

## Genome Mining in *Streptomyces coelicolor*: Molecular Cloning and Characterization of a New Sesquiterpene Synthase

Xin Lin, Russell Hopson, and David E. Cane\*

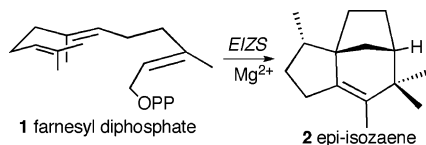
Department of Chemistry, Brown University, Box H, Providence, Rhode Island 02912-9108

Received February 22, 2006; E-mail: David\_Cane@brown.edu

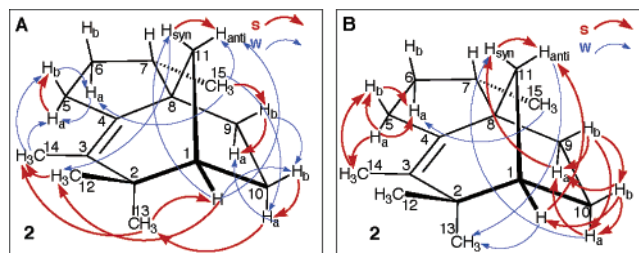
*Streptomyces* are Gram-positive, filamentous, soil-dwelling saprophytic bacteria that are the source of a vast array of bioactive secondary metabolites, including antibiotics, immunosuppressants, and anticancer agents.<sup>1</sup> Among the thousands of such natural products are only a relative handful of cyclic terpenoids. For example the antibiotic pentalenolactone, which is produced by a large number of *Streptomyces* species, is derived from the sesquiterpene hydrocarbon, pentalenene.<sup>2</sup> Pentalenene itself is formed by the cyclization of farnesyl diphosphate (1, FPP), the universal acyclic precursor of more than 300 distinct sesquiterpenes. Pentalenene synthase from *Streptomyces* sp. UC5319 has been cloned and expressed and the structure determined by X-ray crystallography.<sup>3a</sup> We have also recently cloned and expressed a closely related pentalenene synthase from *S. avermitilis*.<sup>3b</sup>

*Streptomyces coelicolor* A3(2) is genetically the most thoroughly studied member of the genus.<sup>4</sup> The 8.7-Mb linear chromosome harbors 7825 open-reading frames (ORFs), among which at least three show significant levels of similarity to pentalenene synthase.<sup>5</sup> The 2-kb SCO6073 (SC9B1.20) encodes a protein of 726 amino acids that catalyzes the cyclization of FPP to a mixture of germacadienol and germacrene D, implicated in the biosynthesis of geosmin, a degraded sesquiterpene that is responsible for the characteristic odor of soil.<sup>6</sup> We now report the molecular cloning and biochemical characterization of a second *S. coelicolor* terpene synthase, SCO5222 (SC7E4.19), that is shown to catalyze the Mg<sup>2+</sup>-dependent cyclization of FPP (1) to a novel sesquiterpene, epi-isozaena (2) (Scheme 1).

### Scheme 1. Cyclization of Farnesyl Diphosphate (1) to Epi-isozaena (2)



The 361 amino acid (aa) SCO5222 protein has 23.8% amino acid sequence identity and 51.1% similarity over 311 aa with the 336-aa pentalenene synthase of *Streptomyces* UC5319. Importantly, the protein encoded by SCO5222 includes the two universally conserved Mg<sup>2+</sup>-binding domains, the aspartate-rich motif **DDRHD**, beginning at Asp99, and the characteristic triad **NDLCSLPKE**, at Asn240. We used PCR to amplify the 1086-bp ORF of SCO5222 *S. coelicolor* cosmid SC7E4 DNA as template, while introducing *Nde*I and *Xho*I restrictions sites at the 5'- and 3'-ends, respectively. The doubly digested DNA amplicon was ligated into the corresponding sites of the expression vector pET28a(+), and the ligation mixture was used to transform competent cells of *Escherichia coli* XL1-Blue. The resultant plasmid, pET28a/SCO5222 was introduced into the T7-RNA polymerase-based expression host, *E. coli* BL21-(DE3), which was grown to an OD<sub>600</sub> of 0.5 and then induced with 0.1 mM IPTG. After further incubation at 20 °C for 18 h, the



