THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 33, Issue of August 15, pp. 31078–31087, 2003
Printed in U.S.A.

Identification, Purification, and Characterization of an Eukaryotic-like Phosphopantetheine Adenylyltransferase (Coenzyme A Biosynthetic Pathway) in the Hyperthermophilic Archaeon *Pyrococcus abyssi**

Received for publication, February 24, 2003, and in revised form, May 7, 2003
Published, JBC Papers in Press, May 19, 2003, DOI 10.1074/jbc.M301891200

Jean Armengaud‡§, Bernard Fernandez‡, Valérie Chaumont‡, Françoise Rollin-Genetet‡, Stéphanie Finet¶, Charles Marchetti‡, Hannu Myllykallio‖, Claude Vidaud‡, Jean-Luc Pellequer‡, Simonetta Gribaldo‖, Patrick Forterre‖, and Pierre Gans‡‡**

From the ‡CEA VALRHO, DSV-DIEP, SBTN, Service de Biochimie post-génomique and Toxicologie Nucléaire, 30207 Bagnols-sur-Cèze, France, the ¶European Synchrotron Radiation Facility, BP 220, 38043 Grenoble, France, the ‖Institut de Genetique et Microbiologie, UMR CNRS 8621, Universite Paris-Sud, 91405 Orsay, France, and the ‡‡Institut de Biologie Structurale CEA-CNRS-UJF, Jean-Pierre Ebel, UMR 5075, Laboratoire de Resonance Magnetique Nucleaire, 38027 Grenoble, France

Although coenzymeA (CoA) is essential in numerous metabolic pathways in all living cells, molecular characterization of the CoA biosynthetic pathway in Archaea remains undocumented. Archaeal genomes contain detectable homologues for only three of the five steps of the CoA biosynthetic pathway characterized in Eukarya and Bacteria. In case of phosphopantetheine adenylyltransferase (PPAT) (EC 2.7.7.3), the putative archaeal enzyme exhibits significant sequence similarity only with its eukaryotic homologs, an unusual situation for a protein involved in a central metabolic pathway. We have overexpressed in *Escherichia coli*, purified, and characterized this putative PPAT from the hyperthermophilic archaeon *Pyrococcus abyssi* (PAB0944). Matrix-assisted laser desorption ionization-time of flight mass spectrometry and high performance liquid chromatography measurements are consistent with the presence of a dephospho-CoA (dPCoA) molecule tightly bound to the polypeptide. The protein indeed catalyzes the synthesis of dPCoA from 4′-phosphopantetheine and ATP, as well as the reverse reaction. The presence of dPCoA stabilizes PAB0944, as it induces a shift from 76 to 82 °C of the apparent T_m measured by differential scanning microcalorimetry. Potassium glutamate was found to stabilize the protein at 400 mM. The enzyme behaves as a monomeric protein. Although only distantly related, secondary structure prediction indicates that archaeal and eukaryal PPAT belong to the same nucleotidyltransferase superfamily of bacterial PPAT. The existence of operational proteins highly conserved between Archaea and Eukarya involved in a central metabolic pathway challenge evolutionary scenarios in which eukaryal operational proteins are strictly of bacterial origin.

Coenzyme A (CoA)[1] is a central compound of the Krebs cycle in all living organisms and an essential cofactor in numerous metabolic pathways. The CoA biosynthetic pathway from pantothenate (vitamin B_5) has been described in Bacteria and Eukarya and consists of five enzymatic steps (Fig. 1). The five enzymes are sequentially pantothenate kinase (EC 2.7.1.33), phosphopantothenoylcysteine synthetase (PPCS, EC 6.3.2.5), phosphopantothenoylcysteine decarboxylase (PPCDC, EC 4.1.1.36), phosphopantetheine adenylyltransferase (PPAT, EC 2.7.7.3), and dephospho-CoA kinase (DPCK, EC 2.7.1.24). The genes encoding these five activities have been identified in some bacteria such as *Escherichia coli* (1–6) and in some Eukarya, such as humans (7–9). These identifications allowed a deeper biochemical characterization of the gene products (10, 11). For three of these enzymes, the three-dimensional structure has been resolved (12–17). However, for the third domain of life, Archaea, no investigation on CoA biosynthetic pathway has yet been reported. *In silico* searches in completely sequenced archaeal genomes using known bacterial and eukaryal proteins involved in the CoA biosynthetic pathway allowed identification of archaeal homologues for only three of the five steps, PPCS, PPCD, and PPAT. Putative archaeal PPCS and PPCDC are homologous to their bacterial counterparts (a putative distantly related eukaryal PPCDC homologue was recently identified with a genomic approach, whereas no PPCS homologue could be highlighted (8)), whereas putative archaeal PPAT are only homologous to their eukaryal counterparts. Whereas it is common that archaeal proteins involved in cell division, metabolic or transport pathways (the so-called operational proteins) resemble more their bacterial counterparts, the reverse situation (an eukaryotic-like operational protein) is quite unusual. Putative archaeal PPAT are homologues to the recently characterized PPAT domains of the bifunctional eukaryotic PPAT/DPCK, involved in the last two steps of the coenzyme A biosynthetic pathway (7–9). We had previously

* This work was supported in part by a grant from the Centre National de la Recherche Scientifique (Program Protéomique et Génie des Protéines, Appel à propositions 2001). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed: CEA VALRHO, DSV-DIEP, SBTN, Service de Biochimie post-génomique and Toxicologie Nucléaire, Marcoule, BP 17171, F-30207 Bagnols-sur-Cèze cedex, France. Tel.: 33-4-66-79-68-02; Fax: 33-4-66-79-19-05; E-mail: armengaud@cea.fr.

** Supported by a grant from the Association de la Recherche sur le Cancer.

[1] The abbreviations used are: CoA, coenzyme A; dPCoA, dephosphocoenzyme A; PPCS, phosphopantothenoylcysteine synthetase; PPCDC, phosphopantothenoylcysteine decarboxylase; PPAT, phosphopantetheine adenylyltransferase; DPCK, dephospho-CoA kinase; HPLC, high performance liquid chromatography; MALDI-TOF, matrix-assisted laser desorption ionization-time of flight; SAXS, small angle x-ray scattering; bis-Tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)-propane-1,3-diol; IMAC, immobilized metal ion adsorption chromatography.

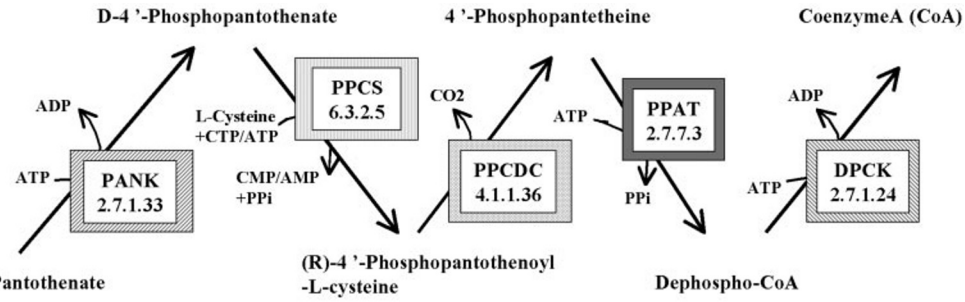

FIG. 1. **CoA biosynthetic pathway and architecture of domains specifying the five enzymatic steps in four organisms.** EC numbers are indicated for each enzymatic step. *PANK*, pantothenate kinase.

characterized this archaeal protein as an interesting target for structural genomic projects (18), and listed it as a PACE (Protein of Archaea Conserved in Eukarya) protein (PACE11), with a general nucleotidyltransferase activity in our PACE data base.[2] To determine whether PACE11 is a PPAT, we have characterized this protein (PAB0944) from the hyperthermophilic archaeon *Pyrococcus abyssi*. Here we describe the purification of a recombinant N terminus His-tagged form of PAB0944 and its biochemical characterization. Indeed, the overexpressed product is shown to exhibit PPAT activity. This is the first biochemical identification of a specific step involved in CoA biosynthesis in Archaea, strongly suggesting that a complete CoA biosynthetic pathway is also operational in the third domain of life.

MATERIALS AND METHODS

Chemical, Biological, and Microbiological Reagents—Most chemicals used in this study were obtained from Sigma and were of analytical grade. Oligonucleotide primers were synthesized and HPLC purified by Genaxys. Restriction enzymes were from New England Biolabs, AmpliTaq DNA polymerase was from Applied Biosystem. P3 medium used in a pH-regulated fermentor consisted of 4 g/liter KH_2PO_4, 4 g/liter K_2HPO_4, 7 g/liter Na_2HPO_4, $12H_2O$, 1.2 g/liter $(NH_4)_2SO_3$, 0.2 g/liter NH_4Cl, 25 g/liter yeast extract, 5 g/liter NZ CAZE Plus supplemented with 4 ml of trace elements solution. The latter consisted of a aqueous solution containing 10 g/liter $FeSO_4$, $7H_2O$, 10 g/liter $CaCl_2$, $2H_2O$, 2.5 g/liter $MnSO_4$, H_2O, 2.5 g/liter $AlCl_3$, $6H_20$, 1 g/liter $CoCl_2$, $6H_2O$, 0.5 g/liter $ZnSO_4$, $7H_2O$, 0.5 g/liter Na_2MoO_4, $2H_2O$, 0.25 g/liter $CuCl_2$, $2H_2O$, 0.125 g/liter H_3BO_3, and 25 ml/liter HCl. Media were supplemented when necessary with the following antibiotics: ampicillin (50 μg/ml), tetracycline (10 μg/ml), and chloramphenicol (20 μg/ml).

