Geranylgeranyl Diphosphate Synthase from *Scoparia dulcis* and *Croton sublyratus*. Plastid Localization and Conversion to a Farnesyl Diphosphate Synthase by Mutagenesis

Worapan SITTHITHAWORN,[a] Naoe KOJIMA,[a] Ekapop VIROONCHATAPAN,[a] Dae-Yeon SUH,[a] Naoko IWANAMI,[a] Toshimitsu HAYASHI,[a] Masaaki NOJI,[b] Kazuki SAITO,[b] Yasuo NIWA,[c] and Ushio SANKAWA*,[a,d]

Faculty of Pharmaceutical Sciences, Toyama Medical and Pharmaceutical University,[a] Toyama 930–0194, Japan, Laboratory of Molecular Biology and Biotechnology, Research Center of Medicinal Resources, Faculty of Pharmaceutical Sciences, Chiba University,[b] Chiba 263–8522, Japan, Laboratory of Plant Cell Technology, Graduate School of Nutritional and Environmental Sciences, University of Shizuoka,[c] Shizuoka 422–8526, Japan, and International Traditional Medicine Research Center, Toyama International Health Complex,[d] Toyama 939–8224, Japan.
Received September 18, 2000; accepted October 12, 2000

cDNAs encoding geranylgeranyl diphosphate synthase (GGPPS) of two diterpene-producing plants, *Scoparia dulcis* and *Croton sublyratus*, have been isolated using the homology-based polymerase chain reaction (PCR) method. Both clones contained highly conserved aspartate-rich motifs (*DDXX*(*XX*)*D*) and their N-terminal residues exhibited the characteristics of chloroplast targeting sequence. When expressed in *Escherichia coli*, both the full-length and truncated proteins in which the putative targeting sequence was deleted catalyzed the condensation of farnesyl diphosphate and isopentenyl diphosphate to produce geranylgeranyl diphosphate (GGPP). The structural factors determining the product length in plant GGPPSs were investigated by constructing *S. dulcis* GGPPS mutants on the basis of sequence comparison with the first aspartate-rich motif (FARM) of plant farnesyl diphosphate synthase. The result indicated that in plant GGPPSs small amino acids, Met and Ser, at the fourth and fifth positions before FARM and Pro and Cys insertion in FARM play essential roles in determination of product length. Further, when a chimeric gene comprised of the putative transit peptide of the *S. dulcis* GGPPS gene and a green fluorescent protein was introduced into *Arabidopsis* leaves by particle gun bombardment, the chimeric protein was localized in chloroplasts, indicating that the cloned *S. dulcis* GGPPS is a chloroplast protein.

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

Geranylgeranyl diphosphate synthase (GGPPS), a member of the short-chain isoprenyl diphosphate synthase family, catalyzes the consecutive condensation of three molecules of isopentenyl diphosphate (IPP) with dimethylallyl diphosphate (DMAPP) to give a C_{20} compound which is a precursor to diterpenes, carotenoids and chlorophylls.[1,2] Meanwhile, farnesyl diphosphate synthase (FPPS) produces a C_{15} compound which is utilized as a precursor in the biosynthesis of sesquiterpenes, steroids and farnesylated proteins, and geranyl diphosphate synthase (GPPS) produces geranyl diphosphate (C_{10}), the key precursor of monoterpene biosynthesis.[3] Genes encoding GPPS, FPPS and GGPPS have been isolated from various organisms. Sequence comparison of these enzymes revealed several conserved regions including two aspartate-rich *DDXX*(*XX*)*D* motifs.[4—6]

X-ray crystallography study of an avian FPPS revealed that these two aspartate rich-motifs, which are located on opposite walls of a large central active-site cavity, act as binding sites for the metal ions that form complexes with the diphosphate moieties of the two substrates.[7] Site-directed mutagenesis and crystallographic studies also indicated that the amino acids in the region around the first aspartate-rich motif (FARM) are essential in determining the ultimate length of the product polyisoprenoid chain. Hence, replacement of Phe-112 and Phe-113, situated at the fifth and fourth amino acid upstream to FARM in the avian FPPS, with smaller side chain amino acids gave FPPS mutants that synthesized longer polyisoprenoid products than FPP.[8] Further, in both cases of bacillus FPPS[9] and archaeal GGPPS,[10] replacements of the aromatic amino acid situated at the fifth amino acid before FARM to a non-aromatic amino acid caused a change of the product specificity. The mutant enzymes were able to catalyze consecutive condensations beyond the destined limit of the original enzymes, and the chain lengths of their final products were dependent on the bulk of the side chain of the substituted amino acid.[11] On the basis of these mutation studies,[12] it was proposed that the mechanism of the product length determination in eukaryotic FPPS is different from those of prokaryotic FPPS. In eukaryotic FPPS, both of the aromatic amino acids at the fourth and fifth positions before FARM directly block the chain elongation greater than C_{15} and provide product specificity. In contrast, the product specificity of prokaryotic FPPS is determined by an aromatic amino acid at the fifth position before FARM and two amino acids inserted in FARM.

