Go your own way: membrane targeting sequences

Stefan Wojcik and Verena Kriechbaumer*

Plant Cell Biology, Department of Biological and Medical Sciences, Oxford Brookes University, Oxford OX3 0BP, UK

*Author for correspondence:
Email: vkriechbaumer@brookes.ac.uk
Tel: +44 (0)1865-488403

Short title: Membrane Targeting Sequences

One sentence summary: This review aims to give an overview on membrane targeting sequences taking into account the connected targeting mechanism and co-factors and focusing on primary targeting to membranes.

Abstract

Distributing proteins to the correct sub-cellular compartments and organelles is crucial for the proper functionality of the proteins as well as for the general function of eukaryotic cells. Cellular targeting is best understood in the case of endoplasmic reticulum (ER) proteins targeted co-translationally via the signal recognition particle (SRP)-mediated pathway but various targeting mechanisms, signals and pathways are in place depending on organism, organelle and protein types to allow for specificity and efficiency of protein targeting.
This review aims to give an overview on membrane targeting sequences taking into account the connected targeting mechanism and co-factors. It focusses on primary targeting to membranes of the endoplasmic reticulum, chloroplast, mitochondrion, peroxisome, nucleus and tonoplast.

Advances Box

- Example sequences that allow for targeting of a protein of interest to a membrane of choice for basic research and biotechnological approaches are available.
- Not all GET homologues have been identified in plants, suggesting some divergence in the pathway for insertion of tail anchored proteins.
- Chaperone protein complexes consisting of specific heat shock protein compositions could be important for membrane protein targeting, and receptors capable of binding the highly conserved C-terminus of Hsp70/Hsp90 can be found on each organelle.
- The main role of targeting factors may be to increase the efficiency of targeting by preventing folding and aggregation.
- Recent discoveries for putative pathways targeting proteins to INM, peroxisome membrane and tonoplast offer insight into the variety and complexity of targeting, as well as identify certain signals relevant for certain pathways.

Introduction

Within eukaryotic cells, distinct functions are distributed across intracellular membrane-bounded organelles. Each of these organelles has a discrete pool of proteins, around 30% of which are membrane-bounded (Krogh et al., 2001). Membrane proteins occupy a wide variety of functions, including channels, pores, receptors, or may be involved in various metabolic processes and signal transduction pathways (Hedin et al., 2011). The vast majority of these proteins within the organelles, even endosymbiotic derived organelles which contain their own genome, originate from nuclear genes and are initially translated in the cytosol. 99% of mitochondrial proteins (Rehling et al., 2004) and more than 95% of chloroplast proteins (Soll, 2002) are encoded by nuclear DNA and imported into the organelles. From this cytosolic pool it is imperative that a protein reaches the correct compartment, and a multitude of mechanisms exist to ensure this specific targeting takes place. These mechanisms largely require the recognition of sequences or motifs within a) the peptide as it is being synthesised (co-translational) or b) the mature protein (post-translational). These are the two main targeting distinctions that determine the mechanism and sequences involved, however, targeting between different organelles, as well as potentially multiple membranes within an organelle increases the requirement for distinct signals and mechanisms for each membrane. Additionally, although membrane targeting is an evolutionarily conserved problem, for some organelles little work has been done on plant models. Plant cells also have the extra requirements of chloroplast and tonoplast membranes. This review aims to give an overview

on specific signal sequences and the mechanisms involved in membrane targeting with a focus on plant cells specifically where data is available and comparing with knowledge from other eukaryotic organisms (Figure 1, Table 1).

Table 1: Select examples of typical signal sequences utilised in targeting to specific organelle membranes. Number corresponds with pathway illustrated in Figure 1, see relevant section for more details and references. INM-SM = inner nuclear membrane sorting motif, NLS = nuclear localisation signal, SA = Signal anchor, TA = Tail anchor, TMD = Transmembrane domain.

	Target Membrane	Type of protein	Source	Typical Signal Sequence/Targeting Motif	Example
1A	ER	SA - Type I		Signal Peptide - N(+)-hydrophobic-C(Polar)	
		SA - Other Types	Cytosol	Signal Anchor - ~20 hydrophobic residues	
1B		TA		Tail Anchor - ~20 hydrophobic residues at C-terminus	
1C		Any	Golgi	C-terminal Dilysine (-KKXX-COOH)	GPAT8 & 9
1D	Golgi		ER	Default (Moderate length TMD)	
1E	PM		Golgi	Default (Long TMD)	
2A	Chloroplast Outer	SA		<0.4 WW Hydrophobicity TMD & C-terminal basic residues	Toc64
		TA	Cytosol	TA of variable length and hydrophobicity & C-terminal basic residues	Toc33, Toc34, OEP9
		β-Barrel (TOC-75)		Bipartate signal - Transit peptide & Poly-glutamine stretch	Toc75
2B	Chloroplast Inner		Cytosol	Transit peptide (Tic40= Ser/pro rich - TMD)	Tic40
2C	Thylakoid		Cytosol	L18 motif & -DPLG-	LHCP
			Stroma	N/A	Cytochrome B5
3A	Mitochondrion Outer	SA		Moderate hydrophobicity & C-terminal basic residues	
		TA	Cytosol	<17aa TMD & C-terminal basic residues	
		β-Barrel		N/A	
3B	Mitochondrion Inner		Cytosol	N/A	
4	Peroxisome		Cytosol	Dilysine, -YLSQLQQHPLRT-, 2 basic clusters	PMP22, APX
			ER	>18aa TMD & C-terminal -RKRMK-	Pex16
5	Inner Nuclear Membrane		ER	NLS or INM-SM	
6	Tonoplast		Cytosol	N/A	CBL2
			ER	Dileucine, or tyrosine motif (-YTRL-)	Ptr2, Fruct4

Co-translational targeting in the secretory pathway

The most common mechanism for the targeting of proteins to the membranes of the secretory pathway is the use of a signal recognition particle (SRP) and subsequent SRP receptor localised in the endoplasmic reticulum (ER) membrane (Alberts et al., 2019) (Fig. 1.1A). The SRP pathway is conserved across all domains of life, with prokaryotes utilising a similar mechanism for plasma membrane targeting. In short: as the polypeptide is being translated, the SRP recognises signal sequences that occur either as an N-terminal cleavable peptide (SP), or an internal non-cleavable signal anchor (SA). It has been shown that binding to the nascent peptide requires a combination of RNA and various small polypeptides. GTPase functionality ensures unidirectionality of targeting (Akopian et al., 2013). Initial binding of SRP in eukaryotes has been suggested to occur via three stages. An initial low-affinity preferential binding to translating ribosomes that takes place regardless of the peptide being synthesized (Flanagan et al., 2003), followed by an increased affinity between SRP and the ribosome nascent chain complex (RNC) when a SA sequence is present in the ribosomal exit tunnel. This primes the RNC for the final step whereby the SRP interacts directly with the SA and

forms the strongest affinity with the complex, to allow efficient translocation to the ER (Berndt et al., 2009). Once bound to the RNC, the SRP halts translation and mediates the translocation of the complex to the ER. Here it transitions to preferential binding of the SRP receptor in a GTP-dependent manner to ensure unidirectionality of movement. Interaction between the SRP and its receptor relinquishes SRP's hold on the RNC, which now resumes translation when attached to the translocon SEC61 (Keenan et al., 2001). From here, the protein will be synthesised as an integral membrane protein.

