Peroxisome biogenesis, protein targeting mechanisms and PEX gene functions in plants

Laura L. Cross a,b,d Heba Talat Ebeed c,1 and Alison Baker a,c,d

a. School of Molecular and Cellular Biology, University of Leeds, Leeds LS2 9JT
b. School of Chemistry, University of Leeds, Leeds LS2 9JT
c. Centre for Plant Sciences, University of Leeds, Leeds LS2 9JT
d. Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds LS2 9JT

1. Permanent address: Botany Department, Faculty of Science, Damietta University, Egypt

Author e-mails: Laura Cross bsltc@leeds.ac.uk
Heba Talat Ebeed fbsheb@leeds.ac.uk; hebeed@du.edu.eg
Alison Baker a.baker@leeds.ac.uk (*Corresponding author)
Abstract
Peroxisomes play diverse and important roles in plants. The functions of peroxisomes are dependent upon their steady state protein composition which in turn reflects the balance of formation and turnover of the organelle. Protein import and turnover of constituent peroxisomal proteins is controlled by the state of cell growth and environment. The evolutionary origin of the peroxisome and the role of the endoplasmic reticulum in peroxisome biogenesis is discussed, as informed by studies of the trafficking of peroxisome membrane proteins. The process of matrix protein import in plants and its similarities and differences with peroxisomes in other organisms is presented and discussed in the context of peroxin distribution across the green plants.

Key words: Peroxin, peroxisome, plant, biogenesis, importomer, membrane

Abbreviations:
AAA; ATPases Associated with various cellular Activities
APEM; Aberrant PEroxisome Morphology
EosFP Eos Fluorescent protein
ERAD; ER Associated Decay
GET; Guided Entry of TA proteins
pER: peroxisomal ER subdomain
PEX; abbreviation for Peroxin, a protein involved in peroxisome formation
PMP; peroxisome membrane protein
PTS1; Peroxisomal targeting signal 1
PTS2; Peroxisomal targeting signal 2
RADAR; Receptor Accumulation and Degradation in the Absence of Recycling
RING; protein domain commonly found in E3 ubiquitin ligases, named for Really Interesting New Gene.
RNAi: RNA interference
TA: tail anchored
T-DNA; Transfer DNA
TMD; Transmembrane domain
TPR; Tetratrico Peptide Repeat
YFP: yellow fluorescent protein
1.0 Introduction

Peroxisomes are single membrane bounded organelles that are found in all major groups of eukaryotes. The initial discovery of a novel type of cellular organelle from rat liver that contained several enzymes involved in the production and degradation of hydrogen peroxide gave the name peroxisomes to these organelles [1]. The study of peroxisomes in plants has a long history. Seminal papers by Harry Beevers established the metabolic route for conversion of fat to carbohydrate in germinating castor bean seeds [2] [3]. These processes were established to be present in organelles termed glyoxysomes [4] which were subsequently shown to contain the enzymes of β-oxidation [5]. Already this work raised questions of how reducing equivalents were balanced and how succinate could be exported to mitochondria. Tolbert and co-workers studied photorespiration in peroxisomes from leaves of C3 and C4 plants and this work too made it apparent that transport of metabolites between peroxisomes, chloroplasts and mitochondria was essential (Figure 1) [6]. Although the importance of transport into and out of peroxisomes was appreciated early on, many years would pass before the membrane proteins responsible for metabolite transport would be identified and characterised. Indeed this is still an active area of research with some transport steps still unaccounted for [7, 8].

Plant peroxisomes show a high degree of tissue specialization and at least four distinct types of this organelle have been described; undifferentiated plant peroxisomes, glyoxysomes which are enriched with enzymes of the fatty acid oxidation and the glyoxylate cycle, leaf peroxisomes, present in photosynthetic tissues, are specialized in the metabolism of glycoglate and host many of the enzymes necessary for photorespiration and the root nodules peroxisomes. Despite this diversity of functions it is now appreciated that all these organelles share common biogenetic mechanisms and are now generically referred to as peroxisomes [9].

It is also clear that the overall composition of proteins within the peroxisome determines its function. The peroxisome does not possess its own genome and acquires matrix proteins from the cytosol, through an import mechanism which allows fully folded proteins to enter the matrix of the peroxisome. The enzymatic content is further remodelled by degradation at the level of both individual enzymes and the whole
organelle (Young and Bartel this volume). This review will focus on mechanisms behind protein targeting to the plant peroxisome membrane and the subsequent import of cargo proteins from the cytosol to the peroxisomal matrix; a process that relies on peroxisome membrane proteins (Figure 1). Proteins which are involved in peroxisomal processes are known as peroxins. Evolutionary origins of these peroxins will be discussed, along with the distribution of fundamental peroxins in plants.

2.0 Biogenesis of peroxisomes in plants

Early models of peroxisome biogenesis invoked an origin from the endoplasmic reticulum [10] and electron microscopic observations that peroxisomes were sometimes seen closely associated with endoplasmic reticulum (ER) (e.g. [11]) were often used as an argument to support this concept. However whether such association also reflected luminal continuity and a biogenetic relationship was hotly debated, as was the significance of reports of glycoproteins in peroxisomes and similarities in membrane protein and phospholipid composition between ER and peroxisomal membranes (reviewed in [12]). Following the discovery of a specialised peroxisome protein import machinery and the ability of peroxisomes to divide, the field as a whole moved from the idea that peroxisomes derived from the ER to the concept that they imported proteins post translationally, but with the capacity to import some components, lipids certainly and possibly some membrane proteins from the ER (reviewed in [13]).

Electron microscopy fails to capture the dynamic nature of peroxisomes. With the advent of live cell imaging it has become apparent that plant peroxisomes move actively on the actin based cytoskeleton [14] [15] and undergo changes of shape, extending tubules termed peroxules [16] particularly under conditions of oxidative stress that generate hydroxyl radicals [17]. Extension of peroxules often reflects underlying ER dynamics, with the peroxules appearing to extend along ER tubules [17]. To address the question of whether there is direct physical connection between ER and peroxisomes in Arabidopsis cells, peroxisomes were labelled with the photo-switchable red-to green EosFP in the apm1 mutant background. The mutants are defective in DRP3a required for peroxisome division [18] and display highly elongated peroxisomes that also mirror underlying RFP marked ER [19]. However despite the close and frequent contact between peroxisomes and ER no evidence for luminal connection was observed [19]. Nevertheless it seems possible that membrane contact sites between ER and
peroxisomes could exist which might aid delivery of membrane lipids and potentially some membrane proteins.

3.0 Sorting of Peroxisome Membrane Proteins (PMPs)

3.1 The PMP import machinery

The machinery for sorting membrane proteins to peroxisomes comprises 3 peroxins, PEX3, PEX19 and PEX16. In *S.cerevisiae* mutants of *pex3* and *pex19* lack peroxisomes, and peroxisomal membrane proteins (PMPs) are either degraded or mis-localised. In mammalian cells an additional peroxin, PEX16, which is absent in *S.cerevisiae*, is also essential for PMP insertion and *pex3, pex16* and *pex19* mutants all lack peroxisomes. In both yeast and mammalian cells, transformation of these mutants with the corresponding wild type gene is sufficient to restore membrane and matrix protein import, and therefore to reconstitute functional peroxisomes. These observations are one of the strongest arguments for the formation of peroxisomes *de novo* from the ER, and indeed PEX3 and PEX19 in Saccharomyces, and PEX3,19 and 16 in mammals all play a role in the formation of peroxisomes from the ER. However PEX3, PEX16 and PEX19 also have roles in direct targeting of membrane proteins to peroxisomes (reviewed in [20] and other articles in this volume). PEX3, PEX16 and PEX19 homologues can be found across the green plants, however PEX16 is absent in diatoms (Table 1). A schematic diagram showing PMP targeting pathways based on composite data from different systems is shown in Figure 1 top panel.

