Archives of Biochemistry and Biophysics
Vol. 383, No. 2, November 15, pp. 178–184, 2000
doi:10.1006/abbi.2000.2061, available online at http://www.idealibrary.com on IDEAL®

Amorpha-4,11-diene Synthase of *Artemisia annua:* cDNA Isolation and Bacterial Expression of a Terpene Synthase Involved in Artemisinin Biosynthesis[1]

Yung-Jin Chang,[2] Seung-Hwan Song,[2] Si-Hyung Park, and Soo-Un Kim[3]
School of Agricultural Biotechnology and the Research Center for New Biomaterials in Agriculture, Seoul National University, Suwon 441-744, Korea

Received March 17, 2000, and in revised form August 7, 2000

***Artemisia annua*, an indigenous plant to Korea, contains an antimalarial sesquiterpene, artemisinin. The first committed step of artemisinin biosynthesis is the cyclization of farnesyl diphosphate by a sesquiterpene synthase to produce an amorphane-type ring system. The aims of this research were to molecularly clone and express amorpha-4,11-diene synthase for metabolic engineering. PCR amplification of genomic DNA with a pair of primers, designed from the conserved regions of sesquiterpene synthases of several plants, produced a 184-bp DNA fragment. This fragment was used in Northern blot analysis as a probe, showing approximately 2.2 kb of a single band. Its sequence information was used to produce 2106 bp of a full-length cDNA sequence including 1641 bp of open reading frame for 546 amino acids (*kcs*12) through a rapid amplification of cDNA ends (RACE). The deduced amino acid sequence displayed 36% identity with 5-*epi*-aristolochene synthase of *Nicotiana tabacum*. A soluble fraction of *Escherichia coli* harboring *kcs*12 catalyzed the cyclization of farnesyl diphosphate to produce a sesquiterpene, which was identified through GC–MS analysis as amorpha-4,11-diene.**

Key Words: **amorpha-4,11-diene synthase; sesquiterpene synthase; rapid amplification of cDNA ends; artemisinin biosynthesis; *Artemisia annua*.**

Artemisinin, produced in sweet wormwood *Artemisia annua* (1), is chemically an amorphane-type sesquiterpene endoperoxide. As a potent antimalarial principle, artemisinin is the last effective defense against sometimes deadly resistant strains of *Plasmodium falciparum* where other common antimalarials have so far failed to work. Nevertheless, on a global scale, artemisinin and its derivatives remain generally unavailable because large-scale isolations of artemisinin from the plant are possible only in a few countries. As such, several research groups have been searching for an alternative method of producing artemisinin in *A. annua* through improved cell culture conditions (2). However, thus far, commercial production of the compound through the manipulation of culture conditions has mostly failed. Therefore, a thorough understanding of the biosynthetic pathway and characterization of the involved enzymes are important for the biological production of artemisinin, such as cell culture and metabolic engineering of the plant.

The first committed step of artemisinin biosynthesis should be the cyclization of farnesyl diphosphate (FDP)[4] by a sesquiterpene synthase to produce a ring system. Although the complete biosynthetic pathway for artemisinin has not yet been established, it has been suggested that artemisinic acid is the biogenic precursor of artemisinin (12, 13). A close structural relationship between artemisinic acid and amorpha-4,11-diene placed the latter as the most plausible candidate for the cyclization product. Detection and partial purification of amorpha-4,11-diene synthase from the plant have been previously reported (14). The low level of amorpha-4,11-diene in the plant and the high

[1] The nucleotide sequence data reported in this article have been deposited with the GenBank/EMBL Database under Accession No. AJ251751.

[2] The first two coauthors share equal credit in this investigation.

[3] To whom correspondence and reprint requests should be addressed. Fax: 82-31-293-8608. E-mail: soounkim@plaza.snu.ac.kr.

[4] Abbreviations used: FDP, farnesyl diphosphate; RACE, rapid amplification of cDNA ends.

0003-9861/00 $35.00

amorphadiene synthase activity are suggested by the authors to be strong evidence that amorpha-4,11-diene is an intermediate in the biosynthesis of artemisinin.

In the biosynthetic pathway of terpenes, various terpene synthases catalyze cyclization reactions, converting a few allylic diphosphates into a surprising array of cyclic products (3). Gene cloning of several sesquiterpene synthases from different plant sources has been reported: 5-*epi*-aristolochene synthase from *Nicotiana tabacum* (4) and *Capsicum annuum* (5), vetispiradiene synthase from *Hyoscyamus muticus* (6) and *Solanum tuberosum* (7), (+)-δ-cadinene synthase from *Gossypium arboreum* (3), and germacrene C synthase from *Lycopersicon esculentum* cv.VFNT Cherry tomato (8). These plant terpene synthases exhibit a significant degree of similarity at the amino acid level (5). The molecular biological aspect in the biogenesis of artemisinin has begun to receive attention recently as reflected by the recent three reports on the terpene synthase genes from the plant. However, one of the genes turned out to be a monoterpene synthase, (3*R*)-linalool synthase (9), and the other two reports were on the identical sesquiterpene synthase producing *epi*-cedrol and cedrol, their presence in the plant under question (10, 11).

