Early steps in plastid evolution: current ideas and controversies

Andrzej Bodył,[1]* *Paweł Mackiewicz,*[2] *and John W. Stiller*[3]

[1]Department of Biodiversity and Evolutionary Taxonomy, Zoological Institute, University of Wrocław, Wrocław, Poland
[2]Department of Genomics, Faculty of Biotechnology, University of Wrocław, Wrocław, Poland
[3]Department of Biology, East Carolina University, Howell Science 551, Greenville, NC, USA

Some nuclear-encoded proteins are imported into higher plant plastids *via* the endomembrane (EM) system. Compared with multi-protein Toc and Tic translocons required for most plastid protein import, the relatively uncomplicated nature of EM trafficking led to suggestions that it was the original transport mechanism for nuclear-encoded endosymbiont proteins, and critical for the early stages of plastid evolution. Its apparent simplicity disappears, however, when EM transport is considered in light of selective constraints likely encountered during the conversion of stable endosymbionts into fully integrated organelles. From this perspective it is more parsimonious to presume the early evolution of post-translational protein import *via* simpler, ancestral forms of modern Toc and Tic plastid translocons, with EM trafficking arising later to accommodate glycosylation and/or protein targeting to multiple cellular locations. This hypothesis is supported by both empirical and comparative data, and is consistent with the relative paucity of EM-based transport to modern primary plastids.

Keywords: carbonic anhydrases; endosymbiosis; evolution; plastids; protein import

Introduction

In accord with the endosymbiotic theory,[1–3] there is overwhelming evidence that plastids evolved from free-living cyanobacteria enslaved by heterotrophic eukaryotic cells. This process, called primary endosymbiosis, resulted in plastids surrounded by two membranes that are characteristic for glaucophytes, red algae, and green algae including their land plant descendants.[4–6] Most molecular phylogenetic analyses recover plastids as monophyletic[7,8] and it is widely accepted that they are derived from a single primary endosymbiosis in the common ancestor of the kingdom Plantae (or Archaeplastida), comprising these three algal/plant taxa[4–6] (but see Refs.[9–11] for alternative viewpoints).

***Correspondence to:** A. Bodył, Department of Biodiversity and Evolutionary Taxonomy, Zoological Institute, University of Wrocław, ul. Przybyszewskiego 63/77, 51-148 Wrocław, Poland.
E-mail: bodyl@biol.uni.wroc.pl

Ancestors of eukaryotes with primary plastids likely were phagotrophic protozoans that regularly captured cyanobacteria as food;[4,12] these engulfed cyanobacteria not only were photosynthetic, but also synthesized many organic compounds (*e.g.*, amino acids, heme, fatty acids[13–15]) and probably fixed nitrogen.[16,17] Therefore, selection favored changes in the host cell that retained endosymbionts as permanent and obligatory factories for production of essential metabolites. Eventually these permanent cyanobacterial endosymbionts were transformed into fully integrated organelles. This process involved two key innovations: (i) transfer of endosymbiont genes to the host nucleus, and (ii) origin of import machinery to move nuclear-encoded proteins back to the plastid through envelope membranes.[18] Modern plastids require between 2100 and 4800 different proteins to function,[19] but plastid genomes contain only 60–200 genes across various photosynthetic lineages.[20] Thus, during establishment of a primary plastid, the vast majority of cyanobacterial genes were either lost (*e.g.*, those encoding the tricarboxylic acid cycle) or transferred to the host nucleus.[20–22]

The fates of transferred cyanobacterial genes were disparate. Some became pseudogenes and were eventually lost, whereas others were adapted to encode cytosolic proteins.[20–22] In a next evolutionary stage, hundreds of these genes acquired targeting signals allowing transport of their protein products into distinct compartments, such as the EM system, mitochondria, and plastids.[23] The large majority of plastid-destined proteins carry N-terminal import signals known as transit sequences/peptides,[24,25] which target their protein products directly to the plastid. An incoming protein reaching the plastid surface is recognized and translocated across the two-membrane envelope by two translocons: (i) the translocon at the outer chloroplast membrane (Toc) and (ii) the translocon at the inner chloroplast membrane (Tic)[26–28] (Figs. 1 and 2). Each of these translocons is composed of a central protein-conducting channel and associated receptor, regulatory, and motor proteins (see Fig. 2). Available data indicate that the Toc–Tic supercomplex has a chimeric origin with some subunits derived from the endosymbiont and others from the host.[27,29,30] Thus, it is an interesting example of evolutionary tinkering with

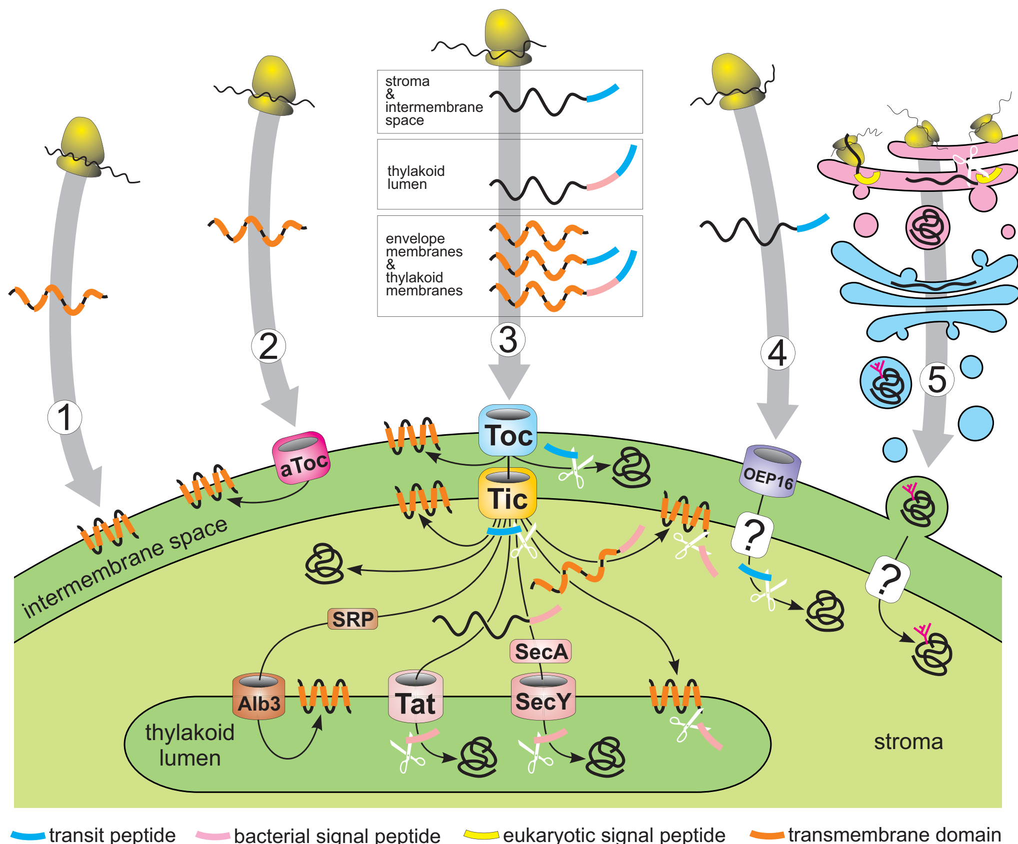

Figure 1. Distinct import routes of nuclear-encoded proteins into primary plastids. Primary plastids import at least 2000 different proteins, which are delivered to distinct sub-compartments (*e.g.*, the outer membrane, intermembrane space, stroma, thylakoid lumen) and carry different kinds of targeting signals. Interestingly, almost all nuclear-encoded plastid proteins are translocated across the envelope membranes of primary plastids with the help of the Toc–Tic super-complex[26–28] (route 3) containing Toc75 and Tic110 channels (see Fig. 2). These translocons participate in the import of the following groups of plastid proteins: (i) outer membrane-residing proteins devoid of N-terminal targeting signals but containing multiple transmembrane domains;[36] (ii) proteins destined to the intermembrane space, which carry typical transit peptides;[37] (iii) inner membrane-embedded proteins with pre-sequences composed of one (transit peptide) or two (transit peptide and signal peptide) domains;[38] (iv) stromal proteins carrying typical transit peptides;[26–28] (v) thylakoid membrane-embedded proteins having pre-sequences with a transit peptide (the Alb3 pathway) or a transit peptide plus signal peptide (the spontaneous pathway) and containing multiple transmembrane domains, some functioning as targeting signals;[39] (vi) proteins destined to the thylakoid lumen with bipartite N-terminal targeting signals composed of a transit peptide followed by a signal peptide.[39] In addition to the classical Toc- and Tic-based import route, alternative trafficking pathways were identified in primary plastids. An alternative Toc translocon (aToc) using a Toc75 homolog known as OEP80[40] (route 2) probably participates in the insertion of some proteins into the outer membrane. Moreover, some proteins appear to be inserted into this membrane spontaneously[41] *via* route 1. It is suggested that the OEP16 channel participates in the import of PORA[42] (route 4). All the above routes represent post-translational pathways; however, recently an EM system-mediated pathway (route 5) was found in higher plants. Proteins using this pathway (*e.g.*, α-carbonic anhydrase[31]) carry signal peptides instead of transit peptides. The proteins are glycosylated in the Golgi apparatus and finally delivered to the plastid surface in vesicles. This targeting pathway resembles protein import into eukaryotic alga-derived plastids (secondary endosymbionts) (see Fig. 6).

pre-existing components. After proteins are imported into the stroma, transit peptides are removed by a peptidase and mature polypeptides are correctly folded with the help of molecular chaperones[26–28] (Fig. 2).

