Sorting pathway and molecular targeting signals for the Arabidopsis peroxin 3

Joanne E. Hunt and Richard N. Trelease*

School of Life Sciences, Cellular and Molecular Biosciences, Arizona State University, Tempe, AZ 85287-4501, USA

Received 15 December 2003

Abstract

Peroxin 3 (Pex3p) has been identified and characterized as a peroxisomal membrane protein in yeasts and mammals. We identified two putative homologs in Arabidopsis (AtPex3p, forms 1 and 2), both with an identical cluster of positively charged amino acid residues (RKHRRK) immediately preceding one of the two predicted transmembrane domains (TMD1). In transiently transformed Arabidopsis and tobacco BY-2 suspension-cultured cells, epitope-tagged AtPex3p (form 2) sorted post-translationally from the cytosol directly to peroxisomes, the first sorting pathway described for any peroxin in plants. TMD1 and RKHRRK were necessary for targeting form 2 to peroxisomes and sufficient for directing chloramphenicol acetyltransferase to peroxisomes in both cell types. The N and C termini of AtPex3p (form 2) extend into the peroxisomal matrix, different from mammal and yeast Pex3 proteins. Thus, two authentic peroxisomal membrane-bound Pex3p homologs possessing a membrane peroxisomal targeting signal, the first one defined for a plant peroxin and for any Pex3p homolog, exist in plant cells.

© 2003 Elsevier Inc. All rights reserved.

Keywords: Arabidopsis; BY-2; Membrane targeting signal; Pex3p; Peroxin; Peroxisome; Transmembrane domain

Peroxisomes are ubiquitous in cells of eukaryotic organisms, constituting a dynamic class of subcellular organelles. Relatively small in diameter (0.1–1.0 μm) and delineated by a single boundary membrane, peroxisomes do not possess their own nucleic acids; consequently, they are solely dependent on the nuclear genome. In all organisms, peroxisomes function in oxidative metabolism that involves oxidases and elimination of H2O2 and reactive oxygen species [1,2]. Peroxisomes also participate in a variety of metabolic pathways depending on the organism, developmental stage, cell/tissue, environment, etc. For example, yeast peroxisomes house inducible enzymes capable of metabolizing alkanes, fatty acids, methanol, uric acid, or primary amines [3]. In mammals, peroxisomes participate in the synthesis of ether-linked phospholipids and sterols, and in the degradation of long chain fatty acids and phytic acid [4,5]. In higher plants, peroxisomes are subdivided according to their main function into four classes: (a) glyoxysomes in oilseed seedlings house enzymes of the glyoxylate cycle and β-oxidation of fatty acids, (b) leaf-type peroxisomes in photosynthetic tissues contain enzymes that participate in photosynthesis, (c) root nodule peroxisomes in differentiated cells of legume nodules possess enzymes essential for ureide biosynthesis, and (d) unspecialized peroxisomes contain the basic enzymes for all peroxisomes, namely catalase and oxidases [6,7].

The most recent focus of peroxisomal research has been the elucidation of biogenetic events related to these organelles [8–10]. Proteins involved in peroxisomal biogenesis are termed peroxins (Pex), most of which are peroxisomal membrane proteins (PMP) [11,12]. Analyses of yeast PEX mutants have been instrumental in discovering putative peroxins in mammals, plants, and trypanosomes [13,14]. The importance of identifying and characterizing peroxins is underscored in humans who suffer from peroxisomal biogenesis disorders resulting in lethal diseases such as Zellweger syndrome, X-linked adrenoleukodystrophy, and Refsum disease [15]. These diseases are attributed to missing or aberrant biogenetic peroxins.
Mature functional peroxisomes house an array of matrix proteins, which possess well-characterized peroxisomal targeting signals (PTSs) [12,16]. They are either a type 1 or type 2 PTS (PTS1 or PTS2), which is recognized by the cytosolic receptors Pex5p and Pex7p, respectively. PTS1 is a carboxy terminal tripeptide composed generally of small-basic-hydrophobic amino acid residues, such as the prototypic serine–lysin–cysteine sequence [12,17], whereas PTS2 is a nonapeptide (R/K)(L/V/I)X5(H/Q)(L/A) found within or at the amino terminus of a different subset of matrix proteins [8,18–20].

