1-1-2006

Novel pathways for glycoprotein import into chloroplasts

L. Faye

H. Daniell

University of Central Florida

Plant Biotechnology Journal (2006) **4**, pp. 275–279 doi: 10.1111/j.1467-7652.2006.00188.x

News and views

Novel pathways for glycoprotein import into chloroplasts

L. Faye[1] and H. Daniell[2]

[1]*CNRS UMR 6037, IFRMP 23, GDR 2590, Université de Rouen, 76821 Mont Saint Aignan, France*
[2]*Department of Molecular Biology and Microbiology, University of Central Florida, Biomolecular Science Building 20, Room 336, Orlando, Florida 32816-2364, USA*

Received 18 January 2006;
revised 19 January 2006;
accepted 19 January 2006.
**Correspondence* (fax +1 407 823 0956;
e-mail daniell@mail.ucf.edu)

Summary

Although the chloroplast contains its own genome, majority of its protein components are encoded by nuclear genes and must be imported post-translationally. In general, proteins synthesized by cytosolic ribosomes are post-translationally targeted to the chloroplast through interactions between their N-terminal transit sequence and protein translocon Toc/Tic complexes in the chloroplast membranes. An alternative pathway that mediates post-translational delivery of proteins to the chloroplast via the secretory pathway was recently described. This pathway provides new opportunities for complementation of the chloroplast protein maturation machinery with chaperones needing endoplasmic reticulum and/or Golgi typical maturations such as N-glycosylation for their biological activity or using chloroplasts as a storage compartment for glycoproteins.

Abbreviations: CAH1; an α-carbonic anhydrase located in the chloroplast stroma of *Arabidopsis thaliana*. RB60; a protein disulfide isomerase localized in plant chloroplast.

Chloroplasts are semiautonomous organelles in the plant cell. However, a vast majority of proteins functioning in the chloroplasts are encoded by the nucleus and imported from the cytoplasm. Plant chloroplast genomes encode only 60–200 proteins (Leister, 2003), while an estimated 2500–3500 proteins functioning in the chloroplasts (Arabidopsis Genome Initiative, 2000) are imported from the cytosol. To date, only a small fraction of the chloroplast proteome has been experimentally characterized in *Arabidopsis* (Peltier *et al*., 2002; Froehlich *et al*., 2003), leaving the rest for speculations and predictions. Most nuclear-encoded chloroplast proteins are synthesized in the cytosol as precursor proteins, bearing an N-terminal transit sequence that contains the targeting information, for transport into the chloroplast. Chloroplast transit peptides are rich in hydrophobic, hydroxylated and positively charged amino acid residues and are deficient in acidic amino acids. Molecular chaperones and other cytosolic factors are involved in the import process: maintaining precursors in an import-competent conformation, guiding them to the organelle and assisting them in their transport through the double-membrane envelope of the higher plant chloroplasts (Glaser and Soll, 2004). The transit peptide interacts specifically with the chloroplast membrane lipids and the translocons of the outer and the inner envelope membranes, Toc (translocon at the outer envelope membrane of the chloroplast) and Tic (translocon at the inner envelope membrane of the chloroplast) protein complexes (Figure 1a). This transit peptide is cleaved off inside the organelle by a highly specific stromal-processing peptidase almost immediately after the N-terminus of the translocating preprotein emerges from the Tic complex into the stroma. The mature proteins are then assembled by molecular chaperones into functional oligomeric protein complexes, whereas the cleaved targeting peptides, potentially destructive for biological membranes, are degraded inside the organelle by a newly identified signal peptide-degrading zinc metalloprotease (Bedard and Jarvis, 2005). A majority of chloroplast-targeted proteins contain a transit peptide that must be recognized by a chloroplast-specific translocation machinery to facilitate their entry into the chloroplasts. However, the outer membrane envelope proteins lack such transit peptides and are targeted to the chloroplast surface via an internal information (Soll and Schleiff, 2004). Interestingly, there are also 'twin' or 'ambiguous' targeting peptides that are recognized by the import apparatus of both chloroplasts and mitochondria (Peeters and Small, 2001; Silva-Filho, 2003). Finally, it has been proposed that other import mechanisms exist (Soll and Schleiff, 2004), but experimental data are very limited.

