Heterologous Expression and Characterization of a “Pseudomature” Form of Taxadiene Synthase Involved in Paclitaxel (Taxol) Biosynthesis and Evaluation of a Potential Intermediate and Inhibitors of the Multistep Diterpene Cyclization Reaction

David C. Williams,* Mark R. Wildung,* Alan Qingwu Jin,† Dolan Dalal,‡ John S. Oliver,‡ Robert M. Coates,† and Rodney Croteau

*Institute of Biological Chemistry, Washington State University, Pullman, Washington 99164-6340;
†Department of Chemistry, University of Illinois, 600 South Matthews Avenue, Urbana, Illinois 61801; and ‡Department of Chemistry, Box H, Brown University, Providence, Rhode Island 02912

Received February 23, 2000, and in revised form April 4, 2000

The diterpene cyclase taxadiene synthase from yew (Taxus) species transforms geranylgeranyl diphosphate to taxa-4(5),11(12)-diene as the first committed step in the biosynthesis of the anti-cancer drug Taxol. Taxadiene synthase is translated as a preprotein bearing an N-terminal targeting sequence for localization to and processing in the plastids. Overexpression of the full-length preprotein in Escherichia coli and purification are compromised by host codon usage, inclusion body formation, and association with host chaperones, and the preprotein is catalytically impaired. Since the transit peptide-mature enzyme cleavage site could not be determined directly, a series of N-terminally truncated enzymes was created by expression of the corresponding cDNAs from a suitable vector, and each was purified and kinetically evaluated. Deletion of up to 79 residues yielded functional protein; however, deletion of 93 or more amino acids resulted in complete elimination of activity, implying a structural or catalytic role for the amino terminus. The pseudomature form of taxadiene synthase having 60 amino acids deleted from the preprotein was found to be superior with respect to level of expression, ease of purification, solubility, stability, and catalytic activity with kinetics comparable to the native enzyme. In addition to the major product, taxa-4(5),11(12)-diene (94%), this enzyme produces a small amount of the isomeric taxa-4(20),11(12)-diene (~5%), and a product tentatively identified as verticillene (~1%). Isotopically sensitive branching experiments utilizing [4-2H]geranylgeranyl diphosphate confirmed that the two taxadiene isomers, and a third (taxa-3(4),11(12)-diene), are derived from the same intermediate taxenyl C4-carbocation. These results, along with the failure of the enzyme to utilize 2,7-cyclogeranylgeranyl diphosphate as an alternate substrate, indicate that the reaction proceeds by initial ionization of the diphosphate ester and macrocyclization to the verticillyl intermediate, followed by a secondary cyclization to the taxenyl cation and deprotonation (i.e., formation of the A-ring prior to B/C-ring closure). Two potential mechanism-based inhibitors were tested with recombinant taxadiene synthase but neither provided time-dependent inactivation nor afforded more than modest competitive inhibition. 

Key Words: Taxol biosynthesis; paclitaxel; taxadiene synthase; Taxus; yew.

Taxadiene synthase from Taxus (yew) species catalyzes the first committed step in the biosynthesis of the anti-cancer drug Taxol and related taxane diterpenes.
noids by the cyclization of the universal precursor of diterpenes, geranylgeranyl diphosphate, to the olefinic product (Fig. 1) (1). This enzyme catalyzes a slow, but apparently not rate limiting, step in the Taxol biosynthetic pathway (2). Taxadiene synthase has been isolated from yew saplings (3) and cell cultures (2), and the enzyme exhibits properties typical of the few diterpene synthases from gymnosperms and angiosperms thus far defined (4). Initial assessment of the mechanism of taxadiene synthase (5) indicates that the reaction involves the ionization and cyclization of geranylgeranyl diphosphate to a transient verticillyl cation intermediate (1), with intramolecular transfer of the C11 proton to C7 to initiate transannular B/C-ring closure to the taxenyl cation (2), followed by deprotonation at C5 to yield the taxa-4(5),11(12)-diene product; however, few other details of the stereochemical mechanism or of the structure of this novel enzyme are known.

