On the mechanisms of receptor-mediated retention of soluble endoplasmic reticulum resident proteins in eukaryotes.

Jonas Chaves Alvim

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds

Faculty of Biological Sciences

November, 2018

The candidate confirms that the work submitted is his/her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Part of the work presented in Chapter 1 and Chapter 2 of the thesis has appeared in publication as follows:

Silva-Alvim, F.A.L. et al., 2018. Predominant Golgi-residency of the plant K/HDEL receptor is essential for its function in mediating ER retention. *The Plant Cell*, p.tpc.00426.2018. Available at: http://www.plantcell.org/lookup/doi/10.1105/tpc.18.00426.

I was responsible for conceiving and designing part of the research, to perform part of the research, to analysed the data and write the article in conjunction to Fernanda A. L. Silva-Alvim and Jurgen Denecke. Other authors have contributed to conceive and design part of the research, to perform part of the research and to analysed part of the data.
In the publication I have specifically designed, performed and analysed the data related to anti-sense inhibition of NbERD2ab (Figure 2B) and was responsible for confocal laser scanning microscopy data used in Figures 5, 7, 8 and 9 of the publication.

Part of the work of this thesis was possible by the contribution of other laboratory members by kindly providing unpublished genetic constructs. The authorship of each construct used is provided in Table 2.

Bee lee assisted with the experiments used in Chapter 1, Figure 8, under my personal supervision. Jack Ranger assisted with the experiments used in Chapter 4, Figure 34, under my personal supervision. Nikkita Bhatia assisted with the experiments used in Chapter 4, figure 42 under my personal supervision.

Acknowledgements

" If you find a path with no obstacles, it probably doesn't lead anywhere" - Frank A. Clark

And indeed the path had obstacles. My greatest gratitude first goes to God above all.

In memoriam to my greatest inspiration, my grandfather Prof. Geraldo Martins Chaves. I hope one day I can walk the same path and find the perfect balance in life and dedicate myself to family and science.

Without the support and guidance of family and friends this work would not have been possible. I would like to address the people who supported me as simple gesture of gratitude.

I would like to thank my family. My parents for the emotional support and advices. In special my wife, Fernanda Aparecida Lima Silva Alvim, for unconditional love and scientific support.

I would like to thank my supervisor Prof. Jürgen Denecke for trusting me, for the good scientific discussions and guidance. I would also like to thank Dr. Carine De Marcos, Dr. Jing An and Lewis Adams for the logistical support as co-workers and for the friendly environment.

I would like to thank professor Sreenivasan Ponnambalam from the University of Leeds for all the support facilitating the studies involving mammalian cell lines.

Finally, I would like to thank all the students that spent time with me in the laboratory and contributed in many ways for this work, James Nye, Jack Ranger, Bee Lee and Nikki Bhatia.

The work in this thesis was in part supported by the European Union, the Biotechnology and Biological Sciences Research Council (BBSRC) and The Leverhulme Trust. I am grateful for the support in the form of a full PhD scholarship given to me by the Conselho Nacional de Desenvolvimento Científico e Tecnológico – Brasil (CNPq 201192/2014-4).

Abstract

Accumulation of soluble proteins in the endoplasmic reticulum (ER) of plants is mediated by a protein receptor termed ER RETENTION DEFECTIVE 2 (ERD2). To study the mechanism for protein accumulation in the ER, I have optimized a previously established bioassay using *Nicotiana benthamiana* protoplasts, which proved to be outstanding. The combined use of gain-of-function assays, complementation assays, anti-sense inhibition and confocal laser scanning microscopy allowed me to show that biologically active fluorescent ERD2 fusions are exclusively detected at the Golgi apparatus and do not show ligand-induced redistribution to the ER. I also show that ERD2 dual ER-Golgi distribution is accompanied to lack-of biological function due to the masking a novel C-terminal di-leucine motif. This motif is shown to not promote rapid ER export, but it prevents recycling from the Golgi apparatus back to the ER. Further analysis revealed that the ERD2 C-terminus is necessary but not sufficient to mediate Golgi residency. ERD2 C-terminus replacement by a canonical KKXX sequence, caused biological inactivity and exclusive ER localisation. Together, the data suggest that Golgi-residency and biological activity are directly linked, which argues strongly against the typical receptor-recycling model that has been accepted for so long. Interestingly, I found an astonishingly high degree of conservation of the receptor amongst eukaryotes, and only the receptor from few species were unable to mediate ER retention of soluble ligands in plants. Combined assays provided the experimental platform to classify further ERD2 mutants into three different functional classes. Finally, I generated further data suggesting the ERD2 may have a role in vacuolar transport and protein turnover. I conclude from my work that the classical recycling model for ERD2-mediated accumulation of soluble proteins in the ER may have to be challenged in the future and that much is to be discovered about the ER-Golgi interface in eukaryotes.

Table of Contents

List of Tables

List of Figures

Abbreviations

Amy	Alpha-amylase
ER	Endoplasmic reticulum
ERD2	ER RETENTION DEFECTIVE 2
ARF	ADP ribosylation factor Adenosine
Rab GTPases	Ras-related in Brain guanosine triphosphatases
SNARE	Soluble N-ethylmaleimide-sensitive factor attachment receptor proteins
BiP	Binding immunoglobulin protein
PDI	protein disulphide isomerase
COPI	Coatomer protein I coated
COPII	Coatomer protein II coated
CSLM	Confocal Laser Scanning Microscopy
EDTA	Ethylenediaminetetraacetic acid
ERAD	ER associated degradation
kDa	Kilodaltons
Man-6-P	Mannose-6-phosphate
PCR	Polymerase chain reaction
PVC	Prevacuolar compartmen
LPVC	Late prevacuolar compartment
SDS-PAGE	Sodium dodecyl sulphate polyacrylamine gel electrophoresis
TGN	Trans-Golgi network
VSR	Vacuolar sorting receptor
BFA	Brefeldin-A
LCEA	last common eukaryotic ancestor
SRP	signal recognition particle
hsps	Heat-shock proteins
IRE1	inositol-requiring enzyme
UPR	unfolded protein response
PERK	protein kinase RNA-like endoplasmic reticulum kinase
ATF6	activating transcription factor 6
EE	early endosomes
LE	late endosomes
PM	plasma membrane
RE	recycling endosomes

ERGIC	ER-Golgi intermediate compartment
GlcNAc	N-acetylglucosamine
CPY	carboxypeptidase Y
MPR	Man-6-P receptors
CCV	Clathrin coated vesicles
CaBPs	ER calcium binding proteins
GPCR	G-protein-coupled-receptor
SNX	Sortin nexin proteins
mVPS	mammalian Vacuolar protein sorting-associated

A **General introduction**

A.1 **The Secretory pathway**

The secretory pathway is a network of organelles that ultimately delivers newly synthesized proteins and membranes to their final destination. It links the endoplasmic reticulum (ER) to either lysosomes/vacuoles or the cell surface, and these proteins will first pass through the Golgi apparatus, followed by several classes of endosomes *en route* to the final location. The vectorial nature of the system was first described from the biochemical perspective (Palade 1975), establishing that soluble proteins destined to the secretory pathway are synthesized on the rough endoplasmic reticulum (ER). During the same time-period, the signal hypothesis was inspired by repeated findings suggesting that secreted proteins are synthesized as larger precursors which are cleaved to smaller so-called mature polypeptides after protein translocation to the ER lumen (Blobel & Dobberstein 1975). Subsequent work firmly established that the cleaved portion represents a specific signal peptide that is cleaved during protein translocation, often long before synthesis of the nascent chain is completed (Walter & Lingappa 1986).

Once translocated and released, proteins will fold in the ER lumen and potentially assemble into their native structures (Vitale & Denecke 1999; Helenius et al. 1992). When proteins are not properly folded they can be assisted by chaperones in their maturation process (Hartl 1996) or be targeted to the so called ER associated degradation (ERAD) pathway for degradation (Brodsky & McCracken 1999). As soon as folding is complete, proteins are transported from the ER via vesicles to the Golgi apparatus (Brandizzi & Barlowe 2013a).

The Golgi apparatus is an important protein sorting station which operates by releasing or receiving cargo via vesicle budding and fusion (Hawes 2005; Emr et al. 2009; Munro 2011). Similar to the Golgi apparatus, endosomes can also carry out a protein sorting function to various locations, including the plasma membrane, the lysosomes/vacuoles or the Golgi apparatus (Lemmon & Traub 2000; Carlton et al. 2005; Cullen 2008). Transport between organelles generally occurs via vesicular membrane carriers (Bonifacino & Glick 2004) although

tubular membrane carriers have also been described (Polishchuk et al. 2009). Finally, the entire pathway also appears to function in reverse, starting with endocytosis at the plasma membrane, followed by retrograde transport through various organelles to ultimately even reach the ER and finally the cytosol, as exemplified by toxin entry into eukaryotic cells (Pelham et al. 1992).

It has become clear that the secretory pathway is one of the most ancient innovations of the emerging eukaryotes and probably evolved before the incorporation of mitochondria and chloroplasts. Many of the key-players of the secretory pathway are highly conserved and were probably present in the last common eukaryotic ancestor (LCEA). Further elaboration may have occurred via gene duplications and subsequent evolution of novel specialised organelles (Dacks et al. 2008), leading to the diverse range of post Golgi organelles involved in exocytic and endocytic protein trafficking found in eukaryotes today (Foresti & Denecke 2008; Paez Valencia et al. 2016).

Due to the complexity of the pathway and the numerous ways in which distinct membrane trafficking steps are regulated to permit bi-directional transport, whilst also maintaining intermediate organelle identity, I will introduce the pathway in a sequential manner, starting from the biochemical perspective from the ER.

A.2 **The early secretory pathway**

Extensive infolding of the plasma membrane is often seen in photosynthetic prokaryotes, which increases the membrane surface for ATP synthesis well beyond that of the plasma membrane. One of several models to explain the origin of the secretory pathway is the need to increase the available membrane surface for protein translocation when cells started to become larger. In today's eukaryotes the ER forms a network of tubes and cisternae that also forms the nuclear envelope, but it is no longer connected with the plasma membrane. In addition, a number of intermediate organelles are positioned between the plasma membrane and the ER, the first of which is the Golgi apparatus. Due to the bi-directional transport that is believed to take place between the ER and the Golgi bodies, this part of the secretory pathway can be classified as the early secretory pathway and is exemplified in Figure 1 (Mironov et al. 2007; Hawes et al. 2015; Robinson et al. 2015).

