# Chapter 14
# The Proteomes of Chloroplasts and other Plastids

**Paul Jarvis**

**Abstract** The post-genomic era of biology has seen a significant shift in focus, from the genes themselves to the proteins they encode. Recent large-scale studies on the proteomes of chloroplasts and others types of plastid have provided significant new insights into the biogenesis, evolution, and functions of these organelles, and have raised some interesting questions. Many of the proteins that define several important sub-organellar compartments (including the envelope and thylakoid membrane systems, the stroma and plastoglobules) have been identified, and this information has been used to make in silico predictions about the entire complement of proteins in each case. Proteomics has revealed that a relatively large number of proteins inside chloroplasts do not possess canonical targeting information (such proteins lack transit peptides for engagement of the general import machinery), and this has led to the elucidation of novel and unusual pathways of chloroplast protein traffic. For example, it is now clear that some Arabidopsis proteins pass through the endoplasmic reticulum and Golgi en route to the chloroplast, and that these proteins may become glycosylated along the way. Comparative studies have been used to characterise organellar proteome changes in response to various environmental cues or genetic perturbations, whilst other approaches have shed light on the oligomerisation and covalent modification of plastidic proteins.

## 14.1 Introduction

The evolution of the modern plant cell involved the acquisition of mitochondria and chloroplasts through endosymbiosis, and it is generally accepted that these organelles are distant relatives of present-day α-proteobacteria and cyanobacteria, respectively (Margulis 1970). Over the course of evolution, the progenitors of mitochondria and chloroplasts relinquished most of their genes to the nuclear genome, so that now >90% of their constituent proteins are translated

Paul Jarvis, Department of Biology, University of Leicester, Leicester LE1 7RH, UK, E-mail: rpj3@le.ac.uk



J. Šamaj and J. Thelen (eds.), *Plant Proteomics*

on cytoplasmic ribosomes, and must engage protein targeting mechanisms that direct them specifically to either organelle (Leister 2003; Bédard and Jarvis 2005; van der Laan et al. 2006). Typically, several thousand different, nucleus-encoded proteins are targeted to these organelles, and it is these proteins that define their distinctive functions.

Chloroplasts are the photosynthetic members of a diverse family of organelles called plastids, which also includes etioplasts (chloroplast precursors that develop in dark-grown, or etiolated, plants), amyloplasts (which are specialised for the storage of starch), and chromoplasts (which accumulate carotenoid pigments) (Whatley 1978; López-Juez and Pyke 2005). Plastids are ubiquitous in plants and algae, and perform numerous essential functions including important steps in the biosynthesis of amino acids, lipids, nucleotides, hormones, vitamins and secondary metabolites, the assimilation of nitrogen and sulphur, and oxygenic photosynthesis (Leister 2003; López-Juez and Pyke 2005). In the latter process, chloroplasts convert energy from sunlight into usable chemical bond energy, and the associated redox reactions lead to the generation of oxygen from water. Chloroplasts are therefore important sites for the production of organic matter and oxygen, and so provide the fuels essential for all higher forms of life (Nelson and Ben-Shem 2004).

Completion of the genome sequencing projects for Arabidopsis, rice and other species, and the development of efficient methods for routine protein identification by mass spectrometry (MS), have together precipitated the onset of the proteomic era; for a discussion of the relevant technologies, see Whitelegge (2003). Due to the overwhelming complexity of cellular proteomes, and the dynamic range limitations associated with analyses on such highly complex mixtures (i.e. the tendency of abundant proteins to mask the presence of less abundant proteins), proteomic studies have tended to focus on isolated subcellular components. In plants, chloroplasts have received considerable attention in this regard (Schröder and Kieselbach 2003; Baginsky and Gruissem 2004; Jarvis 2004; van Wijk 2004; Pan et al. 2005). The value of such proteomic analysis is several-fold: it can confirm the expression and structure of genes predicted based on genome sequence analysis in silico; it can provide information on subcellular and suborganellar protein localisation; it can be used to estimate the abundances of different proteins; and it can even yield information on post-translational modification (PTM) and the composition of multiprotein complexes. Information of this nature will be vital as we move forward into the era of systems biology, especially since it has been estimated that up to 50% of the proteins encoded by the >26,000 genes in the Arabidopsis genome are presently of unknown function (Haas et al. 2005).

The majority of chloroplast proteins bear a cleavable, amino-terminal extension called a transit peptide, which acts as a targeting signal. The transit peptide is recognised by a sophisticated protein import machinery: the so-called TOC and TIC (translocon at the outer / inner envelope membrane of chloroplasts) complexes (Bédard and Jarvis 2005). Since all transit peptides share certain characteristics, it is possible to identify candidate chloroplast proteins through sequence analysis in silico (the TargetP neural network algorithm is a popular method; Emanuelsson et al. 2000). However, transit peptides are not well conserved, and are rather similar

to mitochondrial targeting signals, and so such in silico methods are not completely reliable (Richly and Leister 2004). While it is generally accepted that between 2,000 and 4,000 different proteins are targeted to *Arabidopsis* plastids, there are presently less than 900 *Arabidopsis* entries in one database of experimentally determined plastid proteins (Friso et al. 2004; Peltier et al. 2004a). Thus, it is clear that there is a need for further experimentation, and that the proteomic era of chloroplast biology is far from over.

## 14.2 Proteome Catalogues

Cataloguing aims to identify all of the proteins within a particular cellular or organellar compartment, and so define the functions of that compartment. Although the proteome of a chloroplast is substantially smaller and more manageable than that of an entire cell, it nevertheless comprises several thousand different proteins. Thus, many cataloguing studies have focused on a particular suborganellar compartment: e.g. the internal thylakoid membrane system that bears the photosynthetic complexes, the double-membrane envelope system that surrounds each organelle, the central aqueous matrix, or stroma, and the lipid-containing, thylakoid-associated structures called plastoglobules. Nevertheless, some studies on whole organelles have been done, and these have focused on different plastid types, including chloroplasts, amyloplasts and etioplasts.

### *14.2.1 The Thylakoid Membrane System*

The thylakoids are a complex network of membranous sacks embedded within the stroma of chloroplasts. The membranes themselves harbour the four multiprotein complexes of the photosynthetic light reactions (the photosystems, PSI and PSII, the cytochrome $b_6f$ complex, and the ATP synthase), but also function to form a central aqueous compartment called the thylakoid lumen, which is distinct from the stroma. Difficulties associated with the extraction and analysis of highly hydrophobic membrane proteins have led many thylakoid proteomic studies to focus on lumenal proteins, or proteins peripherally associated with the membranes, which can more easily be resolved by two-dimensional gel electrophoresis (2-DE) (Kieselbach and Schroder 2003). However, the molar ratio between the most abundant and least abundant lumenal proteins has been estimated to be as high as $10^6$, which presents a daunting dynamic-range obstacle to comprehensive analysis (Peltier et al. 2002).

