

A divergent archaeal member of the alkaline phosphatase binuclear metalloenzyme superfamily has phosphoglycerate mutase activity

David E. Graham, Huimin Xu, Robert H. White*

Department of Biochemistry, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0308, USA

Received 11 February 2002; revised 15 March 2002; accepted 20 March 2002

First published online 29 March 2002

Edited by Robert B. Russell

Abstract The hyperthermophilic archaeon *Methanococcus jannaschii* uses several non-canonical enzymes to catalyze conserved reactions in glycolysis and gluconeogenesis. A highly diverged gene from that organism has been proposed to function as a phosphoglycerate mutase. Like the canonical cofactor-independent phosphoglycerate mutase and other members of the binuclear metalloenzyme superfamily, this *M. jannaschii* protein has conserved nucleophilic serine and metal-binding residues. Yet the substrate-binding residues are not conserved. We show that the genes at *M. jannaschii* loci MJ0010 and MJ1612 encode thermostable enzymes with phosphoglycerate mutase activity. Phylogenetic analyses suggest that this gene family arose before the divergence of the archaeal lineage. © 2002 Published by Elsevier Science B.V. on behalf of the Federation of European Biochemical Societies.

Key words: Phosphoglycerate mutase; Glycolysis; Gluconeogenesis; Alkaline phosphatase superfamily; Archaea

1. Introduction

The Embden–Meyerhof–Parnas pathway of glycolysis and gluconeogenesis is one of the most evolutionarily conserved series of reactions [1–3]. Therefore, it was surprising that the first complete genome sequence of an archaeon, *Methanococcus jannaschii*, encoded no recognizable homologs of several enzymes in this pathway [4,5]. It was clear that this autotrophic methanogen had to synthesize carbohydrates by gluconeogenesis, and the relevant enzymatic activities had been demonstrated in *Methanococcus maripaludis* [6,7]. Yet, four crucial enzymes that were highly conserved in the Bacteria and the Eucarya could not be identified in *M. jannaschii*.

The glycolytic and gluconeogenic enzymes ‘missing’ from *M. jannaschii* were 6-phosphofructokinase, fructose-bisphos-

phatase, fructose-bisphosphate aldolase and phosphoglycerate mutase (PGAM) [5]. Subsequently, the three enzymes that produce or consume fructose-1,6-bisphosphate have been identified. A number of euryarchaea, including *M. jannaschii*, have an unusual ADP-dependent glucokinase and phosphofructokinase [8,9]. The fructose bisphosphatase enzyme from *M. jannaschii* has both fructose-1,6-bisphosphatase and inositol monophosphatase activity [10]. And, euryarchaea have a diverged class I fructose-1,6-bisphosphate aldolase [11]. However, no PGAM has been characterized in *M. jannaschii*.

Two non-homologous classes of enzymes catalyze the interconversion of D-2-phosphoglycerate and D-3-phosphoglycerate: 2,3-diphosphoglycerate cofactor-dependent PGAMs (PGAM-d) and 2,3-diphosphoglycerate independent PGAMs (PGAM-i). Members of the superfamily containing PGAM-d share a common phosphohistidine reaction intermediate [12], whereas PGAM-i proteins belong to a binuclear metalloenzyme superfamily whose members form a phosphoserine reaction intermediate [13]. These two versions of PGAM are structurally and mechanistically unrelated, but appear to be functionally interchangeable. Most organisms have homologs of only one class.

Computational analyses by Koonin et al. identified a candidate for the PGAM gene in *M. jannaschii* [14]. This predicted protein sequence, encoded by the *M. jannaschii* locus MJ1612, belongs to the same metalloenzyme superfamily that includes PGAM-i enzymes, alkaline phosphatases, phosphopentomutases, and sulfatases [15]. However, comparison of the MJ1612 protein with the crystal structure model of PGAM-i from *Geobacillus stearothermophilus* shows that only the metal-binding domain of PGAM-i is conserved in MJ1612. Members of the family including MJ1612 are missing numerous active site amino acids that interact with phosphoglycerate in the PGAM-i substrate-binding domain [16]. In protein sequence databases, the MJ1612 protein is currently identified as a phosphonopyruvate decarboxylase, an enzyme required for bialaphos biosynthesis in *Streptomyces hygroscopicus*. However, the true phosphonopyruvate decarboxylase from that organism is unrelated to MJ1612 [17].

