Structure and Mechanism of an Arabidopsis Medium/Long-Chain-Length Prenyl Pyrophosphate Synthase

Fu-Lien Hsieh\textsuperscript{2,3}, Tao-Hsin Chang\textsuperscript{2,3}, Tzu-Ping Ko, and Andrew H.-J. Wang\*  
Institute of Biological Chemistry (F.-L.H., T.-H.C., T.-P.K., A.H.-J.W.) and Core Facilities for Protein Production and X-Ray Structural Analysis (T.-P.K., A.H.-J.W.), Academia Sinica, Taipei 115, Taiwan; and Institute of Biochemical Sciences, National Taiwan University, Taipei 106, Taiwan (F.-L.H., A.H.-J.W.)

Prenyltransferases (PTSs) are involved in the biosynthesis of terpenes with diverse functions. Here, a novel PTS from Arabidopsis (Arabidopsis thaliana) is identified as a trans-type polyprenyl pyrophosphate synthase (AtPPPS), which forms a trans-double bond during each homoallylic substrate condensation, rather than a homomeric C\textsubscript{10}-geranyl pyrophosphate synthase as originally proposed. Biochemical and genetic complementation analyses indicate that AtPPPS synthesizes C\textsubscript{25} to C\textsubscript{10} medium/long-chain products. Its close relationship to other long-chain PTSs is also uncovered by phylogenetic analysis. A mutant of contiguous surface polar residues was produced by replacing four charged surface amino acids with alanines to facilitate the crystallization of the enzyme. The crystal structures of AtPPPS determined here in apo and ligand-bound forms further reveal an active-site cavity sufficient to accommodate the medium/long-chain products. The two monomers in each dimer adopt different conformations at the entrance of the active site depending on the binding of substrates. Taken together, these results suggest that AtPPPS is endowed with a unique functionality among the known PTSs.

Over 55,000 terpenes (isoprenoids), the largest class of plant metabolites, have been identified to be involved in numerous vital biological processes, including growth, development, and response to environmental stresses (Fig. 1; Pichersky et al., 2006; Gershenzon and Dudareva, 2007). Terpenes also have considerable applications as pharmaceuticals, fragrances, and nutritional supplements (Kirby and Keasling, 2009). These diverse compounds are derived from the rather simple universal precursors of linear prenyl pyrophosphates, ranging from C\textsubscript{10} to C\textsubscript{10,000} in the number of carbon atoms, which are synthesized by groups of conserved prenyltransferases (PTSs; Kellogg and Poulter, 1997; Liang et al., 2002). The various chain lengths of these linear prenyl pyrophosphates, reflecting their distinctive physiological functions (Fig. 1), in general are determined by the highly developed active site of PTSs via condensation reactions of allylic substrates (dimethylallyl diphosphate [C\textsubscript{5}-DMAPP], geranyl pyrophosphate [C\textsubscript{10}-GPP], farnesyl pyrophosphate [C\textsubscript{15}-FPP], geranylgeranyl pyrophosphate [C\textsubscript{20}-GGPP]) with corresponding number of isopentenyl pyrophosphates (C\textsubscript{5}-IPP; homoallylic substrate; Liang, 2009). PTSs are generally classified into cis- and trans-type enzymes on the basis of the stereochemistry of double-bound formation between the homoallylic substrate C\textsubscript{5}-IPP and the allylic substrates (Liang et al., 2002). In addition, the two types of PTSs not only differ completely in their primary amino acid sequences and tertiary structures but also utilize distinct mechanisms for substrate binding and catalysis, despite sharing the allylic and homoallylic substrates (Liang, 2009). The sequences of the trans-type PTSs generally have less than 30% conserved amino acids, although they possess a similar protein fold as well as two functional Asp-rich motifs, DD(X)\textsubscript{n}D (in which X encodes any amino acid and \textit{n} = 2 or 4), reflecting the required diversity to achieve their specific condensation reactions (Kellogg and Poulter, 1997; Ogura and Koyama, 1998; Liang et al., 2002).

In plants, C\textsubscript{10}-GPP synthase (GPPS), which catalyzes the condensation of C\textsubscript{5}-DMAPP with C\textsubscript{5}-IPP into C\textsubscript{10}-GPP, is a key enzyme in the C\textsubscript{10}-monoterpene biosynthesis of plant volatiles to attract pollinators, mediators in interplant communication, and secondary metabolites for defense (Kessler and Baldwin, 2001; Gershenzon and Dudareva, 2007). Intriguingly,

\textsuperscript{1} This work was supported by Academia Sinica and the Core Facility for Protein Production and X-Ray Structural Analysis (grant no. NSC97-3112-B-001-035-B4 to A.H.-J.W.). The Protein Crystallography Facility of the National Synchrotron Radiation Research Center is supported by the National Research Program for Genomic Medicine.

\textsuperscript{2} These authors contributed equally to the article.

\textsuperscript{3} Present address: Division of Structural Biology, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, United Kingdom.

\textsuperscript{*} Corresponding author; e-mail ahjwang@gate.sinica.edu.tw.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Andrew H.-J. Wang (ahjwang@gate.sinica.edu.tw).

