Arabidopsis peroxisome proteomics

John D. Bussell¹ *, Christof Behrens², Wiebke Ecke³ and Holger Eubel²*

¹ Australian Research Council Centre of Excellence in Plant Energy Biology, The University of Western Australia, Crawley, WA, Australia
² Institute for Plant Genetics, Leibniz Universität Hannover, Hanover, Germany

*Correspondence: John D. Bussell, Australian Research Council Centre of Excellence in Plant Energy Biology, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia. e-mail: john.bussell@uwa.edu.au

Holger Eubel, Institute for Plant Genetics, Leibniz Universität Hannover, Harnower, Harnew, Germany. e-mail: heubel@genetik.uni-hannover.de

The analytical depth of investigation of the peroxisomal proteome of the model plant Arabidopsis thaliana has not yet reached that of other major cellular organelles such as chloroplasts or mitochondria. This is primarily due to the difficulties associated with isolating and obtaining purified samples of peroxisomes from Arabidopsis. So far only a handful of research groups have been successful in obtaining such fractions. To make things worse, enriched peroxisome fractions frequently suffer from significant organellar contamination, lowering confidence in localization assignment of the identified proteins. As with other cellular compartments, identification of peroxisomal proteins forms the basis for investigations of the dynamics of the peroxisomal proteome. It is therefore not surprising that, in terms of functional analyses by proteome means, peroxisomes are lagging considerably behind chloroplasts or mitochondria. Alternative strategies are needed to overcome the obstacle of hard-to-obtain organellar fractions. This will help to close the knowledge gap between peroxisomes and other organelles and provide a full picture of the physiological pathways shared between organelles. In this review, we briefly summarize the status quo and discuss some of the methodological alternatives to classic organelle proteomic approaches.

Keywords: peroxisome, subcellular localization, protein:protein interaction, free-flow electrophoresis, functional proteomics, targeted quantitation of proteins

INTRODUCTION

Microbodies were discovered in the mid 1950s as a particular structure visible in electron micrographs of mouse kidney and rat liver cells (de Duve and Bauduin, 1966). The organelles were initially characterized by co-precipitating enzymatic activities and were consequently named peroxisomes due to the co-precipitation of oxidases and hydrogen peroxide metabolism (catalase) with the isolated structures (de Duve and Bauduin, 1966). Peroxisomes were discovered in plants in the late 1960s due to their association with enzymes of photorespiration (Tolbert et al., 1968, 1969) and subsequently found in most eukaryotic organisms (de Duve et al., 1969). These early studies showed that peroxisomes primarily housed reactions that yielded reactive oxygen species (ROS), including oxidase reactions of β-oxidation and purine metabolism and enzymes (e.g., catalase) to detoxify ROS. In addition, enzymes of the photorespiratory pathway were found in plant peroxisomes (Figure 1).

Up until about 10 years ago understanding of the basic function of peroxisomes in plants had not changed significantly from those early discoveries. Primarily, enzyme activities that completed already known peroxisomal pathways were identified. These included, for example, peroxisomal thiolase in the β-oxidation pathway of plants (Cooper and Beevers, 1969) which previously had only acyl-CoA oxidase (ACX), β-hydroxyacyl-CoA dehydrogenase (multi-functional protein, MFP) and enoyl-CoA hydratase (MFP) activities defined (de Duve, 1969a). More recently, the availability of the genomic sequence of Arabidopsis thaliana (Arabidopsis Genome Initiative, 2000) provided a significant boost to investigations of peroxisome function. For the first time, a whole genome could be interrogated using algorithms designed for prediction of protein subcellular localization signals (as discussed below). Moreover, comparative genomic tools have meant that homologs from unrelated species could be sought in the Arabidopsis genome sequence.

Experimental approaches aimed at elucidating the protein content of a cellular compartment require its isolation. Compared to other plant compartments, chloroplasts and mitochondria were relatively easy to obtain in fractions of good purity and detailed subcellular proteomes were obtained from those organelles soon after annotation and publication of the Arabidopsis genome. To date, around 800 proteins have been reported from proteome
FIGURE 1 | Summary of major metabolic pathways and processes of plant peroxisomes as identified by proteomics. The peroxisome plays a key role in sequestering reactions that evolve reactive oxygen species (Duve, 1969a). This is highlighted by the diverse oxidases (downward pointing arrows) that generate H2O2, and the detoxification of ROS by catalase and other enzymes. Photorespiration and β-oxidation are emphasized in the center of the figure as the major pathways. The dehydrogenases of β-oxidation are primarily dependent on NAD+ as an electron acceptor. Hydroxypyruvate reductase (HPR) of photorespiration requires NADH and this shift in redox requirements is depicted by the interchange between OAA/NADH on the one hand and malate/NAD+ on the other. NADPH is also required by β-oxidation for pathways including unsaturated FA catabolism and JA synthesis. Peroxisome-localized pathways (e.g., glyoxylate cycle, CoA metabolism, OPPP, Asc-GSH cycle) that can be linked with classic peroxisome metabolism are indicated as regular ellipses. The unconnected (psuedouridine catabolism, MVA pathway, bithionol synthesis) or tentatively connected (phylloquinone synthesis, methylglyoxal detoxification) clouds are new additions to the list of known or proposed peroxisome-localized pathways (Reumann, 2011) and are not readily related to core peroxisome metabolism. KEY: Oxidases: PAO, polyamine oxidase; SOX, sarcosine oxidase; SOX1, sulfite oxidase; GOX, glycolate oxidase; ACX, acyl-CoA oxidase. Substrate classes: FA, fatty acid; IBA, indole-3-butyric acid; OPDA, 12-oxo-phytodienoic acid; CA, cinnamic acid; BCFA, branched chain fatty acid; Product classes: IAA, indole-3-acetic acid; JA, jasmonic acid; BA, benzoic acid; OAA, oxaloacetic acid. Other enzymes and pathways: MDAR, monodehydroascorbate reductase; APX, ascorbate peroxidase; Asc-GSH, ascorbate-glutathione cycle; OPPP, oxidative pentose phosphate pathway; MVA, mevalonate; DHP, NADP-dependent ascorbate dehydrogenase.

