Monoterpene Synthases from Common Sage (Salvia officinalis)*

Mitchell L. Wise‡, Thomas J. Savage§, Eva Katahira, and Rodney Croteau¶

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

Common sage (Salvia officinalis) produces an extremely broad range of cyclic monoterpenes bearing diverse carbon skeletons, including members of the ρ-menthane (1,8-cineole), pinane (α- and β-pinene), thujane (isothujone), camphane (camphene), and bornane (camphor) families. An homology-based polymerase chain reaction cloning strategy was developed and used to isolate the cDNAs encoding three multiproduct monoterpene synthases from this species that were functionally expressed in Escherichia coli. The heterologously expressed synthases produce (+)-bornyl diphosphate, 1,8-cineole, and (+)-sabinene, respectively, as their major products from geranyl diphosphate. The bornyl diphosphate synthase also produces significant amounts of (+)-α-pinene, (+)-camphane, and (±)-limonene. The 1,8-cineole synthase produces significant amounts of (+)- and (−)-α-pinene, (+)- and (−)-β-pinene, myrcene and (±)-sabinene, and the (+)-sabinene synthase produces significant quantities of γ-terpinene and terpinolene. All three enzymes appear to be translated as preproteins bearing an amino-terminal plastid targeting sequence, consistent with the plastidial origin of monoterpenes in plants. Deduced sequence analysis and size exclusion chromatography indicate that the recombinant bornyl diphosphate synthase is a homodimer, whereas the other two recombinant enzymes are monomeric, consistent with the size and subunit architecture of their native enzyme counterparts. The distribution and stereochemistry of the products generated by the recombinant (+)-bornyl diphosphate synthase suggest that this enzyme might represent both (+)-bornyl diphosphate synthase and (+)-pinene synthase which were previously assumed to be distinct enzymes.

The cyclization of the universal precursor geranyl diphosphate to monocular and bicyclic monoterpenes is catalyzed by a group of enzymes termed monoterpene synthases (or cyclases). The biochemical transformation of geranyl diphosphate to cyclic products has been investigated using enzymes from a variety of plants, including both angiosperms (1) and gymnosperms (2–4), and a mechanistic paradigm for these transformations (Scheme 1) is well established (1, 5). Thus, geranyl diphosphate is initially ionized and isomerized to form either (3R)- or (3S)-linyl diphosphate, depending on the particular enzyme. This step permits rotation about the C2–C3 single bond of the bound allylic isomer to the cisoid conformation which, upon subsequent ionization, promotes electrophilic attack by C1 on the C6–C7 double bond, resulting in the formation of the α-terpinyl cation as a central intermediate. Further transformations of this reactive intermediate may be effected by additional intramolecular electrophilic additions, hydrate shifts, or other rearrangements before termination of the sequence by deprotonation of the final cation or capture by an external nucleophile, such as a hydroxyl ion or the diphosphate group. Although the fate of the substrate has been well characterized in numerous monoterpen cyclization reactions, the molecular mechanisms by which the enzymes effect these transformations is still poorly understood.

Culinary sage (Salvia officinalis) produces a number of monoterpenes, including (+)- and (−)-α-pinene, (+)- and (−)-β-pinene, (+)- and (−)-camphene, (+)-sabinene, (+)- and (−)-limonene, myrcene, 1,8-cineole, and (+)-bornyl diphosphate (Scheme 1) (1). Because sage produces this broad range of acyclic, monocyclic, and bicyclic monoterpenes, including several olefin isomers, a cyclic ether and a diphosphate ester, this plant has provided an ideal system for the study of a variety of synthases, all of which utilize the same substrate but produce different products by variations on a single reaction mechanism (1, 5). These include (+)-bornyl diphosphate synthase (the enzyme producing the precursor of (+)-camphor) (6, 7), 1,8-cineole synthase (8), (+)-sabinine synthase (the enzyme producing the precursor of (−)-3-isothujone) (9, 10), and several pinene synthases (11–14). As is typical of monoterpen cyclases (5, 15), many of these enzymes from sage appear to generate multiple products from geranyl diphosphate. Investigations with the partially purified native enzymes have suggested that a single enzyme, termed (+)-pinene synthase (cyclase I), is responsible for the synthesis of both (+)-α-pinene and (+)-camphene, with lesser amounts of (+)-limonene and myrcene, whereas a second enzyme, (−)-pinene synthase (cyclase II), has been shown to produce (−)-α-pinene, (−)-β-pinene, and (−)-camphene, with minor amounts of (−)-limonene, terpinolene, and myrcene (11, 12). More recently, a third synthase from sage, termed cyclase III, has been described which produces a mixture of (+)-α-pinene and (−)-β-pinene, along with minor amounts of myrcene (13, 14). Evidence that these reactions are catalyzed by individual multifunctional enzymes is provided by
co-purification and differential inhibition studies (12), as well as by isotopically sensitive branching experiments (13, 15, 16). Despite considerable effort, the (+)-pinene synthase has never been chromatographically separated from the aforementioned (+)-bornyl diphosphate synthase (17), nor has the (-)-pinene synthase been fully resolved from 1,8-cineole synthase, although stereochemical considerations indicate that the latter two are probably distinct enzyme species (8, 18).