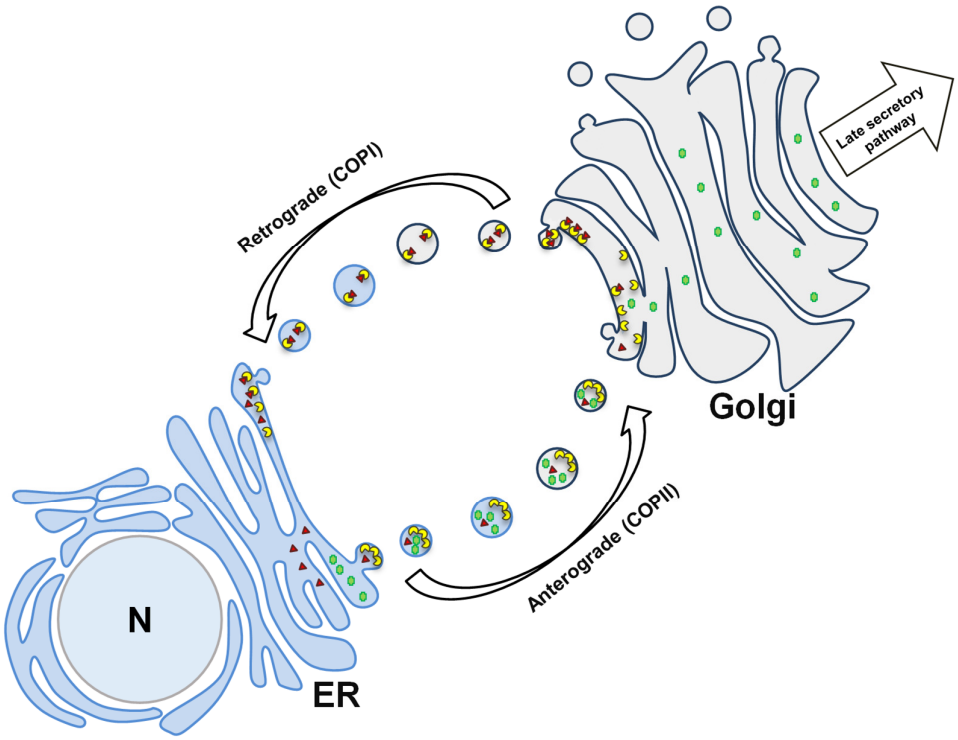


Figure 1 The early secretory pathway and the sorting system of soluble proteins.

The Nucleus (N), Golgi apparatus and Endoplasmic reticulum (ER) are represented. Anterograde from the ER to the Golgi is mediated by COPII vesicles. Retrograde transport from the Golgi to the ER is mediated by COPI vesicles. By bulk-flow ER resident soluble proteins (red triangles) and soluble secreted protein (green spheres) leave the ER using COPII vesicles. Upon arrival at the cis-Golgi ER residents are sorted from proteins destined to the late secretory pathway due to the presence of a tetrapeptide signal motif at their C-terminus, usually a short-sequence (K/HDEL). This signal motif is thought to be recognised by an active receptor protein (yellow half-moon, which binds to the ER resident and via retrograde machinery retrieves them back to the ER. The unbound receptor should then use the anterograde machinery to return to the Golgi apparatus and start a new round of recycling.

A.2.1 The Endoplasmic reticulum morphology, structure and composition

The starting point of the secretory pathway is the organelle knows as Endoplasmic reticulum (ER), which is, amongst many functions, the main station for protein synthesis and folding (Schwarz & Blower 2016). The maintenance of ER structure and sub-domains, as well as the controlling of ER contact sites with different organelles, such as mitochondria and Golgi apparatus, are very complex mechanisms, which are tightly regulated and involve different protein families (for a comprehensive review see Hawes et al., 2015).

The ER structure and sub-domains are a consistent feature amongst eukaryotes and it is generally described as an uninterrupted and interlinked network spreading from the nuclear enveloppe throughout the cell. For animal cells it is sub-divided into the smooth ER and the rough ER based on the absence or presence of ribosomes, respectively (Baumann & Walz 2001). The plant ER does not follow the same split nomenclature but instead the ER is classified into tubular or cisternal, the cisternal has ribosomes on the surface and the tubular does not (Hawes et al. 2015).

Structurally the ER network is consistently formed by either a tubular network, connected by three-way junctions, or by sheet-like structures. These ER structures are drastically different in regards to size, curvature and connections and have different mechanisms controlling their shape (Shibata et al. 2009; Voeltz et al. 2002; English & Voeltz 2013).

As a protein synthesis environment the lumen of the ER is similar to the cytosol in some aspects, particularly with respect to its pH, but which differ from later organelles in the secretory pathway (Shen et al. 2013; Wu et al. 2000). However, compared to the cytosol the calcium concentration is considerably higher in the ER lumen, considered to be one of the main Ca^{2+} storage compartment of eukaryotic cells (Meldolesi & Pozzan 1998). But the most important difference compared to the cytosol is its oxidizing nature (Margittai et al. 2015), which allows the formation of disulphide bridges, which many secretory proteins acquire whilst folding up in this unique environment. The protein folding conditions in the ER are thought to be equivalent to the conditions in the periplasm of gram negative bacteria (Kojima et al. 2013; Herrmann & Riemer 2014).

A.2.2 Protein translocation across the ER membrane

The ER is the gateway to the secretory pathway (Vitale & Denecke 1999), and proteins destined to enter this system of organelles require signals to distinguish them from cytosolic proteins. An ancient ribonucleoprotein complex termed signal recognition particle (SRP) is responsible for the initial targeting of nascent polypeptides to the ER membrane surface (Keenan et al. 2001). In eukaryotes, soluble secreted or vacuolar proteins contain an N-terminal signal peptide (von Heijne 1985) which is recognized by SRP at an early stage of protein synthesis, leading to translational arrest and targeting to the SRP receptor on the ER

membrane (Meyer & Dobberstein 1980; Gilmore 1982). SRP recycles back to the cytosol and the ribosome resumes translation on the translocation pore (Connolly & Gilmore 1989; Johnson & van Waes 1999; Matlack et al. 1998; Pohlschröder et al. 1997). The recycling of the SRP is mediated via an intrinsic GTPase molecular switch (Connolly et al. 1991) and the ribosome seals the translocation pore on the cytosolic side so the protein can only emerge on the lumenal side of the ER membrane. The so-called co-translational translocation offers the advantage to avoid the need for unfolding and refolding as in prokaryotes, although post-translational translocation occurs also in eukaryotes and requires further key-players supporting the pore (for instance, the Sec62/Sec63 complex and BiP in yeast, Rapoport 2007). BiP is thought to act as molecular ratchet to promote post-translational translocation (Rapoport 2007), but it is likely to be essential for co-translational translocation to avoid jamming of the translocation pore. BiP-depletion causes the accumulation of a large number of translocation intermediates (Vogel et al. 1990) suggesting a more general role in this process. Other proteins mediate cleavage of the signal peptide, and also glycosylation, both of which are thought to occur during translocation.

A.2.3 Protein synthesis, folding and final maturation at the ER

The ER lumen is topologically identical to the outside of the cell. Nevertheless, proteins need to fold, assemble and mature into native structures. In addition to soluble proteins, the ER also synthesizes membrane proteins, which may have lumenal as well as cytosolic domains. Following co-translational and post-translational modifications, these proteins will be sorted and distributed to different organelles, but the ER contains a range of quality control mechanisms to prevent misfolded or incompletely folded proteins from leaving the ER (Ellgaard & Helenius 2003; Vitale & Denecke 1999).

There is a diverse group of soluble proteins that can be found in any compartment where protein synthesis is happening, those proteins directly promote the correct folding of nascent proteins and act in the prevention of protein aggregation. Commonly known as the chaperones, and distributed in different families, those soluble proteins act in a quality-control (QC) check which involves protein unfolding, disaggregation and targeting to proteolytic degradation (Ellgaard & Helenius 2003; Kim et al. 2013).

A high number of different proteins have the task to assist in the folding process and ensure that misfolded proteins do not accumulate or advance further in the secretory pathway (Hetz et al. 2011; Ellgaard & Helenius 2003; Kaufman 2004; Rapoport 2007). In short terms, if for any reason a protein is incompletely folded or if it is misfolded it will be recognised and bound to one or more of the different chaperones. Heat shock proteins (Hsp70s, Hsp40s and Hsp90) are the most known and studied family of chaperones, these are found in all cell compartments where protein folding takes place. In addition, the ER also contains a specific set of unique chaperones, including the lectin chaperones (calnexin and calreticulin) and protein disulphide isomerase (PDI). The following sections will introduce the role of different chaperones in mediating the proper conformational changes necessary for a nascent protein to achieve its mature tertiary and sometimes quaternary structure for further transportation across the secretory pathway.

Heat shock proteins (*hsps*)

A vast number of conserved chaperones were initially organized in a multi-family group using as main criteria their molecular weight and gene expression upon a stress condition, in most cases heat due to activation of specific transcription factors and heat shock elements. These were termed heat-shock proteins (*hsps*). However, over time proteins with high homology level, but not necessarily induced by heat-shock, were first termed heat shock cognates (hsc) and later simply included in the *hsps* family. *hsps* have since then been distributed in seven major families, which are: hsp100, hsp90, hsp70, hsp60, hsp40, hsp30, small-hsp. (Kim et al. 2013; De Maio 1999; Lindquist 1986)

A member of the hsp70 family is one of the most studied molecular chaperone and its mechanism of action is extremely well characterized, this protein is called BiP, also known as grp78 or Kar2p. BiP was independently identified as the immunoglobulin heavy chain binding protein (Haas & Wabl 1983) and later grouped into the *hsps*-family as the ER-located HSP70 molecule. The HSP70 group is characterised by two major domains, a catalytic and a binding domain (McKay 1993).

BiP plays a major role in the folding of proteins, binding transiently to newly synthesised ones and more strongly to misfolded substrates (Gething &

Sambrook 1992). The regulation of BiP levels in the lumen of the ER is a complex cascade of events and it is directly involved in different signalling pathways. Although BiP is classified as a *hsps* it is not induced solely by heat stress but instead any condition that promotes misfolding of proteins in the ER, also known as the unfolded protein response (UPR). It was shown before that the reduction of BiP levels in the ER acts as a signal for the UPR and that in yeast the transcription of *KAR2* (BiP) is controlled by three independent *cis*-acting elements related to UPR signal transduction pathway (Gething 1999). In addition, the principle that lumenal BiP levels can directly influence BiP expression has been demonstrated in plants by showing that the overexpression of ectopically expressed BiP downregulates endogenous BiP synthesis (Leborgne-Castel et al. 1999).

Lectin chaperones - Calreticulin and Calnexin

As mentioned before one of the many functions of the ER is to store high concentrations of Ca^{2+}. A class of proteins termed calreticuliuns have been identified by their calcium binding properties due to the presence of an acidic C-terminus. Even though calreticulin plays an important role in within the ER the gene and protein are not as conserved as other ER residents amongst eukaryotes, being absent in yeast (Krause & Michalak 1997). However, yeasts have a molecular chaperone called calnexin with high molecular similarity to calreticulin. Calnexin is a membrane bound protein, with clear calreticulin-like repeated motifs, and many eukaryotes have both calreticulin and calnexin (Bergeron et al. 1994).

Subsequent studies after the initial identification of these proteins have demonstrated that calnexins acts as chaperones and that calreticulin can complement for the lack of calnexins, linking calreticulin to the chaperone role (Krause & Michalak 1997). Unfolded or partially folded monoglucosylated polypeptides can be chaperoned by either the ER soluble calreticulin or membrane bound calnexin by the action of glucosyltransferases, which makes those proteins unique and classified as lectin chaperones (Crofts & Denecke 1998; Del Bem 2011).

Protein disulfide isomerase - PDI

A critical step for stabilization of proteins in their mature and folded form, particularly secreted proteins, is the formation of disulphide bonds linking paired cysteines. The lumen of the ER is equipped with catalysts to enhanced the formation of those bonds as well as to ensure that the folding of the protein is adequate (Fassio & Sitia 2002). A protein named Protein disulfide isomerase (PDI) has been shown to be fundamental for this process acting as both a catalyst and a chaperone, by introducing and rearranging incorrect links (Laboissiere et al. 1995; Wilkinson & Gilbert 2004).