Targeting signals for membrane proteins are distinctively different from those directing proteins to the matrix. Transmembrane domain(s) (TMD) flanked on either side by a stretch of basic amino acid residues generally constitute a membrane peroxisomal targeting signal (mPTS) [21,22]. Peroxisomal targeting of PMP47 in the yeast *Candida boidinii* requires multiple elements for targeting to peroxisomes including: (a) a cytosolic oriented loop of amino acid residues between TMD1 and TMD2, (b) the second of five TMDs, (c) a short matrix loop containing basic amino acids, and (d) a membrane anchoring TMD [23,24]. The human counterpart to PMP47, PMP34, contains two non-overlapping sets of targeting information which are composed of amino or carboxy terminal membrane spanning domains and a cluster of basic amino acid residues [25]. Targeting of human PMP22 was similar in that it required two distinct non-overlapping TMDs and corresponding basic polypeptide loops to efficiently target to GFP peroxisomes [26]. Targeting of PMPs in plants is predicted to be analogous to that observed in yeasts and mammals. For example, the *Arabidopsis* PMP22 mPTS is composed of at least four distinct regions within the N-terminal one-half of PMP22 that includes a positively charged domain and all four TMDs [27]. The mPTS of cottonseed peroxisomal ascorbate peroxidase (APX) is composed of its C-terminal TMD and adjacent basic cluster of amino acid residues [28]. An obvious void in the literature are data describing the necessary and sufficient elements of a mPTS for any plant peroxin.

In this study we report that there are two *Arabidopsis* peroxin 3 homologs (AtPex3p, form 1 and form 2), which are unique to the single forms of Pex3p reported for yeasts [29–32] and mammals [33,34]. In vivo immunofluorescence microscopy of transiently expressed proteins was employed to determine the subcellular localization of AtPex3p (form 2) in *Arabidopsis* and tobacco BY-2 suspension-cultured cells. Experimental results revealed that AtPex3p (form 2) sorts directly from the cytosol to peroxisomes in both cell types. Both forms of AtPex3p possess targeting elements similar to those reported for Pex3 proteins in yeasts and mammals, even though the primary amino acid sequence identity of the *Arabidopsis* forms is less than 19–25% of the yeast and mammalian Pex3 proteins. In vivo experimental results also revealed that the topological orientation of AtPex3p form 2 (Nexo–Cin) differs from that of any of the yeast or mammalian Pex3 proteins that have been examined.

### Materials and methods

**Plasmid constructions.** Molecular biology reagents were purchased from Promega (Madison, WI), New England Biolabs (Beverly, MA) or Takara Biomedicals (Otsu, Shiga, Japan). Standard recombinant DNA procedures were used, and whole-plasmid PCR-based mutagenesis reactions were carried out using *QuickChange* Site-Directed Mutagenesis Kit (Stratgene, La Jolla, CA). Oligonucleotide primers were synthesized by Genetech Biosciences (Tempe, AZ) and dye terminator DNA cycle sequencing was done at the Arizona State University DNA facility (Tempe, AZ).

Construction of pRTL2/BXcmyc was carried out in a two-step process. First, whole-plasmid PCR-based mutagenesis was performed with template prRTL2/myc/BX, which has an in-frame *BanHI* and *XbaI* site found by a single copy of the *myc*-epitope (donated by Robert Mullens, University of Guelph, Guelph, Ontario, Canada), as template DNA. A forward primer (5'-CTGAGAAGATCTGTTGAAGATCCGCAAATAC-3') and a reverse primer (5'-GATTTTGTCGGAGACCTCCTAGTCGAGA-3') introduced a termination codon at the 3' end of the *myc* coding sequence and deleted *BanHI* and *XbaI* sites upstream of the *myc* coding sequence, and deleted the *myc* translation initiation codon to yield prRTL2/BXcmyc.

An expressed sequence tag (EST) of the *Arabidopsis* gene encoding PEX3 (AtPEX3) (form 2) (pU21561) on chromosome 1 of the Columbia ecotype was obtained from the *Arabidopsis Biological Resource Center* (Ohio State University). There was no available EST clone for this gene encoding PEX3 (AtPEX3) (form 2) (pU21561). For construction of pRTL2/myc-BXc, a forward primer (5'-CGACGATATCGCGATCCTTCTATGAGAACAAAGTTGATTTAC-3') and a reverse primer (5'-GAAATCACAATTTTGTCTCAGGATTGCGGTCCGTAGTTG-3') that introduced *BanHI* and *XbaI* sites upstream of the *myc* coding sequence, and deleted the *myc* translation initiation codon to yield prRTL2/BXcmyc.

DNA fusion constructs prRTL2/40AtPEX3-CAT, prRTL2/BXc-AtPEX3, a forward primer (5'-CGCGGATCCGATTTCGTTAGGGTTTGGTTCCGCGGATCCGGGTGTT-3') that replaced the termination codon with an in-frame *BanHI* site, and a reverse primer (5'-GCTCTAGAGCTTTGATGCGGTCCGTAGTTG-3') that introduced an in-frame *XbaI* site at the 3' end of the stop codon. For construction of prRTL2/myc-BXc-AtPEX3, a forward primer (5'-CGCGGATCCGATTTCGTTAGGGTTTGGTTCCGCGGATCCGGGTGTT-3') that replaced the termination codon with an in-frame *XbaI* site was used in the PCR. The resulting PCR products from both reactions were TA-cloned into pCR2.1 (Invitrogen, San Diego, CA), digested with *BanHI* and *XbaI*, and then ligated into *BanHI* and *XbaI* digested pRTL2/myc/BX or pRTL2/BXcmyc to yield prRTL2/myc-BXc-AtPEX3 and prRTL2/myc-AtPEX3.

