Geranylgeranyl Diphosphate Synthases from *Scoparia dulcis* and *Croton sublyratus*. cDNA Cloning, Functional Expression, and Conversion to a Farnesyl Diphosphate Synthase

Naoe Колма,^{*a*} Worapan Sitthithaworn,^{*a*} Ekapop Viroonchatapan,^{*a*} Dae-Yeon Suh,^{*a*} Naoko Iwanami,^{*a*} Toshimitsu Hayashi,^{*a*} and Ushio Sankawa^{*,*a,b*}

Faculty of Pharmaceutical Sciences, Toyama Medical and Pharmaceutical University,^a 2630 Sugitani, Toyama 930–0194, Japan and International Traditional Medicine Research Center, Toyama International Health Complex,^b Toyama 939–8224, Japan. Received March 31, 2000; accepted May 9, 2000

cDNAs encoding geranylgeranyl diphosphate synthase (GGPPS) of two diterpene producing plants, Scoparia dulcis and Croton sublyratus, were isolated using the homology-based polymerase chain reaction method. Both cloned genes showed high amino acid sequence homology (60-70%) to other plant GGPPSs and contained highly conserved aspartate-rich motifs. The obtained clones were functionally expressed in Escherichia coli and showed sufficient GGPPS activity to catalyze the condensation of farnesyl diphosphate (FPP) and isopentenyl diphosphate to form geranylgeranyl diphosphate. To investigate the factor determining the product chain length of plant GGPPSs, S. dulcis GGPPS mutants in which either the small amino acids at the fourth and fifth positions before the first aspartate-rich motif (FARM) were replaced with aromatic amino acids or in which two additional amino acids in FARM were deleted were constructed. Both mutants behaved like FPPS-like enzymes and almost exclusively produced FPP when dimethylallyl diphosphate was used as a primer substrate, and failed to accept FPP as a primer substrate. These results indicate that both small amino acids at the fourth and fifth positions before FARM and the amino acid insertion in FARM play essential roles in product length determination in plant GGPPSs.

Key words geranylgeranyl diphosphate synthase; farnesyl diphosphate synthase; cDNA cloning; site-directed mutagenesis

Geranylgeranyl diphosphate synthase (GGPPS) supplies the essential acyclic precursor geranylgeranyl diphosphate (GGPP) for the biosynthesis of structurally diverse group of compounds including diterpenes. To date, cDNAs encoding GGPPS have been cloned from various organisms ranging from bacteria to higher eukaryotes. Comparison of the amino acid sequences has revealed that GGPPSs known to date contain five distinct regions with highly similar amino acid levels.1) Taking advantage of this similarity, a homologybased polymerase chain reaction (PCR) method was applied to obtain GGPPS cDNAs from two different diterpene-producing plants, Scoparia dulcis L. (Scrophulariaceae) and Croton sublyratus KURZ (Euphorbiaceae). S. dulcis is the source of unique tetracyclic diterpenes, scopadulcic acids A and B and scopadulciol,²⁾ whereas C. sublyratus accumulates a long-chain diterpene plaunotol³⁾ together with kaurene- and clerodane-type diterpenes.⁴⁾

