

Eubacterial Diterpene Cyclase Genes Essential for Production of the Isoprenoid Antibiotic Terpentecin

TOHRU DAIRI,^{1*} YOSHIMITSU HAMANO,¹ TOMOHISA KUZUYAMA,²
NOBUYA ITOH,¹ KAZUO FURIHATA,³ AND HARUO SETO²

*Biotechnology Research Center, Toyama Prefectural University, Toyama 939-0398,¹ and
Institute of Molecular and Cellular Biosciences² and Division of Agriculture
and Agricultural Life Science,³ The University of Tokyo,
Bunkyo-ku, Tokyo 113-0032, Japan*

Received 26 March 2001/Accepted 25 July 2001

A gene cluster containing the mevalonate pathway genes (open reading frame 2 [ORF2] to ORF7) for the formation of isopentenyl diphosphate and a geranylgeranyl diphosphate (GGDP) synthase gene (ORF1) had previously been cloned from *Streptomyces griseolosporeus* strain MF730-N6, a diterpenoid antibiotic, terpentecin (TP) producer (Y. Hamano, T. Dairi, M. Yamamoto, T. Kawasaki, K. Kaneda, T. Kuzuyama, N. Itoh, and H. Seto, *Biosci. Biotech. Biochem.* 65:1627–1635, 2001). Sequence analysis in the upstream region of the cluster revealed seven new ORFs, ORF8 to ORF14, which were suggested to encode TP biosynthetic genes. We constructed two mutants, in which ORF11 and ORF12, which encode a protein showing similarities to eukaryotic diterpene cyclases (DCs) and a eubacterial pentalenene synthase, respectively, were inactivated by gene disruptions. The mutants produced no TP, confirming that these cyclase genes are essential for the production of TP. The two cyclase genes were also expressed in *Streptomyces lividans* together with the GGDP synthase gene under the control of the *ermE constitutive promoter. The transformant produced a novel cyclic diterpenoid, *ent*-clerod-3,13(16),14-triene (terpentetriene), which has the same basic skeleton as TP. The two enzymes, each of which was overproduced in *Escherichia coli* and purified to homogeneity, converted GGDP into terpentetriene. To the best of our knowledge, this is the first report of a eubacterial DC.**

Isoprenoids are the largest single family of compounds found in nature, with over 22,000 known examples (12), and can be classified into several groups based on the number of C₅ units derived from isopentenyl diphosphate (IPP), such as monoterpenes (C₁₀), sesquiterpenes (C₁₅), diterpenes (C₂₀), and triterpenes (C₃₀), etc. (12). These compounds are biosynthesized from the corresponding prenyl diphosphate. Geranyl diphosphate gives rise to monoterpenes, farnesyl diphosphate gives rise to sesquiterpenes, and geranylgeranyl diphosphate (GGDP) gives rise to diterpenes (9, 29) (Fig. 1). In many cases, these prenyl diphosphates undergo a range of cyclizations to produce the parent skeletons of each class, followed by a variety of modifications to give many thousands of different isoprenoid metabolites (9, 29).

A variety of isoprenoid synthases (cyclases), most of which are from plants and fungi, have been purified and extensively studied (12). As for the genes encoding isoprenoid cyclases, more than 30 eukaryotic isoprenoid synthases have been cloned as cDNAs (9, 29). On the other hand, there have been few reports about eubacterial isoprenoid cyclases and genes because the vast majority of isoprenoids are produced by eukaryotes. Pentalenene synthase, a sesquiterpene cyclase from a *Streptomyces* strain (10), and squalene-hopene cyclases, triterpene cyclases from bacteria, are the only examples (22, 32, 34, 35, 42, 46). There are no reports, to the best of our knowledge, about eubacterial monoterpene cyclases and diterpene cyclases (DCs).

We have been studying the biosynthesis of isoprenoid antibiotics produced by actinomycetes. Although actinomycetes produce approximately 70% of all natural compounds, a very limited number of isoprenoid compounds are known to be produced by them (30). The gene cluster containing the mevalonate pathway genes used to synthesize IPP had previously been cloned from *Streptomyces griseolosporeus* strain MF730-N6, a diterpene antibiotic terpentecin (TP) producer (15). The GGDP synthase gene encoding the enzyme catalyzing the formation of GGDP, which is the direct precursor of TP, was also identified in the upstream region of the mevalonate pathway gene cluster (15) (open reading frame 1 [ORF1] in Fig. 2). Considering that the biosynthetic genes for almost all of the antibiotics produced by actinomycetes are known to be clustered in the genomic DNA region (11, 27), the TP biosynthetic genes are also expected to exist in the flanking region of the GGDP synthase gene.

In this paper, we describe the cloning and sequencing analysis of seven genes that were newly found in the flanking regions of the mevalonate pathway gene cluster of strain MF730-N6. In particular, we focused on the two cyclase genes, which encode proteins showing similarities to eukaryotic DCs (ORF11) and a eubacterial pentalenene synthase (ORF12). Mutant constructions to examine if these genes would encode the TP biosynthetic enzymes and heterologous expression of the cyclase genes in *Streptomyces lividans* and *Escherichia coli* are mainly described.

MATERIALS AND METHODS

Chemicals. [α -³²P]dCTP was obtained from Amersham. GGDP was purchased from Sigma. The other chemicals used were all analytical grade.

Bacterial strains. *S. griseolosporeus* strain MF730-N6 (a TP producer), which was formerly classified in the genus *Kitasatosporia* (41), was used for the cloning

* Corresponding author. Mailing address: Biotechnology Research Center, Toyama Prefectural University, Toyama 939-0398, Japan. Phone: 81-766-56-7500. Fax: 81-766-56-2498. E-mail: dairi@pu-toyama.ac.jp.

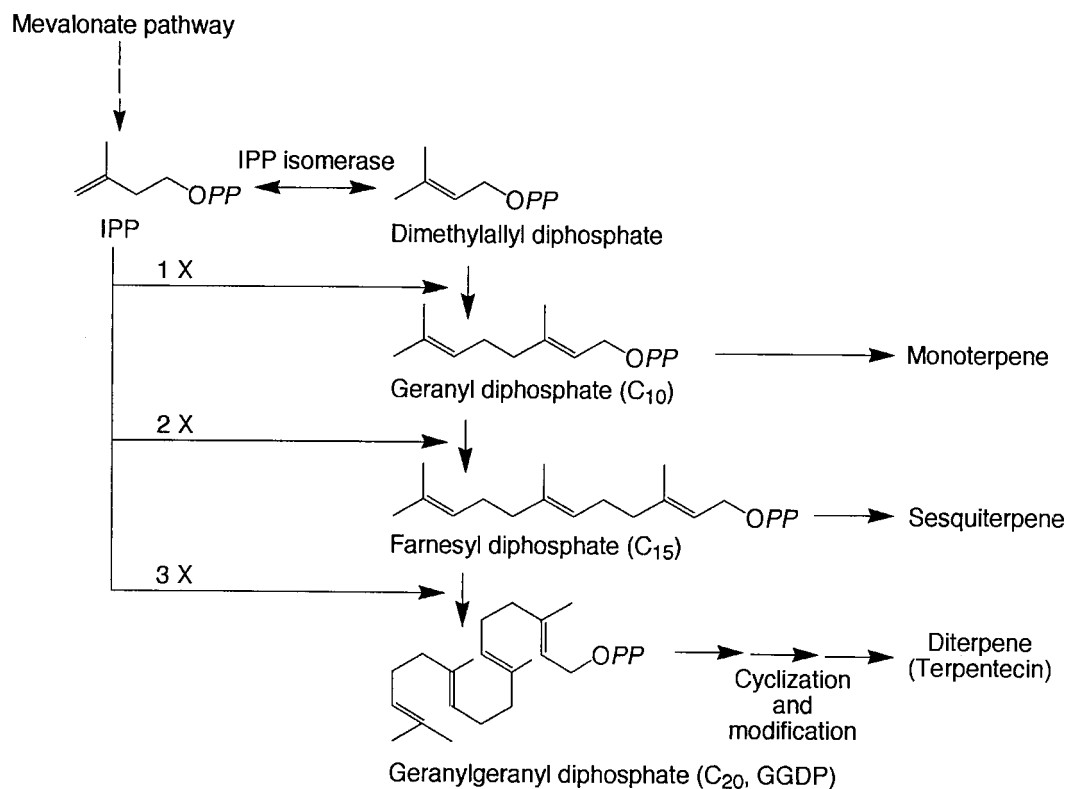


FIG. 1. Formation of different isoprenoid metabolites. *OPP*, diphosphate.

experiment. The media and growth conditions used for strain MF730-N6 were those described by Tamamura et al. (41). *S. lividans* TK23 (17) and pWHM860 (a gift from C. R. Hutchinson, Kosan Biosciences Inc., Hayward, Calif.), a derivative of plasmid pWHM3 (44), were used for heterologous expression of the TP biosynthetic genes. *E. coli* M15/pREP4 and plasmid pQE30 (Qiagen) were used for the expression of His-tagged proteins. *E. coli* JM110 (*rpsL thr leu thi lacY galK ara tonA tsx dam dcm supE44/F' [traD proAB lacI^ΔM15]*) (Toyobo, Osaka, Japan) and plasmids pUC118 and pUC119 were used for sequencing analysis. If necessary, ampicillin was added to the medium to a final concentration of 100 μ g/ml.

