Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 2216–2221, March 1998
Biochemistry

Germacrene C synthase from *Lycopersicon esculentum* cv. VFNT Cherry tomato: cDNA isolation, characterization, and bacterial expression of the multiple product sesquiterpene cyclase

(elemene/terpenoid cyclase/sesquiterpene synthase/farnesyl diphosphate cyclization)

SHEILA M. COLBY*†, JOHN CROCK‡, BARBARA DOWDLE-RIZZO*, PEGGY G. LEMAUX*, AND RODNEY CROTEAU‡§

*U.S. Department of Agriculture/University of California, Berkeley, Plant Gene Expression Center and Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720; and ‡Institute of Biological Chemistry and the Department of Biochemistry and Biophysics, Washington State University, Pullman, WA 99164-6340

Contributed by Rodney Croteau, December 29, 1997

ABSTRACT **Germacrene C was found by GC-MS and NMR analysis to be the most abundant sesquiterpene in the leaf oil of *Lycopersicon esculentum* cv. VFNT Cherry, with lesser amounts of germacrene A, guaia-6,9-diene, germacrene B, β-caryophyllene, α-humulene, and germacrene D. Soluble enzyme preparations from leaves catalyzed the divalent metal ion-dependent cyclization of [1-^{3}H]farnesyl diphosphate to these same sesquiterpene olefins, as determined by radio-GC. To obtain a germacrene synthase cDNA, a set of degenerate primers was constructed based on conserved amino acid sequences of related terpenoid cyclases. With cDNA prepared from leaf epidermis-enriched mRNA, these primers amplified a 767-bp fragment that was used as a hybridization probe to screen the cDNA library. Thirty-one clones were evaluated for functional expression of terpenoid cyclase activity in *Escherichia coli* by using labeled geranyl, farnesyl, and geranylgeranyl diphosphates as substrates. Nine cDNA isolates expressed sesquiterpene synthase activity, and GC-MS analysis of the products identified germacrene C with smaller amounts of germacrene A, B, and D. None of the expressed proteins was active with geranylgeranyl diphosphate; however, one truncated protein converted geranyl diphosphate to the monoterpene limonene. The cDNA inserts specify a deduced polypeptide of 548 amino acids (M_r = 64,114), and sequence comparison with other plant sesquiterpene cyclases indicates that germacrene C synthase most closely resembles cotton δ-cadinene synthase (50% identity).**

Volatile metabolites of *Lycopersicon* species are of interest because of their roles in tomato flavor and in host defense against arthropod herbivores (1–5). Whereas few volatile terpenoids are found in fruit (1, 2), the leaf glandular trichomes produce a rich spectrum of monoterpenes and sesquiterpenes. Nearly 20 monoterpenes, including limonene, have been found in the leaves of the domestic tomato *Lycopersicon esculentum* (6, 7), and most are also present in wild tomato species (7). A number of C13 norsesquiterpenoid glycosidic ethers have been reported in tomato fruit (8); these compounds are derived by degradation of carotenoids (9). The sesquiterpene content of tomato leaf oil varies considerably among species, with caryophyllene and humulene being widespread and reported from *L. esculentum*, *Lycopersicon hirsutum*, *Lycopersicon pimpinellifolium*, *Lycopersicon peruvianum*, *Lycopersicon cheesmanii*, *Lycopersicon chilense*, and *Lycopersicon chumielewski* (7). The epoxides of both of these sesquiterpenes are also present in *L. esculentum*, as is a low level of δ-elemene (6). Various accessions of *L. hirsutum* contain α-copaene, γ-elemene, zingiberene, and α-santalene as major leaf oil sesquiterpenes (7). Germacrenes have not been reported in the genus *Lycopersicon*.

The terpenoid composition of the highly disease-resistant *L. esculentum* cv. VFNT Cherry (a tomato of multispecies pedigree carrying resistance to *Verticillium dahliae, Fusarium oxysporum*, root-knot nematode, tobacco mosaic virus, and *Alternaria* stem canker) has not been examined, although it is an important breeding line (10). VFNT Cherry is a good model system for understanding the molecular basis of sesquiterpene biosynthesis and of glandular trichome-based expression of defense genes from which novel forms of resistance could be developed. In this report, we describe the monoterpene and sesquiterpene composition of VFNT Cherry tomato leaves, leaf sesquiterpene cyclase enzymology, the PCR strategy used to isolate a cDNA encoding germacrene C synthase, and the comparison of this sesquiterpene cyclase to other plant terpenoid cyclases.

MATERIALS AND METHODS

Experimental Materials. *L. esculentum* cv. VFNT Cherry tomato plants (obtained as seed from John Steffens, Cornell University) were propagated from cuttings and grown under conditions as reported (11). *L. esculentum* cv. Better Boy plants were obtained from C.A. Ryan (Washington State University) and similarly propagated. Methods have been reported for the preparation of [1-^{3}H]geranyl diphosphate (GDP; 122 Ci/mol; 1 Ci = 37 GBq) (12), [1-^{3}H]farnesyl diphosphate (FDP; 125 Ci/mol) (13), and [1-^{3}H]geranylgeranyl diphosphate (GGDP; 90 Ci/mol) (14). Terpenoid standards were from our own collection or gifts from Bob Adams (Baylor University) or Larry Cool (University of California, Berkeley). The VFNT Cherry tomato cDNA library, derived from epidermis-enriched young leaf tissue, was also a gift from John Steffens. All other biochemicals were purchased from Sigma or Aldrich, unless otherwise noted.

Volatile Terpene Analysis. Expanding tomato leaves were harvested, frozen in liquid N_2, and then subjected to simultaneous steam distillation and solvent (pentane) extraction (15) using the J & W Scientific apparatus. The resulting pentane phase, collected at 0–4°C, was passed over a column of

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: FDP, farnesyl diphosphate; GDP, geranyl diphosphate; GGDP, geranylgeranyl diphosphate; TEAS, tobacco 5-*epi*-aristolochene synthase.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF035630 and AF035631).

