Archives of Biochemistry and Biophysics
Vol. 375, No. 2, March 15, pp. 261–269, 2000
doi:10.1006/abbi.1999.1669, available online at http://www.idealibrary.com on IDEAL®

Terpenoid Secondary Metabolism in *Arabidopsis thaliana:* cDNA Cloning, Characterization, and Functional Expression of a Myrcene/(*E*)-β-Ocimene Synthase[1]

Jörg Bohlmann,[2] Diane Martin, Neil J. Oldham, and Jonathan Gershenzon
Max Planck Institute for Chemical Ecology, Carl-Zeiss-Promenade 10, 07745 Jena, Germany

Received September 13, 1999, and in revised form November 29, 1999

The *Arabidopsis* genome project has recently reported sequences with similarity to members of the terpene synthase (*TPS*) gene family of higher plants. Surprisingly, several *Arabidopsis* terpene synthase-like sequences (*AtTPS*) share the most identity with *TPS* genes that participate in secondary metabolism in terpenoid-accumulating plant species. Expression of a putative *Arabidopsis* terpene synthase gene, designated *AtTPS03,* was demonstrated by amplification of a 392-bp cDNA fragment using primers designed to conserved regions of plant terpene synthases. Using the *AtTPS03* fragment as a hybridization probe, a second *AtTPS* cDNA, designated *AtTPS10,* was isolated from a jasmonate-induced cDNA library. The partial *AtTPS10* cDNA clone contained an open reading frame of 1665 bp encoding a protein of 555 amino acids. Functional expression of *AtTPS10* in *Escherichia coli* yielded an active monoterpene synthase enzyme, which converted geranyl diphosphate (C_{10}) into the acyclic monoterpenes β-myrcene and (*E*)-β-ocimene and small amounts of cyclic monoterpenes. Based on sequence relatedness, *AtTPS10* was classified as a member of the *TPSb* subfamily of angiosperm monoterpene synthases. Sequence comparison of *AtTPS10* with previously cloned monoterpene synthases suggests independent events of functional specialization of terpene synthases during the evolution of terpenoid secondary metabolism in gymnosperms and angiosperms. Functional characterization of the *AtTPS10* gene was prompted by the availability of *Arabidopsis* genome sequences. Although *Arabidoposis* has not been reported to form terpenoid secondary metabolites, the unexpected expression of *TPS* genes belonging to the *TPSb* subfamily in this species strongly suggests that terpenoid secondary metabolism is active in the model system *Arabidopsis.*

Key Words: **monoterpenoid biosynthesis; monoterpene synthase; terpene cyclase; *Arabidopsis thaliana;* genome; myrcene; ocimene; gene family.**

INTRODUCTION

The *Arabidopsis* genome project has recently reported several sequences with similarity to members of the terpene synthase (*TPS*) gene family of higher plants (1, 2). Terpene synthases use the prenyl diphosphate substrates geranyl diphosphate (GPP)[3] (C_{10}), farnesyl diphosphate (FPP) (C_{15}), and geranylgeranyl diphosphate (GGPP) (C_{20}) to generate the carbon skeletons characteristic of the many terpenoid natural products (3–5). The *TPS* gene family includes monoterpene (C_{10}), sesquiterpene (C_{15}), and diterpene (C_{20}) synthases that participate in secondary metabolism (6). In addition, the terpene synthases copalyl diphosphate synthase (CPS) and *ent*-kaurene synthase (KS) catalyze sequential cyclization reactions converting GGPP to *ent*-kaurene in gibberellin biosynthesis. CPS and KS were cloned and functionally characterized in *Arabidopsis thaliana* (7, 8) and several other plant species (9–12). Terpene synthases participating in secondary metabolism have been cloned as cDNAs from a number

[1] The nucleotide sequence reported in this paper will appear in GenBank under Accession Nos. AF178535 and AF180366.

[2] To whom correspondence should be addressed. E-mail: bohlmann@ice.mpg.de. Fax: ++49-3641-643650.

[3] Abbreviations used: FID, flame ionization detection; nt(s), nucleotide(s); dNTP, deoxyribonucleotide triphosphate; ORF, open reading frame; PCR, polymerase chain reaction; SSPE, saline–sodium phosphate–EDTA; GPP, geranyl diphosphate; FPP, farnesyl diphosphate; GGPP, geranylgeranyl diphosphate; CPS, copalyl diphosphate synthase; KS, *ent*-kaurene synthase; SDS, sodium dodecyl sulfate; BSA, bovine serum albumin; LPP, linalyl diphosphate.

0003-9861/00 $35.00

of species, including the gymnosperms *Abies grandis* (13–17) and *Taxus brevifolia* (18) and members of the Lamiaceae (mint family) (19–22) and the Solanaceae (nightshade family) (23–26). Secondary metabolite terpenoid synthases are involved in the formation of conifer oleoresin (27, 28), floral volatiles (29), phytoalexins (30), and essential oil components of many plant species (29). *Arabidopsis* has not been reported to produce terpenoid secondary metabolites. Thus, the recent identification of *TPS*-like sequences (*AtTPS*) in the *Arabidopsis* genome, which are most similar to members of the *TPS* gene family involved in secondary metabolism, was not anticipated. It is not known whether *AtTPS* genes are transcriptionally expressed and, if so, what biochemical activities are catalyzed by the encoded enzymes.

The *TPS* gene family of higher plants has been divided into six subfamilies (6). Sequence comparison of terpene synthases within and between subfamilies and characterization of the corresponding enzymes demonstrated that sequence relatedness allows identification of members of the three major classes of terpene synthases, monoterpene, sesquiterpene, and diterpene synthases (6). However, sequence relatedness is not sufficient for identification of the exact biochemical functions of these diverse catalysts (6, 27). Expression of active enzymes in a recombinant form is required for functional characterization of new *TPS* genes. This paper reports on the expression of two *TPS* genes in *Arabidopsis*, the cloning and expression in *Escherichia coli* of the *AtTPS10* cDNA, and functional characterization of active AtTPS10 monoterpene synthase as myrcene/(*E*)-β-ocimene synthase.

EXPERIMENTAL PROCEDURES

Substrates, reagents, and cDNA library. [1-^{3}H]Geranyl diphosphate (250 Ci/mol) (31), [1-^{3}H]farnesyl diphosphate (125 Ci/mol) (32), and [1-^{3}H]geranylgeranyl diphosphate (120 Ci/mol) (33) were gifts of Dr. Rodney Croteau, Washington State University, Pullman. Terpenoid standards were from our own collection. All other biochemicals and reagents were purchased from Sigma Chemical Company or Aldrich Chemical Company, unless otherwise noted. The jasmonic acid-induced λgt11 cDNA library prepared from *Arabidopsis thaliana* ecotype C24 (34) was a gift of Dr. Jose J. Sanchez-Serrano, Universidad Autonotia de Tiadrid, Madrid, Spain.

