Chemistry & Biology
Article

Cloning and Heterologous Expression of the Cyclooctatin Biosynthetic Gene Cluster Afford a Diterpene Cyclase and Two P450 Hydroxylases

Seung-Young Kim,[1] Ping Zhao,[1] Masayuki Igarashi,[2] Ryuichi Sawa,[2] Takeo Tomita,[1] Makoto Nishiyama,[1] and Tomohisa Kuzuyama[1,*]
[1]Biotechnology Research Center, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan
[2]Microbial Chemistry Research Center, 3-14-23 Kamiosaki, Shinagawa-ku, Tokyo 141-0021, Japan
*Correspondence: utkuz@mail.ecc.u-tokyo.ac.jp
DOI 10.1016/j.chembiol.2009.06.007

SUMMARY

Cyclooctatin, a diterpene characterized by a 5-8-5 fused ring system, is a potent inhibitor of lysophospholipase. Here we report the cloning and characterization of a complete cyclooctatin biosynthetic gene cluster from *Streptomyces melanosporofaciens* MI614-43F2 and heterologous production of cyclooctatin in *S. albus*. Sequence analysis coupled with subcloning and gene deletion revealed that the minimal cyclooctatin biosynthetic gene cluster consists of four genes, *cotB1* to *cotB4*, encoding geranylgeranyl diphosphate (GGDP) synthase, terpene cyclase (CotB2), and two cytochromes P450, respectively. Incubation of the recombinant CotB2 with GGDP resulted in the formation of cyclooctat-9-en-7-ol, an unprecedented tricyclic diterpene alcohol. The present study establishes the complete biosynthetic pathway of cyclooctatin and provides insights into both the stereospecific diterpene cyclization mechanism of the GGDP cyclase and the molecular bases for the stereospecific and regiospecific hydroxylation.

INTRODUCTION

Cyclooctatin (Figure 1) is a potent inhibitor of lysophospholipase, which catalyzes the hydrolysis of the fatty acid ester bonds of lysophospholipids (Aoyagi et al., 1992). This inhibitor was isolated from the broth of *Streptomyces melanosporofaciens* MI614-43F2 while screening for lead compounds for the development of anti-inflammatory drugs targeting lysophospholipase (Aoyagi et al., 1992).

Cyclooctatin has a unique tricyclic diterpene skeleton (C_{20}) characterized by a 5-8-5 fused ring system (Aoyama et al., 1992). In the biosynthesis of cyclic terpenes such as cyclooctatin, the prenyl chain length is catalytically determined by a polyprenyl diphosphate synthase, the carbon skeleton is generated by terpene cyclase, and other enzymes are involved in chemical modifications such as oxidation, reduction, and other processes. In these multistep biosynthetic processes, terpene cyclases such as geranylgeranyl diphosphate (GGDP) cyclase are frequently key branch point enzymes that catalyze complex cyclizations (Christianson, 2006).

To date, characterized GGDP cyclases have been cloned predominantly from higher plants and fungi (MacMillan and Beale, 1999). Based on the cyclization mechanism, these GGDP cyclases can be classified into four groups: (1) Type A, (2) Type B, (3) Type A-Type B, and (4) Type B-Type A (see Figure S1 available online) (MacMillan and Beale, 1999). The Type A reaction is initiated by ionization of the diphosphate of GGDP to an allylic carbocation, followed by cyclization and deprotonation to an olefin or water capture to form an alcohol. Casbene synthase, which catalyzes the formation of (1*S*,3*R*)-casbene, is representative of this group. Type B cyclization is initiated by protonation at the terminal double bond of GGDP. *ent*-copalyl diphosphate synthase catalyzes the formation of *ent*-copalyl diphosphate via this cyclization mechanism. Taxadiene synthase, which catalyzes the formation of taxa-4,11-diene from GGDP, is representative of GGDP cyclases using the Type A-Type B cyclization mechanism. This enzyme reaction is initiated by ionization of the diphosphate of GGDP and is continued by proton-induced cyclization. Abietadiene synthase catalyzes the formation of abieta-7,13-diene from GGDP by Type B-Type A cyclization, which is initiated by proton-induced cyclization and continued by ionization of the diphosphate of GGDP.

To date, only three GGDP cyclases have been cloned from prokaryotes and characterized: terpentedienol diphosphate synthase from the terpentecin-producing *Kitasatospora griseola* (Hamano et al., 2002), *ent*-copalyl diphosphate synthase from the viguiepinol-producing *Streptomyces* sp. KO-3988 (Kawasaki et al., 2004), and halima-5,13-diene diphosphate synthase from *Mycobacterium tuberculosis* H37Rv3377c (Figure S1) (Nakano et al., 2005). Each of these three GGDP cyclases catalyzes proton-induced cyclization reactions, classified above as Type B. These three cyclases share 25%–34% sequence similarity. In addition, these enzymes have a DXDD/T motif, which is usually found in Type B diterpene cyclases. This motif is believed to provide the acidic proton that initiates cyclization by protonating the terminal double bond.

Very recently, fusicocca-2,10(14)-diene synthase (PaFS), classified as a Type A-Type B cyclase, was cloned from the plant-pathogenic fungus *Phomopsis amygdali* and characterized (Figure S1) (Toyomasu et al., 2007). Interestingly, PaFS is a multifunctional enzyme that consists of a GGDP cyclase domain at the N terminus and a GGDP synthase domain at the C terminus.

