Biosci. Biotechnol. Biochem., **75** (1), 75–81, 2011

Characterization of the Rv3378c Gene Product, a New Diterpene Synthase for Producing Tuberculosinol and (13*R*, *S*)-Isotuberculosinol (Nosyberkol), from the *Mycobacterium tuberculosis* H37Rv Genome

Chiaki NAKANO, Takahiro OOTSUKA, Kazutoshi TAKAYAMA, Toshiaki MITSUI, Tsutomu SATO, and Tsutomu HOSHINO[†]

Department of Applied Biological Chemistry, Faculty of Agriculture and Graduate School of Science and Technology, Niigata University, 2-8050 Ikarashi, Nishi-ku, Niigata 950-2181, Japan

Received August 9, 2010; Accepted October 4, 2010; Online Publication, January 7, 2011
[doi:10.1271/bbb.100570]

The Rv3377c and Rv3378c genes from *Mycobacterium tuberculosis* are specifically found in the virulent *Mycobacterium* species, but not in the avirulent species. The Rv3378c-encoded enzyme produced tuberculosinol 2 (5(6), 13(14)-halimadiene-15-ol), 13*R*-5a and 13*S*-isotuberculosinol 5b (5(6), 14(15)-halimadiene-13-ol) as its enzymatic products from tuberculosinyl diphosphate 3, indicating that the Rv3378c enzyme catalyzed the nucleophilic addition of a water molecule after the release of a diphosphate moiety. The three enzymatic products 2, 5a, and 5b were produced irrespective of the N- and C-terminal His-tagged Rv3378c enzymes, and of the maltose-binding protein fusion enzyme; the product distribution ratio was identical between the enzymes as 1:1 for 2:5, and 1:3 for 5a:5b. The successful separation of 5a and 5b by a chiral HPLC column provided the first complete assignments of ^{1}H- and ^{13}C-NMR data for 5a and 5b. The enzymatic mechanism for producing 2, 5a, and 5b is proposed here, and the optimal catalytic conditions and kinetic parameters, in addition to the divalent metal effects, are described. Site-directed mutagenesis of Asp into Asn, targeted at the DDXXD motif, resulted in significantly decreased enzymatic activity.

Key words: *Mycobacterium tuberculosis*; diterpene; tuberculosinol; isotuberculosinol; nosyberkol

Despite over a century of research, *Mycobacterium tuberculosis* infection is still a leading cause of death worldwide.[1)] Since the completion of the genome project to elucidate the *M. tuberculosis* genome in 1998,[2)] a number of studies on the functional analyses of each open reading frame have been published. The genomes of 21 *Mycobacterium* species have been sequenced to date.[3)] It is interesting to note that the genes homologous to H37 Rv3377c and Rv3378c have only been found in other *Mycobacterium* species that cause tuberculosis.[3–5)]

We have established in our previous studies[4,5)] that the Rv3377c gene encoded a Type B diterpene cyclase to afford tuberculosinyl diphosphate **3** with a halimane scaffold from geranylgeranyl diphosphate (GGPP, **1**). We next attempted to identify the gene product that released a diphosphate moiety from **3**, because naturally occurring cyclic diterpenes are usually found in a PP-released form. These PP-releasing reactions are usually catalyzed by Type A (Class I) diterpene cyclases to yield diterpenes with more extended ring skeletons,[6)] as found in the formation of *ent*-kaurene (tetracyclic) from *ent*-copalyl diphosphate (bicyclic). We presented in 2005 and 2006 preliminary reports at research conferences,[7,8)] stating that the Rv3378c enzyme was responsible for the PP-removal reaction of **3** to produce both tuberculosinol **2** and isotuberculosinol **5**. The structure of isotuberculosinol **5** is identical to that of nosyberkol, which was isolated from the sponge *Raspailia* sp. in 2004 as a single stereoisomer.[9,10)] However, the stereochemistry at C-13 of natural nosyberkol has not been determined.[9)] In contrast, in 2009, Mann *et al.* proposed the tricyclic structure of **4** (named edaxadiene) as the single enzymatic product.[11,12)]

The inconsistency between the findings of Mann *et al.*[11)] and our research group[7,8)] prompted us to undertake a detailed characterization of the Rv3378c enzyme.[11,12)] During the preparation of this manuscript, two research groups independently have reported the total synthesis of (±)-**5** or the synthesis of the partial core of (±)-**4**;[13,14)] they concluded that the edaxadiene structure of **4** is erroneous and that our proposed structure of **5** is the true enzymatic product.[13,14)] However, our results show that the Rv3378c enzyme afforded not only **5** but also **2** as the enzymatic products of **3**.[7,8)] Chemically synthesized **5** consists of a mixture of major and minor diastereomers in a 2:1[13)] or 1:1 ratio,[4)] but the product distribution ratio of enzymatically synthesized **5a** and **5b**, including **2**, has not been reported.[11–14)] The behavior of **5a** and **5b** in a chiral GC column in comparison with those of the related diterpenes, authentic manool (13*R*-**6a**, 13*S*-**6b**) and authentic *ent*-manool (13*R*-**7a**, 13*S*-**7b**), enabled us to assign the stereochemistry of **5a** and **5b**. Moreover, the ratio of **5a** and **5b** was determined to be 1:3. The three

[†] To whom correspondence should be addressed. Fax: +81-25-262-6854; E-mail: hoshitsu@agr.niigata-u.ac.jp

Supplementary Data available on the *Biosci. Biotechnol. Biochem.* web site: supplementary method, amino acid alignment of the Rv3376 and Rv3378c proteins, SDS–PAGE of the expressed Rv3378c protein, MALDI-TOF mass data, EIMS and NMR spectra of **5a** and **5b**, and CD spectra of some mutants and others.