While the existence of some lumenal proteins (e.g. plastocyanin and extrinsic components of PSII) has been known for many years, proteomic analysis has revealed an unanticipated level of complexity. Only a relatively small number of different Arabidopsis lumenal proteins have been identified experimentally, but

information from these sequences has been used to predict the total lumenal proteome in silico, with estimates ranging from ~80 proteins to ~400 proteins (Peltier et al. 2002; Schubert et al. 2002; Kieselbach and Schroder 2003; Westerlund et al. 2003); the actual size of the proteome likely resides somewhere in between these estimates. The predicted proteome contains many unknown or unexpected proteins, suggesting that the compartment has a much broader spectrum of functions than was previously envisaged. In addition to the expected photosynthetic proteins, and those of unknown function, significant numbers of others involved in protein folding, processing and proteolysis, antioxidative defence, and non-photosynthetic redox reactions were present. Some of these proteins likely serve to repair and maintain normal functionality of the photosynthetic complexes, which experience substantial redox stress. The results also revealed that a significant proportion of lumenal proteins (up to 50%, according to one estimate) employ the twin-arginine translocation (TAT) pathway which, unlike the Sec pathway, is able to transport fully folded proteins (Jarvis and Robinson 2004).

The thylakoid membrane itself is dominated by the four photosynthetic complexes, which contain ~100 different proteins in total, and so some early proteomic studies focused on these (Gomez et al. 2002; Whitelegge et al. 2002; Zolla et al. 2002; Whitelegge 2003; Zolla et al. 2003). However, the membrane is also believed to contain many other proteins that are important for the assembly, maintenance and regulation of the photosynthetic apparatus, and which are not themselves components of the complexes. Because strongly hydrophobic proteins are not easily analysed using 2-DE gels, large-scale studies of the thylakoid membrane have employed organic solvent fractionation and other procedures (Friso et al. 2004; Peltier et al. 2004a). These studies achieved near complete coverage of the photosynthetic complexes, but also identified low-abundance components such as those involved in cyclic electron flow around PSI and chlororespiration.

By combining the protein sets derived from the lumenal and membrane proteomic studies mentioned above, and by also including thylakoid proteins identified in classical experiments described in the literature, the total number of experimentally determined thylakoid-associated Arabidopsis proteins is currently about 400 (Peltier et al. 2004a; van Wijk 2004). While some of these may be contaminants from other chloroplast compartments, it is nevertheless interesting to consider the distribution of biological functions: ~30% are involved in photosynthesis, ~25% are of unknown function [including proteins with tetratricopeptide (TPR), pentatricopeptide (PPR), DnaJ and rhodanese domains], ~20% mediate protein translocation, folding, processing and proteolysis, and almost 10% are involved in various aspects of defence against oxidative stress. With respect to protein translocation, it is well documented that nucleus- and chloroplast-encoded proteins are targeted to the thylakoid system through four different targeting pathways, at least three of which are closely related to bacterial transport systems and employ proteinaceous import machineries associated with the thylakoid membrane (Jarvis and Robinson 2004).

### *14.2.2 The Envelope Membrane System*

The envelope is a double-membrane system that forms a semi-permeable barrier between the cytosol and the chloroplast interior. It contains the protein import apparatus responsible for the translocation of nucleus-encoded chloroplast proteins (Bédard and Jarvis 2005), as well as many transporters for the exchange of ions and metabolites (Weber et al. 2005). The envelope also contains a unique biochemical machinery responsible for several important functions, including the synthesis of plastid membrane components and other lipids, and participates in the communication that occurs between the plastid and the nucleus (Jarvis 2003; López-Juez and Pyke 2005).

Like the thylakoid membrane, the envelope proteome is dominated by highly hydrophobic integral membrane proteins, and so proteomic studies of this compartment have employed fractionation procedures other than 2-DE, including differential solvent extraction and multidimensional liquid chromatography (Ferro et al. 2003; Froehlich et al. 2003; Rolland et al. 2003). Up to 50% of the proteins identified as a result of these studies were predicted to have at least one transmembrane domain. Together with the more traditional, single-protein-focused experiments reported in the literature, these proteomic experiments resulted in the identification of more than 400 putative envelope-associated Arabidopsis proteins (Peltier et al. 2004a; van Wijk 2004). Once again, while it is likely that a proportion of these proteins are contaminants from other compartments, it is nonetheless interesting to consider their possible roles: almost 30% of the proteins are of unknown function; 13% mediate protein translocation, folding, processing or degradation; 10% are involved in lipid or fatty acid metabolism; and 9% are transporters of small molecules. Thus, while many of these proteins clearly reflect what are known to be the main functions of the envelope system (functions that are clearly different from those of the thylakoid system discussed earlier), the high proportion of proteins of unknown function indicates that there is still a great deal to be learnt. Interestingly, proteins with similarity to components of the mitochondrial protein import machinery were identified, suggesting the existence of novel, unanticipated protein translocation pathways in the chloroplast envelope (Ferro et al. 2003).

At this point, it should be noted that the intermembrane space, which exists between the two envelope membranes, is considered to be an important suborganellar compartment distinct from the membranes themselves. Unfortunately, very little is known about how proteins are targeted to this space (Kouranov et al. 1999), or indeed about protein targeting to the envelope membranes themselves (Hofmann and Theg 2005; Li and Schnell 2006). While inner envelope membrane proteins possess transit peptides and initially engage the general import machinery (Bédard and Jarvis 2005), much like nucleus-encoded stromal and thylakoidal proteins, little is known about how these proteins reach their final destination. By contrast, the majority of outer envelope membrane proteins do not possess a cleavable targeting signal, and are instead directed to the membrane by virtue of information held in their transmembrane domains. The lack of knowledge concerning envelope targeting mechanisms and signals means that it is presently very difficult to make predictions

concerning the composition of the three individual proteomes in silico. Nevertheless, evidence suggests that the outer membrane is characterised by the presence of beta-barrel proteins, and that the inner membrane is dominated by transporters with multiple, helical transmembrane domains (Koo and Ohlrogge 2002; Schleiff et al. 2003).

By analysing carefully collected sets of integral proteins from the inner envelope and thylakoid membranes (identified on the basis of published information), the respective proteomes were found to have quite different characteristics (Sun et al. 2004). On average, thylakoid proteins were smaller and more acidic than envelope proteins, and additionally contained fewer cysteine residues. The larger average size of the envelope set presumably reflects the presence of numerous transporters with multiple membrane spans (Weber et al. 2005), whereas the pI differences may be related to pH differences between the compartments (protons are concentrated in the thylakoid lumen as part of the photosynthetic mechanism). Cysteine residues have unique properties, including the ability to engage directly in redox reactions, and so the reduced cysteine content of thylakoidal proteins might be a consequence of the redox-associated stresses they experience, and may represent a measure to reduce oxidative damage (Sun et al. 2004). These different characteristics should facilitate the formulation of effective methods for the prediction and discrimination of envelope and thylakoidal proteins in the future.