In an effort to characterize this important enzyme in *A. annua*, we cloned a terpene synthase from genomic DNA using PCR technique with a pair of primers designed from the conserved regions of sesquiterpene synthases of several plants. The sequence of this PCR fragment was used for the rapid amplification of cDNA ends (RACE). Here we report on the isolation of a full-length cDNA sequence of a sesquiterpene synthase and a subsequent functional characterization of the cloned enzyme as amorpha-4,11-diene synthase through GC–MS analysis of the reaction product.

MATERIALS AND METHODS

Plant materials. The leaves of *A. annua* used in this study were harvested in March 1999 from stock (15) maintained in the greenhouse of the College of Agriculture and Life Sciences, Seoul National University, Suwon, Korea.

Polymerase chain reactions. For genomic DNA preparation, young leaves (1 g) were homogenized with 5 ml extraction buffer (7 M urea, 0.35 M NaCl, 50 mM Tris/pH 8.0, 20 mM EDTA, 1% sarkosyl, and 0.6% SDS) (16). Guanidine isothiocyanate was used to obtain total RNA from the leaves of *A. annua* (17). Polyadenylated RNA was subsequently prepared from total RNA using the PolyATract mRNA Isolation System III (Promega, Madison, WI). Primers F7 (5′-CCATGGTGGAAAGATTTGGATTT) and B9 (5′-CCATGGCATAAGCATCGAATGTGTCATC) were selected from the conserved regions of four reported sesquiterpene synthases: *N. tabacum* 5-*epi*-aristolochene synthase (4), *H. muticus* vetispiradiene synthase (6), *L. esculentum* germacrene C synthase (8), and *G. arboreum* (+)-δ-cadinene synthase (3). At the 5′ end of each primer, a *Nco*I site was attached (underlined). PCR amplification was done in a 100-μl total volume containing 100 ng *A. annua* genomic DNA, 0.8 μM each primer, and 2 units *Taq* DNA polymerase (Promega) in a reaction buffer composed of 10 mM Tris–HCl/pH 9.0, 50 mM KCl, 2.5 mM $MgCl_2$, 0.2 mM each deoxyribose trinucleotide, and 0.1% Triton X-100. PCR was cycled 40 times through the following reaction: 2 min (or 4 min at the first step) denaturation at 94°C, 0.5 min annealing at 50°C, and 1.5 min (or 7 min at the final step) extension at 72°C. PCR product, 184 bp long, was cloned into a pGEM-T Easy Vector (Promega) and designated as p76. Sequencing of both strands was done with an ABI PRISM dye terminator cycle sequencing kit and a DNA sequencing system (ABI Model 373A, Applied Biosystems, Perkin–Elmer).

Southern and Northern blot analyses. The genomic DNA (20 μg/lane), digested with *Bam*HI, *Eco*RI, or *Hin*dIII, was separated through 0.8% agarose gel electrophoresis. The gel was depurinated, denatured, and neutralized (17). Polyadenylated RNA (5 μg/lane) was electrophoresed on 1.2% agarose gels containing formaldehyde (17). DNAs in the gel were transferred to N^+ nylon membrane using 6× SSC (1× SSC is 150 mM NaCl and 15 mM Na–citrate/pH 7.0) overnight followed by UV fixation for Southern blottings. The same procedure was followed for Northern blottings. The probe was prepared by a random primer extension of p76 with [α-^{32}P]dATP (3000 Ci/mmol; Amersham–Pharmacia) (17). The membrane for Southern blot analysis was prehybridized with a hybridization solution in 50% formamide (16 mM Na_2HPO_4, 34 mM NaH_2PO_4, 800 mM NaCl, 1 mM EDTA, 1% SDS, 0.5 mg/ml bovine serum albumin, 0.5 mg/ml Ficoll, 0.5 mg/ml PVP, and 50 μg/ml salmon sperm DNA) at 42°C for 1 h and was then hybridized with the probe at 42°C for 16 h. The membrane was washed twice with 2× SSC containing 0.1% SDS at 50°C for 15 min and twice with 0.5× SSC containing 0.1% SDS at 42°C for 15 min, before exposure to X-ray film for 2 days. For Northern blot analysis, the membrane was prehybridized with DIG Easy Hyb solution (Boehringer Mannheim, Indianapolis, IN) at 50°C for 1 h and was then hybridized with the probe at 50°C for 24 h. The membrane was washed four times with 2× SSC containing 0.1% SDS at 50°C for 15 min before exposure to X-ray film for 7 days.

RACE. A double-strand cDNA for 3′-RACE was obtained using the Universal RiboClone cDNA synthesis system (Promega) and AP primer containing UAP (5′-GGCCACGCGTCGACTAC) and poly(dT)$_{17}$ (18). Nested primer IPF (5′-GTCGACTTGAGCCACAGTATTC) and UAP were used for the PCR described above with the modifications of 35 cycles of 1 min denaturation at 94°C, 1 min annealing at 50°C, and 2 min extension at 72°C. A single-strand cDNA for 5′-RACE was obtained using the same RiboClone cDNA synthesis system and poly(dT)$_{15}$ primer. Ligation of the single-strand cDNA with 5′-NV primer, whose 3′- and 5′-ends were, respectively, aminated and phosphorylated, was followed as described elsewhere (18). Nested primer IPB (5′-GTCGACAACAATTCTATCTCTT) and 5′-3 (18) were used for PCR under the same condition as 3′-RACE. At the 5′-end of IPF and IPB, the *Sal*I site was attached (underlined).