3.1.1 PEX19

PEX19 isoforms are present in wheat, Arabidopsis, moss (*Physcomitrella patens*) and the diatom *Phaeodactylum tricornutum* (Table 1). In Arabidopsis there are 2 isoforms of PEX19 (Hadden et al., 2006). Down regulation by RNAi gave different phenotypes compared to those seen in the equivalent mammalian or yeast mutants [21]. In *pex19-1i* and *pex19-2i* plants where the individual genes were knocked down independently slightly enlarged peroxisomes that contained a matrix targeted GFP reporter were observed, but in each mutant the other isoform was expressed, which may explain why a severe phenotype was not observed. Biochemical characterisation of AtPEX19-1 showed that, like PEX19 in other systems, it is a predominantly soluble protein with only a small proportion associated with peroxisome membranes. Recombinant AtPEX19-1 could bind AtPEX10 *in vitro*, consistent with its proposed receptor/chaperone role in
yeasts and mammals [22]. At PEX19-1 and At PEX19-2 could also bind the Arabidopsis peroxisomal ABC transporter COMATOSE [23] which was correctly targeted to peroxisomes when expressed in S.cerevisiae [23, 24]. The human ABCD1-3 transporters ALDP, PMP70 and ALDPR were correctly targeted to peroxisomes in tobacco cells and could be mis-targeted to the nucleus by co-expression with Arabidopsis PEX19-1 to which a nuclear localisation signal had been appended [25]. Together these experiments show that despite the large evolutionary distance, recognition and delivery of PMPs is conserved. However AtPEX19 could not complement the S.cerevisiae pex19∆ mutant, indicating that not all PEX19 functions are conserved [23]. Perhaps this is not surprising given that in S.cerevisiae PEX19 has multiple functions. As well as its role in direct import of membrane proteins, its interaction with PEX3 at the ER is important for the initiation of de novo peroxisome biogenesis (Figure 1 top panel). It is required for the formation of vesicles containing PEX3 from the ER, it interacts with the myosin myo2p and also with PEX14 of the docking translocation machinery (DTM) for matrix protein import (reviewed in [26]).

AtPEX19 appears to form a disulphide bonded dimer both in vitro and in vivo, and in vivo cross-linking showed that it was the dimeric form of PEX19 that preferentially binds cargo proteins [22]. While this property has not been reported for PEX19 of other organisms, it is intriguing to note that the PTS1 receptor PEX5 has also been shown to form a disulphide bonded dimer and it was proposed that the disulphide bonded form had higher affinity for PTS1 cargo [27]. A number of parallels can be drawn between the functions of PEX5 (the receptor for import of matrix proteins, Section 4) and PEX19. Both proteins shuttle between cytosol and peroxisomes. Both have a folded C terminal domain that binds to cargo proteins [28] [29] and an unstructured N terminal domain that interact with docking components; PEX14 in the case of PEX5 [30, 31] and PEX3 in the case of PEX19 [32] [33, 34]. Perhaps we can speculate that redox regulation of both matrix and membrane protein import is one means by which peroxisomes can regulate their activities in concert with metabolism [35].

3.2 PEX3

PEX3 is the membrane bound receptor for PEX19 [32] [33, 34] and is present across the exemplar species surveyed in Table 1 and is encoded by 2 isoforms in Arabidopsis. AtPEX3-2 is proposed to have both N and C termini in the peroxisome [36] but it is difficult to see how that can be the case, given that in other systems the protein has a
single membrane transmembrane segment with N terminal facing the peroxisome lumen while the C terminus is required to be exposed to the cytosol in order to interact with PEX19 (Figure 1A). The deduced topology of AtPEX3-2 was based on the inability to detect N and C terminal epitope tagged versions of AtPEX3-2 when cells were permeabilised with digitonin, which selectively permeabilises the plasma membrane [37] Although this method can be very useful in determining topology, it is prone to false negatives if the protein folds or interacts with other proteins such that the epitope tag is obscured.

PEX3 is considered to be one of the ‘early’ peroxins that is generally accepted to sort from the ER to peroxisomes in mammalian and yeast systems [20] (Figure 1 top panel iii). N terminally myc tagged AtPEX3-2 sorted rapidly to peroxisomes within 2 h when transiently expressed in either Arabidopsis or tobacco suspension cells and was not detected in ER or pER (peroxisomal ER subdomain). [36]. This was interpreted as evidence for direct transport of PEX3 to peroxisomes in plants (Figure 1 top panel ii). However 2h may be too long a time frame to see ER intermediates as ScPex3p was completely localised to peroxisomes within a 45-60 min chase period [38]. The 40 N terminal amino acids of AtPEX3-2 (including the conserved basic cluster and predicted N terminal transmembrane domain (TMD)) could target chloramphenicol acetyl transferase to peroxisomes in both Arabidopsis and tobacco. Deletion and substitution analysis showed the importance of these sequence features [36]. Similarly, the N terminal 45 residues of S.cerevisiae Pex3 were required for peroxisomal targeting (via the ER) [39, 40]. More detailed analysis showed that the TMD and 6 charged residues of the cytosolic domain of ScPex3p are needed for ER targeting while the N terminal luminal part is needed to reach pER and peroxisomes, since removal leads to vacuolar localisation. Within this region the two positively charged regions act redundantly for peroxisomal sorting [38].

In the Arabidopsis pex3i mutant both PEX3 genes were simultaneously down regulated to <20% of wild type expression level but the mutants were not dependent on sucrose for post germinative growth and remained sensitive to 2,4DB. This compound is bioactivated to 2,4D (a herbicide) by β-oxidation, and so retention of sensitivity in mutants is an indication that the peroxisomes retain a functional beta oxidation pathway. The pex3i double knock down had large elongated peroxisomes [21] which is very different to phenotypes of pex3 mutants in mammals and plants. Given the differences
in mutant phenotypes and possible difference in topology, the function of PEX3 in plant systems warrants further study.

3.3 PEX16
In addition to targeting to the peroxisome via the ER in mammals, PEX3 can also insert directly into peroxisome membranes in a PEX19 and PEX16-dependent manner, with PEX16 proposed to function as the membrane receptor [41] (figure 1 top panel ii). However PEX16 is not a universal peroxin as it is absent from S. cerevisiae and P. tricornutum. PEX16 itself traffics to peroxisomes via the ER (Figure 1 top panel iii) (reviewed in [20]). In Arabidopsis a PEX16 homologue (SSE1) was identified through the cloning of the gene defective in the Shrunken SEed 1 (sse1) mutant [42]. As the name suggests this mutant has shrunken seeds due to abnormal deposition of protein and oil bodies during seed development. Despite only low (26% identity) the corresponding SSE1 gene could partially complement the growth of Yarrowia lipolytica pex16 mutant for growth on oleate [42]. In homozygous sse1 embryos the signal from a PTS1 fluorescent reporter was not detected, and a PTS2 reporter labelled small abnormal structures against a diffuse background suggesting normal peroxisomes were absent. SSE1 over expression caused formation of peroxisomal aggregates [43]. As well as these effects on peroxisomes, the sse1 mutant shows reduced fatty acid biosynthesis and production of oil bodies; both ER-dependent processes. The mutant phenotype emphasises the close relationship between ER and peroxisomal metabolic processes, and how disruption of one can impact the other.

SEE1/AtPEX16 itself is targeted to peroxisomes [43] and peroxisomes and ER [44] Transient expression of myc tagged AtPEX16 in Arabidopsis and tobacco suspension cells resulted in co-localisation with ER markers at early time points and with peroxisome markers at late time points. Cold treatment resulted in AtPEX16 accumulation in reticular structures and upon subsequent warming localisation in peroxisomes [45]. A recent study provides evidence that AtPEX16 like its human counterpart can act as a receptor for PMPs [46].

