Biosynthesis of a novel cyclic C_{35}-terpene *via* the cyclisation of a *Z*-type C_{35}-polyprenyl diphosphate obtained from a nonpathogenic *Mycobacterium* species†

Tsutomu Sato,* Atsushi Kigawa, Ryosuke Takagi, Tomomi Adachi and Tsutomu Hoshino

Received 20th May 2008, Accepted 20th June 2008
First published as an Advance Article on the web 19th August 2008
DOI: 10.1039/b808513g

Lipid components from 12 nonpathogenic *Mycobacterium* species were analysed. A novel cyclic C_{35}-terpene, named heptaprenylcycline **1**, was obtained from 3 species, while octahydroheptaprenol **2**, which has 3 *Z*-double bonds, was obtained from 6 species. The amounts of **1** and **2** in the cultured cells increased after the 4- to 6-d stationary phase. The yield of **1** was considerably greater at a higher temperature of 37 °C than at an optimal temperature of 28 °C, while that of **2** remained unchanged at all temperatures. A feeding experiment with D-[1-^{13}C]glucose revealed that **1** was produced *via* isopentenyl diphosphate, which is a metabolite of glycolysis and the methylerythritol phosphate pathway. The conversion of octahydroheptaprenyl diphosphate **2-PP** to **1** was successful by using the cell-free extracts of *M. chlorophenolicum*, demonstrating that **2-PP** is the biosynthetic intermediate of **1**. This is the first example of the biosynthesis of a natural terpene *via* the cyclisation of a linear C_{35}-isoprenoid. The substrate **2-PP** for C_{35}-terpene cyclase has *Z*-type prenyl moieties; however, terpene cyclases usually employ *E*-type isoprenoids. The gene encoding the terpene cyclase that cyclises prenyl diphosphate containing *Z*-double bonds as the natural substrate has not yet been detected. Despite a careful search using the FASTA3 program, we could not detect any gene that is homologous to the known diphosphate-triggered type of mono-, sesqui- and diterpene cyclases in the genome of *M. vanbaalenii*, the DNA sequence of which has recently been elucidated. This suggests that a novel type of terpene cyclase might exist in the nonpathogenic *Mycobacterium* species.

Introduction

The genus *Mycobacterium* includes more than 90 species.[1] Several pathogens causing diseases such as tuberculosis and leprosy have been extensively studied due to their medical importance.[2] On the other hand, a few nonpathogens are likely to be used for bioremediation because of their ability to degrade a wide range of environmentally toxic chemicals such as polycyclic aromatic hydrocarbons (PAHs) and pentachlorophenols (PCPs).[3] The genomes of 10 pathogenic strains (*M. tuberculosis* H37Rv, *M. tuberculosis* CDC1551, *M. tuberculosis* F11, *M. tuberculosis* H37Ra, *M. bovis* AF2122/97, *M. bovis* BCG Pasteur 1173P2, *M. leprae* TN, *M. avium* 104, *M. avium* subsp. *paratuberculosis* K-10 and *M. ulcerans* Agy99) have been analysed. In addition, the genomes of 6 nonpathogens (*M. vanbaalenii* PYR-1, *M. gilvum* PYR-GCK, *Mycobacterium* sp. JLS, *Mycobacterium* sp. KMS, *Mycobacterium* sp. MCS and *M. smegmatis* MC2) have also been elucidated.[4] Recently, we have demonstrated that the Rv3377c gene in the *M. tuberculosis* H37Rv genome encodes a novel diterpene cyclase that plays a role in the production of the halimane skeleton of tuberculosinyl diphosphate (Fig. 1).[5] The Rv3377c

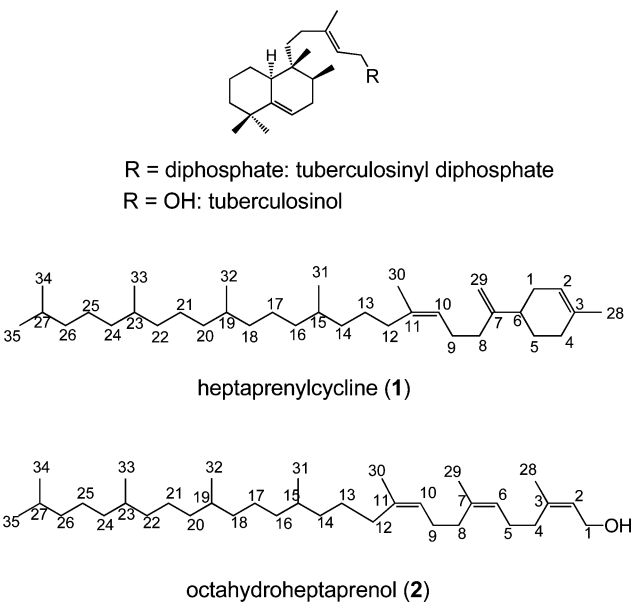

Fig. 1 Structure of tuberculosinyl diphosphate, tuberculosinol, heptaprenylcycline (**1**) and octahydroheptaprenol (**2**).

