SUPPLEMENTARY INFORMATION

Biosynthesis of the earthy odorant geosmin by a bifunctional *Streptomyces coelicolor* enzyme

Jiaoyang Jiang[1], Xiaofei He[1] & David E. Cane[1]*

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

* Corrrespondence: David_Cane@brown.edu
Tel: +1 401 863 3588
Fax: +1 401 863 9368

Contents

Construction of the C-terminal truncated mutant of SCO6073. The C-terminal truncated mutant encoding aa 327-726 of SCO6073 was constructed by using *Pfu* turbo polymerase for PCR amplification of the *Nco*I-*Xho*I fragment of pRW31[1] using the forward primer Met327start, 5′-ATTTTATCCATGGTTATGAACAAGGGGGAGCGGCCC-3′ in combination with the reverse primer rmwhalt4-histag, 5′-TAATAACTCGAGGTGCGTCAGTGCGGGACT-3′. (Underlined bases indicate restriction sites for *Nco*I and *Xho*I, respectively.) The resultant PCR product was purified immediately using a Qiaquick PCR purification kit (Qiagen), eluted with 50–60 μl of 10 mM Tris−HCl, pH 8.5, then double-digested with *Nco*I and *Xho*I at 37 °C for 2.5 h, and repurified with the Qiaquick purification kit. The digested PCR product was ligated into *Nco*I/*Xho*I-digested pET21d using a 6:1 molar ratio of insert:vector and T4 DNA ligase. The ligation mixture was used to transform competent cells of *E. coli* XL1-Blue under standard conditions. The resulting transformants were grown overnight at 37 °C on LB-agar plates containing ampicillin (100 μg/ml). Plasmid DNA was isolated from 18 individual colonies after overnight growth at 37 °C in LB-ampicillin (100 μg/ml) medium and purified by using Qiagen spin miniprep kits. Plasmids were screened by *Pst*I restriction digests and pJJ3 was sequenced to confirm the integrity of the inserted DNA sequences.

Expression of C-terminal truncated mutant of SCO6073, pJJ3. Plamid pJJ3 was used to transform competent cells of *E. coli* BL21(DE3)pLysS. For expression of His_6-tagged pJJ3 protein, a 500-ml culture of *E. coli* BL21(DE3)pLysS/pJJ3 was grown at 37 °C to an OD_{600} of 0.7. The culture was cooled to 18 °C and induced with 0.4 mM IPTG, then incubated overnight at 18 °C. The cells were harvested by centrifugation (7,500*g*) and resuspended in 20 ml of cell lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 % glycerol, 5 mM β-mercaptoethanol, pH 7.5). The cells were disrupted by French Press, and the mixture was centrifuged at 20,200g for 40 min. The supernatant was mixed with 5 ml of Ni-NTA resin, and mixed gently by shaking at 200 rpm on a rotary shaker at 4 °C for 1-2 h. The mixture of pJJ3-containing lysate and Ni-NTA was loaded into a column that was washed with 20 mM imidazole in cell lysis buffer. To elute the desired protein, the column was washed stepwise by 50, 100, 250 and 500 mM imidazole buffer. The desired C-terminal pJJ3 protein was eluted with 100-250 mM imidazole buffer, as monitored by SDS-PAGE. pJJ3 protein exhibited a single band of M_r 46,000, consistent with the predicted M_D of 45,932 Da. After concentration of the C-terminal truncated protein using a Millipore Ultracel YM-10 membrane, a PD-10 column was used to exchange the imidazole buffer for reaction buffer (50 mM Tris-HCl, 20 % glycerol, pH 8.2). Using this method, 2-3 mg of soluble C-terminal truncated pJJ3 protein was isolated, with a purity of >95% by SDS-PAGE. Attempts to express alternative C-terminal truncated mutants corresponding aa 368-726 and aa 383-726 were unsuccessful.

Germacradienol–geosmin synthase mutants. The desired mutations were introduced by using the Stratagene QuikChange II XL Site-Directed Mutagenesis Kit in combination with the mutagenic primers listed in **Supplementary Table 1** and DNA from plasmid pRW31 as template. The identity of each of the mutations was verified by sequencing of the full-length gene. Each of the mutant plasmids was used to transform competent cells of the expression host *E. coli* BL21(DE3)pLysS under standard conditions. The resulting transformants were grown overnight at 37 °C on LB-agar plates containing ampicillin (100 μg/ml) and chloramphenicol (34 μg/ml). For expression of mutant S233A, a 500-ml culture of *E. coli* BL21(DE3)pLysS/pRW31S233A was grown at 37 °C to an OD_{600} of 0.7. The culture was induced with 2 mM IPTG, then incubated for 4 h at 37 °C. The cells were harvested by centrifugation (7,500*g*) and resuspended in 40 ml of cell lysis buffer (50 mM Tris−HCl, 20% glycerol, 1 mM EDTA, pH 8.2). The remainder of the inclusion body solubilisation/refolding procedure followed the previously described protocol[1].

Identification of isolepidozene (6) produced by the S233A mutant of germacradienol–geosmin

synthase. GC–MS analysis of the product mixture resulting from incubation of FPP with the pRW31_S233A mutant showed a pair of overlapping peaks centered at a retention time of 10.24 min (**Supplementary Fig. 4**). One of the overlapping peaks, ret. time 10.23 min, was identified as **4** by comparison of the mass spectrum and retention time with an authentic sample of germacrene D (**4**) produced by the wild-type SCO6073 synthase. The mass spectrum of the second overlapping peak, ret. time 10.25 min ($[M]^+$ *m/z* 204, base peak *m/z* 121; retention index (RI) 1482.77) was assigned to isolepidozene (**6**).

