J Mol Evol (1998) 47:728–738

JOURNAL OF MOLECULAR EVOLUTION

Comparison and Evolutionary Analysis of the Glycosomal Glyceraldehyde-3-Phosphate Dehydrogenase from Different Kinetoplastida

Véronique Hannaert, Fred R. Opperdoes,[1] Paul A.M. Michels[1]

[1] Research Unit for Tropical Diseases, Christian de Duve Institute of Cellular Pathology (ICP), and Laboratory of Biochemistry, Catholic University of Louvain, ICP-TROP/74.39, Avenue Hippocrate 74, B-1200, Brussels, Belgium

Received: 23 February 1998/Accepted: 26 March 1998

Abstract. In this work, we present the sequences and a comparison of the glycosomal GAPDHs from a number of Kinetoplastida. The complete gene sequences have been determined for some species (*Crithidia fasciculata, Herpetomonas samuelpessoai, Leptomonas seymouri,* and *Phytomonas* sp), whereas for other species (*Trypanosoma brucei gambiense, Trypanosoma congolense, Trypanosoma vivax,* and *Leishmania major*), only partial sequences have been obtained by PCR amplification. The structure of all available glycosomal GAPDH genes was analyzed in detail. Considerable variations were observed in both their nucleotide composition and their codon usage. The GC content varies between 64.4% in *L. seymouri* and 49.5% in the previously sequenced GAPDH gene from *Trypanoplasma borreli.* A highly biased codon usage was found in *C. fasciculata,* with only 34 triplets used, whereas in *T. borreli* 57 codons were employed. No obvious correlation could be observed between the codon usage and either the nucleotide composition or the level of gene expression. The glycosomal GAPDH is a very well-conserved enzyme. The maximal overall difference observed in the amino acid sequences is only 25%. Specific insertions and extensions are retained in all sequences. The residues involved in catalysis, substrate, and inorganic phosphate binding are fully conserved, whereas some variability is observed in the cofactor-binding pocket. The implications of these data for the design of new trypanocidal drugs targeted against GAPDH are discussed. All available gene and amino acid sequences of glycosomal GAPDHs were used for a phylogenetic analysis. The division of the Kinetoplastida into two suborders, Bodonina and Trypanosomatina, was well supported. Within the letter group, the *Trypanosoma* species appeared to be monophyletic, whereas the other trypanosomatids form a second clade.

Key words: Kinetoplastida — *Trypanosoma* — *Leishmania* — *Crithidia fasciculata* — *Herpetomonas samuelpessoai* — *Leptomonas seymouri* — Glyceraldehyde-3-phosphate dehydrogenase — Glycosome — Codon usage — Phylogeny

Introduction

The glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase (GAPDH) catalyzes the oxidative phosphorylation of glyceraldehyde 3-phosphate into 1,3-bisphosphoglycerate (Harris and Waters 1976). Each subunit of this homotetrameric enzyme is folded into two separate domains, one of which, the N-terminal domain, binds the cofactor NAD^+. The C-terminal domain is responsible for the specificity of substrate binding and provides the catalytic groups. The intersubunit contacts of the tetramer are formed mainly by the so-called S-loops of the individual subunits, which also contribute to the formation of their NAD^+-binding pockets.

GAPDH has already been studied in some members of the Kinetoplastida. This order of protozoa is defined by the presence of the kinetoplast, a specialized subcom-

Correspondence to: Dr. V. Hannaert; *e-mail:* hannaert@trop.ucl.ac.be

partment of the single mitochondrion, containing a network of catenated DNA minicircles and maxicircles. During their long evolutionary history, the kinetoplastids have developed an impressive variety of life styles and adaptations to parasitism (Vickerman 1994). Two suborders are distinguished within this group: biflagellated Bodonina (including free-living and parasitic organisms) and uniflagellated Trypanosomatina (known only as parasites) (Vickerman 1976). Members' of the latter group are best known as the etiological agents of important diseases of humans and their domestic animals or cultivations: *Trypanosoma* and *Leishmania* cause severe diseases, including sleeping sickness, Chagas disease, and kala-azar, while all *Phytomonas* species are pathogens of plants. In addition to the kinetoplast, these organisms possess another unique organelle: the glycosome. The glycosome is a specialized peroxisome that houses most enzymes of the glycolytic chain (Opperdoes and Borst 1977). However, GAPDH is a glycolytic enzyme that has consistently been detected also in the cytosol, although the function of this isoenzyme is not known yet. In Trypanosomatidae, these two activities could be attributed to different isoenzymes which differ considerably in amino acid sequences and kinetic properties. In *T. brucei*, *T. cruzi*, and *L. mexicana*, the glycosomal GAPDHs are encoded by two tandemly linked identical genes (Michels et al. 1986; Kendall et al. 1990; Hannaert et al. 1992), whereas in the species analyzed, only a single gene coding for the cytosolic enzyme has been detected (Michels et al. 1991; Hannaert et al. 1992). The amino acid sequences of the isoenzymes are only 55% identical. All glycosomal GAPDHs present two striking features compared to their cytosolic counterparts. First, they possess some unique insertions and a C-terminal extension responsible for the relatively high molecular mass. This extension bears the SKL-like tripeptide or PTS-1 (peroxisome-targeting signal type 1) responsible for routing the protein to the organelle. Second, the glycosomal isoenzymes have a higher isoelectric point due to an excess of positively charged amino acids. It has been proposed that these positive residues could play a role in neutralizing the negative charges of the phosphorylated glycolytic metabolites present at high concentrations inside the glycosome (Hannaert and Michels 1994). Enzymological studies also revealed other unique characteristics of the glycosomal GAPDHs, most notably their apparent affinities for both NAD^+ and NADH, which are significantly lower than those measured for all other homologous enzymes (Lambeir et al. 1991; Hannaert et al. 1994). These differences in affinity may result from a looser packing around the adenine ring of NAD^+ for the glycosomal enzymes as shown by the crystal structures of *T. brucei* and *L. mexicana* GAPDHs (Vellieux et al. 1993; Kim et al. 1995). Compared to the human crystal structure (Mercer et al. 1976) more flexible residues are used in the trypanosomatids to wedge the purine ring. In addition, distal from C2 a small hydrophobic cavity is observed, which has been exploited in structure-based drug design studies with *T. brucei* and *L. mexicana* GAPDHs (Verlinde et al. 1994).

Some of the differences between the GAPDH isoenzymes may be attributed to their different localizations within the cell or to a different metabolic function but are due primarily to a different evolutionary origin. Indeed, in *Trypanoplasma borreli*, a representative of the Bodonina, only genes related to the glycosomal genes of Trypanosomatidae could be detected (Wiemer et al. 1995), although GAPDH activity was measured in both the cytosol and the glycosomes. *Euglena gracilis* does not have glycosomes but possesses a cytosolic GAPDH unambiguously homologous to the glycosomal isoenzyme of trypanosomes, sharing with it all the unique insertions (Henze et al. 1995). This free-living flagellated protozoan belongs to the euglenozoa group, one of the earliest-diverging eukaryotic lineages, monophyletically related to the kinetoplastids (Sogin et al. 1989; Hashimoto et al. 1995; Henze et al. 1995). These data suggest that the present-day glycosomal GAPDH was originally a cytosolic enzyme in the ancestor of the Kinetoplastida that, during its evolution, acquired a typical peroxisome-targeting signal at its C-terminal end. The gene coding for the cytosolic GAPDH found in the Trypanosomatidae has been acquired later by horizontal gene transfer. Contrary to the situation seen in *Trypanosoma* and *Leishmania*, the two tandemly linked glycosomal genes of *T. borreli* are 5% different. One of the two polypeptides present some substitutions, notably in the NAD^+-binding domain.

