Proteome Analysis of Bell Pepper (Capsicum annuum L.) Chromoplasts

Muhammad Asim Siddique, Jonas Grossmann, Wilhelm Gruissem and Sacha Baginsky *

Institute of Plant Sciences, ETH Zürich, Universitätstrasse 2, CH-8092 Zürich, Switzerland

We report a comprehensive proteome analysis of chromoplasts from bell pepper (Capsicum annuum L.). The combination of a novel strategy for database-independent detection of proteins from tandem mass spectrometry (MS/MS) data with standard database searches allowed us to identify 151 proteins with a high level of confidence. These include several well-known plastid proteins but also novel proteins that were not previously reported from other plastid proteome studies. The majority of the identified proteins are active in plastid carbohydrate and amino acid metabolism. Among the most abundant individual proteins are capsanthin/capsorubin synthase and fibrillin, which are involved in the synthesis and storage of carotenoids that accumulate to high levels in chromoplasts. The relative abundances of the identified chromoplast proteins differ remarkably compared with their abundances in other plastid types, suggesting a chromoplast-specific metabolic network. Our results provide an overview of the major metabolic pathways active in chromoplasts and extend existing knowledge about prevalent metabolic activities of different plastid types.

Keywords: Biosynthesis — Carotenoid — Chromoplast — Mass spectrometry — Metabolism — Proteomics.

Abbreviations: CCS, capsanthin/capsorubin synthase; DTT, dithiothreitol; MS/MS, tandem mass spectrometry; NSI, nanospray ionization; OPPP, oxidative branch of the pentose phosphate pathway; RP-LC, reverse phase liquid chromatography; SQD1, UDP sulfoquinovose synthase.

Introduction

Plastids are characteristic plant cell organelles that perform essential biosynthetic and metabolic functions. These include photosynthetic carbon fixation, and the synthesis of amino acids, fatty acids, starch and secondary metabolites (Neuhaus and Emes 2000, Lopez-Juez and Pyke 2005). Plastids develop and differentiate into specialized plastid types that can be distinguished by their structure, pigment composition (color) and functional properties. Examples of specialized plastid types include chromoplasts in fruits and petals, amyloplasts in roots and storage tissues, and chloroplasts in photosynthetically active leaf tissues (Neuhaus and Emes 2000, Lopez-Juez and Pyke 2005).

Based on their energy metabolism, plastids can be distinguished as photosynthetic (autotrophic) and non-photosynthetic (heterotrophic) plastid types. Photosynthetic chloroplasts synthesize sugar phosphates that are metabolized by oxidative metabolism to NADPH and ATP. Non-photosynthetic plastid types import sugar phosphates and ATP from the cytosol, which is necessary to sustain their anabolic metabolism.

Although plastids are of significant biological interest (Gewolb 2002), our current understanding of the metabolic functions and capacities of different plastid types is still limited. Proteomics is a powerful approach to map the complete set of plastid proteins and to infer plastid-type specific metabolic functions. Several proteomics studies have been reported from chloroplasts that provide valuable information about their metabolic capacities (Ferro et al. 2002, Peltier et al. 2002, Schubert et al. 2002, Ferro et al. 2003, Froehlich et al. 2003, Kleffmann et al. 2004, Friso et al. 2004, Phinney and Thelen 2005, Majeran et al. 2005, Peltier et al. 2006, reviewed in Baginsky and Gruissem 2004, Baginsky and Gruissem 2006). Chloroplast proteome analysis is nearing saturation, however, because the detection of new proteins is constrained by highly abundant photosynthetic proteins that dominate the proteome of photosynthetically active chloroplasts. To circumvent this constraint and to increase proteome coverage, different plastid types can be used for high-throughput protein identification. For example, heterotrophic plastids do not contain highly abundant photosynthetic proteins and therefore increase the chance of detecting other metabolic activities and regulatory factors. At the same time, the proteomes of different plastid types contain valuable information about plastid type-specific activities.