studies to reside in mitochondria (source: SUBA3; Tanz et al., 2013). Similarly, more than 2100 plastid proteins have been reported (source: SUBA3, queried on November 23, 2012). After the initial cataloging of inventory, recent years have seen mitochondria and chloroplasts being subjected to increasingly detailed functional proteomics with the emphasis shifting from discovery to dynamics (reviewed in Braun and Eubel, 2012). Although the predicted peroxisomal proteome of up to 670 proteins (see below) is clearly simpler than that of chloroplasts (>6000 predicted by ChloroP 1.1; Emanuelsson et al., 1999) or mitochondria (>6000 by TargetP 1.1; Emanuelsson et al., 2000), the difficulties involved in obtaining a pure fraction have severely limited the progress in identifying its true components by MS. It was not until 2007 with significant refinements in peroxisome fractionation techniques and the resultant improvement in the quality of the fractions that large-scale proteomics experiments involving peroxisomes became possible (Reumann et al., 2007, 2009; Eubel et al., 2008).

In this review, we will document peroxisome proteome methodologies before arguing that the basic inventory of the peroxisomal proteome is now reasonably well covered, allowing the move toward quantitative and functional studies.

PROTEOMIC STUDIES OF THE Arabidopsis Peroxisome APPROACHES

In total, five studies have been published with the specified aim of identifying peroxisomal proteins of Arabidopsis (Fukao et al., 2002, 2003; Reumann et al., 2007, 2009; Eubel et al., 2008). The experimental strategies involved in the isolation of organelles and identification of proteins are summarized in Figure 2. Together, these efforts produced a non-redundant list of 204 proteins (source: SUBA, November 23, 2012) but many more are predicted to be located in peroxisomes (Reumann et al., 2004; Reumann, 2011). Conversely, it is likely that at least some of those that have been identified are contaminants.

Nishimura and colleagues’ pioneering studies in Arabidopsis proteomics (Fukao et al., 2002, 2003) reflect the state-of-the-art of proteomics at the beginning of twenty-first century. Using
peptide mass fingerprinting (PMF) of proteins separated by two-dimensional iso-electric focusing/sodium dodecyl sulfate polyacrylamide gel electrophoresis (2D IEF/SDS-PAGE). 29 proteins were identified from leaf peroxisomes of greening Arabidopsis cotyledons (Fukao et al., 2002), while 19 proteins were found in glyoxysomes of etiolated cotyledons (Fukao et al., 2003). The latter study identified glyoxysomal protein kinase 1 (GPK1), a peroxisome-localized protein kinase. The authors were able to show that a number of glyoxysomal proteins are potentially regulated by phosphorylation events (see Section “Protein Modifications” for further details). Interestingly, the overlap between the two studies consisted of only three proteins, suggesting considerable differences in the protein content of Arabidopsis leaf peroxisomes and glyoxysomes. In total, these two studies identified less than 50 proteins. This number includes contaminants as assessed by the combination of high abundances of these proteins in other organelles, the absence of peroxisome targeting signals (PTS; see below), and no obvious relation to expected metabolic activities in peroxisomes.

Since the use of 2D IEF/SDS-PAGE for protein separation actively selects against the identification of membrane proteins for technical reasons, subsequent studies also employed alternative approaches capable of identifying membrane proteins. Reumann et al. (2007) used 2D IEF/SDS-PAGE complimented by shotgun MS to increase the analytical depth of their study of Arabidopsis rosette leaf peroxisomes. In addition, they also established a new purification method employing two successive gradients (Figure 2A). The first gradient was unusual for organelle separations in that the lysate was spun through a zone of Percoll placed on top of three density layers in which the Percoll concentration was successively reduced toward the bottom while sucrose concentration increased concomitantly. The peroxisomes were retrieved from the bottom of this gradient and then further purified through a second gradient made of seven layers of sucrose solutions with increasing density. The peroxisomes formed a clearly visible white band in the lower part of this gradient. The same isolation procedure on the same tissue was used in a follow-up study (Reumann et al., 2009). In this second study, identification of new proteins was facilitated by first running the peroxosomal proteins in a single lane of a SDS gel, then cutting the lane into 16 slices, each of which was submitted to high-resolution tandem MS (Figure 2B). A large proportion of the newly identified putative peroxisomal proteins were then tested for subcellular localization by yellow fluorescent protein fusions (Reumann et al., 2009).

In contrast to the two Reumann studies on Arabidopsis leaf peroxisomes, Eubel et al. (2008) approached elucidation of the proteome of plant peroxisomes by choosing non-green Arabidopsis cell suspension cultures, thereby eliminating chloroplasts as a major source of contamination in the peroxisomal fraction. With the aim of increasing organelle purity even further, a Percoll gradient was followed by free-flow electrophoresis (FFE; reviewed by Islinger et al., 2010; Figure 2A). By using subfractionation of peroxisomes as well as gel (1D and 2D) and non-gel approaches, further novel proteins, including hydrophobic proteins, were discovered by tandem MS (Figure 2B).

