J. Eukaryot. Microbiol., 56(6), 2009 pp. 568–576

DOI: 10.1111/j.1550-7408.2009.00437.x

Transketolase from *Cyanophora paradoxa*: In Vitro Import into Cyanelles and Pea Chloroplasts and a Complex History of a Gene Often, But Not Always, Transferred in the Context of Secondary Endosymbiosis

YAN MA,[a,1] JOHANNES JAKOWITSCH,[a,2] OLIVER DEUSCH,[b] KATRIN HENZE,[b] WILLIAM MARTIN[b] and WOLFGANG LÖFFELHARDT[a]

[a]*Max F. Perutz Laboratories, Department of Biochemistry and Cell Biology, University of Vienna, Dr. Bohrgasse 9, 1030 Vienna, Austria and* [b]*Institut für Botanik III, Heinrich-Heine Universität Düsseldorf, 40225 Düsseldorf, Germany*

ABSTRACT. The glaucocystophyte *Cyanophora paradoxa* is an obligatorily photoautotrophic biflagellated protist containing cyanelles, peculiar plastids surrounded by a peptidoglycan layer between their inner and outer envelope membranes. Although the 136-kb cyanelle genome surpasses higher plant chloroplast genomes in coding capacity by about 50 protein genes, these primitive plastids still have to import >2,000 polypeptides across their unique organelle wall. One such protein is transketolase, an essential enzyme of the Calvin cycle. We report the sequence of the pre-transketolase cDNA from *C. paradoxa* and in vitro import experiments of precursor polypeptides into cyanelles and into pea chloroplasts. The transit sequence clearly indicates the localization of the gene product to cyanelles and is more similar to the transit sequences of the plant homologues than to transit sequences of other cyanelle precursor polypeptides with the exception of a cyanelle consensus sequence at the N-terminus. The mature sequence reveals conservation of the thiamine pyrophosphate binding site. A neighbor-net planar graph suggests that *Cyanophora*, higher plants, and the photosynthetic protist *Euglena gracilis* acquired their nuclear-encoded transketolase genes via endosymbiotic gene transfer from the cyanobacterial ancestor of plastids; in the case of *Euglena* probably entailing two transfers, once from the plastid in the green algal lineage and once again in the secondary endosymbiosis underlying the origin of *Euglena's* plastids. By contrast, transketolase genes in some eukaryotes with secondary plastids of red algal origin, such as *Thalassiosira pseudonana*, have retained the pre-existing transketolase gene germane to their secondary host.

Key Words. Endosymbiotic gene transfer, neighbor-net, plastids, protein import.

THE cyanelles of *Cyanophora paradoxa* are surrounded by a peptidoglycan wall (Pfanzagl et al. 1996). This feature is found among eukaryotes only in glaucocystophyte algae and is one of the strongest hints indicating an endosymbiotic origin of plastids from a cyanobacterial invader. If all plastids descend from a single primary endosymbiotic event, as is currently believed (Gould, Waller, and McFadden 2008; Martin et al. 1998; Moreira, Le Guyader, and Philippe 2000; Rodríguez-Ezpeleta et al. 2005), then all extant plastids, including cyanelles, regardless of their morphology and pigmentation are derived from one and the same ancestral organelle. This ''protoplastid'' had retained the prokaryotic wall and had developed a specific protein import apparatus (Kalanon and McFadden 2008; Steiner and Löffelhardt 2002) for re-importing the products of the genes that had been transferred to the host cell nucleus during the establishment of a stable hereditary endosymbiosis. Thus, the import mechanism at the plastid envelope constitutes an eukaryotic achievement, which does not exclude the recruitment of some suitable cyanobacterial envelope proteins, and should be basically similar for all primary plastids (i.e. chloroplasts, rhodoplasts, and cyanelles). Efficient heterologous import of *Cyanophora* and red algal precursors into isolated pea choroplasts supported this view (Apt, Hoffman, and Grossman 1993; Jakowitsch et al. 1996). However, inverse heterologous import (i.e. import of precursors from higher plants into isolated cyanelles) did not work (Steiner and Löffelhardt 2002).

Previous work on nucleus-encoded precursors to cyanelle proteins has focussed on explaining this discrepancy. The first *Cyanophora* precursors identified showed a rather low positive net charge of the transit peptide (Jakowitsch et al. 1996; Steiner and Löffelhardt 2002) leading to the hypothesis that the negatively charged peptidoglycan layer would be incompatible with highly positively charged transit sequences like those from higher plants (Löffelhardt et al. 1998).

To test this idea further, additional chloroplast proteins are needed as substrate for import into the organelle. One such well-characterized chloroplast protein is transketolase (TKL; EC 2.2.1.1), an essential thiamine pyrophosphate-dependent enzyme that catalyzes two reactions of the Calvin cycle (Flechner et al. 1996): (i) the transfer of the C_2 ketol moiety from D-fructose-6-phosphate to D-glyceraldehyde-3-phosphate yielding D-erythrose-4-phosphate and D-xylulose-5-phosphate; and (ii) the transfer of the C_2 ketol moiety from D-sedoheptulose-7-phosphate to D-glyceraldehyde-3-phosphate yielding D-ribose-5-phosphate and D-xylulose-5-phosphate. The Calvin cycle enzyme from plastids was first characterized at the molecular level in spinach (Flechner et al. 1996). In spinach, TKL also catalyzes the same reactions in the reverse direction in the oxidative pentose phosphate pathway (Teige, Melzer, and Süss 1998), the complete (i.e. cyclic) activity of which is localized exclusively in the plastid (Schnarrenberger, Flechner, and Martin 1995). The crystal structure of the active higher plant homodimer in complex with TPP has been determined and shows similar subunit interactions as the enzyme from yeast (Gerhardt et al. 2003). Like most enzymes of the Calvin cycle in eukaryotes, the gene for TKL is nuclear encoded but was acquired from the cyanobacterial antecedant of plastids via endosymbiotic gene transfer (reviewed by Martin and Schnarrenberger 1997).

Among photosynthetic protists, localization of TKL has not been investigated experimentally. Here we report a cDNA for TKL from *C. paradoxa* with a conspicuous N-terminal extension that differs from the stroma-targeting peptides (STPs) of higher plant chloroplasts while showing a high positive net charge. This N-terminal extension efficiently directs homologous and heterologous in vitro import of the ~74-kDa subunit into cyanelles and pea chloroplasts, respectively. Sequence comparisons of TKLs underscore the role of gene transfer from organelles to the nucleus in the context of endosymbiosis.

[1]Present address: Department of Pathophysiology, Center of Physiology, Pathophysiology and Immunology, Medical University of Vienna, 1090 Vienna, Austria.

[2]Present address: Department of Internal Medicine II, Division of Cardiology, Medical University of Vienna, 1090 Vienna, Austria.

Corresponding Author: W. Löffelhardt; Max F. Perutz Laboratories, Department of Biochemistry and Cell Biology, University of Vienna, Dr. Bohrgasse 9, 1030 Vienna, Austria—Telephone number: +43 1 4277 74802; FAX number: +43 1 4277 9528; e-mail: wolfgang.loeffelhardt@univie.ac.at

MATERIALS AND METHODS

Cloning and sequence analysis. A DNA fragment of 800 bp in length encoding a partial cDNA for *C. paradoxa* TKL, was obtained by PCR. A λ ZAP XR (Stratagene, La Jolla, CA) cDNA library was screened with a digoxigenin (DIG; Boehringer, Mannheim, Germany) -labeled TKL fragment. The positive cDNA clones were isolated. The complete sequences coding for *C. paradoxa* TKL were obtained, analyzed, and compared with known TKL sequences from the EMBL/GenBank and Swiss Prot Databases using Fasta and PileUp programs from the GCG sequence analysis package (Genetics Computer Groups, University of Wisconsin, Madison, WI).

