ACCOUNTS
of chemical research

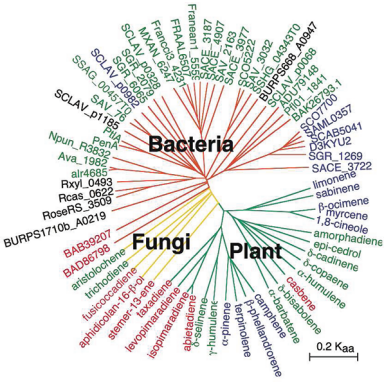
Exploration and Mining of the Bacterial Terpenome

DAVID E. CANE*,‡ AND HARUO IKEDA*,§

‡Department of Chemistry, Box H, Brown University, Providence, Rhode Island 02912-9108, United States, and §Laboratory of Microbial Engineering, Kitasato Institute for Life Sciences, Kitasato University, 1-15-1 Kitasato, Sagamihara, Minami-ku, Kanagawa 252-0373, Japan

RECEIVED ON AUGUST 5, 2011

CONSPECTUS



Tens of thousands of terpenoids are present in both terrestrial and marine plants, as well as fungi. In the last 5−10 years, however, it has become evident that terpenes are also produced by numerous bacteria, especially soil-dwelling Gram-positive organisms such as *Streptomyces* and other Actinomycetes. Although some microbial terpenes, such as geosmin, the degraded sesquiterpene responsible for the smell of moist soil, the characteristic odor of the earth itself, have been known for over 100 years, few terpenoids have been identified by classical structure- or activity-guided screening of bacterial culture extracts. In fact, the majority of cyclic terpenes from bacterial species have only recently been uncovered by the newly developed techniques of "genome mining". In this new paradigm for biochemical discovery, bacterial genome sequences are first analyzed with powerful bioinformatic tools, such as the BLASTP program or Profile Hidden Markov models, to screen for and identify conserved protein sequences harboring a characteristic set of universally conserved functional domains typical of all terpene synthases. Of particular importance is the presence of variants of two universally conserved domains, the aspartate-rich DDXX(D/E) motif and the NSE/DTE triad, (N/D)DXX(S/T)XX(K/R)(D/E). Both domains have been implicated in the binding of the essential divalent cation, typically Mg^{2+}, that is required for cyclization of the universal acyclic terpene precursors, such as farnesyl and geranyl diphosphate.

The low level of overall sequence similarity among terpene synthases, however, has so far precluded any simple correlation of protein sequence with the structure of the cyclized terpene product. The actual biochemical function of a cryptic bacterial (or indeed any) terpene synthase must therefore be determined by direct experiment. Two common approaches are (i) incubation of the expressed recombinant protein with acyclic allylic diphosphate substrates and identification of the resultant terpene hydrocarbon or alcohol and (ii) in vivo expression in engineered bacterial hosts that can support the production of terpene metabolites. One of the most attractive features of the coordinated application of genome mining and biochemical characterization is that the discovery of natural products is directly coupled to the simultaneous discovery and exploitation of the responsible biosynthetic genes and enzymes.

Bacterial genome mining has proved highly rewarding scientifically, already uncovering more than a dozen newly identified cyclic terpenes (many of them unique to bacteria), as well as several novel cyclization mechanisms. Moreover, bioinformatic analysis has identified more than 120 presumptive genes for bacterial terpene synthases that are now ripe for exploration. In this Account, we review a particularly rich vein we have mined in the genomes of two model Actinomycetes, *Streptomyces coelicolor* and *Streptomyces avermitilis*, from which the entire set of terpenoid biosynthetic genes and pathways have now been elucidated. In addition, studies of terpenoid biosynthetic gene clusters have revealed a wealth of previously unknown oxidative enzymes, including cytochromes P450, non-heme iron-dependent dioxygenases, and flavin monooxygenases. We have shown that these enzymes catalyze a variety of unusual biochemical reactions, including two-step ketonization of methylene groups, desaturation−epoxidation of secondary methyl groups, and pathway-specific Baeyer−Villiger oxidations of cyclic ketones.

Introduction

Streptomyces and other Actinomycetes are Gram-positive, soil-dwelling organisms that are prolific producers of an enormous variety of natural products. Until recently, however, reports of the isolation of terpenoid metabolites from Actinomycetes and other bacteria had been surprisingly

Published on the Web 10/31/2011 www.pubs.acs.org/accounts
10.1021/ar200198d

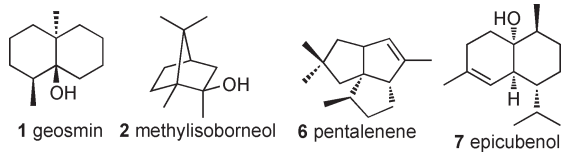
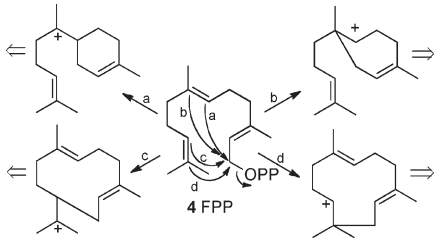


FIGURE 1. Bacterial terpenes.

SCHEME 1. Cyclization of Farnesyl Diphosphate (**4**)



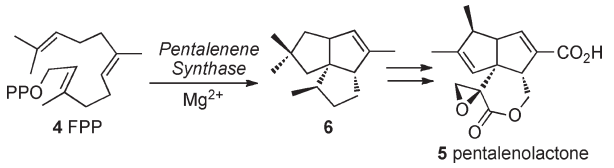
rare. The first terpene to be isolated from *Streptomyces* was the degraded sesquiterpene geosmin (**1**), which is responsible for the characteristic odor of moist soil (Figure 1).[1,2] In addition to detecting geosmin in 17 different species of *Streptomyces*, a cyanobacterium, and a myxobacterium, Gerber also isolated another common metabolite with a characteristic musty, camphor-like odor, identified as 2-methylisoborneol (**2**). In 1981, Bentley reported the first evidence for the likely isoprenoid origin of both geosmin and 2-methylisoborneol produced by *Streptomyces antibioticus*.[3]

The several hundred parent cyclic monoterpene, sesquiterpene, and diterpene hydrocarbons and alcohols are all formed by variations of a common cyclization mechanism initiated by enzyme-catalyzed ionization of the universal acyclic precursors geranyl (**3**, GPP), farnesyl (**4**, FPP), and geranylgeranyl diphosphate (GGPP) (Scheme 1). Intramolecular attack of the resultant allylic cation on the central or distal double bonds followed by well-precedented cationic transformations and eventual quenching of the positive charge by deprotonation or by capture of water can account for the formation of the enormous variety of cyclic terpenes.

Bacterial Terpene Synthases

Pentalenolactone (**5**) is a sesquiterpenoid antibiotic that has been isolated from more than 30 species of *Streptomyces*.[4,5] Pentalenene synthase, which we first isolated from *Streptomcyes exfoliatus* UC5319 and subsequently cloned and expressed in *Escherichia coli*,[6] catalyzes the cyclization of farnesyl diphosphate (**4**) to the parent sesquiterpene hydrocarbon pentalenene (**6**) (Scheme 2). Extensive experiments with stereospecifically labeled FPP have established the detailed mechanism and stereochemistry of the cyclization.[7]

SCHEME 2. Pentalenene Synthase and the Biosynthesis of Pentalenolactone



The crystal structure of pentalenene synthase has also been determined.[8] The only other bacterial terpene synthase to have been isolated by analogous biochemical assay-guided methods is epicubenol (**7**) synthase from *Streptomyces* sp. LL-B7.[9]

"Genome mining" has proven to be a powerful paradigm for the discovery and characterization of natural product biosynthetic genes.[10] The actual biochemical function of a candidate terpene synthase can be assigned by incubation of the recombinant cyclase with the GPP, FPP, or GGPP and identification of the resulting monoterpene, sesquiterpene, or diterpene product. This *in vitro* approach, which simultaneously provides enhanced levels of protein for mechanistic, mutational, and structural studies, can be powerfully complemented by *in vivo* expression of individual genes or clusters of terpenoid biosynthetic genes in suitably engineered host bacteria, followed by identification of the newly generated metabolic products. In this Account, we review our recent application of genome mining to the discovery of new bacterial terpene synthases and the elucidation of terpenoid biosynthetic pathways, focusing on two prototypical Actinomycetes, *S. coelicolor* and *S. avermitilis*.

