The pathway for the synthesis of di-myo-inositol-phosphate (DIP) was recently elucidated on the basis of the detection of the relevant activities in cell extracts of Archaeoglobus fulgidus and structural characterization of products by nuclear magnetic resonance (NMR) (N. Borges, L. G. Gonçalves, M. V. Rodrigues, F. Siopa, R. Ventura, C. Maycock, P. Lamosa, and H. Santos, J. Bacteriol. 188:8128–8135, 2006). Here, a genomic approach was used to identify the genes involved in the synthesis of DIP. Cloning and expression in Escherichia coli of the putative genes for CTP:myo-inositol-1-phosphate cytidylyltransferase and DIPP (di-myo-inositol-1,3′-phosphate-1′-phosphate, a phosphorylated form of DIP) synthase from several (hyper)thermophiles (A. fulgidus, Pyrococcus furiosus, Thermococcus kodakaraensis, Aquifex aeolicus, and Rubrobacter xylanophilus) confirmed the presence of those activities in the gene products. The DIPP synthase activity was part of a bifunctional enzyme that catalyzed the condensation of CTP and myo-inositol-1-phosphate into CDP-myoinositol, as well as the synthesis of DIPP from CDP-myoinositol and L-myo-inositol-1-phosphate. The cytidylyltransferase was absolutely specific for CTP and L-myo-inositol-1-P; the DIPP synthase domain used only L-myo-inositol-1-phosphate as an alcohol acceptor, but CDP-glycerol, as well as CDP-L-myo-inositol and CDP-D-myo-inositol, were recognized as alcohol donors. Genome analysis showed homologous genes in all organisms known to accumulate DIP and for which genome sequences were available. In most cases, the two activities (L-myo-inositol-1-P cytidylyltransferase and DIPP synthase) were fused in a single gene product, but separate genes were predicted in Archaeoglobus fulgidus, Thermotoga maritima, and Hyperthermus butylicus. Additionally, using L-myo-inositol-1-phosphate labeled on C-1 with carbon 13, the stereochemical configuration of all the metabolites involved in DIP synthesis was established by NMR analysis. The two inositol moieties in DIP had different stereochemical configurations, in contradiction of previous reports. The use of the designation di-myo-inositol-1,3′-phosphate is recommended to facilitate tracing individual carbon atoms through metabolic pathways.

Di-myo-inositol phosphate (DIP) is the most widespread organic solute in microorganisms adapted to hot environments (27). This compound was first identified by Scholz and coworkers in the archaean Pyrococcus woesei (28) and was later encountered in many other hyperthermophiles, including members of the genera Methanotorris, Thermococcus, Thermotoga, Aquifex, Pyrodictium, Aeropyrum, Archaeoglobus, Stetteria, and Pyrolobus) (6, 11–13, 15, 16, 27). Additionally, DIP occurs as a minor solute in the thermophilic bacteria Rubrobacter xylanophilus and Persephonella marina (10, 27). Thus far, DIP has never been encountered in organisms with optimal growth temperatures below 60°C, and hence, the assumption that it plays a role in the thermoprotection of cellular components in vivo is often stated (22, 26, 28).

The question concerning the contribution of compatible solutes from hyperthermophiles to the mechanisms of osmo- and thermoadaptation can only be answered once the pathways for the synthesis of these solutes are known in detail, i.e., once the genes, enzymes, substrates, and reaction products are fully characterized. With this goal in mind, our team has elucidated the biosynthetic pathways of mannosylglycerate and characterized the respective genes and enzymes (1, 9, 17). Moreover, we recently reported the routes for the synthesis of DIP, diglycerol-phosphate, and glycerol-phospho-inositol (GPI) in Archaeoglobus fulgidus, but the genetic and biochemical characterization was lacking (2). In this context, the main objective of the present work was to identify the genes responsible for the synthesis of DIP, a canonical solute of hyperthermophiles.

