Sesquiterpene Synthases from Grand Fir (Abies grandis)

Comparision of Constitutive and Wound-Induced Activities, and cDNA Isolation, Characterization, and Bacterial Expression of Δ-Selinene Synthase and η-Humulene Synthase

Christopher L. Steele‡§, John Crock‡, Jörg Bohlmann‡, and Rodney Croteau

From the Institute of Biological Chemistry, Program in Plant Physiology, and Department of Biochemistry and Biophysics, Washington State University, Pullman, Washington 99164-6340

Grand fir (Abies grandis) has been developed as a model system for the study of oleoresin production in response to stem wounding and insect attack. The terpentine fraction of the oleoresin was shown to contain at least 38 sesquiterpenes that represent 12.5% of the terpentine, with the monoterpene sesquiterpenes comprising the remainder. Assays of cell-free extracts from grand fir stem with farnesyldiphosphate as substrate indicated that the constitutive sesquiterpene synthases produced the same sesquiterpenes found in the oleoresin and that, in response to wounding, only two new products were synthesized, Δ-cadinene and (E)-α-bisabolene. A similarity based cloning strategy yielded two new cDNA species from a stem cDNA library that, when expressed in Escherichia coli and the gene products subsequently assayed, yielded a remarkable number of sesquiterpene products. The encoded enzymes have been named Δ-selinene synthase and η-humulene synthase based on the principal products formed; however, each enzyme synthesizes three major products and produces 34 and 52 total sesquiterpenes, respectively, thereby accounting for many of the sesquiterpenes of the oleoresin. The deduced amino acid sequence of the Δ-selinene synthase cDNA open reading frame encodes a protein of 581 residues (at 67.6 kDa), whereas that of the η-humulene synthase cDNA encodes a protein of 593 residues (at 67.9 kDa). The two amino acid sequences are 83% similar and 65% identical to each other and range in similarity from 65 to 67% and in identity from 43 to 48% when compared with the known sequences of monoterpene and diterpene synthases from grand fir. Although the two sesquiterpene synthases from this gymnosperm do not very closely resemble terpene synthases from angiosperm species (52–56% similarity and 26–30% identity), there are clustered regions of significant apparent homology between the enzymes of these two plant classes. The multi-step, multi-product reactions catalyzed by the sesquiterpene synthases from grand fir are among the most complex of any terpenoid cyclase thus far described.

Conifer oleoresin is a mixture of turpentine (monoterpenes (C_{10}) and sesquiterpenes (C_{15}) olefins) and resin (diterpenes (C_{20}) resin acids) that functions in insect defense and in wound sealing (1, 2). Grand fir (Abies grandis) has been developed as a model system for the study of both constitutive and wound-induced oleoresin formation (oleoresinosis). The composition of the monoterpenic olefin and the diterpenic resin acid fractions of grand fir oleoresin has been defined (3), and the induced biosynthesis of these natural products upon stem wounding has been described in detail (2, 4–6). The time course of induction of the monoterpene synthases involved in turpentine formation has been analyzed by immunoblotting techniques, and the process of induced oleoresinosis was thus shown to involve de novo synthesis of these enzymes (5). The cDNA sequence of a diterpene cyclase from grand fir (abietadiene synthase involved in resin acid biosynthesis (7)) has been reported (8), and several cDNA clones encoding monoterpene synthases from this conifer species have recently become available (9).

The sesquiterpene synthases of conifer turpentine have received relatively little experimental attention because they constitute less than 10% of the oleoresin. However, sesquiterpenoid phytoalexins are well known in angiosperm species (10), suggesting that the sesquiterpenes of conifer oleoresin may play a similar role in antibiosis and thus be of greater significance than their lower concentration in resin might otherwise indicate. Sesquiterpenes are produced in the cytosol/endoplasmic reticulum compartment, whereas monoterpenic and diterpenic biosynthesis are compartmentalized in plastids (11), which raises the additional issue of coordinate regulation of oleoresin terpene biosynthesis at several cellular sites. Only a single sesquiterpene synthase, (E)-β-farnesene synthase, from a gymnosperm source, maritime pine (Pinus pinaster), has been reported (12), whereas several sesquiterpene synthases from angiosperms have been described (13–15), and a number of genes encoding sesquiterpene synthases involved in phytoalexin biosynthesis in angiosperms have been isolated (16–18).

To examine the possible role of sesquiterpenes in conifer defense against stem boring insects and their associated fungal pathogens, it is first necessary to examine in greater detail the origin of these oleoresin constituents. In this paper, we describe the sesquiterpene composition of grand fir oleoresin, and the cell-free biosynthesis of these terpenoids from the common isoprenoid intermediate farnesyldiphosphate in extracts from wounded (induced) and nonwounded control (constitutive) sapling stems. In addition, we report on the use of a general cloning strategy (9, 19) in the isolation and functional expres-
Cloning of Conifer Sesquiterpene Synthases

EXPERIMENTAL PROCEDURES

Plant Materials, Substrates, and Reagents—Two-year-old grand fir (Abies grandis Lindl.) saplings were purchased from the Forestry Research Nursery, University of Idaho, Moscow, ID. Saplings were grown in standard potting mix (Sals Inc., Puyallup, WA) with 16 h of light (200–300 microeinsteins/m²s) and a 26 °C day/15 °C night temperature cycle and were fertilized (15:30:15 (N:P:K)) weekly and watered daily. Three grams of needles were used as the enzyme source for determination of constitutive terpene hydrocarbon composition. These stems were chopped into 5–7-cm segments, frozen in liquid N2, and subsequently eluted with 3 x 1 ml of ether to collect any oxygenated products, and an aliquot of each fraction was taken for liquid scintillation counting to determine conversion rate. These monoterpenes and diterpene synthase activity assays were similarly performed as described in detail elsewhere (7, 9, 22).

