

Prediction of Peroxisomal Matrix Proteins in Plants



Sigrun Reumann and Gopal Chowdhary

Abstract Our knowledge of the proteome of plant peroxisomes is far from being complete, and the functional complexity and plasticity of this cell organelle are amazingly high particularly in plants, as exemplified by the model species *Arabidopsis thaliana*. Plant-specific peroxisome functions that have been uncovered only recently include, for instance, the participation of peroxisomes in phyloquinone and biotin biosynthesis. Experimental proteome studies have been proved very successful in defining the proteome of Arabidopsis peroxisomes but this approach also faces significant challenges and limitations. Complementary to experimental approaches, computational methods have emerged as important powerful tools to define the proteome of soluble matrix proteins of plant peroxisomes. Compared to other cell organelles such as mitochondria, plastids and the ER, the simultaneous operation of two major import pathways for soluble proteins in peroxisomes is rather atypical. Novel machine learning prediction approaches have been developed for peroxisome targeting signals type 1 (PTS1) and revealed high sensitivity and specificity, as validated by *in vivo* subcellular targeting analyses in diverse transient plant expression systems. Accordingly, the algorithms allow the correct prediction of many novel peroxisome-targeted proteins from plant genome sequences and the discovery of additional organelle functions. In contrast, the prediction of PTS2 proteins largely remains restricted to genome searches by conserved patterns contrary to more advanced machine learning methods. Here, we summarize and discuss the capabilities and accuracies of available prediction algorithms for PTS1 and PTS2 carrying proteins.

S. Reumann (✉)

Department of Plant Biochemistry and Infection Biology,
Institute of Plant Science and Microbiology, University of Hamburg,
Ohnhorststr. 18, 22609 Hamburg, Germany
e-mail: sigrun.reumann@uni-hamburg.de

G. Chowdhary

KIIT School of Biotechnology, Campus XI, KIIT University,
Bhubaneswar 751024, India
e-mail: gkchowdhary@kiitbiotech.ac.in

© Springer Nature Singapore Pte Ltd. 2018

L. A. del Río and M. Schrader (eds.), *Proteomics of Peroxisomes*,
Subcellular Biochemistry 89, https://doi.org/10.1007/978-981-13-2233-4_5

Keywords Peroxisome · Proteome · Import · Peroxisome targeting signals type 1 and 2

Abbreviations

CML	Calmodulin-like protein
EYFP	Enhanced yellow fluorescent protein
PEX	Peroxin
PTS1/2	Peroxisomal targeting signal type 1/2
TPR	Tetratricopeptide (34-amino acid) repeat
PWM	Position weight matrix

1 Prediction of PTS1 Proteins

All soluble matrix proteins of peroxisomes are nuclear-encoded and synthesized on free cytosolic ribosomes with specific signals that direct them post-translationally into the peroxisomal matrix. The majority of matrix proteins possess a peroxisome targeting signal type 1 (PTS1), which consists of a C-terminal tripeptide such as SKL> (“>” refers to the C-terminal protein end) and auxiliary residues located immediately upstream (Gould et al. 1987, 1989; Swinkels et al. 1992; Kragler et al. 1998; Lametschwandtner et al. 1998). Transport of PTS1 proteins into the peroxisomal matrix is mediated by a set of peroxins encoded by *PEX* genes that are required for peroxisome biogenesis (Distel et al. 1996; Hu et al. 2012; Theodoulou et al. 2013; Baker and Paudyal 2014). In brief, soluble proteins carrying a surface-exposed PTS1 are recognized by the conserved cytosolic receptor, PEX5 (the number reflects the chronology of identification, Hayashi et al. 2005; Kragler et al. 1998; Distel et al. 1996; Wimmer et al. 1998). Cargo-loaded PEX5 diffuses to the peroxisomal membrane and docks to the importomer, which is the central membrane-embedded protein import complex that also enables cargo translocation into the matrix (Rayapuram and Subramani 2006; Meinecke et al. 2010). Interestingly, a second homolog of the tetratricopeptide repeat (TPR) protein family was identified recently in *Saccharomyces cerevisiae*, named PEX9, and was characterized as a specific receptor for a subset of peroxisomal matrix proteins, such as an oleate-inducible malate synthase isoform (Effelsberg et al. 2016; Yifrach et al. 2016).

1.1 Canonical Versus Non-canonical PTS1s

PTS1 tripeptides can be classified into canonical and non-canonical sequences. Canonical plant PTS1 tripeptides confer strong peroxisome targeting efficiency to reporter proteins and match the consensus sequence [SA][RK][LMI] > at all three tripeptide positions (Mullen et al. 1997; Lametschwandtner et al. 1998;

Kragler et al. 1998; Reumann 2004; Lingner et al. 2011). These PTS1 tripeptides and their position-specific individual tripeptide residues occur frequently in higher plant PTS1 proteins and have been experimentally demonstrated to function as strong tripeptides and residues for peroxisome targeting, respectively. Canonical PTS1 tripeptides generally are sufficient for peroxisome targeting and mediate high-affinity binding to PEX5. Nevertheless, upstream amino acid residues have been shown to affect PEX5 affinity also for canonical PTS1 tripeptides (Mullen et al. 1997; Hayashi et al. 1997; Kragler et al. 1998; Reumann 2004; Lingner et al. 2011; Brocard and Hartig 2006; Neuberger et al. 2003a, b; Fodor et al. 2012; Lametschwandtner et al. 1998; Maynard et al. 2004).

