JOURNAL OF BACTERIOLOGY, Nov. 2001, p. 6714–6716 Vol. 183, No. 22
0021-9193/01/$04.00+0 DOI: 10.1128/JB.183.22.6714–6716.2001

Presence of Prokaryotic and Eukaryotic Species in All Subgroups of the PP_i-Dependent Group II Phosphofructokinase Protein Family

MIKLÓS MÜLLER,[1]* JENNIFER A. LEE,[1] PAUL GORDON,[2] TERRY GAASTERLAND,[1] AND CHRISTOPH W. SENSEN[3]

The Rockefeller University, New York, New York 10021[1]; National Research Council of Canada, Institute for Marine Biosciences, Halifax, Nova Scotia, Canada B3H 4H7[2]; and University of Calgary, Faculty of Medicine, Department of Biochemistry and Molecular Biology, Calgary, Alberta, Canada T2N 4N1[3]

Received 14 May 2001/Accepted 24 August 2001

Inorganic pyrophosphate-dependent phosphofructokinase (PP_i-PFK) of the amitochondriate eukaryote *Mastigamoeba balamuthi* was sequenced and showed about 60% identity to PP_i-PFKs from two eubacteria, *Propionibacterium freudenreichii* and *Sinorhizobium meliloti*. These gene products represent a newly recognized lineage of PFKs. All four lineages of group II PFKs, as defined by phylogenetic analysis, contained both prokaryotic and eukaryotic species, underlining the complex evolutionary history of this enzyme.

We have recently extended our studies on glycolytic enzymes of parasitic amitochondriate eukaryotes (20) to the free-living *Mastigamoeba balamuthi* (ATCC 30984) (6) with the goal of comparing the metabolic properties of anaerobic and microaerophilic eukaryotes with dramatically different life styles. This species belongs to the pelobionts, a group of amitochondriate amoeboflagellate protists of uncertain evolutionary position (5, 25). We noted that the sequence of its phosphofructokinase (PFK) showed unexpected characteristics, prompting us to revisit the taxonomic distribution and relationships of various PFK types.

Type A PFK, an enzyme of the glycolytic pathway, phosphorylates fructose 6-phosphate to fructose 1,6-bisphosphate. In most organisms, ATP is the phosphoryl donor (ATP-PFK; EC 2.7.1.11) of the irreversible reaction. A number of protists and plants and some eubacteria contain reversible PFKs, which use inorganic pyrophosphate (PP_i) instead of ATP (PP_i-PFK; EC 2.7.1.90). The assumed significance of PP_i as the phosphoryl donor is reflected in an increase of the ATP yield during glycolysis (16, 26). This notion is supported by the predominant occurrence of PP_i-PFK in organisms living in hypoxic or anoxic environments, which rely on anaerobic glycolysis (17).

The evolutionary history of PFK does not coincide with accepted notions of organismic relationships and points to past gene duplications and lateral gene transfers. Based primarily on sequence characteristics, type A PFKs are currently assigned to three major groups (groups I, II, and III) (22). Group II can be further subdivided into four subgroups, which appear as robust clades in phylogenetic reconstructions (we are using the tentative nomenclature proposed for clades in group II [18]). Closely related organisms may contain close homologs of PFK which use different phosphoryl donors, indicating that enzyme specificity can change relatively easily (2), a conclusion recently confirmed experimentally (7). Some organisms even contain members of two such subgroups (8, 10–12).

* Corresponding author. Mailing address: The Rockefeller University, 1230 York Ave., New York, NY 10021. Phone: (212) 327-8153. Fax: (212) 327-7974. E-mail: mmuller@mail.rockefeller.edu.

Sequence of *M. balamuthi* PP_i-PFK. A random clone of an *M. balamuthi* cDNA library (M. Müller et al., unpublished data), sequenced on both strands, contained a G+C-rich (69.3%) open reading frame putatively encoding a PP_i-PFK of 410 amino acids with an M_r of 44,200. The conceptual translation showed that only 36 codons were used and that 98.2% of the nucleotides in the third position of each codon were G or C. This skewed codon usage, characteristic of protein-encoding genes of this organism (Müller et al., unpublished data), and the presence of a typical eukaryotic poly(A) tail support the origin of this message from authentic *M. balamuthi* DNA. The conceptual translation showed complete colinearity and about 60% amino acid identity to the PP_i-PFK of the gram-positive organism *Propionibacterium freudenreichii* and to the putative PP_i-PFK of the α-proteobacterium, *Sinorhizobium meliloti* (3, 14). The *M. balamuthi* sequence has been deposited in GenBank under accession number AAF70463.