3.2 Other PMPs
3.2.1 PEX10 and PEX2
PEX2, 10 and the third member of the RING complex PEX12 (Section 4.3) are all found in the cross section of plant diversity presented in Table 1. AtPEX2 and AtPEX10 are
most likely involved in receptor ubiquitination (Section 4.3). PEX2 in Yarrowia lipolytica is a glycoprotein and is trafficked to peroxisomes via the ER [47]. However, no evidence could be found for AtPEX2-YFP fusions in ER or pER when the Arabidopsis protein was over expressed in tobacco epidermal cells or the native protein stained by immunofluorescence in suspension culture cells [48]. Similarly AtPEX10-YFP and YFP-AtPEX10 fusions localised to peroxisomes and a diffuse compartment which was distinct from ER and was most likely cytosol. Neither BFA treatment nor use of inhibitory mutants of Sar1 and RabD2a resulted in the relocation of either AtPEX10 or AtPEX2 to the ER, although these treatments had the expected effect on the distribution of a Golgi reporter [48]. While these results were negative, and it is not possible to formally exclude that the proteins trafficked from the ER by a pathway independent of Sar1 or RabD2a and too rapidly for detection at steady state, the simplest interpretation is that AtPEX10 and 2 are inserted by the direct pathway. Contradictory results were obtained when endogenous PEX10 was localised by immunofluorescence in Arabidopsis suspension cultured cells. Staining was only seen in reticular structures and overexpression of tagged constructs resulted in cytosolic labelling [49]. The reason for this discrepancy is unclear. Further cell fractionation experiments reported that AtPEX10 was located in a fraction containing ER markers [49].

3.2.2 PMP22

PMP22 is a non-peroxin peroxisomal membrane protein. It has homology to PxMP2 of mouse which is a channel forming protein in peroxisome membranes that allows the passage of small metabolites [50]. Arabidopsis PMP22 inserts directly into peroxisome membranes in an in vitro system. Insertion is not ATP dependent, but ATP is required to achieve full protease resistance which is interpreted as the achievement of the final functional conformation [51]. A detailed characterisation of the topology and targeting of PMP22 concluded that it spanned the membrane 4 times with both termini in the cytosol [52]. At least 4 regions of the protein were required co-operatively for efficient targeting and insertion into peroxisome membranes. Two regions, a di lysine pair near the N terminus (K7K8) and a region from amino acids 14-26 are similar to targeting signals identified in rat PMP22 [53] while two basic clusters at 49-54 and 82-85 are distinct [52].

3.2.3 TA proteins

Another important class of proteins are the tail anchored (or TA) proteins which are found in all membranes of plant cells [54]. TA proteins as the name suggests are
anchored in the membrane by the C terminus and expose only a short C terminal region to the extra cytoplasmic side of the membrane while the N terminus projects into the cytosol. Examples of TA proteins found in plant peroxisome membranes are ascorbate peroxidase 3 (APX3), mono dehydroascorbate reductase 4 (MDAR4), Fis1a and Fis1b, which recruit the peroxisome division apparatus, and the peroxin APEM9/DAYU which is the membrane anchor for the PEX1/PEX6 complex (Section 4.4). The functional equivalents in humans and fungi have been reported to be targeted via the ER by the Guided entry of TA proteins (Get) pathway [55] (Figure 1 top panel iii) or directly via interaction with PEX19 and PEX3 [56] [57] (Figure 1 top panel i). Conserved hydrophobic segments within the N terminal domain of PEX19 and at the base of PEX3 are proposed to form a relay system of sequential interactions to bring about targeting and insertion of the TA substrate into the peroxisomal membrane [56].

APX3 is one of the best characterised plant PMPs in terms of its trafficking pathway. It is a tail-anchored protein with its N terminal luminal domain exposed to the cytosol where it detoxifies hydrogen peroxide using ascorbate as an electron donor. APX from cotton seed was tagged at the N terminus with a HA epitope tag and expressed transiently in tobacco BY2 cells where it partially co-localised with endogenous catalase by immunofluorescence. However some APX co-localised with DiOC<sub>6</sub> stained membranes described as circular and reticular in appearance, but not with the ER luminal proteins BiP and Calretculin [58]. Native pumpkin APX was also found in both peroxisome and unidentified membranes by immuno-gold electron microscopy and by cell fractionation but was not in rough ER [59]. Native Arabidopsis APX was reported in peroxisomes and rough ER by cell fractionation, but only in peroxisomes by immunofluorescence, unless epitope tagged and over expressed [60]. In vitro assays showed that cotton APX3 could be integrated specifically into microsomal membranes post translationally in an ATP-and chaperone dependent manner [58]. Collectively these results were interpreted as evidence for sorting of APX to a sub compartment of ER distinct from that of ER resident proteins termed the ‘peroxisomal ER subdomain, (pER)’ and the elaboration of a model termed the ‘ER semi-autonomous peroxisome model’ [58] [61].

The targeting signal for cotton APX3 lies within its TM and C terminal domain since the most C terminal 36 residues could sort chloramphenicol acetyl transferase (CAT) to both peroxisomes and circular/reticular membranes [58]. More detailed analysis revealed the
importance of the 8 amino acid luminal tail (RKRMK). The TM domain while important for sorting to peroxisomes and reticular structures could be substituted by an artificial TMD [62]. AKR2A is a soluble protein that has been identified as a protein that interacts with the targeting signal of APX3 [63] but also several other plant TA proteins destined for different membranes. It is known to play a role in import of chloroplast outer envelope proteins [64], thus it is unlikely to confer specificity but has chaperone like activity towards both OEP7 [64] and APX3 [65]. It will be interesting to establish whether plant peroxisomal TA proteins are Get or PEX19 pathway substrates or whether some can use both pathways. In other systems the basis of this selectivity appears to reside in the hydrophobicity of the TMD [56].

3.2.4 Viral proteins
Tombus viruses such as Tomato Bushy Stunt Virus and Cucumber Necrosis Virus hijack peroxisomes for their replication, forming peroxisome-derived multivesicular bodies [66] [67]. The viral encoded p33 protein which is part of the replicase complex contains a signal for PEX19-dependent targeting to peroxisomes in *S.cerevisiae* which can act as a surrogate host for viral replication [68]. Viral replication in *S.cerevisiae* and tobacco is dependent upon Sec39p [69] an ER localised protein that was shown to play a role in ER-to-peroxisome vesicular traffic in *S.cerevisiae* [70]. Over expression of CNVp33 in tobacco resulted in formation of clusters of peroxisomes, similar to those seen in CNV infection [67]. Infection of plants where the ER was labelled with a GFP reporter revealed formation of ER derived vesicles which imported YFP-SKL, however p33 alone could not produce this elaboration of ER membranes [67].

4.0 Peroxisomal matrix protein import
Import of cargo proteins from the cytosol to the matrix of the peroxisome is governed by two pathways, the peroxisomal targeting signal 1 (PTS1)-pathway and the PTS2-pathway (See article by Reumann et al this issue). Although these two pathways begin with distinct receptor-targeting signal binding events, there is co-dependence between them and the separation between the pathways is more ambiguous at all processes downstream from, and including, docking at the peroxisomal membrane [71-73]. PTSs are recognition sequences at the C- (PTS1) or N- (PTS2) terminus of cargo proteins, and these sequences are recognised by specific receptor proteins. The peroxisomal matrix protein import cycle begins with the PTS1/2-cargo protein binding its cognate receptor, followed by docking of this complex at peroxisomal membrane proteins
Translocation of folded cargo is achieved through thermodynamically favourable binding events with PMPs as opposed to ATP hydrolysis. Cargo is released in the peroxisomal matrix and the receptor is recycled to the cytosol for another round of import (Figure 1 lower panel).

4.1 Receptor-targeting signal interactions
There are two peroxins responsible for the recognition of peroxisomal targeting signals (PTSs), peroxin-5 (PEX5) and peroxin-7 (PEX7) (Figure 1 lower panel). PEX7 binds to the PTS2 sequence on cargo proteins to be imported via the PTS2 import pathway. It is important to note here that *S. cerevisiae* PEX7 uses PEX21 and PEX18 as co-receptors where *Arabidopsis* PEX7 uses PEX5 [74, 75]. In rice and humans, there are two splice variants of PEX5 which result in two possible isoforms of the protein: PEX5S and PEX5L [76] [77]. PEX5L contains the PEX7 binding site, so can act as a co-receptor in the PTS2 pathway, whereas PEX5S does not have the ability to bind to PEX7 so can only play a role in the PTS1 import pathway. Arabidopsis, Physcomitrella and wheat appear to have only the PEX5L variant.

