



Crystal Structure of Geranylgeranyl Pyrophosphate Synthase (CrtE) Involved in Cyanobacterial Terpenoid Biosynthesis

Yuchi Feng¹, R. Marc L. Morgan², Paul D. Fraser³, Klaus Hellgardt¹ and Peter J. Nixon^{2*}

¹ Department of Chemical Engineering, Imperial College London, London, United Kingdom, ² Sir Ernst Chain Building-Wolfson Laboratories, Department of Life Sciences, Imperial College London, London, United Kingdom, ³ School of Biological Sciences, Royal Holloway, University of London, Egham, United Kingdom

OPEN ACCESS

Edited by:

Matteo Ballottari,
University of Verona, Italy

Reviewed by:

Eckhard Hofmann,
Ruhr University Bochum, Germany
Hiroshi Noguchi,
Nihon Pharmaceutical University,
Japan

*Correspondence:

Peter J. Nixon
p.nixon@imperial.ac.uk

Specialty section:

This article was submitted to
Plant Biotechnology,
a section of the journal
Frontiers in Plant Science

Received: 18 October 2019

Accepted: 20 April 2020

Published: 25 May 2020

Citation:

Feng Y, Morgan RML, Fraser PD,
Hellgardt K and Nixon PJ (2020)
Crystal Structure of Geranylgeranyl
Pyrophosphate Synthase (CrtE)
Involved in Cyanobacterial Terpenoid
Biosynthesis.
Front. Plant Sci. 11:589.
doi: 10.3389/fpls.2020.00589

Cyanobacteria are photosynthetic prokaryotes that perform oxygenic photosynthesis. Due to their ability to use the photon energy of sunlight to fix carbon dioxide into biomass, cyanobacteria are promising hosts for the sustainable production of terpenoids, also known as isoprenoids, a diverse class of natural products with potential as advanced biofuels and high-value chemicals. However, the cyanobacterial enzymes involved in the biosynthesis of the terpene precursors needed to make more complicated terpenoids are poorly characterized. Here we show that the predicted type II prenyltransferase CrtE encoded by the model cyanobacterium *Synechococcus* sp. PCC 7002 is homodimeric and able to synthesize C20-geranylgeranyl pyrophosphate (GGPP) from C5-isopentenyl pyrophosphate (IPP) and C5-dimethylallyl pyrophosphate (DMAPP). The crystal structure of CrtE solved to a resolution of 2.7 Å revealed a strong structural similarity to the large subunit of the heterodimeric geranylgeranyl pyrophosphate synthase 1 from *Arabidopsis thaliana* with each subunit containing 14 helices. Using mutagenesis, we confirmed that the fourth and fifth amino acids (Met-87 and Ser-88) before the first conserved aspartate-rich motif (FARM) play important roles in controlling chain elongation. While the WT enzyme specifically produced GGPP, variants M87F and S88Y could only generate C15-farnesyl pyrophosphate (FPP), indicating that residues with large side chains obstruct product elongation. In contrast, replacement of M87 with the smaller Ala residue allowed the formation of the longer C25-geranyl-farnesyl pyrophosphate (GFPP) product. Overall, our results provide new structural and functional information on the cyanobacterial CrtE enzyme that could lead to the development of improved cyanobacterial platforms for terpenoid production.

Keywords: isoprenoid, prenyltransferase, site-directed mutagenesis, FARM, phylogenetic analysis

INTRODUCTION

Cyanobacteria are a group of gram-negative photosynthetic prokaryotes that perform oxygenic photosynthesis using sunlight to drive the conversion of CO₂ into a variety of carbon-based compounds. Cyanobacteria are therefore widely considered as promising hosts for the sustainable production of biofuels and commodity and high-value chemicals. Terpenoids are a major class

of secondary metabolites with over 70,000 compounds identified (Vickers et al., 2014) including carotenoids, sterols, steroids, saponins, and hormones, which are widely used in a range of applications, including pharmaceuticals, pesticides, fragrances, flavors, and advanced biofuels (Rabinovitch-Deere et al., 2013).

All terpenoids are synthesized from two isomeric C5 building blocks: isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP), both produced via the 2-C-methyl-D-erythritol-4-phosphate (MEP) pathway in cyanobacteria (Figure 1). Prenyltransferase enzymes catalyze the initial condensation reaction between IPP and DMAPP to give C10-geranyl pyrophosphate (GPP), plus the subsequent addition of IPP molecules to give first C15-farnesyl pyrophosphate (FPP), and then C20-geranylgeranyl pyrophosphate (GGPP), which are the precursors of monoterpenoids, sesquiterpenoids and diterpenoids, respectively (Han et al., 2015). A diverse range of polyprenyl pyrophosphate products (>C25) can be further synthesized through addition of different numbers of IPP to the FPP allylic substrate (Wallrapp et al., 2013).

Prenyltransferases are classified as *cis* (Z) or *trans* (E) depending on the stereochemistry of the carbon-carbon double bonds in the product (Shimizu et al., 1998). Amino acid sequence alignments of different E- and Z-type enzymes suggests that *trans*-prenyltransferases have evolved from a common ancestor, whereas *cis*-prenyltransferases may have a different origin (Chen et al., 1994). In general, *trans*-prenyltransferases generate products up to C50, which can be further classified into short-chain (C10–C25), medium-chain (C30–C35), and long-chain (C40–C50) (Ogura and Koyama, 1998). GPP, FPP, and GGPP are synthesized by geranyl pyrophosphate synthase (GPPS), farnesyl pyrophosphate synthase (FPPS), and geranylgeranyl pyrophosphate synthase (GGPPS), respectively. C25 compounds, formed by geranylgeranyl pyrophosphate synthase (GGPPS), are used in the formation of thermophilic archaeal ether-linked lipids, while hexaprenyl pyrophosphate synthase (HexPPS), heptaprenyl pyrophosphate synthase (HepPPS), octaprenyl pyrophosphate synthase (OPPS), solanesyl pyrophosphate synthase (SPPS), and decaprenyl pyrophosphate synthase (DPPS) generate the side chains needed for the formation of quinones (Wang and Ohnuma, 1999).

All these *trans*-prenyltransferases contain two highly conserved aspartate-rich motifs: the “first aspartate-rich motif” (FARM, DDX_{2–4}D, where X encodes any amino acid) and the “second aspartate-rich motif” (SARM, DDXXD). In previous studies, a sequential ionization-condensation-elimination catalytic mechanism was proposed, in which FARM and SARM coordinate Mg²⁺ ions to promote elimination of the pyrophosphate group from the allylic substrate (such as DMAPP, GPP, and FPP) to generate a carbocation intermediate, which is subsequently attacked by IPP, bound to positively charged residues in the cavity, to form an adduct that is then deprotonated to form a carbon-carbon double bond (Liang, 2009; Chang et al., 2012).

Among the prenyltransferase enzymes, GGPPS has been considered a bottle-neck enzyme that regulates terpenoid biosynthesis (Jiang et al., 2012; Bai et al., 2016). In the case of photosynthetic organisms, most photosynthesis-related

metabolites such as carotenoids and the side chains of chlorophylls and plastoquinones are all derived from GGPP, synthesized by GGPPS (Ruiz-Sola et al., 2016). GGPPS enzymes can be divided into three types: Type-I enzymes contain a large aromatic amino acid at the fourth or fifth position before the FARM; Type-II contain an insertion of two residues in the FARM region but are devoid of aromatic residues at the fourth or fifth position before the FARM; and Type-III are also devoid of aromatic residues just before the FARM but have no extra residues inserted in the FARM (Hemmi et al., 2003; Bouvier et al., 2005; Chang et al., 2006). Structurally, GGPPS enzymes are found mainly as homodimers such as in the archaeon *Geoglobus acetivorans*, the bacterium *Bacteroides thetaiotaomicron*, and the eukaryotes *Plasmodium vivax*, *Saccharomyces cerevisiae*, *Sinapis alba*, and *Arabidopsis thaliana* (Chang et al., 2006; Kloer et al., 2006; Artz et al., 2011; Wallrapp et al., 2013; Wang et al., 2016; Petrova et al., 2018). However, some GGPPS are heterodimers, such as in the cases of *Oryza sativa* and *Mentha piperita* (Chang et al., 2010; Zhou et al., 2017) and higher order oligomers have been described for *Homo sapiens* (Kavanagh et al., 2006).

