Geranylgeranylglyceryl Phosphate Synthase. Characterization of the Recombinant Enzyme from *Methanobacterium thermoautotrophicum*†

Tim Soderberg, Anjun Chen, and C. Dale Poulter*

Department of Chemistry, University of Utah, Salt Lake City, Utah 84112

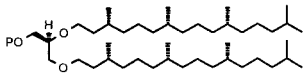
Received June 7, 2001; Revised Manuscript Received September 20, 2001

ABSTRACT: Geranylgeranylglyceryl diphosphate synthase (GGGP synthase) catalyzes alkylation of (*S*)-glyceryl phosphate [(*S*)-GP] by geranylgeranyl diphosphate (GGPP) to produce (*S*)-geranylgeranylglyceryl phosphate [(*S*)-GGGP]. This reaction is the first committed step in the biosynthesis of ether-linked membrane lipids in Archaea. The gene encoding GGGP synthase from *Methanobacterium thermoautotrophicum* was cloned using probes designed from the N-terminal sequence determined from the purified enzyme. The open reading frame, which encoded a protein of 245 amino acids, was inserted into a pET expression vector and expressed in *Escherichia coli*. The recombinant GGGP synthase was purified to homogeneity. The enzyme is active as a homopentamer, as determined by size exclusion chromatography and equilibrium sedimentation experiments. GGGP synthase has optimal activity at 55 °C in pH 8.0 buffer containing 1 mM $MgCl_2$. $V_{max} = 4.0 \pm 0.1$ μmol min^{-1} mg^{-1} ($k_{cat} = 0.34 \pm 0.03$ s^{-1} for pentameric GGGP synthase assuming all subunits are fully active), $K_m^{(S)-GP} = 13.5 \pm 1.0$ μM, and $K_m^{GGPP} = 506 \pm 47$ nM. These steady-state catalytic constants were identical to those for enzyme isolated from cell extracts of *M. thermoautotrophicum* [Chen, A., Zhang, D., and Poulter, C. D. (1993) *J. Biol. Chem. 268*, 21701−21705]. Alignment of seven putative archaeal GGGP synthase sequences revealed a number of highly conserved residues consisting of five aspartate/glutamates, three serine/threonines, two prolines, and five glycines, including a conserved GGG motif.

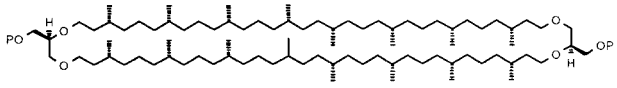
The Archaea are a unique class of organisms that form a third distinct kingdom of life. Early interest in the Archaea stemmed from their ability to thrive in extreme environments, such as high temperatures, high salinity, or low pH, which are inhospitable to most other forms of life. Because their outward physical appearanace resembles the Bacteria, it was originally assumed that these "extremophiles" were unusual members of the bacterial kingdom. This assumption was later challenged after a detailed analysis of ribosomal RNA genetic patterns (*1*).

One of the most striking physiological differences between the Archaea and other organisms is the structure of their membrane lipids (*2*). In Archaea, membrane lipids are based on a core architecture in which branched-chain saturated hydrocarbons—most commonly two C_{20} phytanyl groups—are connected to glycerol through ether linkages. The hyperthermophiles often contain dibiphytanyl tetraethers, where two biphytanylglyceryl units are covalently linked though their hydrocarbon tails (Scheme 1). The branched-chain hydrocarbons and ether linkages in archaeal lipids stand in marked contrast to the straight-chain, ester-linked glyceryl lipids found in eubacterial and eukaryotic membranes. A further distinguishing characteristic of archaeal lipids is the (*S*) absolute stereochemistry of the glyceryl moiety, as opposed to the (*R*) stereochemistry of the diglyceride unit in bacterial and eukaryotic phospholipids.

Scheme 1: Diether and Tetraether Core Structures of Archaeal Lipids[a]



Phytanyl Diether



Biphytanyl Tetraether

[a] Polar head groups are represented by an X.

The biosynthesis of biphytanylglyceryl lipids in Archaea is largely unexplored. The glyceryl unit is derived from (*S*)-glyceryl phosphate [(*S*)-GP].[1] Its enantiomer, (*R*)-GP, has an analogous structural role in eubacterial and eukaryotic lipids (*3*). (*R*)-GP is an important metabolite in species from all three kingdoms and is synthesized by phosphorylation of glycerol and by reduction of dihydroxyacetone phosphate (DHAP). In addition, archaea have an (*S*)-GP dehydrogenase

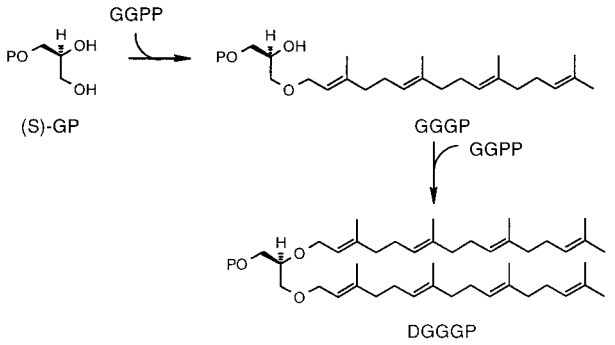
† This work was supported by NIH Grant GM21328. T.S. is a NIH predoctoral trainee (Grant GM08537).

* To whom correspondence should be addressed. Telephone: (801) 581-6685. Fax: (801) 581-4391. E-mail: poulter@chemistry.utah.edu.

[1] Abbreviations: BME, β-mercaptoethanol; DGGGP, digeranylgeranylglyceryl phosphate; DHAP, dihydroxyacetone phosphate; DMAPP, dimethylallyl diphosphate; DMAPP−AMP transferase, dimethylallyl diphosphate:AMP dimethylallyltransferase; DMAPP−tRNA transferase, dimethylallyl diphosphate:tRNA dimethylallyltransferase; FGPP, farnesylgeranyl diphosphate; FPP, farnesyl diphosphate; GGGP, geranylgeranylglyceryl phosphate; GGPP, geranylgeranyl diphosphate; (*S*)-GP, (*S*)-glyceryl phosphate; IPP, isopentenyl diphosphate; IPTG, isopropyl β-D-thiogalactoside; orf, open reading frame; PMSF, phenylmethanesulfonyl chloride; X-gal, 5-bromo-4-chloro-3-indolyl β-D-galactoside.

10.1021/bi0111799 CCC: $20.00
Published on Web 11/09/2001

Scheme 2: Biosynthesis of DGGGP from (*S*)-GP and GGPP



that catalyzes the reduction of DHAP. The dehydrogenase from *Methanobacterium thermoautotrophicum* has been purified and characterized (*4*). Thus far, activity for a (*S*)-GP-specific glycerol kinase has not been detected in any species (*5*).