Abbreviations: GPP, geranyl diphosphate; FPP, farnesyl diphosphate; GGPP, geranylgeranyl diphosphate; MBP, maltose-binding protein

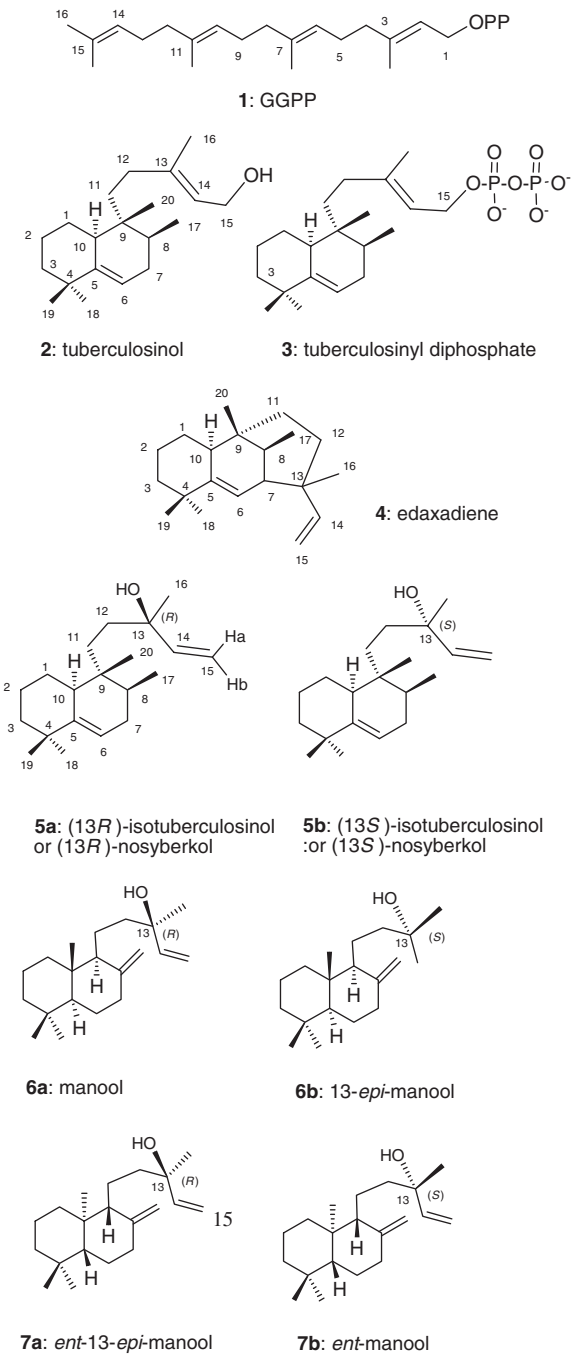


Fig. 1. Chemical Structures Discussed in the Text.

enzymatic products (**2**, **5a**, and **5b**) were confirmed not only by His-tagged proteins but also by the maltose-binding protein (MBP) fusion protein that Mann *et al.* employed.[11] The product distribution ratio of **2**, **5a**, and **5b** was the same between the three enzymes: 1:1 for **2**:**5** and 1:3 for **5a**:**5b**.

We report here the enzymatic reaction mechanism for producing **2**, **5a** and **5b**, and the complete NMR assignments of **5a** and **5b** that were separated by a chiral column for HPLC. The Rv3378c enzyme formed a dimeric structure, as revealed by a gel-filtration method. We additionally describe the detailed enzyme characterization, including kinetic parameters, metal effects on the enzymatic activity, and the results of mutagenesis experiments targeted at the DDXXD motif of the active site.

Materials and Methods

Analytical method. NMR spectra of the enzymatic products were recorded in C_6D_6 by Bruker DMX 600 and DPX 400 spectrometers; the chemical shifts are given in ppm relative to the solvent peak $\delta_H = 7.28$ and $\delta_C = 128.0$ ppm as the respective internal references for the ^{1}H- and ^{13}C-NMR spectra. GC analyses were conducted with a Shimadzu GC-8A chromatograph equipped with a flame ionization detector and a DB-1 capillary column (30 m × 0.25 mm × 0.25 μm; J&W Scientific). GC-MS spectra were measured with a Jeol SX 100 or Jeol JMS-Q1000 GC K9 instrument equipped with a ZB-5 ms capillary column (30 m × 0.25 mm × 0.25 μm; Zebron), using the EI mode operated at 70 eV. High-resolution electron ionization mass spectrometry (HR-EIMS) was performed by using a direct inlet system. The HPLC peaks were monitored at 210 or 214 nm. Specific rotation values were measured with a Horiba SEPA-300 polarimeter. Circular dichroic (CD) spectra were obtained with a Jasco J-725 instrument. TOF/MS data were measured by a Bruker Daltonics autoflex III TOF/TOF instrument with an accelerating voltage of 19.0 kV. The respective chiral HPLC and GC columns were a CHIRALPAK IC (0.46 cm × 25 cm, Daicel Chemical Industries) and a CYCLOSILB capillary column (0.32 mm × 30 m, Agilent Technologies).

Construction of vectors carrying the Rv3378c genes and site-directed mutants, gene expression, enzyme preparation, and the gel-filtration method. Genomic DNA was presented by Colorado State University (TB Research Materials and Vaccine Testing Contract, CO, USA) and used as the template for PCR. To construct the expression vectors for the N- and C-terminal His-tagged (pET16b and pET22b, Novagen) and maltose-binding protein (MBP) fusion Rv3378c enzymes (pMAL-c4X, NEB), the genes were first amplified. Each of the amplified genes was subcloned into the NdeI/BamHI, NdeI/HindIII, and EcoRI/BamHI sites of pET16b, pET22b, and pMAL-c4X respectively. Each of the pET16b and 22b constructs was incorporated into *Escherichia coli* BL21 (DE3) together with the GroE chaperone.[15] The oligonucleotides used to amplify the genes are described in the Supplemental Data (see *Biosci. Biotechnol. Biochem.* Website). The integrity of all of the constructs was confirmed by sequencing.

The transformed *E. coli*, in which the coexpression system (GroE chaperone) was employed, was cultivated at 37 °C in 1 L of an LB medium supplemented with ampicillin (50 μg/mL) and chloramphenicol (7.5 μg/mL). IPTG (0.2 mM) was added to the culture when the optical density had reached 0.6, and cultivation was continued at a lower temperature (25 °C) for 5 h. Expression of the MBP fusion protein was induced at 20 °C for 16 h. The recombinant proteins were purified in an Ni-NTA or amylose column according to the manufacturer's protocol (Qiagen or NEB). The purity of the recombinant proteins was assessed by SDS–PAGE on 10% gel.