One of the hundreds of proteins imported into higher plant plastids is the α-carbonic anhydrase CAH1,[31] which contains five *N*-glycosylation sites for α(1,3)-fucose and β(1,2)-xylose. Interestingly, unlike most plastid-directed proteins, CAH1 carries a typical endoplasmic reticulum (ER) signal peptide instead of a plastid transit peptide,[31] indicating that it is co-translationally translocated into the ER and then targeted to the plastid through the endomembrane (EM) system (Fig. 1). Experiments with brefeldin A, a fungal antibiotic that inhibits Golgi-mediated vesicular trafficking,[32] verified that CAH1 targeting involves both the ER and Golgi apparatus.[31] It remains possible that, after fusion of Golgi-derived vesicles with the outer plastid membrane, CAH1 is moved across the inner membrane using the Tic complex (Fig. 1), but total evidence indicates that its translocation occurs completely Toc independent.[31] Similar evidence exists for the rice

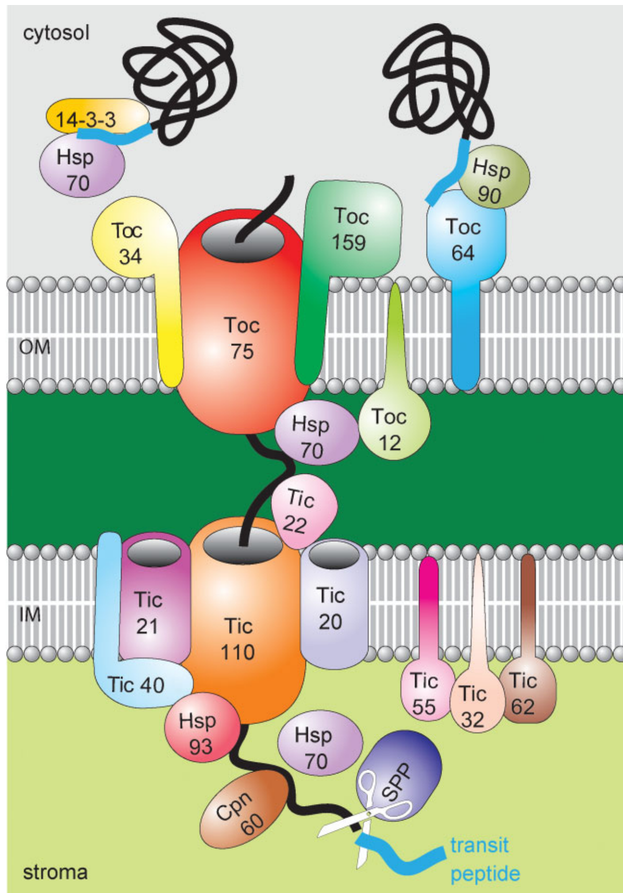

Figure 2. Structure of the Toc–Tic super-complex in higher plant plastids. The core Toc translocon, residing in the outer membrane (OM), is composed of three proteins: Toc34, Toc159, and Toc75.[26–28] Toc34 and Toc159 function as receptors for transit peptides of incoming proteins, whereas Toc75 forms a protein-conducting channel. Toc64 represents an additional receptor protein, but is loosely associated with the core Toc complex.[26–28] Nuclear-encoded plastid proteins are targeted to the distinct Toc receptors by two "guidance" complexes, composed of either 14-3-3 and Hsp70 proteins[43] or only Hsp90 proteins.[44] Toc12, Toc64, and Hsp70, along with Tic22, constitute an inter-envelope space complex that delivers imported proteins to the core Tic complex. In addition to Tic21, the core Tic translocon, existing in the inner membrane (IM), includes the following proteins: Tic20, Tic21, Tic40, and Tic110.[26–28] Tic110 constitutes the main pore for translocation of proteins through the inner plastid membrane, whereas Tic20 and Tic21 could represent additional protein-conducting channels. It also is suggested that Tic21 functions as an iron permease.[45] The stroma-exposed domain of Tic110 appears to cooperate with Tic40 and the stromal chaperone Hsp93 in the formation of a motor machinery pulling imported proteins into the stroma.[26–28] After reaching this sub-compartment, the transit peptide is removed by a stromal processing peptidase (SPP) and the mature protein is correctly folded with the help of Cpn60 and Hsp70.[26–28] Tic55, Tic62, and Tic32 are proposed to be involved in redox regulation of the protein import process.[26–28]

glycoprotein nucleotide pyrophosphate/phosphodiesterase,[33] and ER–Golgi trafficking could occur with other plant glycoproteins as well.[31]

The discovery of ER–Golgi-mediated targeting to primary plastids, combined with the complicated structure of the Toc–Tic super-complex, led to proposals that plastid-targeted proteins were initially imported through the EM system.[31,34,35] While this evolutionary scenario, termed the "early EM trafficking" hypothesis, presents an interesting new perspective on the early evolution of plastid import machinery, it encounters severe obstacles when viewed in the context of recent molecular and phylogenetic data. We propose an alternative scenario for the origin of protein import into primary plastids.

Gradual evolution of Toc and Tic translocons

A hypothesis of "early EM trafficking" rests on the proposition that the complex structures of Toc and Tic in higher plants make them unlikely candidates for transport functions early in the evolution of primary plastids.[34] This reasoning overlooks the likelihood of simpler ancestral forms of these translocons, which could have arisen quickly to provide insertional and/or transport activity. Consider, for example, the Toc complex composed of four receptor-channel subunits in higher plants[26–29] (Fig. 2). One of these proteins, the Toc64 receptor, is absent from red algae,[30] and cross-reaction experiments with heterologous antibodies suggest that glaucophytes are missing another receptor protein, Toc159.[46] Thus, simple parsimony suggests that the common ancestor of primary plastids contained a less complex translocation system (devoid of Toc64 and/or Toc159) than what is found in modern green plants. Following this backward trajectory, we suggest an ancestral outer membrane translocon that contained only Toc75, because this protein has both channel and receptor domains.[47] Higher plant plastids containing Toc75, but with other Toc receptors inactivated chemically[48] or genetically,[49] still can import plastid proteins with typical transit peptides, providing empirical evidence for this proposal. Finally, the outer plastid membrane contains "free" Toc75 proteins that mediate insertion of outer envelope proteins,[36] representing an extant working model for the ancestral state of Toc translocons.

At whatever stage they appeared, the multi-subunit structures of modern plant Toc and Tic complexes must have evolved in a series of steps, each stage providing a clear selective advantage. Without such an assumption it is difficult to imagine how complex translocons could have originated in any evolutionary scenario; nothing is gained by invoking an alternative early EM transport system, except an increase in the complexity of the overall model of plastid evolution.

It is reasonable that the first step in the evolution of the Toc–Tic super-complex was the establishment of protein-conducting channels in envelope membranes of the cyanobacterial endosymbiont. The plastid Toc75 pore evolved from an outer membrane protein 85 (Omp85) already present in the cyanobacterial endosymbiont,[47,50] where its vital function was to insert β-barrel porin-like proteins into the outer membrane.[51,52] In the initial endosymbiont, the opening of this channel was oriented toward the intermembrane space[46,53] and, therefore, could not participate in the import of nuclear-encoded plastid proteins. The latter only became possible after the *omp85/toc75* gene was transferred to the host nucleus, resulting in insertion of its protein product into the outer cyanobacterial membrane in a reverse orientation.[46,53] Moreover, the endosymbiont's inner membrane already contained a Tic20 homolog[27,28,54] belonging to preprotein and amino acid transporter (PRAT) family, along with bacterial channels for branched amino acids and mitochondrial Tim17, Tim23, and Tim22 channels.[55] Tic21 is another candidate for a pre-existing protein-conducting channel in the cyanobacterial inner membrane.[56] The presence of these proteins likely pre-adapted the inner endosymbiont membrane for translocating nuclear-encoded plastid proteins into the stroma. After Omp85/Toc75 was inserted into the outer membrane in a reverse orientation, the endosymbiont would have been capable of translocating inter-membrane space- and inner membrane-residing proteins, and probably stroma-destined proteins as well.

The initial efficiency of plastid protein import was probably low, but gradual addition of receptors (Toc34, Toc159, and Toc64) and regulatory proteins (Tic32, Tic55, and Tic62) increased it step by step.[26–30] Homologs of *toc* and *tic* genes are detectable in all cyanobacterial genomes[30,57] and were clearly present in plastid ancestors, facilitating the evolution of the Toc–Tic machinery. A final improvement was the evolution of "guidance complexes" in the host cytosol, one composed of 14-3-3 proteins and heat shock proteins 70 (HSP70s)[43] and a second involving only HSP90s[44] (Fig. 2); these significantly increased the efficiency of protein delivery to the plastid surface. Clues to the gradual evolution of plastid translocons are also evident for individual Toc and Tic subunits, *e.g.*, the Toc159 receptor (see Fig. 3). Such gradual evolution clearly explains how even complex import systems could have originated, without any additional involvement from the host's EM system (see also Ref.[58]).

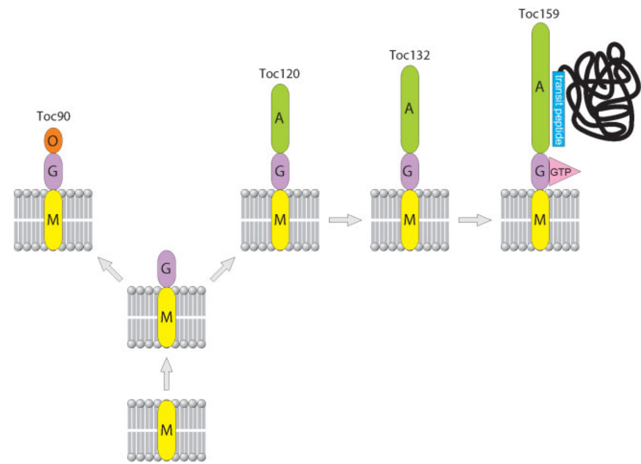

Figure 3. Evolution of Toc159 and its homologs. Toc159 is one of the core Toc receptors for nuclear-encoded plastid-targeted proteins.[26–28] It is also implicated as a molecular motor pushing imported proteins into the Toc75 pore.[59] Toc159 has three distinct domains: a membrane (M) domain, a GTP-binding (G) domain, and an acidic (A) domain that putatively interacts with transit peptides of imported plastid proteins.[26–28] Import studies with *ppi2* mutants without functional Toc159[60] demonstrate that the M domain alone can restore protein import into *Arabidopsis* plastids.[61] This suggests the M domain could represent the ancestral state in the evolution of the Toc159 receptor, with the G and A domains added in two subsequent evolutionary steps. Green plants possess three Toc159 homologs: Toc90, Toc132, and Toc120.[62–64] Toc90 is entirely devoid of the A domain, whereas Toc132 and Toc120 contain different numbers of repeats within this domain. Genetic and biochemical studies suggest that Toc159 and its homologs form at least two distinct Toc translocons in higher plant plastids.[60,62–64] The first of them, containing Toc159, would be responsible for recognition and translocation of highly abundant, photosynthetic proteins, while the alternative translocon with the Toc132 and/or Toc120 receptors would import low-abundance, housekeeping proteins. O, other domain.

Alternative explanations for signal peptides in plastid proteins

The first analyses of the plastid proteome of *Arabidopsis thaliana* suggested that up to 8% of its proteins carry signal peptides.[65] This unexpected finding implied that a large number of plastid proteins could be imported *via* the EM system,[65,66] suggesting support for "early EM trafficking."[34] However, initial estimates of stroma-targeted proteins with signal peptides were apparently inflated by (i) false-positive identifications, (ii) non-plastid contamination, and (iii) outer membrane-targeting N-terminal transmembrane domains that are frequently misidentified as signal peptides. Zybailov et al.[67] estimated that these problems could be relevant for 320–370 of the total 1090 plastid proteins calculated from different proteomic studies. Indeed, when 18 known outer envelope proteins were excluded, only 0.6% of *Arabidopsis* nuclear-encoded plastid proteins were recognized by a signal peptide-predicting program.[67] Moreover, negative controls using cytosolic and mitochondrial proteins produced similar or even higher rates of signal peptide prediction. Thus, it seems unlikely that a significant fraction of plastid proteins are imported through the EM system in land plants or algae with primary plastids.