Phylogenetic analysis. Sequences of PP_i-PFK homologs were retrieved from the National Center for Biotechnology Information protein database. The *S. meliloti* PFK sequence was retrieved from the website of the corresponding genome project (http://sequence.toulouse.inra.fr/meliloti.html) (3). Sequence sampling for group II encompasses the whole diversity present in the nonredundant National Center for Biotechnology Information database. Sampling of group III was restricted to a few species to provide an outgroup. Group I enzymes were not considered. The alignment was performed using the CLUSTAL X program (23) and adjusted visually. Phylogenetic reconstruction was performed with a maximum-likelihood method (PROTML) (1) on 167 shared amino-terminal amino acid positions. Bootstrap proportions were calculated by a resampling of the estimated log likelihood (RELL) values from the maximum-likelihood method (1). Neighbor joining and maximum-parsimony analyses revealed identical groups and subgroups (data not shown).

Evolutionary relationships of group II and III PFK sequences. The phylogenetic tree obtained shows five major clades (Fig. 1). Four clades belong to group II, and one represents group III. The internal branches connecting these clades are relatively long, and the branching pattern is sup-

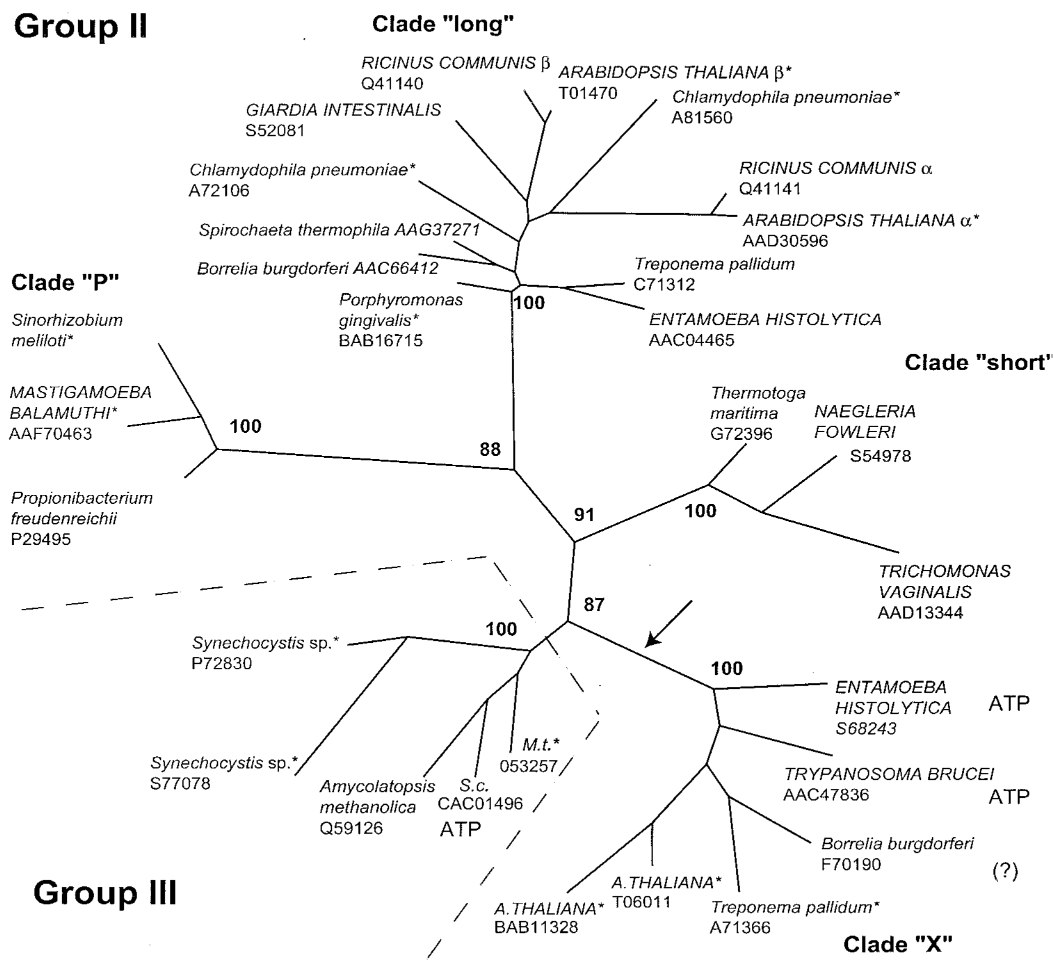

FIG. 1. Unrooted phylogenetic tree of group II and III PP_i-PFKs based on 167 shared amino-terminal residues (between residues 2 and 204 in the *P. freudenreichii* sequence). Taxon selection extended over the complete diversity present in available databases. The arrow denotes a proposed gene duplication. Close paralogs among chlamydiae, actinomycetales, and plants were omitted for clarity. Database accession numbers follow the species names. For the *S. meliloti* sequence, see http://sequence.toulouse.inra.fr/meliloti.html. Eukaryotic species names are in capital letters, and eubacterial species names are in lowercase letters. *A.*, *Arabidopsis*; *M.t.*, *Mycobacterium tuberculosis*; *S.c.*, *Streptomyces coelicolor*. PP_i specificity was established biochemically for unmarked species; asterisks mark species not explored biochemically, "ATP" indicates enzymes with ATP specificity, and the question mark indicates a species for which the heterologously expressed enzyme showed no PFK activity with either PP_i or ATP (9). For the long clade plant enzymes, α is the regulatory subunit and β is the catalytic one. The maximum-likelihood method with local rearrangements was utilized. Numbers at the nodes give RELL bootstrap values (see text).

ported by high bootstrap values. The order of separation of lineages within these clades has not been resolved, however. While this subdivision is based only on the easily aligned amino-terminal parts of the sequences, a global comparison of the complete sequences and an analysis of the lengths and positions of insertions and deletions result in an identical subdivision (18).