Given that the structural peculiarities of the myo-inositol molecule led to great confusion in the nomenclature of inositol derivatives (20), the three possible configurations of DIP are...
The structures of the three possible stereoisomers of DIP. Structure I represents the configuration of DIP as reported by van Leeuwen et al. (30) and designated \(\text{L,\text{-di-myo-inositol-1,1'-phosphate}}\). Compounds I and III are stereoisomers, while compound II is optically inactive. In these representations, the numbering of the inositol atoms is based on the \(\text{L}\) configuration according to the “relaxation of lowest-locant rule” as recommended by the Nomenclature Committee of the International Union of Biochemistry in cases where metabolic relationships should be evidenced (20). A short note on \text{myo-inositol numberlation} is given in the supplemental material.

represented in Fig. 1 for the sake of clarity and easy perception of the structural differences. Throughout this article, stereospecific numbering of the inositol carbon atoms will be used based on the numbering of free \text{myo-inositol}. The carbon atom bearing the axial hydroxyl group is designated \(2\); if the molecule is oriented so that the axial hydroxyl group is above the plane of the molecule, the numbering of carbon atoms goes clockwise (see Fig. S1 in the supplemental material).

The sequence of reactions leading to DIP synthesis has been established on the basis of the detection of the relevant enzymatic activities in cell extracts of \textit{A. fulgidus} and of the structural characterization of intermediates and final products by nuclear magnetic resonance (NMR) (2). \textit{L-\text{myo-inositol-1-P}} is activated to CDP-\text{myo-inositol} via the activity of CTP: \text{myo-inositol-1-P} cytidylyltransferase (IPCT); subsequent condensation of CDP-\text{inositol} and \text{L-\text{myo-inositol-1-P}} yields a phosphorylated form of DIP designated DIPP (\text{di-\text{myo-inositol-1,3'-phosphate-1'-phosphate}}). The extra phosphorylation position was unambiguously assigned. Finally, DIPP is dephosphorylated to yield DIP. Intriguingly, DIP synthesized via this reaction scheme was expected to exhibit the stereochemical configuration of structure II (Fig. 1), in disagreement with the configuration earlier established by van Leeuwen et al. (30) in \textit{P. woesei} (structure I in Fig. 1). This apparent inconsistency required definite resolution.

Even though DIP was discovered more than 15 years ago, the genes and enzymes for the synthesis of the solute remain elusive. In this work, a genomic approach was used to identify the genes involved in DIP biosynthesis. Genes encoding IPCT and DIPP synthase (DIPPS; equivalent to CDP-\text{inositol-1-\text{myo-inositol-1-phosphate transferase}}) from several (hyper)thermophiles were uncovered by functional expression in \textit{Escherichia coli}. Additionally, \textit{\text{L-\text{myo-inositol-1-P}} labeled with carbon 13 was synthesized from \([6-^{13}\text{C}]\)glucose in a coupled, two-step reaction catalyzed by \textit{Thermoproteus tenax} hexokinase and \textit{A. fulgidus \text{L-\text{myo-inositol-1-P}} synthase}. This \textit{\text{L-\text{myo-[1-\text{13}C]}inositol-1-P}} was used as a substrate for recombinant or native enzymes. As a result of this strategy, the configuration of the inositol moieties of DIP was firmly established by tracing the fate of the label through the biosynthetic pathway by NMR analysis of the reaction products.

**MATERIALS AND METHODS**

**Materials.** \(\text{L-\text{glycerol-3-P}}, \text{glucose-6-P}, \text{NAD}^{+}, \text{CTP, ATP, UTP, GTP, CDP-glycerol, glyceral, and myo-inositol were purchased from Sigma-Aldrich (St. Louis, MO)}. \([6-^{13}\text{C}]\)glycerol was obtained from Omiceron Biochemicals, Inc. (Indiana). CDP-\text{myo-inositol} was obtained by chemical synthesis (this work), as well as \textit{\text{L-\text{myo-inositol-1-P}} and \text{myo-inositol-1-P}} (2). CDP-\text{myo-inositol} and \textit{\text{L-[1-\text{13}C]}myo-inositol-1-P} were synthesized enzymatically (this work). DIP from \textit{P. woesei} was obtained from Bitop AG (Witten, Germany), and \(\text{3-(1\text{X-glycerol-})-myo-inositol-phosphate, or GPI, was isolated from \textit{A. fulgidus}}\) (13).

