The proteome map of spinach leaf peroxisomes indicates partial compartmentalization of phylloquinone (vitamin K1) biosynthesis in plant peroxisomes

Lavanya Babujee1,†, Virginie Wurtz2, Changle Ma1,‡, Franziska Lueder1,§, Pradeep Soni3, Alain van Dorselaer2 and Sigrun Reumann1,3,*

1 Georg-August-University of Goettingen, Albrecht-von-Haller-Institute for Plant Sciences, Department of Plant Biochemistry, Justus-von-Liebig-Weg 11, D-37077 Goettingen, Germany
2 Laboratoire de Spectrométrie de Masse Bio-Organique, ECPM, UMR/CNRS 7178, Institut Pluridisciplinaire Hubert CURIE, Universite Louis Pasteur Strasbourg, France
3 Centre for Organelle Research, Faculty of Science and Technology, University of Stavanger, N-4036 Stavanger, Norway

† Present address: Department of Plant Pathology, University of Wisconsin, Madison, WI, USA.
‡ Present address: Section of Molecular Biology, Division of Biological Sciences, University of California, San Diego, La Jolla, CA, USA.
§ Present address: Department of Biochemistry and Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, 30 Flemington Road Parkville, Victoria 3010, Australia.
* To whom correspondence should be addressed. E-mail: sigrun.reumann@uis.no

Received 24 September 2009; Revised 16 December 2009; Accepted 14 January 2010

Abstract

Leaf peroxisomes are fragile, low-abundance plant cell organelles that are difficult to isolate from one of the few plant species whose nuclear genome has been sequenced. Leaf peroxisomes were enriched at high purity from spinach (Spinacia oleracea) and ~100 protein spots identified from 2-dimensional gels by a combination of liquid chromatography–tandem mass spectrometry (LC-MS/MS) and de novo sequencing. In addition to the predominant enzymes involved in photorespiration and detoxification, several minor enzymes were detected, underscoring the high sensitivity of the protein identification. The tryptic peptides of three unknown proteins shared high sequence similarity with Arabidopsis proteins that carry putative peroxisomal targeting signals type 1 or 2 (PTS1/2). The apparent Arabidopsis orthologues are a short-chain alcohol dehydrogenase (SDRa/IBR1, At4g05530, SRL> and two enoyl-CoA hydratases/isomerases (ECHIa, At4g16210, SKL>; NS/ECHId, At1g60550, RLx5HL). The peroxisomal localization of the three proteins was confirmed in vivo by tagging with enhanced yellow fluorescent protein (EYFP), and the targeting signals were identified. The single Arabidopsis isoform of naphthoate synthase (NS) is orthologous to MenB from cyanobacteria, which catalyses an essential reaction in phylloquinone biosynthesis, a pathway previously assumed to be entirely compartmentalized in plastids in higher plants. In an extension of a previous study, the present in vivo targeting data furthermore demonstrate that the enzyme upstream of NS, chloroplastic acyl-CoA activating enzyme isoform 14 (AAE14, SSL>, is dually targeted to both plastids and peroxisomes. This proteomic study, extended by in vivo subcellular localization analyses, indicates a novel function for plant peroxisomes in phylloquinone biosynthesis.

Key words: Acyl-activating enzyme, enoyl-CoA hydratase/isomerase, mass spectrometry, phylloquinone, proteome analysis, short-chain dehydrogenase, subcellular targeting.

Abbreviations: AAE14, acyl-CoA activating enzyme isoform 14; ASB-14, amidosulphobetaine-14; CaMV, cauliflower mosaic virus; CAT, catalase; CFP, cyan fluorescent protein; CHAPS, (3-[3-cholamidopropyl]-dimethylamino)propanesulphate; CHY1, hydroxisobutyryl-CoA hydrolase 1; 2-DE, 2-dimensional gel electrophoresis; DHNA-CoA, 1,4-dihydroxy-2-naphthoate-CoA; DTT, dithiothreitol; ECHI, enoyl-CoA hydratase/isomerase; EST, expressed sequence tag; EYFP, enhanced yellow fluorescent protein; GFP, green fluorescent protein; IEF, isoelectric focusing; IFM, isoelectric point; LC-MS/MS, liquid chromatography–tandem mass spectrometry; MALDI, matrix-assisted laser desorption/ionization; MFP, multifunctional protein; MS, mass spectrometry; NS, naphthoate synthase; PMF, peptide mass fingerprinting; PTD, peroxisome targeting domain; PTS1/2, peroxisomal targeting signal type 1/2; RFP, red fluorescent protein; ROS, reactive oxygen species; SDR, short-chain dehydrogenase/reductase; YFP, yellow fluorescent protein.
Introduction

Technical advances in 2-dimensional gel electrophoresis (2-DE), mass spectrometry, and shotgun proteomics laid the foundation for the first plant organellar proteome studies (Peltier et al., 2000; Kruft et al., 2001; Millar et al., 2001). The proteome in its entirety is the expressed complement of the genome and varies with tissue and organ, the developmental stage, and biotic and abiotic factors. Proteomics technology is, for instance, able to identify novel organellar proteins, characterize protein complexes, and detect post-translational protein modifications on a large scale, which is hardly possible with other methods. The bottleneck of organellar proteomics, however, is the high enrichment of the desired cell compartment.

Peroxisomes are ubiquitous cell organelles that compartmentalize oxidative metabolic reactions by a single boundary membrane. A characteristic property of peroxisomes is their functional flexibility, as their protein content varies depending on the organism, the type of tissue, and the environmental conditions. Peroxisomes are often part of distinct metabolic networks that are spread over different subcellular compartments, including the cytosol, mitochondria, and plastids in plants. For instance, the recycling of 2-phosphoglycolate during photorespiration involves the transmembrane transport steps for intermediate transfer between cell compartments (Reumann and Weber, 2006). Hydrogen peroxide produced by peroxisomal oxidases, as well as other reactive oxygen species (ROS), can be immediately degraded by catalase (CAT) and other peroxisomal antioxidative enzymes (del Rio et al., 2002).

