Enolase from Trypanosoma brucei, from the Ameitochondriate Protist Mastigamoeba balamuthi, and from the Chloroplast and Cytosol of Euglena gracilis: Pieces in the Evolutionary Puzzle of the Eukaryotic Glycolytic Pathway

Véronique Hannaert,* Henner Brinkmann,† Ulrich Nowitzki,‡ Jennifer A. Lee,§ Marie-Astrid Albert,* Christoph W. Sensen,¶ Terry Gaasterland,§ Miklós Müller,§ Paul Michels,* and William Martin‡

*Research Unit for Tropical Diseases, Christian de Duve Institute of Cellular Pathology, Department of Biochemistry, Université catholique de Louvain, Brussels, Belgium; †Institute of Evolutionary Biology, Department of Biology, Universität Konstanz, Konstanz, Germany; ‡Institut für Botanik, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany; §The Rockefeller University, New York; and ¶Institute of Marine Biosciences, National Research Council, Halifax, Nova Scotia, Canada

Genomic or cDNA clones for the glycolytic enzyme enolase were isolated from the amitochondriate pelobiont Mastigamoeba balamuthi, from the kinetoplastid Trypanosoma brucei, and from the euglenid Euglena gracilis. Clones for the cytosolic enzyme were found in all three organisms, whereas Euglena was found to also express mRNA for a second isoenzyme that possesses a putative N-terminal plastid-targeting peptide and is probably targeted to the chloroplast. Database searching revealed that Arabidopsis also possesses a second enolase gene that encodes an N-terminal extension and is likely targeted to the chloroplast. A phylogeny of enolase amino acid sequences from 6 archaeabacteria, 24 eubacteria, and 32 eukaryotes showed that the Mastigamoeba enolase tended to branch with its homologs from Trypanosoma and from the amitochondriate protist Entamoeba histolytica. The compartment-specific isoenzymes in Euglena arose through a gene duplication independent of that which gave rise to the compartment-specific isoenzymes in Arabidopsis, as evidenced by the finding that the Euglena enolases are more similar to the homolog from the eubacterium Treponema pallidum than they are to homologs from any other organism sampled. In marked contrast to all other glycolytic enzymes studied to date, enolases from all eukaryotes surveyed here (except Euglena) are not markedly more similar to eubacterial than to archaeabacterial homologs. An intriguing indel shared by enolase from eukaryotes, from the archaeabacterium Methanococcus jannaschii, and from the eubacterium Campylobacter jejuni maps to the surface of the three-dimensional structure of the enzyme and appears to have occurred at the same position in parallel in independent lineages.

Introduction

The Embden-Meyerhof-Parnas pathway of glycolysis is the backbone of energy metabolism (ATP synthesis) in eukaryotes. Enolase (2-phospho-D-glycerate hydratase; EC 4.2.1.11) catalyzes the Mg2+-dependent dehydration of 2-phosphoglycerate to phosphoenolpyruvate, which is converted by pyruvate kinase into pyruvate with the concomitant generation of ATP in the subsequent, final step of glycolysis. Enolase also catalyzes the reverse reaction during gluconeogenesis, but in addition appears to have been recruited for other functions in at least two lineages: it is a major lens protein of vertebrates (Wistow et al. 1988), and it plays a role in the acquisition of thermal tolerance in yeast (Iida and Yahara 1985). The enzyme from a large variety of organisms, including archaeabacteria, eubacteria, and eukaryotes has been studied and is highly conserved. Although octameric enzymes have been described for Bacillus subtilis (Brown et al. 1998) and Thermotoga maritima (Schurig et al. 1995), in all eukaryotes and many prokaryotes, enolase is biologically active as a dimer, with subunits having an M, of approximately 45,000. Catalytically active dimers may be generated with products of different enolase genes, for example, in Saccharomyces cerevisiae (McAlister and Holland 1982). Crystal structures for the enzyme are known from lobster and S. cerevisiae at 2.2 and 2.0 Å resolution, respectively (Duquerroy, Camus, and Janin 1995; Zhang et al. 1997).

