The Oligomeric Stromal Proteome of Arabidopsis thaliana Chloroplasts

Jean-Benoit Peltier‡§, Yang Cai‡¶, Qi Sun||, Vladimir Zabrouskov‡**, Lisa Giacomelli‡, Andrea Rudella‡, A. Jimmy Ytterberg‡, Heidi Rutschow‡, and Klaas J. van Wijk‡‡

This study presents an analysis of the stromal proteome in its oligomeric state extracted from highly purified chloroplasts of Arabidopsis thaliana. 241 proteins (88% with predicted cTP), mostly assembled in oligomeric complexes, were identified by mass spectrometry with emphasis on distinguishing between paralogues. This is critical because different paralogues in a gene family often have different subcellular localizations and/or different expression patterns and functions. The native protein masses were measured for all identified proteins. Comparison with the few well characterized stromatic complexes from A. thaliana confirmed the accuracy of the native mass determination, and by extension, the usefulness of the native mass data for future in-depth protein interaction studies. Resolved protein interactions are discussed and compared with an extensive collection of native mass data of orthologues in other plants and bacteria. Relative protein expression levels were estimated from spot intensities and also provided estimates of relative concentrations of individual proteins. No such quantification has been reported so far. Surprisingly proteins dedicated to chloroplast protein synthesis, biogenesis, and fate represented nearly 10% of the total stroma protein mass. Oxidative pentose phosphate pathway, glycolysis, and Calvin cycle represented together about 75%, nitrogen assimilation represented 5–7%, and all other pathways such as biosynthesis of e.g. fatty acids, amino acids, nucleotides, tetrapyrroles, and vitamins B₁ and B₂ each represented less than 1% of total protein mass. Several proteins with diverse functions outside primary carbon metabolism, such as the isomerase ROC4, lipoxygenase 2 involved in jasmonic acid biosynthesis, and a carbonic anhydrase (CA1), were surprisingly abundant in the range of 0.75–1.5% of the total stromal mass. Native images with associated information are available via the Plastid Proteome Database. Molecular & Cellular Proteomics 5:114–133, 2006.

Although best known for their role in photosynthesis, chloroplasts (and plastids in general) synthesize many essential compounds such as plant hormones, fatty acids and lipids, amino acids, vitamins, purine and pyrimidine nucleotides, tetrapyrroles, and isoprenoids (1). Chloroplasts are required for nitrogen and sulfur assimilation and contain numerous protein chaperones and assembly factors, peptides, and proteases (1). To facilitate chloroplast gene expression, chloroplasts contain proteins associated with plastid DNA and the plastid transcriptional and translation machinery including many mRNA-binding proteins involved in mRNA processing, stability, and translation (2, 3). Predictions of the plastid proteome by the subcellular localization predictors TargetP (4) and Predotar (5) followed by correction with each reported sensitivity (0.85 and 0.82, respectively) and specificity (0.69 and 0.88, respectively) suggested that all non-green plastid types and chloroplasts together contain between 1707 (for Predotar) and 3454 (for TargetP) proteins (6).

In recent years, proteomic studies together with many detailed “one-protein-at-a-time” studies are collectively beginning to provide a good insight into the chloroplast proteome. The thylakoid and envelope proteomes of chloroplasts from Arabidopsis thaliana have been analyzed in a number of studies, which were reviewed recently (6–9). No in-depth analysis of the highly purified stromal proteome of A. thaliana has been carried out to date but is urgently needed to complete the overview of the chloroplast proteome from A. thaliana rosette leaves. Manual data mining of the A. thaliana chloroplast literature, such as is done at The Arabidopsis Information Resource (TAIR) (1) (www.arabidopsis.org/) and for the Plastid Proteome Database (PPDB; ppdb.tc.cornell.edu/) (10), will further help to provide an

1 The abbreviations used are: TAIR, The Arabidopsis Information Resource; PPDB, Plastid Proteome Database; OPPP, oxidative pentose phosphate pathway; ACCase, acetyl-CoA carboxylase; CN, colorless native; BN, blue native; LOX, lipoxygenase; EF, elongation factor; GS2, glutamine synthase 2; SFB, sedoheptulose/fructose-1,6-biphosphate aldolase; TPI, triose-phosphate isomerase; ROC4, rotamase CyP; CPN, chaperonin; CA, carbonic anhydrase; Rubisco, ribulose-1,5-bisphosphate carboxylase/oxygenase; RBCS, ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit; RBCL, ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; RCA, ribulose-1,5-bisphosphate carboxylase/oxygenase activase; Fd, ferredoxin; Fd-GOGAT, ferredoxin-dependent glutamate synthase; Tricine, N-[2-hydroxy-1,1-bis(hydroxymethyl)methyl]glycine; MDH, malate dehydrogenase; NiR, nitrite reductase; ACP, acyl carrier protein; DMR1, 6,7-dimethyl-8-ribitylumazine; PNPase, polynucleotide phosphorylase; RH, RNA helicase; TAP, tandem affinity purification; RAP, ribosome-associated protein; CTP, chloroplast transit peptide.
accurate overview of the chloroplast proteome.

Several chloroplast-localized biosynthetic pathways are linked to each other with intermediates from one pathway being used in other pathways. In some cases, biosynthetic pathways branch into two different pathways such as heme and chlorophyll biosynthesis at the level of protoporphyrin IX (11). In other cases, several enzymes are shared by different pathways, such as enzymes in the Calvin cycle and oxidative pentose phosphate pathway (OPPP) (12) or in the Calvin cycle and glycolysis (13). It has been demonstrated for several enzymes that specific protein isoforms or functional paralogues specialize in different functions or pathways often located in different subcellular localizations (e.g. cytosol versus chloroplast) or tissues (e.g. root versus shoot) (see e.g. Ref. 14). Thus to understand the regulation of metabolic activity it is important to distinguish between such functional paralogues. Tandem mass spectrometry with high mass accuracy will typically allow distinguishing between such closely related paralogues.

Assembly and disassembly of multisubunit complexes as well as suborganellar localization of enzymes (e.g. thylakoid membrane versus envelope membrane) has been shown to influence flux of different pathways. In some cases, this can lead to so-called “metabolic channeling” (15). Protein-protein interactions can also help to stabilize proteins and to protect against denaturation and proteolysis. These interactions can be homomeric (between identical proteins) or heteromeric (between different proteins). Identifying these protein interactions is needed to truly understand plastid protein functions and plastid metabolic pathways. Given the complexity of the stromal proteome, only a small number of stromal protein complexes in A. thaliana have been characterized. Examples are the heteromeric Clp protease complex of 325–350 kDa with 11 different proteins (16), the 200–240 kDa heterotetramer of ADP-glucose pyrophosphorylase (or glucose-1-phosphate adenylyltransferase) (17), the stromal signal recognition particle (18), and the ~150-kDa heterotetrameric tryptophan synthase (19). Extensive searches of the published literature (this study) did identify a significant number of plastid complexes from a large variety of other plant species than Arabidopsis (spinach, pea, Brassica rapa, potato, barley, etc.), such as the 30 and 50 S ribosomal subunits from spinach chloroplasts (20, 21), plastid pyruvate dehydrogenase (22), a plastid hetero-oligomer of acetyl-CoA carboxylase (ACCase) of 600–800 kDa in soybean (23), and homo-olameric porphobilinogen synthase from pea plastids (24). Currently there is no centralized data deposit of these protein-protein interactions in A. thaliana or other plant species.

Another important aspect of understanding chloroplast function is to determine protein expression levels and molar ratios between different chloroplast proteins. Currently there is little knowledge of the relative accumulation levels of stromal proteins even for the best studied chloroplast pathways. Quantification of molar ratios between proteins in complex proteomes is generally difficult and has not been attempted at any large scale. As will be demonstrated here, we found that native gels, such as colorless native (CN)-PAGE (or blue native (BN)-PAGE) followed by SDS-PAGE, currently provide the most convenient semiquantitative comparison of different protein species.

In this study, we set out to (i) experimentally identify the Arabidopsis stromal proteome with emphasis on distinguishing between paralogues, (ii) determine the approximate and relative accumulation levels of stromal proteins, (iii) determine the native masses of stromal proteins and, where possible, resolve protein interactions, (iv) collect information on plastid protein-protein interactions from A. thaliana or other plant species, and (v) expand the PPDB as a plastid proteome resource for the plant community.

**EXPERIMENTAL PROCEDURES**

**Plant Growth, Protein Preparations, and Protein Separation—** A. thaliana (Col 0) was grown under 10-h light/14-h dark cycles at 25/17 °C. Rosette leaves were collected about 55 days after sowing. These plants were still in their vegetative stage, about 1 week prior to bolting. Intact chloroplasts were isolated, and the native stromal proteome was collected, concentrated, and directly loaded on colorless native gels (CN-PAGE) as described in Ref. 16. The gels were then loaded on Tricine-SDS-PAGE gels (linear gradients with 8–15% acrylamide), focused, and stained with Coomassie Brilliant Blue, silver nitrate, or the fluorescent dye SYPRO Ruby (16).

**Image Analysis and Quantification—** Coomassie-stained gels were scanned with a desktop scanner or high-resolution scanner (Amer sham Biosciences), and SYPRO Ruby-stained gels were scanned using a charge-coupled device camera (FluorS, Bio-Rad). Spot volumes were determined, corrected for background, and normalized to total spot volume using Phoretix software version 2004 (Non-linear, Newcastle, UK). We always verified the correlation between predicted processed molecular mass and experimentally observed mass. In a limited number of cases, more than one protein was identified per gel spot. We first verified whether the identities in a spot could be explained by any background signals (from streaking) from abundant proteins in this gel area. Such identities from background were not quantified. If it appeared that the identities of the proteins in a spot were truly the results of co-migrating proteins and if the MOWSE scores were in a similar range (within 3-fold difference), then each protein was assigned half the spot volume. If spots contained more than one protein with very different MOWSE scores (at least a factor of 3), we removed the protein identification based on the lowest MOWSE score. This was a reasonable strategy because co-migrating proteins were typically of similar molecular mass. In case a protein was identified more than once, we summed all corrected spot volumes for that accession. To facilitate comparison of abundance of different proteins, the spot volume(s) for each accession was divided by the denatured molecular mass for each accession. This resulted in a rough approximation of relative concentration. We do point out that protein abundance was calculated from SYPRO Ruby-stained spot intensities. Because SYPRO Ruby binds preferentially to charged residues (lysine, arginine, and histidine) protein abundance is underestimated or overestimated if the proteins contain few or many charged residues, respectively.

**Protein Identification by Mass Spectrometry, (Un)ambiguous Identification, and Bioinformatics—** Stained protein spots were excised, washed, and digested with modified trypsin, and peptides were extracted automatically (Progest, Genomic Solutions, Ann Arbor, MI). Proteins were identified by peptide mass fingerprinting using a MALDI-TOF mass spectrometer (Voyager DE-STR, Applied Biosystems) and/or by tandem mass spectrometry using a capillary LC-ESI
MS/MS (Q-TOF, Waters) as described in Ref. 25. MS or MS/MS spectra were used to search the predicted A. thaliana proteome downloaded from the TAIR database using an in-house installation of Mascot (www.matrixscience.com). Criteria for positive identification from peptide mass fingerprinting and from MS/MS data are described in Ref. 25. In the analysis of the MS data, an effort was made to distinguish between members of the same gene family or otherwise related gene products. In some cases, peptides were identified that match to more than one protein. Uniquely matching peptides (diagnostic peptides) are then needed to determine which protein is expressed. In case the mass spectrometry data did match ambiguously to more than one protein, these protein identifications are automatically linked within the PPDB database and reported in the tables.