Southern and Northern analysis. Southern and Northern hybridizations were performed as described by Ma et al. (2001). A vacuum blotting apparatus was used for the Southern blotting via downward transfer of DNA from an agarose gel onto an uncharged nylon membrane (Hybond-H; Amersham, Braunschweig, Germany). After treating the gel by three steps with depurination solution (0.25 N HCl), denaturation solution (1.5 M NaCl, 0.5 M NaOH) and neutralization solution (1 M Tris, 2 M NaCl, pH 5.0) sequentially, the transfer was performed with 20 × standard saline citrate (SSC) solution. The membrane was then hybridized with DIG-labeled DNA probe in fresh hybridization buffer (5 × SSC, 50% [v/v] deionized formamide, 2% [w/v] blocking reagent, 0.1% [w/v] *N*-lauroylsarcosine, and 0.02% [w/v] SDS) at 42 °C overnight.

A denaturing agarose–formaldehyde gel electrophoresis of mRNA and total RNA was performed for Northern analysis. The RNA was transferred onto the nylon membrane and hybridized with DIG-labeled probe using the ''high SDS'' hybridization buffer (5 × SSC, 50% [v/v] deionized formamide, 50 mM sodium phosphate, pH 7.0, 2% [w/v] blocking reagent, 0.1% [w/v] *N*-lauroylsarcosine, and 7% [w/v] SDS). The DIG immunological detection system was used for both analyses as recommended by the supplier (Boehringer).

Isolation of pea chloroplasts. Before isolating the intact chloroplasts from pea leaves, 10- to 14-day-old pea plants were placed in dark for 16 h to degrade stored starch, then replaced in light again for 1 h. The fresh tender leaves (~ 40 g) were homogenized in 150 ml ice-cold 1 × GR buffer (0.33 M sorbitol, 2 mM EDTA, 1 mM $MgCl_2$, 1 mM $MnCl_2$, 50 mM HEPES-KOH, pH 7.5 and, before use, adding 300 mg/L BSA, 1 g/L ascorbic acid) by Ultra-Turrax (IKA, Staufen, Germany) for 1 min on ice. The homogenization was repeated twice. The homogenate was centrifuged at 1,500 *g* for 5 min. The chloroplast pellet was resuspended in 1 × GR buffer; each suspension was, respectively, transferred on top of a 50-ml tube containing 25 ml of a Percoll gradient (50% [v/v] Percoll and 50% [v/v] 2 × GR buffer), which was prepared by centrifugation at 30,000 *g* for 15 min. After centrifuging at 15,000 *g* for 5 min, the broken chloroplasts (upper band) were separated from intact chloroplasts (lower band). The upper part was removed and the dark green chloroplast band was transferred to a new 50-ml tube, washed twice with 50 ml ice-cold 1 × GR buffer, and centrifuged at 1,500 *g* for 3 min. The chloroplast pellet was resuspended in 300 µl HS buffer (0.33 mM sorbitol, 50 mM HEPES-KOH, pH 8.0). The quantity of chlorophyll (*a*+*b*) was measured photometrically: 10 µl of the chloroplast sample was mixed with 1 ml of 80% (v/v) acetone and centrifuged at 10,000 *g* for 2 min, the supernatant was transferred to a 1-ml micro-cuvette (1-cm path length) to measure the OD at 663 and 645 nm. The concentration of chlorophyll (*a*+*b*) was calculated according to the formula of $OD_{663} \times 8.02 + OD_{645} \times 20.2 = \mu g$ chlorophyll (*a*+*b*)/ml.

Isolation of cyanelles from *C. paradoxa*. The algal culture of *C. paradoxa* was harvested as described in Jakowitsch et al. (1996). The algal pellet was suspended in 200 ml of 1 × GR buffer, homogenized by Ultra-Turrax three times for 1 min, and, in between, kept on ice for 1 min. The homogenization was repeated twice. The homogenate was filtered through Miracloth (Merck, Nottingham, UK) and centrifuged for 2 min at 2,000 *g*. The pellet was resuspended in 10 ml HS buffer by pipeting slowly, and 4 ml of 40% (v/v) Percoll pad were added to the suspension as underlayer and separated by centrifugation for 5 min at 2,000 *g*. The pellet was gently resuspended in HS buffer and centrifuged for 2 min at 1,500 *g*. The pellets were washed with 40 ml of HS buffer. After centrifuging the suspension, a cyanelle pellet with a transparent supernatant was observed. The pellet was resuspended in 1 ml HS buffer including 1 mM ε-aminocaproic acid and 0.1 mM phenylmethylsulfonylfluoride (or 0.2 mM Pefabloc, Boehringer) as protease inhibitor. The concentration of chlorophyll *a* was measured photometrically.

In vitro protein import. Pre-TKL protein was generated by in vitro transcription/translation from the *C. paradoxa* cDNA cloned into the vector pBAT (Anweiler, Hipskind, and Wirth 1991). A special import buffer was used, containing the main sugar metabolites in *C. paradoxa* and osmotica protective for cyanelle integrity (Jakowitsch et al. 1996). The import reaction mixture (50 µl of a cyanelle suspension containing ≥ 250 µg chlorophyll *a* in import buffer, 10 µl of 100 mM ATP, and 40 µl of ^{35}S-labeled in vitro translation product) was incubated in a 10-ml transparent tube for 30 min at 25 °C in light (30 µmol/photons/m^2/s) with slow shaking to just avoid sedimentation of cyanelles. At the same time, a negative control was set up by incubation on ice in dark. The cyanelles were then washed twice with 1 ml wash buffer (import buffer, 50 mM L-methionine) and centrifuged for 2 min at 2,500 *g*. The cyanelle pellet was resuspended in 400 µl thermolysin buffer (10 mM $CaCl_2$ in wash buffer), and divided into two parts of 200 µl. One part was incubated for 30 min on ice with 10 µl thermolysin solution (2 mg/ml), and 10 µl of 0.2 M EDTA were added to stop the protease reaction. The cyanelles were spun down for 3 min at 2,500 *g* and the pellet was resuspended in 60 µl of 1 × SDS sample buffer and analyzed by a 10% discontinuous SDS-PAGE (Laemmli 1970).

***Euglena* TKL.** A cDNA encoding TKL from *Euglena gracilis* was identified from an EST project (Ahmadinejad, Dagan, and Martin 2007), and a corresponding full-size cDNA (accession number AY738740) was isolated and characterized as previously described (Hoffmeister et al. 2004).

Neighbor-net analysis. TKL homologues were identified by BLAST search of the non-redundant database at GenBank (http://www.ncbi.nlm.nih.gov/), retrieved, and aligned using CLUSTALW (Thompson, Higgins, and Gibson 1994). Gapped positions in the alignment were removed. Protein LogDet distances (LDDist) (Lockhart et al. 1994) between protein coding sequenes were determined with LDDist (Thollesson 2004). NNet planar graphs of splits among protein LDDist were constructed with NNet (Bryant and Moulton 2004) and visualized with Splitstree (Huson and Bryant 2006).

RESULTS

Southern and Northern hybridization of the labeled PCR product of TKL to genomic DNA and polyA$^+$-RNA from *C. paradoxa*, respectively, pointed toward a single copy gene (Fig. 1A) giving rise to a relatively abundant 3.1-kb transcript (Fig. 1B). This indicated a larger 5′-untranslated region and polyA-tail than those present on the 2.6 kb cDNA clone. There, the leader comprised 24 bp only and the polyA-tail (27 bp) appeared to be truncated, too.