Terpene Synthase Bioinformatics

Within the more than 20 Actinomycete genome sequences reported to date are more than 100 presumed or confirmed terpene synthases. Bioinformatic analysis alone has been incapable of assigning a specific biochemical function to most newly recognized terpene synthase genes. Despite the very substantial differences in overall primary amino acid sequence, terpene synthases typically display two highly conserved Mg^{2+}-binding domains: an aspartate-rich motif, **DD**XX(**D**/**E**) or **DD**XXX(**D**/**E**), usually found ~80−120 amino acids (aa) downstream of the N-terminus of microbial synthases, and a second "NSE/DTE" triad, (**N**/**D**)DXX(**S**/**T**)XX(K/R)(**D**/**E**), located 140 ± 5 aa downstream of the aspartate-rich motif.[11] Crystallographic studies of nearly a dozen monoterpene and sesquiterpene synthases from bacterial, fungal, and plant sources have established that these two motifs, which are located at opposite sides of the rim of the deep active site cavity, are responsible for cooperative

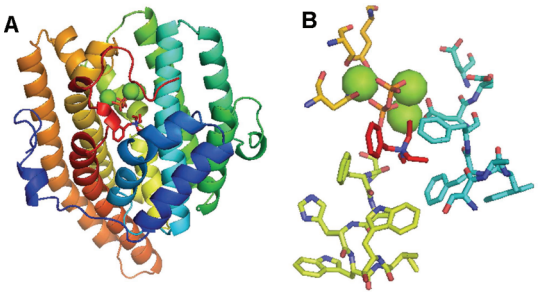


FIGURE 2. Structure of a typical terpene synthase: (A) epi-isozizaene synthase in complex with three Mg^{2+} ions (green spheres), inorganic pyrophosphate (orange), and the benzyltriethylammonium cation (red), showing the universal type I all-α-helical terpene synthase fold; (B) detail of active site, including conserved ^{99}**DD**XX**D** (cyan) and ^{240}**N**XXX**S**XXX**E** (yellow) motifs, as well as aromatic side chains (cyan and lemon), lining the active site cavity.

binding of three divalent cations and the pyrophosphate moiety of the substrate, precisely positioning the acyclic allylic diphosphate substrate and activating it for the ionization that triggers the cyclization cascade (Figure 2).[8,12,13] Despite the significant differences in primary sequence, these terpene cyclases also all share a common α-helical fold. The detailed contour of the cyclase active site is thought to play a major role in the proper folding of the allylic diphosphate substrate and chaperoning the highly reactive carbocationic intermediates so as to control the formation of the characteristic cyclic terpene product.[12]

While local sequence alignment searches using the BLAST algorithm can recognize many presumptive terpene synthases, they may still miss matches of low overall sequence similarity. Hidden Markov models (HMMs) provide a powerful means of codifying the underlying pattern of functional domains and searching for proteins of similar function even in the absence of significant levels of primary sequence similarity.[14,15] The HMM parameters are first estimated from a training set of protein sequences that can include multiple alignments that incorporate three-dimensional structural information. The resulting profile HMM can distinguish members of the relevant protein functional families from nonmembers with a high degree of accuracy.[16−18]

We first used a profile Hidden Markov model from the Pfam database (Pfam entry PF03906), based on the two universally conserved terpene synthase family metal-binding domains, to search the 2008 NCBI database of bacterial genome sequences in order to harvest all the presumptive terpene cyclase sequences.[19] From 1 922 990 predicted proteins, 41 proteins initially selected as strong matches were classified into three distinct groups on the basis of phylogenetic analysis. Group I contained 12 protein sequences

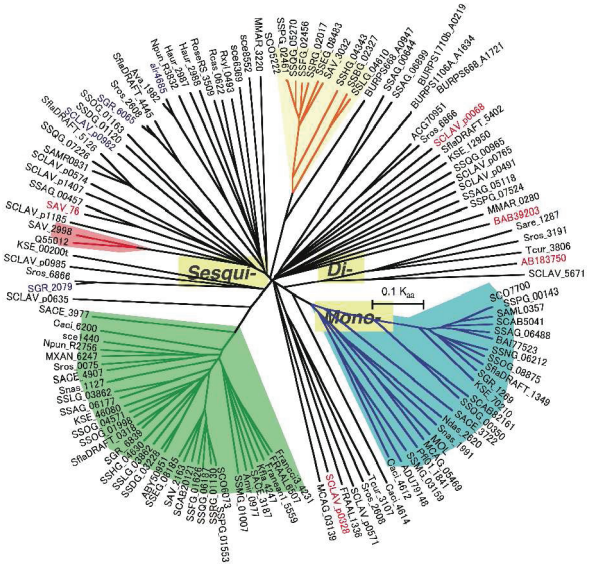


FIGURE 3. Phylogenetic tree of bacterial terpene synthases. Synthases are each labeled by their proton-coding/gene names. The green-colored area indicates a clade of germacradienol/geosmin synthases, the red-colored area corresponds to pentalenene synthases, beige-shaded regions are epi-isozizaene synthases, and the blue-colored zone contains 2-methylisoborneol synthases or 2-methylenebornane synthases. SAV_76 is avermitilol synthase, SCLAV_p0068 (SSCG_03688) is (+)-T-muurolol synthase, SCLAV_0328 (SSCG_02150) is (−)-δ-cadinene synthase, BAB39203 is terpentecin synthase and AB183750 is pimara-9(11),15-diene synthase.

provisionally assigned as monoterpene cyclases. Group II, by far the largest group of 27 protein sequences, included the known sesquiterpene cyclases pentalenene synthase, germacradienol/geosmin synthase, and epi-isozizaene synthase (see below) and close orthologues, as well as several presumptive sesquiterpene synthases of unknown biochemical function. Group III consisted of two diterpene synthase sequences, including a cyclase implicated in terpentecin biosynthesis. These first-pass harvested synthases have since been used to generate a second-generation model specific for bacterial terpene synthases. A phylogenetic tree illustrating the grouping of over 120 candidate bacterial sequences is shown in Figure 3.