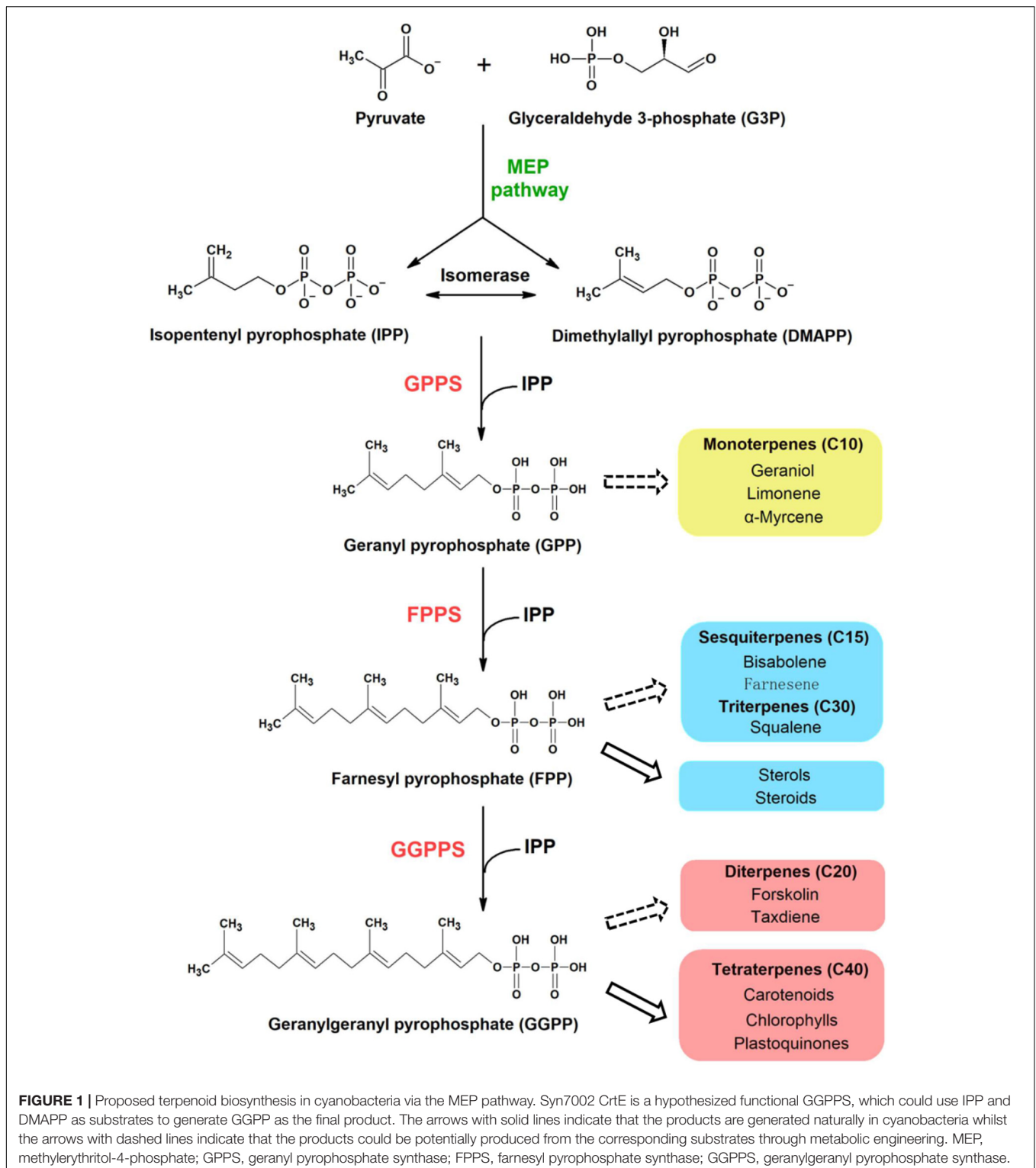
Little is known about the GGPPS enzymes in photosynthetic microorganisms. In cyanobacteria, sequence comparisons have identified a potential GGPPS (Liang et al., 2006), denoted CrtE because of its role in carotenoid biosynthesis (Yen and Marrs, 1976). An early study identified a gene encoding GGPPS in the thermophilic cyanobacterium *Thermosynechococcus elongatus* (Ohto et al., 1999). However, there is still no information on the structure of this enzyme and how product chain length is regulated.

In this study, we report the first crystal structure of GGPPS (CrtE) from a cyanobacterium – in our case the model strain *Synechococcus* sp. PCC 7002 (Syn7002). Based on the structure, site-directed mutagenesis experiments were carried out to investigate the role of specific amino acids in determining the chain length of the product.

MATERIALS AND METHODS

Cloning, Expression, and Purification

The *crtE* gene (SYNPCC7002_A1085) was amplified by PCR using genomic DNA of Syn7002 and primers CrtE_F (5'-CTCCGCGGTGGTATGGTAGTTGCAGA-3') and CrtE_R (5'-CGGAATTCCTTAGTTTTTACGGTTAACGATG-3') (restriction sites are underlined). The PCR products were digested with EcoRI and SacII, and the DNA fragments were cloned into the pHUE expression vector (Catanzariti et al., 2004) to give pHUE_CrtE (Supplementary Dataset 1). Protein expression was performed using *E. coli* BL21(DE3) (New England Biolabs, Ltd). Plasmid Usp2-cc which encodes histidine-tagged deubiquitinase was also transformed into *E. coli* BL21(DE3). The His-tagged recombinant proteins were overexpressed in *E. coli* BL21(DE3) in 3 L of Luria-Bertani medium containing 100 mg L⁻¹ ampicillin by growing to an optical density at 600 nm of 0.6 with 250 rpm at 37°C (Eppendorf Innova 43 incubator shaker) and then inducing with 0.4 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG). After additional overnight incubation at 18°C, the cells were



harvested and disrupted with a T5 cell disruptor (Constant system, GA, United States) using a pressure of 25 kpsi at 5°C in 40 mL of KPN buffer containing 40 mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$ (pH 8.0) and 100 mM NaCl. The homogenate was centrifuged with a Beckman Coulter Optima L-100XP ultracentrifuge and

a Ti45 rotor (Beckman Coulter, Inc., United States) at 172,000 x g for 30 min at 4°C, and the supernatant was recovered as a crude extract. To the 40 mL of supernatant, 2 mL of a 50% slurry of nickel-nitrilotriacetic acid (Ni-NTA) agarose in KPN buffer was added, then placed on a rotary wheel at

4°C for 1.5 h. The lysate/Ni-NTA mixture was centrifuged at 2,000 rpm (805 x g) for 2 min (EppendorfTM 5810 R Centrifuges with a A-4-62 Model rotor), and the Ni-NTA agarose pellet was washed twice with 20 mL of buffer A (25 mM Tris-HCl pH 7.5, 500 mM NaCl) and then twice with 20 mL of buffer A containing 60 mM imidazole. The His-tagged proteins (either deubiquitinase or CrtE fusion protein) were eluted from the Ni-NTA resin in 20 mL of buffer A containing 500 mM imidazole. The purified proteins were dialyzed against 5 L of buffer A at 4°C overnight, and the protein concentration was measured by the Bradford assay using bovine serum albumin (BSA) to generate the standard curve (Zor and Selinger, 1996). Cleavage of the CrtE fusion protein was achieved by incubating with deubiquitinase at a 1:10 molar ratio (enzyme to substrate) overnight at 16°C in 40 mL of buffer A (Catanzariti et al., 2004). Cleaved products were concentrated with a Sartorius 30 K concentrator (Sartorius Stedim Lab Ltd, United Kingdom) at 4,000 rpm (3,220 x g) at 4°C (EppendorfTM 5810 R Centrifuge with a A-4-62 Model rotor) until the final volume reduced to about 1 mL and then purified using an ÄKTA Prime FPLC system (GE Healthcare, United Kingdom) equipped with a HiLoad[®] 16/60 Superdex[®]200 prep grade (pg) (120 mL) column (GE Healthcare, United Kingdom) and equilibrated with 25 mM Tris-HCl (pH 7.5) and 500 mM NaCl buffer in a cold room (4°C). Chromatography was performed at a flow rate of 1 mL min⁻¹ and 2 mL fractions were collected. Absorbance of protein was measured at 280 nm (A280). Fractions contained purified CrtE and His-tagged ubiquitin protein were collected and analyzed by SDS-PAGE and immunoblotting (Michoux et al., 2014). Fractions containing CrtE were concentrated with a Sartorius 30 K MWCO concentrator (Sartorius Stedim Lab Ltd, United Kingdom) at 4,000 rpm, 4°C (EppendorfTM 5810 R Centrifuges with a A-4-62 Model rotor) until the final concentration reached 11.3 mg mL⁻¹. The final predicted protein sequence after cleavage is shown in **Supplementary Dataset 1**.

Crystallization and Data Collection

Crystallization trials were performed using the sitting drop vapor diffusion method at 16°C. Crystallization drops were dispensed using a Mosquito nanoliter high-throughput robot (TTP Labtech Ltd, United Kingdom) and consisted of 200 nL protein solution and 100 nL reservoir solution. Crystal formation was first observed in the condition containing 0.2 M ammonium acetate, 0.1 M sodium acetate (pH 4.6) and 30% (w/v) PEG 4000 in 2 to 5 days (with an approximate size of 60 μm × 40 μm). After 7 days, crystals were observed in another condition containing 0.1 M sodium citrate (pH 5.5) and 20% (w/v) PEG 3000 (with an approximate size of 500 μm × 100 μm). Crystals were flash-cooled in liquid nitrogen after cryoprotection with 30% (w/v) ethylene glycol.

X-ray diffraction data from native crystals were collected at beamline I03 or I04-1 of the Diamond Light Source (DLS), United Kingdom. The data were processed and scaled using SCALA (Evans, 2006) within the xia2-dials package (Winter et al., 2013). In the structural refinements, 5% randomly selected reflections were set aside for calculating R_{free} as a quality monitor (Brünger, 1993).

Structure Determination and Refinement

The structure of CrtE was solved by molecular replacement (MR) using the structure of GGPPS1 large subunit from *Arabidopsis thaliana* (PDB: 5E8L) (Wang et al., 2016) as a starting model (56.4% amino acid sequence identity). MR was performed with Phaser (McCoy et al., 2007). Additional model building was done manually with cycles of WinCoot (Emsley et al., 2010) (Emsley and Cowtan, 2004) and refined in REFMARC5 with using global non-crystallographic symmetry (NCS) restraints (Murshudov et al., 2011). The statistics of the model refinement are summarized in **Table 1**. Protein molecular graphics were generated with PyMOL and electrostatic surface potential map was calculated using the APBS Electrostatics Plugin (DeLano, 2002). In the structure of Syn7002 CrtE with 2 copies in the asymmetric unit (PDB: 6SXL), we identified the following residues in chain A: 8 to 172, 180 to 241 and 263 to 300 of the predicted 302 residues; chain F: 8 to 241 and 253 to 300. In the structure of Syn7002 CrtE with 6 copies in the asymmetric unit, the identified residues of each monomer were similar, covering the range from 9 to 232 and 254 to 297.

