THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 21, pp. 14991–15000, May 26, 2006


# Crystal Structure of Type-III Geranylgeranyl Pyrophosphate Synthase from *Saccharomyces cerevisiae* and the Mechanism of Product Chain Length Determination*

Received for publication, December 2, 2005, and in revised form, February 24, 2006 Published, JBC Papers in Press, March 22, 2006, DOI 10.1074/jbc.M512886200

**Tao-Hsin Chang[‡], Rey-Ting Guo[§], Tzu-Ping Ko[¶], Andrew H.-J. Wang[‡§¶1], and Po-Huang Liang[‡§¶2]**

*From the [‡]Institute of Biochemical Sciences, National Taiwan University, Taipei 106, Taiwan, the [§]Taiwan International Graduate Program, Academia Sinica, Taipei 115, Taiwan, and the [¶]Institute of Biological Chemistry, Academia Sinica, Taipei 115, Taiwan*

**Geranylgeranyl pyrophosphate synthase (GGPPs) catalyzes a condensation reaction of farnesyl pyrophosphate with isopentenyl pyrophosphate to generate $C_{20}$ geranylgeranyl pyrophosphate, which is a precursor for carotenoids, chlorophylls, geranylgeranylated proteins, and archaeal ether-linked lipid. For short-chain *trans*-prenyltransferases that synthesize $C_{10}$–$C_{25}$ products, bulky amino acid residues generally occupy the fourth or fifth position upstream from the first DD*XX*D motif to block further elongation of the final products. However, the short-chain type-III GGPPs in eukaryotes lack any large amino acid at these positions. In this study, the first structure of type-III GGPPs from *Saccharomyces cerevisiae* has been determined to 1.98 Å resolution. The structure is composed entirely of 15 α-helices joined by connecting loops and is arranged with α-helices around a large central cavity. Distinct from other known structures of *trans*-prenyltransferases, the N-terminal 17 amino acids (9-amino acid helix A and the following loop) of this GGPPs protrude from the helix core into the other subunit and contribute to the tight dimer formation. Deletion of the first 9 or 17 amino acids caused the dissociation of dimer into monomer, and the Δ(1–17) mutant showed abolished enzyme activity. In each subunit, an elongated hydrophobic crevice surrounded by D, F, G, H, and I α-helices contains two DD*XX*D motifs at the top for substrate binding with one $Mg^{2+}$ coordinated by $Asp^{75}$, $Asp^{79}$, and four water molecules. It is sealed at the bottom with three large residues of $Tyr^{107}$, $Phe^{108}$, and $His^{139}$. Compared with the major product $C_{30}$ synthesized by mutant H139A, the products generated by mutant Y107A and F108A are predominantly $C_{40}$ and $C_{30}$, respectively, suggesting the most important role of $Tyr^{107}$ in determining the product chain length.**

Five-carbon isopentenyl pyrophosphate (IPP)[3] serves as the building block for the isoprenoid natural products (1). IPP is first converted to its isomer dimethylallyl pyrophosphate (DMAPP) by IPP:DMAPP isomerase (2). Condensation of DAMPP with 1–3 IPP results in $C_{10}$ geranyl pyrophosphate (GPP), $C_{15}$ farnesyl pyrophosphate (FPP), and $C_{20}$ geranylgeranyl pyrophosphate (GGPP) by the corresponding synthases (3). These allylic pyrophosphates, including DMAPP, GPP, FPP, and GGPP, serve as the precursors for hormones, pheromones, visual pigments, photoprotective agents, defensive agents, skeletons of lipid for membranes, and constituents of signal transduction networks (4). From the allylic substrate (FPP in most cases), larger linear polymers can be synthesized through the addition of IPP to the growing chain by a group of prenyltransferases (5). These prenyltransferases either form a *cis* or *trans* double bond during each IPP condensation (6). Two conserved aspartate-rich motifs DD*XX*D, where *X* encodes any amino acid, found in all *trans*-prenyltransferases, coordinate with $Mg^{2+}$ for substrate binding and the subsequent reaction (7–10), whereas an Asp in a P-loop of the *cis*-prenyltransferases is responsible for the similar function (11–13).

In general, *trans*-prenyltransferases synthesize products up to $C_{50}$ in chain lengths, which can be further divided into short-chain ($C_{10}$–$C_{25}$), medium-chain ($C_{30}$–$C_{35}$), and long-chain ($C_{40}$–$C_{50}$) (14) and *cis*-prenyltransferases generate much longer products (15). In the co-crystal structure of FPP synthase (FPPs) with nonreactive thiol analog of the allylic substrate DMAPP and the homoallylic substrate IPP, the first DD*XX*D binds the pyrophosphate head group of the allylic substrate via $Mg^{2+}$ and the second DD*XX*D coordinates with two other $Mg^{2+}$ ions to form a trinuclear cluster for pulling the pyrophosphate leaving group of the allylic substrate (16). A bulky amino acid residue at the fourth or fifth position before the first DD*XX*D motif in helix D (see Fig. 1 for sequence homology) forms a blockage underneath the allylic substrate site to avoid the further elongation of FPP product. Replacement of the large amino acid with a small one removes the blockage and leads to longer products (17–19). For the long-chain $C_{40}$-octaprenyl pyrophosphate synthase (OPPs), its crystal structure shows a small residue at that position and a large amino acid $Phe^{132}$ from another helix (helix G) seals the bottom of the active site crevice (20). Similar to OPPs, medium-chain $C_{30}$ hexaprenyl pyrophosphate synthase (HexPPs) also utilizes L164 from helix G for shielding the final product in the middle of the crevice (10).

It is interesting that the short-chain geranylgeranyl pyrophosphate synthases (GGPPs) from different species contain either large or small residues at the fourth or fifth position before the first DD*XX*D (Fig. 1). By virtue of one IPP condensation reaction with FPP, GGPPs synthesizes $C_{20}$ GGPP, which serves as a precursor for carotenoids, geranylgeranylated proteins, chlorophylls, and archaeal ether-linked lipid for the cytoplasmic membrane (21). These GGPPs are classified into three types based on their amino acid sequences; type-I GGPPs contains a large amino acid at the fourth or fifth position prior to the DD*XX*D motif, but type-II (with insertion of two

* This work was supported by National Science Council Grant NSC94-2311-B-001-034 and a grant from Academia Sinica (to P.-H. L.). The Biological Crystallography Facilities (Taiwan Beamline BL12B2 at SPring-8) were supported by the National Science Council. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

*The atomic coordinates and structure factors (code 2DH4) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).*

[1] To whom correspondence may be addressed: Institute of Biological Chemistry, Academia Sinica, 128 Academia Rd., Taipei 115, Taiwan. Tel.: 886-2-2785-5696 (ext. 5010); Fax: 886-2-2788-9759; E-mail: ahjwang@gate.sinica.edu.tw.

[2] To whom correspondence may be addressed: Institute of Biological Chemistry, Academia Sinica, 128 Academia Rd., Taipei 115, Taiwan. Tel.: 886-2-2785-5696 (ext. 6070); Fax: 886-2-2788-9759; E-mail: phliang@gate.sinica.edu.tw.

[3] The abbreviations used are: IPP, isopentenyl pyrophosphate; DMAPP, dimethylallyl pyrophosphate; GPP, geranyl pyrophosphate; FPP, farnesyl pyrophosphate; GGPP, geranylgeranyl pyrophosphate; FPPs, farnesyl pyrophosphate synthase; OPPs, octaprenyl pyrophosphate synthase; GGPPs, geranylgeranyl pyrophosphate synthase; SeMet, selenomethionine; NTA, nitrilotriacetic acid; HexPPs, hexaprenyl pyrophosphate synthase.