[\textsubscript{W}] The online version of this article contains Web-only data.

[\textsubscript{OA}] Open Access articles can be viewed online without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.110.168799
enzymes possessing GPPS activity have been identified to be either homomeric or heteromeric proteins (Burke et al., 1999; Bouvier et al., 2000; Burke and Croteau, 2002; Tholl et al., 2004; Van Schie et al., 2007; Schmidt and Gershenzon, 2008; Orlova et al., 2009; Wang and Dixon, 2009; Schmidt et al., 2010), in contrast to most homomeric PTSs (Liang, 2009). We previously reported the structure of mint (**Mentha piperita**) heterotetrameric GPPS, composed of two active catalytic large subunits (LSU) and two regulatory non-catalytic small subunits (SSU; Chang et al., 2010), which is distinct from known homomeric PTSs such as C15-FPP synthase (FPPS) and C20-GGPP synthase (GGPPS; Chang et al., 2006; Kavanagh et al., 2006). The LSU is closely akin to the subunit of homomeric PTSs but lacks enzymatic activity on its own, and it requires the interactions with SSU to achieve a functional assembly (Kloer et al., 2006; Chang et al., 2010). The product fidelity of heterotetrameric GPPS is regulated via a regulatory loop in the SSU, which controls the product release from the catalytic LSU (Hsieh et al., 2010).

The homomeric Arabidopsis (**Arabidopsis thaliana**) GPPS (AtGPPS) has been used as a model target for assessing the GPPS activity in angiosperms in the past decade (Bouvier et al., 2000; Lange and Ghassemian, 2003; Van Schie et al., 2007; Orlova et al., 2009; Wang and Dixon, 2009; Schmidt et al., 2010), in contrast to most homomeric PTSs (Liang, 2009). We previously reported the structure of mint (**Mentha piperita**) heterotetrameric GPPS, composed of two active catalytic large subunits (LSU) and two regulatory non-catalytic small subunits (SSU; Chang et al., 2010), which is distinct from known homomeric PTSs such as C15-FPP synthase (FPPS) and C20-GGPP synthase (GGPPS; Chang et al., 2006; Kavanagh et al., 2006). The LSU is closely akin to the subunit of homomeric PTSs but lacks enzymatic activity on its own, and it requires the interactions with SSU to achieve a functional assembly (Kloer et al., 2006; Chang et al., 2010). The product fidelity of heterotetrameric GPPS is regulated via a regulatory loop in the SSU, which controls the product release from the catalytic LSU (Hsieh et al., 2010).

The homomeric Arabidopsis (**Arabidopsis thaliana**) GPPS (AtGPPS) has been used as a model target for assessing the GPPS activity in angiosperms in the past decade (Bouvier et al., 2000; Lange and Ghassemian, 2003; Van Schie et al., 2007; Orlova et al., 2009). A recent study showed that Arabidopsis also expresses a heteromeric GPPS distinct from the homomeric type in their subunit compositions and sequence homology (Supplemental Fig. S1; Wang and Dixon, 2009). This discovery raises two questions. Why does Arabidopsis need to develop two separate types of GPPS? Does Arabidopsis benefit from possessing both types of enzymes? Furthermore, no structure of homomeric GPPS is available so far, and its enzymatic regulation mechanism is difficult to predict from the currently known structures of homomeric PTSs and heteromeric GPPS (Chang et al., 2006, 2010; Hsieh et al., 2010).

In an effort to address the above questions, the putative homomeric AtGPPS was cloned, expressed, and characterized by x-ray crystallography in combination with biochemical and genetic complementation analyses. Surprisingly, our results suggest that this enzyme from Arabidopsis is actually a polyprenyl pyrophosphate synthase (AtPPPS), an unusual PTS generating multiple products with medium/long chain length ranging from C25 to C45, rather than a GPPS as reported previously (Bouvier et al., 2000). Hence, we rename this homomeric AtGPPS to AtPPPS based on its enzymatic activity. These results should provide significant insights into the plant medium/long-chain PTSs and encourage further study to reevaluate the enzymatic functions and physiological roles of angiosperm GPPSs.

**RESULTS**

**Phylogenetic Relationship**

By sequence analyses, homomeric GPPSs in plants have been classified into two groups, one from gym-
nosperms and the other from angiosperms (Bouvier et al., 2000; Burke and Croteau, 2002; Van Schie et al., 2007; Schmidt and Gershenzon, 2008; Schmidt et al., 2010). AtPPPS has high identity, reaching about 70%, to other angiosperm homomeric GPPSs (Supplemental Fig. S2; Supplemental Table S1). However, detailed sequence alignment suggests that AtPPPS is actually similar to the long-chain PTSs (approximately 50% identity), which generate products beyond C_{35} (Fig. 2; Supplemental Fig. S3; Supplemental Table S1).