In a recent paper published in *Nature Cell Biology* (Villarejo *et al*., 2005), a very strong evidence is provided for an

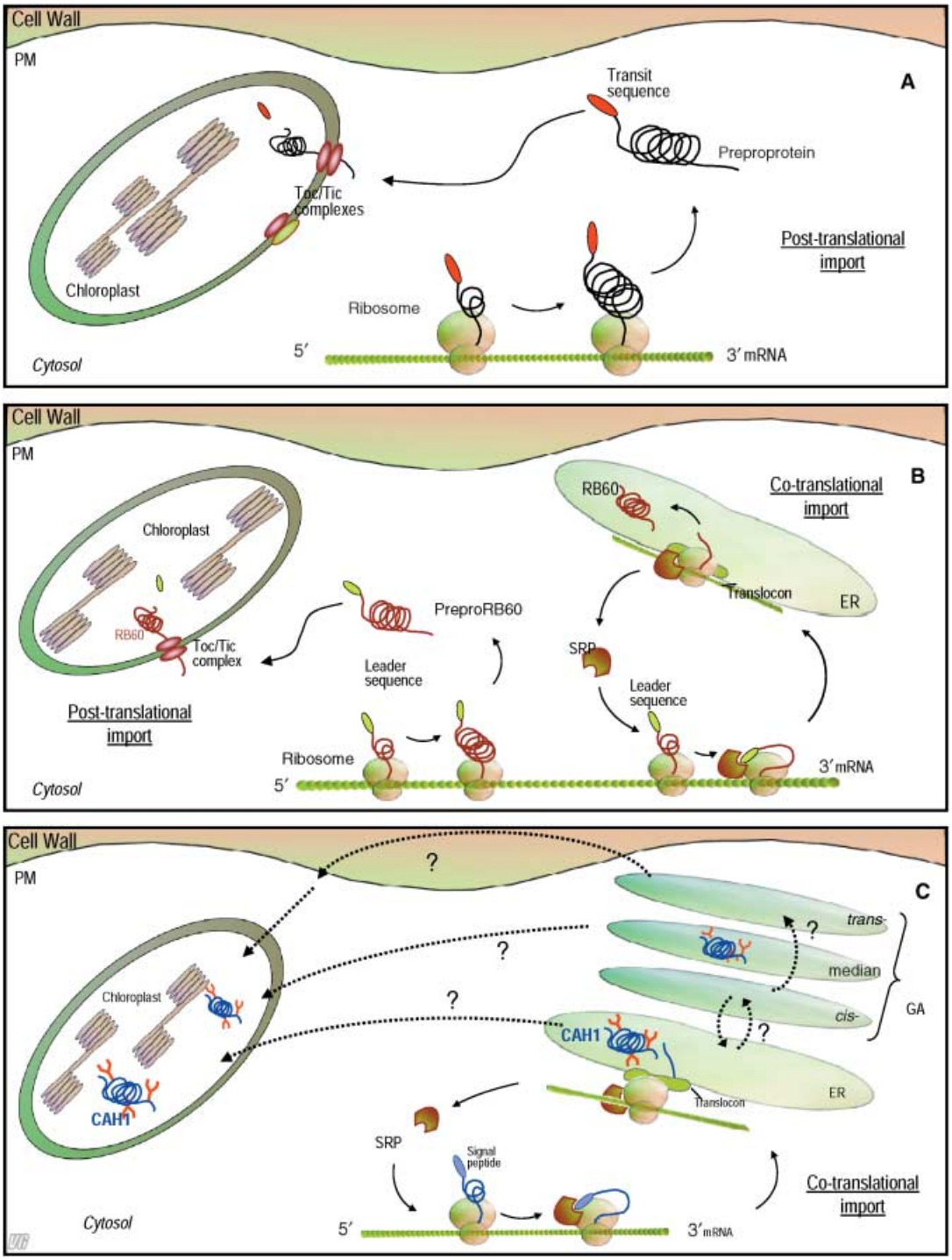

Figure 1 Proteins encoded by nuclear genes are imported into the higher plant chloroplasts by at least two different mechanisms. In the classical protein import process (illustrated in panel A), chloroplast proteins synthesized by cytosolic ribosomes are post-translationally targeted to the chloroplast through interactions between their terminal transit sequence and protein translocon complexes in the outer (Toc apparatus) and inner (Tic apparatus) envelope membranes of the chloroplast. This mechanism is active for proteins exclusively located in the chloroplast and proteins located in the chloroplast and the secretory system. In the latter case, as illustrated by RB60 (Levitan *et al*., 2005) (panel B), the N-terminal leader sequence interacts with either the signal recognition particle (SRP) for an endoplasmic reticulum (ER) co-translational targeting or the Toc/Tic machinery for a post-translational transport from the cytosol to the chloroplast. According to a second transport mechanism, as illustrated with CAH1 in Villarejo *et al*. (2005), a new family of chloroplast proteins is synthesized on membrane-bound ribosomes, co-translationally inserted into the ER lumen via their N-terminal signal peptide and finally targeted to the chloroplast after a transport downstream the secretory system to the late Golgi compartment (panel C). A KDEL endoplasmic reticulum retention signal is present at the C-terminus of RB60. Y (in red); N-glycan (here the use of three of the five potential glycosylation sites of CAH1 should be considered as a working hypothesis).