In vitro CrtE Assays

CrtE prenyltransferase activity was determined *in vitro* as described previously (Jones et al., 2013). The incubation mixture contained, in a total volume of 200 μL, 50 mM Tris-HCl (pH 7.5), 5 mM MgCl₂, 1 mM DTT, 46 μM [1-¹⁴C] IPP (1.85 GBq mmol⁻¹), 0.5% (v/v) Triton X-100 and purified protein, and 10 μM DMAPP. Samples were incubated for 1 h at 30°C and reactions stopped by cooling on ice and the addition of 200 μL of 1 M NaCl. Samples were extracted in 2 mL of water-saturated butanol overnight at -20°C. Butanol extractable products were dephosphorylated by adding potato acid phosphatase (1 unit) (Sigma-Aldrich, United Kingdom) to 1 mL of extract and incubating in 0.2% (v/v) Triton X-100 and 4 mL of methanol in a total volume of 10 mL at 37°C overnight on a shaking platform water bath. Samples were extracted in 10 mL of 10:90 (v/v) diethyl ether/petroleum ether (boiling point 40–60°C) and the organic phase dried under nitrogen. Reaction products (geraniol, farnesol, and geranylgeraniol) were identified by thin-layer chromatography (TLC) using reversed-phase C18 F254S plates (Merck, United Kingdom), with a mobile phase of 19:1 (v/v) acetone/water. The radio-labeled products were identified by autoradiography, and alcohol standards (geraniol, farnesol, geranylgeraniol, and solanesol) (Sigma-Aldrich, United Kingdom) were visualized by exposure to iodine vapor.

Site-Directed Mutagenesis of GGPPS

Reverse PCR was used to linearize the original pHUE_CrtE plasmid with the mutagenic oligonucleotides (shown below), and the re-circulation of the plasmids was achieved by using In-fusion HD Cloning Kit (Thermo Fisher Scientific, United Kingdom). All mutations were confirmed by DNA sequencing (GENWIZ, United Kingdom). The correct constructs were subsequently transformed into *E. coli* BL21 (DE3) for protein expression. The primers designed to

TABLE 1 | Data collection and refinement statistics for the Syn7002 CrtE crystal structures (in brackets are the statistics for the high-resolution shells).

	Syn7002CrtE (six copies, PDB: 6SXN)	Syn7002CrtE (two copies, PDB: 6SXL)
Diffraction source	I04-1, DLS	I03, DLS
Wavelength (Å)	0.9159	0.9763
Temperature (K)	100	100
Detector	PILATUS 6M	PILATUS 6M
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
a, b, c (Å)	102.56, 122.97, 134.19	70.20, 89.47, 107.32
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Resolution range (Å)	68.02–2.66	68.82–2.50
R _{merge}	0.06 (0.75)	0.03 (0.45)
R _{meas}	0.08 (1.1)	0.04 (0.63)
R _{pim}	0.06 (0.75)	0.03 (0.44)
CC1/2	0.99 (0.48)	0.99 (0.77)
Average I/σ(I)	11.76 (1.63)	11.72 (1.94)
Completeness (%)	99.24 (95.35)	99.97 (100.00)
Multiplicity	2.0 (2.0)	2.0 (2.0)
Total no. of reflections	98338 (9428)	48009 (4718)
No. of unique reflections	49329 (4674)	24024 (2358)
Refinement		
R _{work} /R _{free} (%)	22.8/28.5	24.3/30.6
Root-mean-square deviations		
Bond lengths (Å)	0.008	0.009
Bond angles (°)	1.556	1.840
Ramachandran plot (%)		
Most favored regions (%)	96.00	93.47
Allowed regions (%)	3.08	6.33
Outliers (%)	0.92	0.19
Average B-factor	73.73	77.69
Protein	73.76	77.74
Ligands	n/a	114.88
Solvent	57.95	57.92

produce the desired point mutations were: M87Y, 5'-GAG ATGATCCACACCTATTCTTTAATCCATGATGATCTGCC-3' and 5'-GGTGTGGATCATCTCCAGG-3'; S88F, 5'-ATGATC CACACCATGTTTTTAATCCATGATGATCTGCCCG-3' and 5'-CATGGTGTGGATCATCTCCA-3'; M87Y/S88F, 5'-GAG ATGATCCACACCTATTTTTTAATCCATGATGATCTGCCCG -3' and 5'-GGTGTGGATCATCTCCAGG-3'; A96Y, 5'-CATGA TGATCTGCCCTATATGGACAATGACGATCTCCGT-3' and 5'-GGCAGATCATCATGGATTAAG-3'; V161M, 5'-GGGG CCGAAGGCCTCATGGGTGGCCAAGTGGTGGAT-3' and 5'-GAGGCCTTCGGCCCC-3'; A96Y/V161M, 5'-CATGATGAT CTGCCCTATATGGACAATGACGATCTCCGT-3', 5'-GAG GCCTTCGGCCCC -3', 5'-GGGGCCGAAGGCCTCATGG GTGGCCAAGTGGTGGAT-3' and 5'-GGGCAGATCATCA TGGATTAAG-3'; M87A, 5'-GAGATGATCCACACCGCC TCTTTAATCCATGATGATCTGCC-3' and 5'-GGTGTGGAT CATCTCCAGG-3'; K53A, 5'-CTCCTGGCTGGGGGAGCCCC GCTACGCCGATTCTTT-3' and 5'-TCCCCAGCCAGG

AGG-3'. These substituted codons are frequently used in Syn7002. The mutated positions are indicated in boldface type.

Bioinformatic Techniques

Full-length amino-acid sequences of several prenyltransferases (for detailed sequence information, see **Supplementary Dataset 2**) were obtained from the UniProt database¹ and aligned with Aliview software using the MUSCLE program² and displayed with ESPrpt³. GGPPS (CrtE) protein sequences used for phylogenetic analysis and analysis of amino-acid residue conservation were downloaded from UniProt (see **Supplementary Dataset 3**). The phylogenetic tree was constructed using the Neighbor-joining method with the PhyML 3.0⁴. The default parameters of the above software were used. The tree was visualized using Interactive Tree of Life⁵. Analysis of the conservation of GGPPS from cyanobacteria, algae and plants was performed using the ConSurf server⁶.

RESULTS

Comparison of CrtE to Other Prenyltransferases

For Syn7002, three hypothesized prenyltransferases genes have been annotated in the genome sequence in UniProt: *crtE* (SYNPCC7002_A1085), *sdsA* (SYNPCC7002_A0580), and *uppS* (SYNPCC7002_A0099), which are considered to function as GGPPS, SPPS and undecaprenyl pyrophosphate synthase (UPPS) enzymes, respectively. In **Figure 2**, the amino acid sequence of CrtE from Syn7002 is compared to GPPS, FPPS, GGPPS, HexPPS, and SPPS enzymes from various organisms. CrtE, in common with other trans-prenyltransferases, contains the two highly conserved FARM and SARM regions (Tarshis et al., 1996; Narita et al., 1999) and contains a two amino-acid insertion (Pro-95 and Ala-96) in the FARM characteristic of a Type-II GGPPS (Hemmi et al., 2003).

Expression, Purification, and Crystallization of Syn7002 CrtE

To produce CrtE in *E. coli* for structural studies, the CrtE protein was expressed as an N-terminal histidine-tagged ubiquitin fusion protein using the high-level expression vector pHUE (Catanzariti et al., 2004). Recombinant CrtE protein was over-expressed in *E. coli* and purified first using Ni-NTA agarose (**Supplementary Figure 1**) and after cleavage of the His-tagged ubiquitin-CrtE fusion by deubiquitinase (**Supplementary Figure 2**), the liberated CrtE protein was then purified by size-exclusion chromatography and analyzed by SDS-PAGE (**Supplementary Figures 3A,C**). The molecular mass of CrtE was estimated to be about 61 kDa based on the elution