The TKL transit sequence of 90 amino acids (aa) was the longest hitherto found for a cyanelle precursor polypeptide (Table 1).

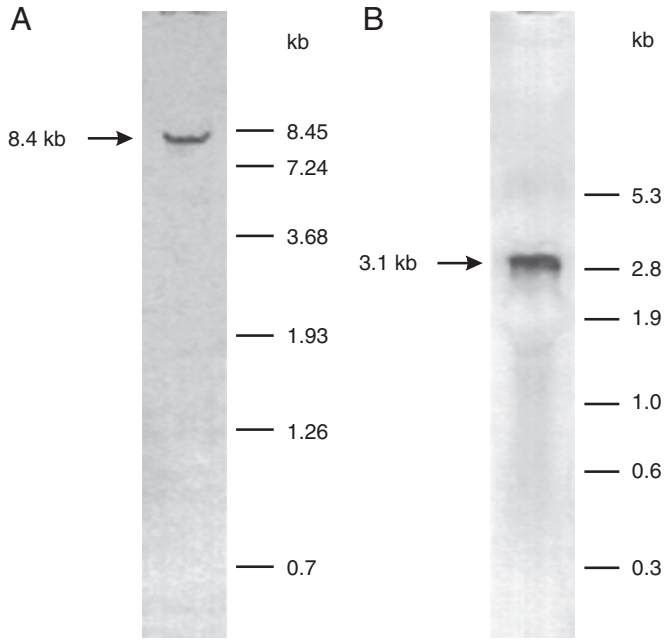

Fig. **1.** (A). Southern analysis of transketolase (TKL) from the glaucocystophyte *Cyanophora paradoxa*. The *C. paradoxa* nuclear DNA digested with *EcoR* I was probed with a digoxigenin-labeled 800-bp TKL fragment. (B). Northern analysis of *C. paradoxa* TKL. The same hybridization results (3.1 kb band) were obtained from total RNA and mRNA.

In contrast to STPs from higher plants, negatively charged amino acids are relatively frequent in *Cyanophora* pre-sequences, which in general results in a moderate to low positive net charge: +2 to +4 (Jakowitsch et al. 1996; Steiner et al. 2000). However, the high number of basic residues in the case of pre-TKL creates a net charge of +8 (despite the presence of four negatively charged amino acids), comparable (or even higher) to what is found in higher plant precursors (Table 1).

Cyanophora paradoxa pre-TKL protein generated by in vitro transcription/translation was imported into isolated cyanelles occurred with high efficiency. The processed mature protein was protease-protected through the cyanelle envelope (Fig. 2, lanes 2 and 3). Import in the dark at 0 °C led to precursor binding and to a low degree of processing. However, the processed band appeared to be susceptible to thermolysin, and thus obviously was not internalized (Fig. 2, lanes 4 and 5). Isolated pea chloroplasts yielded very similar results during heterologous in vitro import (Fig. 3). There was no qualitatively detectable decrease in efficiency of precursor translocation or processing compared with the homologous import experiment.

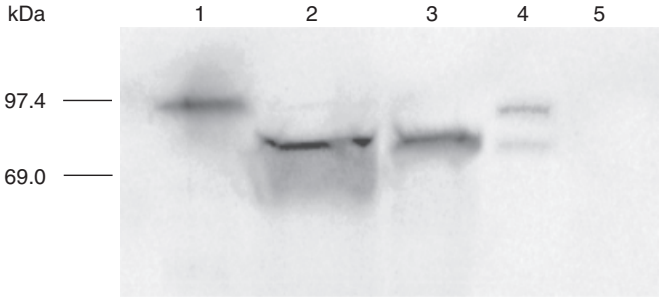

Fig. **2.** Import of pre-transketolase from the glaucocystophyte *Cyanophora paradoxa* into isolated cyanelles. Shown are in vitro translation product (lane 1), import reaction at 25 °C in the light without (lane 2) and with thermolysin treatment (lane 3), import reaction at 0 °C in the dark without (lane 4) and with thermolysin treatment (lane 5).

The 142 aa pre-sequence of *E. gracilis* TKL (Fig. 4A) belongs to class I as defined by Durnford and Gray (2006). The distance (60 aa) between the two hydrophobic domains obeys well the ''60 ± 8 rule''. Other features are also typical for *E. gracilis* chloroplast pre-sequences: acidic amino acid are not as infrequent as in higher plant chloroplast STPs, especially in the region of the transit peptide immediately preceding the stop transfer domain (STD), and a number of basic amino acid immediately after the STD ensure a positive net charge (Durnford and Gray 2006). The estimated cleavage sites for signal peptidase and stroma processing protease (Fig. 4A) were inferred from a compilation of *E. gracilis* STPs according to the criteria used by Durnford and Gray (2006). The mature protein is very similar in sequence to *C. paradoxa* TKL (Fig. 4B), albeit to a slightly lesser extent than enzymes from cyanobacteria and higher plants.

The NNet planar graph of distance splits includes cyanobacterial and plant sequences, homologues from heterotrophic protozoa and fungi, and a large number of eubacteria. The sequences from *C. paradoxa*, *E. gracilis*, and plants form a group that is clearly

Table **1.** Properties of the stroma-targeting peptide of pre-transketolase from the glaucocystophyte *Cyanophora paradoxa*

Position	Feature	Comments
1–8(9)	M(A/S)A**F**VxxVP	Common to all cyanelle transit sequences
2, 10, 12-15, 21, 22, 27, 36, 45, 49, 53, 55, 71, 74, 77, 86	S or T	Weakly conserved in transketolase pre-sequences from other plants
28, 33, 37, 46, 48, 61, 67, 68, 75, 76, 84, 87	R or K	Net charge +8, usually +2 to +4
29, 44, 70, 78	D or E	Somewhat depleted in *C. paradoxa*, very rare in plants
88–90	VAA ↓	Putative processing site

The conserved phenylalanine residue is indicated in bold.

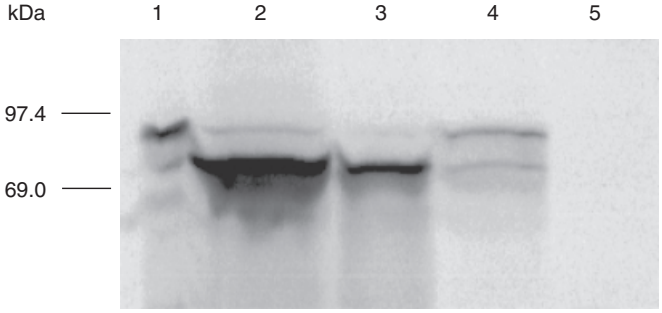

Fig. **3.** Import of pre-transketolase from the glaucocystophyte *Cyanophora paradoxa* into isolated pea chloroplasts. The lanes contain: in vitro translation product (lane 1), import reaction at 25 °C in the light without (lane 2) and with thermolysin treatment (lane 3), import reaction at 0 °C in the dark without (lane 4) and with thermolysin treatment (lane 5).

