ChemPubSoc Europe

DOI: 10.1002/cbic.201402652

CHEMBIOCHEM Full Papers

Identification of a New Diterpene Biosynthetic Gene Cluster that Produces *O*-Methylkolavelool in *Herpetosiphon aurantiacus*

Chiaki Nakano, Misaki Oshima, Nodoka Kurashima, and Tsutomu Hoshino*[a]

Diterpenoids are usually found in plants and fungi, but are rare in bacteria. We have previously reported new diterpenes, named tuberculosinol and isotuberculosinol, which are generated from the *Mycobacterium tuberculosis* gene products Rv3377c and Rv3378c. No homologous gene was found at that time, but we recently found highly homologous proteins in the *Herpetosiphon aurantiacus* ATCC 23779 genome. Haur_2145 was a class II diterpene cyclase responsible for the conversion of geranylgeranyl diphosphate into kolavenyl diphosphate. Haur_2146, homologous to Rv3378c, synthesized (+)-kolavelool through the nucleophilic addition of a water molecule to the incipient cation formed after the diphosphate moiety was released. Haur_2147 afforded (+)-*O*-methylkolavelool from (+)-kolavelool, so this enzyme was an *O*-methyltransferase. This new diterpene was indeed detected in *H. aurantiacus* cells. This is the first report of the identification of a (+)-*O*-methylkolavelool biosynthetic gene cluster.

Introduction

Terpenoids are one of the largest and most diverse classes of natural products, with more than 60000 known members.[1–6] Among the terpenoids, ≈12000 are diterpenes.[5] Diterpenes are biosynthesized from geranylgeranyl diphosphate (GGPP, **1**, C_{20}, Scheme 1). The acyclic substrate **1** is generally cyclized by diterpene cyclases to produce diverse parent cyclic skeletons, which then undergo further modifications by downstream tailoring enzymes to generate numerous diterpenes. Diterpene cyclases are key enzymes responsible for the generation of such a diverse range of diterpenes. Diterpene cyclases are categorized into two different classes—class I and class II—on the basis of their reaction mechanisms and catalytic domains.[1–6] Class I enzymes catalyze cyclization through ionization of the diphosphate component of the polyprenyl diphosphate substrate. The reaction is catalyzed in a conserved α domain consisting of two conserved motifs, the DDXXD motif and the NSE/DTE motif,[1–6] which together are responsible for diphosphate-Mg^{2+}-enzyme complex formation, as well as for the ionization of the substrate. Class II cyclases, on the other hand, trigger cyclization through proton attack on the terminal double bond of the substrate; this reaction occurs in a βγ bidomain. The β domain includes the DXDD motif,[1–6] which functions to donate protons to the terminal double bond, resulting in the initiation of the cyclization reaction.

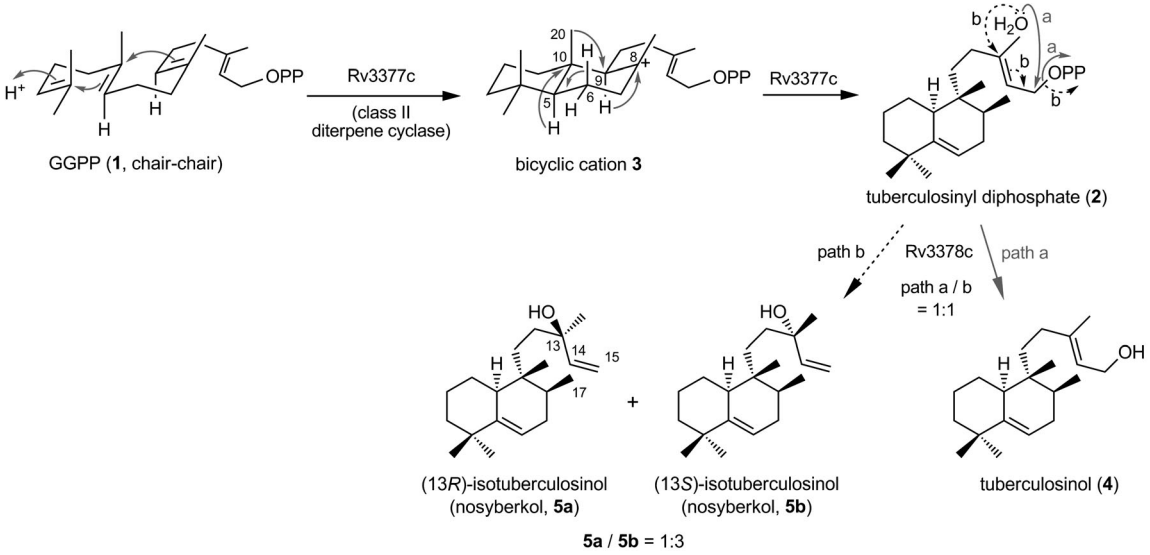
Most diterpenes are found in plants and fungi, and a large number of eukaryotic diterpene cyclases have been studied.[5,6] Bacterial diterpenes, however, are rare, and studies of bacterial diterpene cyclases are very limited,[5,6] although the synthases responsible for biosynthesis of the following diterpenes have been investigated: terpentecin,[7,8] viguiepinol,[9–11] cyclooctatin,[12] cembrane skeleton diterpenes,[13] gibberellin A_3,[14] phenalinolactone A,[15] brasilicardin A,[16] platensimycin,[17] platencin,[17] and isopimara-8,15-dien-19-ol.[18] However, the results of an increasing number of genome-sequencing projects have revealed that the capacity of bacteria to synthesize diterpenes might have been underestimated.[5,19] Identification of the biochemical functions of these terpene cyclases in bacteria might therefore lead to the discovery of new diterpenes.

In previous studies, we characterized diterpene cyclase Rv3377c and diterpene synthase Rv3378c from *Mycobacterium tuberculosis* H37Rv.[20–23] Interestingly, the *Rv3377c* and *Rv3378c* genes are found only in virulent *Mycobacterium* species, and not in avirulent ones.[20–23] *Rv3377c* encodes a diterpene cyclase that converts **1** into tuberculosinyl diphosphate (**2**), with a halimane core (Scheme 1).[20,21] Rv3377c folds **1** into a pre-chair–chair conformation; then the bicyclic cation **3** is generated by protonation-initiated cyclization. 1,2-Shifts of hydrides and a methyl group occur in an antiparallel fashion (H-9→C-8, Me-20→C-9, H-5→C-10), followed by abstraction of H-6_{ax}, introducing the double bond between C-5 and C-6 to give **2** (Scheme 1). Thus, Rv3377c is classified as a class II diterpene cyclase.

Furthermore, we revealed that the flanking gene *Rv3378c* encodes a diterpene synthase, which produces both tuberculosinol (**4**) and isotuberculosinol (**5**; alternative name nosyberkol[24]) from **2** through nucleophilic addition of a water molecule to the carbocation formed after the diphosphate moiety

[a] Dr. C. Nakano, M. Oshima, N. Kurashima, Prof. T. Hoshino
Graduate School of Science and Technology and
Department of Applied Biological Chemistry
Faculty of Agriculture, Niigata University
Niigata 950-2181 (Japan)
E-mail: hoshitsu@agr.niigata-u.ac.jp

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201402652.



Scheme 1. Biosynthetic pathways to tuberculosinol (**4**) and isotuberculosinol (nosyberkol, **5**). The Rv3377c enzyme is a class II diterpene cyclase affording tuberculosinyl diphosphate (**2**) from GGPP (**1**). The Rv3378c enzyme catalyzes the OPP-releasing reaction to afford the incipient C15 cation, which a water molecule attacks to give **4** (path a). Double-bond migration occurs concomitantly to afford a C13-cation, followed by nucleophilic attack by a water molecule, leading to the production of **5** (path b).

has been released (Scheme 1).[22,23] Rv3378c has no sequence homology to any known terpene synthase. It affords both **4** and **5** in a 1:1 ratio; **5** consists of a mixture of the diastereomers (13*R*)-isotuberculosinol (**5a**) and (13*S*)-isotuberculosinol (**5b**) in a 1:3 ratio (Scheme 1).[22]

The *M. tuberculosis* enzymatic products Rv3377c and Rv3378c are essential for the bacterium's survival inside macrophages;[25] they inhibit phagolysosome maturation[26] and macrophage phagocytosis.[22,23] Both enzymes are likely to be new potential targets in the development of new drugs. Furthermore, characterization of the Rv3378c homologue might lead to the further discovery of other new diterpenes and to more detailed understanding of the function of the Rv3378c enzyme family.