The C_{20} phytanyl units found in archaeal membrane lipids are derived from geranylgeranyl diphosphate (GGPP) (*6*). GGPP is constructed from isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) in *M. thermoautotrophicum* by GGPP synthase, a bifunctional short-chain prenyltransferase that makes both farnesyl diphosphate (FPP) and GGPP. The *M. thermoautotrophicum* enzyme was purified and characterized (*7*), and the gene was subsequently cloned and expressed in *Escherichia coli* (*8*). The gene encoding an enzyme responsible for the synthesis of farnesylgeranyl diphosphate (FGPP) was recently cloned from the hyperthermophilic archaeon *Aeropyrum pernix* (*9*). Morii and co-workers (*10*) reported that the membrane lipids of *A. pernix* are composed exclusively of C_{25} diether lipids, presumably derived from FGPP, rather than the more common C_{20} diethers derived from GGPP.

The first committed step in the construction of archaeal membrane lipids is formation of an ether linkage between (*S*)-GP and the isoprenoid diphosphate, usually GGPP. In *M. thermoautotrophicum* two separate enzymes catalyze transfer of geranylgeranyl units to the C3 and C2 oxygen atoms of (*S*)-GP, respectively (Scheme 2) (*3*). The enzyme catalyzing the first reaction, geranylgeranylglyceryl phosphate synthase (GGGP synthase), is cytosolic and has been purified from *M. thermoautotrophicum*, Marburg strain (*11*). The enzyme catalyzing the second prenyltransfer reaction, digeranylgeranylglyceryl phosphate synthase (DGGGP synthase), is membrane bound and has not yet been purified (*11*).

GGGP synthase is an interesting enzyme from at least three standpoints. It lies at the branch point to biosynthesis of the membrane lipids in Archaea. The preliminary report of Chen and co-workers (*11*) suggests that the active form of the enzyme is a multimer of at least four identical subunits, whereas all other prenyltransferases characterized to date are monomers or dimers. Finally, GGGP synthase is a member of a diverse, but not extensively studied, family of prenyltransferases that catalyze the transfer of isoprenoid groups to nonisoprenoid acceptors. Other members of this class include protein prenyltransferases (*12*, *13*), DMAPP−tRNA transferase (*14*), DMAPP−AMP transferase (*15*), and dimethylallyltryptophan synthase (*16*, *17*). The prenyl-transfer reactions catalyzed by protein farnesyltransferase (*13*), DMAPP−tRNA transferase (*14*, *18*, *19*), and dimethylallyltryptophan synthase (*17*) appear to be electrophilic alkylations. We now report cloning of the gene encoding GGGP synthase from *M. thermoautotrophicum* and characterization of the recombinant enzyme.

EXPERIMENTAL PROCEDURES

Materials and General Methods. Frozen cells of *M. thermoautotrophicum*, Marburg strain, were provided by Dr. Lacy Daniels (University of Iowa). Problott PVDF membranes were from Applied Biosystems. [α-^{32}P]dATP (3000 Ci/mol) was from Amersham. Enzymes for DNA cloning were purchased from New England Biolabs and Gibco. All media and culture plates contained 50 μg/mL ampicillin, and plates for blue-white selection also contained 0.2 mM isopropyl β-D-thiogalactoside (IPTG) and 20 μg/mL 5-bromo-4-chloro-3-indolyl β-D-galactoside (X-gal). LB broth contained 5 g of yeast extract, 10 g of tryptone, and 5 g of NaCl in 1 L of water. pBluescript SK(+) was from Stratagene, pTTQ18N was from Amersham, pARC306N was provided by M. Bittner (Biotechnology Division, Amoco Research, Naperville, IL), and pET-11b was from Novagen. (*S*)-GP and (*S*)-[1-^{3}H]GP (4.0 Ci/mol) were available from a previous study (*3*). Butanol-extractable contaminants were removed from stored (*S*)-[1-^{3}H]GP stock solutions by extracting three times with water-saturated 1-butanol (1 mL of 1-butanol/water per 100 μL of GP stock) and once with 1 mL of hexanes to remove residual 1-butanol. GGPP was synthesized as described by Davisson and co-workers (*20*). Stock solutions of GGPP were prepared in 25 mM NH_4HCO_3, and concentrations were determined by phosphate analysis (*21*). Plasmid DNA was purified with the Geneclean system (Bio 101), Wizard Plus Minipreps DNA purification system (Promega), or the Qiagen Plasmid Midi Kit. DNA electrophoresis was conducted with 0.8% agarose gels and visualized by staining with ethidium bromide. Proteins were chromatographed on a Pharmacia FPLC system at 4 °C. Protein concentrations were determined by the method of Bradford (*22*). Protein samples were analyzed by 12% stacking SDS−PAGE on a Bio-Rad Mini-Protein II system and visualized by staining with Coomassie Brilliant Blue. Proteins were dialyzed for 4−12 h at 4 °C using 12000−14000 molecular weight cutoff tubing (Spectrum). *E. coli* strain DH5α was used for all DNA manipulations. DNA oligonucleotide primers were synthesized by the DNA/Peptide Facility of the Huntsman Cancer Institute at the University of Utah. DNA and peptide sequencing was done by the DNA Sequencing Core Facility of the Huntsman Cancer Institute. Negative ion electrospray MS was conducted at the University of Utah Department of Chemistry. Equilibrium sedimentation experiments were carried out at the Hunstman Cancer Institute, University of Utah.

Sequences of Peptides from GGGP Synthase. GGGP synthase was purified to homogeneity as described previously (*11*). To obtain the N-terminal amino acid sequence, 5−10 μg of purified protein was electrophoresed on a 0.75 mm thick SDS−polyacrylamide gel and then electroblotted for 45 min in 10 mM 3-(cyclohexylamino)propanesulfonate, pH 11.0, and 10% methanol on a Problott PVDF membrane using a Mini-Trans Blot cell (Bio-Rad). The membrane was stained in 40% methanol, 10% acetic acid, and 0.1% Coomassie Blue R, destained in 40% methanol and 10%

acetic acid, rinsed with water, and air-dried. The band at ~29 kDa corresponding to GGGPS was excised from the gel and sequenced.

For internal peptide sequences, 6 μg of pure protein was digested with 100 mg of CNBr in 70% formic acid for 24 h in the dark at room temperature (*23*). Peptides were separated by HPLC on a 2 mm × 25 cm Vydac C18 reversed-phase column, using a linear gradient from 0.1% aqueous trifluoroacetic acid to 4:1 acetonitrile:water with 0.1% trifluoroacetic acid. A selection of well-resolved peptides was sequenced.

