Published in final edited form as:

Biochemistry. 2012 April 10; 51(14): 3003–3010. doi:10.1021/bi300109c.

Structure of geranyl diphosphate C-methyltransferase from *Streptomyces coelicolor* and implications for the mechanism of isoprenoid modification†

Mustafa Köksal[§], Wayne K. W. Chou[#], David E. Cane[#], and David W. Christianson^{§,*}

[§]Roy and Diana Vagelos Laboratories, Department of Chemistry, University of Pennsylvania, 231 South 34th Street, Philadelphia, PA 19104-6323 USA

[#]Department of Chemistry, Brown University, Box H, Providence, RI 02912-9108 USA

Abstract

Geranyl diphosphate C-methyltransferase (GPPMT) from *Streptomyces coelicolor* A3(2) is the first methyltransferase discovered that modifies an acyclic isoprenoid diphosphate, geranyl diphosphate (GPP), to yield a non-canonical acyclic allylic diphosphate product, 2-methylgeranyl diphosphate, which serves as the substrate for a subsequent cyclization reaction catalyzed by a terpenoid cyclase, methylisoborneol synthase. Here, we report the crystal structures of GPPMT in complex with GPP or the substrate analogue geranyl-S-thiolodiphosphate (GSPP) along with S-adenosyl-L-homocysteine in the cofactor binding site, resulting from *in situ* demethylation of S-adenosyl-L-methionine, at 2.05 Å and 1.82 Å resolution, respectively. These structures suggest that both GPP and GSPP can undergo catalytic methylation in crystalline GPPMT, followed by dissociation of the isoprenoid product. S-adenosyl-L-homocysteine remains bound in the active site, however, and does not exchange with a fresh molecule of cofactor S-adenosyl-L-methionine. These structures provide important clues regarding the molecular mechanism of the reaction, especially with regard to the face of the 2,3 double bond of GPP that is methylated as well as the stabilization of the resulting carbocation intermediate through cation- π interactions.

Terpenoids, also known as isoprenoids, comprise a ubiquitous family of natural products found in all forms of life. With more than 60,000 terpenoids identified to date, the structural and stereochemical diversity of this family, which includes the steroids, is vast (Dictionary of Natural Products: <http://dnp.chemnetbase.com>). Such molecular diversity arises in large part from the catalytic activity of terpenoid cyclases, which catalyze myriad cyclization reactions utilizing canonical acyclic substrates bearing an integral number of C₅-isoprenoid units, such as C₁₀-geranyl diphosphate, C₁₅-farnesyl diphosphate, and C₂₀-geranylgeranyl diphosphate (1–7). The hydrocarbon skeletons of the resulting (poly)cyclic terpenoids are typically subject to further modification during the course of a biosynthetic pathway, such as oxidation, O-methylation, esterification, or carbon elimination reactions, to yield final products that do not necessarily contain integral numbers of C₅-isoprenoid units. For example, in the biosynthesis of the earthy odorant geosmin, the C₁₅-intermediate

†Supported by US National Institutes of Health Grants GM56838 (D.W.C.) and GM30301 (D.E.C.).

*To whom correspondence should be addressed. D.W.C.: telephone: (215) 898-5714; fax: (215) 573-2201; chris@sas.upenn.edu.

²The atomic coordinates and the crystallographic structure factors of geranyl diphosphate C-methyltransferase from *Streptomyces coelicolor* complexed with Mg²⁺, geranyl-S-thiolodiphosphate, and S-adenosyl-L-homocysteine, and with Mg²⁺, geranyl diphosphate, and S-adenosyl-L-homocysteine, have been deposited in the Protein Data Bank (www.rcsb.org) with accession codes 3VC1 and 3VC2, respectively.

SUPPORTING INFORMATION AVAILABLE Amino acid sequence alignments for GPPMTs and activity assay results for Y51F GPPMT. This material is available free of charge via the Internet at <http://pubs.acs.org>.

germacradienol undergoes an elimination reaction to yield C₁₂-geosmin and a C₃-acetone molecule (8–10). In another example, a common one-carbon modification of plant and fungal steroids is the *S*-adenosyl-L-methionine (SAM)-dependent methylation reaction at C-24 (11, 12). These examples further serve to illustrate that modifications of the hydrocarbon skeleton and carbon stoichiometry in terpenoid biosynthesis generally occur *after* the first committed step of cyclization of a canonical, acyclic isoprenoid substrate.

Another interesting modified terpenoid is the volatile C₁₁-terpenoid 2-methylisoborneol that, along with the C₁₂-terpenoid geosmin, gives rise to the characteristic odor of freshly turned earth (13). While 2-methylisoborneol is identified as a malodorous contaminant in drinking water and infected fish (14–16), it is also responsible for the more pleasing earthy bouquets of Brie and Camembert cheeses (17). It was reported more than thirty years ago that the additional methyl group of 2-methylisoborneol is derived from methionine (18), but it was only recently discovered that the acyclic precursor geranyl diphosphate (GPP) is the substrate for the SAM-dependent C-methylation reaction (Figure 1) (19–21). Thus, the biosynthesis of 2-methylisoborneol provides a rare example of the diversification of terpenoid structure and carbon stoichiometry by covalent modification of the cyclization substrate, rather than the cyclization product, in a terpenoid biosynthetic pathway. Subsequent genetic and biochemical studies have identified two-gene operons in several *Streptomyces* species (20, 21) and cyanobacteria (22) that encode for a SAM-dependent geranyl diphosphate methyltransferase (GPPMT) and 2-methylisoborneol synthase (MIBS). Although there are numerous examples of SAM-dependent methyltransferases that modify proteins, nucleic acids, lipids, and hormones (23), GPPMT is the first example of a methyltransferase that modifies a canonical acyclic isoprenoid diphosphate to enable its subsequent cyclization in a terpenoid biosynthetic pathway.