Because GAPDH is considered a highly relevant protein in terms of drug design and appeared to be a very interesting enzyme from an enzymological and molecular point of view, we decided to analyze this enzyme in other members of the Kinetoplastida representative for the various genera. The sequences so obtained have also been used to infer the phylogenetic relationships between these organisms.

Materials and Methods

Growth of Parasites and DNA Preparations

Phytomonas sp. isolated from *Euphorbia characias* (Dollet et al. 1982), *Crithidia fasciculata* (ATCC 11745), *Herpetomonas samuelpessoai* (ATCC 30252), *Leptomonas seymouri* (ATCC 30220), and promastigotes of *Leishmania major* (MHOM/SU/73/5-ASKH) were grown in vitro at 27°C in the semidefined medium SDM-79, supplemented with 10% heat-inactivated fetal calf serum (GIBCO, UK) (Brun and Schönenberger 1979). High molecular weight DNA was isolated as described by Van der Ploeg et al. (1982). Genomic DNA from *Trypanosoma brucei gambiense* (MA60), *Trypanosoma congolense* (MRC 33 KIBOKO), and *Trypanosoma vivax* were provided by Dr. E. Pays (Free University of Brussels, Belgium).

Cloning and Sequencing of the Glycosomal GAPDH Genes

For *Phytomonas* sp., *C. fasciculata*, *L. seymouri*, and *H. samuelpessoai*, genomic libraries were constructed in *Escherichia coli* MB406, using the phage vector λGEM-11 (Promega, USA) as described previously (Hannaert et al. 1992). The libraries were screened for recombinant clones containing glycosomal GAPDH genes by hybridization with the complete homologous gene of *L. mexicana* (Hannaert et al. 1992) labeled by nick translation. Hybridizations were performed under moderately stringent conditions: 3 × standard saline citrate (SSC)/0.1% sodium dodecyl sulfate (SDS), in the presence of 10% dextran sulfate at 65°C. Posthybridization washes were carried out at 65°C for 1 h with 5 × SSC/0.1% SDS, followed by 40 min with 3 × SSC/0.1% SDS and, finally, by 40 min with 1 × SSC/0.1% SDS. Positive plaques were purified and rescreened. High-titer phage lysates were obtained and DNA was purified from the phages as described (Marchand et al. 1989). Restriction fragments from the recombinant phages were subcloned in the phagemids pTZ18R and pTZ19R (Pharmacia LKB Biotechnology, Sweden). The bacterial host for recombinant plasmids was the *E. coli* strain XL1-blue (Strategene, USA).

For *T. b. gambiense*, *T. congolense*, *T. vivax*, and *L. major*, only partial sequences were obtained by PCR amplification. Based on a comparison of DNA sequences coding for cytosolic and glycosomal GAPDHs from other Kinetoplastida (*T. brucei*, *T. cruzi*, *L. mexicana*, and *Trypanoplasma borreli*), two degenerate oligonucleotides were designed. Primer 1 [5′ CGCGGATCCA(CG)GG(CT)CT(CT)(CA)TCGG(GTC)A(AC)(TG)GAGAT 3′] is homologous to a region specific for the glycosomal genes, whereas primer 2 [5′ CGCGGATCCCC(GTC)AC(GCA)GC(CT)TT(GC)GC(GC)GC(AG)CCAGT 3′] has been chosen in a well-conserved area in both the cytosolic and the glycosomal sequences. The glycosomal genes were amplified in 100 μl of 10 m*M* Tris–HCl, pH 7.4, 50 m*M* KCl, 6 m*M* $MgCl_2$, containing a 200 μ*M* concentration of each of the four deoxynucleotides, 100 pmol of each primer, 1 μg of genomic DNA, and 2.5 U of Taq DNA polymerase (Boehringer Mannheim, Germany). The following program was used for amplifications: first cycle of 3 min at 94°C, 1 min at 55°C, 20 s at 61°C, and 20 s at 72°C; followed by 10 cycles of 30 s at 94°C, 1 min at 55°C, 20 s at 61°C, and 20 s at 72°C; followed by 30 cycles of 30 s at 94°C, 1 min at 50°C, and 1 min at 72°C; and a final cycle of 30 s at 94°C, 1 min at 50°C, and 10 min at 72°C. The amplified fragments were purified, digested with *Bam*H1, and subcloned into the pTZ18R vector.

Nucleotide sequences were determined for both strands using the T7 DNA polymerase kit of Pharmacia and also with an automated sequencer (LI-COR; DNA sequencer model 4000) using the Thermo Sequenase fluorescent-labeled primer cycle sequencing kit of Amersham (UK). For each PCR fragment, the complete sequence was determined using two independent recombinant clones. Nucleotide sequences were deposited in the GenBank database under the following accession numbers: AF047499 (*T. brucei gambiense*), AF047498 (*T. congolense*), AF047500 (*T. vivax*), AF047497 (*L. major*), AF047496 (*Phytomonas* sp.), AF047494 (*H. samuelpessoai*), AF047495 (*L. seymouri*), and AF047493 (*C. fasciculata*).

Phylogenetic Analysis

Phylogenetic relationships were inferred from protein and DNA sequences using distance matrix and maximum-parsimony [PHYLIP package, Version 3.51c (Felsenstein 1993)] and maximum-likelihood methods. The distance method was used after correction for multiple substitutions according to Kimura (1980), in combination with the neighbor-joining method. The maximum-likelihood method with quartet puzzling was used as implicated in the program Puzzle, Version 2.5.1 (Strimmer and von Haeseler 1996), with the evolutionary model according to Dayhoff et al. (1978). Other GAPDH sequences used in this analysis were from *T. brucei* (GenBank accession No. X59955), *T. cruzi* (X52898), *L. mexicana* (X65226), *T. borreli* (X74535), and *Euglena gracilis* (L39772).

Table 1. Number of amino acids, calculated net charge, and (presumed) organelle targeting signal of glycosomal GAPDHs from different Kinetoplastida

	Length of polypeptide (N[a] amino acids)	Charge	C-Terminal tripeptide
T. brucei	359	+11	AKL
T. cruzi	359	+5	ARL
L. mexicana	362	+7	SKM
C. fasciculata	361	+6	SKL
L. seymouri	361	+8	SKM
Phytomonas sp.	361	0	AKL
H. samuelpessoai	361	+10	SKI
T. borreli	363	+8/+6	AKL

Results

Comparison of Kinetoplastid GAPDH Polypeptide Sequences

The glycosomal GAPDH gene sequences of a number of Kinetoplastida (*Trypanosoma brucei brucei*, *Trypanosoma cruzi*, *Leishmania mexicana*, and *Trypanoplasma borreli*) were already available (Michels et al. 1986; Kendall et al. 1990; Hannaert et al. 1992; Wiemer et al. 1995). Based upon well-conserved DNA regions, two degenerate oligonucleotides were selected and synthesized. These primers were used in a polymerase chain reaction to amplify GAPDH gene fragments from genomic DNA of other species: *T. brucei gambiense*, *T. congolense*, *T. vivax*, and *L. major*. The PCR-amplified fragments, representing approximately 54% of the coding region, were cloned and sequenced. For four others species (*Crithidia fasciculata*, *Herpetomonas samuelpessoai*, *Leptomonas seymouri*, and *Phytomonas* sp.), we decided to construct genomic libraries using λ phages and screen them with the homologous gene from *L. mexicana*. For two species (*C. fasciculata* and *H. samuelpessoai*), we first picked up the gene coding for the cytosolic isoenzyme, although the identity between the two GAPDHs in *T. brucei* and in *L. mexicana* is rather limited (55 and 56%, respectively). The nucleotide sequences of the glycosomal genes have been determined and the primary structure of the encoded proteins deduced. The complete polypeptide sequences have almost the same length, varying from 359 to 363 amino acids (Table 1). These small differences could be attributed mainly to variations at both the N and the C termini of the proteins. An optimal alignment of the predicted amino acid sequences of glycosomal GAPDHs from the kinetoplastid species with the *E. gracilis* sequence is pre-

