Pentalenene Synthase. Purification, Molecular Cloning, Sequencing, and High-Level Expression in *Escherichia coli* of a Terpenoid Cyclase from *Streptomyces* UC5319[†,‡]

David E. Cane,* Jae-Kyung Sohng, Carol R. Lamberson, Suzanne M. Rudnicki, Zhen Wu, Matthew D. Lloyd, John S. Oliver, and Brian R. Hubbard

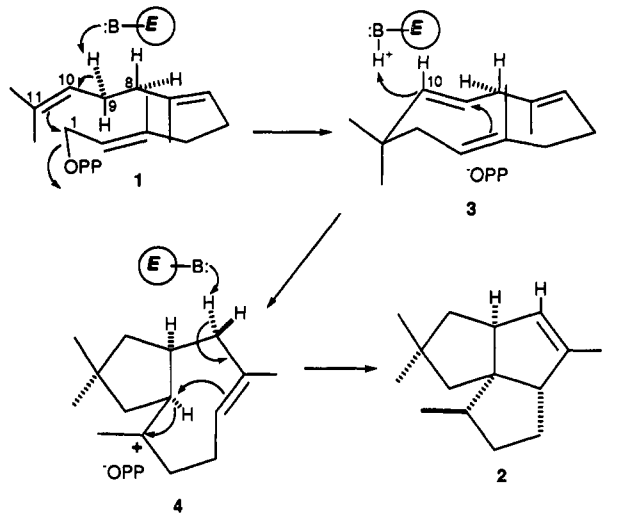
Department of Chemistry, Box H, Brown University, Providence, Rhode Island 02912

Received January 20, 1994; Revised Manuscript Received March 7, 1994[⊗]

ABSTRACT: Pentalenene synthase, which catalyzes the cyclization of farnesyl diphosphate (**1**) to the tricyclic sesquiterpene hydrocarbon pentalenene (**2**), was purified from *Streptomyces* UC5319. A 450-bp hybridization probe, generated by PCR amplification of genomic DNA using primers based on N-terminal and internal tryptic peptide sequence data for pentalenene synthase, was used to screen both plasmid and phage DNA libraries of *Streptomyces* genomic DNA, resulting in the isolation and sequencing of the complete pentalenene synthase gene. PCR was used to insert the pentalenene synthase gene into the T7 expression vector pLM1. Cloning of the resulting construct in the expression host *Escherichia coli* BL21(DE3) gave transformants that expressed pentalenene synthase as greater than 10% of soluble protein. The recombinant enzyme has been purified, and initial physical and kinetic characterization has been performed. The recombinant enzyme appears to be identical in every respect with the native *Streptomyces* synthase and exhibits the following steady-state kinetic parameters: K_m = 0.31 ± 0.05 μM, k_{cat} = 0.32 ± 0.02 s^{-1}, $K_I(PP_i)$ = 3.2 ± 0.6 μM. Both enzymes have an absolute requirement of Mg^{2+} for catalysis and an optimum pH of 8.2–8.4. Both proteins have M_r values of 41–42 kDa, as determined by SDS–PAGE.

Pentalenene synthase catalyzes the cyclization of farnesyl diphosphate (FPP,[1] **1**) to the tricyclic sesquiterpene pentalenene (**2**), the hydrocarbon precursor of the pentalenolactone family of antibiotics which are produced by a variety of *Streptomyces* species (Cane & Tillman, 1983). Extensive investigations with stereospecifically labeled substrates have provided a detailed mechanistic and stereochemical model of pentalenene formation according to which FPP, folded in the conformation illustrated in Scheme 1, undergoes ionization and electrophilic attack of the resultant allylic cation–pyrophosphate anion pair on the distal double bond (Cane et al., 1984, 1990a). Stereospecific removal of the H-9_{re} proton is believed to generate the 11-membered-ring intermediate humulene (**3**). It has been demonstrated that this proton does not exchange with the medium but instead is used to reprotonate the humulene at C-10. Further cyclization generates the protoilludyl cation (**4**), which, upon hydride shift, cyclization, and loss of the proton derived from H-8_{si} of FPP, is converted to pentalenene (Cane et al., 1991; Cane & Weiner, 1994). In principle, a single base would be geometrically competent to mediate the necessary deprotonation–reprotonation–deprotonation processes at C-9, -10, and -8.

Scheme 1: Cyclization of FPP (**1**) to Pentalenene (**2**) through the Intermediacy of Humulene (**3**) and the Protoilludyl Cation **4**



We have previously reported the partial purification and preliminary kinetic characterization of pentalenene synthase from *Streptomyces* UC5319 (Cane & Pargellis, 1987). In common with all other known cyclases generating mono-, sesqui-, and diterpenes, pentalenene synthase requires no cofactors other than Mg^{2+} (Croteau, 1987; Cane, 1990). No further information has been available on the nature of the active site, including the identification of the amino acid residues responsible for proper binding, folding, and ionization of the substrate and stabilization of positive charge during the course of the cyclization, nor has the base which carries out the various proton transfers been identified. We now report the purification of pentalenene synthase to homogeneity, the

[†] This research was supported by National Institutes of Health Grant GM-22172 to D.E.C.

[‡] The 1.48-kb DNA sequence from *Streptomyces* UC5319 which includes the pentalenene synthase gene has been deposited at GenBank and assigned Accession Number U05213.

[⊗] Abstract published in *Advance ACS Abstracts,* April 15, 1994.

[1] Abbreviations: AS, ammonium sulfate; BSA, bovine serum albumin; dsDNA, double-stranded DNA; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid, disodium salt; FPLC, fast protein liquid chromatography; FPP, farnesyl diphosphate; HIC, hydrophobic-interaction chromatography; HPLC, high-pressure liquid chromatography; IPTG, isopropyl β-D-thiogalactopyranoside; LB, Luria–Bertani medium; NTP, nucleotide triphosphate; ORF, open reading frame; PCR, polymerase chain reaction; PEG, poly(ethylene glycol); PMSF, phenylmethanesulfonyl fluoride; PVDF, poly(vinylidene difluoride); rbs, ribosome binding site; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; 20× SSC, 3.0 M NaCl and 0.3 M sodium citrate; ssDNA, single-stranded DNA; TE buffer, 10 mM Tris and 1 mM EDTA, pH 8.0; X-gal, 5-bromo-4-chloro-3-indolyl β-D-galactoside.

0006-2960/94/0433-5846$04.50/0

isolation and sequencing of the corresponding structural gene, and the overexpression of cloned pentalenene synthase as 10% of soluble protein in *Escherichia coli*.

MATERIALS AND METHODS

Materials. *Streptomyces* UC5319 was a gift from the Upjohn Co. (Kalamazoo, MI). Plasmid pLM1 and *E. coli* strain BL21(DE3) (F^- *ompT* r_B^- m_B^- (DE3)) were gifts from Professor Gregory L. Verdine of Harvard University. Library efficiency *E. coli* DH5α competent cells were obtained from Gibco BRL (Grand Island, NY). *E. coli* NM522, helper phage M13KO7, and plasmids pTZ18R and pTZ19R were purchased from Pharmacia (Piscataway, NJ). *E. coli* XL1-Blue and Pfu DNA polymerase were from Stratagene (La Jolla, CA). The vectors EMBL3 and pGEM5Zf(−), *E. coli* NM539, Packagene λ DNA packaging system, and T4 DNA ligase were purchased from Promega (Madison, WI). Lysozyme, DNase I, ribonuclease A, Reactive Blue 4-agarose, and Superose 12 were obtained from Sigma (St. Louis, MO), and Bio-Gel P-6DG was purchased from Bio-Rad (Hercules, CA). DEAE-Sephadex, bulk matrix Q Sepharose, and Sephacryl S-200 were from Pharmacia. [α-^{35}S]dATP (1168 Ci/mmol), [α-^{35}S]dCTP (1039 Ci/mmol), GeneScreen Plus nylon membranes, and Colony/Plaque screen hybridization membranes were purchased from DuPont NEN Research Products (Boston, MA). The coupling of 8-[(3-aminopropyl)-amino]adenosine to CH-Sepharose 4B to give Ado-Sepharose was carried out as described (Trayer et al., 1974; Kajander & Raina, 1981). [1-^{3}H]FPP (16.7 and 71.7 μCi/μmol) was synthesized as previously reported (Cane & Ha, 1988). Oligonucleotides were synthesized in the Division of Biomedical Sciences at Brown University by Charles Sutherland or prepared by Genosys Biotechnologies (The Woodlands, TX). Nitrocellulose membranes (Trans-Blot transfer medium, 0.5 mM) were obtained from Bio-Rad. Poly(vinylidene difluoride) (PVDF) membranes and Ultrafree Pro-Bind and MC-30,000 MW filters were purchased from Millipore (Bedford, MA). Terminal deoxynucleotidyl transferase was purchased from the United States Biochemical Corp. (Cleveland, OH). All other materials used for recombinant DNA manipulations, enzyme assay, and protein purification were analytical grade or higher. All buffers and nutrient broths were prepared with doubly deionized nanopure grade water.