**Plastid Proteome Database and New Interface**—The construction of the PPDB (ppdb.tc.cornell.edu) was originally described in Ref. 25. The PPDB interface was improved, and search functions were expanded since its inception in 2004. Also more detailed curated information can now be accessed. The annotated CN-PAGE gel image and associated experimental, predicted, and other data presented in this study can be accessed via PPDB. PPDB also contains the theoretical analysis of all Arabidopsis entries (currently release 5.0 of ATH1.pep with 29,161 nuclear encoded Arabidopsis proteins as well as the mitochondrial and plastid genomes). Mascot scores, number of matching peptides, and highest peptide score for each identification as well as functional classification are listed.

**Miscellaneous**—Chlorophyll concentrations were determined spectrometrically in 80% acetone (26), and protein determinations were done with the bicinchoninic acid assay (27).

### RESULTS

**Stromal Proteome Identification and Classification**

Purified intact chloroplasts were lysed under non-denaturing conditions, and chloroplast stromal proteins and protein complexes were separated based on native mass using CN-PAGE (28) followed by complete denaturation and separation by SDS-PAGE. Proteins were visualized by Coomassie staining of preparative gels for mass spectrometry analysis (Fig. 1, A and B) or by SYPRO Ruby staining of analytical gels for quantification (Fig. 1, C and D). The smaller figures (Fig. 1, B and D) are from independent preparations and show that these CN-PAGE gel patterns are reproducible, supporting the notion that CN-PAGE gels are an excellent tool for proteome analysis. Proteins were identified by MALDI-TOF MS peptide mass fingerprinting and/or nano-LC-ESI-MS/MS followed by Mascot search against the TAIR database. In the analysis of the MS data, an effort was made to distinguish between members of the same gene family (paralogues). These search results were automatically filtered using an in-house software routine\(^2\) followed by manual verification and additional quality control steps as detailed under “Experimental Procedures.”

241 non-redundant proteins were identified on the CN-PAGE gels (see the supplemental table and PPDB for interactive searches). The identified proteins were classified according to the hierarchical, non-redundant classification system developed for MapMan (29) (gabi.rzpd.de/projects/MapMan) adjusted after in-house manual verification and information from the literature. The MapMan system has 35 main functional categories orBins with a larger number of Subbins (subcategories) (see also ppdb.tc.cornell.edu/mapman.aspx). To our surprise, proteins involved in folding, proteolysis, and sorting (Bins 29.3–29.9) represented the largest functional category in the stroma (14%) closely followed by proteins related to protein synthesis (Bins 29.1 and 29.2) (12%) (Table I). 21% of the identified proteins were involved in secondary metabolism, covering amino acid metabolism (7%), nucleotide synthesis and degradation (4%), tetrapyrrole synthesis (6%), and enzymes involved in synthesis of vitamins B\(_1\) and B\(_2\), isoprenes, jasmonic acid, and lipids/fatty acids (4%). As expected, proteins involved in primary carbon metabolism, such as the Calvin cycle, OPPP, and glycolysis represented a population of significant number (12%) and abundance representing ~76% of the total stroma mass as determined from the relative spot intensities. Enzymes involved in starch synthesis and degradation (Bin 2) were also well represented. The function of 11% of the identified proteins was unknown (see the supplemental table and PPDB).

**The Identified Stromal Proteome Is Pure with High TargetP Prediction Rates**

Cross-correlation of the identified proteins against other plant proteomic studies on A. thaliana plasma membranes (30, 31), vacuole and tonoplast (32, 33), the peroxisome (34), the nucleus (35), the cell wall (36), and the hydrophobic mitochondrial membranes (37) and a dozen other mitochondrial proteome analyses (from www.mitoz.bcs.uwa.edu.au/; see Ref. 38) did not suggest obvious contaminants from non-chloroplast locations. Further analysis suggested two to four potential contaminants from the cytosol and mitochondria as indicated in the supplemental table. In agreement with the low number of contaminants, 88% of the identified nuclear encoded proteins were predicted (TargetP, www.cbs.dtu.dk/services/TargetP/) to be plastid-localized. This is slightly higher than the reported 85% sensitivity (or true positive identification rate) for TargetP (4) and indicative of the high quality of this stromal data set. TargetP prediction for each protein is listed in the supplemental table. Seven proteins were chloroplast-encoded.

**Relation to Other Chloroplast Proteomic Studies**

About 40% (over 100 proteins) were not observed in earlier A. thaliana chloroplast proteomic studies on the thylakoid and envelope membrane and their associated (stromal) proteins (25, 39–46) (for cross-correlation, see the supplemental table and PPDB). Predicted functions of these newly identified proteins include biosynthesis of amino acids, nucleotides, and proteins as well as secondary metabolism (e.g. tetrapyrrole, thiamine, isoprenoids, and hormones), and numerous proteins were without any predicted function. A significant number of proteins were listed in the chloro-
plast analysis described in Ref. 47. However, this dataset is problematic because it appears to contain a large percentage (>40%) of non-chloroplast proteins as also reflected in the low percentage of chloroplast predicted proteins. This is either a consequence of experimental contaminants and/or the result of high rates of “false identifications” during the search with mass spectrometry data (for discussion see Ref. 6).

**Relative Abundance of the Identified Proteome**

An important aspect of understanding plastid function is to determine the expression level and molar ratios between different plastid proteins in addition to their identification. Currently there is very little knowledge of the relative accumulation levels of the proteins in the chloroplast stroma. To this end, we quantified all 287 protein spots from analytical SYPRO Ruby-stained CN-PAGE gels of two independent chloroplast purifications and normalized those to total spot intensity (spot volume) of each gel. Proteins were identified in 251 spots, representing ~99% of the total spot intensity (spot volume). As is apparent from the gel images in Fig. 1, just a handful of protein spots represent a large percentage of the total protein biomass of the stroma. The 23 most abundant stromal proteins (based on normalized spot volume(s)) or based on “relative normalized concentration” (calculated from...
normalized spot volume divided by the experimental molecular mass of the protein) are listed in Table II. Together they represent ~85% of the total stromal protein mass, and their expression covers a dynamic range of 2 orders of magnitude. These proteins were also prominent from independent “shotgun” nano-LC-ESI-MS/MS analysis of the stroma.3 We will discuss these relative abundances further below, in connection with protein interactions and protein function. We do point out, however, that protein abundance was calculated from SYPRO Ruby-stained spot intensity. Because SYPRO Ruby binds preferentially to charged residues (lysine, arginine, and histidine) protein abundance is underestimated or overestimated if the proteins contain few or many charged residues, respectively.

Native Mass, Oligomeric State, and Validation

As a first step in understanding the protein-protein interaction network of the chloroplast stromal proteome, we determined the native mass of the identified proteins on the CN-PAGE gels (Fig. 1 and supplemental table). The gels were calibrated with commercial native standards. The CN-PAGE gels resolved proteins in a molecular mass range from ~20 to ~950 kDa with ribosomes and other large protein complexes accumulating in the stacking gel in the first (native) dimension.

To compare these native mass data with existing data, we extensively searched the literature for mass information on closely related and more distant orthologues in and outside of plastids. Mass information from various species was found for 140 proteins of the 241 identified on our CN-PAGE gels, together representing 82 different monomeric proteins or protein complexes (supplemental table). For ~70% of those 140 identified proteins the native mass deduced from the CN-PAGE gels is in approximate agreement with data from the literature. The native mass and estimated oligomeric state and references for selected proteins are listed in Table III. The complete dataset is available in the supplemental table. Thus our comprehensive experimental native data provide a resource and starting point when searching for potential stromal protein partners. One of the biggest surprises when searching the literature was that about 60% of the proteins were identified as homo-oligomeric complexes or monomers rather than heteromeric complexes. The potential significance will be discussed.

Integration of Native Masses, Relative Expression Levels, and Functions

In the remaining sections, we will highlight novelties and new insights concerning specific proteins and protein complexes in terms of their relative abundance, native state, and functions. To facilitate comparison of relative abundances of different proteins, the spot volume(s) for each accession was divided by its denatured molecular mass, resulting in a relative concentration. Given the inaccuracies and pitfalls of quantification from spot intensities, we simplified these relative concentrations to a scale of 1–5 with each level representing 1 order of magnitude (supplemental table). This provides an immediate impression of the accumulation levels of the different functional categories as a group and for individual proteins within a functional class.

---

**Table I**

| MapMan Bins | Functional classification | Number identified | Number (percentage of total) | Relative amount of protein mass | % |
|-------------|--------------------------|-------------------|-------------------------------|--------------------------------|----|
| 1.1         | Thylakoid                | 9                 | 3.7                           | 0.8                            |    |
| 1.2, 1.3, 1.4, 1.7 | Calvin cycle, OPPP, glycolysis | 30               | 12.4                          | 75.8                           |    |
| 2, 3        | Minor and major CHO      | 12                | 5.0                           | 0.4                            |    |
| 8           | Organic transformation   | 6                 | 2.5                           | 1.4                            |    |
| 12, 14      | Nitrogen, sulfur assimilation | 4                | 1.7                           | 7.6                            |    |
| 13          | Amino acid synthesis and degradation | 17            | 7.1                           | 0.8                            |    |
| 11, 16, 17, 18 | Lipid, hormone, isoprene, vitamin, and cofactor synthesis | 10            | 4.1                           | 1.1                            |    |
| 19          | Tetrapyrroles            | 14                | 5.8                           | 0.5                            |    |
| 20, 21      | Redox and stress         | 20                | 8.3                           | 1.9                            |    |
| 23          | Nucleotide synthesis and degradation | 10            | 4.1                           | 0.5                            |    |
| 27          | RNA                      | 7                 | 2.9                           | 0.6                            |    |
| 29.1 and 29.2 | Protein synthesis       | 29                | 12.0                          | 1.4                            |    |
| 29.3–29.9   | Protein targeting, folding, and degradation | 34            | 14.1                          | 5.2                            |    |
| 24, 26, 28, 35 | Unknown and miscellaneous | 39            | 16.2                          | 1.7                            |    |

---

3 G. Friso, A. Rudella, H. Rutschow, and K. J. van Wijk, unpublished data.
We experimentally identified 27 proteins associated with the Calvin cycle, glycolysis, OPPP, and photorespiration, representing about 76% of the total stromal protein mass (Table I).