Only a few proteomics approaches using heterotrophic plastid types have been reported, including rice etioplasts (von Zychlinski et al. 2005), wheat amyloplasts (Andon et al. 2002, Balmer et al. 2006) and BY-2 cell culture plastids (Baginsky et al. 2004). A comparison of proteins identified from different plastid types confirmed that the proteomes of heterotrophic and autotrophic plastids differ considerably, especially in their distinct energy metabolism. Heterotrophic plastids import metabolites such as ATP and glucose 6-phosphate to feed energy into their essential biosynthetic activities, for example the synthesis of fatty acids from acetate or amino acids.

* Corresponding author: E-mail, sbaginsky@ethz.ch; Fax, +41-1-632-10-79.
Chromoplasts proteomics

Results and Discussion

Isolation and purification of bell pepper chromoplasts

Bell pepper chromoplasts were isolated by Percoll density gradient centrifugation using a step gradient developed with 10–20, 20–30, 30–40, 40–50 and 60% Percoll in GR buffer. During centrifugation, chromoplasts migrate to different Percoll densities in the step gradient (Fig. 1A). In order to assess the relative amount of chromoplasts in each fraction compared with other cell organelles, we isolated material from bands A, B, C, D, E and F and identified the major protein constituents in each band by shotgun MS/MS as described (Materials and Methods) (Fig. 1B). Identified proteins were assigned to a cell organelle on the basis of targeting prediction and literature searches. We found the majority of plastid proteins in band D, which had the highest ratio of peptides detected from plastid proteins and those of other cell organelles. Pt, plastid; Mit, mitochondria; Nu, nucleus; Ot, other.

Fig. 1 Isolation and purification of C. annuum chromoplasts. (A) Chromoplasts were isolated by Percoll density gradient centrifugation as described in Materials and Methods (60–50, 50–40, 40–30, 30–20%). (B) Organelles were isolated from each fraction at the interphases between two density steps, subjected to SDS–PAGE and tryptic digest, and subsequently analyzed by LC-ESI-MS/MS. Chromoplasts in band D had the best ratio of peptides detected from plastid proteins and those of other cell organelles. Pt, plastid; Mit, mitochondria; Nu, nucleus; Ot, other. (C) Light microscopy of organelles isolated from band D revealed chromoplasts of a size of 3–4 µm.
preparation of mostly intact chromoplasts with only minor contaminations from other cell organelles (Fig. 1C).

**Fractionation of chromoplast proteins**

Based on the purity and intactness of chromoplasts in band D (30–40% Percoll interphase), we modified the Percoll density gradient centrifugation protocol in order to yield higher purity chromoplasts and to scale up the isolation procedure. We omitted the 60% Percoll cushion at the bottom of the gradient and instead introduced an additional 15% Percoll step to increase gradient resolution (Fig. 2A). After chromoplast isolation, we devised a serial fractionation procedure to extract and fractionate chromoplast proteins on the basis of their differential solubility. The chromoplast protein extract was fractionated into soluble proteins (OSMO), peripheral membrane proteins (8 M urea) and integral membrane proteins (5% SDS) (Fig. 2B). Soluble proteins made up 76% and peripheral membrane proteins 24% of the total proteins in the extract (Fig. 2C). These values indicated that the isolated chromoplasts contained mostly soluble proteins, suggesting that the organelles were intact after isolation because soluble proteins would leak from broken organelles. Only one major protein was visible in the SDS fraction upon silver staining (Fig. 2A). This protein was identified as capsanthin/capsorubin synthase (CCS), which is involved in the carotenoid biosynthesis pathway (Supplementary Table 2). The small number of integral membrane proteins that were solubilized with 5% SDS suggested that the chloroplast thylakoid membranes had been disassembled and thylakoid membrane proteins degraded during chromoplast development. Thylakoid membrane proteins represent the major protein mass in chloroplasts and require SDS for their solubilization (Kleffmann et al. 2004).