†Current address: Hitachi Software, 601 Gateway Boulevard, Suite 500, South San Francisco, CA 94080.
§To whom reprint requests should be addressed. e-mail: croteau@mail.wsu.edu.

$MgSO_4$-silica gel (Mallinckrodt product SilicAR-60) to remove water and oxygenated compounds and provide the hydrocarbon fraction. To examine the possibility of thermal decomposition during distillation, leaf material was extracted directly with cold (−20°C) pentane and the extract purified as before. To provide germacrene C for NMR studies, the distillate was separated by preparative TLC (20 cm × 20 cm × 1 mm, silica gel G containing 8% $AgNO_3$; plates dried, predeveloped in diethyl ether, and dried again—all steps in the dark) using benzene/hexane/acetonitrile, 60:40:5 (vol/vol), as developing solvent in the dark, with humulene, caryophyllene, and longifolene as standards. Bands were visualized under UV light after spraying with 0.2% dichlorofluorescein, and the triene region ($R_f \sim 0.3$) was separated into 5-mm strips of gel, from which the material was eluted with diethyl ether. Those fractions containing germacrene by GC analysis were concentrated under vacuum, dissolved in a minimum volume of acetonitrile, and then emulsified with one-third volume of H_2O. The emulsion was injected onto an Alltech C_{18} 5U reversed-phase column (4.6 mm i.d. × 250 mm) installed on a Spectra-Physics SP8800 HPLC liquid chromatograph and eluted isocratically with acetonitrile/water, 75:25 (vol/vol), at a flow rate of 1.0 ml/min while monitoring at 220 nm. NaCl was added to fractions containing germacrene C before extraction with pentane, and the extract was passed through a column of $MgSO_4$-silica gel before determination of purity by GC.

Terpene Synthase Isolation and Assay. Immature folded leaves (~15 g) from VFNT Cherry were frozen in liquid N_2, ground to a fine powder, and transferred to a centrifuge tube containing 15 ml of buffer [20 mM 3-(*N*-morpholino)-2-hydroxypropanesulfonic acid, pH 7.0/20 mM 2-mercaptoethanol/10 mM sodium ascorbate/1 mM EDTA/10% (vol/vol) glycerol/0.1% $Na_2S_2O_5$/1.0% polyvinylpyrrolidone ($M_r \sim$ 40,000)/1.0% polyvinylpolypyrrolidone/10 g of Amberlite XAD-4 polystyrene resin beads]. The tube was flushed with N_2, sealed, vigorously agitated, and then centrifuged at 3,200 × *g* for 15 min at 4°C. A 2-ml portion of the supernatant was transferred to a glass Teflon-sealed screw-capped tube, and the contents adjusted to 1 mM $MnCl_2$/10 mM $MgCl_2$/7.3 μM [1-^{3}H]FDP and then overlaid with 1 ml of pentane before closure and incubation at 30°C for 2 h. At the completion of the assay, the reaction mixture was extracted twice with pentane (1 ml) and the combined extract was passed through a $MgSO_4$-silica gel column to provide the terpene hydrocarbon fraction. The incubation mixture was next extracted twice with diethyl ether (1 ml), and the combined ether extract was passed through the corresponding column to provide the oxygenated terpenoids. An aliquot of each fraction was taken for liquid scintillation counting to determine conversion rate, and the remainder was concentrated under vacuum for capillary GC or radio-GC analysis. For radio-GC analysis, authentic carrier standards were added prior to solvent concentration to minimize losses. Boiled controls and incubations without divalent metal ion cofactors were included in all experiments.

PCR-Based Probe Generation and cDNA Library Screening. Comparison of the deduced amino acid sequences of a monoterpene cyclase [spearmint limonene synthase (16)], a sesquiterpene cyclase [tobacco 5-*epi*-aristolochene synthase (TEAS) (17)], and a diterpene cyclase [castor bean casbene synthase (18)] allowed the design of two degenerate oligonucleotide primers for PCR amplification based on conserved domains. The forward and reverse primers [5′-G(A)AIGGIA-(G)AA(G)TTT(C)-AAA(G)GA-3′ and 5′-T(C)TG(T)CAT-A(G)TAA(G)TCIGG(A)-IAG-3′] were used to amplify VFNT Cherry leaf library cDNA by PCR performed with the GeneAmp PCR reagent kit (Perkin–Elmer/Cetus), 0.5 μg of cDNA template, and 50 pmol of each primer per 50-μl reaction. After a hot start at 100°C for 5 min, *Taq* DNA polymerase was added, and the reactions were cycled twice at 97°C (1 min), 55°C (1 min), and 72°C (2 min) and then twice at 94°C (1 min), 53°C (1 min), and 72°C (2 min). While holding denaturing and extension temperatures constant, the annealing temperature was lowered by 2°C each two cycles to 45°C, with a final extension at 75°C for 5 min. PCR products were electrophoresed on a *tris*-(hydroxymethyl)aminomethane/$Na_2B_4O_7$/EDTA/2% agarose gel. Amplicons of the expected size (~770 bp) were isolated by electrophoretic transfer to NA 45 paper (Schleicher & Schuell). The fragments were reamplified, repurified as above, and then ligated into pCR-Script SK(+) by using the Stratagene protocol and transformed into *Escherichia coli* XL-1 Blue (Stratagene) by standard methods (19).