DNA fusion constructs prRTL2/40AtPEX3-CAT, prRTL2/BXc-AtPEX3-CAT, and prRTL2/8AtPEX3-CAT (Table 1) coding for proteins with the N-terminal 40, 16, and 8 amino acid residues of AtPex3p were fused to the N terminus of the passenger protein chloramphenicol acetyltransferase (CAT) and were made as follows. Construction of prRTL2/SmaI-CAT with an in-frame *SmaI* site followed by the CAT coding sequence was described elsewhere [35]. PCR amplification of the 5'-120 base pairs of AtPEX3 in pU21561 included a forward primer.
(5'-GTCCCCGGGATGGATCTTGGGTTTTG3') that introduced an in-frame Smal restriction site and a reverse primer (5'-GGTGGAGGGGTTTTTGGGAGAAA AAAATCACTGG-3') that introduced in an in-frame Smal restriction site.

**Table 1**

<table>
<thead>
<tr>
<th>Protein ID</th>
<th>Amino acid sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>mhc</td>
<td>MEGKLISEEEDL</td>
</tr>
<tr>
<td>CAT</td>
<td>EKKIT...</td>
</tr>
<tr>
<td>nmtcAtPEX3p</td>
<td>MEQKLISEEEDLSDFVRGFWRKHRRKVLYAGCCGLSOGYLLKLYNSHTRLA...</td>
</tr>
<tr>
<td>cmmtcAtPEX3</td>
<td>...NMPEQKLISEEEDL</td>
</tr>
<tr>
<td>40tAtPEX3p-CAT</td>
<td>MDFVGRFWRKHRRKVLYAGCCGLSOGYLLKLYNSHTRLAPGEEKIT...</td>
</tr>
<tr>
<td>16tAtPEX3p-CAT</td>
<td>MDFVGRFWRKHRRKVLYAGCCGLSOGYLLKLYNSHTRLA...</td>
</tr>
<tr>
<td>8AtPEX3p-CAT</td>
<td>MDFVGRFWRKHRRKVLYAGCCGLSOGYLLKLYNSHTRLA...</td>
</tr>
<tr>
<td>nmtcAtPEX3p-RKRRRA6G</td>
<td>MEQKLISEEEDLSDFVRGFWRKHRRKVLYAGCCGLSOGYLLKLYNSHTRLA...</td>
</tr>
<tr>
<td>nntcAtPEX3p-RKRRRA6G</td>
<td>MEQKLISEEEDLSDFVRGFWRKHRRKVLYAGCCGLSOGYLLKLYNSHTRLA...</td>
</tr>
<tr>
<td>nntcAtPEX3pATMD</td>
<td>MEQKLISEEEDLSDFVRGFWRKHRRKVLYAGCCGLSOGYLLKLYNSHTRLA...</td>
</tr>
<tr>
<td>nntcAtPEX3pF3ALF7A</td>
<td>MEQKLISEEEDLSDFVLGRGLWRKHHRRKVLYAGCCGLSOGYLLKLYNSHTRLA...</td>
</tr>
</tbody>
</table>

The N-terminal most TMD1, amino acid residues 14–32 (Tmpred) is highlighted black and the cluster of basic residues is gray.

Biologic bombardment, immunolabeling, and non-confocal epifluorescent, and laser-scanning confocal microscope analyses of Arabidopsis and BY-2 cells. Methods employed for the above procedures (techniques) have been described in detail in other papers from our laboratory [28,36,37]. Briefly, Arabidopsis var. Landsberg erecta and Nicotiana tabacum L., cv. Bright Yellow 2 (BY-2) suspension-cultured cells (4 days post-subculture) were transformed transiently via biolistic bombardment. Cells were harvested 2 and 5–6 h post-bombardment and fixed for 1 h in 4% (w/v) formaldehyde in Arabidopsis or BY-2 transformation buffer. Cell walls of both cell types were perforated and partially digested in 0.1% (w/v) Pectolyase Y-23 (Seishin Pharmaceutical, Tokyo), and Arabidopsis cells also were incubated in 0.1% Cellulase RS (Karan Research Products, Santa Rosa, CA) in phosphate-buffered saline (PBS) for 2 h at 30°C to help minimize clumping of the Arabidopsis cells.