### *14.2.3 The Stroma*

The chloroplast stroma is the aqueous matrix that surrounds the thylakoid membranes and fills the organellar interior. It is the site of the carbon reactions of photosynthesis (the Calvin cycle) and other metabolic pathways, and the location for components of the endogenous genetic system of the plastid. One in silico study estimated that the stroma has the potential to contain over 3,000 different proteins, which is equivalent to ~80% of the total theoretical proteome of plastids (Sun et al. 2004). However, it should be noted that many of these proteins may associate permanently or transiently with the thylakoids or the inner envelope membrane, via protein–protein, electrostatic or hydrophobic interactions, or even lipid anchors, all of which are features that were not screened for in the analysis. Interestingly, evidence suggests that the acetyl-coenzyme A carboxylase complex is envelope-associated, which would place it very close to the site where chloroplast-synthesised fatty acids are used or exported to the cytosol (Rolland et al. 2003). The metabolic channelling advantages of such associations are obvious, and there may be many other similar examples of membrane association that are not easily detected.

Despite its obvious importance, the stroma had not been subjected to large-scale proteomic analysis until recently (Peltier et al. 2006). Using a novel form of 2-DE [involving native gels in the first dimension, instead of isoelectric focusing (IEF) as is more common] a large number of stromal proteins were resolved prior to identification. Well over 200 proteins were identified, and these fall into the following

functional categories: 26% mediate protein synthesis, folding, proteolysis and sorting; 12% are involved in primary carbon metabolism, including Calvin cycle and oxidative pentose phosphate pathway enzymes; 11% are of unknown function; while 7%, 6%, 4% and 4% mediate the biosynthesis of amino acids, tetrapyrroles, nucleotides and lipids, respectively. Interestingly, the relative concentrations of the identified proteins were estimated by quantifying spot intensities on the 2-DE gels. As expected, proteins involved in primary carbon metabolism were found to constitute the majority (~75%) of the total stromal mass. Those dedicated to protein synthesis, biogenesis and fate represented nearly 10%, whereas those involved in nitrogen and sulphur assimilation made up ~8%. Proteins involved in other biosynthetic pathways, such as those for fatty acids, amino acids, nucleotides and tetrapyrroles, each represented less than 1% of the total mass.

Other, more focused proteomic studies of the stromal compartment have also been conducted. For example, in a thorough study of the 70S ribosome of spinach chloroplasts, all of the proteins of the 50S and 30S subunits were mapped by 2-DE and identified (Yamaguchi and Subramanian 2000; Yamaguchi et al. 2000). The ribosome was shown to comprise 59 different proteins (33 in the 50S subunit and 25 in the 30S subunit, as well as a ribosome recycling factor found exclusively in the 70S holocomplex), of which 53 are orthologues of bacterial ribosomal proteins and 6 are plastid-specific proteins. It was proposed that the latter components evolved to perform functions unique to plastid translation and its regulation, such as protein targeting to the thylakoid membrane and the mediation of control by nuclear factors.

Interestingly, two 30S ribosomal subunit proteins were identified as targets for regulation by the stromal thioredoxin system (Balmer et al. 2003). This system is composed of ferredoxin, ferredoxin-thioredoxin reductase, and thioredoxin, and it links light to the regulation of photosynthetic enzymes and other plastidic processes such as lipid biosynthesis. Electrons flow from ferredoxin to thioredoxin, which, in its reduced state, regulates the activity of target proteins through the reduction of specific disulphide groups. To gain a more complete picture of the targets of the system, proteomics strategies have been adopted (Motohashi et al. 2001; Balmer et al. 2003; Hisabori et al. 2005). A key feature of the regulatory mechanism is the formation of a transient, intermolecular disulphide bridge between thioredoxin and the target protein. This linkage is then reduced by a second cysteine residue within thioredoxin, releasing the target protein in its reduced and active form. However, if the second thioredoxin cysteine residue is mutated (e.g. to serine or alanine), reduction of the intermolecular linkage is prevented, locking the target in its association with thioredoxin and providing a convenient method for target identification. Using such approaches, ~30 potential targets have been identified, including components acting in established thioredoxin-regulated pathways (e.g. the Calvin cycle, nitrogen and sulphur metabolism, protein synthesis, and the oxidative pentose phosphate pathway) and others not previously recognised as thioredoxin targets (e.g. tetrapyrrole biosynthesis, protein folding, assembly and degradation, starch degradation, DNA replication and transcription, and plastid division). More recently, it has been reported that the reach of this largely stromal regulatory network extends into the

thylakoid lumen, where it targets components of the photosynthetic electron transport chain (Motohashi and Hisabori 2006). It should be noted, however, that the biological significance of many of these thioredoxin interactions remains to be established (Hisabori et al. 2005).

### *14.2.4 Plastoglobules*

Another plastidic compartment that has been subjected to proteome analysis is the plastoglobule. Plastoglobules are lipid-containing bodies found in different plastid types, including chloroplasts and chromoplasts. Biochemical analyses of these structures revealed that they contain a variety of different lipidic compounds, including galactolipids, fatty acids, carotenoids, tocopherols and plastoquinone, and so they were thought to perform lipid storage functions. Studies using one-dimensional electrophoresis revealed that they contain more than a dozen different proteins, but at the time most were of unknown identity (Kessler et al. 1999). Recently, proteomic analysis led to the identification of about 30 different plastoglobule proteins, and in so doing demonstrated that the role of the plastoglobule is far more complex than was previously envisaged (Vidi et al. 2006; Ytterberg et al. 2006). In addition to the plastid lipid-associated protein (PAP) / fibrillin family of plastoglobule-specific proteins, which are thought to perform a structural role by binding to the surface of the globules and preventing their coalescence, a number of enzymes were identified. The data indicated that the plastoglobules of Arabidopsis chloroplasts play an active role in the synthesis of their lipophilic constituents, including alpha-tocopherol (vitamin E), which has an important anti-oxidative function in the thylakoids. Similarly, the plastoglobules of pepper chromoplasts (which primarily accumulate carotenoids) were found to contain enzymes of carotenoid biosynthesis. In chloroplasts, plastoglobules were shown to be directly and permanently coupled to the thylakoids, which led to the conclusion that their lipidic contents are in a dynamic equilibrium with thylakoid membranes (Austin et al. 2006). This suggests that plastoglobules play an important role in the synthesis of thylakoid constituents.

