineered *Streptomyces* host

Each gene encoding *cldB*, *cllD*, *sclav_p0490* or *sclav_p0491* was amplified by PCR with the cosmid clone containing the *clD* cluster or *sclav_p0490/sclav_p0491* using the following primer pairs: forward for *cldB*: 5'-CGC TCGGGTCTAGACGGAAGGCAGTCATGAGCTCCATATCCC-3' (underlined and bold characters indicate the *XbaI* site and start codon of *cldB*, respectively) and reverse for *cldB*: 5'-TTAATTAAGCAACCTTTGCTCAGGGCCGACGGG CCTGCAGCAG-3' (underlined and bold characters indicate the *BstAPI* site and stop codon of *cldB*, respectively), forward for *cllD*: 5'-TGCCTGCCGCA AAGGTTGAGGGCCGATGACCCGACCC-3' (underlined and bold characters indicate the *BstAPI* site and start codon of *cllD*, respectively) and reverse for *cllD*: 5'-CCCGAGAAGCTTTCACGACTGGTTGGGGCCGCGCCGCG-3' (underlined and bold characters indicate the *HindIII* site and stop codon of *cllD*, respectively), forward for *sclav_p0490*: 5'-CCTCTAGAGAAAGGCAG TCATGCCCGTCGACGTCGGCAC-3' (underlined and bold characters indicate the *XbaI* site and start codon of *sclav_p0490*, respectively) and reverse for *sclav_p0490*: 5'-TTAATTAAGCAACCTTTGCTCAGTCGGCGGGCCCGGT CCGAGC-3' (underlined and bold characters indicate the *BstAPI* site and stop codon of *sclav_p0490*, respectively), or forward for *sclav_p0491*: 5'-TGCTGCC GCAAAGGTTGCGAGGAGTGTCCGTGAGGTCTGGGG-3' (underlined and bold characters indicate the *BstAPI* site and start codon of *sclav_p0491*, respectively) and reverse for *sclav_p0491*: 5'-CCCTCGAGAAGCTTGTACTC CTCCCGCCAGCGGTGCGC-3' (underlined and bold characters indicate the *HindIII* site and stop codon of *sclav_p0491*, respectively). Initial denaturation at 96 °C for 180 s was followed by 25 cycles of amplification (96 °C for 30 s, 60 °C for 30 s and 72 °C for 60 s (*cllD* and *sclav_p0491*) or 100 s (*cldB* and *sclav_p0490*)) and then a final incubation at 72 °C for 5 min. After amplification, the reaction mixture was treated with *DpnI* to remove template DNA. Each amplified DNA was digested with *XbaI* and *BstAPI* (*cldB* and *sclav_p0490*) or *BstAPI* and *HindIII* (*cllD* and *sclav_p0491*). The large *XbaI/HindIII*-cut pKU1021ggs⁷ was ligated with the *XbaI/BstAPI cldB* amplicon and *BstAPI/HindIII cllD* amplicon, *XbaI/BstAPI cllD* amplicon and *BstAPI/HindIII sclav_p0491* amplicon, *XbaI/BstAPI sclav_p0490* amplicon and *BstAPI/HindIII cllD* amplicon, or *XbaI/BstAPI sclav_p0490* and *BstAPI/HindIII sclav_p0491* amplicon to generate pLDT101 (pKU1021ggs::*cldB::cllD*), pLDT102 (pKU1021ggs::*cllD::sclav_p0491*), pLDT103 (pKU1021ggs::*sclav_p0490::cllD*), or pLDT104 (pKU1021ggs::*sclav_p0490::sclav_p0491*), respectively. Each amplified region was confirmed by sequencing. After these integrating plasmids were transformed into *E. coli* GM2929 *hds::Tn10* to prepare unmethylated DNA preparations, the DNAs prepared were introduced into an engineered *S. avermitilis* SUKA22 by PEG-assisted protoplast transformation.¹⁴ Transformants were obtained by the selection of resistance to neomycin and spores of the transformants were grown on YMS¹⁵ agar medium containing 1 µg of neomycin per ml to prepare spores. The spores of *S. avermitilis* SUKA22 transformants carrying pLDT101, pLDT102, pLDT103 or pLDT104 were grown in vegetative medium¹⁶ at 30 °C for 2 days on a reciprocal shaker (200 strokes per min). A 0.15-ml portion of the vegetative culture was inoculated into 15 ml of the production medium¹⁶ in a 125-ml flask. The whole culture was centrifuged at 3000 r.p.m. for 10 min. Sedimented mycelia were extracted with 5 ml of methanol and removed by centrifugation. A 2-µl portion of the methanol extract was directly subjected to GC-MS.

Isolation of diterpenes by heterologous expression in the engineered *Streptomyces* host

Two liters of the culture (20 × 100 ml medium in a 500-ml Erlenmeyer flask) growing *S. avermitilis* SUKA22 carrying pLDT101, pLDT102, pLDT103 or pLDT104 in production medium¹⁶ at 28 °C for 5 days was centrifuged at 6000 r.p.m. for 20 min to separate mycelia and the supernatant. Sedimented mycelia were suspended in 500 ml of water and the suspension was re-centrifuged at 6000 r.p.m. for 20 min. Washed mycelia were extracted twice with 500 ml of methanol and mycelia were then removed by filtration. The methanol extract was concentrated under reduced pressure to make approximately 100 ml. The concentrate was extracted two times with 50 ml of *n*-hexane. The upper organic phases were combined and extracted with 100 ml of 0.1 N NaOH in order to remove free fatty acids. The organic phase was separated, dried over anhydrous Na₂SO₄ and concentrated to yield a brownish oily material (0.5–1.2 g). The oily material was dissolved in 20 ml of *n*-pentane and insoluble materials were removed by filtration. The filtrate was evaporated to reduce the volume and the concentrate obtained was subjected to silica gel column chromatography (20 × 150 mm). The column was developed with *n*-pentane and the eluates were fractionated. The metabolites eluted were monitored by silica gel TLC developed with *n*-hexane. Diterpene compounds on silica gel TLC were detected by heating at 110 °C after spraying 5 w/v% phosphomolybdic acid in ethanol solution. The fractions containing the objective metabolites were combined and concentrated by evaporation. Diterpene hydrocarbon **1** (20.2 mg, $\lambda_{\max}^{\text{MeOH}}$ nm (log ϵ) 232 (4.32), $[\alpha]_{\text{D}}^{25}$ +29.5 (*c* 0.5, CHCl₃)) was obtained as transparent oil from *S. avermitilis* SUKA22 carrying pLDT101. The HR-MS (ESI mode) of **1** showed a molecular ion peak at *m/z* 272.2506 [M]⁺, corresponding to C₂₀H₃₂ (calcd. for 272.2504). Diterpene hydrocarbons **2** (13.4 mg, $\lambda_{\max}^{\text{MeOH}}$ nm (log ϵ) 232 (4.36), $[\alpha]_{\text{D}}^{25}$ +28.0 (*c* 0.5 CHCl₃)) and **3** (7.2 mg, $\lambda_{\max}^{\text{MeOH}}$ nm (log ϵ) 232 (4.35), $[\alpha]_{\text{D}}^{25}$ +3.6 (*c* 0.5 CHCl₃)) were obtained as transparent oil from *S. avermitilis* SUKA22 carrying pLDT102 and pLDT103, respectively. The HR-MS (ESI) of **2** and **3** showed molecular ion peaks at *m/z* 272.2507 [M]⁺ and 272.2506 [M]⁺, respectively, corresponding to C₂₀H₃₂ (calcd. for 272.2504). Diterpene hydrocarbon **4** (8.1 mg, $\lambda_{\max}^{\text{MeOH}}$ nm (log ϵ) 232 (4.32), $[\alpha]_{\text{D}}^{25}$ +3.5 (*c* 0.5 CHCl₃)) was obtained as transparent oil from *S. avermitilis* SUKA22 carrying pLDT104. The HR-MS (ESI) of **4** showed molecular ion peaks at *m/z* 272.2505 [M]⁺, corresponding to C₂₀H₃₂ (calcd. for 272.2504).