alternative pathway that mediates post-translational delivery of nuclear-encoded proteins to the chloroplast via the secretory system. In this study, an *Arabidopsis* α-carbonic anhydrase (CAH1) is used as a model and many evidences for its steady state location in the chloroplast stroma in *Arabidopsis* leaves are provided by subcellular fractionation, immunogold labelling, and expression of CAH1 fused with green fluorescent protein (GFP).

Surprisingly for a chloroplast protein, the analysis of the amino acid sequence deduced from the CAH1 cDNA reveals a putative 24 amino acid-long N-terminal signal peptide (SP) and a consensus signal peptidase cleavage site as commonly observed for soluble proteins co-translationally inserted in the endoplasmic reticulum (ER) lumen.

This putative SP is functional and either targets GFP fused with a KDEL signal to the ER *in vivo* after transient expression in *Arabidopsis* cells or mediates *in vitro* uptake of CAH1 in dog pancreas microsomes in a transcription/translation assay. In contrast, CAH1 is not taken up by pea chloroplasts in an *in vitro* assay as observed for chloroplast proteins having a transit peptide, and hence transport of this protein to the chloroplast was supposed to most likely occur after its SP-mediated insertion in the ER lumen.

Following the observation that most, if not all, of the five potential N-glycosylation sites of CAH1 are occupied in a dog pancreas microsomes transcription/translation assay, the most convincing evidence for a novel protein-targeting pathway from the secretory system to chloroplasts in plants was obtained from the analysis of the *Arabidopsis* CAH1 glycosylation. Indeed, both recombinant CAH1 expressed in *Arabidopsis* protoplasts and natural CAH1 from purified stroma fractions are N-glycosylated with β1,2-xylose and α1,3-fucose containing complex N-glycans. N-glycosylation starts with the addition of a precursor oligosaccharide on specific Asn residues during the co-translational insertion of a protein into the ER lumen. After a trimming of this precursor by glycosidases during the transport of the glycoprotein in the plant secretory system complex N-glycans are further processed by glycosyltransferases and particularly by the β1,2-xylosyltransferase and the α1,3-fucosyltransferase located in the medial and *trans*-Golgi compartments (Fitchette-Lainé *et al*., 1994; Pagny *et al*., 2003; Gomord and Faye, 2004). Here, based on its N-glycan composition it is obvious that CAH1 is transported to the chloroplast stroma via the Golgi apparatus (GA). The involvement of the GA in the transport of CAH1 to the chloroplast is further reinforced when brefeldin A (BFA), a fungal toxin that blocks protein secretion between the ER and the Golgi (Staehelin and Driouich, 1997), is used. Indeed, CAH1-GFP transiently expressed in *Arabidopsis* protoplasts is transported to the chloroplasts, whereas it accumulates in the ER and GA in the presence of BFA. This BFA effect is reversible, so that the CAH1-GFP fusion protein is redirected from ER and Golgi to the chloroplast after BFA removal. Similar results were obtained for native endogenous CAH1 using subcellular fractionation studies in *Arabidopsis* suspension-cultured cells treated with BFA (Villarejo *et al*., 2005).

The presence of a functional signal peptide and complex N-glycans on CAH1 together with its BFA-sensitive transport to the chloroplast indicate that this protein is taken up into the ER where SP cleavage and N-glycosylation occur. After transport through the ER and the GA, where its N-glycans are matured, CAH1 is targeted to the chloroplast (Figure 1c). In fact, CAH1 seems to be a member of a large family of proteins trafficking along the same pathway. Indeed, the chloroplast proteome contains many proteins with a predicted SP (Friso *et al*., 2004; Kleffman *et al*., 2004) and after purification, several stroma proteins react with antibodies specific for plant complex N-glycans (Faye *et al*., 1993; Villarejo *et al*., 2005).

