Review

Wai-Ling Chang, Jürgen Soll and Bettina Bölter*

The gateway to chloroplast: re-defining the function of chloroplast receptor proteins

Abstract: Chloroplast biogenesis often requires a tight orchestration between gene expression (both plastidial and nuclear) and translocation of ~3000 nuclear-encoded proteins into the organelle. Protein translocation is achieved via two multimeric import machineries at the outer (TOC) and inner (TIC) envelope of chloroplast, respectively. Three components constitute the core element of the TOC complex: a β-barrel protein translocation channel Toc75 and two receptor constituents, Toc159 and Toc34. A diverse set of distinct TOC complexes have recently been characterized and these diversified TOC complexes have evolved to coordinate the translocation of differentially expressed proteins. This review aims to describe the recent discoveries relating to the typical characteristics of these distinct TOC complexes, particularly the receptor constituents, which are the main contributors for TOC complex diversification.

Keywords: chloroplast; endosymbiosis; GTPase; protein import; TIC complex; TOC complex.

*Corresponding author: Bettina Bölter, Department Biologie I-Botanik, Ludwig-Maximilians-Universität, Großhadernerstrasse 2–4, D-82152 Planegg-Martinsried, Germany; and Munich Center for Integrated Protein Science (CiPSM), Ludwig-Maximilians-Universität, Feodor-Lynen-Strasse 25, D-81377 Munich, Germany, e-mail: boelter@lrz.uni-muenchen.de

Wai-Ling Chang: Department Biologie I-Botanik, Ludwig-Maximilians-Universität, Großhadernerstrasse 2–4, D-82152 Planegg-Martinsried, Germany; and Munich Center for Integrated Protein Science (CiPSM), Ludwig-Maximilians-Universität, Feodor-Lynen-Strasse 25, D-81377 Munich, Germany

Jürgen Soll: Department Biologie I-Botanik, Ludwig-Maximilians-Universität, Großhadernerstrasse 2–4, D-82152 Planegg-Martinsried, Germany; and Munich Center for Integrated Protein Science (CiPSM), Ludwig-Maximilians-Universität, Feodor-Lynen-Strasse 25, D-81377 Munich, Germany

Introduction

Plastids are a unique set of organelles found in plants with functions covering a wide range of biological processes. Chloroplasts are thus far the best characterized type of plastid. Not only have they served as a platform for oxygenic photosynthesis, a process that is essential to all higher organisms, but they also feature a wondrous array of biosynthetic pathways. These include steps in carbon and nitrogen assimilation, amino acids, lipids, vitamin, hormones and the biosynthesis of secondary metabolites, e.g., terpenoids and porphyrin (Folkes, 1970; Camara, 1984; Browse et al., 1986; Leister, 2003; Takahashi et al., 2005).

The theory on the heredity and evolution of chloroplasts was a topic of major discussion during the late 19th and early 20th century and it has now been widely accepted that chloroplasts – similar to mitochondria – derived from a serial endosymbiosis events; a theory that was first articulated by Mereschkowsky but was later revived and substantiated by Lynn Margulis in the late 1960s (Margulis, 1971; McFadden, 2001). Mitochondria evolved from a single endosymbiosis event that occurred between a host cell and an aerobic α-proteobacterium. Subsequent enslavement of a photosynthetic cyanobacterium by this mitochondriate cell gave rise to the chloroplast, from which a diversity of phototrophic eukaryote lineages emerged (Margulis, 1971). The evolution of chloroplasts occurred at the same time as a progressive loss and massive transfer of the ancient bacterial genes to the host nucleus, rendering the chloroplast a semi-autonomous organelle (Martin and Herrmann, 1998; Timmis et al., 2004). In higher plants, genes of cyanobacterial origin account for only a small percentage of proteins, mainly those involved in translation and photosynthesis; while the majority of the chloroplastic proteins – more than 95% – are encoded in the nucleus of the host cell (Martin and Herrmann, 1998; Martin et al., 2002). Gene relocation from the chloroplast genome typically requires a return ticket for the gene product back to its place of function. A hallmark for this scenario is the development of protein trafficking systems and regulatory networks for the delivery of proteins translated in the cytoplasm back to the compartment of origin within the chloroplast where the proteins perform their function.

Decades of research have started to unravel the riddles of the complex communication networks between...
plastid biogenesis, gene expression/regulation and chloroplast protein import (Pogson et al., 2008; Sakamoto et al., 2008). Plastid biogenesis is a highly intricate process and requires tight coordination of nuclear gene expression and protein import between the host cell and plastids. For example, the development of chloroplasts from undifferentiated proplastids requires a distinct set of photosynthesis-related genes to be expressed at an appropriate level; a process that is triggered by phytochrome, a red-light receptor, in response to light (Terzaghi and Cashmore, 1995), and their subsequent translocation into the plastids. Other plastid types, depending on their metabolic function, require other sets of nuclear-encoded proteins to be assembled into the organelle. Since the plastid protein import apparatus is the point at which most cytosolic proteins enter the organelle, the import mechanism and its constituents (i.e., the import apparatus itself) plays a key regulatory element in bridging the cross-talk between plastids and the nucleus. Indeed the protein import apparatus is indispensable in processes that govern:

- the turnover of nuclear-encoded proteins in the plastids;
- the overall protein profiles during plastid differentiation;
- the status quo of the organelle (i.e., oxidative stress, changes in metabolic requirement, etc.); and
- stoichiometry maintenance of the multimeric complexes in the plastids.

