Insights into chloroplast proteomics: from basic principles to new horizons

Bianca Naumann and Michael Hippler

Abstract

Many proteomic approaches have been employed to investigate the complex and dynamic proteome of the chloroplast. These range from classical methods like one and two dimensional gel electrophoresis to advanced comparative proteomics strategies such as ICAT or SILAC. Mass spectrometry for protein identification or quantitation plays an important role in most of the methods used and is a fast emerging technology in protein biochemistry. Most proteomic studies of the chloroplast focus on the single compartments of this plant organelle, which greatly reduces the complexity of the sample and thus allows for a more complete and detailed analysis of the complex protein composition. The rapidly developing field of comparative proteomics makes it possible to analyze dynamic protein changes caused, for example, by different developmental stages of a plant, by various stress conditions and distinct genetic backgrounds.

1 The art of proteomics

A great challenge of the post genomic era is to understand how genetic information results in the concerted and dynamic action of gene products to generate function. In contrast to a cell's static genome, the proteome is both complex and dynamic. The proteome is defined as the set of all expressed proteins in a cell, tissue or organism (Wilkins et al. 1999). Proteomics can be defined as the systematic analysis of proteins for their identity, abundance, expression pattern, and function. Proteomics permits a global view on dynamics of biological processes by the systematic analysis of expressed proteins and, in particular, of functional protein complexes. The analysis of a proteome is complicated by the fact that the expressed product of a single gene often represents a protein population that may contain a large amount of micro-heterogeneity. Post-translational modifications (PTM), like phosphorylation, acetylation, glycosylation, protease cleavage, lipidation, or ubiquitination may contribute to the expression profile of a protein. The analysis of such complex protein profiles requires methods that allow high resolution protein separation combined with very sensitive methods for protein identification. Mass spectrometry (MS) has become a powerful tool for peptide and protein identification since it allows sensitive, fast and specific measurement, and

Topics in Current Genetics, Vol. 19
R. Bock (Ed.): Cell and Molecular Biology of Plastids
DOI 10.1007/4735_2007_0224 / Published online: 15 May 2007
© Springer-Verlag Berlin Heidelberg 2007

thus allows for recognition of peptides and proteins from complex mixtures (Aebersold and Mann 2003; Domon and Aebersold 2006). A typical workflow for proteomic experiments is depicted in Figure 1. Besides protein identification, recognition of post-translational modifications and protein quantitation are important tasks that can be investigated by mass spectrometric experiments. Today whole suites of potent mass spectrometer (MS) are available to fulfill these tasks (Aebersold and Mann 2003; Domon and Aebersold 2006). It is not our aim to discuss the distinct mass spectrometer options available. In the beginning of this review, we would rather like to address issues that are essential for successful mass spectrometric experiments: (i) peptide ionization allowing the entry of peptides into the mass spectrometer, (ii) peptide mass finger printing (PMF) and tandem mass spectrometry (MS/MS), and (iii) algorithms that permit identification of peptides and in turn proteins from mass spectrometric data.

1.1 Prerequisite for biomolecular mass spectrometry: MALDI and ESI Ionization

To enable mass spectrometric analysis of peptide-molecules, they have to be ionized before they can enter the mass spectrometer. The most common ionization methods for biomolecular mass spectrometry are matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI). MALDI is based on the bombardment of sample molecules with a laser light to induce ionization. The sample is pre-mixed with a highly light-absorbing matrix compound. The matrix absorbs the laser energy and transforms it into excitation energy. This leads to the sputtering of matrix molecules, which drag along the analyte ions from the surface of the mixture, and enables the entry of ionized molecules from an intermediate vacuum region into the analyzer of the mass spectrometer, which is under permanent high vacuum. During electrospray ionization, the sample is dissolved in a polar solvent and pumped through a narrow capillary. A high voltage of 3 to 4 kV is usually applied in between the tip of the capillary, which is positioned within the ionization source of the mass spectrometer and the aperture, which represents the entry point to the high vacuum system. In response to this strong electric field, the sample emerging from the tip is dispersed into an aerosol of highly charged droplets. Ultimately, charged sample ions, free from solvent, are released from the droplets, which pass through a sampling cone into an intermediate vacuum region, and from there through a small aperture into the analyzer of the mass spectrometer.

1.2 Peptide mass finger printing and tandem mass spectrometry

MALDI coupled to time-of-flight (TOF) mass spectrometer instruments are commonly used for large-scale protein identification by the peptide mass mapping technique. Peptide masses are determined for specific spots on the analyzer plate by MALDI-MS and these mass maps are then compared to predicted mass maps

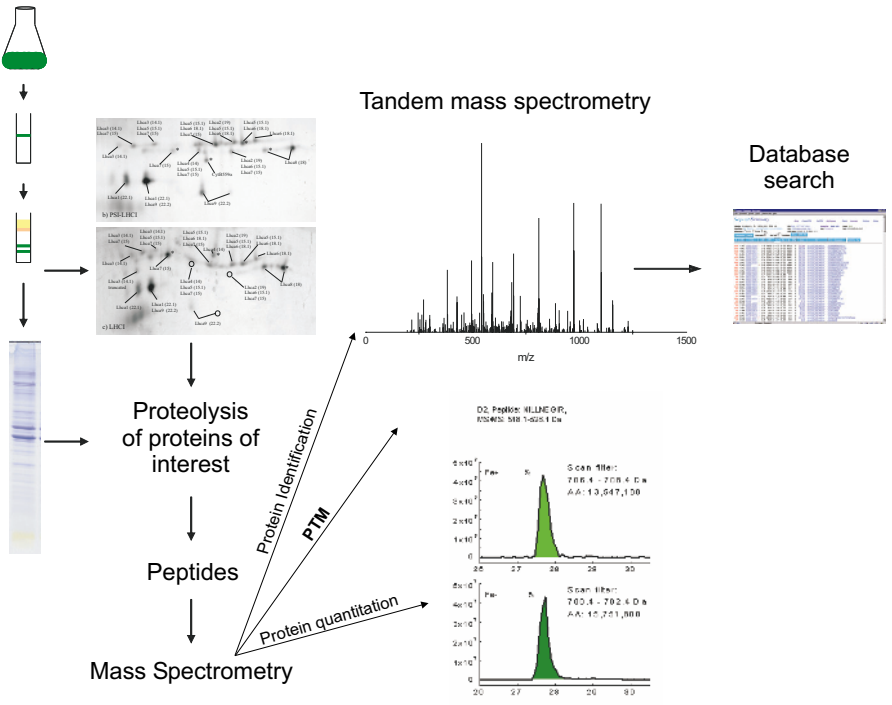

Fig. 1. Typical workflow of a proteomics experiment. Chloroplasts can for example be extracted by centrifugation on a Percoll gradient and than fractionated into their compartments by sucrose gradient centrifugation. Protein samples can than be separated by 1- or 2-dimensional gel electrophoresis and further analyzed with mass spectrometry or tandem mass spectrometry in respect to identification, post-translational modification and quantitation.

in a database to identify the respective protein. MALDI can also be employed to ionize peptides for entry into more complex mass spectrometer, enabling tandem mass spectrometric analyses, such as ion trap, TOF-TOF or quadrupole-TOF mass spectrometer. In contrast to MALDI, ESI allows direct coupling of liquid chromatography (LC) systems to the mass spectrometer. This permits the combination of chromatographic separation of peptides (using on average nano flow rates) and direct elution into the mass spectrometer coupled to mass spectrometric and tandem mass spectrometric analysis of peptides. Besides mass information of the peptide-ion, tandem mass spectrometry produces structural information about specifically selected peptide ions inside the mass spectrometer. MS/MS experiments are performed by colliding a selected ion with inert gas molecules such as argon or helium and subsequent mass measurement of the fragment ions yielding fragmentation mass spectra. Importantly, fragmentation occurs preferably at the peptide bonds resulting in y- and b-type ions that represent fragment ions harboring either C- or N-terminus, respectively. This information can then be assembled to generate structural information regarding the intact molecule and enable direct amino

acid sequencing of peptides. To identify peptides from mass spectrometric and tandem mass spectrometric data, algorithms are available that take advantage of protein and DNA database information to correlate peptide sequences with mass spectrometric data information. In addition *de novo* amino acid sequencing from MS/MS data is feasible. Since today's mass spectrometer become more and more sensitive and faster in data recording, the bottleneck of the mass spectrometric experiment seems to be the evaluation of these data.

1.3 Database searching

A set of distinct peptide masses obtained from proteolytic cleavage (mostly tryptic cleavage) of a protein and subsequent mass spectrometric mass measurement can be used to identify proteins. In this approach such a set of peptide masses is mapped against an *in silico* digest of a protein sequence database. The approach, called peptide mass fingerprinting (Mann et al. 1993; Giddings et al. 2003), is however vulnerable to the complexity of the mixture. With the increase of peptide species in the mixture the possible combinations increase exponentially, thus, making a correct protein assignment difficult. In addition peptide masses may not be unique in sequence databases. Therefore, identification of a protein via PMF depends critically on the mass accuracy and the mass resolution of the mass spectrometer.

The introduction of MS/MS spectra for peptide and protein identification included the fragmentation pattern of a peptide as a supplementary constraint in addition to the peptide mass, thus, rendering the peptide identification and in turn the protein identification more reliable. Sequence tags (Mann and Wilm 1994; Shevchenko et al. 1996) amend the mass of the peptide, as in PMF, with a short partial amino acid sequence, which is determined from the spectrum and its position within the peptide. Thus, four parameters define a sequence tag: a) its mass, b) its partial amino acid sequence, c) the mass before the start of the partial amino acid sequence in the peptide, and d) the remaining mass after the end of the partial amino acid sequence within the peptide. Partial sequences can be searched for in sequence databases, usually presented as plain text-files in fasta-format (Pearson and Lipman 1988). Furthermore, the masses of the resulting fragments of the *in silico* digest of these files are used to filter the results. Another filter is presented through the positioning of the partial amino acid sequence. These three filters are very restrictive and more discriminating than searching with mere masses alone. Therefore, it is in widespread use today, with new developments reported regularly (Bafna and Edwards 2001; Sunyaev et al. 2003; Tabb et al. 2003; Savitski et al. 2005).

Another approach developed around the same time as sequence tagging makes use of the complete MS/MS spectrum (Eng et al. 1994). It uses cross correlation to compare the acquired mass spectra to theoretically derived spectra from sequences in a database. This algorithm, named Sequest, along with Mascot (Perkins et al. 1999), which employs sequence search, ion search, PMF, and introduces a probabilistic based scoring scheme for the first time, are the so called industry stan-

dards for software in this area today. Besides Sequest and Mascot, numerous other tools that match mass spectrometric data to sequence databases are available (for reviews see Kapp et al. 2005; Shadforth et al. 2005). Although database search is able to identify peptides from complex mixtures, it obviously fails if there is no database available, or if other obstacles hinder the identification. In these cases, *de novo* amino acid sequencing may be of use.

1.4 *De Novo* sequencing

De novo sequencing algorithms seek to determine the underlying peptide sequence from the mass spectrometric information alone. The rational behind this approach is that peptides dissociate into predictable fragments. Looking at y-ions alone clearly shows that the difference in-between two consecutive y-ions in a spectrum represents the mass of one or multiple amino acid. Other ion-types may provide additional and supporting information in this scenario. The best case occurs when a complete fragment ion ladder of at least one ion-type is present. The inherent problem in *de novo* sequencing is, however, that it is not known which peak represents which ion-type in a given MS/MS spectrum *a priori*.

A number of *de novo* amino acid sequencing programs have been described and are in use today (Dancik et al. 1999; Fernandez-de-Cossio et al. 2000; Chen et al. 2001; Taylor and Johnson 2001; Bafna and Edwards 2003; Ma et al. 2003). These programs face other limitations. They are usually computational intensive and dependent on high quality spectra (Spengler 2004; Yan et al. 2005). For these reasons, they are quite limited in practice. *De novo* amino acid sequencing information together with mass information could be used for error-tolerant searching of DNA and in particular genomic DNA databases. Therefore, there is a need to connect *de novo* sequencing approaches with database searching algorithms.

