The geranylgeranyl pyrophosphate synthase gene from 
Ginkgo biloba: cloning, characterization and functional identification

Wei Wang1, Min Chen2, Chunxian Yang1, Wanhong Liu1;3, Xiaozhong Lan3, Zhihua Liao1*

1Laboratory of Natural Products and Metabolic Engineering, Chongqing Sweetpotato Research Center, Key Laboratory of Eco-environments in Three Gorges Reservoir Region (Ministry of Education), School of Life Sciences, Southwest University, Chongqing 400715, China.
2Laboratory of Natural Products and Metabolic Engineering, Key Laboratory of Eco-environments in Three Gorges Reservoir Region (Ministry of Education), School of Pharmacy, Southwest University, Chongqing 400715, China.
3Department of Biology, Chongqing University of Science and Technology, Chongqing 400042, China.
4Agricultural and Animal Husbandry College, Tibet University, Linzhi of Tibet 860000, China.

Accepted 30 January, 2009

Geranylgeranyl pyrophosphate synthase (GGDPS, EC: 2.5.1.29) catalyzes the biosynthesis of 20-carbon geranylgeranyl pyrophosphate (GGDP), which is a key precursor for ginkgolides biosynthesis. In order to investigate the role of GGDP synthase in ginkgolides biosynthesis, we cloned, characterized and functionally expressed the GGDP synthase gene from Ginkgo biloba. The genomic DNA fragment of the GGDPS gene of G. biloba (designated GbGGDPS GenBank accession number EF646377) was 2135 bp in length containing an 1176-bp open reading frame (ORF) that encoded a 391-amino acid polypeptide. Comparative analysis showed that GbGGDPS had a high similarity to other plant GGDPSs. Bioinformatic analysis showed that GbGGDPS was an intron-free gene and its deduced polypeptide contained all the five conserved domains and functional aspartate-rich motifs of the polyisoprenyltransferases. By constructing the phylogenetic tree of plant GGDPSs, it was found that plant-derived GGDPSs could be divided into two classes, angiosperm and gymnosperm classes, which might have evolved in parallel from the same ancestor. The homology-based structural modeling showed that GbGGDPS has the typical structure of GGDPS. The tissue expression profiles of GbGGDPS indicated that it could express in roots, stems, leaves, peshes and fruits but at different levels. The highest expression level of GbGGDPS was found in roots and peshes, while the lowest expression level of GbGGDPS was found in leaves. Interestingly, it was found that GbGGDPS had a higher expression level in the treatment of methyl jasmonate. Finally, the coding sequence of GbGGDPS was functionally expressed in Escherica coli in which the β-carotene pathway was reconstructed by genetic complementation, and the transgenic E. coli showed to have an activity of GGDP synthase.

Key words: Ginkgo biloba, geranylgeranyl pyrophosphate synthase, cloning, characterization, expression profile, functional complementation.

INTRODUCTION

Ginkgo biloba is one of the oldest living plant species and often referred to as “a living fossil” on earth, which dates back to more than 200 million years (Jacobs and Browner, 2000). The extracts of ginkgo leaves have been employed for treating cerebrovascular and cardiovascular diseases for centuries because ginkgolides are highly specific and potent platelet-activating factor receptor antagonists (Hosford et al., 1990) that have attracted great commercial interests as pharmaceuticals. And now,
the pharmaceutical ginkgolides are merely produced from the leaves of ginkgo, in which the contents of ginkgolides are very low (van Beek et al., 1991). Huge demand but limited supply makes the price of ginkgolides very high, and many cancer victims therefore cannot afford this drug. It is a hot scientific topic to find alternative sources of ginkgolides. The chemical synthesis of ginkgolides is academically successful, but not amendable for commercial-scale production of ginkgolides (Crimmins et al., 2000); ginkgo cell and tissue culture is used to obtain ginkgolides, but the undifferentiated ginkgo cells and tissue could hardly produce these important pharmaceutical compounds (Laurain et al., 1997). Therefore, metabolic engineering depended on genetically modifying the ginkgolide biosynthetic pathway might lead the production of ginkgolide to a bright future. So, it is necessary to map the ginkgolide pathway at the level of molecular biology.

Ginkgolides belong to plant diterpenoids just like the famous anti-tumor reagent, taxol (van Beek et al., 2002), which are biologically synthesized via the recently unveiled plastidial 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway (Eisenreich et al., 2001). Ginkgolides are derived from 20-carbon geranylgeranyl pyrophosphate (GGDP), the general key precursor for diterpenes. Geranylgeranyl pyrophosphate synthase (GGDPS), the enzyme step of GGDPs is known to be an importantly potential regulating target for metabolic engineering of diterpenes. In the present study, we reported the molecular cloning, characterization and functional identification of a gene encoding GGDPS from G. biloba, which will enable us to map and regulate an important step involved in ginkgolides biosynthetic pathway at the level of molecular biology in the future.