Protein fractions from each serial extraction step were further separated by SDS–PAGE. The gel was cut into 12 pieces, each of which was subsequently subjected to in-gel tryptic digest. The complex peptide mixtures were further fractionated by C18 reversed phase chromatography that was coupled online to an LCQDecaXP ion trap mass spectrometer (LCQDecaXP, Finnigan, San Jose, CA, USA). MS/MS data were first searched against the Viridiplantae subsection of the NCBI non-redundant protein database. Altogether, we identified 151 distinct plastid proteins and 20 putative contaminants (Supplementary Table 2). All identified proteins were manually sorted according to their putative functions. Supplementary Table 2 lists the proteins together with the serial extraction step from which they originated. Most of the proteins were identified in the soluble protein fraction and from the unfractionated protein mixture (86 and 49 proteins, respectively). Abundant proteins were detected in several distinct fractions. Most of the proteins identified in the OSMO fraction are active in carbohydrate and amino acid metabolism (Supplementary Table 2). The urea and SDS fractions contained proteases and some potential transporters, which suggests that these proteins are associated with the plastid membrane system (Supplementary Table 2). The abundant proteins in the SDS fraction are CCS, lethal leaf spot 1-like protein and transketolase 1 (Supplementary Table 2).

**Database-independent identification of chromoplast proteins**

We expected that a database-dependent search strategy would not detect a significant number of chromoplast proteins because bell pepper does not have a large number of protein entries. In order to assess how many peptides were not detected in a standard database search and the proteins they might comprise, we developed a database-independent approach for peptide identification. This approach first used a database-independent MS/MS spectrum quality scoring to identify peptide-derived spectra that were not identified in a standard database.
Several of the proteins that were identified with homology were subsequently scored for their quality, and high quality spectra were then subjected to de novo peptide sequencing. Sequences extracted from MS/MS spectra were submitted to MS BLAST, a tool that allows identification of a peptide by homology. We provide the numbers of spectra that were processed at each step in the workflow.

Functional classification of the identified proteins

Chromoplast development is accompanied by the massive synthesis and accumulation of carotenoids. In pepper, most of these pigments (>95%) accumulate in lipoprotein fibrils, which are supramolecular structures that contain carotenoids, some lipids and fibrillin, a protein that is the main structural component of the fibrillin complex (Deruere et al. 1994). Fibrillin (annotated as plastoglobule-associated protein) is one of the most abundant proteins in bell pepper chromoplasts, since we detected 210 peptides in different protein fractions (Supplementary Table 2). We identified >60% of the enzymes in the carotenoid biosynthesis pathway, suggesting that it is the primary metabolic pathway in pepper chromoplasts (Supplementary Table 2, Fig. 4). CCS was among the most abundant chromoplast enzymes (Table 1). Although several of the identified peptides did not allow CCS to be distinguished from neoxanthin synthase and lycopene β-cyclase, a substantial number of identified peptides unambiguously identified CCS. CCS is a bifunctional enzyme that converts antheraxanthin into capsanthin and violaxanthin into capsorubin (reviewed in Camara et al. 1995, Hirschberg 2001). A chromoplast-specific lycopene β-cyclase (Q9VF32) and a neoxanthin synthase (Q9M424) gave significant peptide alignments in the MS BLAST search but, due to their sequence similarity with CCS, we cannot substantiate their detection because of the error-tolerant sequence alignment (data not shown). Lycopene β-cyclase catalyzes a two-step reaction which creates one β-ionone ring at each end of the lycopene molecule to produce β-carotene, which is a precursor for the synthesis of capsanthin and capsorubin. It is therefore likely that this enzyme is present in the chromoplast extracts. We have not detected lycopene ε-cyclase, which is predominant in chloroplasts but not in chromoplasts (Galpaz et al. 2006).