To test the hypothesis that MJ1612 and its paralog (MJ0010) encode PGAMs, we have cloned and recombinantly expressed these proteins in *Escherichia coli*. Using ¹H-NMR to directly assay the isomerization reaction between D-2-phosphoglycerate and D-3-phosphoglycerate, we show that both proteins have thermophilic PGAM activity. The MJ1612 and MJ0010 enzymes exemplify a new archaeal class of enzymes with PGAM activity (aPGAM), which includes homologs in most archaea and diverse bacteria.

*Corresponding author. Fax: (1)-540-231 9070.
E-mail address: rhwhite@vt.edu. (R.H. White).

Abbreviations: PGAM, phosphoglycerate mutase (EC 5.4.2.1); PGAM-d, 2,3-diphosphoglycerate cofactor-dependent phosphoglycerate mutase; PGAM-i, cofactor-independent phosphoglycerate mutase; aPGAM, archaeal-type phosphoglycerate mutase; cDPG, cyclic-2,3-diphosphoglycerate; RubisCO, ribulose bisphosphate carboxylase/oxygenase

2. Materials and methods

2.1. Cloning and recombinant expression of MJ1612 and MJ0010

The MJ1612 gene was amplified by PCR from genomic DNA, cloned and recombinantly expressed by standard procedures [18] using a forward oligonucleotide primer 5'-GGTCATATGAAAAAGG-GAAAG-3' and a reverse primer 5'-GATGGATCCTTAAGCAC-CAAATTC-3' (Invitrogen). The MJ0010 was similarly amplified using forward primer 5'-GGTCATATGAGGGCTATTTAATTC-3' and reverse primer 5'-GATGGATCCTTATCCCTCTAA-TAAATC-3'. PCR products were digested at the primer-introduced *Nde*I and *Bam*HI sites. The MJ1612 gene was cloned into compatible sites in plasmid vector pET-19b (Novagen) and the MJ0010 gene was cloned into plasmid pT7-7 [19]. DNA sequences were verified by dye-terminator sequencing at the University of Iowa DNA facility. The resulting plasmids were transformed into *E. coli* BL21-Codon-Plus(DE3)-RIL (Stratagene) cells and recombinant protein expression was induced as described previously [18].

2.2. Preparation of heat-purified recombinant proteins

Cells expressing the recombinant MJ1612 protein (300 mg wet weight) were suspended in 3 ml of extraction buffer (50 mM TES/NaOH, 10 mM MgCl₂, 20 mM 2-mercaptoethanol, pH 7.0) and sonicated, followed by centrifugation (10 min, 16000×g, at room temperature). SDS-PAGE analysis of proteins in the resulting pellets and supernatants showed that most of the recombinant protein was present in the pellet of insoluble material. Heating the soluble proteins in the supernatant to 70°C for 10 min denatured most native *E. coli* proteins and these were removed by centrifugation (10 min, 16000×g, at room temperature). The resulting supernatants contained the desired recombinant protein. Similarly prepared extracts from *E. coli* host cells without the recombinant plasmid were used for control experiments. MJ0010 protein was prepared as described for the MJ1612 protein and was also poorly soluble.