The most abundant PTS1 sequences possess the characteristics [small]-[basic]-[hydrophobic]-COOH, but it is becoming clear that this sequence can be quite varied, and can rely heavily on upstream sequence to change the targeting abilities for the PTS1-cargo protein [78-82]. Two novel PTS1 sequences have recently been discovered in *Arabidopsis*, [P]-[S]-[L]-COOH and [K]-[R]-[R]-COOH [83], and another PTS1 sequence that does not fit the consensus, [S]-[A]-[L]-COOH, has been found in Mitogen Activated Protein Kinase (MAPK) phosphatase 1. Interestingly, MAPK phosphatase 1 is only targeted to peroxisomes under conditions of stress, which demonstrates the adaptability of peroxisomes [84]. The probability of a plant PTS1 sequence being targeted to peroxisomes can be predicted using the bioinformatics tool PredPlantPTS1 [81, 85] This has been used alongside *in vitro* binding studies and *in vivo* import experiments and reveals how important the upstream amino acids are in the targeting of the PTS1 sequence.[86] (and see also the article by Reumann et al in this issue). The consensus for a PTS2 sequence is [R/K]-[L/V/I]-X₉-[H/Q]-[L/A] and this sequence is found near the N-terminus of PTS2-cargo proteins [87]. In some organisms, for example the diatom *Phaeodactylum tricornutum*, the PTS2 pathway seems to be completely lost and the PTS1 pathway is responsible for all matrix protein import into peroxisomes (Table 1) [88].
The PTS1 import pathway is the dominant pathway for cargo proteins to enter the plant peroxisome, and the PTS2 receptor PEX7 relies on PEX5, the PTS1 receptor, for efficient PTS2 cargo protein import. This has been found through genetic studies of the two proteins showing that mutation of a site within Arabidopsis PEX5 (Ser318Leu) disrupts PTS2 import but has no apparent effect on PTS1 import \[89\] \[72\]. Mutation of the equivalent residue in PEX5 of Chinese hamster ovary cells (Ser214Phe) also had the same effect of disrupting only PTS2 import \[90\]. It was demonstrated that removal of the first three WxxxF/Y aromatic pentapeptide repeats, implicated in the binding of PEX14 by PEX5, resulted in reduced PTS1 and PTS2 import \[91\] (Figure 2). PTS2 import with both of these protein variants was recovered by the simultaneous expression of a truncated PEX5 variant comprising the N-terminal domain, \(\text{PEX5}^{454}\), which implies that PEX7 relies on the N-terminal domain of PEX5 for PTS2 import \[91\]. The C-terminal domain of PEX5 which is made up of Tetratricopeptide repeats (TPRs) is responsible for PTS1 binding. This has been shown by the expression of only the C-terminal domain of Arabidopsis PEX5. \(\text{In vitro}\) binding studies with this N-terminally truncated variant of PEX5 have demonstrated that binding affinity of a range of PTS1 peptides is effectively unchanged when comparing the full-length protein to the TPR domain alone \[86\].

The receptor for the PTS2 pathway is PEX7. Genetic studies of this protein have shown that PTS2 import is reduced when expression of PEX7 is knocked down \[72, 73, 92\]. Interestingly, a missense mutation in PEX7 (Thr124Ile) results in reduced PTS1 and PTS2 import \[73\] (Figure 2). The first crystal structures for these receptor-targeting signal interactions were solved for Human PEX5C-PTS1 \[93\], and recently for \textit{S. cerevisiae} PEX7-PTS2-PEX21 \[94\]. These have revealed important information regarding the shape of, and the type of interactions occurring at, each of the binding sites. The C-terminal domain of PEX5 binds cargo proteins, via their PTS1 sequence, whereas the N-terminal domain of PEX5 is disordered and is thought to insert into the peroxisomal membrane to aid in cargo protein translocation \[95\]. The disordered nature of the N-terminal domain of PEX5 means that it has only been possible to crystallise the C-terminal domain (PEX5C). The structure of PEX5C-PTS1 reveals that PEX5C is composed of two sets of three TPRs (tetratricopeptide repeats) linked by a hinge region. These two binding faces make up the PTS1-binding site, which is governed by conserved asparagine residues in PEX5C.
making polar contacts with the backbone of the PTS1 sequence. Many other conserved amino acids in the PEX5 binding site are responsible for side chain interactions of the PTS1 sequence, which can be via water molecules [93]. PEX7 is formed of six WD (Trp-Asp) repeats which are arranged in a circular structure to make up the PTS2 binding site. Binding of the PTS2 sequence involves hydrogen bonding and electrostatic interactions of PTS2 with conserved amino acids on loop regions, between the WD repeats, in PEX7. Hydrophobic forces between the PTS2 sequence and both PEX7 and co-receptor PEX21 also look to be important for binding [94].

A phenomenon known as ‘piggyback import’ can occur in the PTS1 and PTS2 import pathway [96, 97]. This is where a protein without a PTS can dimerise with a PTS-cargo protein in order for both proteins to gain access to the peroxisomal matrix. Some of the subunits of an enzyme which seems to be involved in β-oxidation in peroxisomes, protein phosphatase 2A (PP2A), have been found to target to the peroxisome by piggyback import. In this case, one of the subunits of PP2A possesses a PTS1, [S]-[S]-[L]-C terminus, and this sequence targets all interacting subunits of PP2A to the peroxisome [98].

4.2 The peroxisomal membrane docking complex
Cargo-loaded peroxisomal targeting receptors dock at the membrane of the peroxisome (Figure 1). Peroxin-14 (PEX14) and peroxin-13 (PEX13) are the major known docking proteins. It has come to light that PEX14, although playing a major role in facilitating peroxisomal protein import, is not essential for the process in Arabidopsis [99-101] and an obvious PEX14 homologue appears to be missing in diatoms [88] (Table 1). Many of the Arabidopsis PEX14 mutants that have been studied have caused a knockdown of PEX14 protein expression, which results in reduced PTS1 and PTS2 import [100, 101] (Figure 2).

PEX13 is also involved in docking of cargo-loaded receptors. A mutant of PEX13 with a point mutation (Glu243Lys) results in reduced PTS1 and PTS2 import [102] (Figure 2). This site appears to sit within the putative SH3 domain of PEX13 which is thought to be the site at which PEX13 interacts with PEX14. In other studies where PEX13 is truncated or expression is knocked down reduced PTS1 and PTS2 import has also been observed [21, 103] (Figure 2). PEX13 is a crucial protein for cell survival as a T-DNA insertion into the PEX13 gene resulting in a translational frameshift is lethal [104].
Studies of PEX13 have given insight into the order of the import process, for example knockdown of PEX13 expression levels by T-DNA insertion exacerbates the effects of mutation in ‘early acting’ peroxins PEX5 and PEX14. This PEX13 mutation restores the detrimental effects of mutation in ‘late acting’ peroxins PEX1 and PEX6, which suggests that the docking complex, comprised of the ‘early acting’ peroxins, and the receptor recycling complex, comprised of the ‘late acting’ peroxins, may not be directly linked [105].

Interactions between the *Arabidopsis* import receptors and the docking peroxins have been studied *in vitro* using yeast-two hybrid studies, filter binding and pull-down analysis [31, 71, 103]. PEX5 contains WxxxF/Y (Trp-X-X-X-Phe/Tyr) pentapeptide repeats which were found to be crucial in the interaction of PEX5 with PEX14, studied by yeast-two-hybrid with a range of PEX5 constructs. In this study no binding of PEX14 to PEX7 was observed, unlike the mammalian and yeast import systems [71]. Yeast-two-hybrid was also used to investigate binding of *Arabidopsis* PEX13 to PEX5 and PEX7 and it was shown that PEX13 interacts with PEX7 but not PEX5 [103]. Interestingly a PEX13 homologue appears to be absent from diatoms (Table 1) and this fits with the loss of the PTS2 pathway [88].