We propose two alternatives to the "early EM trafficking" hypothesis to explain the occurrence of signal peptides in a small fraction of higher plant plastid proteins. First, signal peptides permit plastid proteins to be targeted to two or more compartments. An example is α-amylase of rice (αAmy3), which is targeted both to the plastid and externally to the cell wall.[68] Second, the presence of signal peptides in some higher plant plastid proteins can be explained by the need for glycosylation in the Golgi apparatus. It is well known that α(1,3)-fucose and β(1,2)-xylose, both characteristic of CAH1 and other plastid glycoproteins,[31] are added specifically within the Golgi apparatus.[69] Thus, plastid proteins carrying these sugar epitopes must be targeted *via* the EM system.

There are three distinct protein families of carbonic anhydrases, designated α, β, and γ.[70] CAH1 protein from *A. thaliana* belongs to the α family,[31] which is widely distributed in eukaryotes and bacteria.[70] Targeting of *Arabidopsis* CAH1 to the plastid is the exception rather than the rule for this protein family. The α-carbonic anhydrase from tobacco is a secretory protein found in nectar but not in plastids.[71] Dioscorin is another member of this family and is deposited in vacuoles of tuber cells of yam (*Dioscorea* spp.).[72] Finally, cell wall-targeted α-carbonic anhydrases were identified in the green algae *Chlamydomonas reinhardtii*,[73] *Chlorella sorokiniana*,[74] and *Dunaliella salina*.[75] In agreement with extracellular and vacuolar localizations of these anhydrases, they all carry typical signal peptides at their N termini.[71–75] Plastid residence has been reported only for one other α-carbonic anhydrase, CAH3 from *C. reinhardtii*,[76] but unlike *Arabidopsis* CAH1, its N-terminal extension is bipartite with a plastid transit peptide followed by a hydrophobic domain[76] as found in other thylakoid lumen-targeted proteins.[39] Carbonic anhydrases targeted to the stroma of higher plant plastids typically belong to the β family.[70] This suggests that *Arabidopsis* CAH1 was adapted secondarily for import into the plastid, probably replacing an original β-carbonic anhydrase (see Gagat et al., manuscript in preparation).

The needs for post-translational modifications and localization to multiple cell compartments provide clear selective advantages for the evolution of EM-mediated targeting to higher plant plastids. Such unique elaborations of trafficking pathways in a handful of proteins are unlikely to be vestiges of the early stages in plastid evolution. Thus, this unconventional, less common import pathway is most easily explained as a derived rather than an ancestral system.

Sugar/inorganic phosphate antiporters

Another argument advanced to support "early EM trafficking"[34] is based on phylogenetic analyses of plastid metabolite transporters.[77] Because primary plastids are involved in diverse metabolic functions,[13–15] they exchange many compounds with the surrounding cytosol and other organelles, such as mitochondria and peroxisomes (see, for example, the photorespiratory pathway in plant cells[14]). To fulfill these requirements, the plastid inner membrane contains several kinds of transporters, including sugar/inorganic phosphate antiporters represented by triose phosphate (TPT), glucose-6-phosphate (GPT), and phosphoenolpyruvate (PPT) translocators.[78,79] In phylogenies inferred from red algae and green plant sequences, TPTs, GPTs, and PPTs form a single clade that is related to sugar/phosphate antiporters residing in the ER and/or Golgi apparatus.[77] Based on these results, it was proposed that all three translocators evolved from an EM transporter *via* gene duplications and, at least initially, were targeted to the plastid through the EM system.[77] However, further analyses are required, as non-plastid transporters on these trees were represented only by ER/Golgi homologs;[77] thus, it is possible that TPTs, GPTs, and PPTs evolved, for example, from a mitochondrial rather than an EM transporter.

More significantly, our computational analyses of the N-terminal extensions of these three transporters from the red alga *Cyanidioschyzon merolae*,[80] which represents an early branch in red algal and primary plastid evolution,[81] indicate that they are typical plastid transit peptides (unpublished results). For example, seven programs that distinguish different kinds of N-terminal targeting signals (PredSL, TargetP, iPSORT, Predotar, PProwler, BLSTM_LOC, TargetLoc) failed to predict signal peptides in any of these sequences. In contrast, programs specifically designed to predict plastid transit peptides (ChloroP and PCLR) identified them in leader sequences from all *C. merolae* transporters. Thus, regardless of their evolutionary ancestor, current evidence suggests that TPTs, GPTs, and PPTs were not imported into plastids *via* the EM, even relatively early in plastid evolution. Finally, the early diversification postulated for these plastid sugar/phosphate antiporters would have faced serious obstacles, if they had been trafficked through the EM system. Mis-targeting of plastid antiporters to alternative EM locations, and of ER/Golgi-localized transporters to the plastid, would have been common. Direct post-translational transport of proteins into early plastids would not have faced this obstacle.

Do eukaryote-derived plastids recapitulate early steps in primary endosymbioses?

Some plastids evolved secondarily from endosymbiotic eukaryotic algae with primary plastids (green or red algae), resulting in three or four surrounding membranes[4–6]

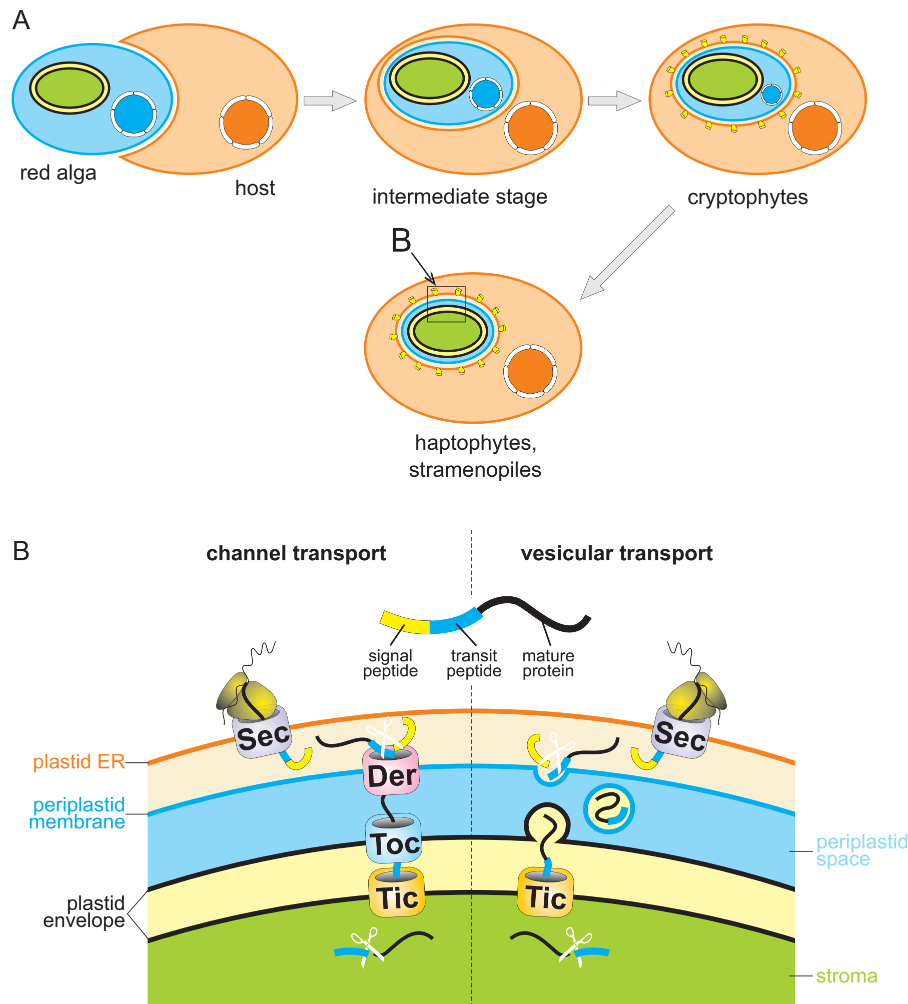

Figure 4. Evolution of secondary plastids and their import apparatus. Secondary plastids evolved from algae with primary plastids (*e.g.*, red algae) that were engulfed by phagocytosis, resulting in their four membrane envelopes.[4–6] Red algal-derived plastids with four envelope membranes are found in three algal lineages: Cryptophyta, Heterokonta, and Haptophyta. Their two innermost membranes (or the plastid envelope) are derived from the red algal primary plastid, whereas the next layer, known as the periplastid membrane, represents the red algal plasmalemma.[82–84] The outermost plastid membrane is derived from the host's phagosome, but now is covered with ribosomes, suggesting that this membrane fused with a rough ER membrane, resulting in the plastid ER (PER) and placement of the entire complex plastid within the ER lumen.[82–84] For this reason, nuclear-encoded plastid proteins in cryptophytes, heterokonts, and haptophytes carry bipartite pre-sequences composed of a signal peptide followed by a transit peptide.[25] The first step in their import is co-translational translocation through the PER membrane dependent on the Sec translocon.[84,91] In the "vesicular" model, transport through the periplastid membrane and the outer membrane of the plastid envelope are mediated by transport vesicles derived from the pinocytotic pathway of the red algal endosymbiont.[91,96] By contrast, the "channel" model postulates that two distinct pore proteins, Der and Toc75, are responsible for these targeting steps.[93,94] Toc75 pre-existed in the outer membrane of the red algal plastid, whereas Der was relocated from the red algal ER to its plasmalemma (= the periplastid membrane). Both models assume that translocation across the inner membrane of the plastid envelope is dependent on the Tic translocon.[91,93,94] The signal peptide is cleaved off during or after translocation across the PER membrane, whereas the transit peptide is removed in the stroma.[82–84] Available data favor the "channel" over the "vesicular" model.[84,93,94]

(Fig. 4A). Euglenids and dinoflagellates have secondary plastids with three membranes, whereas four envelope membranes occur in chlorarachniophytes, cryptophytes, heterokonts, haptophytes, and apicomplexans. In all cases, it is assumed that the two innermost membranes correspond to the primary plastid envelope, while the outermost membrane is derived from the host's phagosomal membrane[82–84] (Fig. 4B).

