THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 290, NO. 35, pp. 21690–21704, August 28, 2015

Structure and Evolution of the Archaeal Lipid Synthesis Enzyme *sn*-Glycerol-1-phosphate Dehydrogenase*

Received for publication, March 3, 2015, and in revised form, July 6, 2015 Published, JBC Papers in Press, July 14, 2015, DOI 10.1074/jbc.M115.647461

Vincenzo Carbone[‡], Linley R. Schofield[‡], Yanli Zhang[‡], Carrie Sang[‡], Debjit Dey[‡], Ingegerd M. Hannus[‡], William F. Martin[§], Andrew J. Sutherland-Smith[¶], and Ron S. Ronimus[‡1]

From [‡]AgResearch Limited, Grasslands Research Centre, Tennent Drive, Private Bag 11008, Palmerston North 4442, New Zealand, [§]Institute for Molecular Evolution, Heinrich Heine University, University of Düsseldorf, 40225 Düsseldorf, Germany, and [¶]Institute of Fundamental Sciences, Massey University, Palmerston North 4442, New Zealand

Background: Archaea synthesize glycerol-based membrane lipids of unique stereochemistry, utilizing distinct enzymology.
Results: The structure of *sn*-glycerol-1-phosphate dehydrogenase (G1PDH), the first step in archaeal lipid synthesis, was determined.
Conclusion: G1PDH is a member of the iron-dependent alcohol dehydrogenase and dehydroquinate synthase superfamily.
Significance: The data contribute to our understanding of the origins of cellular lipids at the divergence of the Archaea and Bacteria.

One of the most critical events in the origins of cellular life was the development of lipid membranes. Archaea use isoprenoid chains linked via ether bonds to *sn*-glycerol 1-phosphate (G1P), whereas bacteria and eukaryotes use fatty acids attached via ester bonds to enantiomeric *sn*-glycerol 3-phosphate. NAD(P)H-dependent G1P dehydrogenase (G1PDH) forms G1P and has been proposed to have played a crucial role in the speciation of the Archaea. We present here, to our knowledge, the first structures of archaeal G1PDH from the hyperthermophilic methanogen *Methanocaldococcus jannaschii* with bound substrate dihydroxyacetone phosphate, product G1P, NADPH, and Zn^{2+} cofactor. We also biochemically characterized the enzyme with respect to pH optimum, cation specificity, and kinetic parameters for dihydroxyacetone phosphate and NAD(P)H. The structures provide key evidence for the reaction mechanism in the stereospecific addition for the NAD(P)H-based pro-*R* hydrogen transfer and the coordination of the Zn^{2+} cofactor during catalysis. Structure-based phylogenetic analyses also provide insight into the origins of G1PDH.

All life comprises cells with lipid membranes separating the cell from the external environment and permitting chemiosmotic energy harnessing (1). Understanding how membranes evolved is crucial to understanding how energy harnessing and life itself came into existence. Stable and readily synthesized cell lipids could have conferred on their hosts a number of clear and crucial advantages including the ability to accumulate metabolites and intracellular enzymes to high concentrations (2), an increased ability to use concentration gradients for production of ATP (3), a reduced rate of lateral gene transfer (4), and protection against viruses (5). Efficient replication of cells would have required rapid synthesis of membranes and would have aided the development of more stable early archaeal and bacterial lineages.

The rotor-stator ATPase that harnesses ion gradients for chemiosmotic coupling is universal among all prokaryotes (1, 6, 7), but the membranes that maintain those ion gradients are not. Bacterial membranes consist of fatty acid esters with glycerol 3-phosphate, the stereochemistry of which contrasts with that of archaeal membranes based on isoprenoid ethers linked to *sn*-glycerol 1-phosphate (G1P).[2] The domain-specific differences in lipid synthesis between Archaea and Bacteria reflect their very ancient divergence, and despite numerous lateral gene transfers between Archaea and Bacteria, lipid chemistry has remained a stable, vertically inherited trait within each prokaryotic domain (7, 8). The evolutionary process that gave rise to this prokaryotic membrane dichotomy, the so-called "lipid divide," is still poorly understood (9). Two main alternatives are currently debated (10). In the first, the last universal common ancestor (LUCA) possessed genes for biosynthesis of both G1P isoprenoid and glycerol 3-phosphate fatty acyl lipid types, and the extant bacterial-archaeal differences are due to differential loss (10–12). In the second, the LUCA did not have genetically encoded lipid biosynthesis, although it might have had geochemically synthesized lipids, and the evolutionary invention of distinct bacterial and archaeal lipids occurred independently in the stem lineages that gave rise to the Achaea and Bacteria (3, 7, 13). The first alternative predicts that both lipid synthetic pathways were early evolutionary inventions, whereas the second proposes that life arose within geochemically formed inorganic compartments

* This work was supported by the Pastoral Greenhouse Gas Research Consortium, the Royal Society of New Zealand, the New Zealand Synchrotron Group, and the Australian Synchrotron Foundation investor access program. The authors declare no conflicts of interest.
The atomic coordinates and structure factors (codes 4RGV, 4RFL, and 4RGQ) have been deposited in the Protein Data Bank (http://wwpdb.org/).
[1] To whom correspondence should be addressed: Animal Nutrition and Health, AgResearch Ltd., Grasslands Research Centre, Tennent Dr., Private Bag 11008, Palmerston North 4442, New Zealand. Tel.: 64-6-351-8036; Fax: 64-6-351-8032; E-mail: ron.ronimus@agresearch.co.nz.

[2] The abbreviations used are: G1P, *sn*-glycerol 1-phosphate; ADH, alcohol dehydrogenase; DHAP, dihydroxyacetone phosphate; DHQS, dehydroquinate synthase; G1PDH, G1P dehydrogenase; GDH, glycerol dehydrogenase; LUCA, last universal common ancestor; MJ, *M. jannaschii*; TCEP, tris(2-carboxyethyl)phosphine; Bistris propane, 1,3-bis[tris(hydroxymethyl)methylamino]propane.

and predicts that lipid synthesis arose later during primordial evolution but before the emergence of free living cells.

One key to resolving how the lipid divide arose is understanding the evolution of the non-homologous archaeal G1P dehydrogenase (G1PDH) and bacterial glycerol 3-phosphate dehydrogenase that convert non-enantiomeric dihydroxyacetone phosphate to the G1P or glycerol 3-phosphate stereoisomers of glycerol phosphate, respectively, that form the backbone of the domain-distinct lipid types. Describing the structure and evolution of archaeal G1PDH is thus important for understanding how the transition from the LUCA led to the archaeal domain. Here we present the first x-ray crystal structures for archaeal G1PDH to our knowledge. The *Methanocaldococcus jannaschii* (MJ) G1PDH structure provides novel insight on binding of its DHAP substrate, enantiomeric product G1P, NADP(H) coenzyme, and Zn^{2+} cofactor coordination state, revealing details of the stereospecific pro-*R* hydrogen reaction. We also biochemically characterized the MJ G1PDH with respect to kinetics with its substrates (DHAP and NAD(P)H) and to our knowledge are the first to have critically examined the effects of metal ions on G1PDH activity. In addition, comparative structural and structure-based phylogenetic analyses contribute further to our understanding of the origin of the archaeal G1PDH and therefore that of the archaeal domain of life.

Experimental Procedures

General Methods—Electrophoresis was performed with 12% Mini-PROTEAN® TGX™ precast gels (Bio-Rad) using low range SDS-PAGE molecular weight standards (Bio-Rad) and Coomassie Brilliant Blue R-250 staining. Protein concentrations were determined by the method of Bradford (14) using bovine serum albumin as a standard. The pH of buffers was adjusted at room temperature. All pH values are reported as at the temperature of use and allow for ΔpK_a/°C. When required, metal ions were removed from solutions by treatment with Chelex 100 chelating resin (Bio-Rad). DHAP, NADPH, and NADH were purchased from Sigma-Aldrich. All other materials were at least analytical grade quality.

Cloning, Protein Expression, and Purification—MJ was obtained from the German Collection of Microorganisms and Cell Cultures GmbH (DSMZ) (Germany; DSM 2661), grown in RM02 medium (15), and DNA-extracted using Chelex InstaGene matrix (Bio-Rad) according to the manufacturer's suggested protocol. PCR primers were used to amplify the gene for cloning into the expression vectors pET151D TOPO and pET100D TOPO (Invitrogen), which add an N-terminal 6-residue histidine affinity purification tag and a peptidase cleavage site (recombinant tobacco etch virus and enterokinase, respectively; Life Technologies). The forward primer was 5′-caccatgattatagtcacaccaagatatac-3′, and the reverse primer was 5′-ttaaataactcctgtttcttcagcc-3′. The PCR utilized Hercules II high fidelity DNA polymerase (2.5 units; Stratagene) in a 35-cycle reaction: 95 °C for 2 min followed by 35 cycles of 94 °C for 15 s, 56 °C for 15 s, and 68 °C for 1 min and 15 s. The dNTP concentration was 300 μM, and the primer concentrations were each 0.2 μM. The final extension was at 68 °C for 2 min. The specificity of the PCRs was checked by agarose gel electrophoresis, and the products were gel-purified using a Wizard gel DNA extraction kit (Promega). The concentrations of the purified PCR products were quantified using a NanoDrop spectrophotometer (NanoDrop) and TOPO-cloned into expression vectors using the manufacturer's suggested method. Single colonies were checked for inserts in the correct orientation and length using colony PCR incorporating the expression vector T7 forward primer (5′-taataatacgactcactataggg-3′) and the above reverse primer for the MJ G1PDH gene. Single replicate colonies were used to generate plasmid DNA for sequence verification and subsequent transformation of the *Escherichia coli* expression strain Rosetta 2 (DE3) (Novagen/Life Technologies).

Single colonies of *E. coli* Rosetta 2 (DE3) cells containing either the MJ G1PDH pET151D-based plasmid or the MJ G1PDH pET100D-based plasmid were precultured for ~16 h in 10 ml of Luria-Bertani (LB) medium containing 100 μg ml^{-1} ampicillin and 34 μg ml^{-1} chloramphenicol at 37 °C. 10 ml of preculture was used to inoculate 700 ml of 2× yeast extract-tryptone medium containing 100 μg ml^{-1} ampicillin and 34 μg ml^{-1} chloramphenicol in a 2-liter baffled flask. Induction was initiated with isopropyl β-D-1-thiogalactopyranoside at 1 mM when the cell culture reached a density of 0.45 at 600 nm. Cells were grown for ~8 h with vigorous shaking at 28 °C, cooled, and stored overnight at 4 °C before harvesting (16,000 × *g* for 15 min at 4 °C), freezing, and storage at −20 °C. Cell pellets were thawed and resuspended in 4–5 volumes of lysis buffer (50 mM Tris, pH 7.4 (pET151D) or pH 8.2 (pET100D) containing 2 mM β-mercaptoethanol, 300 mM NaCl, 10 mM imidazole, 1% (v/v) Triton X-100, and 10 mM $MgCl_2$). Complete EDTA-free protease inhibitor (Roche Applied Science) was added as a stock solution following the manufacturer's instructions. Lysozyme (Sigma-Aldrich, L8676) was added to a final concentration of 1 mg ml^{-1} followed by gentle agitation for 30–60 min on ice. DNase (Sigma-Aldrich, D5025) and RNase (Sigma-Aldrich, R4642) were added to a final concentration of 5 μg ml^{-1} each followed by gentle agitation for 30–60 min on ice. Cell debris was removed by centrifugation (16,000 × *g* for 15 min at 4 °C). The hexahistidine-tagged enzyme was purified from the cell-free extract using nickel affinity chromatography. The filtered enzyme was applied (1 ml min^{-1}) to a 6% CL-Nickel ChroMatrix™ resin (Jena Bioscience, Germany) column (2.5 × 18 cm) equilibrated with 20 mM Tris, pH 7.4 (pET151D) or pH 8.2 (pET100D) containing 2 mM β-mercaptoethanol, 300 mM NaCl, and 20 mM imidazole. The column was washed with the equilibration buffer before fractions were eluted with a linear gradient of 20–250 mM imidazole (3 ml min^{-1}). Fractions were examined by SDS-PAGE, and those containing protein of the expected molecular mass were pooled and concentrated using a stirred ultrafiltration cell (Amicon) with a 10-kDa-nominal molecular mass limit membrane. The imidazole was removed and buffer-exchanged using a Bio-Gel® P-6DG (Bio-Rad) column (2.5 × 17 cm) equilibrated with 20 mM MOPS, pH 7.0 (pET151D) or 10 mM Tris, pH 8.1 (pET100D) containing 2 mM tris(2-carboxyethyl)phosphine (TCEP) and 150 mM KCl (pET151D) or 125 mM KCl (pET100D) (1 ml min^{-1}). Fractions containing protein were collected and concentrated either using a stirred ultrafiltration cell (Amicon) with a 10-kDa-nominal molecular mass limit membrane (pET151D) or a 20-ml Vivaspin 10-kDa-molecular mass cutoff sample concentrator

(GE Healthcare) (pET100D). All chromatographic steps were performed at room temperature using a BioLogic LP system (Bio-Rad) with detection at 280 nm.