2.3. ¹H-NMR analysis of PGAM activity

D-2-Phosphoglycerate trisodium salt (Sigma, 5.6 μmol, 2 mg) was dissolved in 1 ml of ²H₂O (99.8 atom% ²H) and 1.8 μmol of [2,2,3,3-²H₄]-3-trimethylsilylpropionate was added and the pH adjusted to 7.0. The sample was then divided into two equal portions. To one portion was added 20 μl of the heated *E. coli* extract containing MJ1612 enzyme and to the other portion was added 20 μl of heated *E. coli* extract without the recombinant protein. The samples were incubated for 10 min at 60°C and then placed in a 100°C water bath and reduced to a volume of ca. 30 μl by evaporation with a stream of nitrogen gas. The resulting liquid was dissolved in 0.5 ml of the ²H₂O. ¹H-NMR spectroscopy was performed on these samples using a 500 MHz JEOL Eclipse 500 NMR spectrometer. Activity of the MJ0010 enzyme was analyzed by the same method.

Known phosphoglycerate samples were purchased from Sigma and analyzed by ¹H-NMR under the same conditions. ¹H-NMR resonances for the trisodium salt of D-2-phosphoglyceric acid were (δ) 4.48 (1H, septet, *J*_{H-2→H-3} = 2.9 Hz, *J*_{H-2→H-3'} = 5.6 Hz, *J*_{H-2→P} = 8.4 Hz, H-2); 3.91 (1H, dd, *J*_{H-3→H-2} = 2.9 Hz, *J*_{H-3→H-3'} = 11.7 Hz, H-3); and 3.80 (1H, dd, *J*_{H-3'→H-2} = 5.6 Hz, *J*_{H-3'→H-3} = 11.7 Hz, H-3'). For the disodium salt of D-3-phosphoglyceric acid, resonances were (δ) 4.19 (1H, octet, *J*_{H-2→P} = 1.1 Hz, *J*_{H-2→H-3} = 2.8 Hz, *J*_{H-2→H-3'} = 6.0 Hz, H-2); 4.07 (1H, octet, *J*_{H-3→H-2} = 2.8 Hz, *J*_{H-3→P} = 11.7 Hz, *J*_{H-3→H-3'} = 11.0 Hz, H-3); and 3.95 (1H, octet, *J*_{H-3'→P} = 5.0 Hz, *J*_{H-3'→H-2} = 6.0 Hz, *J*_{H-3'→H-3} = 11.0 Hz, H-3').

2.4. Alignment of aPGAM protein sequences with homologs from the alkaline phosphatase superfamily

The translated sequence of *M. jannaschii* gene MJ1612 (sp|Q59007) was used to query the non-redundant protein database at the National Center for Biotechnology Information using the BLASTP program (v. 2.2.1) [20] with the BLOSSUM62 matrix and default gap costs for existence-11 and extension-1. Members of the aPGAM family were identified in the archaea *Aeropyrum pernix* (dbj|BAA80616.1), *Archaeoglobus fulgidus* (gb|AAB89499.1 and gb|AAB89821.1), *Methanobacterium thermoautotrophicum* (gb|AAB86064.1 and gb|AAB84924.1), *M. maripaludis* (gb|AAD47607.1), *M. jannaschii* (sp|Q60326), *Pyrobaculum aerophilum* (gb|AAL64112.1), *Pyrococcus horikoshii* (dbj|BAA29105.1), *Sulfolobus solfataricus* (gb|AAK40745.1) and *Thermoplasma acidophilum* (emb|CAC11555.1). Bacterial homo-

logs were found in *Aquifex aeolicus* (gb|AAC06791.1), *Deinococcus radiodurans* (gb|AAF11775.1), *S. hygroscopicus* (dbj|BAA93685.1), and *Thermotoga maritima* (gb|AAD36837.1). Eucaryal homologs were found in *Arabidopsis thaliana* (dbj|BAB02608.1) and *Dictyostelium discoideum* (cDNAs gb|AU071980.1, gb|AU061080.1 and gb|C93983.1).

Additional members of the aPGAM family were identified in incomplete genomic sequences of *Ferroplasma acidarmanus* and *Methanosarcina barkeri* from the US DOE Joint Genome Institute (<http://www.jgi.doe.gov>); *M. jannaschii* aPGAM sequences were aligned to the *G. stearothermophilus* PGAM-i sequence using the CLUSTALW program (v. 1.82) [22]. The structural alignment figure was created using the ESPript program (v2.0 beta) (<http://prodes.toulouse.inra.fr/ESPrpt/>). Alignments for phylogenetic analyses were created using the CLUSTALW program.