The largest group of identified proteins (25 proteins, 580 identified peptides) is involved in carbohydrate metabolism (Fig. 5). We identified 208 peptides from transketolase, 67 peptides from transaldolase, 104 peptides from the plastid phosphoglycerate kinase, 32 peptides from glyceraldehyde-3-phosphate dehydrogenase and 11 peptides from triose phosphate isomerase (Supplementary Table 2).
Isolated bell pepper chromoplasts have high hexose kinase activity, indicating that chromoplast carbohydrate metabolism relies to some extent on import of hexoses that may be converted into hexose phosphates after import (Camara et al. 1995). Our data support this view because we detected fructokinase and a hexose transporter in the chromoplast preparation (Supplementary Table 2). The latter was originally identified as a hexose exporter, but recent data suggest an additional function in the import of hexoses from the cytosol (Butowt et al. 2003). We also identified glucose-6-phosphate isomerase, which enables chromoplasts to convert fructose-6-phosphate into glucose-6-phosphate to feed the OPPP. The OPPP is the major route of NADP reduction in heterotrophic plastid types and generates precursors for, for example, the shikimate pathway and nucleotide metabolism (reviewed in Neuhaus and Emes 2000). We did not detect glucose-6-phosphate dehydrogenase but found 6-phosphogluconate dehydrogenase from the oxidative branch of the OPPP. In addition to a hexose transporter, we also identified a triose phosphate/phosphate translocator, suggesting that both hexoses and triose phosphates are imported into chromoplasts. In contrast, in other heterotrophic plastid types such as etioplasts and BY2 plastids, we identified either a triose Table 1 Proteins from bell pepper chromoplasts that were not identified in a standard database search but with the modified strategy described here

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Annotation</th>
<th>Protein score</th>
<th>Peptide score</th>
<th>Query sequence</th>
<th>Retrieved protein</th>
<th>Organelle</th>
<th>In plprot</th>
</tr>
</thead>
<tbody>
<tr>
<td>O22774</td>
<td>Chloroplast outer envelope 86-like</td>
<td>314</td>
<td>105</td>
<td>SHVVQQSL GQAVGDLR</td>
<td>SHIVQVSIG QAVGDLR</td>
<td>Chloro Y</td>
<td></td>
</tr>
<tr>
<td>Q8LB69</td>
<td>Thylakoid formation 1</td>
<td>77</td>
<td>77</td>
<td>TVLQEL LVQH</td>
<td>TVLQELIVQQH</td>
<td>Chloro Y</td>
<td></td>
</tr>
<tr>
<td>O24024</td>
<td>Plastid-lipid-associated protein</td>
<td>352</td>
<td>68</td>
<td>ADSFYGTNR</td>
<td>ADSFYGTNR</td>
<td>Chloro Y</td>
<td></td>
</tr>
<tr>
<td>Q9SUJ7</td>
<td>Copper/zinc superoxide dismutase</td>
<td>83</td>
<td>83</td>
<td>BXXVEL LTIQESK</td>
<td>RAEIVELI TQLESK</td>
<td>Chloro Y</td>
<td></td>
</tr>
<tr>
<td>Q9C9T</td>
<td>Hypothetical protein</td>
<td>96</td>
<td>96</td>
<td>BSPAEAGY SEGGLNA</td>
<td>RSPAEAGY SEGGLNA</td>
<td>Chloro Y</td>
<td></td>
</tr>
<tr>
<td>NP_179318</td>
<td>Homoserine kinase</td>
<td>98</td>
<td>98</td>
<td>ANLGPNLDM LGCAVDGGLDG</td>
<td>ANLGPGFDGFL GCACVDGGLG</td>
<td>Chloro Y</td>
<td></td>
</tr>
<tr>
<td>Q8RU74</td>
<td>Dehydroquinate synthase</td>
<td>87</td>
<td>87</td>
<td>BAPSQAPTVVE VDLPT</td>
<td>KASSQAPTV VEVDLGT</td>
<td>Chloro Y</td>
<td></td>
</tr>
<tr>
<td>Q94F93</td>
<td>Putative 3-ketoacyl-ACP dehydratase</td>
<td>87</td>
<td>87</td>
<td>BYPAMPTVMDL NQLR</td>
<td>RYEAFPTV MDINQIR</td>
<td>Chloro Y</td>
<td></td>
</tr>
</tbody>
</table>