**MATERIALS AND METHODS**

**Plant materials**

All tissue materials including roots, stems, leaves, peshes and fruits were excised from G. biloba grown in the Medical Plants Garden of Southwest University (Chongqing, China). The cells of G. biloba were cultured and treated according to a previous report (Gong et al., 2005). Plant tissues were immersed in liquid nitrogen immediately after excision and preserved in a –70°C ultra low temperature refrigerator for RNA and DNA extraction.

**Strain and plasmids**

*E. coli* XL1-Blue was used for the general construction of plasmids. The pAtipiTrc harboring the Arabidopsis IPI gene (Cunningham and Elisabeth, 2000) was digested with Pst I to get rid of it and then recycled with T4 ligase. The modified plasmid was designed as pTrc that has the resistance of ampicillin that was used as vectors for expressing the recombinant GGDPs of ginkgo. The plasmid pACCARS5&crlE, which contains the gene cluster crtX (zeaxanthin glucosytrans-ferase), crfY (lycopene cyclase), crf (phytoene desaturase), crfB (phytoene synthase) and crtZ (β-carotene hydroxylation) encoding the carotenoid biosynthetic enzymes (Zhu et al., 1997), and pTRC-GbGGDPS, which contains GbGGDPS gene, were used for color production to test GGDP synthase activity.

**RNA isolation**

The genomic DNA of G. biloba was extracted with the CTAB method. The RNA isolating kit provided by TianGen (Beijing, China) was used in total RNA extraction. The quality and concentration of the RNA was checked by agarose gel electrophoresis (EC250-90, E-C Apparatus Corporation) and spectrophotometer (WFZUV-2100, Unico, Shanghai, China) Instruments Inc.) analysis and the RNA samples were stored in a –70°C ultra low temperature refrigerator prior to RACE and semi-quantitative one-step RT-PCR.

**Cloning of the genomic DNA of GbGGDPS**

Firstly, a core fragment of GbGGDPS was isolated with a pair of primers (FC: 5’-CTAGTTTTGTCTGAGGCAATATAATC-3’) from the genomic DNA of G. biloba. The PCR was carried out by denaturing the DNA at 94°C for 6 min, which was followed by 29 cycles of amplification (94°C for 50 min, 55°C for 45 s and 72°C for 2 min) and by extension at 72°C for 6 min. The first genomic DNA of GbGGDPS was isolated with the Universal GenomeWalker kit (Clontech, USA). Using the genomic DNA of G. biloba, four individual pools (Dral, EcoRV, PvuII, Stul) of uncloned, adaptor-ligated (GW Adaptor provided by CLONTECH) genomic DNA fragments were constructed according to the manufacturer’s instruction (CLONTECH). The four individual pools were then used as the libraries to isolate the genome-specific sequences of GbGGDPS by the PCR-based genome walking method. For the first PCR amplification, Dral, EcoRV, PvuII and Stul pools were used respectively as templates, AP1 (5’-CTATACGACTCCTATAGGC-3’) and GGPPSSW-1 (5’-CACCCTCTGGACGACAGTAACC-3’) as primers for upstream genomic sequence isolation and AP1 and GGPPSSW-1 (5’-CCGATGTGTGGGTCCTTTTGTGTC-3’) as primers for downstream sequence isolation. The PCR was conducted under the following conditions: 7 cycles (25 s at 94°C, 3 min at 72°C), followed by 32 cycles (25 s at 94°C follow by 3 min at 67°C) and then 6 min of extension at 67°C. The first amplified PCR products (50-fold dilution) were respectively used as templates for the nested PCR amplification of the genomic-specific sequences of GbGGDPS. AP2 (5’-ACTATAAGGACCCGTTGT-3’) and GGPPSSW-2 (5’-TGCCTGTCATTGCAGCTCG-3’) were used as the nested primers for the nested PCR amplification of the upstream genomic-specific sequence, and AP2 and GGPPSSW-2 (5’-GTGGGTCTTTCATTCACTGAGG-3’) as the nested primers for the nested PCR amplification of the downstream genomic-specific sequence. The nested PCR was carried out under the following conditions: 5 cycles (25 s at 94°C follow by 3 min at 72°C), followed by 20 cycles (25 s at 94°C follow by 3 min at 67°C) and then 6 min of extension at 67°C. The amplified PCR product was purified and cloned into T-easy vector (Promega) followed by sequencing. After assembling the core fragment, the upstream and downstream sequences a genomic sequence of interest was produced and the physical genomic DNA of interest was amplified by PCR using PF (5’-GTGGGTCTTTCATTCACTGAGG-3’) and RP (5’-CTGATATTCACCAAACTGAAACTCA-3’) as primers. When confirmed by sequencing, the genomic DNA sequence of GbGGDPS was submitted to GenBank.