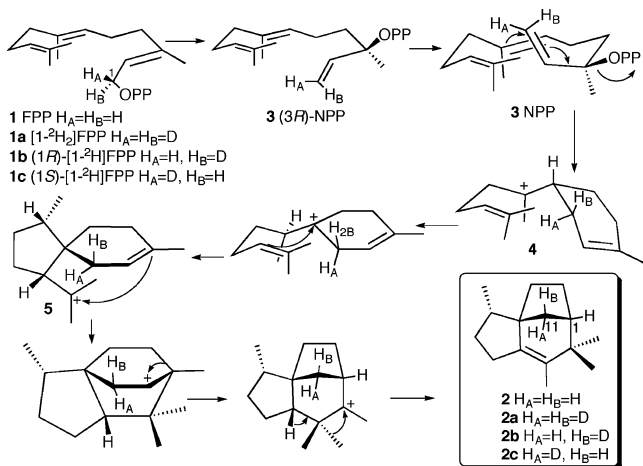
**Figure 1.** NOESY (A) and COSY (B) correlations for epi-isozaena (2). Correlations with H-6b ( $\delta$  1.77) and with H-7 ( $\delta$  1.77) are not assigned due to signal overlap. s, strong; w, weak.

resultant soluble N-terminal His<sub>6</sub>-tag protein was harvested and purified to homogeneity by Ni<sup>2+</sup>-affinity chromatography to yield 13 mg protein/L culture with *M<sub>D</sub>* (MALDI-TOF) 43665 ± 22 (calcd His<sub>6</sub>-tag-SCO5222 43650). Incubation of the purified SCO5222 protein with FPP (1) and analysis by GC-MS revealed the formation of a single sesquiterpene hydrocarbon product, C<sub>15</sub>H<sub>24</sub>, *m/z* 204.

The steady-state kinetic parameters were determined by carrying out a series of 15-min incubations at 30 °C in pH 6.5 PIPES using 83 nM to 2.49 μM [1-<sup>3</sup>H]-FPP (164 mCi/mmol), giving a *k<sub>cat</sub>* of 0.049 ± 0.001 s<sup>-1</sup> and a *K<sub>m</sub>* (FPP) of 147 ± 14 nM.<sup>7</sup>

Preparative-scale (500-mL) incubations were carried out for 18 h at 30 °C using 12.6 mg (30 μmol) of FPP and a total of 0.6 μmol of recombinant SCO5222 protein, added in equal batches at 0, 6, and 12 h. The hydrocarbon product was extracted into HPLC-grade pentane and purified by SiO<sub>2</sub> chromatography, yielding 3–4 mg of 2 that was >98% pure by GC-MS. The structure and stereochemistry of the enzymatic cyclization product were determined by a combination of <sup>1</sup>H, <sup>13</sup>C, <sup>1</sup>H-<sup>1</sup>H COSY, NOESY, HMQC, and HMBC NMR (Figure 1). The presence of a fully substituted double bond in 2 was indicated by the observation of two quaternary olefinic carbons at 127.4 (C-3) and 143.0 ppm (C-4) in the 75.48 MHz <sup>13</sup>C NMR spectrum and the absence of any signals due to olefinic protons in the 300 MHz <sup>1</sup>H NMR spectrum. Four methyl groups were readily identified: a geminal methyl pair (H-12,  $\delta$  0.99; C-12, 28.3 ppm and H-13,  $\delta$  0.96; C-13, 25.0 ppm) attached to a quaternary carbon (C-2, 40.4 ppm); an allylic methyl (H-14,  $\delta$  1.42; C-14, 12.8 ppm) attached to the olefinic C-3 carbon at 127.4 ppm; and a secondary methyl (H-15,  $\delta$  0.90, *d*, *J* = 6.3 Hz; C-15, 14.0 ppm) attached to a C-H (H-7,  $\delta$  1.77, *m*; C-7, 39.7 ppm). The <sup>1</sup>H-<sup>1</sup>H COSY spectrum revealed long-range couplings between the allylic H-14 methyl and each of the allylic H-5 methylene protons at  $\delta$  2.06 and 2.2, both of which were correlated with C-5 (27.3 ppm). The HMBC spectrum also revealed cross-peaks between each of the geminal methyl protons and the olefinic C-3 signal at 127.4 ppm as well as a second set of cross-peaks with the C-1 methine (47.1 ppm; H-1,  $\delta$  1.81). The secondary methyl (H-15) also showed a three-bond HMBC cross-peak to C-8 (52.6 ppm). The remaining four methylene carbons each carried a