Strains, Plasmids, and DNA Manipulation—E. coli strain DH5α (Invitrogen) was used for cloning and both *E. coli* strains BL21-Codonplus(DE3)-RIL (Stratagene) and Rosetta(DE3)pLysS (Novagen) were checked for protein overproduction. Vector pCR T7/NT-topo (Invitrogen) containing a T7 promoter, His_6 tag, and Xpress epitope was used for cloning and expression. The usual DNA manipulations were done according to standard procedures. PCR fragments were extracted from agarose gel by means of QiaexII gel extraction kit (Qiagen) and plasmid DNA was purified with Qiawell 8 kit (Qiagen), as recommended by the supplier. Nucleotide sequencing of both DNA strands was carried out by using a Quick Start Beckman Coulter DTCS sequencing kit with double-stranded templates. Samples containing fluorescence-labeled dideoxynucleotide terminators were cleaned by ethanol precipitation and processed on a CEQ2000 XL Beckman Coulter capillary-based automated sequencer. Sequences were compiled and analyzed with the Vector NTI software package (Informax Inc).

Cloning and Overexpression of PAB0944 Gene—The synthetic oligonucleotide primers oAA98 (5′-gctagcATGAAGTTCAAGAAAGTTGTC-3′) and oAA99 (5′-ggatccTTACTCTCTTTTTTATTGGGTT-3′) contain engineered *Nhe*I and *Bam*HI sites (restriction sites underlined in the primer sequences and nucleotides not present in the original sequence are shown by lowercase), respectively. They were used for PCR amplification of *PAB0944* with *P. abyssi* total DNA as template. The 483-bp fragment obtained after amplification was electrophoretically separated on a 1.5% SeaKem GTG-agarose gel and extracted prior cloning into pCR T7/NT-topo. The resulting plasmid pSBTN-AB31 was selected for correct orientation by restriction mapping using *Bam*HI and *Eco*RI restriction enzymes. As expected, the PCR fragment was shown by sequencing to be in-frame with the His tag and Xpress sequences, both located at the N terminus. Plasmid pSBTN-AB31 was subjected to digestion by *Nhe*I and ligated with T4 DNA ligase. The plasmid pSBTN-AB30 obtained after transformation with the ligation mixture resulted in suppression of the Xpress epitope. Both constructs were verified by DNA sequencing to ascertain the integrity of the nucleotide sequence. Hyperexpression of the recombinant *PAB0944* construct was achieved either with *E. coli* BL21(DE3)Codonplus-RIL or Rosetta(DE3)pLysS strains, both freshly transformed with pSBTN-AB30. Cultures were carried out at 30 °C in LB medium in baffled flasks or P3 medium supplemented with 30 g/liter of glucose in a 4-liter Labfors fermentor (Infors). They were induced with 1 mM isopropyl-1-thio-β-D-galactopyranoside when $A_{600\ nm}$ reached 0.6 (flasks) or 2.5 (fermentor). After 3 h of induction, cells were harvested by centrifugation (4,000 $\times$ *g*, 20 min, 4 °C) and washed once with 50 mM Tris-HCl buffered at pH 8.0 and maintained at 4 °C. Cell pellets (typically 118 g of wet material for a 4-liter culture) obtained by a second centrifugation were frozen in liquid nitrogen and kept several days at −80 °C.

Purification of Recombinant His-tagged PAB0944 Protein—The purification of recombinant PAB0944 was performed from packed cells equivalent to 4 liters of culture. Buffer A consisted of 50 mM Na_2HPO_4/KH_2PO_4 buffer (pH 7.2). The pellet was thawed on ice and resuspended in 300 ml of buffer A containing 10 mM NaCl. The cells were disrupted by sonication with a total energy delivered of 131 kJ. The cell extract was then centrifuged at 30,000 $\times$ *g* for 30 min at 4 °C to remove cellular debris and aggregated proteins. The supernatant was subjected to a 20-min heat treatment using a water bath maintained at 60 °C, and immediately centrifuged a second time at 30,000 $\times$ *g* for 30 min at 4 °C. Subsequent steps were performed at room temperature. NaCl was added to the supernatant to reach a final concentration of 100 mM. The sample was applied at a flow rate of 2.8 ml/min onto a XK 26 $\times$ 20 column (Amersham Biosciences) containing 50 ml of chelating Sepharose Fast Flow (Amersham Biosciences) and prelabely loaded with 200 mM $NiSO_4$, washed with Milli-Q water, and equilibrated with Buffer A containing 150 mM NaCl (Buffer B) and 50 mM imidazole. The column was washed at a flow rate of 4.9 ml/min with 4 column volumes of buffer B and then developed with a 8-column volume linear gradient from 50 to 300 mM imidazole in buffer B using an Äkta Purifier 100 FPLC system (Amersham Biosciences). The His-tagged PAB0944 protein was eluted as a 310-ml fraction at ~250 mM imidazole. The protein sample was concentrated to a volume of 30 ml by means of three Pellicon XL Ultracel-PLCCC5 (3 $\times$ 50 cm^2) filtration units assembled on a Labscale TFF system (Amicon, Millipore) and then dialyzed overnight at 4 °C against Buffer A containing 50 mM NaCl and 1 mM EDTA. The purified protein was stored at −80 °C at a concentration of 1.0 mg/ml. Eventually, a second chromatography step was carried out. The sample was melted and dialyzed overnight at 4 °C against 20 mM KH_2PO_4/Na_2HPO_4 buffered at pH 7.2 (Buffer C). The supernatant obtained after centrifugation at 30,000 $\times$ *g* for 15 min at 4 °C was applied onto a 6-ml Resource-S ion exchange column (30 $\times$ 16 mm, 15 μm) from Amersham Biosciences, previously equilibrated with buffer C and operated at a flow rate of 3 ml/min. After a 10-column volume wash with Buffer C, proteins were resolved with a 20-column volume linear gradient from 0 to 500 mM NaCl in Buffer C. Recombinant PAB0944 was eluted at ~400 mM NaCl and desalted by overnight dialysis against Buffer A containing 50 mM NaCl and 1 mM EDTA. When necessary, the purified protein was concentrated on a Centriplus YM3 or microcon YM3 filtration units (Amicon, Millipore). Buffer exchange was performed either by dialysis or by gel filtration by means of Micro Bio-Spin P-30 chromatography columns (Bio-Rad).

[2] www-archbac.u-psud.fr/projects/pace/paceproteins.html.

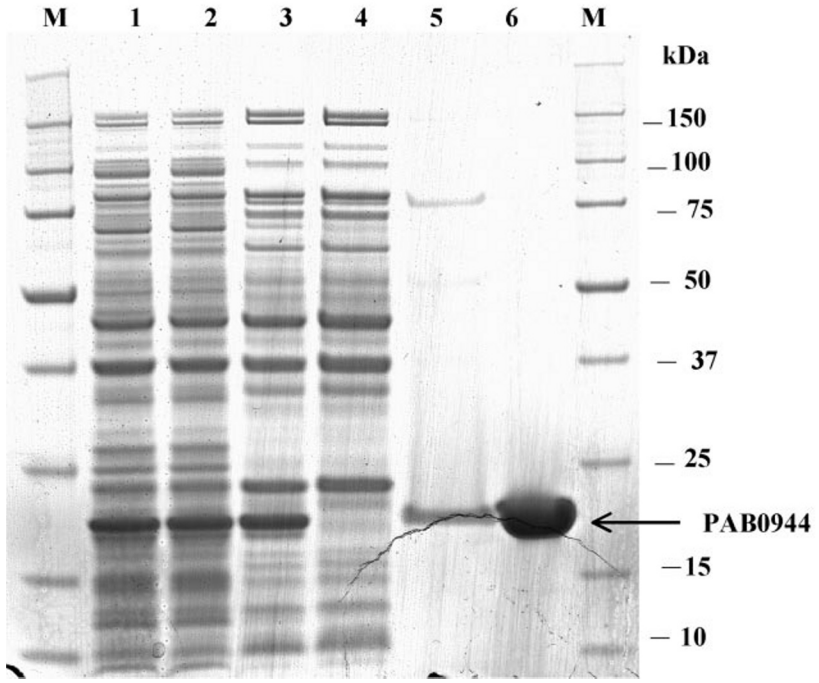

FIG. 2. **SDS-PAGE analysis of expression and purification of recombinant PAB0944.** SDS-PAGE was performed on a 4–12% gradient Novex bis-Tris acrylamide gel (Invitrogen) and stained with BlueSafe Coomassie solution (Bio-Rad). *Lane M*, molecular weight markers; *lane 1*, cell-free extract of *E. coli* Rosetta(DE3)(pLysS)(pSBTN-AB30); *lane 2*, soluble proteins from cell-free extract; *lane 3*, soluble proteins after the 60 °C heat treatment; *lane 4*, proteins not retained by IMAC chromatography; *lane 5*, total proteins eluate from the IMAC column; *lane 6*, Resource-S eluate. Bands corresponding to recombinant PAB0944 polypeptide are indicated with the *arrow*.