This SP-dependent pathway to the chloroplast is different from the one recently described for RB60, a protein disulfide isomerase from *Chlamidomonas reinhardtii* (Figure 1b). RB60 contains a 50 amino acid N-terminal sequence sufficient for its targeting either to the ER or to the chloroplast (Levitan *et al*., 2005). Interestingly, RB60 and probably also rice α-amylase (αAmy3) (Chen *et al*., 2004) show a dual targeting to both the ER and the chloroplasts via two N-terminal extension-dependent and mutually exclusive mechanisms. In contrast, the transport of CAH1 to the chloroplast via the ER and Golgi is strongly reminiscent of an ancestral pathway for protein targeting to the three or four membrane-bound plastids in *Euglena,* apicomplexan parasites such as *Plasmodium*, and diatoms (Slavikova *et al*., 2005). In apicomplexans, the transport of cytoplasmically synthesized proteins to plastids depends on a bipartite N-terminal presequence consisting of an SP that targets proteins to the secretory pathway adjacent to a transit peptide acting as a second signal involved in transport from the secretory pathway to the plastids (van Dooren *et al*., 2001,Waller and McFadden, 2006).

Although the putative CAH1 SP is 24 amino acids long; in their study, Villarejo *et al*. (2005) have fused a 40 amino acid-long N-terminal sequence with GFP–KDEL to show that SP is functional and sufficient for targeting to the ER. The SP–GFP–KDEL fusion is retained in the ER after its co-translational insertion in this compartment. Unfortunately, it is impossible to infer from these results whether or not the 40 amino acid-long N-terminal sequence of CAH1 contains a bipartite presequence that is sufficient for targeting of a reporter protein to

the chloroplast via the secretory system. Alternatively, as observed for *Euglena* and dinoflagellate, it is also possible that the necessary additional information required for CAH1 targeting from the secretory system to the chloroplast is contained in the C-terminal end of this protein that presents some of the signatures found in the transit peptide used in the ancestral pathway.

In the next few years, further studies will be required to define more precisely the pathways and signals required for targeting of a protein, first to the secretory system and next to the chloroplast in higher plants. Whereas the composition of CAH1 N-glycans clearly shows that this glycoprotein is transported from the ER to the medial or *trans*-Golgi cisternae, we do not know if the proteins are transported directly from the Golgi to the chloroplast or via another organelle (plasma membrane (PM), ER) However, the connections frequently observed between the ER and chloroplast membranes in higher plants do not allow to speculate that CAH1 is transported from the ER to the Golgi and then back to the ER before translocation to the chloroplast, as suggested in Villarejo *et al*. (2005). Indeed, based on immunolocalization and subcellular fractionation analysis, glycoproteins with β1,2-xylose- and α1,3-fucose-containing complex N-glycans are not detected in the plant cell ER (Fitchette *et al*., 1994; Pagny *et al*., 2003). If continuity between ER and chloroplast membranes probably does not explain protein targeting from the secretory system to the chloroplast as described for diatoms, CAH1 and other chloroplast proteins of the same family should contain at least two signals, and a transport route to the chloroplast involving vesicles must also be considered. Indeed, it was shown that SP mediates the entry of CAH1 into the endomembrane system. This first signal is cleaved cotranslationally so that in the absence of a second signal, CAH1 would follow a default pathway of secretion and accumulate in the intercellular space. Consequently, a second signal and a new receptor-mediated transport system directing CAH1-containing vesicles budding from the Golgi or the PM to the chloroplast still remain to be elucidated. A more detailed structural analysis of glycans N-linked to CAH1 coupled with immunoelectron microscopy immunocalization would have provided further information on the intracellular transport of this glycoprotein to the chloroplast. For instance, the detection of Lewis a epitopes (Fitchette *et al*., 1997) in CAH1 N-glycans would argue for either a new mechanism of cargo sorting and carrier formation for a *trans*-Golgi network to chloroplast transport or a secretion prior to an endocytose-mediated transport to the chloroplast.