For preparative incubations, the assay was scaled to 5 ml containing the total protein extracted from 3 x 1 ml of ether to collect any oxygenated products, and an aliquot of each fraction was taken for liquid scintillation counting to determine conversion rate. These monoterpenes and diterpene synthase activity assays were similarly performed as described in detail elsewhere (7, 9, 22).

Liquid scintillation counting was performed in 10 ml of toluene: ethanol (7:3, v:v) containing 0.4% (w:v) Omnifluor (NEL Life Science Products) at a H counting efficiency of 43%. Protein concentrations were determined by the method of Bradford (25) using the Bio-Rad reagent and bovine serum albumin as standard.

cDNA Isolation, 5'-RACE, and Expression of Sesquiterpene Synthases—Fundamentally, the cDNA library had been described (8), and the details of hybridization probe generation and library screening are reported elsewhere (9). In summary, hybridization probes for terpenoid synthases were generated by PCR using degenerate oligonucleotide primers designed from conserved amino acid sequences (designated in boldface in Fig. 7) of several monoterpene, sesquiterpene, and diterpene synthases from angiosperm species (19). DNA from a λ phage cDNA library, constructed from mRNA isolated from wounded grand fir sapling stems (8), was purified and used as template for PCR reactions (26). Four unique, 110-bp fragments were amplified, cloned, and shown to be cycle-like in sequence, and they were designated probes 1, 2, 4, and 5. Upon screening of the cDNA library, probes 4 and 5 hybridized, respectively, to two unique cDNA species designated ag4.30 and ag5.9; the location of each probe is doubly underlined in the sequences illustrated in Fig. 7.

Since neither of the cDNA isolates encoded a starting methionine, 5'-RACE was carried out using the Marathon cDNA amplification kit (CLONTECH) following by the manufacturer's protocol with slight modification (see below). Total RNA was extracted from 60 saplings (2-year-old; 6 or 8 days after wounding) by scale-up of a published procedure (27). Poly(A)¹ RNA was isolated using Oligotex beads and the spin column procedure described by Qiagen. To prevent RNA secondary structure features from obstructing full-length cDNA synthesis, three different cDNA synthesis reactions were performed by first denaturing the RNA at 42 or at 50 °C or by treatment with methylmercury hydroxide (28). Twice as much mRNA (2 μg) as recommended in the protocol was used, and PCR amplification was performed as described, indicating that a low amount of poly(A)¹ may be tolerated. The respective cDNA-specific reverse PCR primers were 5'-TCGCGAACCT-TGAGAAGTGGCTGCAG-3' for ag4.30 and 5'-GCTATCGATGCGGTTGTTTCAGC-3' for ag5.9. The resulting PCR products were cloned into the pT7Blue-vector (Novagen) following standard procedures, and they were partially sequenced to reveal in each case a putative starting methionine codon, thus indicating that successful 5'-RACE syntheses

1 The abbreviations used are: GLC, gas-liquid chromatography; MS, mass spectrometry (spectra); aa, amino acid(s); bp, base pair(s); PCR, polymerase chain reaction; RACE, rapid amplification of cDNA ends.
had occurred. Full-length representatives were generated by designing 5’- and 3’-specific PCR primers for each cDNA for subsequent high fidelity amplification. The 5’-specific primers were designed with a BamHI restriction endonuclease site immediately upstream of the starting methionine codon for each cDNA (5’-GGAGGATCCATGCGTGTAATTCTCGT-3’ for ag4.30 and 5’-TGGTACCATGCGTCGGCTTTGTGC-CTGTATC-3’ for ag5.9). The 3’-specific primers were designed to encompass the stop codon; the ag4.30 primer included an XhoI site, whereas the ag5.9 primer included an EcoRI site (5’-AAAAGTCTCGAG-ATATATTATTGCC-3’ for ag4.30 and 5’-TATGATATTCTCAATAG-TGCCGGGGAC-3’ for ag5.9) to facilitate ligation into the pGEX-4T-1 expression vector (Pharmacia Biotech Inc.). PCR reactions were performed at 94 °C for 1 min, 50 °C for 1 min, and 72 °C for 6 min for 30 cycles followed by a 5-min final extension period at 72 °C, using Pfu polymerase and the buffer described by the manufacturer (Stratagene). The resulting DNA fragments were sequenced by cloned standard methods, first into pBluescript (SK+) (Stratagene) and then into pGEX (Pharmacia) vectors designated as pGAG4 and pGAG5. For further subcloning of cDNAs into the pSBETa vector for high level expression (29), inserts of pGAG4 and pGAG5 were amplified by PCR (Stratagene Pfu polymerase as above) using primer combinations 4-NdeI (5’-CTGGGTCCCGCGATGGCTGAGT-3’) and 4-BamHI (5’-CTACAACCAAGGAGGCTTATTTTCACTGC-3’) with pGAG4, and 5-NdeI (5’-CTGGTCTCCGCTCGGACATGCTGCTAG-3’) and 5-BamHI (5’-GTCGGATCCATGCTCCTCAAATAGGCACGG-3’) with pGAG5. The PCR products were digested with the above indicated restriction enzymes, purified by ultrafiltration, and then ligated into NdeI/BamHI-digested pSBETa to yield plasmid pSB4AG and pSB5AG, respectively.

The original isolates, ag4.30 and ag5.9, and their full-length cDNA representatives, inserts of ag4 in pGAG4 and pSBAG4, and of ag5 in pGAG5 and pSBAG5, were entirely sequenced on both strands via primer walking using the dye-terminator-cycle sequencing method (Applied Biosystems) on a ABI 373 DNA Sequencer Stretch instrument at the Washington State University Laboratory for Biotechnology and Bioanalysis. Sequence analysis was done using programs from the Wisconsin Package Version 9.0 of the Genetics Computer Group (30).

