Characterization of Periplastidal Compartment–Targeting Signals in Chlorarachniophytes

Yoshihisa Hirakawa,[1] Gillian H. Gile,[2] Shuhei Ota,[1] Patrick J. Keeling,[2] and Ken-ichiro Ishida*,[1]

[1]Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Japan

[2]Canadian Institute for Advanced Research and Department of Botany, University of British Columbia, Vancouver, BC, Canada

*Corresponding author: E-mail: ishida@sakura.cc.tsukuba.ac.jp.

Associate editor: Martin Embley

Abstract

Secondary plastids are acquired by the engulfment and retention of eukaryotic algae, which results in an additional surrounding membrane or pair of membranes relative to the more familiar primary plastids of land plants. In most cases, the endocytosed alga loses its eukaryotic genome as it becomes integrated, but in two algal groups, the cryptophytes and chlorarachniophytes, the secondary plastids retain a vestigial nucleus in the periplastidal compartment (PPC), the remnant eukaryotic cytoplasm between the inner and the outer membrane pairs. Many essential housekeeping genes are missing from these reduced genomes, suggesting that they are now encoded in the host nucleus and their products are targeted to the PPC. One such nucleus-encoded, PPC-targeted protein, the translation elongation factor like (EFL) was recently identified in chlorarachniophytes. It bears an N-terminal–targeting sequence comprising a signal peptide and a transit peptide–like sequence (TPL) similar to the plastid-targeted proteins of chlorarachniophytes as well as a hydrophilic C-terminal extension rich in lysine and aspartic acid. Here, we characterize the function of the N- and C-terminal extensions of PPC-targeted EFL in transformed chlorarachniophyte cells. Using green fluorescent protein as a reporter molecule, we demonstrate that several negatively charged amino acids within the TPL are essential for accurate targeting to the PPC. Our findings further reveal that the C-terminal extension functions as a PPC retention signal in combination with an N-terminal plastid-targeting peptide, which suggests that plastid and PPC proteins may be sorted in the PPC.

Key words: chlorarachniophyte, nucleomorph, periplastidal compartment, plastid targeting, secondary endosymbiosis, transit peptide.

Introduction

Plants and algae acquired their diverse plastids by multiple endosymbiotic events (for reviews, see McFadden 2001; Keeling et al. 2004). The plastids of land plants, green algae, red algae, and glaucophytes are surrounded by two membranes and are all derived from a single primary endosymbiosis with a cyanobacterium (Moreira et al. 2000; Rodríguez-Ezpeleta et al. 2005). In other algal groups, plastids are surrounded by three or four membranes and are descended from eukaryotic endosymbionts (Cavalier-Smith 2000; Keeling 2004). The plastids of cryptophytes, haptophytes, apicomplexans, heterokonts, and most photosynthetic dinoflagellates are derived from a single or multiple secondary endosymbioses between red algae and eukaryotes; how many endosymbioses have occurred is less clear (Keeling 2009), whereas the plastids of chlorarachniophytes and euglenophytes are derived from endosymbionts of two different green algae (Ishida et al. 1997; Rogers et al. 2007).

As a photosynthetic endosymbiont becomes a plastid, many of its genes are lost or transferred to the nuclear genome of the host (Martin et al. 1998; Bock and Timmis 2008). The products of transferred genes that are still required by the plastids must therefore be targeted back to them across multiple envelope membranes (van Dooren et al. 2001; Ishida 2005; Nassoury and Morse 2005). In the case of primary plastids, nucleus-encoded plastid proteins typically carry a plastid-targeting signal called a transit peptide (TP) as an N-terminal extension (Bruce 2001; Steiner et al. 2005). These precursor proteins (preproteins) are posttranslationally transported across two membranes into the plastids, and this process is mediated by molecular machines in the outer and inner envelope membranes, referred to as translocons at the outer/inner envelope membrane of chloroplasts (TOC/TIC; Jarvis and Robinson 2004). In organisms with secondary plastids, plastid-targeted preproteins typically have a bipartite-targeting sequence consisting of a signal peptide (SP) followed by a transit peptide–like sequence (TPL) at their N-termini (Patron and Waller 2007). The SP is a hydrophobic sequence that targets proteins cotranslationally to the endoplasmic reticulum (ER), which is in some cases continuous with the outermost membrane of the plastid (Bolte et al. 2009). The TPL is necessary for targeting proteins across the remaining two or three membranes into the plastid stroma (Patron and Waller 2007; Bolte et al. 2009). TPLs and TPs are highly variable in sequence, but they typically carry an overall positive charge.

Although most secondary plastids have completely lost their eukaryotic genomes, the secondary plastids of cryptophytes and chlorarachniophytes retain vestigial nuclei called nucleomorphs. Each of these plastids is surrounded

by four membranes, and the nucleomorph is located between the inner and the outer pairs of plastid membranes in the periplastidal compartment (PPC), which is the remnant cytoplasm of the endosymbiont (Archibald 2007). Nucleomorph genomes in both cryptophytes and chlorarachniophytes encode predominantly housekeeping genes along with a handful of plastid-targeted proteins (Douglas et al. 2001; Gilson et al. 2006; Lane et al. 2007), but many essential genes for DNA replication and protein synthesis are missing. These housekeeping genes are therefore expected to have transferred to the host nuclear genome and acquired targeting information in order to be sent to the PPC. The secondary plastids of cryptophytes and chlorarachniophytes therefore consist of two major compartments, each of which requires a distinct set of proteins encoded by the host nucleus.