Department of Applied Biological Chemistry, Faculty of Agriculture, and Graduate School of Science and Technology, Niigata University, Ikarashi, Niigata, 950-2181, Japan. E-mail: satot@agr.niigata-u.ac.jp; Fax: +81-25-262-6854

† Electronic supplementary information (ESI) available: Amount of C_{35} terpenes in the *M. chlorophenolicum* cells cultured under different durations and temperatures, comparison of ^{13}C NMR spectrum of ^{13}C-labelled **1** with natural compound **1** and structure of C_{35}-terpenes found in nature, search for the terpene cyclase gene from the *M. vanbaalenii* genome, sesquiterpene cyclase utilising farnesyl diphosphate with *Z*-double bonds as the natural substrate, NMR spectra of **1** and NMR spectra of **2**. See DOI: 10.1039/b808513g

Table 1 Nonpathogenic *Mycobacterium* species analysed in this study and their type, which was determined on the basis of their yields of compounds **1** and **2**

Species	1	2	Type
M. chlorophenolicum	○	○	I
M. thermoresistibile	○	○	
M. vanbaalenii	○	○	
M. aichiense	–	○	II
M. smegmatis	–	○	
M. parafortuitum	–	○	
M. agri	–	–	III
M. aurum	–	–	
M. diernhoferi	–	–	
M. moriokaense	–	–	
M. nonchromogenicum	–	–	
M. pulveris	–	–	

○: detected and –: not detected.

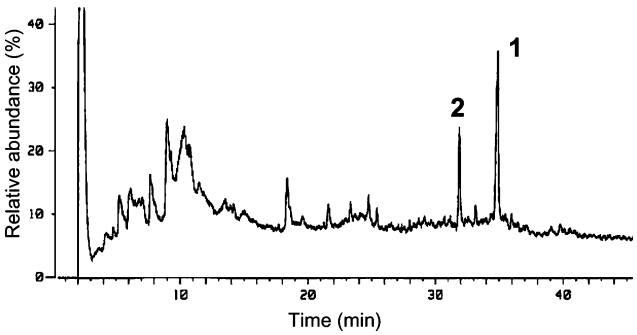

Fig. 2 GC-MS chromatogram displaying the total ions in nonsaponifiable lipids. The data of *M. chlorophenolicum* are shown as a representative example. The retention time of tuberculosinol was 11.2 min under the same conditions.

homologous genes are limited to other *Mycobacterium* species causing tuberculosis (*M. tuberculosis* CDC1551, *M. tuberculosis* F11, *M. tuberculosis* H37Ra and *M. bovis*) and are not found in the other species. On the basis of the findings of transposon-insertion experiments, Pethe *et al.* reported that Rv3377c was indispensable for survival in macrophages.[6] Therefore, we have proposed that the Rv3377c gene may be responsible for the pathogenicity of tuberculosis.[5]

In order to confirm this hypothesis, lipid constituents from 12 nonpathogenic *Mycobacterium* species (Table 1) were analysed. We detected a novel monocyclic C_{35}-terpene **1** from 3 species and octahydroheptaprenol **2** from 6 species (Fig. 1). Herein, we demonstrate that **1** is biosynthesised by the cyclisation of a linear octahydroheptaprenyl diphosphate containing *Z*-double bonds; the cyclisation of this molecule has never been reported hitherto.

Results and discussion

Lipid analyses of nonpathogenic *Mycobacterium* spp

According to the method described in the experimental section, the hexane extract of the nonsaponifiable lipids from the cells and the ethyl acetate (EtOAc) extract from the broth filtrate were obtained from the 12 mycobacteria cultures; both the extracts were subjected to gas chromatography-mass spectrometry (GC-MS). Tuberculosinol was not detected (Fig. 2); this may support the theory that tuberculosinol is produced only by pathogenic species and that the Rv3377c enzyme encoding diterpene cyclase may be involved in the pathogenicity of these bacteria. In contrast, compounds **1** and/or **2** were detected in the nonsaponifiable fraction obtained from the cells of several species (Fig. 2 and Table 1). The 12 species were categorised on the basis of the production of **1** and **2** as shown in Table 1: (type I) both **1** and **2** were produced by *M. chlorophenolicum*, *M. thermoresistibile* and *M. vanbaalenii*; (type II) only **2** was produced by *M. aichiense*, *M. smegmatis* and *M. parafortuitum* and (type III) both **1** and **2** were not produced by *M. agri*, *M. aurum*, *M. diernhoferi*, *M. moriokaense*, *M. nonchromogenicum* and *M. pulveris*.

Isolation and structural determination of 1 and 2

Among the mycobacteria examined, *M. chlorophenolicum* produced the largest amounts of **1** and **2**. Thus, we employed *M. chlorophenolicum* cells for the isolation of **1** and **2**. After a 50 L cultivation of *M. chlorophenolicum*, the cells were saponified with 15% KOH–MeOH at 80 °C for 30 min and were then extracted with *n*-hexane. The nonsaponifiable lipids were subjected to silica gel column chromatography, followed by reverse-phase high-performance liquid chromatography (HPLC); this process yielded 5.5 mg and 5.1 mg of pure **1** and **2**, respectively. The structure of **1** was determined by detailed nuclear magnetic resonance (NMR) data analysis (DEPT, COSY, HOHAHA, NOESY, HMQC and HMBC) and electron impact mass spectrometry (EIMS). The complete NMR assignments are shown in Table 2. The molecular formula of **1** was assigned to be $C_{35}H_{64}$ on the basis of high resolution (HR)-EIMS. As shown in Fig. 3, a partial structure (right side) of compound **1** (C-1–C-14 and C-28–C-30) was mainly assigned according to the data obtained from HMBC and COSY spectra. A clear NOE was observed between H-10 (δ_H 5.41, t, *J* 6.8) and Me-30 (δ_H 1.87, s), indicating that H-10 and Me-30 were arranged in the *Z*-configuration. The comparison of a chemical shift of C-12 (32.42 ppm) with that of C-12 in ficaprenol[7] (*Z*: *ca.* 32.0 ppm, *E*: *ca.* 40.0 ppm), which was previously reported, further supported the *Z*-geometry. The structure of the isobutyl moiety on the left side of compound **1** (C-26, C-27, C-34 and C-35) was determined by HMBC and COSY (Fig. 3) spectral analyses. The carbon signals, which were yet to be assigned, demonstrated 3 methyl groups at δ_C 24.94–25.26 (blank square), 3 methylene groups at δ_C 24.94–25.26 (blank circle), 2 methylene groups at δ_C 37.71–37.91 (filled circle) and 3 methyne groups at δ_C 33.12–33.24 (filled square) (Fig. 3). HMBC data demonstrated that the 3 isopentyl moieties represented by dotted squares in Fig. 3 were linked to each other. The complete structure of compound **1** could be elucidated by connecting the 3 partial structures obtained from the HMBC cross peaks between H-12 (δ_H 2.21, m) and C-14 (δ_C 37.28) and between H-26 (δ_H 1.64, m) and C-25 (δ_C 25.26). The stereochemistries at C-6, C-15, C-19 and C-23 remain to be elucidated. Compound **1** is novel and its proposed name is heptaprenylcycline.