¹<https://www.uniprot.org/>

²<https://www.ebi.ac.uk/Tools/msa/muscle/>

³<http://esprpt.ibcp.fr/ESPrpt/ESPrpt/>

⁴<http://www.atgc-montpellier.fr/phyml/>

⁵<https://itol.embl.de/>

⁶<https://consurf.tau.ac.il/>

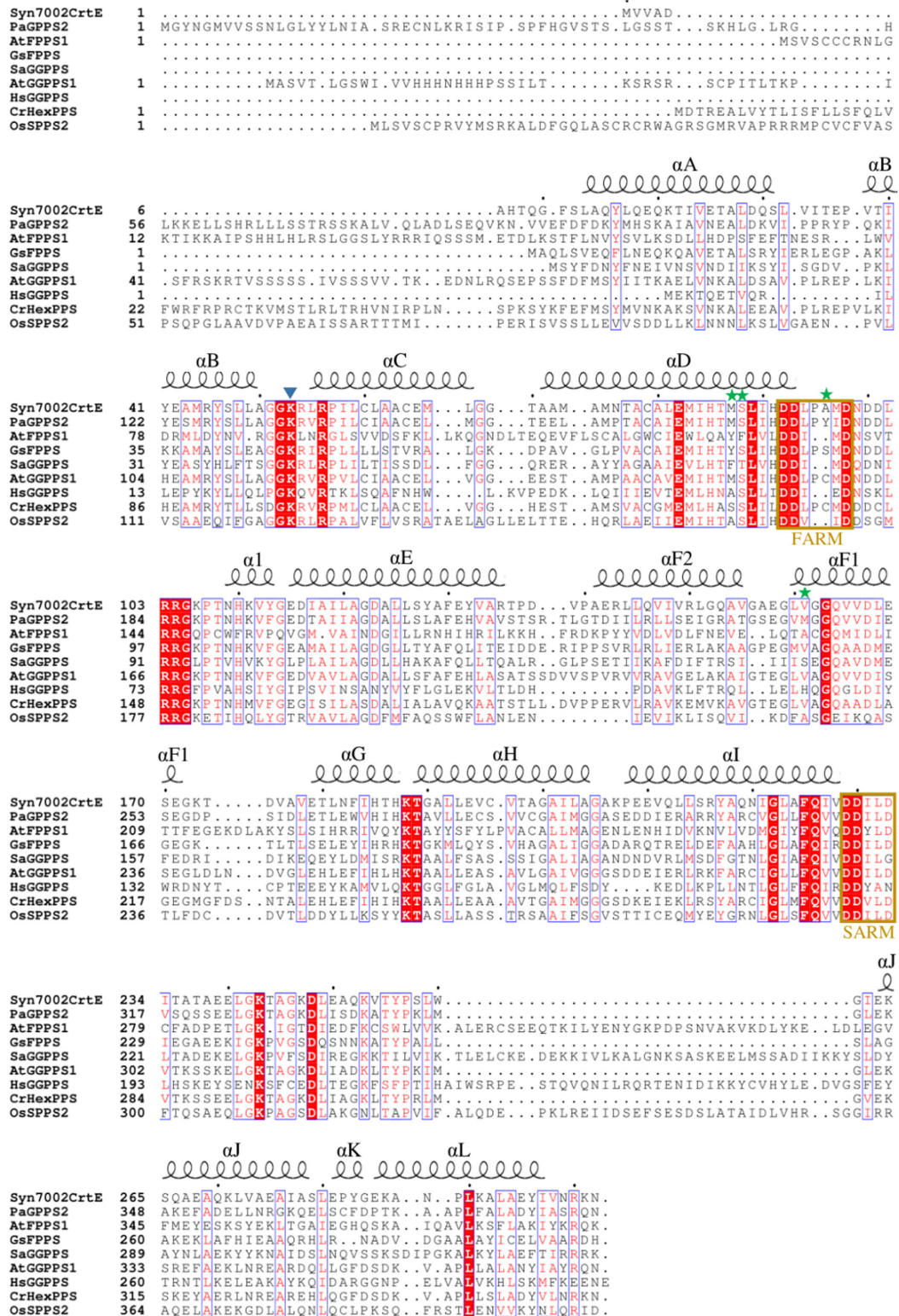
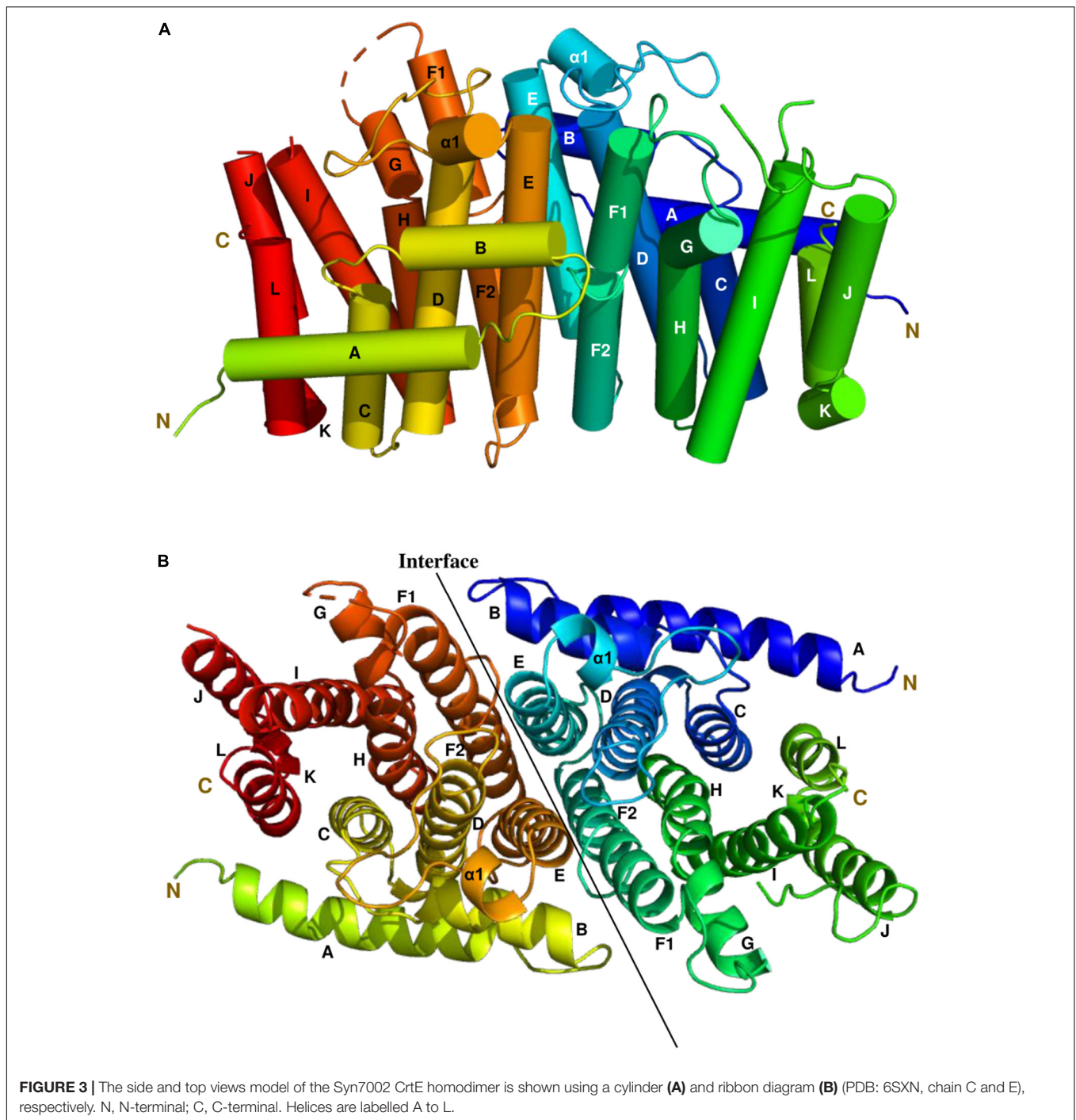


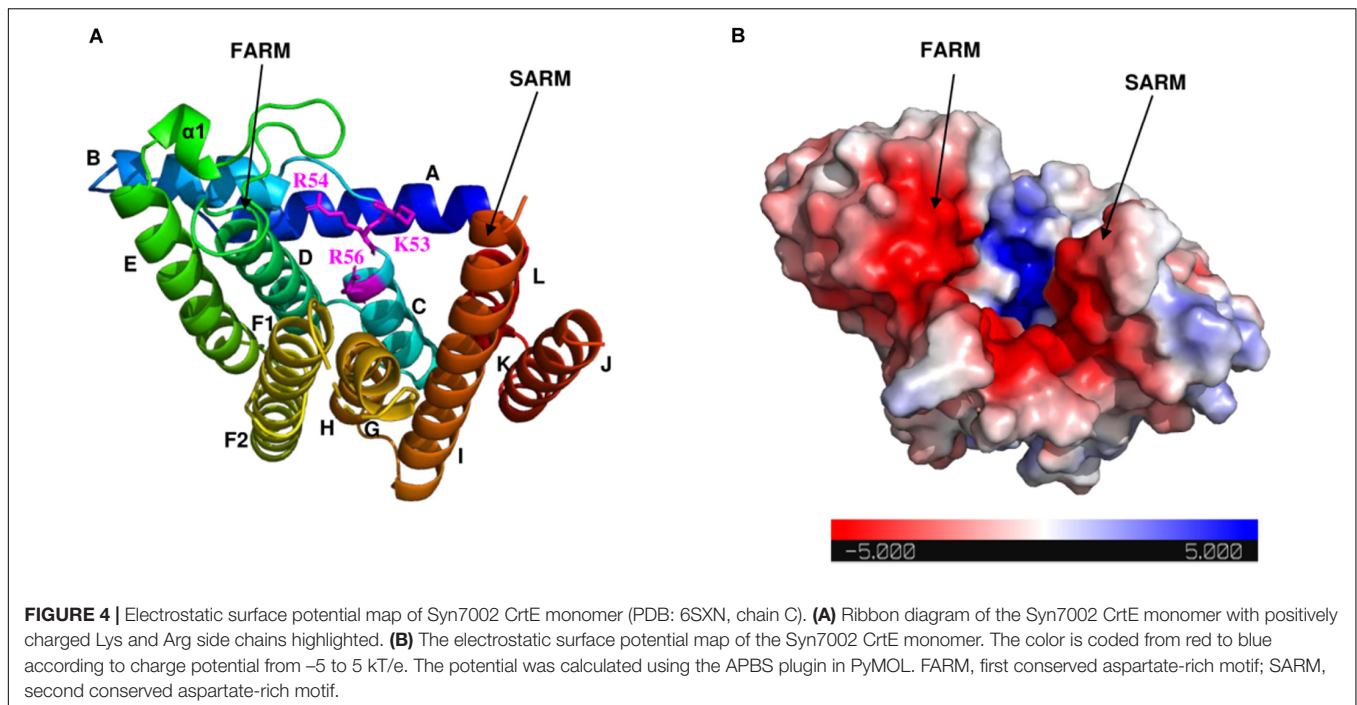
FIGURE 2 | Alignment of amino acid sequences of Syn7002, *Synechococcus* sp. PCC 7002 CrtE; Pa, *Picea abies* GPPS; At, *Arabidopsis thaliana* FPPS1 and GGPPS1; Gs, *Geobacillus stearothermophilus* FPPS; Sa, *Sulfolobus acidocaldarius* GGPPS; Hs, *Homo sapiens* GGPPS; Cr, *Capsella rubella* HexPPS; and Os, *Oryza sativa* SPPS2. Strictly conserved residues are highlighted by using red background and conservatively substituted residues are boxed. The secondary structure α -helices of Syn7002 CrtE are shown above the aligned sequences. The blue triangle indicates the Lys residue predicted to bind to IPP; the green stars indicate the amino acid residues mutated in this study.



volume (Supplementary Figure 3B), close to the predicted mass of 65 kDa for a dimer.

