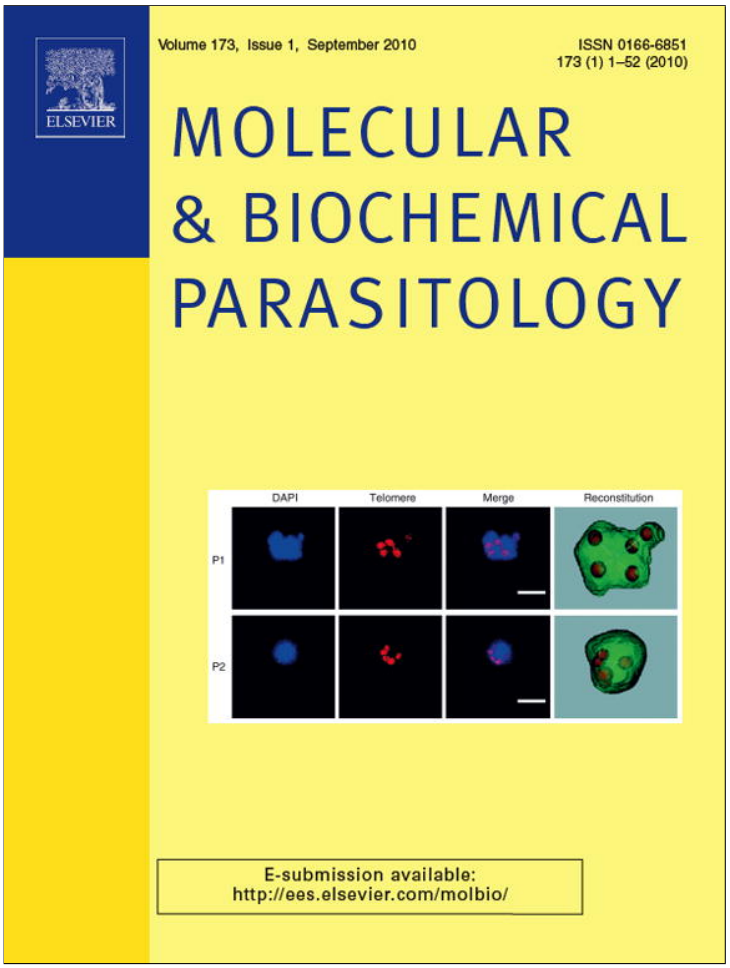

Molecular & Biochemical Parasitology 173 (2010) 43–47

Contents lists available at ScienceDirect

Molecular & Biochemical Parasitology

Short communication

YCF45 protein, usually associated with plastids, is targeted into the mitochondrion of *Trypanosoma brucei*

Jiří Týč, Shaojun Long, Milan Jirků, Julius Lukeš*

Biology Centre, Institute of Parasitology and Faculty of Sciences, University of South Bohemia, České Budějovice (Budweis), Czech Republic

ARTICLE INFO

Article history:
Received 1 March 2010
Received in revised form 30 April 2010
Accepted 5 May 2010
Available online 12 May 2010

Keywords:
Trypanosoma
Plastid
Mitochondrion
Targeting
YCF45
Horizontal gene transfer

ABSTRACT

YCF45 belongs to a family of proteins of unknown function usually located in the chloroplast of plants. Its highly conserved homologues were found in the genomes of several *Trypanosoma* and *Leishmania* species. HA_3-tagging of the YCF45 protein with the start codon as annotated in the Gene^{DB} revealed its cytosolic localization in the cultured procyclic stage of *Trypanosoma brucei*. However, when a more upstream located start codon was used in another HA_3-tagged construct, the resulting protein was targeted to the mitochondrion. We propose that YCF45 was acquired by an ancestral trypanosomatid by horizontal gene transfer and in the absence of a plastid was re-targeted to the mitochondrion.