Supplementary Table 1. Mutagenic primers for *S. coelicolor* germacradienol–geosmin synthase (GS)

Enzyme	Plasmid	Wild-type and mutant primer sequence[a]
WT GS	pRW31	GTGGGTCTTCTTCTTCGACGACCACTTCCTGGAGAAGTACA
D86E	pRW31D86E	GTGGGTCTTCTTCTTC**GAA**GACCACTTCCTGG
L90D	pRW31L90D	CTTCTTCGACGACCACTTC**GAC**GAGAAGTACA
WT GS	pRW31	AACGACCTCTTCTCCTACCAGCGCGAG
S233A	pRW31S233A	AACGACCTCTTC**GCC**TACCAGCGCGAG
WT GS	pRW31	CGACGTCCTCACCTCGCGGCTGC
T271A	pRW31T271A	CGACGTCCTC**GCC**TCGCGGCTGC
S272A	pRW31S272A	CGACGTCCTCACC**GCG**CGGCTGC
WT GS	pRW31	GGCACGTACGGCGACGACTACTACCCG
D455N/ D456N	pRW31D455N/D456N	GGCACGTACGGC**AATAAT**TACTACCCG
WT GS	pRW31	GCGTGCCTGCTCAACGACGTCTTCTCCTACCAGAAGGAGATCGAGTACGAG
N598L	pRW31N598L	GCGTGCCTGCTC**CTA**GACGTCTTCTCC
D599L	pRW31D599L	TGCCTGCTCAAC**CTA**GTCTTCTCCTAC
S602A	pRW31S602A	AACGACGTCTTC**GCA**TACCAGAAGGAG
E606Q	pRW31E606Q	TCCTACCAGAAG**CAA**ATCGAGTACGAG

a. Nucleotide substitutions are highlighted in **boldface**.

Supplementary Table 2. Distribution of products from incubation of FPP with C-terminal domain mutants of full-length germacradienol–geosmin synthase.

Protein (μM)		FPP (μM)	Product Distribution (%)			
			geosmin (**1**)	germacradienol (**3**)	germacrene D (**4**)	octalin (**5**)
WT (pRW31)	9.2	82	13	74	10	3
D455N/D456N	8.6	49	ND[a]	90	10	trace
N598L	13.8	61	ND	81	16	3
D599L	8.3	42	ND	85	13	2

S602A	11.2	53	ND	87	11	2
E606Q	8.3	55	ND	89	11	trace

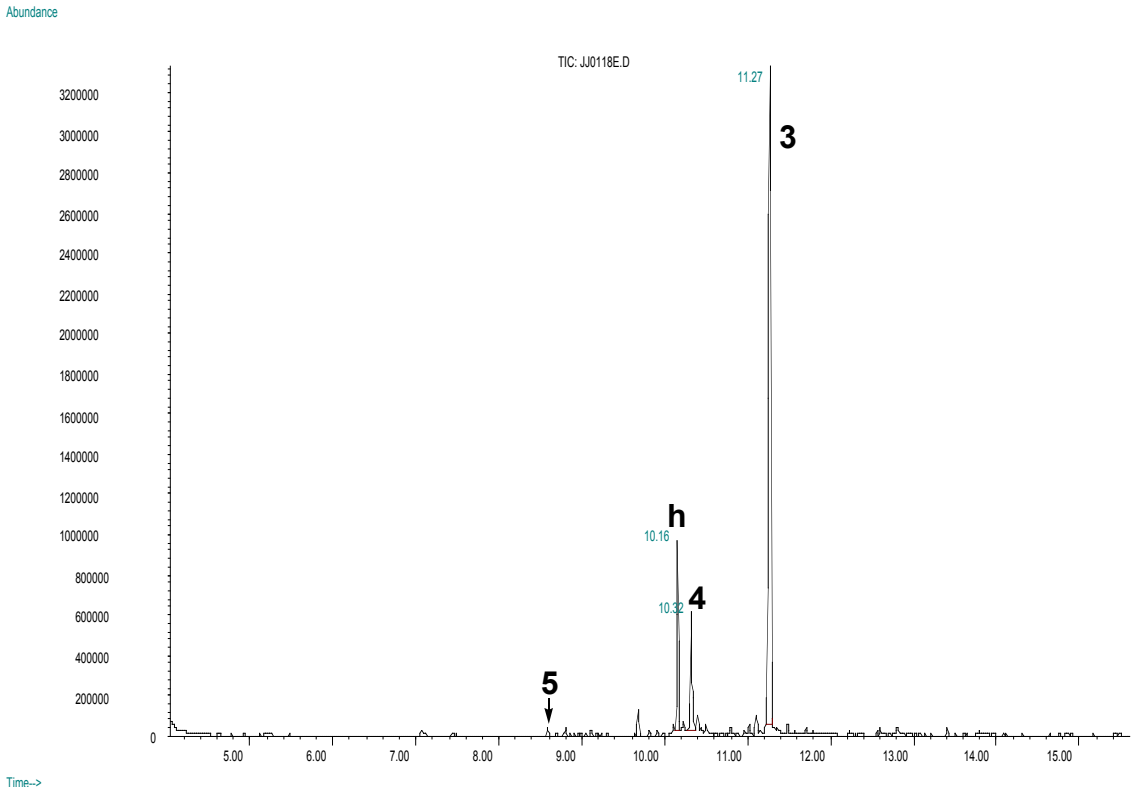
[a] ND, not detected.

Supplementary Table 3. Sequence comparison of *S. coelicolor* SCO6073 germacradienol–geosmin synthase (Q9X839, CYC2_STRCO) with bacterial orthologs. (For sequence alignments, see **Supplementary Figure 5**.)