The relative concentrations of these 27 proteins ranged from level 1 to 5 with the small and large subunits of Rubisco (RBCS and RBCL, respectively) at level 1, the other enzymes of the Calvin cycle mostly at levels 2 and 3, and specific enzymes of glycolysis and OPP at levels 3–5. We also identified 2-phosphoglycolate phosphatases 1 and 2 involved in photorespiration at level 3 as a dimer. The Rubisco complex needs to be activated by the reversible carbamylation of a lysine residue in RBLC (Lys-201) followed by rapid binding of magnesium. This process is regulated by Rubisco activase (RCA) (48). The relative concentration of RCA (At1g73110)
| Accession   | Laboratory annotation | Oligomeric state | Oligomeric state reference | Organism      | Molecular mass | Relative normalized concentration (×100) | MapMan Bin |
|-------------|------------------------|------------------|----------------------------|---------------|----------------|----------------------------------------|------------|
| At5g46290.1 | Ketoacyl-ACP synthase I (Lipid Gene Database; Beisson) | Dimer            | 124                        | E. coli       | 50 64 129       | 0.00                                   | 11.1.03 lipid metabolism, FA synthesis and FA elongation, ketoacyl-ACP synthase (Beisson) |
| At1g29900.1 | Carbamoylphosphate synthase | Heterodimer      | 63, 125                    | A. aeolicus, E. coli | 130 123 174    | 0.19                                   | 13.1.2.3 amino acid metabolism, synthesis, glutamate family, arginine; 23.1.1 nucleotide metabolism, synthesis, pyrimidine synthesis |
| At3g27740.1 | Carbamoyl phosphate synthetase small subunit | Heterodimer      | 63, 125                    | A. aeolicus, E. coli | 47 101 163     | 1.24                                   | 13.1.2.3 amino acid metabolism, synthesis, glutamate family, arginine; 23.1.1 nucleotide metabolism, synthesis, pyrimidine synthesis |
| At3g22890.1 | ATP sulfurylase (ATPS1) | Tetramer         | 62                         | Spinach       | 51 52 148       | 0.09                                   | 14.01 sulfur assimilation, APS |
| At2g02500.1 | 4-Diphosphocytidyl-2-C-methyl-erythritol synthase (SPD), mevalonate-independent | Likely a dimer | This study A. thaliana | 34            |                | 0.05                                   | 16.1 secondary metabolism, isoprenoids |
| At3g45140.1 | Lipoxigenase AtLOX2, plastid | Monomer         | This study A. thaliana     | 102 34 120    |                | 7.53                                   | 17.7.1 hormone metabolism, jasmonate, synthesis-degradation |
| At2g44050.1 | DMRL synthase (COST1) (vitamin B2 synthesis) | Homohexamer      | 68–70                      | E. coli and Bacillus, spinach | 24 17 738     | 3.59                                   | 18 cofactor and vitamin metabolism |
| At1g03475.1 | Coproporphyrinogen III oxidase | Homodimer        | 71                         | A. thaliana   | 41 35 82        | 2.80                                   | 19 tetrapyrrole synthesis |
| At1g48520.1 | Glu-tRNA(Gln) amidotransferase subunit B (GATB or GLU-ADT subunit B) | Homodimer        | 72                         | C. reinhardtii | 61 199         | 0.03                                   | 19 tetrapyrrole synthesis |
| At1g69740.1 | Porphobilinogen synthase-1 (3-aminolevulinic acid dehydratase-1) (ALAD-1) | Homo-octamer     | 24                         | Pea           | 47 41 343       | 0.16                                   | 19 tetrapyrrole synthesis |
| At2g40490.1 | Uroporphyrinogen decarboxylase (UDP) | Homodimer        | 126                        | Tobacco       | 44 40 126       | 0.30                                   | 19 tetrapyrrole synthesis |
| At3g25660.1 | Glutamyl-tRNA amidotransferase subunit A | Homodimer        | 72, 127                    | C. reinhardtii, wheat | 57 43 191     | 0.05                                   | 19 tetrapyrrole synthesis |
| At5g24300.1 | Starch synthase | Homo- and heterodimer | 54, 55                    | Potato, maize | 72 74 157       | 0.62                                   | 2.1.2.02 major CHO metabolism, synthesis, starch, starch synthesis |
| At4g09020.1 | Glycoside hydrolase family 13, similar to isoamylase | Homo- and heterodimer | 54, 55                    | Potato, maize | 109 96 234      | 0.68                                   | 2.1.2.04 major CHO metabolism, synthesis, starch, debranching, starch phosphorylase |
| At3g29320.1 | Starch phosphorylase-1 | Homo- and heterodimer | 54, 55                    | Potato, maize | 109 96 234      | 0.68                                   | 2.2.2.02 major CHO metabolism, degradation, starch, starch phosphorylase |
| A1g103680.1 | Thioredoxin m1 (Trxm1) | Interacts with different targets | 129                       | A. thaliana   | 20 14 96, 89, 48 | 0.10                                   | 21.1 redox, thioredoxin |
| A1g50320.1  | Thioredoxin X (TrxX) | Interacts with different targets | 130                       | A. thaliana   | 20 13 154       | 0.04                                   | 21.1 redox, thioredoxin |
| A1g52730.1 | Thioredoxin F1 (TrxF1) | Interacts with different targets | 129                       | A. thaliana   | 19 13 15, 212, 170, 183 | 0.01                                   | 21.1 redox, thioredoxin |
| A1g20370.1 | Glutaredoxin protein family | Unknown          | None                       | x             | 19 13 34        | 0.27                                   | 21.4 redox, glutaredoxins |
| A1g311630.1 | 2-Cys peroxiredoxin A (PrxA) | Monomer, dimer, or decamer | 77, 78                    | Pea, barley     | 29 21 360,219,119, 110, 58 | 29.59                                   | 21.5 redox, periredoxins |
| Accession | Laboratory annotation | Oligomeric state | Oligomeric state reference | Organism | Molecular mass | Relative normalized concentration (x100) | MapMan Bin |
|-----------|-----------------------|-----------------|----------------------------|----------|----------------|------------------------------------------|------------|
| At3g52960.1 | Peroxiredoxin II E (PxE) | Probably dimer | This study | A. thaliana | 25 20 84 | 0.24 | 21.5 redox.periredoxins |
| At1g74260.1 | Phosphoribosylformylglycinamidine synthase | Unknown | None | x | 152 142, 39, 114, 93 | 0.04 | 23.1.2 nucleotide metabolism.synthesis.purine |
| At2g35040.1 | Phosphoribosylaminimidazolecarboxamide formyltransferase | Unknown | None | x | 65 60 222, 213, 38, 187 | 0.09 | 23.1.2 nucleotide metabolism.synthesis.purine |
| At5g48960.1 | Similar to 5'-nucleotidase | Unkown | None | x | 73 67 166, 65 | 0.01 | 23.2 nucleotide metabolism.degradation |
| At5g63310.1 | NDPK2, stromal | Hexamer | 81 | Spinach | 26 19 133 | 1.02 | 23.4 nucleotide metabolism.phosphotransfer and pyrophosphatases |
| At1g35420.1 | Dielactate hydrase family | Monomer | 96 | Pseudomonas sp. B13 | 34 30 28 | 0.22 | 26.1 misc.misc2 |
| At3g03710.1 | 3'-5' exoribonuclease | Homotrimer? | 89, 90 | Spinach | 100 94 410 | 0.55 | 27.1 RNA.processing |
| At5g26742.1 | DEAD box RNA helicase (RH3), mRNA etiolated seedling | Unknown | None | x | 81 45 980 | 0.03 | 27.1 RNA.processing |
| At1g77060.1 | Isocitrate lyase and CoA biosynthesis domain | Unkown | None | x | 37 33 119, 147 | 0.01 | 27.3.99 RNA.regulation of transcription.unclassified |
| At3g63140.1 | Ribosome-associated protein (RAP1) | Inter acts with 70S ribosome | 84 | C. reinhardtii | 44 78 107, 126, 946, 224 | 0.58 | 27.4 RNA.RNA binding |
| At1g75350.1 | 50S ribosomal protein L31 chloroplast | Hetero-59-mer | 20 | Spinach | 16 13 212 | 0.01 | 29.2.1 protein.synthesis.chloroplast-plastid |
| At5g16520.1 | Vpp1, mutant HCF155, PsPα-like | Homohexameter (60-mer)? | 131 | A. thaliana | 36 30 932 | 0.68 | 29.3 protein.targeting |
| At2g20890.1 | THF1, thylakoid formation 1 | Unknown | None | x | 34 27 270 | 0.00 | 29.3 protein.targeting |
| At5g55220.1 | Trigger factor | Dimer | 95 | E. coli | 82 60 146, 163 | 0.07 | 29.3 protein.targeting |
| At5g66520.1 | Zinc oligopeptidase A (M3 family) | Unkown | None | x | 89 80 107 | 0.07 | 29.5.07 protein.degradation.metalloprotease |
| At1g36390.1 | GrpE-2 | Dimer/tetramer and in complex with HSP70 | 93 | C. reinhardtii | 31 24 154, 197 | 1.95 | 29.6 protein.folding |
| At2g28000.1 | CPN60-α-1 | Heterotetradecamer and HMM complex with CPN21 when ADP or ATP added | 132 | Narcissus pseudonarcissus | 62 57 803 | 2.88 | 29.6 protein.folding |
| At2g44650.1 | CPN10-1 | Heteromultimer and multiple forms | 133 | A. thaliana | 15 11 170 | 1.50 | 29.6 protein.folding |
| At3g13470.1: At1g55490.1: At5g56500.1 | CPN60-1,2,3 | Heterododecamer and HMM complex with CPN21 when ADP or ATP added | 132 | N. pseudonarcissus | 63 46 807 | 2.24 | 29.6 protein.folding |
| At5g62030.1 | Peptidylprolyl isomerase ROC4 | Unkown | None | x | 28 36 107 | 72.78 | 29.6 protein.folding |
| At4g24280.1 | cpHSP70-1 (DnaK homologue) | Heterotrimer | 93 | C. reinhardtii | 77 68 123 | 7.26 | 29.6 protein.folding |
| At5g17710.1 | GrpE-1 | Dimer/tetramer and in complex with HSP70 | 93 | C. reinhardtii | 35 21 154, 202 | 1.27 | 29.6 protein.folding |
| At5g20720.1 | CPN21 (also CPN20) | Multiple form and complexed with CPN60 | 133, 134 | A. thaliana, spinach | 27 66 171 | 27.30 | 29.6 protein.folding |
| At5g49910.1 | cpHSP70-2 (DnaK homologue) | Heterotrimer | 93 | C. reinhardtii | 77 99 215 | 0.32 | 29.6 protein.folding |
| At5g50920.1 | ClpC1 | Dimer | 16 | A. thaliana | 103 59 199 | 2.50 | 29.6 protein.folding |
| At5g66530.1 | Aldose 1-epimerase | Dimer | 56 | Aspergillus niger | 34 30 61 | 0.31 | 3.5 minor CHO metabolism.others |
**TABLE III—continued**

| Accession<sup>a</sup> | Laboratory annotation<sup>b</sup> | Oligomeric state<sup>c</sup> | Oligomeric state reference<sup>d</sup> | Organism<sup>e</sup> | Molecular mass (kDa) | Relative concentration (<100) | MapMan Bin<sup>f</sup> |
|-----------------------|----------------------------------|-----------------------------|-------------------------------------|-----------------|---------------------|--------------------------|------------------|
| At1g11430.1           | DAG protein-Related              | Unknown                     | None                                | x               | 26                  | 0.06                    | 33.99 development.unspecified |
| At1g21440.1           | Carboxyvinyl-carboxyphosphonate phosphorimutase | Unknown                     | None                                | x               | 36                  | 0.00                    | 35.1 not assigned.no ontology |
| At1g09340.1           | Ribosome-associated protein (RAP38) | Hetero-59-mer              | C. reinhardtii                      | 84              | 43                  | 0.69                    | 35.2 not assigned.unknown |
| At3g47520.1           | Malate dehydrogenase (NAD), plastid | Dimer                      | Sorghum                             | 57              | 42                  | 1.31                    | 8.2.09 TCA/org. transformation.other organic acid transformations.cyt MDH |
| At5g58330.1           | Malate dehydrogenase (NADP)      | Dimer                       | Sorghum                             | 57              | 48                  | 3.52                    | 8.2.09 TCA/org. transformation.other organic acid transformations.cyt MDH |

<sup>a</sup> Accession number from TAIR. Note that only gene model (typically .1) is indicated. However, in most cases the mass spectrometry data also matched the other gene models; such details can be found in PPDB.