A cDNA encoding the taxadiene synthase from *Taxus brevifolia* has been obtained by a homology-based PCR cloning method (6). Like the monoterpene synthases and other diterpene synthases (7), taxadiene synthase is translated as a preprotein bearing an N-terminal targeting sequence for localization to plastids where proteolytic processing of the transit peptide occurs and where the mature form of the enzyme functions in the first step of Taxol biosynthesis. The preprotein forms of the terpene synthases are often sufficiently active to permit identification of clones by functional heterologous expression from standard cloning vectors (7), and the taxadiene synthase was confirmed by this means (6). However, in high level expression systems, the presence of the transit peptide of the terpene synthases often contributes to the formation of inclusion bodies and can compromise purification of the soluble form(s) by promoting aggregation and/or the tight association with host chaperones and other proteins (8). Additionally, the preproteins may be catalytically impaired relative to their native counterparts. Soluble taxadiene synthase has been overexpressed recently as a His-tagged, thioredoxin fusion of the preprotein (9). The enzyme is catalytically capable when expressed in this form; however, such preprotein fusions are unsuitable for detailed mechanistic and crystallographic study.

Deciphering the transit peptide-mature protein cleavage junction of terpene synthase preproteins from sequence information alone is not possible because of the considerable variation in primary sequence of these targeting peptides (10–13), and studies with the native enzymes to define this feature directly have indicated that most are N-terminally blocked (8, 14). Moreover, detailed mass spectrometric analysis of a prototype terpene synthase, limonene synthase, indicated that plastidial proteolytic processing to the native form in this case was quite imprecise (8). Therefore, to prepare a "pseudomature" form of such a terpene synthase requires the empirical construction and testing of a series of truncated versions of the target protein. In

---

**FIG. 1.** Cyclization of geranylgeranyl diphosphate to taxa-4(5),11(12)-diene and subsequent elaboration to Taxol. Ac, acetyl group; Bz, benzoyl group; OPP, diphosphate moiety.
this paper, we describe the construction, heterologous expression and purification of a series of truncated forms of the preprotein, report on the kinetic properties, solubility and stability of these “pseudomature” forms of taxadiene synthase, and discuss the mechanistic implications of several new findings revealed by studies with the recombinant enzyme.

**MATERIALS AND METHODS**

**Substrates, standards, and inhibitors.** (E,E,E)-[1-3H]Geranylgeranyl diphosphate (120 Ci/mol) was prepared as described previously (15). (−)-Taxa-(4)5,11(12)-diene and (−)-taxa-4(20),11(12)-diene (16) were gifts from Robert M. Williams, Colorado State University (Fort Collins, CO). All other biochemicals and reagents were purchased from Sigma Chemical Co. or Aldrich Chemical Co., unless otherwise noted.

(−)-2,7-Cyclogeranylgeranyl diphosphate (3, Fig. 2) and its deuterium labeled form [1,1-2H2]3 were prepared from 2,7-cyclogeranylgeraniol (17) by conversion to the respective chlorides and displacement with pyrophosphate trianion. Reduction of methyl 2,7-cyclogeranylgeranoate with LiAl2H4 (17) afforded the corresponding labeled alcohol (305 mg, 77%, 99% d 2 by 1H NMR integration) after purification by flash chromatography (18) on silica gel. The 1H and 13C NMR spectra (CDCl3, 400 MHz) were identical to those of the unlabeled alcohol except for the following: δH 4.07 and 4.18 (CH2OH) missing; δC 58.0 (quintet, J = 21.7 Hz, CD2OH). Conversions to the unlabeled and labeled allylic chlorides (75 mg, 95%; 97 mg, 92%; respectively) were carried out by the Meyers procedure (19), using 6 equiv CH3SO2Cl, 5 equiv LiCl, and 8.8 equiv of collidine (DMF, 0°C, 65 min). The unstable chlorides, contaminated by about 10% of the methyleneencyclohexene elimination product, were characterized by 1H NMR, and 13C NMR spectra, and used without purification. Selected NMR data (benzene-d6, 400 MHz) for 2,7-cyclogeranylgeranyl chloride: δ 1.04, 1.56, 1.60 (3s, 3H, CH3), 3.83 and 3.93 (ABq, 2H, J = 11.4 CH2Cl), 5.24 (m, 2H, vinyl H); δ 41.2 (CH2Cl). Reaction of the chlorides with (Bu4N)3HP2O7 (1.08 equiv) in CH3CN (1 ml) containing powdered 3 Å molecular sieve was carried out for 48 h at room temp. according to literature procedures (20, 21). Purification by ion exchange chromatography, selective precipitation, and cellulose chromatography afforded 3 (16 mg, 26%) and [1,1-2H2]3 (7.9 mg, 10%) as white solids, which were characterized by TLC, 1H, and 31P NMR spectra, and negative ion FAB-MS. Data for 3: 1H NMR (D2O 0.06% (w/v) EDTA 1 concentrated NH4OH to pH 9, 400 MHz) δ 0.88 (s, 3 H, CH3), 1.41 (s, 6 H, 2 × CH3), 1.48 (s, 3 H, CH3), 1.59 (s, 3 H, CH3), 1.04–1.20 (m, 2 H, CH2), 1.21–1.34 (m, 1 H), 1.42–1.54 (m, 1 H), 1.56–1.68 (m, 1 H), 1.68–1.94 (m, 8 H, CH2), 4.25 and 4.35 (ABq, 2 H, J = 9.4 Hz, CH2O), 4.63 (s, NH4), 4.89 (2 H, vinyl H); 31P NMR (D2O 0.06% (w/v) EDTA + concentrated NH4OH to pH 9, 162 MHz) δ 2.6 (d, J = 20.8 Hz), 10.2 (d, J = 20.8 Hz). Low resolution negative ion FAB-MS: m/z 449.22 (M−). Spectra of [1,1-2H2]3 were identical to those of 3 except for the following: 1H NMR: δ 4.25 and 4.35 (CH2OPP) absent; negative ion FAB MS: low resolution: m/z 451.2 (M−); high resolution Calcd for C20H33D2O7P2, 451.198359; Found, 451.197500.