Both putative sesquiterpene cyclase cDNAs were expressed in bacterial strains Escherichia coli XL1-Blue/pGAG4, E. coli XL1-Blue/pGAG5, E. coli BL21(DE3)/pSBAG4, and E. coli BL21(DE3)/pSBAG5. Bacteria were grown to $A_{600}$ = 0.5 at 37 °C in 5 ml or 100 ml of LB medium supplemented with 100 μg of ampicillin/ml or 30 μg of kanamicin/ml as determined by the vector. 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 (2000 × g, 10 min) and resuspended in either 1 or 5 ml of sesquiterpene synthase assay buffer. Cells were disrupted by sonication (Braun-Sonic 2000 with microprobe at maximum power for 15 s at 0–4 °C), and the homogenates were cleared by centrifugation (18,000 × g, 10 min).

Preparative assays were employed to generate product for GLC-MS analysis as before, with quantification of composition via the total ion current chromatogram.

RESULTS

Oleoresin Sesquiterpenes—Grand fir has been utilized as a model system for the study of induced oleoresin production in conifers in response to wounding and insect attack (19). The monoterpene and diterpenoid components of the oleoresin have been defined, and several of the responsible monoterpene and diterpenoid synthases have been purified and characterized and the corresponding cDNAs isolated to provide tools for examining the regulation of this defense response (6–9, 21). The third component of the oleoresin, the sesquiterpenes, has not been defined. The results of our investigations into the regulation of sesquiterpene synthesis in Grand fir revealed that the product derived from farnesyl diphosphate, the universal precursor of sesquiterpenoids (32, 33), was a labeled olefin fraction (12 × 2 ml assays yielded −1 nmol of product) that upon radio-GLC analysis (Fig. 2a) was shown to contain the same spectrum of sesquiterpenes found in the oleoresin. Enzyme extracts were similarly prepared from sapling stems 9 days after wounding and were assayed as before (2 × 2 ml assays yielded 1.8 nmol of product). Radio-GLC analysis of the olefin fraction revealed the presence of an apparently single component with retention time very similar to that of $\delta$-cadinenone (Fig. 2b). Partial purification of the extract from induced saplings to eliminate traces of endogenous oleoresin, followed by preparative-scale assay, provided sufficient material for capillary GLC-MS analysis. This higher resolution method revealed that the product derived from farnesyl diphosphate by the induced enzyme(s) consisted of two components that were identified as $\delta$-cadinenone and (E,E)-germacrene B. Differential loss of the $\delta$-cadinenone synthase activity during storage (data not shown) suggested that $\delta$-cadinenone and (E,E)-germacrene were the products of two different enzymes. Boiled controls, and control reactions without farnesyl diphosphate, confirmed that both the constitutive and inducible sesquiterpene synthase activities observed were enzymatic and substrate-dependent. The $K_m$ value for [1-3H]farnesyl diphosphate with the partially purified inducible sesquiterpene synthase was determined to be about 0.4 μM. It is of interest, and of probable physiological significance, that the constitutive sesquiterpene synthase activities differ from the wound-induced enzyme activities in product composition. A similar phenomenon has been previously observed with the constitutive and wound-inducible monoterpene and diterpene synthases of this tissue (6, 21).

2 C. L. Steele, S. Katoh, J. Bohmann, and R. Croteau, manuscript submitted for publication.

![Fig. 1. Structures of the six major sesquiterpene olefins of grand fir turpentine. The percentage contribution to the total sesquiterpene fraction is indicated in parentheses. The stereochemistry illustrated is relative.](image-url)
cDNA Isolation and Expression—A similarity-based PCR cloning strategy for terpenoid synthases (9) yielded two different truncated cDNA species, ag4.30 and ag5.9, upon screening a grand fir stem cDNA library. The full-length forms, ag4 and ag5, were acquired by a 5′-RACE technique, and both of these, and the original cDNA isolates, were sequenced completely on both strands. Sequence similarity to other terpenoid synthases of plant origin (see below) and the apparent lack of an encoded translation within the 120 individual trees used in cDNA library construction.

Both ag4 and ag5 were expressed in E. coli using pGEX and pSBET (29) vectors; the latter encodes a tRNA for rare arginine codon usage in E. coli that is common in higher plants. Levels of expressed enzyme activity were 50 to 100 times higher in extracts from E. coli BL21(DE3)/pSBAG4 and E. coli BL21(DE3)/pSBAG5 as compared with extracts from E. coli XL1-Blue/pGAG4 and E. coli XL1-Blue/pGAG5. Therefore, for subsequent product identification, recombinant enzymes were produced with pSBET expression constructs in E. coli BL21(DE3). Large scale incubation with [1-3H]farnesyl diphosphate, followed by isolation of the derived olefins and GLC-MS analysis (Fig. 3), revealed the enzyme encoded by ag4 to produce mainly γ-selinene, for which this synthase is named, along with germacrene B and guai-6,9-diene as major products. In addition, 17 other sesquiterpene olefins were identified by GLC-MS, and another 14 products which displayed the characteristic sesquiterpene olefin mass spectral pattern (m/z 204 (P′), 189 (P′-CH₃), and 161 (P′-CH₂) were detected for a total of at least 34 different sesquiterpene products (Table I).