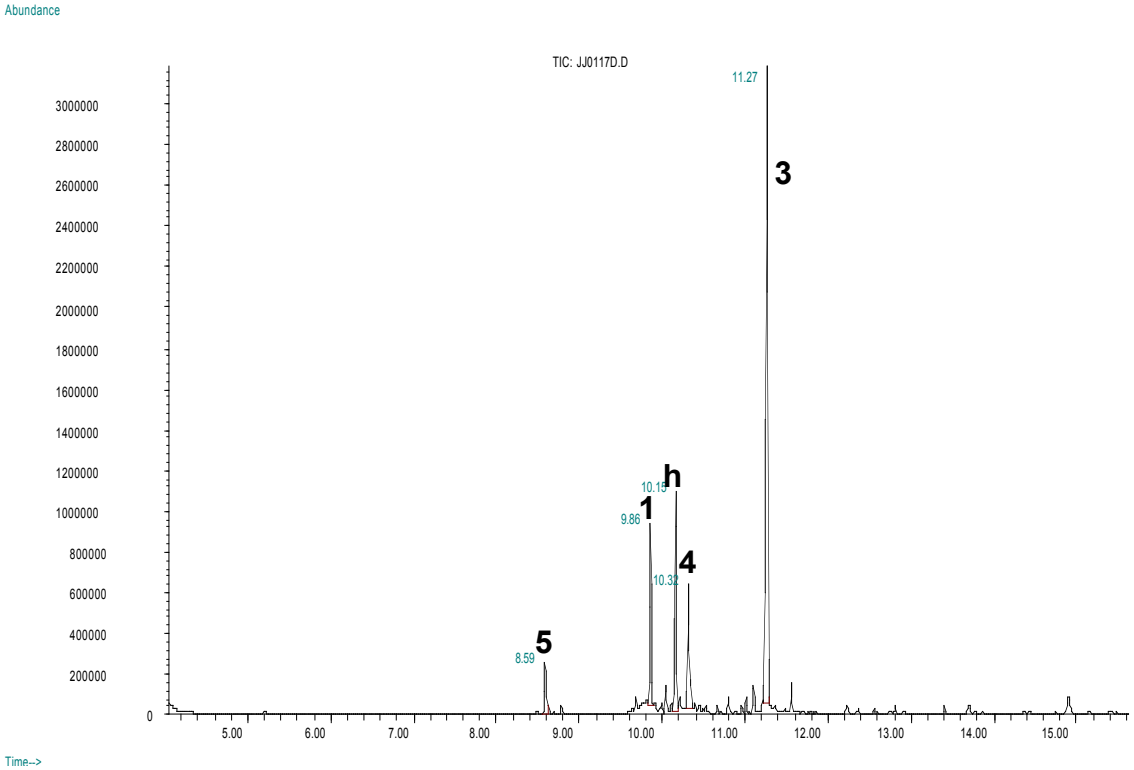
Organism	UniProt ID	aa	Identity (%)	Similarity (%)
Streptomyces scabies	[a]	738	78	85
Streptomyces avermitilis	Q82L49	725	76	85
Streptomyces peucetius ATCC 27952	spterp13[b]	732	64	74
Frankia sp. Strain Cc13	Q2J565	751	60	72
Saccharopolyspora erythraea NRL 2338	A4FEI8	758	58	70
Frankia alni ACN14a	Q0RBQ4	758	58	70
Frankia sp. EAN1pec	Q3WJX6	750	59	72
Myxococcus xanthus strain DK 1622	Q1CYZ7	755	57	72
Saccharopolyspora erythraea NRL 2338	A4FJE8	763	56	69
Stigmatella aurantiaca DW4/3-1	Q09A24	704	55	67
Saccharopolyspora erythraea NRL 2338	A4FGS3	732	45	57

[a] *S. scabies* chromosome, nt 2284449-2282248; Sanger Centre, http://www.sanger.ac.uk/Projects/S_scabies/.

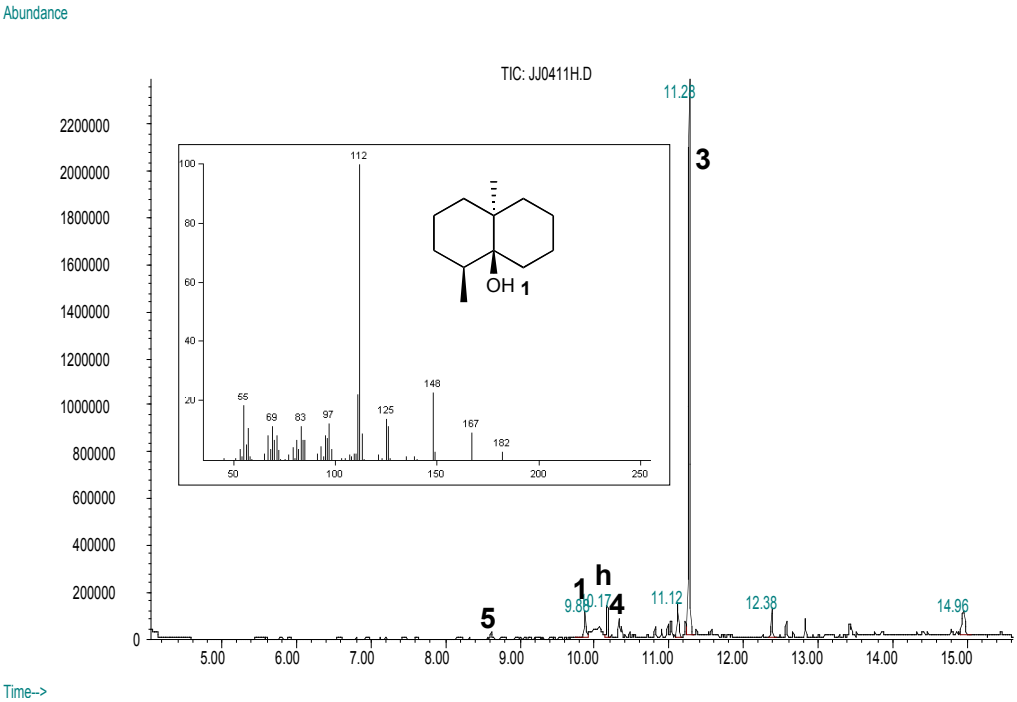
[b] *S. peucetius* genome sequence. Cf. Parajuli, N., Basnet, D.B., Chan Lee, H., Sohng, J.K. & Liou, K. Genome analyses of *Streptomyces peucetius* ATCC 27952 for the identification and comparison of cytochrome P450 complement with other *Streptomyces*. *Arch. Biochem. Biophys.* **425**, 233-241 (2004).



Supplementary Figure 1. Incubation of FPP (82 μM) with recombinant N-terminal domain (pRW22p, 5.3 μM) at 30 °C for 6.25 h. GC–MS (TIC): **3**, germacradienol; **4**, germacrene D; **5**, octalin; **h**, α-humulene (**16**) internal standard.