PCR Cloning of GGGP Synthase Gene Fragments.[2] Four fully degenerate PCR primers were constructed on the basis of the peptide sequences of GGGP synthase obtained as described above. The primers contained 5′-terminal *Eco*RI or *Hin*dIII restriction sites (underlined). Primers sequences were as follows: sense primer NSP, 5′ TC GAA TTC ATG AAR GTA GAR GAY TAY TTY CAY 3′; antisense primer NAP, 5′ AT AAG CCT GC DAT YTC IAC NGC YTC YTC 3′; antisense primer 1AP, 5′ AT AAG CTT CC IAA RAA YTC NGC NGC CAT 3′; antisense primer 2AP, 5′ AT AAG CTT TC YTC IGG NAC RTG YTG 3′ (I is deoxyinosine, R is A or G, Y is C or T, D is G, A, or T, and N is G, A, T, or C).

PCR reaction mixtures (100 μL) contained 1 mM deoxynucleotide, 5 μL of DMSO, 500 ng of *M. thermoautotrophicum* genomic DNA, 5 μg of primer pair NSP/NAP, NSP/1AP, or NSP/2AP, and 8 units of Taq DNA polymerase. The initial denaturation step was 4 min at 94 °C, followed by 30 cycles of PCR at 94 °C, 2 min at 37 °C, and 3 min at 72 °C. A 500 bp PCR product from primer pair NSP/2AP was purified on a 0.8% agarose gel and extracted with a QIAEX gel extraction kit (Qiagen). Initial attempts to ligate the PCR product into the *Eco*RI and *Hin*dIII sites of pBluescript SK(+) failed. Instead, gel-purified DNA (1.6 μg) was phosphorylated for 1 h at 37 °C with 20 units of T4 polynucleotide kinase. DNA polymerase I Klenow fragment (9 units) was added, the concentrations of each deoxynucleotide were adjusted to 0.3 mM, and the solution was incubated at 37 °C for 30 min. The 500 bp blunt-ended fragment was purified by Geneclean and ligated into the *Sma*I site of pBluescript SK(+) at room temperature overnight using 400 units of T4 DNA ligase. *E. coli* DH5α competent cells were transformed with the ligation mixture and plated onto LB/Amp/IPTG/X-gal plates. White colonies were screened by restriction analysis using *Bam*HI and *Hin*dIII, and positive clones were sequenced. Plasmid pACII-174 contained a 500 bp insert (cN-2) encoding a polypeptide that matched the determined peptide sequences.

Colony Lift Hybridization. The partial gene sequence cN-2 in pACII-174 was amplified by PCR using the same conditions as above with primers TSP (5′ ATG AAG GTG GAA GAT TA 3′) and TAP (5′ CGC ACC CAG GGC CTG GGC 3′). After purification by Geneclean, the DNA probe was labeled with ^{32}P by primer extension using the Prime-It Gold kit (Stratagene).

Colonies were transferred to a Nylon-1 membrane (Gibco). The membrane was placed on an agar plate until it was completely wet (1−2 min). The membrane was removed and dried on 3MM filter papers. Cells on the membrane were lysed by floating the membranes with colonies facing up on the surface of a solution of 1.5 M NaCl and 0.5 M NaOH for 5−10 min. The membrane was dried on 3MM papers and neutralized by soaking for 5−10 min in 3 M NaOAc (pH 5.5). The membrane was dried again on 3MM papers, equilibrated in 2× SSC solution for 5 min, and air-dried. It was then baked in a 80 °C vacuum oven for 90 min to immobilize the DNA. The membrane was then wetted in a 5× SSC solution and incubated with rocking at 42 °C for 4 h in 20−40 mL of prehybridization solution (5× SSC, 5× Denhardt's reagent, 0.5% SDS, 50% formamide) containing 2 mg of salmon sperm ssDNA. The prehibridization solution was removed, and 20−40 mL of hybridization solution (5× SSC, 0.5% SDS, 50% formamide) was added. The labeled DNA probe was boiled for 5 min in 200 mL of 10 mg/mL salmon sperm ssDNA and added to the hybridization solution immediately. The sample was rocked at 42 °C for 12−24 h. The radioactive solution was removed, and the membranes were washed twice for 5 min at room temperature with 100 mL of 0.1% SDS and 1× SSC solution, twice for 15 min at 42 °C with 100 mL of 0.1% SDS and 0.1× SSC, and finally once for 30 min at 65 °C with 100 mL of 0.1% SDS and 0.1% SSC. The membranes were wrapped in cellophane, and positive colonies were visualized by exposing Crones-4 X-ray film (DuPont) to the membranes for 24 h at −70 °C.

Cloning of the Full-Length GGGP Synthase Gene. Genomic DNA (4 μg) isolated from *M. thermoautotrophicum* cells was completely digested by incubating with 0.2 unit of *Eco*RI at 37 °C for 30 min. The restriction fragments were purified by Geneclean and ligated into the dephosphorylated *Eco*RI site of pBluescript SK(+) by incubation with 400 units of T4 DNA ligase at 16 °C for 8 h. *E. coli* XL1-blue supercompetent cells (Stratagene) were transformed with 5 μL of ligation mixture and incubated in 1 mL of LB at 37 °C for 1 h before being plated on five LB/Amp/IPTG/X-gal plates. The plates were incubated overnight at 37 °C, and those that contained a confluent lawn of colonies were screened by colony life hybridization using the ^{32}P-labeled probe cN-2. Colonies from areas giving positive signals were removed from the plates, resuspended in 1 mL of LB, and replated after a series of dilutions. Plates containing 200−500 well-separated colonies were chosen for screening by a second round of colony lift hybridization. Plasmids isolated from positive isolated colonies from the second round were digested with *Eco*RI and analyzed by Southern blot hybridization using the cN-2 probe. A positive plasmid, pACIII-37, was sequenced and found to contain a complete open reading frame (orf) with N-terminal and internal regions corresponding to the amino acid sequences determined for GGGP synthase.

Recombinant GGGP Synthase. Three regions in pACIII-37 were altered by site-directed mutagenesis using the Muta-Gene Phagemid in vitro mutagenesis kit (Bio-Rad). An *Nde*I restriction site was introduced that overlapped the start codon for the GGGP synthase orf, and a *Sal*I site was introduced downstream from the stop codon. In addition, an internal *Nde*I site in the GGGP synthase orf was inactivated by a silent mutation. The reconstructed gene was excised by restriction digest with *Nde*I and *Sal*I and ligated into the same sites of plasmids pTTQ18N and pARC306N by incubation with 8 units of DNA ligase at 16 °C overnight. The two

[2] Much of the cloning work described in this report was completed before the first publication of a complete archaeal genomic sequence.

resulting plasmids, pACIII-73 and pACIII-74, respectively, were sequenced to verify the presence of the full-length gene. Because subsequent expression trials with pACIII-73 and pACIII-74 failed, the GGGP synthase gene was then subcloned into the pET-11b expression vector (Novagen). The *Sal*I site of plasmid pACIII-73 was changed to a *Bam*HI site by site-directed mutagenesis using the Gene Editor system (Promega), and the GGGP synthase gene was excised by digestion with *Nde*I and *Bam*HI and ligated into identical sites in pET-11b. The resulting plasmid, pTJSIII-124, was sequenced to verify the presence of the full-length gene.

