BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Structure of Geranyl Diphosphate *C*-Methyltransferase from *Streptomyces coelicolor* and Implications for the Mechanism of Isoprenoid Modification

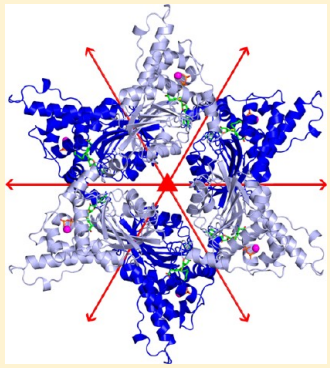
Mustafa Köksal,[†] Wayne K. W. Chou,[‡] David E. Cane,[‡] and David W. Christianson*[,†]

[†]Roy and Diana Vagelos Laboratories, Department of Chemistry, University of Pennsylvania, 231 South 34th Street, Philadelphia, Pennsylvania 19104-6323, United States

[‡]Department of Chemistry, Brown University, Box H, Providence, Rhode Island 02912-9108, United States

Supporting Information

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 noncanonical 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 or 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 about 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 60000 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 C_5-isoprenoid units, such as C_{10}-geranyl diphosphate, C_{15}-farnesyl diphosphate, and C_{20}-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 C_5-isoprenoid units. For example, in the biosynthesis of the earthy odorant geosmin, the C_{15}-intermediate germacradienol undergoes an elimination reaction to yield C_{12}-geosmin and a C_3-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 C_{11}-terpenoid 2-methylisoborneol that, along with the C_{12}-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 30 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 a SAM-dependent geranyl diphosphate methyltransferase (GPPMT) and 2-methylisoborneol synthase (MIBS).[a] Although there are numerous examples of SAM-dependent methyltransferases that modify proteins,

Received: January 24, 2012
Revised: March 9, 2012
Published: March 28, 2012

   dx.doi.org/10.1021/bi300109c | *Biochemistry* 2012, 51, 3003−3010

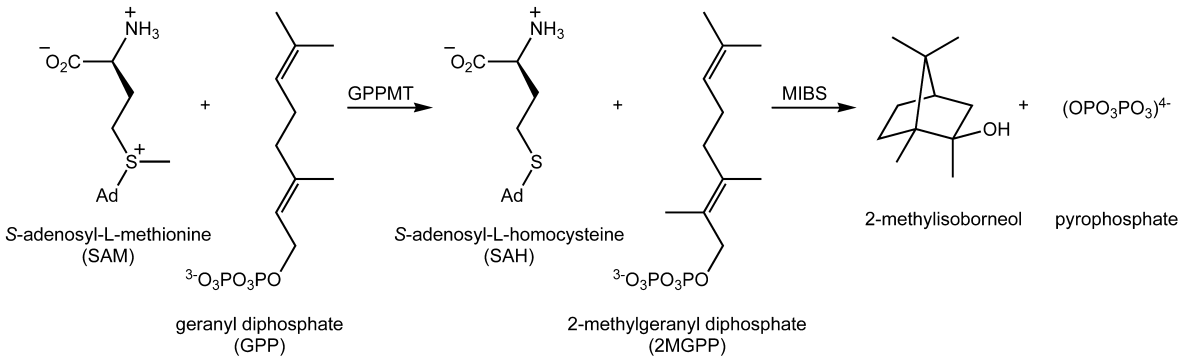


Figure 1. Reactions catalyzed by geranyl diphosphate *C*-methyltransferase (GPPMT) and methylisoborneol synthase (MIBS) in the biosynthesis of 2-methylisoborneol (Ad = adenosyl).

nucleic acids, lipids, and hormones,[23] GPPMT is the first example of a methyltransferase that modifies a canonical acyclic isoprenoid diphosphate to allow its subsequent cyclization in a terpenoid biosynthetic pathway.