Non-canonical PTS1 tripeptides generally carry one non-canonical amino acid residue at one tripeptide position (e.g., TRL>, SDL>, and SRV>, non-canonical residues underlined). Nearly all experimentally verified plant PTS1 tripeptides identified to date follow the pattern that one low-abundance PTS1 residue (denoted as x, y, or z) is combined with two high-abundance PTS1 tripeptide residues (x[KR] [LMI]>, [SA]y[LMI]>, [SA][KR]z>). Importantly, this PTS1 classification into canonical and non-canonical tripeptides is simplified and reflects the present status of experimental results and predictions. For instance, SNV> was also validated as a functional plant PTS1 tripeptide, carrying Asn (pos. -2) and Val (pos. -3) and, hence, two low abundance residues (Skoulding et al. 2015).

Non-canonical PTS1 tripeptides alone generally represent weak signals and often require auxiliary targeting-enhancing patterns (e.g., basic residues) for functionality. These are located immediately upstream of the tripeptide and are often kingdom-specific (Neuberger et al. 2003b; Lametschwandtner et al. 1998; Kragler et al. 1998; Ma and Reumann 2008). According to present knowledge, 35 functional plant PTS1 tripeptide residues have been reported. The residues are distributed in the following manner: ([SAPCFVGT LKIQ] [RKNMSLHGTFPQCYDA] [LMIVYF]>), leading to twelve (pos. -3), 17 (pos. -2), and six (pos. -1) allowed amino acid residues in plant PTS1 tripeptides. The targeting strength of PTS1 tripeptides could be classified by in vivo subcellular targeting analyses into three categories: strong, moderate and weak. This classification is based on the time required for the PTS1 to target a reporter protein such as enhanced yellow fluorescent protein (EYFP) to peroxisomes. Further details on this topic are summarized below Sect. (1.3) and available in the authors' publication (Skoulding et al. 2015).

1.2 Prediction Algorithms for PTS1 Proteins

Similar to fungi and animal PTS1s, plant PTS1s exhibit a conserved pattern in the primary sequence level that can be utilized to predict novel peroxisomal proteins by computational approaches. The PTS1 pattern with characteristic features includes the PTS1 tripeptide and several amino acids immediately upstream of the tripeptide. Global biochemical properties and N-terminal targeting information of the protein

can sometimes be added to the prediction models. By utilizing a suitable PTS1 prediction approach in combination with genome information for a species of interest, peroxisomal proteomes of PTS1 proteins can now be predicted in a straightforward manner.

Prediction models are validated for their accuracy by calculation of their prediction sensitivity and specificity. The sensitivity is usually determined as the ratio between correctly predicted peroxisomal proteins (true positives) and the number of all known peroxisomal proteins. The specificity can be assessed by dividing the number of true positives by the number of all (true and falsely) predicted peroxisomal proteins (for more details, see Reumann et al. 2016). Prediction models are usually trained on the larger subset of “training” example sequences, while the accuracy (i.e. sensitivity and specificity) is estimated on a so-called test set of the remaining “unseen” example sequences. In general, the prediction accuracy strongly increases with the size and sequence diversity of the set of example sequences.

In the past decades, several approaches for sequence-based prediction of PTS1 proteins were presented. The first approach developed by Nakai and Kanehisa (1992) was based on overall characteristic amino acid content and a conserved motif ([SA][KRH]L as defined in Gould et al. 1989). Due to a limited set of positive example sequences, the prediction accuracy of the later developed webserver PSORT remained low. The prediction algorithms of PSORTII and WoLF PSORT were based on a larger set of training sequences but did not improve the accuracy significantly (Nakai and Kanehisa 1992; Horton et al. 2007). The PTS1predictor (<http://mendel.imp.ac.at/pts1/>) built in 2003 is still leading in the field and is based on characteristic structural and functional features of more than 300 PTS1 sequences from metazoa, fungi and plants (Neuberger et al. 2003a, b). The algorithm takes the twelve C-terminal amino acids into consideration and evaluates both sequence conservation and structural properties. For plants, however, only a general prediction model is available, contrary to taxa-specific algorithms for metazoa and fungi. Further PTS1 prediction approaches comprise the PeroxiP method (Emanuelsson et al. 2003; discontinued) and the PTS1Prowler algorithm (Hawkins et al. 2007), which was later integrated into the PProwler server (Boden and Hawkins 2005). For details, the reader is referred to our previous publication (Reumann et al. 2016).

The first plant-specific prediction approach for PTS1-containing proteins was published by our group (Lingner et al. 2011), followed by presentation of the public web server PredPlantPTS1 (<http://ppp.gobics.de/>, Reumann et al. 2012). For development of the prediction model, a large set of plant PTS1 sequences homologous to known *A. thaliana* PTS1 sequences was manually identified in protein and EST databases and was manually verified. Positive and negative example sequences were analyzed by a discriminative machine learning model without any restrictions on the tripeptide pattern. The 14 C-terminal amino acids were found to contain discriminative properties. We confirmed the high prediction accuracy of the algorithm by in vivo subcellular targeting analyses of PTS1 decapeptides and full-length proteins fused N-terminally to reporter proteins.

Most importantly (because most challenging in terms of PTS1 protein prediction), several novel peroxisomal proteins bearing non-canonical PTS1 tripeptides were newly identified since publication of the algorithm (Lingner et al. 2011; Kataya et al. 2015a, b, 2016; Kataya and Reumann 2010; Chowdhary et al. 2012; for review see Reumann and Bartel 2016). Notably, the use of a large number of positive and negative example sequences allowed the statistically founded deduction of so-called posterior probabilities (or balanced targeting probability) for peroxisomal targeting between 0 and 100%, which are easier to interpret. Moreover, these balanced posterior probabilities of PTS1 peptides were found to correlate well with experimentally measured binding affinities to Arabidopsis PEX5 (Skoulding et al. 2015).