For preparation of the 13-cyclopropylidene analog of geranylgeranyl diphosphate (4, Fig. 2), geranylgeraniol (5 mmol) was first converted to the acetate ester (22) (91.5% yield following flash chromatography (18)) and then to the 14,15-epoxide by treatment with 3-chloroperoxybenzoic acid in CH2Cl2 (23). The epoxide (2 mmol) was transformed to 14,15-dihydroxygeranylgeranyl acetate (25% yield).
by treatment with HClO₄, and the dial then cleaved with excess peroxide to afford the geranylgeranyl alcohol. Purification was by flash chromatography (18). Coupling of the terminal aldehyde with methyltriphenylphosphonium bromide (29) provided 16-methylidenegeranylgeranyl acetate in 37% yield following flash chromatography (18). Purification, and the ester was hydrolyzed as before (22), converted to the corresponding chloride (30), and then diphosphorylated with ATP and MgSO₄-silica gel column chromatography step was dispensed with (this step had no influence on the assay), and the pentane extract was passed through a short column of MgSO₄-silica gel prior to aliquot counting by LSC and GC-MS analysis for verification of taxane olefin production. For convenience in most assays, the taxadiene synthase eluted at a concentration of 50 mM MgCl₂ (pH 6.8). The taxadiene synthase eluted from hydroxyapatite (the different truncations elute at different phosphate concentrations on this column) was applied directly to a Pharmacia FPLC 10/10 column packed with DEAE Sepharose (Sigma, St. Louis, MO) that was previously equilibrated with buffer A until the baseline stabilized (UV monitoring at A₅₄₀, and then eluted with a 150 mM linear gradient from 0 to 300 mM NaCl (buffer B plus 1 M NaCl). Taxadiene synthase eluted at −150 mM NaCl. The enzyme preparation was next applied directly to a Pharmacia FPLC 10/10 column packed with ceramic hydroxyapatite (Bio-Rad) that was previously equilibrated with buffer A. The column was washed with buffer A until the baseline stabilized and then eluted with a 100 mM linear gradient from 0 to 100 mM potassium phosphate (pH 6.8). The taxadiene synthase eluted from hydroxyapatite (the different truncations elute at different phosphate concentrations on this column) was applied directly to a Pharmacia FPLC 10/10 column packed with DEAE Sepharose (Sigma, St. Louis, MO) that was previously equilibrated with buffer A. This weak anion-exchange column was washed with buffer A until the baseline stabilized and then eluted with a 100 mM linear gradient from 0 to 1 M NaCl in buffer B. Taxadiene synthase eluted at −350 mM NaCl, and was desalted into buffer A using a Centrprep YM50 centrifugal concentrator (Millipore, Bedford, MA) prior to the standard assay. The course of the purification was monitored throughout by enzyme assay and SDS–PAGE (33). Protein concentration was determined by the Bradford assay using a standard curve calibrated with purified, recombinant limonene synthase (34).