In the vast majority of eukaryotes studied to date, glycolysis is a cytosolic pathway. This is always true in eukaryotes that lack organellar compartments involved in ATP synthesis, such as Giardia lamblia (Müller 1998) or Mastigamoeba balamuthi, an amoeboid flagellate protist that possesses a strikingly barren cytoplasm and lacks morphologically recognizable mitochondria (Chavéz, Balamuth, and Gong 1986; Brugerolle 1993). However, among other protists, there are some notable exceptions to be found with regard to the compartmentation of glycolysis. In Chlamydomonas reinhardtii, for example, all glycolytic enzyme activities studied to date were found to be localized in the chloroplast, rather than in the cytosol (Schnarrenberger et al. 1990). A better-known and well-studied exception is the kinetoplastids (trypanosomes and related organisms), in which most of the gly-
colytic pathway is compartmentalized in specialized microbodies, glycosomes (Opperdoes and Borst 1977; Hannaert and Michels 1994; Clayton and Michels 1996). In these organisms, only the last three enzymes of the pathway leading to pyruvate production, including enolase, are found in the cytosol. In relatives of the kinetoplastids, the euconoids, which possess plastids but not glycosomes, some of the enzymes that are common to glycolysis and the Calvin cycle occur as distinct chloroplast-cytosol isoenzyme pairs (Kitaoka et al. 1989; Henze et al. 1995; Plaumann et al. 1997), as is also found in many higher plants (Martin and Schnarrenberger 1997). Several key glycolytic enzymes from trypanosomes and amitochondriate protists differ in their regulatory properties from those of other eukaryotes, e.g., yeast or vertebrates, differences that appear to correlate with the lifestyle of these organisms (Mertens, van Schaftingen, and Müller 1992; Mertens 1993; Hannaert and Michels 1994; Michels and Hannaert 1994; Clayton and Michels 1996; Bakker et al. 1997; Park, Schofield, and Edwards 1997).

The study of the evolution of pathways requires sequences for all of the enzymes involved. For the kinetoplastid Trypanosoma brucei, all glycolytic enzymes except enolase have been cloned. For amitochondriate protists and Euglena, whose glycolytic pathways differ from kinetoplastids not only with respect to compartmentation, but also with respect to regulation of the enzymes involved (Kitaoka et al. 1989; Bakker et al. 1997), fewer sequences for glycolytic enzymes are known. Here we report the molecular analysis of cytosolic enolase from the amitochondriate protist M. balamuthi and T. brucei and two isoenzymes of enolase from Euglena gracilis, one of which possesses a putative N-terminal plastid-targeting peptide. A protein phylogeny of enolase sequences from eubacterial, archaeabacterial, and eukaryotic sources is presented that reveals insights into the complex evolutionary history and subcellular compartmentation of this glycolytic enzyme.

Material and Methods
Cloning and Molecular Analysis of T. brucei Enolase

A search in the EST subset of the GenBank nucleic acid database revealed a 369-bp T. brucei rhodesiense sequence (accession number AA023828; cDNA clone number T3246) coding for an amino acid sequence of which 122 residues showed 55% identity with the C-terminal part of human enolase. Based on this sequence, two oligonucleotides containing a BamH1 site (italicized) were designed: 5′-CGCGGATCCCACAATTGG-CAGTATTGAAGACCC-3′ and 5′-CGCGGATCCCCTCTGTCTCACACTGCGG-3′. PCR amplification was performed using genomic DNA from T. brucei stock 427 as template with initial denaturation for 5 min at 95°C, followed by 30 cycles of 30 s at 95°C, 1 min at 60°C, and 1 min at 72°C and, finally, 10 min of incubation at 72°C. A major amplified product of 300 bp was cloned into pTZ19R (Amersham Pharmacia Biotech, Sweden), sequenced, and used to screen a genomic library of T. brucei in Escherichia coli in λGEM11 (Promega) (Michels et al. 1991). Hybridization was performed at stringent conditions as described previously (Kohl et al. 1996). Six positive recombinant phages were purified and rescreened. A 2.5-kb PstI fragment from one of the hybridizing phages was subcloned into pTZ19R and sequenced on both strands.

Cloning and Molecular Analysis of E. gracilis Enolase Isoenzymes

Axenic E. gracilis cultures (SAG 1224-5/25) were grown as described (Henze et al. 1995) under a 14:10 h light:dark regime. Isolation of mRNA through two rounds of oligo-dT cellulose chromatography and synthesis of cDNA for amplification was performed as described (Henze et al. 1995) but without the NotI adapters. Degenerate primers 1F (5′-GAYWSMGIGGIAAYC-CIACTIGIGAR-3′) and 3R (5′RTCYTCITYTCCISS-WICKRTG-3′) were designed against the conserved amino acid sequence motifs “DSRNPNPTEV” and “HRSGETED” from the N-terminal and C-terminal regions, respectively, of an alignment of enolase sequences extracted from GenBank. Amplification was performed for 35 cycles of 2 min at 94°C, 1 min at 50°C, and 2 min at 72°C in 25 µl containing 10 µM of each primer, 50 µM dNTP, 1 mM MgCl2, 10 ng cDNA, and 0.5 U Taq polymerase (Perkin-Elmer) using the supplier’s buffer. The 1.1-kb fragment was subcloned and used to screen 300,000 recombinants of an Euglena cDNA library (Henze et al. 1995). From 150 hybridizing positives, 15 were subcloned and sequenced. One corresponded to pEgEno29. The other 14 represented an mRNA encoding a different enolase sequence, the largest clone of which was designated pSoEno02.