To clone the genes, degenerate oligonucleotide primers were designed according to the regions in which amino acids are highly conserved among the known plant GGPPSs. Nested PCR with the same sets of primers successfully amplified the core fragments of GGPPS cDNAs from S. dulcis and C. sublyratus. Specific amplification of each clone using the rapid amplification of the cDNA end (RACE) method⁵⁾ was carried out to obtain the 5' and 3' termini of the cDNA. The open reading frames (ORFs) of the GGPPS genes from S. dulcis and C. sublyratus encode proteins of 368 and 351 amino acid residues, respectively. The two clones exhibit 70% amino acid identity to each other. Sequence comparison with other GGPPSs from Arabidopsis thaliana L., Capsicum annuum L., Catharanthus roseus G. DON, and Sinapis alba L. revealed a high level of similarity (60-70%) throughout the entire coding regions (Fig. 1). The cloned sequences also contain two conserved aspartate-rich motifs that have been reported to be important in substrate binding and catalysis.^{6,7)} Analysis of the amino acid sequences suggests that the amino terminal parts of GGPPSs cloned from S. dulcis and C. sublyratus likely encode a transit peptide for chloroplast import. The 50-60 amino terminal residues of the two clones exhibit a low degree of similarity, yet they share common features of plastidial transit peptides in that they are rich in serine, threonine, and small hydrophobic residues but contain few acidic residues.⁸⁾ Previous study of the production of diterpenes in leaf organ cultures of S. dulcis revealed that the concentration of diterpenes increased parallel to the differentiation of green leaves.⁹⁾ Furthermore, the labeling pattern of [1-¹³C] glucose incorporated into the diterpenes in shoot cultures¹⁰ demonstrated that they are biosynthesized via a mevalonate-independent pathway that has been shown to function in chloroplast.¹¹⁾ In C. sublyratus, an electron microscope study showed that the main diterpene plaunotol accumulated in the leaf.¹²⁾ These data suggest that the biosynthesis of these diterpenes is catalyzed by enzymes present in chloroplasts. As the cloned GGPPSs bear transit peptides for chloroplast imports, they are likely to be involved in the biosyntheses of these diterpenes in chloroplasts.

To confirm that the obtained clones encode GGPPSs, the ORFs were first PCR-amplified with gene-specific 5'- and 3'-flanking primers and subcloned into the Escherichia coli expression vector pET32a(+) (Novagen) to give pET-SDG-GPPS and pET-CSGGPPS for the S. dulcis and C. sublyratus GGPPSs, respectively. The plasmids were then transformed into E. coli AD494 (DE3)pLysS cells (Novagen). After isopropyl- β -D-thiogalactopyranoside (IPTG) induction, the crude protein extracts containing heterologously expressed thioredoxin-fusion GGPPS were used to test for enzyme activity according to the method described by Ogura et al.¹³⁾ Both pET-SDGGPPS and pET-CSGGPPS transformants exhibited sufficient GGPPS activity to catalyze the condensation of farnesyl diphosphate (FPP) and isopentenyl diphosphate (IPP) to form GGPP, whereas control E. coli failed to produce GGPP in noticeable amounts. Specific activities of S. dulcis and C. sublyratus GGPPSs determined from the radioactivity of GGPP were 71 and 66 pmol/h/mg, respectively, indicating that the obtained clones were catalytically active genes encoding GGPPSs.

Interestingly, in spite of higher amino acid sequence homology, the expression level of pET-SDGGPPS in *E. coli* was apparently higher than that of pET-CSGGPPS judging



Fig. 1. Alignment of Amino Acid Sequences of GGPP Synthases from Plant Origin

The plants from which GGPP synthases were obtained are indicated in the left column. Abbreviations and Genbank Accession Nos. are: Ara, *Arabidopsis thaliana*, Z99708; Cap, *Capsicum annuum*, X80267; Cat, *Catharanthus roseus*, X92893; Sin, *Sinapis alba*, X98795; and two plants from this work Sco, *Scoparia dulcis*, AB034250; and Cro, *Croton sublyrarus*, AB034249. Putative transit peptides are underlined. The broken lines indicate the first and second aspartate-rich motifs (FARM, Asp139-Asp145 in *S. dulcis* and SARM, Asp277-Asp281). Identical amino acid residues are boxed in black and similar residues are shaded in gray. Gaps are inserted to maximize homology.

from SDS-PAGE analysis of the crude extracts (data not shown). Earlier expression studies with *Taxus canadensis* GGPPS in yeasts¹⁴⁾ and spearmint limonene synthase in *E. coli*¹⁵⁾ demonstrated that a higher level of heterologous expression could be achieved upon removal of the transit sequences at the 5'-ends. In line with these studies, GGPPSs from *S. dulcis* and *C. sublyratus* were also expressed as 5'-truncated proteins.