Taxadiene synthase assay and enzyme characterization. The standard assay was performed in 25 mM Hepes (pH 8.0) containing 10% (v/v) glycerol in the presence of 1 mM MgCl₂ and [1-H]-geranylgeranyl diphosphate as substrate as previously described (3). The reaction products were extracted into pentane, and the pentane extract was passed through a short column of MgSO₄-silica gel prior to aliquot counting by LSC and GC-MS analysis for verification of taxane olefin production. For convenience in most assays, the MgSO₄-silica gel column chromatography step was dispensed with (this step had no influence on the assay), and the pentane extract was concentrated and analyzed directly by GC-MS via external calibration with authentic taxadiene. The latter assay protocol was employed to evaluate the conversion of the alternate substrate, 2,7-cyclogeranylgeranyl diphosphate (3), and the catalytic turnover of the two potential mechanism-based inhibitors (4 and 7). GC-MS analysis was performed on a Hewlett-Packard 6890–7293 MSD system using a 30 m fused silica capillary 5MS column. Samples were separated using cool-on-column injection (40°C) with temperature programming to 320°C at 20°C/min and He as carrier gas; 70 eV spectra were recorded.

Conversion of the substrate analog 2,7-cyclogeranylgeranyl diphosphate (3) was evaluated at a concentration of 50 μM (i.e., saturation with the normal geranylgeranyl substrate); the two mechanism-based inhibitors (4 and 7) were similarly tested. To examine the time-dependent inactivation of taxadiene synthase, the enzyme was incubated at 31°C with 50 μM concentrations of either the
cyclopropylidene analog (4) or the vinyl analog (7). At 30-min intervals, aliquots of the preparations were removed, concentrated, re-suspended, and re-concentrated twice in buffer A using a Centriprep concentrator to remove residual inhibitor, and then assayed for activity with [1-^3H]geranylgeranyl diphosphate as substrate using the standard protocol. Untreated preparations that were similarly processed served as the controls. For kinetic analysis, standard assays were performed at ten substrate concentrations ranging from 0.5 to 50 μM [1-^3H]geranylgeranyl diphosphate. After correction for very minor solvolytic background rates (buffer controls), the averaged values for triplicate analyses at each substrate concentration were plotted and evaluated using the Enzyme Kinetics Program (Trinity Software, Inc.).

The stability of the enzyme in buffer A, in the presence of 50 μM substrate, was determined at room temperature with purified preparations overlayed with hexane to exclude air and continuously extract the olefin product, the formation of which was conveniently monitored by aliquot counting. The solubility of taxadiene synthase was determined in buffer A at room temp. by sequential ultrafiltration (using a Centriprep YM50 concentrator) and centrifugation at 13,000g to remove aggregated protein, while monitoring protein concentration by dye binding assay (33) and UV absorption at A_280.

RESULTS AND DISCUSSION

Truncation of Taxadiene Synthase

The cDNA for taxadiene synthase encodes a preprotein with a calculated size of 98.3 kDa (862 amino acids) (6), compared with a value of about 80 kDa determined for the native enzyme by SDS–PAGE and gel permeation chromatography (3). Plastidial processing of the preprotein appears to result in removal of the targeting peptide and N-terminal modification which blocks Edman degradative sequencing (6). Thus, it is not possible to determine the proteolytic cleavage site and the exact size of the native enzyme by this means, in order to construct a recombinant mature protein by precise truncation of the cDNA. A similar problem was recently addressed in the case of the monoterpene cyclohexane–4S-limonene synthase (35) by systematic truncation of the cDNA, and expression and kinetic evaluation of the resulting enzymes (8). A very tractable “pseudomature” form of limonene synthase was prepared by truncation immediately upstream of a pair of tandem arginine residues (R58R59), a sequence element that was ultimately shown to participate in the preliminary substrate isomerization step that is specifically and universally required for monoterpene cyclization reactions (8). Alignment of the deduced taxadiene synthase sequence with those of several monoterpene synthases and other terpenoid synthases (7), in an attempt to define the corresponding cleavage site, was not fruitful. The taxadiene synthase sequence contains an unusual insertional element of about 200 amino acid residues (6, 7) that results in placement of this corresponding site some 300 amino acid residues downstream of the starting methionine to afford a protein of about 65 kDa, far smaller than the native enzyme. Thus, in producing a recombinant “pseudomature” form of taxadiene synthase for detailed mechanistic and structural studies, there was little choice but to construct, express, and test a series of truncations surrounding the approximate proteolytic processing site based upon rather imprecise predictive methods (10–13).

For this purpose, the full-length preprotein and truncations corresponding to removal of 60, 79, 93, 113, and 126 residues from the N-terminus (designated M1, M60, M79, M93, M113, and M126, respectively) were prepared in pSBET (31), a vector that has proved to be highly suitable for heterologous expression in E. coli of plant genes that specify arginine by rare codon usage. Low temperature, long duration expression from this vector in E. coli BL21(DE3) (20°C, 14 h) yielded functional, soluble protein in the case of M1, M60, and M79 (confirmed by assay); SDS–PAGE of the soluble protein expressed from M93, M113, and M126 verified the presence of the corresponding recombinant, but inactive, forms of taxadiene synthase (data not shown).