Exactly how the cargo crosses the membrane is still unclear. PEX5 has the capacity to insert into membranes, and a dynamic transient pore which can reach up to 9 nm in diameter containing PEX5 and PEX14 has been demonstrated to form in the yeast system, which is presumed to allow cargo to enter the matrix of the peroxisome [106]. In the absence of PEX14, it could be postulated that PEX5 can form PEX5-only pores by self-oligomerisation as low levels of PTS1-cargo import are still observed in the absence of PEX14 in both *Arabidopsis* and yeast [100, 107]. It was shown in the mammalian import system that the binding of PEX5 to PEX14 causes release of the non-canonical PTS1 cargo catalase [108] however recent work in *Arabidopsis* has shown that the interaction between PEX5 and PTS1 cargo is unchanged when increasing concentrations of PEX14 are titrated into the mixture [31, 108]. The finding that the PTS1 and PTS2 pathways are co-dependent in *Arabidopsis*, and that *Arabidopsis* PEX14 appears to cause PTS2 cargo unloading in pull-down of the cytosolic fraction of *Arabidopsis* followed by immunoblotting, seems to suggest that the two pathways are more interlinked in plants than in other organisms [31, 73]. However it has been shown
that PTS1 import in Arabidopsis can be selectively targeted by a small molecule inhibitor [109].

When docking and cargo translocation have taken place, the receptor(s) must be recycled to the cytosol for subsequent rounds of cargo import. Ubiquitination of the PTS receptor(s), is an important part of the recycling process and crucial components in this process are largely responsible for the transfer of ubiquitin onto target molecules.

4.3 Ubiquitination machinery for the PTS receptor(s)

In order for multiple rounds of import to be carried out, the PEX5 receptor must be recycled to the cytosol (Figure 1). This process is governed by the monoubiquitination of PEX5 on a cysteine residue (in yeast and mammals). This cysteine is conserved in plant PEX5. There is also the option for PEX5 degradation by polyubiquitination of lysine residues near the N terminus, which directs PEX5 through the RADAR (Receptor Accumulation and Degradation in the Absence of Recycling) pathway [110]. Ubiquitination requires an E1 ubiquitin-activating enzyme, which transfers ubiquitin to an E2 ubiquitin-conjugating enzyme. This allows the transfer of ubiquitin to the target protein in the presence of an E3 ligase [111].

The ubiquitination machinery for receptor recycling in plants has not been extensively studied, however it has been found that the Arabidopsis peroxin-4-peroxin-22 (PEX4-PEX22) complex can restore function of yeast peroxisomes in which PEX4 and PEX22 are deficient [112]. Yeast PEX4 is the E2 ubiquitin-conjugating enzyme which comes into close proximity to PEX12 of the PEX2/10/12 RING-finger (E3 ligase) complex to allow the catalysis of ubiquitin transfer from PEX4 to PEX5 [113, 114]. The function of PEX4 is enhanced by the cytosolic domain of PEX22 [115]. Knockdown of PEX22 expression in Arabidopsis by T-DNA insertion enhances effects caused by PEX4 knockdown [112]. As Arabidopsis and yeast PEX4 and PEX22 seem to be interchangeable, and Arabidopsis PEX22 has also been found to be an enhancer of PEX4 [112], it seems plausible that that the plant ubiquitin machinery for recycling or degradation of PEX5 works in a similar way to the yeast system. A point mutation towards the C-terminus of PEX4 (Pro123Leu) resulted in reduced PTS2 protein processing, which is used as a proxy for import in vivo [112], and knockdown of PEX4 expression by RNA interference results in reduced PTS1 and PTS2 import [21] (Figure 2). Both PTS1 and PTS2 import, therefore, seem to rely on the recycling machinery so perhaps PEX5 is not only a co-receptor for the import
process, but also involved in PTS2 receptor recycling. PEX4 homologues are found across all the species in Table 1. A diatom homologue of PEX22 was not detected so whether and how PEX4 is attached to diatom peroxisomes is unknown.

Peroxin-2 (PEX2), peroxin-10 (PEX10) and peroxin-12 (PEX12) form the RING complex, so named because each of these peroxins contains a RING-finger domain. These peroxins are also broadly distributed in plants (Table 1). The RING-finger domains of the three proteins from *Arabidopsis* have been isolated and studied to elucidate function, and it has been found that all three possess E3 ligase activity [116]. Recent research into the *Arabidopsis* RING complex has revealed that PEX2 and PEX10 seem to have synergistic effects. When there is a point mutation in PEX2 (pex2-1; Arg161Lys), a defect in PTS1 import can be detected. This defect is ameliorated through PEX2 overexpression but not PEX10 overexpression. A truncation mutant of PEX10 (pex10-2) has the same effect as pex2-1 in that PTS1 import is defective. PTS1 import is rescued by overexpression of PEX10, but not by overexpression of PEX2. Intriguingly, the two single mutants pex2-1 and pex10-2 show defective PTS1 import with no effect on PTS2 import, but when the two mutants are combined PTS2 import is also affected. Overall, this shows that the two proteins PEX2 and PEX10 appear to be linked in function but they do not have the same effect in the cell [117].

PEX2 and PEX10 appear to have additional roles beyond protein import into peroxisomes. The RING finger domain of PEX2 (PEX2-RING) has recently been shown to translocate to the nucleus and interact with a transcription factor HY5 when expressed alone. A similar effect is seen when PEX2 contains a missense mutation just upstream of the RING domain in the mutant *ted3* (Val275Met) (Figure 2). It is hypothesised that PEX2-RING can substitute for the E3 ligase complex COP1/DET1, which catalyses the transfer of ubiquitin to positive regulators of photomorphogenesis to target their degradation [118, 119]. PEX10 also has additional functions, aside from acting in protein import, for example its roles in cuticle biosynthesis [120] and organelle adhesion[121]. The PEX2/10/12 RING complex is vital for peroxisome function and therefore cell survival, which has been shown though genetic studies. Mutants of the three proteins (pex2-T-DNA; pex10-1; pex12-T-DNA), in which frame shift results from T-DNA insertion, are all lethal to the plant [119, 122-124]. Truncation of PEX10 (pex10-W313*) also results in a lethal phenotype [121]. Knockdown of expression of the three RING
peroxins by RNA interference (pex2i, pex10i and pex12i) all result in reduced PTS1 and PTS2 import [21] (Figure 2).

Following the ubiquitination of the PTS receptor(s), PEX5 can either be labelled for recycling to the cytosol (cysteine-monoubiquitinated) or labelled for degradation by the proteasome (lysine-polyubiquitinated). It has been reported that the degradation of PEX5 could require DSK2a and DSK2b, proteins which interact with PEX2 and PEX12 and link substrate ubiquitination with proteasomal degradation in other systems [116]. A quality control mechanism also appears to be in place for PEX7, as the GTPase rabE1c has been found to be involved in PEX7 degradation in Arabidopsis [125]. To meet either fate of recycling or degradation, PEX5 must be exported from the peroxisomal membrane. For receptor recycling, a membrane-anchored AAA (ATPases associated with various cellular activities) ATPase complex is required.

4.4 The receptor recycling complex

Three peroxins, peroxin-1 (PEX1), peroxin-6 (PEX6) and aberrant peroxisome morphology-9 (APEM9) are required for the PTS receptor(s) to be recycled back to the cytosol for subsequent rounds of peroxisomal import (Figure 1). The AAA ATPase complex is comprised of PEX1 and PEX6. PEX1 and PEX6 are anchored to the membrane by APEM9 binding to PEX6, and this complex seems to be important for both the PTS1 and PTS2 import cycles in Arabidopsis. This has been found through genetic studies of the three peroxins. When expression of PEX1, PEX6 or APEM9 is knocked down separately (pex1i, pex6i, apem9i), the result is reduced PTS1 and PTS2 import [21, 126] (Figure 2).