CHALLENGES AND PITFALLS
Arabidopsis has been the preferred option for subcellular plant proteomics. Isolates of moderate to good quality can readily be obtained for plastids, mitochondria, plasma membrane, and other organelles and compartments from green tissue or non-green suspension cell cultures. Unfortunately isolating good peroxisomal fractions is notoriously difficult in Arabidopsis. The reasons for this are likely to be due to some or all of the following: (1) the high content of secondary metabolites found in Brassicaceae potentially interferes with isolation (Kaur and Hu, 2011; Reumann, 2011); (2) peroxisomes are considered to be present in much lower numbers in cells than, for example, mitochondria or chloroplasts (e.g., see Germain et al., 2011; Palma et al., 2009); (3) peroxisomes are thought to be particularly fragile in vitro (Palma et al., 2009; Reumann, 2011); (4) losses can be expected to occur due to stresses imposed on the organelles during the isolation procedure; and (5) peroxisomes often physically interact with other organelles (e.g., chloroplasts; Reumann, 2011) and, because of
their fragility, using enough force to break these associations may lead to additional peroxisome damage. As such, highly specialized methods were developed for the isolation of organelles from this species that require sophisticated equipment (e.g., FFE) and/or highly optimized procedures as well as detailed knowledge about the subcellular organelles associated with them. Low yields are typical for peroxisome preparations from Arabidopsis; this further reduces the error margins for successful preparation of peroxisomes and compromises the purity of peroxisome fractions.

The advantages inherent to Arabidopsis (genetic resources, short generation time, easy cultivation, research history) at least partly outweigh the difficulties associated with isolating its peroxisomes and persistence has eventually resulted in acceptable proteomic outcomes. We suggest that a big step forward for peroxosomal proteomics would be if a more generally accessible and well-described isolation technique for Arabidopsis were available. Also, we envisage that species such as spinach, pumpkin, soy, and castor oil bean that were all used as early plant models for peroxisome studies will again rise to be of prominence in peroxisome proteomic studies due to the increasing availability of sequenced, annotated genomes and their amenability to be used as model species.

THE Arabidopsis PEROXISOMAL PROTEOME

The Arabidopsis 2010 peroxisome project compiled a “parts list” of 133 confirmed Arabidopsis peroxisomal proteins, most of them validated by localization of fluorescent reporter fusion proteins. Although curation of this list has ended, an updated version compiled Arabidopsis peroxisome proteome with that of rice has recently been published (Kaur and Hu, 2011). This list includes 163 proteins that fulfill at least two of the criteria of (a) having been identified by MS, and having had their peroxisomal localization supported by either (b) the presence of a PTS or (c) localization of fluorescent reporter fusion proteins. As might be expected, this list contains proteins involved in β-oxidation of fatty acids, auxiliary β-oxidation pathways, photorespiration, ROS detoxification, the glyoxylate cycle, branched chain amino acid metabolism, and peroxisome proliferation. It also contains a considerable number of “less traditional” peroxisomal proteins and others with unknown functions that are so far used as few matrix protein import components are represented. These groups of proteins are summarized in Figure 1 and will be discussed below in terms of their proteomic coverage and the scope to improve this coverage.

It is particularly noteworthy that the main peroxisomal pathway of Arabidopsis (Figure 1), as well as almost all of the secondary pathways represented by these 163 proteins, are also among the putative rice peroxisome proteins, at least as evidenced by the presence of rice homologs possessing peroxisome targeting signals (Kaur and Hu, 2011). Such conservation between plant genomes as diverged as rice and Arabidopsis suggests that this set of proteins can be taken to comprise the core plant peroxisomal proteome. Indeed it is fair to say that (occasional future surprises notwithstanding) the basic function and proteome of plant peroxisomes is now well established and that this fundamental knowledge will foster more advanced proteomic studies in the future.

PEROXISOME EVOLUTION, BIOGENESIS, AND PROTEIN IMPORT

Debate on the origin of peroxisomes has centered on two competing hypotheses. Thus, peroxisomes could either have originated as discrete (single-) membrane-bound structures in primitive eukaryotes or alternatively as an engulfed endosymbiont similar to nascent mitochondria and plastids (de Duve, 1969b).

The evolution of peroxisomes remained a hotly debated field well into the 2000s. It was only the recent discovery in baker’s yeast (Saccharomyces cerevisiae) of de novo peroxisome biogenesis from endoplasmic reticulum (ER) that provided strong evidence against an endosymbiont hypothesis for their origin (Hoepfner et al., 2005). As well as budding directly from ER (van der Zand et al., 2012), peroxisomes can proliferate by division of existing organelles and they frequently receive vesicles from the ER that add to the peroxisomal membrane and carry proteins destined for the peroxisomes. The current model of peroxisome biogenesis in plants considers both options of organelle genesis and, accordingly, is referred to as the “ER-semi-autonomous peroxisome maturation and replication model” (Mullen and Telse, 2006). According to this model, in addition to the direct import of proteins synthesized in the cytosol, proteins can also be transported to peroxisomes directly from the ER via the proposed ER/peroxisome intermediate compartments (ERPICs). Understanding of the basics of protein import machinery in plants followed discoveries in mammalian and yeast systems and was largely dictated by the availability of genome sequences that facilitated gene mining for homologs from these systems (Baker and Spakke, 2005). Genes encoding the essential import and biogenesis machinery are largely conserved across kingdoms indicating that these processes are ancient evolutionary innovations. Peroxisome de novo biogenesis involves peroxins PEX3, PEX16, and PEX19 that are all at some time associated with the peroxisomal membrane. Division in Arabidopsis involves five different isoforms of PEX11(a–e), at least two fertilization-independent seed (FIS) proteins and three dynamin-related proteins (DRPs) which again are at some time peroxisome-localized (Lingard and Telse, 2006; Lingard et al., 2008; Hu et al., 2012).