PCR-based probe generation. Two rounds of PCR were performed each in a total volume of 50 μl containing 20 mM Tris/HCl (pH 8.4), 50 mM KCl, 5 mM $MgCl_2$, 200 μM of each dNTP, 0.1 μM of primer 2-1 (5′-AGG TTT ACT CTC TCT ATA CGA AGC-3′), 0.1 μM of primer 2-2 (5′-AAT AAC TCT CTG TGA TTC GGT CTC-3′), 1 U of *Taq* polymerase (BRL), and 5 μl of template. As template in the primary PCR, 5 μl of the *Arabidopsis* cDNA library (2×10^8 pfu/ml) was used. Five microliters of the completed primary reaction served as template in a secondary amplification. The temperature program for PCR was denaturation at 95°C for 2 min, followed by 35 cycles of amplification employing the following temperature program in a Gradient 96 Robocycler (Stratagene): 1 min at 95°C, 1 min at 45°C, 1 min at 72°C. PCRs were analyzed by agarose gel electrophoresis (35). A 392-bp amplicon of the secondary reaction was ligated into pCR2.1-TOPO (Invitrogen), and transformed into *E. coli* TOP10F′ cells (Invitrogen). Plasmid DNA was prepared from individual transformants and the inserts were sequenced (DyeDeoxy Terminator Cycle Sequencing, Applied Biosystems). A 392-bp insert sequence was identified and was designated as probe *AtTPS03*.

Library screening. Fifty nanograms of probe *AtTPS03* was amplified by PCR, purified by agarose gel electrophoresis and extraction using the Qiagen (Hilden, Germany) gel extraction kit, randomly labeled with [α-^{32}P]dATP (36), and used to screen replica filters of 2×10^5 plaques of the jasmonic acid-induced *Arabidopsis* cDNA library plated on *E. coli* Y1090r$^-$. Hybridization was performed for 18 h at 50°C in 3× SSPE and 0.1% SDS. Filters were washed three times for 10 min at 50°C in 3× SSPE with 0.1% SDS and exposed for 17 h to Kodak BioMax-MS film at −80°C. Thirty clones yielding positive signals with probe *AtTPS03* were purified through a second round of hybridization. Inserts of purified λgt11 clones were amplified by PCR using primers λgt11-F (5′-GGT GGC GAC GAC TCC TGG AGC CCG-3′) and λgt11-R (5′-TTG ACA CCA GAC CAA CTG GTA ATG-3′). PCR was performed in a total volume of 50 μl containing 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 20 mM Tris/HCl (pH 8.75), 2 mM $MgSO_4$, 0.1% Triton X-100, 5 μg BSA, 200 μM of each dNTP, 0.1 μM of primer λgt11-F, 0.1 μM of primer λgt11-F, 1 U of high-fidelity cloned *Pfu* polymerase (Stratagene), and 10 μl of phage solution. After a denaturation step at 95°C for 2 min, 30 cycles of amplification were performed employing the following temperature program in a Gradient 96 Robocycler: 1 min at 94°C, 1 min at 60°C, 2 min at 72°C, followed by a 10-min final extension at 72°C. PCR-amplified cDNA inserts were ligated into pCR-Blunt (Invitrogen) and transformed into *E. coli* TOP10 cells (Invitrogen). Plasmid DNA was prepared from individual transformants and the inserts were sequenced.

cDNA expression in E. coli and enzyme assays. The *AtTPS10* insert fragment of plasmid pCR-Blunt-J1 was subcloned in frame into the pSBETa expression vector (37). Restriction sites *Nde*I and *Bam*HI for subcloning a version of cDNA *AtTPS10* without the presumptive plastidial targeting peptide (38) were introduced by sticky-end PCR (39). Fragments were amplified by PCR using primer 1-J1-RR (5′-TAT GCG TCG TTC TGC GAA CTA TCA GCC-3′) in combination with primer 3-J1-RR (5′-CTC AAT CTA AAG GAA TCG GAT TG-3′) and primer 2-J1-RR (5′-TGC GTC GTT CTG CGA ACT ATC AGC C-3′) in combination with primer 4-J1-RR (5′-GAT CCT CAA TCT AAA GGA ATC GGA TTG-3′). PCRs were performed in volumes of 100 μl containing 20 mM Tris/HCl (pH 8.8), 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100, 5 μg BSA, 200 μM of each dNTP, 0.1 μM of each primer, 2.5 U of high-fidelity Turbo *Pfu* polymerase (Stratagene), and 10 ng plasmid pCR-Blunt-J1. After a denaturation step at 95°C for 2 min, 30 cycles of amplification were performed employing the following temperature program in a Gradient 96 Robocycler (Stratagene): 30 s at 95°C, 30 s at 50°C, 2 min at 72°C, followed by a 10-min final extension at 72°C. The resulting PCR products were purified by agarose gel electrophoresis and extraction using the Qiagen gel extraction kit. Purified PCR products were adjusted to concentrations of 10 ng/μl and equal volumes were combined, denatured at 94°C for 10 min, and reannealed at 25°C for 30 min. One hundred nanograms of reannealed PCR products were ligated into *Nde*I/*Bam*HI-digested pSBETa to yield plasmid pSBET-*AtTPS10*(R). Plasmid pSBET-*AtTPS10*(R) was then transformed into *E. coli* BL21(DE3). For functional expression, bacterial strain *E. coli* BL21(DE3)/pSBET-*AtTPS10*(R) was grown to $A_{600} = 0.5$ at 37°C in 5 ml of LB medium supplemented with 30 μg kanamycin/ml. Cultures were then induced by addition of 1 mM isopropyl-1-thio-β-D-galactopyranoside and grown for another 12 h at 20°C. Cells were harvested by centrifugation and disrupted by sonication using a Model HD2070, MS72 sonicator (Bandeln, Berlin, Germany) at 20% power for 15 s, and the resulting soluble enzyme preparations were assayed for monoterpene, sesquiterpene, and diterpene synthase