Since the euglenoids, a sister group of kinetoplastid flagellates, harbor a functional three-membraned secondary green plastid [1,2], it is not surprising that a homologous organelle was searched for also in trypanosomatids and other Kinetoplastida. If these parasitic protists ever carried a plastid, any remnant proteins or pathways would represent an excellent drug target, as is the case in *Plasmodium* and other apicomplexan parasites [3]. Electron microscopy was instrumental to show that a typical plastid is missing in *Trypanosoma* species [4], yet a possibility remained that the common predecessor of kinetoplastids and euglenoids once contained a plastid that was lost in the kinetoplastids after both lineages split [5–8]. An alternative scenario postulates that the organelle was engulfed by the euglenoids in the course of their separate evolutionary history [9,10].

Molecular phylogenetic analyses of selected *Trypanosoma brucei* genes supported the existence of a plastid that was lost in all extant trypanosomatids [6,7], but subsequent availability of three whole trypanosomatid genomes did not reveal any substantial evidence for this notion [2,11]. Moreover, conclusions based on extensive morphological analysis favoured acquisition of the chloroplast only by the euglenoid lineage [10].

The contentious case of a plastid remnant was revived by the unexpected finding of YCF45 in the genomes of *Trypanosoma cruzi*, *T. brucei* and *Leishmania major* [12], and also in *Leishmania braziliensis* and *Leishmania infantum* (Fig. 1A). Furthermore, a recent study claims that the N-terminus of one of the mitochondrial iron-containing superoxide dismutases resembles targeting sequence used by secondary plastids [13]. YCF stands for h*Y*pothetical *C*hloroplast open reading *F*rame, an abbreviation that brings together a group of conserved open reading frames, which in most organisms are one way or another associated with chloroplast and for which a function is presently unknown [14]. Members of this artificial assembly are either encoded by the chloroplast genome [15], or respective genes have been transferred into the nucleus [16]. While a few YCF genes have been assigned a function [17], not a single homologue of YCF45 has been functionally studied so far. Taking into consideration the statement of Opperdoes and Michels [12] that YCF45 was found only in the genomes of chloroplast-containing eukaryotes and in trypanosomatids, it either represents a genuine remnant of this organelle, or has been acquired via lateral (=horizontal) gene transfer by the common ancestor of the *Trypanosoma* and *Leishmania* species.

We have initiated the study of trypanosomatid YCF45 by phylogenetic analysis. Alignment of several plant YCF45 genes with their trypanosomatid homologues revealed an unexpectedly high conservation throughout the central part of the gene. For a 307 amino acids-long part of the protein, sequence identity between the *T. brucei* and *Arabidopsis* and *Synechocystis* YCF45 proteins ranges from 37% to 39% (amino acid sequence similarity is 56% and 57%, respectively), which are surprisingly high values hinting for conservation of function. There is also a very high sequence similarity

* Corresponding author. Tel.: +420 38 7775416; fax: +420 38 5310388.
E-mail address: jula@paru.cas.cz (J. Lukeš).

0166-6851/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.molbiopara.2010.05.002

