

Protein Targeting into the Complex Plastid of Cryptophytes

Sven B. Gould,¹ Maik S. Sommer,¹ Katalin Hadfi,¹ Stefan Zauner,¹ Peter G. Kroth,² Uwe-G. Maier¹

¹ Cell Biology, Philipps-University Marburg, Karl-von-Frisch Straße 8, 35042 Marburg, Germany
² Department of Biology, University of Konstanz, Universitätsstraße 10, 78457 Konstanz, Germany

Received: 25 April 2005 / Accepted: 25 July 2005 [Reviewing Editor: Dr. Yves Van de Peer]

The cryptophyte Guillardia theta harbors Abstract. a plastid surrounded by four membranes. This turns protein targeting of nucleus-encoded endosymbiont localized proteins into quite a challenge, as the respective precursors have to pass either all four membranes to reach the plastid stroma or only the outermost two membranes to enter the periplastidal compartment. Therefore two sets of nuclear-encoded proteins imported into the endosymbiont can be distinguished and their topogenic signals may serve as good indicators for studying protein targeting and subsequent transport across the outermost membranes of the cryptophyte plastid. We isolated genes encoding enzymes involved in two different biochemical pathways, both of which are predicted to be localized inside the periplastidal compartment, and compared their topogenic signals to those of precursor proteins for the plastid stroma, which are encoded on either the nucleus or the nucleomorph. By this and exemplary in vitro and in vivo analyses of the topogenic signal of one protein localized in the periplastidal compartment, we present new data implicating the mechanism of targeting and transport of proteins to and across the outermost plastid membranes. Furthermore, we demonstrate that one single, but conserved amino acid is the triggering key for the discrimination between nucleus-encoded plastid and periplastidal proteins.

Key words: Secondary plastids — Periplastidal compartment — Nucleomorph — Protein transport — Cryptophytes

Introduction

Plastids of land plants and green and red algae as well as of glaucocystophytes are surrounded by a double membrane, the plastid envelope (Martin et al. 2002; Cavalier-Smith 2003; McFadden and van Dooren 2004). Import of proteins into these so-called primary plastids, is accomplished in land plants by two translocons termed Toc and Tic (translocators of the <u>o</u>uter and <u>inner chloroplast membrane</u>, respectively), with the specificity for unidirectional transport of nucleusencoded plastid proteins (reviewed in Soll and Schleiff 2004).

Secondary or complex plastids are surrounded by one or two additional membranes and have evolved by the engulfment and subsequent reduction of a phototrophic alga containing primary plastids (secondary endosymbiosis [see Cavalier-Smith 2000; McFadden 2001; Prechtl and Maier 2002]). Some important algal groups harbor complex organelles, such as the peridinin-containing dinoflagellates and the phototrophic euglenoids with plastids surrounded by three membranes or heterokont algae, haptophytes, and the apicomplexa with four-membrane bounded plastids (Stoebe and Maier 2002, Cavalier-Smith 2003). Moreover, cryptophytes and chlorarachniophytes, both also possessing four-membrane bounded plast

Correspondence to: Uwe-G. Maier; email: maier@staff.uni-marburg.de

ids, are morphological intermediates, as in these species a remnant of the cytoplasm of the secondary endosymbiont is maintained between the outer and the inner membrane pair of the complex plastid (Maier et al. 2000). This cytoplasm, the periplastidal compartment, is devoid of typical eukaryotic compartments, e.g., mitochondrion and Golgi apparatus, but still harbors 80S ribosomes and a pigmy nucleus, the nucleomorph (Maier et al. 2000, Douglas et al. 2001; Cavalier-Smith 2002; Gilson and McFadden 2002).

Establishing an obligate endosymbiosis always includes a loss of the majority of the endosymbiont genes (Delwiche and Palmer 1997) or a transfer to the nucleus of the host cell (Martin and Herrmann 1998; Martin et al. 1998, 2002). As a major part of the gene products encoded by these transferred genes has to be reimported into the endosymbiont, evolution of a complex protein-transport and targeting system to the "destination compartment" was an essential prerequisite for the establishment of the organelles (Cavalier-Smith 1999). For example, the plastid genome of the cryptophyte Guillardia theta harbors genes for only 147 putative proteins and 36 RNA genes (Douglas and Penny 1999) and the nucleomorph encodes an additional 30 plastid proteins (Douglas et al. 2001), while at least a 10-fold amount of different proteins is expected to be necessary for its solar power-driven plastid (Delwiche and Palmer 1997; Martin et al. 2002) and therefore have to be reimported after being synthesized in the host's cytoplasm.

