Published in final edited form as:
Environ Sci Technol. 2008 November 1; 42(21): 8027–8032.

Isolation and Characterization of the Gene Associated with Geosmin Production in Cyanobacteria

Steven Giglio[1,2,*], **Jiaoyang Jiang**[3], **Christopher P Saint**[1,4], **David Cane**[3], and **Paul T Monis**[1,2]

[1] Australian Water Quality Centre, SA Water Corporation, Salisbury, SA 5108, Australia

[2] Department of Pharmacy and Medical Sciences, University of South Australia, SA 5095, Australia

[3] Department of Chemistry, Box H, Brown University, Providence, Rhode Island 02912-9108, USA

[4] School of Earth and Environmental Sciences, The University of Adelaide, Adelaide, SA 5005, Australia

Abstract

Geosmin is a secondary metabolite responsible for earthy tastes and odors in potable water supplies. Geosmin continues to be a challenge to water utility management regimes and remains one of the most common causes of consumer complaints, as the taste of "dirty" water may suggest a failed disinfection regime and that the water may be unsafe to drink. Although cyanobacteria have been reported to be largely responsible for these taste and odor events, the answer as to how or why geosmin is produced has eluded researchers. We describe here for the first time the mechanism by which geosmin is produced in a model cyanobacterium, *Nostoc punctiforme* PCC 73102 (ATCC 29133), which we demonstrate utilizes a single enzyme to catalyze the cyclization of farnesyl diphosphate to geosmin. Using this information, we have developed a PCR-based assay that allows the rapid detection of geosmin-producing cyanobacteria. This test may be utilized to confirm and track the emergence of taste and odor-producing cyanobacteria in any given water body and thus can be used as an early warning system by managers of water bodies that may suffer from adverse taste and odor episodes.

INTRODUCTION

Globally, cyanobacteria are nuisance organisms in fresh water supplies and have been primarily responsible for taste and odor episodes as well as toxic incidents (1–4). Consumer taste and odor complaints to water utilities are high, and are second only to those concerning chlorinous taints. One of the commonest complaints, that of an "earthy" or musty odor, is largely a result of geosmin and/or 2 methylisoborneol (MIB) produced by cyanobacteria. The human taste and odor sensitivity threshold for geosmin is an extraordinarily low 10 ng/L (10 ppt) (5) and, although geosmin has no known adverse side effects, consumers associate the detection of this compound with poorly treated water that might be unsafe to drink. The detection and management of these geosmin-producing cyanobacteria is therefore of paramount importance to water utilities. Currently, the only options to deal with this issue are dosing water bodies with environmentally unfriendly chemicals such as copper sulfate in order to control algal blooms, as well as the use of powdered activated carbon in water treatment plants to remove

*Corresponding author. Mailing address: Australian Water Quality Centre, PMB 3, Salisbury, South Australia, Australia 5108. Phone +618 82590312. Fax +61 8 82590228. steven.giglio@sawater.com.au.

Supporting Information available: This information is available free of charge via the Internet at http://pubs.acs.org.

the responsible taste and odor compounds. Recent work has also suggested that biological filtration using aged sand filters may represent a suitable alternative to remove these undesired compounds (6).

The volatile metabolite geosmin is produced by a wide range of microorganisms, including most species of *Streptomyces* and a variety of other actinomycetes, myxobacteria and cyanobacteria, as well as certain fungi and selected higher plants such as liverworts and beets. Although geosmin was first identified by Gerber in 1965 (7), the biochemical pathway of geosmin production remained obscure for many years. In 1981, Bentley first proposed that geosmin was likely a degraded sesquiterpene, based on the observed incorporation of labeled acetate by cultures *of Streptomyces antibioticus* (8). Within the last several years, however, the mechanism of geosmin biosynthesis in *Streptomyces* and myxobacteria (9–11) has been elucidated in detail. Despite earlier confusion as to how many enzymes were involved, and indeed if the mechanism of geosmin production was different in different organisms, it is now known that in *S. coelicolor* A3(2) a single 726-amino acid protein, encoded by the 2181-bp SCO6073 gene (*cyc2*) (12,13), catalyzes the Mg^{2+}-dependent cyclization of farnesyl diphosphate (FPP), the universal acyclic precursor of all sesquiterpenes, to a mixture of geosmin and the sesquiterpene alcohol germacradienol, accompanied by smaller amounts of the bicyclic hydrocarbons germacrene D and 8,10-dimethyl-1-octalin (14–17) (Figure 1). The *S. coelicolor* geosmin/germacradienol synthase is in fact a bifunctional enzyme in which both the N-terminal and C-terminal halves show a high degree of sequence similarity to the well-studied 336-amino acid sesquiterpene synthase, pentalenene synthase (16). Experiments with individually expressed recombinant proteins corresponding to the N-terminal and C-terminal domains have shown that the N-terminal half of the synthase catalyzes the cyclization of FPP to a 85:15 mixture of germacradienol and germacrene D, accompanied by traces of the octalin, while the C-terminal domain catalyzes the Mg^{2+}-dependent cyclization–fragmentation of germacradienol to geosmin, with release of the 2-propanol side chain as acetone (14,16). Site-directed mutagenesis experiments have confirmed that the N- and C-terminal domains each harbor catalytically independent active sites (16).