For Southern blot analysis, 10 μg of each isolated genomic DNA (19) from VFNT Cherry leaf and spearmint leaf (as control) was digested with *Eco*RI, *Hin*dIII, and *Bam*HI restriction enzymes, electrophoresed on *tris*-(hydroxymethyl-)aminomethane/$Na_2B_4O_7$/EDTA/0.8% agarose, and transferred to a Magna membrane (Micron Separations, Westboro, MA). The amplicon from above was excised with *Bam*HI and *Not*I, and 40 ng was radiolabeled with [α-^{32}P]dATP (Random-Primed labeling kit, U.S. Biochemical), purified over Sephadex G-50, then added to 20 ml of hybridization buffer containing 0.9 M NaCl/60 mM NaH_2PO_4/6 mM EDTA/5× Denhardt's solution/0.5% SDS/denatured salmon sperm DNA (50 μg/ml)/10% dextran sulfate. Prehybridization at 65°C (4 h) was followed by hybridization at 65°C overnight. The stringency wash was performed at 65°C in 0.3 M NaCl/20 mM NaH_2PO_4/2 mM EDTA buffer with 0.5% SDS (20 min). Blots were exposed to Kodak XAR-5 film for 12 h at −70°C.

For cDNA library screening, the purified *Not*I/*Bam*HI-digested amplicon from above was radiolabeled with ^{32}P as before and used to screen 7.5×10^5 cDNA clones (plated at 2.5×10^4 plaque-forming units per 150-mm plate) from a VFNT Cherry leaf λ cDNA library. Hybridization was conducted on nitrocellulose membranes (Schleicher & Schuell) in hybridization buffer without Denhardt's solution at 65°C (16 h). Blots were washed for two 5-min periods in 0.9 M NaCl/90 mM sodium citrate/6 mM EDTA buffer at 20°C, then for one 20-min in 0.9 M NaCl/90 mM sodium citrate/6 mM EDTA buffer with 1% SDS at 37°C, and for two 30-min periods at 65°C. Filters were exposed to Kodak XAR-5 film as above. Thirty-one positive plaques were purified through four additional cycles of hybridization, excised *in vivo* as Bluescript II phagemids, and used to infect *E. coli* XL1-Blue, from which plasmids were prepared on a purification column (Qiagen, Chatsworth, CA). The clones were digested with *Rsa*I, and the fragments were electrophoresed to sort into groups; insert size was determined by PCR using T3 and T7 promoter primers. DNA sequencing (both strands) was performed by Retrogen (San Diego, CA) or by automated DyeDeoxy Terminator cycle sequencing (Applied Biosystems 373 Sequencer). DNA sequence analysis employed programs from the Genetics Computer Group (University of Wisconsin Package, version 8.0.1-Unix), and searches were done at the National Center for Biotechnology Information by using the BLAST network service to search standard databases.

Expression of Terpenoid Synthase Activity. To evaluate functional expression of putative terpenoid synthases, *E. coli* XL1-Blue cells harboring the selected phagemid were grown, induced, harvested, and extracted as described (11), and the extracts were assayed for mono-, sesqui-, or diterpene synthase activity as before with [1-^{3}H]GDP, FDP, or GGDP as the respective substrate (no pentane overlay was used in assays containing GGDP). To obtain sufficient product for chromatographic analysis, the bacterial preparations were scaled to 400 ml, lysozyme was added (5 μg/ml, on ice 20 min) prior to sonication, and the resulting enzyme preparation was cleared by centrifugation as before and incubated with substrate overnight at 30°C. The terpenoid products were isolated for

radio-GC and GC-MS analysis and determination of conversion rate as before.

Instrumental Analyses. Liquid scintillation spectrometry and radio-GC and GC-MS analyses have been described (11). NMR analysis was performed on a Varian Unity*Plus* (599.89 MHz ^{1}H, 150.85 MHz ^{13}C) spectrometer using a Nalorac 3-mm dual ^{1}H/^{13}C probe. The experiments were run at 21°C with 100 μg of germacrene C in 120 μl of C_6D_6. ^{1}H and ^{13}C chemical shifts are reported in δ (ppm) by using tetramethylsilane as an internal standard. All two-dimensional data were obtained by using the hypercomplex phase-sensitive method (20). The double quantum filter homonuclear correlated spectrum was recorded with the standard pulse sequence (21) at a spectral window of 4,181.5 Hz. The 256 t_1 increments of 112 scans were sampled in 2,048 complex data points. Linear prediction to the 1,024 complex data points in the F_1 domain was used. Zero-filling to 2,048 complex points in the F_1 domain and Gaussian weighting functions were applied in both dimensions prior to 2K $\times$ 2K double Fourier transformation. The heteronuclear multiple quantum correlation spectrum was measured with the pulse sequence (22) at a proton spectral window of 4,181.5 Hz and carbon spectral window of 19,870.8 Hz. The 512 t_1 increments of 256 scans were sampled in 2,048 complex data points. Linear prediction to the 2,048 complex data points in the F_1 domain was used. Zero filling to 4,096 complex points in the F_1 domain and Gaussian weighting functions were applied in both dimensions prior to 2K $\times$ 4K double Fourier transformation. Similar parameters were used for the heteronuclear multiple bond correlation spectrum that was acquired with the pulse sequence of Bax and Summers (23).

RESULTS AND DISCUSSION

Volatile Terpenes of Tomato Leaf. Preliminary studies, in which leaf extracts were compared with steam distillates by standard GC-MS methods (hot injector) and by cool on-column injection, revealed that the target metabolite germacrene C (Fig. 1) was prone to thermal rearrangement to a compound with shorter GC retention time but nearly identical mass spectrum. This compound was absent in the pentane leaf extract when analyzed by cool (35°C) on-column injection but was present when the same pentane extract was injected onto a hot (230°C) injector (Fig. 2). Comparison of this compound with authentic δ-elemene (Fig. 1) from black pepper oleoresin yielded an identical retention time and mass spectrum (data not shown). Heating purified germacrene C in a sealed container to 130°C for 1 h resulted in complete conversion to δ-elemene, which was confirmed by GC-MS. Conversion of germacrene C to δ-elemene has been reported by Morikawa

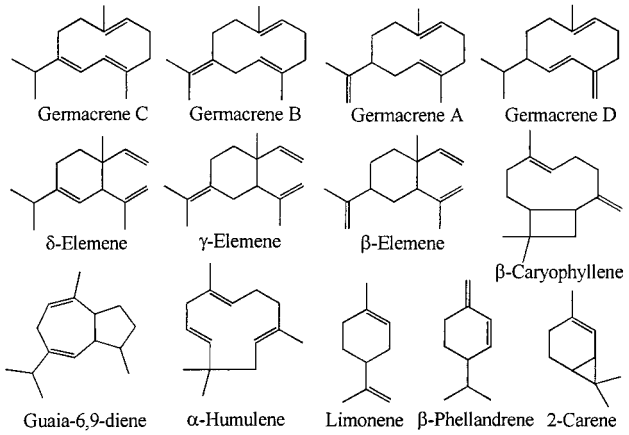

FIG. 1. Sesquiterpene and monoterpene olefins of tomato essential oil. The elemenes formed by Cope rearrangement are shown beneath the corresponding germacrenes.

