Compartment-Specific Isoforms of TPI and GAPDH are Imported into Diatom Mitochondria as a Fusion Protein: Evidence in Favor of a Mitochondrial Origin of the Eukaryotic Glycolytic Pathway

Marie-Françoise Liaud, Christiane Lichtlé,† Kirk Apt,‡ William Martin,* and Rüdiger Cerff**

*Institute of Genetics, University of Braunschweig, Braunschweig, Germany; †Laboratoire de Photorégulation et Dynamique des Membranes Végétales, Ecole Normale Supérieure, Paris, France; and ‡Martek Biosciences, Columbia, Maryland

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and triosephosphate isomerase (TPI) are essential to glycolysis, the major route of carbohydrate breakdown in eukaryotes. In animals and other heterotrophic eukaryotes, both enzymes are localized in the cytosol; in photosynthetic eukaryotes, GAPDH and TPI exist as isoenzymes that function in the glycolytic pathway of the cytosol and in the Calvin cycle of chloroplasts. Here, we show that diatoms—photosynthetic protists that acquired their plastids through secondary symbiotic engulfment of a eukaryotic rhodophyte—possess an additional isoenzyme each of both GAPDH and TPI. Surprisingly, these new forms are expressed as an TPI-GAPDH fusion protein which is imported into mitochondria prior to its assembly into a tetrameric bifunctional enzyme complex. Homologs of this translational fusion are shown to be conserved and expressed also in nonphotosynthetic, heterokont-flagellated oomycetes. Phylogenetic analyses show that mitochondrial GAPDH and its N-terminal TPI fusion branch deeply within their respective eukaryotic protein phylogenies, suggesting that diatom mitochondria may have retained an ancestral state of glycolytic compartmentation that existed at the onset of mitochondrial symbiosis. These findings strongly support the view that nuclear genes for enzymes of glycolysis in eukaryotes were acquired from mitochondrial genomes and provide new insights into the evolutionary history (host-symbiont relationships) of diatoms and other heterokont-flagellated protists.

Introduction

Higher plants synthesize carbohydrates from CO_2 in plastids via the Calvin cycle and light-driven electron transport. Reduced carbon compounds are exported to the cytosol, where they are oxidized to pyruvate through the glycolytic pathway. Pyruvate is imported into mitochondria for oxidation back to CO_2 via the Krebs cycle, yielding ATP from oxidative phosphorylation. This pattern of compartmentalized energy metabolism carries the unmistakable imprint of the symbiotic origins of organelles: Photosynthesis in plastids is an inheritance from cyanobacteria, and respiratory carbohydrate breakdown in mitochondria is an inheritance from α-proteobacteria. But organelle genomes encode almost none of the enzymes involved in this biochemistry. Rather, the corresponding genes involved were transferred from organelle genomes to the nucleus during the course of evolution (Andersson et al. 1998; Martin and Herrmann 1998). Whereas photosynthesis and mitochondrial carbohydrate breakdown in eukaryotes are obviously acquisitions through symbiosis, the origin of the glycolytic pathway in the cytosol that biochemically connects them is a much more difficult issue.

Traditionally, the glycolytic pathway in the cytosol has been viewed as the original and ancestral form of eukaryotic energy metabolism: fermentative ATP-synthesis (Margulis 1970; Whatley, John, and Whatley 1979; Blackstone 1995). This view is consistent with the unproven premise that the host cell which acquired mitochondria was a heterotrophic, anaerobic, fermenting eukaryote that arose from within the archaebacteria (Doolittle 1998). In line with that reasoning, contemporary eukaryotes that lack ATP-producing organelles, like the diplomonad *Giardia lamblia,* do indeed obtain their ATP solely through glycolysis in the cytosol (Müller 1998) and have been suspected of being direct descendants of such a hypothetical host (Cavalier-Smith 1987; Sogin et al. 1989). However, recent phylogenetic analyses indicate that amitochondriate protists harbor nuclear genes of mitochondrial descent and thus once possessed mitochondria in their evolutionary history but lost them subsequently (Embley and Hirt 1998; Hashimoto et al. 1998; Müller 1998).

In particular, all glycolytic enzymes of the eukaryotic cytosol analyzed to date, including those of amitochondriate protists, are much more similar to eubacterial homologs than they are to archaebacterial homologs. This is the case for cytosolic glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (Martin et al. 1993; Cerff 1995), for cytosolic triosephosphate isomerase (TPI) (Keeling and Doolittle 1997), and for several other glycolytic enzymes (Brown and Doolittle 1997; Henze et al. 1998; Martin and Herrmann 1998; Siebers, Klenk, and Hensel 1998). Such findings indicate that these eukaryotic nuclear genes for glycolytic enzymes are acquisitions from ancestral mitochondrial genomes, whereby the genes were transferred to the nucleus, but the proteins are not reimported into the organelle that donated the gene.

The observation that eukaryotes possess a cytosolic glycolytic pathway of eubacterial origin is difficult to account for under classical formulations of the endosymbiont hypothesis (Doolittle 1998). It can be readily explained, however, under the premises that the host of mitochondrial symbiosis was not a heterotroph at all, but rather was a strict autotroph that acquired, through se-

Key words: glyceraldehyde 3-phosphate dehydrogenase, triosephosphate isomerase, phosphoglycerate kinase, multienzyme complex, mitochondrial glycolysis, algal photosynthesis, diatoms, oomycetes, cryptomonads, stramenopiles, cell evolution.

Address for correspondence and reprints: Rüdiger Cerff, Institut für Genetik, Technische Universität Braunschweig, Spielmannstrasse 7, D-38106 Braunschweig, Germany. E-mail: r.cerff@tu-bs.de.

Mol. Biol. Evol. 17(2):213–223. 2000

lection, its glycolytic pathway and heterotrophic capacity from the mitochondrial symbiont via endosymbiotic gene transfer (Martin and Müller 1998). If the eukaryotic glycolytic pathway is, indeed, an acquisition from the mitochondrial genome, it follows that some eukaryotes should have retained at least segments of glycolysis in the donor organelle. But neither is any glycolytic enzyme encoded in mitochondrial DNA (Gray, Burger, and Lang 1999), nor has any glycolytic enzyme ever been found to be localized within mitochondria (Fothergill-Gilmore and Michels 1993; Müller 1998; Mackenzie and McIntosh 1999) although some glycolytic enzymes have been found to associate with the outer mitochondrial membrane (Srere 1987), some even with the outer surface of free-living bacteria (Pancholi and Fischetti 1998).

Here, we show that two essential glycolytic enzymes, GAPDH and TPI, are localized as active proteins in the mitochondria of the diatom *Phaeodactylum tricornutum,* a photosynthetic, heterokont-flagellated protist. The enzymes are imported into the organelle as a TPI-GAPDH translational fusion with the help of a single transit peptide. Homologs of this translational fusion are shown to be conserved and expressed also in nonphotosynthetic, heterokont-flagellated, oomycete protists which belong, together with diatoms, to a larger group of heterokont organisms known as stramenopiles (see *Discussion*). Evidence for GAPDH and TPI in mitochondria of heterokonts provides a previously missing metabolic link between nuclear genes for glycolytic enzymes and the organelle from whose genome they were acquired.

Materials and Methods

Nomenclature

The designations for individual GAPDH genes and proteins follow the guidelines of the Plant Gene Nomenclature Commission (mnemonic/numeric system of the Mendel Database). All GAPDH genes and proteins start with the mnemonic term ''gap'' followed by a number (''gap1'', ''gap2,'' etc.) or a letter (''GapC,'' ''GapA,'' etc.) or both (''GapC1,'' ''GapC2,'' etc.). All gene names are written in italics, with the first letter being lowercase (*gap1*) or upper case (*GapC*) for bacterial/organellar and nuclear encoded genes, respectively. All GAPDH proteins, whether encoded by bacterial/organellar or nuclear genes, have the same names as their corresponding genes but are written in roman type and with the first letter uppercase, e.g., proteins/enzymes Gap1, GapC, GapC1, etc. The generic terms GapC-I (plastid C-I), GapC-III (mitochondrial C-III), and GapC-II/C-IV (cytosolic C-II/C-IV) were chosen to designate protein/gene classes comprising homologs from different organisms; e.g., class GapC-III includes GapC3 from diatoms, GapC2 from *Achlya bisexualis,* and GapC from *Phytophtora infestans* (see fig. 5).