**Organisms and growth conditions.** \textit{Pyrococcus furiosus} strain 3635, \textit{A. fulgidus} strain 7324, \textit{R. xylanophilus} strain 99417, and \textit{T. maritima} strain 31092 were obtained from Deutsche Sammlung von Mikroorganismen und Zellkulturen, Braunschweig, Germany. \textit{Thermococcus kodakareniscus} strain KOD1 and \textit{Aquifex aeolicus} strain VES biomasses were kindly provided by T. Imanaka (Kyoto University, Japan) and H. Huber (University of Regensburg, Germany), respectively. \textit{P. furiosus}, \textit{A. fulgidus}, and \textit{R. xylanophilus} were cultivated as previously described by Martins and Santos (14), Borges et al. (2), and Empadinhas et al. (10), respectively. \textit{T. maritima} was grown in medium 343 as described in the Deutsche Sammlung von Mikroorganismen und Zellkulturen.

**Enzyme assays.** The activity of DIPPS was determined in cell extracts of \textit{E. coli} harboring ipct/dipps genes. The reaction mixtures, in a total volume of 0.5 ml containing 20 mM Tris-HCl (pH 7.6), 10 mM MgCl\(_2\), 5 mM of the putative substrates, and cell extract (around 12 mg of total protein), were incubated for 1 h at 80°C (or 45°C for the enzyme from \textit{R. xylanophilus}). The reaction products were analyzed by \(^{31}\text{P-NMR spectroscopy. For details, see Borges et al. (2).}

**Determination of DIP stereochimeristry.** The stereochemical configuration of the inositol moieties in intermediate metabolites and end products was determined by \(^{31}\text{P-NMR and }^{31}\text{C-NMR spectroscopy using \textit{\text{L-[1-\text{13}C]}myo-inositol-1-P}} as a substrate for the recombinant enzymes. The stereochemistry of DIP was also investigated in cell extracts of natural producers (\textit{A. fulgidus}, \textit{A. xylanophilus}, and \textit{T. maritima}). Except for the labeled substrate provided here, the assays were carried out as previously described (2).

**Partial purification of native \textit{A. fulgidus} DIPPS.** Cells of \textit{A. fulgidus} harvested during the late exponential phase of growth (optical density at 600 nm = 0.35) and suspended in Tris-HCl (20 mM; pH 7.6) containing 5 mM MgCl\(_2\), 5 mM of the putative substrates, and cell extract (around 12 mg of total protein), were incubated for 1 h at 80°C (or 45°C for the enzyme from \textit{R. xylanophilus}). The reaction products were analyzed by \(^{31}\text{P-NMR spectroscopy. For details, see Borges et al. (2).}

Cloning and expression of putative ipct/dipps genes. Chromosomal DNAs from \textit{A. fulgidus}, \textit{P. furiosus}, \textit{T. kodakareniscus}, \textit{R. xylanophilus}, and \textit{A. aeolicus} were isolated according to the method of Ramakrishnan and Adams (21). The ipct/dipps genes from \textit{P. furiosus} (PF1058), \textit{A. fulgidus} (AF0263), and \textit{A. aeolicus} (aq.1367) were amplified by PCR using Pfu DNA polymerase (Fermentas) and
cloned in pET23a plasmid following standard protocols (25). The homologous genes from *T. kodakaraensis* (TK2279) and *R. xylanophilus* (EFS23341) were cloned in pTRC99a and pET30a, respectively. Many sequencing errors were found in the gene Rxyl_1212 (http://www.jgi.doe.gov/), with around 10% of the residues incorrectly determined. Therefore, the correct sequence was determined using *Pfu* DNA polymerase and submitted to GenBank. *E. coli* BL21(DE3) cells bearing the constructs were grown at 37°C in LB medium with ampicillin (100 μg/ml) to an optical density at 600 nm of 0.5 and induced with 1 mM IPTG (isopropyl-β-D-thiogalactopyranoside) for 4 h. Kanamycin (30 μg/ml) was used for the selection of *E. coli* BL21(DE3) bearing the *R. xylanophilus* gene. The cells were harvested, suspended in Tris-HCl (20 mM; pH 7.6) containing 10 mM MgCl₂, and disrupted in a French press; cell debris was removed by centrifugation (30,000 × g; 4°C; 1 h). Total protein was estimated by the Bradford method (3).