DNA isolation and manipulation. Plasmids from *E. coli* were prepared by using a Qiagen Plasmid Kit (QIAGEN, Inc., Chatsworth, Calif.). All restriction enzymes, T4 ligase, and calf intestinal alkaline phosphatase were obtained from Toyobo and used in accordance with the manufacturer's protocols. Transformation of *E. coli* with plasmid DNA by electroporation was performed under standard conditions by using a BTX ECM 600 electroporation system (Biotechnologies and Experimental Research, Inc., San Diego, Calif.). The transformation protocols used for *S. lividans* and *S. griseoloporeus* were essentially the same as those described by Hopwood et al. (17) and Dairi et al. (13), respectively. Other general procedures were performed as described by Maniatis et al. (26).

Sequence analysis. A cosmid clone, pSG003 (15), which carried the mevalonate pathway gene cluster and its flanking region, was used for sequencing analysis. Several restriction enzyme-digested fragments were subcloned to pUC118 or pUC119. After construction of a series of plasmids, sequencing was done by the dideoxy-chain termination method of Sanger et al. (37) with an automated DNA sequencer (Li-cor model 4000L). Finally, the nucleotide sequence of an 8.5-kb *Bam*HI fragment was determined.

Disruption of ORF11 and ORF12. A 5.9-kb *Sac*I fragment carrying the ORF11 and ORF12 genes was subcloned into pGEM5Zf(+) (Promega, Madison, Wis.) to give pCYC1. The plasmid was digested with *Sma*I, which existed in the inserted DNA, and the site was changed into a *Hind*III site by a linker to introduce a frameshift mutation into the ORF11 gene. The plasmid was digested with *Sac*I, and the resultant fragment was subcloned into the same sites of pEN101 (encodes thiostrepton resistance) (13) to give pCYC1FS. Plasmid pCYC1 was digested with *Bsr*EII, treated with the Klenow fragment of DNA polymerase I to fill in the sticky ends of this site, and self-ligated, in which process ORF12 was inactivated by a frameshift mutation. The plasmid was digested with

*Sac*I, and the resultant fragment was subcloned into the same sites of pEN101 to give pCYC2FS. Plasmids pCYC1FS and pCYC2FS were independently introduced into strain MF730-N6, and thiostrepton-resistant colonies were selected. After protoplasting and regeneration of the transformant, thiostrepton-sensitive colonies were collected. Among them, a frameshifted mutation in ORF11 or ORF12 was confirmed by Southern blot analysis.

Complementation of the cyclase gene-inactivated mutant. To complement the ORF11- and ORF12-inactivated mutants, two plasmids, pEN-ORF11 and pEN-ORF12, which carried the ORF11 gene and the ORF12 gene, respectively, were constructed. To obtain the entire ORF11 gene without the excessive flanking region, PCR amplification was carried out under standard conditions. The 5' and 3' primers with an additional restriction site (underlined) had the respective sequences 5'-ATCAGCCGACGCTTCTAGACGCTCCCCGTC-3' (ORF11-N) and 5'-AAGCCGCGAGCAAGCTTTCGGCATCC-3' (ORF11-C). After sequence confirmation, the *Xba*I-*Hind*III fragment was inserted into the same sites of pWHM860 to give pWHM-ORF11, in which the ORF11 gene was expressed under the control of the *ermE** promoter. The plasmid was digested with *Hind*III, and the site was changed into a *Bgl*II site by a linker. The plasmid, which had two *Bgl*II sites (one is the site newly created by the linker, and the other existed in the upstream region of the *ermE** promoter), was digested with *Bgl*II, and the resultant fragment was subcloned into the same site of pEN101 (13) to give pEN-ORF11. Plasmid pEN-ORF12 was constructed by the same method as pEN-ORF11 by using the primers 5'-ATGGGCAAGGACCTCTAGACGCCTTCCGG-3' (ORF12-N) and 5'-GTAACGGACGGCAAGCTTCTCGGATCGAGC-3' (ORF12-C). Plasmids pEN-ORF11 and pEN-ORF12 were introduced into the ORF11-inactivated mutant and the ORF12-inactivated mutant, respectively, and TP productivity was examined.

Expression of the cyclase genes in *S. lividans*. A 1.2-kb *Bam*HI-*Hinc*II fragment carrying the GGDP synthase gene was inserted into the *Bam*HI-*Sma*I sites of pGEM7Z (Promega). The *Bam*HI-*Xba*I fragment (the latter site existed in multilinker sites) was subcloned into the same site of pWHM860 to construct pWHM-GGDP. The entire cyclase genes (ORF11 and ORF12) was amplified by PCR with the ORF11-N primer and the ORF12-C primer. After sequence confirmation, the *Xba*I-*Hind*III fragment was inserted into the same sites of pWHM-GGDP to give pWHM-TER1, in which the GGDP synthase gene and the two cyclase genes were expressed under the control of the *ermE** promoter.

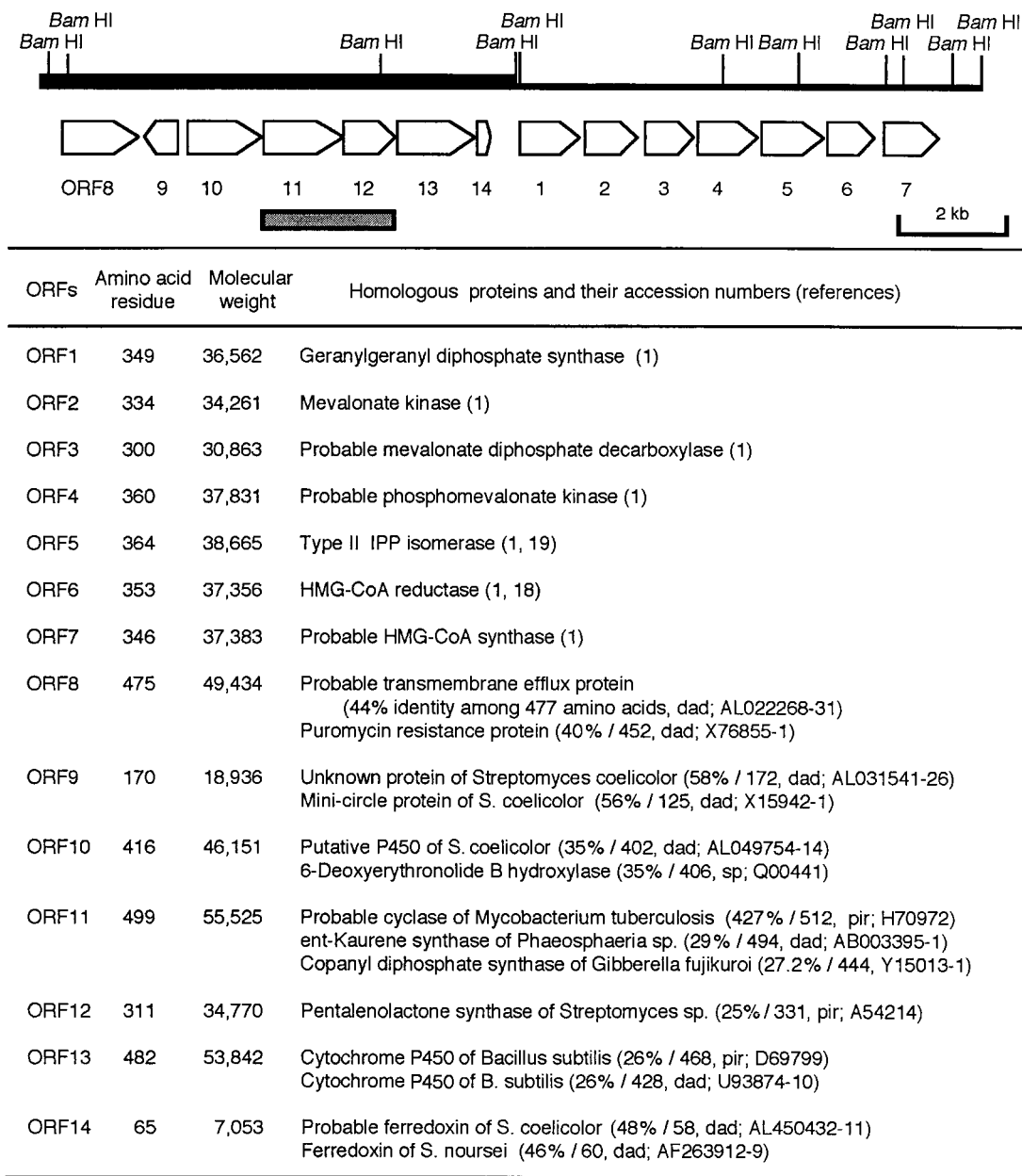


FIG. 2. ORFs deduced by sequencing analysis. The thick and thin bars represent the DNA fragments used for sequencing analysis in this study and in the previous study, respectively. A gray box shows the DNA fragment used for Northern blot analysis. The arrows indicate the extents and directions of the ORFs. CoA, coenzyme A; HMG, 3-hydroxy-3-methylglutaryl; dad, DNA Data Bank of Japan; sp, Swissprot; pir, Protein Identification Resource.