Identification and Sequencing of cDNA and Genomic Clones

Screening of libraries, sequencing of positive clones, and RACE analysis were performed as described (Liaud et al. 1997). cDNA libraries were from *P. tricornutum* (Apt, Sukenik, and Grossman 1995), *Odontella sinensis* (Pancic and Strotmann 1993), and *A. bisexualis* (Bhattacharya, Stickel, and Sogin 1991). Clones for *GapC1, TPI-GapC3, PGK1,* and *PGK2* from *Phaeodactylum* were isolated from a genomic library (Bhaya and Grossman 1993). Genomic clones encoding *PGK1* and *PGK2* from *Phaeodactylum* were identified with a probe amplified to span the entire PGK gene from the cyanobacterium *Synechocystis* PCC6803. To localize intron positions in the genomic sequences, partial cDNA sequences of the corresponding transcripts were synthesized by reverse transcription–polymerase chain reaction (RT-PCR) using poly $(A)^+$ from *Phaeodactylum* and primers complementary to the flanking exons.

Overexpression and Purification of Antigens and Preparation of Polyclonal Antibodies

Purified antigens GapC1 and GapC3 from *Phaeodactylum* were obtained from proteins overexpressed in *Escherichia coli* using the QIA*express* system (QIAGEN). PCR fragments spanning mature subunits GapC1 and GapC3 were prepared from the *GapC1* and *TPI-GapC3* genomic clones as templates with *Bam*HI or *Hin*dIII (underlined) clamped primers: 5′-GGATCCATGTCGATGGCTACCGG-3′ and 5′-AAGCTTAGGCCTTGATCTTG-3′ for GapC1, and 5′-GGATCCATGCCGGTCAATATCGGAATC-3′ and 5′-AAGCTTCATTTGGATTTCTCCATG-3′ for GapC3. The PCR products were digested and cloned as N-terminal 6xHis fusions into pQE30. The constructs were expressed in *E. coli* strain M15 at 37°C, induced by the addition of 2 mM IPTG to the bacterial culture. The His-tagged proteins were affinity purified under denaturing conditions as described (QIAGEN manual, protocol 9) and dialyzed against phosphate buffer. Monospecific polyclonal antibodies against purified GapC1 and GapC3 were prepared by BioScience, Göttingen, Germany.

Embedding of Cells and Immunocytochemistry

Phaeodactylum Böhling cells were grown on airlift in F/2 medium at 15°C, with a 12:12 h light-dark cycle and a light intensity of 80 $mE/m^2/s$ and harvested in the midlogarithmic phase by gentle centrifugation. Cells were fixed at 4°C for 60 min in 2% glutaraldehyde in 0.2 M sodium phosphate buffer (pH 7.4) with 0.4 M saccharose and embedded after dehydration in LRWhite medium grade resin (London Resin Company) as described (Lichtlé, McKay, and Gibbs 1992). Anti-GAPDH antisera were used at dilutions of 1/500 (GapC3) and 1/1,000 (GapC1) overnight, and the secondary antibody (goat anti-rabbit 10-nm gold particles, Bio Cell Gold Conjugates) was used at a dilution of 1:30 for 60 min. Membranes were contrasted by overnight exposure to osmium tetroxyde vapors, followed by uranyl acetate staining. Sections were observed with a Jeol CX2 electron microscope at 80 kV.

Native Molecular Mass Determination and Assay of TPI and GAPDH Enzymes

Böhling cells were grown at 25°C on airlift F/2 medium (SIGMA) complemented with a vitamin mix-

ture (SIGMA), with a 14:10 h light-dark cycle and a light intensity of 180 μE/m²/s. Cells were harvested by centrifugation and resuspended in 1 volume of 100 mM MOPS/NaOH (pH 6.9), 200 mM NaCl, 7 mM EDTA, 7 mM DTE, 0.05 mM NAD, and 1/100 volume of a protease inhibitor cocktail (SIGMA). Extracts were centrifuged at 100,000 × *g* for 1 h. Extracts (100 μl sample) were FPLC-chromatographed at 0.2 ml/min on Superdex 200 (separation range 10–600 kDa) equilibrated with 100 mM Tris/Cl (pH 7.8), 200 mM NaCl, 2 mM DTE, 1 mM EDTA, and 0.2 mM NAD. Fractions (0.2 ml) were collected and assayed for TPI and NAD- and NADP-dependent GAPDH activities. Molecular mass standards were ribonuclease A (13.7 kDa), chymotrypsinogen (25 kDa), ovalbumin (43 kDa), transferrin (81 kDa), aldolase (158 kDa), catalase (232 kDa), ferritin (440 kDa), and thyroglobulin (669 kDa).

Enzyme assays were performed at 20°C and monitored spectrophotometrically at 334 nm in a total volume of 1 ml. One unit of activity corresponds to 1 μmol of NAD(P)H oxidized per minute. TPI assays (1 ml) contained 100 μmol triethanolamine (pH 7.6), 1 μmol glyceraldehyde-3-phosphate, 0.15 μmol NADH, and 2 units glycerol-3-phosphate dehydrogenase (Beisenherz 1955). GAPDH assays (1 ml) contained 100 mmol Tris-Cl (pH 7.8), 4.5 μmol 3-phosphoglycerate, 2 μmol ATP, 0.15 μmol NAD(P)H, 8 μmol $MgSO_4$, 2 μmol DTE, 1 μmol EDTA, and 1.8 U of 3-phosphoglycerate kinase (Cerff 1982).

Immunoblotting

Extracts were resolved on a 12% SDS polyacrylamide gel and blotted onto an ECL nitrocellulose membrane with a semidry transfer apparatus (PHASE, Germany). The membrane was blocked for 1 h in TBS (20 mM Tris-HCl [pH 7.6], 137 mM NaCl) containing 2% blocking reagent (Boehringer), rinsed three times for 10 min with TBS, incubated with the anti-GapC3 antiserum (1:1,000 dilution in TBS) for 1 h at room temperature, and rinsed as before. Antigen-antibody complexes were detected using a 1:1,000 dilution (in TBS, 1 h incubation) of a donkey anti-rabbit secondary antibody conjugated with a horseradish peroxidase (Amersham Life Sciences Ltd.) and rinsed as before. Cross-reacting proteins were visualized using an enhanced chemiluminescence kit (Amersham) following the manufacturer's guidelines.

The membrane was submerged in a stripping buffer (100 mM 2-mercaptoethanol, 2% SDS, 62.5 mM Tris-HCl [pH 6.7]) at 60°C for 30 min, rinsed twice for 10 min in TBS, and reprobed with the anti-GapC1 antiserum (1:1,000 dilution in TBS) in essentially the same way as described above.

Phylogenetic Data Analysis

GAPDH and TPI amino acid sequences were retrieved from GenBank, aligned with the new sequences using CLUSTAL W (Thompson, Higgins, and Gibson 1994), and further aligned by hand. Phylogenetic trees were constructed by the neighbor-joining method (Saitou and Nei 1987) using Dayhoff distances with the PROTDIST, NEIGHBOR, and SEQBOOT (100 replicates) tools of PHYLIP, version 3.572c (Felsenstein 1998). Parsimony analyses of the same alignment were performed with PAUP, version 3.1.1 (Swofford 1993), using random-addition heuristic searches on 100 samples. The same data set was also subjected to maximum-likelihood analyses (Adachi and Hasegawa 1996) with the NJ DIST and local rearrangement options of ProtML and by using the JTT-F substitution matrix.

Results

Isolation and Sequence Analysis of cDNAs and Genes Encoding GAPDH, TPI, and PGK Enzymes from Diatoms and Oomycetes

We screened libraries of the diatoms *P. tricornutum* and *O. sinensis* with various GAPDH probes. We identified three distinct GapC transcripts expressed in both organisms, either as full-size cDNAs (*GapC1, GapC2,* and *GapC3* in *Odontella, GapC2* in *Phaeodactylum*) or as intron-containing genomic clones that are expressed as processed transcripts, as shown by RT-PCR (*GapC1* and *GapC3* in *Phaeodactylum*). For comparison with related nonphotosynthetic protists, we also isolated full-size cDNAs corresponding to diatom *GapC2* and *GapC3* from the oomycete *A. bisexualis.*

Surprisingly, the *GapC3* gene was found to be part of a contiguous translational fusion encoding an N-terminal triosephosphate isomerase (TPI) and a C-terminal GAPDH (alias GapC3) polypeptide. The two glycolytic enzymes are linked by a short, flexible, and positively charged hinge region (fig. 1*A*). In both diatoms, this TPI-GapC3 fusion is preceded by a positively charged presequence (fig. 1*B*) resembling mitochondrial targeting signals (Schatz and Dobberstein 1996) as identified by the PSORT computer program (http://psort.nibb.ac.jp:8800/). Two potential processing sites detected by PSORT for each sequence suggest that the putative targeting peptides are either 20/21 or 31/32 residues long for *Phaeodactylum*/*Odontella,* respectively. In the oomycete *A. bisexualis,* we found homologs of the diatom *TPI-GapC3* and *GapC2* genes. The deduced TPI-GAPDH polypeptide from *Achlya* also has a putative mitochondrial targeting signal which, however, is only 13 residues long (fig. 1*B*). The genomic sequences of *GapC1* and *TPI-GapC3* from *Phaeodactylum* each possess a single intron in the region encoding the presequences of *GapC1* and *TPI-GapC3* (fig. 1*B* and *C*), as identified by comparison with the RT-PCR cDNA sequence.