In cryptophytes, several nucleus-encoded PPC proteins have been identified, and these preproteins have an N-terminal bipartite-targeting sequence consisting of an SP and a TPL similar to those of plastid-targeted preproteins (Gould, Sommer, Hadfi, et al. 2006). The characteristic by which the cryptophyte cell can distinguish plastid- from PPC-targeted preproteins is believed to be the presence or absence of an aromatic amino acid (most often phenylalanine) at the +1 position of the TPL. Generally, presence of the aromatic residue results in plastid stromal localization, whereas absence of this residue leads to proteins residing in the PPC (Gould, Sommer, Kroth, et al. 2006). A similar characteristic is also found in diatoms (Kilian and Kroth 2005), haptophytes (Patron et al. 2006), dinoflagellates (Patron et al. 2005), and certain apicomplexans (Ralph et al. 2004), and it is considered to be a shared ancestral feature derived from the red algal endosymbiont of these lineages (Patron and Waller 2007; Sommer et al. 2007).

In chlorarachniophytes, a putative nucleus-encoded, PPC-targeted protein, the translation elongation factor like (EFL) has been characterized from several species. These proteins also carry an N-terminal bipartite-targeting sequence, consisting of an SP and a TPL, similar to those of plastid-targeted preproteins (Rogers et al. 2004; Gile and Keeling 2008). In addition, PPC-targeted EFL has a novel hydrophilic C-terminal extension consisting of mainly lysine and aspartic acid residues (Gile and Keeling 2008). As with cryptophytes, the chlorarachniophyte cell must be able to distinguish plastid proteins from PPC proteins, but this mechanism is not yet understood.

In order to determine which aspects of the targeting peptides are most important for PPC localization, we have investigated the localization pattern of green fluorescent protein (GFP) fused to the SP and TPL of PPC-targeted EFL preprotein in transformed chlorarachniophyte cells. We show that targeting information from PPC-targeted EFL results in GFP localizing to a structure adjacent to the chlorophyll autofluorescence, and we confirm by immunoelectron microscopy that this is the PPC. By replacing the negatively charged amino acids of the TPL with neutral and positively charged residues, we determine that negatively charged amino acids within the TPL are essential for PPC targeting. In addition, we reveal that the C-terminal extension of the EFL preprotein serves as a novel PPC retention signal. These features together allow the chlorarachniophyte cell to accurately sort PPC- and plastid-targeted proteins.

Materials and Methods

Plasmid Constructions

In order to construct pBnEFL97 + GFP and pGsEFL120 + GFP, complementary DNA fragments encoding the N-terminal bipartite-targeting sequences of *Bigelowiella natans* EFL gene (EU810324) and *Gymnochlora stellata* EFL gene (EU810329) were amplified by polymerase chain reaction (PCR), and each fragment was inserted between *Hind*III and *Nco*I sites of the pLaRGfp + mc vector (Hirakawa et al. 2009). In order to introduce amino acid substitutions into the TPL of BnEFL preprotein, a PCR-based site-directed mutagenesis technique (Higuchi et al. 1988) was used. Each of the final PCR products was subsequently inserted to the pLaRGfp + mc vector to generate pBnEFL-E50A-E51A-E66A + GFP, pBnEFL-E50R-E51R-E66R + GFP, and pBnEFL-E50D-E51D-E66D + GFP. In order to insert the C-terminal D/K domain sequence of BnEFL at the C-terminus of GFP, we used a splicing by overlapped extension by PCR (SOE by PCR) technique (Horton et al. 1989). The fragments encoding the D/K domain, GFP, and BnAtpD62 + GFP were amplified, and each two fragments were recombined using the SOE by PCR technique. Each of the resulting fragments was inserted into the pLaRGfp + mc vector replacing the *gfp* gene to generate pGFP + D/K and pBnAtpD62 + GFP + D/K. Every primer set used for PCR are listed in supplementary table S1, Supplementary Material online. All plasmids were cloned in the DH5α strain of *Escherichia coli* and purified using a QIAprep Spin Miniprep Kit (QIAGEN, Tokyo, Japan). These constructs were subsequently sequenced to ensure correct construction.

Transient Transformation of the Chlorarachniophyte Cells by Microparticle Bombardment

To prepare for transformation, *Lotharella amoebiformis* (CCMP2058; Ishida et al. 2000) was grown at 20 °C under white illumination (80–100 µmol photons·m^{-2}·s^{-1}) on a 12-h light/12-h dark cycle in 500 ml Erlenmeyer flasks containing 300 ml of Erd-Schreiber Modified (ESM) medium (Kasai et al. 2009). We used a Biolistic PDS-1000/He Particle Delivery System (Bio-Rad, Hercules, CA) for the transformation of *L. amoebiformis* cells. The cells were then bombarded in the best condition as described previously (Hirakawa et al. 2008). After the bombardment, the cells were immediately transferred into new plastic plates with 10 ml of fresh ESM medium and incubated under the conditions described above.

Observation of GFP Fluorescence

After bombardment of 24–48 h, transiently transformed GFP-expressing cells were detected and isolated under

a Leica DMR fluorescent microscope (Leica, Wetzlar, Germany) with L5 filter set (excitation filter was BP 480/40 nm and suppression filter was BP 527/30 nm). Confocal imaging was performed using an inverted Zeiss LSM 510 laser scanning microscope (Carl Zeiss AG, Jena, Germany). GFP fluorescence and plastid autofluorescence were detected with a 505- to 530-nm band pass filter and a 585-nm long pass filter, respectively, in the excitation line of a 488-nm argon laser and 543-nm He/Ne laser using single-track mode.