A cyclization scheme to account for the generation of this remarkable number of products has been formulated (Fig. 4) and begins with the ionization and subsequent isomerization of (E,E)-farnesyl diphosphate to the corresponding tertiary allylic isomer, nerolidyl diphosphate, a process known to occur in the biosynthesis of other sesquiterpenes (33). The route to the principal product, γ-selinene, is based upon the formation of other eudesmane sesquiterpenes (32, 33, 35) and is initiated by ionization of the transoid-conformer of nerolidyl diphosphate, with C-10 to C-1 (farnesyl backbone numbering) ring closure, to afford the germacrane skeleton from which deprotonation yields germacrenes A and B or (following hydride shift) germacrenes C and D. These germacrenes may be released by the synthase (they account for 28.3% of the olefins generated) or they may be reprotonated to allow for additional ring closures (at least 58% of the carbon flux involves reprotonation). Reprotonation at C-6 restricts carbon flow to the eudesmane-structural types, whereas protonation at C-3 directs flux to the azulene (guaiane)-structural types. After protonation, the fate of the intermediates generated is determined by the position of the terminal double bond formed during germacrane biosynthesis. Although only the products of C-3 protonation of germacrene A and C were found, it is likely that both germacrene B and D also undergo protonation at C-3 to yield products that are not yet identified (Table I). Each of the germacrenes yields a product identified by protonation at C-6 (Fig. 4). Closure of the transoid-nerolidyl cation to the C₁₁ macrocyclic cation provides for formation of both α-humulene and β-caryophyllene (Fig. 4). Similar closure of the cisoid-nerolidyl cation also allows formation of β-caryophyllene, the only identified olefin from this synthase that can be produced via either cisoid- or transoid-nerolidyl diphosphate.

Carbon flow through the cisoid-nerolidyl diphosphate pathway accounts for approximately 13% of the products identified (Table I). Besides closure to the C₁₁ macrocycle, C-6 to C-1 closure of this intermediate may also occur to form β-hisabolenol, in a reaction analogous to the formation of the monoterpenes limonene by limonene synthase (34); no other isomers of hisabolenol were detected. The C-10 to C-1 closure leads to an (Z,E)-intermediate that may deprotonate to form (Z,E)-germacrene B. Although this compound has not been reported from a natural source, it is analogous to the naturally occurring helminthogermacrene A or (Z,E)-germacrene A (36). (Z,E)-Germacrene B was tentatively identified as an enzyme product based on an exact mass spectral match to (E,E)-germacrene B and a similar GLC retention time (data not shown). Although the mass spectrum of γ-elemene is very similar to that of germacrene B, and the former is the Cope rearrangement product of germacrene B, the GLC retention time of γ-elemene is much earlier than that of germacrene B. Likewise, although bicyclergermacrene (the tricyclic, dimethylcycloperyl analog of germacrene B) has a mass spectrum similar to (E,E)-germacrene B, the GLC retention time did not match that of the putative
FIG. 3. Total ion chromatograms of the sesquiterpene products derived from farnesyl diphosphate by δ-selinene synthase (ag4) (A) and γ-humulene synthase (ag5) (H). In the column below each total ion chromatogram are the mass spectra and retention times for the numbered peaks and the spectra and retention times for the corresponding authentic standards.
Cloning of Conifer Sesquiterpene Synthases

Products are listed in order of their abundance and were identified by matching GLC retention time and mass spectrum to authentic standards. Compounds labeled as (tent.) were tentatively identified based on the mass spectrum alone.

### Table I

| Sesquiterpene (ag4) | %  | γ-Humulene synthase (ag5) | %  |
|---------------------|----|---------------------------|----|
| δ-Selinene          | 25.3 | γ-Humulene                | 28.6  |
| E,E-Germacrene B    | 17.4 | Sibirene                  | 15.1  |
| Guai-6,9-diene       | 9.7  | Longifolene               | 11.8  |
| Germacrene A        | 6.7  | β-Himachalene             | 7.2   |
| δ-Amorphene          | 6.4  | γ-Himachalene             | 5.8   |
| Unknown              | 4.7  | α-Himachalene             | 4.8   |
| Unknown              | 4.4  | β-Bisabolene              | 3.9   |
| Germacrene C         | 3.4  | α-Longipinene             | 3.4   |
| α-Amorphene          | 2.7  | Sativene                  | 3.1   |
| Unknown              | 2.6  | α-Ylangene                | 2.7   |
| α-Selinene           | 1.7  | β-Gurjunene               | 2.0   |
| β-Caryophyllene      | 1.5  | γ-Bisabolene              | 1.9   |
| δ-Cadinene           | 1.4  | β-Longipinene             | 1.5   |
| Unknown              | 1.4  | (E)-β-Farnesene           | 1.3   |
| Unknown              | 1.3  | (E)-α-Bisabolene          | 0.9   |
| (Z,E)-Germacrene B (tent.) | 1.3 | δ-Amorphene               | 0.5   |
| Seli-3,7(11)-diene   | 1.2  | α-Amorphene               | 0.5   |
| Germacrene D         | 0.8  | β-Ylangene                | 0.3   |
| α-Humulene           | 0.7  | 1,5,9-Trimethylcyclododeca-1,5,9-triene (tent. 2 isomers) | 0.2 |
| β-Bisabolene         | 0.7  | Longicyclene              | 0.2   |
| Sibirene             | 0.7  | β-Cubebene                | 0.1   |
| α-Guaene             | 0.2  | α-Copaene                 | 0.1   |
| α-Cadinene           | 0.2  | α-Cubebene                | 0.1   |
| α-Ylangene           | 0.1  | (Z,E)-Germacrene B (tent.) | 0.1 |
| α-Copaene            | 0.1  | Cyclosativene             | <0.1  |
| 9 unknowns, each <1% | 3.4  | 26 unknowns, each <1%     | 4.1   |

(Z,E)-germacene B. Given the mass spectrometric and GLC retention data, and the fact that the enzyme can deprotonate the (E,E)-intermediate to form (E,E)-germacene, it seems likely that (Z,E)-germacene is, in fact, a product of the δ-selinene synthase.