In this report, we describe the homology-based cloning, and subsequent sequencing and heterologous expression, of three monoterpene synthase cDNA genes from sage, the recombinant enzymes from which produce three different major types of cyclic monoterpene products, (+)-sabinene (a bicyclic olefin), 1,8-cineole (a bicyclic ether), and (+)-bornyl diphosphate (a bicyclic diphosphate ester), respectively (Scheme 1). Comparison of the sizes, subunit architectures, and product distributions of these multiple-product enzymes clarifies the assignment of specific catalytic functions to defined monoterpene synthases, and the deduced sequences provide information on the relatedness of these enzymes within the species and to other terpenoid synthases of plant origin. Additionally, comparison of the primary structures of these mechanistically different monoterpene synthases from the same plant, the first examples of this type thus far available, allows preliminary assessment of active site interactions and provides the foundation for more detailed study of structure-function relationships.

EXPERIMENTAL PROCEDURES

Plant Materials, Substrate, and Reagents—Sage plants (S. officinalis L.) were grown from seed as described previously (19). [1-3H]Geranyl diphosphate (250 Ci/mol) was prepared by an established method (8). Terpenoid standards were from our own collection. Unless otherwise stated, all reagents were obtained from Sigma or Aldrich Chemical Co. DNA sequences were assembled and analyzed using GCG software (20).