For Northern blots, inserts of pEgEno02 and pSoEno29 were purified by electrolution. Five micrograms of polyA+ mRNA purified through two rounds of oligo-dT cellulose affinity chromatography was loaded per lane on the same formaldehyde gel, electrophoresed as described (Sambrook, Fritsch, and Maniatis 1989), and transferred to Hybond N (Amershams) membranes according to the manufacturer’s protocol. Lanes were stained, photographed, cut, and hybridized separately at 70 °C in 3 × SSPE, 0.1% SDS, 0.02% (w/v) PVP, 0.02% (w/v) Ficoll 400, 50 µg/ml polyadenylic acid, and 10 ng/ml of probe labeled to 5 × 107 cpm/µg. Washing was at 70 °C in 2 × SSPE, 0.1% SDS.

Spinach Enolase

Using cDNA from 10-day-old spinach seedlings, the primers 1F and 3R, described above for Euglena, were used under the same conditions to amplify a 1.1-kb fragment from spinach. The amplification product was cloned as above and used to screen a spinach cDNA library prepared as described (Nowitzki et al. 1998). Ten independent positive clones were identified, subcloned, and shown by terminal sequencing to represent the same mRNA. The longest of these, pSoEno8.3, was sequenced.
Enolase from *M. balamuthi*

*Mastigamoeba balamuthi* identical to the American Type Culture Collection isolate ATCC 30984, was obtained from Dr. Frederick Schuster (Brooklyn College of the City University of New York, Brooklyn, N.Y.). cDNA clone PHR60, containing the complete ORF for enolase, was identified in our ongoing *M. balamuthi* cDNA sequencing project, to be described elsewhere. The insert in this clone was sequenced on both strands by primer walking.

**Sequence Handling and Phylogenetic Analysis**

From sequences extracted from GenBank, an alignment was produced with CLUSTAL W (Thompson, Higgins, and Gibson 1994) that was manually refined using the program ED of the MUST package (Philippe 1993). Regions of uncertain alignment were omitted from the analysis, leaving 400 amino acid positions for analysis. Protein phylogeny was inferred using protein maximum-likelihood (ML) as implemented in MOLPHY, version 2.3 (Adachi and Hasegawa 1996), using local rearrangement starting with the neighbor-joining (NJ) tree and using the best tree obtained from 1,000 PROTML topologies generated using the quicksearch option of MOLPHY using the JTT-F model. Distance analyses were performed with the MUST package using the Kimura correction and the NJ method (Saitou and Nei 1987) using 1,000 bootstrap replicates. Parsimony analysis was performed using PAUP* with 500 bootstrap replicates and 10 times random addition (Swofford 1999). The significance of different ML topologies was tested using the Kishino-Hasegawa test implemented in MOLPHY. Accession numbers for all sequences and the alignment are available via anonymous ftp from 134.169.70.80/ftp/pub/incoming/enolase.

**Molecular Modeling**

The structure of yeast enolase (Wedekind, Reed, and Rayment 1994) was downloaded from the Protein Data Bank (http://www.rcsb.org/pdb/) and visualized using RASMOL (http://www.umass.edu/microbio/rasmol/) on a Macintosh computer.

**Results**

Enolase from *T. brucei*

Based on an EST sequence of a *T. brucei rhodesiense* cDNA fragment available in public databases, oligonucleotides were designed to amplify a homologous hybridization probe from genomic DNA of *T. brucei*. The amplified fragment was identical in sequence to that of the partial cDNA of *T. b. rhodesiense* present in the EST database and was used to screen a lambdaGE11 *T. brucei* genomic DNA library. Among hybridizing clones, restriction mapping revealed a 2.5-kb *Pst*I fragment that contained an open reading frame (ORF) for enolase encoding a polypeptide of 428 amino acids (excluding the initiator methionine) with a molecular mass of 46,461, a pl of 6.2, and a net charge of −5.

Enolase from the Amitochondriate Protist *M. balamuthi*

Of the approximately 200 clones sequenced from a random cDNA library of *M. balamuthi*, six clones encoded a typical enolase. Of these, one clone was completely sequenced and found to code for a putative enolase of 438 amino acid residues with a molecular mass of 47,596, a pl of 6.7, and a net charge of −3.