Protein identifications were accepted if at least one individual peptide alignment exceeded a threshold score of 75 and if the protein derived from higher plants. Column 'in plprot' indicates whether the protein was identified in previous plastid proteome analyses (Kleffmann et al. 2006). The column 'organelle' refers to the targeting prediction by TargetP (Emanuelsson et al. 2000).
phosphate/phosphate translocator (von Zychlinski et al. 2005) or a hexose phosphate transporter (Baginsky et al. 2004), but not both.

During chromoplast development, the photosynthetic energy metabolism is changed into a non-photosynthetic, heterotrophic energy metabolism, which is accompanied by the degradation of abundant photosynthetic proteins. For example, the total activities of Rubisco and phosphoribulokinase, two key enzymes of the Calvin cycle, decrease by 80% during chromoplast development.
We did not detect phosphoribulokinase but high amounts of Rubisco [174 (LSU) and 31 (SSU) detected peptides] and ribulose phosphate 3-epimerase (10 detected peptides), indicating that the degradation of Calvin cycle enzymes was not fully completed (Supplementary Table 2) (reviewed in Camara et al. 1995). We found eight proteins with a potential function in protein degradation, four of which belong to the family of Clp proteases, which may be involved in the degradation of chloroplast proteins that are not required in chromoplasts (Supplementary Table 2) (Sakamoto 2006). The other proteases included a metalloprotease of unknown function, the FtsZ protease that is involved in plastid division, and the FtsH protease that degrades the PSII reaction center protein D1 during moderate heat stress (Sakamoto et al. 2006, Yoshioika et al. 2006).

In addition to the down-regulation of photosynthetic functions, enzymes that synthesize amino acids decrease in relative abundance during chromoplast development (Gallardo et al. 1988, Camara et al. 1995). It can be expected that degradation of chloroplast proteins fills the plastid amino acid reservoir, which would make de novo synthesis of amino acids during the transition phase unnecessary. However, recent reports and our results suggest that not all amino acid-synthesizing enzymes are down-regulated during chromoplast development. For example, cysteine synthase accumulates to relatively high levels in pepper chromoplasts (Roemer et al. 1992), a finding that is supported by our data (Supplementary Table 2). Additionally, we identified a substantial number of peptides from aspartate aminotransferase (39 detected peptides) and ferredoxin-dependent glutamate synthase (Fd-GOGAT, 64 detected peptides) (Fig. 5, Supplementary Table 2). Work in tomato showed that Fd-GOGAT is an abundant enzyme in green and in ripe fruits (Gallardo et al. 1993) and presumably responsible for the re-assimilation of nitrogen after protein and chlorophyll degradation (Gallardo et al. 1993). Together, our results suggest that down-regulation of amino acid biosynthesis follows a coordinated program. The important question remains why some amino acid-synthesizing enzymes are retained longer than others during chromoplast development.