**Characterization of recombinant enzymes.** The substrate specificity of inositol-1-P cytidylyltransferase was studied in assays using CTP and each of the following alcohol donors: myo-inositol, L- myo-inositol-1-P, D- myo-inositol-1-P, glycerol, and α,β-glycerol-3-P; additionally, CTP, ATP, GTP, and UTP were examined as putative nucleotidyl donors in combination with L- myo-inositol-1-P. The substrate specificity of DIPPS was investigated using myo-inositol, L- myo-inositol-1-P, D- myo-inositol-1-P, glycerol, and α,β-glycerol-3-P as alcohol acceptors and CDP-L- myo-inositol, CDP-D- myo-inositol, and CDP-glycerol as alcohol donors.

**Preparation of CDP-L- myo-inositol and [1-13C]myo-inositol-1-P.** CDP-L- myo-inositol was produced by using cell extracts of *E. coli* BL21(DE3) harboring the gene EFS23341 (*R. xylanophilus*). The cell extract was incubated at 45°C with CTP and L- myo-inositol-1-P. After 1 h, the reaction mixtures were treated with cyanogen bromide (30 min at 37°C) to dephosphorylate residual myo-inositol-1-P. The sample was lyophilized, and the residue was dissolved in 1-propanol–25% ammonia (1:1.5 [vol/vol]) and applied to a silica gel S column. The same solvent system was used for elution. CDP-L- myo-inositol eluted first and was followed by myo-inositol. The fractions containing CDP-L- myo-inositol were pooled after analysis by 1H- and 31P-NMR for purity assessment.

CDP-L- myo-inositol and [1-13C]myo-inositol-1-P were produced from D-[6-13C]glucose by coupling hexokinase from *T. maritima* and myo-inositol-1-P synthase (IPS) from *A. fulgidus*.

The gene was cloned into pET23a; PET11cHX, containing the hx gene, was kindly donated by Bettina Siebers (University Duisburg-Essen, Germany). *E. coli* BL21(DE3) cells bearing the constructs were grown and induced as described above. Hexokinase and IPS were partially purified by heat treatment: 30 min at 90°C and 15 min at 60°C, respectively. The synthesis of [1-13C]myo-inositol-1-P was performed in a reaction mixture containing 6 mM D-[6-13C]glucose, 6 mM ATP, 6 mM NAD⁺, 10 mM MgCl₂, 3 μg of hexokinase, and 3 μg IPS in 20 mM Tris-HCl (pH 7.6). After 1 h of incubation at 70°C, the labeled compound was purified by anion-exchange chromatography (2).

**NMR spectroscopy.** The identification of metabolites involved in the synthesis of DIP and GIP was accomplished by using [1H, 13C]-, and 1H-31P-NMR on a Bruker DRX500 (Bruker, Rheinerten, Germany). Proton, carbon, and phosphorus chemical shifts are relative to 3-(trimethylsilyl)propanesulfonylic acid (at 0 ppm) in methanol (at 49.3 ppm), or 85% H₃PO₄ (at 0 ppm), respectively. During this work, we found that the assignment of the 13C resonances of myo-inositol reported in the literature is incorrect (4). Therefore, the assignment of the myo-inositol resonances was performed by running heteronuclear multiple quantum 1D- and 1L-2,3,4,5,6-penta-oxy-inositol was produced by using cell extracts of *E. coli* BL21(DE3) harboring the gene EF523341 (http://www.jgi.doe.gov), with around 10% of the amino acid sequence information that would lead to the identification of the gene in the genome sequence of this organism. We failed to reach a satisfactory purification yield but noticed that DIPPS activity was not separated from the IPCT activity even after four chromatographic steps (Q-Sepharose [pH 7.6 and 8.1], SP-Sepharose, and phenyl-Sepharose). The observed co-purification of the two activities led to the hint that they could be present in a single polypeptide chain. Further evidence in support of this hypothesis came from the analysis of the *A. fulgidus* genome (http://www.sanger.ac.uk/cgi-bin/Pam). Three of them (AF1143, AF1744, AF2044, and AF2299) whose predicted products belong to the family of proteins containing a domain characteristic of CDP-alcohol phosphatidylylycerol synthases (http://www.sanger.ac.uk/cgi-bin/Pam) are annotated as coding for enzymes implicated in the synthesis of phospholipids, leaving two genes (AF2063 and AF2299) as good candidates to encode DIPPS. Gene AF2063, encoding a bifunctional protein with high homology to nucleotidyltransferases (N-terminal domain) and to CDP-alcohol phosphatidylylycerol synthases (C-terminal domain), appeared especially promising. Importantly, genes homologous to AF2063 were found in all the genomes of organisms known to accumulate DIP (*P. furiosus*, *Pyrococcus horikoshi*, *T. kodakaraensis*, *A. aeolicus*, *Aeropyrum pernix*, *T. maritima*, and *R. xylanophilus*). Furthermore, genes encoding myo-inositol-1-P synthase, the first enzyme in DIP synthesis, were found in the immediate flanking region in *T. kodakaraensis*, *R. xylanophilus*, *T. maritima*, and *A. pernix*, substantiating the suspected connection between AF2063 and DIP synthesis. Therefore, we proceeded with cloning and expression of this gene in *E. coli* with the objective of validating the predicted functions.