Here, we report the determination of the X-ray crystal structure 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 about 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, 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 *S. coelicolor* with a 20-residue N-terminal hexahistidine tag and linker (GPPMT) in the pET28a plasmid (Novagen Inc.) 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 of Terrific Broth with 50 mg of kanamycin at 37 °C. At an OD_{600} of 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 6000*g* for 10 min, producing approximately 7 g pellet/L 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, and 3 mM β-mercaptoethanol (BME)] containing 100 μM phenylmethanesulfonyl fluoride and an EDTA-free Complete Protease Inhibitor Cocktail tablet (Roche Diagnostics, Indianapolis, IN). Cells were disrupted by sonication on ice with a large probe at medium power (five times with 1 min on and 1 min off). Cell debris was cleared by centrifugation twice at 20000*g* for 30 min. The clear supernatant was applied to a pre-equilibrated 5 mL HisTrap column (GE Healthcare) at a flow rate of 1 mL/min using an ÄKTAprime plus FPLC system (GE Healthcare Bio-Sciences AB). The loaded column was washed three times: first with 10 column volumes of buffer E, second with 10 column volumes of buffer E and 25 mM imidazole, and finally with 10 column volumes of buffer E and 50 mM imidazole. The GPPMT protein was eluted with a 50 mL gradient from 50 to 250 mM imidazole in buffer E at a flow rate of 2.5 mL/min. Selected fractions were combined; the 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) 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 $MgCl_2$, and 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 h before crystallization experiments. Typically, a 1 μL 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, and 1 mM monoterpenoid ligand] was added to a 1 μL drop of precipitant solution [100 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) (pH 7.5), 25% polyethylene glycol 3350, and 200 mM $(NH_4)_2SO_4$ for cocrystallization with GSPP; 100 mM Bis-Tris (pH 6.5), 25% polyethylene glycol 3350, and 200 mM $(NH_4)_2SO_4$ for cocrystallization with GPP] and equilibrated against a 100 μL reservoir of precipitant solution at 21 °C. Crystals appeared as rectangular prisms within 1 day and grew to maximal dimensions of 200 μm × 100 μm × 100 μm in 1−2 weeks. Crystals were flash-cooled in liquid nitrogen after being transferred 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 and 2.05 Å resolution, respectively, at the National Synchrotron Light Source (beamline X29A) using incident radiation with λ = 1.075 Å. Crystals of the complex with GSPP belonged to space group $P2_1$ with the following unit cell parameters: a = 97.03 Å, b = 102.90 Å, c = 203.58 Å, $\alpha = \gamma$ = 90°, β = 99.59°. Crystals of the complex with GPP belonged to space group $P2_1$ with the following unit cell parameters: a = 98.13 Å, b = 103.25 Å, c = 204.13 Å, $\alpha = \gamma$ = 90°, and β = 99.05°. Crystals of both complexes contained 12 monomers in

the asymmetric unit with a Matthews coefficient (V_M) of 2.41 Å^3/Da (solvent content of 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, and 10% glycerol for 6 h and equilibrated with a 250 μL reservoir of the cryoprotectant solution at 21 °C prior to being flash-cooled 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 λ = 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 listed in Table 1.

Table 1. Data Collection and Refinement Statistics

	GPPMT·SAH·GSPP	GPPMT·SAH·GPP
Data Collection		
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[a]	10.2 (2.1)	10.2 (2.1)
R_{merge}[a,b]	0.106 (0.536)	0.109 (0.574)
Refinement		
R_{work}/R_{free}[c]	0.167/0.194	0.181/0.222
no. of protein atoms[d]	26658	26281
no. of solvent atoms[d]	1904	1165
no. of ligand atoms[d]	653	397
rmsd		
bonds (Å)	0.016	0.007
angles (deg)	1.6	1.0
dihedral angles (deg)	17.9	15.9
improper dihedral angles (deg)	1.6	1.2
average B factor (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]The numbers in parentheses refer to the outer shell of data. [b]$R_{merge} = \sum|I - \langle I\rangle|/\sum I$, where I is the observed intensity and $\langle I\rangle$ is the average intensity calculated from replicate data. [c]$R_{work} = \sum||F_o| - |F_c||/\sum|F_o|$ for reflections contained in the working set, and $R_{free} = \sum||F_o| - |F_c||/\sum|F_o|$ for reflections contained in the test set held aside during refinement (1% of the total number of reflections). $|F_o|$ and $|F_c|$ are the observed and calculated structure factor amplitudes, respectively. [d]Per asymmetric unit.