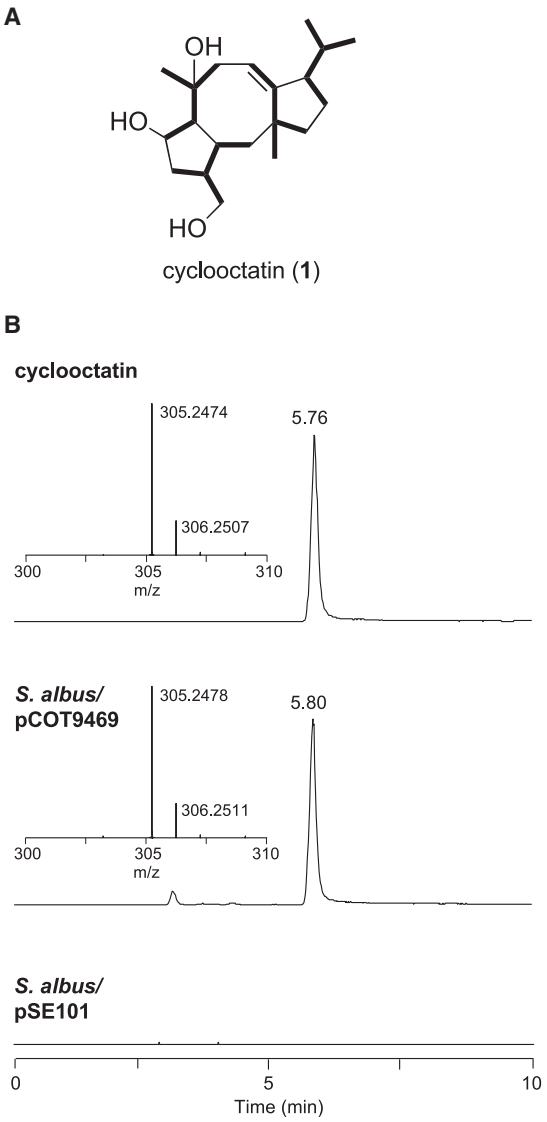


Figure 1. Identification of the Complete Cyclooctatin Biosynthetic Gene Cluster

(A) Structure of cyclooctatin. Bold lines represent isoprene units.

(B) LC-MS analysis of the cyclooctatin production of an *S. albus* transformant harboring pCOT9469. Mass chromatogram (*m/z*: 305.2444–305.2506) and the corresponding high resolution mass spectrum (inset) for the ethyl acetate extract of broth from the transformant. *S. albus*/pSE101 was also analyzed as a negative control.

Due to the similarity of the carbon skeleton of fusicocca-2,10(14)-diene and cyclooctatin (Toyomasu et al., 2007), it was expected that fusicocca-2,10(14)-diene also be a key intermediate in the synthesis of cyclooctatin and that a PaFS-like GGDP cyclase be involved in cyclooctatin biosynthesis. However, no biosynthetic intermediates of cyclooctatin have been identified and no biosynthetic genes for cyclooctatin have been cloned. In addition, the absolute configuration of cyclooctatin has not been determined.

Here, we report the cloning of the cyclooctatin biosynthetic gene cluster from *S. melanosporofaciens* MI614-43F2, its functional expression in the heterologous host *S. albus*, and the structural elucidation of two novel tricyclic diterpene intermediates. Characterization of a GGDP cyclase for cyclooctatin biosynthesis revealed the molecular basis for an unprecedented cyclization of GGDP. This cyclase is the first prokaryotic enzyme reported to synthesize a tricyclic diterpene alcohol. Identification and functional characterization of two cytochromes P450 for cyclooctatin biosynthesis are also described. These results establish the complete biosynthetic pathway of cyclooctatin.

RESULTS AND DISCUSSION

To isolate the complete cyclooctatin gene cluster from *S. melanosporofaciens* MI614-43F2, we prepared a cosmid library of its genomic DNA. Cloning of 35 to 45 kb *Sau*3AI-*Sau*3AI fragments followed by λ phage transduction yielded a library of 3.6×10^3 cosmid-containing clones of *Escherichia coli* XL1-Blue MRF'. A 0.5 kb DNA fragment containing the partial GGDP synthase gene, expected to be involved in the biosynthesis of the diterpene cyclooctatin, was amplified by PCR from the MI614-43F2 genome using conserved sequences of GGDP synthases. Five apramycin-resistant clones that hybridized with the GGDP synthase gene were found by colony hybridization. Of the five cosmids, three were eliminated because they had the same insert DNA.

Next, the remaining two cosmids, pCOT2861 and pCOT9469, were transformed into *S. albus* to verify that they contained all the genes required for cyclooctatin biosynthesis. These transformants, *S. albus*/pCOT2861 and *S. albus*/pCOT9469, were cultured and the broths were extracted with ethyl acetate. Coupled liquid chromatography high-resolution mass spectrometry (LC-MS) analysis of the extracts revealed cyclooctatin ([M ($C_{20}H_{34}O_3$)+H–H_2O]$^+$ *m/z* 305.2478, $C_{20}H_{33}O_2$, calcd 305.2481) in the broth of *S. albus*/pCOT9469 (Figure 1B). This heterologous expression experiment in *S. albus* unequivocally demonstrated that the cosmid clone pCOT9469 contains all the genes required for the production of cyclooctatin in a 34 kb genomic insert.

We sequenced pCOT9469 using a shotgun approach. Analysis of the 34 kb DNA insert revealed 24 open reading frames, whose organization is shown in Table S1 and Figure S2. Among them, we found four putative cyclooctatin biosynthetic genes encoding GGDP synthase (CotB1), terpene cyclase (CotB2), and two putative cytochromes P450 (CotB3 and CotB4). CotB2 showed no significant similarity to any protein in databases searched using the programs BLAST and FASTA. However, the position-specific iterated BLAST program revealed that CotB2 contains terpene cyclase domains, including a polyprenyl diphosphate-Mg^{2+} binding site. In addition, since antibiotic biosynthesis genes cloned from actinomycetes are usually clustered in the genomic DNA region, we assumed that CotB2 encodes the GGDP cyclase for cyclooctatin biosynthesis. ORF5 was also deduced as a further putative cytochrome P450 enzyme from sequence comparison. However, this ORF is likely a pseudogene because it lacks most of the N-terminal sequence compared to many previously characterized cytochromes P450. Thus, we did not include orf5 in the following heterologous expression experiments.