Biochemical Analytical Methods—Protein concentration was determined either by the Bradford method using the Bio-Rad Protein Assay (Bio-Rad) and bovine serum albumin as standard or by measuring the absorbance at 280 nm. Molar absorption coefficients of 28,000 and 18,000 M^{-1} cm^{-1} were estimated at 265 (λ_{max}) and 280 nm, respectively, from UV-visible spectra after determination of protein concentration by total amino acid analysis. The samples were dried and then hydrolyzed at 110 °C in constant boiling HCl containing 1% (v/v) phenol, for 24 h under reduced pressure. Amino acids were analyzed on a Beckman amino acid analyzer model 6300, with the standard sodium citrate eluting buffer system. The analyzer was calibrated using a standard solution of amino acids between 500 and 5000 pmol. N-terminal amino acid sequencing was performed by Edman degradation on 7 pmol of recombinant His-tagged PAB0944 by means of a Procise protein sequencing system (Applied Biosystems) coupled with a 140C microgradient pump system (Applied Biosystems) and a Series 200 UV-visible detector (PerkinElmer Life Sciences). UV-visible absorption spectra were acquired on a Cary 50 scan UV-visible spectrophotometer (Varian) operated at 25 °C.

Circular Dichroism and Secondary Structure Prediction—Circular dichroism spectra were recorded at 25 °C between 190 and 250 nm on a Jobin-Yvon CD6 spectro-dichrograph, using a quartz cuvette of 1-mm path length, with a 2-s integration time for each 0.5-nm step and a bandwidth of 2 nm. Two spectra of purified His-tagged PAB0944 at 7 μM in 2 mM KH_2PO_4/Na_2HPO_4 buffer (pH 7.2) containing 2 mM NaCl were averaged and corrected from the baseline for buffer solvents. Spectra were analyzed using the program K2D[3] described by Andrade *et al.* (19). Secondary structure predictions were obtained through the PSIpred v2.4 web-interfaced facilities[4] described by McGuffin *et al.* (20) or through the IBPC consensus program.[5]

Nuclear Magnetic Resonance Spectroscopy—Unlabeled PAB0944 sample was prepared at a concentration of 0.1 mM in a 10 mM KH_2PO_4/Na_2HPO_4 buffer (pH 7.2), and 10% D_2O. The sample was degassed, argon-saturated, and sealed before analysis. One-dimensional NMR spectra were recorded at various temperatures between 25 and 55 °C on a Varian Inova 400 MHz spectrometer equipped with a triple-resonance (^1H, ^{13}C, ^{15}N) probe including shielded z-gradients. Spectra were treated with the Felix97 software (Accelrys).

Determination of Native Molecular Mass by Gel Filtration—The native molecular mass of His-tagged PAB0944 was estimated by gel filtration chromatography on a Superdex 75 gel packed into a HR10/30 column (Amersham Biosciences) with a final bed volume of 23 ml. The column was equilibrated at room temperature at a flow rate of 1.0 ml/min with 50 mM Tris/HCl buffer (pH 8.0), containing 400 mM potassium glutamate and eluted with the same buffer. Protein standards used to calibrate the column were ribonuclease A (15.8 kDa), chymotrypsinogen A (21.2 kDa), ovalbumin (49.4 kDa), and albumin (69.8 kDa), all from Amersham Biosciences. Exclusion limit was evaluated with dextran blue 2000 (Amersham Biosciences). Samples consisting of 350 μl of PAB0944 at 0.285 and 0.855 mg/ml were injected. Specific absorption at 280 and 266 nm were followed and the ratio $A_{280\ nm}/A_{266\ nm}$ was found consistent with the elution of the major peak being PAB0944 in the two runs.

Mass Spectrometry—Matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry was carried out with a Biflex IV instrument (Brüker Daltonik). Protein samples were applied to the target using sinapinic acid prepared as saturated solution in 30% acetonitrile, 70% Milli-Q water, and 0.1% trifluoroacetic acid as matrix. Mass spectra were recorded in the positive ion mode from ~250 laser flashes. The instrument was calibrated for determination of entire protein masses using either a mixture of chymotrypsin and bovine serum albumin, or apomyoglobin and aldolase. The calibrants were mixed with the samples (internal calibration) or used separately (external calibration). Tryptic digested samples were desalted by means of a $ZipTip_{C18}$ pipette tip (Millipore) according to the protocol specified by the manufacturer. Tryptic samples were then mixed directly onto the target with an equal volume of a saturated solution of α-cyano-4-hydroxycinnamic acid in 30% acetonitrile containing 0.1% trifluoroacetic acid. In this case, the instrument was externally calibrated using a pepmix calibration kit (Brüker Daltonik) and then, internally calibrated using some of the theoretical peptide masses. Spectra were evaluated with the BioTools package (Brüker Daltonik) and identification of PAB0944 was done using the MASCOT search engine (Matrix Science).

Differential Scanning Calorimetry—To determine the transition temperature T_m of His-tagged PAB0944, microcalorimetry measurements were carried out with a high-sensitivity differential scanning microcalorimeter VP-DSC from MicroCal. Prior to the calorimetric analysis, the samples were diluted into the reference buffers used in the experiment (10 mM KH_2PO_4/Na_2HPO_4 buffer (pH 7.2), containing 1 mM EDTA and 10 mM NaCl or 50 mM Tris/HCl buffer (pH 8.0), containing 400 mM potassium glutamate), and then degassed. Recombinant His_6-tagged PAB0944 was analyzed at a protein concentration comprised between 45 and 70 μM as determined by spectrophotometry. The calorimetric scans were carried out between 40 and 85 °C with a heating scan rate of 50 K/h in 0.51 ml of cells, against the reference buffer.

HPLC Analysis—Samples were analyzed using an Agilent 1100 Series reverse-phase HPLC system equipped with G1315B diode array detector, G1322A degasser and G1311A quaternary pump, G1329A autosampler, and sample cooler unit operated at 4 °C. The different

[3] www.embl-heidelberg.de/~andrade/k2d.
[4] bioinf.cs.ucl.ac.uk/psipred.
[5] npsa-pbil.ibcp.fr.

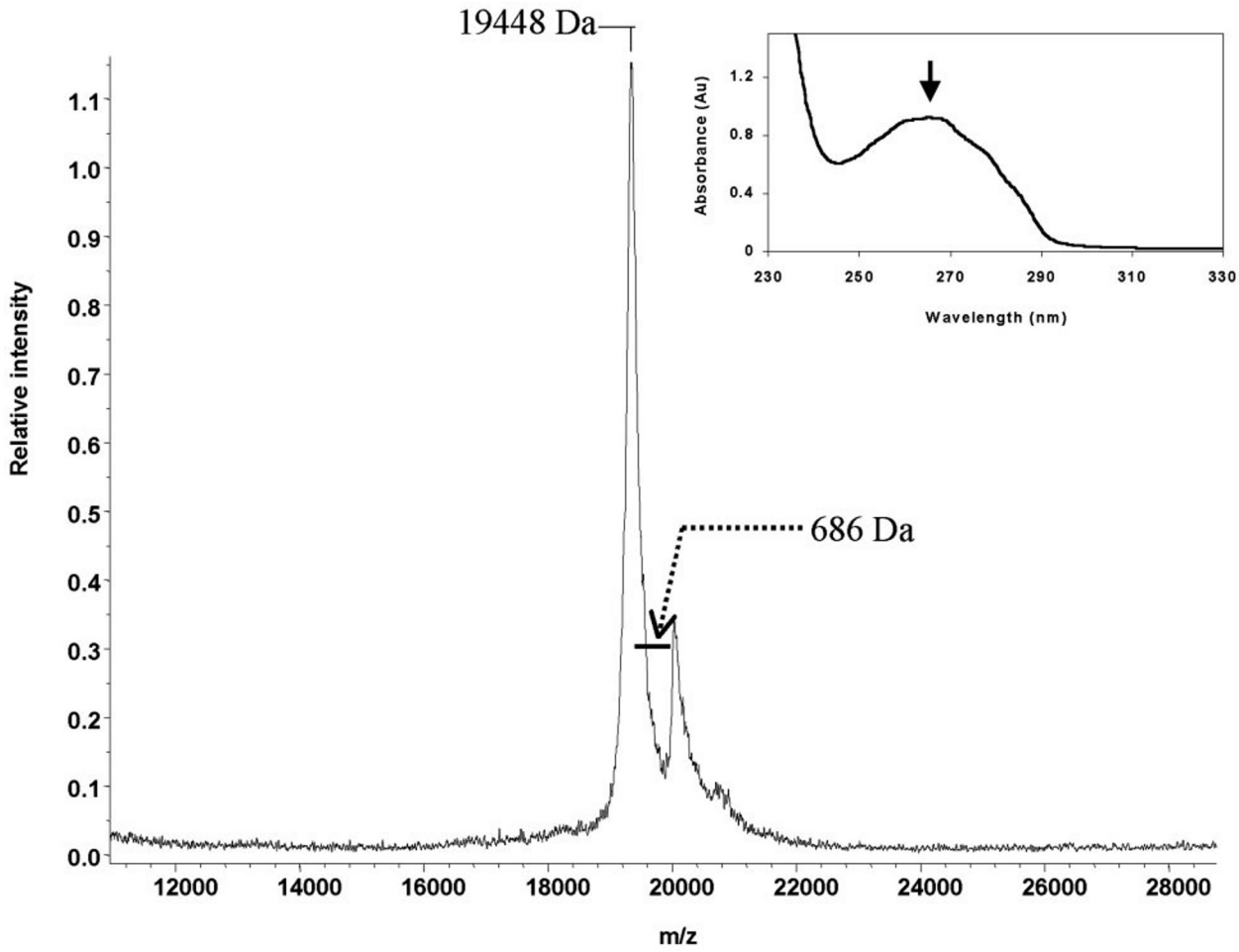

FIG. 3. **Molecular mass characterization of purified PAB0944 by MALDI-TOF.** The protein samples were mixed with sinapinic acid. Measurements from 250 laser shots were averaged. Detection and analysis of cations was performed on a Brüker Reflex III MALDI-TOF mass spectrometer. The distance between the two main peaks was estimated in two separate experiments. The UV-visible absorption spectrum of recombinant PAB0944 is shown in the *inset*. The spectrum was acquired in a 1-cm path length cuvette between 230 and 330 nm at a scan rate of 60 nm/min with a 0.5-nm step. Baseline was corrected against buffer. Protein sample was at a concentration of 33 μM in 10 mM KH_2PO_4/Na_2HPO_4 buffer (pH 7.2) and containing 10 mM NaCl. Maximum at 265 nm is indicated by a *vertical arrow*.

