J Mol Evol (1998) 47:739–750

JOURNAL OF MOLECULAR EVOLUTION

© Springer-Verlag New York Inc. 1998

The Pyrophosphate-Dependent Phosphofructokinase of the Protist, *Trichomonas vaginalis,* and the Evolutionary Relationships of Protist Phosphofructokinases

Emmanuel Mertens,[1,*] Uri S. Ladror,[2,†] Jennifer A. Lee,[1] Anya Miretsky,[1] Andrea Morris,[1,‡] Catherine Rozario,[1,§] Robert G. Kemp,[2] Miklós Müller[1]

[1] The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA
[2] Department of Biological Chemistry, Finch University of Health Sciences/The Chicago Medical School, North Chicago, IL 60064, USA

Received: 5 December 1997 / Accepted: 18 March 1998

Abstract. The pyrophosphate-dependent phosphofructokinase (PP_i-PFK) of the amitochondriate protist *Trichomonas vaginalis* has been purified. The enzyme is a homotetramer of about 50 kDa subunits and is not subject to allosteric regulation. The protein was fragmented and a number of peptides were sequenced. Based on this information a PCR product was obtained from *T. vaginalis* gDNA and used to isolate corresponding cDNA and gDNA clones. Southern analysis indicated the presence of five genes. One open reading frame (ORF) was completely sequenced and for two others the 5′ half of the gene was determined. The sequences were highly similar. The complete ORF corresponded to a polypeptide of about 46 kDa. All the peptide sequences obtained were present in the derived sequences. The complete ORF was highly similar to that of other PFKs, primarily in its amino-terminal half. The *T. vaginalis* enzyme was most similar to PP_i-PFK of the mitochondriate heterolobosean, *Naegleria fowleri.* Most of the residues shown or assumed to be involved in substrate binding in other PP_i-PFKs were conserved in the *T. vaginalis* enzyme. Direct comparison and phylogenetic reconstruction revealed a significant divergence among PP_i-PFKs and related enzymes, which can be assigned to at least four distantly related groups, three of which contain enzymes of protists. The separation of these groups is supported with a high percentage of bootstrap proportions. The short *T. vaginalis* PFK shares a most recent common ancestor with the enzyme from *N. fowleri.* This pair is clearly separated from a group comprising the long (>60-kDa) enzymes from *Giardia lamblia, Entamoeba histolytica pfk2,* the spirochaetes *Borrelia burgdorferi* and *Trepomena pallidum,* as well as the α- and β-subunits of plant PP_i-PFKs. The third group (''X'') containing protist sequences includes the glycosomal ATP-PFK of *Trypanosoma brucei, E. histolytica pfk1,* and a second sequence from *B. burgdorferi.* The fourth group (''Y'') comprises cyanobacterial and high-G + C, Gram-positive eubacterial sequences. The well-studied PP_i-PFK of *Propionibacterium freudenreichii* is highly divergent and cannot be assigned to any of these groups. These four groups are well separated from typical ATP-PFKs, the phylogenetic analysis of which confirmed relationships established earlier. These findings indicate a complex

* *Present address:* SmithKline Beecham Biologicals, 89 Rue de l'Institut, 1325 Rixensart, Belgium
† *Present address:* Department 464, Building AP10, Abbott Laboratories, Abbott Park, IL 20009, USA
‡ *Present address:* Department of Molecular Biology, Princeton University, Princeton, NJ 08544, USA
§ *Present address:* 140-71 34th Avenue, Flushing, NY 11354, USA

GeneBank database accession numbers of new nucleotide sequences: *Tvpfk1* (combination of a gDNA sequence through residue E210 in the putative translation and a cDNA sequence from residue G14 in the putative translation—through the 3′ end), AF044973; *Tvpfk2* and *Tvpfk3* (partial 5′ gDNA sequences through residue E210 in the putative translation), AF053370 and AF053371, respectively. Noncoding regions not discussed here are included in the entries.

Correspondence to: M. Müller; *e-mail:* mmuller@rockvax.rockefeller.edu

history of a key step of glycolysis in protists with several early gene duplications and possible horizontal gene transfers.

Key words: Amitochondriate protist — Enzyme evolution — Glycolysis — Parabasala — Phosphofructokinase — PP_i — *Trichomonas*

Introduction

Phosphofructokinase (fructose 6-phosphate 1-phosphotransferase; PFK) phosphorylates fructose-6-phosphate to fructose-1,6-bisphosphate and thus occupies a key position in glycolysis (Fothergill-Gilmore and Michels 1993). The two major types of this enzyme differ in phosphoryl donor specificity. The more widespread ATP-PFK (EC 2.7.1.11) utilizes ATP and catalyzes an irreversible catabolic reaction. In a number of organisms, however, the phosphoryl donor is inorganic pyrophosphate (PP_i). PP_i-PFKs (EC 2.7.1.90) catalyze a reversible reaction of low-free energy change and can function in both glycolysis and gluconeogenesis. The use of PP_i instead of ATP in the phosphorylation of fructose-6-phosphate probably increases the efficiency of glycolysis and can be of importance to organisms which lack electron transport-linked phosphorylation and rely on glycolysis as a major source of energy (Wood 1977; Mertens 1991, 1993). This is the case for most amitochondriate (Müller 1988; Mertens 1993) and some mitochondriate (Peng and Mansour 1992; Denton et al. 1994, 1996) protists and probably for green plants under anaerobic conditions (Mertens 1991).

Recent studies indicate that all PFKs belong to the same kinase superfamily and are probably descendants of one ancestral protein (Wu et al. 1991). This superfamily encompasses at least two families (PFK-A and PFK-B). The PFK-A family contains most PFKs, both ATP- and PP_i-linked enzymes. Its members are diverse in subunit size, number of subunits in the mature protein, and susceptibility to various effectors, primarily adenine nucleotides and fructose-2,6-bisphosphate (Fothergill-Gilmore and Michels 1993). The PFK-B family comprises the minor PFK of *Escherichia coli* and a number of kinases that phosphorylate other sugar phosphates (Wu et al. 1991). Sequence comparisons and phylogenetic reconstruction indicate that all members of the PFK-A family form a monophyletic group. The PP_i-dependent and ATP-linked enzymes were found to separate into two well-defined subgroups (Ladror et al. 1991; Alves et al. 1996), respectively, although some recently explored enzymes defy a clear assignment. The PP_i-dependent group has been classified into two types based on molecular size and susceptibility to fructose-2,6-bisphosphate (Mertens 1991, 1993; Li and Phillips 1995). Type I enzymes are found in eubacteria and various protists, are homomeric, and are not affected by fructose-2,6-bisphosphate. Type II enzymes are present in plants (Carlisle et al. 1990; Todd et al. 1995) and also in the protist, *Euglena gracilis* (Miyatake et al. 1986), and consist of catalytic and regulatory subunits. The latter probably arose through a duplication and subsequent modification of the gene coding for the catalytic subunit and are responsible for the allosteric effect of fructose-2,6-bisphosphate on the enzyme (Poorman et al. 1984).

PP_i-PFKs have recently been characterized from several protists (Mertens et al. 1989, 1993; Mertens 1990; Peng and Mansour 1992; Denton et al. 1994; Phillips and Li 1995; Li and Phillips 1995; Deng et al. 1998; reviewed by Mertens 1993). These enzymes are not regulated by fructose-2,6-bisphosphate and thus can be assigned to Type I. However, they show an unexpectedly high level of divergence in their size and subunit structure. Sequences determined recently from several protists (Rozario et al. 1995; Huang et al. 1995; Wessberg et al. 1995; Bruchhaus et al. 1996; Michels et al. 1997; Deng et al. 1998) and compared with new data for several eubacteria confirm this high diversity and raise interesting questions concerning the evolutionary history of the homomeric PFKs.

The present paper describes the purification and primary structure of the PP_i-PFK (Mertens et al. 1989) of *Trichomonas vaginalis*, an amitochondriate parabasalid protist. The enzyme was found to be related to other PFKs, with the closest relationship to the PP_i-PFK of *Naegleria fowleri*, a heterolobosean (Wessberg et al. 1995), a relationship already apparent on earlier tentative phylogenetic reconstruction based on a partial sequence (Müller 1998). These two enzymes comprise a well-defined PP_i-PFK subfamily. Comparison of all available PP_i-PFK sequences reveals the existence of at least four subfamilies, three of which contain enzymes from protists.

Materials and Methods

Organism. Trichomonas vaginalis (strain NIH-Cl; ATCC 30001) was cultured in tryptose–yeast extract–maltose medium, pH 6.4, supplemented with 10% heat-inactivated horse serum (Diamond 1957). Cells from late-logarithmic phase cultures were collected with centrifugation, washed with phosphate-buffered saline, pH 6.4 (Doran 1957), and stored at −70°C.

