Chrysanthemyl diphosphate synthase: Isolation of the gene and characterization of the recombinant non-head-to-tail monoterpene synthase from *Chrysanthemum cinerariaefolium*

Susan B. Rivera*, Bradley D. Swedlund†‡, Gretchen J. King†§, Russell N. Bell†‡, Charles E. Hussey, Jr.†, Donna M. Shattuck-Eidens†‡, Wislawa M. Wrobel†, Galen D. Peiser†, and C. Dale Poulter*¶

*Department of Chemistry, University of Utah, Salt Lake City, UT 84112; and †Agridyne Technologies, Salt Lake City, UT 84112

Edited by Rodney B. Croteau, Washington State University, Pullman, WA, and approved January 18, 2001 (received for review November 15, 2000)

Chrysanthemyl diphosphate synthase (CPPase) catalyzes the condensation of two molecules of dimethylallyl diphosphate to produce chrysanthemyl diphosphate (CPP), a monoterpene with a non-head-to-tail or irregular c1′-2-3 linkage between isoprenoid units. Irregular monoterpenes are common in *Chrysanthemum cinerariaefolium* and related members of the Asteraceae family. In *C. cinerariaefolium,* CPP is an intermediate in the biosynthesis of the pyrethrin ester insecticides. CPPase was purified from immature chrysanthemum flowers, and the N terminus of the protein was sequenced. A *C. cinerariaefolium* λ cDNA library was screened by using degenerate oligonucleotide probes based on the amino acid sequence to identify a CPPase clone that encoded a 45-kDa preprotein. The first 50 aa of the ORF constitute a putative plastidial targeting sequence. Recombinant CPPase bearing an N-terminal polyhistidine affinity tag in place of the targeting sequence was purified to homogeneity from an overproducing *Escherichia coli* strain by Ni^{2+} chromatography. Incubation of recombinant CPPase with dimethylallyl diphosphate produced CPP. The diphosphate ester was hydrolyzed by alkaline phosphatase, and the resulting monoterpene alcohol was analyzed by GC/MS to confirm its structure. The amino acid sequence of CPPase aligns closely with that of the chain elongation prenyltransferase farnesyl diphosphate synthase rather than squalene synthase or phytoene synthase, which catalyze c1′-2-3 cyclopropanation reactions similar to the CPPase reaction.

Over 30,000 isoprenoid compounds have now been identified (1). The vast majority of these compounds have "regular" 1′-4, or head-to-tail, linkages between isoprenoid units (2). The pathway for biosynthesis of regular isoprenoids from the basic five-carbon precursors isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) is well established. More complex hydrocarbon skeletons are typically generated by prenyl transfer reactions, where successive molecules of IPP are added to a growing allylic diphosphate chain through 1′-4 linkages (3, 4). "Irregular," or non-head-to-tail, isoprenoids are encountered less frequently. The most prominent examples are the 1′-1 isoprenoid compounds squalene (SQ) and phytoene (PT), which are precursors for biosynthesis of sterols and carotenoids (2). However, the greatest variety of non-head-to-tail structures are found in monoterpenes from closely related species of sagebrush indigenous to the Great Basin of the western United States (2). These plants belong to the Asteraceae family and produce irregular monoterpenes with 1′-1 (5), 1′-2 (5), 1′-3 (6), c1′-2-3 (7, 8), c2–1′-3 (6, 9–12), and c1′-1-2 (6) skeletons. Epstein and Poulter (9) proposed a scheme for the biosynthesis of irregular monoterpenes in which the first step is the condensation of two molecules of DMAPP to give chrysanthemyl diphosphate (CPP) with a c1′-2-3 structure as shown in Fig. 1.

C_{30} and C_{40} cyclopropylcarbinyl diphosphates with structures similar to CPP are well-established intermediates in the 1′-1 coupling reaction that produce SQ and PT (2). Squalene synthase (SQase), is the first pathway-specific enzyme in cholesterol biosynthesis. The enzyme catalyzes two reactions: a c1′-2-3 condensation of two molecules of farnesyl diphosphate to form presqualene diphosphate (PSPP) and a reductive rearrangement of PSPP to generate SQ. The first reaction is similar to the condensation of two molecules of DMAPP to give CPP catalyzed

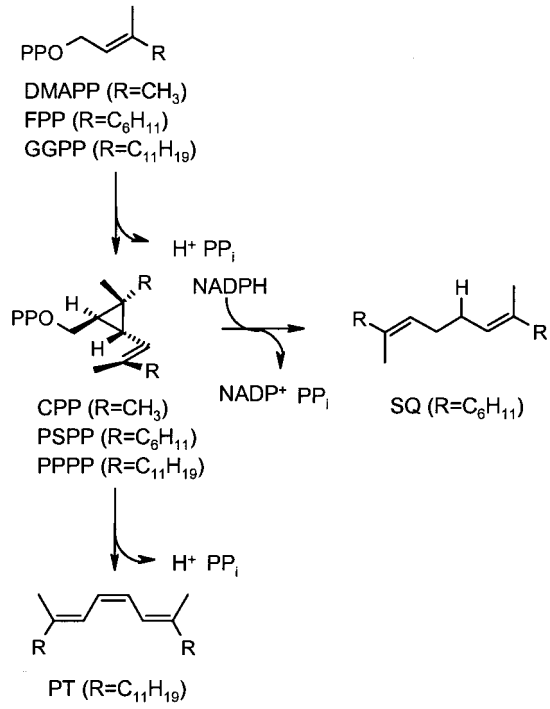

Fig. 1. Reactions catalyzed by CPPase, SQase, and PTase.

BIOCHEMISTRY

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: β-ME, 2-mercaptoethanol; CPPase, chrysanthemyl diphosphate synthase; CPP, chrysanthemyl diphosphate; DMAPP, dimethylallyl diphosphate; FPPase, farnesyl diphosphate synthase; IPP, isopentenyl diphosphate; PSPP, presqualene diphosphate; PT, phytoene; PTase, phytoene synthase; SQ, squalene; SQase, squalene synthase.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. I13995).