To resolve this controversy, we further analyzed the phylogenetic relationships between AtPPPS and other plant PTSs (Fig. 3). The phylogenetic tree shows that AtPPPS is evolutionarily more closely related to the long-chain PTSs (e.g. C_{35}-solanesyl pyrophosphate synthase [SPPS] and C_{30}-decaprenyl pyrophosphate synthase) than to the short-chain PTSs (e.g. GGPPS and FPPS). An exception, naturally, is the grouping with other angiosperm GPPSs. Previous studies suggest that the active site of PTSs has been exquisitely developed to control their substrate and product specificities (Ohnuma et al., 1996; Tarshis et al., 1996; Guo et al., 2004; Sun et al., 2005; Chang et al., 2006). Therefore, the specifically conserved amino acid sequence of PTSs has been used to predict the chain length of the final product (Kellogg and Poulter, 1997; Ogura and Koyama, 1998; Liang, 2009). Our results imply that AtPPPS and long-chain PTSs may have similar functions.

### Catalytic Activity

To investigate the genuine enzymatic function of AtPPPS, we expressed and purified the protein in a pseudomature form by removing the plastid targeting sequence (Supplemental Table S2). Its activity was subsequently measured using four allylic substrates (C_{5}-DMAPP, C_{10}-GPP, C_{15}-FPP, and C_{20}-GGPP) in the presence of C_{5}-[{^14}C]IPP. Surprisingly, the reaction yielded a broad spectrum of multiple products ranging from C_{35} to C_{45} (Fig. 4). Except for C_{5}-DMAPP, AtPPPS can recognize the other three allylic substrates and react them with C_{5}-[{^14}C]IPP, resulting in similar multiple product distribution patterns having C_{35} as the major product (Fig. 4; Table I). In the subsequent time-course assay, multiple products were detected simultaneously (i.e. not sequentially) in the chain elongation process (Supplemental Fig. S4). This observation further indicates that the products of medium/long chain lengths as synthesized by AtPPPS (Fig. 4)

**Figure 2.** Structure-based multiple sequence alignment. The alignment of the deduced amino acid sequences of Arabidopsis PTSs includes PPS (AtPPPS, AT2G43630), SPPS isoform 1 (AtSPPS1; AT1G78510), SPPS isoform 2 (AtSPPS2; AT1G17050), GGPPS isoform 1 (AtGGPPS1; AT1G49530), GGPPS isoform 11 (AtGGPPS11; AT4G36810), FPPS isoform 1 (AtFPPS1; AT5G47770), and FPPS isoform 2 (AtFPPS2; AT4G17190). Regions of AtPPPS corresponding to the a-helices are denoted by purple cylinders. Identical and similar amino acid residues are shaded in black and gray, respectively. Dashes indicate the sequence gaps introduced to optimize the amino acid sequence alignment. The conserved functional motifs, DD(XX)D, are boxed in cyan. All sequences presented here have the N-terminal signal peptides omitted.
are not a result of the longer reaction time. Additionally, our results also imply that the released products having longer chain lengths than C25 would have a lower frequency of rebinding to the active site for further product elongation. Based on its product distribution, we rename this enzyme as a PPPS. Intriguingly, most PTSs are monofunctional enzymes that exclusively synthesize single chain-length products (Tarshis et al., 1996; Ogura and Koyama, 1998; Guo et al., 2004; Sun et al., 2005; Chang et al., 2006; Kloer et al., 2006; Hsiao et al., 2008). The dimerization interface is mainly contributed by the respective helices F and G (Fig. 5A). Each subunit is composed of 16 antiparallel \( \alpha \)-helices that surround the active site, with the two conserved DD(\( X \))nD motifs facing each other on helices D and J (Fig. 5A).}

**Overall Structure and Active Site**

To further understand its function, we determined the crystal structure at 2.6-\( \AA \) resolution of the wild-type AtPPPS in its apo form, denoted WT-AtPPPS (Fig. 5A; Table II). While WT-AtPPPS shares less than 30% sequence identity with other PTSs, it clearly adopts the conserved all-\( \alpha \)-helix fold of PTSs (Tarshis et al., 1996; Liang, 2009). A stable homodimer was also detected by gel filtration analysis in a protein concentration-independent manner (Supplemental Fig. S5), consistent with previous findings that most PTSs exist as homodimers under physiological conditions (Guo et al., 2004; Sun et al., 2005; Chang et al., 2006; Kloer et al., 2006; Hsiao et al., 2008). The active-site region of WT-AtPPPS is embedded with highly conserved catalytic amino acid residues to be implemented in its enzymatic reaction. The consensus catalytic mechanism of PTSs has been demonstrated to be a set of sequential ionization-condensation-elimination reactions: the homoallylic substrate attacks the allylic substrate, which forms a carbocation intermediate by removing its inorganic pyrophosphate group, with concomitant removal of a proton from the adduct...
The binding of the allylic substrate is mainly contributed by the DD(X)_nD motif, Mg^{2+} ions, and the associated water molecules. The homoallylic substrate is bound in a positively charged pocket surrounded by residues Arg-54, His-100, and Arg-117.