A BLAST search showed that Haur_2145 and Haur_2146 from *Herpetosiphon aurantiacus* ATCC 23779 have high amino acid sequence similarity with Rv3377c and Rv3378c, respectively. The complete genome sequence of *H. aurantiacus* was reported in 2011, and it appears to contain many biosynthetic genes for secondary metabolites.[27] *H. aurantiacus* is a Gram-negative and filamentous gliding predatory bacterium.[27]

Haur_2145 (521 amino acids) shows 29% amino acid sequence identity with the class II diterpene cyclase Rv3377c and possesses the DXDD motif (**D**G**DD**, from Asp311 to Asp314, Figure S1 in the Supporting Information). Haur_2146 (289 amino acids) shows 25% amino acid sequence identity with Rv3378c (Figure S2). In addition to these diterpene synthase homologues, Haur_2147 (283 amino acids) has high similarity to many methyltransferases that use *S*-adenosyl-L-methionine (SAM) as a methyl donor. *Haur_2145*, *Haur_2146*, and *Haur_2147* are encoded in the same direction. These genes are predicted to form an operon. We have characterized the function of the diterpene biosynthetic gene cluster in *H. aurantiacus* by in vitro analysis: the biosynthetic gene cluster is responsible for the production of (+)-*O*-methylkolavelool. The actual production of (+)-*O*-methylkolavelool by *H. aurantiacus* was experimentally demonstrated. This is the first discovery of (+)-*O*-methylkolavelool and its biosynthetic enzymes. We also discuss here the biosynthetic mechanism for the formation of (+)-*O*-methylkolavelool.

Results

In vitro analysis of Haur_2145

To elucidate the function of Haur_2145, the protein was overexpressed in *Escherichia coli* BL21(DE3) under the control of the *cspA* promoter on the pColdI vector, in which the fusion protein with a His tag at its N terminus can be expressed. The expressed Haur 2145 proteins almost formed inclusion bodies. Coproduction of the Haur_2145 proteins and the GroE chaperone[28] in *E. coli* was effective for solubilizing the recombinant protein (Figure 1). The recombinant protein thus prepared was purified by nickel-nitrilotriacetic acid (Ni-NTA) affinity chromatography to give a single SDS-PAGE band (Figure 1).

The DXDD motif is present in the Haur_2145 protein (Figure S1), but the DDXXD and NSE/DTE motifs are not found; this suggests that this enzyme may be categorized as a class II cyclase. To validate this assumption, the purified Haur_2145 protein was incubated with **1** in the presence of $MgCl_2$, and the incubation mixture was treated with acid phosphatase to hydrolyze the diphosphate moiety. The *n*-hexane extract of the reaction mixture was analyzed by GC-MS. In the *n*-hexane extract obtained without acid phosphatase treatment, no product was detected (Figure 2A). On the other hand, the new product **6** was detected as a single compound in the *n*-hexane extract that was obtained from the phosphatase treatment mixture (Figure 2B). These results strongly indicate that Haur_

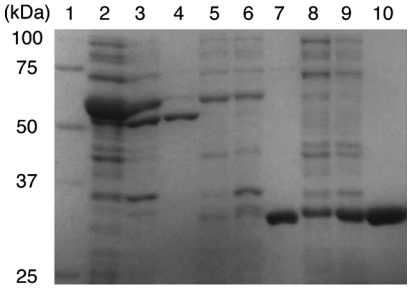


Figure 1. SDS-PAGE of overproduced and purified Haur_2145, Haur_2146, and Haur_2147. Lane 1: molecular weight marker. Lanes 2, 5, and 8: the soluble fractions of the crude extracts from *E. coli* BL21(DE3) harboring pColdI-*Haur_2145* and pT-GroE, pColdI-*Haur_2146* and pT-GroE, and pColdI-*Haur_2147*, respectively. Lanes 3, 6, and 9: the insoluble fractions of the crude extracts from *E. coli* BL21(DE3) harboring pColdI-*Haur_2145* and pT-GroE, pColdI-*Haur_2146*, and pT-GroE, and pColdI-*Haur_2147*, respectively. Lanes 4, 7, and 10: purified Haur_2145, Haur_2146, and Haur_2147, respectively.

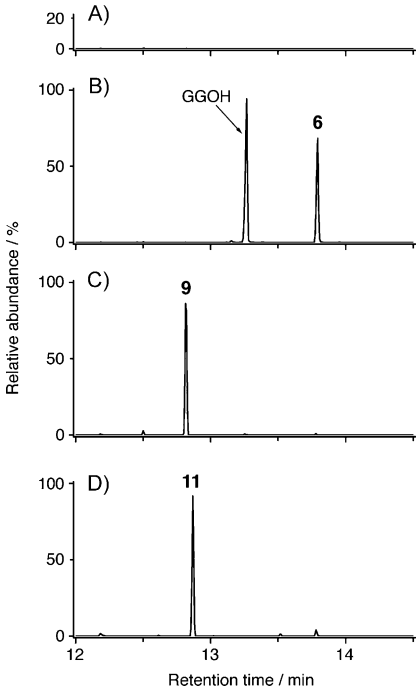


Figure 2. GC-MS analysis of the *n*-hexane extracts from the in vitro Haur_2145, Haur_2146, and Haur_2147 reaction mixtures. Incubation of Haur_2145: A) with GGPP (**1**), and B) with **1** followed by acid phosphatase treatment. Geranylgeraniol (GGOH, a peak of 13.2 min) is a hydrolytic product of **1**. C) Incubation of Haur_2146 with Haur_2145 and **1**. D) Incubation of Haur_2147 with Haur_2145, Haur_2146, **1**, and SAM. The retention times of **6**, **9**, and **11** were 13.7, 12.7, and 12.8 min, respectively.

2145 does not catalyze the OPP-releasing reaction. The Haur_2145 reaction did not proceed in the absence of $MgCl_2$, and farnesyl diphosphate (C_{15}) was not converted as a substrate (data not shown).

To determine the structure of **6**, 10 mg of **1** were incubated with the purified enzyme, and the reaction mixture was treated with acid phosphatase. The lipophilic materials were extracted from the incubation mixture with *n*-hexane. Product **6** was isolated by silica gel column chromatography with isocratic elution with *n*-hexane/EtOAc (100:0 to 100:5), followed by reversed-phase (C_{18}) HPLC (CH_3CN), yielding 2.2 mg of **6**. The molecular formula of **6** was determined to be $C_{20}H_{34}O$ by HRMS (EI).

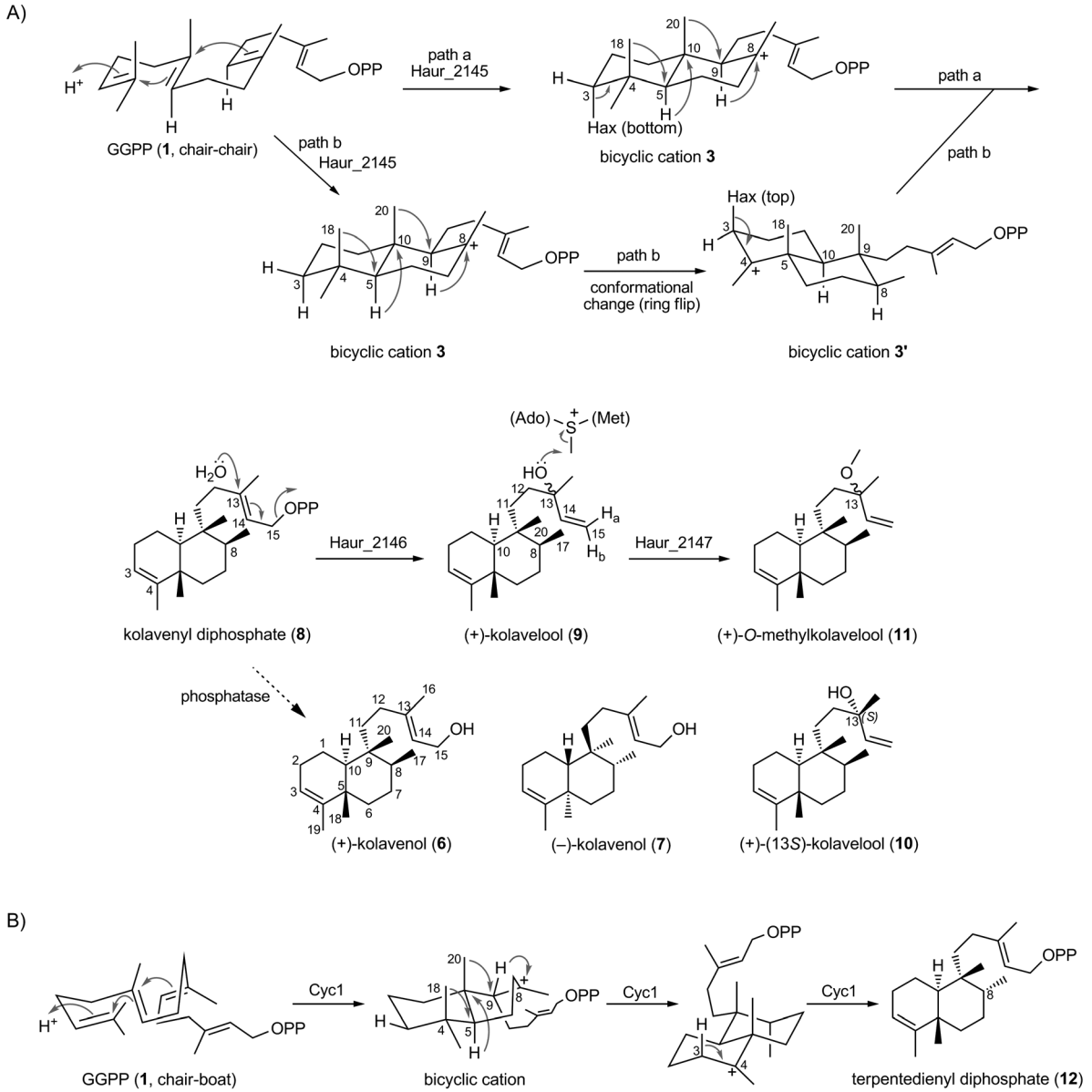
The ^{13}C NMR and DEPT spectra of **6** indicated the presence of two double bonds ($\delta_C = 120.5$ (d; C-3), 122.9 (d; C-14), 141.0 (s; C-13) 144.6 ppm (s; C-4)), five methyl carbon atoms, seven methylene carbon atoms including one primary alcoholic carbon ($\delta_C = 59.51$ ppm (t; C-15)), four methine carbon atoms, and four quaternary carbon atoms. The ^1H NMR spectrum showed a doublet methyl ($\delta_H = 0.78$ ppm (d, $J = 6.4$ Hz, 3H; Me-17)), two singlet methyl groups ($\delta_H = 0.69$ (s, 3H; Me-20), 0.97 ppm (s, 3H; Me-18)), two olefinic methyl groups ($\delta_H = 1.56$ (s, 3H; Me-19), 1.66 ppm (s, 3H; Me-16)), and two olefinic protons ($\delta_H = 5.17$ (brs, 1H; H-3), 5.37 ppm (t, $J = 6.9$ Hz, 1H; H-14)). The olefinic methyl signal ($\delta_H = 1.56$ ppm (s; Me-19)) had definitive HMBC crosspeaks with C-3 ($\delta_C = 120.5$ ppm (d)) and C-4 ($\delta_C = 144.6$ ppm (s)), thus indicating that the first double bond was situated between C-3 and C-4.