A

```
1   MALLGNKNESTFKWVAMGVGAGVVAAVLLSATHDNALYTVPATRVTSIS   49

50  AVNPVTYQRVNFDNVLQAAPSASVENQSEIYAASGHNVTGLAAVLALPL   98

99  AAAVGFLFGRSKKSTVAMATYSGKPIEVVSLGEKPQLPKAERRSTALQG  147
```

B

```
Egra: 148 AHAGLIINR--TGPMTD-DELCVNTIRFLAVDGVNKANSGHPGAPMGQAPIAHVLWNEEM 204
          A  G +IN   T  MT  D+ C+N+IRFLA+D + KA SGHPG PMG AP+++VL+NE M
Cpar: 90  ASTGFVINNQATAAMTPVDDKCINSIRFLAIDAIEKAKSGHPGLPMGCAPMSYVLFNEFM 149

Egra: 205 KYNPKNPYFVNRDRFVLSSGHGCMLQYALLHLTGYDDVTLDQLKSFRQWGSKTPGHPENF 264
          K+NPKNP +++RDRFVLS+GHGCMLQYALL+LTGYD V ++ +K+FRQW S+ PGHPENF
Cpar: 150 KFNPKNPDWIDRDRFVLSAGHGCMLQYALLYLTGYDSVGIEDIKTFRQWESECPGHPENF 209

Egra: 265 ETRGVEVTTGPLGMGISNAVGLAAAEAHLAAVYNKPGHTLIDHYTYTIAGDGCFQEGISH 324
           T+G+EVTTGPLG GI+ AVGLA  EAHLAA +NKP   LIDHYTY I GDGC  EG++
Cpar: 210 VTKGIEVTTGPLGQGIAQAVGLAMGEAHLAARFNKPDVKLIDHYTYVIMGDGCNMEGVAA 269

Egra: 325 EACSYAGHLKLGKLIAFYDDNNITIDGETSLSFTEDVAKRYEAYGWQVLKVADGNTDVNG 384
          EA S AGH  LG LIA YDDN I+IDG T +SFTEDV KRY++YGW  + V DGNTD+N
Cpar: 270 EAASLAGHYGLGNLIALYDDNEISIDGNTDISFTEDVTKRYQSYGWHTVVVEDGNTDINA 329

Egra: 385 IRKAIAQAKAEKNKPTLIMVKTVIGYGAPTKANSHDAHGAPLGKDEAAAARKNLGWEFGE 444
          IRKAIA+AKA  +KP+L+ ++T IGYG+P KANS+  HGA LG  E  A R+NL W +
Cpar: 330 IRKAIAEAKAVTDKPSLVTIRTTIGYGSPNKANSYAVHGAALGDKEVDATRQNLNWPYAP 389

Egra: 445 FEIPEQALNTFRQAIPRGAAVEAEWNKRFEAYKQAYPELAKQFQDTVLDNKLPEGWEKAL 504
          FEIPE+A+N +R+AIP+G + E EWNK+F  YK  YP+ A  F+  V+  +LP  WEK L
Cpar: 390 FEIPEEAMNKWREAIPKGKSAEDEWNKKFAEYKAKYPQEAADFEKYVMKKELPANWEKCL 449

Egra: 505 PTY-KAEDKAFATRINSQKCINALAPVLPGFMGGSADLAPSNMTLMECTGDFLAGQYENR 563
          P Y  A D   ATRI +   +NA+A  +P F+GGSADLA SNMTL++   DF     E R
Cpar: 450 PVYDPATDAGDATRILTGNTLNAIADAVPTFLGGSADLASSNMTLLKKYADFQKTSPEGR 509

Egra: 564 NFRFGVREFGMGAVANALALHKSGIIPYCATFLIFSDYMRNAIRIAALSQAGTIFVMTHD 623
          N RFGVREF M A+AN L LH SG+IPY ATFL+FSDYMR A+R+++LS+  TI+V+THD
Cpar: 510 NLRFGVREFAMAAIANGLHLHPSGLIPYGATFLVFSDYMRAAMRLSSLSKCRTIYVLTHD 569

Egra: 624 SVALGEDGPTHQPVEIIASLRLIPQLAVVRPCDGNETSGAYKMAVLRSNGIGKPGISGRP 683
          S+  GEDGPTHQP+E +A+ R +P   V RPCDGNE SGAYK+AV
Cpar: 570 SIGAGEDGPTHQPIEHLAAHRAVPNTFVFRPCDGNEVSGAYKVAV----------EE 616

Egra: 684 KTFPTLLALSRQVLPNQKGSSIDAVAKGGYTIQD--CEGKPDLILIGTGSEVQLCIESAA 741
          +  P+L+ L+RQ +P   G+SI   AKG Y + D    GKPDLIL+GTGSEV LC+ +A
Cpar: 617 RETPSLMILTRQKIPTLDGTSIANTAKGAYVLSDNSTNGKPDLILMGTGSEVHLCVNAAE 676

Egra: 742 ALAKEGKKVRVVSMPCTEFFDEQPKEYRDSVLIPGVP-AVSVEAGVTAGWQKY---SHAQ 797
          A+ KEGK VRVVSMP  E F+ Q  EY  SV       A++VEA  + GW KY     A
Cpar: 677 AIRKEGKTVRVVSMPSWELFERQSAEYEASVFPKDCKRALAVEAASSFGWHKYFGDEGAM 736

Egra: 798 VGIDSFGASAPGDKCMKEFGMTVENVVATAKSLL 831
          V ID FGASAPGDK M+EFG T ENV+A AK LL
Cpar: 737 VSIDGFGASAPGDKLMQEFGFTTENVIAQAKKLL 770
```

Fig. **4.** (**A**). The N-terminal sequence of the *Euglena gracilis* transketolase (TKL) precursor is given preceding the BLAST alignment starting with pos. 148. ▼ Putative processing site of signal peptidase. ↓ Putative processing site of stroma processing protease. The two hydrophobic domains in the signal peptide and the transit peptide (stop transfer domain), respectively, are underlined. (**B**). BLAST alignment showing the identity score (56%) between the mature TKLs from both algae.