To further unveil details in the product elongation region of AtPPPS, we sought to solve the structure of WT-AtPPPS in complex with its ligand. Despite extensive efforts, WT-AtPPPS failed to crystallize in a ligand-bound form. Crystallization is generally believed to involve the free energy change upon assembly of the protein molecules into a crystal lattice. Hence, replacing high-entropy polar amino acids on the protein surface with Ala residues has been used to reduce the entropy, in favor of crystal contact formation (Derewenda, 2004; Goldschmidt et al., 2007). To enhance the crystallizability of AtPPPS, we used the Surface Entropy Reduction prediction server (http://nihserver.mbi.ucla.edu/SER/) to analyze the primary sequence and located several flexible polar residues with high entropy values. We then constructed several mutants according to the prediction and finally obtained a mutant of contiguous surface polar residues, denoted SM-AtPPPS, by replacing four residues (Glu-178, Gln-179, Glu-281, and Lys-282; Fig. 2) on the protein surface with Ala residues to facilitate crystal lattice formation. Although the mutations reduced the enzymatic activity slightly when C10-GPP was used as the allylic substrate, the overall functional activity and product distribution pattern of SM-AtPPPS remained comparable with those of WT-AtPPPS (Supplemental Fig. S6; Table I).

The SM-AtPPPS crystal solved at 2.65-Å resolution contains an octamer (chain A–H) as its asymmetric unit, comprising four identical dimers related by three orthogonal noncrystallographic 2-fold axes and expressing a tetrahedral 222 symmetry (Fig. 5B; Table II). Those four mutations generate additional intermolecular crystal contacts both within the asymmetric unit and between different octamers in the unit cell (Supplemental Fig. S7). Although the two monomers in each dimer adopt distinct conformations, a dimer as the basic assembly unit is consistent with WT-AtPPPS and other known structures of PTSs (Tarshis et al., 1996; Guo et al., 2004; Hosfield et al., 2004; Sun et al., 2005; Chang et al., 2006; Kavanagh et al., 2006; Kloer et al., 2006). Judging by its gel filtration chromatography profile, SM-AtPPPS also exists as a dimer in solution (data not shown). The electron density map clearly shows that a different monomer is bound with either Mg^{2+} ions, an inactive C5-IPP analog isopentenyl thiolopyrophosphate (C5-IPSP), and C15-FPP or Mg^{2+} ions, inorganic pyrophosphate, and C15-FPP in its active site (Fig. 5C; Supplemental Fig. S8). The crystal structure of homodimeric Sinapis alba GGPPS was found to contain different ligands bound to different subunits as well (Kloer et al., 2006). The N-terminal residues and some surface loops were disordered.

Further structural analyses show that the aliphatic tail of C15-FPP is located in a large hydrophobic cleft starting with the active site cavity and connecting with the elongation cavity (EC) adjacent to the dimer interface (Fig. 6A). Previous studies suggest that the regulation of product chain length specificity and substrate selectivity is determined by the size of the tunnel-shaped cleft of PTSs, since the product elongation extends along the EC tunnel (Ohnuma et al., 1996; Tarshis et al., 1996; Guo et al., 2004; Sun et al., 2005; Chang et al., 2006). The bound substrate models and the critical amino acid residues in the active site of AtPPPS are similar to those of all other known PTSs, as verified by the biochemical and crystallographic studies (Liang et al., 2002; Liang, 2009). The residues located on

Table 1. Product distributions of WT-AtPPPS, SM-AtPPPS, and AtPPPS(I99F/V162F) in the presence of a range of allylic substrates

The radioactivity of each product was normalized by the numbers of C5-[14C]IPP incorporated. Further experimental details can be found in "Materials and Methods.”
helices D, E, and G that surround the EC tunnel are highly conserved among long-chain PTSs (Fig. 2).

Overall, these structural studies of AtPPPS allow an unambiguous identification of a cavity to accommodate longer products beyond C₁₀-GPP (Figs. 5C and 6A; Supplemental Fig. S8).

Comparison of the apo Form and the Ternary Complex

Superposition of WT-AtPPPS and SM-AtPPPS allows the identification of three notable regions with significant conformational changes (Supplemental Fig. S9). First, the disordered region of helices D to F shows ligand-binding-induced conformational changes to act as a gate for substrate entry and product release, consistent with previous studies (Sun et al., 2005; Kloer et al., 2006). Second, the region connecting helices A and C has extensive conformational change. Helix B becomes an ordered structure when the C₅-IPP substrate is bound (Supplemental Fig. S9). Therefore, this highly mobile region may be induced to become ordered by the binding of C₅-IPP. Third, the orientation of the first N-terminal helix A protrudes into the top of the other subunit and seems to be involved in regulating the conformational change during the catalytic reaction (Supplemental Figs. S8 and S9). This is in accordance with an alternating catalytic mechanism in the dimer (i.e. when one subunit is in action, the other subunit is empty in its active site; Sun et al., 2005; Kloer et al., 2006). The alternating mechanism is also reflected in the asymmetric binding of different ligands to different protein subunits of the homodimeric enzyme. This kind of enzymatic regulation mechanism may be used to control the steps of substrate entry and product release. Hence, these observations of the crystal structure further explain why the basic functional unit of PTSs is a dimer instead of a monomer.
The Mechanism of Product Elongation

To investigate how the EC tunnel accommodates the long-chain products, two hydrophobic residues, Ile-99 on helix D and Val-162 on helix G, in the middle of the EC tunnel are substituted by larger Phe residues to serve as a new floor to block the product chain elongation beyond C20-GGPP (Fig. 6A). AtPPPS(I99F/V162F) generates C20-GGPP as the major product, plus a small amount of farnesylgeranyl pyrophosphate (C25-FGPP; Fig. 6B; Table I). The shorter product chain length is consistent with the reduced size of the EC tunnel as a result of the mutations. As expected, the mutant enzyme showed lower activity to recognize the C20-GGPP as the allylic substrate to implement the chain-elongation reaction, while the activity for C15-FPP was largely unaffected (Fig. 6B). The altered active-site structure might have an unfavorable effect in retaining the short-chain intermediate; therefore, the enzymatic activity for reacting C10-GPP with C5-IPP was reduced.