Cytosolic and Plastid-Targeting Peptide-Bearing Isoforms of Enolase in *Euglena*

Grown in rich medium in the light under aerobic conditions, *Euglena* expresses two different genes for enolase. The sequence of pEGEno02 is 1,523 bp long, including a 21-base polyA stretch, and encodes an ORF of 431 amino acids with a molecular mass of 46,501, a pl of 5.79, and a net charge of −7. pEGEno29 is 1,622 bp long with a 10-base polyA tail and encodes an ORF of 485 amino acids that lacks a start codon near the N-terminus of pEGEno2. Instead, the reading frame extends beyond the start codon of pEGEno2 by at least 53 amino acids. This region extends into, but terminates in, a highly hydrophobic region (fig. 1). The plastid targeting-peptides of nuclear-encoded chloroplast proteins in *Euglena* are roughly 120–150 amino acids long and possess two highly hydrophobic regions (Plaumann et al. 1994).
alignment and Phylogenetic Analysis

The alignment of enolase sequences revealed two length polymorphisms with notable phylogenetic distribution, as shown in figure 3 (see Discussion). One thousand trees were generated with the quicksearch option of MOLPHY and evaluated using the JTT-F model. Using the Kishino-Hasegawa test implemented in MOLPHY, 585 of the trees were not significantly worse than the best tree found, which had a log likelihood (lnL) of \(-26,714.8\). Using that topology for local rearrangement with the JTT-F model, the topology in figure 4 was found, the lnL of which (\(-26,671.97\)) was slightly, but insignificantly, better than the lnL value for the starting topology. Using the NJ tree of Kimura distances as the starting topology, an ML tree was found that also was not significantly different at the 5% level from the topology in figure 4.

Discussion

Indels in Enolase Sequences and Evidence for Plastid Isoforms in *Euglena* and *Arabidopsis*

Enolase is one of the most conservatively evolving glycolytic enzymes (Fothergill-Gilmore and Michels 1993). With the availability of many new enolase sequences, including those from sequenced genomes, it was of interest to address the gene phylogeny of this enzyme. We extracted enolase homologs from the databases and produced an alignment encompassing 62 sequences and 400 sites. Overall, enolase sequences are highly conserved, showing roughly 40% amino acid identity across eukaryotes, eubacteria, and archaebacteria. Two notable indels found in enolase amino acid sequences are shown in figure 3.

The first of these, at positions 161–165 of our alignment, underscores the specificity of a five-aminoc acid insertion previously noted to be shared in cytosolic enolase from higher plants and enolase from the apicomplexans *Plasmodium falciparum* (Read et al. 1994) and *Toxoplasma gondii* (Dzierzinski et al. 1999). This indel suggests that apicomplexan parasites, which possess plastids (McFadden et al. 1996; McFadden and Roos 1999), probably obtained their nuclear enolase gene from the photosynthetic symbiont (Read et al. 1994; Dzierzinski et al. 1999), a view that receives additional support from the current analysis, due to the highly restricted phylogenetic distribution of the indel.
Interestingly, two enolase sequences from plants, that from *Chlamydomonas* and a new enolase sequence from the *Arabidopsis* genome project (*Arabidopsis* 2 in fig. 3), do not share the insertion at positions 161–165. The *Arabidopsis* 2 protein also carries a transit peptide, suggesting that it represents a plastid isoform. Many higher plant tissues are known to possess distinct and separable chloroplast-cytosol isoenzymes of enolase, for example, developing *Ricinus* seeds (Miernyk and Dennis 1992) and developing barley leaves (Hoppe et al. 1993), but the plastid enzyme is usually not highly expressed in fully green tissues. A previous study indicated the lack of a plastid enolase in *Arabidopsis* (van der Straeten et al. 1991), but it seems likely that the plastid enzyme is expressed only in early stages of *Arabidopsis* development, as is the case in barley (Hoppe et al. 1993). Like the probable plastid isoform from *Arabidopsis*, the *Chlamydomonas* enolase also lacks the five–amino acid insertion characteristic of plant (and apicomplexan) cytosolic enolase. Furthermore, a study of subcellular distribution of several glycolytic enzymes in *Chlamydomonas* has shown that only a single isoform for each of the glycolytic enzymes glucose-6-phosphate isomerase (GPI), FBA, phosphoglycerate kinase (PGK), and TPI exists and that each is localized in the chloroplast (Schnarrenberger et al. 1990). Although enolase