In Vitro and *in Vivo* Investigations of Terpene Synthases and Terpenoid Biosynthetic Pathways

The majority of *Streptomyces* proteins, including terpene synthases, can be well-expressed in *E. coli* using either native or synthetic genes. Incubation of the purified recombinant terpene synthase with FPP, GPP, or GGPP generates a cyclized terpene hydrocarbon or alcohol whose structure can be assigned by GC−MS comparison with known

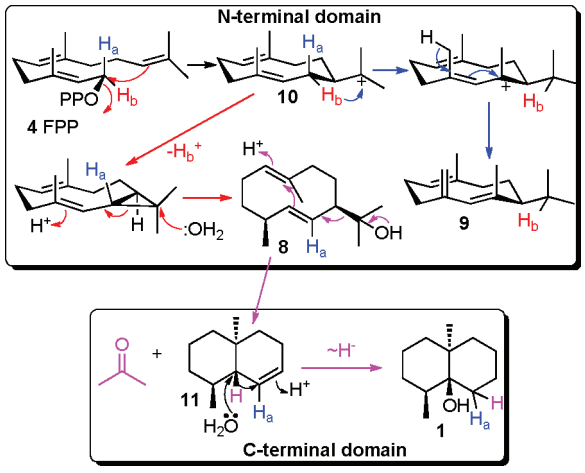
standards or by detailed NMR analysis. A complementary *in vivo* approach is to use a metabolically engineered *E. coli* host designed for the efficient expression of terpenoid metabolites.[20] We have also developed engineered strains of *S. avermitilis* from which ~1 MB of DNA harboring the majority of natural product biosynthetic genes has been deleted.[21−23] These so-called "SUKA" mutants are robust natural product factories that retain the capability of the parent industrial *S. avermitilis* strain to synthesize biological building blocks and export biosynthetic end-products. Indeed, the SUKA mutants are excellent hosts for the introduction of individual biosynthetic genes or even entire biosynthetic gene clusters from *Streptomyces* or even eukaryotic sources, allowing determination of their biochemical function. The resultant transformants can also be used for the production of multi-milligram quantities of otherwise inaccessible, complex biosynthetic intermediates.

Case Studies

I. Terpene Synthases of *Streptomyces coelicolor* A3(2) Geosmin Synthase. The 8.6-Mb linear genome of *S. coelicolor* A3(2), usually considered *the* model *Streptomyces* organism, harbors 7825 predicted genes,[24] of which three correspond to predicted terpene synthases. The 2181-bp *sco6073* gene encodes an unusually large 726-aa protein in which both the N-terminal (366 aa) and C-terminal (339 aa) halves showed ~30% identity to pentalenene synthase. The N-terminal domain harbors variants of the characteristic conserved Mg^{2+}-binding domains, ^{86}**DD**HFL**E** and ^{229}**ND**LF**S**YQR**E**, while the C-terminal half displays an unusual ^{455}**DD**YYP motif, as well as a canonical ^{598}**ND**VF**S**YQK**E** sequence. We initially found that the full-length recombinant protein expressed in *E. coli* catalyzed the Mg^{2+}-dependent cyclization of FPP to an 85:15 mixture of (4*S*,7*R*)-germacra-1(10)*E*,5*E*-diene-11-ol (**8**) and (7*S*)-germacrene D (**9**) (Scheme 3).[25,26] Incubations with chirally deuterated FPP established that both sesquiterpene products resulted from partitioning of a common germacradienyl cation intermediate **10**.[27]

Independently, Gust et al. reported that in-frame deletion of the entire *sco6073* gene resulted in complete loss of production of geosmin.[28] The combined *in vitro* biochemical and *in vivo* molecular genetic evidence thus established that the *sco6073* gene encodes a germacradienol synthase and that formation of germacradienol must be an essential step in the formation of geosmin. We then unexpectedly discovered that incubation of FPP with germacradienol synthase also produced geosmin.[29] Using both the recombinant C-terminal domain of SCO6073 protein and the full-length protein with site-specific mutations in either of the two

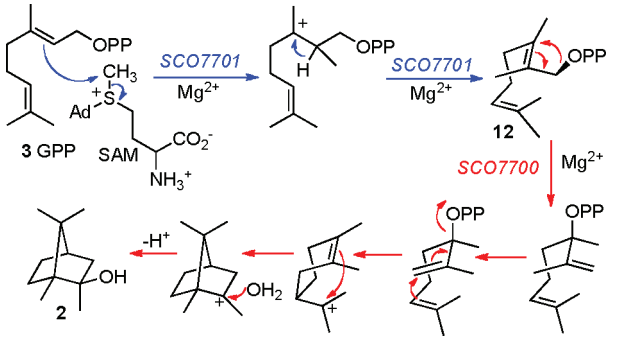
SCHEME 3. Cyclization of Farnesyl Diphosphate by the Bifunctional Geosmin Synthase, SCO6073



Mg^{2+}-binding motifs of the C-terminal domain, we demonstrated that geosmin synthase is in fact a bifunctional protein (Scheme 3).[30] The N-terminal domain catalyzes the cyclization of FPP to germacradienol (**8**) and germacrene D (**9**), while the C-terminal domain is responsible for the proton-initiated sequential retro-Prins fragmentation of germacradienol (**8**) with loss of the 2-propanol side chain as acetone to give the intermediate octalin **11**, which is then converted to geosmin (**1**).[25,31,32]

The geosmin synthase gene is highly conserved, with more than 50 orthologues found in the genomes of a variety Actinomycetes, myxobacteria, and cyanobacteria that display 45−99% identity to SCO6073.[33,34] Incubation of recombinant SAV_2163 protein (GeoA) from *S. avermitilis* with FPP produced a mixture of germacradienol and geosmin while the *geoA* mutant no longer produced geosmin.[35] Curiously, the erythromycin producer *Saccharopolyspora erythraea* harbors *three* distinct geosmin synthase orthologues. The production of geosmin and 2-methylisoborneol by cyanobacteria is responsible for episodes of unpleasant taste and odor in public water supplies as well as an off-taste in fish and other products of aquaculture. We have identified the geosmin synthase of the model cyanobacterium *Nostoc punctiforme* by incubation of the recombinant protein with FPP.[36] Despite the ubiquity of geosmin-producing microorganisms and a variety of imaginative suppositions as to the possible role of this volatile organic metabolite (for example in attracting various species, from earthworms to camels),[37] there has been only limited experimental support for these speculations. For example, glass eels (*Anguila anguila*), which migrate from salt water to fresh water, are strongly

SCHEME 4. Biosynthesis of 2-Methylisoborneol

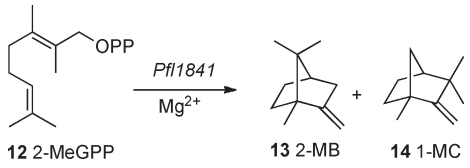


attracted by geosmin which is thus thought to serve as an important inland water marker (although this would have no obvious evolutionary benefit to *Streptomyces*).[38]

2-Methylisoborneol Synthase. The 1323-bp *sco7700* gene encodes a 440-aa protein with <20% identity to pentalenene synthase that incorporates variants of the two characteristic terpene synthase Mg^{2+}-binding motifs, an unusual ^{197}**DD**CYC**ED** acidic motif and a more conventional downstream NSE triad ^{345}**ND**LY**S**YTK**E**. Intriguingly, the 3′-end of the *sco7700* gene is separated by only 16-bp from the downstream coding sequence, *sco7701*, which had been annotated only as a generic *C*-methyltransferase. Recalling that Bentley had reported that 2-methylisoborneol is labeled by [^{14}C]-methyl]methionine,[3] it appeared that we might be dealing with a two-gene biosynthetic operon for the formation of 2-methylisoborneol (**2**). Indeed, incubation of a reconstituted mixture of recombinant SCO7700 and SCO7701 with GPP (**3**) and *S*-adenosyl-L-methionine (SAM) gave 2-methylisoborneol (**2**) as the major product (Scheme 4).[39] Incubation of GPP and SAM with recombinant SCO7701 gave exclusively the previously unknown acyclic substrate (*E*)-2-methylgeranyl diphosphate (**12**, 2-MeGPP). Finally, direct incubation of synthetic **12** with SCO7700 gave 2-methylisoborneol (**2**). These results are consistent with the reported incorporation of labeled mevalonate and methionine into 2-methylisoborneol by the myxobacterium *Nannocystis exedens*, as well as the isolation of 2-methylgeraniol from this organism.[40]