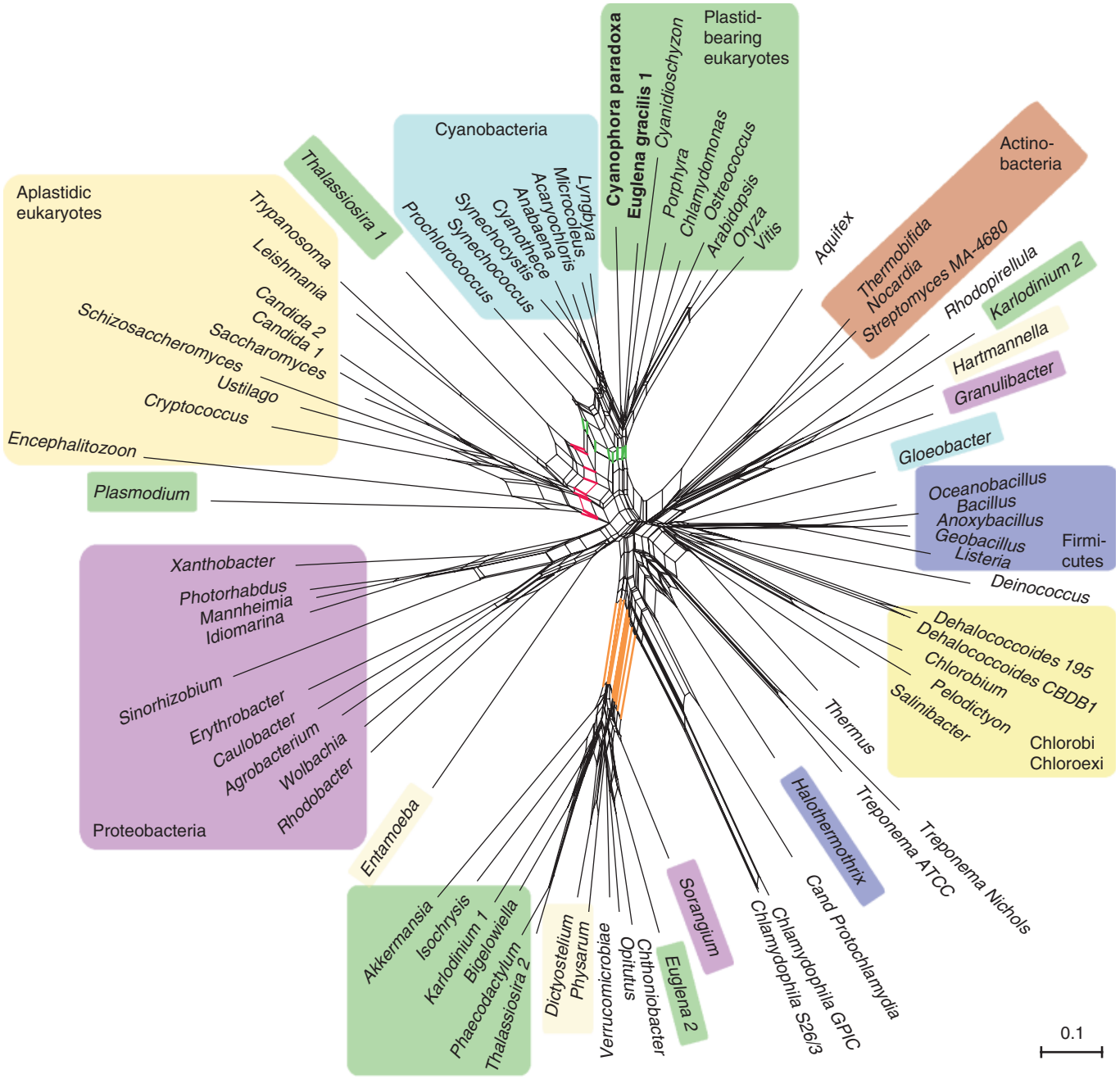

Fig. **5.** Neighbor-net graph of transketolase (TKL) homologues from plants, algae, and eubacteria. Sequences from eukaryotes with plastids are indicated with green shading, those from cyanobacteria in cyan, those from α-proteobacteria in purple, and those from eukaryotes that lack plastids in tan. The split (bipartition) uniting TKL sequences from cyanobacteria, eukaryotes with primary plastids, and *Euglena* (*Euglena 1*) is indicated with green highlighting. The split uniting TKL sequences from fungi, *Plasmodium*, *Thalassiosira*, and kinetoplastids (the host lineage for the *Euglena* secondary symbiosis), is indicated with pink highlighting. The Verrumicrobia/Proteobacteria branch (see text) is indicated with orange highlighting. Sequences were retrieved from GenBank (http://www.ncbi.nlm.nih.gov/) and from genome projects through TIGR (http://www.tigr.org). A complete list of sequence sources, species/strain names and accession numbers is given in Table S1.

most similar to cyanobacterial homologues, supporting a single cyanobacterial origin of cyanelles and plastids (highlighted in green, Fig. 5). *Cyanophora* TKL shares 60–61% aa identity with homologues from higher plants and 56–66% aa identity with homologues from cyanobacteria. The TKL homologues from trypanosomatids, fungi, and the protists *Thalassiosira* and *Plasmodium*, which possess secondary, rhodophyte-derived plastids, are distinct from those of higher plants, cyanobacteria, and *Euglena* (highlighted in red, Fig. 5). Notably, TKL homologues from animals and ciliates are related to, but highly distinct from, plant and fungal TKL (Flechner et al. 1996; Mittenhuber 2001) (see also supporting information, Fig. S1), and they share greater similarity to TKL homologues from archaebacteria, actinobacteria, and several sequenced eubacterial genomes (Fig. 6).

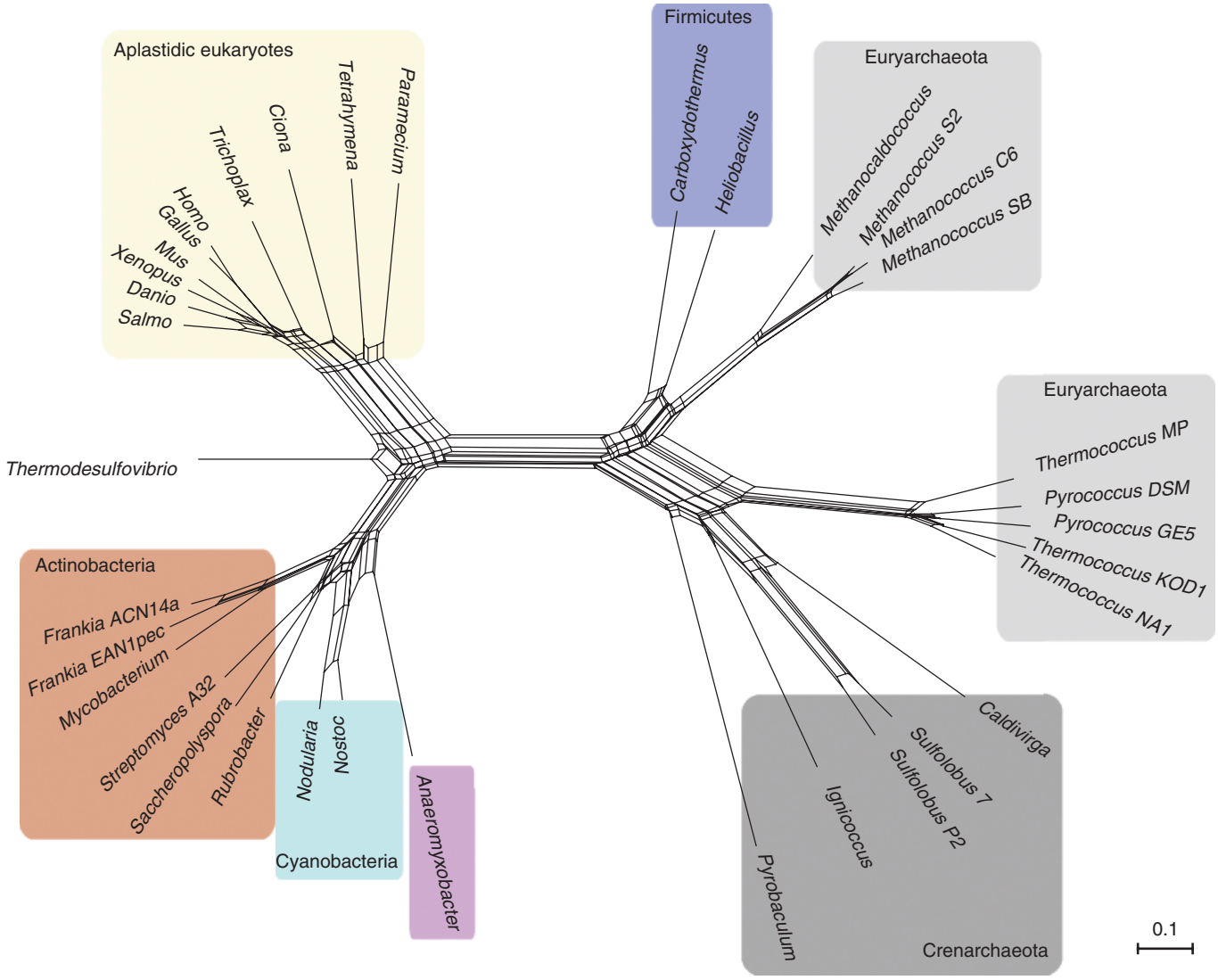

Fig. **6.** Neighbor-net graph of transketolase homologues from animals and archaebacteria. Sequences from eukaryotes that lack plastids are indicated with tan shading, those from crenarchaeotes in gray, those from euryarchaeotes in light gray. Homologues from eubacterial genomes are shaded as labelled in the figure. Sequences were retrieved from GenBank (http://www.ncbi.nlm.nih.gov/) and from genome projects through TIGR (http://www.tigr.org). A complete list of sequence sources, species/strain names and accession numbers is given in Table S1.