Cell extracts of *E. coli* BL21(DE3) harboring the gene AF2063 were examined for the presence of IPCT/DIPPS activities using CTP and L- myo-inositol-1-P as substrates. DIPPS and vestigial amounts of CDP-inositol were identified by 31P-NMR as products in the reaction mixture (Fig. 2a). We verified that these products were not formed in extracts of *E. coli* BL21(DE3) cells harboring the empty plasmid. These results showed that AF2063 encodes a bifunctional enzyme able to synthesize DIP from CTP and L- myo-inositol-1-P via CDP-inositol.

**Characterization of recombinant enzymes.** Research effort was directed to purifying the recombinant enzyme from *A. fulgidus*, but all purification strategies (heat treatment and/or liquid chromatography) applied to extracts of *E. coli* BL21(DE3) harboring the gene AF2063 led to loss of the DIPPS activity, i.e., only the cytidylyltransferase activity was detected after heat treatment or chromatography. The gene was also expressed in *E. coli* Rosetta2(DE), which carries a plasmid containing the DNA sequences for codons rarely used in *E. coli*, with equal or better results. Expression of homologs from *T. kodakaraensis*, *P. furiosus*, and *A. aeolicus* led to similar results. In view of the apparent instability of the second activity, the
subsequent characterization of the enzyme activities was performed in cell extracts of *E. coli* BL21(DE3).

The recombinant enzyme from *R. xylanophilus* showed distinctive behavior in that the DIPPS activity was never detected; only the cytidylyltransferase activity was observed. The enzyme synthesized CDP-1-myoinositol from CTP and 1-myoinositol-1-P with high yield, and this property was exploited to produce CDP-1-myoinositol required for the substrate specificity assays (see below). It was also verified that myo-inositol was not a substrate for this enzyme.

Hereafter, the bifunctional enzyme encoded by gene AF0263 (and homologs) will be designated bifunctional DIPPS and the corresponding fused gene *ipt/dipps*. The substrate specificity of the bifunctional DIPPS from *A. fulgidus* was investigated in regard to both cytidylyltransferase and DIPPS activities. The cytidylyltransferase was absolutely specific for 1-myoinositol as a substrate for this enzyme.

Hereafter, the bifunctional enzyme encoded by gene AF0263 (and homologs) will be designated bifunctional DIPPS and the corresponding fused gene *ipt/dipps*. The substrate specificity of the bifunctional DIPPS from *A. fulgidus* was investigated in regard to both cytidylyltransferase and DIPPS activities. The cytidylyltransferase was absolutely specific for 1-myoinositol as a substrate for this enzyme.

Hereafter, the bifunctional enzyme encoded by gene AF0263 (and homologs) will be designated bifunctional DIPPS and the corresponding fused gene *ipt/dipps*. The substrate specificity of the bifunctional DIPPS from *A. fulgidus* was investigated in regard to both cytidylyltransferase and DIPPS activities. The cytidylyltransferase was absolutely specific for 1-myoinositol as a substrate for this enzyme.

Hereafter, the bifunctional enzyme encoded by gene AF0263 (and homologs) will be designated bifunctional DIPPS and the corresponding fused gene *ipt/dipps*. The substrate specificity of the bifunctional DIPPS from *A. fulgidus* was investigated in regard to both cytidylyltransferase and DIPPS activities. The cytidylyltransferase was absolutely specific for 1-myoinositol as a substrate for this enzyme.