The data clearly show that each of the four clades of group II PFKs encompasses both eubacterial and eukaryotic species. Eubacterial species in each clade represent diverse lineages, but the overall taxonomic diversity of group II enzymes is rather limited. Clade X contains sequences from several species that are also present in the long clade. So far, no archaebacterial sequences were noted in PFK group II. While the products of a number of sequences analyzed remain to be studied, available information indicates that the long, short, and P (new designation) clades correspond to PP_i-specific enzymes (10, 18) and that clade X contains ATP-specific PFKs (8, 19).

Since only the amino-terminal half of the molecule could be confidently aligned across all groups, the information available for phylogenetic analysis was limited. The robust separation of major groups and clades, however, seems to be sufficient to suggest two major events in the evolution of the enzymes of group II. The first event was probably a gene duplication that separated clade X and the lineage leading to the three other clades (Fig. 1). This duplication was probably accompanied by a change of enzyme specificity in one of the branches. The substrate specificity of the common ancestor remains to be established, however. The second set of events led to differentiation between clades long, short, and P by marked changes in the overall sequence structure but without further change in phosphoryl donor specificity.

The presence of two PFK genes for the chlamydial and plant enzymes within the long clade reveals further gene duplications within this clade. In plants, gene duplication led to the emergence of catalytic (β) and regulatory (α) subunits (4, 24). A similar duplication and functional change also occurred in group I, which contains the classical ATP-linked enzymes (21). The functional significance of the chlamydial paralogs remains unknown.

The limited and peculiar taxonomic distribution of group II sequences makes a coherent reconstruction of events leading to the observed phenomena a daunting task. The relationships seen in the phylogenetic reconstruction do not coincide with accepted organismic relationships. One must account for the presence of both eubacteria and eukaryotes in each of the four clades of group II PFK genes as well as for the existence of sequences from the same organisms that fall into separate clades. While both early gene duplications and subsequent differential losses (15) and lateral gene transfers (13) have probably contributed to the current picture, only a significantly larger taxonomic sampling and functional characterization of the proteins encoded will permit a convincing reconstruction of the peculiar history of PP_i-PFK homologs.