Wang et al. (2017) recently presented another computational model for the prediction of plant PTS1 proteins. A major difference compared to the above-mentioned machine learning methods, is the authors' claim that also the residues located distantly of the PTS1 tripeptide (between pos. -30 and -15) contained discriminative features distinct from non-peroxisomal proteins (Wang et al. 2017). The prediction model called PPero is publicly available (<https://biocomputer.bio.cuhk.edu.hk/PP/>).

1.3 Prediction and Analysis of Peroxisome Targeting Efficiency

For several reasons, it is often desirable to predict the efficiency at which proteins are targeted to peroxisomes, as outlined previously (Reumann et al. 2016). In vivo subcellular targeting analyses are the gold standard for studying protein localization in peroxisomes to date, and several suitable transient expression systems have been established for in vivo subcellular targeting analyses (Reumann et al. 2016). Only very few studies, however, have also addressed targeting efficiency and were shown to be suited to yield semi-quantitative results. Onion epidermal cells, for instance, used for long-term expression studies over several days of cold incubation allowed the observation of weak peroxisome targeting (Lingner et al. 2011). In the same expression system, it was possible to even resolve significant differences in strong peroxisome targeting efficiency for two canonical PTS1 decapeptides terminating with either SRM> or SRI> after very short expression times (Skoulding et al. 2015).

Thermodynamic in vitro analyses of binding constants are a valuable complementary method to obtain quantitative data of cargo-PEX5 interactions. In fluorescence anisotropy-based assays the affinity of synthetic PTS1 peptides to recombinant PEX5 is determined in a competition experiment, in which a constant, fluorescently labelled peptide bound to PEX5 is replaced by diverse PTS1 peptides of interest (Gatto et al. 2000, 2003; Maynard and Berg 2007). We carried out a systematic comparative analysis of in silico predictions, in vivo subcellular localization data and in vitro thermodynamic binding constant analyses for one model

PTS1 decapeptide and its cytosolic receptor PEX5 and several amino acid residue point mutations of the PTS1 (Skoulding et al. 2015). A good correlation was found between the two experimental methods and the prediction scores. While in vivo subcellular localizations studies turned out to be more sensitive, thermodynamic binding assays yielded quantitative results and allowed a finer discrimination between similar PTS1 peptides (Skoulding et al. 2015). The finding that the position weight matrix (PWM) prediction scores and posterior probabilities also predict the efficiency of protein import into plant peroxisomes is valuable because both experimental methods are laborious and time-consuming compared to the application of prediction tools.

2 PTS2 Nonapeptide Definition and Prediction of PTS2 Proteins

The second targeting signal of peroxisomal matrix proteins is the so-called PTS2 (Swinkels et al. 1991; Osumi et al. 1992). The major targeting information of the PTS2 is included in a conserved nonapeptide of the prototype RL_x5HL located in the N-terminal domain. Four residues of the nonapeptide are highly conserved and spaced by five rather variable residues (Kato et al. 1996, 1998). Interestingly, the number of known PTS2 proteins is rather small in most organisms. In plants, however, as exemplified for Arabidopsis, the number of known PTS2 proteins is with approx. 20 matrix proteins relatively high.

The targeting pathway from the cytosol to the peroxisomal matrix uses pathway-specific PEX proteins in the beginning, before thought to merge with the PTS1 pathway at the peroxisomal membrane. The cytosolic receptor for PTS2 proteins is PEX7, a soluble protein with six WD40 domains. Contrary to PEX5, which is sufficient to target PTS1 containing proteins to the peroxisomal membrane, many organisms require one or two additional co-receptors for proper targeting of PTS2 containing proteins. For instance, *S. cerevisiae* needs PEX18 and PEX21 (Purdue et al. 1998), while in other fungi PTS2 protein import depends on PEX20 (an ortholog of PEX21). In plants and mammals co-receptors of the PEX18/20/21 family have not been reported yet, but PTS2 protein import by PEX7 requires the long version of PEX5 with its PEX7 interaction domain, implying that PEX5 takes over the function of the PEX7 co-receptor in these kingdoms (Dodt et al. 2001; Woodward and Bartel 2005; Ramon and Bartel 2010; Khan and Zolman 2010). The Erdmann group (Ruhr-University Bochum, Germany) recently characterized electrophysiologically a distinct PTS2-specific pore, which consisted of the PTS2 co-receptor PEX18 and the PEX14/Pex17-docking complex as major constituents and also allowed import of folded PTS2 proteins (Montilla-Martinez et al. 2015). Contrary to the PTS1 pore, the reconstituted PTS2 channel was constitutively present in an open state. The new results question the previous concept according to which both import pathways were thought to converge at the peroxisomal

membrane (Montilla-Martinez et al. 2015). Contrary to PTS1 proteins, which are not processed in the matrix, the PTS2 domain is cleaved upon import into peroxisomes by a trypsin-like endopeptidase, referred to as DEG15 in plants (TYSND1 in mammals, Helm et al. 2007; Schuhmann et al. 2008). Dimerization of AtDEG15 was shown to be mediated by the Calmodulin-like protein, AtCML3 (Dolze et al. 2013).