For plastids with four envelope membranes, the "periplastid" membrane, localized between the primary plastid envelope and the outermost membrane, is widely held to be the modified plasmalemma of the engulfed eukaryotic

alga[82–84] (Fig. 4B). Nuclear-encoded plastid proteins targeted to secondary plastids carry complex pre-sequences composed of a signal peptide followed by a transit peptide[25] (Fig. 4B), and their targeting pathway involves the EM system.[82–84] In secondary plastids where the outermost membrane is devoid of ribosomes (*e.g.*, euglenids), nuclear-encoded proteins are delivered to the plastid surface in vesicles budding off the ER or Golgi apparatus,[85] whereas in plastids with the outermost membrane covered with ribosomes (*e.g.*, cryptophytes) they are translocated directly into the space between the outermost and periplastid membranes.[86] It quickly became evident that protein translocation across their outermost membrane proceeds co-translationally (requiring a signal peptide and the Sec translocon), whereas transport through the two innermost membranes occurs post-translationally (involving a transit peptide and the Toc–Tic translocons).[82–84]

''Early EM trafficking'' into primary plastids has been modeled on EM-mediated targeting to secondary plastids.[34] Before discussing the serious problems with this analogy below, we note that transformation of initial co-translational import into a post-translational system has never occurred in any of numerous lineages with secondary plastids.[82–84,87] It appears that such a transformation would be traumatic or even lethal for the host cell. Loss of the outermost, phagosome-derived membrane, a pre-requisite for evolution of a post-translational system, would imprison nuclear-encoded proteins in ER- or Golgi-derived vesicles that could not fuse with the new plastid outer membrane (Fig. 5A). Thus, their further import into the stroma would be impossible, disrupting all vital plastid functions.

All available evidence suggests that when ER and/or Golgi-mediated targeting was canalized during establishment of new endosymbiotic relationships, it became exceptionally stable through subsequent evolution.[82–84,87] Thus, assuming secondary endosymbioses as the model, if nuclear-encoded proteins had originally been imported into primary plastids through the EM system, these plastids should still be surrounded by three membranes and retain this initial step of ER–Golgi-mediated targeting. It is interesting that for years Cavalier-Smith[12,88] used similar logic to ''early EM trafficking'' to argue that three-membrane euglenid and dinoflagellate plastids evolved directly from the primary cyanobacterial endosymbionts and that they represented an ancestral stage of primary plastid evolution. Both hypotheses have been widely rejected.[89,90]

A very similar scenario to an ''early EM trafficking'' hypothesis was proposed earlier by Kilian and Kroth[91] to explain the evolution of protein import into secondary plastids with four envelope membranes. The mechanism of protein passage across the periplastid membrane remained unclear for some time. Two alternative models were developed: The first (or channel) model postulated the use of pore-forming proteins such as Toc75 relocated from the endosymbiont's plastid,[83] the host mitochondrion-derived Tim23,[92] or Der relocated from an endosymbiont's ER membrane[93,94] (Fig. 4B). Der forms a channel in the ER-associated degradation (ERAD) system linked to the Sec translocon that participates in retro-translocation of misfolded proteins into the cytosol, where they are degraded by proteasomes.[95] The alternative (or vesicular) model for protein translocation across the periplastid membrane argued that transport vesicles pinch off the periplastid membrane, then fuse with the primary plastid's outer membrane to liberate imported proteins into the periplasmic space[91,96] (Fig. 4B). From there proteins could be translocated into the stroma *via* the Tic translocon.

To explain the evolution of this intraplastidal protein import pathway, it was hypothesized that proteins were initially imported into primary plastids through the EM system, because such a pathway required the presence of SNARE proteins (responsible for membrane recognition and fusion[97]) in the outer membrane of the red or green algal plastid.[91] Consequently, pinocytotic vesicles budding off the endosymbiont plasmalemma would be able to fuse with the outer membrane of its plastid and establish an intra-plastidial transport route for nuclear-encoded proteins. Recent experimental work by Uwe Maier's group, however, demonstrates that protein translocation across the periplastid membrane is mediated by Der[93,94] rather than through vesicular fusions (Fig. 4B). Thus, all current evidence indicates that even secondary endosymbioses have not evolved the kind of vesicular transport on which an early evolution of EM trafficking to primary plastids has been modeled.

Does the outer plastid membrane contain alternative protein-conducting channels to the Toc system?

One of the ideas that led to an ''early EM trafficking'' hypothesis was the suggested presence of alternative protein-conducting channels to the typical Toc75 pore in the outer membrane of primary plastids.[34] Two proteins have been cited in support of this view: Tic32[98] and chloroplast envelope quinone oxidoreductase (ceQORH).[99,100] Both reside in the inner plastid membrane[98–100] but, in contrast to most proteins targeted to this membrane,[38] they contain internal, uncleaved targeting signals[98–100] rather than typical N-terminal, cleavable transit peptides.[38] In addition, competition experiments for the Toc complex, using precursors of RuBisCO, ferredoxin, and the 33-kDa subunit of oxygen evolving complex (OE33) did not affect the import of Tic32 and ceQORH.[98–100] Based on these results, it was suggested that these two proteins are imported through uncharacterized Toc75-independent channels.[98–100]

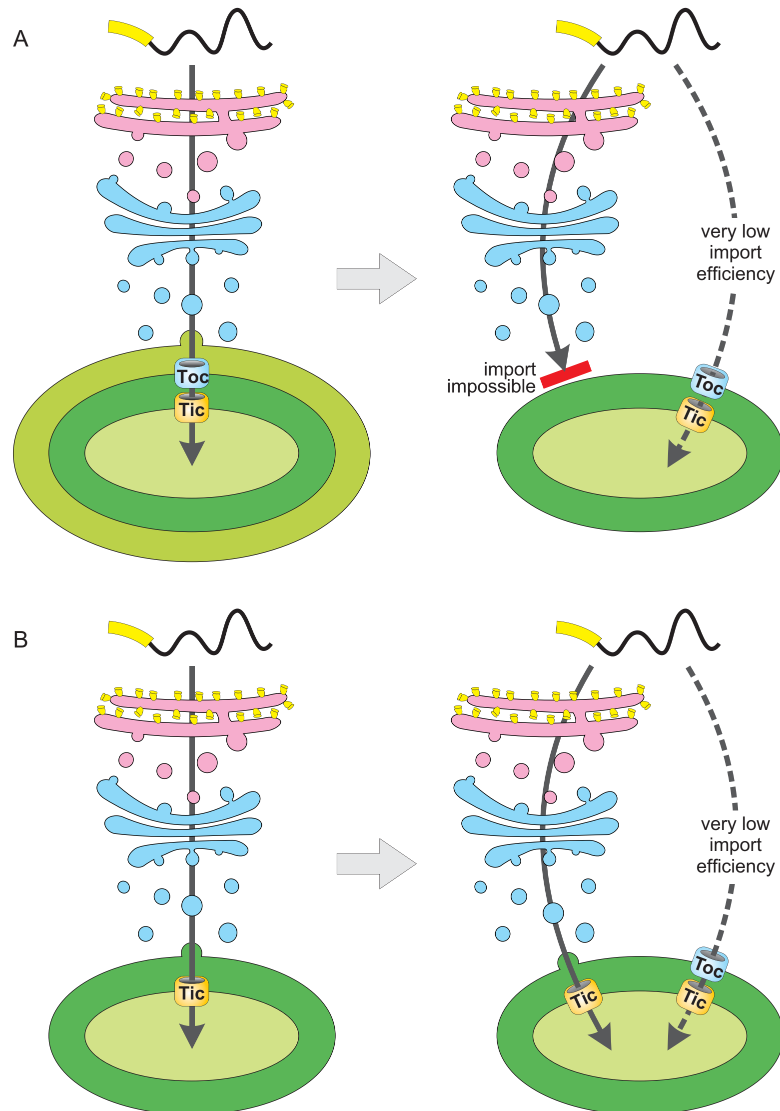

Figure 5. Two alternative evolutionary scenarios explaining transformation of a co-translational to a post-translational transport system in the evolution of primary plastids. The first model (**A**) postulates that primary plastids were initially surrounded by three membranes, with the inner and middle membranes derived from the envelope membranes of the cyanobacterial endosymbiont, and the outer membrane from the host's phagosome.[13,88] Thus, at this initial stage of evolution, they would have resembled eukaryote-derived plastids of euglenids and dinoflagellates.[82–84] Their nuclear-encoded proteins carried signal peptides that targeted them to the plastid *via* the EM system involving the ER and Golgi apparatus.[85,90] After fusion of Golgi-derived pre-plastid vesicles with the outer plastid membrane, imported proteins were translocated to the stroma with the help of the Toc–Tic super-complex.[85,90] In a next evolutionary stage, however, primary plastids lost the outer, phagosome-derived membrane for an unknown reason, leading to exposition of the Toc translocon to the cytosol and permitting subsequent evolution of a direct post-translational import mechanism.[88] After such a loss, however, import of hundreds of nuclear-encoded proteins would have been prevented by the inability of pre-plastid vesicles to fuse with the new outer plastid, cyanobacterial-derived membrane. These obstacles could be bypassed by postulating modifications of signal peptides to transit peptides (to prevent their interaction with the signal recognition particle (SRP) complex and co-translational translocation into the ER[88]), but it is unlikely that such changes would have occurred nearly simultaneously in perhaps several thousand plastid-targeted proteins as required under this scenario. Moreover, changes permitting post-translational transport would have produced a much less effective system than the original and highly adapted EM-based transport, thereby disturbing vital plastid functions performed by affected proteins. Thus, selection almost certainly would have rejected individual changes that could eventually result in loss of the outermost membrane, as has been the case in extant algal lineages with three membrane-bound plastids. The alternative model (**B**) postulates that primary plastids, from the beginning, were surrounded by two membranes and that the original import mechanism for protein transport across the outer membrane was based on the ER–Golgi-mediated pathway.[34] In a next evolutionary stage this co-translational system would be transformed to a post-translational one. As with the previous model, however, it is hard to imagine how such a conversion could have proceeded. Modifications of signal peptides could result in a post-translational protein import, but such import would be ineffective and would preclude import through the EM system. Consequently, there would likely have been strong selection against modifications of signal peptides.