EDTA Treatment of G1PDH—To remove metal ions for enzyme assays, G1PDH was treated with EDTA as follows. A 2-ml Vivaspin 10-kDa-molecular mass cutoff sample concentrator was prewashed two times with buffer A (20 mM MOPS, pH 7.0, 2 mM TCEP, 150 mM KCl (or 150 mM NaCl when determining the effect of NaCl on activity)). 50 μl of G1PDH (prepared using pET151D) was added to the prewashed concentrator, diluted to 200 μl with buffer A also containing 2 mM EDTA, mixed, and incubated for 20 min at room temperature. The volume was concentrated to 50 μl by centrifugation (10,000 $\times g$ for 2 min at 20 °C), and the dilution-concentration process was repeated three times before reconstitution in buffer A (EDTA-free, Chelex-treated) and incubation for 5 min at room temperature.

G1PDH Activity Assays—Spectrophotometric measurements and calculation of initial velocity were performed using a Cary 100 UV-visible spectrophotometer (Agilent Technologies) with a thermostated cuvette holder using 1- or 0.5-cm-path length stoppered quartz cuvettes. The consumption of NADPH or NADH in the reaction of G1PDH was monitored at 366 nm. The extinction coefficient of NADPH was determined to be 2,530 $\text{M}^{-1}\ \text{cm}^{-1}$ at 366 nm and 65 °C in 50 mM Bistris propane buffer at pH 7.8. Initial rates of reaction were measured within 2 min and were determined by a least square fit of the initial rate data. K_m and k_{cat} values were determined by fitting the data to the Michaelis-Menten equation using GraFit (16). Activity was measured at 65 °C because of the decomposition of DHAP and NAD(P)H above 70 °C (17). One unit of activity is defined as the conversion of 1 μmole of NAD(P)H to NAD(P)^+/min at the stated temperature. All assays were performed in triplicate. Metal ions were removed from all buffer and reagent solutions by treatment with Chelex 100 chelating ion-exchange resin.

The standard reaction mixture contained (unless otherwise stated) 50 mM Bistris propane buffer at pH 7.8, 100 mM KCl, 0.1 mM $ZnCl_2$, 3 mM DHAP, and 0.15 mM NAD(P)H and was preincubated at 65 °C for 7–10 min. The reaction was initiated by addition of 0.276 μM EDTA-treated G1PDH (pET151D). The total volume of each assay was 200 μl. Assays to determine the effect of pH on activity contained 50 mM Bistris propane, pH 5.38–8.78 or citrate, pH 4.0–6.5. All divalent metal salts used in assays to restore activity to the EDTA-treated G1PDH were dissolved in water that had been pretreated with Chelex. The final concentration of divalent metal salt was 0.1 mM in the assay reaction mixture. Activity was also measured in the absence of divalent metal ions. Assays for the determination of DHAP kinetic parameters contained 0.05–20.0 mM DHAP and 0.15 mM NAD(P)H. Assays for the determination of NAD(P)H kinetic parameters contained 0.035–0.3 mM NADPH or 0.1–0.45 mM NADH and 10 mM DHAP. Additionally, assays for the determination of NADH kinetic parameters contained double the concentration of G1PDH (0.552 μM).

Molecular Mass Determination—The native molecular mass was determined by gel filtration chromatography using a BioLogic DuoFlow QuadTec 10 chromatography system (Bio-Rad). A filtered sample of G1PDH (400 μl) at a concentration of 0.9–3.7 mg ml^{-1} was applied to a Superdex 200 (GE Healthcare) column (1 $\times$ 59 cm) using a 1-ml sample loop. The column was eluted with 50 mM MOPS, pH 7.0 containing 2 mM TCEP and 0.15 M KCl at a flow rate of 0.6 ml min^{-1}. A standard curve was generated using commercial standards (Sigma-Aldrich).

Sequence and Phylogenetic Analyses—Unrooted phylogenetic analyses were generated using MEGA version 6 (18, 19). The structure-based phylogenetic tree for G1PDHs, glycerol dehydrogenases (GDHs), alcohol dehydrogenases (ADHs), and dehydroquinate synthases (DHQSs) was constructed using the structure-based alignment program PromalS3D and then implemented in MEGA6 (20). The PromalS3D alignment parameters were as follows: identity threshold above which fast alignment is applied, 0.6; weight for constraints derived from sequences, 1; weight for constraints derived from homologues with structures, 1; weight for constraints derived from input structures, 1.

Crystallization—A number of initial crystallization conditions for G1PDH were identified using the Morpheus screen from Molecular Dimensions (UK). The lead precipitant condition (30% (v/v) poly(ethylene glycol) methyl ether 550/polyethylene glycol 20,000 and 0.1 M Tris, pH 8.5) was further optimized using the Silver Bullets Bio screen from Hampton Research. The apo (MJ G1PDH (pET151D); 10.7 mg ml^{-1} in 20 mM MOPS, pH 7.0, 2.0 mM TCEP, and 150 mM KCl; Silver Bullets Bio condition G1) and binary (MJ G1PDH (pET100D); 14.5 mg ml^{-1} in 10 mM Tris, pH 8.1, 2.0 mM TCEP, 125 mM KCl, and 2.79 mM NADPH; Silver Bullets Bio condition H5) forms of G1PDH were grown at 21 °C using the sitting drop method. The ternary complex (MJ G1PDH (pET151D); 10.5 mg ml^{-1} in 20 mM MOPS, pH 7.0, 2.0 mM TCEP, 150 mM KCl, and 2.05 mM NADPH) was achieved by growing crystals in precipitant containing 30% (v/v) poly(ethylene glycol) methyl ether 550/polyethylene glycol 20,000 and 0.1 M HEPES, pH 7.6 and then looping crystals for 1–2 s through a solution of mother liquor containing cryoprotectant and 10 mM DHAP.

Data Collection and Structure Determination—All diffraction data sets were collected at the Australian Synchrotron MX1 and MX2 beamlines using Blu-Ice (21) and processed with XDS (22) and SCALA (23) from flash cooled crystals (100 K) in mother liquor containing 25% (v/v) ethylene glycol as cryoprotectant for the binary and ternary complexes and perfluoropolyether oil for the apo structure. Exposure time, oscillation range, crystal-detector distance, and beam attenuation were adjusted to optimize the collection of data to the maximum resolution possible ranging from 2.45 to 2.20 Å. Initial phases for G1PDH were determined in CCP4 (24) by the molecular replacement program Phaser (25) using a polyalanine model prepared by CHAINSAW (26) of the crystal structure of GDH from *Clostridium acetobutylicum* (Protein Data Bank code 3CE9). Structural idealization and restrained refinement were carried out using REFMAC5 (27) with local non-crystallographic symmetry restraints used throughout the refinement process. Translation/libration/screw and restrained refinement were used in the final cycles of refinement for the apo structure (28). Difference Fourier maps ($2F_o - F_c$ and $F_o - F_c$)

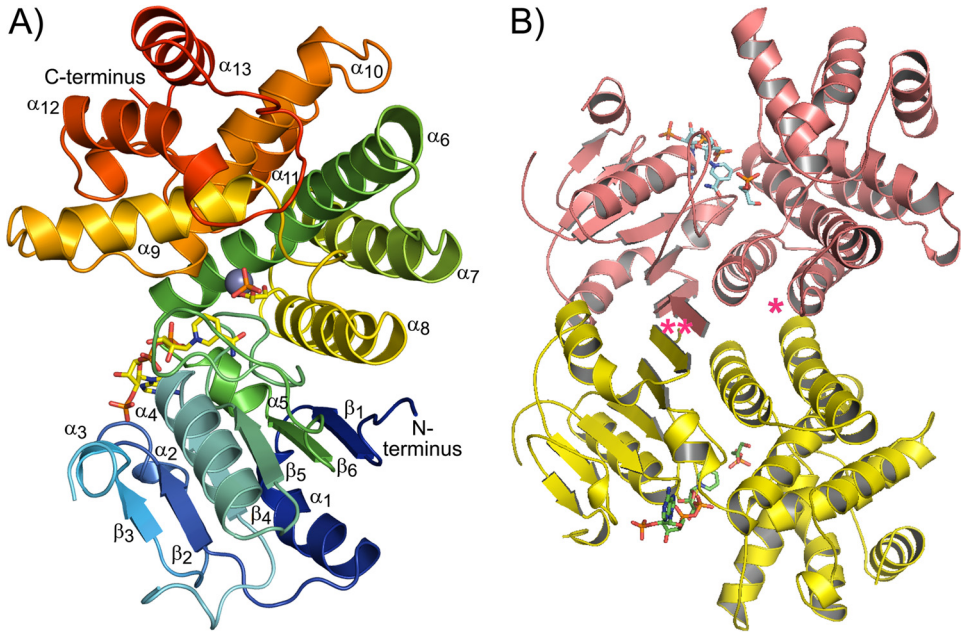

FIGURE 1. **Ribbon representations of MJ G1PDH.** *A*, the monomer with coenzyme and substrate molecules in *yellow* and Zn^{2+} in *gray*. *B*, the observed crystallographic dimer of MJ G1PDH. The interactions facilitated by the N-terminal domain and the hydrophobic interactions maintained at the interface of α7 and α8 are indicated with ** and *, respectively.

were visualized in Coot (29) and enabled the addition of amino acid side chains, substrates, and solvent molecules, revealing clear density for the G1PDH structure. Solvent content was estimated at 2.16, 2.78, and 2.47 $Å^3$ Da^{-1} (30) for the apo, binary, and ternary complexes, respectively. Further data collection and final refinement statistics are listed in Table 1. All figures were prepared using PyMOL and CCP4mg version 2.7.3 (31). The atomic coordinates and structure factors (codes 4RGV (apo), 4RFL (binary), and 4RGQ (ternary)) have been deposited in the Protein Data Bank.

Results

Structural Analysis of the MJ G1PDH—Three MJ G1PDH structures (apo, binary, and ternary) were determined to a maximum resolution of 2.20 Å (Fig. 1, *A* and *B* and Table 1). The ternary and binary complexes revealed four molecules in the asymmetric unit, two molecules in the apo structure, and well defined density for the metal ion (Zn^{2+}; Fig. 2, *A* and *B*), the substrate/product (DHAP/G1P; ternary complex only; Fig. 2, *A*, *B*, and *C*), and coenzyme (NADPH; Fig. 2*D*). The only regions of poorly defined side chain density are solvent-exposed and not associated with the active site (Glu^{54}–Lys^{64} and Arg^{115}-Gln^{116}).

G1PDH possesses two distinct structural domains separated by a deep binding cleft occupied by NADPH, substrate (or product), a single K^+ (binary and ternary complexes only), and at the center Zn^{2+}. The structure of MJ G1PDH at the N-terminal domain (residues 1–137) forms an atypical Rossman fold consisting of a six-stranded parallel β-sheet surrounded by five α-helices arranged in a $\beta_1\alpha_1\beta_2\alpha_2\beta_3\alpha_3\beta_4\alpha_4\beta_5\beta_6\alpha_5$ pattern. Strands β_5 and β_6 are connected by a hairpin loop (residues 100–125) that partially bifurcates the center of the enzyme, whereas α_5 sits apart from the traditional Rossmann fold architecture and atop β-strands 1, 5, and 6. The end of α_5 also marks the beginning of the C-terminal and catalytic domain (residues 138–335) made up exclusively of eight α-helices and two small α-helical turns between α_{10} to α_{11} and α_{11} to α_{12}.