Using our standard 1-mL tube method [36], cell membranes were permeabilized with 0.3% (v/v) Triton X-100 (Sigma–Aldrich) in PBS for 15 min. To determine the in vivo topology of transiently expressed myc-epitope-tagged AtPex3p in organellar membranes, BY-2 cells were permeabilized with digitonin, rather than with Triton X-100, following methods previously described [35,38]. Briefly, BY-2 cells (4 days post-subculture) were suspended in an equal volume of 2× growth medium (without 2,4-D) for 15 min were spread onto filter papers moistened with 2× growth medium for microprojectile bombardment. Bombarded cells were fixed and treated as described above, and then differentially permeabilized with digitin (25 μg/mL) (Sigma). Formed, fixed, and permeabilized cells were incubated for 1 h each with primary and secondary antibodies in PBS following procedures described previously [35,36]. Antibody sources and concentrations used were as follows: rabbit anti-cottonseed catalase IgGs at 1:500 (v/v) or 1:1000 (v/v) (Kunce et al., 1998), mouse anti-myc monoclonal antibody at 1:500 (v/v) (purified 9E10, Covance, Berkeley, CA), mouse anti-CAT (chloramphenicol acetyltransferase) monoclonal antibody (undiluted culture medium) at 1:1 (v/v) (provided by S. Subramani, University of California, San Diego, CA), mouse anti-HA epo monoclonal antibody at 1:500 (v/v) (Boehringer–Mannheim, Indianapolis, IN), mouse anti-n-tubulin monoclonal antibody 1:500 (Amsersham, Arlington Heights, IL), goat anti-rabbit Cy5 at 1:500 (v/v), goat anti-rabbit rhodamine 1:1500 (v/v), and goat anti-mouse Cy2 1:500 (v/v) (Jackson Immuno Research Laboratories, Westgrove, PA).

The cells were observed and analyzed using two different microscope imaging setups. A Zeiss Axioplan 100 epifluorescence microscope (Carl Zeiss, Thornwood, NY) equipped with a CoolSnap ES CCD camera (Roper Scientific, Tucson, AZ) was used for Fig. 6. Laser confocal images were collected using a Leica DM RBE microscope equipped with Leica TCS NT scanning head (Leica, Heidelberg, Germany). All captured images were adjusted for brightness and contrast.
Results

Two putative peroxin 3 proteins in *Arabidopsis* (AtPex3p) were identified through a protein–protein BLAST (blastp) search using sequences of yeast and mammalian Pex3p homologs. The two forms were designated as forms 1 and 2. Table 2 lists all of the putative Pex3p homologs and compares their predicted molecular masses (kDa) and the percent identities of their primary amino acid sequences relative to AtPex3p forms 1 and 2. The two *Arabidopsis* forms are 72% identical. Remarkably low percent identities to form 1 and form 2 are apparent for all three known mammalian Pex3 proteins (22–25% and 19–25%, respectively) and for the four yeast Pex3 proteins (19–20% and 15–21%, respectively). Interestingly, the mammalian and yeast Pex3 proteins exhibit a relatively high degree of identity among themselves as evidenced by the narrow range (15–25%) of identity to the *Arabidopsis* Pex3p forms.

Alignments of all Pex3 proteins revealed that each possessed a cluster of basic amino acid residues within 10 residues of their N terminus. Fig. 1 shows an amino acid sequence alignment of selected portions of the N-terminal regions of the two putative *Arabidopsis* forms and the human and *Hansenula* Pex3 proteins. The latter two are included in Fig. 1 because they have the highest percent identities to the two *Arabidopsis* forms of the mammalian and yeast species. Moreover, published data are available on their putative targeting signals within the N-terminal regions of these proteins (see Discussion). All of the mammalian and plant Pex3 proteins (examples in Fig. 1) have a conserved sequence (K/R, K/R, H, K/R, K/R, and K) of six residues within the cluster that is located 8–9 residues from the N terminus. On the other hand, yeast Pex3 proteins have a recognizable, but varied, cluster of basic amino acid residues ranging 4–6 amino acid residues in length, within 4–9 residues from the N terminus. Clusters in all Pex3 proteins are nearby or adjacent to a predicted transmembrane domain (TMD). Computer programs predicted the presence of two membrane spanning helices or TMDs for both AtPex3 proteins. Fig. 1 shows that the first of two predicted TMDs, TMD1, is positioned at or near amino acid residues 14–32, and that the second TMD, TMD2, occurs at amino acid residues 121–143. Computer predictions indicate that the N terminus of

<table>
<thead>
<tr>
<th>Organism</th>
<th>NCBI Accession</th>
<th>kDa</th>
<th>% Identity</th>
<th>% Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Arabidopsis thaliana</em></td>
<td>NP_566598</td>
<td>42</td>
<td>100</td>
<td>72</td>
</tr>
<tr>
<td>(form 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Arabidopsis thaliana</em></td>
<td>AAM70527</td>
<td>42</td>
<td>72</td>
<td>100</td>
</tr>
<tr>
<td>(form 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Homo sapiens</em></td>
<td>P56889</td>
<td>42</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td><em>Mus musculus</em></td>
<td>Q9QXY9</td>
<td>42</td>
<td>22</td>
<td>20</td>
</tr>
<tr>
<td><em>Rattus norvegicus</em></td>
<td>Q9JKK4</td>
<td>42</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td><em>Pichia pastoris</em></td>
<td>Q92262</td>
<td>52</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td><em>Hansenula polymorpha</em></td>
<td>AAC49471</td>
<td>52</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td><em>Yarrowia lipolytica</em></td>
<td>AA033094</td>
<td>48</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td><em>Saccharomyces cerevisiae</em></td>
<td>NP_010616</td>
<td>48</td>
<td>19</td>
<td>15</td>
</tr>
</tbody>
</table>