Intensive research during recent years using complementary approaches, including forward and reverse genetics, has revealed an unexpected metabolic complexity of plant peroxisomes. For instance, plant peroxisomes play a physiological role in the biosynthesis of the signalling molecule jasmonic acid, β-oxidation of indole butyric acid (IBA), and sulphur and polyamine metabolism (for a review, see Kaur et al., 2009). Moreover, evidence is emerging from recent studies that peroxisomes have important functions in specific defence mechanisms, conferring resistance against pathogen attack (for a review, see Kaur et al., 2009). Additional physiological functions of plant peroxisomes are probably as yet uncovered.

Given their fragility in vitro, their limited abundance under standard growth conditions, and their pronounced physical association with mitochondria and chloroplasts, proteomic studies of peroxisomes have been reported only in the past few years, for example in Saccharomyces cerevisiae (Schafer et al., 2001; Marelli et al., 2004), in mammals (Kikuchi et al., 2004; Wiese et al., 2007), and in plants (for a review, see Saleem et al., 2006; Palma et al., 2009). Proteomic analyses of plant peroxisomes are made difficult by the fact that the genomes of only a few plant species have been sequenced to date, none of which represents established model organisms for the isolation of leaf peroxisomes. Also, large expressed sequence tag (EST) sequencing projects have not been realized for plant species such as Spinacia, from which leaf peroxisomes can be enriched in high quality and considerable quantity (Yu and Huang, 1986; Lopez-Huertas et al., 1995).

Complementarily or alternatively to experimental proteomic studies, peroxisomal matrix proteins can be predicted from genome sequences by taking advantage of the sequence conservation of matrix targeting signals (Emanuelsson et al., 2003; Kamada et al., 2003; Neuberger et al., 2003; Reumann, 2004; Reumann et al., 2004; Boden and Hawkins, 2005). Peroxisomal matrix proteins are synthesized on free ribosomes and generally targeted to the matrix by one of two conserved targeting sequences, namely the C-terminal peroxisomal targeting signal type 1 (PTS1) of the prototype SKL> (where > indicates the C-terminal end), or the cleavable PTS2 nonapeptide of the prototype RLx3HL (Gould et al., 1987; Swinkels et al., 1991). When these conserved targeting peptides are applied to genomic screens, most known and many novel matrix proteins can be identified. The AraPerox database (www3.uis.no/AraPeroxV1), for instance, currently lists ~400 candidate proteins that are potentially targeted to peroxisomes in Arabidopsis thaliana (Reumann et al., 2004, 2007). However, because these PTS peptides and auxiliary or even essential targeting elements surrounding the small targeting peptides have not yet been precisely defined, peroxisomal targeting cannot be predicted with sufficient accuracy, reiterating the need for experimental approaches to reveal the proteome of peroxisomes.

In the present study, a previously utilized method to purify leaf peroxisomes from Spinacia oleracea L. was extended by the addition of a second density gradient to allow for proteomic analysis. De novo sequencing of three unknown proteins permitted the identification of the corresponding Arabidopsis homologues. The Arabidopsis cDNAs were cloned by reverse transcription-PCR (RT-PCR), and peroxisomal localization was verified for fusion proteins with enhanced yellow fluorescent protein (EYFP) by fluorescence microscopy in epidermal cells of Allium cepa L. The peroxisomal localization of naphthoate synthase (NS), which catalyses an essential step in phyloquinone biosynthesis, prompted the investigation of the subcellular targeting of the enzyme immediately upstream in this pathway, chloroplastic acyl-CoA activating enzyme isoform 14 (AAE14), in greater detail. Upon N-terminal fusion of the protein with EYFP, placing the predicted PTS1 of AAE14, SSL>, at the free C-terminus, the fusion protein indeed targeted plant peroxisomes. The data thus indicate that at least two reactions of phyloquinone biosynthesis have been transferred from chloroplasts to peroxisomes during the evolution of higher plants.

Materials and methods

Plant growth

Spinach plants (S. oleracea L., cv. Monopa) were grown hydroponically in nutrient solution under a light intensity of ~300–350 µmol
m⁻² s⁻¹ (9/15 h light/dark cycle, 21 °C/18 °C day/night temperature) for up to 8 or 12 weeks (Reumann et al., 1994).

Isolation of leaf peroxisomes

The basic protocol described by Yu and Huang (1986) that involves a Percoll density gradient separation of sedimented leaf peroxisomes in 0.25 M sucrose was adapted for the large-scale isolation of leaf peroxisomes for preparative purposes (Reumann et al., 1995). The stability and purity of leaf peroxisomes was further improved by the addition of protease inhibitors, by increasing the concentration of Percoll from 45% to 48% (v/v), and by the addition of a second density gradient. After Percoll density gradient centrifugation and washing, the peroxisome suspension was homogenized using a tight-fitting Potter–Elvehjem homogenizer, layered on a discontinuous sucrose density gradient [small layers of 2 ml each of 18, 25, and 35% (w/w) sucrose over a linear 40–50% (w/w) gradient on a 60% (w/w) sucrose cushion (all in 10 mM HEPES-NaOH pH 7.5, 1 mM EDTA)], and spun for 2 h at 83,000 g (SW 28 rotor, Beckman Coulter ultracentrifuge). The peroxisome band, visible as a yellowish-white layer, was harvested, supplemented with protease inhibitors, and frozen at −80 °C until use.

2-DE and SDS-PAGE

Protein content was estimated using bovine serum albumin (BSA) as standard (Lowry et al., 1951; Bradford, 1976). Prior to 2-DE, proteins from peroxisomes were generally precipitated using methanol and chloroform (Wessel and Flugge, 1984). The proteins were solubilized for 2 h in 350 μl of immunoelectrofocusing (IEF) buffer [7 M urea, 2 M thiourea, 4% (w/v) CHAPS (3-[3-cholamidopropyl]dimethylammonio)-propanesulfate], 0.5% (v/v) IPG buffer, 3 mg ml⁻¹ dithiothreitol (DTT), and trace amounts of bromophenol blue, and the solution was clarified by centrifugation (20 min, 20,000 g). Immobile 18 cm pH gradient (IPG) strips were actively rehydrated with the protein solution, and IEF was carried out for 20 h for a total of 60 kVh (IPG Phor, APB, Sweden). The proteins were then reduced and alkylated by equilibration for 12 min each in buffers [6 M urea, 50 mM TRIS-HCl pH 8.8, 30% (v/v) glycerol, 2% (w/v) SDS, and 0.001% (w/v) bromophenol blue] containing either 10 mg ml⁻¹ DTT or 25 mg ml⁻¹ iodoacetamide, respectively. The equilibrated strips were positioned over vertical polyacrylamide gels of the desired concentration and sealed in place with 1% (w/v) agarose prepared in 4% (v/v) acetic acid. The flow rate was 5 ml h⁻¹. Post-electrophoretic staining of the gels was carried out according to described protocols using colloidal Coomassie solution (Herbert et al., 2001).

