Cloning and Functional Characterization of the Germacradienol Synthase (spterp13) from *Streptomyces peucetius* ATCC 27952

Ghimire, Gopal Prasad¹, Tae-Jin Oh¹, Hei Chan Lee¹, Byung-Gee Kim², and Jae Kyung Sohn¹*

¹Institute of Biomolecule Reconstruction (IBR), Department of Pharmaceutical Engineering, SunMoon University, Asan 336-708, Korea
²School of Chemical and Biological Engineering, Seoul National University, Seoul 151-742, Korea

Received: December 21, 2007 / Accepted: January 30, 2008

Sequence analysis of the metabolically rich genome of *Streptomyces peucetius* ATCC 27952 revealed a 2,199 bp sesquiterpene alcohol (germacradienol) synthase-encoding gene from the germacradienol synthase/terpene cyclase gene cluster. The gene was named spterp13, and its putative function is as a germacradienol synthase/terpene cyclase. The amino acid sequence of Spterp13 shows 66% identity with SAV2163 (GeoA) from *S. avermitilis* MA-4680 and 65% identity with SCO6073 from *S. coelicolor* A3(2), which produces germacradienol/geoamin. The full-length recombinant protein was heterologously expressed as a his-tagged fusion protein in *Escherichia coli*, purified, and shown to catalyze the Mg²⁺-dependent conversion of farnesyl diphosphate to the germacradienol, which was verified by gas chromatography/mass spectrometry.

**Keywords:** Geosmin, germacradienol, heterologous expression, sesquiterpene cyclase, *Streptomyces peucetius*.

Terpenoid or isoprenoid metabolites represent one of the largest and most diverse groups of natural products that have physiological and commercial importance, which include antibiotics, hormones, antigenic agents, flavor and odor constituents, and pigments [12, 24, 25]. Over 20,000 terpenoids have been reported, and most of them have been isolated from plants or fungi, and very few of them have been isolated from prokaryotes. Terpene synthases (also known as terpene cyclases) are responsible for the formation of over 20,000 different products that stem from linear isoprenoid precursors [14]. Although *Streptomyces* are a rich source of secondary metabolites of all structural classes, no more than a dozen terpenoid metabolites have been obtained from them [2, 3]. Only a few of the thousands of useful natural products produced by *Streptomyces* are cyclic terpenoids [19]. Many streptomycetes produce earthy or musty odors [23].

The characteristic odor of moist soil is due to the presence of two terpene derivatives, 2-methylisoborneol and geosmin [5, 7]. Geosmin is a well-known odorous metabolite that is produced by a wide variety of microorganisms, including the genera actinomycetes, myxobacteria, cyanobacteria, fungi, and by higher plants such as liverworts and red beets [20]. Geosmin has an exceptionally low detection threshold (on the order of 10 to 100 parts per trillion), and is associated with the characteristic odor of freshly turned earth and with the unpleasant off-flavors in water, wine, fish, and other foodstuffs [6, 15]. Farnesyl diphosphate is the immediate precursor of cyclic sesquiterpenes [8]. There are several biosynthetic pathways of isoprenoid synthesis in microorganisms, one or more of which may lead to the production of germacradienol and geosmin [4, 11] (Fig. 1).

**Fig. 1.** Simplified biosynthetic scheme for the formation of 2-methylisoborneol (MIB) and geosmin in streptomycetes and myxobacteria.

---

*Corresponding author
Phone: 82-41-536-2246; Fax: 82-41-544-2919; E-mail: sohn@sunmoon.ac.kr
We recently determined the genome sequence of the 8.7 Mb genome of Streptomyces peucetius ATCC 27952, which produces clinically important anthracycline chemotherapeutic agents of the polyketide class of antibiotics, daunorubicin and doxorubicin [21]. From the genome mining of S. peucetius, we found one open reading frame of 2,199 bp, named spterp13, which encodes a putative protein of 732 aa with significant similarity to S. avermitilis MA-4680 (SAV2163, GeoA) and S. coelicolor A3(2) (SCO6073) [1]. We expect that Spterp13 acts as a germacradienol/geosmin synthase because sequence alignment of the deduced amino acids of the Spterp13 protein with the known germacradienol/geosmin synthases provides strong bootstrap support for heterologous expression. To investigate the function of spterp13, the spterp13 gene was cloned and heterologously expressed in Escherichia coli, and an in vitro enzyme assay was carried out. Our results demonstrate that Spterp13 is a key enzyme for the cyclization of farnesyl diphosphate (FPP) to germacradienol (Fig. 2).