In the meantime, the Kitasato group had carried out an independent analysis of bacterial 2-methylisoborneol biosynthesis.[19] Significantly, the majority of the 12 genes for the group I proteins tentatively identified as monoterpene synthases by the profile HMM search (Figure 3) were flanked by a gene encoding a predicted SAM-dependent *C*-methyltransferase. Indeed, five of the strains harboring these gene pairs produced 2-methylisoborneol, while a sixth, *Micromonospora olivasterospora*, generated the corresponding

SCHEME 5. Biosynthesis of 2-Methylenebornane (**13**) and 1-Methylcamphene (**14**)

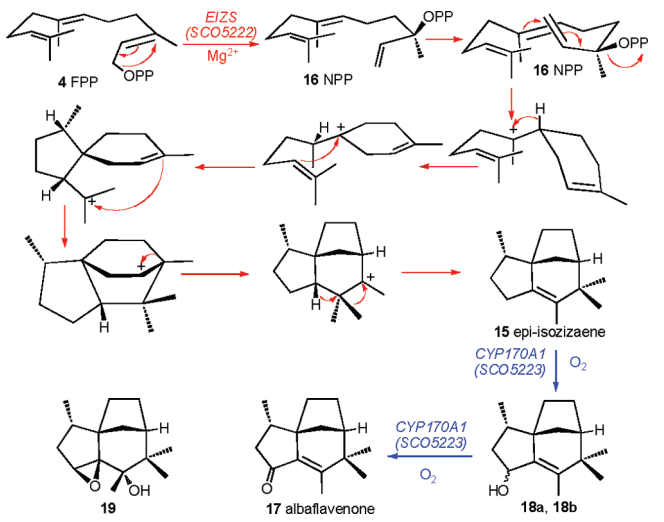


dehydrated homo-monoterpene, 2-methylenebornane (**13**, 2-MB). Furthermore, when each of the two gene segments from *Streptomyces ambofaciens*, *Streptomyces lasaliensis*, and *Sac. erythraea* were inserted into *S. avermitilis* SUKA16 downstream of the strong, constitutive *rpsJ* (*sav4925*) promoter, the resultant exconjugants each produced 2-methylisoborneol. Finally, incubation of recombinant *S. lasaliensis* 2-methylisoborneol synthase and the 2-MeGPP synthase with GPP and SAM gave 2-methylisoborneol (**2**), while incubation of the same substrates in the absence of the monoterpene cyclase, followed by phosphatase treatment, gave (*E*)-2-methylgeraniol. The preliminary X-ray crystallographic analysis of the *S. lasaliensis* 2-MeGPP synthase has recently been reported.[41] A new profile HMM for the region around two conserved metal-binding motifs of microbial terpene synthases also identified a single open reading frame of 397 aa within the genome sequence of the cyanobacterium *Pseudanabaena limonetica*.[42] Heterologous expression of the corresponding gene product has confirmed the identity of the cyanobacterial 2-methylisoborneol synthase, as well as the coupled 2-MeGPP synthase.

We have recently found that the Pfl_1841 protein of *Pseudomonas fluorescens* PfO-1 catalyzes the cyclization of 2-MeGPP (**12**) to a mixture of 2-methylenebornane (**13**) and 1-methylcamphene (**14**) (Scheme 5).[43] A truncated recombinant form of *M. olivasterospora* 2-methylenebornane synthase, which is clustered in the phylogenetic tree with Pfl_1841 (Figure 3), also catalyzes the formation of 2-methylenebornane from 2-MeGPP (H. Ikeda, unpublished).

Epi-isozizaene Synthase and the Biosynthesis of Albaflavenone. The *S. coelicolor* A3(2) *sco5222* gene encodes a 361-aa protein with only 24% identity to *S. exfoliatus* pentalenene synthase. The encoded SCO5222 protein harbors two simple variants of the conserved Mg^{2+}-binding domains, an aspartate-rich ^{99}**DD**RH**D** and the downstream triad ^{240}**N**DLC**S**LPK**E**. Incubation of the recombinant protein with FPP gave as the major product a novel tricyclic sesquiterpene hydrocarbon identified as (+)-epi-isozizaene (**15**) (Scheme 6).[44] Using stereospecifically deuterated FPP showed that the cyclization to give **15** takes place with net retention of configuration at C-1 of FPP through the

SCHEME 6. Biosynthesis of Epi-isozizaene (**15**) and Albaflavenone (**17**)



demonstrated intermediacy of (3*R*)-nerolidyl diphosphate (**16**, NPP) (Scheme 6).[44,45] The 1.60 Å resolution X-ray crystal structure of recombinant *S. coelicolor* A3(2) epi-isozizaene synthase in complex with three Mg^{2+} ions, inorganic pyrophosphate, and the carbocation analogue benzyltriethylammonium cation (BTAC) showed the universal class I α-helical terpene synthase fold in a closed conformation with an active site contour that closely complements the natural epi-isozizaene product (Figure 2).[46]

The *S. coelicolor* A3(2) *sco5222* gene shares a four-nucleotide ATGA transcriptional overlap at its 3′-end with the *sco5223* gene encoding the cytochrome P450 CYP170A1. Purified recombinant CYP170A1 catalyzed an unusual series of two consecutive allylic oxidations to convert epi-isozizaene (**15**) to the camphoraceous antibiotic albaflavenone (**17**) by way of an epimeric mixture of albaflavenols (**18a**, **18b**) (Scheme 6).[47] GC-MS analysis of *S. coelicolor* A3(2) cultures established the time-dependent appearance of epi-isozizaene (**15**), both albaflavenols (**18a**, **18b**), and albaflavenone (**17**), while disruption of the CYP170A1 gene abolished formation of all three oxidation products. The crystal structure of CYP170A1 revealed that epi-isozizaene binds in the active site in two orientations with respect to the heme (Figure 4).[48]

The orthologous SAV_3032 protein of *S. avermitilis* also catalyzed the expected cyclization of FPP to epi-isozizaene.[23] Importantly, *sav3032* also shares a four-nucleotide ATGA overlap with the downsteam *sav3031* gene encoding a 456-aa cytochrome P450, CYP170A2. When this *sav3032* gene was placed under control of the *rpsJ* promoter in *S. avermitilis* SUKA16, the resultant transformants

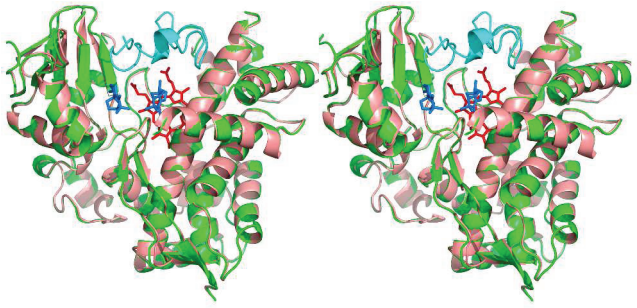


FIGURE 4. Stereoview of overlaid ligand-free CYP170A1 (SCO5223, salmon) and ligand-bound structures (marine) (heme in red; ligand epi-isozizaene in blue). The closed BC loop in the epi-isozizaene complex structure is highlighted in cyan. Reproduced with permission from Figure 1 of ref 48. Copyright 2009 American Society for Biochemistry and Molecular Biology.

accumulated epi-isozizaene. Coexpression of the *sav3031* gene for CYP170A2 generated both (4*R*)- and (4*S*)-albaflavenol (**18a**, **18b**) and albaflavenone (**17**), as well as a previously unknown epoxy-alcohol **19**.