Although C20-GGPP from the AtPPPS(I99F/V162F) mutant can be detected by in vitro assay, it remains to be validated under in vivo conditions. The genetic complementation method (Zhu et al., 1997; Engprasert et al., 2004; Chang et al., 2010) was employed to investigate whether this mutant exhibits the GGPPS activity in vivo. The \textit{crt} gene cluster of \textit{Pantoea ananatis}, responsible for biosynthesis of the yellow pigment carotenoid except \textit{crtE}, which encodes GGPPS, was constructed into pACCAR25\textsubscript{D}crtE (Fig. 6C; Zhu et al., 1997). \textit{Escherichia coli} does not possess an intrinsic GGPPS gene; therefore, the \textit{E. coli} cells carrying the pACCAR25\textsubscript{D}crtE vector and the empty vector pET-16 cannot accumulate the yellow pigment (Fig. 6D). In contrast, transformants harboring pACCAR25\textsubscript{D}crtE and pET-32 that contains the \textit{Saccharomyces cerevisiae} GGPPS or AtPPPS(I99F/V162F) gene showed a visible yellow color, and the extracted pigments were further measured by optical absorption (Fig. 6D; Supplemental Fig. S10). Taken together, these results confirm that

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Apo Form (WT-AtPPPS)</th>
<th>Ternary Complex (SM-AtPPPS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ligands</td>
<td>–</td>
<td>Mg\textsuperscript{2+}/C5-IPSP/C15-FPP</td>
</tr>
<tr>
<td>Data collection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>F\textsubscript{222}</td>
<td>P\textsubscript{61}</td>
</tr>
<tr>
<td>Unit cell dimensions (Å)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>148.51</td>
<td>115.96</td>
</tr>
<tr>
<td>b</td>
<td>150.13</td>
<td>115.96</td>
</tr>
<tr>
<td>c</td>
<td>176.11</td>
<td>385.87</td>
</tr>
<tr>
<td>Resolution (Å\textsuperscript{a})</td>
<td>30.00-2.60 (2.69-2.60)</td>
<td>30.00-2.65 (2.74-2.65)</td>
</tr>
<tr>
<td>No. of unique reflections\textsuperscript{b}</td>
<td>30,232 (2,979)</td>
<td>83,662 (8,206)</td>
</tr>
<tr>
<td>Redundancy\textsuperscript{c}</td>
<td>3.9 (3.5)</td>
<td>3.7 (2.6)</td>
</tr>
<tr>
<td>Completeness (%)\textsuperscript{d}</td>
<td>99.5 (98.8)</td>
<td>98.8 (97.4)</td>
</tr>
<tr>
<td>Average I/I(0)\textsuperscript{e}</td>
<td>18.3 (2.5)</td>
<td>28.3 (2.1)</td>
</tr>
<tr>
<td>R\textsubscript{merge} (%)\textsuperscript{f}</td>
<td>7.1 (48.9)</td>
<td>4.8 (49.5)</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of reflections\textsuperscript{g}</td>
<td>28,230</td>
<td>79,172</td>
</tr>
<tr>
<td>R\textsubscript{work}/R\textsubscript{free} (%)\textsuperscript{h}</td>
<td>20.6/25.6</td>
<td>22.2/27.4</td>
</tr>
<tr>
<td>r.m.s. deviations\textsuperscript{i}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.685</td>
<td></td>
</tr>
<tr>
<td>B value (Å\textsuperscript{c}/no. of atoms)</td>
<td>48.0/4,586</td>
<td>32.2/19,484</td>
</tr>
<tr>
<td>Protein</td>
<td>48.0/4,586</td>
<td>32.2/19,484</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>–</td>
<td>47.9/248</td>
</tr>
<tr>
<td>Water</td>
<td>53.5/373</td>
<td>47.8/721</td>
</tr>
<tr>
<td>Ramachandran plot (%)\textsuperscript{j}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Most favored regions</td>
<td>93.1</td>
<td>92.8</td>
</tr>
<tr>
<td>Allowed regions</td>
<td>6.5</td>
<td>6.4</td>
</tr>
<tr>
<td>Generously allowed</td>
<td>0.4</td>
<td>0.7</td>
</tr>
<tr>
<td>Disallowed</td>
<td>0</td>
<td>0.1</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Values in parentheses correspond to the highest resolution data shell. \textsuperscript{b}R\textsubscript{merge} = \sum \langle I(h) \rangle - \langle I(h) \rangle, where \langle I(h) \rangle is the measured diffraction intensity and the summation includes all observations. \textsuperscript{c}All positive reflections were used in the refinement. \textsuperscript{d}R\textsubscript{work} is the R factor = \sum (F_{o} - F_{c})/\sum F_{o}, where \textit{F}_{o} and \textit{F}_{c} are observed and calculated structure factors, respectively. \textsuperscript{e}R\textsubscript{free} is the R factor calculated using 5% of the data that were excluded from the refinement. \textsuperscript{f}The r.m.s. deviation is the root mean square deviation from ideal geometry of protein. \textsuperscript{g}Ramachandran plots were generated with PROCHECK (Laskowski et al., 1993). For the ternary complex of SM-AtPPPS, three residues, Asn-123 (chain D), Leu-30 (chain D), and Leu-30 (chain F), are located in the disallowed region.
the double mutant produces C$_{20}$-GGPP as the major product both in vitro and in vivo (Fig. 6, B and D).