Bacterial expression and enzyme assays. Based on the sequences from RACE, a pair of START (5′-CGGGATCCATGTCACTTACAGAAG) and STOP (5′-CGGAATTCTCATATACTCATAGGA) primers, which included the restriction sites (underlined) for *Bam*HI and *Eco*RI, respectively, were designed to obtain a PCR product of the open reading frame. The PCR product (1641 bp) was cloned into a pET-5a expression vector (Promega), and the resulting plasmid was designated as *kcs*12. *Escherichia coli* BL21(DE3)pLysS cells (Promega) with or without the plasmid *kcs*12 were grown overnight at 37°C in an LB medium supplemented with ampicillin (100 μg/ml). The overnight culture (0.5 ml) was used to inoculate 50 ml of the fresh medium. Each culture was grown at 37°C with vigorous agitation until A_{600} reached 0.5, induced with 1 mM isopropyl-β-D-thiogalactosidase, and allowed to grow further for 3 h at 37°C. The suspended cells were centrifuged (1000*g*, 15 min, 4°C), and the pellet was resuspended in 0.5 ml of lysis buffer (50 mM Tris/pH 8.0, 1 mM EDTA, 100 mM NaCl, and 1 mM dithiothreitol) and disrupted by sonication (Sonic & Materials VC-600 microprobe at maximum power for three 20-s bursts with a 1-min chilling period on ice between bursts). Protein concentration was determined through the Bradford method (19) using bovine serum albumin as the standard.

An aliquot of the bacterial extract was incubated in the reaction

mixture containing 200 mM Tris/pH 7.5, 40 mM $MgCl_2$, 0.2 μM [1-^{3}H]FDP (108 mCi/mmol, NEN, Boston, MA), and 36 μM FDP (Sigma, St. Louis, MO). After incubation for 1 h at 35°C, the reaction mixture was extracted with hexane. The hexane extract was allowed to pass through a small silica gel filter constructed with a Pasteur pipet and silica gel 50G (Merck, Darmstadt, Germany) to remove any polar materials such as unreacted FDP or farnesol generated by the phosphatase activity. The radioactivity remaining in the extract, representing the sesquiterpene product, was determined by scintillation counting. The activity was expressed as nmol of sesquiterpene product $\cdot$ h^{-1} $\cdot$ mg $protein^{-1}$. SDS–PAGE on a 10% running gel was carried out using the Laemmli method (20). The gel was stained with Coomassie brilliant blue.

For the identification of the expressed enzyme reaction products, a large-scale culture was carried out. Crude enzyme extracted from 50 ml culture was incubated with cold 36 μM FDP at 37°C overnight. The reaction mixture was further extracted with hexane, filtered through the silica gel column, and concentrated into a minimum volume for GC–MS analysis. The product was identified through comparison of its retention time and mass spectrum with those of the authentic standard (see below). GC–MS conditions were as follows: HP5890 series II system (Hewlett–Packard, Germany) equipped with an AX505WA mass spectrometer (JEOL, Japan) operating in an electron impact mode; injector temperature, 250°C; flow rate, 2 ml/min; split ratio, 5:1; oven temperature, 100°C for 2 min, 5°C/min increase to 150°C, 10°C/min increase to 250°C, and 250°C for 2 min; N_2 carrier gas; and DB-5 column, 30 m $\times$ 0.25 μm $\times$ 0.25 mm.

Synthesis of amorpha-4,11-diene. Artemisinic acid (1.5 g), isolated from the methanol extract of *A. annua* (dry wt 500 g) as described previously (15, 21), was directly reduced to the corresponding alcohol with diisobutyl aluminum hydride (22). The alcohol was tosylated with *p*-toluenesulfonyl chloride (23) and then reduced with lithium aluminum hydride (24) to yield amorpha-4,11-diene. The product was subsequently isolated through preparative silica gel TLC (15% $AgNO_3$; developing solvents, hexane:benzene:diethyl ether = 50:50:1; R_f = 0.75), and the structure was confirmed through NMR and MS analyses. NMR spectra were measured on an LA400 NMR spectrometer (400 MHz for ^{1}H and 100 MHz for ^{13}C; JEOL). The complete assignments of the spectral data are as follows (14, 25).

