JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

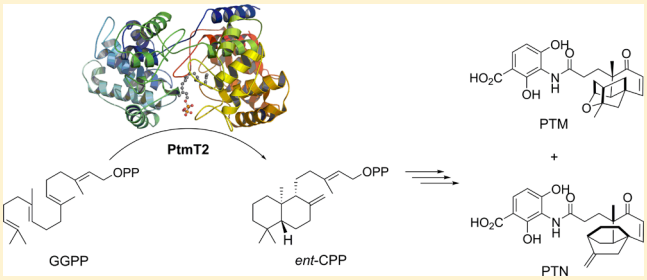
# Structure of the *ent*-Copalyl Diphosphate Synthase PtmT2 from *Streptomyces platensis* CB00739, a Bacterial Type II Diterpene Synthase

Jeffrey D. Rudolf,[†,#] Liao-Bin Dong,[†,#] Hongnan Cao,[‡] Catherine Hatzos-Skintges,[§] Jerzy Osipiuk,[§] Michael Endres,[§] Chin-Yuan Chang,[†] Ming Ma,[†] Gyorgy Babnigg,[§] Andrzej Joachimiak,[§] George N. Phillips, Jr.,[‡] and Ben Shen*[,†,∥,⊥]

[†]Department of Chemistry, The Scripps Research Institute, Jupiter, Florida 33458, United States
[‡]Department of Biosciences, Rice University, Houston, Texas 77005, United States
[§]Midwest Center for Structural Genomics and Structural Biology Center, Biosciences Division, Argonne National Laboratory, Argonne, Illinois 60439, United States
[∥]Department of Molecular Therapeutics, The Scripps Research Institute, Jupiter, Florida 33458, United States
[⊥]Natural Products Library Initiative, The Scripps Research Institute, Jupiter, Florida 33458, United States

Supporting Information

**ABSTRACT:** Terpenoids are the largest and most structurally diverse family of natural products found in nature, yet their presence in bacteria is underappreciated. The carbon skeletons of terpenoids are generated through carbocation-dependent cyclization cascades catalyzed by terpene synthases (TSs). Type I and type II TSs initiate cyclization via diphosphate ionization and protonation, respectively, and protein structures of both types are known. Most plant diterpene synthases (DTSs) possess three $\alpha$-helical domains ($\alpha\beta\gamma$), which are thought to have arisen from the fusion of discrete, ancestral bacterial type I TSs ($\alpha$) and type II TSs ($\beta\gamma$). Type II DTSs of bacterial origin, of which there are no structurally characterized members, are a missing piece in the structural evolution of TSs. Here, we report the first crystal structure of a type II DTS from bacteria. PtmT2 from *Streptomyces platensis* CB00739 was verified as an *ent*-copalyl diphosphate synthase involved in the biosynthesis of platensimycin and platencin. The crystal structure of PtmT2 was solved at a resolution of 1.80 Å, and docking studies suggest the catalytically active conformation of geranylgeranyl diphosphate (GGPP). Site-directed mutagenesis confirmed residues involved in binding the diphosphate moiety of GGPP and identified DxxxxE as a potential $Mg^{2+}$-binding motif for type II DTSs of bacterial origin. Finally, both the shape and physicochemical properties of the active sites are responsible for determining specific catalytic outcomes of TSs. The structure of PtmT2 fundamentally advances the knowledge of bacterial TSs, their mechanisms, and their role in the evolution of TSs.




## INTRODUCTION

Terpenoids are the largest and most structurally diverse family of natural products found in nature with over 65 000 known compounds (http://dnp.chemnetbase.com). Diterpenoids, with carbon skeletons originating from the C20 precursor geranylgeranyl diphosphate (GGPP), are well represented with ~18 000 members, most of which are produced by plants and fungi. Conversely, diterpenoids of bacterial origin are rare and thus understudied, but recent genomics, bioinformatics, and experimental studies suggest bacterial diterpenoids are underrepresented in the current knowledge of the terpenome.[1−3]

Linear terpenoid precursors are cyclized by terpene synthases (TSs, also known as terpene cyclases) in a regiospecific and stereospecific manner.[1,4,5] The cyclization cascades, terminated by proton elimination or nucleophilic capture, generate the vast array of carbon scaffolds found in terpenoid natural products. TSs are categorized into two distinct classes based on their strategies to initiate catalysis and the aspartate-rich sequence motifs responsible for initiation. Type I TSs initiate cyclization by generating a carbocation intermediate from heterolytic cleavage of the diphosphate moiety ("ionization initiation"). Type I TSs contain DDxxD and "NSE/DTE" motifs which coordinate three $Mg^{2+}$ ions and facilitate ionization of the diphosphate group.[6] Type II TSs generate the reactive carbocation and initiate cyclization by protonation of a double bond or epoxide ring ("protonation initiation"). Type II TSs possess DxDD motifs that work in concert to activate the "middle" aspartate for protonation.[7] Type II TSs that act on prenyl diphosphates leave the diphosphate intact, allowing the

Received: April 27, 2016
Published: August 4, 2016




DOI: 10.1021/jacs.6b04317
J. Am. Chem. Soc. 2016, 138, 10905−10915

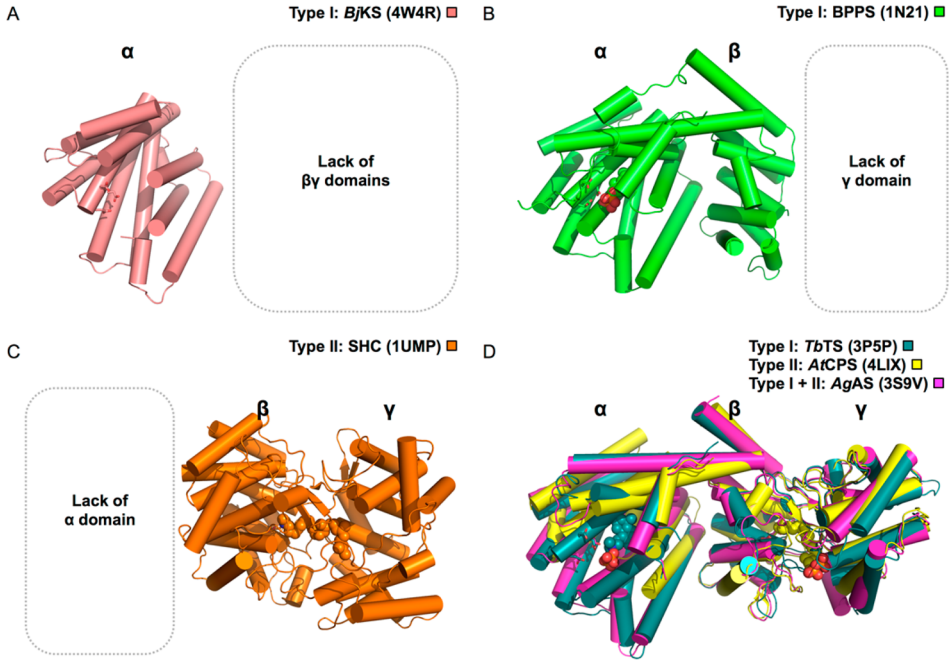



**Figure 1.** Structures of selected terpene synthases (TSs). (A) The type I DTS kaurene synthase from *Bradyrhizobium japonicum* (*Bj*KS, PDB entry 4W4R,[55] pink) is a single isoprenoid domain, termed the $\alpha$-domain. (B) Many plant type I TSs, including bornyl diphosphate synthase (BPPS, PDB entry 1N21,[56] green), are comprised of an $\alpha\beta$-didomain consisting of an $\alpha$-domain and a catalytically inactive $\alpha$-helical $\beta$-domain. (C) Bacterial type II TTSs, such as squalene-hopene cyclase (SHC, PDB entry 1UMP,[57] orange), adopt an $\alpha$-helical $\beta\gamma$ fold, forming an interfacial active site. (D) Plant TSs can also be comprised of all three domains, forming an $\alpha\beta\gamma$ architecture, and be type I, type II, or bifunctional TSs; examples include taxadiene synthase from *Taxus brevifolia* (*Tb*TS, PDB entry 3P5P,[11] teal), *ent*-copalyl diphosphate synthase from *Arabidopsis thaliana* (*At*CPS, PDB entry 4LIX,[44] yellow) and abietadiene synthase from *Abies grandis* (*Ag*AS, PDB entry 3S9 V,[13] magenta), respectively. The DDxxD (type I) and DxDD (type II) motifs are shown as sticks and active site ligands, if present, are shown as spheres. Structures are depicted spacially with a fixed relative position to highlight the missing domains.

products to be used as substrates of type I TSs, further increasing the structural diversity of the carbon skeleton.

The structures of several type I TSs from plants, fungi, and bacteria are known. TSs are known to adopt the $\alpha$-helical "isoprenoid" fold, first observed for avian farnesyl diphosphate synthase.[8] Bacterial and fungal type I TSs are a single isoprenoid domain, termed the $\alpha$-domain, while many plant type I TSs fold into an $\alpha\beta$ didomain consisting of the $\alpha$-domain and a catalytically inactive $\alpha$-helical $\beta$-domain (Figures 1A and 1B).[5] Structurally, type II TSs adopt an $\alpha$-helical $\beta\gamma$ fold, forming an active site and positioning the DxDD motif at the interface of the two domains (Figure 1C).[9] The structures of only a few type II TSs are known and only one structure of a bacterial type II TS has been reported, the tri-TS (TTS) squalene-hopene cyclase (SHC) from *Alicyclobacillus acidocaldarius*.[10] Plant TSs commonly possess all three domains, forming an $\alpha\beta\gamma$ architecture (Figure 1D).[5] These TSs can be type I (with nonfunctional $\beta\gamma$ domains),[11] type II (with nonfunctional $\alpha$ domains),[12] or bifunctional with both type I and type II catalytic activities.[13] The structures of type II di-TSs (DTSs) from bacteria ($\beta\gamma$) are of particular interest as they provide an excellent opportunity to assess the structural, and sequence, similarities between bacterial and plant DTSs.