Immunoelectron Microscopy

After bombardment of 24 h, transiently transformed *L. amoebiformis* cells were fixed for 2 h at 4 °C in 3% paraformaldehyde/0.5% glutaraldehyde/0.25 M sucrose in PHEM buffer (60 mM PIPES/25 mM HEPES/10 mM EGTA/2 mM $MgCl_2$, pH 7.4). An isolated transformant exhibiting GFP fluorescence was micropipetted onto a poly-L-lysine coated coverslip (18 × 18 mm); it was then dehydrated for 5 min in each increment of the graded ethanol series (20%, 40%, 60%, and 80%), followed by infiltration with LR White:ethanol gradients of 1:2 for 1 h, 1:1 for 1 h, and 2:1 for 1 h, and 100% for 12 h. All dehydration and infiltration steps were performed at 4 °C. The coverslip was placed on gelatin capsules filled with LR White, and it was polymerized at 58 °C for 24 h. The polymerized block was removed from the coverslip and sectioned on a Reichert Ultracut S ultramicrotome (Leica) using a diamond knife. Gold sections were collected onto Formvar-coated copper mesh or one-slot grids. Before immunogold labeling, sections on the grids were blocked with a blocking solution (5% normal goat serum/2.5% skim milk/0.1% NaN_3 in phosphate-buffered saline [PBS]) for 1 h at room temperature. The grids were then incubated in 25 µl of anti-GFP primary antibody (JL-8; Clontech, Mountain View, CA), diluted 1:25 with PBS for 2 h at 30 °C. The grids were washed 20 times with PSB supplemented with 0.05% Tween-20 for 5 min on drops of PBS. The rinsed grids were then incubated on 30 µl anti-mouse immunoglobulin G secondary antibody (Sigma, St Louis, MO) conjugated with 10 nm gold particles (diluted 1:20 with PBS) for 1 h at 30 °C. The labeled grids were rinsed with PBS and Milli-Q water, followed by staining with uranyl acetate for 10 min; the ultrathin sections were then observed under a JEM-1010 transmission electron microscope (JEOL, Tokyo, Japan) at 80 kV.

TPL Charge Prediction

In order to provide a basis of comparison between the TPLs of PPC- and plastid-targeted preproteins, the overall charge of each sequence was estimated using the Peptide Property Calculator (Innovagen, Lund, Sweden). Assumptions and the equation used can be viewed at http://www.innovagen.se/custom-peptide-synthesis/peptide-property-calculator/peptide-property-calculator-notes.asp. Each TPL region of N-terminal bipartite-targeting sequences in 31 plastid-targeted preproteins (*B. natans* expressed sequence tag [EST] data from Rogers et al. 2004) was predicted by the neural networks method of SP prediction server SignalP (http://www.cbs.dtu.dk/services/

SignalP/; Nielsen et al. 1997) and the chloroplast TP prediction server ChloroP (http://www.cbs.dtu.dk/services/ChloroP/; Emanuelsson et al. 1999). We also estimated the TPL cleavage site of each preprotein using an alignment of mature regions in homologous proteins of other organisms (land plants, green algae, and cyanobacteria). In 19 preproteins, different cleavage sites were estimated by these two predictions: ChloroP and alignment prediction (supplementary fig. S1, Supplementary Material online). For the net charge calculation, we used TPL sequences predicted by the alignment predictions.

Results and Discussion

N-Terminal Bipartite-Targeting Signals Are Necessary and Sufficient for PPC Targeting in Chlorarachniophytes

The cellular location of putatively PPC-targeted EFL in chlorarachniophytes was previously inferred on the basis of sequence and phylogenetic evidence and suggestive evidence from immunolocalizaion that excluded a cytosolic location (Gile and Keeling 2008). In order to confirm the PPC localization, we constructed two plasmids that express GFP fused with the N-terminal bipartite-targeting sequence of EFL preproteins from *B. natans* and *G. stellata* (BnEFL97 + GFP and GsEFL120 + GFP). Live cells of *L. amoebiformis* were then transformed with each of these plasmid constructs. In both cases, GFP fluorescence appeared in a distinct pattern adjacent to, but not overlapping with, chlorophyll autofluorescence of plastids (fig. 1A and B). Fluorescence was restricted to a small spot near the base of the bulbous pyrenoid where the two plastid lobes meet, where the nucleomorph is known to reside (Ishida et al. 2000), which is consistent with a PPC localization. This fluorescence pattern is similar to the blob-like structures reported from diatoms expressing GFP fused to PPC-targeting peptides (Gould, Sommer, Kroth, et al. 2006). In order to observe this localization more precisely, we also performed an immunogold localization using an anti-GFP antibody on a cell transformed with pBnEFL97 + GFP. Most gold particles were accumulated in a space between the second and the third outermost plastid envelope membranes where the nucleomorph resides, confirming that the location of observed GFP fluorescence is the PPC (fig. 1C and D). These observations indicate that the N-terminal bipartite-targeting sequences of EFL preproteins are sufficient for delivering proteins to the PPC.

Negatively Charged Amino Acids within the TPL Are Significant for the Localization of PPC-Targeted EFL