Here, we report the X-ray crystal structure determination of the 33-kDa GPPMT from *Streptomyces coelicolor* A3(2) complexed with SAM that has been demethylated *in situ* to give *S*-adenosyl-L-homocysteine (SAH) and substrate GPP at 2.05 Å resolution. We also report the structure of the complex with SAH and substrate analogue geranyl-*S*-thiolodiphosphate (GSPP) at 1.82 Å resolution. Together, these structures provide important clues regarding the mechanism of the unusual isoprenoid methylation reaction catalyzed by GPPMT. Although these are the first structures of an isoprenoid methyltransferase to be reported in the Protein Data Bank (www.rcsb.org), we note that the crystallization and preliminary X-ray crystallographic study of GPPMT from *Streptomyces lasaliensis* was recently reported (24).

EXPERIMENTAL PROCEDURES

Isoprenoid Diphosphate Ligands

Geranyl-*S*-thiolodiphosphate (GSPP) and geranyl diphosphate (GPP) were purchased from Echelon Biosciences Inc.

Expression and Purification of Geranyl Diphosphate Methyltransferase (GPPMT)

A clone of geranyl diphosphate methyltransferase from *Streptomyces coelicolor* with a 20-residue N-terminal hexahistidine tag and linker (GPPMT) in pET28a plasmid (Novagen Inc, USA) was described previously (21). Protein was expressed using *Escherichia coli* BL21 (DE3) cells (Stratagene Inc.). Transformed cell cultures were grown in 2 L flasks containing 1 L Terrific Broth medium with 50 mg kanamycin at 37° C. At OD₆₀₀ = 0.8 – 0.9, cultures were equilibrated at 18° C and expression was induced by 0.2 mM isopropyl-1-thio-β-D-galactopyranoside for 16 h. Cells were harvested by centrifugation at 6000g for 10 min, producing ~7 g pellet per liter of culture. A 35 g pellet was suspended in 50 ml of buffer E

(50 mM K_2HPO_4 (pH 8.0), 300 mM NaCl, 10% (v/v) glycerol, 3 mM β -mercaptoethanol (BME)) containing 100 μ M phenylmethylsulfonyl fluoride and an EDTA-free Complete Protease Inhibitor Cocktail tablet (Roche Diagnostics, IN). Cells were disrupted by sonication on ice with a large probe at medium power, $5 \times (1 \text{ min on} + 1 \text{ min off})$. Cell debris was cleared by centrifugation twice at $20,000g$ for 30 min. The clear supernatant was applied to a pre-equilibrated 5 mL HisTrap column (GE Healthcare, USA) at a flow rate of 1 mL/min using an ÄKTAprime plus FPLC system (GE Healthcare Bio-Sciences AB, Sweden). The loaded column was washed three times: first with 10 column volumes of buffer E, second with 10 column volumes of buffer E plus 25 mM imidazole, and finally with 10 column volumes of buffer E plus 50 mM imidazole. The GPPMT protein was eluted with a 50 mL gradient of 50–250 mM imidazole in buffer E at a flow rate of 2.5 mL/min. Selected fractions were combined, precipitate was filtered, and soluble protein was concentrated to a 20 mL volume (approximately 10 mg/mL). The protein was applied as 4×5 mL samples to a Superdex 200 preparative grade 26/60 size exclusion column (GE Healthcare Bio-Sciences AB, Sweden) with buffer S (50 mM piperazine- N,N' -bis(2-ethanesulfonic acid) (PIPES) (pH 6.7), 20% (v/v) glycerol, 5 mM BME, 15 mM $MgCl_2$, 100 mM NaCl). Fractions from the size exclusion column were combined and concentrated to 22 mg/mL in buffer S.

Crystallization of GPPMT

GPPMT was crystallized in the presence of SAM and monoterpenoid ligands GSPP or GPP by the sitting drop vapor diffusion method; we were unable to crystallize the protein in the absence of SAM. Accordingly, GPPMT was incubated at 4°C in the presence of 10 mM SAM and 1 mM monoterpenoid ligand for 2 hours before crystallization experiments. Typically, a 1 μ L drop of protein solution [6 mg/mL GPPMT, 50 mM PIPES (pH 6.7), 20% glycerol, 5 mM BME, 15 mM $MgCl_2$, 100 mM NaCl, 10 mM SAM, 1 mM monoterpenoid ligand] was added to a 1 μ L drop of precipitant solution [100 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) (pH 7.5), 25% polyethylene glycol 3350, 200 mM $(NH_4)_2SO_4$ for cocrystallization with GSPP; 100 mM Bis-Tris (pH 6.5), 25% polyethylene glycol 3350, 200 mM $(NH_4)_2SO_4$ for cocrystallization with GPP] and equilibrated against a 100 μ L reservoir of precipitant solution at 21°C . Crystals appeared as rectangular prisms within one day and grew to maximal dimensions of $200 \mu\text{m} \times 100 \mu\text{m} \times 100 \mu\text{m}$ in 1–2 weeks. Crystals were flash-cooled in liquid nitrogen after transfer to a cryoprotectant solution consisting of the mother liquor augmented with 15% glycerol.