Enzyme Purification. All purification steps were performed at 4°C. A frozen pellet from a 10-L culture was homogenized in 50 m*M* Tris buffer, pH 7.5, containing 5 μg/ml leupeptin, with a Potter–Elvejhem device. The homogenate (about 250 ml) was centrifuged for 10 min at 20,000*g*. To the supernatant, polyethyleneglycol flakes were added (to 10%, w/v). After dissolution and 30 min of standing, the solution was centrifuged as above and polyethylene glycol was again added to a 20% (w/v) final concentration. After 30 min of standing, the solution was centrifuged as before and the resulting pellet was resuspended in 25 m*M* histidine buffer, pH 6.0, containing 5 μg/ml leupeptin (Buffer A).

The solution (about 100 ml) was loaded on a QAE-Sepharose

(Pharmacia, Piscataway, NJ) column (150-ml bed volume) at a flow rate of 1 ml/min. The column was washed with 80 ml Buffer A and subsequently eluted with a linear NaCl gradient (NaCl, 0–250 m*M;* 100 + 100 ml) in Buffer A. The enzyme eluted as a single peak at about 100 m*M* NaCl. The most active fractions were pooled, ultrafiltrated, and diluted thrice in a Millipore CX-30 ultrafiltration device against 20 m*M* MOPS buffer, pH 7.5 (Buffer B), containing 2 m*M* dithiothreitol, 2 μg/ml leupeptin, and 0.1 m*M* EDTA.

The ultrafiltrate (80 ml) was loaded on a Phospho-Ultrogel (Bio-Rad Laboratories, Hercules, CA) column (50-ml bed volume), pre-equilibrated with Buffer B. After washing with 50 ml Buffer B, a linear NaCl gradient (0–400 m*M;* 100 + 100 ml) was applied. PP_i-PFK eluted as a single peak at about 225 m*M* NaCl. The most active fractions were pooled (8 ml) and diluted in about 20 ml Buffer C containing 1 μg/ml leupeptin, 2 m*M* dithiothreitol, 0.02% Triton X-100, and 25 m*M* Tris, pH 7.0.

The pooled material was loaded on a smaller QAE-Sepharose column (10-ml bed volume) previously equilibrated with Buffer C. PP_i-PFK was eluted as a single peak and was apparently homogeneous, as indicated by a single band in SDS-PAGE analysis of the active fractions.

Cleavage with Trypsin and Peptide Sequencing. Approximately 250 μg of the purified *T. vaginalis* PP_i-PFK was dissolved in 5 μl of a solution consisting of 8 *M* urea, 5 m*M* dithiothreitol, and 0.4 *M* ammonium bicarbonate, pH 7.5, and incubated for 15 min at 50°C. After the protein solution was cooled to room temperature, the sulfhydryl groups were modified by reaction with 15 m*M* vinyl pyridine for 2 h at room temperature. The mixture was diluted to 0.2 ml with water. Trypsin (treated with tosylphenylalanylchloromethylketone, from Cooper–Worthington) was added at a ratio of 1:25 (w:w protease to enzyme) and incubated for 22 h at 30°C. The mixture was then lyophilized, and the dried material taken up in 0.06% trifluoroacetic acid. Peptides were separated by Applied Biosystems reversed-phase HPLC equipped with an Alltech C-18 (150 × 2.1-mm) column. Elution was carried out with a linear water/acetonitrile gradient containing 0.06% trifluoroacetic acid (0.75% acetonitrile/min) at a flow rate of 0.15 ml/min. Absorbance was measured at 214 nm. Over 30 distinct peaks were detected, of which 13 were further purified by HPLC and sequenced. Automated sequencing of peptides (0.1–1.5 nmol) was performed in an Applied Biosystems 470A gas-phase protein sequencer. The liberated amino acid derivatives were identified and quantified by an on-line Applied Biosystems 120A HPLC with a PTH C18 column.

Other Biochemical Methods. PP_i-PFK activity was measured in the glycolytic direction by a spectrophotometric assay as described elsewhere (Mertens et al. 1989).

Gel sizing was performed on a Sephadex G-200 (Pharmacia, Piscataway, NJ) column (1.4-cm diameter, 60-ml bed volume) calibrated with high and low molecular weight standards (Bio-Rad). The column buffer was 25 m*M* Tris, pH 8.0, with 2 μg/ml leupeptin, 200 m*M* NaCl; the flow rate was 10 ml/h and the injected volume was 1.5 ml. Polyacrylamide gel electrophoresis (PAGE) and Coomassie blue staining were performed by standard methods.

Cloning Procedures. Some of the tryptic fragments revealed sufficient similarities to PFKs of other organisms to permit the design of degenerate oligonucleotide primers for cloning purposes. With a primer pair (sense, 5′-TGG TTY CTY GTN GAR GC-3′; antisense, 5′-GCN ACN GGR TCN GCR CA-3′) corresponding to amino acids 180–184 and 312–316 in the final sequence, an approximately 400-bp PCR product was obtained using cDNA prepared from *T. vaginalis* total RNA as template. This product was cloned into pBluescript KS+ and sequenced. The sequence obtained showed an overall similarity to PFKs from other organisms as well as the presence of sequences corresponding to two of the tryptic peptides.

A λZAP II cDNA library and a λZAPII *Eco*RI gDNA library of *T. vaginalis* (strain NIH-C1, ATCC 30001) provided by Dr. Patricia J. Johnson (University of California, Los Angeles, Medical School) were used. The cDNA library was screened with the amplified and biotin-labeled PCR product as a probe. The phagemid was excised with the use of R408 helper phage according to manufacturer's instructions.

Since all the cDNA clones isolated lacked the amino-terminal end of the ORF, a gDNA library was also screened. Screening this time was done with a ^{32}P-labeled PCR product representing most of the ORF. This was obtained with nondegenerate primers (sense, 5′-AGGTTTCAAGTACCTTTGCAC-3′, and antisense, 5′-AGCGCAACGAAGTTCGTAACC-3′, corresponding in the final sequence to amino acids 43–50 and 306–312.

DNA Sequencing. The cloned PCR product and the excised phagemids carrying an insert were sequenced with the dideoxy chain termination method using Sequenase II DNA polymerase. In the latter part of this study automated sequencing was performed by the Protein Sequencing Facility at the Rockefeller University. Both strands of the clones were sequenced by ''primer walking.''

Restriction Analysis. T. vaginalis gDNA was digested with the restriction enzymes *Acc*1, *Hin*dIII, *Nde*1, and *Xba*1, size fractionated in 1.2% agarose gels, and blotted to a nitrocellulose membrane (Schleicher & Schuell). The membrane was hybridized with the probe used in the screening of the gDNA library.

Evaluation of the Sequence. Sequences were assembled in the Eyeball Sequence Editor (Cabot and Beckenbach 1989). The derived amino acid sequences were compared with those of other PFKs assembled in the program MUST (Philippe 1993). Database accession numbers are listed in the figure legends. The alignment was done by visual adjustment with the use of conserved sequence motifs present throughout the molecule. Phylogenetic trees were reconstructed with neighbor-joining (NJ) and maximum-likelihood (ML) methods (NJ and PROTML in the MOLPHY package (Adachi and Hasegawa 1996) and with the maximum-parsimony (MP) method [PROTPARS in the PHYLIP package (Felsenstein 1996)].