---

**FIG. 3.**—Conspicuous indels in enolase sequences. Position numbers in the alignment are indicated. Gaps are shown as dashes. Enolase sequences from completely sequenced genomes are indicated by asterisks.
Fig. 4.—Phylogeny of enolase protein sequences constructed by maximum likelihood (ML) (Adachi and Hasegawa 1996). Numbers at nodes indicate bootstrap proportions, upper values indicate Kimura distances with neighbor-joining (1,000 replicates), and lower numbers indicate parsimony (see Materials and Methods). Solid dots indicate branches that were found in all 585 of the PROTML trees identified from quicksearch of MOLPHY that were not significantly worse than the best tree found at the 5% level. The scale bar at lower left indicates 10 substitutions. Sequences presented in this paper are given in boldface type. An asterisk indicates that the probable chloroplast enolase isoforms from *Chlamydomonas*, *Arabidopsis*, and *Euglena* have not been directly localized to the plastid (see text). Eubacterial groups as recognized by rRNA systematics are indicated.
was not assayed in that study, the lack of cytosolic isoenzymes for those glycolytic activities indicates that glycolysis in light-grown *Chlamydomonas* is primarily localized in the plastid, suggesting that the *Chlamydomonas* enolase is also a plastid isofrom, consistent with its shared lack of the insertion at positions 161–165. Since the *Chlamydomonas* enolase sequence in the database lacks about 50 residues of the mature subunit, it is possible that the complete clone may possess a transit peptide.

Another notable indel occurs at positions 305–306 of the alignment. All enolase sequences from eukaryotic sources lack the motif “PG” (or its variants) at this position, with the notable exception of the Euglena isoenzymes. All prokaryotic sequences surveyed possess the PG motif, with the exception of the archaeabacterium *Methanococcus* and the eubacterium *Campylobacter*. However, the positions of *Methanococcus*, *Campylobacter*, and eukaryotic enolases in the gene phylogeny (fig. 4) suggest that this indel does not represent a unique event in enolase evolution, but rather that this deletion has occurred several times independently. That independent indel events can occur at this position is furthermore seen by the unique event in the *Chlorobium* 1 sequence (fig. 3).

**Conflict Between Indel Distribution and Protein Phylogeny in Plasmodium Enolase**

The conspicuous indel in figure 3 that indicates common ancestry of the *Plasmodium* and plant cytosolic enolases is not consistent with the position of *Plasmodium* in the phylogeny of figure 4. Instead of branching with plant cytosolic enolase, the *Plasmodium* sequence tends to assume a basal position among eukaryotic homologs. Also in distance and parsimony analyses, the *Plasmodium* sequence did not branch with the plant homologs, but rather assumed a basal position, whereby it branched above the branch bearing *Trypanosoma*, *Mastigamoeba*, and *Entamoeba* in parsimony analysis and branched with *Trypanosoma* and *Entamoeba* when *Mastigamoeba* was excluded in PROTML.

The nature of the indel at positions 161–165 (fig. 3) suggests that *Plasmodium* enolase does indeed share a common ancestry with the plant cytosolic sequences and that its position in figure 4 may be a reconstruction artifact. Overall, the position of the *Plasmodium* sequence was rather unstable, and it seems likely that its basal position in figure 4 is due to a long-branch attraction artifact and is hence a misplacement. Using the Kishino-Hasegawa test, PROTML trees in which *Plasmodium* was forced with the plant sequences were not worse than the tree in figure 4 at the 5% level. In parsimony analyses of smaller data sets, consisting of 45 sequences through exclusion of many eubacterial sequences, the *Plasmodium* sequence shared a common branch with its homologs from higher plants, but with a bootstrap proportion of only 46%. In an analysis of the enolase sequence from the apicomplexan *T. gondii*, Dzierszinski et al. (1999) also found that with smaller data sets, the apicomplexan enolase sequences branched with plant homologs in parsimony analyses.

Among the 585 trees found with the quicksearch option of MOLPHY that were not significantly worse at the 5% level than the best tree found, only 23 grouped *Plasmodium* with the plant sequences. The position of *Plasmodium* enolase is thus difficult to resolve in phylogenetic inference from sequence data, with the conflict between the specific indel (fig. 3) and the position of *Plasmodium* enolase in the phylogeny suggesting that it is simply misplaced in figure 4. In light of the plastid that has been found in apicomplexans (McFadden et al. 1996; McFadden and Roos 1999), a link between apicomplexan and plant enolase, as evidenced by the indel at positions 161–165, is not surprising. Indeed, in analyses of a different glycolytic enzyme, GPI, sequences from apicomplexans did branch specifically with their plant homologs in PROTML analyses (Nowitzki et al. 1998; Dzierszinski et al. 1999).