1.5 Linking database searching and *de novo* sequencing

The GenomicPeptideFinder (GPF) connects *de novo* sequencing with database search (Fig. 2) (Allmer et al. 2006). The aim of GPF is to employ mass spectrometric data for genomic data mining. It enables detection of intron-split and/or alternatively spliced peptides from MS/MS data when deduced from genomic DNA (Allmer et al. 2004). As depicted in Figure 2, prior to GPF search, mass spectra are submitted to *de novo* amino acid sequencing by PEAKS (Ma et al. 2003). The predictions are converted to queries for GPF and searched against the six-frame translation of a genomic DNA database. For this error tolerant search small sub sequences of the *de novo* prediction are mapped to the six-frame translation of a genomic database. The proximity of a match, usually 2100 base pairs upstream and downstream, is investigated in more detail. This time shorter sequence fragments are searched in the extended region. All matches are tried out whether they, when they are joined, define a tryptic peptide that would explain the precursor

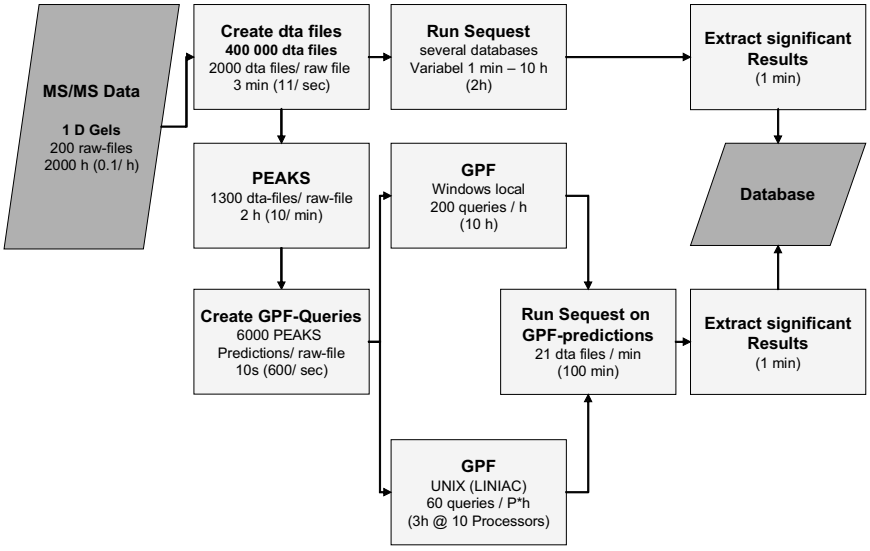

Fig. 2. Computational processing of the MS/MS spectra acquired by mass spectrometry. Certain data associated with each process such as processing time is presented in the boxes above, each box representing a distinct process. Processing times were calculated for one PC if not indicated otherwise. Most triply charged dta-files were not submitted to *de novo* prediction analysis. The GPF core is the same in both PC and UNIX distribution. Most GPF processing was done on the LINIAC Cluster (University of Pennsylvania, Philadelphia, USA). Figure taken from Allmer et al. (2006).

mass within the error of the mass spectrometer used. Joining the matched sequences allows splicing out of intervening sequences. In order to allow for sequencing errors of the *de novo* algorithm, the intervening sequence is checked along the reading frames of the bordering matches whether they can explain the precursor mass, if completely removing it, renders the mass too low. All *de novo* sequence predictions are used in this fashion. All resulting peptides are stored in a fasta-file and are submitted to database search. The new sequences function as the database and they are correlated using Sequest against the original mass spectra which gave rise to the *de novo* predictions.

This approach was used to study the thylakoid proteome of *Chlamydomonas reinhardtii* (Allmer et al. 2006). The concerted action of Sequest and GPF allowed identification of 2622 distinct peptides. In total 448 peptides were identified by GPF analysis alone including 98 intron-split peptides, resulting in the identification of novel proteins, improved annotation of gene models, and evidence of alternative splicing. It is predictable that the combination of *de novo* sequencing from MS/MS spectra in conjunction with error-tolerant GPF performance will be of

help to explore nuclear gene structures and identify alternative splicing in eukaryotic organism with complex genomes.

1.6 Strategies for the analysis of proteome dynamics

Traditionally, a standard technique for proteome analysis combines protein separation by high-resolution (isoelectric focusing (IEF)/SDS-PAGE) two-dimensional gel electrophoresis (2-DE) with mass spectrometry or tandem MS identification of selected protein spots. The 2-DE technique has been used for the separation, detection and quantification of individual proteins present in a complex sample in combination with mass spectrometry and database searching for the identification of the separated proteins (as reviewed in Aebersold and Mann 2003; Gorg et al. 2004; Wittmann-Liebold et al. 2006).
In 2-DE, proteins are separated in first and second dimension according to their isoelectric point and molecular mass, respectively. In the first dimension proteins are fractionated by isoelectric focusing. Hereby separation of proteins is achieved through electrophoresis in a pH gradient gel system (using a gel strip with embedded pH gradient). Proteins migrate in the gel according to their charge at the respective pH and will accumulate at their isolelectric point (IP) where the positive and negative charges of the peptide are balanced so that they do not display a charge to the outside and do not migrate in an electric field anymore. In the second dimension proteins from the gel strip are run into a SDS-PAGE and separated according to their molecular mass. Before separation, proteins in the gel strip are treated with sodium dodecyl sulfate (SDS) along with other reagents to ensure that they are denatured and carry an appropriate negative charge. After separation, the gel is stained (i.e. Coomassie brilliant blue, silver, fluorescence dyes) and further analyzed. Protein spots of interest can be excised and digested with a site-specific protease (often trypsin). The resulting peptides are further investigated by mass spectrometry. This combination of methods is employed as a tool to detect and dissect dynamic changes in the proteome of a cell or tissue in response to changes in the physiological environment, the developmental state or internal perturbations, such as mutations. Fluorescence 2-DE Difference Gel Electrophoresis (DIGE) (Unlu et al. 1997) represents a new development in 2-DE. The use of multiple distinct fluorescent dyes to label protein samples prior to 2-DE PAGE allows multiple samples to be co-separated and visualized on one 2-DE gel.

Although, classical two-dimensional gel electrophoresis is a powerful tool, it faces a number of limitations especially when it comes to the separation of highly hydrophobic membrane proteins and proteins that possess basic isoelectric points. Hydrophobic proteins tend to precipitate at their isoelectric point in the non-detergent isoelectric focusing. In addition, they are often very heterogeneous in their physico-chemical properties what makes it difficult to achieve comprehensive, reproducible and comparable protein maps of membrane fractions (Ephritikhine et al. 2004).

New experimental approaches that are independent of 2-DE have been developed recently to overcome these limitations and allow comparative analysis of a

protein between experimental and control samples in "solution", enabling a quantitative overview of the dynamically altered proteome.

Differential isotopic labeling strategies can also be employed to distinguish proteins from control and experimental conditions. Besides crosslinking of proteins isolated from cells grown under different conditions with isotopically labeled and unlabeled chemical probes (ICAT, isotope-coded affinity tag) (Gygi et al. 1999), proteins can be metabolically labeled with stable isotopes by growing cells in isotopically enriched media (SILAC, stable isotope-labeling of amino acids in cell culture). Experimental and control cell pools are then mixed, digested with enzymes and analyzed by LC-MS/MS for protein quantification (Oda et al. 1999; Ong et al. 2002). Moreover, mass spectrometry can be used to achieve absolute quantitation of proteins. For this purpose, proteotypic peptides that distinctively recognize a protein can be chemically synthesized holding stable isotopes (e.g. ^{13}C, ^{15}N etc.) at a single amino acid so that their masses will differ from the mass of the analyte, thereby permitting differentiation by MS and MS/MS methods. The absolute concentration of a protein can be calculated from the signal intensities derived from the analyte and from those of the internal standard (Zhu and Desiderio 1996). Tryptic peptides derived from the analyzed proteins and synthetic isotopically labeled internal standards were employed in absolute quantification of proteins in solution (Barr et al. 1996; Barnidge et al. 2003) and, recently, in-gel (Gerber et al. 2003).

In recent years alternative approaches have been developed that make use of the coupling between liquid chromatography and tandem MS (LC-MS/MS) (as reviewed in (Peng and Gygi 2001). The power of such a strategy can for example be illustrated by the identification of more than 70 proteins from the yeast ribosome in a single analysis. This approach was performed by analyzing tryptic peptides, which derived from digestion of the whole complex, by multi-dimensional liquid chromatography (mudPIT) coupled to MS/MS (Link et al. 1999).

Our current understanding of the organization of a proteome-wide interaction network points to its enormous complexity. It has become apparent that on average, every fourth protein in a proteome might be shared between protein complexes of different function. Two rather impressive proteome approaches were described for systematic analyses of components of multi-protein complexes from baker yeast *Saccharomyces cerevisiae* (Gavin et al. 2002; Ho et al. 2002). In one approach about 10% of predicted yeast proteins were used as baits to discover protein-protein interactions. Fascinatingly, 3,617 associated proteins were identified by mass spectrometry, covering about 25% of the yeast proteome (Ho et al. 2002). In the other approach distinct genes were tagged with an expression cassette encoding for protein A and the calmodulin binding protein. Protein complexes that contain a tagged protein could be isolated by tandem-affinity purification and the individual components be analyzed by MS/MS (Gavin et al. 2002).

The importance and power of proteomics for the exploration of plant proteomes and in particular plastid proteomics will be discussed in depth in the following section.

2 Proteomics of the chloroplast and its compartments

The chloroplast is a highly dynamic and complex cell organelle and a major part of its metabolism is involved in photosynthesis and related energy producing processes. However, it also produces amino acids and lipids as well as secondary metabolites like isoprenoids. A chloroplast can be divided into several compartments: the double layered envelope membrane, the soluble stroma and the thylakoid membrane enclosing the lumen. Despite of its endosymbiotic origin, most of the chloroplast proteins are nuclear encoded (about 90%). Plastidic genetic material is organized in so-called nucleoids and expressed via the chloroplast's transcription and translation machinery. Most of the chloroplast proteins synthesized in the cytosol are imported via the Toc and Tic (translocon at the outer/inner envelope membrane) translocation machinery. They contain an N-terminal chloroplast transit peptide (cTP) that is necessary for recognition at the outer membrane and is cleaved off after the passage into the stroma. Proteins targeted to the lumen of the thylakoids have to overcome a second barrier: the thylakoid membrane. Therefore, they carry a bipartite transit peptide, the more N-terminal region determining the chloroplast targeting. Once the initial transit peptide is cleaved, the formerly masked second transit peptide can be assessed and determines the next target, for example the thylakoid lumen (lTP).

To cope with the large amount of functions in cell metabolism, chloroplasts require a large amount of enzymes and other multi protein complexes. Therefore, a number of about 3000-4000 predicted chloroplast proteins in *Arabidopsis* is easily imaginable (Peltier et al. 2002; Schubert et al. 2002; Kieselbach and Schroder 2003; Baginsky and Gruissem 2004; Kleffmann et al. 2004; Richly and Leister 2004; Sun et al. 2004; van Wijk 2004). Only one proteomic study so far took on the difficult task to analyze the complete chloroplast proteome. In a study on *Arabidopsis thaliana* Kleffmann et al. (2004) were able to improve the dynamic range by using multi dimensional chromatography and additional enrichment of envelope membrane proteins to identify a set of 636 proteins, mostly associated with energy production or metabolic processes. But even with advanced proteomic strategies, the analysis of the whole chloroplast proteome is a difficult operation and therefore the chloroplast proteome is often divided in different subproteomes composed of the proteins from its subcompartments.

2.1 Envelope membranes

The chloroplast is a closed compartment located in the cytosol of the cell. It is completely surrounded by a double membrane, the so-called envelope. Since a wide range of different metabolic processes take place in the chloroplast, it is necessary for it to be functionally fully integrated into the plant cell, requiring an unproblematic exchange of metabolites and signals with the other compartments. Therefore, the envelope membrane is the site of many transport systems that facilitate the transport of not only carbohydrates but also phosphates, amino acids, protons and different metal-ions (Joyard et al. 1998; Rolland et al. 2003).

About 90% of the chloroplast proteins are synthesized in the cytosol and must be transported through the envelope membranes. This is achieved with the Toc/Tic protein import machinery. The Toc proteins, located in the outer membrane, are able to recognize chloroplast transit peptides and guide the immature proteins through the outer envelope. Here the Tic complex takes over and translocates the preprotein into the stroma where the TP is cleaved and chaperones fold the proteins into their functional conformation (Bedard and Jarvis 2005). The envelope membranes are also involved in the production and metabolism of different lipids such as structural membrane constituents, carotenoids and prenylquinones. Lipid compounds can also be further metabolized to signal molecules that are active in, for example, growth regulation or plant defense (Joyard et al. 1998; Rolland et al. 2003).