This review will concentrate on several key aspects of the Translocon at the Outer envelope of Chloroplast (TOC) complex that are fundamental to protein trafficking in plastids. Special emphasis will also be given to the emerging role of TOC receptor families in regulating the entry of nuclear-encoded preproteins at the surface of chloroplasts. Due to space limitations, events which take place at the TIC complex are beyond the scope of this review and will only briefly be mentioned. Comprehensive updates of the TIC complex have been discussed in detail in several recent reviews (Benz et al., 2009; Kovacs-Bogdan et al., 2010; Li and Chiu, 2010).

### The molecular framework of the TOC complex

Protein translocation is achieved by means of two multimeric protein complexes (the TOC and TIC complexes) located at the outer (TOC) and inner (TIC) envelopes of chloroplast (Soll, 2002). Among the first components of the chloroplast import machinery to be identified were three components of the outer membrane of pea chloroplasts (Schnell and Blobel, 1993; Waegemann and Soll, 1995). These include a β-barrel membrane channel, Toc75 (Schnell et al., 1994; Keegstra and Cline, 1999), and associated its associated receptor constituents Toc159 and Toc34 (Hirsch et al., 1994; Kessler, Blobel et al., 1994; Seedorf et al., 1995). Both receptors are integral GTPases at the outer membrane. They are unique to plastids, in which they are responsible for the recognition and faithful selection of their characteristic set of nuclear-encoded preproteins at the outer envelopes. Together Toc159, Toc34 and Toc75 form a stable TOC core complex capable and sufficient for preprotein translocation in artificial lipid vesicles in vitro (Schleiff et al., 2003a). Quantitative isolation and detailed structural analysis of the import complexes revealed an ‘oligomeric form’ of the Toc75/Toc34/Toc159 TOC core complex with an apparent molecular mass between 550 kDa and 1 MDa (Schleiff et al., 2003b; Kikuchi et al., 2006). The proteolytically-sensitive Toc159 receptor might possibly account for the drastic size shift of the TOC core complex. Kikuchi and coworkers were able to demonstrate co-migration of intact Toc159 together with the rest of the TOC core components in the blue native polyacrylamide gel electrophoresis isolated 1 MDa complex using size exclusion chromatography. This variation of molecular mass also favors the idea of the existence of multiple copies of the TOC core complex constituents, which is further supported by the reported Toc75:Toc34:Toc159 stoichiometry of 4:4:1 (Schleiff et al., 2003a). The TOC core complex is an approximately circular particle, as revealed by two-dimensional electron microscopy. This particle is enclosed by a dense outer ring with a central ‘finger’ domain that divides the central cavity into four pores. The reported structure of TOC core complex correlates well with the observed stoichiometry. Whether the TOC translocation channel is formed by four copies of Toc75 or a single Toc75 with four exposed soluble domains facing the intermembrane space (IMS) still remains to be determined.

Other auxiliary components of the TOC complex include an accessory receptor, Toc64 and a novel J-domain containing protein Toc12. Toc64 was first reported as a 64 kDa protein, which co-purifies on sucrose-domain gradients with isolated TOC complex from pea chloroplasts (Sohrt and Soll, 2000). It has a short N-terminal hydrophobic transmembrane anchor, a central region with homology to amidases, and three tetratricopeptide repeats (TPRs) at its C-terminus which are exposed to the cytosol. In vitro cross-linking and co-immune precipitation experiments with antibodies against Toc64 revealed...
a close association of the receptor with TOC and TIC components as well as translocation precursor proteins (Sohrt and Soll, 2000; Becker et al., 2004b). It was later revealed that Toc64 only transiently associates with the TOC core complex (Schleiff et al., 2003b) and functions in providing a docking site for Hsp90-affiliated preproteins via its TPR domain (Qbadou et al., 2006), indicating a function in post-translational protein translocation across the chloroplast outer envelopes. TPR domain-containing proteins, which are found in all organisms in cellular compartments are known to contribute to post-translational protein import by the interaction with cytosolic chaperones (Feldheim and Schekman, 1994; Young et al., 2003; Schlegel et al., 2007). Post-translational protein translocation involving TPR-containing proteins and cytosolic chaperones has been well documented for mammalian and yeast receptors Tom20, Tom34 and Tom70 (Abe et al., 2000; Young et al., 2003; Chew et al., 2004; Saitoh et al., 2007; Faou and Hoogenraad, 2012). In plants, apart from Toc64, TPR-mediated preprotein translocation has thus far been reported for the mitochondrial outer envelope protein, OM64 and recently, the endoplasmic reticulum (ER) TPR-docking protein, AtTPR7. Both proteins have been shown to interact with chaperones and preproteins (Chew et al., 2004; Lister et al., 2007; Schweiger et al., 2012). Intriguingly, both OM64 and AtTPR7 can functionally complement yeast mitochondrial Atom70 and Δsec71, respectively, thus hinting at a potential function in post-translational protein translocation across the mitochondria and ER membrane.

In light of the recent data, a clear role of Toc64, particularly in the protein translocation process, could not be concluded since Arabidopsis Toc64 (AtToc64) and Physcomitrella patens Toc64 (PpToc64) loss-of-function mutants do not show deleterious effects on chloroplast biogenesis nor on plant development (Rosenbaum Hofmann and Theg, 2005; Aronsson et al., 2007), indicating that the chaperone recognition process by the receptor can be bypassed. Likewise, neither the loss of the mitochondrial OM64 and ER AtTPR7 results in any visible phenotype (Lister et al., 2004; Schweiger et al., 2012).