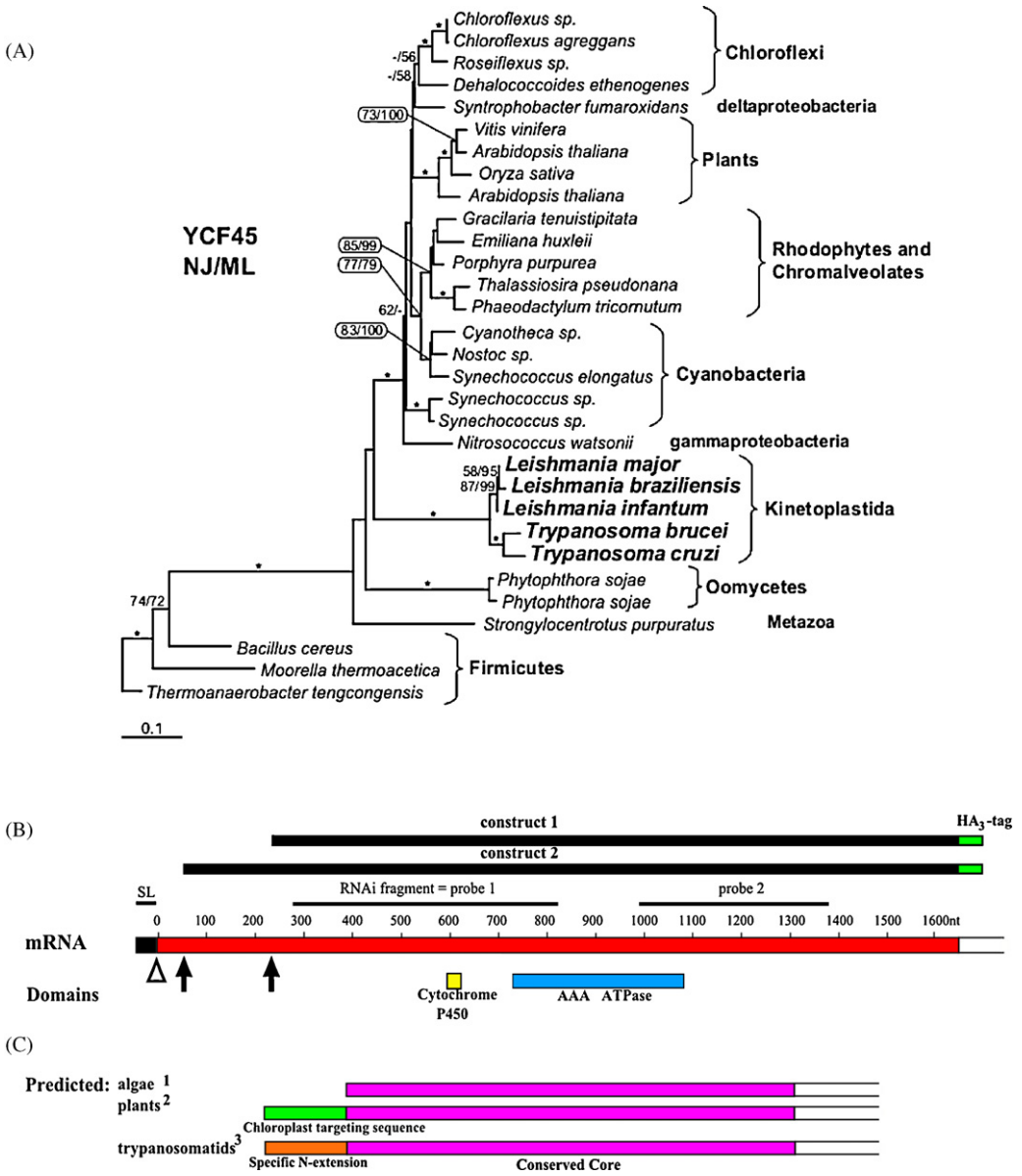

Fig. 1. (A) Phylogeny of YCF45. Phylogenetic tree was inferred from YCF 45 amino acid sequences that were aligned using the Kalign program at http://www.ebi.ac.uk/kalign/. Ambiguously aligned regions as well as gaps were excluded from further analysis. Tree was computed using Neighbor-joining (weighted) as implemented in AsaturA program, with cutoff value = 1.163. Appropriate ML tree (log *lk* = −9477.97519) was computed by WAG model with discrete gamma distribution in 4 + 1 categories as implemented in PhyML. Particular substitutional model was chosen by PROTTEST according to AIC. ML bootstraps were computed from 200 replicates, NJ bootstraps from 1000 replicates. Numbers above branches indicate NJ bootstrap support/maximum likelihood bootstraps. Asterisks indicate bootstrap support higher than 90%. (B) Schematic representation of YCF45 mRNA in *Trypanosoma brucei*. Arrowhead points at the splice acceptor site, black arrows indicate alternative start codons. Above the mRNA regions used for Northern analysis (probes 1 and 2) and for RNA interference (probe 1) are shown. SL stands for splice leader RNA sequence. The uppermost black lines indicate two different YCF45 constructs with the C-terminal HA_3 tags. (C) Schematic view of N-termini of YCF45 from distantly related organisms: [1]in algae, YCF45 is encoded in the chloroplast genome and thus lacks any targeting presequence; [2]in plants, YCF45 is nucleus-encoded and is targeted into the chloroplast via a plastid targeting signal; [3]in trypanosomatids, the YCF45 homologue is nucleus-encoded and contains a specific N-terminal extension, which is likely a mitochondrion-targeting signal.