In a previous study, we found that the plastid-targeting peptide of the ATP synthase delta subunit (AtpD) can direct proteins to the PPC if as few as three of its five positively charged amino acids are substituted with neutral residues (Hirakawa et al. 2009). This observation implies that the net charge of the TPL might play a role in
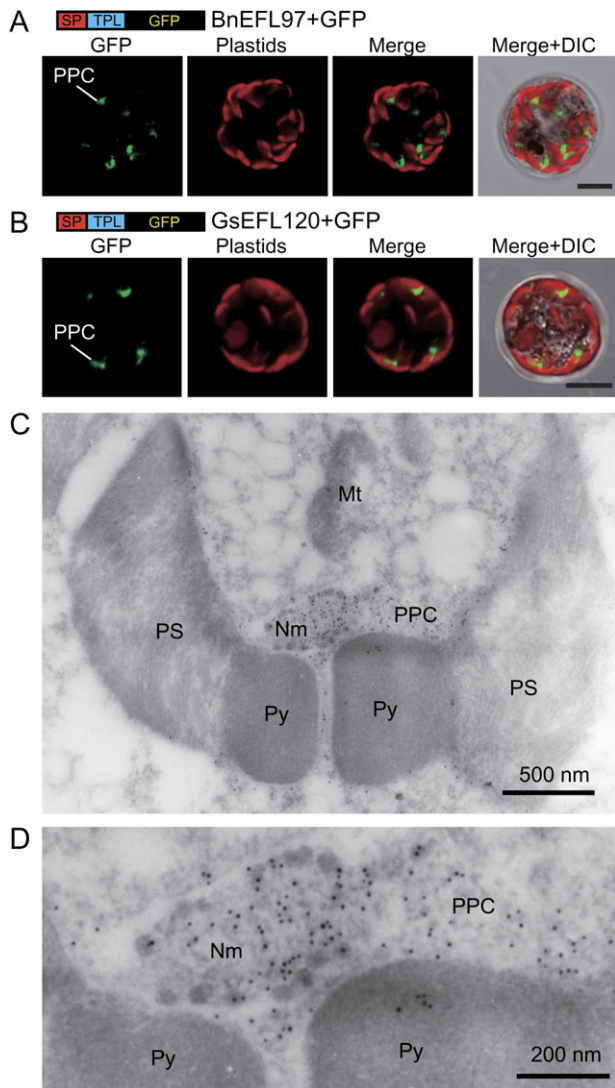
FIG. 1. Localization of GFP fused with putative PPC-targeting peptides in *Lotharella amoebiformis* cells. Confocal images labeled "GFP" and "plastids" show GFP localization (green) and chlorophyll autofluorescence (red), respectively. (*A*) Localization of GFP fused with the N-terminal bipartite-targeting sequence of BnEFL (BnEFL97 + GFP) in the putative PPC; scale bar: 5 μm. (*B*) Localization of GFP fused with the N-terminal bipartite-targeting signal of GsEFL (GsEFL120 + GFP) in the putative PPC. (*C*) Immunogold localization of GFP in *L. amoebiformis* transformed with pBnEFL97 + GFP. Gold particles were accumulated between inner and outer pairs of plastid membranes, near the nucleomorph (Nm), but not in the PS or the Py, indicating PPC localization. (*D*) Enlarged region of fig. 1C showing the PPC with nucleomorph and part of the bilobed Py. Mt, mitochondrion; PS, plastid stroma; Py, pyrenoid.

the discrimination of plastid- from PPC-targeted proteins in chlorarachniophyte cells. Further supporting this idea, each TPL of PPC-targeted EFL preproteins (BnEFL and GsEFL) has several negatively charged residues in addition to its positively charged residues, resulting in near neutrality at pH 7.0 (the calculated net charges are −0.8 and −0.9, respectively), whereas the plastid-targeting TPLs characterized from 31 plastid-targeted preproteins of

B. natans have a pronounced net positive charge of +5.6 ± 2.6 (supplementary fig. S2, Supplementary Material online). Therefore, we have undertaken a substitution analysis of these negatively charged residues in the TPL from *B. natans* PPC-targeted EFL. The TPL of BnEFL has three glutamic acid residues (*E*) at positions 50, 51, and 66 from the N-terminus of this preprotein (fig. 2*A*). We substituted these three amino acids in the BnEFL97 + GFP construct with neutral alanine residues (*A*), positively charged arginine residues (*R*), or negatively charged aspartic acid residues (*D*) to generate BnEFL-E50A-E51A-E66A + GFP (calculated net charge of TPL is +2.2 at pH 7.0), BnEFL-E50R-E51R-E66R + GFP (+5.2), and BnEFL-E50D-E51D-E66D + GFP (−0.8) and observed their localization in *L. amoebiformis* cells. For the two plasmid constructs in which glutamic acid residues were substituted by alanine or arginine residues, transformed cells exhibited GFP fluorescence mainly in the plastid stroma, though some fluorescence was also observed in the PPC (fig. 2*B* and *C*). On the other hand, GFP fluorescence was observed in the PPC when cells were transformed with pBnEFL-E50D-E51D-E66D + GFP (fig. 2*D*). These observations are consistent with our previous findings that a minimum number of positively charged residues is necessary for targeting the plastid stroma, although the detectable PPC fluorescence suggests that another factor or factors may also contribute to stromal localization. Our results support the idea that an overall positive charge may be the means by which chlorarachniophyte cells distinguish between plastid and PPC preproteins.
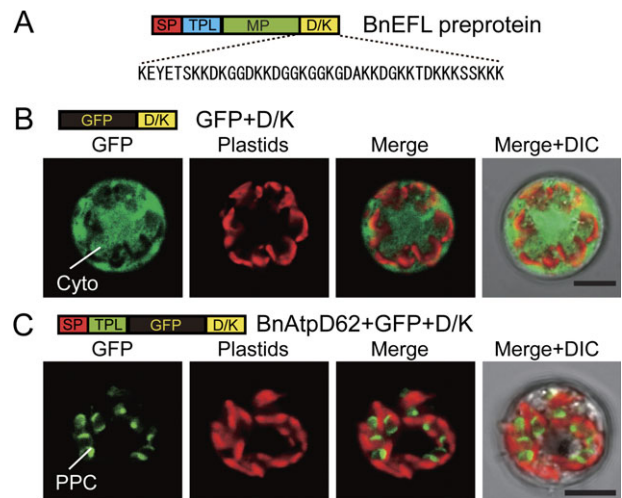
In the red algal–derived plastids of cryptophytes, plastid- and PPC-targeted proteins also carry N-terminal bipartite-targeting sequences consisting of an SP and a TPL. However, both classes of TPL carry an overall positive charge, and they are instead distinguished from one another by the presence or absence of an aromatic residue (typically phenylalanine) at the +1 position of the TPL (Gould, Sommer, Hadfi, et al. 2006; Gould, Sommer, Kroth 2006). A phenylalanine residue at or near the N-terminus of the TP is also found in glaucophytes, red algae, and other algae with secondary plastids derived from a red alga (Patron and Waller 2007). In diatoms, it has been demonstrated that the phenylalanine residue in TPLs is exchangeable with other aromatic residues (tryptophan and tyrosine) or leucine residue (Gruber et al. 2007). When the phenylalanine residue is replaced with a non-aromatic, nonleucine residue, plastid-targeting peptides have been observed to direct proteins to the PPC of diatom (Gruber et al. 2007) and vice versa similar to the reciprocal mistargeting of charge altered TPLs that we have observed in chlorarachniophytes. Because this important aromatic amino acid is not a characteristic of green algae, we would expect that the ancestor of chlorarachniophytes must have come up with a different mechanism to distinguish between plastid- and PPC-targeted proteins, and now it is apparent that at least part of this mechanism is the overall charge of the TPL.
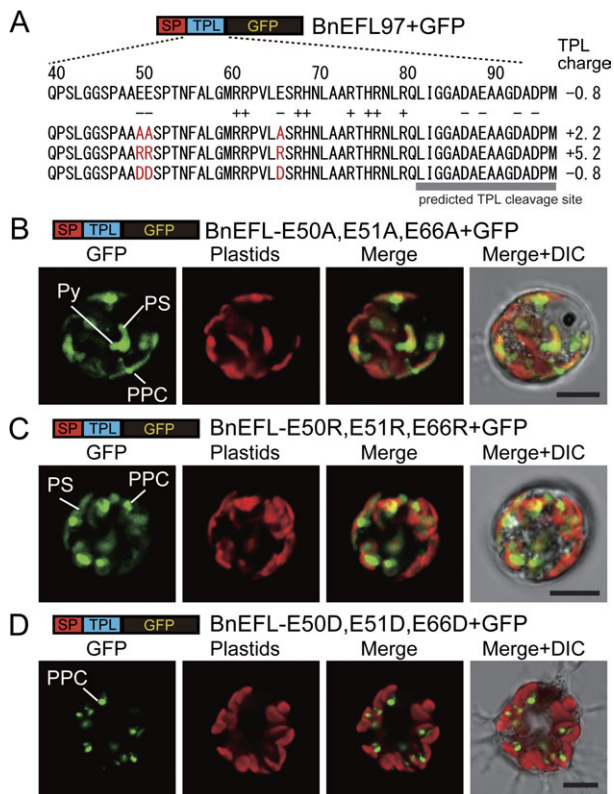
Fig. 2. In vivo targeting of GFP fusion proteins having substitutions at three positively charged amino acids in the TPL of BnEFL preprotein. (A) Amino acid sequences in the TPLs of BnEFL + GFP, BnEFL-E50A-E51A-E66A + GFP, BnEFL-E50R-E51R-E66R + GFP, and BnEFL-E50D-E51D-E66D + GFP. Numbers above the sequence indicate amino acid positions from the N-terminus of this protein, and signs, + and −, indicate positively and negatively charged amino acids, respectively. The numbers of right side of sequences show calculated TPL net charge at pH 7.0. (B) Localization of BnEFL-E50A-E51A-E66A + GFP in PS, Py, and some PPC. (C) Localization of BnEFL-E50R-E51R-E66R + GFP in PS and some PPC. (D) Localization of BnEFL-E50R-E51R-E66R + GFP in the PPC; scale bar: 5 μm. PS, plastid stroma; Py, pyrenoid.