Isolation of a compound produced by an *S. lividans* transformant. *S. lividans* harboring pWHM-TER1 was grown in many 300-ml Erlenmeyer flasks containing SK-no. 2 medium (30 ml) (14) and thiostrepton (10 µg/ml). Fermentation was carried out for 7 days at 30°C with agitation (200 rpm). The culture broth (15 liters) was centrifuged, and the precipitated mycelial cake was suspended in 15 liters of acetone. After vigorous shaking, the suspension was filtered and the acetone filtrate was concentrated to dryness in vacuo. The dried material was dissolved in 200 ml of chloroform and water (1:1). After centrifugation to separate the emulsion, the organic layer was recovered. The aqueous layer was extracted twice with chloroform. The combined organic layer was evaporated to dryness under reduced pressure. The dried material was dissolved in a small volume of chloroform and acetone (1:1) and subjected to chromatography on thin-layer chromatography plates (silica gel 60F₂₅₄; Merck) developed in hexane. The material (*R_f* 0.47 to 0.68) was subjected to extraction with acetone, followed

by filtration and concentrated to dryness. The dried material was dissolved in a small volume of acetonitrile and then fractionated by preparative high-performance liquid chromatography (HPLC; Merck Mightisil RP-8 column, 250 by 20 mm; mobile phase, 100% acetonitrile; flow rate, 5 ml/min; detection wavelength, 210 nm).

NMR spectroscopy. ¹H and ¹³C nuclear magnetic resonance (NMR) spectra were recorded at 500 and 125 MHz, respectively, by using a JEOL A500 spectrometer. One- and two-dimensional experiments (correlation spectroscopy [COSY], nuclear Overhauser and exchange spectroscopy [NOESY], heteronuclear single quantum coherence [HSQC], heteronuclear multi quantum coherence [HMQC], and a difference decoupling experiment) were performed at ambient temperature. The sample (4.4 mg) was dissolved in 0.35 ml of CDCl₃.

Mass spectroscopy. High-resolution electron ionization (EI) mass spectra were obtained by using a JEOL HX-110 mass spectrometer.

TABLE 1. ^1H - and ^{13}C -NMR data of terpenetriene and *ent*-clerod-3,13(16),14-triene

Group no. ^a , group	Terpenetriene		<i>ent</i> -clerod- 3,13(16), 14-triene
	C-13	H-1 ^b	C-13
1, CH ₂	17.85	ca. 1.42 (m), ca. 1.64 (m)	18.3
2, CH ₂	26.85	ca. 1.97 (m), ca. 2.04 (m)	26.9
3, CH=	120.24	5.143 (br. s)	120.4
4, C=	144.55		144.5
5, C	38.31		38.2
6, CH ₂	30.20	ca. 1.435 (13, 13, 6), ca. 1.46 (m)	38.8
7, CH ₂	25.59	1.299 (14, 3, 3, 3), ca. 1.95 (13, 4, 4)	27.5
8, CH	35.16	ca. 1.65 (m)	38.3
9, C	37.69		38.8
10, CH	45.13	ca. 1.43 (m)	46.5
11, CH ₂	38.31	1.179 (13, 13, 5), ca. 1.63 (m)	37.5
12, CH ₂	24.57	2.136 (14, 14, 5), 2.185 (14, 14, 5)	24.7
13, C=	147.75		147.6
14, CH=	139.24	6.372 (18, 11)	139.0
15, CH ₂ =	113.02	5.242 (18, br), 5.043 (11, br)	112.9
16, CH ₂ =	115.24	4.99 (s)	115.5
17, CH ₃	14.88	0.96 (7)	16.1
18, CH ₃	20.63	1.051 (s)	20.0
19, CH ₃	18.01	1.582 (br. s)	16.0
20, CH ₃	20.38	0.968 (s)	18.4

^a The numbering corresponds to that in Fig. 3 and 4.

^b Values in parentheses represent ^1H - ^{13}C coupling constants in hertz.

Structure determination. The molecular formula of terpenetriene was determined to be C₂₀H₃₂ by high-resolution EI mass spectral data (observed M_r , 272.2547; calculated M_r , 272.2504), which indicated the degree of unsaturation of terpenetriene to be 5. The ^{13}C -NMR spectral data of terpenetriene (Table 1) showed the presence of four CH₃, six CH₂, two CH, two C, two CH₂=, two CH=, and two C= groups. These data, together with its molecular formula, indicated that the hydrocarbon terpenetriene is a diterpenoid with two rings. The ^1H - and ^{13}C -NMR spectral data obtained by phase-sensitive COSY, HMQC, and heteronuclear multiple-bond connectivity (HMBC) experiments (Fig. 3 and Table 1) disclosed partial structures compatible with the planar structure shown in Fig. 3, although conclusive evidence was not available due to the overlapping of proton signals at around 1.44 and 1.65 δ .

In view of the carbon skeleton of the diterpenoid TP and its biosynthetic formation mechanism, we assumed that terpenetriene might be a diterpenoid possessing the same carbon skeleton as TP. Thus, we compared the ^{13}C -NMR spectral data of terpenetriene with those of *ent*-clerod-3,13(16),14-triene (31), which was found by a database search to have the same planar structure. As summarized in Table 1, the carbon chemical shifts of ring A (C-1 to -5), its appended methyls (C-18 and -19), and the side chain (C-12 to -16) were completely identical between terpenetriene and *ent*-clerod-3,13(16),14-triene. The chemical shifts of the remaining carbons of ring B (C-6 to -10) and its appended methyls (C-17 and -20), however, are different, suggesting that they are stereoisomers of each other at C-8 and/or C-9. A strong nuclear Overhauser effect (NOE) observed between the two methyls C-18 and -20 but not between the methyl C-18 and methylene C-11 of terpenetriene confirmed that these two methyl groups are both in an axial orientation. Thus, the only difference between terpenetriene and *ent*-clerod-3,13(16),14-triene was the stereochemistry at C-8, with the methyl C-17 of terpenetriene being in the axial orientation, as shown in Fig. 4. This conclusion was corroborated by a large upfield shift of C-6 caused by a strong γ effect of the axial methyl group (C-17). Another conclusive evidence was obtained by a difference-decoupled spectrum prepared by irradiation of C-17 methyl protons. H-8 was observed as a triplet with $J = 3.5$ Hz, showing that this proton was in an equatorial orientation. Thus, the structure of terpenetriene was determined to be a stereoisomer at C-8 of *ent*-clerod-3,13(16),14-triene, as shown in Fig. 4. The identical stereostructures of terpenetriene and TP strongly suggest that terpenetriene is a biosynthetic intermediate of TP and its absolute stereochemistry is reasonably assumed as shown in Fig. 4 based on the established absolute stereochemistry of TP (43).

Overproduction of the cyclases in *E. coli*. Two sets of primers were used to amplify ORF11 and ORF12. The primers were designed as follows (restriction sites are underlined): ORF11, 5'-CGGGATCCAAGGACCGCGCTGCCGACCCG-3' (forward) and 5'-CGCGGATCCTCAGTACCTGCCACGACGG

C-3' (reverse); ORF12, 5'-CGGGGTACCCCGACGCGATCGAGTTCGA G-3' (forward) and 5'-CGGGGTACCTCAGCGGTAGCGGTTCGTCTC-3' (reverse). PCR was carried out under standard conditions. After sequence confirmation, a 1.5-kb *Bam*HI fragment (ORF11) and a 940-bp *Kpn*I fragment (ORF12) were each inserted into the same site of pQE30. Plasmids pQE30-ORF11 and pQE30-ORF12, in which recombinant proteins were expressed as N-terminal six-His-tagged fusion proteins, were selected.

E. coli M15(pREP4) harboring pQE30-ORF11 or pQE30-ORF12 was grown at 37°C in Luria broth with appropriate antibiotics. Expression of the recombinant protein was induced by adding 0.1 mM isopropyl- β -D-thiogalactopyranoside (IPTG) when the optical density at 600 nm reached about 0.8. Cultivation was continued for additional 12 h at 18°C. Purification of His-tagged recombinant proteins was done in accordance with the manufacturer's (Qiagen) protocols. Purified proteins were analyzed by sodium dodecyl sulfate (SDS)-10% polyacrylamide gel electrophoresis (PAGE).