1,3-Hydride shift in the (Z,E)-germacryl cation, followed by C-6 to C-1 ring closure and subsequent deprotonation, provides α- and δ-cadinene and the diastereomers α- and δ-amorphene. The cadinyl cation may further cyclize to the tricyclic diastereomers, α-ylangene and α-copaene, through ring closure from C-2 to either the si or re face of C-7. Apparently, attack from C-3 of the cadinyl cation is forbidden, since this would yield cubebenes or sativenes which have not been detected as enzymatic products. It is noteworthy that many of the minor products of the δ-selinene synthase also occur as minor products in the volatiles of guava (Psidium guajava) leaves, which is also a source of δ-selinene (37). It may be that the co-occurrence of this spectrum of minor products reflects a common reaction mechanism for the δ-selinene synthases of grand fir and guava.

Although the array of products generated by δ-selinene is remarkable, an even more bewildering spectrum of sesquiterpene olefins is produced by the synthase encoded by ag5. γ-Humulene, the principal olefin for which this synthase has been named, sibirene, and longifolene were identified by GLC-MS analysis (Fig. 3) as major products. In addition, 23 other sesquiterpenes were identified, along with 26 unknown sesquiterpene olefins, for a total of 52 different products (Table I). These products range in complexity from the simple acyclic olefin (E)-β-farnesene to the complex tetracyclic olefins longicyclene and cyclosativene.

Whereas δ-selinene synthase appears to utilize both cisoid- and transoid-nerolidyl diphasphate as cyclization intermediates, the γ-humulene synthase may utilize exclusively cisoid-nerolidyl diphasphate (Fig. 5). Thus, with δ-selinene synthase, most of the germacrones produced bear a trans-C2,C3-double bond, thereby revealing the transoid nature of the nerolidyl diphasphate intermediate. By contrast, none of the 26 identified products of γ-humulene synthase bear a trans-C2,C3-double bond, suggesting that the reaction cascade (Fig. 5) proceeds entirely from cisoid-nerolidyl diphosphate, even in the formation of the acyclic olefin (E)-β-farnesene and of sibirene and β-gurjunene which, as eudesmanes, are usually considered to originate via the transoid intermediate (32, 33, 35).

In the proposed reaction scheme (Fig. 5), C-6 to C-1 closure followed by deprotonation will produce the various bisabolenes in a manner analogous to the corresponding monoterpene cyclization (34). All remaining cyclizations catalyzed by this enzyme require C-11 to C-1 closure to form a C11 macrocycle (humulene cation) or C-10 to C-1 closure to produce a C10 macrocycle in (Z,E)-configuration (Fig. 5). Approximately 62% of the reaction flux is shuttled through the C11 macrocycle and 25% is directed through the C10 macrocycle. The C10 macrocycle can be deprotonated directly to form (Z,E)-germacrene B, which is also produced by δ-selinene synthase. Hydride shift and deprotonation of the cyclodecadienyl cation from the C3-methyl (C-15) leads to germacrene D, which, following reprovocation at C-6, allows C-2, C-7 ring closure to the eudesmane skeleton of sibirene. In the production of β-gurjunene, cyclopropane ring formation precedes protonation (at C-6) and C-2, C-7 closure, and is followed by an additional hydride shift and methyl migration before deprotonation to complete the reaction. In the formation of sibirene and β-gurjunene, the original cis-configuration of the macroyclic intermediate is obscured in the transformations.

If the 1,3-hydride shift in the (Z,E)-C10 macrocycle is followed by C-6 to C-1 closure, the amorphenes (but apparently not the diastereomeric cadinenes) can be formed by deprotonation, whereas subsequent C-2 (or C-3) to C-7 closure via the remaining double bond yields copaenes and ylangenes, as with δ-selinene synthase, or sativene and cyclosativene. The formation of sativene, and possibly cyclosativene, must occur via Wagner-Meerwein rearrangement involving a cation-induced 1,2-carbon shift, analogous to the 1,2-methyl shift that occurs in the formation of β-gurjunene.
In the case of the C_{11} macrocyclic cation, 1,3-hyride shift and deprotonation from C-3 will afford γ-humulene (Fig. 5), in a manner analogous to the formation of germacrene D from the C_{15} macrocycle. Ring closure from C-6 to C-1 of the humulyl cation provides the himachalyl cation, from which three alternatives for deprotonation yield α-, β- and γ-himachalenes in which the cisoid conformation at C-2, C-3 of the original nerolidyl intermediate is preserved. Ring closure from C-2 to C-7 of the himachalyl cation leads to α- and β-longipinene, whereas closure from C-3 to C-7, with Wagner-Meerwein rearrangement, leads to longifolene and longicyclene. These additional ring closures, via the double bond of the bicyclo[5:4:0](himachalyl) cation, are analogous to the closures of the bicyclo[4:4:0](cadinyln/amarophyl) cation involved in the formation of the copaenes and sativenes. Similarly, the Wagner-Meerwein rearrangement involved in the formation of sativene (and possibly...
Cyclosativene is analogous to that which occurs in the formation of longifolene (and possibly longicyclene). Zavarin and co-workers (38) have previously identified γ-humulene in the cortical oleoresin of many *Abies* species, including *A. grandis*, and have suggested that this macrocyclic olefin is a by-product of longifolene biosynthesis.