<sup>b</sup> Internal annotation in PPDB.

<sup>c</sup> Oligomeric state as reported in the literature.

<sup>d</sup> Oligomeric state in *A. thaliana* or other species.

<sup>e</sup> Organism used in the cited reference(s).

<sup>f</sup> Calculated molecular mass of the precursor proteins as reported in PPDB.

<sup>g</sup> Calculated molecular mass of the processed proteins as reported in PPDB.

<sup>h</sup> Experimental native mass as determined from the CN-PAGE gels. If proteins were identified in more than one spot (e.g. due to aggregation, smearing, or truly different oligomeric states), these additional native masses are also listed.

<sup>i</sup> Relative concentration (multiplied by 100) in the stroma as determined by normalize spot volume divided by experimental mass. If more than one protein was identified in a spot, the spot volume was divided by the number of its protein components.

<sup>j</sup> Functional assignment according to the MapMan Bin classification (29) and reported in PPDB.
was about 70 times less than RBCL and RBCS. Rubisco N-methyltransferase catalyzing N-methylation of RBCL (needed for an active enzyme) was ~14,000 times less abundant than Rubisco.

Many of the enzymes in these three pathways form hetero- or homo-oligomeric complexes. The 550-kDa Rubisco complex is the most abundant and well studied example and forms a hetero-oligomer with eight small and eight large subunits. RCA was found in different native mass complexes (330, 480, 600, and 770 kDa and as a >950-kDa complex in the stacking gel) and is known to transiently associate with the Rubisco complex (for a review, see Ref. 49). Sedoheptulose/fructose-bisphosphate aldolase 1 and 2 (SFBA-1 or -2) (At2g21330 and At4g38970) were both found at 178 kDa, suggesting a homo- or heterotetrameric state. These aldolases are reported to form stromal homotetramers, whereas aldolases outside the chloroplast form dimers (50, 51).

Minor and Major Carbohydrate Metabolism (Bins 2 and 3)—Starch and minor carbohydrate metabolism contribute to the carbohydrate storage and sugar diversity in plastids. We identified 12 proteins in this functional category with relative expression levels ranging from levels 3 to 5. These included two ADP-dependent pfkB carbohydrate kinases and Sexl involved in starch phosphorylation control (52, 53). Glucan phosphorylase-1 (At3g29320) (~110 kDa), involved in phosphorolytic starch breakdown, forms a homodimer or heterodimer (54, 55). We identified it on the stroma native gels at 233 kDa, corresponding to a dimer. α isoamylase, a starch debranching enzyme (At4g09020) of the hydrolytic or amylo-lytic pathway, was found in the stroma as a monomer. Aldose 1-epimerase (At5g66530) catalyzes the interconversion of the α- and β-anomers of hexose sugars such as glucose and galactose. Aldose 1-epimerase (35 kDa) acts as a dimer in Aspergillus niger (56) and was indeed found on the CN-PAGE gel at 61 kDa.

Tricarboxylic Acid Cycle and Organic Transformation (Bin 8)—This functional class was represented by just five proteins in agreement with the fact that these are mostly mitochondrial functions (tricarboxylic acid cycle) or present in different subcellular compartments (carbonic anhydrases). We identified two types of malate dehydrogenases (MDHs), differing in the choice of cofactor (NADP or NAD) and activation mechanism. Chloroplast NADP-MDH (At5g58330) catalyzes the reduction of oxaloacetate into l-malate and is involved in the export of reducing power from the chloroplast to the cytosol (the malate valve). We identified it predominantly at 323 kDa and to a lesser extent at 154 kDa. Chloroplast NADP-MDH (At3g47520; 37 kDa) was suggested to form a homodimer (57), and we identified it at 96 kDa. The accumulation level of the NADPH form was higher than the NAD form in agreement with enzyme activity assays on purified Arabidopsis chloroplasts (58). The Arabidopsis genome encodes some 15 different carbonic anhydrases. We identified an abundant (“at expression level 2”) carbonic anhydrase (CA1; At3g01500) at multiple native masses between ~100 and 360 kDa in agreement with observations of homo-, octa-, and decameric states (59, 60). A second paralogue, CA2 (At5g14740), was identified ambiguously with CA1 at low levels (level 5).

Nitrogen and Sulfur Assimilation (Bins 12 and 14)—Plastids play a vital functional role in sulfur and nitrogen assimilation with both elements used in amino acid biosynthesis. Chloroplasts import nitrite, which is converted by Fd-dependent nitrite reductase (NiR) into (toxic) ammonia followed by the glutamate- and ATP-dependent conversion into glutamine by glutamine synthase 2 (GS2) and subsequent conversion into glutamate by Fd- and α-ketoglutarate-dependent glutamate synthase (Glu1 or Fd-GOGAT-1). We identified and quantified each of these three key enzymes (At2g15620, At5g35630, and At5g04140) as a monomer (NiR) and at different native masses (GS2 and GOGAT). GS2 appears to have a high relative concentration in particular as compared with NiR. This is logical because 90% of the glutamine synthesized in leaf chloroplasts is derived from photorespiration rather than by chloroplast import of nitrate and subsequent reduction to nitrite (for discussion, see Ref. 61).

ATP sulfurylase (ATPS1) catalyzes the formation of adenosine 5-phosphosulfate from inorganic sulfate and ATP. Four paralogues (ATPS1–4) are predicted to localize to non-green and/or green plastids in A. thaliana. Cytosolic and chloroplast isoforms were purified from spinach leaves, and their native mass was about 170 kDa as determined by gel filtration subsequent to other fractionation steps (62). We detected ATPS1 (At3g22890; ~49 kDa) at low levels (level 5) in the stroma with a native mass between 129 and 147 kDa.

Amino Acid Synthesis and Degradation (Bin 13)—We identified 17 proteins involved in amino acid metabolism, representing about 0.8% of the stromal protein mass. They were typically expressed at levels 3 and 4. Literature searches for their oligomeric state suggest that most of these accumulate as dimers and trimers (between ~100 and 150 kDa) in different plant species, and indeed these 17 proteins were found in this mass range (supplemental table). As an example, we mention the carbamoylphosphate synthetase large (At1g29900; 130 kDa) and small (At3g27740; 45 kDa) subunits in Aquifex aeolicus and in Escherichia coli, they were reported to form a heterodimer of 171 kDa (63). In appears that this oligomeric state is conserved in A. thaliana chloroplasts because we found both subunits with a native mass between 162 and 173 kDa, corresponding to a heterodimer.

Synthesis of Lipids (Bin 11), Hormones (Bin 16), Isoprenoids (Bin 17), and Cofactors and Vitamins (Bin 18)—We identified six proteins involved in lipid/fatty acid biosynthesis, all of them at low levels (mostly level 5), clearly less abundant than proteins involved in e.g. amino acid biosynthesis. These include acetyl-CoA synthetase (acetyl-CoA ligase) generating acetyl-CoA from acetate (typically produced from glycolysis in mitochondria), ACCase, and one of the three ketoacyl-ACP synthases, KAS1, as well as two desaturases, stearoyl-ACP...
desaturases 1 and 2. KAS1 (At5g46290; 50 kDa) is an essential enzyme involved in the construction of unsaturated fatty acid carbon skeletons, and we identified it at 130 kDa. A Case, catalyzing the first committed reaction of de novo fatty acid biosynthesis, forms a heterotetrameric enzyme with plastid-coded subunit β-carboxyltransferase, biotin carboxy carrier, biotin carboxylase, and α-carboxyltransferase. We did not identify these additional subunits most likely because they are primarily associated with the inner envelope membrane (64).

We identified two proteins in plastid isoprenoid biosynthesis from the non-mevalonic acid or methyl-α-erythritol 4-phosphate pathway. These were 4-diphosphocytidyl-2C-methyl-α-erythritol synthase (ISPD or CMS) and GcpE/IspG/Cib4 immediately downstream in the pathway. Both accumulated at “level 4.” GcpE/IspG was only recently localized to plastids (65). ISPD (At2g02500; ~30 kDa) was identified at 68 kDa on the CN-PAGE gel, probably forming a dimer.

Plant hormones such as jasmonic acid, gibberellic acid, and abscisic acid and products of the terpenoid pathway are, in part, synthesized in plastids. We identified lipooxygenase LOX2 (At3g45140; 102 kDa) involved in jasmonic acid synthesis with a predominant native mass of 116 kDa, corresponding to a monomer. It is surprisingly abundant, here quantified with a relative concentration of level 3. We also identified it with high Mascot scores in an earlier study as a thylakoid-associated protein (43). It was shown that chloroplast-localized LOX2 is required for the wound-induced synthesis of the plant growth regulator jasmonic acid in leaves (66).

Cofactors and vitamins are synthesized in plastids. We identified two plastid enzymes involved in thiamine/vitamin B, biosynthesis, Thi1 (level 3) and ThiC (level 4), both accumulating in complexes of 200 and 157 kDa, respectively (supplemental table). To our knowledge, ThiC has not been identified earlier in chloroplasts. Interestingly we found that Thi1 was heavily oxidized as evidenced by high levels of methionine oxidation in the mass spectrometer. This high level of methionine oxidation was clearly specific for Thi1.

Many enzymes in plastids are activated and deactivated through oxidation/reduction reactions via the thioredoxin system. The thioredoxin family consists of nine proteins grouped in four clusters (m1,2,3,4; x; y1,2; and f1,2) (73–75). We identified five thioredoxins on the CN-PAGE gels; TrxM1 was found at around 90 kDa, TrxM2 and TrxX were found at 154 kDa, and TrxX1 was found at both 154 and 212 kDa. Most likely they associate transiently with one or more different enzymes, but given the proteome complexity and the resolution of the native gels, we could not identify their respective partners. This is not surprising because recent affinity studies using modified thioredoxins as bait have shown that the thioredoxins interact with 50 or more chloroplast stromal proteins covering a wide range of functions (76).

We identified 14 proteins involved in oxidative stress responses. Some of them were quite abundant, such as peroxiredoxins A and B at level 2. Peroxiredoxin IIE (At3g52960) was identified at a native mass of 84 kDa, whereas the abundant peroxiredoxins A/B were found at multiple native masses in agreement with reports from the literature (77–79). We also identified several members of the ascorbate and glutathione defense systems, such as monodehydroascorbate reductase and dehydroascorbate reductase-2, involved in recycling oxidized ascorbate, and glutathione reductase. The native data suggest that they might interact with each other in an 80–100 kDa complex, corresponding to a heterodimer.