Crystallographic Data Collection and Processing

Crystals of the GPPMT complexes with GSPP and GPP diffracted X-rays to 1.82 Å and 2.05 Å resolution at the National Synchrotron Light Source (beamline X29A) using incident radiation with $\lambda = 1.075$ Å. Crystals of the complex with GSPP belonged to space group $P2_1$ with unit cell parameters $a = 97.03$ Å, $b = 102.90$ Å, $c = 203.58$ Å, $\alpha = \gamma = 90^\circ$, $\beta = 99.59^\circ$. Crystals of the complex with GPP belonged to space group $P2_1$ with unit cell parameters $a = 98.13$ Å, $b = 103.25$ Å, $c = 204.13$ Å, $\alpha = \gamma = 90^\circ$, $\beta = 99.05^\circ$. Crystals of both complexes contained 12 monomers in the asymmetric unit with Matthews coefficient $V_M = 2.41$ Å³/Da (solvent content = 49%). For experimental phasing, crystals of the complex with GSPP were soaked in 90 mM HEPES (pH 7.5), 22.5% polyethylene glycol 3350, 180 mM $(NH_4)_2SO_4$, 2 mM methylmercury chloride, 10% glycerol for 6 h and equilibrated with a 250 μ L reservoir of the cryoprotectant solution at 21°C prior to flash-cooling in liquid nitrogen. Single wavelength anomalous dispersion (SAD) data were collected from these crystals at the National Synchrotron Light Source (beamline X29A) using incident radiation with $\lambda = 1.005$ Å. These crystals were isomorphous with those of the native complex and diffracted X-rays to 2.1 Å resolution. Diffraction data were processed with HKL2000 (25). Data collection and reduction statistics are recorded in Table 1.

Structure Determination of GPPMT

The initial electron density map of the GPPMT complex with GSPP was phased by the single isomorphous replacement with anomalous scattering (SIRAS) method using the anomalous signal up to 2.6 Å resolution. Search and refinement of 24 mercury sites, density modification, and experimentally-phased electron density map calculation were performed using the program HKL2MAP (26). A total of 272 out of 312 protein residues per monomer were built manually using the experimentally-phased electron density map and propagated through non-crystallographic symmetry (NCS) operators to build all 12 monomers in the asymmetric unit with the program COOT (27). The resulting structure was refined against the 1.82 Å and 2.05 Å resolution data sets obtained from isomorphous crystals of the GSPP and GPP complexes, respectively. Initial rigid body refinement, iterative cycles of positional refinement, and refinement of grouped and then individual atomic *B*-factors were performed using PHENIX (28); manual model rebuilding was performed using COOT (27). No NCS restraints were used during refinement. In both structures, electron density corresponding to the cofactor clearly indicated that SAM had been demethylated to yield *S*-adenosyl-*L*-homocysteine (SAH). Water molecules, Mg²⁺ ions, SAH, and monoterpenoid ligands were included in later cycles of refinement. A total of 275–290 out of 312 residues are present in the final models of the 12 monomers in the asymmetric units of the GPPMT•SAH•GSPP and GPPMT•SAH•GPP complexes. The N-terminal hexahistidine tag, its linker segment, and the first 17 residues (M1-P17) of 11 monomers were disordered and are absent in the final model (monomer D contains T3-P17). Refinement statistics are recorded in Table 1; Ramachandran plot statistics were calculated with PROCHECK (29). Simulated annealing omit maps were calculated with CNS (30). Protein structure figures were prepared with the graphics program PyMol (<http://www.pymol.org>) and labeled for publication using PhotoshopCS.

Preparation and Assay of Y51F GPPMT

The Y51F GPPMT mutant was prepared using the QuickChange protocol (Qiagen Inc., USA) with the mutagenic primers CTATCACCACCACTTCGGCATCGGCCCGG (sense) and CGGGGCCGATGCCGAAGTGGTGGTGATAG (antisense). The mutant gene was verified by DNA sequencing, and the mutant protein was expressed and purified as described above for wild-type GPPMT. After purification on the metal ion affinity column, sample was dialyzed into the assay buffer described below and used for activity measurements.