**DISCUSSION**

The plant GPPS-encoding genes have been identified in both gymnosperms and angiosperms (Burke et al., 1999; Bouvier et al., 2000; Burke and Croteau, 2002; Tholl et al., 2004; Van Schie et al., 2007; Schmidt and Gershenzon, 2008; Wang and Dixon, 2009; Schmidt et al., 2010). Interestingly, enzymes exhibiting this catalytic activity can be further classified into homomeric and heteromeric proteins. In contrast to the studies of homomeric proteins (Chang et al., 2006; Liang, 2009), the crystal structure of a heteromeric GPPS and its enzymatic regulation mechanism were elucidated very recently (Chang et al., 2010; Hsieh et al., 2010). The production of C$_{10}$-GPP is the key branching point in the C$_{10}$-monoterpene biosynthesis by which the plant volatiles with the critical bioactivities involved in plant growth, development, and defense are made (Pichersky et al., 2006; Gershenzon and Dudareva, 2007).

The model plant Arabidopsis is generally believed to be a self-pollinating plant. However, several pieces of evidence support that insect-mediated cross-pollination also happens in the wild population (Jones, 1971; Snape and Lawrence, 1971; Davis et al., 1998). Arabidopsis has been confirmed to synthesize plant volatiles and emits a range of these compounds from its flower (Aharoni et al., 2003; Chen et al., 2003). Remarkably, the emission of volatiles is a major feature of most insect-pollinated flowers (Dudareva and Pichersky, 2000). In addition, the C$_{10}$-monoterpene could also protect the reproductive organs from pathogen attack or oxidative damage (Wu et al., 2006). Consequently, the presence of homomeric GPPS in Arabidopsis can be responsible for providing the C$_{10}$-monoterpene in the critical metabolism of C$_{10}$-monoterpenes. On the other hand, Wang and Dixon (2009) have also identified a new plastidic Arabidopsis heteromeric GPPS, comprising SSU (AtSSU) and GGPPS isoform 11 (AtGGPPS11).

Here, we showed that the homomeric GPPS in Arabidopsis should be a novel enzyme to generate multiple products with medium/long chain lengths, rather than a GPPS as reported previously (Bouvier et al., 2000). The previous study used C$_5$-DMAPP and C$_5$-[14C]IPP in a ratio of 2:1. A homolog from tomato (Solanum lycopersicum) was also proposed to possess the GPPS activity when the ratio of C$_5$-DMAPP to C$_5$-[14C]IPP was 2.5:1 (Van Schie et al., 2007). In contrast, we used various ratios of 1:15, 1:14, 1:13, and 1:12 for C$_5$-DMAPP to C$_5$-[14C]IPP, C$_{10}$-GPP to C$_5$-[14C]IPP, C$_{15}$-FPP to C$_5$-[14C]IPP, and C$_{20}$-GGPP to C$_5$-[14C]IPP, respectively, to ensure sufficient homoallylic substrates (C$_5$-IPP) for the condensation reaction (see “Materials and Methods”) and determined the preferred allylic substrates of AtPPPS. Even in this assay condition, two unexpected products of C$_{15}$-FPP and C$_{20}$-GGPP were
detected by using radio-gas chromatography (Van Schie et al., 2007). Hence, if sufficient C₅-[¹⁴C]IPP is provided for the continued enzymatic reaction, the short-chain products would turn out to become the medium/long-chain products.

Although AtPPPS showed some GPPS activity when the isotope signal was measured prior to performing thin-layer chromatography, it was barely detectable and insignificant. By reacting C₅-[¹⁴C]IPP with the other three allylic substrates (C₁₀-GPP, C₁₅-FPP, and C₂₀-GGPP), similar multiple medium/long-chain product distribution patterns were observed (Fig. 4). It is also consistent with our phylogenetic analysis that AtPPPS is closely related to long-chain PTSs and with the previous studies that the long-chain PTSs generally prefer using C₁₅-FPP as the allylic substrate instead of C₅-DMAPP (Kellogg and Poultier, 1997; Ogura and Koyama, 1998; Liang et al., 2002). The long-chain PTSs possess a long hydrophobic tunnel, which has higher affinity for intermediates with a longer aliphatic tail. Because C₅-DMAPP has the shortest tail, it would be harder for C₅-DMAPP to remain in the long hydrophobic tunnel and easier to escape from the active site into the bulk solvent than the other allylic substrates. The subtle balance between substrate binding and product release seems to determine the product chain length distribution.