Enzymatic synthesis of terpenetriene in vitro. The reaction mixture contained 50 mM potassium phosphate (pH 6.5), 10 mM MgCl₂, 5 mM 2-mercaptoethanol, 0.1% Tween 80, 40 μM GGDP, and 150 μg each of the purified ORF11 and ORF12 proteins per ml. The reaction was carried out at 30°C for 3 h. The product was extracted with chloroform and analyzed by reversed-phase HPLC. The analytical conditions were essentially the same as those described above, except for the column (Mightisil RP-18 GP, 250 by 4.6 mm) and the flow rate (1 ml/min).

Hybridization. Northern blot hybridization with a ^{32}P -labeled DNA fragment made by nick translation (5×10^8 cpm/mg) was done as described previously (14). The filter was washed twice with a buffer containing 0.3 \times SSC (1 \times SSC is 0.15 M NaCl plus 0.015 sodium citrate) and 0.2% SDS for 30 min at 68°C. Total RNA was isolated from *S. griseolosporeus* strain MF730-N6 grown at 30°C for 2 days in production medium as described previously (14).

Nucleotide sequence accession number. The DNA sequence of an 8.5-kb *Bam*HI fragment containing the upstream region of the GGDP synthase has been deposited in the DDBJ, EMBL, and GenBank databases under accession number AB048795.

RESULTS

Nucleotide sequences of the flanking region of the mevalonate pathway gene cluster. We have recently cloned the gene cluster containing the mevalonate pathway genes for the formation of IPP and the gene encoding the enzyme catalyzing the formation of GGDP, which is the direct precursor of TP (15). In many cases, antibiotic biosynthetic genes cloned from

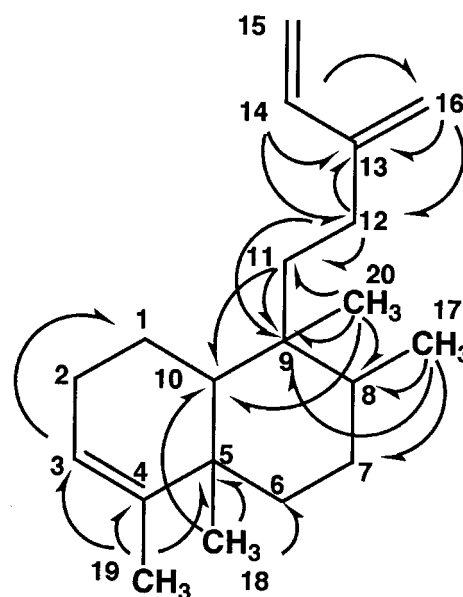


FIG. 3. Structure of terpenetriene revealed by NMR spectral analysis. Arrows indicate ^1H - ^{13}C long-range couplings observed by HMBC experiments.

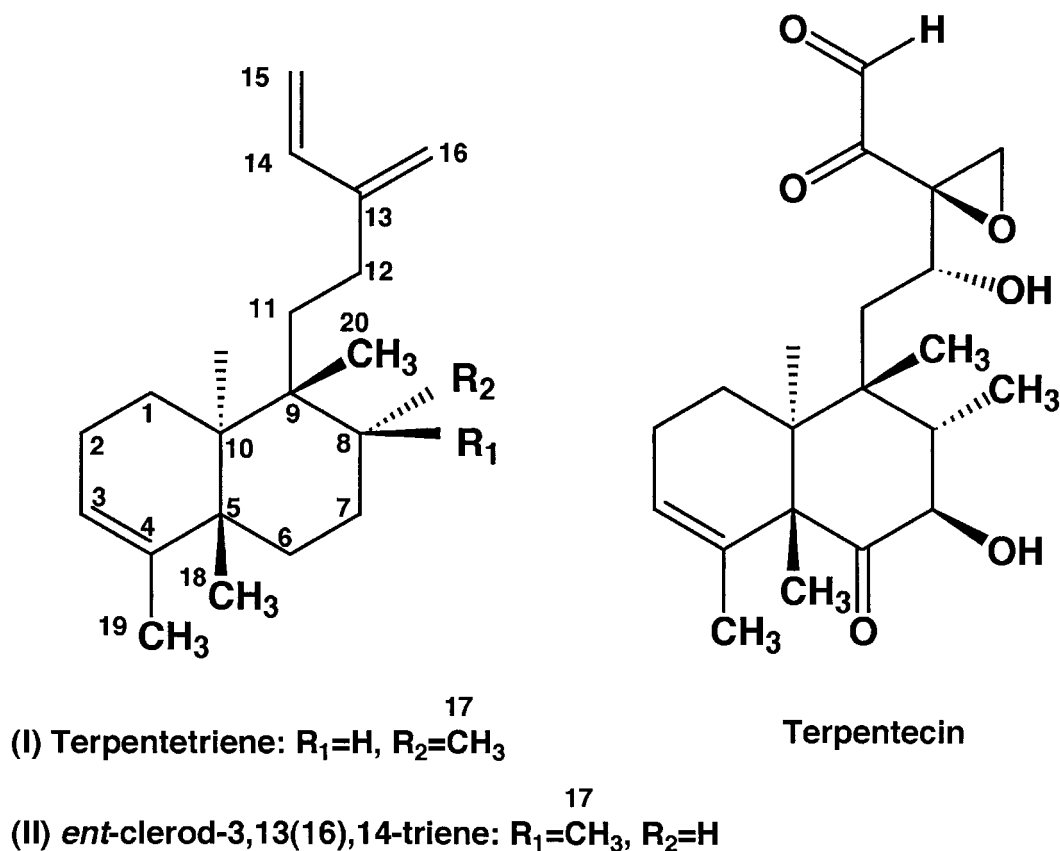


FIG. 4. Structures of terpenetriene (I), *ent*-clerod-3,13(16),14-triene (II), and TP.

actinomycetes are clustered in the genomic DNA region (11, 27). Therefore, the TP biosynthetic genes are also expected to exist in the flanking region of the GGDP synthase gene. To examine this possibility, the DNA sequence of an 8.5-kb *Bam*HI fragment containing the upstream region of the GGDP synthase gene was determined (Fig. 2). Computer analysis of the DNA sequence by Frame Analysis (6) showed seven ORFs (ORF8 to ORF14) in the same direction (Fig. 2), except for ORF9. ORF11 and ORF12 appeared to be translationally coupled.

To understand the functions of individual ORFs deduced by DNA sequencing, we searched the databases with their translated products by means of the sequence similarity search programs FASTA (33) and BLAST (3). The results are summarized in Fig. 2. In brief, each of the ORFs had significant similarity to efflux proteins responsible for antibiotic resistance (ORF8), an unknown protein found in the genomic DNA of *Streptomyces coelicolor* (ORF9), P450-like hydroxylation proteins (ORF10), DCs from eukaryotes (ORF11), a pentalenene synthase of *Streptomyces* sp. (ORF12), P450-like hydroxylation proteins (ORF13), and ferredoxin (ORF14), respectively. TP was previously reported to be synthesized from GGDP after successive cyclization, hydroxylation, and epoxidation (18), suggesting that the ORFs found in this study would encode TP biosynthetic genes.

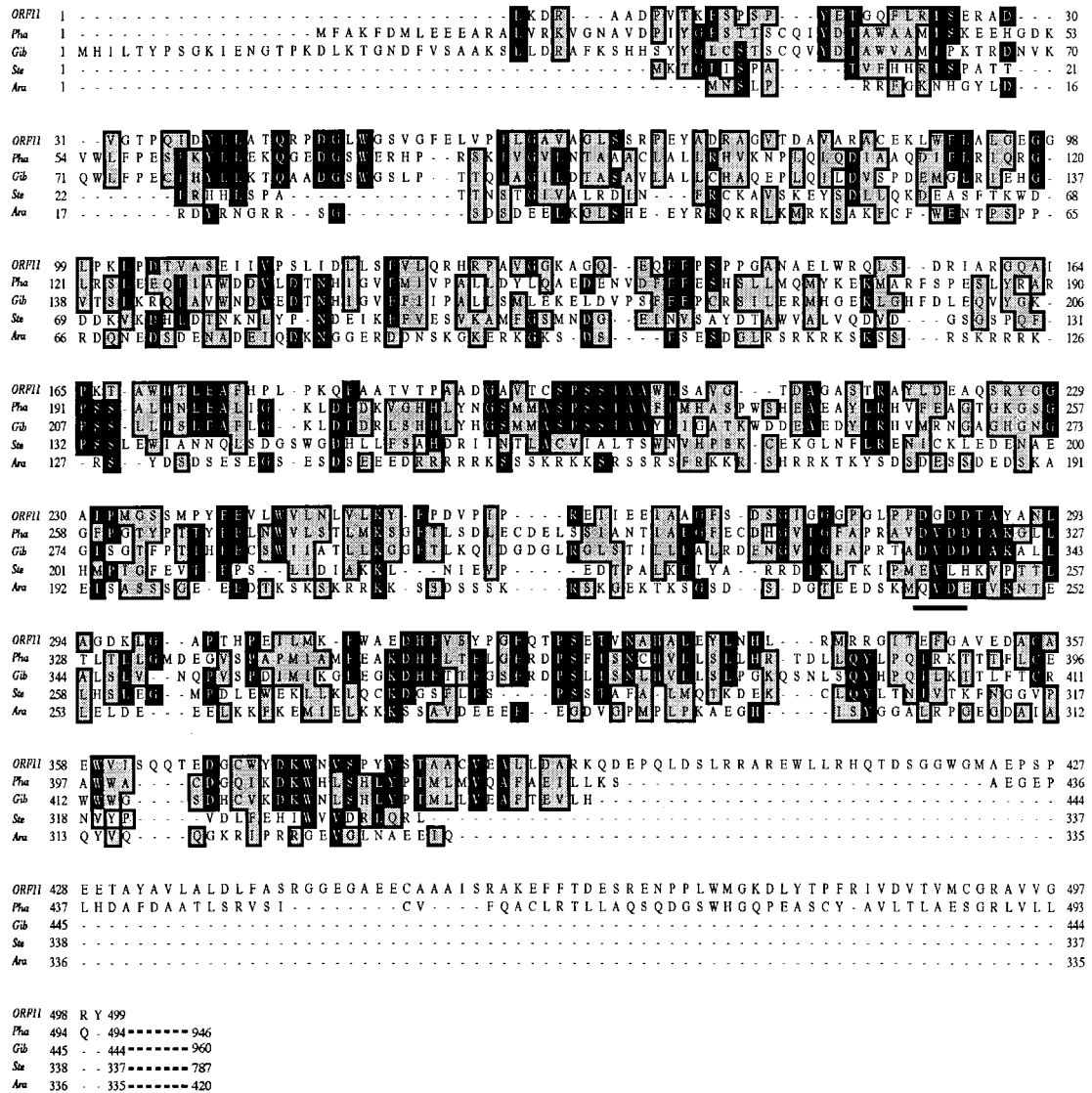
Features of ORF11 and ORF12 products deduced by primary structures. The predicted amino acid sequence of the