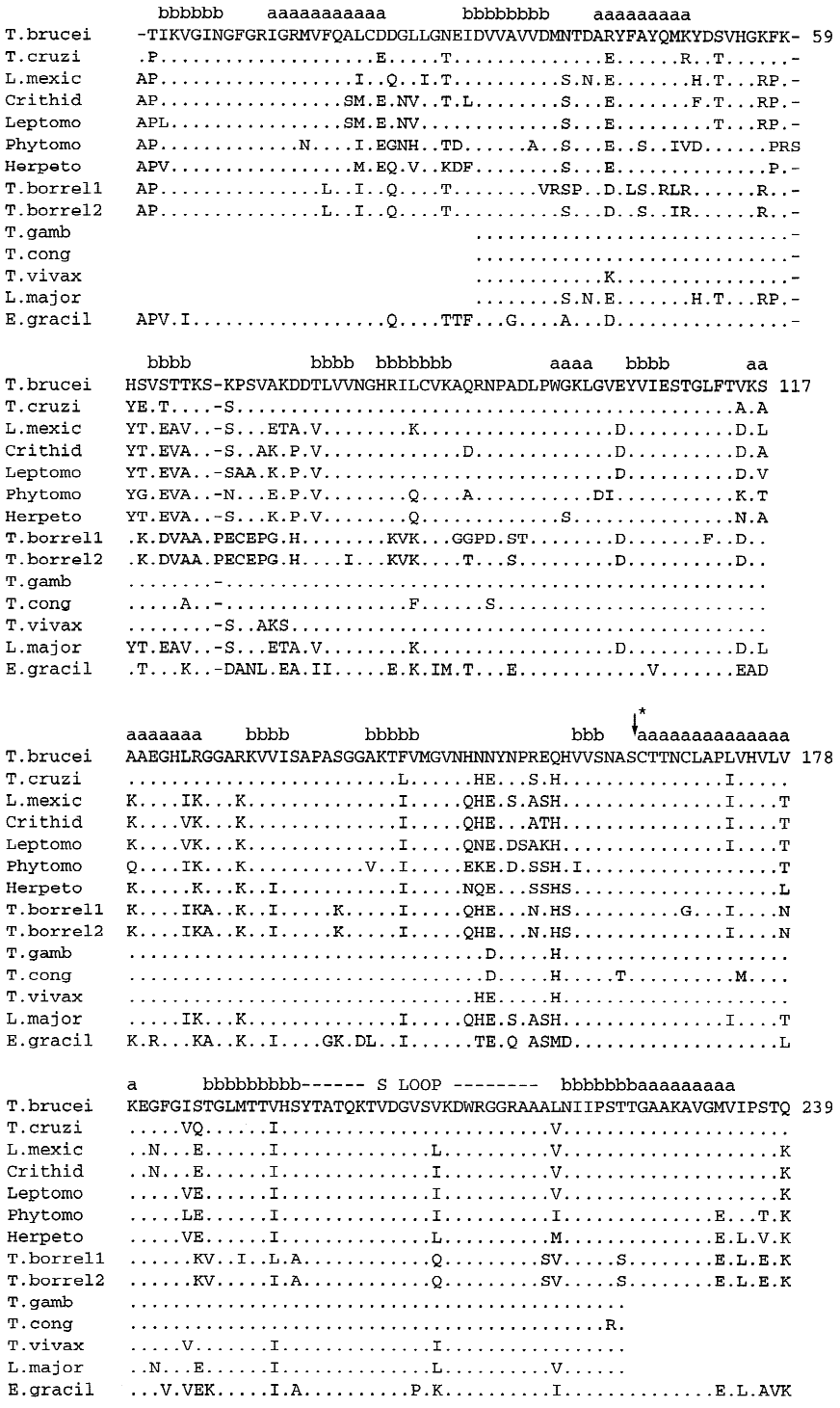

Fig. 1. Multiple alignment of the deduced amino acid sequences of glycosomal GAPDHs of Kinetoplastida and the closely related *E. gracilis* homologue. *Dots* represent amino acids that are identical to those at the corresponding positions in the *T. brucei* sequence. *Hyphens* indicate the absence of amino acids at the corresponding positions. Elements of regular secondary structure are shown as *a* (α-helix) and *b* (β-strand) and the S-loop is labeled. The active-site cysteine is indicated by the *asterisk*, and the boundary between the N- and the C-terminal domains by the *arrow*. The numbering is according to the *T. brucei* enzyme.

sented in Fig. 1. The GAPDH sequences from *Phytomonas* sp, *C. fasciculata*, *H. samuelpessoai*, and *L. seymouri* represent typical glycosomal proteins with specific peptide insertions at the same positions as described previously for the *T. brucei*, *T. cruzi*, *L. mexicana*, and *T. borreli* proteins. All enzymes are positively charged with the exception of the *Phytomonas* GAPDH, which is neutral, and all sequences present a short C-terminal extension carrying a PTS-1 motif (Table 1). The percentage of amino acid identity between the different

```
                bbbbbbbb      bbbbbbbb       aaaaaaaaaaaaa       bbbb
T.brucei    GKLTGMAFRVPTADVSVVDLTFIATRDTSIKEIDAALKRASKTYMKNILGYTDEELVSADF 300
T.cruzi     ......S.....P.........T.A.....Q.............G.............
L.mexic     ......S.....P.........R.Q.....Q...K.I.K.AQ....G...F.........
Crithid     ......S.....P.........R.......Q.......K..Q....G...F.......S..
Leptomo     ......S.....P.........R.......Q.......K.......G...F.........
Phytomo     ............P.........T..............K.V ..I.G..SF.......T..
Herpeto     ......S.....P.........T.AK........E.F.........G..S........S..
T.borrel1   ............P.........TT.K..............AAS.LRG..DISK.....T..
T.borrel2   ............P.........TT.K..............AAS.LRG..DISK.....T..
E.gracil    ............P.........L.EK........SL..K..Q..L.G...F.......T..

                bbbbaaaaaa          bbbbbbbbb aaaaaaaaaaaaaaaaaa
T.brucei    ISDSRSSIYDSKATLQNNLPNERRFFKIVSWYDNEWGYSHRVVDLVRHMAARD--RAAKL 357
T.cruzi     .N.N.................K.............................SK.--.S.R.
L.mexic     .N.N...V............G.K....V.........A..........Y...K.-AASS.M
Crithid     .N.A.................G.K....V...................F.G.K.-RSSS..
Leptomo     .N.A.................G.K....V...................F.G.K.-RASS.M
Phytomo     .N.NH.................K....L.................IIY..K..-AAS...
Herpeto     YNNNN................G.K....V..............N.LSF.IKK.-RAGS.I
T.borrel1   .HNPN......L........T.K....V...........N.......F.NSK.SKCH...
T.borrel2   .HNPN......L........T.K....V...........N.......F.NSK.SKSH...
E.gracil    VH.N.......L........G.K.L..V...........N.....LK..SGN
```

Fig. 1. Continued.