The basic components of the protein import machinery of primary plastids have been identified and functionally characterized in land plants. However, nearly nothing is known about possible components of the translocons in either red algae or complex plastids. Sequencing the genomes of the red alga Cyanidioschyzon merolae (Matsuzaki et al. 2004) and that of the diatom Thalassiosira pseudonana (Armbrust et al. 2004) led to the impression that protein transport across the plastid envelope of complex plastids of red algal origin might not require a TOC translocon. So far only genes encoding proteins that are somehow similar to Tic components have been identified (McFadden and van Dooren 2004). It can be assumed that mechanisms similar to those found in red algae are used for the import of nucleomorph-encoded plastid proteins into the stroma of cryptophytic plastids. These nucleomorph-encoded proteins as well as those imported into primary plastids are synthesized as a preprotein with an Nterminal extension, the transit peptide (Soll and Schleiff 2004), which is cleaved off by a stromal peptidase (Richter and Lamppa 2003), and thereby the mature protein is released for further subplastidal localization (van Dooren et al. 2001).

Independent of the phylogenetic origin of the secondary endosymbiont and its host, nucleus-encoded proteins with a destination in a secondary plastid are marked by a bipartite N-terminal topogenic signal (BTS), in which a signal sequence precedes a transit peptide (Waller et al. 1998; Kroth and Strothmann 1999; Sulli et al. 1999; Wastl and Maier 2000; Kroth 2002; Archibald et al. 2003; Nassoury et al. 2003 Patron et al. 2005). In vitro and in vivo studies showed that the signal sequence is responsible for targeting nucleus-encoded plastid proteins into either the secretory pathway or the space between the two outermost membranes of plastids with four surrounding membranes (Waller et al. 1998, 2000; Lang et al. 1998; DeRocher et al. 2000; Ishida et al. 2000; Wastl and Maier 2000; Apt et al. 2002; Kilian and Kroth 2004). In the case of the three-membrane-bounded plastids of dinoflagellates and euglenophytes, a tripartite signal peptide drives the precursor into the ER membrane (Sulli et al. 1999; Nassoury et al. 2003, Slavikova et al. 2005; Patron et al. 2005). Regarding the transport into four-membrane bounded plastids of chromists, the signal sequence of the precursor is cleaved after passing the first membrane (Apt et al. 2002), thereby leading to a preprotein with the ability to reach the stroma of the plastid with an N-terminal transit peptide only (Cavalier-Smith 2003).

For crossing the periplastidal membrane of complex plastids (PPM) with four surrounding membranes, two models are discussed, a vesicle shuttling mechanism and a specific protein transport via a translocon (Gibbs 1979; Cavalier-Smith 1999). Whereas in the vesicle shuttling model vesicles bud from the periplastidal membrane (PPM) and fuse with the outer plastid envelope membrane (Kroth 2002), a TOC-related import channel in the PPM is predicted in the translocon model (van Dooren et al. 2000; Cavalier-Smith 1999, 2003). An important implication is that in the vesicle shuttling model no TOCs are definitely needed. This is suggested by a recent report from McFadden and van Dooren (2004), which indicates that no TOC-encoding genes are present in the available genomic sequences of organism with secondary plastids of red algal origin (Plasmodium falciparum and Thalassiosira pseudonana).

When analyzing the mechanism of transport across the second outermost membrane, the uniqueness of cryptophytes and chlorarachniophytes—the periplastidal compartment (PPC)—is of major advantage, as the targeting of two sets of imported proteins can be distinguished: proteins for the PPC, which have to pass only the two outer membranes, and nucleus-encoded plastid proteins crossing all four membranes (Douglas et al. 2001; van Dooren et al. 2001). In cryptomonads starch is synthesized and deposited in the PPC (Douglas et al. 2001; Cavalier-Smith 2002) and not in the plastid itself as in the green lineage. Additionally, in the progenitor of the secondary endosymbiont of cryptophytes, the floridean starch (Meeuse et al. 1960) is synthesized from UDP-glucose instead of ADP-glucose (Viola et al. 2001). In the course of coevolution of two eukaryotic partners in secondary endosymbiosis, the deposition of carbohydrate compounds is retained in its localization in cryptophytes, whereas in heterokonts and haptophytes it is relocated into the host cytoplasm. As in cryptophytes the enzymes for starch biosynthesis are encoded in the cell nucleus, they need a topogenic signal that allows them to pass only the two outer membranes of the endosymbiont and, as a consequence, are excellent indicators for the evolution of these topogenic signals in general. At least parts of another biochemical pathway, carotenoid synthesis, were also predicted to be localized inside the PPC (Douglas et al. 2001), as geranylgeranyltransferase was found to be encoded on the nucleomorph without a N-terminal extension, which is typically found in all nucleomorph-encoded plastid proteins.