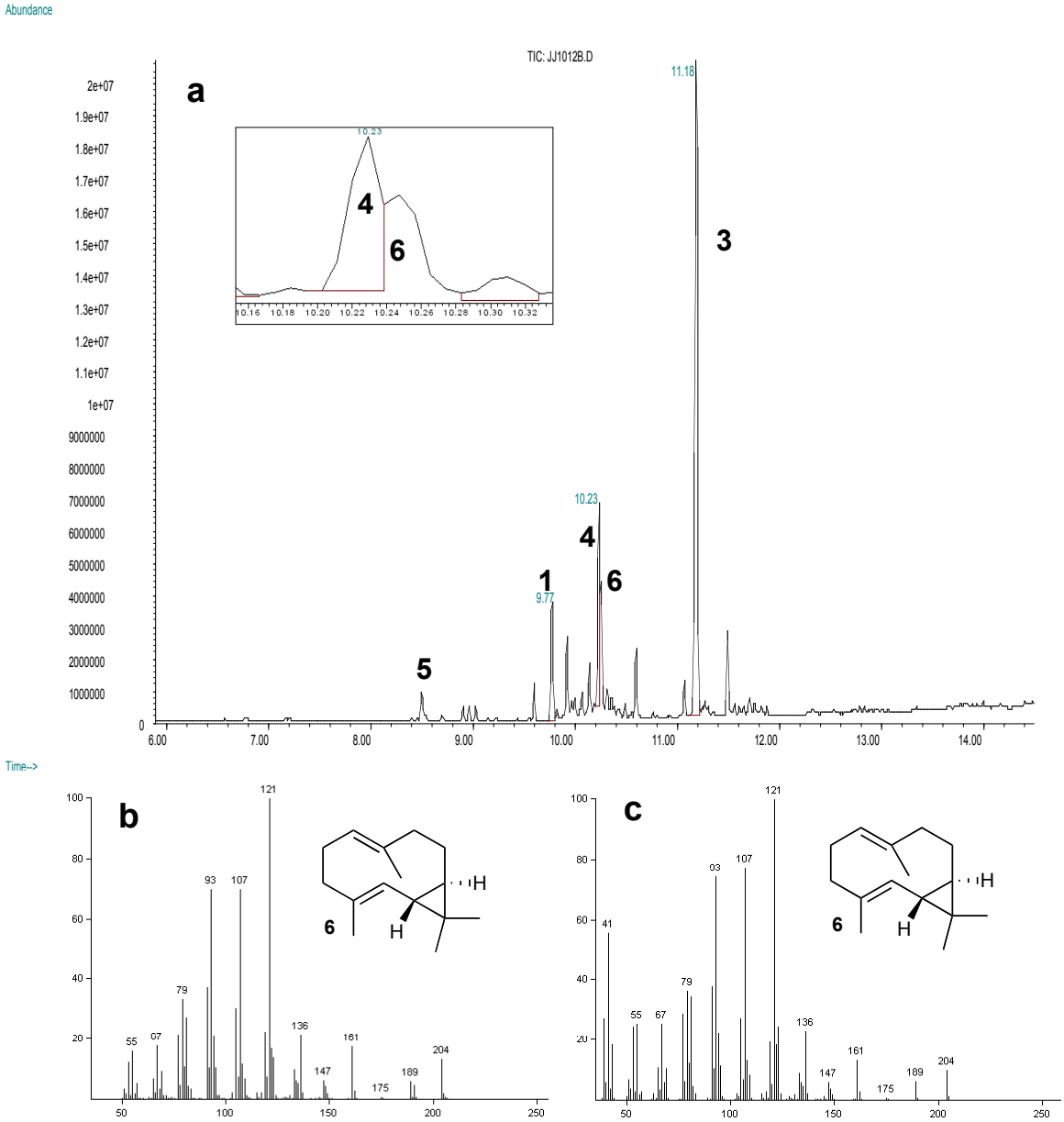


Supplementary Figure 2. Incubation of FPP (**2**) (82 μM) with recombinant N-terminal domain (pRW22p, 5.4 μM) and C-terminal domain (pJJ3p, 5.5 μM) at 30 °C for 6.25 h. GC–MS (TIC): **1**, geosmin; **3**, germacradienol; **4**, germacrene D; **5**, octalin; **h**, α-humulene (**16**) internal standard.



Supplementary Figure 3. Incubation of germacradienol (**3**) (containing germacrene D and octalin **5**, in 50 μl MeOH) with recombinant C-terminal domain (3.1 μM) and 4 mM Mg^{2+}. GC–MS (TIC): **1**, geosmin; **3**, germacradienol; **4**, germacrene D; **5**, octalin; **h**, α-humulene (**16**) internal standard.

Inset: Mass spectrum of geosmin (**1**) ret. time 9.88 min.