It has also been shown that the closely related GeoA (SAV2163) protein of *S. avermitilis* (78% sequence identity and 85% similarity) catalyzes the same biochemical reaction (18), while more than a dozen known or presumed geosmin synthases deduced from a variety of *Streptomyces*, *Frankia*, *Saccharopolyspora*, and myxobacterial genome sequences share correspondingly high levels of sequence conservation over all 730–740 amino acids (45–75% identity, 57–85% similarity). In all these proteins, the N-terminal region contains two strictly conserved motifs, a, **DD**HFL**E** sequence, typically 80–100 amino acids, from the N-terminus and a **N**D(L/I)F**S**Y(Q/E)R**E** motif approximately 140-amino acids downstream, corresponding to the universally conserved aspartate-rich DDXXD motif and NSE triad (**N/D**)DXX(**S/T**)XXX**E,** respectively, that are found in all sesquiterpene synthases and that are known to be involved in the binding of the essential cofactor Mg^{2+} (19,20). Similarly the C-terminal half of geosmin synthase has a canonical variant of the aspartate rich motif, **DD**YYP, as well as a downstream **N**D(V/I/L)F**S**YQK**E** variant of the NSE triad.

Following the initial biochemical characterization of the SCO6073 gene (13) and its implication in geosmin biosynthesis (12), Ludwig *et al* (21) reported the use of PCR to isolate homologous genes from an environmental geosmin-producing *Phormidium sp.* that were similar in sequence both SCO6073 and SAV2163. Although they demonstrated that these genes were expressed in the parent cyanobacteria and speculated on the possibility of a “geosmin operon”, they did not directly correlate the expression with geosmin production nor did they explicitly demonstrate that these genes were functionally responsible for geosmin production in the *Phormidium* isolate examined.

The *npun02003620* gene in the reported genome sequence of the cyanobacterium *Nostoc punctiforme* PCC 73102 (ATCC 29133), encodes a hypothetical protein ZP_00109187 with 55% amino acid sequence similarity to the N-terminal region of SCO6073, including the presence of the universally conserved DDHFLE and NDLFSYQRE motifs characteristic of this class of enzyme. The predicted protein, however, consists of only 630 amino acids, approximately 100 amino acids shorter than the SCO6073 protein or any other known or predicted geosmin synthase. Equally importantly, although the C-terminal half of ZP_00109187 also is predicted to harbour a typical DDYFP motif, the essential NSE triad is apparently absent.

We describe here experiments that provide a new foundation for our understanding of geosmin production in cyanobacteria. We first examine the model cyanobacterium *Nostoc punctiforme* PCC 73102 (ATCC 29133) and demonstrate that the hypothetical protein ZP_00109187 encoded by *npun02003620* is in fact a truncated protein that, while catalyzing the conversion of FPP to germacradienol, is incapable of supporting geosmin formation. We further demonstrate that the apparent truncation is the result of a single, but critical, sequencing error in the published DNA sequence and that the corrected open reading frame corresponds to a fully functional geosmin synthase, dubbed NPUNMOD, that is of similar length and sequence to the *S. coelicolor* SCO6073 enzyme and all other geosmin synthase proteins. Having established the identity and biochemical function of the *Nostoc* geosmin synthase gene, we then utilize this information to design a PCR-based diagnostic tool for the detection of geosmin-producing cyanobacteria. Collectively, these results provide the fundamental step forward for understanding taste and odor episodes and provide a powerful tool that can be used to a) predict whether an emerging cyanobacterial bloom will be a geosmin producer and assist in designing strategies to limit its effects in the short term; b) track the emergence of taste and odor producers in the longer term; and c) anticipate what trends may occur in the future.

EXPERIMENTAL

Culture of cyanobacteria and DNA extraction

Nostoc punctiforme PCC 73102 (ATCC 29133), a known geosmin producer, was maintained in ATCC medium #819 according to the provided product information sheet. The following geosmin- and 2-methylisoborneol (MIB)-producing isolates, confirmed by GC-MS, were a generous gift from G. Izaguirre: *Pseudoanabaena limnetica* (MIB producer), *Anabaena laxa* (geosmin producer), *Nostoc sp.* UTAH12–18b (geosmin producer), and *Phormidium calcicola* (geosmin and MIB producer). GC-MS of additional environmental isolates used in this study were performed using fresh cultures of cyanobacteria (not normalized for cell number). The above isolates were grown in BG-11 medium and maintained under standard conditions at 25°C. Before DNA extraction, cultures were subcultured on starch-casein agar for the detection of possible contaminating geosmin-producing actinomycete bacteria. All subcultures were negative for actinomycete organisms. DNA was extracted using the Qiagen DNA Mini Spin kit according to manufacturer's instructions, using 1 ml of cyanobacterial culture, with the addition of an overnight proteinase K incubation at 56°C.