Predictions of envelope membrane components are difficult to make since a lot of these proteins do not contain cTP. Ferro et al. (2002) were able to determine typical properties for internal envelope membrane proteins, like a strong hydrophobicity based on several transmembrane domains (TMD), a pI larger than 8.8 and a Res/TM value (amino acid residues/TMD) of less than 100. Their prediction with ChloroP (Emanuelsson et al. 1999) and a manual check on these criteria resulted in 136 potential envelope proteins in *Arabidopsis*. Koo and Ohlrogge (2002) used a combination of TargetP (Emanuelsson et al. 2000), the TMD predictor TMHMM and a manual rejection of known thylakoid proteins to predict a number of 541 inner envelope candidate proteins. A new prediction approach was used by Schleiff et al. (2003) based on the idea that most of the outer envelope proteins are embedded in the membrane by a β-barrel structure. Their prediction was based on the combination of a computational β-barrel analysis, the determination of the isoelectric point, a TargetP analysis and a manual selection. This resulted in a pool of 891 putative outer envelope proteins. The candidate proteins derived from these predictions can be seen as starting points to design experimental approaches to characterize the envelope proteome. But experimental proteomic analysis of integral membrane proteins has always proven to be difficult because of the hydrophobic nature of the proteins and their highly dynamic expression. Therefore, a variety of extraction strategies such as the solubilization of proteins with organic solvents such as chloroform/methanol mixtures or treatments of membranes with alkaline substances or salts but also different fractionation strategies like SCX columns have been used (Ephritikhine et al. 2004; Rolland et al. 2006). Subsequent protein identification was mostly done by nano/LC coupled mass spectrometry (Ferro et al. 2000, 2002, 2003; Froehlich et al. 2003). A wide variety of proteins have been identified with these different approaches. A head count done by Peltier et al. (2004b) resulted in a number of 429 identified proteins located in the chloroplast envelope of *Arabidopsis*. Many of these proteins have no known function yet another big part of them works in protein translocation or metabolism. Froehlich et al. (2003) were able to identify many components of the *Arabidopsis* Toc/Tic protein import complex in a large scale proteomic analysis also including proteases and chaperones associated with these complexes. A blue native PAGE (BN-PAGE) analysis in pea revealed the molecular organization of the Toc core complex consistent of Toc159, Toc75 and Toc34 (Kikuchi et al.

2006) with an estimated size of about 800-1000 kDa and a stoichiometry of 1:3:3 (Toc159:Toc75:Toc34). It also became clear that the A-domain of Toc159 is involved in stabilizing the association of Toc34 with the complex. As for the Tic complex, the analysis by BN-PAGE uncovered a new subunit, Tic62, and demonstrated that this subunit together with Tic110 and Tic55 forms a core protein complex in the Tic translocon (Kuchler et al. 2002). As mentioned above, not only proteins have to be transported through the envelope membranes. Therefore, it is not surprising that all proteome studies of the envelope so far yielded in a large number of identified transporters for metabolites, ions or other organic components. Very abundant in most studies are, just to mention a few, oxoglutarate/malate-, phosphate/triosephosphate-, sugar- or ABC-type transporters (Seigneurin-Berny et al. 1999; Ferro et al. 2003; Froehlich et al. 2003). Another large group of identified proteins is involved in lipid metabolism. Among these are synthases, desaturases and acyltransferases that metabolize fatty acids, glycerolipids, pigments or prenylquinones. Noteworthy are also enzymes like the allene oxide synthase that is involved in the metabolism of oxylipins, which are signal components in plant growth and defense reactions. Other identified proteins were, for example, components involved in the response of the plant to oxidative stress like superoxide dismutase or ascorbate peroxidase (Seigneurin-Berny et al. 1999; Ferro et al. 2003; Froehlich et al. 2003). In conclusion, these findings demonstrate that the envelope membranes are not only an important transport machinery, but additionally represent a specialized and essential part in chloroplast metabolism.

2.2 Stroma and chloroplast ribosome

The chloroplast stroma is enclosed by the envelope membranes. Most importantly, this compartment is the site of the light independent photosynthetic reaction, the Calvin cycle, but also the oxidative pentose phosphate pathway and glycolysis are located here. The stroma contains the chloroplast DNA as well as its complex translational machinery in the form of the 70S chloroplast ribosomes. The stromal proteome is considered to be quite intricate and a prediction made by Sun et al. (2004) resulted in the number of 3387 putative stromal proteins. Due to it's complexity, most studies of the chloroplast stroma focused on single protein components and so far only one large scale proteomics study aimed to investigate the complete *Arabidopsis* stromal proteome (Peltier et al. 2006). A two dimensional approach was used combining colorless-native PAGE (CN-PAGE) in the first and SDS-PAGE in the second dimension. Gels were subjected to analysis with MALDI-TOF or LC-ESI-MS/MS as well as to a semi quantitative approach using staining with CyproRuby followed by image analysis. A number of 241 nonredundant proteins was identified and sorted into functional categories. Interestingly, a significant part of the proteins (26%) was involved in protein metabolism such as synthesis, folding, sorting, and proteolysis. As expected numerous proteins associated with carbon metabolism were also detected (12%). These enzymes from the Calvin cycle, the oxidative pentose phosphate pathway and glycolysis made up about three quarters of the stromal protein mass. Another 21% of the pro-

teins had functions in nucleotide synthesis and degradation, amino acid metabolism or tetrapyrrole synthesis. It was not possible to assign a function to 11% of the identified proteins. Additionally, the authors extensively searched existing literature to determine the oligomeric state of the identified proteins in order to demonstrate the importance of functional paralogues in complex metabolic pathways.

2.2.1 The chloroplast ribosomes

The proteome of chloroplast ribosomes has an important role in the expression of chloroplast encoded genes and was therefore investigated in great detail. 2-DE-PAGE coupled with protein sequencing, HPLC, LC/MS and mass spectrometric analysis showed that the 70S ribosomes of spinach chloroplasts contain no less than 59 proteins; 33 in the 50S and 25 in the 30S subunit as well as a 70S complex associated ribosome recycling factor (Yamaguchi and Subramanian 2000; Yamaguchi et al. 2000). The 30S subunit that is responsible for mRNA binding and initiation of translation contains 21 *E. coli* orthologues as well as four plastid specific ribosomal proteins. In the 50S subunit, where the peptide synthesis takes place, 31 *E. coli* orthologues as well as two plastid specific ribosomal proteins were identified. These plastid specific ribosomal proteins termed PSRP1-6 are proposed to perform functions unique to plastid translation. The subsequent analysis of the chloroplast ribosomes of the unicellular green algae *Chlamydomonas rheinhardtii* using SDS-PAGE, LC MS/MS and mudPIT analysis resulted in the identification of 21 30S and 28 50S proteins as well as two proteins (RAP 38 and RAP 41) solely associated with the 70S complex (Yamaguchi et al. 2002, 2003). As for the 30S subunit in addition to several *E. coli* orthologues, the authors identified a PSRP3 homologue from spinach as well as a novel S1- domain containing protein named PSRP7. They were also able to determine that due to N-terminal extensions or other inserted sequences, three of the *Chlamydomonas* 30S ribosomal proteins are unusually large. In contrast to that, the composition of the 50S subunit is more conserved (27 *E. coli* orthologues, one spinach homologue, similar sizes). The authors, therefore, concluded that the differences in the composition of the 30S subunit might be related to unique features of *Chlamydomonas* chloroplast translational regulation, especially concerning mRNA discrimination at the 30S complex. In contrast to that, the enzymatic function of forming peptidyl bonds between amino acids in the 50S subunit is more conserved between organisms.

2.2.2 Plastoglobuli

Plastoglobuli (PG) are an additional component of the chloroplast stroma representing lipid containing structures that are mostly attached to the thylakoid membranes. They are known to accumulate α-tocopherol, plastoquinone, and triacylglycerols. Recently, the plastoglobuli proteome of *Arabidopsis* chloroplasts from WT plants grown under different light conditions and from the *clpr2-1* mutant that over accumulates PG were analyzed. NanoLC-ESI-MS/MS and a stable isotope labeling strategy were used to reveal the protein composition and acquire a func-

tional model of these structures as well as to characterize their function in chloroplast metabolism (Ytterberg et al. 2006). It became clear that plastoglobuli contain a specific proteome mostly consisting of proteins from the fibrillin family forming the coating of the particles. In addition to these, several proteins involved in lipid metabolism, quinone synthesis and regulation as well as a number of aldolases involved in Calvin cycle and/or glycolysis could be identified. The authors, therefore, concluded that PG are not only storage facilities, but have a defined function in several metabolic pathways. They state that PG represent a connection inbetween the thylakoid and the inner envelope membrane in the context of the metabolism of small molecules essential in thylakoid function and protection such as tocopherols and quinones. The important role of PG in tocopherol metabolism was also demonstrated in a parallel study in *Arabidopsis* (Vidi et al. 2006). The authors could show the localization of the tocopherol cyclase VTE1 within the PG using MS/MS, immunogold, and fluorescence labeling. In addition, they were able to identify a number of unclassified proteins and proteins from the plastid lipid associated protein/fibrillin-like family. Proteins involved in chloroplast processes like sugar and abscisic acid metabolism as well as jasmonic acid biosynthesis were also found.

2.2.3 The ferredoxin/thioredoxin system

The ferredoxin/thioredoxin system of the chloroplast is a key component of the regulation of photosynthetic enzymes in response to light facilitated through a number of redox processes. In an approach to find potential thioredoxin targets in the chloroplast, Balmer et al. (2003) bound thioredoxin f and m where one of the active Lys residues was replaced by Ser to a column in order to trap interacting stromal proteins from spinach. Specific proteins eluted from the column by addition of reducing equivalent (DTT) were then analyzed by 2-DE gels and MS. This led to the identification of several known along with a large number of so far unknown thioredoxin targets. It could be established that thioredoxins are also involved in regulation of so far unrecognized processes like isoprenoid, tetrapyrrole, or vitamin biosynthesis, protein assembly, folding, and degradation as well as processes involved in carbohydrate metabolism and DNA replication and transcription. In a subsequent study, the same group used a similar strategy including an affinity chromatography with WT thioredoxin f to trap proteins forming protein complexes based on protein/protein interactions (Balmer et al. 2004). The data revealed 27 so far unrecognized partners for protein/protein interaction with thioredoxin and indicated that not all of the thioredoxin targets identified to date that are able to form covalent interactions are able to interact electrostatically. The interaction partners cover a wide variety of chloroplast functions like Calvin cycle, translation, protein assembly and folding or other biosynthetic processes. The authors concluded that the formation of electrostatic complexes may help in the efficient transfer of electrons from photosystem I over the ferredoxin/thioredoxin system to the target proteins that can then be differentially regulated.

2.2.4 The Clp protease complex

An important component of the chloroplast stroma is the Clp protease complex, which, for example, degrades misfolded or unassembled proteins in an ATP dependent manner. Peltier et al. (2001; 2004a) identified and characterized a Clp protease complex of about 350 kDa in the stroma of *Arabidopsis* by employing BN-PAGE, CN-PAGE and native IEF/SDS-PAGE in combination with MALDI and ESI-MS. It could be shown that the complex is partially associated with the thylakoid membrane and consists of eleven different Clp proteins, five of which are serin-type proteases (ClpP1, 3-6) present in 1-3 copies per complex, four are non-proteolytic (ClpR1-4) and two have chaperone functions (ClpS1, 2) (Peltier et al. 2001, 2004a). This data, in addition to a detailed analysis of a ClpR2 deficient mutant of *Arabidopsis* with comparative quantitation using iTRAQ (isobaric tags for relative and absolute quantitation), and other approaches indicate a central role of the Clp protease complex proteins in plastid homeostasis as well as in chloroplast biogenesis and plant development (Peltier et al. 2004a; Rudella et al. 2006).