Although chromoplasts actively transcribe plastid-encoded genes, we identified only four proteins with a function in plastid gene expression. The low coverage of the gene expression system can be explained by the low abundance of its constituents. This has been reported for most plastid types analyzed so far and may be especially critical for fully differentiated chromoplasts (Kleffmann et al. 2006). The detection of elongation factor Tu is not surprising because it is one of the most abundant plastid proteins and has been identified in all plastid types analyzed to date (Kleffmann et al. 2006). More surprisingly, we detected the ‘translation inhibitor protein’, suggesting that it may have a specific function in chromoplasts. ‘Translation inhibitor protein’ is a homolog of endoribonuclease L-PSP, which was first identified as a single-stranded RNA-specific endonuclease in rat (Oka et al. 1995). The L-PSP homolog could be one of the endonucleases that degrade mRNAs from plastid-encoded photosynthetic proteins that are transcribed but not required in fully developed chromoplasts. A dedicated analysis of the plastid mRNA degradation system is necessary to shed light on the exact molecular function of the L-PSP homolog in plastids.

Comparison of plastid proteomes

In order to identify proteins that are common to all plastid types and those that are specific for chromoplasts, we BLAST-searched all 151 identified chromoplast proteins against 2,650 experimentally verified plastid
proteins that are available in plastids (Kleffmann et al. 2004). Nine chromoplast proteins were not previously reported from proteome analyses with other plastid types. Among them is the plastid ‘stay green protein 1’, which is also induced during senescence and potentially involved in Chl catabolism or protein degradation. Its exact role, however, is unknown. Other enzymes that until now were exclusively identified in chromoplasts include β-carotene hydroxylase and CCS, which are both active in the carotenoid biosynthesis pathway. We furthermore detected the sulfolipid-synthesizing enzyme UDP sulfoquinovose synthase (SQD1), which catalyzes the transfer of sulfite to UDP-glucose. Recent data suggested that SQD1 forms a stable complex with Fd-GOGAT via the FMN-binding domain of Fd-GOGAT, which led the authors to propose that Fd-GOGAT channels sulfite to SQD1 (Shimojina et al. 2005) (Supplementary Table 2).

All other chromoplast proteins were already identified in one or several other plastid types (Fig. 6). Seventy-nine of the proteins were identified in all four plastid types analyzed to date, suggesting that they constitute a plastid core proteome for housekeeping functions, such as amino acid and carbohydrate metabolism, as well as protein metabolism (Supplementary Table 3). We compared the peptide numbers for the five most abundant chromoplast proteins represented in the core proteome with their numbers found in proteome studies of other plastid types. The five proteins were glyceraldehyde 3-phosphate dehydrogenase, transketolase (carbohydrate), Clp protease (proteases), elongation factor Tu (gene expression) and Fd-dependent glutamate synthase (amino acid synthesis). These key proteins are highly abundant in all four plastid types. Based on the number of detected peptides, transketolase is one of the most abundant enzymes common to all different plastid types analyzed to date (Fig. 6) (Kleffmann et al. 2006). Transketolase transfers keto-groups between sugars and as such enables cross-talk between anabolic and catabolic plastid carbohydrate metabolism.

In conclusion, the chromoplast proteome reported here is characterized by a prevalence of carbohydrate-metabolizing activities and a predominant carotenoid biosynthetic pathway. Except for nine proteins, all other identified proteins were already reported to be present in other plastid types. Nonetheless, the analysis of the chromoplast proteome revealed some remarkable differences in the relative abundances of the identified proteins compared with other plastid types, supporting the view that a plastid-specific metabolic network exists in plants. Further analyses of metabolite fluxes within these networks should clarify their activity and functions, which will provide new insights into plastid functions in specialized plant cells.