A *TPI-GapC3* homolog (designated *TigA*) was recently reported from the oomycete *Phytophthora infestans* (Unkles et al. 1997). The gene was suggested to encode a cytosolic protein; neither enzyme activity nor localization were studied. A close inspection of the published *Phytophthora TigA* sequence revealed an additional peptide extending 13 codons beyond the presumptive initiator ATG (fig. 1*B*) that might correspond to a mitochondrial transit peptide. All four derived TPI-GapC-III sequences (for generic terms, see *Materials and Methods* and below) are highly conserved with re-

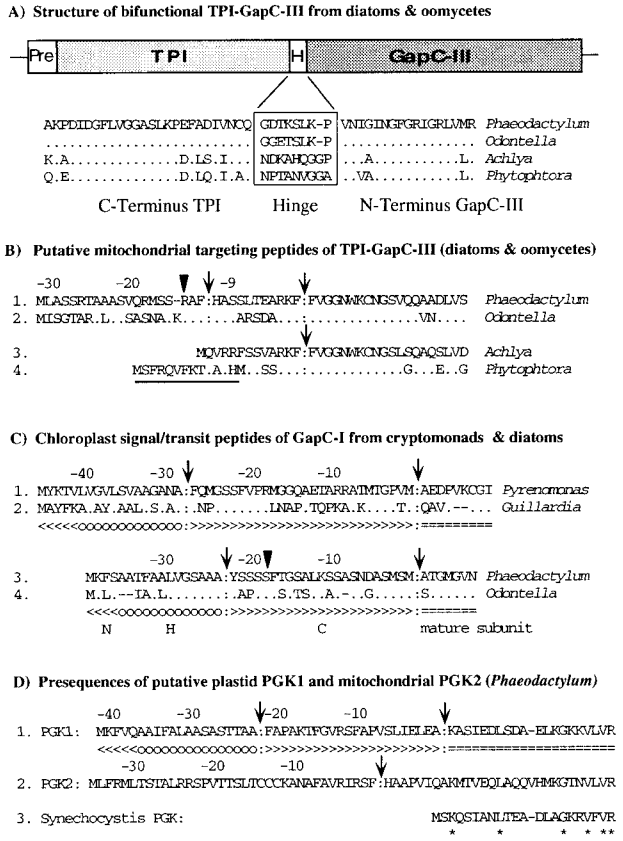

FIG. 1.—General structure of TPI-GapC-III and N-terminal presequences of TPI-GapC-III, GapC-I, and PGK from diatoms and oomycetes. *A,* General structure of the derived TPI-GapC-III precursor polypeptide. ''Pre'' designates N-terminal presequence. TPI and GapC-III are connected by a flexible hinge region (H) of eight or nine residues, as shown below the schematic drawing. Residues identical to the reference sequence are indicated by dots, indels by dashes. *B,* Putative mitochondrial targeting peptides of TPI-GapC-III from diatoms and oomycetes. Arrows indicate processing sites as predicted by the PSORT computer program. The underlined sequence in line 4 indicates an N-terminal sequence not shown in the original publication (Unkles et al. 1997). The position of the phase 2 intron interrupting the presequence of *Phaeodactylum* TPI-GapC-III is indicated by a closed arrowhead (Arg-14/2). *C,* GapC1 signal/transit peptides from diatoms resemble those of cryptomonads. Symbols <<<, ooo, >>>, and === designate the N-terminal, hydrophobic, and C-terminal domains of the presequences and the N-terminus of the mature subunit, respectively. Arrows indicate the putative cleavage sites for signal peptides and transit peptides, respectively (Liaud et al. 1997). The position of the phase 2 intron interrupting the presequence of *Phaeodactylum* GapC-I is indicated by a closed arrowhead (Ser-19/2). *D,* Topological affinities of presequences from *Phaeodactylum* PGKs as predicted by comparison with the presequence of chloroplast GapC-I (PGK1) and by PSORT (PGK2), respectively. Asterisks indicate residues conserved between PGKs of *Phaeodactylum* and *Synechocystis.*

spect to residues implicated in substrate binding and catalysis (Banner et al. 1975; Olsen, Moras, and Rossmann 1975). The GapC-III domain is also strongly conserved with respect to subunit interface residues, indicating that it forms a tetramer, like all other glycolytic GAPDH enzymes known.

The *GapC1* genes isolated from diatoms are orthologs to the *GapC1* gene recently characterized from cryptomonads (Liaud et al. 1997). They contain a bipartite signal/transit peptide necessary for precursor-protein import across the four membranes that surround both diatom and cryptomonad plastids (fig. 1*C*). The deduced GapC1 proteins also contain sequence signatures diagnostic for dual cosubstrate specificity with NADP and NAD. As in cryptomonads, the usually invariant residues Asp32, Gly187, and Pro188, which determine NAD-specificity in most GAPDH enzymes (including GapC2 and GapC3 from diatoms and their homologs in oomycetes), have been substituted with Ala32, Ser187, and Ser188, suggesting that GapC1 can use NADPH in addition to NADH (Clermont et al. 1993; Liaud et al. 1997). The *GapC-II* genes of diatoms and oomycetes encode typical NAD-specific enzymes of cytosolic glycolysis that lack a recognizable presequence.

By using a heterologous probe from the cyanobacterium *Synechocystis* PCC6803, we also isolated genomic clones from *Phaeodactylum* encoding two distinct phosphoglycerate kinases, PGK1 and PGK2, both of which possess N-terminal presequences, suggesting that they might also be imported into plastids and mitochondria, respectively (fig. 1*D*). The two sequences are strongly conserved with respect to functional residues. They are 52% identical and share 59% and 58% of their amino acids, respectively, with *Synechocystis* PGK.

Immunolocalization of GapC3 and GapC1 in Mitochondria and Plastids

In order to determine the intracellular compartmentation of TPI-GapC3 and GapC1, we expressed GapC3 and GapC1 proteins in *E. coli,* raised antisera against the purified proteins, and performed immunogold labeling experiments on sections of *Phaeodactylum* cells (see *Materials and Methods*). GapC3 label is specific and is confined to the mitochondrial matrix space (fig. 2). GapC1 is clearly localized in plastids (fig. 3) and appears to be predominantly associated with thylakoid membranes and the pyrenoid. These findings establish the compartmentation of the products, indicating that the *Phaeodactylum* mitochondrial and plastid transit peptides shown in figure 1 are functional. The anti-GapC1 and anti-GapC3 antisera do not detectably cross-react (fig. 4*B*).

Native Molecular Mass and Western Analysis of TPI-GapC3 and GapC1 from *Phaeodactylum*

Native TPI in eukaryotes is a ~50-kDa dimer, and native GAPDH (both NAD- and NADP-specific forms) is a ~150-kDa tetramer. Since TPI-GapC3 has a deduced monomeric mass of 63 kDa, size exclusion chromatography was performed to investigate the organization of these activities in extracts from *Phaeodactylum* cells harvested 4 h after the onset of light. Figure 4*A* shows elution profiles of TPI, NAD-GAPDH, and NADP-GAPDH activities from Superdex 200. TPI activity elutes as two distinct peaks at ~240 and ~45 kDa. NAD-GAPDH activity is also divided into two peaks, of which one coelutes with the fast TPI peak and the other coelutes with a single peak of NADP-GAPDH activity at ~150 kDa.

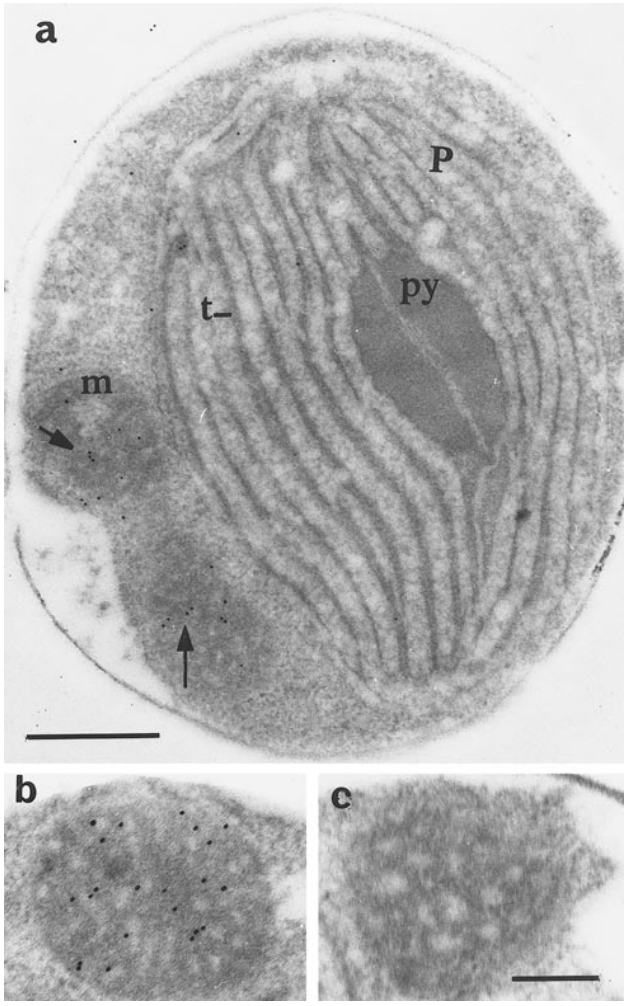

FIG. 2.—Immunogold localization of enzyme GapC3 in *Phaeodactylum* mitochondria. Colloidal gold particles are concentrated (arrows) in the mitochondrial matrix space (m). The chloroplast (P) with pyrenoid (py) and thylakoids (t) and the cytoplasm are not labeled. The transmission electron micrographs show a total cell (*a*) and details of a labeled mitochondrium (*b*). In the control experiment (*c*), the anti-GapC3 antiserum was replaced by a preimmune antiserum. Scale bars in *a* and *c* represent distances of 0.5 and 0.2 μm, respectively.