### *14.2.5 Whole Organelles*

An extensive study of the whole chloroplast proteome of Arabidopsis was recently reported (Kleffmann et al. 2004). Using a comprehensive series of fractionation procedures to overcome dynamic range limitations, a total of ~600 different proteins were identified. Interestingly, more than 30% of these proteins are of unknown function. In a recent, similarly comprehensive study of the Arabidopsis mitochondrial proteome, almost 20% of the identified proteins were of unknown function (Heazlewood et al. 2004; see also Chap. 15 by Millar, this volume), and so it seems

that we have some way to go before the functions of these two organelles are fully understood. The chloroplast study achieved nearly complete identification coverage for major metabolic pathways, such as the Calvin cycle, but only partial coverage for pathways that are not abundantly expressed (Kleffmann et al. 2004). Parallel RNA profiling experiments revealed a correlation between transcript level and protein abundance for some metabolic pathways, but not others, implying the utilisation of different regulatory mechanisms in different pathways.

Proteome studies have also been conducted on amyloplasts from wheat endosperm (Andon et al. 2002; Balmer et al. 2006a). Amyloplasts are non-photosynthetic plastids specialised for the synthesis and long-term storage of starch. In mature wheat seeds, over 80% of the kernel volume is occupied by starchy endosperm, which is dominated by amyloplasts. By comparing MS data with sequence information from wheat and related cereals, particularly rice, a total of ~400 different proteins were identified (Andon et al. 2002; Balmer et al. 2006a). As expected, most of the enzymes of starch biosynthesis were detected. However, the results revealed a surprisingly broad spectrum of biosynthetic capabilities: like chloroplasts, amyloplasts were found to be endowed with enzymes for the assimilation of nitrogen and sulphur, and for the biosynthesis of amino acids, fatty acids and tetrapyrroles. Nevertheless, when a profile of the functions of the identified amyloplast proteins was compared with a corresponding profile for the chloroplast proteome, some significant differences were observed. For example, the amyloplast proteome contains a substantially higher proportion of proteins involved in carbon, nitrogen and sulphur metabolism, and transport processes. By contrast, the chloroplast proteome contains proportionally more proteins of unknown function, which presumably reflects the more complex set of activities mediated by photosynthetic plastids. Interestingly, while several components of the TOC/TIC import machinery were identified, no proteins of the plastidic ribosome were detected, suggesting that most proteins needed by the developing amyloplast are encoded in the nucleus. This makes sense, since the plastid genome is dominated by genes for components of the photosynthetic apparatus.

It was recently reported that, as in chloroplasts, a fully functional thioredoxin regulatory network is operational in amyloplasts (Balmer et al. 2006b). However, ferredoxin is not reduced by light in this case, but instead by metabolically generated NADPH (via ferredoxin-NADP reductase). As with the chloroplast system, potential targets were identified using a proteomics approach, and were found to act in a range of different processes, including starch metabolism, the biosynthesis of amino acids and lipids, and transport processes. It was suggested that the thioredoxin system might enable amyloplasts to indirectly perceive and respond to light, through the generation of reducing power (leading to reduced thioredoxin) upon the arrival of newly synthesised photosynthate in sink tissues. This would enable the amyloplasts to couple their activities with photosynthesis taking place in leaves.

An analysis of etioplasts in dark-grown rice plants revealed a proteome broadly consistent with what one would expect of plastids in heterotrophic tissue, along with some novel functions (von Zychlinski et al. 2005). Interestingly, whereas the data for amyloplasts and the heterotrophic plastids of tobacco culture cells indicated that

these organelles import primarily hexose phosphates (Andon et al. 2002; Baginsky et al. 2004; Balmer et al. 2006a), etioplasts were found to contain triose phosphate translocators, but no hexose phosphate translocator. Thus, while the metabolic functions of etioplasts are similar to those of other heterotrophic plastids, it would appear that etioplasts nevertheless share significant similarities with chloroplasts (the organelles into which they might ultimately develop), since chloroplasts also transport primarily three-carbon sugars (Weber et al. 2005). Consistent with this notion, over 70% of the 240 identified rice etioplast proteins were found to have clear homologues in the Arabidopsis chloroplast proteome (von Zychlinski et al. 2005).

## 14.3 Protein Targeting Issues

Surprisingly, when the ~600 proteins identified in the whole chloroplast proteome study mentioned above were analysed using the TargetP program (Emanuelsson et al. 2000), only ~60% were predicted to have a chloroplast transit peptide (Kleffmann et al. 2004). Of the remainder, ~40 were predicted to have a mitochondrial presequence, ~50 were predicted to have a signal peptide for translocation into the endoplasmic reticulum (ER), and ~140 were predicted to have no cleavable targeting signal at all. Many of these "misplaced" chloroplast proteins were shown to be of cyanobacterial origin, or are encoded by low abundance transcripts, and so it seems unlikely that they are all simply contaminants from other cellular compartments. In addition, while it is doubtful that the TargetP predictions are accurate for all of these proteins (Emanuelsson et al. 2000; Richly and Leister 2004), it seems equally unlikely that they are wholly incorrect, and so the data suggested that protein targeting to chloroplasts may be more complex than was previously envisaged (Jarvis 2004; Bédard and Jarvis 2005). The existence of mitochondrial proteins with non-canonical targeting signals is well documented (van der Laan et al. 2006), and so it is much less surprising that TargetP predicted only ~50% of the proteins identified in Arabidopsis mitochondria (Heazlewood et al. 2004).

Until recently, all nucleus-encoded chloroplast proteins were thought to arrive in the organelle via one of two post-translational targeting mechanisms: (1) active transport of transit peptide-bearing proteins through the TOC and TIC import complexes (Bédard and Jarvis 2005); (2) direct insertion into the cytosolically exposed outer envelope membrane (Hofmann and Theg 2005). The former mechanism mediates the import of numerous proteins destined for interior locations within chloroplasts (such as the inner envelope membrane, stroma and thylakoids) whereas the latter is exclusively associated with integral proteins of the outer membrane. Thus, it had been assumed that all nucleus-encoded chloroplast interior proteins must have a cleavable, amino-terminal transit peptide.