It has been postulated that all plant TSs share a common evolutionary origin[14,15] and plant DTSs, particularly bifunctional DTSs possessingthe $\alpha\beta\gamma$ architecture, may represent an ancestral-like version of plant TSs.[16] It was additionally hypothesized that the $\alpha\beta\gamma$ architecture arose from the fusion of ancestral (i.e., soil bacterial) genes encoding the discrete $\alpha$ and $\beta\gamma$ proteins.[9] The evolution of both bacterial and plant TSs, however, remains an open question. Analysis of the genome of *Selaginella moellendorffii*, a nonseed vascular plant, revealed that it possessed both typical seed plant-like and microbial-like DTSs, suggesting an alternative hypothesis that TSs in plants may have more than one evolutionary origin.[17]

PtmT2 is a type II DTS involved in the biosynthesis of the bacterial diterpenoids platensimycin (PTM) and platencin (PTN), two fatty acid synthase inhibitors and promising drug leads (Figure 2A). Based on bioinformatics and sequence analysis, PtmT2 is proposed to catalyze the cyclization of GGPP to *ent*-copalyl diphosphate (*ent*-CPP).[18] *ent*-CPP is then channeled into (16*R*)-*ent*-kauran-16-ol or *ent*-atiserene by two distinct type I DTSs specific for PTM or PTN biosynthesis, respectively.[18,19] *ent*-CPP synthases are found in plants, fungi, and bacteria, and the structure of *ent*-CPP synthase from *Arabidopsis thaliana* (*At*CPS with an $\alpha\beta\gamma$ architecture) is known.[12] With no structures of bacterial type II DTSs for comparison, it is unknown if and how the structures of plant DTSs mimic those of bacterial origin, what structural similarities of *ent*-CPP synthases were preserved or conserved during evolution, and what drives the regio- and stereochemical control of the cyclization cascade in bacterial type II DTSs.

Here, we report the first structure of a type II DTS from bacteria. PtmT2, from *Streptomyces platensis* CB00739,[20] was experimentally verified as an *ent*-CPP synthase, its structure determined by X-ray crystallography at a resolution of 1.80 Å, and its substrate and product binding conformations modeled using docking studies. In addition, we performed structure-based site-directed mutagenesis to investigate the residues involved in binding $Mg^{2+}$ and the diphosphate moiety of


DOI: 10.1021/jacs.6b04317
J. Am. Chem. Soc. 2016, 138, 10905−10915

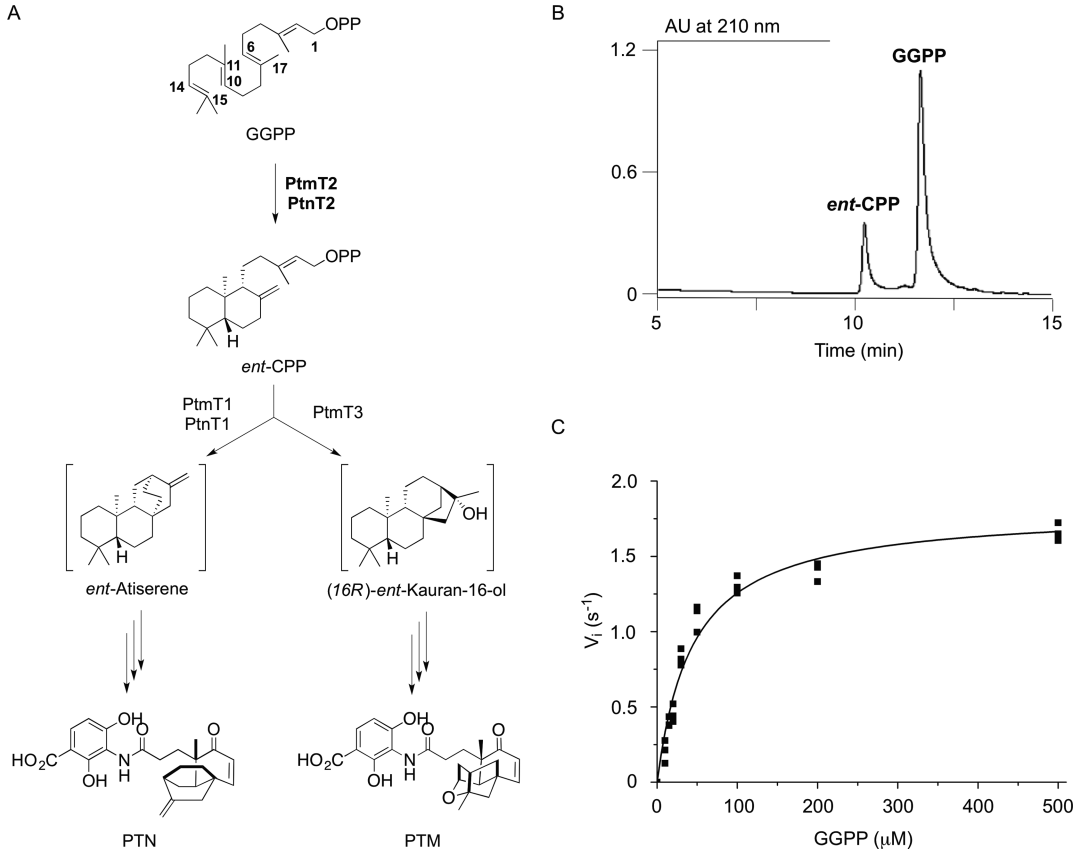



**Figure 2.** PtmT2 catalyzes the cyclization of GGPP into *ent*-CPP. (A) *ent*-CPP is the most advanced common biosynthetic intermediate in the biosynthesis of PTM and PTN. (B) HPLC chromatogram of the production of *ent*-CPP by incubation of GGPP with PtmT2. The structure of *ent*-CPP was purified and confirmed by NMR. (C) Michaelis–Menten kinetic curve for wild-type PtmT2 and GGPP.

GGPP. The structure of PtmT2 advances the field of bacterial TSs, their mechanisms, and their place in the evolution of TSs.

## MATERIALS AND METHODS

**Gene Cloning.** Strains and plasmids used in this study are listed in Tables S1 and S2, respectively. Sesame was used as the laboratory information system for project information and reporting to the PSI Target Track Database.[21] The full-length *ptmT2* gene from *S. platensis* CB00739 (NCBI accession, *ptm* gene cluster KJ189771; PtmT2 protein AIW55555, amino acid residues 1–533) was amplified by PCR from genomic DNA with Q5 DNA polymerase (NEB) following the manufacturer's protocols using the 739T2_F and 739T2_R primers (Table S3). The PCR product was purified, treated with T4 polymerase, and cloned into pBS3080[22] according to ligation-independent procedures[23] to afford pBS12043. For site-directed mutagenesis of *ptmT2*, the *ptmT2* gene from pBS12043 was amplified in two steps by primer extension[24] using the 739T2_F and 739T2_R primers with internal primers containing the desired mutation(s) (Table S3). The mutant *ptmT2* genes were then cloned into pBS3080 as described above yielding pBS12044–pBS12057. Truncated *ptmT2* (amino acid residues 9–528) was amplified with KOD Hot Start DNA polymerase (Novagen) in amplification buffer supplemented with betaine to a final concentration of 2.5 M using the 739xtalT2_F and 739xtalT2_R primers (Table S3). The PCR product was treated and cloned into pMCSG68[25] as described above to afford pBS12058.

**Gene Expression and Protein Production and Purification.** For enzyme activity assays, pBS12043, harboring the full length *ptmT2* gene, was transformed into *E. coli* BL21(DE3) (Life Technologies) and grown in 1 L of lysogeny broth (LB) at 37 °C with shaking at 250 rpm until an $OD_{600}$ of 0.6 was reached. The culture was cooled to 4 °C, gene expression was induced with the addition of 0.25 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and the cells were grown overnight at 18 °C with shaking. After harvesting the cells by centrifugation at 4000*g* for 15 min at 4 °C, the pellet was resuspended in lysis buffer (100 mM Tris, pH 8.0, containing 300 mM NaCl, 15 mM imidazole, and 10% glycerol), lysed by sonication, and centrifuged at 15 000*g* for 20 min at 4 °C. The supernatant containing PtmT2 was purified by nickel-affinity chromatography using an ÄKTA FPLC system (GE Healthcare Biosciences) equipped with a HisTrap column. The resultant protein with an N-terminal $His_6$-tag was desalted using a HiPrep desalting column (GE Healthcare Biosciences) and concentrated using an Amicon Ultra-15 concentrator (Millipore) in 50 mM Tris, pH 7.8, containing 100 mM NaCl, 50 mM KCl, and 5% glycerol. Protein concentrations were determined from the absorbance at 280 nm using a molar absorptivity constant ($\varepsilon_{280}$ = 97 860 $M^{-1}$ $cm^{-1}$). Individual aliquots of PtmT2 were stored at −80 °C until use. Each of the PtmT2 site-directed mutants was produced and purified as described above.

For protein crystallization, pBS12058, harboring the truncated *ptmT2* gene, was transformed into *E. coli* BL21(DE3)-Gold (Stratagene) and grown in 1 L of enriched M9 medium[26] at 37 °C with shaking at 200 rpm until an $OD_{600}$ of 1.0 was reached. Methionine biosynthetic inhibitory amino acids (25 mg $L^{-1}$ each of L-valine, L-isoleucine, L-leucine, L-lysine, L-threonine, L-phenylalanine) and 90 mg $L^{-1}$ of L-selenomethionine (SeMet, Orion Enterprises) were added to the culture, which was then cooled to 4 °C for 60 min. Gene expression was induced with 0.5 mM IPTG and the cells were grown overnight at 18 °C with shaking. After harvesting the cells by centrifugation at 4500*g* for 25 min at 4 °C, the pellet was resuspended in lysis buffer (50 mM HEPES, pH 8.0, containing 500 mM NaCl, 20 mM imidazole, 10 mM β-mercaptoethanol, and 5% glycerol), lysed, and purified using Ni-NTA Immobilized Metal Affinity Chromatography (IMAC 1) and the ÄKTAxpress system (GE Healthcare Biosciences) as previously described.[27] The N-terminal $His_6$-tag was then cleaved from purified PtmT2 using recombinant $His_7$-tagged


DOI: 10.1021/jacs.6b04317
J. Am. Chem. Soc. 2016, 138, 10905–10915

TEV protease. After an additional IMAC step (IMAC 2) to remove the protease, affinity tag, and uncut PtmT2, the resultant cut PtmT2 was concentrated using an Amicon Ultra-15 concentrator (Millipore) in 20 mM HEPES, pH 8.0, containing 250 mM NaCl, and 2 mM DTT. The concentration of pure protein used for crystallization was 128 mg $mL^{-1}$. Individual aliquots of purified protein were stored at −80 °C until use.