Amorpha-4,11-diene. ^{1}H-NMR (400 MHz, $CDCl_3$) δ, 0.88 (3H, *d*, 6.3 Hz, H-15), 0.97 (1H, *dddd*, 3.4, 10.8, 12.1, 12.7 Hz, H-9ax), 1.27 (1H, *dddd*, 3.4, 11.0, 12.1, 12.7 Hz, H-8ax), 1.32 (1H, *m*, H-2), 1.40 (1H, *ddd*, 3.9, 6.3, 10.8 Hz, H-10), 1.52 (1H, *dddd*, 1.4, 3.4, 3.4, 12.7 Hz, H-8eq), 1.56 (1H, *m*, H-2ax), 1.60 (3H, *s*, H-14), 1.67 (1H, *dddd*, 3.4, 3.4, 3.6, 12.7 Hz, H-9eq), 1.74 (3H, *s*, H-12), 1.79 (1H, *m*, H-1), 1.87 (1H, *m*, H-3), 1.95 (1H, *m*, H-7), 1.97 (1H, *m*, H-3), 1.99 (1H, *m*, H-7), 2.55 (1H, *br. s*, H-6), 4.64 (1H, *s*, H-13), 4.87 (1H, *s*, H-13), 5.06 (1H, *s*, H-5); ^{13}C-NMR (100 MHz, $CDCl_3$) δ, 19.8 (C-15), 22.6 (C-14), 23.6 (C-12), 25.8 (C-8), 26.1 (C-2), 26.5 (C-3), 27.9 (C-10), 35.4 (C-9), 37.6 (C-6), 41.8 (C-1), 47.6 (C-7), 109.8 (C-13), 120.9 (C-5), 134.6 (C-4), 148.0 (C-11); EI–MS (70 eV) *m/z* 204 ([M]$^+$, 83), 119 (100).

RESULTS

Partial sequence of terpene synthase. A fragment of amorpha-4,11-diene synthase gene from *A. annua* was amplified by PCR employing the following strategy. Based on the consensus sequences of the four reported plant sesquiterpene synthases (3, 4, 6, 8), several primers for PCR were designed. Among them, a pair of primers, F7 and B9, produced a PCR product from the genomic DNA under the PCR condition described above. The size of the PCR product was 184 bp long, suggesting that no intron existed between F7 and B9 primers as predicted from the intron positions of *N. tabacum* 5-*epi*-aristolochene synthase (4). This major PCR product, designated as p76, was separated on 1.0% agarose gel electrophoresis (Fig. 1A), cloned into a pGEM-T vector, and finally sequenced. The comparison of the nucleotide sequence, 133 bp long excluding the consensus sequence of the primers, and the deduced amino acid sequence of p76 with those of *N. tabacum* 5-*epi*-aristolochene synthase displayed 56.4 and 52.3% identities, respectively. For expression and size determination of the terpene synthase mRNA in *A. annua*, Northern analysis probed with p76 was carried out, showing an mRNA band of approximately 2.2 kb long from the leaves of *A. annua* (Fig. 1B). For the determination of genomic structure, Southern analysis was performed with p76 as a probe on the genomic DNA. The DNA digested with *Bam*HI, *Eco*RI, and *Hin*dIII showed single bands of 13.1, 7.8, and 5.9 kb long, respectively (Fig. 1C), suggesting the presence of a single copy gene in the genome of *A. annua*. Based on these data, this p76 was suggested, with a high possibility, to be a fragment of a terpene synthase gene with an unknown function.

Identification of a full-length cDNA. The sequences of a 1291-bp-long 3′-RACE (sequence readings from five colonies) and 878-bp long 5′-RACE products (sequence readings from three colonies) overlapped exactly with each other in 31 bases and included the partial nucleotide sequence of p76. Only 3 among the

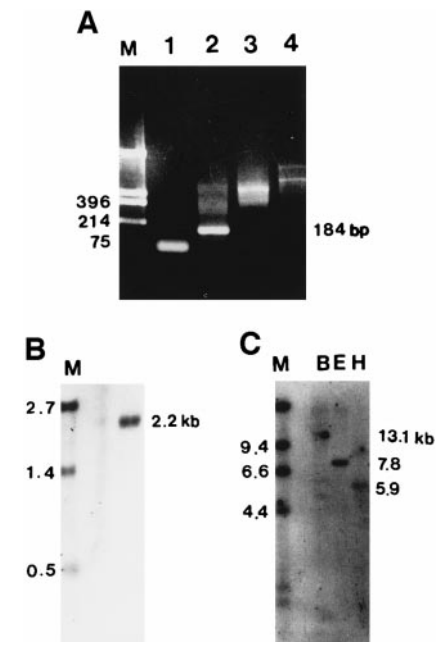

FIG. 1. Northern and Southern hybridizations with a p76 probe. (A) Lane 1, control PCR without *A. annua* genomic DNA template; lane 2, PCR of genomic DNA with a pair of primers, F7 and B9, producing a 184-bp-long PCR product; lane 3, PCR of genomic DNA with primer F7; lane 4, PCR of genomic DNA with primer B9. (B) Northern analysis with a p76 probe. (C) Southern analysis of genomic DNA digested with *Bam*HI (B), *Eco*RI (E), and *Hin*dIII(H).

17 bases in F7 primer and 4 among the 22 in B9 primer, excluding the restriction site in each primer, were different from the determined sequence, as these primers were based on the consensus sequences of sesquiterpene synthase genes in other plants. Combination of 3′- and 5′-RACE sequence information yielded a full-length cDNA sequence of 2106 bp long. A consensus for the plant translation start (G/CAANNATGG) was found between positions 27 and 37. Downstream of the starting codon (ATG) is a stop codon (TGA) in frame with the open reading frame. A typical poly(A) signal (AATAAT/A) occurred twice between positions 1778 and 1786.