Co-translational signal sequences are not only responsible for the targeting of the protein to the ER but are also integral for the protein assuming correct topology, and thus the type of membrane protein will have differing signal sequences. The structure of the SP has been well characterized and is comprised of three domains with; an N-terminal positively charged region (1-5 amino acids), followed by a hydrophobic core (7-15 amino acids) and a C-terminal polar region (3-7 amino acids) (von Heijne, 1990). SA's comprising the transmembrane domains (TMD) again are very generic across proteins, consisting simply of around 20 hydrophobic residues. Typically, whichever flanking region contains positively charged residues will result in that terminus facing the cytosol (Hartmann et al., 1989). For multi-spanning proteins, this rule applies to the first SA in the protein, and the presence of latter SA and stop-transfer (ST) sequences (another hydrophobic sequence), will generate the correct topology. Other factors such as the folding state (Denzer et al., 1995) and the N-terminal charge:hydrophobic domain length ratio (Sakaguchi et al., 1992) are also implicated in the generation of correct topology. SA membrane proteins are divided into different types (Fig. 2); Type I are single spanning membrane proteins that contain a cleavable SP which orients the N-terminus to be lumenal at the translocon complex and attaches to the bilayer prior to being cleaved by relevant peptidases (von Heijne, 1990). Following this an internal ST sequence prevents further translocation of the protein, and the C-terminus of the polypeptide is subsequently synthesised in the cytosol. Type II and III membrane proteins only contain a single SA sequences, and are therefore single membrane spanning proteins with no cleavable peptide. Type II are orientated via the presence of basic residues N-terminal of the TMD, and the N-terminus of the protein residing in the cytosol. Type III is the opposite, with a luminal facing N-terminus. This is an example of the positive inside rule with inside referring to the cytosol (von Heijne, 1990). Type IVA and IVB are multi-spanning membrane proteins, with the first TMD following the same rules as other transmembrane proteins. IVA proteins contain an N-terminal cytosolic facing portion and IVB a C-terminal one. This rule of orientation is generally well accepted, however, recent studies have highlighted more specific sequence features that determine topology, such as for example the cysteine and tyrosine inside preference (Baker et al., 2017).

Without further mechanisms, proteins bound to ER membranes would be transported to the plasma membrane (Fig. 1.1E) or tonoplast via anterograde vesicular transport. Therefore, as well as having an initial signal sequence to ensure targeting to the ER, it is also imperative the polypeptide contains a retention sequence (Fig. 1.1D). Plant ER membrane proteins utilise the C-terminally positioned dilysine (KKXX) motif (Jackson et al, 1990). These signals are present across Eukarya, and found in ER-resident plant proteins such as glycerol-3-phosphate acyltransferases 8 and 9 (Gidda et al., 2009) and tomato Cf-9 (Benghezal et al., 2000). In mammalian cells, ER membrane proteins such as the p24 family of proteins (Contreras et al., 2004) and others (Zerangue et al., 2001) interact with components of the COPI vesicles such as coatomer subunits (Cosson and Letourner, 1994), as well as the ARF GTPase directly via the dilysine motif; this could explain how this motif is responsible for formation of COPI vesicles and subsequent retrograde transport. Other retention signals include the Arginine RXR motif found in type II ER membrane proteins, the hydrophobic pentapeptide motif, and the atypical dilysine KK-COOH motif (reviewed in Gao et al., 2014). Similarly, the C-terminal Golgi body retention signal KXD/E was first identified in the Arabidopsis EMP family of proteins and has since been found in Golgi-resident proteins across Eukarya (Gao et al., 2012). It is speculated that the composition of COPI coatomer isoforms and distinct subpopulations of COPI vesicles are responsible for the capacity for KKXX and KXD/E motifs to sequester membrane proteins in distinct compartments (Donohoe et al., 2007). It is also suggested that longer TMD's favour localisation to latter parts of the secretory pathway and truncating a TMD can result in re-targeting to earlier compartments (Yang et al., 1997).

Post-translational targeting

All protein targeting to organelles without the use of the SRP-mediated pathway occurs in a post-translational manner after translation is complete. This is the case for all targeting to organelles other than the ER including chloroplasts, mitochondria and peroxisomes and some ER proteins. This targeting is less well understood than the co-translational pathways and for example, the N-terminal signal sequences that direct proteins into the chloroplast and mitochondria are interestingly very similar, and some signal sequences are capable of dual organellar targeting (Chew et al., 2003).

Tail-anchored proteins as a model system to study post-translational protein targeting

In the absence of specific signal peptides, tail-anchored (TA) proteins rely on a single C-terminal TMD to act both as a targeting sequence as well as a membrane anchor. TA proteins are produced in full cytosolically, before translocation to the required membrane where they are anchored via their C-terminus, with the active N-terminal domain facing the cytosol

(Borgese et al., 2007; High and Abell, 2004) (Fig 1.1B). In Arabidopsis there are 454 TA protein genes with predicted 520 splice variants. The majority of these TA proteins are of unknown localisation (75%) and function (66%). 13% localise to the ER, 5% each to plastids and mitochondria (Kriechbaumer et al., 2009). Likewise, in rice and potato, the majority of predicted TA proteins (508 and 912) are of unknown localisation (Manu et al., 2018). 10-20% localise to mitochondria and plastids each, and around 3% localise to the secretory pathway (Manu et al., 2018). Analysis of the targeting of specific TA proteins has identified features rather than specific signal sequences that are responsible for targeting to specific organelles. This includes the length of the TMD and charged residues in the flanking regions (Hwang et al., 2004; Maggio et al., 2007; Wattenberg et al., 2007; Kriechbaumer et al., 2009). In addition PEX19 is capable of recognising specific sequence motifs in the tail-anchor of PEX26 leading to its peroxisomal targeting. As there are not many peroxisomal proteins known, it is unclear whether this is a general mechanism (Halbach et al., 2006). The most extensively studied TA proteins cytochrome b5 family proteins (CYB5). Different CYB5 isoforms localise to the mitochondrion or chloroplast outer membranes or the ER. One distinguishing characteristic that determines their localisation appears to be their hydrophobicity profile which could potentially influence the interaction between the TMD with the differing contents of the lipid bilayers (Pedrazzini, 2009). Depending on organism, protein and target organelle, hydrophobicity scores, length of TMD and frequencies of specific residues such as Leucine within the targeting sequence are all impactful on the localisation of the protein (Manu et al., 2018).

TA membrane insertion pathways

In yeast and mammalian cells, membrane insertion of TA proteins is suggested to be mediated via the Guide Entry of Tail-anchor (GET) pathway (Stefanovic and Hedge, 2007; Favalora et al., 2008). In short: the fully synthesized protein is shielded from cytosolic degradation via the pre-targeting complex consisting of Sgt2, Get4 and Get5. These transfer the TA protein to Get3, which then shuttles the complex to Get1/2. Get3 consists of a homodimer with a nucleotide binding domain, Zn^{2+}, and TA protein binding domains. Nucleotide exchange and hydrolysis of ATP occurs during transfer to the Get1/2 complex. The TA is inserted into the membrane and ADP is released (Denic et al., 2013). Despite strong understanding in yeast the same pathway cannot be extrapolated completely to other branches of eukaryotic life. The components of the plant pre-targeting complex are currently not completely understood but are presumed to consist of currently unidentified Get2/5 and SGt2 homologues. However, multiple Get3 paralogs in Arabidopsis, which appear to be responsible for the differential targeting to the ER (Get3a), plastids (Get3b), and mitochondria (Get3c) (Xing et al., 2017). As the expression of GET pathway proteins causes pleiotropic growth defects, alternative

pathways for TA insertion and additional functions of GET in plants are suggested (Xing et al., 2017). It is suggested that Get3b and Get3c are unlikely to be involved in TA insertion, whereas GET3a is capable of interacting with Get1 and Get4 at the ER to mediate insertion into the bilayer (Zhuang et al., 2017).

Insertion of TA proteins in eukaryotes has also been reported to occur spontaneous (Brambillasca et al., 2006) or via chaperone protein mediated mechanisms (Rabu et al., 2008). In plants, a class of chaperone receptors bearing tetratricopeptide repeat (TPR) domains and capable to specifically bind the highly conserved C-terminus of Hsp70 and/or Hsp90 can be found on each organelle (Schlegel et al., 2007). This might implicate these chaperone receptors in a universal role in protein targeting (reviewed in Kriechbaumer et al., 2012) but the role of these chaperone receptors is not fully understood; chaperone receptors have been shown to increase the efficiency of protein targeting in plants and yeast (Qbadou et al., 2006; Kriechbaumer and Abell, 2012; Young et al., 2003). It is not clear if they also contribute to targeting specificity. It is suggested that chaperones might provide targeting specificity by recognizing features of the bound protein they deliver and form unique signatures through distinct combinations of chaperones which are then recognized by membrane receptors at the destination organelle. As an example, post-translational targeting of the M13 procoat into ER microsomal membranes requires specifically the cytosolic Hsp70, Hsc70 and is not interchangeable with other Hsp70s different chaperone classes (Zimmermann et al., 1988). This chaperone specificity could be due to sequences in M13 but also due to specific ER receptors.