Physicochemical analysis

The crude and purified preparations of metabolites were analyzed by GC-MS (Shimadzu GC-17 A; 70 eV, electron ionization, positive ion mode; 30 m × 0.25 mm InterCap 5 capillary column (diphenyl (5%)-dimethyl (95%) polysiloxane; GL Sciences, Tokyo, Japan) using a temperature program of ~50–280 °C and temperature gradient of 20 °C/min). The compounds produced were evaluated by comparisons with the mass spectra of the corresponding reference compounds in the GC-MS database, the NIST/EPA/NIH MS Library (2014 version). 1D and 2D NMR (¹H: 500 MHz; ¹³C: 125 MHz, in CDCl₃) spectra were obtained on a JEOL JNM-ECP 500 FT NMR system. Abbreviations in ¹H NMR signals were as follows: singlet (s), multiplet (m), doublet (d), triplet (t), double doublet (dd), double triplet (dt) and broad singlet (brs) signals. HR-MS by the ESI mode were obtained on a JEOL JMS-700 Mstation. Optical rotations were recorded on a Horiba SEPA-300 polarimeter.

RESULTS

Searching pairs of type-A and type-B diterpene synthases from public databases

In our previous studies, more than 260 presumptive terpene synthases were discovered in public databases and in-house databases using a search method based on the use of hidden Markov models and the protein families database (Pfam) search.⁷ Of these, the type-A diterpene synthases with *E*-value ranges from 6 × 10⁻⁹ to 2.9 × 10⁻⁵⁰, terpenete-3,13,15-triene synthase (311-aa; BAB39207) and pimara-9 (11),15-diene synthase (295-aa; BAD86798), located downstream of the type-B diterpene synthases, terpentenedienyl diphosphate synthase (499-aa; BAB39206) of terpentenic-producing *Kitasatospora griseola*

MF730-N6 and *ent*-copalyl diphosphate synthase (511-aa; BAD86797) of viguiepinol-producing *Streptomyces* sp. KO-3988, respectively, had already been characterized. We recently characterized new genes encoding the labda-8(17),12(*E*),14-triene synthases, CldD (329-aa; BAR97453) of *S. cylabdanicus* K04-0144 and RmnD (322-aa; BAR97463) of *S. anulatus* GM95, which were also located downstream of genes encoding the (+)-copalyl diphosphate synthases, CldB (532-aa; BAR97451) and RmnB (549-aa; BAR97461), respectively. Since the genes encoding each pair of type-A and type-B diterpene synthases formed an operon, a new pair of type-A and type-B diterpene synthases were searched for in the public databases using a combination of the recent bacterial type-A terpene synthase hidden Markov models and bacterial type-B diterpene synthase hidden Markov models, as reported previously.¹⁷ The pairs of diterpene synthases selected by their genes were located in the same operon or gene cluster. In consideration of the pair of type-A and type-B diterpene synthases in the diterpene synthase family from bacterial proteins in the public databases, Sare_1288 (531-aa)/Sare_1287 (295-aa) of *Salinispora arenicola* CNS-205, SCLAV_5672 (529-aa)/SCLAV_5671 (311-aa) and SCLAV_p0490 (545-aa)/SCLAV_p0491 (343-aa) of *Streptomyces clavuligerus* ATCC 27064, AHK61133 (311-aa)/AHK61132 (510-aa) of *S. griseus* CB00830 (viguiepinol, oxaloterpins C and E producer), and Sros_3192 (560-aa)/Sros_3191 (295-aa) of *Streptosporangium roseum* DSM 43021 were selected as the combinations of pairs of type-A and type-B diterpene synthases.

In order to evaluate these diterpene synthase pairs, the deduced amino acid sequences of type-A and type-B diterpene synthases were aligned and subjected to a phylogenetic analysis. Type-A diterpene synthases were classified into at least two groups (group 1: Sros_3191, Sare_1287, SCLAV_5671, AHK61133 and BAD86798; group 2: SCLAV_p0491, PmnD, CldD and BAB39207) in Figure 1a. Since two pimara-9(11),15-diene synthases, AHK61133 and BAD86798, belong to group 1, the other synthases, Sros_3191, Sare_1287 and SCLAV_5671, may exhibit similar catalytic activities. SCLAV_5671 was located in the pimara-9(11),15-diene synthase clade. The presumptive type-A diterpene synthase, SCLAV_p0491, was very similar to the labda-8(17),12(*E*),14-triene synthases, RmnD and CldD. All predicted type-A terpene synthases possessed a pair of characteristically conserved metal-binding domains consisting of the acidic amino acid-rich motif DW[LQ]VD (Sros_3191 and Sare_1287), DD[HRY] [TVW]E (SCLAV_5671, AHK61132, BAD86798 and BAB39207) or DD[AMV]H[GV]E (SCLAV_p0491, CldD and RmnD) located within 65–114 amino acids of the N-terminus, while the downstream triad Nxxx[ST]xxx[DE] found other 140 amino acids downstream, closer to the C-terminus (Figure 2a). The phylogenetic analysis of type-B synthases shown in Figure 1b suggested that the synthases were classified into three groups (group 3: Sros_3192, Sare_1288, SCLAV_5672, AHK61132 and BAD86797; group 4: SCLAV_p0490, RmnB and CldB; group 5: BAB39206). Five synthases, including the *ent*-copalyl diphosphate synthases, AHK61132 and BAD86797, belonged to the group 3 clade. Since SCLAV_5672 was similar to AHK61132 and BAD86797, the presumptive synthase may exhibit *ent*-copalyl diphosphate synthase activity. On the other hand, the presumptive synthase, SCLAV_p0490, was classified into two (+)-copalyl diphosphate synthases, RmnB and CldB, suggesting that SCLAV_p0490 exhibited similar catalytic activity to these synthases. Sequence alignment with the predicted type-B diterpene synthases revealed that the motifs Dx[D]T[AS] and Qxx[DNS]G[SG]W, which were proposed to stabilize an intermediate cation during the cyclization process and mediate substrate binding by the chelation of divalent metal ions,^{18,19} respectively, were highly conserved in all