components from the enzymatic reaction were resolved on a Supelcosil LC-18 column (50 × 4.6 mm, 5 μm) from Supelco protected by a pre-column filter and a Supelguard LC-18 guard column (20 × 4.6 mm, 5 μm). Isocratic separation was performed with a filtered aqueous solvent system containing 90 mM NaH_2PO_4/Na_2HPO_4 buffer (pH 5.5), 7.2 mM tetrabutylammonium bromide, and 10% acetonitrile at a flow rate of 1.5 ml/min as described by Daugherty *et al.* (8). Dithiothreitol was added to cooled samples to reach a final concentration of 1 mM and 40 μl of sample were analyzed. Absorbance was monitored between 190 and 500 nm with the diode array detector. Identities of ATP, dPCoA, CoA, and 4′-phosphopantetheine were confirmed by comparison of elution time and UV spectrum with those of authentic standards. Molar areas were determined using a calibration with these standards in the same chromatographic conditions.

Enzyme Assays—The reverse PPAT reaction assay was first assessed at 25 °C by coupling the release of ATP to the reduction of NADP to NADPH and detection at 340 nm as described previously (21). dPCoA was incubated at a concentration of 200 μM with 2 mM PP_i, 5 mM glucose, 1 mM NADP, 2 units of hexokinase, 1 unit of glucose-6-phosphate dehydrogenase, 2 mM $MgCl_2$, and 1.7 μg/ml purified His-tagged PAB0944 in 50 mM Tris/HCl buffer (pH 8.0). Change in absorbance at 340 nm was monitored over 5 min.

To follow the enzymatic activity at higher temperature and with various conditions of substrates and buffers, samples were prepared into thin-wall 0.2-ml tubes and were incubated in a Tgradient thermocycler (Biometra) operated with a 110 °C pre-heated lid to avoid any solvent evaporation and thus, undesired variation of product concentration. In this case, dPCoA was incubated at a concentration of 5–500 μM with 2 mM PP_i, 2 mM $MgCl_2$, and variable amounts (typically 1.7 μg/ml) of purified His-tagged PAB0944 in 50 mM Tris/HCl buffer (pH 8.0, 25 °C) containing 400 mM potassium glutamate. After reaction, samples were immediately cooled in an ice-melting water bath and kept at 4 °C until HPLC analyzed. The components of the enzyme assays were resolved by direct HPLC analysis, monitoring at 255 nm formation of ATP, and concomitant consumption of dPCoA.

The forward phosphopantetheine adenylyltransferase activity assay was verified by detecting the release of dPCoA. Briefly, 1 mM 4′-phosphopantetheine was incubated with 0.1 mM ATP, 4 mM dithiothreitol, 2 mM $MgCl_2$ and various amount of purified His-tagged PAB0944 in 50 mM Tris/HCl buffer (pH 8.0) containing 400 mM potassium glutamate. Samples were treated as above and products from the reaction were resolved by reverse-phase chromatography as described above.

Small Angle X-ray Scattering (SAXS)—The SAXS experiments were carried out using the high brilliance beamline ID2 at the European Radiation Synchrotron Facility in Grenoble, France (22). The experiments in solution were performed at 20 °C using the ID2 flow-through cell (2 mm-diameter and 10 μm-thickness quartz capillary, from GLAS) operated under vacuum that could be filled and rinsed *in situ*. The sample volume was 40 μl and the sample concentration was either 75 or 86 μM for His-tagged PAB0944 in 50 mM Tris/HCl buffer (pH 8.0), with and without 400 mM potassium glutamate, respectively. Data were collected using a two-dimensional detector (x-ray Image Intensifier coupled to an ESRF-developed FReLoN CCD camera, 1024 × 1024 pixels 2 × 2 binned). The sample to detector distances were 2.5 and 1 m, yielding a q-increment per channel of 0.004168 and 0.010082 nm^{-1}, respectively. The amplitude of the scattering vector **s** is defined as $Q = 2\pi s = 4\pi \sin\theta/\lambda$, where λ is the wavelength of the X-rays and 2θ is the scattering angle. With $\lambda = 0.0995$ nm, the total useful Q range was $0.08544 < Q < 6.5\ nm^{-1}$. The exposure time was 100 or 500 ms (3 to 10 frames for the protein, 10 to 20 frames for the buffer with 5 s pause

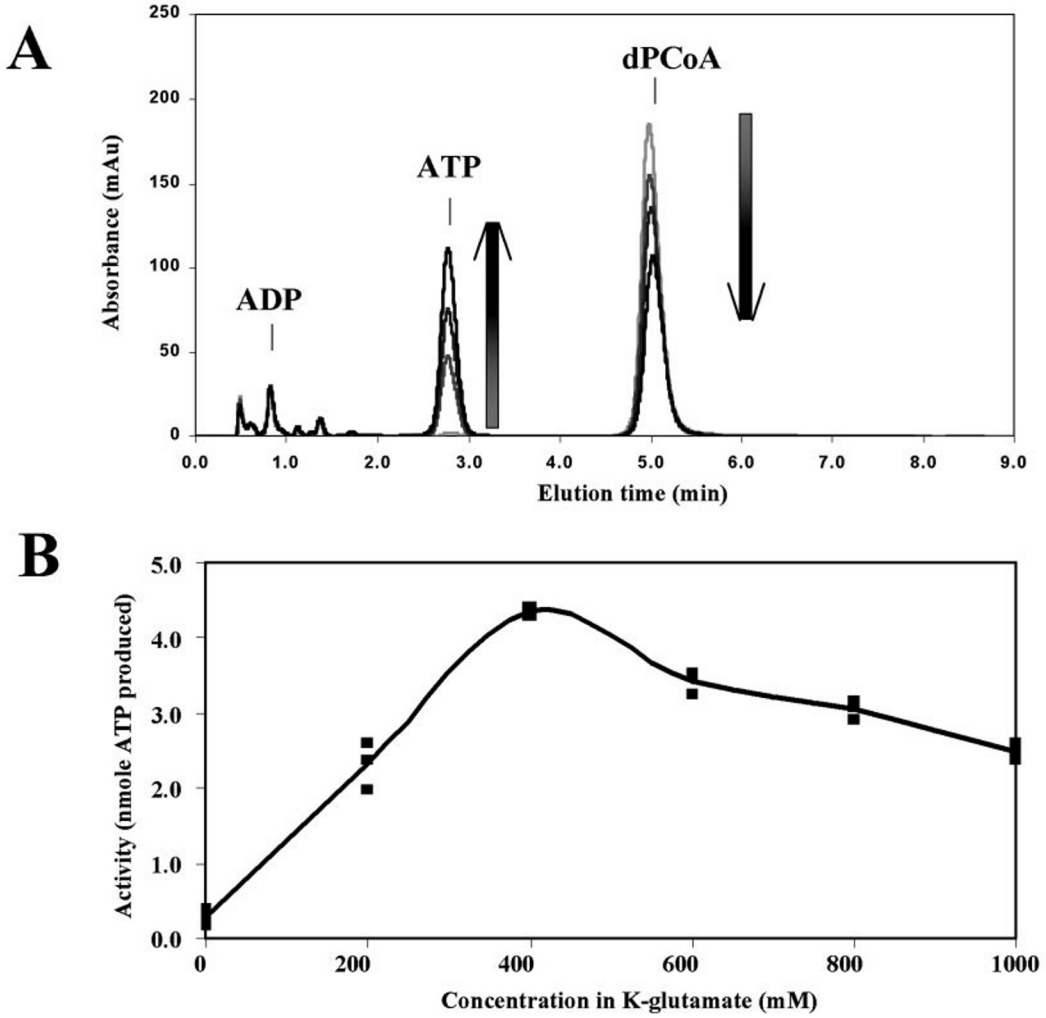

FIG. 4. **Reverse PPAT activity assay.** *A*, synthesis of 4′-phosphopantetheine and ATP from PP_i and dPCoA analyzed using an HPLC-based assay. Samples were incubated for 0 (*light line*), 1 (*medium line*), 2 (*dark line*), and 4 (*darkest line*) min in the presence of PAB0944, diluted in fresh dithiothreitol solution, and subjected to isocratic separation on a Supelcosil LC-18 reverse-phase column. Pure dPCoA, ATP, ADP, and phosphopantetheine solutions resolved in similar conditions were used as standards. Absorbance at 255 nm was used to follow the elution of these compounds. *Arrows* indicate the evolution over time of the concentration of the two detectable compounds. *B*, influence of potassium glutamate on reverse PPAT activity. The reverse PPAT reaction was carried out in 50 mM Tris/HCl buffer (pH 8.0) containing 0–1000 mM potassium glutamate. PAB0944 (88 nM) was incubated at 85 °C with 0.25 mM dPCoA, 2 mM inorganic phosphate, 2 mM $MgCl_2$. The total activity after 4 min from 40 μl of the reaction was estimated in triplicate assays by HPLC measurements.

between each frame). After the two-dimensional data reduction, the curves were averaged when no radiation damage or bubbles were observed.