The ~240-kDa peak of TPI and NAD-GAPDH corresponds to the elution expected for a tetrameric TPI-GapC3 fusion protein. Western blots from denaturing gels of crude extracts (fig. 4*B*) probed with anti-GapC3 reveal a minor, larger band corresponding to the TPI-GapC3 translational fusion (63 kDa) and a major band (sometimes a close doublet) of about 40 kDa, with the molecular mass expected of a discrete GapC3 subunit posttranslationally cleaved from TPI-GapC3 at the hinge region. Probing fraction F1 (fig. 4*A*) with anti-GapC3 reveals a 63- kDa protein corresponding in size to uncleaved TPI-GapC3, whereas fraction F2 again contains a 40-kDa protein with the size expected for a single discrete GapC3 subunit. The 63- and 40-kDa GapC3 bands were detected in extracts ground in liquid nitrogen and immediately transferred to boiling Laemmli buffer, so they are unlikely to be preparation artifacts. Since GapC3 immunogold label was detected only in the mitochondrion (fig. 2), the 40-kDa GapC3 band in F2 (fig. 4*B*) indicates that TPI-GapC3 is subject to partial cleavage inside the mitochondrion.

The NADP-GAPDH peak at ~150 kDa corresponds to a GapC1 tetramer, which should accept both NADP and NAD as cosubstrate (Liaud et al. 1997). It is expected to comigrate with NAD-specific tetramers of cytosolic GapC2 and processed mitochondrial GapC3, respectively. The ~45-kDa TPI peak corresponds to the size expected for the monofunctional, dimeric TPI enzyme that would result from cleavage of TPI-GapC3 or from chloroplast and/or cytosolic TPI isoenzymes that were not cloned from *Phaeodactylum*. Apart from these positive qualitative findings, no conclusions can be drawn from figure 4*A* with respect to relative amounts and specific activities of the different enzyme species tested in *Phaeodactylum* extracts.

Gene Phylogeny of Compartment-Specific GAPDH and TPI Enzymes

To trace the evolutionary history of the mitochondrial TPI-GapC-III protein, we performed neighbor-joining (NJ), maximum-parsimony (MP), and maximum-likelihood (ML) analyses of GAPDH and TPI in the context of their chloroplast and cytosolic counterparts from photosynthetic and nonphotosynthetic eukaryotes and their corresponding homologs in prokaryotes (see *Materials and Methods*). All methods produced similar results for both proteins; only the NJ trees are shown.

The overall topology of the GAPDH tree coincides with previous analyses in four general attributes. (1) The eukaryotic nuclear genes are more similar to eubacterial than to archaebacterial homologs. (2) The NADP-dependent, chloroplast-specific GapA/B subunits of higher plants and green and red algae are most similar to eubacterial Gap2 homologs, which occur only in cyanobacteria. (3) The eukaryotic GapC subunits (usually NAD-specific) share their greatest similarity to proteobacterial Gap1 homologs. (4) Eukaryotic GAPDH gene diversity is a distinct and limited subset of eubacterial *gap* gene diversity, the former being an inheritance of the latter mainly through endosymbiosis and gene transfer to the nucleus (Martin et al. 1993; Figge et al. 1999).

The GAPDH genes and mRNAs that we found in diatoms, oomycetes, and cryptomonads belong to the *GapC* subfamily and fall into four different categories for which we chose the generic designations GapC-I (plastid C-I), GapC-III (mitochondrial C-III), and GapC-II/C-IV (cytosolic C-II/C-IV) to allow intertaxon comparisons. Type GapC-I is shared by diatoms and cryptomonads, both of which are photosynthetic heterokont-flagellated protists (Cavalier-Smith 1993; Corliss 1994). Type GapC-III is shared by diatoms and oomycetes, as is type GapC-II. The related GapC-IV of cryptomonads shares several sequence characteristics with, but does not specifically branch with, GapC-II in these analyses.

The overall topology of the TPI tree (fig. 5*B*) also coincides with previous analyses in that eukaryotic TPI is much more similar to eubacterial TPI genes than it is to archaebacterial homologs (Keeling and Doolittle 1997). Diatom and oomycete TPI, which is N-terminally

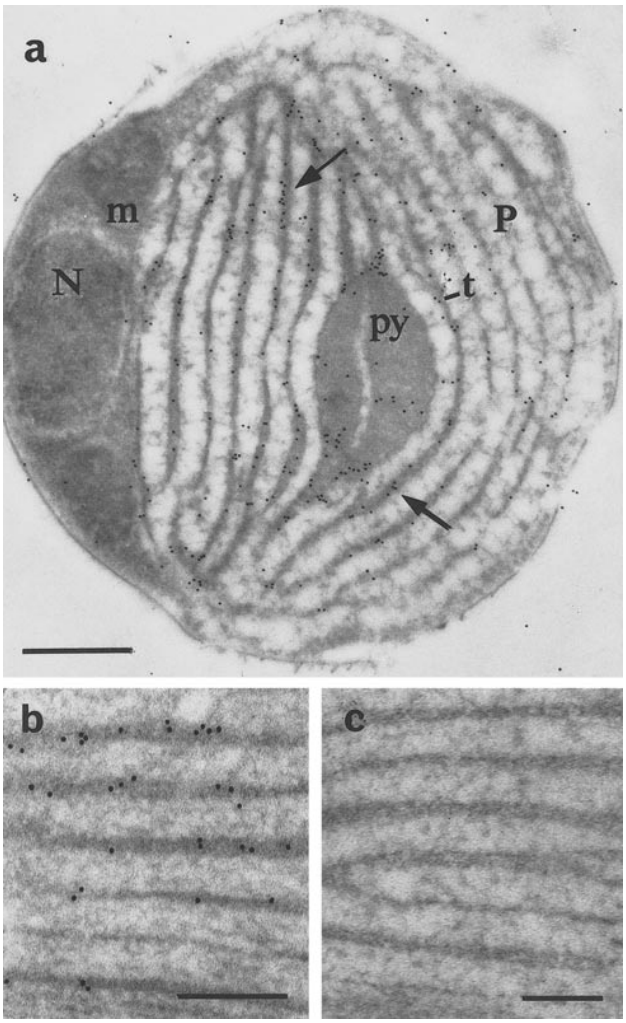

FIG. 3.—Immunogold localization of enzyme GapC1 in *Phaeodactylum* plastids. As shown in micrographs *a* and *b*, gold grains are associated mainly with plastid thylakoids (t) and the pyrenoid (py), while other parts of the cell, including mitochondria (m) and the nucleus (N), are unlabeled. Scale bars represent distances of 0.5 μm in *a* and 0.2 μm in *b* and *c*.