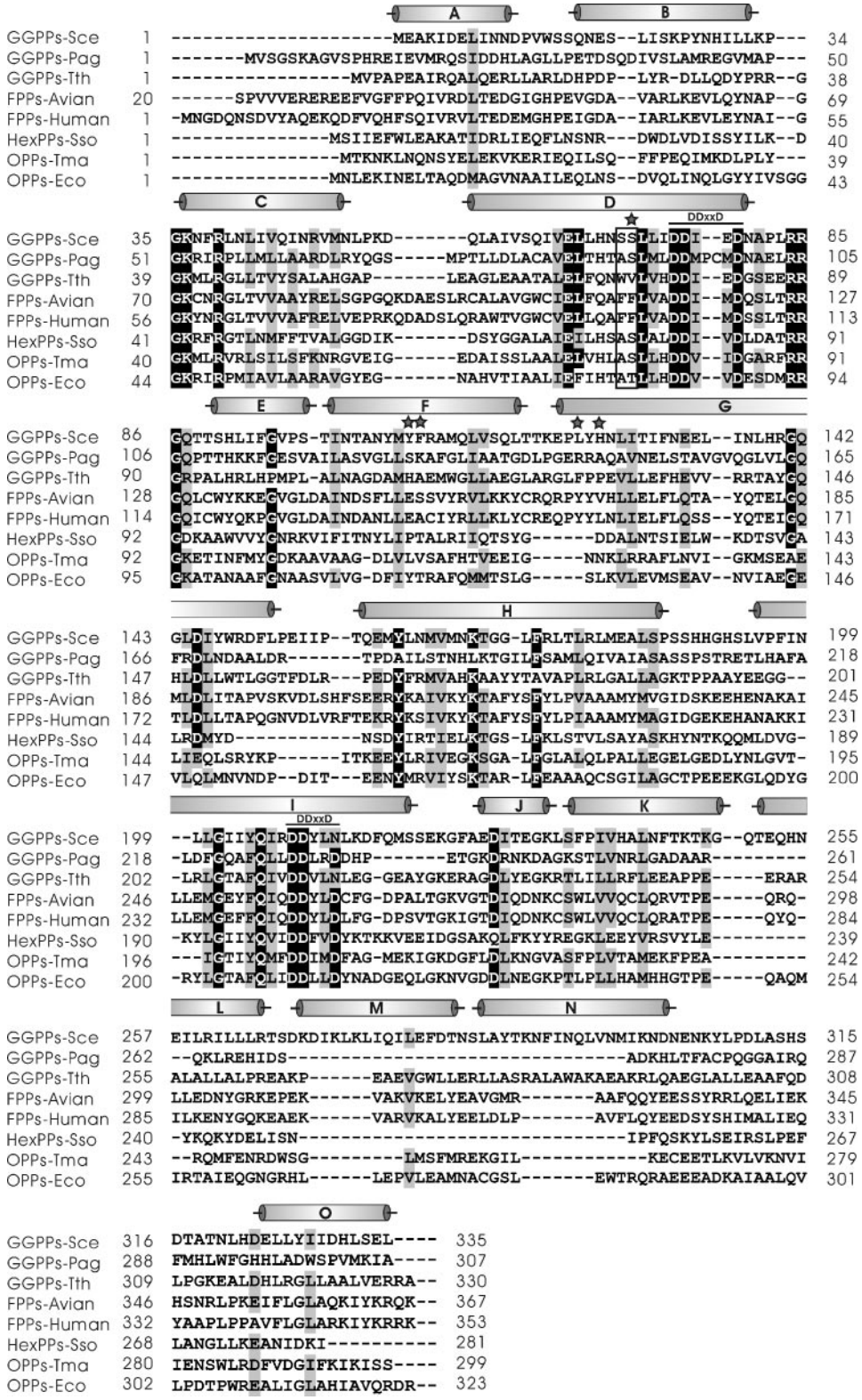


FIGURE 1. **Alignment of amino acid sequences of *S. cerevisiae* type-III GGPPs (*GGPPs-Sce*), *Pantoea agglomerans* type-II GGPPs (*GGPPs-Pag*), *T. thermophilus* type-I GGPPs (*GGPPs-Tth*; Protein Data Bank accession number 1WMW), the avian FPPs (Protein Data Bank accession number 1FPS), human FPPs (Protein Data Bank accession number 1YQ7 and 1YV5), *S. solfataricus* HexPPs (HexPPs-Sso, Protein Data Bank accession number 2AZJ), *T. maritima* OPPs (OPPs-Tma, Protein Data Bank accession number 1V4E), and *E. coli* OPPs (OPPs-Eco).** Identical and similar amino acid residues are shown in *black* and *gray shadow*, respectively. The regions of the 15 helices of GGPPs-Sce are indicated at the *top*. The *stars* indicate the amino acid residues mutated in this study. The *box* denotes the fourth and fifth amino acid residues, which are small in long-chain OPPs, medium-chain HexPPs, and type-II and -III GGPPs but large in short-chain FPPs and type-I GGPPs.

amino acids within the first Asp-rich motif) and type-III GGPPs found in eukaryotes (except plants) have a small amino acid at the corresponding position (see Fig. 1) (22). To rationalize the mechanism of product chain length determination for the type-III GGPPs, we solved the crystal structure of *Saccharomyces cerevisiae* GGPPs to 1.98 Å resolution as reported here. While bearing a common structural feature with other *trans*-prenyltransferases, GGPPs structure reveals a unique orientation of its N terminus. The helix A of type-III GGPPs protrudes into the other subunit, which seems to be involved in the tight dimer formation. Mutants truncated by deleting this helix and the following loop region were examined for dimer formation as presented here. On the basis of the structure, we carried out site-directed mutagenesis studies to identify the amino acids critical for product chain length determination. Together with recently

solved type-I GGPPs,[4] our structures enhance the understanding of the chain elongation mechanisms for the short-chain *trans*-prenyltransferases.

## EXPERIMENTAL PROCEDURES

*Materials*—Radiolabeled [$^{14}$C]IPP (55 mCi/mmol) was purchased from Amersham Biosciences. Nonlabeled FPP and seleno-DL-methionine were obtained from Sigma. Reverse-phase TLC plates were purchased from Merck. *PfuTurbo* DNA polymerase was obtained from Invitrogen. The plasmid miniprep kit, DNA gel extraction kit, and $Ni^{2+}$-NTA resin were purchased from Qiagen. Potato acid phosphatase (2 units/mg) was purchased from Roche Applied Science. Factor Xa and the protein expression kit (including the pET32Xa/LIC vector and competent JM109 and BL21 cells) were obtained from Novagen. The QuikChange site-directed mutagenesis kit was obtained from Stratagene. All commercial buffers and reagents were of the highest grade.

*Expression and Purification of GGPPs*—The gene encoding GGPPs was cloned from *S. cerevisiae* genomic DNA by using PCR and the forward primer 5′-ggtattgagggtcgcatggaggccaagatagatg-3′ and the reverse primer 5′-agaggagagttagagcctcacaattcggataagtgg-3′ into the pET32Xa/LIC vector. In this construct, we included a five-residue sequence (MTKNK) from *Thermotoga maritima* OPPs between the N terminus of GGPPs and the Factor Xa cleavage sequence to expose the protease site for tag removal. The recombinant GGPPs plasmid was then used to transform *Escherichia coli* JM109 competent cells that were streaked on a Luria-Bertani (LB) agar plate containing 100 $\mu$g/ml ampicillin. Ampicillin-resistant colonies were selected from the agar plate and grown in 5 ml of LB culture containing 100 $\mu$g/ml ampicillin overnight at 37 °C. The correct construct by sequencing was subsequently transformed to *E. coli* BL21 (*DE*3) for protein expression. The 60-ml overnight culture of a single transformant was used to inoculate 6 liters of fresh LB medium containing 100 $\mu$g/ml ampicillin. The cells were grown to $A_{600}$ = 0.6 and induced with 1 mM isopropyl $\beta$-thiogalactopyranoside at 16 °C. After 16 h, the cells were harvested by centrifugation at 7,000 × *g* for 15 min to collect the cell paste.

The enzyme purification was conducted at 4 °C. Cell paste was suspended in 75 ml of lysis buffer containing 25 mM Tris-HCl, pH 7.5, and 150 mM NaCl. Cell lysate was prepared with a French pressure cell press (AIM-AMINCO Spectronic Instruments) and centrifuged at 17,000 × *g* to remove cell debris. The cell-free extract was loaded onto a $Ni^{2+}$-NTA column, which had been previously equilibrated with lysis buffer. The column was washed with 10 mM imidazole followed by 20 mM imidazole-containing buffer. His-tagged GGPPs eluted with 100 mM imidazole was dialyzed twice against 3 liters of buffer (25 mM Tris-HCl, pH 7.5, and 150 mM NaCl) and then subjected to Factor Xa digestion to remove the tag. The mixture was then passed through another $Ni^{2+}$-NTA column, and subsequently untagged GGPPs was eluted with 10 mM imidazole-containing buffer and then dialyzed twice against 3 liters of buffer (25 mM Tris-HCl, pH 7.5, and 150 mM NaCl) for storage. SDS-PAGE analysis was used to check the purity (>95%) of GGPPs and its mutants.

*Preparation of Selenomethionine-labeled GGPPs*—To produce selenomethionine (SeMet)-labeled GGPPs for x-ray phase solving, the plasmid that GGPPs gene was transformed into a BL21 (*DE*3)-competent cell. The single transformant was first grown overnight at 37 °C in 50 ml of LB medium containing 100 $\mu$g/ml ampicillin. The cells were harvested by centrifugation at 7,000 × *g* for 15 min. The paste suspended in SeMet minimal medium (24) was then used to inoculate 6 liters of SeMet minimal medium with the supplement of 100 $\mu$g/ml seleno-DL-methionine. Isopropyl $\beta$-thiogalactopyranoside was added to a final concentration of 1 mM for induction after the cell density reached $A_{600}$ = 0.4. The cells were grown for another 96 h at 16 °C and then harvested by centrifugation. The subsequent purification procedure for SeMet GGPPs was the same as for the native GGPPs.

**TABLE 1**
**Data collection and refinement statistics of the *S. cerevisiae* native and SeMet-labeled GGPPs crystals**

The unit cell dimensions for native are as follows: $a$ = 48.75, $b$ = 116.33, and $c$ = 129.77 Å (two molecules per asymmetric unit). Dimensions for SeMet-labeled GGPPs are as follows: $a$ = 48.84, $b$ = 116.67, $c$ = 129.76 Å (two molecules per asymmetric unit).

| Parameters | Values | |
|---|---|---|
| | SeMet | Native |
| **Data set** | | |
| Wavelength | 0.9796 | 1.1274 |
| Space group | $P2_12_12_1$ | |
| Resolution range (Å)[a] | 30–2.4 (2.49–2.40) | 50–1.98 (2.05–1.98) |
| No. of reflections | | |
| Unique | 28,379 (2829) | 50,750 (4280) |
| Observed | 302,189 (28,290) | 254,446 (17,548) |
| Completeness (%)[a] | 94.4 (95.8) | 96.2 (82.7) |
| $R_{merge}$ (%)[a] | 8.6 (46.1) | 5.3 (52.8) |
| $I/\sigma(I)$[a] | 26.15 (6.09) | 28.04 (2.48) |
| Selenium sites | 20 | |
| Figure of merit | 0.32 | |
| **Refinement** | | |
| $R_{factor}$ (%)[a] | | 18.90 (30.57) |
| $R_{free}$ (%)[a] | | 24.32 (35.70) |
| Deviations | | |
| Bond lengths (Å) | | 0.018 |
| Bond angles (degrees) | | 1.66 |
| Average B factors (Å) | | |
| 644 residues | | 30.76 |
| 578 water molecules | | 46.67 |
| 2 Mg ions | | 27.09 |
| Ramachandran plot | | |
| Residues in most favored regions (%) | | 94.1 |
| Residues in additionally allowed regions (%) | | 5.9 |

[a] The values in the parentheses are for the highest resolution shells.