The position of the second double bond was also established, thanks to clear HMBC crosspeaks between the vinylic methyl group ($\delta_H = 1.66$ ppm (s; Me-16)) and C-13 ($\delta_C = 141.0$ ppm (s)) and C-14 ($\delta_C = 122.9$ ppm (d)). A clear HMBC crosspeak between H-14 ($\delta_H = 5.37$ ppm (t, $J = 6.9$ Hz, 1H)) and C-15 ($\delta_C = 59.51$ ppm (t)) suggested that the hydroxy group was attached to C-15. The doublet methyl group ($\delta_C = 15.96$ ppm (q; C-17)) was attached to C-8 ($\delta_C = 36.35$ ppm (d)), as was shown by the clear $^1H,^1H$ spin–spin coupling between the doublet methyl group and H-8 ($\delta_H = 1.44$ ppm (m, 1H)) in the $^1H,^1H$ COSY spectrum.

In the COSY or HOHAHA spectra, the following four spin networks were clearly observed: H-3/H-2/H-1/H-10, H-6/H-7/H-8/Me-17, H-11/H-12, and H-14/H-15. The following key HMBC crosspeaks were found: Me-19/C-5, Me-18/C-4, Me-18/C-5, Me-18/C-6, Me-18/C-10, Me-20/C-8, Me-20/C-9, Me-20/C-10, Me-20/C-11, and Me-16/C-12.

The above NMR data verified that product **6** was a bicyclic diterpene alcohol based on a clerodane skeleton (Scheme 2A and Figure S3a). The clear NOEs for Me-17/Me-20, Me-18/Me-20, and H-6α/H-8, together with the lack of NOEs for H-10/Me-18 and H-10/Me-20, supported the relative stereochemistry of the bicyclic scaffold shown in Scheme 2A and Figure S3a. These data definitively demonstrated that product **6** was the bicyclic diterpene alcohol known as kolavenol[29,30] (Scheme 2A and Figure S3a). The NMR data for **6** are in good agreement with the reported data for (−)-kolavenol (**7**, ref. [30] $[\alpha]_D^{20} = -42.9$, Scheme 2A),[29] the absolute configuration of which had been verified by its chemical synthesis.[31] The specific rotation value of **6** was $[\alpha]_D^{20} = +51.9$ ($c = 0.20$, $CHCl_3$); hence it was positive for the isolated **6**, and negative for **7**, thus indicating that **6** was the enantiomer of **7** (Scheme 2A and Figure S3a).

Because Haur_2145 is not involved in the ionization of the diphosphate moiety, the true enzymatic product of Haur_2145 would be (+)-kolavenyl diphosphate (**8**, Scheme 2A). To confirm this, further incubation of 2 mg of **1** with purified Haur_2145 enzyme was conducted. Product **8** was extracted from the incubation mixture with *n*-butanol and isolated by re-



Scheme 2. Biosynthetic pathway to (+)-*O*-methylkolavelool (**11**). A) Haur_2145 catalyzes a proton-initiated cyclization reaction to produce kolavenyl diphosphate (**8**). (+)-Kolavenol (**6**) is a hydrolytic product of **8**, formed by acid phosphatase treatment. (−)-Kolavenol (**7**) is the enantiomeric form of **6**. Haur_2146 catalyzes the conversion of **8** into (+)-kolavelool (**9**) by addition of a water molecule after ionization and allylic double bond migration. (+)-(13*S*)-Kolavelool (**10**) has been isolated from the Japanese liverwort *Jungermannia infusca*. Finally, Haur_2147 produces (+)-*O*-methylkolavelool (**11**) from **9** by transfer of a methyl group from SAM to a hydroxy group. B) The reaction mechanism of the transformation of **1** into terpentedienyl diphosphate (**12**). This reaction is mediated by Cyc1.

versed-phase (C_{18}) HPLC (CH_3CN/25 mM NH_4HCO_3 4:6). The lyophilized white solid material was subjected to NMR spectroscopy. In the ^{1}H-decoupled ^{31}P NMR spectrum, the following two chemical shifts were observed: $\delta_p = -8.00$ and -4.05 ppm. The signals were coupled to each other with a $^2J_{P,P}$ coupling constant of 19.7 Hz (Figure S4b). The ^{1}H,^{31}P HMBC spectrum verified that the diphosphate group is linked to the OH at the C-15 position of **6**, because a clear crosspeak (Figure S4c) was observed between $\delta_P = -8.00$ ppm and H-15 ($\delta_H = 4.47$ ppm). It was thus established that the true product of the Haur_2145 enzyme was kolavenyl diphosphate (**8**, Scheme 2A).

In vitro analysis of Haur_2146

Haur_2146 had amino acid sequence homology with Rv3378c (Figure S2). Thus, we predicted that Haur_2146 would catalyze the ionization of the diphosphate moiety of **8**. N-terminal His-tagged Haur_2146 was coproduced with GroE chaperone in *E. coli* for solubilization of the recombinant protein and purified by Ni-NTA affinity chromatography (Figure 1). A mixture of Haur_2146 and Haur_2145 enzymes was incubated with **1** to examine whether Haur_2146 could react with kolavenyl diphosphate (**8**). The unidentified product **9** was detected by GC-MS in the *n*-hexane extract of the incubation mixture (Fig-

ure 2C), thus suggesting that **9** is the OPP-released form. When Haur_2146 was incubated with the acyclic **1**, the ionization product of **1** was not detected (data not shown), so the OPP-releasing (diphosphate-releasing) reaction was specific to cyclic **8**, with no reaction in the case of acyclic **1**.

Two milligrams of **1** were incubated with the mixed enzymes Haur_2145 and Haur_2146. The *n*-hexane extract from the incubation mixture was subjected to SiO_2 column chromatography (*n*-hexane/EtOAc 100:0 to 100:5), followed by normal-phase HPLC column chromatography (*n*-hexane/propan-2-ol 100:0.2), to yield **9** in a pure state. The molecular formula of **9** was determined to be $C_{20}H_{34}O$ by HRMS (ESI).

Detailed NMR analyses showed that **9** had the same bicyclic clerodane skeleton as **6**. The strong NOEs for H-8 (δ_H = 1.40 ppm (m, 1H))/H-10 (δ_H = 1.28 ppm (br d, J = 12.6 Hz, 1H)) of **9** further supported the α-configuration of H-10. In addition to the bicyclic core, a vinyl group was observed in the ^{1}H NMR spectrum: H_b-15 (δ_H = 5.04 ppm (dd, $J_{H_b15,H14}$ = 10.7 Hz, J_{H_a15,H_b15} = 0.9 Hz, 1H)), H_a-15 (δ_H = 5.18 ppm (dd, $J_{H_a15,H14}$ = 17.4 Hz, J_{H_a15,H_b15} = 0.9 Hz, 1H)), and H-14 (δ_H = 5.86 ppm (dd, $J_{H_a15,H14}$ = 17.4 Hz, $J_{H_b15,H14}$ = 10.7 Hz, 1H)). The ^{13}C NMR and DEPT spectra of **9** showed that the C-13 signal resonated at 73.46 ppm as a singlet, thus indicating that the tertiary alcohol group was positioned at C-13. Clear HMBC crosspeaks of Me-16/C-12, Me-16/C-13, Me-16/C-14, Me-20/C-11, and H-12/C-11 were observed; therefore, a 3-methylpent-1-en-3-ol moiety was connected to the bicyclic skeleton. The specific rotation value of **9** was $[\alpha]_D^{18}$ = +22.4 (c = 0.03, $CHCl_3$). These results indicated that **9** was (+)-kolavelool[32] (Scheme 2A and Figure S5a).