In the ternary complex, a single Zn^{2+} ion is bound in each G1PDH molecule forming bonds (with average distances) to His^{226} (Nϵ2; 2.06 Å), His^{247} (Nϵ2; 2.24 Å), and Asp^{148} (Oδ1; 2.17 Å) and in monomers C and D to a single water molecule (HOH; 1.99 and 2.03 Å, respectively). Three molecules of DHAP (monomers A, C, and D; Fig. 2*A*) and one molecule of G1P (monomer B; Fig. 2*B*) are observed in the active sites. The 2-carbonyl of DHAP also coordinates with Zn^{2+} via an ion-dipole interaction (distance, 2.01–2.09 Å) and to a lesser extent so does the 3-hydroxyl moiety (monomers A and C; 2.54–2.74 Å), whereas for the bound G1P product, the 2-hydroxyl is 2.72 Å from Zn^{2+}, and the two conformers of the 3-hydroxyl moiety of G1P are 2.08 and 2.83 Å away. An overall composite picture of DHAP binding is shown in Fig. 2*C* where hydrogen bonds are formed by the hydroxyl moiety of DHAP with the side chain carboxylate of Asp^{105}, the main chain carbonyl of Ala^{221}, and a K^+ ion buried in a deep cleft adjacent to the bound Zn^{2+}. DHAP phosphate binding is maintained by a series of hydrogen bond contacts with the side chain atoms of Ser^{118} (Oγ) and Gln^{116} (Nϵ2) (belonging to the N-terminal hairpin loop), Ser^{218} (Oγ), Ser^{222} (Oγ), and His^{230} (Nϵ2) (of α_9) and Arg^{310} (Nη1) (present on the loop following α_{12}). Temperature factors for the fully occupied Zn^{2+} were close to those of the coordinating atom of Asp^{148} and the carbonyl of DHAP in the ternary structure, whereas those of the histidine and coordinating water molecule were slightly lower. We also observed variable coordination numbers and changes in ideal geometry of Zn^{2+} (assuming a coordinate distance of <2.5 Å; Fig. 3).

The binary and ternary G1PDH structures also revealed well defined density for coenzyme (NADPH), although density was

TABLE 1
Data collection and refinement statistics
ref., reflections; RMSD, root mean square deviation.

	Apo G1PDH	*Binary Complex*	*Ternary Complex*
Space group	P 1 21 1	P 1	P 1
Unit Cell parameters			
a, b, c (Å)	49.76, 59.55, 119.8	59.39, 72.11, 101.7	59.38, 71.9, 101.7
α, β, γ (°)	90, 90.39, 90	77.62, 79.58, 75.63	77.52, 79.54, 75.6
Data Collection Statistics			
Wavelength (Å)	0.91840	0.91160	0.91160
Temperature (K)	100K	100K	100K
Resolution Range (Å)	46.069 - 2.455	46.68 - 2.20	46.66 - 2.22
No. of observed ref.*	80057(9663)	306528(16788)	294987(42057)
No. of unique ref.*	25465(3089)	78703(4374)	76018(10895)
R_{sym}*[1]	0.058(0.551)	0.104(0.827)	0.114(0.778)
R_{pim}*[2]	0.038(0.366)	0.061(0.488)	0.067(0.457)
Completeness (%)*	98.4(98.5)	98.3(95.8)	98.2(96.5)
Multiplicity*	3.1(3.1)	3.9(3.8)	3.9(3.9)
I/σ(I)*	14.5(2.0)	13.6(1.9)	10.3(1.8)
Refinement Statistics			
Resolution range (Å)	46.1-2.45	46.7-2.20	46.66-2.22
All reflections used	25922	80009	77389
Size R_{free} set (%)	5	5	5
All reflections (R_{free})	1318	4013	3887
R-values			
R_{cryst} (%)	19.28	18.49	18.22
R_{free} (%)	23.07	22.45	22.87
Matthews coefficient ($Å^3$ Da^{-1})	2.16	2.78	2.47
Solvent content (%)	43.03	55.81	50.28
Mean B Factor ($Å^2$)			
Protein	61.59	34.89	36.36
Water	60.33	33.98	35.54
Zinc	59.04	35.19	49.73
Magnesium	24.99		
Ethylene glycol		29.42	35.15
Sodium		32.21	
Potassium		37.01	42.22
Coenzyme		30.35	54.85
DHAP			48.46
G1P			24.93
RMSD			
Bonds (Å)	0.0148	0.0143	0.0142
Angles (°)	1.5256	1.5767	1.5914
Ramachandran Plot			
favored regions (%)	97.2	97.7	97.6
additionally allowed regions (%)	2.8	2.3	2.4

*Data in the highest resolution shell is in parentheses (*Apo* G1PDH 2.56-2.45 Å, *Binary* complex 2.25-2.20 Å and *Ternary* complex 2.35-2.22 Å). [1]$R_{sym} = \frac{\sum_{hkl}\sum_j |I_{hkl,j} - \langle I_{hkl}\rangle|}{\sum_{hkl}\sum_j I_{hkl,j}}$.
[2]The precision indicating merging R factor value $R_{pim} = \frac{\sum_{hkl}\sqrt{\frac{1}{n-1}}\sum_{j=1}^{n} |I_{hkl,j} - \langle I_{hkl}\rangle|}{\sum_{hkl}\sum_j I_{hkl,j}}$.

discontinuous to some degree for the nicotinamide ring. The conserved GGG*XXX*D motif (32) makes a number of hydrogen bond contacts with the coenzyme including the pyrophosphate bridge, ribose, and 3′-amine of the nicotinamide ring (Fig. 2*D*). Additional coenzyme-binding residues include Asn^{104} (pyrophosphate bridge and ribose sugar), Ser^{103} (pyrophosphate bridge), and Val^{112} (3′-amine of the nicotinamide ring). The 2′-phosphate moiety of NADP(H) forms hydrogen bond contacts with Tyr^{52}, Thr^{39}, and Asn^{38}, whereas the adenine moiety forms hydrogen bond contacts with Thr^{100} and in some instances with Tyr^{42} via a water molecule. The orientation of NADPH in the ternary complex (including the *anti* conformation of the nicotinamide ring) indicates pro-*R* specificity, confirming a previous prediction based on modeling using a GDH (32). The preference of G1PDH for NADPH is 3 times that for NADH (Table 2).

Structural alignment (Fig. 4) suggests that coenzyme preference may be due in part to Gly^{36} immediately adjacent to the 2′-phosphate of NADP in G1PDH. The absence of a side chain creates a pocket that can accommodate the phosphate moiety, whereas for example in GDH (Protein Data Bank code 1JQ5 (33)) or iron-containing ADH (Protein Data Bank code 3JZD), these positions are occupied by bulkier Asp and Thr residues, respectively, which in turn are found to make hydrogen bond contacts with the ribose of NAD(H).

The four molecules in the asymmetric unit showing the positions and binding interactions of DHAP, NADPH, Zn^{2+}, and product G1P enable a structure-based analysis of the enzyme-catalyzed stereospecific reaction (Fig. 3). The ternary complex reveals a possible stepwise mechanism affecting, but not limited to, the number and position of water molecules in the active site, the movement of the nicotinamide ring into the pro-*R* position, the change in rotameric state of the coenzyme-binding residue Asn^{104}, and the DHAP phosphate-binding residues Gln^{116} and Arg^{310}. Molecule C marks the beginning of the catalytic process; Zn^{2+} is coordinated to His^{226}, His^{247}, and Asp^{148}; one water molecule; and the carbonyl of DHAP (and to some degree with the hydroxyl). The nicotinamide ring of NADPH is not in the pro-*R* position and is 3.35 Å from the side chain amine of Asn^{104} (with the ribose sugar). Arg^{310} and Ser^{222} form the shortest hydrogen bond interactions with the phosphate group. In monomer D, NADPH has moved into the pro-*R* position (3.13-Å distance between the carbonyl of DHAP and C4 of NADPH), moving closer to the amine of Asn^{104} (3.24 and 3.31 Å away from the pyrophosphate bridge and ribose sugar, respectively). The phosphate group of DHAP is now coordinated closest to residues Ser^{222} and His^{230}, and the hydroxyl moiety no longer coordinates tightly with Zn^{2+}. In monomer A, there is no discernible density for a coordinated water molecule bound to Zn^{2+}; however, a new water molecule is seen in the active site 2.93 Å from the carbonyl of DHAP and 2.96 Å from C4 of NADPH. The phosphate of DHAP now coordinates with the side chain amine of Gln^{116} in addition to histidine, arginine, and the series of serines that encircle the phosphate group. The coenzyme moves closer to Asn^{104} (3.16 Å from the pyrophosphate bridge and 3.26 Å from the ribose sugar). Monomer B shows a bound G1P molecule and the C4 of the coenzyme molecule (still in the pro-*R* position) 3.05 Å from the hydroxyl of G1P (O3), which no longer coordinates Zn^{2+}. No water molecules are observed in the immediate vicinity of G1P, and Asn^{104} is 3.12 Å away from the pyrophosphate bridge of NADP.

These observations suggest the following points. First, the sole role of His^{226}, His^{247}, and Asp^{148} is to coordinate the active site Zn^{2+}. However, this coordination is influenced by bound substrate and coenzyme molecules. Second, a pro-*R* hydride transfer/relay system is in operation during catalysis, mediated by water and completed by the polarization of the carbonyl of DHAP by Zn^{2+}. It has been suggested that Zn^{2+} may also help polarize a key hydroxyl group on the substrate as part of the reaction mechanism (32, 34), and we did observe an initial coordination of the hydroxyl to Zn^{2+}. Third, Asn^{104} may influence affinity for the coenzyme in G1PDH and the rate of reaction via its interaction with NADPH. It should also be noted that in some monomers of G1PDH undefined density was seen around the side chain carbonyl of Asn^{104}, suggesting some type of oxidation or reduction event. The structure also explains previous assertions of a secondary ion-binding site (occupied by a K^+) and the role of Asp^{105} in coordinating the hydroxyl of G1P in the active site (32, 35). Previous studies suggested that the metal coordination residues do not function in substrate binding (35), and in the case of bound DHAP, we observed no interaction

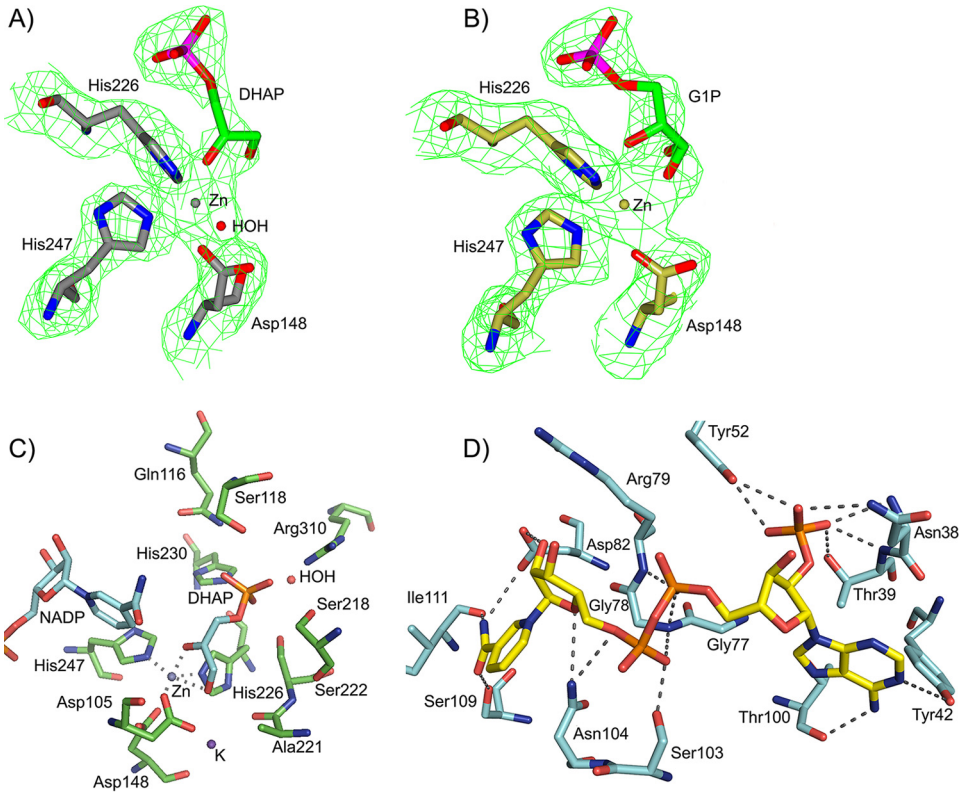

FIGURE 2. **Representations of the ternary complex of MJ G1PDH.** Omit maps ($F_o - F_c$) set to 2.5 σ of the substrate DHAP (*A*) and product G1P (*B*) coordinated to Zn^{2+} (*centered gray* and *yellow spheres*) and the conserved residues of His^{226}, His^{247}, and Asp^{148} are shown. *C*, the hydrogen bond interactions mediated by bound DHAP in the active site (composite diagram); contacts with Zn^{2+} are indicated with *dotted lines*. *D*, hydrogen bond interactions (indicated with *dotted lines*) formed by NADPH (*yellow* carbon chain backbone) with G1PDH (*cyan* carbon chain backbone) in the ternary complex.

with either histidine or aspartic acid residues. However, bound G1P could form a hydrogen bond via its 2-hydroxyl with the carboxylate of Asp^{148} (Oδ1; 3.30 Å).