The high conservation level of PDIs is noticeable since it has been identified in a variety of eukaryotes allowing for comparative studies of biological function and complementation as well as phylogenetic analysis (McArthur et al. 2001; Ostermeier et al. 1996). PDI and related proteins have also been identified in plants and comprise a very vast family of proteins, including 22 orthologs in *Arabidopsis thaliana*, forming many phylogenetic groups (Houston et al. 2005).

A.2.4 The unfolded protein response (UPR) and ER stress

The inappropriate folding of proteins is prone to happen due to different factors (Duwi Fanata et al. 2013), and very often will cause various defects, being directly related to many diseases in all eukaryotic organisms (Hartl 2017; Hartl 1996). The unfolded protein response (UPR) is the second step of a cascade of events, occurring after the cell reduces the protein synthesis, triggered by the accumulation of misfolded proteins, commonly termed ER stress (Ron & Walter 2007; Hartl & Hayer-Hartl 2009; Gardner et al. 2013).

UPR was originally studied in yeast and that process is very well characterized, it is believed that key proteins act together to detect and respond to ER stress. Studies with different organisms revealed that UPR is well conserved amongst eukaryotes, but presenting different complexity levels, thus it is very likely that most of the proteins involved are also conserved (for review consult, Patil & Walter 2001).

As mentioned before, in an ER stress situation, sensors monitor the concentration of BiP, but also detects the presence of unfolded proteins, which will induce a response to correct any problems. There are different models trying to explaining how those sensors are activated or deactivated (Ron & Walter 2007). A protein called inositol-requiring enzyme (Ire1p or IRE1) is of extreme

importance since it operates as the main sensor for UPR and is conserved amongst eukaryotes. IRE1 is a transmembrane protein that stays in the ER and it has different domains, including a luminal sensing domain, a cytosolic kinase domain and an RNAse domain (Cox et al. 1993; Chen & Brandizzi 2013; Hetz et al. 2011).

Eventually the ER stress caused by unfolded proteins will either induce the release of secondary proteins to rebalance the unfolding problems, for instance the chaperones previously discussed, or the problematic proteins will be retained in the ER for later degradation in an alternative complex process called ERAD (ER-associated degradation). Lastly, if ER stress is prolonged it could trigger cell-death, which is much less conserved mechanism throughout eukaryotes (Chen & Brandizzi 2013; for reviews consult, Smith et al. 2011; Tabas & Ron 2011).

The UPR process and how IRE1 acts as a sensor has been extensively studied in both yeast and mammals. In mammalian cells IRE1 is only one of three sensors, the other ones being called PERK (protein kinase RNA-like endoplasmic reticulum kinase) and ATF6 (activating transcription factor 6) (Harding et al. 1999; Haze et al. 1999). Acting together with IRE1 they maintain the internal balance of proteins in the ER. In plants there is no homologue of PERK, however there is for ATF6 and they are called bZIP28 and bZIP17 (Iwata & Koizumi 2012).

It is very important to highlight the statement made by Chen & Brandizzi (2013) asserting that although IRE1 is highly conserved its sensing and regulatory mechanisms have been found to be remarkably divergent, the first one being unidentified in plants. Another relevant piece of the puzzle is that in the before mentioned organism both isoforms are ubiquitously expressed differing from the mammalian equivalents and the results from knock-outs are not alike.

A.2.5 Export from the ER

Once proteins have acquired their final conformation and passed the various quality control mechanisms, they can be transported to the Golgi apparatus, the next station in the pathway. Most proteins will exit the ER in transport vesicles from so-called ER exit sites (daSilva et al. 2004). Some proteins contain ER export signals that facilitate active transport (Nishimura et al. 1999) whilst others

simply progress in a more passive manner, also known as *bulk-flow* (Wieland et al. 1987; Denecke et al. 1990). The transport of proteins and lipids from the ER towards the Golgi complex is called anterograde or biosynthetic transport in the secretory pathway, and it is thought to matched by an equally important transport in the opposite direction that is called retrograde (Lee et al. 2004). The retrograde route is thought to make sure that valuable proteins that support basic ER functions in protein translocation, folding and modification are not lost by unspecific *bulk-flow*.

It is clear that the two organelles need to work closely together to sort ER residents from other proteins destined to late organelles and secreted proteins. However, the bi-directional traffic between the ER and the Golgi is not only essential to retrieve ER resident proteins, but also to maintain the integrity and structural identity of the two organelles themselves. Membranes lost from the ER in the form of new vesicles need to be replaced via recycling (Harter & Wieland 1996). Similarly, newly arriving vesicles at the Golgi would cause this organelle to grow, but the retrograde transport machinery prevents this from happening. The recycling principle is based on a very large number of regulatory proteins, protein coats, adaptor molecules and other proteins that facilitate vesicle budding and fusion that will be discussed in much further detail below.

A.3 **The late secretory pathway**

The late secretory pathway contains two major branches, one of which leading to the plasma membrane (PM), the other to the lytic organelles termed lysosomes in mammals and vacuoles in plants, protists and fungi. This branched nature, together with the fact that the endocytic route from the PM also leads to a range of directions, imposes a much higher complexity level compared to the early stages of protein sorting between the ER and the Golgi alone. Due to the fact that the bi-directional transport occurs between most of the organelles in this pathway, it is no simple matter to study individual transport steps because the various routes are strongly interconnected.

A.3.1 Organelles of the late secretory pathway

There is also a close connection between the early and late secretory pathway because the Golgi apparatus is part of both. In mammalian cells, post-Golgi

organelles include the early endosomes (EE), the late endosomes (LE) and the recycling endosomes (RE), and transport is mediated by vesicular or tubular membrane carriers. Compared with the ER-Golgi interface, post-Golgi trafficking is much less understood and nomenclature of endosomal compartments differs for the plant field (Rojo & Denecke 2008). In the late secretory pathway endocytosis will happen at the plasma membrane level and the transport from that structure to other organelles is mediated by clathrin coated vesicles. Conventionally it is believed that the transport from the Golgi onwards is also mediated by the same type of vesicles and that it happens at the latest region of that organelle, known as the *trans*-Golgi network (Traub & Kornfeld 1997). Generally, the transport from the Golgi to the termination points can happen directly or indirectly, via intermediate endosomal compartments. In plants the prevacuolar compartment (PVC) is distinguished from the late pre-vacuolar compartment (LPVC) that could eventually fuse with the vacuole (Foresti et al. 2010), but the opinions regarding the nature of the recycling endosome and its relation to the early endosome are currently divided (Uemura 2016; De Marcos Lousa et al. 2012).

A.3.2 The Golgi apparatus: connecting organelles

Due to its central position in the pathway, the Golgi apparatus can be considered the main sorting point from the biosynthetic perspective. As mentioned before, the Golgi stacks need to be in perfect synchrony because the organelle is effectively the crossroads for the exocytic and endocytic pathways; receiving and sending cargo from/to the ER, plasma membrane (PM) and other endosomal compartments. In most of eukaryotic organisms the Golgi has a very similar and conserved morphology consisting of a conglomerate of many membranous stacks, which are maintained very close together by structural proteins (Mironov & Pavelka 2008; Mowbrey & Dacks 2009).

Currently two models speculate on how the Golgi stacks are formed and can maintain their integrity via either an intra-Golgi bi-directional trafficking machinery or purely by a polarized retrieval vs progression of cargoes. The vesicular transport model supports the idea that two independent COPI populations are present and act in counter-flux. A COPI intra-Golgi antegrade population and a COPI retrograde population responsible for intra-Golgi and Golgi-ER transport.

This model was long proposed (Rothman & Wieland 1996) and it is has been extensively tested over the years until today (Dunlop et al. 2017).

The second model, knows as the cisternal maturation model was observed in yeast (Losev et al. 2006) and strongly questioned because it couldn't be observed in mammalian cells, although observed in other organisms (Stefanic et al. 2009). This model proposes that ER-derived COPII vesicles containing secreted, vacuolar and ER resident cargoes fuse with each other and retrieval of ER-residents would progressively deplete the Golgi cisternae of these molecules and thus explain the observed Golgi polarity. This means that the composition of the cis-cisternae of the Golgi is very similar to that of the ER lumen in the second model. However, this model is too simple and ignores influx from post-Golgi organelles. To illustrate established concepts and to highlight thorny open questions and prominent disagreements, the remainder of this introduction focuses on the transport events that direct proteins to different locations in the pathway.

A.4 **Protein transport through the secretory pathway**

To mediate vesicle transport and segregate proteins to different routes requires the interplay of many proteins. The following sub-sections will be dedicated to discuss general protein sorting principles and introduce the major gene families and/or complexes that mediate protein trafficking.

A.4.1 ***Bulk-flow*** **and active transport of proteins from the ER**

There are two different models used to explain how proteins are transported from the ER to the later components of the secretory pathway. Early research into protein sorting signals was carried out at a time when it was still unclear why different proteins were secreted at different rates.

A number of landmark studies led to the concept of *bulk-flow*, which states that soluble proteins in the secretory pathway do not require signals to reach the plasma membrane. The *bulk-flow* model was firstly proposed after results shown by Wieland et al. (1987) using synthetic peptides as secretory markers. These were introduced into the cell by diffusion across the cell membranes but trapped in the ER lumen by glycosylation, rendering them membrane impermeable so that they can only reach the plasma membrane via the secretory pathway. The

authors demonstrated that the rate of secretion of those markers is considerably faster than of the classic secreted proteins and concluded that secretion is a default mechanism and would be unselective or passive.

This concept was strongly supported by the discovery of ER retention signals (Munro & Pelham 1987), short C-terminal tetrapeptides such as KDEL and HDEL, deletion of which resulted in the secretion of the truncated proteins. Similar results were obtained with vacuolar proteins deprived of their vacuolar sorting signal (Machamer & Rose 1987). The combined results suggested that secretion is the default pathway for soluble proteins, whilst arrival or retention at specific locations, such as the lysosome/vacuole or the endoplasmic reticulum depend on the presence of sorting signals (Johnson et al. 1987; Valls et al. 1987; Figura & Hasilik 1986; Munro & Pelham 1987).

The default pathway was first demonstrated in plants by introducing cytosolic prokaryotic proteins to the ER lumen by fusion to an N-terminal signal peptide (Denecke et al. 1990). The resulting proteins were secreted at different rates, according to the ability to diffuse or interact with stationary components in the ER. Essentially the experiment indicated that proteins would be normally exported/secreted and that retention would be achieved/mediated by signals (Denecke et al. 1992).

Although the discovery and analysis of ER retention signals strongly supported the *bulk-flow* model, the default concept for secretion was not generally accepted. The active transport model postulates that proteins are selected by ER export signals, the different affinities of which explain different transport rates observed for different proteins. *In vitro* generated COPII vesicles appeared to contain selective cargo compositions (Barlowe et al. 1994), and transport studies on viral transmembrane proteins suggested a crucial role of di-acidic signals (DXE) in the cytosolic tail domain for ER export (Nishimura & Balch 1997). In plants, experiments *in vivo* using point mutagenesis and cell live image approaches have also shown that the di-acidic export signal plays very important role in type I, type II and multi-spanning membrane proteins (Hanton et al. 2005). Two ER-export motifs are commonly founded in mammalian cells, controlling type I membrane proteins localisation, a diphenylalanine (FF) export motif in the C-terminus or a di-acidic signal (Nufer et al. 2002; Contreras et al. 2000). Based on those studies is appropriate to say that the active transport model has been

well documented in the literature and proved that if the object of study is a membrane spanning protein the model is very suitable to explain the secretion, sorting and retention. However, it should be noted that DXE and FF signals are seen on the cytosolic side of membrane proteins only.