activity using the appropriate radiolabeled substrates as described previously (14). A modified buffer for monoterpene synthase extraction and assays consisted of 50 mM Tris/HCl, pH 7.0, 10 mM $MgCl_2$, 1 mM $MnCl_2$, 10% (v/v) glycerol, 5 mM dithiothreitol. In the case of the monoterpene synthase and sesquiterpene synthase assays, the assay mixture (1 ml) was overlaid with 1 ml of pentane to trap volatile products. In all cases, after incubation at 30°C for 1 h, the reaction mixture was extracted with pentane (3 × 1 ml) and the combined extract was passed through a 1.5-ml column of anhydrous $MgSO_4$ and silica gel (Sigma, 60 Å) to provide the terpene hydrocarbon fraction free of oxygenated metabolites. To collect any oxygenated products, assay mixtures were subsequently extracted with 3 × 1 ml of diethyl ether. To test for the formation of phosphorylated products (21), the aqueous phase was then treated with excess potato grade III apyrase, adjusted to pH 8.0–8.5, and treated with excess calf intestinal alkaline phosphatase to hydrolyze monoterpenol diphosphate esters. The incubation mixtures were subsequently extracted with 3 × 1 ml of diethyl ether, and water was removed from combined extracts using anhydrous $MgSO_4$. Aliquots of each fraction were taken for liquid scintillation counting to determine conversion rate. To obtain sufficient product for analysis by GC-FID and GC-MS, the enzyme assay was scaled up by a factor of 20 and assays were incubated for 5 h. Controls for product formation independent of *AtTPS10* cDNA were performed using extracts of *E. coli* BL21(DE3) transformed with plasmid pSBETa.

Monoterpenes were quantified by GC-FID analysis on a Hewlett–Packard 6890 using a DB-WAX (methyl dodecanoate, J&W Scientific) capillary column (0.25 mm i.d. × 30 m with 0.25-μm film). A Cyclodex B (chiral β-cyclodextrin, J&W Scientific) capillary column (0.25 mm i.d. × 30 m with 0.25-μm film) was used for chiral separation of monoterpene products. Split injections at a ratio of 5:1 were made at an injector temperature of 220°C and a column flow of 2 ml He min^{-1}. The oven was operated at an initial temperature of 40°C (3-min hold) which was then increased to 80°C at 3°C min^{-1} followed by a 15°C min^{-1} ramp until 240°C (1-min hold). Compounds were identified by comparing retention times with those of authentic standards and the peak areas were integrated by Hewlett–Packard Chemstation software. GC-MS analysis was performed on a Hewlett–Packard 6890 GC–MS system (70 eV) using a Cyclodex B capillary column (0.25 mm i.d. × 30 m with 0.25-μm film). Splitless injections were made at an injector temperature of 220°C. The oven was programmed from 40°C (10-min hold) to 57°C at 3°C min^{-1}, followed by 0.1°C min^{-1} to 62°C, then 0.5°C min^{-1} to 65°C, and finally 15°C min^{-1} to 230°C (2-min hold) with a constant flow of 1 ml min^{-1} He. Mass spectra were analyzed using Hewlett–Packard Chemstation software and were compared with those of authentic standards. Stereochemistry was assigned based on matching of retention times with enantiomerically pure standards.

Sequence analysis. Inserts of all recombinant plasmids were completely sequenced on both strands via primer walking using the DyeDeoxy Terminator Cycle Sequencing method (Applied Biosystems). Sequence analysis was done using programs from the Wisconsin Package Version 9.1, Genetics Computer Group (GCG), Madison; CLUSTALW (http://www2.ebi.ac.uk/clustalw/); and TREECON Version 1.3b (Yves van de Peer, University of Antwerp, Belgium).

RESULTS AND DISCUSSION

cDNA Cloning of AtTPS03 and AtTPS10

The *TPS* gene family of higher plants represents a probable monophyletic group of highly specialized monoterpene, sesquiterpene, and diterpene synthases of primary and secondary metabolism (6). Recently, the *Arabidopsis* genome project reported several genes with sequence similarity to members of the *TPS* gene family (1, 2). Only two terpene synthases involved in the biosynthesis of gibberellins, CPS and KS, which catalyze the two-step cyclization of GGPP, have been previously cloned and characterized from *Arabidopsis* (7, 8). The newly discovered *Arabidopsis TPS*-like genes show only low similarities to CPS and KS, and are more closely related to sequences of monoterpene, sesquiterpene, and diterpene synthases of secondary metabolism described from other plant species. At the initiation of this study, three different *Arabidopsis* gene sequences with significant similarities to members of the *TPS* gene family were available in GenBank. Following a previously introduced nomenclature (14) and following the numbering system suggested by Aubourg *et al.* (1), the putative *Arabidopsis TPS* genes were designated *AtTPS01* (Accession No. Y11187, nucleotides 6093–8668), *AtTPS02* (Accession No. Z97341, nucleotides 164983–170504), and *AtTPS03* (Accession No. Z97341, nucleotides 172598–178554). Expression of these, or any other, *TPS* genes with similarity to known terpenoid synthases of secondary metabolism has not yet been demonstrated in *Arabidopsis*. To test for *AtTPS* gene expression and to identify a source for cDNA isolation, several *Arabidopsis* cDNA libraries made from whole plants, including rosette leaves, stems, and flowers, from flower buds (40), or from jasmonate-induced plants (34), were screened using PCR. Libraries were adjusted to equal titers of 2×10^8 pfu/ml. Two regions of the putative terpene synthase genes *AtTPS02* and *AtTPS03* were selected as primer binding sites for PCR (Fig. 1). These positions are highly conserved among *TPS* genes of plant origin (6) and identical for *AtTPS02* and *AtTPS03*. With all template libraries tested, two rounds of PCR were necessary to yield detectable amounts of a 392-bp amplicon (GenBank Accession No. AF180366), the sequence of which was identical to the genomic sequence of *AtTPS03* (Fig. 1). No amplicon was detected after primary PCR, suggesting a low level of representation of the genes *AtTPS02* and *AtTPS03* in these libraries. The 392-bp *AtTPS03* probe was subsequently used for filter hybridization of 2×10^5 cDNAs of the jasmonate-induced library, yielding a single strong hybridization signal, J1, and 29 weaker signals. All clones yielding positive signals with probe *AtTPS03* were purified through a second round of hybridization. cDNA inserts of purified phage clones were then amplified by PCR using high-fidelity polymerase, subcloned into plasmid vector pCR-Blunt, and partially sequenced from both ends. Sequence analysis revealed similarities to *TPS* genes only for cDNA clone J1, which was subsequently designated *AtTPS10*. The *AtTPS10* cDNA (GenBank Accession No. AF178535) is not identical to the genes *AtTPS01*, *AtTPS02*, and *AtTPS03*. However, a genomic