Cloning and expression of npun020003620 and NPUNMOD proteins, incubation with farnesyl pyrophosphate, and GC-MS analysis

The DNA sequence corresponding to *npun020003620* was amplified using the primers npunstart1 (5′-ATTTTATCCATGGTTATGCAACCCTTTGAACTGCCAGAA-3′) and npunhalt1 (5′-TAATAACTCGAGTTATGGATTTCGCCCTCG-3′), while the full length natural NPUNMOD gene was amplified with primers npunstart1 and npunhalt4 (5′-TAATAACTCGAGTAATTGACCGAGTAATGAC-3′), inserting *Nco*I and *Xho*I restriction sites (underlined) for the npunstart1 and npunhalt primers respectively. Fragments were PCR-

amplified using proofreading Elongase Taq polymerase (Invitrogen) as described by the manufacturer, using 35 cycles consisting of 94°C for 30 s, 55°C for 30 s, and 68°C for 120 s, with a final hold at 4°C until needed. The products were digested with *Nco*I and *Xho*I (Promega, USA) and ligated into a similarly digested pET21d(+) expression vector (Novagen). The previously described protocols for propagation, expression, and purification of recombinant *S. coelicolor* geosmin synthase were followed for the ZP_00109187 and NPUNMOD proteins (12,14), except that transformants were induced with 1 mM IPTG for 2 h at 35°C. Successful transformants were sequenced using Big Dye Terminator sequencing. Recombinant ZP_00109187 protein, which did not carry a His_6-tag, was obtained in soluble form, while the initially generated recombinant NPUNMOD protein, carrying a C-terminal His_6-tag, was obtained as insoluble inclusion bodies that could not be resolubilized in active form despite the use of a wide variety of re-folding conditions. To remove the His_6-tag, a stop codon was introduced by site directed mutagenesis using the QuikChange® Site-Directed Mutagenesis Kit (Stratagene, USA) using the mutagenic primers NPM2_58-notag (5′-TTACTCGGTCAA**TGA**TTACTCGAGCAC-3′) and NPM2_58-notaga (5′-GTGCTCGAGTAA**TCA**TTGACCGAGTAA-3′) (stop codon in ***bold***, *Xho*I restriction sites underlined). The sequence of NPM2-58-pJJ32 was confirmed by resequencing (University of California, Davis Sequencing Facility, Davis, CA, USA). Protein expression from *E coli* BL21 (DE3)pLysS once again gave insoluble inclusion bodies that could not be properly solubilized under a variety of denaturation–refolding conditions. We therefore co-expressed the NPM2-58 protein along with the chaperone protein combination GroES/GroEL by co-transformation of the two plasmids, NPM2-58-pJJ32 (ampicillin-resistant) and pGro12 (kanamycin-resistant) into *E coli* BL21(DE3)pLysS, selecting a single colony displaying ampicillin-kanamycin-chloramphenicol multiresistance. Cells from this colony were used to inoculate LB broth containing 100 μg/ml ampicillin, 50 μg/ml kanamycin and 34 μg/ml chloramphenicol (LBAKC) and the overnight culture was transferred to 500 ml of LBAKC broth and incubated at 37 °C and 250 rpm for 2 h. Chaperone expression was induced by addition of arabinose to a final concentration of 4 mg/ml and the desired protein expression was then induced by addition of IPTG to a final concentration of 0.4 mM. Incubation was continued for an additional 18 h at 18 °C, 250 rpm. The cells were harvested by centrifugation and resuspended in 30 ml of lysis buffer (20 mM Tris-HCI, 10% glycerol, 0.1 mM DTT, pH 8.0). The cells were disrupted by sonication and the cell lysate was clarified by centrifugation. The supernatant was purified by 12% ammonium sulfate precipitation followed by purification on a 25-ml *n*-butyl-Sepharose column that had been pre-equilibrated in buffer A (0.5 M ammonium sulfate, 50 mM NaH_2P0_4, 0.1 mM DTT, pH 7.0). After loading of the supernatant onto the column, the resin was washed with 60 ml of buffer A followed by a 180- ml linear gradient from buffer A to buffer B (50 mM NaH_2P04, 0.1 mM DTT, pH 7.0). The purified protein eluted in buffer B as a 1:1 mixture with chaperone protein, as determined by SDS-PAGE. The apparent molecular weight (M_r) of 82 kDa of the desired protein is very close to the theoretical MW 85 kDa. Incubation of the purified protein, free of all contaminating proteins, with FPP and subsequent GC-MS analysis were performed using the procedures described by Jiang et al (15).