*Crystallization and Data Collection*—Native and SeMet-labeled GGPPs were crystallized using the hanging drop method from Hampton Research (Laguna Niguel, CA) by mixing 2 $\mu$l of the GGPPs solution (10–12 mg/ml in 25 mM Tris-HCl, pH 7.5, 150 mM NaCl, and 0.1% Triton X-100) with 2 $\mu$l of the mother liquor (0.08 M $CH_3COONa$, 0.145 M $(NH_4)_2SO_4$, 13% polyethylene glycol 4000, 7–9% glycerol, and 7–9% 1,2-propanediol) and equilibrating with 500 $\mu$l of the mother liquor. Within 7 days at room temperature, crystals grew to dimensions of about 0.4 × 0.15 × 0.15 mm. The x-ray diffraction data for the native GGPPs were collected to 1.98 Å resolution at beam line BL17B2 of the National Synchrotron Radiation Research Center (NSRRC, Hsinchu, Taiwan), and the SeMet-labeled GGPPs data were collected to 2.4 Å at 100 K on the Taiwan Contract BL12B2 station at SPring-8 (Hyogo, Japan) by employing the multiple-wavelength anomalous diffraction method. The diffraction data were processed and scaled by using the program HKL2000 (25). These GGPPs crystals belong to the space group $P2_12_12_1$, with typical unit cell parameters of $a$ = 48 Å, $b$ = 116 Å, and $c$ = 129 Å. Each asymmetric unit contained a dimeric GGPPs molecule.

*Structure Determination and Refinement*—The multiple-wavelength anomalous diffraction data sets in the resolution range of 30 to 2.4 Å were collected at wavelengths of 1.0332 Å (low remote), 0.9799 Å (inflection point), 0.9796 Å (peak), and 0.9537 Å (high remote) at and near the selenium absorption edge. Using SOLVE (26), 20 selenium sites

[4] Crystal structure of geranylgeranyl diphosphate synthase from *T. thermophilus*, deposited by K. Suto, K. Nishio, Y. Nodake, K. Hamada, M. Kawamoto, N. Nakagawa, S. Kuramitu, and K. Miura, Riken Structural Genomics Proteomics and Initiative (Rsgi) on July 21, 2004.

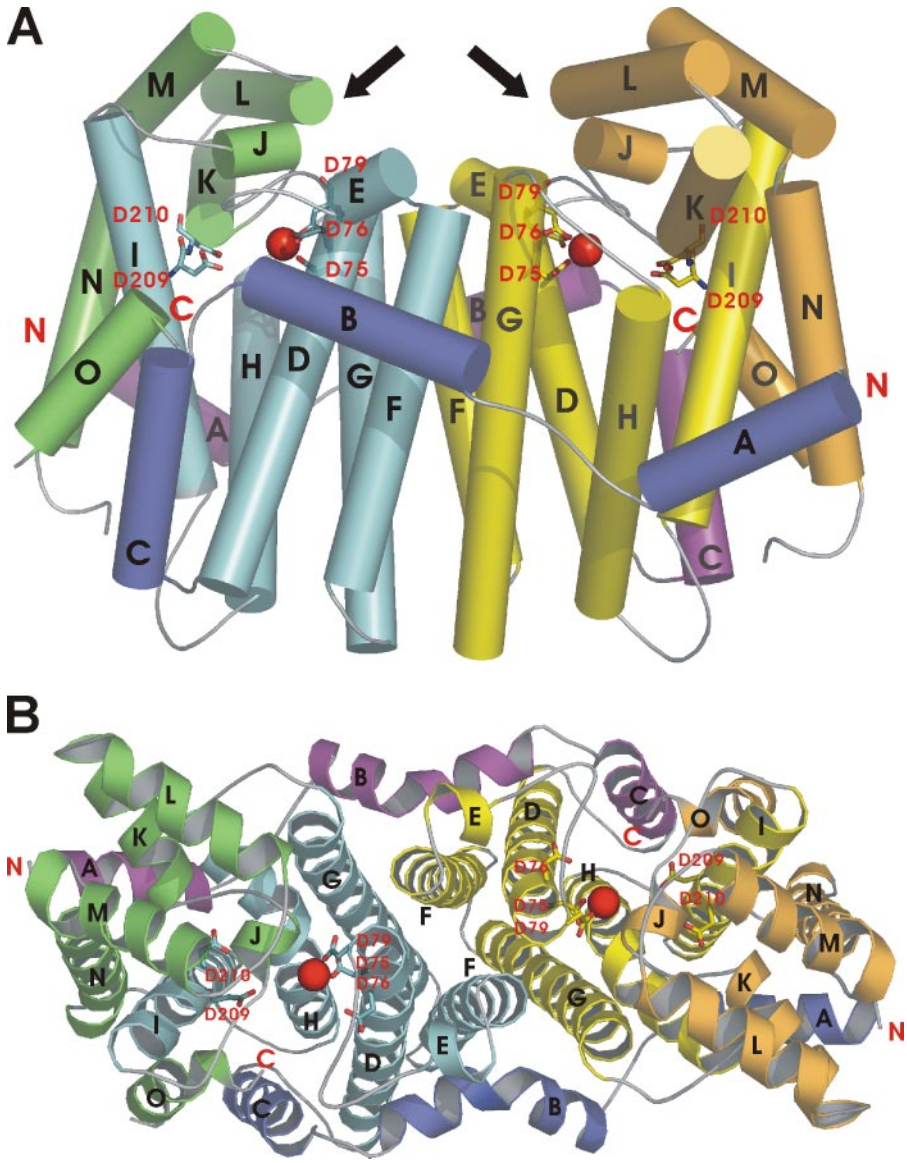


FIGURE 2. **The structural model of *S. cerevisiae* GGPPs shown in *cylinder* and *ribbon diagrams*.** *A*, N-terminal α-helices in the outer layer are shown in *blue* and *purple* for the two individual subunits, whereas the six helices in the inner layers are shown in *cyan* and *yellow*, respectively. Two small peripheral domains are *colored green* and *orange*. The *black arrows* indicate the location of the active site with the Asp residues in the two DD*XX*D motifs shown as *sticks* in *cyan* and *yellow*. Two identical subunits are associated into a dimer by forming a four-layer helix bundle (F and G helixes) and a four layer helix arm (A and B helixes). $Mg^{2+}$ ions observed near the DD*XX*D motifs are shown as *red balls*. *B*, the top view model of *S. cerevisiae* GGPPs is shown as a *ribbon diagram*. The *color* and *styles* shown here are the same as above. These images were produced using PyMol.

in the asymmetric unit were located, and the phase angles were determined in the single wavelength anomalous diffraction method using the peak wavelength data. Further iterative cycles of density modification, model building, and refinement were carried out with RESOLVE (27), Xtalview (28), and the CNS program (29), respectively. $R_{free}$ values were calculated by using 5% reflections. The stereochemical quality of the refined structure was checked with the program PROCHECK (30). Each monomer of the refined structure encompassed 322 of the 335 total residues of the GGPPs sequence, with a small disordered region of residues 315–327. 578 well ordered water molecules were also located and included in the model. Refinement statistics are summarized in Table 1. All of the structural diagrams were drawn by using the PyMol software (available on the World Wide Web at pymol.sourceforge.net/) and GRASP (31).

*Site-directed Mutagenesis of GGPPs*—GGPPs mutants were prepared by using the QuikChange site-directed mutagenesis kit in conjunction with the *S. cerevisiae* GGPPs gene template in the pET32Xa/LIC vector. The mutagenic primers used were prepared by MDBio Inc. (Taiwan). The mutagenic oligonucleotides for performing site-directed mutagenesis are 5′-caccgcaaattatatgGCGttcagagccatgcaacttg-3′ for Y107A, 5′-caccgcaaattatatgtatGCGagagccatgcaacttg-3′ for F108A, 5′-caccgcaaattatatgGCGGCGagagccatgcaacttg-3′ for Y107A/F108A, 5′-cgatttcaacgaagaaGCGatcaatctacatagg-3′ for L135A, 5′-gaagaattgatcaatctaGCGaggggacaaggcttgg-3′ for H139A, and 5′-gctcttgcataattccTACctttaatcgacgatatag-3′ for S71Y (the mutated nucleotides are underlined and shown in capital letters). The multiple mutants were constructed with the combination of the above mutagenic primers. The mutations were confirmed by sequencing the entire GGPPs mutant gene of the plasmid obtained from the overnight culture. The correct construct was subsequently transformed to *E. coli* BL21(*DE*3) for protein expression and purification. The purity (>95%) of each purified mutant GGPPs was checked by SDS-PAGE.