The majority of peroxisome matrix proteins are imported directly from the cytosol by import machinery that recognizes PTS on the proteins. Two types of PTS peptides are responsible for protein targeting to peroxisomes. About 75% of peroxisome-targeted proteins have a so-called PTS1: a C-terminal tripeptide comprised of a non-polar residue in position-1, a basic amino acid in position-2 and a small and uncharged amino acid in position-3 (Reumann, 2004; Lingner et al., 2011). Serine-lysine-leucine (SKL) is the canonical PTS1 but many possible amino acid combinations can function as a PTS1. Improvements in proteomic techniques (e.g., by increased sensitivity of mass spectrometers which produce higher proteome coverage) and confirmation of targeting by other methods (such as targeting of fluorescent fusion proteins) have led to considerable refinement of the definition of the plant PTS1, recognition of the importance of the sequestered nature of the tripeptide and expansion in the range...
of permissible PTS1 sequences (Reumann et al., 2012). The sec-
type of peroxisome targeting signal, PTS2, is a nonapeptide
usually located within the 20–30 N-terminal residues of pro-
teins that utilize this signal (Reumann, 2004). The consensus
PTS2 sequence consists of two pairs of conserved amino acids
separated by five non-conserved amino acids and in plants is
R[ILQ]x5HL (Kato et al., 1998). As is the case for PTS1s, more
recent studies have refined and expanded the range of import-
capable PTS2 peptides (e.g., Simkin et al., 2011). PTS2 sequences
may also in rare instances be found in the body of the pro-
tein (Reumann et al., 2007; Lambert et al., 2010) and include
alternative residues at the second and ninth positions (Kaur and
Hu, 2011).

PTS1 tripeptides are recognized by the PEX3 protein that docks
with its cargo and is then imported by membrane-bound assembly
of PEX13 and PEX14, which form the core import channel. PTS2
import utilizes PEX7 that must first interact with PEX5 before
import via the same channels (reviewed in Hu et al., 2012). In both
cases, the PEX5/PEX7-cargo complex is imported as a whole into
the peroxisomal matrix, with PEX5/PEX7 recycled to the cytosol by
a mechanism involving ubiquitination of PEX5 and that utilizes
membrane-bound proteins PEX22, PEX2, PEX16, PEX12, and
APEM9 as well as other PEX proteins (PEX4, PEX6, and PEX11)
that are tethered on the cytosolic side to various protein components
of the receptor recycling machinery. After import into peroxisomes,
PTS2 sequences are cleaved by the trypsin-like DEG15 protease
(Schuhmann et al., 2008); such modifications do not occur with
PTS1 signals.

Proteomic analysis by MS has to date been almost singu-
larly unsuccessful at detecting peroxisome-localized components
of biogenesis, division and import pathways in Arabidopsis. Of
the at least 12 protein import-related proteins, only PEX14 was
identified in recent Arabidopsis studies (Reumann et al., 2007,
2009; Eubel et al., 2008). Similarly, with the exception of PEX11
isoforms (see Reumann et al., 2007, 2009; Eubel et al., 2008), per-
oxisome biogenesis and division related proteins have not been
found by MS-proteomics. Possible reasons for the lack of success
in isolating most of the membrane-bound proteins involved in
peroxisome biogenesis and protein import in plants are their very
low abundances, life stage specific expression, transient associ-
tions with peroxisomes, difficulties in the isolation and analysis
of peroxisome membranes, and selection against the identification
of membrane proteins due to technical limitations (e.g., the use of 2D
IEF/SDS-PAGE).

PROTEINS OF THE CORE PEROXISOME METABOLISM

Peroxisome metabolism is important to all stages of plant growth,
including seed development, germination, general growth, and
senescence. Peroxisomes can be seen essentially as organelles
that sequester reactions producing reactive oxygen species and facil-
itate benign detoxification of ROS. Other reactions that occur
in peroxisomes have almost always been shown to comprise a
“service-industry” that recycles reaction intermediates and co-
products. Peroxisomes are primarily β-oxidation (β-oxidation
activities such as acyl-activating enzymes (AAEs), ACX, MFP, 3-ketoacid-CoA thioloase (KAT), and the many sub-
unitary enzymes in these multigene families such as single-function
hydratases, epimerases, and short chain dehydrogenases. This
figure of 25% does not include enzymes outside the core path-
way such as those of ROS detoxification, malate dehydrogenase
(MDH), peroxisomal adenine nucleotide carrier (ANT), enzymes
of CoA metabolism, and others that are necessary to sustain β-
oxidation and that collectively would clearly significantly increase
this proportion.