Fig. 3. In vivo targeting of GFP fusion proteins having the C-terminal D/K domain of BnEFL preprotein. (A) Amino acid sequence of the C-terminal D/K domain abounding in lysine (K) and aspartic acid (D) residues. (B) Localization of GFP fused with the D/K domain at the C-terminus (GFP + DK) in the cytoplasm (Cyto). (C) Localization of plastid-targeted GFP, BnAtpD62 + GFP, fused with the D/K domain (BnAtpD62 + GFP + D/K). Fluorescence appears in the PPC; scale bar: 5 μm.

Novel C-Terminal Retention Signal of the PPC-Targeted EFL

In addition to the N-terminal bipartite-targeting peptide, the PPC-targeted EFL preproteins of chlorarachniophytes carry a curious C-terminal extension that is not present in cytosolic EFL proteins (fig. 3A). This hydrophilic stretch of approximately 30 residues is rich in lysine (K) and aspartic acid (D). We constructed plasmids encoding two different GFP fusion proteins carrying the C-terminal D/K domain of B. natans to test whether this sequence has any function in PPC targeting. First, we transformed L. amoebiformis cells with a plasmid-encoding GFP fused with the D/K domain at its C-terminus (GFP + D/K). We observed GFP fluorescence in the cytoplasm of the transformed cells, but not in the plastids or the PPC (fig. 3B). The localization of this GFP fusion protein was not different from that of only GFP. Next, we transformed cells with a construct including the bipartite-targeting sequence

of B. natans AtpD preprotein, a known plastid-targeting peptide (Hirakawa et al. 2009) at the N-terminus of GFP + D/K (BnAtpD62 + GFP + DK). Interestingly, in the cells transformed with pBnAtpD62 + GFP + D/K, GFP fluorescence was observed in the PPC, not in the plastid stroma (fig. 3C). These findings indicate that the C-terminal D/K domain has no ability to target GFP into the PPC or plastid stroma but suggest that it is able to retain plastid-targeted GFP in the PPC. This is the first time that a C-terminal extension has been implicated in any aspect of plastid targeting, and it suggests not only that the distinction between plastid- and PPC-targeted proteins can involve both the TPL and the C-terminal extension but also that sorting of plastid from PPC preproteins may occur in the PPC.