Each of the functional taxadiene synthases was purified by the combination of hydroxyapatite chromatography (two steps) and anion-exchange chromatography (two steps) to provide essentially homogenous preparations (>99% pure) of the M60 and M79 truncations by SDS–PAGE (data not shown). Both of these proteins were expressed to levels of 30–40% of total bacterial soluble protein, with little formation of inclusion bodies. Minor adjustments in the expression and purification protocol have now allowed the routine preparation of homogeneous forms of both enzymes in yields of 30–40 mg/L (M60) and 10–20 mg/L (M79) of bacterial culture. Because of the instability of the full-length preprotein (M1) and the tendency to form inclusion bodies and otherwise aggregate, this form of taxadiene synthase (expressed to levels of about 10% of total bacterial soluble protein) could be brought to only about 50% purity by the same fractionation protocol; however, this was sufficient for accurate kinetic characterization of the enzyme. Kinetic evaluation of each functional form of taxadiene synthase by several graphical plotting procedures revealed the truncated versions, M60 and M79, to be superior to the preprotein M1 and to the previously expressed thioredoxin-fusion of the preprotein (9), and comparable to the native enzyme based on rough estimates of kinetic constants from previous studies (2, 3) (Table I). The observed K_m for the recombinant enzymes is somewhat higher than that estimated for the native enzyme (Table I) but still in the low μM range typical for this class of enzymes (4, 14, 36, 37). Given the turnover numbers of the M60 and M79 enzymes, and the evaluation of approximate rates and protein contents of the crude bacterial extracts, it is apparent that the “inactive” truncations of taxadiene synthase, i.e., M93, M113, and M126, could not have possessed k_cat values in ex-
The activity was negligible. The kinetic constants are averages of triplicate analyses with SE less than 10% of the indicated value. Kinetic Constants for the Native Enzyme, Full-Length Preprotein, Thioredoxin Fusion, and Truncated Versions of Taxadiene Synthase

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$K_m^a$ (μM)</th>
<th>$k_{cat}^b$ (s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native enzyme (ref. 3)</td>
<td>3.0</td>
<td>0.0106</td>
</tr>
<tr>
<td>Thioredoxin fusion (ref. 9)</td>
<td>3.1</td>
<td>0.0025</td>
</tr>
<tr>
<td>M1 preprotein</td>
<td>10.5</td>
<td>0.0024</td>
</tr>
<tr>
<td>M60</td>
<td>15.9</td>
<td>0.0102</td>
</tr>
<tr>
<td>M79</td>
<td>14.0</td>
<td>0.0008</td>
</tr>
<tr>
<td>M93</td>
<td>&lt;0.00001</td>
<td>&lt;0.00001</td>
</tr>
<tr>
<td>M113</td>
<td>&lt;0.00001</td>
<td>&lt;0.00001</td>
</tr>
<tr>
<td>M126</td>
<td>&lt;0.00001</td>
<td>&lt;0.00001</td>
</tr>
</tbody>
</table>

* No entry indicates a $K_m$ value could not be determined because the activity was negligible.

* Kinetic constants are averages of triplicate analyses with SE less than 10% of the indicated value.

The minor diterpene olefin products of taxadiene synthase were subsequently evaluated for stability and solubility at room temperature as a prelude to crystallographic studies. The M60 form, at 10 mg/ml in 25 mM Mops buffer (pH 8.0), was stable at room temperature for at least 10 days (with less than 30% loss of activity) when maintained in the presence of substrate (50 μM) and Mg$^{2+}$ cofactor (1 mM) and with the continuous removal of product by the use of a hexane overlay. Under these conditions, the M79 version was much less stable, losing about 70% activity over a 2-day period at room temperature. The M60 form of taxadiene synthase, at room temperature in a buffer consisting of 25 mM Mops (pH 6.8), containing 10% (v/v) glycerol and 1 mM DTT, could be concentrated to at least 20 mg/ml without precipitation. The M79 form of the enzyme underwent visible aggregation under these conditions when the protein concentration exceeded 5 mg/ml. Thus, based on expression yield, ease of purification, kinetic constants, stability, and solubility, the M60 truncation provided a highly useful "pseudomature" form of taxadiene synthase.