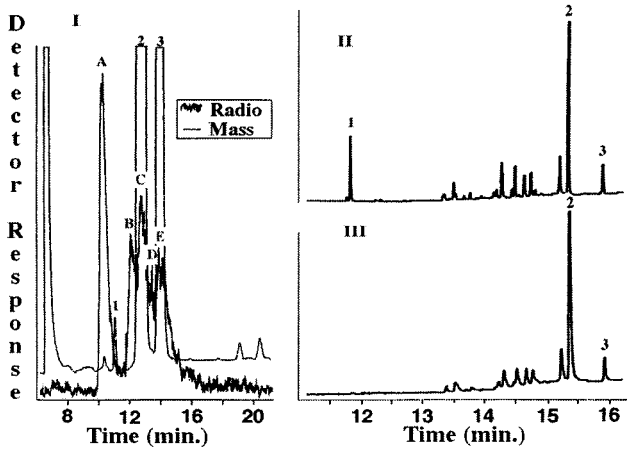

FIG. 2. (*I*) Radio-GC analysis of the olefinic products generated from [1-^{3}H]FDP by an enzyme preparation from VFNT Cherry leaves. Radiolabeled products are δ-elemene (peak A), β-caryophyllene (peak C), and α-humulene (peak E). The authentic standards are longifolene (peak 1), β-caryophyllene (peak 2), and α-humulene (peak 3). (*II*) Capillary GC analysis of partially purified germacrene C with injector temperature at 230°C. (*III*) Capillary GC analysis of the same sample with injector at 40°C. The numbered peaks in *II* and *III* correspond to δ-elemene (peak 1), germacrene C (peak 2), and germacrene B (peak 3).

and Hirose (24). The sesquiterpene fraction of VFNT Cherry leaves (3.2% of total volatile olefins) contained germacrene C (66%), germacrene A (7%), guaia-6,9-diene (7%), germacrene B (6%), β-caryophyllene (6%), germacrene D (4%), α-humulene (4%), and β-elemene (1%). Azulane (guaia-6,9-diene) and germacrane skeletal types have not been reported previously in tomatoes. By contrast, the sesquiterpene fraction of leaves from the commercial tomato variety Better Boy (3.3% of total volatile olefins) contained mostly β-caryophyllene (71%), germacrene C (15%), and α-humulene (9%). The sesquiterpene composition of Better Boy agrees with previous reports on the content of *L. esculentum* leaf oil (6, 7), except that no δ-elemene was found by on-column injection; germacrene C has not been previously observed in *Lycopersicon*.

To confirm identity, a sample of germacrene C (>96% purity by GC) from VFNT Cherry leaf oil was prepared for NMR analysis. Spectra of the putative germacrene C was consistent with the germacrene C structure. The following chemical shifts (δ) were recorded for ^{13}C NMR (see Fig. 5 for farnesyl numbering system): C1, 122.58; C2, 130.41; C3, 127.46; C4, 40.48; C5, 28.27; C6, 125.72; C7, 141.49; C8, 40.60; C9, 32.22; C10, 145.69; C11, 37.26; C12, 22.45; C13, 22.81; C14, 21.09; C15, 16.97. Protons on these carbons produced the following δ values and multiplets: C1, 6.349, bd, $J_{1,11} = 1.0$, $J_{1,9} = 1.6$; C2, 5.326, dt, $J_{1,2} = 9.34$, $J_{2,15} = 0.9$; C4^{ax}, 1.719; C4^{eq}, 2.126; C5^{ax}, 1.954; C5^{eq}, 2.083; C6, 4.880; C8^{ax}, 1.868; C8^{eq}, 2.517; C9^{ax}, 2.355; C9^{eq}, 1.986; C11, 2.288; C12, 1.039; C13, 1.056; C14, 1.176, dd, $J_{14,8eq} = 0.5$; C15, 1.547.

Germacrene B was identified by GC-MS comparison with the authentic standard. Although retention times of the standard on two different columns (Alltech AT-1000 and Hewlett Packard 5MS) exactly matched those of the putative germacrene B, the mass spectrum matched γ-elemene. This apparent spectral match was noted by earlier workers (25); however, the retention time of γ-elemene is much shorter than that of germacrene B. γ-Elemene has been reported from *Lycopersicon* (4, 6, 7) but was not found in either Better Boy or VFNT Cherry. Although no germacrene B was found in Better Boy, consistent with published analyses, it was found in VFNT Cherry.

The monoterpene composition of VFNT Cherry leaves was quantitatively and qualitatively very similar to that of Better Boy leaves. Leaves of both cultivars contained monoterpenes as the principal volatile olefins (91.4%), with β-phellandrene (52% of total) followed by 2-carene (16%) and limonene (10%) as the most abundant components (Fig. 1). The monoterpene content of *L. esculentum* has been examined in some detail by other workers (6, 7).

