

Chloroplast transit peptides: structure, function and evolution

Barry D. Bruce

It is thought that two to three thousand different proteins are targeted to the chloroplast, and the 'transit peptides' that act as chloroplast targeting sequences are probably the largest class of targeting sequences in plants. At a primary structural level, transit peptide sequences are highly divergent in length, composition and organization. An emerging concept suggests that transit peptides contain multiple domains that provide either distinct or overlapping functions. These functions include direct interaction with envelope lipids, chloroplast receptors and the stromal processing peptidase.

The genomic organization of transit peptides suggests that these domains might have originated from distinct exons, which were shuffled and streamlined throughout evolution to yield a modern, multifunctional transit peptide. Although still poorly characterized, this evolutionary process could yield transit peptides with different domain organizations. The plasticity of transit peptide design is consistent with the diverse biological functions of chloroplast proteins.

Barry Bruce is in the Dept of Biochemistry, Cellular and Molecular Biology Department, The Center for Legume Research, University of Tennessee-Knoxville, The Graduate Program in Genome Science and Technology, University of Tennessee and Oak Ridge National Laboratory. E-mail: bbruce@utk.edu

The family of plant organelles, collectively known as plastids, are widely accepted to have evolved from free-living cyanobacteria through the process of endosymbiosis. Although a modern plastid still retains a semi-autonomous genome, its coding capacity has been reduced to only 100–200 genes. The now classic assumption of the endosymbiotic theory, which was initially articulated by Weeden¹, is that the proteins encoded by the genes newly transferred from the endosymbiont to the host genome will return to the organelle from which they originated. Considering that a typical cyanobacteria contains ~3200 genes², several thousand gene products must be targeted back to the plastid to enable the same level of metabolic complexity in a modern plant cell as existed in the ancestral cyanobacteria. The vast majority of these proteins are targeted back into the chloroplast as a precursor protein whose transport is facilitated by the acquisition of an N-terminal extension referred to as a transit peptide.

Although the first demonstration of precursor transport into chloroplasts was shown over two decades ago^{3,4}, only now is this area of cell biology becoming well understood. Many excellent reviews have been published recently on the evolution of plastids⁵, the evolution of organelle genomes⁶, the mechanism of gene transfer from organelles to the nucleus⁷ and the mechanism of protein import into chloroplasts^{8,9}.

Proteins destined to plastids and other organelles share in common the requirement for 'new' sequence information to facilitate their correct trafficking within the cell. Although in most cases this information resides in a cleavable, N-terminal sequence often collectively referred to as signal sequence, the different organelle-targeting sequences have distinct properties and names: 'signal peptides' for the endoplasmic reticulum, 'presequences' for the mitochondria and 'transit peptides' for chloroplasts and other plastids. This review focuses on recent progress in dissecting the role of the stromal-targeting domain of chloroplast transit peptides. I will consider briefly the multitude of distinct functions that transit peptides perform, provide an update on the limited structural information of a number of transit peptides and finally give some ideas on their evolution. Figure 1 provides a schematic illustration of the chloroplast translocation apparatus and a working model of different transit peptide-mediated interactions that must occur during successful protein import. This model is based on multiple observations and contributions from several laboratories and has been simplified for the purposes of this review.

The pool size of chloroplast precursors

With the recent progress in genomic sequencing efforts, it is not surprising that the number of potential chloroplast precursors identified is growing rapidly. Although certainly not up to date, the CHLPEP database¹⁰ contains sequences for nearly 300 different transit peptides. A similar database today might contain well over a thousand different transit peptides. However, it is difficult to know how many different precursors are targeted to plastids during the life span of a typical plant. Recently, ChloroP, a neural network-based method of predicting transit peptides was used to analyse the 715 *Arabidopsis* proteins found in SWISS-PROT¹¹. ChloroP showed that 13–22% of the proteins contained potential chloroplast transit peptides. Extrapolating to the entire *Arabidopsis* genome^{12,13}, using even the lower value (13%), gives a predicted number of chloroplast-targeted precursors of 2900–3500, depending upon the estimated size of the genome. However, it is both interesting and reassuring that analysis of the existing *Arabidopsis* genomic databases with the best computational tools currently available confirms the original predictions on the number of chloroplast precursors required to permit plastid metabolic complexity to approach that of a free-living cyanobacterium^{12,13}.