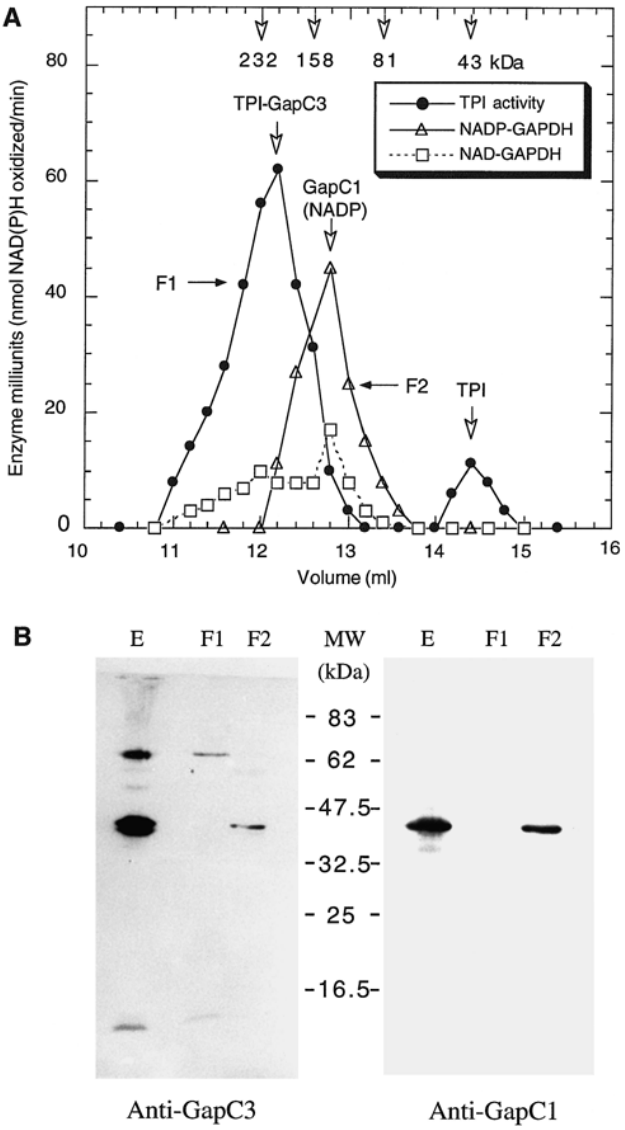

FIG. 4.—Native molecular mass determination and western analysis of bifunctional TPI-GapC3 in extracts from *Phaeodactylum* cells. *A*, Activity profiles of TPI, NAD-GAPDH, and NADP-GAPDH in fractions (0.2 μl) from an extract (100 μl) submitted to FPLC gel filtration on Superdex 200 (separation range 10–600 kDa). Arrows above the profiles indicate the positions and sizes of the following external standards (from left to right): catalase, aldolase, transferrin, and ovalbumin. F1 and F2 designate fractions which were analyzed by immunoblotting. *B*, Immunoblots showing an extract (lane E) and FPLC fractions F1 and F2 probed with anti-GapC3 (left panel) and anti-GapC1 (right panel), respectively. Molecular mass markers are indicated.

fused to GapC-III, also reflects a common ancestry for these heterokont-flagellated protists.

In contrast to a previous study (Keeling and Doolittle 1997), our analyses did not reveal a specific association of eukaryotic TPI with the TPI from the α group of proteobacteria (*Rhizobium etli*; fig. 5*B*). In the case of GAPDH, no Gap1/GapC homolog has yet been found in α-proteobacteria (Figge et al. 1999). Ideally, one would like to see genes argued to be of mitochondrial origin branch specifically with α-proteobacterial homologs. However, proteins as short as TPI (~250 amino acids) and GAPDH (~330 amino acids) are subject to the limitations of phylogenetic information contained within individual proteins (Brown and Doolittle 1997; Keeling and Doolittle 1997; Martin et al. 1998), sampling of eubacterial gene diversity through symbiosis (Martin and Herrmann 1998), and phylogenetic methods themselves (Philippe and Laurent 1998; Figge et al. 1999). Thus, the proteobacterial ancestry of these genes is easier to demonstrate than is their specifically α-proteobacterial ancestry.

Discussion

The diatoms are an extremely diverse group of microalgae that constitute a major component of marine phytoplankton, colonizing the whole ocean down to depths to which photosynthetically available radiation penetrates. Although remarkably successful, diatoms emerged comparatively late in evolution, with the oldest

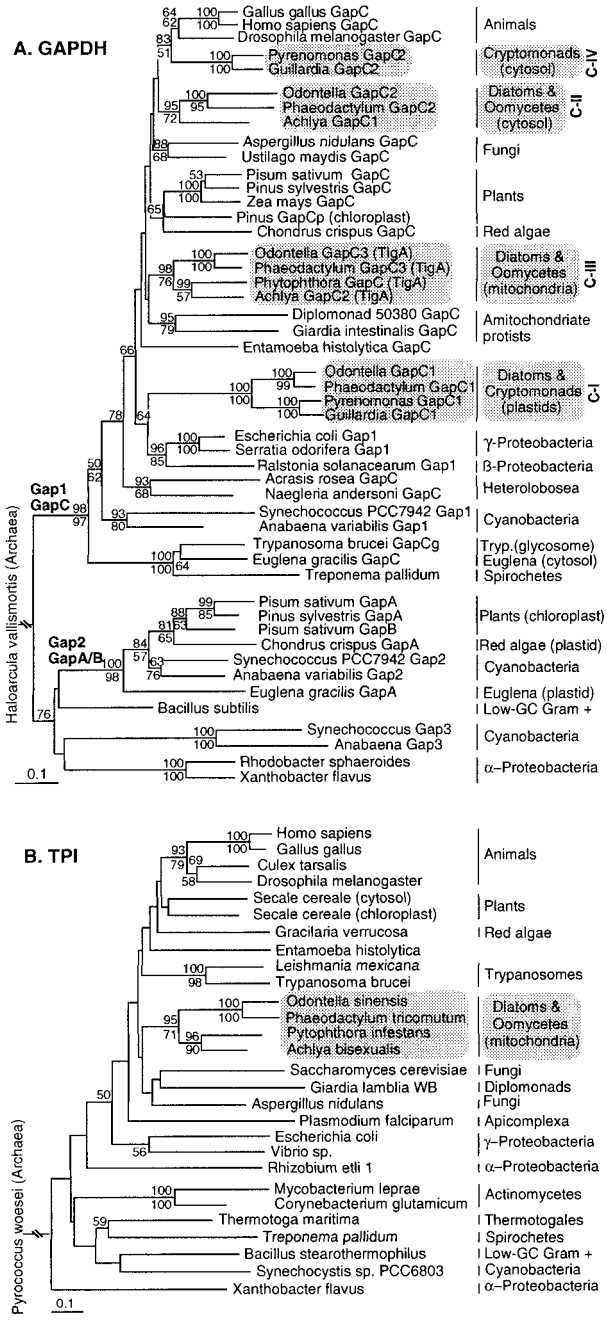

FIG. 5.—Protein phylogeny of compartment-specific GAPDH and TPI enzymes. *A,* Tree of GAPDH protein sequences constructed by the neighbor-joining (NJ) method using Dayhoff distances, rooted to class I GAPDH from the archaebacterium *Haloarcula vallismortis.* Two major branches, corresponding to the subfamilies Gap1-GapC (mainly glycolytic) and Gap2-GapA/B (exclusively photosynthetic) are indicated. The four separate GapC branches leading to the diatom, oomycete, and cryptomonad enzymes are shaded. These are designated here C-I (plastids), C-III (mitochondria), and C-II/C-IV (cytosolic) (see text). Numbers at branches indicate the number of times that a branch was detected out of 100 bootstrap replicates for neighbor joining (upper value) or parsimony (lower value); only values above 50% are indicated. The scale bar indicates 0.1 amino acid substitutions per site. The positions of the GAPDH types C-I–C-IV were similar in the ProtML tree except that *Entamoeba* and fungi (*Aspergillus, Ustilago*) branched at the basis of the C-II and C-IV/animal lineages, respectively. *B,* Tree of TPI protein sequences constructed as in *A,* rooted to the enzyme from the archaebacterium *Pyrococcus woesii.* Numbers and designations are as in *A.*

fossil records known being from the lower Cretaceous (about 120 MYA). In rRNA phylogenies, diatoms belong to a heterogeneous group of eukaryotes known as stramenopiles (Patterson 1989; Medlin et al. 1997; Van de Peer and De Wachter 1997), which typically posses two different-sized (heterokont) flagella, one of which is a tinsel flagellum. Some members of this group, such as the parasitic and saprophytic fungal-like oomycetes, are nonphotosynthetic and possess heterokont flagella only in motile stages (Cavalier-Smith 1993; Corliss 1994). In photosynthetic stramenopiles (heterokont algae) such as diatoms, brown algae, and chrysophytes, the plastids are surrounded by two additional membranes known as the periplastid endoplasmic reticulum (ER). The resulting compartment between the periplastid ER and the plastid envelope is termed the periplastid space. In chlorarachniophytes and cryptomonads, which contain periplastid ERs (Corliss 1994) but do not belong to the stramenopiles in rRNA trees (Cavalier-Smith 1993; Van de Peer and De Wachter 1997), the periplastid space harbors remnants of an engulfed cell's nucleus, the nucleomorph (Gilson, Maier, and McFadden 1997), indicating that these algae originated via eukaryote-eukaryote endosymbioses between a phagotrophic host cell and an unicellular microalga (Gibbs 1981; Sitte 1993; McFadden and Gilson 1995; Douglas 1998). Although diatoms, like most photosynthetic heterokont-flagellated protists, have not preserved a nucleomorph, molecular data indicate that their four membrane-bounded plastids descend from a eukaryotic red alga (Martin et al. 1998; Leitsch, Kowallik, and Douglas 1999). To see if traces of this complex symbiotic history might be preserved in the compartmentation of their energy metabolism, we investigated the compartment-specific isoforms of GAPDH and TPI in diatoms.