Nucleotide Synthesis and Degradation (Bin 23)—Plastids are a major site for pyrimidine and purine nucleotide synthesis, and indeed we identified 12 enzymes in these pathways (all at levels 3 and 4), corresponding to about 0.5% of stromal protein mass. Two of them are shared with amino acid biosynthesis. NDPK2 (At5g63310) is a nucleotide-diphosphate kinase involved in nucleotide metabolism (transfer of phosphate from NTP to NDP). NDPK2 (~20 kDa) was reported to form a homohexamer (80, 81) in agreement with an observed native mass of 133 kDa on the CN-PAGE gels. NDPK2 was predicted by TargetP and Predotar to be plastid-localized and was purified from spinach chloroplasts (81). Curiously NDPK2 is proposed to be a signal transducer in phytochrome-mediated light signaling, co-localizing with phytochrome in the nucleus (80, 82). In light of the purification from spinach chloroplasts and its significant accumulation level in A. thaliana chloroplasts in this study, it seems less likely that NDPK2 is localized in the nucleus.
RNA (Bin 27) and Protein Synthesis and Protein Fate (Bin 29)—We were quite surprised to identify so many proteins (67) assigned to Bins 27 and 29, together representing about 7% of the stromal mass. The relative concentration among these proteins spanned 3–4 orders of magnitude with ROC4 (At3g62030; 20 kDa) as the most abundant protein. ROC4 is an abundant stromal peptidylprolyl isomerase with demonstrated in vitro rotamase activity, but its role is unclear (83). ROC4 was found predominantly around 110 kDa and may form a hexamer.

Spinach plastid 70 S ribosomes are composed of more than 60 proteins and have a native mass of around 2 MDa (20, 21). The stromal 70 S ribosomes migrated just into the CN-PAGE gel with low amounts of other large complexes (Fig. 1, A and B). We analyzed 21 protein spots in this gel area, but the analysis was not exhaustive. Nevertheless we identified some 11 30 S subunits, 10 50 S subunits, and a plastid-specific ribosome-associated protein (PSRP2). Ribosome-associated proteins RAP41 (At3g63140) and RAP38 (At1g09340), originally identified in C. reinhardtii ribosomes (84), were each found at three different locations of the stromal CN-PAGE native gels: (i) in a complex larger than 950 kDa most likely associated with 70 S ribosomes, (ii) at 224 kDa, and (iii) at 106–126 kDa. At 224 kDa, the only obvious potential partners are ribosomal protein L5 (At4g01310) and ribosomal protein L31 (At1g75350) (Fig. 2A). At 106–126 kDa, no obvious partners were found, suggesting the possibility that RAP38 and RAP41 form a heterotrimer.

We identified elongation factors, EF-Tu-1 (At4g20360; 45 kDa), ET-G (At1g62750; 78 kDa), and EF-P (At3g08740; 21 kDa) as well a new elongation factor typeA/bipA like protein (At5g13650; 69 kDa). The BipA-like protein was found in a complex over 950 kDa likely interacting with ribosomes. In E. coli, BipA is required specifically for the expression of the transcriptional modulator Fis and binds to ribosomes at a site that coincides with that of elongation factor G (85). Elongation factor EF-Tu-1 found at multiple native masses ranging from 110 to over 950 kDa was identified several times with high Mascot scores (up to 732) and was abundant (level 2), suggesting that it has additional functions. Indeed an orthologue in maize plastids was suggested to also serve as a chaperone in particular during heat stress (86, 87).

We identified several proteins involved in mRNA binding and processing/degradation. Polynucleotiducleotide nucleotidyltransferase (At3g03710; ~95 kDa) acts as a 3′–5′ exoribonuclease. We identified it in a stromal complex of 410 kDa. This protein shows sequence homology to the polynucleotide phosphorylase (PNPase) of the E. coli degradosome. It possibly acts as a homotetramer (88–90) and not as a heterooligomer as shown in other non-plant euukaryotes. We also identified At5g48960 (~65 kDa) encoding for a putative 5′-nucleotidase with unknown function; it was detected at 166 kDa in the stroma. At5g26742 encodes a putative DEAD box RNA helicase (RH3) and was identified on the CN-PAGE stromal gel in a molecular mass complex of >950 kDa. A plastid-localized RH3 was identified in tobacco, and a loss of function mutant resulted in variegated leaves and abnormal roots and flowers (91). In E. coli, a DEAD-RNA helicase is part of the “degradosome” along with the PNPase, the endoribonuclease RNase E, and the glycolytic enzyme enolase (92). RH3 in chloroplasts does not seem to be associated with the PNPase or At5g48960 mentioned above.

Chaperones cpHSP70-1 (At4g24280) and cpHSP70-2 (At5g49910) are most likely in a complex of ~200 kDa with GrpE1 (At5g17710) and GrpE2 (At1g36390) (Fig. 2B). In addition, GrpE1 and GrpE2 seem to form a complex at 150 kDa without cpHSP70, but a potential pitfall is that transketolase is a major spot possibly masking cpHSP70. Indeed when analyzing the CN-PAGE gel of the peripheral thylakoid proteins, where transketolase is less abundant, cpHSP70 was detected close to transketolase at ~150 kDa and most likely forms a second type of complex with GrpE1–2 (not shown). cpHSP70-1 and cpHSP70-2 (mostly the -1 isoform) also form complex at 123 kDa, but we did not detect any GrpE in this...
native mass range. GrpE and cpHSP70 have been shown to form complexes of ~120 and 230 kDa in chloroplasts of C. reinhardtii, and GrpE also forms homo-oligomers (93).

We identified a well resolved CPN60 complex at 806 kDa with \( \alpha \) (At2g28000) and \( \beta \) subunits (At5g56500, At3g13470, and At1g55490) in an approximate 1:1 ratio. CPN10 (At2g44650) or CPN20 (At5g20720) were not part of this 800-kDa complex but were found in complexes of 150–170 kDa. A trigger factor-like protein (At5g55220; ~60 kDa) was identified most likely as a dimer. Trigger factor in the cytosol of E. coli prevents misfolding and aggregation of nascent proteins emerging from the 70 S ribosome (94). The structure of “free” trigger factor (without ribosomes) of Vibrio cholerae was determined at 2.5-Å resolution as a dimer (95).

The ClpP/R core protease complex was detected at 325–350 kDa. It is a hexadecamer composed of two rings of seven subunits with ClpP1,3,4,5,6 and ClpR1,2,3,4 in one or more copies associated with two chaperone-like proteins named ClpS1 and -S2 (16). Oligopeptidase A in the M3 peptidase family (At5g65620; 88 kDa) and a zinc-metalloprotease of the M16 family (At3g19170; 120 kDa) both have unknown function and accumulated as monomers.

Miscellaneous Function and “Unknown” Proteins (Bins 24, 26, 28, and 35)—A significant number of the identified A. thaliana proteins currently have no obvious function (supplemental table). Here we highlight just a few and comment on their native mass, abundance, and functions in non-plant species.

Denelactone hydrolase is a monomeric enzyme in bacteria involved in the degradation of aromatic compounds such as chlorocatechol and has been crystallized from Pseudomonas (96). Nothing is known about its substrate in plants. We found a dienelactone hydrolase at 27.5 kDa (At1g35420; ~30 kDa) as monomer. At1g21440 and At1g77060 are ~33-kDa paralogues and encode for a putative plastid-localized mutase likely involved in the formation of C–P bonds. In Streptomyces hygroscopicus it is involved in the biosynthesis of phosphonates (97, 98). We identified both proteins in the same spot at a native mass of 119 kDa; it is quite possible that they form trimers. At1g77060 was also identified at 147 kDa. The biosynthesis of phosphonate has not been investigated in plants, and the only report found in the literature concerns the accumulation of synthase mRNA in senescing carnation (Dianthus caryophyllus) flower petals in response to ethylene (99).

The DAG gene (At1g11430) encodes a 20-kDa protein with no obvious functional domains involved in the development of chloroplasts (100). We identified the DAG protein in a high molecular mass complex of >950 kDa. The protein THF1 (At2g20890) involved in thylakoid biogenesis was shown to be present in chloroplasts as well as in plastid-associated stromules (101). We originally identified the 25-kDa THF1 protein under denaturing conditions in the stripped thylakoid proteome (43) and in this study as a 270-kDa complex in the stroma. Interacting protein partners are unknown.
of the proteome (cleavable ICAT and iTRAQ) (for reviews, see Refs. 103–106). However, these techniques cannot be used to compare accumulation levels of different proteins within a sample and are thus not applicable to answer questions raised in the current study. A new approach has been introduced recently in which isotope-labeled peptides matching to individual proteins are added as internal standards into complex mixtures. This technique (assigned AQUA) is most appropriate for comparing expression levels of a small number of known proteins, and it is not suitable for quantification of unidentified proteins (107). Denaturing IEF using commercial IPG strips is not quantitative when comparing stoichiometries of different protein species because of underestimation of high molecular mass proteins and proteins with significant hydrophobicity or extreme pl. Here we show that native gels, although not perfect, currently provide the most convenient semiquantitative comparison of larger sets of protein species.

This study provides an overview of the steady state accumulation levels of stromal proteins and provides insight in the relative concentrations of different metabolic pathways. For instance, the data clearly suggest that components of protein synthesis, folding, and proteolysis are generally more abundant than enzymes involved in amino acid synthesis. Enzymes in fatty acid biosynthesis are general of lower abundance in the stroma than those of amino acid synthesis possibly because the first are typically associated with the inner chloroplast envelope. The Calvin cycle, OPPP, and glycolysis together represent some 75% of the total mass with protein synthesis, folding, and biogenesis representing about 6.5% and nitrogen and sulfur assimilation representing an additional 7.5%. This leaves less than 10% of the total stromal mass for all other stromal functions. The large investments in carbon metabolism and the Calvin cycle in particular reflect the inefficiency of Rubisco as well as the high metabolic flux of reduced carbohydrates. The large protein mass dedicated to protein synthesis (ribosomes, RNA-binding proteins, and elongation factors), protein folding, and proteolysis is likely a reflection of the abundance of the chloroplast-encoded large subunit of Rubisco and chloroplast-encoded proteins of the thylakoid photosynthetic apparatus. The abundance of the chaperones is likely also a reflection of the need for folding and assembly of the thousands of nuclear encoded proteins imported into the stroma. The identified proteins span at least an expression range of 5 orders of magnitude.

It is interesting to contrast the proteome of chloroplasts with those of non-green plastids. Non-green plastids exist in roots, non-green flowers, fruits, and seeds. Two shotgun proteomic studies were published concerning (non-green) wheat amyloplasts and tobacco (colorless) BY-2 cell culture plastids (108, 109). It is hard to compare these identified proteomes to the current chloroplast proteome analysis because these studies did not quantify protein accumulation but also because wheat, tobacco, and A. thaliana are quite distant in evolution with relatively low levels of sequence homology. Nevertheless the most striking difference between the BY-2 plastid proteome and the chloroplast proteome is that a significantly higher percentage of identified proteins is involved in amino acid metabolism in the BY-2 cells (25 versus 7%).