DISCUSSION

A crucial distinction between cyanelle and chloroplast transit sequences appears to be the N-terminal consensus sequence M(A/S)AFVxxVP in the former, which is centered around a conserved phenylalanine residue (Table 1). Parallels to higher plant STPs are proline and/or glycine residues (position 11) at the end of the N-terminal hydrophobic domain (Bruce 2000), a positive net charge, a high content of hydroxylated amino acids, and moderately conserved processing sites (Emanuelsson, Nielsen, and von Heijne 1999).

The invariant phenylalanine residue in the N-terminal domain of cyanelle transit peptides (Steiner and Löffelhardt 2002) that was also found in the case of pre-TKL might play a role in the heterologous import enigma. It is not encountered in STPs from higher plants and is thought to fulfil a crucial function in cyanelle import (Steiner and Löffelhardt 2005; Steiner et al. 2005a; Wunder et al. 2007). The organelle wall constitutes no barrier for the import of an estimated 2,000 precursor polypeptides (Abdallah, Leister, and Salamini 2000) from the cytosol into cyanelles. It does not seem either to be the reason for the thus far unsuccessful import experiments of heterologous (chloroplast) precursors into cyanelles.

To date, two *Cyanophora* precursors were tested and both were imported into isolated pea chloroplasts: pre-FNR (ferredoxin-$NADP^+$ oxidoreductase) (Jakowitsch et al. 1996) and pre-cytochrome c_6 (Steiner et al. 2005b). This also now applies for *Cyanophora* pre-TKL, which is imported with an efficiency comparable to that for the homologous system. It seems that cyanelle precursors from *C. paradoxa* are, in general, readily translocated and processed by the import apparatus of pea chloroplasts. This suggests the presence of a similar Toc-Tic type translocon (Soll and Schleiff 2004) in the membranes of both plastid types. However, the cyanelle import apparatus requires a phenylalanine residue in the N-terminal domain of the transit sequence and higher plant plastid precursors are only imported efficiently into cyanelles when it is engineered in this very position (Steiner and Löffelhardt 2005). The present findings are at odds with the pre-

vious suggestion (Löffelhardt et al. 1998) that the high positive charge of higher plant transit peptides might inhibit their import into cyanelles.

The phylogenetic comparison of plastid genes (Martin et al. 1998) and the comparison of plastid genome organization (Löffelhardt, Bohnert, and Bryant 1997; Reith and Munholland 1995; Stoebe and Kowallik 1999) led to the current view that there was a single primary endosymbiotic event. This would imply the monophyletic origin of the kingdom Plantae, which is supported through phylogenetic analysis of nuclear genes (Moreira et al. 2000; Rodríguez-Ezpeleta et al. 2005) and which necessitates a homologous protein import mechanism for all primary plastids. The plastid import machinery of the green and red plastid lineages is homologous, providing strong support for a single common origin of these plastids, as the genome sequence of the red alga *Cyanidiodoschyzon merolae* clearly revealed (Matsuzaki et al. 2004). Evidence for conservation of the components of the cyanelle import machinery with that of red and green plastids can be anticipated. Preliminary results point toward Toc75 and Tic110 as the minimal components of a primordial import apparatus in the cyanelle envelope (Yusa, Steiner, and Löffelhardt 2008).

A notable aspect of TKL phylogeny is the distinctly different isoform possessed by animals and archaebacteria, which share only about 40% aa identity with homologues from photosynthetic eukaryotes. The animal homologues share an average 47.8% aa identity with TKL homologues from archaebacterial genomes, which are split into N-terminal and C-terminal domains. In the case of *Methanocaldococcus jannaschii* the corresponding proteins are Q58094 and Q58092. So far, both TKL isoforms are found among eukaryotes, cyanobacteria, actinobacteria, and proteobacteria, but we identified no genomes in which both TKL types co-occurred.

The basal position of cyanelles between chloroplasts and free-living cyanobacteria is apparent from phylogenetic analyses of many traits (Jakowitsch et al. 1996; Ma et al. 2001; Martin et al. 1998) and is consistent with shared similarities among TKL sequences. Glaucocystophytes are not only the most ancient among the phototrophic eukaryotes known to date but are also living fossils in that they have retained prokaryotic features not found among other recent phyla, such as the peptidoglycan wall and, possibly, the carboxysomes. These Rubisco-containing microcompartments are involved in cyanobacterial CO_2 fixation in the context of the inorganic carbon concentrating mechanism (Raven 2003).

Although TKL in eukaryotes with primary plastids is clearly of cyanobacterial origin, it is not encoded in any chloroplast genome, suggesting that it was transferred to the nucleus early in evolution. Consequently, the *Euglena* TKL gene must have undergone two transfers, once from the primary plastid to the nucleus in the green lineage, and once from the nucleus of the green secondary symbiont to the nucleus of the secondary host. Several similar transfers have been previously reported (Ahmadinejad et al. 2007; Archibald et al. 2003; Deane et al. 2008; Henze et al. 1995), but genome-wide analyses of such transfers among completely sequenced genomes of various eukaryotes with secondary plastids that would reveal quantitative estimates for the frequency of such events during evolution are still lacking. By contrast, genome wide estimates for the frequency of gene transfers from cyanobacteria to the nucleus in the context of primary symbiosis indicate a very substantial endosymbiotic contribution to the complement of plant nuclear genes, on the order of 10–20%, depending upon the species sampling and methodological factors (Deusch et al. 2008; Lane and Archibald 2008). However, for genes common to symbionts and hosts during endosymbiosis, it has long been observed that usually only one persists, as they are functionally redundant (Martin and Schnarrenberger 1997). In this context it is noteworthy that the diatom *Thalassiosira* and the alveolate *Plasmodium*, both of which harbor plastids of red algal ancestry (Gould et al. 2008), have retained the TKL gene of their host lineage rather than having acquired the gene from their endosymbiont, as in the case of *Euglena*.

Rogers et al. (2007) reported EST-derived sequences with sequence similarity to TKL from several protists and discussed aspects of hypothetical subcellular localization of the putative protein products, but no evidence for import or localization of the putative gene products was provided. They found that one group of TKL-related sequences of unknown function in some photosynthetic protists and *Dictyostelium* was closely related to the genes in chlamydiae (Rogers et al. 2007). While overlooking earlier studies of TKL gene evolution (Flechner et al. 1996; Martin and Schnarrenberger 1997; Matsuzaki et al. 2004; Mittenhuber 2001) and subcellular localization of TKL activity in photosynthetic eukaryotes (Schnarrenberger et al. 1995), Rogers et al. (2007) noted that with better sampling, that result may change, which it has (Fig. 5). The chlamydial branch of Rogers et al. (2007) linking the functionally uncharacterized TKLs from EST projects, including the sequence labeled ''*Euglena 2*,'' is now a Verrumicrobia/Proteobacteria branch. It is also worth noting that other sequences link *Euglena* genes with *Dictyostelium* homologues (Torrents et al. 2006), the possible evolutionary significance of which is not the topic of this paper.