2.3 Thylakoid membrane

Oxygenic photosynthesis is the predominant function of the thylakoid membrane within the chloroplast. To perform this function and transform light energy into chemical energy in the form of ATP and NADPH, the tylakoid membrane contains four large multisubunit complexes involved in photosynthetic electron transfer and ATP synthesis. These are photosystem II (PSII), the cytochrome b_6f ($cytb_6f$) complex, photosystem I (PSI), and the ATP synthase. These complexes are distributed between two distinct membrane types. The stacked grana lamellae contain most of the PSII as well as $cytb_6f$ complexes, whereas the stroma lamellae contain most of the PSI as well as $cytb_6f$ and ATP synthase complexes (Timperio et al. 2004). In addition, the thylakoid membrane harbors proteins involved in assembly and maintenance of the lipid bilayer and the proteins therein. These include proteins for folding, incorporation, modification and degradation of the photosynthetic complexes as well as components of a complex transport machinery (Friso et al. 2004). The thylakoid proteome is a highly dynamic system since it requires the ability to adapt to changing environmental conditions such as increasing light intensities or changing temperature in order to especially protect the photosynthetic machinery from damage (Aro et al. 2005). Analysis of the thylakoid proteome proved to be difficult due to the strong prominence of the photosynthetic proteins that enhance the problem of dynamic resolution (van Wijk 2004) and due to the strong hydrophobicity of the integral membrane proteins (Whitelegge et al. 2006). Therefore, a wide range of approaches was employed to identify and analyze the components of the membrane and its associated proteins. The most comprehensive investigations of the thylakoid proteome involved combinations of several fractionation strategies to avoid the problems mentioned above. Ultracentrifugation, aqueous polymer two phase partitioning and two dimensional SDS-PAGE combined with MS were applied in a study of the *Synechocystis sp. PCC*

6803 thylakoid proteome (Srivastava et al. 2005). This yielded a pure integral membrane fraction where 76 proteins could be identified. Only 14 of these had transmembrane domains whereas the other proteins were peripherally located, most likely on the cytosolic side of the membrane. With the applied strategies the authors were able to not only resolve the abundant photosynthetic proteins, but also a multitude of proteins having functions in protein sorting mechanisms, pigment biosynthesis, protein folding or hypothetical proteins with yet unknown function. In a literally "in depth analysis" of the thylakoid proteome of *Arabidopsis* salt, detergent, and organic solvent extraction in combination with different multidimensional protein separation techniques, enzymatic, and non-enzymatic protein cleavages, MALDI and ESI-MS as well as bioinformatic approaches were used to generate an overview of peripheral and integral thylakoid proteins (Friso et al. 2004). Three distinct subproteomes could be separated: a peripheral, a peripheral but tightly associated and an integral membrane proteome. The peripheral subproteome was combined with proteins previously identified in a study of the lumenal proteome from the same group (Peltier et al. 2002) resulting in a set of 99 proteins that were dominated by the proteins of the oxygen evolving complex from PSII but also contained a number of unique new proteins. The hydrophobic fractions included 134 proteins with 76 of them having one or more known or predicted transmembrane domains. In the complete set of 198 non-redundant proteins, many (42%) photosynthetic proteins could be detected representing 85% of the known proteins in the four multisubunit complexes. In addition, a large number (15%) of new proteins with unknown function was identified such as two rubredoxins, a metallochaperone and a new DnaJ-domain protein. Other proteins found were involved in translation, metabolism or protein fate as well as in the protection from oxidative stress. The combination of extraction strategies used here allowed the detection of low abundant as well as small and very hydrophobic proteins that can hardly be resolved with standard 2-DE gel approaches. In a subsequent study, the same group developed a faster fractionation strategy employing three phase partitioning (TPP) of salt stripped thylakoids combined with RP-nano-LC-ESI-MS/MS. The authors again combined the data from the TPP approach with the analysis mentioned before (Friso et al. 2004) as well as their results of the lumenal and peripheral proteome of the thylakoids (Peltier et al. 2002) to achieve a combined dataset of more than 300 proteins. Whereas all other fractionation strategies mostly identified photosynthetic or other abundant thylakoid proteins, the TPP revealed a whole new level of lower abundant proteins, of which 50% were unknown. Others were involved in chlorophyll/prenyl lipid biosynthesis, protein sorting or degradation, stress defense or signaling. These studies show the significant improvement of dynamic resolution by the improvement of fractionation strategies and, concomitant with that, an increasing understanding of the thylakoid membrane proteome of higher plants. The combined datasets including functional assignments as well as other data from chloroplast proteomic studies from the van Wijk group can be accessed via the Plastid Proteome Database at Cornell (http://ppdb.tc.cornell.edu/). Still there are other approaches employed to characterize the thylakoid proteome. In a recent publication, Allmer et al. (2006) were able to create a comprehensive overview on the thylakoid proteome of

Chlamydomonas reinhardtii by combining SDS-page and ESI MS/MS of purified thylakoid membrane fractions obtained by ultracentrifugation with a high throughput genomic data mining strategy. This resulted in the detection of numerous low abundant proteins consisting of, for example, a novel light-harvesting protein, the STT7 kinase and a DegP-like protease along with numerous proteins with unknown function.

2.3.1 Analysis of the photosynthetic machinery

Numerous studies of thylakoid proteins focused on the proteins of the photosynthetic machinery to elucidate the composition of the different supercomplexes and their response towards changing environmental conditions. BN-PAGE is a method that is often employed to investigate the native state and thereby the interaction and the composition of the photosynthetic complexes. Combined with a second dimension SDS-PAGE and MS, this approach additionally enables the identification of subunits of the complexes. Studies in several higher plants gave insights in the supramolecular organization of the photosynthetic structures and revealed the existence of supercomplexes of PSI and PSII as well as of different forms of the photosystemI-light harvesting complexI (PSI-LHCI) and the PSII core complex. The employment of native gels also made it possible to resolve dimeric LHCI as well as monomeric and trimeric light harvesting complexII (LHCII), thus, supporting the results from diverse X-ray crystallographic analyses as well as the subunit organization and composition of the cytochrome b_6f complex and the ATP synthase (Heinemeyer et al. 2004; Ciambella et al. 2005; Granvogl et al. 2006). Still BN-PAGE analysis struggle to detect very small and hydrophobic proteins most likely due to the fact that these proteins have less tryptic cleavage sites and therefore often escape mass spectrometric detection (Granvogl et al. 2006). Also the dynamic resolution is not sufficient enough to detect very low abundant proteins making the discovery of unknown proteins very difficult. However, BN-PAGE represents a good alternative to the resolution of sucrose gradient for the separation of protein complexes. Additionally, this method can be used to show alterations in the protein complex structure in response to environmental factors or the metabolic state of the cell as was shown for *Chlamydomonas* cell grown in either photoautotrophic or photoheterotrophic conditions (Rexroth et al. 2003).

PSI-LHCI Complex. High performance liquid chromatography, used after prefractionation of the photosynthetic complexes through solubilisation and either differential or ultra centrifugation, can also be employed to analyze the thylakoid protein composition on an intact molecule level (Huber et al. 2004; Timperio et al. 2004). In this context, Zolla and Timperio (2000) were able to resolve most of the components of the spinach PSI-LHCI complex and to determine their molecular masses. In addition, it was shown on five dicotyledonous and four monocotyledonous plants that all species investigated possessed isoforms of the Lhca1 protein (Zolla et al. 2002). With classical denaturating one and two dimensional gel electrophoresis and MS/MS Storf et al. (2004) were for the first time able to identify the very low expressed Lhca5 gene product on the protein level. The authors showed the existence of several isoforms of all Lhca proteins (Lhca1-4) in tomato,

demonstrating the presence of different populations of PSI in higher plants. A two dimensional SDS-PAGE approach was used to separate even highly hydrophobic proteins from the green alga *Chlamydomonas* when sample preparation was adapted towards the specific requirements of membrane proteins and resulted in a 2-DE map resolving LHCI and LHCII components in addition to other thylakoid proteins (Hippler et al. 2001). Using this improved 2-DE technique and mass spectrometry in a subsequent publication, Stauber et al. (2003) were able to establish a detailed 2-DE map of the *Chlamydomonas* light harvesting proteins demonstrating the expression of nine different Lhca as well as eight Lhcb proteins. In addition, a differential modification of Lhcbm3 and Lhcbm6 could be shown. According to these studies, it appeared that the LHCI of *Chlamydomonas* is significantly larger than that of higher plants. In an analysis of extracted PSI-LHCI and LHCI complexes by western blotting and 2-DE, Takahashi et al. (2004) additionally demonstrated that the core of the LHCI complex in *Chlamydomonas* is different from the dimeric structures of that of higher plants and forms a stable oligomeric complex that is able to assemble in the absence of the PSI core. It could further be demonstrated that three of the Lhca proteins (Lhca2, 3, 9) are only able to associate with the LHCI when they are stabilized by the presence of the PSI core what might be due to a functional role of these subunits in excitation energy transfer.

PSII-LHCII Complex. The PSII-LHCII complex was also object of detailed analysis of its structure and function using a number of different proteomic approaches. One study used SDS-PAGE in combination with protein sequencing and MS to analyze a highly purified PSII-LHCII complex containing a HIS-tag on the *psbB* gene product from *Synechocystis* (Kashino et al. 2002). This resulted in the detection of all known PSII subunits and some novel proteins representing potential candidates for the functional regulation of PSII. Using RP-HPLC-ESI-MS analysis on several different dicotyledonous and monocotyledonous plants, it was possible to characterize unique chromatographic patterns for each species. It became clear, that in monocots the LHCII complex appeared as mono- and trimers, whereas in dicots the trimeric form seems predominant. In addition, several isoforms of Lhcb1, 3, and 6 were detected that were predicted from the sequences of the multigene families coding for these proteins, but could not be resolved well with traditional gel based approaches (Huber et al. 2001; Zolla et al. 2003). These isoforms may have a role in the adaptation of PSII to different light conditions.

2.3.2 Post-translational modifications

Post-translational modifications, especially phosphorylation of photosynthetic membrane proteins, also play an essential role in the redistribution of excess light energy from the LHCII in PSII to the PSI in a process known as state transition that includes a migration of LHCII proteins from the grana to the stroma thylakoid regions (Vener et al. 2001; Timperio and Zolla 2005). Employing mass spectrometry on thylakoid membranes from *Arabidopsis* that were enriched for phosphoproteins using immobilized metal affinity chromatography (IMAC), Vener et al. (2001) were able to map phosphorylation sites of the central photosynthetic proteins including LHCII, D1, D2, CP43 and PsbH. In a subsequent analysis using

this approach, they were able to identify in addition to the previously known phosphoproteins three more phosphorylation sites in a peptide from CP29, an expressed membrane protein, and for the first time in a PSI protein, namely PsaD (Hansson and Vener 2003). It became clear that the phosphorylation sites of all these *Arabidopsis* proteins are located at threonine residues near the N-terminus of the protein (Vener et al. 2001; Hansson and Vener 2003). In an analysis of the trypsin shaved thylakoid membrane of *Chlamydomonas*, enriched for phosphoproteins Turkina et al. (2004) detected a very unusual LHCII protein. It was demonstrated that the CP29 in its mature form still contained its transit peptide and solely showed a N-terminal methionine excision as well as a phosphorylation and acetylation site. This might represent an evolutionary compromise to keep the TP and in turn its functionally important phosphorylation site (Turkina et al. 2004). It is of note that the phosphorylation sites of *Chlamydomonas* Lhcbm proteins are also closer to the N-terminus as expected from predicted cleavage sites of the transit peptides (Stauber et al. 2003; Turkina et al. 2006b). Interestingly, N-terminal processed forms of Lhcbm proteins in *C. reinhardtii* exist, which lack these phosophorylation site, suggesting a novel type of regulation for *Chlamydomonas* Lhcbm proteins (Stauber et al. 2003). RP-HPLC can also be used to detect PTMs as shown on the PSII-LHCII complexes of pea and spinach where phosphorylation of for example D1, D2, CP43, two Lhcbs and PsbH could be resolved (Gomez et al. 2002). This approach was also capable of actually showing the migration of LHCII proteins like Lhcb2 and several isoforms of Lhcb1 from the grana to the stroma regions (Timperio and Zolla 2005). But it became clear that the phosphorylation state of the proteins was not the determining factor for the movement of the proteins, since some migrated in their unmodified form, but rather structural changes in the thylakoid organization due to PTMs as well as the pigment composition might be the cause for the migration of LHCII proteins to the PSI. Another important feature of phosphorylation is the maintenance of PSII by controlling the turnover of its reaction center proteins (Vener et al. 2001). This has, for example, been demonstrated by employing BN-gels to visualize the photoinhibition repair cycle on the basis of different PSII complexes represented in the native gel (Aro et al. 2005). With the employment of a wide variety of proteomic approaches including diverse fractionation strategies it became possible to rapidly increase the understanding of the complex organization, function, and regulation of the chloroplast thylakoid membrane and the complexes of oxygenic photosynthesis located therein. The rapidly evolving field of proteomics makes it possible to slowly overcome the limitations created by the physico-chemical character of a protein, its post-translational modification as well as dynamic resolution problems caused by the high abundance of the photosynthetic proteins.