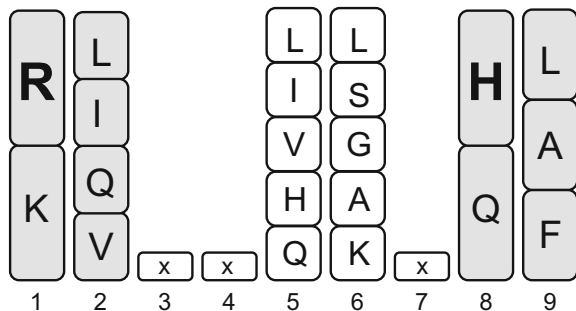
Initial PTS2 analyses by sequence conservation and site-directed mutagenesis revealed that the first two and the last two positions are most conserved in PTS2 nonapeptides in all organismal groups. According to present knowledge, pos. 1 and 8 of the PTS2 nonapeptide are nearly constant with Arg and His, respectively. Both residues are only rarely replaced each by single alternatives, namely Arg by Lys (pos. 1), thus showing a requirement for a positively charged residue, and His by Gln (pos. 8, Fig. 1). Four and three possible hydrophobic residues can occur at pos. 2 (L, I, Q or V) and pos. 9 (L, A or F), respectively (Fig. 1; Petriv et al. 2004).

Initially, the five middle residues were considered highly variable and flexible, any lacking sequence conservation in different orthologous groups. However, advanced computational analyses revealed a preference also at these positions for certain residues. A preference for hydrophobic residues was found also at pos. 5 (L, V, I, H or Q) and 6 (L, S, G, A or K, Petriv et al. 2004). Moreover, with increasing knowledge of the peroxisomal proteome additional PTS2 proteins were identified and an extended consensus PTS2 motif was deduced that included all known PTS2 nonapeptides ([RK][LVIQ]_{x2}[LVIHQ][LSGAK]_x[HQ][LAF], Fig. 1; Petriv et al. 2004).

Kunze et al. (2011) added structural characteristics of the PTS2 receptor and the PTS2 proteins to prediction algorithms by performing mutational studies of PTS2. Using the PTS2 of human thiolase as model nonapeptide, the authors revealed that bulky aliphatic amino acids are essential at pos. 5 for a functional PTS2, while both positively and negatively charged residues at the same position rendered the signal non-functional (Kunze et al. 2011). At pos. 4 the amino acid preference and mutational effect was similar for negatively charged residues.

The similarity between peroxisomal PTS2 and mitochondrial presequences had been early noticed, and single amino acid mutations in the PTS2 domain, such as H-to-R/L (pos. 8), could redirect reporter genes to mitochondria (Osumi et al.

Fig. 1 Graphical presentation of the general PTS2 motif (Petriv et al. 2004). The four most conserved residues of the PTS2 nonapeptide are shaded gray



1992). Even single point mutations in the x_5 sequence, such as the introduction of a basic residue at pos. 4 or 5, directed the reporter protein partially to mitochondria (Kunze et al. 2011).

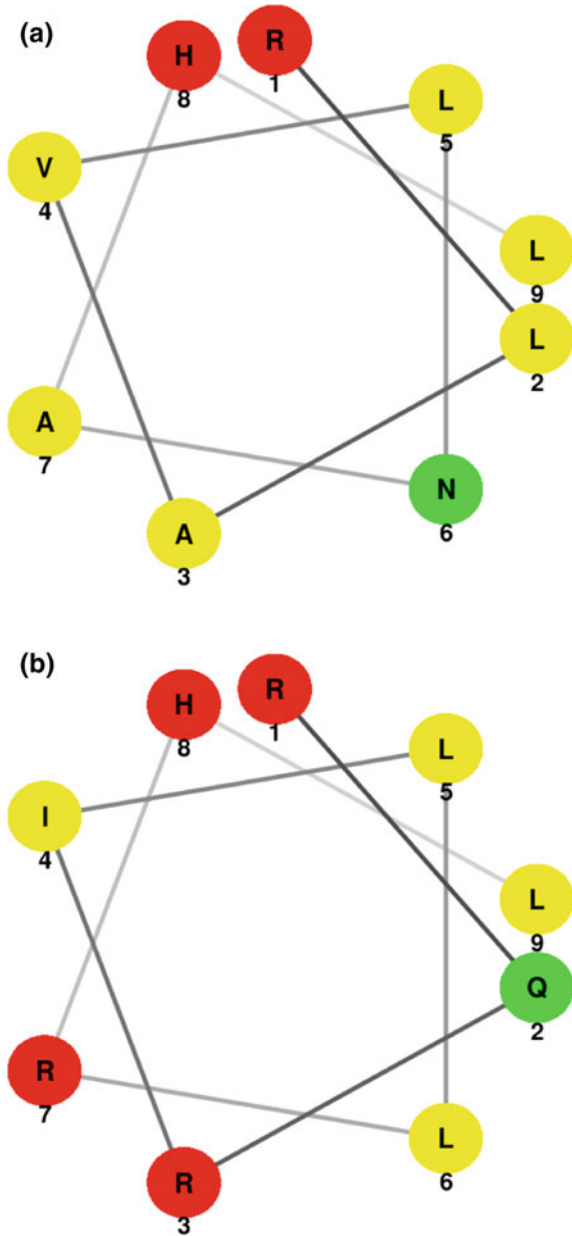
The secondary structure of PTS2 nonapeptides remained long unknown. The hypothesis that the PTS2 nonapeptide forms an α -helix (Reumann 2004; Fig. 2) was strongly supported by the fact that the mutation of a hydrophobic residue to the helix breaking residue, proline, at pos. 6 abolished peroxisome targeting (Kunze et al. 2011; Fig. 2). By generating a homology-based structural model of PEX7, Kunze et al. (2011) could show that human PEX7 formed a groove with an evolutionary conserved charge distribution complementary to the PTS2. The predicted PTS2-PEX7 interaction site was confirmed by mammalian two-hybrid studies. Based on all these PTS2 characteristics, the authors developed a computational screening method and identified a fourth PTS2 protein for mammals, namely potassium channel interacting protein 4 (RV x_5 HL, Kunze et al. 2011).