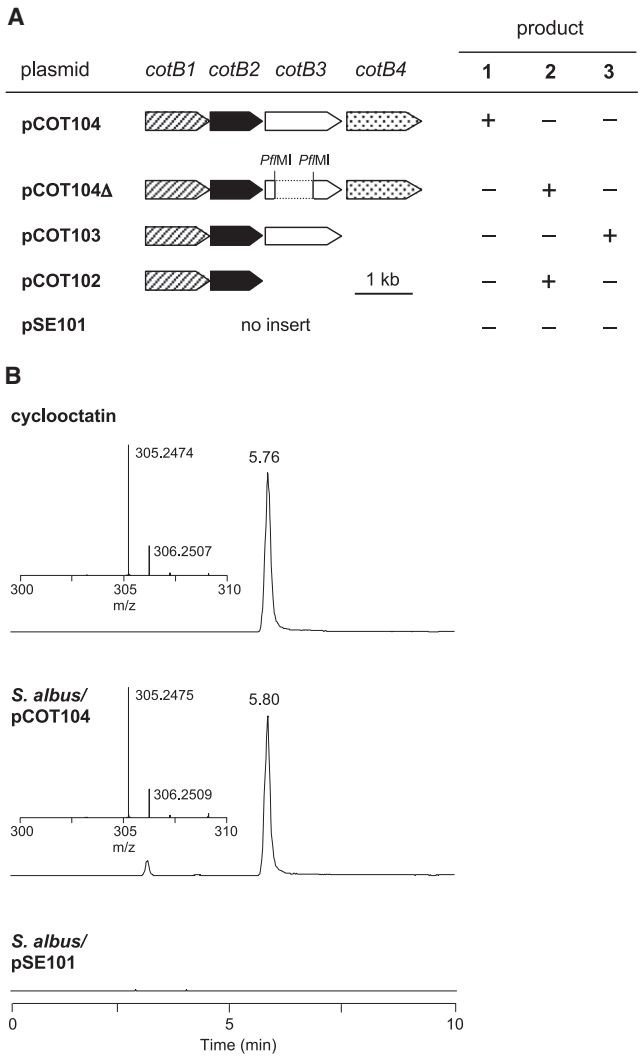


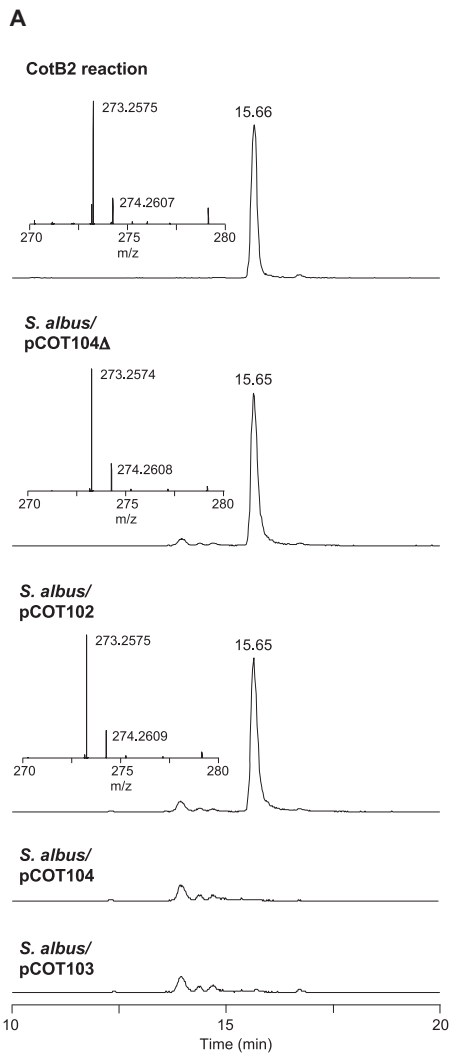
Figure 2. Identification of the Minimal Cyclooctatin Biosynthetic Gene Cluster

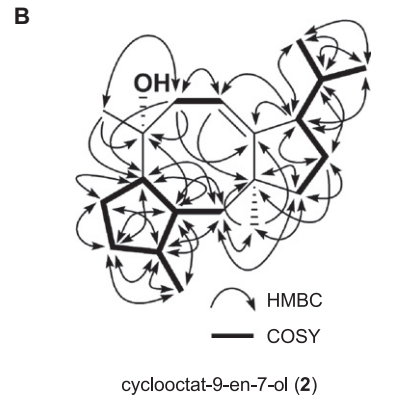
(A) Plasmids used to determine the minimal cluster. Descriptions of each plasmid are listed in Table S2. The function of the *cotB1*, *cotB2*, *cotB3*, and *cotB4* gene products was deduced as GGDP synthase, terpene cyclase, and two cytochromes P450, respectively, from sequence comparisons. The production or nonproduction of cyclooctatin **1**, compound **2**, and compound **3** in the strains carrying different plasmids is indicated with + and – symbols. (B) LC-MS analysis of cyclooctatin production of an *S. albus* transformant harboring pCOT104. Mass chromatogram (*m/z*: 305.2444–305.2506) and the corresponding high resolution mass spectrum (inset) of the ethyl acetate extract of broth from the transformant. *S. albus*/pSE101 was also analyzed as a negative control.

To verify that the four *cotB* genes encode enzymes required for the biosynthesis of cyclooctatin, we constructed a series of plasmids (pCOT102, pCOT103, pCOT104, and pCOT104Δ) by subcloning four DNA fragments containing the *cotB* genes into the *Streptomyces-E. coli* shuttle vector pSE101 (Figure 2A and Table S2). The *cotB* genes in these plasmids are most likely transcribed as a single operon by read-through from the promoter of the replication protein in the vector pSE101. The *S. albus* transformants harboring each plasmid were cultured and the broths were then extracted with ethyl acetate. LC-MS analysis of the extracts showed that cyclooctatin ([M ($C_{20}H_{34}O_3$)+H–H_2O]$^+$ *m/z* 305.2475, $C_{20}H_{33}O_2$, calcd 305.2481) was produced only by *S. albus*/pCOT104 (Figure 2B), thereby allowing the identification of a minimal gene cluster of 4.8 kb required for cyclooctatin biosynthesis.

To elucidate the enzymatic function of CotB2, we overexpressed the *cotB2* gene in *E. coli* and characterized the gene product. The molecular mass of CotB2 was estimated to be 34 kDa by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and 78 kDa by gel filtration chromatography, suggesting that CotB2 is likely a homodimer. Functional analysis of the CotB2 recombinant enzyme, using geranyl diphosphate (C_{10}), farnesyl diphosphate (C_{15}), or GGDP (C_{20}) as a substrate, showed that only GGDP is converted into an unidentified compound (**2**) with a retention time of 15.66 min on high performance liquid chromatography (HPLC) in the presence of $MgCl_2$ (Figure 3A). Elimination of $MgCl_2$ from the GGDP reaction mixture resulted in no formation of **2**, suggesting that the CotB2 enzyme is an ionization-dependent terpene cyclase.