General Methods. Restriction endonuclease digestions, DNA ligations, preparation and transformation of competent cells, plasmid minipreps, Southern hybridizations, and other standard recombinant DNA manipulations were carried out according to published procedures (Sambrook et al., 1989; Cane et al., 1993b). Labeling of the 450-bp probe generated by PCR was performed by random priming using the Pharmacia oligolabeling kit and [α-^{32}P]dCTP (3000 Ci/mmol) to give an oligonucleotide with an activity of 1×10^9 cpm/μg. Genomic DNA was isolated from *Streptomyces* UC5319 by the method (procedure 1) of Hopwood et al. (1985). Methods for PCR, SDS–PAGE, densitometric analysis of dried SDS–PAGE gels, liquid scintillation, and N-terminal sequencing were as previously described (Cane et al., 1993b). 2-D gel electrophoresis was run in the first dimension by denaturing isoelectric focusing using a Bio-Rad Model 150A tube gel electrophoresis apparatus and in the second dimension by SDS–PAGE (12% T, 2.7% C; 20 cm × 20 cm × 0.75 mm). Isoelectric focusing was carried out on a Pharmacia Phast-gel system, using Bio-Rad IEF standards. Electrophoretic transfers were performed on a Hoefer TE 22 Mighty Transfer unit using a Hoefer TE 51 power supply. Sequencing of both single-stranded and double-stranded DNA templates was performed by the dideoxy chain termination method (Sanger et al., 1977) using the Sequenase 2.0 kit (U.S. Biochemicals, Cleveland, OH). Parallel incubations using 7-deaza-dGTP or dITP in place of dGTP were used to resolve ambiguities resulting from compressions. Protein concentrations were determined by the Bradford dye-binding assay (Bradford, 1976; Spector, 1978) (Bio-Rad) with bovine plasma γ-globulin or bovine serum albumin as standards. High-pressure fast protein liquid chromatography was performed with Waters Model 501 pumps and a Model 680 automated gradient controller with an in-line Model 440 absorbance detector or on a Pharmacia Automated FPLC system. *Streptomyces* fermentations were carried out in a New Brunswick G25 gyrotory shaker and in preparative 10-L cylindrical shake flasks on a horizontal shaker. Cell disruption of *Streptomyces* was performed using a Bead-Beater (Biospec Products, Bartlesville, OK) or a Dyno-Mill Model KDL bead mill with batch or continuous flow cells. Preparative centrifugations were done on a Sorvall RC-5 superspeed refrigerated centrifuge at 4 °C using GSA, GS-3, and SS-34 rotors or on a Beckman Model L8-70M ultracentrifuge. Small-scale centrifugations were performed with an Eppendorf Model 5415 centrifuge. Enzyme incubations were carried out in a GCA/Precision Scientific Thelco 184 constant temperature water bath. Enzyme concentration was accomplished with Amicon Model 12 or Model 202 ultrafiltration cells and YM-30 membranes (Amicon, Beverly, MA). Small-scale concentration was performed with YM-30 Centricons (Amicon). Automated Edman degradation was carried out in an ABI Model 477A protein sequencer by Dr. William Lane at the Harvard Microchemistry Facility. Radioisotopic determinations were performed in 5 mL of Opti-Fluor cocktail on a Beckman Model LS 5801 liquid scintillation counter. Sequence homology searches of DNA [(GenBank, Release 78.0 (8/93); EMBL (Modified), Release 35.0 (6/93); PIR-Nucleic, Release 36.0 (3/90)] and protein [(PIR-Protein, Release 37.0 (6/93); SWISS-PROT, Release 26.0 (7/93)] databases utilized the TFastA and Fasta programs of the Genetics Computer Group, Inc., Sequence Analysis Software Package, version 7.3-UNIX. (Devereux et al., 1984).

Pentalenene Synthase Assay. Native pentalenene synthase was assayed as previously described (Cane et al., 1984), while recombinant enzyme was assayed by a modification of the earlier procedure using 50 mM Tris-HCl, 5 mM $MgCl_2$, 5 mM 2-mercaptoethanol, 1 mM EDTA, 0.2 mM PMSF, 20% (v/v) glycerol, pH 8.2 at 30 °C, and 2 μM [1-^{3}H]FPP (**1**) in a total volume of 250 μL in 12 × 75 mm glass test tubes. Assays were initiated by the addition of 10 μL of protein solution (diluted in buffer) to buffer and FPP, and assay mixtures were overlaid with 0.5 mL of hexane. After 5 min the incubations were terminated by the addition of 100 μL of 100 mM EDTA, pH 8.0. The samples were vortex mixed for 20 s, and a further 0.5 mL of hexane was added. The hexane layer was loaded onto a small silica gel column and expelled with a nitrogen stream into a scintillation vial containing 5 mL of Opti-Fluor. The column was washed with a further 0.5 mL of hexane. Recombinant protein which had been purified by anion-exchange chromatography on Q-Sepharose was free of phosphatase activity, and the hexane extract was directly transferred into the scintillation vial.

The buffer for kinetic studies was identical to the assay buffer except for the addition of 0.2 mM benzamidine–HCl and the substitution of 1 mM DTT for 2-mercaptoethanol. Assays were conducted in duplicate for all concentrations of

Table 1: Purification of Pentalenene Synthase from *Streptomyces* UC5319

purification step	total protein (mg)	total activity (nmol/min)	specific activity (nmol/min/mg)	purification (x-fold)	yield (%) (net yield)
70% AS ppt	3635	2370	0.67	1	
45% AS	884	1700	1.92	2.9	72 (72)
methyl agarose	185	1330	7.2	10.8	78 (56)
DEAE	75.6	1200	15.9	23.7	90 (51)
Sephacryl S-200	26.6	842	37.9	56.6	69 (36)
Ado-Sepharose	3.2	467	143	213	55 (19.7)
methyl HIC	0.41	133	324	484	29 (5.6)

[1-^{3}H]FPP (**1**) (0.174, 0.348, 0.700, 1.05, 1.40, and 1.75 μM final concentration). Assay mixtures were prewarmed for 2 min before the reaction was initiated with 50 μL of enzyme (30 ng of protein per assay) and were overlaid with 0.5 mL of hexane. Blank assays were initiated using buffer. After 5 min the reaction was terminated by addition of 200 μL of 100 mM EDTA, pH 8.0. The samples were vortex mixed, and the hexane extracts were analyzed as described above. Inhibition studies on inorganic pyrophosphate were carried out in the presence of 0, 20, and 50 μM disodium dihydrogen pyrophosphate. In each case the data was analyzed by double-reciprocal plots and fitted directly to the relevant kinetic expressions by nonlinear regression (Cane et al., 1990b) to determine the steady-state kinetic parameters K_m, K_I, and V_{max}. Investigations into the effect of metal ions were carried out in an identical manner, except that the metal ion was added to each assay tube immediately before the preincubation period to avoid non-enzymatic hydrolysis of FPP.

Purification of Pentalenene Synthase from Streptomyces UC5319 (Table 1). Vegetative cultures (100 mL) of *Streptomyces* UC5319 were grown for 2.5 days in 500-mL shake flasks without baffles at 26 °C and 300 rpm (Cane et al., 1984). The vegetative mycelia were used to inoculate four 10-L cylindrical flasks (inoculum 1%, v/v) fitted with glass inlet tubes for aeration and each containing 4 L of production medium (8 g of NaCl, 5 g of $CaCO_3$, 16.75 g of casein (acid hydrolysate), 13 g of bactodextrose, 8 g of blackstrap molasses, 80 g of dextrin, and 4 L of nanopure water, pH 7.4). The flasks were shaken on a reciprocating shaker at 25 cycles per min and 32 °C with aeration at 6 L/min (8 psi). After 2.5 days, active aeration was discontinued and shaking was continued for an additional 24 h (Cane & Pargellis, 1987). After a total of 3.5 days, the mycelia were collected by centrifugation using a GS-3 rotor at 11000*g* for 10 min at 4 °C. The packed mycelia were successively washed with water, 1 M KCl, 0.8 M NaCl and degassed *Streptomyces* breaking buffer (50 mM KH_2PO_4, 10% glycerol, 5 mM 2-mercaptoethanol, and 5 mM EDTA, pH adjusted to 7.2 with 10% NaOH). The yield of washed mycelia was 245–270 g from 16 L of culture.

Cells (245–270 g) were suspended in 600–700 mL of breaking buffer in preparation for cellular disruption. Immediately prior to breaking, PMSF was added to 0.5 mM to protect against proteolysis. Cell breakage was achieved using a DYNO-Mill flow-through bead mill equipped with a 600-mL grinding chamber containing 520 mL of 0.1-mm glass beads and cooled with −20 °C glycerol (Neslab HX-150 series cooling glycerol recirculators). A thick suspension of cells was continuously pumped into the grinding chamber using a peristaltic pump at around 3 L/h. After passage of the cellular suspension into the DYNO-Mill, the grinding chamber was flushed with fresh breaking buffer until the effluent was clear. The cellular homogenate was clarified by centrifugation at 11000*g* for 20 min at 4 °C, and trypsin inhibitor (5 mg), pepstatin A (5 mg), and leupeptin (1 mg) were added. All procedures subsequent to the cell breakage were performed at 4 °C.

70% Ammonium Sulfate Precipitation. Fractionation was performed by addition of crushed, solid ammonium sulfate to 70% saturation (0.472 g/mL) over 1 h with stirring. The extract was stirred slowly for an additional 2 h and centrifuged at 16000*g* for 20 min. The precipitated protein was resuspended in a minimum volume of degassed column buffer (50 mM Tris-HCl, pH 8.2, 20% (v/v) glycerol, 0.02% sodium azide, 1 mM EDTA, and 5 mM 2-mercaptoethanol) and 45% saturated ammonium sulfate (0.277 g/mL). The suspension was centrifuged at 104000*g* for 60 min.

Hydrophobic-Interaction Chromatography. Methyl agarose gel equilibrated in column buffer and 45% saturated ammonium sulfate was poured into a glass column (2.5 × 20 cm, 100 mL). The protein sample in 45% saturated ammonium sulfate solution was applied to the column at ca. 25 mL/h. The column was washed with ca. 25 mL of column buffer at 45% saturated (2.1 M) ammonium sulfate. A linear gradient of 2.1–0.0 M ammonium sulfate over 600 mL was used to elute pentalenene synthase, and fractions of 300 drops were collected and assayed for pentalenene synthase activity and protein concentration. The fractions containing pentalenene synthase activity (110–140 mL; 0.55–0.40 M ammonium sulfate) were pooled.

DEAE Sephadex A-25 Anion-Exchange Chromatography. The sample from the methyl agarose pool was dialyzed against column buffer with Spectra/POR membrane tubing (Spectrum Medical Industries; diameter, 10 cm; 6000–8000 MW cutoff) and loaded onto a DEAE-Sephadex A-25 column (2.5 × 20 cm, 100 mL). After equilibration with 25 mL of column buffer at 25 mL/h, the protein was eluted with a linear gradient of 0–0.5 M NaCl over 700 mL. Fractions (300 drops) were assayed for synthase activity and protein concentration. The fractions containing pentalenene synthase (150–190 mL, 0.22–0.35 M NaCl) were pooled.

Gel Filtration. Enzyme from the DEAE column was concentrated by ultrafiltration using YM-30 ultrafiltration membranes at 40 psi of nitrogen gas to 2–4 mL. The concentrated synthase solution was applied to a Sephacryl S-200 column (2.5 cm × 88 cm, 430 mL) and eluted with column buffer (64 mL/h). Each fraction (250 drops) was assayed for protein content and pentalenene synthase activity, and fractions 26–29, with the highest specific activity of pentalenene synthase, were combined.

Ado-Sepharose Column Chromatography. A 10–15-mL portion of the enzyme from the Sephacryl gel filtration was applied to a column of Ado-Sepharose (1.5 × 8 cm, 14 mL) preequilibrated with column buffer. The column was washed with 80 mL of column buffer followed by a linear gradient of 0–0.4 M KCl over 180 mL at 10 mL/h. Fractions (200 drops) were assayed for pentalenene synthase activity and protein concentration. The synthase activity eluted at

approximately 0.2 M KCl, immediately following a major protein contaminant at 0.13 M KCl.