‡Present address: Myriad Genetics, Salt Lake City, UT 84108.

§Present address: Obstetrics and Gynecology, School of Medicine, University of Utah, 50 N. Medical Drive RM 2B200, Salt Lake City, UT 84112.

¶To whom reprint requests should be addressed. E-mail: Poulter@chemistry.utah.edu.

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

by chrysanthemyl diphosphate synthase (CPPase). In the second reaction, PSPP rearranges and is reduced to give SQ. The mechanism for formation of PT from geranylgeranyl diphosphate is closely related. In this case, two molecules of geranylgeranyl diphosphate condense to give prephytoene diphosphate, which then rearranges to PT (13).

Although SQase and phytoene synthase (PTase) have been studied extensively, virtually nothing is known about CPPase. We now report the isolation and characterization of the gene for CPPase from *Chrysanthemum cinerariaefolium*, a member of the Asteraceae family that produces CPP as an intermediate in the biosynthesis of the naturally occurring pyrethrin insecticides, construction of an *Escherichia coli* strain for production of recombinant enzyme, purification and characterization of the enzyme, and product studies that conclusively establish the function of the enzyme.

Experimental Procedures

Plant Materials, Substrates, and Reagents. Plant material was obtained from *C. cinerariaefolium* plants grown in a greenhouse. *E. coli* XA90 and pKEN2 were provided by Gregory Verdine (Harvard University). Synthesis of oligonucleotide primers, DNA sequencing, and analytical ultracentrifugation was performed at the Huntsman Cancer Institute Core Facility at the University of Utah or by Genosys Biotechnologies. Racemic *trans*-chrysanthemol was obtained from Aldrich. (1*R*,3*R*)-Chrysanthemol was synthesized by $LiAlH_4$ reduction of (1*R*,3*R*)-chrysanthemic acid (7). [1-^{14}C]DMAPP (55 mCi/mmol) was purchased from American Radiolabeled Chemicals (St. Louis), and unlabeled DMAPP, geranyl diphosphate, and IPP were synthesized by the procedure of Davisson *et al.* (14).

General Procedures. Standard molecular biology protocols were followed (15). Protein samples were analyzed by SDS/PAGE and stained with Coomassie blue R according to the procedure of Laemmli (16). Protein concentrations were determined by using BSA as a standard (17).

Synthesis of (1*R*,3*R*)-CPP. (1*R*,3*R*)-Chrysanthemol was phosphorylated by a modification of a procedure reported by Danilov *et al.* (18). *N*-tetrabutylammonium dihydrogen phosphate (726 mg, 2.14 mmol) was added to a solution of (1*R*,3*R*)-chrysanthemol (300 mg, 1.97 mmol) and trichloroacetonitrile (2.16 ml, 21.5 mmol) in 25 ml of acetonitrile. The mixture was stirred for 90 min and concentrated *in vacuo*. The residual oil was dissolved in water and passed through a column of 83 meq of Dowex AG50W-X8 resin (Aldrich, NH_4^+ form) and lyophilized. The residual white powder was dissolved in 25 mM NH_4HCO_3 and extracted three times with 1:1 (vol/vol) acetonitrile/isopropyl alcohol. The material was purified by flash chromatography on cellulose with a gradient from acetonitrile to 3:2 acetonitrile/25 mM NH_4HCO_3 to give 489 mg (31%) of a white solid; ^{1}H NMR (300 MHz, D_2O/ND_4OD) δ 0.97 (1H, ddd, $J = 9$ Hz, $J = 5$ Hz, $J = 5$ Hz), 1.04 (3H, s), 1.12 (3H, s), 1.32 (1H, dd, $J = 9$ Hz, $J = 5$ Hz), 1.67 (3H, s), 1.68 (3H, s), 3.81 (1H, ddd, $J = 9$ Hz, $J = 9$ Hz, $J = 5$ Hz), 4.14 (1H, ddd, $J = 5$ Hz, $J = 9$ Hz, $J = 5$ Hz), 4.98 (1H, d, $J = 9$ Hz); ^{31}P NMR (300 MHz, D_2O/ND_4OD) δ −10.5 (1P, d, $J_{P,P} = 22.0$ Hz), −8.2 (1P, d, $J_{P,P} = 22$ Hz). The ^{1}H NMR spectrum is consistent with a less well-resolved spectrum published previously (19).

Purification of Wild-Type CPPase. Plants of *C. cinerariaefolium* were grown at 15–23°C with an 18-h light/6-h dark cycle. Seedlings were initially induced to flower by a 6-week cold period at 4°C before they were transferred to the greenhouse. Plants were encouraged to continue flowering by harvesting all flower buds each week and by division of the plants. Buds were harvested as the flower petals began to emerge but before the petals had fully emerged and dropped to a horizontal position. Stem material was removed, and the buds were immediately placed on ice. Unless otherwise stated, subsequent purification steps were carried out at 4°C. Buds were disrupted in 100 mM Tris, pH 7.6/100 mM ascorbic acid/100 mM sucrose/100 mM 2-mercaptoethanol (β-ME)/5 mM $MgSO_4$/1 mM PMSF/0.4 g of insoluble polyvinylpolypyrrolidone. Grinding was done in two steps: an initial coarse grinding for 30 s in a blender followed by 2 min in a Sorvall Tissue Mizer. The disrupted material was spun for 15 min at 15,000 × *g*. The supernatant was removed, and the pellet was extracted with buffer and spun. The combined supernatants were filtered through Miracloth (Calbiochem). For each 20 ml of filtrate, 1 g of Whatman cell debris remover was added, the suspension was mixed thoroughly, and the material was allowed to stand without stirring for 10 min before it was passed through a Whatman 541 filter. The filtrate was adjusted to 30% saturation with $(NH_4)_2SO_4$, stirred for 30 min, and spun for 15 min at 15,000 × *g*. The supernatant was brought to 55% saturation, stirred for 30 min, and spun for 20 min at 15,000 × *g*. The pellet was stored at −70°C or used directly for further purification.