Geosmin synthase PCR, cloning of PCR products and sequencing

Geosmin synthase PCR (G-PCR1) mastermix consisted of 2.5 mM $MgCl_2$, 1X PCR buffer, 200 μM dNTP, 300 μM each of forward primer 250F (5′-TTCTTCGACGAYCACTTCC-3′) and reverse primer 971R (5′-CCCTYGTTCATGTARCGGC-3′), 5% dimethyl sulphoxide, 1 U of Platinum Taq DNA polymerase (Invitrogen), and 2 μl of extracted cyanobacterial DNA. Reactions were run on a Perkin Elmer 2400 thermal cycler with an initial denaturation step of 95°C for 5 min, followed by 55 cycles of 95°C for 30 s, 55°C for 30 s, and 72°C for 120 s, with a final extension step of 10 min at 72°C. Samples were then run on a 1% agarose gel with 10 μl of SYBR Safe (Invitrogen) for 45min at 80V and bands were visualized using a Dark reader transilluminator (Clare Chemical Research, Dolores, CO, USA). For selected sample

bands of expected size (743 bp), DNA was extracted using a Qiagen Qiaquick Gel extraction kit according to the manufacturer's instructions. The purified PCR product was then cloned into vector PCR 2.1 TOPO® according to the manufacturer's instructions (TOPO cloning kit, Invitrogen), and subsequently sequenced using Big Dye Terminator sequencing reactions.

A second geosmin synthase PCR (G-PCR2) was developed for real time PCR screening of other cyanobacterial DNA samples. Reactions were first optimized by standard thermal cycling as described above, and then 2.5 μM SYT09 (Invitrogen) was incorporated into the mastermix. Primers used for G-PCR2 were: 288AF (5′-AACGACCTGTTCTCCTA-3′), and 288AR (5′-GCTCGATCTCATGTGCC-3′), generating an amplicon of 288 bp. Real time PCR was performed using a Corbett Research Rotorgene 6000HRM (Corbett Research, Australia) under the same conditions as above, except that the final extension step was replaced with a DNA melting analysis from 75–95°C, with data being acquired every degree with a 5 s hold at each step. All data were acquired on the "Green" channel, with excitation at 470 nm and emission at 510 nm.

RESULTS AND DISCUSSION

Demonstration of geosmin synthase in the model cyanobacterium *Nostoc punctiforme*

The reported 1893-bp *npun02003620* gene of *Nostoc punctiforme* PCC 73102 (ATCC 29133), currently annotated as encoding a hypothetical protein (ZP_00109187), has 55% amino acid sequence similarity to the N-terminal region of both known geosmin synthases, *S. coelicolor* A3(2) SCO6073 and *S. avermitilis* SAV2163, thereby suggesting the likely biochemical function of this cyanobacterial gene may be similar to the N-terminal mediated reactions of the reported *Streptomyces*. The reported open reading frame is, however, almost 300 bp shorter than either the SCO6073 or SAV2163 gene or any of the homologous Actinomycete and myxobacterial geosmin synthase genes. The predicted ZP_00109187 protein also lacks the essential NSE triad in the C-terminal domain that is found in all other geosmin synthases as well as in all other known terpene synthases. To assess the biochemical function of the reported *npun020003620* open reading frame, we generated recombinant ZP_00109187 protein based on the start and stop codons predicted by the deposited sequence, resulting in a 71KDa protein of the expected size as determined by SDS-PAGE, that we named NJ2. Incubation of recombinant NJ2 with FPP yielded a mixture of germacradienol (94%) and germacrene D (6%) accompanied by a trace of the octalin, but without any detectable geosmin according to the standard GC-MS analysis. The generation of germacradienol indicates that the N-terminal half of the NJ2 protein has been properly folded and has the expected germacradienol/germacrene D synthase activity, consistent with the demonstrated properties of the homologous N-terminal domain of *S. coelicolor* SCO6073 geosmin synthase. On the other hand, the absence of geosmin production is consistent with the apparent truncation of the C-terminal domain and the absence of a functional active site for geosmin formation. Indeed the biochemical properties of the recombinant *Nostoc* NJ2 protein are similar to those of the previously reported truncated mutant derived from the N-terminal half of *S. coelicolor* geosmin synthase, as well as variants of full-length SCO6073 protein carrying mutations in the essential C-terminal DDYYP or NSE triad regions, all of which could convert FPP to germacradienol and germacrene D but were completely defective in geosmin formation.

Comparison of the experimentally determined DNA sequence of the 3′-region of the PCR-amplified NJ2 construct with the published *npun02003620* sequence revealed that the sequence recorded in the *Nostoc punctiforme* genome database contains an extra "T" nucleotide at position 33677, resulting in a false frame shift with the predicted insertion of the isoleucine at amino acid position 616, as well as a premature stop codon after amino acid 630 of the deduced protein ZP_00109187. Excluding this extraneous T nucleotide relieves the implied frame shift and eliminates the erroneous stop codon. The corrected open reading frame encodes an

additional 126 amino acids, corresponding to a predicted length of 756 amino acids (85 kDa). Most importantly, the full-length C-terminal domain of the corrected protein sequence, termed NPUNMOD, now has the universally conserved DDYFP and NDVFSYQKE motifs that are found in the *Stretptomyces* geosmin synthases SCO6073 and SAV2163 (Figure 2) as well as all other Actinomycete and myxobacterial orthologs (GenBank accession FJ010202 and FJ010203 for NJ2 and NPUNMOD). Although the PCR-amplified DNA sequence encoding recombinant NJ2 does retain the C-terminal NSE triad just before the C-terminus, the premature truncation presumably prevents proper assembly of the active site of the C-terminal half of the geosmin synthase.