GAPDH sequences calculated from the alignment in Fig. 1 is given in Table 2. A very high percentage of identity (more than 90%) is observed between the two *Trypanosoma* species as well as between *C. fasciculata* and *L. seymouri.* The GAPDH sequence of *T. borreli* is, with values of about 75%, more distantly related to the Trypanosomatidae but clearly homologous.

Nucleotide Composition and Codon Usage of the GAPDH Genes

The overall nucleotide composition as well as the frequency of the nucleotides at the third codon position and the frequencies of each codon triplet was determined for the different GAPDH genes and is presented in Tables 3 and 4. In all species the GAPDH genes are relatively GC-rich except in *T. borreli,* where all nucleotides are equally well represented. Compared to the other species. *T. brucei* shows a lower GC content due to a smaller percentage of G. When the frequencies of the third base of the codons used are compared with the overall nucleotide composition, it is immediately apparent that the codon usage of the GAPDH genes of all species is biased. However, the nature of this bias is not the same in all Kinetoplastida: *T. borreli* favors a T or a C at the third position, whereas in *T. brucei* some bias exists only toward C and against A. A similar but much more pronounced effect can be observed for all other species. Moreover, these organisms also show preference for codons with a G and not a T in the wobble position. This bias is particularly strong in *C. fasciculata* and *L. seymouri,* where the GC content of the third base is higher than 95%. A remarkable feature of these two species, but also of *H. samuelpessoai,* is that the only triplet ending with an A is the stop codon (TAA), while a T is rarely incorporated at the degenerate position (in 9, 21, and 38 codons for *C. fasciculata, L. seymouri,* and *H. samuelpessoai,* respectively). Even for the arginine and leucine codons, in which the first position is degenerate, C is preferred at this position. These three species have almost-identical codon usage preferences which could not be attributed only to the GC preference at the third base position. Indeed, the order of usage for a group of degenerate codons is the same (except for proline and alanine, for which the order of preference in *H. samuelpessoai* is reversed compared to that in both other species). Another noteworthy feature of these three species is the highly biased usage of the 61 possible coding triplets compared to the other species: between 34 and 39 codons are used in these three genes whenever the other species employ more than 51 triplets (Table 4). In GAPDH of *C. fasciculata* 10 amino acids are each encoded by only one triplet. Except for the fact that these three insect parasites have a higher-biased codon usage, they show preferences very similar to those in *L. mexicana* and *Phytomonas* sp. We do not present here a detailed analysis of the codon usage of *L. mexicana* and both *Trypanosoma* species, since a number of such studies, all including the GAPDH data, have already been published (Michels 1986; Parsons et al. 1991; Alonso et al. 1992; Alvarez et al. 1994). Concerning the partial sequences, *L. major* shows the same codon usage as *L. mexicana,* whereas the preferences of the three *Trypanosoma* species are very close to that of *T. brucei* (data not shown). It is evident from the data in Table 4 that *T. borreli* is apart from the others, with a lower bias (57 triplets are used).

Analysis of Amino Acid Substitutions in the GAPDH Proteins

Based on the three-dimensional structures of the *T. brucei* and *L. mexicana* GAPDHs (Vellieux et al. 1993; Kim et al. 1995), amino acid substitutions in the various ki-

Table 2. Percentage of identity between the amino acid sequences of glycosomal GAPDHs from different Kinetoplastida

	T. bruc.	*T. cruzi*	*L. mex.*	*C. fasc.*	*L. sem.*	*Phyto.*	*H. samu.*	*T. bor.*
T. brucei	100	90.2	80.3	81.2	81.7	78.4	78.4	75.3
T. cruzi		100	84.3	85.7	85.4	80.6	82.3	76.7
L. mexicana			100	90.4	90.2	82.3	82.9	77.5
C. fasciculata				100	96.2	83.1	85.1	78.4
L. seymouri					100	83.1	85.4	77.2
Phytomonas sp.						100	82.6	75.6
H. samuelpessoai							100	77.0
T. borreli 1								100

Table 3. Nucleotide composition and frequencies of nucleotides at the third codon position of genes coding for glycosomal GAPDHs in Kinetoplastida

	Frequency of nucleotides (%)					Frequency of nucleotides at the third codon position (%)				
	A	G	T	C	GC	A	G	T	C	GC
T. brucei	21.65	27.15	21.65	29.55	56.70	8.76	25.77	23.19	42.27	68.04
T. cruzi	20.79	33.16	18.21	27.83	60.99	5.15	40.72	16.49	37.62	78.34
L. mexicana	21.48	33.85	16.67	28.00	61.85	3.60	45.88	11.34	39.17	85.05
C. fasciculata	19.80	33.80	15.00	31.40	65.20	0.28	46.41	2.49	50.82	97.23
L. seymouri	19.89	34.80	15.75	29.56	64.36	0.28	47.79	4.70	47.23	95.02
Phytomonas sp.	22.28	30.20	18.70	28.82	59.02	4.42	37.29	13.81	44.48	81.77
H. samuelpessoai	20.72	31.12	18.14	30.02	61.14	0.28	40.88	10.50	48.34	89.22
T. borreli 1	25.68	25.51	24.82	23.97	49.48	18.04	18.55	34.54	28.86	47.41

netoplastid sequences have been analyzed in a structural and functional context. The *T. borreli* GAPDH encoded by gene 1 has not been included in this study since this polypeptide is anomalous in several respects (Wiemer et al. 1994). In all species analyzed, the residues involved in catalysis (C165, H193), in substrate-phosphate binding (T196, A197, T198, R248), and in inorganic phosphate binding (S164, T166, T167, T225, G226, A227) are fully conserved. A comparison of the three-dimensional structures of the parasites' GAPDHs and the human enzyme has shown that the NAD^+-binding region is well conserved around the nicotinamide part of the cofactor, involved in the reaction mechanism. In contrast, the residues closest to the adenosine moiety of NAD^+ appear to be more loosely packed around the cofactor. The cleft and pockets so created have already been exploited for the design of selective inhibitors (Verlinde et al. 1994). Three residues in this particular region are variable in the Kinetoplastida. In *Phytomonas* an alanine is found at position 36 instead of a valine in the other GAPDHs. This residue does not contact NAD^+ directly but is very close to C2 of the adenosine part of the molecule. Residue 39 is an asparagine in the *Trypanosoma* species and a serine in all the others. This amino acid has an important position in the selectivity cleft for drug design, although it does not interact with the cofactor (Verlinde et al. 1994). At position 90, different amino acids are found (Q, D, A, T). In *T. brucei* and *L. mexicana* GAPDHs the glutamine makes a hydrogen bond to N6 of the NAD^+ adenine via its backbone carbonyl but not through the side chain. A good conservation is observed when comparing the subunit interface residues of the sequences: 82% identity (excluding the interface residues involved in catalysis, substrate, and NAD^+ binding), whereas the amino acids located at the surface present only 54% identity.