Kinetic parameters ($V_{\max}$ and K_M) were obtained for GPP and SAM (*S*-[methyl-³H]-SAM, PerkinElmer) under initial velocity conditions. Values for the variable substrate GPP (0.5 – 30 μM) were determined at a saturating SAM concentration (60 μM; 16.2 Ci/mol), and similarly the values for variable SAM (0.5 – 30 μM) were measured at a saturating concentration of GPP (60 μM). The reactions were carried out in 0.5 mL of assay buffer (50 mM PIPES (pH 7.0), 20% (vol/vol) glycerol, 10 mM MgCl₂, 100 mM NaCl, 5 mM BME) with 1.9–2.7 μM of wild-type or Y51F GPPMT, and were incubated at 30 °C for 5 min. Reactions were quenched with 0.5 mL 2N HCl in 83% ethanol. The resulting mixtures were overlaid with 1.0 mL pentane, incubated at 30 °C for another 10 min to hydrolyze the acid-labile pyrophosphate, and then neutralized by adding 0.35 mL of 10% NaOH. Reaction mixtures were extracted with 3×1.0 mL of pentane, with each extract passed through a 2 cm column of silica gel (in a Pasteur pipette) into a scintillation vial containing 7 mL of Opti-Fluor. The column was subjected to a final wash of 1.0 mL of diethyl ether, collected into the same vial and counted by liquid scintillation. Kinetic constants were determined by using the Kaleidagraph program and fitting the data to the Michaelis-Menten equation.

RESULTS AND DISCUSSION

Structure of GPPMT

GPPMT crystallizes as a trimer of dimers, i.e., a hexamer, although no monomer-monomer interface (including crystal packing interfaces) meets the criteria established for contact surface area expected for stable protein oligomers (31) (Figure 2a). GPPMT elutes as monomer in size exclusion chromatography (data not shown), so we conclude that the monomer is the biologically active form of the protein *in vivo*. Even so, NCS relationships among the 12 monomers in the asymmetric unit of the crystal are noteworthy. In each hexamer, three 2-fold NCS axes (relating the top half of the hexamer to the bottom half) are perpendicular to two nearly coincident 3-fold NCS axes. The trimers are oriented in such a way that the 2-fold NCS axes are offset by $\sim 60^\circ$. The 3-fold NCS axes of the two hexamers in the asymmetric unit are oriented nearly parallel to the *a*-axis but deviate away from the *a*-axis and from each other by $\sim 10^\circ$.

GPPMT shares low sequence identity ($\sim 20\%$) with SAM-dependent methyltransferases of known structure. Nonetheless, GPPMT contains all structural components of the Rossmann fold, which is adopted by the family of SAM-dependent *C*-, *N*-, or *O*-methyltransferases (23, 32) (Figure 2b, c). In addition to the core α/β structure, which comprises a 7-stranded β sheet sandwiched between 5 α helices, GPPMT contains three α helices at the N-terminus, an α helix inserted between $\beta 5$ and αE , and two α helices inserted between $\beta 6$ and $\beta 7$ (secondary structure nomenclature is outlined in ref. 23). The SAM binding site is located exclusively in the N-terminal region of the protein and is defined by residues from $\alpha Z 2$ and polypeptide loops following $\beta 1$, $\beta 2$, $\beta 3$, and $\alpha Z 3$. The GPP/GSPP binding site is encapsulated by residues from $\alpha Z 3$, $\alpha D 1$, $\alpha E 1$, $\alpha E 2$ and the loops following $\alpha Z 2$, $\alpha Z 3$, $\beta 4$, and $\beta 5$ (Figure 2b).

Methyltransferase structures that have highest sequence identity to GPPMT are those of rebeccamycin sugar 4'-*O*-methyltransferase RebM from *Lechevaleria aerocolonigenes* (20% sequence identity) and mycolic acid cyclopropane synthase MmaA2 from *Mycobacterium tuberculosis* (22% sequence identity). Although the core α/β structures of RebM (33) and MmaA2 (PDB ID: 1TPY) are very similar to the core α/β structure of GPPMT (the r.m.s. deviations of core C_α atoms are 1.3 Å and 1.9 Å, respectively), auxiliary secondary structure elements differ: RebM is missing the N-terminal helices $\alpha Z 1$, $\alpha Z 2$, $\alpha Z 3$, helix αZ is shorter, and helices $\alpha D 1$, $\alpha E 1$, and $\alpha E 2$ are longer; MmaA2 is missing the N-terminal helix $\alpha Z 1$, helix αZ is shorter, helices $\alpha D 1$, $\alpha E 1$, and $\alpha E 2$ are longer, and additional short α -helices are inserted between strand $\beta 5$ and helix $\alpha D 1$, and between helix $\alpha E 1$ and helix $\alpha E 2$.

Although GPPMT crystals were grown in the presence of SAM and GPP or GSPP, the active site of GPPMT contains the intact substrate GPP or substrate analogue GSPP and the product form of the cofactor, *S*-adenosyl-L-homocysteine (SAH), resulting from the enzyme-catalyzed demethylation of SAM (Figure 3). Possibly, a molecule of GPP/GSPP initially binds in the active site and is methylated *in situ*, with the 2-methylated product subsequently dissociating. This allows for the binding of a second, intact GPP/GSPP molecule. However, the product cofactor SAH remains trapped in crystalline GPPMT since it binds more deeply in the active site. This could account for the observation of the "arrested" enzyme-substrate-product cofactor complex. Notably, sufficient electron density to model GPP is observed only in monomers A, D, H, and J; other monomers do not appear to bind GPP to a significant extent. Moreover, GPP molecules (and associated Mg^{2+} ion and coordination waters) modeled in these monomers are characterized by poor electron density and high *B*-factors, implying a high degree of disorder or reduced occupancy. However, GSPP is bound with full occupancy in all monomers, suggesting that GPPMT binds GSPP with higher affinity than GPP under the respective crystallization conditions. In some of the monomers

that do not bind GPP (monomers B, C, F, K, and L), helix $\alpha Z3$ and the preceding loop, which enclose GPP binding site in monomers A, D, H, and J, are disordered (Figure 4). This segment is ordered in monomers E, G, and I, even though these monomers do not bind GPP. This observation suggests that helix $\alpha Z3$ and the preceding loop open and close to facilitate substrate binding and product dissociation. A similar scenario involving SAM binding and SAH dissociation may be suggested for helix $\alpha Z2$, which encloses the cofactor binding cavity.