Determination of the Structure 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 HKL2MAP.[26] A total of 272 of 312 protein residues per monomer were built manually using the experimentally phased electron density map and propagated through noncrystallographic symmetry (NCS) operators to build all 12 monomers in the asymmetric unit with COOT.[27] The resulting structure was refined against the 1.82 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, Mg^{2+} ions, SAH, and monoterpenoid ligands were included in later cycles of refinement. A total of 275−290 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 listed 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 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.) with the mutagenic primers CTATCACC-ACCACTTCGGCATCGGCCCCG (sense) and CGGGG-CCGATGCCGAAGTGGTGGTGATAG (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, the 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*-^{3}H]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 conducted in 0.5 mL of assay buffer [50 mM PIPES (pH 7.0), 20% (v/v) glycerol, 10 mM $MgCl_2$, 100 mM NaCl, and 5 mM BME] with 1.9−2.7 μM wild-type or Y51F GPPMT, and the mixtures were incubated at 30 °C for 5 min. Reactions were quenched with 0.5 mL of 2 N HCl in 83% ethanol. The resulting mixtures were overlaid with 1.0 mL of pentane, incubated at 30 °C for an additional 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 pipet) 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, and the sample was collected into the same vial and counted by liquid scintillation. Kinetic constants were determined by using Kaleidagraph 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

dx.doi.org/10.1021/bi300109c | *Biochemistry* 2012, 51, 3003−3010

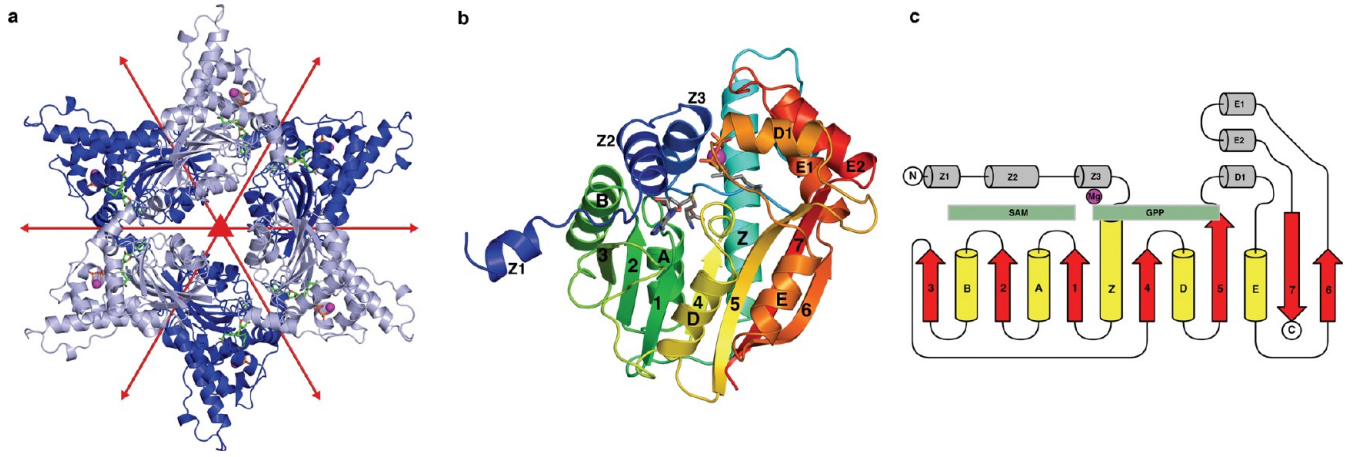


Figure 2. (a) Hexameric quaternary structure of GPPMT in the crystal. The buried surface areas are ~1000 and ~900 Å^2 for dimers across the 2-fold noncrystallographic axes and ~70 Å^2 for trimers across the 3-fold noncrystallographic axis (dark/light blue). NCS symmetry elements are denoted with red symbols. (b) Ribbon drawing of GPPMT rainbow-colored from the N-terminus (blue) to the C-terminus (red). SAH and GSPP 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 colored red, flanking α helices yellow, and additional α helices beyond the core Rossmann fold gray. The general regions of the SAM and GPP/GSPP binding sites are shown as green boxes, and the Mg^{2+} binding site is represented as a magenta sphere.