Phylogenetic Analysis

Phylogenetic analyses were carried out on the partial nucleotide as well as protein sequences of all 14 taxa and using *E. gracilis* as the outgroup. The results of a maximum-likelihood analysis based on either DNA or protein sequences are shown in Figs. 2A and B. All trypanosomes behaved as a monophyletic group with high quartet-puzzling and bootstrap support. Within this trypanosome clade, *T. cruzi* separated before *T. vivax* and the African trypanosomes, which formed a separate clade of closely related taxa. Although the branching order of *T. cruzi*, *T. vivax*, and the African trypanosomes was well established, the relationships within the African trypanosomes *T. brucei*, *T. b. gambiense*, and *T. congolense*, could not be determined. The other trypanosomatids formed the second clade in which *Leptomonas* and *Crithidia* appeared monophyletic with all methods used, but not always strongly supported by bootstrap analysis or quartet puzzling. *Phytomonas* and *Herpetomonas* behaved as either monophyletic or paraphyletic and different methods gave different branching orders for these

Table 4. Codon usage of glycosomal GAPDH genes in Kinetoplastida[a]

	T.b.	*T.c.*	*L.m.*	*Phyto.*	*C.f.*	*L.s.*	*H.s.*	*T.bo.*
Phe								
TTT	0	4	2	2	1	0	1	6
TTC	12	7	9	9	12	12	12	5
Leu								
TTA	0	0	0	0	0	0	0	1
TTG	1	2	2	3	0	0	0	5
CTT	5	7	3	2	1	2	3	3
CTC	12	5	3	7	4	4	4	6
CTA	1	0	0	2	0	0	0	1
CTG	4	8	9	6	14	12	14	8
Ile								
ATT	8	9	4	3	0	1	2	7
ATC	10	8	18	22	19	17	14	12
ATA	0	1	0	2	0	0	0	2
Met								
ATG	10	10	11	6	11	12	11	5
Val								
GTT	8	2	1	5	2	1	5	10
GTC	11	9	7	5	3	5	8	7
GTA	0	2	2	0	0	0	0	13
GTG	20	26	30	26	35	37	27	9
Ser								
TCT	4	4	4	3	1	1	5	6
TCC	8	5	8	9	10	9	10	11
TCA	2	2	0	0	0	0	0	2
TCG	4	5	6	7	9	9	8	2
AGT	2	2	0	0	0	0	0	1
AGC	7	7	8	7	8	7	8	5
Pro								
CCT	2	0	2	0	0	0	1	7
CCC	3	6	2	8	2	1	7	5
CCA	4	1	1	0	0	0	0	3
CCG	1	5	8	6	12	11	6	1
Thr								
ACT	9	3	3	5	0	1	2	9
ACC	6	8	8	8	13	7	10	6
ACA	5	2	1	4	0	0	0	10
ACG	7	16	16	12	15	18	12	0
Ala								
GCT	8	3	4	8	3	4	8	12
GCC	17	8	13	10	9	10	12	11
GCA	4	5	3	2	0	0	0	6
GCG	5	17	15	12	19	19	9	0
Tyr								
TAT	3	0	1	3	0	0	0	6
TAC	8	12	10	8	10	11	13	4
His								
CAT	1	2	0	1	0	0	0	8
CAC	9	9	10	9	9	8	7	3
Gln								
CAA	2	0	1	1	0	0	0	1
CAG	5	8	8	5	7	7	8	5
Asn								
AAT	3	3	2	1	0	0	0	8
AAC	14	11	12	15	14	14	18	8
Lys								
AAA	4	1	1	0	0	0	0	13
AAG	20	21	27	28	28	30	33	18
Asp								
GAT	8	5	3	6	1	3	4	9
GAC	14	16	18	17	22	20	14	13
Glu								
GAA	2	1	1	1	0	0	0	4
GAG	9	14	13	16	14	14	16	11
Cys								
TGT	0	1	1	1	0	0	0	2
TGC	4	3	3	3	4	4	4	3
Trp								
TGG	4	4	4	4	4	4	4	4
Arg								
CGT	8	6	2	3	0	0	3	5
CGC	10	10	12	8	15	15	8	7
CGA	0	0	0	0	0	0	0	0
CGG	2	2	0	0	0	0	0	1
AGA	0	0	0	0	0	0	0	1
AGG	0	1	0	1	0	0	0	0
Gly								
GGT	14	7	6	7	0	4	4	12
GGC	13	18	22	16	30	27	26	11
GGA	1	4	1	3	0	0	0	9
GGG	1	1	0	3	0	0	0	0
Stop								
TAA	1	1	0	1	1	1	1	1
TAG	0	0	0	0	0	0	0	0
TGA	0	0	1	0	0	0	0	0
No. of codons used	54	55	51	52	34	36	39	57

[a] *T.b.*, *Trypanosoma brucei*; *T.c.*, *Trypanosoma cruzi*; *L.m.*, *Leishmania mexicana*; *Phyto.*, *Phytomonas* sp.; *C.f.*, *Crithidia fasciculata*; *L.s.*, *Leptomonas seymouri*; *H.s.*, *Herpetomonas samuelpessoai*; *T.bo.*, *Trypanoplasma borreli* (gene 1).

two taxa, all with low bootstrap values, indicating that the precise order of events in this part of the tree could not be precisely determined. Identical results were obtained when the 10 taxa were used for which full-length sequences have been determined.

Discussion

Using the glycosomal GAPDH gene from *Leishmania mexicana* as a probe, we have cloned and sequenced the homologous genes from four species belonging to the order of Kinetoplastida: *Crithidia fasciculata*, *Leptomonas seymouri*, *Herpetomonas samuelpessoai*, and *Phytomonas* sp. Moreover, partial sequences have also been obtained by PCR for *Leishmania major*, *Trypanosoma brucei gambiense*, *Trypanosoma congolense*, and *Trypanosoma vivax*.

When analyzing the different GAPDH gene sequences, we found that *L. mexicana*, *C. fasciculata*, *L. seymouri*, *H. samuelpessoai*, and *Phytomonas* sp. share similar codon usage, different from the other species. Moreover, the sequences from *C. fasciculata*, *L. seymouri*, and *H. samuelpessoai* each exhibit a more biased codon usage toward the same set of preferred codons: only 34, 36, and 39 of the 61 possible coding triplets are used, respectively. Such a strong bias has also been found for the highly expressed GAPDH genes in yeast