The two-gene operon for albaflavenone biosynthesis is highly conserved and apparently widely distributed, with orthologous pairs of transcriptionally coupled synthase/P450 genes with 56−100% identity evident in 10 species of *Streptomyces*, nearly half of all those with reported genome sequences.[23]

II. Terpene Synthases of *Streptomyces avermitilis* MA-4680. *Streptomyces avermitilis* is responsible for the production of the widely used anthelmintic polyketide avermectins. The 9.03-Mb linear genome of *S. avermitilis* harbors four terpene synthases among its complement of 7575 predicted proteins.[49] In addition to geosmin synthase (SAV_2163)[40] and epi-isozizaene synthase (SAV_3032),[23] these include a newly discovered cyclase, avermitilol synthase (SAV_76), and a pentalenene synthase (PtlA, SAV_2998).

Avermitilol Synthase. The *sav76* gene of *S. avermitilis* encodes a 335-aa protein with 35% identity to *S. exfoliatus* pentalenene synthase harboring the conserved aspartate-rich ^{80}**DD**QF**D** and the triad ^{239}**N**DVY**S**LEK**E**. Other than a single predicted ortholog from *Streptomyces* sp. Mg1 (78% identity), there are no other significant matches among predicted bacterial gene products. Incubation of the purified recombinant SAV76 protein with FPP gave a mixture consisting of a novel sesquiterpene alcohol, avermitilol (**20**, 85%), accompanied by the known isomer viridiflorol (**21**, 3%), as well as germacrenes A (**22**, 1%) and B (**23**, 5%) (Scheme 7).[22] The mechanism and stereochemistry of this unusual cyclization has been experimentally established. Cultures of *S. avermitilis* SUKA17 harboring *sav76* under

SCHEME 7. Cyclization of FPP to Avermitilol (**20**)

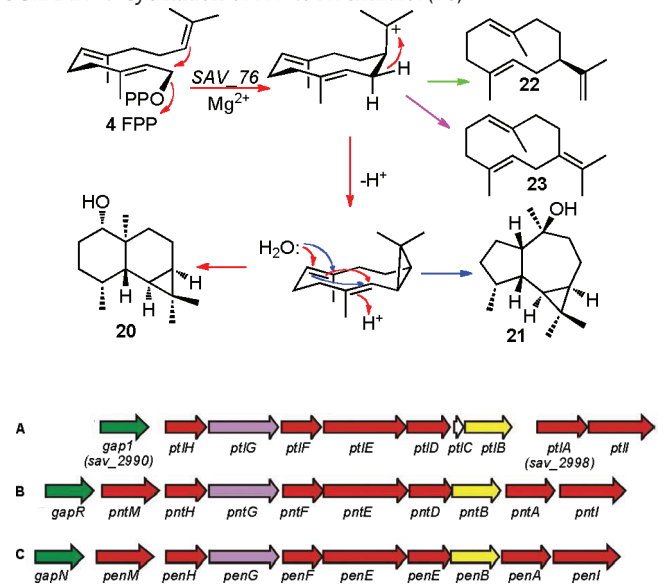
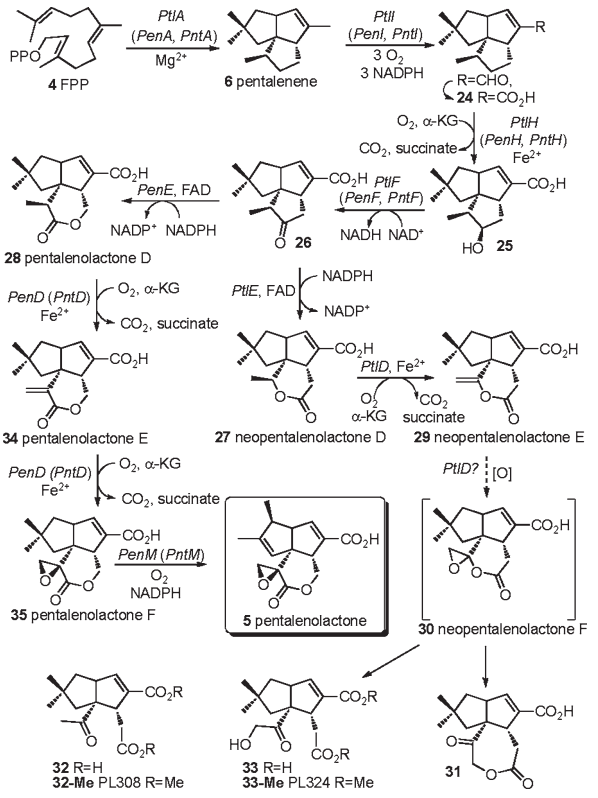


FIGURE 5. Neopentalenolactone and pentalenolactone biosynthetic gene clusters: (A) *S. avermitilis ptl* gene cluster for neopentalenolactone biosynthesis; (B) *Streptomyces arenae pnt* gene cluster for pentalenolactone biosynthesis; (C) *S. exfoliatus pen* gene cluster for pentalenolactone biosynthesis. Biosynthetic genes in red, pentalenolactone-insensitive *gapdh* in green, and FPP synthase in yellow.

control of the *rpsJ* promoter produced avermitilol (15%) along with viridiflorol (2%) and the oxidized derivative of **20**, avermitilone (67%).

Pentalenene Synthase and Discovery of the Neopentalenolactone Biosynthetic Pathway. The *sav_2998* gene of *S. avermitilis* encodes a 336-aa protein (PtlA) with 76% identity to the well-characterized pentalenene synthase of *S. exfoliatus* UC5319. Incubation of recombinant PtlA with FPP gave pentalenene (**6**).[50] The *ptlA* gene is itself located within a 13.4-kb *ptl* gene cluster encoding 13 predicted protein coding sequences (CDSs) that we initially thought might represent the set of biosynthetic genes for pentalenolactone itself (Figure 5). Indeed, we established that *sav2990* (*gap1*), located at the 5′-end of the cluster, is a resistance gene that encodes a pentalenolactone-insensitive glyceraldehyde-3-phosphate dehydrogenase.[50] By systematic expression of each of the individual CDSs of the *ptl* gene cluster, we demonstrated that PtlI (SAV_2999) is a cytochrome P450 catalyzing the three-step oxidation of pentalenene to pentalenal and, almost certainly, to 1-deoxypentalenic acid (**24**),[51] that PtlH (SAV_2991) is an α-ketoglutarate- and non-heme iron-dependent dioxygenase that mediates the 11-β-hydroxylation of **24** to give **25**,[52,53] and that PtlF (SAV_2993) catalyzes the NAD^+-dependent dehydrogenation of **25** to the cyclopentanone **26** (Scheme 8).[54]

SCHEME 8. Biosynthesis of Pentalenolactone (**5**) and Neopentalenolactone F (**30**)



Unexpectedly, however, incubation of **26** with the flavin-dependent Baeyer–Villiger monooxygenase PtlE (SAV_2994) gave the previously unknown metabolite neopentalenolactone D **(27)**, an isomer of the expected product pentalenolactone D (**28**).[55] We have also established that PtlD (SAV_2995) is an α-ketoglutarate- and non-heme iron-dependent dioxygenase that catalyzes the desaturation of **28** to a second novel metabolite, neopentalenolactone E (**29**) (Scheme 8).[56] The *S. avermitilis ptl* gene cluster is thus likely responsible for biosynthesis of a previously unknown metabolite, predicted to be neopentalenolactone F (**30**). In fact, cultures of *S. avermitilis* SUKA5 were found to produce neopentalenoketolactone (**31**), most likely resulting from rearrangement of **30**.[55] When the *ptl* gene cluster was placed under control of the strong constitutive *ermE* promoter, the resulting cultures of SUKA16 accumulated two new metabolites, **32** and **33**, characterized as the derived methyl esters PL308 (**32-Me**) and PL324 (**33-Me**), which are presumably formed by hydrolysis of **29** and **30**, respectively. Notably, the *S. avermitilis* Δ*ptlD* mutant was blocked in formation of **29**, **32**, and **33** but still produced neopentalenolactone D (**27**), while the corresponding double deletion mutant, *S. avermitilis* Δ*ptlE*Δ*ptlD*, lacked **27**, accumulating instead its precursor, ketone **26** (Scheme 8).