Few studies, however, have been reported concerning what structural factor(s) determines the product chain length in plant GGPPSs. To this end, GGPPSs from two diterpene-producing plants, *Scoparia dulcis* and *Croton sublyratus*, were cloned, functionally expressed in *Escherichia coli*,[13] and the mutants were analyzed for product specificity. A comparison of amino acids around the FARMs of FPPSs and GGPPSs of plant origin revealed two distinctive points. First, amino acids situated at the fourth and fifth positions upstream to FARM in plant FPPSs are aromatic, whereas the corresponding residues are non-aromatic amino acids in plant GGPPSs.

* To whom correspondence should be addressed. e-mail: sankawa@toyama-pref-ihc.or.jp

Second, plant FPPSs do not contain any amino acid insertion in FARM, yet plant GGPPSs contain two additional amino acids. On the basis of this sequence information, *S. dulcis* GGPPS mutants were constructed by replacing the small amino acids at the fourth and fifth positions before FARM with aromatic amino acids or by deleting two additional amino acids in FARM.

In addition, the localization of *S. dulcis* GGPPS into plastids was demonstrated by a transient expression of fusion protein comprised of *S. dulcis* N-terminal region with jellyfish green fluorescent protein (GFP)[14] in *Arabidopsis* leaves.

Experimental

Cloning of *S. dulcis* and *C. sublyratus* GGPPS cDNAs Total RNA was isolated from two week-old seedlings of *Scoparia dulcis* and young leaves of *Croton sublyratus*, respectively, by phenol-SDS procedure and lithium chloride precipitation. Single-strand cDNAs were synthesized from 10 μg of total RNA with an oligo$(dT)_{17}$ primer and reverse transcriptase (Superscript II, BRL) according to the manufacturer's protocol. After RNase H treatment, the resulting single-strand cDNA mixtures were used as templates for polymerase chain reaction (PCR).

The same sets of degenerate oligonucleotide primers corresponding to the highly conserved amino acid sequences of the known plant GGPPSs (Fig. 1) were used for PCR-amplification of the core cDNA fragments. The first PCR was performed with the external set of primers GGPP433S and GGPP1085A (Table 1) using Ex-Taq DNA polymerase (Takara). The PCR program used throughout this study was 94 °C, 1 min, 43 °C, 1 min, 72 °C, 2 min for 30 cycles and final extension at 72 °C, 10 min. After filtration with Suprec-02 filter (Takara) to remove the primers, a nested PCR was performed with the internal set of primers, GGPP464S and GGPP952A, under the same PCR conditions except that the annealing temperature (T_a) was 50 °C. The approx. 470 bp PCR products thus obtained were subcloned into pT7Blue Blunt or T-vector (Novagen) and the resulting plasmids were used to transform *Escherichia coli* strain NovaBlue (Novagen). DNA sequencing was performed with the dRhodamine Terminator Cycle Sequencing Ready Reaction kit (Perkin-Elmer).

The 3′-end of cDNAs was amplified using the rapid amplification of cDNA ends (RACE) method.[15] Specific sense primers were synthesized for both *S. dulcis* and *C. sublyratus* GGPPSs based on the core sequences obtained (Table 1). The sense primers were: for *S. dulcis* GGPPS, 3′SD-S1 for the first PCR and 3′SD-S2 for the second PCR; for *C. sublyratus* GGPPS, 3′CS-S1 for the first PCR and 3′CS-S2 for the second PCR. An anchor primer, RACE17 was used as the antisense primer. Likewise, 5′-RACE was performed with a 5′-RACE system of Gibco BRL following the manufacturer's protocol. Template *S. dulcis* cDNAs were prepared with Superscript II reverse transcriptase (50 min at 42 °C) and 5′SD-A1 as a primer. Poly(C) tail was then added to the 3′-ends of the resultant DNA mixture to create an abridged anchor primer (AAP) binding site. The first and second PCR reactions for *S. dulcis* GGPPS were performed with primers 5′SD-A2 and AAP, and with primers of 5′SD-A3 and the abridged universal amplification primer (AUAP), respectively. For *C. sublyratus* GGPPS, the primer 5′CS-A1 was used during reverse transcription. The first and second PCR reactions were performed with primers 5′CS-A2 and AAP, and with primers of 5′CS-A3 and AUAP, respectively. The nucleotide sequences obtained in this study have been deposited in the EMBL/GenBank/DDBJ databank (*S. dulcis* GGPPS, AB034250; *C. sublyratus* GGPPS, AB034249).

Heterologous Expression of Intact and Truncated GGPPSs *S. dulcis* and *C. sublyratus* GGPPS genes were expressed as both intact proteins and truncated proteins in which putative transit peptides were removed. First, the open reading frames (ORF) of the cloned cDNAs were PCR-amplified with 5′- and 3′-flanking primers (Table 1). For intact *S. dulcis* GGPPS, the 5′- and 3′-flanking primers were SD-S1 and SD-A1051, respectively; for intact *C. sublyratus* GGPPS, they were CS-S1 and CS-A1104, respectively. Restriction sites were introduced immediately upstream of the ATG codon and downstream of the stop codon (Table 1). Similarly, cDNAs for truncated GGPPSs were PCR-amplified with corresponding full-length cDNAs as templates. For each truncated GGPPS, a new 5′-flanking primer was designed in such a way as to remove the putative transit peptide at the 5′-end. Thus, for the truncated *S. dulcis* GGPPS, SD-S167 indicating the truncation site at nucleotide 167 in the ORF was used together with SD-A1051 during the PCR amplification. Likewise, for the truncated *C. sublyratus* GGPPS, primers CS-S216 and CS-A1104 were used. Therefore, the truncated *S. dulcis* GGPPS starts with ^{57}Tyr at its N-terminal, whereas the truncated *C. sublyratus* GGPPS starts with ^{73}Leu at its N-terminal.