The first evidence for a more complicated picture of chloroplast protein biogenesis was provided by analyses of the *Arabidopsis* envelope proteome, which identified a protein with homology to quinone oxidoreductases from bacteria, yeast and

animals (Miras et al. 2002; Ferro et al. 2003); the protein was termed ceQORH, for chloroplast envelope quinone oxidoreductase homologue. Intriguingly, although alignments of ceQORH with its bacterial, yeast and animal counterparts revealed no amino-terminal extension, or transit peptide, the protein was nevertheless found associated with the inner envelope membrane (inner membrane proteins normally have a transit peptide). Further studies revealed that the extreme amino-terminus of ceQORH is not required for efficient chloroplast targeting, and instead identified an internal sequence of ~40 residues that controls localisation (Miras et al. 2002). Although this internal targeting signal does not bear any obvious resemblance to standard transit peptides, the protein may nevertheless follow the TOC/TIC-mediated import route. Mitochondrial proteins with internal targeting signals also exist, and, while these proteins are recognised by a different primary receptor, they do pass through the same core translocon complex as proteins with cleavable pre-sequences (van der Laan et al. 2006). However, in the absence of any relevant experimental data, it remains to be determined how ceQORH gains access to the chloroplast interior. More recently, another inner membrane protein, IEP32 (inner envelope protein of 32 kDa), was identified and found to lack a canonical transit peptide (Nada and Soll 2004). Import of this protein was shown to proceed at low ATP concentrations and without the assistance of key components of the TOC machinery, suggesting that it follows an import pathway distinct from that used by precursors with transit peptides.

The identification of so many proteins with predicted signal peptides (for ER translocation) in the whole chloroplast proteome was a surprising result (Kleffmann et al. 2004). While some of these proteins may have been contaminants from other compartments, or proteins with amino-terminal transmembrane domains that target them to the outer envelope membrane (such domains are frequently misidentified as signal peptides; Hofmann and Theg 2005), it seems unlikely that they can all be explained away in this fashion. Intriguingly, close physical associations between the ER and the outer envelope membrane have been documented over many years (Crotty and Ledbetter 1973; Whatley et al. 1991), and recent microscopy studies have shown that the envelope exhibits profound structural fluidity (Kwok and Hanson 2004). Furthermore, biochemical interactions between the ER and envelope membrane systems are an essential part of normal lipid metabolism (Awai et al. 2006). Chloroplast protein traffic through the endomembrane system is well documented in algae that have complex plastids; i.e. plastids that are surrounded by three or four membranes, instead of the usual two, and which were derived from algae with simple plastids through secondary endosymbioses (van Dooren et al. 2001). Chloroplast proteins in these species typically have a bipartite targeting signal, comprised of an amino-terminal signal peptide fused to a more-or-less standard chloroplast transit peptide. The signal peptide directs the chloroplast precursor into the ER, where it is removed, and the protein then passes through the endomembrane system until it arrives at the plastid, at which point the transit peptide mediates chloroplast import in the usual fashion (van Dooren et al. 2001). This type of targeting pathway makes sense in these organisms, due to the complex nature of their plastids and the likely autogenous origin of the outer organellar membrane, but would seem unnecessary in higher plants.

Nevertheless, indirect evidence suggesting the existence of a protein transport pathway to chloroplasts through the ER and Golgi in plants has existed for some time: i.e. glycosylated proteins and proteins with predicted signal peptides have been found to localise in plastids (Gaikwad et al. 1999; Chen et al. 2004; Asatsuma et al. 2005). More recently, firm evidence for such a targeting pathway was presented (Villarejo et al. 2005). The Arabidopsis carbonic anhydrase 1 (CAH1) protein, which controls hydration of carbon dioxide, was found to localise in the chloroplast stroma, despite the fact that it was strongly predicted to possess a signal peptide. Intriguingly, CAH1 could not be imported directly by isolated pea chloroplasts, but was instead taken up by pancreatic microsomes and concomitantly processed to its mature size. The protein was predicted to have several acceptor sites for *N*-linked glycosylation, and proteome analysis led to the identification of glycosylated CAH1, as well as several other glycoproteins, in the chloroplast stroma. Because the detected glycans are added only in the Golgi, these data implied the existence of a chloroplast protein transport pathway through the Golgi. Indeed, application of brefeldin A, a commonly used fungal agent that interferes with Golgi-mediated vesicle traffic, was found to obstruct the targeting of CAH1, causing it to arrest within the endomembrane system. The elucidated pathway was proposed to represent an ancestral targeting mechanism that predominated during early evolution (before the development of the now dominant TOC/TIC system), and which for some reason has been retained for a few proteins (Villarejo et al. 2005); the necessity for dual targeting of some proteins, to both plastids and compartments on the secretory pathway, may account for the retention of the ancestral mechanism in some cases (Chen et al. 2004; Asatsuma et al. 2005). It seems likely that many of the other chloroplast proteins with predicted signal peptides follow the same targeting route as CAH1 (Jarvis 2004; Kleffmann et al. 2004).

That some of the proteins identified within the whole chloroplast proteome are predicted to have mitochondrial presequences is much less surprising (Kleffmann et al. 2004). Presequences and chloroplast transit peptides share many similarities and are difficult to distinguish (Bhushan et al. 2006), and it is now becoming increasingly clear that many proteins are dual-targeted to both chloroplasts and mitochondria (Peeters and Small 2001; Duchêne et al. 2005). The identification of chloroplast proteins predicted to have no cleavable targeting signal can be explained in a number of different ways. Some of these may be outer envelope membrane proteins that insert directly into the membrane (Hofmann and Theg 2005), whereas others may have internal targeting signals for chloroplast localisation, like ceQORH or IEP32 (Miras et al. 2002; Nada and Soll 2004). The remainder are presumably contaminants, or proteins that have been classified incorrectly by TargetP (Emanuelsson et al. 2000; Richly and Leister 2004).

The emerging picture of protein targeting to chloroplasts is increasingly complex, demonstrating that transit peptide prediction in silico cannot provide a complete description of the chloroplast proteome. The existence of dual-targeted proteins adds an additional level of complexity (Levitan et al. 2005), further emphasising the need for experimental determination of protein localisation and the value of proteomic analysis.

## 14.4 Comparative Proteomics

As well as studies that simply catalogue the proteins present in a particular subcellular or suborganellar compartment, comparative proteomics has been employed with considerable success. For example, changes in the organellar proteome were studied during de-etiolation or greening (Lonosky et al. 2004), and the responses of the lumenal, stromal and plastoglobular proteomes to low temperature or light stress have been characterised (Giacomelli et al. 2006; Goulas et al. 2006; Ytterberg et al. 2006). In another example, chloroplasts isolated from Arabidopsis mutants lacking different TOC protein import receptor isoforms were compared with wild-type chloroplasts (Kubis et al. 2003, 2004). Different groups of chloroplast proteins were found to be selectively deficient in different receptor mutants, indicating that the different TOC receptor isoforms likely possess a degree of preprotein recognition specificity (Jarvis and Robinson 2004; Bédard and Jarvis 2005). The data suggested that at least two different import pathways operate in plastids (one for highly abundant components of the photosynthetic apparatus, and another for much less abundant housekeeping proteins). The existence of these separate import pathways may help to prevent deleterious competition effects between preproteins (the housekeeping proteins might otherwise be out-competed), or play a role in the differentiation of different plastid types.