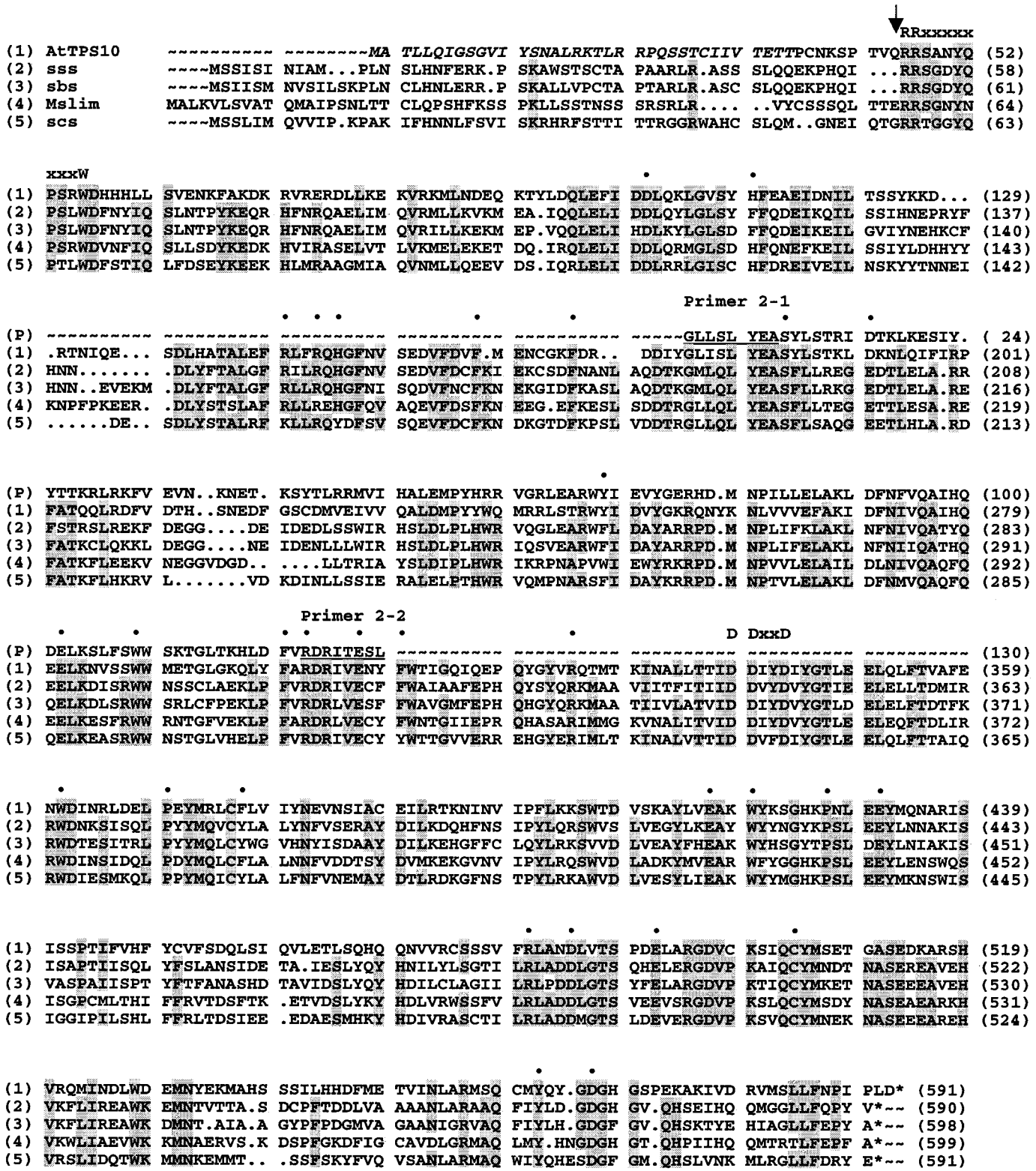


FIG. 1. Alignment of deduced amino acid sequences of monoterpene synthases of the *TPSb* group of the plant *TPS* gene family. (1) *Arabidopsis thaliana*, myrcene/(*E*)-β-ocimene synthase *AtTPS10*, (2) *Salvia officinalis*, (+)-sabinene synthase (21); (3) *S. officinalis*, (+)-bornyl diphosphate synthase (21); (4) *Mentha spicata*, (−)-limonene synthase (19); (5) *S. officinalis*, 1,8-cineole synthase (21). The alignment was created using the PILEUP program. Residues that are identical among at least four of the five monoterpene synthases are shaded in