Haur_2146 is therefore a diterpene synthase that produces (+)-kolavelool (**9**) from **8**. In a previous study, we had established that **5** produced by the Rv3378c enzymatic reaction was a diastereomeric mixture of (13*R*)-**5** and (13*S*)-**5** in a 1:3 ratio (Scheme 1).[22] ^{1}H NMR analysis of the (13*R*)- and (13*S*)-**5** mixture showed that the stronger signals for H-14, H_a-15, and Me-17 were accompanied by weaker peaks for the corresponding protons (1/3 intensity) and that the spin–spin coupling patterns of these small peaks were the same as those of the large peaks.[22] No satellite peaks corresponding to H-14, H_a-15, H_b-15, and Me-17 were observed in the ^{1}H NMR spectrum of **9**, thus indicating that **9** has a single configuration at C-13. The specific rotation of (+)-(13*S*)-kolavelool (**10**) was reported to be $[\alpha]_D^{18}$ = +59.1 (c = 0.68, $CHCl_3$).[32] The difference between the numerical values of the specific rotation for **9** ($[\alpha]_D^{18}$ = +22.4) and for **10** ($[\alpha]_D^{18}$ = +59.1) suggests that the configuration at C-13 of **9** might be inferred to be *R*, but other analytical methods including X-ray crystallographic analysis of **9** would be required in order to establish the C-13 stereochemistry without ambiguity.

In vitro analysis of Haur_2147

A BLAST search with Haur_2147 revealed sequence homology with methyltransferases. The recombinant Haur_2147 protein was expressed in *E. coli* as an N-terminal His-tagged protein and purified by Ni-NTA affinity chromatography (Figure 1). In order to test for methyltransferase activity of **9**, we incubated **1** with supplementation of the methyl donor SAM and the three mixed enzymes Haur_2145, Haur_2146, and Haur_2147. Figure 2D shows that the new product **11** resulted from this incubation experiment.

To prepare product **11** in sufficient amounts for a structural analysis, 2 mg of **1** were incubated under the conditions described above. SiO_2 column chromatography (*n*-hexane/EtOAc 100:0 to 100:5), followed by normal-phase HPLC (*n*-hexane/propan-2-ol 100:0.06), gave 0.5 mg of pure **11**. The molecular formula of **11** was determined to be $C_{21}H_{36}O$ by HRMS (ESI), thus indicating that a C_1 unit was appended to **11**. ^{1}H and ^{13}C NMR analysis suggested the presence of a methoxy group (δ_H = 3.12 ppm (s, 3H); δ_C = 49.92 ppm (q)). The methoxy signal had a clear HMBC crosspeak with C-13 (δ_C = 77.62 ppm (s)), thus suggesting that the methoxy group was connected to C-13. Further analyses of NMR data indicated that the structure of **11** was consistent with *O*-methylated **9** (Scheme 2A and Figure S6a). The specific rotation value of **11** was $[\alpha]_D^{20}$ = +102.3. Therefore, **11** was determined to be (+)-*O*-methylkolavelool. Incubation of **1** with Haur_2147 and with the mixed enzymes Haur_2145 and Haur_2147 in the presence of SAM did not generate methylated **1** or methylated **8**, respectively (data not shown). Thus, this methyltransferase selectively accepts the cyclic **9** as a substrate to produce (+)-*O*-methylkolavelool (**11**, Scheme 2A).

In vivo production of (+)-*O*-methylkolavelool by *H. aurantiacus*

To examine whether **11** is produced in vivo by *H. aurantiacus*, the bacterial species was cultured in CY medium, and the lipophilic materials were extracted with EtOAc. The lipophilic materials were fractionated on the basis of compound polarity by SiO_2 column chromatography. The fractionations were performed stepwise with use of mixtures of *n*-hexane and EtOAc (100:0 to 100:5), and each fraction obtained was analyzed by GC-MS (Figure 3). In the fraction eluted with *n*-hexane/EtOAc (100:2), the peak corresponding to authentic **11** was found (Figure 3); the retention time (12.8 min, Figure 3) and MS fragment of the peak were identical to those of **11** (Figure S7). From this result, we concluded that *Haur_2145*, *Haur_2146*, and *Haur_2147* were not cryptic genes and that they did indeed function to generate (+)-*O*-methylkolavelool in *H. aurantiacus*.

Discussion

In this study, we demonstrated that the gene cluster consisting of *Haur_2145*, *Haur_2146*, and *Haur_2147* is responsible for the synthesis of (+)-*O*-methylkolavelool (**11**) in *H. aurantiacus* (Scheme 2A). Haur_2145 encodes a diterpene cyclase that catalyzes the synthesis of kolavenyl diphosphate (**8**) from GGPP (**1**). Haur_2146 converts **8** into (+)-kolavelool (**9**). Haur_2147 functions to catalyze *O*-methylation of the hydroxy group in **9**, resulting in the production of **11**.

Scheme 2A shows the proposed cyclization mechanism leading to **8**. GGPP (**1**) is folded into a pre-chair–chair "normal"

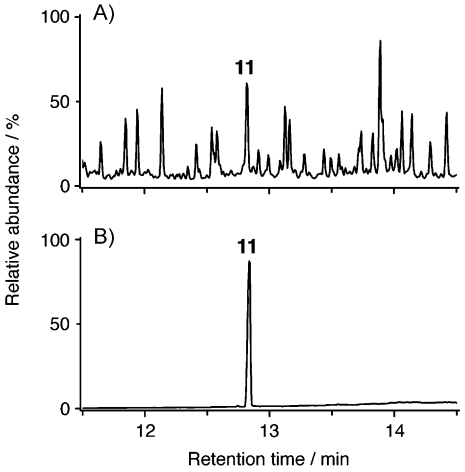


Figure 3. GC-MS analysis of terpenes produced by *H. aurantiacus*. GC-MS chromatograms of partially purified samples of: A) *H. aurantiacus* cell extracts, and B) (+)-*O*-methylkolavelool (**11**). The peak at 12.8 min in (A) was identified as **11** (peak at 12.8 min in (B)).

conformation.[33] The terminal double bond of **1** is protonated by an acidic proton provided by the aspartic acid in the DXDD motif to initiate the cyclization, and the 6/6-fused bicyclic **3** with C-8 cation is generated by Markovnikov closure. A series of 1,2-shifts of hydride and methyl group occur in an antiparallel fashion (H-9→C-8, Me-20→C-9, H-5→C-10, Me-18→C-5). Next, elimination of H-3 occurs to introduce the double bond between C-3 and C-4. Abstraction of the *bottom* H-3_{ax} (path a) could give **8**. The bicyclic skeleton with C-4 cation underwent a conformation change (ring flip) to afford **3′**, followed by abstraction of the *top* H-3_{ax} to afford **8** (path b). To ascertain which pathway is operative, isotope-labeling experiments are necessary. In the case in which the incubation experiment is conducted in D_2O, the deuterium atom would remain at C-3 in the case of path a, whereas the deuterium atom would be lost in that of path b. Incubation experiments with D_2O and isotope-labeled **1** will be reported in due course.

The Cyc1 enzyme from *Kitasatospora griseola*, a bacterial class II diterpene cyclase, produces the bicyclic terpentedienyl diphosphate (**12**, Scheme 2B),[8] which possesses the same clerodane scaffold as **8**. However, the stereochemistry at C-8 in terpentedienyl diphosphate (**12**) is opposite to that in **8**, thus indicating that the folding conformation of **1** differs between the syntheses of **12** and **8**: a chair-boat conformation is adopted in the case of **12**,[34] whereas a chair–chair conformation is adopted in that of **8**, although the 1,2-migration reaction of hydride and methyl groups, followed by deprotonation, occurs in a similar fashion (Scheme 2B). Haur_2145 was the first identified kolavenyl diphosphate synthase. Haur_2145 folded **1** into a pre-chair–chair "normal" conformation to produce (+)-kolavenyl diphosphate (**8**). On the other hand, (−)-kolavenyl diphosphate (**7**-PP), the enantiomeric form of **8**, is generated from the pre-chair–chair-"antipodal" conformation of **1**. The folding conformations for diterpene biosyntheses are discussed in detail by Peters.[33] It is not surprising that enantiomers—(+)- and (−)-copalyl diphosphates, for example—should both exist in nature. However, no information on how the diterpene cyclases differentiate the production of both enantiomeric diterpenes from the perspective of the enzymes' structures is currently available. Class II diterpene cyclases require metal ions for optimal activity;[8,11,20,21,35–37] they are thought to bind both with the diphosphate group and with the acidic amino acid of the cyclase enzyme.[35–37] Recently, the crystal structure of *ent*-copalyl diphosphate (*ent*-CDP) synthase, a class II diterpene cyclase, from *Arabidopsis thaliana* has been solved.[36,37] The E211 residue of this *ent*-CDP synthase is located near the diphosphate group of the substrate and is thought to be a ligand for Mg^{2+}.[36,37] Bacterial class II diterpene cyclases display somewhat high sequence similarity (<30%) to the N-terminal halves of eukaryotic class II diterpene cyclases.[5] The glutamic acid residue corresponding to E211 of the *ent*-CDP synthase is also conserved in bacterial class II diterpene cyclases (Figure S1), thus suggesting that this glutamic acid residue might act as a substrate binding in coordination with a divalent metal ion.