In-gel digestion for automated application

In-gel digestion with trypsin was performed according to published methods (Jeno et al., 1995) modified for use with a robotic digestion system (MassPREP, Micromass, Manchester, UK). Each gel slice was destained by washing with 50 μl of 25 mM ammonium hydrogen carbonate followed by washing with 50 μl of acetonitrile. This step was repeated twice. The gel pieces were then dehydrated with acetonitrile and dried at 60 °C, prior to the addition of modified trypsin (Promega, Madison, WI, USA; 10 μl at a concentration of 12.5 ng μl⁻¹ in 25 mM ammonium hydrogen carbonate). The digestion was performed at 37 °C overnight followed by elution in 5 μl of 30% (v/v) water/65% (v/v) acetonitrile/5% (v/v) formic acid.

ZipTip® Pipette Tips (Millipore Corporation) were used for the rapid purification and concentration of peptides prior to mass spectrometry. The C18 columns (ZipTips) were pre-wetted by washing (3 × 10 μl) with 10 μl of 50% (v/v) acetonitrile. The tips were equilibrated with 1% (v/v) formic acid (3 × 10 μl). Next, 10 μl of 1% formic acid was added to the peptide extract. The peptides were bound to the equilibrated C18 column by drawing the sample slowly through it. The columns were washed with the loading solvent (3 × 10 μl) to remove the non-peptide impurities. The desalted peptides were concentrated by eluting in 2–5 μl of 60% (v/v) acetonitrile/1% (v/v) formic acid.

Matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS)

Mass measurements were carried out on a Biflex (Bruker, Wissenburg, Germany) matrix-assisted laser desorption time-of-flight mass spectrometer (MALDI-TOF). A saturated solution of α-cyano-4-hydroxyquinamic acid in acetone was used as the matrix. The first layer of fine matrix crystals was obtained by spreading and fast evaporation of 0.5 μl of matrix solution. On this fine layer of crystals, a droplet of 0.5 μl of aqueous formic acid (5% v/v) was deposited. Afterwards, 0.5 μl of the peptide sample was added, and a second 0.25 μl droplet of saturated matrix solution (in 50% water/50% acetonitrile) was added. The preparation was dried under vacuum and washed in 1 μl of 5% (v/v) formic acid. All mass spectra were internally calibrated with trypsin autolysis peaks. The resulting peptide mass fingerprints (PMFs) were searched against a local copy of the non-redundant database SWISS-PROT (http://www.expasy.ch/prot) using the MASCOT (Perkins et al., 1999) search program. The parameters used in the search were as follows: peptide mass tolerance 50 ppm, one missed cleavage, carboxymethylated cysteine, methionine oxidation, and N-terminal acetylation.

Liquid chromatography–tandem mass spectrometry (LC-MS/MS)

Samples were injected into a CapLC System (Waters, Milford, MA, USA) equipped with an autosampler, gradient, and auxiliary pump. A volume of 6.4 μl was injected via ‘microliter pickup’ mode and desalted online through a 300 μm×5 mm C₁₈ trapping cartridge (LC Packings, San Francisco, CA, USA). The samples were desalted at a high flow rate of 30 μl min⁻¹ for 3 min. The peptides were separated on a 75 μm×15 cm, 3 μm particle size, 100 Å pore size C₁₈ PepMap™ column (LC Packings, CA, USA) prior to introduction into the mass spectrometer. A typical reversed-phase was used from low to high organic over ~35 min. Mobile phase A was 0.1% formic acid, and B was 95% acetonitrile and 0.1% formic acid. The flow rate was 5 μl min⁻¹. The system utilized a split flow, resulting in a column flow rate of ~400–500 μl min⁻¹. MS/MS data were obtained using a Q-ToF-2 (Micromass) fitted with a Z-spray nanoflow electrospray ion source. The mass spectrometer was operated in positive ion mode with a potential of 3500 V applied to the nanoflow probe body. The collision energy was determined ‘on the fly’ based on the mass and charge state of peptide. Charge state recognition was used to switch into MS/MS mode only for double- and triple-charged ions. Several trypsin autolysis ions were excluded. The data were processed by Protein Lynx Version (Micromass) to generate searchable peak lists. Initial protein identifications were made by the correlation of uninterpreted tandem mass spectra to entries in SWISS-PROT using Global Server (Version 1.1, Micromass).

Gene cloning by RT-PCR and subcellular localization analysis in Allium cepa L.

Total RNA was isolated from different tissues of A. thaliana cv. Columbia using the Invisorb Spin Plant Mini Kit (Invitek GmbH, Berlin, Germany). Full-length cDNAs of a short-chain dehydrogenase/reductase (SDRAs, At4g05530, 254 residues, SRL+) and two monofunctional enoyl-CoA hydratases/isomerases (ECH1a, At4g16210, 265 residues; SKL>, NS/ECHId, At1g60550, 337 residues, RLxJHL) were isolated from RNA using appropriate oligonucleotide primers (Supplementary Table SII available at JXB online). Total RNA was converted to single-stranded cDNA by reverse transcriptase (Superscript II, Invitrogen, Karlsruhe, Germany).
Germany) and used as the template for PCR using a proofreading DNA polymerase (Thermozyme, Invitrogen, Karlsruhe, Germany). Amplification products were subcloned into pGEMT using the pGEM^T Easy Vector System (Promega, Madison, WI, USA) and sequenced. Site-directed mutagenesis (PTS2 of NS, RLx5HL to RLx5VL) was performed using PfuUltra high-fidelity DNA polymerase for mutagenic primer-directed replication of both plasmid strands of NS-EYFP in the pCAT vector using the Quick-Change II site-directed mutagenesis kit (Stratagene; Supplementary Table SII). The cDNA of AAE14 was obtained using the Arabidopsis Biological Resource Center (ABRC, OH, USA). Gene-specific subcloning primers are listed in Supplementary Table SII.