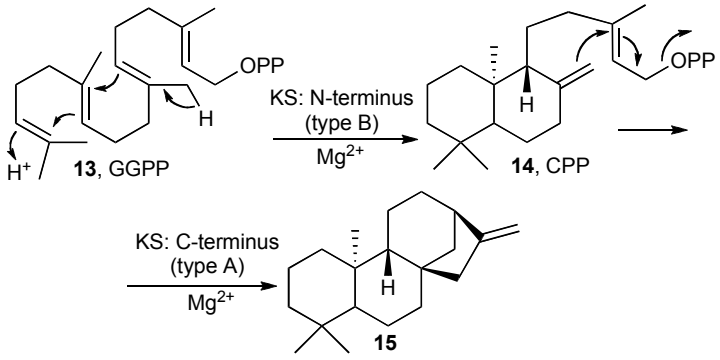
Supplementary Figure 4. Incubation of FPP (**2**) (165 μM) with 7.3 μM GS mutant S233A at 30 °C for 7.25 h. **a.** GC–MS (TIC), Inset: Overlapping peaks at 10.23 min (germacrene D (**4**)) and 10.25 min (**6**). **b.** Mass spectrum of isolepidozene (**6**) ret. time 10.25 min, RI (retention index) 1483. The RI was calculated by the virtual carbon number method[2], using germacrene D (virtual carbon number 14.79) and germacradienol (virtual carbon number 16.33), both present in the sample, as the internal references. **c.** Mass spectrum of isolepidozene (**6**) from MassFinder 3.0 Library, RI 1483. The corresponding *cis*-fused dimethylcyclopropyl stereoisomer, *E*,*E*-bicyclogermacrene, which has a mass spectrum very similar to **6**, has an RI value of 1494 and is therefore excluded.

```
CLUSTAL W (1.83) multiple sequence alignment

Q1CYZ7|Q1CYZ7_MYXXD      MSTAKNKQPFELPDFYVPWPARLNPNLEGARVHSKAWARELGIIGRPKDGSAPEIWSEAK 60
Q09A24|Q09A24_STIAU      ------------------------------------------------------------
Q9X839|CYC2_STRCO        ----MTQQPFQLPHFYLPHPARLNPHLDEARAHSTTWAREMGMLE------GSGVWEQSD 50
Q82L49|Q82L49_STRAW      -----MTQPFQLPHFYMPYPARLNPHLDEARAHSTRWARGMGMLE------GSGIWEQSD 49
Q2J565|Q2J565_FRASC      ------MQPFTLPEFYVPYPARLNPNLEQARVHSRAWADEMEMIDSPQH--GTAIWTEAD 52
Q0RBQ4|Q0RBQ4_FRAAA      ------MQPFTLPEFYVPYPARLSPHLEQAREHSREWARAMEMIDTPQH--GIAIWTERD 52
Q3WJX6|Q3WJX6_9ACTO      ------MQPFTLPEFYLPYPPRLNPNLEHARVHSRAWAGEMEMIDVPQD--GVAIWSGQD 52
A4FEI8|A4FEI8_SACEN      ------MQPFQQPEFYMPYPARLNPNLERAREHSKAWACAMDMIDVPQE--GTLIWDEND 52
A4FGS3|A4FGS3_SACEN      MPAPQQRQPYRLPAFYLPRPARLNPDLEAARARSRRWAEEMGMLGSRAEPEGEQVWTRED 60

Q1CYZ7|Q1CYZ7_MYXXD      FDAMDYALLCAYTHPEAPGPELDLVTDWYVWVFYFDDHFLELYKRPQDQVGAKAYLDRLP 120
Q09A24|Q09A24_STIAU      ---MDYALLCAYTHPEAPSLELDLVTDWYVWVFYFDDHFLDVYKRTQDQVGAREYLDRLP 57
Q9X839|CYC2_STRCO        LEAHDYGLLCAYTHPDCDGPALSLITDWYVWVFFFDDHFLEKYKRSQDRLAGKAHLDRLP 110
Q82L49|Q82L49_STRAW      LDAHDYGLLCAYTHPDCDGPALSLITDWYVWVFFFDDHFLETFKRTQDREGGKAYLDRLP 109
Q2J565|Q2J565_FRASC      FDAHDYALLCAYTHPDSVSRKLDLVTDWYVWVFYFDDHFLELYKRSHDMAGARAYLDRLP 112
Q0RBQ4|Q0RBQ4_FRAAA      LDAHDYALLCAYTHPDATADRLNLITDWYVWVFYFDDHFLELYKRSHDLAGARAYLDRLP 112
Q3WJX6|Q3WJX6_9ACTO      FDSHDYALLCAYTHPDADEARLDLITDWYVWVFYFDDHFLEVYKRGRDVAGARRYLDRLR 112
A4FEI8|A4FEI8_SACEN      FDSHDYALLCAYTHPDADGPMLDLITDWYVWVFYFDDHFVELYKRNPDLAGAKEYLDRLP 112
A4FGS3|A4FGS3_SACEN      FDRHDYALLCAYAHPDASAPALELITGWYVWAFFFDDHFLARYKRTGDVDGARAHLLGLA 120

Q1CYZ7|Q1CYZ7_MYXXD      LFMPVDPAATPPPPTNPVEAGLLDLWNRTVPSRSMAWRRRFFESTKHLLDESSWELSNIS 180
Q09A24|Q09A24_STIAU      AFMPVDLSAAPPTPTNPVERGLADLWARTVPTKSEAWRRRFFESTKSLLEESNWELNNIS 117
Q9X839|CYC2_STRCO        LFMPLDDAAGMPEPRNPVEAGLADLWTRTVPAMSADWRRRFAVATEHLLNESMWELSNIN 170
Q82L49|Q82L49_STRAW      LFMPLDLSAPVPEPENPVEAGLADLWARTVPAMSADWRKRFAVSTEHLLNESLWELSNIN 169
Q2J565|Q2J565_FRASC      AFMPVDGEITE-TPTNPVERGLADLWTRTVPERSADWRRRFAVSTKNLLDESLWELANIN 171
Q0RBQ4|Q0RBQ4_FRAAA      AFMPVDGEITE-EPSNPVERGLADLWTRTVPARSADWRARFAVSTRNLLDESLWELENIN 171
Q3WJX6|Q3WJX6_9ACTO      LFMPVEGAVTA-EPANPVERGLADLWSRTVPDRTPAWRRRFATSTRHLLDESLWELANID 171
A4FEI8|A4FEI8_SACEN      AFMPVEGPITA-EPTNPVERGLADLWQRTVPARTADWRRRYAENTKHLLDESLWELSNIS 171
A4FGS3|A4FGS3_SACEN      ELMPVGPSDAAPAATGPVERGLADLWVRTAPEVPARWLVRFAASTRELLENRLRELTGTS 180

Q1CYZ7|Q1CYZ7_MYXXD      DRRVSNPIEYIEMRRKVGGAPWSANLVEHAVFAEVPDRVAASRPMRVLKDTFSDAVHLRN 240
Q09A24|Q09A24_STIAU      ERRVSNPIEYIEMRRKVGGAPWSADLVEHAVFAEIPARIAASRPMTVLKDTFSDGVHLRN 177
Q9X839|CYC2_STRCO        EGRVANPVEYIEMRRKVGGAPWSAGLVEYAT-AEVPAAVAGTRPLRVLMETFSDGVHLRN 229
Q82L49|Q82L49_STRAW      EGRIANPVEYIEMRRKVGGAPWSAGLVEYAT-AEVPAAVAGSRPLRVLMETFSDGVHLRN 228
Q2J565|Q2J565_FRASC      AGRLANPIEYVEMRRKVGGAPWSANLVEHAADAEVPAQVAATRPLQVLRDTFADAVHLRN 231
Q0RBQ4|Q0RBQ4_FRAAA      AARLSNPIEYIEMRRKVGGAPWSANLVEHAADAEVPARVAATRPLQVLRDTFADAVHLRN 231
Q3WJX6|Q3WJX6_9ACTO      ENRLANPVEYIEMRRKVGGAPWSANLVEHAADAEVPDAIAATRPAQVLRDTFSDAIHLRN 231
A4FEI8|A4FEI8_SACEN      RNRLSNPIEYIEMRRKVGGAPWSANLVEHAVDSEVPAAIASARPMQVLRDTFSDAVHLRN 231