**Scheme 1. Conversion of geranyl diphosphate to the monoterpene of sage.** Formation of the monocyclic and bicyclic products requires preliminary isomerization of geranyl diphosphate to linalyl diphosphate. Acyclic products can be formed from either geranyl diphosphate or linalyl diphosphate. OPP denotes the diphosphate moiety.
tissue) from 3-week-old plants were ground to a fine powder in liquid nitrogen and extracted into buffer composed of 200 mM Tris-HCl (pH 8.5), 300 mM LiCl, 5 mM thionoic, 1 mM uracilribocarboxyl acid, 10 mM dithiothreitol, and 10 mM EDTA, and containing 1% (v/v) polyvinylpyrrolidone (Mw ~ 40,000). Total RNA thus extracted was prepared by phenol-chloroform extraction. After subsequent precipitation with 50% ethanol, the RNA was subjected to reverse transcriptase (RT) PCR using a Superscript II reverse transcriptase (Life Technologies, Inc.) and primers 2F and 3R. For comparison of the cDNA sequences of a monoterpene synthase (23), a sesquiterpene synthase (24), and a diterpene synthase (25) of angiosperm origin, thus three, regions of deduced amino acid sequence (corresponding to residues 180–187, 197–203, and 380–387 of the limesiphone synthase from spearment (23)) were employed to design primers corresponding to 1F, 5′-ATG(A/G)AA/C/TG/G/A/G/A/GAiAGGI(G/A)G/A/TATT/GA/GA/ TACTA/C/G/A/G/G/A/G: 2F, 5′-ATG/T/T/C/TCA/G/A/C/T/T/T/TTT/GA/ GGA/G/G/C-3′; and 3R 5′-CTTG/C/TG/TG/G/AGG/CG/G/A/TAG/A/TAC/G/G/T/CTC-3′. Using purified sage leaf cDNA library phage as template (5 or at 1 × 10^6 plaque-forming units/ml), PCR was performed under a wide range of amplification conditions (26, 27) in a total volume of 50 µl containing 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 2.5 mM MgCl2, 0.2 mM dNTPs, 0.5 µM of each primer, and 0.5 µl of Taq polymerase (Life Technologies, Inc.). Analysis of the PCR reactions of products by agarose gel electrophoresis (26) indicated that only the combination of primers 2F and 3R amplified a discrete product of approximately 600 bp, which was ligated into pT7Blue (Novagen), and transformed into E. coli NovaBlue cells. Plasmid DNA was prepared from 32 individual transformants; seven of these had inserts of the predicted size (~400 bp). These inserts were partially sequenced (DyeDeoxy Terminator Cycle Sequencing; Applied Biosystems) to confirm molecular weights (29), which were then corrected for the engineered XhoI and the soluble supernatant prepared as before. Purification of the resulting fusion proteins was attempted using the glutathione-Sepharose affinity column according to the manufacturer's instructions (Pharmacia). Of the three expressed monoterpene synthases (see below), only one (SBS) bound to the matrix but, even in this case, affinity-based purification proved to be unreliable. Therefore, partial purification of the heterologously expressed synthases was achieved by ion-exchange chromatography on Q-diethylaminoethyl-cellulose (Whatman DE-52) using a 0–400 mM NaCl gradient. The partially purified preparations were desalted by repeated ultrafiltration and dialysis using an Amicon Centriprep 30 concentrator (30 kDa cutoff) and the appropriate assay buffer. The PCR-expressed fusion proteins were also subjected to gel permeation chromatography by loading T7 promoter primers, and transformed clones containing an insert. The size of each cDNA insert was determined by PCR using T3 and T7 promoter primers, and transformed clones containing an insert. The size of each cDNA insert was determined. Reverse transcription (RT) PCR was used to screen each cDNA library in an attempt to isolate full-length cDNA sequences encoding the corresponding native forms. The pGEX constructs in E. coli XL1-Blue-MRF cells yielded expression to assay for monoterpene synthase activity. The sonicates were cleared by centrifugation, as described previously (21). Poly(A)+ mRNA was isolated by chromatography on oligo(dT)-cellulose (Qiagen) and 6.3 µg of the resulting mRNA was used to construct a zAPII cDNA library according to the manufacturer's instructions (Stratagene).

A general strategy for the homology-based PCR cloning of terpenoid synthases of higher plant origin has been described (22). Based on the comparison of the cDNA sequences of a monoterpene synthase (23), a sesquiterpene synthase (24), and a diterpene synthase (25) of angiosperm origin, thus three, regions of deduced amino acid sequence (corresponding to residues 180–187, 197–203, and 380–387 of the limesiphone synthase from spearment (23)) were employed to design primers corresponding to 1F, 5′-ATG(A/G)AA/C/TG/G/A/G/A/GAiAGGI(G/A)G/A/TATT/GA/GA/ TACTA/C/G/A/G/G/A/G: 2F, 5′-ATG/T/T/C/TCA/G/A/C/T/T/T/TTT/GA/ GGA/G/G/C-3′; and 3R 5′-CTTG/C/TG/TG/G/AGG/CG/G/A/TAG/A/TAC/G/G/T/CTC-3′. Using purified sage leaf cDNA library phage as template (5 or at 1 × 10^6 plaque-forming units/ml), PCR was performed under a wide range of amplification conditions (26, 27) in a total volume of 50 µl containing 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 2.5 mM MgCl2, 0.2 mM dNTPs, 0.5 µM of each primer, and 0.5 µl of Taq polymerase (Life Technologies, Inc.). Analysis of the PCR reactions of products by agarose gel electrophoresis (26) indicated that only the combination of primers 2F and 3R amplified a discrete product of approximately 600 bp, which was ligated into pT7Blue (Novagen), and transformed into E. coli NovaBlue cells.