Materials and Methods

Isolation of C. annuum chromoplasts

Red fruits (~230 g) of Capsicum annuum L were bought from a local market. Chromoplasts were isolated as described previously by Hadjeb et al. (1988) with minor modifications, e.g. using Percoll as the separation medium for density gradient centrifugation. The fruit was cut into small pieces of about 1 cm² and blended in ice-cold GR buffer (1 mM NaP₂O₇, 50 mM HEPES, 330 mM sorbitol, 2 mM EDTA, 1 mM MgCl₂, 2 mM MnCl₂, 2 mM dithiothreitol (DTT), pH 6.8) using five bursts of 5–10 s each with a common Waring blender. The solution was subsequently filtered through four layers of Miracloth and the filtrate was subsequently centrifuged at 800 g for 10 min. The supernatant was discarded and the pellet was washed three times with GR buffer. The pellet was resuspended in 2 ml of a GR mix (1 mM NaP₂O₇, 50 mM HEPES, 330 mM sorbitol, 2 mM EDTA, 1 mM MgCl₂, 2 mM
MnCl₂, 2 mM DTT, pH 6.8) and loaded on top of a Percoll density gradient developed with 2.5 ml each of 20–30, 30–40, 40–50 and 50–60% Percoll in 1× GR medium and centrifuged at 7,000×g for 30 min. Six chromoplast-containing bands were recovered at the interphases of 60–50%, 50–40%, 40–30%, 30–20%, 20–upper side and from the top of the gradient (supernatant). These bands were designated as A, B, C, D, E and F, and the material therein was washed twice in GR buffer and subsequently stored at −80°C for further analyses. For the characterization of chromoplasts, the recovered bands (A, B, C, D, E and F) were examined by light microscopy using an AxioPlan 2 microscope (Zeiss, Wetzikon, Switzerland) to check their purity and intactness. Band D contained the highest amount of plastids as compared with the other bands.

Fractionation of proteins

With the isolated chromoplasts, we employed a multidimensional protein fractionation strategy to increase the dynamic range of the proteome analysis. A serial extraction procedure was used for protein fractions as described previously, but with minor modifications (Baginsky et al. 2004, Kleffmann et al. 2004). The proteins were solubilized using three different buffer compositions for soluble proteins (OSMOP), peripheral membrane proteins (8 M urea) and integral membrane proteins (5% SDS). The buffer compositions for each step were as follows: OSMO—40 mM Tris–HCl pH 8.0, 5 mM MgCl₂, 1 mM DTT and 2× protease inhibitor cocktail (Roche Diagnostics GmbH, Germany, 2× the supplier’s recommended concentration); 8 M urea—8 M urea, 20 mM Tris-base, 5 mM MgCl₂, 20 mM DTT and 2× protease inhibitor; 5% SDS—40 mM Tris base, 5% SDS, 40 mM DTT and 2× protease inhibitor. We also analyzed the complete mixture of proteins (crude) in order to increase the amount of chromoplast proteins that can be detected. The complete chromoplast protein mixture (crude) derived from resuspension of isolated chromoplasts in 10× Laemmli buffer (0.625 M Tris, 10% SDS, 20% β-mercaptoethanol, 10% glycerin, 0.5% bromophenolblue pH 6.8). Proteins from all fractions were further fractionated by SDS-PAGE (Laemmli 1970).

In-gel digestion and mass spectrometric analysis

Prior to mass spectrometric protein identification, all protein samples were directly subjected to SDS-PAGE by loading them onto 10 cm long homogeneous 10% polyacrylamide gels (Laemmli 1970). After electrophoresis, the gels were cut into 12 pieces. Proteins in each section were immediately subjected to in-gel tryptic digest according to the protocol described by Shevchenko et al. (1996). Before mass spectrometric peptide analysis, all tryptic peptides were separated by reversed phase-liquid chromatography (RP-LC) on a C18 matrix. Peptides were separated on laboratory-made capillary columns with an inner diameter of 75 μm (length 8 cm, BGB Analytik AG, Böckten, Switzerland) for nanospray ionization (NSI) packed with C18 reversed phase material (Magic C₁₈ resins; 5 μm, 200 A pore; Michrom BioResources, Auburn, CA, USA). The peptide mixture was resolved in buffer A (5% acetonitrile, 0.5% formic acid), centrifuged at 16,000×g to pellet small remaining gel pieces and subsequently cleaned with ZipTips. The samples were loaded manually or with an auto sampler onto the RP column and the bound peptides were eluted with an increasing concentration of acetonitrile from 5 to 80% acetonitrile in the mobile phase in 2 h at a flow rate of 300 nL min⁻¹. RP-LC was coupled online to an LCQDecaXP ion trap mass spectrometer (Thermo Finnigan, San Jose, CA, USA) equipped with an NSI source. The mass spectrometric analysis was performed in the positive ion mode and peptides were ionized with a spray voltage of 2.0–2.8 kV. Data were acquired in the data-dependent scan mode with one full scan (MS) and four subsequent data-dependent MS/MS scans with the four most intense ions as determined in the MS scan. Dynamic exclusion was enabled to allow the MS/MS measurement of the same parent mass twice and excluding it thereafter for 1 min.