To conclude, whilst ER export signals have been identified for membrane spanning proteins (Nufer et al. 2002), there is no evidence for export signals of soluble proteins, supporting the *Bulk-flow model* (for review consult Vitale & Denecke 1999). The export of proteins from the ER continues to be controversial and with new information emerging every year. We should not discard the possibility that both models act together to achieve a maximum efficiency transport (for review consult Vitale & Denecke 1999; Pimpl & Denecke 2000).

A.4.2 Vesicle mediated protein traffic

Proteins and membranes shuttle between these organelles via the so-called transport vesicles, protein-coated membrane spheres that are short-lived and highly enriched in protein-cargo (Bonifacino & Glick 2004). Protein selection for specific vesicle shuttles is mediated by sorting receptors, a class of membrane spanning proteins that bind to ligands in the lumen and interfaces with cytosolic proteins to form vesicle coats. Vesicles bud from the donor membrane, are then transported to a specific acceptor organelle, and upon fusion with its delimiting membrane, cargo and receptors are released. Finally, the receptors dissociate from their ligands and are recycled in another type of vesicle to return to the original donor compartment and start a new transport cycle (Pearse & Bretscher 1981; Dancourt & Barlowe 2010).

Two different types of vesicles, coat protein complex I and II (COPI and COPII), are responsible for retrograde and anterograde transport at the ER-Golgi interface (Paul & Frigerio 2007). This transport cycle is difficult to study, because when COPI transport is altered, the anterograde COPII pathway is disturbed as well. This was shown by the use of Brefeldin A (BFA), which is a drug that directly influences the activity of a guanine nucleotide exchange factor that controls the GTPase ARF1, an extremely important protein necessary for the formation of COPI vesicles (protein ARF1p; Pimpl et al. 2000). The use of dominant-negative mutants of ARF1 reveals a similar interdependence of COPI and COPII mediated vesicle pathways (Stefano et al. 2006), suggesting that anterograde and

retrograde membrane trafficking critically depend on each other for the efficient recycling of transport machinery components.

COPI and COPII vesicle formation is triggered by two different mechanisms, controlled by the ARF1 and Sar1 subfamily GTPases respectively that act as molecular switches. In addition, Type I p24 proteins have been identified as important components of COPI and COPII vesicles and could also act as cargo receptors (for review consult Paul & Frigerio 2007).

It is important to mention an intermediate organelle, consisting of a complex membrane system, found between the ER and the *cis*-Golgi in mammalian cells, known as ERGIC (Appenzeller-Herzog & Hauri 2006). The ERGIC-53 protein, a membrane protein characteristic from this compartment, is highly conserved and found in all types of mammalian and animal cells, but not in plants (Hauri et al. 2000). It is possible that plants have a fundamentally different ER-Golgi transport cycle, and that the function of the mammalian ERGIC is carried out by the *cis*-cisternae of the Golgi apparatus (Phillipson et al. 2001; Robinson et al. 2015).

A.4.2.1 Coated protein complex I (COPI) – Discovery, mechanisms and components

The so called cell-free system, proposed more than thirty years ago by Rothman's group, was very important in the early studies aiming to understand, what they believed to be, the traffic of proteins within the Golgi apparatus (Rothman et al. 1984; Balch et al. 1984). The assay proposed by the group employed the VSV-G protein (vesicular stomatitis virus encoded glycoprotein G) as a cargo molecule and a glycosylation defective mutant (15B CHO cells) missing GlcNAc transferase I as tool to measure cargo progression in the pathway. This allowed the group to measure the transfer of G-proteins between successive Golgi compartments by using Golgi fractions from VSV-infected 15B mutants (called "donor fraction") and mixing them *in vitro* with Golgi fractions from non-infected wild type cells (called the "acceptor fraction"). Incorporation of ^{3}H-GlcNAc, from pre radio tagged UDP-^{3}H-GlcNAc, into the VSV-G-protein in 15B mutants allowed to test the potential vesicle trafficking between the two different fractions of Golgi membranes (Balch et al. 1984).

VSV-G transport studies revealed the presence of clathrin coated and non-clathrin coated vesicle structures at the *trans*-Golgi cisternae, the latter of which

contained the VSV-G protein and could represent a new type of *bulk-flow* carrier for transport to the cell surface as opposed to the endosomes (Orci et al. 1986; Griffiths et al. 1985). The Coatomer or “Golgi coat promoter” is a complex of seven cytosolic proteins, which was firstly described and identified by Waters et al. (1991) as the molecular basis for the formation of non-clathrin coated vesicles *in vitro.* The identification of this complex was possible after the proposition of a novel technique, which allow the large-scale purification of the vesicles and further identification of the individual components (Malhotra et al. 1989). In the following years it was shown that the depletion of coatomer can influence the ratio of budding from Golgi cisternae *in vitro*, indicating that an interaction between ARF (ADP ribosylation factor) and coatomer could be one necessary step to deform the membrane and form the vesicle (Orci et al. 1993).

The first studies lead to the belief that non-clathrin vesicles were involved in the exocytic pathway, since galactose residues were believe to be added in the *trans* compartments of Golgi and N-acetylglucosamine (GlcNAc) in the *medial* cisternae (Balch et al. 1984). However, subsequent work revealed that coatomer coated vesicles mediate retrograde transport, from the Golgi back to the ER. Firstly, it was shown that the addition of the cytoplasmic domain of an yeast ER resident membrane protein (WBP1), which contains a di-lysine motif, was capable of redirecting a mammalian secreted protein (Tac antigen) acting as an ER retention signal, the motif was shown to directly interact with coatomer components (Cosson & Letourneur 1994). These findings are not in conflict with the earlier VSV-G protein modification assays that indicated vesicle trafficking, as modification of glycans could have been due to arrival of GlcNAc transferase I from later compartments, causing the modification of radiolabelled VSV-G-protein in earlier compartments. Furthermore, it was shown that in general the majority of proteins recycled from the Golgi back to the ER are fully folded proteins, consistent with the idea of exported proteins recycled back from the Golgi apparatus (Lee et al. 2004).

After the discovery of a different type of non-clathrin coated vesicle carrier, which mediates the ER to Golgi anterograde transport, the coatomer coated vesicles were re-named as COPI (Barlowe et al. 1994). The formation of the COPI vesicles is achieved via a complex interaction network between activating proteins, exchange factors and lipids. More specifically the cascade of events

the leads to the formation of COPI starts with the activation of the GTPase ARFp1, which is found in the cytosol, by the guanosine nucleotide exchange factor Gea1/2p and ARNO leading to the recruitment of the coatomer complex and ARFGAP1 which promotes ARF-mediated GTP hydrolysis (Hsu & Yang 2009). It is important to mention that the coatomer complex can interact with other proteins that maybe are necessary for the vesicle formation, for instance proteins from the p23/24 families (Lavoie et al. 1999). *In situ* localization of those vesicles and *in vitro* formation in a plant system was previously shown and it was demonstrated that coatomer is recruited to the donor membrane and the vesicles formation occurs mostly in the *cis* region of the Golgi (Pimpl et al. 2000).

The retrieval mechanism from the Golgi to the ER demands a very well-regulated system. For that reason, various molecular targeting signals can be found. Transmembrane proteins that need to be retrieved to the ER share a di-lysine KKXX motif that can interact with COPI coat components (Cosson & Letourneur 1994), an aromatic ϕ-XX-K/R/D/E-ϕ-COOH (McCartney et al. 2004) and a di-arginine motif XXRR or RXR (Zerangue et al. 1999) are some examples. Recent studies have shown that the last mentioned motif could instead be related to Golgi retention (for review consult Banfield 2011). And lastly, ER soluble proteins have a KDEL motif in mammals (Munro & Pelham 1987) and related signals, and are generally thought to be retrieved from the Golgi back to the ER, as discussed before (for review consult, Robinson 2003).

The puzzling argument for different sub-populations of COPI vesicles

The initial discovery and characterization of COPI components lead a constant flow of publications every year aiming to explain all the interactions and cascade of events that take place during vesicle budding and fusion, respectively from and to different organelles. The main mechanism controlling COPI vesicles is now very well established, notwithstanding, discussions about the possibility of these vesicles being involved also in intra-Golgi transport has emerged in the late 90s and is still debated (Rabouille & Klumperman 2005; Cottam & Ungar 2012).

The initial proposal for a bi-directional traffic mediated by independent COPI types was based in the observation that both anterograde and retrograde cargo could be present in these vesicles, which were not exclusively located at the cis-

Golgi. These observations were made by the use of electron microscopy and immunocytochemistry and targeting three molecules, proinsulin VSV-G proteins and the KDEL receptor and using the so called cell-free system (Orci et al. 1997).

Lanoix et al. (2001) were responsible for conducting an important experiment observing intra-Golgi anterograde traffic mediated by COPI vesicles. Their observations, in conjunction to early ones as pointed by Pelham (2001), culminate in the proposal that sub-populations of COPI exist. Their study was conducted in mammalian cell lines *in vitro* and took advantage of Golgi markers believed to be accumulated at different sub-domains of this organelle (cis, medial and trans stacks). That work tracked how these proteins progress thought the stacks and saw that the sorting of Man II, p24s and GS28 happens in a selective manner into more than one group of COPI, interestingly these vesicles did not contained anterograde cargo.

The existence of different COPI sub-populations was also corroborated by the observation that different members of a family of coiled-coil proteins called Golgins have more affinity to act as tethering factor for specific COPI types, in summary CASP or p115 golgins were not found to interact with the same vesicles (Malsam et al. 2005).

Years later after the first observation proposing the existence of different sub-populations COPI vesicles, researchers have presented new evidence for this dual distribution. An *in vivo* study conducted in two different organisms, an alga and a higher plant, and by using electron tomography analysis combined with immunogold labelling followed by a structural characterization using six criteria of vesicles. This study has shown that the retrograde traffic from the Golgi to the ER and intra-Golgi transport is mediated by two types of COPI vesicles, COPIa and COPIb respectively (Donohoe et al. 2007).

A.4.2.2 Coated protein complex II (COPII) – Discovery, mechanisms and components

Using an ingenious genetic screen to isolate a large collection of yeast mutants defective in protein secretion, scientists were able to identify many important genes controlling anterograde transport in the secretory pathway (Novick et al. 1980). These secretory mutants failed to export invertase and acid phosphatase, exhibited strongly reduced growth rates and a higher density compared to wild

type yeasts, which permitted a simple but effective enrichment method for the isolation of secretory mutants. The results revealed that at non-permissive temperatures the so-called *SEC* mutants accumulate excess of ER-like membranes and were classified into 23 complementation groups each representing a gene product involved in the control of protein secretion in yeast.

Follow-up studies lead to the identification of the genes involved in the formation/fusion of the new vesicles, and the proposition of a model for protein transport from the ER to the Golgi apparatus (Kaiser & Schekman 1990; Rexach & Schekman 1991; Barlowe et al. 1994). Nine *sec* mutant strains were analysed, grown at 24°C and then shifted to 37°C for fixation and further analyses. As expected, it was possible to verify the accumulation of vesicles when grown at non-permissive temperature. The strains were separated in different classes accordantly to the number, size, density and position of the vesicles. Genes interactions were tested proposing for the first time the biochemical function of the *SEC* genes (Kaiser & Schekman 1990). Rexach & Schekman (1991) show a vesicular intermediate between the ER and the Golgi, in a yeast cell-free *in vitro* transport reaction. The results suggested that vesicle budding from the ER, targeting and fusion to the Golgi membranes requires a set of genes, amongst which Sec12 and Sec23 are required for vesicle budding but not Golgi targeting and fusion (Barlowe et al. 1994). The *in vitro* isolated vesicles were non-clathrin coated but distinct from the earlier described coatomer, which then led to the proposed name of COPI for coatomer and COPII for the new anterograde ER to Golgi vesicle class.