*Kinetic Parameters for Mutant GGPPs*—For enzyme activity measurements, each mutant GGPPs (0.2 μM Y107A, F108A, H139A, Y107A/F108A, Y107A/H139A, L135A/H139A, F108A/H139A, Y107A/F108A/H139A, or Y107A/F108A//L135A/H139A) was used. The reaction was initiated in 200 μl of solution containing 100 mM Hepes (pH 7.5), various concentrations of FPP and [$^{14}$C]IPP as specified below, 50 mM KCl, 0.5 mM $MgCl_2$, and 0.1% Triton X-100 at 25 °C. The enzyme concentration used in all experiments was determined from its absorbance at 280 nm ([cepsilon] = 20,340 $M^{-1}$ $cm^{-1}$). The measurements of the kinetic parameters for the wild type and mutants followed our published pro-

cedure (32, 33). For IPP $K_m$ determinations, 20 $\mu$M FPP was utilized to saturate the enzyme, and IPP concentrations from 0.25 to 400 $\mu$M varied with the $K_m$ of IPP for each mutant were employed. For FPP $K_m$ and $k_{cat}$ measurements, 0.25–32 $\mu$M FPP was used along with 50 $\mu$M [$^{14}$C]IPP, except for the mutants of Y107A/H139A, F108A/H139A, Y107A/F108A/H139A, and Y107A/F108A/L135A/H139A, where 250 $\mu$M [$^{14}$C]IPP was used. For Y107A/H139A, Y107A/F108A/H139A, and Y107A/F108A/L135A/H139A, the enzyme reaction was initiated by adding 0.1 $\mu$M mutants to measure the FPP $K_m$ and $k_{cat}$. GPP $K_m$ measurements for wild type and S71Y used 0.25–50 $\mu$M GPP along with 50 $\mu$M [$^{14}$C]IPP. To measure the initial rate, 40-$\mu$l portions of the reaction mixture were periodically withdrawn within 10% substrate depletion and then mixed with 10 mM EDTA for reaction termination. The radiolabeled products were then extracted with 1-butanol, and the radioactivities associated with aqueous and butanol phases were separately quantitated by using a Beckman LS6500 scintillation counter. Data of initial rates *versus* substrate concentrations were analyzed by nonlinear regression of the Michaelis-Menten equation using the KaleidaGraph computer program (Synergy software) to obtain $K_m$ and $V_{max}$ values. The $k_{cat}$ was calculated from $V_{max}/[E]$. Steady-state activity of the S71Y mutant was assayed under the same condition except that GPP instead of FPP was used.

*Product Analysis*—The GGPPs reaction containing 1 $\mu$M enzyme (wild-type or mutant GGPPs), 10 $\mu$M FPP, 150 $\mu$M [$^{14}$C]IPP, 0.1% Triton X-100, 0.5 mM $MgCl_2$, and 50 mM KCl in 100 mM Hepes buffer (pH 7.5) was incubated for 6 h at 25 °C. Ten mM EDTA was used to terminate the enzyme reaction. For identification of S71Y product, GPP instead of FPP was used as the allylic substrate. The radiolabeled products were extracted with 1-butanol. The 1-butanol was then evaporated, and the 20% 1-propanol solution containing 4.4 units/ml acidic phosphatase, 0.1% Triton X-100, 50 mM sodium acetate (pH 4.7) was used to convert the products to the corresponding alcohols according to the reported procedure (32). After the pyrophosphate hydrolysis catalyzed by acidic phosphatase was completed, the polyprenols were extracted with *n*-hexane. The polyprenols were separated on reversed-phase TLC using acetone/water (19:1) as mobile phase. The radiolabeled products were identified by autoradiography using a bioimaging analyzer FUJIFILM BAS-1500 (Japan) according to their $R_f$ (retention factor) values reported (22).

*Construction of N-terminal Truncated Mutants*—For constructing the genes to express the truncated GGPPs without the first N-terminal 17 amino acids Δ(1–17) or without the first 9 amino acids Δ(1–9), a common reverse primer, 5′-agaggagagttagagcctcacaattcggataagtgg-3′, was used with the forward primer 5′-ggtattgagggtcgccaaaatgaaagcttgatttc-3′ for Δ(1–17) or 5′-ggtattgagggtcgcaataatgatcctgtttggtc-3′ for Δ(1–9). The PCR-amplified products from the template were inserted into the pET32Xa/LIC vector as described above. The deletion was confirmed by sequencing, and correct construct was subsequently transformed to *E. coli* BL21 (*DE*3) for protein expression. The truncated proteins were purified using a $Ni^{2+}$-NTA column, and their activities were measured as the other mutant enzymes described above.

*Gel Filtration Experiments*—The molecular mass of GGPPs was determined on a Superdex 75 10/300 GL High Performance column (1 × 30 cm; Amersham Biosciences) by comparing the elution volume of the GGPPs (1 and 0.2 mg/ml) with those of protein molecular mass standards, including catalase (232 kDa), aldolase (158 kDa), albumin (67 kDa), ovalbumin (43 kDa), and ribonuclease A (13.7 kDa). A buffer of 25 mM Tris-HCl, pH 7.5, and 150 mM NaCl was used to elute the proteins at a flow rate of 0.5 ml/min.

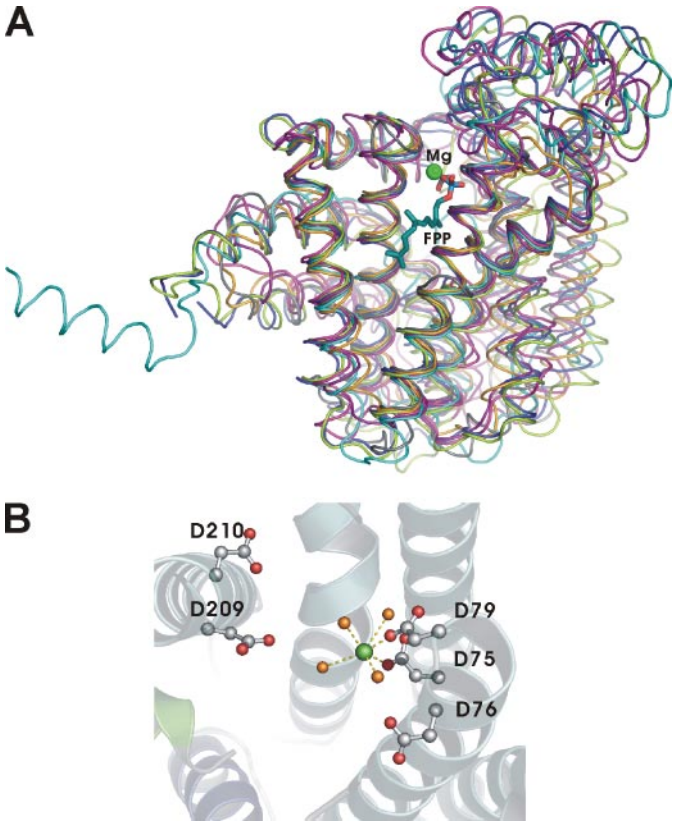


FIGURE 3. *A*, seven homologous proteins, including avian FPPs (in *green*) with FPP bound, *E. coli* FPPs (Protein Data Bank accession number IRQI; shown in *gray*), human FPPs (in *blue*), *T. thermophilus* type-I GGPPs (in *magenta*), *S. cerevisiae* type-III GGPPs (in *cyan*) containing one $Mg^{2+}$ ion (in *green*), *S. solfataricus* HexPPs (in *orange*), and *T. maritima* OPPs (in *purple*) are superimposed by the SPDB viewer (23). The FPP and $Mg^{2+}$ ion observed in the superimposed structures are shown with a *ball-and-stick model*. According to the structures, the N-terminal 17 residues of *S. cerevisiae* type-III GGPPs adopt opposite orientation compared with the other prenyltransferases. *B*, $Mg^{2+}$ ion (in *green*) in the active site is octahedrally coordinated with two carboxylate groups from $Asp^{75}$ and $Asp^{79}$ of the first DD*XX*D and four water molecules (in *orange*). Coordination bonds are shown as *yellow dotted lines*. The $Mg^{2+}$ ion binds with the pyrophosphate of FPP and plays a catalytic role according to the previous studies. These images were generated using PyMol.

## RESULTS

*Overall Structure of S. cerevisiae GGPPs*—The crystal structure of *S. cerevisiae* GGPPs has been determined to 1.98 Å resolution. As shown in Fig. 2, *A* and *B*, each asymmetric unit of the crystal unit cell contains one GGPPs dimer, and two identical subunits are associated into a dimer by forming a four-layer helix bundle using helices F and G and a four-layer helix arm using helices A and B. The refined structure of type-III GGPPs in complex with one magnesium ion in each subunit contains amino acid residues 2–314 and 328–340. The electron density map between residues 315 and 327 at the C terminus is not clearly visible. The structure contains 15 $\alpha$ helices, nine of them (helices A–I) surrounding a large central cavity. Helices F and G are involved in the dimer formation, with the major stabilization coming from the helices F ($\alpha$ chain)–F ($\beta$ chain) and F ($\alpha$ chain)–G ($\beta$ chain) intersubunit hydrophobic interactions and hydrogen bonding. The side chain of $Met^{111}$ is stacked with that of $Met^{111}$ from the other subunit, so it is different from the *T. maritima* OPPs, which uses aromatic amino acid of $Phe^{117}$ for hydrophobic interaction (9), and *Sulfolobus solfataricus* HexPPs, which uses the crossing stacking interaction between $Trp^{136}$ from monomer A and $Pro^{114}$ from monomer B (10).