The molecular formula of **2** was assigned to be $C_{35}H_{66}O$ on the basis of HR-EIMS data. The detailed 2D-NMR analysis revealed that **2** was octahydroheptaprenol, as shown in Fig. 1. Compounds **3** and **4** were isolated from *M. smegmatis*[8] (Fig. 4), and compound **2** was obtained by acid hydrolysis of **3** and **4**. The NMR assignment of **2** (Table 3), however, has not been reported. NOE correlation

Table 2 NMR assignments of **1** in C_6D_6 and abundance of ^{13}C-labeled **1**

No.	δ_H (mult., J in Hz)	δ_C (mult.)	^{13}C Enrichment[b]
1	2.11 (m); 2.29 (m)	31.83 (t)	**4.5**
2	5.56 (1H, br t)	121.26 (d)	1.1
3	—	133.43 (s)	1.1
4	1.97 (2H, m)	30.96 (t)	1.0
5	1.63 (m)	28.68 (t)	**3.1**
6	2.23 (2H, m)	40.15 (d)	1.6
7	—	154.06 (s)	1.5
8	2.26 (2H, m)	35.63 (t)	1.0
9	2.41 (2H, m)	27.18 (t)	**4.0**
10	5.41 (1H, t, 6.8)	125.37 (d)	1.4
11	—	135.59 (s)	1.7
12	2.21 (2H, m)	32.42 (t)	1.0
13	1.57 (m); 1.62 (m)	25.84 (t)	**2.8**
14	1.29 (m); 1.51 (m)	37.28 (t)	1.1
15	1.61 (m)	33.12 (d)[a]	1.0
16	1.29 (m); 1.51 (m)	37.71 (t)[a]	1.0
17	1.42 (m); 1.54 (m)	24.95 (t)[a]	**4.1**
18	1.29 (m); 1.51 (m)	37.84 (t)[a]	1.6
19	1.61 (m)	33.22 (d)[a]	1.5
20	1.29 (m); 1.51 (m)	37.85 (t)[a]	1.6
21	1.42 (m); 1.54 (m)	24.97 (t)[a]	**4.1**
22	1.29 (m); 1.51 (m)	37.90 (t)[a]	1.6
23	1.61 (m)	33.24 (d)[a]	1.1
24	1.29 (m); 1.51 (m)	37.91 (t)[a]	1.7
25	1.42 (m); 1.54 (m)	25.26 (t)[a]	**4.5**
26	1.64 (m)	28.32 (t)	1.7
27	1.31 (m)	39.73 (d)	1.2
28	1.75 (3H, s)	23.62 (q)[a]	**4.5**
29	5.03 (1H, s); 5.05 (1H, s)	107.80 (t)	**3.0**
30	1.87 (3H, s)	23.62 (q)[a]	**2.9**
31	1.06 (3H, d, 6.8)	19.91 (q)[a]	**3.0**
32	1.07 (3H, d, 6.8)	19.98 (q)[a]	**3.1**
33	1.08 (3H, d, 6.7)	20.03 (q)[a]	**3.4**
34	1.04 (3H, d, 6.6)	22.789 (q)[a]	**2.6**
35	1.04 (3H, d, 6.6)	22.89 (q)[a]	1.0

[a] The assignments may be exchangeable. [b] The values for enrichments were determined by comparison of the relative peak intensities of the corresponding carbons in labelled and unlabelled spectra. The bold type shows the ^{13}C-enriched position (>2.5-fold).

and the chemical shifts of C-4, C-8 and C-12 (32.42, 32.50 and 32.57 ppm)[7] suggested that all the 3 prenyl residues in **2** have *Z*-configurations. The stereochemistries at C-15, C-19 and C-23 could not be determined. Compound **2** was not detected in the nonsaponified (no alkaline-treatment) samples obtained from all the 12 *Mycobacterium* species. The octahydroheptaprenyl moieties of compounds **3** and **4** are stable under alkaline conditions,[8] strongly indicating that an unknown compound possessing **2** as the moiety exists in type I and II species.