Chloroplast membrane targeting

Perhaps the most complicated topic pertains to that of chloroplast targeting, with three membranes consisting of; the outer chloroplast envelope, inner chloroplast envelope and the thylakoid membrane. Each of these have a distinct protein and lipid composition, with the majority of proteins coming from nuclear encoded genes and thus require specific mechanisms and pathways of insertion depending on the nature of the protein itself. Soluble stromal targeted proteins rely on N-terminal transit peptides (TP) of variable length and sequence, with little specify other than that of a low number of acidic residues (Bruce, 2000); however, more specific subgroups and motifs related to specific aspects of targeting within the TP have been identified (Lee et al., 2008). Import into the stroma is mediated by GTP-dependent activity of two translocon components – translocon of the outer envelope (TOC) and translocon of the inner envelope (TIC). The mechanism in which transit peptides are targeted and imported to the stroma is reviewed in Li and Chiu (2010).

Outer chloroplast membrane proteins

The mechanism for outer chloroplast membrane protein integration will vary depending on the type of protein and its location (Kim et al., 2019) SA proteins do not contain a TP, and instead are post-translationally targeted via the cytosolic Ankyrin repeat-containing protein 2 (AKR2) (Bae et al., 2008) which recognises the signal anchor at the ribosomal exit tunnel and acts as a chaperone during transport (Fig. 1.2A). ARK2 subsequently targets the protein to the outer envelope membrane (OEM) by interaction with specific lipid components of the OEM such as MGDG and PG's. Insertion into the membrane is then aided by Toc75 (Kim et al., 2014; Bae et al., 2008). Interaction with ARK2 occurs via the signal domain, consisting of a TMD of moderate hydrophobicity (<0.4 on W&W scale), as well as a C-terminal positively charged region (CPR) consisting of a minimum of three arginine or lysine residues (Lee et al., 2011). One example that includes these sequence examples is that of Toc64, whose TMD and C-terminal dilysine residues are essential and sufficient for chloroplast targeting. It is speculated that the positive residues prevent SRP binding and trafficking to the ER during translation, and the mature TMD is specific for interaction with AKR2 or other chaperones for correct targeting (Lee et al., 2004).

Chloroplast targeting of TA membrane proteins appears to occur in a similar manner to that of SA proteins, with a requirement of ARK2 and HSP70 acting as chaperones and Toc75 for insertion (Kim et al., 2019). The signal sequence again consists of a TMD followed by a C-terminal sequence. However, some proteins such as Toc33 and Toc34 require an intact N-terminal GTPase domain, whilst others such as OEP9 only require an intact TMD and CPR. Hydrophobicity levels appear to have little effect on this targeting so the specifics of the targeting sequence are relatively plastic between individual proteins (Dhanoa et al., 2010). The exact mechanisms behind these distinct targeting pathways is unclear. Due to the localisation of GET3 isoforms at endosymbiotic organelles, as well as their incapability of rescuing AtGet3a mutant defects in root hair growth, it is suggested that Get3b may be required for TA protein targeting to plastids, and Get3C to mitochondria (Zhuang et al., 2017).

The final category of OEM chloroplast proteins is that of the β-barrels. Seemingly none of them contains TP's and their method of targeting is not well categorised. One exception is that of Toc75 which is unique in that it contains a bipartite signal, consisting of a generic transit peptide followed by polyglutamine stretch (Tranel and Keegstra, 1996). This allows Toc75 to be translocated from the cytosol with the N-terminus reaching the stroma, where the TP can then cleave, retaining Toc75 in the intermembrane space. Following this, the poly-glutamine then stretch is cleaved, an essential step for Toc75 entering the OEM (Inoue, et al., 2005).

How Toc75 is specifically integrated into the membrane following this processing is still unclear, nor is much understood about how the majority of β-barrel proteins are targeted to the OEM.

Inner chloroplast membrane proteins

Similar to that of the outer membrane, multiple targeting pathways have been suggested for inner chloroplast membrane proteins (Lee et al., 2017) (Fig. 1.2B). The majority of inner envelope membrane (IEM) proteins contain TP's, and thus are likely translocated via TOC-TIC machineries. During the translocation at the TIC complex, either the TMD's of the mature protein being translocated can signal the release and insertion from the TIC complex via a stop-transfer mechanism, or the protein is fully translocated to the stroma (Viana, et al., 2010). Membrane proteins such as Tic40 and Tic110 that reach the stroma must undergo post-import insertion, and are suggested to use the SEC2 complex (Li et al., 2017). Tic40 contains a bipartite sorting signal comprised both a TMD and a serine/proline rich domain N-terminal to the TMD, although this is not sufficient to target other proteins to the IEM upon fusion of this signal sequence (Tripp et al., 2007).

Thylakoid membrane proteins

Essential components of the photosynthetic machinery are thylakoid membrane proteins. These can either be encoded by the nuclear genome and imported post-translationally to the stroma prior to insertion into the thylakoid membrane, or are encoded via the chloroplast genome and inserted co-translationally via a chloroplast SRP (csSRP), which is distinct from that of its cytosolic counterpart (Henry, 2010) (Fig. 1.2C). Thylakoid membrane proteins that undergo post-translational insertion proteins are imported to the membrane either via the cpSPR, or spontaneous insertion (Aldridge et al., 2009). Interestingly, csSRP is not only required for co-translational insertion, but is also utilised for post-translational insertion. One example is that of the light harvesting chlorophyll-binding protein (LHCP), which when in the stroma forms a complex with cpSRP43 and cpSRP54 subunits (Falk and Sinning, 2010). Interaction occurs between cpSRP43 and an 18 amino acid motif in LHCP called L18 which contains 4 residues (–DPLG-) essential for this interaction and subsequent targeting (Tu et al., 2000) (Stengel et al., 2008). Following this, the complex binds to the thylakoid membrane-bounded cpFtsY acting as the target (Kogata et al., 1999). LHCP is integrated into the membrane via the integrase activity of albino3 (ALB3), a homologue of the yeast mitochondrial Oxa1p, (Moore et al., 2000). An SRP/ALB3-independent pathway of spontaneous insertion has been shown for other non-LHCP, nuclear encoded thylakoid membrane proteins such as components of the Tat translocase, photosystems I and II, although specific details regarding

the mechanism or signal sequences are not well understood (Woolhead et al., 2001; Shünemann, 2007).

Similarly, the csSRP molecule is also essential for insertion of co-translationally transported proteins derived from the chloroplast genome (Hristou et al., 2019). Although interaction with csSRP and thylakoid membrane proteins has been demonstrated for proteins such as cytochrome b6 (Króliczewski et al., 2016) and the D1 protein of photosystem II (Kim et al., 1991) no specific signals have been identified asides from general hydrophobicity-mediated interaction with csSRP54 and the nascent chain leaving the ribosome (van Wijk and Nilsson, 2002). The general import process operates largely in the same manner as described for LHCP, requiring the integrase activity of ALB3 and other protein components such as cpFtsY. Mutants affecting this pathway usually result in a reduction in chloroplast content and varying severity (Henry et al., 2007).