APEM9, when disrupted by T-DNA insertion at either the N- or C-terminal portion of the protein causing a frameshift, results in a lethal phenotype. This highlights the importance of receptor recycling to the peroxisomal matrix protein import process. Missense mutations have also been studied in APEM9 and PEX6. In APEM9, a missense mutation (apem9-1; Gly278Glu) results in reduced PTS1 and PTS2 import [126, 127]. In PEX6, a mutation in the C-terminal half of the protein (pex6-1; Arg766Gln) reduces PEX5 levels whereas a mutation in the N-terminal half of PEX6 (pex6-2; Leu328Phe) does not seem to affect PEX5 levels [101, 117, 128] (Figure 2). This could be due to the location of the
mutation, or it could be that the less conservative substitution in \textit{pex6-1} results in a more defective phenotype.

Recent electron micrographs of the PEX1/6 complex from yeast have revealed that this complex is a hexamer comprising a trimer of dimers, with each dimer consisting of PEX1 and PEX6. This work has shown that rotational movement of the complex in response to ATP appears to be responsible for the export of PEX5 from the membrane, and that this export may require partial or complete unfolding of PEX5 [129]. The receptor recycling stage of plant peroxisomal matrix protein import still requires unravelling and it will be exciting to see how this process, and all involved in import, works in plants and how this differs from mammals and fungi.

5.0 Evolutionary Origin and distribution of peroxins in plants

The biogenesis of peroxisomes should reflect their evolutionary origin. The presence of a common set of proteins involved in peroxisomes biogenesis and proliferation supports the idea of a single evolutionary origin of peroxisomes. However, the nature of this evolutionary origin, like the mechanism of biogenesis has been a matter of debate. Another property that seems to be shared by all types of peroxisomes is the division machinery, as it has been shown to be largely conserved in yeast, plant and mammalian peroxisomes and involves, at least, a dynamin-like protein and a TPR (Tetratrico Peptide Repeat)-motif containing protein that serves as a membrane anchor [130]. As discussed in section 2, early models considered that peroxisomes were formed from the endomembrane system [131] based on the close interactions between peroxisomes and the endoplasmic reticulum (ER). Conversely, the ability of peroxisome to divide and import proteins post-translationally [132] supports the endosymbiotic origin. However, this idea has been challenged by many findings that point to very tight relationships between the ER and the biogenesis of peroxisomes (Section 2.0). Independent evidence for an evolutionary link between peroxisomes and the ER was provided by phylogenetic studies that showed homologous relationships between components of the peroxisomal import machinery and those of the ER-associated decay (ERAD) pathway [133] [134]. Recently, a new hypothetical model for the origin of peroxisomes in which they are considered evolutionary off-shoots of the endomembranous system, rather than a result of bacterial endosymbiosis was proposed [135]. This scenario proposed the avoidance of toxic by-products of lipid metabolism as an initial driving force for the separation of the
organelles, this is based on the facts that fatty acid metabolism is the most widespread function in peroxisomes and related organelles and some of the oxidative enzymes are involved in pathways, such as the synthesis of poly-unsaturated fatty acids that in some species require both ER and peroxisomal steps [136].

The genes encoding the essential import and biogenesis machinery are largely conserved across species indicating that these processes have the same ancient evolutionary origin [133]. Most peroxisomal proteins are of eukaryotic origin including PEX proteins which are involved in peroxisome biogenesis and organisation [133]. Interestingly, five of the six most ancient PEX proteins (PEX1, PEX2, PEX4, PEX5, PEX6, and PEX10) show functional similarity with the ERAD system, which pulls proteins from the ER membrane and ubiquitinylates them in preparation for degradation in the proteasome [137] (see sections 4.3 and 4.4). Also, it is proposed that matrix proteins have been recruited to peroxisomes over evolutionary time and may be of prokaryotic or eukaryotic origin [133]. While about 17-18% have alpha-proteobacterial origin, these proteins seemed to be retargeted from mitochondria rather than evolving endosymbiotically as the beta-oxidation pathway has at least one protein (Fox2p) of alpha-proteobacterial descent, indicating that the capacity to degrade long-chain fatty acids was acquired after mitochondrial endosymbiosis.

In contrast to the relatively large amount of information available for higher plants, little is known about the diversity of functions of peroxisomes in unicellular plants. The presence of peroxisomes has been reported in several green and red algae [138, 139]. The highly adaptable enzymatic content of peroxisome across species indicates a high level of evolutionary plasticity. Some peroxisomal enzymes can only be found in a very narrow range of species. Examples are several key enzymes for the production of penicillin, which are restricted to a few fungal genera such as Penicillium [140] and methylotrophic yeast species (e.g Candida boidinii, Pichia. pastoris) which have enzymes necessary for methanol metabolism [141]. In contrast, other peroxisomal pathways as the β-oxidation of fatty acids and enzymes responsible for oxidative stress response are much more widespread. The distribution of typical peroxisomal enzymes in algae is unusual. Enzymes of the β-oxidation pathway, such as thiolase and acyl-CoA oxidase, may be found in peroxisomes only (Mougeotia), in mitochondria only (Bumilleriopsis, Vaucheria, Pyramimonas), or in both peroxisomes and mitochondria (Eremosphaera, Platymonas,
Heteromastix, Pedinomonas) [142-144]. Mitochondria in this alga also possess some enzymes of the glyoxylate cycle and all the enzymes for fatty acid β-oxidation. Thus, algal peroxisomes may have physiological roles quite different from their higher-plant counterparts.

Analysis of genome sequences shows the presence of PEX genes in plants and other organisms however, this presence is diverse in the different taxonomic groups. For instance, the yeast S. cerevisiae possesses a set of biogenesis proteins of which homologues for 13 have not yet been found in plants or mammals [134], although they are conserved among fungi [145]. In higher plants, PEX genes are functionally conserved [146]. Table 1 shows a survey of peroxins in selected green plants. Arabidopsis thaliana is selected as a representative dicot since most information about gene function is from this species. For comparison wheat (Triticum aestivum) is taken as a monocot example since wheat is one of the most important global food crops. Unlike diploid Arabidopsis, wheat is hexaploid so it is interesting to explore the changes of peroxin complement during the development of this species. Physcomitrella patens (a moss), is selected as a representative of the bryophytes a group of plants that colonised the land about 450mya and for which high quality genome and transcriptome data are available. Finally the model diatom Phaeodactylum tricornutum is taken as a major group of algae that are among the most important of the phytoplankton.

Table 1 lists the complement of PEX genes in Arabidopsis and homologous genes in a moss (Physcomitrella patens) and wheat deduced by bioinformatics. The complement of PEX genes in Physcomitrella and wheat has not previously been described. The completion and annotation of the Physcomitrella genome opens the door to comparative genomics approaches aimed at understanding the evolution and potential functions of genes such as PEX genes, using Arabidopsis as a reference genome (Arabidopsis Genome Initiative, 2000; http://www.arabidopsis.org/). The analysis of Physcomitrella protein sequences revealed the presence of 27 homologs for all Arabidopsis PEX proteins. The Physcomitrella genome contains two homologs for PEX7, PEX14, PEX19 and PEX22 and three homologs for PEX3. The ease with which mutants can be obtained by homologous recombination in Physcomitrella and the relatively simple cellular organisation which is very suitable for microscopy makes it a very good system for exploring peroxin function. In wheat we have identified 48 homologs of PEX genes;
there are three homologs in wheat (genome A, B and D) for most PEX proteins and sometimes one is missing from one genome as in PEX2, PEX4 and PEX10 while there is only one homolog for PEX7 and PEX13. However, the genome of *Phaeodactylum tricornutum* seems to lack all genes encoding proteins specific for the PTS2 import pathway, the most important of which is the PTS2 receptor protein PEX7 as well as obvious homologues of components of the docking complex PEX13 and PEX14 [88].