Mechanistic Implications

In the active site of GPPMT, one oxygen atom from each phosphate group of GPP/GSPP coordinates to a single Mg^{2+} ion, thereby forming a 6-membered ring chelate complex; the octahedral coordination polyhedron of this Mg^{2+} ion is completed by the side chain carboxamide group of N37 on helix $\alpha Z3$ and three water molecules. The substrate diphosphate group additionally engages in ionic and hydrogen bond interactions with the side chains of R34, H49, Y51, and R260 (Figure 5). Of these residues, only R34, which is replaced by aspartate or asparagine, is not conserved among GPPMTs from various microorganisms (Figure S1). Mg^{2+} -coordinated water molecules donate hydrogen bonds to the side chain carboxylate group of E81 and the main chain carbonyl groups of V36, H48, and H50. The hydrocarbon tail of GPP/GSPP extends away from SAH and binds in a groove defined by the side chains of W29, Y51, E173, M176, Y177, I218, F222, C224, I226, F273, Y277, F282, and Y284 (Figure 5). All of these residues are conserved among GPPMTs except for F222, which appears as phenylalanine in 4 GPPMTs and tyrosine in 11 GPPMTs (Figure S1).

The SAM and GPP/GSPP binding sites are located near the junction of the N- and C-terminal parts of the central β sheet, respectively. The SAM-binding cavity and the GPP/GSPP binding cavity are continuous; at the junction, H49, E173, and the side chains of W29 and Y177 surround the cavity. At the junction of the two cavities is a water molecule, which is presumably displaced by the binding of an intact SAM molecule. H49 is part of a histidine triplet on the $\alpha Z3$ - αZ loop, which is conserved among GPPMTs (Figure S1). On the opposite side of GPP/GSPP, F222 is positioned to make van der Waals interactions with the 2,3 π bond of GPP/GSPP. The aromatic side chain of F222 is ideally positioned to stabilize a tertiary carbocation intermediate at the substrate C-3 atom through cation- π interactions. As mentioned previously, the aromatic ring of F222 is conserved as phenylalanine or tyrosine in all known GPPMTs (Figure S1). It is interesting to note that F222 is also conserved as F200 in MmaA2 and its homologue, cyclopropane mycolic acid synthase (CmaA1). These methyltransferases catalyze SAM-dependent cyclopropanation of mycolic acids via a similar catalytic mechanism involving a carbocation intermediate (34). Even though a nearby tyrosine residue of CmaA1 is proposed to interact with the carbocation intermediate, F200 is also close to the carbocation and could play a similar role in carbocation stabilization as proposed for F222 of GPPMT (35).

Presuming that the conformation of SAH is comparable to that of SAM, the 2,3 π bond of the GPP/GSPP substrate is closest to the reactive methyl group of SAM. However, the 2,3 π system is not in the optimal orientation for electrophilic attack by the reactive methyl group of SAM: the C-4 methyl group of GPP/GSPP is tilted towards SAM, which could be an artifact of SAH binding instead of SAM. Although the side chains of conserved residues Y51 and E173 are close to the C-2 atom of GPP/GSPP, the side chain of E173 is on the same face of the 2,3 π bond as the cofactor, and hence would be on the wrong face of the carbocation intermediate to assist the final deprotonation at C-2 leading to product formation (however, E173 might provide electrostatic stabilization for the carbocation intermediate). The phenolic hydroxyl group of Y51 is located on the opposite face of the 2,3 π bond, but is somewhat distant from C-2 (3.5 Å). While Y51 would be properly oriented to assist in the

final deprotonation at C-2, catalytic activity measurements of Y51F GPPMT reveal only a ~2-fold reduction in k_{cat} and no effect on K_M for either GPP or SAM (Figure S2), indicating that Y51 is not obligatory for catalysis. A stereospecific proton acceptor may not be required in the final step of catalysis, given the superacidity of the final carbocation intermediate with $\text{pK}_a \sim -5$.

Based on the GPPMT structure, the following reaction mechanism is suggested (Figure 6). Upon GPP binding, complete octahedral coordination of Mg^{2+} by the substrate diphosphate group, N37 of helix $\alpha 3$, and 3 water molecules leads to the fully closed active site conformation. Electrophilic attack of the reactive methyl group of SAM at the 2,3 π bond of GPP then yields SAH and the C-3 tertiary carbocation, which is stabilized by cation- π interactions with the side chain of F222 and by electrostatic interactions with the side chain of E173. A 60° conformational change about the 2,3 σ bond subsequently aligns the C2-H bond with the empty p orbital on C-3, thereby enabling proton elimination from C-2 to form the 2,3 π bond of 2MGPP.