COPII vesicle formation, on the ER membrane, starts with the activation of SAR1p by the guanine nucleotide exchange factor Sec12p, which leads to the recruitment of the Sec23p-Sec24p complex and subsequently the Sec13p-Sec31p complex to the cytosolic face of the ER membrane forcing its deformation (Tang et al. 2005; Dancourt & Barlowe 2010). Further studies have shown that the formation of those vesicles is restricted to specific areas of the ER, known as ER exit sites, which are dedicated to the export from ER to Golgi (ERES; daSilva et al. 2004; Tang et al. 2005; Lee et al. 2004; Matheson et al. 2006). Hanton et al. (2007) have shown that there is a very close relationship between *de novo* formation of ERES regions and the COPII vesicles, since the vesicles are recruited to previously existing ERES regions to start the formation

of new export sites. Finally, a more recent model called “hug-and-kiss” aims to elucidate the delivery of proteins from the ERES to the cis-Golgi. Using high speed and resolution microscopy it was shown that, in the model organism *S. cerevisiae*, cis-Golgi and ERES can approximate to such level that the exchange of cargo by COPII vesicles can happen without vesicle dissociation to the cytosol, but instead upon the contact the cages collapses releasing their contents (Kurokawa et al. 2014).

There are a number of similarities between the mechanisms for protein sorting in COPI and COPII vesicles, in both cases it is based on the presence of peptide signals in the cargo molecules. Nonetheless, the signals recognised by COPII are different, which makes sense to avoid miss-sorting. Matheson et al. (2006) summarizes some signals that influence the anterograde transport, which are: a di-acidic DXE motif, a di-basic R/K-X-R/K and a di-hydrophobic YF motif. These classes of signals mediate direct interactions with the COPII coat in the cytosol, a feature denied to soluble proteins which are confined to the ER lumen. The mechanism responsible for the ER export of soluble proteins in the lumen of the secretory pathway is most likely by passive diffusion into a growing COPII vesicle lumen, in other words by *bulk-flow* (Denecke et al., 1990, Phillipson et al., 2001). This does not rule out that certain soluble proteins may have export signals, but these remain to be identified, together with a lumenal domain of a receptor that may interact with it and then trigger COPII vesicle budding. It is entirely possible that such motifs exist, which makes the continuous studies in the early secretory pathway very important.

A.4.2.3 p24 proteins

A family of low molecular weight (~25kDa), single transmembrane domain proteins, termed p24 proteins, were found to be abundant constituents of COPI and COPII vesicles and thought to play an important role in the ER-Golgi interface (Lavoie et al. 1999; Kaiser 2000; Strating et al. 2009) and in particular COPI vesicles formation (Gao et al. 2014). Other proposed functions of p24 proteins in the early secretory pathway include cargo protein selection/packaging and biogenesis/maintenance of the Golgi apparatus (Paul & Frigerio 2007). In yeast 8 members of this proteins can be found, 10 in humans and in *Arabidopsis thaliana* 11 coding sequences were found. Those were classified in two subfamilies (p24δ and p24β; Chen et al. 2012).

The proteins can be found and are abundant along the early secretory pathway, localizing differently depending of the subfamilies and the organism (Chen et al. 2012; Langhans et al. 2008). Much has been suggested about possible functions for those proteins, however there are many questions to be answered (for review consult Dancourt & Barlowe 2010). In *Arabidopsis thaliana* it was first shown that p24δ5 and p24β2 localises on the ER and Golgi, respectively; and that the interaction between them can facilitate their traffic between Golgi-ER. A later study revealed that p24δ9 and p24β3 again have different localisations and it was proposed that p24 proteins in *Arabidopsis* could form heteromeric complexes, which would allow the coupled traffic of these proteins (Montesinos et al. 2012; Montesinos et al. 2013). Recently, Aniento's group (Montesinos et al. 2014) proposed a direct relation between the ERD2 and the p24δ5 protein, suggesting that the former facilitate the retrograde transport of the H/KDEL receptor. The same group has also suggested that a quadruple *knock-out* of p24s from the δ-1 family can directly influence the distribution of ERD2 along the secretory pathway, causing its accumulation at the Golgi apparatus and disrupt the retention of its ligands leading to increased secretion of BiP (Pastor-Cantizano et al. 2018). In the same paper the authors also suggest that the absence of those P24s can have a direct influence in the unfolded protein response (UPR) and one of the COPII subunits (*SEC13A*).

Recently, a new model has been proposed offering a possible mechanism directly linking P24 proteins with COPII vesicles and ER retention. To generate these data Ma et al. (2017) combined *in vitro* binding assays with the drug 4-phenylbutyrate (4-PBA) and p24 proteins. The drug in question has been previously used to possibly restore trafficking of misfolded proteins (Balch et al. 2008; Tveten et al. 2007). That new model introduces the idea that it is possible to block the COPII machinery using 4-PBA because that drug competes with p24 proteins and has higher binding affinity to the COPII coat, ultimately leading to the mis-packaging of ER residents.

Studies around p24 proteins in plants have been published in the past few years demonstrating that much knowledge is being produced and many more questions are still to be answered, and the proposed direct interaction of p24 with ERD2 could be a very interesting subject to be more extensively tested.

A.4.2.4 Rab GTPases and SNAREs

After vesicle budding from the donor membrane, protein coats are thought to be recycled to allow membrane fusion with the acceptor membrane and to start new rounds of vesicle budding. However, further proteins are required to mediate vesicle transport and docking to the correct acceptor membrane. This is carried out by RAB GTPases (Ras-related in Brain guanosine triphosphatases) and SNAREs (Soluble N-ethylmaleimide-sensitive factor attachment receptor proteins) (Bonifacino & Glick 2004).

The GTPases are very important molecules that play different roles within the cell, including intracellular trafficking (Bourne et al. 1990). The RAS superfamily can be distinguished in two different families, in our context the more important family would be the YPT/Rab family, which directly influences the secretory pathway (for review consult Segev 2001). The Rab family of proteins, which has more than 60 members in human, are extremely important proteins acting as "*switches*" controlling protein traffic (Stenmark 2009). With the exception of the GTPases Sar1 and ARF1, which form a small subgroup specialised in vesicle budding from the ER membrane and the Golgi membrane respectively (see Chapter 5), the vast majority of these low molecular weight GTPases control later and more diverse steps in the vesicle transport reactions, including membrane tethering and association with the cytoskeleton (Kelly et al. 2012). The Ras-family of proteins was first identified in yeast, YPT gene family and YPT1 protein (Schmitt et al. 1986). Few years later a human homologue was characterized, Rab1 (Segev et al. 1988), and in the same study these proteins were shown to be associated with the secretory machinery.

In addition to Rab GTPases, SNAREs are important proteins controlling the actual membrane fusion process. Many different proteins can be found in this family, mostly having a long cytosolic N-terminal domain and a C-terminal transmembrane anchored domain. The initial classification and distribution of SNARE proteins in subfamilies was guided by the structural presence of a variety of SNARE motifs. SNAREs were initially distribuited in two large groups, vesicle SNARES (v-SNARES) and target SNARES (t-SNARES), based on their topology and function. Later an additional classification was proposed in addition to the first one, this time based on the amino acid composition and structural features

of the proteins, splitting SNARES into R-SNARES and Q_a/Q_b/Q_c-SNARES (Fasshauer et al. 1998).

SNARES function is ruled by a general machinery involving few fundamental steps, including: vesicle budding from donor membrane, vesicle fusion to the membrane of target compartment and finally the recycling of SNARES. This process involves the recruitment of tethering factors and SNARES to a budding vesicle, followed by the formation of t-SNARE complex. Upon vesicle fusion, a tetrameric SNARE complex is formed between t-SNAREs and v-SNARE. Finally, the presence of accessory proteins, such as SNAPs, will release the SNAREs and allow for these to be used in a new round of budding-fusion (Malsam & Sollner 2011).

A subfamily, termed Syntaxins, is a group of SNAREs with high homology and can be found in many organisms, including in *Arabidopsis thaliana* (Sanderfoot et al. 2000). Suppressor screens in yeast have shown that overexpressed syntaxin Sed5 can inhibit ER export and secretion (Hardwick & Pelham 1992). The homologue of this gene in *Arabidopsis thaliana* is the gene SYP31, which can also inhibit secretion in tobacco protoplasts (Bubeck et al. 2008). Inhibitory SNARE function was proposed for other syntaxins (Foresti et al. 2006) but functional studies in plants are mostly lacking. A systematic study of sub-cellular localisations of a diverse range of SNAREs revealed interesting insights, including the existence of a single syntaxin for the ER, but multiple syntaxins for the plasma membrane and the vacuolar route (Uemura et al. 2004). A proposal for the relationship between ER exits and ER import sites has been made (Lerich et al. 2012) but further research is necessary to complete the full cycle of transport events at the ER-Golgi interface.

A.5 Sorting signals for soluble and membrane spanning proteins

One of the main questions in cell biology is related to the understanding of the complete transport cycle of protein sorting receptors, the exact roadmap of transport between organelles and the way in which organelles maintain their identity. The sequential nature of protein transport through the secretory pathway includes protein translocation across the ER membrane, export to the Golgi apparatus and then further delivery to either the plasma membrane or the

vacuoles/lysosomes (Palade 1975). Soluble proteins are transported through organelles via vesicles and are secreted unless specific sorting signals direct them to intracellular locations. (Palade 1975; Brandizzi & Barlowe 2013b; Lee et al. 2004). However, a large number of proteins are not translocated fully across the ER membrane but instead become embedded in the ER membrane, and they are also transported to various places in the cell. The transport of soluble proteins is very well characterized, but the sorting, delivering and retention of membrane proteins seems to be much more elaborate (Rojo & Denecke 2008).

Not all proteins are targeted to the two end locations of the pathway and a number of soluble and membrane proteins have to reside in various intracellular organelles of the pathway, including the ER, the Golgi apparatus and the endosomes. This means that in addition to segregating secreted and vacuolar proteins, the system must have mechanisms to mediate high steady state levels of resident proteins in specific organelles, despite the constant influx and efflux of membranes and cargo. In addition, the transport of soluble proteins is mediated by membrane spanning receptor molecules, and these must also contain sorting signals to navigate the secretory pathway.

A.5.1 Organelle retention motifs for membrane spanning proteins

Each of the two main organelles comprising the early secretory pathway need to maintain a different set of specific membrane proteins to ensure that their structural and functional capabilities are properly working. For that reason, different signals have been proposed to either interact with COPI vesicles or COPII vesicles and mediate the retention of proteins at the ER or Golgi, respectively (Gao et al. 2014; Cosson & Letourneur 1994; Contreras et al. 2004). The interaction of short motifs with the COPI components has been long proposed and experimentally tested over the years.