Analysis and interpretation of mass spectrometric data

MS/MS data were interpreted according to the standards put forward by Carr et al. (2004). MS/MS data were analyzed by the SEQUEST software using the Viridiplantae subsection of the NCBI non-redundant protein database (http://www.ncbi.nlm.nih.gov/) (March 200 version 5). All RAW files were pooled and subjected to statistical analysis with Peptide- and ProteinProphet (Nevižhskii et al. 2003). ‘.dta’ files were created by the SEQUEST software for every MS/MS scan with a total ion count of at least 5 × 10⁴, minimal peak count of 35, and a precursor ion mass in the range of 300–2000 m/z. Data were searched against the above database (restricted to tryptic peptides) without modifications, allowing a parent mass error tolerance of 2 amu and daughter ion error tolerance of 0.8 amu.

Database-independent protein identification

We first analyzed the mass spectrometric data from 83 LCQ analytical runs altogether to identify those spectra that were left unassigned by the standard database search described above despite originating from true peptide fragmentation. Collision-induced peptide dissociation (CID) generates spectra with a set of well defined characteristics that distinguish them from low-quality noise spectra or non-protein contaminant fragmentation. For this purpose, we employed QUALSCORE, a tool that was developed in house (Nevižhskii et al. 2006). We analyzed 83 mass spectrometer runs that altogether gave rise to 302,430 spectra, 4,787 of which could be assigned to a peptide by a standard database search. Of the remaining spectra that were left unassigned, 8,666 were rated as high quality (score >0). These spectra were subsequently submitted to PepNovo (Frank and Pevzner 2005) to extract an amino acid sequence exclusively from the information contained in the MS/MS spectra using PepNovo default parameters (http://peptide.ucsd.edu/pepnovo.py). Only those PepNovo results were accepted that received a mean reliability score of at least 0.5. These sequences were further processed to make them compatible with MS BLAST searches.

An additional amino acid letter B was added to the N-terminus of the peptide (B in MS BLAST can be either lysine or arginine). Gaps were filled with X, XX, XXX or XXXX (X in MS BLAST can be any amino acid) according to the assumed gap length. In cases where a gap was difficult to define, the sequence was written in all possible combinations. The queries were then loaded into MS BLAST (Shevchenko et al. 2001) via the web interface (http://dove.embl-heidelberg.de/Blast2/msblast.html) and searched against the NCBI non-redundant protein database using MS BLAST default parameters. MS BLAST results were carefully filtered manually and only those identifications that derived from higher plants were accepted that exceeded a threshold of 75. Most of the identified proteins were already identified during the database search and essentially reproduced the database findings. Additionally, a significant number of trypsin and keratin spectra were identified that escaped the original database detection. Altogether, eight chromoplast proteins were identified anew (see Results).
Supplementary material

Supplementary material mentioned in the article is available to online subscribers at the journal website www.pcp.oxfordjournals.org.

Acknowledgments

The authors thank the Functional Genomics Center Zürich for assistance, and Dr. Johannes Fütterer for critical comments. We are grateful for the generous support of the VELUX foundation to A.S. and for ETH funds that were provided to W.G. and S.B.

References


(Received October 16, 2006; Accepted November 8, 2006)