Although γ-humulene synthase appears to be restricted to utilization of the cisoid-nerolidyl diphosphate intermediate, as opposed to δ-selinene synthase which utilizes both cisoid and transoid forms, the former is able to catalyze formation of the greater number and more diverse products, including acyclic, monocyclic, bicyclic, tricyclic, and tetracyclic types, as well as olefins produced by Wagner-Meerwein rearrangements. It is worth noting that, despite the remarkable number of different sesquiterpene skeletal types generated by these two enzymes, deprotonation in each set of structures occurs from the same

![Diagram of proposed mechanism for sesquiterpene formation by γ-humulene synthase](image.png)

**FIG. 5.** Proposed mechanism for sesquiterpene formation by γ-humulene synthase (ag5). Numbering of the farnesyl system is indicated, and OPP denotes the diphosphate moiety. Structures enclosed in boxes indicate products that have been identified by GLC-MS. WMR indicates a Wagner-Meerwein rearrangement.
few carbons of the common substrate, providing at least some measure of uniformity between the two and suggesting the involvement of a limited number of enzyme bases in catalysis. The γ-humulene synthase, in particular, catalyzes several very complex reaction cascades and generates far more products than any terpenoid synthase thus far described, and it along with δ-selinene synthase account for many of the constitutively produced cortical sesquiterpenes. However, the cDNA species encoding synthases responsible for the formation of several of the more abundant constitutive sesquiterpenes of grand fir oleoresin (e.g., α-muurolene, α-copaene, α-cubenene, and β-caryophyllene) have not yet been acquired. Although (E)-α-bisabolene is a product of γ-humulene synthase, and δ-cadinene is a product of δ-selinene synthase, these two sesquiterpenes account for only a small fraction of the many olefins generated by these two synthases. Thus, γ-humulene synthase and δ-selinene synthase cannot be responsible for the wound-induced production of (E)-α-bisabolene and δ-cadinene.

**Sesquiterpene Synthase Characterization**—The $K_e$ values for [1-$^3$H]farnesyl diphosphate with γ-humulene synthase and δ-selinene synthase were estimated to be about 4.5 and 1.5 μM, respectively. The metal ion requirements of γ-humulene synthase and δ-selinene synthase were also evaluated, as cofactor specificity is often characteristic of the different terpenoid synthase types (39). δ-Selinene synthase shows a distinct preference for Mg$_{2+}$; the maximum rate with Mn$_{2+}$ is less than 10% of that with Mg$_{2+}$ at saturation. By contrast, γ-humulene synthase can utilize Mg$_{2+}$ or Mn$_{2+}$ with comparable velocities in the cyclization reaction. For both enzymes, the $K_m$ value for Mg$_{2+}$ is about 125 μM and for Mn$_{2+}$ about 25 μM. Neither of the sesquiterpene synthases requires K$^+$ or other monovalent cation for activity. The monoterpen synthases from conifers require Mn$_{2+}$ or Fe$_{2+}$ for activity, but Mg$_{2+}$ fails to support catalysis (9, 40), and these enzymes also exhibit an absolute requirement for a monovalent cation, with K$^+$ preferred (9, 39).

Substrate specificity of these sesquiterpene synthases was evaluated by comparing farnesyl diphosphate to geranyl diphosphate (C$_{10}$) and geranylglycerol diphosphate (C$_{20}$) at saturation as precursors of the respective terpene olefins. Both enzymes failed to generate detectable olefinic products from geranylglycerol diphosphate, whereas both synthesized monoterpenes from geranyl diphosphate at roughly half the rate of sesquiterpene biosynthesis from farnesyl diphosphate. The identities of the monoterpeno synthases produced by the sesquiterpene synthases were determined by GLC-MS. Limonene is the principal monoterpeno product of both synthases (see Fig. 6 for product structures and amounts), with most of the other products being made in roughly comparable proportions by each, with the notable exception of (Z)-ocimene which is a major product of δ-selinene synthase (22.4%) but is not detectable as a product of γ-humulene synthase. γ-Humulene synthase produces detectable amounts of camphene, the biosynthesis of which requires a Wagner-Meerwein rearrangement as with longifolene and sativene; δ-selinene synthase does not produce this monoterpene.

**Sequence Analysis**—The δ-selinene synthase cDNA encodes a protein that is 581 amino acids in length with a predicted molecular weight of 67,625, and the γ-humulene synthase cDNA encodes a protein of 593 residues with a predicted molecular weight of 67,937 (Fig. 7). The γ-humulene synthase sequence contains a stop codon in frame with the putative initiation methionine at −21 bp of the 89-bp 5′-untranslated region, whereas the δ-selinene synthase sequence is truncated at −12 bp. The nucleotide sequence surrounding the putative starting ATG of both sesquiterpene synthase genes is conserved and resembles that which surrounds the initiating methionine of other plant genes (41). These data support the proposed location of the initiation sites and, thus, the identification of both cDNAs as sesquiterpene synthases, since the predicted molecular weights are appropriate for this class of cytosolic enzymes (16, 18) which lack a plastidial targeting peptide found in both monoterpeno synthase and diterpene synthase preproteins (8, 9, 34, 42).

Comparison of the deduced amino acid sequences indicates that they are 83% similar and 65% identical to each other (Fig. 7), and both sesquiterpene synthases show about 66% similarity and 45% identity when compared with other terpenoid synthases from conifers, including abietadiene synthase (diterpene) from grand fir (8), myrcene synthase, pinene synthase, and limonene synthase (monoterpenes) from this species (9), and the taxadiene synthase (diterpene) from Pacific yew (43). Comparison of these gymnosperm sesquiterpene synthases to a range of terpenoid synthases from angiosperms reveals similarities of about 56% and identities in the range of 30%. Interestingly, the gymnosperm sesquiterpene synthases show a greater resemblance to gymnosperm monoterpen synthases and diterpene synthases than they do to angiosperm sesquiterpene synthases, suggesting an ancient divergence of the gymnosperm tspsd family (9), which consists of monoterpeno, sesquiterpene, and diterpene synthases. Comparison with microbial sesquiterpene synthases evidenced no significant similarity. For example, the δ-selinene and γ-humulene synthases show 45–46% similarity and 18–22% identity when compared with trichodiene synthase from Fusarium sporotrichioides (44).