The Bluescript plasmids expressed in E. coli XL1-Blue-MRF cells were transformed into E. coli SOLR cells according to the Stratagene protocol.
Cloning and Expression of Monoterpene Synthases from Sage

RESULTS AND DISCUSSION

Similarity-based Cloning of Monoterpene Synthases from S. officinalis—Because S. officinalis (culinary or common sage) produces such a broad structural variety of monoterpenes, this species has been utilized extensively for studies on the enzy-
mology, stereochemistry, and mechanism of monoterpene cy-
clication reactions (1, 5). Structural analyses of the responsible
monoterpene synthases, and more detailed study of the cycli-
ze mechanisms, require the isolation of cDNA sequences en-
coding these target enzymes. Protein purification from sage, as
the basis for cDNA isolation, has been of limited success (17)
because of the number of synthases present and their similar-
ity in physical properties (32), and thus far has not permitted a
reverse genetic approach to cloning of any of the monoterpene
synthases from this species.

As a possible alternative to protein-based cloning of terpene
synthases, a homology-based PCR strategy was recently pro-
posed (22) that was developed by comparison of the deduced
amino acid sequences of cDNAs encoding a monoterpene syn-
thase (23), a sesquiterpene synthase (24), and a diterpene
synthase (25) of phylogenetically distant angiosperm species.
Three conserved regions of sequence were identified that ap-
ppeared to be useful for the design of degenerate PCR primers.
Two of these primers ultimately amplified a 600-bp fragment
using cDNA from a sage leaf library as template. Cloning and
sequencing showed the amplified products to correspond to two
distinct sequence groups, both of which showed similarity to
sequences of cloned terpene synthases, but only one of which
hybridized strongly to a 2-kilobase target upon Northern blot
analysis of sage leaf mRNA. This more efficient probe was
utilized to screen the sage leaf cDNA library, from which 77
positive phagemids were purified. Size selection of the purified
and in vivo excised clones yielded a subset of 44 with insert size
greater than 1.6 kilobases, and these were expressed in E. coli
XLI-Blue cells and the resulting extracts were assayed for
functional monoterpene synthase activity by monitoring the
conversion of [1-1H]geranyl diphosphate to monoterpene ole-
fins, oxygenated monoterpenes, and monoterpenyl diphosphate
esters.

Nine functionally active clones were identified by this
means, four types of which showed unique product profiles.
Two cDNA clones, of which the clone designated 3C6 was most
active in expression, yielded an enzyme in the corresponding
identity with the authentic standard. Full spectra were re-
corded for major reaction products which were identified by comparison of retention times to authentic standards and by comparison of spectra to those of the NSB75K library using the G1033A NIST probability
based matching algorithm. Chiral phase separations were performed on
a Hewlett-Packard 5890 GC by split injection (80:1) on a 30-m cyclo-
dex-B capillary column (J & W Scientific) using H2 as carrier at 0.6
ml/min and temperature programming from 70 to 200 °C at 10 °C/min
with flame ionization detection. Compound identification was based on
retention time identity with the authentic standard.

Additional clones, of which clone 3F25 was seemingly the most
active in expression, yielded E. coli extracts capable of trans-
forming geranyl diphosphate to sabinene as the dominant ole-
fin product. This acquisition was named SSS (sage sabinene
synthase), with correspondence assigned to the native (+)-
sabinene synthase which catalyzes the cyclization to the bicy-
clic olefin precursor of (−)-isothujone (9, 10, 36). The last func-
tional clone (3M13) yielded an expressed activity that catalyzed
the conversion of geranyl diphosphate to the monocyclic olefin
limonene as essentially the only product.

Clones 3C6, 3B5, 3F2P, and 3M13 were fully sequenced (see
Fig. 1 for deduced amino acid alignments), whereupon it was
shown that the insert of 3M13 was a 5′-truncated version of
3P25 (data not shown), resulting in the translation of a short-
ened fusion protein starting at residue 106 of the sabinene
synthase (SSS) that apparently leads to derailment of the nor-
mal bicyclization to instead yield the monocyclic product limo-
none. The remaining 35 positive clones, that were inactive in
expressing a functional monoterpene synthase activity, were
partially sequenced from the 5′ terminus to search for addi-
tional truncated versions or inserts that were out of frame for
proper expression. This approach revealed three truncated and/or out of frame versions of SBS, 20 additional versions of
SCS, and 11 versions of SSS, while unveiling one new clone
(number 3F5 and designated SUS). This clone showed exten-
sive sequence homology to 3C6, 3B5, and 3F2P but encoded a
premature stop codon (Fig. 1) that would result, because of this
cloning artifact, in translational truncation toward the car-
boxyl terminus. A functional form of the SUS protein has not
yet been obtained.