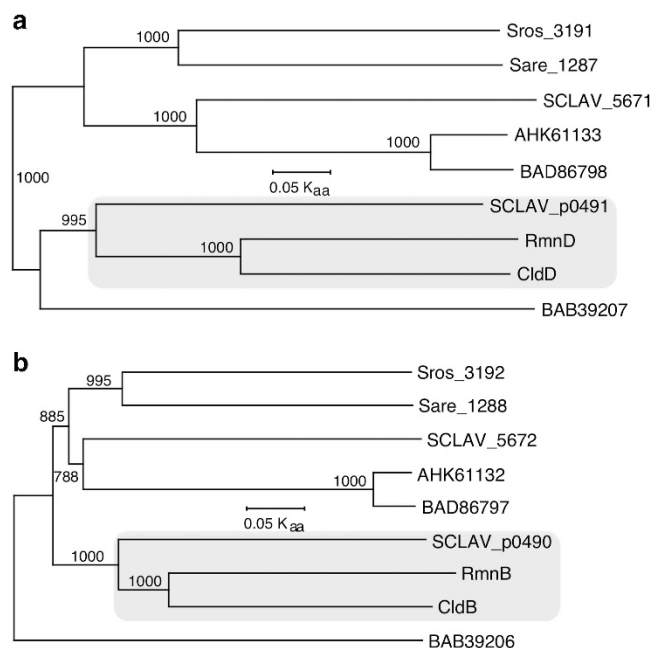


Figure 1 Phylogenetic analysis of (a) type-A diterpene synthases and (b) type-B diterpene synthases from Actinomycetales microorganisms. A phylogenetic analysis of aligned sequences was conducted using the bootstrap method (bootstrap number: 1000; seed number: 111) of CLUSTALW version 2.10 ([ftp://ftp.ebi.ac.uk/pub/software/clusterw2/2.1/](http://ftp.ebi.ac.uk/pub/software/clusterw2/2.1/)). A bootstrap tree was drawn by njplot (<http://doua.prabi.fr/software/njplot>). Abbreviations: AHK61132 (510-aa, accession # AHK61132), AHK61133 (311-aa, accession # AHK61132), BAB39206 (499-aa, accession # BAB39206), BAB39207 (311-aa, accession # BAB39207), BAD86797 (511-aa, accession # BAD86797), BAD86798 (295-aa, accession # BAD86798), CldB (532-aa, accession # BAR97451), CldD (329-aa, accession # BAR97453), RmnB (549-aa, accession # BAR97461), RmnD (322-aa, accession # BAR97463), Sare_1287 (295-aa, accession # ABV97190), Sare_1288 (531-aa, accession # ABV97191), SCLAV_5671 (311-aa, accession # EFG10738), SCLAV_5672 (529-aa, accession # EFG10739), SCLAV_p0490 (545-aa, accession # EFG03980), SCLAV_p0491 (343-aa, accession # EFG03981), Sros_3191 (295-aa, accession # ACZ86136) and Sros_2192 (560-aa, accession # ACZ86137).

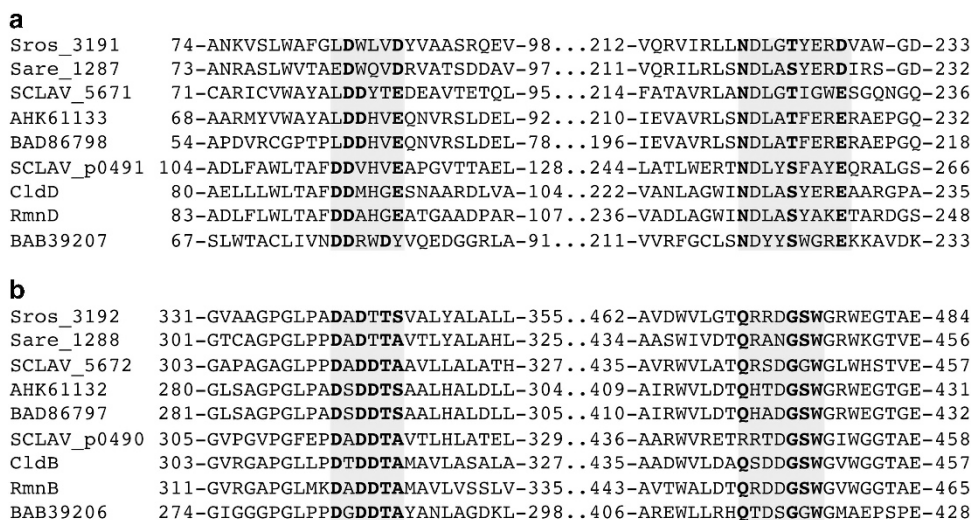


Figure 2 Alignment of amino acid sequences of (a) type-A diterpene synthases and (b) type-B diterpene synthases. Shadow boxes in the upper alignment indicate metal-binding domains consisting of an acidic amino acid-rich motif (left box) and downstream triad (right box). Shadow boxes in the lower alignment indicate an acidic amino acid-rich motif (left box) and downstream highly conserved motif (right box).

predicted type-B diterpene synthases, except for SCLAV_p0490, which possessed two motifs ³¹⁵DADDTA and ⁴⁴⁴RRTDGSW (Figure 2b).

A phylogenetic analysis of these aligned sequences of type-A and type-B diterpene synthases revealed that the pairs of Sros_3191/Sros_3192, Sare_1287/Sare_1288, SCLAV_5671/SCLAV_5672, AHK61133/AHK61132 and BAD86798/BAD86797 possessed the catalytic function for the formation of the tricyclic diterpene skeleton including pimara-9(11),15-diene, whereas the pairs of SCLAV_p0491/SCLAV_p0490, RmnD/RmnA, CldD/CldB and BAB39207/BAB39206 are involved in the formation of the bicyclic diterpene skeleton including labda-8(17),12(E),14-triene and terpen-3,13,15-triene. The presumptive diterpene synthases, SCLAV_p0491 and SCLAV_p0490, were classified into the clade of labda-8(17),12(E),14-triene synthases, RmnD and CldD, and the clade of (+)-copalyl diphosphate synthases, RmnB and CldB, respectively. Therefore, the pair of SCLAV_p0491 and SCLAV_p0490 may be involved in the formation of labdane-type diterpenes.