Structural alignment of MJ G1PDH using Dali-lite (36) (Fig. 4 and Table 3) identified homology with a number of metal-binding GDHs and ADHs as well as DHQSs. These often dimeric or multimeric enzymes possess common N- and C-terminal macroarchitectures including the α-helical array at the C terminus and a Rossmann fold-type element at the N terminus despite low sequence identity (15–33%). Analysis of the Rossmann fold showed that G1PDH has contracted β_3, α_2, and α_3 secondary structural elements (residues 37–69) with shorter loops and a smaller β-hairpin loop (residues 100–121) when compared with its structural homologues. The orientation of α_3 is also unique, running perpendicular, and not antiparallel, with β_3. Sequence alignment of MJ G1PDH with other archaeal species suggests that these contracted elements in MJ G1PDH may reflect thermal adaption, limiting the surface exposure of the enzyme suited for activity in hydrothermal environments. G1PDHs of non-thermophilic archaeal species do not possess these contractions in their homologous Rossmann fold loop regions and β-hairpin loop. A number of these secondary structure-related contractions may also be responsible for the monomeric solution state for MJ G1PDH observed by gel filtration chromatography experiments. MJ G1PDH lacks a number of N- and C-terminal structural elements ascribed to its homologues that enable the formation of biochemically active dimers, octamers, and decamers (33, 37) including an elongated β-hairpin loop, which is critical in forming the dimer interface in DHQS from *Thermus thermophilus* (38). QtPISA analysis (39) of the MJ G1PDH crystal structure indicated a positive protein interaction and free energy gain upon the formation of a MJ G1PDH dimer, revealing some potential for complex formation in solution. Among only the ADHs and best exemplified by the 1,3-propanediol oxidoreductase (Protein Data Bank code 3BFJ) exists an additional α-helical domain and elongated loop in the C terminus (between α_{10} and α_{11} of G1PDH). This region, containing conserved residues, facilitates the stabilization of the quaternary structure via ionic and hydrogen bond interactions with corresponding residues of neighboring dimers to help form a decamer formed by a pentamer of dimers (Protein Data Bank code 3BFJ (37)). However, within G1PDH, we observed intersubunit contacts similar to some homodimeric ADHs (40, 41) (Fig. 1*B*) including contacts between the short antiparallel β-sheet at the N terminus (residues 1–11) in the crystal structure. The side chain of Arg^7 also forms hydrogen bonds with the carbonyls of Ala^{123} (on the elongated loop connecting β_5 and β_6) and Thr^5. Although these remain the only obvious polar contacts between the monomers, we also observed a hydrophobic patch at the interface of α_7 and α_8 of opposing monomers by Ala^{175}, Ile^{176}, Phe^{177}, and Ile^{181} on α_7 and residues 210–215 on α_8 (Fig. 4). Previous biochemical characterizations of archaeal G1PDHs have shown the enzyme to be multimeric (17, 42).

The dehydrogenase superfamily displays variable metal binding capabilities dependent entirely upon the makeup of

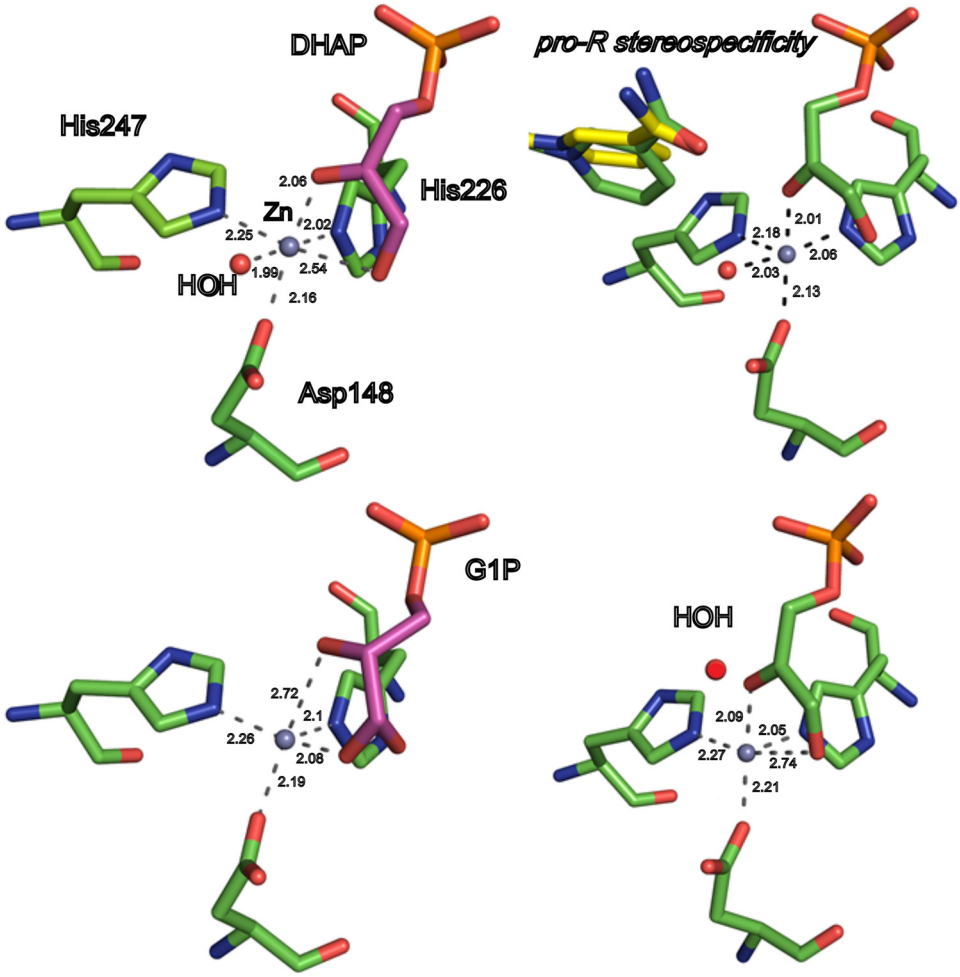

FIGURE 3. **The coordination of Zn^{2+} and structure-based catalytic mechanism in the ternary complex of G1PDH.** The mechanism starts from the *top left* with substrate DHAP and proceeds in a clockwise manner ultimately producing *sn*-G1P. Water molecules are indicated with *red spheres*, and Zn^{2+} is indicated as a *gray sphere*. Distances are represented as *dashed lines* measured in Å.

TABLE 2

Kinetic constants of MJ G1PDH

G1PDH was EDTA-treated and assayed in the presence of 0.1 mM Zn^{2+}. Data are presented as means ± S.E.

Substrate	K_m	V_{max}	k_{cat}	k_{cat}/K_m
	mM	*μmol min⁻¹ mg⁻¹*	*min⁻¹*	*mM⁻¹ min⁻¹*
DHAP (NADPH)	2.05 ± 0.19	4.09 ± 0.11	167 ± 5	81.6
DHAP (NADH)	1.11 ± 0.06	2.11 ± 0.03	86.4 ± 1.1	77.7
NADPH	0.0432 ± 0.0090	4.73 ± 0.27	194 ± 11	4480
NADH	0.141 ± 0.061	4.88 ± 0.81	200 ± 33	1420

their metal-coordinating residues within the active site (Fig. 5, *A* and *B*). Zn^{2+} is utilized by G1PDH, GDH, and DHQS, whereas the ADHs utilize Fe^{2+} or Zn^{2+}. Interestingly, the only characterized bacterial G1PDH shows much higher activity when expressed in *E. coli* with added Ni^{2+} rather than Zn^{2+} (43), whereas the closest structural homologue of MJ G1PDH is from *C. acetobutylicum* (Protein Data Bank code 3CE9); although annotated as a Zn^{2+}-binding GDH, it is more likely to be a G1PDH (44). The location of the catalytic metal coordination/binding site is almost identical among the enzymes listed in Table 3 and is maintained via two strictly conserved histidine residues (on conserved helices α_9 and α_{10}) and a mostly conserved aspartic acid on α_6 (Fig. 5, *A* and *B*). It should be noted that the metal-coordinating Asp residue is an Asn in Protein Data Banks codes 3HL0, 3JZD, and 3IV7. 3-Dehydroquinate synthase (Protein Data Bank code 3ZOK) replaces this same residue with a glutamate and, unlike its GDH counterparts, demonstrates a markedly lower affinity for Zn^{2+} (45). Excluding the GDHs, a fourth metal-coordinating residue is observed downstream of Asp^{148} among the small molecule alcohol dehydrogenases, and Fe^{2+} is the preferred metal. Our structural alignment supports previous assertions (37) that this fourth coordinating residue is a histidine among the Gram-negative bacteria (Protein Data Bank codes 3BFJ, 1RRM, 3OWO, 1VLJ, 1VHD, 3RF7, and 1OJ7) on our list and a glutamine in Gram-positive bacteria (iron-binding 1,3-propanediol dehydrogenase from *Oenococcus oeni*; Protein Data Bank code 4FR2 (40)).