Table 1. Oligonucleotide Primers Used in This Study

Primer	Sequence (5′→3′) (T_a, °C)
	Core fragments
GGPP433S	GARATGATICAYACYATGTC (^{145}EMIHTM) (43)
GGPP464S	GAYYTICCYTGIATGGAYAAY (^{155}DLPCMD) (50)
GGPP952A	AARTCYTTCCCWGCRGTTTTYC (^{310}GKTAGKD) (50)
GGPP1085A	ARWGCMAWYAAIGGAGCA (^{361}APLLALA) (43)
	3′-RACE
3′SD-S1	^{481}GGAGAGGACGTCGCTGTCCTCGC (64)
3′SD-S2	^{500}TCGCCGGCGACGCTCTTCTT (56)
3′CS-S1	^{502}CGTGGTAAGCCGACGAATCA (55)
3′CS-S2	^{548}TTCTTGCAGGTGATGCGCTTCT (58)
RACE17	GACTCGAGTCGACATCGA
	5′-RACE
5′SD-A1	^{851}TTAGTAACATCAAGAATATCATCCACCACC (42)
5′SD-A2	^{821}TGAAACAACAGCCCAATACACCTGGCAAACTTCC (55)
5′SD-A3	^{519}AAGAAGAGCGTCGCCGGCGA (52)
5′CS-A1	^{660}AATCGCCTTAGCCAATTCA (42)
5′CS-A2	^{633}CCGAACAATTCTGACAGGAGAAAC (52)
5′CS-A3	^{572}GCAAGAAGCGCATCACCTGCAAGA (50)
AAP	GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG
AUAP	GGCCACGCGTCGACTAGTAC
	Heterologous expression
SD-S1	$^{-15}$AGACCAACAGGATCC*ATG*AGCCTT (*Bam*H I) (50)
SD-S167	^{160}ACCTCCGGATCCTTCGACTTCAAGAAGTAC*ATG*CTC (*Bam*H I) (50)
SD-A1051	^{1071}CATCCTCCAAAGCTT*TCA*ATTCTCCC (*Hin*d III) (50)
CS-S1	$^{-9}$AAAGCAA*CCATGG*GTTCTGTGAATTTAGGTTC (*Nco* I) (50)
CS-S216	^{210}TTCTT*CCATGG*GTTTTGATTTCAAATCTTACATGG (*Nco* I) (50)
CS-A1104	^{1120}GTACTACGGATCC*TTA*GTTTTGCCTG (*Bam*H I) (50)
	Site directed mutagenesis
SD-AYF	^{420}GTCGTCGTGCATGAG**GTAGAA**GGTGTGGATCAT (50)
SD-APC	^{447}GAGATGGTCGTTGTCCATGAGGTCGTCGTGCATGAG (50)
SD-SPC	^{406}CTCATGCACGACGACCTCATGGACAACGACCATCTC (50)
	Chloroplast targeting
SD-SGFP	$^{-15}$AGACCAACAGTCGAC*ATG*AGCCTT (*Sal* I) (55)
SD-AGFP	^{264}GTGGATCTCCATGGGTTCTTTGAGGGA (*Nco* I) (55)

The amino acid sequences corresponding to the core fragment primers are those of *A. thaliana* GGPPS. The restriction enzyme sites are underlined, the 5′-start and 3′-stop codons are in italics, and mutated nucleotides in the SD-AYF primer are in bold. The primers SD- and CS- are for *S. dulcis* and *C. sublyratus*, respectively. The number in parenthesis indicates annealing temperature (T_a) during PCR.

The amplified PCR products were gel-purified, digested with restriction enzymes and subcloned into an *E. coli* expression vector, pET32a(+) (Novagen), which gives the expressed protein as a thioredoxin-His-tagged fusion protein. *E. coli* AD494 (DE3)pLysS cells (Novagen) transformed with expression plasmids were grown and overexpression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 20 h at 25 °C according to the manufacturer's protocol. Soluble fractions obtained following the manufacturer's recommendations were directly used as enzyme sources and the protein concentrations were determined with the Bio-Rad protein microassay using bovine serum albumin as standard.[16]

Site-Directed Mutagenesis of *S. dulcis* GGPPS The synthetic mutagenic primers used for mutant construction were SD-AYF, SD-APC, and SD-SPC (Table 1). Mutant YF (Met134Tyr/Ser135Phe) was constructed by the PCR megaprimer method[17] using the full-length *S. dulcis* GGPPS clone as template. After the first PCR with primers SD-S1 and SD-AYF, a second PCR was performed with SD-A1051 and the first PCR products (megaprimer), giving the full-length mutant clone. For the construction of mutant PC (ΔPro142/ΔCys143), a PCR was performed with primers SD-S1 and SD-APC and the full-length *S. dulcis* GGPPS clone as template. A separate PCR was also performed with primers SD-SPC and SD-A1051. Then, the resulting PCR products were combined and used as template for the next