Fig. 1. N-terminal sequence alignment of predicted Pex3p homologs. The *H. polymorpha* Pex3p (HpPex3p) and *Homo sapiens* Pex3p (HsPex3p) sequences were aligned with the putative *Arabidopsis* Pex3p (AtPex3p) forms 1 and 2 homologs using the ClustalW 1.82 program. The conserved clusters of basic residues highlighted in gray, and the putative TMDs in black, were predicted using TmPred (EMBnet-CH) prediction program.
the AtPex3p proteins faces the inside of an organelle/membrane compartment.

The full-length EST clone of AtPex3p (form 2) was subcloned into our plant expression vector pRTL2. The DNA encoded for an N-terminal myc-epitope tag (pRTL2/myc-AtPEX3, Table 1) and was expressed transiently for varying times in Arabidopsis and BY-2 suspension-cultured cells. The subcellular localization of the gene product, myc-AtPex3p, was examined in both cell types using immunofluorescence microscopy. Controls such as omission and substitution of primary antibodies were included (data not shown). Figs. 2A–D and G–J show that after 2 and 5 h post-bombardment, myc-AtPex3p colocalizes with endogenous peroxisomal catalase in both cell types. Peroxisomes in Arabidopsis cells are characteristically larger and fewer in number per cell than those in BY-2 cells [36,39]. Notably, at 2 h post-bombardment, myc-AtPex3p is observed solely in peroxisomes, indicative of direct sorting to peroxisomes. Examples of an indirectly sorted protein are shown in Figs. 2E, F, K, and L. Hemagglutinin (HA) epitope-tagged peroxisomal membrane ascorbate peroxidase (HA-APX) was expressed in Arabidopsis and BY-2 cells.

In both cell types (Figs. 2E and K), HA-APX is present in two intracellular immunofluorescent compartments, one observed as punctate structures and the other as reticular structures. Punctate HA-APX is colocalized with punctate catalase (Figs. 2F and L), whereas the reticular HA-APX is not colocalized with endogenous catalase (compare Figs. 2E with F and K with L). Taken together, these time-course results show that myc-AtPex3p sorts within 2 h (the earliest time that it is detectable) directly to peroxisomes.

Fig. 3 shows the targeting of modified chloramphenicol acetyltransferase (CAT) proteins transiently expressed in Arabidopsis and BY-2 cells. CAT-SKL with a C-terminal serine–lysine–leucine (SKL), a prototypical type 1 peroxisomal targeting signal 1 (PTS1), was employed as a control to demonstrate targeting of a CAT protein directly to peroxisomes in both cell types. Figs. 3A, B, I, and J show a punctate colocalization in a CAT-SKL-transformed cell with endogenous catalase in both cell types. The 40 N-terminal amino acid residues of AtPex3p contain a cluster of six basic amino acid residues (RKHRRK) and the N-terminal-most TMD (TMD1) (Table 1). Figs. 3C, D, K, and L show that these 40 N-terminal amino acid residues of AtPex3p appended to the N terminus of CAT are sufficient to target CAT to peroxisomes in both cell types. Comparison of the peroxisomal images in transformed cells
and in nearby non-transformed cells in the same picture frame (e.g., Figs. 3D and L) shows that peroxisomes in \textit{Arabidopsis} cells are aggregated. Such aggregation, more pronounced in \textit{Arabidopsis} cells, is commonly observed and has been reported previously [38].

To further define possible sufficient targeting residues of AtPex3p, the gene coding for 40AtPex3p-CAT was truncated to produce genes coding for 16AtPex3p-CAT and 8AtPex3p-CAT, which eliminated the TMD1 only and both the TMD1 and RKHRRK, respectively (Table 1). Figs. 3E, F, M, and N show that 16AtPex3p-CAT does not target to peroxisomes in either cell type. Instead, the fusion CAT protein mislocalizes to the cytosol in \textit{Arabidopsis} cells (Fig. 3E) and to an unidentified non-peroxisomal punctate structure in BY-2 cells (Fig. 3M). The results indicate that the 8 N-terminal residues plus the RKHRRK cluster are not sufficient (without the TMD1) to target CAT to peroxisomes (Figs. 3E, F, M, and N). The fusion protein, 8AtPex3p-CAT, as expected also mislocalizes to the cytosol (Figs. 3G and O) in both cell types, indicating that the 8 N-terminal residues of AtPex3p are not sufficient for targeting CAT to peroxisomes. Note that there is no apparent aggregation of peroxisomes in \textit{Arabidopsis} and BY-2 cells expressing 16AtPex3p-CAT or 8AtPex3p-CAT (Figs. 3F, N, H, and P, respectively). These results provide further evidence that the CAT fusion proteins did not end up in the peroxisomal membranes (see Discussion).