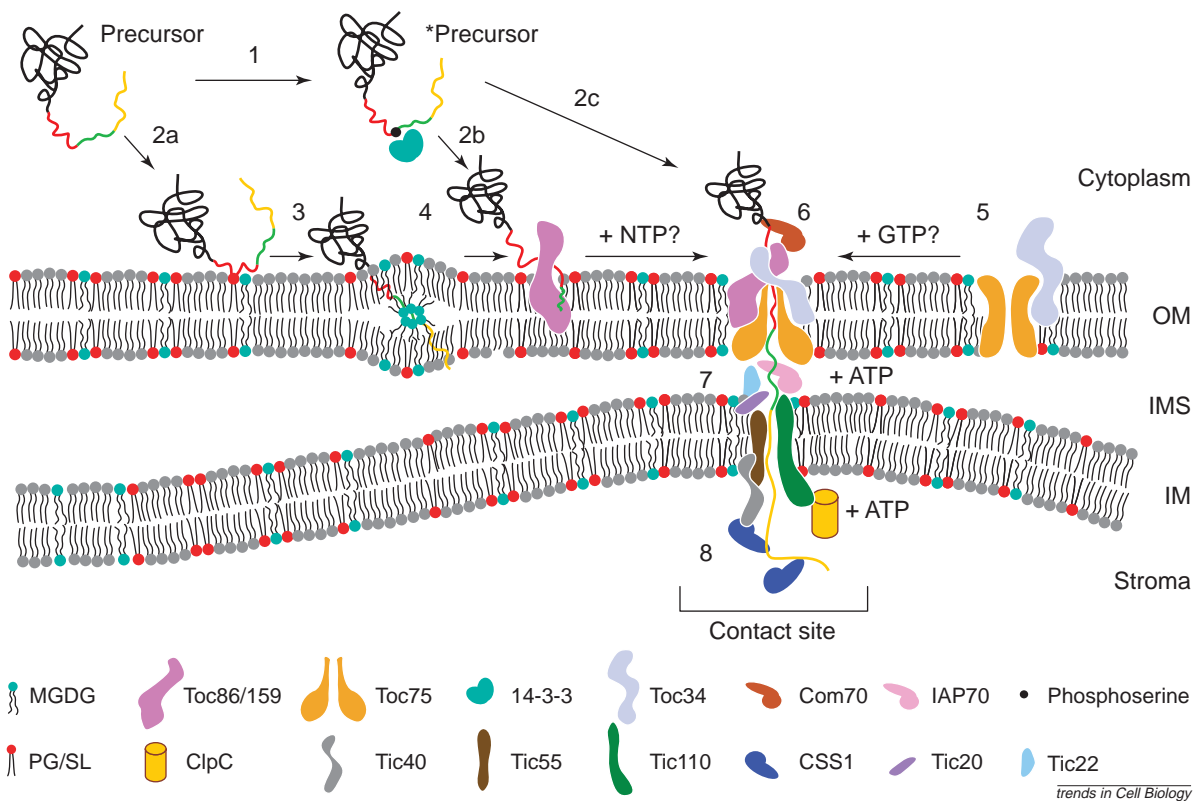


FIGURE 1

General import pathway for plastid precursor import. Three hypothetical domains of the transit peptide are shown in red, green and yellow. Multiple steps of transit peptide-mediated protein import are shown by the following numbers: (1) Interaction of the precursor containing a phosphorylated serine in the transit peptide (*precursor) with soluble factors such as the molecular chaperone 14–3–3 protein in the cytoplasm. (2a) Partitioning of the precursor out of the cytoplasm on to the chloroplast surface through a direct NTP-independent interaction of the transit peptide with the chloroplast-specific lipids monogalactosyldiacylglycerol (MGDG), sulfolipid (SL) and phosphatidylglycerol (PG). (2b and 2c) Direct interaction of the precursor with Toc components, facilitated by association with the 14–3–3 molecular chaperone. (3) Peptide–lipid interactions resulting in reciprocal changes in both the transit peptide structure (shown as a green helix) and the lipid phase preference of the envelope (shown as an inverted micelle). (4) Recognition and interaction of transit peptide with Toc86/159 receptor. (5 and 6) Lateral movement and/or association of Toc86/159 with Toc34, resulting in the creation of a membrane contact site containing both the inner and outer envelope translocons. This also illustrates the sequential or concurrent GTP-driven insertion of the transit peptide into Toc75. (7) Precursor translocation across the outer envelope membrane by a push–pull mechanism using the ATP-dependent molecular motor(s) Com70 and/or IAP70. (8) Precursor translocation across the inner envelope membrane by a push–pull mechanism using the ATP-dependent molecular motor(s) IAP70 and/or CSS1.

Structural analysis of transit peptides

In contrast to the rapid progress in elucidating the primary structure of chloroplast transit peptides, only limited information is available concerning the structure of transit peptides. The lack of structural information is not only a result of limited investigation but also might reflect a fundamental property of transit peptides. Experimental results reveal that, in an aqueous environment, transit peptides are largely unstructured^{14–16}, reinforcing an earlier proposal that they have evolved to maximize the potential to form a random coil¹⁷. Attempts to measure the structure of transit peptides have utilized membrane-mimetic solvents such as TFE (2,2,2-trifluoroethanol) and aqueous buffers containing detergent micelles. Analysis of transit peptides in these solvents by circular dichroism spectrometry^{14–16,18} has demonstrated that transit peptides might contain significant α -helical structure(s); for example, in the transit peptide for prSSU (SStp), both the N- and the C-termini exhibit α -helical structure¹⁴. Although no single mechanism accounts

for the ability of TFE to preferentially stabilize peptides in a helical conformation, it is clear that this helix stabilization is not indiscriminate but does indeed reflect the underlying structural preferences of a given peptide sequence¹⁹.

Attempts to refine the identity and placement of transit peptide structural elements have used multidimensional NMR on either synthetic or recombinant transit peptides. To date, the only two structures reported are for the ferredoxin²⁰ and Rubisco activase¹⁶ transit peptide from the algae *Chlamydomonas*. Figure 2 shows the lowest-energy structures for these transit peptides. Both of these structures were determined in the presence of TFE and shown to contain a helix and a random coil. However, the order of these two motifs is reversed. Ferredoxin has an α -helix at its N-terminus from position A₂–V₁₃, followed by an unstructured C-terminal domain of ~19 amino acids. By contrast, the activase peptide exhibits an unstructured N-terminus of ~15 residues, followed by an α -helix from position A₁₈–L₃₀.

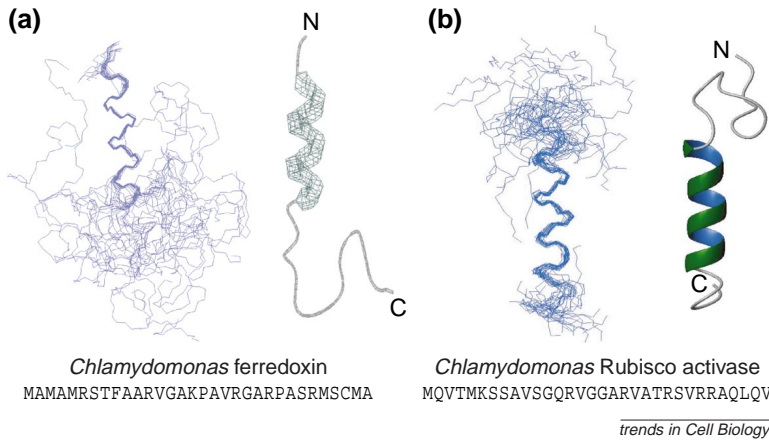


FIGURE 2

NMR structures for two *Chlamydomonas* transit peptides. The left panel includes an overlay of the 27 lowest-energy structures of the 32-residue ferredoxin transit peptide, and one isolated structure that denotes the N- and C-termini. The right panel shows an overlay of the 22 lowest-energy structures of the 32-residue Rubisco activase transit peptide, and one isolated structure that denotes the N- and C-termini. Note the inverted organization of these two peptides, helix-coil for ferredoxin versus coil-helix for activase.