Conclusive evidence for the PTS2 forming an amphipathic α -helix (Fig. 2), similar to mitochondrial presequences and plastidic transit peptides (Kunze and Berger 2015), was provided by structural analyses. Pan et al. (2013) determined the structure of the ternary complex of *S. cerevisiae* PEX7, the C-terminal domain of PEX21 and the PTS2 from thiolase at 1.8 Å resolution. Accordingly, PEX7 forms a ring structure with a seven-bladed propeller fold formed by the typical WD40 repeats and acts as a platform for binding of both PEX21 and PTS2 cargo. Both receptors form a binding groove in a cooperative manner for the amphipathic α -helix of the PTS2 (Pan et al. 2013).

Prediction methodologies for PTS2 proteins can be classified into simpler motif-based methods and more advanced machine learning methods. Motif-based methods are solely based on the detection of short peptides included in the applied motif, which can be relaxed or specific (stringent, see above). Bodén and Hawkins (2006) combined different motifs in a hierarchical manner from relaxed to stringent. Their PTS2 motif included both “positive” and “negative” properties. The authors claimed that their prediction method had a discriminative accuracy exceeding previously manually curated motifs and could be used to screen genomic data for putative peroxisomal proteins. Applied to the Arabidopsis genome, 76 putative PTS2 proteins were identified (Bodén and Hawkins 2006). Unfortunately, the Arabidopsis proteins were not published, and a public prediction webserver was not created.

Machine learning methods require a large and diverse dataset of positive example sequences to discriminate between PTS2-specific and other protein-specific conserved features. Due to the low number of PTS2 proteins in most organisms (except for plants) and due to the lack of a sufficiently large training data set of positive example sequences, true machine learning methods are not available yet for the prediction of PTS2 proteins. However, in due course of time, (i) as the peroxisomal proteome knowledge will get deeper and richer, (ii) as more peroxisomal PTS2 proteins will become known and (iii) as more genome sequence information will become available, the training data set of positive PTS2 protein example sequences will steadily increase. This development will facilitate the establishment of robust accurate PTS2 protein prediction algorithms for plants,

Fig. 2 Helical wheel presentation for two PTS2 of Arabidopsis proteins. NetWheels (<http://lbqp.unb.br/NetWheels/>) was applied to show the positioning of the nine residues of the PTS2 nonapeptide in the amphipathic α -helix. **a** Arabidopsis citrate synthase (At3g58740, CSY1, RLAVLNAHL) and **b** thiolase (At1g04710, KAT1/PKT4, RQRILLRHL) serve as examples of plant PTS2 proteins. The nonapeptide residues are numbered below the circles (from 1–9). Polar residues are colored red (basic residues), blue (acidic) and green (uncharged), and nonpolar residues are shown in yellow. To mimic a 3-dimensional view from the top into the helix, the lines indicating the peptide bonds are shown as a color gradient from black (beginning of the peptide) to light gray (end of the peptide)



which will most likely also be well applicable to all other eukaryotes that possess the import route of the PTS2 pathway into peroxisomes. Key to successful PTS2 protein prediction will also be the integration of quantitative affinity data between PEX7 and its PTS2 cargo as well as structural data into the prediction models.

3 Conclusions and Future Perspectives

It is well established that the PTS1 is the predominant targeting signal for peroxisome import of matrix proteins. The larger number of positive PTS1 example sequences and the signal's precise position at the C-terminus made it possible to develop successful prediction algorithms. Regarding the PTS2, the restriction to small data sets of positive example sequences, the signal's flexibility in primary structure and its positional flexibility in the N-terminal domain made it difficult to develop accurate prediction algorithms for PTS2 proteins up to now. However, in due course of time, as the peroxisome proteome resources will become richer, the number of known PTS2 proteins will increase. Even more rapidly, the number of fully sequenced eukaryotic genomes increases exponentially, leading to a significant increase of orthologous PTS2-containing sequences per newly identified PTS2 protein. These facts altogether will increase and improve the quantity and quality (e.g. diversity) of the dataset of positive PTS2 protein example sequences, which will facilitate the development of robust PTS2 protein prediction algorithms in the near future.

The peroxisome is the only organelle having two different types of targeting signals for soluble proteins of the organelle matrix, while mitochondria, plastids and the ER evolved only one type of targeting signal, namely a presequence, a transit peptide and a signal peptide, respectively. It is presently unknown why eukaryotes evolved and maintained two different import pathways for peroxisomal matrix proteins and whether one of them is superior, for instance, in terms of import efficiency or specificity, or substrate range and size. A model for the sequential evolution of the two import pathways for peroxisomal matrix proteins has been proposed, starting with the evolution of the PTS2 import pathway and being followed by the PTS1 import pathway for soluble proteins into peroxisomes. (Reumann et al. 2016). It will be interesting to validate this model, for instance by the detection of cargo intermediates of both pathway in ancient organisms.