We obtained CrtE crystals in two different crystallization conditions (Supplementary Figures 3D,E). The crystal formed in 0.2 M ammonium acetate, 0.1 M sodium acetate (pH 4.6) and 30% (w/v) PEG 4000 had 6 copies in the asymmetric unit, and the crystal generated in 0.1 M sodium citrate (pH 5.5) and 20% (w/v) PEG 3000 contained 2 copies in the asymmetric unit. In both cases CrtE crystals had the same

space group - $P2_12_12_1$. Among the structurally characterized prenyltransferase enzymes, the predicted CrtE from Syn7002 shows 56.3% sequence identity with the large subunit of *A. thaliana* GGPPS1 (PDB: 5E8L). This was successfully used as a starting model to determine the structure of CrtE by molecular replacement. In the present study, we determined the structures of Syn7002 CrtE with 6 copies in the asymmetric unit in its apo form to a resolution of 2.66 Å (PDB: 6SXN) and with 2 copies in the asymmetric unit in its apo form



to a resolution of 2.50 \AA (PDB: 6SXL) (Table 1). Structure-based sequence alignments between the CrtE homodimer in PDB:6SXN (chain C and E) and PDB:6SXL (chain A and B) showed a root mean square difference (RMSD) of 0.471 \AA over 520 residues (2531 atoms superposed) performed using PyMOL (Supplementary Figure 4). The three CrtE dimers (CE, BD, and AF) in PDB: 6SXN have corresponding RMSD of 0.256 \AA (CE and BD, 2791 atoms superposed), 0.278 \AA (CE and AF, 2827 atoms superposed) and 0.413 \AA (AF and BD, 2953 atoms superposed), respectively. The RMSD between the CrtE monomer in PDB: 6SXN (chain C) and PDB:6SXL (chain F) is 0.420 \AA over 258 residues (1825 atoms superposed).

Overall Structure of Syn7002 CrtE

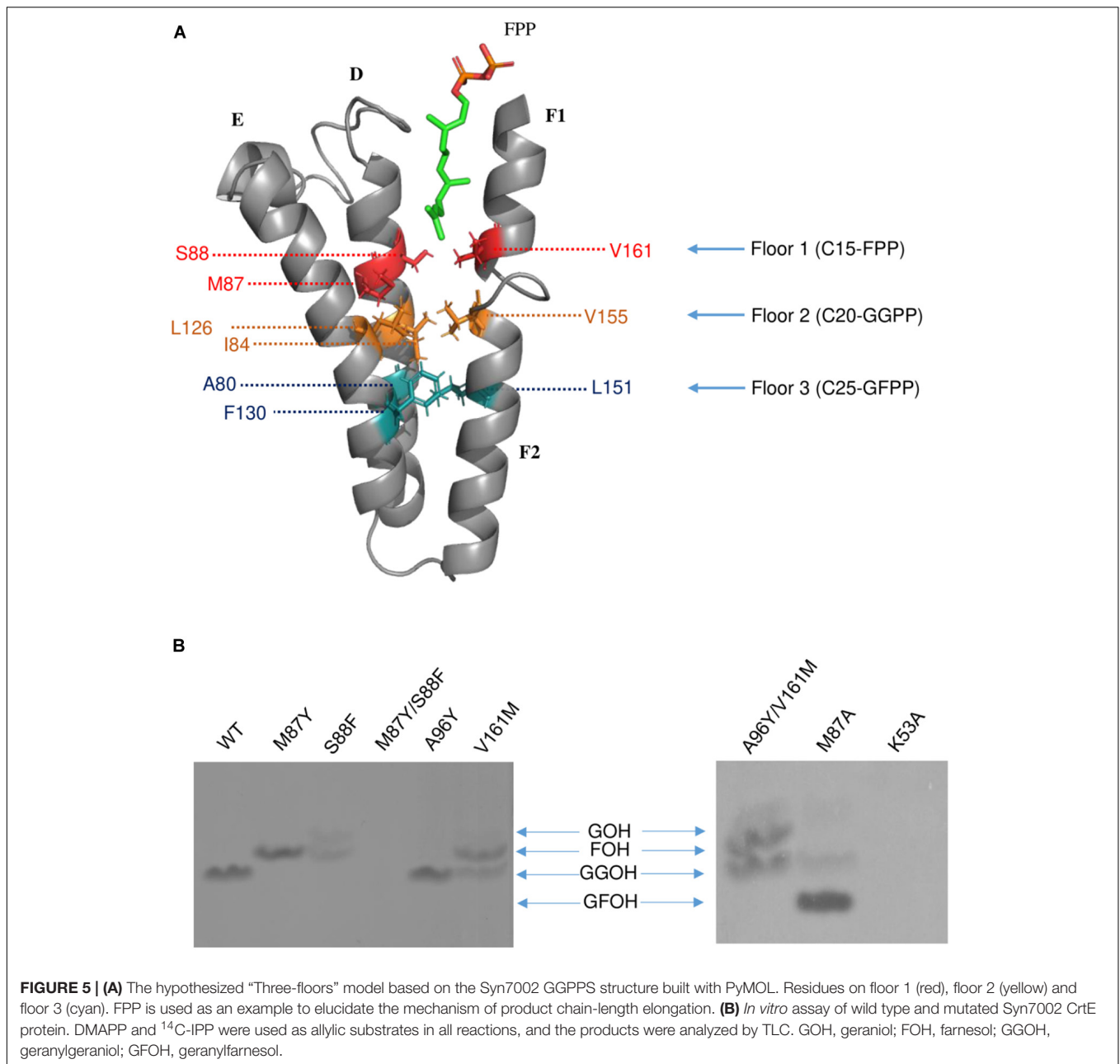
The most complete structure was obtained with PDB: 6SXN chain C, in which all 302 residues apart from the first eight N-terminal residues, twenty-two residues from 233 to 254 and the last five C-terminal residues could be identified in the electron density map. Each CrtE monomer is composed of 14 α -helices (A to L, and $\alpha 1$) joined by connecting loops (Figure 3). Helix F is broken in the middle to give helices F1 and F2, respectively. A short helix $\alpha 1$ (Asn-109-Tyr-113) was observed between helix D and E. A similar structure has been reported for *M. piperita* GGPPS and *Sinapis alba* GGPPS, and is thought to be involved in the protein conformational changes needed for substrate binding and product release (Kloer et al., 2006; Chang et al., 2010). The two aspartate-rich motifs, FARM and SARM, are located in helix D and H, respectively, forming a large catalytic cavity for allyl substrate binding (Tarshis et al., 1996; Guo et al., 2004). α -helices (D, E, F1, and F2) form a large tunnel-shaped pocket which is hypothesized to be involved in prenyl chain elongation according to previous structural studies (Wallrapp et al., 2013;

Wang et al., 2016). The inner surface of the pocket is filled with hydrophobic amino acids (Ala-80, Ile-84, Met-87, Leu-126, Phe-130, Leu-151, Val-155, and Leu-160) which, based on previous studies on prenyltransferase enzymes, are predicted to determine the chain length of the final product (Tarshis et al., 1996). In addition, residues in helices E, F1 and F2 at the homodimer interface are non-polar. Notably, the side chain of Phe-130 in helix E is involved in a π - π interaction with its counterpart in the other subunit (Supplementary Figure 5), which also has been found in *A. thaliana* GFPPS and *Sulfolobus solfataricus* HexPPS (Sun et al., 2005; Wang et al., 2016).

As shown in the electrostatic surface potential map (Figure 4), there are a number of positively charged amino acid residues (Lys-53, Arg-54, and Arg-56) located at the bottom of the catalytic cavity, which are proposed to bind the IPP substrate (Wang et al., 2016). The negatively charged FARM and SARM regions sit atop of the cavity flanking the positively charged bottom. Binding sites for Mg^{2+} and DMAPP could not be obtained in this study, but have been reported in previous structural studies of GGPPS from *S. cerevisiae*, *Plasmodium falciparum*, and *Corynebacterium glutamicum* (Guo et al., 2007; No et al., 2012; Wallrapp et al., 2013).

