Heteromeric and Homomeric Geranyl Diphosphate Synthases from Catharanthus roseus and Their Role in Monoterpene Indole Alkaloid Biosynthesis

Avanish Rai, Shachi S. Smita, Anup Kumar Singh, Karuna Shanker and Dinesh A. Nagegowda

a CSIR–Central Institute of Medicinal and Aromatic Plants Research Centre, Bangalore, 560065, India
b CSIR–Central Institute of Medicinal and Aromatic Plants, Lucknow, 226015, India

ABSTRACT Catharanthus roseus is the sole source of two most important monoterpene indole alkaloid (MIA) anti-cancer agents: vinblastine and vincristine. MIAs possess a terpene and an indole moiety derived from terpenoid and shikimate pathways, respectively. Geranyl diphosphate (GPP), the entry point to the formation of terpene moiety, is a product of the condensation of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) by GPP synthase (GPPS). Here, we report three genes encoding proteins with sequence similarity to large subunit (CrGPPS.LSU) and small subunit (CrGPPS.SSU) of heteromeric GPPSs, and a homomeric GPPSs. CrGPPS.LSU is a bifunctional enzyme producing both GPP and geranyl geranyl diphosphate (GGPP), CrGPPS.SSU is inactive, whereas CrGPPS is a homomeric enzyme forming GPP. Co-expression of both subunits in Escherichia coli resulted in heteromeric enzyme with enhanced activity producing only GPP. While CrGPPS.LSU and CrGPPS showed higher expression in older and younger leaves, respectively, CrGPPS.SSU showed an increasing trend and decreased gradually. Methyl jasmonate (MeJA) treatment of leaves significantly induced the expression of only CrGPPS.SSU. GFP localization indicated that CrGPPS.SSU is plastidial whereas CrGPPS is mitochondrial. Transient overexpression of AmGPPS.SSU in C. roseus leaves resulted in increased vindoline, immediate monomeric precursor of vinblastine and vincristine. Although C. roseus has both heteromeric and homomeric GPPS enzymes, our results implicate the involvement of only heteromeric GPPS with CrGPPS.SSU regulating the GPP flux for MIA biosynthesis.

Key words: C. roseus; monoterpene indole alkaloids; geranyl diphosphate synthase; heteromer; homomer; large subunit; small subunit; vindoline.

INTRODUCTION

Plants produce a vast array of natural products also called secondary metabolites as a means of self-defense against herbivores and pathogens. Alkaloids are one such class of compounds produced by about 20% of plant species and are mostly derived from amino acids. Monoterpenoid indole alkaloids (MIAs) represent one of the largest classes of alkaloids consisting of ~3000 different compounds, some of which have been shown to possess powerful pharmacological activities (Facchini and De Luca, 2008). The MIAs have been reported to be present in eight different plant families, but are most commonly found in Apocynaceae, Loganiaceae, and Rubiaceae families (Facchini and De Luca, 2008). Madagaskar Periwinkle (C. roseus) belonging to Apocynaceae family is a medicinal plant of enormous importance, as it produces more than 130 MIAs (van der Heijden et al., 2004) and it is the only source of two most important anti-cancer dimeric MIAs—vinblastine and vincristine—and one of their monomeric precursors—vindoline (Liscombe and O'Connor, 2011). The dimeric MIAs vinblastine and vincristine, which accumulate in C. roseus leaves at very low amounts (0.0003–0.01%), are condensation products of vindoline and catharanthine monomers (Liscombe and O'Connor, 2011). Approximately 35 intermediates and more than 30 enzymes are responsible for the biosynthesis of bisindole alkaloids involving different
subcellular compartments and require exclusive transport of pathway intermediates (St-Pierre et al., 1999; van der Heijden et al., 2004). The complex multi-step MIA biosynthetic pathway is under strict regulation of developmental, environmental, organo-, and cell-specific controls (Aerts et al., 1994; Vázquez-Flota and De Luca, 1998; St-Pierre et al., 1999; van der Heijden et al., 2004; Rischer et al., 2006; Wei, 2010). The biosynthesis of strictosidine, which serves as the central precursor for all downstream steps leading to MIAs, involves condensation of shikimate pathway-derived indole moiety, tryptamine, and terpene pathway-derived moiety, secologanin (Facchini and De Luca, 2008) (Figure 1). Precursor feeding experiments have shown that the formation of monoterpene moiety is rate-limiting in MIA biosynthesis (Arvy et al., 1994; Hedhili et al., 2007). Geranyl diphosphate (GPP), the universal precursor for all monoterpenes (Dudareva et al., 2006), is the entry point for the formation of secologanin (Figure 1). GPP is formed by the condensation of dimethylallyl diphosphate (DMAPP) and isopentenyl diphosphate (IPP) in a reaction catalyzed by GPP synthase (GPPS, EC 2.5.1.1) (Ogura and Koyama, 1998), which belongs to the family of short-chain prenyltransferases (Figure 1). GPP is subsequently converted to geraniol, 10-hydroxygeraniol, loganin, and ultimately to secologanin via multiple enzymatic steps (Liu et al., 2007). The short-chain prenyltransferases also include farnesyl diphosphate (FPP) synthase (FPPS; EC 2.5.1.10) and GGPP synthase (GGPPS; EC 2.5.1.30) (Tholl et al., 2004). All these short-chain prenyltransferases function at the branch points of isoprenoid metabolism and play regulatory role in controlling IPP flux into various terpenoid families (Gershenzon and Croteau, 1993; Liang et al., 2002).

GGPPSs have been characterized in few plant species and exist as homomorphic and heteromorphic structures (Nagegowda, 2010). The homomeric GGPPSs are reported in two gymnosperm species (Abies grandis and Picea abies) (Burke and Croteau, 2002a, 2002b; Schmidt and Gershenzon, 2007, 2008; Schmidt et al., 2010) and four angiosperm species (Arabidopsis thaliana, Lycopersicon esculentum, Quercus robur, and Phalaenopsis bellina) (Bouvier et al., 2000; van Schie et al., 2007; Hsiao et al., 2008; Schmidt and Gershenzon, 2008). The heteromeric GGPPSs have been reported only in angiosperms that include Mentha piperita, Antirrhinum majus, Clarkia breweri, and Humulus lupulus, all of which produce large amounts of monoterpenes in specific organs such as trichomes and flower petals (Burke et al., 1999; Tholl et al., 2004; Wang and Dixon, 2009). Structurally, the heteromeric GGPPS is composed of a non-catalytic small subunit (SSU) and a large subunit (LSU), which could be either inactive (Burke et al., 1999) or function as GGPPS on its own (Tholl et al., 2004; Wang and Dixon, 2009), and interaction between the two subunits results in an active heteromeric GGPPS. So far, little is known about the type of GGPPS present in MIA biosynthesizing plants despite the fact that this enzyme catalyzes the branch-point reaction leading to GPP, which serves as substrate for the monoterpene moiety of all MIAs. It has been reported that homomeric GGPPSs in gymnosperms and SSUs of heteromeric GGPPSs regulate monoterpene biosynthesis (Tholl et al., 2004; Hsiao et al., 2008; Wang and Dixon, 2009; Schmidt et al., 2010). Moreover, in tobacco, the endogenous pool of GPP was shown to be limiting for monoterpene biosynthesis (Orlova et al., 2009). In C. roseus also, GGPPS may play a regulatory role in controlling the flux towards formation of MIAs. Hence, it is important to know the type(s) and the role of GGPPS in MIA biosynthesis for eventual metabolic engineering for improved accumulation of MIAs.