Compound **2** was isolated from the CotB2 reaction mixture by ethyl acetate extraction and its molecular formula was deduced to be $C_{20}H_{34}O$ by positive high-resolution mass spectrometry ([M+H–H_2O]$^+$ *m/z* 273.2583, $C_{20}H_{33}$, calcd 273.2577). This is consistent with a diterpene alcohol with four degrees of unsaturation. The ^{1}H nuclear magnetic resonance (NMR) spectra (Figures S4 and S5) displayed signals arising from two fully substituted methyl groups (δ 1.12 and 1.19 [each 3H, s]), three doublet methyl groups (δ 0.77, 0.94 [each 3H, d, J = 6.8 Hz], and 0.87 [3H, d, J = 7.6 Hz]), and one olefinic proton (δ 5.21 [1H, dd, J = 10.3, 8.3 Hz]), while 17 remaining protons appeared in the region δ 1.20–2.70. The ^{13}C NMR spectrum (Figure S6) showed 20 carbon resonances, with the two lower field signals at δ 153.1 and 117.8 ppm clearly indicating the existence of a trisubstituted (>C = CH–) double bond. Directly bonded carbon and hydrogen atoms were assigned from the heteronuclear single quantum coherence (HSQC) spectrum (Table 1). A quaternary carbon signal at δ 76.2 indicated the existence of a tertiary hydroxyl group while the higher field region contained signals for one quaternary carbon, five methine, six methylene, and five methyl carbons. Comparison of the ^{1}H and ^{13}C NMR spectroscopic data (Table 1) of **2** with those of cyclooctatin (**1**) (Aoyama et al., 1992) suggested a 5-8-5 tricyclic structure for **2**. Extensive NMR spectroscopic analysis, including COSY and HMBC (Figure 3B), allowed us to elucidate the structure of **2** and showed that it has a hydroxyl group at C-7 and a double bond at C-9 and C-10. Compound **2** was therefore named cyclooctat-9-en-7-ol. Its relative configuration (2*S**, 3*S**, 6*S**, 7*R**, 11*S**, and 14*S**) was deduced on the basis of NOESY correlations and further confirmed by molecular modeling using molecular mechanics (MM2) calculations within ChemDraw Ultra Ver. 9.0 (CambridgeSoft) (Figure 3C). The NOESY spectrum showed NOE correlations of H-2/H-3, H-1/(H-12b, H-18), H-6/(H-1, H-19), H-9/H-15, H-13a, H-16, H-14/(H-13b, H-17), and H-19/(H-5b, H-9), indicating that H-6, H-14, Me-18, and Me-19 are in a β orientation. In addition, NOESY exhibited correlations of H-2/H-8/H-20 and H-20/H-12a, in good agreement with the α configurations of H-2 and Me-20 (Figure 3C). These results suggest that the *cotB2* gene encodes a novel diterpene cyclase, cyclooctat-9-en-7-ol synthase.





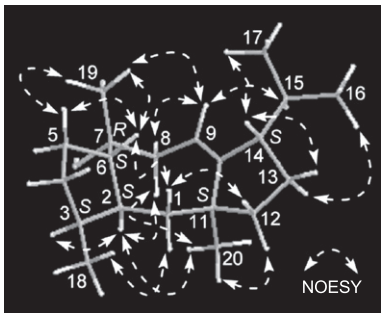


Figure 3. Identification and Structure Determination of Cyclooctat-9-en-7-ol (2)

(A) Mass chromatograms (*m/z*: 273.2550–273.2604) and the corresponding mass spectra (inset) of a reaction product generated from GGDP by the recombinant CotB2 and of the ethyl acetate extracts of the broths from *S. albus*/pCOT104Δ and *S. albus*/pCOT102. Cyclooctat-9-en-7-ol **2** was not detected in the broths from *S. albus*/pCOT104 and *S. albus*/pCOT103.

(B) HMBC and COSY correlations in **2**.

(C) Conformation calculated by MM2 and key NOESY correlations in **2**.

Compound **2** was detected in the broths of *S. albus*/pCOT102 and *S. albus*/pCOT104Δ, but not *S. albus*/pCOT103 (Figures 2A and 3A), strongly suggesting that CotB3 catalyzed the hydroxylation of **2** at C-5 or C-18 to give an unidentified intermediate of cyclooctatin. To identify the intermediate leading to cyclooctatin from **2**, we further analyzed the broth from *S. albus*/pCOT103 by LC-MS. A new product **3** with a retention time of 11.58 min (Figure 4A) was purified from the broth using silica gel chromatography and its molecular formula was deduced to be $C_{20}H_{34}O_2$ by positive high-resolution mass spectrometry ($[M+H-H_2O]^+$ *m/z* 289.2522, $C_{20}H_{33}O$, calcd 289.2526; $[M+H-(H_2O \times 2)]^+$ *m/z* 271.2418, $C_{20}H_{31}$, calcd 271.2426). The ^1H spectra (Figures S7 and S8) and ^{13}C NMR spectrum (Figure S9) closely resembled those of **2**, except for the signals corresponding to the five-membered ring located on C-2/C-6 (Table 1). Instead of a methylene signal at C-5 in **2**, an oxygen-bearing methine signal (δ_H 4.43 [br dd, 4.8, 3.4], δ_C 75.5) was observed in **3**, indicating an additional hydroxyl group at C-5 in **3**. Extensive analysis of COSY and HMBC NMR spectra (Figure 4B) determined the structure of **3** to be cyclooctat-9-en-5,7-diol (**3**); the relative configuration (2*S**, 3*S**, 5*S**, 6*R**, 7*R**, 11*S**, and 14*S**) was deduced in the same manner as for **2**. The β configuration for H-5 was established from the NOE correlations of H-5 with H-6 and H-19 (Figure 4C).

Taken together, identification of the two biosynthetic precursors to cyclooctatin, **2** and **3**, helped elucidate the complete biosynthetic pathway for cyclooctatin (Figure 5). That is, CotB1 synthesizes GGDP from isopentenyl diphosphate and dimethylallyl diphosphate, and then CotB2 catalyzes the stereospecific cyclization of GGDP to give **2**. Next, the cytochrome P450, CotB3, likely catalyzes the stereospecific hydroxylation of **2** at C-5 to form **3**. Finally, the other cytochrome P450, CotB4, catalyzes the hydroxylation of **3** at C-18 to yield the final product, cyclooctatin. In addition, since the relative configuration of cyclooctatin must be the same as that of **3**, we also deduced the relative configuration (2*S**, 3*S**, 5*S**, 6*R**, 7*R**, 11*S**, and 14*S**) of cyclooctatin. This is the first determination of the relative stereochemistry of cyclooctatin.