Hydrophobic Interaction Chromatography. Hydrophobic interaction chromatography was carried out using a prepacked MP7 HIC FPLC column (50 × 7.8 mm, Bio-Rad) on a Waters HPLC system. The synthase from the Ado-Sepharose column, concentrated using Centricon YM-30 membranes (4–5 times), was added to an equal volume of saturated ammonium sulfate column buffer and then subjected to methyl HIC fast protein liquid chromatography. The column was washed with high-salt buffer, and protein eluted with a 45–0% ammonium sulfate gradient over 35 mL at 0.5 mL/min, followed by 7 mL of column buffer. Fractions of 1 mL were collected. Pentalenene synthase activity eluted in 0.4 M ammonium sulfate.

Estimation of M_r of Native Pentalenene Synthase by Non-Denaturing Gel Electrophoresis. Non-denaturing gel electrophoresis (Ornstein, 1964; Davis, 1964) was carried out with 7 cm × 8 cm × 0.75 mm polyacrylamide gels (12% T, 2.7% C) using egg albumin (45 000) and bovine serium albumin (66 200) as molecular weight standards. Homogeneous native pentalenene synthase was electrophoresed at 30 mA/cm^2 constant current until the bromophenol blue tracking dye had migrated almost to the bottom of the gel. Gel slices of 3 mm each were excised and incubated in 150 μL of column buffer at 4 °C for 4–5 h, and then portions were assayed for pentalenene synthase activity. The remaining gel was stained with Coomassie blue and silver to visualize pentalenene synthase, BSA, and ovalbumin.

N-Terminal Sequencing of Native Pentalenene Synthase. Pentalenene synthase from Ado-Sepharose column chromatography was subjected to SDS–PAGE (12.5%, 20 × 20 cm, 6 μg per lane), transferred to an Immobilon-P PVDF membrane by electroelution (Matsudaira, 1987), and subjected to N-terminal sequencing.

Internal Amino Acid Sequence Analysis of Native Pentalenene Synthase. Pentalenene synthase purified by methyl HIC FPLC was subjected to SDS–PAGE (12.5%, 20 × 20 cm, 3 μg per lane), transferred to nitrocellulose membranes by electroelution, and submitted to the Harvard Microchemistry Facility for tryptic digestion, HPLC purification, and amino acid sequence analysis of the major tryptic peptides.

PCR Amplification of Genomic DNA. For the PCR amplification of portions of the pentalenene synthase gene, two primers were designed based on the N-terminal sequence and one of the internal tryptic peptide sequences, taking into account known *Streptomyces* codon preferences (Bibb et al., 1984; Wright & Bibb, 1992). A typical 100-μL reaction mixture contained 10× buffer A (10 μL consisting of 100 mM Tris-HCl (pH 8.3), 500 mM KCl, 15 mM $MgCl_2$, and 0.1% (w/v) gelatin), sterile water (32.4 μL), 25% glycerol (32.0 μL), a mixture of four dNTPs, each at 1.25 mM (16.0 μL), 10 mM $MgCl_2$ (3.0 μL), primer **F-1** (3.0 μL, 57 ng/μL, 20 pmol), primer **R-3** (2.1 μL, 82 ng/μL, 20 pmol), *Streptomyces* UC5319 genomic DNA (1 μL, 50 ng), and Taq DNA polymerase (0.5 μL, 2.5 units). The reactions were performed in 0.5-mL Eppendorf tubes which were immersed in ice during addition of all the reagents. The reagents were added in the order listed, mixed, and centrifuged. The enzyme was added last. The mixture was overlaid with mineral oil (75 μL) and immediately placed in the thermal cycler before the following program was started: (a) 95 °C, 3.0 min (denaturation); (b) 73 °C, 2.5 min (annealing and extension), and 95 °C, 1.5 min (denaturation), 35 cycles; (c) 73 °C, 7.0 min (final extension); (d) 4 °C (hold). The reaction product was examined by running 8 μL of the reaction mixture on a 1.5% agarose minigel. (The same reaction conditions were used to generate the 300-bp PCR product from primers **F-1** and **R-2**. Both reactions were very sensitive to reaction conditions and buffer composition and had an absolute requirement for 8% glycerol.) The typical yield was ca. 500 ng of DNA per 100 μL of reaction mixture. The combined products of 8–10 reactions were extracted twice with 1 vol of chloroform to remove trace amounts of mineral oil and precipitated by the addition of 0.1 vol of 3 M sodium acetate and 2 vol of ethanol. The pellet was collected, washed with 70% ethanol, dissolved in 60 μL of TE buffer (Hopwood et al., 1985), and purified on a 10% non-denaturing polyacrylamide gel. The band containing DNA of the appropriate length was excised and incubated overnight at 27 °C in TE buffer, and the DNA was purified by passage through a Waters C-18 Sep-Pak column (Millipore). After lyophilization, the 450-bp PCR product was taken up in 100 μL of TE buffer.

Preparation of a Sized Genomic Plasmid Library. On the basis of the hybridization of the 450-bp probe to a single 2.8-kb *Bgl*II fragment at 0.2× SSC at 70 °C, 37.4 μg of *Streptomyces* UC5319 genomic DNA was digested with a total of 40 units of *Bgl*II at 37 °C for 4 h. The digested DNA was subjected to electrophoresis on a 1.1% agarose gel, and the DNA between 2.5 and 4.0 kb was isolated and redissolved in TE buffer (30 ng/uL).

The sized *Bgl*II fragments (90 ng) were ligated into dephosphorylated, *Bam*HI-digested pUC18, using a vector: insert ratio of 1:1, by incubation with T4 DNA ligase at 16 °C overnight, and the ligation mix was used to transform library efficiency competent cells of *E. coli* DH5α. The resulting transformation mixture was plated out on 10 LB agar plates containing 25 μg/mL ampicillin, 0.1 mM IPTG, and 0.004% X-gal to give 7380 white colonies, which were blotted to Gene Screen nylon membranes (Sambrook et al., 1989) and screened by hybridization with the ^{32}P-labeled 450-bp probe at 0.2× SSC and 70 °C, as for the genomic Southerns. Positively hybridizing colonies were replated to give single colonies and rescreened by Southern blotting of *Eco*RI-linearized plasmid minipreps (Kieser, 1984). A single clone was isolated, which harbored a 5.6-kb plasmid, designated pCRL1000, carrying a 2.8-kb *Bgl*II insert which hybridized strongly to the 450-bp probe. Restriction digests indicated that this insert contained only the first ca. 850 bp of the pentalenene synthase gene.

Sequencing of the 2.8-kb BglII Insert. For sequencing, the *Bgl*II insert was subcloned into the phagemids pTZ18R and pTZ19R. pCRL1000 was digested with *Eco*RI and *Hin*dIII followed by *Sca*I to cleave the pUC18-derived DNA into two smaller fragments of 1.7 and 0.9 kb. The 2.8-kb fragment was isolated by electrophoresis on 0.7% agarose and ligated with T4 DNA ligase into the corresponding *Eco*RI/*Hin*dIII sites of pTZ18R and pTZ19R, and the individual ligation mixes were used to transform *E. coli* NM522. The resulting plasmid constructs were designated pCRL1201 and pCRL1100, respectively. The presence of the desired inserts was verified in each case by PCR analysis using primers **F-1** and **R-3**. Growth of recombinant strains of *E. coli* NM522 harboring phagemids pCRL1201 and pCRL1100 in the presence of helper phage M13KO7, according to the supplier's protocols, generated single-stranded DNA. ssDNA corresponding to each strand of the insert was sequenced by primer walking from 240 bp upstream of the ATG start codon to the internal *Bgl*II site.

Construction of Streptomyces UC5319 EMBL3 Library. *Streptomyces* UC5319 genomic DNA was partially digested with *Sau*3a and size fractionated by 10–40% sucrose gradients. Fractions containing DNA fragments of 9–23 kb were pooled and ligated with λEMBL3 DNA which had been previously digested with *Eco*RI and *Bam*HI. The ligation mixture was packaged using the Packagene λ DNA system. The packaged phage were then grown on the restrictive P2 host *E. coli* NM539 and amplified.

The genomic library of *Streptomyces* UC5319 in λEMBL3 was screened by plaque hybridization with the labeled 450-bp PCR probe. The library was transfected into *E. coli* NM539 and transferred to Colony/Plaque screen hybridization membranes according to the recommendations of the manufacturer. Hybridization was performed at 65 °C for 16–18 h. Positive plaques containing the desired clones were isolated and plaque purified. Restriction maps and Southern blots of these clones using the 450-bp probe were compared to genomic DNA and to the previously cloned 2.8-kb *Bgl*II fragment containing ca. 80% of the pentalenene synthase gene. A clone with an insert of 15.5 kb of *Streptomyces* UC5319 genomic DNA that contained a 3.9-kb *Nco*I fragment which positively hybridized to the 450-bp probe was designated as λEMBL3-4-13.

The 3.9-kb *Nco*I fragment of λEMBL3-4-13 was subcloned into pGEM5Zf(–). The λEMBL3-4-13 DNA (1 μg) was digested with *Nco*I (10 units) for 2 h at 37 °C. The reaction was terminated by heating to 65 °C for 20 min, followed by removal of the protein by centrifuging the reaction mixture through an Ultrafree Probind filter. pGEM5Zf(–) was digested with *Nco*I in a similar manner and dephosphorylated with calf intestinal alkaline phosphatase. The *Nco*I-digested λEMBL3-4-13 DNA was combined with vector (100 ng) and ligated with T4 DNA ligase (8 units) at 16 °C for 17 h. The DNA was combined and heated to 55 °C for 5 min and placed on ice prior to the addition of enzyme and ATP. The ligated DNA was used to transform *E. coli* XL1Blue (Ausubel et al., 1987). Transformants were selected on LB agar plates containing carbenicillin (100 μg/mL) and IPTG/X-gal. Twenty-one white colonies were selected and analyzed by restriction mapping and Southern blottting. One clone was found to contain the 3.9-kb *Nco*I fragment and was designated pE34-135. Restriction enzyme digestion of pE34-135 with *Nco*I showed that the 10-kb plasmid contained not only the 3.9-kb fragment of interest but a second *Streptomyces*-derived *Nco*I fragment of ca. 3 kb. Since this extra DNA did not interfere with either subsequent sequencing or PCR amplification, no attempt was made to remove it.