Toc12 was initially identified as a novel J domain-containing protein from the pea chloroplast outer membrane proteome and was believed to mediate the transfer of preproteins across the intermembrane space (IMS) of chloroplast together with Toc64 and the imsHsp70 (Becker et al., 2004b; Qbadou et al., 2007). The existence of Toc12 in IMS, however, has been controversially discussed in the recent years (Becker et al., 2004a; Chen et al., 2010; Chiu et al., 2010). Toc12 was annotated as a J-domain protein named AtJ8 in Arabidopsis and as a truncated clone of pea PsJ8b (Chen et al., 2010; Chiu et al., 2010). Based largely on in vitro biochemical data, both works came to the assumption that AtJ8/PsJ8b is a soluble component of the chloroplast stroma as opposed to the previously published IMS localization; however, the experimental setup was not designed to allow differentiation between stroma and IMS. Further, immunodetection using an antibody specific against the DnaJ domain of PsJ8b in pea fail to detect the protein in all tested chloroplast compartments (Chiu et al., 2010). Which leads us to the question whether the in vitro localization data could be substantiated in vivo, at least in the case of pea. AtJ8 knockout mutant plants did not show a severe effect on plant development, but revealed a decreased rate of CO₂ assimilation and activity of Rubisco, which was most likely due to down-regulated Rubisco activase. The stability of PSII seemed to be effected, but further investigation of several photosynthetic parameters disclosed only a slight influence of AtJ8 deletion, implying that AtJ8/PsJ8b/Toc12 might not be involved in preprotein translocation across the IMS. Further possible roles of AtJ8/PsJ8b/Toc12 will have to be clarified in the future.

The completion of the Arabidopsis thaliana genome sequencing project, combined with the establishment of a large insertion mutant library marked a milestone in Arabidopsis functional genomics and reverse genetics. In less than a decade, novel components of the chloroplast protein import machinery in Arabidopsis have been unmasked. A rather complicated picture of the simple version of protein import machinery observed in Pisum sativum (pea) was also revealed.

Most of the TOC components in Arabidopsis, notably Toc34, Toc75 and Toc159, are encoded by gene families rather than unique genes. The Arabidopsis genome encodes two paralogs of Toc34 (atToc33 and atToc34); Guntesohn et al., 2000), and four paralogs of Toc159 (atToc159, atToc132, atToc120 and atToc90; Bauer et al., 2000; Hiltbrunner et al., 2004; Kubis et al., 2004) and Toc75 (atToc75-III, atToc75-IV, atToc75-1 and atToc75V/ atOep80) (Eckart et al., 2002; Baldwin et al., 2005). The discovery of multiple arrays of isoforms in Arabidopsis promises to make the functional characterization of each individual component of the chloroplast import system challenging. In accordance with the diversity of import substrates (preproteins) in higher plants, in current models these different isoforms are involved in the assembly of distinct but structurally similar translocation complexes, each responsible for a varying set of client proteins (Ivanova et al., 2004).
Recognition of preproteins at the chloroplast surface

Preprotein recognition and translocation initiation at the TOC complex are mediated by two receptors, Toc159 and Toc34. Deriving from a common ancestor, Toc34 and its homologs were the first to branch off, and later, after the addition of an extension at the N-terminus, Toc90 and the larger TOC receptors Toc120, Toc132 and Toc159 emerged (Hofmann and Theg, 2003). Interestingly, both Toc159 and Toc34 belong to a distinct plant family of GTPases of eukaryotic origin (Reumann et al., 2005). Alignment-guided secondary structural analysis revealed that the core of the TOC receptor’s G-domain resembles the basic structure of other translation factor-related (TRAFAC) family members, such as Ras GTPases (Aronsson and Jarvis, 2011). This analysis places the TOC receptors specifically in the septin-like superfamily within the TRAFAC class of GTPases (Aronsson and Jarvis, 2011).

In general, members of the Toc159/Toc34 superfamily share a typical domain structure organization (Figure 1). Toc34 is mostly composed of its GTPase-binding domain (the G-domain) and is anchored to the outer envelope of chloroplasts by a short hydrophobic patch at the C-terminus of the protein (Kessler et al., 1994; Seedorf et al., 1995). Toc159, on the other hand, can be subdivided into three functional domains: an acidic amino acid extension at the N-terminus (A-domain), a central GTP-binding domain (G-domain) and a C-terminal membrane anchoring domain of 52 kDa (M-domain) (Muckel and Soll, 1996). Both the A- and G-domain are exposed to the cytosol while the M-domain substantiates the membrane anchoring (Hirsch et al., 1994). While the G-and M-domain between the different homologs exhibit a relatively high homology, their A-domain has fairly low sequence conservation (Ivanova et al., 2004). The N-terminal A-domain is also the most variable region of the Toc159 receptor families both in length and primary structure. This structural remodeling (i.e., domain enlargement and negative charge introduction) suggests functional specialization and optimization, yet the exact functional relevance of the A-domain thus far remains largely unknown. Recently interest in the role of the A-domain has begun to sparkle (see the Construction underway section).