Determination of Native Mass and Deduction of Oligomeric State of Chloroplast Proteins—In this study we present a “snapshot” of the oligomeric chloroplast proteome of A. thaliana, isolated during the first half of the 10-h light period from complete rosettes of 6-week-old plants. The CN-PAGE gels were reproducible across different stromal preparations, suggesting that the native state of the stromal proteome visualized on the gel was relatively stable under the conditions used. We chose to use CN gels originally developed by Schägger et al. (28) rather than the more popular BN gels (110) because the high concentrations of Coomassie in the BN-PAGE sample buffer destabilized a significant fraction of the soluble complexes (not shown). Although BN-PAGE has been successful for separation of membrane complexes (111–114), it is less suited for native mass determination of the soluble stroma. The effective native mass range on the CN-PAGE gels was about 10–950 kDa. It should be noted that very little destabilization of complexes took place during separation in the first dimension as evidenced by the near absence of “streaking” in the first dimension. The most significant difference with studies reported in the literature is that the native stromal proteome was separated under low salt conditions while keeping the protein concentrations high until the point of entering the native gel. Most other studies analyzed “one protein (complex) at a time” and involved typically multiple chromatography steps (ion exchange and gel filtration) often combined with ammonium salt precipitations and sometimes co-immunoprecipitations followed by N-terminal Edman sequencing or Western blots.

About 10% of the 241 proteins identified in this study were found in a monomeric state. More than 85% of the proteins were essentially (within a ~10-kDa mass range) found at a single native mass. The remaining 15% were found at multiple native masses either because they were abundant and trace amounts of streaking were enough to identify the protein by the sensitive mass spectrometry measurements or because they likely interacted with different partners or formed different homo-oligomeric states. Examples of the latter explanation include thioredoxin f at 155 and 212 kDa, uridylyltransferase-related identified as a ~30-kDa monomer and at 287 and 721 kDa, 3-β-hydroxy-Δ5-steroid dehydrogenase as a monomer and at 112 kDa, and chloroplast NAD malate dehydrogenase at 323 and 154 kDa.

Searching the literature for native mass information for each of the identified stromal proteins or their homologues, we were surprised to find that ~60% of the complexes reported in the literature are homomeric rather than heteromeric. These complexes are involved in a wide variety of metabolic activity without any particular bias for metabolic function. In a few cases, it appeared that these complexes are homomeric in
The Oligomeric Chloroplast Stroma

cellular chloroplasts but heteromeric in non-plant species. For example, aspartate transcarbamoylase in pea chloroplasts is a homotrimer, whereas in other eukaryotes it is part of a multifunctional enzyme containing glutamine-dependent carbamoylphosphate synthetase and dihydroorotase activities (115). It is not clear whether there is a more general trend for a higher degree of homo-oligomerization in plastids than elsewhere in the plant or in other non-plant species.

The prevalence of homomeric complexes over heteromeric complexes in the published literature is striking and warrants examination. Does this reflect a bias due to the methodology used (mostly non-denaturing chromatography involving salt) and proteins studied, or does this truly reflect the protein interaction network in plastids? If indeed homo-oligomers are prevailing over heteromers, what could be the benefit for plastid function? It might be easier to regulate protein complex concentrations if they consist of only one gene product. Maybe many interactions in vivo are in fact between different homo-oligomers forming larger hetero-oligomers, but these interactions are more dynamic and less stable than those between homomers themselves. Currently it is difficult to evaluate any of these explanations. Different types of protein-protein interaction studies are required as well as in vivo monitoring of protein-protein interactions using techniques such as fluorescence resonance energy transfer and bioluminescence resonance energy transfer (116–118).

In recent years, several large scale protein interaction studies in yeast (119, 120) and E. coli (121) using tandem affinity purification (TAP) tags and mass spectrometry were carried out. These experiments were focused on identifying hetero-oligomers, whereas homo-oligomers were not recognized because the identification step involved only denaturing SDS-PAGE. No native masses of the tagged complex were measured in these studies. It would be interesting to determine how many of these TAP-tagged proteins in yeast and E. coli (for which no heteromeric interaction was found) in fact form homo-oligomeric structures.

Protein Stoichiometry, Native Mass, and Paralogue Identification of Enzymes in the Calvin Cycle: Data Integration and Interactive Figure—As an example as to how the experimental data, prediction data, and published literature can be integrated, we generated a diagram for the Calvin cycle, one of the most studied plant pathways (Fig. 3). The experimental data concerning the relative concentrations and native state of the Calvin cycle enzymes and the regulators CP12, Rubisco activase, and Rubisco methyltransferase were integrated in Fig. 3. Because this figure contains a fair amount of detail, we also incorporated the figure into the PPDB, allowing retrieval of associated information for each accession. All A. thaliana proteins and paralogues associated with the Calvin cycle were collected using information from two metabolic databases, Aracyc (www.arabidopsis.org/tools/aracyc/) and MapMan, as well as published literature. The presence of chloroplast transit peptides was predicted using TargetP. Accessions in purple have no predicted cTP, whereas those in green have predicted cTPs. Proteins identified on the CNPAGE gel of the stromal proteome are underlined. Enzymes boxed in orange and in blue are shared with OPPP and glycolysis, respectively.

A significant number of paralogues to Calvin cycle enzymes are present in the A. thaliana genome but were not identified in the stroma. Several of these paralogues (in particular for SFBA) are predicted to be located outside the plastid and are likely cytosolic, most likely involved in cytosolic glycolysis and OPPP. The expression and accumulation of unidentified paralogues predicted to be plastid-localized d-ribose-5-phosphate isomerase, phosphoribulokinase, phosphoglycerate kinase, GAPDH, SFBA, and transketolase are interesting. It is possible that these paralogues are expressed either at low levels (several orders of magnitude lower as for the identified Calvin cycle enzymes) in chloroplasts or, more likely, are important in non-green plastids. Unpublished data show that several are expressed in non-green root or flower plastids in different Brassica sp.4 Interestingly the Rubisco activase paralogue At1g773110 was not identified in the stroma; however, this protein was identified with high Mascot scores in our earlier study on the stripped and denatured thylakoid proteome (43). It is not clear how this tight thylakoid association should be interpreted. All other Calvin cycle enzymes were predicted to be plastid-localized (by TargetP) with the exception of triose-phosphate isomerase 2 (TPI-2) (At3g55440). It was identified unambiguously with high MOWSE scores (~200) in MS/MS; given the high purity of the stromal samples, this indicates that TPI-2 is likely chloroplast-localized and that the cTP prediction is incorrect.

As is indicated in Fig. 3, all but three of the Calvin cycle enzymes are shared with glycolysis or with the OPPP. A number of paralogues for several of these enzymes show specificity to one pathway. Examples are GAPDH with GAPDH-Cp-1 (At1g16300) involved in glycolysis and GAPDH-A1,A2 and -B (At1g12900, At3g26650, and At1g42970) involved in the Calvin cycle (for a review, see Ref. 122).

Conclusions—This study provides a first semiquantitative overview of the oligomeric state of the stromal proteome of A. thaliana chloroplasts and relates this information to plastid functions. This collective dataset is a major step forward in the analysis of the stromal proteome. These experimental observations were compared with published literature on stromal chloroplast complexes. Data are accessible via the PPDB. The integration of literature, experimental paralogue identification, and protein quantification for the Calvin cycle highlights the value of the information accumulated in this study and the challenge for understanding metabolic pathways. To truly understand chloroplast metabolic function, these protein measurements need to be integrated with information of their post-translational modifications, binding of substrate mole-

4 J.-B. Peltier, G. Friso, and K. J. van Wijk, unpublished data.
molecules, and measurements of metabolic flux and metabolite levels as is eloquently discussed in a recent review (123). Systematic TAP tagging of chloroplast proteins and measurements of protein interactions under different conditions are needed to truly establish stromal protein interactions networks in particular for proteins of low abundance. New technologies are now in place to advance each of these different aspects.

Acknowledgments—We thank in particular Wojciech Majeran for discussions and Giulia Friso for help with mass spectrometry.

* This work was supported by the National Science Foundation, division of Molecular and Cellular Biochemistry (Grant 090942), the United States Department of Agriculture (Biochemistry) (Grant 2003-35100-13579), and New York State Office of Science, Technology, and Research. Large scale data collection at Cornell University was conducted using the resources of the Cornell Theory Center, which receives funding from Cornell University, New York State, federal agencies, foundations, and corporate partners. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ The on-line version of this article (available at http://www.mcponline.org) contains supplemental material.

¶ Present address: Laboratoire de protéomique 2, Place Viala 34060, Montpellier cedex 1, France.

‖ Present address: The Research Inst. for Children, Children’s Hospital, 200 Henry Clay Ave., New Orleans, LA 70118.

** Present address: Thermo Electron Corp., 355 River Oaks Parkway, San Jose, CA 95134.

‡‡ To whom correspondence should be addressed: Dept. of Plant Biology, Emerson Hall 3rd floor, Tower Rd., Cornell University, Ithaca, NY 14853, Tel.: 607-254-1211; Fax: 607-255-5407; E-mail: kv35@cornell.edu.

---

FIG. 3. Protein stoichiometry, native mass, and paralogue identification of enzymes in the Calvin cycle. All A. thaliana proteins associated with the Calvin cycle were collected using information from two metabolic databases (Aracyc) and the MapMan Bin system as well as published literature. The presence of chloroplast transit peptides was predicted using TargetP. The relative concentration of each protein in the stromal proteome (identified from the CN-PAGE gels) was calculated based on the stained spot volumes divided by the experimentally determined denatured masses. An interactive figure that allows retrieval of more detailed information can be found in PPDB at ppdb.tc.cornell.edu. Accessions in purple have no predicted cTP, whereas those in green have predicted cTPs. The proteins identified on the CN-PAGE gel are underlined. Enzymes boxed in orange and in blue are shared with OPPP and glycolysis, respectively.
The Oligomeric Chloroplast Stroma

REFERENCES

1. Buchanan, B., Gruissem, W., and Jones, R. L. (eds) (2000) Biochemistry and Molecular Biology of Plants, American Society of Plant Physiolo-
gists, Rockville, MD

2. Monde, R. A., Schuster, G., and Stern, D. B. (2000) Processing and degradation of chloroplast mRNA. Biochimie (Paris) 82, 573–582

3. Barkan, A., and Goldschmidt-Clermont, M. (2000) Participation of nuclear genes in chloroplast gene expression. Biochimie (Paris) 82, 559–572

4. Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G. (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J. Mol. Biol. 300, 1003–1016

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

6. van Wijk, K. J. (2004) Plastid proteomics. Plant Physiol. Biochem. 42, 963–977

7. Ephritikhine, G., Ferro, M., and Rolland, N. (2004) Plant membrane proteo-

8. Whitelegge, J. P. (2004) Mass spectrometry for high throughput quanti-

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

10. Sun, Q., Emanuelsson, O., and van Wijk, K. J. (2004) Analysis of curated and predicted plastid subproteomes of Arabidopsis. Subcellular com-
partmentalization leads to distinctive proteome properties. Plant Physiol. 135, 723–734