Table 1. Purification of *Trichomonas vaginalis* PP_i-PFK

Step	Protein (mg)	Activity (U)	Specific activity (U/mg)	Purification (-fold)	Yield (%)
Crude extract	2378	599	0.25	1	100
Polyethylene glycol (10–20%)	209	276	1.32	5.3	48
QAE-Sepharose	74	171	2.3	9.2	29
P-Ultrogel	9.2	118	13	52	20
QAE-Sepharose	2.0	112	56	224	19

Results

Protein Purification and Amino Acid Sequencing

The procedure used to purify PP_i-PFK from the amitochondriate ciliate *Isotricha prostoma* (Mertens et al. 1989) was applied to *T. vaginalis* (Table 1). The procedure yielded a product with an over 200-fold increased specific activity. SDS-PAGE revealed one dominant band with a molecular mass of 50 kDa. Gel filtration on a calibrated column showed a molecular mass of 200

Table 2. Tryptic peptides from the purified PP_i-PFK of *Trichomonas vaginalis* and their position in the ORF

Peptide/ peak No.	Sequence	Gene and position
1/25, 26	SAGHLALGMAEASGAH	1–3, 189–204
2/23	ELGYGAIDAFKL	1,[a] 325–335
3/14	VYMWR	1,[a] 383–387
4/4	ITPK	1,[a] 301–304
5/21	WFLVEAMGR	1–3, 180–188
6/16–4, 16–3	SLGMEIIR	1,[a] 161–168
7/16–2	CADPVAFDAVYTR	1,[a] 312–324
8/18–1	LFGNREPPTDPH-GHILLDD	1,[a] 261–279
9/7–3, 7–4, 7–5	NLMVDSK	1–3, 169–175
10/7–4, 7–2	***K***SPELQENVR[b]	1, 88–97
11/10	KIGYEL	1,[a] 305–310
12/19a	TIDNDLPLP***A***DQSTFG-FHTAR[b]	1 and 2, 140–160
12/19b	TIDNDLPLP***S***DQSTFG-FHTAR[b]	3, 140–160
13/22	YFLTIGGDDTASSAVSVA-***Q***GMDXNE[b]	2 and 3, 107–131

[a] Only *pfk1* sequenced in this area.
[b] Residues in boldface italics differ in the different genes.

kDa for the native enzyme. Accordingly, *T. vaginalis* PP_i-PFK is a homotetramer.

A number of peptide fragments were obtained from the purified product and 13 of these were sequenced (Table 2). Several of the longer sequences were sufficiently similar to regions of various ATP-PFKs to establish their position within the *T. vaginalis* protein. Based on this information, we designed degenerate oligonucleotides to obtain a PCR product corresponding to a significant part of the coding sequence.

DNA Sequencing and Southern Analysis

All of the cDNA clones isolated lacked the area corresponding to the amino terminus of the protein, with the longest one short by only 13 amino acid residues. To obtain the complete open reading frame (ORF), several gDNA clones were also isolated and three of these were sequenced. These contained the initiator codon and a stretch of the 5′ portion of the gene (not discussed here) but all ended at an *Eco*RI site corresponding to the conserved E211. The three sequences showed a low polymorphism. One of the partial sequences was identical to the cDNA clone for the full 591-nt overlap and was assumed to represent the same gene, designated *pfk1*. The complete ORF for the *T. vaginalis pfk1* gene thus is a composite of the longest cDNA clone and this gDNA *Eco*RI fragment. The ORF comprises 426 codons and has AAA as the termination codon. In the overlapping region the cDNA and gDNA clones were colinear, showing the absence of introns, as noted also for other protein coding genes of *T. vaginalis.*

Southern blotting showed 5 hybridizing bands for gDNA digested with *Acc*1, *Hin*dIII, and *Nde*1 and 10 bands after *Xba*1 digestion (results not shown). There is an *Xba*I site in the ORF. These results indicate the presence of five *pfk* genes in the *T. vaginalis* genome.

Derived Amino Acid Sequence

The putative product (Fig. 1) of the complete ORF in *pfk1* was of a molecular mass of 46,534 Da, in good agreement with the 50-kDa value estimated from SDS-PAGE. The complete and the two partial sequences (*pfk2* and *pfk3*) differed at the amino acid level by 2–3% from each other (not shown). Sequences of all tryptic peptides were accounted for in the conceptual translations of the nucleotide sequences (residues in boldface italics in Fig. 1 and Table 2), adding up to 31% of the full ORF. With the exception of peptide 13, all others are colinear with the sequence of *pfk1*. Peptide 12 resolved into two peptides that differed at position 10, where either alanine (*pfk1* and *pfk2*) or serine (*pfk3*) was found. This corresponds to a difference between *pfk3* and the two others. Peptide 13 corresponds to *pfk2* and *pfk3*. These data indicate that the purified protein contained the products of more than one *pfk* gene of this organism.

Alignment of PFK Sequences from Various Organisms

The amino acid sequence of *T. vaginalis* PP_i-PFK was aligned with all available PP_i-PFK and PP_i-PFK-related sequences and selected ATP-PFK ones. A representative subset of these is shown in Fig. 1. In view of the greater similarity found with the catalytic amino-terminal half of the metazoan ATP-linked enzymes, we did not include the carboxy-terminal regulatory domains of these. Although a global alignment was not easily achieved, conserved areas provided a number of clear guideposts. The alignment proposed was robust in certain parts and less so in others. The data are in agreement with the notion repeatedly stressed that the amino-terminal half of the molecule, through the MGR tripeptide (residues 194–196 in *T. vaginalis*), is more conserved than the carboxy-terminal half (Ladror et al. 1991; Rozario et al. 1995). The increasing number of related sequences available made the alignment of the carboxy-terminal half of the divergent PFK sequences easier than in the past. The alignment proposed here differs from earlier ones that searched for homologous stretches between *Propionibacterium freudenreichii* PP_i-PFK and the ATP-linked enzymes (Ladror et al. 1991; Xu et al. 1994) but is close to the one shown for the *Trypanosoma brucei* glycosomal enzyme (Michels et al. 1997). It is likely that further adjustments will be possible in the carboxy-terminal half of the molecule, primarily in the region of a long indel.