To verify peroxisomal localization in plant cells, the cDNAs were subcloned into plant expression vectors (Supplementary Table SII). Those encoding proteins with predicted PTS1s (SDRs, SRL, SKL) were cloned in-frame with and downstream of the EYFP gene in the plant expression vector pCAT-YFP-C, whereas NS/ECHId carrying a predicted N-terminal PTS2 (RLx5HL) was cloned in-frame with and upstream of EYFP in pCAT-YFP-Ni4. The expression of both fusion genes was under the control of a double 35S cauliflower mosaic virus (CaMV) promoter (Fulda et al., 2002; Ma et al., 2006). Onion epidermal cells were transformed by cell bombardment with gold particles coated with plasmid DNA, and the subcellular localization of the fusion proteins was analysed by fluorescence microscopy (Olympus BX51, SDRa, ECHIa, NS/ECHId, Ma et al., 2006; Nikon TE-2000U, AAE14). In double-labelling experiments, a PTS2 sequence fused to cyan fluorescent protein (ECFP) was used as a peroxisomal marker (Fulda et al., 2002; Ma et al., 2006). For analysis of PTS2 proteins, another peroxisomal marker, red fluorescent protein (RFP)-SKL, was used (Matre et al., 2009).

Results

Isolation of spinach leaf peroxisomes and protein identification

The bottleneck of organellar proteomics is generally the isolation of highly purified organelles from organisms with fully sequenced genomes. Different protocols have been established for the isolation of leaf peroxisomes from model plants of biochemical research. These methods, however, could not be transferred to Arabidopsis, probably due to high concentrations of secondary metabolites and a pronounced physical association of leaf peroxisomes with mitochondria and plastids in Brassicaceae. Investigations were therefore carried out to determine whether leaf peroxisomes can be isolated from S. oleracea L. at a sufficient purity for in-depth proteome analyses and whether sufficient sequence information can be obtained from representative tryptic peptides by LC-MS/MS and de novo sequencing, to facilitate the in silico identification of the corresponding Arabidopsis orthologues.

Leaf peroxisomes from S. oleracea L. isolated by Percoll density gradient centrifugation according to reported protocols are largely devoid of chloroplastic and mitochondrial contaminations, as indicated by sensitive biochemical methods and estimated by chlorophyll content and fumarase activity, respectively (Yu and Huang, 1986; Reumann et al., 1995). Their initial proteome analyses of such a fraction of enriched leaf peroxisomes by 2-DE, however, identified prominent protein spots as the small and large subunits of ribulose-bisphosphate carboxylase/oxygenase (Rubisco) and other plastidic enzymes. Measurement of the plastidic enzyme NADP-dependent glyceraldehyde dehydrogenase (NADP-GAPDH) confirmed a partial co-purification of a non-green plastid variant with leaf peroxisomes (data not shown) that most probably represented partially differentiated plastids. To improve leaf peroxisome purification, the Percoll concentration was increased. Moreover, the washed organelle fraction was laid on top of a second density gradient, i.e. a linear high-resolution sucrose gradient, and subjected to ultracentrifugation. As analysed by sucrose fractionation, the plastid-like organelles had a slightly lower density than the leaf peroxisomes (Supplementary Fig. S1 at JXB online). A significant separation of these plastid-like organelles from leaf peroxisomes was achieved, and their contamination was reduced to a minimum (Fig. 1, Supplementary Table SII at JXB online).

Proteins of highly enriched spinach leaf peroxisomes were analysed by 2-DE. Optimal protein solubilization and representation on 2-D gels was achieved in the presence of thiourea and the zwitterionic detergent CHAPS alone or in combination with amidosulfobetaine-14 (ASB14). The majority of the leaf peroxisomal proteins represented on 2-D gels had a slightly acidic to strongly alkaline isoelectric point (IEP) (pH 6.5–9.5). Individual protein spots were spotted to tryptic in-gel digestion and the peptides analysed by a combination of MALDI, LC-MS/MS, and de novo sequencing by electrospray ionisation/tandem mass spectrometry (ESI-MS/MS). For MALDI-TOF-MS, at least five matching peptides with an m/z error <50 ppm were considered minimal requirements for reliable protein identification. For nano-LC-MS/MS, only identifications presenting high-quality MS/MS spectra (MS/MS ion scores >40) were retained.

Some protein spots were identified as plastidic enzymes, for instance the large subunit of RubisCO and Rubisco activase (Supplementary Table SI at JXB online), indicating the persistence of minor contamination by partially differentiated plastid-like organelles. In contrast, neither proteins from mitochondria, which are the major contaminant of peroxisomes from mammals and fungi, nor endoplasmic reticulum (ER) or nuclear proteins were detected. Known proteins of plant peroxisomes were identified based on high sequence coverage of the full-length polypeptide by tryptic peptides and high sequence similarity of the peptides with homologues cloned from various plant species (Supplementary Table SI). In many cases, the proteins of highest sequence similarity derived from Arabidopsis, but proteins from other plant species (e.g. Citrullus lanatus, Fritillaria agrestis, and Medicago sativa) were also detected (Supplementary Table SI). The most dominant protein spots included two enzymes from the antioxidative system, CAT and ascorbate peroxidase (APX), and several enzymes involved in the photosynthetic C3 cycle [glycolate oxidase (GOX), serine-glyoxylate aminotransferase, glutamate-glyoxylate aminotransferase (GGT), hydroxypropyruvate reductase (HPR), and malate dehydrogenase]. Several protein spots were identified as enzymes involved in fatty acid
β-oxidation or the glyoxylate cycle, including, for instance, long-chain acyl-CoA synthetase isoform 6, thiolase, citrate synthase, isocitrate lyase (ICL), and malate synthase (Fig. 1, Supplementary Table SI). Other proteins included 12-oxo-phytodienoic acid reductase isoform 3, which is involved in the biosynthesis of jasmonic acid, and sulphite oxidase. Some enzymes were found in multiple spots arranged in a horizontal line, indicating different IEPs but similar molecular masses. These proteins included some proteins that are encoded by small gene families in *Arabidopsis*, such as CAT (three genes in *Arabidopsis*), GOX (three genes in *Arabidopsis*), and GGT (two genes in *Arabidopsis*). Additionally, the same horizontal spot pattern was observed for some proteins that are encoded by single genes in *Arabidopsis*, including HPR, ICL, and APX, suggesting that these proteins may differ in IEP due to post-translational modifications.