RESULTS

Putative Function of PAB0944 and Purification of the His-tagged Recombinant Form—Recently, the human PPAT protein domain was identified from its association with a DPCK domain and on biochemical evidence (7–9). This PPAT domain is universally present in Eukarya and it is homologous to putative archaeal nucleotidyltransferases (PACE 11), whereas it harbors no homologues in Bacteria. To determine whether these archaeal homologues are indeed PPAT, the corresponding protein from *P. abyssi* (PAB0944) was produced as an N-terminal His-tagged recombinant form. An *E. coli* T7-based overexpression system was constructed using for *PAB0944* as first codon the most probable ATG located at 1380906 (numbering referred to whole *P. abyssi* genomic sequence). Apparently, genome annotators have considered that the upstream TTG codon was encoding the initial methionine. However, TTG codons are used as initiators only in a few open reading frames. Moreover, the two initial methionines are not conserved in archaeal homologues, and at least in *E. coli*, the first formylmethionine would probably be processed by the methionine-aminopeptidase leaving only one of the two residues. For these reasons, the N terminus sequence of PAB0944 would be rather Met-Lys-Phe- than Met-Met-Lys-Phe-. Therefore, the His tag construct was based on the former sequence.

The *P. abyssi* chromosomal gene comprises a particularly high rate of *E. coli* rare codons: 32 over a sequence of 156 amino acids (21%) and three tandem rare codon triple repeats. For this reason two *E. coli* strains, BL21-Codonplus(DE3)-RIL (Stratagene) and Rosetta(DE3)pLysS (Novagen), were checked for protein overexpression. These two strains carry extra tRNA genes for arginine, isoleucine, and leucine codons commonly found in many Archaea including *Pyrococcus* but rarely used by *E. coli*, the latter containing in addition an extra tRNA gene for proline codons. No major difference in expression was detected with the two strains. An overexpressed product with an apparent molecular weight of ~19,000 was obtained upon addition of isopropyl-1-thio-β-D-galactopyranoside (Fig. 2, *lane 1*). The PAB0944 polypeptide was mainly produced as a soluble protein (Fig. 2, *lane 2*). This was the case whatever the temperature used during induction (37 or 30 °C).

Because of the usual growth conditions of *P. abyssi,* the

protein was supposed to be thermostable. The crude extract was heated at different temperatures and the remaining soluble proteins were analyzed on SDS-PAGE. PAB0944 was shown to remain soluble until at least 70 °C and precipitated at temperature above 80 °C. The crude extract heated to 60 °C led to the precipitation of some contaminants from the host without any significant lost of PAB0944 (Fig. 2, *lane 3*). This behavior simplified the design of the purification strategy. A two-step purification protocol including a 20-min heat treatment at 60 °C and IMAC chromatography was first developed. The degree of purity obtained after IMAC chromatography was verified by means of SDS-PAGE. The main IMAC fraction was judged almost pure. An additional Resource-S ion exchange chromatography (Fig. 2, *lane 6*) was carried out for removing the last minor contaminants from some of the IMAC fractions. The choice of this last chromatography was based on the large excess of basic residues in the polypeptide sequence, the calculated isoelectric point of the His-tagged form being 10.1.

A dPCoA Molecule Is Tightly Bound to the His-tagged PAB0944—For determination of the protein N terminus sequence, Edman degradation was carried out on purified protein over 7 cycles. The sequence obtained, Met-Arg-Gly-Ser-His-His-His, clearly indicates that the initial methionine had not been removed in *E. coli*. Such a result was expected for this construct because the second amino acid residue in the polypeptide chain is an arginine residue that sterically hinders the action of the methionine-aminopeptidase (23, 24). To make sure that the recombinant protein had not undergone some unexpected modification, the purified sample was subjected to mass spectrometric analysis (Fig. 3). An experimental mass of 19,448 Da was measured for the polypeptide. This experimental value is in the range of the expected polypeptide mass calculated from the amino acid sequence of His-tagged PAB0944, including the unprocessed initial methionine (19,435 Da). Another significant signal was detected higher than the major peak (centered at a distance from the main signal between 675 and 697 atomic units depending on two separate sets of measurements). This signal could neither be attributed to matrix (sinapinic acid) or salt adducts, nor to another polypeptide. The presence of a detectable contaminant from *E. coli* of almost the same size as PAB0944 was indeed excluded as a tryptic digest of the sample analyzed by MALDI-TOF could lead only to the identification of PAB0944.

Although no amino acid sequence similarity is detected between CoaD from *E. coli* and PAB0944 a similar function may be fulfilled by the two proteins, based on the common function of CoaD and human DPCK/PPAT. Geerlof *et al.* (5) have shown that homologous overexpression of an enterobacterial *coaD* gene in *E. coli* led to the production of a protein containing a CoA molecule. The UV spectrum measured for His-tagged PAB0944 is reported in the *inset* of Fig. 3. The absorbance of the polypeptide at 280 nm is surprisingly high: 18,000 M^{-1} cm^{-1} compared with the theoretical molar absorption coefficient of 6,400 M^{-1} cm^{-1} expected from the residues content (4 tyrosines, no tryptophan). The maximum expected absorption at 276 nm is shifted at 265 nm. This indicates the presence of an UV-absorbing chromophore bound to the protein. By analogy with CoaD produced in *E. coli*, the presence of a metabolite from the CoA synthetic pathway (CoA or dPCoA) trapped into the polypeptide was therefore suspected. The mass difference (686 atomic units in average) between the two peaks measured by MALDI-TOF mass spectrometry of purified sample corresponds to the theoretical mass of 687.6 Da of dPCoA indicating that the chromophore is a dPCoA molecule, rather than a CoA molecule (theoretical mass of 767.6 Da). The presence and identity of this ligand was confirmed by a different method. A

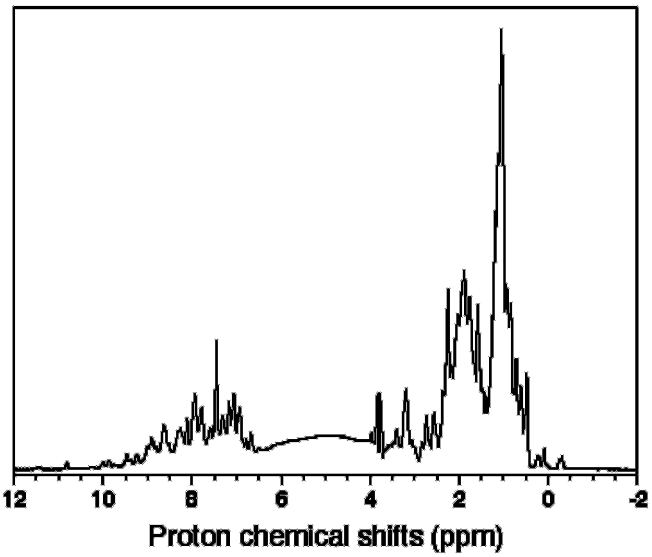

FIG. 5. **One-dimensional NMR spectrum of recombinant PAB0944.** The spectrum of PAB0944 in 10 mM KH_2PO_4/Na_2HPO_4 buffer (pH 7.2) and containing 10 mM NaCl was acquired at 55 °C and 400 MHz. The spectral width was 5600 Hz and the number of accumulations was 512. The observation of resolved peaks in particular below 0 and above 9 ppm indicates a stable three-dimensional structure.

sample of recombinant PAB0944 was heat-treated for several minutes at 99 °C and then subjected to centrifugation. In these experimental conditions, the protein is irreversibly denatured and precipitates. The supernatant was concentrated 5-fold by evaporation in a speed-vacuum concentrator and analyzed by HPLC. The products released by this treatment were resolved on a LC18 reverse phase column and compared with standards leading to unequivocal identification of dPCoA as chromophore (data not shown). At the difference of CoaD purified from *E. coli*, extensive dialysis of the sample against 50 mM citrate buffer (pH 5.0) remained unsuccessful for removal of the ligand.

PPAT Activity—Reverse PPAT reaction assay was first followed by spectrophotometry measurements at 25 °C in a reaction where the release of ATP is coupled to the reduction of NADP to NADPH. Experimental conditions were the same as those described by Worrall and Tubbs (21). A low but significant increase of absorbance at 340 nm was observed, thus indicating that the recombinant protein was able to transform dPCoA into ATP and 4′-phosphopantetheine. The nature of the products released during the reaction was confirmed by HPLC measurements and comparison with authentic standards. To follow enzymatic activity at higher temperature, reaction conditions were simplified by omitting glucose, NADP, hexokinase, and glucose-6-phosphate dehydrogenase. $MgCl_2$ was added at 1 mM in all reaction buffers as this cation is generally implicated into the catalytic mechanism of nucleotidyltransferases (17) and was found here to stabilize the protein. Enzymatic kinetics were then checked only by reverse phase chromatography. It has been first verified that in assay conditions without protein the substrates and products of the reaction were stable at 80 °C over the first 60 min. As illustrated in Fig. 4, in the presence of catalytic quantities of recombinant PAB0944 disappearance of dPCoA was observed and was always found concomitant to appearance of ATP. In these assays, 4′-phosphopantetheine was not detected because of its low UV absorbance. An important increase of activity was measured when temperature was shifted from 25 °C to higher values. A specific activity of 0.17, 0.43, 5.15 (±5%) μmol of ATP produced per min per mg of purified PAB0944 was measured at 25, 37, and 70 °C, respec-

tively. The forward PPAT reaction was also confirmed by HPLC measurement. Specific appearance of dPCoA was determined concomitantly with the disappearance of ATP.