In this work, we report functional characterization of heteromeric and homomorphic GGPPSs from C. roseus. We show that CrGGPPS.SSU by itself is a bifunctional G(G)PPS producing both GPP and GGPP, and, in combination with CrGGPPS.LSU, it functions as LSU forming only GPP, whereas CrGGPPS falls into homomeric GGPPS class and produces GGPP as the sole product. The biochemical characterization and MeJA elicitation studies, and analysis of subcellular localization indicated the involvement of CrGGPPS.SSU in regulating the flux for MIA biosynthesis, whereas the other two CrGGPPS.LSU and CrGGPPS could be involved in the formation of plastidial GPP/GGPP for monoterpene biosynthesis and mitochondrial GPP for other

---

**Figure 1.** Simplified View of MIA Biosynthesis in *C. roseus.*

Full and dashed arrows indicate single and multiple enzymatic steps, respectively. The key branch-point step by geranyl diphosphate synthase (GPPS) has been highlighted. DMAPP, dimethylallyl diphosphate; GES, geraniol synthase; GGPP, geranylgeranyl diphosphate; GGPPS, GGPP synthase; GPP, geranyl diphosphate; IPP, isopentenyl diphosphate; PER, peroxidase; SLS, secologanin synthase; STR, strictosidine synthase; TDC, tryptophan decarboxylase; TS, tryptophan synthase.
isoprenoid biosynthesis, respectively. This is the first report of characterization of both heteromeric and homomeric GPPS enzymes from the same plant.

RESULTS AND DISCUSSION

Isolation of C. roseus cDNAs Encoding GPP Synthases

In plants, the enzyme responsible for GPP formation is known to exist either as homomeric or as heteromeric forms (Burke et al., 1999; Bouvier et al., 2000; Burke and Croteau, 2002a, 2002b; Tholl et al., 2004; Schmidt and Gershenzon, 2007; van Schie et al., 2007; Schmidt and Gershenzon, 2008; Orlova et al., 2009; Chang et al., 2010). To identify and investigate the role of GPPS in C. roseus MIA biosynthesis, a search for potential GPPS sequences was carried out initially by blast search analysis against the NCBI (www.ncbi.nlm.nih.gov) database and later against the recently released MPGR (http://mpgrcentral.tamu.edu) database. A search for potential gene encoding GPPS.LSU yielded four potential GGPPSs. The first sequence, named CrGGPPS1 (for C. roseus GGPPS), is represented by a sequence submitted in GenBank (GGPPS; JF896104) and two contigs annotated as GGPPS in MPGR (cra_locus_1452_iso_1_len_1087_ver_3 and cra_locus_1452_iso_2_len_1656_ver_3). This gene has been recently shown by genetic complementation as a functional GGPPS (Thabet et al., 2012). The second sequence referred to as CrGGPPS2 is represented by three contigs (cra_locus_6993_iso_3_len_1491_ver_3, cra_locus_6993_iso_2_len_1416_ver_3, and cra_locus_6993_iso_1_len_1405_ver_3) with an annotation as GGPPS. The third and the fourth sequences, named CrGGPPS3 and CrGGPPS4, are represented by single contigs cra_locus_6454_iso_5_len_849_ver_3 and cra_locus_12153_iso_7_len_1107_ver_3, respectively. Of these CrGGPPS1, CrGGPPS2, and CrGGPPS4 encoded full-length open reading frames, whereas CrGGPPS3 was missing the 3’ region, which was obtained by rapid amplification of cDNA ends (RACE).

Sequence Relatedness of CrGGPPS and Plant GGPPS/LSU Proteins

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CrGGPPS1</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmGGPPS.LSU</td>
<td>74 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIGPPS.LSU</td>
<td>70 69 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MgGGPPS.LSU</td>
<td>69 68 69 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CrGGPPS2</td>
<td>65 64 68 62 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NkGGPPS5</td>
<td>63 57 68 63 74 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CrGGPPS4</td>
<td>62 59 60 58 60 57 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AtGGPPS6</td>
<td>58 58 60 57 58 56 54 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NkGGPPS2</td>
<td>55 57 57 56 57 55 52 52 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CrGGPPS3</td>
<td>46 44 42 42 52 39 52 43 39 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The alignment of sequences of different homomeric and heteromeric enzymes showed the presence of two highly conserved aspartate-rich regions, designated as the first aspartate-rich motif (FARM, DDX(2–4)D) and second aspartate-rich motif (SARM, DDXXD) (where D indicates Asp, X indicates any residue) essential for catalytic function and substrate binding (Koyama et al., 1996; Wang and Ohnuma, 2000) whereas CrGGPPS.SSU lacks both DD(X)D motifs (Tholl et al., 2004). Further comparisons of sequences highlight the presence of two conserved CxxxC motifs (where ‘x’ can be alanine, leucine, isoleucine, valine, glycine, or serine) in CrGGPPS.SSU similarly
to all other SSUs and one such conserved motif in CrGPPS. LSU similarly to other LSUs/GGPPSs. Like other characterized homomeric GPPSs, CrGPPS lacks the CxxxC motif (Figures 2 and 3). The CxxxC motifs are critical for physical interaction between both subunits of plant heterodimeric GPPSs (Wang and Dixon, 2009). Prediction programs like TargetP 1.1, PREDOTAR, ChloroP, and MitoProt indicated the presence of a transit peptide in CrGPPS.LSU and CrGPPS.SSU, suggesting their plastidial localization, and a signal peptide for mitochondrial localization in CrGPPS (Supplemental Table 1).
Heterologous Expression in *Escherichia coli* and Functional Characterization of Heteromeric CrGGPPS

In orchid *P. bellina*, an active homodimeric GPPS with similarity to the SSU class of proteins contributes to monoterpene emission. This protein lacks FARM and SARM motifs, but has GPPS activity. CrGGPPS.SSU significantly homologous to all other SSUs displays 33% sequence identity to PbGGPPS from *P. bellina* (Hsiao et al., 2008). To know whether CrGGPPS.SSU has any activity, CrGGPPS.SSU CDNA with (His)$_6$-tag was expressed and purified (Figure 4D, lane 2). The purified recombinant CrGGPPS.SSU had no detectable prenyltransferase activity with IPP and DMAPP (Figure 4E and 4F, panel 2). This inactive nature of SSU is similar to previously reported SSUs from *M. piperita* (Burke et al., 1999), *A. majus*, *C. breweri* (Tholl et al., 2004), *H. Lupulus*, and *A. thaliana* AtSSU-II, which constitutes a distinct clade of trans-prenyltransferases and contains first conserved DDX$_{2-4}$D motif and two CxxxC motif but lacks SARM motif (Wang and Dixon, 2009).