Unknown proteins of spinach leaf peroxisomes

Sequence information of three unknown proteins of spinach leaf peroxisomes permitted the identification of *Arabidopsis* homologues, which had unknown functions at the time of investigation. One of the three proteins is a member of the superfamily of short-chain alcohol dehydrogenases (spot 104, SDRa, 254 residues, At4g05530). Two other proteins are novel monofunctional enoyl-CoA hydratases/isomerases (ECHIs) of the superfamily of peroxisomal ECHIs that contains five known plant peroxisomal members (Reumann *et al.*, 2004). The first ECHI, referred to as isoform a (spot 103, ECHIa, 264 residues, At4g16210), belongs to clade VI of putative plant enoyl-CoA hydratases (Reumann *et al.*, 2004). The second enoyl-CoA hydratase/isomerase is annotated as naphthoate synthase (spot 63, NS/ECHId, 337 residues, At1g60550) and represents the single orthologue of cyanobacterial MenB, which participates in phylloquinone biosynthesis (Johnson *et al.*, 2001; Gross *et al.*, 2008). Remarkably, all three novel *Arabidopsis* proteins that share significant sequence similarity with the tryptic spinach proteins carry predicted targeting signals for peroxisomes: either C-terminal PTS1 tripeptides (SDRa, SRL>; ECHIa, SKL>) or a putative PTS2 nonapeptide in the N-terminal domain (NS/ECHId, RLx5HL). All three PTS peptides have previously been classified as major PTS peptides and are high-probability predictors of peroxisomal targeting in unknown proteins (Reumann, 2004).

---

**Fig. 1.** Proteomic analysis of leaf peroxisomes isolated from *Spinacia oleracea* L. Proteins of isolated spinach leaf peroxisomes were separated by 2-DE and identified by mass spectrometry. Three different gels with optimal resolution in different areas are shown (A–C). Proteins spots labelled by asterisks (*) were identified from another gel (see Supplementary Table SI at *JXB* online).
Subcellular localization analysis of EYFP fusion proteins in vivo

To provide conclusive evidence for the localization of these unknown proteins in peroxisomes, the cDNAs of the Arabidopsis homologues were cloned by RT-PCR. The cDNAs encoding proteins with putative C-terminal PTS1 tripeptides (SDRa and ECHIa) or an N-terminal PTS2 nonapeptide (NS/ECHId) were fused at their 5' or 3' ends, respectively, in-frame with EYFP in the plant expression vector pCAT. The expression of these fusion genes was under the control of two copies of the 35S promoter of the CaMV (Fulda et al., 2002; Ma et al., 2006). Onion epidermal cells were transformed by bombardment with gold particles coated with plasmid DNA, and the subcellular localization of the fusion proteins was analysed by fluorescence microscopy upon transient gene expression.

The fusion protein EYFP-SDRa, which carried a predicted PTS1 tripeptide (SRL>), was targeted to small subcellular structures that moved quickly along cytoplasmic strands in living cells in single transformants (Fig. 2A1). These punctate structures coincided with peroxisomes labelled with a control fusion protein containing the PTS2 domain of glyoxysomal malate dehydrogenase (gMDH) and ECFP (gMDH-CFP; Fulda et al., 2002), as shown in double transformants expressing EYFP-SDRa and gMDH-CFP.

Fig. 2. In vivo subcellular localization analysis of novel leaf peroxisomal enzymes tagged with EYFP in onion epidermal cells. The cDNAs of four Arabidopsis proteins that were homologous to novel matrix proteins from spinach leaf peroxisomes were fused, depending on the predicted presence of PTS1 or PTS2, in-frame either N-terminally (SDRa, At4g05530, SRL>; ECHIa, At4g16210, SKL>; AAE14, At1g30520, SSL>) or C-terminally (NS/ECHId, At1g60550, RLx5HL) to enhanced yellow fluorescent protein (EYFP) under the control of a double 35S CaMV promoter. Onion epidermal cells were transformed by bombardment with gold particles coated with plasmid DNA, and protein localization was analysed by fluorescence microscopy. For imaging, either YFP- (A1, B, C1, D, E, F1, G, H, I1, J), CFP- (A2, C2, F2, I2), or RFP-specific filters were used (F2). In I2, cyan fluorescence is shown as red for the image overlay (I3).
simultaneously (Fig. 2A). To investigate whether the predicted C-terminal tripeptide of SDRa, SRL>, was indeed a functional PTS1, the predicted peroxisome targeting domain (PTD) of SDRa, comprising the C-terminal 10 amino acid residues, was fused to EYFP. The resulting construct, referred to as EYFP–PTD(SDRa), was localized to punctate subcellular structures, demonstrating that SRL> is the PTS1 of SDRa (Fig. 2B). Likewise, EYFP–ECHIa, with the accessible C-terminal tripeptide SKL>, was localized to subcellular structures that coincided with peroxisomes (Fig. 2C). Consistent with the targeting prediction, EYFP appended C-terminally with the 10 amino acid domain of ECHIa was detected in organelle-like structures, thereby identifying SKL> as the PTS1 of ECHIa (Fig. 2D).

To investigate the subcellular localization of NS/ECHId carrying the predicted PTS2 RLx5HL, a reporter protein fusion construct was created, leaving the N-terminal domain of NS/ECHId accessible. NS/ECHId–EYFP was clearly localized to peroxisome-like structures that were distinguishable in size and shape from plastids (Fig. 2E). In double transformants expressing two PTS2 proteins simultaneously, i.e. NS/ECHId–EYFP and gMDH–CFP, cyan fluorescent organelles could barely be detected next to numerous bright-yellow fluorescent peroxisomes (data not shown). The most likely explanation is that the higher expression level of NS/ECHId–EYFP and competition between both proteins for binding to the cytosolic PTS2 receptor, Pex7p, reduced peroxisomal import of gMDH–CFP below the detection limit. However, in double transformants expressing NS/ECHId–EYFP and RFP–SKL, the yellow and red fluorescent organelles co-localized, conclusively demonstrating that NS–EYFP was localized to peroxisomes (Fig. 2F).