Competent cells of *E. coli* strain Epicurian coli BL21-CodonPlus (DE3)-RIL (Strategene) were transformed with plasmid pTJSIII-124 and grown in LB medium containing 100 μg/mL ampicillin and 34 μg/mL chloramphenicol at 30 °C with vigorous aeration. When the cell density reached $OD_{600} = 1.0$, IPTG was added to a final concentration of 0.5 mM to induce expression, and the culture was allowed to grow for an additional 18 h before cells were harvested by centrifugation. Cells were resuspended in buffer A (50 mM potassium phosphate, pH 7.5, 5 mM $MgCl_2$, 5 mM BME, 200 mM NaCl) plus 1 mM PMSF and disrupted by sonication on ice in two sequences of 30 s each. Cell debris was cleared by centrifugation at 12000*g* for 10 min, and the supernatant was incubated at 60 °C for 10 min. Precipitated proteins were pelleted by centrifugation, and the heat−centrifugation cycle was repeated. The supernatant was then loaded onto a 1.5 × 25 cm DE52 column (Whatman) that had been preequilibrated with buffer A, and the column was washed with buffer A at a flow rate of 1.0 mL/min until UV absorbance at 280 nm reached baseline level. The column was then eluted with a 200 mL linear gradient of 200−600 mM NaCl at a flow rate of 1.0 mL/min. Fractions containing GGGP synthase, as determined by SDS−PAGE, were pooled, dialyzed overnight against buffer A, and concentrated to approximately 5 mL with Centriprep 30 concentrators (Amicon). The sample was loaded onto a Pharmacia 16/60 Superdex 200 size exclusion column that had been preequilibrated with buffer A. The column was eluted with buffer A at a flow rate of 0.5 mL/min, and fractions were pooled from individual UV-active peaks. The peak eluting at 64 mL contained active enzyme. The pooled fractions from this peak were dialyzed against two changes of 1 L of buffer B (50 mM Tris, pH 7.5, 10 mM $MgCl_2$, 5 mM BME) and concentrated with Centriprep 30 and Centricon 30 concentrators (Amicon) to approximately 2 mg/mL. Protein was stored at 4 °C until needed. Storage at −20 °C in buffer B containing 10% glycerol resulted in diminished activity.

Assay for GGGP Synthase. The enzyme was assayed at 55 °C in a buffer containing 50 mM Tris, pH 8.0 (at 55 °C), 10 mM $MgCl_2$, 5 mM BME, and 0.2% Tween 80 (Boehringer Mannheim). Stock solutions of enzyme were prepared in assay buffer, and stock solutions of (*S*)-[^{3}H]GP (4.0 Ci/mol) and GGPP were prepared in 25 mM NH_4HCO_3/0.2% Tween 80. In a 1.7 mL microcentrifuge tube, assay buffer (180 μL) and enzyme solution (10 μL) were incubated at 55 °C for 5 min, and the reaction was started by addition of 10 μL of a solution containing (*S*)-[^{3}H]GP and GGPP. Assay solutions were incubated at 55 °C for 1 min and quenched by addition of 100 μL of 0.5 M EDTA. Saturated NaCl (0.5 mL) was added, and the contents were then transferred to a glass culture tube containing 2 mL of water-saturated 1-butanol. The contents were mixed thoroughly by vortexing for 10 s and allowed to stand for 10 min to allow the 1-butanol and aqueous layers to separate. The radioactivity in 1 mL of 1-butanol was determined by liquid scintillation spectrometry. Control assays were run in an identical manner except that assay buffer was substituted for enzyme solution. The efficiency of the 1-butanol extraction procedure was assessed by performing a second extraction of the aqueous layer with 2 mL of water-saturated 1-butanol. Radioactivity in the 1-butanol layer of the second extraction was at background levels.

K_m for (*S*)-GP was determined from initial velocities measured at a saturating concentration of GGPP (20 μM), 6.8 or 13.6 nM enzyme, and varied concentrations of (*S*)-[1-^{3}H]GP (8−250 μM, 4 Ci/mol). Controls were run by substituting buffer for enzyme at several different concentrations of (*S*)-[1-^{3}H]GP. K_m for GGPP was determined at a saturating concentration of (*S*)-[1-^{3}H]GP (100 μM), 1.7 or 6.3 nM enzyme, and varying concentrations of GGPP (0.5−10 μM). Kinetic constants were calculated using Grafit software (*24*).

Native Molecular Mass Determination. A 0.5 mL sample of purified protein (~2 mg/mL) was loaded onto a Pharmacia 16/60 Superdex 200 size exclusion column. The column was eluted at 0.5 mL/min with 50 mM potassium phosphate, pH 7.5, 5 mM $MgCl_2$, 5 mM BME, and 150 mM NaCl. Protein standards were chymotrypsinogen (25 kDa), ovalbumin (43 kDa), BSA (67 kDa), aldolase (158 kDa), catalase (232 kDa), and thyroglobulin (669 kDa). The apparent native molecular mass of GGGP synthase was calculated by plotting the log of molecular mass of each of the standards versus K_{av}, where $K_{av} = (V_e - V_0)/(V_t - V_0)$. V_e = elution volume of the protein, V_0 = void volume (39 mL), and V_t = bed volume (120 mL).

Protein samples for equilibrium sedimentation experiments were diluted into a buffer containing 50 mM Tris, pH 7.5, 10 mM $MgCl_2$, and 150 mM NaCl. Ten scans of absorbance at 280 nm vs radial distribution were collected with a step size of 0.001 cm after sedimentation equilibrium had been reached at 20 °C. Samples were run at nine different monomer concentrations ranging from 25 to 0.16 μM at 10000*g* and 14000*g*.

Analysis of Products by Electrospray Mass Spectrometry. A 1 mL preparative scale reaction was carried out at 55 °C in 10 mM Tris, pH 8.0, and 1 mM $MgCl_2$, containing 225 μM (*S*)-GP (unlabeled) and 80 μg of GGGPS. GGPP was added slowly over 1 h to a final concentration of 565 μM (2.5 equiv), and the reaction was incubated at 55 °C for an additional 1 h. The white precipitate that formed was removed by centrifugation, air-dried, and redissolved in a 1:1 acetonitrile/water mixture for MS analysis.