Haur_2146, a homologue of Rv3378c, synthesizes **9** from **8**. (+)-(13*S*)-Kolavelool (**10**, Scheme 2A) has been isolated from the Japanese liverwort *Jungermannia infusca*,[32] but not from bacteria. Elimination of the diphosphate moiety of **8** could result in the incipient C-15 carbocation. The C-13=C-14 to C-14=C-15 double-bond rearrangement process would generate the C-13 cation, from which nucleophilic attack by a water molecule would produce the alcohol **9**, possessing a single stereocenter at C-13. On the other hand, Rv3378c synthesizes a mixture of **4**, (13*R*)-**5**, and (13*S*)-**5**.[22] This indicates that the regiospecificity of the addition of a water molecule by Haur_2146 is more strictly controlled than that of Rv3378c. Very recently, the crystal structure of Rv3378c was solved by an international collaboration including ourselves.[38] The overall fold of the Rv3378c protein shows close structural homology to the ζ fold that is seen in *cis*-head-to-tail prenyl diphosphate synthases, which are involved in cell wall biosynthesis.[38] For undecaprenyl diphosphate synthase in *E. coli*, D26 and R30 are involved both in the binding of Mg^{2+} with farnesyl diphosphate and in the ionization of the diphosphate, but the DDXXD motif is not present.[38,39] These two residues correspond to D34 and R38 in Rv3378c (Figure S2), and are predicted to play a crucial role in the ionization of the diphosphate moiety in the formation of **5**, on the basis of the crystal structure and mutational analysis.[38] The DDXXD sequence is present in Rv3378c, but is not adjacent to the active site.[38] The mechanism of the ionization of **2** catalyzed by Rv3378c is similar to that of *cis*-prenyl diphosphate synthases.[38] From the crystal structure and mutational analysis, the residues Y51 and Y90 in Rv3378c play a role in the formation of **5**, whereas Y51 is mainly involved in formation of **4**.[38] The residues equivalent to R38, Y51, and Y90 in Rv3378c are conserved in Haur_2146, but the D34 residue is not conserved in Haur_2146 (Figure S2). Furthermore, the DDXXD sequence is not present in Haur_2146. The crystal structure and mutational analyses of Haur_2146 will be required to address the differences between Haur_2146 and Rv3378c in terms of product specificity and catalytic mechanism. Very recently, Layre et al. isolated 1-tuberculosinyladeno-

sine from the lipid fraction of *M. tuberculosis* H37Rv;[40] this indicates that Rv3378c catalyzes the dephosphorylation of tuberculosinyl diphosphate (**2**) and the transfer reaction of the moiety to adenosine.[40] However, tuberculosinol and isotuberculosinol have also indeed been detected from extracts of *M. tuberculosis* H37Rv cells.[41] It is likely that Rv3378c is a bifunctional enzyme, producing 1-tuberculosinyladenosine and tuberculosinol/isotuberculosinol. *O*-Methylkolavelool was detected in extracts of *H. aurantiacus* cells (Figure 3); this indicates strongly that Haur_2146 also catalyzes the nucleophilic addition of a water molecule to the C-13 cation generated by the ionization of the diphosphate moiety of **8** to produce **9**, as shown in Scheme 2A. Haur_2149, located in the neighborhood of Haur_2145 to Haur_2147, is tentatively assigned as a 5′-nucleotidase-domain-containing protein; this suggests that Haur_2146 might catalyze the formation of kolavenyladenosine in a similar way to Rv3378c. To validate this assumption, we are planning to perform further experiments to determine whether kolavenyladenosine is produced or not.

Haur_2147 affords **11** from **9** through the transfer of a methyl group from SAM to a hydroxy group (Scheme 2A). Several bicyclic diterpenes containing the 3-methylpent-1-en-3-ol moiety have been isolated from nature; they include **5**,[24,41] manool,[42,43] *ent*-manool,[44,45] vitexifolin A,[46] and sclareol.[47,48] However, to the best of our knowledge, no bicyclic diterpene containing a 3-methoxy-3-methylpent-1-ene moiety has previously been isolated. Compound **11** is thus an unprecedented diterpene; Haur_2147 is the first enzyme that catalyzes the methylation of the hydroxy group of the 3-methylpent-1-en-3-ol side chain. We have succeeded in the enzymatic syntheses of the above-mentioned diterpenes[23] and also of bicyclic diterpenes containing the buta-1,3-diene moiety.[49] In vitro treatment of these bicyclic diterpene alcohols with Haur_2147 could lead to the generation of the unnatural *O*-methylated diterpenes. This trial is now in progress in our laboratory.

Conclusions

In conclusion, we have identified a new diterpene biosynthetic gene cluster synthesizing (+)-*O*-methylkolavelool (**11**) in *H. aurantiacus*. Haur_2145 is a new diterpene cyclase that produces kolavenyl diphosphate (**8**) from GGPP (**1**). The diterpene synthase Haur_2146 catalyzes a reaction involving the addition of water after the ionization of the diphosphate moiety of **8** to produce (+)-kolavelool (**9**). Finally, Haur_2147 acts as an *O*-methyltransferase to produce the new diterpene **11** from **9**. Note that **11** was found in *H. aurantiacus* cells; this suggests that the diterpene biosynthetic gene cluster in *H. aurantiacus* does indeed function in (+)-*O*-methylkolavelool biosynthesis. This is the first identification of (+)-*O*-methylkolavelool (**11**) and its biosynthetic gene cluster in nature. Rv3377c and Rv3378c from *M. tuberculosis* are attractive targets for the development of new antituberculosis drugs, and comparison of the crystal structures and catalytic mechanisms of these diterpene synthases from *M. tuberculosis* and from *H. aurantiacus* should therefore provide further information to facilitate drug development.

Experimental Section

General analytical methods: NMR spectra were recorded in $CDCl_3$ with a DMX 600 spectrometer (Bruker). Chemical shifts in ^1H and ^{13}C NMR spectra (ppm) are relative to 7.24 ($CHCl_3$) and 77.0 ppm ($CDCl_3$), respectively. Coupling constants (*J*) are given in Hz. The ^1H and ^{31}P NMR spectra of kolavenyl diphosphate (**8**) were measured in D_2O containing 25 mM NH_4HCO_3, with a DMX 600 spectrometer. The chemical shifts of ^1H are referenced to the 3-(trimethylsilyl)propionic acid sodium salt (TSP) signal, whereas those of ^{31}P were measured relative to 85% H_3PO_4, by using external referencing. Gas chromatography-time-of-flight mass spectrometry (GC-TOFMS) analysis was performed with a JEOL (Tokyo, Japan) JMS-T100GCV instrument under electronic impact at 70 eV with a Zebron ZB-5 MS capillary column (30 m × 0.25 mm inside diameter, 0.25 μm film thickness, Phenomenes, Torrance, CA), with the oven temperature being elevated from 60 (3 min hold) to 260 °C (20 °C min^{-1}). HRMS was performed with a JMS-T100GCV instrument and use of the electron ionization (EI) mode or with a JMS-T100LP (JEOL) instrument and use of the electrospray ionization (ESI) mode. The specific rotation values were measured with a SEPA-300 high-sensitivity polarimeter (Horiba, Kyoto, Japan) at 20 or 18 °C in $CHCl_3$ as a solvent. The product weights were measured with a microbalance (M1-20A, Chyo, Japan, accuracy range ±0.02 mg), and the measurement temperature was kept constant by use of a CTS-134A temperature control system (IWAKI, Tokyo, Japan, ±0.05 °C).

Bacterial strains, plasmids, and media: *E. coli* strains JM109 and BL21(DE3) as well as plasmids pUC119 and pColdI were purchased from Takara Bio (Shiga, Japan). *H. aurantiacus* ATCC 23779 was obtained from the American Type Culture Collection. *H. aurantiacus* was cultured in CY medium, based on casitone (0.5%) and yeast extract (0.1%).

Construction of plasmids: The *H. aurantiacus* genomic DNA was used as a template for PCR. The absence of undesired alterations during PCR was confirmed by nucleotide sequencing. A 1.6 kb DNA fragment containing the *Haur_2145* coding sequence was amplified by PCR with primer I (5′-GCG TCT AGA *CAT* ***ATG*** AGC TTA ATT GTT G-3′ (NdeI site italicized and the start codon of *Haur_2145* shown in boldface)) and primer II (5′-GCG AAG CTT *GGA TCC* TCA TAA TTT TAA TTC TCC-3′ (BamHI site italicized)). The resultant PCR product was digested with NdeI and BamHI and cloned between the NdeI and BamHI sites of pColdI, resulting in pColdI-*Haur_2145*.

A 0.9 kb DNA fragment containing the *Haur_2146* coding sequence was amplified by PCR with primer III (5′-GCG GAA TTC *CAT* ***ATG*** AGG CCC ACG CCC ACG TTA G-3′ (EcoRI site underlined, NdeI site italicized, and the start codon of *Haur_2146* shown in bold)) and primer IV (5′-GCG *GGA TCC* CTA CAA CTC TGA TTC AAT GCT G-3′ (BamHI site italicized)). The amplified DNA fragment was cloned between the EcoRI and BamHI sites of pUC119 to give pUC119-*Haur_2146*. The NdeI-BamHI fragment excised from pUC119-*Haur_2146* was cloned between the NdeI and BamHI sites of pColdI to give pColdI-*Haur_2146*.