CONCLUSIONS

Although the utilization of non-canonical modified isoprenoid substrates such as 2MGPP by terpenoid cyclases is not commonly observed in nature, terpenoid cyclases can sometimes catalyze reactions in the laboratory with synthetically modified substrates such as fluorinated and methylated sesquiterpenoids (36–39). The recent discovery of cyclases such as MIBS and 2-methylenebornane synthase, which utilize a naturally-modified non-canonical monoterpene substrate generated by GPPMT, illustrates a new strategy for the diversification of terpenoid structure in which the isoprenoid substrate is covalently modified *before* the cyclization reaction (20–22, 40). The modification of isoprenoid substrates can redirect biosynthetic pathways by altering structure-stability relationships of carbocation intermediates that influence critical steps such as hydride shifts, methyl migrations, and other alkyl transfers common in terpenoid cyclase mechanisms. GPPMT, which generates a methylated isoprenoid substrate for a subsequent cyclization reaction in a terpenoid biosynthetic pathway, is the first methyltransferase discovered that modifies a terpenoid cyclase substrate to expand the chemical diversity of a biosynthetic pathway. Possibly, the structure-based engineering of this novel enzyme may lead to additional biosynthetic strategies to further enhance the diversity of the terpenome.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Abbreviations

BME	β -mercaptoethanol
GPP	geranyl diphosphate
GPPMT	geranyl diphosphate C-methyltransferase
GSPP	geranyl-S-thiolodiphosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
2MGPP	2-methylgeranyl diphosphate
MIBS	2-methylisoborneol synthase
NCS	non-crystallographic symmetry

PIPES	piperazine- <i>N,N'</i> -bis(2-ethanesulfonic acid)
SAD	single wavelength anomalous dispersion
SAH	<i>S</i> -adenosyl-L-homocysteine
SAM	<i>S</i> -adenosyl-L-methionine
SIRAS	single isomorphous replacement with anomalous scattering

Acknowledgments

We thank the National Synchrotron Light Source at Brookhaven National Laboratory for beamline access.

References

1. Davis EM, Croteau R. Cyclization enzymes in the biosynthesis of monoterpenes, sesquiterpenes, and diterpenes. Top. in Curr. Chem. 2000; 209:53–95.
2. Wendt KU, Schulz GE, Corey EJ, Liu DR. Enzyme mechanisms for polycyclic triterpene formation. Angew. Chem. Int. Ed. 2000; 39:2812–2833.
3. Christianson DW. Structural biology and chemistry of the terpenoid cyclases. Chem. Rev. 2006; 106:3412–3442. [PubMed: 16895335]
4. Christianson DW. Unearthing the roots of the terpenome. Curr. Op. Chem. Biol. 2008; 12:141–150.
5. Allemann RK. Chemical wizardry? The generation of chemical diversity in terpenoid biosynthesis. Pure Appl. Chem. 2008; 80:1791–1798.
6. Aaron JA, Christianson DW. Trinuclear metal clusters in catalysis by terpenoid synthases. Pure. Appl. Chem. 2010; 82:1585–1597. [PubMed: 21562622]
7. Cane DE, Ikeda H. Exploration and mining of the bacterial terpenome. Acc. Chem. Res. 2011 in press.
8. 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]
9. 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]
10. 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]
11. Nes WD. Enzyme mechanisms for sterol C-methylations. Phytochemistry. 2003; 64:75–95. [PubMed: 12946407]
12. Nes WD. Biosynthesis of cholesterol and other sterols. Chem. Rev. 2011; 111:6423–6451. [PubMed: 21902244]
13. Buttery RG, Garibaldi JA. Geosmin and methylisoborneol in garden soil. J. Agric. Food Chem. 1976; 24:1246–1247.
14. Martin JF, Bennett LW, Graham WH. Off-flavor in the channel catfish (*Ictalurus punctatus*) due to 2-methylisoborneol and its dehydration products. Water Sci. Technol. 1988; 20:99–105.
15. McCallum R, Pendleton P, Schumann R, Trinh MU. Determination of geosmin and 2-methylisoborneol in water using solid-phase microextraction and gas chromatography-chemical ionisation/electron impact ionisation-ion-trap mass spectrometry. Analyst. 1998; 123:2155–2160. [PubMed: 10209901]
16. Jüttner F, Watson SB. Biochemical and ecological control of geosmin and 2-methylisoborneol in source waters. Appl. Environ. Microbiol. 2007; 73:4395–4406. [PubMed: 17400777]
17. Karahadian C, Josephson DB, Lindsay RC. Volatile compounds from *Penicillium* sp. contributing musty-earth notes to Brie and Camembert cheese flavors. J. Agric. Food Chem. 1985; 33:339–343.