Sequence Alignments of Putative Archaeal GGGP Synthases. Putative amino acid sequences for GGGP synthase from seven archaeal species were identified from the Institute for Genomic Research (TIGR) database and aligned with ClustalW using the interface provided by Baylor College of Medicine (http://dot.imgen.bcm.tmc.edu:9331/multi-align/multi-align.html). The organisms and primary loci numbers are as follows: *M. thermoautotrophicum*, MTH552; *Methanococcus jannaschii*, MJ1250; *Pyrococcus horikoshii*, PH1124;

Pyrococcus abyssi, PAB0735; *A. pernix*, APE0621; *Thermoplasma acidophilum*, Ta0995; *Archaeoglobus fulgidus*, AF0403.

RESULTS

Peptide Sequences for GGGP Synthase. An N-terminal peptide sequence of 35 amino acids and four internal sequences of 7−16 amino acids were determined from wild-type GGGP synthase purified from *M. thermoautotrophicum*, Marburg strain (*11*). The N-terminal sequence was MKVED YFHDI LGERK IHLTL IDPEE QTPEE AVEIA, and the internal sequences were (1) (M)AAEFLG, (2) (M)RLFY LEAGS GAPQH VPEE, (3) (M)GYLV VEPGG TVGXV GDTKP VPXNN P, and (4) (M)SLLN SNNPY XIIGA QALXA P, where X represents an unassigned amino acid. The methionines in parentheses were assumed as a consequence of the CNBr cleavage. Underlined regions were used to design degenerate PCR primers.

PCR Cloning of a Fragment of the GGGP Synthase Gene.[2] Of the various combinations tried, the primer pair NSP/2AP (see Experimental Procedures) produced a single 500 bp PCR product from genomic *M. thermoautotrophicum* DNA. After attempts to ligate the fragment into the *Eco*RI and *Hin*dIII sites of pBluescript SK(+) failed, the fragment was instead cloned into the *Sma*I site of the same plasmid by blunt-end ligation. The insert in this clone was sequenced and was found to encode the amino acid sequences that matched those determined for purified GGGP synthase. The insert was amplified by PCR to give the DNA, designated cN-2, used to probe a genomic library of *M. thermoautotrophicum* DNA.

Cloning of the Full-Length GGGP Synthase Gene. A *M. thermoautotrophicum* genomic library was constructed in pBluescript SK(+) from *Eco*RI fragments of fully digested genomic DNA. More than 1000 colonies from the library were screened by two rounds of colony lift hybridization using the ^{32}P-labeled cN-2 probe. Fifteen of the colonies contained a 3.6 kb insert that hybridized to the probe on a Southern blot. One of these, designated pACIII-37, was sequenced and found to contain an orf encoding the N-terminal and internal amino acid sequences for GGGP synthase. This orf was excised from pACIII-37 and subcloned into expression vectors pTTQ18N, pARC306N, and pET-11b to create pACIII-73, pACIII-74, and pTJSIII-124, respectively. Initial attempts to express the gene in pACIII-73 and pACIII-74 gave crude extracts with GGGP synthase activities of only ∼1 nmol min^{-1} mg^{-1}, and no visible band that clearly corresponded to recombinant GGGP synthase on SDS−PAGE. However, induced cultures of *E. coli* strain Epicurian coli BL21-CodonPlus (DE3)-RIL (Stratagene) transformed with pTJSIII-124 gave specific activities for GGGP synthase as high as 1.5 μmol min^{-1} mg^{-1} and a correspondingly large band for GGGP synthase (Figure 1). The genome of the CodonPlus strain contains extra copies of *argU*, *ileY*, and *leuW*, which encode the tRNAs that recognize AGA and AGG (arginine), AUA (isoleucine), and CUA (leucine). These codons are used frequently by Archaea and Eukaryotes and only rarely in *E. coli*. In general, better expression was obtained at lower growth temperatures (25 or 30 °C) and longer incubation times (∼18 h) after induction with IPTG.

Purification of Recombinant GGGP Synthase. GGGP synthase was purified in three steps: heat precipitation, DE52 anion-exchange chromatography, and Superdex 200 size exclusion chromatography. Incubation of clarified cell extracts at 60 °C for 10 min resulted in the precipitation of a substantial amount of endogenous *E. coli* protein. Most of the remaining protein contaminants were removed by DE52 chromatography. Size exclusion chromatography gave an active multimeric enzyme free from larger protein aggregates, inactive monomeric protein, and unidentified lower MW contaminants.

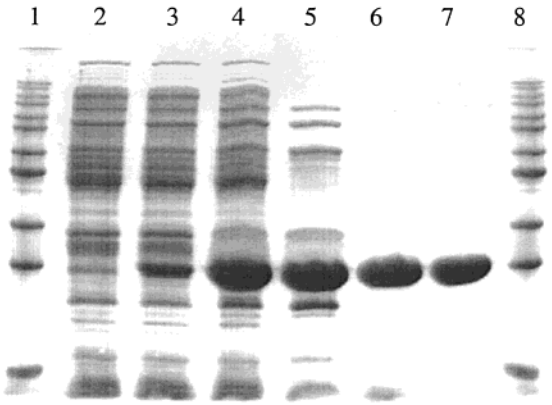


FIGURE 1: SDS−PAGE of samples from the purification of GGGP synthase. Lanes 1 and 8: 10 kDa molecular mass markers (the lowest band corresponds to 20 kDa). Lane 2: Supernatant from a crude extract of *E. coli* BL21-CodonPlus/pET-11b. Lane 3: Supernatant from a crude extract of *E. coli* BL21-CodonPlus/pTJSIII-124, uninduced. Lane 4: Supernatant from a crude extract of *E. coli* BL21-CodonPlus/pTJSIII-124, induced. Lane 5: GGGP synthase after heat purification. Lane 6: GGGP synthase after DE52 chromatography. Lane 7: GGGP synthase after size exclusion chromatography.

Characterization of Recombinant GGGP Synthase. (i) Quaternary Structure of Recombinant GGGP Synthase. Monomeric GGGP synthase has a molecular mass of 25.8 kDa as calculated from the translated amino acid sequence and runs at approximately 29 kDa on SDS−PAGE. Chen and co-workers (*11*) reported that wt GGGP synthase was most likely a pentamer, based on gel filtration chromatography on a calibrated column. We found that recombinant GGGP synthase eluted as a single peak on a calibrated Superdex column at a retention volume corresponding to a molecular mass of 150 ± 25 kDa. Because the Superdex 200 16/60 column used in these experiments would not clearly resolve the individual oligomers in this molecular weight range, we also conducted equilibrium centrifugation experiments. The absorbance vs radius data at nine different protein concentrations fit well to a model of a single oligomeric state with a native molecular mass of 137 kDa, indicating that the catalytically active enzyme is a pentamer (Figure 2).