Product Profiles of Recombinant Synthases—Since the for-
mation of multiple products from geranyl diphosphate is a
common, if unusual, feature of the monoterpene synthases (1, 5),
the product profiles of the recombinant enzymes were ex-
amined in detail by radio-GC and GC-MS. Recombinant sabi-
nene synthase (SSS) produces exclusively monoterpene olefins,
which by radio-GC analysis (Fig. 2) were identified as sabinene
(63%), γ-terpinene (21%), terpinolene (7.0%), limonene (6.5%),
and myrcene (2.5%). The major products of this enzyme (sabi-
nene and γ-terpinene) are formed by a cyclization mechanism
involving a 1,2-hydride shift in the α-terpinyl cation interme-
diate (Scheme 1). Identification of sabinene as the major com-
ponent was confirmed by retention time identity with an au-
thentic standard on GC-MS and by comparison of the mass
spectra of the various metabolite fractions (Figs. 2 and 3) to
produce both oxygenated monoterpenes (1,8-cineole, 79%, with
a few percent α-terpinene) and a mixture of olefins (~20%).
Confirmation of 1,8-cineole as the major product was carried
out by GC-MS to demonstrate identity of retention time and
mass spectrum to the authentic standard: m/z 136 (P+ 14%),
94 (P−42, 14%), 93 (P−43, base peak), 91 (P−45, 40%), 80 (P−56,
10%), 79 (P−57, 26%), 77 (P−59, 36%), and 69 (P−67, 9%).
Chiral phase capillary GC analysis demonstrated the biosyn-
thetic sabinene to be coincident with authentic (+)-sabinene
(data not shown); however, the (−)-enantiomer was not avail-
able for analysis to confirm the absolute configuration of this
product. Previous studies have shown that cell-free extracts
from sage produce only the (−)-enantiope of sabinene from ger-
anyl diphosphate (9, 10), supporting the assignment of the
(−)-stereoisomer in this case; the other principal olefinic pro-
ducts of SSS are achiral.

Cineole synthase (SCS) was shown by aliquot counting and
radio-GC of the various metabolite fractions (Figs. 2 and 3) to
produce such a broad structural variety of monoterpenes, this
species (6) that produces the first dedicated intermediate in
(+)-cineole biosynthesis (7, 33, 34). Four clones, of which
clone 3BS yielded the highest activity, expressed a synthase in
bacterial extracts that converted geranyl diphosphate to 1,8-
cineole as the major product. This acquisition was designated
SCS (sage 1,8-cineole synthase) and considered to represent
the native 1,8-cineole synthase, an enzyme for which the me-
chanism of cyclization has been studied in detail (8, 35). Two
confirmation, and quantification of the olefins (Fig. 4) as (1)-a-pinene (5.5% of total products), (2)-a-pinene (0.9%), myrcene (2.9%), sabinene (2.6%, presumably the (1)-enantiomer), (1)-b-pinene (2.7%), (2)-b-pinene (4.1%), (1)-limonene (1.1%), and (2)-limonene (0.4%). The stereochemistry of the enzymatic transformation leading to 1,8-cineole has been examined (8) and shown to involve the cyclization of the bound intermediate (3R)-linalyl diphosphate in anti,endo-conformation, i.e. the same overall stereochemistry required for the production of (+)-a-pinene, (+)-b-pinene, and (+)-limonene (12). The formation of the (2)-series of antipodes must therefore occur via the extended anti,exo-conformation. This apparent loss of stereoselectivity in the production of minor amounts of the olefin by-products may be a consequence of the fact that the enzyme, expressed as the pGEX fusion of the preprotein of the native synthase, was proteolytically processed by the E. coli
host to a form that could compromise substrate and intermediate binding conformations (see below).