The deduced amino acid sequence (546 aa) of the open reading frame (1641 bp) showed calculated p*I* and M_r of 5.59 and 63.9 kDa, respectively. It did not encode a plastid targeting sequence, suggesting an exclusive localization of this sesquiterpene synthase in cytosol (26, 27). Alignment of the amino acid sequence of the sesquiterpene synthase with the known synthases showed relatively high homologies: 36 and 41% identity to tobacco 5-*epi*-aristolochene synthase (4) and cotton (+)-δ-cadinene synthase (3), respectively (Fig. 2). Comparison of the synthase with two known terpene synthases from *A. annua*, *epi*-cedrol synthase (10) and linalool synthase (9), revealed 51 and 27% identity, respectively.

Bacterial expression and characterization. To identify the function of this isolated cDNA, the open reading frame of the putative terpene synthase was expressed via PCR amplification and ligation into an expression vector. The PCR with a pair of START and STOP primers produced *kcs*12, which was cloned into the pET-5a expression vector. An insoluble fraction of the transformed *E. coli* harboring *kcs*12 showed the presence of a 63.5-kDa protein band through SDS–PAGE analysis (Fig. 3). The enzyme reaction of the soluble fraction of the transformed *E. coli* with FDP as a substrate produced amorpha-4,11-diene, which was not detected in the control harboring an empty vector (data not shown). Total ion chromatogram through GC–MS analysis showed respective retention time and molecular ion of 9.3 min and *m*/*z* 204, the retention time coinciding with the authentic standard prepared from artemisinic acid (Fig. 4). Selective ion monitoring at *m*/*z* 204 (M^+) of the reaction product also showed a single peak with the same retention time as the authentic sample. Reaction with [^{3}H]FDP showed a specific activity of 0.212 nmol · h^{-1} · mg $protein^{-1}$.

DISCUSSION

The efforts to clone the terpene synthase genes from *A. annua* resulted in the isolation of genes for *epi*-cedrol and linalool synthases (9–11). Our successful cloning of the amorpha-4,11-diene synthase gene could be attributed to the proper sampling season of March, at which time most of the other inducible synthase genes were probably silent due to the lack of elicitation (5). An elicitation could induce a nondiscriminative expression enhancement of the otherwise silent genes, which would result in a lower probability of cloning the amorpha-4,11-diene synthase gene. Our preliminary data indicated a constitutive presence of the synthase mRNA in the leaves from March through August (unpublished data). Another contribution to the successful cloning in this work could have come from the regional variations of *A. annua* strains. The involvement of the enzyme reaction product amorpha-4,11-diene in artemisinic acid biosynthesis is now accepted from several lines of evidence, aside from the structural similarity to artemisinic acid. Radiolabeling study with the whole cell extract showed that artemisinic acid is a major intermediate metabolite in the biosynthetic route from mevalonic acid to artemisinin (12). In addition, it was proposed that the amorphadiene might be present in the biogenic route to artemisinic acid biosynthesis (10). Direct detection of the hydrocarbon metabolite and the amorphadiene synthase activity from the plant extract suggest the intermediacy of amorpha-4,11-diene in the artemisinin biosynthesis of *A. annua* (14). GC–MS analysis of the hexane extract of *A. annua* leaves in this laboratory also showed the presence of amorpha-4,11-diene (data not shown).

A cyclization scheme to account for the generation of amorpha-4,11-diene has been proposed (Fig. 5), beginning with the ionization and subsequent isomerization to make *E*,*Z*-farnesyl cation (28). The double bond at C10—C11 makes electrophilic attack at C1 forming the first ring closure between C1 and C10, causing the development of a positive charge on C11 (*cis*-germacryl cation). 1,3-Hydride shift takes place to form a positive charge at C1 and induces the second electrophilic attack from the double bond at C6—C7. This electrophilic reaction then causes the second ring closure between C1 and C6 producing cadinyl cation with a charge on C7, which would subsequently form amorpha-4,11-diene through 1,5-hydride shift (Fig. 5, pathway b). An alternative pathway involving bisabolyl carbocation (Fig. 5, pathway a) has not yet received experimental support (12). With the synthase in our hands, it would be possible to distinguish one mechanism from the other. Specifically, for example, deuterium labeling at C1 would discern pathway a from pathway b by determining the position of deuterium after the putative hydride shift.

Homology at 51% between *epi*-cedrol synthase (10, 11) and amorpha-4,11-diene synthase, both from *A. annua*, was higher than the homology at 41% between amorpha-4,11-diene synthase and cotton (+)-δ-cadinene synthase with identical ring systems. Through X-ray crystallography study, the mechanism of tobacco *epi*-aristolochene synthase was determined (29). A lid of helical loop over the active site was known to be held by the hydrophobic interaction between