Another interesting study compared the soluble stromal proteomes of mesophyll cell and bundle-sheath cell chloroplasts in maize, a plant that utilises the $C_4$ photosynthetic mechanism (Majeran et al. 2005). The data not only revealed differential accumulation of photosynthetic carbon metabolism enzymes consistent with our understanding of $C_4$ photosynthesis, but also shed light on how other plastidic functions are distributed between the two cell types. For example, enzymes involved in nitrogen assimilation and the biosynthesis of lipids and tetrapyrroles were found predominantly in mesophyll cell chloroplasts, whereas those for starch biosynthesis were more abundant in bundle sheath chloroplasts. Many of these differences can be explained by consideration of the fundamental differences that exist between the two types of chloroplast. For instance, nitrogen assimilation has a high requirement for energy (ATP) and reducing power (NADPH), both of which are in short supply in the bundle sheath cell chloroplasts due to the absence of linear photosynthetic electron flow and the high demands of the Calvin cycle; it therefore makes sense to concentrate this process in mesophyll chloroplasts. Similarly, it is quite logical that starch biosynthetic enzymes should be located in the vicinity of the Calvin cycle, which operates exclusively in the bundle sheath cells, since this is the source of new photosynthate for starch synthesis.

## 14.5 Multiprotein Complexes and Protein Modification

Many proteins do not function alone, but as a part of multiprotein complexes of varying complexity. In order to truly understand the functions of the plastidic proteome, these protein–protein interactions must be identified and characterised.

Furthermore, complexes usually represent functional units, and so hypothetic functions can be assigned to some proteins via their physical association with others of known function. Thus, experimental procedures that maintain the oligomeric status of protein complexes are desirable, since they preserve important interaction information. Several major protein complexes of chloroplasts have been purified to homogeneity and analysed individually; examples include the photosystems (Szabo et al. 2001; Zolla et al. 2002), the cytochrome $b_6f$ complex (Whitelegge et al. 2002), the 70S ribosomal subunits (see Sect. 14.2.3), and a caseinolytic protease (Clp) complex (Peltier et al. 2004b). One approach that has recently come to the fore is native electrophoresis. This provides an effective alternative to IEF as a first dimension in 2-DE analysis, and has the advantage that it preserves oligomeric status. This method has been used to study protein complexes in the thylakoids and the stroma (Granvogl et al. 2006; Peltier et al. 2006).

Protein modification is another important issue that must be taken into consideration. The molecular mass of an intact protein defines its native covalent state, and so its accurate measurement can reveal modifications mediated either post-transcriptionally, through processes such as RNA editing, or post-translationally (Whitelegge 2003; van Wijk 2004). Several studies have revealed covalent modifications of plastidic proteins, including acetylation, glycosylation, palmitoylation, phosphorylation, and N-terminal methionine excision (Yamaguchi et al. 2000; Gomez et al. 2002; Ferro et al. 2003; Giglione et al. 2003; Villarejo et al. 2005). Such modifications may influence the activity, interactions or stability of the protein, or anchor it to a membrane. So far, many proteomic studies have employed a "bottom-up" approach, in which the proteins are first broken up into manageable, characteristic fragments (e.g. using the protease trypsin) prior to MS analysis. The disadvantage of this "peptide mass fingerprinting" strategy is that much of the information inherent in the intact protein (including its overall size and covalent status) may be lost. An alternative, attractive approach is "top-down" proteomics, in which the intact mass of the protein is determined prior to fragmentation (Whitelegge 2003; Zabrouskov et al. 2003; van Wijk 2004). This approach can therefore reveal any modifications that the protein has undergone, as well as its identity.

## 14.6 Concluding Remarks

The various studies described above have demonstrated the power and utility of plastid proteome analysis, and it is anticipated that proteomics will continue to form an essential component of chloroplast research in the future. In light of the emerging complexity of plastid protein traffic, revealed in part through proteomics, it seems that it will be necessary to develop new prediction tools for the identification of chloroplast proteins, to revise current estimates of the size of the Arabidopsis plastid proteome based on in silico analysis, and to reassess the notion that many plant nuclear genes inherited from the cyanobacterial endosymbiont

encode proteins that are not actually targeted back to the chloroplast (Martin et al. 2002). The further and more extensive application of proteomics will play a significant role in achieving these objectives. In conjunction with complementary technologies such as transcriptomics and metabolomics, proteomics will enable the application of unified, systems-based approaches, leading ultimately to a complete and accurate description of the constitution and functioning of the organelles upon which we all depend.

**Acknowledgements** I thank Enrique López-Juez for helpful comments on the manuscript, and gratefully acknowledge financial support from the Royal Society Rosenheim Research Fellowship and the Biotechnology and Biological Sciences Research Council (BBSRC).

## References

Andon NL, Hollingworth S, Koller A, Greenland AJ, Yates JR 3rd, Haynes PA (2002) Proteomic characterization of wheat amyloplasts using identification of proteins by tandem mass spectrometry. Proteomics 2:1156–1168

Asatsuma S, Sawada C, Itoh K, Okito M, Kitajima A, Mitsui T (2005) Involvement of alpha-amylase I-1 in starch degradation in rice chloroplasts. Plant Cell Physiol 46:858–869

Austin JR 2nd, Frost E, Vidi PA, Kessler F, Staehelin LA (2006) Plastoglobules are lipoprotein subcompartments of the chloroplast that are permanently coupled to thylakoid membranes and contain biosynthetic enzymes. Plant Cell 18:1693–1703

Awai K, Xu C, Lu B, Benning C (2006) Lipid trafficking between the endoplasmic reticulum and the chloroplast. Biochem Soc Trans 34:395–398

Baginsky S, Gruissem W (2004) Chloroplast proteomics: potentials and challenges. J Exp Bot 55:1213–1220

Baginsky S, Siddique A, Gruissem W (2004) Proteome analysis of tobacco bright yellow-2 (BY-2) cell culture plastids as a model for undifferentiated heterotrophic plastids. J Proteome Res 3:1128–1137

Balmer Y, Koller A, del Val G, Manieri W, Schurmann P, Buchanan BB (2003) Proteomics gives insight into the regulatory function of chloroplast thioredoxins. Proc Natl Acad Sci USA 100:370–375

Balmer Y, Vensel WH, DuPont FM, Buchanan BB, Hurkman WJ (2006a) Proteome of amyloplasts isolated from developing wheat endosperm presents evidence of broad metabolic capability. J Exp Bot 57:1591–1602

Balmer Y, Vensel WH, Cai N, Manieri W, Schurmann P, Hurkman WJ, Buchanan BB (2006b) A complete ferredoxin/thioredoxin system regulates fundamental processes in amyloplasts. Proc Natl Acad Sci USA 103:2988–2993