It is known that plants contain multiple GGPPS or GGPPS-related enzymes, in which some function as LSU of heteromeric GPPS (Tholl et al., 2004; Orlova et al., 2009; Wang and Dixon, 2009). The isolated sequences clearly suggest the scenario of multiple GGPPS enzymes in *C. roseus* similar to *A. thaliana* (Okada et al., 2000). However, one of these may act as LSU in combination with SSU to form a functional heteromer. Hence, to identify the GPPS.LSU, the open reading frame of CrGGPPS.SSU without (His)$_6$-tag was co-expressed individually with CrGGPPSs (lacking transit peptide) carrying a (His)$_6$-tag. Earlier studies with GPPS and GGPPS have shown optimal expression of soluble proteins only after deleting the N-terminal transit-peptide sequence (Burke and Croteau, 2002a). Purification of co-expressed recombinant proteins and their subsequent sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) analysis revealed the interaction of only CrGGPPS1 among all CrGGPPSs with CrGGPPS.SSU, which is evident from co-purified bands of both subunits (Figure 4A, lane 1). Moreover, the prenyltransferase activity assay with [1-3H]-IPP and DMAPP confirmed the formation of an active heteromeric complex containing CrGGPPS1 and CrGGPPS.SSU (Figure 4B).

Although all CrGGPPSs possess one CxxxC motif similar to other characterized LSUs that is critical for interaction with SSU, only CrGGPPS1 interacted with CrGGPPS.SSU, suggesting that the CxxxC motif is necessary but not sufficient for physical interaction (Wang and Dixon, 2009). It is not surprising that the other three CrGGPPSs did not show interaction with CrGGPPS.SSU, as it was also reported in the case of *Arabidopsis* GGPPSs that AtGGPPS6 and AtGGPPS11 possess one CxxxC motif, but only the latter interacted with AtSSU (Wang and Dixon, 2009). Moreover, AtSSU also interacted with *H. Lupulus* GPPS.LSU (Wang and Dixon, 2009). The pairwise percent- age analysis of deduced amino acid sequences of CrGGPPSs with other LSUs and GGPPSs indicated that only CrGGPPS1 is closely related to LSUs (Table 1). Also, in tobacco, two GGPPS proteins interacted with AmGPPS.SSU (Orlova et al., 2009), indicating that further studies are required to understand

**Figure 3.** Multiple Sequence Alignment of Plant Homomeric GPP Synthases.

The characterized GPPSs from *L. esculentum* (LeGPPS; DQ286930), *A. thaliana* (AtGPPS; Y17376), *P. abies* (PaIDS3; EU432047), and *Q. robur* (QrIDS1; CAC20852) are included in sequence alignment along with *C. roseus* (CrGPPS, JX417185). Conserved residues are in black (identical in at least three out of six sequences shown) and residues in gray are similar in at least two of the six sequences shown. The two conserved Asp-rich motifs (FARM and SARM) are indicated by solid line. The black inverted triangle indicates the truncation site for CrGPPS expressed in *E. coli* as a pseudomature form.
Since CrGGPPS1 interacted with CrGGPPS.SSU, we hereafter refer to this as CrGGPPS.LSU.

It has been reported in heteromeric GPPS that LSU alone could be either inactive (Burke et al., 1999) or function as a bona fide GGPPS enzyme on its own (Tholl et al., 2004; Wang and Dixon, 2009). Hence, to identify the function of CrGGPPS.LSU, a truncated version starting from second methionine (M-27) was expressed in E. coli as a (His)_6-tagged protein. The assay with soluble extracts containing recombinant (His)_6-CrGGPPS.LSU proteins using [1-^3H]-IPP and DMAPP as substrates and Mg^{2+} as cofactor detected prenyltransferase activity. This activity was further confirmed using purified (His)_6-CrGGPPS.SSU proteins (Figure 4D, lane 1, and Figure 4E, panel 1). Product verification of (His)_6-CrGGPPS.LSU enzyme assay using IPP and DMAPP revealed the formation of both geranyl (G) and farnesyl (F) pyrophosphates.
GPP and GGPP (~2:1) indicating the bifunctional nature of this enzyme (Figure 4F, panel 1). The bifunctional CrGPPS. LSU could provide both GPP and GGPP pool. GPP could be used as a precursor for terpene moiety formation, and GGPP could be used for biosynthesis of diterpene, chlorophyll, phylloquinone, and plastoquinone (Lange and Ghassemian, 2003) or in protein prenylation like geranylgeranylation (Gerber et al., 2009). As GGPPSs are known to have an ability to use other allylic diphosphates such as GPP and FPP in addition to DMAPP (Ogura and Koyama, 1998; Takaya et al., 2003; Schmidt et al., 2010), enzyme assays and product verification showed that CrGPPS.LSU produced GGPP by utilizing both GPP and FPP in the presence of IPP, suggesting that reaction proceeds by sequential incorporation of IPP units (Figure 4G). The GGPPS activity of CrGPPS.LSU was further confirmed by in vivo genetic complementation (Figure 5). The transformants carrying pACCAR25ΔcrtE and human GGPPS (HsGGPPS, positive control) are expected to accumulate yellow pigmentation, confirming the substitution of crtE, whereas cells harboring pACCAR25ΔcrtE and CrGPPS.SSU/pET28a empty vector were used as negative controls. Cotransformation of CrGPPS.LSU (which is the same as CrGGPPS reported by Thabet et al., 2012) with pACCAR25ΔcrtE resulted in accumulation of yellow pigmentation (Figure 5B), confirming the in vitro assay data that CrGPPS.LSU has a GGPPS activity. Although, M. piperita heteromeric GPPS produced both GPP and GGPP (Tholl et al., 2004; Orlova et al., 2009) whereas LSU from H. lupulus GPPS formed GPP, FPP, and GGPP (Wang and Dixon, 2009). It is known that most short-chain prenyltransferases produce only a single main product (GPP or FPP or GGPP) but some homomeric prenyltransferases found in Zea mays and P. abies, and heteromeric GGPPS from H. lupus and M. Piperita, make products with one more or one fewer C5 units than the main product (Cervantes-Cervantes et al., 2006; Schmidt and Gershenzon, 2008; Wang and Dixon, 2009). Unlike other bifunctional enzymes stated above, CrGPPS.LSU is functionally similar to a recently reported phylogenetically distant gymnosperm PaIDS1 from P. abies, which showed bifunctional activity forming products GPP and GGPP (9:1 ratio) that differ from each other in size by more than one C5 unit (Schmidt et al., 2010). In this context, PaIDS1 and CrGPPS.LSU belong to an unusual class of short-chain prenyltransferases that make two major products with a difference of more than one C5 unit (Schmidt et al., 2010). This is the first report demonstrating the ‘catalytic promiscuity’ of an angiosperm prenyltransferase where the enzyme has the ability to catalyze an adventitious secondary activity (GGPS) at the active site responsible for the primary activity (GGPPS) (Copley, 2003; Schmidt et al., 2010). This ‘catalytic promiscuity’ of CrGPPS.LSU suggests its possible involvement in primary as well as secondary metabolism during the course of growth and development of C. roseus. The overexpression of snapdragon GPPS.SSU changed the chain length specificity of endogenous GGPPSs into GPPS leading to enhanced
monoterpene formation, but at the expense of primary metabolism (Oriova et al., 2009). It was suggested that the in vivo formation of heteromeric GPPS complex between introduced snapdragon GPPS.SSU and tobacco GGPPSs redirected the metabolic flux to GPP formation consequently depleting the GGPP pool needed for primary metabolism (Oriova et al., 2009). Thus, efforts through metabolic engineering of dual G(G)PP synthase could increase both GPP and GGPP pool without affecting the primary metabolism in C. roseus.