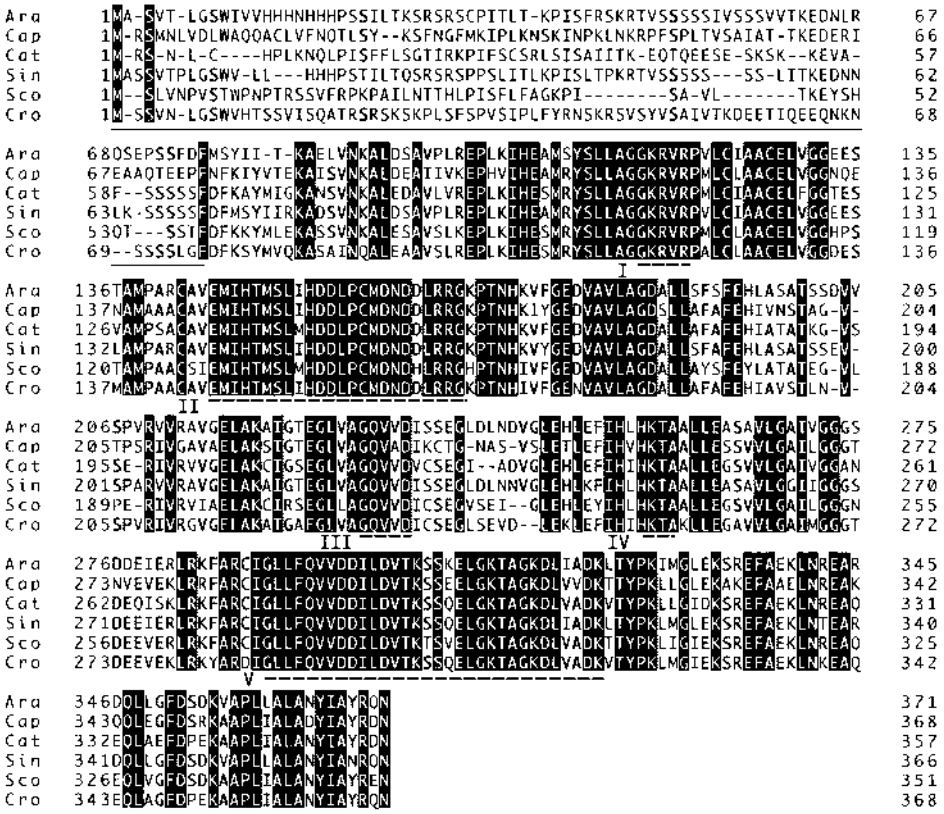

Fig. 1. Alignment of Amino Acid Sequences of Plant GGPP Synthases

Abbreviations and Genbank accession no. are: Ara, *Arabidopsis thaliana*, Z99708; Cap, *Capsicum annuum*, X80267; Cat, *Catharanthus roseus*, X92893; Sin, S*inapis alba*, X98795; Sco, *Scoparia dulcis*, AB034250; Cro, *Croton sublyrarus*, AB034249. Putative transit peptides are underlined. The broken lines indicate the highly conserved domains, I—V. The first and second aspartate-rich motifs (FARM, Asp139—Asp145 and SARM, Asp277—Asp281 in *S. dulcis*) are located in domains II and V, respectively. Identical amino acid residues are boxed in black and similar residues are shaded in gray. Gaps are inserted to maximize homology.

PCR with SD-S1 and SD-A1051 as primers to give a full-length DNA fragment with desired deletions. Mutant YP (Met134Tyr/Ser135Phe/ΔPro142/ΔCys143) was constructed using the same PCR-based method as for the mutant PC except that the full-length mutant YF DNA fragment was used as template for the first two PCR reactions. The mutagenesis was confirmed by DNA sequencing. After subcloning into pET32a(+), the mutant expression plasmids were used to transform *E. coli* AD494(DE3)pLysS cells for functional expression.

GGPP Synthase Assay and Product Analysis The incubation mixture (1 ml) contained 50 mM iodoacetamide, 5 mM $MnCl_2$, 50 μM FPP, 25 μM [4-^{14}C]IPP (1.85 kBq) and *E. coli* soluble fraction (800 μg total protein) as the enzyme source. After incubation at 37 °C for 3 h, the reaction was stopped by alkaline phosphatase hydrolysis.[18] The reaction products were analyzed by TLC on a reverse-phase LKC 18 silica gel 60 plate (Whatman) developed with acetone/water (9 : 1). The TLC plate was exposed on a Fuji film BAS-II imaging plate at room temperature and analyzed with a Fuji BAS 2000 Mac bioimaging analyzer. Authentic geraniol, (all-*E*)-farnesol and (all-*E*)-geranylgeraniol were used to identify the reaction products.