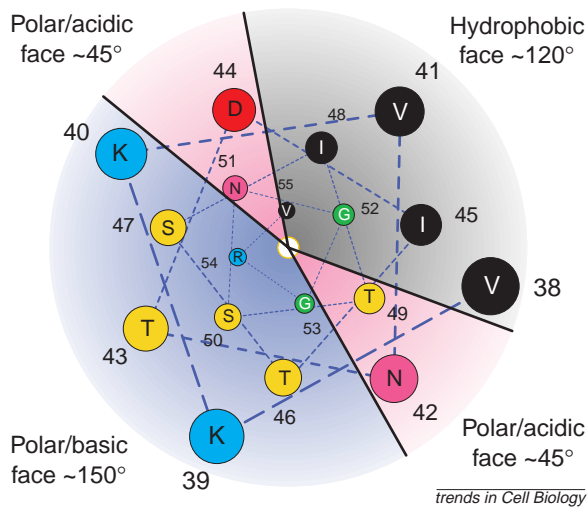


FIGURE 3

Helical wheel projection of prSSU transit peptide amino acids 38–55. The C-terminal region of the prSSU transit peptide is an example of an amphipathic sequence within a chloroplast transit peptide. However, unlike mitochondrial presequences, this amphipathicity is largely the result of the selective placement of the hydroxylated amino acids serine and threonine on one face of this potential α -helix. Black indicates hydrophobic amino acids. Red/pink represents acidic/polar amino acids. Blue represents basic amino acids. Yellow represents hydroxylated amino acids. Green represents the helix-breaking residues glycine/proline.

Unfortunately, the only NMR data for higher-plant transit peptides are for *Silene ferredoxin* (trFd). Since plant transit peptides are much longer than algal transit peptides, this work required that the full-length peptide be expressed and labelled with stable isotopes in *Escherichia coli*¹⁵. These investigators evaluated the structure in both TFE and a

more physiological micellar system that contained monogalactosyldiacylglycerol (MGDG) and an anionic detergent, dodecylphosphoglycol²¹. Insertion/association of trFd with these micelles introduced two helical domains (S_{10} – L_{13} and G_{30} – F_{34}) in an otherwise unstructured peptide. Further relaxation experiments indicated that, although the whole peptide backbone is flexible, the central proline-rich region (P_{15} – P_{26}) is highly unstructured, suggesting that this region of the trFd requires extreme flexibility for its function. Suggestions for functions for this region include a role as a flexible linker that disrupts or separates other functional regions, or alternatively this region could act as a recognition motif for some later step in protein import, as was suggested earlier²².

Although only a few transit peptides have been analysed by NMR and circular dichroism, some general conclusions are apparent. In aqueous solution, transit peptides are largely unstructured. When placed in either a more hydrophobic solvent or, upon insertion into micelles, one or more regions of the transit peptide become α -helical. Although algal transit peptides appear to contain only a single helical domain, both the higher plant ferredoxin and small subunit transit peptide contain two discontinuous α -helical domains. In addition, although both transit peptides and presequences contain regions that are amphipathic, the amphipathicity in transit peptides is determined by hydroxylated amino acids instead of the basic residues seen in presequences. Although the helix-coil-helix organization of higher-plant transit peptides might be a universal feature, the position and degree of amphipathicity of these two helices could vary. For instance, in trFd, NMR indicates that the N-terminal helix is amphipathic, whereas, in SStp, the C-terminal region is predicted to be amphipathic (Fig. 3).

Such membrane-induced secondary structures in transit peptides could define an otherwise ‘silent’ recognition element(s) for the import machinery. An attractive feature of this hypothesis is that the environment at the chloroplast surface might induce a ‘common’ conformation in transit peptides, enabling a single receptor at the outer chloroplast envelope, such as Toc86/159, to bind to and facilitate transport of potentially thousands of different precursors²³. Although two related receptors, Toc120 and Toc 132, were identified recently, structural element(s) in transit peptides might still be required to ensure the high level of fidelity in targeting plastid-destined precursors²⁴. To firmly establish the generality of these structural motifs, it will be necessary to extend these structural studies to a wider range of chloroplast transit peptides.

Domain structure and modular organization

Although early investigation into the primary sequences of transit peptides identified three major blocks of amino acid homology²⁵, the concept of ‘homology blocks’ was challenged and shown not to describe transit peptides as a group²⁶. However, that report and others concluded that stromal-targeting transit peptides do contain three distinct regions: an

uncharged N-terminal domain of ~10 residues beginning with MA- and terminating with a G/P, a central domain lacking acidic residues but enriched in S/T and, finally, a C-terminal domain enriched in arginines and potentially forming an amphiphilic β strand²⁷.

Part of the variability of transit peptide sequences might be explained by the recent discovery that the import receptor of the outer envelope, Toc86/159, has two related receptors, Toc120 and Toc132, that together make up a small gene family²⁴. Although the basis of interaction between transit peptides and this family of receptors is still unknown, it has been proposed that Toc86/159 functions primarily in the import of photosynthesis-related precursors into chloroplasts, whereas Toc120 and Toc132 might function in the import of precursors to other types of plastids (amyloplasts, chromoplasts, leucoplasts, etc.) that are not active in photosynthesis²⁴. This hypothesis could partially explain why transit peptides as a group fail to comply with the original 'homology block' model of transit peptides. As a more detailed understanding of the interaction between transit peptides and these receptors emerges, we might be able to reclassify transit peptides into specific subgroups. Further 'bioinformatic' analysis of these subgroups of transit peptides might identify currently unrecognized regions of homology or functional similarity.

Lipid interacting activity

Without exception, membranes that are active in protein translocation contain significant levels of lipids that strongly prefer to adopt a non-bilayer structure. In plastids, MGDG has a wedge-like molecular shape and prefers to form an H_{II} phase when isolated. Several reports suggest that non-bilayer-forming lipids are required at one or more steps during protein translocation^{28,29}. In support of this idea, several studies have demonstrated that the interaction between transit peptides and artificial bilayers is dependent upon lipid composition. Specifically, this interaction is clearly increased when the artificial membrane contains MGDG^{30,31}. Although other studies have demonstrated that the ferredoxin transit peptide-lipid interaction is dependent on the presence of the anionic lipids PG and sulfoquinosyldiacylglycerol (SL), a significant interaction still occurs with monolayers containing MGDG³².