A 0.9 kb DNA fragment containing the *Haur_2147* coding sequence was amplified by PCR with primer V (5′-GCG GAA TTC *CTC GAG* **ATG** ACA GTC CTA TCA ACC G-3′ (EcoRI site underlined, XhoI site italicized, and the start codon of *Haur_2147* shown in bold)) and primer VI (5′-GCG CTG CAG *GGA TCC* TTA GCC ACG TTT ACG CGC GAC-3′ (BamHI site italicized)). The amplified DNA fragment was cloned between the EcoRI and BamHI sites of pUC119 to give pUC119-*Haur_2147*. The XhoI-BamHI fragment excised from pUC119-*Haur_2147* was cloned between the XhoI and BamHI sites of pColdI to give pColdI-*Haur_2147*.

Production and purification of recombinant proteins: To produce N-terminal His_6-tagged fusion proteins, pColdI-*Haur_2145* and pColdI-*Haur_2146* were individually transformed into *E. coli* BL21(DE3) harboring pT-GroE,[28] and the transformant was grown at 37 °C in lysogeny broth (LB; 1 L) supplemented with ampicillin (50 μg mL^{-1}) and chloramphenicol (3.75 μg mL^{-1}). *E. coli* BL21(DE3) harboring pColdI-*Haur_2147* was grown at 37 °C in LB (1 L) supplemented with ampicillin (50 μg mL^{-1}). When the OD_{600} had reached 0.4, the culture was kept at 15 °C for 30 min. IPTG (final 10 μM) was added, and cultivation was continued for 24 h at 15 °C. Cells were harvested by centrifugation and resuspended in buffer A (Tris·HCl (50 mM), NaCl (300 mM), imidazole (5 mM), glycerol (10%), pH 8.0, 30 mL)). After sonication, cell debris was removed by centrifugation at 10000 *g* for 20 min. Recombinant His-tagged protein was purified over Ni-NTA agarose (Qiagen) according to the manufacturer's instructions, with the exception that glycerol (10%) was added to all buffers. The purified Haur_2145 and Haur_2146 proteins were independently dialyzed against buffer B (Tris·HCl (50 mM), glycerol (10%), pH 8.0). The purified Haur_2147 protein was subjected to Sephadex G-10 (GE Healthcare) gel filtration chromatography in buffer B. The purities of the recombinant proteins were checked by SDS-PAGE on 12% gel. The protein concentration was determined by a Bio-Rad Protein Assay (Bio-Rad, Tokyo, Japan) with bovine serum albumin as a standard.

In vitro enzyme assay: For the analysis of the Haur_2145 enzymatic product, the reaction mixture, consisting of Tris·HCl (pH 7.5, 50 mM), $MgCl_2$ (5 mM), GGPP (44 μM), and Haur_2145 (2.4 μM) in a total volume of 2.5 mL, was incubated at 30 °C for 24 h. To hydrolyze the diphosphate moiety of the product, the phosphatase reaction mixture (7.5 mL), containing Haur_2145 reaction mixture (2.5 mL), acetate buffer (0.2 M, pH 5.6), propan-2-ol (20%, *v/v*), and acid phosphatase (0.25 mg mL^{-1}, Sigma), was then further incubated at 37 °C for 12 h. For the analysis of the Haur_2146 enzymatic product, the reaction mixture, consisting of Tris·HCl (pH 7.5, 50 mM), $MgCl_2$ (5 mM), GGPP (44 μM), Haur_2145 (2.4 μM), and Haur_2146 (8.8 μM) in a total volume of 2.5 mL, was incubated at 30 °C for 24 h. For the analysis of the Haur_2147 enzymatic product, the reaction mixture, consisting of Tris·HCl (pH 7.5, 50 mM), $MgCl_2$ (5 mM), GGPP (44 μM), SAM (100 μM), Haur_2145 (2.4 μM), Haur_2146 (8.8 μM), and Haur_2147 (8.5 μM) in a total volume of 2.5 mL, was incubated at 30 °C for 24 h. After addition of MeOH (4 mL for Haur_2145 or 2 mL for Haur_2146 and Haur_2147 reaction mixtures), the product was extracted with *n*-hexane (2×3 mL). The *n*-hexane layer was concentrated to dryness. The crude materials were dissolved in *n*-hexane (200 μL) and subjected to GC-MS analysis (2 μL).

Isolation of diterpenes: To isolate product **6**, the reaction mixture (500 mL), containing Tris·HCl buffer (pH 7.5, 50 mM), $MgCl_2$ (5 mM), GGPP (10 mg, 44 μM), and Haur_2145 protein (2.5 μM), was incubated at 30 °C for 24 h. The phosphatase reaction mixture (1500 mL), containing Haur_2145 reaction mixture (500 mL), acetate buffer (0.2 M, pH 5.6), propan-2-ol (20%, *v/v*), and acid phosphatase (0.25 mg mL^{-1}), was then further incubated at 37 °C for 12 h. After addition of MeOH (750 mL), the lipophilic materials were extracted with *n*-hexane (3×600 mL). The *n*-hexane layer was dried over Na_2SO_4 and concentrated to dryness. The crude materials were dissolved in a small amount of *n*-hexane. The enzymatic product was purified by SiO_2 column chromatography with *n*-hexane/EtOAc (100:0 to 100:5) followed by reversed-phase HPLC (CAPCELL PAK C18 MG S5 column, 250×15 mm; SHISEIDO, Tokyo, Japan; mobile phase CH_3CN; flow rate 3 mL min^{-1}; detection at 210 nm), yielding **6** (2.2 mg).

To purify product **8** formed by the Haur_2145 protein, the reaction mixture (100 mL), containing Tris·HCl buffer (pH 7.5, 50 mM), $MgCl_2$ (5 mM), GGPP (2 mg, 44 μM), and Haur_2145 protein (2.5 μM), was incubated at 30 °C for 24 h. After the incubation, EDTA (pH 8.0, 0.5 M, 2 mL) was added to terminate the reaction. Product **8** was extracted with *n*-butanol (40 mL×3), followed by concentration to dryness. The dried material was dissolved in a small volume of CH_3CN/25 mM NH_4HCO_3 (4:6) and then fractionated by reversed-phase HPLC (CAPCELL PAK C18 MG S5 column, 250×15 mm; SHISEIDO; mobile phase CH_3CN/25 mM NH_4HCO_3 4:6; flow rate 3 mL min^{-1}; detection at 210 nm). After the removal of the organic solvent, the aqueous layer was lyophilized to give **8** (0.3 mg) as a white powder.

The Haur_2146 reaction mixture (100 mL), containing Tris·HCl buffer (pH 7.5, 50 mM), $MgCl_2$ (5 mM), GGPP (2 mg, 44 μM), Haur_2145 protein (2.5 μM), and Haur_2146 protein (18.6 μM), was incubated at 30 °C for 24 h. The Haur_2147 reaction mixture (100 mL), containing Tris·HCl buffer (pH 7.5, 50 mM), $MgCl_2$ (5 mM), GGPP (2 mg, 44 μM), SAM (100 μM), Haur_2145 protein (2.5 μM), Haur_2146 protein (18.6 μM), and Haur_2147 (7.7 μM), was incubated at 30 °C for 24 h. After addition of MeOH (50 mL), the lipophilic materials were extracted with *n*-hexane (3×50 mL). The *n*-hexane layer was dried over Na_2SO_4 and concentrated to dryness. The crude materials were dissolved in a small amount of *n*-hexane. The enzymatic product was partially purified by SiO_2 column chromatography with *n*-hexane/EtOAc (100:0 to 100:5). Further purification of the product was performed by SiO_2 HPLC (Inertsil SIL 100 A, 7.6× 250 mm; GL Science, Tokyo, Japan; flow rate 3 mL min^{-1}; detection at 210 nm) with *n*-hexane/propan-2-ol (100:0.2 for **9** or 100:0.06 for **11**), yielding **9** (0.4 mg) and **11** (0.5 mg). The structures of the diterpenes were determined by NMR, with use of ^1H, ^{13}C, DEPT, ^1H–^1H COSY, HOHAHA, NOESY, HSQC and HMBC experiments, HRMS (EI) or HRMS (ESI), and specific rotation analysis.

Production of 11 by *H. aurantiacus*: *H. aurantiacus* was grown in CY medium (100 mL×4) at 30 °C for 60 h. Cells were harvested by centrifugation and resuspended in saturated aq. NaCl solution, then disrupted by sonication. Compounds of low polarity were then extracted from the crude cell lysate with EtOAc. The organic layer was dried over Na_2SO_4 and concentrated to dryness. The crude materials were partially purified by SiO_2 column chromatography with *n*-hexane/EtOAc (100:0 to 100:5) and then analyzed by GC-MS.