Construction of a Chimeric Gene Encoding a Fusion Protein of the Putative Transit Peptide of *S. dulcis* GGPPS with Green Fluorescent Protein (GFP) A 252-bp fragment of the N-terminal region of the *S. dulcis* GGPPS cDNA was PCR-amplified with SD-SGFP and SD-AGFP as sense and antisense primers, respectively (Table 1). The resulting PCR products were subcloned into a GFP expression vector, CaMV35S-sGFP(S65T)-NOS3′ which contains a cauliflower mosaic virus ^{35}S-labeled promoter as its 5′-end, GFP as a reporter protein and nopaline synthetase terminator as a transcription terminator.[14] The resulting plasmid was subjected to particle gun bombardment on *Arabidopsis* tissues and individual leaves were visualized with a fluorescent microscope as described previously.[19]

Results and Discussion

Cloning of *S. dulcis* and *C. sublyratus* GGPPS cDNAs The ORF of the GGPPS gene from *S. dulcis* encodes a protein of 351 amino acid residues whereas the ORF of *C. sublyratus* GGPPS gene encodes a protein of 368 amino acid residues. The two clones exhibit 70% amino acid identity to each other. Sequence comparison with other plant GGPPSs from *A. thaliana*, *C. annuum*, *C. roseus*, and *S. alba* revealed a high level (60—70%) of similarity throughout the entire coding regions. However, sequence similarity with procaryotic GGPPSs from *Thermus thermophilus*, *Archaeoglobus fulgidus*, *Saccharomyces cerevisae* and *Gibberella fujikuroi* was low (15—30%). Five regions with highly conserved residues,[6] designated I—V, were found in both clones (Fig. 1). Regions II and V are rich in negatively charged aspartate and positively charged arginine or lysine residues. These sequences correspond to the first and second aspartate-rich motifs (FARM and SARM, respectively) which have been proposed to be the diphosphate binding sites.[4]

N-terminal Transit Peptide Domain Sequence alignment (Fig. 1) revealed that the N-terminal sequences of plant GGPPSs including *S. dulcis* and *C. sublyratus* GGPPSs exhibited a low degree of similarity, yet they shared common features of chloroplast transit peptides[20] in that 1) the N-terminal domains of 10 amino acids of the putative transit sequences were uncharged amino acids; 2) the central regions (amino acids 11—43 in *S. dulcis* GGPPS and amino acids 11—46 in *C. sublyratus* GGPPS) contained high amounts of the hydroxy amino acids serine and threonine and contained the positively charged amino acids lysine, arginine and histidine; 3) the C-terminal parts had a high potential to form amphiphilic β-strands. It is also noteworthy that a computer program (TargetP, www.cbs.dtu.dk/services/TargetP-1.0/), which

predicts the subcellular localization and its cleavage site of a given protein by a computational neural network-based method,[21] predicted that both *S. dulcis* and *C. sublyratus* GGPPSs would be probably localized in chloroplasts.

Functional Expression of Intact and Truncated GGPPSs in *E. coli* To confirm that the obtained clones encode catalytically active GGPPSs, the cloned genes were heterologously expressed as thioredoxin-fusion proteins in *E. coli*. Both transformants exhibited significant GGPPS activity to catalyze the condensation of FPP and IPP to form GGPP, whereas control *E. coli* cells transformed with pET32a(+) failed to produce GGPP in noticeable amounts (Fig. 2B). Specific activities of *S. dulcis* and *C. sublyratus* GGPPSs determined from the radioactivity of GGPP were 71 and 66 pmol/h/mg, respectively. Interestingly, in spite of higher amino acid sequence identity, the expression level of *S. dulcis* GGPPS in *E. coli* was apparently higher than that of *C. sublyratus* GGPPS judging from sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of the crude extracts (Fig. 2A). A prominent protein band of 49 kDa corresponding to the thioredoxin-fusion protein was observed with *S. dulcis* GGPPS transformed *E. coli* extracts.

Earlier expression studies with *Taxus canadensis* GGPPS in yeasts[22] and spearmint limonene synthase in *E. coli*[23] demonstrated that a higher level of heterologous expression could be achieved upon removal of the plastidial transit sequences at the N-ends. In line with these studies, GGPPSs from *S. dulcis* and *C. sublyratus* were also expressed as mature (truncated) proteins in which the putative transit peptides at the N-ends were removed. It has been suggested that a change in secondary structures between a β-sheet and a more helical structure marks a transit peptide cleavage site.[20] Sequence alignment (Fig. 1) and predicted secondary structures (with PREDATOR)[24] suggested the putative cleavage sites are located in the vicinity of Phe-58 and Phe-75 of *S. dulcis* GGPPS and *C. sublyratus* GGPPS, respectively. When expressed as thioredoxin-fusion proteins in *E. coli*, both truncated *S. dulcis* and *C. sublyratus* GGPPSs (approx. 44 kDa) were expressed in higher yields than the intact enzymes as judged by SDS-PAGE (Fig. 2A). The truncated proteins also exhibited GGPPS activity indicating that the putative transit peptides at the amino terminal regions are not essential for the enzyme activity (Fig. 2B). Specific activities of the truncated *S. dulcis* and *C. sublyratus* GGPPSs determined from the radioactivity of GGPP were 79 and 74 pmol/h/mg, respectively.