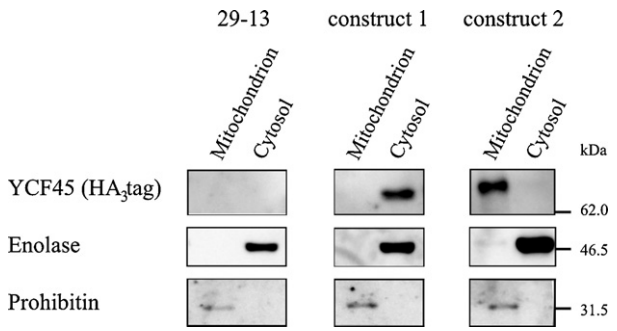

Fig. 2. Intracellular localization of two different inducibly expressed HA_3-tagged YCF45 proteins in the *T. brucei* procyclic stage. Immunoblot analysis of the HA_3-tagged protein (detected using antibody against the HA_3 tag) in the cytosolic and mitochondrial fractions extracted from the parental 29-13 cells (left panel), cells containing construct 1, which starts with the annotated start codon (central panel), and cells containing construct 2, which starts with the newly identified upstream start codon (right panel) (also see Fig. 1B). Polyclonal antibodies against prohibitin and enolase were used as mitochondrial and cytosolic markers, respectively. Cellular fractions were obtained by digitonin fractionation as described elsewhere [19].