Studies employing amino acid-modifying reagents have implicated histidine, cysteine, and arginine residues in catalysis by terpene synthases from angiosperms and gymnosperms (7, 13, 14, 39, 40, 45–47). Comparisons of all published terpene synthase sequences reveal Cys$_{507}$ (with reference to γ-humulene synthase) as the only universally conserved cysteine residue, and His$_{95}$ and His$_{141}$ as the only conserved histidines. However, comparisons between only the sesquiterpene synthases reveal an additional conserved histidine residue at position 283. Comparison among the sesquiterpene synthases also shows that eight arginine residues are conserved at positions 136, 139, 243, 247, 306, 329, 364, and 485 of γ-humulene synthase, a surprising number considering the broad taxo-
design PCR primers are indicated in boldface. Conserved histidine, cysteine, and arginine residues (from comparison of all sesquiterpene synthase sequences published to date) are indicated by single arrows. Regions of sequence that were used to design PCR primers are indicated in boldface, and the amino acids encoded by the nucleotide probes that were used to isolate the corresponding full-length cDNAs are indicated by double arrows. The analysis of the sesquiterpene fraction of grand fir oleoresin reported here for the Rocky Mountain ecotype agrees well with a previous analysis of this material from the coastal ecotype (38) with but minor differences between the former (19% germacrene B without detectable β-elemene) and the latter (8% β-elemene without detectable germacrene B). The discrepancy is likely the result of misidentification due to methodology (identification of β-elemene by retention time only) and has been rectified by recent re-analysis of the oleoresin of the coastal ecotype.

Upon wound stomping, two sesquiterpene synthase activities are induced, one for the increased production of a prominent constitutive component (δ-cadinene; see Fig. 1) and one for the production of a very minor sesquiterpene of the constitutive oleoresin (Eβ-o-bisabolene). This situation is reminiscent of that observed with the constitutive and inducible monoterpene synthases of grand fir (4, 6, 21).

Although δ-selinene synthase and γ-humulene synthase are capable of producing monoterpene when presented with geranyl diphosphate, several lines of evidence indicate that these enzymes are, in fact, sesquiterpene synthases. First, the corresponding cDNA species do not appear to encode preproteins bearing a plastidial transit peptide characteristic of monoterpene (and diterpene) synthases but rather mature proteins of a size typical of this class of cytosolic enzymes. Second, the divalent and monovalent ion requirements do not resemble those of sesquiterpene synthases to produce monoterpenes but rather those of other sesquiterpene synthases. Finally, the acyclic monoterpenes (ocimenes) produced by δ-selinene synthase and γ-humulene synthase from geranyl diphosphate are not found in the turpentine fraction of grand fir oleoresin (3, 22, 38). The accumulated evidence therefore clearly supports the identification of these enzymes as sesquiterpene synthases. Since sesquiterpene biosynthesis occurs in the cytosol where the precursor farnesyl diphosphate is also synthesized, whereas the monoterpene synthases are compartmentalized within plastids where the precursor geranyl diphosphate also arises (11, 34, 55, 56), the ability of the sesquiterpene synthases to produce monoterpenes in vitro may simply represent the adventitious utilization of a substrate that is never encountered in vivo and against which there is no evolutionary pressure to discriminate. It now seems likely that the adventitious utilization of geranyl diphosphate by the sesquiterpene synthases accounts, in part, for the relatively high level of limonene synthase activity observed in crude stem extracts of grand fir (21).

The ability of terpene synthases to produce multiple products has been well documented (14, 39, 40, 47, 57) and may be a consequence of the unusual electrophilic reaction mechanisms employed by this enzyme type (11, 35, 58) that may also represent an evolutionary adaptation for the production of the maximum number of terpene products using the minimum genetic and enzymatic machinery (59). Nevertheless, the production of 34 different sesquiterpenes by δ-selinene synthase

3 L. Cool, personal communication.
and 52 discrete sesquiterpenes by γ-humulene synthase, by variations upon several different cyclization routes, is quite remarkable. The reaction cascade catalyzed by γ-humulene synthase is particularly complex in generating (by deprotonation) stable olefinic end products corresponding to many of the proposed carbocationic intermediates of each cyclization route (Fig. 5). Significantly, the essential elements of these cyclization schemes have been delineated by Arigoni and collaborators (32, 60, 61) via a series of elegant in vivo labeling studies directed toward the biosynthesis of longifolene and sativene in the fungi Helminthosporium victoriae or sativum or the gymnosperm Pinus ponderosa. Additionally, in vivo studies with (5R)- and (5S)-[5-3H]mevalonate provided evidence, based upon the observation of isotopically sensitive branching (62), that the formation of (−)-longifolene and (−)-sativene was catalyzed by a single enzyme. The isolation and functional expression of the γ-humulene synthase cDNA reported here provides direct and unequivocal proof for this earlier, prescient biosynthesis of the γ-systems, the means for examining structure-function relationships in these mechanistically fascinating catalysts. These cDNAs should also provide access to genomic clones to allow comparison of intron/exon structure of these genes to their angiosperm counterparts (40, 65). Grand fir is the first plant from which cDNA species encoding representative monoterpene, sesquiterpene, and diterpene synthases have been isolated (8, 9). Although the sequence comparisons are in themselves instructive, the availability of these clones should permit a more highly refined understanding of oleoresinosis and lead to the manipulation of this defensive secretion in the protection of conifer species against the devastating environmental and economic effects of bark beetle predation (1, 2, 66, 67).