Targeting of mitochondrial membrane proteins

Being both derived from endosymbiotic organisms, and consisting at least of outer and inner membranes, much of the targeting to mitochondrial membranes occurs via analogous mechanisms to that of the TOC-TIC translocons. In the mitochondria these are translocase of the outer membrane (TOM) and translocase of the inner membrane (TIM) (Neupert, 1997; reviewed in Wiedemann and Pfanner, 2017) (Fig. 1.3). Typically, matrix destined proteins will contain pre-sequences, the mitochondrial equivalent of transit peptides, which are identified and then cleaved by the TOM-TIM translocon following successful translocation. Mitochondrial membrane proteins are mostly derived from nuclear encoded genes (Schleiff and Becker, 2011), and come in various topologies and structures thus relying on multiple signals and pathways for correct targeting. Being analogous to the chloroplast, it is often difficult to discriminate or predict a signal that may result in preferential localisation to either chloroplast or mitochondria, and little work has been done specifically on plant mitochondrial targeting, instead heavily relying on information derived from mammalian and yeast studies. Distinguishing the factors that determine SA protein localisation between chloroplast and mitochondrion outer membranes has proven difficult due to the similarities in their target sequences. Likewise, for chloroplast SA proteins a moderate hydrophobicity TMD and a minimum of three basic residues C-terminal of the TMD are enough to prevent targeting to the ER (Lee et al., 2011). This is consistent in mammalian and yeast mitochondrial outer membrane proteins such as Tom70, although yeast mitochondrial proteins can tolerate basic residue substitution to serine, but not to acidic residues (Waizenegger et al., 2003). Beyond

this, specific SA signal sequences that determine between chloroplast and mitochondria have not been identified.

A similar issue is present for that of TA outer mitochondrial membrane proteins, with the majority of work being done in mammalian and yeast systems. These proteins are similarly targeted as with any other organelle, requiring a C-terminal TMD which acts both as the signal and the anchor for the fully mature protein to be inserted into the membrane. Factors of the signal sequence that determine between ER and mitochondria targeting have been well established. Mitochondrial TA proteins typically contain a shorter hydrophobic segment and require basic residues flanking this TMD, such as the 17aa long TMD with C-terminal twin arginine residues observed in mammalian VAMP1-B that is essential for its unique mitochondrial targeting (Isenmann et al., 1998). Other examples include that of Tom20 which again requires both a TMD of moderate hydrophobicity and flanking basic residues to result in outer membrane targeting (Kanaji et al., 2000). This flanking of basic residues is effective enough for BCL-xl with 2 -RK- C-terminal of the TMD to localise endogenously to the outer mitochondrial membrane, whereas BCL-2 with -HK-, only one basic amino acid, is targeted across other membranes (Kaufmann et al., 2003). This di-basic motif is also consistent with plant mitochondrial proteins, with CYB5–D exclusively localising to the outer mitochondrial membrane not present in ER localised CYB5 isoforms (Hwang et al., 2004).

Little is understood about β-barrel mitochondrial proteins, besides the fact that like chloroplast β-barrel proteins, most of them if not all do not contain cleavable targeting sequences, nor specific signals. Tom40 is imported to the TOM complex via the interaction with Tom22 and Tom70 and insertion is based on the formation of specific intermediates based on several factors surrounding the complex (Rapaport and Neupert, 1999). Interestingly bacterial β-barrel proteins are capable of localising to outer mitochondrial membranes when expressed in eukaryotic cells (Walther et al., 2009). Together, this could suggest that evolutionary conserved aspects of the whole protein structure are more important for targeting rather than specific sequences or cleavable peptides. Initial binding to Tom22 and Tom70 is a consistent first step for any insertion into mitochondrial membranes or further translocation. Little has been described for the import of inner mitochondrial membrane proteins, besides that of ADP/ATP carriers. Following binding with Tom70 and Tom22, Tim9 and Tim10 bind to the hydrophobic segments of the carrier proteins and shuttle the carrier protein to the TIM22 complex where insertion to the inner membrane occurs (Sirrenberg et al., 1996). Aside from general hydrophobic interactions, no specific sequences have been highlighted as essential for this specific targeting, and once again more complex factors involving the full length protein

and components of multiple TMD's, as well as even membrane charge are vital for the insertion of inner membrane proteins (Truscott et al., 2003).

Peroxisomal targeting

The targeting of peroxisomal membrane proteins (PMP) is mediated via two pathways; either imported directly from the cytosol, or via an ER-derived compartment (Mayerhofer, 2016) (Fig. 1.4). Exploring the example of Arabidopsis PMP22, 4 membrane peroxisome targeting signal (mPTS) motifs have been identified for targeting to the peroxisomal membrane (Murphy et al., 2003). 2 motifs acting in parallel are a dilysine motif at residues 7 and 8, as well as a -YLSQLQQHPLRTK- motif between residues 14 and 26. This is consistent with the presence of a Y-x_3-I-x_3-P-x_3-K found in mammalian PMP22 which also acts as a targeting signal (Pause et al., 2000). 2 basic clusters –KIQIRR- (amino acid 49-54) and –KGKK- (amino acid 82-85) were also found to be important for membrane targeting, and similar motifs can be found in peroxisomal ER (pER) derived membrane proteins such as the –RKRMK- motif found in ascorbate peroxidases (APX) which functions as a signal sequence. The sorting signals for peroxisomal membrane-bounded APX are within its C-terminal tail (Mullen and Trelease, 2000). Individually each motif is not sufficient to allow targeting to the peroxisome, and there is a more complex mechanism involving the correct folding, association with chaperones and receptors, as well as the correct topology of the TMD's and motifs to ensure correct targeting (Murphy et al., 2003). The pER subdomain was first identified by the presence of labelled PMP showing signal in domains of the ER, as well as a lack of labelling occurring in peroxisomes when secretion was blocked via brefeldin A treatment (BFA) (Mullen et al., 1999). It was later understood that this targeting to the pER and subsequently the peroxisomes requires the presence of a positively charged motif (-RKRMK-) at the very C-terminus of the protein, as well as a highly hydrophobic TMD at least 18 amino acids long (Mullen and Trelease, 2000). Similarly, other PMP's have been shown to localise to the pER as well as peroxisomes, such as Arabidopsis PEX16; here both TMD's and clusters of basic residues are required for localisation (Karnik and Trelease, 2007).

Nuclear membrane targeting

Targeting of soluble nucleoplasm destined proteins has been well characterised. Nuclear localisation signals (NLS) are recognised by importin (Imp) proteins, and subsequent RanGTP cycles ensure unidirectional transport through the nuclear pore complex (NPC), which connects the inner (INM) and outer (ONM) nuclear membranes (Pouton et al., 2007). Being contiguous with the ER, targeting of membrane proteins to the outer nuclear envelope is thought to only require specific targeting mechanisms that are relevant for ER targeting (Ref).

Conversely, multiple mechanisms have been suggested for the targeting of INM proteins (Fig. 1.5). Non-signal-mediated methods include the diffusion retention model which suggests that membrane proteins can simply diffuse from one membrane to the other via the NPC and are maintained in the INM via interaction partners found within the nucleoplasm (Ungricht et al., 2015). Another putative mechanism is the formation of vesicles that traffic between the ONM and INM, but this is not well characterised and not widely observed (Johnson and Baines, 2011). Other mechanisms include the utilisation of NLS's and/or other INM sorting motifs (INM-SM), for specific metazoan and yeast examples of proteins and signals, as well as more detail on these mechanisms (reviewed in Katta et al., 2014). Recently it has been shown that this NLS mediated pathway is utilised in plants, with the fusion of yeast SV40 monopartate NLS (-PKKKRKV-) (Kalderon et al., 1984) as well as bipartite and plant specific NLS's, fused to a tail-anchored ER protein result in localisation to the INM (Groves et al., 2019).