Sequencing of pE34-135. The remainder of the pentalenene synthase ORF was sequenced by primer walking using [α-^{35}S]-dCTP and pE34-135 DNA containing the 3.9-kb *Nco*I fragment. Template dsDNA, obtained by plasmid maxipreps followed by 13% PEG precipitation, was subjected to alkali denaturation or heat denaturation prior to sequencing. In some cases it was advantageous to reduce sequencing gel banding artifacts resulting from premature termination by a final treatment with terminal deoxynucleotidyl transferase and dNTPs (Fawcett & Bartlett, 1990, 1991).

Construction of the Pentalenene Synthase Overproducer E. coli BL21(DE3)/pZW05. Plasmid pE34-135 was digested with *Nco*I, and the resulting DNA was used directly as the template for PCR, using primers **PS-1** and **PS-2**. Conditions for PCR amplification and product purification were identical to those previously described (Cane et al., 1993b) except that the annealing step in each cycle was carried out at 65 °C for 2 min. The typical yield was 1–2 μg of amplified DNA per 100 μL of reaction mixture. The purified DNA was digested sequentially with *Eco*RI and *Hin*dIII, repurified using Pro-Bind and Ultrafree-MC 30 000 MW filter sets, and ligated with T4 DNA ligase into pLM1 previously treated with *Eco*RI and *Hin*dIII. The ligation product was used to transform competent cells of *E. coli* XL1-blue, and the transformation mix was plated out on LB agar plates containing 60 μg/mL of ampicillin. Plasmid minipreps on several of the resulting single colonies led to isolation of pZW05 containing the desired insert, as established initially by agarose gel electrophoresis of linearized DNA. pZW05 was then used to transform *E. coli* BL21(DE3).

Expression of Pentalenene Synthase by E. coli BL21(DE3)/pZW05. Six colonies of the above-generated transformants were screened for pentalenene synthase and found to have the desired cyclase activity. For each assay, a single colony of *E. coli* BL21(DE3)/pZW05 was used to inoculate 5 mL of LB medium containing 100 μg/mL ampicillin in a 15-mL tube. The culture was incubated overnight at 37 °C to an OD of 1.0. IPTG was then added to a final concentration of 0.5 mM, and the cultures were incubated further at 37 °C for 2 h. A 1.5-mL sample of the culture broth was withdrawn and centrifuged for 1 min at 14 000 rpm. The pellet was taken up in 75 μL of lysis buffer (50 mM Tris, pH 8.0, 1 mM EDTA, and 100 mM NaCl) and 5 μL of a solution of lysozyme (10 μg/μL). After being vortexed, the mixture was incubated at room temperature for 20 min and then frozen at –80 °C overnight. The cell lysate was thawed in warm water, and a 10-μL aliquot was assayed for pentalenene synthase activity. The whole cell lysate (70 μL) was analyzed by SDS–PAGE on a 10% gel.

GC–MS Analysis of Pentalenene Formation. The cell lysate of *E. coli* BL21(DE3)/pZW05 (50 μL) was incubated with [1-^{3}H]FPP (16.7 mCi/mmol, 3.60×10^5 dpm) in buffer T (1 mL, pH 8.4) at 30 °C for 1 h. The olefinic product was extracted by pentane and purified by passage through a 4-cm silica gel column in a Pasteur pipet. The recovered activity was 8.9×10^4 dpm, corresponding to 2.5 nmol of product (25% conversion). The extract was concentrated to ca. 50 μL on a rotatory evaporator at 0 °C, and an estimated one-fifth of the concentrate was used for GC–MS analysis on a Hewlett-Packard 5898 gas chromatograph mass spectrometer using a 12 m × 0.2 mm × 0.33 μm thick crosslinked methyl silicone gum capillary column in electron ionization mode and a temperature program of 80–250 °C over 13 min. Enzymatically generated pentalenene (*m/e* 204) eluted at 5.99 min at a temperature of 158 °C and was identical in retention time and mass fragmentation pattern to authentic (±)-pentalenene[2] (Cane & Tillman, 1983; Ohfune et al., 1976; Misumi et al., 1979).

Purification of Overexpressed Pentalenene Synthase. *E. coli* cells from 4 × 100 mL of LBA media were harvested 3 h postinduction and washed with 0.9% (w/v) NaCl. They were then resuspended in lysis buffer (50 mM Tris-HCl, pH 8.2, 10% (v/v) glycerol, 1 mM EDTA, 1.0 mM PMSF, 1.0 mM benzamidine, 5 mM 2-mercaptoethanol, and 10 mM $MgCl_2$) and stored at –80 °C. The thawed cell suspension was diluted to 20 mL with buffer and stirred in a 50-mL glass beaker in the presence of lysozyme (1 mg/mL) and DNase 1 (0.03 mg/mL). After 30 min, 200 μL of 10% (v/v) aqueous Triton X-100 was added and the extract was stirred for a

[2] We thank Mr. Clifford M. Bryant for preparation of the reference sample of (±)-pentalenene.

further 30 min. The extract was then sonicated with the microtip probe (5 min; power, 5; 50% duty cycle) and clarified by centrifugation for 10 min at 12000*g*.

The protein extract was diluted with 20 mL of deionized water, loaded onto a Q-Sepharose column (2.5 × 6 cm, 30 mL) equilibrated in column buffer supplemented with 0.2 mM PMSF and 0.2 mM benzamidine, and eluted with a 0–0.5 M NaCl gradient over 300 mL. Fractions 34–40 were pooled (34 mL, 0.175–0.25 M NaCl) and stored at –80 °C. These fractions were concentrated in 50- and 10-mL Amicon stirred cells fitted with YM-30 membranes to ca. 1 mL (3.7 mg/mL), and 2 × 100 mL was further purified by passing through an Affigel Blue column (0.7 mL) equilibrated with column buffer. Fractions were analyzed by SDS–PAGE.

Fractions containing pentalenene synthase from the Affigel blue columns were pooled (ca. 4 mL) and loaded onto a Mono-Q 5/5 anion-exchange column. Protein was eluted with a 0–1 M NaCl gradient over 25 mL, and fractions of 1 mL were collected. Pentalenene synthase, which eluted at 0.4 M NaCl, was homogenous after this purification step, as judged by SDS–PAGE.

N-Terminal Sequencing of Recombinant Pentalenene Synthase. Purified pentalenene synthase was electrophoresed on a 12.5% SDS–PAGE minigel (6 μg per well), blotted to a PVDF membrane at 20 V for 2 h, and submitted for N-terminal sequencing.

RESULTS

Purification of Pentalenene Synthase from Streptomyces UC5319. Although the previously developed partial purification of pentalenene synthase (Cane & Pargellis, 1987) had been adequate for initial characterization of the cyclase, it was evident that a simple scale-up and improvement of this method would be incapable of providing the necessary quantities of homogeneous enzyme for further study. Several alternative purification strategies which took advantage of the relative hydrophobicity of pentalenene synthase were therefore explored, resulting ultimately in the development of a six-step protocol which provided homogeneous pentalenene synthase in 5.5% overall yield (Table 1).

The mycelia from 16 L of a 3.5-day fermentation culture of *Streptomyces* UC5319 were disrupted by grinding with glass beads, and the supernatant from the resulting crude cell extract was precipitated with 70% saturated ammonium sulfate. The redissolved pellet was subjected to hydrophobic-interaction chromatography on methyl agarose, resulting in the removal of significant quantities of highly pigmented material which were retained on the column and which otherwise interfered with purification on other media. Active pentalenene synthase could not be recovered from ethyl agarose or longer alkyl agaroses. Fractionation by anion exchange on DEAE-Sephadex followed by gel filtration on Sephacryl S-200 effected an overall 60-fold enrichment. The fractions with the highest pentalenene synthase activity, which showed a major protein band by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE, Figure 1) between 40 000 and 45 000, were pooled and further purified by affinity chromatography on adenosyl-(1,3-diaminopropyl)-Sepharose. The latter step served to remove a major protein impurity which could not be resolved from pentalenene synthase by other chromatographic methods. 2-D gel electrophoretic analysis of the most active fractions revealed the presence of three components, of which the two major proteins displayed very similar molecular weight (M_r 42.5 kDa) and p*I* (5.0–5.1). Final purification was achieved by hydrophobic-interaction HPLC using a methyl HIC column. The recovered pentalenene synthase was homogeneous by SDS–PAGE using silver staining and displayed a specific activity of 325 nmol of pentalenene min^{-1} (mg of protein)$^{-1}$, repesenting a ca. 490-fold purification relative to the original 70% ammonium sulfate precipitate. The entire purification procedure afforded a total of ca. 400 μg of purified pentalenene synthase from 16 L of *Streptomyces* UC5319 culture.

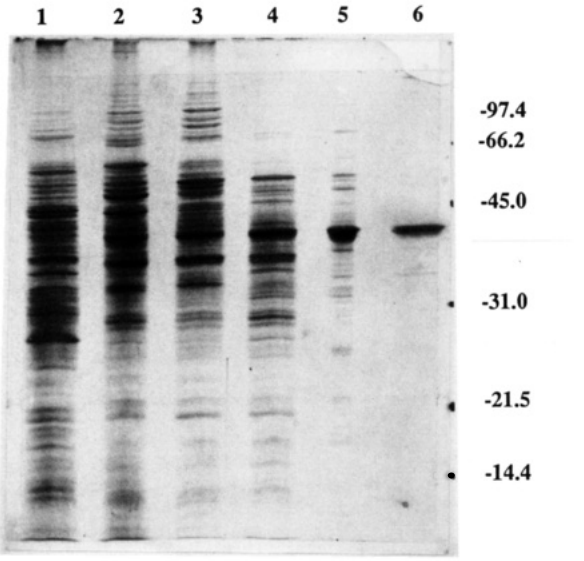


FIGURE 1: SDS–PAGE (12%) analysis of purification of pentalenene synthase from *Streptomyces* UC5319 (molecular weight standards (in kDa): rabbit muscle phosphorylase B, 97.4; BSA, 66.2; hen egg white ovalbumin, 45.0; bovine carbonic anhydrase, 31.0; soybean trypsin inhibitor, 21.5; hen egg white lysozyme, 14.4). Lane 1, 45–75% AS pellet, redissolved in 45% AS; lane 2, methyl agarose column pool; lane 3, DEAE-Sephadex column pool; lane 4, Sephacryl S-200 pool; lane 5, Ado-Sepharose column pool; lane 6, methyl HIC FPLC pool.