**Materials and Methods**

**Bacterial Strains, Media, and Culture Conditions**

S. peucetius ATCC 27952 was grown in 50 ml of R2YE medium at 28°C in a shaking incubator for 5 days. E. coli XL1-Blue MRF and E. coli BL21 (DE3) (Stratagene, La Jolla, CA, U.S.A.) were used as hosts for plasmid preparation and expression, respectively [18, 22]. pET32a(+) (Novagen, U.S.A.) was used as an expression vector [17]. E. coli strains were grown in Luria Bertani (LB) medium, both in liquid culture and agar plates, with appropriate antibiotics for selection. Synthetic oligonucleotides were synthesized by GenoTech (Daejeon, Korea). A polymerase chain reaction (PCR) kit and enzymes were purchased from Takara (Japan). All chemicals were purchased from Sigma unless otherwise stated. Database searches were performed by using the BLAST, FASTA, and CLUSTALW programs.

**DNA Manipulation**

Genomic DNA was isolated from S. peucetius using the Kirby mix procedure [13]. The spterp13 (2.2 kb) was obtained using the primers TS-t (5'-CAT GAA TTC CGA AAG GCC CTG CCG CAC-3') and TS-r (5'-ATT AAG CTT ATC CCG GGC CAG CTC GGC-3'). PCR was carried out in a thermal cycler (Takara, Japan) under the following conditions: 94°C for 15 min, followed by 30 cycles of 94°C for 1 min, 62°C for 1 min, 72°C for 1 min, and finally 72°C for 7 min. The amplification mixture was prepared in a total volume of 20 µl containing 5 µl of PCR premix [Buffer (10 mM Tris-HCl, pH 9.0, 1.5 mM MgCl2, and 40 mM KCl)], 1 mM dNTP, and Taq DNA polymerase] with dimethyl sulfoxide (DMSO).

All of the PCR products were extracted, purified, and ligated into the pGEM-T easy vector (Promega, U.S.A.). The ligated mixture was transformed into E. coli XL1-Blue MRF by heat pulse transformation. Positive clones containing the recombinant plasmid DNA were isolated and sequenced prior to cloning into expression vectors to verify that no mutation had been introduced during PCR amplification. A fragment containing EcoRI-HindIII was recovered from one such clone and then subcloned into the pET32a(+) vector (Novagen, U.S.A.) for construction of pGGG. The purified pGGG was then transformed into E. coli BL21 (DE3) for heterologous expression.

**Expression and Purification of Spterp13**

E. coli BL21 (DE3) harboring pGGG was grown in 3 ml of LB culture medium. For the expression of Spterp13, E. coli-pGGG was transferred from an overnight culture to 400 ml of fresh LB containing ampicillin (100 µg/ml) at 37°C and 250 rpm. At an OD600 of 0.6, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.7 mM, and the incubation was continued at 37°C for 8 h. The induced culture was harvested, and the cell pellet was resuspended in 40 ml of lysis buffer [50 mM Tris-HCl, 0.5 mM EDTA, 10% glycerol (v/v), pH 8.2]. Lysozyme (1 mg/ml) and β-mercaptoethanol were added to the cell suspension, which was then incubated at 32°C, with shaking at 100 rpm for 15 min, followed by the addition of MgCl2 (5 mM) and Triton X-100 (0.1%) after 10 min. All subsequent procedures were carried out at 4°C. The soluble fraction was collected by centrifugation at 10,000 rpm for 15 min and purified by Co2+-affinity chromatography (TALON Purification Kit, Clontech, Takara Bio Company, U.S.A.) according to the instructions supplied by the manufacturer. The purified fractions were desalted by dialysis and concentrated by Centricron with Ultracel YM-10. The purity of the recombinant protein was evaluated by 12% SDS-PAGE, as described by Laemmli [16].