Tonoplast targeting

Without alternative signals that result in the retention of membrane proteins targeted to the secretory pathway, the default end point is the plasma membrane. However, the most common route for tonoplast membrane proteins is via the secretory pathway via multiple distinct routes including Golgi-dependent and -independent routes (reviewed in Rojas-Pierce, 2013) (Fig. 1.6). Specific motifs and sequences that have been identified to discriminate and target effectively to either protein storage or lytic vacuole and avoid plasma membrane targeting. Early studies of bean α-Tonoplast intrinsic protein (TIP) identified that the sixth TMD and 18 C-terminal residues could target a reporter protein to the tonoplast in tobacco cells (Höfte and Chrispeels, 1992). Other larger cytosolic domains of proteins such as VAMP7 (Uemura, et al., 2005), and AtPK1 (Maîtrejean et al., 2001) are also capable of signalling tonoplast targeting, and fusions of these cytosolic domains can redirect PM localised proteins to the tonoplast (Pedrazzini et al., 2013). Beyond larger domains found in tonoplast residing proteins, specific domains are featured such as dileucine ([D/E]X$_{3-5}$L[L/I]) motifs in AtPTR2,4 and 6 (Komarova et al., 2012), as well as ESL1 (Yamada et al., 2010) and tyrosine based (YTRL) domains found in AtFruct4 (Jung, et al., 2011). These signals are conserved across Eukarya for vacuole targeting in yeast or endosome targeting in mammals and are typically found in close proximity to either TMD's or termini. They are recognised by the adapter group of proteins which results in vesicle formation in post-Golgi trafficking to the tonoplast (Bonifacino and Traub, 2003). As well as these secretory pathway derived routes, targeting directly from the cytosol has been documented for calcineurin B-like (CBL) 2 proteins, on which targeting is mediated by a "tonoplast targeting sequence" consisting of the 19 most N-

terminal residues. This sequence was sufficient to localise a reporter and other CBL proteins to the tonoplast as part of a fusion protein (Tang et al., 2012).

Prospects and outlook

Membrane proteins can occupy a variety of functions that are essential in producing a functional organelle with compartmentalised functions. Therefore, targeting of proteins to specific membranes is highly important and regulated via a multitude of mechanism. Targeting sequences are mostly comprised of hydrophobic segments that are either cleaved as part of targeting or both initiate targeting and subsequently exist as transmembrane domains of the mature, embedded protein. Minor variations in the properties of these hydrophobic segments, properties of the flanking residues, as well as the presence of other motifs that may interact with specific chaperones or translocon components most likely enable targeting to specific membranes.

Exploiting membrane protein targeting signals and mechanism can be of great value to biotechnological approaches. Often the cytoplasm is an inhospitable environment for protein accumulation and due to the presence of vacuoles, cytoplasmic space in plant cells especially is limited. Hence targeting of proteins for biotechnological aspects such as enzyme complexes for production or recombinant protein products to organellar membranes as for example the ER, chloroplasts, tonoplast or oil bodies can be of advantage. Membrane anchoring of enzyme complexes can also be beneficial for complex assembly and stability. TA proteins have the potential to colocalise functional enzyme complexes such as the mitochondrial TOM complex or to specific organelle surfaces. Furthermore half-life and yields of proteins can be increased with TA addition as shown for tobacco expression of the HIV gene product and antiviral factor Nef (Barbante et al., 2008). This is suggested to be due to Nef being less susceptible to cytosolic degradation processes. Targeting of enzyme cascades to the ER may result in proteins directly accumulating in storage bodies formed within the ER lumen (seeds in cereal crops) or be transported to specialised protein storage vacuoles (legume seeds). Such research could have great potential for manipulating or increasing protein and lipid productivity in plants. Plant cells have been successfully glycol-engineered to mimic human glycosylation pathways for production of recombinant proteins. Examples include the generation catalytic domains of mammalian *N*-acetylglucosaminyltransferase (GnT) paired with plant fucosyltransferase localisation signals which target the protein to the Golgi apparatus allowing human-like *N*-glycosylation of recombinant protein (Nagels et al., 2011). Furthermore, mammalian like *N*- and *O*-glycosylation of Human erythropoietin was achieved via the transient expression in *Nicotiana benthamiana* of multiple Golgi localised mammalian and

Drosophila enzymes such as α2,6-sialytransferase, proving the system is viable for potential pharmaceutical production of recombinant proteins (Castilho et al., 2012).

Outstanding questions

- Besides generic targeting sequences, what are the consequences of fully folded proteins and presence of other domains on their targeting?
- How essential are chaperone complexes for post-translational targeting and what are they comprised of?
- How are membrane proteins integrated into the membranes via soluble intermediates or from a translocon complex?
- As targeting of membrane proteins is a fundamental process that occurs in all forms of eukaryotic organisms, most of the pathways are evolutionarily conserved across Eukarya. As shown throughout this review, for most yeast or mammalian proteins, equivalent mechanisms and homologous proteins key to these processes have been identified in plants. However, in the GET pathway for example, there is a lack of identifiable GET2 and GET5 homologues (Xing et al., 2017). This, amongst other areas such as plant specific mitochondrial membrane targeting suggests further work is needed to fully elucidate the mechanisms behind membrane protein targeting in plant cells, rather than filling gaps in our understanding with knowledge from yeast and mammalian studies.
- How do mitochondrial membrane proteins as well as mechanisms such as the GET pathway work specifically in plant cells?
- What is the mechanism of β-barrel protein targeting for outer chloroplast and mitochondrial membranes?
- Are whole organism or *in vitro* approaches with reduced components preferential in protein targeting research?

Acknowledgements

The authors thank Oxford Brookes University for PhD funding for S.W.

Figure legends

Figure 1: Schematics showing a summary of general routes of import for membrane proteins. Numbers correspond with signal sequence details in Table 1. 1A) Co-translational targeting of signal anchored proteins destined for the secretory pathway, mediated via the signal recognition receptor (SRP). 1B) Post-translational targeting of tail-anchored proteins similarly for the secretory pathway, mediated by the GET pathway and other mechanisms. 1C) COPII mediated anterograde transport to the Golgi bodies. 1D) COPI mediated retrograde transport, utilising the dilysine ER retention motif. 1E) Vesicular transport from Golgi to the plasma membrane. 2A) Targeting of outer chloroplast membrane proteins via multiple mechanisms. Translocon of the outer chloroplast membrane (TOC) mediates import of both soluble stromal proteins and is also involved in the insertion of outer chloroplast membrane proteins. 2) Translocon of the inner chloroplast (TIC) membrane is again required for soluble stromal protein trafficking, but also that of inner chloroplast membrane proteins. 2C) Thylakoid membrane protein insertion, post or co-translational both require chloroplast SRP (csSRP). 3A) and 3B), analogous to the TOC-TIC mechanism for chloroplast, the translocon of the outer (TOM) and inner (TIM) mitochondrial membranes mediate insertion of both soluble matrix targeted proteins, and also are involved in the insertion of mitochondrial membrane proteins. 4) Targeting of peroxisomal membrane proteins, either from ER-derived peroxisomal ER (pER) or direct from the cytosol. 5) Inner nuclear envelope targeting requires translocation from ER/outer nuclear envelope to the inner via multiple mechanisms, some requiring nuclear localisation signals (NLS) others not. 6) Targeting of tonoplast membrane proteins, similar to peroxisomes can occur via an ER derived pathway or directly from the cytosol.

Figure 2: Illustration of type I, type II, type III and type IV membrane proteins that are co-translationally inserted into the ER via the SRP.
Type I shows the presence of both an N-terminal cleavable signal peptide (SP) as well as an internal signal anchor domain (SA). The rest of the proteins only contain a SA domain and are oriented via the presence of basic residues either C- or N-terminal of the transmembrane domain (TMD). Stop-transfer (ST) signals are where the protein is reinserted to the membrane to introduce a subsequent TMD following translation of the first SA.

Bibliography

Akopian, D., Shen, K., Zhang, X., and Shan, S. O. (2013). Signal recognition particle: an essential protein-targeting machine. *Annu Rev Biochem,* 82, 693-721.

Alberts, B., Hopkin, K., Johnson, A., Morgan, D., Raff, M., Roberts, K. and Walter, P. (2019) Essential Cell Biology. 5th Edition. New York: W.W. Norton & Company.

Aldridge, C., Cain, P., and Robinson, C. (2009). Protein transport in organelles: Protein transport into and across the thylakoid membrane. *The FEBS Journal,* 276, 1177-86.