(ii) Optimal Conditions for Enzyme Activity. Optimal conditions for recombinant GGGP synthase activity were essentially identical to those reported for the native enzyme (*11*). The enzyme had maximal activity at 55 °C in buffer, pH 6−9, containing at least 1 mM Mg^{2+}. Under these conditions, the enzyme is stable for at least 10 min. GGGP synthase was less active when Mg^{2+} was replaced by Mn^{2+} or Zn^{2+}.

(iii) Steady-State Kinetic Constants for Recombinant GGGP Synthase. Under the assay conditions described in Experimental Procedures, GGGP synthase was stable. The

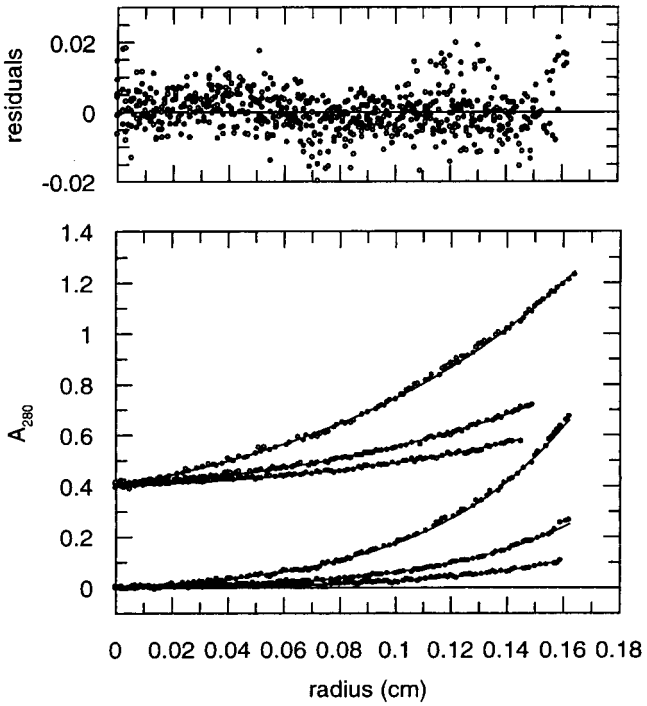


FIGURE 2: Equilibrium centrifugation data for GGGP synthase. Experiments were done at monomer concentrations of 25, 12.5, and 6.25 μM. A_{280} values are referenced to an arbitrary radius r_0. Protein samples were centrifuged at 10000*g* (solid circles) and 8000*g* (open circles; A_{280} values offset for clarity). The quality of fit to a pentamer model with a native molecular mass of 137 kDa was assessed by plotting residuals.

rate of GGGP synthesis was linear for at least 75 s and 20% conversion for low GGPP concentrations (0.5 μM) and up to 90 s and 16% conversion for low GP concentrations (7 μM GP, 20 μM GGPP). At higher substrate concentrations (5 μM GGPP, 100 μM GP), the reaction was linear at least up to 180 s and 20% conversion of GGPP. Under optimal conditions (55 °C in 50 mM Tris, pH 8.0, 10 mM $MgCl_2$, 5 mM BME, 0.2% Tween 80), $V_{max} = 4.0 \pm 0.1$ μmol min^{-1} mg^{-1} ($k_{cat} = 0.34 \pm 0.03$ s^{-1} for pentameric GGGP synthase).

Table 1: Kinetic Data for Recombinant and Native GGGPS[a]

protein	k_{cat} (s^{-1})	K_m^{GGPP} (μM)	K_m^{GP} (μM)	T_{max} (°C)	pH_{max}
native	0.35 ± 0.04	4.1 ± 1.1	41 ± 5	50−65	6.0−7.5
recombinant	0.34 ± 0.03	0.51 ± 0.05	13.5 ± 1.0	50−60	6.0−9.0

[a] Native enzyme was isolated from cell-free extracts of *M. thermoautotrophicum* (Chen et al., 1993). Recombinant enzyme was produced in *E. coli* (this study).

Michaelis constants were $K_m^{(S)-GP} = 13.5 \pm 1.0$ μM and $K_m^{GGPP} = 506 \pm 47$ nM. Values for k_{cat} and $K_m^{(S)-GP}$ determined in this study for the recombinant enzyme were very similar to those reported by Chen and co-workers (*11*). However, K_m^{GGPP} for native GGGP synthase was 8-fold higher than we found for the recombinant enzyme. Upon discovering this difference, we carefully checked the purity and concentration of our substrates. Perhaps the difference is the result of a mutation introduced during PCR cloning, but at this point we have no definitive evidence to explain the difference. These data are summarized in Table 1.

Product Analysis by Electrospray Mass Spectometry. A negative ion electrospray mass spectrum of the product from a large-scale reaction had a prominent peak at $m/z = 443.2$, as expected for GGGP. A peak at $m/z = 715$ was not seen, indicating that diether DGGGP was not formed in this reaction, even in incubations when excess GGPP was used. In a control reaction in which enzyme was omitted, the peak at $m/z = 443.2$ was absent.