Yields of 1 and 2 under different culture conditions

Species that produce **1** are categorised as type I species. These exhibited special abilities that were not found in type II and III species. *M. chlorophenolicum* and *M. vanbaalenii* can degrade PCPs and PAHs, respectively.[3] *M. thermoresistibile* can survive and propagate at elevated temperatures of up to 55 °C.[9] It is interesting to elucidate the functions of compound **1** in *Mycobacterium* cells. As the first step towards the functional analysis of **1**, we determined the amounts of **1** and **2** produced in *M. chlorophenolicum* cells under varied culture conditions—time and temperature. The yields of **1** and **2** increased after the 4- to 6-d stationary phase. Similar results have been observed for many secondary metabolites produced by micro-organisms. Moreover, the yield of compound **1** was considerably greater at a higher temperature of 37 °C than at an optimal growth temperature (28 °C), while the amount of **2** remained unchanged at all temperatures. A similar phenomenon has been noted during the analysis of mycolic acid: in several *Mycobacterium* species, the amount and composition of mycolic acid varied at higher temperatures.[9,10] The adaptive changes in mycolic acids in response to temperature changes are assumed to play a role in the maintenance of suitable membrane function.[9,10] It was proposed that **3** may be involved in the transport of mycolic acids through the plasma membrane in order to constitute the cell-wall component of a covalently linked peptidoglycan–arabinogalactan–mycolate framework.[8] Since both **1** and **3** may be biosynthesised *via* a common intermediate octahydroheptaprenyl diphosphate **2-PP** (Fig. 4), the function of **1** may be closely related to that of mycolic acid. However, further experiments are required to establish the functions of **1**.

Biosynthesis of 1

As shown in Fig. 4, we propose that the pathway for the biosynthesis of **1** is as follows. The carbon framework of **1** is produced by the head-to-tail condensations of 7 isopentenyl diphosphate (IPP) units, followed by hydrogenation reactions that occur at 4 double bonds to give rise to **2-PP** and a diphosphate-triggered type of cyclisation to form the complete structure of compound **1**. The detection of **2** strongly supports the existence of the intermediate **2-PP**, because **2** is produced by the dephosphorylation of **2-PP**. It has been reported that mannosyl-β-1-phosphoisoprenoid **5** bearing a C_{32} polyprenyl-like moiety with a structure analogous to those of **3** and **4**, is found in *M. tuberculosis*, and that the C_{32} moiety may be produced by polyketide synthase.[11] The proposed elongation reaction is primed by a fatty acyl group and proceeds by alternate condensations of methylmalonate and malonate in repeated cycles to yield a carboxylic acid product.[11] In order to identify the biosynthetic pathway involved in the production of **1**, we carried out a feeding experiment using D-[1-^{13}C]glucose. *M. chlorophenolicum* was cultured in a 10 L medium with 10% isotopic abundance of D-[1-^{13}C]glucose, and 0.2 mg of ^{13}C-labelled **1** was isolated. ^{13}C NMR analysis revealed that the isotopic abundant signals (Table 2) of **1** were in agreement with the position of the filled circles shown in Fig. 4; this unambiguously demonstrated

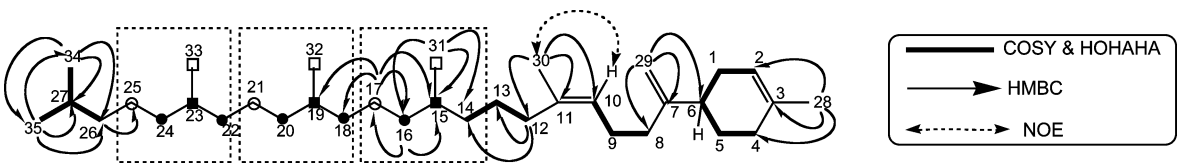

Fig. 3 Significant correlations in the NMR spectra of compound **1**.