These results suggest two possible interactions between transit peptides and membrane lipids: an initial ionic interaction between the anionic phospholipids and the basic amino acids of the transit peptide, and a second interaction that involves the galactose headgroups of the glycolipids and the hydroxylated amino acids of the transit peptide. The latter could involve direct hydrogen bonding between the hydroxyl groups of the transit peptide and the galactose. Alternatively, the formation of 'new' hydrogen bonds between the hydroxylated amino acids and tightly bound water molecules at the membrane interface could result in a reduced hydration shell of the galactolipids¹⁴. This transit

peptide-induced 'competition' for water molecules would favour the non-bilayer-forming tendency of MGDG and thereby promote transient lipid polymorphisms. Localized H_{II} phase conformations (Fig. 1, Step 3) could be a crucial step in initiating protein targeting/translocation.

Attempts to map the region(s) of transit peptides involved in lipid interaction have identified domains at both the N- and C-termini of the few transit peptides that have been investigated. Analysis of the SStp indicates that primarily the last 10 amino acids of the C-terminal region of the transit peptide are responsible for interaction with artificial bilayers containing MGDG³¹. This region is within the predicted amphipathic α -helix containing the hydroxylated amino acids serine and threonine (Fig. 3). Analysis of the transit peptide for plastocyanin indicates that the lipid-interacting domain is within the first 10 amino acid residues of the thylakoid targeting domain, which is C-terminal to the stromal targeting domain¹⁸. By contrast, a detailed study using deletion mutants of trFd have shown that the N-terminal region primarily interacts with MGDG, whereas the C-terminus mediates recognition of negatively charged phospholipids²².

The ability of one or more regions of a transit peptide to interact directly with chloroplast lipids (Fig. 1, Step 2a) could provide the initial ATP-independent step in precursor association with the chloroplast surface. The unusually low protein content of the outer envelope would provide a two-dimensional surface that is defined predominantly by its lipid composition³³. This observation coupled with the apparent lack of obligate cytosolic factors (see below) suggests that the initial partitioning of precursor molecules from the site of synthesis to the chloroplast surface might rely heavily on the recognition of plastid-specific lipids. In addition, the ability of transit peptides to gain secondary structure in a membrane-mimetic environment suggests that transit peptide-lipid interactions are reciprocal in nature. These 'new' membrane-induced structural motifs (Fig. 1, Step 4) might facilitate a secondary interaction with one or more components of the translocation apparatus. As the structure of more transit peptides in different environments becomes available, the physicochemical basis of the transit peptide-lipid interaction should become clearer.

Transit peptide-receptor interactions

As is shown in Fig. 1, the precursor is recognized by one or more chloroplast-specific receptors. Although most models depict this interaction as being mediated predominantly by the transit peptide, there is evidence implicating that the mature domain might in some way modulate this interaction³⁴. Most of the detailed information concerning interaction between chloroplast precursors and components of the translocation apparatus has come from studying the two precursors prSSU and prFd by photochemical crosslinking³⁵⁻³⁷. In the absence of an ATP source, two components of the translocation apparatus Toc86 and Toc75 crosslink

to the precursor, presumably through the transit peptide^{35,36}. This crosslinked species represents a bound intermediate where Toc86/159 functions as an acceptor for the transit peptide/precursor. When the concentration of ATP was increased, the crosslinking occurred predominately between the precursor and Toc75^{35,36}. In addition, an inner membrane protein has been identified, Tic20, that probably contacts the incoming precursor in the inner membrane space^{36,37}. This crosslinked form of the precursor represents a true translocation intermediate (Fig. 1, Steps 6 and 7) that has crossed the outer envelope. In none of these studies has the specific region of either the transit peptide or the translocation component been identified.

A recent, systematic, attempt to explore interactions between the transit peptide and the translocation apparatus utilized multiple mutants of the prFd in which internal portions of the transit peptide were deleted³⁷. This study utilized the same label-transfer crosslinking strategy used in the previous studies^{35,36}. The authors demonstrated that, although the whole transit peptide is important for the overall efficiency of binding and import, the N-terminal region (residues 6–14) was crucial for initial interaction with the translocation apparatus and potentially contains essential information for interacting with Toc86/159. A second region (residues 15–25) was found to contribute to interaction with Toc86/159; however, this region is responsible primarily for interaction with components of the inner membrane, such as Tic20. Surprisingly, no region within the transit peptide was found to interact specifically with Toc75. The authors conclude that the information for interaction with the outer and inner membrane components resides in two separate, but partly overlapping, domains in the first 25 amino acids of the ferredoxin transit peptide. It will require similar detailed analysis of several transit peptides before there is sufficient information to evaluate how these different domains interact with the receptors in general.

Recently, as illustrated in Fig. 1, the outer envelope protein Toc34 has also been shown to interact directly with the chloroplast precursor prSSU³⁸. Toc34 is a GTPase (as is Toc86/159)^{39,40}. The interaction between prSSU and Toc34 is through a phosphorylated form of the transit peptide. Moreover, the affinity of interaction between Toc34 and the transit peptide is mediated by GTP binding to Toc34, with the precursor being released or inserted into the translocon upon nucleotide hydrolysis or exchange³⁸. Although the details of the coordinate interaction of Toc34 and Toc86/159 are still unclear and difficult to differentiate through *in situ* experiments⁴¹, it appears that Toc34 is involved somehow in 'handing over' the transit peptide/precursor to the translocon (Fig. 1, Steps 5 and 6).

In vivo analysis of transit peptide activity

In addition, to the *in vitro* analysis of different regions of the transit peptide in receptor interaction, recent efforts have begun to evaluate the role of different regions of the transit peptide *in vivo*.

This work has involved transgenic work in *Chlamydomonas*^{42,43} and *Arabidopsis*⁴⁴. The *in vitro* and *in vivo* effects of mutations in the transit peptides of ferredoxin^{22,44}, plastocyanin^{42,45} and the gamma subunit of CF1 have been compared⁴³. Surprisingly, the results derived from these two approaches did not agree, suggesting that *in vivo* and *in vitro* analyses probe different aspects of the chloroplast protein import process. However, some general conclusions can be drawn:

- most transit peptide mutations affect *in vitro* import more severely than *in vivo* import;
- deletions in the transit peptide N-termini eliminate *in vivo* import, suggesting that this region might be required to compete successfully with other precursors for import *in vivo*; and
- severely impaired precursors might still accumulate *in vivo* to a sufficient level to maintain viability.