Small TOC GTPases: Arabidopsis Toc34 receptor families

Similar to the homologs of Toc159, the two respective homologs of Toc34 in Arabidopsis (atToc33 and atToc34) also display different developmental expression profiles. atTOC33 is expressed at very high levels in young, rapidly expanding photosynthetic tissues, whereas atTOC34 is expressed at low levels throughout development (Gutensohn et al., 2000; Kubis et al., 2004; Jarvis, 2008). Evidence for the functional specificity of these receptors to the TOC complex has emerged from the analysis of the atToc33 and atToc34 T DNA knockout mutants. The atToc33 knockout mutant (plastid protein import 1: ppi1) appears uniformly pale during the first two weeks of development, suggesting a role in early chloroplast biogenesis, presumably during the expansion of cotyledons (Gutensohn et al., 2000; Bauer et al., 2001). These phenotypic defects, however, were restored to that of the wild-type.
in mature plants (Gutensohn et al., 2000). Similar to the ppi2 mutant, the ppi1 mutant renders a down-regulation of photosynthetic genes but the 'housekeeping' genes are unaffected (Kubis et al., 2004). Isolated of photosynthetic genes but the 'housekeeping' genes are similar to those of wild-type. Despite the apparent effect of this mutation on root development, root plastids appeared to be very similar to those of wild-type Arabidopsis (Constan et al., 2004). It is safe to assume that both atToc33 and atToc34 are functionally equivalent in Arabidopsis, as evidenced by the viability of each respective knockout mutant (Jarvis et al., 1998; Constan et al., 2004). Furthermore, the ectopic expression of both atToc33 and atToc34 under the control of a constitutive promoter is able to fully complement the ppi1 mutant (Jarvis et al., 1998).

Large TOC GTPases: Arabidopsis Toc159 receptor families

Toc159 was initially identified as an 86 kDa fragment due to its high susceptibility to proteolysis (Kessler et al., 1994; Bolter et al., 1998; Waegemann and Soll, 1995). Toc159 has been proposed to be involved in initial preprotein binding (Chen et al., 2000). A detailed understanding of the targeting and insertion of this important group of proteins is definitely on the horizon. Correct subcellular sorting and membrane anchoring of Toc159 relies on the vital information that is invariably found at the C-terminal segment of the protein (Muckel and Soll, 1996). For the initial docking and proper integration of Toc159 into the TOC complex, intrinsic Toc159 GTPase activities as well as the interaction of its M-domain with both the G-domain of Toc34 and Toc75 are essential (Bauer et al., 2002; Schleiff et al., 2002; Wallas et al., 2003). Characterization of the T-DNA insertion mutant of Arabidopsis Toc159 (atToc159), ppi2 mutant, showed that the differentiation of plastid into plastids is arrested, resulting in an albino phenotype (Bauer et al., 2000): in other words, the plant cannot develop photoautotrophically. The M-domain, the function of which is to anchor the protein in the outer membrane and to assemble the TOC core complex, was demonstrated to partially complement the preproteins' import defect in ppi2 mutant (Lee et al., 2003). The accumulation and expression level of photosynthesis-related proteins were drastically decreased. This did not appear to be the case for non-photosynthetic plastid proteins. This observation led to the proposal that proteins from this class are imported by other members of the large Toc GTPase family, namely atToc132 and atToc120 (Bauer et al., 2000; Kubis et al., 2004). These different receptors were assembled into structurally-distinct translocation complexes that exhibit unique preprotein binding properties that differ to that of atToc159 (Figure 2; Bauer et al., 2000; Ivanova et al., 2004; Kubis et al., 2004). Interestingly, single knock-out mutants of atToc132, atToc120 or atToc90 show no specific phenotypes, indicating their subsidiary function in plastid biogenesis (Hiltbrunner et al., 2004; Kubis et al., 2004). However, the double mutant atToc132/atToc120 is seedling lethal (Ivanova et al., 2004), demonstrating a phenotype similar to ppi2 (Kubis et al., 2004), which confirms that these two proteins are redundant and constitute a structural and functional subclass of protein import receptors. The double knock-out mutants of atToc90/atToc132 and atToc90/atToc120 are viable even though atTOC90 expression level is high throughout development of WT plants (Kubis et al., 2004), indicating that there is little redundancy between the atToc90 and atToc132/atToc120 subgroups. Interestingly, over-expression of atToc90 was able to partially complement ppi2 mutant (Infanger et al., 2011). The accumulation of photosynthetic proteins in the ppi2 genetic background could (Hiltbrunner et al., 2004) be restored by the atToc90 over-expressor line – strongly hinting to a potential involvement in the atToc159-targeting pathway (Infanger et al., 2011).

Construction underway: the ambiguous role of Toc159 family A-domain

As aforementioned, the function of the A-domain of Toc159 is still under investigation, largely due to its dispensable function in chloroplast biogenesis (Lee et al., 2003). It is highly improbable that the A-domain has evolved and was conserved throughout evolution without having functional significance. In all likelihood, the A-domains of the Toc159 homologs represent major determinants of distinct pathways for protein import into plastids (Ivanova et al., 2004). The selectivity of the different receptors (atToc159, atToc132 and atToc132) towards preproteins was altered when their respective A-domains were swapped (Inoue et al., 2010). Similarly, an atToc132GM over-expressor line (a construct lacking the A-domain) was able to partially complement the import defects in ppi2 mutant, but in a non-discriminating fashion towards the different classes of preproteins (Inoue et al., 2010). These observations
clearly suggest that the A-domain of Toc159 receptor families confers a certain degree of selectivity to the distinct TOC core complexes. Further, the isolated A-domain behaves as an intrinsically disordered protein (Figure 3; Richardson et al., 2009). This places them in the category of natively unstructured proteins (Hernandez Torres et al., 2007). Many disordered regions are associated with protein–protein interaction and surprisingly implicated in an array of regulatory functions in eukaryotic cells (i.e., control of cell cycle and the regulation of transcription and translation; Dyson and Wright, 2005). In agreement with the concept that reversible protein phosphorylation is central to the regulation of most aspects of cell function (Johnson, 2009), many disordered regions present in proteins are indeed regulated by phosphorylation (Dyson and Wright, 2005). Phosphoproteomic profiling of A. thaliana proteins from several independent studies revealed that the A-domain as well as the full-length Toc159 is phosphorylated (de la Fuente van Bentem et al., 2008; Reiland et al., 2009). Similarly, cell fractionation followed by in vitro phospospecific staining further demonstrated that full-length Toc159 and the free A-domain were indeed both phosphoproteins (Agne et al., 2010). As such, the regulation of the A-domain via phosphorylation is not surprising, as it coincides nicely with the reported phosphoregulation of the intrinsically disordered proteins (Dyson and Wright, 2005; Johnson, 2009). Taken together, these findings strongly hint at a complex regulation of A-domain function that is important for the maintenance of preprotein selectivity at the TOC translocons.