Bornyl diphosphate synthase (SBS) was shown, by radio-GC evaluation of all metabolite fractions, to produce principally bornyl diphosphate (75%), as demonstrated by enzymatic hydrolysis of this product followed by separation of the derived borneol from the residual geraniol (liberated from the substrate) and from lesser amounts of nonenzymatic solvolysis products (also generated from geranyl diphosphate in the course of the analysis) (Fig. 3). The production of bornyl diphosphate by this recombinant enzyme was also demonstrated directly by radio-high performance liquid chromatographic analysis of the aqueous reaction mixture using an ion-paring, reversed-phase chromatography protocol previously established for the separation of prenyl diphosphate esters (data not shown) (37). GC-MS analysis of the derived borneol confirmed the identity of this product by comparison of retention time and mass spectrum to the authentic standard: \(m/z\) 154 (P⁺, <1%), 139 (P⁺-15, 8%), 121 (P⁺-33, 6%), 110 (P⁺-44, 18%), 96 (P⁺-58, 8%), 95 (P⁺-59, base peak), 93 (P⁺-61, 7%), 67 (P⁺-87, 8%), and 55 (P⁺-99, 8%). Additionally, chiral phase capillary GC analysis of the derived borneol demonstrated the exclusive presence of the (+)-antipode (data not shown), as expected based on studies with the corresponding native enzyme (6, 38, 39). The recombinant (+)-bornyl diphosphate synthase was also shown, by radio-GC analysis of the olefin fraction (Fig. 2) and chiral phase GC analysis (Fig. 4), to produce a series of olefins (25% of total product) identified as (+)-α-pinene (3.4% of total product), (+)-camphene (9.5%), (-)-camphene (0.5%), (+)-limonene (3.9%), (-)-limonene (3.9%), (+)-terpinolene (2.1%), and myrcene (1.5%). Since formation of the (+)-olefin series is mechanistically related to the formation of (+)-bornyl diphosphate via the anti,endo-cyclization of the intermediate (3R)-linalyl diphosphate (18, 38-41) (Scheme 1), the generation of small amounts of the antipodal (-)-camphene and (-)-limonene by the recombinant cyclase again suggests some loss of stereochemic fidelity in the overall reaction sequence.

**Sequence Analysis**—Alignment of the deduced amino acid sequences (Fig. 1) of SBS clone 3C6 (2025 bp, with an open reading frame of 1794 nucleotides encoding 599 amino acids for a protein of 69.3 kDa and calculated pI of 5.79), and SUS clone 3F25 (1911 bp, with an open reading frame of 1773 nucleotides encoding 591 amino acids for a protein of 69.3 kDa and calculated pI of 6.06), SCS clone 3B5 (1968 bp, with an open reading frame of 1773 nucleotides encoding 591 amino acids for a protein of 69.4 kDa and calculated pI of 6.06), and SSS clone 3F5 (2022 bp), with the published sequences (Fig. 1) of SBS clone 3C6 (2025 bp, with an open reading frame of 1794 nucleotides encoding 599 amino acids for a protein of 69.3 kDa and calculated pI of 5.79), and SSS clone 3F25 (1911 bp, with an open reading frame of 1773 nucleotides encoding 591 amino acids for a protein of 69.3 kDa and calculated pI of 6.06), and for a crude enzyme extract from isolated sage oil gland secretory cells. The bottom tracing is the response of the mass detector to authentic standards of limonene (1), 1,8-cineole (2), trans-sabinene hydrate (3), linalool (4), α-pinene (5), borneol (6), nerol (7), and geraniol (8). Geraniol is derived from hydrolysis of the substrate by contaminating phosphatases, and nerol and linalool are derived from nonenzymatic solvolysis of the substrate during the course of the reaction.
recombinant spearmint limonene synthase preprotein immediately upstream of these tandem arginines yields a fully functional "pseudomature" form of the enzyme, whereas truncation downstream from this element severely impairs activity.

Downstream of the tandem arginines are several regions of homology, including the highly conserved (I, L, or V)DXXD motif (e.g. residues Ile539, Asp355 of SBS) found in virtually all deduced sequences for enzymes that utilize prenyl diphosphate substrates (50, 51). This aspartate-rich element is now generally recognized as a binding site for the metal ion chelated to SBS. The active site peptide LQLYEASFLL, previously isolated from the co-purified (52) has suggested p-cation interactions as a means of stabilizing such reactive intermediates in enzyme-catalyzed reactions, and the concept has received some support from both sequence analysis and x-ray crystallographic studies (59–61). While far from conclusive, the number of conserved aromatic residues displayed among the monoterpene synthases does suggest the possibility of such a role for aromatic side chains in these electrophilic cyclizations.