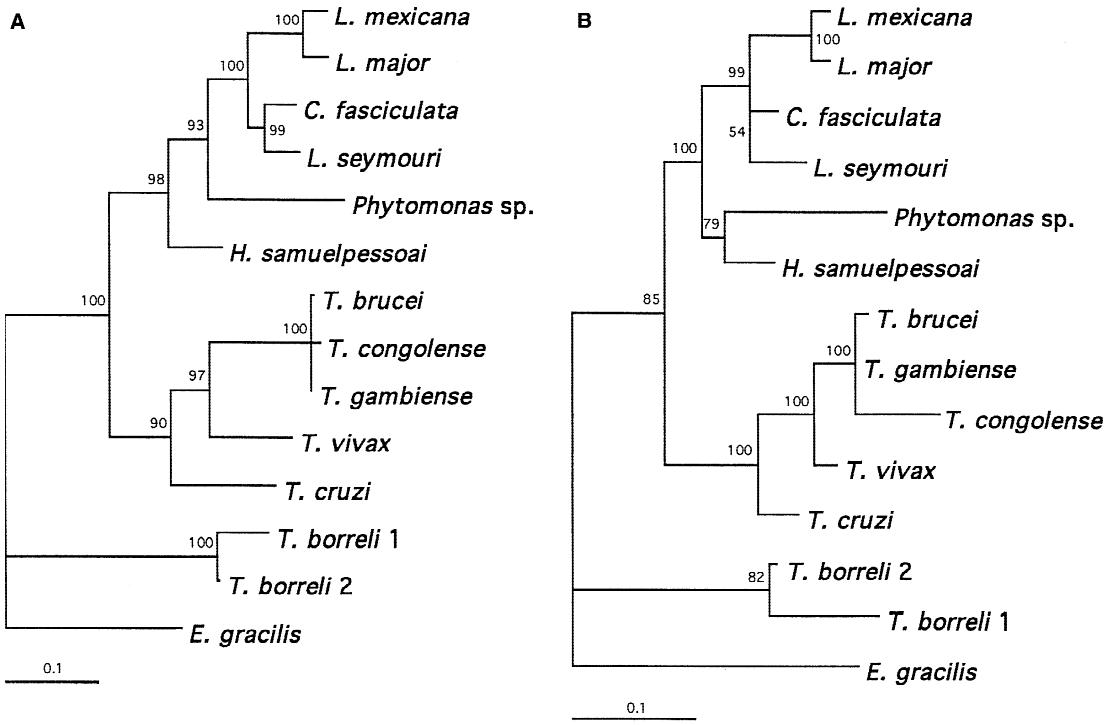

Fig. 2. Maximum-likelihood phylogenetic trees constructed from glycosomal GAPDH sequences of kinetoplastid protozoa. Trees were constructed on the basis of nucleotide **A** and deduced amino acid **B** sequences. The trees were rooted with the gene/amino acid sequence of the corresponding GAPDH of *E. gracilis.* Values indicated at the nodes represent estimates of support (percentages) for each internal branch based on 1000 quartet puzzling steps. For references to the methodology and the source of sequences used for the alignment and tree construction, see Materials and Methods.

(29 codons) and *E. coli* (40 codons) (Bennetzen and Hall 1982; Branlant and Branlant 1985). In both organisms, the preferred codons correspond with the most abundant isoacceptor tRNA species in the cell. In *T. brucei*, where GAPDH represents 0.5% of the total cellular proteins, the gene is composed of 54 different triplets (Misset et al. 1986; Michels et al. 1986). However, Alvarez et al. (1994) have shown that for *T. brucei* and for *T. cruzi*, the level of gene expression is not reflected in an extreme codon bias but, rather, in a trend toward the use of G- and C-ending triplets in the highly expressed genes. The same authors have shown that also in *Leishmania* species, genes which are strongly expressed present stronger codon biases. This conclusion could not be extended to *C. fasciculata:* the highly expressed gene coding for metalloproteinase and the dihydrofolate reductase–thymidylate synthase gene expressed at a very low level show the same codon usage and biases. Although the level of expression of the glycolytic enzymes has not been determined in *C. fasciculata*, we know that a relatively high glucose consumption rate has been measured in this organism [38 nmol min^{-1} (g protein)$^{-1}$]. This value is lower than what has been determined for the bloodstream form of *T. brucei* [80 nmol min^{-1} (g protein)$^{-1}$], which relies only an aerobic fermentation of glucose for its energy supply, but higher than in the procyclic insect form [9 nmol min^{-1} (g protein)$^{-1}$], also capable of oxidative phosphorylation (Michels et al. 1997). However, a high glucose consumption rate has also been measured for the *Phytomonas* sp. [50 nmol min^{-1} (g protein)$^{-1}$], which does not show such a strong codon bias. This organism, parasitizing the carbohydrate-rich fluids of plants, secretes cellulose-starch-, and pectin-degrading enzymes into the external medium and utilizes glucose, fructose, and mannose as the major energy substrates (Sanchez-Moreno et al. 1992). *C. fasciculata*, a monoxenous trypanosomatid of mosquitoes, on the contrary, is not able to degrade trehalose, the most abundant sugar in insect hemolymph (Cosgrove 1963). Moreover, the carbohydrate content of its environment, the insect's midgut, is in general quite low, whereas other oxidizable substrates such as amino acids and fatty acids are abundant. Sugars are present in blood and plants saps, the insect food, but disappear from the intestine soon after the meal (Evans and Brown 1972). From these considerations, it does not seem likely that *C. fasciculata* and, probably also, *Leptomonas* and *Herpetomonas*, which also live in the intestine of their insect hosts, rely heavily on carbohydrates for their energy supply and therefore should not need a high level of expression of their glycolytic enzymes. However, when we analyze the gene coding for another glycolytic enzyme of *C. fasciculata*, phosphoglycerate kinase, we can also observe a strong bias (only 43 different triplets are used) (Swinkels et al. 1992). The great similarity between these two glycosomal genes, in both the degree and the

direction of their codon preferences, suggests that this usage pattern has functional significance and is not merely a statistical anomaly. But this characteristic may not necessarily correlate with the level of expression of the genes. When analyzing the GAPDH genes from five parabasalid species, Viscogliosi and Müller (1998) also noticed a striking species-dependent range of the number of codons used (from 32 to 59), although all these organisms are glycolytic, probably with high glucose fluxes.

All the GAPDHs present the typical characteristics for glycosomal proteins: the presence of peptide insertions compared to the cytosolic enzymes, a high positive net charge (except the *Phytomonas* enzyme), and a PTS-1 motif. As expected, residues important for catalysis and substrate binding are fully conserved. Some variability is observed in the cofactor-binding pocket. However, the general structure of this region and, more precisely, the three areas that have been exploited for the design of selective inhibitors are probably maintained. The small hydrophobic pocket created by Val36 should not be drastically modified in *Phytomonas* by the presence of an alanine. The second area of interest in front of Leu112 is present in all GAPDHs. Finally, the hydrophobic cleft around the adenosine ribose 2′-hydroxyl group composed of residues 36–39 is also well conserved. The variability observed for residue 39 (asparagine or serine) had already been identified when inhibitors were tested on the *T. brucei* and *L. mexicana* enzymes and appeared not to prevent the inhibitory action (Verlinde et al. 1994). These results are encouraging regarding the development of future drugs. However, some caution is warranted because the variability suggests the possibility that variant parasites which are less sensitive already exist in nature and that resistance could spread upon introduction of the new drugs. Therefore, a search for more new targets, enabling combined drug therapy, seems relevant.

If the subunit interfaces of the proteins are reasonably well conserved, this is not the case for the surface residues. We noted already that the surface charge is not essential for the functioning of GAPDH inside the glycosome because it varies considerably. It seems now that the overall nature of the surface is less important. This also implies that it is unlikely that drugs could be targeted efficiently to surface residues. Resistance would, indeed, spread quickly due to the existing variability.