Detection of diterpenes by enzymatic synthesis

At least 17 genes encode type-A mono-, sesqui- and diterpene synthases, including *sclav_p0491*, in *S. clavuligerus* ATCC 27064; however, when grown under various culture conditions, the microorganism did not produce any detectable terpenes.⁷ Therefore, these two genes were also defined as a cryptic state in this microorganism and the function of either gene product, SCLAV_p0490 or SCLAV_p0491, was not expected from cultural metabolites. In order to verify the biochemical functions of SCLAV_p0490 and SCLAV_p0491 and similar diterpene synthases, CldB and CldD, their recombinant enzymes were prepared in *E. coli*. Two recombinant type-A synthases, CldD and SCLAV_p0491, were obtained as soluble forms, whereas recombinant CldB and SCLAV_p0490 both formed insoluble proteins under the various induction conditions (concentration of isopropyl- β -D-thiogalactopyranoside, growth phase for induction and the incubation temperature after induction). Furthermore, although a recombinant C-terminal His₆-tagged protein was prepared, the resultant proteins were still insoluble. These insoluble proteins were dissolved in 8 M urea containing 5 mM DTT at 37 °C for 20 h, and a very small amount of soluble proteins was

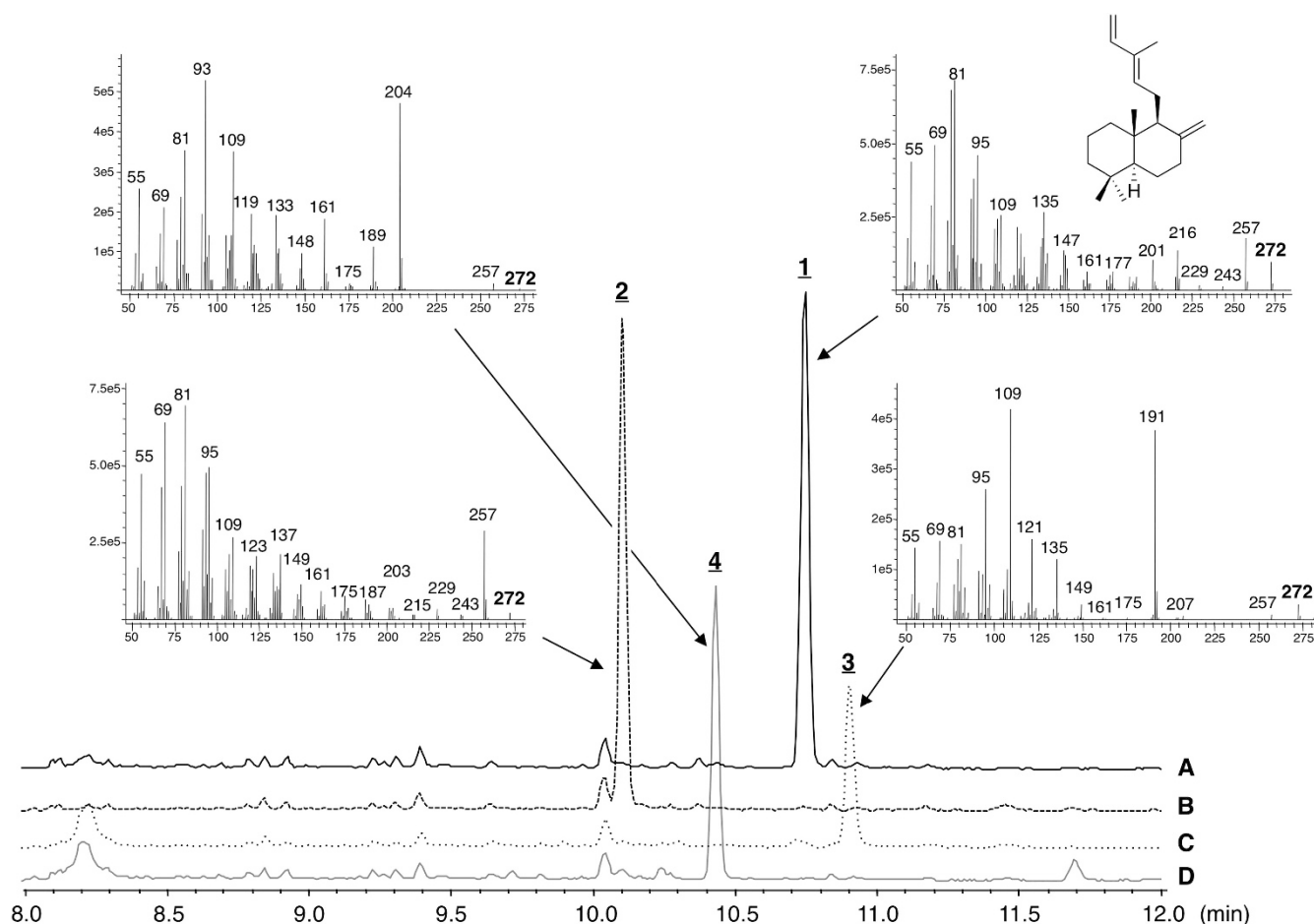


Figure 3 GC-MS analysis of *n*-pentane extracts of recombinant enzyme reactions. After the enzyme reactions containing type-B and type-A diterpene synthases (A: CldB and CldD; B: CldB and SCLAV_p0491; C: SCLAV_p0490 and CldD; and D: SCLAV_p0490 and SCLAV_p0491) were performed at 30 °C for 20 h, they were stopped by the addition of EDTA, and the products were extracted with *n*-pentane. A 5- μ l portion of the *n*-pentane extract was subjected to GC-MS. The reaction products of CldB and CldD were identical to labda-8(17),12(*E*),14-triene.

recovered by dialysis against buffer containing 1 mM DTT followed by centrifugation to remove insoluble proteins. The reaction products of four combinations of pairs of type-A and type-B diterpene synthases (CldD/CldB, CldD/SCLAV_p0490, SCLAV_p0491/CldB and SCLAV_p0491/SCLAV_p0490, respectively) were examined. As shown in Figure 3, the elution time (10.74 min) and MS fragmentation pattern of the enzymatic product (**1**) generated by the catalytic reaction of CldD and CldB in the presence of magnesium ions, IPP, DMAPP and GGPP synthase (CldA) were identical to those of the authentic sample of labda-8(17),12(*E*),14-triene (biformene), which was prepared from the mycelial products of *S. avermitilis* SUKA22 carrying pCLD104¹³ (*clmA-cldB-cldD*). The enzymatic reaction product by SCLAV_p0491/SCLAV_p0490 was eluted at 10.42 min, with a molecular ion peak at *m/z* 272 [M^+], which was presumed to be diterpene hydrocarbon, and MS fragmentation patterns different from those of **1**. Other two-enzymatic reaction products were eluted at 10.12 min (**2**) from CldD/SCLAV_p0490 and at 10.91 min (**3**) from SCLAV_p0491/CldB, and the MS fragmentation patterns of these two products with a molecular ion peak at *m/z* 272 [M^+] were different from those of **1** and **4**. These results indicated that all enzymatic products were diterpene hydrocarbons and these three diterpene hydrocarbons (**2**, **3** and **4**) were not identical to **1**. Since the soluble forms of the type-B diterpene synthases, CldB and SCLAV_p0490, were inefficiently obtained from the T7-based *E. coli* expression

system, it was difficult to prepare a large amount of recombinant enzyme for structural analysis of these diterpene hydrocarbons (**2**, **3** and **4**). The preparation of each product was accomplished by the heterologous expression of two genes encoding type-A and type-B diterpene synthases using the *Streptomyces* expression system.