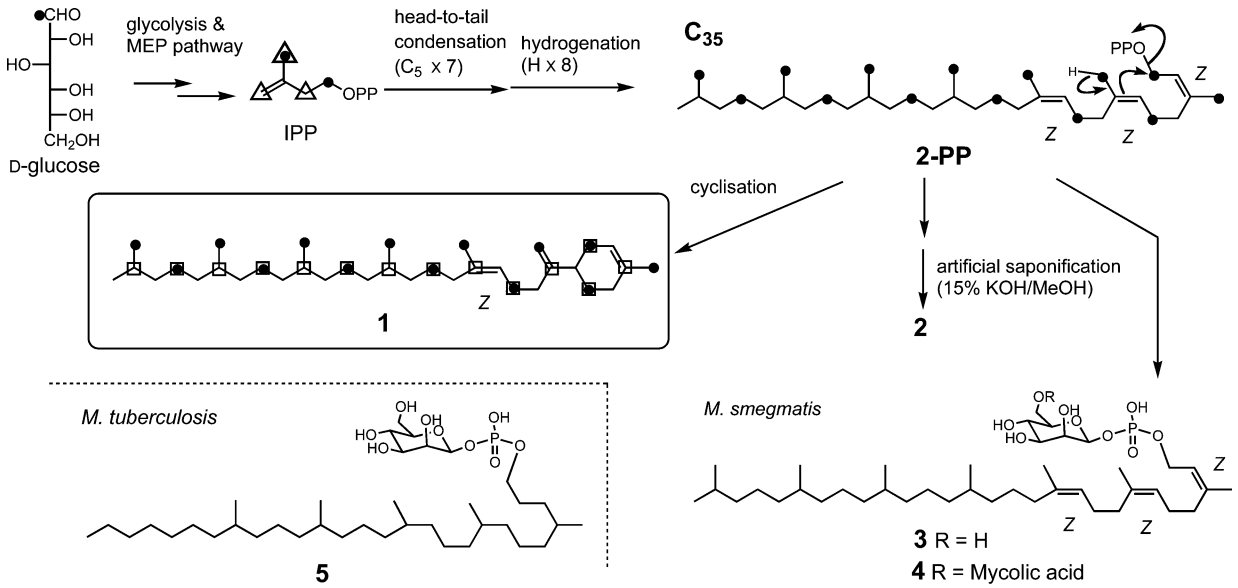

Fig. 4 The proposed pathway for the biosynthesis of C_{35}-terpenes from *Mycobacteium* species. The filled circle indicates the ^{13}C-enriched positions at which compound **1** could be produced *via* IPP by glycolysis of D-[1-^{13}C]glucose and the MEP pathway. The blank square of **1** indicates the ^{13}C-enriched positions at which **1** could be produced using methylmalonate and malonate by polyketide synthase. The blank triangle of IPP indicates the ^{13}C-enriched positions at which IPP could be produced by glycolysis of D-[1-^{13}C]glucose and the mevalonate pathway.[11] Compounds **3** and **4** were obtained from *M. smegmatis*.[4] It has been proposed that compound **5** may be biosynthesised by polyketide synthase in *M. tuberculosis*.[8]

Table 3 NMR assignments of **2** in C_6D_6

No.	δ_H (mult., *J* in Hz)	δ_C (mult.)
1	4.14 (2H, br dd)	59.03 (t)
2	5.53 (1H, t, 6.8)	125.97 (d)
3	—	139.14 (s)
4	2.17 (2H, m)	32.42 (t)[a]
5	2.24 (m)	26.79 (t)[a]
6	5.29 (1H, t, 6.8)	125.23 (d)
7	—	135.77 (s)
8	2.24 (m)	32.50 (t)[a]
9	2.32 (2H, m)	26.86 (t)[a]
10	5.39 (1H, t, 6.8)	125.23 (d)
11	—	135.77 (s)
12	2.24 (m)	32.57 (t)[a]
13	1.57 (m)	25.94 (t)
14	1.31 (m); 1.48 (m)	37.36 (t)[a]
15	1.60 (m)	33.25 (d)[a]
16	1.31 (m); 1.48 (m)	37.71 (t)[a]
17	1.41 (m); 1.54 (m)	24.96 (t)[a]
18	1.31 (m); 1.48 (m)	37.85 (t)[a]
19	1.60 (m)	33.21 (d)[a]
20	1.31 (m); 1.48 (m)	37.85 (t)[a]
21	1.41 (m); 1.54 (m)	24.96 (t)[a]
22	1.31 (m); 1.48 (m)	37.85 (t)[a]
23	1.60 (m)	33.25 (d)[a]
24	1.31 (m); 1.48 (m)	37.90 (t)[a]
25	1.41 (m); 1.54 (m)	25.25 (t)[a]
26	1.35 (m)	39.73 (t)
27	1.65 (m)	28.32 (d)
28	1.77 (3H, s)	23.48 (q)
29	1.85 (3H, s)	23.55 (q)
30	1.87 (3H, s)	23.62 (q)
31	1.07 (3H, d, 6.6)	19.93 (q)[a]
32	1.07 (3H, d, 6.6)	19.97 (q)[a]
33	1.09 (3H, d, 6.7)	20.02 (q)[a]
34	1.04 (3H, d, 6.6)	22.78 (q)[a]
35	1.04 (3H, d, 6.6)	22.88 (q)[a]

[a] The assignments may be exchangeable.

that compound **1** was produced from IPP, which is produced by glycolysis and the methylerythritol phosphate (MEP) pathway.[12] If **1** was biosynthesised by polyketide synthase or the mevalonate pathway, a different incorporation pattern should be observed as shown in Fig. 4.

The bioconversion of **2-PP** to compound **1** is anticipated. Compound **2** was diphosphorylated using the method described by Davisson *et al.*[13] The product **2-PP** was then incubated with the cell-free extracts of *M. chlorophenolicum.* As shown in Fig. 5, the amount of **1** significantly increased by the addition of **2-PP** (>7-fold) compared with that of the control (heat denaturation of the cell-free extracts or no addition of **2-PP**); thus, **2-PP** would be the biosynthetic intermediate of **1**, as illustrated in Fig. 4. To the best of our knowledge, the known C_{35}-terpenes include 8 cyclic members: plagiospirolides

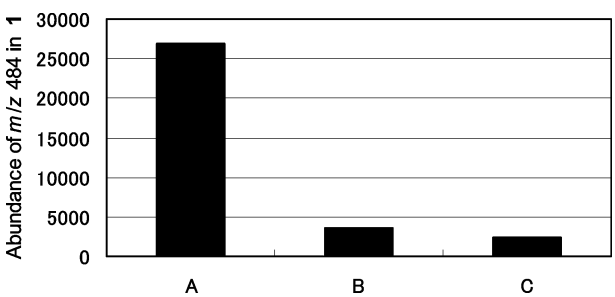

Fig. 5 The amount of **1** produced by the incubation of **2-PP** with the cell-free extracts of *M. chlorophenolicum.* The relative amount of **1** was estimated by the abundance of the molecular ion *m/z* 484. A: 1 mg of **2-PP** was incubated with the cell-free extracts. B: 1 mg of **2-PP** was incubated with heat-denatured cell-free extracts. C: no substrate was incubated with the cell-free extracts.