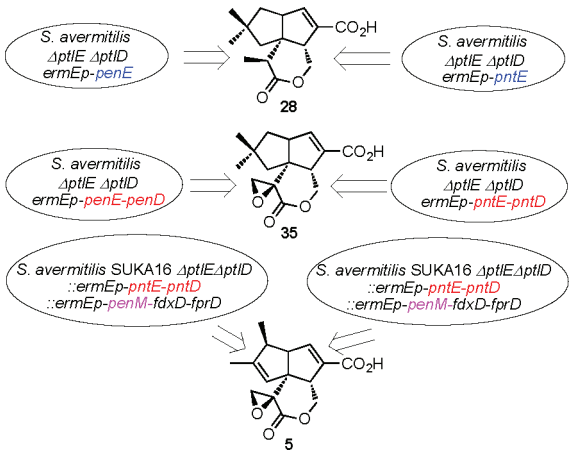


FIGURE 6. Engineered biosynthesis of pentalenolactone metabolites by complementation of *S. avermitilis* SUKA Δ*ptlE*Δ*ptlD* mutants.

III. Pentalenolactone Biosynthesis in *S. exfoliatus* UC5319 and *Streptomyces arenae* TU469. We have cloned and sequenced the complete pentalenolactone biosynthetic gene clusters from two established producers of the antibiotic, *S. exfoliatus* UC5319 and *S. arenae* TU469.[56] Each of the two clusters harbors a homologous set of 10 unidirectionally transcribed CDSs, including close orthologues of each of the previously characterized gene products in the *S. avermitilis ptl* gene cluster (Figure 5).

The recombinant enzymes PenE and PntE from *S. exfoliatus* and *S. arenae*, respectively, each catalyze the flavin-dependent Baeyer–Villiger oxidation of **26** exclusively to pentalenolactone D (**28**). Introduction of the corresponding *penE* and *pntE* genes into the double deletion mutant, *S. avermitilis* Δ*ptlE*Δ*ptlD*, resulted in heterologous production of **28** (Figure 6). Recombinant PenD and PntD each catalyzed the α-ketoglutarate- and non-heme iron-dependent, two-step desaturation/epoxidation of **28** to pentalenolactones E (**34**) and F (**35**) (Scheme 8). The corresponding *S. exfoliatus* Δ*penD* and *S. arenae* Δ*pntD* deletion mutants both accumulated **28** but were unable to produce either **34**, **35**, or **5**, while complementation of the *S. avermitilis* Δ*ptlE*Δ*ptlD* double deletion mutant with either *penE* plus *penD* or *pntE* plus *pntD* led to heterologous production of pentalenolactone F (**35**) (Figure 6).

The recombinant cytochromes P450, PenM, and PntM, each catalyzed the unusual oxidative rearrangement of **35** to pentalenolactone (**5**) (Scheme 8).[57] In support of these observations, the *S. exfoliatus* Δ*penM* and *S. arenae* Δ*pntM* deletion mutants both accumulated **35** and were blocked in the formation of pentalenolactone, which could be restored by complementation with either *penM* or *pntM*. Similarly, complementation of the *S. avermitilis* Δ*ptlE*Δ*ptlD*::*ermE-pntE-pntD* mutant with *penM* or *pntM* also resulted in the production of pentalenolactone (**5**) (Figure 6).

Conclusions

Mining of the rapidly emerging number of bacterial genome sequences has already uncovered a treasure trove of new terpene synthases,[58–63] previously unknown terpene metabolites, and novel variations on the universal terpene cyclization mechanisms, while providing powerful genetic and structural tools to investigate the biological roles of these ubiquitous yet still poorly understood natural products. The combined application of bioinformatic analysis, microbial genetics, mechanistic enzymology, and structural biology should continue to uncover the interrelationships among terpene synthase primary sequence, functional domain organization, protein structure, and cyclization mechanism and specificity, for not only bacterial but fungal and plant terpene synthases as well.

The work carried out in each of our laboratories was supported by NIH grant GM30301 (D.E.C.) and by a Grant-in-Aid for Scientific Research on Innovative Areas from MEXT Japan, from JSPS 2031-0122 and from the Institute for Fermentation, Osaka, Japan (H.I.). We each thank the dedicated co-workers who carried out the studies that are described, as well as Prof. Satoshi Omura for his advice and encouragement in the study of the S. avermitilis terpene synthases.

BIOGRAPHICAL INFORMATION

David E. Cane was born in New York in 1944. After undergraduate study at Harvard (B.A., 1966), he earned his Ph.D. in 1971 from Harvard University for research in synthetic organic chemistry with Prof. E. J. Corey. After postdoctoral study in natural products biosynthesis and bioorganic chemistry with Prof. Duilio Arigoni at the Eidgenössiche Technische Hochschule in Zürich, he joined the faculty of Brown University in 1973 where he is now Vernon K. Krieble Professor of Chemistry and Professor of Biochemistry. His research interests span the fields of organic chemistry and biochemistry and are focused on the biochemical origins of naturally occurring substances, including antibiotics, vitamins, and common flavor and odor compounds. He is particularly interested in unraveling the mechanistic and stereochemical details of multistep biochemical transformations catalyzed by a single enzyme (for example, a terpene synthase) or a multifunctional enzyme (such as a modular polyketide synthase) in which the reaction intermediates remain sequestered in the enzyme active site or sites, either non-covalently or covalently. His awards include the Ernest Guenther Award (1985), the Cope Scholar Award (2000), and the Repligen Award (2005) of the American Chemical Society. He is a Fellow of the American Association of the Advancement of Science.

Haruo Ikeda was born in Tokyo, Japan (1954). He received his B.S. (1977) and M.S. (1979) in pharmaceutical sciences at Kitasato University. He obtained his Ph.D. (1982) in pharmaceutical sciences from Kitasato University, where he studied biosynthesis of 16-membered macrolide antibiotics produced by *Streptomyces* spp. with Professor Satoshi Omura. He began studying *Streptomyces* genetics with Professors Sir David A. Hopwood and Keith F. Chater at the Department of Genetics, John Innes Institute, U.K., as a postdoctoral fellow working on the development of actinophage vectors and characteristics of the gene involving catabolite regulation in *Streptomyces*. He then took a faculty position in the School of Pharmaceutical Sciences, Kitasato University (1983−2002), and is a now a full Professor (2002−present) at Kitasato Institute for Life Sciences and the Graduate School of Infection Control Sciences, Kitasato University. His research interests are in the study of the biosynthesis of microbial secondary metabolites using bioinformatics and in engineering and designing the *Streptomyces* chromosome for the evaluation and optimization of secondary metabolism. He received the Japan Award for Scientific Promotion (1991) and the Japan Award (1999) from the Society for Actinomycetes and the Sumiki-Umezawa Memorial Award of the Japan Antibiotics Research Association (2000).

FOOTNOTES

*E-mail addresses: David_Cane@brown.edu; ikeda@ls.kitasato-u.ac.jp.

REFERENCES

1 Berthelot, M.; André, G. Sur l'odeur propre de la terre. *Compt. Rend.* **1891**, *112*, 598–599.

2 Gerber, N. N. Geosmin, from microorganisms, is *trans*-1,10-dimethyl-*trans*-9-decalol. *Tetrahedron Lett.* **1968**, 2971–2974.

3 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.

4 Martin, D. G.; Slomp, G.; Mizsak, S.; Duchamp, D. J.; Chidester, C. G. The structure and absolute configuration of pentalenolactone (PA 132). *Tetrahedron Lett.* **1970**, 4901–4904.