Analysis of the Amino Acid Sequence for GGGP Synthase. The gene for GGGP synthase from *M. thermoautotrophicum*, Marburg strain, encodes a protein with 245 amino acids. After completion of much of the genetic work described in this report, the total genome sequence of *M. thermoautotrophicum*, ΔH, was reported (*25*). The putative GGGP synthase gene from the ΔH strain also encodes a protein of 245 amino acids. There is 88% identity between the two DNA sequences and 94% identity between the translated amino acid sequences. The amino acid sequence of GGGP

```
      1            15 16           30 31           45 46           60 61           75 76           90
1 PH1124   -MVGELRIGKVEKYI HEKLEK-KKLHFVLI DPDDTSPEVAGKLAR VCEELGVDAIMVGGS TGAEGEVLDNVVRSI KDNSSLPVILFPGSH  87
2 PAB0735  ----MVKIGKVEMYI NEELESGKKLHFVLI DPDDTHPEMAGRIAE LCENVGVNAIMVGGS TGAEGEMLDNVVKAI KESSSLPVILFPGSH  86
3 MTH552   --------MKVEDYF HDILRE-RKIHLTLI DPEEQTPEEAVEIAR AAIRGGTDGIMLGGS T-TDSSELDNTARAL RENIDVPIILFPGNT  80
4 Ta0995   -------MMTVLEDM LRKTRN-GKVHMTLI DPGAKPPQECARIAE EAEMAGTDFIMVGGS TDIDSRAMDEAISAI KAKTDLKVIIFPGSS  82
5 APE0621  MARLAVKRRRLLEKL LERRSR-GRLHFTLI DPDKTGPGEAGEIAA RAAEAGSDAILVGGS IGVTFEETDGVVKAA K-RSGLPVILFPGGH  88
6 MJ1250   ---MKIKIGKVEKRL NQIIEEGAVYLTLL DPE----EENIEEIA ENVKDYADAIMVGGS I--GIVNLDETVKKI KKITKLPIILFPGNV  81
7 AF0403   --------------- ---MRWRKWRHITKL DPDRTN---TDEIIK AVADSGTDAVMISGT QNVTYEKARTLIEKV S-QYGLPIVVEPSDP  68
                                        * **             *           *   * **                 * *** *

          91          105 106         120 121         135 136         150 151         165 166         180
1 PH1124  GGISRYADAIFFMSL LNSRNPFFITGAQAL GAFTVKK---FGIEP IPMAYIVVEPGE--- -TVGWVGDARPIPRH KPKLAAAYALAGQYL 170
2 PAB0735 SGISKYADAIFFMSL LNSRNPFFITGAQAL GAFTVKR---YGLEP IPMAYIIVEPGE--- -TVGWVGDAKPIPRH KPKLAAAYALAGQYL 169
3 MTH552  TGVSRYADAIFFMSL LNSTNPYWIIGAQAL GAATVKK---MGIEA LPMGYLVVEPGG--- -TVGWVGDTKPVPRN KPDIAAAYAMAAEFL 163
4 Ta0995  LMISPKADAIFFMSL LNSGSLEYVVGHQVK AAIPLSA---MKIEK IPMAYLVFDPGM--- -TVGRVGKAHLIPRD DEKTALSYALAAQYF 165
5 APE0621 TNASRHADAVLFLTV MNSDNPYYIVQAQIL GAPLALK---LGLEA IPTSYIIVGYGG--- -AAGFVARARPIPYE KPELAALHALAGAMM 171
6 MJ1250  DGLSRYADAVFYMSL MNSANTYWVVTAPTL GAITILK---YNLEP IPMAYLCIEPAKKT- -AVGYVGEIREIPQN KPKITAMYCLSAKFF 166
7 AF0403  SNVVYDVDYLFVPTV LNSADGDWITGKHAQ WVRMHYENLQKFTEI IESEFIQIEGYIVLN PDSAVARVTKALCNI DKELAASYALVGEKL 158
            *   ** *   **  **    **          *           *  *   **              *      *        *

          181         195 196         210 211         225 226         240 241         255 256         270
1 PH1124  GMRLVYLEAGSGAPE PVPEEMVRVVKSVID VP-------LIVGGG IKSGEQAKKLIKSGA DIIVTGTAIEKAKSL EEARKRLEAIRNGVF V----- 253
2 PAB0735 GMRLVYLEAGSGSPE PVPPEMVRIVKSVID VP-------LIVGGG IRTGDQVRELTKAGA DIIVTGTAIESTKSI DEAKRKLEEIRRGLK ----- 252
3 MTH552  GMRLFYLEAGSGAPE HVPEEMIALVKRCTD QI-------LIVGGG IRSGEDAARVAGAGA DVVVTGTVVENSDNV EDKIREIVEGMGSV- ----- 245
4 Ta0995  GFRLVYFEAGSGSPY HVGENVVRRVKQELD IP-------VIVGGG IRTPEAAKALAQAGA DMIVTGTIAERSVNV YEALHPIVESIKEVG ISKIQ 248
5 APE0621 GGRIIYLEAGSGAPK PVPPEAVAASRKLVD AAGYGGEVLLTVGGG VRTPEAARMLAEAGA DVLVTGTLAEES--- PGKLADVVEAFKSA- ----- 257
6 MJ1250  GMRWAYLEAGSGASY PVNNETIALSKKLSG IN-------IIVGGG IRKPEIAYEKVLAGA DAIVTGNLLEENPKA VEMMYDAIKKAGKEK ----- 249
7 AF0403  FNLPIIYIEYSGTYG --NPELVAEVKKVLD KAR------LFYGGG IDSREKAREMLRYAD TIIVGNVIYEKG--- ---IDAFLETLP--- ----- 231
            *    *          *   *                 *** *   * *             **   * *
```

FIGURE 3: Alignment of amino acid sequences for putative GGGP synthases from seven archaebacteria. Sequences were identified from a BLAST search with the *M. thermoautotrophicum* sequence as a query. Line 1: *P. horikoshii*. Line 2: *P. abyssi*. Line 3: *M. thermoautotrophicum*. Line 4: *T. acidophilum*. Line 5: *A. pernix*. Line 6: *M. jannaschii*. Line 7: *A. fulgidus*. Asterisks denote identical or similar residues. Bold letters in the *M. thermoautotrophicum* sequence denote conserved residues other than alanine, valine, leucine, and isoleucine. Numbers on the top line of each grouping are global sequence numbers.

synthase from *M. thermoautotrophicum*, strain ΔH, was used in a BLAST search to identify putative GGGP synthase sequences from six other archaeal species. It is likely that the sequence from *A. pernix* actually encodes a farnesylgeranylglyceryl phosphate synthase, based on the observation that the membrane lipids in this species contain C_{25} hydrocarbons (*9*, *10*). An alignment of the sequences reveals significant homology (Figure 3), including five acidic aspartate/glutamate residues and a basic lysine/arginine residue. In addition, there are two conserved prolines, three conserved serine/threonines, three conserved aromatic residues, and five conserved glycines, including a completely conserved GGG motif.

DISCUSSION

One of the most characteristic molecular features that distinguish the Archaea from other organisms is the structure of their membrane lipids. Until recently, little was known about the enzymes responsible for biosynthesis of archaeal lipids. Synthesis of GGGP and DGGGP, the first two committed intermediates in archaeal lipid biosynthesis, was first detected in cell free preparations of *M. thermoautotrophicum* (*3*), where it was found that GGGP synthase was cytosolic and DGGGP synthase was associated with the membrane. Subsequently, the *M. thermoautotrophicum* GGGP synthase was purified (*11*).