Characterization of Pentalenene Synthase. To reconfirm that the purified synthase was indeed generating pentalenene, the methyl HIC-purified enzyme was incubated with [1-^{3}H]-FPP and the resulting hydrocarbon, after dilution with unlabeled (±)-pentalenene, was reacted with OsO_4 (Cane & Tillman, 1983; Cane et al., 1984). Each of the derived diastereomeric *cis*-diols recrystallized to constant activity, thereby confirming the identity of the enzymatic cyclization product as pentalenene. On the basis of electrophoretic mobility on SDS–PAGE, the molecular weight of the purified pentalenene synthase was assigned as M_r 41.7 ± 0.5 kDa. When analyzed by non-denaturing polyacrylamide gel electrophoresis, the maximum pentalenene synthase activity was centered at ca. 50 kDa, thereby establishing that the native pentalenene synthase is a monomer.[3]

Amino Acid Sequence Determination. The N-terminal sequence of native pentalenene synthase was determined by automated Edman degradation for the first 27 amino acids of the methyl HIC-purified protein, which had been subjected to SDS–PAGE and blotted to a PVDF membrane (Table 2). Analogous blotting to nitrocellulose and digestion with trypsin, followed by microsequencing of the derived HPLC-purified peptides, gave sequences for three tryptic fragments (Table 2). Neither the N-terminal nor the internal tryptic peptides

[3] Electrophoresis of less pure preparations of pentalenene synthase on native gels with varying acrylamide concentrations and analysis of the relative gel mobility by Ferguson plots (Ferguson, 1964; Hendrick & Smith, 1968; Chrambach & Rodbard, 1971) indicated an apparent M_r of 51 kDa.

Table 2: N-Terminal (A) and Internal Tryptic Sequences (B–D) of Pentalenene Synthase and Sequences of Derived Oligonucleotides (**F-1**, **R-1**, **R-2**, **R-3**)[a]

N-Terminus

```
      G           G                *
5'-CCC-CAG-GAC-GTC-GAC-TTC-CAC-ATC-CC-3'          F-1
```

A. Pro-Gln-Asp-Val-Asp-Phe-His-Ile-Pro-Leu-Pro-Gly-Arg-Gln-Ser-Pro-Asp-His-Ala-Arg-Ala-Glu-Ala-Glu-Gln-Leu-Ala

```
3'-TG-CGC-GCC-CGC-CTC-CGC-CTC-GTC-GA-5'      R-1
      G   G   G       G
```

Tryptic Fragments

B. Phe-Tyr-Pro-His-Ala-Thr-Gly-Ala-Asp-Leu-Asp-Leu-Gly-Val-Asp-Leu-Met

C. Gln-Leu-Thr-Asp-Gln-Val-Ala-Ala-Ala-Leu-Asp-Gly-Pro-Leu-Pro-Asp-Thr-Ala-Pro-Pro-Ile-Ala-Trp-Gly-Phe

```
3'-TGG-CGC-GGC-GGC-TAG-CGC-ACC-CCG-AA-5'      R-2
      G   G   G       G ***
```

D. Asn-Tyr-Phe-Asp-Gly-Tyr-Val-Asp-Glu-Ala-Glu-Ser-Arg

```
3'-ATG-AAG-CTG-CCG-ATG-CAC-CTG-CTC-CG-5'      R-3
          *       G       *
```

[a] Asterisks denote regions of mismatch between the oligonucleotide and the actual sequence. In tryptic fragment C, His-118 has been misidentified as Trp.

showed any significant homologies to known proteins in the DNA or protein databases.

Molecular Cloning of Pentalenene Synthase. PCR was used to amplify specific 77-, ca. 350-, and ca. 450-bp segments of *Streptomyces* UC5319 genomic DNA by using degenerate oligonucleotide primers chosen on the basis of preferred codon usage in *Streptomyces* (Bibb et al., 1984; Wright & Bibb, 1992) (Table 2). For amplification of the 450-bp product, the forward (sense) primer (**F-1**) was a 26-mer based on the first nine amino acids of pentalenene synthase and containing [C,G] degeneracies in the third base of the first and fourth codons. The reverse (antisense) primer (**R-3**) was a 26-mer based on the Tyr-Phe...Glu-Ala segment of one of the tryptic peptides and containing a single [G,C] mixture for the Gly anticodon. Partial sequencing of the 450-bp product confirmed that the PCR product encoded the corresponding N-terminal amino acid and internal tryptic fragments.

The 450-bp PCR product was next used to probe restriction digests of genomic DNA obtained from *Streptomyces* UC5319. In Southern blots at high stringency (0.2× SSC, 70 °C), ^{32}P-oligolabeled 450-bp PCR product hybridized strongly to a number of single-restriction fragments, including 2.8-kb *Bgl*II, 3.9-kb *Nco*I, and 7–9-kb *Bcl*I fragments. Both *Eco*RI and *Hin*dIII digests gave rise to bands of >23 kb. On the basis of these observations, a partial genomic library was constructed by inserting size-fractionated (2.5–4.0 kb) *Bgl*II-digested *Streptomyces* DNA into the *Bam*HI site of pUC18 and transforming *E. coli* DH5α. Colony hybridization and screening with the 450-bp PCR probe led to isolation of a single clone harboring a plasmid (pCRL1000) with a 2.8-kb insert which hybridized strongly to the probe DNA in Southern blots. Restriction endonuclease analysis and comparison with the partial sequence of the 450-bp PCR product indicated that this insert contained only approximately the first 850 bp of the putative pentalenene synthase gene (data not shown). This inference was confirmed by subcloning the 2.8-kb insert as an *Eco*RI–*Hin*dIII fragment into both pTZ18R and pTZ19R. These constructs were each sequenced from the *Hin*dIII site to ca. 240 bp upstream of the pentalenene synthase start codon (total ca. 1080 bp). As expected, on the basis of the apparent protein molecular weight and the inferred size of the gene, no stop codon was found between the start codon and the *Bgl*II restriction site located 835 bp downstream.

In order to obtain the entire pentalenene synthase gene, an EMBL3 library was constructed using sized (9–23 kb) DNA obtained from a partial *Sau*3a digest of *Streptomyces* UC5319 genomic DNA. Screening with the previously used 450-bp PCR probe led to the isolation of a strongly hybridizing plaque containing a 15.5-kb insert. Limited restriction mapping of this insert combined with Southern hybridization experiments utilizing the 450-bp probe identified a 3.9-kb *Nco*I fragment which was identical in size to the *Nco*I restriction fragment previously observed in genomic Southerns. The 3.9-kb *Nco*I fragment prepared from the plaque DNA was therefore subcloned into the appropriate site of pGEM-5Zf(−). The remaining portion of the pentalenene synthase gene downstream of the internal *Bgl*II site was then sequenced by primer walking on both strands by the dideoxy method.

The combined sequence information (Figure 2) revealed that the pentalenene synthase gene was encoded by an ORF of 1014 bp, corresponding to 337 amino acids with a predicted M_D of 38 002, compared to the previously estimated M_r of 41 700 deduced from the behavior of the native *Streptomyces* protein on SDS–PAGE. The sequence of the ORF, which lies 12 nt downstream of a putative *Streptomyces* ribosome binding site (rbs, AAGGA) (Bibb & Cohen, 1982) and upstream of a 9-bp inverted repeat (presumptive transcription terminator) (Deng et al., 1987; Pulino & Jimenez, 1987; Rosenberg & Court, 1979), conformed to established patterns of GC bias and preferred codon usage in *Streptomyces* (Bibb et al., 1984; Wright & Bibb, 1992) (Figure 3). The deduced amino acid sequence included the previously determined 27

```
-240  ggaggaactgcggcgcgtggccggcctgatcgagggcgccggcggccgggactgggcgctggccgaggccgcccggcggatggccctggc -151
-150  cgaggacgcgttgcgcggtgtcgagctggacgccggcggcgcgggatgaactcctcgcgctcgggcggttcgtcgtcgaccgcgaggtct  -61
 -60  gagaccgggccggccccctacccgaatgtccctccgcaacccacccccaaggatttcctcATGCCCCAGGACGTCGACTTCCACATACCC   30
                                                                    M  P  Q  D  V  D  F  H  I  P     10

  31  CTGCCGGGTCGGCAGAGCCCGGATCACGCGCGGGCCGAGGCCGAGCAGCTCGCCTGGCCGCGGTCGCTCGGGCTCATCAGGTCCGACGCG  120
  11  L  P  G  R  Q  S  P  D  H  A  R  A  E  A  E  Q  L  A  W  P  R  S  L  G  L  I  R  S  D  A     40

 121  GCGGCCGAGCGTCATCTGCGCGGCGGCTACGCCGACCTGGCCTCCCGTTTCTACCCCCACGCCACCGGCGCCGACCTGGACCTGGGCGTC  210
  41  A  A  E  R  H  L  R  G  G  Y  A  D  L  A  S  R  F  Y  P  H  A  T  G  A  D  L  D  L  G  V     70

 211  GACCTGATGTCGTGGTTCTTCCTCTTCGACGACCTCTTCGACGGGCCGCGCGGCGAGAACCCGGAGGACACCAAACAGCTCACCGACCAG  300
  71  D  L  M  S  W  F  F  L  F  D  D  L  F  D  G  P  R  G  E  N  P  E  D  T  K  Q  L  T  D  Q    100

 301  GTCGCGGCGGCGCTGGACGGCCCGCTCCCCGACACCGCGCCGCCCATCGCCCACGGCTTCGCCGACATCTGGCGGCGCACCTGTGAGGGC  390
 101  V  A  A  A  L  D  G  P  L  P  D  T  A  P  P  I  A  H  G  F  A  D  I  W  R  R  T  C  E  G    130

 391  ATGACCCCCGCCTGGTGCGCGCGCAGCGCCCGGCACTGGCGGAACTACTTCGACGGGTACGTCGACGAAGCGGAGAGCCGGTTCTGGAAC  480
 131  M  T  P  A  W  C  A  R  S  A  R  H  W  R  N  Y  F  D  G  Y  V  D  E  A  E  S  R  F  W  N    160

 481  GCCCCCTGCGACTCCGCCGCGCAGTACCTGGCCATGCGCCGGCACACCATCGGGGTGCAGCCGACGGTGGACCTCGCCGAGCGCGCGGGC  570
 161  A  P  C  D  S  A  A  Q  Y  L  A  M  R  R  H  T  I  G  V  Q  P  T  V  D  L  A  E  R  A  G    190

 571  CGCTTCGAGGTGCCGCACCGGGTCTTCGACAGCGCCGTGATGTCCGCGATGCTCCAGATCGCCGTCGACGTCAACCTGCTGCTCAACGAC  660
 191  R  F  E  V  P  H  R  V  F  D  S  A  V  M  S  A  M  L  Q  I  A  V  D  V  N  L  L  L  N  D    220

 661  ATCGCCTCCCTGGAGAAGGAGGAGGCCCGCGGCGAGCAGAACAACATGGTCATGATCCTGCGGCGCGAGCACGGCTGGTCCAAGAGCCGC  750
 221  I  A  S  L  E  K  E  E  A  R  G  E  Q  N  N  M  V  M  I  L  R  R  E  H  G  W  S  K  S  R    250

 751  AGCGTCTCCCACATGCAGAACGAGGTGCGCGCCCGCCTGGAGCAGTACCTCCTGCTGGAGTCCTGCCTGCCCAAGGTCGGCGAGATCTAC  840
 251  S  V  S  H  M  Q  N  E  V  R  A  R  L  E  Q  Y  L  L  L  E  S  C  L  P  K  V  G  E  I  Y    280

 841  CAGCTCGACACGGCCGAACGCGAGGCGCTGGAGCGGTATCGCACGGACGCGGTGCGCACGGTGATCCGCGGTTCCTACGACTGGCACCGG  930
 281  Q  L  D  T  A  E  R  E  A  L  E  R  Y  R  T  D  A  V  R  T  V  I  R  G  S  Y  D  W  H  R    310

 931  TCCTCGGGCCGCTACGACGCCGAGTTCGCGCTCGCCGCCGGCGCCCAGGGCTACCTGGAGGAGCTCGGCAGCAGCGCCCACTAGggcctc 1020
 311  S  S  G  R  Y  D  A  E  F  A  L  A  A  G  A  Q  G  Y  L  E  E  L  G  S  S  A  H  *          340

1021  gtacccggccgtcgcggcggctgggctgtcaacagatctgtagagacgggaagttgagaaactcatgaccgagcagaccttcaccgcggg 1110

1111  agccgcccggggccctcccgtcgtgggacacgcctgcagatgatgcgtcaccagtgaacttcatgactcgctgtcggccacgaactggtc 1200

1201  gagatcagatcgggccacacggctacgtcgaccaccgact  1240
```