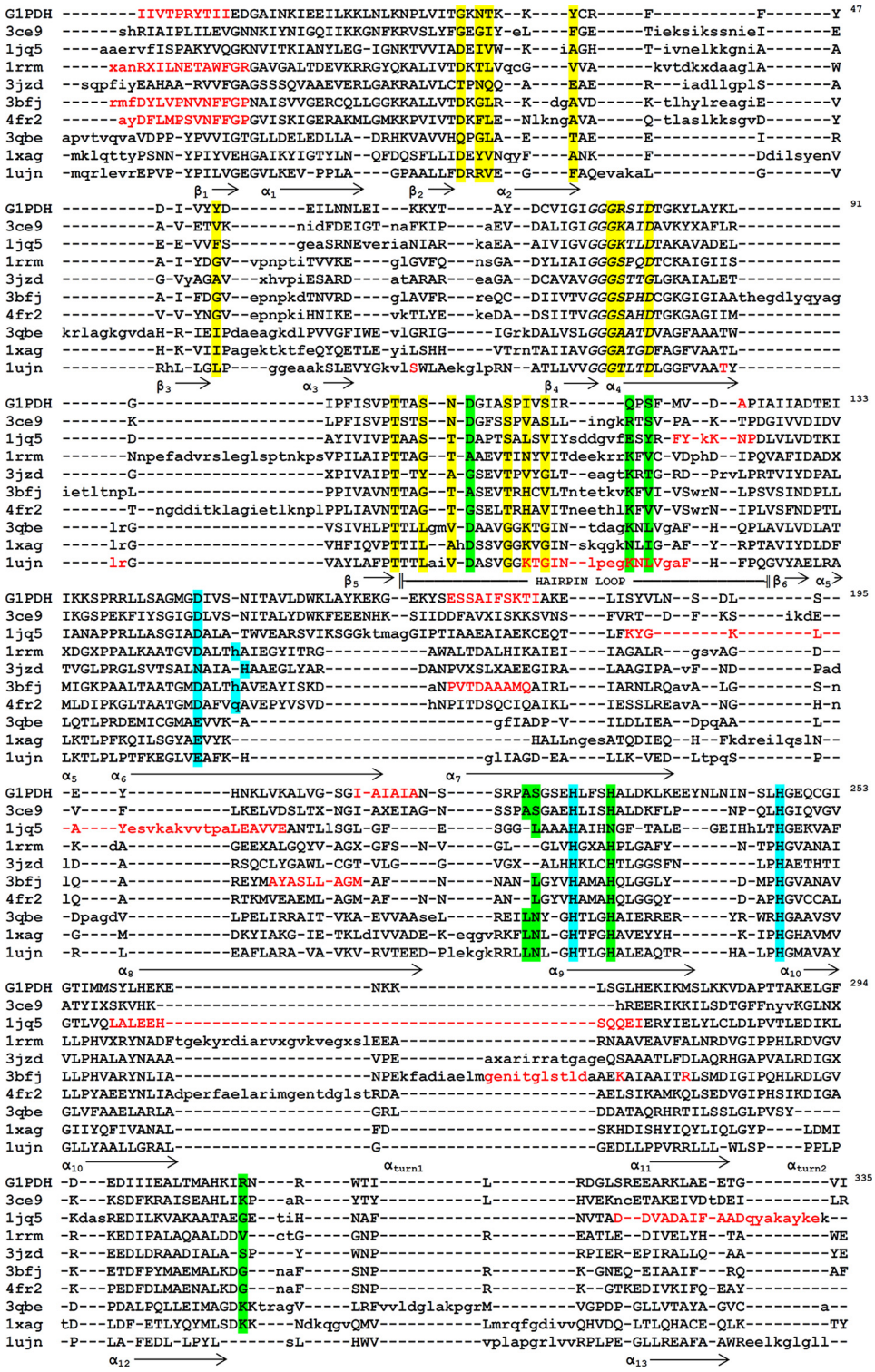

FIGURE 4. **Expanded gap version of Dali-lite pairwise structural alignment of MJ G1PDH (36).** Selected members of the larger enzyme superfamily including GDHs (Protein Data Bank codes 3CE9 and 1JQ5), ADHs (Protein Data Bank codes 1RRM, 3JZD, 3BFJ, and 4FR2), and DHQSs (Protein Data Bank codes 3QBE, 1XAG, and 1UJN) were used in the alignment. Secondary structural elements and residue numbering correspond to G1PDH. Residues that coordinate metals are highlighted in *blue*. The coenzyme binding motif is in *italics*, whereas residues that interact with coenzyme (NADP(H)) and substrate (DHAP) with respect to MJ G1PDH are highlighted in *yellow* and *green*, respectively. Reported intersubunit contacts between monomers are in *red*. *Uppercase lettering* indicates structurally equivalent positions with G1PDH, whereas *lowercase* indicates insertions relative to G1PDH.

Biochemical Characterization of the MJ G1PDH—In vitro analysis of MJ G1PDH revealed optimal specific activity between pH 6.5 and 7.5 (Fig. 6*A*) with an optimum concentration of ~150 mM KCl producing a more than 400% increase in activity compared with an absence of KCl (Fig. 6*B*). The enzyme showed much higher activity with K^+ compared with Na^+ but

TABLE 3

Top structural alignment hits from the Dali-based structural alignment of MJ G1PDH (36)

MR, maleylacetate reductase; POR, 1,3-propanediol oxidoreductase.

Organism	Class	Protein Data Bank code-monomer	Z-score[a]	r.m.s.d.[b]	lali[c]	%id[d]
C. acetobutylicum ATCC 824	GDH	3CE9-A	39.1	2.0	310	34
Sinorhizobium meliloti	GDH	3UHJ-A	31.7	2.6	305	21
Geobacillus stearothermophilus	GDH	1JQ5-A	30.3	2.5	300	21
Serratia plymuthica A30	GDH	4MCA-A	30.1	2.7	302	21
Schizosaccharomyces pombe	GDH	1TA9-B	29.7	2.7	306	20
T. maritima	GDH	1KQ3-A	29.5	2.6	301	26
E. coli	Lactaldehyde reductase	1RRM-A	27.2	3.1	301	16
Ralstonia eutropha	Fe-ADH	3JZD-A	26.9	2.9	302	15
Rhizobium sp. MTP-10005	MR	3W5S-A	26.8	3.2	300	15
Klebsiella pneumoniae	POR	3BFJ-A	26.7	3.1	302	18
Agrobacterium tumefaciens	MR	3HL0-A	26.8	3.0	302	15
Zymomonas mobilis	ADH 2	3OWO-A	26.4	3.2	302	16
O. oeni	POR	4FR2-A	26.4	3.2	302	20
Corynebacterium glutamicum	ADH IV	3IV7-A	26.0	3.1	303	16
T. maritima	Butanol dehydrogenase	1VLJ-B	25.2	3.5	299	18
T. maritima	Fe-ADH	1VHD-A	25.1	3.2	294	18
Geobacillus thermoglucosidasius	ADH	3ZDR-A	24.2	3.2	299	16
Shewanella denitrificans	Fe-ADH	3RF7-A	24.2	3.4	290	17
E. coli	Hypothetical oxidoreductase YqhD	1OJ7-A	23.6	3.9	300	17
Actinidia chinensis	DHQS	3ZOK-D	23.0	3.5	294	16
Mycobacterium tuberculosis	DHQS	3QBE-A	22.9	2.9	287	14
Aspergillus nidulans	DHQS	1NVB-B	22.3	3.3	290	17
Streptomyces hygroscopicus	Cyclase	4P53-A	21.0	2.9	277	16
Staphylococcus aureus	DHQS	1XAG-A	20.9	3.3	282	17
Bacillus circulans	2-Deoxy-*scyllo*-inosose synthase	2GRU-A	20.3	3.2	283	18
Staphylococcus aureus	DHQS	1XAH-A	19.5	3.2	264	16
Vibrio cholerae	DHQS	3OKF-A	19.4	3.4	280	18
T. thermophilus	DHQS	1UJN-A	18.8	3.4	271	14
Helicobacter pylori	DHQS	3CLH-A	17.6	3.4	255	21

[a] A measure of the statistical significance of the result relative to an alignment of random structures.
[b] Root mean square deviation of α-carbon atoms.
[c] Number of aligned residues.
[d] Sequence identity between the two chains.

similar activities for the anions Cl^- and formate (Fig. 6*B*). EDTA-treated G1PDH in the absence of Zn^{2+} showed less than 2% activity when compared with untreated enzyme in the presence of Zn^{2+}. However, when assayed with Zn^{2+} present, activity of the EDTA-treated enzyme was greater than that of the untreated enzyme. Activity of the EDTA-treated G1PDH increased rapidly with increasing Zn^{2+} concentrations (Fig. 6*C*) with an optimum concentration of about 0.1 mM. Although this observation suggests a strict dependence on Zn^{2+} for catalysis, a range of divalent metal ions were also tested for their ability to restore activity to the EDTA-treated G1PDH, and activity (highest to lowest) was observed for Co^{2+}, Mn^{2+}, Mg^{2+}, and Cd^{2+}. Interestingly, activity with Co^{2+} was higher than that of Zn^{2+} (about 150%) and lower for the other divalent ions (~20%). There was little to no activation by Ca^{2+}, Ni^{2+}, Sr^{2+}, Cu^{2+}, Fe^{2+}, and Ba^{2+}. G1PDH followed Michaelis-Menten kinetics using the substrate DHAP (Fig. 7). The apparent kinetic constants for G1PDH at 65 °C and pH 7.85 are shown in Table 2. NADPH was the preferred substrate compared with NADH with both K_m values and k_{cat}/K_m values differing by a factor of 3.

The apparent molecular mass of the purified recombinant G1PDH pET151D was 44 kDa as determined by gel filtration chromatography and 43 kDa by SDS-PAGE, whereas that of G1PDH pET100D was 49 and 47 kDa, respectively. These values are close to those of 41,058 and 41,401 Da predicted for the His-tagged G1PDH pET151D and G1PDH pET100D proteins, respectively (368 and 371 amino acids), and indicate that G1PDH pET151D and G1PDH pET100D are both monomeric.

G1PDHs have been previously characterized biochemically from three thermophilic archaea (17, 35, 42, 46–48), and the enzyme reaction has been shown to follow an ordered bi-bi reaction mechanism with the reaction favoring the production of *sn*-glycerol 1-phosphate (17, 42). Biochemically, the MJ G1PDH is similar to the G1PDHs described from *Methanothermobacter thermautotrophicus* ΔH and *Aeropyrum pernix* K1 with a near-neutral pH optimum, which is substantially higher than the pH 6.2 optimum for G1PDH from the acidophilic *Sulfolobus tokodaii* (48). A specific requirement for Zn^{2+} has been found for the *A. pernix* G1PDH with an optimum concentration of 0.5–1.0 mM (35). Additionally, atomic absorption analysis has shown the *A. pernix* and *M. thermautotrophicus* G1PDHs to both contain Zn^{2+} and in the case of the *Aeropyrum* enzyme 0.81 mol of Zn^{2+} per monomer and small amounts of magnesium and manganese (less than 0.05 mol of metal ion per monomer; Refs. 35 and 47).

The MJ G1PDH K_m and k_{cat}/K_m values for NADPH and NADH show a preference for NADPH. Similarly, the *M. thermautotrophicus* enzyme K_m and k_{cat}/K_m values show a preference for NADPH (factors of 5 and ~3, respectively) (17). Specific activity values of the enzyme from *S. tokodaii* also show a preference for NADPH by a factor of ~3 (48). In contrast, the *A. pernix* enzyme shows a slight preference for NADH if K_m values are compared but a preference for NADH by a factor of 5 when k_{cat}/K_m values are compared (41). The MJ G1PDH K_m value for