Sesquiterpene Synthase Activity in Leaf Extracts. To examine the origin of the germacrenes, a soluble protein extract was prepared from young VFNT Cherry leaves and assayed for sesquiterpene synthases with [1-^{3}H]FDP as substrate. Radio-GC analysis (250°C injector) of the sesquiterpene olefins generated (~3% conversion of substrate) revealed the presence of tritium-labeled olefins coincident with δ-elemene (33%), caryophyllene (22%), and humulene (19%), with at least two other sesquiterpenes (total of 26%) (Fig. 2). None of these cyclic olefins was produced in boiled controls, confirming their enzymatic origin. The formation of these products was dependent on the presence of the divalent metal ion cofactors, consistent with the established behavior of all known sesquiterpene synthases (26). No oxygenated sesquiterpenoids were produced from [1-^{3}H]FDP, other than farnesol and nerolidol (minor), which were derived via endogenous phosphatases and/or nonenzymatic solvolysis. Caryophyllene and humulene oxides have been reported in tomato leaves (6), but these metabolites are likely formed by subsequent oxygenation of the corresponding olefins. Thus, the sesquiterpene synthases of *Lycopersicon* yield olefinic products consistent with the volatile oil content of VFNT Cherry leaves.

Terpenoid Synthase cDNA Isolation and Characterization. By using the VFNT Cherry leaf library cDNA as template, the two degenerate PCR primers amplified a gene fragment of the expected size. This 767-bp fragment was shown to resemble TEAS (74% identity) and vetispiradiene synthase (78% identity) (17, 27), two sesquiterpene cyclases from related solanaceous plants. Southern blot hybridization of the PCR product to multiply-digested VFNT Cherry and spearmint (control) DNA revealed the probe to hybridize to three or four bands of tomato DNA, suggesting a small gene family; the probe did not recognize the spearmint DNA control (data not shown).

From high-stringency screening of the tomato cDNA library, 31 positive clones were isolated. Restriction mapping indicated all clones to be representatives of the same gene family, and 16 of these phagemids containing inserts of more than 1.8 kbp were used to transform *E. coli*. Nine clones (pLE7.1, 11.3, 12.1, 14.1, 14.2, 15.4, 16.3, 17.5, and 20.3) expressed sesquiterpene synthase activity capable of converting [1-^{3}H]FDP to radiolabeled olefin(s); clones pLE11.3 and pLE14.2 were most active, followed by clone pLE20.3. Radio-GC analysis of the products of all active clones revealed the same pattern of sesquiterpene olefins. Detailed GC-MS analysis of the products generated by clones pLE11.3, 14.2, and 20.3 confirmed the identical distribution pattern and permitted the identification of germacrene C (64%), germacrene A (18%), germacrene B (11%), and germacrene D (7%) (Fig. 3), all products that were identified in the volatile fraction of VFNT Cherry leaves or generated in the cell-free assay. However, because of the thermal decomposition of germacrene C while on column (note the rising baseline preceding germacrene C elution in Fig. 3), the amount of germacrene C actually produced may be twice that indicated. Control experiments established that sesquiterpene formation from FDP in extracts of transformed *E. coli* was insert-dependent and required divalent cation, substrate, and functional enzyme (boiled extracts were inactive). The only oxygenated sesquiterpenes generated in transformed *E. coli* preparations were farnesol and nerolidol, derived via endogenous phosphatases and/or nonenzymatic solvolysis; the formation of both was independent of the cDNA insert. Although caryophyllene and humulene are present in the volatile complex of VFNT Cherry leaves and are produced from FDP by extracts of this tissue, neither sesquiterpene was produced in detectable amounts by the recombinant enzyme. Therefore, these products must arise from a different synthase(s).

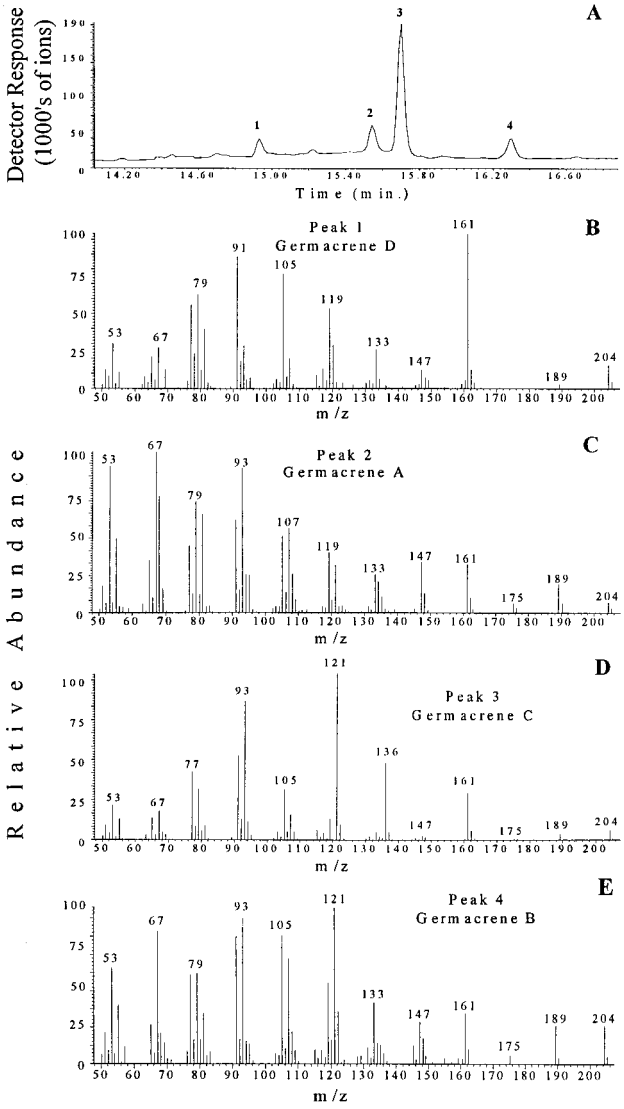

FIG. 3. GC-MS analysis of the major products generated from farnesyl diphosphate by the recombinant germacrene C synthase. (*A*) Total ion chromatogram; note the rising baseline preceding germacrene C (peak 3) due to thermal decomposition to δ-elemene while on column. (*B–E*) Mass spectra of the sesquiterpene products generated by germacrene C synthase.