18. Bentley R, Meganathan R. Geosmin and methylisoborneol biosynthesis in *Streptomyces*. FEBS Lett. 1981; 125:220–222. [PubMed: 7227551]
19. Dickschat JS, Nawrath T, Theil V, Kunze B, Müller R, Schulz S. Biosynthesis of the off-flavor 2-methylisoborneol by the myxobacterium *Nannocystis exedens*. Angew. Chem. Int. Ed. 2007; 46:8287–8290.
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. USA. 2008; 105:7422–7427. [PubMed: 18492804]
21. Wang C-M, Cane DE. Biochemistry and molecular genetics of the biosynthesis of the earthy odorant methylisoborneol in *Streptomyces coelicolor*. J. Am. Chem. Soc. 2008; 130:8908–8909. [PubMed: 18563898]
22. Wang Z, Xu Y, Shao J, Wang J, Li R. Genes associated with 2-methylisoborneol biosynthesis in cyanobacteria: isolation, characterization, and expression in response to light. PLoS ONE. 2011; 6(4):e18665. [PubMed: 21490938]
23. Martin JL, McMillan FM. SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold. Curr. Op. Struct. Biol. 2002; 12:783–793.
24. 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 Cryst. 2011; F67:417–420.
25. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 1997; 276:307–326.
26. Pape T, Schneider TR. HKL2MAP: a graphical user interface for phasing with SHELX programs. J. Appl. Cryst. 2004; 37:843–844.
27. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr. 2010; D66:486–501.
28. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. 2010; D66:213–221.
29. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: A program to check the stereochemical quality of protein structures. J. Appl. Cryst. 1993; 26:283–291.
30. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang J-S, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR System: A new software suite for macromolecular structure determination. Acta Crystallogr. 1998; D54:905–921.
31. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 2007; 372:774–797. [PubMed: 17681537]
32. Rao ST, Rossmann MG. Comparison of super-secondary structures in proteins. J. Mol. Biol. 1973; 76:241–256. [PubMed: 4737475]
33. Singh S, McCoy JG, Zhang C, Bingman CA, Phillips GN Jr, Thorson JS. Structure and mechanism of the rebeccamycin sugar 4'-O-methyltransferase RebM. J. Biol. Chem. 2008; 283:22628–22636. [PubMed: 18502766]
34. Liao R-Z, Georgieva P, Yu J-G, Himo F. Mechanism of Mycolic Acid Cyclopropane Synthase: A Theoretical Study. Biochemistry. 2011; 50:1505–1513. [PubMed: 21241051]
35. Huang CC, Smith CV, Glickman MS, Jacobs WR Jr, Sacchettini JC. Crystal structures of mycolic acid cyclopropane synthases from *Mycobacterium tuberculosis*. J. Biol. Chem. 2002; 277:11559–11569. [PubMed: 11756461]
36. Cane DE, Tsantrizos YS. Aristolochene synthase. Elucidation of the cryptic germacrene A synthase activity using the anomalous substrate dihydrofarnesyl diphosphate. J. Am. Chem. Soc. 1996; 118:10037–10040.
37. Miller DJ, Yu F, Allemann RK. Aristolochene synthase-catalyzed cyclization of 2-fluorofarnesyl-diphosphate to 2-fluorogermacrene A. ChemBioChem. 2007; 8:1819–1825. [PubMed: 17683054]

38. Miller DJ, Yu F, Knight DW, Allemann RK. 6- and 14-Fluoro farnesyl diphosphate: mechanistic probes for the reaction catalysed by aristolochene synthase. *Org. Biomol. Chem.* 2009; 7:962–975. [PubMed: 19225680]
39. Vedula LS, Zhao Y, Coates RM, Koyama T, Cane DE, Christianson DW. Exploring biosynthetic diversity with trichodiene synthase. *Arch. Biochem. Biophys.* 2007; 466:260–266. [PubMed: 17678871]
40. Chou WK, Ikeda H, Cane DE. Cloning and characterization of Pfl_1841, a 2-methylenebornane synthase in *Pseudomonas fluorescens* PfO-1. *Tetrahedron.* 2011; 67:6627–6632. [PubMed: 21804650]

Figure 1. Reactions catalysed by geranyl diphosphate C-methyltransferase (GPPMT) and methylisoborneol synthase (MIBS) in the biosynthesis of 2-methylisoborneol (Ad = adenosyl).

Figure 2.

(a) Hexameric quaternary structure of GPPMT in the crystal. Buried surface areas are $\sim 1000 \text{ \AA}^2$ and $\sim 900 \text{ \AA}^2$ for dimers across the 2-fold non-crystallographic axes and $\sim 70 \text{ \AA}^2$ for trimers across the 3-fold non-crystallographic axis (dark/light blue). NCS symmetry elements are indicated by red symbols. (b) Ribbon drawing of GPPMT rainbow-colored from N-terminus (blue) to C-terminus (red); SAM and GSPG are shown as stick figures; the Mg^{2+} ion appears as a magenta sphere. (c) Topology diagram of GPPMT (numbering as outlined in ref. 23). Strands of the central β sheet are red, flanking α helices are yellow, additional α helices beyond the core Rossmann fold are grey. The general regions of the SAM and GPP/GSPG binding sites are indicated by green boxes, and the Mg^{2+} binding site is represented by a magenta sphere.

Figure 3.