FIGURE 2: Pentalenene synthase sequence. DNA flanking the pentalenene synthase gene is in lowercase type. The deduced amino acid sequence is shown below the nucleotide coding sequence. The +1 position in the nucleotide sequence corresponds to the A in the translational start codon. A putative rbs at −8 is underlined, and the 9-nt inverted repeat, corresponding to a possible transcriptional terminator, is double-underlined. The internal *Bgl*II site at nt-833 is shown in italics. Amino acid sequences corresponding to experimentally determined N-terminal or internal tryptic peptides are dot-underlined. The aspartate-rich and base-rich domains are indicated by underlines under the amino acids of the domains.

amino acid N-terminal sequence as well as the three previously observed internal tryptic peptides.

Overexpression in E. coli of Pentalenene Synthase. The identity of the isolated gene was fully confirmed by expression of recombinant pentalenene synthase in *E. coli*. Based on our previous success in expressing both trichodiene synthase (Cane et al., 1993a) and aristolochene synthase (Cane et al., 1993b) in *E. coli* using the T7 expression vector pLM1 (Sodeoka et al., 1993), the same strategy was used to express pentalenene synthase, thereby confirming the identity of the gene and providing sufficient quantities of the protein for further mechanistic and structural studies.

The start primer (**PS-1**) for PCR amplification of the pentalenene synthase coding region was a 56-mer oligonucleotide containing an *Eco*RI site (underlined) followed by the T7 gene 10 rbs and translational spacer element (italics), immediately upstream of the ATG start codon (bold) and the next eight amino acid codons of pentalenene synthase: [5′-d(TACCGAGCTCGAATTC*AGGAGATATACAT***ATG**-CCCCAGGACGTCGACTTCCACATA)-3′]. The reverse or halt primer (**PS-2**) was a 38-mer containing a *Hin*dIII site (underlined) followed by 24 bases complementary to the stop codon (bold) and seven C-terminal amino acid codons of the pentalenene synthase gene: [5′-(dCAGTGAATAAGCTT-**CTA**GTGGGCGCTGCTGCCGAGCTC)-3′]. DNA from plasmid pE34-135, digested with *Nco*I, was used as the template for PCR amplification, resulting in the generation of a single PCR product of the expected size (1057 bp). This PCR product was digested with *Eco*RI and *Hin*dIII and ligated into the corresponding sites of pLM1 (Scheme 2).

The resulting ligation product was used to transform *E. coli* XL1-blue. The plasmid pZW05 containing the desired insert was isolated and used to transform the expression host *E. coli* BL21(DE3), which harbors a prophage carrying the gene for T7 RNA polymerase under control of the lacUV5 promoter (Studier et al., 1990; Studier & Moffatt, 1986). Single colonies of the resulting transformants were individually analyzed, and of six samples examined, all contained plasmids of the expected size. Resequencing of pZW05 confirmed that no mutations had been introduced during the PCR amplification and subcloning of the pentalenene synthase coding sequence.

The six single colonies of *E. coli* BL21(DE3)/pZW05 were used to investigate the production of active pentalenene synthase. Each of them was used to inoculate 5 mL of LB medium containing 100 μL/mL ampicillin. Individual cultures were grown at 37 °C until the OD_{595} reached 1.0–1.6. Following induction with IPTG and incubation for 2 h at 37 °C, aliquots of each culture were treated with lysozyme and found to contain pentalenene synthase activity. SDS–PAGE analysis of total cellular protein revealed the presence of an intense protein band of M_r 40.5 kDa, which was absent from extracts of a control culture of *E. coli* BL21(DE3) transformed with pLM1 (data not shown). The amount of pentalenene synthase produced in these cultures was estimated to be about 90 μg/(mL of culture) by comparing the intensity of the 40.5-kDa band with that of the corresponding lysozyme band at 14.4 kDa. Analysis of both the soluble and insoluble components established that a significant proportion (>80%)

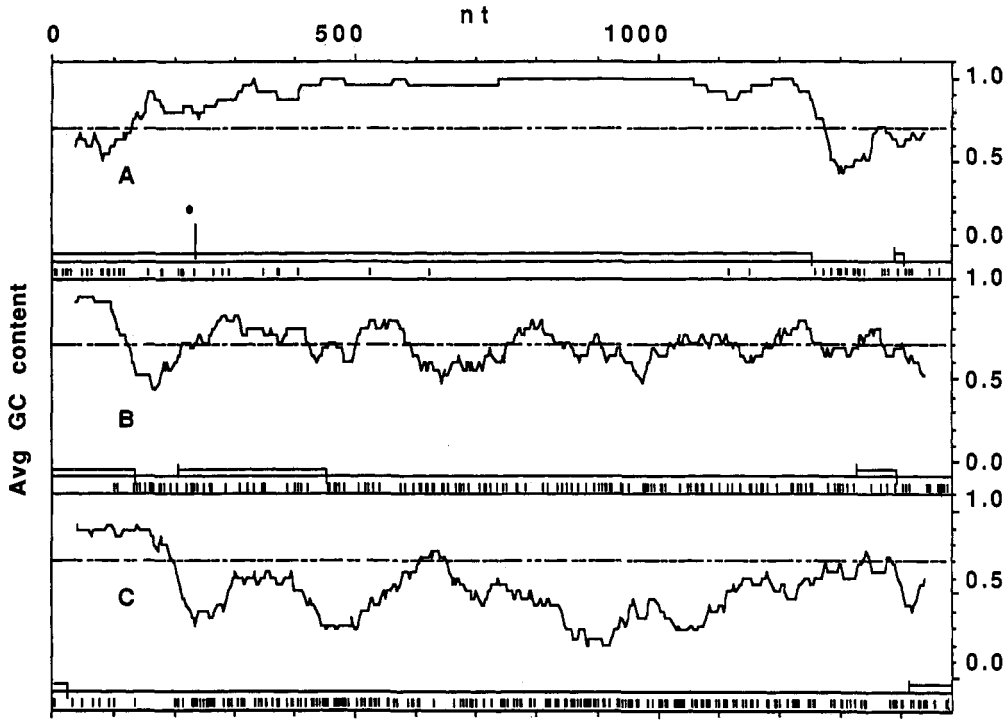
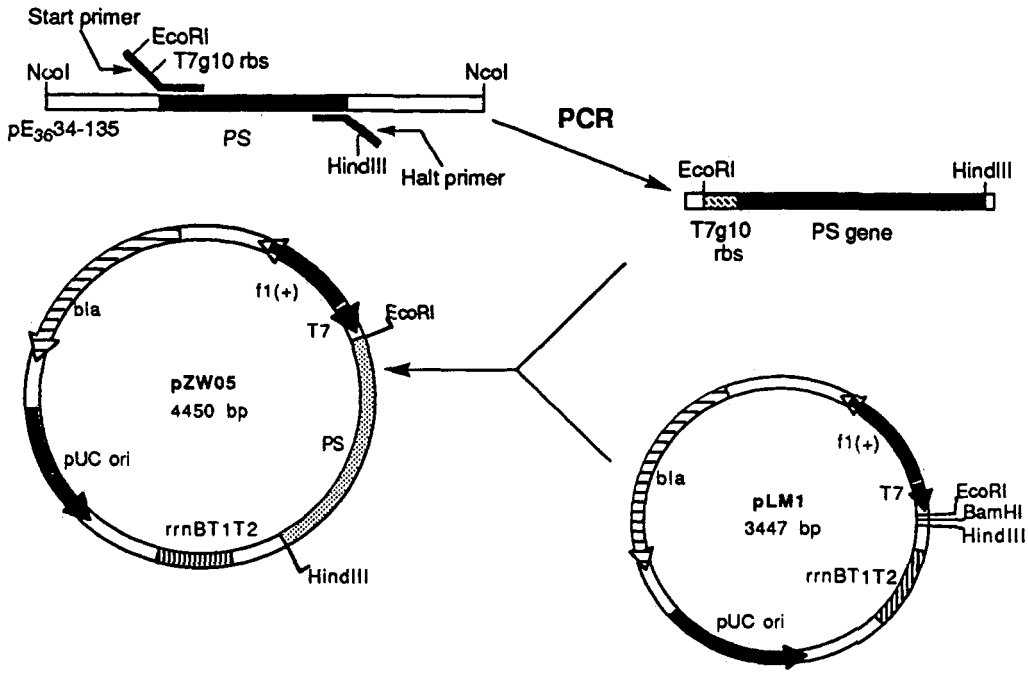


FIGURE 3: GC bias and codon preference analysis of pentalenene synthase sequence. The CodonPreference program of the GCG Sequence Analysis Software Package was used to analyze average GC content and codon usage in each of three reading frames, using a codon bias window of 25 codons, a rare codon display threshold of 0.1, and a comparison table generated from *Streptomyces* coding sequences from genes with average GC content of 70–74%. For each reading frame, the graph of third-position GC bias plots the fraction of the third position in each codon that is either G or C. Position 241 on the horizontal axis corresponds to nt +1 in Figure 2. Panel A gives the average GC content of the third base in each codon of the pentalenene synthase ORF, panel B gives the average GC content of the first base in each codon of the pentalenene synthase ORF, and panel C gives the average GC content of the second base in each codon of the pentalenene synthase ORF. At the bottom of each panel, rare *Streptomyces* codons for each reading frame are indicated by vertical lines and ORFs are indicated by horizontal lines, with stop codons marked by vertical lines which intersect the horizontal axis. The actual ATG start codon is indicated by an asterisk and a vertical line in panel A.