The gel-filtration method was employed to determine the quaternary structure, using a G3000SWXL (7.8 mm × 300 mm) column (Tosoh, Japan) that had been equilibrated with a 0.1 M potassium phosphate buffer (pH 7.0) containing 0.2 M NaCl. Glutamate dehydrogenase (290 kDa), lactate dehydrogenase (142.0 kDa), enolase (67 kDa), myokinase (32.0 kDa), and cytochrome *c* (12.4 kDa) were used as molecular weight standards.

MALDI-TOF/MS analyses of the N- and C-terminal His-tagged Rv3378c proteins. Tryptic and chymotryptic digestion was performed by standard procedures. Each sample used for the MS analyses was prepared according to the published protocol.[16]

Enzyme assay of the Rv3378c protein. Enzymatic reactions were carried out under catalytically optimal conditions to estimate the kinetic parameters of the N-terminal His-tagged enzyme; the reaction mixture (2.5 mL), which contained a 50 mM Tris–HCl buffer (pH 7.0), 1 mM $MgCl_2$, 0.1% (w/v) Triton X-100, 50 μM of **3**, and 30 μg of the purified Rv3378c protein, was incubated at 45 °C for 2 h. The reaction was quenched by adding 15% KOH/MeOH, before heating at 70 °C for 15 min. The lipophilic material was extracted with hexane, and was then subjected to a GC analysis. Squalene was used as an internal standard. Enzymatic assays were conducted with various concentra-

tions of **3**, ranging from 20 μM to 100 μM, and the kinetic parameters were deduced by non-linear least-squares fitting to the Michaelis–Menten equation.

*Enzymatic syntheses of **2** and **5** by the MBP fusion Rv3378c enzyme.* The incubation conditions were essentially the same as those reported by Mann *et al.*[11] Substrate **3** (50 μg) was incubated at 30 °C with the purified protein (0.575 mg) in a solution containing 50 mM HEPES at pH 7.5, 10% glycerol, 1 mM KCl, and 10 mM $MnCl_2$. The GC-MS analysis we conducted showed two peaks for **2** and **5** that appeared in a 1:1 ratio (Supplemental Fig. S11).

*Spectroscopic data for **2**, **5a**, and **5b**.* Product **2**. The detailed NMR and MS data have been given in the previous paper[4] (see also Supplemental Fig. S5b).

Product **5a**. NMR data (Supplemental Figs. S7 and S8). ^{1}H-NMR (600 MHz, C_6D_6,) δ: 0.831 (Me-20, 3H, s), 0.935 (Me-17, d, $J = 6.7$ Hz), 1.19 (H-1, m), 1.213 (Me-19, 3H, s), 1.242 (Me-16, 3H, s), 1.264 (Me-18, 3H, s), 1.39 (H-3, m), 1.45 (H-11, m), 1.50 (H-12, m), 1.52 (H-3, m), 1.61 (H-8, m), 1.62 (H-12, m), 1.68 (H-2, 2H, m), 1.94 (H-1, m), 1.96 (H-7, 2H, m), 160 (H-11, m), 2.38 (H-10, bd, $J = 12.5$ Hz), 5.08 (Hb-15, dd, $J = 10.7$, 1.3 Hz), 5.30 (Ha-15, dd, $J = 17.2$, 1.3 Hz), 5.67 (H-6, bs), 5.90 (H-14, dd, $J = 17.2$, 10.7 Hz); ^{13}C-NMR (150 MHz, C_6D_6) δ: 15.28 (C-17, q), 16.48 (C-20, s), 22.55 (C-2, t), 27.63 (C-1, t), 28.22 (C-16, q), 29.24 (C-19, q), 30.06 (C-18, q), 30.39 (C-11, t), 32.04 (C-7, t), 33.66 (C-8, d), 35.46 (C-12, t), 36.26 (C-4, s), 36.95 (C-9, s), 40.20 (C-10, d), 41.23 (C-3, t), 72.93 (C-13, s), 111.4 (C-15, t), 116.8 (C-6, d), 145.7 (C-14, d), 146.2 (C-5, s). EIMS (%) (Supplemental Fig. S7a) *m/z*: 80 (75), 119 (48), 136 (23), 175 (13), 191 (100), 272 (0.92), 290 (M^+, 0.03). HR-EIMS *m/z*: found, 290.2603 (M^+, $C_{20}H_{34}O$ requires 290.2610). $[\alpha]_D^{25}$ +33.0 (*c* 0.05, EtOH).

Product **5b**. NMR data (Supplemental Figs. S9 and S10). ^{1}H-NMR (600 MHz, C_6D_6) δ: 0.831 (Me-20, 3H, s), 0.959 (Me-17, d, $J = 6.7$ Hz), 1.19 (H-1, m), 1.211 (Me-19, 3H, s), 1.242 (Me-16, 3H, s), 1.265 (Me-18, 3H, s), 1.38 (H-3, m), 1.40 (H-11, m), 1.52 (H-12 and H-3, 3H, m), 1.55 (H-11, m), 1.62 (H-8, m), 1.69 (H-2, 2H, m), 1.92 (H-1, m), 1.96 (H-7, 2H, m), 2.36 (H-10, bd, $J = 12.6$ Hz), 5.08 (Hb-15, dd, $J = 10.7$, 1.3 Hz), 5.30 (Ha-15, dd, $J = 17.3$, 1.3 Hz), 5.67 (H-6, bs), 5.90 (H-14, dd, $J = 17.3$, 10.7 Hz); ^{13}C-NMR (150 MHz, C_6D_6) δ: 15.24 (C-17, q), 16.47 (C-20, s), 22.57 (C-2, t), 27.64 (C-1, t), 28.26 (C-16, q), 29.25 (C-19, q), 30.06 (C-18, q), 30.38 (C-11, t), 32.03 (C-7, t), 33.68 (C-8, d), 35.47 (C-12, t), 36.26 (C-4, s), 36.95 (C-9, s), 40.20 (C-10, d), 41.23 (C-3, t), 72.94 (C-13, s), 111.4 (C-15, t), 116.8 (C-6, d), 145.7 (C-14, d), 146.1 (C-5, s). EIMS (%) (Supplemental Fig. S9a) *m/z*: 80 (78), 119 (52), 136 (21), 175 (12), 191 (100), 272 (0.77), 290 (M^+, 0.03). HR-EIMS *m/z*: found, 290.2603 (M^+, $C_{20}H_{34}O$ requires 290.2610). $[\alpha]_D^{25}$ +32.1 (*c* 0.05, EtOH).