To characterize the product formation by the heteromeric CrGPPS.SSU/CrGPPS.LSU complex, the co-purified proteins were used for prenyltransferase assays, which showed GPP as the exclusive product (Figure 4C). This is consistent with earlier reports on heteromeric GPPS from A. majus (Tholl et al., 2004) and in contrast to observations with heteromeric GPPS from H. lupulus and M. piperita that produced both GPP and GGPP and trace amounts of FPP (Wang and Dixon, 2009; Chang et al., 2010). These results indicated that catalytically active GPPS in C. roseus is present as a heteromer, which consists of inactive CrGPPS.SSU as a small subunit and an active bifunctional G(G)PPS as a large subunit. The interaction between CrGPPS.LSU and CrGPPS.SSU modified the chain length specificity of CrGPPS.LSU resulting in the exclusive production of GPP from IPP and DMAPP (Figure 4C). It has been shown in the structure of heteromeric GPPS from M. piperita that SSU restricts the connection between active site cavity and elongation cavity thereby limiting the catalytic reaction to go beyond C10-GPP (Chang et al., 2010). Similarly, in C. roseus, it is possible that the interaction of CrGPPS.SSU alters the catalytic site of CrGPPS.LSU in their subunit architecture and confined the enzyme specificity to the formation of only GPP (Figure 4C). Moreover, the complex formation suggests that CrGPPS.SSU has a strong affinity to CrGPPS.LSU and could control CrGPPS.LSU homomerization. Our results show that CrGPPS.LSU serves as a functional GGPPS, which may be involved in primary metabolism, but to some extent could also contribute to secondary metabolism with its ability to release a substantial portion of the intermediate GPP, similarly to PaIDS1 (Schmidt et al., 2010). However, formation of a heteromer with CrGPPS.SSU resulting in highly efficient GPPS could contribute to the major redirection of flux to secondary metabolism.

Heterologous Expression and Functional Characterization of Homomeric CrGPPS

For functional characterization of homomeric CrGPPS, the open reading frame was cloned into expression vector pET28a having an N-terminal (His)$_n$-tag. The cloned construct was transformed into E. coli and the recombinant CrGPPS was expressed as a (His)$_n$-tag fusion protein. The full-length recombinant CrGPPS protein was completely insoluble (data not shown). A truncated version of CrGPPS starting from second methionine (Met$_{100}$–CrGPPS), and containing a (His)$_n$-tag, was expressed in E. coli. The truncation was based on earlier reports of A. grandis GPPS and LeGPS where truncation did not affect the catalytic activity and product profile (Burke and Croteau, 2002a; van Schie et al., 2007). The assay with purified recombinant CrGPPS proteins using [1-3H]-IPP and DMAPP as substrates showed prenyltransferase activity (Figure 6A and 6B). While the purified recombinant CrGPPS produced GPP as the sole product with IPP and DMAPP as substrates (Figure 6C), it produced trace amounts of GGPP in the presence of either GPP or FPP along with IPP (Figure 6D). Both CrGPPS.LSU and CrGPPS showed preference for Mg$^{2+}$ over Mn$^{2+}$ and K$^+$ as cofactor similarly to other prenyltransferases (Supplemental Figure 1). The formation of only GPP by CrGPPS with IPP and DMAPP as substrates is in contrast to other homomeric GPPSs of the same clade represented by LeGPS, QrIDS1, and PaIDS3, which produced substantial
amounts of GPP, FPP, and GGPP (van Schie et al., 2007; Schmidt and Gershenzon, 2008), whereas an Arabidopsis prenyltransferase that was thought to be a GPPS earlier (Bouvier et al., 2000) but shown to be a polyprenyl pyrophosphate synthase (PPPS) recently produced no product (Hsieh et al., 2011). Also, a homomeric PbGDPS lacking substrate binding motifs (FARM and SARM) catalyzed GPP and trace FPP formation with IPP and DMAPP (Hsiao et al., 2008). However, a homomeric GPPS, PaIDS2 belonging to a different clade has the same product profile as that of CrGPPS with GPP as the sole product (Schmidt and Gershenzon, 2008). The production of a trace amount of GGPP by CrGPPS when incubated with GPP/FPP and IPP prompted us to check whether CrGPPS has any in vivo GGPPS activity. The genetic complementation assay showed no detectable yellow pigmentation in transformants carrying CrGPPS and pACCAR25ΔcrtE constructs similarly to negative controls (Figure 5B), suggesting that CrGPPS is a bona fide GPPS. Interestingly, none of the homomeric GPPSs present in the same clade along with CrGPPS is involved in monoterpene biosynthesis and it is evident from earlier studies that GPPSs from tomato (van Schie et al., 2007), P. abies (PaIDS3), and Q. robur (QrIDS1) were shown to be involved in diterpene formation (Schmidt and Gershenzon, 2008).

### Spatio-Temporal Expression of CrGPPS Encoding Genes

The biosynthesis of MIAs in *C. roseus* is regulated tissue-specifically and developmentally (De Luca and St-Pierre, 2000; Facchini, 2001; Roepke et al., 2010). The complex MIA biosynthetic pathway is compartmentalized in multiple cell types of different organs such as shoots, leaves, and flowers (St-Pierre et al., 1999). To determine the contribution of *CrGPPS* genes in MIA formation, the mRNA expression was analyzed in different tissues and developmental stages, and in leaves treated with MeJA (a known alkaloid pathway regulator) by quantitative real-time PCR using gene-specific primers. With respect to tissue specificity, *CrGPPS.LSU* showed highest expression in stem and leaf followed by flower, with least expression in siliques and roots (Figure 7A). *CrGPPS.SSU* exhibited highest expression in flowers followed by stem, leaf, and siliques, with least expression in roots (Figure 7B). *CrGPPS* was constitutively expressed in all tissues tested (Figure 7C). In earlier studies, the gene encoding SSUs of all characterized heteromeric GPPSs and a homomeric GPPS from orchid have shown tissue-specific expression in flowers (Tholl et al., 2004; Hsiao et al., 2008) or in trichomes (Wang and Dixon, 2009) that are the sites of monoterpene biosynthesis and emission. However, *CrGPPS.SSU* showed broad tissue specificity, with highest expression in flowers, followed by stem and leaf, indicating its possible involvement in floral monoterpene emission as well as in MIA formation. Young developing leaves are the sites of MIA biosynthesis and regulation of the MIA pathway at cell-, development-, and organ-specific levels, which suggests that the pathway gene expression is coupled to the secretory mechanisms of vindoline and catharanthine (De Luca and St-Pierre, 2000; Roepke et al., 2010). The accumulation of vindoline and catharanthine has been shown to increase with leaf age during early leaf
development (up to third-leaf age) and gradually declined reaching to ~40% of the maximum in older leaves (Roepke et al., 2010). When expression levels of mRNAs for CrGPPS genes were analyzed in leaves of different developing stages (from the first pair to the ninth pair of developing leaves), CrGPPS.SSU showed an increasing trend, reaching a maximum in the fifth-leaf pair, and declined slightly thereafter (Figure 8A), which more or less corresponded with the accumulation profile of vindoline and catharanthine (Roepke et al., 2010). In contrast, the expression of CrGPPS.LSU and CrGPPS showed reverse trends, with CrGPPS.LSU showing increased expression in older leaves and CrGPPS3 with higher expression in younger leaves (Figure 8A).