Hereafter, the bifunctional enzyme encoded by gene AF0263 (and homologs) will be designated bifunctional DIPPS and the corresponding fused gene *ipt/dipps*. The substrate specificity of the bifunctional DIPPS from *A. fulgidus* was investigated in regard to both cytidylyltransferase and DIPPS activities. The cytidylyltransferase was absolutely specific for 1-myoinositol as a substrate for this enzyme.

Hereafter, the bifunctional enzyme encoded by gene AF0263 (and homologs) will be designated bifunctional DIPPS and the corresponding fused gene *ipt/dipps*. The substrate specificity of the bifunctional DIPPS from *A. fulgidus* was investigated in regard to both cytidylyltransferase and DIPPS activities. The cytidylyltransferase was absolutely specific for 1-myoinositol as a substrate for this enzyme.

Hereafter, the bifunctional enzyme encoded by gene AF0263 (and homologs) will be designated bifunctional DIPPS and the corresponding fused gene *ipt/dipps*. The substrate specificity of the bifunctional DIPPS from *A. fulgidus* was investigated in regard to both cytidylyltransferase and DIPPS activities. The cytidylyltransferase was absolutely specific for 1-myoinositol as a substrate for this enzyme.

Hereafter, the bifunctional enzyme encoded by gene AF0263 (and homologs) will be designated bifunctional DIPPS and the corresponding fused gene *ipt/dipps*. The substrate specificity of the bifunctional DIPPS from *A. fulgidus* was investigated in regard to both cytidylyltransferase and DIPPS activities. The cytidylyltransferase was absolutely specific for 1-myoinositol as a substrate for this enzyme.

Hereafter, the bifunctional enzyme encoded by gene AF0263 (and homologs) will be designated bifunctional DIPPS and the corresponding fused gene *ipt/dipps*. The substrate specificity of the bifunctional DIPPS from *A. fulgidus* was investigated in regard to both cytidylyltransferase and DIPPS activities. The cytidylyltransferase was absolutely specific for 1-myoinositol as a substrate for this enzyme.

Hereafter, the bifunctional enzyme encoded by gene AF0263 (and homologs) will be designated bifunctional DIPPS and the corresponding fused gene *ipt/dipps*. The substrate specificity of the bifunctional DIPPS from *A. fulgidus* was investigated in regard to both cytidylyltransferase and DIPPS activities. The cytidylyltransferase was absolutely specific for 1-myoinositol as a substrate for this enzyme.
A. aeolicus. Therefore, in all cases examined, the correct structure of DIP as established in this work is represented by structure II in Fig. 1.

**DISCUSSION**

Recently, we unequivocally established the biosynthetic pathway of DIP in *A. fulgidus* and showed that CDP-inositol and the phosphorylated form of DIP are intermediate metabolites in DIP synthesis. Furthermore, the presence of IPCT, DIPPS, and DIPP phosphatase activities was demonstrated in extracts of *A. fulgidus* and *P. furiosus* (2). As a logical extension of this work, we set about characterizing the genes and enzymes ascribed to this biosynthetic pathway. A genomic approach combined with hints originating from our attempts to purify the relevant activities led to the proposal of gene AF0263 as a strong candidate to encode the bifunctional IPCT/DIPPS. The final proof came from the functional expression of the gene in *E. coli*: the recombinant protein in cell extracts catalyzed the synthesis of DIPP from L-myo-inositol-1-P and CTP via CDP-inositol. Moreover, cloning and expression of the homologous genes from *P. furiosus*, *A. aeolicus*, and *T. kodakaraensis* revealed identical functions.

A BLAST search in public databases with the bifunctional IPCT/DIPPS from *A. fulgidus* resulted in the identification of homologous proteins in several cultured organisms, all of them known to thrive in hot environments (see Fig. S2 in the supplemental material). The two activities, cytidylyltransferase and DIPPS, are encoded in a single polypeptide chain in *Pyrococcus* spp., *T. kodakaraensis*, *R. xylanophilus*, *P. marina*, and *A. aeolicus*, while separate genes were found in *Hyperthermus butylicus*, *A. pernix*, and *T. maritima*. Shortly before the manuscript was submitted for publication, the demonstration of this function for the genes of *T. maritima* appeared in the literature (23).