Approximately 300 mg of precipitated protein was resuspended in 35 ml of 20 mM imidazole, pH 7.0/5 mM $MgSO_4$/0.5 mM DTT. The suspension was spun at 100,000 × *g* for 45 min, the supernatant was passed through a 0.45-μm cellulose acetate filter, and the filtrate was passed through a KWIK desalting column (Pierce). The eluate was adjusted to 75 mM NaCl and applied to a Toyopearl DEAE-650 M (Toso Haas) column equilibrated in 20 mM imidazole, pH 7.0/5 mM $MgSO_4$/0.5 mM DTT/75 mM NaCl. Protein was eluted with a 160-ml linear gradient from 75 to 200 mM NaCl. Active fractions were brought to 2 M NaCl and chromatographed on phenyl-Sepharose CL-4B equilibrated with 20 mM imidazole, pH 7.0/5 mM $MgSO_4$/0.5 mM DTT/2 M NaCl. Protein was eluted in steps with 5-ml portions of 1.0 M, 0.5 M, 0.3 M, 0.2 M, and 0 M NaCl in 20 mM imidazole, pH 7.0. Peak activity eluted with the 0 mM NaCl fraction. The enzyme was stored as 30% glycerol stocks at −20°C.

The protein was applied to a Protein-Pak DEAE 8-h column (Waters) equilibrated with 20 mM Bis-Tris buffer, pH 6.3/5 mM $MgSO_4$/0.5 mM DTT and eluted with 0 to 100 mM NaCl over 10 min, followed by a linear gradient from 100 mM NaCl to 300 mM NaCl in Bis-Tris buffer. Active fractions were pooled and passed through a KWIK desalting column to remove NaCl and exchange Bis-Tris buffer for 1 mM phosphate, pH 6.8, containing 5 mM $MgSO_4$/0.5 mM DTT. The solution was applied to hydroxylapatite HCA (Rainin Instruments) equilibrated with 1 mM phosphate buffer, pH 6.8, and eluted with a 1 mM to 200 mM gradient of phosphate, pH 6.8.

A portion of the protein, which had been purified to apparent homogeneity as judged by SDS/PAGE, was lyophilized and resuspended in 20% acetonitrile/0.1% trifluoroacetic acid. The sample was chromatographed on a C4 column (Vydac, Hesperia, CA) reverse phase column with a gradient from 1:5 to 1:19 acetonitrile/0.1% triethylamine:water.

Tryptic Digestion of CPPase. A solution of 100 pmol of purified protein in 100 μl of 20 mM NH_4HCO_3 was adjusted to 4 M urea and 1.6 mM DTT and heated for 15 min at 95°C. Imidazole was added to 5 mM, and the mixture was allowed to stand at room temperature for 15 min. An equal volume of water was added, followed by sequencing grade trypsin (Promega) to a final concentration of 1.5 mg/ml. After incubation overnight at 37°C, the peptides were chromatographed on a C18 (Vydac) column by using a linear gradient from 1:18 A (0.1% trifluoroacetic acid):B (0.1% trifluoroacetic acid in 95% acetonitrile).

N-Terminal Protein Sequencing. HPLC-hydroxylapatite or reverse phase purified CPPase was subjected to SDS/PAGE. The band at 40 kDa was excised and electroblotted to poly(vinylidene difluoride) membranes. Electroblotted proteins were stained with Coomassie blue, and bands containing ≈100 pmol of CPPase were eluted and analyzed by microsequencing (20). CPPase that had been only purified by HPLC-hydroxylapatite was also sequenced.

Construction of a *C. cinerariaefolium* cDNA Library and Characterization of a CPPase Clone. A cDNA library was constructed from poly(A) selected RNA (Dynabeads, Dynal, Great Neck, NY) isolated from flower buds of *C. cinerariaefolium*. The cDNA was synthesized with a Zap cDNA synthesis kit (Stratagene) and cloned into the *Eco*RI/*Xho*I site of the Lambda Zap vector.

An oligonucleotide probe was constructed based on the region of the N-terminal protein sequence with the least degeneracy (QFMQVYET). The probe, CARTTYATGCARGTNTAYGARAC (A, adenosine; C, cytidine; G, guanosine; T, thymidine; R, purine; Y, pyrimidine; N, inosine), was 5′ end-labeled with [γ-^{32}P]ATP and used to screen the cDNA library. DNA from ≈106 plaques was transferred to duplicate sets of nitrocellulose filters and hybridized initially under low stringency (49°C, 6× SSC with 0.5% SDS overnight, followed by three washes for 10 min each at 45°C with 6× SSC and 0.1% SDS). One plaque hybridized with the probe. After plaque purification and rescreening, restriction analysis revealed a 1.6-kb *Eco*RI/*Xho*I fragment that was excised and incorporated as phagemids into XLl pBluescript SK$^+$ cells with the Stratagene Exassist/Solr system. Plasmid DNA was extracted from cells, purified with anion-exchange resin (Qiagen, Chatsworth, CA), and sequenced.

Recombinant Poly(His)-CPPase. An *Nde*I restriction site was introduced by PCR after the putative plastidial targeting sequence in CPPase (21) to generate a truncated gene encoding the post-translationally modified mature protein isolated from plants. The ORF was subcloned into pMPM3B, a derivative of pKEN2, to generate pSCB1. Oligonucleotides 5′-TATG*CATCATCATCATCATCAT*ATAGAGGGGCG-3′ and 3′-AC*GTAGTAGTAGTAGTAGTA*TATCTCCCCGCAT-5′ were ligated into the *Nde*I site of pSCB1 to form pSCB4. The resulting insert contained codons for six histidines (italics) and a factor Xa recognition site (underlined). The correct orientation of the insert was determined by restriction enzyme screening and verified by sequence analysis.