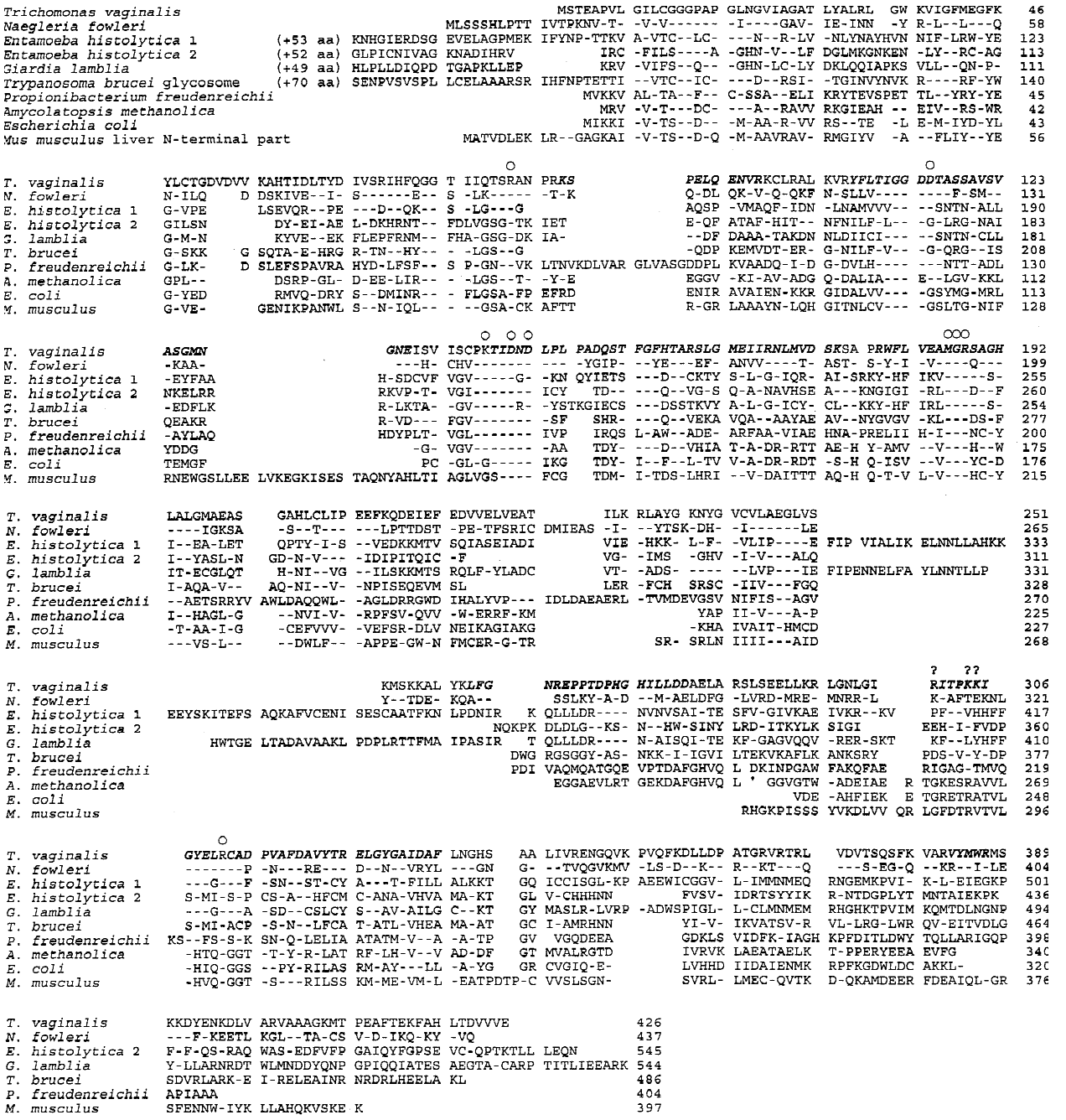

Fig. 1. Alignment of *Trichomonas vaginalis* PP_i-phosphofructokinase (*Tvpfk1*) sequence with selected phosphofructokinase sequences from other organisms. Long amino-terminal extensions of certain sequences were omitted. *Dashes* represent residues identical to those in the *T. vaginalis* sequence; *empty spaces* represent positions with a gap. *Open circles* above the alignment indicate putative catalytically important residues discussed in the text, while *question marks* indicate the three basic residues, one of which is the putative homologue of a catalytically important residue in other PFKs. In the *T. vaginalis* sequence residues in boldface italics represent the peptides obtained from the purified protein (Table 2). Database accession numbers: *Amicolatopsis methanolica*, U31277; *Entamoeba histolytica* (*pfk1*), X82173; *E. histolytica* (*pfk2*), AF013986; *Escherichia coli*, X02519; *Giardia lamblia*, U12337; *Mus musculus* liver, L27699; *Naegleria fowleri*, U11733; *Propionibacterium freudenreichii*, A41169; *Trypanosoma brucei* glycosome, AF088186.

A number of clearly defined blocks are apparent in all sequences analyzed. These parts can be aligned with those parts of the well-studied enzyme from *E. coli*, which exhibit a well-defined secondary structure (Shirakihara and Evans 1988) and also to homologous parts of the extensively explored PP_i-PFK of *P. freudenreichii* (Ladror et al. 1991; Green et al. 1992, 1993) (Fig. 1). These relatively conserved areas were separated from each other by highly divergent stretches, where some sequences often had deletions or insertions.

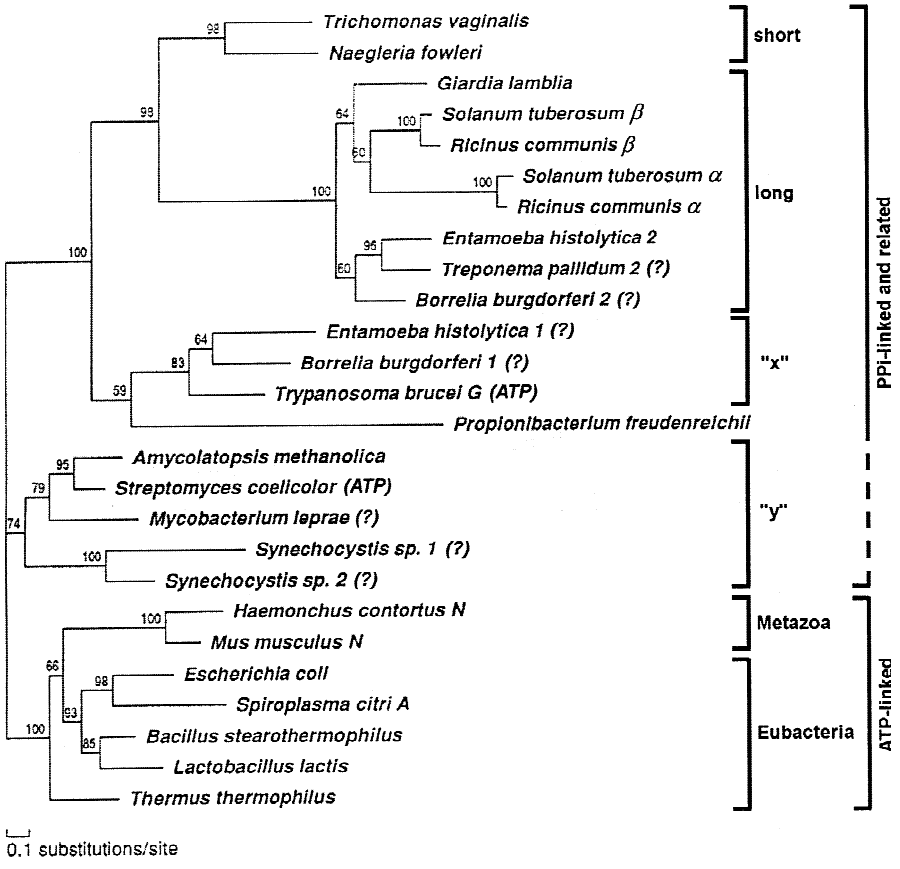

Fig. 2. Unrooted phylogenetic reconstruction of PP_i-dependent and selected ATP-dependent phosphofructokinases. Two hundred forty-seven shared residues were analyzed after exclusion of all indels and the carboxy-terminal parts of the sequences beyond amino acid residue 351 of *Tvpfk1*. Maximum-likelihood (PROTML) method. The tree was obtained with the local rearrangement option (−r) from a neighbor-joining tree. The JTT model of amino acid substitution was used. *Numbers at the nodes* represent bootstrap proportions. ATP, ATP-specific enzymes; ?, unknown donor specificity. Database accession numbers for sequences not included in the figure: *Bacillus stearothermophilus,* P00512; *Borrelia burgdorferi* sequences 1 and 2, AE00783; *Haemonchus contortus,* M59805; *Lactobacillus lactis,* Q07636; *Mycobacterium leprae,* Z99263; *Ricinus comunis* α- and β-subunits, Z32850 and Z32849; *Solanum tuberosum* α- and β-subunits, M55190 and M55191; *Spiroplasma citri,* P20275; *Streptomyces coelicolor,* U51728; *Synechocystis* sp. sequences 1 and 2, D64005 and D90901; *Thermus thermophilus,* P21777. The *Treponema pallidum* sequence was obtained from the Institute of Genomic Research, Bethesda, MD.

Amino Acid Residues Involved in Catalysis

The significant overall differences between individual PFK molecules notwithstanding, a remarkable conservation was seen of residues that may be important for catalytic activity (Fig. 1). In the amino-terminal half of the *T. vaginalis* PP_i-PFK, these could be readily identified. Site-directed mutagenesis studies of the *P. freudenreichii* enzyme (Green et al. 1993; Xu et al. 1994) have identified a number of these residues, most of which can be seen in all PP_i-PFKs. The two catalytically important aspartate residues recognized in *P* freudenreichii PP_i-PFK, and present in all PP_i-PFKs, are found within the motif TIDND (residues 140 to 144) of the *T. vaginalis* enzyme. Immediately preceding this sequence is K139 residue, which has also been shown to be important in catalysis (Xu et al. 1994) in the *P. freudenreichii* PFK and to be present in other PP_i-PFKs. Another motif of interest is the sequence GGDD (positions 112–115). The corresponding region in *E. coli* ATP-PFK is GGDG, which is in close proximity to the substrate binding site (Shirakihara and Evans 1988). The GGDD sequence or its variation GGED is found in most members of the PP_i-PFK enzyme family, but in some enzymes belonging to the PP_i-PFK-related group, the GGDG motif is present. These include the ATP-specific enzyme of *T. brucei* (Michels et al. 1997) and the short *E. histolytica* PP_i-PFK, an enzyme with very low catalytic activity (Bruchhaus et al. 1996). The sequence MGR (positions 194–196) is found in all PP_i-PFKs and also in most ATP-PFKs. Furthermore, the arginine of this tripeptide has been recently found by mutagenesis to be important for the activity of *N. fowleri* PP_i-PFK (Hinds et al. 1998). Another basic residue that has been shown in *N. fowleri* PP_i-PFK to be critical is at the position corresponding to R83 in *T. vaginalis* and is occupied by a basic residue in all PFKs (Hinds et al. 1998) with the exception of the two *Synechocystis* sp. sequences, where it is displaced by one residue downstream.

Alignments of the carboxy-terminal half of the enzyme with those of other PFKs require many more insertions and deletions, particularly when making the alignment to ATP-PFKs (Fig. 1). This makes some of the comparisons less obvious. Nonetheless, one can identify homologues of several of the basic residues that have been shown by mutagenesis to be important for substrate binding by *P. freudenreichii* (Xu et al. 1994) and *E. coli* (Berger and Evans 1990; Zheng and Kemp 1995). R311 of *T. vaginalis* can be aligned in all PFKs, and its homologues in *P. freudenreichii* (R326) and *E. coli* (R252)

have been found to be crucial to the binding of fructose-6-phosphate (Xu et al. 1994; Zheng and Kemp 1995). The alignment of *P. freudenreichii* K315, which has been shown to be important for fructose-6-phosphate binding (Green et al. 1992), with a basic residue in *T. vaginalis* is more difficult. The *T. vaginalis* and *N. fowleri* sequences show significant deletions in this area, suggesting that their secondary structure will be different from the other enzymes. This position is occupied by either arginine, histidine, or lysine in all PFKs. In *T. vaginalis*, it could be aligned with R300, K304, or K305, depending on the placement of a deletion (question marks in Fig. 1). Without functional studies on the *T. vaginalis* enzyme, it cannot be predicted, which will turn out to play a role in the binding of the substrate.