None of the clones expressed detectable diterpene synthase activity with GGDP as substrate. Only one germacrene C synthase clone, pLE11.3, yielded a protein capable of converting [1-^{3}H]GDP to a monoterpene product, identified as limonene by radio-GC analysis (data not shown). The germacrene synthase activity expressed from clone pLE11.3 exceeded the limonene synthase activity by a factor of 10, when measured in the same enzyme preparation at saturating levels of the corresponding prenyl diphosphate substrate and Mg^{2+} cofactor. Sequence analysis (see below) provides a rationale for this observation of bifunctional cyclization activity.

Sequence Analysis. Clones pLE11.3, pLE14.2, and pLE20.3 were selected for complete nucleotide sequence analysis based on functional expression and comparatively large insert size. These cDNAs are 1,948, 2,022, and 1,878 bp long, respectively, and contain single ORFs of 1,622, 1,647, and 1,647 bp,

```
  1 MAASSADKCR PLANFHPSVW GYHFLSYTHE ITNQEKVEVD EYKETIRKML
 51 VETCDNSTQK LVLIDAMQRL GVAYHFDNEI ETSIQNIFDA SSKQNDNDNN
101 LYVVSLRFRL VRQQGHYMSS DVFKQFTNQD GKFKETLTND VQGLLSLYEA
151 SHLRVRNEEI LEEALTFTTT HLESIVSNLS NNNNSLKVEV GEALTQPIRM
201 TLPRMGARKY ISIYENNDAH HHLLLKFAKL DFNMLQKFHQ RELSDLTRWW
251 KDLDFANKYP YARDRLVECY FWILGVYFEP KYSRARKMMT KVLNLTSIID
301 DTFDAYATFD ELVTFNDAIQ RWDANAIDSI QPYMRPAYQA LLDIYSEMEQ
351 VLSKEGKLDR VYYAKNEMKK LVRAYFKETQ WLNDCDHIPK YEEQVENAIV
401 SAGYMMISTT CLVGIEEFIS HETFEWLMNE SVIVRASALI ARAMNDIVGH
451 EDEQERGHVA SLIECYMKDY GASKQETYIK FLKEVTNAWK DINKQFERPT
501 EVPMFVLERV LNLTRVADTL YKEKDTYTNA KGKLKNMINS ILIESVKI*
```

FIG. 4. Deduced amino acid sequence of germacrene C synthase encoded by the cDNA insert of pLE20.3 (accession no. AF035630). Residues in boldface type are the conserved aspartate-rich motif involved in binding the divalent metal ion-chelated substrate. Underlined residues indicate the region of the cDNA to which the probe was directed. Double underlined residues are changed to S, ST, and P, respectively, in the cDNA insert pLE14.2 (accession no. AF035631).

respectively. Clones pLE11.3 and pLE14.2 contain the longest 3′ untranslated regions but shorter 5′ untranslated regions than does pLE20.3. Clones pLE14.2 and pLE20.3 are in different reading frames with respect to the vector β-galactosidase start site; however, each contains a stop codon (−12 and −29 nucleotides from the initial cDNA ATG, respectively) in the 5′ untranslated region that is in-frame with the β-galactosidase start site. Both pLE14.2 and pLE20.3 inserts contain at the start site a purine XXAUGG motif, which provides high recognition as a secondary start site in polycistronic messages. Thus, it is likely that the germacrene synthases derived from these clones are translated as proteins free of vector-derived peptide and that their apparent difference in enzyme activity (3-fold) is probably due to the extra distance the ribosome must traverse to locate the second start site in the pLE20.3 message compared with that of pLE14.2, resulting in lower translation efficiency. Clones pLE20.3 and pLE14.2 bear several single base differences that result in four amino acid changes near the carboxyl terminus (Fig. 4). Because VFNT Cherry is a F_5 selection and, therefore, not entirely homozygous, it is unclear whether pLE20.3 and pLE14.2 represent different alleles or different transcripts derived from two germacrene C synthase loci.

The only clone capable of expressing a monoterpene (limonene) synthase activity (pLE11.3) is truncated by 25 nucleotides into the 5′ end of the ORF, thereby deleting the starting methionine and the following 7 residues and also changing the 9th residue. The cDNA sequence for the remaining ORF is identical to that of clone pLE20.3, which does not express detectable monoterpene synthase activity. The first residue (Arg) encoded entirely by the pLE11.3 cDNA insert is also the first conserved residue among the known plant terpenoid synthases, and this codon is in-frame with the vector β-galactosidase start site. Thus, the expressed product derived from pLE11.3 is probably a fusion protein in which the amino-terminally truncated germacrene synthase is fused to a 42-residue vector-derived peptide. The ability of clone pLE11.3 to express limonene synthase activity with GDP as substrate is apparently the result of this cloning artifact and so is not relevant to limonene production *in vivo*. The use of GDP as an alternate substrate by a sesquiterpene synthase has been reported for three other plant sesquiterpene synthases (11, 28), which are apparently translated in the native form; however, the present observation indicates that substrate use can be altered by modification of the N terminus. Back and Chappell (29) have conducted domain-swapping experiments in which domains from TEAS were switched with domains from vetispiradiene synthase. They found that chimeric enzymes with new product compositions could be produced; however, they did not test these hybrid enzymes for acceptance of alternate substrates (either GDP or GGDP).

Comparison of the deduced amino acid sequence of clone pLE20.3 (Fig. 4) with those of other plant sesquiterpene synthases reveals a significant degree of similarity. The sequence of tomato germacrene C synthase is most similar to δ-cadinene synthase from cotton (Malvaceae) in showing 68% similarity and 50% identity (30). The defined sesquiterpene synthases from other solanaceous plants, vetispiradiene synthase from *Hyoscyamus muticus* (27) and TEAS (17), each exhibit 64% similarity and 45% identity to germacrene C synthase at the amino acid level. Germacrene C synthase is least like (*E*)-β-farnesene synthase from peppermint (45% similarity and 34% identity) (11) and δ-selinene and γ-humulene synthases from grand fir (41% similarity and 29% identity vs. 40% similarity and 27% identity, respectively) (28). It is interesting that, although the reaction mechanisms of TEAS, vetispiradiene synthase, and δ-selinene synthase have been postulated to proceed through a germacrene intermediate (28, 29), these are not more similar to germacrene C synthase in sequence than are other known sesquiterpene synthases (28, 30) that produce structurally unrelated products.