Moreover, the crystal structures of AtPPPS, in its apo and ligand-bound forms, indicate that its substrate-binding cleft is capable of accommodating products larger than C₁₀-GPP. The two point mutations of AtPPPS(I99F/V162F) in the cleft lead to the short-chain-length product of C₂₀-GGPP, with consistent results both in vitro and in vivo. Our results not only clarify the originally thought homomeric GPPS to be AtPPPS but also help explain the actual role of Arabidopsis heteromeric GPPS in the process of C₁₀-monoterpane biosynthesis.

Nevertheless, two other questions remain to be investigated. First, does the regulation system of heteromeric GPPS in Arabidopsis act like other plant heteromeric GPPSs to modulate the distribution of C₁₀-GPP and C₂₀-GGPPS and display the tissue-specific expression pattern (Tholl et al., 2004; Orlova et al., 2009; Wang and Dixon, 2009)? Second, what is the exact biological role of AtPPPS in plants? As shown in previous studies of cellular compartments (Bouvier et al., 2000), AtPPPS has been identified to be capable of transporting into nongreen plastids and chloroplast photosynthetic cells. Additionally, plant SPPSs, closely related to AtPPPS by our phylogenetic analysis, are generally considered to be employed in generating the long-chain prenyl products to serve as the terpene side chains of ubiquinone and plastoquinone, essential components of the electron transport machinery (Hirooka et al., 2003; Jun et al., 2004). Soll et al. (1985) also reported that the final steps of plastoquinone biosynthesis are implemented on the inner envelope of chloroplasts. Judging by these findings, it is tempting to suggest that AtPPPS may play a role in Arabidopsis ubiquinone and plastoquinone biosynthesis. However, we cannot exclude that AtPPPS may also function in other terpene biosyntheses, such as gibberellin, carotenoid, and chlorophyll. It is thus an open question regarding the physiological roles of such multiple medium/long-chain products generated by AtPPPS. It remains to be investigated whether these multiple products can serve as precursor pools to precisely balance diterpene, triterpene, tetra-terpene, and poly-terpene metabolisms, because the dysfunction in the terpene biosynthesis has been reported to have a deleterious effect on plant growth and development (Orlova et al., 2009). We also survey sequences similar to AtPPPS by using the conventional sequence homology search to provide insight for future investigations (Supplemental Table S3). In the end, the role played by AtPPPS remains to be clarified, and our findings encourage reevaluating its enzymatic function in the complex system of metabolite biosyntheses.

CONCLUSION

Taken together, the crystal structures of AtPPPS in combination with phylogenetic analysis and both in vitro and in vivo biochemical assays have clarified the role of this enzyme, which was thought to be a GPPS in previous studies (Bouvier et al., 2000), to be an unusual PTS that synthesizes multiple medium/long-chain products. Our results, along with the identification of a heteromeric GPPS from Arabidopsis (Wang and Dixon, 2009), suggest that the precursor C₁₀-GPP for C₁₀-monoterpane biosynthesis in Arabidopsis may be provided only by heteromeric GPPS. The integrated approach as described here can also be a good example of how gene functions in plant terpene biosynthesis are unraveled.

MATERIALS AND METHODS

Cloning and Mutagenesis

The sequence information of AtPPPS was downloaded from the National Center for Biotechnology Information database with the accession number ATCC34630. The PCR product without its plastid-targeting sequences was amplified by PCR from the Arabidopsis (Arabidopsis thaliana) cDNA libraries using the primers LIC-TEV-At-F and LIC-At-R and cloned into pET-32 Xa/LIC (Novagen; Supplemental Table S2). The forward primer LIC-TEV-At-F includes a tobacco etch virus (TEV) protease cleavage site allowing for N-terminal fusion tag removal. To enhance the nickel-resin-binding affinity, the vector WT-AtPPPS/pET-32 was constructed to insert one additional His₁₀ tag by site-directed mutagenesis. The TEV protease was cloned into pET-51 Ek/LIC (Novagen) by using primers of LIC-TEV-F and LIC-TEV-R (Supplemental Table S2). Other mutant constructs, SM-AtPPPS/pET-32 and AtPPPS(I99F/V162F)/pET-32, used WT-AtPPPS/pET-32 as template and were prepared by site-directed mutagenesis. The primers are shown in Supplemental Table S2.