The N-terminal domain of the bifunctional IPCT/DIPPS showed high similarity to the nucleotidyltransferase family (pfam00483), comprising enzymes that catalyze the transfer of nucleotides onto phosphosugars, like UTP:glucose-1-P uridylyltransferase, CTP:glucose-1-P cytidylyltransferase, or GTP:mannose-1-P guanylyltransferase. The C-terminal domain of IPCT/DIPPS contains the conserved motif (D-G-2[X]-A-R-8[X]-

---

FIG. 3. 13C-NMR and 31P-NMR spectra showing the formation of 13C-labeled products in a cell extract of *A. fulgidus* after incubation at 80°C. Traces a and c represent spectra of the reaction mixture to which CTP and L-[1-13C]myo-inositol-1-phosphate were added; traces b and d represent spectra of the same sample after treatment with alkaline phosphatase. A schematic representation of DIPP is also shown. Spectra were run at a temperature of 25°C. inos-C1, L-inos-1P-C1, DIP-C1, DIPP-C1, and CDP-inos-C1 designate resonances due to carbon 1 of the inositol moiety in myo-inositol, L-myo-inositol-1-P, DIP, DIPP, and CDP-L-myo-inositol. The splitting of the carbon (or phosphorus) resonances denotes two-bond couplings with phosphorus (or carbon) nuclei.
G-3[X]-D-3[X]-D) that is characteristic of the CDP-alcohol phosphatidyltransferase family (pfam01066). These enzymes catalyze the displacement of CMP from CDP-alcohol (diacylglycerol is the most common alcohol) to a second alcohol with formation of a phosphodiester bond and concomitant hydrolysis of pyrophosphate. The phosphatidylinositol synthase from yeast is a well-characterized representative of this family, having 23% identity with the amino acid sequence of *A. fulgidus* DIPPS (19). Hydropathic profiles predict the presence of three transmembrane segments in the yeast enzyme (19), and similar

![Diagram](image1)

**FIG. 4.** Pathway for DIP biosynthesis using L-[1-13C]myo-inositol-1-phosphate as a labeled precursor. The asterisks highlight the positions of labeled carbon atoms in the several intermediate metabolites and end products as established in this work from 13C- and 31P-NMR analyses of reaction mixtures containing the bifunctional enzyme IPCT/DIPPS (see the text for details). Dashes are used to designate the carbon atoms in the phosphorylated inositol moiety.

![Diagram](image2)

**FIG. 5.** 13C-NMR spectra showing the formation of 13C-labeled products resulting from incubation at 80°C of an extract of *E. coli*, bearing the ipct/dipps gene from *P. furiosus*, with CTP and L-[1-13C]myo-inositol-1-phosphate, before (a) and after (b) treatment of the final reaction mixture with alkaline phosphatase. The inset shows a schematic representation of di-myoinositol-1,3′-phosphate-1′-phosphate. The spectra were run at a temperature of 25°C. Symbols are as in Fig. 2.
features are found in the DIPPS domains identified in our study: the stretches comprising residues 272 to 294, 334 to 356, and 393 to 435 have sufficient length and hydrophobicity to span the membrane. The predicted membrane association of the DIPPS domain in the bifunctional IPCT/DIPPS provides a plausible explanation for our failure to purify the enzyme.

The reaction scheme established for DIP synthesis points to a configuration of the inositol units as displayed in structure II (Fig. 1), in disagreement with the stereochemical configuration earlier determined in DIP isolated from Pyrococcus spp. (30). This intriguing inconsistency was definitely resolved in the present work. By using labeled 1·[1-3C]-myo-inositol-1-P as a substrate for the bifunctional, recombinant enzyme or for the enzymes in cell extracts of the natural producers, the 3C label was traced through the biosynthetic pathway, and we showed that the two inositol moieties in DIP have the configuration represented by structure II (Fig. 1) in all the organisms examined (A. fulgidus, P. furiosus, A. aeolicus, T. kodakaraensis, and T. maritima). We propose that henceforth the designation di-myoinositol-1,3′-phosphate should be used to highlight the fact that the phosphate group is linked to stereochemically distinct carbons of the two inositol groups and to facilitate spotting individual carbon atoms through metabolic pathways.