Classification and Phylogenetic Reconstruction of PP_i-PFKs

Several PFK sequences related to PP_i-specific enzymes have been determined recently and prompted us to reexamine their relationships by direct comparison and phylogenetic reconstruction. We analyzed a complete alignment of all 19 sequences available for PP_i-specific and related enzymes. A number of these are derived from gene sequences, thus neither their donor specificity nor their physiological function is known. For comparison, seven eubacterial and metazoan ATP-linked enzymes were included. Visual inspection of the alignment suggested the existence of at least four separate groups of PP_i-PFK related enzymes, a conclusion borne out by further analysis. These groups are described in some detail below.

The pairwise amino acid identity between various sequences of the whole set (considering 247 shared residues) ranged from 20 to 90% (Table 3). Intergroup amino acid identity values were, with few exceptions, below 35% for this dataset, which was restricted to the most conserved parts of the sequences. The number of clearly homologous positions within each of the four groups significantly exceeded the number of those recognizable in the alignment of the full set. Amino acid identity scores determined for each of these groups separately (Table 3) confirmed the conclusions drawn from fewer shared positions. Phylogenetic reconstructions with ML (Fig. 2) and PM methods (not shown) also recovered the four subgroups of PP_i-linked and related enzymes. These were clearly separated from the bacterial and animal ATP-linked enzyme group by a well-supported [100% bootstrap proportion (BP)] node, indicating their sharing a most recent common ancestor, well documented in the literature (Alves et al. 1996). The latter group was used as an outgroup in the phylogenetic reconstructions, but the trees obtained are unrooted.

The first subgroup (short) consisted of two sequences, the *T. vaginalis* sequence reported here and that of *N. fowleri*. Both enzymes are of known PP_i-specificity (Mertens et al. 1989, 1993b). The two sequences were colinear throughout their entire length, with only a few short indels (Fig. 1), and showed over 50% amino acid identity. In phylogenetic reconstructions they shared a nearest common ancestor. The corresponding node had strong support (98% BP).

The second group (long) comprised a number of longer sequences. In addition to the protists *G. lamblia* (Rozario et al. 1995) and *E. histolytica* (*pfk2*) (Deng et al. 1998), it included sequences from two spirochaete species [*Treponema pallidum* and *Borrelia burgdorferi pfk2* (Fraser et al. 1997)] as well as genes for the catalytic β-subunits and the regulatory α-subunits of plant PP_i-PFKs (Carlisle et al. 1990; Todd et al. 1995). The two protist sequences correspond to PP_i-linked PFKs with glycolytic function (Li and Phillips 1995; Deng et al. 1998). The products of plant genes have also been characterized functionally, while the products of the spirochaete genes are as yet unknown. Disregarding their amino-terminal extensions of different lengths, members of this group were also colinear to each other. They differed, however, from other sequences in having a long (about 80-residue) insertion in their carboxy-terminal half. The intragroup amino acid identity ranged from 39 to 58%, with the pairs of plant subunits even more similar to each other. This group also forms a well-supported (100% BP) clade, connected to the rest of the tree by a rather long internal node. The internal relationships within this group are not robust.

The third group (provisionally designated group ''X'') consisted of three seemingly unrelated sequences, the ATP-linked glycosomal enzyme of *T. brucei* (Michels et al. 1997), a second enzyme of *E. histolytica* (*pfk1*), which possibly does not have a glycolytic function (Deng et al. 1998), and a second protein of unknown function from *T. burgdorferi* (*pfk1*) (Fraser et al. 1997). The robustness of this group is indicated by the relatively high intragroup amino acid identity of about 40% and high (89% BP) support in phylogenetic reconstructions. These sequences also have an insertion, though shorter, in the area where the long sequences have one. Similarities between the insertions of the two groups are minimal.

The fourth group (provisionally designated group ''Y'') contained two closely related sequences from the cyanobacterium *Synechocystis* sp. (Kaneko et al. 1996), and sequences from high-G + C, Gram-positive eubacteria, two actinomycetes (Alves et al. 1996, 1997), and *Mycobacterium leprae*. Donor specificity is known only for the closely related actinomycete pair. Interestingly one is ATP- and the other PP_i-linked. The intragroup amino acid identities for this group are 33 to 70%, with

Table 3. Percentage pairwise amino acid identities between phosphofructokinases of various species: below diagonal, data for all sequences studied (247 shared residues); above diagonal, data for the individual enzyme groups

	Short		Long								''X''		
Enzyme group (shared residues)	*T.v.*	*N.f.*	G.1.	*E.h.2*	*B.b.2*	*T.p.2*	*S.t.*β	*R.c.*β	*S.t.*α	*R.c.*α	*E.h.1*	*B.b.1*	*T.b.*
Short (416)													
Trichomonas vaginalis		**52.2**											
Naegleria fowleri	**54.3**												
Long (461)													
Giardia lamblia	34.4	31.6		**45.6**	**49.2**	**47.7**	**51.2**	**48.8**	**39.0**	**39.0**			
Entamoeba histolytica 2	34.0	27.9	**55.5**		**53.4**	**57.3**	**50.5**	**50.3**	**57.5**	**57.5**			
Borrelia burgdorferi 2	35.6	32.0	**61.1**	**64.4**		**54.2**	**49.7**	**50.3**	**36.9**	**37.3**			
Treponema pallidum 2	33.2	28.3	**57.1**	**68.4**	**63.2**		**52.3**	**52.3**	**38.5**	**40.1**			
Solanum tuberosum β	35.2	32.8	**61.1**	**57.6**	**57.5**	**59.9**		**86.6**	**42.3**	**43.2**			
Ricinus communis β	34.8	31.2	**57.9**	**58.7**	**58.3**	**59.1**	**88.7**		**41.9**	**42.5**			
Solanum tuberosum α	37.5	30.0	**48.6**	**44.9**	**46.6**	**48.2**	**50.2**	**50.2**		**86.3**			
Ricinus communis α	26.7	29.6	**48.2**	**44.5**	**47.0**	**49.8**	**51.4**	**50.6**	**90.3**				
Group ''X'' (431)													
Entamoeba histolytica 1	35.6	34.4	29.1	27.1	28.7	29.5	29.5	28.3	25.9	26.7		**39.0**	**38.0**
Borrelia burgdorferi 1	30.8	30.8	27.1	29.1	30.0	30.0	30.0	29.1	24.3	23.9	**51.0**		**39.9**
Trypanosoma brucei G	26.4	33.2	29.1	27.5	30.8	31.5	28.7	28.7	23.9	24.3	**51.4**	**52.6**	
Propionibacterium freudenreichii	27.5	27.9	23.5	24.7	23.9	26.3	25.1	24.3	21.5	22.1	32.4	30.8	30.0
Group ''Y'' (336)													
Amycolatopsis methanolica	36.8	34.4	28.7	27.5	30.4	30.4	29.1	27.5	25.5	24.7	38.1	34.4	37.7
Synechococystis sp. 1	28.3	28.7	23.9	23.9	26.7	24.7	27.1	27.1	22.7	21.5	30.4	28.5	27.5
Synechococystis sp. 2	30.0	27.5	27.5	27.5	30.0	27.1	27.1	26.7	23.9	22.3	31.6	31.2	32.8
Streptomyces coelicolor	35.2	34.4	27.5	29.2	30.4	30.0	30.4	28.3	25.5	24.7	36.8	34.4	37.2
Mycobacterium leprae	34.4	31.6	27.5	23.9	26.7	27.5	27.5	28.3	25.9	25.9	36.8	36.0	34.8
ATP-linked (317)													
Escherichia coli	30.4	28.7	25.5	26.5	27.5	26.7	30.4	28.7	23.1	22.7	28.7	29.1	28.7
Thermus thermophilus	33.2	33.6	28.3	29.1	28.3	30.0	30.4	30.1	27.1	25.1	34.8	32.4	35.7
Bacillus stearothermophilus	31.6	34.8	27.5	27.9	27.5	28.7	31.6	30.4	27.9	26.7	33.2	33.2	34.0
Lactobacillus lactis	30.8	33.6	25.1	25.9	23.9	25.5	29.5	27.9	23.9	22.7	31.2	31.6	29.1
Spiroplasma citri	27.1	26.3	24.3	23.9	26.7	24.3	27.5	26.7	19.4	19.0	29.1	25.1	31.2
Haemonchus contortus	27.1	28.7	22.6	23.1	23.5	23.5	23.9	23.1	19.0	20.2	32.0	27.1	29.5
Mus musculus	30.4	29.1	23.9	24.3	25.5	24.3	25.5	24.7	21.5	21.9	30.0	26.7	30.0

one of the *Synechocystis* sp. sequences being the most divergent. Support for two subgroups in phylogenetic reconstructions is strong, but not for their sharing a most recent common ancestor.