**Synthesis of GGPP.** GGPP was synthesized according to previously reported procedures (SI).[28,29]

**Enzymatic Activity of PtmT2.** Preliminary incubations were performed in 50 mM Tris, pH 7.8, containing 1 mM $MgCl_2$, 1 mM $\beta$-mercaptoethanol, 2 mM GGPP, and 1 $\mu$M PtmT2 in a total volume of 100 $\mu$L. After incubation at 30 °C for 10 min, 100 $\mu$L of methanol was added to quench the reaction. The reaction mixture was then centrifuged and 20 $\mu$L of the supernatant were injected and analyzed by HPLC. Each sample was run on an Agilent 1260 HPLC system equipped with a YMC-Triart $C_{18}$ column (250 × 4.6 mm, 5 $\mu$m) using a solvent gradient (1 mL $min^{-1}$) of 20−70% acetonitrile in 25 mM $NH_4HCO_3$. Substrate and product were detected by monitoring 210 nm with a photodiode array detector. GGPP and *ent*-CPP were eluted with retention times of 11.9 and 10.4 min, respectively.

The reaction conditions for PtmT2 were optimized by monitoring *ent*-CPP production using the HPLC method described above. Buffers (Tris, MOPS, citrate), pH (4.5−9.0), temperatures (22, 30, and 37 °C), and selected divalent metals ($Mg^{2+}$, $Mn^{2+}$, $Co^{2+}$, $Ni^{2+}$, and $Zn^{2+}$, with EDTA as a negative control) were all tested for improved PtmT2 activity. Each divalent metal and EDTA were tested at 1 mM. For large-scale production of *ent*-CPP, 20 mg of GGPP was incubated with 5 $\mu$M PtmT2 in 50 mM MOPS, pH 6.5, containing 1 mM each of $MgCl_2$ and $\beta$-mercaptoethanol. After incubation at 30 °C for 1.5 h, 20 mL of methanol was added, and the reaction mixture was centrifuged. The supernatant was evaporated to 5 mL in vacuo at RT and purified by HPLC to yield 8.1 mg of *ent*-CPP (41%). The NMR spectra of *ent*-CPP matched the literature (SI). NMR experiments were performed on a Bruker Avance III Ultrashield 700 at 700 MHz for $^1H$ and 175 MHz for $^{13}C$ and a Bruker Avance Ultrashield 400 at 162 MHz for $^{31}P$.

**Kinetic Studies of PtmT2.** All kinetics assays were performed in citrate buffer, pH 6.0, containing 1 mM $MgCl_2$, 1 mM $\beta$-mercaptoethanol, 10% glycerol, 20 nM PtmT2, and GGPP concentrations of 10, 15, 20, 30, 50, 100, 200, and 500 $\mu$M in a total volume of 100 $\mu$L. Each reaction was incubated at 30 °C for 10 min and quenched with 100 $\mu$L of methanol. After centrifugation, the reaction mixture was analyzed by HPLC as described above and the integrated area under curve (AUC) at 210 nm was calculated. A standard curve of *ent*-CPP was used to convert AUC into the amount of product formed. Each kinetic assay was performed in triplicate and the data was fit to the Michaelis−Menten equation (GraphPad Prism 6.07). Due to limit of detection of *ent*-CPP by HPLC analysis and significant substrate inhibition, relative activities, rather than full kinetic analysis, of all PtmT2 mutants were determined as described above using a GGPP concentration of 50 $\mu$M. For mutants with slower turnovers, enzyme concentration was increased (up to 0.5 $\mu$M) to facilitate product detection, and rates were correspondingly adjusted to account for the changes in enzyme concentration.

**Analytical Size-Exclusion Chromatography.** The molecular weight (MW) and monomeric state of PtmT2 in solution was determined by analytical size-exclusion chromatography using an SRT SEC-150 (7.8 × 250 mm) column (Sepax Technologies) connected to a Dionex HPLC system (Thermo Scientific Dionex). The column was pre-equilibrated with two column volumes of 20 mM HEPES, pH 8.0, containing 250 mM NaCl, and calibrated with ribonuclease A (13.7 kDa), carbonic anhydrase (29 kDa), ovalbumin (44 kDa), conalbumin (75 kDa), and aldolase (158 kDa). The chromatography was carried out at 22 °C at a flow rate of 1 mL $min^{-1}$. The calibration curve of $K_{av}$ versus log(MW) was prepared using the equation $K_{av} = V_e - V_o/(V_t - V_o)$, where $V_e$, $V_o$, and $V_t$ is the elution volume, column void volume, and total bed volume, respectively. Data analysis was performed using Chromeleon Client 6.80 SR10 Build 2818 software (ThermoScientific Dionex).

**Protein Crystallization.** SeMet-labeled PtmT2 was screened for crystallization conditions using a Mosquito liquid dispenser (TTP Labtech) and a sitting drop vapor diffusion technique in 96-well CrystalQuick plates (Greiner Bio-one). The protein was screened against the MCSG 1−4 screens (Microlytic) at 16 °C. For each condition, 0.4 $\mu$L of protein (128 mg $mL^{-1}$) and 0.4 $\mu$L of crystallization formulation were mixed and then equilibrated against 140 $\mu$L of the reservoir solution. Crystals appeared under a number of conditions; those harvested from 0.1 M bis-tris propane:NaOH, pH 7.0 containing 1.5 M ammonium sulfate were used for structure solution at a resolution of 2.2 Å. Optimizations of this conditions were then set up manually in a 1 $\mu$L reservoir:1 $\mu$L protein ratio hanging drops, while varying the concentrations of the precipitant each reservoir and including 5 mM $MgSO_4$ and 2 mM of either GGPP, geranylgeranyl thiodiphosphate (GGSPP), or *ent*-CPP. Triangular shaped crystals appeared after four months at 16 °C in 0.1 M bis-tris propane:NaOH, pH 7.0 containing 1.5 M ammonium sulfate, 5 mM $MgSO_4$, and 2 mM GGSPP. Crystals selected for data collection were soaked in the crystallization buffer supplemented with 25% glycerol and flash-frozen in liquid nitrogen.

**Data Collection, Structure Determination, and Refinement.** Single-wavelength X-ray diffraction data were collected at 100 K temperature at the 19-ID beamline[30] of the Structural Biology Center at the Advanced Photon Source at Argonne National Laboratory using the program SBCcollect. The intensities were integrated and scaled with the HKL-3000 suite.[31] The structure was determined by single-wavelength anomalous dispersion (SAD) phasing using the AutoSol/AutoBuild phasing pipeline[32] from the PHENIX suite.[33] Several rounds of manual adjustments of structure models and refinements using Coot[34] and Refmac[35] from the CCP4 suite,[36] respectively, were done. The stereochemistry of the structure was validated using the PHENIX suite,[33] incorporating MolProbity tools.[37] Although the best diffracting crystals were grown in the presence of $MgSO_4$ and GGSPP, neither $Mg^{2+}$ nor the ligand was observed in the active site. A summary of data collection and refinement statistics is given in Table 1. Figures were prepared using PyMOL (Schrödinger, LLC).

**Computational Modeling of PtmT2 Ligand Complexes.** The ligand and receptor grid map files were prepared using AutoDockTools 1.5.6. Waters and ligands were removed from the PDB file (5BP8) of PtmT2 and selenomethionine residues were changed to methionines. Each of the six methionine residues are sufficiently distant from the active site and >12 Å from the docked ligands. The SMILES file of GGPP was imported into the eLBOW program of the Phenix suite to perform an energy minimization using the semi-empirical quantum mechanical AM1 method.[33] An isomeric SMILES file of *ent*-CPP, which explicitly specifies the chirality of carbon centers and *cis*/*trans* double bonds, was downloaded under ligand code ECP from PDB and energy minimized using the same AM1 method to yield the correct conformer structure. The default parameters of AutoDockTools 1.5.6 were used to add Gasteiger charges and to add and merge hydrogen atoms for the ligand and protein models.[38] Docking receptor map files were generated for a grid box of 60 × 50 × 40 points with default grid spacing of 0.375 Å, which covers the entire space of the active site pocket and solvent accessible substrate channel predicted based on the apparent structural conservation of the type II TS family. Docking experiments were performed using AutoDock 4.2.5.1.[38] Two hundred independent runs were performed for each ligand using the Lamarckian Genetic Algorithm (LGA) with a maximum of 100 million energy evaluations and 27 000 generations. Specifically, the local search frequency was set to default (0.06) and the population size of each generation was set at 350. All parameters defining initial orientation and location of the ligand were set to random. All single C−C bonds in the ligands were treated as rotatable, all double and intracyclic bonds were treated as nonrotatable, and the protein model was treated as rigid. Any conformational changes involving the existing network of interactions between residues forming the active site pocket in PtmT2 might invoke a concerted movement of multiple residues to a first approximation and beyond the accuracy of the current modeling method. Clustering of ligand poses was performed with a default RMSD tolerance of 2.0 Å based on


DOI: 10.1021/jacs.6b04317
J. Am. Chem. Soc. 2016, 138, 10905−10915

**Table 1. Data Collection and Refinement Statistics**

| data set | PtmT2 |
|---|---|
| PDB ID | 5BP8 |
| Space group | *H*32 |
| Unit cell | $a = b = 132.87$ Å, $c = 190.75$ Å; $\alpha = \beta = 90°$, $\gamma = 120°$ |
| Wavelength (Å) | 0.9792 |
| Resolution range (Å)[a] | 42.5−1.80 (1.83−1.80) |
| Unique reflections | 59533 (2749) |
| Multiplicity | 9.2 (5.4) |
| Completeness (%) | 99.4 (93.3) |
| $I/\sigma_I$ | 32.6 (2.0) |
| Wilson B-factor | 27.1 |
| $R_{merge}$ (%)[b] | 7.7 (63.5) |
| $R_{meas}$ (%)[c] | 8.2 |
| $CC_{1/2}$ (Å$^2$) | − (0.78) |
| CC*[d] | 0.937 |
| Phasing and refinement | |
| Resolution (Å) | 42.5−1.80 (1.847−1.80) |
| Number of reflections | 56537 (4078) |
| $R_{work}/R_{free}$ (%) | 15.6/18.8 (23.9/24.0) |
| Number of atoms | 4358 |
| Residues | |
| Macromolecules | 523 |
| Ligand/ion | 6 |
| Water | 530 |
| RMSD bond (Å) | 0.011 |
| RMSD angle (deg) | 1.42 |
| RMSD chiral (Å) | 0.079 |
| Ramachandran plot (%)[e] | |
| Favored | 98.3 |
| Allowed | 99.8 |
| Disallowed | 0.2 |
| Average B-factor (Å$^2$) | 27.9 |
| Macromolecules | 26.6 |
| Sulfates | 42.4 |
| Ethylene glycol | 31.1 |
| Water | 36.9 |