As is typical of the sesquiterpene synthases of plant origin (11, 17, 27, 28, 30), tomato germacrene C synthase appears to lack an amino-terminal organelle-targeting sequence. Therefore, the enzyme must be directed to the cytoplasm, which is the site of sesquiterpene biosynthesis; monoterpenes and diterpenes are synthesized in plastids (16). Translation of the germacrene synthase cDNA yields a deduced protein with a molecular weight of 64,114 and a pI of 5.82. The aspartate-rich motif (DDXXD) found in most prenyltransferases and terpenoid cyclases, and thought to play a role in substrate binding (31), is also present in germacrene C synthase (Fig. 4).

One of the degenerate sequences of the forward PCR primer used in this work is found at the predicted location within the germacrene synthase cDNA; however, the region that is conserved in other known terpenoid cyclases, and to which the reverse PCR primer was designed, is not conserved in germacrene C synthase (Fig. 4). Thus, the template that generated the hybridization probe was not germacrene C synthase (the probe exhibits only 58% identity with germacrene C synthase at the nucleotide level) but probably another sesquiterpene cyclase that has, thus far, eluded capture.

Cyclization Mechanism. Germacrene C synthase from tomato, like many other terpenoid cyclases of plant origin (11, 28), is capable of producing multiple reaction products, probably as a consequence of the highly reactive carbocationic intermediates that are generated from FDP (26). The electrophilic cyclization reaction is postulated to proceed by the initial ionization of the diphosphate ester (Fig. 5). Capture of the diphosphate anion at C3 of the resulting carbocation may occur to form nerolidyl diphosphate that may simply reionize to the original transoid carbocation to permit C1,C10 closure to the germacryl skeleton. Two alternatives for deprotonation of the C_{10} macrocyclic carbocation yield germacrenes A or B. A 1,3-hydride shift in the macrocycle and deprotonation by two alternate routes produces germacrenes D or C, whereas a 1,2-hydride shift with alternative deprotonations can generate germacrenes C or B. Upon heating, germacrene C undergoes Cope rearrangement to δ-elemene (see Figs. 2 and 5). Similarly, β-elemene is the Cope rearrangement product of germacrene A. Although β-elemene was not observed as a biosynthetic product of germacrene C synthase, this olefin was detected in VFNT Cherry leaf volatiles. γ-Elemene (the Cope rearrangement product of germacrene B) was not detected as a biosynthetic product or as a metabolite of VFNT Cherry leaves, but it has been reported in tomato by other workers (7, 32).

Recently, the crystal structure of TEAS has been solved (33) and detailed sequence comparison with germacrene C synthase

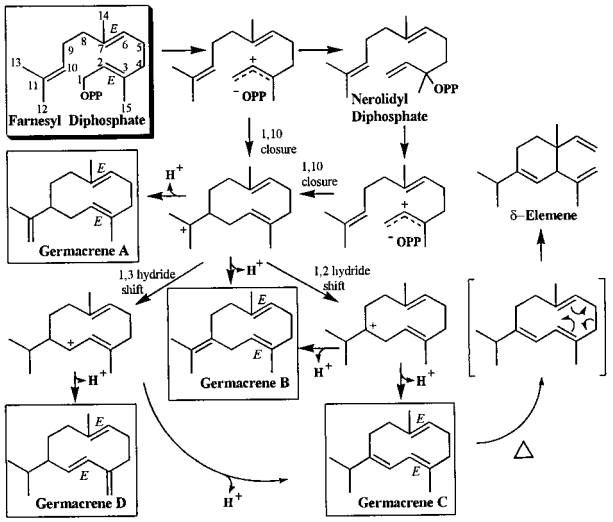

FIG. 5. Proposed mechanism for the formation of germacrenes from FDP. OPP denotes the diphosphate moiety. δ-Elemene is not a direct enzyme product but is produced by thermal rearrangement of germacrene C.

may reveal the structural basis for many of the catalytic steps in germacrene formation. In TEAS, upon binding and subsequent ionization of FDP, closure of an α-helical loop over the active site is facilitated by interactions between Trp^{273} and Tyr^{527} to form a pocket that shields the resulting carbocation intermediate from quenching by solvent water. These two residues are conserved as Trp^{272} and Tyr^{527} in germacrene C synthase and, if a common synthase folding is assumed (33), they likely play an identical role. Loop closure also positions residues Arg^{264} and Arg^{441} of the aristolochene synthase near C1 of the substrate and this, along with the coordination of substrate-bound metal ions by the DDXXD motif, helps direct the metal ion-chelated diphosphate anion away from the newly formed carbocation. These basic residues are conserved as Arg^{263} and Arg^{442} in the germacrene synthase. Several of the backbone carbonyls and Thr^{403} direct attack of C10 on the C1 cationic center to form a germacryl intermediate in the aristolochene synthase. Although there is no direct homolog of Thr^{403} in the germacrene synthase (only Ser^{401}), positioning of Tyr^{527} (in both synthases) near C11 stabilizes the macrocyclic carbocation formed and allows Asp^{525} (in both synthases) to deprotonate at C13 to yield germacrene A. In the germacrene synthase, several lysine residues upstream of Asp^{525} may discourage proton elimination and permit the hydride shift(s) needed to generate germacrenes C and D. In TEAS, reprotonation of the germacrene intermediate at C6 to yield the bicyclic eudesmane nucleus is facilitated by a catalytic triad of Asp^{444}, Tyr^{520}, and Asp^{525}. The protonation cycle is initiated by Asp^{444}, and the cationic center formed at C3 after ring closure is stabilized by the hydroxyl of Thr^{403}. No reprotonation of germacrene can occur in the germacrene synthase because the initiating aspartate has been altered to Asn^{445} and the hydroxyl of Ser^{401} may be too distant to stabilize the developing positive charge on C3; thus, the macrocyclic germacrenes are released as terminal products. With the development of an efficient functional expression system for germacrene C synthase, these mechanistic inferences can be tested by mutagenesis and the structure of the enzyme examined directly. The isolation of the germacrene C synthase cDNA also provides the means for genetic engineering of sesquiterpenoid-based plant defenses and refines the development of probes for identification of other terpenoid synthases in tomato.