In this study we determined the differences in topogenic signals for host-encoded proteins targeted in either the plastid stroma or the PPC by comparing their topogenic targeting signals among each other and to those of plastid proteins encoded on the nucleomorph. By close analysis of a topogenic signal for one PPC located protein by in vivo and in vitro experiments, we come to the conclusion that a receptor-mediated import mechanism is used for both types of nuclear-encoded proteins crossing the PPM of cryptophytes and that one single aromatic amino acid is crucial for their discrimination.

Materials and Methods

Bioinformatics

Sequences were screened for signal peptide encoding regions by SignalP v3.0 (http://www.cbs.dtu.dk/services/SignalP/). Putative transit peptides were predicted with TargetP v1.01 (http://www. cbs.dtu.dk/services/TargetP/#submission) and iPSORT (http:// hc.ims.u-tokyo.ac.jp/iPSORT/). For blast searches, http://www. ncbi.nlm.nih.gov/BLAST/ and http://blast.genome.jp/ were used.

Sequence Analyses

Genomic sequences were obtained by performing standard polymerase chain reactions using *G. theta* total DNA and specific 5'and 3'-oligonucleotides derived from EST data. All PCR products were subcloned into pGEM-T (Promega) and sequenced on a LI-COR DNA Sequencer 4200 using Thermo Sequenase fluorescent labeled primer cycle sequencing kit with 7-deaza-dGTP (Amersham Pharmacia Biotech) and analyzed with Sequencher Software v4.0.5 (Gene Codes Corporation).

In Vitro Translation and Microsome Assay

Partial UGG-transferase from nucleotide 1 to nucleotide 777 (AJ784213) for in vitro translation was amplified via reverse tran-

scription-PCR from G. theta total RNA using 5'-CTCGA-GATGCGCCGTTCCGTTCTGTC-3' and 5'-TCTAGAATTACA TGGGTACCAAGGC-3' oligonucleotides and M-MuLV reverse transcriptase (Fermentas). For signal peptide processing, the TnT Coupled Reticulocyte Lysate System and Canine Pancreatic Micosomal Membranes from Promega were used following the manufacturer's manual, using 2 µl of microsomal membranes and Redivue PRO-MIX ([³⁵S]methionine and [³⁵S]cysteine with 1000 Ci/mmol; Amersham Pharmacia Biotech). For the protease protection assay, external translation product was digested with 0.1 µg/µl thermolysin (Sigma-Aldrich Biochemicals) treatment in the presence of 2 mM CaCl₂ for 15 min on ice. Thermolysin-treated fractions were then inhibited by adding EDTA to a final concentration of 10 mM, washed with 50 µl PBS, and concentrated by short centrifugation at 20,000g. Sample buffer was added to the remaining microsomes and analyzed by SDS-PAGE and subsequent Storage Phosphor Screen (Amersham Pharmacia Biotech).

In Vivo Import Experiments and GFP Fluorescence Analysis

Phaeodactylum tricornutum transformations were performed as described by Apt et al. (1996) with pPhaT1 plasmid containing the GFP constructs. Amino acid substitution from serine to phenylalanine of the UDP-AF presequence construct was performed by a point mutation of TCC to TTC at position 80. Analysis of transfected diatoms was performed with a confocal laser scanning microscope Leica TCS SP2 using HCX PL APO $40\times/1.25-0.75$ oil CS or PL APO $63\times/1.32-0.60$ oil Ph3 CS objectives. GFP and chlorophyll fluorescence was excited at 488 nm, filtered with a beam splitter (TD 488/543/633), and detected by two different photomultiplier tubes with bandwidths of 500–520 and 625–720 nm for GFP and chlorophyll fluorescence, respectively.

Results

As the nuclear genome of *G. theta* is not yet sequenced, we consulted our current EST project to find nucleus-located genes encoding endosymbiont proteins. For further analyses we first filtered those full-length transcripts which, by neural network tools like SignalP and TargetP, were predicted to bear an N-terminal bipartite topogenic signal (BTS), composed of a signal peptide at the N-terminus which is followed by a transit peptide. Furthermore, these topogenic signal regions needed to precede the mature part of a protein, which was determined by standard blast searches and protein databases.