Currently, no other PPC-targeted protein is known to carry a C-terminal D/K-rich peptide. Only one other putative PPC-targeted protein has been identified in chlorarachniophytes to date, the eukaryotic translation initiation factor 1 (eIF1) from G. stellata, and this protein lacks a C-terminal extension (Gile and Keeling 2008). Notably, the TPL of eIF1 has three negatively charged and no positively charged residues, conferring an estimated net charge of −2 at pH 7.0 (supplementary fig. S2, Supplementary Material online). This raises the possibility that targeting is a sum of various components, so that a TPL with a negative charge, such as that of eIF1, can be adequately distinguished from plastid-targeted proteins but that a near-neutral TPL, such as that of EFL, requires a C-terminal extension in order to be retained in the PPC with a high efficiency. This would suggest that GFP constructs with the N-terminal leaders from EFL preproteins should be targeted to the plastid at some level, which was not observed, but whether

a low level of leakage could be detected is not known. Overall, the C-terminal extension is a unique and interesting feature of this targeting pathway, but at present, it remains to be seen whether it is a common distinguishing factor of PPC-targeted proteins or a unique feature of PPC-targeted EFL.

How Does Sorting of PPC- from Plastid-Targeted Proteins Occur?

Here and in a previous study (Hirakawa et al. 2009), we have shown that substitution of positively charged residues in a plastid-targeting TPL leads to a PPC localization and alteration of negatively charged residues in a PPC-targeting TPL leads to a plastid stromal localization. In accordance with this observation, the TPLs of plastid- and PPC-targeted preproteins are highly similar in amino acid composition and hydropathy as well as a motif sequence previously shown to be important for plastid targeting (Gile and

Keeling 2008; Hirakawa et al. 2009). Together these observations suggest that plastid- and PPC-targeting TPLs are recognized and transported to the plastid by the same system. How are the plastid- and PPC-targeted preproteins sorted after reaching the plastids? In addition to the PPC retention activity, we have shown that for the C-terminal extension, the positive charge requirement for TPLs to pass through the inner pair of plastid membranes suggests that the sorting of these preproteins might take place in the PPC. In land plants, positively charged residues in TPs are thought to be essential for recognition by the import receptor, TOC 159, which carries a negatively charged domain on the plastid outer membrane (Jarvis and Robinson 2004). A homologous import receptor might be expected to reside in the third outermost membrane of chlorarachniophytes (equivalent to the outer plastid membrane of plants) where it would be capable of discriminating plastid preproteins from PPC preproteins on the basis
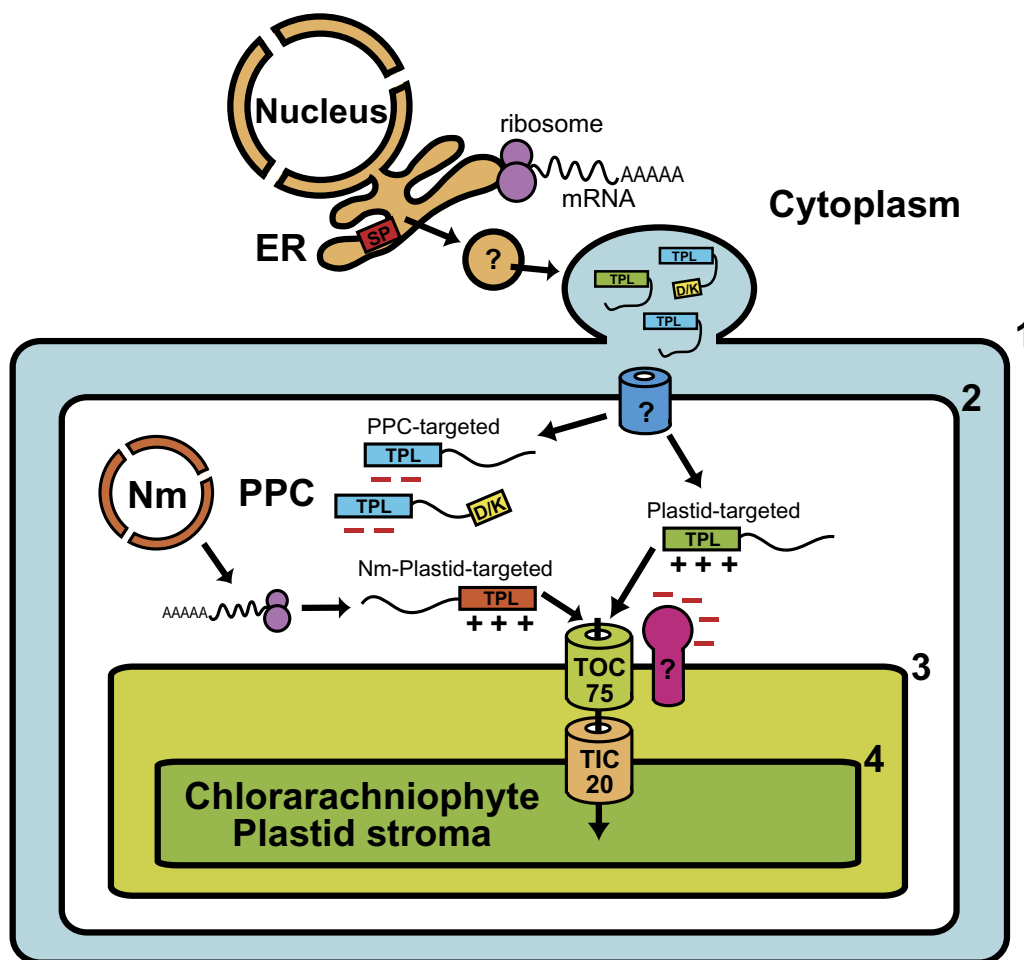
FIG. 4. A model for plastid- and PPC targeting in chlorarachniophytes. Nucleus-encoded preproteins having an N-terminal bipartite-targeting sequence are first cotranslationally transported into the ER concomitant with the removal of the SP in the ER lumen. Subsequently, these preproteins are transported by vesicles to the second outermost membrane (periplastid membrane) of the plastid. Once in the PPC, preproteins are sorted according to the charge of their TPL. Positively charged TPLs of plastid-targeted preproteins are recognized by a yet unidentified import receptor and transported into the plastid stroma possibly via homologs of TOC75 and TIC20 that are encoded in the nucleomorph (Nm). Neutral and negatively charged TPLs of PPC-targeted preproteins are not recognized by the import receptor and accumulate in the PPC. The C-terminal D/K domain may serve to improve fidelity of PPC localization. Nucleomorph-encoded, plastid-targeted preproteins have a positively charged TPL and may enter the stroma via TOC75 and TIC20 like nucleus-encoded, plastid-targeted preproteins.