A–D,[14] ferrugicadinol,[15] ferrugieudesmol,[15] tetraprenylated-α-curcumene[16] and tetraprenylated-β-curcumene[16] and 5 acyclic members.[8,17–19] Since the known cyclic C_{35}-terpenes are hybrids of sesquiterpene and diterpene,[14–16] compound **1** would be the first example of a terpene that can be biosynthesised *via* the cyclisation of a linear C_{35}-isoprenoid such as **2-PP**.[20]

The reaction of head-to-tail condensation, hydrogenation, and cyclisation in the biosynthesis of **1** (Fig. 4) would be catalysed by *Z*-prenyltransferase, prenylreductase and terpene cyclase, respectively. Since a complete genome sequence of *M. vanbaalenii*, which has been categorised as a type I species in this paper, was released on the website http://genome.jgi-psf.org by the Joint Genome Institute in 2006, we attempted to search a candidate gene for the biosynthesis of compound **1** in the *M. vanbaalenii* genome using the FASTA3 program of Genome Information Broker (http://gib.genes.nig.ac.jp/). The homologs of *Z*-prenyltransferase and prenylreductase having a very low *E* value were detected (*e.g.* Mvan_3822 *vs.* undecaprenyl diphosphate synthase from *Micrococcus luteus* [$E = 2.1 \times 10^{-39}$] and Mvan_2055 vs. geranylgeranyl reductase from *Synechocystis* sp. PCC6803 [$E = 3.6 \times 10^{-9}$]), while any enzyme homologous to the known diphosphate-triggered type of mono-, sesqui- and diterpene cyclases could not be detected (*E* value > 0.01).[21] This result suggested that a novel type of terpene cyclase might exist in type I nonpathogenic *Mycobacterium* species. The C_{35}-terpene cyclases utilise **2-PP**, which contains *Z*-type prenyl moieties, while the standard cyclases utilise *E*-polyprenyl diphosphate. To the best of our knowledge, there is only one report on the use of *Z*-prenyl diphosphate as a natural substrate by terpene cyclases: Nabeta *et al.* reported that sesquiterpene cyclases producing (–)-γ-cadinene and germacrene D from *Heteroscyphus planus* utilised (2*Z*,6*E*)-farnesyl diphosphate.[22] However, the gene coding for *Z*-terpene cyclase has not yet been detected. It has been reported that not only the primary but also the tertiary structure of *Z*-prenyltransferase is considerably different from that of the *E*- prenyltransferase.[23] Therefore, the C_{35}-terpene cyclase that is situated downstream of the *Z*-prenyltransferase in the biosynthetic pathway might differ structurally from the known *E*-terpene cyclase.

Conclusion

The known cyclic C_{35}-terpenes can be regarded as the derivatives of sesqui- and diterpenes. In contrast, compound **1** cannot belong to the well-known terpene family including mono-, sesqui-, di-, sester-, tri- and tetraterpenes because it is biosynthesised from the linear C_{35}-isoprenoid. The function of the rare C_{35}-terpene **1** in the *Mycobacterium* cell may be interesting, because mycobacteria adjust the amount of compound **1** produced according to their growth phase or surrounding temperature. In addition, the diterpene of (–)-axinyssene is known to be a terpene hydrocarbon having a cyclic skeleton similar to that of compound **1**, which shows mild toxicity against acute promyelocytic leukemia HL-60 cell line.[24] Thus, the bioactivity of **1** in other organisms might also be interesting. Furthermore, the genome analysis of *M. vanbaalenii* revealed that the C_{35}-terpene cyclase producing **1** may be novel. Recently, unusual enzymes and pathways for terpene biosynthesis have been discovered. These are represented by a new class of prenyltransferases having no similarity with the known prenyltransferases,[25] a chimera diterpene synthase having both prenyltransferases and terpene cyclase activity,[26] MEP pathway[12] and a new biosynthetic pathway for menaquinone production.[27] Therefore, a careful analysis of both terpene metabolites and genome sequences may be useful in detecting a new biosynthetic enzyme.

Experimental

Analytical method

NMR spectra were recorded on a Bruker DMX 600 spectrometer at 600 MHz for proton and at 125 MHz for carbon. GC was performed on a Shimadzu GC-8A instrument equipped with a flame-ionisation detector and a DB-1 capillary column (30 m × 0.25 mm; J & W Scientific. Inc.). GC-MS was performed on an SX 102 spectrometer (JEOL) by electronic impact at 70 eV with a DB-1 capillary column or a JMS-Q1000 GC K9 (JEOL) by electronic impact at 70 eV with a ZB-5 ms capillary column (30 m × 0.25 mm; Zebron). HRMS was performed using a spectrometer equipped with a direct inlet system. The specific rotation was measured at 25 °C by using a Horiba SEPA-300 polarimeter.

Standard culture conditions

The following 12 nonpathogenic *Mycobacterium* species were employed in the study: *M. agri* JCM 6377, *M. aichiense* JCM 6376, *M. aurum* JCM 6366, *M. chlorophenolicum* JCM 7439, *M. diernhoferi* JCM 6371, *M. moriokaense* JCM 6375, *M. nonchromogenicum* JCM 6364, *M. parafortuitum* JCM 6367, *M. pulveris* JCM 6370, *M. smegmatis* JCM 6386, *M. thermoresistibile* JCM 6362 and *M. vanbaalenii* JCM 13017. All species were reciprocally shake-cultured at an optimal temperature (28 °C for *M. chlorophenolicum* and *M. vanbaalenii* and 37 °C for the others) until the stationary growth phase was achieved (*M. chlorophenolicum*, 4 d; remaining species, 3–14 d); they were shake-cultured in 3000 mL Sakaguchi flasks, each containing 1000 mL of medium containing 1% polypeptone (Nihon Seiyaku Co.), 0.5% yeast extract (Oxoid Co.), 0.5% malt extract (Difco Co.), 0.5% casamino acids (Nihon Seiyaku Co.), 0.2% glycerol (Wako Co.), 0.005% Tween 80 (Wako Co.) and 0.1% $MgSO_4 \cdot 7H_2O$ (Kanto Chemical Co.).