Scheme 2: Construction of the Expression Vector pZW05 from the 3.9-kb *Nco*I Fragment, Containing the Pentalenene Synthase Gene, and pLM1, with the T7 Gene 10 Promoter and Translational Leader (T7)



of the desired 40.5-kDa protein was present in the soluble fraction.

In order to confirm the identity of recombinant pentalenene synthase, the crude cell lysate was incubated with FPP. Analysis of the resulting pentane extract by capillary GC–MS indicated the presence of a single component which was identical in retention time, parent molecular ion (*m/e* 204), and EI mass spectrum to a synthetic sample of (±)-pentalenene.

Purification of Recombinant Pentalenene Synthase. In order to optimize production of active pentalenene synthase, the effects of different concentrations of IPTG, time of

Table 3: Purification of Recombinant Pentalenene Synthase

purification step	total protein (mg)	total activity (nmol/min)	specific activity (nmol/min/mg)	purification (x-fold)	yield (%)[a] (net yield)
crude extract	70	3800	54	1	100
Q-Sepharose	3.5	564	150	2.9	15 (15)
Affigel Blue	0.35	81	239	10.8	77 (11.6)
Mono-Q 5/5	0.14	124	864	23.7	153 (17.7)

[a] Calculated from loadings onto the column at each step, based on activity measurements.

induction, and length of incubation following induction, as well as methods of cell lysis, were investigated. These studies showed that optimum enzyme activity was produced when 0.5 mM IPTG was used to induce cultures which had achieved a cell density corresponding to OD_{595} 0.6–0.7. Specific and total pentalenene synthase activities peaked at 3.0 and 3.5 h postinduction, respectively. For purification of pentalenene synthase, cells were harvested after 3 h by centrifugation and washed with 0.9% (w/v) NaCl to yield typically 0.3–0.6 g of cells per 100 mL of LB media. The washed cells were lysed by treatment with lysozyme for 1 h at room temperature in the presence of protease inhibitors, 10% (v/v) glycerol, and DNAse I, as well as the detergent Triton X-100, which was added 30 min into the incubation. After brief sonication and clarification by centrifugation, the resulting extracts typically contained 60–80 mg of protein per gram of cells and displayed specific activities up to 55 nmol min^{-1} (mg of protein)$^{-1}$.

Recombinant pentalenene synthase was purified by a three-step procedure (Table 3). Crude cell extracts[4] were subjected to Q-Sepharose anion-exchange chromatography followed by passage of the purest cyclase fractions through an Affigel Blue column. Although the latter step did not effect a significant additional purification, as judged by the increase in apparent specific activity, it did serve to remove a few higher molecular weight contaminants. Final purification was achieved by Mono Q 5/5 anion-exchange chromatography. The recovered pentalenene synthase was homogeneous as judged by SDS–PAGE and displayed a specific activity of 864 nmol of pentalenene min^{-1} (mg of protein)$^{-1}$, about 1.5 times that previously observed for the purest preparations of native *Streptomyces* pentalenene synthase and corresponding to a k_{cat} of 0.32 ± 0.02 s^{-1}. The K_m of 0.31 ± 0.05 μM for FPP was in good agreement with that previously determined with partially purified preparations of native enzyme (Cane & Pargellis, 1987). Finally, the identity of the cyclization product generated using homgeneous pentalenene synthase was reconfirmed as pentalenene by capillary GC–MS analysis.

Characterization of Recombinant Pentalenene Synthase. The mobility of recombinant pentalenene synthase upon SDS–PAGE corresponded to a M_r of 40.5 kDa, in close agreement with the value of 41.7 kDa previously assigned to the native *Streptomyces* protein analyzed under the same conditions. Gel filtration of the Mono-Q-purified protein on a precalibrated Superose-12 column gave an apparent M_r of 53.4 ± 2.8 kDa, which, although high with respect to the calculated subunit M_D of 38 000 Da, nonetheless compares favorably with the M_r values of 50–51 kDa which had been estimated for the native *Streptomyces* pentalenene synthase by non-denaturing gel electrophoresis,[3] thereby reconfirming the monomeric nature of the enzyme.[5] The pH optimum for recombinant pentalenene synthase was 8.4–8.6, again consistent with that determined for the native enzyme. The apparent isoelectric point of the Mono-Q-purified recombinant protein, determined by isoelectric focusing, was 5.2, in excellent agreement not only with that established for native enzyme by 2-D electrophoresis (p*I* 5.0–5.1) but with the p*I* value of 5.13 calculated directly from the predicted amino acid sequence.[6] N-terminal sequencing of the first five amino acids of the PVDF-blotted protein revealed a sequence identical both with that predicted from the corresponding DNA sequence, after removal of the N-terminal fMet, and with that determined for native pentalenene synthase.

The metal ion dependence of recombinant pentalenene synthase was typical of that exhibited by sesquiterpene and monoterpene synthases (Cane, 1990; Croteau, 1987). When a series of divalent metal ions were incubated at 5 mM concentration with Mono-Q-purified enzyme, Mg^{2+} gave the highest observed levels of activity, as expected, although Mn^{2+} also was effective in supporting cyclization. As with the native enzyme, Mn^{2+} appeared to be inhibitory at higher concentrations (Cane & Pargellis, 1987); conversion of FPP to pentalenene was negligible in the presence of 15 mM Mn^{2+}. Surprisingly, Fe^{2+} gave low but detectable levels of conversion to hydrocarbon products, although it was not established whether or not a competing slow non-enzymatic solvolysis was involved. The other metal ions which were tested (Ca^{2+}, Co^{2+}, Cu^{2+}, and Zn^{2+}) gave no detectable conversion to product. Finally, inorganic pyrophosphate was found to be a competitive inhibitor of pentalenene synthase, with a K_I of 3.23 ± 0.59 μM [cf. Cane and Pargellis (1987)].

DISCUSSION

Cloning Strategy. A commonly used strategy for the isolation of structural genes is to use amino acid sequence information obtained from the purified protein to design oligonucleotide probes for the screening of the appropriate DNA libraries. Unfortunately, the inherent degeneracy of the genetic code can necessitate the use of mixed primers, of which the primer with the correct sequence often comprises only a minor component, even with probes of only 20 nt, resulting in weak or false hybridization signals, or both. Various strategies have been developed to overcome this problem, including the use of deoxyinosine in the third position of each codon (Saito et al., 1986) or the construction of "guessmers" chosen on the basis of preferred codon usage. The high (73%) GC content of *Streptomyces* as well as the well-documented

[4] Preliminary experiments indicated that although the crude cyclase could be concentrated by ammonium sulfate fractionation, the resuspended and desalted pellet from the 60–80% fraction had suffered a ca. 5-fold decrease in specific activity compared to the original crude extract.

[5] The deviation between the observed and expected behaviors upon gel filtration, as well as the discepancy between the mobility upon SDS–PAGE and the calculated molecular weight of the monomeric subunit, suggests that the native enzyme may have an elongated, nonglobular shape. Such behavior is well-precedented. (For example, isopentenyl diphosphate isomerase: M_D = 33.25 kDa based on the gene sequence; M_r = 39–40 kDa based on SDS–PAGE (Anderson et al. 1989).)

[6] The p*I* was calculated from the deduced amino acid sequence using the Peptidesort program of the GCG Sequence Analysis Software Package.

Table 4: Comparison of Aspartate-Rich Domains in Terpenoid Synthases

synthase	sequence	reference
pentalenene synthase	FLDDFLD	this paper
trichodiene synthase (*F. sporotrichioides*)	VLDDSKD	Hohn and Beremand (1989)
trichodiene synthase (*G. pulicaris*)	VLDDSSD	Hohn and Desjardins (1991)
aristolochene synthase	LIDDVLE	Proctor and Hohn (1993)
epi-aristolochene synthase	IVDDTFD	Facchini and Chappell (1992)
casbene synthase	LIDDTID	C. A. West, pers. commun.
limonene synthase	VIDDIYD	Colby et al. (1993)

bias in codon usage which results in 90–95% GC content in the third codon (Bibb et al., 1984; Wright & Bibb, 1992) allows, in principle, the design of longer probes of lower inherent degeneracy (Hori et al., 1988). Unfortunately, this same higher GC content also carries with it the problem of false hybridization under conditions of lower stringency. We therefore chose to use N-terminal and internal amino acid sequences to design degenerate oligonucleotides, not for use directly as hybridization probes but as primers for amplification of genomic DNA by PCR. After oligolabeling, *the resulting dsDNA amplification products, which can range in size from 50 bp to more than 500 bp, are in fact ideal hybridization probes*, since they are 95–100% identical to the target sequences, even allowing for significant numbers of mismatches in the original oligonucleotide mix.[7,8] As a result, Southern blots as well as colony or plaque hybridizations can be carried out at very high stringency, thereby minimizing false positives and allowing more reliable screening of plasmid or phage DNA libraries. By using appropriate combinations of primers based on more than one internal sequence, PCR amplification can not only generate the longest possible probe but also establish the relative ordering of several different internal peptides.