Regulation of Product Chain Length

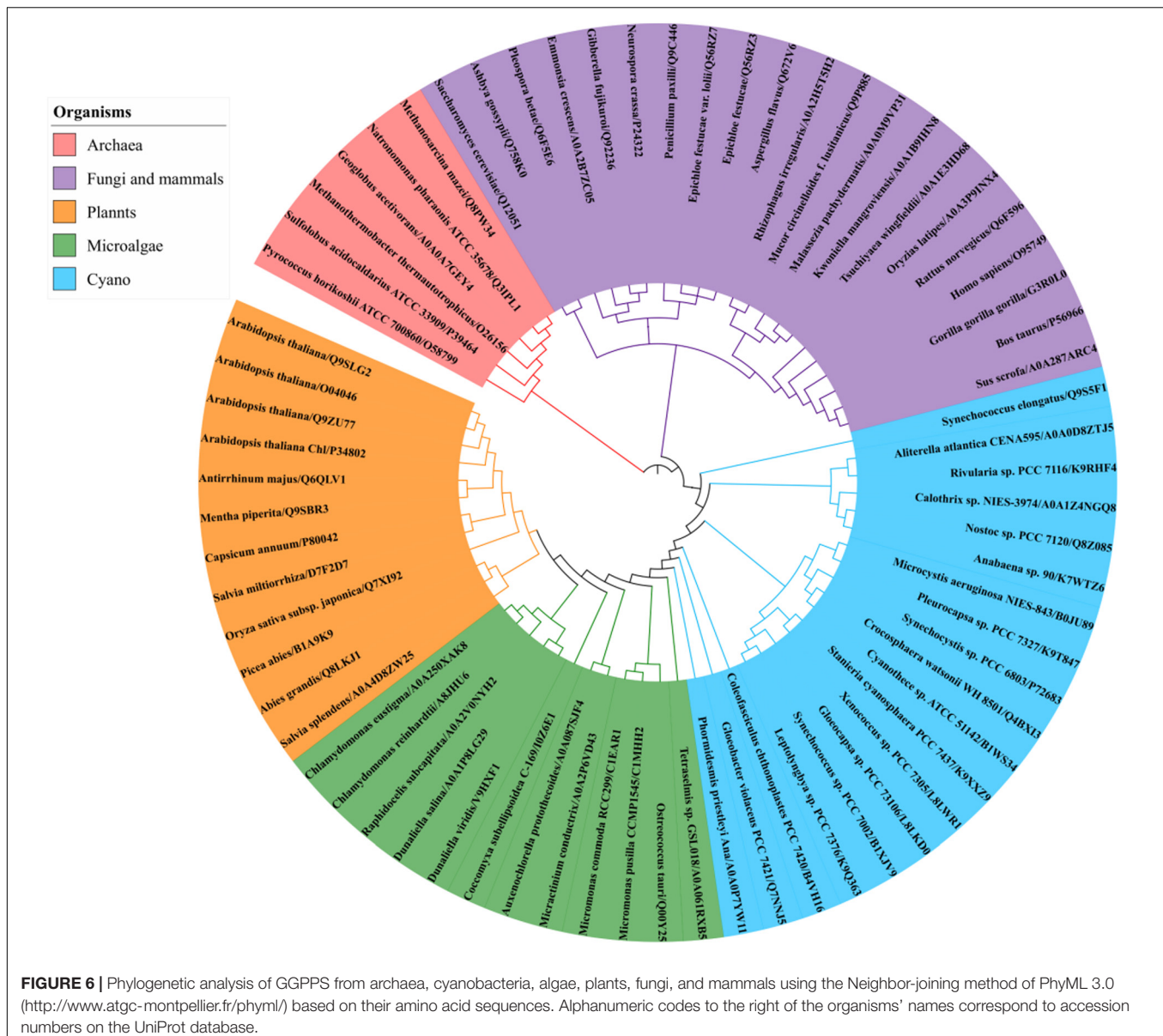
Previous studies have demonstrated that the side chains of residues in the hydrophobic pocket (formed by helices D, E, F1, and F2) could form a “floor” that blocks product elongation and so determine the product chain length (Sun et al., 2005; Han et al., 2015). Based on the structure of wild type Syn7002 CrtE, an equivalent “floor” model was built, and amino-acid residues predicted to determine product chain length were identified (Figure 5A). Site-directed mutagenesis of CrtE was



performed to test the first “floor” of this model; therefore, Met-87 and Ser-88 on helix D, and Val-161 on helix F1, were replaced by residues with larger (Phe and Tyr) or smaller (Ala) side chains.

As pure CrtE WT and CrtE WT fusion gave the same result in the TLC chromatography profile (**Supplementary Figure 6**), the variant His-tagged CrtE fusion proteins were directly used for their enzyme assays after purification using immobilized Ni-affinity chromatography and dialysis. Enzyme assays were conducted with purified wild-type and variant fusion proteins (**Supplementary Figure 7**) using $[1-^{14}\text{C}]$ IPP and DMAPP as substrates as described in materials and methods. The radio-labeled products were dephosphorylated, and the generated

alcohols were observed on TLC (**Figure 5B**). Under the same reaction conditions, the final product of wild type CrtE was C20-GGPP. Variants M87Y and S88F both shifted the final product from GGPP to C15-FPP, but S88F produced less FPP and some C10-GPP compared to the products of M87Y. However, no product was detected in the mutated enzyme (M87Y/S88F) in which both residues were changed, possibly because the enzyme kinetics were dramatically affected, or the final product was volatile. When Met-87 was mutated to Ala, somewhat surprisingly, the longer geranylfarnesyl pyrophosphate molecule (GFPP, C25) was generated as a major product, with traces of GGPP. The second residue before the conserved GQ motif (Val-161) on helix D, which is also hypothesized to be involved in the



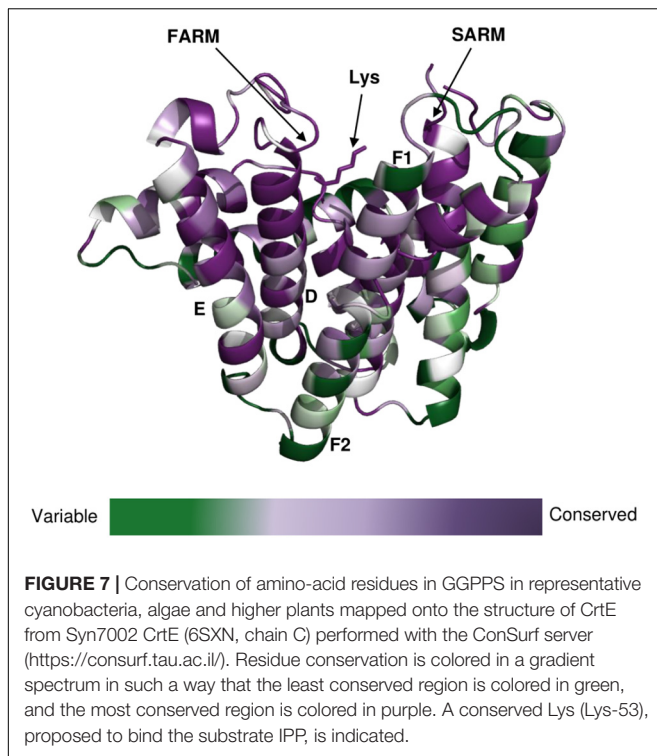
first floor formation (Hemmi et al., 2003), was mutated to Met, and found to produce similar amounts of FPP and GGPP.

Type-II Syn7002 CrtE has two extra residues (Pro-95 and Ala-96) inserted within the FARM region. These two residues are located in the loop between helices D and E, and spatially on the top of the elongation pocket in the structure. According to the sequence alignment in **Figure 1**, Pro-95 in the FARM is highly conserved in Type-II prenyltransferases, but the second additional residue shows more variability. For instance, PaGGPPS has a bulky Tyr residue at the second position, whereas the other Type-II prenyltransferases (Syn7002 CrtE, GsFPPS, AtGGPPS, and CrHexPPS) have a smaller residue such as Cys, Ser and Ala. To determine whether a bulky residue at this position could interrupt the elongation reaction, to give rise to GPP, we mutated CrtE Ala-96 to Tyr, but GGPP was still formed. These data indicate that the size of the second extra amino acid in the

FARM does not critically affect product chain-length elongation. We also tried to mimic PaGGPPS by making a version of CrtE with A96Y/V161M, so that the mutated enzyme has the same first “floor” and the same two residue insertion as that found in PaGGPPS. But like the V161M single variant, similar amounts of FPP and GGPP were produced and there was no evidence for formation of GPP. Mutation of the predicted IPP binding site (Lys-53) was achieved by replacing it with Ala but no products were detected.

Phylogenetic Relationship of Photosynthetic Organisms GGPPSs

A BLAST search of the UniProt online protein database of archaea, cyanobacteria, algae, higher plants, fungi and mammal sequences (**Supplementary Dataset 3**) using Syn7002 CrtE



(GGPPS) was conducted to study the evolutionary relationships. A phylogenetic analysis (**Figure 6**), based on the degree of sequence similarity, revealed that all the GGPPSs have evolved from a common ancestor and form three distinct clades (Type-I GGPPS in archaeobacteria, Type-II GGPPS in photosynthetic organisms, and Type-III GGPPS found in fungi and mammals). GGPPS from photosynthetic organisms (Type-II) and GGPPS from fungi and mammals (Type-III) belong to different clades but may share a common ancestor.

The clade of photosynthetic organisms was divided into three clusters (cyanobacteria, algae, and plants). Analysis of the sequence conservation of Type-II GGPPSs was conducted based on sequence alignments and the determined Syn7002 GGPPS structure (PDB: 6SXN) using the ConSurf server (**Figure 7**). The essential functional domains such as short helix $\alpha 1$, FARM, SARM, the ligand-binding region and the residues in the product elongation tunnel are highly conserved, which indicates that GGPPSs in photosynthetic organisms have a very close evolutionary relationship and are likely to display similar mechanisms for catalysis and for controlling product chain length.

DISCUSSION

In this study, we have demonstrated that the CrtE enzyme encoded by the model cyanobacterium Syn7002 codes for a Type-II GGPPS. A phylogenetic analysis indicates that Type-II GGPPS is widely found in photosynthetic organisms. Cyanobacteria and algae encode one GGPPS (CrtE) and higher plants usually have multiple GGPPS enzymes, for example

A. thaliana has 10 GGPPSs (Ruiz-Sola et al., 2016). Uniquely, the thermophilic cyanobacterium *T. elongatus* possesses a FPPS in addition to a GGPPS (Ohto et al., 1999). Syn7002 GGPPS is a homodimer and its three-dimensional structure is most similar to the large subunit of heterodimeric *A. thaliana* GGPPS1. Heterodimeric and heterotetrameric GGPPSs are found in higher plants, with the large subunit involved in catalysis and the small subunit in product release (Chang et al., 2010; Wang et al., 2016).

Syn7002 GGPPS contains the two highly conserved aspartate-rich motifs (FARM and SARM), which are important for substrate binding and catalysis in prenyltransferases (Song and Poulter, 1994; Koyama et al., 1996). In most trans-prenyltransferase structures, dimer interactions have been observed between helices E and F (F1 and F2) (Guo et al., 2004; Noike et al., 2008; Han et al., 2015; Wang et al., 2016). Van der Waals interactions and stacking (π - π) interactions between the hydrophobic residues on these two helices contribute to dimerization (Sun et al., 2005). Previous studies have demonstrated that a positively charged residue Lys at the bottom of the catalytic cavity is responsible for IPP binding: Lys-36 in *S. cerevisiae* GGPPS (PDB: 2E8U), Lys-60 in *Plasmodium vivax* GGPPS (3LDW), and Lys-47 in *Trypanosoma brucei* FPPS (PDB: 3DYF) (Guo et al., 2007; Zhang et al., 2009; Artz et al., 2011). In addition, Koyama et al. (1996) replaced Lys-47 with Ile in *B. stearothermophilus* FPPS and found the affinity for substrate IPP was decreased. In Syn7002 GGPPS, replacement of Lys-53 by Ala led to no detectable formation of product. Arg-56, which is highly conserved in prenyltransferases, is also ideally placed to interact with IPP, as also observed for *S. cerevisiae* GGPPS (Guo et al., 2007).