Bae, W., Lee, Y., Kim, D., Lee, J., Kim, S., Sohn, E. and Hwang, I. (2008). AKR2A-mediated import of chloroplast outer membrane proteins is essential for chloroplast biogenesis. *Nature Cell Biology*, 10, 220-227.

Baker, J., Wong, W., Eisenhaber, B., Warwicker, J. and Eisenhaber, F. (2017). Charged residues next to transmembrane regions revisited: “Positive-inside rule” is complemented by the “negative inside depletion/outside enrichment rule”. *BMC Biology*, 15, 66.

Benghezal, M., Wasteneys, G. O., Jones, D. A. (2000). The C-terminal dilysine motif confers endoplasmic reticulum localization to type I membrane proteins in plants. *The Plant Cell*, 12, 1179-201.

Berndt, U., Oellerer, S., Zhang Y., Johnson, A. E., and Rospert, S. (2009). A signal-anchor sequence stimulates signal recognition particle binding to ribosomes from inside the exit tunnel. *PNAS,* 106, 1398-1403.

Bonifacino, J. S. and Traub, L. M. (2003). Signals for sorting of transmembrane proteins to endosomes and lysosomes. *Annual Review of Biochemistry*, 72, 395-447.

Borgese, N., Brambillasca, S., and Colombo, S. (2007) How tails guide tail-anchored proteins to their destinations. *Curr Opin Cell Biol,* 19, 368–375.

Brambillasca, S., Yabal, M., Makarow, M., and Borgese, N. (2006). Unassisted translocation of large polypeptide domains across phospholipid bilayers. *J Cell Bio*, 175, 767-777.

Bruce, B. D. (2000). Chloroplast transit peptides: structure, function and evolution. *Trends in Cell Biology*, 10, 440-7.

Castilho, A., Neumann, L., Daskalova, S., Mason, H. S., Steinkellner, H., Altmann, F., and Strasser, R. (2012). Engineering of sialylated mucin-type O-glycosylation in plants. *The Journal of Biological Chemistry*, 287, 36518–26.

Chew, O., Rudhe, C., Glaser, E., Whelan, J. (2003). Characterization of the targeting signal of dual-targeted pea glutathione reductase. *Plant Mol Biol*, 53, 341–356.

Contreras, I., Yang, Y., Robinson, D. G., and Aniento, F. (2004). Sorting signals in the cytosolic tail of plant p24 proteins involved in the interaction with the COPII Coat. *Plant and Cell Physiology,* 45, 1779-1786.

Cosson, P., and Letourner, F. (1994). Coatomer interaction with dilysine endoplasmic reticulum retention motifs. *Science,* 263, 1629-1631.

Denic, V., Dötsch, V., and Sinning, I. (2013). Endoplasmic reticulum targeting and insertion of tail-anchored membrane proteins by the GET pathway. *Cold Spring Harb Perspect Biol,* 5, a013334.

Denzer, A. J., Nabholz, C. E., and Spiess, M. (1995). Transmembrane orientation of signal-anchor proteins is affected by the folding state but not the size of the N-terminal domain. *EMBO J,* 14, 6311-17.

Dhanoa, P., Richardson, L., Smith, M., Gidda, S., Henderson, M., Andrews, D. and Mullen, R. (2010). Distinct pathways mediate the sorting of tail-aAnchored proteins to the plastid outer envelope. *PLoS ONE*, 5, e10098.

Donohoe, B. S., Kang, B. H., and Staehelin, L. A. (2007). Identification and characterization of COPIa- and COPIb-type vesicle classes associated with plant and algal Golgi. *PNAS*, 104, 163-168.

Falk, S., and Sinning, I. (2010). cpSRP43 is a novel chaperone specific for light-harvesting chlorophyll a,b-binding proteins. *J Biol Chem*, 285(28), 21655-61.

Favaloro, V., Spasic, M., Schwappach, B., and Dobberstein, B. (2008). Distinct targeting pathways for the membrane insertion of tail-anchored (TA) proteins. *J Cell Sci*, 121, 1832-40.

Flanagan, J., Chen, J., Miao, Y., Shao, Y., Lin, J., Bock, P. and Johnson, A. (2003). Signal recognition particle binds to ribosome-bound signal sequences with fluorescence-detected subnanomolar affinity that does not diminish as the nascent chain lengthens. *Journal of Biological Chemistry*, 278, 18628-37.

Gao, C., Cai, Y., Wang, Y., Kang, B., Aniento, F., Robinson, D. and Jiang, L. (2014). Retention mechanisms for ER and Golgi membrane proteins. *Trends in Plant Science*, 19, 508-15.

Gao, C., Yu, C., Qu, S., San, M., Li, K., Lo, S. and Jiang, L. (2012). The Golgi-localized Arabidopsis endomembrane protein12 contains both endoplasmic reticulum export and Golgi retention signals at its c terminus. *The Plant Cell*, 24, 2086-104.

Gidda, S., Shockey, J., Rothstein, S., Dyer, J. and Mullen, R. (2009). Arabidopsis thaliana GPAT8 and GPAT9 are localized to the ER and possess distinct ER retrieval signals: Functional divergence of the dilysine ER retrieval motif in plant cells. *Plant Physiology and Biochemistry*, 47, 867-79.

611 Halbach, A., Landgraf, C., Lorenzen, S., Rosenkranz, K., Volkmer-Engert, R., Erdmann, R.,
612 and Rottensteiner, H. (2006). Targeting of the tail-anchored peroxisomal membrane proteins
613 PEX26 and PEX15 occurs through C-terminal PEX19-binding sites. *J Cell Sci*, 119, 2508–17.
614
615 Hartmann, E., Rapoport, T. A., Lodish, H. F. (1989). Predicting the orientation of eukaryotic
616 membrane spanning proteins. *PNAS,* 86, 5786-90.
617
618 Hedin, L. E., Illergård, K., and Elofsson, A. (2011). An introduction to membrane proteins. *J*
619 *Proteome Res,* 10, 3324-31.
620
621 Henry, R. L. (2010). SRP: adapting to life in the chloroplast. *Nature Structural & Molecular*
622 *Biology,* 17, 676-7.
623
624 High, S., Abell, B. M. (2004). Tail-anchored protein biosynthesis at the endoplasmic reticulum:
625 the same but different. *Biochem Soc Trans*, 32, 659–62.
626
627 Höfte, H. and Chrispeels, M. J., (1992). Protein sorting to the vacuolar membrane. *The Plant*
628 *Cell*, 4, 995-1004.
629
630 Hood, E., Love, R., Lane, J., Bray, J., Clough, R., Pappu, K., Drees, C., Hood, K., Yoon, S.,
631 Ahmad, A. and Howard, J. (2007). Subcellular targeting is a key condition for high-level
632 accumulation of cellulase protein in transgenic maize seed. *Plant Biotechnology Journal*, 5,
633 709-19.
634
635 Hristou, A., Gerlach, I., Stolle, D., Neumann, J., Bischoff, A., Dünschede, B., Nowaczyk, M.,
636 Zoschke, R. and Schunemann, D. (2019). Ribosome-associated chloroplast SRP54 enables
637 efficient co-translational membrane insertion of key photosynthetic proteins. *The Plant Cell*,
638 31, 2734-50.
639
640 Hwang, Y., Pelitire, S., Henderson, M., Andrews, D., Dyer, J. and Mullen, R. (2004). Novel
641 targeting signals mediate the sorting of different isoforms of the tail-anchored membrane
642 protein cytochrome b5 to either endoplasmic reticulum or mitochondria. *The Plant Cell*, 16,
643 3002-19.
644
645 Inoue, K., Baldwin, A., Shipman, R., Matsui, K., Theg, S. and Ohme-Takagi, M. (2005).
646 Complete maturation of the plastid protein translocation channel requires a type I signal
647 peptidase. *Journal of Cell Biology*, 171, 425-30.
648
649 Isenmann, S., Khew-Goodall, Y., Gamble, J., Vadas, M. and Wattenberg, B. (1998). A splice-
650 isoform of vesicle-associated membrane protein-1 (VAMP-1) contains a mitochondrial
651 targeting signal. *Molecular Biology of the Cell*, 9, 1649-60.
652
653 Jackson, M., Nilsson, T. and Peterson, P. (1990). Identification of a consensus motif for
654 retention of transmembrane proteins in the endoplasmic reticulum. *The EMBO Journal*, 9,
655 3153-62.
656
657 Johnson, D. and Baines, J. (2011). Herpesviruses remodel host membranes for virus egress.
658 *Nature Reviews Microbiology*, 9, 382-94.