6.0 Conclusions and future perspectives

One process in particular that has not yet been fully elucidated in any system is that of cargo unloading. The most likely candidates for this are the docking peroxins, PEX13 and PEX14. It has been demonstrated that PTS1-PEX5-PEX7-PTS2 interact by pull-down from the cytosolic fraction of Arabidopsis, but with AtPEX14 only PTS1-PEX5-PEX7 are pulled down, which suggests a role for PEX14 in PTS2 cargo unloading [31]. However any models will have to take into account the apparent absence of the docking peroxins from the diatoms ([88]and Table 1)

It has been suggested in the mammalian import system that the RING complex acts as part of the importomer and that export of the PEX5 protein from the membrane is somehow linked to the cargo unloading process [147, 148]. This is highlighted by the finding that levels of PEX12 are enhanced when PEX13 is truncated [103], and also APEM9 has been found to interact with PEX13 [127]. Three models for the dynamics of the PEX5 protein in cargo delivery and export into the cytosol have been proposed for the mammalian PTS1 import system [147]. These models state that PEX5 could be pulled from the membrane by PEX1/6 in a process twinned with cargo unloading, or the two processes could be separate. Another model is that the two processes are cooperatively coupled: as a monoubiquitinated PEX5 is extracted from the membrane, this allows a cargo-loaded PEX5 protein to release its cargo into the peroxisomal matrix. PEX5:PEX14 have been found in a 1:5 ratio under natural conditions [149], and a 1:1 ratio when export of PEX5 is blocked, which supports the cooperatively coupled model [106].

Establishing the role of the ER in peroxisome biogenesis has proved difficult and controversial as the experiments are technically challenging and often require very careful interpretation. Results with over expressed and tagged proteins can be
potentially misleading, however expression of proteins at native levels may be insufficient to detect, and if detected by immunofluorescence, is critically dependent on the specificity of the antibody. Likewise inferences drawn from mutants (gain or loss of function) must be treated with caution since mutations can have pleiotropic effects, and when not lethal, cells may adapt to the consequences of the mutation. Peroxisomes and ER share an intimate connection at the metabolic level and perturbation of either could impact on the other. In yeast and in fibroblast cells the viability of mutants that lack peroxisomes allows experiments on the early time points of peroxisome recovery which are not possible in plants, and time resolved live imaging experiments support the conclusion that PEX3 traffics first to the ER and is subsequently sorted to peroxisomes [150] [39]. Such direct evidence for a trafficking route from ER to peroxisome is missing for any plant protein. However the ability of CNV to elaborate ER-derived membranes in tobacco into vesicles which contain a peroxisomal matrix reporter [67] and the dependence of viral replication in peroxisomes on an ER protein, Sec39, which implicated in the budding of pre peroxisomes from the ER [69] is suggestive, as is the evolutionary relationship of core peroxins with the ERAD pathway [133] [134]. The evidence for APX3 trafficking from ER to peroxisomes, whilst not wholly conclusive is strong. It is also becoming clear from studies in other organisms that both direct and indirect routes are possible for some peroxisome membrane proteins [20]; PEX16 in mammalian cells being a case in point. Perhaps this is also possible for certain PMPs in plants such as PEX16 and PEX10 where there is contradictory evidence of localisation and trafficking? As well as possible delivery of proteins from ER to peroxisome, an interesting converse situation is the recent demonstration that the oil body triacylglycerol lipase SDP1 is targeted to oil bodies via peroxisomes and transferred by peroxule-like extensions [151].

Peroxisomes and therefore their biogenesis are highly dynamic and regulated by environmental conditions. How environmental signals and internal metabolic states are translated at the molecular level to drive peroxisome biogenesis, membrane protein targeting and peroxisome turnover is beginning to be explored but there is much still to be learned.

Acknowledgements
The authors acknowledge the financial support of the Leverhulme Trust RPG-2012-516 (AB & LLC), BBSRC grants C19029, 24/P13265 and BB/E013740/1 (AB) and a
Fellowship from the Egyptian Cultural Affairs and Mission Sector MOHE (HTE). We thank Dr S. L. Warriner for comments on the manuscript and Dr A Cuming and Mr S. Stevenson for help with the Physcomitrella bioinformatics.
References

[48] I.A. Sparkes, C. Hawes, A. Baker, AtPEX2 and AtPEX10 are targeted to peroxisomes independently of known endoplasmic reticulum trafficking routes, Plant Physiology, 139 (2005) 690-700.


[142] K. Backeshoff, H. Stabenau, Peroxisomes in the alga Vaucheria are neither the leaf peroxisomal or the glyoxysomal type, Botanica Acta, 103 (1990) 190-196.
Figure legends

**Figure 1** Upper panel Schematic diagram of peroxisomal membrane protein trafficking.

As information from studies from plant systems are very limited, the diagram incorporates information from studies on yeast and mammalian systems which may not be identical. i shows the direct import pathway for both multispan and tail anchored PMPs. These interact with PEX19 via their mPTS in the cytosol; subsequently PEX19 docks with PEX3 on the peroxisome membrane leading to insertion of the PMP. Arabidopsis PMP22 inserts directly into peroxisomes in vitro [51] and in vivo [52] and Arabidopsis PEX19 binds multiple PMPs [22]. ii. Direct insertion of PEX3 in a PEX19 and PEX16 dependent manner. This pathway has been described in mammalian cells [41] but has not been examined in any plant system. iii. In humans and fungi some tail anchored (TA) proteins are inserted first in the R by the Get complex. PEX16 and PEX3 are delivered to the ER and are delivered to peroxisomes via vesicular traffic. This route has been characterised in mammalian cells [20]. PEX19 is required for PEX3 exit from the ER in *S.cerevisiae* (which lacks PEX16). At PEX16 has been reported in peroxisomes and ER and modification with a signal anchor sequence resulted in longer retention of AtPEX16 and retention of PEX3 and PMP34 in the ER consistent with a role as a PMP receptor [46]. It is not excluded that PEX16 could also insert directly into peroxisomes (dotted arrow and ?). As PEX16 is not a universal peroxin the pathways shown in ii and iii are not representative of all organisms.

**Middle panel** Metabolic interrelationships between peroxisomes, mitochondria and chloroplasts. Peroxisomes and mitochondria are both oxidative organelles but only in mitochondria is reduction of oxygen coupled to ATP synthesis. Peroxisomes in plants are the sole site of fatty acid oxidation but export succinate or citrate to mitochondria for further metabolism. Leaf peroxisomes house the photorespiratory C2 cycle which involves metabolite exchange with both chloroplasts and mitochondria. The function of peroxisomes is determined by the complement of resident proteins which are established by import and modified by turnover.

**Lower panel** Scheme for the import of proteins into the matrix of the plant peroxisome. Cargo proteins, which contain either a PTS1 (C-terminal) or a PTS2 (N-
terminal) sequence, are bound by their cognate receptors PEX5 or PEX7, respectively. This then targets cargo proteins to the peroxisomal membrane for import and cargo unloading to occur. PTS2 is cleaved from cargo in the peroxisomal matrix. When cargo proteins have been released into the peroxisomal matrix, PEX5 and possibly PEX7 are recycled back into the cytosol to begin another round of import. Accumulation of PEX5 at the peroxisomal membrane can result in lysine polyubiquitination, which results in the degradation of PEX5 via the proteasome in a pathway which could involve the PEX2/PEX12-binding proteins DSK2a and DSK2b. PTS, peroxisomal targeting signal; PEX, peroxin; Ub, ubiquitin; APEM, aberrant peroxisome morphology; DSK, ubiquitin domain-containing protein.

Figure 2 Representation of key mutations or alterations found or applied to each of the proteins involved in peroxisomal matrix protein import in plants. Closed triangle, truncation (caused by nonsense point mutation or transfer DNA (T-DNA) insertion); open triangle, T-DNA insertion resulting in alternate transcript (with STOP codon not reached as a result of the insertion); black circle, missense point mutation; bracket underlining protein, this part of the protein was spliced out; black arrow, expression of the protein knocked down (by T-DNA insertion or RNA interference); red arrow, knock out mutant; aa, amino acids; SH3, SRC homology; TM, transmembrane; RING, really interesting new gene. ARAMEMNON [152] was used to predict transmembrane regions for all proteins except APEM9, where TMPred was used [153].