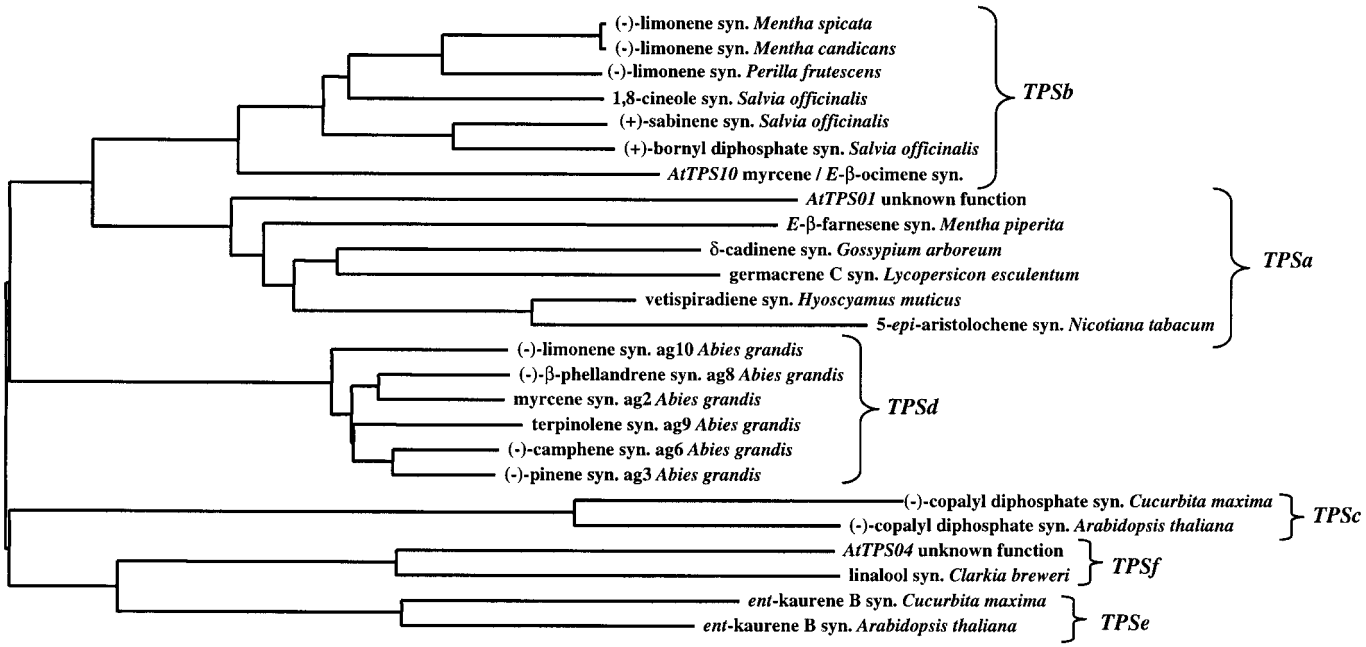


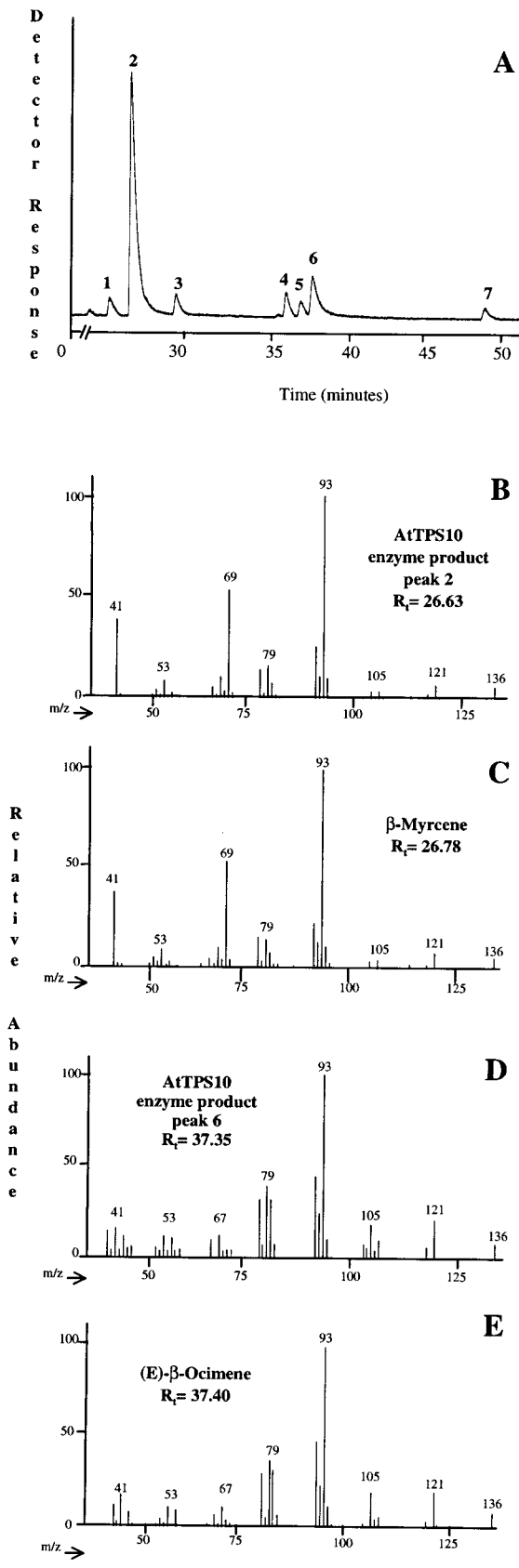
FIG. 2. Sequence relatedness of *AtTPS10* [myrcene/(*E*)-*β*-ocimene synthase] and other members of the plant *TPS* gene family. A phylogenetic tree of the plant *TPS* gene family was reconstructed using the neighbor joining method (48) and using the programs CLUSTALW to create the multiple sequence alignment and TREECON to calculate sequence relatedness. The *TPS* gene family is subdivided into six subfamilies *TPSa* through *TPSf* (6). *Arabidopsis* sequences, which resemble members of the *TPS* gene family of higher plants, are represented in five of the six *TPS* subfamilies. The *Arabidopsis TPSc* and *TPSe* genes encode the diterpene synthases, which catalyze the two-step cyclization of geranylgeranyl diphosphate to *ent*-kaurene via (−)-copalyl diphosphate. (−)-Copalyl diphosphate synthase and *ent*-kaurene synthase are the initial steps in the biosynthesis of gibberellic acid-derived phytohormones. *AtTPS10* is a member of the *TPSb* group of angiosperm monoterpene synthases. Functional expression of the *AtTPS10* cDNA and characterization of recombinant enzyme identified the *AtTPS10* gene product as myrcene/(*E*)-*β*-ocimene synthase. *AtTPS10* is the first functionally characterized *Arabidopsis TPS* gene encoding a secondary metabolite terpene synthase. The *Arabidopsis* myrcene/(*E*)-*β*-ocimene synthase is only distantly related to *Abies grandis* myrcene synthase, ag2 (14), a member of the *TPSd* group of gymnosperm terpenoid synthases. Other putative *Arabidopsis* terpene synthases resemble most closely members of the *TPSa* subgroup of angiosperm sesquiterpene synthases [e.g., *AtTPS01,* GenBank Accession No. Y11187 (1)] and linalool synthase in the *TPSf* subgroup (*AtTPS04,* GenBank Accession No. AC002294).

sequence identical to the *AtTPS10* cDNA was submitted to GenBank (Accession No. AC005967) after this study was completed.

Sequence Analysis of AtTPS10

The deduced 555-amino-acid sequence of the longest ORF (1665 bp) of the partial *AtTPS10* cDNA closely resembles the predicted proteins of the genomic sequences of *AtTPS02* (62.0% similarity, 52.6% identity) and *AtTPS03* (66.1% similarity, 56.0% identity). Overall sequence relatedness places *AtTPS10* into the *TPSb* group of the *TPS* gene family (6) (Fig. 2). Genes of the *TPSb* group have been previously described only from members of the Lamiaceae, and the corresponding enzymes have all been identified as monoterpene cyclases, including limonene synthase (19, 20), bornyl diphosphate synthase, 1,8-cineole synthase, and sabinene synthase (21). Comparisons of the amino acid sequence deduced from the *AtTPS*10 cDNA with its corresponding genomic sequence and with sequences of other terpene synthases of the *TPSb* group indicated that the *AtTPS10* cDNA is truncated at the 5′ terminus (Fig. 1). The truncation represents 36 amino acid residues of a presumptive transit peptide region at the N-terminus of the deduced protein. All previously described monoterpene and diterpene synthases of the

gray. A cDNA fragment of the *A. thaliana AtTPS03* gene (P) was amplified using primer 2-1 (underlined region) and primer 2-2 (underlined region) and used as a hybridization probe. The arrow indicates the site of truncation of monoterpene synthases for removal of transit peptides (17, 38). Dots indicate highly or absolutely conserved residues of members of the plant *TPS* gene family (6). Positions of the absolutely conserved RRX_8W motif of monoterpene synthases and the absolutely conserved aspartate rich region DDXXD of terpene synthases are marked. The truncated region of the transit pepetide of *AtTPS10,* which was deduced from the genomic sequence, is shown in italics.