**Enolases from Amitochondriate Protists Branch with the *Trypanosoma* Homolog**

The detection of enolase transcripts indicates that the core metabolism of *M. balamuthi* involves an Embden-Meyerhof-Parnas–type glycolysis. This conclusion is supported by the identification of cDNA clones from the EST project encoding PP\(_i\)-linked phosphofructokinase (PP\(_i\)-PFK), class II fructose 1,6-bisphosphate aldolase, GAPDH, and phosphoglycerate mutase (PGM) (unpublished data). The organism contains no mitochondrion- or hydrogenosome-like structures (Chavéz, Balamuth, and Gong 1986; Brugerolle 1993) and grows in a cysteine-rich anaerobic medium, and lactate is one of its major end products (unpublished data). These characteristics permit an assignment of *M. balamuthi* to type I amitochondriate protists characterized by lack of compartmentalized energy metabolism (Martin and Müller 1998; Müller 1998).

*Mastigamoeba balamuthi* is a member of the amoeboflagellate group of pelobionts, which also contains the giant amitochondriate amoeba *Pelomyxa palustris*. Because of its simple cytological makeup, this group of eukaryotes is often regarded as “primitive” and is placed with other mitochondrion-lacking eukaryotes in some taxonomic schemes (Cavalier-Smith 1991, 1993). The evolutionary position of these organisms, however, remains unresolved (Embley and Hirt 1998). While some data indicate an ancestral (basal) position of this group among eukaryotes (Stiller, Dufield, and Hall 1998), others suggest a derived, nonbasal position (Hinkle at al. 1994). The *Mastigamoeba* sequence branches with *Trypanosoma* and *Entamoeba* enolase also in distance and parsimony analyses. The limited taxonomic sample provides no further insight into the evolution of enolase in these protists.

**Gene Duplications for the Origin of Chloroplast-Cytosol Enolase Isoenzymes**

As discussed above, available molecular and biochemical data suggest that the putative transit peptide-
bearing enolases from *Arabidopsis* and *Chlamydomonas* likely represent plastid isoforms of the enzyme. The position of these isoforms in figure 4 suggests that the genes for the chloroplast-specific isoforms were not acquired by the nucleus from the cyanobacterial ancestor of plastids, but, rather, arose through gene duplication involving the cytosolic isoform and acquisition of a transit peptide, as is the case for many compartment-specific enzymes that were common to hosts and symbionts during the origin of organelles (Martin and Schnarrenberger 1997). This duplication appears to have taken place relatively early in plant evolution, probably prior to the origin of the insertion at positions 161–165 that is specific to the plant cytosolic isoforms (under the reasonable premise that the *Plasmodium* sequence is misplaced in the phylogenetic analyses).

### Enolase from *Euglena* Is an Exception

Many genes studied to date underscore a common ancestry of trypanosomes and euglenids. These include nuclear-encoded ribosomal RNA (Sogin 1994), tubulin (Levasseur, Meng, and Bouck 1994), Hsp60 (Yasuhiro and Simpson 1997), and the ER-specific protein calreticulin (Navazio et al. 1998), in addition to mitochondrial-encoded cytochrome oxidase subunit I (Tessier et al. 1997; Yasuhiro and Simpson 1997). Enolase is thus an exception in that it clearly does not reflect the common ancestry of the kinetoplastid (*Trypanosoma*) and euglenid lineages (fig. 4).

In contrast to trypanosomes, euglenids possess plastids. These are surrounded by three membranes, suggesting that *Euglena* may have acquired its plastids by secondary symbiosis through engulfment of a eukaryotic green alga (Gibbs 1978), a view that is supported by chloroplast genome phylogeny (Martin et al. 1998). As in the case of higher plants and *Chlamydomonas*, the cytosolic and plastid-targeting peptide-bearing forms of *Euglena* enolase also must have arisen through gene duplication. Yet, the source of the eukaryotic nuclear gene which gave rise to that duplication is different from that of all other eukaryotic enolase genes. Enolase genes in *Euglena* are much more similar to the homolog from the spirochaete *Treponema* than they are to enolase from any other source currently sampled. This is a mildly surprising, but by no means unprecedented, finding. The nuclear gene for cytosolic GAPDH in *Euglena* also branches quite robustly with its homolog from the *Treponema* genome (Figge et al. 1999). Further affinities between spirochaete and euglenozoan (kinetoplastid and euglenid) glycolytic enzymes are found for PFK, for which the *Trypanosoma* homolog branches with *Borrelia* and, curiously, as with enolase, *Entamoeba* homologs (Michels et al. 1997; Mertens et al. 1998).