A well-established and recognized motif for type I membrane proteins is the –K(X)KXX, originally discovered in mammals (Nilsson et al. 1989) but it can also be found in different proteins present in plant cells, especially in the members of the p24 family (Gao et al. 2014; Chen et al. 2012). It has been shown before that a di-lysine motif can mediate the retention of membrane proteins in the ER and therefore is of extreme importance for ER residents (Jackson et al. 1990) via its possible direct interaction with coatomer (Letourneur et al. 1994; Cosson &

Letourneur 1994). The discovery of the KKXX motif was one of the main arguments to question the earlier suggested role of COPI mediated transport in anterograde intra-Golgi transport.

More recently a new retention motif, which seems to be specific for Golgi localisation, has been proposed to operate in plants and other organisms and to interact with COPI vesicles. The KXD/E motif has been identified in different proteins from plants, humans and yeast. In *Arabidopsis* this motif is highly conserved in the EMPs (Endomembrane proteins 1 to 12). More interestingly is that fact that the transplantation of this C-terminal signal caused the redistribution of other proteins to the Golgi apparatus (Gao et al. 2012; Gao et al. 2014). The exact mechanism responsible for the retention in this highly dynamic organelle is as yet unknown.

Finally, the retention via terminal arginine based motifs, for Type II membrane proteins, has also been proposed to play an interesting and contradictory role since it is possibly mediating ER retention (Boulaflous et al. 2009). However this signal can be found in different Golgi resident proteins as well (Banfield 2011; Tu et al. 2008) and further research is needed to characterise the underlying mechanisms for organelle retention.

A.5.2 Alternative models to mediate retention of membrane proteins

There is a large number of membrane proteins that do not seem to contain an evident motif to cause their organelle retention and different models have been proposed over the years to explain signal-independent protein localisation (for review consult Banfield, 2011).

The retention of membrane proteins at the Golgi apparatus is especially difficult to understand considering that this organelle is constantly receiving proteins from both the ER (biosynthetic pathway) and the Plasma membrane (endocytic pathway). Apart from that it also needs to maintain its structural and functional integrity, because of that is highly possible that multiple mechanism operates together to keep the steady-state of Golgi specific proteins intact. Indirect retention mechanisms have been suggested to be caused by protein aggregation and formation of hetero/homo -dimers (Tu et al. 2008), (Kin recognition model - Nilsson et al., 1993).

Organelle retention could also be based on exclusion mechanisms, for instance by matching the transmembrane length to the thickness of specific organelle membranes. In this respect it is noteworthy to mention that the apical plasma membrane was shown as the thickest of all membranes whilst the ER membrane is thinner in comparison (Mitra et al. 2004). Short transmembrane domains may thus be incompatible with post ER membrane thickness and could be excluded, but only very few studies have been conducted (Brandizzi et al. 2002) and a working model is yet to be formulated. Differences in the lipid-composition of membranes could therefore have a profound influence on the distribution of membrane proteins (Schaecher et al. 2008; Sharpe et al. 2010; Patterson et al. 2008; Brandizzi 2002).

It is entirely possible that a combination of different factors is required to achieve optimal subcellular localisation of certain classes of molecules (Rojo & Denecke 2008), as illustrated by the multiple set of sorting signals found on the plant vacuolar sorting receptor (daSilva et al. 2006; Foresti et al. 2010; Gershlick et al. 2014).

A.5.3 Receptor mediated retention of soluble proteins in the secretory pathway

Soluble proteins in the secretory pathway cannot directly interact with cytosolic coat proteins and rely on membrane spanning sorting receptors to form interfaces. The discovery of signals for cell retention, for instance in the ER (Munro & Pelham 1987) or the vacuole (Valls et al. 1987) strongly supported the proposal that secretion occurs by default (Wieland et al. 1987), via passive diffusion or *bulk-flow*. The search for receptors involved biochemical methods, leading to the discovery of the mannose-6-phosphate receptor (MPR) for lysosomal protein sorting in mammals (Sahagian et al. 1981) or vacuolar sorting receptors in plants (Kirsch et al. 1994). Genetic approaches in yeast led to mutants defective in vacuolar sorting (VPS10 - Marcusson et al. 1994) and ER retention (ERD2, ER retention defective 2 - Semenza et al. 1990). In this chapter the best-known classes of sorting signals and their sorting receptors will be introduced. Whilst vacuolar protein sorting has evolved differently in plants, yeasts and fungi, the signals and the machinery involved in ER retention appear to be extremely conserved, and this topic will be discussed separately as it leads directly to the research work explored in this dissertation.

A.5.3.1 Targeting of proteins to endosomal compartments - Mannose 6-phosphate receptors

Mammalian lysosomal proteins can be separated from secreted proteins by the addition of a mannose 6-phosphate residue in a pre-Golgi compartment, which allows for the dependent targeting of these proteins to their final destination. This process is receptor mediated, which can recognize the specific target residue and consequently binding to ligands, in what is called the Man-6-P recognition system (Figura & Hasilik 1986). The recognition of ligands occurs in the latest stacks of the Golgi and upon binding and export in clathrin coated vesicles, the receptor will release ligands when it arrives in a pre-lysosomal compartment due to the pH difference of the organelles that allows for complex to dissociate (Duncan & Kornfeld 1988).

In most mammals this transport is carried out by two distinct Man-6-P receptors, MPRs. Interestingly, it has been shown that both receptors, CI-MPR and CD-MPR, are able to bind to Man-6-P glycoproteins, but the receptors have different affinities to subsets of ligands (Sleat & Lobel 1997). The receptors also differ in size, 46 and 300 kDa for CD-MPR and CI-MPR respectively, but share the same topological structure and sequence homology in their lumenal domains (Kornfeld 1992).

The retrieval of MPRs from the pre-lysosomal compartments after cargo releasing was poorly understood until genetic approaches identified the retromer complex in yeast (see below). It is composed of three mVPS proteins (mVPS35, mVPS39 and mVPS267) and two Sortin nexin proteins (SNX1 and SNX2). This complex is considered to be crucial since its interaction and consequently recycling of CI-MPR allows for the receptors to start a new round of sorting ligands (Seaman 2005). Recycling from the early-endosome back to the *trans*-Golgi by retromer is a well-accepted mechanism to explain this fundamental step of MPR's life-cycle, but the regulation of retromer complex is still a matter a much debate (Burd & Cullen 2014).

A.5.3.2 Targeting of proteins to the yeast vacuole

After the discovery and initial characterization of lysosomal targeting mediated by MPRs two independent research groups decided to investigate post-Golgi specific targeting in the simple eukaryote *S. cerevisiae*. Unlike animals yeasts have a lysosome-like vacuole, which also receive specific hydrolases, such as

carboxypeptidase Y (CPY). For that reason, both research groups decided to study how this particular protein is sorted from proteins designated to be secreted.

CPY targeting to the yeast vacuole was examined by hybridization and a putative vacuolar targeting sequence was isolated, deletion of this signal from wild-type CPY lead to its missorting (Johnson et al. 1987). In parallel, via *cis*-acting mutation of the gene encoding for CPY, missorting of the protein was also observed and linked to the N-terminal propeptide of CPY, which was assigned different functions including vacuolar targeting (Valls et al. 1987).

The discovery of a specific signal controlling the traffic of CPY to the vacuole led to the conclusion that a putative receptor to recognise this signal was due to be discovered. However, it took many years until the isolation of the first protein to act as the elusive receptor. A large number of candidates where found via different genetic screenings of yeast, for instance *vps* and *pep* mutants (for review, Bowers & Stevens 2005). A mutational analysis of candidates genes belonging to the "*vps* mutants" collection looking for mutants that would specifically affect CPY targeting to the vacuole. This second screen brought to attention one particular gene, the VPS10 gene and its gene product Vps10p. A meticulous study showed that the protein subcellular localization as well as its structure were as expected, but mostly importantly that the Vps10p had high binding affinity to CPY in a chemical cross-linking experiment (Marcusson et al. 1994). Further work on the sorting of VPS10 led to the discovery of the retromer complex (Seaman et al. 1998).

A.5.3.3 Vacuolar sorting in plants

Analysis of vacuolar sorting signals in plants started in the late 80s and was based on the comparison of secreted and vacuolar variants of plant hydrolases involved in defence reactions against microbial pathogens (Shinshi et al. 1988). It soon resulted in the discovery of C-terminal and N-terminal processed peptides that contain specific vacuolar sorting signals. Some other signals were shown to be part of the mature protein and internally localised (Matsuoka & Neuhaus 1999). The search for the plant vacuolar sorting receptor was carried out by a pure biochemical affinity approach, in which clathrin coated vesicles from developing peas were detergent-solubilised and passed over an affinity column

carrying one of the better characterised vacuolar sorting signals, leading to the purification of BP80, a binding protein of 80 kDa (Kirsch et al. 1994). The *in vitro* binding properties of this protein were further confirmed and characterised (Kirsch et al. 1996), leading to the cloning of the plant vacuolar sorting receptor (VSR).

With the establishment that vacuolar targeting was mediated by interactions between specific signal-motifs and cargo receptors the next step was to investigate how the receptor executes its function. For that it was necessary to answer fundamental questions, such as, where does ligand-binding occur? How is ligand-release triggered? How is the receptor traffic controlled?

It was though that BP80 would bind to ligands at the Golgi and releasing of cargo would occur at a pre-vacuolar compartment (PVC), avoiding acid degradation of the receptor, similarly to MPRs (Seaman 2005). To allow for this model to work BP80 would need to efficiently undergo anterograde transport after binding to its ligands and retrograde transport after ligand-release. This cycle was thought to be signal mediated as well. A detailed *in vivo* study revealed that indeed, as predicted, BP80 C-terminus contains multiple signals to allow for the efficient transport of the receptor in between organelles (daSilva et al. 2006), including ER export, transport to the PVC, recycling from the PVC (Foresti et al. 2010) and endocytosis (Gershlick et al. 2014) in the case of receptor mis-targeting.

A general model for plant VSRs was established and it was assumed that the receptor would bind to ligands at the later Golgi stacks and finally releasing it at the PVC, which would eventually fuse to the vacuole (daSilva et al. 2006; Martinière et al. 2013). However, years later it was demonstrated that the PVC was not the latest stage of maturation before vacuolar fusion and that CCVs could bud-off from either the Golgi cisternae or from the TGN, carrying VSR-ligand complexes. Fusion of CCVs to the PVC would cause VSR complex dissociation and finally PVC would mature into a cargo-mainly compartment, termed LPVC, which does not accept CCVs anymore (Foresti et al. 2010).

Our understanding of the complex mechanism for vacuolar targeting in plants, particularly due to the presence of a lytic and protein storage vacuole (Jiang et al. 2001), is far from being accomplished. Like so, in recent studies it was proposed that plant vacuolar receptors release their cargo in a much earlier stage

of the secretory pathway, the *trans*-Golgi network (TGN) (Früholz et al. 2018; Künzl et al. 2016). Further work is required to establish which of the conflicting models will prevail.

A.6 Retention of soluble proteins in the Endoplasmic reticulum

The ER lumen is a highly concentrated protein solution with gel-like properties (Koch 1987; Booth & Koch 1989). Most of the protein content originates from a low number of very abundant ER resident chaperones rather than the secretory or vacuolar products found in transit. One of the most conserved processes in the secretory pathway is the retention of soluble proteins in the ER. As discussed before, by default, luminal proteins can become incorporated in COPII vesicles and leave the ER frequently (Denecke et al. 1990; Crofts et al. 1999; Phillipson et al. 2001). To explain accumulation in the ER, a mechanism for molecular recognition in the Golgi was suggested three decades ago (Pelham 1988) to identify and retrieve the escaped proteins for transport back to the ER when they carry out their main function in the ER lumen. This section will focus on how these findings were originally made, and how this led to the aims of my dissertation work.