S.M.C. and J.C. share equal credit in this investigation. We thank Gerald Still for support and encouragement, John Steffens for the gift of the cDNA library, Gerhard Munske for nucleotide sequencing, Hiroko Ichii for technical assistance, Ron Buttery for helpful discussion, Bob Adams and Larry Cool for gifts of sesquiterpene standards, Jaraslov Zajicek for obtaining the NMR spectra, Joyce Tamura-Brown for typing the manuscript, and Thom Koehler for raising the plants. This work was supported in part by U.S. Department of Agriculture National Research Institute (NRI) Grant 93–37301-9504 (to S.M.C.), U.S. Department of Agriculture Agricultural Research Center Contingency Funds for Whitefly Research (to P.G.L.), National Institutes of Health Grant GM31354 (to R.C.), and Project 0268 (to R.C.) from the Agricultural Research Center, Washington State University.

1. Buttery, R. G. & Ling, L. C. (1993) *Am. Chem. Soc. Symp. Ser.* **52,** 22–34.
2. Buttery, R. G., Takeoka, G., Teranishi, R. & Ling, L. C. (1990) *J. Agric. Food Chem.* **38,** 2050–2053.
3. Carter, C. D., Gianfagna, T. J. & Sacalis, J. N. (1989) *J. Agric. Food Chem.* **37,** 1425–1428.
4. Lin, S. Y. H., Trumble, J. T. & Kumamoto, J. (1987) *J. Chem. Ecol.* **13,** 837–850.
5. Carter, C. D., Sacalis, J. N. & Gianfagna, T. J. (1989) *J. Agric. Food Chem.* **37,** 206–210.
6. Buttery, R. G., Ling, L. C. & Light, D. L. (1987) *J. Agric. Food Chem.* **35,** 1039–1042.
7. Lundgren, L., Norelius, G. & Stenhagen, G. (1985) *Nord. J. Bot.* **5,** 315–320.
8. Marlatt, C., Ho, C. T. & Chien, M. (1992) *J. Agric. Food Chem.* **40,** 249–252.
9. Isoe, S., Katsumura, S. & Sakan, T. (1973) *Helv. Chim. Acta* **56,** 1514–1516.
10. Jones, R. A., Smith, P. G., Millet, A. H. & Kimble, K. A. (1980) *Hortic. Sci.* **15,** 98.
11. Crock, J., Wildung, M. & Croteau, R. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 12833–12838.
12. Croteau, R., Alonso, W. R., Koepp, A. E. & Johnson, M. A. (1994) *Arch. Biochem. Biophys.* **309,** 184–192.
13. Dixit, V. M., Laskovics, F. M., Noall, W. I. & Poulter, C. D. (1981) *J. Org. Chem.* **46,** 1967–1969.
14. LaFever, R. E., Stofer-Vogel, B. & Croteau, R. (1994) *Arch. Biochem. Biophys.* **313,** 139–149.
15. Maarse, H. & Kepner, R. E. (1970) *J. Agric. Food Chem.* **18,** 1095–1011.
16. Colby, S. M., Alonso, W. R., Katahira, E. J., McGarvey, D. J. & Croteau, R. (1993) *J. Biol. Chem.* **268,** 23016–23024.
17. Facchini, P. J. & Chappell, J. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 11088–11092.
18. Mau, C. J. D. & West, C. A. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 8497–8501.
19. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
20. States, D. J., Heberkorn, R. A. & Ruben, D. J. (1982) *J. Magn. Reson.* **48,** 286–293.
21. Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R. & Wuthrich, K. (1983) *Biochem. Biophys. Res. Commun.* **117,** 479–485.
22. Bax, A., Griffey, R. H. & Hawkins, B. L. (1983) *J. Magn. Reson.* **55,** 501–505.
23. Bax, A. & Summers, M. F. (1986) *J. Am. Chem. Soc.* **108,** 2093–2094.
24. Morikawa, K. & Hirose, Y. (1969) *Tetrahedron Lett.* **22,** 1799–1801.
25. Clark, B. C., Chamblee, T. S. & Iacobucci, G. A. (1987) *J. Agric. Food Chem.* **35,** 514–518.
26. Cane, D. E. (1990) *Chem. Rev.* **90,** 1089–1103.
27. Back, K. & Chappell, J. (1995) *J. Biol. Chem.* **270,** 7375–7381.
28. Steele, C. L., Crock, J., Bohlmann, J. & Croteau, R. (1998) *J. Biol. Chem.* **273,** 2078–2089.
29. Back, K. & Chappell, J. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 6841–6845.
30. Chen, X. Y., Chen, Y., Heinstein, P. & Davisson, V. J. (1996) *Arch. Biochem. Biophys.* **324,** 255–266.
31. Marrero, O. F., Poulter, C. D. & Edwards, P. A. (1992) *J. Biol. Chem.* **267,** 21873–21878.
32. Weston, P. A., Johnson, D. A., Burton, D. A. & Snyder, J. C. (1989) *J. Am. Soc. Hortic. Sci.* **114,** 492–498.
33. Starks, C. M., Back, K., Chappell, J. & Noel, J. P. (1997) *Science* **277,** 1815–1820.