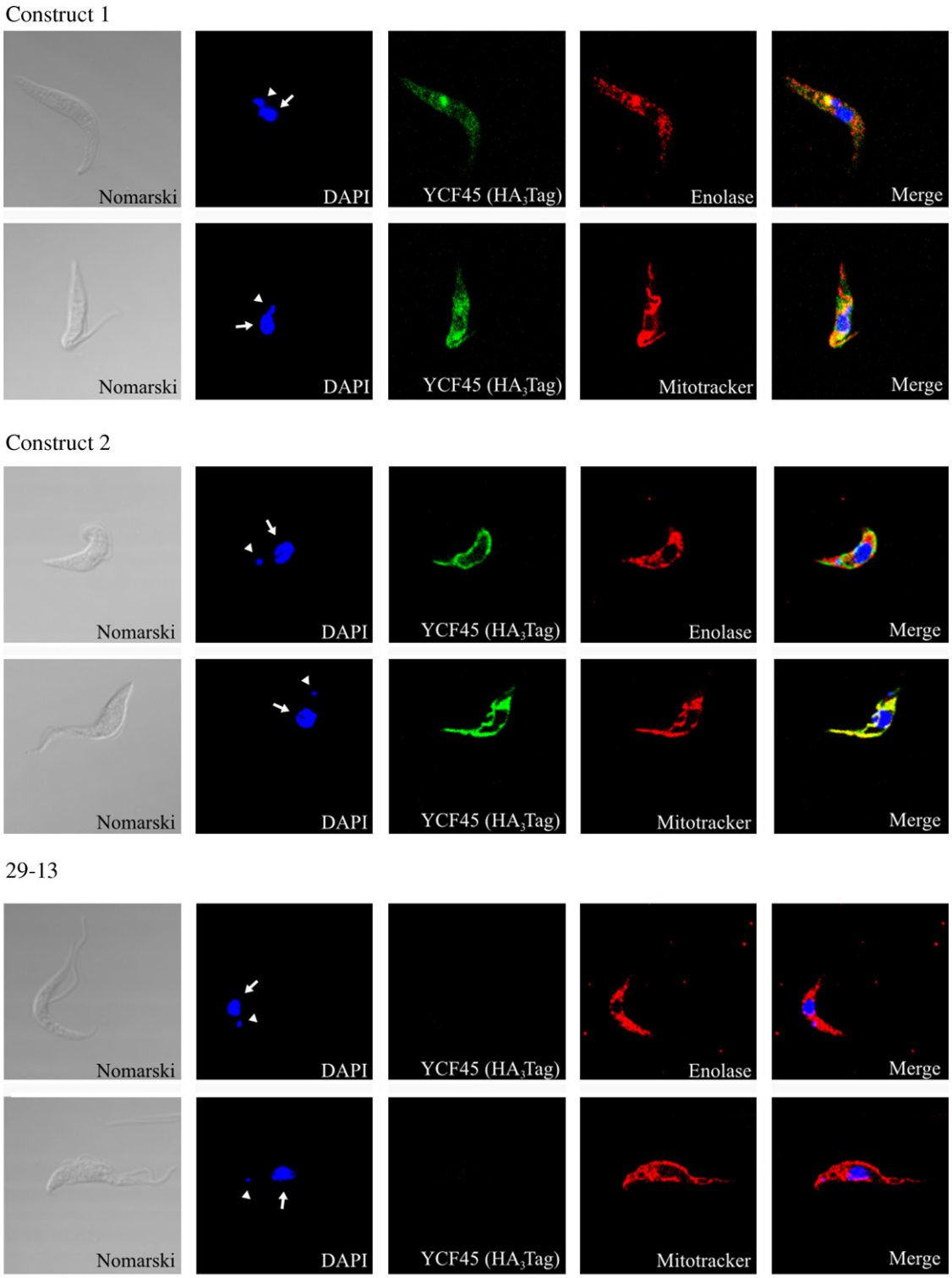

Fig. 3. Immunolocalization of two inducibly expressed HA_3-tagged YCF45 protein forms (constructs 1 and 2) in procyclic *T. brucei*. Mitotracker red and enolase were used as mitochondrial and cytosolic markers, respectively. The signal of construct 1 (upper panel) overlaps with cytosolic enolase, while the signal of construct 2 (middle panel) co-localizes with mitotracker. The parental 29-13 cells (lower panel) lacking HA_3-tagged protein show no signal. Paraformaldehyde-fixed cells permeabilized in methanol were processed as described elsewhere [18], and treated with α-HA_3 tag mouse monoclonal antibody (Sigma) and α-enolase rabbit polyclonal antibody, followed by incubation with goat α-mouse antibody (green) and goat α-rabbit antibody (red). Alternatively, mitotracker staining was performed by incubating living cells in medium containing 100 nM mitotracker red prior to fixation and immunodecoration with α-HA_3 tag mouse monoclonal antibody. Finally, the cells were mounted in Vectashield containing DAPI. Arrowhead and arrow indicate kinetoplast DNA and nuclear DNA, respectively. Cells were imaged using an OLYMPUS FV1000 confocal microscope equipped with the Fluoview v1.7 software. (For interpretation of the references to color in the figure caption, the reader is referred to the web version of the article.)

among the *L. major*, *T. brucei* and other trypanosomatid YCF45 genes (Supplementary Fig. 1).

Same as all other YCF45 genes, the *T. brucei* homologue contains the AAA-type ATPase and cytochrome P450 domains (Fig. 1B). The gene (Tb927.6.3350) is located on chromosome 6 and the protein has a predicted mass of 52.7 kDa. For phylogenetic analysis we have extended the dataset used by Opperdoes and Michels [12] by newly available sequences and additional outgroups. Moreover, we have found putative homologues of YCF45 also in oomycetes and firmicutes, which implies that this protein may not be associated exclusively with the chloroplast. Subsequent analysis showed that the available trypanosomatid YCF45 genes constitute a strongly monophyletic clade that branches off between oomycetes and cyanobacteria (Fig. 1A).

Next, we attempted down-regulation of the YCF45 transcript by RNA interference (RNAi) into the *T. brucei* 29-13 procyclic and 920 bloodstream strains following protocols described elsewhere [18]. Unfortunately, although in the procyclics we were able to detect inducibly transcribed double stranded RNA, in both stages the sensitivity of our Northern analyses was insufficient to detect the YCF45 mRNA, indicating that the transcript is extremely rare. Upon RNAi induction, no growth phenotype was observed in either stage (data not shown).