Interaction with soluble targeting factors

In the targeting of proteins to organelles, soluble targeting factors have been shown to help maintain proteins in a transport-competent conformation, as well as provide an increased fidelity of targeting⁴⁶. However, it is thought that plastid-destined precursors do not require a soluble targeting factor. This notion is consistent with the transit peptide–lipid interaction being the initial organelle-specific interaction between a precursor and the plastid¹⁴.

New evidence is emerging that suggests a role of soluble factors in chloroplast protein import. Transit peptides may contain a motif that is specifically recognized by a member of the 14-3-3 class of molecular chaperones⁴⁷. The role of these chaperones in protein transport has already been established since one member, 'mitochondrial import stimulating factor' (MSF), has been shown to facilitate the delivery of import-competent precursors to the mitochondrial translocation apparatus⁴⁸. Recently it was shown that chloroplast transit peptides, when phosphorylated, could possibly function as a recognition domain for a 14-3-3 protein^{49,50}. This involvement has been advanced by the recent demonstration that, when translated in a wheat germ lysate, prSSU exists as a high-molecular-mass complex containing a 14-3-3 protein and an Hsp70. The complex maintains the precursor in a highly import-competent state, such that it imports *in vitro* into chloroplasts at a rate three to fourfold higher than that of the free precursor⁴⁷. Whether this complex is responsible for targeting prSSU to a specific envelope receptor (Fig. 1, Step 2b), similar to the role of MSF in mitochondria protein targeting⁴⁸, or alternatively enables the precursor to bypass one or more of the early steps in translocation (Fig. 1, Step 2c), making the import process more efficient, is not known. However, the fact that purified, recombinant chloroplast precursors can be imported *in vitro*, in the absence of the 14-3-3-Hsp70 complex, at rates comparable to requisite physiological targeting rates argues that interaction with a soluble component it is not an obligate step in the translocation process^{34,51}.

Recognition by Hsp70 molecular chaperones

For almost a decade, it has been proposed that the high random-coil content that is predicted for chloroplast transit peptides functions to provide a substrate for the Hsp70 class of molecular chaperones¹⁷. Although Hsp70 chaperones have been proposed to function as the molecular motor driving precursor translocation into the ER and mitochondria⁵², there has been little support for chloroplasts utilizing an Hsp70 as the ATP-dependent molecular motor. Therefore, even though chloroplasts contain three Hsp70 homologues⁵³, current models depict an Hsp100 (ClpC) as the motor driving precursor import⁸. However, recent evidence clearly indicates that the transit peptide for prSSU interacts both *in vitro* and in *E. coli* with Hsp70 proteins⁵⁴. Furthermore, combination of statistical algorithms and biochemical analysis *in vitro* identified a specific domain of the prSSU transit peptide that interacts with the chloroplast stromal Hsp70, CSS1 (Ref. 55). When these algorithms were applied to transit peptides in the CHLPEP database¹⁰, >75% of the transit peptides were also predicted to contain an Hsp70 recognition domain within the N-terminal third of the transit peptide⁵⁵. Interestingly, this work also demonstrated that many transit peptides contained a second, lower-affinity, site that was positioned in the central third of the transit peptide. In prSSU, the positions of these two sites were spaced appropriately to enable one transit peptide to engage Hsp70 simultaneously on both the *cis* and *trans* side of either the outer or inner envelope membrane. Involvement of Hsp70 as the chloroplast translocation motor is supported by the recent discovery that one component of the inner envelope translocator, Tic40, is related to the Hsp70-interacting protein Hip⁵⁶. This finding suggests that Tic40 might function to recruit Hsp70 to the inner envelope, where the chaperone utilizes ATP hydrolysis to drive the movement of a precursor protein across one or both envelope membranes. The ability of transit peptides to act as substrates for Hsp70 proteins provides a surprising functional similarity between transit peptides and mitochondrial presequences as presequences have also been shown to contain Hsp70-recognition domains at their N-terminus⁵⁷.

Origin of the chloroplast translocation apparatus and transit peptides

Insight into the origin of the chloroplast translocation apparatus has recently benefited from the completion of the genome sequence² of the cyanobacterium *Synechocystis*, as discussed in several reviews^{58–61}. Analysis of this sequence indicates that three translocation components, Toc75, Tic22 and Tic20, evolved from existing proteins in the cyanobacterial genome. The best characterized of these homologues, SynToc75, is speculated to have been transferred from the cyanobacteria into the 'host' nucleus yet retained its function as a transport protein. However, upon insertion into the plastid outer membrane, Toc75 adopted an inverted topology relative to that which it possessed when in the cyanobacterial cytoplasm, such that it could now function as an importer of molecules into the plastid, as is shown in Fig. 4a.

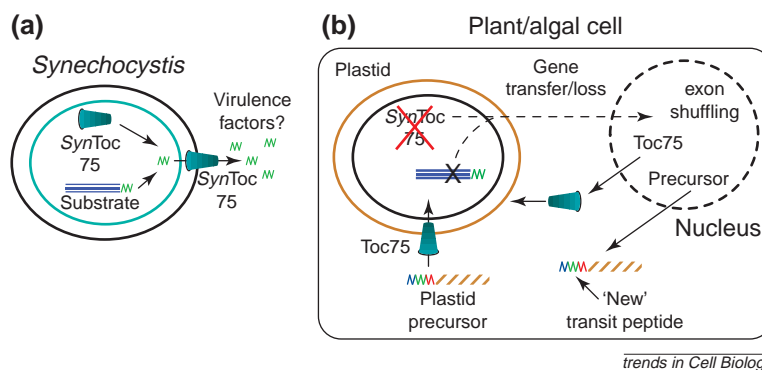


FIGURE 4

Model demonstrating possible evolutionary origin of Toc75 and modern transit peptides. (a) Possible role of SynToc75 in *Synechocystis* in facilitating the secretion of virulence factors. The substrates for this transporter might include sequences (shown in green) that are recognized by the SynToc75 and might provide the evolutionary origin of at least one domain of the modern chloroplast transit peptide. (b) Movement of genes encoding both SynToc75 and substrate(s) out of the cyanobacteria into the new host genome. The Toc75 protein is now inserted back into the plastid outer membrane in such a way that it now facilitates protein import back into the organelle. Some of the original information recognized by SynToc 75 (shown in green) is rearranged by exon shuffling, eventually yielding a modern transit peptide with the addition of new functional domains (shown in red and blue) positioned at the N-terminus of the 'new' precursor. This model is based on some of the ideas originally presented by G. McFadden⁶¹.