Pentalenene Synthase Sequence Analysis. The deduced amino acid sequence for pentalenene synthase showed no significant similarity to any known sequences in the DNA and protein databases, including a number of previously reported sesquiterpene synthase genes (Hohn & Bereman, 1989; Hohn & Dejardins, 1991; Proctor & Hohn, 1993; Facchini & Chappell, 1992) as well as several related prenyl transferases (Ashby & Edwards, 1990) and a monoterpene synthase (Colby et al., 1993). On the other hand, the short aspartate-rich sequence LFDDLFD beginning at amino acid 78 conforms to the consensus sequence (I,L,V)XDDXX(D,E), which was first noted by Edwards (Ashby & Edwards, 1990; Ashby et al., 1992) in several prenyl transferases which mediate the allylic diphosphate condensation reactions of isoprenoid chain elongation, reactions with close mechanistic analogy to terpenoid cyclization. It is intriguing, therefore, that this same sequence has subsequently been found in all monoterpene, sequiterpene, and diterpene synthase genes characterized to date (Table 4). Although Edwards has suggested that this acidic motif may be involved in divalent metal binding, this has yet to be explicitly demonstrated. Nonetheless, site-directed mutagenesis has been used to show that replacement of the third asparate (D247) by glutamate in domain II of rat FPP synthase had no significant effect on the observed prenyl transferase activity, while substitution of glutamate for the first aspartate (D243) resulted in a 90-fold reduction in V_{max} and an increase in K_m for isopentenyl diphosphate (Marrero et al., 1992). The corresponding D244E mutant exhibited only a 10-fold decrease in the observed V_{max} (Joly & Edwards, 1993). By contrast, analogous mutations in the aspartates of domain I resulted in near 1000-fold reductions in V_{max} in both the D104E and D107E mutants, but almost no change in V_{max} for the D103E mutant. In none of these mutants was the K_m for the allylic substrate geranyl diphosphate significantly affected. Interestingly, an analogous VADDTAE sequence has also been shown to be present in rat oxidosqualene cyclase, an enzyme catalyzing cationic polyolefin cyclizations with no requirement at all for divalent metal ions (Abe & Prestwich, 1994). Even more intriguing is the observation that inactivation of rat oxidosqualene cyclase by an active site directed affinity label results in equal labeling of *both* aspartates. Experiments are in progress in our own laboratory to establish the functional significance of this acidic motif in pentalenene synthase and other cloned sesquiterpene synthases.

Brems and Rilling have reported the use of photoaffinity labeling to identify a putative active site arginine in avian FPP synthase within the sequence ERYK (Brems et al., 1981). A similar base-rich motif (ERYR) is evident in pentalenene synthase as well as in trichodiene synthases from both *Fusarium sporotrichioides* (RRYR) (Hohn & Beremand, 1989) and *Gibberella pulicaris* (HRYK) (Hohn & Desjardins, 1991). On the other hand, no homologous basic sequences appear to be present in aristolochene synthase (Proctor & Hohn, 1993).

In summary, the structural gene for pentalenene synthase from *Streptomyces* UC5319 has been cloned, sequenced, and expressed at high levels in *E. coli*. The purified recombinant pentalenene synthase, which is identical to the native *Streptomyces* enzyme in all respects, is now conveniently available in substantial quantities for further enzyme mechanistic and protein structural studies.

ACKNOWLEDGMENT

We would like to particularly thank Professor David Hopwood of the John Innes Institute in Norwich, England, for providing the opportunity to D.E.C. to spend part of a Guggenheim Fellowship in his laboratory, and Professor Hopwood's collaborators for numerous helpful suggestions and invaluable advice. We would also like to thank Professor David H. Sherman for helpful suggestions and critical reading of the manuscript.

REFERENCES

Abe, I., & Prestwich, G. D. (1994) *J. Biol. Chem. 269*, 802–804.

Anderson, M. S., Muehlbacher, M., Street, I. P., Proffitt, J., & Poulter, C. D. (1989) *J. Biol. Chem. 264*, 19169–19175.

Ashby, M. N., & Edwards, P. A. (1990) *J. Biol. Chem. 265*, 13157–13164.

Ashby, M. N., Kutsunai, S. Y., Ackerman, S., Tzagoloff, A., & Edwards, P. A. (1992) *J. Biol. Chem. 267*, 4128–4136.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., & Struhl, K. (1987) *Current Protocols in Molecular Biology*, Vol. 1, p 1.8.3, John Wiley, New York.

Bibb, M. J., & Cohen, S. N. (1982) *MGG, Mol. Gen. Genet. 187*, 265–277.

Bibb, M. J., Findlay, P. R., & Johnson, M. W. (1984) *Gene 30*, 157–166.

Bradford, M. (1976) *Anal. Biochem. 72*, 248–254.

[7] Comparison of the sequences of primers **F-1**, **R-1**, **R-2**, and **R-3** with the actual genomic sequences indicates a number of mismatches, noted in Table 2. These mismatches may account for the very narrow window for clean amplification of genomic DNA using the mixed primers.

[8] In recent years, this technique has been more and more frequently used to screen both genomic and cDNA libraries. (Cf. Seidel et al. (1992), Ropp et al. (1993), and Proctor and Hohn (1993).)

Brems, D. N., Bruenger, E., & Rilling, H. C. (1981) *Biochemistry 20*, 3711–3718.
Cane, D. E. (1990) *Chem. Rev. 90*, 1089–1103.
Cane, D. E., & Tillman, A. M. (1983) *J. Am. Chem. Soc. 105*, 122–124.
Cane, D. E., & Pargellis, C. (1987) *Arch. Biochem. Biophys. 254*, 421–429.
Cane, D. E., & Ha, H. (1988) *J. Am. Chem. Soc. 110*, 6865–6870.
Cane, D. E., & Weiner, S. W. (1994) *Can. J. Chem. 72*, 118–127.
Cane, D. E., Abell, C., & Tillman, A. M. (1984) *Bioorganic Chemistry 12*, 312–328.
Cane, D. E., Oliver, J. S., Harrison, P. H. M., Abell, C., Hubbard, B. R., Kane, C. T., & Lattman, R. (1990a) *J. Am. Chem. Soc. 112*, 4513–4524.
Cane, D. E., Pawlak, J. L., Horak, R. M., & Hohn, T. M. (1990b) *Biochemistry 29*, 5476–5490.
Cane, D. E., Abell, C., Harrison, P. H. M., Hubbard, B. R., Kane, C. T., Lattman, R., Oliver, J. S., & Weiner, S. W. (1991) *Philos. Trans. R. Soc., B. 332*, 123–129.
Cane, D. E., Wu, Z., Oliver, J. S., & Hohn, T. M. (1993a) *Arch. Biochem. Biophys. 300*, 416–422.
Cane, D. E., Wu, Z., Proctor, R. H., & Hohn, T. M. (1993b) *Arch. Biochem. Biophys. 304*, 415–419.
Chrambach, A., & Rodbard, D. (1971) *Science 172*, 440–451.
Colby, S. M., Alonso, W. R., Katahira, E., McGarvey, D. J., & Croteau, R. (1993) *J. Biol. Chem. 268*, 23016–23024.
Croteau, R. (1987) *Chem. Rev. 87*, 929–954.
Davis, B. J. (1964) *Ann. N.Y. Acad. Sci. 121*, 404–427.
Deng, Z., Kieser, T., & Hopwood, D. A. (1987) *Nucleic Acids Res. 15*, 2665–2675.
Devereux, J., Haeberli, P., & Smithies, O. (1984) *Nucleic Acids Res. 12*, 387–395.
Facchini, P. J., & Chappell, J. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 11088–11092.
Fawcett, T. W., & Bartlett, S. G. (1990) *Bio/Technology 9*, 46–49.
Fawcett, T. W., & Bartlett, S. G. (1991) *Comments*, Vol. 17, No. 4, p 19, United States Biochemical Corp., Cleveland, OH.
Ferguson, K. A. (1964) *Metabolism 13*, 985–1002.
Hedrick, J. L., & Smith, A. J. (1968) *Arch. Biochem. Biophys. 126*, 155–164.
Hohn, T. M., & Beremand, P. D. (1989) *Gene 79*, 131–138.
Hohn, T. M., & Desjardins, A. E. (1992) *Mol. Plant-Microbe Interact. 5*, 249–256.
Hopwood, D. A., Bibb, M. J., Chater, K. F., Kieser, T., Bruton, C. J., Kieser, H. M., Lydiate, D. J., Smith, C. P., Ward, J. M., & Schrempf, H. (1985) *Genetic Manipulation of Streptomyces, A Laboratory Manual*, John Innes Foundation, Norwich, U.K.
Hori, M., Sakata, N., Niino, Y., Makabe, O., Hamada, M., Mizuno, S., & Hotta, K. (1988) *J. Antibiot. 41*, 1462–1470.
Joly, A., & Edwards, P. A. (1993) *J. Biol. Chem. 268*, 26983–26989.
Kajander, E. O., & Raina, A. M. (1981) *Biochem. J. 193*, 503–512.
Kieser, T. (1984) *Plasmid 12*, 19–36.
Marrero, P. F., Poulter, C. D., & Edwards, P. A. (1992) *J. Biol. Chem. 267*, 21873–21878.
Matsudaira, P. (1987) *J. Biol. Chem. 262*, 10035–10038.
Misumi, S., Ohtsuka, T., Ohfune, Y., Sugita, K., Shirahama, H., & Matsumoto, T. (1979) *Tetrahedron Lett.* 31–34.
Ohfune, Y., Shirahama, H., & Matsumoto, T. (1976) *Tetrahedron Lett.* 2869–2872.
Proctor, R. H., & Hohn, T. M. (1993) *J. Biol. Chem. 268*, 4543–4548.
Pulido, D., & Jimenez, A. (1987) *Nucleic Acids Res. 15*, 4227–4240.
Ropp, J. D., Gunsalus, I. C., & Sligar, S. G. (1993) *J. Bacteriol. 175*, 6028–6037.
Rosenberg, M., & Court, D. (1979) *Annu. Rev. Genet. 13*, 319–353.
Saito, S., Takahashi, H., Saito, H., Arai, M., & Murao, S. (1986) *Biochem. Biophys. Res. Commun. 141*, 1099–1103.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning, A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sanger, F., Nicklen, S., & Coulsen, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Seidel, H. M., Pompliano, D. L., & Knowles, J. R. (1992) *Biochemistry 31*, 2598–2608.
Sodeoka, M., Larson, C. J., Chen, L., LeClaire, K. P., & Verdine, G. L. (1993) *Bioorg. Med. Chem. Lett. 3*, 1095–1100.
Spector, T. (1978) *Anal. Biochem. 86*, 142–146.
Studier, F. W., & Moffatt, B. A. (1986) *J. Mol. Biol. 189*, 113–130.
Studier, F. W., Rosenberg, A. H., Dunn, J. J., & Dubendorff, J. W. (1990) in *Methods in Enzymology: Gene Expression Technology* (Goeddel, D. V., Ed.), Vol. 185, pp 60–89, Academic Press, New York.
Trayer, I. P., Trayer, H. R., Small, D. A. P., & Bottomley, R. C. (1974) *Biochem. J. 139*, 609–623.
Wright, F., & Bibb, M. J. (1992) *Gene 113*, 55–65.