Diatom Mitochondria Harbor a Bifunctional TPI-GAPDH Complex and Probably Also a Separate Monofunctional PGK Isoform

Diatom mitochondria are unique to date in that they possess isoforms of the glycolytic enzymes TPI and GAPDH, which enter the organelle in the form of a TPI-GapC3 fusion protein. Although not yet confirmed by cell fractionation data, this is clearly indicated by our immunogold labeling experiments (fig. 2) and by the presence of positively charged presequences resembling mitochondrial targeting signals (fig. 1). Genes for TPI-GapC3 homologs (TPI-GapC-III) are also present in the nonphotosynthetic oomycetes *Achlya* and *Phytophthora* (fig. 1*A* and *B*; Unkles et al. 1997). The presence of TPI-GapC-III in mitochondria has so far been shown only for the diatom *Phaeodactylum* (fig. 2), but it is likely to be the case also for the corresponding polypeptides of oomycetes, since their N-terminal presequences, although shorter than those from diatoms, display a fairly clear-cut topogenic preference for this organelle.

The major part of TPI activity and at least half of the NAD-GAPDH activity in *Phaeodactylum* extracts is associated with a high molecular mass fraction of >200 kDa (fig. 4), roughly corresponding to the size expected for a tetramer of TPI-GapC3 fusion polypeptides. Al-

though the three-dimensional structure of the active fusion has yet to be clarified, the well-known structures of GAPDH and TPI provide clues. All active GAPDH enzymes studied to date are tetramers, while TPI, with only one exception among the archaebacteria (Kohlhoff, Dahm, and Hensel 1996), is active as a dimer. Furthermore, the N- and C-termini of monofunctional TPI and GAPDH are solvent exposed and accessible to chain extensions (Banner et al. 1975; Skarzynski and Wonacott 1988). Thus, native TPI-GapC3 could consist of a classical GAPDH core tetramer bearing two diametrically opposed TPI dimers.

Associations between soluble glycolytic enzymes have often been reported (for review, see Srere 1987), but only one other translational fusion is known. In the thermophilic eubacterium *Thermotoga maritima,* a PGK-TPI fusion is encoded by the *gap-pgk-tpi* operon; the active enzyme is composed of four identical subunits, each consisting of covalently linked PGK-TPI (Schurig et al. 1995; Beaucamp, Schurig, and Jaenicke 1997). In contrast to diatom TPI-GapC3, the two reading frames of *Thermotoga*'s bifunctional enzyme are not separated by a spacer region; rather, they overlap in the PGK stop codon and require a programmed translational frameshift (−1) for the synthesis of the PGK-TPI fusion (Schurig et al. 1995), the constituents of which fold independently (Beaucamp, Schurig, and Jaenicke 1997).

The two distinct PGK genes we found in a *Phaeodactylum* gene library encode enzymes with characteristic N-terminal extensions (fig. 1*D*). Assuming that the oxidative decarboxylation of pyruvate occurs in *Phaeodactylum* mitochondria, as in all mitochondria and hydrogenosomes studied to date (Müller 1998), mitochondrial localization of TPI and GAPDH in diatoms would only make sense if the remaining glycolytic sequence to pyruvate (PGK, phosphoglycerate mutase, enolase, and pyruvate kinase) were also targeted to the organelle, in line with the mitochondrial targeting predicted for *Phaeodactylum* PGK2 (fig. 1*D*).

Taken together, these data favor the view that in diatoms the glycolytic flux from triosephosphate to pyruvate passes mainly through mitochondria, while the cytosolic isoenzymes may play different roles, perhaps in gluconeogenesis and the oxidative pentose phosphate pathway. It is also possible that in diatoms photosynthetic triosephosphate flows straight from the chloroplast to mitochondria via specific export/import translocators, thereby providing a direct metabolic link between the Calvin cycle and the citric acid cycle.

Chloroplast GapC-I: Metabolic Constraints for NADP-Dependent Function

Diatoms employ an atypical GAPDH enzyme in their Calvin cycle, GapC-I, which is an early-diverging member of the GapC subfamily (fig. 5*A*) and was first described for the nucleomorph-bearing cryptomonads (Liaud et al. 1997). This is in contrast to all other photosynthetic eukaryotes studied to date, which use the classical marker enzyme of chloroplasts, NADP-dependent GapA/B, a nuclear-encoded chloroplast enzyme of cyanobacterial origin (see branch GapA/B in fig. 5; Figge et al. 1999 and references therein). While all other eukaryotic GapC enzymes are NAD-specific and usually located in the cytosol, GapC-I of diatoms is localized in plastids (fig. 3) and displays NADP-dependent activity (fig. 4*A*) by virtue of three key amino acids diagnostic for dual NADP/NAD cosubstrate specificity (see *Results* and Liaud et al. [1997]).

The evolution of NADP-dependent GapC-I illustrates the biochemical complexity inherent in secondary symbiosis and endosymbiotic gene transfer. Starting simply, cyanobacteria use NADP-dependent GapA (Gap2) in their Calvin cycles, probably because photosynthetic electron transport yields NADPH, not NADH (Koksharova et al. 1998). Engulfment of a cyanobacterium gave rise to the primary plastids found in the red, green, and glaucocystophyte lineages (Douglas 1998), all of which have preserved cyanobacterial photosynthesis and the original GapA enzyme (Cerff 1995; Figge et al. 1999). The *GapA* gene was transferred from the plastid genome to the nucleus during evolution, and the gene product is reimported into the plastid with the help of a transit peptide.

Secondary symbiosis complicates matters. In the case of *Euglena,* the *GapA* gene of cyanobacterial origin was transferred a second time from the nucleus of the chlorophytic symbiont to the nucleus of the engulfing host; the corresponding protein traverses three plastid membranes to return to its original compartment (Henze et al. 1995; Figge et al. 1999). The plastids of diatoms and cryptomonads both descend from rhodophytic endosymbionts (Douglas 1998; Martin et al. 1998). Although rhodophyte plastids import a typical cyanobacterial GapA (Liaud et al. 1994; Ragan and Gutell 1995), the *GapA* gene was not successfully transferred to either diatom or cryptomonad nuclei. Rather, a preexisting GapC protein acquired NADP activity and the targeting signals necessary for import across the four plastid membranes, thereby replacing the function of the rhodophytic symbiont's GapA (fig. 5*A*). The production of NADPH by photosystem I in plastids could have provided a selective pressure sufficient to drive fixation of the mere three substitutions (Ala32, Ser187, and Ser188) determining NADP/NAD bispecificity in GapC-I, and, hence, the functional replacement and loss of GapA in these heterokont-flagellated protists.

Cytosolic GapC-II and mitochondrial GapC-III support a common ancestry of diatoms and oomycetes, consistent with rRNA phylogeny. However, diatoms are also evolutionarily linked to cryptomonads by their common possession of GapC-I (fig. 5*A*), which is surprising since their host lineages are distantly related in rRNA phylogenies (Cavalier-Smith 1993; Van de Peer and De Wachter 1997). This could reflect evidence for common ancestry of these heterokont-flagellated host lineages that is not apparent in rRNA phylogeny.

Diatom Mitochondria: A Missing Metabolic Link of Organelle Evolution?

In diatoms, we observe the classical GAPDH enzyme of the cytosol (GapC-II), an atypical NADP-dependent enzyme in the four membrane-bounded plastids

that was selected from an NAD-specific ancestor (GapC-I), and a novel NAD-specific enzyme (GapC-III) in the mitochondrion that is translated and imported into the organelle as a fusion protein with TPI. All three GAPDH enzymes are specifically related to proteobacterial homologs of glycolytic function, and the same is true for TPI at the N-terminus of GapC-III. In terms of the evolution of compartmentalized energy metabolism in eukaryotes, this is a surprising, but not inexplicable, set of findings.

The common proteobacterial ancestor of mitochondria and hydrogenosomes unquestionably possessed respiratory pyruvate metabolism, and many of the components of the respiratory chain are still encoded in mitochondrial DNA (Gray, Burger, and Lang 1999). As a free-living bacterium, it therefore should have possessed a glycolytic pathway as well (Brown and Doolittle 1997; Doolittle 1998; Martin and Müller 1998). However, no trace of glycolysis has previously been found in either mitochondria or hydrogenosomes (Fothergill-Gilmore and Michels 1993; Müller 1998). Localization of GAPDH and TPI in diatom mitochondria could, in principle, reflect a rare conservation of the ancestral state of glycolytic compartmentation that existed at the onset of mitochondrial symbiosis. In support of this view are the findings (1) that GapC-III and its N-terminal TPI fusion tend to branch deeply within their respective eukaryotic protein phylogenies (fig. 5); (2) that independent data indicate a mitochondrial origin of the nuclear genes for both proteins (Martin et al. 1993; Cerff 1995; Keeling and Doolittle 1997); (3) that the fusion predates the separation of the oomycete and diatom lineages sampled (fig. 1*A*); and (4) that in both lineages, the fusion bears a presequence corresponding to a transit peptide (diatoms) and a putative transit peptide (oomycetes), respectively (fig. 1*B*). The (putative) transit peptides themselves are not conserved, but this is not surprising because such sequences evolve quite rapidly and rarely share detectable sequence similarity (Schatz and Dobberstein 1996).