*Active Site*—Two conserved DD*XX*D motifs located on helices D and I, respectively, near the opening of a deep cleft, are proposed to be the

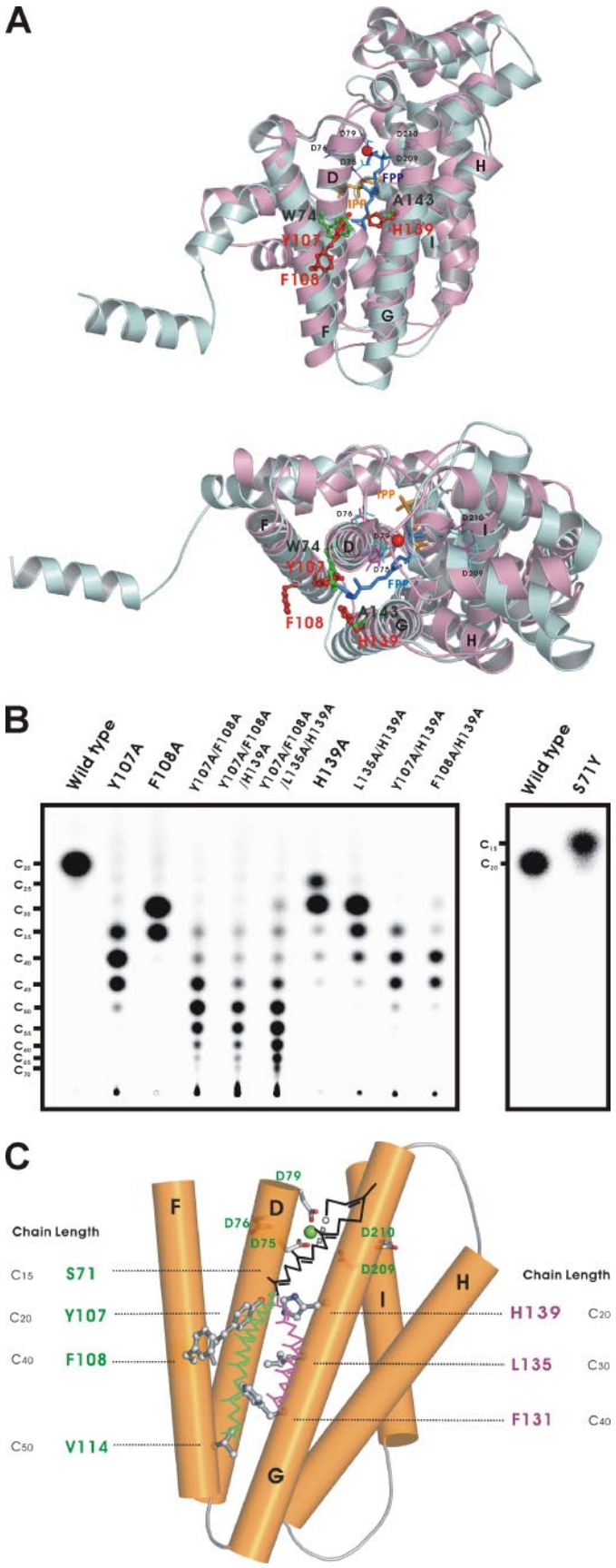


FIGURE 4. *A*, superimposed structures of *S. cerevisiae* type-III GGPPs (in *cyan*) and *T. thermophilus* type-I GGPPs (in *pink*). The FPP (in *blue*) from avian FPPs and the IPP (in *orange*) from *E. coli* FPPs shown with a *stick model* are also superimposed. $Trp^{74}$ and $Ala^{143}$ of type-I GGPPs shown in a *stick-and-ball model* with *green color* have similar spatial location as $Tyr^{107}$ and $His^{139}$ shown in *red* of type-III GGPPs, respectively. $Trp^{74}$ at the fifth position before the first DD*XX*D motif in helix D of type-I GGPPs can be predicted to be a critical residue for chain length determination. *B*, products synthesized by wild-type and mutant GGPPs. Y107A, F108A, and H139A generated $C_{40}$ and $C_{30}$ products, which are longer than the $C_{20}$ synthesized by the wild type. $Tyr^{107}$ plays a more important role in chain length determination. S71Y synthesizes shorter product $C_{15}$, a single condensation between the bound GPP and IPP. Their product distributions are summarized in Table 3. *C*, the molecular ruler mechanism for chain length determination catalyzed by the type-III GGPPs. The final product of $C_{20}$ GGPP shown in *black* is shielded by $Tyr^{107}$ and $His^{139}$ at the *bottom*. The H139A mutation has yielded smaller space for product elongation through helix G (in *purple*). However, Y107A provides more space to accumulate longer product ($C_{40}$) along helix F (in green). The substitution of $Ser^{71}$, the fourth amino acid prior to the first DD*XX*D motif, with a large residue of Tyr caused the formation of $C_{15}$.

substrate-binding pocket (Fig. 2*A*). In the superimposed three-dimensional structures of this GGPPs with other *trans*-prenyltransferases (Fig. 3*A*), this region is coincident with the active site proposed for FPPs. From the potential surface diagram (not shown), this proposed active site pocket is surrounded by the polar (mostly positively charged) residues, including $Arg^{84}$, $Arg^{85}$, $His^{68}$, $Lys^{37}$, and $Arg^{40}$, which are probably responsible for attracting if not directly interacting with the negatively charged pyrophosphate head groups of the substrates. As revealed by a more focused view of this active site (Fig. 3*B*), a $Mg^{2+}$ ion is coordinated by $Asp^{75}$ and $Asp^{79}$ from the first DD*XX*D and four water molecules. The distances of coordination bonds from the $Mg^{2+}$ to $Asp^{75}$, $Asp^{79}$, and four water molecules are 2.06, 2.0, 2.5, 3.08, 3.07, and 2.12 Å, respectively, in $\alpha$ subunit and 2.03, 2.06, 2.3, 2.33, 2.8, and 3.04 Å, respectively, in $\beta$ subunit. This $Mg^{2+}$ supposedly coordinates with the pyrophosphate group of FPP. There are two short $\alpha$ helices, helix E ($Ser^{95}$–$Phe^{99}$) and helix J ($Glu^{231}$–$Glu^{235}$) on the outer surface of the $\alpha$-cone. The region including the loop between D and E, the loop between I and J, helix J, and the loop between J and K are not part of the active site, but it may switch the top region of the active site between open and closed forms, as shown in FPPs (16) and HexPPs structures (10). The open form may allow the substrates to enter and the final product to release, and the closed form makes the tight binding of the substrates with the enzyme (16).

*The Amino Acids Responsible for Controlling Product Chain Length*—The active site crevice is surrounded by five $\alpha$-helices (helices D, F, G, H, and I). Underneath the substrate-binding site, three large amino acid residues, including $Tyr^{107}$ and $Phe^{108}$ on the side of helix F and $His^{139}$ on another side of helix G, occupy the bottom portion of the elongated crevice (Fig. 4*A*). These residues are probably responsible for blocking the further chain elongation of the $C_{20}$ products. In fact, $His^{139}$ has been previously suggested to be important in controlling the product chain length for the type-III GGPPs, since H139A formed larger products than $C_{20}$ (22).

*Reaction Kinetics and Final Products of Different Mutants*—To test the above hypothesis, we replaced these large residues $Tyr^{107}$, $Phe^{108}$, and $His^{139}$ with Ala and examined the chain lengths of their products. Less than 5-fold changes of enzyme $k_{cat}$ values for the mutants as compared with the wild type (see Table 2) were observed. However, the FPP $K_m$ of Y107A/F108A/H139A was decreased 9-fold, indicating that this triple mutant has higher FPP affinity. None of these single mutations or the mutations on other sites has caused significantly reduced FPP $K_m$. The increase of IPP $K_m$ values by 120, 65, 22, 340, and 170-fold compared with the wild type were observed in the mutants of Y107A/F108A/H139A, Y107A/F108A/L135A/H139A, L135A/H139A, Y107A/H139A, and F108A/H139A, respectively, indicating significantly lower affinity of these mutants with IPP. However, under the high concentrations of FPP (10 $\mu$M) and IPP (150 $\mu$M), which were used to yield the products by all of the mutants as shown below, the activities were not significantly changed (less than 5-fold). Further increase of the substrate

**TABLE 2**

**Kinetic parameters of wild-type and mutant *S. cerevisiae* GGPPs**

The steady-state kinetics of wild-type and mutant GGPPs were assayed with FPP and [$^{14}$C]IPP, and S71Y was assayed with GPP and [$^{14}$C]IPP at pH 7.5 and 25 °C.