Fig. 4 shows subcellular localizations of altered versions of \textit{nnmyc-AtPex3p} (form 2) transiently expressed in \textit{Arabidopsis} and BY-2 cells. To ascertain the necessary role of the amino acid residues within the entire basic cluster (RKHRRK), the residues were changed to glycines (GGGGGG) and the resulting proteins (\textit{nnmyc-AtPex3p-RKHRRK}\textsubscript{6G}, Table 1) were expressed in both cell types (Figs. 4A and G). This mutation had the effect of \textit{nnmyc-AtPex3p-RKHRRK}\textsubscript{6G} mislocalizing to the cytosol in both \textit{Arabidopsis} and BY-2 cells. To test the hypothesis that three of the six basic residues (RRK) preceding the TMD1 were essential for targeting, they
were changed to glycine residues and the resulting protein (\(\text{nmyc-AtPex3p-RKHRRKD}3G\), Table 1) was expressed in both cell types (Figs. 4C and I). In Arabidopsis cells, \(\text{nmyc-AtPex3p-RRRK}3G\) mislocalized to the cytosol (not colocalized in Figs. 4C and D), indicating that the RRK residues of RKHRRK are necessary for targeting AtPex3p to peroxisomes. In contrast, \(\text{nmyc-AtPex3p-RRRK}3G\) in BY-2 cells sorted to peroxisomes (colocalized in Figs. 4I and J), indicating that RRK residues in this context were not necessary for targeting in BY-2 cells. To determine the necessity of TMD1 for targeting of AtPex3p to peroxisomes, transformed cells expressing \(\text{nmyc-AtPex3p-RKHRRKD}6G\) (Table 1) were examined (Figs. 4E and K). \(\text{nmyc-AtPex3p-DTMD}\) mislocalized to the cytosol in both cell types, indicating the necessity of the TMD1 adjacent to the basic cluster of residues.

Comparisons of the amino acid residues 1–14 within the N terminus between forms 1 and 2 reveal that only two amino acid residues differ within this region (compare forms 1 and 2 in Fig. 1). To test the hypothesis that the same N-terminal 14 residues would target both AtPex3p forms to peroxisomes, the two different amino acid residues in \(\text{nmyc-AtPex3p-RKHRRK}3G\) (C,I) and \(\text{nmyc-AtPex3p-DTMD}\) (E,K), respectively. All constructs, except RRK3G in BY-2 cells (I), are mislocalized to the cytosol. Confocal projection micrographs. Bar = 5 \(\mu\m).
Arabidopsis and BY-2 cells and colocalizes with endogenous catalase (Figs. 5B and D). These results indicate that the N terminus of AtPex3p (form 1) is sufficient to target AtPex3p (form 2) directly to peroxisomes in both cell types.

Lee et al. [35] reported that treatment of BY-2 cells with digitonin (25 μg/mL) selectively permeabilized the plasma membrane versus the peroxisomal membrane. Our normal treatment of cells with Triton X-100 permeabilizes all cellular membranes. Fig. 6 shows BY-2 cells transiently express N- or C-terminal myc epitope-tagged AtPex3p (form 2) in differentially permeabilized BY-2 cells. Figs. 6A, B, G, and H illustrate transformed cells treated with Triton X-100. Both epitope-tagged versions of AtPex3p (nmyc-AtPex3p and cmyc-AtPex3p, Table 1) (Figs. 6A and G) are colocalized with endogenous peroxisomal catalase (Figs. 6B and H). Figs. 6C, D, I, and J show digitonin-treated cells from the same batch of bombarded cells. Fluorescence labeling attributable to nmyc-AtPex3p (Fig. 6C), cmyc-AtPex3p (Fig. 6I), or to endogenous catalase (Figs. 6D and J) is not observed in the cells. These results indicate that the applied anti-catalase and anti-myc antibodies do not reach their antigenic sites within the peroxisomal matrix because the peroxisomal boundary membranes were not detergent permeabilized. The applied antibodies reaching the cytosol outside of the peroxisomes are shown in Figs. 6E, F, K, and L. Bombarded cells treated with digitonin were co-immunolabeled for α-tubulin and peroxisomal catalase. Immunofluorescent α-tubulin in the cytoplasm is clearly observed (Figs. 6E and K), whereas peroxisomal catalase is not detected (Figs. 6F and L), indicating again that the peroxisomal membranes are not permeabilized with digitonin. These data are interpreted to indicate that the myc-epitope-tagged N and C termini of AtPex3p (form 2) are inaccessible to applied myc antibodies that reach in the cytoplasm of digitonin treated cells, thus revealing that both termini are facing the peroxisomal matrix. Taken together with the predictions that AtPex3p (forms 1 and 2) possess two TMDs within their 143 N-terminal residues (Fig. 1), it is likely that AtPex3p is anchored in the boundary...
membrane by TMD1 with residues 33–121 facing the cytosol, and TMD2 spanning the peroxisomal membrane again such that the majority of AtPex3p and its C terminus reside in the peroxisomal matrix.