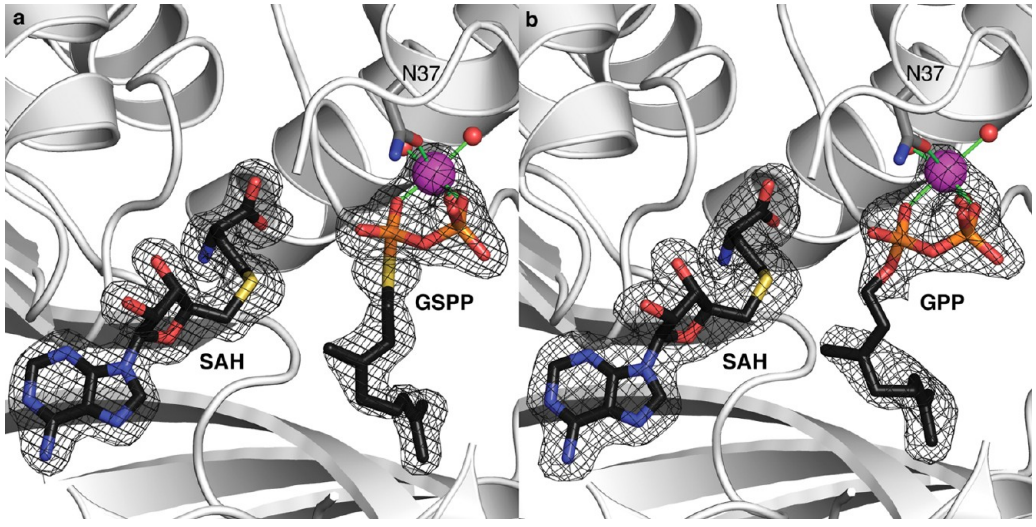


Figure 3. (a) Simulated annealing omit map (black meshwork, contoured at 4σ) showing the binding of SAH, GSPP, and Mg^{2+} to the active site of monomer D of GPPMT. SAH, GSPP, GPP, and the side chain of N37 coordinated to Mg^{2+} are displayed as stick figures. Ligand atoms are color-coded as follows: black for carbon, blue for nitrogen, red for oxygen , yellow for sulfur, and orange for phosphorus. The Mg^{2+} ion and coordinated water molecules are shown as magenta and red spheres, respectively. Certain parts of the structure were omitted from the foreground for the sake of 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. The color coding is identical to that outlined for panel a.

interface (including crystal packing interfaces) meets the criteria established for contact surface area expected for stable protein oligomers[31] (Figure 2a). GPPMT elutes as a monomer via 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 ~60°. 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 ~10°.

GPPMT shares a low level of sequence identity (~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 seven-stranded β sheet sandwiched between five α helices, GPPMT contains three α helices at the N-terminus, an α helix inserted between β5 and αE, and two α helices inserted between β6 and β7 (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 αZ2 and polypeptide loops following β1, β2, β3, and αZ3. The GPP/GSPP binding site is encapsulated by residues from αZ3, αD1,

dx.doi.org/10.1021/bi300109c | *Biochemistry* 2012, 51, 3003−3010

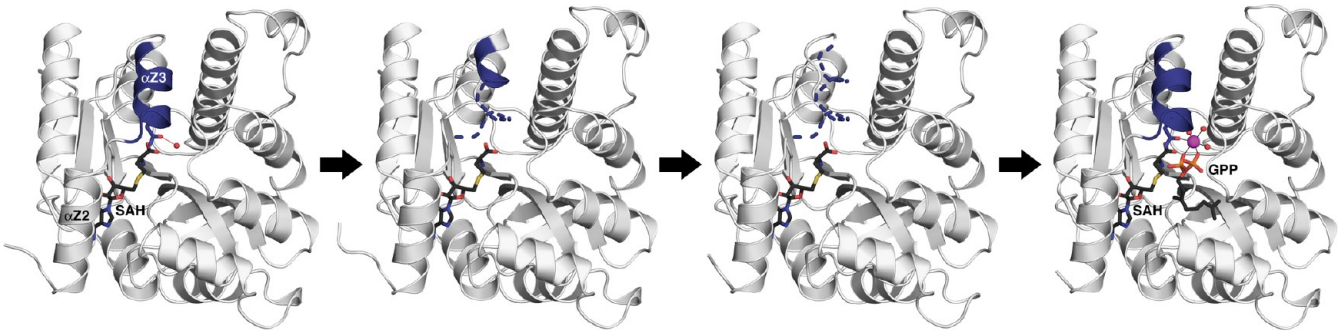


Figure 4. Conformational changes of flexible helix αZ3 may allow substrate binding and product dissociation in the active site of GPPMT. From left to right, helix αZ3 and the preceding loop are ordered (dark blue) when there is no substrate in the active site (monomer E); they are partially or completely disordered when there is no substrate in the active site (monomers F and C, respectively), and they are ordered when there is substrate in the active site (monomer J). The ribbon drawing represents the ordered parts of the structure; dashed lines represent the disordered segments. SAH and GPP are displayed as black stick figures; the Mg^{2+} ion is represented as a magenta sphere, and water molecules are represented as small red spheres. The side chain of N37 on helix αZ3 (stick figure) hydrogen bonds with a solvent molecule in the absence of substrate but coordinates to the Mg^{2+} ion when substrate is bound.