ORF11 product has significant homology with those of the N-terminal halves of DCs from plants and fungi: 29% identity over 494 amino acids with the *ent*-kaurene synthase (KS) from *Phaeosphaeria* sp., 27% identity over 444 amino acids with the (-)-copalyl diphosphate synthase (CPS) from *Gibberella fujikuroi*, 22% identity over 337 amino acids with the CPS from *Stevia rebaudiana*, 23% identity over 335 amino acids with the KS from *Arabidopsis thaliana*, 24% identity over 346 amino acids with the taxadiene synthase from *Taxus brevifolia*, and 21% identity over 383 amino acids with the KS from *Pisum sativum*. Sequence alignment with eukaryotic DCs revealed that the motifs QXXDGSW and DXDDTA, which were proposed to stabilize an intermediate cation during the cyclization process and to mediate substrate binding by chelation of divalent metal ions, respectively (9, 29), are conserved in the ORF11 product (⁴³QRPDGLW, ²⁸⁴DGDDTA) (Fig. 5).

The sole protein homologous to the ORF12 product was the pentalenene synthase cloned from *Streptomyces* sp. strain UC5319 (5) (Fig. 5). These proteins share 25% sequence identity over 331 amino acids. The ORF12 product has the DDX XD motif (9, 29), which was also suggested to be involved in the coordination of divalent metal ions for substrate binding (⁷⁷DDRWD).

Construction of ORF11- and ORF12-inactivated strains. To explore the possibility that the ORF11 and ORF12 genes code for cyclases responsible for TP biosynthesis, we constructed two mutant strains in which ORF11 or ORF12 was inacti-

(A)



(B)

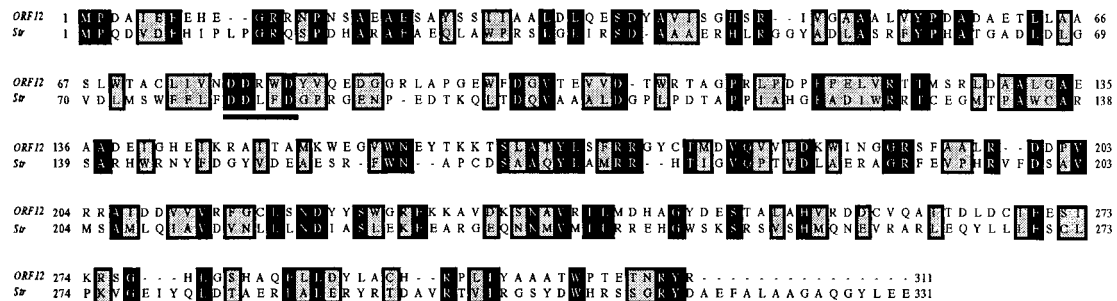


FIG. 5. Alignment of the ORF11 and ORF12 products with homologous proteins. Multiple alignments of the ORF11 product with eukaryotic DCs (A) and an alignment of the ORF12 product with pentalenene synthase (B) are shown. Identical (black) and similar (gray) amino acid residues are boxed. Aspartate-rich motifs (DXDD and DDXXD) are underlined. The numbers on the left and right are the residue numbers for each protein. Abbreviations and accession numbers: *Pha*, KS of *Phaeosphaeria* sp., GenBank accession no. AB003395-1; *Gib*, CPS of *G. fujikuroi*, GenBank accession no. Y15013-1; *Ste*, CPS of *S. rebaudiana*, GenBank accession no. AF034545-1; *Ara*, KS of *A. thaliana*, GenBank accession no. AC004044-7; *Str*, pentalenolactone synthase of *Streptomyces* sp., GenBank accession no. A54214.