Overproduction and Purification of Recombinant CPPase. *E. coli* strain XA90 was transformed with pSCB4. Overnight cultures of single transformants in Luria Bertani (LB) medium containing 100 μg of ampicillin per ml were used to inoculate 500 ml of LB medium containing 100 μg of ampicillin per ml. Cells were incubated at 30°C with shaking at 250 rpm and induced with isopropyl β-D-thiogalactoside to a final concentration of 0.5 mM when the $OD_{600} \approx 0.5$. Cells were harvested 4 h after induction by centrifugation at 40,000 × *g* for 10 min. The resulting paste was resuspended in 20 ml of buffer containing 1 mM EDTA/1 mM EGTA/1 mM PMSF/10 μg/ml leupeptin/5 μg/ml pepstatin/10 mM β-ME/50 mM Tris, pH 7.5, and disrupted by sonication. The suspension was spun at 40,000 × *g* for 20 min. The supernatant was adjusted to 50% saturation with $(NH_4)_2SO_4$ and cleared by centrifugation. The pellet was resuspended in 5 ml of lysis buffer (500 mM NaCl/10 mM imidazole/5 mM β-ME/50 mM phosphate buffer, pH 7.7) and loaded at 0.1 ml/min onto a Ni^{2+}-NTA Silica column (Qiagen) that had been equilibrated with lysis buffer. The column was washed with 40 ml of lysis buffer and eluted with a 0 to 300 mM gradient of imidazole over 70 ml. The eluent was concentrated by centrifugation with an Amicon Centriprep filter and dialyzed against 30 mM Hepes, pH 7.5, containing 5 mM β-ME.

Product Analysis. Purified enzyme (30 μg) was added to 50 mM Tris, pH 7.8/5 mM $MgCl_2$/0.5 mM DTT/2 mM DMAPP to a final volume of 0.5 ml. The mixture was incubated from 15 min to 2 h at 30°C and then heated at 95°C for 2 min. Glycine (110 μl of 500 mM, pH 10.5), 5 mM $ZnCl_2$, and 40 units of calf alkaline phosphatase (Sigma) were added to the cooled solution, and the mixture was incubated at 37°C for 30 min. Approximately 0.5 g of NaCl was added, and the mixture was extracted with 1 ml of *t*-butyl methyl ether. A 1-μl portion of the extract was analyzed by gas chromatography on an Innowax capillary column (Hewlett–Packard) with a temperature program from 70°C to 150°C at 10°C per min. Samples for GC/MS were resolved on a 30-m DB-1 Supelcowax capillary column. The mass spectra of eluted peaks were compared with published spectra for *cis*- and *trans*-chrysanthemol (22).

Farnesyl Diphosphate Synthase (FPPase) Assay. CPPase was assayed for FPPase activity by the acid lability method (23) in 200 μl of 40 mM bicyclo[2.2.1]hept-5-en-2,3-dicarboxylic acid (pH 7.0) containing 2 mM $MgCl_2$, 2 mg/ml BSA, 20 mM β-ME, [1-^{14}C]IPP (1 mM, 5 μCi/μmol) and 5 mM allylic substrate (geranyl diphosphate or DMAPP).

CPPase TLC Assay. Reactions were carried out in 20 μl of 35 mM Hepes (pH 7.5) containing 10 mM $MgCl_2$ and 0.5 mM DTT and with various concentrations of [1-^{14}C]DMAPP (1.0 μCi/μmol). The reactions were initiated by the addition of 5–10 μg of CPPase and incubated for 10–40 min before a 5-μl portion was removed and spotted on a 20 × 20-cm 60 Å silica plate (Merck). The plates were developed in 25:15:4:2 (vol/vol) $CHCl_3$/methanol/water/acetic acid and dried for 10 min with a gentle stream of air. The TLC plates were imaged (Molecular Dynamics Storm 840 PhosphorImager), and the resulting autoradiogram was analyzed to determine amount of substrate that had been converted to product. Data were corrected for background signal with a nonenzyme control.

pH Dependence. Reactions were incubated at 30°C in 50 μl of poly buffer (24) consisting of 100 mM Tris/50 mM Mes/50 mM acetic acid and containing 5 mM $MgCl_2$, 0.5 mM DTT, 207 μM [1-^{14}C]DMAPP (2.5 μCi/μmol), and 20 μg of purified CPPase from 4.5 to 8.5.

Molecular Mass Determination. Size exclusion chromatography was performed on a Superdex 200 (Amersham Pharmacia) gel filtration column. The column was equilibrated with running buffer consisting of 50 mM bicyclo[2.2.1]hept-5-en-2,3-dicarboxylic acid, pH 7.0/10 mM β-ME/10 mM $MgCl_2$ and calibrated with chymotrypsinogen A (35 kDa), ovalbumin (43 kDa), BSA (66 kDa), and aldolase (158 kDa). The molecular mass of CPPase was calculated from a linear calibration curve of the logarithm of the molecular mass vs. elution volume obtained for the standards.

Sedimentation equilibrium experiments were performed on a Beckman Optima XL-A analytical ultracentrifuge at 19,000 rpm over a temperature range of 4–20°C. P6-spin columns (Bio-Rad) were used to exchange samples of CPPase into buffer A (40 mM bicyclo[2.2.1]hept-5-en-2,3-dicarboxylic acid, pH 7.0/5 mM $MgCl_2$/0.5 mM DTT) or buffer B (buffer A containing 100 mM NaCl). Three different CPPase concentrations, ranging from 0.10 to 0.32 mg/ml, were used to load two 6-channel 12-mm-thick Charcoal-epon centerpieces, which were surrounded by matched quartz windows (25). Ten scans of absorbance vs. radial distribution were collected with a step size of 0.001 cm after sedimentation equilibrium had been reached at 4, 10, 15, and