There are, undoubtedly, many surprises still in store, and our understanding of the regulatory role of A-domain, particularly its significance in the translocation process, is still in its infancy. One of the major gaps in our knowledge concerns the precise nature of the signal that is necessary for targeting the different classes of preproteins. In other words, how do these different receptors distinguish their own client preproteins from the others?

The diverse nature (i.e., the variable size, no apparent secondary structure and consensus sequences) of

Figure 2  Hypothetical model of two distinct TOC core complexes and two independent channels at the TIC translocon.
The Toc159 and Toc34 family members assemble into distinct translocons with the Toc75 channel. The different TOC receptor isoforms mediate the recognition of distinct classes of nucleus-encoded preproteins (red and yellow) to maintain the proper levels of functional classes of proteins that are required for the biogenesis and homeostasis of the organelle (Ivanova et al., 2004; Inoue et al., 2010). Also depicted is the hypothetical model of Tic20 and Tic110 channels in the inner envelope (IE) of chloroplasts. After tranlocating through the outer envelope (OE), preproteins are imported either via Tic110 or Tic20 through the IE. Tic110 is thought to form a homodimer with a total of eight amphipathic transmembrane helices forming the translocation channel and four hydrophobic α-helices involved in the insertion into the membrane (Lubeck et al., 1996; Balsera et al., 2009). The proposed Tic20 channel is depicted as a homo-oligomer with a proposed molecular mass of >700 kD but only three molecules are drawn for simplicity. Due to the low overall abundance of Tic20 (Kovacs-Bogdán et al., 2011), it might be responsible for importing a smaller and distinct subset of preproteins.
transit peptides certainly promise to be very challenging in the functional characterization of A-domains. Recent advances, however, have demonstrated that subgroups of transit peptides contain distinct motifs that could alter their import efficiency and receptor specificity (Lee et al., 2008; Lee et al., 2009a). Therefore, the finding that the Toc159 family A0domains are natively unstructured proteins is highly significant. Many natively unstructured proteins have a large surface area under physiological conditions (Dyson and Wright, 2005). This makes them a perfect platform for simultaneous interaction with several binding partners (Dyson and Wright, 2005). The predominant unordered structure of the A-domain, as well as its 50% coverage of the total length of the protein within the Toc159 family (with exception of atToc90), leaves it in a good position to facilitate interaction with multiple motifs within transit peptides. In addition, many natively unstructured proteins undergo transitions to a more stable secondary or tertiary structure upon binding to their target proteins (Dyson and Wright, 2005). Hence, the presence of distinct targeting motifs in the different classes of preproteins coupled with the subsequent substrate (preprotein)-binding that might induce subtle conformational changes may reflect the reported differential recognition between the different Toc159 receptors. Is this a daring hypothesis? The A-domain of both atToc132 and atToc159 adopt a certain but distinct secondary structure, as measured in circular dichroism spectra, upon incubation with trifluoroethanol under normal physiological temperature and pH (i.e., 25°C and pH 7.6; Richardson et al., 2009). Whether these observations reflect the actual situation in vivo and whether the A-domain will behave in a similar fashion in the presence of preproteins remains to be established.

There is a recent paper by Bischof and co-workers which utilizes a combination of quantitative proteomics and transcriptomics approach to investigate the role of Toc159 receptor in the assembly of chloroplast proteome using Arabidopsis WT and ppi2 mutant plants (=Toc159 deficient plants). Both genotypes (WT and ppi2 mutant) demonstrated an efficient accumulation both photosynthetic as well as non-photosynthetic proteins (Bischof et al., 2011), which now leads to the notion that the substrate specificity of Toc159 is not only limited to the photosynthetic proteins. The decrease in accumulation of photosynthetic proteins in ppi2 mutant reported previously (Bauer et al., 2001; Hiltbrunner