Based on the structure of **2**, we propose that the CotB2-catalyzed reaction uses the Type A-Type B cyclization mechanism (Figure 5). This mechanism involves ionization of GGDP and C1 to C-11 bond formation, followed by ring closure via *re*-face attack at C-10 to give a carbocation (**4**). Next, **4** is converted to a carbocation (**5**) by two consecutive 1,2-hydride shifts, from C-10 to C-14 and from C-14 to C-15. Then, deprotonation of **5** occurs at C-9 to give an olefin (**6**), followed by a C-6 to C-2 bond formation, initiated by the *re*-face protonation of **6** at C-3, to form a carbocation (**7**). Finally, **7** reacts with the water nucleophile to yield the CotB2 reaction product **2**.

To date, crystal structures of terpene cyclases such as 5-epi-aristolochene synthase from tobacco (C_{15}) (Starks et al., 1997), aristolochene synthases (C_{15}) from *Penicillium roqueforti* (Caruthers et al., 2000) and *Aspergillus terreus* (Shishova et al., 2007), and pentalenene synthase (C_{15}) from *Streptomyces* sp. UC5319 (Lesburg et al., 1997) have been determined, and metal ion binding sequences essential for the cyclization reactions have been suggested (Christianson, 2006). Alignment (Corpet, 1988; Poirot et al., 2003) of cyclooctat-9-en-7-ol synthase

Table 1. NMR Spectral Data for Cyclooctat-9-en-7-ol (2), Cyclooctat-9-en-5,7-diol (3), and Cyclooctatin (1)

	2[a]		3[a]		1[b]	
#	δ_C	δ_H (mult., *J* in Hz)	δ_C	δ_H (mult., *J* in Hz)	δ_C	δ_H (mult., *J* in Hz)
1	45.8 (t)	1.56 (br d, 13.0), 1.21 (t, 13.0)	44.7 (t)	1.57 (br d, 13.1), 1.05 (t, 13.1)	45.6 (t)	1.68 (br d, 12.8), 1.20 (t, 12.8)
2	38.1 (d)	2.12 (m)	36.4 (d)	2.37 (m)	35.8 (d)	2.56 (m)
3	40.0 (d)	1.99 (m)	35.6 (d)	2.56 (m)	44.9 (d)	2.61 (m)
4	34.7 (t)	1.63 (m), 1.23 (m)	44.0 (t)	1.64 (dd, 12.7, 5.2), 1.38 (dd, 12.7, 3.4)	39.7 (t)	1.71 (br dd, 12.6, 5.0), 1.38 (dt, 12.6, 3.4)
5	26.7 (t)	1.65 (m), 1.51 (m)	75.5 (d)	4.43 (br dd, 4.8, 3.4)	75.7 (d)	4.44 (br dd, 5.0, 3.4)
6	52.9 (d)	1.94 (br dd, 8.3, 5.5)	57.0 (d)	1.95 (t, 4.8)	58.0 (d)	1.97 (t, 5.0)
7	76.2 (s)		77.3 (s)		78.4 (s)	
8	41.3 (t)	2.67 (br t, 11.6, 11.0), 1.98 (dd, 12.4, 7.6)	41.8 (t)	2.69 (br t, 11.7, 11.0),1.94 (m)	42.2 (t)	2.72 (br t, 11.6),1.91 (dd, 12.8, 7.4)
9	117.8 (d)	5.21 (ddd, 10.3, 7.6, 2.0)	117.9 (d)	5.23 (ddd, 10.3, 7.6, 2.0)	119.1 (d)	5.28 (ddd, 10.3, 7.4, 2.2)
10	153.1 (s)		153.3 (s)		154.5 (s)	
11	44.8 (s)		44.69 (s)		45.9 (s)	
12	45.5 (t)	1.59 (m), 1.40 (m)	45.6 (t)	1.59 (m), 1.37 (m)	46.6 (t)	1.59 (m), 1.42 (m)
13	23.5 (t)	1.53 (m), 1.36 (m)	23.3 (t)	1.51 (m), 1.36 (m)	24.3 (t)	1.56 (m), 1.38 (m)
14	53.9 (d)	2.28 (m)	53.9 (d)	2.26 (m)	55.1 (d)	2.30 (m)
15	29.1 (d)	1.80 (m)	29.0 (d)	1.80 (m)	30.2 (d)	1.83 (m)
16	17.5 (q)	0.77 (d, 6.8)	17.4 (q)	0.76 (d, 6.9)	17.8 (q)	0.79 (d, 6.6)
17	22.2 (q)	0.94 (d, 6.8)	22.2 (q)	0.93 (d, 7.6)	22.5 (q)	0.96 (d, 6.6)
18	15.6 (q)	0.87 (d, 7.6)	15.4 (q)	0.95 (d, 7.6)	63.4 (t)	3.66 (dd, 10.8, 7.4), 3.55 (dd, 10.8, 6.8)
19	26.6 (q)	1.12 (s)	26.6 (q)	1.35 (s)	26.7 (q)	1.33 (br s)
20	25.0 (q)	1.19 (s)	24.9 (q)	1.21 (s)	25.2 (q)	1.25 (s)

[a] Measured in $CDCl_3$.
[b] The NMR data listed (CD_3OD) were reported previously by Aoyama et al. (1992).

(C_{20}) with these characterized terpene cyclases and the fusicocca-2,10(14)-diene synthase (C_{20}) domain of PaFS revealed two possible metal binding motifs: the "aspartate-rich" motif ^{110}DDMD and the "NSE/DTE" motif ^{220}NDFYSYDRE in cyclooctat-9-en-7-ol synthase, both of which are conserved in ionization-dependent terpene cyclases (Figure S10) (Christianson, 2006). Presumably, the motifs trigger ionic cleavage of the C-O bond in the substrate GGDP. However, the alignment could not predict the amino acid residues involved in the protonation of C-3 and the hydroxylation at C-7; both reactions are essential steps for GGDP cyclization catalyzed by cyclooctat-9-en-7-ol synthase. Further insights into the structural basis for the unique diterpene cyclization will require crystal structures of cyclooctat-9-en-7-ol synthase complexed with the substrate or the product.