Discussion

It is well-established that peroxisomal membrane and matrix proteins are synthesized on free ribosomes and then are sorted post-translationally either directly or indirectly from the cytosol to pre-existing peroxisomes [13,21]. Sorting of PMPs via both pathways is evident in plant cells. For example, Murphy et al. [27] show in their immunofluorescence microscopy study that Arabidopsis PMP22 sorts directly to BY-2 cell peroxisomes. On the other hand, Mullen et al. [37,38] and Lisenbee et al. [36,39] present compelling evidence in their microscopy and cell fractionation studies that cottonseed peroxisomal APX sorts indirectly through pER to Arabidopsis and BY-2 cell peroxisomes.

Pex3 proteins sort to peroxisomes through different intracellular pathways depending upon the organism. For example, in several yeast species, Pex3p is reported to sort to peroxisomes via the ER. Bascom et al. [32] immunodetected overexpressed myc-epitope-tagged Yl-Pex3p in the ER/perinuclear regions of Yarrowia cells. Baerends et al. [30] showed that the 16 N-terminal amino acid residues of HpPex3p appended to a reporter protein directed this fusion protein to the ER and nuclear membranes. Similarly, Salomons et al. [40] observed accumulation of HpPex3p in the ER when Hansenula cells were treated with brefeldin A (BFA), which interferes with the formation of coated vesicles. In contrast, South et al. [41] found no evidence for involvement of ER in the sorting of HsPex3p to human fibroblast peroxisomes. Endogenous HsPex3p was detected only in Nycodenz-gradient isolated peroxisomes, and transiently expressed HsPex3p–myc was immunodetected as early as 1 h post-microinjection exclusively in peroxisomes [41].

Prior to the current study, the sorting pathway had not been elucidated for any plant putative Pex3p homolog or for any other putative peroxin homolog in plants. This statement excludes from consideration Pex5p and Pex7p, which are soluble receptor peroxins. Results of our time-course experiments indicate that nmyc-AtPex3p sorts directly to Arabidopsis and BY-2 peroxisomes. At all post-bombardment time points tested, the epitope-tagged AtPex3p was observed exclusively in peroxisomes. After relatively long post-bombardment time points (e.g., 9 h), nmyc-AtPex3p localized solely to aggregated peroxisomes (colocalized with endogenous catalase, data not shown). This observation is consistent with the localization described for overexpressed Yarrowia lipolytica Pex3p in clustered Yarrowia peroxisomes [32]. At the earliest time that nmyc-AtPex3p could be immunodetected (2 h post-bombardment), it also was found solely in peroxisomes (Figs. 2A, B, G, and H). Peroxisomal APX (HA-epitope tagged APX) served as a positive control to show that a non-peroxisomal sorting compartment, indicative of indirect sorting, could be immuno-identified at an early stage of post-bombardment (Figs. 2E, F, K, and L). Thus, Arabidopsis Pex3p appears to sort directly to peroxisomes as does human Pex3p, rather than indirectly via the ER as do yeast Pex3 proteins. The definitive localization of AtPex3p (form 2) to peroxisomes within cells of two plant species (Arabidopsis and tobacco) supports our hypothesis that AtPex3p (form 2) is a plant peroxin and therefore is a homolog to the other (yeast and mammalian) Pex3 proteins.

The mPTS described herein is the only one that has been described for a plant peroxin or for any Pex3p homolog. As exemplified in Fig. 1, nearly all Pex3 proteins possess within the first 10 amino acid residues a positively charged stretch of residues in close proximity to a hydrophobic region. Kammerer et al. [33] and Ghaedi et al. [34] found that the N-terminal 40 amino acid residues of human and rat Pex3p, which included these regions, were sufficient to target GFP directly to peroxisomes. Our results are consistent with these results, i.e., the N-terminal 40 amino acid residues of AtPex3p, which also possesses both regions, are sufficient to direct a passenger protein, in our case CAT, to peroxisomes (Fig. 3). These data indicate that the mPTS for the plant and mammalian Pex3 proteins is composed of consecutive basic residues nearby a predicted TMD within the N-terminus of these homologs.