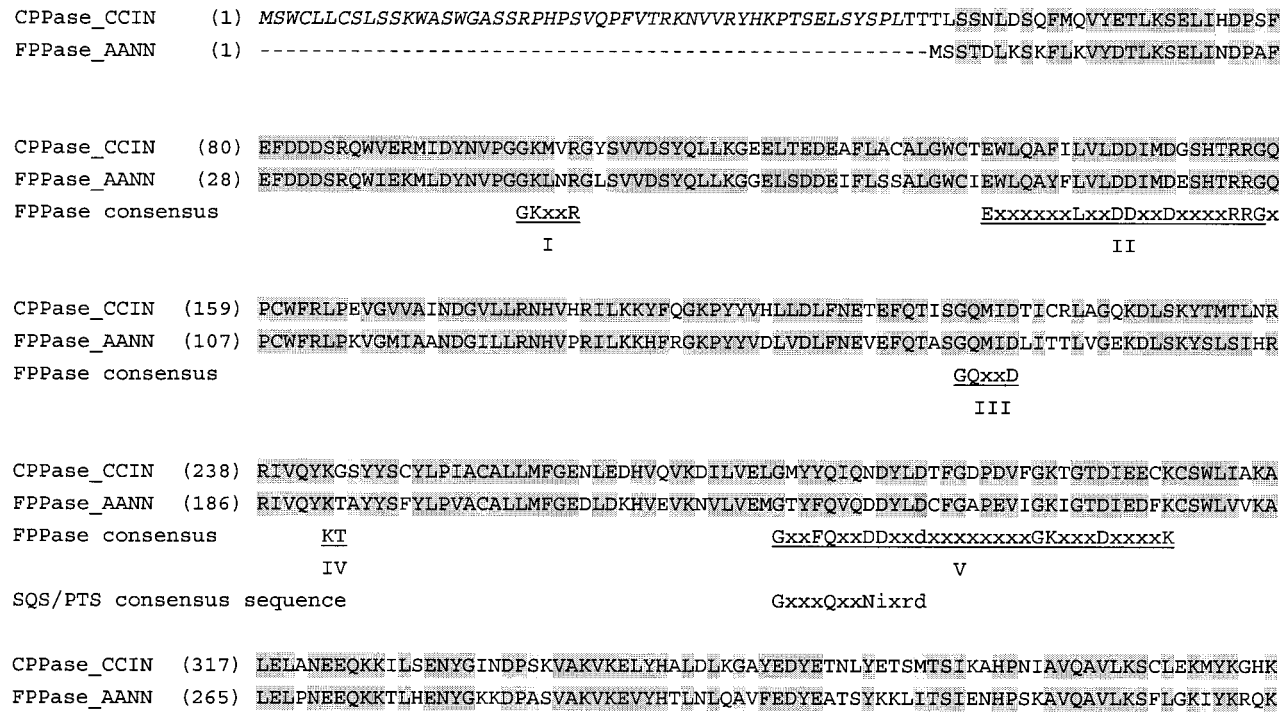

Fig. 2. Alignment CPPase with FPPases from the Asteraceae family. The designations correspond to the following: CPPase_CCIN, *C. cinerariaefolium* CPPase (I13995) and FPPase_AANN, *A. annua* FPPase (U36376). The putative targeting peptide of CPPase is italicized. The FPPase consensus sequence highlights the five conserved domains identified in FPPase from a variety of organisms (3, 27). Prosite (PS01045) coupled with block sequence analysis aligned an SQase/PTase consensus sequence with domain V of FPPase. Two of the CPPase residues (Asn-285 and Asp-293) that do not match amino acids conserved in FPPase align with conserved residues in the SQase/PTase consensus sequence.

20°C. The scans were averaged, and data were corrected for window aberrations and partial masking of light pulses by subtraction of a 360-nm scan from each 280-nm scan (25). The density of each solution and partial specific volume for each temperature (26) was calculated in SEDNTERP (version 1.01). The partial specific volume at 20°C and the density of buffer B were 0.7344 ml/g and 1.004, respectively. Various models of association were used to fit the absorbance vs. radial data, and the "goodness of fit" was assessed from residual plots.

Results

Purification of CPPase from *C. cinerariaefolium*. CPPase was purified in four chromatographic steps to give a single polypeptide band on silver-stained SDS/PAGE with a molecular mass of ≈41 kDa. Activity was lost between each purification step unless the protein was stored as an $(NH_4)_2SO_4$ precipitate, where activity was stable for several months. Because samples rapidly lost activity after DEAE purification, they were rapidly processed through the phenyl-Sepharose step (see *Experimental Procedures*). After chromatography on phenyl Sepharose, the enzyme could be stored at −20°C in 30% glycerol for ≈72 h with only a 20% decrease in activity.

Isolation of cDNA Clones and Sequence Determination. The N-terminal sequence and four tryptic fragments from CPPase were used to construct degenerate oligonucleotide probes for a *C. cinerariaefolium* cDNA library. Sequence data from the initial positive clone were compared with the N-terminal and internal peptide sequences to ensure that the correct construct had been isolated. A new probe was constructed based on the sequence of the first clone, and the library was rescreened at higher stringency. After screening ≈5 million additional plaques, five more phages were isolated that hybridized strongly with the N-terminal probe. Variations in length and in the regions flanking the putative ORF indicated that six independent clones had been isolated.

The N-terminal sequence of the native protein began 50 aa downstream from a putative initiating methionine. Subtraction of the mass for these 50 aa from 45,358 Da gave a deduced molecular mass of 39,681 for processed CPPase, in agreement with the value of ≈41 kDa estimated by SDS/PAGE. These observations indicate that CPPase is posttranslationally modified by removal of an N-terminal leader. The amino acid sequence of CPPase has a high degree of similarity with members of the FPPase family, including the enzyme from *Artemisia annua*, a member of the Asteraceae family, as shown in Fig. 2 (see also ref. 27). Lower degrees of similarity were seen with terpene cyclases and other terpene synthases. Surprisingly, little similarity was seen with SQase or PTase, enzymes that catalyze a similar 1′-2-3 cyclopropanation reaction.