| *S. cerevisiae* GGPPs | $k_{cat}$ | $K_m^{FPP}$ | $K_m^{IPP}$ | Rel $k_{cat}$[a] |
|---|---|---|---|---|
| | $s^{-1}$ | $\mu$M | | -fold |
| Wild type | $(2.5 \pm 0.4) \times 10^{-2}$ | 3.2 ± 0.3 | 0.8 ± 0.2 | 1 |
| Y107A | $(5.6 \pm 0.2) \times 10^{-2}$ | 1.7 ± 0.2 | 4.7 ± 0.9 | 2.2 |
| F108A | $(4.7 \pm 0.2) \times 10^{-2}$ | 2.1 ± 0.4 | 2.7 ± 0.6 | 1.9 |
| Y107A/F108A | $(4.9 \pm 0.3) \times 10^{-2}$ | 2.6 ± 0.5 | 3.3 ± 0.5 | 2.0 |
| Y107A/F108A/H139A | $(2.5 \pm 0.1) \times 10^{-2}$ | 0.4 ± 0.1 | 97 ± 26 | 1.0 |
| Y107A/F108A/L135A/H139A | $(8.4 \pm 0.7) \times 10^{-3}$ | 1.0 ± 0.2 | 51 ± 12 | 0.3 |
| H139A | $(4.2 \pm 0.5) \times 10^{-2}$ | 2.5 ± 1.1 | 5.1 ± 1.7 | 1.7 |
| L135A/H139A | $(6.4 \pm 0.5) \times 10^{-2}$ | 2.7 ± 0.7 | 17 ± 3 | 2.6 |
| Y107A/H139A | $(3.9 \pm 0.3) \times 10^{-2}$ | 1.1 ± 0.5 | 267 ± 42 | 1.6 |
| F108A/H139A | $(1.6 \pm 0.2) \times 10^{-2}$ | 5.1 ± 2.2 | 136 ± 43 | 0.6 |
| Wild type[b] | $(1.3 \pm 0.1) \times 10^{-2}$ | 0.9 ± 0.1[c] | 1.3 ± 0.2 | 1[d] |
| S71Y[b] | $(3.4 \pm 0.3) \times 10^{-3}$ | 1.7 ± 0.8[c] | 169 ± 54 | 0.3[d] |

[a] $k_{cat}$ relative to that of the wild type.
[b] Reaction with GPP and [$^{14}$C]IPP.
[c] $K_m$ for GPP.
[d] $k_{cat}$ relative to using GPP as the substrate.

**TABLE 3**

**Product distribution of the wild-type and mutant type-III GGPPs from *S. cerevisiae***

The enzyme-catalyzed condensation reactions were performed with 10 $\mu$M FPP and 150 $\mu$M [$^{14}$C]IPP at pH 7.5 and 25 °C. The radioactivity of each product was normalized by the number of [$^{14}$C] IPP incorporated.

| *S. cerevisiae* GGPPs | Product | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $C_{20}$ | $C_{25}$ | $C_{30}$ | $C_{35}$ | $C_{40}$ | $C_{45}$ | $C_{50}$ | $C_{55}$ | $C_{60}$ | $C_{65}$ | $C_{70}$ |
| | % | % | % | % | % | % | % | % | % | % | % |
| Wild type | 100 | | | | | | | | | | |
| Y107A | | | | 24.5 | 49.8 | 22.7 | 3.0 | | | | |
| F108A | | | 74.2 | 25.8 | | | | | | | |
| Y107A/F108A | | | | 12.4 | 11.1 | 25.4 | 33.1 | 13.0 | 3.8 | 1.3 | |
| Y107A/F108A/H139A | | | | 17.7 | 11.1 | 13.3 | 27.6 | 20.4 | 6.2 | 2.1 | 1.6 |
| Y107A/F108A/L135A/H139A | | | 13.6 | 9.4 | 5.1 | 10.1 | 21.0 | 21.8 | 11.1 | 5.5 | 2.4 |
| H139A | | 34.8 | 57.8 | 3.9 | 3.5 | | | | | | |
| L135A/H139A | | | 72.8 | 18.0 | 7.5 | 1.7 | | | | | |
| Y107A/H139A | | | | 31.0 | 34.7 | 28.6 | 5.6 | | | | |
| F108A/H139A | | | 17.1 | 14.3 | 40.8 | 23.9 | 3.8 | | | | |

concentrations (up to 20 $\mu$M FPP and 320 $\mu$M [$^{14}$C]IPP) did not lead to different final products (data not shown).

The final products of wild-type and mutant enzymes at 1 $\mu$M concentration with the amount of 10 $\mu$M FPP and 150 $\mu$M [$^{14}$C]IPP were obtained. Under the same reaction condition, wild-type GGPPs generated a single $C_{20}$ product, but Y107A and F108A, which are associated with helix F, generated larger $C_{40}$ and $C_{30}$ as major products, respectively (see Fig. 4*B*, *left*, and Table 3, $C_{35}/C_{40}/C_{45}/C_{50}$ = 25:50:23:3 for Y107A, $C_{30}/C_{35}$ = 74:26 for F108A). Y107A/F108A double mutant generated even larger $C_{50}$ major product ($C_{35}/C_{40}/C_{45}/C_{50}/C_{55}$ = 12:11:25:33:13). Additional replacement of His$^{139}$ with Ala on another side (helix G) did not significantly increase the product chain length but slightly increased the percentage of $C_{55}$ product ($C_{35}/C_{40}/C_{45}/C_{50}/C_{55}$ = 18:11:13:28:20 for Y107A/F108A/H139A triple mutant). The quartet mutant Y107A/F108A/L135A/H139A still generated $C_{50}$ and $C_{55}$ as major products, but the largest product reached $C_{70}$ (Fig. 4*B*, *left*, and Table 3).

Simultaneous replacement of His$^{139}$ and Leu$^{135}$ on helix G with small Ala (L135A/H139A) produced $C_{30}$ as a major product and $C_{35}$ ($C_{30}/C_{35}$ = 73:18; Fig. 4*B*, *left*, and Table 3), whereas H139A single mutant produced $C_{25}$ and $C_{30}$ with $C_{30}$ as the major product ($C_{25}/C_{30}$ = 35:58). The double mutant Y107A/H139A generated $C_{40}$ as a major product ($C_{35}/C_{40}/C_{45}$ = 31:35:29; Fig. 4*B*, *left*, and Table 3); so did the F108A/H139A mutant ($C_{30}/C_{35}/C_{40}/C_{45}$ = 17:14:41:24).

Compared with the previously identified His$^{139}$ from mutagenesis study (22), Tyr$^{107}$ plays a more dominant role in determining product chain length. According to our crystal structure, the side chains of Tyr$^{107}$ and His$^{139}$ point into the tunnel interior from opposite $\alpha$-helixes (Fig. 4*A*, *top*, for the side view). From the top view (*bottom* of Fig. 4*A*), Tyr$^{107}$ and His$^{139}$ are well positioned to act as the "double floor" in the hollow tunnel, with Tyr$^{107}$ occupying the major part.

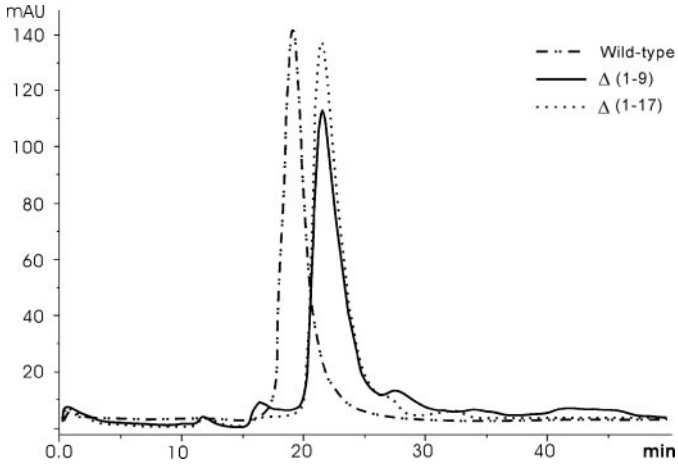


FIGURE 5. **Size column elution profile for determining the molecular species of wild type (*dotted and dashed line*), Δ(1–9) (*solid line*), and Δ(1–17) GGPPs (*dotted line*).** The wild type (1 and 0.2 mg/ml) was eluted as a dimer, but Δ(1–9) and Δ(1–17) GGPPs (1 mg/ml) were both eluted later, corresponding to a monomer. Only the wild type contained an extra five amino acids in the N terminus for tag cleavage (see "Experimental Procedures"). *mAU*, $10^{-3}$ × absorbance unit.