Archaeal enzymes are usually found more stable in high ionic strength solution and in the presence of osmolyte such as betaine, or salts such as sodium or potassium glutamate (25, 26). For this reason, reverse PPAT activity was measured for PAB0944 in 50 mM Tris/HCl buffer (pH 8.0) containing different concentrations of potassium glutamate (Fig. 4). The maximal activity was obtained when the K^+ concentration was 400 mM, a value near that of intracellular K^+ concentration in *Pyrococcus* (27). In the assay conditions, the activity drastically decreases in absence of this salt while concentrations higher than 400 mM although not optimal have less effect on the activity.

In Vitro Stability—To investigate the thermostability of PAB0944, differential scanning microcalorimetry was carried out for different buffer conditions: 50 mM KH_2PO_4/Na_2HPO_4 buffer (pH 7.2); 50 mM KH_2PO_4/Na_2HPO_4 buffer (pH 7.2), containing 400 mM potassium glutamate; 50 mM Tris/HCl buffer (pH 8.0) at 25 °C, containing 400 mM potassium glutamate. Thermal unfolding was found in all these conditions to be a irreversible process. Hence, only an apparent midpoint melting temperature could be determined. The presence of two distinct forms in the recombinant PAB0944 sample was observed in the calorimetric melting curve because a shoulder was visible at ~75 °C in addition to the maximum centered at 82 °C. These two distinct forms may probably correspond to polypeptide bound to dPCoA and free-polypeptide. dPCoA-free enzyme was obtained by incubating the purified protein at 60 °C with 2 mM PP_i for 20 min, prior Sephadex G-25 gel filtration. HPLC analysis of the supernatant of the 99 °C heated reaction solution show total disappearance of dPCoA and concomitant appearance of ATP. Measurements in both phosphate buffers indicate that the dPCoA-free enzyme exhibits a T_m of 76 °C. No noticeable influence on the thermal stability of PAB0944 was observed in the presence of 400 mM potassium glutamate. A higher T_m (82 °C) was observed when 0.2 mM dPCoA was added into the buffer. A role of thermal stabilization is therefore supposed for the dPCoA ligand. Although Tris/HCl buffer is not adapted to such measurements (dpH/dT = −0.028 units/°C) but was used in enzymatic assays, we checked that similar T_m values could be obtained in the absence or presence of 0.2 mM dPCoA (74 and 82 °C, respectively). The T_m measured for the recombinant product (76 and 82 °C in the presence of the dPCoA ligand) is relatively low for a protein originating from a hyperthermophilic microorganism (28). This behavior may be explained by the addition of the His_6 tag that may affect protein stability or by a drastic change of protein environment from *P. abyssi* cytoplasm to *in vitro* buffer conditions.

Structural Information—Archaeal, eukaryal, and bacterial PPAT all belong to the cytidylyltransferase superfamily, which includes structurally related (CATH 3.40.50.620) but functionally diverse nucleotide-binding enzymes (29). Members of this superfamily harbor a common dinucleotide-binding fold (or canonical Rossmann-binding fold (30)) consisting of a five-stranded parallel β−pleated sheet (16). To estimate the content of secondary structure elements in PAB0944, near-UV CD spectroscopy was performed. The recombinant PAB0944 protein presents negative ellipticity in the near-UV with minima at 208 and 218 nm (data not shown). Deconvolution of the CD spectrum using the K2D neural software leads to an estimation of the content of secondary structure elements of about 35% of α-helix and 20% of β-sheets. Prediction of His-tagged PAB0944 secondary structures gives similar global values (34–42% of α-helix and 15–18% of β-sheets) in fair agreement with the CD data. The predicted secondary structure shows a pattern of α-helices and β-strands (as discussed below) consistent with its inclusion in the cytidylyltransferase superfamily. The one-dimensional NMR spectra of PAB0944 (Fig. 5) not only corroborates the presence of a large extent of secondary structures but also indicates the existence of a stable three-dimensional structure. This is illustrated in Fig. 5 by the large dispersion of the resonances in particular in the amide region (above 9 ppm) and below 1 ppm. No significant changes were observed in the resonance positions between 25 and 55 °C, but a decrease of the line width, in agreement with the PAB0944 thermostability observed in this temperature range.

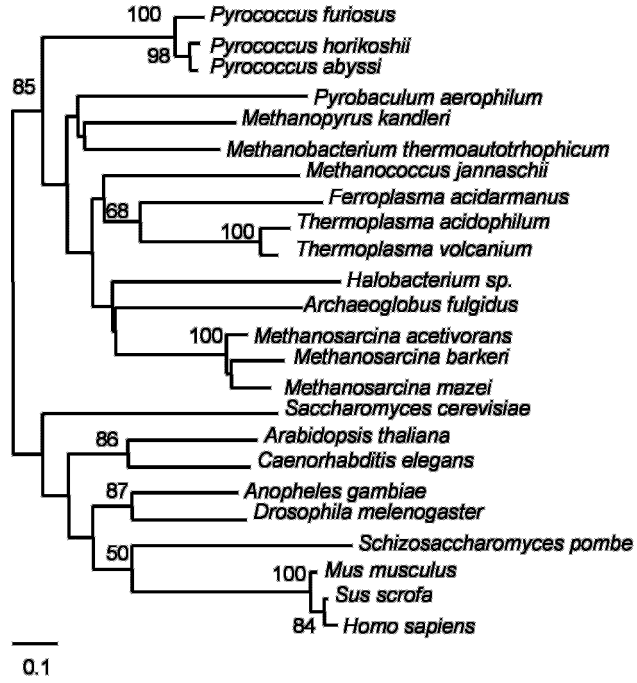

FIG. 6. **Phylogenetic analysis of archaeal and eukaryal PPATs.** Archaeal and eukaryal homologues were obtained from public data bases (www.ncbi.nlm.nih.gov/) by PsiBlast searches. Multiple alignments were performed by ClustalW and manually refined. Following removal of a few ambiguously aligned regions, a data set was assembled comprising 24 sequences totaling 127 amino acid positions. An unrooted evolutionary distance tree was constructed based on Kimura distances and the neighbor joining tree reconstruction algorithm. Bootstrap confidence levels at nodes were computed by the RELL method. Values lower than 50% are not displayed. *Scale bar* represents unit of amino acid substitutions per position. Accession numbers are as follows: *Archaeoglobus fulgidus* NP_071031.1; *Anopheles gambiae* EAA11969.1; *Arabidopsis thaliana* NP_179417.1; *Caenorhabditis elegans* NP_490766.2; *Drosophila melanogaster* NP_647985.1; *Ferroplasma acidarmanus* ZP_00001331.1; *Halobacterium* species NP_444219.1; *Homo sapiens* NP_079509.3; *Methanosarcina acetivorans* NP_618428.1; *Methanosarcina barkeri* ZP_00076768.1; *Methanococcus jannaschii* NP_248024.1; *Methanopyrus kandleri* NP_614787.1; *Methanosarcina mazei* NP_632482.1; *Methanobacterium thermoautotrophicum* NP_276996.1; *Mus musculus* NP_082172.2; *Pyrococcus abyssi* NP_127102.1; *Pyrobaculum aerophilum* NP_127102.1; *Pseudomonas aeruginosa* NP_558918.1; *Pyrococcus furiosus* NP_578813.1; *Pyrococcus horikoshii* NP_142583.1; *Saccharomyces cerevisiae* NP_011793.1; *Schizosaccharomyces pombe* NP_594334.1; *Sus scrofa* AAM19997.1; *Thermoplasma acidocaldarius* NP_394650.1; *Thermoplasma volcanium* NP_110922.1.

PAB0944 Behaves as a Monomeric Protein—All PPAT enzymes characterized from Bacteria and Eukarya have been shown to be multimeric enzymes (5, 21). To verify whether a similar behavior is harbored by archaeal PPAT, the native molecular mass of PAB0944 was determined in the presence of 400 mM potassium glutamate. Size-exclusion chromatography on a Superdex 75 calibrated column revealed that PAB0944 eluted as a peak centered at 14.3 ml in the assay conditions correspond-