Recombinant full-length NPUNMOD protein was shown to be a fully functional geosmin synthase by incubation with FPP under the standard conditions. GC-MS analysis of the hexane-soluble extract confirmed the formation of a typical product mixture consisting of geosmin (4%), germacradienol (78%), germacrene D (8%), and octalin (1%) as well (*E*)-nerolidol (9%) (Supplementary information Figure S1). The formation of nerolidol is unusual and may result from some degree of improper folding of the recombinant geosmin synthase or by chaperone-mediated solvolysis of FPP.

Development of a geosmin synthase screening tool

Having identified the corrected, full-length geosmin synthase gene of *Nostoc punctiforme*, we next used the specific sequence information to develop a PCR-based screening procedure to evaluate the presence of geosmin synthase genes in other cyanobacterial strains. To this end we used a pair of partially degenerate primers and G-PCR1 to amplify the target geosmin synthase genes of several cyanobacterial isolates that were directly confirmed as geosmin producers by GC-MS. The PCR primers were designed so that the forward 250F primer sequence would be anchored on the conserved aspartate-rich motif of the N-terminal domain while the 971R reverse primer flanked the NSE triad region normally found 140 amino acids downstream of the aspartate-rich motif. Using this procedure we were able to amplify 743-bp DNA fragments from geosmin-producing *Anabaena laxa*, *Nostoc sp.* UTAH12-18b, and *Phormidium calcicola*, while the diagnostic fragment was absent in the geosmin non-producing organism *Pseudoanabaena limnetica* (Supplementary information Figure S2). Positive PCR products were cloned and sequenced as described earlier. All positive amplicons encoded the strictly conserved Mg^{2+}-binding motifs, DDHFLE and NDLFSYQRE, that are found in the N-terminal domain of all geosmin synthases (Supplementary information Figure S3).

We next developed a rapid real-time PCR (G-PCR2) for screening cyanobacterial isolates, using primers internal to the G-PCR1 products. Seventeen such isolates were screened using this PCR assay, which was run in parallel with GC-MS analyses of isolate suspensions. Table 1 demonstrates that G-PCR2 is able to detect additional isolates of *Phormidium*, *Anabaena*, and *Geitlerinema sp*, all of which were GC-MS positive for geosmin. Moreover, the real-time G-PCR2 protocol did not amplify DNA from any GC-MS geosmin-negative isolates such as *Planktothrix*, *Oscillatoria*, and *Pseudoanabaena sp.* clearly implying that the presence of a geosmin synthase gene is invariably associated with geosmin production by the host, thereby validating the utility of the real-time PCR gene detection as a reliable diagnostic assay for cyanobacterial geosmin producers.

The use of G-PCR2 and melting curve analysis has the additional advantage of differentiating among geosmin-producing species. In this PCR protocol, the positive *Geitlerinema*, *Phormidium*, and *Anabaena* species produce different characteristic melting profiles (Figure 3). For example, *Anabaena circinalis* gives rise to a single peak with a melting temperature (*T*m) of 84.5°C, while *Anabaena laxa* has a more complex melting pattern with two *T*ms; one at 84.5°C and another at 82.3°C. A similarly complex melting pattern is seen for all *Phormidium sp*, with *T*ms at 84.5 and 87.5°C and also for *Geitlerinema* that has *T*ms at 84.5 and 86.5°C.

The use of melting curve analysis for species identification and genotyping is becoming increasingly popular, especially with the availability of *T*m prediction software such as POLAND and MELTSIM (22). Furthermore, the observation of differential *T*ms from a single primer pair has been used elsewhere to characterise several *Naegleria* isolates in a diagnostic setting (23), suggesting that the combination of real-time PCR for detection of geosmin synthase genes and *T*m analysis may turn out to be similarly useful if validated further.

In conclusion, we have demonstrated that the production of geosmin in cyanobacteria is due to the presence of a single gene encoding the geosmin synthase enzyme. We have also developed a diagnostic geosmin synthase PCR protocol that promises to be a valuable tool for use by water utilities in the detection of organisms responsible for geosmin production in any given water body. The flexibility and portability of real-time PCR equipment has previously been used to track toxic cyanobacteria in the field (24), and therefore it is foreseeable that this technology may also be used for mobile monitoring of geosmin-producing cyanobacteria and could assist with current dosing protocols by pinpointing specific problem areas and thereby increasing the efficacy of treatment. The quantification of the abundance of geosmin synthase genes in a water body may become a valuable input into predictive modelling of water storages when coupled with chemical, physical, and physiochemical values, and these data may be used to predict the occurrence of taste and odor episodes before they become an operational problem.