TPSb and *TPSd* groups have transit peptides of approximately 50–60 amino acids (6, 38, 41) which facilitate the import of these nuclear encoded gene products into plastids, a process that involves cleavage of the preproteins to the mature active enzymes. Although the precise cleavage site is not yet known for any terpene synthase preprotein, truncation of monoterpene synthases upstream of a conserved tandem arginine motif (RRX_8W) has been demonstrated to result in fully active enzymes (17, 38). The tandem arginine motif (RRX_8W) is conserved in the deduced protein of *AtTPS10*, as are other previously described motifs common to the *TPSa*, *TPSb*, and *TPSd* subfamilies of plant terpene synthases (6) (Fig. 1). The molecular weight of the *AtTPS10*-encoded preprotein was predicted based on the corresponding genomic sequence (GenBank Accession No. AC005967) to be 69,278 Da, and to have a p*I* of 6.01, similar to previously characterized monoterpene synthases of the *TPSb* and *TPSd* groups. Truncation of the *AtTPS10*-encoded protein immediately upstream of the RRX_8W motif results in a protein of predicted molecular weight 64,417 Da and p*I* of 5.56. Since the deletion in the *AtTPS10* deduced protein is within the putative transit peptide, the truncated cDNA was expected to yield an active terpene synthase enzyme when expressed in *E. coli*.

Expression of AtTPS10 in E. coli and Characterization of Myrcene/Ocimene Synthase

The AtTPS10 cDNA was subcloned into the pSBETa expression vector, yielding pSBET-*AtTPS10*(R) and expressed in *E. coli* BL21(DE3). The expression vector employs the T7 RNA polymerase promoter and contains the *argU* gene for tRNAs using the arginine codons AGA and AGG, which are rare in *E. coli* but commonly found in plant terpene synthases (17, 38). The protein deduced from *AtTPS10* contains 16 arginine residues encoded by the AGA or AGG codon. Recombinant AtTPS10 protein was expressed as a "pseudo-mature" protein as described by Williams *et al.* (38), which is devoid of the transit peptide and contains

FIG. 3. Identification of the monoterpene products derived from geranyl diphosphate *in vitro* by recombinant *Arabidopsis thaliana* AtTPS10 myrcene/(*E*)-*β*-ocimene synthase. (A) Total ion chromatogram of reaction products separated on a chiral phase (*β*-cyclodextrin) column and identified as tricyclene (peak 1), myrcene (peak 2), 2-carene (peak 3), (−)-limonene (peak 4), (+)-limonene (peak 5), and (*E*)-*β*-ocimene (peak 6). Peak 7 represents an unknown monoterpene. (B) and (C) illustrate mass spectra and retention times of the principal AtTPS10 enzyme product, *β*-myrcene (B), and of the authentic *β*-myrcene standard (C). (D) and (E) illustrate mass spectra and retention times of the AtTPS10 product (*E*)-*β*-ocimene (D) and of the authentic (*E*)-*β*-ocimene standard (E).

a starting methionine immediately upstream of the RRX_8W motif (Fig. 1). Deletion of the transit peptide can improve expression of plant terpene synthases in *E. coli* and reduce the formation of intractable inclusion bodies without affecting product outcome (17, 38). Extracts of induced *E. coli* BL21(DE3)/pSBET-*AtTPS10*(R) were assayed for monoterpene, sesquiterpene, and diterpene synthase activity using the corresponding ^{3}H-labeled C_{10}, C_{15}, and C_{20} prenyl diphosphate substrates (14). Enzymatic production of terpene olefins in extracts of 100 ml culture of *E. coli* BL21(DE3)/pSBET-*AtTPS10*(R) was observed only using [1-^{3}H]GPP as substrate at the level of 0.5–1.0 nmol/h. Extracts prepared from *E. coli* BL21(DE3)/pSBETa and heat-denatured extracts of BL21(DE3)/pSBET-*AtTPS10*(R) served as controls for monoterpene formation independent of recombinant AtTPS10 enzyme. These extracts did not yield detectable amounts of monoterpene product. Monoterpenes generated by recombinant AtTPS10 enzyme were analyzed by GC-FID and GC-MS (Fig. 3) on both normal and chiral phase (β-cyclodextrin) capillary columns. Compounds were identified by comparison of retention times and mass spectra with those of authentic standards (Fig. 3), and quantified based on GC-FID response (not shown). The major monoterpene product of the AtTPS10 enzyme was identified as β-myrcene (56% of total hydrocarbon product), followed by (*E*)-β-ocimene (20%) and minor amounts (each less than 5%) of the cyclic monoterpenes (−)-limonene, (+)-limonene, 2-carene, and tricyclene and an unknown, possibly novel, monoterpene (Fig. 3). Because of the dominance of the two acyclic monoterpene products, myrcene and β-ocimene, which distinguishes this enzyme from the single product myrcene synthase from grand fir (17), the enzyme encoded by *AtTPS10* is designated as a myrcene/(*E*)-β-ocimene synthase. No oxygenated or phosphorylated monoterpene products were detected. A genomic sequence for *AtTPS10* was recently reported with the annotation "putative limonene cyclase" (GenBank Accession No. AC005967). However, functional analysis of the gene product demonstrated that limonene, in the form of its two stereoisomers, is only a minor product of this monoterpene synthase. The cloning of the *AtTPS10* cDNA and expression of the enzyme, as described here, provide a clear illustration of the fact that sequence relatedness is insufficient to determine the catalytic activity of new members of the *TPS* gene family. Sequence relatedness and diagnostic sequence elements, such as the RRX_8W motif and putative transit peptide sequences, allow placement of new members of the *TPS* gene family into subfamilies, such as the *TPSb* group of angiosperm monoterpene synthases. Identification of the exact catalytic function, however, requires functional characterization.