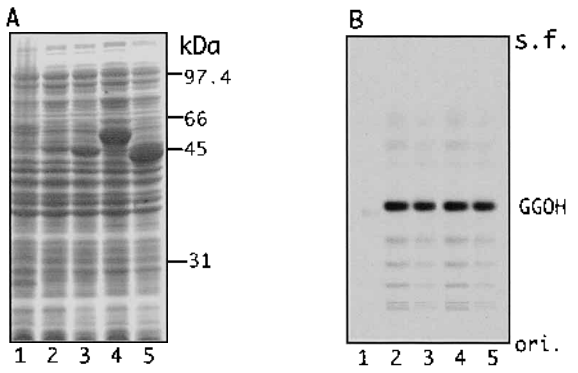

Fig. 2. SDS-PAGE (A) and Radio TLC (B) of the *E. coli* Extracts Containing the Full-Length and Truncated GGPPSs from *S. dulcis* and *C. sublyratus*

(A) Proteins of soluble *E. coli* fractions were separated on 12% (w/v) polyacrylamide mini-slab gels under reducing conditions followed by staining with Coomassie Brilliant Blue R250. The positions of molecular mass markers (in kDa) are shown at the right. (B) Radio-TLC of the GGPPS reaction products. The enzyme activities of the intact and truncated GGPPSs were determined with FPP and [4-^{14}C]IPP as substrates. The reaction product ((all-*E*)-geranylgeraniol, GGOH) was identified using authentic sample. ori., TLC origin; s.f., solvent front. In both (A) and (B), lane 1, control *E. coli* transformed with the pET32a(+) vector; lane 2, full-length *C. sublyratus* GGPPS; lane 3, truncated *C. sublyratus* GGPPS; lane 4, full-length *S. dulcis* GGPPS; lane 5, truncated *S. dulcis* GGPPS.

Subcellular Localization of *S. dulcis* GGPPS To determine the subcellular localization of *S. dulcis* GGPPS, a localization study was performed using jellyfish green fluorescent protein carrying the S65T mutation, sGFP(S65T)[14] as a reporter protein. A 252-base pair PCR-amplified fragment encoding 57 amino acids of the putative transit peptide and an additional 27 amino acids of the *S. dulcis* GGPPS gene was fused in frame to the N-terminus of sGFP(S65T). The fusion protein GGPPS-sGFP(S65T) was detected in intact tissues after delivering the construct into *Arabidopsis* leaves by particle gun bombardment. The observed localization patterns of the construct containing the putative *S. dulcis* GGPPS transit peptide (Fig. 3A) were similar to those exhibited by the transit peptide of the ribulose-1,5-bisphosphate carboxylase small subunit polypeptide of *Arabidopsis* (35SΩ-TP-sGFP(S65T)) (Fig. 3B), which had previously been demonstrated to target chloroplasts.[19] The patterns of the plastidic proteins were distinct from the patterns observed with the mitochondrion-localizing isoform C of spinach cysteine synthase (CS-C-GFP)[19] (Fig. 3C), and also distinct from those observed with the control vector, sGFP(S65T), containing no transit peptide (Fig. 3D). These results confirmed that the N-terminal sequence of *S. dulcis* GGPPS encodes a transit signal for targeting chloroplasts and that *S. dulcis* GGPPS is a plastidic protein.

Earlier study of the production of diterpenes in *S. dulcis* leaf organ cultures revealed that the concentration of the diterpenes increased in parallel to differentiation of green leaves.[25] Furthermore, the labeling pattern of [1-^{13}C] glucose incorporated into the diterpenes in *S. dulcis* shoot cultures[26]

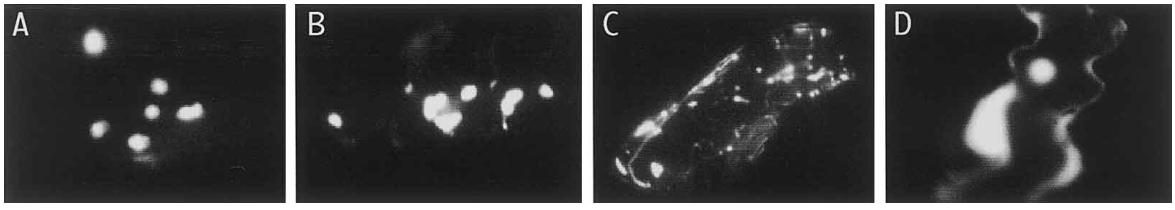

Fig. 3. Subcellular Localization of GFP(S65T) Containing the N-Terminal Putative Transit Peptide of *S. dulcis* GGPPS

Panels show fluorescence signal pattern exhibited by (A) GGPPS-sGFP(S65T), (B) 35SΩ-TP-sGFP(S65T) as the positive control of plastid targeting, (C) CS-C-GFP as the positive control of mitochondria targeting, and (D) sGFP(S65T) as the positive control for localization in cytosol and partly in nucleus.

demonstrated that they are biosynthesized *via* a non-mevalonate pathway which has been shown to function in chloroplasts.[27] Taken together, these results suggest that *S. dulcis* GGPPS is involved in the diterpene biosynthesis in chloroplasts.