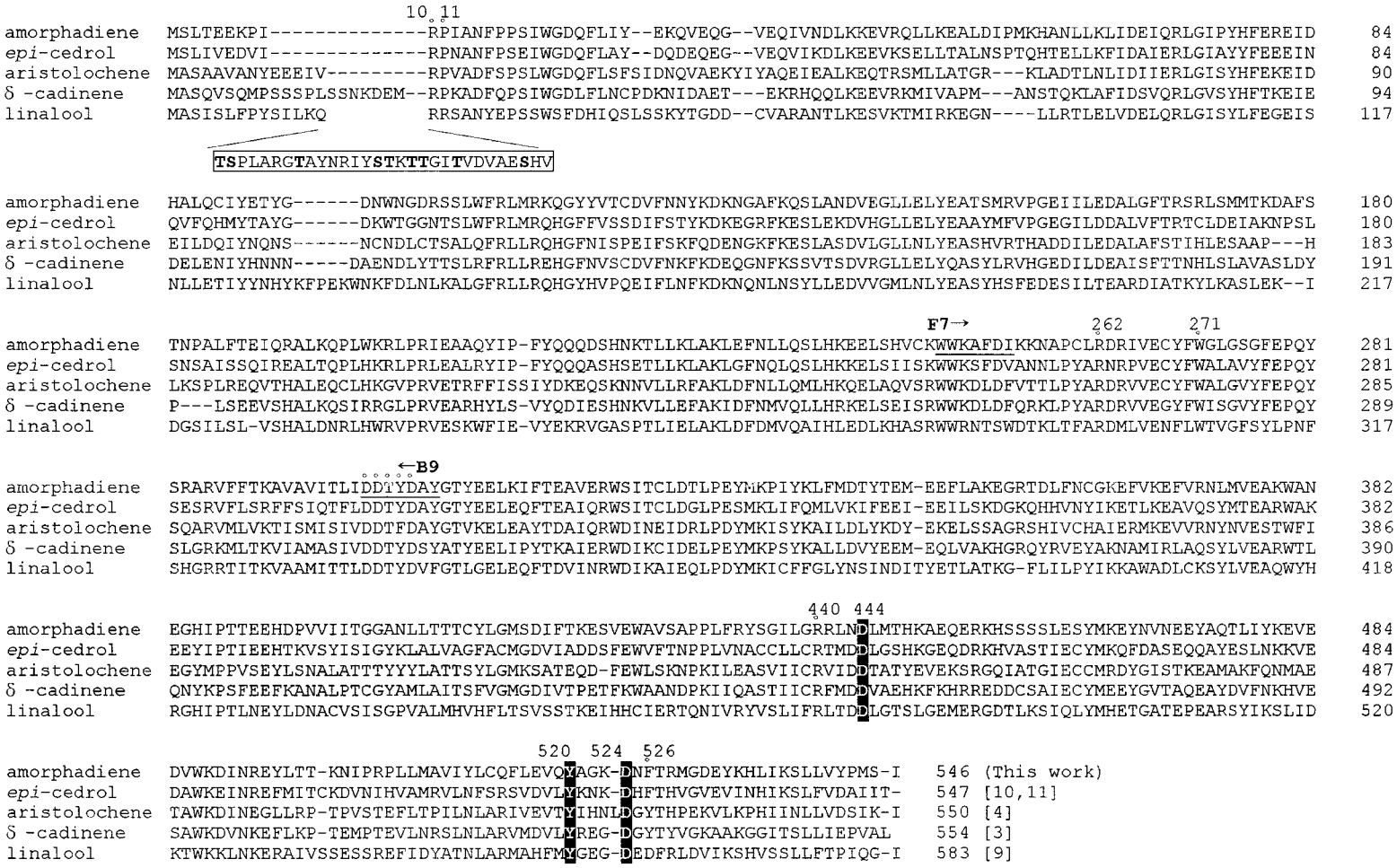

FIG. 2. Alignment of the deduced amino acid sequences of the plant mono- and sesquiterpene synthases. Amino acid residues of amorpha-4,11-diene synthase discussed in the text are labeled with open circles and residue numbers. Amino acids involved in the catalytic triad (29) are shown in black boxes. Plastid-targeting sequences, common features of monoterpene synthases, are open boxed. Regions for the pair of PCR primers, F7 and B9, are underlined with forward and backward arrows, respectively, for polymerization direction.

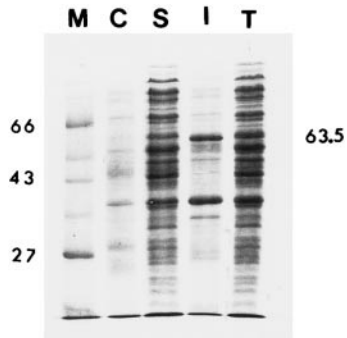

FIG. 3. Ten percent SDS–polyacrylamide gel electropherogram of the enzyme extract from transformed *E. coli* harboring *kcs*12 (S, soluble fraction; I, insoluble fraction; T, total fraction) and the control harboring an empty vector (C). A protein band with M_r of 63.5 kDa is clearly shown. (M) Molecular weight marker.

Trp^{273} and Tyr^{527}, which were conserved as Trp^{271} and Phe^{526} in amorpha-4,11-diene synthase, respectively. Carbocation intermediates, through a complex cyclization cascade, are stabilized by the nucleophilicity of the indole ring of Trp^{271} in amorpha-4,11-diene synthase. Upon loop closure, positive charge on Arg^{262} and Arg^{440} in amorpha-4,11-diene synthase facilitates ionization of the substrate FDP with the assistance of Mg^{2+} coordinated by the aspartate-rich metal binding motif (DDxxD, positions 299 through 303 in amorpha-4,11-diene synthase). The Arg^{10}-Pro^{11} pair motif, common to all cloned sesquiterpene synthases, is also conserved in amorpha-4,11-diene synthase. The catalytic triad Asp^{444}, Tyr^{520}, and Asp^{525} corresponds to the positions of Asp^{444}, Tyr^{520}, and Asp^{524} of this enzyme, respectively.