To generate a set of nucleus-encoded plastid proteins (NuPPs) we first focused on the identification of conserved genes encoding proteins and enzymes being unique and essential for photosynthesis. Identification of nucleus-encoded proteins located in the PPC proved to be more complicated, because several housekeeping functions, which are still present inside the periplastidal compartment (e.g., protein biosynthesis or DNA replication), are also elementary in the host cell. Nevertheless, factors for maintaining the PPC need to be imported from the host's cytosol by passing the two outer membranes (Douglas et al. 2001; van Dooren et al. 2001). Hence, such proteins



Fig. 1. Representative predicted topogenic signals of nuclear- and nucleomorph-encoded proteins. Cathepsin H, cyclophilin B, and a protein disulfide isomerase (*catH*, *cycB*, *psi*) with a signal peptide (SP) are cotranslationally directed into the secretory system (SEC). Proteins located in the periplastidal compartment (PPC), UGG-transferase, uroporphyrinogen decarboxylase, and isopentenyl-diphosphate δ -isomerase (*ugg-t*, *hemE*, *iddi*), harbor a bipartite topogenic signal (BTS). Plastid (P)-localized proteins, e.g., phen-ylalanine-tRNA synthetase, light harvesting complex protein, and glyceraldehyde-3-phosphate dehydrogenase (*mpheS*, *lhcc13*, *gap*-

may be encoded in the host nucleus, and there at least twice, one copy for the host and one for the endosymbiont. This could be managed either by two different genes or by alternative splicing of one single message. However, in both cases clear discrimination of the two copies could be difficult. Fortunately at least one biochemical pathway is known to be exclusive for the PPC, starch synthesis and storage.

Conserved Topogenic Signals for Plastid Proteins

For Guillardia theta it was already known that nucleus-encoded plastid proteins possess a BTS with a cleavable signal peptide at the N-terminus followed by a transit peptide (Wastl and Maier 2000). Our EST data supplied several more NuPPs including proteins from the photosynthetic apparatus and the Calvin cycle and some stromal proteins, all showing the same characteristic BTS (Fig. 1). A detailed analysis of the BTS indicated that a conserved motif is evident at the border of signal and transit peptides from cryptophytes: the predicted cleavage site in most cases between an alanine and a phenylalanine, with an alanine at positions -1 and -3 making up an AXAF motif. With minor exception a phenylalanine at position +1, hence the first amino acid of the transit peptide, was observed (see Fig. 1). One exception was a tyrosine in the BTS of the plastidic phosphoglycerate kinase (data not shown). A similar conserved signal was described for nucleus-encoded plastid proteins of other secondarily evolved organisms as well (Cavalier-Smith 2003; Kilian and Kroth

DH), have a BTS. The photosystem II reaction center H protein, rubredoxin, and the translocator of the outer chloroplast membrane (*hcf136, rub, tic22*) possess only a transit peptide (TP), as they are plastid proteins encoded in the nucleomorph. The +1 amino acid of the signal peptide cleavage site of the nuclear-encoded proteins is in boldface. Note that all plastid proteins share a phenylalanine at their N-terminal transit peptide, which is not present in PPC proteins. Only the first 20 amino acids of the transit peptides are shown.

2004). In addition, we studied secretory proteins from G. *theta* to compare their signal peptides to the signal peptides of NuPPs and also observed an AXA in most cases, but as expected, neither a transit peptide nor the amino acid phenylalanine preceding the mature protein was evident (Fig. 1).

It should be noted that the transit peptides of the few definitely known plastid proteins encoded on the nucleomorph genome of *G. theta* share a small but obvious similarity, which is the short motif FXN right at their N-terminus (Cavalier-Smith 2003). Therefore the amino acid phenylalanine is conserved at the very N-terminal part of the transit peptides, whether they are nucleus or nucleomorph encoded.

A Bipartite Topogenic Signal for Crossing the Outer Two Membranes

The coding capacity of the nucleomorph does not cover all proteins required for the starch synthesis pathway, which is known to be localized inside the PPC (Douglas et al. 2001). Therefore enzymes like UDP glucose–starch glycosyl transferase (UGGtransferase), a transferase which links glucose to starch, needs to be imported from the host cytosol. In order to isolate the gene for the nucleus-encoded periplastidal protein UGG-transferase, we screened our EST and genomic libraries. A 51 bp long spliceosomal intron with the typical GT/AG borders, divides the 100 amino acid long BTS encoding region from the mature protein encoding region. A similar characteristic arrangement of the BTS and mature



Fig. 2. In vitro signal peptide processing of a periplastidal compartment-localized protein. Translation of the UGG-transferase presequence with rabbit reticulocyte lysate (RRL; lane 1) in the presence of canine microsomal membranes (CMM; lane 2) revealing a functional and processed signal peptide of the UGG-t preprotein. Protein is internalized into the microsomes, shown by the protease protection assay (lane 3) with Trition X-100 in lane 4.

protein-encoding region was recently shown for nucleus-encoded plastid protein sequences of *P. tricor-nutum* (Kilian and Kroth 2004) as well. Therefore this feature might indicate an evolutionary mechanism by which the topogenic signals needed for the reimport of plastid proteins have evolved.