MeJA is a known inducer of MIA pathway resulting in the induced production of MIAs in C. roseus seedlings (Aerts et al., 1994), in hairy root cultures (Rijhwani and Shanks, 1998), in detached leaves (El-Sayed and Verpoorte, 2005), and in cell cultures (Rischer et al., 2006). Moreover, MeJA stimulates expression of enzymes associated with transcription factor-controlled accumulation of terpenoids and also is known to modulate expression of genes of early steps of terpene biosynthesis (reviewed in Hemmerlin et al., 2012). To determine the role of CrGPPS genes in the induced formation of MIAs, transcript levels were measured over a 12-h time course after leaves were treated with MeJA. Relative expression levels of CrGPPS genes were compared with those of control (Figure 8B). Unlike CrGPPS.LSU and CrGPPS that remained relatively constant after MeJA treatment, CrGPPS.SSU expression showed a rapid induction in a biphasic kinetic manner, with an immediate induction after 1 h followed by a decrease at 4 h and 8 h and dramatically peaking at 12 h (Figure 8B). A similar biphasic induction pattern was previously observed for octadecanoid-responsive Catharanthus AP-2 domain ORCA2 (Menke et al., 1999), ORCA3 (van der Fits and Memelink, 2001), and TDC and STR transcripts (Dutta et al., 2007). Also, ESMB (Early Steps in Monoterpene Biosynthesis) genes such as 1-deoxy-D-xylulose 5-phosphate synthase (DXS2) (Chahed et al., 2000), 1-deoxy-D-xylulose 5-phosphate reductase (DXR) (Veau et al., 2000), 2C-methyl-D-erythrol-2,4-cyclodiphosphate synthase (MECS), and 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase (HDS), and geraniol 10-hydroxylase (G10H) have shown induced expression in response to MeJA (Rischer et al., 2006; Oudin et al., 2007). Moreover, a genome-wide transcript profiling by cDNA-amplified fragment-length polymorphisms (AFLP) with metabolic profiling of MeJA elicited C. roseus cell cultures showed an induced transcript accumulation of ORCA2 and ORCA3 and of genes encoding ESMB enzymes such as MECS, HDS, GPPS, G10H, and 10-hydroxygeraniol oxidoreductase (10HGO) (Rischer et al., 2006), in which the AFLP fragment annotated as GPPS in their study corresponded to CrGPPS.SSU. Endogenous GPPS enzyme activity was increased after MeJA application in P. Abies, suggesting that GPP produced by this enzyme is an important substrate for monoterpene biosynthesis in oleoresin production (Martin et al., 2002). The results
of spatio-temporal gene expression analysis combined with MeJA induction experiments suggested that only CrGPPS.SSU could be involved in regulating GPP availability for MIA biosynthesis. This conclusion is in agreement with previous studies that expression of SSU but not LSU regulates monoterpene biosynthesis (Tholl et al., 2004; Orlova et al., 2009; Wang and Dixon, 2009). Since it is known that catharanthine and vindoline production occurs in young developing leaves (Roepke et al., 2010), their relative transcript abundance in leaves of the third developmental stage was measured, which showed that, of the three genes, CrGPPS.LSU exhibited the highest expression, followed by CrGPPS and with lowest expression for CrGPPS.SSU (Figure 8C). Relative to CrGPPS.LSU, the low level of CrGPPS.SSU expression in contrast to other reported highly expressing SSUs of snapdragon, mint, and hop suggests that, in C. roseus, GPP formation is regulated in a different manner.

Effect of Transient Overexpression of AmGPPS.SSU in C. roseus Leaves

The SSU of heteromeric GPPS is known to interact with phylogenetically distant GGPPS proteins and modify the chain length specificity resulting in enhanced GPPS activity (Burke and Croteau, 2002b; Tholl et al., 2004; Wang and Dixon, 2009; Orlova et al., 2009). To determine whether AmGPPS.SSU interacts with CrGPPS.LSU, constructs carrying CrGPPS.LSU with a (His)_6-tag and AmGPPS.SSU without (His)_6-tag were co-expressed in E. coli and purified by Ni-NTA chromatography. The SDS–PAGE analysis (Figure 9A) and activity assays showed that CrGPPS.LSU interacted with AmGPPS.SSU, resulting in an active heteromeric GPPS that produced exclusively GPP from IPP and DMAPP (Figure 9B). In heteromeric GPPSs, SSU acts as a regulatory unit in controlling the flux to GPP for monoterpene production in specialized tissues such as trichomes of Mentha and hop, and petals of snapdragon flowers (Tholl et al., 2004; Orlova et al., 2009; Wang and Dixon, 2009). Since AmGPPS.SSU interacted with CrGPPS.LSU in vitro resulting in heteromeric GPPS, we transiently overexpressed AmGPPS.SSU in C. roseus leaves by Agroinfiltration. Transient overexpression of TDC and STR1 has been successfully demonstrated earlier in C. roseus leaves by Agroinfiltration (Di Fiore et al., 2004). Analysis of mRNA expression in C. roseus leaves infiltrated with Agrobacterium carrying a AmGPPSS.SSU overexpression construct and control leaves infiltrated with Agrobacteria alone (Figure 9D). This suggested that the introduced AmGPPS.SSU interacted with the endogenous CrGPPS.LSU, resulting in increased endogenous GPPS activity leading to improved availability of GPP for MIA biosynthesis. Thus, GPPS, being the

Figure 9. Transient Overexpression of AmGPPS.SSU in C. roseus Leaves.
(A) SDS−PAGE analysis showing co-purified band of heteromeric complex containing CrGPPS.LSU and AmGPPS.SSU.
(B) TLC analysis of reaction products from in vitro assays using the CrGPPS.LSU (panel 1) and CrGPPS.LSU/AmGPPS.SSU complex (panel 2) in presence of IPP and DMAPP. Semi-quantitative RT–PCR analysis (C) and vindoline content (D) in C. roseus leaves infiltrated with Agrobacterium carrying a AmGPPSS.SSU overexpression construct and control leaves infiltrated with Agrobacterium carrying empty vector. Analysis was carried out 48 h post infiltration.
branch-point enzyme for monoterpenoid moiety formation, could control the flux towards MIA formation.

Subcellular Localization of C. roseus GPPS Proteins

The biochemical characterization of CrGPPS proteins showed that they may participate in the formation of GPP in vivo—a process that is generally accepted to take place in plastids by utilizing IPP and DMAPP derived from the MEP pathway (Eisenreich et al., 1997). In silico analysis of CrGPPS amino acid sequences using various prediction programs revealed that CrGPPS.LSU and CrGPPS.SSU were predicted to be localized in the chloroplast, whereas CrGPPS had a high possibility of mitochondrial localization (Supplemental Table 1). Moreover, analysis of sequences of characterized homomeric GPPSs from A. thaliana, L. esculentum, Q. robur, P. abies (PaIDS2), and A. grandis also predicted their localization in mitochondria (Supplemental Table 1). The localization of CrGPPS.LSU has been shown to be plastidial in a previous study (Thabet et al., 2012), where they have annotated it as CrGGPPS. To determine the subcellular localization of the other two proteins, the first 100- and 142-aa sequences of CrGPPS.SSU and CrGPPS, respectively, were fused to the N-terminus of 326-sGFP under the control of CaMV 35S promoter. Transient overexpression of corresponding fusion proteins in A. thaliana protoplasts exhibited the GFP fluorescence in chloroplasts for CrGPPS.SSU (Figure 10). The plastidial localization of both subunits of heteromeric CrGPPS is consistent with earlier reports from SSU of A. majus and LSU of GPPSs from M. piperita and H. lupulus (Tholl et al., 2004; Turner and Croteau, 2004; Wang and Dixon, 2009).