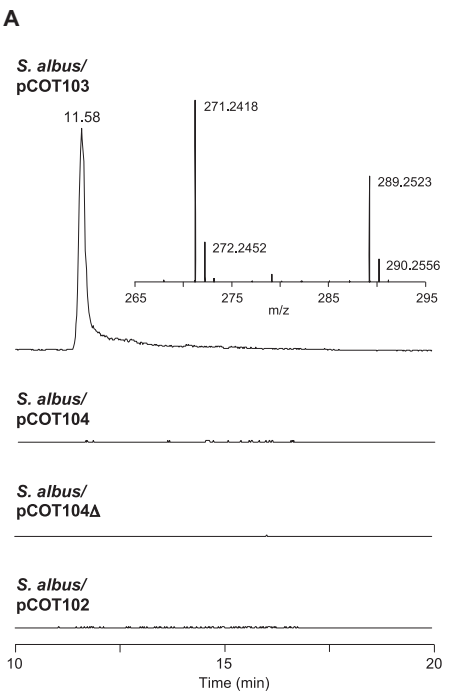
SIGNIFICANCE

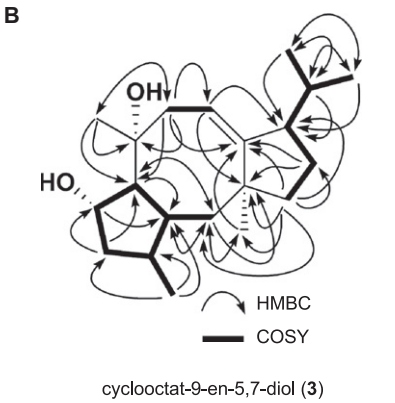
We have cloned a complete cyclooctatin biosynthetic gene cluster from *S. melanosporofaciens* MI614-43F2 and characterized the heterologous production of cyclooctatin in *S. albus*. The minimal cluster was identified to be located in a 4.8 kb region and contained four genes, *cotB1–4*. Identification and structural elucidation, including the relative configurations of the new biosynthetic intermediates, established the complete biosynthetic pathway of cyclooctatin. Biochemical studies on CotB2 revealed this diterpene cyclase to be the first prokaryotic enzyme known to synthesize a tricyclic diterpene alcohol (2). Presumably, synthesis is initiated by ionization of the diphosphate of GGDP and is continued by proton-induced cyclization. This reaction mechanism is similar to that of taxol biosynthesis by taxadiene synthase. Very recently, we determined the crystal structure of the apo form of CotB2 (S.-Y.K., T.T., M.N., and T.K., unpublished data). Thus, a combination of the crystal structure and additional biochemical data on CotB2 may provide insights into the structural basis that governs the mechanism of stereoselective cyclization catalyzed by diterpene cyclases, including taxadiene synthase. In addition, future structure-based enzyme engineering may be able to generate more diverse cyclic terpenoids with novel bioactivities.

EXPERIMENTAL PROCEDURES

Bacterial Strain, Plasmids, and Culture Conditions

S. melanosporofaciens MI614-43F2, isolated by the Microbial Chemistry Research Center (Tokyo, Japan), was used as a cyclooctatin producer in this study (Aoyagi et al., 1992). Authentic cyclooctatin was obtained by fermentation of *S. melanosporofaciens* MI614-43F2 as described previously (Aoyama et al., 1992). Cosmid vector pOJ446 (Bierman et al., 1992) and plasmid vector pSE101 (Dairi et al., 1995) were used to clone the cyclooctatin biosynthetic gene cluster. Cosmids and plasmids used in this study are listed





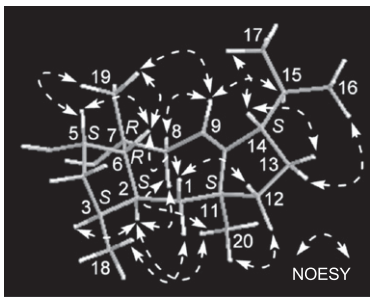


Figure 4. Identification and Structure Determination of Cyclooctat-9-en-5,7-diol (3)

(A) Mass chromatogram (*m/z*: 289.2497–289.2555) and the corresponding mass spectrum (inset) of the ethyl acetate extract of the broth from *S. albus*/pCOT103. Cyclooctat-9-en-5,7-diol **3** was not detected in the broths from *S. albus*/pCOT104, *S. albus*/pCOT104Δ, and *S. albus*/pCOT104.

(B) HMBC and COSY correlations in **3**.

(C) Conformation calculated by MM2 and key NOESY correlations in **3**.

in Table S2. All *S. albus* transformants were grown in tryptone soya broth (Kanto Chemical), containing 50 μg/ml apramycin or 50 μg/ml thiostrepton, at 30°C for 3 days by rotary shaking (250 rpm).

PCR Amplification of a GGDP Synthase Gene Probe from the MI614-43F2 Genome

Several homologous regions of GGDP synthases were found in *Kitasatospora griseola* (accession number AB037907) (Dairi et al., 2001) and *Streptomyces* sp. KO3988 (accession number AB183750) (Kawasaki et al., 2004). Two amino acid sequences (LL/IHDDVMD and FQL/IRDDLL) are highly conserved among them and thus the corresponding forward oligonucleotide primer, pGGDPS1, (5′-CTSMTSCACGACGACGTSATGGAC-3′) and the reverse primer, pGGDPS2 (5′-SAGSAGGTCGTCSCGSAKCTGGAA-3′), were synthesized (Operon Biotechnologies). The letters S, M, and K in these primers mean G or C, A or C, and G or T, respectively. PCR was carried out in PCR buffer (Roche Diagnostics). The PCR-amplified 0.5 kb DNA fragment was used as the DNA probe for cosmid library screening.