Results and Discussion

Expression of the Rv3378c-encoded enzymes and the enzymatic products

We have revealed in our previous studies[4,5] that the Rv3377c enzyme catalyzed the conversion of **1** into **3**. Naturally occurring diterpenes are usually isolated in a PP-released form. We searched for the gene(s) homologous to terpene synthases that were involved in the H37Rv genome, but no homologous gene other than Rv3377c could be found by a BLAST search. We postulated that the flanking gene, Rv3376 or Rv3378c, may encode a dephosphorylating enzyme, because both gene products possess the characteristic DDXXD motif that is highly conserved in diterpene cyclases[6] (Supplemental Fig. S1). Both genes were amplified by PCR and cloned into the *Nde*I and *Bam*HI sites of the pET16b vector. Figure 2 shows the SDS–PAGE results for the Rv3378c protein expressed in *E. coli* BL21 (DE3). The expressed proteins almost formed inclusion bodies. However, coexpression of the Rv3378c proteins and GroE chaperones[15] in *E. coli* was effective for obtaining the active soluble enzyme. Purification in an Ni-NTA affinity column afforded a single band on SDS–PAGE (Fig. 2 and Supplemental Fig. S2). Since the crude cell-free extract obtained from *E. coli* had phosphatase activity,[4,5] we completely purified the protein to examine the enzymatic reaction. Incubating **3** with the purified N- and C-terminal His-tagged Rv3378c enzymes resulted in two peaks of **2** and **5** in an approximate 1:1 ratio as shown in the GC trace (Fig. 3A). It should be noted that **2** was not available from the mixture incubated in the absence of the purified Rv3378c enzyme, definitively indicating that **2**

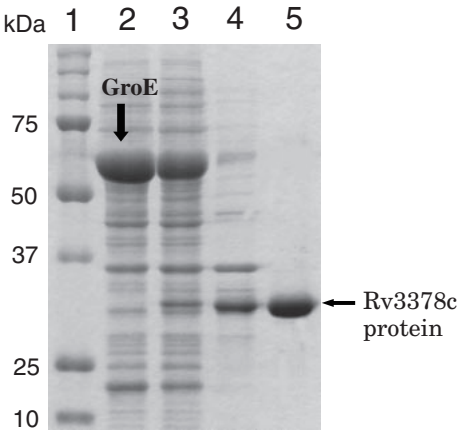


Fig. 2. SDS–PAGE of the Rv3378c-Encoded Protein Expressed by pET-16b Carrying the Rv3378c Gene in *E. coli* BL21 (DE3).

Lane 1, molecular mass marker; Lane 2, pET-16b/BL21 (DE3) + GroE (no incorporation of the Rv3378c gene); Lane 3, pET-16b-Rv3378c/BL21 (DE3) + GroE, soluble protein; Lane 4, pET-16b-Rv3378c/BL21 (DE3) + GroE, insoluble protein; Lane 5, pET-16b-Rv3378c/BL21 (DE3) + GroE, purified in an Ni-NTA column.

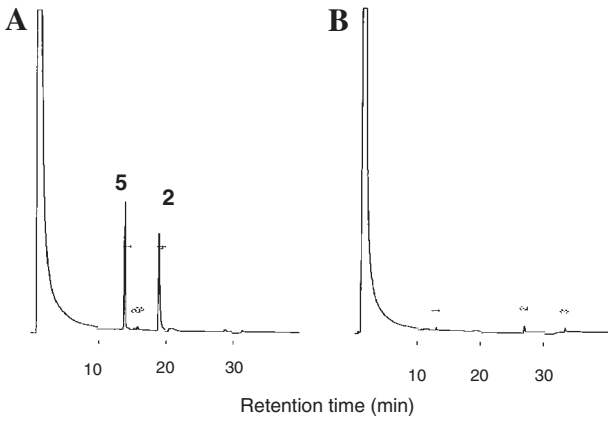


Fig. 3. GC Traces of the Hexane Extract from the Reaction Mixture of **3** with the Purified N-Terminal His-Tagged Rv3378c Enzyme (A) and without the Rv3378c Enzyme (B).

Two peaks of **2** and **5** appeared as the enzymatic products (A), but no production of **2** was apparent without this enzyme, indicating that no hydrolysis reaction of **3** to give **2** had occurred in the absence of this enzyme and under the work-up conditions. Similar GC chromatograms were obtained for the C-terminal His-tagged and MBP fusion enzymes, indicating that the ratio of **2** to **5** was approximately 1:1 for the three expressed enzymes. The GC conditions were as follows: capillary column, DB-1 (0.25 mm × 30 m); initial temperature, 170 °C; final temperature, 270 °C; rate, 3 °C/min; carrier gas pressure, 0.5 kg/cm^2.

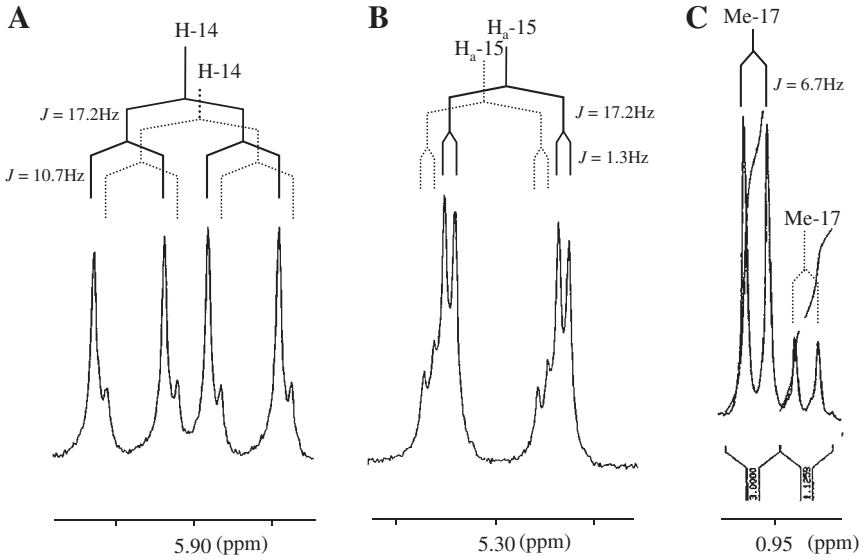


Fig. 4. ^{1}H-NMR Spectrum of Isolated **5** (a mixture of **5a** and **5b**) in Benzene d_6 (partial regions, 600 MHz). A, H-14; B, H_a-15; and C, Me-17. The large peaks were accompanied by small peaks in a 3:1 ratio.