At face value, the surprising affinity of nuclear-encoded *Euglena* enolases to the *Treponema* homolog could reflect an outright lateral (i.e., nonendosymbiotic) gene transfer from spirochaetes to the *Euglena* lineage. If so, then future sampling of eukaryotic groups should reveal that the acquisition is specific to the euglenid lineage. If, on the other hand, other, more distantly related, eukaryotic lineages are also found to possess enolase genes that branch with the *Treponema* homolog, then the acquisition could be traced more deeply into eukaryotic phylogeny, in which case differential loss could be invoked to explain the distribution of enolase genes among eukaryotes. In analogy, a similarly “odd” aldolase gene was recently found in *Giardia* that is closely related to α-proteobacterial homologs, whereas other eukaryotic aldolase genes described to date are not (Henze et al. 1998). As in the case of *Euglena* enolase, it is possible that other eukaryotes will be found to possess the same “odd” FBA gene as *Giardia* does.

Thus, further sampling of protist lineages is needed before a strong case for or against horizontal enolase gene transfer from prokaryotes to *Euglena* in a context that did not involve the origin of organelles can be argued. Several previous claims of outright prokaryote-eukaryote horizontal gene transfer (i.e., not involving the origins of chloroplasts or mitochondria) argued on the basis of small species samples have turned out, with time, to be just as easily or more easily explained by gene transfer in the context of the origins of organelles (Martin and Schnarrenberger 1997; Nowitzki et al. 1998), warranting caution when only few protist sequences are available for comparison. The finding that prokaryotes tend to exchange genes via horizontal transfer (Lawrence and Ochman 1998; Vorholt et al. 1999) lends an additional degree of complexity to the problem of trying to identify the source of eukaryotic nuclear genes (Brown and Doolittle 1997; Martin and Schnarrenberger 1997; Doolittle 1999). The non-rRNA-like phylogeny of some eubacterial enolases in figure 4—for example, a lack of common branching for the two *Chlorobium* enolases, the sequences from two spirochaetes, or the four sequences from low-GC Gram positives—would not exclude the possibility that some transfer of enolase genes among eubacteria has occurred.

### Eukaryotic Enolase Is an Exception Among Glycolytic Enzymes

Previous phylogenetic studies of glycolytic/gluconeogenetic enzymes have revealed that most of the glycolytic pathway in the eukaryotic cytosol consists of enzymes that are more similar to eubacterial homologs than they are to archaeabacterial homologs. This is true for GAPDH (Markos, Miretsky, and Müller 1993; Henze et al. 1995; Wiemer et al. 1995), PGK (Brinkmann and Martin 1996; Adje, Opperdoes, and Michels 1998), FBA (Plaumann et al. 1997; Henze et al. 1998), FBP (Martin and Schnarrenberger 1997), TPI (Keeling and Doolittle 1997; Vorholt et al. 1999), GPI (Nowitzki et al. 1998), and GAP (Lawrence and Ochman 1998; Vorholt et al. 1999).

In marked contrast to those enzymes, eukaryotic enolase sequences are not more similar to eubacterial than to archaeabacterial homologs, with the exception of *Euglena*. Rather, they tend to branch near the homologs from *Pyrococcus horihoshii, Pyrococcus furiosus*, and *Methanococcus jannaschii*. However, the position of the branch bearing these three archaeabacterial sequences basal to eukaryotic enolase in figure 4 is not strongly
Enolase is an exception among eukaryotic glycolytic enzymes in that it does not show markedly more similarity to eubacterial homologs than it does to archaeabacterial homologs (except the enolases from *Euglena*). This is in contrast to other enzymes of energy metabolism in eukaryotes, which appear to be acquisitions from eubacteria, probably from the antecedent of mitochondria, even in eukaryotes that now lack mitochondria (Martin and Müller 1998). However, enolase is not the only exception. For example, transketolase in some eukaryotes is very similar to methanogen homologs, whereas most eukaryotic transketolases are more similar to eubacterial homologs (Martin and Schnarrenberger 1997). Furthermore, an important enzyme of the ATP-generating pathway downstream of pyruvate in *Giardia*, acetyl-CoA synthase, is also more similar to archaeabacterial than to eubacterial homologs (Sánchez et al. 1999). Current hypotheses for the origin of eukaryotes postulate that the host that acquired the mitochondrion was a descendant or a member of the archaeabacteria (Doolittle 1998). With only one exception (Martin and Müller 1998), all models for the origin of mitochondria assume that the host was a heterotrophic fermenting organism that obtained its ATP through glycolysis, a view that is very difficult to reconcile with the finding that eukaryotes, including amitochondriate...
forms, generally possess eubacterial glycolytic enzymes. Current data indicate that the origin of mitochondria occurred very early in eukaryotic evolution, if not at its root (Embley and Hirt 1998; Gray, Burger, and Lang 1999), and that the origin of the glycolytic pathway that is the basis of the heterotrophic lifestyle in eukaryotes may trace to the origin of mitochondria (Martin and Müller 1998). In agreement with that view, some protists possess nuclear-encoded glycolytic enzymes that are imported into mitochondria (Liaud et al. 2000). The evolution of the glycolytic pathway entails aspects of gene phylogeny, protein compartmentation, and enzyme structure and function (Fothergill-Gilmore and Michels 1993). Further study of glycolytic enzymes, particularly from protists, should lead to a more complete picture of the complex evolutionary history of this pathway in eukaryotes.