Now that the geosmin synthase gene in cyanobacteria has been identified, exploration of key regulatory mechanisms controlling geosmin production as a function of life-cycle and environmental conditions is now also possible. Such studies could provide insight into strategies to better control the production of geosmin in water storages, and eliminate the need to use environmentally controversial control methods such as copper sulfate dosing. Control and mitigation of taste and odor episodes is a frustratingly common event for water utilities and it is foreseeable that the tools described herein may be used as an adjunct to current monitoring programs to help better engineer a timely response[1].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding from the South Australian Water Corporation is gratefully acknowledged. Research in the laboratory of D. E. C. was supported by grant GM30301 from the U. S. National Institutes of Health.

Literature Cited

1. Burlingame GA, Dann RM, Brock GL. A case study of geosmin in Philadelphia's water. Journal AWWA 1986:56–61.(Management and operations)
2. Carmichael WW, Li R. Cyanobacteria toxins in the Salton Sea. Saline Systems 2006;2:5. [PubMed: 16623944]
3. Izaguirre G, Taylor WD. Geosmin and 2-methlyisoborneol production in a major aqueduct system. Wat Sci Tech 1995;31:41–48.
4. van Apeldoorn ME, van Egmond HP, Speijers GJ, Bakker GJ. Toxins of cyanobacteria. Mol Nutr Food Res 2007;51:7–60. [PubMed: 17195276]

[1]After the original submission of this manuscript, Agger *et al* (epub ahead of print *J Bac* doi:10.1128/JB.00759-08) also reported the correction of the sequencing error in *npun02003620* and describe the expression of the full length recombinant protein. Although they reported the conversion of FPP to germacradienol and germacrene D, there were unable, however, to detect the formation of geosmin as outlined in this manuscript.

5. Cook D, Newcombe G, Sztajnbok P. The application of powdered activated carbon for MIB and geosmin removal: predicting PAC doses in four raw waters. Water Res 2001;35:1325–1333. [PubMed: 11268853]
6. Ho L, Hoefel D, Bock F, Saint CP, Newcombe G. Biodegradation rates of 2-methylisoborneol (MIB) and geosmin through sand filters and in bioreactors. Chemosphere 2007;66:2210–2218. [PubMed: 17005238]
7. Bentley R, Meganathan R. Geosmin and methylisoborneol biosynthesis in streptomycetes. Evidence for an isoprenoid pathway and its absence in non-differentiating isolates. FEBS Lett 1981;125:220–222. [PubMed: 7227551]
8. Dickschat JS, Bode HB, Mahmud T, Muller R, Schulz S. A novel type of geosmin biosynthesis in myxobacteria. J Org Chem 2005;70:5174–5182. [PubMed: 15960521]
9. Dickschat JS, Wenzel SC, Bode HB, Muller R, Schulz S. Biosynthesis of volatiles by the myxobacterium *Myxococcus xanthus*. Chembiochem 2004;5:778–787. [PubMed: 15174160]
10. Schulz S, Fuhlendorff J, Reichenbach H. Identification and synthesis of volatiles released by the myxobacterium *Chondromyces crocatus*. Tetrahedron 2004;60:3863–3872.
11. Gerber NN, Lechavalier HA. Geosmin, an earthy smelling substance isolated from actinomyces. Appl microbiol 1956;13(6):935–938. [PubMed: 5866039]
12. Gust B, Challis GL, Fowler K, Kieser T, Chater KF. PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. Proc Natl Acad Sci U S A 2003;100:1541–1546. [PubMed: 12563033]
13. Cane DE, Watt RM. Expression and mechanistic analysis of a germacradienol synthase from *Streptomyces coelicolor* implicated in geosmin biosynthesis. Proc Natl Acad Sci U S A 2003;100:1547–1551. [PubMed: 12556563]
14. Jiang J, Cane DE. Geosmin biosynthesis. Mechanism of the fragmentation-rearrangement in the conversion of germacradienol to geosmin. J Am Chem Soc 2008;130:428–429. [PubMed: 18095692]
15. Jiang J, He X, Cane DE. Geosmin Biosynthesis. *Streptomyces coelicolor* Germacradienol/Germacrene D Synthase Converts Farnesyl Diphosphate to Geosmin. J Am Chem Soc 2006;128:8128–8129. [PubMed: 16787064]
16. Jiang J, He X, Cane DE. Biosynthesis of the earthy odorant geosmin by a bifunctional *Streptomyces coelicolor* enzyme. Nat Chem Biol 2007;3:711–715. [PubMed: 17873868]
17. Nawrath T, Dickschat JS, Muller R, Jiang J, Cane DE, Schulz S. Identification of (8S,9S,10S)-8, 10-Dimethyl-1-octalin, a Key Intermediate in the Biosynthesis of Geosmin in Bacteria. J Am Chem Soc 2008;130:430–431. [PubMed: 18095691]
18. Cane DE, He X, Kobayashi S, Omura S, Ikeda H. Geosmin biosynthesis in *Streptomyces avermitilis*. Molecular cloning, expression, and mechanistic study of the germacradienol/geosmin synthase. J Antibiot (Tokyo) 2006;59:471–479. [PubMed: 17080683]
19. Christianson DW. Structural biology and chemistry of the terpenoid cyclases. Chem Rev 2006;106:3412–3442. [PubMed: 16895335]
20. Komatsu M, Tsuda M, Omura S, Oikawa H, Ikeda H. Identification and functional analysis of genes controlling biosynthesis of 2-methylisoborneol. Proc Natl Acad Sci U S A 2008;105:7422–7427. [PubMed: 18492804]
21. Ludwig F, Medger A, Bornick H, Opitz M, Lang K, Gottfert M, Roske I. Identification and expression analyses of putative sesquiterpene synthase genes in *Phormidium sp*. and prevalence of geoA-like genes in a drinking water reservoir. Appl Environ Microbiol 2007;73:6988–6993. [PubMed: 17827321]
22. Rasmussen JP, Saint CP, Monis PT. Use of DNA melting simulation software for in silico diagnostic assay design: targeting regions with complex melting curves and confirmation by real-time PCR using intercalating dyes. BMC Bioinformatics 2007;8:107. [PubMed: 17391531]
23. Robinson BS, Monis PT, Dobson PJ. Rapid, sensitive, and discriminating identification of *Naegleria spp*. by real-time PCR and melting-curve analysis. Appl Environ Microbiol 2006;72:5857–5863. [PubMed: 16957204]
24. Rasmussen JP, Giglio S, Monis PT, Campbell RJ, Saint CP. Development and field testing of a real-time PCR assay for cylindrospermopsin-producing cyanobacteria. J Appl Microbiol 2008;104:1503–15. [PubMed: 18179541]