was an enzymatic product and that **3** was not hydrolyzed to yield **2** under these incubation and work-up conditions (Fig. 3B). Furthermore, the Rv3378c enzymes yielded no product when GPP, FPP, or **1** was used as the substrate. The Rv3378c enzyme therefore accepted **3** as the true substrate. In contrast, the alternative flanking gene product, the Rv3376-encoded enzyme, which was expressed as an N-terminal His-tagged protein and completely purified in an Ni-NTA column, afforded **2** from **3**. Moreover, this enzyme hydrolyzed GPP, FPP, and GGPP to respectively yield geraniol, farnesol, and geranylgeraniol. This non-specificity for substrates suggests that the Rv3376 enzyme was a phosphatase.

*Structural determination of **2**, **5a**, and **5b** produced by the Rv3378c enzyme*

A large-scale incubation was conducted to isolate enzymatic products **2** and **5**. Diphosphate **3** (3.3 mg) was incubated at 30 °C for 20 h with the purified enzyme (37.5 mg; the N-terminal His-tagged protein) that had been obtained from a 3-L culture of transformed *E. coli*. The reaction was terminated by adding 15% KOH/MeOH and heating at 70 °C for 20 min. The lipophilic materials were extracted with hexane. SiO_2 column chromatography, eluting with hexane/EtOAc (approximately 100:0–100:5, gradient elution) afforded 1.0 mg of **2** and 1.1 mg of **5** in a 98.7% total yield. Product **2** was confirmed to be tuberculosinol by comparing the NMR and GC-MS results with our published data (Supplemental Fig. S5b).[4,5] Figure 4 depicts the ^{1}H-NMR spectrum of **5** which had been completely purified by normal-phase HPLC. Small peaks appeared around Me-17, Ha-15, and H-14; the intensity of each of these small peaks was approximately one-third that of the large peaks, and the spin-spin coupling patterns of these small peaks were the same as those of the large peaks. This finding suggests that isolated **5** consisted of two stereoisomers in a 1:3 ratio, differing from chemically synthesized **5** (2:1 or 1:1).[4,13] GC-MS in a chiral column led to the separation of two peaks, **5a** and **5b** (Fig. 5A), and each of the EIMS data for the separated peaks were

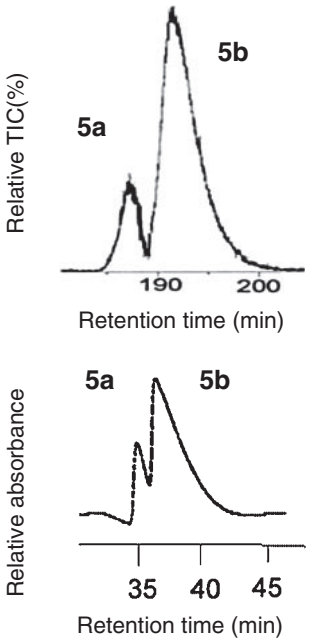
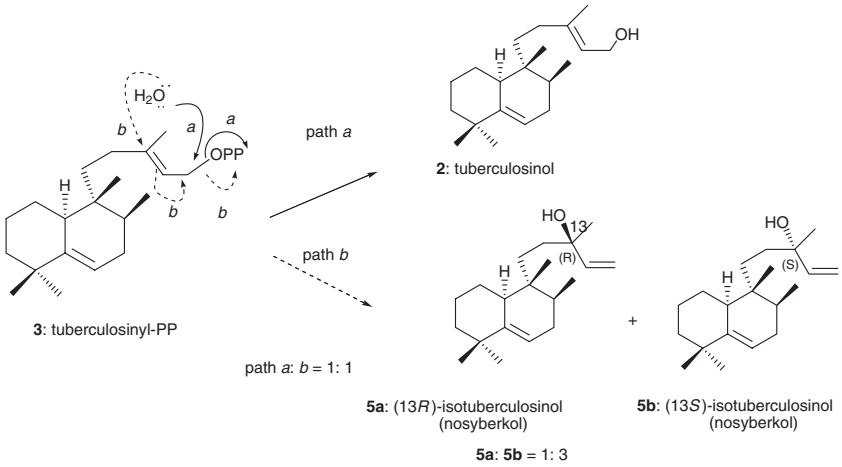


Fig. 5. Separation of Diastereomers **5a** and **5b** in GC-MS (A) and HPLC (B) Chiral Columns.

GC-MS conditions (total ions monitored): column, CYCLOSILB capillary; injection temp., 250 °C; initial temp., 130 °C; flow rate 0.01 °C/min; carrier gas, 0.5 kg/cm^2. HPLC conditions (UV monitored): column, Daicel CHIRALPAK IC; 210 nm detection; mobile phase, hexane:THF (100:0.05). The first and the second peaks were respectively assigned to (13*R*)-**5a** and (13*S*)-isotuberculosinol **5b** (see the text).

identical, including the fragment ions (Supplemental Figs. S7a and S9a). In a similar fashion, a chiral HPLC column exhibited two peaks of **5a** and **5b** in a 1:3 ratio (Fig. 5B). Pure **5a** and **5b** were subjected to NMR analyses, including DEPT, ^{1}H-^{1}H COSY, NOESY, HMQC and HMBC (Supplemental Figs. S7b–k and S9b–e). The chemical shifts of Me-17 were markedly different between **5a** and **5b** (Supplemental Fig. S6a), with a higher field resonance for **5a** than for **5b**. The chemical shifts of **5a** and **5b** were almost identical for the bicyclic core, but the chemical shifts of H-12 for **5a**