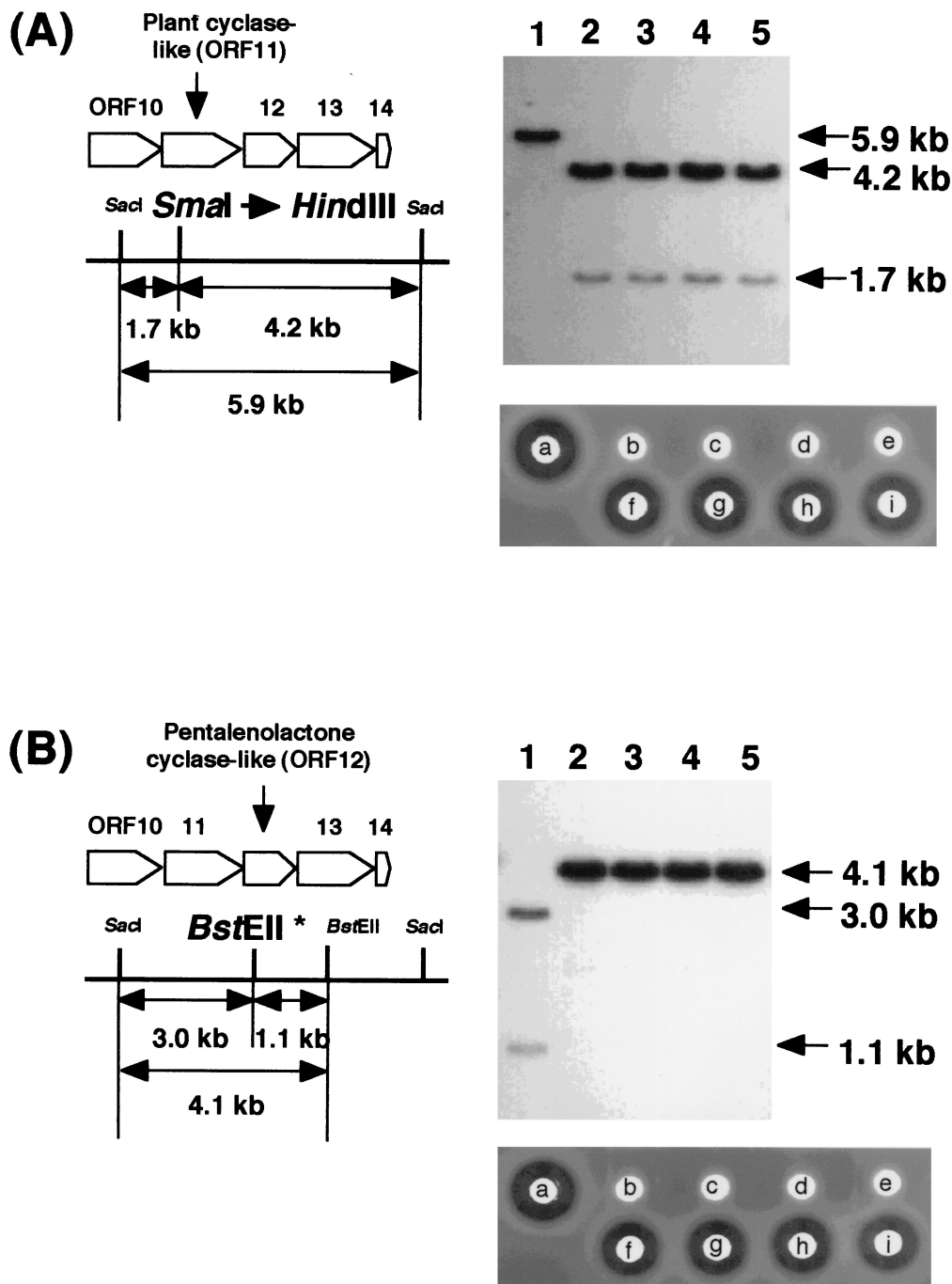


FIG. 6. Southern blot analyses of the genomic DNAs of the ORF11- and ORF12-inactivated strains and bioassay for the TP productivity of these mutants and their complemented transformants. *SacI-HindIII* digests of DNAs from ORF11-inactivated strains (A, lanes 2 to 5) and *SacI-BstEII* digests of DNAs from ORF12-inactivated strains (B, lanes 2 to 5) were, respectively, subjected to Southern blot analyses with a ³²P-labeled 5.9-kb *SacI* fragment as a probe. *SacI-HindIII* (A) and *SacI-BstEII* digests (B) of DNAs from *S. griseolosporeus* strain MF730-N6 (wild type) are shown in lane 1. Arrows indicate the sizes of the hybridized fragments. The hybridized fragments are schematically shown. Bioassays for the TP production of the wild-type strain (A, a), the ORF11-inactivated (A, b to e) and ORF12-inactivated (B, b to e) strains, the ORF11-inactivated strain harboring pEN-ORF11 (A, f to i), and the ORF12-inactivated strain harboring pEN-ORF12 (B, f to i) are shown.

vated by the gene replacement technique originally developed for *Streptomyces* strains by Anzai et al. (4). Two plasmids, pCYC1FS and pCYC2FS, in which ORF11 and ORF12 were inactivated by a frameshift mutation, respectively, were constructed. After introduction of each plasmid into the TP pro-

ducer, protoplasts from thiostrepton-resistant colonies were prepared and regenerated on RSG medium (13). Genomic DNAs of the regenerated thiostrepton-sensitive (plasmid-free) colonies were prepared, and double digested with *SacI-HindIII* (ORF11-inactivated strain) or *SacI-BstEII* (ORF12-inactivat-

ed strain) and hybridized by Southern blotting using a 5.9-kb *SacI* fragment carrying ORF11 and ORF12 as a probe (Fig. 6). In our experiments, about half of the regenerated colonies lost plasmids and the ORF11- and ORF12-inactivated strains emerged from plasmid-cured colonies at a frequency of about 10%. Four strains were randomly isolated from each of the ORF11- and ORF12-inactivated strains to examine TP productivity. Both types of mutants produced no TP, as judged by bioassay (Fig. 6) and HPLC analysis (data not shown). As described below, the ORF11 and ORF12 genes were suggested to be polycistronically transcribed with the downstream genes by Northern blot analysis. Therefore, to confirm that the loss of TP productivity of the ORF11- or ORF12-inactivated strain is not due to polar effects on the downstream genes, complementation experiments were carried out. Two plasmids, pEN-ORF11 and pEN-ORF12, which carried the ORF11 and ORF12 genes, respectively, were constructed and introduced into the corresponding mutants. As shown in Fig. 6, each of the mutants recovered TP productivity, confirming that the loss of TP productivity was due not to a polar effect of the frameshift mutation but to disruption of the ORF11 or ORF12 gene.

Expression of the cyclase genes in *S. lividans*. Considering that both the ORF11 and ORF12 products were essential for TP production and had significant similarities to isoprenoid cyclases, these proteins were expected to catalyze the cyclization of GGDP. To confirm this possibility and to elucidate the structure of the reaction product formed by these enzymes, heterologous expression of the ORF11 and ORF12 genes was employed. We constructed plasmid pWHM-TER1, in which the GGDP synthase, ORF11, and ORF12 genes were expressed under the control of the constitutive *ermE** promoter (7, 8, 38). We transformed *S. lividans* TK23 with the plasmid and investigated the production of new compounds in the transformant. *S. lividans* TK23 harboring pWHM-TER1 was cultivated in liquid medium, and the metabolites were analyzed by HPLC. A new compound, which was eluted with a retention time longer than that of geranylgeraniol, was specifically detected in the culture broth of the transformant harboring pWHM-TER1. We propose to name this new compound terpenetriene based on structural similarities to TP and the presence of three double bonds. The product was purified, and its structure was determined (see Materials and Methods and Fig. 4).

Enzymatic synthesis of terpenetriene in vitro. To verify the enzymatic function of ORF11 and ORF12, these ORF products were overproduced in *E. coli*. Soluble protein extracts from *E. coli* harboring pQE30-ORF11, pQE30-ORF12, or pQE30 (no insert) were analyzed by SDS-PAGE. His-tagged ORF11, with a molecular mass of about 55 kDa, and His-tagged ORF12, with a molecular mass of about 37 kDa, which were in good agreement with the values calculated from the deduced amino acid sequences of the enzymes, were expressed at a high level (Fig. 7, lanes 3 and 4). The expressed proteins were then purified (Fig. 7, lanes 5 and 6) and used for an enzyme assay. Incubation of both of the purified recombinant proteins with GGDP resulted in the formation of terpenetriene (Fig. 8), confirming that the ORF11 and ORF12 products were the DCs.

Northern blot analysis of DC genes. The ORF11 and ORF12 genes were shown to be essential for TP production by gene

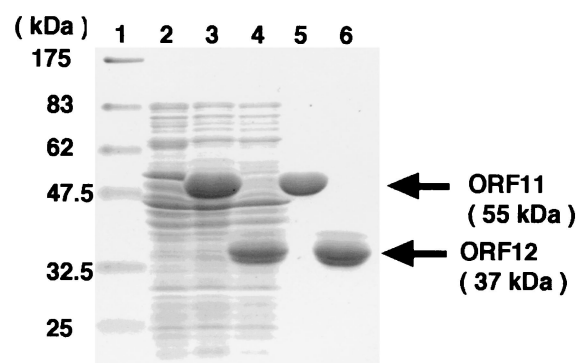


FIG. 7. Electrophoresis of overproduced and purified cyclases. Molecular mass markers (lane 1); total soluble proteins from the *E. coli* harboring pQE30 (lane 2), pQE30-ORF11 (lane 3), or pQE30-ORF12 (lane 4); purified, His-tagged ORF11 (lane 5); and purified, His-tagged ORF12 (lane 6) were analyzed by SDS-10% PAGE. Proteins were stained with Coomassie brilliant blue R-250.

disruption experiments (Fig. 6). In many cases, antibiotic biosynthetic genes cloned from actinomycetes are clustered in the genomic DNA region (11, 27). Therefore, other ORFs found in this study were also expected to code for TP biosynthetic enzymes. To study this possibility, Northern blot analysis was performed with the *XbaI-HindIII* fragment carrying the two cyclase genes as a probe. As shown in Fig. 9, a message of 6.5 kb was specifically detected. Considering that the mevalonate pathway gene cluster (ORF1 to ORF7) seemed to be polycistronically transcribed (15), we surmised that the genes encoding ORF10 to ORF14 were also polycistronically transcribed and that these gene products would participate in the biosynthesis of TP.

DISCUSSION

By sequencing analysis of the flanking region of the mevalonate pathway gene cluster cloned from a diterpenoid antibiotic (TP) producer, *S. griseolosporus* strain MF730-N6, we have found seven additional ORFs, which were suggested to encode the TP biosynthetic gene cluster. Among the ORFs identified by the sequencing analysis, the two genes encoding proteins homologous to isoprenoid cyclases were extensively analyzed and confirmed to be DCs by heterologous expression and in vitro experiments. The ORF11 and ORF12 products catalyzed the cyclization of GGDP to form terpenetriene, a compound with the same planar structure as *ent*-clerod-3,13 (16),14-triene, which was previously isolated from *Jungmannia infusca* (31). However, the compound isolated in this study had a structure with stereochemistry different from that of *J. infusca*, showing that our compound is a novel diterpenoid. The genes for the eubacterial isoprenoid cyclases, the pentalenolactone synthase from *Streptomyces* sp. strain UC5319 (10) and the squalene-hopene cyclases from *Bacillus acidocaldarius* (32), *Rhodospseudomonas palustris* (22), *Zymomonas mobilis* (35), *Bradyrhizobium japonicum* (34), *Methylococcus capsulatus* (42), and *Alicyclobacillus acidocaldarius* (46), have been characterized. However, none of the genes encoding DCs have been isolated from a prokaryote. Thus, this is the first report of the cloning and identification of DCs of eubacterial origin.