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

Bhushan S, Kuhn C, Berglund AK, Roth C, Glaser E (2006) The role of the N-terminal domain of chloroplast targeting peptides in organellar protein import and mis-sorting. FEBS Lett 580:3966–3972

Chen MH, Huang LF, Li HM, Chen YR, Yu SM (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

Crotty WJ, Ledbetter MC (1973) Membrane continuities involving chloroplasts and other organelles in plant cells. Science 182:839–841

Duchêne AM, Giritch A, Hoffmann B, Cognat V, Lancelin D, Peeters NM, Zaepfel M, Marechal-Drouard L, Small ID (2005) Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in *Arabidopsis thaliana*. Proc Natl Acad Sci USA 102:16484–16489

Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol 300:1005–1016

Ferro M, Salvi D, Brugiere S, Miras S, Kowalski S, Louwagie M, Garin J, Joyard J, Rolland N (2003) Proteomics of the chloroplast envelope membranes from *Arabidopsis thaliana*. Mol Cell Proteomics 2:325–345

Friso G, Giacomelli L, Ytterberg AJ, Peltier JB, Rudella A, Sun Q, Wijk KJ (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 JE, Wilkerson CG, Ray WK, McAndrew RS, Osteryoung KW, Gage DA, Phinney BS (2003) Proteomic study of the *Arabidopsis thaliana* chloroplastic envelope membrane utilizing alternatives to traditional two-dimensional electrophoresis. J Proteome Res 2:413–425

Gaikwad A, Tewari KK, Kumar D, Chen W, Mukherjee SK (1999) Isolation and characterisation of the cDNA encoding a glycosylated accessory protein of pea chloroplast DNA polymerase. Nucleic Acids Res 27:3120–3129

Giacomelli L, Rudella A, van Wijk KJ (2006) High light response of the thylakoid proteome in arabidopsis wild type and the ascorbate-deficient mutant *vtc2-2*. A comparative proteomics study. Plant Physiol 141:685–701

Giglione C, Vallon O, Meinnel T (2003) Control of protein life-span by N-terminal methionine excision. EMBO J 22:13–23

Gomez SM, Nishio JN, Faull KF, Whitelegge JP (2002) The chloroplast grana proteome defined by intact mass measurements from liquid chromatography mass spectrometry. Mol Cell Proteomics 1:46–59

Goulas E, Schubert M, Kieselbach T, Kleczkowski LA, Gardestrom P, Schröder W, Hurry V (2006) The chloroplast lumen and stromal proteomes of *Arabidopsis thaliana* show differential sensitivity to short- and long-term exposure to low temperature. Plant J 47:720–734

Granvogl B, Reisinger V, Eichacker LA (2006) Mapping the proteome of thylakoid membranes by de novo sequencing of intermembrane peptide domains. Proteomics 6:3681–3695

Haas BJ, Wortman JR, Ronning CM, Hannick LI, Smith RK Jr, Maiti R, Chan AP, Yu C, Farzad M, Wu D, White O, Town CD (2005) Complete reannotation of the *Arabidopsis* genome: methods, tools, protocols and the final release. BMC Biol 3:7

Heazlewood JL, Tonti-Filippini JS, Gout AM, Day DA, Whelan J, Millar AH (2004) Experimental analysis of the Arabidopsis mitochondrial proteome highlights signaling and regulatory components, provides assessment of targeting prediction programs, and indicates plant-specific mitochondrial proteins. Plant Cell 16:241–256

Hisabori T, Hara S, Fujii T, Yamazaki D, Hosoya-Matsuda N, Motohashi K (2005) Thioredoxin affinity chromatography: a useful method for further understanding the thioredoxin network. J Exp Bot 56:1463–1468

Hofmann NR, Theg SM (2005) Chloroplast outer membrane protein targeting and insertion. Trends Plant Sci 10:450–457

Jarvis P (2003) Intracellular signalling: the language of the chloroplast. Curr Biol 13: R314–R316

Jarvis P (2004) Organellar proteomics: chloroplasts in the spotlight. Curr Biol 14:R317–R319

Jarvis P, Robinson C (2004) Mechanisms of protein import and routing in chloroplasts. Curr Biol 14:R1064–R1077

Kessler F, Schnell D, Blobel G (1999) Identification of proteins associated with plastoglobules isolated from pea (*Pisum sativum* L.) chloroplasts. Planta 208:107–113

Kieselbach T, Schroder WP (2003) The proteome of the chloroplast lumen of higher plants. Photosynth Res 78:249–264

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

Koo AJ, Ohlrogge JB (2002) The predicted candidates of Arabidopsis plastid inner envelope membrane proteins and their expression profiles. Plant Physiol 130:823–836

Kouranov A, Wang H, Schnell DJ (1999) Tic22 is targeted to the intermembrane space of chloroplasts by a novel pathway. J Biol Chem 274:25181–25186

Kubis S, Baldwin A, Patel R, Razzaq A, Dupree P, Lilley K, Kurth J, Leister D, Jarvis P (2003) The Arabidopsis *ppi1* mutant is specifically defective in the expression, chloroplast import, and accumulation of photosynthetic proteins. Plant Cell 15:1859–1871

Kubis S, Patel R, Combe J, Bédard J, Kovacheva S, Lilley K, Biehl A, Leister D, Ríos G, Koncz C, Jarvis P (2004) Functional specialization amongst the Arabidopsis Toc159 family of chloroplast protein import receptors. Plant Cell 16:2059–2077

Kwok EY, Hanson MR (2004) Stromules and the dynamic nature of plastid morphology. J Microsc 214:124–137

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

Levitan A, Trebitsh T, Kiss V, Pereg Y, Dangoor I, 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

Li M, Schnell DJ (2006) Reconstitution of protein targeting to the inner envelope membrane of chloroplasts. J Cell Biol 175:249–259

Lonosky PM, Zhang X, Honavar VG, Dobbs DL, Fu A, Rodermel SR (2004) A proteomic analysis of maize chloroplast biogenesis. Plant Physiol 134:560–574

López-Juez E, Pyke KA (2005) Plastids unleashed: their development and their integration in plant development. Int J Dev Biol 49:557–577

Majeran W, Cai Y, Sun Q, van Wijk KJ (2005) Functional differentiation of bundle sheath and mesophyll maize chloroplasts determined by comparative proteomics. Plant Cell 17:3111–3140

Margulis L (1970) The origin of eukaryotic cells. Yale University Press, New Haven

Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc Natl Acad Sci USA 99:12246–12251

Miras S, Salvi D, Ferro M, Grunwald D, Garin J, Joyard J, Rolland N (2002) Non-canonical transit peptide for import into the chloroplast. J Biol Chem 277:47770–47778