[a]Numbers in parentheses are values for the highest-resolution shell. [b]$R_{merge} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - \overline{I}(hkl)|/\Sigma_{hkl}\Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the *i*th observation of reflection *hkl*, and $\overline{I}(hkl)$ is the weighted average intensity for all observations *i* of reflection *hkl*. [c]$R_{meas} = \Sigma_{hkl}(N/(N-1)^{1/2})\Sigma_i|I_i(hkl) - \overline{I}(hkl)|/\Sigma_{hkl}\Sigma_i I_i(hkl)$. [d]$CC^* = (2CC_{1/2}/(1 + CC_{1/2}))^{1/2}$. [e]As defined by MolProbity.

the heteroatoms of ligands only. Clusters were ranked by the lowest estimated free energy of binding for each cluster. Free energy of binding is defined as the sum of final intermolecular energy (sum of the van der Waals, hydrogen bonding, and desolvation energy terms) and torsion free energy.[38] The lowest binding energy conformation of each ligand of the #1 ranked cluster was shown using PyMOL.

## RESULTS

**PtmT2 Is an *ent*-CPP Synthase.** The diterpenoid moieties of PTM and PTN are derived from the *ent*-kaurene and *ent*-atiserene scaffolds, respectively, supporting *ent*-CPP as the most advanced common intermediate in the biosynthesis of PTM and PTN (Figure 2A).[18] BLAST analysis of the proteins encoded by the *ptm* gene cluster from *S. platensis* CB00739 revealed that PtmT2 resembles the biochemically characterized *ent*-CPP synthase from the viguiepinol gene cluster found in *Streptomyces* sp. KO-3988 (46% identity, 58% similarity).[39] We cloned and expressed *ptmT2* from *S. platensis* CB00739[20] and overproduced PtmT2 in *E. coli* BL21 (DE3) (Figure S2). In contrast to *ent*-CPP synthase from *S.* sp. KO-3988, which appeared to form a homodimer,[39] PtmT2 ($His_6$-tag was removed) is a monomer in solution as determined by size-exclusion chromatography (Figure S3). Purified PtmT2 was incubated with GGPP and analysis of the reaction mixture by HPLC revealed the disappearance of GGPP and a concomitant increase of one new peak (Figure 2B). Isolation and characterization of this enzymatic product by NMR analysis confirmed that PtmT2 cyclizes GGPP into *ent*-CPP (SI).

The PtmT2 catalyzed reaction was optimized by measuring *ent*-CPP production in various conditions including different buffers, pH values, temperatures, and divalent cations. Optimized conditions for PtmT2 catalysis was established to be 50 mM MOPS, pH 6.5, containing 1 mM $MgCl_2$ (Figure S4A). GGPP was instable in MOPS buffer, however, and therefore citrate buffer was used for subsequent assays. As previously seen for type II DTSs,[39−41] formation of *ent*-CPP was inhibited when PtmT2 was incubated in the presence of 1 mM EDTA, suggesting divalent cations were necessary for activity. The highest enzyme activity was seen using $Mg^{2+}$, although incubations with other divalent cations were also active (Figure S4B). Using optimized conditions, the kinetic properties of PtmT2 were evaluated. The rate ($k_{cat}$) and Michaelis constant ($K_m$) were determined using a nonlinear fit of initial velocities versus [GGPP] (Figure 2C). The values of $k_{cat}$ and $K_m$ were determined to be $1.8 \pm 0.1$ s$^{-1}$ and $44 \pm 5$ $\mu$M ($k_{cat}/K_m = 4.1 \times 10^4$ s$^{-1}$ M$^{-1}$), respectively, correlating well with reported values for *At*CPS in saturated $Mg^{2+}$ conditions ($k_{cat} = 1.8$ s$^{-1}$)[41] and *ent*-CPP synthase from *S.* sp. KO-3988 ($K_m = 13.7$ $\mu$M),[39] although the $K_m$ of the former and the $k_{cat}$ of the latter were ∼65-fold smaller and ∼55-fold slower, respectively, than that of PtmT2. Although substrate inhibition of other TSs have been reported,[39,41] no substrate inhibition was seen for wild-type (WT) PtmT2 up to 500 $\mu$M GGPP (Figure 2C).

**Structure of a Bacterial *ent*-CPP Synthase.** We determined the crystal structure of PtmT2 at a resolution of 1.80 Å. Data collection and refinement statistics can be found in Table 1. The overall architecture of PtmT2 is a double $\alpha$-barrel type II TS fold ($\beta\gamma$) reminiscent of the TTSs SHC and human oxidosqualene cyclase (hOSC) (Figure 3A).[10] The bacterial PtmT2 also mirrors the $\beta\gamma$ domains of plant *ent*-CPP synthase (*A. thaliana*, *At*CPS)[12] and abietadiene synthase [*A. grandis*, *Ag*AS, the $\beta\gamma$ domains of *Ag*AS form (+)-CPP] (Figure 3B).[13] The $\beta\gamma$ structure of PtmT2, however, more closely resembles the $\beta\gamma$ domains of *At*CPS and *Ag*AS than those of SHC and hOSC (Tables S4−S5, Figure 3), an unsurprising discovery considering the substrate preferences and catalytic similarities between these DTSs, as well as the absence of the membrane-associating helix found in SHC and hOSC. As previously predicted,[9] the bacterial type II DTS structure does not contain a domain analogous to the nonfunctional, vestigial type I TS $\alpha$-domain of *At*CPS or the functional $\alpha$-domain of *Ag*AS (Figure 3B).

As with SHC and the plant DTSs, the $\gamma$ domain (residues P34−T262) of PtmT2 is inserted between the first ($\alpha$1) and second ($\alpha$11) helices of the $\beta$ domain (residues A11−A33 and P263−A525, Figure S5). These two domains form an interfacial, solvent-accessible active site cavity that places the $D^{311}$xDD motif at the base of the cavity (Figures 3B and 3C). The $N_\delta$ of H359 is 2.6 Å away from one of the oxygens in the carboxylate of D313, and likely directs and/or activates the


DOI: 10.1021/jacs.6b04317
J. Am. Chem. Soc. 2016, 138, 10905−10915

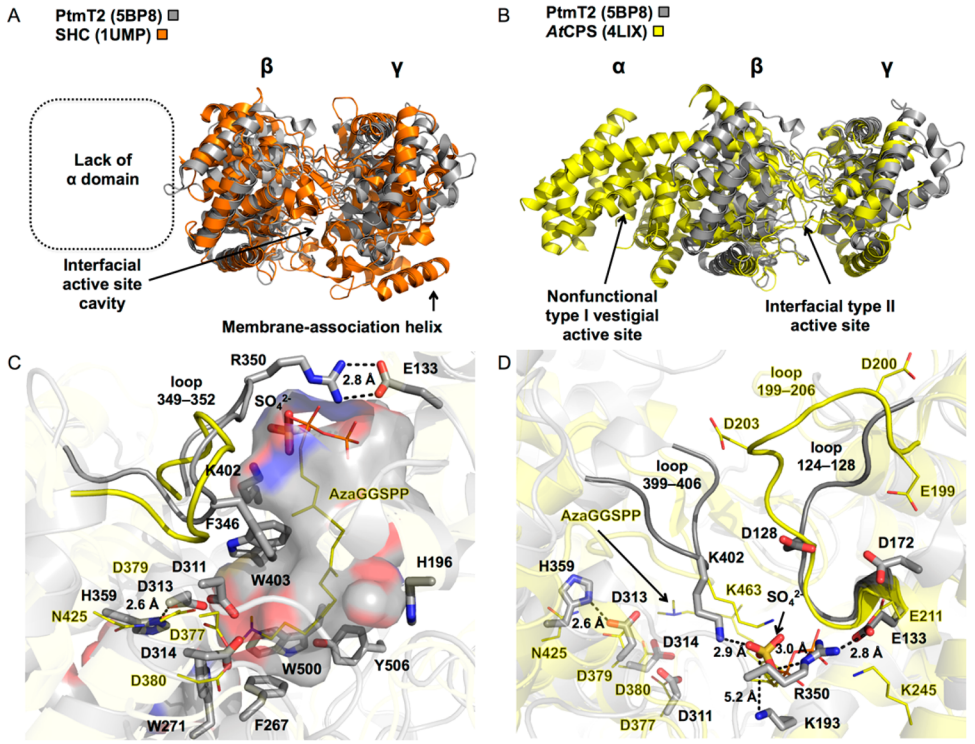



**Figure 3.** Overall structure of PtmT2, active site cavity, and proposed residues that bind the diphosphate moiety of GGPP or $Mg^{2+}$. (A) Superimposition of PtmT2 (PDB entry 5BP8, gray) with the bacterial type II TTS SHC (PDB entry 1UMP,[37] orange) showing the $\alpha$-helical $\beta\gamma$ fold that forms an interfacial active site cavity. Both PtmT2 and SHC lack the $\alpha$-domain. PtmT2 also lacks the membrane-association element found in SHC. (B) Superimposition of PtmT2 (gray) with plant type II DTS *At*CPS (PDB entry 4LIX,[44] yellow) showing the conserved $\beta\gamma$ fold and interfacial type II active site. PtmT2 lacks the $\alpha$-domain and thus the type I active site. (C) The interfacial active site cavity of PtmT2 highlighting the $D^{311}$xDD motif and H359 at the "bottom" of the cavity, the aromatic residues lining the cavity, and key residues (E133, R350, K402, loop 349−352) shaping the "top" of the cavity. Loop 399−406, which helps to form the "closed" conformation of the active site, is behind the binding pocket and is not shown for clarity (see panel D). The substrate mimic azaGGSPP and key homologous residues and loops from *At*CPS (yellow sticks) are superimposed for comparison. An omit map of the active site in (C) can be found in Figure S8. (D) A "top-down" view of the active site of PtmT2 depicting the residues proposed for diphosphate or $Mg^{2+}$ binding. PtmT2 does not possess the proposed EDxxD motif from plants (in loop 199−206), but instead has three negatively charged residues, D128 (in loop 124−128), E133, and D172, that surround a likely spot for a $Mg^{2+}$ ion coordinated to the diphosphate moiety of GGPP. These residues may be locked in a $Mg^{2+}$-binding conformation, as opposed to the EDxxD motif in plants. Positively charged K193, R350, K402 (in loop 399−406) are in close proximity to the expected diphosphate moiety position, which is partially mimicked in PtmT2 by a sulfate ion. The DxDD motif and H359 lie at the "bottom" of the active site cavity, away from the diphosphate and $Mg^{2+}$ ion binding location. The substrate mimic azaGGSPP and key homologous residues and loops from *At*CPS (yellow sticks), are superimposed for comparison.