Unlike CrGPPS.LSU and CrGPPS.SSU, CrGPPS–GFP fusion protein exhibited a punctuate pattern of fluorescence indicating mitochondrial localization (Figure 10), which is consistent with all prediction programs (Supplemental Table 1).

Phylogenetic Relationships of CrGPPSs

Genes encoding GPPS enzymes have been isolated and characterized in select members of plants belonging to both angiosperms and gymnosperms, which are mostly known to produce high amounts of monoterpenes. Based on their sequence comparison and prenyltransferase activity, it has

Figure 10. Confocal Laser Scanning Microscopy of Transiently Expressed CrGPPS.SSU–GFP and CrGPPS–GFP Fusions in Arabidopsis Protoplasts. GFP fusion constructs are shown on the left, and the corresponding transient expression in the protoplast is shown on the right. Green fluorescent protein fluorescence detected in the green channel is shown in the ‘Green’ column; the ‘Transmission’ column shows light-microscopy images of intact protoplasts. Scale bars = 50 μm. The numbers in the fusion constructs correspond to amino acid positions.
been suggested that plant GPPSs might have multiple origins during the course of evolution (Tholl et al., 2004) with an ability to produce monoterpenes in specialized organs (Wang and Dixon, 2009). Although GPPS enzymes have been characterized in angiosperm and gymnosperms, the types of GPPS enzymes and its/their role in MIA producing plant is not yet reported. A phylogenetic analysis of CrGPPS proteins with all characterized GPPS sequences revealed three major clades of isoprenyl diphosphate synthases (IDS): IDS-a, IDS-b, and IDS-c (Figure 11). IDS-a is subdivided into IDS-a1 and IDS-a2 based on the enzyme activity and evolutionary relatedness. IDS-a1 clade consists of enzymes from angiosperms and is represented by LSUs of heteromeric GPPS and GGPPS (Burke et al., 1999; Tholl et al., 2004; Wang and Dixon, 2009; Orlova et al., 2009). All candidates of IDS-a1 produced GGPP as the major product with the exception of CrGPPS.LSU, which predominantly produced GPP (Figure 11). IDS-a2 representing homomeric GPPS is found only in conifers, which includes GPPSs from A. grandis (Burke and Croteau, 2002a) and P. abies (Schmidt and Gershenzon, 2008; Schmidt et al., 2010). Interestingly, all enzymes of IDS-a2 produced GPP as the predominant product in addition to FPP and GGPP with the exception of PaIDS1, which is reported to be a bifunctional G(G)PPS producing substantial amounts of GPP and GGPP. CrGPPS.LSU and PaIDS1, being in two different sub-clades, display the catalytic promiscuity having bifunctional enzymatic activity producing GPP and GGPP. CrGPPS falls

![Figure 11](http://mplant.oxfordjournals.org/article/fig/11.png)

**Figure 11.** Phylogenetic Analysis of Amino Acid Sequences of Homomeric and Heteromeric Plant GPPSs. A phylogenetic tree was constructed by the neighbor-joining method using default settings of MEGA 4.0 (Tamura et al., 2007). The branch length of the line indicates evolutionary distances, and numbers represent the confidence of the phylogenetic tree calculated by bootstrap analysis from 1000 replicates. The abbreviation and accession number used in constructing the phylogenetic tree are given in the ‘Methods’ section. Products generated by different short-chain prenyltransferases are represented by G/F/GG in order of their abundance. Products generated by heteromeric complexes involving SSUs are shown in italics. IDS, isoprenyl diphosphate synthase. G, F, and GG represent GPP, FPP, and GGPP, respectively.
into IDS-b clade that includes three angiosperm GPPSs from Arabidopsis (AtGPPS), L. esculentum (LeGPS), and Q. rubur (QrIDS1), and one gymnosperm GPPS from P. abies (PaiDS3), which produced FPP and or GGPP in addition to GPP (Bouvier et al., 2000; van Schie et al., 2007; Schmidt and Gershenzon, 2008). AtGPPS was recently shown to be a PPPS in contrast to the earlier report (Hsieh et al., 2011). However, CrGPPS from the same IDS-b clade produced only GPP and showed mitochondrial localization (Figure 10 and Figure 11). Interestingly, prediction analyses revealed that, like CrGPPS, the GPPSs of A. grandis (AgGPPS1) of clade-IDSa2 and all proteins of clade IDS-b seem to be localized to mitochondria (Supplemental Table 1). This suggested the possible role of IDS-a2 and IDS-b clade of enzymes in associated terpene biosynthesis pathways in plant cell. The IDS-c1 is represented by inactive SSU-I class of heteromeric GPPS, whereas an unusual homomeric GPPS from orchid (P. bellina) lacking a DD(X)2–4D motif with close relatedness to SSUs has been clustered as IDS-c2 clade (Hsiao et al., 2008). The IDS-c3 clade contains a unique A. thaliana protein (AtSSU-II) with two CxxxC motifs and a FARM motif but lacking a SARM motif (Wang and Dixon, 2009) (Figure 11). AtSSU-II showed no activity but modified the chain length specificity of AtGGPPS11 and HIGPPS.LSU, which is similar to the function of other characterized SSUs (Wang and Dixon, 2009). Our phylogenetic analysis demonstrated that IDS-c clade, consisting of members of IDS-c1, IDS-c2, and IDS-c3, could have evolved from a common ancestor in which a homodimeric GPPS falling into IDS-c2 clade is active whereas SSU-I (IDS-c1) and SSU-II (IDS-c3) are catalytically inactive. This could be due to the exceptionally high speciation rate of orchids (Tsai and Chen, 2006) that would have altered the inactive SSU by modification of a few amino acids, resulting in the required PPPS activity. This is evident with the presence of an alternate EAEEV motif of PbGPPS, which is located in the equivalent position of SARM (Hsiao et al., 2008). The significance of presence of SSU-II in plants is currently unknown but a systematic characterization of this class in different plants should provide the evidence for its role in plant metabolism.

Conclusions
The study of complex MIA biosynthesis and its regulation remains a great challenge in plant secondary metabolism. The functional characterization of CrGPPSs in this study provided insights into the formation of GPP in C. roseus. The CrGPPS.LSU might have evolved to produce both GPP and GGPP in C. roseus to maintain relative amounts of precursors for both primary as well as secondary metabolism. The plastid targeted CrGPPS.SSU and CrGPPS.LSU form active heteromeric GPPS to catalyze efficient production of GPP in vivo and provide better availability of this substrate to geraniol synthase, whereas homomeric CrGPPS might be involved in mitochondrial isoprenoid biosynthesis for the production of ubiquinones. The temporal and spatial expression analysis of CrGPPS genes and transient overexpression implies that CrGPPS.SSU could function in coordinated redirection of the metabolic flux, thereby acting as a primary regulator to a certain extent in MIA biosynthesis. Taken together, our results suggest that, under normal conditions, a basal level of the GPP pool is provided by both CrGPPS.LSU (a bifunctional G(G)PPS) and heteromeric CrGPPS.LSU/CrGPPSSSU towards MIA formation in C. roseus. However, under biotic/abiotic stress, the induced expression of SSU could result in an elevated level of GPPS activity by LSU and SSU interaction, thereby increasing the GPP pool towards MIA biosynthesis.