The structures of the *G. stearothermophilus* PGAM-i protein (pdb|1EQJ) and *E. coli* alkaline phosphatase (pdb|1B8J) were aligned using the FSSP database [21]. *M. jannaschii* aPGAM sequences were aligned to the *G. stearothermophilus* PGAM-i sequence using the CLUSTALW program (v. 1.82) [22]. The structural alignment figure was created using the ESPript program (v2.0 beta) (<http://prodes.toulouse.inra.fr/ESPrpt/>). Alignments for phylogenetic analyses were created using the CLUSTALW program.

2.5. Inference of the aPGAM phylogeny

From the alignment of 27 aPGAM protein sequences, ambiguously aligned positions near insertions or deletions were removed and the remaining 390 positions were deemed to be confidently aligned. The phylogeny of these genes was inferred using maximum likelihood criteria to evaluate 500 trees identified by maximum parsimony methods, as described previously [23]. Bootstrap probabilities were estimated using the resampling estimated log-likelihood method [24] with the 500 candidate trees. For comparison, the phylogeny was inferred from the same alignment using the neighbor-joining algorithm with

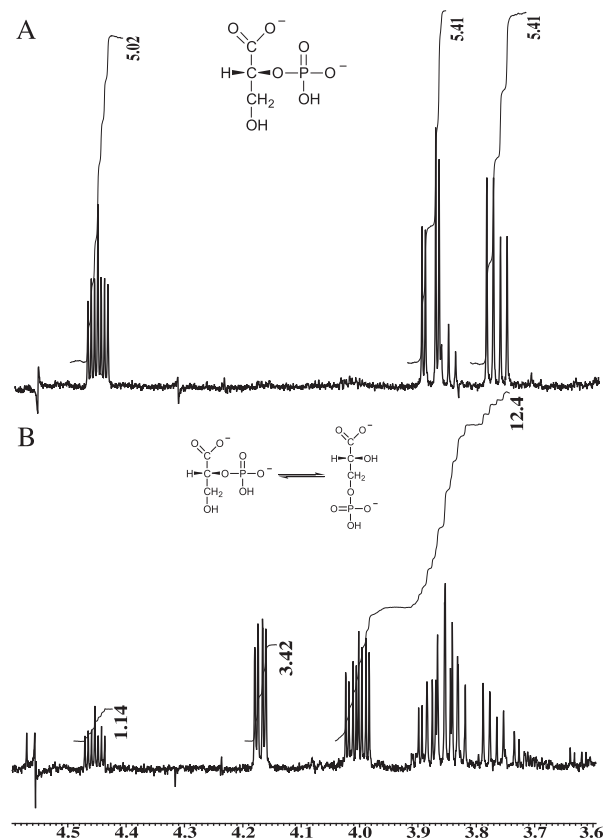


Fig. 1. ¹H-NMR (500 MHz) analysis of D-2-phosphoglycerate incubated with heated cell extract from *E. coli* (A) or heated cell extract from *E. coli* expressing recombinant MJ1612 protein (B). Resonance intensities were measured relative to the 3-trimethylsilylpropionate standard (not shown).

1000 bootstrap replicates, implemented by the Seqboot, ProtDist, Neighbor and Consense programs [PHYLP (phylogeny inference package), v. 3.6a2.1; J. Felsenstein, Department of Genetics, University of Washington, Seattle, 2001].