Expression of the pair of type-A and type-B diterpene synthases in the engineered *Streptomyces* host

We previously developed an efficient expression system for the gene encoding terpene synthase in an engineered *Streptomyces* host,^{7,14,17,20} and discovered many novel terpene compounds from bacterial origins.⁸ Furthermore, we examined the overexpression of the *cldB* and *cldB* pair by co-expressing with a gene encoding GGPP synthase, which was controlled by the constitutively expressed promoter in *S. avermitilis* SUKA22, and a large amount of **1** was obtained from the transformants.¹³ Each gene with its ribosome-binding sequence was prepared by PCR amplification. All gene pairs, *cldB/cldB*, *cldB/sclav_p0490*, *sclav_p0491/cldB* and *sclav_p0491/sclav_p0490*, were introduced downstream of the *cldB*-encoding GGPP synthase in pKU1012,²¹ the operon of which was controlled by the constitutively expressed promoter of ribosomal protein S10 (P_{rps}),²¹ in order to obtain pLDT101, pLDT102, pLDT103 and pLDT104, respectively. As shown in Figure 4, all transformants produced a single component of diterpene hydrocarbon in the mycelium and each peak was identical to

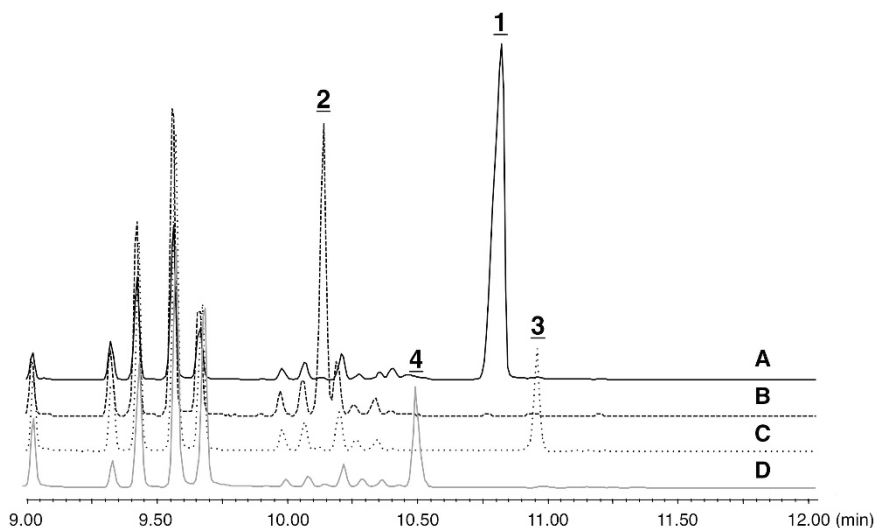


Figure 4 GC-MS analysis of *n*-hexane extracts of *S. avermitilis* SUKA22 transformants carrying (A) pLDT101 (*cldA::cldB::cldD*), (B) pLDT102 (*cldA::cldB::sclav_p0491*), (C) pLDT103 (*cldA::sclav_p0490::cldD*) and (D) pLDT104 (*cldA::sclav_p490::sclav_p0491*). Peaks eluted from 9.0 to 9.7 min were identical to middle chain fatty acids.

Table 1 ^1H and ^{13}C NMR data of **4**

	δ_{C}	δ_{H} (multi, J in Hz)
1	39.2	1.81 (1H, m), 0.96 (1H, m)
2	18.9	1.58 (1H, m), 1.42 (1H, m)
3	42.4	1.38 (1H, m), 1.15 (1H, m)
4	33.3	—
5	50.3	1.19 (1H, m)
6	23.9	1.98 (1H, m), 1.72 (1H, m)
7	122.4	5.36 (1H, brs)
8	135.4	—
9	54.9	1.68 (1H, m)
10	36.9	—
11	29.8	1.98 (1H, m), 1.70 (1H, m)
12	54.9	1.68 (1H, m), 1.67 (1H, m)
13	147.1	—
14	139.0	6.36 (1H, dd, $J=10.5, 6.9$)
15	113.3	5.23 (1H, d, $J=10.6$), 5.05 (1H, d, $J=7.4$)
16	115.9	4.99 (2H, d, $J=10.6$)
17	22.3	1.72 (3H, s)
18	33.1	0.85 (3H, s)
19	21.9	0.86 (3H, s)
20	13.6	0.73 (3H, s)

that of the enzymatic reaction described in Figure 3. The productivity of **1** by *S. avermitilis* SUKA22 transformants carrying pLDT101 was similar to that of transformants carrying pCLD104, as previously reported.¹³ The *S. avermitilis* SUKA22 transformant carrying pLDT101, pLDT102, pLDT103 or pLDT104 was grown in a 2-l culture and the structure of each diterpene hydrocarbon was determined by detailed ^1H and ^{13}C NMR analyses including DEPT, ^1H - ^1H COSY, ^1H - ^1H NOESY, ^1H - ^{13}C HMQC and ^1H - ^{13}C HMBC, as measured in CDCl_3 .

The diterpene hydrocarbon **1** was obtained from *S. avermitilis* SUKA22 carrying pLDT101 and the structure of **1** was identical to labda-8(17),12(*E*),14-triene (biformene) based on comparisons with the ^1H and ^{13}C NMR spectra of previous studies¹³ and published data.²² Two diterpene hydrocarbons generated by the expression of the

heterogeneous combination of genes encoding type-A and type-B diterpene synthases, **2** and **3**, were obtained from *S. avermitilis* SUKA22 carrying pLDT102 and pLDT103, respectively. In consideration of NMR spectral data (Supplementary Figures S1–S5), the structure of **2** was identical to the known diterpene labda-8(17),13(16),14-triene (sclarene; from the foliage of *Podocarpus cunninghamii*).²³ Although the molecular formula of **3** was identical to those of **1** and **2**, the MS fragmentation pattern of **3** was different (Figure 3). The structure of **3** was identical to the known diterpene labda-7,12(*E*),14-triene by comparisons with published data on plant metabolites (from the stem bark of *Croton oblongifolius*)²⁴ and spectra of **3** (Supplementary Figures S6–S10).

A diterpene hydrocarbon that possesses the characteristics of **4** has not yet been described in the literature, indicating that **4** is a novel diterpene. The MS fragmentation pattern of **4** was similar to that of **3** (Figure 3) and the molecular formula was estimated to be $\text{C}_{20}\text{H}_{32}$ for a diterpene hydrocarbon with five degrees of unsaturation. The ^1H NMR spectrum of **4** showed two pairs of two olefinic protons, H-15 (δ_{H} 5.23, d, $J=10.6$ Hz, 1H and 5.05, d, $J=7.4$ Hz, 1H) and H-16 (δ_{H} 4.99, d, $J=10.6$ Hz, 2H), assignable to exomethylenes. An additional olefinic proton (δ_{H} 5.36, brs, 1H) was observed. An analysis of the ^{13}C NMR and DEPT spectra of **4** showed 20 resolved signals and confirmed the presence of four methyls, six sp^3 methylenes, two sp^2 methylenes, two sp^3 methines, two sp^2 methines, and two sp^2 and two sp^3 quaternary carbons (Table 1). A comparison between the **1** and **4** data of ^1H and ^{13}C NMR spectra (Supplementary Figure S11) indicated that the following four different points were observed: the first was that the C-16 of **4** was assigned as an exomethylene (sp^2 methylene), whereas that of **1** was a methyl carbon; the second was that the C-12 of **4** was assigned as a sp^3 methylene carbon, whereas that of **1** was an sp^2 methine carbon; the third was that the C-17 of **4** was assigned as a methyl carbon, whereas that of **1** was an exomethylene carbon; and the fourth was that C-7 was assigned as an sp^2 methine, whereas that of **1** was an sp^2 methylene carbon. Three double bonds accounted for the three degrees of unsaturation, indicating that **4** possessed a bicyclic skeleton. The connectivity of protons and carbons was assigned by an analysis of the ^1H - ^{13}C HMQC spectrum (Supplementary Figure S13), as shown in Figure 5.