METHODS

Plant Material and MeJA Treatment
C. roseus cv. Dhawal (National Gene Bank, CSIR–CIMAP, India) plants were grown under normal greenhouse conditions. For MeJA treatments, 95% pure MeJA (Sigma-Aldrich) was added to a 5% (w/v) sucrose solution to a final concentration of 200 mM. Excised leaves of the third developmental stage were placed in a Petri plate containing the MeJA/sucrose or DMSO (the MeJA solvent) as a control. Samples were collected at 0, 1, 4, 8, and 12 h, and stored at –80°C until further use.

Chemicals and Radiochemicals
[1-3H]-IPP (1.85 MBq) was purchased from American Radiolabeled Chemicals (www.arincusa.com). Unlabeled IPP, DMAPP, GPP, and FPP were purchased from Echelon Research Laboratories (www.echelon-inc.com). Terpenoid standards and reagents were purchased from Sigma-Aldrich (www.sigmaaldrich.com) unless otherwise noted.

RNA Isolation and cDNA Synthesis
About 100 mg of tissue was ground in liquid nitrogen and total RNA was extracted using a Spectrum™ Plant Total RNA Kit (Sigma-Aldrich, USA) according to the manufacturer’s instructions. To amplify the 3′ end of CrGPPS.SSU, 3′ RACE cDNA was prepared with 5 μg total RNA using a Superscript III RT Module (Invitrogen). Then, 3′ RACE PCR was carried out using a GeneRacer™ Kit (Invitrogen) with 3′ RACE cDNA along with gene-specific reverse and nested primers and GeneRacer 3′ and 3′ nested primers (Supplemental Table 2). The amplified fragment was cloned into pJET1.2/ vector and the sequence was confirmed by nucleotide sequencing. cDNA for amplifying open reading frames of CrGPPSS and for real-time PCR analysis was prepared by using a RevertAid™ H Minus First Strand cDNA Synthesis Kit (Fermentas International Inc., Canada). For real-time qRT–PCR, the DNA-free total RNA (5 μg) was used for first-strand cDNA synthesis with Oligo(dT)18 primers using RevertAid H Minus Reverse Transcriptase. A linear range of cDNA was used for real-time PCR. Real-time qRT–PCR was performed in a 96-well plate using the Applied Biosystems 7900HT Fast Real-Time PCR System (PE Applied Biosystems, www.appliedbiosystems.com) with SYBR green fluorescent dye using respective forward and reverse primers for each gene (Supplemental Table 2). Fold change differences in gene
expression were analyzed using the comparative cycle threshold (Ct) method (Applied Biosystems). Each data point represents the mean of two independent biological replicates and three technical replicates. Real-time qRT–PCR conditions were as follows: 94°C for 10 min for one cycle, followed by 40 cycles of 94°C for 15 s, 54°C for 15 s, and 72°C for 15 s.

**Phylogenetic Analysis**

The deduced amino acid sequence were aligned using MEGA 4.0 (Tamura et al., 2007) with default settings. A phylogenetic tree was constructed with the neighbor-joining method using default settings of MEGA (Tamura et al., 2007). Bootstrap values were calculated from 1000 replicates. Shading of amino acid sequence alignment was done by BOXSHADE 3.21 version (www.ch.embnet.org/software/BOX_form.html).

**Cloning of GPPS Genes and Generation of Overexpression Constructs**

The open reading frame of each CrGPPS was amplified with a forward and a reverse primer consisting of the primers that included a start and stop codon along with restriction sites for cloning into pET28a vector. The primer combinations used are mentioned in Supplemental Table 2. PCR was performed using Platinum Taq DNA polymerase High Fidelity (Invitrogen) and PCR products were gel-purified using a gel extraction kit (Fermentas). PCR products were cloned into pET 1.2/blunt cloning vector (CloneJET PCR cloning kit, Fermentas) and transformed into E. coli XL1 for plasmid amplification. After restriction digestion and gel extraction, resulting fragments were sub-cloned into the expression vector pET28a (Novagen) downstream and in-frame of (His)_6-tag or pET32a for non-(His)_6-tag clones. Positive clones were selected and the resulting constructs were confirmed by restriction digestion and nucleotide sequencing.

**Expression and Purification of Recombinant Proteins**

For functional expression, E. coli Rosetta-2 competent cells were transformed with recombinant plasmids and pET28a/pET32a lacking an insert (control). Induction, harvesting, and protein purification by affinity chromatography on nickel-nitrilotriacetic acid agarose (Bio-Rad, www.bio-rad.com) were performed as described by Nagegowda et al. (2008). Briefly, a single colony was used to inoculate 25 ml Luria-Bertani medium with 37 mg ml⁻¹ chloramphenicol and 50 mg ml⁻¹ ampicillin (for pET32a constructs) or 50 mg ml⁻¹ kanamycin (for pET28a constructs) or with both antibiotics (for co-expression of both pET32a/pET28a constructs), which were grown overnight at 37°C. Five ml of these cultures were transferred to 1000 ml of the same medium and continued to grow at 37°C until an absorbance of 0.5 at OD 600 nm. Cultures were then induced by the addition of isopropyl-1-thio-β-D-galactopyranoside (IPTG) to a final concentration of 0.4 mM and grown for an additional 18 h at 18°C. Protein purification by affinity chromatography on nickel-nitrilotriacetic (Ni-NTA) acid agarose (Bio-Rad) was performed. Protein concentration was determined using the Bradford method (Bradford, 1976).

**Prenyltransferase Assays and Product Identification**

Prenyltransferase assays were performed as described previously (Orlova et al., 2009) except for the use of 40 μM [1-³H]-IPP instead of 14C-IPP. The reaction was performed in a final volume of 100 μl containing 40 μM[1-³H]-IPP (50 mCi mmol⁻¹) and 40 μM DMAPP in assay buffer (25 mM MOPS, pH 7.0, 10% [v/v] glycerol, 2 mM DTT, and 10 mM MgCl₂). For the identification of reaction products, larger-scale assays were performed in a final volume of 200 μl containing 80 μM-IPP and 80 μM DMAPP in assay buffer (25 mM MOPS, pH 7.0, 10% [v/v] glycerol, 2 mM DTT, and 10 mM MgCl₂). Assays were performed for 6 h at 30°C. To stop the assay and hydrolyze all diphosphate esters (including un-reacted substrate as well as products), 200 μl of a solution containing 2 units of bovine intestine alkaline phosphatase (18 units mg⁻¹; Sigma-Aldrich) and 2 units of potato aroylase (25.2 units mg⁻¹; Sigma-Aldrich) in 0.2 M Tris-HCl, pH 9.5, were added to samples followed by overnight incubation at 30°C. After enzymatic hydrolysis, the resulting prenyl alcohols were extracted with 1 ml hexane and the hexane fraction was concentrated to 25 μl. The products were separated on reversed-phase TLC plates (TLC Silica gel 60 RP-18 F₂₅₄S; MERCK, Germany). Chromatography was performed using a methanol:water (95:5 v/v) mobile phase, and spots were visualized by exposure of TLC plates to iodine vapor. Triplicate assays were performed for all data points. Enzyme assays were performed with one of following cations present in the assay buffer at final concentrations of 10 mM: Mg⁺², Mn⁺², or K⁺. All results represent an average of three independent assays.