In addition to its use for the hyper-expression of several vaccine antigens (Daniell *et al*., 2001; Molina *et al*., 2004; Watson *et al*., 2004; Koya *et al*., 2005), transgenic chloroplasts have been used for the production of therapeutic proteins, such as human serum albumin (Fernandez-San Millan *et al*., 2003), magainin, a broad-spectrum topical agent, a systemic antibiotic, a wound-healing stimulant and a potential anticancer agent (DeGray *et al*., 2001), interferons, and insulin-like growth factors (Leelavathi and Reddy, 2003; Daniell *et al*., 2004, 2005). These examples show that the chloroplast contains a machinery that allows correct folding and disulfide bond formation, resulting in fully functional proteins. Despite such rapid progress in the use of this organelle for plant molecular pharming, no glycoprotein has been expressed in transgenic chloroplasts, although N- or O-glycosylation is required for stability and functionality of many proteins. The discovery of a pathway where proteins are imported into the chloroplast after their transport downstream the secretory system offers new opportunities to improve the activity of recombinant pharmaceuticals made in chloroplasts by complementation of the maturation machinery of this organelle with proteins that need glycosylation and/or other ER/Golgi post-translational maturations to be fully active. Alternatively, the chloroplast could be used as storage compartment for glycoproteins made and matured in the secretory system.

Acknowledgements

Research in the laboratory of L. Faye is mainly supported by grants from CNRS and the University of Rouen and in the Daniell laboratory by the USDA 3611-21000-017-00D and NIH R01GM 63879. We sincerely thank Veronique Gomord for the art work in Figure 1.

References

Bedard, J. and Jarvis, P. (2005) Recognition and envelope translocation of chloroplast preproteins. *J. Exp. Bot*. **56**, 2287–2320.

Chen, M.H., Huang, L.F., Li, H.M., Chen, Y.R. and Yu, S.M. (2004) Signal peptide-dependent targeting of a rice alpha-amylase and cargo proteins to plastids and extracellular compartments of plant cells. *Plant Physiol*. **135**, 1367–1377.

Daniell, H., Carmona-Sanchez, O. and Burns, B.B. (2004) Chloroplast-derived vaccine antibodies, biopharmaceuticals, and edible vaccines in transgenic plants engineered via the chloroplast genome. In *Molecular Farming*, Chapter 8 (Schillberg, S., ed.), pp. 113–133. Berlin, Germany: Wiley–VCH Verlag.

Daniell, H., Chebolu, S., Kumar, S., Singleton, M. and Falconer, R. (2005) Chloroplast-derived vaccine antigens and other therapeutic proteins. *Vaccine*, **23**, 1779–1783.

Daniell, H., Lee, S.B., Panchal, T. and Wiebe, P.O. (2001) Expression of cholera toxin B subunit gene and assembly as functional oligomers in transgenic tobacco chloroplasts. *J. Mol. Biol*. **311**, 1001–1009.

DeGray, G., Rajasekaran, K., Smith, F., Sanford, J. and Daniell, H. (2001) Expression of an antimicrobial peptide via the chloroplast genome to control phytopathogenic bacteria and fungi. *Plant Physiol*. **127**, 852–862.

van Dooren, G.G., Schwartzbach, S.D., Osafune, T. and McFadden, G.I. (2001) Translocation of proteins across the multiple membranes of complex plastids. *Biochim. Biophys. Acta*. **1541**, 34–53.

Faye, L., Gomord, V., Lainé, A.-C. and Chrispeels, M.J. (1993) Selection of antibodies specific for fucose or xylose containing N-linked glycans from higher plants. *Anal. Biochem*. **209**, 104–108.

Fernandez-San millan, A., Mingo-Castel, A., Miller, M. and Daniell, H. (2003) A chloroplast transgenic approach to hyper-express and purify human serum albumin, a protein highly susceptible to proteolytic degradation. *Plant Biotechnol. J*. **1**, 71–79.

Fitchette-Lainé, A.C., Gomord, V., Cabanes, M., Michalski, J.C., Saint-Macary, M., Foucher, B., Cavelier, B., Hawes, C., Lerouge, P. and Faye, L. (1997) N-glycans harboring the Lewis a epitope are expressed at the surface of plant cells. *Plant J*. **12**, 101–107.

Fitchette-Lainé, A.-C., Gomord, V., Chekkafi, A. and Faye, L. (1994) Distribution of xylosylation and fucosylation in the plant Golgi apparatus. *Plant J*. **5**, 673–682.

Friso, G., Giacomelli, L., Ytterberg, A.J., Peltier, J.B., Rudella, A., Sun, Q. and Wijk, K.J. (2004) In-depth analysis of the thylakoid membrane proteome of *Arabidopsis thaliana* chloroplasts: new proteins, new functions, and a plastid proteome database. *Plant Cell*. **16**, 478–499.

Froehlich, J.E., Wilkerson, C.G., Ray, W.K., McAndrew, R.S., Osteryoung, K.W., Gage, D.A. and Phinney, B.S. (2003) Proteomic study of the *Arabidopsis thaliana* chloroplastic envelope membrane utilizing alternatives to traditional two-dimensional electrophoresis. *J. Proteome Res*. **2**, 413–425.