However, Inaba and Schnell[28] pointed out that competitor proteins used in these assays engage the classical Toc159 receptor, and that Tic32 and ceQORH could use alternative Toc translocons containing Toc75 but having alternative Toc159 homologs such as Toc132, Toc120, and Toc90 (see also Refs.[62–64]). Moreover, if Tic32 turned out to be imported through an alternative protein-conducting channel (its import is not inhibited by spermine, which blocks the Toc75 pore[98]), this would not preclude its initial translocation through Toc75. In support of this view, its targeting information is localized near the N-terminus,[98] suggesting it could have evolved from a typical transit peptide. Finally, the Tom40 channel in the outer mitochondrial membrane,[101] which is homologous to Toc75,[102] translocates proteins with typical N terminal, cleavable transit peptides, and also those with internal, non-cleavable signals.[103]

Another candidate for a Toc75-independent pore is outer envelope protein 16 (OEP16), a cation-selective high-conductance channel with remarkable selectivity for amino acids and amines[104] (Fig. 1). Import experiments performed by Steffen Reinbothe's group provided evidence for OEP16-dependent import of protochlorophyllide oxidoreductase A (PORA),[42,105] although these results were questioned or reinterpreted by other researchers.[106,107] Even if OEP16 participates in the import of PORA, this will not suggest that it played a role in early plastid evolution. Along with mitochondrial Tim17, Tim22, and Tim23 channels, OEP16 belongs to the PRAT family[55] and it has been proposed that this protein evolved from a Tim channel that was relocated from the inner mitochondrial membrane to the outer plastid membrane.[58] Moreover, this relocation probably occurred only in the green algal/plant lineage, because the nuclear genome of the red alga *C. merolae*[80] does not contain a recognizable OEP16 homolog. Finally, the presence of a Toc-independent channel capable of translocating proteins across the cyanobacterial endosymbiont's outer membrane would favor evolution of post-translational rather than co-translational import of nuclear-encoded proteins.

All available data clearly indicate that Toc75 plays the central role in importing a wide variety of plastid proteins (Fig. 1). It not only enables efficient translocation of stromal proteins with classical transit peptides,[26–28] but also proteins destined for the intermembrane space[37] and the plastid inner membrane.[38] These latter proteins carry bipartite pre-sequences composed of a transit peptide followed by a hydrophobic domain, which functions as an export signal or a stop-transfer sequence.[38] The Toc75 pore also constitutes an entrance site for proteins imported into the thylakoid membrane as well as the lumen.[39] Some of these proteins carry bipartite pre-sequences, and those targeted to the thylakoid membrane also possess multiple additional hydrophobic membrane-spanning domains.[39] Finally, Toc75 participates in the insertion of proteins into the outer plastid membrane[36] and an additional Omp85/Toc75 homolog was identified in the outer membrane of primary plastids[40] (see also Fig. 1). The centrality of Toc75 and its homologs in transporting the broad diversity of plastid-directed proteins strongly supports the hypothesis that it was engaged as the primary outer membrane channel very early in the evolution of host cell to plastid targeting.

Alternative scenario for the evolution of the protein import apparatus of primary plastids

Cyanobacteria are Gram-negative bacteria with an envelope composed of a plasmalemma and an additional outer membrane.[108] Because the cyanobacterial ancestor of plastids very likely entered the host cell *via* phagocytosis, it was originally surrounded by three membranes: the host's phagosomal membrane and the two envelope membranes of the endosymbiont.[4,12,88] Modern primary plastids have only a two-membrane envelope, meaning that one of these three membranes, most likely either the phagosomal or the outer cyanobacterial membrane, was eliminated during their evolution. Which of these membranes was lost is, however, unclear because the current outer plastid membranes have features of both eukaryotic phagosomal and bacterial outer membranes.[109] Given that peptidoglycan is still present in the glaucophyte plastid,[110] it is most reasonable to assume that the ancestral primary plastid lost the phagosomal membrane (in support of this view see Ref.[111]). We suggest a scenario in which uncoordinated division of the endosymbiont and the phagosome resulted in regular escapes of endosymbionts into the host cytosol. During these escapes the outer cyanobacterial membrane could have acquired some lipids and proteins from the phagosomal membrane (a kind of membrane mutation[4]), thereby accounting for its chimeric bacterial-eukaryotic nature.

The question arises, however, at what stage it was more likely for the phagosomal membrane to be lost. One hypothesis, originally formulated by Cavalier-Smith[12,88] and briefly discussed in a previous section, posits this loss relatively late in the evolution of primary plastids (see also Fig. 5A). Such a scenario is, however, exceedingly unparsimonious for two main reasons: First, the phagosomal membrane initially surrounding each engulfed cell was incapable of both permanent growth and division,[112] two features needed to establish a permanent endosymbiosis.[53] It could grow intermittently through fusions with pre-lysosomal vesicles, but these would result in digestion of the endosymbiont. Thus, permanent incorporation of the phagosomal membrane into the primary plastid envelope required the origin of new mechanisms for its controlled growth and division.[112] Later, only after all these innovations had originated, this modified (now symbiosomal) membrane

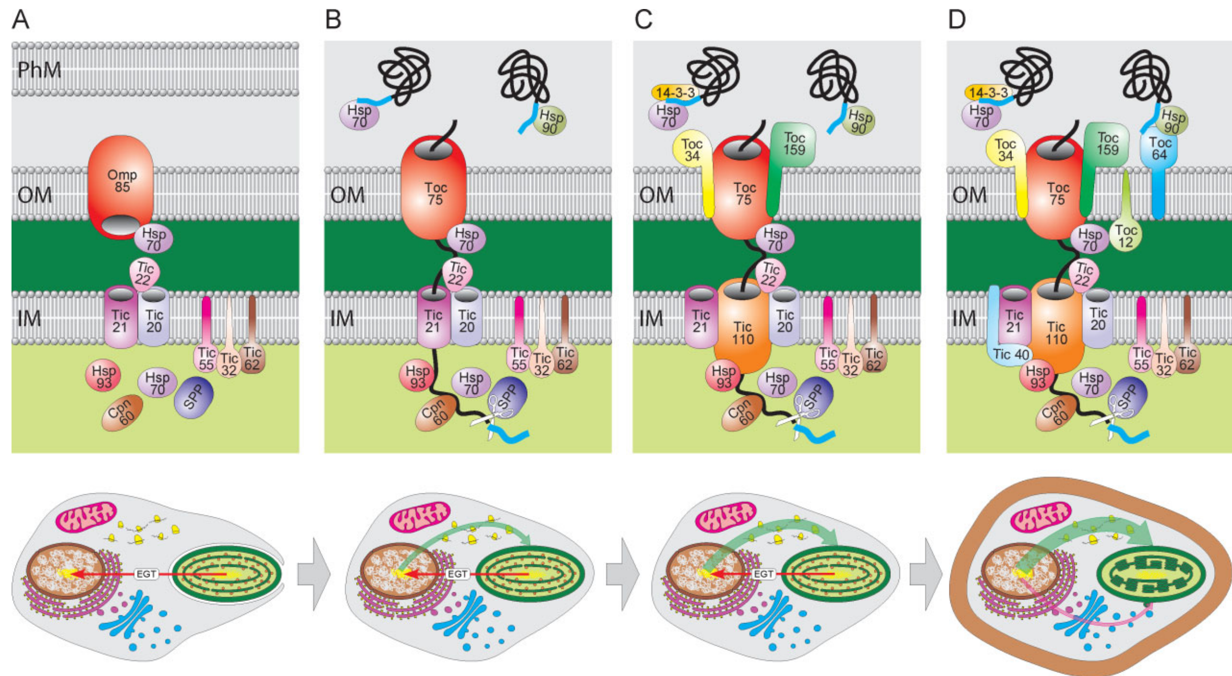

Figure 6. Hypothetical stages in the evolution of the Toc–Tic-based import system of primary plastids. (**A**) Cyanobacteria were regularly captured as food and placed in phagosomes where they were digested.[4,12] Thus, their cells were initially surrounded by three membranes: the phagosomal memebrane (PhM) derived from the host's EM system, the outer membrane (OM) and the inner membrane (IM) of the cyanobacterium. After the digestion in phagosomes, numerous cyanobacterial genes migrated to the nucleus of the eukaryotic host *via* endosymbiotic gene transfer (EGT).[20–22] These events, happening before the cyanobacterium was established as a stable endosymbiont, significantly improved its further transformation to a primary plastid. The engulfed cyanobacteria already possessed many proteins that were ancestors of present components of Toc and Tic translocons.[27,29,30,57] Therefore, import based on these translocons could evolve very quickly. (**B**) After adaptation of the endosymbiont to the host environment, the phagosomal membrane was lost very early. The previous EGT of a gene coding for Omp85 (finally transformed to Toc75) enabled insertion of this protein in the reverse orientation into the outer endosymbiont membrane[46,53] and translocation of nuclear-encoded plastid proteins into the intermembrane space. Because the inner cyanobacterial membrane contained homologs of Tic20 and Tic21, which probably were pre-adapted to translocate proteins encoded by the host nucleus,[54,56] the imported proteins could cross the inner envelope membrane and reach the stroma (green arrow). (**C**) In a next evolutionary stage, two regulatory receptor proteins, Toc34 and Toc159, were added to Toc75, creating the Toc translocon.[27,29,30] Moreover, Tic110 was inserted into the inner membrane, becoming the main translocation pore.[27,29,30] In contrast to Tic20 and Tic21, it possessed a large stroma-exposed domain providing a binding site for Tic40 and Hsp93. Inclusion of new proteins in the Toc–Tic super-complex resulted in more effective import of proteins encoded by cyanobacterial genes transferred to the host's nucleus (indicated by a wider green arrow). The organization of Toc and Tic translocons shown assumes content of proteins presumed for the common ancestor of red and green lineages. (**D**) Additional components of Toc and Tic translocons, such as Toc64, Toc12, and Tic40, could have been added in higher plants, further improving protein recognition on the plastid surface and translocation across the envelope membranes.[27,29,30] These changes were accompanied by the evolution of two "guidance" complexes in the cytosol: one composed of Hsp70 and 14-3-3 proteins[43] and the other involving only Hsp90 proteins.[44] Finally, and apparently only in the green plant lineage, an alternative import of some proteins (*e.g.*, α-carbonic anhydrase[31]) *via* EM system evolved (pink arrow).

was eliminated for unknown reasons. Second, if primary plastids had been surrounded by three membranes well into their evolution, their nuclear-encoded proteins would have carried signal peptides targeting them to the plastid *via* the EM system (Fig. 5A), as currently found in eukaryotic alga-derived plastids (see Fig. 4 and the previous discussion). After the phagosomal membrane was lost, however, nuclear-encoded plastid proteins were still packaged in Golgi-derived pre-plastid vesicles that could not fuse with the cyanobacterial outer membrane now exposed to the cytosol (Fig. 5A). This would have prevented plastid biogenesis and there are no obvious or even proposed selective advantages to explain such traumatic changes.

A much more parsimonious hypothesis is that loss of the phagosomal membrane happened very early in the evolution of primary plastids, *i.e.*, before establishment of any import system for nuclear-encoded proteins (Fig. 6). There are at least three clear advantages to this hypothesis. First, transport systems for diverse compounds would have had to evolve (or be elaborated) in only two envelope membranes rather than three, and both could be adapted from pre-existing endosymbiont transporters. Second, intracellular cyanobacteria devoid of phagosomal membranes could have effectively divided in the host cytosol (their plasmalemma and outer membrane already possessed abilities to both permanently grow and divide[112]), encouraging their stability as permanent

endosymbionts. Third, loss of the phagosome would have permitted direct import of many cyanobacterial proteins that were already encoded by the host nucleus.