Jung, C., Lee, G., Jang, M., Lee, M., Lee, J., Kang, H., Sohn, E. and Hwang, I. (2011). Identification of sorting motifs of AtβFruct4 for trafficking from the ER to the vacuole through the Golgi and PVC. *Traffic*, 12, 1774-92.

Kanaji, S., Iwahashi, J., Kida, Y., Sakaguchi, M. and Mihara, K. (2000). Characterization of the signal that directs Tom20 to the mitochondrial outer membrane. *Journal of Cell Biology*, 151, 277-88.

Karnik, S. and Trelease, R. (2007). Arabidopsis peroxin 16 trafficks through the ER and an intermediate compartment to pre-existing peroxisomes via overlapping molecular targeting signals. *Journal of Experimental Botany*, 58(7), 1677-93.

Kaufmann, T., Schlipf, S., Sanz, J., Neubert, K., Stein, R. and Borner, C. (2003). Characterization of the signal that directs Bcl-xL, but not Bcl-2, to the mitochondrial outer membrane. *Journal of Cell Biology*, 160, 53-64.

Keenan, R. J., Freymann, D. M., Stroud, R. M., and Walter, P. (2001). The signal recognition particle. *Annual Review of Biochemistry*, 70, 755-75.

Kim, J., Klein, P. G., and Mullet, J. E. (1991). Ribosomes pause at specific sites during synthesis of membrane-bound chloroplast reaction center protein D1. *The Journal of Biological Chemistry,* 266, 14931-8.

Kim, D. H., Park, M. J., Gwon, G. H., Silkov, A., Xu, Z. Y., Yang, E. C., Song, S., Song, K., Kim, Y., Yoon, H. S., Honig, B., Cho, W., Cho, Y., and Hwang, I. (2014). An ankyrin repeat domain of AKR2 drives chloroplast targeting through coincident binding of two chloroplast lipids. *Dev Cell,* 30, 598–609.

Kim, J., Na, Y., Park, S., Baek, S., and Kim, D. (2019). Biogenesis of chloroplast outer envelope membrane proteins. *Plant Cell Reports*, 38, 783-92.

Kogata, N., Nishio, K., Hirohashi, T., Kikuchi, S. and Nakai, M. (1999). Involvement of a chloroplast homologue of the signal recognition particle receptor protein, FtsY, in protein targeting to thylakoids. *FEBS Letters*, 447, 329-33.

Komarova, N., Meier, S., Meier, A., Grotemeyer, M., and Rentsch, D. (2012). Determinants for Arabidopsis peptide transporter targeting to the tonoplast or plasma membrane. *Traffic*, 13, 1090-105.

Kriechbaumer, V., Shaw, R., Mukherjee, J., Bowsher, C. G., Harrison, A. M., and Abell, B. M. (2019). Subcellular distribution of tail-anchored proteins in Arabidopsis. *Traffic*, 10, 1753-64.

Kriechbaumer, V., and Abell, B. M. (2012). Chloroplast envelope protein targeting fidelity is independent of cytosolic components in dual organelle assays. *Front Plant Sci*, 3, 148.

Kriechbaumer, V., von Löffelholz, O. and Abell, B.M. (2012). Chaperone receptors: guiding proteins to intracellular compartments. *Protoplasma,* 249, 21–30.

Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *Journal of Molecular Biology*, 305, 567-80.

Króliczewski, J., Piskozub, M., Bartoszewski, R. and Króliczewska, B. (2016). ALB3 Insertase mediates cytochrome b6 co-translational import into the thylakoid membrane. *Scientific Reports*, 6, 34557.

Lee, D. W., Lee, J. and Hwang, I. (2017). Sorting of nuclear-encoded chloroplast membrane. *Current Opinion in Plant Biology*, 40, 1-7.

Lee, Y. J. Sohn, E. J., Lee, Kwang. Hee., Lee, D. W., and Hwang, I. (2004). The transmembrane domain of AtToc64 and its C-terminal lysine-rich flanking region are targeting signals to the chloroplast outer envelope membrane. *Molecules and Cells*, 17, 281-91.

Li, Y., Martin, J., Aldama, G., Fernandez, D. and Cline, K. (2017). Identification of putative substrates of SEC2, a chloroplast inner envelope translocase. *Plant Physiology*, 173, 2121-37.

Lee, D. W., and Hwang, I. (2018). Evolution and design principles of the diverse chloroplast transit peptides. *Mol Cells*, 41(3), 161-7.

Lee, J., Lee, H., Kim, J., Lee, S., Kim, D. H., and Kim, S. (2011). Both the hydrophobicity and a positively charged region flanking the C-terminal region of the transmembrane domain of signal-anchored proteins play critical roles in determining their targeting specificity to the endoplasmic reticulum or endosymbiotic organelles in Arabidopsis cells. *The Plant Cell*, 23, 1588-607,

Lee, W. D., Kim, J. K., Lee, S., Choi, S., Kim, S., and Hwang, I. (2008). Arabidopsis nuclear-encoded plastid transit peptides contain multiple sequence subgroups with distinctive chloroplast-targeting sequence motifs. *The Plant Cell*, 20, 1603-22.

Li, H. M., and Chiu, C. C. (2010). Protein transport into chloroplasts. *Annual Review of Plant Biology*, 61, 157-80.

Maîtrejean, M., Wudick, M., Voelker, C., Prinsi, B., Mueller-Roeber, B., Czempinski, K., Pedrazzini, E. and Vitale, A. (2011). Assembly and sorting of the tonoplast potassium channel AtTPK1 and its turnover by internalization into the vacuole. *Plant Physiology*, 156, 1783-96.

Maggio, C., Barbante, A., Ferro, F., Frigerio, L., and Pedrazzini, E. (2007). Intracellular sorting of the tail-anchored protein cytochrome b5 in plants: a comparative study using different isoforms from rabbit and Arabidopsis. *J Exp Bot*, 58, 1365–79.

Manu, M. S., Ghosh, D., Chaudhari, B. and Ramasamy, S. (2018). Analysis of tail-anchored protein translocation pathway in plants. *Biochemistry and Biophysics Reports*, 14, 161-167.

Mayerhofer, P. (2016). Targeting and insertion of peroxisomal membrane proteins: ER trafficking versus direct delivery to peroxisomes. *Biochimica et Biophysica Acta*, 1863, 870-80.

Moore, M., Harrison, M., Peterson, E. and Henry, R. (2000). Chloroplast Oxa1p homolog Albino3 is required for post-translational integration of the light harvesting chlorophyll-binding protein into thylakoid membranes. *Journal of Biological Chemistry*, 275, 1529-32.

Mullen, R. T., and Trelease, R. N. (2000). The sorting signals for peroxisomal membrane-bound ascorbate peroxidase are within its C-terminal tail. *J Biol Chem,* 275, 16337-44.

Mullen, R. T., Lisenbee, C., Miernyk, J. and Trelease, R. N. (1999). Peroxisomal membrane ascorbate peroxidase is sorted to a membranous network that resembles a subdomain of the endoplasmic reticulum. *The Plant Cell*, 11, 2167-85.

Murphy, M., Phillipson, B., Baker, A. and Mullen, R. T. (2003). Characterization of the targeting signal of the Arabidopsis 22-kD integral peroxisomal membrane protein. *Plant Physiology*, 133, 813-28.

Nagels, B., Van Damme, E. J. M., Pabst, M., Callewaert, N., and Weterings, K. (2011). Production of complex multiantennary N-Glycans in *Nicotiana benthamiana* plants. *Plant Physiology*, 155, 1103-12.