Since the splice acceptor site of the *T. brucei* YCF45 gene cannot be unambiguously established by *in silico* analysis, we undertook its determination by splice leader (SL) RACE PCR. RNA isolated from the *T. brucei* 29-13 cells was reverse transcribed into cDNA using Superscript RNAse H^- Reverse Transcriptase (Invitrogen) and oligo $(dT)_{12–18}$ (Invitrogen) accordingly to the manufacturer's instructions. Next, using this cDNA as a template for nested PCR with two YCF45-specific primers (Rv1 – GAAAGACCAC AATCCTGAGA; Rv2 – ACAAGACATC GAAGTGAGAG), in combination with primers TbSL1 (AACTAACGCT ATTATTAGAA CAGTTTC) and TbSL2 (TATTATTAGA ACAGTTTCTG TACTATATTG) derived from the SL RNA sequence, an amplicon containing the predicted N-terminus of YCF45 was obtained, proving that the amplification was specific. This allowed us to identify the splice acceptor site, which is located 237 bp upstream from the predicted start codon (Fig. 1B). This result also confirmed the presence of YCF45 mRNA in the procyclic cells, albeit at a very low level. In order to establish the intracellular localization of the YCF45 protein in the absence of specific antibodies, we have used the tagging strategy. The full-length YCF45 gene, with the start and stop codons as annotated in the $Gene^{DB}$, was HA_3-tagged and cloned into the pJH54 vector, which allows inducible expression of the protein (construct 1) (Fig. 1B). The HA_3 tag was attached to the C-terminus in order not to interfere with the putative organellar import signal located at the N-terminus. Subcellular fractions, obtained by digitonin treatment as described previously [19], were immunoprobed with the anti-HA_3 monoclonal antibody. Bulk of the tagged protein was present in the cytosol (Fig. 2) and this result was confirmed by immunolocalization of the tagged protein in fixed procyclic trypanosomes (Fig. 3; upper pannel). Thus, we concluded that YCF45 is a cytosolic protein.

However, a careful comparison with the *T. cruzi* YCF45 sequence followed by re-analysis of the cDNA fragment amplified with the YCF45-specific and SL RNA-derived primers lead to the identification of another putative start methionine that was not considered during the annotation of the *T. brucei* genome. This start codon is located 183 bp upstream from the annotated one (Fig. 1B). Surprisingly, the *in silico* prediction using Mitoprot and TargetP programs attached 99.5% and 97.4% probabilities of import of this protein into the mitochondrion, respectively, which is dramatically different from the originally annotated version of YCF45, for which a cytosolic localization was predicted (data not shown). Therefore, another HA_3-tagged construct was generated in the pJH54 vector, containing the N-terminally extended version of the YCF45 gene (construct 2) (Fig. 1B). Indeed, in this case the tagged protein was targeted into the mitochondrion of procyclic *T. brucei*, as shown by Western analysis of the cytosolic and mitochondrial fractions (Fig. 2). Strict mitochondrial localization was confirmed by immunolocalization, where the HA_3-tagged YCF45 in construct 2 co-localized with mitotracker (Fig. 3; middle panel). We therefore conclude that the start codon of the YCF45 protein is 183 bp upstream from the one hitherto predicted, and that the protein is imported into the mitochondrion of the procyclic trypanosomes.

Mitochondrial localization is certainly unexpected for a protein predicted to be almost invariably confined to the chloroplast in plants which, moreover, was viewed as the most tangible remnant of the lost plastid of trypanosomatids [12]. Our results strongly indicate that the ancestral trypanosomatid obtained YCF45 from an unknown source by lateral gene transfer. Since the mitochondrial and chloroplast import signals share similar sequence motifs and a subset of proteins is known to be dual targeted into both organelles [20,21], the newly acquired protein was targeted to the mitochondrion. It is also possible that this targeting was enabled by the use of an alternative start codon [22].

Besides experimentally correcting *in silico* assigned N-terminus of YCF45, we also show that this (hap-hazardously) acquired protein remained surprisingly conserved in these flagellates and likely fulfils an as yet unidentified, under cultivation conditions non-essential, function in their mitochondrion.

Acknowledgements

We thank Miroslav Oborník for help with phylogenetic analysis and Lucka Novotná for assistance with confocal microscope. This work was supported by the Student Grant Agency to J.T., the Grant Agency of the Czech Republic 204/09/1667, the Ministry of Education of the Czech Republic (2B06129, LC07032 and 6007665801), and the Praemium Academiae award to J.L.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.molbiopara.2010.05.002.