Ser$^{71}$ (the fourth amino acid before the first DD*XX*D in type-III GGPPs) is found in the upper portion of the active site crevice. When this residue was substituted with a larger residue of Tyr, S71Y GGPPs became an FPPs to produce $C_{15}$ FPP as a final product when using $C_{10}$

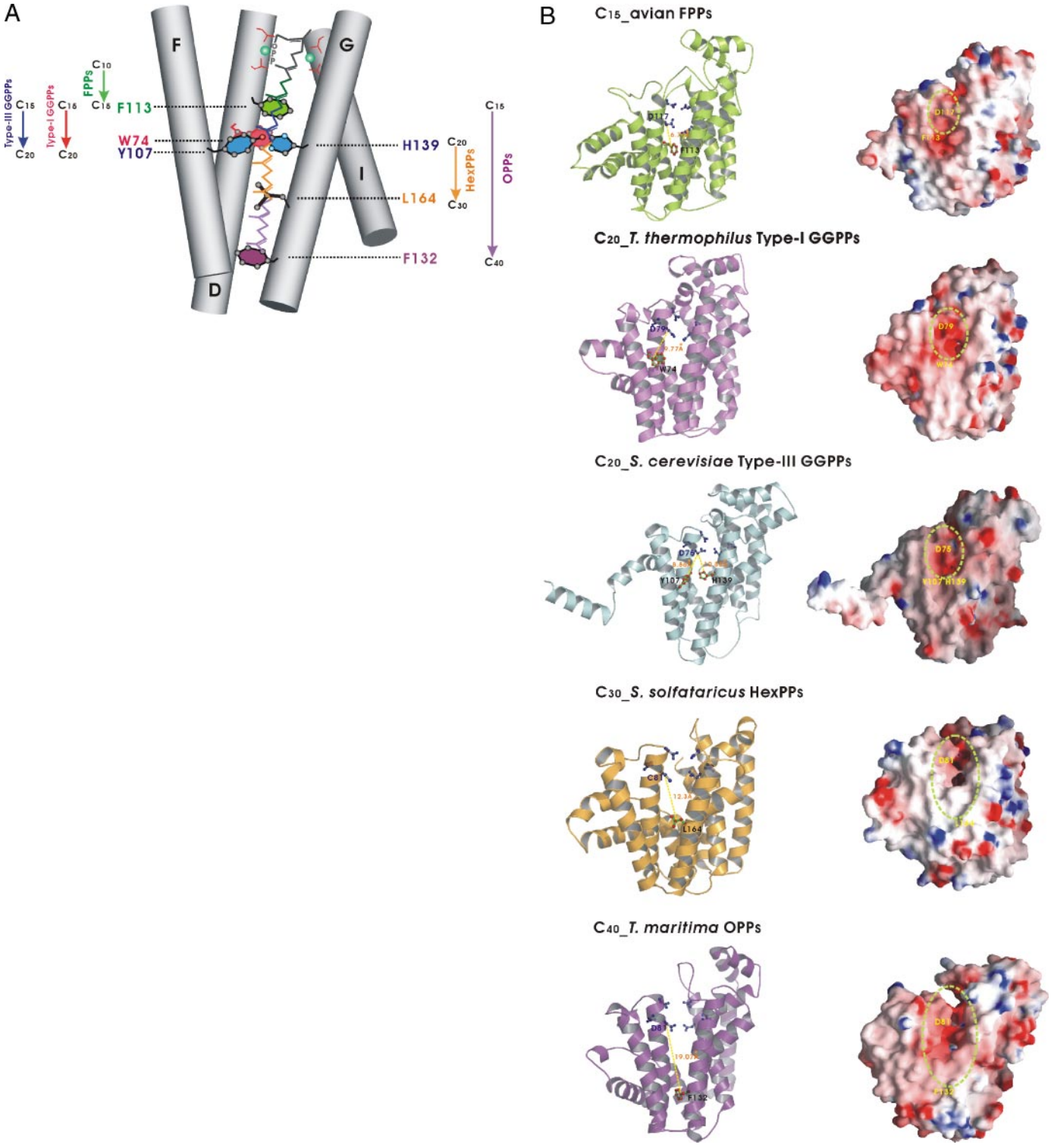


FIGURE 6. *A*, the cylindrical diagram showing the essential amino acids for product chain length determination from short-, medium-, and long-chain *trans*-prenyltransferases. The common active site crevice surrounded by helices D, F, and G with the bulky amino acids at the bottom for different *trans*-prenyltransferases are displayed. The key residues for chain length determination are Phe$^{113}$ for $C_{15}$-avian FPPs, Trp$^{74}$ for $C_{20}$-*T. thermophilus* type-I GGPPs, Tyr$^{107}$ and His$^{139}$ for $C_{20}$-*S. cerevisiae* type-III GGPPs, Leu$^{164}$ for $C_{30}$-*S. solfataricus* HexPPs, and Phe$^{132}$ for $C_{40}$-*T. maritima* OPPs, which are shown in *green*, *magenta*, *cyan*, *brown*, and *purple*, respectively. *B*, the active site cavity of avian FPPs, *T. thermophilus* type-I GGPPs, *S. cerevisiae* type-III GGPPs, *S. solfataricus* HexPPs, and *T. maritima* OPPs shown in *green*, *magenta*, *cyan*, *orange*, and *purple*, respectively. *Right*, in a *ribbon diagram*, the DD*XX*D motifs and the key residues for chain length determination are colored with *blue* and *gray* and *red* and *green* in a *ball-and-stick model*, respectively. The distances between the first Asp of the first DD*XX*D motif and Phe$^{113}$, Trp$^{74}$, Tyr$^{107}$/His$^{139}$, Leu$^{164}$, and Phe$^{132}$ are 6.77, 9.77, 8.86/10.85, 12.3, and 19.07 Å for FPPs, type-I GGPPs, type-III GGPPs, HexPPs, and OPPs, respectively. *Left*, electrostatic surface potential diagrams are *color-coded* from *red* to *blue* according to charge potential from −15 to 15 $k_BT$ (where $k_B$ Boltzmann constant and $T$ is temperature in Kelvin). The scope of the *green dashed circles* is restricted from the first Asp of the first DD*XX*D motif at the *top* to the bulky resides for chain length determination at the *bottom*. The active site tunnels of these *trans*-prenyltransferases are deeper and wider with the increased chain lengths of the ultimate products. These diagrams were generated with PyMol and GRASP (31).

GPP as the allylic substrate to react with one IPP (Fig. 4*B*, *right*). Under the same reaction conditions, wild-type enzyme catalyzes $C_{20}$ GGPP formation. This is consistent with the fourth amino acid before the first conserved DD*XX*D is at the correct position to shield the product of FPPs.

*The Molecular Ruler Mechanism*—As shown in the superimposed structures (Fig. 4*A*), $Tyr^{107}$ and $His^{139}$ of type-III GGPPs in fact reside at the almost identical spatial positions as $Trp^{74}$ (the fifth amino acid before the first DD*XX*D) and $Ala^{143}$ of type-I GGPPs, respectively. It can be expected that $Trp^{74}$ in type-I GGPPs controls the product chain length, although this has not yet been experimentally proven. This supports our hypothesis that $Tyr^{107}$ is the most critical amino acid in determining the product chain length of type-III GGPPs, since $Ala^{143}$ in the type-I GGPPs is a small amino acid that could not form a sufficient blockage, and the large $Trp^{74}$ (corresponding to $Tyr^{107}$ in type-III GGPPs) should play a dominant role in chain length determination. However, it is the "double-floor" by $Tyr^{107}$ and $His^{139}$ for the type-III GGPPs but a "single floor" for the type-I GGPPs. According to our data, a molecular ruler mechanism for regulation of product chain lengths of type-III GGPPs is proposed in Fig. 4*C*. When the large side chain of $Tyr^{107}$ is removed, the chain length of Y107A reaches $C_{40}$ at the position of $Phe^{108}$. By the double replacement of $Tyr^{107}$ and $Phe^{108}$ with small Ala, $C_{50}$ product blocked by $Val^{114}$ is generated. This chain elongation pathway is shown in *green* (Fig. 4*C*). On the other way of chain elongation shown in *purple*, a $C_{30}$ product blocked by $Leu^{135}$ is formed when $His^{139}$ is substituted with Ala. The mutants with removal of both $His^{139}$ and $Leu^{135}$ side chains still generated $C_{30}$, although $C_{40}$ can be reached, probably due to the partial blockage by $Tyr^{107}$ and $Phe^{131}$. On the other hand, mutation S71Y (the fourth amino acid prior to the first DD*XX*D motif of type-III GGPPs) created a blockage at the upper part of the active site crevice, and the mutant produced shorter $C_{15}$ FPP as the final product. Our data also suggest that the fourth amino acid before the first DD*XX*D is critical for the product chain length of $C_{15}$-FPPs, and the fifth amino acid is important for type-I $C_{20}$-GGPPs.

*The Function of N-terminal Helix A in Dimer Formation*—As shown in Fig. 3*A*, an unexpected feature in the crystal structure of type-III GGPPs is the orientation of the first N-terminal 17 amino acids, which protrude into and hold the other subunit. In order to elucidate the possible function of the N-terminal residues in dimer formation, the truncated Δ(1–9) and Δ(1–17) GGPPs, by deleting the first 9 amino acids (helix A) and 17 amino acids (helix A plus a loop), were constructed. According to size exclusion chromatography (Fig. 5), both of the truncated proteins were a monomer at 1 mg/ml, whereas the wild-type GGPPs formed a dimer at 1 mg/ml and lower protein concentration (0.2 mg/ml). Δ(1–9) showed a 325-fold lower activity compared with the wild type, but Δ(1–17) had no detectable activity. The CD spectra of wild type, Δ(1–9), and Δ(1–17) were similar (data not shown), indicating the deletion did not change the secondary structures. A subtle change in active site conformation must occur to account for the activity reduction for the truncated mutants.