of TPL charge. In this scenario, both plastid- and PPC-targeted preproteins are first imported into the PPC, where they are distinguished by their ability or inability to cross the third outermost membrane. No homolog of TOC 159 has yet been identified in chlorarachniophytes, however, though homologs of two other plastid translocon components, TOC75 and TIC20, are encoded by the nucleomorph genome of *B. natans* (Gilson et al. 2006). In heterokonts, cryptophytes, and apicomplexans, which harbor red algal–derived plastids, on the other hand, the distinguishing factor between plastid and PPC preproteins is the phenylalanine (*F*) residue at the +1 position of the TPL (Gould, Sommer, Kroth, et al. 2006, Gruber et al. 2007; Patron and Waller 2007). Recently, a putative second outermost membrane (periplastid membrane) translocon in diatoms, a PPC-specific Der1 protein complex, has been shown to interact with only PPC-targeting TPLs that lack the first phenylalanine residue (Hempel et al. 2009). This study suggests that the PPC preproteins are withheld by the Der1 complex from further transport across the inner pair of plastid membranes and that the Der1 complex is involved in discriminating the PPC- from plastid-targeted preproteins right after these preproteins pass through the periplastid membrane. In chlorarachniophytes, no such protein complex has been found, but it is also conceivable that negatively charged residues of PPC-targeting TPLs or the C-terminal D/K domain interact with an unknown protein/protein complex that is present in the periplastid membrane or PPC in chlorarachniophytes and that this interaction impedes the further transport of PPC preproteins across the inner pair of plastid membranes.

Emerging Model of Plastid and PPC Targeting in Chlorarachniophytes

Here, we have shown that the N- and C-terminal extensions of PPC-targeted EFL preproteins in chlorarachniophytes function as a PPC-targeting signal and a PPC retention signal, respectively, in the chlorarachniophytes. The N-terminal extension consists of a SP, for targeting to the ER, and a TPL that is similar to those of plastid-targeted preproteins but without a net positive charge that is necessary and sufficient for plastid stroma localization. The C-terminal extension is a hydrophilic sequence mainly consisting of lysine (*K*) and aspartic acid (*D*) residues, and it has the ability to prevent plastid-targeted proteins from reaching the stroma. Based on these and previous findings (Hirakawa et al. 2009), we propose a model for plastid- and PPC targeting in chlorarachniophytes (fig. 4).

Nucleus-encoded preproteins with an N-terminal bipartite-targeting sequence are first targeted to the ER by their SPs, where they are cotranslationally transported into the ER lumen. After cleavage of the SPs, the preproteins, now with monopartite-targeting peptides, are delivered to the plastid. The similarity of plastid- and PPC-targeting TPLs and their ability to target each other's compartment when altered in charge suggest that a single mechanism could be able to target both classes of preprotein from the ER to the plastid. This stage of transport

likely occurs in vesicles because the plastid outermost membrane lacks ribosomes and is not continuous with the ER (Hibberd and Norris 1984; Ishida et al. 2000). On arriving at the plastid, both PPC and plastid preproteins are transported into the PPC, where they are sorted according to their ability or inability to cross the third outermost membrane. Thus, preproteins with a neutral or negatively charged TPL, and/or a C-terminal D/K-rich extension, remain in the PPC, whereas plastid preproteins, with a positively charged TPL, are able to travel into the plastid stroma. Finally, a few plastid proteins are encoded in the nucleomorph genome (Gilson et al. 2006), and they are targeted to the stroma by a positively charged N-terminal TPL (Gile and Keeling 2008). Whether these proteins use the same translocons as nucleus-encoded plastid stromal proteins remains unknown, but so far, only one putative TOC and one TIC protein have been identified. Further work is needed to clarify this issue and what proteins affect the sorting and import of PPC and plastid proteins, but this will be aided by the upcoming and eagerly awaited nuclear genome sequence of the chlorarachniophyte *B. natans*.

Supplementary Material

Acknowledgments

References

Archibald JM. 2007. Nucleomorph genomes: structure, function, origin and evolution. *BioEssays* 29:392–402.

Bock R, Timmis JN. 2008. Reconstructing evolution: gene transfer from plastids to the nucleus. *BioEssays* 30:556–566.

Bolte K, Bullmann L, Hempel F, Bozarth A, Zauner S, Maier UG. 2009. Protein targeting into secondary plastids. *J Eukaryot Microbiol.* 56:9–15.

Bruce BD. 2001. The paradox of plastid transit peptide: conservation of function despite divergence in primary structure. *Biochim Biophys Acta.* 1541:2–21.

Cavalier-Smith T. 2000. Membrane heredity and early chloroplast evolution. *Trends Plant Sci.* 5:174–182.

Douglas S, Zauner S, Fraunholz M, Beaton M, Penny S, Deng LT, Wu X, Reith M, Cavalier-smith T, Maier UG. 2001. The highly reduced genome of an enslaved algal nucleus. *Nature* 410: 1091–1096.

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.

Gile GH, Keeling PJ. 2008. Nucleus-encoded periplastid-targeted EFL in chlorarachniophytes. *Mol Biol Evol.* 25:1967–1977.

Gilson PR, Su V, Slamovits CH, Reith ME, Keeling PJ, McFadden GI. 2006. Complete nucleotide sequence of the chlorarachniophyte nucleomorph: nature's smallest nucleus. *Proc Natl Acad Sci U S A.* 103:9566–9571.

Gould SB, Sommer MS, Hadfi K, Zauner S, Kroth PG, Maier UG. 2006. Protein targeting into the complex plastid of cryptophytes. *J Mol Evol.* 62:674–681.