References

[1] Adl SM, Simpson AGB, Farmer MA, et al. The new higher level classification of eukaryotes with emphasis on the taxonomy of protist. J Euk Microbiol 2005;52(5):399–451.
[2] Simpson AGB, Stevens JR, Lukeš J. The evolution and diversity of kinetoplastid flagellates. Trends Parasitol 2006;22:168–74.
[3] Waller RF, McConville MJ, McFadden GI. More plastids in human parasites? Trends Parasitol 2004;20:54–7.
[4] Vickerman K. The diversity of the kinetoplastid flagellates. In: Lumsden WHR, Evans DA, editors. Biology of the Kinetoplastida. London: Academic Press; 1976. p. 1–34.
[5] Martin W, Borst P. Secondary loss of chloroplast in trypanosomes. Proc Natl Acad Sci USA 2003;100:765–7.
[6] Hannaert V, Bringaud F, Opperdoes FR, Michels PAM. Evolution of energy metabolism and its compartmentation in Kinetoplastida. Kinetopl Biol Dis 2003;2:11.
[7] Hannaert V, Saavedra E, Duffieux F, et al. Plant-like traits associated with metabolism of *Trypanosoma* parasites. Proc Natl Acad Sci USA 2003;100:1067–71.
[8] Cavalier-Smith T. Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. J Euk Microbiol 1999;46:347–66.
[9] Kivic PA, Walne PL. An evaluation of a possible phylogenetic relationship between the Euglenophyta and Kinetoplastida. Orig Life 1984;13:269–88.
[10] Leander BS. Did trypanosomatid parasites have photosynthetic ancestors? Trends Microbiol 2004;12:251–8.
[11] El-Sayed NM, Myler PJ, Blandin G, et al. Comparative genomics of trypanosomatid parasitic protozoa. Science 2005;309:404–9.
[12] Opperdoes FR, Michels P. Horizontal gene transfer in trypanosomatids. Trends Parasitol 2007;23:470–6.
[13] Bodyl A, Mackiewicz P. Were class C iron-containing superoxide dismutases of trypanosomatid parasites initially imported into a complex plastid? A hypoth-

esis based on analyses of their N-terminal targeting signals. Parasitology 2008;135:1101–10.
[14] Hallick RB, Bairoch A. Proposals for the naming of chloroplast genes. III. Nomenclature for open reading frames encoded in chloroplast genomes. Plant Mol Biol Rep 1994;12(2):S29–30.
[15] Steane DA. Complete nucleotide sequence of the chloroplast genome from the Tasmania blue gum. *Eucalyptus globulus* (Myrtaceae). DNA Res 2005;12:215–20.
[16] Hackett JD, Yoon HS, Soares MB, et al. Migration of the plastid genome to the nucleus in a peridinin dinoflagellate. Curr Biol 2004;14:213–8.
[17] Ohtsuka M, Oyabu J, Kashino Y, Satoh K, Koike H. Inactivation of *ycf33* results in an altered cyclic electron transport pathway around photosystem I in *Synechocystis* sp. PCC6803. Plant Cell Physiol 2004;45:1243–51.
[18] Hashimi H, Čičová Z, Novotná L, Wen YZ, Lukeš J. Kinetoplastid guide RNA biogenesis is dependent on subunits of the mitochondrial RNA binding complex 1 and mitochondrial RNA polymerase. RNA 2009;15:588–99.
[19] Long S, Jirků M, Mach J, et al. Ancestral roles of eukaryotic frataxin: mitochondrial frataxin function and heterologous expression of hydrogenosomal *Trichomonas* homologs in trypanosomes. Mol Microbiol 2008;69:94–109.
[20] Peeters N, Small I. Dual targeting to mitochondria and chloroplast. Biochim Biophys Acta 2001;1541:54–63.
[21] Carrie C, Giraud E, Whelan J. Protein transport in organelles: dual targeting of proteins to mitochondria and chloroplasts. FEBS J 2009;276:1187–95.
[22] Danpure CJ. How can the products of a single gene be localized to more than one intracellular compartment? Trends Cell Biol 1995;5:230–8.