Figure 3  The A-domains of the Toc159 receptors are predicted to be predominantly unstructured. FoldIndex (Prilusky et al., 2005) was used to predict the intrinsic disordered region of the Toc159 family A-domain. The regions predicted to be disordered are shaded in gray.
residual import in \( \text{Infanger et al., 2011} \), it is highly improbable that the accumulation of Toc159 client proteins in the over-expression of Toc90 could partially restore the Toc120 receptors in the observed by Bischof and co-workers in their study, was the accumulation of proteins in the mutant plastids which catalyze the exchange of GDP for GTP forms; and guanine nucleotide-exchange factors (GEFs), the hydrolytic capacity of GTPases to promote GDP-bound molecules: GTPase-activating proteins (GAPs), which enhance states to interact with specific classes of regulator molecules. The 'empty' state (usually) serves as a transient piece has been presented in the recent work by Oreb et al., 2009). As demonstrated by size-exclusion chromatography, binding of preprotein does not directly stimulate GDP dissociation of the TOC GTPase cycle and its impact on the translocation process. In this context, several possibilities have been proposed. An early hypothesis is that each monomer of the psToc34 or atToc33/34 receptor acts as GAP for its respective partner, thus activating its cognate GTPase activity (Sun et al., 2002). An arginine residue (also designated as arginine finger) is often found at the dimerization cleft of these GAPs. Indeed, a critical arginine residue (R133 in psToc34 and R130 in atToc33) was present in the TOC receptor dimer interface. This positional feature, however, is structurally important for dimerization but has little effect on binding and the overall structure of the monomer (Weibel et al., 2003), consequently overruling the idea that the monomer of psToc34 or atToc33/34 acts as self-activating GAP. Others have reported that preproteins, specifically the transit peptide, stimulate GTP hydrolysis of the TOC receptors (Jelic et al., 2002; Lee et al., 2009b). The interrelation between dimerization and TOC GTPase cycles and the subsequent downstream events, however, was not established in these studies. This missing puzzle piece has been presented in the recent work by Oreb and co-workers (2011), where it is demonstrated that the dimerization kinetics of Toc34 dimer is altered upon preprotein binding. The GDP dissociation constant of the dimer is much lower than that of the monomer (Oreb et al., 2011). As demonstrated by size-exclusion chromatography, binding of preprotein does not directly stimulate GDP release from each individual monomer (Oreb et al., 2011). Hence, a comparable function of the transit peptide and GEF is excluded by the authors as nucleotide exchange is apparently a downstream effect from the transit peptide-dependent psToc34/atToc33 dimer dissociation event.

**The paradox of TOC GTPase cycles**

The G-domains of the TOC GTPases exhibit classical motifs of nucleotide binding and hydrolysis characteristic to that of many GTPases (Bourne et al., 1991). GTPases are often regarded as 'molecular switches' that regulate a myriad of cellular processes in eukaryotic cells (e.g., cell proliferation, signaling, protein trafficking and synthesis; Vernoud et al., 2003). GTPases typically switch through three common conformational stages: GDP-bound, 'empty' and GTP-bound states, with the former being the inactive condition. The 'empty' state (usually) serves as a transient intermediate to replace GDP with GTP in the guanine nucleotide-binding site of the GTPase (Gasper et al., 2009). Most often, the switching function of the GTPases is governed by the differing abilities of its conformational states to interact with specific classes of regulator molecules: GTPase-activating proteins (GAPs), which enhance the hydrolytic capacity of GTPases to promote GDP-bound forms; and guanine nucleotide-exchange factors (GEFs), which catalyze the exchange of GDP for GTP \textit{in vivo} (Aronsson and Jarvis, 2011).

Although it is generally clear that the TOC GTPase receptors control transit peptide recognition and the initial stages of membrane translocation, the molecular details that link GTPase activity with receptor function remains vague. The three-dimensional structure from pea Toc34 (psToc34) disclosed a homodimeric configuration with bound GDP at the dimerization interface (Sun et al., 2002). Such structural arrangement is strikingly reminiscent of a GTPase with its corresponding GAP (Gasper et al., 2009). The unifying principle of the TOC receptors is that psToc34 or \textit{Arabidopsis} Toc33/34 (atToc33/34) dimerizes in a nucleotide-dependent manner and GDP binding moderates the interaction between the G-domains of the respective monomer. Dimerization has henceforth been proposed as an important regulatory aspect that controls the entry of preproteins to the membrane translocation channel of the TOC translocon (Sun et al., 2002). This structural observation does not allow any conclusion to be drawn concerning the underlying mechanism of dimerization of the TOC GTPase cycle and its impact on the translocation process. In this context, several possibilities have been proposed. An early hypothesis is that each monomer of the psToc34 or atToc33/34 receptor acts as GAP for its respective partner, thus activating its cognate GTPase activity (Sun et al., 2002). An arginine residue (also designated as arginine finger) is often found at the dimerization cleft of these GAPs. Indeed, a critical arginine residue (R133 in psToc34 and R130 in atToc33) was present in the TOC receptor dimer interface. This positional feature, however, is structurally important for dimerization but has little effect on binding and the overall structure of the monomer (Weibel et al., 2003), consequently overruling the idea that the monomer of psToc34 or atToc33/34 acts as self-activating GAP. Others have reported that preproteins, specifically the transit peptide, stimulate GTP hydrolysis of the TOC receptors (Jelic et al., 2002; Lee et al., 2009b). The interrelation between dimerization and TOC GTPase cycles and the subsequent downstream events, however, was not established in these studies. This missing puzzle piece has been presented in the recent work by Oreb and co-workers (2011), where it is demonstrated that the dimerization kinetics of Toc34 dimer is altered upon preprotein binding. The GDP dissociation constant of the dimer is much lower than that of the monomer (Oreb et al., 2011). As demonstrated by size-exclusion chromatography, binding of preprotein does not directly stimulate GDP release from each individual monomer (Oreb et al., 2011). Hence, a comparable function of the transit peptide and GEF is excluded by the authors as nucleotide exchange is apparently a downstream effect from the transit peptide-dependent psToc34/atToc33 dimer dissociation event.
Heterodimerization between Toc34 and Toc159 has also been observed biochemically (Wallas et al., 2003; Weibel et al., 2003; Yeh et al., 2007; Rahim et al., 2009) and is achieved via conserved arginine residues in Toc34 (R133 in psToc34 and R130 in atToc33) with the G-domain of Toc159 (Schleiff et al., 2002). Thus, homodimerization of the Toc159 receptor is theoretically possible, however, experimental data to support this possibility are currently lacking. An attractive current hypothesis envisaged that the heterodimerization event could represent an additional regulatory element in the TOC GTPase cycle or that the transition between homodimeric and heterodimeric state acts as trigger for translocation initiation.