Expression and Purification of CPPase. Recombinant His-tagged CPPase was purified from cell-free homogenates in two steps with $(NH_4)_2SO_4$ precipitation and Ni^{2+} affinity chromatography. The recombinant protein was greater than 95% pure as judged by SDS/PAGE and had a specific activity of 160 nmol $h^{-1} \cdot mg^{-1}$.

Product Analysis. The structure of CPP formed upon incubation of CPPase with DMAPP was verified by hydrolysis of the diphosphate with alkaline phosphatase and GC/MS analysis of the resulting monoterpene alcohol. Chrysanthemol was not detected if the phosphatase step was omitted. Although a trace amount of yomogi alcohol (22) was detected in CPPase incubations longer than 1.5 h, control experiments indicated that the alcohol was from the nonenzymatic hydrolysis of CPP (28).

Incubations of CPPase and radiolabeled IPP with DMAPP or

geranyl diphosphate gave no evidence for formation of a 1′-4 product. The upper limit of FPPase activity in a purified sample of CPPase was 800-fold lower than a cell-free preparation of *A. annua* (29), 850-fold lower than a cell-free preparation from *Gossypium arboreum* (30), and 2,000-fold lower than purified recombinant avian FPPase. Thus, CPPase seems to be incapable of catalyzing 1′-4 chain elongation despite its significant sequence homology to FPPase.

Molecular Mass Determination of Recombinant CPPase. Size exclusion chromatography gave a molecular mass for CPPase of ≈66 kDa. This value is substantially higher than the mass of 41 kDa estimated by gel filtration but lower than expected for a homodimer. A more accurate determination was then made by analytical sedimentation equilibrium ultracentrifugation. Data were collected at varying temperatures and three different concentrations of CPPase. To obtain a rigorous fit, the data obtained for each of the three different loading concentrations at a given temperature were fit simultaneously, and a plot for each temperature and buffer condition reflecting three different protein concentrations was generated. Eight plots were obtained, and in each case the data were best described by a monomer–dimer model. The calculated molecular mass was 41.2 kDa, in agreement with the value of 41,218 kDa calculated for the affinity-tagged protein. The estimated equilibrium constant for the dimerization of CPPase was $\approx 2.0 \times 10^6\ M^{-1}$.

TLC Assay and Steady-State Kinetic Constants. A TLC assay with elution by $CHCl_3$:methanol:water:acetic acid was developed to resolve DMAPP and CPP. In this solvent, CPP migrates faster than DMAPP, and the corresponding alcohols migrate with the solvent front. We detected a small amount of radioactivity at the solvent front for all of our samples, including control samples of DMAPP before incubation. Although allylic and cyclopropylcarbinyl diphosphates are sensitive to acid, the fast-moving radioactivity does not increase with longer exposure to the TLC solvent and seems to be an impurity in DMAPP.

Like the FPPase, CPPase requires a divalent metal ion. Maximal rates were seen for Mg^{2+} at concentrations between 1 and 10 mM. Mn^{2+} served as a substitute, but the specific activity of the enzyme was 2-fold lower at the optimal Mn^{2+} concentration of 1 mM. CPPase had a broad pH optimum from 6.5 to 8.0. The DMAPP concentration required for half-maximal velocity, $[S]_{0.5}^{DMAPP}$, was 600 ± 150 μM, and k_{cat} was 0.5 min^{-1} with the enzyme preparation of highest specific activity. When the poly(His) tag was removed from CPPase with factor Xa, the kinetic constants did not change significantly.

Discussion

Except for the sterol and carotenoid pathways, little is known about the biosynthesis of irregular isoprenoids at the genetic or at the enzyme level. The greatest variety of structures for irregular isoprenoids is found in plants, especially the monoterpenes. It is generally assumed that monoterpene biosynthesis in plants occurs in plastids via the recently discovered methylerythritol phosphate pathway (1). Typically, the nuclear gene products are directed to plastids by specific amino-terminal peptide sequences, where they are processed to mature forms (31, 32). Although plastidial targeting sequences for plant isoprenoid enzymes share little homology with each other or with known plastid targeting sequences (31), they are typically rich in serine, threonine, and small hydrophobic residues and contain few acidic residues. In contrast, peptides targeted to mitochondria have positively charged and hydrophobic residues mixed throughout their lengths (33). The 50-aa N-terminal sequence in CPPase shown in Fig. 2 is consistent with a plastidial target, although we have not established the final destination for the enzyme.

The flowers of *C. cinerariaefolium* produce potent pyrethrin insecticides, a family of six monoterpene esters with low mammalian toxicity (34, 35). The acid portion of the esters is either chrysanthemic or pyrethric acid, and the alcohol portion is rethrolone, cinerolone, or jasmolone. Chrysanthemic and pyrethric acids are irregular monoterpenes with c1′-2-3 linkages between the two dimethylallyl units. These molecules are presumably synthesized from DMAPP to give CPP, followed by hydrolysis of the diphosphate moiety and oxidation of chrysanthemol to give the acids. Chrysanthemol has been found in the essential oils of other members of the Asteraceae family (7, 8).

Epstein and Poulter (9) recognized that CPP could rearrange to a variety of other irregular monoterpenes and proposed that CPP has a central role in the biosynthesis of these molecules. Model studies subsequently demonstrated the feasibility of their proposal (2, 28). As illustrated in Fig. 1, cyclopropylcarbinyl diphosphates are also intermediates in the biosynthesis of the irregular 1′-1 linkages found in the sterol and carotenoid pathways. CPP, PSPP, and prephytoene diphosphate all have the same absolute stereochemistry at the common stereocenters in their cyclopropane rings. CPPase only catalyzes the cyclopropanation reaction, whereas SQase and PTase also catalyze the rearrangements of PSPP and prephytoene diphosphate to 1′-1 structures.