A4FGS3|A4FGS3_SACEN      RCGVPNPVDHIAMRREAGGASWSAALVEYAAGSEVPDVVARSRPMRVLRDSFCDGVHLRN 240

Q1CYZ7|Q1CYZ7_MYXXD      DLFSYEREILEEGELSNGVLVMEKFLNISPPSAAHLVNEVLTSRLQQFENTVLTELPSLF 300
Q09A24|Q09A24_STIAU      DLFSYQREIQEEGELANCVLVFEKFLNVDAQRAANLVNEVLTSRLQQFENTALTELPSLF 237
Q9X839|CYC2_STRCO        DLFSYQREVEDEGELSNGVLVLETFFGCTTQEAADLVNDVLTSRLHQFEHTAFTEVPAVA 289
Q82L49|Q82L49_STRAW      DLFSYQREVEEEGELSNGVLVLETFFGCTTQEAAETVNDILTSRLHQFEHTALTEVPALA 288
Q2J565|Q2J565_FRASC      DLFSYQREVEEEGELSNGVLVIERFLGCGTQEAADTVNDLLTSRLHQFEHTAVTELPAVL 291
Q0RBQ4|Q0RBQ4_FRAAA      DLFSYEREVTEEGELSNGVLVVERFLDIDTQAAADTVNDLLTSRLHQFEHTAATELPAVL 291
Q3WJX6|Q3WJX6_9ACTO      DLFSYQREVQEEGELSNGVLVLERFLDCPTQQAADAVNDLLTSRLHQFEHTALTELPPVL 291
A4FEI8|A4FEI8_SACEN      DLFSYQREVQDEGELSNSVLVFEKFLDCSTQDAADTVNDLLTSRLHQFEHTALTEVPALL 291
A4FGS3|A4FGS3_SACEN      DIFSYPRETSEEGELGNGVLVVERFFDTDPQEAADTVNDLLTSRLHQFENVTLTELPAMF 300

Q1CYZ7|Q1CYZ7_MYXXD      VEFGLNPVEQAQVLTYVRGLQDWQSGGHEWHMRSSRYMNKGSGGAG-------------- 346
Q09A24|Q09A24_STIAU      EENALNPVERAHVLTYVRGLQDWQSGGHEWHMRSSRYMNKGAGGAGDTD----------- 286
Q9X839|CYC2_STRCO        LEKGLTPLEVAAVGAYTKGLQDWQSGGHEWHMRSSRYMNKGERPLAGWQA---------- 339
Q82L49|Q82L49_STRAW      LEKGLTPPEVAAVAAYARGLQDWQSGGHEWHLRSSRYMNEGALSQKRPFG---------- 338
Q2J565|Q2J565_FRASC      EEHGVDPGSRLEVLAYVKGLQDWQSGGHEWHLRSSRYMNRAVAPESGE------------ 339
Q0RBQ4|Q0RBQ4_FRAAA      DEHAIDPAGRLAALAYIKGLQDWQSGGHEWHLRSSRYMNREATPDAVPPGLGPLAGLGGT 351
Q3WJX6|Q3WJX6_9ACTO      DEHGVTPTARRDVLAYVKGLQDWQAGGHEWHMRSSRYMNAESGATGPVP----------- 340
A4FEI8|A4FEI8_SACEN      DENGVDPQGRLAVLGYVKGLQDWQSGGHEWHIRSSRYMNEGLVEQSALAGQSAPGQPALP 351
A4FGS3|A4FGS3_SACEN      EEHGLSPVERADVLDYVKGLQDWQSGAHEWHLRSGRYAVPGGAEPREPRR---------- 350
```

```
Q1CYZ7|Q1CYZ7_MYXXD      ------------GFFLGPNGLGTSAARLPQSPTALGLTRLKNFSHVPYQPVGPVKLPKFY 394
Q09A24|Q09A24_STIAU      ------------GLPLGLSGLGLSAVRFPFSASALGLNRFKSFTHTPYMPVGPVKLPKFY 334
Q9X839|CYC2_STRCO        ------------------LTGPGTSAADVGALLADAVAQRARSYTYVPFQKVGPSVIPDIR 382
Q82L49|Q82L49_STRAW      ------------------LSAIGTSAADLRGLLADAGAERLRRYTHVPFQKVGPSRIPDFH 381
Q2J565|Q2J565_FRASC      ----------LSGLLG-LTGLGTSAARIVPSLVTTTPRRIRSFTHIPHQIVGPLRHPDFC 388
Q0RBQ4|Q0RBQ4_FRAAA      GSLVPAA--GLPGIPG-IPSLGTSAIQVLPSLLATAPRRIRSFANVPFRLVGPTPLPEFY 408
Q3WJX6|Q3WJX6_9ACTO      -----------GSLPGDATGLGTSAVRIAASLLATAPARMRAFTHVPHQVVGPVKLPAFY 389
A4FEI8|A4FEI8_SACEN      QSAPDGTGPATQPVLGGPTGLGTSAARIVQSLLSTAPQRIRSFTHTPYEPAGPIRMPEIY 411
A4FGS3|A4FGS3_SACEN      -------------FLSGPHGLGTSSSHLGSLLRTVRPGLPIPHGQLRYARIA---VPAMS 394

Q1CYZ7|Q1CYZ7_MYXXD      MPYSTKPSPHLDAARRDSKAWARRMGMLDVLPGVPGGYIWDDHKFDVADVALCGALIHPH 454
Q09A24|Q09A24_STIAU      MPYSTSVSPHLDAARRHSKEWARQMGMLDSLPGLPGVYIWDDHKFDVADVALCGALIHPE 394
Q9X839|CYC2_STRCO        MPYPLELSPALDGARRHLSEWCREMGILS--EG-----VWDEDKLESCDLPLCAAGLDPD 435
Q82L49|Q82L49_STRAW      MPFQVELSPHLEGARARLTPWMHSTGMLQ--EG-----VWDEDKLTAYDLPLCSAGLDPD 434
Q2J565|Q2J565_FRASC      MPFSTGQSPHLDASRRENIIWARAVGMLDPIPG-----IWDEHKLRAFDFALCSAGIHPD 443
Q0RBQ4|Q0RBQ4_FRAAA      LPYTTGLSPHLDSSRRAIIPWARSMGMLDRVPG-----IWDEHKLWSYDFALCSAGIHPD 463
Q3WJX6|Q3WJX6_9ACTO      MPFTTGESRHLAAARHNIVEWSAAVGFLDPVPG-----IWDEHKLRAADFALCSAAIHPN 444
A4FEI8|A4FEI8_SACEN      MPFDLSLSPHLDVCRENTAAWARAMGIFDDVPR-----VWDENQMRGYDLPLCSAGLDPD 466
A4FGS3|A4FGS3_SACEN      SPHPVRTNPQVGTVRAHAKEWARRMGMLDGSG------VWTANVFDAADFGQFSAMAHPD 448

Q1CYZ7|Q1CYZ7_MYXXD      ATAAQLNLSSCWLVWGTYA**DDYFP**AFYGHTKDMAGAKVFNARLALFVPEDAGAVVPPPTN 514
Q09A24|Q09A24_STIAU      ASAEQLNLTACWLVWGTYA**DDYFP**AFYGYTRDMAGAKLFNARLSAFMPDGPCTAVP--TN 452
Q9X839|CYC2_STRCO        ATQDQLDLASGWLAFGTYG**DDYYP**LVYGHRRDLAAARLTTTRLSDCMPLD-GEPVPPPGN 494
Q82L49|Q82L49_STRAW      ATPDELDLSSRWLAWGTYG**DDYYP**MVFGPRRDLAAAKLCTRRLSACMPVD-GEEVPAPVN 493
Q2J565|Q2J565_FRASC      ATLPELNLTTDWLTWGTYA**DDYYP**VIFGRTRDILGAKVCNARLSEFMPLD-SPVTAVPAN 502
Q0RBQ4|Q0RBQ4_FRAAA      ATADELDLTTAWLTWGTYG**DDYYP**VIFGASRNLAAAKLCNERLRLFMPVD-GPLTEPPVN 522