3. Results and discussion

3.1. Recombinant MJ0010 and MJ1612 have PGAM activity

Both the MJ0010 and MJ1612 proteins are expressed at high levels in an *E. coli* host; however, most of the recombinant protein is insoluble in cell-free extracts. Portions of the recombinant proteins that remain soluble resist denaturation by heating. Although *E. coli* has at least two PGAM genes [25], the native enzymes are not heat stable. The ¹H-NMR analysis in Fig. 1A shows that a sample of 2-phosphoglycerate incubated with heated *E. coli* extract contains no detectable resonances for 3-phosphoglycerate. This spectrum is identical

to that observed before the addition of the extract except for the buffer resonances (not shown). In contrast, 2-phosphoglycerate incubated for 10 min at 60°C in the presence of heated *E. coli* extract expressing recombinant MJ1612 enzyme forms 3-phosphoglycerate, shown in Fig. 1B. The decrease in the intensity of the 2-phosphoglycerate resonances corresponded to the observed increase in the intensity of the 3-phosphoglycerate resonances. This is most clearly observed from the changes in intensity of the H-2 resonances at (δ) 4.46 and 4.17 for the 2-phosphoglycerate and 3-phosphoglycerate, respectively. Addition of 3-phosphoglycerate to the sample increased only the intensity the 3-phosphoglycerate resonances. The conversion of 75% of the initial 2-phosphoglycerate to 3-phosphoglycerate is consistent with the thermodynamic equilibrium favoring 3-phosphoglycerate, as observed for canonical PGAMs [26]. Because the sum of the resonance intensities of the two isomers remained constant, we conclude