The similarity between CPPase and FPPase is striking! At the primary sequence level, 70% of the amino acids in CPPase and the cytosolic FPPase from *A. annua*, a closely related plant in the Asteraceae family (Fig. 2), are identical, and 80% are identical or strongly similar. In addition, both enzymes are homodimers. Presumably, the three-dimensional structures of the two enzymes are also similar, including many features in the catalytic

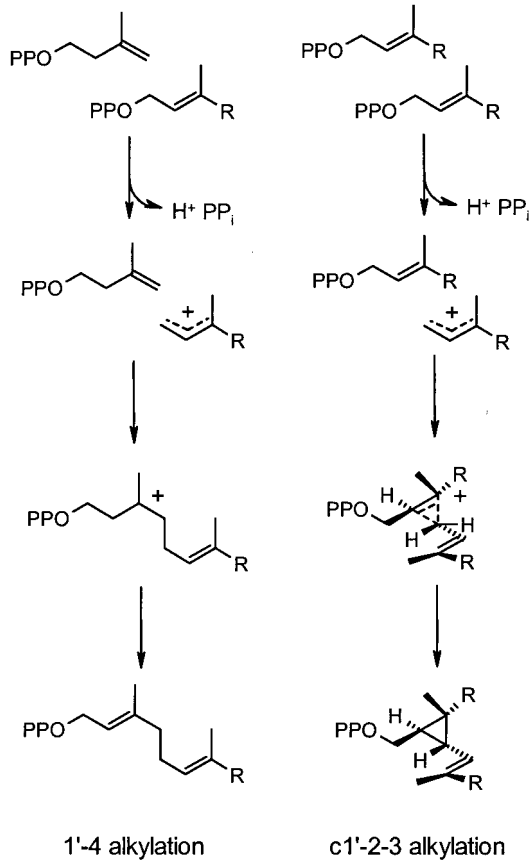

Fig. 3. Electrophilic mechanisms for 1′-4 chain elongation and c1′-2-3 cyclopropanation.

site. FPPase contains five characteristic highly conserved regions that cluster around the active site (36). These are present in CPPase, with the exception of an asparagine replacement of an aspartate residue in the second of the two aspartate-rich DDXXD motifs in FPPase that bind the diphosphate residues of the substrates through Mg^{2+} bridges.

The mechanism of the chemical reaction for 1′-4 elongation catalyzed by FPPase is a stepwise electrophilic alkylation of the double bond in IPP by a carbocation generated from the allylic diphosphate substrate (37, 38) as illustrated in Fig. 3. Over the years, several mechanisms have been proposed for the cyclopropanation reaction (39–45). Although some of the earlier proposals have succumbed to the sands of time and are now regarded as unlikely, none enjoys a significant amount of experimental support. However, the substantial structural similarity between CPPase and FPPase provides strong ''genetic'' evidence for an electrophilic c1′-2-3 cyclopropanation (2) that closely parallels the electrophilic alkylation for 1′-4 elongation.

The amino acid sequences of SQase and PTase, the other two enzymes that catalyze intermolecular c1′-2-3 cyclopropanation reactions, have only a very limited similarity to those of CPPase or FPPase. In fact, the only noticeable motif is a GXXXQXX-DDXXD sequence at the second aspartate-rich region in FPPase and CPPase that corresponds to a GXXXQXXNIXRD sequence in SQase and PTase (see Fig. 2). Interestingly, the sequence for the second aspartate-rich motif in CPPase is GMYYQIQNDYLD, where the asparagine that replaces the first aspartate in the DDXXD motif is in the same location as the conserved asparagine in the SQase and PTase sequences. However, the D → N replacement is not an absolute requirement for cyclopropanation. Casbene synthase, an enzyme that catalyzes a related intramolecular cyclopropanation of the distal double bond in geranylgeranyl diphosphate, has the typical DDXXD motif for the allylic diphosphate binding site. The amino acid sequence of casbene synthase most closely aligns with other terpene cyclases that do not catalyze cyclopropanation reactions and has only limited similarity with the FPPases, geranylgeranyl diphosphate synthases, SQases, and PTases (46).

A recent x-ray structure of human SQase (47) reveals that the enzyme has the ''prenyltransferase fold'' first seen for FPPase (36). Two class I sesquiterpene cyclases whose x-ray structures were recently reported, 5-*epi*-aristolochene synthase (48) from tobacco and pentalenene synthase (49) from *Streptomyces*, also have the prenyltransferase fold. Thus, it is likely that the 1′-4 chain elongation enzymes related to FPPase, the c1′-2-3 cyclopropanating enzymes CPPase, SQase, and PTase, and the class I cyclases all belong to a superfamily of enzymes in the isoprenoid pathway whose members catalyze electrophilic alkylations.

This work was supported by National Institutes of Health Grant GM 25521.