Due to the lack of consensus sequence, the exact locations of the cleavage sites for chloroplast transit peptides could not be determined. Nonetheless, sequence alignment with other plant GGPPSs (Fig. 1) suggested putative cleavage sites to be located in the vicinity of Phe-58 and Phe-75 of S. dulcis GGPPS and C. sublyratus GGPPS, respectively.^{8,14)} Each truncated GGPPS was functionally expressed in E. coli by the same methods used for intact proteins. SDS-PAGE analysis showed prominent bands corresponding to the truncated S. dulcis GGPPS and truncated C. sublyratus GGPPS. As expected, the truncated proteins exhibited sufficient GGPPS activity to catalyze the formation of GGPP with FPP and IPP as substrates. Specific activities of truncated S. dulcis and C. sublyratus GGPPSs determined from the radioactivity of GGPP were 79 and 74 pmol/h/mg, respectively. The results suggested that for S. dulcis GGPPS and C. sublyratus GGPPS, the putative signal peptides at the amino terminal regions are not essential for the enzyme activity.

Farnesyl diphosphate synthase (FPPS) is a prenyl enzyme producing FPP (C_{15}) which is one isoprene unit shorter than GGPP (C_{20}). It has been shown that amino acids in the region

around the first aspartate-rich motif (DDXX(XX)D, FARM) plays an essential role in determining the product specificity of all FPPSs.¹⁶⁾ Comparison of amino acids around the FARMs of FPPSs and GGPPSs of plant origin revealed two different points. First, amino acids situated at the fourth and fifth positions before FARM in FPPSs are aromatic amino acids, whereas the corresponding residues are non-aromatic amino acids in GGPPSs. Second, FPPSs do not contain any amino acid insertion in FARM, yet GGPPSs contain two additional amino acids. To make mutants producing FPP instead of GGPP as a major product, three GGPPS mutants, mutant YF, mutant PC, and mutant YP, were constructed by site-directed mutagenesis. Due to its higher enzyme activity, *S. dulcis* GGPPS was chosen for this mutagenesis study.

In the mutant YF, the small amino acids Met and Ser were replaced with the aromatic amino acids Tyr and Phe. The mutant PC was also constructed by deleting the two additional amino acids Pro and Cys in FARM. Combination of these two mutants gave the mutant YP. Thus the mutant YP had aromatic amino acids at the fourth and fifth positions before FARM and did not have any amino acid insertion in FARM (Fig. 2A). With dimethylallyl diphosphate (DMAPP) and isopentenyl diphosphate (IPP) as substrates, all these GGPPS mutants produced FPP as the major product along with small amounts of geranyl diphosphate (GPP) (Fig. 2B). On the other hand, the mutants were not able to utilize FPP and therefore no major prenyl compounds could be detected when FPP was used as a primer substrate (Fig. 2C). The radioactivities of GPP, FPP, and GGPP in the reaction of the



Fig. 2. Amino Acid Sequences around FARMs of Wild-Type and Mutant S. dulcis GGPPSs and A. thaliana FPPS

Mutated amino acid residues are underlined and FARMs are boxed (A). TLC radiochromatograms of the alcohols obtained by enzymatic hydrolysis of the products produced by the GGPPS mutants. The enzyme activities of the mutants were measured with [1-¹⁴C]IPP and either DMAPP (B) or FPP (C) as the primer substrate and the reaction products were analyzed by reversed-phase LKC-18 TLC. Lane 1, untransformed pET vector; lane 2, wild-type GGPPS; lane 3, mutant YF; lane 4, mutant PC; lane 5, mutant YP. Spots of authentic alcohols: GOH, geraniol; FOH, (all-*E*)-farnesol; GGOH, (all-*E*)-geranylgeraniol. ori., origin; s.f., solvent front.