Neupert, W. (1997). Protein import into mitochondria. *Annual Review of Biochemistry*, 66, 863-917.

Pause, B., Saffrich, R., Hunziker, A., Ansorge, W. and Just, W. (2000). Targeting of the 22 kDa integral peroxisomal membrane protein. *FEBS Letters*, 471, 23-8.

Pedrazzini, E., Komarova, N. Y., Rentsch, D., and Vitale, A. (2013). Traffic routes and signals for the tonoplast. Traffic, 14, 622-8.

Pedrazzini, E. (2009). Tail-anchored proteins in plants. *J Plant Bio,* 52, 88-101.

Qbadou, S., Becker, T., Mirus, O., Tews, I., Soll, J., and Schleiff, E. (2006). The molecular chaperone Hsp90 delivers precursor proteins to the chloroplast import receptor Toc64. *EMBO J*, 25, 1836-47.

Rabu, C., Wipf, P., Brodsky, J. and High, S. (2008). A precursor-specific role for Hsp40/Hsc70 during tail-anchored protein integration at the endoplasmic reticulum. *Journal of Biological Chemistry*, 283, 27504-13.

Rapaport, D. and Neupert, W. (1999). Biogenesis of Tom40, core component of the Tom complex of mitochondria. *J Cell Bio*, 146, 321-32.

Rehling, P., Brandner, K., Pfanner, N. (2004). Mitochondrial import and the twinpore translocase. *Nat Rev Mol Cell Biol,* 5, 519-30.

802 Rojas-Pierce, M. (2013). Targeting of tonoplast proteins to the vacuole. *Plant Science*, 211,
803 132-6.
804
805 Sakaguchi, M., Tomiyoshi, R., Kuroiwa, T., Mihara, K. and Omura, T. (1992). Functions of
806 signal and signal-anchor sequences are determined by the balance between the hydrophobic
807 segment and the N-terminal charge. PNAS, 89, 16-9.
808
809 Schlegel, T., Mirus, O., von Haeseler, A., and Schleiff, E. (2007). The tetratricopeptide repeats
810 of receptors involved in protein translocation across membranes. *Mol Biol Evol*, 24, 2763-74
811
812 Schleiff, E. and Becker, T. (2011). Common ground for protein translocation: access control
813 for mitochondria and chloroplasts. *Nature Reviews Molecular Cell Biology*, 12, 48-59.
814
815 Sirrenberg, C., Bauer, M., Guiard, B., Neupert, W. and Brunner, M. (1996). Import of carrier
816 proteins into the mitochondrial inner membrane mediated by Tim22. *Nature*, 384, 582-585.
817
818 Shünemann, D. (2007). Mechanisms of protein import into thylakoids of chloroplasts.
819 *Biological Chemistry*, 388, 907-15.
820
821 Soll, J. (2002). Protein import into chloroplasts. *Current Opinion in Plant Biology*, 5, 529-35.
822
823 Stefanovic, S. and Hegde, R. (2007). Identification of a targeting factor for posttranslational
824 membrane protein insertion into the ER. *Cell,* 128, 1147-59.
825
826 Stengel, K. F., Holdermann, I., Cain, P., Robinson, C., Wild, K. and Sinning, I. (2008).
827 Structural basis for specific substrate recognition by the chloroplast signal recognition particle
828 protein cpSRP43. *Science*, 321, 253-6.
829
830 Tang, R., Liu, H., Yang, Y., Yang, L., Gao, X., Garcia, V., Luan, S. and Zhang, H. (2012).
831 Tonoplast calcium sensors CBL2 and CBL3 control plant growth and ion homeostasis through
832 regulating V-ATPase activity in Arabidopsis. *Cell Research*, 22, 1650-65.
833
834 Tranel, P. J. and Keegstra, K. (1996). A novel, bipartite transit peptide targets OEP75 to the
835 outer membrane of the chloroplastic envelope. *The Plant Cell*, 8, 2093-104.
836
837 Tripp, J., Inoue, K., Keegstra, K. and Froehlich, J. E. (2007). A novel serine/proline-rich domain
838 in combination with a transmembrane domain is required for the insertion of AtTic40 into the
839 inner envelope membrane of chloroplasts. *The Plant Journal*, 52, 824-38.
840
841 Truscott, K. N., Brandner, K. and Pfanner, N. (2003). Mechanisms of protein import into
842 mitochondria. *Current Biology*, 13, 326-37.
843
844 Tu, C., Peterson, E., Henry, R. and Hoffman, N. (2000). The L18 domain of light-harvesting
845 chlorophyll proteins binds to chloroplast signal recognition particle 43. *Journal of Biological*
846 *Chemistry*, 275, 13187-90.
847
848 Uemura, T., Sato, M. H. and Takeyasu, K. (2005). The longin domain regulates subcellular
849 targeting of VAMP7 in Arabidopsis thaliana. *FEBS Letters*, 579, 2824-46.

Waizenegger, T., Stan, T., Neupert, W. and Rapaport, D. (2003). Signal-anchor domains of proteins of the outer membrane of mitochondria. *The Journal of Biological Chemistry*, 278, 42064-71.

Walther, D., Papic, D., Bos, M., Tommassen, J. and Rapaport, D. (2009). Signals in bacterial β-barrel proteins are functional in eukaryotic cells for targeting to and assembly in mitochondria. *PNAS*, 106, 2531-6.

Wattenberg, B. W., Clark, D., Brock, S. (2007). An artificial mitochondrial tail signal/anchor sequence confirms a requirement for moderate hydrophobicity for targeting. *Biosci Rep*, 27, 385–401.

Woolhead, C., Thompson, S., Moore, M., Tissier, C., Mant, A., Rodger, A., Henry, R. and Robinson, C. (2001). Distinct Albino3-dependent and -independent pathways for thylakoid membrane protein insertion. *Journal of Biological Chemistry*, 276, 40841-6.

van Wijk, K. J., and Nilsson, R. (2002). Transient interaction of cpSRP54 with elongating nascent chains of the chloroplast-encoded D1 protein; 'cpSRP54 caught in the act'. *FEBS Letters*, 524, 127-33.

Viana, A. A., Li, M. and Schnell, D. J. (2010). Determinants for stop-transfer and post-import pathways for protein targeting to the chloroplast inner envelope membrane. *J Biol Chem*, 285, 12948-60.

von Heijne, G. (1990). The Signal Peptide*. J. Membrane Biol,* 115, 195-201.

Wiedemann, N., and Pfanner, N. (2017) Mitochondrial machineries for protein import and assembly. *Annual Review of Biochemistry*, 86, 685-714.

Xing, S., Mehlhorn, D., Wallmeroth, N., Asseck, L., Kar, R., Voss, A., Denninger, P., Schmidt, V., Schwarzländer, M., Stierhof, Y., Grossmann, G. and Grefen, C. (2017). Loss of GET pathway orthologs in Arabidopsis thaliana causes root hair growth defects and affects SNARE abundance. *PNAS*, 114, 1544-53.

Yamada, K., Osakabe, Y., Mizoi, J., Nakashima, K., Fujita, Y., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2009). Functional analysis of an Arabidopsis thaliana abiotic stress-inducible facilitated diffusion transporter for monosaccharides. *Journal of Biological Chemistry*, 285, 1138-46.

Yang, M., Ellenberg, J., Bonifacino, J. S., and Weissman, A. M. (1997). The transmembrane domain of a carboxyl-terminal anchored protein determines localization to the endoplasmic reticulum. *The Journal of Biological Chemistry*, 272, 1970-5.

Young, J. C., Hoogenraad, N. J., and Hartl, F. U. (2003). Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70. *Cell*, 112, 41-50.

Zerangue, N., Malan, M., Fried, S., Dazin, P., Jan, Y., Jan, L. and Schwappach, B. (2001). Analysis of endoplasmic reticulum trafficking signals by combinatorial screening in mammalian cells. *PNAS,* 98, 2431-6

Zhuang, X., Chung, K. P., and Jiang, L. (2017). Targeting tail-anchored proteins into plant organelles. *PNAS,* 114, 1762-4.