GGGP synthase belongs to a diverse class of prenyltransferases that catalyze alkylation of electron-rich centers in nonisoprenoid acceptors. In this case the acceptor unit is a hydroxyl group in (*S*)-GP. Other prenyltransferases in this class alkylate amino groups [DMAPP−tRNA transferase (*14*) and DMAPP−AMP transferase (*15*)], thiols (protein prenyltransferase) (*13*), and aromatic rings (dimethylallyltryptophan synthase) (*16*, *17*). Despite the similarities of the chemistry of the reactions catalyzed by these enzymes, there is little homology among their various amino acid sequences. All lack the characteristic DDxxD aspartate-rich motif found in short-chain isoprene diphosphate synthases and terpene cyclases (*26*−*29*). This motif is thought to play a critical role during catalysis of prenyl transfer by binding and activating the diphosphate group of the allylic substrate through salt bridges with Mg^{2+} ions. It appears that the prenyltransferases with nonisoprenoid acceptors employ different methods for binding and activating the allylic diphosphate. The X-ray structure of protein farnesyltransferase (*12*) indicates that the diphosphate group of FPP interacts with the side chains of lysine, histidine, arginine, and tyrosine residues, and it appears that DMAPP−tRNA transferase might have a similar diphosphate-binding environment (*19*).

Most prenyltransferases are monomers, homodimers, or heterodimers. Except for the very unusual pentameric structure of GGGPase, the only other example of a higher oligomer is dimethylallyl diphosphate:aspulvinone dimethylallyltransferase, which was proposed to be a hexamer on the basis of gel filtration studies (*30*).

We confirmed that GGGP synthase catalyzes transfer of a single geranylgeranyl unit to glyceryl phosphate, even after prolonged incubation in the presence of excess GGPP. The only peak seen for products by electrospray mass spectrometry was at $m/z = 443$, corresponding to the molecular ion for GGGP. No molecular ion was detected at $m/z = 715$ for DGGGP. These observations agree with a previous report (*3*) that the activities for GGGP and DGGGP synthesis resided with two different enzymes that separated into cytosolic and membrane-bound fractions, respectively.

An alignment of putative GGGP synthase sequences from seven archaeal genomes (Figure 3) reveals the existence of several conserved aspartate and glutamate residues. Although none of the aspartates or glutamates constitute motifs similar to the DDxxD sequence seen in many of the enzymes that catalyze prenyl transfer to carbon−carbon double bonds, conserved arginine/lysine, serine/threonine and tyrosine residues raise the possibility that the diphosphate binding pocket in GGGP synthase resembles that of protein farnesyltransferase. In addition, two prolines and five glycines, including a GGG motif, are conserved in GGGP synthase. At this point additional structural studies are needed to characterize the architecture of the active site of this enzyme and to identify the residues that are important for substrate binding and catalysis.

ACKNOWLEDGMENT

We thank Dr. Lacy Daniels for providing cell paste of *M. thermoautotrophicum*, Dr. Elliott Rachlin for electrospray mass spectra, and Dr. David Myszka for conducting the equilibrium sedimentation experiments.

REFERENCES

1. Woese, C. R., and Fox, G. E. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5088−5090.
2. Koga, Y., Nishihara, M., Morii, H., and Akagawa-Matsushita, M. (1993) *Microbiol. Rev. 57*, 164−182.
3. Zhang, D. L., and Poulter, C. D. (1993) *J. Am. Chem. Soc. 115*, 1270−1277.
4. Nishihara, M., and Koga, Y. (1997) *J. Biochem. 122*, 572−576.
5. Nishihara, M.., Yamazaki, T., Oshima, T., and Koga, Y. (1999) *J. Bacteriol. 181*, 1330−1333.
6. Poulter, C. D. Aoki, T., and Daniels, L. (1988) *J. Am. Chem. Soc. 110*, 2620−2624.
7. Chen, A., and Poulter, C. D. (1993) *J. Biol. Chem. 268*, 11002−11007.
8. Ohto, C., Nakane, H., Hemmi, H., Ohnuma, S., Obata, S., and Nishino, T. (1998) *Biosci. Biotechnol. Biochem. 62*, 1243−1246.
9. Tachibana, A., Yano, Y., Otani, S., Nomura, N., Sako, Y., and Taniguchi, M. (2000) *Eur. J. Biochem. 267*, 321−328.
10. Morii, H., Yagi, H., Akutsu, H., Nomura, N., Sako, Y., and Koga, Y. (1999*) Biochim. Biophys. Acta 1436*, 426−436.
11. Chen, A., Zhang, D., and Poulter, C. D. (1993) *J. Biol. Chem. 268*, 21701−21705.
12. Strickland, C. L., Windsor, W. T., Syto, R., Wang, L., Bond, R., Wu, Z., Schwartz, H. V., Beese, L. S., and Weber, P. C. (1998) *Biochemistry 37*, 16601−16611.
13. Harris, C. M., and Poulter, C. D. (2000) *Nat. Prod. Rep. 17*, 137−144.
14. Moore, J. A., and Poulter, C. D. (1997) *Biochemistry 36*, 604−614.
15. Blackwell, J. R., and Horgan, R. (1993) *Phytochemistry 34*, 1477−1481.
16. Tsai, H. F., Wang, H., Gebler, J. C., Poulter, C. D., and Schardl, C. L. (1995) *Biochem. Biophys. Res. Commun. 216*, 119−125.
17. Gebler, J. C., Woodside, A. B., and Poulter, C. D. (1992) *J. Am. Chem. Soc. 114*, 7354−7360.
18. Moore, J. A., Mathis, J. R., and Poulter, C. D. (2000) *Biochim. Biophys. Acta 1479*, 166−174.

19. Soderberg, T., and Poulter, C. D. (2001) *Biochemistry 40*, 1734−1740.
20. Davisson, V. J., Woodside, A. B., Neal, T. R., Stremler, K. E., Muehlbacher, M., and Poulter, C. D. (1986) *J. Org. Chem. 51*, 4768−4779.
21. Reed, B. C., and Rilling, H. C. (1976) *Biochemistry 15*, 3739−3745.
22. Bradford, M. M. (1976) *Anal. Biochem. 72*, 248−254.
23. Matsudaira, P. T. (1989) *A Practical Guide to Protein and Peptide Purification for Microsequencing*, Academic Press, San Diego.
24. Leatherbarrow, R. J. (1998) *Grafit*, Version 4.0, Erithicus Software Ltd., Staines, U.K.
25. Smith, D. R., et al. (1997) *J. Bacteriol. 179*, 5088−5090.
26. Ashby, M. N., and Edwards, P. A. (1990) *J. Biol. Chem. 265*, 13157−13164.
27. Lesburg, C. A., Zhai, G., Cane, D. E., and Christianson, D. W. (1997) *Science 277*, 1820−1824.
28. Starks C. M., Back, K., Chappell, J., and Noel, J. P. (1997) *Science 277*, 1815−1820.
29. Wendt, K. U., Poralla, K., and Schulz, G. E. (1997) *Science 277*, 1811−1815.
30. Takahashi, I., Ojima, N., Ogura, K., and Seto, S. (1978) *Biochemistry 17*, 2696−2702.

BI0111799