Preparation and GC-MS analysis of the lipid fraction obtained from the cells and of the broth filtrate

The cells and broth filtrate of each *Mycobacterium* sp. obtained from the 1 L culture medium were separated by centrifugation (6000 × *g* at 4 °C for 10 min). After lyophilisation, the cells were saponified with 15% methanolic KOH (600 mL) at 80 °C for 30 min. The nonsaponifiable lipids were extracted with *n*-hexane (600 mL × 3); the broth filtrate was extracted with EtOAc (1 L × 3). GC-MS (JEOL SX 102) was performed under the following conditions: an injection temperature of 290 °C and an oven temperature of 180–270 °C at an increment of 3 °C min^{-1}.

Isolation of compounds 1 and 2

M. chlorophenolicum was cultured in a 50 L medium at 28 °C for 4 d. The nonsaponifiable lipids from lyophilised cells (dry weight, 195 g) were extracted as described above and concentrated by

vacuum drying to obtain 5.1 g of the residue. The *n*-hexane extract was partially purified by silica gel (500 g) column chromatography with *n*-hexane and EtOAc. The fraction eluted with *n*-hexane (144 mg) was separated by subjecting it to silica gel (40 g) column chromatography with *n*-hexane in order to obtain 5.5 mg of pure compound **1**. The fraction eluted with EtOAc (4.5 g) was subjected to silica gel (300 g) column chromatography *via* gradient elution using a mixed solvent comprising *n*-hexane and EtOAc (100 : 1 → 100 : 5). The complete purification of compound **2** (5.1 mg) was achieved by reverse-phase HPLC (Shiseido, CAPCELL PAK C18) with MeOH–H_2O (100 : 5).

Instrumental data of compound 1

Colourless oil; $[\alpha]_D^{25}$ +80.5 (*c* 0.35, $CHCl_3$); NMR assignment is shown in Table 2. EIMS: *m/z* 484 (66), 387 (59), 189 (31), 161 (86), 135 (39), 121 (29), 119 (55), 107 (28), 94 (82), 81 (64), 71 (65), 67 (74) and 57 (100). HR-EIMS: *m/z* 484.5014 $[M]^+$ (calculated for $C_{35}H_{64}$, 484.5008).

Instrumental data of compound 2

Colourless oil; $[\alpha]_D^{25}$ +134.09 (*c* 0.022, $CHCl_3$); NMR assignment is shown in Table 3. EIMS: *m/z* 502 (14), 484 (28), 416 (15), 189 (23), 161 (36), 135 (29), 122 (47), 109 (52), 95 (75), 81 (100), 71 (76), 69 (80) and 57 (94). HR-EIMS: *m/z* 502.5122 $[M]^+$ (calculated for $C_{35}H_{66}O$, 502.5114).

Analysis of the yield of C_{35}-terpenes produced under different culture conditions

M. chlorophenolicum was cultured in a 1 L medium under varied culture conditions—duration (1–8 d at 28 °C) or temperature (20 °C, 28 °C or 37 °C for 8 d). The abundance of *m/z* 484 was measured by GC-MS for estimating the yield of products **1** and **2**.

Feeding experiment with D-[1-^{13}C]glucose

D-[1-^{13}C]Glucose (99% isotopic abundance) was purchased from Cambridge Isotope Laboratories, Inc. (USA). *M. chlorophenolicum* was grown at 28 °C for 3 d. A medium containing 0.5% yeast extract (Oxoid Co.), 0.5% malt extract (Difco Co.), 1% $CaCO_3$ (Wako Co.), 0.9% D-glucose (Wako Co.) and 0.1% D-[1-^{13}C]glucose was used. ^{13}C-labelled **1** was isolated using a similar method as that employed for natural compound **1** as described above, yielding 0.2 mg of pure compound **1**. The ^{13}C NMR spectra of ^{13}C-labelled **1** and natural compound **1** were compared as shown in the ESI.†

Preparation of cell-free extracts and incubation with the 2-PP substrate

M. chlorophenolicum was cultured at 28 °C for 60 h in a 6 L medium and centrifuged at 6000 × *g* at 4 °C for 20 min. After washing the cells with 50 mM of methyl-3-aminopropanesulfonic acid (MOPS; pH 7.9), the washed cells (wet weight, 41.4 g) were suspended in 100 mL of 50 mM MOPS (pH 7.9) containing 5 mM dithiothreitol, 0.3% Triton X-100, 10 mM KF and 0.25 M sucrose, followed by sonication at 4 °C for 30 min to prepare the cell-free extract; the supernatant thus obtained was used as the enzyme source. The prepared cell-free homogenate (5 mL) was added to the solution (0.3 mL) containing 1 mg substrate (**2-PP**), 10 mM $MgCl_2$ and 10 mM $MnCl_2$. **2-PP** was synthesised from the isolated compound **2** according to the method described by Davisson *et al.*[13] After incubation at 28 °C for 12 h, the reaction was quenched by adding 5 mL of MeOH. The reaction mixture was extracted with *n*-hexane (10 mL × 3) and analysed by GC-MS (JMS-Q1000 GC K9) under the following conditions: an injection temperature of 290 °C and an oven temperature of 180–270 °C at an increment of 3 °C min^{-1}. The abundance of the molecular ion *m/z* 484 was measured for estimating the relative amount of product **1**. These incubation experiments were carried out twice and similar results were obtained on both occasions.