Q3WJX6|Q3WJX6_9ACTO      ATAAELDLTTGWLTWGTYA**DDLYP**VLYGRTRDLAGARACTERLKELMPVE-PGPLPVPVG 503
A4FEI8|A4FEI8_SACEN      ATPEELDLSAAWLTWGTYG**DDYYP**RVFGRTLDMAGARACNARLKELMPVE-SAPATAPVT 525
A4FGS3|A4FGS3_SACEN      SPGPELELVNDWHVWGWFF**DDFFT**EVFKRSRNRAGAEAFLARLPGFMPAD-TRRTPAPAN 507

Q1CYZ7|Q1CYZ7_MYXXD      PVERGLADLWARTTEGVTPASRSLFRKAILDMTESWVWELANQIQNRIPDPIDYVEMRRQ 574
Q09A24|Q09A24_STIAU      PVEHGLADLWARTAGPMTDNARRLFRKAIQDMTASWLWELANQIQNRIPDPVDYVEMRRK 512
Q9X839|CYC2_STRCO        AMERSLIDLWVRTTAGMTPEERRPLKKAVDDMTEAWLWELSNQIQNRVPDPVDYLEMRRA 554
Q82L49|Q82L49_STRAW      GMERGLIDLWAITTAEMTPDERRTFRASVDVMTESWVWELSNQLQHRIPDPIDYLEMRRA 553
Q2J565|Q2J565_FRASC      ALERGLADLWTRTTETMAPGARETFRGTVEVMIDSWLWELANQAQNRIPDPIDYIEMRRA 562
Q0RBQ4|Q0RBQ4_FRAAA      ALERGLADLWERTGAGMEPAARATFRRTIEVMIDSWLWELANQAHNRIPDPVDYLEMRRA 582
Q3WJX6|Q3WJX6_9ACTO      GLERGLADLWPRTTRDMTPDSRRTFRRTVCIMLDSWQWELANQAQNRIPDPVDYIEMRRR 563
A4FEI8|A4FEI8_SACEN      PLERGLADLWARTAGPMPVETRRRFRAAVDTMIDSWLWELHNQHLNRIPDPVDYFEMRRR 585
A4FGS3|A4FGS3_SACEN      PVERGLADLWARSTPVLAPRLRRRFPEHVRNFVGSWLWELDNLIQNRVSDPVDYLRMRRR 567

Q1CYZ7|Q1CYZ7_MYXXD      TFGSDLTMSLSRLAHG---DALPPEVFHTRPIRSLENSAADYACLI**NDVFSYQKE**IEFEG 631
Q09A24|Q09A24_STIAU      TFGSDLTMSLSRLAHG---DAIPQEIFHTRPVRGLENSAADYACLT**NDIFSYQKE**IEYEG 569
Q9X839|CYC2_STRCO        TFGSDLTLGLCRAGHG---PAVPPEVYRSGPVRSLENAAIDYACLL**NDVFSYQKE**IEYEG 611
Q82L49|Q82L49_STRAW      TFGADLTLSLCRVGHG---PKVPPEIYRSGPVRSLENAAVDYGMLI**NDVFSYQKE**IEYEG 610
Q2J565|Q2J565_FRASC      TFGSDLTMSLARLARLAQEQTVPPEIYRTRPIQALENAAADYACLL**NDVFSYQKE**IQFEG 622
Q0RBQ4|Q0RBQ4_FRAAA      TFGSDLTMSLCRLAR---WHSVPAEVFGTRPLRALENAAADYACLL**NDIFSYQKE**IQFEG 639
Q3WJX6|Q3WJX6_9ACTO      TFGSDLTMSLSRLGHG---RSVPPEIYGTRPIRALENSAADYSCLL**NDIFSYQKE**IQFEG 620
A4FEI8|A4FEI8_SACEN      TFGSDLTISLAKFSHG---EAVPPEIYRTRTIRNMENSAIDYATML**NDVFSYRKE**IEYEG 642
A4FGS3|A4FGS3_SACEN      TGGSAFRGALARHTLG---AGLAPAVFDTPEMRALHENWADVGPLR**NDLFSYHKE**VDRET 624

Q1CYZ7|Q1CYZ7_MYXXD      ELNNGVLVVQRFLDLDPARAVSVVNDLMTARMQQFEYIIANELEPLARNFNLDGKAQDKL 691
Q09A24|Q09A24_STIAU      ELNNGVLVVQRFLEIEPPQAVEIVNDLMTARMRQFEHTVKMELPLLIRSTGLDAKAQEKL 629
Q9X839|CYC2_STRCO        EIHNAVLVVQNFFGVDYPAALGVVQDLMNQRMRQFEHVVAHELPVVYDDFQLSEEARTVM 671
Q82L49|Q82L49_STRAW      EVHNAILVVQNFFGCDYPTALGVINDLMTQRMHQFEHVAAHELPLLYKDFKLPQEVRDIM 670
Q2J565|Q2J565_FRASC      EIHNCVLVVENFLDCDRERALAVVNDLMTSRIRQFEHIVAHELPALFDSFALDASARQAL 682
Q0RBQ4|Q0RBQ4_FRAAA      EIHNCVLVVENFLDCDRGRAVEVVNALMTARMRQFEHVVDRELPDLFDRLDLDGEARAAI 699
Q3WJX6|Q3WJX6_9ACTO      EIHNCVLVFQNFLGCGAERAIGVVNDLMTARLREFEHVVDVELPALFDTYELTEEARDVL 680
A4FEI8|A4FEI8_SACEN      EVHNAVLVVRNFLDCDQDRAFEIVGDLMTARMKQFQYTVDDELPVLCEDFGLSSESRAVL 702
A4FGS3|A4FGS3_SACEN      EVTNGVLAVQRFFDCGLQQAAAVVADLAEVRLRRFTAVAEQELPALAHRFEPGRAPREEL 684
```

```
Q1CYZ7|Q1CYZ7_MYXXD      KQYVQKLQWWMSGVLIWHQTVDRYKEFELRASRK-----------LAPRLSSGPTGLGTS 740
Q09A24|Q09A24_STIAU      RTYVEKLQRWMCGVLRWHMTVDRYKEFELRNTRKPRRGGWEDPRDGAPPRPASRRSLGAT 689
Q9X839|CYC2_STRCO        RGYVTDLQNWMAGILNWHRNVPRYKAEYLAGRTHGFL---------PDRIPAPPVPRSSP 722
Q82L49|Q82L49_STRAW      DGYVVELQNWMSGILKWHQDCHRYGAADLARRAHGFV---------PDRAPSAPFTAWAA 721
Q2J565|Q2J565_FRASC      LGYARELQNWLAGILRWHEGTHRYEESELRYHPAAG-----------VRPFGGPTGLGTS 731
Q0RBQ4|Q0RBQ4_FRAAA      VSYARELQNWLAGILRWHQGTHRYEEAELRYHPAAD-----------RRPFGSPTGLGTS 748
Q3WJX6|Q3WJX6_9ACTO      RGYVGELKSWLAGVLRWHQGTRRYDEAELRHHPAVG-----------VRPFGGPVGLGTS 729
A4FEI8|A4FEI8_SACEN      TRYADELRDWMSGILNWHRECVRYKDEDLRHDAVSQG---------LAALLRGPSGLGTS 753
A4FGS3|A4FGS3_SACEN      DRYVRGLHDWLAGELAWSQVTGRYREPSVSAVGADLP--------------AAPLGITGA 730

Q1CYZ7|Q1CYZ7_MYXXD      AARITSLFANLRSGA------ 755
Q09A24|Q09A24_STIAU      GAEVEKKLEKSGSST------ 704
Q9X839|CYC2_STRCO        ALTH----------------- 726
Q82L49|Q82L49_STRAW      PVAR----------------- 725
Q2J565|Q2J565_FRASC      SAHVRPRPAAAAGAAGDSEM- 751
Q0RBQ4|Q0RBQ4_FRAAA      AADVR-RLASR---------- 758
Q3WJX6|Q3WJX6_9ACTO      AADIRRALSGKSGQPTALTGS 750
A4FEI8|A4FEI8_SACEN      AVELR---------------- 758
A4FGS3|A4FGS3_SACEN      AG------------------- 732
```