Motohashi K, Hisabori T (2006) HCF164 receives reducing equivalents from stromal thioredoxin across the thylakoid membrane and mediates reduction of target proteins in the thylakoid lumen. J Biol Chem 281:35039–35047

Motohashi K, Kondoh A, Stumpp MT, Hisabori T (2001) Comprehensive survey of proteins targeted by chloroplast thioredoxin. Proc Natl Acad Sci USA 98:11224–11229

Nada A, Soll J (2004) Inner envelope protein 32 is imported into chloroplasts by a novel pathway. J Cell Sci 117:3975–3982

Nelson N, Ben-Shem A (2004) The complex architecture of oxygenic photosynthesis. Nat Rev Mol Cell Biol 5:971–982

Pan S, Carter CJ, Raikhel NV (2005) Understanding protein trafficking in plant cells through proteomics. Expert Rev Proteomics 2:781–792

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

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

Peltier JB, Ytterberg AJ, Sun Q, van Wijk KJ (2004a) New functions of the thylakoid membrane proteome of *Arabidopsis thaliana* revealed by a simple, fast, and versatile fractionation strategy. J Biol Chem 279:49367–49383

Peltier JB, Ripoll DR, Friso G, Rudella A, Cai Y, Ytterberg J, Giacomelli L, Pillardy J, van Wijk KJ (2004b) Clp protease complexes from photosynthetic and non-photosynthetic plastids and

mitochondria of plants, their predicted three-dimensional structures, and functional implications. J Biol Chem 279:4768–4781

Peltier JB, Cai Y, Sun Q, Zabrouskov V, Giacomelli L, Rudella A, Ytterberg AJ, Rutschow H, van Wijk KJ (2006) The oligomeric stromal proteome of *Arabidopsis thaliana* chloroplasts. Mol Cell Proteomics 5:114–133

Richly E, Leister D (2004) An improved prediction of chloroplast proteins reveals diversities and commonalities in the chloroplast proteomes of Arabidopsis and rice. Gene 329:11–16

Rolland N, Ferro M, Seigneurin-Berny D, Garin J, Douce R, Joyard J (2003) Proteomics of chloroplast envelope membranes. Photosynth Res 78:205–230

Schleiff E, Eichacker LA, Eckart K, Becker T, Mirus O, Stahl T, Soll J (2003) Prediction of the plant beta-barrel proteome: a case study of the chloroplast outer envelope. Protein Sci 12:748–759

Schröder WP, Kieselbach T (2003) Update on chloroplast proteomics. Photosynth Res 78:181–193

Schubert M, Petersson UA, Haas BJ, Funk C, Schröder WP, Kieselbach T (2002) Proteome map of the chloroplast lumen of *Arabidopsis thaliana*. J Biol Chem 277:8354–8365

Sun Q, Emanuelsson O, van Wijk KJ (2004) Analysis of curated and predicted plastid subproteomes of Arabidopsis. Subcellular compartmentalization leads to distinctive proteome properties. Plant Physiol 135:723–734

Szabo I, Seraglia R, Rigoni F, Traldi P, Giacometti GM (2001) Determination of photosystem II subunits by matrix-assisted laser desorption/ionization mass spectrometry. J Biol Chem 276:13784–13790

Van der Laan M, Rissler M, Rehling P (2006) Mitochondrial preprotein translocases as dynamic molecular machines. FEMS Yeast Res 6:849–861

Van Dooren GG, Schwartzbach SD, Osafune T, McFadden GI (2001) Translocation of proteins across the multiple membranes of complex plastids. Biochim Biophys Acta 1541:34–53

van Wijk KJ (2004) Plastid proteomics. Plant Physiol Biochem 42:963–977

Vidi PA, Kanwischer M, Baginsky S, Austin JR, Csucs G, Dormann P, Kessler F, Brehelin C (2006) Tocopherol cyclase (VTE1) localization and vitamin E accumulation in chloroplast plastoglobule lipoprotein particles. J Biol Chem 281:11225–11234

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, 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

Von Zychlinski A, Kleffmann T, Krishnamurthy N, Sjolander K, Baginsky S, Gruissem W (2005) Proteome analysis of the rice etioplast: metabolic and regulatory networks and novel protein functions. Mol Cell Proteomics 4:1072–1084

Weber AP, Schwacke R, Flügge UI (2005) Solute transporters of the plastid envelope membrane. Annu Rev Plant Biol 56:133–164

Westerlund I, Von Heijne G, Emanuelsson O (2003) LumenP – a neural network predictor for protein localization in the thylakoid lumen. Protein Sci 12:2360–2366

Whatley JM (1978) A suggested cycle of plastid developmental interrelationships. New Phytol 80:489–502

Whatley JM, McLean B, Juniper BE (1991) Continuity of chloroplast and endoplasmic reticulum membranes in *Phaseolus vulgaris*. New Phytol 117:209–217

Whitelegge JP (2003) Thylakoid membrane proteomics. Photosynth Res 78:265–277

Whitelegge JP, Zhang H, Aguilera R, Taylor RM, Cramer WA (2002) Full subunit coverage liquid chromatography electrospray ionization mass spectrometry (LCMS+) of an oligomeric membrane protein: cytochrome b(6)f complex from spinach and the cyanobacterium *Mastigocladus laminosus*. Mol Cell Proteomics 1:816–827

Yamaguchi K, Subramanian AR (2000) The plastid ribosomal proteins. Identification of all the proteins in the 50 S subunit of an organelle ribosome (chloroplast). J Biol Chem 275:28466–28482

Yamaguchi K, von Knoblauch K, Subramanian AR (2000) The plastid ribosomal proteins. Identification of all the proteins in the 30 S subunit of an organelle ribosome (chloroplast). J Biol Chem 275:28455–28465

Ytterberg AJ, Peltier JB, van Wijk KJ (2006) Protein profiling of plastoglobules in chloroplasts and chromoplasts. A surprising site for differential accumulation of metabolic enzymes. Plant Physiol 140:984–997

Zabrouskov V, Giacomelli L, van Wijk KJ, McLafferty FW (2003) A new approach for plant proteomics: characterization of chloroplast proteins of *Arabidopsis thaliana* by top-down mass spectrometry. Mol Cell Proteomics 2:1253–1260

Zolla L, Rinalducci S, Timperio AM, Huber CG (2002) Proteomics of light-harvesting proteins in different plant species. Analysis and comparison by liquid chromatography-electrospray ionization mass spectrometry. Photosystem I. Plant Physiol 130:1938–1950

Zolla L, Timperio AM, Walcher W, Huber CG (2003) Proteomics of light-harvesting proteins in different plant species. Analysis and comparison by liquid chromatography-electrospray ionization mass spectrometry. Photosystem II. Plant Physiol 131:198–214