We thank Gordona Bothe (GATC-Biotechnology GMBH, Constance, Germany) for providing the *S. meliloti* sequence before its publication, Frederick Schuster (Brooklyn College, City University of New York, Brooklyn) for the *M. balamuthi* culture, Rama K. Singh and his team (NRC, Institute for Marine Biosciences, Halifax, Canada) for the DNA sequencing, Robert Kemp (The Chicago Medical School, North Chicago, Ill.) for suggestions and permission to refer to his paper before its publication, and William Martin (Heinrich-Heine Universität, Düsseldorf, Germany) and Lidya B. Sanchez and Dorothy V. Moore of the New York laboratory for comments on the manuscript.

This research was supported by grants to M.M. from the National Science Foundation (MCB9615659) and the National Institutes of Health (AI 11942).

REFERENCES

1. **Adachi, J., and M. Hasegawa.** 1996. MOLPHY version 2.3: programs for molecular phylogenetics based on maximum likelihood. Computer Science Monographs, no. 28, Institute of Statistical Mathematics, Tokyo, Japan.
2. **Alves, A. M. C. R., G. J. W. Euverink, M. J. Bibb, and L. Dijkhuizen.** 1997. Identification of ATP-dependent phosphofructokinase as a regulatory step in the glycolytic pathway of the actinomycete *Streptomyces coelicolor* A3(2). Appl. Environ. Microbiol. **63:**956–961.
3. **Capela, D., F. Barloy-Huber, J. Gouzy, G. Bothe, F. Ampe, J. Batut, P. Boistard, A. Becker, M. Boutry, E. Cadieu, S. Dréano, S. Gloux, T. Godrie, A. Goffeau, D. Kahn, E. Kiss, V. Lelaure, D. Masuy, T. Pohl, D. Portetelle, B. Purnelle, U. Ramsperger, C. Renard, F. Galibert, M. Vandenbol, S. Weidner, and F. Galibert.** 2001. Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. Proc. Natl. Acad. Sci. USA **98:**9877–9882.
4. **Carlisle, S. M., S. D. Blakele, S. M. Hemmingsen, S. J. Trevanion, T. Hiyoshi, N. J. Kruger, and D. T. Dennis.** 1990. Pyrophosphate-dependent phosphofructokinase. Conservation of protein sequence between the α- and β-subunits and with the ATP-dependent phosphofructokinase. J. Biol. Chem. **265:**18366–18371.
5. **Cavalier-Smith, T.** 1991. Archamoebae: the ancestral eukaryotes? BioSystems **25:**25–38.
6. **Chavez, L. A., W. Balamuth, and T. Gong.** 1986. A light and electron microscopical study of a new, polymorphic free-living amoeba, *Phreatamoeba balamuthi* n.g., n. sp. J. Protozool. **33:**397–404.
7. **Chi, A., and R. G. Kemp.** 2000. The primordial high energy compound: ATP or inorganic pyrophosphate. J. Biol. Chem. **275:**35677–35679.
8. **Chi, A. S., Z. Deng, R. A. Albach, and R. G. Kemp.** 2001. The two phosphofructokinase gene products of *Entamoeba histolytica*. J. Biol. Chem. **276:** 19974–19981.
9. **Deng, Z., D. Roberts, X. Wang, and R. G. Kemp.** 1999. Expression, characterization, and crystallization of the pyrophosphate-dependent phosphofructo-1-kinase of *Borrelia burgdorferi*. Arch. Biochem. Biophys. **371:**326–331.
10. **Ding, Y.-H. R., R. S. Ronimus, and H. W. Morgan.** 2001. *Thermotoga maritima* phosphofructokinases: expression and characterization of two unique enzymes. J. Bacteriol. **183:**791–794.
11. **Fraser, C. M., S. Casjens, W. M. Huang, G. G. Sutton, R. Clayton, R. Lathigra, O. White, K. A. Ketchum, R. Dodson, E. K. Hickey, M. Gwinn, B. Dougherty, J.-F. Tomb, R. D. Fleischmann, D. Richardson, J. Peterson, A. R. Kerlavage, J. Quackenbush, S. Salzberg, M. Hanson, R. van Vugt, N. Palmer, M. D. Adams, J. Gocayne, J. Weidman, T. Utterback, L. Whattey, L. McDonald, P. Artiach, P. Bowman, S. Garland, C. Fujii, M. D. Cotton, K. Horst, K. Roberts, B. Hatch, H. O. Smith, and J. C. Venter.** 1997. Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. Nature **390:** 580–586.