**Bioinformatic analyses**

Comparative and bioinformatic analyses of GbGGDPS were carried out online at the websites (http://www.ncbi.nlm.nih.gov and http://www.expasy.org). The nucleotide sequence, deduced amino
acid sequence and ORF (open reading frame) encoded by GbGGDPS were analyzed and the sequence comparison was conducted through a database search using the BLAST program. The subcellular location was predicted by TargetP (Emanuelsson et al., 2007). The multiple alignments of GbGGDPS and GGDPs from other plant species were aligned with CLUSTAL W (Thompson et al., 1997) using default parameters. A phylogenetic tree was constructed using MEGA version 3 (Kumar et al., 2004) from CLUSTAL W alignments. The neighbor-joining method (Saitou and Nei, 1987) was used to construct the tree. The homology-based 3-D structural modeling of GbGGDPS was accomplished by Swiss-Modeling (Arnold et al., 2006). WebLab ViewerLite was used for 3-D structure displaying.

Tissue expression profile

Semi-quantitative one-step RT-PCR was carried out to investigate the expression profile of GbGGDPS in different tissues including roots, stems, leaves, peshes, fruits and cells (without MeJA-treated) of G. biloba. Aliquots of 0.5 µg total RNA extracted from roots, stems, leaves and peshes of G. biloba were used as templates in the one-step RT-PCR reaction with the forward primer FC and RC specific to the cDNA of GbGGDPS using the one-step RNA PCR kit (Takara, Kyoto, Japan). Amplifications were performed under the following conditions: 50°C for 30 min, 94°C for 2 min followed by 25 cycles of amplification (94°C for 50 s, 55°C for 50 s and 72°C for 2 min). Meanwhile, the RT-PCR reaction for the house-keeping gene (18S rRNA gene) using the specific primers, 18SF (5'-GTGACAATGGAACTGGAATGG-3') and 18SR (5'-CTCTAGA CTAGTTTTGTCTGAAGGCAATATAATC -3') and then 18SR (5'-CTCTAGA CTAGTTTTGTCTGAAGGCAATATAATC -3') and rexggdps (5'-GGATCCATGGCTGCCAGTGCAATGAC -3') and 18SR (5'-CTCTAGA CTAGTTTTGTCTGAAGGCAATATAATC -3') and then a 569-bp fragment was amplified from the PCR amplification products generated from the Pvu II pool. By assembling the upstream, core and downstream fragments, a 2135-bp genomic DNA of GbGGDPS was produced and confirmed by PCR and sequencing (Figure 1). The genomic GbGGDPS sequence was submitted to GenBank and assigned an accession number, EF646377.

The ORF finding analysis showed that the GbGGDPS contained an 1176-bp coding sequence encoding a 391-amino-acid polypeptide with a calculated molecular mass of 42.5 kDa and an isolectric point of 5.99 that were similar with the reported plant GGDPs from Taxus species (Liao et al., 2005). The coding sequence of GbGGDPS was not interrupted by intron like the genomic organization of other plants including Taxus (Liao et al., 2005) and Capsicum (Badillo et al., 1995).

Bioinformatic analysis

The deduced amino acid sequence of GbGGDPS was submitted to NCBI for BLAST searching and the results showed that GbGGDPS had high similarities with GGDPs from other plant species, such as Abies grandis (72% identities), Picea abies (72% identities), Taxus species (72% identities) and Adonis aestivalis (67% identities). Thus, the BLAST analysis results indicated that GbGGDPS belonged to the GGDPs family. The subcellular prediction analysis by TargetP suggested that GbGGDPS had a 79-aa transit peptide that directed GbGGDPS into plastids and this was consistent with the fact that terpenoids including ginkgolides was synthesized in plastids (Eisenreich et al., 2001). Based on the multiple alignments, it was found that all aligned plant GbGGDPSs had a plastidial transit peptide at the N terminus, which was not present in prokaryotic GGDPs (Liao et al., 2005). Furthermore, five domains (from I to V), found in the same relative locations as described for GGDPs (Chen et al., 1994), were also identified in all the aligned sequences (Figure 2). Domains II and V contained aspartate-rich motifs, which were proposed to be diphosphate-binding sites important for the catalytic activity of GGDPs (Chen et al., 1994). Using MEGA
Figure 1. Nucleotide and deduced amino acid sequence of GbGGDPS. The coding sequence was shown in capital and bold letters. The stop codon was marked with an asterisk.