(a) Simulated annealing omit map (black meshwork, contoured at 4σ) showing the binding of SAH, GSPP, and Mg^{2+} to the active site of monomer D of GPPMT. SAH, GSPP, GPP, and the side chain of N37 coordinated to Mg^{2+} are displayed as stick figures. Ligand atoms are color-coded as follows: C = black, N = blue, O = red, S = yellow, P = orange. The Mg^{2+} ion and coordinated water molecules are shown as magenta and red spheres, respectively. Certain parts of the structure are omitted from the foreground of the figure for clarity. (b) Simulated annealing omit map (contoured at 3σ) showing the binding of SAH, GPP, and Mg^{2+} to the active site of monomer J of GPPMT. Weak electron density for GPP and Mg^{2+} -coordinated water molecules suggests partial occupancy, which is accounted for by higher *B*-factors compared with neighboring residues. Color-coding is identical to that outlined in (a).

Figure 4.

Conformational changes of flexible helix $\alpha Z3$ may enable substrate binding and product dissociation in the active site of GPPMT. Left to right: helix $\alpha Z3$ and the preceding loop are ordered (dark blue) when there is no substrate in active site (monomer E); they are partially or completely disordered when there is no substrate in active site (monomers F and C, respectively); they are ordered when there is substrate in active site (monomer J). Ribbon drawing represents the ordered parts of the structure; broken lines represent the disordered segments. SAH and GPP are displayed as black stick figures, the Mg^{2+} ion is represented by a magenta sphere, and water molecules are represented by small red spheres. The side chain of N37 on helix $\alpha Z3$ (stick figure) hydrogen bonds with a solvent molecule in the absence of substrate but coordinates to the Mg^{2+} ion when substrate is bound.

Figure 5.

Stereographic view showing the interactions between the diphosphate group and the isoprenoid group of GSPP and the residues defining the substrate binding pocket of GPPMT. SAH, GSPP, and side chains are displayed as stick figures. Atoms are color-coded as follows: C = black (ligand) or gray (protein), N = blue, O = red, S = yellow, P = orange. The Mg^{2+} ion and water molecules are displayed as magenta and red spheres, respectively. Helix αZ3 is omitted from the figure for clarity. Metal coordination and hydrogen bond interactions are indicated by solid black and dashed red lines, respectively. Interactions between the hydroxyl group of Y51 and the C-2 atom of GSPP (3.5 Å), the side chain of F222 and the C-3 atom of GSPP (4 Å), and the $\text{S}\delta$ atom of SAH and the water molecule at the junction of cofactor and substrate binding cavities (3.7 Å) are indicated by blue dashed lines.

Figure 6.

Proposed catalytic mechanism of GPPMT for the SAM-dependent methylation at C-2 of GPP, yielding SAH and 2MGPP (Ad = adenosyl). Hydrogen bonds with the substrate diphosphate group as observed in the crystal structures are shown as dotted red lines. For clarity, diphosphate-metal coordination and hydrogen bond interactions with R34, H49, and R260 are omitted. The cation- π interaction between the F222 side chain and the tertiary carbocation intermediate is shown as a dashed blue line and is based on the substrate binding geometry illustrated in Figure 5. The substrate binding conformation must be sufficiently flexible to allow the 60° conformational change about the 2,3 σ bond which enables proton elimination and formation of the 2,3 π bond of 2MGPP.

Table 1

Data Collection and Refinement Statistics

Complex	GPPMT•SAH•GSPP	GPPMT•SAH•GPP
<i>A. Data Collection</i>		
Incident wavelength (Å)	1.075	1.075
Resolution range (Å)	50.0–1.82	50.0–2.05
No. of reflections (total/unique)	1067694 / 348804	847434 / 246191
Completeness ^a (%)	99.1 (98.7)	97.4 (97.3)
Redundancy ^a	3.1 (2.9)	3.4 (3.2)
I/σ	10.2 (2.1)	10.2 (2.1)
R _{merge} ^b	0.106 (0.536)	0.109 (0.574)
<i>B. Refinement</i>		
R _{work} /R _{free} ^c	0.167 / 0.194	0.181 / 0.222
Protein atoms ^d	26,658	26,281
Solvent atoms ^d	1904	1165
Ligand atoms ^d	653	397
R.m.s. deviations		
Bonds (Å)	0.016	0.007
Angles (°)	1.6	1.0
Dihedral angles (°)	17.9	15.9
Improper dihedral angles (°)	1.6	1.2
Average B factors (Å ²)		
Main chain	24	27
Side chain	27	30
Ligand	24	25
Solvent	29	27
Ramachandran plot		
Allowed (%)	93	93.6
Additionally allowed (%)	6.7	6.2
Generously allowed (%)	0.3	0.2
Disallowed (%)	0	0

^a Number in parentheses refer to the outer shell of data.

^b $R_{\text{merge}} = \sum |I - \langle I \rangle| / \sum I$, where I is the observed intensity and $\langle I \rangle$ is the average intensity calculated from replicate data.

^c $R_{\text{work}} = \sum |F_{\text{O}} - F_{\text{C}}| / \sum F_{\text{O}}$ for reflections contained in the working set, and $R_{\text{free}} = \sum |F_{\text{O}} - F_{\text{C}}| / \sum F_{\text{O}}$ for reflections contained in the test set held aside during refinement (1% of the total number of reflections). $|F_{\text{O}}|$ and $|F_{\text{C}}|$ are the observed and calculated structure factor amplitudes, respectively.

^d Per asymmetric unit.