GAPDH and TPI often occur in the same operon in eubacteria (Henze et al. 1995). Therefore, the TPI-GapC-III fusion could either reflect a direct inheritance of gene order from the proteobacterial ancestor of mitochondria or a fusion specific to lineages of heterokont-flagellated protists. In the first case, TPI-GAPDH clusters or fusions should eventually be found among proteobacteria; in the second case, their occurrence should be restricted to specific eukaryotic lineages. The ability of individual genes to accurately reflect deep phylogenetic relationships among eukaryotes is a debated issue (Embley and Hirt 1998; Philippe and Laurent 1998). TPI-GAPDH and similar gene clusters (Leitsch, Kowallik, and Douglas 1999) can help to untangle the conundrum of protist evolution.

Chloroplasts and mitochondria are integral to energy metabolism in most eukaryotes. During the symbiotic origins of organelles, many hundreds of genes were transferred from organelle genomes to the nucleus. A substantial proportion of the encoded products acquired a transit peptide, permitting import into the organelle from which the gene was donated, whereas others took up residence in the cytosol or were rerouted to other compartments. Although the first α-proteobacterial genome sequenced, that of the highly reduced obligate parasite *Rickettsia prowazekii*, is devoid of glycolytic enzymes (Andersson et al. 1998), current views on the origin of mitochondria posit that the α-proteobacterial ancestor of the organelle possessed a glycolytic pathway (Martin and Müller 1998; Gray, Burger, and Lang 1999). Whether further glycolytic enzymes can be found in diatom mitochondria, whether other groups of eukaryotes possess glycolytic enzymes in their mitochondria, and whether the condition in *Phaeodactylum* might reflect a relic of ancient eukaryotic energy metabolism are important questions that warrant further investigation of compartmentalized energy metabolism in eukaryotes.

Acknowledgments

We are grateful to P. G. Kroth-Pancic (Botanical Institute, Düsseldorf, Germany) and D. Bhattacharya (Max-Planck-Institute, Göttingen, Germany) for providing cDNA libraries of *Odontella sinensis* and *Achlya bisexualis*, respectively. We thank Nicole Frahm and Armaz Aschrafi for their help with the characterization of *GapC1* and *PGK1* from *Phaeodactylum*, respectively, Günter Schwarz (Botanical Institute, Braunschweig) for help with FPLC, and Ulrike Brandt and Stephanie Escher for technical assistance. We also thank Sergey Shestakov (Insitute of Genetics, Moscow) and Reinhard Hehl (Institute of Genetics, Braunschweig) for critical comments on an earlier version of the manuscript. This work was funded by the Deutsche Forschungsgemeinschaft (CE 1/15 and CE 1/24). The nucleotide sequence data reported will appear in the DDBJ/EMBL/GenBank International Nucleotide Sequence Database under the accession numbers AF063106, AF063107 (GapC1 and TPI-GapC3 of *A. bisexualis*), AF063800, AF063801, AF063802 (GapC1, GapC2, and TPI-GapC3 of *O. sinensis*), AF063803, AF063805, AF063804 (GapC1, GapC2, and TPI-GapC3 of *P. tricornutum*), AF108451, and AF108452 (PGK1 and PGK2 of *P. tricornutum*).

LITERATURE CITED

ADACHI, J., and M. HASEGAWA. 1996. MOLPHY. Version 2.3. The Institute of Statistical Mathematics, Tokyo.

ANDERSSON, S. G., A. ZOMORODIPOUR, J. O. ANDERSSON, T. SICHERITZ-PONTEN, U. C. ALSMARK, R. M. PODOWSKI, A. K. NASLUND, A. S. ERIKSSON, H. H. WINKLER, and C. G. KURLAND. 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature **396**:133–140.

APT, K. E., A. SUKENIK, and A. R. GROSSMAN. 1995. The ER chaperone BiP from the diatom *Phaeodactylum* (PGR95-061). Plant Physiol. **109**:339.

BANNER, D. W., A. C. BLOOMER, G. A. PETSKO et al. (12 coauthors). 1975. Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 angstrom resolution using amino acid sequence data. Nature **255**:609–614.

BEAUCAMP, N., H. SCHURIG, and R. JAENICKE. 1997. The PGK-TIM fusion protein from *Thermotoga maritima* and its constituent parts are intrinsically stable and fold independently. Biol. Chem. **378**:679–685.

BEISENHERZ, G. 1955. Triosephosphate isomerase from calf muscle. Methods Enzymol. **1**:387–391.

BHATTACHARYA, D., S. K. STICKEL, and M. L. SOGIN. 1991. Molecular phylogenetic analysis of actin genic regions from *Achlya bisexualis* (Oomycota) and *Costaria costata* (Chromophyta). J. Mol. Evol. **33**:525–536.

BHAYA, D., and A. R. GROSSMAN. 1993. Characterization of gene clusters encoding the fucoxanthin chlorophyll proteins of the diatom *Phaeodactylum tricornutum.* Nucleic Acids Res. **21**:4458–4466.

BLACKSTONE, N. W. 1995. A units-of-evolution perspective on the endosymbiont theory of the origin of the mitochondrion. Evolution **49**:785–796.

BROWN, J. R., and W. F. DOOLITTLE. 1997. Archaea and the prokaryote-to-eukaryote transition. Microbiol. Mol. Biol. Rev. **61**:456–502.

CAVALIER-SMITH, T. 1987. Eukaryotes with no mitochondria. Nature **326**:332–333.

———. 1993. Kingdom protozoa and its 18 phyla. Microbiol. Rev. **57**:953–994.

CERFF, R. 1982. Separation and purification of NAD- and NADP-linked glyceraldehyde-3-phosphate dehydrogenases from higher plants. Pp. 683–694 *in* M. EDELMAN, R. B. HALLICK, and N.-H. CHUA, eds. Methods in chloroplast molecular biology. Elsevier/North Holland, Amsterdam.

———. 1995. The chimeric nature of nuclear genomes and the antiquity of introns as demonstrated by the GAPDH gene system. Pp. 205–227 *in* M. GÕ and P. SCHIMMEL, eds. Tracing biological evolution in protein and gene structures. Elsevier, Amsterdam.

CLERMONT, S., C. CORBIER, Y. MELY, D. GERARD, A. WONACOTT, and G. BRANLANT. 1993. Determinants of coenzyme specificity in glyceraldehyde-3-phosphate dehydrogenase: role of the acidic residue in the fingerprint region of the nucleotide binding fold. Biochemistry **32**:10178–10184.

CORLISS, J. O. 1994. An interim utilitarian (''user-friendly'') hierarchical classification and characterization of the protists. Acta Protozool. **33**:1–51.

DOOLITTLE, W. F. 1998. A paradigm gets shifty. Nature **392**: 15–16.

DOUGLAS, S. E. 1998. Plastid evolution: origins, diversity, trends. Curr. Opin. Genet. Dev. **8**:655–661.

EMBLEY, T. M., and R. P. HIRT. 1998. Early branching eukaryotes? Curr. Opin. Genet. Dev. **8**:624–629.

FELSENSTEIN, J. 1998. PHYLIP (phylogeny inference package). Version 3.572c. Distributed by the author, Department of Genetics, University of Washington, Seattle.

FIGGE, R. M., M. SCHUBERT, H. BRINKMANN, and R. CERFF. 1999. Glyceraldehyde-3-phosphate dehydrogenase gene diversity in eubacteria and eukaryotes: evidence of intra- and inter-kingdom gene transfer. Mol. Biol. Evol. **16**:429–440.

FOTHERGILL-GILMORE, L. A., and P. A. M. MICHELS. 1993. Evolution of glycolysis. Prog. Biophys. Mol. Biol. **59**:105–235.

GIBBS, S. P. 1981. The chloroplasts of some algal groups may have evolved from endosymbiotic eukaryotic algae. Ann. N.Y. Acad. Sci. **361**:193–208.

GILSON, P. R., U. G. MAIER, and G. I. MCFADDEN. 1997. Size isn't everything: lessons in genetic miniaturisation from nucleomorphs. Curr. Opin. Genet. Dev. **7**:800–806.

GRAY, M. W., G. BURGER, and B. F. LANG. 1999. Mitochondrial evolution. Science **283**:1476–1481.