Related to this proposal, sequences contained within the cyanobacterial genome might also have provided the source of coding information for modern transit peptides. One interesting possibility is that the 'signal' that targeted proteins for secretion via SynToc75 could have given rise to some or all of the information now contained within a modern transit peptide (Fig. 4b). The emerging notion that transit peptides are organized as functional domains indicates that transit peptides emerged by shuffling of existing cyanobacterial exons, which through selective pressure could yield a transit peptide capable of targeting and translocating the nucleus-encoded cyanobacterial proteins back into the plastid. The recent report that 'early' exons would encode protein modules of 15–30 amino acids in length suggests that the average transit peptide could have evolved through the linking of three separate exons⁶². In support of this hypothesis, the transit peptides of several chloroplast precursors are still in fact encoded by three distinct exons^{63,64}. A very similar evolutionary mechanism has been demonstrated for a plant mitochondria presequence that originated from the acquisition of three exons from a separate donor protein⁶⁵. The fact that this multiple-exon organization of the transit peptide might not be characteristic of all plastid precursors suggests that, over time, genetic streamlining resulted in the loss of introns from the transit peptide coding region⁶⁶. Moreover, the process of exon shuffling might yield transit peptides with alternative domain organizations, as has been observed for the two *Chlamydomonas* transit peptides shown above (coil–helix vs helix–coil; Fig. 2)^{16,20}.

Concluding remarks

Taking all the structural data into account, it can be seen that transit peptides contain substantial regions of flexibility. Inserted within these extended, flexible structures are regions with helix-forming potential that become 'activated' when exposed to membrane-mimetic environments. Higher plants appear to require at least two helical domains, whereas the shorter algal transit peptides appear to contain only one. The basis for an additional sequence requirement in plants is not known but probably reflects the requirement for increased targeting fidelity and/or efficacy in a more complex cellular environment. Overlaid on top of these structural motifs are the multiple and potentially overlapping recognition elements for cytosolic factors, envelope receptors, molecular chaperones and the stromal processing protease. The ability of one short peptide sequence to satisfy simultaneously all of these various roles indicates both an exquisite evolutionary refinement and the possibility of a given sequence to reflect alternative properties when exposed to different chemical environments. This type of functional 'schizophrenia' could be a common trait of highly evolved targeting sequences in general. Moreover, the plasticity in design of a given transit peptide and for transit peptides as a group is ironically fitting for an organelle named in recognition of its own plasticity in form and function.

References

- 1 Weeden, N.F. (1981) Genetic and biochemical implications of the endosymbiotic origin of the chloroplast. *J. Mol. Evol.* 17, 133–139
- 2 Kaneko, T. *et al.* (1996) Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* 3, 109–136
- 3 Chua, N.H. and Schmidt, G.W. (1978) Post-translational transport into intact chloroplasts of a precursor to the small subunit of ribulose-1,5-bisphosphate carboxylase. *Proc. Natl. Acad. Sci. U. S. A.* 75, 6110–6114
- 4 Highfield, P.E. and Ellis, R.J. (1978) Synthesis and transport of the small of ribulose bisphosphate carboxylase. *Nature* 271, 420–424
- 5 Douglas, S.E. (1998) Plastid evolution: origins, diversity, trends. *Curr. Opin. Genet. Dev.* 8, 655–661
- 6 Gray, M.W. (1999) Evolution of organellar genomes. *Curr. Opin. Genet. Dev.* 9, 678–687
- 7 Martin, W. and Herrmann, R.G. (1998) Gene transfer from organelles to the nucleus: how much, what happens, and Why? *Plant Physiol.* 118, 9–17
- 8 Keegstra, K. and Cline, K. (1999) Protein import and routing systems of chloroplasts. *Plant Cell* 11, 557–570
- 9 Chen, X. and Schnell, D.J. (1999) Protein import into chloroplasts. *Trends Cell Biol.* 9, 222–227
- 10 von Heijne, G. *et al.* (1991) CHLPEP- A database of chloroplast transit peptides. *Plant Mol. Biol. Rep.* 9, 104–126
- 11 Emanuelsson, O. *et al.* (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* 8, 978–984
- 12 Lin, X. *et al.* (1999) Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. *Nature* 402, 761–768
- 13 Mayer, K. *et al.* (1999) Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*. *Nature* 402, 769–777
- 14 Bruce, B.D. (1998) The role of lipids in plastid protein transport. *Plant Mol. Biol.* 38, 223–246
- 15 Wienk, H.L. *et al.* (1999) The structural flexibility of the preferred transit peptide. *FEBS Lett.* 453, 318–326
- 16 Krimm, I. *et al.* (1999) A coil-helix instead of a helix-coil motif can be induced in a chloroplast transit peptide from *Chlamydomonas reinhardtii*. *Eur. J. Biochem.* 265, 171–180
- 17 von Heijne, G. and Nishikawa, K. (1991) Chloroplast transit peptides. The perfect random coil? *FEBS Lett.* 278, 1–3