References

- Baker A, Paudyal R (2014) The life of the peroxisome: from birth to death. *Curr Opin Plant Biol* 22:39–47. <https://doi.org/10.1016/j.pbi.2014.09.003>
- Boden M, Hawkins J (2005) Prediction of subcellular localization using sequence-biased recurrent networks. *Bioinformatics* 21(10):2279–2286
- Bodén M, Hawkins JC (2006) Evolving discriminative motifs for recognizing proteins imported to the peroxisome via the PTS2 pathway. <https://doi.org/10.1109/cec.2006.1688653>
- Brocard C, Hartig A (2006) Peroxisome targeting signal 1: is it really a simple tripeptide? *Biochim Biophys Acta* 1763(12):1565–1573
- Chowdhary G, Kataya AR, Lingner T, Reumann S (2012) Non-canonical peroxisome targeting signals: identification of novel PTS1 tripeptides and characterization of enhancer elements by computational permutation analysis. *BMC Plant Biol* 12(1):142. <https://doi.org/10.1186/1471-2229-12-142> (doi:1471-2229-12-142 [pii])

- Distel B, Erdmann R, Gould SJ, Blobel G, Crane DI, Cregg JM, Dodt G, Fujiki Y, Goodman JM, Just WW, Kiel JA, Kunau WH, Lazarow PB, Mannaerts GP, Moser HW, Osumi T, Rachubinski RA, Roscher A, Subramani S, Tabak HF, Tsukamoto T, Valle D, van der Klei I, van Veldhoven PP, Veenhuis M (1996) A unified nomenclature for peroxisome biogenesis factors. *J Cell Biol* 135(1):1–3
- Dodt G, Warren D, Becker E, Rehling P, Gould SJ (2001) Domain mapping of human PEX5 reveals functional and structural similarities to *Saccharomyces cerevisiae* Pex18p and Pex21p. *J Biol Chem* 276(45):41769–41781. <https://doi.org/10.1074/jbc.M106932200> (M106932200 [pii])
- Dolze E, Chigri F, Howing T, Hierl G, Isono E, Vothknecht UC, Gietl C (2013) Calmodulin-like protein AtCML3 mediates dimerization of peroxisomal processing protease AtDEG15 and contributes to normal peroxisome metabolism. *Plant Mol Biol* 83(6):607–624. <https://doi.org/10.1007/s11103-013-0112-6>
- Effelsberg D, Cruz-Zaragoza LD, Schliebs W, Erdmann R (2016) Pex9p is a new yeast peroxisomal import receptor for PTS1-containing proteins. *J Cell Sci* 129(21):4057–4066. <https://doi.org/10.1242/jcs.195271>
- Emanuelsson O, Elofsson A, von Heijne G, Cristobal S (2003) In silico prediction of the peroxisomal proteome in fungi, plants and animals. *J Mol Biol* 330(2):443–456
- Fodor K, Wolf J, Erdmann R, Schliebs W, Wilmanns M (2012) Molecular requirements for peroxisomal targeting of alanine-glyoxylate aminotransferase as an essential determinant in primary hyperoxaluria type 1. *PLoS Biol* 10(4):e1001309. <https://doi.org/10.1371/journal.pbio.1001309>
- Gatto GJ Jr, Geisbrecht BV, Gould SJ, Berg JM (2000) Peroxisomal targeting signal-1 recognition by the TPR domains of human PEX5. *Nat Struct Biol* 7(12):1091–1095
- Gatto GJ Jr, Maynard EL, Guerrero AL, Geisbrecht BV, Gould SJ, Berg JM (2003) Correlating structure and affinity for PEX5: PTS1 complexes. *Biochemistry* 42(6):1660–1666
- Gould SG, Keller GA, Subramani S (1987) Identification of a peroxisomal targeting signal at the carboxy terminus of firefly luciferase. *J Cell Biol* 105(6 Pt 2):2923–2931
- Gould SJ, Keller GA, Hosken N, Wilkinson J, Subramani S (1989) A conserved tripeptide sorts proteins to peroxisomes. *J Cell Biol* 108(5):1657–1664
- Hawkins J, Mahony D, Maetschke S, Wakabayashi M, Teasdale RD, Boden M (2007) Identifying novel peroxisomal proteins. *Proteins* 69(3):606–616
- Hayashi M, Aoki M, Kondo M, Nishimura M (1997) Changes in targeting efficiencies of proteins to plant microbodies caused by amino acid substitutions in the carboxy-terminal tripeptide. *Plant Cell Physiol* 38(6):759–768
- Hayashi M, Yagi M, Nito K, Kamada T, Nishimura M (2005) Differential contribution of two peroxisomal protein receptors to the maintenance of peroxisomal functions in *Arabidopsis*. *J Biol Chem* 280(15):14829–14835
- Helm M, Luck C, Prestele J, Hierl G, Huesgen PF, Frohlich T, Arnold GJ, Adamska I, Gorg A, Lottspeich F, Gietl C (2007) Dual specificities of the glyoxysomal/peroxisomal processing protease Deg15 in higher plants. *Proc Natl Acad Sci U S A* 104(27):11501–11506
- Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K (2007) WoLF PSORT: protein localization predictor. *Nucleic Acids Res* 35 (Web Server issue): W585–587. <https://doi.org/10.1093/nar/gkm259>
- Hu J, Baker A, Bartel B, Linka N, Mullen RT, Reumann S, Zolman BK (2012) Plant peroxisomes: biogenesis and function. *Plant Cell* 24(6):2279–2303. <https://doi.org/10.1105/tpc.112.096586> (doi:tpc.112.096586 [pii])
- Kataya AR, Heidari B, Lillo C (2015a) Protein phosphatase 2A regulatory subunits affecting plant innate immunity, energy metabolism, and flowering time—joint functions among B β subfamily members. *Plant Signal Behav* 10(5):e1026024. <https://doi.org/10.1080/15592324.2015.1026024>
- Kataya AR, Reumann S (2010) *Arabidopsis* glutathione reductase 1 is dually targeted to peroxisomes and the cytosol. *Plant Signal Behav* 5(2):171–175 (doi:10527 [pii])