**Scheme 2.** Mechanism of the Cyclization of Farnesyl Diphosphate (1) to Epi-isozizaene (2)



pair of diastereotopic protons, which were identified by their characteristic sets of cross-peaks in the HMQC, COSY, and NOESY spectra and were assigned on the basis of the observed coupling constants and the NOESY spectrum.<sup>8</sup> The configuration of the secondary methyl group in **2** was unambiguously assigned as anti to the C-11 methylene bridge, based on the strong 600 MHz NOESY cross-peak between H-15 ( $\delta$  0.90) and H-9b ( $\delta$  1.35). A weaker NOESY cross-peak was also observed between H-15 and H-11<sub>anti</sub> ( $\delta$  1.47), as well as a faint correlation to H-9a ( $\delta$  1.16). On the other hand, there was no detectable NOE interaction between H-15 and H-11<sub>syn</sub> ( $\delta$  1.39), which itself was unambiguously assigned on the basis of its NOESY cross-peak with the H-12 *exo*-methyl singlet at  $\delta$  0.99.

Although the enzymatic cyclization product **2** has not previously been isolated as a natural product, the structure and relative stereochemistry of **2** correspond to the known compound (–)-epi-isozizaene, originally synthesized by formic acid-catalyzed rearrangement of (+)-zizaene.<sup>9,10</sup> Enzymatically generated (+)-(1*S*,7*S*,8*R*)-epi-isozizaene is enantiomeric to (–)-**2** prepared from (+)-zizaene, based on the  $[\alpha]_D^{21} +21^\circ$  ( $\text{CHCl}_3$ ,  $c = 0.10$ ) ((–)-**2**,  $[\alpha]_D -17^\circ$  ( $\text{CHCl}_3$ ,  $c = 0.10$ )<sup>9a</sup>).

The formation of epi-isozizaene by cyclization of FPP can be readily explained by the mechanism illustrated in Scheme 2. Ionization and isomerization of FPP (**1**) will give (3*R*)-nerolidyl diphosphate (NPP, **3**). Rotation about the newly generated C-2/C-3 bond generates the corresponding *cisoid* NPP conformer which can undergo ionization and cyclization to form bisabolyl cation **4**. Following 1,2-hydride shift and spirocyclization, the resultant acorenyl cation **5** can undergo further cyclization, ring contraction, methyl migration, and deprotonation to yield (+)-epi-isozizaene (**2**). The proposed mechanism was tested by incubation of individual samples of  $[1,1-^2H_2]$ -FPP (**1a**), (1*R*)- $[1-^2H]$ -FPP (**1b**), and (1*S*)- $[1-^2H]$ -FPP (**1c**),<sup>2</sup> with recombinant epi-isozizaene synthase and NMR analysis of the resultant labeled **2**. Epi-isozizaene (**2a**) derived from  $[1,1-^2H_2]$ -FPP (**1a**) lacked the characteristic H-11 signals at  $\delta$  1.39 and 1.47, as well as both the corresponding pair of cross-peaks in the COSY spectrum. Similarly, both of the cross-peaks to C-11 (36.9 ppm) were absent from the HMQC spectrum. By contrast, cyclization of (1*R*)- $[1-^2H]$ -FPP (**1b**) gave **2b** lacking H-11<sub>anti</sub> ( $\delta$  1.47) while displaying a broad singlet for H-11<sub>syn</sub> shifted upfield to  $\delta$  1.36 due to the isotope effect of the geminal deuterium. **2b** also had a single H-11<sub>syn</sub>/C-11 cross-peak in the HSQC spectrum, shifted upfield 0.4 ppm in the <sup>13</sup>C dimension to 36.5 ppm by the

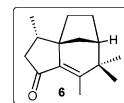
$\{^2H^{13}C\}$  isotope effect.<sup>11</sup> Similarly, the sample of **2c** derived from (1*S*)- $[1-^2H]$ -FPP (**1c**) lacked H-11<sub>syn</sub>, while displaying H-11<sub>anti</sub> as a broadened doublet ( $J = 5.3$  Hz,  $\delta$  1.45). The COSY spectrum of **2c** showed the expected H-11<sub>anti</sub>/H-1 cross-peak, while the HSQC spectrum retained a single cross-peak between H-11<sub>anti</sub> and C-11 (36.3 ppm). The combined labeling results are completely consistent with the stereochemical model illustrated in Scheme 2, confirming the expected net retention of configuration in the displacement of the pyrophosphate group of FPP that reflects the intermediacy of (3*R*)-NPP (**3**).