Results obtained with yeast Pex3p homologs provide important data for defining the Pex3p mPTS. Weimer et al. [31] found that the 40 N-terminal amino acid residues of Pichia Pex3p (PpPex3p) also were sufficient to target GFP to peroxisomes, but this appended polypeptide did not possess any significant hydrophobic domain. Baerends et al. [30] reported that the 37 N-terminal amino acid residues of Hansenula Pex3p (HpPex3p) were sufficient to target catalase (minus its PTS, thus it served as a passenger protein) to peroxisomes; however, this polypeptide possessed a predicted TMD. Biochemical data from cell fractionation studies on HpPex3p in peroxisomal membranes provided significant insight into these apparent discrepancies. Hann et al. [42] reported that HpPex3p in peroxisomal membranes was not extracted with high salt or alkaline carbonate solutions, but was completely removed from these membranes with urea [43]. Using intact organelles, they found that the HpPex3p was digested with Proteinase K (minus detergent), indicating that HpPex3p did not possess a peroxisomal membrane spanning region(s). Thus, the question arises as to whether an authentic TMD is an essential element of the mPTS for Pex3 proteins?
Knowledge of the topological orientation of Pex3 proteins in the peroxisomal membranes is useful to help answer this question. Interestingly, the topology varies among organisms. The N terminus of ScPex3p and HsPex3p is exposed to the peroxisomal matrix with its C terminus facing the cytosol [31,33], HpPex3p is peripherally bound on the cytosolic face of the peroxisomal membrane [44], the N and C termini of RnPex3p face the cytosol [34], and both termini of AtPex3p face the peroxisomal matrix (Fig. 6). It is likely that both forms of AtPex3p possess the same topological orientation based on the identical hydrophathy predictions between forms 1 and 2, and the identical sequence of amino acid residues of TMD2 within both forms.

Therefore, an authentic TMD may not constitute a necessary element of the targeting signal for all Pex3 proteins. Instead, a hydrophobic domain may be a required element of the mPTS for integral or peripheral Pex3p PMPs. Within this context, the mPTS has not been completely defined for any of the yeast or mammalian Pex3 proteins; sufficient, but not necessary, constituents have been described for a putative signal. Moreover, it is not known whether there is a separate or overlapping signal(s) for targeting of the yeast Pex3p proteins to ER (or nuclear membranes) as has been described for ScPex15p [45] cottonseed APX [28].

The predicted TMD1 of AtPex3p (form 2) was necessary, but not sufficient, for targeting AtPex3p to peroxisomes in both cell types (Fig. 4E, F, K, and L). Our data, however, do not reveal definitively whether AtPex3p is a matrix-facing peripheral PMP or a double membrane-spanning integral PMP with its N and C termini facing the peroxisomal matrix. Nevertheless, we conclude that the N-terminal mPTS for Arabidopsis Pex3p (form 2, and probably form 1) is composed of three or more basic residues and an adjacent predicted/authentic TMD, which is capable of targeting the protein directly to the peroxisomal boundary membrane.

Pex3 proteins seem to function in the early stages of peroxisomal biogenesis; however, the primary role of Pex3 proteins seems to differ within yeasts and mammals. Faber et al. [46] provided evidence that HpPex3p in Hansenula polymorpha participates in the formation of pre-peroxisomal vesicles at the ER/perinuclear region followed by the recruitment of nascent PMPs such as Pex14p. In contrast, South et al. [41] concluded from their studies with inhibitors of COPI and COPII that HsPex3p did not associate with the ER or play any role in potential pre-peroxisomal vesiculation/origination events. Similarities between AtPex3p (form 2) and HsPex3p, such as their mPTS and direct sorting to peroxisomes, may indicate that AtPex3p does not function in pre-peroxisome membrane biosynthesis/recruitment, but functions at the peroxisomal membrane in the assembly of peroxisomal protein import machinery. The availability of appropriate Arabidopsis mutants will be required to test this hypothesis.

Acknowledgments

We thank Dr. Robert Mullen for constructing pRTL2/nmycBX and Dr. Cayle Lisenbee for providing confocal images of BY-2 HA-APX transformed cells. Special thanks are extended to Ms. Sheetal Karnik and Dr. Scott Bingham for their advice and insightful discussions on various aspects of the research. Confocal microscopic studies were performed in the ASU W.M. Keck Biomaging Laboratory. This work was supported by NSF Grant MCB-0091826 and in part by the William and Myriam Pennington Foundation.

References


[18] R.C. Flynn, R.T. Mullen, R.N. Trelease, Mutational analyses of a type 2 peroxisomal targeting signal that is capable of directing...


[43] Y. Elgersma, L. Kwast, M. van den Berg, W.B. Snyder, B. Distel, S. Subramani, H.F. Tabak, Overexpression of Pex15p, a phosphorylated peroxisomal integral membrane protein required for peroxisome assembly in S. cerevisiae, causes proliferation of the endoplasmic reticulum membrane, EMBO J. (1997) 7326–7341.