An analysis of ^1H - ^1H COSY (Supplementary Figure S12) revealed four partial structures: C-1–C-3, C-5–C-7, C-9–C-11 and C-14–C-15 (Figure 5), and the ^1H - ^{13}C long-range coupling of 2 *J* and 3 *J* observed in the ^1H - ^{13}C HMBC experiments (Supplementary Figure S14) provided the following information: the cross peaks from H-14 (δ_{H} 6.36) to C-15 (δ_{C} 113.3) from H-15 (δ_{H} 5.23) to C-13 (δ_{C} 147.1) and C-14 (δ_{C} 139.0), from H-16 (δ_{H} 4.99) to C-12 (δ_{C} 54.9), C-13 and C-14, from H-17 (δ_{H} 1.72) to C-7 (δ_{C} 122.4), C-8 (δ_{C} 135.4) and C-9 (δ_{C} 54.9), H-18 (δ_{H} 0.85) to C-3 (δ_{C} 42.4), C-4 (δ_{C} 33.3), C-5 (δ_{C} 50.3) and C-19 (δ_{C} 21.9), from H-19 (δ_{H} 0.86) to C-3, C-4, C-5 and C-18 (δ_{C} 33.1), and from H-20 (δ_{H} 0.73) to C-1 (δ_{C} 39.2), C-5, C-9 and C-10 (δ_{C} 36.9) supported the structure shown in Figure 5. The stereochemical assignment of two methyl protons (H₃-19 and H₃-20) was deduced from ^1H - ^1H NOESY data (Supplementary Figure S15), whereas the decaline skeleton from C-1–C-10 was assigned a chair–chair configuration based on the observed correlation between H-5 and H-1, H-3 and H-9. Thus, the structure of the novel diterpene hydrocarbon (**4**) was elucidated as labda-7,13(16),14-triene.

DISCUSSION

Terpenoid metabolites are a relatively minor fraction in bacterial natural products. Since diterpenoid metabolites are extremely rare in bacterial metabolites, studies on the isolation of diterpenoid metabolites, diterpene synthases and their genes are limited. On the other hand, diterpenoid metabolites and diterpene synthases, which have been cloned predominantly from higher plants and fungi, have been extensively examined.⁹ Many bi- and tricyclic diterpenes of plants have been generated by the cyclization of GGPP under type-B and type-A diterpene synthases. In most of the type-A and type-B diterpene synthases of plant and fungal origins, metal-binding domains consisting of the acidic amino acid-rich motif sequences, -DDxxD- and -DxDD-, have been observed in type-A and type-B diterpene synthases, respectively. Furthermore, the highly conserved sequence, -SAYDTAW- (the function of which remains unknown) located within 80–150 aa of the N-terminus, was found in all diterpene synthases of plants, whereas the corresponding motif was not observed in any bacterial diterpene synthases. The type-A terpene synthases of bacterial, fungal and plant origins possess a pair of characteristically conserved metal-binding domains consisting of the acidic amino acid-rich motif -DDxxx[DE]- or -DDxxxE-, located within 60–120 aa (bacterial and fungal terpene synthases) or 230–270 aa (plant terpene synthases) of the N-terminus, and a three-amino acid triad -Nxxx[ST]xxxE-, found 140 aa downstream, closer to the C-terminus. Since the two synthases of Sros_3191 and Sare_1287 displayed the incomplete motif, -DxxxD-, these proteins may not catalyze ionization by the

removal of a diphosphate residue. Two type-A diterpene synthases, AHK61133 and BAD86798, which were classified into group 1 in the phylogenetic analysis in Figure 2, have been characterized as pimara-9(11),15-diene synthases and the acidic amino acid-rich motif -DDxxE- was present in both synthases. The same motif sequence was also observed in SCLAV_5671. On the other hand, the labda-8(17),12(*E*),14-triene synthases, CldD and RmnD, possessed the slightly longer motif, -DDxxxE-, which was also observed in SCLAV_p0491. Furthermore, another characteristically conserved metal-binding domain, the triad, was present in all bacterial type-A diterpene synthases, and was classified into two groups. The downstream triad -NxxxSxxxE- in the C-terminus region was present in four synthases: the labda-8(17),12(*E*),14-triene synthases, CldD and RmnD, and terpenete-3,13,15-triene synthases, BAB39207 and SCLAV_p0491. Another downstream triad motif -NxxxTxxxE- was in the pimara-9(11),15-diene synthases, AHK61133 and BAD86798, and SCLAV_5671. Two conserved motifs of type-B synthases, an acidic amino acid-rich motif -DxDDT[AS]- and downstream motif -QxxDGSW-, were present and conserved in all type-B synthases. Pimara-9(11),15-diene synthases displayed the acidic amino acid-rich motif -DxDDTS-, whereas the -DxDDTA- motif was observed in the two labda-8(17),12(*E*),14-triene synthases, CldB and RmnB, and SCLAV_p0490. Another downstream motif, -QxxDGSW-, was present and conserved in the pimara-9(11),15-diene synthases and labda-8(17),12(*E*),14-triene synthases. Based on these alignment and phylogenetic analyses of type-A and type-B diterpene synthases, a pair of presumptive diterpene synthases SCLAV_p0491 and SCLAV_p0490 may be involved in the generation of labdane-type bicyclic diterpenes.

The structural elucidation of diterpene products generated by the heterologous expression of genes encoding presumptive labdane-type diterpene synthases provided the biochemical functions of gene products. Since SCLAV_p0490/SCLAV_p0491 appeared to be involved in the formation of labdane-type diterpenes by a phylogenetic analysis of aligned sequences, the function of each gene product was examined. Protonation of the C-14,15-double bond of GGPP by type-B diterpene synthases, followed by the attack of C-10 on C-15, then C-7 on C-11 (Figure 6), gave four possible products ((+)-copalyl, *ent*-copalyl, *syn*-copalyl and *syn-ent*-copalyl diphosphates were generated from the chair–chair–'normal', chair–chair–'antipodal', chair–boat–'normal' and chair–boat–'antipodal' conformation of GGPPs, respectively), depending on the conformation of the *prochiral* substrate. The normal chair–chair conformation of GGPP gave the C8-carbonium ion of (+)-copalyl diphosphate with the normal *anti,anti* absolute stereochemistry in the generation of bicyclic labdane-type diterpenes. However, the antipodal chair–chair conformation of GGPP gave the C8-carbonium ion of *ent*-copalyl diphosphate with the enantiomeric *anti,anti* absolute stereochemistry in the generation of the tricyclic diterpenes, pimaradiene and *ent*-kaurene. Thus, CldB and RmnB, which are involved in the formation of the bicyclic labdane-type diterpene, labda-8(17),12(*E*),14-triene, have been defined as (+)-copalyl diphosphate synthases, while AHK61132 and BAD86797, which are involved in the formation of the tricyclic diterpene, pimara-9(11),15-diene, are *ent*-copalyl diphosphate synthases. Four different bicyclic labdane-type diterpene hydrocarbons were produced by the heterologous expression of four pairs of genes encoding type-A and type-B diterpene synthases: *cldB/cldD*, *cldB/sclav_p0491*, *sclav_p0490/cldD* and *sclav_p0490/sclav_p0491*, respectively. Transformants carrying the pair of genes *cldB/cldD* or *cldB/sclav_p0491* produced bicyclic labdane-type diterpene hydrocarbons consisting of the 8(17)-methylenedecaline skeleton, while bicyclic diterpene hydrocarbons with the 7,8-dehydrodecaline skeleton were produced by