The extensively studied PP_i-PFK from the eubacterium *P. freudenreichii* (Ladror et al. 1991; Green et al. 1992, 1993) is so divergent as to resist all attempts to assign it to any subgroup. Outside of the highly conserved blocks, its alignment is uncertain. It has insertions at regions where none of the other sequences have one. Its long branch on phylogenetic reconstructions indicates an accelerated rate of evolution and a high level of mutational saturation. Its changing position in the trees (not shown) suggests that it was subject to a long branch attraction artifact.

The relationships of the four groups to each other are less clearly defined. Although some of the internal nodes have high bootstrap support, most are short and probably mutationally saturated. Nonetheless, the data indicate a close sister-group relationship between the long and the short PP_i-PFK groups and a similar relationship between the common ancestor of these two and group ''X.'' As mentioned above, the position of *P. freudenreichii* remains unresolved. The fourth group (''Y'') presents a somewhat different situation. Although it is consistently recovered in phylogenetic reconstructions, the node connecting it to the other lineages has low bootstrap support. The intragroup amino acid identity for this group does not exceed the similarity values between its members and members of the ATP-PFK group. Since of the five members of this group, three have not been characterized, one is ATP-specific, and only one is PP_i-dependent, it seems prudent to withhold the decision whether this group is a genuine member of the PP_i-linked family or not. The reconstructions are unrooted, leaving open the nature of the ancestral enzyme for this enzyme family.

Discussion

Our results show that the amitochondriate flagellated protist, *Trichomonas vaginalis,* contains a tetrameric PP_i-PFK. Its PP_i specificity and unresponsiveness to fructose-2,6-bisphosphate (Mertens et al. 1989) places it in the Type 1 group of PP_i-PFK enzymes. It is simpler in its structure than the allosterically regulated ATP-linked

Table 3. Extended

	''Y''					ATP-linked						
P.f.	*A.m.*	*S.1*	*S.2*	*S.c.*	*M.1.*	*E.c.*	*T.t.*	*B.s.*	*L.1.*	*S.c.*	*H.c.*	*M.m.*
26.7		**34.0**	**42.3**	**70.3**	**55.1**							
23.9	**42.9**		**26.9**	**33.9**	**32.4**							
22.7	**47.4**	**52.2**		**43.2**	**40.8**							
27.9	**74.1**	**43.0**	**38.6**		**45.7**							
27.5	**57.9**	**40.1**	**47.4**	**60.7**								
25.5	45.0	38.9	43.3	54.3	42.1		**47.0**	**54.9**	**54.6**	**47.6**	**37.5**	**30.4**
26.3	55.1	35.6	42.5	53.4	48.6	**53.0**		**57.4**	**49.8**	**41.6**	**41.0**	**42.9**
25.1	50.6	39.7	45.3	49.4	47.0	**57.9**	**64.4**		**61.5**	**46.7**	**41.6**	**45.4**
25.5	47.8	41.3	42.9	47.0	42.5	**60.3**	**56.7**	**68.8**		**43.2**	**40.0**	**40.0**
26.3	40.9	40.5	40.5	41.3	38.5	**52.6**	**47.0**	**51.4**	**48.6**		**35.6**	**35.3**
23.9	39.3	31.6	36.0	39.3	39.3	**45.5**	**47.0**	**49.8**	**47.4**	**40.9**		**64.4**
27.5	40.5	35.6	37.2	42.5	42.1	**47.0**	**47.7**	**52.6**	**47.4**	**40.5**	**70.0**	

and PP_i-linked enzymes of fungi, animals, and plants (Fothergill-Gilmore and Michels 1993).

Southern hybridization and sequencing results revealed the presence of at least five very similar copies of the *pfk* gene in the *T. vaginalis* genome. Divergence of the upstream noncoding areas suggests that these represent separate loci and not an allelic polymorphism (not shown). The peptide data suggest that more than one of these genes is expressed. The subunit composition of the active tetramer, however, remains to be established. Multiple copies, often more than two, have been reported for a number of genes of *T. vaginalis* and other trichomonads (e.g., Lahti et al. 1992, 1994; Hrdý and Mùller 1995a, 1995b; Bui and Johnson 1996; Viscogliosi et al. 1996; Viscogliosi and Müller 1998). The origin and significance of this phenomenon remain to be elucidated.

The few PFKs characterized from protists, whether PP_i-linked (Reeves 1968; Mertens et al. 1989, 1993; Mertens 1990; Peng and Mansour 1992; Denton et al. 1994) or ATP-linked (Cronin and Tipton 1985), have subunit sizes ranging from 42 to 67 kDa. Although larger than the 35-kDa eubacterial subunit, this size still can accommodate only one substrate binding site and one phosphoryl-donor binding site and thus corresponds to the simplest types of PFKs known. The active oligomers probably are also composed of identical or almost-identical subunits. The catalytic and regulatory parts or subunits of the complex enzymes of fungi, multicellular animals, and plants (Fothergill-Gilmore and Michels 1993) probably arose through early gene duplications (Poorman et al. 1984; Kruger and Hammond 1988; Carlisle et al. 1990; Todd et al. 1995).

The high sequence similarity of the PFKs from the hydrogenosome-containing *T. vaginalis* and the mitochondriate *N. fowleri* is clearly reflected in their sister-group relationship. Interesting the enzyme is unregulated in *T. vaginalis* (Mertens et al. 1989) and regulated by AMP in *N. fowleri* (Mertens et al. 1993b). Only one additional glycolytic enzyme has been sequenced from both groups, glyceraldehyde-3-phosphate dehydrogenase (Markoš et al. 1993; Roger et al. 1996). This is, however, highly divergent in the two species and reveals separate evolutionary histories. These two species belong to two unrelated protist lineages, the parabasalids and the heterolobosea (Cavalier-Smith 1996), and the close relationship of their PP_i-PFKs is probably due only to the small sample size.

Adding the sequence of the *T. vaginalis* PP_i-PFK to

recent data on the enzymes of other protists (Deng et al. 1998; Rozario et al. 1995; Wessberg et al. 1995; Huang et al. 1995; Bruchhaus et al. 1996; Michels et al. 1997) and new sequence information emerging from genome projects provided an opportunity to define tentatively the overall relationships within the PP_i-PFK enzyme family and establish the position of protist sequences within it. At present four subgroups can be delimited, which accommodate in monophyletic groups all known sequences with the exception of that of *P. freudenreichii.* Three of these are well supported by bootstrap values: short enzymes, long enzymes, and group ''X,'' which contains a set of enzymes with unusual or not-well-defined functions. The fourth one (group ''Y''), of cyanobacterial and high-G + C, Gram-positive eubacterial sequences, shares a last common ancestor with the first three, but its monophyly is less well supported. The database is very limited and it is likely that further sequences will increase the membership of various subfamilies and possibly also lead to the recognition of further ones.

The enlarged data set challenges earlier suggestions for a linear evolutionary history of PP_i-PFK, which regarded *A. methanolica* the ancestral eubacterial form and depicted a progression through several steps of increasing structural and regulatory complexity to culminate in the plant enzymes (Alves et al. 1996). The subfamily containing *A. methanolica* is indeed the first branch on the PP_i-PFK tree, but the subsequently emerging subfamilies do not reveal such a progression. Most importantly, the glycolytically active PP_i-PFKs of protists belong to two separate monophyletic clades (short and long enzymes), which are in sister-group relationship to each other. It can be anticipated that the short enzyme subfamily will accommodate also the PP_i-PFKs of several alveolates as indicated by the subunit size of the enzymes from the apicomplexans, *Toxoplasma gondii* (Peng and Mansour 1992) and *Eimeria tenella* (Denton et al. 1994), and from the ciliate, *Isotricha prostoma* (Mertens et al. 1989).