A.6.1 Identifying the mechanism for ER retention of soluble proteins

In an important key-study based on the realization that several ER residents share a C-terminal tetrapeptide KDEL sequence, deletion and transplantation experiments demonstrated that this sequence is necessary and sufficient to mediate protein accumulation in the ER (Munro & Pelham 1987). Soon afterwards it was proposed that ER accumulation was not caused by true retention, but instead dependent on continuous retrieval from the Golgi apparatus. The mobility of proteins in the ER was shown to be independent of the presence of KDEL (Ceriotti & Colman 1988), thus ruling out binding to an ER resident receptor. In addition, KDEL-tagging of the lysosomal enzyme cathepsin D did not prevent post-translational modifications that normally take place in the Golgi apparatus, but yet mediated ER accumulation (Pelham 1988). The data were consistent with the presence of a receptor that binds KDEL proteins in the Golgi apparatus, mediates vesicle trafficking back to the ER, followed by release of the KDEL protein.

Further studies revealed that closely related tetrapeptides can also function in this process and may differ between eukaryotic kingdoms (Pelham et al. 1988; Pelham 1990; Denecke et al. 1992). Zagouras and Rose (1989) and Herman et. al (1990), working with different organisms and proteins, presented results showing that in some cases the insertion of the carboxy-terminal sequence is not the only necessary condition to induce the absolute retention in the ER, and probably other features in the protein structure are necessary for this.

A.6.2 Discovery of the K/HDEL receptor

The receptor that binds the tetrapeptides KDEL or HDEL was identified via an elegant genetic screen in *Saccharomyces cerevisiae* and is encoded by ER retention defective 2 (Semenza et al. 1990), a protein with 7 predicted transmembrane domains which controls the specificity of ER retention (Lewis et al. 1990; Townsley et al. 1993).

The search for the receptor was based on the assumption that yeast would use a similar mechanism as animal cells. A candidate gene that encoded an endogenous ER resident of yeast was the Karyogamie 2 (Kar2) gene product, which is the functional homologue of GRP78 (BiP). However, instead of a C-terminal KDEL, it carried the much-related tetrapeptide HDEL at its C-terminus (Pelham et al. 1988). Interestingly, the HDEL motif was unable to redirect chicken lysozyme to the ER in COS cells, suggesting that the ligand-binding specificity of the yeast machinery is slightly different.

In order to isolate ER retention defective yeast mutants, yeast invertase was replaced by invertase-HDEL by recombination, leading to a strain that was unable to grow on sucrose medium as secreted invertase is required to utilize this carbon source. ER-retention-defective yeasts would secrete the engineered HDEL protein and thus re-gain the ability to utilize sucrose (Pelham et al. 1988; Hardwick et al. 1990). Characterization of several candidate genes cloned by complementation of the mutants with wild type genes to regain ER retention led to the characterization of the ERD2 gene product (Semenza et al. 1990).

It was shown that ERD2p controlled the specificity and capacity of HDEL-mediated retention in yeast (Lewis et al. 1990). To test if the ERD2 gene product was the real receptor, experiments testing its specificity difference between species have been carried and it was shown that *Kluyveromyces lactis* BiP bears

a DDEL sequence at its C-terminus, which was surprising since it indicates divergence between very similar species. In the experiment different levels of retention and secretion were observed when comparing *S. cerevisiae* invertase bearing different motifs (SEKDEL, FEHDEL, YFDDEL or SEEDLN). Both *S. cerevisiae* and *K. lactis* ERD2 genes, normal and over-expressed, showed and proved the divergence on the receptors and that indeed the ERD2 gene encodes the receptor for those motifs since the ERD2 gene of *K. lactis* apparently slightly changes the retention system in *S. cerevisiae* (Lewis et al. 1990).

The mechanism of recycling ligands back to the ER from the early Golgi was initially demonstrated by organelle fractionation of yeast overexpressing HDEL-tagged preproalpha factor protein, which acquired Golgi modifications but was found in ER fractions (Dean & Pelham 1990). Even though the retention system is very efficient, in the same study it was shown to be saturable due to secretion of endogenous HDEL proteins caused by overexpression of HDEL-tagged preproalpha. It is notable that those proteins encoded by the ERD2 genes are highly conserved between species and in general the functionality is to recognize the C-terminal tetrapeptide sequences because soon afterwards the initial discovery a ERD2 homologue was identified in humans and shown to specifically recognize the KDEL motif (Lewis & Pelham 1990).

In more recent studies on the ER retention system in mammalian cells, which have three different ERD2 genes (Raykhel et al. 2007; Alanen et al. 2011), it was shown that not only the four residues in the tetrapeptide C-terminal motif are important for the ER retention, but at least the last six amino acids can determine the efficiency of it. Even more interesting was the fact that apparently there is a “hierarchy” between the multiple motifs, controlling the interactions of different receptors based on a gradient of strength. At least 24 variations are known and all carry the consensus sequence [KRHQSA]-[DENQ]-E-L. The more recent discovered third variation of the human ERD2 (ERD23) recognizes motif variations that do not follow the previous mentioned consensus, an unexpected result that opens even more possibilities to work and study those genes, both in plants and in other kingdoms (Alanen et al. 2011; Raykhel et al. 2007; Hulo et al. 2006). These results have been re-evaluated using a different organism, in that case *Saccharomyces cerevisiae*. The research data demonstrate in a comprehensive analysis of ERS (ER retention sequence) that the amino acid

residues in the position -5 and -6 of the C-terminal sequence seem to be important in determining the retention of ER luminal proteins (Mei et al. 2017).

A.6.3 Discovery of ERD2 plant homologues

The first plant homologue was identified serendipitously, via random cDNA library sequencing of expressed sequence tags (Lee et al. 1993). The protein was reported to complement the yeast mutant, and C-terminal fluorescent fusion labelled the Golgi and the ER in tobacco leaves (Boevink et al. 1998).Comparison of eight variants of the KDEL motif looking for functionality divergence in plant cells revealed that KDEL, HDEL and RDEL sequence work equally well in plants and share a similar epitope recognized by a monoclonal antibody (Denecke et al. 1992; Hadlington & Denecke 2000).

In contrast to yeast and mammals, which contain only a small variety of ERD2 proteins, there are seven different ERD2-related genes in plants. Those are divided into two different classes and the genes included in the class I (ERD2a and ERD2b – Figure 2), are more closely related to the yeast and mammalian genes. The other 5 genes resemble ERD2 but contain N-terminal extensions (Hadlington & Denecke 2000). Functional studies of plant ERD2 were based on similar C-terminal ERD2 fusion proteins to the one used for its localisation (Montesinos et al. 2014; Xu & Liu 2012; Xu et al. 2012; Li et al. 2009).

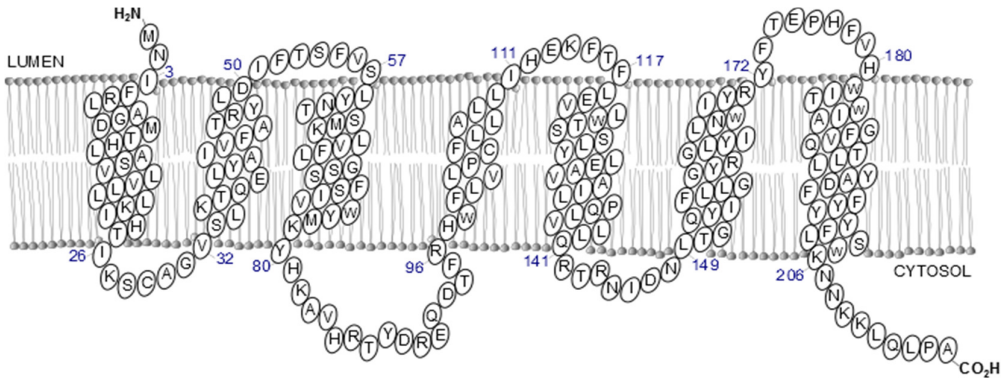


Figure 2 Proposed topology of the protein encoded by one of the two ERD2 genes from *Arabidopsis thaliana* (ERD2b).

ERD2 gene product is proposed to be a seven transmembrane protein with the N-terminus facing the lumen and the C-terminus facing the cytosol. A schematic representation based on the prediction created by JPred4 and previous models from the literature for *Homo sapiens* ERD2 (Townsley et al. 1993).

Following the same model of ERD2 genes from yeast and mammalian, recent results showed that the transport of proteins bearing the K/HDEL motif in plants could be remarkably similar, where an anterograde transport mediated via COPII vesicles and a retrograde mediated by COPI vesicles act together. It has being shown that in mammals there are two different populations of COPI vesicles and they have different functionalities, but for plants, as stated by Gao et al. (2014), the mechanisms that rule how the possible different populations of COPI vesicles work are still unclear.

The role of the ERD2 encoded protein in the general context of the organism is still poorly understood and demands further investigation. New research has recently suggested that the ERD2a and ERD2b of *N. benthamiana* participate in programmed cell death and the hypersensitive response (Xu et al. 2012).

ERD2 function has also been proposed to be linked to G proteins, both in mammals and plants, influencing forward transport from the Golgi to the plasma membrane (Giannotta et al. 2012; Wang et al. 2018)

It is interestingly to notice that ERD2 copies belonging to class I, ERD2a and ERD2b, diverge more within one plant species than when compared against different species. This divergence is exemplified in Figure 3 and it could indicate that an early duplication event took place, but the reasons for that are still unclear.

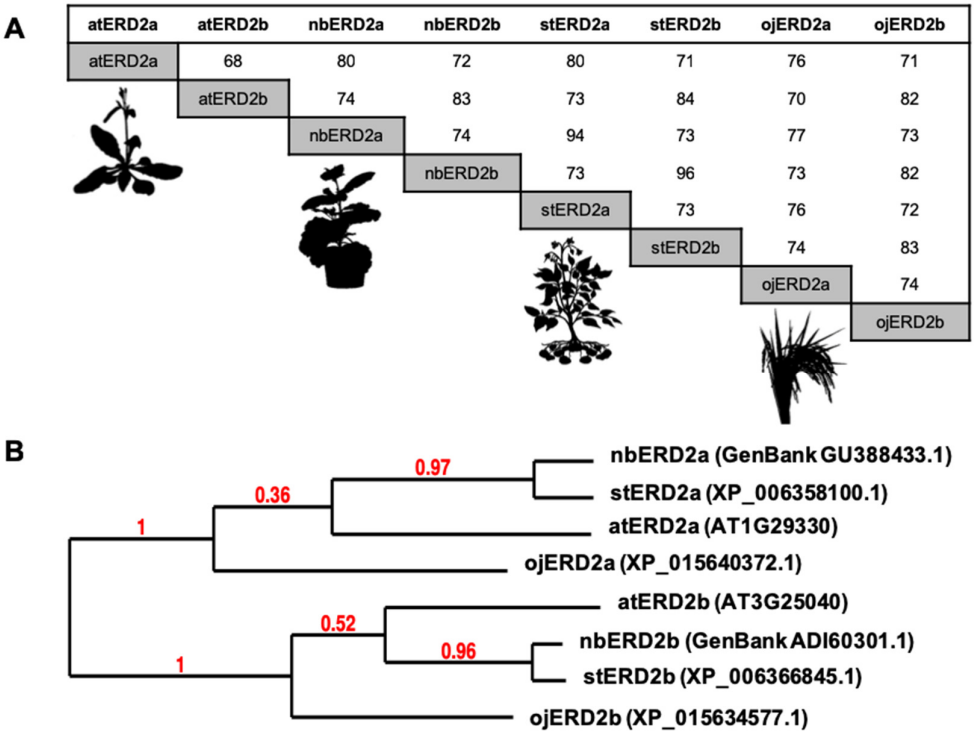


Figure 3 ERD2a and ERD2b gene products diverge more in the same specie than amongst different plants. Example of how the homology level of two ERD2 proteins can vary in within one species.