Once seedlings are established, the primary role of plant perox-
isomes is their participation in the photosynthetic pathway. This
process serves the detoxification of 2-phosphoglycerate produced
by the oxygenase reaction of RuBisCO and the salvage of the bulk
of its carbon atoms. While the photosynthetic pathway is spa-
tially split between chloroplasts, mitochondria, and peroxisomes,
the majority of the enzymes directly involved in the recovery of
phosphoglycerate (after removal of the P2 group in the chloroplast)
are peroxisomal. Six enzymes of photosynthesis (including catalase) are located in the peroxisome and at least six different substrates have to be transported across the peroxisome membrane for photosynthesis to function properly. This renders peroxisomes the central hub in a biochemical pathway that is so important for the central carbon metabolism of plants that, in terms of carbon flux, it is surpassed only by photosynthesis (Bausch et al., 2010) and it is no surprise that these proteins are prominently represented in proteomic studies. Despite substantial efforts to reduce the ostensibly wasteful process of photosynthesis (e.g., Kebesh et al., 2007; Peterhansel and Maurino, 2011), it is becoming increasingly clear that it is a necessary prerequisite for photosynthesis to function properly. This renders peroxisomes provisionally, or at least six different sub- enzymes of pseudouridine catabolism (pxPfkB/At1g49350, At1g43340, At1g65520, At2g66050, and At1g89670) are listed in the PhosPhAt database4. Furthermore, PMP38/PXN (At2G39970, a peroxisome membrane-localized NAD+ transporter, Bernhardt et al., 2012), was also found to be phosphorylated (Eubel et al., 2008). We suggest that targeted proteomics could play an important role in analyzing mutants in these protein modification and phosphorylation/kinase genes to determine substrates and extent of processing.

**PEROXISOME PROTEOMICS IN THE FUTURE**

With the knowledge gained from previous studies, Arabidopsis peroxisome proteomics can be expected to venture in the following directions: (1) search for novel peroxisomal proteins and (2) detailed analyses of peroxisomal proteins in respect to dynamic changes of protein abundances, modifications of peroxisomal proteins, and their potential interactions with other proteins to form temporary or stable protein complexes. Several routes with the potential to further populate the list of plant peroxisomal proteins are discussed and these approaches may be tied to functional proteomic characterization as outlined below.

**IDENTIFICATION OF NOVEL PEROXISOMAL PROTEINS**

Refinement of peroxisome isolation methodology

In general, allocating proteins to the peroxisomal compartment by experimental data (MS or other methods such as localization of fluorescent protein fusions) provides greater confidence

http://phosphat.mpimp-golm.mpg.de/phosphat.html (queried on March 8, 2013)

4 http://phosphat.mpimp-golm.mpg.de/phosphat.html (queried on March 8, 2013)
than predictions alone. However, often MS data and even fluorescent reporter proteins produce false-positive results. In the case of MS, the presence of some of the proteins in the data is likely to be due to contamination of the peroxisomal fraction with other organelles. Usually, proteins from contaminating organelles are present in low numbers in the target fraction, but nonetheless they may swamp the analytical sensitivity needed to isolate low abundance proteins of the target organelle. Alternative means are therefore required to isolate organelles to a higher level of purity, both to reduce contamination and to increase the chance of finding rare proteins. One such strategy could be the use of FFE. FFE has been used to further purify mitochondria from yeast and plants, as well as rat peroxisomes (reviewed in Islinger et al., 2010). It has also been successfully employed for the isolation of peroxisomes from Arabidopsis cell suspension cultures. Starting from material heavily contaminated with mitochondria, FFE was able to reduce the mitochondrial content by a factor of 5 (based on oxygen electrode assays) while increasing the concentration of peroxisomes threefold (Islinger et al., 2008). Although FFE will face a second major source of contamination in fractions prepared from green leaves (plastids in addition to mitochondria), this approach might also improve the purity of peroxisomes prepared from green plant tissue. In combination with the gradient(s) established for leaf material (Reumann et al., 2007), FFE has the potential to deliver peroxisomes with very low levels of contamination by mitochondria or plastids.

While FFE might be able to reduce contamination, it cannot compensate for the low yields typical of plant peroxisome preparations. The use of more starting material in existing protocols will inevitably alter the physiology of the plant cell and its organelles, they may represent a viable option to increase yield if contamination in fractions prepared from green leaves (plastids in addition to mitochondria), this approach might also improve the purity of peroxisomes prepared from green plant tissue. In combination with the gradient(s) established for leaf material (Reumann et al., 2007), FFE has the potential to deliver peroxisomes with very low levels of contamination by mitochondria or plastids.

Bioinformatic approaches to predict novel peroxisomal proteins

Bioinformatics has provided considerable insight into likely upper limits to the size of the Arabidopsis peroxisomal proteome. Up to 542 proteins are predicted to contain a PTS1 sequence (SUBA3, November 23, 2012; Kaur and Hu, 2011). The number of potential peroxisomal proteins has expanded considerably in the last 10 years. Reumann et al. (2004) provided the first comprehensive database of putative Arabidopsis peroxisome proteins (Arapoex5), listing 284 proteins based on prediction of PTS1 and PTS2 targeting peptides. Arapeox was updated in 2008 to 440 proteins, including another 110 proteins with the newly demonstrated PTS1 signals SSL, SSI, ASL, and AKI (Liszenbee et al., 2005; Reumann et al., 2007). PEX proteins, demonstrated membrane proteins and proteins that are imported using non-standard targeting peptides (e.g., catalase and sarcosine oxidase; Goyer et al., 2004; Osshima et al., 2008). The number of predicted proteins has further increased by refinements to (in particular) PTS1 prediction algorithms (Ma and Reumann, 2008; Lingner et al., 2011; Chowdhury et al., 2012). These studies have taken particular consideration of the context (upstream) of the putative PTS1s and thus identified weak, non-canonical PTS sequences that supported import of proteins into plant peroxisomes. Extensive testing of 23 newly predicted PTS1 motifs suggested unforeseen diversity in plant peroxisome-import competent C-terminal tripeptides (Lingner et al., 2011). The majority of these were tested by fusing enhanced yellow fluorescent protein (EYFP) to the 10 terminal amino acid residues of the predicted proteins or, in a few cases, to the N-terminus of full-length proteins.