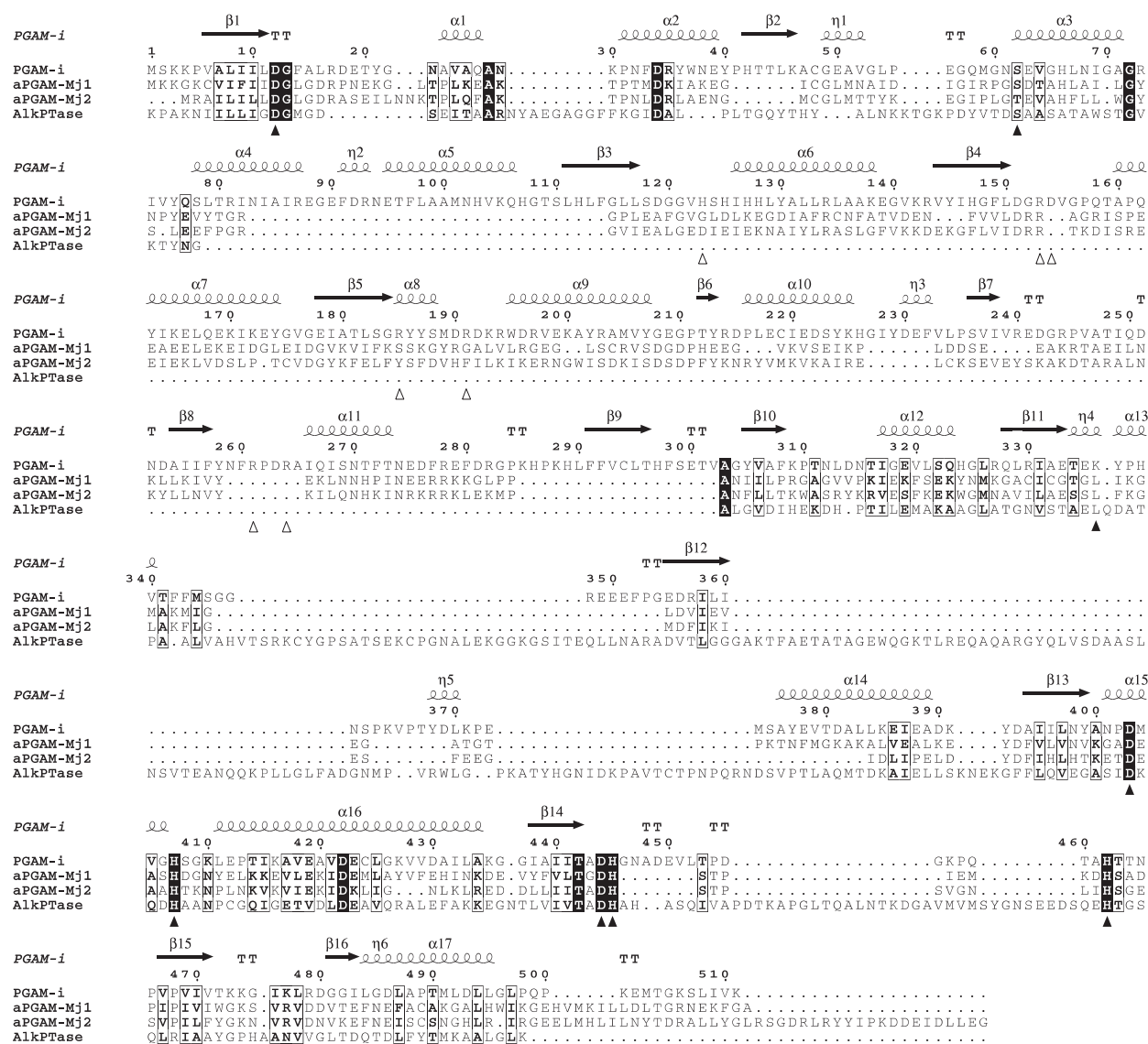


Fig. 2. Alignment of members of the alkaline phosphatase superfamily. PGAM-i from *G. stearothermophilus* is aligned with aPGAM paralogs from *M. jannaschii* MJ1612 (aPGAM-Mj1) and MJ0010 (aPGAM-Mj2) and alkaline phosphatase from *E. coli*. Positions of identically conserved residues are in inverted text and similarly conserved residues are outlined. The secondary structure of the *G. stearothermophilus* enzyme, derived from the crystal structure model [16], is indicated above the alignment as α-helices (α), 3₁₀-helices (η), β-strands (β) and β-turns (TT). Arrowheads below the alignment indicate active site residues identified in the *G. stearothermophilus* protein structure: closed arrowheads mark residues in the phosphatase domain and open arrowheads mark residues in the phosphotransferase domain [16].

that no additional products were generated. Under similar conditions, the MJ0010 protein converted 29% of the initial 2-phosphoglycerate to 3-phosphoglycerate.

3.2. Sequence alignment with the alkaline phosphatase superfamily

The crystal structure model of *G. stearothermophilus* PGAM-i revealed two domains: a catalytic phosphatase domain coordinates two Mn^{2+} ions (amino acids 1–76 and 311–511) and a substrate-binding domain forms hydrogen bonds with substrate and product (amino acids 77–310) [16]. The alignment in Fig. 2 shows that the metal-binding ligands and nucleophilic serine residue are conserved throughout the alkaline phosphatase superfamily. Although both PGAM-i and aPGAM sequences share a unique substrate-binding domain, aPGAM sequences lack the basic residues that interact with phosphoglycerate isomers. Therefore, the mechanism of substrate binding must be different in aPGAM proteins.

3.3. Phylogeny of the aPGAM family

Phylogenetic analyses of aPGAM and PGAM-i sequences confirm our expectation that these modular genes have been frequently exchanged across organismal lineages [27]. Nevertheless, there is consistency in the class of PGAM found within lineages. Bacterial members of the low %G+C Gram-positive group, Mycoplasmataceae, γ -proteobacteria, ϵ -proteobacteria and cyanobacteria all appear to have vertically

inherited PGAM-i genes (data not shown). Plants and several deeply branching eukaryotes also have PGAM-i homologs. Two euryarchaea, *M. barkeri* and *Halobacterium* sp. may have acquired PGAM-i homologs from bacteria. Most yeast, animals, high %G+C Gram-positive bacteria and some pathogenic microbes have PGAM-d homologs [28].