In silico analyses revealed that the BTS consists of a 26-amino acid-long signal peptide (SignalP) followed by an amino acid stretch predicted to be a chloroplast transit peptide by iPSORT. A geranylgeranyl-transferase was identified to be encoded on the nucleomorph without a transit peptide (Douglas et al. 2001). Therefore it was suggested that at least parts of the carotenoid synthesis pathway are also located inside the PPC. Based on our investigation of the UGG-transferase topogenic signal, we were able to identify two other nucleus-located genes encoding proteins involved in the carotenoid synthesis harboring a BTS, the isopentenyl-diphosphate δ -isomerase (iddi; AJ937542) and the uroporphyrinogen decarboxylase (hemE; AJ937543). This indicated that all identified nucleus-encoded proteins of the PPC are encoded as a preprotein with a BTS.

In Vitro Analysis of a Predicted Signal Sequence of a PPC Protein

In order to demonstrate that the presequence of the UGG-transferase harbors a cleavable signal peptide, we initially performed an in vitro translation assay of the first 777 bp of the UGG-transferase gene and studied import into canine microsomal membranes. These experiments show that the N-terminal UGG-transferase fragment is imported into the microsomes and processed as predicted (see Fig. 2). The mass difference observed by SDS-PAGE between nonprocessed and processed protein fits the predicted mass of the signal peptide of 2.7 kDa. This finding together with our earlier results (Wastl and Maier 2000) demonstrates that N-terminal topogenic signals of

nucleus-encoded proteins for either the PPC or the plastid are generally designed in the same functional manner, indicating that the first import step of both PPC and plastid proteins at the ER membranes might be identical. This made it inevitable to characterize the preproteins more intensively to find the necessary difference to discriminate between the two groups of proteins.

A Single Conserved Amino Acid Inside the Transit Peptide Determines Localization

As already mentioned, the transit peptides of plastid proteins, whether encoded on the nucleus or nucleomorph, show one striking similarity: the phenylalanine at the very N-terminus. In order to test whether the second part of the BTS of the UGG-transferase is functional as a transit peptide, we analyzed the presequence in vivo. Moreover, we introduced an amino acid substitution, which changed the cleavage site motif between signal peptide and transit peptide from the wild-type SNA^VS (UGG-AS) into SNA^VF (UGG-AF) and fused these two 101-amino acid-long bipartite topogenic signals to GFP.

We chose the diatom Phaeodactylum tricornutum for in vivo GFP fusion protein localization studies, because G. theta cannot be selectively transformed and P. tricornutum-also belonging to the chromista—shows a similar complex plastid regarding the four membranes as the cryptophyte G. theta. UGG-AS directed the GFP to a concentrated spot, which is most likely to be localized within the diatom envelope membrane at the center of the plastid (Fig. 3). A similar localization was shown for GFP constructs of plastid proteins from *P. tricornutum*, when changing the cleavage site of the their BTS from, e.g., ASA^VF to ASA[♥]M. These localizations were termed "blob"like structures (Kilian and Kroth 2005) and proposed to consist of small vesicles. In contrast, the UGG-AF fusion protein showed a perfect colocalization with the autofluorescense of the plastid (see Fig. 3). Therefore the single amino acid substitution of serine to phenylalanine was sufficient for relocalization of the GFP fusion protein into the plastid and demonstrates that the BTS of the periplastidal compartment-localized UGG-transferase reassembles not only a signal peptide, but also a transit peptide, which becomes functional by changing only one amino acid.

Discussion

For every eukaryotic cell it is inevitable to optimize and establish functional communication and exchange of material between its compartments. This also holds true for organisms evolved in secondary endosymbiosis with plastids surrounded by four



Fig. 3. One single amino acid inside the bipartite topogenic signal (BTS) of the periplastidal compartment protein UGG-transferase determines different in vivo localizations. The wt-BTS directs the preprotein into a concentrated spot at the center of the plastid (row 1) of Phaeodactvlum tricornutum, A single amino acid substitution of serine to phenylalanine, which changes the cleavage site of the BTS from SNA▼S to SNA▼F, relocates the preprotein into the plastid (row 2). GFP without a topogenic signal is expressed in the cytosol (control, row 3). Bars represent 10 µm.

membranes. One major building block of evolution is to recycle and modify existing systems for new duties. All tested nucleus-encoded plastid proteins of secondary endosymbionts harbor a BTS, which consists in the N-terminal part of a signal peptide. In the case of chromists, the first step across the outermost membrane seems to be cotranslational via the Sec61 complex. This may be indicated by different in vivo and in vitro experiments (Bhaya and Grossman et al. 1991; Lang et al. 1998; Wastl and Maier 2000; Ishida et al. 2000; Apt et al. 2002). In cryptophytes, heterokonts, and haptophytes, the outermost membrane forms a continuum with the ER, called cER (chloroplast ER). Thus, after cotranslational import, nucleus-encoded proteins for the endosymbiont have to be distinguished from proteins for the exocytotic pathway. A topogenic signal for this discrimination of proteins could be the existing transit peptide. In vivo experiments clearly revealed that a transit peptide is necessary not only for import of proteins into the secondary plastids of chromists, but also for that of apicomplexa (e.g., DeRocher et al. 2000; Waller et al. 2000; Kilian and Kroth 2004). However, these experiments did not help to determine whether the transit peptide is the true topogenic signal to differentiate between endosymbiont proteins and those for exocytosis.