While it is possible that the peroxisomal proteome might contain more matrix proteins than are currently predicted, it is equally clear that some predicted peroxisomal proteins are unlikely to localize to the organelle in vivo. As a trivial example, the plastid genome encoded rpoC2 (RNA pol) has a strong PTS1 like sequence (SRI) at its C-terminus. In total, 204 proteins have been assigned to the peroxisomal compartment by MS but only 97 of these are found in the list of ~670 proteins predicted (or confirmed by other means) to reside in peroxisomes. This small overlap (<15% of the predicted and <50% of the MS-detected proteins) implies that both the prediction and detection of peroxisomal proteins suffer from false-positive results. It also seems likely that there will be more surprises in the form of unexpected proteins to be assigned to this organelle. Further research and refinement of prediction algorithms (especially for PTS2 sequences) will yield more candidate proteins for the peroxisomal proteome. These bioinformatic works represent major advances in setting the outer boundaries for the plant peroxisomal proteome.

An alternative bioinformatic approach has been to mine the wealth of publicly available transcriptome data for genes that are...
co-transcribed with genes encoding proteins for core peroxisome functions. This approach was taken in Wizniewski et al. (2009) to show that numerous previously uncharacterized PTS-encoding β-oxidation genes followed similar patterns of transcriptional expression to other, well-characterized β-oxidation genes. Such data mining for proteins involved in other areas of peroxisome biology could help to provide clues for timing of expression and the function of novel or thus-far undetected putative peroxisomal proteins.

Genetic resources

Two classic screens for mutants affected in peroxisome function have been used extensively and both screens can reveal mutants compromised in β-oxidation (Hayaishi et al., 1998; Zelterman et al., 2000). The first uses the peroxisome-localized conversion of pro-ayins [IBA or 2,4-dichlorobutyruric acid (2,4-DB)] into active forms [IAA or 2,4-dichloroacetic acid (2,4-D)]. Mutants in these pathways exhibit root elongation when grown on media containing IBA or 2,4-DB, whereas growth of wild-type roots is inhibited. The second screen utilizes the requirement for β-oxidation to release carbon from fatty acids to fuel germination and seedling establishment. Lesions in this pathway result in seedlings that either do not germinate or fail to establish unless they are grown on media that is supplemented with a sugar carbon source. These early forward genetics studies identified a number of single gene mutants of large effect in these pathways (CTN/PLEX1/PEP3, KAT2/PED1, etc.). However, many β-oxidation proteins are encoded by gene families that exhibit functional redundancy; and this necessitates targeted (reverse genetic) generation of double- and potentially higher order mutants (e.g., ACX, long chain acyl-CoA synthetase (LACS) families).

Two studies have taken a brute force approach to genetic characterization of putative peroxisome genes. In the first (Wisniewski et al., 2009), all available mutants for 16 newly predicted (or otherwise uncharacterized) genes with similarity to known β-oxidation genes (as reported in Arabidopsis; Reumann et al., 2004) were screened for response to growth on IBA, 2,4-DB, and sugar-free media. This study yielded new genes in auxin metabolism pathways, but suggested that there were no new single gene knockouts of peroxisome proteins that would display a sugar dependence phenotype. The study also showed that the new auxin pathway genes followed a transcriptional pattern common to many β-oxidation genes. Secondly, Kaur and Hu (2011) hunted at a large, unpublished study that took a similar approach with about 50 predicted peroxisome genes and that involved various biochemical, physiological, and cell biological assays aimed at documenting the role of these genes in embryogenesis, peroxisomal protein import, and defense response.

Identification of proteins interacting with bona fide peroxisomal proteins

An alternative experimental approach to defining localization is to identify interacting proteins, since these are very likely to be localized in the same compartment. By using an approach not biased by previously defined parameters for peroxisomal targeting, the verification of peroxisome localization of interacting proteins potentially leads to the discovery of so far unknown plant peroxisome proteins and new definitions of functional PTS sequences. A suite of techniques is available to isolate protein complexes. Antibody- or affinity-tag-based techniques such as co-immunoprecipitation (co-IP) or co-precipitation have been flagged to be of likely utility in expanding proteomic knowledge (Yates, 2000). By these methods, a protein of known localization is used as bait and incubated with protein extracts from fractionated samples to permit interactions and formation of complexes. Subsequently, an antibody against the known peroxisomal protein (e.g., bound to the resin of an affinity chromatography column) is used to pull down the complex, which can be dematured and run on a gel or subject directly to MS detection. An extension of these methods, tandem affinity purification (TAP), has shown particular promise in proteomic applications. TAP involves a two-step affinity purification method that may significantly reduce the incidence of non-specific binding, thus resulting in greater purity of the isolate (Rigaut et al., 1999). The potential for these methods to be used in expanding documentation of the plant peroxisomal proteome is discussed briefly below.