general acid for protonation of C14 in GGPP. In the active site of SHC, as in PtmT2, a histidine hydrogen bonds with the catalytic aspartate, while plant *ent*-CPP synthases use asparagine residues instead of histidines, as seen for N425 in *At*CPS (Figure 3C). Sequence alignment of all functionally characterized type II DTSs of bacterial origin that use GGPP as a substrate reveals that this histidine is conserved and may be an indicator, along with the DxDD motif, of bacterial type II DTS function (Figure S5). Both Bra4 and PlaT2, bacterial DTSs that cyclize epoxyGGPP,[42,43] have glutamines, instead of histidines, in this position.

Matching the hydrophobic character of the diterpene chain, seven aromatic residues (H196, F267, W271, F346, W403, W500, and Y506), along with four aliphatic residues (I130, I261, A304, and L308) form the hydrophobic walls of the active site cavity (Figure 3C). Near the entrance of the cavity, on opposite sides, two lysine residues (K193 and K402) position their side chains toward each other, potentially interacting with the negatively charged diphosphate moiety (Figures 3C and 3D). K245 and K463 in *At*CPS, although spatially distinct from the corresponding residues in PtmT2, was similarly proposed to interact with the diphosphate moiety.[12] Also at the cavity opening, three negatively charged residues (D128, E133 and D172) seem poised to stabilize the diphosphate moiety by coordinating $Mg^{2+}$ (Figures 3C and 3D). Although the crystals were grown in the presence of $MgSO_4$, no $Mg^{2+}$ ions were found in the electron density. A bulge in loop E349−P352 from the $\beta$-domain protrudes across the entrance of the active site, allowing the side chain of R350 to form an ionic interaction (2.8 Å) with the side chain of E133 ($\gamma$-domain) (Figure 3C). It is unknown whether R350 lies in this position in the absence of GGPP or $Mg^{2+}$, but a $SO_4^{2-}$ ion, which may indicate where the negatively charged diphosphate moiety binds, was found bound to R350 and K402, suggesting a potential substrate recognition role for these positively charged residues.

Despite no ligand bound in the active site of PtmT2, the active site is in the "closed" conformation, previously seen in *At*CPS,[5] due to the W399−S406 loop folding down toward the active site opening (Figure 3D). This "closed" conformation suggests that the binding pocket of PtmT2 can support the catalytically active conformation of GGPP.[13] The side chains of K402 and W403, which are nearest to the active site cavity, are


DOI: 10.1021/jacs.6b04317
J. Am. Chem. Soc. 2016, 138, 10905−10915

in slightly different conformations than those of *At*CPS (K463 and W464) bound with the substrate analogue, although the extended conformation of the GGPP mimic (aza-14,15-dihydroGGSPP) in *At*CPS does not appear to mimic the precyclization conformation.[12]

**Active Site Mutations Distinguishing Catalytically Relevant Residues of Bacterial Type II DTSs.** Key questions remain unanswered regarding how and where the catalytically important $Mg^{2+}$ ion binds to type II DTSs. In *Ag*AS, a highly acidic EDxxD motif, chemically equivalent to the DDxxD $Mg^{2+}$-binding motif found in type I TSs, was proposed as the $Mg^{2+}$-binding motif.[9] However, in the crystal structure of *At*CPS, the equivalent $E^{199}$DxxD motif was too far (18 Å) from the diphosphate moiety, barring a substantial conformational change, to coordinate a $Mg^{2+}$ ion that also interacts with the substrate.[12] Another residue in *At*CPS, E211, was proposed to coordinate $Mg^{2+}$ given its proximity to the diphosphate moiety, but site-directed mutagenesis of this residue unaffected its $K_m$ while severely (500-fold) decreased its $k_{cat}$.[44] If E211 indeed coordinates the $Mg^{2+}$ ion, this enzyme-metal-substrate interaction may stabilize the transition state(s) of cyclization rather than contribute to substrate binding.

PtmT2 does not contain an EDxxD motif. The EDxxD motif-containing loop in *At*CPS (E199–M206) is replaced with a shorter and significantly less negatively charged loop (G124–D128) in PtmT2 (Figure 3D). D128, a conserved residue in characterized bacterial type II DTSs (Figure S5) but conspicuously absent in *At*CPS and *Ag*AS, as well as D172, an unconserved residue in DTSs, reside near the entrance to the active site cavity (Figure 3D). We hypothesized that D128 and D172, along with E133, the analogous residue to E211 from *At*CPS, work together to coordinate the $Mg^{2+}$ ion near the entrance of the active site cavity where the diphosphate moiety is positioned. We were also interested in whether the semiconserved residues K193 and R350, and the highly conserved K402 contribute to substrate (i.e., diphosphate) binding and activity.

All mutants generated by site-directed mutagenesis, except K402A, retained the ability to transform GGPP into *ent*-CPP, albeit at lower specific activities in all but two mutants (Figure 4, Table S6). Undetectable product formation in K402A implies that K402 is essential for substrate binding and/or activity. The restoration of activity in K402R demonstrates that a positively charged side chain is vital for binding and/or activity and supports K402 interaction with the diphosphate moiety of GGPP. The positively charged K193 and R350 residues also appear to be involved, but in a minor role. Of the three residues (D128, E133, and D172) proposed to form the $Mg^{2+}$-binding motif, mutagenesis of only D128 and E133 severely affected activity. Functional group deletion in the D128A and E133A mutants implicate a role for these negatively charged residues. Functional group replacement in D128E boosted activity 10-fold relative to D128A; similar functional replacement did not affect activity in E133D relative to E133A. The D172A and D172E mutants showed little to no change compared to WT PtmT2, suggesting they are not involved in $Mg^{2+}$-binding or the cyclization reaction. The double (D128A-E133A) and triple (D128A-E133A-D172A) mutants showed no further decreases in activity compared with the D128A and E133A mutants. Full kinetic characterization of the D128A, E133A, and D128-E133A mutants revealed that the $K_m$ values for GGPP were relatively (2-fold) unchanged, while the $k_{cat}$ values dropped significantly (~30–40-fold, and similar levels as the specific activities mentioned above) compared with those of WT PtmT2 (Figure S6). Finally, while no substrate inhibition was seen for WT PtmT2, high concentrations ($K_i$ = 206–420 $\mu$M and 2–4-fold greater than $K_m$) of GGPP inhibited product formation, indicative that the mutated residues altered the preferred substrate and/or $Mg^{2+}$ binding conformations. The exact mechanism of substrate inhibition in these mutants, however, remains unclear.

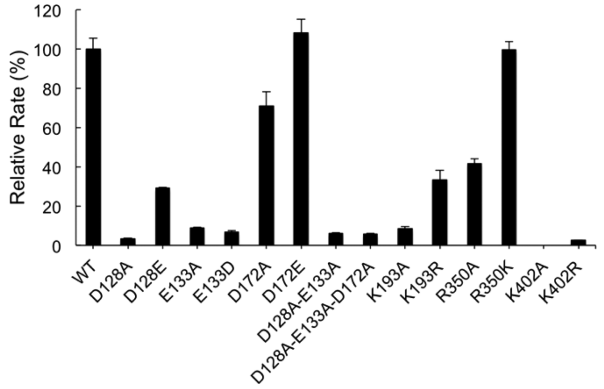



**Figure 4.** Relative activities of *ent*-CPP formation by PtmT2 mutants. No *ent*-CPP was detected using PtmT2 K402A. See Table S6 for a tabular depiction of the relative activities.

**Ligand Docking Studies Suggesting the Catalytically Active Conformation of GGPP.** In order to model the cyclization reaction catalyzed by PtmT2, *ent*-CPP and GGPP complexes were computed using AutoDock, allowing noncyclic single bonds to be rotatable in the ligands and keeping protein residues rigid. We identified a highly populated and lowest energy conformational cluster of the *ent*-CPP complex comprised of 111 structurally similar poses (RMSD ≤ 2.0 Å) out of 200 independent runs (~56%, Figure 5A). In addition, we identified the lowest energy conformation of the GGPP complex from the #1 ranked cluster of 21 similar poses (~11%, Figure 5B). The lowest energy substrate and product complexes show high similarities between the positioning and orientations of the aliphatic chains along the substrate cavity, diphosphate moiety locations, and interactions with active site residues. The GGPP conformer is oriented such that the distances between the carbons that will form C–C bonds during cyclization, C6–C11 and C10–C15, are 4.4 and 4.2 Å, respectively. A carboxylate oxygen of the general acid D313 is near C14 of both GGPP and *ent*-CPP (3.2 and 3.0 Å, respectively), an ideal position for protonation-initiated cyclization. These lowest binding energy computational models are consistent with the stereospecific cyclization reaction of converting GGPP into *ent*-CPP (Figure 5). Overall, C14 of GGPP is protonated by D313 in cooperation with H359, cyclization ensues resulting in a chair–chair conformation of the decalin ring with aromatic residues nearby and stabilizing the transient carbocations, and *ent*-CPP is formed via deprotonation of the C17 methyl (Figure 5C). Deprotonation is likely catalyzed by a water molecule hydrogen bonded to H196 and D503, as previously seen for *At*CPS.[12] Indeed, the crystal structure of PtmT2 shows electron density for a water molecule hydrogen bonded to H196 and D502. The donated proton from D313 is likely regenerated via another water molecule that is hydrogen bonded to Y409 (Figure 5C), in a similar manner as was proposed for SHC.[45,46] Reprotonation of