2.4 Thylakoid lumen

The thylakoid lumen is the space enclosed by the thylakoid membranes. It is known that it contains proteins involved in oxygenic photosynthesis like the extrinsic subunits of PSII PsbO, PsbP, and PsbQ that function in the stabilization of

the water oxidizing complex as well as the electron carrier protein plastocyanin. With the increased use of 2-DE separation techniques combined with mass spectrometry it became clear that the lumen harbors a lot more proteins than the ones involved directly in primary aspects of photosynthesis (Kieselbach et al. 1998). After the genome of *Arabidopsis thaliana* was sequenced the efforts to identify these proteins increased and yielded in the identification of so far about 40 to 50 proteins from *Arabidopsis*, spinach, and pea (Kieselbach and Schroder 2003; Sun et al. 2004). A part of these proteins was related to the already known subunits of PSII. Isoforms with unknown functions of PsbO and plastocyanin were found as well as many proteins with a PsbP domain and the PSII assembly factor Hcf136 (Kieselbach et al. 2000; Peltier et al. 2000, 2002; Schubert et al. 2002). The xanthophyll cycle enzyme violaxanthin deepoxidase as well as different peroxidases have an important function in the protection of PSII from oxidative stress and were identified in several studies (Peltier et al. 2002; Schubert et al. 2002). The lumenal proteome was also shown to be rich in putative immunophilins. They belong to the cyclophilin-type peptidy-prolyl *cis-trans* isomerases (PPIases) or to the FKBP-type PPIases that are involved in protein folding and might as well work in chaperoning or have regulatory functions (Peltier et al. 2002; Schubert et al. 2002; Kieselbach and Schroder 2003). Numerous proteases of different types were identified in the lumen as well. Their roles include, among others, the processing of the D1 protein (Schubert et al. 2002; Kieselbach and Schroder 2003). A surprisingly large group of proteins with a pentapeptide repeat were also found under the lumenal proteins, but their sequences did not show any other known functional domains (Peltier et al. 2000, 2002; Schubert et al. 2002; Kieselbach and Schroder 2003). Proteins located on the thylakoid periphery of the lumen were additionally investigated by Peltier et al. (2000; 2002) and contained a large number of fibrillins, which might function in carotenoid storage and a ClpS1 protease from the thylakoid associated Clp protease complex.

To help in the identification of possible lumenal proteins, efforts were also put in the prediction of the complete lumenal proteome. Predictions of proteins with plastidal and lumenal transit peptides made with TargetP and SignalP (Nielsen et al. 1997) resulted in only 80 putative proteins for *Arabidopsis* (Schubert et al. 2002; Kieselbach and Schroder 2003) and at least 200 candidate proteins for *Arabidopsis* and pea (Peltier et al. 2000, 2002). The difficulties with the prediction programs used led to the development of LumenP (Westerlund et al. 2003), a neural network predictor especially designed to identify lumenal target sequences. A combined prediction using TargetP and LumenP on the *Arabidopsis* open reading frames resulted in 417 proteins to be potentially located in the thylakoid lumen. An independent comparison of predictions from SignalP with predictions from LumenP yielded in a number of 285 and 291 lumenal proteins, respectively. Still, only 150 proteins were predicted by both algorithms and from the 100 proteins that were different between the programs, only 26 overlapped with the 53 experimentally shown lumenal proteins, which the authors collected from literature (Sun et al. 2004). In the course of the combined efforts to investigate the proteome of the lumen, it became clear that in order to transport proteins from the chloroplast stroma through the thylakoid membrane into the lumen, the delta pH dependent

TAT pathway (requiring a twin Arginine motive in the transit peptide) plays a much more pronounced role than the ATP driven Sec pathway, and additionally that the twin arginine motive makes a prediction of a lumenal protein more reliable than any other amino acid sequence in the lTP (Westerlund et al. 2003; Sun et al. 2004). It is noteworthy that all types of predictions have the tendency to be erroneous and have to be verified with experimental data.

3 Predictions and collections of the chloroplast proteome

The knowledge of the subcellular localization of so far uncharacterized proteins can provide valuable information to elucidate their function in specific metabolic processes and to get more insights into cellular functions as a whole (Heazlewood et al. 2005). For this reason, efforts were made to develop software tools like SignalP (Nielsen et al. 1997), ChloroP (Emanuelsson et al. 1999), TargetP (Emanuelsson et al. 2000), or Predotar (Small et al. 2004) to predict N-terminal transit peptides which determine the subcellular localization of the proteins as well as their putative cleavage sites. The major drawback of this strategy is that even TargetP, the currently most successful prediction program, has a prediction efficiency of only 70-85% for vascular plant gene products, whereas its prediction accuracy is even less when it comes to green alga (Kleffmann et al. 2006). This high number of erroneous predictions has several reasons. One is that transit peptides are usually not well conserved and quite diverse in length and amino acid composition. Some proteins of the outer envelope membrane are also known to have no target peptide at all (Jarvis 2004; Richly and Leister 2004; Lunn 2006). Another problem is the limited number of proteins that are usually used to train the predictors making it not surprising that results can be inconsistent when using different prediction tools (Heazlewood et al. 2005). Even though most of the used training sets contain a subset of algae proteins, predictions for these are often especially erroneous. The problem here might be based on the fact that algal transit peptides are on average 32 amino acids shorter than those of vascular plants and therefore more difficult to predict with the current mixed training sets used (Gomez et al. 2003).

Nevertheless, in addition to the limitations of computer algorithms, the targeting of some proteins will never be recognizable with such tools. The existence of proteins without target sequences and the employment of unconventional import strategies such as via the secretory pathway and dual targeting, demonstrates the need for alternative approaches to elucidate organelle targeting (for reviews see: Millar et al. 2006; Radhamony and Theg 2006). Such limitations will result in an underestimation of organellar proteomes as, for example, the proteome of the chloroplast. Still computer based predictions of the possible localization of a protein can be very useful since they can give indications on the presence of low abundant proteins that would not be detectable in an experimental proteomic approach due to dynamic range limitations (Newton et al. 2004). Nevertheless, experimental proof is essential to verify the predicted location for these proteins.

With the sequencing of the *Arabidopsis thaliana* genome, this plant became the primary subject to predict chloroplast proteins with estimated numbers ranging

from about 3000 to 4000 proteins for the complete organelle (Peltier et al. 2002; Schubert et al. 2002; Kieselbach and Schroder 2003; Baginsky and Gruissem 2004; Kleffmann et al. 2004; Richly and Leister 2004; Sun et al. 2004; van Wijk 2004). These collected sets of protein predictions were then used to predict the localization of proteins within the different compartments of the chloroplast (see chapters in the article). Sun et al. (2004) made a large effort to improve these strategies by collecting a set of about 250 proteins with different compartment locations from published data. The analysis of biochemical properties such as the cysteine content, the protein-size or the number of transmembrane domains revealed that, for example, the cysteine content of proteins in the thylakoid membranes or the lumen is much lower than in proteins of the envelope membranes. The authors were also able to show differences in the size, pI and number of transmembrane domains within proteins from different sub compartments. These characterizations can be helpful in determining the location of a protein and to further optimize prediction programs.

Recently, a number of databases were created containing huge collections of experimental as well as in silico data to provide information on the proteomes of cellular as well as organellar components, especially from *Arabidopsis*.

Heazlewood et al. (2006) developed SUBA, a subcellular database that combines information from mass spectrometric and fluorescent protein experiments, as well as several other databases, to a total of more than 6700 nonredundant proteins that can be assigned to ten distinct locations within the cell. The SUBA database gives the user the possibility to build protein sets out of published or newly imported data including the results of different prediction programs and compare these with the database entries (http://www.plantenergy.uwa.edu.au/applications/suba/index.php). plprot, developed by Kleffmann et al. (2006) is a database that contains collected experimental data on different plastid types like chloroplasts, etioplasts, and undifferentiated plastids. It contains 2043 partially redundant protein entries from *Arabidopsis*, rice, and BY2 cells including a variety of information on the single proteins. It also features a plastid type comparison that makes it possible to compare proteins within the different datasets (http://www.plprot.ethz.ch/). A comprehensive chloroplast specific database, PPDB, is available from the lab of KJ van Wijk (Friso et al. 2004) (http://ppdb.tc.cornell.edu/). It combines predicted and experimentally identified proteins from all suborganellar locations, including also information on fluorescent protein experiments from the SUBA database. Protein entries include annotations, predicted and experimentally determined molecular and biophysical properties, as well as information on protein-protein interactions, results from comparative proteomics studies, functional classifications, and schematics on biochemical pathways.

4 Comparative proteomics

With the increasing amount of information about the various subproteomes of cell organelles, it became clear that in order to understand complex biological processes it is not only important to understand gene expression but also to analyze complex protein expression patterns that are strongly influenced by post-transcriptional and post-translational modification processes (Steen and Pandey 2002). Therefore, much effort was put into the comparison of protein abundance in different sets of samples that derived from, for instance, different developmental stages of the organelle or from cells under diverse biotic or abiotic stress conditions. Several methods used rely on gel systems like the classical two-dimensional SDS-PAGE with isoelectric focusing as a first dimension or the difference gel electrophoresis. Since these methods cope with the known problems of gel based protein separation, like the difficulties in the resolution of strong hydrophobic membrane proteins, non-gel-based methods were developed that employ stable isotopes like ICAT, SILAC, or iTRAQ (isobaric tags for relative and absolute quantitation). Nevertheless, in chloroplast proteomics, most comparative studies so far rely on 2-DE and concomitant image analysis and mass spectrometry.

4.1 Plant and chloroplast development

Several studies used proteomics to investigate developmental stages of whole leaves as done in rice (Zhao et al. 2005) or on senescent leaves of white clover (Wilson et al. 2002). This included the analysis of a chloroplast fraction and demonstrated the organized breakdown of organelles and macromolecules and increased levels of proteins involved in remobilization of nutrients that are relocated to developing plant parts. 2-DE and a detailed data evaluation were used to characterize chloroplast biogenesis in maize that led to the identification of 26 unique spots on gels from 5 different time points. These proteins were mostly from the light reaction and the carbon assimilation cycle of photosynthesis but also chaperones and other metabolic enzymes (Lonosky et al. 2004). In a plant cell, a fully differentiated chloroplast can, dependent on the tissue it is located in, serve a variety of functions as for example known from mesophyll and bundle sheath cells in C4 plants. This aspect was analyzed in detail using maize chloroplast stroma for 2-DE and image analysis as well as differential labeling with cleavable ICAT and a comparison of unlabeled stroma proteins by LC-ESI-MS (Majeran et al. 2005) (Figure 3). The three approaches proved to be complementary and resulted in the identification of 400 proteins from a wide variety of pathways and a detailed overview of differential protein accumulation in chloroplasts from mesophyll or bundle sheath cells that are mainly due to the metabolic differentiation of the two cell types. The authors additionally provided evidence for a differential regulation of plastid gene expression, protein biogenesis and protein fate and presented a number of so far unknown proteins that are specifically expressed in one tissue and probably have central functions in the C4 plant metabolism.

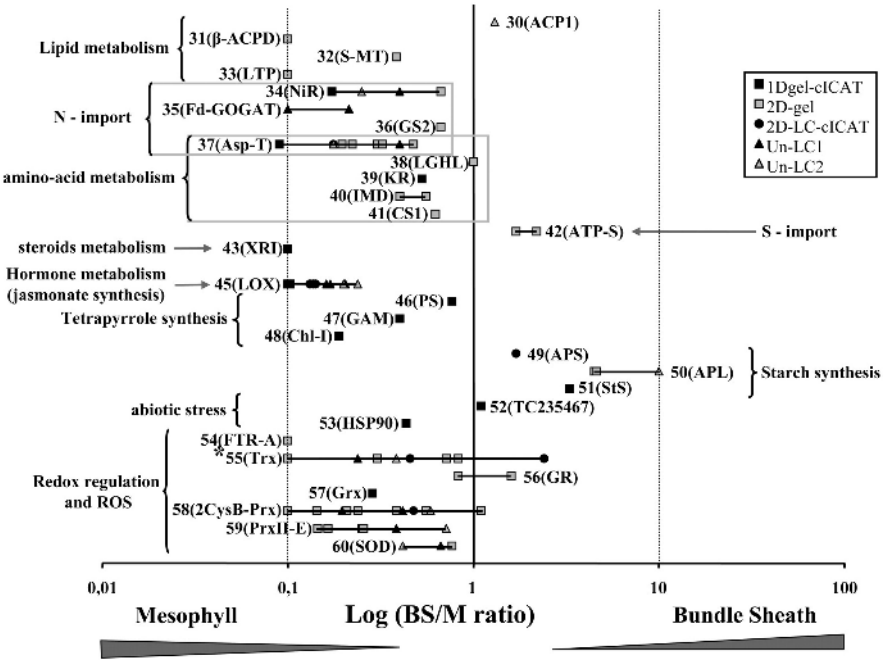

Fig. 3. Overview of bundle sheath: mesophyll ratios of selected proteins. Values were determined by three complementary methods, 2DE and image analysis, differential labeling with cleavable ICAT, and a comparison of unlabeled stroma proteins by LC-ESI-MS. Taken with permission from Majeran et al. (2005).