Based on CLUSTAL W alignments to construct the phylogenetic tree, the phylogenetic analysis showed that plant-derived GGDPSs could be divided into two classes, those of angiosperms and gymnosperms (Figure 3). The result was consistent with our previous reports (Liao et al., 2004). This suggested that plant GGDPSs evolved in parallel and could have originated from the same ancestor.

The homology-based structural modeling of GbGGDPS was performed by Swiss-Modeling on the basis of the Sinapis Alba GGDPS crystal structure (Kloer et al., 2006) and displayed by WebLab ViewerLite (Figure 4). GbGGDPS displays a typical GGDPS structure, which was composed of numerous α-helices connected with random coils. The helices were circled with a cave in which the aspartate-rich motifs were localized.

**Tissue expression profile**

To investigate the expression profile of GbGGDPS in different tissues of *G. biloba*, total RNA was isolated from different tissues including roots, stems, leaves, peshes, fruits and cells (with/without MeJA-treated) and subjected to semi-quantitative one-step RT-PCR using FC and RC as primers. The 18S rRNA gene expression in all the detected tissues was used as an internal control that showed no significant difference (Figure 5). The result showed that GbGGDPS expression could be detected in all tested tissues including roots, stems, leaves, peshes and fruits but at different levels (Figure 5). The highest expression level of GbGGDPS was found in roots, followed by peshes, stems and fruits; and the lowest expression level of GbGGDPS was found in leaves. Further, it was also found that GbGGDPS had a higher expression when treated with MeJA (Figure 6). The previous research reported that MeJA could induce the over production of diterpenes including taxol (Jennewein and Croteau, 2004) and ginkgolides (Gong et al., 2005). In taxol biosynthetic pathway, GGDPS was a key enzyme gene for metabolic engineering of ginkgolides.
Figure 2. Alignments of amino acid sequences in the five conserved domains of plant GGPPSs. The identical and conserved amino acid residues are shown in white capital letters against the black and gray background, respectively, and the conserved amino acid residues, in white capital letters and gray background. The five conserved domains are designated by Roman numerals (I–V). The highly conserved aspartate-rich motifs in Domain II (DDXXXD) and Domain V (DDXXD) are boxed. Sources of GGPPS: AtGGPPS, Arabidopsis thaliana (GenBank Accession No. P34802); CcGGPPS, Clastos creticus (GenBank Accession No. AAM21638); MpGGPPS, Mentha peperita (GenBank Accession No. AAF08793); SaGGPPS, Sinapis alba (GenBank Accession No. CAA67330); TeGGPPS, Tagetes erecta (GenBank Accession No. AAG10424); AgGGPPS, Abies grandis (GenBank Accession No. AAL17614); GbGGDPS, Ginkgo biloba (GenBank Accession No. AY371321); and TmGGPPS, Taxus media (GenBank Accession No. AY453404).
Figure 3. The phylogenetic tree of plant GGDPSs.

Figure 4. 3-D structure of GbGGDPS from homology-based structural modeling. The columns represent α-helices; the arrow plates represent β-sheets; the ropes represent random coils. The small balls showed the motifs composed of amino acids.

Figure 5. Tissue expression profile of GbGGDPS. The upper lane is GbGGDPS in different organs including roots, stems, leaves, peshes and fruits; the lower lane is the 18s rRNA gene used as internal control.

Figure 6. MeJA-inducible GbGGDPS expression profile. The upper lane is GbGGDPS expression in the treatment of MeJA; the lower lane is the 18s rRNA gene used as internal control. MeJA: without MeJA; 24 h MeJA: MeJA 24 h; 48 h MeJA: MeJA 48 h.

Functional identification of GbGGDPS

In the color complementation assay, E. coli strain XL1-Blue was used to test if GbGGDPS encoded the anticipated functional enzyme, GGDPS from G. biloba. On the media containing ampicillin (150 μg/mL) and chloromyc-
In summary, a functional gene encoding GGDPS (GbGGDPS) was cloned, characterized and functionally identified from G. biloba. It will facilitate the unveiling of the biosynthesis of ginkgolides at the level of molecular biology and provide a candidate gene for metabolic engineering of ginkgolides.

ACKNOWLEDGEMENTS

This research was financially supported by the NSFC Project (30500303). The authors are thankful to Dr. Misawa in the Kirin Company of Japan for his kindly providing pACCAR25ΔcrtE.

REFERENCES


Sagami H, Kurisaki A, Ogura K (1993). Formation of dolichol from dehy-
drodolichol is catalyzed by NADPH-dependent reductase localized in microsomes of rat liver. J. Biol. Chem. 268: 10109-10113.