**Germacradienol Synthase Assay**

The enzyme assay was performed at 30°C in 1 ml of buffer [50 mM Tris-HCl, pH 8.2, containing 10% glycerol (v/v), 10 mM MgCl2, 0.2 mM β-mercaptoethanol] and 800 nM FPP with 100 ng of purified protein. The mixture was immediately overlaid with 1.5 ml of HPLC-grade pentane for 3 h. The reaction mixtures were extracted twice with 2 ml of pentane/dichloromethane (5:1), and the combined organic layers were dried over MgSO4 (anhydrous), concentrated, filtered, and subjected to gas chromatography/mass spectrometry (GC/MS) analysis.

**GC/MS Analysis**

A 5-µl portion of the concentrated extract was analyzed by GC/MS (Shimadzu GC-17A, 70 eV, EI, positive-ion mode; 30 m×0.25 mm bond capillary column, 5% phenyl/methylsilicon), using a temperature program of 50 to 280°C, and a temperature gradient of 20°C per min). The incubation yielded two products: germacradienol (rt: 8.314 min, [M]+ m/z=222, base peak m/z=82) and a sesquiterpene hydrocarbon (rt: 8.322 min, [M]+ m/z=204, base peak m/z=107).

**Nucleotide Sequence Accession Number**

The sequence of spterp13 reported here has been deposited in the GenBank under the accession number ABY50951.
RESULTS AND DISCUSSION

Homology-based Structure and Function of Spterp13

The Spterp13 encodes a protein of 732 aa with a molecular mass of 80.63 kDa. The overall G+C content of the gene is 70%, which is characteristic of the genes in Streptomyces. Spterp13 shows high sequence similarity with the terpene cyclase genes in the database (data not shown). For example, it has the closest homology with SAV2163 (GenBank Accession No. NP_823339) of S. avermitilis MA-4680, having a 66% identity according to BLAST. It also exhibits 65% identity with SCO6073 (GenBank Accession No. NP_630182) from S. coelicolor A3(2), 60% identity with terpene synthase (GenBank Accession No. YP_483306) from Frankia sp. Ce13, and 59% identity from Saccharopolyspora erythraea NNRL 2338.

Several bacterial and fungal terpene synthases have been characterized to date and share many features such as a conserved aspartate-rich DDHFL motif with a high proportion of aromatic amino acids and a downstream NSF triad NDLSFYQRE, the conserved Mg$^{2+}$-binding domains of 330 to 400 aa in length, which correspond to a subunit Ml of 35 to 45 kDa [10]. The 732 aa Spterp13 of S. peuceus is unusual in the sense that it is about twice the size of a typical terpene cyclase. Notably, both the N- and C-terminal halves of the Spterp13 show significant sequence similarity to the germacradienol/geosmin cyclases, SAV2163 (725 aa) of S. avermitilis MA-4680 and SCO6073 (726 aa) of S. coelicolor A3(2). Like SAV2163 and SCO6073, Spterp13 shows variation of an aspartate-rich domain, with a DDHFL motif in the N-terminal half and an unusual DDYYP motif in the C-terminal half. Both halves also have NSF motifs: NDLSFYQRE and NDLSFYQKE with the repetition of an NSF motif, NDLLTSRVQQFE, in the N-terminal half. The organization of protein domains and conserved Mg$^{2+}$-binding motifs in Spterp13, SAV2163, and SCO6073 are also similar (Fig. 3). The active-site residues are marked by germacradienol synthase taking the references of SAV2163 (GenBank) from S. avermitilis MA-4680 and SCO6073 from S. coelicolor A3(2) [1, 10]. Therefore, based on aa alignment and conserved domains, we suggest that Spterp13 can serve as a germacradienol synthase, similar to SCO6073 and SAV2163.