**Supplementary Material**

Enolase sequences reported in this paper have been deposited in GenBank under the accession numbers AF152348 (*T. brucei*), AJ272112 (*Euglena chloroplast*), AJ272111 (*Euglena cytoso*), AJ271719 (spinach cytoso), and AF205070 (*Mastigamoeba*).

**Acknowledgments**

A culture of *M. balamuthi* was kindly provided by Dr. Frederick Schuster (Brooklyn College of the City University of New York, Brooklyn, N.Y.), M.M., P.M., and W.M. thank the National Science Foundation (DEB-9615659), the Interuniversity Poles ofAttraction Programme—Belgium State, Prime Minister’s Office—Federal Office for Scientific, Technical and Cultural Affairs, and the Deutsche Forschungsgemeinschaft, respectively, for financial support. We thank Christiane Lindemann for excellent technical assistance. The contributions of the authors in this paper were as follows: V.H., *T. brucei* cloning and sequencing, writing; H.B., phylogenetic analysis; M.-A.A., *T. brucei* cloning and sequencing; U.N., *Euglena* and spinach cloning and sequencing; J.A.L., *Mastigamoeba* cloning; C.W.S., *Mastigamoeba* sequencing; T.G., *Mastigamoeba* sequence analysis; M.M., P.M., and W.M., interpretation and writing.

**LITERATURE CITED**


Keeling, P. J., and W. F. Doollittle. 1997. Evidence that eu-


3-phosphate dehydrogenases of Trypanosoma brucei and Leishmania mexicana mexicana and expression of the try-


Langer, D., J. Hain, P. Thuriaux, and W. Zillig. 1995. Transcrip-


Martin, W., and C. Schinarrenberger. 1997. The evolution of the Calvin cycle from prokaryotic to eukaryotic chro-


Mertens, E., U. S. Ladrón, J. A. Lee, A. Miretsky, A. Morris, C. Rozario, R. J. Kemp, and M. Müller. 1998. The pyrophosphate-dependent phosphofructokinase of the pro-


Mertens, E., E. van Schaftingen, and M. Muller. 1992. Pyruvate kinase from Trichomonas vaginalis, an allosteric enzyme stimulated by ribose 5-phosphate and glyceral 3-


Michels, P. A. M., M. Marchand, L. Kohl, S. Allert, R. K. Wierenga, and F. R. Oppenrodes. 1991. The cytosolic and glycosomal isoenzymes of glyceraldehyde-3-phosphate dehydrogenase in Trypanosoma brucei have a distinct evo-


Mo, Y., B. G. Harris, and R. W. Gracy. 1973. Triosephos-


Mukai, K., M. Yoshida, H. Toyosaki, Y. Yao, S. Wakabayashi, and H. Matsubara. 1989. An atypical heme-

binding structure of cytochrome c1 of Euglena gracilis mitochon-


Navazio, L., C. Nardi, B. Baldan, P. Dainese, A. C. Fitchett,

E. W. Martin, and P. Mariani. 1998. Functional conserva-


solc glucose-6-phosphate isomerase: sampling eubacte-

rial gene diversity in eukaryotic chromosomes through sym-


Oppenrodes, F. R., and P. Borst. 1977. Localization of nine glycolytic enzymes in a microbody-like organelle in Try-


ruvate kinase is present in Giardia intestinalis. Exp. Para-

sitol. 87:153–156.


Read, M., K. E. Hicks, P. F. Sims, and J. E. Hyde. 1994. Molecular characterisation of the enolase gene from the hu-


tory Press, New York.


Geoffrey McFadden, reviewing editor

Accepted February 14, 2000