Product Characterization

Each construct which yielded a functional cyclase (M1, M60, M79) was evaluated for product formation (olefins and oxygenated diterpenes) by coupled GC-MS analysis of pentane extracts of the corresponding large-scale as-

In this case, the use of MgSO$_4$-silica gel was avoided and cool-on-column injection was employed to minimize the possibility of acid-catalyzed or thermal rearrangement of the product(s). Each of the functional enzymes, M1, M60 and M79, was shown to generate a qualitatively and quantitatively identical product set (illustrated for the M60 enzyme in Fig. 3). The major product (retention time 10.71 min of Fig. 3A) was confirmed as taxa-4(5),11(12)-diene (94.0%), as expected, by comparison of retention time and mass spectrum (Fig. 3B) to those of the authentic standard prepared by total synthesis (16). One minor product (retention time 10.50 minutes of Fig. 3A) was similarly confirmed as the isomeric taxa-4(20),11(12)-diene (4.8%) by comparison of retention time and mass spectrum (Fig. 3C) to the authentic standard (16). The second, very minor product (retention time 10.08 of Fig. 3A) was tentatively identified as verticillene (1.2%, presumably the 3(4),7(8),11(12)-isomer derived from intermediate 1 of Fig. 1) on the basis of an authentic compound (spectrum kindly provided by Professor G. Pattenden, University of Nottingham, UK).

Since it was independently demonstrated that both the 4(5),11(12)- and 4(20),11(12)-isomers of taxadiene were entirely stable under the conditions of the analysis (5), it can be assumed that the minor olefins are, in fact, authentic products of the taxadiene synthase reaction formed by alternative deprotonation steps. Thus, small amounts of verticillene could arise by direct deprotonation of the verticillyl C12-carbocation (1 of Fig. 1), which consequently aborts the normal intramolecular hydrogen transfer from C11 to C7 to initiate B/C-ring closure to the taxane skeleton (Fig. 1). The formation of only traces of verticillene is consistent with preliminary mechanistic studies with taxadiene synthase (5) that have demonstrated the internal transfer of the hydrogen atom to occur with high efficiency (>90%) and, thus, to imply tight binding of the verticillyl intermediate. In the course of the principal reaction route, the intramolecular migration and ring closure steps provide the taxenyl C4-carbocation (2), which offers three possibilities for deprotonation to terminate the cyclization cascade. In approximately 94% of the reaction cycles, proton loss occurs from C5 to yield the major product taxa-4(5),11(12)-diene. In the infrequent deprotonation occurs from the C20 methyl group (~5%) to give the 4(20)-isomer. Proton loss from the bridgehead methine (C3) to give the 3(4)-isomer was not observed (i.e., trace levels of product).

The minor diterpene olefin products of taxadiene synthase documented here were not observed previously in studies with the native enzyme from Taxus stems (3) or cell cultures (2) because too little material was available for detailed analysis compared to the amounts now accessible via the recombinant forms. The analysis of large-
scale, crude preparations of taxadiene synthase from induced T. cuspidata suspension cell cultures (38) has since verified the biosynthesis of the identical olefin product set by this native enzyme. Analysis of the oxygenated diterpenes produced from geranylgeranyl diphosphate by the recombinant, purified forms of taxadiene synthase revealed the presence of only traces of geranyllinalool, geranylnerol and geranylgeraniol produced solvolytically in the course of the assay.

To examine in greater detail the alternative deprotonations catalyzed by taxadiene synthase, we employed the deuterium-labeled substrate (4R)-[4-2H1]geranylgeranyl diphosphate (84 mol %) to exploit the phenomenon of isotopically sensitive branching (39, 40). Isotopically sensitive branching experiments have been utilized extensively to evaluate monoterpene cyclization reactions (41–45). In the present case, the deuterium-dependent slowing of the C5-deprotonation of a common intermediate (the taxenyl cation 2) should promote the alternative means of terminating the reaction. Comparison of the product mixture derived from the deuterated substrate (Fig. 4A) relative to that generated from the normal substrate (Fig. 3A) revealed a decreased proportion of taxa-4(5),11(12)-diene (77.6%), with an increase in the formation of the 4(20),11(12)-isomer (11.1%) and the production of the tentatively identified 3(4),11(12)-isomer (Fig. 4B, 9.2%) (which is produced at only trace levels with the normal substrate) but without change in the formation of verticillene (1.1%). GC-MS analysis of the products was entirely consistent with the assumed mechanism of taxadiene synthase (1, 5), which predicts loss of the deuterium in formation of taxa-4(5),11(12)-diene but retention of the deuterium in the other olefins. These results confirm that all of the taxadiene isomers are derived from the taxenyl C4-carbocation intermediate (2) and establish that the 5β-proton (Hre) is eliminated in forming the 4,5-double bond of the major product. That the rate of formation of verticillene is not altered by the deuterium substitution is expected, since verticillene is formed prior to the isotopically sensitive branch that comprises the terminating deprotonation step of the reaction. A rough estimate of the kinetic isotope effect for the C5-deprotonation gives a kH/kD of

The synthesis of this stereospecifically deuterium-labeled compound will be described elsewhere in the context of additional mechanistic studies with this enzyme.