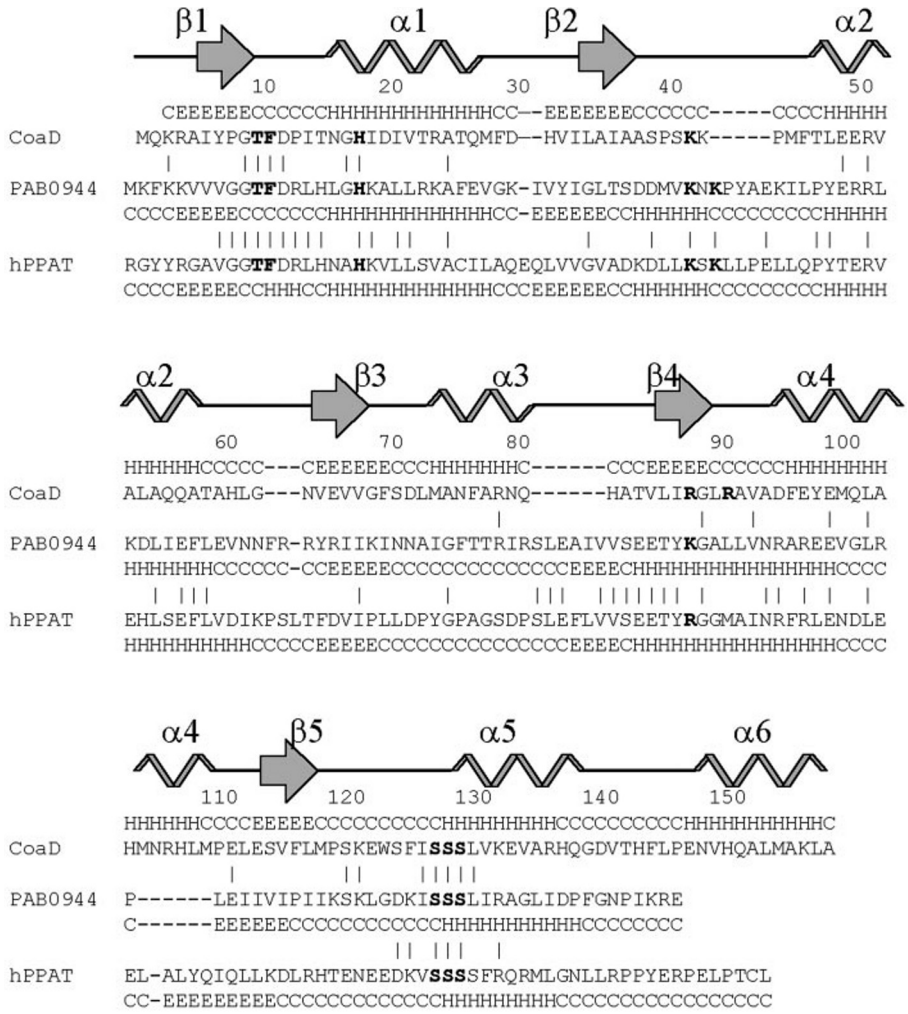

FIG. 7. **Sequence comparison between PAB0944 from *P. abyssi*, CoaD from *E. coli*, and the human PPAT domain.** Prediction of secondary structure elements of PAB0944 were made through the Psipred web-interfaced software (Footnote 4) with the 1b6t structure of CoaD from *E. coli* as model. The alignment between the three polypeptides was manually corrected. Numbering is indicated for the CoaD sequence from *E. coli*. Important residues cited under "Discussion" are indicated in *boldface*. β-Sheets (*E*) are represented by *arrows* and α-helices (*H*) are numbered based on the CoaD structure. Identical residues between PAB0944 and CoaD (14%) as well as between PAB0944 and the human PPAT domain (32%) are indicated by *I*.

ing to an apparent molecular mass of 20 kDa. This elution profile indicates that the protein behaves as a monomer.

SAXS gives structural information on inhomogeneities of electron density. It is a convenient tool to probe the low-resolution structure of macromolecules in solution with characteristic dimensions of about 10 to a few thousand Ångstroms. This technique gives access to the mean particle size (with the radius of gyration, R_g) and the molecular mass. The Guinier approximation, valid for dilute solutions in the Guinier region where $QR_g < 1$ (31), allows to express the scattering intensity $I(Q) \approx I(0) \exp(-Q^2R_g^2/3)$, where I(0) is the forward intensity. The radius of gyration and I(0) are obtained, respectively, from the slope and the intercept of the linear fit of $\ln(I(Q))$ *versus* Q^2 at low angles and the intensity at zero angle scaled by the protein concentration is proportional to its molecular mass. Recombinant PAB0944 was analyzed in two buffer conditions: 50 mM KH_2PO_4/Na_2HPO_4 buffer (pH 7.2) and 50 mM KH_2PO_4/Na_2HPO_4 buffer (pH 7.2) containing 400 mM potassium glutamate. Values for R_g of 19 and 19.4 were measured in these two buffers, respectively. The deduced corresponding molecular masses are 21.0 and 17.1 kDa in the two conditions. These values are in agreement with a monomeric state for the recombinant protein.

DISCUSSION

The *PAB0944* gene from the euryarchaeote *P. abyssi* was cloned into an overexpression system and a N-terminal His-tagged form of the encoded protein was purified. The protein overproduced by *E. coli* retains a dPCoA ligand through purification. This protein exhibits a PPAT activity. The PAB0944 protein represents the first enzyme characterized in the archaeal pathway of CoA biosynthesis. The gene coding for this enzyme is present in all completely sequenced archaeal genomes, except those of *Sulfolobus* species and *Aeropyrum pernix*. Archaeal PPATs are homologous to eukaryotic PPATs, although in Eukarya these proteins are fused to DPCK to form a bifunctional protein. Phylogenetic analysis shows that archaeal and eukaryal PPAT form two distinct monophyletic groups (Fig. 6), suggesting that these proteins are orthologous; *i.e.* they have diverged from an ancestral PPAT that was present in the last common ancestor of Archaea and Eukarya.

Eukaryal and archaeal PPAT have no noticeable sequence similarity with bacterial PPAT (encoded by the *coaD* gene in *E. coli*). However, secondary structure prediction suggests that all PPAT belongs to the same superfamily of nucleotidyltransferases (Fig. 7). All these proteins are structurally based on a canonical dinucleotide-binding fold composed of two Rossmann folds. Alignment of the predicted β-sheets in archaeal/eukaryal PPAT on those of the structurally well characterized *E. coli* PPAT, elements constituting the β-sheet core of the structure, suggests several major differences between archaeal/eukaryal and bacterial versions of PPAT (Fig. 7). First, an insertion of five residues takes place in archaeal/eukaryal PPAT between the second β-sheet (β2) and the second α-helix (α2) of bacterial PPAT. Although automatically predicted to form an α-helix (Fig. 7), because of local environment constraints this insertion was built as an external loop in the three-dimensional model shown in Fig. 8. Second, the last α-helix (α6) of bacterial PPAT

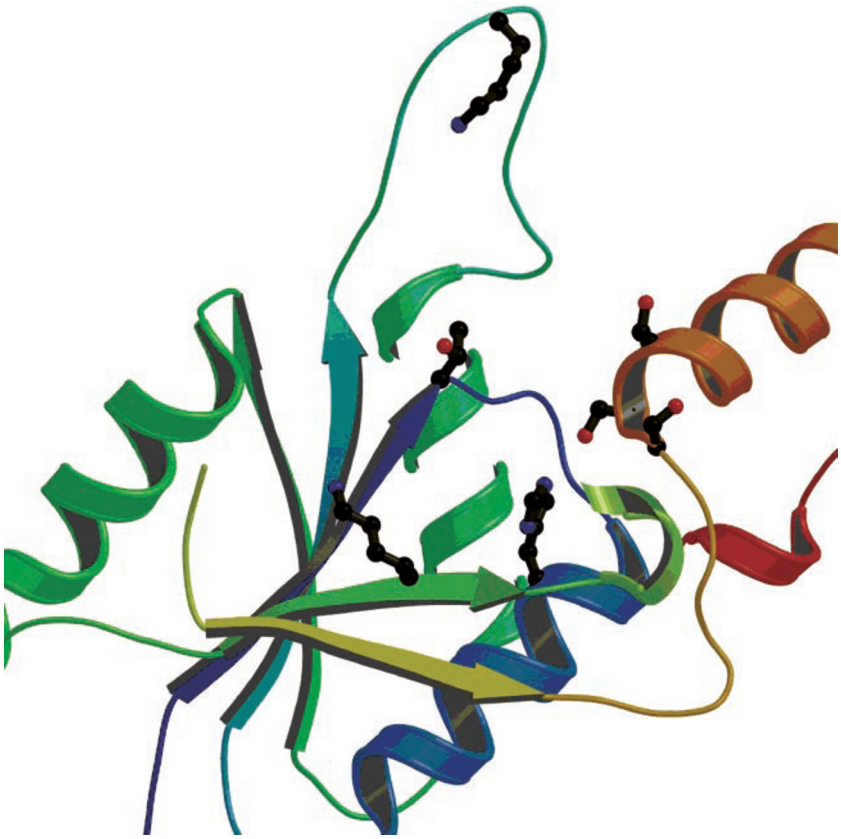

FIG. 8. **Ribbon diagram of the PAB0944 comparative model based on the homologous template CoaD of *E. coli*.** Sequence alignment was manually adjusted and the model was built using published procedures (34). The model is color-coded according to a rainbow gradient where *blue* represents the N terminus and *red* the C terminus. Critical side chains involved in putative binding to dPCoA are drawn as *balls and sticks*. The figure was prepared with the MOLSCRIPT program (35) and rendered using RASTER3D (36).

is absent in archaeal/eukaryal homologues. Last, the third α-helix (α3) in bacterial PPAT seems to be absent in archaeal/eukaryal PPAT.

Three-dimensional structure studies on *E. coli* PPAT (17) have identified several residues essential for substrates binding or catalysis. His_{18} plays a key role in state transition stabilization and belongs to the T/H*X*GH motif found in all the members of the nucleotidyltransferases superfamily. Residues Phe^{11}, Arg^{91}, Ser^{128}, Ser^{129}, and Ser^{130} are involved in binding of ATP, Thr^{10}, and Lys^{42} are implicated in its orientation, and Arg^{88} interacts with the 4′-phosphate of phosphopantothenate. Except for the N-terminal T/H*X*GH motif, classical alignment methods do not allow the identification of these residues. In contrast, the alignment using the predicted secondary structures places correctly the second part of the ATP-binding site (ISSSL motif). A three-dimensional comparative model of the *P. abyssi* PPAT (Fig. 8) indicates that Lys^{46} in *P. abyssi* PPAT may play the role of *E. coli* PPAT Lys^{42} in the orientation of ATP. In our model, Lys^{103} in *P. abyssi* PPAT may replace the important residue Arg^{88} (conserved in all the members of this superfamily) in interaction with phosphopantothenate. These considerations strongly suggest that *P. abyssi* PPAT and its archaeal and eukaryal homologues share the same global fold rather than other nucleotidyltransferases, but significantly differ from the bacterial model.