1. Eisenreich, W., Schwartz, M., Cartayrade, A., Arigoni, D., Zenk, M. H. & Bacher, A. (1998) *Chem. Biol.* **5,** R221–R233.
2. Poulter, C. D. (1990) *Acc. Chem. Res.* **23,** 70–77.
3. Kellogg, B. A. & Poulter, C. D. (1997) *Curr. Opin. Chem. Biol.* **4,** 570–578.
4. Koyama, T. & Ogura, K. (1998) in *Comprehensive Natural Products Chemistry: Isoprenoids*, ed. Cane, D. E. (Elsevier Science, Oxford), Vol. 2, pp. 73–96.
5. Gunawardena, K. (1994) Ph.D. dissertation (Univ. of Utah, Salt Lake City).
6. Epstein, W. W., Gaudioso, L. A. & Brewster, G. B. (1984) *J. Org. Chem.* **49,** 2748–2754.
7. Alexander, K. & Epstein. W. W. (1975) *J. Org. Chem.* **40,** 2576.
8. Epstein, W. W., Klobus, M. A. & Edison, A. S. (1991) *J. Org. Chem.* **56,** 4451–4456.
9. Epstein, W. W. & Poulter, C. D. (1973) *Phytochemistry* **12,** 737–747.
10. Shaw, J., Noble, T. & Epstein, W. W. (1975) *J. Chem. Soc. Chem. Commun*., 590–591.
11. Noble, T. A. & Epstein, W. W. (1977) *Tetrahedron Lett*., 3931–3932.
12. Epstein, W. W. & Gaudioso, L. A. (1979) *J. Org. Chem.* **44,** 3113–3117.
13. Gregonis, D. E. & Rilling, H. C. (1974) *Biochemistry* **13,** 1538–1542.
14. Davisson, V. J., Woodside, A. B., Neal, T. R., Stremler, K. E., Muehlbacher, M. & Poulter, C. D. (1986) *J. Org. Chem.* **51,** 4768–4779.
15. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
16. Laemmli, U. K. (1970) *Nature (London)* **227,** 680–685.
17. Bradford, N. M. (1976) *Anal. Biochem.* **72,** 248–254.
18. Danilov, L. L., Druzhinina, T. N., Kalinchuk, N. A., Maltsev, S. D. & Shibaev, V. N. (1989) *Chem. Phys. Lipids* **51,** 191–203.
19. Epstein, W. W. & Rilling, H. C. (1970) *J. Biol. Chem.* **245,** 4597–4605.
20. Matsudaira, P. T. (1990) *Methods Enzymol.* **182,** 602–613.
21. Ellenberger, S. R., Peiser, G. D., Bell, R. N., Hussey, C. E., Shattuck-Eidens, D. M. & Swedlund, B. D. (1995) U.S. Patent 5,443,978.
22. Epstein, W. W., McGee, L. R., Poulter, C. D. & Marsh. L. L. (1976) *J. Chem. Eng. Data* **21,** 500–502.
23. Bartlett, D. L., King, C. H. & Poulter, C. D. (1985) *Methods Enzymol.* **110,** 171–184.
24. Ellis, K. J. & Morrison, J. F. (1982) *Methods Enzymol.* **87,** 405–426.
25. Johnson, M. L., Correia, J. J., Yphantis, D. A. & Halvorson, H. R. (1981) *Biophys. J.* **36,** 575–588.
26. Laue, T. M., Shah, B. D., Ridgeway, T. M. & Pelletier, S. L. (1992) in *Analytical Ultracentrifugation in Biochemistry and Polymer Science,* eds. Harding, S. E., Rowe, A. J. & Horton, J. C. (R. Soc. Chem., Cambridge, U.K.), pp. 90–125.
27. Chen, A., Kroon, P. A. & Poulter, C. D. (1994) *Protein Sci.* **3,** 600–607.
28. Poulter, C. D., Marsh, L. L., Hughes, J. M., Argyle, J. C., Satterwhite, D. M., Goodfellow, R. J. & Moesinger, S. G. (1977) *J. Am. Chem. Soc*. **99,** 3816–3823.
29. Matsushita, Y., Kang, W. & Charlwood, B. V. (1996) *Gene* **172,** 207–209.
30. Liu, C., Heinstein, P. & Chen X. (1999) *Mol. Plant–Microbe Interact.* **12,** 1095–1104.
31. Wise, M. L. & Croteau, R. (1998) in *Comprehensive Natural Products Chemistry: Isoprenoids*, ed. Cane, D. E. (Elsevier Science, Oxford), Vol. 2, pp. 97–153.
32. Turner, G., Gershenzon, J., Nielson, E. E., Froehlich, J. E. & Croteau, R. (1999) *Plant Physiol.* **120,** 879–886.
33. Keegstra, K. & von Heijne, G. (1992) in *Plant Gene Research: Cell Organelles*, ed. Herrmann, R. G. (Springer, New York), pp. 353–370.
34. Crowley, M. P., Inglis, H. S., Snarey, M. & Thain, E. M. (1961) *Nature (London)* **191,** 281–282.
35. Crombie, L. (1995) in *Pyrethrum Flowers*, eds. Casida, J. E. & Quistad, G. B. (Oxford Univ. Press, New York), pp. 123–193.
36. Tarshis, L. C., Yan, M., Poulter, C. D. & Sacchettini, J. C. (1994) *Biochemistry* **33,** 10871–10877.
37. Poulter, C. D., Wiggins, P. L. & Le, A. T. (1981) *J. Am. Chem. Soc.* **103,** 3926–3927.
38. Davisson, V. J., Neal, T. R. & Poulter, C. D. (1993) *J. Am. Chem. Soc.* **115,** 1235–1245.
39. Saunders, M., Vogel, P., Hagen, E. L. & Rosenfeld, J. (1973) *Acc. Chem. Res.* **6,** 53–59.
40. Altman, L. J., Kowerski, R. C. & Laungani, D. R. (1978) *J. Am. Chem. Soc.* **100,** 6174–6182.
41. Rilling, H. C., Poulter, C. D., Epstein, W. W. & Larsen, B. R. (1971) *J. Am. Chem. Soc.* **93,** 1783–1785.
42. Beytia, E., Qureshi, A. A. & Porter, J. W. (1973) *J. Biol. Chem.* **248,** 1856–1867.
43. Van Tamelen, E. E. & Schwartz, M. A. (1971) *J. Am. Chem. Soc*. **93,** 1780–1782.
44. Trost, B. M. & Biddlecom, W. G. (1973) *J. Org. Chem.* **38,** 3438–3439.
45. Cohen, T., Herman, G., Chapman, T. M. & Kuhn, D. (1974) *J. Am. Chem. Soc.* **96,** 5627–5628.
46. Mau, C. J. D. & West, C. A. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 8497–8501.
47. Pandit, J., Danley, D. E., Schulte, G. K., Mazzalupo, S., Pauly, T. A., Hayward, C. M., Hamanaka, E. S., Thompson, J. F. & Harwood, H. J., Jr. (2000) *J. Biol. Chem.* **275,** 30610–30617.
48. Starks, C. M., Back, K., Chappell, J. & Noel, J. P. (1997) *Science* **277,** 1815–1820.
49. Lesburg, C. A., Zhai, G., Cane, D. E. & Christianson, D. W. (1997) *Science* **277,** 1820–1824.