The (†)-bornyl diphosphate synthase from sage has previously been shown to be inhibited by the "active serine"-directed reagent diisopropyl fluorophosphate, a characteristic not shared by other monoterpene synthases (6). Because of the unique utilization of the substrate diphosphate moiety as the terminating nucleophile by this enzyme (62, 63), it was hypothesized that a serine residue may be involved in binding and transfer of the diphosphate function in the course of the reaction. Sequence comparison of SBS with the other two monoterpene synthases of sage reveals five unique serine residues at positions 67, 255, 302, 454, and 469 (the serine at position 454 is conserved in monoterpene synthases from other species). Two of these serines (at positions 67, 255, 302) are within conserved regions and are, therefore, obvious

| A.g. myrc | A.g. limo | C.b. lino | P.f. limo | M.s. limo | S.o. born | S.o. cine | S.o. sabi |
|----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 71.2     | 66.1      | 24.3      | 30.2      | 30.1      | 31.0      | 31.2      | 31.2      |
| 63.5     | 28.0      | 30.9      | 32.3      | 33.6      | 31.4      | 31.6      |
| 24.4     | 28.4      | 29.6      | 31.9      | 29.6      | 31.4      |
| 22.3     | 25.3      | 22.3      | 25.6      | 25.0      |
| 65.3     | 50.8      | 54.5      | 52.5      |
| 49.8     | 53.3      | 52.1      |
| 50.3     | 69.7      |
| 53.3     |

3 D. C. Williams, E. J. Katahira, D. J. McGarvey, and R. Croteau, submitted for publication.
targets for selective covalent modification with radiolabeled diisopropyl fluorophosphate and directed mutagenesis studies.

Properties of Recombinant Synthases—Calibrated gel permeation chromatography of the pGEX fusion form of SBS revealed a single peak of activity at an elution volume corresponding to an $M_r \sim 200,000$, indicating that the expressed fusion protein (corresponding to a molecular weight of about $2 \times 96,300$) was a functional dimer. Treatment of the SBS protein with thrombin to remove the glutathione S-transferase fusion tag, followed by re-chromatography, indicated a decrease in molecular weight to approximately 135,000, consistent with the loss of the 27-kDa transferase peptide from each subunit at a calculated molecular weight of 69,300 for the preprotein. Further correction of the molecular mass to account for the transit peptide would yield a dimer of about 120 kDa which corresponds roughly to the native dimer molecular mass of both (+)-bornyl diphosphate synthase and (+)-pinene synthase from sage (6, 12), two enzymes which have never been satisfactorily resolved as distinct species. Although a dimeric quaternary structure is not unique to these two synthases, the vast majority of the monoterpene synthases characterized to date are monomeric (1, 5). The product profile of SBS clone 3CS is qualitatively similar to the combination of both of these enzymes (i.e. (+)-bornyl diphosphate and the (+)-series of α-pinene and related olefins) (6, 12), although the quantitative distributions do not exactly match, and the stereochemistry of the olefin products is anomalous. Thus, as indicated previously, (+)-bornyl diphosphate, (+)-α-pinene, (+)-camphene, and (+)-limonene arise via the same overall cyclization stereochemistry, and these enantiomers are produced exclusively from geranyl diphosphate by the native (+)-bornyl diphosphate and (+)-pinene synthase activities (18, 38–41). The small amounts of (−)-limonene and (−)-camphene formed by the recombinant enzyme are attributed to antipodal cyclizations via abnormal, extended substrate conformations, as the phenomenon has been described previously, especially when using neryl diphosphate (the cis-analog of geranyl diphosphate) as an alternate substrate (40, 64). The geranyl substrate, however, was verified as >99% pure, thereby eliminating this possibility in the present instance and suggesting that loss of stereochemical fidelity (to the extent of 5% of the total product mixture) may be attributed to the presence of the glutathione S-transferase fusion peptide plus transit peptide which may alter substrate binding directly, or indirectly by compromising subunit assembly. To address this question, as well as the correct identity of the SBS protein, will require the detailed assessment of truncated enzymes that more closely resemble the native form.