A “three-floor” model has been proposed to explain how the length of the final product is controlled in prenyltransferases (Wang et al., 2016). The size of residues at each floor determines the length of the final product. For example, “Floor one” determines if the final product is C15-FPP, “Floor two” if it is C20-GGPP and “Floor three” if it is C25-GFPP (**Figure 5A**). From these models and corresponding site-directed mutagenesis experiments, it is generally believed that the size of the residues located four and five positions before the FARM determines the chain length of the product, with a large side chain forming a “floor” that blocks further product elongation (Ohnuma et al., 1996; Narita et al., 1999; Guo et al., 2004; Sun et al., 2005; Szkopi and Danuta, 2005; Ling et al., 2007; Han et al., 2015). According to our structural models, Syn7002 CrtE has two small residues and one medium sized Met residue on the first floor, and its second floor is formed from Ile-84, Leu-126 and Val-155 (**Figure 5A**). Surprisingly, replacement of Met-87 on the first floor with Ala generated a longer product C25-GFPP, suggesting that Met-87 has a role in forming the first floor but its side chain also indirectly impacts the second floor. An identical result was also observed previously for the equivalent Met to Ala mutation in AtGGPPS1 (Wang et al., 2016) and, indeed, the predicted first floor residues are exactly the same in Syn7002 CrtE and AtGGPPS1 and the second floor residues are very similar with Ile, Leu, and Val found in Syn7002 CrtE and Ile, Leu, and Ile in AtGGPPS1.

The residues before another conserved motif G(Q/E) on helix F1 are also considered to regulate chain termination (Hemmi et al., 2003; Kawasaki et al., 2003). Our sequence alignment (Figure 2) demonstrated that the size of the second residue before the G(Q/E) motif in the short-chain prenyltransferases (like PaGPPS and HsGGPPS) is larger than that in the medium/long-chain prenyltransferases. In addition, our Syn7002 GGPPS structure showed that the second residue (Val-161) before the conserved GQ motif is more likely to generate a “floor” with Met-87 and Ser-88 according to its spatial position. Mutating Val-161 to Met partially blocked the reaction and generated more FPP and less GGPP compared to the products of wild type GGPPS. A similar mutagenesis study was performed with *S. acidocaldarius* HexPPS: a point mutation at this position (S140H, S140F, and S140Y) resulted in the shortening of the final products (Hemmi et al., 2003). Collectively, these results indicate that the second residue before the conserved G(Q/E) motif is also involved in determining product chain length. The third amino acid (L151) before the G(Q/E) motif of *A. thaliana* GGPPS1 has also been reported to be important for “floor” formation, with mutation of this residue (L151F) leading to the formation of shorter products (GPP and FPP) and less GGPP (Wang et al., 2016).

Cyanobacteria have been considered as green microbial cell factories for producing high value-added products. However, the yield of short chain terpenoids is limited by the small pool sizes of their precursors GPP and FPP (Kiyota et al., 2014), as the major product of CrtE is GGPP. We have shown here that mutation of CrtE can alter product specificity. In principle, it is possible to use this knowledge to make cyanobacterial strains with larger pools of FPP or GFPP which could be useful platforms for the production of sesquiterpenoids, triterpenoids and sesterterpenoids.

DATA AVAILABILITY STATEMENT

Two protein structures were elucidated and submitted to the RCSB Protein Data Bank: the structures of Syn7002 CrtE with

six copies in the asymmetric unit in its apo form (PDB accession: 6SXN) and with two copies in the asymmetric unit in its apo form (PDB accession: 6SXL).

AUTHOR CONTRIBUTIONS

YF performed the experimental work. RM collected the X-ray diffraction data and solved the structures. PF contributed reagents, materials, and analysis tools for TLC analysis. YF, RM, PF, and PN designed the experiments. All authors listed have contributed to data interpretation, preparation, and writing of the manuscript.

FUNDING

This work was provided by the European Union’s Horizon 2020 Research and Innovation Program project PHOTOFUEL under grant agreement number 640720. YF was supported by the China Scholarship Council.

ACKNOWLEDGMENTS

We are extremely grateful to Chris Gerrish (Royal Holloway, University of London) for help with the TLC assays and to James W. Murray and Ciaran McFarlane (Imperial College London) for help with data collection and structure refinement.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2020.00589/full#supplementary-material>

REFERENCES

- Artz, J. D., Wernimont, A. K., Dunford, J. E., Schapira, M., Dong, A., Zhao, Y., et al. (2011). Molecular characterization of a novel geranylgeranyl pyrophosphate synthase from plasmodium parasites. *J. Biol. Chem.* 286, 3315–3322. doi: 10.1074/jbc.M109.027235
- Bai, C., Capell, T., Berman, J., Medina, V., Sandmann, G., Christou, P., et al. (2016). Bottlenecks in carotenoid biosynthesis and accumulation in rice endosperm are influenced by the precursor-product balance. *Plant Biotechnol. J.* 14, 195–205. doi: 10.1111/pbi.12373
- Bouvier, F., Rahier, A., and Camara, B. (2005). Biogenesis, molecular regulation and function of plant isoprenoids. *Prog. Lipid Res.* 44, 357–429. doi: 10.1016/j.plipres.2005.09.003
- Brünger, A. T. (1993). Assessment of phase accuracy by cross validation: the free R value. methods and applications. *acta crystallogr. Sect. D Biol. Crystallogr.* 49, 24–36. doi: 10.1107/s0907444992007352
- Catanzariti, A.-M., Soboleva, T. A., Jans, D. A., Board, P. G., and Baker, R. T. (2004). An efficient system for high-level expression and easy purification of authentic recombinant proteins. *Protein Sci.* 13, 1331–1339. doi: 10.1110/ps.04618904
- Chang, K. M., Chen, S. H., Kuo, C. J., Chang, C. K., Guo, R. T., Yang, J. M., et al. (2012). Roles of amino acids in the *Escherichia coli* octaprenyl diphosphate synthase active site probed by structure-guided site-directed mutagenesis. *Biochemistry* 51, 3412–3419. doi: 10.1021/bi30069j
- Chang, T. H., Guo, R. T., Ko, T. P., Wang, A. H. J., and Liang, P. H. (2006). Crystal structure of type-III geranylgeranyl pyrophosphate synthase from *Saccharomyces cerevisiae* and the mechanism of product chain length determination. *J. Biol. Chem.* 281, 14991–15000. doi: 10.1074/jbc.M512886200
- Chang, T.-H., Hsieh, F.-L., Ko, T.-P., Teng, K.-H., Liang, P.-H., and Wang, A. H.-J. (2010). Structure of a heterotetrameric geranyl pyrophosphate synthase from Mint (*Mentha piperita*) reveals intersubunit regulation. *Plant Cell* 22, 454–467. doi: 10.1105/tpc.109.071738
- Chen, A., Dale Poulter, C., and Kroon, P. A. (1994). Isoprenyl diphosphate synthases: protein sequence comparisons, a phylogenetic tree, and predictions of secondary structure. *Protein Sci.* 3, 600–607. doi: 10.1002/pro.5560030408
- DeLano, W. L. (2002). *PyMOL*. Available online at: <http://www.pymol.org>.

- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 60, 2126–2132. doi: 10.1107/S0907444904019158
- Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501. doi: 10.1107/S0907444910007493
- Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 62, 72–82. doi: 10.1107/S0907444905036693
- Guo, R.-T., Cao, R., Liang, P.-H., Ko, T.-P., Chang, T.-H., Hudock, M. P., et al. (2007). Bisphosphonates target multiple sites in both cis- and trans-prenyltransferases. *Proc. Natl. Acad. Sci. U.S.A.* 104, 10022–10027. doi: 10.1073/pnas.0702254104
- Guo, R. T., Kuo, C. J., Chou, C. C., Ko, T. P., Shr, H. L., Liang, P. H., et al. (2004). Crystal structure of octaprenyl pyrophosphate synthase from hyperthermophilic thermotoga maritima and mechanism of product chain length determination. *J. Biol. Chem.* 279, 4903–4912. doi: 10.1074/jbc.M310161200
- Han, X., Chen, C. C., Kuo, C. J., Huang, C. H., Zheng, Y., Ko, T. P., et al. (2015). Crystal structures of ligand-bound octaprenyl pyrophosphate synthase from *Escherichia coli* reveal the catalytic and chain-length determining mechanisms. *Proteins Struct. Funct. Bioinforma.* 83, 37–45. doi: 10.1002/prot.24618
- Hemmi, H., Noike, M., Nakayama, T., and Nishino, T. (2003). An alternative mechanism of product chain-length determination in type III geranylgeranyl diphosphate synthase. *Eur. J. Biochem.* 270, 2186–2194. doi: 10.1046/j.1432-1033.2003.03583.x
- Jiang, M., Stephanopoulos, G., and Pfeifer, B. A. (2012). Toward biosynthetic design and implementation of *Escherichia coli*-derived paclitaxel and other heterologous polyisoprene compounds. *Appl. Environ. Microbiol.* 78, 2497–2504. doi: 10.1128/AEM.07391-11
- Jones, M. O., Perez-Fons, L., Robertson, F. P., Bramley, P. M., and Fraser, P. D. (2013). Functional characterization of long-chain prenyl diphosphate synthases from tomato. *Biochem. J.* 449, 729–740. doi: 10.1042/bj20120988
- Kavanagh, K. L., Dunford, J. E., Bunkoczi, G., Russell, R. G. G., and Oppermann, U. (2006). The crystal structure of human geranylgeranyl pyrophosphate synthase reveals a novel hexameric arrangement and inhibitory product binding. *J. Biol. Chem.* 281, 22004–22012. doi: 10.1074/jbc.M602603200
- Kawasaki, T., Hamano, Y., Kuzuyama, T., Itoh, N., Seto, H., and Dairi, T. (2003). Interconversion of the product specificity of type I eubacterial farnesyl diphosphate synthase and geranylgeranyl diphosphate synthase through one amino acid substitution. *J. Biochem.* 133, 83–91. doi: 10.1093/jb/mvg002
- Kiyota, H., Okuda, Y., Ito, M., Hirai, M. Y., and Ikeuchi, M. (2014). Engineering of cyanobacteria for the photosynthetic production of limonene from CO₂. *J. Biotechnol.* 185, 1–7. doi: 10.1016/j.jbiotec.2014.05.025
- Kloer, D. P., Welsch, R., Beyer, P., and Schulz, G. E. (2006). Structure and reaction geometry of geranylgeranyl diphosphate synthase from *Sinapis alba*. *Biochemistry* 45, 15197–15204. doi: 10.1021/bi061572k
- Koyama, T., Tajima, M., Sano, H., Doi, T., Koike-Takeshita, A., Obata, S., et al. (1996). Identification of significant residues in the substrate binding site of *Bacillus stearothermophilus* farnesyl diphosphate synthase. *Biochemistry* 35, 9533–9538. doi: 10.1021/bi960137v
- Liang, C., Zhao, F., Wei, W., Wen, Z., and Qin, S. (2006). Carotenoid biosynthesis in cyanobacteria: Structural and evolutionary scenarios based on comparative genomics. *Int. J. Biol. Sci.* 2, 197–207. doi: 10.7150/ijbs.2.197
- Liang, P. H. (2009). Reaction kinetics, catalytic mechanisms, conformational changes, and inhibitor design for prenyltransferases. *Biochemistry* 48, 6562–6570. doi: 10.1021/bi900371p
- Ling, Y., Li, Z. H., Miranda, K., Oldfield, E., and Moreno, S. N. J. (2007). The farnesyl-diphosphate/geranylgeranyl-diphosphate synthase of *Toxoplasma gondii* is a bifunctional enzyme and a molecular target of bisphosphonates. *J. Biol. Chem.* 282, 30804–30816. doi: 10.1074/jbc.M703178200
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* 40, 658–674. doi: 10.1107/S0021889807021206
- Michoux, F., Boehm, M., Bialek, W., Takasaka, K., Maghlaoui, K., Barber, J., et al. (2014). Crystal structure of CyanoQ from the thermophilic cyanobacterium *Thermosynechococcus elongatus* and detection in isolated photosystem II complexes. *Photosynth. Res.* 122, 57–67. doi: 10.1007/s11120-014-0010-z
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., et al. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 67, 355–367. doi: 10.1107/S0907444911001314
- Narita, K., Ohnuma, S., and Nishino, T. (1999). Protein design of geranyl diphosphate. *J. Biochem.* 126, 566–571.
- No, J. H., De Macedo Dossin, F., Zhang, Y., Liu, Y. L., Zhu, W., Feng, X., et al. (2012). Lipophilic analogs of zoledronate and risedronate inhibit Plasmodium geranylgeranyl diphosphate synthase (GGPPS) and exhibit potent antimalarial activity. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4058–4063. doi: 10.1073/pnas.1118215109
- Noike, M., Katagiri, T., Nakayama, T., Koyama, T., Nishino, T., and Hemmi, H. (2008). The product chain length determination mechanism of type II geranylgeranyl diphosphate synthase requires subunit interaction. *FEBS J.* 275, 3921–3933. doi: 10.1111/j.1742-4658.2008.06538.x
- Ogura, K., and Koyama, T. (1998). Enzymatic aspects of isoprenoid chain elongation. *Chem. Rev.* 98, 1263–1276. doi: 10.1021/cr9600464
- Ohnuma, N. T., Nishinof, T., Narita, K., Ishida, C., Takeuchi, Y., Nishino, T., et al. (1996). A role of the amino acid residue located on the fifth position before the first aspartate-rich motif of farnesyl diphosphate synthase on determination of the final product. *J. Biol. Chem.* 271, 30748–30754. doi: 10.1074/jbc.271.48.30748
- Ohto, C., Ishida, C., Nakane, H., Muramatsu, M., Nishino, T., and Obata, S. (1999). A thermophilic cyanobacterium *Synechococcus elongatus* has three different class I prenyltransferase genes. *Plant Mol. Biol.* 40, 307–321. doi: 10.1023/A:1006295705142
- Petrova, T. E., Boyko, K. M., Nikolaeva, A. Y., Stekhanova, T. N., Gruzdev, E. V., Mardanov, A. V., et al. (2018). Structural characterization of geranylgeranyl pyrophosphate synthase GACE1337 from the hyperthermophilic archaeon *Geoglobus acetivorans*. *Extremophiles* 22, 877–888. doi: 10.1007/s00792-018-1044-5
- Rabinovitch-Deere, C. A., Oliver, J. W. K., Rodriguez, G. M., and Atsumi, S. (2013). Synthetic biology and metabolic engineering approaches to produce biofuels. *Chem. Rev.* 113, 4611–4632. doi: 10.1021/cr300361t
- Ruiz-Sola, M. Á., Barja, M. V., Rodríguez-Concepción, M., Coman, D., Beck, G., Colinas, M., et al. (2016). *Arabidopsis* geranylgeranyl diphosphate synthase 11 Is a hub isozyme required for the production of most photosynthesis-related isoprenoids. *New Phytol.* 209, 252–264. doi: 10.1111/nph.13580
- Shimizu, N., Koyama, T., and Ogura, K. (1998). molecular cloning, expression, and purification of undecaprenyl diphosphate synthase. *J. Biol. Chem.* 273, 19476–19481. doi: 10.1074/jbc.273.31.19476
- Song, L., and Poulter, A. C. D. (1994). Yeast farnesyl-diphosphate synthase: Site-directed mutagenesis of residues in highly conserved prenyltransferase domains I and II (sitedirected mutants/kinetic cntn/immunoafflity chromatography). *Biochemistry* 91, 3044–3048. doi: 10.1073/pnas.91.8.3044
- Sun, H.-Y., Guo, R.-T., Liang, P.-H., Wang, A. H.-J., Ko, T.-P., Kuo, C.-J., et al. (2005). Homodimeric hexaprenyl pyrophosphate synthase from the thermoacidophilic crenarchaeon *Sulfolobus solfataricus* displays asymmetric subunit structures. *J. Bacteriol.* 187, 8137–8148. doi: 10.1128/JB.187.23.8137-8148.2005
- Szkopi, A., and Danuta, P. (2005). Farnesyl diphosphate synthase; regulation of product. *Acta Biochim. Pol.* 52, 45–55.
- Tarshis, L. C., Proteau, P. J., Kellogg, B. A., Sacchettini, J. C., and Poulter, C. D. (1996). Regulation of product chain length by isoprenyl diphosphate synthases. *Proc. Natl. Acad. Sci. U.S.A.* 93, 15018–15023. doi: 10.1073/pnas.93.26.15018
- Vickers, C. E., Bongers, M., Liu, Q., Delatte, T., and Bouwmeester, H. (2014). Metabolic engineering of volatile isoprenoids in plants and microbes. *Plant, Cell Environ.* 37, 1753–1775. doi: 10.1111/pce.12316
- Wallrap, F. H., Pan, J.-J., Ramamoorthy, G., Almonacid, D. E., Hillerich, B. S., Seidel, R., et al. (2013). Prediction of function for the polyprenyl transferase subgroup in the isoprenoid synthase superfamily. *Proc. Natl. Acad. Sci. U.S.A.* 110:E1196–E1202. doi: 10.1073/pnas.1300632110
- Wang, C., Chen, Q., Fan, D., Li, J., Wang, G., and Zhang, P. (2016). Structural analyses of short-chain prenyltransferases identify an evolutionarily conserved GGPPS clade in brassicaceae plants. *Mol. Plant* 9, 195–204. doi: 10.1016/j.molp.2015.10.010

- Wang, K., and Ohnuma, S. (1999). Chain-length determination mechanism of isoprenyl diphosphate synthases and implications for molecular evolution. *Trends Biochem. Sci.* 24, 445–451. doi: 10.1016/S0968-0004(99)01464-4
- Winter, G., Lobley, C. M. C., and Prince, S. M. (2013). Decision making in xia2. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 69, 1260–1273. doi: 10.1107/S0907444913015308
- Yen, H. C., and Marrs, B. (1976). Map of genes for carotenoid and bacteriochlorophyll biosynthesis in *Rhodospseudomonas capsulata*. *J. Bacteriol.* 126, 619–629.
- Zhang, Y., Cao, R., Yin, F., Hudock, M. P., Guo, R., Mukherjee, S., et al. (2009). Lipophilic bisphosphonates as dual farnesyl / geranylgeranyl diphosphate synthase inhibitors: an X-ray and NMR investigation. *J. Am. Chem. Soc.* 131, 5153–5162. doi: 10.1021/ja808285e
- Zhou, F., Wang, C.-Y., Gutensohn, M., Jiang, L., Zhang, P., Zhang, D., et al. (2017). A recruiting protein of geranylgeranyl diphosphate synthase controls metabolic flux toward chlorophyll biosynthesis in rice. *Proc. Natl. Acad. Sci. U.S.A.* 114, 6866–6871. doi: 10.1073/pnas.1705689114
- Zor, T., and Selinger, Z. (1996). Linearization of the Bradford protein assay increases its sensitivity: theoretical and experimental studies. *Anal. Biochem.* 2, 302–308. doi: 10.1006/abio.1996.0171
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Copyright © 2020 Feng, Morgan, Fraser, Hellgardt and Nixon. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.