Co-IP has been used in mammalian cells (rat liver) to isolate and confirm interaction of a matrix protein (β-bifunctional enzyme, β-BFE) with catalase (Makkar et al., 2006). Likewise, it was used to demonstrate interactions between a subset of yeast (S. cerevisiae) PEX proteins (Eckert and Johnson, 2003). However, these clearly do not represent high-throughput discovery studies. In a more generic approach, antibodies against a PTS could be used in pull-down experiments. For example, anti-luciferase-PTS1 antibody detects multiple proteins on western blots of purified peroxisome fractions from rat livers (Gould et al., 1990). Using such an antibody in co-IP studies could conceivably yield new proteins, but the antibody is likely to be specific to the particular C-terminal tripeptide (SKI, in this case; Gould et al., 1990) and the growing diversity of PTS1 sequences as documented above may limit this approach. Isolation of PTS2-containing proteins would require another suite of antibodies. TAP has been promoted as a high-throughput method for protein complex discovery. An early application of TAP was a large-scale analysis of protein complexes in yeast (S. cerevisiae) in which TAP tags were directly attached to the C-termini of 1739 proteins (Asp et al., 2001). Isolation of PTS2-containing proteins would require a different suite of antibodies.
S. cerevisiae is well suited to such TAP analysis because homologous recombination can be used to generate large libraries of strains expressing the tagged proteins at approximately native levels (under control of endogenous promoters), and N-terminal tagging could just as well be used to preserve endogenous C-terminal targeting signals. Unfortunately, translating such approaches to Agrobacterium and other plants has been problematic, not least because there is no method for homologous recombination of constructs into the genome. TAP tagging in plants thus requires Agrobacterium-mediated transformation of individually cloned constructs into wild-type plants, or into mutants in which the native gene has been knocked out to preclude competition of the untagged endogenous protein for binding. Nevertheless, TAP tagging has been promoted for protein complex discovery in plants and the method adapted for plant-specific application (Van Leene et al., 2008). Plant cell cultures have been successfully used with TAP methodology to identify cell cycle component interacting proteins but the process had numerous drawbacks including its complexity, susceptibility to false negatives (due to low abundance, transient expression, or absence of likely interactors in cell cultures) and false positives to highly "sticky" non-specific interactions (Van Leene et al., 2011).

Data on protein-protein interactions of peroxisomal proteins can also be found in recent global interactome studies. Databases such as the IntAct molecular interaction database (Lee et al., 2010; Kerrien et al., 2012) or the third version of the Arabidopsis Subcellular Database (SUBA3, Tan et al., 2013), that now incorporates protein interactions, can be queried to detect interaction partners of peroxisomal proteins. Kaur and Hu (2011) have summarized the currently confirmed Arabidopsis peroxisomal proteome into a list of 163 proteins (see above). By interrogating SUBA3 (November 23, 2012) with this list, we identified a non-redundant set of 133 proteins with claims for interactions with the original set (Figure 3A; Table 1). Of these, the subcellular location was already deduced from fluorescent reporter fusion proteins. Nine of them were exclusively or non-exclusively assigned to peroxisomes. MS assigned 30 other proteins to their respective intracellular locations, and nine of these were also found in peroxisomes. For the remaining 68 proteins, no experimental GFP and/or MS data are stored in SUBA3; these proteins are thus candidates for peroxisomal localization. Eight of them were predicted to be peroxisomal by PredPlantPTSn1 (Reumann et al., 2012) or Multiloci2 (Blum et al., 2009) and are therefore very likely imported into peroxisomes. On this basis, they should now no longer be considered as candidates but as established peroxisomal proteins (Table 1). The remaining 60 proteins were used as queries to interrogate SUBA3 for their interaction partners. Using the putative subcellular localization of the returned interaction partners, as given by the consensus of all localization data for each protein stored in the SUBA3 (SUB-Acorn), we calculated the percentage of the interacting proteins that are known to be peroxisomal compared to those localizing to other cellular compartments. For 13 candidate proteins, at least a third of the interaction partners consisted of peroxisomal proteins, while this number shrank to nine if a cut-off value of 50% was applied (Figure 3B; Table 1). Of these, six proteins had only a single interacting partner and thus by definition

### Functional Characterization of the Peroxisomal Proteome

There is more to proteomics than a mere stocktake of the protein content of an organelle. The studies by Fukao et al. (2002, 2003) have clearly shown that we can expect differences in peroxisomes isolated from cotyledons performing autotrophic and heterotrophic metabolism. Except for these early studies (that have more of a qualitative than quantitative character), comparative studies of the plant peroxisomal proteome are, to our knowledge, non-existent. Clearly, obtaining results on changing protein abundances would have a strong impact on our current view of peroxisomes as cellular organelles and their reactions to changing physiological conditions. MS-based comparative studies are traditionally performed by using quantitative approaches such as isobaric labeling (isobaric tags for relative and absolute quantitation (iTRAQ)), tandem mass tag (TMT), heavy mass tags (isotope coded affinity tags, ICAT) and, to a lesser degree, stable isotope labeling (stable isotope labeling by amino acids in cell culture, SILAC). These techniques are quite sensitive and allow quantitation of several thousand proteins at a time. However, due to the relatively low number of peroxisomes within cells and the resulting low average abundance of peroxisomal proteins in cell lysates, coverage of peroxisomal proteins present only in low copy numbers still presents a challenge to quantitative shotgun proteomics. Therefore, isolating peroxisomes from two or more plant populations is still deemed necessary to achieve a better coverage of peroxisomal proteins in shotgun comparative approaches.