Comparison of the phylogenetic tree for Spterp13 and various terpene cyclases showed a distinct relationship between Spterp13 and other germacradienol synthase/geosmin synthase/terpene cyclase (Fig. 4). The SAV2163 and SCO6073 proteins, which were already characterized as germacradienol/geosmin synthases, occupied one group, whereas the other terpene cyclases were separated into another group.

PCR Amplification, Overexpression, and Purification of Recombinant Spterp13

The 2,199 bp DNA fragment was amplified from the total DNA of S. peuceus by PCR. The amplified DNA was initially cloned into the pGEM-T easy vector. The fragment digested with EcoRI and HindIII was ligated with the pET32a(+) expression vector to prepare pGGGS. E. coli was transformed with pGGGS in order to express spterp13 as the 6×his-tagged fusion protein. The expression of E. coli BL21 (DE3)-pGGGS was optimized in order to obtain a soluble protein. The induction with IPTG (0.7 mM) at 37°C for 8 h led to the excessive formation of soluble proteins. The soluble fraction was purified, desalted, and concentrated as described in Materials and Methods. The molecular mass of the recombinant Spterp13 was approximately 98 kDa, which is consistent with the predicted molecular mass of Spterp13 plus his-tag determined by SDS-PAGE analysis (Fig. 5).
Characterization of Recombinant Spterp13

The protein encoded by Spterp13 was overexpressed in E. coli. To confirm that the purified synthase was indeed generating germacradienol, the recombinant protein was incubated with FPP in the presence of Mg\(^{2+}\) as described in the Materials and Methods section, and the resulting products were analyzed by GC/MS. The GC/MS gave two products: germacradienol (r.t. 8.314 min, [M\(^+\)] m/z=222, base peak m/z=82) as a major product, and a detectable amount of sesquiterpene hydrocarbon of unassigned structure (r.t. 8.322 min, [M\(^+\)] m/z=204, base peak m/z=107) (Figs. 6A and 6B). The structure of germacradienol was verified by comparison with the spectra of the corresponding reference compounds in the Automatic Mass Spectral Deconvolution and Identification System (AMDIS) program (Version 2.65) and also by comparing the GC/MS spectrum of germacradienol produced by S. coelicolor SCO6073 [9]. Since the same GC/MS was detected in the enzymatic conversion of FPP by SAV2163 from S. avermitilis MA-4680, we called it a sesquiterpene hydrocarbon by taking the reference of the Cane group [1].

We also tried to isolate compounds from the wild strain but no products that are biosynthetically related to germacradienol synthase were detected even though they were screened for secondary metabolite production. We assumed spterp13 to be a “silent” gene under the laboratory condition administrated. The combined evidence from sequence analysis and enzymatic conversion conclusively established that the spterp13 gene encodes a germacradienol synthase and that this enzyme catalyzes an essential step in the biosynthesis of geosmin. The N- and C-terminal domains of Spterp13 are similar to those of the SCO6073 and SAV2163 proteins. It has been reported that the S. coelicolor SCO6073 N-terminal domain catalyzes the Mg\(^{2+}\)-dependent cyclization of FPP to germacradienol/germacrene D, whereas the highly homologous C-terminal domain, previously thought to be catalytically silent, catalyzes the Mg\(^{2+}\)-dependent germacradienol/geosmin. Presently, the study in SCO6073 proved that the N- and C-terminal domains each harbor distinct and independently functioning active sites [10]. Although geosmin was not detected in the in vitro assay of the full-length recombinant Spterp13 protein, the formation of germacradienol seems to be a committed step in the formation of geosmin, the characteristic odoriferous degraded sesquiterpene of several Streptomyces species. Owing to its unusual size and repeated conserved motifs, further experiments are needed to establish the function of the individual halves between the N- and C-terminals of Spterp13.

Acknowledgments

The work was supported by the Basic Research Program of the Korea Science and Engineering Foundation, Grant No. R01-2006-000-10234-0, and by a medium-term strategic technology development grant from the Republic of Korea to J.K.S.
REFERENCES