Our interest concerns functional and evolutionary aspects of *Cyanophora* TKL and the inheritance of those specific homologues during secondary endosymbiosis. *Cyanophora* TKL is imported into plastids, like the enzymatically active homologues from higher plants, the evolution (Flechner et al. 1996), subcellular localization (Schnarrenberger et al. 1995), and crystal structure (Gerhardt et al. 2003) of which have been characterized previously. The components involved in protein import into the two membrane-bounded chloroplasts of higher plants are reasonably well circumscribed (Kalanon and McFadden 2008), and the core of these components seem to be present and operative in cyanelles as well (Wunder et al. 2007; Yusa et al. 2008). However, in the case of secondary plastids surrounded by more than two membranes, which necessitates bipartite pre-sequences (van Dooren et al. 2001; Schwartzbach, Osafune, and Löffelhardt 1998), the situation is more diverse, though, in our belief, the final steps of translocation into the stroma also occur via Toc/Tic-related translocons (Steiner and Löffelhardt 2005). In *Euglena* (three membranes), plastid targeting involves the Golgi (Slavikova et al. 2005) and a hydrophobic domain in the transit sequence while in the malaria parasite *Plasmodium* (four membranes), belonging to the alveolates, plastid targeting is Golgi independent (Tonkin et al. 2006). In *Plasmodium* and in the chlorophyll *c*-containing chromists, whose plastids are also surrounded by four membranes, the ER-associated degradation (ERAD)-like pathway appears to mediate protein translocation across the second outermost plastid membrane in species examined for this trait thus far (Hempel et al. 2007; Sommer et al. 2007). However, no evidence for the participation of the ERAD-like pathway in plastid protein import has been reported for *Euglena*. Clearly, the evolution of protein import was an essential step allowing endosymbionts to be converted into organelles during evolution, and further study of the underlying mechanisms should improve our understanding of the quantitative and qualitative contributions of endosymbiosis to cell evolution.

ACKNOWLEDGMENTS

This work was supported by grants from the Austrian ''Fonds zur Förderung der Wissenschaftlichen Forschung'' (P14573 and

P19683, to W.L.), and by a grant from the Deutsche Forschungsgemeinschaft (within SFB-Tr1) to W.M.

LITERATURE CITED

Abdallah, F., Leister, D. & Salamini, F. 2000. A prediction of the size and evolutionary origin of the proteome of chloroplasts of *Arabidopsis*. *Trends Plant Sci.*, **5**:141–142.

Ahmadinejad, N., Dagan, T. & Martin, W. 2007. Genome history in the symbiotic hybrid *Euglena gracilis*. *Gene*, **402**:35–39.

Anweiler, A., Hipskind, R. A. & Wirth, T. 1991. A strategy for efficient in vitro translation of cDNAs using the rabbit β-globin leader sequence. *Nucleic Acids Res.*, **19**:3750.

Apt, K. E., Hoffman, N. E. & Grossman, A. J. 1993. The γ-subunit of R-phycoerythrin and its possible mode of transport into the plastids of red algae. *J. Biol. Chem.*, **268**:16208–16215.

Archibald, J. M., Rogers, M. B., Toop, M., Ishida, K. & Keeling, P. J. 2003. Lateral gene transfer and the evolution of plastid-targeted proteins in the secondary plastid-containing alga *Bigelowiella natans*. *Proc. Natl. Acad. Sci. USA*, **100**:7678–7683.

Bruce, B. D. 2000. Chloroplast transit peptides: structure, function and evolution. *Trends Cell Biol.*, **10**:440–447.

Bryant, D. & Moulton, V. 2004. Neighbor-Net: an agglomerative method for the construction of planar phylogenetic networks. *Mol. Biol. Evol.*, **21**:255–265.

Deane, J. A., Fraunholz, M., Su, V., Maier, U. G., Martin, W., Durnford, D. G. & McFadden, G. I. 2008. Evidence for nucleomorph to host nucleus gene transfer: light-harvesting complex proteins from cryptomonads and chlorarachniophytes. *Protist*, **151**:239–252.

Deusch, O., Landan, G., Roettger, M., Gruenheit, N., Kowallik, K. V., Allen, J. F., Martin, W. & Dagan, T. 2008. Genes of cyanobacterial origin in plant nuclear genomes point to a heterocyst-forming plastid ancestor. *Mol. Biol. Evol.*, **25**:748–761.

van Dooren, G. G., Schwartzbach, S. D., Osafune, T. & McFadden, G. I. 2001. Translocation of proteins across multiple membranes of complex plastids. *Biochim. Biophys. Acta*, **1541**:34–53.

Durnford, D. G. & Gray, M. W. 2006. Analysis of *Euglena gracilis* plastid-targeted proteins reveals different classes of transit sequences. *Eukaryot. Cell*, **5**:2079–2091.

Emanuelsson, O., Nielsen, H. & von Heijne, G. 1999. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.*, **8**:978–984.

Flechner, A., Dreßen, U., Westhoff, P., Henze, K., Schnarrenberger, C. & Martin, W. 1996. Molecular characterization of transketolase (EC 2.2.1.1) active in the Calvin cycle of spinach chloroplasts. *Plant Mol. Biol.*, **32**:475–484.

Gerhardt, S., Echt, S., Busch, M., Freigang, J., Auerbach, G., Bader, G., Martin, W., Bacher, A., Huber, R. & Fischer, M. 2003. Structure and properties of an engineered transketolase from maize. *Plant Physiol.*, **132**:1941–1949.

Gould, S. B., Waller, R. F. & McFadden, G. I. 2008. Plastid evolution. *Annu. Rev. Plant. Biol.*, **59**:491–517.

Hempel, F., Bozarth, A., Sommer, M. S., Zauner, S., Przyborski, J. M. & Maier, U. G. 2007. Transport of nuclear-encoded proteins into secondarily evolved plastids. *Biol. Chem.*, **388**:899–906.

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

Hoffmeister, M., van der Klei, A., Rotte, C., van Grinsven, K. W. A., van Hellemond, J. J., Henze, K., Tielens, A. G. M. & Martin, W. 2004. *Euglena gracilis* rhodoquinone:ubiquinone ratio and mitochondrial proteome differ under aerobic and anaerobic conditions. *J. Biol. Chem.*, **279**:22422–22429.

Huson, D. H. & Bryant, D. 2006. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.*, **23**:254–267.

Jakowitsch, J., Neumann-Spallart, C., Ma, Y., Steiner, J. M., Schenk, H. E. A., Bohnert, H. J. & Löffelhardt, W. 1996. In vitro import of pre-ferredoxin-NADP+-oxidoreductase from *Cyanophora paradoxa* into cyanelles and into pea chloroplasts. *FEBS Lett.*, **381**:153–155.

Kalanon, M. & McFadden, G. I. 2008. The chloroplast protein translocation complexes of *Chlamydomonas reinhardtii*: a bioinformatic comparison of Toc and Tic components in plants, green algae and red algae. *Genetics*, **179**:95–112.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**:680–685.

Lane, C. E. & Archibald, J. M. 2008. The eukaryotic tree of life: endosymbiosis takes its TOL. *Trends Ecol. Evol.*, **23**:268–275.

Lockhart, P. J., Steel, M. A., Hendy, M. D. & Penny, D. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. *Mol. Biol. Evol.*, **11**:605–612.

Löffelhardt, W., Bohnert, H. J. & Bryant, D. A. 1997. The complete sequence of the *Cyanophora paradoxa* cyanelle genome (*Glaucocystophyceae*). *Plant Syst. Evol.*, **11**(Suppl.):149–162.

Löffelhardt, W., Steiner, J. M., Ma, Y. & Jakowitsch, J. 1998. Protein translocation into and within an armoured plastid: the cyanelle of *Cyanophora paradoxa*. *In*: Garab, G. (ed.), Photosynthesis: Mechanisms and Effects. Kluwer Academic Publishers, Dordrecht, the Netherlands. **IV**:3057–3061.