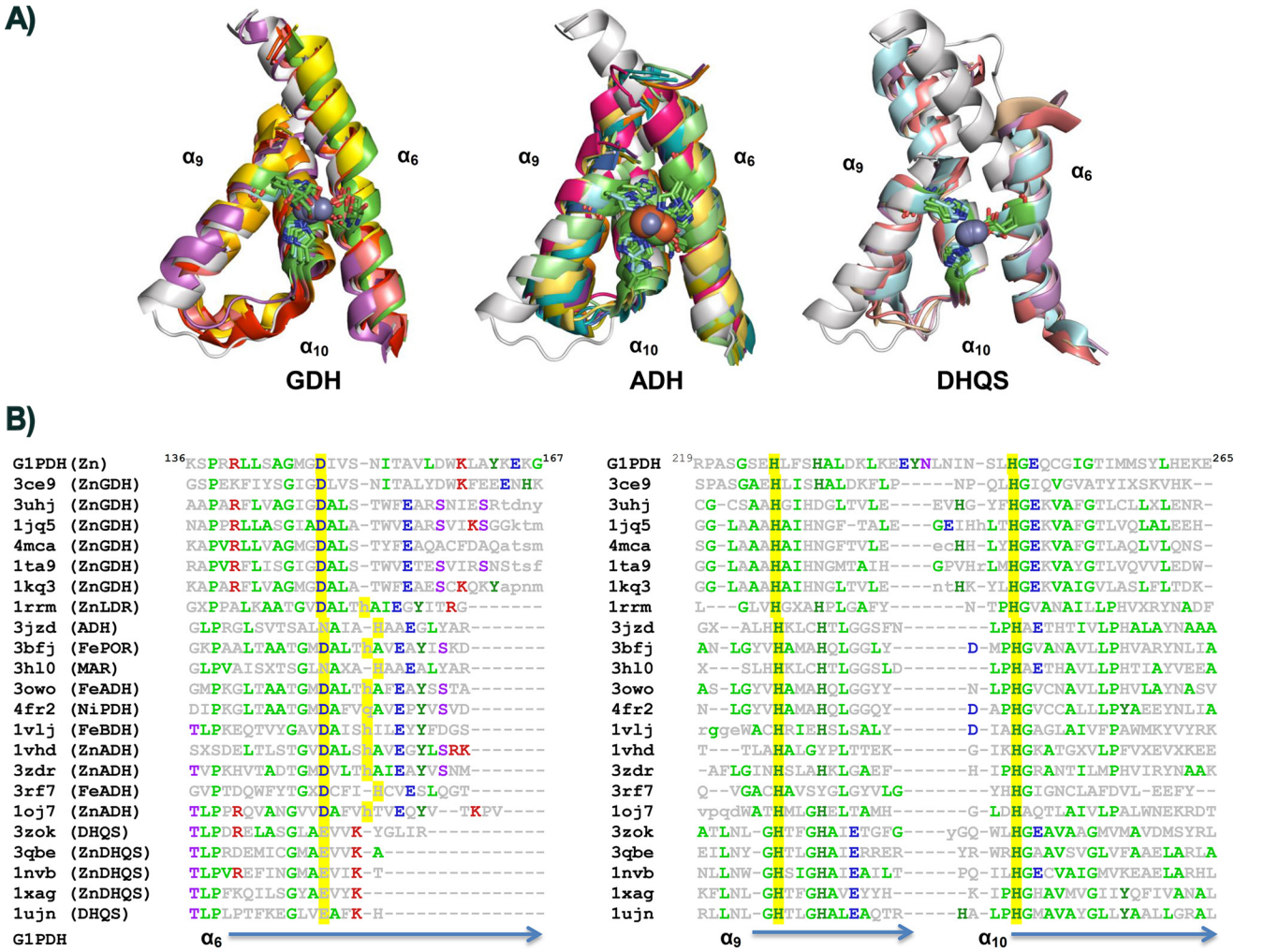

FIGURE 5. **Metal-coordinating residues in the dehydrogenase superfamily.** *A*, a ribbon representation of the active site domain of MJ G1PDH (in *white*) and metal-binding residues Asp[148], His[226], and His[247] (in *blue* as *sticks*) superimposed on key enzymes with respect to activity within the larger superfamily including GDHs, ADHs (including lactaldehyde dehydrogenase (*LDR*), maleylacetate reductase (*MAR*), 1,3-propanediol oxidoreductase (*POR*), butanol dehydrogenase (*BDH*), and 1,3-propanediol dehydrogenase (*PDH*)), and DHQSs. Bound metals if present are listed and shown as *spheres* (Zn^{2+} in *gray*, Fe^{2+} in *orange*, and Ni^{2+} in *green*). *B*, an expanded gap version of the Dali-lite (36) pairwise structural alignment of MJ G1PDH. Secondary structural elements and residue numbering correspond to G1PDH. Residues that coordinate metals are highlighted in *yellow*. *Uppercase lettering* indicates structurally equivalent positions with G1PDH, whereas *lowercase* indicates insertions relative to G1PDH.

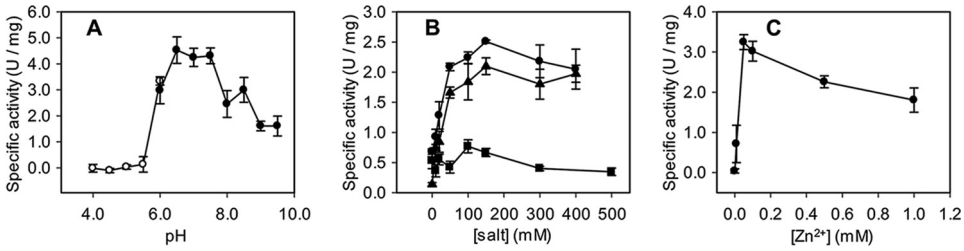

FIGURE 6. **Properties of MJ G1PDH.** *A*, pH dependence. ○, citrate buffer; ●, Bistris propane buffer. *B*, effect of [salt] on activity. ●, KCl; ▲, potassium formate; ■, NaCl. *C*, effect of [Zn^{2+}] on activity. Specific activity (units/mg) of EDTA-treated G1PDH (in the presence of 0.1 mM Zn^{2+}) is shown. Assays were performed in triplicate. *Error bars* represent S.D.

the substrate DHAP with NADPH is 2.05 mM, and the K_m values for DHAP with the preferred NAD(P)H cofactor for the three previously characterized G1PDHs are similar: *M. thermautotrophicus*, 0.58 mM (NADPH); *S. tokodaii*, 0.47 mM (NADPH); and *A. pernix*, 0.46 mM (NADH) (17, 42, 48). The V_{max} for MJ G1PDH for the substrate DHAP with NADPH was 4.09 μmol min^{-1} mg^{-1} at 65 °C, which is 2 orders of magnitude lower than that of 323 μmol min^{-1} mg^{-1} for the *M. thermautotrophicus* enzyme (NADPH) and 2 orders of magnitude higher than that of 32.8 nmol min^{-1} for the *S. tokodaii* enzyme (NADPH) (17, 48). The *M. thermautotrophicus* enzyme was the only G1PDH assayed at the growth temperature of the source organism. Hence, the V_{max} and k_{cat} values of the other three characterized enzymes could be expected to be higher if assayed at the growth temperature of the source organisms. The V_{max} for MJ G1PDH is very similar to the specific activity of

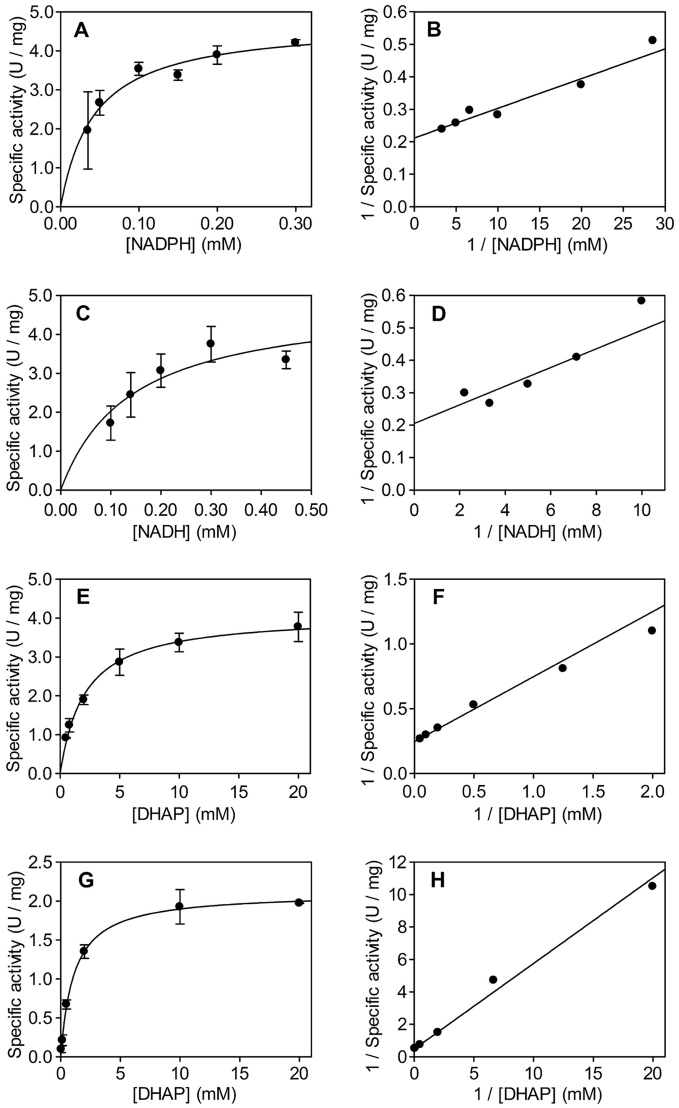

FIGURE 7. **Michaelis-Menten (*left*) and Lineweaver-Burk (*right*) plots for EDTA-treated G1PDH.** *A*, and *B*, NADPH. *C*, and *D*, NADH. *E*, and *F*, DHAP (NADPH). *G*, and *H*, DHAP (NADH). Assays were performed in triplicate in the presence of 0.1 mM Zn^{2+}. *Error bars* represent S.D.

the *A. pernix* enzyme (3.22 μmol min^{-1} mg^{-1}), and the k_{cat} values of the two enzymes are also very similar (167 (NADPH) and 154.3 min^{-1} (NADH), respectively) (42).

Phylogenetic and Structure-based Sequence Analyses—The structure of the MJ G1PDH displays a similar overall fold to related enzymes of the DHQSs, GDHs, and ADHs, indicating common ancestry within the superfamily despite limited sequence identity. Based on structural similarity (root mean square distances between the superimposed structures main chains) and structure-based sequence identities, the order of similarity of the individual families starting from G1PDH is GDH, ADH, and then DHQS (Figs. 4 and 5, *A* and *B*, and Table 3). The Dali-lite analysis and PromalS3D structure-based phylogenetic tree containing archaeal and bacterial G1PDHs, DHQSs, GDHs, and ADHs separate each enzyme type into distinct and highly supported clades (≥98%; Fig. 8) (20, 36). The phylogenetic analysis also suggests that archaeal G1PDHs are closest to GDHs, and this is supported by a cluster of orthologous groups (COG) assignment of COG0371, which contains GDHs and G1PDHs (45, 49). The tree in Fig. 8 is in general agreement with previously described sequence-only based trees (32, 50) particularly with respect to the clear and strongly supported groupings of the individual superfamily member enzyme types. The archaeal G1PDH branching pattern concurs with 16S rRNA gene and ribosomal protein gene phylogenies (51–53) albeit with some discrepancies and with low bootstrap support (<50%) in some cases. For example, the euryarchaeal sequences (*e.g.* methanogens, Thermoplasmatales, Thermococcales, and Halobacteriales) group together. In addition, members of the so-called TACK group, the Thaumarchaeota, Aigarchaeota, Crenarchaeota sequences, and the single korarchaeal sequence, group together with 53% support (54). The Thaumarchaeota form a highly supported group (94% support). Although some bacterial G1PDH-like sequences were identified in BLASTP searches and form two distinct clades, these are relatively few in number and were found almost entirely in Firmicutes and high G + C content Actinobacteria. The very limited and sporadic distribution of G1PDH-like sequences in bacteria is suggestive of genes that have been acquired through lateral gene transfer (50). One of the bacterial G1PDH-like clades has high bootstrap support (99%), is more deeply branching, contains the actinobacterial species *Streptomyces* and *Thermobifida*, and tends to have shorter branch lengths than the other bacterial clade, which contains mostly Firmicutes but also contains sequences from *Thermotoga maritima* and Proteobacteria (*Anaplasma centrale*). The branch lengths for this latter bacterial G1PDH-like clade, which contains *C. acetobutylicum* (Protein Data Bank code 3CE9), tend to be longer than those for archaeal G1PDHs and actinobacterial G1PDH-like sequences, suggesting a more rapid rate of evolution. The universal presence of G1P in Archaea and the structure-based phylogeny presented here indicate that G1PDH was present in the archaeal cenancestor (50). The presence of G1PDH sequences in Archaea and in some deeply branching bacterial clades (*e.g.* Firmicutes and Actinobacteria) could be taken as evidence that G1PDH was present in the LUCA, although the distribution of bacterial G1PDHs is quite limited (10). In contrast, DHQSs are widely present in Crenarchaea but are lacking in many other archaea (55), particularly the Euryarchaea and the Thaumarchaeota, Aigarchaeota, Crenarchaeota, and Korarchaeota (TACK) superphylum (54). The essential role of DHQS catalyzing the second step in aromatic amino acid synthesis in bacteria and some archaea (*e.g.* Crenarchaea) potentially suggests an early origin for this enzymatic activity (55). GDHs and ADHs are found only sporadically in the Archaea, and GDHs have been considered to represent an ancient capability enhancing the utilization of glycerol (Fig. 8) (10, 33, 50, 56). Conversely, GDHs and ADHs are widely present in Bacteria (50, 57). The limited and somewhat irregular distribution of GDHs and ADHs in Archaea is reminiscent of having been acquired through lateral gene transfer events (50).