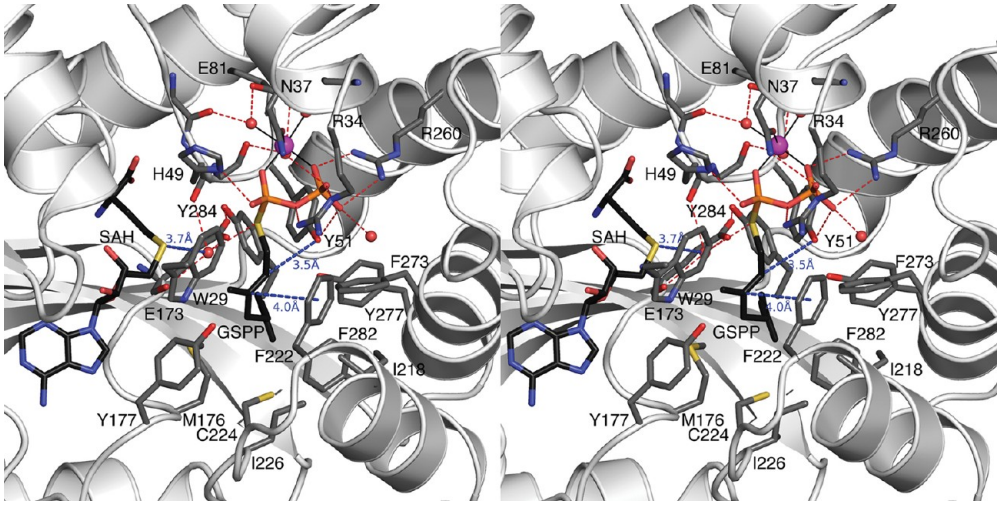


Figure 5. Stereographic view showing the interactions between the diphosphate group and the isoprenoid group of GSPP and the residues defining the substrate binding pocket of GPPMT. SAH, GSPP, and side chains are displayed as stick figures. Atoms are color-coded as follows: black (ligand) or gray (protein) for carbon, blue for nitrogen, red for oxygen, yellow for sulfur, and orange for phosphorus. The Mg^{2+} ion and water molecules are displayed as magenta and red spheres, respectively. Helix αZ3 has been omitted from the figure for the sake of 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 Sδ atom of SAH and the water molecule at the junction of cofactor and substrate binding cavities (3.7 Å) are represented as blue dashed lines.

αE1, and αE2 and the loops following αZ2, αZ3, β4, and β5 (Figure 2b).

Methyltransferase structures that are most identical in sequence to GPPMT are those of rebeccamycin sugar 4′-*O*-methyltransferase RebM from *Lechevaleria aerocolonigenes* (20% identical) and mycolic acid cyclopropane synthase MmaA2 from *Mycobacterium tuberculosis* (22% identical). Although the core α/β structures of RebM[33] and MmaA2 (Protein Data Bank entry 1TPY) are very similar to the core α/β structure of GPPMT (the rmsds of core Cα atoms are 1.3 and 1.9 Å, respectively), auxiliary secondary structure elements differ: RebM is missing N-terminal helices αZ1, αZ2, and αZ3; helix αZ is shorter; and helices αD1, αE1, and αE2 are longer; MmaA2 is missing N-terminal helix αZ1, helix αZ is shorter, helices αD1, αE1, and αE2 are longer, and additional short α helices are inserted between strand β5 and helix αD1 and between helix αE1 and helix αE2.