Ma, Y., Jakowitsch, J., Maier, T. L., Bayer, M. G., Müller, N. E., Schenk, H. E. A. & Löffelhardt, W. 2001. ATP citrate lyase of the glaucocystophyte alga *Cyanophora paradoxa* is a cytosolic enzyme: characterisation of the large subunit at the cDNA and genome level. *Mol. Genet. Genom.*, **266**:231–238.

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

Martin, W., Stoebe, B., Goremykin, V., Hansmann, S., Hasegawa, M. & Kowallik, K. 1998. Gene transfer to the nucleus and the evolution of chloroplasts. *Nature*, **393**:162–165.

Matsuzaki, M., Misumi, O., Tadasu, S.-I., Maruyama, S., Takahara, M., Miyagishima, S.-Y., Mori, T., Nishida, K., Yagisawa, F., Nishida, K., Yoshida, Y., Nishimura, Y., Nakao, S., Kobayashi, T., Momoyama, Y., Higashiyama, T., Minoda, A., Sano, M., Nomoto, H., Oishi, K., Hayashi, H., Ohta, F., Nishizaka, S., Haga, S., Miura, S., Morishita, T., Kabeya, Y., Terasawa, K., Suzuki, Y., Ishii, Y., Asakawa, S., Takano, H., Ohta, N., Kuroiwa, H., Tanaka, K., Shimizu, N., Sugano, S., Sato, N., Nozak, H., Ogasawara, N., Kohara, Y. & Kuroiwa, T. 2004. Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature*, **428**:653–657.

Mittenhuber, G. 2001. Phylogenetic analyses and comparative genomics of vitamin B_6 (pyridoxine) and pyridoxal phosphate biosynthesis pathways. *J. Mol. Microbiol. Biotechnol.*, **3**:1–20.

Moreira, D., Le Guyader, H. & Philippe, H. 2000. The origin of red algae and the evolution of chloroplasts. *Nature*, **405**:69–72.

Pfanzagl, B., Zenker, A., Pittenauer, E., Allmaier, G., Martinez-Torrecuadrada, J., Schmid, E. R., de Pedro, M. A. & Löffelhardt, W. 1996. Primary structure of cyanelle peptidoglycan of *Cyanophora paradoxa*: a prokaryotic cell wall as part of an organelle envelope. *J. Bacteriol.*, **178**:332–339.

Raven, J. A. 2003. Carboxysomes and peptidoglycan walls of cyanelles: possible physiological functions. *Eur. J. Phycol.*, **38**:47–53.

Reith, M. & Munholland, J. 1995. Complete nucleotide sequence of the *Porphyra purpurea* chloroplast genome. *Plant Mol. Biol. Reptr.*, **13**:333–335.

Rodríguez-Ezpeleta, N., Brinkmann, H., Burey, S. C., Roure, B., Burger, G., Löffelhardt, W., Bohnert, H. J., Philippe, H. & Lang, B. F. 2005. Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. *Curr. Biol.*, **15**:1325–1330.

Rogers, M. B., Watkins, R. F., Harper, J. T., Durnford, D. G., Gray, M. W. & Keeling, P. J. 2007. A complex and punctate distribution of three eukaryotic genes derived by lateral gene transfer. *BMC Evol. Biol.*, **7**:89.

Schnarrenberger, C., Flechner, A. & Martin, W. 1995. Enzymatic evidence indicating a complete oxidative pentose phosphate in the chloroplasts and an incomplete pathway in the cytosol of spinach leaves. *Plant Physiol.*, **108**:609–614.

Schwartzbach, S. D., Osafune, T. & Löffelhardt, W. 1998. Protein import into cyanelles and complex chloroplasts. *Plant Mol. Biol.*, **38**: 247–263.

Slavikova, S., Vacula, R., Fang, Z. W., Ehara, T., Osafune, T. & Schwartzbach, S. D. 2005. Homologous and heterologous reconstitution of Golgi to chloroplast transport and protein import into the complex chloroplasts of *Euglena*. *J. Cell Sci.*, **118**:1651–1661.

Soll, J. & Schleiff, E. 2004. Protein import into chloroplasts. *Nat. Rev. Mol. Cell. Biol.*, **5**:198–208.

Sommer, M. S., Gould, S. B., Lehmann, P., Gruber, A., Przyborski, J. M. & Maier, U. G. 2007. Der1-mediated pre-protein import into the periplastid compartment of chromalveolates? *Mol. Biol. Evol.*, **24**:1–11.

Steiner, J. M. & Löffelhardt, W. 2002. Protein import into cyanelles. *Trends Plant Sci.*, **6**:72–77.

Steiner, J. M. & Löffelhardt, W. 2005. Protein translocation into and within cyanelles. *Mol. Memb. Biol.*, **22**:123–132.

Steiner, J. M., Yusa, F., Pompe, J. A. & Löffelhardt, W. 2005a. Homologous protein import machineries in chloroplasts and cyanelles. *Plant J.*, **44**:646–652.

Steiner, J. M., Serrano, A., Allmaier, G., Jakowitsch, J. & Löffelhardt, W. 2000. Cytochrome c_6 from *Cyanophora paradoxa*. Characterization of the protein and the cDNA of the precursor and import into isolated cyanelles. *Eur. J. Biochem.*, **267**:4232–4241.

Steiner, J. M., Berghöfer, J., Yusa, F., Pompe, J. A., Klösgen, R. B. & Löffelhardt, W. 2005b. Conservative sorting in a primitive plastid: the cyanelle of *Cyanophora paradoxa*. *FEBS J.*, **272**:987–998.

Stoebe, B. & Kowallik, K. 1999. Gene-cluster analysis in chloroplast genomics. *Trends Genet.*, **15**:344–347.

Teige, M., Melzer, M. & Süss, K.-H. 1998. Purification, properties and in situ localization of the amphibolic enzymes D-ribulose 5-phosphate 3-epimerase and transketolase from spinach chloroplasts. *Eur. J. Biochem.*, **252**:237–244.

Thollesson, M. 2004. LDDist: a Perl module for calculating LogDet pairwise distances for protein and nucleotide sequences. *Bioinformatics*, **20**:416–418.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. 1994. CLUSTALW: 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.

Tonkin, C. J., Struck, N. S., Mullin, K. A., Stimmler, L. M. & McFadden, G. I. 2006. Evidence for Golgi-independent transport from the early secretory pathway to the plastid in malaria parasites. *Mol. Microbiol.*, **61**:614–630.

Torrents, E., Trevisiol, C., Rotte, C., Hellman, U., Martin, W. & Reichard, P. 2006. *Euglena gracilis* ribonucleotide reductase: the eukaryotic class II enzyme and the antiquity of eukaryotic B_{12}-dependence. *J. Biol. Chem.*, **281**:5604–5611.

Wunder, T., Martin, R., Löffelhardt, W., Schleiff, E. & Steiner, J. M. 2007. The invariant phenylalanine of precursor proteins discloses the importance of Omp85 for protein translocation into cyanelles. *BMC Evol. Biol.*, **7**:236.

Yusa, F., Steiner, J. M. & Löffelhardt, W. 2008. Evolutionary conservation of dual Sec translocases in the cyanelles of *Cyanophora paradoxa*. *BMC Evol. Biol.*, **8**:304.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Joint NNet graph of transketolase homologues shown separately in Fig. 5 and Fig. 6, constructed as described in Material and Methods. The complete list of sequence sources, species/strain names and accession numbers is given in Table S1.

Table S1. Complete list of 115 protein sequences used for phylogenetic reconstruction of transketolase homologues. Some bacterial sequences consist of a N- and a C-terminal section in which cases accession numbers for both sections are provided.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Received: 11/30/08, 03/14/09, 06/08/09; accepted: 06/10/09