Members of the aPGAM family are found in every complete archaeal genome sequence, except for *Halobacterium* sp. (Fig. 3). Crenarchaea have one homolog each, whereas several euryarchaea (*Methanococcus* spp., *A. fulgidus* and *M. thermoautotrophicum*) have two paralogs. Microorganisms from some of the earliest diverged bacterial lineages, *A. aeolicus*, *T. maritima* and *D. radiodurans*, also have specifically related aPGAM homologs, suggesting that this gene arose early in microbial evolution.

The function of the two aPGAM paralogs in *M. jannaschii* is unknown, although vertebrates and *E. coli* differentially express PGAM isozymes [25,29]. D-3-Phosphoglycerate occupies a key role in *M. jannaschii* central metabolism not only in glycolysis and gluconeogenesis, but also as the precursor for serine biosynthesis [6] and the product of an unusual ribulose biphosphate carboxylase/oxygenase (RubisCO) [30]. *M. thermoautotrophicum* cells produce a polyanionic solute, cyclic-2,3-diphosphoglycerate (cDPG), that is derived from 2-phosphoglycerate [31]. Although *M. jannaschii* lacks the cDPG synthase enzyme required to make cDPG, it has a functional homolog of the 2-phosphoglycerate kinase required to make the 2,3-diphosphoglycerate precursor (White, unpublished data).

The experiments described in this paper characterize the ‘missing’ PGAM activity in *M. jannaschii*, but reiterate several outstanding questions about *M. jannaschii* central metabolism: what are the functions of RubisCO, 2-phosphoglycerate kinase and two PGAM paralogs in this organism?

Acknowledgements: We thank Tom Glass for assistance with NMR analysis. This work was supported by U.S. National Science Foundation Grant MCB 9985712 to R.H.W.

References

- [1] Selig, M., Xavier, K.B., Santos, H. and Schönheit, P. (1997) Arch. Microbiol. 167, 217–232.
- [2] Fothergill-Gilmore, L.A. (1986) Trends Biochem. Sci. 11, 47–51.
- [3] Dandekar, T., Schuster, S., Snel, B., Huynen, M. and Bork, P. (1999) Biochem. J. 343, 115–124.
- [4] Bult, C.J. et al. (1996) Science 273, 1017–1140.
- [5] Selkov, E., Maltsev, N., Olsen, G.J., Overbeek, R. and Whitman, W.B. (1997) Gene 197, GC11–GC26.
- [6] Simpson, P.G. and Whitman, W.B. (1993) in: Methanogenesis: Ecology, Physiology Biochemistry and Genetics (Ferry, J.G., Ed.), pp. 445–472, Chapman and Hall, New York.
- [7] Yu, J.-P., Ladapo, J. and Whitman, W.B. (1994) J. Bacteriol. 176, 325–332.
- [8] Verhees, C.H., Tuininga, J.E., Kengen, S.W., Stams, A.J., van der Oost, J. and de Vos, W.M. (2001) J. Bacteriol. 183, 7145–7153.
- [9] Sakuraba, H., Yoshioka, I., Koga, S., Takahashi, M., Kitahama, Y., Satomura, T., Kawakami, R. and Ohshima, T. (2002) J. Biol. Chem., published online as manuscript C200059200.
- [10] Stec, B., Yang, H., Johnson, K.A., Chen, L. and Roberts, M.F. (2000) Nat. Struct. Biol. 7, 1046–1050.
- [11] Siebers, B., Brinkmann, H., Dörr, C., Tjaden, B., Lilie, H., van der Oost, J. and Verhees, C.H. (2001) J. Biol. Chem. 276, 28710–28718.
- [12] Jedrzejewski, M.J. (2000) Prog. Biophys. Mol. Biol. 73, 263–287.