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 because 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

dx.doi.org/10.1021/bi300109c | *Biochemistry* 2012, 51, 3003–3010

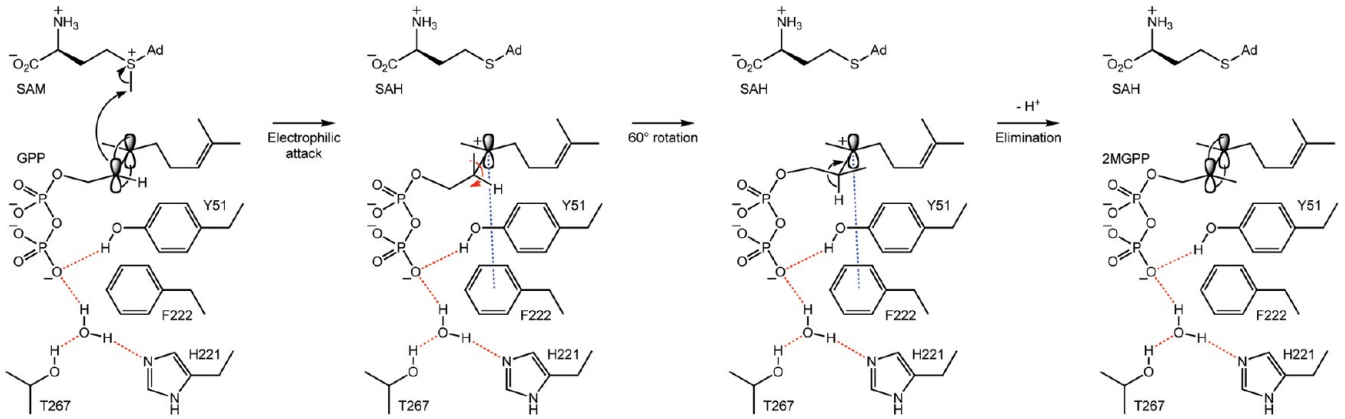


Figure 6. Proposed catalytic mechanism of GPPMT for the SAM-dependent methylation at C-2 of GPP, yielding SAH and 2MGPP (Ad = adenosyl). Hydrogen bonds with the substrate diphosphate group as observed in the crystal structures are shown as red dotted lines. For the sake of clarity, diphosphate−metal coordination and hydrogen bond interactions with R34, H49, and R260 have been 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 that allows proton elimination and formation of the 2,3 π bond of 2MGPP.

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 αZ3 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 αZ3 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 αZ2, 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 six-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 αZ3 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 with aspartate or asparagine, is not conserved among GPPMTs from various microorganisms (Figure S1 of the Supporting Information). 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 four GPPMTs and tyrosine in 11 GPPMTs (Figure S1 of the Supporting Information).

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 αZ3−αZ loop, which is conserved among GPPMTs (Figure S1 of the Supporting Information). 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 of the Supporting Information). 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 toward 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 of the Supporting Information), indicating

dx.doi.org/10.1021/bi300109c | *Biochemistry* 2012, 51, 3003−3010

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 a pK_a of approximately −5.

On the basis of the GPPMT structure, the following reaction mechanism is suggested (Figure 6). When GPP binds, complete octahedral coordination of Mg^{2+} by the substrate diphosphate group, N37 of helix αZ3, and three 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 C-2−H bond with the empty *p* orbital on C-3, thereby allowing proton elimination from C-2 to form the 2,3 π bond of 2MGPP.

CONCLUSIONS

Although the utilization of noncanonical 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 noncanonical monoterpenoid 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,[a] 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 for further enhancing the diversity of the terpenome.

ASSOCIATED CONTENT

Supporting Information

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.

Accession Codes

The atomic coordinates and the crystallographic structure factors of geranyl diphosphate *C*-methyltransferase from *S. coelicolor* complexed with Mg^{2+}, geranyl *S*-thiolodiphosphate, and *S*-adenosyl-L-homocysteine and with Mg^{2+}, geranyl diphosphate, and *S*-adenosyl-L-homocysteine have been deposited in the Protein Data Bank as entries 3VC1 and 3VC2, respectively.

AUTHOR INFORMATION

Corresponding Author

*Telephone: (215) 898-5714. Fax: (215) 573-2201. E-mail: chris@sas.upenn.edu.