The recent data have indeed shed new light on the mechanism underlying the regulation of TOC GTPases; still, our current understanding of this unique cycle and its significance in the translocation process is far from being complete. These findings call into question the prevailing ‘missing’ factors that fill the gap in the GTPase cycle. At such, additional factors for activation (GAPs or co-activating GAPs) and GEFs would be essential for precise mapping of the interaction network within the TOC translocation complex.

The pivotal role of phosphorylation on TOC receptors

Phosphorylation, in general, is a widespread regulatory principle of many protein complexes, which are part of wide arrays of cellular processes (Seebach et al., 2005; Wang and Johnsson, 2005). The TOC receptors can be post-translationally modified by kinases present in the hosting membranes both in vitro and in vivo (Sveshnikova et al., 2000; Jelic et al., 2002). Arabidopsis Toc34 receptor was shown to be phosphorylated in a similar fashion to P. sativum Toc34 receptor in in vitro studies (Sveshnikova et al., 2000; Fulgos and Soll, 2002; Jelic et al., 2003). Surprisingly, the results revealed that in Arabidopsis only one isoform of Toc34 (atToc33) is phosphorylated, the second isoform (atToc34) was not (Jelic et al., 2003). Since both isoforms acted in distinct targeting pathways involved in transporting specific subsets of proteins (see below), this observation postulated that the different receptors within the distinct TOC complexes might be differentially regulated. Phosphorylation specifically imposes a negative influence on the TOC receptors, particularly their GTPase activity (Sveshnikova et al., 2000; Fulgos and Soll, 2002; Jelic et al., 2002) and the assembly of Toc34 with the TOC core complex into the outer membrane (Oreb et al., 2008).

The functional relevance of TOC receptor phosphorylation is demonstrated via the mutant mimicking phosphorylation (atToc33<sub>S181E</sub> Aronsson et al., 2006). The actual impact of phosphorylation on TOC receptors has been discussed controversially in the literature; in one study, the atToc33<sub>S181E</sub> mutant indeed demonstrated a reduced GTPase activity with lower affinity for preproteins in vitro (Aronsson et al., 2006; Oreb et al., 2007). Furthermore, complementation studies of the atToc33 knockout mutant, ppi1, with three different over-expressor transgene atToc33 mutants (atToc33<sub>S181E/S118A/S181D</sub>) demonstrated efficient complementation of the ppi1 phenotype in all tested criteria (i.e., chloroplast protein import efficiency, chloroplast ultra structure, overall plant growth and development, chlorophyll concentration and light stress tolerance), indicating phosphorylation of atToc33 at serine 181 is functionally dispensable in vivo (Aronsson et al., 2006). Oreb and co-workers (2007) showed in a subsequent study that the photosynthetic performance of over-expressor atToc33<sub>S181E</sub> ppi1 transgenic lines is slightly reduced. This observation, however, only occurs at an early developmental stage under heterotrophic growth conditions, correlating nicely with the requirement for atToc33 in the early steps of chloroplast biogenesis (Kubis et al., 2004). More recently, phosphorylation of atToc33 was reported to negatively regulate its homodimerization and heterodimerization with atToc159 as well as its assembly into the TOC core complex (Oreb et al., 2008).

Multiple import pathways

The prevailing idea of multiple import pathways is now seemingly possible, largely due to functional selectivity between members of the Toc159 and Toc34 families in Arabidopsis that demonstrate differential import characteristic of photosynthetic and non-photosynthetic plastid proteins. Evidence that Toc159 forms a TOC complex that is distinct from that of Toc132 and Toc120 in Arabidopsis comes from in vitro co-immunoprecipitation studies with detergent extracts of total chloroplast membranes (Ivanova et al., 2004). Both Toc132 and Toc120 were found to form a single TOC complex together, further supporting the functional redundancy within this subclass of receptors (Ivanova et al., 2004). While atToc33 co-immunoprecipitates mainly with atToc159, atToc34 forms a complex together with atToc132/atToc120. (Ivanova et al., 2004). This observation led to the notion that the TOC core complex in Arabidopsis comprises either combination – atToc159/33/75 or atToc132/120/34/75 – whose
functions are reflected by their individual receptor diversities (Ivanova et al., 2004). Cross-reactivity was observed for atToc34 between atToc159 and atToc132/120 complexes. However, the authors reasoned that the functional redundancy of Toc34 receptors in Arabidopsis might contribute to this experimental observation (Jarvis et al., 1998; Ivanova et al., 2004). Only one functional Toc75 homolog (atToc75-III) was found in complex with atToc159/33 and atToc132/120/33 (Ivanova et al., 2004). These functionally-distinct import routes seem to converge at the TIC complex via atTic110 (Ivanova et al., 2004). However, the differential expression of Tic20 may suggest that they represent alternative TIC translocation complexes that contribute to the distinct targeting pathways together with the different TOC core complex (Inaba et al., 2005; see below). Multiple structural and functionally-distinct TOC core complexes have thus far been reported in Arabidopsis. These distinct TOC complexes are, however, not only exclusive to higher plants. As demonstrated by bioinformatic analysis, the diversities in TOC receptors was also observed among other species, including Spinacia oleracea (spinach; Voigt et al., 2005), Oryza sativa (rice; Kubis et al., 2004), P. patens (moss; Hofmann and Theg, 2004) and Picea abies (spruce; Fulgosi et al., 2005).