During the last decade, the development of novel molecular evolutionary methods has provided new means of assessing the kinetoplastid relationships and thus of unraveling the origin and evolution of parasitism. The majority of the studies on the phylogenetic inferences within the order Kinetoplastida were based on comparison of nuclear and kinetoplastid ribosomal RNA gene sequences. All rooted trees suggest that *T. brucei* and *T. cruzi* diverged before the lineage leading to *Crithidia, Leishmania,* and other trypanosomatid genera which are closely related. However, depending on the number, the origin, or the alignment of the sequences, the genus *Trypanosoma* appears to be either paraphyletic (Gomez et al. 1991; Fernandes et al. 1993; Landweber and Gilbert 1994; Maslov et al. 1994, 1996; Du et al. 1994; Lukes et al. 1994) or monophyletic (Hernandez et al. 1990; Berchtold et al. 1994; Marche et al. 1995). Recently, Lukes et al. (1997) suggested that the paraphyletic tree topology may be an artifact caused by a high rate of substitutions in the rRNA genes of the *T. brucei* and outgroup lineages. Previously we already proposed a phylogenetic tree based on GAPDH sequences suggesting the monophyly of the genus *Trypanosoma* (Hannaert et al. 1992). Comparisons of other protein coding genes also showed that *T. brucei* and *T. cruzi* are more closely related to one another than either species is to *Crithidia* or *Leishmania* (Hannaert 1995, Alvarez et al. 1996). Unfortunately, the number of available sequences for each of these proteins is, in general, limited to species belonging to the genus *Trypanosoma, Leishmania,* and *Crithidia.*

GAPDH is considered a suitable phylogenetic marker, because of its ubiquitous distribution and slow rate of evolution (Fothergill-Gilmore and Michels 1993). Most Kinetoplastida contain two GAPDH isoenzymes, one directly involved in glycolysis and located in the glycosomes and the other present in the cytosol. We used the glycosomal GAPDH to determine the phylogenetic relationships of these protists since this isoenzyme must have been the original one, already present in the common ancestor of the Kinetoplastida, whereas the other one was acquired by the lineage at a later stage (Wiemer et al. 1995). This analysis was performed with the DNA and amino acid sequences, either with the 10 complete or with the 14 partial kinetoplastid sequences, and the GAPDH of *E. gracilis* as the outgroup. Phylogenetic trees were constructed with several methods, and their reliability was tested by bootstrapping. Each of the methods resulted in essentially similar branching orders, with the exception of the *Phytomonas* and *Herpetomonas* lineages, of which the relationship is not reliably resolved by any reconstruction. They appear as either monophyletic or paraphyletic, always with little support. This problem was already encountered by others (Marche et al. 1995). More related sequences would probably be needed to draw conclusions about their phylogenetic position. Consistent statistical support was, however, obtained for the monophyly of all *Trypanosoma* species examined even with nucleotide sequences, indicating that the slightly different G+C content and codon usage used by *T. brucei* do not have a significant influence. The difficulty to establish relationships among the African trypanosomes is probably due to the high level of identity between the sequences (between 96 and 99%). Nevertheless, the Salivarian trypanosomes (*T. vivax, T. brucei, T. congolense,* and *T. gambiense*) form a well-supported cluster separating after *T. cruzi,* a representa-

tive of the Stercorarian trypanosomes in agreement with other trees (Lukes et al. 1997). As in other studies, the Bodonina is the first group to branch off during the kinetoplastid evolution. Thereafter, the family Trypanosomatidae is divided into two major clades: the *Trypanosoma* genus on one side and all the other genera (*Leishmania, Crithidia, Herpetomonas, Leptomonas,* and *Phytomonas*) on the other side. The inclusion of additional kinetoplastid sequences might enable us to define further the phylogenetic relationships within each of these major clades, to reduce the uncertainties associated with the position of certain branches, and to retrace in more detail the adaptation of all these organisms to their parasitic life style.

Acknowledgments. The authors express their gratitude to Ms. Dominique Cottem for her assistance in culturing the different species. We are grateful to Dr. Etienne Pays (Free University, Brussels, Belgium) for kindly providing some genomic DNA used in this study and to Drs. Christophe Verlinde and Stephen Suresh (University of Washington, Seattle) for their contribution in the mapping of the amino acid sequences into known three-dimensional GAPDH structures. We also thank Drs. Miklos Müller (Rockefeller University, New York) and Christophe Verlinde for critical comments on the manuscript. This investigation received financial support from the Belgian State-Prime Minister's office—Science Policy Programming (Grant 93/98-163).

References

Alonso G, Guevara P, Ramirez J (1992) Trypanosomatidae codon usage and GC distribution. Mem Inst Oswaldo Cruz Rio de Janeiro 87:517–523

Alvarez F, Robello C, Vignali M (1994) Evolution of codon usage and base contents in Kinetoplastid protozoans. Mol Biol Evol 11:790–802

Alvarez F, Cortinas MN, Musto H (1996) The analysis of protein coding genes suggests monophyly of *Trypanosoma.* Mol Phylogenet Evol 5:333–343

Bennetzen JL, Hall BD (1982) Codon selection in yeast. J Biol Chem 257:3026–3031

Berchtold M, Philippe H, Breunig A, Brugerolle G, Koenig H (1994) The phylogenetic position of *Dimastigella trypaniformis* within the parasitic kinetoplastids. Parasitol Res 80:672–679

Branlant G, Branlant C (1985) Nucleotide sequence of the *Escherichia coli gap* gene. Different evolutionary behavior of the NAD^+-binding domain and of the catalytic domain of D-glyceraldehyde-3-phosphate dehydrogenase. Eur J Biochem 150:61–66

Brun R, Schönenberger M (1979) Cultivation and *in vitro* cloning of procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. Acta Trop 36:289–292

Cosgrove WB (1963) Carbohydrate utilization by trypanosomatids from insects. Exp Parasitol 13:173–177

Dayhoff MO, Schwartz RM, Orcutt BC (1978) A model of evolutionary change in proteins. In: Dayhoff MO (ed) Atlas of protein sequence structure, Vol 5, Suppl 3. National Biomedical Research Foundation, Washington DC, pp 345–352

Dollet M, Cambrony D, Gargani D (1982) Culture axénique *in vitro* de *Phytomonas* sp. (Trypanosomatidae) d'Euphorbe, transmis par *Stenocephalus agalis* scop. (Coreide). CR Acad Sci 295:547–550

Du Y, Maslov DA, Chang K-P (1994) Monophyletic origin of β-division proteobacterial endosymbionts and their coevolution with insect trypanosomatid protozoa *Blastocrithidia culicis* and *Crithidia* spp. Proc Acad Natl Sci USA 91:8437–8441

Evans DA, Brown RC (1972) The utilization of glucose and proline by culture forms of *Trypanosoma brucei.* J Protozool 19:686–690

Felsenstein J (1993) PHYLIP (phylogeny inference package), Version 3.5c. Department of Genetics, University of Washington, Seattle

Fernandes AP, Nelson K, Beverley SM (1993) Evolution of nuclear ribosomal RNAs in kinetoplastid protozoa: perspectives on the age and origins of parasitism. Proc Natl Acad Sci USA 90:11608–11612

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

Gomez E, Valdes AM, Pinero D, Hernandez R (1991) What is a genus in the Trypanosomatidae family? Phylogenetic analysis of two small rRNA sequences. Mol Evol Biol 8:254–259

Hannaert V (1995) La glycéraldéhyde-3-phosphate déshydrogénase: une enzyme-clé chez les Kinetoplastida. Thèse de doctorat. Université Catholique de Louvain, Belgium

Hannaert V, Michels PAM (1994) Structure, function and biogenesis of glycosomes in Kinetoplastida. J Bioenerg Biomembr 26:205–212

Hannaert V, Blaauw M, Kohl L, Allert S, Opperdoes FR, Michels PAM (1992) Molecular analysis of the cytosolic and glycosomal glyceraldehyde-3-phosphate dehydrogenase in *Leishmania mexicana.* Mol Biochem Parasitol 55:115–126

Hannaert V, Callens M, Opperdoes FR, Michels PAM (1994) Purification and characterization of the native and the recombinant *Leishmania mexicana* glycosomal glyceraldehyde-3-phosphate dehydrogenase. Eur J Biochem 225:143–149

Harris JJ, Waters M (1976) Glyceraldehyde-3-phosphate dehydrogenase. In: Boyer PD (ed) The enzymes. Academic Press, New York, Vol 13, pp 1–49