For analyzing protein targeting and import across the PPM of secondary plastids, cryptophytes do have the advantage that the remnant cytoplasm of the secondary endosymbiont is maintained. Nucleus-encoded proteins with a PPC localization are not expected to pass the plastid envelope, and therefore a transit peptide is theoretically not needed. By analyzing a gene encoding a protein for the endosymbiont's exclusive synthesis machinery, starch synthesis, we demonstrated that, contrary to our prediction, a transit peptide region is present in the UGG-transferase, although this enzyme does not pass the plastid envelope. As one can expect that nucleus-encoded PPC proteins (e.g., UGG-transferase) were nucleus encoded (today's nucleomorph) when a red algal cell was enslaved in secondary endosymbiosis, in this early stage the enzymes were not encoded as a preprotein with a transit peptide. Consequently, by transferring the genes to the host nucleus in cryptophytes, they did not encode it. Therefore the transit peptide of G. theta UGG-transferase preprotein cannot be a relic of its history; instead it has evolved by structural and mechanistically needs, which is also supported by the exon-intron structure of the gene. As the geranylgeranyl-transferase was found to be encoded on the nucleomorph without an N-terminal extension, it was also assumed that the carotenoid precursor synthesis might also be localized inside the PPC. Using the criteria of the UGG-transferase, we detected, in addition, iddi and hemE, both having an N-terminal BTS, which supports the earlier suggestion of Douglas et al. (2001) that carotenoid precursors are synthesized inside the PPC, but does not finally prove it.

The in silico prediction of a BTS in the UGGtransferase was proven to be correct in in vivo and in vitro experiments (see Figs. 1 and 2). We demonstrated that the signal peptide directs the protein into microsomes, indicating a functional topogenic signal at the N-terminus of this protein. Moreover, by fusing the BTS to GFP and stable transfection of

P. tricornutum, we observed a "blob"-like structure as described recently (Kilian and Kroth 2005). This blob-like structure indicates import into the endosymbiont, but not release of the protein into the stroma. Therefore, these experiments may reflect a situation comparable to the import into the PPC in cryptomonads. Our results clearly show the crucial importance of one specific amino acid-the phenylalanine-at the N-terminus of the predicted transit peptide. This amino acid seems to be the triggering key for the transport step across the plastid envelope inside the PPC and allows discrimination between plastid- and PPC-located proteins. This idea is supported by the fact that nucleomorph-encoded plastid proteins share a phenylalanine at their N-terminus. The importance of this amino acid is additionally supported by findings of Steiner and Löffelhardt (2002, 2005), who demonstrated that in glaucocystophytes, whose plastids are surrounded by two membranes, a phenylalanine at the third or fourth position of the transit peptide of nucleus-encoded plastid proteins is essential and to a certain degree can only be replaced by tyrosine (Steiner and Löffelhardt 2005) inevitable. Furthermore, recent investigations of secondary evolved organisms demonstrated that a phenylalanine-motif is conserved in the transit peptide not only in diatoms and dinoflagellates (Kilian and Kroth 2005; Patron et al. 2005; Chaal and Green 2005), but also in apicomplexa (Ralph et al. 2004). Taken together, the evidence indicates that a transit peptide is found in all precursor proteins intended to enter the endosymbiont and may be responsible for recognition of an endosymbiont protein inside the ER. Moreover, as the topogenic signals are exchangeable in cryptophytes and diatoms, protein import into the complex plastid may be conserved in these organisms and therefore might reflect a common origin of organisms with a red algal endosymbiont.

It seems that a BTS is necessary for the import of all proteins into the endosymbiont of cryptophytes, whether their final destination is the plastid or the PPC. However, at the moment we cannot determine whether proteins with a plastid or a PPC destination in cryptophytes are recognized by independent translocon complexes in the second outermost membrane, as proposed by Cavalier-Smith (2003), or if receptor coupled endocytosis is used for import, in which a transit peptide is the ligand for the receptor. It should be noted that in the vesicle-mediated transport model, PPC-located proteins must either use a different subclass of vesicles or escape them, before fusing with the outer chloroplast membrane. Furthermore an existing translocator in the third membrane is supported by the fact that nucleomorph-encoded proteins are capable of entering isolated plastids of G. theta with only the plastid envelope remaining (Wastl and Maier

2000), and their transit peptides also show the N-terminal phenylalanine to be conserved.