11. Cornah, J. E., Terry, M. J., and Smith, A. G. (2003) Green or reed: what stops the traffic in the tetrapyrrole pathway? Trends Plant Sci. 8, 231–239

12. Kruger, N. J., and van Schaeven, A. (2003) The oxidative pentose phos-

13. Plaxton, W. C. (1996) The organization and regulation of plant glycolysis. Annu. Rev. Plant Physiol. Plant Mol. Biol. 47, 185–214

14. Esposito, S., Guerrier, G., Vona, V., Di Martino Riganò, V., Cartagna, S., and Riganò, C. (2005) Glutamate synthase activities and protein changes in relation to nitrogen nutrition in barley: the dependence on different plastidic glucose-6P dehydrogenase isoforms. J. Exp. Bot. 56, 55–64

15. Winkel, B. S. (2004) Metabolic channeling in plants. Annu. Rev. Plant Biol. 55, 85–107

16. Peltier, J. B., Ripoll, D. R., Friso, G., Rudella, A., Cai, Y., Ytterberg, J., Mooney, B. P., Miernyk, J. A., and Randall, D. D. (2002) The complex fate of a-ketoacids. Annu. Rev. Plant Biol. 53, 357–375

17. Kervinen, J., Dunbrack, R. L., Jr., Litwin, S., Martins, J., Scarrow, R. C., Volin, M., Yeung, A. T., Yoon, E., and Jaffe, E. K. (2000) Porphobilinogen synthase from pea: expression from an artificial gene, kinetic charac-
terization, and novel implications for subunit interactions. Biochemistry 39, 9018–9029

18. Kervinen, J., Dunbrack, R. L., Jr., Litwin, S., Martins, J., Scarrow, R. C., Volin, M., Yeung, A. T., Yoon, E., and Jaffe, E. K. (2000) Porphobilinogen synthase from pea: expression from an artificial gene, kinetic charac-
terization, and novel implications for subunit interactions. Biochemistry 39, 9018–9029

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

20. Porra, R. J., Thompson, F. A., and Kriedemann, P. E. (1989) Determina-
tion of accurate extinction coefficients and simultaneous equations for assaying chlorophyll a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. Biochim. Biophys. Acta 975, 384–394

21. Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goede, N. M., Olson, B. J., and Physik, D. C. (1985) Measurement of protein using bicinchoninic acid. Anal. Biochem. 150, 76–85

22. Schaggher, H., Cramer, W. A., and van Jagow, G. (1994) Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. Anal. Biochem. 217, 220–230

23. Thimm, O., Blasing, O., Gibson, Y., Nagel, A., Meyer, S., Kruger, P., Selbig, J., Muller, L. A., Rhee, S. Y., and Stitt, M. (2004) MAMPAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J. 37, 914–939

24. Marmagne, A., Rouet, M. A., Ferro, M., Rolland, N., Alcon, C., Joyard, J., Garin, J., Barbier-Brygoo, H., and Ephritikhine, G. (2004) Identification of new intrinsic proteins in Arabidopsis plasma membrane proteome. Mol. Cell. Proteomics 3, 675–691

25. Nuhse, T. S., Stensballe, A., Jensen, O. N., and Peck, S. C. (2003) Large-scale analysis of in vivo phosphorylated membrane proteins by immobilized metal ion affinity chromatography and mass spectrometry. Mol. Cell. Proteomics 2, 1234–1243

26. Shimaoka, T., Onishi, M., Sazuka, T., Mitsuhashi, N., Hara-Nishimura, I., Shimazaki, K., Maeshima, M., Yokota, A., Tomizawa, K., and Mimura, T. (2004) Isolation of intact vacuoles and proteomic analysis of tonoplast from suspension-cultured cells of Arabidopsis thaliana. Plant Cell Physiol. 45, 672–683

27. Szponarski, W., Sommerer, N., Boyer, J. C., Rossignol, M., and Gratras, B. (2004) Large-scale characterization of integral proteins from Arabidop-
sis vacuolar membrane by two-dimensional liquid chromatography. Proteomics 4, 397–406

28. Fukao, Y., Hayashi, M., and Nishimura, M. (2002) Proteomic analysis of leaf peroxisomal proteins in greening cotyledons of Arabidopsis thali-
a. Plant Cell Physiol. 43, 689–696

29. Bae, M. S., Cho, E. J., Choi, E. Y., and Park, O. K. (2003) Analysis of the Arabidopsis nuclear proteome and its response to cold stress. Plant J. 36, 652–663

30. Borderies, G., Jamet, E., Lafitte, C., Rossignol, M., Seigneurin-Berny, D., Monsarrat, B., Esquerre-Tugaye, M. T., Boudet, A., and Pont-Lezica, R. (2003) Proteomics of loosely bound cell wall proteins of Arabidopsis thaliana cell suspension cultures: a critical analysis. Electrophoresis 24, 3421–3432

31. Brugiere, S., Kowalski, S., Ferro, M., Seigneurin-Berny, D., Miras, S., Salvi, D., Ravelan, S., D’Herin, P., Garin, J., Bourguignon, J., Joyard, J., and Rolland, N. (2004) The hydrophobic proteome of mitochondrial membranes from Arabidopsis cell suspensions. Phytochemistry 65, 1693–1707

32. Heazlewood, J. L., Tonti-Filippini, J. S., Gout, A. M., Day, D. A., Whelan, J., and Millar, A. H. (2004) Experimental analysis of the Arabidopsis mito-
chondrion. Plant Physiol. 138, 224–233

33. Borderies, G., Jamet, E., Lafitte, C., Rossignol, M., Seigneurin-Berny, D., Monsarrat, B., Esquerre-Tugaye, M. T., Boudet, A., and Pont-Lezica, R. (2003) Proteomics of loosely bound cell wall proteins of Arabidopsis thaliana cell suspension cultures: a critical analysis. Electrophoresis 24, 3421–3432

34. Brugiere, S., Kowalski, S., Ferro, M., Seigneurin-Berny, D., Miras, S., Salvi, D., Ravelan, S., D’Herin, P., Garin, J., Bourguignon, J., Joyard, J., and Rolland, N. (2004) The hydrophobic proteome of mitochondrial membranes from Arabidopsis cell suspensions. Phytochemistry 65, 1693–1707

35. Heazlewood, J. L., Tonti-Filippini, J. S., Gout, A. M., Day, D. A., Whelan, J., and Millar, A. H. (2004) Experimental analysis of the Arabidopsis mito-
chondrion. Plant Physiol. 138, 224–233
Grunwald, D., Garin, J., Joyard, J., and Rolland, N. (2002) Integral membrane proteins of the chloroplast envelope: identification and subcellular localization of new transporters. Proc. Natl. Acad. Sci. U. S. A. 99, 11487–11492

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

43. Pelletier, J. B., Ytterberg, A. J., Sun, Q., and van Wijk, K. J. (2004) 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

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

45. Schubert, M., Petersson, U. A., Haas, B. J., Funk, C., Schroder, W. P., and Kieselbach, T. (2002) Proteome map of the chloroplast lumen of Arabidopsis thaliana. J. Biol. Chem. 277, 8354–8365

46. Gomez, S. M., Nishio, J. I., Faull, K. F., and Whitelegge, J. P. (2002) The chloroplast grana proteome defined by intact mass measurements from liquid chromatography mass spectrometry. Mol. Cell. Proteomics 1, 46–59

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

48. Mates, C. J., Hudson, G. S., van Caemmerer, S., Evans, J. R., and Andrews, T. J. (1993) Reduction of ribulose biphosphate carboxylase active sites in tobacco (Nicotiana tabacum) by antisense RNA reduces ribulose biphosphate carboxylase carboxylation and impairs photosynthesis. Plant Physiol. 102, 1119–1128

49. Spreitzer, R. J., and Salvucci, M. E. (2002) Rubisco: structure, regulatory interactions, and possibilities for a better enzyme. Annu. Rev. Plant Biol. 53, 469–493

50. Pelzer-Reith, B., Wiegand, S., and Schnarrenberger, C. (1994) Plastid membrane proteins of the chloroplast envelope: identification and subcellular localization of new transporters. J. Biol. Chem. 269, 8059–8064

51. Dietz, K. J. (2003) Plant peroxiredoxins. Annu. Rev. Plant Physiol. Plant Mol. Biol. 54, 93–107

52. Jahn, D., Kim, Y. C., Ishino, Y., Chen, M. W., and Soll, D. (1990) Purification and functional characterization of the Glu-tRNA(Gln) amidotransferase from Chlamydomonas reinhardtii. J. Biol. Chem. 265, 8059–8064

53. Jordan, D. B., Bacot, K. O., Carlson, T. J., Kessel, M., and Vilaitan, P. V. (1999) Plant riboflavin biosynthesis. Cloning, chloroplast localization, expression, purification, and partial characterization of spinach lumazine synthase. J. Biol. Chem. 274, 22114–22121

54. Fornasari, M. S., Laplagne, D. A., Frankel, N., Cauerhff, A. A., Goldbaum, F. A., and Echave, J. (2002) Proteomics uncovers proteins interacting electrostatically with thioredoxin system in lumazine synthase. Mol. Biol. Evol. 21, 97–107

55. Persson, K., Schneider, G., Jordan, D. B., Vilaitan, P. V., and Sandalova, T. (1999) Crystal structure analysis of a pentameric fungal and an icoshedral plant lumazine synthase reveals the structural basis for differences in assembly. Protein Sci. 8, 2355–2365

56. Konig, J., Lotte, K., Plessow, R., Brockhinke, A., Baier, M., and Dietz, K. J. (2003) Reaction mechanism of plant 2-Cys peroxiredoxin. Role of the C-terminal segment and the quaternary structure. J. Biol. Chem. 278, 24409–24420

57. Wood, Z. A., Schroder, E., Robin Harris, J., and Pooler, L. B. (2003) Structure, mechanism and regulation of peroxiredoxins. Trends Biochem. Sci. 28, 32–40

58. Shen, Y., Kim, J. I., and Song, P. S. (2005) NDK2 as a signal transducer in the phytochrome-mediated light signaling. J. Biol. Chem. 280, 5740–5747

59. Yang, L. M., and Lamppa, G. K. (1996) Rapid purification of a chloroplast nucleoside diphosphate kinase using CoA-affinity chromatography. Biochim. Biophys. Acta 1294, 99–102

60. Choi, G., Yi, H., Lee, J., Kwon, Y. K., Soh, M. S., Shin, B., Luka, Z., Hahn, T. R., and Song, P. S. (1999) Phytochrome signalling is mediated through nucleoside diphosphate kinase 2. Nature 401, 610–613
The Oligomeric Chloroplast Stroma

83. Lipponer, V., Chou, I. T., Scott, S. V., Ettinger, W. F., Theg, S. M., and Gasser, C. S. (1994) Cloning and characterization of chloroplast and cytosolic forms of cyclophilin from Arabidopsis thaliana. J. Biol. Chem. 269, 7863–7868

84. Yamaguchi, K., and Subramanian, A. R. (2003) Proteomic identification of all plastid-specific ribosomal proteins in higher plant chloroplast 30S ribosomal subunits. Eur. J. Biochem. 270, 190–205