Gel permeation chromatography of SCS revealed a single peak of activity at an elution volume corresponding to an $M_r$ of 72,000, whereas SSS gave two peaks of activity, an aggregated form eluting in the void volume and a second corresponding to an $M_r$ of 60,000. Both of these molecular weights are significantly lower than those predicted from pGEX expression-based fusion of the glutathione S-transferase (27 kDa) with the respective preproteins (both SCS and SSS – 96 kDa). Thrombin treatment was without influence on the gel permeation chromatographic behavior of these enzymes, indicating the absence of the glutathione S-transferase peptide tag and rationalizing the previously observed inability of the recombinant SCS and SSS enzymes to bind to the glutathione-affinity column. Inspection of the 5′-sequences of the corresponding pGEX constructs showed both to be free of in-frame stop codons that might have permitted polycistrionic translation of the preprotein devoid of the glutathione S-transferase peptide. The apparent truncation was therefore attributed to proteolytic processing of the recombinant SCS and SSS in the E. coli host to proteins that more closely resemble the preprotein forms of the native, monomeric 1,8-cineole synthase (8, 35) and (+)-sabinene synthase (9, 10) of sage. Similar proteolytic processing of a recombinant limonene synthase preprotein from spearmint has been observed previously in this E. coli host (23).

1,8-Cineole synthase has never been satisfactorily separated from the aforementioned (−)-pinene synthase from sage but, in this instance, the product distribution of SCS does not match the product distribution of (−)-pinene synthase, either quantitatively, qualitatively, or in stereochemical terms, since the reactions catalyzed are of the opposite antipodal series (8, 11, 18). However, the product distribution of SCS shows some parallels with that of the recently described cyclase III which produces (+)-α-pinene and (+)-β-pinene (13, 14). Even here, the match is not perfect and the production of anomalous products of the antipodal (−)-series (<6% of total) again suggests that substrate binding interactions may be compromised. To assess the latter possibility, the $K_m$ values for SCS (7.0 μM), SSS (7.4 μM), and SBS (3.0 μM) were determined. These values are likely somewhat high because the recombinant enzymes were not purified sufficiently to remove all contaminating phosphatases that result in some depletion of the substrate geranyl diphosphate. While the calculated $K_m$ values compare reasonably well with the literature values of 1.1 μM (8), 2.0 μM (9), and 2.0 μM (18, 41), respectively, for the corresponding native enzymes, they are sufficiently higher to suggest at least subtle alteration in binding capacity of the recombinant forms.

Although the combination of products generated by the three recombinant enzymes described in this paper represent many of the monoterpene synthases synthesized by extracts of sage oil gland cells, several lines of evidence suggest that more monoterpene synthases remain to be acquired from this species. Thus, clone 3F5 (SUS) has not yet been functionally expressed, one of the two probes generated by PCR has not yet located the corresponding cDNA (one probe matches SUS), and an active site peptide (17) has not yet been matched to a cDNA sequence. The lack of the full set of synthases, coupled to the production of multiple products, including abnormal enantiomers, by the extant recombinant forms, has obscured the assignment of cDNA sequences to their native monoterpene synthase counterparts. This is most vividly illustrated by clone 3C6 which is clearly the recombinant form of (+)-bornyl diphosphate synthase and may also represent the enzyme responsible for the activity previously assigned to (−)-pinene synthase. Nevertheless, it is the great mechanistic diversity of this group of enzymes, originating from a single plant species yet producing a cyclic ether, diphosphate ester, and numerous olefins, that makes this set of catalysts unique and so biochemically appealing. Although the enzymatic generation of multiple products is an unusual and complex phenomenon, it does provide a most convenient and very powerful reporting device for communicating the influence of directed mutagenic change on the cyclization reaction. Defining how the similarities and differences of these enzymes relate to the control of catalytic channeling to the different products can be expected to provide the foundation for a detailed, molecular level understanding of structure-function relationships in the monoterpene synthases.

Acknowledgments—We thank Gerhard Munske of the Washington State University Laboratory for Bioanalysis and Biotechnology for primer synthesis and nucleotide sequencing, and Joyce Tamura-Brown for typing the manuscript.

REFERENCES

1. Croteau, R. (1987) Chem. Rev. 87, 929–954
2. Lewinsdor, E., Gijzen, M., and Croteau, R. (1992) Arch. Biochem. Biophys. 293, 167–173
3. Savage, T. J., Hatch, M. W., and Croteau, R. (1994) J. Biol. Chem. 269, 4012–4020
4. Savage, T. J., Ichii, H., Hume, S. D., Little, D. B., and Croteau, R. (1995) Arch.