Gould SB, Sommer MS, Kroth PG, Gile GH, Keeling PJ, Maier UG. 2006. Nucleus-to-nucleus gene transfer and protein retargeting into a remnant cytoplasm of cryptophytes and diatoms. *Mol Biol Evol.* 23:2413–2422.

Gruber A, Vugrinec S, Hempel F, Gould SB, Maier UG, Kroth PG. 2007. Protein targeting into complex diatom plastids: functional characterisation of a specific targeting motif. *Plant Mol Biol.* 64:519–530.

Hempel F, Bullmann L, Lau J, Zauner S, Maier UG. 2009. ERAD-derived preprotein transport across the second outermost plastid membrane of diatoms. *Mol Biol Evol.* 26:1781–1790.

Hibberd DJ, Norris RE. 1984. Cytology and ultrastructure of *Chlorarachnion reptans* (Chlorarachniophyta divisio nova, Chlorarachniophyceae classis nova). *J Phycol.* 20:310–330.

Higuchi R, Krummel B, Saiki RK. 1988. A general method of *in vitro* preparation and specific mutagenesis of DNA fragments: study of protein and DNA interaction. *Nucleic Acids Res.* 16:7351–7367.

Hirakawa Y, Kofuji R, Ishida K. 2008. Transient transformation of a chlorarachniophyte alga, *Lotharella amoebiformis* (Chlorarachniophyceae), with *uid*A and *egfp* reporter genes. *J Phycol.* 44:814–820.

Hirakawa Y, Nagamune K, Ishida K. 2009. Protein targeting into secondary plastids of chlorarachniophytes. *Proc Natl Acad Sci U S A.* 106:12820–12825.

Horton RM, Hunt HD, Ho SN, Pullen JK, Pease LR. 1989. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* 77:61–68.

Ishida K. 2005. Protein targeting into plastids: a key to understanding the symbiogenetic acquisitions of plastids. *J Plant Res.* 118:237–245.

Ishida K, Cao Y, Hasegawa M, Okada N, Hara Y. 1997. The origin of chlorarachniophyte plastids, as inferred from phylogenetic comparisons of amino acid sequences of EF-Tu. *J Mol Evol.* 45:682–687.

Ishida K, Ishida N, Hara Y. 2000. *Lotharella amoeboformis* sp. nov.: a new species of chlorarachniophytes from Japan. *Phycol Res.* 48:221–229.

Jarvis P, Robinson C. 2004. Mechanisms of protein import and routing in chloroplasts. *Curr Biol.* 14:R1064–R1077.

Kasai F, Kawachi M, Erata M, Yumoto K, Sato M, Ishimoto M. 2009. NIES-collection list of strains, 8th edition. *Jpn J Phycol (Sôrui).* 57:220.

Keeling PJ. 2004. Diversity and evolutionary history of plastids and their hosts. *Am J Bot.* 91:1481–1493.

Keeling PJ. 2009. Chromalveolates and the evolution of plastids by secondary endosymbiosis. *J. Eukaryot. Microbiol.* 56:1–8.

Keeling PJ, Archibald JM, Fast NM, Palmer JD. 2004. Comment on "the evolution of modern eukaryotic phytoplankton. *Science* 306:2191b

Kilian O, Kroth PG. 2005. Identification and characterization of a new conserved motif within the presequence of proteins targeted into complex diatom plastids. *Plant J.* 41:175–183.

Lane CE, van den Heuvel K, Kozera C, Curtis BA, Parsons BJ, Bowman S, Archibald JM. 2007. Nucleomorph genome of *Hemiselmis andersenii* reveals complete intron loss and compaction as a driver of protein structure and function. *Proc Natl Acad Sci U S A.* 104:19908–19913.

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

McFadden GI. 2001. Primary and secondary endosymbiosis and the origin of plastids. *J Phycol.* 37:951–959.

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

Nassoury N, Morse D. 2005. Protein targeting to the chloroplasts of photosynthetic eukaryotes: getting there is half the fun. *Biochim Biophys Acta.* 1743:5–19.

Nielsen H, Engelbrecht J, Brunak S, von Heijne G. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* 10:1–6.

Patron NJ, Waller RF. 2007. Transit peptide diversity and divergence: a global analysis of plastid targeting signals. *BioEssays.* 29:1048–1058.

Patron NJ, Waller RF, Archibald JM, Keeling PJ. 2005. Complex protein targeting to dinoflagellate plastids. *J Mol Biol.* 348:1015–1024.

Patron NJ, Waller RF, Keeling PJ. 2006. A tertiary plastid uses genes from two endosymbionts. *J Mol Biol.* 357:1373–1382.

Ralph SA, Foth BJ, Hall N, McFadden GI. 2004. Evolutionary pressures on apicoplast transit peptides. *Mol Biol Evol.* 21:2183–2194.

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

Rogers MB, Archibald JA, Field MA, Li C, Striepen B, Keeling PJ. 2004. Plastid-targeting peptides from the chlorarachniophyte *Bigelowiella natans*. *J Eukaryot Microbiol.* 51:529–535.

Rogers MB, Gilson PR, Su V, McFadden GI, Keeling PJ. 2007. The complete chloroplast genome of the chlorarachniophyte *Bigelowiella natans*: evidence for independent origins of chlorarachniophyte and euglenid secondary endosymbionts. *Mol Biol Evol.* 24:54–62.

Sommer MS, Gould SB, Lehmann P, Gruber A, Przyborski JM, Maier UG. 2007. Der1-mediated preprotein import into the periplastid compartment of chromalveolates? *Mol Biol Evol.* 24:918–928.

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

van Dooren GG, Schwartzbach SD, Osafune T, McFadden GI. 2001. Translocation of proteins across the multiple membranes of complex plastids. *Biochim Biophys Acta* 1541:64–53.