In *C. sublyratus*, an electron microscope study by De-Eknamkul and Potduang[28] showed that the main diterpene, plaunotol (18-hydroxy geranylgeraniol)[29] was accumulated in the leaf chloroplasts (Fig. 4), suggesting that the plaunotol biosynthesis is catalyzed by the enzymes present in chloroplasts. As *C. sublyratus* GGPPS contains the putative chloroplast targeting sequence, it is likely to be involved in the biosynthesis of plaunotol in chloroplasts.

Conversion of *S. dulcis* GGPPS to a FPPS by Mutagenesis On the basis of the difference in plant GGPPS and FPPS sequences (see the Introduction section), three *S. dulcis* GGPPS mutants were constructed, namely mutant YF, mutant PC and mutant YP. The mutant YF contained Met and Ser instead of Tyr and Phe at the fourth and fifth positions upstream to FARM. The mutant PC was 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 5A).

When the reactions were carried out with crude soluble fractions and DMAPP (25 μM) and [1-^{14}C]IPP (25 μM) as substrates, similar product distributions were observed in all mutants (Fig 5B). Clearly, FPP was the major product produced by all mutants (Table 2). The amounts of FPP produced by the mutant YF and mutant YP were more than two times higher than that produced by the mutant PC. When the reaction was carried out using FPP (25 μM) as a primer substrate instead of DMAPP, the mutants were not able to utilize FPP and no major prenyl compounds could be detected, whereas the wild-type enzyme produced GGPP from FPP as expected (Fig 5C). These data clearly indicated that either: 1) replacement of small amino acids at the fourth and fifth positions upstream to FARM by aromatic amino acids or 2) deletion of two amino acids in FARM of *S. dulcis* GGPPS was sufficient to convert a plant GGPPS to a FPPS.

Earlier mutagenesis studies of bacillus FPPS[10] and archaeal GGPPS[9,12] showed that the ultimate length of their products is determined by the amino acid at the fifth position upstream to FARM. Both bacillus FPPS and archaeal GGPPS possess an aromatic amino acid of Tyr and Phe, respectively, at that position. Unlike the plant GGPPSs, archaeal GGPPS lacks an amino acid insertion in FARM. Mutagenesis around the FARM region of archaeal GGPPS[12] demonstrated that GGPPS could produce FPP as a major product by conversion

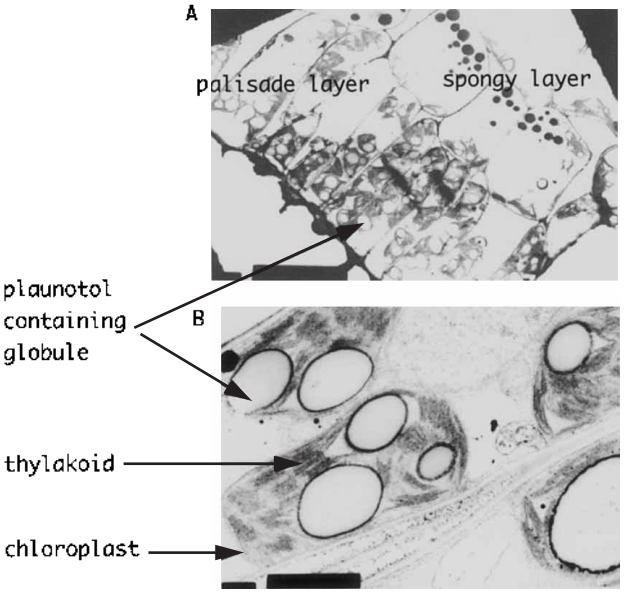

Fig. 4. Electron Micrographs of a Transverse Section of *C. sublyratus* Leaf Magnified 2000 Times (A) and Palisade Mesophyll Cells Magnified 15400 Times (B) Showing Oil Globules of Plaunotol in the Chloroplasts.

Table 2. Product Distribution of Wild-Type and Mutant *S. dulcis* GGPPSs

Enzyme	GPP		FPP		GGPP	
	Activity (pmol/h/mg)	%	Activity (pmol/h/mg)	%	Activity (pmol/h/mg)	%
Wild-type	2.7	6.5	6.8	16.2	42.2	100
Mutant YF	9.3	22.1	23.3	55.3	3.0	7.1
Mutant PC	5.5	13.0	8.8	20.9	1.0	2.3
Mutant YP	8.8	20.9	19.1	45.2	3.7	8.8

Amounts of GGPP produced by the wild-type GGPPS were taken as 100%.