FIG. 3. Capillary GC-MS analysis of the diterpene olefin products of the truncated (M60) recombinant taxadiene synthase. The GC profile (A, total ion current) and mass spectra of the olefins generated from geranylgeranyl diphosphate are illustrated for taxa-4(5),11(12)-diene (B), taxa-4(20),11(12)-diene (C), and a product tentatively identified as verticillene (D). Analytical conditions are described under Materials and Methods. Identification of the taxadiene isomers is based on comparison of retention times and mass spectra to the authentic standards.
about 2. The formation of multiple products is a common feature of terpenoid cyclization reactions (46, 47). For example, some sesquiterpene synthases produce in excess of 30 different olefin products (48). The fidelity of taxadiene synthase in the production of largely a single product (taxa-4(5),11(12)-diene at 94%) is, in fact, notable.

The truncation studies indicate that the N-terminus of taxadiene synthase is essential for catalysis (i.e., there is complete loss of function by truncation between aa79 and aa93 (see Table I)), but does not significantly influence product outcome in the distribution of olefins (at least for M1, and truncations M60 and M79). The M60, M79, M93, M113, and M126 versions of taxadiene synthase were intended to represent forms of the preenzyme from which the transit peptide was deleted and for which, contrary to observation, the minimum mechanistic consequences were anticipated. Structural studies with the monoterpen cyclase (−)-limonene synthase (8) and the sesquiterpene cyclase 5-epi-aristolochene synthase (49) indicate the active sites of these enzymes to be largely derived from the C-terminal domain; however, in both cases, the N-terminus of the mature enzyme appears to play a functional role. The structural and/or mechanistic rationale for the abrupt loss of activity at truncation M93 is presently unclear.

Inhibitors and Alternative Substrates

The conversion of geranylgeranyl diphosphate to the taxane skeleton by initial macrocyclization and bridging to form a verticillyl intermediate followed by proton-induced closure of the B/C-rings (Fig. 1) was first proposed in 1966 (50). The evidence, thus far, based largely on the cyclization to taxa-4(5),11(12)-diene catalyzed by taxadiene synthase, supports this mechanistic proposal (1–3). However, exogenous (±)-verticillene is not detectably incorporated into taxadiene by the enzyme (5), nor do verticillene and its various epoxide derivatives undergo such transannular cyclizations in chemical model reactions (51). Nevertheless, the apparent production of trace levels of verticillene in the enzymatic reaction is consistent with the proposed mechanism, and with the tight binding of such a transient verticillyl intermediate required to enforce a conformation conducive to the subsequent cyclization to the B/C-ring system and to effect the efficient intramolecular hydrogen transfer step (5) (Fig. 1).

Another plausible mechanism for taxadiene formation could involve formation of the C-ring first, with the generation of 2,7-cyclogeranylgeranyl diphosphate (3, Fig. 2) as a transient intermediate instead of the verticillyl structure. Completion of the cyclization could then occur by similar intramolecular alkylation of the terminal double bond in the sidechain followed by bridging and intramolecular C11 to C3 proton transfer (based on available deuterium labeling studies and MS analysis of the derived taxadiene product (5), this scenario is indistinguishable from the previously proposed hydrogen transfer from C11 to C7), and final proton elimination from C5 to form taxa-4(5),11(12)-diene. Structural precedent for the proposed cyclogeranylgeranyl-type intermediate (3) in this alternative route is provided by the natural occurrence of α- and β-cyclogeraniols (52), picrocrocin (53), and safranal and related monoterpenes (53–55) in various plant sources. Furthermore, similar proton-initiated cyclizations of geranylgeranyl diphosphate in the olefinic chain distal to the diphosphate ester group are well-established in the biosynthesis of the diterpenoid resin acids and gibberelins (4, 56).