Upon deletion of the putative N-terminal targeting domain of ~40 residues from NS/ECHId, the shortened fusion protein (NS/ECHIdAN–EYFP) remained cytosolic (Fig. 2G). To investigate the subcellular targeting of NS/ECHId by the predicted PTS2 in more detail, the N-terminal domain was altered by site-directed mutagenesis. The exchange of the essential histidine residue to valine (H20→V) was predicted to render the putative PTS2 dysfunctional. The mutated construct remained cytosolic (NS/ECHIdAPTS2–EYFP, Fig. 2H). These data demonstrate that the nonapeptide RLx5HL is indeed the PTS2 of NS/ECHId and that the enzyme is, unlike AAE14 (see below), exclusively targeted to peroxisomes.

Thus, all three Arabidopsis homologues of the novel proteins identified in spinach leaf peroxisomes were verified to be peroxisome localized as EYFP fusion proteins in vivo, and the predicted PTS1/2 signals were conclusively demonstrated to be responsible for peroxisomal targeting.

NS is encoded by a single gene in Arabidopsis that is orthologous to MenB from cyanobacteria, chlorobi, and red algae (Gross et al., 2008; Kim et al., 2008). The enzyme catalyzes an essential step in phylloquinone biosynthesis, the conversion of o-succinylbenzoyl-coenzyme A to 1,4-dihydroxy-2-naphthoyl-CoA (DHNA-CoA) (Fig. 3). Based on gene clustering in cyanobacteria, algae, and higher plants, and experimental subcellular localization data for most enzymes (see Discussion), phylloquinone biosynthesis was previously thought to be exclusively compartmentalized in plastids in higher plants. The present peroxisomal localization data on the Arabidopsis orthologue (see above) cast doubt on this view. Kim et al. (2008) showed that AAE14 catalyzes the reaction upstream of that of MenB, the activation of o-succinyl benzoate to o-succinyl benzoyl-CoA (Fig. 3). When fused C-terminally with green fluorescent protein (GFP) (AAE14-GFP), AAE14 localizes to chloroplasts (Kim et al., 2008). However, AAE14 ends with SLL>, which has recently been characterized as a functional PTS1 (Reumann et al., 2007), suggesting that the enzyme might be dually targeted to both plastids and peroxisomes. These data prompted the investigation of the subcellular localization of AAE14 in more detail as an N-terminal EYFP fusion protein (EYFP–AAE14) that permitted the C-terminal tripeptide of AAE14, SLL>, to function as a subcellular targeting signal. The fusion protein indeed localized to small punctate structures that were clearly distinct from plastids. The identity of the EYFP-labelled organelles as peroxisomes was demonstrated by double labelling the peroxisomes with CFP (Fig. 2I). Moreover, the C-terminal 10 amino acid residues of AAE14 directed EYFP to small subcellular organelles indicative of peroxisomes, identifying SLL> as the PTS1 of AAE14 (Fig. 2J). In summary and in an extension of the study by Kim et al. (2008), the present in vivo targeting data demonstrate that AAE14 is dually targeted to both plastids and peroxisomes. The extended subcellular localization analyses of this proteomic study thus indicate that at least two steps of phylloquinone biosynthesis are compartmentalized in peroxisomes in higher plants (Fig. 3).

Discussion

Proteomic studies are generally restricted to organisms with sequenced genomes to facilitate protein identification from high-resolution 2-D gels by mass spectrometry. Even though apparent tissue limitations of the diminutive Arabidopsis weed can be overcome by large-scale plant growth, it was not possible to isolate leaf peroxisomes of satisfactory quality and quantity from Arabidopsis using methods established for other plants species or Arabidopsis itself. In a parallel approach, an attempt was therefore made to analyse the proteome of leaf peroxisomes from S. oleracea L.

The isolates of leaf peroxisomes from spinach following published methods (Yu and Huang, 1986) were virtually free of chloroplasts and mitochondria, as indicated by sensitive measurements based on the chlorophyll content and the activity of the mitochondrial marker enzyme fumarase (Supplementary Fig. S1 at JXB online). However, the leaf peroxisomes were still contaminated by partially differentiated ‘proplastid-like’ organelles that contained RubisCO and Calvin cycle enzymes. Due to their similar equilibrium density in Percoll and sucrose density gradients, these
proplastids and leaf peroxisomes were difficult to separate. The plastidic contamination was reduced to a minimum by the application of two successive density gradients, but it could not be completely eliminated, a problem shared by recent studies of the peroxisomal proteome (Fukao et al., 2002, 2003; Reumann et al., 2007, 2009; Arai et al., 2008; Eubel et al., 2008). As a result, some plastidic proteins were still present on the 2-D gels.

In vivo subcellular localization analyses are thus essential to validate the putative peroxisomal localization of candidate proteins identified in enriched spinach leaf peroxisomes.

Remarkably, a comparison of the proteomic data of this study with recent studies of leaf peroxisomes from Arabidopsis (Reumann et al., 2007, 2009) shows that some proteins were identified in leaf peroxisomes from Spinacia but not in Arabidopsis, for instance malate synthase and ICL. These data support the necessity of analysing organelles from different plant species before drawing generalized conclusions about organelle constituents (Fukao et al., 2002, 2003; Reumann et al., 2007, 2009; Arai et al., 2008; Eubel et al., 2008).

Since plant peroxisomes cannot be isolated in absolute purity, additional independent lines of evidence were provided that the novel proteins indeed reside in peroxisomes. One efficient method to verify protein targeting to peroxisomes by the PTS1/2 pathways is in silico analysis for conservation of the predicted PTS in homologous plant proteins and ESTs (Reumann, 2004). For instance, all full-length protein sequences of ECHIa homologues from higher plants and from Physcomitrella terminate with putative
PTS1s. Likewise, the PTS2 nonapeptide of NS/ECHId has previously been shown to be highly conserved (Reumann et al., 2007, see also Supplementary Fig. S2 at JXB online).

Many apparent orthologues of AAE14 terminate with SSL> or SNL>, with two sequences even carrying the prototypical PTS1s, SRL> and SKL>. Only a single EST carries a C-terminal tripeptide, CSL> (Manihot esculenta), that has not (yet) been described as a functional plant PTS1 (Fig. 4). In summary, all four candidate proteins carry conserved PTS peptides, strongly supporting peroxisomal protein functions across higher plant species.