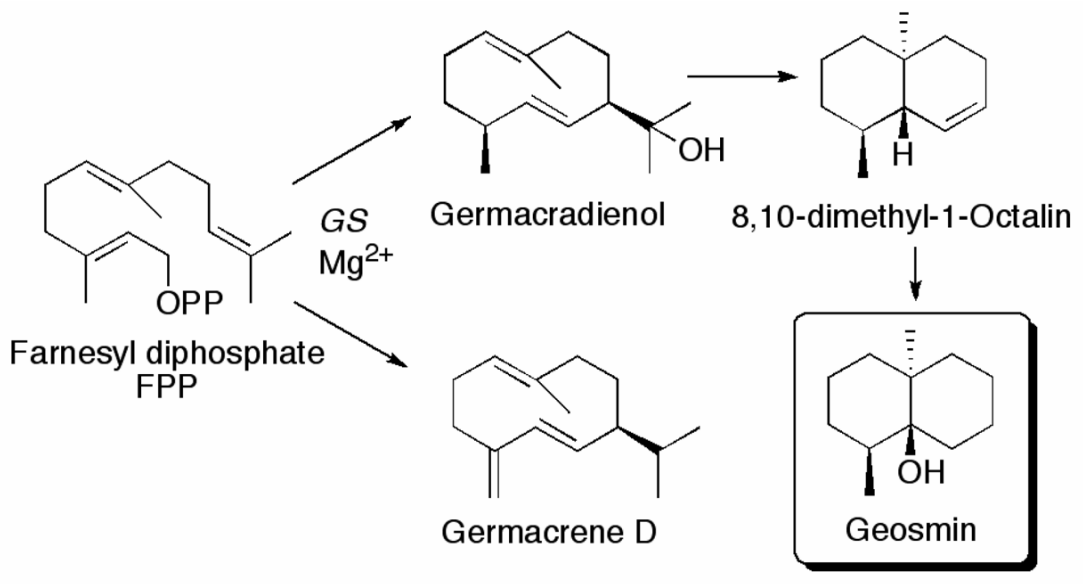

Figure 1.
Geosmin synthase(GS)-catalyzed cyclization of farnesyl diphosphate (FPP) to geosmin, germacradienol, germacrene D, and 8,10-dimethyl-1-octalin. Germacradienol and germacrene D are formed by the N-terminal domain of the bifunctional protein, and geosmin is generated from germacradienol via the octalin by the C-terminal domain.

Figure 2.

Alignment of proteins from *Nostoc punctiforme* (published ZP_00109187, experimental NJ2, and experimental NPUNMOD), *Streptomyces coelicolor* A3(2) (SCO6073), and *Streptomyces avermitilis* (SAV2163). Black boxed residues indicate signature terpene synthase motifs; the grey boxed residue in ZP_00109187 details where the sequencing error was detected ("*" indicates exact sequence match,":" indicates moderate match, "." Indicates low match)