Scheme 1. Enzymatic Mechanism for the Formation of **2**, **5a**, and **5b** Catalyzed by the Rv3378c-Encoded Enzyme. The ratio of paths *a* to *b* was approximately 1:1, and the production ratio of **5a** to **5b** was 1:3.

and **5b** were significantly different: 1.50 (1H, m) and 1.62 (1H, m) for **5a** (Supplemental Fig. S7i) and 1.52 (2H, m) for **5b** (Supplemental Fig. S9e). These results indicate that **5a** and **5b** had opposing stereochemistry at C-13. The structure of edaxadiene **4** has five CH_2 groups, while **5a** and **5b** include six. Six CH_2 groups were found in the DEPT 135 pulse sequences of both **5a** and **5b** (Supplemental Figs. S7d and S9d), further supporting the enzymatic product being **5** and not **4**. ^{1}H-^{1}H correlation networks were observed between the following protons: H-10/H-1/H-2/H-3, H-10/H-6 (long-range spin-spin coupling), and H-6/H-7/H-8/Me-17. Key HMBC results revealed the following: Me-18/C-3, Me-18/C-5; Me-19/C-3, Me-19/C-5; H-7/C-6, H-7/C-5; Me-17/C-7, Me-17/C-8, Me-17/C-9; and Me-20/C-8, Me-20/C-10, Me-20/C-11. The lack of NOE between H-10 and Me-20 and clear NOE between Me-17 and Me-20 supported the relative stereochemistry of the bicyclic scaffold shown in Fig. 1. These data (Supplemental Figs. S8 and S10) definitively demonstrate the involvement of the halimane carbocycle in **5a** and **5b**. In addition to the bicyclic scaffold, a vinyl group was found in the ^{1}H-NMR spectra of **5a** and **5b**: H_b-15 (1H, δ_H 5.08, dd; $J_{Hb15\text{-}H14} = 10.7$ Hz, $J_{Ha15\text{-}Hb15} = 1.3$ Hz), H_a-15 (1H, δ_H 5.30, dd; $J_{Ha15\text{-}H14} = 17.2$ Hz, $J_{Ha15\text{-}Hb15} = 1.3$ Hz), and H-14 (1H, δ_H 5.90, dd; $J_{Ha15\text{-}H14} = 17.2$ Hz, $J_{Hb15\text{-}H14} = 10.7$ Hz). The chemical shift of C-13 was observed at 72.93 ppm (s), indicating that the tertiary alcohol group was positioned at C-13. In addition, the following HMBC cross peaks were found: Me-16/C-12, Me-16/C-13, and Me-16/C-14. The NMR analyses of enzymatic products **5a** and **5b** therefore further support the structure of **5** that had been validated by chemical syntheses.[13,14] The retention time of authentic 13*R*-manool **6a** in the chiral GC column was shorter than that of **6b**. Similarly, the first peak in the chiral GC column was *ent*-13*R*-manool **7a**, while the second peak was *ent*-13*S*-**7b** (results to be published in a subsequent paper). These findings strongly indicate that the first and second peaks could be respectively assigned to the 13*R*- and 13*S*-isomers, given that their bicyclic skeletons were identical. Thus, **5a** is assigned as the 13*R*-isomer, while **5b** is the 13*S*-isomer. The complete NMR assignments are described in the Materials and Methods section.

We also examined the enzymatic reaction with the same MBP fusion enzyme as that employed by Mann *et al.*[11] Incubation conducted under the same conditions as those reported by Mann *et al.*, including the buffer composition and temperature,[11] also resulted in the production of the three products, **2**, **5a**, and **5b**, with the same product distribution ratio as that of the N- and C-terminal His-tagged proteins: 1:1 for **2** and **5** and 1:3 for **5a** and **5b** (Supplemental Figs. S11 and S12). However, Mann *et al.* have reported that the MBP fusion enzyme afforded **4** as a single product,[11,12] while other workers who have reported the chemical synthesis of **5** have also not provided evidence for the production of **2** by the Rv3378c enzyme.[13,14]

The structures of enzymatic products **2**, **5a**, and **5b** indicate that the Rv3378c enzyme catalyzed the nucleophilic addition of a water molecule to the carbocations generated after the release of the diphosphate (Scheme 1). The OPP moiety could have been eliminated by enzyme action to allow the formation of the incipient C(15)-carbocation, which underwent nucleophilic attack by a water molecule to the C(15)-cation, yielding **2** (path *a*). The allylic rearrangement of C(13)-C(14) to C(14)-C(15) could generate the C(13)-cation, followed by a water attack, respectively yielding diastereomers **5a** (13*R*) and **5b** (13*S*) *via re-* and *si*-face attack (path *b*).

Characterization of the Rv3378c enzyme

The molecular mass of the Rv3378c protein, which was estimated from the SDS–PAGE gel mobility, was significantly lower (approximately 20%) than the calculated value, irrespective of N- and C-terminal His-tagging or the MBP fusion enzymes (Supplemental Fig. S3). However, TOF/MS analyses of the tryptic and chymotryptic digests of the N- and C-terminal His-tagged proteins (Supplemental Fig. S4 and Table S1) show that the Rv3378c proteins were actually full-length in size with 68% sequence coverage credibility. The Rv3378c protein used in the present study therefore did not cause any problems for characterization of this enzyme. A significant difference between the calculated and apparent molecular masses is frequently encountered; for example, for fission yeast proteins, the observed difference has been reported to be at the level of 10–30%.[17]

To examine whether the enzymatically active form had a quaternary structure, the assay mixture, which contained the N-terminal His-tagged protein, was subjected to gel-filtration HPLC. The plots of log MW against retention time (Fig. 6) show that the estimated molecular mass was 60 kDa, suggesting that the Rv3378c enzyme formed a dimeric structure. The eluted Rv3378c enzyme was incubated with **3**, yielding **2** and (**5a** + **5b**). The Rv3378c enzyme was therefore eluted in its active form.