Supplementary Figure 5. ClustalW Alignment of presumptive germacradienol–geosmin synthases with *S. coelicolor* SCO6073. Q1CYZ, *Myxococcus xanthus* strain DK 1622; Q09A24, *Stigmatella aurantiaca* DW4/3-1; Q9X839, *S. coelicolor* A3(2); Q82L49, *S. avermitilis*; Q2J565, *Frankia* sp. Strain Cc13; Q0RBQ4, *Frankia alni* ACN14a; Q3WJX6, *Frankia* sp. EAN1pec; A4FEI8 and A4FGS3, *Saccharopolyspora erythraea* NRL 2338. Conserved Mg^{2+}-binding motifs are highlighted in bold.



Supplementary Scheme 1. Cyclization of geranylgeranyl diphosphate (**13**, GGPP) to copalyl diphosphate (**14**, CPP) and *ent*-kaurene (**15**) by bifunctional fungal *ent*-kaurene synthase (KS).

Supplementary References

1. Cane, D.E. & Watt, R.M. Expression and mechanistic analysis of a germacradienol synthase from *Streptomyces coelicolor* implicated in geosmin biosynthesis. *Proc. Natl. Acad. Sci. U S A* **100**, 1547-1551 (2003).

2. Harangi, J. Retention index calculation without n-alkanes – the virtual carbon number. *J. Chromatogr. A* **993**, 187-195 (2003).