12. **Fraser, C. M., S. J. Norris, G. M. Weinstock, O. White, G. G. Sutton, R. Dodson, M. Gwinn, E. K. Hickey, R. Clayton, K. A. Ketchum, E. Sodergren, J. M. Hardham, M. P. McLeod, S. Salzberg, J. Peterson, H. Khalak, D. Richardson, J. K. Howell, M. Chidambaram, T. Utterback, L. McDonald, P. Artiach, C. Bowman, M. D. Cotton, C. Fujii, S. Garland, B. Hatch, K. Horst, K. Roberts, M. Sandusky, J. Weidman, H. O. Smith, and J. C. Venter.** 1998. Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. Science **281:**375–388.
13. **Groisman, F. A., and H. Ochman.** 1996. Pathogenicity islands: bacterial evolution in quantum leaps. Cell **87:**791–794.
14. **Ladror, U., L. Gollapudi, R. L. Tripathi, S. L. Latshaw, and R. G. Kemp.** 1991. Cloning, sequencing and expression of pyrophosphate-dependent phosphofructokinase from *Propionibacterium freudenreichii*. J. Biol. Chem. **266:**16550–16555.
15. **Martin, W., and C. Schnarrenberger.** 1997. The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: a case study of functional redundancy in ancient pathways through endosymbiosis. Curr. Genet. **32:**1–18.
16. **Mertens, E.** 1991. Pyrophosphate-dependent phosphofructokinase, an anaerobic glycolytic enzyme? FEBS Lett. **285:**1–5.
17. **Mertens, E.** 1993. ATP versus pyrophosphate: glycolysis revisited in parasitic protists. Parasitol. Today **9:**122–126.
18. **Mertens, E., U. S. Ladror, J. A. Lee, A. Miretsky, A. Morris, C. Rozario, R. G. Kemp, and M. Müller.** 1998. The pyrophosphate-dependent phosphofructokinase of the protist, *Trichomonas vaginalis*, and the evolutionary relationships of protist phosphofructokinases. J. Mol. Evol. **47:**739–750.
19. **Michels, P. A. M., N. Chevalier, F. R. Opperdoes, M. H. Rider, and D. Rigden.** 1997. The glycosomal ATP-dependent phosphofructokinase of *Trypanosoma brucei* must have evolved from an ancestral pyrophosphate-dependent enzyme. Eur. J. Biochem. **250:**698–704.
20. **Müller, M.** 1998. Enzymes and compartmentation of core energy metabolism of anaerobic protists—a special case in eukaryotic evolution?, p. 109–131. *In* G. H. Coombs, K. Vickerman, M. A. Sleigh, and A. Warren (ed.), Evolutionary relationships among protozoa. Kluwer, Dordrecht, The Netherlands.
21. **Poorman, R. A., A. Randolph, R. G. Kemp, and R. L. Heinrickson.** 1984. Evolution of phosphofructokinase—gene duplication and creation of new effector sites. Nature **309:**467–469.
22. **Siebers, B., H.-P. Klenk, and R. Hensel.** 1998. PP_i-dependent phosphofructokinase from *Thermoproteus tenax*, an archeal descendant of an ancient line in phosphofructokinase evolution. J. Bacteriol. **180:**2137–2143.
23. **Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins.** 2000. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25:**4876–4882.
24. **Todd, J. F., S. D. Blakeley, and D. T. Dennis.** 1995. Structure of the genes encoding the α- and β-subunits of castor pyrophosphate-dependent phosphofructokinase. Gene **152:**181–186.
25. **Walker, G., A. G. B. Simpson, V. Edgcomb, M. L. Sogin, and D. J. Patterson.** 2001. Ultrastructural identities of *Mastigamoeba punctachora, Mastigamoeba simplex* and *Mastigella commutans* and assessment of hypotheses of relatedness of the pelobionts (Protista). Eur. J. Protistol. **37:**25–49.
26. **Wood, H. G.** 1977. Some reactions in which inorganic pyrophosphate replaces ATP and serves as a source of energy. Fed. Proc. **36:**2197–2205.