Discussion

We have presented here three structures of an archaeal G1PDH from the hyperthermophilic marine archaeal spe-

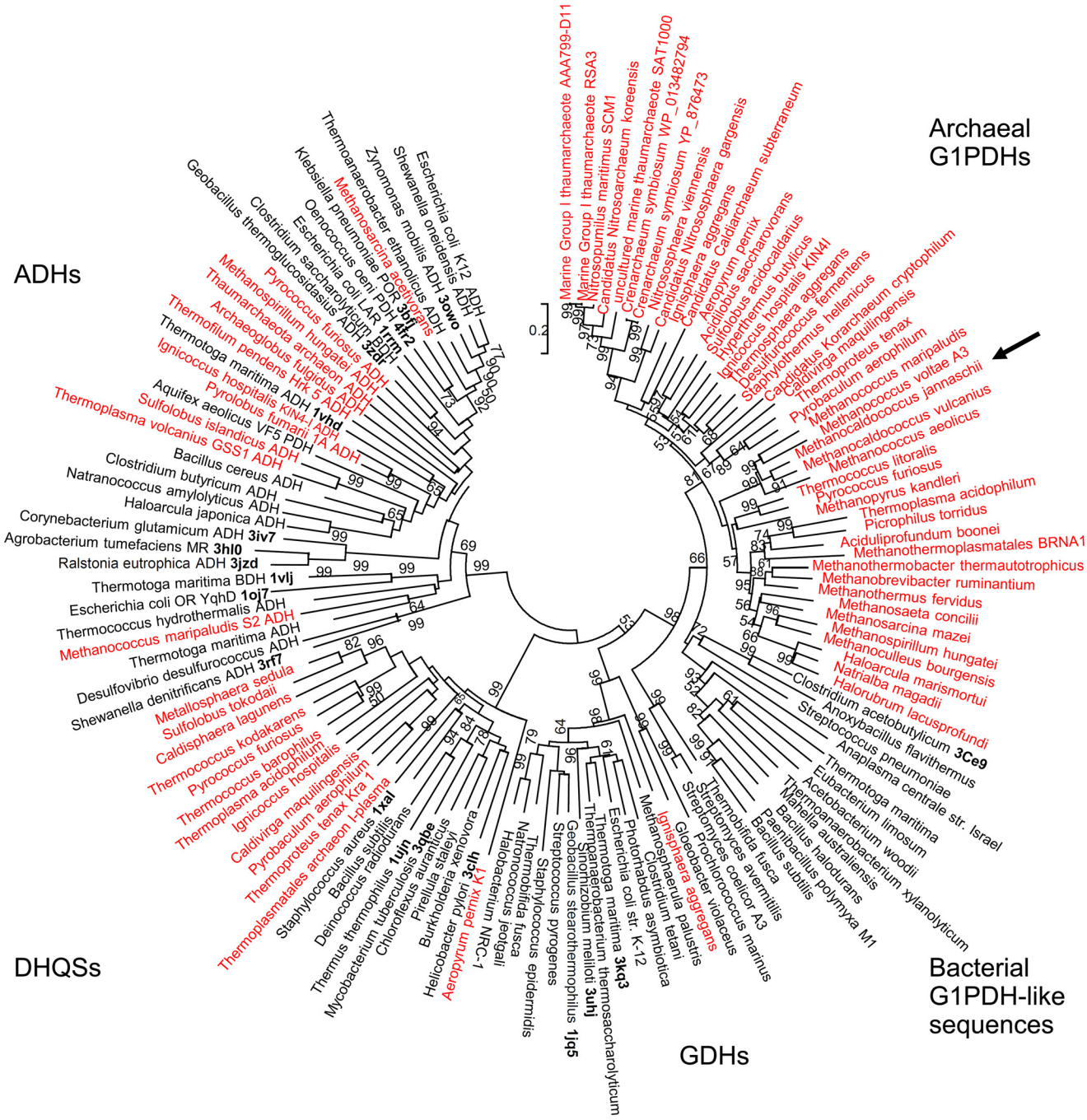

FIGURE 8. **Phylogenetic tree (circular format) of G1PDHs, DHQSs, ADHs, and GDHs.** The neighbor joining treeing method (implemented in MEGA6) based on a PromalS3D structure-based amino acid sequence alignment was used (20). Archaeal sequences are shown in *red font*, and bacterial sequences are shown in *black*. The tree incorporated 159 residues from 133 amino acid sequences. Protein Data Bank codes for available structures are included as part of the enzyme name on the tree and are in *bold font*. The *black arrow* indicates the position of MJ G1PDH. Bootstrap validation values below 50% are not shown (total of 500 bootstraps).

cies *M. jannaschii*. Analysis of the ternary structure in particular lends support to the steps involved in the catalytic mechanism of G1PDH and provides new insight into the roles of the Zn^{2+} cofactor and the NADPH coenzyme during conversion of the substrate DHAP into the stereospecific *sn*-glycerol 1-phosphate product. Phylogenetic and structure-based sequence analysis using the new archaeal G1PDH structure confirms that G1PDHs are part of the larger structurally related superfamily containing four clades of metal- and NAD(P)H-dependent dehydrogenases (G1PDHs, GDHs, DHQSs, and ADHs) and provides insight into the origins of G1PDH.

The distribution of G1PDHs, DHQSs, ADHs, and GDHs in Archaea and Bacteria suggests that at least one ancestral sequence for this metallodehydrogenase superfamily was present in the LUCA (50). Despite rare exceptions, in isolated lineages (44), the use of different glycerol stereoisomers for lipid backbones is a very robust domain-specific trait (7, 8, 10, 56–59). The bacterial glycerol-3-phosphate dehydrogenase is structurally unrelated to the archaeal G1PDH and belongs to

the 6-phosphogluconate dehydrogenase C-terminal domain-like dehydrogenase structural classification of protein (SCOP) superfamily containing UDP-glucose 6-dehydrogenases and 3-hydroxyacyl-CoA dehydrogenases, both of which are widely distributed among Archaea and Bacteria (44, 50). This suggests that at least one ancestral member of the 6-phosphogluconate dehydrogenase-like superfamily was also present in the LUCA at the same time as the G1PDH-GDH-ADH-DHQS superfamily ancestor (10). The presence of ancestral sequences for both superfamilies in the LUCA followed by the creation of G1PDH- and glycerol-3-phosphate dehydrogenase-specific clades thereof in the ancestors of Archaea and Bacteria, respectively, supports the scenario whereby domain-specific lipids arose through differential gene loss (10). This is also supported by the broad presence of CDP-alcohol phosphatidyltransferases, which catalyze the addition of polar headgroups (serine, glycerol, and *myo*-inositol) to produce intact phospholipids in both prokaryotic domains (60–62), and suggests that at least one ancestor of CDP-based phospholipid synthesis was present in the LUCA (10, 62, 63). Biochemical and phylogenomic analyses have suggested that isoprenoids and fatty acid synthesis genes might have been present in the LUCA (10, 50, 64). However, a recent extensive phylogenomic analysis of the presence of fatty acid synthesis genes in Archaea contradicts the latter suggestion, instead indicating that in those archaea containing fatty acid synthesis genes (a chimeric pathway with both bacterial-like and archaeal genes) most of the genes were likely acquired from bacteria (65). Gene distributions across the archaeal-bacterial divide can reflect either presence in the LUCA or later origins followed by interdomain lateral gene transfer (66, 67) whereby distinctions between the two are not always easy.

A common theme of proposals of early membrane evolution is that lipids were synthesized abiotically (geochemically) at first followed by biological synthesis underpinned by genes, thus leading to homochiral membranes (3, 9–11, 47, 68). Evidently, the origin of stereospecific lipid membranes entailed independent evolutionary pathways and postdated the origin of genes but preceded the divergence of the bacterial and archaeal lineages. Once established in the ancestors of the two domains, the lipid trait remained stable except at the origin of eukaryotes (69). That the energetic harnessing of chemiosmotic gradients across membranes via an ATPase is more conserved than the synthesis of the lipids themselves favors an abiotic source of lipids at the dawn of cellular evolution (7, 70, 71). Recent analyses of possible evolutionary scenarios of early membrane-based bioenergetics that incorporate a predictive quantitative model for estimating available free energy using geochemical proton gradients and sodium-proton antiporters provide support for this scenario (7).

We have presented the first structural characterization of an archaeal G1PDH, an enzyme hypothesized to have played a critical role in the speciation of Archaea (10, 13). The structures and biochemical characterization have provided new catalytic insight explaining the pro-*R* stereospecific reaction mechanism to produce G1P for archaeal lipid synthesis and have contributed to improved structural, biochemical, and phylogenetic comparisons.

Author Contributions—R. S. R., V. C., L. R. S., and A. J. S.-S. conceived and coordinated the study and wrote the paper. R. S. R. and D. D. purified DNA from *M. jannaschii* and cloned the genes into the expression vector. Y. Z., C. S., D. D., I. M. H., L. R. S., and R. S. R. purified the protein, performed crystal screens, and performed the biochemical analyses. V. C. performed crystal screens and determined the structures with help from A. J. S.-S. R. S. R. prepared the phylogenetic tree in Fig. 8. W. F. M. helped analyze the phylogenetic data and helped write the evolutionary aspects of the paper. All authors reviewed the results and approved the final version of the paper.

Acknowledgments—The x-ray data collection was undertaken on the MX1 and MX2 beamlines at the Australian Synchrotron, Victoria, Australia. We acknowledge the help of the New Zealand Synchrotron Group. We thank Mark Aspin of the Pastoral Greenhouse Gas Research Consortium for support. We also thank Adrian Cookson, Graeme Attwood, and Peter Janssen of AgResearch Ltd. for critical review of the manuscript and Pauline Hunt for help with formatting of Figs. 2, 7, and 8.

References

1. Mayer, F., and Müller, V. (2014) Adaptations of anaerobic archaea to life under extreme energy limitation. *FEMS Microbiol. Rev.* **38,** 449–472
2. Koga, Y. (2012) Thermal adaptation of the archaeal and bacterial lipid membranes. *Archaea* **2012,** 789652
3. Martin, W., and Russell, M. J. (2003) On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **358,** 59–83; discussion 83–55
4. Boucher, Y., Douady, C. J., Papke, R. T., Walsh, D. A., Boudreau, M. E., Nesbø, C. L., Case, R. J., and Doolittle, W. F. (2003) Lateral gene transfer and the origins of prokaryotic groups. *Annu. Rev. Genet.* **37,** 283–328
5. Koonin, E. V. (2009) On the origin of cells and viruses: primordial virus world scenario. *Ann. N.Y. Acad. Sci.* **1178,** 47–64
6. Lane, N., Allen, J. F., and Martin, W. (2010) How did LUCA make a living? Chemiosmosis in the origin of life. *BioEssays* **32,** 271–280
7. Sojo, V., Pomiankowski, A., and Lane, N. (2014) A bioenergetic basis for membrane divergence in archaea and bacteria. *PLos Biol.* **12,** e1001926
8. Nelson-Sathi, S., Sousa, F. L., Roettger, M., Lozada-Chávez, N., Thiergart, T., Janssen, A., Bryant, D., Landan, G., Schönheit, P., Siebers, B., McInerney, J. O., and Martin, W. F. (2015) Origins of major archaeal clades correspond to gene acquisitions from bacteria. *Nature* **517,** 77–80
9. Koga, Y. (2011) Early evolution of membrane lipids: how did the lipid divide occur? *J. Mol. Evol.* **72,** 274–282
10. Lombard, J., López-García, P., and Moreira, D. (2012) The early evolution of lipid membranes and the three domains of life. *Nat. Rev. Microbiol.* **10,** 507–515
11. Wächtershäuser, G. (2003) From pre-cells to Eukarya—a tale of two lipids. *Mol. Microbiol.* **47,** 13–22
12. Lombard, J., and Moreira, D. (2011) Origins and early evolution of the mevalonate pathway of isoprenoid biosynthesis in the three domains of life. *Mol. Biol. Evol.* **28,** 87–99
13. Koga, Y., Kyuragi, T., Nishihara, M., and Sone, N. (1998) Did archaeal and bacterial cells arise independently from noncellular precursors? A hypothesis stating that the advent of membrane phospholipid with enantiomeric glycerophosphate backbones caused the separation of the two lines of descent. *J. Mol. Evol.* **46,** 54–63
14. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72,** 248–254
15. Wedlock, D. N., Pedersen, G., Denis, M., Dey, D., Janssen, P. H., and Buddle, B. M. (2010) Development of a vaccine to mitigate greenhouse gas emissions in agriculture: vaccination of sheep with methanogen fractions