The presence of eubacterial and eukaryotic sequences in two of the four subfamilies suggests that an early gene duplication, possibly antedating the emergence of eukaryotes, produced two paralogous lineages, the short and long PP_i-PFKs. In various lineages descendants of either the short or the long ancestor assumed a key role in catabolism. In *E. histolytica,* just as in *G. lamblia* and the plants, descendants of the long form became the major type, while in other protist lineages, parabasalids (*T. vaginalis*) and heterolobosea (*N. fowleri*), the short ones became fixed. Selective retention of one or another paralogous gene in the glycolytic pathway has been noted for other enzymes as well (Markoš et al. 1993; Henze et al. 1995; Martin and Schnarrenberger 1997). The third subfamily (group ''X'') is particularly enigmatic. It includes genes coding for a second, different PP_i-PFK-related enzyme from two organisms that contain genes for the typical long enzyme. While the spirochaete *B. burgdorferi pfk1* product has not been characterized, the expressed product of *E. histolytica pfk1* had only negligible PP_i-PFK activity (Bruchhaus et al. 1996), thus its functional role remains to be elucidated. These findings again point to another early gene duplication with the arising paralogous genes inherited together by some lineages, with the long one retaining its original role and the short one probably undergoing a functional change. The third member of this group, the glycosomal ATP-PFK of *Trypanosoma brucei* does function in glycolysis (Cronin and Tipton 1985) and is assumed to have acquired its ATP specificity secondarily (Michels et al. 1997). Biochemical characterization of all members of group ''X'' will be needed to gain a clearer view of the functional significance of this peculiar group.

It remains to be seen whether coexisting long and group ''X'' genes will also be found in *G. lamblia.* Earlier studies by one of us (E.M.) (Mertens 1990), and unpublished data quoted by Mertens (1991) indicated that its enzyme is a dimer of 43 kDa subunits, while the product of the sequenced putative gene is 63 kDa (Rozario et al. 1995). These findings left open the possibility that the active enzyme of this organism is not the one coded by the gene sequenced. Recently a purified active enzyme was found to be a monomer of 63–67 kDa (Li and Phillips 1995), probably corresponding to the putative translation product. The discrepancy in these results remains unexplained. Proteolytic degradation could account for the smaller size of the enzyme purified earlier, but one cannot dismiss the possibility that *G. lamblia* also has two expressed genes and that, in the two studies, one or the other was purified and studied.

In conclusion, even in the few sequences from protists, PP_i-PFK presents a bewildering diversity, which points to interesting directions, but does not permit us to paint a consistent picture of their evolutionary history and, perhaps more significantly, of the extent of their functional diversity and the number of occasions when changes in phosphoryl donor or substrate specificity occurred. It remains to be elucidated how much of this diversity is due to selective retention of paralogous gene lineages as documented for other glycolytic enzymes (Markoš et al. 1993; Henze et al. 1995; Martin and Schnarrenberger, 1997) or are due to lateral gene transfers occurring later in evolution (Doolittle et al. 1990; Smith et al. 1992; Rosenthal et al. 1997).

Acknowledgments. We acknowledge the generous help of Drs. Patricia J. Johnson (Los Angeles, CA) and John M. Logsdon (Halifax, NS) by making the cDNA and gDNA libraries of *T. vaginalis* available and giving advice. We also thank Drs. Hervé Philippe, Philippe Lopez, and Henner Brinkmann (Orsay, France) for valuable comments and help in retrieving sequences from genome databases and The Institute of Genomic Research (Bethesda, MD) for making sequence data available prior to publication. The MUST package was kindly supplied by Dr. Philippe, MOLPHY by Drs. Jun Adachi and Masami Hasegawa (Tokyo, Japan), and PHYLIP by Dr. Joseph Felsenstein (Seattle, WA).

This study was supported by U.S. Public Health Service Grants AI 11942 to M.M. and AI 34527 to R.G.K.

Addendum. After completion of the manuscript a partial sequence of another plant (*Prunus armeniaca*–apricot) PP_i-PFK became available in GenBank (U93272). This sequence is not part of the long PP_i-PFK group but of the ''X'' group. This finding indicates a broader distribution of members of this group and underlines the need for their detailed functional exploration and for an extension of the taxonomic sampling.

References

Adachi J, Hasegawa M (1996) MOLPHY Version 2.3: programs for molecular phylogenetics based on maximum likelihood. Computer Science Monographs, No. 28. Institute of Statistical Mathematics, Tokyo, pp 1–150

Alves AMCR, Meijer WG, Vrijbloed JW, Dijkhuizen L (1996) Characterization and phylogeny of the pfp gene of *Amycolatopsis methanolica* encoding PPi-dependent phosphofructokinase. J Bacteriol 178:149–155

Alves AMCR, Euverink GJW, Bibb MJ, Dijkhuizen L (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

Berger SA, Evans PR (1990) Active-site mutants altering the cooperativity of *E. coli* phosphofructokinase. Nature 343:575–576

Bruchhaus I, Jacobs T, Denart M, Tannich E (1996) Pyrophosphate-dependent phosphofructokinase of *Entamoeba histolytica:* molecular cloning, recombinant expression and inhibition by pyrophosphate analogues. Biochem J 316:57–63

Bui ETN, Johnson PJ (1996) Identification and characterization of [Fe]-hydrogenases in the hydrogenosome of *Trichomonas vaginalis*. Mol Biochem Parasitol 76:305–310

Cabot EL, Beckenbach AT (1989) Simultaneous editing of multiple nucleic acid and protein sequences with ESEE. Comput Appl Biosci 5:233–234

Carlisle SM, Blakeley SD, Hemmingsen SM, Trevanion SJ, Hiyoshi T, Kruger NJ, Dennis DT (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

Cavalier-Smith T (1996) Amoeboflagellates and mitochondrial cristae in eukaryote evolution: megasystematics of the new protozoan subkingdoms Eozoa and Neozoa. Arch Protistenkd 147:237–258

Cronin CN, Tipton KF (1985) Purification and regulatory properties of phosphofructokinase from *Trypanosoma* (*Trypanozoon*) *brucei brucei.* Biochem J 227:113–124

Deng Z, Huang M, Singh K, Albach RA, Latshaw SP, Chang K-P, Kemp RG (1998) Cloning and expression of the gene for the active PPi-dependent phosphofructokinase of *Entamoeba histolytica.* Biochem J 329:659–664

Denton H, Thong K-W, Coombs GH (1994) Eimeria tenella contains a pyrophosphate-dependent phosphofructokinase and a pyruvate kinase with unusual allosteric regulators. FEMS Microbiol Lett 115: 87–91

Denton H, Brown SMA, Roberts CW, Alexander J, McDonald V, Thong K-W, Coombs GH (1996) Comparison of phosphofructokinase and pyruvate kinase activities of *Cryptosporidium parvum, Eimeria tenella* and *Toxoplasma gondii.* Mol Biochem Parasitol 76:23–29

Diamond LS (1957) The establishment of various trichomonads of animals and man in axenic cultures. J Parasitol 43:488–490

Doolittle RF, Feng DF, Anderson KL, Alberro MR (1990) A naturally occurring horizontal gene transfer from a eukaryote to a prokaryote. J Mol Evol 31:383–388

Doran DJ (1957) Studies on trichomonads. I. The metabolism of Tritrichomonas foetus and trichomonads from the nasal cavity and cecum of swine. J Protozool 4:182–190

Felsenstein J (1996) PHYLIP (Phylogeny inference package), version 3.57. University of Washington, Seattle

Fothergill-Gilmore LA, Michels PAM (1993) Evolution of glycolysis. Progr Biophys Mol Biol 59:105–235

Fraser CM, Casjens S, Huang WM, Sutton GG, Clayton R, Lathigra R, White O, Ketchum KA, Dodson R, Hickey EK, Gwinn M, Dougherty B, Tomb J-F, Fleischmann RD, Richardson D, Peterson J. Kerlavage AR, Quackenbush J, Salzberg S, Hanson M, van Vugt R, Palmer N, Adams MD, Gocayne J, Weidman J, Utterback T, Whattey L, McDonald L, Artiach P, Bowman P, Garland S, Fujii C, Cotton MD, Horst K, Roberts K, Hatch B, Smith HO, Venter JC (1997) Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi.* Nature 390:580

Green PC, Latshaw SP, Ladror US, Kemp RG (1992) Identification of critical lysyl residues in the pyrophosphate-dependent phosphofructo-1-kinase of *Propionibacterium freudenrechii.* Biochemistry 31:4815–4821

Green PC, Tripathi RL, Kemp RG (1993) Identification of active site residues in pyrophosphate-dependent phosphofructo-1-kinase by site-directed mutagenesis. J Biol Chem 268:5085–5088

Henze K, Badr A, Wettern M, Cerff R, Martin W (1995) A nuclear gene of eubacterial origin in *Euglena gracilis* reflects cryptic endosymbioses during protist evolution. Proc Natl Acad Sci USA 92: 9122–9126