Conclusive evidence for peroxisome localization of the unknown proteins was provided by in vivo subcellular targeting analysis. It was demonstrated by double labelling subcellular organelles that SDRa is targeted to peroxisomes in vivo. Similar data have been obtained for Arabidopsis SDRa (Eubel et al., 2008; Wiszniewski et al., 2009) and for the orthologue from soybean (Gm SDRa; Arai et al., 2008). By fusing the 10 C-terminal amino acid residues of SDRa to EYFP, it was further confirmed that the predicted PTS, SRL>, indeed targets SDRa to peroxisomes. The characterization of the targeting signals of predicted PTS1 proteins is important because it enables those proteins to be included as positive control sequences in the development of PTS1 prediction algorithms (T Lingner, G E Antonicelli, P Meinicke, and S Reumann, unpublished data).

Members of the superfamily of SDRs form a universal superfamily of metabolically important oxidoreductases. In Arabidopsis leaf peroxisomes, three further SDR homologues were identified, including the predicted DECR orthologue (SDRb; Reumann et al., 2007). The oxidoreductase SDRa is the plant orthologue of vertebrate peroxisomal NADP-dependent carboxyl reductase, which participates in arachidonic acid metabolism (Fransen et al., 1999). Co-factor specificity for NADP appears to be conserved in SDRa (Reumann et al., 2007). In an ethylmethane sulphonate (EMS) screen, a strong IBA-response mutant, ibr1, was recently isolated that is deficient in IBR1/SDRa (Zolman et al., 2008). Like ibr3 (Zolman et al., 2007), the mutant does not have apparent defects in peroxisomal fatty acid β-oxidation, indicating that IBR1/SDRa is specifically involved in shortening the side chain of IBA to indole acetic acid (IAA) rather than the β-oxidation of straight-chain fatty acids. These present data complement this study by confirming that IBR1/SDRa is indeed a peroxisomal enzyme.

Furthermore, it was verified that ECHIa is targeted to peroxisomes in vivo, as shown previously (Eubel et al., 2008). By fusing its C-terminal 10 residues to EYFP, it was confirmed that the predicted PTS, SKL>, is indeed the peroxisomal targeting signal of ECHIa. Well-characterized members of the Arabidopsis superfamily of peroxisomal ECHIs are the multifunctional proteins MFP2 and AIM1 and the branched-chain amino acid catabolic enzyme CHY1 (Zolman et al., 2001; for a review, see Graham and Eastmond, 2002). Additional members of this protein family include Δ(3,5),Δ(2,4)-dienoyl-coenzyme A (CoA) isomerase (Goepfert et al., 2005) and Δ(3),Δ(2)-enoyl-CoA isomerases 1 and 2 (Goepfert et al., 2008), both of which participate in the degradation of unsaturated fatty acids through the β-oxidation cycle. ECHIa is plant-specific among eukaryotes and is closely related to some poorly characterized prokaryotic ECHIs. The physiological function of ECHIa in plant peroxisomal metabolism remains to be elucidated.

Among plant peroxisomal isomerases, NS/ECHId has a unique role. The fusion protein NS–EYFP, with a predicted N-terminal PTS2 (RLxHL), was localized to small subcellular organelles that were clearly distinguishable in size and shape from non-green plastids with their characteristic stromuli in onion epidermal cells (Reumann et al., 2007). Upon removal of the N-terminal domain, NS/ ECHIdAN–EYFP remained cytosolic. The peroxisomal proteomics and subcellular localization of NS, supplemented by PTS2 conservation in plant orthologues (Reumann et al., 2007; Supplementary Fig. S2 at JXB online), demonstrate that NS/ECHId is a plant peroxisomal protein that is targeted to peroxisomes by the PTS2 pathway. Moreover, NS lacks a predicted transit peptide and has not been identified in any chloroplast proteome study (SUBA II; Heazlewood et al., 2007). Site-directed mutagenesis of

![Fig. 4. PTS1 conservation analysis in AAE14 orthologues. Orthologous protein sequences were retrieved from the protein and EST databases at NCBI by homology analysis (Reumann, 2004), aligned by ClustalX, and shaded by Genedoc. Selected plant species with the most divergent C-terminal PTS1 tripeptides are shown. At, Arabidopsis thaliana; Aa, Artemisia annua; Bm, Burma mangrove; Bn, Brassica napus; Ca, Capsicum annuum; Cc, Cenchrus ciliaris; Gm, Glycine max; Ha, Helianthus annuus; Hc, Hedychium centranthoides; Hv, Hordeum vulgare; Lc, Lotus corniculatus; Ls, Lactuca serriola; Me, Manihot esculenta; Mt, Medicago trunculata; Os, Oryza sativa; Pt, Poncirus trifoliata; Ptri, Populus trichocarpa; Pv, Phaseolus vulgaris; Rc, Ricinus communis; Sb, Sorghum bicolor; Tae, Triticum aestivum; Th, Thellungiella halophila; Tp, Triphysaria pusilla; Vv, Vitis vinifera; Zm, Zea mays.](https://academic.oup.com/jxb/article-abstract/61/5/1441/444046/1449?download=true)
RLx_{5}^{H}L to RLx_{5}^{V}L in the putative PTS2 did not significantly lower the prediction score for plastid targeting (TargetP: wild type, 0.543; mutated version, 0.508; Emanuelsson et al., 2000) but rendered the fusion protein entirely cytosolic. Therefore, it is concluded that NS is exclusively localized to peroxisomes.

The annotation of NS/ECHId as naphthoate synthase is derived from facultative anaerobic bacteria and cyanobacteria, where the orthologues are involved in the biosynthesis of menaquinone (vitamin K2) and phyloquinone (vitamin K1), respectively (Johnson et al., 2001; see also references therein). Phyloquinone, a substituted 1,4-naphthoquinone with an 18-carbon-saturated phytyl tail, functions as a bound one-electron carrier cofactor at the A1 site of the pathway (MenF, D, H, and C), which convert chorismate to o-succinylbenzoate (Fig. 3), are nuclear encoded in higher plants and localized to chloroplasts; the MenD, H, and C domains are encoded as a trifunctional fusion protein in Arabidopsis (Gross et al., 2006). Thirdly, MenA/ABC4 and MenG are also plastid localized (Shimada et al., 2005; Lohmann et al., 2006). For MenB, neither subcellular localization nor functional data have been reported for any plant homologue to date. In light of the lack of any evidence for the existence of additional MenB paralogues in Arabidopsis or additional transcriptional or translational variants, it is concluded that Arabidopsis NS/ECHId is orthologous to MenB from red algae, chlorobi, cyanobacteria, and anaerobic bacteria (Gross et al., 2008).