Acknowledgments—We thank Larry Cool, Ron Binder, Robert Adams, Margaret Eisenberg, and Rudolf Hopp for the gifts of sesquiterpene reference standards; Jack Alexander of the Harvard Arnold Arboretum for the gift of A. alba tissue; Joyce Tamura-Brown for typing the manuscript; and Thom Koehler for raising the trees. We thank Eva Katahira for technical assistance and Gerhard Musanke of the Laboratory for Biotechnology and Bioanalysis for DNA sequencing. We are especially grateful to Larry Cool (Forest Products Laboratory, University of California, Berkeley) for expert assistance in the identification of the sesquiterpenes from grand fir oleoresin.

REFERENCES

1. Johnson, M., and Croteau, R. (1987) ACS Symp. Ser. 325, 76–91
2. Gijzen, E., Lewinsohn, E., and Croteau, R. (1993) ACS Symp. Ser. 525, 8–22
3. Lewinsohn, E., Savage, T. J., Gijzen, M., and Croteau, R. (1993) Phytochemistry 34, 220–225
4. Lewinsohn, E., Gijzen, M., and Croteau, R. (1992) ACS Symp. Ser. 497, 8–17
5. Gijzen, M., Lewinsohn, E., and Croteau, R. (1992) Arch. Biochem. Biophys. 294, 670–674
6. Funk, C., Lewinsohn, E., Vogel, B. S., Steele, C. L., and Croteau, R. (1994) Plant Physiol. (Bethesda) 106, 999–1005
7. LaFever, R. E., Vogel, B. S., and Croteau, R. (1994) Arch. Biochem. Biophys. 313, 139–149
8. Vogel, B. S., Wildung, M., Vogel, G., and Croteau, R. (1996) J. Biol. Chem. 271, 32062–32069
9. Bohlmann, J., Steele, C. L., and Croteau, R. (1997) J. Biol. Chem. 272, 21784–21792
10. Threlfall, D. R., and Whitehead, I. M. (1991) in Ecological Chemistry and Biochemistry of Plant Terpenoids (Harborne, J. B., and Tomas-Barberan, F. A., eds) pp. 159–208, Clarendon Press, Oxford
11. Gerashenkov, J., and Croteau, R. (1993) in Lipid Metabolism in Plants (Moore, F. S., Jr., ed) pp. 340–388, CRC Press, Inc., Boca Raton, FL
12. Salin, F., Pauly, G., Charon, J., and Gleizes, M. (1995) J. Plant Physiol. 146, 203–209
13. Dehal, S. S., and Croteau, R. (1988) Arch. Biochem. Biophys. 261, 346–356
14. Munch, S. L., and Croteau, R. (1996) Arch. Biochem. Biophys. 282, 58–64
15. Belingheri, L., Cartayrade, A., Pauly, G., and Gleize, M. (1992) Plant Sci. (Shannon) 84, 129–136
16. Parchin, P. J., and Chappell, J. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 11088–11092
17. Back, K., and Chappell, J. (1995) J. Biol. Chem. 270, 7375–7381
18. Chen, C.-Y., Chen, Y. H., Heinze, P., and Davidson, V. J. (1995) Arch. Biochem. Biophys. 324, 255–266
19. Steele, C. L., Lewinsohn, E., and Croteau, R. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 4146–4150
20. Croteau, R., Alonso, W. R., Koepp, A. E., and Johnson, M. A. (1994) Arch. Biochem. Biophys. 309, 184–192
21. Gijzen, M., Lewinsohn, E., and Croteau, R. (1991) Arch. Biochem. Biophys. 289, 267–273
22. Lewinsohn, E., Gijzen, M., Savage, T. J., and Croteau, R. (1991) Plant Physiol. (Bethesda) 96, 38–43
23. Croteau, R., and Canes, D. E. (1985) Methods Enzymol. 110, 383–405
24. Satterwhite, D. M., and Croteau, R. (1988) J. Chromatogr. 452, 61–73
25. Bradford, M. M. (1976) Anal. Biochem. 72, 248–254
26. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1989) Current Protocols in Molecular Biology, Greene Publishing Associates and Wiley-Interscience, New York
27. Lewinsohn, E., Steele, C. L., and Croteau, R. (1994) Plant Mol. Biol. Rep. 12, 20–25
28. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Laboratory Manual. Cold Spring Harbor, NY
29. Schenck, P. M., Baumann, S., Mattes, R., and Steinbiss, H.-H. (1995) Biotechniques 19, 196–200
30. Genetics Computer Group (1996) Program Manual for the Wisconsin Package, Version 9.0, Genetics Computer Group, Madison, WI
31. Lewinsohn, E., Gijzen, M., and Croteau, R. (1991) Plant Physiol. (Bethesda) 96, 44–49
32. Cane, D. E. (1981) in Biosynthesis of Isoprenoid Compounds (Porter, J. W., and Spurgeon, S. L., eds) Vol. 1, pp. 263–374, John Wiley & Sons, Inc., New York
33. Cane, D. E. (1990) Chem. Rev. 90, 1089–1103
34. Colby, S. M., Alonso, W. R., Katahira, E. J., McGarvey, D. J., and Croteau, R. (1993) J. Biol. Chem. 268, 20366–20374
35. Cane, D. E. (1992) Ciba Found. Symp. 171, 163–183
36. Winter, R. E. K., Dorn, F., and Arigoni, D. (1980) J. Org. Chem. 45, 4786–4789
37. Sagrero-Nieves, L., Bartley, J. P., Provis-Schwede, A. (1993) Flavour Fragrance J. 8, 135–137
38. Smedman, L. A., Snajberk, K., Zavarin, E., and Mon, T. R. (1969) J. Org. Chem. 34, 4164–4168
39. Savage, T. J., Hatch, M. W., and Croteau, R. (1994) J. Biol. Chem. 269, 4012–4020
40. Lewinsohn, E., Gijzen, M., and Croteau, R. (1994) Arch. Biochem. Biophys. 293, 167–173

Cloning of Conifer Sesquiterpene Synthases