The synthase thus obtained was found to have a M_r of 63.5 kDa through SDS–electrophoresis, whereas the partially purified enzyme reported by Boumeester *et al.* (14) has a M_r of 56 kDa as determined by gel filtration. It is not certain whether this discrepancy is due to the compactness of the native protein conformation with a high α-helix content, which was shown through X-ray crystallography analysis (29). Research to purify and characterize the expressed enzyme is underway. Special emphasis will be put on the identification of the cyclization inter-

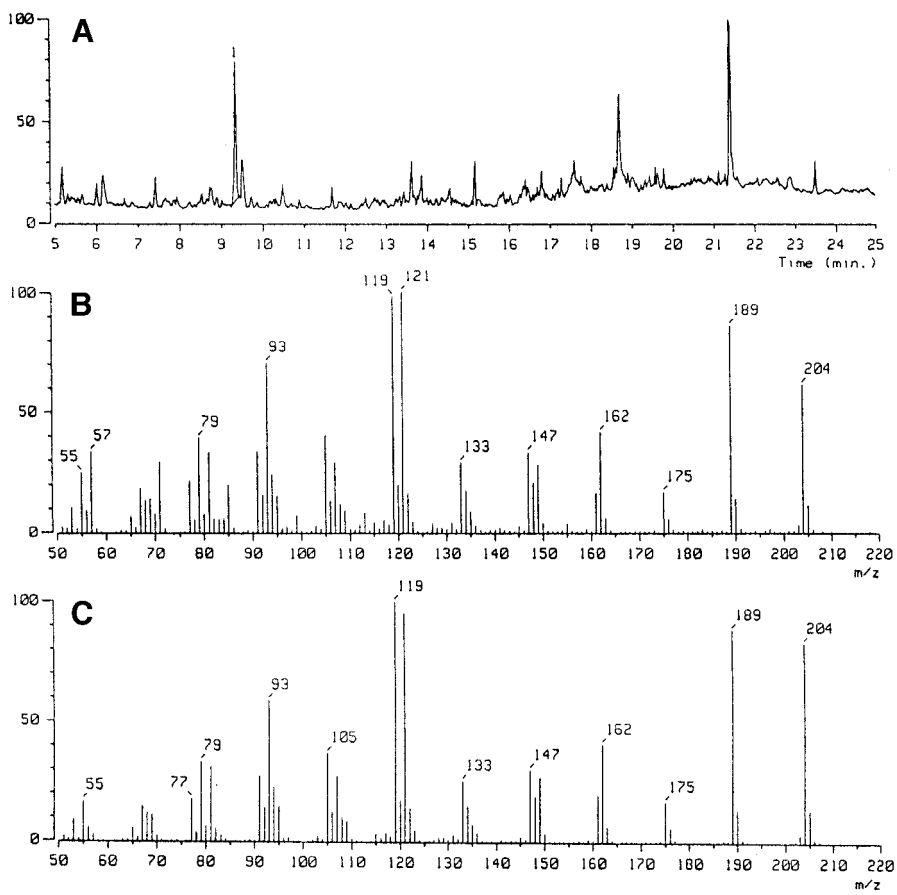

FIG. 4. GC–MS analysis of the reaction product of amorpha-4,11-diene synthase. (A) Gas chromatogram of the enzyme reaction product using FDP as a substrate. (B) Mass spectrum of the amorpha-4,11-diene synthase reaction product derived from FDP (R_t = 9.3 min). (C) Mass spectrum of the authentic amorpha-4,11-diene (R_t = 9.3 min).

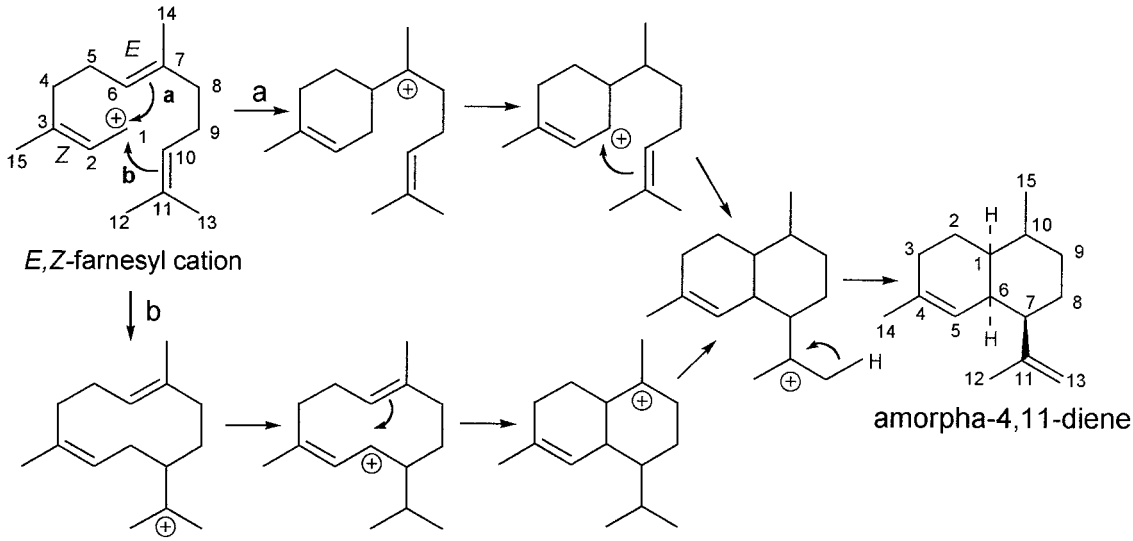

FIG. 5. Proposed cyclization scheme for the generation of amorpha-4,11-diene. Ionized *E,Z*-farnesyl cation cyclizes to either bisabolyl cation (pathway a) or *cis*-germacryl cation (pathway b). 1,3-Hydride shift takes place to form the second cyclization, eventually producing amorpha-4,11-diene.