Although we do not know the detailed reaction mechanism catalyzed by the ORF11 and ORF12 products, both of the

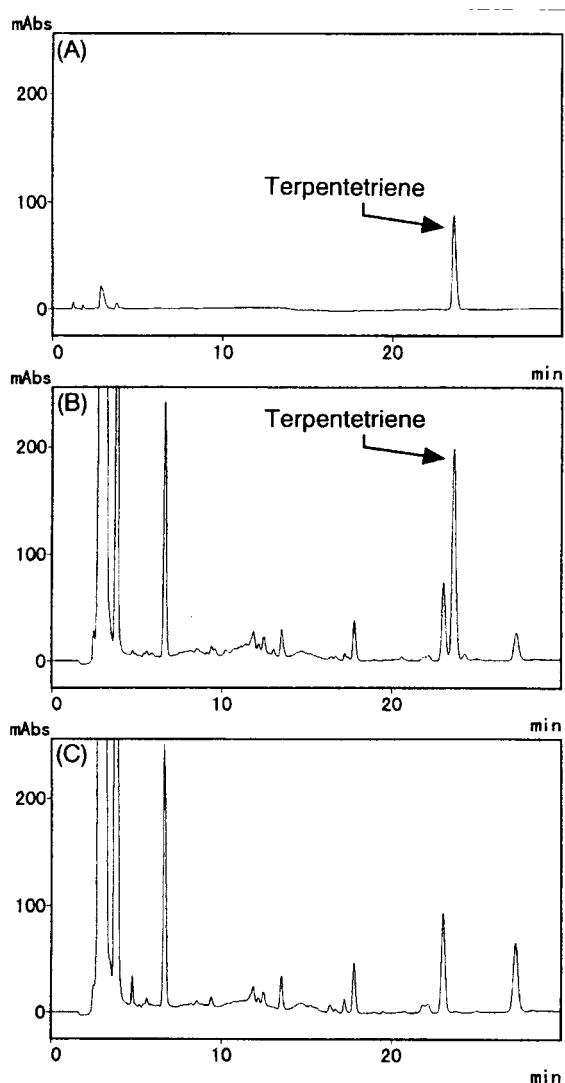


FIG. 8. HPLC analysis of the product generated by recombinant cyclases. An authentic terpenetriene standard (A) and chloroform extracts of the reaction products generated by the purified, His-tagged ORF11 and ORF12 products in the presence of GGDP (B) or the absence of GGDP (C) were analyzed by reversed-phase HPLC. mAbs, milliabsorbance.

proteins were essential for terpenetriene production because both the ORF11- and ORF12-inactivated strains produced no TP and neither the ORF11 product nor the ORF12 product alone could catalyze the cyclization of GGDP to produce terpenetriene (data not shown). DCs are classified into two major types with respect to their modes of cyclization (25). One type of reaction is initiated by ionization of GGDP to an allylic carbocation, followed by cyclization and deprotonation to the olefin (9, 29). Casbene synthase (28) and taxadiene synthase (23, 24) are representatives of this class. The other type of reaction is initiated by protonation at the 14, 15 double bond of GGDP (2, 16, 40) in a manner similar to that of the squalene-hopene cyclase from *Alicyclobacillus acidocaldarius* (1, 32). Croteau and coworkers suggested that the presence of the DXDD or DDXXD motif in DCs might be used to determine which mode of cyclization would be involved in the reaction

catalyzed by uncharacterized DCs (45). They found that the N-terminal region of the abietadiene synthase, which contains the DXDD motif, resembles those of enzymes catalyzing protonation-initiated isoprenoid cyclization, whereas the C-terminal region of the cyclase, which has the DDXXD motif, resembles those of enzymes catalyzing ionization-dependent isoprenoid cyclization. The following fact supports their assumption. In higher plants, two different DCs, CPS and KS, participate in the biosynthesis of (-)-kaurene, the precursor of the gibberellin plant hormone (16). CPS, which contains the DXDD motif and lacks the DDXXD motif (2, 5, 36, 39, 40), is known to catalyze the protonation-initiated cyclization of GGDP to (-)-copalyl diphosphate. A separate enzyme, kaurene synthase B, which has the DDXXD motif and no DXDD motif (36, 47, 48), was shown to transform this intermediate to (-)-kaurene via ionization-dependent cyclization. The predicted amino acid sequences of ORF11 and ORF12 have the DXDD and DDXXD motifs, respectively. Therefore, the ORF11 product catalyzes cyclization by protonation of the terminal double bond of GGDP, resulting in the formation of an intermediate with a phosphodiester, as does CPS; the ORF12 product completes the reaction sequence by ionization-dependent cyclization, as does KS. Kawaide et al. also reported that the DVDD motif of the KS from the fungus *Phaeosphaeria* sp. strain L487, which is a bifunctional enzyme catalyzing the two-step cyclization reaction from GGDP to *ent*-kaurene via (-)-copalyl diphosphate, is essential for CPS activity (20, 21). However, they also suggested that the DDVLD motif is involved in both the CPS and KS reactions (20, 21). Therefore, to elucidate the detailed reaction mechanisms catalyzed by the ORF11 and ORF12 products, more studies are required to clarify the function of each the ORF

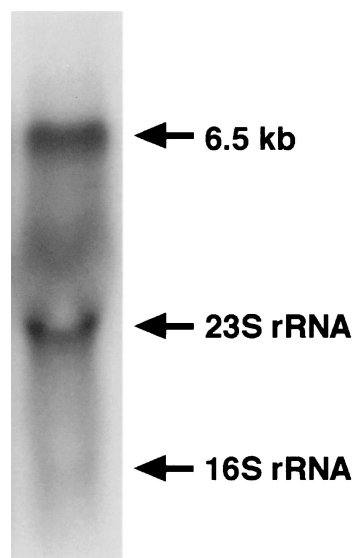


FIG. 9. Northern blot analysis of the gene cluster essential for TP production. Total RNA was isolated from *S. griseolosporeus* strain MF730-N6 grown at 30°C for 2 days and analyzed by Northern blotting using the *Xba*I-*Hind*III fragment carrying the two cyclase genes as a probe (Fig. 1). The arrowheads indicate the positions of 23S and 16S rRNAs and the sizes of hybridized fragments. A background signal due to hybridization to degraded RNAs occurred.

products. These studies are now in progress, and the results will be reported in the near future.

ACKNOWLEDGMENTS

We thank Y. Asakawa of Tokushima Bunri University for ^1H - and ^{13}C -NMR spectra of *ent*-clerod-3,13(16),14-triene, H. Kobayashi of the University of Tokyo for collecting MS spectra, M. Hamada of the Institute of Microbial Chemistry for providing the TP producer, and Y. Ono for excellent technical assistance.

This work was supported in part by RFTF (JSPS-RFTF96I00301) from JSPS and by a Grant-in-Aid for Scientific Research (C) to T.D. from JSPS.