4.2 Biotic stress

Comparative proteomics was also used to characterize the answers of plants to biotic or abiotic stress factors. The response of *Arabidopsis* to a pathogen attack of *Pseudomonas syringae* pv. *tomato*, for example, was shown to induce phosphorylation of leaf proteins whose abundances could be compared using iTRAQ (Jones et al. 2006a). The attack also led to the induction of characteristic protein changes in the total soluble leaf proteome. In addition, chloroplast and mitochondrial fractions that mainly contained defense related antioxidants and metabolic enzymes were analyzed with 2-DE (Jones et al. 2006b). In a study on *Nicotiana benthamiana*, thylakoid membranes 2-DE enabled the identification of the PsbO and PsbP proteins as well as at least four different isoforms respectively. When plants were inoculated with the Spanish strain of pepper mild mottle virus PsbP proteins selectively decreased dramatically in abundance as compared to PsbO proteins, which can be seen as a counteraction to the virus infection by regulating photosynthetic activity probably as a basic defense mechanism (Perez-Bueno et al. 2004).

4.3 Abiotic stress

4.3.1 Light and temperature

Abiotic stress can be caused by several environmental conditions. For plants, light is the most important requirement for life but excess light can severely damage the cells especially with the accumulation of reactive oxygen species (ROS). As was shown for *Arabidopsis* chloroplasts by 2-DE and image analysis the largest damage can be observed in the protein complexes involved in photosynthesis or in protein metabolism whereas proteins with increased abundance under high light conditions were mainly defense related like scavengers for ROS or chaperones (Phee et al. 2004). In a later study on the effects of high light, Giacomelli et al. (2006) analyzed the *Arabidopsis* thylakoid lumenal and peripheral proteomes as well as the thylakoid associated plastoglobuli from wild type and the vtc2-2 mutant, containing only 20-30% of WT level of ascorbate, using proteomics and physiological experiments. Seven proteins were found upregulated in both genotypes under high light conditions: YCF37, four members of the fibrillin protein family, Fru-biphosphate aldolase-1 and a flavin reductase-related protein, of which the latter three types of proteins were located in the PG. The authors, therefore, concluded that the PG are probably involved in the synthesis and accumulation of α-tocopherol and quinones which are major antioxidants and that breakdown of carotenoids and turnover of lipids/fatty acids might also be a plastoglobuli associated process. High light response of the soluble proteome led to the upregulation of only the Ser type IV thylakoid protease SPPA whereas many other proteases showed no significant response. Since ascorbate is an important antioxidant, the analysis of the ascorbate-mutant was expected to reveal an additional level of specific stress responses to high light, but finally, ascorbate deficiency showed only small effects on the stress reaction since the mutant, compared to the WT, only differentially accumulated proteins that belonged to known stress response functions like the superoxide dismutases or some chaperones. In a study on whole *Chlamydomonas* cells, including two very high light resistant mutant strains, Forster et al. (2006) were able to identify new candidate proteins possibly involved in high light resistance like a DEAD box RNA helicase-like protein, NAB1 and RB38 where the latter two might be especially involved in the stability and function of the LHC and PSII under high light conditions. The effects of different light and temperature conditions on thylakoid pigment binding proteins, especially of Lhcb family members, was investigated in maize (Caffarri et al. 2005). Physiological analysis showed that plants grown in low temperature had an increased photoinhibitory damage, and in combination with high light, also an increased non-photochemical quenching. Low temperature and high light additionally increased the LHCII content whereas PSII core and PSI-LHCI proteins decreased in abundance. 2-DE gels also demonstrated a differential accumulation of several individual Lhcb1-3 proteins under different growth conditions indicating that diverse LHCII isoforms corresponding to multiple Lhcb1-3 genes might be needed for the acclimation to changing light and temperature conditions. Also low temperature alone can lead to significant changes in the chloroplast proteome.

Goulas et al. (2006) characterized lumenal and stromal proteins of *Arabidopsis* plants that were grown at low temperatures for different periods of time using a DIGE approach (Fig. 4). Most changes in the protein pattern could be observed after ten days in the cold. Changes in the stromal proteome included the upregulation of Rubisco whereas the other Calvin cycle enzymes decreased. An increase was also observed for some proteins from PSI and several enzymes related to oxidative stress. In the lumen only a few proteins like PsbP1 and PsbO2 increased abundance in the cold and some immunophilins showed variable responses.

4.3.2 CO_2 and iron

Another important aspect in plant life is the availability of carbon dioxide in order to sustain oxygenic photosynthesis. Under low CO_2 conditions, especially algae induce CO_2 concentrating mechanisms (CCM). However, even before these are completely developed, carbon dioxide limitation leads to a specific redox-dependent phosphorylation of two proteins, Lci5 and UEP (unknown expressed protein) that could be characterized in extrinsic thylakoid protein preparations using IMAC and ESI/TOF in *Chlamydomonas* chloroplasts (Turkina et al. 2006a). The data also indicated that thylakoids might contain a redox-dependent protein kinase specifically activated in the early stages of CCM.

Since plants are carrying out photosynthesis, they are highly dependent on metal components like iron or copper that function as cofactors in many enzyme complexes. Changes occurring in the thylakoid membrane at the onset of iron deficiency were analyzed in detail in several studies. Moseley et al. (2002) employed fluorescence emission analysis, immunoblots and 2-DE gel electrophoresis on thylakoid membranes of the green alga *Chlamydomonas*. They could demonstrate that iron deficiency leads to a functional uncoupling of the LHCI antenna from the PSI core leading to an impaired efficiency in the excitation energy transfer. Furthermore, adaptation to iron deficiency leads to a distinct and highly coordinated remodeling of the photosynthetic antenna complexes of LHCI and to a pronounced decrease in the abundance of PSI. Importantly, the remodeling of the photosynthetic apparatus became evident before a chlorotic phenotype was visible.

In a following publication Naumann et al. (2005) used a SILAC strategy to further characterize the remodeling process of the LHCI in *Chlamydomonas* and showed that, whereas Lhca5, 1, 7, and 8 are reduced in abundance, Lhca4 and 9 are induced. They also demonstrated that the N-terminal processing of Lhca3 occurs at a functionally assembled PSI-LHCI complex and could therefore be regarded as a key regulatory step in the remodeling process (see Fig. 4). Employing the SILAC approach, it was shown that the onset of iron deficiency additionally leads to alterations in the abundance of a variety of thylakoid proteins not directly involved in primary photosynthetic processes (Naumann et al., unpublished data). Iron deficiency was also analyzed on thylakoid membranes of sugar beet with IEF-SDS-PAGE as well as BN-SDS-PAGE in combination with mass spectrometry (Andaluz et al. 2006) (Fig. 4). In this study, it was demonstrated that iron deficiency leads to a pronounced decrease in proteins involved in photosynthetic

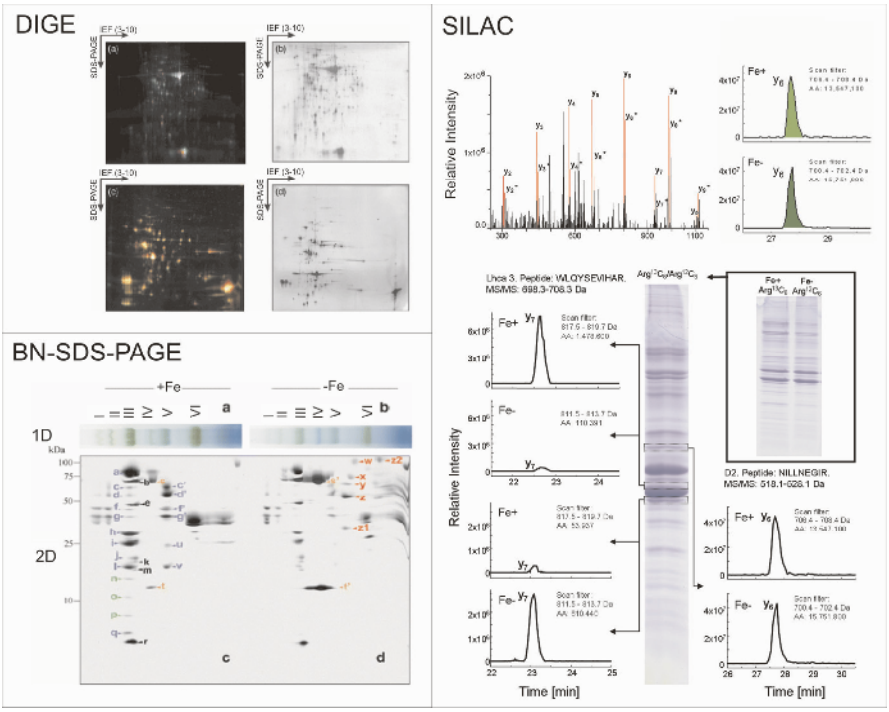

Fig. 4. Examples of comparative proteomics approaches, DIGE analysis of stromal and lumenal proteins of *Arabidopsis* plants, grown at low temperatures for different periods of time (Goulas et al. 2006), BN-SDS-PAGE of thylakoid membranes from Fe-sufficient and Fe-deficient samples from sugar beet (Andaluz et al. 2006), and fragmentation spectrum and elution profile for abundance calculation of a SILAC labeling experiment with labeled arginine on thylakoids from *Chlamydomonas* cells grown under Fe-limiting or standard conditions (Naumann et al. 2005). All figures taken with permission.

electron transport, whereas Calvin cycle enzymes partially increase in abundance. Also increased were proteins from biosynthetic pathways as well as stress related proteins. Comparing the two methods employed, it became clear that they are complementary since BN-SDS-PAGE is better suited for the resolution of hydrophobic membrane proteins, whereas IEF-SDS-PAGE has a higher resolving power. The use of proteomics to investigate plastid protein dynamics, adaptation, and acclimation to diverse stress conditions demonstrate that it is an emerging key technique to tackle these kinds of questions in plastid biology. A short summary of methods used in comparative proteomic approaches as well as their field of application is presented in Table 1; selected examples are additionally illustrated in Figure 4.