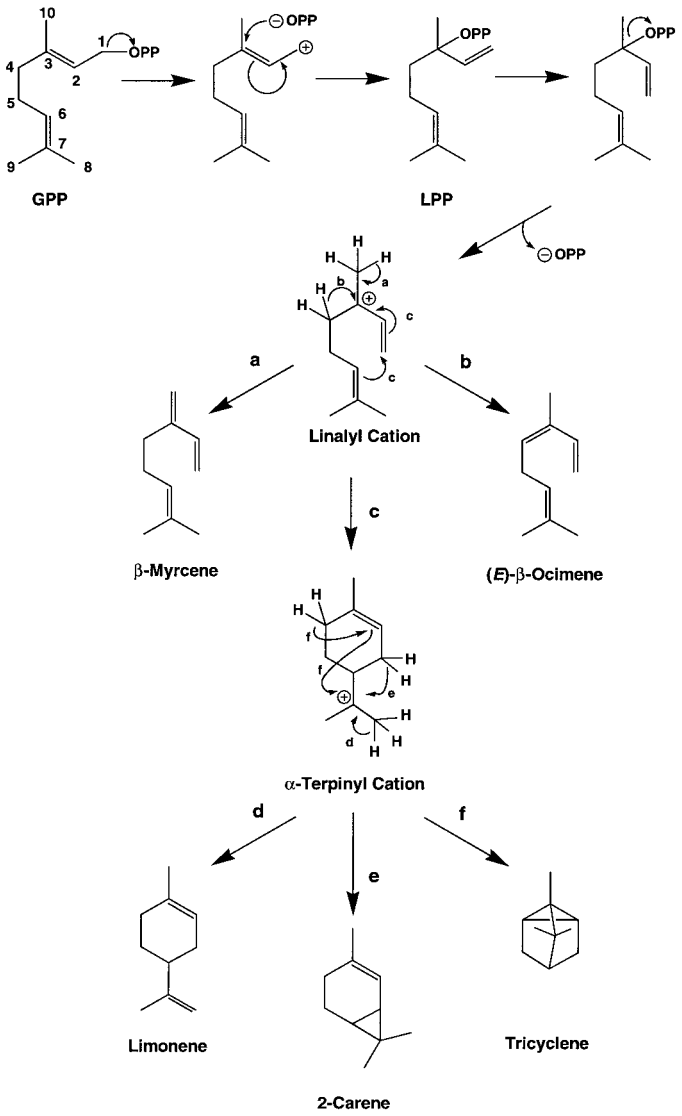


FIG. 4. Proposed reaction mechanism for the conversion of geranyl diphosphate (GPP) to myrcene and (*E*)-β-ocimene. OPP denotes the diphosphate moiety. Formation of monoterpenoids is initiated by the ionization of GPP and subsequent isomerization to form linalyl diphosphate (LPP). After ionization of LPP, the resulting linalyl carbocation can undergo deprotonation at the C3 methyl group (a) to form myrcene or at the C4 methylene group (b) to form (*E*)-β-ocimene. Formation of the monocyclic and bicyclic minor products of AtTPS10, limonene, 2-carene, and tricyclene proceeds via the α-terpinyl cation.

The catalytic mechanism of monoterpene synthases proceeds through either a (3*R*)- or (3*S*)-linalyl diphosphate intermediate, yielding products of two opposing stereochemical series (3). The formation of both (−) and (+) antipodes of limonene at a ratio of 3:2 by the *AtTPS10* gene product indicates that this enzyme does not possess complete stereochemical fidelity, a trait also found in several other monoterpene synthases

(17). Nevertheless, the acyclic nature of the major AtTPS10 products precludes determination of the major stereochemical mode of catalysis without administration of additional labeled substrates. The coproduction of myrcene and (*E*)-β-ocimene by a single terpene synthase is not surprising given the close biosynthetic relationship of these two acyclic monoterpene olefins (Fig. 4). The proposed reaction mechanism for all monoterpene synthases is thought to be initiated by the ionization of GPP and subsequent isomerization to form (3*R*)- or (3*S*)-linalyl diphosphate (LPP) as an enzyme-bound intermediate (3). After ionization of LPP, the resulting linalyl carbocation can undergo deprotonation at the C3 methyl group (path a) to form myrcene or at the C4 methylene group (path b) to form (*E*)-β-ocimene. Thus, both acyclic products of this new terpene synthase can be seen to arise from a single carbocationic intermediate.

The *Arabidopsis* myrcene/ocimene synthase is the first functionally characterized monoterpene synthase of the *TPSb* group that was not cloned from a member of the Lamiaceae, and represents the only known member of this group that forms mainly acyclic monoterpenes. A myrcene synthase previously cloned from the conifer *Abies grandis* is a member of the distantly related *TPSd* group (14) (Fig. 2), which comprises monoterpene, sesquiterpene, and diterpene synthases of gymnosperm origin. Cloning and functional characterization of *AtTPS10* confirms that functional specialization of monoterpene synthases evolved independently in angiosperms and gymnosperms (6).

Arabidopsis has not been reported to produce myrcene, ocimene, or other terpenoid secondary metabolites which occur in many plant species as components of essential oils, resins, or floral scents (42). However, other members of the Brassicaceae are known to emit myrcene and ocimene from foliar or floral tissues (43–45). Certain species, including members of the Pinaceae, Myrtaceae, Rutaceae, Lamiaceae, and Asteraceae, possess specialized anatomical structures for terpenoid biosynthesis and accumulation, such as resin ducts, secretory cavities, and glandular trichomes. The majority of plant species, including *Arabidopsis*, do not develop conspicuous compartments for terpenoid formation and storage, and may therefore produce terpenoid secondary metabolites only in small quantities, e.g., in floral tissues, or in response to environmental challenges, such as pathogen attack and damage caused by herbivores (46, 47). Based on the surprising expression of *TPSb* genes in *Arabidopsis*, additional investigations are now in progress to determine the spatial and temporal patterns of expression of members of the *AtTPS* gene family and to identify terpenoid secondary metabolites in *Arabidopsis*.

ACKNOWLEDGMENTS

We thank Tina Letsch for technical assistance, Domenica Schnabelrauch for DNA sequencing, the Ohio *Arabidopsis* Biological Resource Center for cDNA libraries, Dr. Jose Sanchez-Serrano for the jasmonate-induced cDNA library and Dr. Rodney Croteau for substrates. This work was supported by the Max Planck Gesellschaft.