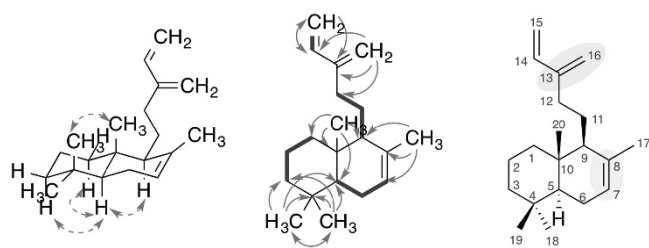


Figure 5 Structure of labda-7,13(16),14-triene (**4**). The left panel shows ^1H - ^1H NOESY (dashed arrows) data for each compound. The middle panel shows ^1H - ^1H COSY (bold lines) and ^1H - ^{13}C HMBC (arrows) and double arrows indicate a cross peak in ^1H - ^{13}C HMBC.

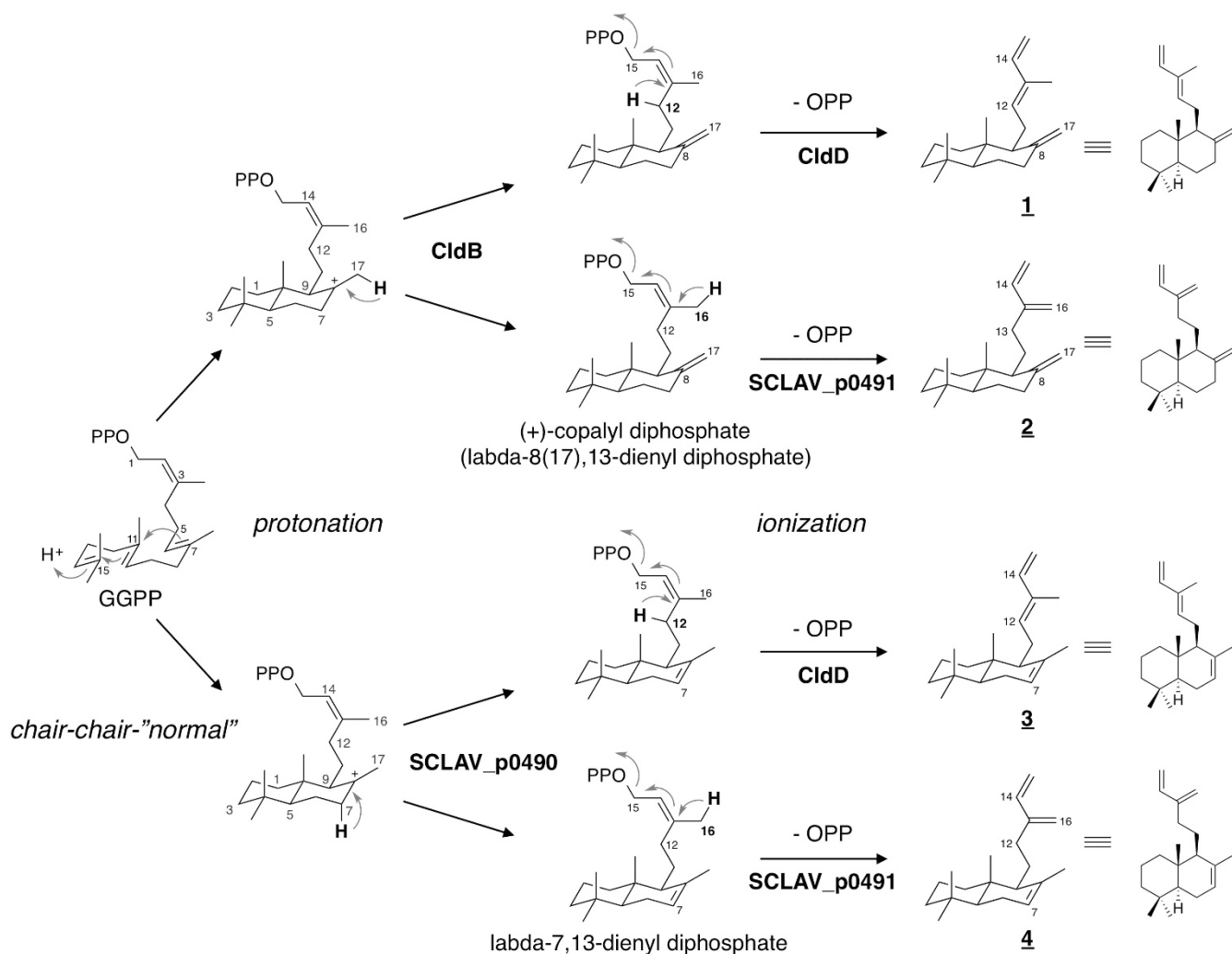


Figure 6 Cyclization mechanisms underlying the generation of labdane-type bicyclic diterpene hydrocarbons **1**, **2**, **3** and **4** by combinations of type-B and type-A diterpene synthases (CldB/CldD, CldB/SCLAV_p0491, SCLAV_p0490/CldD and SCLAV_p0490/SCLAV_p0491).

transformants carrying *sclav_p0490/cldD* or *sclav_p0490/sclav_p0491* (Figure 5). These results indicated that CldB catalyzed the generation of (+)-copalyl diphosphate from GGPP through protonation at the C-14,15-double bond of GGPP, in which the C8-carbonium ion of (+)-copalyl diphosphate was generated, and (+)-copalyl diphosphate was then formed by the release of a proton at H-17 in the C8-carbonium ion (Figure 6). On the other hand, SCLAV_p0490 catalyzed the generation of labda-8(17),13-dienyl diphosphate by a similar reaction, in which the diterpene diphosphate, labda-7,13-dienyl diphosphate, was generated by the release of a proton at H-7 in the C8-carbonium ion of (+)-copalyl diphosphate (Figure 6). The assignment of the actual biochemical functions of these type-B diterpene synthases, CldB and SCLAV_p0490, requires direct experimental verification.