**Acknowledgment.** We thank Dr. Tun-Li Shen for assistance in obtaining the mass spectra, Prof. Keith Chater of the John Innes Centre for a gift of *S. coelicolor* cosmid SC7E4, Dr. Gregory L. Challis of the University of Warwick for providing *S. coelicolor* M145, and Prof. Niels H. Andersen of the University of Washington for providing the original NMR and IR spectra of (–)-epi-isozizaene and (+)-isozizaene. This work was supported by NIH Grant GM30301 to D.E.C.

**Supporting Information Available:** Sequence alignments, expression plasmid construction, protein expression and assay, NMR spectroscopic data for **2**, including labeling by deuterated FPP, and full citation for ref 5. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## References

- Hodgson, D. A. *Adv. Microb. Physiol.* **2000**, *42*, 47–238.
- Cane, D. E.; Oliver, J. S.; Harrison, P. H. M.; Abell, C.; Hubbard, B. R.; Kane, C. T.; Lattman, R. *J. Am. Chem. Soc.* **1990**, *112*, 4513–4524.
- (a) Cane, D. E.; Sohng, J.-K.; Lamberson, C. R.; Rudnicki, S. M.; Wu, Z.; Lloyd, M. D.; Oliver, J. S.; Hubbard, B. R. *Biochemistry* **1994**, *33*, 5846–5857; Lesburg, C. A.; Zhai, G.; Cane, D. E.; Christianson, D. W. *Science* **1997**, *277*, 1820–1824. (b) Tetzlaff, C. N.; You, Z.; Cane, D. E.; Takamatsu, S.; Omura, S.; Ikeda, H. *Biochemistry* **2006**. In press.
- Hopwood, D. A. *Microbiology* **1999**, *145*, 2183–2202.
- Hopwood, D. A.; et al. *Nature* **2002**, *417*, 141–147.
- (a) Cane, D. E.; Watt, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 1547–1551; He, X.; Cane, D. E. *J. Am. Chem. Soc.* **2004**, *126*, 2678–2679. (b) Gust, B.; Challis, G. L.; Fowler, K.; Kieser, T.; Chater, K. F. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 1541–1546.
- Substitution of 10 mM  $Mn^{2+}$  or  $Fe^{3+}$  for  $Mg^{2+}$  gave relative rates of formation of **2** of 72% and 85%, respectively. Other divalent cations were significantly less effective:  $Zn^{2+}$  (9%),  $Fe^{2+}$  (8%),  $Co^{2+}$  (4%),  $Cu^{2+}$  (4%), and  $Ni^{2+}$  (1%).
- The observed coupling constants were fully consistent with the conformation of **2** calculated using MM+ with HyperChem 7.5 and HyperNMR. See Supporting Information for detailed NMR assignments and spectra.
- (a) Andersen, N. H.; Smith, S. E.; Ohta, Y. *J. Chem. Soc., Chem. Commun.* **1973**, 447–448. The <sup>1</sup>H NMR spectrum of enzymatically generated **2** matched that of synthetic **2**, and was readily distinguished from that of isozizaene. See Supporting Information for details. (b) Rearrangement of (–)- $\alpha$ -cedrene in superacid also gives (–)-epi-isozizaene, whose reported <sup>1</sup>H and <sup>13</sup>C NMR spectra are identical to those determined for enzymatically generated **2**. (Polovinka, M. P.; Korzhagina, D. V.; Shcherbukhin, V. V.; Gatilov, Y. V.; Rybalova, T. V.; Zefirov, N. S.; Barkhash, V. A. *Tetrahedron Lett.* **1995**, *36*, 8093–8096.) Because Andersen had shown, however, that the secondary methyl group of epi-isozizaene can epimerize under strong acid conditions, the configuration of the C-15 methyl group could not be assumed a priori to be the same as that in (–)- $\alpha$ -cedrene.
- Although epi-isozizaene (**2**) has not been previously reported as a natural product, the corresponding 5-keto derivative, albaflavenone (**6**), of unassigned absolute configuration, has recently been isolated from a species of *S. albidoflavus* (Gurtler, H.; Pedersen, R.; Anthoni, U.; Christophersen, C.; Nielsen, P. H.; Wellington, E. M. H.; Pedersen, C.; Bock, K. *J. Antibiot.* **1994**, *47*, 434–439.) Albaflavenone is reported to have antibiotic activity and an “earthy, camphor-like” odor. We have as yet been unable to detect either **2** or albaflavenone by GC–MS in the volatile extracts of either submerged or surface cultures of *S. coelicolor* M145.



- Garson, M. J.; Staunton, J. *Chem. Soc. Rev.* **1979**, 539–561.

JA061292S