- Kataya AR, Schei E, Lillo C (2015b) MAP kinase phosphatase 1 harbors a novel PTS1 and is targeted to peroxisomes following stress treatments. *J Plant Physiol* 179:12–20. <https://doi.org/10.1016/j.jplph.2015.03.002>
- Kataya AR, Schei E, Lillo C (2016) Towards understanding peroxisomal phosphoregulation in *Arabidopsis thaliana*. *Planta* 243(3):699–717. <https://doi.org/10.1007/s00425-015-2439-5>
- Kato A, Hayashi M, Kondo M, Nishimura M (1996) Targeting and processing of a chimeric protein with the N-terminal presequence of the precursor to glyoxysomal citrate synthase. *Plant Cell* 8(9):1601–1611
- Kato A, Takeda-Yoshikawa Y, Hayashi M, Kondo M, Hara-Nishimura I, Nishimura M (1998) Glyoxysomal malate dehydrogenase in pumpkin: cloning of a cDNA and functional analysis of its presequence. *Plant Cell Physiol* 39(2):186–195
- Khan BR, Zolman BK (2010) pex5 Mutants that differentially disrupt PTS1 and PTS2 peroxisomal matrix protein import in *Arabidopsis*. *Plant Physiol* 154(4):1602–1615. <https://doi.org/10.1104/pp.110.162479>
- Kragler F, Lametschwandtner G, Christmann J, Hartig A, Harada JJ (1998) Identification and analysis of the plant peroxisomal targeting signal 1 receptor NtPEX5. *Proc Natl Acad Sci U S A* 95(22):13336–13341
- Kunze M, Berger J (2015) The similarity between N-terminal targeting signals for protein import into different organelles and its evolutionary relevance. *Front Physiol* 6:259. <https://doi.org/10.3389/fphys.2015.00259>
- Kunze M, Neuberger G, Maurer-Stroh S, Ma J, Eck T, Braverman N, Schmid JA, Eisenhaber F, Berger J (2011) Structural requirements for interaction of peroxisomal targeting signal 2 and its receptor PEX7. *J Biol Chem* 286(52):45048–45062. <https://doi.org/10.1074/jbc.M111.301853>
- Lametschwandtner G, Brocard C, Fransen M, Van Veldhoven P, Berger J, Hartig A (1998) The difference in recognition of terminal tripeptides as peroxisomal targeting signal 1 between yeast and human is due to different affinities of their receptor Pex5p to the cognate signal and to residues adjacent to it. *J Biol Chem* 273(50):33635–33643
- Lingner T, Kataya AR, Antonicelli GE, Benichou A, Nilssen K, Chen X-Y, Siemsen T, Morgenstern B, Meinicke P, Reumann S (2011) Identification of novel plant peroxisomal targeting signals by a combination of machine learning methods and in vivo subcellular targeting analyses. *Plant Cell* 23(4):1556–1572. <https://doi.org/10.1105/tpc.111.084095>
- Ma C, Reumann S (2008) Improved prediction of peroxisomal PTS1 proteins from genome sequences based on experimental subcellular targeting analyses as exemplified for protein kinases from *Arabidopsis*. *J Exp Bot* 59(13):3767–3779. <https://doi.org/10.1093/jxb/ern221> (doi:ern221 [pii])
- Maynard EL, Berg JM (2007) Quantitative analysis of peroxisomal targeting signal type-1 binding to wild-type and pathogenic mutants of Pex5p supports an affinity threshold for peroxisomal protein targeting. *J Mol Biol* 368(5):1259–1266
- Maynard EL, Gatto GJ Jr, Berg JM (2004) Pex5p binding affinities for canonical and noncanonical PTS1 peptides. *Proteins* 55(4):856–861. <https://doi.org/10.1002/prot.20112>
- Meinecke M, Cizmowski C, Schliebs W, Kruger V, Beck S, Wagner R, Erdmann R (2010) The peroxisomal importomer constitutes a large and highly dynamic pore. *Nat Cell Biol* 12(3):273–277. <https://doi.org/10.1038/ncb2027> (doi:ncb2027 [pii])
- Montilla-Martinez M, Beck S, Klumper J, Meinecke M, Schliebs W, Wagner R, Erdmann R (2015) Distinct pores for peroxisomal import of PTS1 and PTS2 proteins. *Cell Rep* 13(10):2126–2134. <https://doi.org/10.1016/j.celrep.2015.11.016>
- Mullen RT, Lee MS, Flynn CR, Trelease RN (1997) Diverse amino acid residues function within the type 1 peroxisomal targeting signal. Implications for the role of accessory residues upstream of the type 1 peroxisomal targeting signal. *Plant Physiol* 115(3):881–889
- Nakai K, Kanehisa M (1992) A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* 14(4):897–911
- Neuberger G, Maurer-Stroh S, Eisenhaber B, Hartig A, Eisenhaber F (2003a) Motif refinement of the peroxisomal targeting signal 1 and evaluation of taxon-specific differences. *J Mol Biol* 328(3):567–579