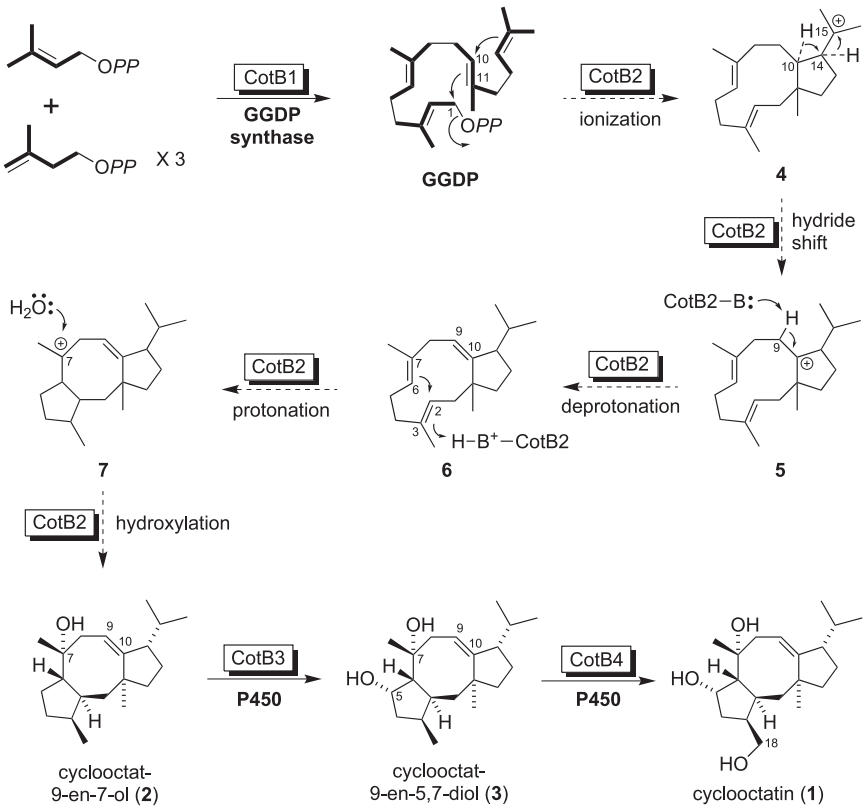


Figure 5. Biosynthesis of Cyclooctatin

Dotted arrows represent the CotB2-catalyzed GGDP cyclase reaction.

Cosmid Library Construction, Screening, and Sequencing

Total DNA from MI614-43F2, which was prepared as described previously (Kieser et al., 2000), was partially digested with *Sau*3AI, followed by separation using agarose gel electrophoresis. DNA fragments larger than 20 kb were ligated to a BamHI- and phosphatase-treated pOJ446 cosmid vector to give a cosmid library of MI614-43F2. This cosmid library was screened by colony hybridization with DNA fragments containing the partial GGDP synthase gene mentioned above as a probe. A positive cosmid clone, pCOT9469, was sequenced by the shotgun method (Genotech, Inc.) and annotated with position-specific iterated BLAST and FRAMEPLOT 4.0 (Ishikawa and Hotta, 1999).

Construction of pCOT Plasmids

PCR amplification using pCOT9469 and oligonucleotides for ligation into the *Streptomyces-E. coli* shuttle vector pSE101 was carried out with the forward primer, pCOT-N (5′-GGATATCTAGATTTCGACGTGGAGGAGTCAG-3′ [XbaI site underlined]), and the reverse primer, pCOT-C1 (5′-GCCCAAGCTTCAGCGTGGTTCCGGAATCATCC-3′ [HindIII site underlined]), to generate pCOT104 containing CotB1–4. The resulting pCOT104 was digested with *Pfl*MI (New England Biolabs), whose recognition sites were in the targeted *cotB3* gene, and then self-ligated to construct pCOT104Δ. This deletion plasmid therefore lacked only *cotB3*. Construction of pCOT103 containing CotB1–3 was done using the forward primer, pCOT-N, and the reverse primer, pCOT-C2 (5′-CGGGAAAGCTTCAGCGCGGCTCGCACACCATGG-3′ [HindIII site underlined]) in the same manner as for pCOT104. Construction of pCOT102 containing CotB1–2 was also done using the forward primer, pCOT-N, and the reverse primer, pCOT-C3 (5′-CCCCAAGCTTTCACTGGATGCGAGAGTTGACG-3′ [HindIII site underlined]) by using the same protocol as above. Primers used in this study are shown in Figure S3.

Identification of Cyclooctatin and the Biosynthetic Intermediates

Cyclooctatin and the intermediates were identified using a Thermo Fisher Accela HPLC system linked to a Thermo Fisher Scientific LTQ Orbitrap XL mass spectrometer. Compounds were loaded onto and separated on a CAPCELL PAK C18 UG120 column (2.0 ϕ × 150 mm, 5 μm, column temperature 40°C; Shiseido), using a flow rate of 0.2 ml/min with a linear solvent gradient of 50%–100% acetonitrile in water (with 0.01% trifluoroacetic acid) over a 20 min period. Mass spectrometry analysis was performed by electrospray ionization in positive ion mode. The structures of the intermediates were analyzed from their mass and NMR spectral data (600 MHz; JEOL ECA-600). Optical rotations were measured with a JASCO model P-1030 polarimeter.

Functional Analysis for Diterpene Cyclase CotB2

PCR amplification using pCOT9469 and oligonucleotides for ligation into the *E. coli* expression vector pHIS8 (Jez et al., 2000) was carried out with the forward primer 5′-GAGGAATTCATGACGACAGGACTTTC (EcoRI site underlined) and the reverse primer 5′-CCCCAAGCTTTCACTGGATGCGAG (HindIII site underlined) to generate pHIS8CotB2. The pHIS8CotB2 construct was transformed into *E. coli* BL21(DE3) and the recombinant CotB2 protein was purified as previously described (Jez et al., 2000). The cyclase reaction mixture (500 μl) contained 50 mM Tris-HCl (pH 7.5), 1 mM $MgCl_2$, 50 μM GGDP, farnesyl diphosphate or geranyl diphosphate (Davisson et al., 1985), and 10 μg/ml purified CotB2 protein. After incubation of the mixture at 30°C for 3 hr, 500 μl ethyl acetate was added to extract the reaction products. The organic extract was then evaporated under vacuum to dryness and the residue was reconstituted with 100 μl of methanol in preparation for HPLC analysis.

Purification of Cyclooctat-9-en-7-ol (2)

Large-scale production of the CotB2-catalyzed reaction product **2** in a total volume of 200 ml was carried out as described above for the enzyme assay. The reaction mixtures were incubated at 30°C overnight and then extracted twice with 200 ml ethyl acetate. After drying over Na_2SO_4, the organic layer was evaporated in vacuo to yield 3.0 mg of sufficiently pure oily product **2**. Optical rotation of **2**, $[\alpha]_D^{22}$ + 82.2 ± 0.5° (*c* 0.47, CH_3OH).

Purification of Cyclooctat-9-en-5,7-diol (3)

The supernatant from the entire *S. albus*/pCOT103 culture broth (10 liters) was extracted twice with ethyl acetate. After drying over Na_2SO_4, the organic layer was evaporated in vacuo to yield 3 g of oily extract. The extract was applied to a silica gel column (20 × 150 mm; Wako Pure Chemical Industries) in chloroform and eluted successively with 100% chloroform and chloroform/methanol mixtures. Each fraction was analyzed by LC-MS as described above. Compound **3** was eluted with 50:1 chloroform/methanol. Further purification was achieved by preparative HPLC with PEGASIL ODS (20 × 250 mm; Senshu Chemical), using an isocratic elution of 90% methanol at a flow rate of 8 ml/min and monitoring at 203 nm. **3** (4.2 mg) was obtained as an oil. Optical rotation of **3**, $[\alpha]_D^{22}$ + 40.7 ± 1.6° (*c* 0.10, CH_3OH).