Type-A diterpene synthases catalyze the generation of diterpene hydrocarbons or alcohols from copalyl diphosphate. This reaction is initiated by the enzyme-catalyzed ionization of copalyl diphosphate to form an allylic cation and the resultant cation is converted to a diterpene hydrocarbon by deprotonation. Transformants carrying the pair of genes *cldb/cldD* or *sclav_p0490/cldD* produced bicyclic diterpene hydrocarbons possessing two olefins at C-12 and C-14 in the

side chain of the labdane skeleton, while bicyclic diterpene hydrocarbons possessing two olefins at C-13(16) and C-14 in the side chain of the labdane skeleton were produced by transformants carrying *cldb/sclav_p0491* or *sclav_p0490/sclav_p0491* (Figure 5). The ionization of (+)-copalyl diphosphate (generated by CldB) or labda-7,13-dienyl diphosphate (generated by SCLAV_p0490) by CldD was then quenched by the deprotonation at H-12, whereas the cation was quenched by deprotonation at H-16 in the reaction of SCLAV_p0491 (Figure 6).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank K Miyano and T Nishide for their technical assistance in this study. This work was supported by a research Grant-in-Aid for Scientific Research on Innovative Areas from the Ministry of Education, Culture, Sports, Science and Technology of Japan (H.I.), and a research Grant-in-Aid for 'Project focused on developing key technology of discovering and manufacturing drug for next-generation treatment and diagnosis' from the Ministry of Economy, Trade and Industry of Japan (H.I.).

- 1 Gerber, N. N. Geosmin, an earthy-smelling substance isolated from actinomycetes. *Biotechnol. Bioeng.* **9**, 321–327 (1967).
- 2 Gerber, N. N. A volatile metabolite of actinomycetes, 2-methylisoborneol. *J. Antibiot.* **22**, 508–509 (1969).
- 3 Izaguirre, G., Hwang, C. J., Krasner, S. W. & McGuire, M. J. Geosmin and 2-methylisoborneol from cyanobacteria in three water supply systems. *Appl. Environ. Microbiol.* **43**, 708–714 (1982).
- 4 Jüttner, F. in *The Cyanobacteria* (eds Fay, B. & Baalen, C. V.) 453–469 (Elsevier, Amsterdam, 1987).
- 5 Wu, J.-T. & Jüttner, F. Differential partitioning of geosmin and 2-methylisoborneol between cellular constituents in *Oscillatoria tenuis*. *Arch. Microbiol.* **150**, 580–583 (1988).
- 6 Dickschat, J. S., Wenzel, S. C., Bode, H. B., Müller, R. & Schulz, S. Biosynthesis of volatiles by the myxobacterium *Myxococcus xanthus*. *ChemBioChem.* **5**, 778–787 (2004).
- 7 Yamada, Y. *et al*. Terpene synthases are widely distributed in bacteria. *Proc. Natl Acad. Sci. USA* **112**, 857–862 (2015).
- 8 Yamada, Y. *et al*. Novel terpenes generated by heterologous expression of bacterial terpene synthase genes in an engineered *Streptomyces* host. *J. Antibiot.* **68**, 385–394 (2015).
- 9 MacMillan, J. & Beale, M. H. in *Comprehensive Natural Products Chemistry: Isoprenoids Including Carotenoids and Steroids*, Vol. 2 (ed. Cane D. E.) 217–243 (Elsevier, Amsterdam, Netherlands, 1999).
- 10 Ikeda, C., Hayashi, Y., Itoh, N., Seto, H. & Dairi, T. Functional analysis of eubacterial *ent-copalyl* diphosphate synthase and pimara-9(11),15-diene synthase with unique primary sequences. *J. Biochem.* **141**, 37–45 (2007).
- 11 Xie, P. *et al*. Biosynthetic potential-based strain prioritization for natural product discovery: A showcase for diterpenoid-producing actinomycetes. *J. Nat. Prod.* **77**, 377–387 (2014).
- 12 Dairi, T. *et al*. Eubacterial diterpene cyclase genes essential for production of the isoprenoid antibiotic terpentecin. *J. Bacteriol.* **183**, 6085–6094 (2001).
- 13 Ikeda, H., Shin-ya, K., Nagamitsu, T. & Tomoda, H. Biosynthesis of mercapturic acid derivative of the labdane-type diterpene, cyclabdan that potentiates imipenem activity against methicillin-resistant *Staphylococcus aureus*: cyclabdan is generated by mycothiol-mediated xenobiotic detoxification. *J. Ind. Microbiol. Biotechnol.* (e-pub ahead of print 27 October 2015; doi: 10.1007/s10295-015-1694-6).
- 14 Komatsu, M., Uchiyama, T., Omura, S., Cane, D. E. & Ikeda, H. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc. Natl Acad. Sci. USA* **107**, 2646–2651 (2010).
- 15 Ikeda, H., Kotaki, H. & Omura, S. Genetic studies of avermectin biosynthesis in *Streptomyces avermitilis*. *J. Bacteriol.* **169**, 5615–5621 (1987).
- 16 Cane, D. E., He, X., Kobayashi, S., Omura, S. & Ikeda, H. Geosmin biosynthesis in *Streptomyces avermitilis*. Molecular cloning, expression, and mechanistic study of the germacradienol/geosmin synthase. *J. Antibiot.* **59**, 471–479 (2006).
- 17 Yamada, Y., Cane, D. E. & Ikeda, H. in *Natural Product Biosynthesis by Microorganisms and Plants, Part A, Methods in Enzymology*, Vol. 515 (ed. Hopwood D. A.) Ch. 7, 123–166 (Elsevier Inc. Academic Press, New York, 2012).
- 18 McCaskill, D. & Croteau, R. Prospects for the bioengineering of isoprenoid biosynthesis. *Adv. Biochem. Eng. BioTechnology* **55**, 107–146 (1997).
- 19 Bohlmann, J., Meyer-Gauen, G. & Croteau, R. Plant terpenoid synthases: molecular biology and phylogenetic analysis. *Proc. Natl Acad. Sci. USA* **95**, 4126–4133 (1998).
- 20 Komatsu, M., Tsuda, M., Omura, S., Oikawa, H. & Ikeda, H. Identification and functional analysis of genes controlling biosynthesis of 2-methylisoborneol. *Proc. Natl Acad. Sci. USA* **105**, 7422–7427 (2008).
- 21 Komatsu, M. *et al*. Engineered *Streptomyces avermitilis* host for heterologous expression of biosynthetic gene cluster for secondary metabolites. *ACS Synth. Biol.* **2**, 384–396 (2013).
- 22 Carman, R. M. & Grant, P. K. Biformene. *J. Chem. Soc.* **1961**, 2187–2191 (1961).
- 23 Ahond, A. & Gastambide, B. Préparation et identification des sclarènes. *Bull. Soc. Chim. Fr* **12**, 4533–4536 (1967).
- 24 Roengsumran, S., Petsom, A., Sommit, D. & Vilaivan, T. Labdane diterpenoids from *Croton oblongifolius*. *Phytochem.* **50**, 449–453 (1999).

Supplementary Information accompanies the paper on The Journal of Antibiotics website (<http://www.nature.com/ja>)