- Neuberger G, Maurer-Stroh S, Eisenhaber B, Hartig A, Eisenhaber F (2003b) Prediction of peroxisomal targeting signal 1 containing proteins from amino acid sequence. *J Mol Biol* 328(3):581–592
- Osumi T, Tsukamoto T, Hata S (1992) Signal peptide for peroxisomal targeting: replacement of an essential histidine residue by certain amino acids converts the amino-terminal presequence of peroxisomal 3-ketoacyl-CoA thiolase to a mitochondrial signal peptide. *Biochem Biophys Res Commun* 186(2):811–818
- Pan D, Nakatsu T, Kato H (2013) Crystal structure of peroxisomal targeting signal-2 bound to its receptor complex Pex7p-Pex21p. *Nat Struct Mol Biol* 20(8):987–993. <https://doi.org/10.1038/nsmb.2618> (nsmb.2618 [pii])
- Petriv OI, Tang L, Titorenko VI, Rachubinski RA (2004) A new definition for the consensus sequence of the peroxisome targeting signal type 2. *J Mol Biol* 341(1):119–134. <https://doi.org/10.1016/j.jmb.2004.05.064>
- Purdue PE, Yang X, Lazarow PB (1998) Pex18p and Pex21p, a novel pair of related peroxins essential for peroxisomal targeting by the PTS2 pathway. *J Cell Biol* 143(7):1859–1869
- Ramon NM, Bartel B (2010) Interdependence of the peroxisome-targeting receptors in *Arabidopsis thaliana*: PEX7 facilitates PEX5 accumulation and import of PTS1 cargo into peroxisomes. *Mol Biol Cell* 21(7):1263–1271. <https://doi.org/10.1091/mbc.E09-08-0672> (doi: E09-08-0672 [pii])
- Rayapuram N, Subramani S (2006) The importomer—a peroxisomal membrane complex involved in protein translocation into the peroxisome matrix. *Biochim Biophys Acta* 1763(12):1613–1619
- Reumann S (2004) Specification of the peroxisome targeting signals type 1 and type 2 of plant peroxisomes by bioinformatics analyses. *Plant Physiol* 135(2):783–800
- Reumann S, Bartel B (2016) Plant peroxisomes: recent discoveries in functional complexity, organelle homeostasis, and morphological dynamics. *Curr Opin Plant Biol* 34:17–26. <https://doi.org/10.1016/j.pbi.2016.07.008>
- Reumann S, Buchwald D, Lingner T (2012) PredPlantPTS1: a web server for the prediction of plant peroxisomal proteins. *Front Plant Sci* 3:194. <https://doi.org/10.3389/fpls.2012.00194>
- Reumann S, Chowdhary G, Lingner T (2016) Characterization, prediction and evolution of plant peroxisomal targeting signals type 1 (PTS1s). *Biochim Biophys Acta* 1863(5):790–803. <https://doi.org/10.1016/j.bbamcr.2016.01.001>
- Schuhmann H, Huesgen PF, Gietl C, Adamska I (2008) The DEG15 serine protease cleaves peroxisomal targeting signal 2-containing proteins in *Arabidopsis*. *Plant Physiol* 148(4):1847–1856. <https://doi.org/10.1104/pp.108.125377>
- Skoulding NS, Chowdhary G, Deus MJ, Baker A, Reumann S, Warriner SL (2015) Experimental validation of plant peroxisomal targeting prediction algorithms by systematic comparison of in vivo import efficiency and in vitro PTS1 binding affinity. *J Mol Biol* 427(5):1085–1101. <https://doi.org/10.1016/j.jmb.2014.12.003>
- Swinkels BW, Gould SJ, Bodnar AG, Rachubinski RA, Subramani S (1991) A novel, cleavable peroxisomal targeting signal at the amino-terminus of the rat 3-ketoacyl-CoA thiolase. *EMBO J* 10(11):3255–3262
- Swinkels BW, Gould SJ, Subramani S (1992) Targeting efficiencies of various permutations of the consensus C-terminal tripeptide peroxisomal targeting signal. *FEBS Lett* 305(2):133–136
- Theodoulou FL, Bernhardt K, Linka N, Baker A (2013) Peroxisome membrane proteins: multiple trafficking routes and multiple functions? *Biochem J* 451(3):345–352. <https://doi.org/10.1042/BJ20130078>
- Wang J, Wang Y, Gao C, Jiang L, Guo D (2017) PPero, a computational model for plant PTS1 type peroxisomal protein prediction. *PLoS ONE* 12(1):e0168912. <https://doi.org/10.1371/journal.pone.0168912>
- Wimmer C, Schmid M, Veenhuis M, Gietl C (1998) The plant PTS1 receptor: similarities and differences to its human and yeast counterparts. *Plant J* 16(4):453–464

- Woodward AW, Bartel B (2005) The Arabidopsis peroxisomal targeting signal type 2 receptor PEX7 is necessary for peroxisome function and dependent on PEX5. *Mol Biol Cell* 16(2): 573–583
- Yifrach E, Chuartzman SG, Dahan N, Maskit S, Zada L, Weill U, Yofe I, Olender T, Schuldiner M, Zalckvar E (2016) Characterization of proteome dynamics during growth in oleate reveals a new peroxisome-targeting receptor. *J Cell Sci* 129(21):4067–4075. <https://doi.org/10.1242/jcs.195255>