Table 1. Overview of approaches used in comparative proteomics experiments

Proteomic approach	Protein identification	Topic	Reference
2DE and image analysis	MALDI-TOF	Chloroplast biogenesis (Maize)	Lonosky et al. (2004)
	MALDI-TOF, LC-MS/MS, IB*	Leaf development (Rice)	Zhao et al. (2005)
	MALDI-TOF	Leaf senenscence (White clover)	Wilson et al. (2002)
	N-terminal sequencing, IB	OEC after pathogen attack (*Nicotiana benthamiana*)	Perez-Bueno et al. (2004)
	LC-MS/MS	Defence proteome after pathogen attack (*Arabidopsis*)	Jones et al. (2006b)
	LC-MS/MS, IB	Effects of iron deficiency on thylakoid proteome (*Chlamydomonas*)	Moseley et al. (2002)
	MALDI-TOF (/TOF)	Effects of iron deficiency on thylakoid proteome (*Beta vulgaris*)	Andaluz et al. (2006)
	LC-MS/MS, IB	Effects of high light on thylakoid proteome (*Arabidopsis*)	Giacomelli et al. (2006)
	MALDI-TOF	Effects of high light on chloroplast proteome (*Arabidopsis*)	Phee et al. (2004)
	MALDI-TOF	Effects of high light on chloroplast proteome (*Chlamydomonas*)	Forster et al. (2006)
	IB	Effects of diverse light and temperature on Lhcb gene products (Maize)	Caffarri et al. (2005)
	MALDI-TOF, LC-MS/MS	Proteomes of bundle sheath and mesophyll cells (Maize)	Majeran et al. (2005)
DIGE	MALDI-TOF (/TOF)	Effects of low temperature on chloroplast lumen and stromal proteome (*Arabidopsis*)	Goulas et al. (2006)
IMAC	ESI-MS/MS, IB	Phosphorylation of extrinsic thylakoid proteins under CO_2 limitation (*Chlamydomonas*)	Turkina et al. (2006a)
iTRAQ	LC-MS/MS	Defense phosphoproteome of leaves after pathogen attack (*Arabidopsis*)	Jones et al. (2006a)
BN-SDS-PAGE	MALDI-TOF (/TOF)	Effects of iron deficiency on thylakoid proteome (*Beta vulgaris*)	Andaluz et al. (2006)
SILAC	LC-MS/MS, IB	Remodeling of PSI in iron deficiency (*Chlamydomonas*)	Naumann et al. (2005)
cICAT and comparative Ion chromatography with LC-MS	LC-MS/MS	Proteomes of bundle sheath and mesophyll cells (Maize)	Majeran et al. (2005)

*IB: immunoblot

5 Conclusion

In the recent years, many proteomic studies of the chloroplast revealed a large number of proteins with yet unknown function. This calls for further analysis to elucidate the function and role in physiological processes of these gene products by reverse genetics experiments. It also became clear that comparative proteomics is an important tool to analyze dynamic changes in the chloroplast proteome. It is foreseeable that absolute quantitation, especially of individual subunits stemming from multiprotein complexes will open new perspectives in analyzing structures and functions in response to changing cellular environment. In this context, it will also be important to focus on the pronounced dynamic and variability of a proteome specifically due to post-translational modifications. Proteomics will also play a role in increasing the understanding of protein/protein interactions, and in combination with detailed gene expression studies help to improve our knowledge of the biology in complex cell organelles like the chloroplast in a whole systems perspective.

Acknowledgement

We are grateful to Mia Terashima and Dr. Jens Allmer for fruitful discussions and critical reading of the manuscript.

References

Aebersold R, Mann M (2003) Mass spectrometry-based proteomics. Nature 422:198-207

Allmer J, Markert C, Stauber EJ, Hippler M (2004) A new approach that allows identification of intron-split peptides from mass spectrometric data in genomic databases. FEBS Lett 562:202-206

Allmer J, Naumann B, Markert C, Zhang M, Hippler M (2006) Mass spectrometric genomic data mining: Novel insights into bioenergetic pathways in *Chlamydomonas reinhardtii*. Proteomics 6:6207-6220

Andaluz S, Lopez-Millan AF, De Las Rivas J, Aro EM, Abadia J, Abadia A (2006) Proteomic profiles of thylakoid membranes and changes in response to iron deficiency. Photosynth Res 89:141-155

Aro EM, Suorsa M, Rokka A, Allahverdiyeva Y, Paakkarinen V, Saleem A, Battchikova N, Rintamaki E (2005) Dynamics of photosystem II: a proteomic approach to thylakoid protein complexes. J Exp Bot 56:347-356

Bafna V, Edwards N (2001) SCOPE: a probabilistic model for scoring tandem mass spectra against a peptide database. Bioinformatics 17:3-21

Bafna V, Edwards N (2003) On *de novo* Interpretation of Tandem Mass Spectra for Peptide Identification. RECOMB 2003:9-18

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

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, Koller A, Val GD, Schurmann P, Buchanan BB (2004) Proteomics uncovers proteins interacting electrostatically with thioredoxin in chloroplasts. Photosynth Res 79:275-280

Barnidge DR, Dratz EA, Martin T, Bonilla LE, Moran LB, Lindall A (2003) Absolute quantification of the G protein-coupled receptor rhodopsin by LC/MS/MS using proteolysis product peptides and synthetic peptide standards. Anal Chem 75:445-451

Barr JR, Maggio VL, Patterson DG Jr, Cooper GR, Henderson LO, Turner WE, Smith SJ, Hannon WH, Needham LL, Sampson EJ (1996) Isotope dilution--mass spectrometric quantification of specific proteins: model application with apolipoprotein A-I. Clin Chem 42:1676-1682

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

Caffarri S, Frigerio S, Olivieri E, Righetti PG, Bassi R (2005) Differential accumulation of Lhcb gene products in thylakoid membranes of *Zea mays* plants grown under contrasting light and temperature conditions. Proteomics 5:758-768

Chen T, Kao MY, Tepel M, Rush J, Church GM (2001) A dynamic programming approach to de novo peptide sequencing via tandem mass spectrometry. J Comput Biol 8:325-337

Ciambella C, Roepstorff P, Aro EM, Zolla L (2005) A proteomic approach for investigation of photosynthetic apparatus in plants. Proteomics 5:746-757

Dancik V, Addona TA, Clauser KR, Vath JE, Pevzner PA (1999) *De novo* peptide sequencing via tandem mass spectrometry. J Comput Biol 6:327-342

Domon B, Aebersold R (2006) Mass spectrometry and protein analysis. Science 312:212-217

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

Emanuelsson O, Nielsen H, von Heijne G (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci 8:978-984

Eng JK, Mccormack AL, Yates JR (1994) An approach to correlate tandem mass-spectral data of peptides with amino-acid-sequences in a protein database. J Am Soc Mass Spectrom 5:976-989

Ephritikhine G, Ferro M, Rolland N (2004) Plant membrane proteomics. Plant Physiol Biochem 42:943-962

Fernandez-de-Cossio J, Gonzalez J, Satomi Y, Shima T, Okumura N, Besada V, Betancourt L, Padron G, Shimonishi Y, Takao T (2000) Automated interpretation of low-energy collision-induced dissociation spectra by SeqMS, a software aid for *de novo* sequencing by tandem mass spectrometry. Electrophoresis 21:1694-1699

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

Ferro M, Salvi D, Riviere-Rolland H, Vermat T, Seigneurin-Berny D, Grunwald D, Garin J, Joyard J, Rolland N (2002) Integral membrane proteins of the chloroplast envelope:

identification and subcellular localization of new transporters. Proc Natl Acad Sci USA 99:11487-11492

Ferro M, Seigneurin-Berny D, Rolland N, Chapel A, Salvi D, Garin J, Joyard J (2000) Organic solvent extraction as a versatile procedure to identify hydrophobic chloroplast membrane proteins. Electrophoresis 21:3517-3526

Forster B, Mathesius U, Pogson BJ (2006) Comparative proteomics of high light stress in the model alga *Chlamydomonas reinhardtii*. Proteomics 6:4309-4320

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

Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon AM, Cruciat CM, Remor M, Hofert C, Schelder M, Brajenovic M, Ruffner H, Merino A, Klein K, Hudak M, Dickson D, Rudi T, Gnau V, Bauch A, Bastuck S, Huhse B, Leutwein C, Heurtier MA, Copley RR, Edelmann A, Querfurth E, Rybin V, Drewes G, Raida M, Bouwmeester T, Bork P, Seraphin B, Kuster B, Neubauer G, Superti-Furga G (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature 415:141-147

Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci USA 100:6940-6945

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

Giddings MC, Shah AA, Gesteland R, Moore B (2003) Genome-based peptide fingerprint scanning. Proc Natl Acad Sci USA 100:20-25

Gomez SM, Bil KY, Aguilera R, Nishio JN, Faull KF, Whitelegge JP (2003) Transit peptide cleavage sites of integral thylakoid membrane proteins. Mol Cell Proteomics 2:1068-1085

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

Gorg A, Weiss W, Dunn MJ (2004) Current two-dimensional electrophoresis technology for proteomics. Proteomics 4:3665-3685

Goulas E, Schubert M, Kieselbach T, Kleczkowski LA, Gardestrom P, Schroder 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

Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 17:994-999

Hansson M, Vener AV (2003) Identification of three previously unknown in vivo protein phosphorylation sites in thylakoid membranes of *Arabidopsis thaliana*. Mol Cell Proteomics 2:550-559

Heazlewood JL, Tonti-Filippini J, Verboom RE, Millar AH (2005) Combining experimental and predicted datasets for determination of the subcellular location of proteins in *Arabidopsis*. Plant Physiol 139:598-609

Heazlewood JL, Verboom RE, Tonti-Filippini J, Small I, Millar AH (2006) SUBA: the *Arabidopsis* Subcellular Database. Nucleic Acids Res 35:D213-D218

Heinemeyer J, Eubel H, Wehmhoner D, Jansch L, Braun HP (2004) Proteomic approach to characterize the supramolecular organization of photosystems in higher plants. Phytochemistry 65:1683-1692

Hippler M, Klein J, Fink A, Allinger T, Hoerth P (2001) Towards functional proteomics of membrane protein complexes: analysis of thylakoid membranes from *Chlamydomonas reinhardtii*. Plant J 28:595-606

Ho Y, Gruhler A, Heilbut A, Bader GD, Moore L, Adams SL, Millar A, Taylor P, Bennett K, Boutilier K, Yang L, Wolting C, Donaldson I, Schandorff S, Shewnarane J, Vo M, Taggart J, Goudreault M, Muskat B, Alfarano C, Dewar D, Lin Z, Michalickova K, Willems AR, Sassi H, Nielsen PA, Rasmussen KJ, Andersen JR, Johansen LE, Hansen LH, Jespersen H, Podtelejnikov A, Nielsen E, Crawford J, Poulsen V, Sorensen BD, Matthiesen J, Hendrickson RC, Gleeson F, Pawson T, Moran MF, Durocher D, Mann M, Hogue CW, Figeys D, Tyers M (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. Nature 415:180-183

Huber CG, Timperio AM, Zolla L (2001) Isoforms of photosystem II antenna proteins in different plant species revealed by liquid chromatography-electrospray ionization mass spectrometry. J Biol Chem 276:45755-45761

Huber CG, Walcher W, Timperio AM, Troiani S, Porceddu A, Zolla L (2004) Multidimensional proteomic analysis of photosynthetic membrane proteins by liquid extraction-ultracentrifugation-liquid chromatography-mass spectrometry. Proteomics 4:3909-3920

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

Jones AM, Bennett MH, Mansfield JW, Grant M (2006a) Analysis of the defence phosphoproteome of *Arabidopsis thaliana* using differential mass tagging. Proteomics 6:4155-4165

Jones AM, Thomas V, Bennett MH, Mansfield J, Grant M (2006b) Modifications to the *Arabidopsis* defense proteome occur prior to significant transcriptional change in response to inoculation with pseudomonas syringae. Plant Physiol 142:1603-1620

Joyard J, Teyssier E, Miege C, Berny-Seigneurin D, Marechal E, Block MA, Dorne AJ, Rolland N, Ajlani G, Douce R (1998) The biochemical machinery of plastid envelope membranes. Plant Physiol 118:715-723

Kapp EA, Schutz F, Connolly LM, Chakel JA, Meza JE, Miller CA, Fenyo D, Eng JK, Adkins JN, Omenn GS, Simpson RJ (2005) An evaluation, comparison, and accurate benchmarking of several publicly available MS/MS search algorithms: sensitivity and specificity analysis. Proteomics 5:3475-3490

Kashino Y, Lauber WM, Carroll JA, Wang Q, Whitmarsh J, Satoh K, Pakrasi HB (2002) Proteomic analysis of a highly active photosystem II preparation from the cyanobacterium *Synechocystis* sp. PCC 6803 reveals the presence of novel polypeptides. Biochemistry 41:8004-8012

Kieselbach T, Bystedt M, Hynds P, Robinson C, Schroder WP (2000) A peroxidase homologue and novel plastocyanin located by proteomics to the *Arabidopsis* chloroplast thylakoid lumen. FEBS Lett 480:271-276

Kieselbach T, Hagman, Andersson B, Schroder WP (1998) The thylakoid lumen of chloroplasts. Isolation and characterization. J Biol Chem 273:6710-6716

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

Kikuchi S, Hirohashi T, Nakai M (2006) Characterization of the preprotein translocon at the outer envelope membrane of chloroplasts by blue native PAGE. Plant Cell Physiol 47:363-371

Kleffmann T, Hirsch-Hoffmann M, Gruissem W, Baginsky S (2006) plprot: a comprehensive proteome database for different plastid types. Plant Cell Physiol 47:432-436

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

Kuchler M, Decker S, Hormann F, Soll J, Heins L (2002) Protein import into chloroplasts involves redox-regulated proteins. Embo J 21:6136-6145

Link AJ, Eng J, Schieltz DM, Carmack E, Mize GJ, Morris DR, Garvik BM, Yates JR 3rd (1999) Direct analysis of protein complexes using mass spectrometry. Nat Biotechnol 17:676-682