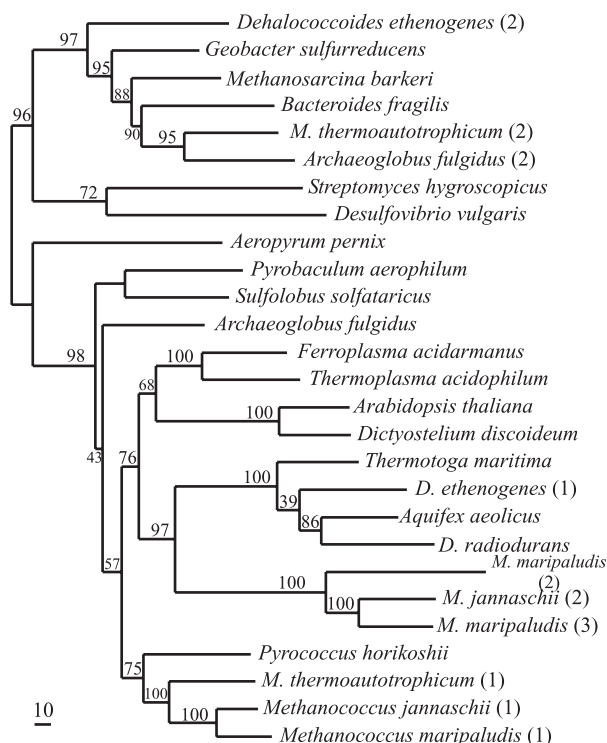


Fig. 3. Phylogeny of aPGAM homologs. Paralogous genes are distinguished by numbers placed after the organism name. Estimated bootstrap percentages are indicated for significantly supported branches. The phylogeny of the same sequences inferred by the neighbor-joining method is substantially similar to this tree. The two trees differ in their ordering of several branches with low bootstrap support, particularly the crenarchaeal and *Thermoplasma*/eukaryote branches. The scale bar represents 10 amino acid replacements per 100 positions.

- [13] Jedrzejewski, M.J. and Setlow, P. (2001) *Chem. Rev.* 101, 607–618.
- [14] Koonin, E.V., Mushegian, A.R., Galperin, M.Y. and Walker, D.R. (1997) *Mol. Microbiol.* 25, 619–637.
- [15] Galperin, M.Y., Bairoch, A. and Koonin, E.V. (1998) *Protein Sci.* 7, 1829–1835.
- [16] Jedrzejewski, M.J., Chander, M., Setlow, P. and Krishnasamy, G. (2000) *J. Biol. Chem.* 275, 23146–23153.
- [17] Nakashita, H., Kozuka, K., Hidaka, T., Hara, O. and Seto, H. (2000) *Biochim. Biophys. Acta* 1490, 159–162.
- [18] Graupner, M., Xu, H. and White, R.H. (2000) *J. Bacteriol.* 182, 4862–4867.
- [19] Tabor, S. and Richardson, C.C. (1985) *Proc. Natl. Acad. Sci. USA* 82, 1074–1078.
- [20] Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) *Nucleic Acids Res.* 25, 3389–3402.
- [21] Holm, L. and Sander, C. (1996) *Science* 273, 595–603.
- [22] Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) *Nucleic Acids Res.* 22, 4673–4680.
- [23] Bunjun, S. et al. (2000) *Proc. Natl. Acad. Sci. USA* 97, 12997–13002.
- [24] Kishino, H., Miyata, T. and Hasegawa, M. (1990) *J. Mol. Evol.* 31, 151–160.
- [25] Fraser, H.I., Kvaratskhelia, M. and White, M.F. (1999) *FEBS Lett.* 455, 344–348.
- [26] Grisolia, S. (1962) *Methods Enzymol.* 5, 236–242.
- [27] Woese, C.R. (2000) *Proc. Natl. Acad. Sci. USA* 97, 8392–8396.
- [28] Rigden, D.J., Bagyan, I., Lamani, E., Setlow, P. and Jedrzejewski, M.J. (2001) *Protein Sci.* 10, 1835–1846.
- [29] Zhang, J. et al. (2001) *Gene* 264, 273–279.
- [30] Watson, G.M.F., Yu, J.-P. and Tabita, F.R. (1999) *J. Bacteriol.* 181, 1569–1575.
- [31] Evans, J.N., Tolman, C.J., Kanodia, S. and Roberts, M.F. (1985) *Biochemistry* 24, 5693–5698.