- 18 Endo, T. *et al.* (1992) The chloroplast-targeting domain of plastocyanin transit peptide can form a helical structure but does not have a high affinity for lipid bilayers. *Eur. J. Biochem.* 207, 671–675
- 19 Buck, M. (1998) Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins. *Q. Rev. Biophys.* 31, 297–355
- 20 Lancelin, J.M. *et al.* (1994) NMR structures of ferredoxin chloroplast transit peptide from *Chlamydomonas reinhardtii* promoted by trifluoroethanol in aqueous solution. *FEBS Lett.* 343, 261–266
- 21 Wienk, H.L. *et al.* (2000) Structure, dynamics, and insertion of a chloroplast targeting peptide in mixed micelles. *Biochemistry* 39, 8219–8227
- 22 Pilon, M. *et al.* (1995) Functional domains of the ferredoxin transit sequence involved in chloroplast import. *J. Biol. Chem.* 270, 3882–3893
- 23 Bolter, B. *et al.* (1998) A protein import receptor in pea chloroplasts, Toc86, is only a proteolytic fragment of a larger polypeptide. *FEBS Lett.* 441, 59–62
- 24 Bauer, J. *et al.* (2000) The major protein import receptor of plastids is essential for chloroplast biogenesis. *Nature* 403, 203–207
- 25 Karlin-Neumann, G.A. and Tobin, E.M. (1986) Transit peptides of nuclear-encoded chloroplast proteins share a common amino acid framework. *EMBO J.* 5, 9–13
- 26 von Heijne, G. *et al.* (1989) Domain structure of mitochondrial and chloroplast targeting peptides. *Eur. J. Biochem.* 180, 535–545
- 27 Claros, M.G. *et al.* (1997) Prediction of N-terminal protein sorting signals. *Curr. Opin. Struct. Biol.* 7, 394–398
- 28 Endo, T. and Schatz, G. (1988) Latent membrane perturbation activity of a mitochondrial precursor protein is exposed by unfolding [published erratum appears in *EMBO J.* (1988) 7, 1915]. *EMBO J.* 7, 1153–1158
- 29 Reitevold, A.G. *et al.* (1995) Non-bilayer lipids are required for efficient protein transport across the plasma membrane of *E. coli*. *EMBO J.* 14, 5506–5513
- 30 van 't Hof, R. *et al.* (1991) Lipid-peptide interactions between fragments of the transit peptide of ribulose-1,5-bisphosphate carboxylase/oxygenase and chloroplast membrane lipids. *FEBS Lett.* 291, 350–354
- 31 Pinnaduwa, P. and Bruce, B.D. (1996) *In vitro* interaction between a chloroplast transit peptide and chloroplast outer envelope lipids is sequence-specific and lipid class-dependent. *J. Biol. Chem.* 271, 32907–32915
- 32 van't Hof, R. *et al.* (1993) The transit sequence mediates the specific interaction of the precursor of ferredoxin with chloroplast envelope membrane lipids. *J. Biol. Chem.* 268, 4037–4042
- 33 Douce, R. and Joyard, J. (1990) Biochemistry and function of the plastid envelope. *Annu. Rev. Cell Biol.* 6, 173–216
- 34 Dabney-Smith, C. *et al.* (1999) The C terminus of a chloroplast precursor modulates its interaction with the translocation apparatus and PIRAC. *J. Biol. Chem.* 274, 32351–32359
- 35 Perry, S.E. and Keegstra, K. (1994) Envelope membrane proteins that interact with chloroplastic precursor proteins. *Plant Cell* 6, 93–105
- 36 Ma, Y. *et al.* (1996) Two components of the chloroplast protein import apparatus, IAP86 and IAP75, interact with the transit sequence during the recognition and translocation of precursor proteins at the outer envelope. *J. Cell Biol.* 134, 315–327
- 37 Rensink, W.A. *et al.* (2000) The transit sequence of ferredoxin contains different domains for translocation across the outer and inner membrane of the chloroplast envelope. *J. Biol. Chem.* 275, 10265–10271
- 38 Sveshnikova, N. *et al.* (2000) Toc34 is a preprotein receptor regulated by GTP and phosphorylation. *Proc. Natl. Acad. Sci. U. S. A.* 97, 4973–4978
- 39 Kessler, F. *et al.* (1994) Identification of two GTP-binding proteins in the chloroplast protein import machinery. *Science* 266, 1035–1039
- 40 Seedorf, M. *et al.* (1995) A constituent of the chloroplast import complex represents a new type of GTP-binding protein. *Plant J.* 7, 401–411
- 41 Heins, L. *et al.* (1998) The protein translocation apparatus of chloroplast envelopes. *Trends Plant Sci.* 3, 56–61
- 42 Kindle, K.L. (1998) Amino-terminal and hydrophobic regions of the *Chlamydomonas reinhardtii* plastocyanin transit peptide are required for efficient protein accumulation *in vivo*. *Plant Mol. Biol.* 38, 365–377
- 43 Kindle, K.L. and Lawrence, S.D. (1998) Transit peptide mutations that impair *in vitro* and *in vivo* chloroplast protein import do not affect accumulation of the gamma-subunit of chloroplast ATPase. *Plant Physiol.* 116, 1179–1190
- 44 Rensink, W.A. *et al.* (1998) Domains of a transit sequence required for *in vivo* import in *Arabidopsis* chloroplasts. *Plant Physiol.* 118, 691–699

Acknowledgements

I thank Roz Miltenberger and members of the Bruce Laboratory for their critical comments and for carefully reading the manuscript. I also thank Jean Marc Lancelin, Hans Wienk and Ben de Kruijff for sharing their unpublished results. This work has been supported by grants from the NSF Cell Biology Program.