85. Owens, R. M., Pritchard, G., Skipp, P., Hood, M., Connell, S. R., Nierhaus, K. H., and O’Connor, C. D. (2004) A dedicated translation factor controls the synthesis of the global regulator Fis. EMBO J. 23, 3375–3385

86. Bhadula, S. K., Elthon, T. E., Habben, J. E., Helentjaris, T. G., Jiao, S., and Owens, R. M., Pritchard, G., Skipp, P., Hodey, M., Connell, S. R., Nierhaus, K. H., and O’Connor, C. D. (2004) Domain analysis of the chloroplast polynucleotide phosphorylase reveals discrete functions in RNA degradation, polyadenylation, and ribosome heterogeneity with exosome proteins. Plant Cell 15, 2003–2015

87. Baginsky, S., Sheteiman-Kotler, A., Liveanu, V., Yehudai-Resheff, S., Beliaouei, M., Settlage, R. E., Shabanowitz, J., Hunt, D. F., Schuster, G., and Gussem, W. (2001) Chloroplast PNPase exists as a homo-multimer enzyme complex that is distinct from the Escherichia coli degradosome. RNA 7, 1464–1475

88. Hayes, R., Kudla, J., Schuster, G., Gabay, L., Maliga, P., and Gruissem, W. (1996) Chloroplast mRNA 3′–end processing by a high molecular weight protein complex is regulated by nuclear encoded RNA binding proteins. EMBO J. 15, 1132–1141

89. Wang, Y., Duby, G., Purnelle, B., and Boutry, M. (2000) Tobacco VDL gene encodes a plastid DEAD box RNA helicase and is involved in chloroplast differentiation and plant morphogenesis. Plant Cell 12, 2129–2142

90. Carposius, A. J. (2002) The Escherichia coli RNA degradosome: structure, function and relationship in other ribonucleolytic multienzyme complexes. Biochem. Soc. Trans. 30, 150–155

91. Schroda, M., Vallon, O., Whitelegge, J. P., Beck, C. F., and Wollman, F. A. (2003) A cradle for new proteins: trigger factor as a nucleation factor (EF-Tu) in a heat-tolerant maize line. Planta 212, 359–366

92. Rao, D., Momcilovic, I., Kobayashi, S., Callegari, E., and Ristic, Z. (2004) Chaperone activity of recombinant maize chloroplast protein synthesis elongation factor, EF-Tu in a heat-tolerant maize line. Planta 212, 359–366

93. Bhadula, S. K., Elthon, T. E., Habben, J. E., Helentjaris, T. G., Jiao, S., and Owens, R. M., Pritchard, G., Skipp, P., Hodey, M., Connell, S. R., Nierhaus, K. H., and O’Connor, C. D. (2004) Domain analysis of the chloroplast polynucleotide phosphorylase reveals discrete functions in RNA degradation, polyadenylation, and ribosome heterogeneity with exosome proteins. Plant Cell 15, 2003–2015

94. Owens, R. M., Pritchard, G., Skipp, P., Hood, M., Connell, S. R., Nierhaus, K. H., and O’Connor, C. D. (2004) A dedicated translation factor controls the synthesis of the global regulator Fis. EMBO J. 23, 3375–3385

95. Ludlam, A. V., Moore, B. A., and Xu, Z. (2004) The crystal structure of tobacco bright yellow-2 (BY-2) cell culture plastids as a model for undifferentiated heterotrophic plastids. J. Proteome Res. 3, 1128–1137

96. Ristic, Z. (2001) Heat-stress induced synthesis of chloroplast protein synthesis elongation factor (EF-Tu) in a heat-tolerant maize line. EMBO J. 20, 265–273

97. Pathak, D., and Ollis, D. (1990) Refined structure of dienelactone hydrolyase at 1.8 Å. J. Mol. Biol. 214, 497–525

98. Wang, Y., Duby, G., Purnelle, B., and Boutry, M. (2000) Tobacco VDL gene encodes a plastid DEAD box RNA helicase and is involved in chloroplast differentiation and plant morphogenesis. Plant Cell 12, 2129–2142

99. Wang, H., Brandt, A. S., and Woodson, W. R. (1993) A flower senescence-resistance component in a plasma membrane microdomain. Proc. Natl. Acad. Sci. U. S. A. 102, 3135–3140

100. Ho, Y., Gruhler, A., Heilbut, A., Bader, G. D., Moore, L., Adams, S. L., Andersen, J. R., Johansen, L. E., Hansen, L. H., Jespersen, H., Podtelejnikov, A., Nielsen, E., Crawford, J., Poulsen, V., Sorensen, B. D., Matthiesens, J., Hendrickson, R. C., Gleeson, F., Pawson, T., Moran, M. F., Durocher, D., Mann, M., Hogue, C. W., Figeys, D., and Tyers, M. (2002) Proteomic characterization of wheat amyloplasts using identification of proteins by tandem mass spectrometry. Proteomics 2, 1156–1168

101. Schagger, H., Aquila, H., and von Jagow, G. (1989) Coomassie blue-stained sodium dodecyl polyacrylamide gel electrophoresis for direct visualization of polypeptides during electrophoresis. Anal. Biochem. 173, 201–205

102. Heinmeyer, J., Eubel, H., Wehmhoner, D., Jansch, L., and Braun, H. P. (2004) Proteomic approach to characterize the supramolecular organization of photosystems in higher plants. Phytochemistry 65, 1683–1692

103. Singh, P., Jansch, L., Braun, H. P., and Schmitz, U. K. (2000) Resolution and structural characterization of mitochondrial and chloroplast membrane protein complexes from green leaves of potato on blue-native polyacrylamide gels. Indian J. Biochem. Biophys. 37, 59–66

104. Williamson, C. L., and Slocum, R. D. (1994) Molecular cloning and characterization of the pyrB1 and pyrB2 genes encoding aspartate transcarbamoylase in pea (Pisum sativum L.). Plant Physiol. 105, 377–384

105. Bhat, R. A., Miklis, M., Schmelzer, E., Schulze-Lefert, P., and Panstruga, R. (2005) Recruitment and interaction dynamics of plant penetration resistance components in a plasma membrane microdomain. Proc. Natl. Acad. Sci. U. S. A. 102, 3135–3140

106. Kluge, C., Seidel, T., Bolte, S., Sharma, S. S., Hametzik, M., Satiat-Jeunemaitre, B., Ross, J., Sauer, M., Goldrick, D., and Dietz, K. J. (2004) Subcellular distribution of the V-ATPase complex in plant cells, and in vivo localisation of the 100 kDa subunit VHA-a within the complex. Plant Physiol. 134, 1156–1168

107. Subramanian, C., Xu, Y., Johnson, C. H., and von Arnim, A. G. (2004) In vivo detection of protein-protein interaction in plant cells using BRET. Methods Mol. Biol. 284, 271–286

108. Gavin, A. C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J. M., Michon, A. M., Cruciat, C. M., Remor, M., Hofert, C., Schedler, M., Brajenovic, M., Rufthner, H., Merino, A., Kleijn, K., Dudak, M., Dickson, D., Rudi, T., Grau, V., Bauch, A., Bastuck, S., Huser, B., Leutheu, C., Heurtier, M. A., Copley, R. R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G., and Superti-Furga, G. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature 415, 141–147
(2002) Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. *Nature* 415, 180–183

121. Butland, G., Peregrin-Alvarez, J. M., Li, J., Yang, W., Yang, X., Canadien, V., Starostine, A., Richards, D., Beattie, B., Krogan, N., Davey, M., Parkinson, J., Greenblatt, J., and Emili, A. (2005) Interaction network containing conserved and essential protein complexes in Escherichia coli. *Nature* 433, 531–537

122. Graciet, E., Lebreton, S., and Gontero, B. (2004) Emergence of new regulatory mechanisms in the Benson-Calvin pathway via protein-protein interactions: a glyceraldehyde-3-phosphate dehydrogenase/CP12/phosphoribulokinase complex. *J. Exp. Bot.* 55, 1245–1254

123. Fernie, A. R., Geigenberger, P., and Stitt, M. (2005) Flux an important, but neglected, component of functional genomics. *Curr. Opin. Plant Biol.* 8, 174–182

124. Olsen, J. G., Kadziola, A., von Wettstein-Knowles, P., Siggaard-Andersen, M., and Larsen, S. (2001) Structures of β-ketoacyl-acyl carrier protein synthase I complexed with fatty acids elucidate its catalytic machinery. *Structure (Camb.)* 9, 233–243

125. Meister, A. (1989) Mechanism and regulation of the glutamine-dependent carbamyl phosphate synthetase of Escherichia coli. *Adv. Enzymol. Relat. Areas Mol. Biol.* 62, 315–374

126. Martins, B. M., Grimm, B., Mock, H. P., Richter, G., Huber, R., and Messerschmidt, A. (2001) Tobacco uroporphyrinogen-III decarboxylase: characterization, crystallization and preliminary X-ray analysis. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 57, 1709–1711

127. Ratinaud, M. H., Thomes, J. C., and Julien, R. (1983) Glutaryl-tRNA synthetases from wheat. Isolation and characterization of three dimeric enzymes. *Eur. J. Biochem.* 135, 471–477

128. Fujita, N., Kubo, A., Francisco, P. B., Jr., Nakakita, M., Harada, K., Minaka, N., and Nakamura, Y. (1999) Purification, characterization, and cDNA structure of isoamylase from developing endosperm of rice. *Planta* 208, 283–293

129. Balmer, Y., Koller, A., del Val, G., Manieri, W., Schurmann, P., and Buchanan, B. B. (2003) Proteomics gives insight into the regulatory function of chloroplast thioredoxins. *Proc. Natl. Acad. Sci. U. S. A.* 100, 370–375

130. Collin, V., Issakidis-Bourguet, E., Marchand, C., Hirasawa, M., Lancelin, J. M., Knaff, D. B., and Migniac-Maslow, M. (2003) The Arabidopsis plastidial thioredoxins: new functions and new insights into specificity. *J. Biol. Chem.* 278, 23747–23752

131. Aseeva, E., Ossenbuhl, F., Eichacker, L. A., Wanner, G., Soll, J., and Voitknecht, U. C. (2004) Complex formation of Vipp1 depends on its α-helical PsbA-like domain. *J. Biol. Chem.* 279, 35535–35541

132. Bonk, M., Tadros, M., Vandekerckhove, J., Al-Babili, S., and Beyer, P. (1996) Purification and characterization of chaperonin 60 and heat-shock protein 70 from chromoplasts of Narcissus pseudonarcissus. *Plant Physiol.* 111, 931–939

133. Sharkia, R., Bonshtien, A. L., Mizrahi, I., Weiss, C., Niv, A., Lustig, A., Vitlanen, P. V., and Azem, A. (2003) On the oligomeric state of chloroplast chaperonin 10 and chaperonin 20. *Biochim. Biophys. Acta* 1651, 76–84

134. Baneyx, F., Bertsch, U., Kalbach, C. E., van der Vies, S. M., Soll, J., and Gatenby, A. A. (1995) Spinach chloroplast cpn21 co-chaperonin possesses two functional domains fused together in a toroidal structure and exhibits nucleotide-dependent binding to plastid chaperonin 60. *J. Biol. Chem.* 270, 10695–10702