induces antibodies that block methane production *in vitro*. *N. Z. Vet. J.* **58,** 29–36
16. Leatherbarrow, R. J. (2009) *GraFit*, version 7, Erithacus Software Ltd., Horley, UK
17. Nishihara, M., and Koga, Y. (1997) Purification and properties of *sn*-glycerol-1-phosphate dehydrogenase from *Methanobacterium thermoautotrophicum*: characterization of the biosynthetic enzyme for the enantiomeric glycerophosphate backbone of ether polar lipids of Archaea. *J. Biochem.* **122,** 572–576
18. Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* **30,** 2725–2729
19. Hall, B. G. (2013) Building phylogenetic trees from molecular data with MEGA. *Mol. Biol. Evol.* **30,** 1229–1235
20. Pei, J., Kim, B. H., and Grishin, N. V. (2008) PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res.* **36,** 2295–2300
21. McPhillips, T. M., McPhillips, S. E., Chiu, H. J., Cohen, A. E., Deacon, A. M., Ellis, P. J., Garman, E., Gonzalez, A., Sauter, N. K., Phizackerley, R. P., Soltis, S. M., and Kuhn, P. (2002) Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines. *J. Synchrotron Radiat.* **9,** 401–406
22. Kabsch, W. (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 133–144
23. Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62,** 72–82
24. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67,** 235–242
25. McCoy, A. J. (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D Biol. Crystallogr.* **63,** 32–41
26. Stein, N. (2008) CHAINSAW: a program for mutating PDB files used as templates in molecular replacement. *J. Appl. Crystallogr.* **41,** 641–643
27. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67,** 355–367
28. Winn, M. D., Murshudov, G. N., and Papiz, M. Z. (2003) Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol.* **374,** 300–321
29. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132
30. Matthews, B. W. (1968) Solvent content of protein crystals. *J. Mol. Biol.* **33,** 491–497
31. McNicholas, S., Potterton, E., Wilson, K. S., and Noble, M. E. (2011) Presenting your structures: the CCP4mg molecular-graphics software. *Acta Crystallogr. D Biol. Crystallogr.* **67,** 386–394
32. Daiyasu, H., Hiroike, T., Koga, Y., and Toh, H. (2002) Analysis of membrane stereochemistry with homology modeling of *sn*-glycerol-1-phosphate dehydrogenase. *Protein Eng.* **15,** 987–995
33. Ruzheinikov, S. N., Burke, J., Sedelnikova, S., Baker, P. J., Taylor, R., Bullough, P. A., Muir, N. M., Gore, M. G., and Rice, D. W. (2001) Glycerol dehydrogenase: structure, specificity, and mechanism of a family III polyol dehydrogenase. *Structure* **9,** 789–802
34. Carpenter, E. P., Hawkins, A. R., Frost, J. W., and Brown, K. A. (1998) Structure of dehydroquinate synthase reveals an active site capable of multistep catalysis. *Nature* **394,** 299–302
35. Han, J. S., and Ishikawa, K. (2005) Active site of Zn^{2+}-dependent *sn*-glycerol-1-phosphate dehydrogenase from *Aeropyrum pernix* K1. *Archaea* **1,** 311–317
36. Holm, L., Kääriäinen, S., Wilton, C., and Plewczynski, D. (2006) Using Dali for structural comparison of proteins. *Curr. Protoc. Bioinformatics* **Chapter 5,** Unit 5.5
37. Marçal, D., Rêgo, A. T., Carrondo, M. A., and Enguita, F. J. (2009) 1,3-Propanediol dehydrogenase from *Klebsiella pneumoniae*: decameric quaternary structure and possible subunit cooperativity. *J. Bacteriol.* **191,** 1143–1151
38. Sugahara, M., Nodake, Y., Sugahara, M., and Kunishima, N. (2005) Crystal structure of dehydroquinate synthase from *Thermus thermophilus* HB8 showing functional importance of the dimeric state. *Proteins* **58,** 249–252
39. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372,** 774–797
40. Elleuche, S., Fodor, K., Klippel, B., von der Heyde, A., Wilmanns, M., and Antranikian, G. (2013) Structural and biochemical characterisation of a NAD^+-dependent alcohol dehydrogenase from *Oenococcus oeni* as a new model molecule for industrial biotechnology applications. *Appl. Microbiol. Biotechnol.* **97,** 8963–8975
41. Moon, J. H., Lee, H. J., Park, S. Y., Song, J. M., Park, M. Y., Park, H. M., Sun, J., Park, J. H., Kim, B. Y., and Kim, J. S. (2011) Structures of iron-dependent alcohol dehydrogenase 2 from *Zymomonas mobilis* ZM4 with and without NAD^+ cofactor. *J. Mol. Biol.* **407,** 413–424
42. Han, J. S., Kosugi, Y., Ishida, H., and Ishikawa, K. (2002) Kinetic study of *sn*-glycerol-1-phosphate dehydrogenase from the aerobic hyperthermophilic archaeon, *Aeropyrum pernix* K1. *Eur. J. Biochem.* **269,** 969–976
43. Guldan, H., Sterner, R., and Babinger, P. (2008) Identification and characterization of a bacterial glycerol-1-phosphate dehydrogenase: Ni^{2+}-dependent AraM from *Bacillus subtilis*. *Biochemistry* **47,** 7376–7384
44. Alarcon, D. A., Nandi, M., Carpena, X., Fita, I., and Loewen, P. C. (2012) Structure of glycerol-3-phosphate dehydrogenase (GPD1) from *Saccharomyces cerevisiae* at 2.45 Å resolution. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **68,** 1279–1283
45. Mittelstädt, G., Negron, L., Schofield, L. R., Marsh, K., and Parker, E. J. (2013) Biochemical and structural characterisation of dehydroquinate synthase from the New Zealand kiwifruit *Actinidia chinensis*. *Arch. Biochem. Biophys.* **537,** 185–191
46. Nishihara, M., and Koga, Y. (1995) *sn*-Glycerol-1-phosphate dehydrogenase in *Methanobacterium thermoautotrophicum*: key enzyme in biosynthesis of the enantiomeric glycerophosphate backbone of ether phospholipids of archaebacteria. *J. Biochem.* **117,** 933–935
47. Koga, Y., Sone, N., Noguchi, S., and Morii, H. (2003) Transfer of pro-R hydrogen from NADH to dihydroxyacetonephosphate by *sn*-glycerol-1-phosphate dehydrogenase from the archaeon *Methanothermobacter thermautotrophicus*. *Biosci. Biotechnol. Biochem.* **67,** 1605–1608
48. Koga, Y., Ohga, M., Tsujimura, M., Morii, H., and Kawarabayasi, Y. (2006) Identification of *sn*-glycerol-1-phosphate dehydrogenase activity from genomic information on a hyperthermophilic archaeon, *Sulfolobus tokodaii* strain 7. *Biosci. Biotechnol. Biochem.* **70,** 282–285
49. Tatusov, R. L., Koonin, E. V., and Lipman, D. J. (1997) A genomic perspective on protein families. *Science* **278,** 631–637
50. Peretó, J., López-García, P., and Moreira, D. (2004) Ancestral lipid biosynthesis and early membrane evolution. *Trends Biochem. Sci.* **29,** 469–477
51. Pace, N. R. (2009) Mapping the tree of life: progress and prospects. *Microbiol. Mol. Biol. Rev.* **73,** 565–576
52. Pester, M., Schleper, C., and Wagner, M. (2011) The Thaumarchaeota: an emerging view of their phylogeny and ecophysiology. *Curr. Opin. Microbiol.* **14,** 300–306
53. Gribaldo, S., and Brochier-Armanet, C. (2006) The origin and evolution of Archaea: a state of the art. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **361,** 1007–1022
54. Guy, L., and Ettema, T. J. (2011) The archaeal 'TACK' superphylum and the origin of eukaryotes. *Trends Microbiol.* **19,** 580–587
55. Grochowski, L. L., and White, R. H. (2008) Promiscuous anaerobes: new and unconventional metabolism in methanogenic archaea. *Ann. N.Y. Acad. Sci.* **1125,** 190–214
56. Soderberg, T. (2005) Biosynthesis of ribose-5-phosphate and erythrose-4-phosphate in archaea: a phylogenetic analysis of archaeal genomes. *Archaea* **1,** 347–352
57. Boucher, Y., Kamekura, M., and Doolittle, W. F. (2004) Origins and evolution of isoprenoid lipid biosynthesis in archaea. *Mol. Microbiol.* **52,** 515–527
58. Guldan, H., Matysik, F. M., Bocola, M., Sterner, R., and Babinger, P. (2011) Functional assignment of an enzyme that catalyzes the synthesis of an archaea-type ether lipid in bacteria. *Angew. Chem. Int. Ed. Engl.* **50,** 8188–8191

59. Koga, Y. (2014) From promiscuity to the lipid divide: on the evolution of distinct membranes in Archaea and Bacteria. *J. Mol. Evol.* **78,** 234–242
60. Sciara, G., Clarke, O. B., Tomasek, D., Kloss, B., Tabuso, S., Byfield, R., Cohn, R., Banerjee, S., Rajashankar, K. R., Slavkovic, V., Graziano, J. H., Shapiro, L., and Mancia, F. (2014) Structural basis for catalysis in a CDP-alcohol phosphotransferase. *Nat. Commun.* **5,** 4068
61. Nogly, P., Gushchin, I., Remeeva, A., Esteves, A. M., Borges, N., Ma, P., Ishchenko, A., Grudinin, S., Round, E., Moraes, I., Borshchevskiy, V., Santos, H., Gordeliy, V., and Archer, M. (2014) X-ray structure of a CDP-alcohol phosphatidyltransferase membrane enzyme and insights into its catalytic mechanism. *Nat. Commun.* **5,** 4169
62. Lombard, J., López-García, P., and Moreira, D. (2012) Phylogenomic investigation of phospholipid synthesis in archaea. *Archaea* **2012,** 630910
63. Jain, S., Caforio, A., Fodran, P., Lolkema, J. S., Minnaard, A. J., and Driessen, A. J. (2014) Identification of CDP-archaeol synthase, a missing link of ether lipid biosynthesis in Archaea. *Chem. Biol.* **21,** 1392–1401
64. Nakatani, Y., Ribeiro, N., Streiff, S., Gotoh, M., Pozzi, G., Désaubry, L., and Milon, A. (2014) Search for the most 'primitive' membranes and their reinforcers: a review of the polyprenyl phosphates theory. *Orig. Life Evol. Biosph.* **44,** 197–208
65. Dibrova, D. V., Galperin, M. Y., and Mulkidjanian, A. Y. (2014) Phylogenomic reconstruction of archaeal fatty acid metabolism. *Environ. Microbiol.* **16,** 907–918
66. Treangen, T. J., and Rocha, E. P. (2011) Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. *PLoS Genet.* **7,** e1001284
67. Doolittle, W. F. (2000) Uprooting the tree of life. *Sci. Am.* **282,** 90–95
68. Kandler, O. (1998) in *Thermophiles: the Keys to Molecular Evolution and the Origin of Life?* (Wiegel, J., and Adams, W. W., eds) pp. 19–32, Taylor and Francis, Philadelphia
69. Williams, T. A., Foster, P. G., Cox, C. J., and Embley, T. M. (2013) An archaeal origin of eukaryotes supports only two primary domains of life. *Nature* **504,** 231–236
70. Sousa, F. L., Thiergart, T., Landan, G., Nelson-Sathi, S., Pereira, I. A., Allen, J. F., Lane, N., and Martin, W. F. (2013) Early bioenergetic evolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **368,** 20130088
71. Martin, W. F., Sousa, F. L., and Lane, N. (2014) Evolution. Energy at life's origin. *Science* **344,** 1092–1093