## DISCUSSION

In this study, we solved the three-dimensional structure of type-III GGPPs from *S. cerevisiae*. This enzyme shares mostly the same folding as the available structures of avian FPPs (8), *Thermus thermophilus* type-I GGPPs,[4] *S. solfataricus* HexPPs (10), and *T. maritima* OPPs (9), although it shares only 22, 23, 22, and 25% sequence identity and 39, 40, 45, and 45% similarity, respectively, with those prenyltransferases. Compared with the type-I GGPPs, the superposition is better in helices D, E, F, G, H, I, K, and N with root mean square deviation = 1.53 Å for the fitting of 648 Cα atoms of the dimeric type-III GGPPs. The most different feature in the crystal structures of type-III GGPPs is the location of the N-terminal helix A, which embraces the other subunit. Deletion of the first nine amino acids (the helical region) is sufficient to dissociate the dimer into monomer, supporting the importance of the N-terminal helix (helix A) in dimer formation. For other *trans*-prenyltransferases, however, their N-terminal helix is located within its own subunit, and the dimer is formed mainly through the interfacial interaction from helices E and F (*e.g.* OPPs and HexPPs) without the contribution from this N-terminal helix. Our data also indicate the monomer is remarkably less active, although each monomer contains a separate active site. To our knowledge, this is the first case in which a dimeric prenyltransferase can be disrupted into a monomer by a mutation (a deletion here). The structure of the monomer will be further studied.

It was mysterious that unlike the other short-chain *trans*-prenyltransferases, the fourth and the fifth amino acids prior to the first DD*XX*D motif (the FPP binding site) of type-III GGPPs are not large amino acids. On the basis of crystal structure, we have replaced some large amino acids at the bottom of the active site crevice to test their role in chain length determination. Some mutants show different kinetic parameters (smaller FPP $K_m$ and larger IPP $K_m$ values) compared with the wild type. This was also observed for the mutants of FPPs that became GGPPs by random mutagenesis (18). The calculation of kinetic parameters followed our previous reports on the kinetics of prenyltransferases (32, 33). Although the prenyltransferases and their mutants produced various products, all of the products can be extracted with 1-butanol. We also demonstrated that each IPP condensation step has a similar rate constant (32, 33), so the $k_{cat}$ could represent the IPP condensation rate constant. Here the $k_{cat}$ values of the GGPP mutants, which made multiple products, were calculated from the initial rates of the substrate depletion and should represent the steady-state activity for forming their favorite elongated products due to the removal of the blockage by replacing the large amino acids at the bottom of active crevice with a small Ala. The final products of the mutants described here were also generated under sufficient quantities of FPP and [$^{14}$C]IPP.

From the structure and mutagenesis analysis presented above, we discovered an important residue, $Tyr^{107}$, in addition to the previously identified $His^{139}$ (22), which seals the bottom of the active site crevice to control the product chain length. The substitution of the large $Tyr^{107}$ and $His^{139}$ with a smaller Ala removes the "floor" of the crevice, thereby allowing for the formation of longer chain length products. This kind of "floor" has been found in other *trans*-prenyltransferases and also the protein prenyltransferases, which catalyze the transfer of farnesyl or geranylgeranyl group into the Cys of the CAA*X* motif (where A is an aliphatic amino acid) of the protein substrates, indicating a common mechanism of molecular ruler (34). However, distinct from other prenyltransferases, the "floor" in type-III GGPPs is a "double floor" constituted of two large amino acids. As illustrated in Fig. 6*A*, the "floor" is located at helix D for FPPs and type-I GGPPs but in helix F and G for type-III GGPPs and in helix G for OPPs and HexPPs. The key residues for chain length determination are $Phe^{113}$ for $C_{15}$-avian FPPs, $Trp^{74}$ for $C_{20}$-*T. thermophilus* type-I GGPPs, $Tyr^{107}$ and $His^{139}$ for $C_{20}$-*S. cerevisiae* type-III GGPPs, $Leu^{164}$ for $C_{30}$-*S. solfataricus* HexPPs, and $Phe^{132}$ for $C_{40}$-*T. maritima* OPPs (Fig. 6*A*). According to these known three-dimensional structures, the distances between the first Asp of the first DD*XX*D motif and $Phe^{113}$, $Trp^{74}$, $Tyr^{107}$/$His^{139}$, $Leu^{164}$, and $Phe^{132}$ for FPPs, Type-I GGPPs, Type-III GGPPs, HexPPs, and OPPs, respectively, are increasing with the longer product chain lengths (Fig. 6*B*, *left*). The areas inside the *circled cavity* of these prenyltransferases are also larger

with the increasing size of the ultimate products catalyzed by these enzymes (Fig. 6*B*, *right*).

Type-I GGPPs utilizes $Trp^{74}$ at helix D, and type-III GGPPs uses $Tyr^{107}$ and $His^{139}$ at different helices, F and G, to control the product chain length although both enzymes make the same product. Based on the different strategies in sealing the active site, our data suggest that type-I GGPPs may evolve from FPPs by shifting the large amino acid from the fourth to fifth position before the first DD*XX*D to shield the product, whereas type-III GGPPs and medium-chain and long-chain prenyltransferases may evolve from the common ancestor. Together, the data significantly extend our understanding on the biosynthesis of short-chain polyprenyl molecules.

*Acknowledgment—The synchrotron data collection was conducted at the Biological Crystallography Facilities (Taiwan Beamline BL12B2 at SPring-8).*

## REFERENCES


1. Kellogg, B. A., and Poulter, C. D. (1997) *Curr. Opin. Chem. Biol.* **1,** 570–578
2. Ogura, K., Koyama, T., and Sagami, H. (1997) *Subcell. Biochem.* **28,** 57–88
3. Liang, P. H., Ko, T. P., and Wang, A. H.-J. (2002) *Eur. J. Biochem.* **269,** 3339–3354
4. Sacchettini, J. C., and Poulter, C. D. (1997) *Science* **277,** 1788–1789
5. Wang, K., and Ohnuma, S. (1999) *Trends. Biochem. Sci.* **24,** 445–451
6. Szkopinska, A., and Plochocka, D. (2005) *Acta Biochim. Pol.* **52,** 45–55
7. Chen, A., Kroon, P. A., and Poulter, C. D. (1994) *Protein Sci.* **3,** 600–607
8. Tarshis, L. C., Yan, M., Poulter, C. D., and Sacchettini, J. C. (1994) *Biochemistry* **33,** 10871–10877
9. Guo, R. T., Kuo, C. J., Chou, C. C., Ko, T. P., Shr, H. L., Liang, P. H., and Wang, A. H.-J. (2004) *J. Biol. Chem.* **279,** 4903–4912
10. Sun, H. Y., Ko, T. P., Kuo, C. J., Guo, R. T., Chou, C. C., Liang, P. H., and Wang, A. H.-J. (2005) *J. Bacteriol.* **187,** 8137–8148
11. Fujihashi, M., Zhang, Y.-W., Higuchi, Y., Li, X.-Y., Koyama, T., and Miki, K. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 4337–4342
12. Chang, S. Y., Ko, T. P., Chen, A. P., Wang, A. H.-J., and Liang, P. H. (2004) *Protein Sci.* **13,** 971–978
13. Guo, R. T., Ko, T. P., Chen, A. P.-C., Kuo, C. J., Wang, A. H.-J., and Liang, P. H. (2005) *J. Biol. Chem.* **280,** 20762–20774
14. Ogura, K., and Koyama, T. (1998) *Chem. Rev.* **98,** 1263–1276
15. Ko, T. P., Chen, Y. K., Robinson, H., Tsai, P. C., Gao, Y. G., Chen, A. P., Wang, A. H.-J., and Liang, P. H. (2001) *J. Biol. Chem.* **276,** 47474–47482
16. Hosfield, D. J., Zhang, Y., Dougan, D. R., Broun, A., Tari, L. W., Swanson, R. V., and Finn, J. (2004) *J. Biol. Chem.* **279,** 8526–8529
17. Tarshis, L. C., Proteau, P. J., Kellogg, B. A., Sacchettini, J. C., and Poulter, C. D. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 15018–15023
18. Ohnuma, S., Nakazawa, T., Hemmi, H., Hallberg, A. M., Koyama, T., Ogura, K., and Nishino, T. (1996) *J. Biol. Chem.* **271,** 10087–10095
19. Ohnuma, S., Narita, K., Nakazawa, T., Ishida, C., Takeuchi, Y., Ohto, C., and Nishino, T. (1996) *J. Biol. Chem.* **271,** 30748–30754
20. Guo, R. T., Kuo, C. J., Ko, T. P., Chou, C. C., Liang, P. H., and Wang, A. H.-J. (2004) *Biochemistry* **43,** 7678–7686
21. Jiang, Y., Proteau, P., Poulter, D., and Ferro-Novick, S. (1995) *J. Biol. Chem.* **270,** 21793–21799
22. Hemmi, H., Noike, M., Nakayama, T., and Nishino, T. (2003) *Eur. J. Biochem.* **270,** 2186–2194
23. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis* **18,** 2714–2723
24. Guerrero, S. A., Hecht, H. J., Hofmann, B., Biebl, H., and Singh, M. (2001) *Appl. Microbiol. Biotechnol.* **56,** 718–723
25. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
26. Terwilliger, T. C. (2003) *Methods Enzymol.* **374,** 22–37
27. Terwilliger, T. C. (2001) *Acta Crystallogr. Sect. D* **57,** 1763–1775
28. McRee, D. E. (1999) *J. Struct. Biol.* **125,** 156–165
29. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., GrosseKunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D* **54,** 905–921
30. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
31. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins Struct. Funct. Genet.* **11,** 281–296
32. Kuo, T. H., and Liang, P. H. (2002) *Biochim. Biophys. Acta* **1599,** 125–133
33. Pan, J. J., Chiu, S. T., and Liang, P. H. (2000) *Biochemistry* **39,** 10936–10942
34. Long, S. B., Casey, P. J., and Beese, L. S. (1998) *Biochemistry* **37,** 9612–9618