This point is further corroborated by the experimental differences observed between *P. abyssi* and *E. coli* PPAT. The latter is purified with a CoA molecule that can be removed by gel filtration after extensive dialysis against citrate buffer. In contrast, when purified from *E. coli*, *P. abyssi* PPAT scavenges a dPCoA molecule that cannot be removed using the same procedure. Whether this difference is because of conditions of expression (media and *E. coli* strains are different) or affinity of each protein for the CoA biosynthetic pathway intermediates remains unclear.

Gel filtration experiments as well as SAXS spectroscopy indicate that *P. abyssi* PPAT behaves as a monomer in solution. In contrast, the *E. coli* PPAT is hexameric in solution and in solid state as observed by sedimentation equilibrium experiments and x-ray crystallography, respectively (5, 16). The PPAT protein from another bacterium, *Brevibacterium ammoniagenes*, has been shown to behave as a trimeric protein (32). Thus, the fact that *P. abyssi* PPAT behaves as a monomeric protein further differentiates archaeal and bacterial PPAT. This behavior also distinguishes the *P. abyssi* PPAT from its eukaryal homologue because the bifunctional human PPAT/DPCK protein is a homodimer (21). The monomeric state and thermostability of the *P. abyssi* PPAT should facilitate the structural analysis of this family of proteins.

The CoA synthetic pathway has been thoroughly described in Bacteria, and more recently in Eukarya (7–9). In Archaea, of the five enzymatic steps necessary in this lower pathway starting from pantothenate, only two may be deduced with a good level of confidence from comparative genomics (Fig. 1). The bifunctional bacterial PPCS/PPCDC enzyme, responsible of L-cysteine ligation to phosphopantothenate and decarboxylation of the resulting product, is also present in Archaea as a bifunctional protein (COG0452), and resides in two separate proteins in Eukarya. At present, no homologues of the first enzyme of the pathway, pantothenate kinase (COG1072), have been identified in Archaea. The last enzyme of the pathway, DPCK, is fused to PPAT into a bifunctional enzyme in some Eukarya such as humans. This DPCK domain is quite well conserved (COG0237) in Eukarya and Bacteria (Fig. 1), but Psi-Blast analysis was unsuccessful in identifying the corresponding proteins in Archaea. In *P. abyssi*, several candidates for such activity may be proposed, such as PAB1012 and PAB1725 (14% identity with DPCK) both proposed to belong to COG0237. From analysis of the genetic environment of all the archaeal homologues of *PAB0944* (PPAT activity in *P. abyssi*) and

PAB1897 (PPCS/PPCDC activity in *P. abyssi*), no further indications on possible archaeal candidates for pantothenate kinase and DPCK activities can be proposed. These considerations suggest that the proteins performing pantothenate kinase and DCPK activities remain to be identified in Archaea, and may exhibit major three-dimensional structure differences with their corresponding proteins from Bacteria and Eukarya.

In recent comparative genomic studies, it was shown that the phylogenetic distribution pattern of proteins across the three domains of life is somewhat correlated with their function. It was found that informational proteins (involved in replication, transcription and translation) are usually closely related between Archaea and Eukarya, whereas operational proteins (involved in metabolic pathways, transport or cell division) are often closely related between Eukarya and Bacteria (33). It is thus unusual that an archaeal operational protein, such as PPAT, is "eukaryotic-like." The specific relatedness of many eukaryal and bacterial operational proteins has been interpreted in two ways: either it reflects gene transfers from the ancestral mitochondrial and chloroplast genomes to the nucleus, or it testifies for an even more ancient evolutionary event, *i.e.* the original association of an archaeal cell (proto-nucleus) and a bacterial cell (proto-cytoplasm) to give birth to the first eukaryotes. In such a scenario, all eukaryotic operational proteins would have originated from Bacteria. The existence of an operational archaeal protein that is eukaryotic-like, such as PPAT, is clearly at odds with this scenario, and indicates that Eukarya and Archaea have inherited both informational and operational proteins from their last common ancestor.

Acknowledgments—We gratefully acknowledge J.-P. Andrieu (IBS) for performing the analysis of total amino acids, L. Bellanger (SBTN) for useful advices on media for high-cell density cultures, I. Dany and A. Dedieu (SBTN) for performing several MALDI-TOF mass spectrometry experiments, J.-C. Gaillard (SBTN) for N-terminal amino acid sequencing of the protein, Y. Zivanovic for helpful discussions, and Eric Quéméneur for support and critical reading of the manuscript.

REFERENCES

1. Song, W. J., and Jackowski, S. (1992) *J. Bacteriol.* **174,** 6411–6417
2. Kupke, T., Uebele, M., Schmid, D., Jung, G., Blaesse, M., and Steinbacher, S. (2000) *J. Biol. Chem.* **275,** 31838–31846
3. Strauss, E., Kinsland, C., Ge, Y., McLafferty, F. W., and Begley, T. P. (2001) *J. Biol. Chem.* **276,** 13513–13516
4. Clementz, T., and Raetz, C. R. (1991) *J. Biol. Chem.* **266,** 9687–9696
5. Geerlof, A., Lewendon, A., and Shaw, W. V. (1999) *J. Biol. Chem.* **274,** 27105–27111
6. Mishra, P., Park, P. K., and Drueckhammer, D. G. (2001) *J. Bacteriol.* **183,** 2774–2778
7. Zhyvoloup, A., Nemazanyy, I., Babich, A., Panasyuk, G., Pobigailo, N., Vudmaska, M., Naidenov, V., Kukharenko, O., Palchevskii, S., Savinska, L., Ovcharenko, G., Verdier, F., Valovka, T., Fenton, T., Rebholz, H., Wang, M. L., Shepherd, P., Matsuka, G., Filonenko, V., and Gout, I. T. (2002) *J. Biol. Chem.* **277,** 22107–22110
8. Daugherty, M., Polanuyer, B., Farrell, M., Scholle, M., Lykidis, A., de Crecy-Lagard, V., and Osterman, A. (2002) *J. Biol. Chem.* **277,** 21431–21439
9. Aghajanian, S., and Worrall, D. M. (2002) *Biochem. J.* **365,** 13–18
10. Kupke, T. (2002) *J. Biol. Chem.* **277,** 36137–36145
11. Song, W. J., and Jackowski, S. (1994) *J. Biol. Chem.* **269,** 27051–27058
12. Yun, M., Park, C. G., Kim, J. Y., Rock, C. O., Jackowski, S., and Park, H. W. (2000) *J. Biol. Chem.* **275,** 28093–28099
13. Obmolova, G., Teplyakov, A., Bonander, N., Eisenstein, E., Howard, A. J., and Gilliland, G. L. (2001) *J. Struct. Biol.* **136,** 119–125
14. O'Toole, N., Barbosa, J. A., Li, Y., Hung, L. W., Matte, A., and Cygler, M. (2003) *Protein Sci.* **12,** 327–336
15. Izard, T., Geerlof, A., Lewendon, A., and Barker, J. J. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55,** 1226–1228
16. Izard, T., and Geerlof, A. (1999) *EMBO J.* **18,** 2021–2030
17. Izard, T. (2002) *J. Mol. Biol.* **315,** 487–495
18. Matte-Tailliez, O., Zivanovic, Y., and Forterre, P. (2000) *Trends Genet.* **16,** 533–536
19. Andrade, M. A., Chacon, P., Merelo, J. J., and Moran, F. (1993) *Protein Eng.* **6,** 383–390
20. McGuffin, L. J., Bryson, K., and Jones, D. T. (2000) *Bioinformatics* **16,** 404–405
21. Worrall, D. M., and Tubbs, P. K. (1983) *Biochem. J.* **215,** 153–157
22. Nayaraman, T., Diat, O., and Bösecke, P. (2001) *Nucl. Instr.* **467,** 1005–1009
23. Dalboge, H., Bayne, S., and Pedersen, J. (1990) *FEBS Lett.* **266,** 1–3
24. Hirel, P. H., Schmitter, M. J., Dessen, P., Fayat, G., and Blanquet, S. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86,** 8247–8251
25. Hethke, C., Bergerat, A., Hausner, W., Forterre, P., and Thomm, M. (1999) *Genetics* **152,** 1325–1333
26. Lamosa, P., Burke, A., Peist, R., Huber, R., Liu, M. Y., Silva, G., Rodrigues-Pousada, C., LeGall, J., Maycock, C., and Santos, H. (2000) *Appl. Environ. Microbiol.* **66,** 1974–1979
27. Scholz, S., Sonnenbichler, J., Schafer, W., and Hensel, R. (1992) *FEBS Lett.* **306,** 239–242
28. Vieille, C., and Zeikus, G. J. (2001) *Microbiol. Mol. Biol. Rev.* **65,** 1–43
29. Bork, P., Holm, L., Koonin, E. V., and Sander, C. (1995) *Proteins* **22,** 259–266
30. Rossmann, M. G., Liljas, A., Branden, C.-I., and Banaszak, L. J. (1975) in *The Enzymes* (Boyer, P. D., ed) 3rd Ed., Vol. 11, Academic Press, New York
31. Guinier, A., and Fournet, G. (1955) *Small-angle Scattering of X-rays*, John Wiley, New York
32. Martin, D. P., and Drueckhammer, D. G. (1993) *Biochem. Biophys. Res. Commun.* **192,** 1155–1161
33. Jain, R., Rivera, M. C., and Lake, J. A. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 3801–3806
34. Chen, S. W. W., and Pellequer, J.-L. (2003) *Curr. Med. Chem.*, in press
35. Kraulis, P. J. (1991) *J. Appl. Cryst.* **24,** 946–950
36. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol.* **277,** 505–525