REFERENCES

- Abe, I., M. Rohmer, and G. D. Prestwich. 1993. Enzymatic cyclization of squalene and oxidosqualene to sterol and triterpenes. *Chem. Rev.* **93**:2189–2206.
- Ait-Ali, T., S. M. Swain, J. B. Reid, T. P. Sun, and Y. Kamiya. 1997. The LS locus of pea encodes the gibberellin biosynthesis enzyme *ent*-kaurene synthase A. *Plant J.* **11**:443–454.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403–410.
- Anzai, H., Y. Kumada, O. Hara, T. Murakami, T. Ito, E. Takano, S. Imai, A. Satoh, and K. Nagaoka. 1988. Replacement of *Streptomyces hygroscopicus* genomic segments with *in vitro* altered DNA sequences. *J. Antibiot.* **41**:226–233.
- Bensen, R. J., G. S. Johal, V. C. Crane, J. T. Tossberg, P. S. Schnable, R. B. Meeley, and S. P. Briggs. 1995. Cloning and characterization of the maize An1 gene. *Plant Cell* **7**:75–84.
- Bibb, M. J., P. R. Findlay, and M. W. Johnson. 1984. The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. *Gene* **30**:157–166.
- Bibb, M. J., G. R. Janssen, and J. M. Ward. 1985. Cloning and analysis of the promoter region of the erythromycin resistance gene (*ermE*) of *Streptomyces erythraeus*. *Gene* **38**:215–226.
- Bibb, M. J., J. White, J. M. Ward, and G. R. Janssen. 1994. The mRNA for the 23S rRNA methylase encoded by the *ermE* gene of *Saccharopolyspora erythraea* is translated in the absence of a conventional ribosome-binding site. *Mol. Microbiol.* **14**:533–545.
- Bohlmann, J., G. Meyer-Gauen, and R. Croteau. 1998. Plant terpenoid synthases: molecular biology and phylogenetic analysis. *Proc. Natl. Acad. Sci. USA* **95**:4126–4133.
- Cane, D. E., J. K. Sohng, C. R. Lamberson, S. M. Rudnicki, Z. Wu, M. D. Lloyd, J. S. Oliver, and B. R. Hubbard. 1994. Pentalenene synthase. Purification, molecular cloning, sequencing, and high-level expression in *Escherichia coli* of a terpenoid cyclase from *Streptomyces* UC5319. *Biochemistry* **33**:5846–5857.
- Chater, K. F. 1990. The improving prospects for yield increase by genetic engineering in antibiotic-producing streptomycetes. *Bio/Technology* **8**:115–121.
- Connolly, J. D., and R. A. Hill. 1992. Dictionary of terpenoids. Chapman & Hall, New York, N.Y.
- Dairi, T., Y. Motohira, T. Kuzuyama, S. Takahashi, N. Itoh, and H. Seto. 2000. Cloning of the gene encoding 3-hydroxy-3-methylglutaryl coenzyme A reductase from terpenoid antibiotic-producing *Streptomyces* strains. *Mol. Gen. Genet.* **262**:957–964.
- Dairi, T., T. Nakano, K. Aisaka, R. Katsumata, and M. Hasegawa. 1995. Cloning and nucleotide sequence of the gene responsible for chlorination of tetracycline. *Biosci. Biotech. Biochem.* **59**:1099–1106.
- Hamano, Y., T. Dairi, M. Yamamoto, T. Kawasaki, K. Kaneda, T. Kuzuyama, N. Itoh, and H. Seto. 2001. Cloning of a gene cluster encoding enzymes responsible for the mevalonate pathway from a terpenoid-antibiotic-producing *Streptomyces* strain. *Biosci. Biotech. Biochem.* **65**:1627–1635.
- Hedden, P., and Y. Kamiya. 1997. Gibberellin biosynthesis: enzymes, genes and their regulation. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **48**:431–460.
- Hopwood, D. A., M. J. Bibb, K. F. Chater, T. Kieser, C. J. Bruton, H. M. Kieser, D. J. Lydiate, C. P. Smith, J. M. Ward, and H. Schreppe. 1985. Gene manipulation of *Streptomyces*, a laboratory manual. The John Innes Foundation, Norwich, United Kingdom.
- Isshiki, K., T. Tamamura, T. Sawa, H. Naganawa, T. Takeuchi, and H. Umezawa. 1986. Biosynthetic studies of terpentecin. *J. Antibiot.* **34**:1634–1635.
- Kaneda, K., T. Kuzuyama, M. Takagi, Y. Hayakawa, and H. Seto. 2001. An unusual isopentenyl diphosphate isomerase found in the mevalonate pathway gene cluster from *Streptomyces* sp. strain CL190. *Proc. Natl. Acad. Sci. USA* **98**:932–937.
- Kawaide, H., R. Imai, T. Sassa, and Y. Kamiya. 1997. *Ent*-kaurene synthase from the fungus *Phaeosphaeria* sp. L487. cDNA isolation, characterization, and bacterial expression of a bifunctional diterpene cyclase in fungal gibberellin biosynthesis. *J. Biol. Chem.* **272**:21706–21712.
- Kawaide, H., T. Sassa, and Y. Kamiya. 2000. Functional analysis of the two interacting cyclase domains in *ent*-kaurene synthase from the fungus *Phaeosphaeria* sp. L487 and a comparison with cyclases from higher plants. *J. Biol. Chem.* **275**:2276–2280.
- Kleemann, G., R. Kellner, and K. Poralla. 1994. Purification and properties of the squalene-hopene cyclase from *Rhodospseudomonas palustris*, a purple non-sulfur bacterium producing hopanoids and tetrahymanol. *Biochim. Biophys. Acta* **1210**:317–320.
- Koepp, A. E., M. Hezari, J. Zajicek, B. S. Vogel, R. E. LaFever, N. G. Lewis, and R. Croteau. 1995. Cyclization of geranylgeranyl diphosphate to taxadiene (5,11(12)-diene) is the committed step of taxol biosynthesis in Pacific yew. *J. Biol. Chem.* **270**:8686–8690.
- Lin, X., M. Hezari, A. E. Koepp, H. G. Floss, and R. Croteau. 1996. Mechanism of taxadiene synthase, a diterpene cyclase that catalyzes the first step of taxol biosynthesis in Pacific yew. *Biochemistry* **35**:2968–2977.
- MacMillan, J., and M. H. Beale. 1999. In D. Barton, K. Nakanishi, and O. Meth-Cohn (ed.), *Comprehensive natural products chemistry*, vol. 2. Diterpene biosynthesis, p. 217–243. Elsevier, Oxford, United Kingdom.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Martin, J. F., and P. Liras. 1989. Organization and expression of genes involved in the biosynthesis of antibiotics and other secondary metabolites. *Annu. Rev. Microbiol.* **43**:173–206.
- Mau, C. J. D., and C. A. West. 1994. Cloning of casbene synthase cDNA: evidence for conserved structural features among terpenoid cyclases in plants. *Proc. Natl. Acad. Sci. USA* **91**:8497–8501.
- McCaskill, D., and R. Croteau. 1997. Prospects for the bioengineering of isoprenoid biosynthesis. *Adv. Biochem. Eng. Bio/Technology* **55**:107–146.
- Miyado, S. 1993. Research on antibiotic screening in Japan over the last decade: a producing microorganism approach. *Actinomycetologica* **7**:100–106.
- Nagashima, F., S. Takaoka, and Y. Asakawa. 1998. Diterpenoids from the Japanese liverwort *Jungermannia infusca*. *Phytochemistry* **49**:601–608.
- Ochs, D., C. Kaletta, K. D. Entian, A. Beck-Sickinger, and K. Poralla. 1992. Cloning, expression, and sequencing of squalene-hopene cyclase, a key enzyme in triterpenoid metabolism. *J. Bacteriol.* **174**:298–302.
- Pearson, W. R., and D. J. Lipman. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85**:2444–2448.
- Perzl, M., P. Muller, K. Poralla, and E. L. Kannenberg. 1997. Squalene-hopene cyclase from *Bradyrhizobium japonicum*: cloning, expression, sequence analysis and comparison to other triterpenoid cyclases. *Microbiology* **143**:1235–1242.
- Reipen, I. G., K. Poralla, H. Sahm, and G. A. Sprenger. 1995. *Zymomonas mobilis* squalene-hopene cyclase gene (shc): cloning, DNA sequence analysis, and expression in *Escherichia coli*. *Microbiology* **141**:155–161.
- Richman, A. S., M. Gijzen, A. N. Starratt, Z. Yang, and J. E. Brandle. 1999. Diterpene synthesis in *Stevia rebaudiana*: recruitment and up-regulation of key enzymes from the gibberellin biosynthetic pathway. *Plant J.* **19**:411–421.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
- Schmitt-John, T., and J. W. Engels. 1992. Promoter constructions for efficient secretion expression in *Streptomyces lividans*. *Appl. Microbiol. Biotechnol.* **36**:493–498.
- Smith, M. W., S. Yamaguchi, T. Ait-Ali, and Y. Kamiya. 1998. The first step of gibberellin biosynthesis in pumpkin is catalyzed by at least two copalyl diphosphate synthases encoded by differentially regulated genes. *Plant Physiol.* **118**:1411–1419.
- Sun, T. P., and Y. Kamiya. 1994. The *Arabidopsis* GA1 locus encodes the cyclase *ent*-kaurene synthetase A of gibberellin biosynthesis. *Plant Cell* **6**:1509–1518.
- Tamamura, T., T. Sawa, K. Isshiki, T. Masuda, Y. Homma, H. Inuma, H. Naganawa, M. Hamada, T. Takeuchi, and H. Umezawa. 1985. Isolation and characterization of terpentecin, a new antitumor antibiotic. *J. Antibiot.* **38**:1664–1669.
- Tippelt, A., L. Jahnke, and K. Poralla. 1998. Squalene-hopene cyclase from *Methylococcus capsulatus* (Bath): a bacterium producing hopanoids and steroids. *Biochim. Biophys. Acta* **1391**:223–232.
- Uosaki, Y., S. Kawada, H. Nakano, Y. Saitoh, and H. Sano. 1993. UCT4B, a new antitumor antibiotic with topoisomerase II mediated DNA cleavage activity. *Structure determination*. *J. Antibiot.* **46**:235–240.
- Vara, J., M. Lewandowska-Skarbek, Y.-G. Wang, S. Donadio, and C. R. Hutchinson. 1989. Cloning of genes governing the deoxysugar portion of the erythromycin biosynthesis pathway in *Saccharopolyspora erythraea* (*Streptomyces erythraeus*). *J. Bacteriol.* **171**:5872–5881.
- Vogel, B. S., M. R. Wildung, G. Vogel, and R. Croteau. 1996. Abietadiene synthase from grand fir (*Abies grandis*). cDNA isolation, characterization, and bacterial expression of a bifunctional diterpene cyclase involved in resin acid biosynthesis. *J. Biol. Chem.* **271**:23262–23268.
- Wendt, K., K. Poralla, and G. E. Schulz. 1997. Structure and function of a squalene cyclase. *Science* **277**:1811–1815.
- Yamaguchi, S., T. Saito, H. Abe, H. Yamane, N. Murofushi, and Y. Kamiya. 1996. Molecular cloning and characterization of a cDNA encoding the gibberellin biosynthetic enzyme *ent*-kaurene synthase B from pumpkin (*Cucurbita maxima* L.). *Plant J.* **10**:203–213.
- Yamaguchi, S., T. P. Sun, H. Kawaide, and Y. Kamiya. 1998. The GA2 locus of *Arabidopsis thaliana* encodes *ent*-kaurene synthase of gibberellin biosynthesis. *Plant Physiol.* **116**:1271–1278.