Funding

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

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

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

ABBREVIATIONS

BME, β-mercaptoethanol; GPP, geranyl diphosphate; GPPMT, geranyl diphosphate *C*-methyltransferase; GSPP, geranyl *S*-thiolodiphosphate; HEPES, 4-(2-hydroxyethyl)-1-piperazine-ethanesulfonic acid; 2MGPP, 2-methylgeranyl diphosphate; MIBS, 2-methylisoborneol synthase; NCS, noncrystallographic symmetry; PIPES, piperazine-*N*,*N*′-bis(2-ethanesulfonic acid); rmsd, root-mean-square deviation; SAD, single-wavelength anomalous dispersion; SAH, *S*-adenosyl-L-homocysteine; SAM, *S*-adenosyl-L-methionine; SIRAS, single isomorphous replacement with anomalous scattering.

ADDITIONAL NOTE

[a]See following paper in this issue, dx.doi.org/10.1021/bi201827a.

REFERENCES

(1) Davis, E. M., and Croteau, R. (2000) Cyclization enzymes in the biosynthesis of monoterpenes, sesquiterpenes, and diterpenes. *Top. Curr. Chem. 209*, 53−95.

(2) Wendt, K. U., Schulz, G. E., Corey, E. J., and Liu, D. R. (2000) Enzyme mechanisms for polycyclic triterpene formation. *Angew. Chem., Int. Ed. 39*, 2812−2833.

(3) Christianson, D. W. (2006) Structural biology and chemistry of the terpenoid cyclases. *Chem. Rev. 106*, 3412−3442.

(4) Christianson, D. W. (2008) Unearthing the roots of the terpenome. *Curr. Opin. Chem. Biol. 12*, 141−150.

(5) Allemann, R. K. (2008) Chemical wizardry? The generation of chemical diversity in terpenoid biosynthesis. *Pure Appl. Chem. 80*, 1791−1798.

(6) Aaron, J. A., and Christianson, D. W. (2010) Trinuclear metal clusters in catalysis by terpenoid synthases. *Pure Appl. Chem. 82*, 1585−1597.

(7) Cane, D. E., and Ikeda, H. (2012) Exploration and mining of the bacterial terpenome. *Acc. Chem. Res. 45*, 463−472.

(8) Jiang, J., He, X., and Cane, D. E. (2006) Geosmin biosynthesis. *Streptomyces coelicolor* germacradienol/germacrene D synthase converts farnesyl diphosphate to geosmin. *J. Am. Chem. Soc. 128*, 8128−8129.

(9) Jiang, J., He, X., and Cane, D. E. (2007) Biosynthesis of the earthy odorant geosmin by a bifunctional *Streptomyces coelicolor* enzyme. *Nat. Chem. Biol. 3*, 711−715.

(10) Jiang, J., and Cane, D. E. (2008) Geosmin biosynthesis. Mechanism of the fragmentation-rearrangement in the conversion of germacradienol to geosmin. *J. Am. Chem. Soc. 130*, 428−429.

(11) Nes, W. D. (2003) Enzyme mechanisms for sterol C-methylations. *Phytochemistry 64*, 75−95.

(12) Nes, W. D. (2011) Biosynthesis of cholesterol and other sterols. *Chem. Rev. 111*, 6423−6451.

(13) Buttery, R. G., and Garibaldi, J. A. (1976) Geosmin and methylisoborneol in garden soil. *J. Agric. Food Chem. 24*, 1246−1247.

(14) Martin, J. F., Bennett, L. W., and Graham, W. H. (1988) Off-flavor in the channel catfish (*Ictalurus punctatus*) due to 2-methylisoborneol and its dehydration products. *Water Sci. Technol. 20*, 99−105.