Hashimoto T, Nakamura Y, Kamaishi T, Adachi J, Nakamura F, Okamoto K, Hasegawa M (1995) Phylogenetic place of kinetoplastid protozoa inferred from a protein phylogeny of elongation factor la. Mol Biochem Parasitol 70:181–185

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

Hernandes R, Rios P, Valdes AM, Pinero D (1990) Primary structure of *Trypanosoma cruzi* small-subunit ribosomal RNA coding region: comparison with other trypanosomatids. Mol Biochem Parasitol 41:207–212

Kendall G, Wilderspin AF, Ashall F, Miles MA, Kelly JM (1990) *Trypanosoma cruzi* glycosomal glyceraldehyde-3-phosphate dehydrogenase does not conform to the ''hotspot'' topogenic signal model. EMBO J 9:2751–2758

Kim H, Feil IK, Verlinde CL, Petra PH, Hol WG (1995) Crystal structure of glycosomal glyceraldehyde-3-phosphate dehydrogenase from *Leishmania mexicana:* implications for structure-based drug design and a new position for the inorganic phosphate binding site. Biochemistry 34:14975–14983

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120

Lambeir A-M, Loiseau AM, Kuntz DA, Vellieux FM, Michels PAM, Opperdoes FR (1991) The cytosolic and glycosomal glyceraldehyde-3-phosphate dehydrogenase from *Trypanosoma brucei.* Kinetic properties and comparison with homologous enzymes. Eur J Biochem 198:429–435

Landweber LF, Gilbert W (1994) Phylogenetic analysis of RNA editing: a primitive genetic phenomenon. Proc Natl Acad Sci USA 91:918–921

Lukes J, Arts GJ, Van den Burg J, De Haan A, Opperdoes F, Sloof P, Benne R (1994) Novel pattern of editing regions in mitochondrial transcripts of the cryptobiid *Trypanoplasma borreli.* EMBO J 13: 5086–5098

Lukes J, Jirku M, Dolezel D, Kral'ova I, Hollar L, Maslov DA (1997) Analysis of ribosomal RNA genes suggests that trypanosomes are monophyletic. J Mol Evol 44:521–527

Marchand M, Kooystra U, Wierenga RK, Lambeir AM, Van Beeumen J, Opperdoes FR, Michels PAM (1989) Glucosephosphate isomerase from *Trypanosoma brucei.* Cloning and characterization of the gene and analysis of the enzyme. Eur J Biochem 184:455–464

Marche S, Roth C, Philippe H, Dollet M, Baltz T (1995) Characterization and detection of plant trypanosomatids by sequence analysis of the small subunit ribosomal RNA gene. Mol Biochem Parasitol 71:15–26

Maslov DA, Avila HA, Lake JA, Simpson L (1994) Evolution of RNA editing in kinetoplastid protozoa. Nature 365:345–348

Maslov DA, Lukes J, Jirku M, Simpson L (1996) Phylogeny of trypanosomes as inferred from the small and large subunit rRNAs: implications for the evolution of parasitism in the trypanosomatid protozoa. Mol Biochem Parasitol 75:197–205

Mercer WD, Winn SI, Watson HC (1976) Twinning in crystals of human skeletal muscle D-glyceraldehyde-3-phosphate dehydrogenase. J Mol Biol 104:277–283

Michels PAM (1986) Evolutionary aspects of trypanosomes: analysis of genes. J Mol Evol 24:45–52

Michels PAM, Poliszczak A, Osinga KA, Misset O, Van Beeumen J, Wierenga RK, Borst P, Opperdoes FR (1986) Two tandemly linked identical genes code for the glycosomal glyceraldehyde-phosphate dehydrogenase in *Trypanosoma brucei.* EMBO J 5:1049–1056

Michels PAM, Marchand M, Kohl L, Allert S, Wierenga RK, Opperdoes FR (1991) The cytosolic and glycosomal isoenzymes of glyceraldehyde-3-phosphate dehydrogenase in *Trypanosoma brucei* have a distant evolutionary relationship. Eur J Biochem 198:421–428

Michels PAM, Hannaert V, Bakker BM (1997) Glycolysis of Kinetoplastida. In: Hide G, Mottram JC, Coombs GH, Holmes PH (eds) Trypanosomiasis and leishmaniasis. CAB International, Oxford, pp 133–148

Misset O, Bos OJM, Opperdoes FR (1986) Glycolytic enzymes of *Trypanosoma brucei:* simultaneous purification, intraglycosomal concentrations and physical properties. Eur J Biochem 157:441–453

Opperdoes FR, Borst P (1977) Localization of nine glycolytic enzymes in a microbody-like organelle in *Trypanosoma brucei:* the glycosome. FEBS Lett 143:360–364

Parsons M, Stuart K, Smiley BL (1991) *Trypanosoma brucei:* analysis of codon usage and nucleotide composition of nuclear genes. Exp Parasitol 73:101–105

Sanchez-Moreno M, Lasztity D, Coppens I, Opperdoes FR (1992) Characterization of carbohydrate metabolism and demonstration of glycosomes in a *Phytomonas* sp. isolated from *Euphorbia characias.* Mol Biochem Parasitol 54:185–200

Sogin M, Gunderson J, Elwood H, Alonso R, Peattie D (1989) Phylogenetic meaning of the kingdom concept: an unusual ribosomal RNA from *Giardia lamblia.* Science 243:75–77

Strimmer K, von Haeseler A (1996) Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. Mol Biol Evol 13:964–969

Swinkels BW, Loiseau A, Opperdoes FR, Borst P (1992) A phosphoglycerate kinase-related gene conserved between *Trypanosoma brucei* and *Crithidia fasciculata.* Mol Biochem Parasitol 50:69–78

Van der Ploeg LHT, Bernards A, Rijsewijk FAM, Borst P (1982) Characterization of the DNA duplication-transposition that controls the expression of two genes for variant surface glycoproteins in *Trypanosoma brucei.* Nucleic Acids Res 10:593–609

Vellieux FMD, Hajdu J, Verlinde CLMJ, Groendijk H, Read RJ, Greenhough TJ, Campbell JW, Kalk KH, Littlechild JA, Watson HC, Hol WGJ (1993) Structure of glycosomal glyceraldehyde-3-phosphate dehydrogenase from *Trypanosoma brucei* determined from Laue data. Proc Natl Acad Sci USA 90:2355–2359

Verlinde CLMJ, Callens M, Van Calenbergh S, Van Aerschot A, Herdewijn P, Hannaert V, Michels PAM, Opperdoes FR, Hol W (1994) Selective inhibition of trypanosomal glyceraldehyde-3-phosphate dehydrogenase by protein structure based design: toward new drug for the treatment of sleeping sickness. J Med Chem 37:3605–3613

Vickerman K (1976) The diversity of the kinetoplastid flagellates. In: Lumsden WHR, Evans DA (eds) Biology of the Kinetoplastida. Academic Press, London, pp 11–34

Vickerman K (1994) The evolutionary expansion of the trypanosomatid flagellates. Int J Parasitol 24:1317–1331

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

Wiemer EAC, Hannaert V, van den IJssel P, Van Roy J, Opperdoes FR, Michels PAM (1995) Molecular analysis of glyceraldehyde-3-phosphate dehydrogenase in *Trypanoplasma borelli:* an evolutionary scenario of subcellular compartmentation in Kinetoplastida. J Mol Evol 40:443–454