An unrooted phylogenetic tree based on the amino acid sequences of the bifunctional IPCT/DIPPS showed significant topological differences compared with the 16S rRNA tree (31). Although the limited number of sequences available hampers a reliable comparison, it is interesting that only two clusters emerged in the protein-based tree (Fig. 6). In the predominantly euryarchaeote cluster, the Thermococcales form the expected tight cluster and the position of A. fulgidus suggests an early separation from the common ancestor of the genera Thermococcus and Pyrococcus. Curiously, the bacterial species A. aeolicus and P. marina also fall in this “euryarchaeote” cluster. The second cluster comprises the crenarchaeote members, A. pernix and H. butylicus, but also the bacteria T. maritima and R. xylanophilus. It is tempting to speculate that bacterial IPCT/DIPPS either have evolved from distinct ancestors or the respective genes were acquired from archaea by lateral gene transfer. However, analysis of the GC contents and the codon compositions of bacterial ipct/dipps genes did not provide evidence for lateral gene transfer, since their GC contents and codon usages were in line with the properties found in the bulk genes of the respective genomes (data not shown). It is noteworthy that the lack of unusual codon composition or GC content does not rule out the occurrence of lateral gene transfer, since it is conceivable that genes acquired by this means will adapt to the background genomic codon usage during evolution.

The cytidylyltransferase domain (C-terminal) of recombinant IPCT/DIPPS was highly specific for CTP and 1·myo-inositol-1-P, while the DIPPS domain showed some plasticity in regard to the alcohol-activated donor. All the DIPPS examined (i.e., A. fulgidus, P. furiosus, A. aeolicus, and T. kodakaraensis) used both CDP-inositol and CDP-glycerol in combination with 1·myo-inositol-1-P. Therefore, either DIPP or GPIP can be end products of this activity, depending on the specific CDP-alcohol provided (CDP-inositol or CDP-glycerol, respectively). Thus far, the compatible solute GPI has been found only in species of the genera Archaeoglobus and Aquifex (13). Thus, the observation that all DIPPS examined in this work are also able to catalyze GPI synthesis makes us wonder why GPI does not occur in members of other hyperthermophilic genera, namely, Pyrococcus or Thermococcus. A plausible explanation could be related to insufficient levels of CDP-glycerol, the substrate, besides 1·myo-inositol-1-P, that is needed for the synthesis of GPI (2). In A. aeolicus, one of the two organisms known to accumulate GPI, the genes encoding glycerol-3-P cytidylyltransferase and DIPPS are organized in the same operon-like structure and probably under the control of the same promoter. In the other GPI-accumulating organism, A. fulgidus, this close spatial organization is not found, with the gene encoding glycerol-3-P cytidylyltranserase and DIPPS organized in the same operon-like structure and probably under the control of the same promoter, and the gene encoding glycerol-3-P cytidylyltranserase is located elsewhere. Nevertheless, CDP-glycerol is expected to be highly available in the cytoplasm of this organism, since CDP-glycerol is also the precursor for the synthesis of diglycerol phosphate, a prominent component of the solute pool of A. fulgidus (11).

In conclusion, this work characterized the key genes and enzymes for the synthesis of DIP, the most widespread compatible solute in hyperthermophiles. Furthermore, it was shown that the same genes are used in the synthesis of GPI, a
solute chemically related to DIP and also confined to (hyper)thermophiles. The gene encoding the phosphatase activity that dephosphorylates the phosphorylated intermediates DIPP and GPiP was not characterized; however, the sequence information available for the uncultured organism GZfios13E1 revealed a gene encoding three remarkably expressive activities: cytidylyltransferase, DIPPS, and phosphatase. The hypothesis that the last domain bears the activity that dephosphorylates DIPP appears highly plausible and deserves further investigation. The results disclosed in the present work represent an important milestone in the elucidation of the regulation of DIP biosynthesis and its role in the thermoadaptation of hyperthermophiles.

ACKNOWLEDGMENTS


We thank Vera Lopes and Carla P. Almeida for technical support.

REFERENCES