A promising alternative to this classical approach may be the use of targeted quantitation of proteins. Using targeted quantitation, low abundance proteins are detectable in tryptic digests of, for example, leaf homogenates without the requirement of first obtaining peroxisome isolated. Targeted absolute or relative quantitation is most commonly performed by employing selected reaction monitoring (SRM; Gerber et al., 2003; DeSouza et al., 2008; Zhi et al., 2011), a technique that originated from the targeted analysis of small molecules and has also become available for peptide analysis. In SRM, only a few peptides specific for the target protein are considered in the MS analysis, resulting in short duty cycles which save analytical time when compared to data-dependent analyses approaches. Thus the elution peaks of the target peptides can be monitored more closely, resulting in higher accuracy quantitation, especially for low abundance peptides. Modern mass spectrometers and knowledge of the retention times of the selected peptides allow the quantitation of up to several hundred transitions in one event and for this reason the technique is often referred to as multiple reaction monitoring (MRM). Unfortunately, establishing transitions for MRM is a labor-intensive process. However, once this has been

---

**Table 1**

<table>
<thead>
<tr>
<th>Protein</th>
<th>Subcellular Location</th>
<th>Identified in Peroxisomes</th>
<th>MS Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein 1</td>
<td>Peroxisome</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Protein 2</td>
<td>Peroxisome</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Protein 3</td>
<td>Peroxisome</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Protein 4</td>
<td>Peroxisome</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Protein 5</td>
<td>Peroxisome</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Protein 6</td>
<td>Peroxisome</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Protein 7</td>
<td>Peroxisome</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Protein 8</td>
<td>Peroxisome</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Protein 9</td>
<td>Peroxisome</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

*Figure 3A and 3B are not shown in this text representation.*

---

**Figure 3A**

**Figure 3B**

---

*The full text of the document is available on the website.*

---

www.frontiern.org
FIGURE 3 | Subcellular localizations of proteins interacting with confirmed peroxisomal proteins. One hundred and sixty-three proteins listed by Kaur and Hu (2011) were queried for interacting proteins listed in the SUBA3 database. This returned 133 proteins (Table 1) that were then interrogated for their subcellular localization. Thirty-five proteins were allocated to their respective cellular compartment by GFP analyses, 30 by proteomics studies (MS/MS), and eight proteins possessed localization data on the basis of high-confidence predictions. For 60 proteins, no localization was obtainable and therefore they are considered candidate peroxisome proteins (A). In case proteins were assigned to a location by more than one approach, their listing in Table 1 was done according to the following priorities: GFP > MS/MS > predictions. SUBA was then interrogated with the 60 candidate proteins to determine the localization of their interaction partners, as summarized in the SUBAcon (B). Proteins are sorted by the percentage of peroxisomal interactors, regardless of the overall number of interacting proteins. For six proteins, the bait protein was the only interactor. For some proteins on the list of Kaur and Hu (2011) SUBAcon states a non-peroxisomal localization. Therefore, 13 proteins do not appear to have a single peroxisomal interaction partner although at least the bait protein is expected here.
Table 1 | SUBA3 localization data of proteins interacting with confirmed peroxisomal proteins according to Kaur and Hu (2011).

| AT1G23790.1 | AT5G25420.1 | AT4G17760.1 | AT4G21160.1 | AT5G01500.1 | AT5G19920.1 | AT5G46030.1 | AT5G51300.1 | AT5G557840.1 | AT5G644160.1 | AT5G65480.1 | AT5G67750.1 |
| AT3G12110.1 | AT3G14415.1 | AT3G14420.1 | AT3G21865.1 | AT3G26900.1 | AT3G58740.1 | AT4G14440.1 | AT5G56220.1 | AT5G65683.1 |

Apart from the peroxisome-specific reasons to employ MRM assays for the quantitation of proteins, it might also prove advantageous for other organelles. Isolation of the target compartment of a cell is a process that usually takes several hours, during which time unforeseen changes to metabolites, membranes and also to proteins may occur. Additionally, because centrifugation is often used, organelar subpopulations characterized by extreme densities or sizes might constitute the final isolate, differing somewhat from the situation found in vivo. Again, due to their fragility, this might affect peroxisomes more severely than other, more stable compartments. Thus, the development of these newer methods to reduce the isolation time and organelar stress prior to analysis will only serve to enhance the value of and possibilities for dynamic proteome studies.

CONCLUSION

Peroxisomal proteomics in Arabidopsis is seriously hampered by limited access to isolated organelles. Due to the low number of peroxisomes in plant cells, most peroxisomal proteins do not rank among the high-abundant proteins. This prevents good coverage of the peroxisomal proteome in shotgun proteomics. Increasingly good predictions of peroxisomal location by different routes will become available in the future. However, experience gained in the past on false-positive allocation of proteins to this organelle has shown that predictions are only suitable for generating candidates for the peroxisomal proteome and that experimental validation of predicted results is still necessary and will most likely remain so. Most of the experimental evidence will be obtained by fluorescent reporter fusion protein assays, but other approaches also lend themselves for this purpose. Laboratories geared toward proteomic studies could use targeted quantitative analysis of candidate proteins from isolated peroxisomes and fractions purified to lesser degrees of homogeneity in order to show enrichment of candidate proteins in peroxisomes.

In order for peroxisomal proteomics to reach the standard of the studies performed on plastids or mitochondria, technical improvements are of paramount importance. Such improvements will be obtained either in the purity and yield of peroxisomal isolates, or by alternative analytical methods to increase proteome coverage. Moreover, investigation of the dynamic changes occurring in the plant peroxisomal proteome may require a completely different approach. Since MS-technology is developing rapidly and advances in the isolation of peroxisomes are comparatively slow, recent and future developments in MS in combination with alternative strategies for identifying members of the plant peroxisomal proteome will be key to a better understanding of the functions and dynamics of this important compartment of the plant cell.

ACKNOWLEDGMENTS

This work was supported by the Australian Research Council (grant number CE0561495). We acknowledge support by the Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of Leibniz Universität Hannover.
REFERENCES


Arabidopsis peroxisomes pro tease


Lisenbee, C. S., Lingard, M. J., and Bussell et al. Arabidopsis


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 28 December 2012; accepted: 02 April 2013.

Copyright: 2013 Bussell, Behrens, Ecke W and Zabel. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution, and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

April 2013 | Volume 4 | Article 101 | 13