This last point is frequently overlooked in evolutionary scenarios for the origin of plastids. Because ancestors of algae with primary plastids were phagotrophic protozoans, their nuclei undoubtedly contained numerous horizontally transferred cyanobacterial genes, even before establishment of a permanent endosymbiosis[113] (Fig. 6). This could have provided a direct selective advantage for losing the phagosomal membrane. It would have allowed outer plastid membrane-residing proteins (*e.g.*, OEP7, OEP24, OEP37[114]) to gain direct access to primary plastids; it is known that most of these proteins can insert spontaneously into the membrane without help from Toc75 or other transport proteins.[41,115] At the same time, or soon thereafter, Omp85/Toc75 could have been inserted into the outer membrane from the host cytosol in a reverse orientation[46,53] (Fig. 6), facilitating additional protein insertions into the outer membrane, and enabling import of proteins into the intermembrane space, the inner membrane, and probably the stroma.

It has been shown that as many as 5% of proteins encoded by bacterial genomes contain cryptic transit peptides[116] (see also Refs.[117,118]). Moreover, even cyanobacterial genes devoid of such signals could acquire them quickly, either from random sequences dispersed in the host genome[117] or from pre-existing mitochondrial transit sequences.[119,120] If such genes already resided in the host nucleus, their encoded proteins could be imported into primary plastids immediately after insertion of Omp85/Toc75, because cyanobacterial homologs themselves contain a receptor domain (see the previous discussion). Thus, there were far fewer obstacles to establishing import of proteins with typical transit peptides very early in primary plastid evolution.

It is evident that, after the early loss of the phagosome, a direct origin of the Toc–Tic-based import system (see Fig. 6 for details) would have been simpler than evolving ER–Golgi-mediated targeting.[34] Accepting the less parsimonious ''early EM trafficking'' scenario requires postulating (i) acquisition of signal peptides by hundreds of plastid proteins, (ii) origin of a specific class of Golgi-derived vesicles to deliver proteins to the plastid, and (iii) evolution of a Golgi sorting signal in each of numerous plastid proteins. For some reason, this highly derived targeting and sorting system then would have been lost for almost all plastid proteins, in favor of Toc-based translocation (Fig. 5A). During this process, each individual plastid signal peptide had to be modified into a plastid transit peptide or perhaps replaced by an existing mitochondrial transit peptide. Such an incipient post-translational import mechanism would almost certainly have been less efficient than the original, well-adapted co-translational system (Fig. 5A). Therefore, these changes would have been purged by selection without some alternative and unexplained advantage. Finally, if initial protein targeting through the outer primary plastid membrane was *via* the EM system, a mechanism must be advanced to explain how so many Toc-encoding genes were retained during this prolonged period of EM trafficking. What was their selective advantage and why were they conserved so strongly as to retain transport function until they once again were needed, far into the future of the plastid endosymbiosis?

Conclusion

The scenario we propose is consistent with the fact that almost all plastid proteins are targeted post-translationally and carry easily recognizable N-terminal transit peptides[24–30] (Fig. 1). This is far more parsimoniously explained as the ancestral rather than derived state. It is possible that Golgi-derived vesicles could fuse with the outer membrane in ancient plastids, given its chimeric nature, enabling some early exploitation of this pathway. However, based on its identification to date only in higher plants,[31,33,68] and the absence of a viable explanation of how selection could have favored its replacement by a post-translational machinery, ER–Golgi-based transport is implausible as the ancestral mechanism that permitted wholesale movement of endosymbiont genes into the nucleus. It is much more likely to be a derived pathway that evolved to handle Golgi-modified[31,33] or dually targeted[68] proteins that were later refinements of the integration of plastids into total cellular function.

Beyond plastids and mitochondria, there are many known endosymbionts in eukaryotes;[121–126] they range from transient associations[121] to highly reduced, organelle-like entities with genomes comparable in size to those of plastids.[126] Such independent endosymbioses offer insights into how host cell–endosymbiont communication develops. For example, the 160-kb genome of *Carsonella ruddii*, a proteobacterial primary endosymbiont of the psyllid insect *Pachpsylla venusta*, encodes only 182 predicted genes and none for DNA replication and other vital informational processes.[126] Given the difficulty of transferring numerous essential genes to the germ cell line in animals, it is hypothesized that *C. ruddii* endosymbionts import mitochondrial proteins encoded by the host nucleus.[127] The mitochondrion almost certainly predated any primary plastid endosymbiosis[58,128] and many modern proteins exhibit dual targeting to plastids and mitochondria.[129–131] Thus, *Carsonella* provides a viable model for how protein import could have first developed in primary plastids, and clearly supports a post-translational over an EM-based system.

One additional endosymbiosis bears special consideration in light of our major arguments: an independent lineage of primary plastids is found in the filose, thecate amoeba

Paulinella chromatophora.[120,132] The completely sequenced genome of *Paulinella* plastids has undergone a drastic reduction, losing approximately 75% of its original gene complement, including many with essential functions.[133] A search of the complete sequence of the *Paulinella* plastid genome yields no homologs of Omp85/Toc75, Tic20, Tic22, and Tic55, which suggests they could have been transferred to the host nucleus.[57]

Indeed, Nakayama and Ishida[134] found that the *psaE* gene (encoding the photosystem I reaction center subunit IV) now resides in the *Paulinella* host nucleus. Interestingly, *psaE* does not appear to encode any N-terminal targeting signal. This suggests that its protein product is post-translationally imported into *Paulinella* plastids using an uncharacterized targeting signal, perhaps one localized in the mature protein as discussed previously for the higher plant ceQORH protein. It also is possible that the *psaE* gene contains an unidentified intron and additional open reading frame (ORF) for a targeting signal.[134] If this hypothetical ORF encoded a signal peptide, it would indicate that a primitive EM-based transport has evolved in *P. chromatophora*. If so, our evolutionary scenario predicts that endosymbiont *toc*-like genes have been lost completely and should not be present in the nuclear genome. Conversely, if a transit peptide eventually is found, we should expect to find a Toc–Tic-based transport system in *Paulinella* plastids. In that case, essential *toc* and *tic* genes missing from the endosymbiont should be encoded in the *Paulinella* nucleus and the EM transport should be absent at this stage of plastid evolution. Further analysis of the *Paulinella* endosymbiosis can provide direct empirical tests of the selective arguments that underlie our hypothesis of the early evolution of primary plastid protein import.

Acknowledgments: We are very grateful to two anonymous reviewers for their excellent comments and suggestions. A. B. is supported by funds from Wrocław University, grant BS/1018/2009.

References

1. **Schimper AFW.** 1883. Über die entwicklung der chlorophyllkörner und farbkörper. *Bot Zeitung* **41**: 105–14, 121–31, 137–46, 153–62.
2. **Mereschkowsky C.** 1905. Über natur and ursprung der chromatophoren im pflanzenreiche. *Biol Zentralbl* **25**: 593–604.
3. **Margulis L.** 1970. Origin of Eukaryotic Cells. New Haven: Yale University Press.
4. **Cavalier-Smith T.** 2000. Membrane heredity and early chloroplast evolution. *Trends Plant Sci* **5**: 174–82.
5. **Palmer JD.** 2003. The symbiotic birth and spread of plastids: how many times and whodunit? *J Phycol* **39**: 4–11.
6. **Archibald JM.** 2009. The puzzle of plastid evolution. *Curr Biol* **19**: R81–8.
7. **Rodríguez-Ezpeleta N, Brinkmann H, Burey SC,** et al. 2005. Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. *Curr Biol* **15**: 1325–30.
8. **Reyes-Prieto A, Bhattacharya D.** 2007. Phylogeny of nuclear-encoded plastid-targeted proteins supports an early divergence of glaucophytes within Plantae. *Mol Biol Evol* **24**: 2358–61.
9. **Howe CJ, Barbrook AC, Nisbet RE,** et al. 2008. The origin of plastids. *Philos Trans R Soc Lond B Biol Sci* **363**: 2675–85.
10. **Stiller JW.** 2007. Plastid endosymbiosis, genome evolution and the origin of green plants. *Trends Plant Sci* **12**: 391–6.
11. **Nozaki H, Iseki M, Hasegawa M,** et al. 2007. Phylogeny of primary photosynthetic eukaryotes as deduced from slowly evolving nuclear genes. *Mol Biol Evol* **24**: 1592–5.
12. **Cavalier-Smith T.** 1982. The origins of plastids. *Biol J Linn Soc* **17**: 289–306.
13. **Weeden NF.** 1981. Genetic and biochemical implications of the endosymbiotic origin of the chloroplast. *J Mol Evol* **17**: 133–9.
14. **Tetlow IJ, Rawsthorne S, Raines C,** et al. 2005. Plastid metabolic pathways. In Møller SG. ed; Plastids (Annual Plant Reviews 13). Oxford: Blackwell Publishing. p 60–125.
15. **Borza T, Popescu CE, Lee RW.** 2005. Multiple metabolic roles for the nonphotosynthetic plastid of the green alga *Prototheca wickerhamii*. *Eukaryot Cell* **4**: 253–61.
16. **Kneip C, Lockhart P, Voss C,** et al. 2007. Nitrogen fixation in eukaryotes – new models for symbiosis. *BMC Evol Biol* **7**: 55.
17. **Deusch O, Landan G, Roettger M,** et al. 2008. Genes of cyanobacterial origin in plant nuclear genomes point to a heterocyst-forming plastid ancestor. *Mol Biol Evol* **25**: 748–61.
18. **Cavalier-Smith T, Lee JJ.** 1985. Protozoa as hosts for endosymbioses and the conversion of symbionts into organelles. *J Protozool* **32**: 376–9.
19. **Richly E, Leister D.** 2004. An improved prediction of chloroplast proteins reveals diversities and commonalities in the chloroplast proteomes of *Arabidopsis* and rice. *Gene* **329**: 11–6.
20. **Timmis JN, Ayliffe MA, Huang CY,** et al. 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat Rev Genet* **5**: 123–35.
21. **Bock R, Timmis JN.** 2008. Reconstructing evolution: gene transfer from plastids to the nucleus. *BioEssays* **30**: 556–66.
22. **Kleine T, Maier UG, Leister D.** 2008. DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. *Annu Rev Plant Biol* **60**: 115–38.
23. **Martin W, Rujan T, Richly E,** et al. 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**: 12246–51.
24. **Bruce BD.** 2001. The paradox of plastid transit peptides: conservation of function despite divergence in primary structure. *Biochim Biophys Acta* **1541**: 2–21.
25. **Patron NJ, Waller RF.** 2007. Transit peptide diversity and divergence: a global analysis of plastid targeting signals. *BioEssays* **29**: 1048–58.
26. **Soll J, Schleiff E.** 2004. Protein import into chloroplasts. *Nat Rev Mol Cell Biol* **5**: 198–208.
27. **Reumann S, Inoue K, Keegstra K.** 2005. Evolution of the general protein import pathway of plastids (review). *Mol Membr Biol* **22**: 73–86.
28. **Inaba T, Schnell DJ.** 2008. Protein trafficking to plastids: one theme, many variations. *Biochem J* **413**: 15–28.
29. **Gross J, Bhattacharya D.** 2009. Revaluating the evolution of the Toc and Tic protein translocons. *Trends Plant Sci* **14**: 13–20.
30. **Kalanon M, McFadden GI.** 2008. The chloroplast protein translocation complexes of *Chlamydomonas reinhardtii*: a bioinformatic comparison of Toc and Tic components in plants, green algae and red algae. *Genetics* **179**: 95–112.
31. **Villarejo A, Buren S, Larsson S,** et al. 2005. Evidence for a protein transported through the secretory pathway *en route* to the higher plant chloroplast. *Nat Cell Biol* **7**: 1224–31.
32. **Ritzenthaler C, Nebenführ A, Movafeghi A,** et al. 2002. Reevaluation of the effects of brefeldin A on plant cells using tobacco Bright Yellow 2 cells expressing Golgi-targeted green fluorescent protein and COPI antisera. *Plant Cell* **14**: 237–61.
33. **Nanjo Y, Oka H, Ikarashi N,** et al. 2006. Rice plastidial N-glycosylated nucleotide pyrophosphatase/phosphodiesterase is transported from the ER-Golgi to the chloroplast through the secretory pathway. *Plant Cell* **18**: 2582–92.