Spectroscopic data for products 6, 9, and 11

Product 6: Oil; $[\alpha]_D^{20} = +51.9$ ($c = 0.20$, $CHCl_3$), error rage having been estimated to be ±0.5; ^1H NMR (600 MHz, $CDCl_3$): $\delta = 0.69$ (s, 3H; Me-20), 0.78 (d, $J = 6.4$ Hz, 3H; Me-17), 0.97 (s, 3H; Me-18), 1.15 (ddd, $J = 12.4$, 12.4, 4.0 Hz, 1H; H-6), 1.33 (m, 1H; H-10), 1.35 (m, 1H; H-11), 1.38 (m, 2H; H-7), 1.41 (m, 1H; H-1), 1.44 (m, 1H; H-8), 1.45 (m, 1H; H-11), 1.56 (m, 1H; H-1), 1.56 (s, 3H; Me-19), 1.66 (s, 3H; H-16), 1.68 (ddd, $J = 12.8$, 3.2, 3.2 Hz, 1H; H-6), 1.79 (ddd, $J = 12.8$, 12.8, 4.7 Hz, 1H; H-12), 1.85 (ddd, $J = 12.7$, 12.7, 4.1 Hz, 1H; H-12), 2.01 (m, 2H; H-2), 4.12 (d, $J = 6.9$ Hz, 2H; H-15), 5.17 (br s, 1H; H-3), 5.37 ppm (t, $J = 6.9$ Hz, 1H; H-14); ^{13}C NMR (150 MHz, $CDCl_3$): $\delta = 15.96$ (q; C-17), 16.50 (q; C-16), 17.93 (q; C-19), 18.31 (t; C-1), 18.38 (q; C-20), 19.98 (q; C-18), 26.91 (t; C-2), 27.58 (t; C-7), 32.89 (t; C-12), 36.35 (d; C-8), 36.85 (t; C-11), 36.93 (t; C-6), 38.25 (s; C-5), 38.68 (s; C-9), 46.53 (d; C-10), 59.51 (t; C-15), 120.5 (d; C-3), 122.9 (d; C-14), 141.0 (s; C-13), 144.6 ppm (s; C-4); MS (EI): *m/z* (%): 69 (37), 81 (42), 95 (100), 107 (73), 120 (61), 189 (63), 257 (6), 290 (3) $[M]^+$; HRMS (EI): *m/z* calcd for $C_{20}H_{34}O$: 290.26096 $[M]^+$; found: 290.26074.

Product 9: Oil; $[\alpha]_D^{18} = +22.4$ ($c = 0.03$, $CHCl_3$); ^1H NMR (600 MHz, $CDCl_3$): $\delta = 0.69$ (s, 3H; Me-20), 0.75 (d, $J = 5.9$ Hz, 3H; Me-17), 0.96 (s, 3H; Me-18), 1.14 (ddd, $J = 12.8$, 12.8, 4.0 Hz, 1H; H-6), 1.23 (m, 1H; H-11), 1.23 (s, 3H; Me-16), 1.28 (br d, $J = 12.6$ Hz, 1H; H-10), 1.35 (m, 2H; H-12), 1.35 (m, 1H; H-11), 1.36 (m, 2H; H-7), 1.38 (m, 1H; H-1), 1.40 (m, 1H; H-8), 1.52 (m, 1H; H-1), 1.55 (s, 3H; Me-19), 1.67 (ddd, $J = 12.6$, 4.0, 4.0 Hz, 1H; H-6), 1.98 (m, 2H; H-2), 5.04 (dd, $J = 10.7$, 0.9 Hz, 1H; H_b-15), 5.16 (br s, 1H; H-3), 5.18 (dd, $J = 17.4$, 0.9 Hz, 1H; H_a-15), 5.86 ppm (dd, $J = 17.4$, 10.7 Hz, 1H; H-14); ^{13}C NMR (150 MHz, $CDCl_3$): $\delta = 15.92$ (q; C-17), 17.93 (q; C-19), 18.28 (t; C-1), 18.46 (q; C-20), 19.95 (q; C-18), 26.88 (t; C-2), 27.55 (t; C-7), 27.81 (q; C-16), 31.87 (t; C-11), 35.37 (t; C-12), 36.22 (d; C-8), 36.92 (t; C-6), 38.21 (s; C-5), 38.41 (s; C-9), 46.44 (d; C-10), 73.46 (s; C-13), 111.7 (t; C-15), 120.5 (d; C-3), 144.5 (s; C-4), 145.3 ppm (d; C-14); MS (EI): *m/z* (%): 71 (44), 81 (51), 95 (100), 107 (87), 121 (54), 189 (61), 257 (16), 272 (4) $[M-H_2O]^+$; HRMS (ESI) (negative ion mode): *m/z* calcd for $C_{20}H_{33}O$: 289.25314 $[M-H]^-$; found: 289.25266.

Product 11: Oil; $[\alpha]_D^{20} = +102.3$ ($c = 0.02$, $CHCl_3$); ^1H NMR (600 MHz, $CDCl_3$): $\delta = 0.69$ (s, 3H; Me-20), 0.76 (d, $J = 6.0$ Hz, 3H; Me-17), 0.96 (s, 3H; Me-18), 1.13 (ddd, $J = 12.9$, 12.9, 3.8 Hz, 1H; H-6), 1.19 (s, 3H; Me-16), 1.22 (m, 1H; H-11), 1.30 (m, 1H; H-10), 1.32 (m, 1H; H-11), 1.32 (m, 1H; H-12), 1.36 (m, 2H; H-7), 1.39 (m, 1H; H-1), 1.40 (m, 1H; H-12), 1.41 (m, 1H; H-8), 1.51 (m, 1H; H-1), 1.55 (s, 3H; Me-19), 1.66 (ddd, $J = 12.9$, 3.8, 3.8 Hz, 1H; H-6), 1.99 (m, 2H; H-2), 3.12 (s, 3H; Me-21), 5.10 (br d, $J = 17.7$ Hz; 1H, H_a-15), 5.15 (br s, 1H; H-3), 5.15 (br d, $J = 10.9$ Hz, 1H; H_b-15), 5.70 ppm (dd, $J = 17.7$, 10.9 Hz, 1H; H-14); ^{13}C NMR (150 MHz, $CDCl_3$): $\delta = 15.96$ (q; C-17), 17.93 (q; C-19), 18.36 (t; C-1), 18.47 (q; C-20), 19.96 (q; C-18), 21.40 (q; C-16), 26.89 (t; C-2), 27.57 (t; C-7), 31.57 (t; C-11), 32.42 (t; C-12), 36.25 (d; C-8), 36.92 (t; C-6), 38.23 (s; C-5), 38.45 (s; C-9), 46.37 (d; C-10), 49.92 (q; C-21), 77.62 (s; C-13), 114.6 (t; C-15), 120.4 (d; C-3), 143.1 (d; C-14), 144.6 ppm (s; C-4); MS (EI): *m/z* (%): 69 (13), 81 (20), 85 (100), 95 (53), 107 (31), 189 (26), 257 (8), 272 (2) $[M-CH_3OH]^+$; HRMS (ESI) (positive ion mode): *m/z* calcd for $C_{21}H_{36}NaO$: 327.26639 $[M+Na]^+$; found: 327.26657.

Acknowledgements

We thank Prof. Y. Asakawa and Dr. F. Nagashima (Tokushima Bunri University) for valuable suggestions regarding kolavelool. We also thank Prof. A. F. Barrero (University of Granada, Spain) for valuable suggestions about kolavenol. This work was financially supported by Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Young Scientists (B) (25850079 to C.N.) and for Scientific Research (C) (25450150 to T.H.).

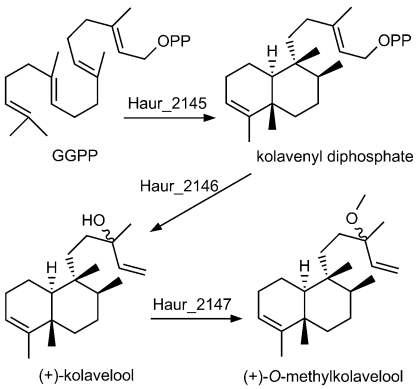
Keywords: diterpene cyclases · diterpene synthases · *Herpetosiphon aurantiacus* · natural products · terpenoids