Our data demonstrate that the transit peptide of the BTS could be the general clue for how endosymbiont proteins are sorted out from those for the exocytosis pathway, perhaps utilizing the same mechanism whether plastid or PPC protein, and further discrimination then occurs inside the PPC by the presence or absence of a phenylalanine at the Nterminus. Further investigations are needed to clarify the detailed molecular mechanism of transport and sorting of nuclear-encoded endosymbiont proteins.

Acknowledgments. The authors thank Doris Ballert for help in diatom transfromation. Additionally, we thank Prof. E. Neuhaus (Kaiserslautern, Germany) and Steven Ball (Villeneuve d'Ascq, France) for helpful discussions on starch metabolism in cryptophytes. This work was supported by the DFG (project SFB593 to U.-G. Maier and project KR 1661/3-1 to P.G. Kroth). Accession numbers are as follows: AJ784213, AJ937541, AJ937542, AJ937544, AJ937545, AJ937546.

References

- Apt KE, Kroth-Pancic PG, Grossmann AR (1996) Stable nuclear transformation of the diatom *Phaeodactylum tricornutum*. Mol Gen Genet 252:572–579
- Apt KE, Zaslavkaia L, Lippmeier JC, Lang M, Kilian O, Wetherbee R, Grossman AR, Kroth PG (2002) Characterization of diatom multipartite plastid targeting signals. J Cell Sci 115:4061–4069
- Archibald JM, Rogers MB, Toop M, Ishida K, Keeling PJ (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
- Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou S, Allen AE, Apt KE, Bechner M, Brzezinski MA, Chaal BK, Chiovitti A, Davis AK, Demarest MS, Detter JC, Glavina T, Goodstein D, Hadi MZ, Hellsten U, Hildebrand M, Jenkins BD, Jurka J, Kapitonov VV, Kroger N, Lau WW, Lane TW, Larimer FW, Lippmeier JC, Lucas S, Medina M, Montsant A, Obornik M, Parker MS, Palenik B, Pazour GJ, Richardson PM, Rynearson TA, Saito MA, Schwartz DC, Thamatrakoln K, Valentin K, Vardi A, Wilkerson FP, Rokhsar DS (2004) The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. Science 306:79–86
- Bhaya D, Grossman A (1991) Targeting proteins to diatom plastids involves transport through an endoplasmic reticulum. Mol Gen Genet 229:400–404
- Cavalier-Smith T (1999) Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate and sporozoan plastid origins and the eukaryotic family tree. J Eukary Microbiol 46:347–366
- Cavalier-Smith T (2000) Membrane hereditiy and early chloroplast evolution. Trends Plant Sci 5:174–182
- Cavalier-Smith T (2002) Nucleomorphs: enslaved algal nuclei. Curr Opin Microbiol 2002(6):612–619, 2002
- Cavalier-Smith T (2003) Genomic reduction and evolution of novel genetic membranes and protein-targeting machinery in eukaryote-eukaryote chimaeras (meta-algae). Philos Trans R Soc London B 358:109–134