5 Takeuchi, S.; Ogawa, Y.; Yonehara, H. The structure of pentalenolactone (PA-132). *Tetrahedron Lett.* **1969**, 2737–2740.

6 Cane, D. E.; Sohng, J.-K.; Lamberson, C. R.; Rudnicki, S. M.; Wu, Z.; Lloyd, M. D.; Oliver, J. S.; Hubbard, B. R. Pentalenene synthase. Purification, molecular cloning, sequencing and high-level expression in *Escherichia coli* of a terpenoid cyclase from *Streptomyces* UC5319. *Biochemistry* **1994**, *33*, 5846–5857.

7 Cane, D. E.; Oliver, J. S.; Harrison, P. H. M.; Abell, C.; Hubbard, B. R.; Kane, C. T.; Lattman, R. The Biosynthesis of pentalenene and pentalenolactone. *J. Am. Chem. Soc.* **1990**, *112*, 4513–4524.

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

9 Cane, D. E.; Tandon, M.; Prabhakaran, P. C. Epicubenol synthase and the enzymic cyclization of farnesyl diphosphate. *J. Am. Chem. Soc.* **1993**, *115*, 8103–8106.

10 Corre, C.; Challis, G. L. New natural product biosynthetic chemistry discovered by genome mining. *Nat. Prod. Rep.* **2009**, *26*, 977–986.

11 Felicetti, B.; Cane, D. E. Aristolochene synthase: Mechanistic analysis of active site residues by site-directed mutagenesis. *J. Am. Chem. Soc.* **2004**, *126*, 7212–7221.

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

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

14 Krogh, A.; Brown, M.; Mian, I. S.; Sjolander, K.; Haussler, D. Hidden Markov models in computational biology. Applications to protein modeling. *J. Mol. Biol.* **1994**, *235*, 1501–1531.

15 Eddy, S. R. Profile Hidden Markov models. *Bioinformatics (Oxford, England)* **1998**, *14*, 755–763.

16 Sonnhammer, E. L.; Eddy, S. R.; Birney, E.; Bateman, A.; Durbin, R. Pfam: Multiple sequence alignments and HMM-profiles of protein domains. *Nucleic Acids Res.* **1998**, *26*, 320–322.

17 Bateman, A.; Coin, L.; Durbin, R.; Finn, R. D.; Hollich, V.; Griffiths-Jones, S.; Khanna, A.; Marshall, M.; Moxon, S.; Sonnhammer, E. L.; Studholme, D. J.; Yeats, C.; Eddy, S. R. The Pfam protein families database. *Nucleic Acids Res.* **2004**, *32*, D138–141.

18 Finn, R. D.; Mistry, J.; Tate, J.; Coggill, P.; Heger, A.; Pollington, J. E.; Gavin, O. L.; Gunasekaran, P.; Ceric, G.; Forslund, K.; Holm, L.; Sonnhammer, E. L.; Eddy, S. R.; Bateman, A. The Pfam protein families database. *Nucleic Acids Res.* **2010**, *38*, D211–222.

19 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.

20 Harada, H.; Yu, F.; Okamoto, S.; Kuzuyama, T.; Utsumi, R.; Misawa, N. Efficient synthesis of functional isoprenoids from acetoacetate through metabolic pathway-engineered *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **2009**, *81*, 915–925.

21 Komatsu, M.; Uchiyama, T.; Omura, S.; Cane, D. E.; Ikeda, H. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 2646–2651.

22 Chou, W. K.; Fanizza, I.; Uchiyama, T.; Komatsu, M.; Ikeda, H.; Cane, D. E. Genome mining in *Streptomyces avermitilis*: Cloning and characterization of SAV_76, the synthase for a new sesquiterpene, avermitilol. *J. Am. Chem. Soc.* **2010**, *132*, 8850–8851.

23 Takamatsu, S.; Lin, X.; Nara, A.; Komatsu, M.; Cane, D. E.; Ikeda, H. Characterization of a silent sesquiterpenoid biosynthetic pathway in *Streptomyces avermitilis* controlling epi-isozizaene and albaflavenone biosynthesis and isolation of a new oxidized epi-isozizaene metabolite. *Microb. Biotechnol.* **2011**, *4*, 184–191.

24 Bentley, S. D.; Chater, K. F.; Cerdeno-Tarraga, A. M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Kieser, H.; Harper, D.; Bateman, A.; Brown, S.; Chandra, G.; Chen, C. W.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Huang, C. H.; Kieser, T.; Larke, L.; Murphy, L.; Oliver, K.; O'Neil, S.; Rabbinowitsch, E.; Rajandream, M. A.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Wietzorrek, A.; Woodward, J.; Barrell, B. G.; Parkhill, J.; Hopwood, D. A. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **2002**, *417*, 141–147.

25 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.* **2003**, *100*, 1547–1551.

26 Gansser, D.; Pollak, F. C.; Berger, R. G. A sesquiterpene alcohol from *Streptomyces citreus* CBS 109.60. *J. Nat. Prod.* **1995**, *58*, 1790–1793.

27 He, X.; Cane, D. E. Mechanism and stereochemistry of the germacradienol/germacrene D synthase of *Streptomyces coelicolor* A3(2). *J. Am. Chem. Soc.* **2004**, *126*, 2678–2679.

28 Gust, B.; Challis, G. L.; Fowler, K.; Kieser, T.; Chater, K. F. 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.

29 Jiang, J.; He, X.; Cane, D. E. Geosmin biosynthesis. *Streptomyces coelicolor* germacradienol/germacrene D synthase converts farnesyl diphosphate to geosmin. *J. Am. Chem. Soc.* **2006**, *128*, 8128–8129.

30 Jiang, J.; He, X.; Cane, D. E. Biosynthesis of the earthy odorant geosmin by a bifunctional *Streptomyces coelicolor* enzyme. *Nat. Chem. Biol.* **2007**, *3*, 711–715.

31 Jiang, J.; Cane, D. E. Geosmin biosynthesis. Mechanism of the fragmentation-rearrangement in the conversion of germacradienol to geosmin. *J. Am. Chem. Soc.* **2008**, *130*, 428–429.

32 Nawrath, T.; Dickschat, J. S.; Muller, R.; Jiang, J.; Cane, D. E.; Schulz, S. Identification of (8*S*,9*S*,10*S*)-8,10-dimethyl-1-octalin, a key intermediate in the biosynthesis of geosmin in bacteria. *J. Am. Chem. Soc.* **2008**, *130*, 430–431.

33 Singh, B.; Oh, T. J.; Sohng, J. K. Exploration of geosmin synthase from *Streptomyces peucetius* ATCC 27952 by deletion of doxorubicin biosynthetic gene cluster. *J. Ind. Microbiol. Biotechnol.* **2009**, *36*, 1257–1265.

34 Oliynyk, M.; Samborskyy, M.; Lester, J. B.; Mironenko, T.; Scott, N.; Dickens, S.; Haydock, S. F.; Leadlay, P. F. Complete genome sequence of the erythromycin-producing bacterium *Saccharopolyspora erythraea* NRRL23338. *Nat. Biotechnol.* **2007**, *25*, 447–453.

35 Cane, D. E.; 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.

36 Giglio, S.; Jiang, J.; Saint, C. P.; Cane, D. E.; Monis, P. T. Isolation and characterization of the gene associated with geosmin production in cyanobacteria. *Environ. Sci. Technol.* **2008**, *42*, 8027–8032.

37 Simons, P. Camels act on a hump. *The Guardian* 6 Mar 2003.

38 Tosi, L.; Sola, C. Role of geosmin, a typical inland water odour, in guiding glass eels *Anguilla anguilla* (L.) migration. *Ethology* **1993**, *95*, 177–185.