Protein Expression and Purification

WT-AtPPPS/pET-32 was transformed to Escherichia coli BL21 (DE3) (Novagen) and induced with 0.1 mM isopropyl β-D-thiogalactopyranoside in Luria-Bertani medium containing 100 μg mL⁻¹ carbenicillin at 10°C for 24 h.
Cell pellets were harvested and resuspended in extraction buffer (50 mM Tris, pH 8.5, 40 mM imidazole, 0.75 mM NaN3, 25% [v/v] glycerol, 0.2 mM sorbitol, 10 mM MgCl₂, 2 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP), 2 ng/mL 1-benzonase (Novagen), and Protease Inhibitor Cocktail (Roche)). Cell lysate was prepared by Cell Disruption Solutions (Constant Systems) and centrifuged at 38,000 rpm (Beckman Ti45) for 60 min at 4°C to remove cell debris. After filtration by using Steripor (Millipore), the supernatant was loaded onto a nickel-nitriotropic acid agarose column (GE Healthcare) pre-equilibrated with extraction buffer. The column was washed with wash buffer (25 mM Tris, pH 8.0, 20 mM imidazole, 0.5 mM NaN3, 10% [v/v] glycerol, 10 mM MgCl₂, and 1 mM TCEP) following 10% [v/v] elute buffer (25 mM Tris, pH 8.0, 250 mM imidazole, 0.5 mM NaN3, 10% [v/v] glycerol, 10 mM MgCl₂, and 2 mM TCEP) and eluted with a linear gradient to 100% [v/v] elute buffer. The isolated sample was dialyzed twice against 5 L of buffer (10 mM Tris, pH 8.0, 250 mM imidazole, 0.5 mM NaN3, 10% [v/v] glycerol, 10 mM MgCl₂, and 2 mM TCEP) and concentrated to 20% volume. The sample was used for the crystallization trials.

**Enzymatic Assay**

The biochemical assays followed similar protocols as described previously (Guo et al., 2004; Sun et al., 2005; Chang et al., 2006, 2010; Hsieh et al., 2010). The substrate mixtures (10 μM C₅-DMA with 150 μM C₁₀-[¹⁴C]IPP, 10 μM C₁₀ GPP with 140 μM C₁₀-[¹⁴C]IPP, 10 μM C₁₀-FFP with 130 μM C₁₀-[¹⁴C]IPP, and 10 μM C₁₀-GPP with 120 μM C₁₀-[¹⁴C]IPP) were incubated with 1 μM enzyme for 24 h at 25°C in the reaction buffer (100 mM HEPES, pH 7.5, 0.1% [v/v] Triton X-100, 5 mM MgCl₂, and 50 mM KCN) for the determination of product chain length. The products were separated by thin-layer chromatography using acetonewater (29:1) as the mobile phase.

**Phylogenetic Analysis**

The full-length amino acid sequences were aligned by using ClustalW (Thompson et al., 1994). The evolution history was inferred using the neighbor-joining method (Saitou and Nei, 1987). The percentage of replicate trees in which the corresponding taxa were recovered at least 1000 replicates is shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolution distances used to infer the phylogenetic tree. The phylogenetic tree was made by using MEGA4 (Tamura et al., 2007).

**Reagents and Chemicals**

**Cell Wall Assay**

The cell wall material was extracted from 3-week-old whole plants using the procedure of Mozetic et al. (1985). The extracted material was dissolved in 1 mL of 0.1 M NaOH and neutralized with 1 M HCl to pH 7.0. The cell wall material was extracted, and the concentration of carotenoids was determined by absorbance at 450 nm (Perkin-Elmer Lambda Biod).
deposited in the Protein Data Bank (www.rcsb.org) with codes 3APZ and 3AQG, respectively.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. Multiple sequence alignment of heteromeric and homomeric AtGPPSs.

Supplemental Figure S2. Multiple sequence alignment of plant GPPSs.

Supplemental Figure S3. Amino acid sequence alignment of plant PTs.

Supplemental Figure S4. Time-course assays of substrate and product specificities of WT-AtPPPS.

Supplemental Figure S5. Analytic gel filtration of WT-AtPPPS.

Supplemental Figure S6. Products synthesized by SM-AtPPPS.

Supplemental Figure S7. The additional crystal contacts are used to facilitate crystal lattice formation.

Supplemental Figure S8. The architectures of SM-AIPPPS individual dimers and electron density maps for the ligands.

Supplemental Figure S9. Subunit comparisons of WT-AtPPPS and SM-AIPPPS.

Supplemental Figure S10. Optical absorption was used to quantify the accumulated yellow carotenoid in E. coli carrying pACCAR25lε1F and respective vectors.

Supplemental Table S1. Full-length amino acid sequence relatedness of AtPPPSs with GPPSs from angiosperms and SPPSs.

Supplemental Table S2. The primers used to construct the following clones and mutants of AtPPPSs.

Supplemental Table S3. The BLAST relatedness of AtPPPSs and other plant PTs.

ACKNOWLEDGMENTS

We thank Dr. Makoto Kawamukai (Shimane University) for providing the construct pACCAR25lε1F, the Photon Factory and Spring-8 for beam time allocations, and Dr. Cheng-Chung Lee (Academia Sinica) for collecting the SM-AIPPPS diffraction data.

Received November 2, 2010; accepted January 6, 2011; published January 10, 2011.

LITERATURE CITED


Davis AR, Pylatuik JD, Paradis JC, Low NH (1998) Nectar-carbohydrate production and composition vary in relation to nectary anatomy and location within individual flowers of several species of Brassicaceae. Planta 205: 305–318


Arabidopsis Polyprenyl Pyrophosphate Synthase


1089