DOI: 10.1021/jacs.6b04317
J. Am. Chem. Soc. 2016, 138, 10905–10915

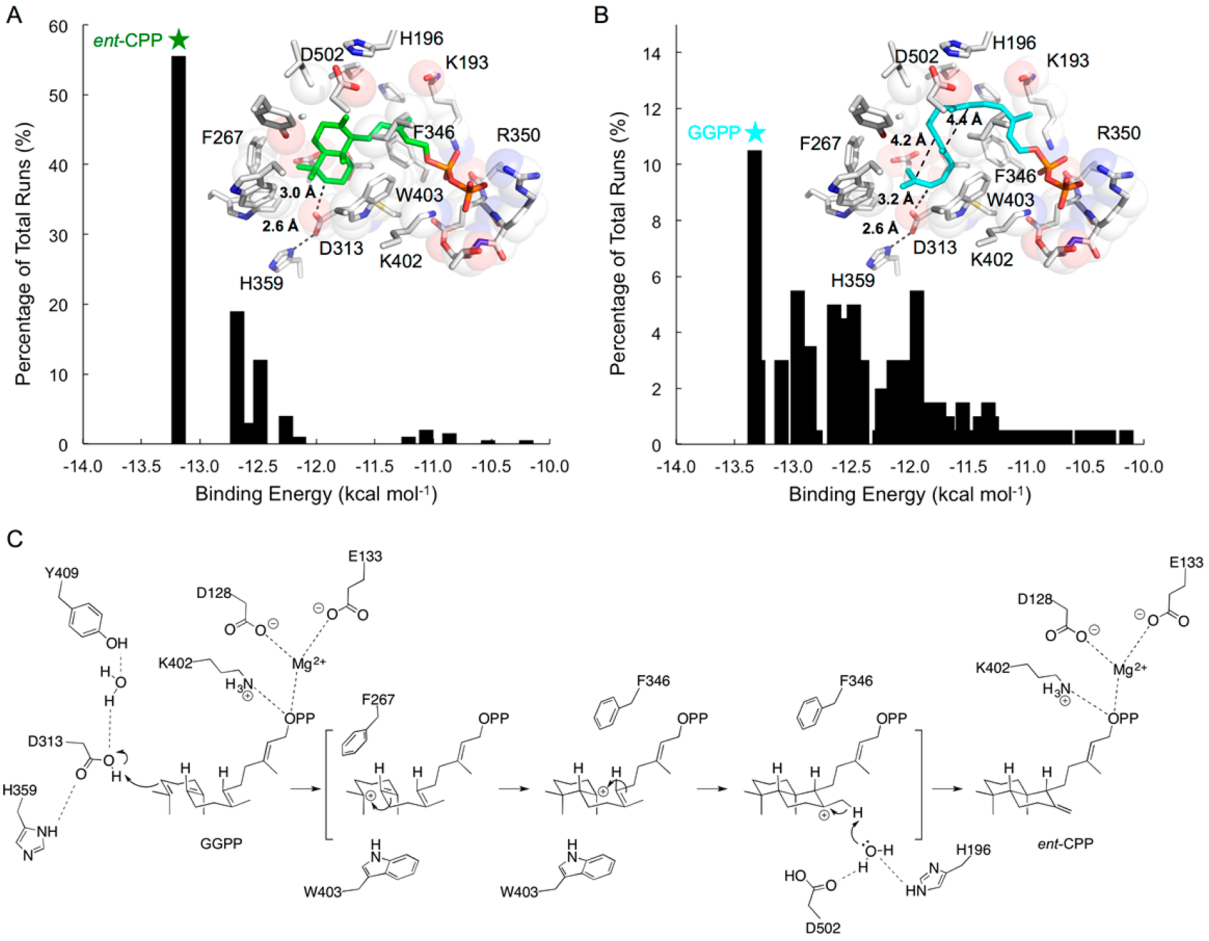



**Figure 5.** Catalytically active conformation of GGPP sets up the *ent*-CPP cyclization cascade. (A and B) Histograms of *ent*-CPP (product) and GGPP (substrate) conformations bound in the active site, respectively, generated using AutoDock. Insets are the lowest binding energy conformation of each ligand (bars labeled with stars in histogram). The ligand and corresponding active site residues are shown as sticks. Surrounding protein atoms within 4 Å of the ligand are displayed as transparent spheres. Selected residues involved in protonation, carbocation stabilization, and diphosphate coordination are labeled. Distances between D313 and H359, the proposed proton acceptor (C14) and donor (D313), and the C−C bond forming atoms in GGPP are shown. (C) Proposed cyclization mechanism highlighting formation of *ent*-CPP through protonation-initiation, stabilization of the transient carbocations by adjacent aromatic residues, and deprotonation. Only selected aromatic residues were shown to highlight carbocation stabilization, although several additional aromatic residues are also present. A water molecule hydrogen bonded to Y409 is proposed to reprotonate D313 after cyclization. The geometry of the carbon atoms in GGPP and each transition state intermediate are depicted to highlight the chair−chair conformation of *ent*-CPP formation.

the catalytic acid, D379, in *At*CPS was proposed to occur through a proton shuttle in a polar channel connecting D379 to another aspartic acid residue.[44]

## ■ DISCUSSION

PtmT2, the *ent*-CPP synthase for PTM and PTN biosynthesis from *S. platensis* CB00739, one of the few bacterial DTSs to be functionally characterized and the first bacterial type II DTS to be structurally characterized, serves as a model for understanding diterpene cyclization in bacteria. For example, given the identical catalytic outcomes of PtmT2 and the *ent*-CPP synthase from *Bradyrhizobium japonicum* (*Bj*CPS) it is likely the overall structures, active site cavity shapes, and catalytic residues are highly conserved between these two DTSs, even with differences in their sequences (31% identity, Figure S5) and bacterial sources (*S. platensis* is Gram-positive, while *B. japonicum* is Gram-negative). In fact, the DxDD motif, activating H359, and K402 are all strictly conserved between the two enzymes. The newly proposed $Mg^{2+}$-binding motif $D^{128}$xxxxE motif in PtmT2 has a single residue (alanine) insertion in *B. japonicum*, but retains the aspartate and glutamate residues (i.e., DxxxxxE) (Figure S5).

Determination of the $Mg^{2+}$-binding site in type II DTSs has been elusive, even after extensive site-directed mutagenesis,[9,12] computational,[9] and structural studies.[12,13,44] The structure of PtmT2 revealed a major difference in a loop near the active site entrance and the proposed diphosphate and $Mg^{2+}$ binding positions. This loop, containing a DxxxxE motif, was significantly shorter and less flexible than the corresponding EDxxD-motif containing loop in plant *ent*-CPP synthases. The bacterial DxxxxE motif is proposed to lock the $Mg^{2+}$ ion in a catalytically competent binding position, whereas the plant loop may need a significant conformational change for proper coordination. Site-directed mutagenesis supported the $D^{128}$xxxxE motif, as well as K402, as key players in the activity of PtmT2.

The emergence of the $\alpha\beta\gamma$ architecture found in type II TSs from plants is hypothesized to be due to the fusion of ancestral genes encoding the discrete $\alpha$ and $\beta\gamma$ proteins.[9] This bifunctional enzyme is a likely common ancestor for many


DOI: 10.1021/jacs.6b04317
J. Am. Chem. Soc. 2016, 138, 10905−10915

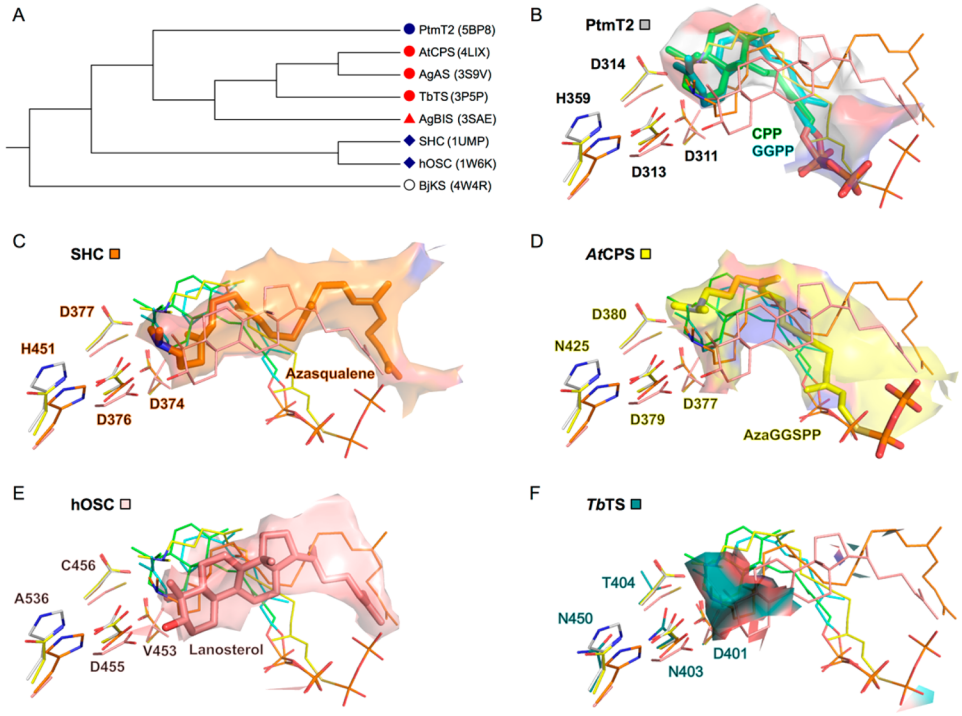



**Figure 6.** Structural comparisons of type I, type II, and bifunctional terpene synthases. (A) Structural similarities of the $\beta\gamma$-domains of selected terpene synthases. This dendrogram is based on Q-scores (shown in Table S5) calculated by PDBeFold[58] and generated using DendroUPGMA (genomes.urv.cat/UPGMA).[59] Labels indicate DTS (circles), TTS (diamonds), sesquiterpene synthase (triangle), $\beta\gamma$ architecture (blue), $\alpha\beta\gamma$ architecture (red), and $\alpha$ architecture (white). For proteins with $\alpha\beta\gamma$ architecture, the $\alpha$-domain was removed before Q-score calculation. The type I DTS *Bj*KS ($\alpha$-domain) was used as an outgroup. (B–F) Solvent accessible surfaces of the active site cavities of PtmT2 (PDB entry 5BP8), SHC (1UMP[57]), *At*CPS (4LIX[44]), hOSC (1W6K[47]), and *Tb*TS (3P5P[11]). Each active site cavity is shown with their DxDD motif or corresponding residues and their complexed ligands, i.e., *ent*-CPP (green) or GGPP (cyan) for PtmT2, azasqualene (orange) for SHC, azaGGSPP (yellow) for *At*CPS, lanosterol (pink) for hOSC. *Tb*TS does not have a ligand in the type II active site as it is a type I DTS with a nonfunctional type II active site. Structures were aligned based on the DxDD motifs (or corresponding residues). Carbon atoms and cavities are color-coded for each structure (PtmT2, gray; SHC, orange; *At*CPS, yellow; hOSC, pink; *Tb*TS, teal), with constant colors for noncarbon atoms (oxygen, red; nitrogen, blue; phosphorus, orange; sulfur, yellow). The ligands corresponding to the active site cavity are shown as sticks. Ligands in the aligned structures, the DxDD motif, and their corresponding residues are shown as thin lines.