In higher plants, all nine reactions of phyloquinone biosynthesis have previously been demonstrated or assumed to be compartmentalized exclusively in chloroplasts for several reasons. First, several men genes, including menB, are plastid encoded in red algae, and the gene products are unlikely to be exported. Secondly, the first four enzymes of the pathway (MenF, D, H, and C), which convert chorismate to o-succinylbenzoate (Fig. 3), are nuclear encoded in higher plants and localized to chloroplasts; the MenD, H, and C domains are encoded as a trifunctional fusion protein in Arabidopsis (Gross et al., 2006). Thirdly, MenA/ABC4 and MenG are also plastid localized (Shimada et al., 2005; Lohmann et al., 2006). For MenB, neither subcellular localization nor functional data have been reported for any plant homologue to date. In light of the lack of any evidence for the existence of additional MenB paralogues in Arabidopsis or additional transcriptional or translational variants, it is concluded that Arabidopsis NS/ECHId is orthologous to MenB in Arabidopsis and in most other higher plant species.

The localization of NS/ECHId to peroxisomes sheds new light on the compartmentalization of phyloquinone biosynthesis in higher plants and raises further questions, regarding, for instance, the evolution of phyloquinone compartmentalization. In red algae such as Cyanidioschyzon merolae and Cyanidium caldarium, the (single) homologue of NS is still plastid encoded and thus plastid localized. After the divergence of rhodophyta, the menB gene was transferred to the nucleus. The acquisition of an N-terminal extension and a PTS2 seems to have happened twice independently in the lineage of streptophyta and chlorophyta. While the orthologues of both Ostreococcus (RVx_{5}^{H}HV) and higher plants carry N-terminal extensions harbouring predicted or related PTS2s, those of Chlamydomonas and Physcomitrella are shorter, similar to those of prokaryotic homologues (Mycobacterium, Pelodictyon; Supplementary Fig. S2 at JXB online), and lack PTS2-like nonapeptides. Thus, the NS gene inherited from cyanobacteria has been transferred from the plastid to the nuclear genome at a rather late stage of streptophyta evolution, and its protein product has probably been redirected from plastids to peroxisomes via an intermediate, cytosolic stage, as found nowadays in Physcomitrella.

MenE/AAE14 is the ligase that activates o-succinylbenzoate to o-succinyl-CoA in phyloquinone biosynthesis. Even though the presence of a transit peptide is not predicted for AAE14, the C-terminal full-length fusion protein (AAE14-EYFP) and a GFP with the first 120 residues of AAE14 appended to its N-terminus localized to chloroplasts (Kim et al., 2008). However, similarly to NS, AAE14 has not been identified in any of the numerous high-sensitivity proteomic analyses of chloroplasts (SUBA II). Moreover, analogously to peroxisomal fatty acid catabolism and jasmonic acid biosynthesis, it might be postulated that o-succinylbenzoate, rather than its CoA derivative, is imported into peroxisomes and activated in the matrix (Fig. 3) by transport and reaction processes similar to those for straight-chain fatty acids and OPDA by LACS6/7 and OPCL1, respectively (Fulda et al., 2004; Koo et al., 2006). Indeed, both full-length AAE14 and EYFP appended with the C-terminal 10 residues of AAE14 localized to peroxisomes (Fig. 2), indicating that AAE14 is dually targeted to both chloroplasts and peroxisomes. Future studies need to address the regulation and possible species specificity of the dual targeting of AAE14. The present proteomic and localization data thus indicate that at least two enzymes of phyloquinone biosynthesis are peroxisomal, and that phyloquinone biosynthesis is partially compartmentalized in peroxisomes in higher plants (Fig. 3).

The peroxisomal localization of AAE14 and NS/ECHId raises the question of whether additional enzymes involved in phyloquinone biosynthesis might be localized in peroxisomes and which intermediates are transported across the peroxisomal and the chloroplast envelope membranes. The enzyme immediately upstream of AAE14, MenC, is part of the trifunctional PHYLLO locus and is chloroplast localized. The enzyme downstream of NS, DHNA-CoA thioesterase, has recently been characterized biochemically and cloned from Synechocystis but lacks closely related homologues in higher plants (Widhalm et al., 2009). In light of the recent identification of four small thioesterases in Arabidopsis leaf peroxisomes and the peroxisomal targeting prediction of additional paralogues (Reumann et al., 2004, 2007, 2009), DHNA-CoA thioesterase may well be a peroxisomal enzyme that releases DHNA in the peroxisomal matrix, which is then transported across the peroxisomal membrane back to the chloroplast for attachment of the phytol chain and completion of phyloquinone biosynthesis (Fig. 3).

Supplementary data
Supplementary data are available at JXB online.

**Figure S1.** Improved separation of Percoll-purified leaf peroxisomes and proplastid-like organelles from Spinacia oleracea L. by a second sucrose density gradient.
Figure S2. Multiple sequence alignment of NS/MenB orthologues.

Table S1. Proteins identified in isolated leaf peroxisomes from *Spinacia oleracea* L. by mass spectrometry.

Table SII. Oligonucleotide primers used for cDNA subcloning.

Acknowledgements

We thank Aline Benichou, Dr Xiong-Yan Chen, Alena Liavonchanka, and Nora Valeur for experimental contributions, Gilles Basset for comments, and the Arabidopsis Biological Resource Center (ABRC) for provision of the cDNA clone of AAE14. The research was supported by a grant from the Deutsche Forschungsgemeinschaft (RE1304/2), a Dorothea-Erxleben stipend from the government of Lower Saxony (to SR), and UiS funding.

References


AAE14 gene encodes the Arabidopsis \( o \)-succinylbenzoyl-CoA ligase.


that are required for responses to pro-auxins. *Plant Molecular Biology* **69**, 503–515.