ACCESSION NUMBERS

The nucleotide sequence of the 5 kb DNA fragment containing the cyclooctatin biosynthetic gene cluster has been deposited in the DDBJ/EMBL/GenBank nucleotide sequence database and assigned the accession number AB448947.

SUPPLEMENTAL DATA

Supplemental Data include ten figures and two tables and can be found with this article online at http://www.cell.com/chemistry-biology/supplemental/S1074-5521(09)00205-1.

ACKNOWLEDGMENTS

We are grateful to Bradley S. Moore of the University of California, San Diego, for providing *S. albus* and pOJ446. Also, we are grateful to Tohru Dairi of the Toyama Prefectural University for pSE101 and for fruitful discussions about the reaction mechanism of CotB2. This work was supported by a Grant-in-Aid for Scientific Research (21380071 to T.K.) from the Japan Society for the Promotion of Sciences.

Received: January 27, 2009
Revised: June 9, 2009
Accepted: June 16, 2009
Published: July 30, 2009

REFERENCES

Aoyagi, T., Aoyama, T., Kojima, F., Hattori, S., Honma, Y., Hamada, M., and Takeuchi, T. (1992). Cyclooctatin, a new inhibitor of lysophospholipase, produced by *Streptomyces melanosporofaciens* MI614-43F2. Taxonomy, production, isolation, physico-chemical properties and biological activities. J. Antibiot. (Tokyo) *45*, 1587–1591.

Aoyama, T., Naganawa, H., Muraoka, Y., Aoyagi, T., and Takeuchi, T. (1992). The structure of cyclooctatin, a new inhibitor of lysophospholipase. J. Antibiot. (Tokyo) *45*, 1703–1704.

Bierman, M., Logan, R., O'Brien, K., Seno, E.T., Rao, R.N., and Schoner, B.E. (1992). Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces spp*. Gene *116*, 43–49.

Caruthers, J.M., Kang, I., Rynkiewicz, M.J., Cane, D.E., and Christianson, D.W. (2000). Crystal structure determination of aristolochene synthase from the blue cheese mold, *Penicillium roqueforti*. J. Biol. Chem. *275*, 25533–25539.

Christianson, D.W. (2006). Structural biology and chemistry of the terpenoid cyclases. Chem. Rev. *106*, 3412–3442.

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. *16*, 10881–10890.

Dairi, T., Aisaka, K., Katsumata, R., and Hasegawa, M. (1995). A self-defense gene homologous to tetracycline effluxing gene essential for antibiotic production in *Streptomyces aureofaciens*. Biosci. Biotechnol. Biochem. *59*, 1835–1841.

Dairi, T., Hamano, Y., Kuzuyama, T., Itoh, N., Furihata, K., and Seto, H. (2001). Eubacterial diterpene cyclase genes essential for production of the isoprenoid antibiotic terpentecin. J. Bacteriol. *183*, 6085–6094.

Davisson, V.J., Woodside, A.B., and Poulter, C.D. (1985). Synthesis of allylic and homoallylic isoprenoid pyrophosphates. Methods Enzymol. *110*, 130–144.

Hamano, Y., Kuzuyama, T., Itoh, N., Furihata, K., Seto, H., and Dairi, T. (2002). Functional analysis of eubacterial diterpene cyclases responsible for biosynthesis of a diterpene antibiotic, terpentecin. J. Biol. Chem. *277*, 37098–37104.

Ishikawa, J., and Hotta, K. (1999). FramePlot: a new implementation of the frame analysis for predicting protein-coding regions in bacterial DNA with a high G + C content. FEMS Microbiol. Lett. *174*, 251–253.

Jez, J.M., Ferrer, J.L., Bowman, M.E., Dixon, R.A., and Noel, J.P. (2000). Dissection of malonyl-coenzyme A decarboxylation from polyketide formation in the reaction mechanism of a plant polyketide synthase. Biochemistry *39*, 890–902.

Kawasaki, T., Kuzuyama, T., Kuwamori, Y., Matsuura, N., Itoh, N., Furihata, K., Seto, H., and Dairi, T. (2004). Presence of copalyl diphosphate synthase gene in an actinomycete possessing the mevalonate pathway. J. Antibiot. (Tokyo) *57*, 739–747.

Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F., and Hopwood, D.A. (2000). Practical *Streptomyces* Genetics (Norwich, UK: John Innes Foundation).

Lesburg, C.A., Zhai, G., Cane, D.E., and Christianson, D.W. (1997). Crystal structure of pentalenene synthase: mechanistic insights on terpenoid cyclization reactions in biology. Science *277*, 1820–1824.

MacMillan, J., and Beale, M.H. (1999). Diterpene biosynthesis. In Comprehensive Natural Products Chemistry (vol. 2), Isoprenoids Including Carotenoids and Steroids, D.E. Cane, ed. (Amsterdam, Netherlands: Elsevier), pp. 217–243.

Nakano, C., Okamura, T., Sato, T., Dairi, T., and Hoshino, T. (2005). *Mycobacterium tuberculosis* H37Rv3377c encodes the diterpene cyclase for producing the halimane skeleton. Chem. Commun. (Camb.) *8*, 1016–1018.

Poirot, O., O'Toole, E., and Notredame, C. (2003). Tcoffee@igs: A web server for computing, evaluating and combining multiple sequence alignments. Nucleic Acids Res. *31*, 3503–3506.

Shishova, E.Y., Di Costanzo, L., Cane, D.E., and Christianson, D.W. (2007). X-ray crystal structure of aristolochene synthase from *Aspergillus terreus* and evolution of templates for the cyclization of farnesyl diphosphate. Biochemistry *46*, 1941–1951.

Starks, C.M., Back, K., Chappell, J., and Noel, J.P. (1997). Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-aristolochene synthase. Science *277*, 1815–1820.

Toyomasu, T., Tsukahara, M., Kaneko, A., Niida, R., Mitsuhashi, W., Dairi, T., Kato, N., and Sassa, T. (2007). Fusicoccins are biosynthesized by an unusual chimera diterpene synthase in fungi. Proc. Natl. Acad. Sci. USA *104*, 3084–3088.