current plant TSs,[14,15] although recent evidence suggests that there may be more than one evolutionary origin for plant TSs.[17] With structurally characterized examples of bacterial type I TSs ($\alpha$) and type II TTSs ($\beta\gamma$), as well as bifunctional DTSs ($\alpha\beta\gamma$) and mono- and sesqui-TSs ($\alpha\beta$ or $\alpha$) from plants, bacterial DTSs is a current knowledge gap in the structural puzzle. PtmT2, being the first bacterial type II DTS to be structurally characterized, is the one of the missing links in the structural evolution of TSs.

To try to understand the sequence-structure evolution of TSs, a sequence-based phylogenetic tree and structure-based dendrogram of the $\beta\gamma$ domains of relevant TSs were constructed. The sequences of PtmT2 and other characterized bacterial type II DTSs are more similar to those of the TTSs (SHC and hOSC) than the sequences of the $\beta\gamma$ domains of plant DTSs (Figure S7). This may suggest that the structure of PtmT2 would also be more similar to TTSs, a reasonable proposal since both PtmT2 and TTSs possess only the $\beta\gamma$ architecture. However, the structure of PtmT2 more closely resembles the $\beta\gamma$ domains of plant DTSs, and not with the TTSs (Figure 6A). The lack of a membrane association helix in the DTSs is partially responsible for this discrepancy; however, the structure, and consequently the active site, is more indicative of the function of TSs than the sequence alone. This basic idea presumably rings true in all TSs, and especially TSs of bacterial origin. TSs in bacteria, and DTSs in particular, are notorious for their diverse amino acid sequences, which severely limits their ability to be detected using current bioinformatics techniques.[1,3]

The commonly accepted mechanism for TSs is substrate activation (i.e., protonation or ionization) and stabilization of the transition state carbocations by aromatic and polar residues.[47,48] Quantum mechanics calculations of carbocation rearrangements have been exceptional at calculating geometries and relative energies of cyclization transition states, yielding insights into the mechanisms and stereospecificities of TSs;[49] however, it is not clear whether TSs are under thermodynamic or kinetic control.[50] PtmT2 additionally functions as a molecular template, previously described for type I TSs,[51,52] by favorably folding the substrate in a specific and catalytically competent conformation. Thus, PtmT2, and TSs in general, likely select specific reaction routes by exploiting the contour of the active site cavity to stabilize the substrate into the desired transition state or intermediate conformation that gives rise to stereospecific cyclization.

To further understand how the structures of TSs have evolved and the implications in their catalytic mechanisms, the structure of PtmT2 was aligned with other structurally characterized type II TSs and the active site cavities, bound- or modeled-ligands, and conserved DxDD motifs (VxDC in hOSC) were compared. Each type II TS (see Figure S1 for the reactions catalyzed by the selected type II TSs) shows a distinct


DOI: 10.1021/jacs.6b04317
J. Am. Chem. Soc. 2016, 138, 10905–10915

active site cavity shape, which likely favors each stereospecific cyclization cascade as evidenced by the solvent-accessible surface boundary (Figures 6B−6F). The low B-factor values of the residues forming the active site cavities in these TSs, especially those of PtmT2, suggest that the overall shape of the active site are relatively rigid and do not undergo major conformational changes for catalysis. As seen in the structure-based dendrogram of the $\beta\gamma$ domains of TSs (Figure 6A), functionally similar type II TSs adopt structurally similar active site cavity shapes (i.e., PtmT2 ∼ *At*CPS, SHC ∼ hOSC). Interestingly, soluble type II DTSs (PtmT2 and *At*CPS; Figures 6B and 6D) and membrane-bound type II TTSs (SHC and hOSC; Figures 6C and 6E) display different shapes of their active site cavities despite (i) higher sequence homologies between bacterial type II DTSs and TTSs, (ii) PtmT2, SHC, and hOSC all possessing $\beta\gamma$ architectures, while *At*CPS contains both type I and type II TS domains and an $\alpha\beta\gamma$ architecture, and (iii) PtmT2, *At*CPS, and SHC all contain the conserved DxDD motif, while hOSC not only has replaced the DxDD motif with VxDC, but also lacks the conserved activating His/Asn. Furthermore, taxadiene synthase (*Tb*TS), an $\alpha\beta\gamma$ type I DTS that contains a nonfunctional $\beta\gamma$ domain and lacks the DxDD motif (DxNT), lost its type II active site cavity during evolution (Figure 6F). It is unclear when *Tb*TS lost its type II DTS activity and if loss of the DxDD motif preempted collapse of the active site cavity, or vice versa, or they simultaneously evolved. These observations underline the importance of both the shape and physicochemical properties of the active site in determining specific catalytic outcomes of TSs. Two structurally similar proteins, with no significant sequence homology, that catalyze identical chemical reactions, as in the case of PtmT2 and *At*CPS, may support a convergent evolution model for *ent*-CPP synthases in bacteria and plants.

Although stereochemical control can be achieved by TSs including PtmT2, this is often not the case as evidenced by leaky alternative cyclization cascades that result in a variety of side products.[53,54] It should be noted that while the *ent*-CPP-bound complex showed dominant conformational clusters (Figure 5A), GGPP showed weaker discrimination between conformers (Figure 5B). These weak discriminations may account for the alternative cyclization cascades in other TSs; however, most of the higher energy conformers of GGPP-bound to PtmT2 displayed catalytically ineffective distances or orientations to D313.

The extreme discrepancy in total numbers of diterpenoids found in plants/fungi vs bacteria suggests either bacteria do not have the necessary biosynthetic machinery (i.e., DTSs), have not evolved DTSs for secondary metabolism on a comparable scale to that of plants and fungi, or the community at large has been unable to develop efficient means of diterpenoid discovery in bacteria. Gibberellin biosynthesis in bacteria supports bacteria as the ancestral carriers of DTSs and therefore bacteria appear to have had ample access to DTSs for secondary metabolism evolution. Yet, the diverse nature of diterpenoids in plants dominates the chemical space compared with the handful of bacterial diterpenoid scaffolds. Have plants evolved to produce these diverse scaffolds while bacterial DTS evolution remained stagnant? PTM and PTN, two natural products of bacterial origin, offer a unique perspective. *ent*-Kauranol and *ent*-atiserene, derived from the common scaffold of *ent*-CPP for PTM and PTN biosynthesis, respectively, are processed and modified in ways not seen in any other organism or biosynthetic pathway. This supports that bacteria have taken their own path to creating structural diversity and that more, novel diterpenoid natural products from bacteria are waiting to be discovered.

## CONCLUSION

Terpene synthases are ubiquitous enzymes that attract considerable attention given that their complex mechanisms generate diverse carbon skeletons in a regio- and stereoselective manner. Accordingly, diterpenoids and DTSs from plants and fungi have been and continue to be widely studied. Diterpenoids and DTSs of bacterial origin, due to their scarcity, remain underappreciated. PtmT2, the *ent*-CPP synthase for PTM and PTN biosynthesis from *S. platensis* CB00739, one of the few bacterial DTSs to be functionally characterized and the first bacterial type II DTS to be structurally characterized, serves as a model for understanding diterpene cyclization in bacteria. The structure of PtmT2 also gives insight into the conserved, or preserved, evolution of active sites in DTSs of bacterial and plant origin. Discovering the differences between DTSs of various organisms and how these DTSs have evolved will prove to be invaluable in not only understanding how DTSs, and TSs in general, function, but also in the continued search for DTSs, and thereby the discovery of diterpenoids, in bacteria.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b04317.

Synthesis of GGPP; $^{1}H$, $^{13}C$, and $^{31}P$ NMR data for *ent*-CPP; strains, plasmids, and primers used in this study (Tables S1−S3); sequence and structure comparisons of selected TSs (Table S4); Q-score matrix of $\beta\gamma$-domains of selected TSs (Table S5); relative activities of PtmT2 mutants (Table S6); enzymatic reactions of TSs discussed in this study (Figure S1); SDS-PAGE of PtmT2 (Figure S2); size-exclusion chromatography of PtmT2 (Figure S3); PtmT2 reaction optimization (Figure S4); sequence alignment of functionally characterized type II DTSs from bacteria (Figure S5); kinetic characterization of five PtmT2 mutants (Figure S6); phylogenetic analysis of PtmT2 with selected TSs (Figure S7); electron density map for selected active site residues in PtmT2 (Figure S8) (PDF)

## AUTHOR INFORMATION


### Corresponding Author

*shenb@scripps.edu

### Author Contributions

#J.D.R. and L.-B.D. contributed equally to this work.


### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank Prof. C. Dale Poulter at the University of Utah for the generous gift of GGSPP. This work is supported in part by the National Institute of General Medical Sciences Protein Structure Initiative Grants GM094585 (AJ) and GM098248 (G.N.P.), and National Institutes of Health Grants GM109456 (G.N.P.) and GM114353 (BS). The use of Structural Biology Center beamlines at the Advanced Photon Source was



DOI: 10.1021/jacs.6b04317
J. Am. Chem. Soc. 2016, 138, 10905−10915

supported by U.S. Department of Energy, Office of Biological and Environmental Research grant DE-AC02-06CH11357 (AJ).