HASHIMOTO, T., L. B. SANCHEZ, T. SHIRAKURA, M. MÜLLER, and M. HASEGAWA. 1998. Secondary absence of mitochondria in *Giardia lamblia* and *Trichomonas vaginalis* revealed by valyl-tRNA synthetase phylogeny. Proc. Natl. Acad. Sci. USA **95**:6860–6865.

HENZE, K., A. BADR, M. WETTERN, R. CERFF, and W. MARTIN. 1995. A nuclear gene of eubacterial origin in *Euglena gracilis* reflects cryptic endosymbioses during protist evolution. Proc. Natl. Acad. Sci. USA **92**:9122–9126.

HENZE, K., H. G. MORRISON, M. L. SOGIN, and M. MÜLLER. 1998. Sequence and phylogenetic position of a class II aldolase gene in the amitochondriate protist, *Giardia lamblia.* Gene **222**:163–168.

KEELING, P. J., and W. F. DOOLITTLE. 1997. Evidence that eukaryotic triosephosphate isomerase is of alpha- proteobacterial origin. Proc. Natl. Acad. Sci. USA **94**:1270–1275.

KOHLHOFF, M., A. DAHM, and R. HENSEL. 1996. Tetrameric triosephosphate isomerase from hyperthermophilic Archaea. FEBS Lett. **383**:245–250.

KOKSHAROVA, O., M. SCHUBERT, S. SHESTAKOV, and R. CERFF. 1998. Genetic and biochemical evidence for distinct key functions of two highly divergent GAPDH genes in catabolic and anabolic carbon flow of the cyanobacterium *Synechocystis* sp. PCC 6803. Plant Mol. Biol. **36**:183–194.

LEITSCH, C. I. W., K. V. KOWALLIK, and S. DOUGLAS. 1999. The atpA gene cluster of *Guillardia theta* (Cryptophyta): a piece in the puzzle of chloroplast genome evolution. J. Phycol. **35**:115–122.

LIAUD, M. F., U. BRANDT, M. SCHERZINGER, and R. CERFF. 1997. Evolutionary origin of cryptomonad microalgae: two novel chloroplast/cytosol-specific GAPDH genes as potential markers of ancestral endosymbiont and host cell components. J. Mol. Evol. **44**:S28–S37.

LIAUD, M. F., C. VALENTIN, W. MARTIN, F. Y. BOUGET, B. KLOAREG, and R. CERFF. 1994. The evolutionary origin of red algae as deduced from the nuclear genes encoding cytosolic and chloroplast glyceraldehyde-3-phosphate dehydrogenases from *Chondrus crispus.* J. Mol. Evol. **38**:319–327.

LICHTLÉ, C., R. M. L. MCKAY, and S. P. GIBBS. 1992. Immunogold localization of photosystem I and photosystem II light-harvesting complexes in cryptomonad thylakoids. Biol. Cell **74**:187–194.

MCFADDEN, G., and P. GILSON. 1995. Something borrowed, something green: lateral transfer of chloroplasts by secondary endosymbiosis. TREE **10**:12–17.

MACKENZIE, S., and L. MCINTOSH. 1999. Higher plant mitochondria. Plant Cell **11**:571–586.

MARGULIS, L. 1970. Origin of eukaryotic cells. Yale University Press, New Haven, Conn.

MARTIN, W., H. BRINKMANN, C. SAVONA, and R. CERFF. 1993. Evidence for a chimeric nature of nuclear genomes: eubacterial origin of eukaryotic glyceraldehyde-3-phosphate dehydrogenase genes. Proc. Natl. Acad. Sci. USA **90**:8692–8696.

MARTIN, W., and R. G. HERRMANN. 1998. Gene transfer from organelles to the nucleus: how much, what happens, and why? Plant Physiol. **118**:9–17.

MARTIN, W., and M. MÜLLER. 1998. The hydrogen hypothesis for the first eukaryote. Nature **392**:37–41.

MARTIN, W., B. STÖBE, V. GOREMYKIN, S. HANSMANN, M. HASEGAWA, and K. V. KOWALLIK. 1998. Gene transfer to the nucleus and the evolution of chloroplasts. Nature **393**: 162–165.

MEDLIN, L. K., W. H. C. F. KOOISTRA, D. POTTER, G. W. SAUNDERS, and R. A. ANDERSON. 1997. Phylogenetic relationships of the 'golden algae' (haptophytes, heterokont chromophytes) and their plastids. Pp. 187–219 *in* D. BHATTACHARYA, ed. Origins of algae and their plastids. Springer Verlag, Vienna, New York.

MÜLLER, M. 1998. Enzymes and compartmentation of core energy metabolism of anaerobic protists—a special case in

eukaryotic evolution? Pp. 109–132 *in* G. H. COOMBS, K. VICKERMANN, M. A. SLEIGH, and A. WARREN, eds. Evolutionary relationships among protozoa. Chapman Hall, London.

OLSEN, K. W., D. MORAS, and M. G. ROSSMANN. 1975. Sequence variability and structure of D-glyceraldehyde-3-phosphate dehydrogenase. J. Biol. Chem. **250**:9313–9321.

PANCHOLI, V., and V. A. FISCHETTI. 1998. Alpha-enolase, a novel strong plasmin(ogen) binding protein on the surface of pathogenic streptococci. J. Biol. Chem. **273**:14503–14515.

PANCIC, P. G., and H. STROTMANN. 1993. Structure of the nuclear encoded gamma subunit of CF0CF1 of the diatom *Odontella sinensis* including its presequence. FEBS Lett. **320**:61–66.

PATTERSON, D. J. 1989. Stramenopiles: chromophytes from a protistan perspective. Pp. 357–379 *in* J. C. GREEN, B. S. C. LEADBEATER, and W. L. DIVER, eds. The chromophyte algae: problems and perspectives. Clarendon, Oxford, England.

PHILIPPE, H., and J. LAURENT. 1998. How good are deep phylogenetic trees? Curr. Opin. Genet. Dev. **8**:616–623.

RAGAN, M. A., and R. R. GUTELL. 1995. Are red algae plants? Bot. J. Linn. Soc. **118**:81–105.

SAITOU, N., and M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

SCHATZ, G., and B. DOBBERSTEIN. 1996. Common principles of protein translocation across membranes. Science **271**: 1519–1526.

SCHURIG, H., N. BEAUCAMP, R. OSTENDORP, R. JAENICKE, E. ADLER, and J. R. KNOWLES. 1995. Phosphoglycerate kinase and triosephosphate isomerase from the hyperthermophilic bacterium *Thermotoga maritima* form a covalent bifunctional enzyme complex. EMBO J. **14**:442–451.

SIEBERS, B., H. P. KLENK, and R. HENSEL. 1998. PPi-dependent phosphofructokinase from *Thermoproteus tenax,* an archaeal descendant of an ancient line in phosphofructokinase evolution. J. Bacteriol. **180**:2137–2143.

SITTE, P. 1993. Symbiogenic evolution of complex cells and complex plastids. Eur. J. Protistol. **29**:131–143.

SKARZYNSKI, T., and A. J. WONACOTT. 1988. Coenzyme-induced conformational changes in glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus.* J. Mol. Biol. **203**:1097–1118.

SOGIN, M. L., J. H. GUNDERSON, H. J. ELWOOD, R. A. ALONSO, and D. A. PEATTIE. 1989. Phylogenetic meaning of the kingdom concept: an unusual rRNA from *Giardia lamblia.* Science **243**:75–77.

SRERE, P. A. 1987. Complexes of sequential metabolic enzymes. Annu. Rev. Biochem. **56**:89–124.

SWOFFORD, D. L. 1993. PAUP: phylogenetic analysis using parsimony. Version 3.1.1. Laboratory of Molecular Systematics, Smithsonian Institution, Washington, D.C.

THOMPSON, J. D., D. G. HIGGINS, and T. J. GIBSON. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**:4673–4680.

UNKLES, S. E., J. M. LOGSDON JR., K. ROBISON, J. R. KINGHORN, and J. M. DUNCAN. 1997. The tigA gene is a transcriptional fusion of glycolytic genes encoding triose-phosphate isomerase and glyceraldehyde-3-phosphate dehydrogenase in oomycota. J. Bacteriol. **179**:6816–6823.

VAN DE PEER, Y., and R. DE WACHTER. 1997. Evolutionary relationships among the eukaryotic crown taxa taking into account site-to-site rate variation in 18S rRNA. J. Mol. Evol. **45**:619–630.

WHATLEY, J. M., P. JOHN, and F. R. WHATLEY. 1979. From extracellular to intracellular: the establishment of mitochondria and chloroplasts. Proc. R. Soc. Lond. B Biol. Sci. **204**: 165–187.

GEOFFREY MCFADDEN, reviewing editor

Accepted October 11, 1999