39 Wang, C. M.; Cane, D. E. Biochemistry and molecular genetics of the biosynthesis of the earthy odorant methylisoborneol in *Streptomyces coelicolor*. *J. Am. Chem. Soc.* **2008**, *130*, 8908–8909.

40 Dickschat, J. S.; Nawrath, T.; Thiel, V.; Kunze, B.; Muller, R.; Schulz, S. Biosynthesis of the off-flavor 2-methylisoborneol by the myxobacterium *Nannocystis exedens*. *Angew. Chem., Int. Ed. Engl.* **2007**, *46*, 8287–8290.

41 Ariyawutthiphan, O.; Ose, T.; Tsuda, M.; Gao, Y.; Yao, M.; Minami, A.; Oikawa, H.; Tanaka, I. Crystallization and preliminary X-ray crystallographic study of a methyltransferase involved in 2-methylisoborneol biosynthesis in *Streptomyces lasaliensis*. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2011**, *67*, 417–420.

42 Giglio, S.; Chou, W. K.; Ikeda, H.; Cane, D. E.; Monis, P. T. Biosynthesis of 2-methylisoborneol in cyanobacteria. *Environ. Sci. Technol.* **2011**, *45*, 992–998.

43 Chou, W. K. W.; Ikeda, H.; Cane, D. E. Cloning and characterization of Pfl_1841, a 2-methylenebornane synthase in *Pseudomonas fluorescens* Pf0-1. *Tetrahedron* **2011**, *67*, 6627–6632.

44 Lin, X.; Hopson, R.; Cane, D. E. Genome mining in *Streptomyces coelicolor*: Molecular cloning and characterization of a new sesquiterpene synthase. *J. Am. Chem. Soc.* **2006**, *128*, 6022–6023.

45 Lin, X.; Cane, D. E. Biosynthesis of the sesquiterpene antibiotic albaflavenone in *Streptomyces coelicolor*. Mechanism and stereochemistry of the enzymatic formation of epi-isozizaene. *J. Am. Chem. Soc.* **2009**, *131*, 6332–6333.

46 Aaron, J. A.; Lin, X.; Cane, D. E.; Christianson, D. W. Structure of epi-isozizaene synthase from *Streptomyces coelicolor* A3(2), a platform for new terpenoid cyclization templates. *Biochemistry* **2010**, *49*, 1787–1797.

47 Zhao, B.; Lin, X.; Lei, L.; Lamb, D. C.; Kelly, S. L.; Waterman, M. R.; Cane, D. E. Biosynthesis of the sesquiterpene antibiotic albaflavenone in *Streptomyces coelicolor* A3(2). *J. Biol. Chem.* **2008**, *283*, 8183–8189.

48 Zhao, B.; Lei, L.; Vassylyev, D. G.; Lin, X.; Cane, D. E.; Kelly, S. L.; Yuan, H.; Lamb, D. C.; Waterman, M. R. Crystal structure of albaflavenone monooxygenase containing a moonlighting terpene synthase active site. *J. Biol. Chem.* **2009**, *284*, 36711–36719.

49 Omura, S.; Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Takahashi, C.; Shinose, M.; Takahashi, Y.; Horikawa, H.; Nakazawa, H.; Osonoe, T.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M. Genome sequence of an industrial microorganism *Streptomyces avermitilis*: deducing the ability of producing secondary metabolites. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 12215–12220.

50 Tetzlaff, C. N.; You, Z.; Cane, D. E.; Takamatsu, S.; Omura, S.; Ikeda, H. A gene cluster for biosynthesis of the sesquiterpenoid antibiotic pentalenolactone in *Streptomyces avermitilis*. *Biochemistry* **2006**, *45*, 6179–6186.

51 Quaderer, R.; Omura, S.; Ikeda, H.; Cane, D. E. Pentalenolactone biosynthesis. Molecular cloning and assignment of biochemical function to PtlI, a cytochrome P450 of *Streptomyces avermitilis*. *J. Am. Chem. Soc.* **2006**, *128*, 13036–13037.

52 You, Z.; Omura, S.; Ikeda, H.; Cane, D. E. Pentalenolactone biosynthesis. Molecular cloning and assignment of biochemical function to PtlH, a non-heme iron dioxygenase of *Streptomyces avermitilis*. *J. Am. Chem. Soc.* **2006**, *128* (20), 6566–6567.

53 You, Z.; Omura, S.; Ikeda, H.; Cane, D. E.; Jogl, G. Crystal structure of the non-heme iron dioxygenase PtlH in pentalenolactone biosynthesis. *J. Biol. Chem.* **2007**, *282*, 36552–36560.

54 You, Z.; Omura, S.; Ikeda, H.; Cane, D. E. Pentalenolactone biosynthesis: Molecular cloning and assignment of biochemical function to PtlF, a short-chain dehydrogenase from *Streptomyces avermitilis*, and identification of a new biosynthetic intermediate. *Arch. Biochem. Biophys.* **2007**, *459*, 233–240.

55 Jiang, J.; Tetzlaff, C. N.; Takamatsu, S.; Iwatsuki, M.; Komatsu, M.; Ikeda, H.; Cane, D. E. Genome mining in *Streptomyces avermitilis*: A biochemical Baeyer−Villiger reaction and discovery of a new branch of the pentalenolactone family tree. *Biochemistry* **2009**, *48*, 6431–6440.

56 Seo, M. J.; Zhu, D.; Endo, S.; Ikeda, H.; Cane, D. E. Genome mining in *Streptomyces*. Elucidation of the role of Baeyer−Villiger monooxygenases and non-heme iron-dependent dehydrogenase/oxygenases in the final steps of the biosynthesis of pentalenolactone and neopentalenolactone. *Biochemistry* **2011**, *50*, 1739–1754.

57 Zhu, D.; Seo, M. J.; Ikeda, H.; Cane, D. E. Genome mining in *Streptomyces*. Discovery of an unprecedented P450-catalyzed oxidative rearrangement that is the final step in the biosynthesis of pentalenolactone. *J. Am. Chem. Soc.* **2011**, *133*, 2128–2131.

58 Agger, S. A.; Lopez-Gallego, F.; Hoye, T. R.; Schmidt-Dannert, C. Identification of sesquiterpene synthases from Nostoc punctiforme PCC 73102 and Nostoc sp. strain PCC 7120. *J. Bacteriol.* **2008**, *190*, 6084–6096.

59 Hu, Y.; Chou, W. K. W.; Hopson, R.; Cane, D. E. Genome mining in *Streptomyces clavuligerus*: Expression and biochemical characterization of two new cryptic sesquiterpene synthases. *Chem. Biol.* **2011**, *18*, 32–37.

60 Nakano, C.; Kim, H.-K.; Ohnishi, Y. Identification and characterization of the linalool/nerolidol synthase from *Streptomyces clavuligerus*. *ChemBioChem* **2011**, *12*, 2403–2407.

61 Nakano, C.; Kim, H. K.; Ohnishi, Y. Identification of the first bacterial monoterpene cyclase, a 1,8-cineole synthase, that catalyzes the direct conversion of geranyl diphosphate. *ChemBioChem* **2011**, *12*, 1988–1991.

62 Nakano, C.; Horinouchi, S.; Ohnishi, Y. Characterization of a novel sesquiterpene cyclase involved in (+)-caryolan-1-ol biosynthesis in *Streptomyces griseus*. *J. Biol. Chem.* **2011**, *286*, 27980–27987.

63 Nakano, C.; Kudo, F.; Eguchi, T.; Ohnishi, Y. Genome mining reveals two novel bacterial sesquiterpene cyclases: (−)-Germacradien-4-ol and (−)-epi-α-bisabolol synthases from *Streptomyces citricolor*. *ChemBioChem* **2011**, *12*, 2271–2275.