[1] D. W. Christianson, *Chem. Rev.* **2006**, *106*, 3412 – 3442.
[2] D. W. Christianson, *Curr. Opin. Chem. Biol.* **2008**, *12*, 141 – 150.
[3] E. Oldfield, F.-Y. Lin, *Angew. Chem. Int. Ed.* **2012**, *51*, 1124 – 1137; *Angew. Chem.* **2012**, *124*, 1150 – 1163.
[4] R. Cao, Y. Zhang, F. M. Mann, C. Huang, D. Mukkamala, M. P. Hudock, M. E. Mead, S. Prisic, K. Wang, F.-Y. Lin, T.-K. Chang, R. J. Peters, E. Oldfield, *Proteins Struct. Funct. Bioinf.* **2010**, *78*, 2417 – 2432.
[5] M. J. Smanski, R. M. Peterson, S.-X. Huang, B. Shen, *Curr. Opin. Chem. Biol.* **2012**, *16*, 132 – 141.
[6] H.-Y. Gong, Y. Zeng, X.-Y. Chen, *Nat. Prod. Bioprospect.* **2014**, *4*, 59 – 72.
[7] T. Dairi, Y. Hamano, T. Kuzuyama, N. Itoh, K. Furihata, H. Seto, *J. Bacteriol.* **2001**, *183*, 6085 – 6094.
[8] Y. Hamano, T. Kuzuyama, N. Itoh, K. Furihata, H. Seto, T. Dairi, *J. Biol. Chem.* **2002**, *277*, 37098 – 37104.
[9] T. Kawasaki, T. Kuzuyama, Y. Kuwamori, N. Matsuura, N. Itoh, K. Furihata, H. Seto, T. Dairi, *J. Antibiot.* **2004**, *57*, 739 – 747.
[10] T. Kawasaki, Y. Hayashi, T. Kuzuyama, K. Furihata, N. Itoh, H. Seto, T. Dairi, *J. Bacteriol.* **2006**, *188*, 1236 – 1244.
[11] C. Ikeda, Y. Hayashi, N. Itoh, H. Seto, T. Dairi, *J. Biochem.* **2007**, *141*, 37 – 45.
[12] S.-Y. Kim, P. Zhao, M. Igarashi, R. Sawa, T. Tomita, M. Nishiyama, T. Kuzuyama, *Chem. Biol.* **2009**, *16*, 736 – 743.
[13] A. Meguro, T. Tomita, M. Nishiyama, T. Kuzuyama, *ChemBioChem* **2013**, *14*, 316 – 321.
[14] D. Morrone, J. Chambers, L. Lowry, G. Kim, A. Anterola, K. Bender, R. J. Peters, *FEBS Lett.* **2009**, *583*, 475 – 480.
[15] C. Dürr, H.-J. Schnell, A. Luzhetskyy, R. Murillo, M. Weber, K. Welzel, A. Vente, A. Bechthold, *Chem. Biol.* **2006**, *13*, 365 – 377.
[16] Y. Hayashi, N. Matsuura, H. Toshima, N. Itoh, J. Ishikawa, Y. Mikami, T. Dairi, *J. Antibiot.* **2008**, *61*, 164 – 174.
[17] M. J. Smanski, Z. Yu, J. Casper, S. Lin, R. M. Peterson, Y. Chen, E. Wendt-Pienkowski, S. R. Rajski, B. Shen, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 13498 – 13503.
[18] M. Xu, M. L. Hillwig, A. L. Lane, M. S. Tiernan, B. S. Moore, R. J. Peters, *J. Nat. Prod.* **2014**, *77*, 2144 – 2147.
[19] P. Xie, M. Ma, M. E. Rateb, K. A. Shaaban, Z. Yu, S.-X. Huang, L.-X. Zhao, X. Zhu, Y. Yan, R. M. Peterson, J. R. Lohman, D. Yang, M. Yin, J. D. Rudolf, Y. Jiang, Y. Duan, B. Shen, *J. Nat. Prod.* **2014**, *77*, 377 – 387.
[20] C. Nakano, T. Okamura, T. Sato, T. Dairi, T. Hoshino, *Chem. Commun.* **2005**, 1016 – 1018.
[21] C. Nakano, T. Hoshino, *ChemBioChem* **2009**, *10*, 2060 – 2071.
[22] C. Nakano, T. Ootsuka, K. Takayama, T. Mitsui, T. Sato, T. Hoshino, *Biosci. Biotechnol. Biochem.* **2011**, *75*, 75 – 81.
[23] T. Hoshino, C. Nakano, T. Ootsuka, Y. Shinohara, T. Hara, *Org. Biomol. Chem.* **2011**, *9*, 2156 – 2165.
[24] A. Rudi, M. Aknin, E. Gaydou, Y. Kashman, *J. Nat. Prod.* **2004**, *67*, 1932 – 1935.
[25] K. Pethe, D. L. Swenson, S. Alonso, J. Anderson, C. Wang, D. G. Russell, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 13642 – 13647.
[26] F. M. Mann, M. Xu, X. Chen, D. B. Fulton, D. G. Russell, R. J. Peters, *J. Am. Chem. Soc.* **2009**, *131*, 17526 – 17527.
[27] H. Kiss, M. Nett, N. Domin, K. Martin, J. A. Maresca, A. Copeland, A. Lapidus, S. Lucas, K. W. Berry, T. Glavina Del Rio, E. Dalin, H. Tice, S. Pitluck, P. Richardson, D. Bruce, L. Goodwin, C. Han, J. C. Detter, J. Schmutz, T. Brettin, M. Land, L. Hauser, N. C. Kyrpides, N. Ivanova, M. Göker, T. Woyke, H.-P. Klenk, D. A. Bryant, *Stand. Genomic Sci.* **2011**, *5*, 356 – 370.
[28] T. Yasukawa, C. Kanei-Ishii, T. Maekawa, J. Fujimoto, T. Yamamoto, S. Ishii, *J. Biol. Chem.* **1995**, *270*, 25328 – 25331.
[29] A. Ohsaki, L. T. Yan, S. Ito, H. Edatsugi, D. Iwata, Y. Komoda, *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2889 – 2892.
[30] A. Sawabe, K. Minemoto, S. Ouchi, T. Okamoto, *J. Jpn. Oil Chem. Soc.* **1999**, *48*, 139 – 142.
[31] E. Piers, J. Y. Roberge, *Tetrahedron Lett.* **1992**, *33*, 6923 – 6926.
[32] F. Nagashima, M. Suzuki, S. Takaoka, Y. Asakawa, *J. Nat. Prod.* **2001**, *64*, 1309 – 1317.
[33] R. J. Peters, *Nat. Prod. Rep.* **2010**, *27*, 1521 – 1530.
[34] T. Eguchi, Y. Dekishima, Y. Hamano, T. Dairi, H. Seto, K. Kakinuma, *J. Org. Chem.* **2003**, *68*, 5433 – 5438.
[35] S. Prisic, R. J. Peters, *Plant Physiol.* **2007**, *144*, 445 – 454.
[36] M. Köksal, H. Hu, R. M. Coates, R. J. Peters, D. W. Christianson, *Nat. Chem. Biol.* **2011**, *7*, 431 – 433.
[37] M. Köksal, K. Potter, R. J. Peters, D. W. Christianson, *Biochim. Biophys. Acta Gen. Subj.* **2014**, *1840*, 184 – 190.
[38] H.-C. Chan, X. Feng, T.-P. Ko, C.-H. Huang, Y. Hu, Y. Zheng, S. Bogue, C. Nakano, T. Hoshino, L. Zhang, P. Lv, W. Liu, D. C. Crick, P.-H. Liang, A. H.-J. Wang, E. Oldfield, R.-T. Guo, *J. Am. Chem. Soc.* **2014**, *136*, 2892 – 2896.
[39] K.-H. Teng, P.-H. Liang, *Bioorg. Chem.* **2012**, *43*, 51 – 57.
[40] E. Layre, H. J. Lee, D. C. Young, A. J. Martinot, J. Buter, A. J. Minnaard, J. W. Annand, S. M. Fortune, B. B. Snider, I. Matsunaga, E. J. Rubin, T. Alber, D. B. Moody, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 2978 – 2983.
[41] L. Prach, J. Kirby, J. D. Keasling, T. Alber, *FEBS J.* **2010**, *277*, 3588 – 3595.
[42] A. F. Barrero, J. F. Sanchez, E. J. Alvarez-Manzaneda, M. M. Dorado, A. Haidour, *Phytochemistry* **1993**, *32*, 1261 – 1265.

[43] K. Bruns, *Tetrahedron Lett.* **1970**, *11*, 3263 – 3264.
[44] F. Nagashima, H. Tanaka, Y. Asakawa, *Phytochemistry* **1996**, *42*, 93 – 96.
[45] A. Urzua, E. Tojo, J. Soto, *Phytochemistry* **1995**, *38*, 555 – 556.
[46] M. Ono, T. Yanaka, M. Yamamoto, Y. Ito, T. Nohara, *J. Nat. Prod.* **2002**, *65*, 537 – 541.
[47] D. V. Banthorpe, J. T. Brown, G. S. Morris, *Phytochemistry* **1990**, *29*, 2145 – 2148.
[48] A. Ulubelen, G. Topcu, C. Eri, U. Sönmez, M. Kartal, S. Kurucu, C. Bozok-Johansson, *Phytochemistry* **1994**, *36*, 971 – 974.
[49] C. Nakano, T. Hoshino, T. Sato, T. Toyomasu, T. Dairi, T. Sassa, *Tetrahedron Lett.* **2010**, *51*, 125 – 128.

Received: November 9, 2014
Published online on ■■ ■■, 0000

Unique bacterial diterpene: A diterpene biosynthetic gene cluster consisting of a diterpene cyclase, a diterpene synthase, and a methyltransferase was found in the *Herpetosiphon aurantiacus* genome. This study demonstrated that this gene cluster is responsible for (+)-*O*-methylkolavelool biosynthesis. This is the first discovery of (+)-*O*-methylkolavelool from nature.



*C. Nakano, M. Oshima, N. Kurashima, T. Hoshino**

■■ – ■■

Identification of a New Diterpene Biosynthetic Gene Cluster that Produces *O*-Methylkolavelool in *Herpetosiphon aurantiacus*

These are not the final page numbers! ↗↗