34. **Bhattacharya D, Archibald JM, Weber AP,** et al. 2007. How do endosymbionts become organelles? Understanding early events in plastid evolution. *BioEssays* **29**: 1239–46.
35. **Radhamony RN, Theg SM.** 2006. Evidence for an ER to Golgi to chloroplast protein transport pathway. *Trends Cell Biol* **16**: 385–7.
36. **Tu SL, Chen LJ, Smith MD,** et al. 2004. Import pathways of chloroplast interior proteins and the outer-membrane protein OEP14 converge at Toc75. *Plant Cell* **16**: 2078–88.
37. **Vojta L, Soll J, Bölter B.** 2007. Protein transport in chloroplasts – targeting to the intermembrane space. *FEBS J* **274**: 5043–54.
38. **Firlej-Kwoka E, Strittmatter P, Soll J,** et al. 2008. Import of preproteins into the chloroplast inner envelope membrane. *Plant Mol Biol* **68**: 505–19.
39. **Schünemann D.** 2007. Mechanisms of protein import into thylakoids of chloroplasts. *Biol Chem* **388**: 907–15.
40. **Inoue K, Potter D.** 2004. The chloroplastic protein translocation channel Toc75 and its paralog OEP80 represent two distinct protein families and are targeted to the chloroplastic outer envelope by different mechanisms. *Plant J* **39**: 354–65.
41. **Hofmann NR, Theg SM.** 2005. Chloroplast outer membrane protein targeting and insertion. *Trends Plant Sci* **10**: 450–7.
42. **Reinbothe S, Quigley F, Springer A,** et al. 2004. The outer plastid envelope protein Oep16: role as precursor translocase in import of protochlorophyllide oxidoreductase A. *Proc Natl Acad Sci USA* **101**: 2203–8.
43. **May T, Soll J.** 2000. 14-3-3 proteins form a guidance complex with chloroplast precursor proteins in plants. *Plant Cell* **12**: 53–64.
44. **Qbadou S, Becker T, Mirus O,** et al. 2006. The molecular chaperone Hsp90 delivers precursor proteins to the chloroplast import receptor Toc64. *EMBO J* **25**: 1836–47.
45. **Duy D, Wanner G, Meda AR,** et al. 2007. PIC1, an ancient permease in *Arabidopsis* chloroplasts, mediates iron transport. *Plant Cell* **19**: 986–1006.
46. **Steiner JM, Yusa F, Pompe JA,** et al. 2005. Homologous protein import machineries in chloroplasts and cyanelles. *Plant J* **44**: 646–52.
47. **Ertel F, Mirus O, Bredemeier R,** et al. 2005. The evolutionarily related beta-barrel polypeptide transporters from *Pisum sativum* and *Nostoc* PCC7120 contain two distinct functional domains. *J Biol Chem* **280**: 28281–9.
48. **Chen K, Chen X, Schnell DJ.** 2000. Initial binding of preproteins involving the Toc159 receptor can be bypassed during protein import into chloroplasts. *Plant Physiol* **122**: 813–22.
49. **Hust B, Gutensohn M.** 2006. Deletion of core components of the plastid protein import machinery causes differential arrest of embryo development in *Arabidopsis thaliana*. *Plant Biol* **8**: 18–30.
50. **Reumann S, Davila-Aponte J, Keegstra K.** 1999. The evolutionary origin of the protein-translocating channel of chloroplastic envelope membranes: identification of a cyanobacterial homolog. *Proc Natl Acad Sci USA* **96**: 784–9.
51. **Voulhoux R, Tommassen J.** 2004. Omp85, an evolutionarily conserved bacterial protein involved in outer-membrane-protein assembly. *Res Microbiol* **155**: 129–35.
52. **Gentle IE, Burri L, Lithgow T.** 2005. Molecular architecture and function of the Omp85 family of proteins. *Mol Microbiol* **58**: 1216–25.
53. **McFadden GI.** 1999. Endosymbiosis and evolution of the plant cell. *Curr Opin Plant Biol* **2**: 513–9.
54. **Chen X, Smith MD, Fitzpatrick L,** et al. 2002. *In vivo* analysis of the role of atTic20 in protein import into chloroplasts. *Plant Cell* **14**: 641–54.
55. **Rassow J, Dekker PJ, van Wilpe S,** et al. 1999. The preprotein translocase of the mitochondrial inner membrane: function and evolution. *J Mol Biol* **286**: 105–20.
56. **Teng YS, Su YS, Chen LJ,** et al. 2006. Tic21 is an essential translocon component for protein translocation across the chloroplast inner envelope membrane. *Plant Cell* **18**: 2247–57.
57. **Bodył A, Mackiewicz P, Stiller JW.** 2009. Comparative genomic studies suggest that the cyanobacterial endosymbionts of the amoeba *Paulinella chromatophora* possess an import apparatus for nuclear-encoded proteins. *Plant Biol* (in press).
58. **Cavalier-Smith T.** 2006. Origin of mitochondria by intracellular enslavement of a photosynthetic purple bacterium. *Proc Biol Sci* **273**: 1943–52.
59. **Schleiff E, Jelic M, Soll J.** 2003. A GTP-driven motor moves proteins across the outer envelope of chloroplasts. *Proc Natl Acad Sci USA* **100**: 4604–9.
60. **Bauer J, Chen K, Hiltbunner A,** et al. 2000. The major protein import receptor of plastids is essential for chloroplast biogenesis. *Nature* **403**: 203–7.
61. **Lee KH, Kim SJ, Lee YJ,** et al. 2003. The M domain of atToc159 plays an essential role in the import of proteins into chloroplasts and chloroplast biogenesis. *J Biol Chem* **278**: 36794–805.
62. **Ivanova Y, Smith MD, Chen K,** et al. 2004. Members of the Toc159 import receptor family represent distinct pathways for protein targeting to plastids. *Mol Biol Cell* **15**: 3379–92.
63. **Kubis S, Patel R, Combe J,** et al. 2004. Functional specialization amongst the *Arabidopsis* Toc159 family of chloroplast protein import receptors. *Plant Cell* **16**: 2059–77.
64. **Hiltbrunner A, Grünig K, Alvarez-Huerta M,** et al. 2004. AtToc90, a new GTP-binding component of the *Arabidopsis* chloroplast protein import machinery. *Plant Mol Biol* **54**: 427–40.
65. **Kleffmann T, Russenberger D, von Zychlinski A,** et al. 2004. The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions. *Curr Biol* **14**: 354–62.
66. **Jarvis P.** 2004. Organellar proteomics: chloroplasts in the spotlight. *Curr Biol* **14**: R317–9.
67. **Zybailov B, Rutschow H, Friso G,** et al. 2008. Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PLoS ONE* **3**: e1994.
68. **Chen MH, Huang LF, Li HM,** et al. 2004. Signal peptide-dependent targeting of a rice alpha-amylase and cargo proteins to plastids and extracellular compartments of plant cells. *Plant Physiol* **135**: 1367–77.
69. **Lerouge P, Cabanes-Macheteau M, Rayon C,** et al. 1998. N-glycoprotein biosynthesis in plants: recent developments and future trends. *Plant Mol Biol* **38**: 31–48.
70. **Moroney JV, Bartlett SG, Samuelsson G.** 2001. Carbonic anhydrases in plants and algae. *Plant Cell Environ* **24**: 141–53.
71. **Carter CJ, Thornburg RW.** 2004. Tobacco Nectarin III is a bifunctional enzyme with monodehydroascorbate reductase and carbonic anhydrase activities. *Plant Mol Biol* **54**: 415–25.
72. **Conlan S, Griffiths LA, Turner M,** et al. 1998. Characterisation of the yam tuber storage protein dioscorin. *J Plant Physiol* **153**: 25–31.
73. **Fukuzawa H, Fujiwara S, Yamamoto Y,** et al. 1990. cDNA cloning, sequence, and expression of carbonic anhydrase in *Chlamydomonas reinhardtii*: regulation by environmental CO_2 concentration. *Proc Natl Acad Sci USA* **87**: 4383–7.
74. **Satoh A, Iwasaki T, Odani S,** et al. 1998. Purification, characterization and cDNA cloning of soluble carbonic anhydrase from *Chlorella sorokiniana* grown under ordinary air. *Planta* **206**: 657–65.
75. **Fisher M, Gokhman I, Pick U,** et al. 1996. A salt-resistant plasma membrane carbonic anhydrase is induced by salt in *Dunaliella salina*. *J Biol Chem* **271**: 17718–23.
76. **Karlsson J, Clarke AK, Chen Z-Y,** et al. 1998. A novel alpha-type carbonic anhydrase associated with the thylakoid membrane in *Chlamydomonas reinhardtii* is required for growth at ambient CO_2. *EMBO J* **17**: 1208–16.
77. **Weber AP, Linka M, Bhattacharya D.** 2006. Single, ancient origin of a plastid metabolite translocator family in Plantae from an endomembrane-derived ancestor. *Eukaryot Cell* **5**: 609–12.
78. **Weber AP.** 2004. Solute transporters as connecting elements between cytosol and plastid stroma. *Curr Opin Plant Biol* **7**: 247–53.
79. **Weber AP, Fischer K.** 2007. Making the connections – the crucial role of metabolite transporters at the interface between chloroplast and cytosol. *FEBS Lett* **581**: 2215–22.
80. **Matsuzaki M, Misumi O, Shin-I T,** et al. 2004. Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature* **428**: 653–7.
81. **Yoon HS, Müller KM, Sheath RG,** et al. 2006. Defining the major lineages of red algae (Rhodophyta). *J Phycol* **42**: 482–92.
82. **Kroth PG.** 2002. Protein transport into secondary plastids and the evolution of primary and secondary plastids. *Int Rev Cytol* **221**: 191–255.