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

Lunn JE (2006) Compartmentation in plant metabolism. J Exp Bot 58:35-47

Ma B, Zhang K, Hendrie C, Liang C, Li M, Doherty-Kirby A, Lajoie G (2003) PEAKS: powerful software for peptide *de novo* sequencing by tandem mass spectrometry. Rapid Commun Mass Spectrom 17:2337-2342

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

Mann M, Hojrup P, Roepstorff P (1993) Use of mass spectrometric molecular weight information to identify proteins in sequence databases. Biol Mass Spectrom 22:338-345

Mann M, Wilm M (1994) Error-tolerant identification of peptides in sequence databases by peptide sequence tags. Anal Chem 66:4390-4399

Millar AH, Whelan J, Small I (2006) Recent surprises in protein targeting to mitochondria and plastids. Curr Opin Plant Biol 9:610-615

Moseley JL, Allinger T, Herzog S, Hoerth P, Wehinger E, Merchant S, Hippler M (2002) Adaptation to Fe-deficiency requires remodeling of the photosynthetic apparatus. Embo J 21:6709-6720

Naumann B, Stauber EJ, Busch A, Sommer F, Hippler M (2005) N-terminal processing of Lhca3 Is a key step in remodeling of the photosystem I-light-harvesting complex under iron deficiency in *Chlamydomonas reinhardtii*. J Biol Chem 280:20431-20441

Newton RP, Brenton AG, Smith CJ, Dudley E (2004) Plant proteome analysis by mass spectrometry: principles, problems, pitfalls and recent developments. Phytochemistry 65:1449-1485

Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10:1-6

Oda Y, Huang K, Cross FR, Cowburn D, Chait BT (1999) Accurate quantitation of protein expression and site-specific phosphorylation. Proc Natl Acad Sci USA 96:6591-6596

Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1:376-386

Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. Proc Natl Acad Sci USA 85:2444-2448

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

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, Friso G, Kalume DE, Roepstorff P, Nilsson F, Adamska I, van Wijk KJ (2000) Proteomics of the chloroplast: systematic identification and targeting analysis of lumenal and peripheral thylakoid proteins. Plant Cell 12:319-341

Peltier JB, Ripoll DR, Friso G, Rudella A, Cai Y, Ytterberg J, Giacomelli L, Pillardy J, van Wijk KJ (2004a) 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, Ytterberg AJ, Sun Q, van Wijk KJ (2004b) 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, Ytterberg J, Liberles DA, Roepstorff P, van Wijk KJ (2001) Identification of a 350-kDa ClpP protease complex with 10 different Clp isoforms in chloroplasts of *Arabidopsis thaliana*. J Biol Chem 276:16318-16327

Peng J, Gygi SP (2001) Proteomics: the move to mixtures. J Mass Spectrom 36:1083-1091

Perez-Bueno ML, Rahoutei J, Sajnani C, Garcia-Luque I, Baron M (2004) Proteomic analysis of the oxygen-evolving complex of photosystem II under biotec stress: Studies on *Nicotiana benthamiana* infected with tobamoviruses. Proteomics 4:418-425

Perkins DN, Pappin DJ, Creasy DM, Cottrell JS (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20:3551-3567

Phee BK, Cho JH, Park S, Jung JH, Lee YH, Jeon JS, Bhoo SH, Hahn TR (2004) Proteomic analysis of the response of *Arabidopsis* chloroplast proteins to high light stress. Proteomics 4:3560-3568

Radhamony RN, Theg SM (2006) Evidence for an ER to Golgi to chloroplast protein transport pathway. Trends Cell Biol 16:385-387

Rexroth S, Meyer zu Tittingdorf JM, Krause F, Dencher NA, Seelert H (2003) Thylakoid membrane at altered metabolic state: challenging the forgotten realms of the proteome. Electrophoresis 24:2814-2823

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, Ephritikhine G, Marmagne A, Ramus C, Brugiere S, Salvi D, Seigneurin-Berny D, Bourguignon J, Barbier-Brygoo H, Joyard J, Garin J (2006) A versatile method for deciphering plant membrane proteomes. J Exp Bot 57:1579-1589

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

Rudella A, Friso G, Alonso JM, Ecker JR, van Wijk KJ (2006) Downregulation of ClpR2 leads to reduced accumulation of the ClpPRS protease complex and defects in chloroplast biogenesis in *Arabidopsis*. Plant Cell 18:1704-1721

Savitski MM, Nielsen ML, Zubarev RA (2005) New data base-independent, sequence tag-based scoring of peptide MS/MS data validates Mowse scores, recovers below threshold data, singles out modified peptides, and assesses the quality of MS/MS techniques. Mol Cell Proteomics 4:1180-1188

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

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

Seigneurin-Berny D, Rolland N, Garin J, Joyard J (1999) Technical advance: differential extraction of hydrophobic proteins from chloroplast envelope membranes: a subcellular-specific proteomic approach to identify rare intrinsic membrane proteins. Plant J 19:217-228

Shadforth I, Crowther D, Bessant C (2005) Protein and peptide identification algorithms using MS for use in high-throughput, automated pipelines. Proteomics 5:4082-4095

Shevchenko A, Jensen ON, Podtelejnikov AV, Sagliocco F, Wilm M, Vorm O, Mortensen P, Boucherie H, Mann M (1996) Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. Proc Natl Acad Sci USA 93:14440-14445

Small I, Peeters N, Legeai F, Lurin C (2004) Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics 4:1581-1590

Spengler B (2004) *De novo* sequencing, peptide composition analysis, and composition-based sequencing: a new strategy employing accurate mass determination by fourier transform ion cyclotron resonance mass spectrometry. J Am Soc Mass Spectrom 15:703-714

Srivastava R, Pisareva T, Norling B (2005) Proteomic studies of the thylakoid membrane of *Synechocystis* sp. PCC 6803. Proteomics 5:4905-4916

Stauber EJ, Fink A, Markert C, Kruse O, Johanningmeier U, Hippler M (2003) Proteomics of *Chlamydomonas reinhardtii* light-harvesting proteins. Eukaryot Cell 2:978-994

Steen H, Pandey A (2002) Proteomics goes quantitative: measuring protein abundance. Trends Biotechnol 20:361-364

Storf S, Stauber EJ, Hippler M, Schmid VH (2004) Proteomic analysis of the photosystem I light-harvesting antenna in tomato (*Lycopersicon esculentum*). Biochemistry 43:9214-9224

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

Sunyaev S, Liska AJ, Golod A, Shevchenko A (2003) MultiTag: multiple error-tolerant sequence tag search for the sequence-similarity identification of proteins by mass spectrometry. Anal Chem 75:1307-1315

Tabb DL, Saraf A, Yates JR 3rd (2003) GutenTag: high-throughput sequence tagging via an empirically derived fragmentation model. Anal Chem 75:6415-6421

Takahashi Y, Yasui TA, Stauber EJ, Hippler M (2004) Comparison of the subunit compositions of the PSI-LHCI supercomplex and the LHCI in the green alga *Chlamydomonas reinhardtii*. Biochemistry 43:7816-7823

Taylor JA, Johnson RS (2001) Implementation and uses of automated *de novo* peptide sequencing by tandem mass spectrometry. Anal Chem 73:2594-2604

Timperio AM, Huber CG, Zolla L (2004) Separation and identification of the light harvesting proteins contained in grana and stroma thylakoid membrane fractions. J Chromatogr A 1040:73-81

Timperio AM, Zolla L (2005) Investigation of the lateral light-induced migration of photosystem II light-harvesting proteins by nano-high performance liquid chromatography electrospray ionization mass spectrometry. J Biol Chem 280:28858-28866

Turkina MV, Blanco-Rivero A, Vainonen JP, Vener AV, Villarejo A (2006a) CO2 limitation induces specific redox-dependent protein phosphorylation in *Chlamydomonas reinhardtii*. Proteomics 6:2693-2704

Turkina MV, Kargul J, Blanco-Rivero A, Villarejo A, Barber J, Vener AV (2006b) Environmentally modulated phosphoproteome of photosynthetic membranes in the green alga *Chlamydomonas reinhardtii*. Mol Cell Proteomics 5:1412-1425

Turkina MV, Villarejo A, Vener AV (2004) The transit peptide of CP29 thylakoid protein in *Chlamydomonas reinhardtii* is not removed but undergoes acetylation and phosphorylation. FEBS Lett 564:104-108

Unlu M, Morgan ME, Minden JS (1997) Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. Electrophoresis 18:2071-2077

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

Vener AV, Harms A, Sussman MR, Vierstra RD (2001) Mass spectrometric resolution of reversible protein phosphorylation in photosynthetic membranes of *Arabidopsis thaliana*. J Biol Chem 276:6959-6966

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

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

Whitelegge JP, Laganowsky A, Nishio J, Souda P, Zhang H, Cramer WA (2006) Sequencing covalent modifications of membrane proteins. J Exp Bot 57:1515-1522

Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, Hochstrasser DF (1999) Protein identification and analysis tools in the ExPASy server. Methods Mol Biol 112:531-552

Wilson KA, McManus MT, Gordon ME, Jordan TW (2002) The proteomics of senescence in leaves of white clover, *Trifolium repens* (L.). Proteomics 2:1114-1122

Wittmann-Liebold B, Graack HR, Pohl T (2006) Two-dimensional gel electrophoresis as tool for proteomics studies in combination with protein identification by mass spectrometry. Proteomics 6:4688-4703

Yamaguchi K, Beligni MV, Prieto S, Haynes PA, McDonald WH, Yates JR 3rd, Mayfield SP (2003) Proteomic characterization of the *Chlamydomonas reinhardtii* chloroplast ribosome. Identification of proteins unique to the 70 S ribosome. J Biol Chem 278:33774-33785

Yamaguchi K, Prieto S, Beligni MV, Haynes PA, McDonald WH, Yates JR 3rd, Mayfield SP (2002) Proteomic characterization of the small subunit of *Chlamydomonas*

reinhardtii chloroplast ribosome: identification of a novel S1 domain-containing protein and unusually large orthologs of bacterial S2, S3, and S5. Plant Cell 14:2957-2974

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

Yan B, Pan C, Olman VN, Hettich RL, Xu Y (2005) A graph-theoretic approach for the separation of b and y ions in tandem mass spectra. Bioinformatics 21:563-574

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

Zhao C, Wang J, Cao M, Zhao K, Shao J, Lei T, Yin J, Hill GG, Xu N, Liu S (2005) Proteomic changes in rice leaves during development of field-grown rice plants. Proteomics 5:961-972

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 (2000) High performance liquid chromatography-electrospray mass spectrometry for the simultaneous resolution and identification of intrinsic thylakoid membrane proteins. Proteins 41:398-406

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

Hippler, Michael
Institute of Plant Biochemistry and Biotechnology, University of Muenster, Hindenburgplatz 55, 48143 Muenster, Germany
mhippler@uni-muenster.de

Naumann, Bianca
Institute of Plant Biochemistry and Biotechnology, University of Muenster, Hindenburgplatz 55, 48143 Muenster, Germany

List of abbreviations

2-DE: two dimensional gel electrophoresis
BN-PAGE: blue-native polyacrylamide gel electrophoresis
CCM: CO_2 concentrating mechanism
CN-PAGE: colorless-native polyacrylamide gel electrophoresis
cTP: chloroplast transit peptide

$cytb_6f$: cytochrome b_6f complex
DIGE: fluorescence two dimensional Difference Gel Electrophoresis
ESI: electrospray ionization
GPF: genomic peptide finder
HPLC: high performance liquid chromatography
IB: immunoblot
ICAT: isotope-coded affinity tag
IEF: isoelectric focusing
IMAC: immobilized metal affinity chromatography
IB: immunoblot
IP: isoelectric point
LC: liquid chromatography
LHCI: light harvesting complex I
LHCII: light harvesting complex II
lTP: lumenal transit peptide
MALDI: matrix assisted laser desorption ionisation
MS/MS: tandem mass spectrometry
MS: mass spectrometry, mass spectrometer
mudPIT: multidimensional protein identification technology
PAGE: polyacrylamide gel electrophoresis
PG: plastoglobuli
PMF: peptide mass finger printing
PSI: photosystem I
PSII: photosystem II
PTM: post-translational modification
ROS: reactive oxygen species
RP: reversed phase
SCX: strong cation exchange
SDS: sodium dodecyl sulfate
SILAC: stable isotope-labeling of amino acids in cell culture
TMD: transmembrane domain
TOF: time of flight
TPP: three phase partitioning
WT: wild type