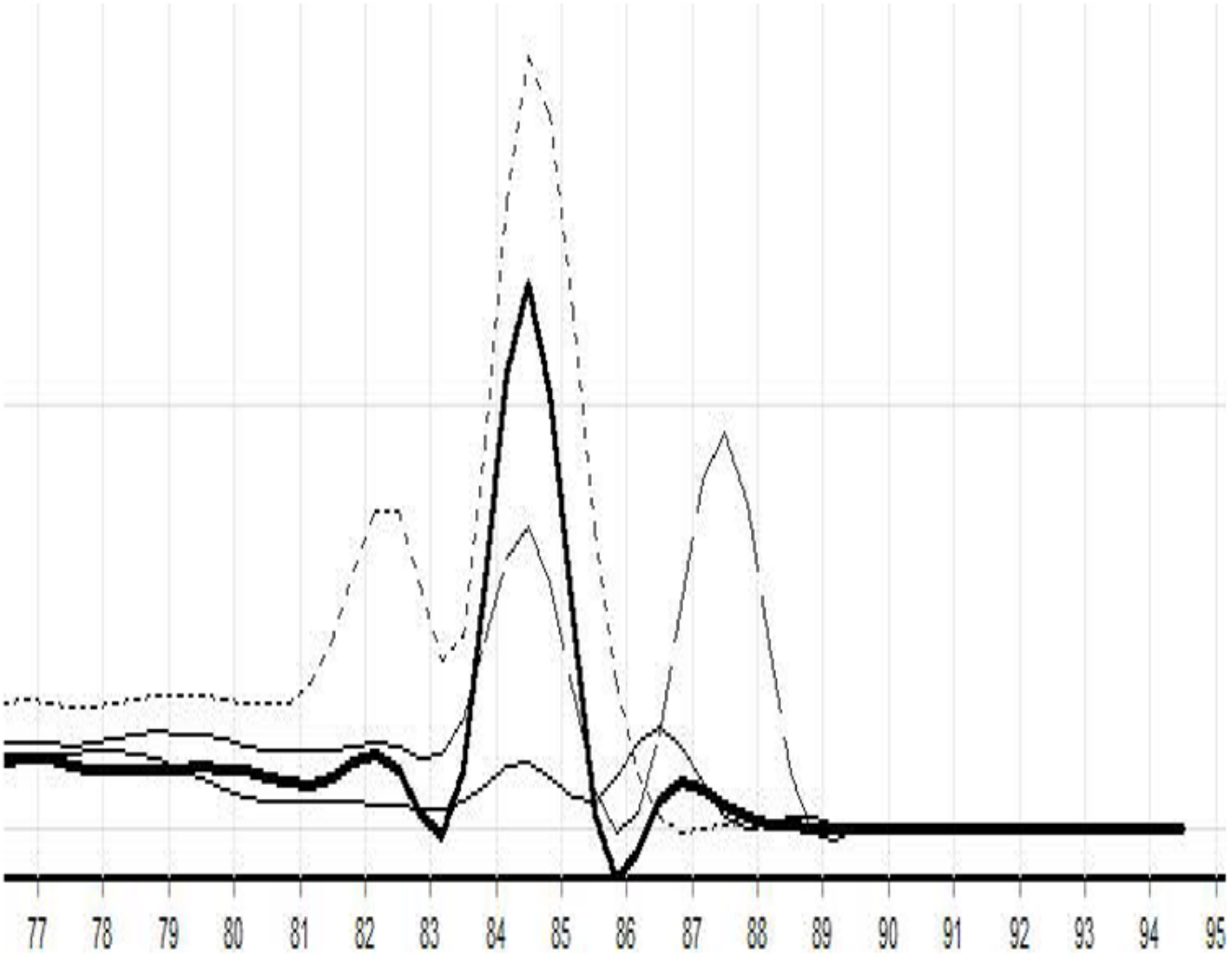

Figure 3.
Melting curve analysis of G-PCR2 products. *Anabaena circinalis* (thick line), *Anabaena laxa* (dotted), *Phormidium sp* (large dashes), and *Geitlerinema* (thin line) all displaying differentiated melting curves.

Table 1

Comparison of a geosmin synthase PCR assay and GC-MS analysis using environmental cyanobacterial isolates

Isolate	Origin	GEOSMIN (ng/L)	G-PCR2
Oscillatoria LM603	Castiac lake, USA	<2	−
Oscillatoria	Hope Valley, SA	<2	−
Oscillatoria	Hope Valley, SA	<2	−
Phormidium sp	Happy Valley, SA	112	+
Phormidium sp	Happy Valley, SA	21	+
Phormidium sp	Happy Valley, SA	4	+
Phormidium sp	Happy Valley BR, SA	47	+
Phormidium sp	Happy Valley, SA	67	+
Phormidium sp	Happy Valley, SA	553	+
Planktothrix sp DVL1003C	Diamond Valley, USA	<2	−
Anabaena circinalis ANA346B	Myponga, SA	152	+
Pseudoanabaena sp	Happy Valley, SA	<2	−
Pseudoanabaena sp	Happy Valley, SA	<2	−
Geitlerinema	Little Para, SA	3113	+
Phormidium 007D	Hope Valley, SA	198	+
Phormidium 005E	Hope Valley, SA	60	+
Phormidium 012G	Hope Valley, SA	21	+