mediates leading to amorpha-4,11-diene to elucidate the enzyme reaction mechanism. This work is expected to open the way to enhancement of the biosynthetic efficiency of artemisinin in *A. annua* through metabolic engineering.

ACKNOWLEDGMENTS

This work was supported by the Korea Science and Engineering Foundation (KOSEF 981-0608-040-2) and the Brain Korea 21 program administered by the Graduate School of Agricultural Biotechnology, S.N.U., through the support of the Ministry of Education, Korea. Y.J.C. is grateful for a fellowship provided by KOSEF through the Research Center for New Biomaterials in Agriculture.

REFERENCES

1. Klayman, D. L. (1985) *Science* **228,** 1049–1055.
2. Weathers, P. J., Cheetham, R. D., Follansbee, E., and Teoh, K. (1994) *Biotechnol. Lett.* **16,** 1281–1286.
3. Chen, X. Y., Chen, Y., Heinstein, P., and Davisson, V. J. (1995) *Arch. Biochem. Biophys.* **324,** 255–266.
4. Facchini, P. J., and Chappell, J. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 11088–11092.
5. Back, K., He, S., Kim, K. U., and Shin, D. H. (1998) *Plant Cell Physiol.* **39,** 899–904.
6. Back, K., and Chappell, J. (1995) *J. Biol. Chem.* **270,** 7375–7381.
7. Yoshioka, H., Yamada, N., and Doke, N. (1999) *Plant Cell Physiol.* **40,** 993–998.
8. Colby, S. M., Crock, J., Dowdle-Rizzo, B., Lemaux, P. G., and Croteau, R. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 2216–2221.
9. Jia, J.-W., Crock, J., Lu, S., Croteau, R., and Chen, X.-Y. (1999) *Arch. Biochem. Biophys.* **372,** 143–149.
10. Mercke, P., Crock, J., Croteau, R., and Brodelius, P. E. (1999) *Arch. Biochem. Biophys.* **369,** 213–222.
11. Hua, L., and Matsuda, S. P. T. (1999) *Arch. Biochem. Biophys.* **369,** 208–212.
12. Akhila, A., Kumkum, R., and Thakur, R. S. (1990) *Phytochemistry* **29,** 2129–2132.
13. Sangwan, R. S., Agarwal, K., Luthra, R., Thakur, R. S., and Sangwan, S. N. (1993) *Phytochemistry* **34,** 1301–1302.
14. Bouwmeester, H. J., Wallaart, T. E., Janssen, M. H. A., van Loo, B., Jansen, B. J. M., Posthumus, M. A., Schmidt, C. O., De Kraker, J.-W., König, W. A., and Franssen, M. C. R. (1999) *Phytochemistry* **52,** 843–854.
15. Kim, S.-U., and Choi, J.-H. (1993) *J. Nat. Prod.* **56,** 859–863.
16. Shure, M., Wessler, S., and Fedoroff, N. (1983) *Cell* **35,** 225–233.
17. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, NY.
18. McPherson, M. J., Hames, B. D., and Taylor, G. R. (1995) *in* PCR 2: A Practical Approach, pp. 89–118, IRL Press, Oxford Univ. Press, NY.
19. Bradford, M. (1976) *Anal. Biochem.* **72,** 248–254.
20. Laemmli, U. K. (1970) *Nature* **227,** 680–685
21. Kim, S.-U., and Lim, H.-J. (1989) *J. Kor. Agric. Chem. Soc.* **32,** 178–179.
22. Daniewski, A. R., and Wojciechowska, W. (1982) *J. Org. Chem.* **47,** 2993–2995.
23. Kabalka, G. W., Varma, M., Varma, R. S., Srivastava, P. C., and Knapp, F. F., Jr. (1986) *J. Org. Chem.* **51,** 2386–2388.
24. Zorbach, W. W., and Tio, C. O. (1961) *J. Org. Chem.* **26,** 3543–3545.
25. Bohlmann, F., Gerke, T., Jakupovic, J., King, R. M., and Robinson, H. (1984) *Phytochemistry* **23,** 1183–1184.
26. Bohlmann, J., Meyer-Gauen, G., and Croteau, R. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 4126–4133.
27. McGarvey, D. J., and Croteau, R. (1995) *Plant Cell* **7,** 1015–1026.
28. Dewick, P. M. (1997) *in* Medicinal Natural Products: A Biosynthetic Approach, pp. 172–184, Wiley, New York.
29. Starks, C. M., Back, K., Chappell, J., and Noel, J. P. (1997) *Science* **277,** 1815–1820.