Table 1 PEX gene homologues in Arabidopsis thaliana, Physcomitrella patens, Triticum aestivum and Phaeodactylum tricornutum.
The set of Arabidopsis PEX gene sequences were obtained from the database maintained at The Arabidopsis Information Resource (http://www.arabidopsis.org). Wheat homologues were identified by reciprocal blast searches of the wheat genome assembly v2.2 (http://phytozome.jgi.doe.gov/pz/portal.html#!search?show=BLAST&method=Org_Taestivum_er) using Arabidopsis PEX genes sequences as queries. Physcomitrella sequences were identified using the JGI website (http://phytozome.jgi.doe.gov/pz/portal.html#!search?show=BLAST&method=Org_Ppate ns). A blast cutoff e value of $10^{-23}$ was used and retrieved.
protein sequences were aligned using the Clustal W program. To confirm each predicted protein sequence as a PEX protein family member sharing a common domain with the corresponding Arabidopsis protein sequence (http://prosite.expasy.org/) was used. Phaeodactylum tricornutum was from [88].
<table>
<thead>
<tr>
<th>Function</th>
<th>Name</th>
<th>Arabidopsis thaliana</th>
<th>Synonym</th>
<th>P. patens (Moss)</th>
<th>Triticum aestivum (Wheat)</th>
<th>Phaeodactylum tricornutum (Diatom)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Receptor for PTS1 proteins</td>
<td>PEX5</td>
<td>At5g56290</td>
<td></td>
<td>Phpat.014G016800</td>
<td>Traes_7DS_820643E76</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Traes_7AS_A21A0DB5B</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Traes_7BS_0B1F58B15</td>
<td></td>
</tr>
<tr>
<td>Receptor for PTS2 proteins</td>
<td>PEX7</td>
<td>At1g29260</td>
<td></td>
<td>Phpat.015G008600</td>
<td>Traes_3B_6B3B42539</td>
<td>N</td>
</tr>
<tr>
<td>Receptor docking</td>
<td>PEX13</td>
<td>At3g07560</td>
<td>APEM2/AMC</td>
<td>Phpat.011G074100</td>
<td>Traes_2DL_B959053A5</td>
<td>N</td>
</tr>
<tr>
<td>Receptor docking</td>
<td>PEX14</td>
<td>At5g62810</td>
<td>PED2</td>
<td>Phpat.009G079100</td>
<td>Traes_1BS_40116F0E7</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phpat.015G080800</td>
<td>Traes_1DS_463CB0B78</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Traes_1AS_39A586E3E</td>
<td></td>
</tr>
<tr>
<td>E2 ligase</td>
<td>PEX4</td>
<td>At5g25760</td>
<td></td>
<td>Phpat.026G021700</td>
<td>Traes_5DL_A4EBA3572</td>
<td>47555</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Traes_5AL_0C019E2F5</td>
<td></td>
</tr>
<tr>
<td>Membrane anchor of PEX4</td>
<td>PEX22</td>
<td>At3g21865</td>
<td></td>
<td>Phpat.003G063400</td>
<td>Traes_2AL_F66956C73</td>
<td>N</td>
</tr>
<tr>
<td>RING complex</td>
<td>PEX2</td>
<td>At1g79810</td>
<td>TED3</td>
<td>Phpat.003G071900</td>
<td>Traes_7AL_1D1A5DBBC</td>
<td>49301</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Traes_7BL_516B17B13</td>
<td></td>
</tr>
<tr>
<td>RING complex</td>
<td>PEX10</td>
<td>At2g26350</td>
<td></td>
<td>Phpat.013G066000</td>
<td>Traes_2AS_E219FAA81</td>
<td>47516</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Traes_2DS_2DDDB6CF9</td>
<td></td>
</tr>
<tr>
<td>RING complex</td>
<td>PEX12</td>
<td>At3g04460</td>
<td>APEM4</td>
<td>Phpat.017G021300</td>
<td>Traes_1BL_A7BB3A6BD</td>
<td>49405</td>
</tr>
<tr>
<td>AAA complex</td>
<td>PEX1</td>
<td>At5g08470</td>
<td>Embryo defective 2817</td>
<td>Phpat.008G077200</td>
<td>Traes_7AS_CB2A9275C</td>
<td>Traes_7BS_192A9C1F7</td>
</tr>
<tr>
<td>AAA complex</td>
<td>PEX6</td>
<td>At1g03000</td>
<td></td>
<td>Phpat.017G081000</td>
<td>Traes_2BL_55117179B</td>
<td>Traes_2DL_B3CD23EBC</td>
</tr>
<tr>
<td>Membrane anchor for AAA complex</td>
<td>APEM9</td>
<td>At3g10572</td>
<td>DAYU ABERRANT PEROXISOME MORPHOLOGY 9</td>
<td>Phpat.009G022500</td>
<td>Traes_7AL_819E90CB</td>
<td>Traes_7BL_AF7AE4AB3</td>
</tr>
<tr>
<td>Membrane protein import?</td>
<td>PEX16</td>
<td>At2g45690</td>
<td>SSE1</td>
<td>Phpat.020G026000</td>
<td>Traes_6AS_6EE0B0E1D</td>
<td>Traes_6BS_841A98B87</td>
</tr>
<tr>
<td>Membrane protein import?</td>
<td>PEX3</td>
<td>At3g18160</td>
<td>AtPEX3-2</td>
<td>Phpat.024G043900</td>
<td>Traes_5DS_BB388ED7C</td>
<td>Traes_5AS_6B30155C7</td>
</tr>
<tr>
<td>Membrane protein import</td>
<td>PEX19</td>
<td>At3g03490</td>
<td>AtPEX19-1</td>
<td>Phpat.002G071900</td>
<td>Traes_6AL_B41FCBEB</td>
<td>Traes_6BL_81E60BF7D</td>
</tr>
<tr>
<td>Peroxisome division/proliferation</td>
<td>PEX11a</td>
<td>At1g47750</td>
<td>AtPEX19-2</td>
<td>Phpat.002G050500</td>
<td>Traes_4DL_4DC47F0501</td>
<td>Traes_4AS_C1AAF027B</td>
</tr>
<tr>
<td>Peroxisome division/proliferation</td>
<td>PEX11b</td>
<td>At3g47430</td>
<td></td>
<td>Phpat.026G017000</td>
<td>Traes_2BL_B974A912B</td>
<td>Traes_2AL_FB2B6601D</td>
</tr>
<tr>
<td>Peroxisome division/proliferation</td>
<td>PEX11c</td>
<td>At2g45740</td>
<td></td>
<td>Phpat.019G070900</td>
<td>Traes_4AL_D9FFAA1A</td>
<td>Traes_4BL_DD7569D22</td>
</tr>
<tr>
<td>Peroxisome division/proliferation</td>
<td>PEX11d</td>
<td>At3g61070</td>
<td></td>
<td>Phpat.018G038300</td>
<td>Traes_4DL_0971895E8</td>
<td>Traes_7AS_D51B7852F</td>
</tr>
<tr>
<td>Peroxisome division/proliferation</td>
<td>PEX11e</td>
<td>At1g01820</td>
<td></td>
<td></td>
<td>Traes_5AL_13E38EF75</td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>Domain/Feature Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX5</td>
<td>- PEX14 binding (WxxxF/Y repeats)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- PEX7 binding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- TPR domain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>728 aa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX7</td>
<td>- WD-40 repeat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>317 aa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX14</td>
<td>- PEX5 binding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Predicted TM region</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Coiled-coil region</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>507 aa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX13</td>
<td>- PEX7 binding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Putative SH3 domain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>304 aa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX4</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>157 aa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX22</td>
<td>- Predicted TM region</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>283 aa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX2</td>
<td>- Predicted TM region</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- RING domain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>333 aa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX10</td>
<td>- Predicted TM region</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- RING domain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>381 aa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX12</td>
<td>- Predicted TM region</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Partial RING domain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>393 aa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>APEM9</td>
<td>- Predicted TM region</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>333 aa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX1</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1,030 aa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX6</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>941 aa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Highlights**

- Peroxisomes use 2 types of targeting signal for matrix protein import
- The receptors bind at the peroxisome membrane, deliver their cargo and are recycled
- The import machinery is comprised mainly of peroxisomal membrane proteins (PMPs).
- PMPs can insert post translationally but some may traffic to peroxisomes via the ER
- Proteins involved in peroxisome biogenesis are conserved across plant species