**In Vivo Genetic Complementation Assay**

The pACCAR25ΔcrtE (Kainou et al., 1999) which contain the gene cluster crtX, crtY, crtI, crtb, and crtZ encoding carotenoid biosynthetic enzymes except crtE encoding GGPPS were used for determining in vivo GGPPS activity. The pET28a–CrGPPS constructs or HsGGPPS/pBH (as positive controls) and pACCAR25ΔcrtE were co-transformed into E. coli BL21 (DE3) with respective antibiotics (chloramphenicol for pACCAR25ΔcrtE; ampicillin for HsGGPPS and pBH; kanamycin for pET28a construct of CrGPPS.LSU, CrGPPS.SSU, and CrGPPS) and plated on LB plates to grow overnight at 37°C. A single colony from plates containing transformed colonies was picked and used for growth and induction as mentioned in the earlier section. To quantify carotenoids, pellets were obtained after centrifugation and dissolved in 90% (v/v) acetone to extract pigments. The concentration of carotenoid was determined by absorption at 450 nm using a Bio-Rad SmartSpec Spectrophotometer (Chang et al., 2010).

**Transient Overexpression of AmGGPPS.SSU in C. roseus Leaves by Agroinfiltration**

The plasmid pEF1.LIS–AmGGPPS.SSU (Orlova et al., 2009) containing the coding region of Snapdragon (A. majus) GGPS. SSU under C. breweri LIS promoter (1038 bp) (Orlova et al.,
2006) was used for transient expression in *C. roseus* leaves following the protocol of Long et al. (2009). Leaves from the third developmental stage of *C. roseus* were vacuum-infiltrated with *A. tumefaciens* strain GV3101 containing pEF1. LIS–AmGPPs.SSU and pEF1–LIS vector (control) for 15 min and kept for 36 h in the dark in 5% sucrose. The DNA-free total RNA from infiltrated leaves was isolated as described in the ‘Methods’ section. The expression of introduced AmGPPs.SSU was analyzed by semi-quantitative RT–PCR. Extraction of alkaloids from infiltrated leaves and determination of vindoline content by HPLC were carried out using the protocol reported earlier (Singh et al., 2000; Gupta et al., 2005). Briefly, shade-dried leaves (1 g) were powdered and extracted thrice with 90% ethanol (3 x 30 ml, 12 h each time). The alcohol extract was filtered, concentrated to 10 ml, then acidified with 10 ml of 3% HCl and washed thrice with 30 ml hexane. The aqueous portion was basified with ammonia to pH 8.5 and extracted thrice using 30 ml chloroform. The chloroform extract was washed with water, dried over sodium sulfate, and concentrated under vacuum. The residue was re-dissolved in methanol. Methanolic extract and pre-filtered solvents were subjected to HPLC analysis using a Shimadzu LC-8A gradient HPLC equipped with LC-8A pumps, manual injector valve, SPD-M10Avp PDA detector, and RP-18e reversed-phase chromolith performance HPLC column.

Bioinformatic Analysis

The predictions of protein subcellular localization were performed using PSORT (http://psort.hiroshima-u.ac.jp/), Predotar (http://urgi.versailles.inra.fr/predotar/predotar.html), ChloroP 1.1 (www.cbs.dtu.dk/services/ChloroP/), MitoProt (http://ihg.gsf.de/ihg/mitoprot.html), and TargetP 1.1 (www.cbs.dtu.dk/services/TargetP/).

Subcellular Localization Studies

The subcellular localization of CrGPPSs was studied by generating green fluorescent fusion proteins in-frame with *XbaI/BamHI* cloning sites of the p326–SGFP vector containing the CaMV 35S promoter. Full-length ORF of CrGPPSs and their truncated variants were amplified with Platinum Taq DNA polymerase High Fidelity (Invitrogen) using respective primers described in Supplemental Table 2. Amplified fragments were cloned in-frame with the 5’ end of the GFP coding sequence to generate the CrGPPSs–GFP fusion protein and then sequenced to confirm accuracy of fusions. Arabidopsis protoplasts were prepared and transformed as described previously (Sheen, 2002; Nagegowda et al., 2008). DNA from each construct (10 µg) was used for PEG-mediated transformation of 100 µl of ice-cold protoplasts. p326–SGFP and p326–RbTP–SGFP were used as controls for cytosolic and plastidial localization. Transient expression of GFP fusion proteins was observed 16–20 h after transformation using a LSM 510 Meta confocal microscope with a 405 laser-Carl Zeiss in the Central Imaging and Flow Cytometry Facility at the National Centre for Biological Sciences (NCBS), Bangalore, India.

Accession Numbers

The GenBank accession numbers for the sequences mentioned in this article are as follows: AgGPPS1, AF513111; AgGPPS2, AF513112; AgGPPS3, AF513113; AmGPPSSSU, AAS82859; AmGPPS.SSU, AAS82860; AtGPPPS6, At3g14530; AtGPPS, Y17376; AtSSU-II, At4g38460; CbGPPS.SSU, AYS53745; CrGPPS2, JX417186; CrGPPS3, JX417187; CrGPPS4, KC288140; CrGPPS, JX417185; CrGPPS.SSU, JX417184; HISSU, ACQ90681; HILLSU, ACQ90682; LeGPPS, DQ286930; MpGPPSSSU, AF182828; MpGPPS.SSU, AF182827; NtGPPS1, ADD9735; NtGPPS2, ADD9735; PaID51, ACZ55751; PaID52, ACA21458; PaID53, ACA21459; PbgPPS, EU023907; and QrtID51, CAC20852.

SUPPLEMENTARY DATA

Supplementary Data are available at Molecular Plant Online.

FUNDING

DAN thanks the Department of Biotechnology, Government of India, for the Ramalingaswami Fellowship (BT/HRD/35/24/2006). A.R. and A.K.S. are the recipients of a Research Fellowship from the University Grants Commission, and the Council of Scientific and Industrial Research, respectively.

ACKNOWLEDGMENTS

The authors are grateful to Dr C.S. Nautiyal, Director CSIR-CIMAP. D.A.N. expresses his sincere gratitude to Professor Ram Rajasekharan for his support and encouragement. Our special thanks go to Dr Ajit Shasany for sharing reagents. The authors are also thankful to Drs Natalia Dudareva, Inhwan Hwang, and Makoto Kawamukai for AmGPPS.SSU overexpression constructs, 326–SGFP, and pACCAR25ΔcrtE vectors, respectively. Confocal images were acquired at the Central Imaging and Flow Cytometry Facility, NCBS, Bangalore. We thank anonymous reviewers for their comments and suggestions. No conflict of interest declared.

REFERENCES


SUMMARY
This work suggests that, in Catharanthus roseus, the inactive small subunit of heteromeric GPP synthase (GPPS) interacts with bifunctional G(G)PPS and redirects the metabolic flux, thus acting as primary regulator of monoterpene indole alkaloid biosynthesis, whereas homomeric GPPS could be involved in ubiquinone formation.