REFERENCES

1. Aubourg, S., Takvorian, A., Cheron, A., Kreis, M., and Lecharny, A. (1997) *Gene* **199,** 241–253.
2. The EU Arabidopsis Genome Project (1998) *Nature* **391,** 485–488.
3. Wise, M. L., and Croteau, R. (1999) *in* Comprehensive Natural Products Chemistry: Isoprenoid Biosynthesis (Cane, D. E., Ed.), Vol. 2, pp. 97–153, Elsevier Science, Oxford.
4. Cane, D. E. (1999) *in* Comprehensive Natural Products Chemistry: Isoprenoid Biosynthesis (Cane, D. E., Ed.), Vol. 2, pp. 155–200, Elsevier Science, Oxford.
5. McMillan, J., and Beale, M. H. (1999) *in* Comprehensive Natural Products Chemistry: Isoprenoid Biosynthesis (Cane, D. E., Ed.), Vol. 2, pp. 217–243, Elsevier Science, Oxford.
6. Bohlmann, J., Meyer-Gauen, G., and Croteau, R. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 4126–4133.
7. Sun, T. P., and Kamiya, Y. (1994) *Plant Cell* **6,** 1509–1518.
8. Yamaguchi, S., Sun, T.-P., Kawaide, H., and Kamiya, Y. (1998) *Plant Physiol.* **116,** 1271–1278.
9. Bemsen, B. J., Johal, G. S., Crane, V. C., Tossberg, J. T., Schnable, P. S., Meeley, R. B., and Briggs, S. P. (1995) *Plant Cell* **7,** 75–84.
10. Yamaguchi, S., Saito, T., Abe, H., Yamane, H., Murofushi, N., and Kamiya, Y. (1996) *Plant J.* **10,** 203–213.
11. Ait-Ali, T., Swain, S. M., Reid, J. B., Sun, T. P., and Kamiya, Y. (1997) *Plant J.* **11,** 443–454.
12. Smith, M. W., Yamaguchi, S., Ait-Ali, T., and Kamiya, Y. (1998) *Plant Physiol.* **118,** 1411–1419.
13. Stofer Vogel, B., Wildung, M., Vogel, G., and Croteau, R. (1996) *J. Biol. Chem.* **271,** 23262–23268.
14. Bohlmann, J., Steele, C. L., and Croteau, R. (1997) *J. Biol. Chem.* **272,** 21784–21792.
15. Steele, C. L., Crock, J., Bohlmann, J., and Croteau, R. (1998) *J. Biol. Chem.* **273,** 2078–2089.
16. Bohlmann, J., Crock, J., Jetter, R., and Croteau, R. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 6756–6761.
17. Bohlmann, J., Phillips, M., Ramachandrian, V., Katoh, S., and Croteau, R. (1999) *Arch. Biochem. Biophys.* **368,** 232–243.
18. Wildung, M., and Croteau, R. (1996) *J. Biol. Chem.* **271,** 9201–9204.
19. Colby, S. M., Alonso, W. R., Katahira, E., McGarvey, D. J., and Croteau, R. (1993) *J. Biol. Chem.* **268,** 23016–23024.
20. Yuba, A., Yazaki, K., Tabata, M., Honda, G., and Croteau, R. (1996) *Arch. Biochem. Biophys.* **332,** 280–287.
21. Wise, M. L., Savage, T. J., Katahira, E., and Croteau, R. (1998) *J. Biol. Chem.* **273,** 14891–14899.
22. Crock, J., Wildung, M. R., and Croteau, R. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 12833–12838.
23. Facchini, P. J., and Chappell, J. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 11088–11092.
24. Back, K., and Chappell, J. (1995) *J. Biol. Chem.* **270,** 7375–7381.
25. Back, K, He, S., Kim, K. U., and Shin, D. H. (1998) *Plant Cell Physiol.* **39,** 899–904.

26. Colby, S. M., Crock, J., Dowdle-Rizzo, B., Lemaux, P. G., and Croteau, R. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 2216–2221.
27. Bohlmann, J., and Croteau, R. (1999) *in* Insect-plant Interactions and Induced Plant Defense (Chadwick, D. J., and Goode, J. A., Eds.), Novartis Found. Symp. 223, pp. 132–149, Wiley, Chichester.
28. Phillips, M. A., and Croteau, R. B. (1999) *Trends Plant Sci.* **4,** 184–190.
29. McCaskill, D., and Croteau, R. (1999) *Nature Biotechnol.* **17,** 31–36.
30. Threlfall, D. R., and Whitehead, I. M. (1991) *in* Ecological Chemistry and Biochemistry of Plant Terpenoids (Harborne, J. B., and Thomas-Barberan, F. A., Eds.), pp. 159–208, Clarendon Press, Oxford.
31. Croteau, R., Alonso, W. R., Koepp, A. E., and Johnson, M. A. (1994) *Arch. Biochem. Biophys.* **309,** 184–192.
32. Dehal, S. S., and Croteau, R. (1988) *Arch. Biochem. Biophys.* **261,** 346–356.
33. LaFever, R. E., Stofer Vogel, B., and Croteau, R. (1994) *Arch. Biochem. Biophys.* **313,** 139–149.
34. Titarenko, E., Rojo, E., Leon, J., and Sanchez-Serrano, J. (1997) *Plant Physiol.* **115,** 817–826.
35. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., pp. 7.20–7.43, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
36. Feinberg, A. P., and Vogelstein, B. (1984) *Anal. Biochem.* **137,** 266–267.
37. Schenk, P. M., Baumann, S., Mattes, R., and Steinbiss, H.-H. (1995) *BioTechniques* **19,** 196–198.
38. Williams, D. C., McGarvey, D. J., Katahira, E. J., and Croteau, R. (1998) *Biochemistry* **37,** 12213–12220.
39. Zeng, G. (1998) *BioTechniques* **25,** 206–208.
40. Weigel, D., Alvarez, J., Smyth, D. R., Yanofsky, M. F., and Meyerowitz, E. M. (1992) *Cell* **69,** 843–859.
41. McGarvey, D. J., and Croteau, R. (1995) *Plant Cell* **7,** 1015–1026.
42. Connolly, J. D., and Hill, R. A. (1991) Dictionary of Terpenoids, Chapman & Hall, London.
43. Tollsten, L., and Bergstrom, G. (1988) *Phytochemistry* **27,** 4013–4018.
44. Mattiacci, L., Dicke, M., and Posthumus, M. A. (1994) *J. Chem. Ecol.* **20,** 2229–2247.
45. McEvan, M., and MacFarlane Smith, W. H. (1998) *Clin. Exp. Allergy* **28,** 332–338.
46. Takabayashi, J., and Dicke, M. (1996) *Trends Plant Sci.* **1,** 109–113.
47. Pare, P. W., and Tumlinson, J. H. (1999) *Plant Physiol.* **121,** 325–331.
48. Saitou, N., and Nei, M. (1987) *Mol. Biol. Evol.* **4,** 406–420.