Hinds RM, Xu J, Walters DE, Kemp RG (1998) The active site of PPi-dependent phosphofructokinase based on site-directed mutagenesis and molecular modeling. Arch Biochem Biophys 349:47–52

Hrdý I, Müller M (1995a) Primary structure and eubacterial relationships of the pyruvate:ferredoxin oxidoreductase of the amitochondriate eukaryote, *Trichomonas vaginalis*. J Mol Evol 41:388–396

Hrdý I, Müller M (1995b) Primary structure of the hydrogenosomal malic enzyme of *Trichomonas vaginalis* and its relationship to homologous enzymes. J Eukar Microbiol 42:593–603

Huang M, Albach RA, Chang K-P, Tripathi RL, Kemp RG (1995) Cloning and sequencing a putative pyrophosphate-dependent phosphofructokinase gene from *Entamoeba histolytica.* Biochem Biophys Acta 1260:215–217

Kaneko T, Sato S, Kotani H, Asamizu E, Nakamura Y, Miyajima N, Hirosawa M, Sugiura M, Sasamoto S, Kimura T, Hosouchi T, Hatsuno A, Muraki A, Nakazaki N, Naruo K, Okumura S, Shimpo S, Takeuchi C, Wada T, Watanabe A, Yamada M, Yasuda M, Takata S (1996) Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. DNA Res 3:109–136

Kruger NJ, Hammond JBW (1988) Molecular comparison of pyrophosphate- and ATP-dependent fructose 6-phosphate 1-phosphotransferase from potato tuber. Plant Physiol 86:645–648

Ladror U, Gollapudi L, Tripathi RL, Latshaw SL, Kemp RG (1991) Cloning, sequencing and expression of pyrophosphate-dependent phosphofructokinase from *Propionibacterium freudenreichii.* J Biol Chem 266:16550–16555

Lahti CJ, d'Oliveira CE, Johnson PJ (1992) β-Succinyl-coenzyme A synthetase from *Trichomonas vaginalis* is a soluble hydrogenosomal protein with an amino-terminal sequence that resembles mitochondrial presequences. J Bacteriol 174:6822–6830

Lahti CJ, Bradley PJ, Johnson PJ (1994) Molecular characterization of the α-subunit of *Trichomonas vaginalis* hydrogenosomal succinyl CoA synthetase. Mol Biochem Parasitol 66:309–318

Li Z, Phillips NFB (1995) Pyrophosphate-dependent phosphofructokinase from *Giardia lamblia:* purification and characterization. Prot Express Purif 6:319–328

Markoš A, Miretsky A, Müller M (1993) A glyceraldehyde 3-phosphate dehydrogenase with eubacterial features in the amito-

chondriate eukaryote, *Trichomonas vaginalis*. J Mol Evol 37:631–643

Martin W, Schnarrenberger C (1997) The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: a case study of functional redundancey in ancient pathways through endosymbiosis. Curr Genet 32:1–18

Mertens E (1990) Occurrence of pyrophosphate:fructose-6-phosphate 1-phosphotransferase in *Giardia lamblia* trophozoites. Mol Biochem Parasitol 40:147–150

Mertens E (1991) Pyrophosphate-dependent phosphofructokinase, an anaerobic glycolytic enzyme? FEBS Lett 285:1–5

Mertens E (1993) ATP versus pyrophosphate: glycolysis revisited in parasitic protists. Parasitol Today 9:122–126

Mertens E, Van Schaftingen E, Muller M (1989) Presence of a fructose-2,6-bisphosphate-insensitive pyrophosphate:fructose-6-phosphate phosphotransferase in the anaerobic protozoa *Tritrichomonas foetus, Trichomonas vaginalis* and *Isotricha prostoma.* Mol Biochem Parasitol 37:183–190

Mertens E, De Jonckheere J, Van Schaftingen E (1993) Pyrophosphate-dependent phosphofructokinase from the amoeba, *Naegleria fowleri*, an AMP-sensitive enzyme. Biochem J 292:797–803

Michels PAM, Chevalier N, Opperdoes FR, Rider MH, Rigden D (1997) The glycosomal ATP-dependent phosphofructokinase of *Trypanosoma brucei* must have evolved from an ancestral pyrophosphate dependent enzyme. Eur J Biochem 250:698–704

Miyatake K, Enomoto T, Kitakoa S (1986) Fructose-2,6-bisphosphate activates pyrophosphate: D-fructose 6-phosphate 1-phosphotransferase from *Euglena gracilis.* Agr Biol Chem 50: 2417–2418

Müller M (1988) Energy metabolism of protozoa without mitochondria. Annu Rev Microbiol 42:465–488

Müller M (1998) Enzymes and compartmentation of core energy metabolism of anaerobic protists—a special case in eukaryotic evolution?. In: Coombs GH, Vickerman K, Sleigh MA, Warren A (eds) Evolutionary relationships among protozoa. Chapman & Hall, London, pp 109–131

Peng Z-Y, Mansour TE (1992) Purification and properties of a pyrophosphate-dependent phosphofructokinase from *Toxoplasma gondii.* Mol Biochem Parasitol 54:223–230

Philippe H (1993) MUST, a computer package of Management Utilities for Sequences and Trees. Nucleic Acids Res 21:5264–5272

Phillips NFB, Li Z (1995) Kinetic mechanism of pyrophosphate-dependent phosphofructokinase from *Giardia lamblia.* Mol Biochem Parasitol 73:43–51

Poorman RA, Randolph A, Kemp RG, Heinrickson RL (1984) Evolution of phosphofructokinase—gene duplication and creation of new effector sites. Nature 309:467–469

Reeves RE (1968) A new enzyme with the glycolytic function of pyruvate kinase. J Biol Chem 243:3202–3204

Roger AJ, Smith MW, Doolittle RF, Doolittle WF (1996) Evidence for the Heterolobosea from phylogenetic analysis of genes encoding glyceraldehyde-3-phosphate dehydrogenase. J Eukar Microbiol 43: 475–485

Rosenthal B, Mai Z, Caplivksi D, Ghosh S, de la Vega H, Graf T, Samuelson J (1997) Evidence for the bacterial origin of genes encoding fermentation enzymes of the amitochondriate protozoan parasite *Entamoeba histolytica.* J Bacteriol 179:3736–3747

Rozario C, Smith MW, Muller M (1995) Primary sequence of a putative pyrophosphate-linked phosphofructokinase gene of *Giardia lamblia.* Biochim Biophys Acta 1260:218–222

Shirakihara Y, Evans PR (1988) Crystal structure of the complex of phosphofructokinase from *Escherichia coli* with its reaction products. J Mol Biol 204:973–994

Smith MW, Feng D-F, Doolittle R (1992) Evolution by acquisition: the case for horizontal gene transfers. Trends Biochem Sci 17:489–493

Todd JF, Blakeley SD, Dennis DT (1995) Structure of the genes encoding the α- and β-subunits of castor pyrophosphate-dependent phosphofructokinase. Gene 152:181–186

Viscogliosi E, Müller M (1998) Phylogenetic relationships of the glycolytic enzyme, glyceraldehyde-3-phosphate dehydrogenase, in trichomonad flagellates. J Mol Evol (in press)

Viscogliosi E, Durieux I, Delgado-Viscogliosi P, Bayle D, Dive D (1996) Phylogenetic implications of iron-containing superoxide dismutase genes from trichomonad species. Mol Biochem Parasitol 80:209–214

Wessberg KL, Skolnick S, Xu J, Marciano-Cabral F, Kemp RG (1995) Cloning, sequencing and expression of the pyrophosphate-dependent phosphofructo-1-kinase from *Naegleria fowleri*. Biochem J 307:143–149

Wood HG (1977) Some reactions in which inorganic pyrophosphate replaces ATP and serves as a source of energy. Fed Proc 36:2197–2205

Wu L-F, Reizer A, Reizer J, Cai B, Tomich JM, Saier Jr MH (1991) Nucleotide sequence of the Rhodobacter capsulatus fruK gene, which encodes fructose-1-phosphate kinase: evidence for a kinase superfamily including both phosphofructokinases of *Escherichia coli.* J Bacteriol 173:3117–3127

Xu J, Green PC, Kemp RG (1994) Identification of basic residues involved in substrate binding and catalysis by pyrophosphate-dependent phosphofructokinase from *Propionibacterium freudenreichii.* J Biol Chem 269:15553–15557

Zheng R-L, Kemp RG (1995) Phosphofructo-1-kinase: role of charge neutralization in the active site. Biochem Biophys Res Commun 214:765–770