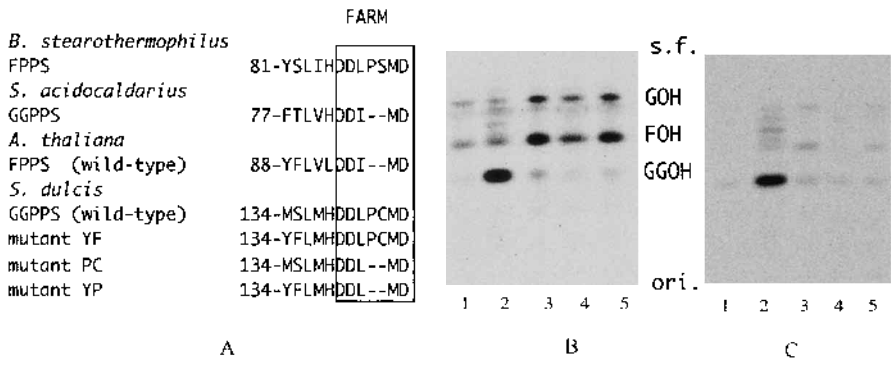

Fig. 5. Amino Acid Sequences Around FARMs of *Bacillus stearothermophilus* FPPS, *Sulfolobus acidocaldarius* GGPPS, *A. thaliana* FPPS, and Wild-Type and Mutant *S. dulcis* GGPPSs

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-^{14}C]IPP and either DMAPP (B) or FPP (C) as the primer substrate. The reaction products were analyzed by reversed-phase LKC-18 TLC. Lane 1, control *E. coil* transformed with 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., TLC origin; s.f., solvent front.

of Thr situated at the fourth position before FARM to Phe. Further, insertion of two amino acids in the archaeal GGPPS FARM also caused the mutant GGPPS produce FPP in a manner dependent on the substrate concentration. These results suggest that plant and archaeal GGPPSs employ different mechanisms to determine the chain length of the product. Therefore, 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 and the insertion in FARM caused the enzyme to produce FPP instead of GGPP.

Acknowledgments 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

1) Ogura K., Koyama T., *Chem. Rev*., **98**, 1263—1276 (1998).
2) Wang K., Ohnuma S. I., *Trends Biochem. Sci.*, **24**, 445—451 (1999).
3) Burke C. C., Wildung M. R., Croteau R., *Proc. Natl. Acad. Sci. U.S.A*., **96**, 13062—13067 (1999).
4) Ashby M. N., Edwards P. A., *J. Biol. Chem*., **265**, 13157—13164 (1990).
5) Marrero P. F., Poulter C. D., Edwards P. A., *J. Biol. Chem*., **267**, 21873—21878 (1992).
6) Chen A., Kroon P. A., Poulter C. D., *Protein Sci.*, **3**, 600—607 (1994).
7) Tarshis L. C., Yan M., Poulter C. D., Sacchettini J. C., *Biochemistry*, **33**, 10871—10877 (1994).
8) Tarshis 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).
9) Ohnuma S., Nakazawa T., Hemmi H., Hallberg A., Koyama T., Ogura K., Nishino T., *J. Biol. Chem*., **271**, 10087—10095 (1996).
10) Ohnuma S., Hirooka K., Hemmi H., Ishlda C., Ohto C., Nishino T., *J. Biol. Chem*., **271**, 18831—18837 (1996).
11) Ohnuma S., Narita K., Nakazawa T., Ishida C., Takeuchi Y., Ohto C., Nishino T., *J. Biol. Chem*., **271**, 30748—30754 (1996).
12) Ohnuma S., Hirooka K., Ohto C., Nishino T., *J. Biol. Chem.*, **272**, 5192—5198 (1997).
13) Kojima N., Sitthithaworn W., Viroonchatapan E., Suh D.-Y., Iwanami N., Hayashi T., Sankawa U., *Chem. Pharm. Bull*., **48**, 1101—1103 (2000).
14) Chiu W., Niwa Y., Zeng W., Hirano T., Kobayashi H., Sheen J., *Curr. Biol*., **6**, 325—330 (1996).
15) Frohman M. A., Dush M. K., Martin G. R., *Proc. Natl. Acad. Sci. U.S.A*., **85**, 8998—9002 (1988).
16) Bradford M. M., *Anal. Biochem*., **72**, 248—254 (1976).
17) Ke S. H., Madison E. L., *Nucleic Acids Res.*, **25**, 3371—3372 (1997).
18) Ogura K., Nishino T., Shinka T., Seto S., *Methods Enzymol*., **110**, 167—171 (1985).
19) Noji M., Inoue K., Kimura N., Gouda A., Saito K., *J. Biol. Chem*., **273**, 32739—32745 (1998).
20) von Heijne G., Steppuhn J., Herrmann G. R., *Eur. J. Biochem.*, **180**, 535—545 (1989).
21) Emanuelsson O., Nielsen H., von Heijne G., *Protein Sci*., **8**, 978—984 (1999).
22) Hefner J., Ketchum E. B., Croteau R., *Arch. Biochem. Biophys*., **360**, 62—74 (1998).
23) Williams D. C., McGarvey J. D., Katahira E. J., Croteau R., *Biochemistry*, **37**, 12213—12220 (1998).
24) Frishman D., Argos P., *Protein Eng.*, **9**, 133—142 (1996).
25) Hayashi T., Kasahara K., Sankawa U., *Phytochemistry*, **46**, 517—520 (1997).
26) Hayashi T., Asai T., Sankawa U., *Tetrahedron Lett.*, **40**, 8239—8243 (1999).
27) Lichtenthaler H. K., Schender J., Disch A., Rohmer M., *FEBS Lett*., **400**, 271—274 (1997).
28) De-Eknamkul W., Potduang B., *Science Asia*, 2000 (in press).
29) 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).