(±)-2,7-Cycd[1-2H2]geranylgeranyl diphosphate was therefore prepared and tested at saturating concentrations (50 μM, with 1 mM Mg2+) as a substrate for taxadiene formation with the purified M60 version of the synthase under established assay conditions. No detectable taxadiene, or other diterpene olefin, was produced in the incubation of cyclogeranylgeranyl
diphosphate with the synthase under conditions which efficiently yielded ample product (>100 pmol) with the normal geranylgeranyl substrate; these results argue against the alternative pathway to taxadiene involving 2,7-cyclogeranylgeranyl diphosphate as an intermediate. However, negative results regarding the intervention and conversion of potential intermediates of the reaction (e.g., both verticillene and cyclogeranylgeranyl diphosphate) do not eliminate the possible involvement of the corresponding, tightly bound, transient intermediate that is not accessible to, or exchangeable with, exogenous material. Thus, on binding the geranylgeranyl diphosphate substrate, taxadiene synthase could undergo conformational alteration as a means of shielding subsequently formed carbocation intermediates from water. Such conformational change could as a consequence also prevent the egress of any stable intermediate from the active site, and disallow the productive binding from solution of such an exogenous intermediate (to either conformational form). Similar stable, but tightly bound, intermediates, especially olefinic intermediates, are thought to participate in related cyclizations in the monoterpene (46, 57), sesquiterpene (37, 58) and diterpene (15, 47) series, yet direct evidence for the participation of such neutral intermediates of the reaction cycle has been similarly difficult to obtain (59). Conversely, the participation of certain diphosphorylated intermediates in terpenoid cyclization reactions (linalyl diphosphate in the monoterpene series (36, 46), nerolidyl diphosphate in the sesquiterpene series (58), and copalyl diphosphate in the diterpene series (4, 47)) has been amply demonstrated by direct testing of these “alternate” substrates. These latter results support the argument that the failure of taxadiene synthase to utilize cyclogeranylgeranyl diphosphate indicates that this cyclic isomer of the normal substrate is not an intermediate of the cyclization reaction.

In a final set of experiments, two potential, mechanism-based inhibitors (4 and 7) of the taxadiene cyclization reaction were tested with the purified M60 version of the enzyme. These substrate analogs are designed to reposition and/or stabilize positive charge in intermediate(s) generated in the course of the normal cyclization reaction (Fig. 2), with the anticipation that alteration in charge configuration of the carbocation will result in alkylation of a proximate basic residue of the synthase. This general approach has been utilized to inactivate (and specifically label) certain monoterpene synthases (60, 61), and has been widely, and successfully, employed to probe the structure and mechanism of triterpene synthases (62).

In the present instance, neither the cyclopropylidene analog (4) nor the vinyl homolog (7) at a concentration of 100 μM exhibited time-dependent inactivation of taxadiene synthase when incubated under standard conditions (at 1 mM Mg^2+) with the M60 version of the enzyme. However, when tested as substrates under standard assay conditions, the cyclopropylidene analog (4) did yield an olefinic product (with parent ion at m/z 270 and base peak at m/z 122) consistent with formation of the cyclopropylidene analog of taxadiene (6) at a rate corresponding to about 30% of that observed for total olefin formation with the normal geranylgeranyl substrate. The vinyl homolog (7) was not utilized, yielding no detectable olefinic (e.g., 9) or oxygenated diterpenoid products when incubated with taxadiene synthase. Thus, it appeared that the enzyme escaped the possibility of alkylation by the cyclopropylidene analog (4) by conduct of the normal reaction, albeit at a somewhat compromised rate, to produce the corresponding 16,17-cyclotaxadiene derivative (6). In the case of the vinyl-substituted homolog (7), it appeared that this substrate, modified by the additional methylenedioxy, did not undergo normal, productive binding to initiate the ionization and macrocyclization steps required to unmask the novel, potentially reactive, carbocation (8 of Fig. 2). A preliminary evaluation of all of the substrate analogs (3, 4, and 7) as presumptive competitive inhibitors of the normal cyclization reaction of [1-^3H]geranylgeranyl diphosphate (at 30 μM) revealed all three to be relatively weak inhibitors of taxadiene synthase (M60 version), with approximate Ks values of 90 μM (3), 120 μM (4) and 100 μM (7), respectively. These results suggest that none of the substrate analogs is readily recognized by taxadiene synthase, which exhibits a relatively low Km value of 15.0 μM for the normal geranylgeranyl substrate that the proposed reaction mechanism suggests must be precisely aligned at the active site for the cyclization to occur (5). Further studies to understand active site structure and the mechanism of action of this unusual catalyst are in progress.9

ACKNOWLEDGMENTS
We thank C. Sanchez, D. Pouchnik, and G. Munske for nucleotide sequencing and primer synthesis, B. Long for assistance with the protein expression, and Joyce Tamura for typing the manuscript.

REFERENCES