Acknowledgements

This work was supported by a Grant-in-Aid for Young Scientists (B) from Japan Society for the Promotion of Science (no. 19780085) and a grant for the promotion of the Niigata University Research Project.

Notes and references

1 G. Devulder, M. Pe' rouse de Montclos and J. P. Flandrois, *Int. J. Syst. Evol. Microbiol.*, 2005, **55**, 293; J. P. Euze′by's List of Bacterial Names with Standing in Nomenclature (http://www.bacterio.cict.fr/m/mycobacterium.html).
2 T. Hussain, *Crit. Rev. Microbiol.*, 2007, **33**, 15.
3 S. Brandt, A. P. Zeng and W. D. Deckwer, *Biotechnol. Bioeng.*, 1999, **65**, 93; J. Moody, J. P. Freeman, P. P. Peter and C. E. Cerniglia, *Appl. Environ. Microbiol.*, 2004, **70**, 340.
4 P. R. Marri, J. P. Bannantine and G. B. Golding, *FEMS Microbiol. Rev.*, 2006, **30**, 906; Genome Information Broker (http://gib.genes.nig.ac.jp).
5 C. Nakano, T. Okamura, T. Sato, T. Dairi and T. Hoshino, *Chem. Commun.*, 2005, 1016.
6 K. Pethe, D. L. Swenson, S. Alonso, J. Anderson, C. Wang and D. G. Russell, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 13642.
7 Y. Tanaka and M. Takagi, *Biochem. J.*, 1979, **183**, 163.
8 K. Takayama, H. K. Schnoes and E. J. Semmler, *Biochim. Biophys. Acta*, 1973, **316**, 212; G. S. Besra, T. Sievert, R. E. Lee, R. A. Slayden and P. J. Brennan, *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**, 12735; B. A. Wolucka and E. Hoffmann, *Glycobiology*, 1998, **8**, 955.
9 L. Kremer, Y. Guerardel, S. S. Gurcha, C. Locht and G. S. Besra, *Microbiology*, 2002, **148**, 3145.
10 S. Toriyama, I. Yano, M. Masui, E. Kusunose, M. Kusunose and N. Akimori, *J. Biochem.*, 1980, **88**, 211; T. Baba, K. Kaneda, E. Kusunose, M. Kusunose and I. Yano, *J. Biochem.*, 1989, **106**, 81.
11 I. Matsunaga, A. Bhatt, D. C. Young, T. Cheng, S. J. Eyler, G. S. Besra, V. Briken, S. A. Porcelli, C. E. Costello, W. R. Jacobs, Jr and D. B. Moody, *J. Exp. Med.*, 2004, **200**, 1559.
12 M. Rohmer, *Pure Appl. Chem.*, 2003, **75**, 375.
13 V. J. Davisson, A. B. Woodside and C. D. Poulter, *Methods Enzymol.*, 1985, **110**, 130.
14 J. Sporle, H. Becker, M. P. Gupta, M. Veith and V. Huch, *Tetrahedron*, 1989, **45**, 5003; J. Sporle, H. Becker, N. S. Allen and M. P. Gupta, *Phytochemistry*, 1991, **30**, 3043.
15 C. Hseih, M. Tseng, Y. Shao, J. Chang, C. Kuo, C. Chang and Y. Kuo, *J. Nat. Prod.*, 2006, **69**, 1611.
16 K. Boroczky, H. Laatsch, I. Wagner-Dobler, K. Stritzke and S. Schulz, *Chem. Biodiversity*, 2006, **3**, 622.
17 M. J. L. Hoefs, J. S. Sinninghe Damste and J. W. de Leeuw, *Org. Geochem.*, 1995, **23**, 263.
18 B. P. Manquin, J. A. Morgan, J. Ju, T. Muller-Spath and D. S. Clark, *Extremophiles*, 2004, **8**, 13.
19 V. Costantino, E. Fattorusso, A. Mangoni, M. D. Rosa and A. Ianaro, *Tetrahedron*, 2000, **56**, 1393.
20 J. D. Connolly and R. A. Hill, *Dictionary of terpenoids*, Chapman & Hall, New York, 1992; S. D. Barton, K. Nakanishi and O. Meth-Cohn,

Comprehensive Natural Products Chemistry, Elsevier, Oxford, United Kingdom, 1999; P. M. Dewick, *Nat. Prod. Rep.*, 2002, **19**, 181.
21 W. R. Pearson, *Methods Mol. Biol.*, 2000, **132**, 185.
22 K. Nabeta, M. Fujita, K. Katayama and T. Takasawa, *J. Chem. Soc., Perkin Trans. 1*, 1997, 2065.
23 S. Takahashi and T. Koyama, *Chem. Rec.*, 2006, **6**, 194.
24 K. Kodama, R. Higuchi, T. Miyamoto and R. W. Van Soest, *Org. Lett.*, 2003, **5**, 169.
25 F. Pojer, E. Wemakor, B. Kammerer, H. Chen, C. T. Walsh, S. Li and L. Heide, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 2316.
26 T. Toyomasu, M. Tsukahara, A. Kaneko, R. Niida, W. Mitushashi, T. Dairi, N. Kato and T. Sassa, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 3084.
27 H. Seto, Y. Jinnai, T. Hiratsuka, M. Fukawa, K. Furihata, N. Itoh and T. Dairi, *J. Am. Chem. Soc.*, 2008, **130**, 5614.