Glaser, E. and Soll, J. (2004) Targeting signals and import machinery of plastids and mitochondria. In *Molecular Biology and Biotechnology of Plant Organelles*, Chapter 14 (Daniell, H. and Chase, C., eds), pp. 385–417. Dortrecht, the Netherlands: Springer.

Gomord, V. and Faye, L. (2004) Posttranslational modifications of therapeutic proteins in plants. *Curr. Opin. Plant Biol.* **7**, 171–181.

Kleffmann, T., Russenberger, D., von Zychlinski, A., Christopher, W., Sjolander, K., Gruissem, W. and Baginsky, S. (2004) The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions. *Curr. Biol*. **14**, 354–362.

Koya, V., Moayeri, M., Leppla, S.H. and Daniell, H. (2005) Plant-based vaccine: mice immunized with chloroplast-derived anthrax protective antigen survive anthrax lethal toxin challenge. *Infect. Immun.* **73**, 8266–8274.

Leelavathi, S. and Reddy, V.S. (2003) Chloroplast expression of His-tagged GUS fusions: a general strategy to overproduce and purify foreign proteins using transplastomic plants as bioreactors. *Mol. Breed*. **11**, 49–58.

Leister, D. (2003) Chloroplast research in the genomic age. *Trends Genet*. **19**, 47–56.

Levitan, A., Trebitsh, T., Kiss, V., Pereg, Y., Dangoor, I. and Danon, A. (2005) Dual targeting of the protein disulfide isomerase RB60 to the chloroplast and the endoplasmic reticulum. *Proc. Natl Acad. Sci. USA*. **102**, 6225–6230.

Molina, A., Herva-Stubbs, S., Daniell, H., Mingo-Castel, A.M. and Veramendi, J. (2004) High yield expression of a viral peptide animal vaccine in transgenic tobacco chloroplasts. *Plant Biotechnol. J*. **2**, 141–153.

Pagny, S., Bouissonnié, F., Sarkar, M., Follet-Gueye, M.L., Driouich, A., Schachter, H., Faye, L. and Gomord, V. (2003) Structural requirements for *Arabidopsis* beta 1,2xylosyltransferase activity and targeting to the Golgi. *Plant J*. **33**, 189–203.

Peeters, N. and Small, I. (2001) Dual targeting to mitochondria and chloroplasts. *Biochim. Biophys. Acta*. **1541**, 54–63.

Peltier, J.B., Emanuelsson, O., Kalume, D.E., Ytterberg, J., Friso, G., Rudella, A., Liberles, D.A., Soderberg, L., Roepstorff, P., von Heijne, G. and van Wijk, K.J. (2002) Central functions of the lumenal and peripheral thylakoid proteome of Arabidopsis determined by experimentation and genome-wide prediction. *Plant Cell*. **14**, 211–236.

Silva-Filho, M.C. (2003) One ticket for multiple destinations: dual targeting of proteins to distinct subcellular locations. *Curr. Opin. Plant Biol*. **6**, 589–595.

Slavikova, S., Vacula, R., Fang, Z., Ehara, T., Osafune, T. and Schwartzbach, S.D. (2005) Homologous and heterologous reconstitution of Golgi to chloroplast transport and protein import into the complex chloroplasts of *Euglena*. *J. Cell Sci*. **118**, 1651–1661.

Soll, J. and Schleiff, E. (2004) Protein import into chloroplasts. *Nat. Rev. Mol. Cell Biol*. **5**, 198–208.

Staehelin, L.A. and Driouich, A. (1997) Brefeldin A effects plants (Are different Golgi responses caused different sites action?). *Plant Physiol*. **114**, 401–403.

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, **408**, 796–815.

Villarejo, A., Buren, S., Larsson, S., Dejardin, A., Monne, M., Rudhe, C., Karlsson, J., Jansson, S., Lerouge, P., Rolland, N., von Heijne, G., Grebe, M., Bako, L. and Samuelsson, G. (2005) Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast. *Nat. Cell Biol*. **7**, 1124–1131.

Waller, R.F. and McFadden, G.I. (2006) The apicoplast. *Curr. Issues Mol. Biol*. **348**, 1015–1024.

Watson, J., Koya, V., Leppla, S. and Daniell, H. (2004) Expression of *Bacillus anthracis* protective antigen in transgenic chloroplasts of tobacco, a non-food/feed crop. *Vaccine*, **22**, 4374–4384.