The enzyme activity was estimated by measuring the total amounts of the products (**2**, **5a**, and **5b**) with a GC instrument. The effects of metal ions on the enzyme activity were also examined. Figure 7A shows the relationship between the specific activity and concentration of $MgCl_2$. No detectable amounts of **2**, **5a**, and **5b** were produced in the absence of Mg^{2+}, indicating that Mg^{2+} was essential for the enzyme activity. The enzyme activity reached a maximum at a 1 mM Mg^{2+} concentration; however, this activity did not decrease at concentrations greater than 1 mM. Additionally, the influence of the other divalent metal ions, $FeCl_2$, $CuCl_2$, $MnCl_2$, $ZnCl_2$, and $NiCl_2$, each at a concentration of 1.0 mM yielded respective activity levels of 20%, 17%, 13%, 10%, and 35%, when compared to the activity for Mg^{2+}.

The optimum temperature was determined to be 45 °C, but no activity was detected at temperatures greater than 50 °C (Fig. 7B). The enzyme was active in a pH range of 5.5–9.0, with optimal activity at pH 7.0 (Fig. 7C).

Supplementation of the detergents (Triton X-100 and Tween-80 at concentrations of 0–1.0%) and dithiothreitol (concentrations of 0–10 mM) had no effect on the enzyme activity.

Figure 7D shows that the enzyme reaction underwent no substrate inhibition. The kinetic parameters were measured under optimal catalytic conditions. The specific activity was 0.68 nmol · min^{-1} · mg^{-1}. Steady-state parameters were determined from Fig. 7D by fitting the curve to $v = V_{max}[S]/(K_m + [S])$. The values of K_m and k_{cat} were calculated to be 27.1 μM and 3.1×10^{-2} min^{-1}, respectively. The catalytic efficiency was 1.14×10^{-3} min^{-1} μM^{-1} (1.91×10 s^{-1} · M^{-1}).

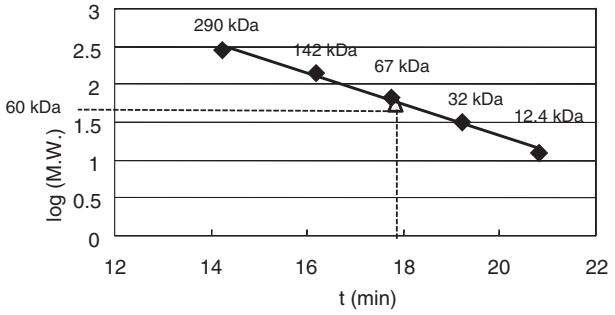


Fig. 6. Molecular Mass Calibration Curve Based on the Retention Times of Standard Proteins Obtained by Gel-Filtration HPLC.

The conditions and standard proteins are described in the Materials and Methods section. The molecular mass of the N-terminal His-tagged Rv3378c protein was determined to be approximately 60 kDa from this curve.

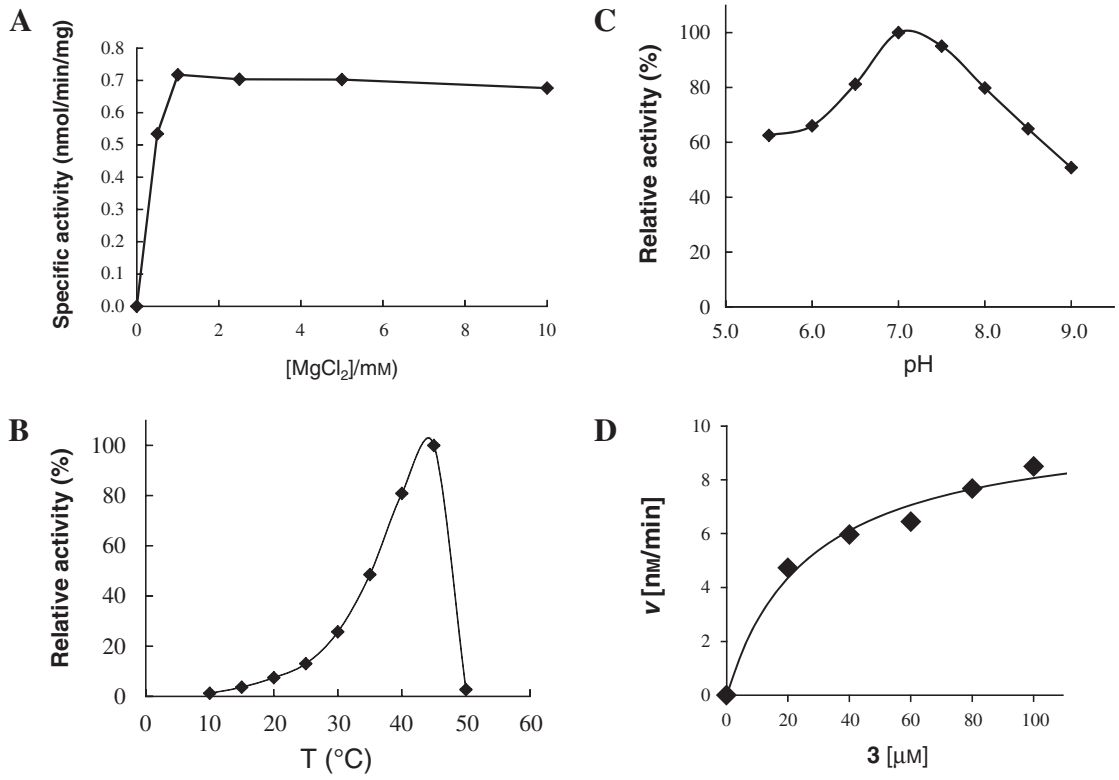


Fig. 7. Enzyme Activities of the Rv3378c Protein under Various Incubation Conditions.

A, Effect of Mg^{2+} concentration on the enzymatic activity, with $MgCl_2$ concentrations of 0.5, 1.0, 2.5, 5, and 10 mM. B, Determination of the optimum temperature. C, Determination of the optimum pH value: pH 5.5–6.0, Mes buffer (50 mM); pH 6.5, Mops buffer (50 mM); pH 7.0–9.0, Tris–HCl buffer (50 mM). D, Reaction velocity against substrate concentration. Incubation for all of the experiments described here was carried out at 45 °C for 120 min as follows: the purified enzyme (350 μg), **3** (25 μM), $MgCl_2$ (1.0 mM), and Triton X-100 (0.1% w/v) in a Tris buffer (pH 7.0, 50 mM, 2.5 mL total volume). Mg^{2+} was omitted from these standard conditions to determine the optimal Mg^{2+} concentration, its concentration being varied. The kinetic parameters were determined by using 30 μg of the purified Rv3378c enzyme (Fig. 7D).