(15) McCallum, R., Pendleton, P., Schumann, R., and Trinh, M. U. (1998) 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 123*, 2155−12160.
(16) Jüttner, F., and Watson, S. B. (2007) Biochemical and ecological control of geosmin and 2-methylisoborneol in source waters. *Appl. Environ. Microbiol. 73*, 4395−4406.
(17) Karahadian, C., Josephson, D. B., and Lindsay, R. C. (1985) Volatile compounds from *Penicillium* sp. contributing musty-earthy notes to Brie and Camembert cheese flavors. *J. Agric. Food Chem. 33*, 339−343.
(18) Bentley, R., and Meganathan, R. (1981) Geosmin and methylisoborneol biosynthesis in *Streptomyces*. *FEBS Lett. 125*, 220−222.
(19) Dickschat, J. S., Nawrath, T., Theil, V., Kunze, B., Müller, R., and Schulz, S. (2007) Biosynthesis of the off-flavor 2-methylisoborneol by the myxobacterium *Nannocystis exedens*. *Angew. Chem., Int. Ed. 46*, 8287−8290.
(20) Komatsu, M., Tsuda, M., Omura, S., Oikawa, H., and Ikeda, H. (2008) Identification and functional analysis of genes controlling biosynthesis of 2-methylisoborneol. *Proc. Natl. Acad. Sci. U.S.A. 105*, 7422−7427.
(21) Wang, C.-M., and Cane, D. E. (2008) Biochemistry and molecular genetics of the biosynthesis of the earthy odorant methylisoborneol in *Streptomyces coelicolor*. *J. Am. Chem. Soc. 130*, 8908−8909.
(22) Wang, Z., Xu, Y., Shao, J., Wang, J., and Li, R. (2011) Genes associated with 2-methylisoborneol biosynthesis in cyanobacteria: Isolation, characterization, and expression in response to light. *PLoS One 6*, e18665.
(23) Martin, J. L., and McMillan, F. M. (2002) SAM (dependent) I AM: The S-adenosylmethionine-dependent methyltransferase fold. *Curr. Opin. Struct. Biol. 12*, 783−793.
(24) Ariyawutthiphan, O., Ose, T., Tsuda, M., Gao, Y., Yao, M., Minami, A., Oikawa, H., and Tanaka, I. (2011) Crystallization and preliminary X-ray crystallographic study of a methyltransferase involved in 2-methylisoborneol biosynthesis in *Streptomyces lasaliensis*. *Acta Crystallogr. F67*, 417−420.
(25) Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307−326.
(26) Pape, T., and Schneider, T. R. (2004) HKL2MAP: A graphical user interface for phasing with SHELX programs. *J. Appl. Crystallogr. 37*, 843−844.
(27) Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D66*, 486−501.
(28) Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D66*, 213−221.
(29) Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr. 26*, 283−291.
(30) Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR System: A new software suite for macromolecular structure determination. *Acta Crystallogr. D54*, 905−921.
(31) Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol. 372*, 774−797.
(32) Rao, S. T., and Rossmann, M. G. (1973) Comparison of super-secondary structures in proteins. *J. Mol. Biol. 76*, 241−256.
(33) Singh, S., McCoy, J. G., Zhang, C., Bingman, C. A., Phillips, G. N. Jr., and Thorson, J. S. (2008) Structure and mechanism of the rebeccamycin sugar 4′-*O*-methyltransferase RebM. *J. Biol. Chem. 283*, 22628−22636.
(34) Liao, R.-Z., Georgieva, P., Yu, J.-G., and Himo, F. (2011) Mechanism of Mycolic Acid Cyclopropane Synthase: A Theoretical Study. *Biochemistry 50*, 1505−1513.
(35) Huang, C. C., Smith, C. V., Glickman, M. S., Jacobs, W. R. Jr., and Sacchettini, J. C. (2002) Crystal structures of mycolic acid cyclopropane synthases from *Mycobacterium tuberculosis*. *J. Biol. Chem. 277*, 11559−11569.
(36) Cane, D. E., and Tsantrizos, Y. S. (1996) Aristolochene synthase. Elucidation of the cryptic germacrene A synthase activity using the anomalous substrate dihydrofarnesyl diphosphate. *J. Am. Chem. Soc. 118*, 10037−10040.
(37) Miller, D. J., Yu, F., and Allemann, R. K. (2007) Aristolochene synthase-catalyzed cyclization of 2-fluorofarnesyl-diphosphate to 2-fluorogermacrene A. *ChemBioChem 8*, 1819−1825.
(38) Miller, D. J., Yu, F., Knight, D. W., and Allemann, R. K. (2009) 6- and 14-Fluoro farnesyl diphosphate: Mechanistic probes for the reaction catalysed by aristolochene synthase. *Org. Biomol. Chem. 7*, 962−975.
(39) Vedula, L. S., Zhao, Y., Coates, R. M., Koyama, T., Cane, D. E., and Christianson, D. W. (2007) Exploring biosynthetic diversity with trichodiene synthase. *Arch. Biochem. Biophys. 466*, 260−266.
(40) Chou, W. K., Ikeda, H., and Cane, D. E. (2011) Cloning and characterization of Pfl_1841, a 2-methylenebornane synthase in *Pseudomonas fluorescens* PfO-1. *Tetrahedron 67*, 6627−6632.