 Table 1. Relative Activity of the Mutated GGPPS to Wild-Type GGPPS

 Determined from Radioactivities of GPP, FPP, and GGPP

Enzyme	Relative activity (%)		
	GPP	FPP	GGPP
Wild-type	100	100	100
Mutant YF	340	340	7
Mutant PC	200	129	2
Mutant YP	320	278	8

Amounts of GPP, FPP, and GGPP formed in the reaction of wild-type GGPPS were taken as 100%.

mutant YF, mutant PC, and mutant YP based on the specificity of the wild-type GGPPS are shown in Table 1. Relative formation of GPP, FPP, and GGPP fell in the range of 200-340%, 120-340% and 2-8%, respectively. These results clearly demonstrated that either: 1) replacement of small amino acids at the fourth and fifth positions before FARM with aromatic amino acids or 2) deletion of two amino acids in the FARM of S. dulcis GGPPS was sufficient to convert a plant GGPPS into a FPPS. A previous mutagenesis study around the FARM region of archaeal bacteria Sulfolobus acidocaldarius GGPPS¹⁶ showed that conversion of Thr situated at the fourth position before FARM to Phe was sufficient to allow the GGPPS produce FPP as a major product. Unlike plant GGPPSs, archaeal GGPPS lacks amino acid insertion in FARM. This suggests that plant and archaeal GGPPSs employ different mechanisms to control the chain length of the products. In plant GGPPSs, both small amino acids at the fourth and fifth positions before FARM and the insertion in FARM play essential roles in product length determination. On the other hand, in archaeal GGPPS, having a bulky aromatic amino acid at the fifth and a small amino acid at the fourth position before FARM was essential in chain length determination. This is the first report of successful conversion of a plant GGPPS into an enzyme producing a shorter prenyl compound by site-directed mutagenesis.

Acknowledgment This work was supported by a Grants-in-Aid for Scientific Research (B) (No. 09044212) from the Ministry of Education, Science, Sports and Culture, Japan.

References and Notes

- 1) Chen A., Kroon P. A., Poulter C. D., Protein Sci., 3, 600-607 (1994).
- Hayashi T., Kishi M., Kawasaki M., Shimizu M., Suzuki S., Yoshizaki M., Morita N., Tezuka Y., Kikuchi T., Berganza L. H., Ferro E., Basualdo I., *Tetrahedron Lett.*, 28, 3693–3696 (1987).
- Ogiso A., Kitazawa E., Kurabayashi M., Sato A., Takahashi S., Noguchi H., Kuwano H., Kobayashi S., Mishima H., *Chem. Pharm. Bull.*, 26, 3117–3123 (1978).
- 4) Kitazawa E., Ogiso A., *Phytochemistry*, **20**, 287–289 (1981) and references cited therein.
- Frohman M. A., Dush M. K., Martin G. R., Proc. Natl. Acad. Sci. U.S.A., 85, 8998—9002 (1988).
- 6) Song L., Poulter C. D., Proc. Natl. Acad. Sci. U.S.A., 91, 3044–3048 (1994).
- Tarchis L. C., Proteau P. J., Kellogg B. A., Sacchettini J. C., Poulter C. D., Proc. Natl. Acad. Sci. U.S.A., 93, 15018–15023 (1996).
- von Heijne G., Steppuhn J., Herrmann G. R., *Eur. J. Biochem.*, 180, 535–545 (1989).
- Hayashi T., Kasahara K., Sankawa U., *Phytochemistry*, 46, 517–520 (1997).
- Hayashi T., Asai T., Sankawa U., *Tetrahedron Lett.*, 40, 8239–8243 (1999).
- Lichtenthaler H. K., Schwender J., Disch A., Rohmer M., *FEBS Lett.*, 400, 271–274 (1997).
- 12) De-Eknamkul W., personal communication.
- Ogura K., Nishino T., Shinka T., Seto S., *Methods Enzymol.*, 110, 167–171 (1985).
- 14) Hefner J., Ketchum E. B., Croteau R., Arch. Biochem. Biophys., 360, 62—74 (1998).
- Williams D. C., McGarvey J. D., Katahira E. J., Croteau R., *Biochemistry*, 37, 12213—12220 (1998).
- 16) Ohnuma S., Hirooka K., Ohto C., Nishino T., J. Biol. Chem., 272, 5192—5198 (1997).