Table 1. Kinetic Parameters for the Wild-Type and Mutants Targeted at the DDXXD Motif of the Rv3378c Enzyme

	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m ($min^{-1}/\mu M$)	Relative activity (%)
Wild-type	27.1	3.12×10^{-2}	1.15×10^{-3}	100
D81N	39.6	1.08×10^{-2}	2.73×10^{-4}	24
D82N	25.4	5.99×10^{-3}	2.36×10^{-4}	21
D85N	27.3	1.15×10^{-2}	4.21×10^{-4}	37
D81N/D82N	—	—	—	—
D81N/D85N	—	—	—	—
D82N/D85N	—	—	—	—

—, No enzyme activity

The Rv3378c enzyme had a characteristic $DDXXD_{81–85}$ motif, as already described. The three Asp residues were mutated into Asn to assess the function of this motif, Table 1 summarizing the results of the mutagenesis experiments. These site-directed mutations resulted in significantly decreased activity. The double mutants had no activity. Therefore, the Asp residues of this sequence were responsible for releasing the diphosphate group from **3** through Mg^{2+} chelation.[6] No change in the CD spectra for these mutants was apparent, indicating that the structural alteration of the protein architecture did not occur by these mutations (Supplemental Fig. S13).

We examined whether **2** and/or **5** could be detected from the cultured cells of avirulent *Mycobacterium* species; 12 species were investigated, but these compounds were not found in any of them.[18] The Rv3377c and Rv3378c genes have only been found in virulent *Mycobacterium* species.[4,5] Both enzymes are therefore likely to be a new potential target for the development of new drugs. We have recently found that **2** inhibited phagocytosis of opsonized zymosan particles by human macrophage-like cells, as well as **5**. Furthermore, a strong synergistic effect for biological activity was observed in the case of the coexistence of **2** and **5**.[7] Mann *et al.* and other researchers have never referred to the production of **2** as an enzymatic product,[11–13] but a significant amount of **2** (**2**:**5** = 1:1) was detected *in vitro* in our experiments. Further studies are necessary to demonstrate that **2** and **5** are produced *in vivo* or to detect them as natural products in *M. tuberculosis* cells.

Acknowledgments

This work was supported by grant-aid given to T. H. for scientific research from the Ministry of Education, Culture, Sports, Science and Technology of Japan (no. 18380001).

References and Notes

1) Andries K, Verhasselt P, Guillemont J, Göhlmann HWH, Neefs J-M, Winkler H, Gestel JV, Timmerman P, Zhu M, Lee E, Williams P, de Chaffoy D, Huitric E, Hoffner S, Cambau E, Truffot-Pernot C, Lounis N, and Jarlier V, *Science*, **307**, 223–227 (2005).
2) Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry III CE, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Krogh A, McLean J, Moule S, Murphy L, Oliver K, Osborne J, Quail MA, Rajandream M-A, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S, and Barrell BG, *Nature*, **393**, 537–544 (1998).
3) DNA data bank of Japan http://www.ddbj.nig.ac.jp/
4) Nakano C, Okamura T, Sato T, Dairi T, and Hoshino T, *Chem. Commun.*, 1016–1018 (2005).
5) Nakano C and Hoshino T, *ChemBioChem*, **10**, 2060–2071 (2009).
6) Christianson DW, *Chem. Rev.*, **106**, 3412–3442 (2006).
7) Nakano C, Okamura T, Sato T, Hara T, Dairi T, Toyomasu T, Sassa T, and Hoshino T, Symposium paper, 48th Tennen Yuukikagoubutu Toronkai (Symposium on the Chemistry of Natural Products), Sendai, October, 2006, Japan, pp. 19–24 (in Japanese): http://ci.nii.ac.jp/naid/110006682624/en
8) Nakano C, Sato T, and Hoshino T, Abstract, 49th Koryo, Terupen oyobi Seiyu Kagaku ni kansuru Toronkai (Symposium on the Chemistry of Tepenes, Essential Oils, and Aromatics), Fukui, November, 2005, pp. 247–249 (in Japanese).
9) Rudi A, Aknin M, Gaydon E, and Kashman Y, *J. Nat. Prod.*, **67**, 1932–1935 (2004).
10) Either **5a** or **5b** corresponds to nosyberkol, and the alternative one is a new compound. The compound names of both nosyberkol and isotuberculosinol are cited in the ref. 13. In this paper, we prefer to use isotuberculosinol rather than nosyberkol in the text, because we have used the name of isotuberculosinol since the first finding of **5** by us in 2004.
11) Mann FM, Xu M, Chen X, Fulton DB, Russell DG, and Peters RJ, *J. Am. Chem. Soc.*, **131**, 17526–17527 (2009).
12) Mann FM, Prisic S, Hu H, Xu M, Coates RM, and Peters RJ, *J. Biol. Chem.*, **284**, 23574–23579 (2009).
13) Maugel N, Mann FM, Hillwig ML, Peters RJ, and Snider BB, *Org. Lett.*, **12**, 2626–2629 (2010).
14) Spangler JE, Carson CA, and Sorensen EJ, *Chem. Sci.*, **1**, 202–205 (2010).
15) Yasukawa T, Kanei-Ishii C, Maekawa T, Fujimoto J, Yamamoto T, and Ishii S, *J. Biol. Chem.*, **270**, 25328–25331 (1995).
16) John JPP, Anrather D, Pollak A, and Lubec G, *Proteins*, **64**, 543–551 (2006).
17) Shirai A, Matsuyama A, Yashiroda Y, Hashimoto A, Kawamura Y, Arai R, Komatsu Y, Horinouchi S, and Yoshida M, *J. Biol. Chem.*, **283**, 10745–10752 (2008).
18) Sato T, Kigawa A, Takagi R, Adachi T, and Hoshino T, *Org. Biomol. Chem.*, **6**, 3788–3794 (2008).