83. **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 Lond B Biol Sci* **358**: 109–33.
84. **Bolte K, Bullmann L, Hempel F,** et al. 2009. Protein targeting into secondary plastids. *J Eukaryot Microbiol* **56**: 9–15.
85. **Sulli C, Fang Z, Muchhal U,** et al. 1999. Topology of *Euglena* chloroplast protein precursors within endoplasmic reticulum to Golgi to chloroplast transport vesicles. *J Biol Chem* **27**: 457–63.
86. **Wastl J, Maier UG.** 2000. Transport of proteins into cryptomonads complex plastids. *J Biol Chem* **275**: 23194–8.
87. **Bodył A, Moszczyński K.** 2006. Did the peridinin plastid evolve through tertiary endosymbiosis? A hypothesis. *Eur J Phycol* **41**: 435–48.
88. **Cavalier-Smith T.** 1993. The origin, losses and gains of chloroplasts. In Lewin RA. ed; Origins of Plastids: Symbiogenesis, Prochlorophytes, and the Origins of Chloroplasts. New York and London: Chapman & Hall. p 291–348.
89. **Rogers MB, Gilson PR, Su V,** et al. 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.
90. **Patron NJ, Waller RF, Archibald JM,** et al. 2005. Complex protein targeting to dinoflagellate plastids. *J Mol Biol* **348**: 1015–24.
91. **Kilian O, Kroth PG.** 2003. Evolution of protein targeting into ''complex'' plastids: the ''secretory transport hypothesis''. *Plant Biol* **5**: 350–8.
92. **Bodył A.** 2004. Evolutionary origin of a preprotein translocase in the periplastid membrane of complex plastids: a hypothesis. *Plant Biol* **6**: 513–8.
93. **Sommer MS, Gould SB, Lehmann P,** et al. 2007. Der1-mediated preprotein import into the periplastid compartment of chromalveolates? *Mol Biol Evol* **24**: 918–28.
94. **Spork S, Hiss JA, Mandel K,** et al. 2009. An unusual ERAD-like complex is targeted to the apicoplast of *Plasmodium falciparum*. *Eukaryot Cell* **8**: 1134–45.
95. **Ismail N, Ng DT.** 2006. Have you HRD? Understanding ERAD is DOAble. *Cell* **126**: 237–9.
96. **Gibbs SP.** 1979. The route of entry of cytoplasmically synthesized proteins into chloroplasts of algae possessing chloroplast ER. *J Cell Sci* **35**: 253–66.
97. **Jena BP.** 2008. Assembly and disassembly of SNAREs in membrane fusion. *Methods Cell Biol* **90**: 157–82.
98. **Nada A, Soll J.** 2004. Inner envelope protein 32 is imported into chloroplasts by a novel pathway. *J Cell Sci* **117**: 3975–82.
99. **Miras S, Salvi D, Ferro M,** et al. 2002. Non-canonical transit peptide for import into the chloroplast. *J Biol Chem* **277**: 47770–8.
100. **Miras S, Salvi D, Piette L,** et al. 2007. Toc159- and Toc75-independent import of a transit sequence-less precursor into the inner envelope of chloroplasts. *J Biol Chem* **282**: 29482–92.
101. **Sherman EL, Taylor RD, Go NE,** et al. 2006. Effect of mutations in Tom40 on stability of the translocase of the outer mitochondrial membrane (TOM) complex, assembly of Tom40, and import of mitochondrial preproteins. *J Biol Chem* **281**: 22554–65.
102. **Löffelhardt H, von Haeseler A, Schleiff E.** 2007. The β-barrel shaped polypeptide transporter, an old concept for precursor protein transfer across membranes. *Symbiosis* **44**: 33–42.
103. **Neupert W, Herrmann JM.** 2007. Translocation of proteins into mitochondria. *Annu Rev Biochem* **76**: 723–49.
104. **Pohlmeyer K, Soll J, Steinkamp T,** et al. 1997. Isolation and characterization of an amino acid-selective channel protein present in the chloroplastic outer envelope membrane. *Proc Natl Acad Sci USA* **94**: 9504–9.
105. **Pollmann S, Springer A, Buhr F,** et al. 2007. A plant porphyria related to defects in plastid import of protochlorophyllide oxidoreductase A. *Proc Natl Acad Sci USA* **104**: 2019–23.
106. **Jarvis P, Soll J.** 2001. Toc, Tic, and chloroplast protein import. *Biochim Biophys Acta* **1541**: 64–79.
107. **Philippar K, Geis T, Ilkavets I,** et al. 2007. Chloroplast biogenesis: the use of mutants to study the etioplast-chloroplast transition. *Proc Natl Acad Sci USA* **104**: 678–83.
108. **Cavalier-Smith T.** 2006. Rooting the tree of life by transition analyses. *Biol Direct* **1**: 19.
109. **Inoue K.** 2007. The chloroplast outer envelope membrane: the edge of light and excitement. *J Int Plant Biol* **49**: 1100–11.
110. **Pfanzagl B, Löffelhardt W.** 1999. *In vitro* synthesis of peptidoglycan precursors modified with N-acetylputrescine by *Cyanophora paradoxa* cyanelle envelope membranes. *J Bacteriol* **181**: 2643–7.
111. **Koike H, Ikeda Y, Yusa F,** et al. 2007. Isolation and characterization of outer and inner envelope membranes of cyanelles from a glaucocystophyte, *Cyanophora paradoxa*. *Photosynth Res* **93**: 45–53.
112. **Cavalier-Smith T.** 1987. The simultaneous symbiotic origin of mitochondria, chloroplasts, and microbodies. *Ann N Y Acad Sci* **503**: 55–71.
113. **Larkum AW, Lockhart PJ, Howe CJ.** 2007. Shopping for plastids. *Trends Plant Sci* **12**: 189–95.
114. **Duy D, Soll J, Philippar K.** 2007. Solute channels of the outer membrane: from bacteria to chloroplasts. *Biol Chem* **388**: 879–89.
115. **Schleiff E, Klösgen RB.** 2001. Without a little help from 'my' friends: direct insertion of proteins into chloroplast membranes? *Biochim Biophys Acta* **1541**: 22–33.
116. **Lucattini R, Likic VA, Lithgow T.** 2004. Bacterial proteins predisposed for targeting to mitochondria. *Mol Biol Evol* **21**: 652–8.
117. **Baker A, Schatz G.** 1987. Sequences from a prokaryotic genome or the mouse dihydrofolate reductase gene can restore the import of a truncated precursor protein into yeast mitochondria. *Proc Natl Acad Sci USA* **84**: 3117–21.
118. **Ueda M, Fujimoto M, Arimura S,** et al. 2008. Presence of a latent mitochondrial targeting signal in gene on mitochondrial genome. *Mol Biol Evol* **25**: 1791–3.
119. **Silva-Filho MD, Wieërs MC, Flügge UI,** et al. 1997. Different *in vitro* and *in vivo* targeting properties of the transit peptide of a chloroplast envelope inner membrane protein. *J Biol Chem* **272**: 15264–9.
120. **Bodył A, Mackiewicz P, Stiller JW.** 2007. The intracellular cyanobacteria of *Paulinella chromatophora*: endosymbionts or organelles? *Trends Microbiol* **15**: 295–6.
121. **Thancker RW.** 2005. Impacts of shading on sponge-cyanobacteria symbioses: a comparison between host-specific and generalist associations. *Integr Comp Biol* **45**: 369–76.
122. **Rumpho ME, Worful JM, Lee J,** et al. 2008. Horizontal gene transfer of the algal nuclear gene *psbO* to the photosynthetic sea slug *Elysia chlorotica*. *Proc Natl Acad Sci USA* **105**: 17867–71.
123. **Kodama Y, Fujishima M.** 2009. Localization of perialgal vacuoles beneath the host cell surface is not a prerequisite phenomenon for protection from the host's lysosomal fusion in the ciliate *Paramecium bursaria*. *Protist* **160**: 319–29.
124. **Kneip C, Voss C, Lockhart PJ,** et al. 2008. The cyanobacterial endosymbiont of the unicellular algae *Rhopalodia gibba* shows reductive genome evolution. *BMC Evol Biol* **8**: 30.
125. **Nikoh N, Nakabachi A.** 2009. Aphids acquired symbiotic genes via lateral gene transfer. *BMC Biol* **7**: 12.
126. **Nakabachi A, Yamashita A, Toh H,** et al. 2006. The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. *Science* **314**: 267.
127. **Tamames J, Gil R, Latorre A,** et al. 2007. The frontier between cell and organelle: genome analysis of *Candidatus* Carsonella ruddii. *BMC Evol Biol* **7**: 181.
128. **Martin W, Müller M.** 1998. The hydrogen hypothesis for the first eukaryote. *Nature* **392**: 37–41.
129. **Peeters N, Small I.** 2001. Dual targeting to mitochondria and chloroplasts. *Biochim Biophys Acta* **1541**: 54–63.
130. **Duchêne AM, Giritch A, Hoffmann B,** et al. 2005. Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* **102**: 16484–9.
131. **Pujol C, Maréchal-Drouard L, Duchêne AM.** 2007. How can organellar protein N-terminal sequences be dual targeting signals? *In silico* analysis and mutagenesis approach. *J Mol Biol* **369**: 356–67.
132. **Archibald JM.** 2006. Endosymbiosis: double-take on plastid origins. *Curr Biol* **16**: R690–2.
133. **Nowack EC, Melkonian M, Glöckner G.** 2008. Chromatophore genome sequence of *Paulinella* sheds light on acquisition of photosynthesis by eukaryotes. *Curr Biol* **18**: 410–8.
134. **Nakayama T, Ishida K.** 2009. Another acquisition of a primary photosynthetic organelle is underway in *Paulinella chromatophora*. *Curr Biol* **19**: R284–5.