## REFERENCES


(1) Smanski, M. J.; Peterson, R. M.; Huang, S.-X.; Shen, B. *Curr. Opin. Chem. Biol.* **2012**, *16*, 132.
(2) Cane, D. E.; Ikeda, H. *Acc. Chem. Res.* **2012**, *45*, 463.
(3) Yamada, Y.; Kuzuyama, T.; Komatsu, M.; Shin-ya, K.; Omura, S.; Cane, D. E.; Ikeda, H. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 857.
(4) Christianson, D. W. *Chem. Rev.* **2006**, *106*, 3412.
(5) Gao, Y.; Honzatko, R. B.; Peters, R. J. *Nat. Prod. Rep.* **2012**, *29*, 1153.
(6) Aaron, J. A.; Christianson, D. W. *Pure Appl. Chem.* **2010**, *82*, 1585.
(7) Prisic, S.; Xu, J.; Coates, R. M.; Peters, R. J. *ChemBioChem* **2007**, *8*, 869.
(8) Tarshis, L. C.; Yan, M.; Poulter, C. D.; Sacchettini, J. C. *Biochemistry* **1994**, *33*, 10871.
(9) Cao, R.; Zhang, Y.; Mann, F. M.; Huang, C.; Mukkamala, D.; Hudock, M. P.; Mead, M. E.; Prisic, S.; Wang, K.; Lin, F.-Y.; Chang, T.-K.; Peters, R. J.; Oldfield, E. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 2417.
(10) Wendt, K. U.; Poralla, K.; Schulz, G. E. *Science* **1997**, *277*, 1811.
(11) Koeksal, M.; Jin, Y.-H.; Coates, R. M.; Croteau, R.; Christianson, D. W. *Nature* **2011**, *469*, 116.
(12) Koeksal, M.; Hu, H.; Coates, R. M.; Peters, R. J.; Christianson, D. W. *Nat. Chem. Biol.* **2011**, *7*, 431.
(13) Zhou, K.; Gao, Y.; Hoy, J. A.; Mann, F. M.; Honzatko, R. B.; Peters, R. J. *J. Biol. Chem.* **2012**, *287*, 6840.
(14) Bohlmann, J.; Meyer-Gauen, G.; Croteau, R. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 4126.
(15) Trapp, S. C.; Croteau, R. B. *Genetics* **2001**, *158*, 811.
(16) Keeling, C. I.; Dullat, H. K.; Yuen, M.; Ralph, S. G.; Jancsik, S.; Bohlmann, J. *Plant Physiol.* **2010**, *152*, 1197.
(17) Li, G.; Kollner, T. G.; Yin, Y.; Jiang, Y.; Chen, H.; Xu, Y.; Gershenzon, J.; Pichersky, E.; Chen, F. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 14711.
(18) Smanski, M. J.; Yu, Z.; Casper, J.; Lin, S.; Peterson, R. M.; Chen, Y.; Wendt-Pienkowski, E.; Rajski, S. R.; Shen, B. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 13498.
(19) Rudolf, J. D.; Dong, L.-B.; Huang, T.; Shen, B. *Mol. BioSyst.* **2015**, *11*, 2717.
(20) Hindra; Huang, T.; Yang, D.; Rudolf, J. D.; Xie, P.; Xie, G.; Teng, Q.; Lohman, J. R.; Zhu, X.; Huang, Y.; Zhao, L.-X.; Jiang, Y.; Duan, Y.; Shen, B. *J. Nat. Prod.* **2014**, *77*, 2296.
(21) Zolnai, Z.; Lee, P. T.; Li, J.; Chapman, M. R.; Newman, C. S.; Phillips, G. N., Jr.; Rayment, I.; Ulrich, E. L.; Volkman, B. F.; Markley, J. L. *J. Struct. Funct. Genomics* **2003**, *4*, 11.
(22) Lohman, J. R.; Bingman, C. A.; Phillips, G. N.; Shen, B. *Biochemistry* **2013**, *52*, 902.
(23) Aslanidis, C.; De Jong, P. J. *Nucleic Acids Res.* **1990**, *18*, 6069.
(24) Ho, S. N.; Hunt, H. D.; Horton, R. M.; Pullen, J. K.; Pease, L. R. *Gene* **1989**, *77*, 51.
(25) Eschenfeldt, W. H.; Stols, L.; Millard, C. S.; Joachimiak, A.; Donnelly, M. I. *Methods Mol. Biol.* **2009**, *498*, 105.
(26) Stols, L.; Millard, C. S.; Dementieva, I.; Donnelly, M. I. *J. Struct. Funct. Genomics* **2004**, *5*, 95.
(27) Kim, Y.; Babnigg, G.; Jedrzejczak, R.; Eschenfeldt, W. H.; Li, H.; Maltseva, N.; Hatzos-Skintges, C.; Gu, M.; Makowska-Grzyska, M.; Wu, R.; An, H.; Chhor, G.; Joachimiak, A. *Methods* **2011**, *55*, 12.
(28) Roe, S. J.; Oldfield, M. F.; Geach, N.; Baxter, A. *J. Labelled Compd. Radiopharm.* **2013**, *56*, 485.
(29) Davisson, V. J.; Woodside, A. B.; Neal, T. R.; Stremler, K. E.; Muehlbacher, M.; Poulter, C. D. *J. Org. Chem.* **1986**, *51*, 4768.
(30) Rosenbaum, G.; Alkire, R. W.; Evans, G.; Rotella, F. J.; Lazarski, K.; Zhang, R. G.; Ginell, S. L.; Duke, N.; Naday, I.; Lazarz, J.; Molitsky, M. J.; Keefe, L.; Gonczy, J.; Rock, L.; Sanishvili, R.; Walsh, M. A.; Westbrook, E.; Joachimiak, A. *J. Synchrotron Radiat.* **2006**, *13*, 30.
(31) Minor, W.; Cymborowski, M.; Otwinowski, Z.; Chruszcz, M. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2006**, *62*, 859.
(32) Terwilliger, T. C.; Grosse-Kunstleve, R. W.; Afonine, P. V.; Moriarty, N. W.; Zwart, P. H.; Hung, L. W.; Read, R. J.; Adams, P. D. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2008**, *64*, 61.
(33) Adams, P. D.; Afonine, P. V.; Bunkoczi, 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.; Zwart, P. H. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 213.
(34) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 486.
(35) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1997**, *53*, 240.
(36) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 235.
(37) Chen, V. B.; Arendall, W. B., III; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 12.
(38) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. *J. Comput. Chem.* **2009**, *30*, 2785.
(39) Ikeda, C.; Hayashi, Y.; Itoh, N.; Seto, H.; Dairi, T. *J. Biochem.* **2007**, *141*, 37.
(40) Hamano, Y.; Kuzuyama, T.; Itoh, N.; Furihata, K.; Seto, H.; Dairi, T. *J. Biol. Chem.* **2002**, *277*, 37098.
(41) Prisic, S.; Peters, R. J. *Plant Physiol.* **2007**, *144*, 445.
(42) Hayashi, Y.; Matsuura, N.; Toshima, H.; Itoh, N.; Ishikawa, J.; Mikami, Y.; Dairi, T. *J. Antibiot.* **2008**, *61*, 164.
(43) Duerr, C.; Schnell, H.-J.; Luzhetskyy, A.; Murillo, R.; Weber, M.; Welzel, K.; Vente, A.; Bechthold, A. *Chem. Biol.* **2006**, *13*, 365.
(44) Koeksal, M.; Potter, K.; Peters, R. J.; Christianson, D. W. *Biochim. Biophys. Acta, Gen. Subj.* **2014**, *1840*, 184.
(45) Wendt, K. U.; Lenhart, A.; Schulz, G. E. *J. Mol. Biol.* **1999**, *286*, 175.
(46) Wendt, K. U.; Schulz, G. E.; Corey, E. J.; Liu, D. R. *Angew. Chem., Int. Ed.* **2000**, *39*, 2812.
(47) Thoma, R.; Schulz-Gasch, T.; D'Arcy, B.; Benz, J.; Aebi, J.; Dehmlow, H.; Hennig, M.; Stihle, M.; Ruf, A. *Nature* **2004**, *432*, 118.
(48) Dougherty, D. A. *Science* **1996**, *271*, 163.
(49) Tantillo, D. J. *Nat. Prod. Rep.* **2011**, *28*, 1035.
(50) Vedula, L. S.; Rynkiewicz, M. J.; Pyun, H.-J.; Coates, R. M.; Cane, D. E.; Christianson, D. W. *Biochemistry* **2005**, *44*, 6153.
(51) Shishova, E. Y.; Di Costanzo, L.; Cane, D. E.; Christianson, D. W. *Biochemistry* **2007**, *46*, 1941.
(52) Christianson, D. W. *Curr. Opin. Chem. Biol.* **2008**, *12*, 141.
(53) Lodeiro, S.; Xiong, Q.; Wilson, W. K.; Kolesnikova, M. D.; Onak, C. S.; Matsuda, S. P. T. *J. Am. Chem. Soc.* **2007**, *129*, 11213.
(54) Steele, C. L.; Crock, J.; Bohlmann, J.; Croteau, R. *J. Biol. Chem.* **1998**, *273*, 2078.
(55) Liu, W.; Zheng, Y.; Feng, X.; Huang, C.-H.; Chen, C.-C.; Cui, Y.; Li, J.; Guo, R.-T.; Nakano, C.; Hoshino, T.; Bogue, S.; Oldfield, E.; Ko, T.-P.; Wang, I.; Hsu Shang-Te, D. *Sci. Rep.* **2014**, *4*, 6214.
(56) Whittington, D. A.; Wise, M. L.; Urbansky, M.; Coates, R. M.; Croteau, R. B.; Christianson, D. W. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 15375.
(57) Reinert, D. J.; Balliano, G.; Schulz, G. E. *Chem. Biol.* **2004**, *11*, 121.
(58) Krissinel, E.; Henrick, K. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2256.
(59) Garcia-Vallve, S.; Palau, J.; Romeu, A. *Mol. Biol. Evol.* **1999**, *16*, 1125.




DOI: 10.1021/jacs.6b04317
*J. Am. Chem. Soc.* 2016, 138, 10905−10915