- 45 Lawrence, S.D. and Kindle, K.L. (1997) Alterations in the *Chlamydomonas* plastocyanin transit peptide have distinct effects on *in vitro* import and *in vivo* protein accumulation. *J. Biol. Chem.* 272, 20357–20363
- 46 Schatz, G. and Dobberstein, B. (1996) Common principles of protein translocation across membranes. *Science* 271, 1519–1526
- 47 May, T. and Soll, J. (2000) 14-3-3 proteins form a guidance complex with chloroplast precursor proteins in plants. *Plant Cell* 12, 53–64
- 48 Komiya, T. *et al.* (1997) Binding of mitochondrial precursor proteins to the cytoplasmic domains of the import receptors Tom70 and Tom20 is determined by cytoplasmic chaperones. *EMBO J.* 16, 4267–4275
- 49 Waegemann, K. and Soll, J. (1996) Phosphorylation of the transit sequence of chloroplast precursor proteins. *J. Biol. Chem.* 271, 6545–6554
- 50 Muslin, A.J. *et al.* (1996) Interaction of 14-3-3 with signaling proteins is mediated by the recognition of phosphoserine. *Cell* 84, 889–897
- 51 Pilon, M. *et al.* (1992) Kinetic analysis of translocation into isolated chloroplasts of the purified ferredoxin precursor. *FEBS Lett.* 302, 65–68
- 52 Pilon, M. and Schekman, R. (1999) Protein translocation: how Hsp70 pulls it off. *Cell* 97, 679–682
- 53 Marshall, J.S. *et al.* (1990) Identification of heat shock protein hsp70 homologues in chloroplasts. *Proc. Natl. Acad. Sci. U. S. A.* 87, 374–378
- 54 Ivey, R.A. and Bruce, B.D. (2000) *In vivo* and *in vitro* interaction of DnaK and a chloroplast transit peptide. *Cell Stress Chaperones* 5, 62–71
- 55 Ivey, R.A., 3rd *et al.* (2000) Identification of a hsp70 recognition domain within the rubisco small subunit transit peptide. *Plant Physiol.* 122, 1289–1300
- 56 Stahl, T. *et al.* (1999) Tic40, a new 'old' subunit of the chloroplast protein import translocon. *J. Biol. Chem.* 274, 37467–37472
- 57 Zhang, X.P. *et al.* (1999) Interaction of mitochondrial presequences with DnaK and mitochondrial hsp70. *J. Mol. Biol.* 288, 177–190
- 58 Bolter, B. *et al.* (1998) Origin of a chloroplast protein importer. *Proc. Natl. Acad. Sci. U. S. A.* 95, 15831–15836
- 59 Reumann, S. and Keegstra, K. (1999) The endosymbiotic origin of the protein import machinery of chloroplast envelope membranes. *Trends Plant Sci.* 4, 302–307
- 60 Reumann, S. *et al.* (1999) The evolutionary origin of the protein-translocating channel of chloroplast envelope membranes: identification of a cyanobacterial homolog. *Proc. Natl. Acad. Sci. U. S. A.* 96, 784–789
- 61 McFadden, G.I. (1999) Endosymbiosis and evolution of the plant cell. *Curr. Opin. Plant Biol.* 2, 513–519
- 62 de Souza, S.J. *et al.* (1996) Intron positions correlate with module boundaries in ancient proteins. *Proc. Natl. Acad. Sci. U. S. A.* 93, 14632–14636
- 63 Quigley, F. *et al.* (1988) Intron conservation across the prokaryote-eukaryote boundary: structure of the nuclear gene for chloroplast glyceraldehyde-3-phosphate dehydrogenase from maize. *Proc. Natl. Acad. Sci. U. S. A.* 85, 2672–2676
- 64 Gregerson, R.G. *et al.* (1994) Genomic structure, expression and evolution of the alfalfa aspartate aminotransferase genes. *Plant Mol. Biol.* 25, 387–399
- 65 Long, M. *et al.* (1996) Exon shuffling and the origin of the mitochondrial targeting function in plant cytochrome c1 precursor. *Proc. Natl. Acad. Sci. U. S. A.* 93, 7727–7731
- 66 Liaud, M.F. *et al.* (1990) Differential intron loss and endosymbiotic transfer of chloroplast glyceraldehyde-3-phosphate dehydrogenase genes to the nucleus. *Proc. Natl. Acad. Sci. U. S. A.* 87, 8918–8922

Orchestrating cell division

In the April to May transition of this year, a conference in Roscoff, France* paid tribute to Yoshio Masui one of the founders of the cell-cycle field and was a forum for discussion of some of the central features that cause cells to transit from one cell-cycle stage to another. Although this meeting was billed as 'The Cell Division Cycle', the choice of speakers clearly pointed to an intense interest in the architecture and regulation of M phase.

Are centrosomes an essential structure?

Centrosomes are thought to be required for the assembly of the mitotic spindle and thus for correct chromosome segregation. This view has been challenged lately by a number of studies that highlight alternative mechanisms for the formation of the spindle and the maintenance of its bipolarity. Whether or not there is an absolute need for a centrosome is controversial, and work presented by K. Sluder (Worcester, USA) fuelled the debate. He described a delicate microsurgical procedure developed to cut cells in two just between the nucleus and centrosome. The resulting acentrosomal karyoplasts are able to form a bipolar spindle and can go through a round of chromosomal division but cannot complete the final stages of cytokinesis. These binucleate acentrosomal karyoplasts are unable to enter the next cell cycle and consequently arrest in G1 phase.

Kirsten C. Sadler and Maria do Carmo Avides

Other observations do point to a central role of the centrosome in the establishment of a functional bipolar mitotic spindle. As an example, *Drosophila* mutants in key centrosomal components display gross spindle morphological and functional defects. One such mutant, *dd4*, a mutant in the *spc98* *Drosophila* homologue dGRIP91, has abnormal mitotic spindles with a reduced number of microtubules (D. Glover, Cambridge, UK; Fig. 1).

It is often said that a book is better than a movie, but, in cell biology, it is just the opposite. Many presentations contained spectacular movies of gyrating spindles, dancing centrosomes and dividing chromosomes. For example, the movies shown by M. Bornens (Paris, France) that accompanied a recent publication were certainly more entertaining and informative than the corresponding print version¹. GFP-labelled centrin labels the mother centriole more brightly than the daughter, providing a means to discriminate the two as they perform their separation dance during anaphase. The daughter centriole leaps and twitches around the unmoving mother and, if she

dances near the midbody, the connection between two recently formed daughter cells is severed and the cells are free to move apart.

Until recently, the process of centrosome duplication has been poorly understood, and the first insights came from work on the budding yeast spindle pole body (SPB). It soon became apparent that the proteins involved in SPB duplication had functional homologues in higher eukaryotes. One such example is the *S. cerevisiae* kinase MPS1, also required for the spindle integrity checkpoint². Two informative MPS1 mutant alleles, *mps1-7* and *mps1-8*, were described by M. Winey (Boulder, USA). *Mps1-7* mutants are defective in the spindle checkpoint and show abnormalities in meiotic chromosome segregation, implying a new function for MPS1 during meiosis. *Mps1-8* defects are exclusively on the duplication of the SPB and relate to the structure of the half-bridge.

A link between regulatory and structural cell-cycle events was provided by E. Nigg (Martinsreid, Germany). The nuclear cell cycle and the centrosome cycle appear to be linked through

Kirsten Sadler is in the Dept of Molecular Biology and Genetics, Bosphorus University, Istanbul, Turkey; and Maria do Carmo Avides is in the Dept of Genetics, University of Cambridge, Cambridge, UK CB23EH.
E-mail: kirsten_sadler_phd98@post.harvard.edu

*The Jacques Monod Conference on the Cell Division Cycle; Roscoff, France; 30 April – 3 May, 2000. Organized by Eric Nigg and Bernard Maro.