- Chaal BK, Green BR (2005) Protein import pathways in 'complex' chloroplasts derived from secondary endosymbiosis involving a red algal ancestor. Plant Mol Biol 57:333–342
- Delwiche CF, Palmer JD (1997) The origin of plastids and their spread via secondary endosymbiosis. Plant Syst Evol 11 (Suppl):53–86
- DeRocher A, Hagen CB, Froehlich JE, Feagin JE, Parsons M (2000) Analysis of targeting sequences demonstrates that trafficking to the *Toxoplasma gondii* plastid branches off the secretory system. J Cell Sci 113:3969–3977
- Douglas SE, Penny S (1999) The plastid genome of the cryptophyte alga, *Guillardia theta*: complete sequence and conserved synteny groups confirm its common ancestry with red algae. J Mol Evol 48:236–244
- Douglas SE, 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
- Gibbs SP (1979) The route of entry of cytoplasmatically synthesized into chloroplasts of algae possessing chloroplast ER. J Cell Sci 35:235–266
- Gilson PR, McFadden GI (2002) Jam packed genomes—a preliminary, comparative analysis of nucleomorphs. Genetica 115:13–28
- Ishida K, Cavalier-Smith T, Green BR (2000) Endomembrane structure and the chloroplast protein targeting pathway in *Heterosigma akashiwo* (Raphidophyceae, Chromista) J. Phycol 36:1135–1144
- Kilian O, Kroth PG (2004) Presequence acquisition during secondary endocytobiosis and the possible role of introns. J Mol Evol 58:712–721
- 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
- Kroth PG (2002) Protein transport into secondary plastids and the evolution of primary and secondary plastids. Int Rev Cytol 221:191–255
- Kroth PG, Strotmann H (1999) Diatom plastids: secondary endocytobiosis, plastid genome and protein import. Phys Plant 107:136–141
- Lang M, Apt KE, Kroth PG (1998) Protein transport into "complex" diatom plastids utilizes two different targeting signals. J Biol Chem 273:30973–30978
- Maier UG, Douglas S, Cavalier-Smith T (2000) The nucleomorph genomes of cryptophytes and chlorarchniophytes. Protist 151:103–109
- Martin W, Herrmann RH (1998) Gene transfer from organelles to the nucleus: How much, what happens, and why? Plant Physiol 118:9–17
- 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
- Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D (2002) Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc Natl Acad Sci USA 99:12246–12251
- Matsuzaki M, Misumi O, Shin-I T, Maruyama S, Takahara M, Miyagishima SY, 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, Nozaki H, Ogasawara N, Kohara Y, Kuroiwa T (2004) Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. Nature 428:653–657
- McFadden GI (2001) Primary and secondary endosymbiosis and the origin of plastids. J Phycol 37:951–959
- McFadden GI, van Dooren GG (2004) Evolution: red algal genome affirms a common origin of all plastids. Curr Biol 14:514– 516
- Meeuse BJD, Andries M, Wood JA, Wood M (1960) Floridean starch. J Exp Bot 11:129–140
- Nassoury N, Cappadocia M, Morse D (2003) Plastid ultrastructure defines the protein import pathway in dinoflagellates. J Cell Sci 116:2867–2874
- Patron NJ, Waller RF, Archibald JM, Keeling PJ (2005) Complex protein targeting to dinoflagellate plastids. J Mol Biol 348:1015–1024
- Prechtl J, Maier UG (2002) Zoology meets Botany: establishing intracellular organelles by endosymbiosis. Zoology 104:284–289
- Ralph SA, Foth BJ, Hall N, McFadden GI (2004) Evolutionary pressures on apicoplast transit peptides. Mol Biol Evol 21:2183– 94
- Richter S, Lamppa GK (2003) Structural properties of the chloroplast stromal processing peptidase required for its function in transit peptide removal. J Biol Chem 278:39497– 39502
- Slavikova S, Vacula R, Fang Z, Ehara T, Osafune T, Schwartzbach SD (2005) Homologous and heterologous reconstitution of Golgi to chloroplast transport and protein import into the complex chloroplasts of Euglena. J Cell Sci 118:1651–61
- Soll J, Schleiff E (2004) Protein import into chloroplasts [Review]. Nature Rev Mol Cell Biol 5:198–208
- Steiner JM, Löffelhardt W (2002) Protein import into cyanelles. Trends Plant Sci 7:72–77
- Steiner JM, Berghofer J, Yusa F, Pompe JA, Klosgen RB, Loffelhardt W (2005) Conservative sorting in a primitive plastid. The cyanelle of *Cyanophora paradoxa*. FEBS J 272:987–98
- Stoebe B, Maier UG (2002) One, two, three: nature's tool box for building plastids. Protoplasma 219:123–130
- Sulli C, Fang Z, Muchhal U, Schwartzbach SD (1999) Topology of Euglena Chloroplast protein precursors within endoplasmic reticulum to golgi to chloroplast transport vesicles. J Biol Chem 274:457–463
- Van Dooren GG, Waller RF, Joiner KA, Roos DS, McFadden GI (2000) Traffic jams: protein transport in Plasmodium falciparum. Parasitol Today 16:421–427
- Van Dooren GG, Schwartzbach SD, Osafune T, McFadden GI (2001) Translocation of proteins across the multiple membranes of complex plastids. Biochim Biophys Acta 1541:34–53
- Viola R, Nyvall P, Pedersen M (2001) The unique features of starch metabolism in red algae. Proc R Soc London 268:1417– 1422
- Waller RF, Keeling PJ, Donald RG, Striepen B, Handman E, Lang-Unnasch N, Cowman AF, Besra GS, Roos DS, McFadden GI (1998) Nuclear-encoded proteins target to the plastid in *Toxoplasma gondii* and *Plasmodium falciparum*. Proc Natl Acad Sci USA 95:12352–12357
- Waller RF, Reed MB, Cowman AF, McFadden GI (2000) Protein trafficking to the plastid of *Plasmodium falciparum* is via the secretory pathway. EMBO J 19:1794–1802
- Wastl J, Maier UG (2000) Transport of proteins into cryptomonads complex plastids. J Biol Chem 275:23194–23198