The structurally- and functionally-distinct TOC complexes therefore represent adaptation strategies with diverse gene-expression profiles during plastid differentiation in plants. In this context, multiple import pathways allow efficient targeting of ‘housekeeping’ vs. highly expressed proteins (which are required for specialized metabolic functions) in the dynamic nature of plastids (Inaba et al., 2005).

The TIC complex: a short glimpse

Translocation of preproteins into chloroplasts also requires the passage through the inner membrane, a process that is facilitated by the TIC complex. In most cases, preproteins are thought to be translocated simultaneously through both TOC and TIC complexes. Some components of the TIC complex have been identified and extensively characterized over the years; but the question regarding the precise nature of the TIC channel remains enigmatic. Three conserved membrane-spanning proteins – Tic110, Tic20 and Tic21 – were proposed as candidates for the inner membrane translocation channel. The latter was proposed by Teng and co-workers (2006) as the third potential protein-conducting channel at the inner membrane of chloroplasts. This notion was disputed in another study, where the reported Tic21 (Teng et al., 2006) most likely represents an ancient metal permease, which regulates iron uptake and metal homeostasis in chloroplast and not a protein conducting channel (Duy et al., 2007). Despite the occurrence of several protein-conducting channel candidates for the inner membrane of the chloroplast, considerable lines of evidence clearly pinpoint Tic110 as the central subunit of the TIC complex, forming a high-conductance cation-selective channel (Heins et al., 2002; Balsera et al., 2009). Electrophysiology measurements indicate a pore size of 1.7 nm, similar to that of Toc75. Tic110 is encoded by a single gene copy in A. thaliana and is constitutively expressed in all tissues, indicating an indispensable role in plastid biogenesis (Inaba et al., 2005). Homozygous T DNA insertion lines of Tic110 are embryo lethal, further establishing the role of Tic110 in plant viability. Additionally, Tic110 contains binding sites for stromal Hsp93 and Cpn60 (Kessler and Blobel, 1996; Jackson et al., 1998; Inaba et al., 2003). Both chaperones function as part of the import motor, providing a driving force for translocation, as well as folding of the imported proteins in the stroma. This also accounts for the additional energy expenditure for the translocation of proteins across the inner membrane.

The evidence of more than one TIC channel constituent would lead to a hypothesis that the TIC complex comprises of at least two translocation channels: Tic110 as the core translocation channel with Tic20 forming a distinct channel (Kovacs-Bogdan et al., 2011), independent of Tic110 (Figure 3). It has been proposed that both translocation channels might be involved in the translocation of different subsets of proteins, mirroring the translocation system at the inner membrane of mitochondria where Tim22 and Tim23 form distinct translocation channels responsible for importing different sets of proteins (Mokranjac and Neupert, 2010).

Concluding remark

It is abundantly clear that many unanswered questions remain in relation to the diversities of the protein-targeting pathway in chloroplasts. The discovery of multiple distinct import complexes raises several questions concerning:
- the assembly of these structurally-distinct TOC complexes; and
- how the information of the different classes of preproteins is relayed to these different TOC complexes.
The research also highlights the targeting pathways for those proteins with internal targeting sequence (Miras et al., 2002, 2007; Nada and Soll, 2004). The question remains as to whether this specific subset of proteins is transported by the distinct atToc159/atToc33–atToc132/120/34 receptor complexes or via an as yet uncharacterized protein translocation complex. As preprotein recognition and translocation at the outer chloroplast membrane require a tight, GTP-dependent regulation on both Toc159 and Toc34 receptors, the missing elements in each of these processes (i.e., the complex cycle of GTPase and negative regulation by GTP, such as kinases and phosphatases as well as co-GAPs and GDP/GTP GEFs, will definitely provide a beneficial insight into how these regulatory processes relate to the protein translocation process in chloroplast.

Received June 21, 2012; accepted August 1, 2012

References


De Gruyter


Wai-Ling Chang studied Biochemistry at the University of Malaya in Malaysia and obtained her bachelor's degree in 2005. In order to gain a broader insight into the different aspects of molecular and cell biology, she later joined the master programme in Molecular and Cell Biology at the University of Heidelberg and graduated in 2009. Following her passion in basic research, she is currently working in the group of Prof. Soll in LMU Munich as a PhD student, where she is expected to finish in March 2013.

Bettina Bölter studied biology at the CAU in Kiel and obtained her diploma in 1997. She stayed on for a PhD in cell biology, which was finished in 2000. Afterwards Dr. Bölter joined Prof. F.U. Hartl’s group at the MPI for Biochemistry in Martinsried as a postdoc. To gain experience in more applied sciences she switched to the institute of pathology at the UKSH in Lübeck in 2003, where she was responsible for molecular diagnostics. Since the fascination for basic research and especially chloroplast biology remained, Dr. Bölter accepted a position as a group leader at the LMU Munich and completed her habilitation in 2010.

Prof. Jürgen Soll studied biochemistry and obtained his PhD from the Technical University of Hannover. After postdoctoral studies at UC Berkeley he obtained his habilitation at the University of Munich. Soll held several professor positions at the universities of Saarbrücken and Kiel before he moved to his current Chair in Plant Biochemistry in 2001 at the University of Munich.