The comparison between ERD2a and ERD2b from four different plant species revealed that ERD2a is more similar to ERD2s from other plants than to ERD2b from the same species. A) Direct comparison of amino acids identity level of ERD2a and ERD2b from *Arabidopsis thaliana* (atERD2), *Nicotiana benthamiana* (nbERD2), *Solanum tuberosum* (stERD2) and *Oryza sativa* Japonica (ojERD2). B) Phylogenetic three constructed using the ERD2a and ERD2b from four species shows a clear split into two defined clades of ERD2 isoforms. Analysis via *Phylogeny.fr* platform (Dereeper et al. 2008)

A.6.4 Functional, topological and localization studies of the receptor

In mammalian cells, the ERD2 gene product is mostly localised to the Golgi apparatus (Lewis & Pelham 1990) and causes retrieval of soluble ER proteins from the Golgi (Pelham et al. 1988). Although not described in detail, it is assumed that ERD2 function involves four fundamental steps: 1) ligand-binding in the Golgi; 2) retrograde transport via COPI vesicles; 3) ligand-release in the ER and 4) anterograde transport back to the Golgi.

The complete structure and functional analysis of the corresponding ERD2 homologue in mammalian cells was carried out by systematic point mutagenesis, followed by *in vitro* KDEL-peptide binding experiments and fluorescence

microscopy (Townsley et al. 1993). This revealed a number of key-residues in the primary sequence that were important in ERD2 activity. Interestingly, neither cell viability nor the KDEL-mediated retention efficiency *in vivo* were used as bio-assays to study the mutations. Ligand-binding was scored by a KDEL peptide-binding assay using total microsomal membranes from cells over-expressing specific ERD2 mutants. Since the experiments were not carried out with pure components, direct binding of the peptide to the ERD2 gene product was not strictly proven, as an intermediate adaptor could be present and not ruled out. In addition, the investigators also scored the re-distribution of ERD2 from the Golgi to the ER by KDEL-overdose and interpreted any change as evidence for a transport defect. The assays also could not score defective ligand-release, nor could they distinguish between impaired ER export or increased Golgi-recycling. A second study (Janson et al. 1998) has also inquired the *in vitro* binding activity of the receptor, this time using synthetic cellulose-bound overlapping peptides of ERD2 and tested the interaction with two ER calcium binding proteins (CaBPs). As a control they demonstrated the saturation with KDEL containing peptides inhibits the interaction with CaBPs. Researchers concluded that a specific N-terminal sequence (^{22}KIWK^{25}) at ERD2 was responsible for the positive interactions and that phosphorylation of substrates did not influence the results.

Janson and colaborators (1998) identified two putative binding regions in ERD2 core, the strong ^{22}KIWK^{25} sequence mentioned before and a weaker region between amino acids 43 and 57. Their results are in contrast to a previous publication from Pelham's group where they claim to have identified ERD2 binding pocket via site-directed sulfhydryl labelling allowing for the spatial analysis of the receptor (Scheel & Pelham 1998). In their paper they have shown that the removal of four inconvenient endogenous cysteines does not influence the expression of the receptor and it allows for the replacement of systematically chosen amino acids for cysteines. It was demonstrated that ERD2 topology is resembling a G-coupled receptor and that four amino acids (Arg-5, Asp-50, Tyr-162, and Asn-165) are important for reaction with thiol-specific reagents when compared side-by-side to the cysteine less version and could possibly form a hydrophilic binding pocket.

More recently the binding region of the receptor was once again studied and a third hypothesis was presented stating that the high amount of aromatic residues

in the luminal loops of the yeast ERD2 could direct interact with the aromatic residues of HDEL and mediate retention (Mei et al. 2017)

Hsu et al. (1992) demonstrated that a Brefeldin-like effect can be obtained by overexpressing an ERD2-like protein (ELP-1) in different cell types, leading to a redistribution of Golgi membranes to form an ER-Golgi super-compartment. As proposed by the authors the overexpression of this gene could be inducing a hyper-retrograde transport. Not least the effects of the ERD2 deletion isolated by Semenza et al. (1990) proved that this gene is essential for growth, suggesting a secondary function. However, Hardwick et al. (1992) showed later that in the absence of the ERD2 gene in yeast a different family of genes called SED (Suppressors of the *erd2*-deletion) are responsible to allow the growth of the cells. Townsley et al. (1994) also support the previous idea, through point mutation analysis of ERD2 they show that viability and cell growth are strictly related to the capacity of ligand-binding and recycling of the ERD2.

Formation of COPI vesicles, leading to the initiation of retrograde transport could be influenced/induced by the accumulation of cargo and/or receptor molecules at the *cis* region of the Golgi. Townsley et al. (1993) were the first ones to raise the possibility of ligand-induced or oligomerization of ERD2, triggering its recovery from Golgi to the ER. Majoul et al. (2001) proposed a model where the free KDEL-receptor is found in a monomeric state whereas ligand binding triggers oligomerization and induces the formation of the COPI vesicle in live cells.

As mentioned in previous sections, the interaction of membrane proteins with COPI and COPII vesicles has been shown to be mediated by the presence of sorting motifs. Stornaiuolo et al. (2003) raised the possibility that the interaction of KDEL-KDELr-COPI occurs with higher affinity than between the KKXX motif and COPI. ERD2 has a KKXX motif-like signal; however, this motif is not typically positioned, since it is not the last four amino acids of the protein, and it is not highly conserved between species. The KKXX motif is important to induce the retention of Type I membrane proteins in the ER, but it should not influence the retention of a membrane spanning-protein in the Golgi (Vincent et al. 1998).

A new model was introduced in the recent years (Cancino et al. 2014; Pulvirenti et al. 2008) which involves a cascade system triggered by the KDEL-receptor

whereby, upon binding to the ligands, it would induce vesicle traffic and consequently maintain the balance of membranes of the two organelles by the interaction of ERD2, G-proteins and Protein kinase A (PKA).

As pointed before, and it is true until today, one of the biggest mysteries around the life-cycle of ERD2 is related to how the receptor maintains its Golgi localization (Pfeffer 2007). ERD2 is thought to shuttle between the ER and the Golgi, but how exactly the receptor gets integrated into COPI/COPII vesicles is unclear. Perhaps a KKXX-like motif could be due to be discovered. Another intriguing question is how does ERD2 not progress through the Golgi cisternae? A DXE-like signal, which induces the retention in the Golgi, would provide an explanation, but again such signal is still to be discovered.

Neither of those classical signals used by other membrane spanning proteins seems to be evidently present in the sequence of ERD2, independently of the organism. Interesting is also the fact that it has been shown before that the addition of a KKXX motif to the C-terminus of ERD2 can completely redistribute the receptor to the ER without affecting its *in vitro* ligand-binding properties (Townsley et al. 1993).

Another possible theory that could help explain how ERD2 is targeted and retained in the Golgi is based on previously studies which show that the physical-chemical characteristic of membrane spanning-proteins could control its localization (Ronchi et al. 2008), but in matter of fact although a nice concept it has not been used in the past to explain ERD2 Golgi residency.

Regarding the topology of the receptor it is believed that the ERD2 gene encodes for an integral membrane protein with seven transmembrane domains, similar to rhodopsin, which is essential for growth and to regulate the retention capacity of molecules accordantly to its expression levels (Semenza et al. 1990). ERD2 appears to show similarities with the G-protein-coupled-receptor (GPCR) family as it has 7 predicted transmembrane domains, a cytosolic C-terminus and its steady state distribution is reported to shift upon ligand binding (Lewis & Pelham 1992a; Capitani & Sallese 2009; Griffiths et al. 1994). An alternative model of six transmembrane domains has been proposed, where both N-terminus and C-terminus are exposed to the cytosol. The initial proposition for that alternative conformation was based on the strategy of tagging ERD2 at its N-terminus with

a myc-tag and the methodical addition of a glycosylatable reporter at putative loops and subsequent analysis of protein size-shifts via western blot. If the reporter was exposed to lumenal size a shift was expected due to the glycosylation of the molecule (Singh et al. 1993). This topological conformation of the receptor was later supported via a redox-based analysis using redox-sensitive GFP fusions (Brach et al. 2009).

As it can be seen, although ERD2 was discovered nearly three decades ago, very little is known about the exact steps in its transport cycle. Signals controlling ERD2 transport between the ER and the Golgi, as well as mechanisms that prevent post Golgi trafficking of ERD2 remain elusive to date. Albeit different studies claiming to have identified the region of the receptor controlling ligand-interaction there is no consensus and the results are in disagreement with each other. Finally, the topology of the receptor is still under debate and the field lack a study demonstrating the crystal structure of the protein.

A.7 **Open questions and research objectives**

Even though the ERD2 gene product was identified more than 25 years ago, there are still many open questions regarding its biological function (Pfeffer 2007). Although the recycling of KKXX proteins from the Golgi apparatus was shown to involve direct binding to COPI coat proteins (Cosson & Letourneur 1994), it is yet unclear if ERD2 uses COPI-mediated transport to mediate its function (Hsu et al. 1992). There is currently no information about the signals that control retrograde Golgi to ER transport of ERD2 and anterograde recycling from the ER back to the Golgi apparatus.

The ERD2 gene product does not carry a typical ER export signal, such as an FF motif, and a DXE motif found in the second cytosolic loop of *Arabidopsis thaliana* ERD2 is not conserved across other kingdoms (Cole et al. 2008). Its potential role in mediating active export via the interaction with sec24A-sec24B of COPII coated vesicles (Otsu et al. 2013) remains to be shown. Cabrera et al. (2003) have shown that ERD2 C-terminus does not promote Golgi localisation when fused to a plasma membrane protein. However, in the same work it is proposed that this region of the receptor apparently plays a very important role promoting interaction with coatomer and ARFGAP by the phosphorylation of a C-terminal serine (Cabrera et al. 2003). Even less is known about the mechanism

by which ERD2 accumulates in the Golgi without leaking to post Golgi compartments (Pfeffer 2007). Different models have been proposed such as the protein aggregation/Kin-recognition, transmembrane domain-mediated retention and Golgi membrane lipid composition-based retention (for review consult Banfield 2011).

Different tagging strategies have been used in the past to generate a fluorescent ERD2 molecule, ranging from N-terminal or C-terminal tagging (Boevink et al. 1998) to intra-molecular tagging at the1st predicted lumenal loop (Li et al. 2009). The latter revealed a Golgi-only localisation (Li et al. 2009) whilst by far the most frequently used C-terminal fluorescent fusions show a dual Golgi-ER localisation (Montesinos et al. 2014). In sharp contrast, N-terminal fluorescent ERD2 fusions were detected only in the ER (An 2015). Using an *in vivo* assay, work from a previous PhD dissertation showed that neither C-terminal not N-terminal ERD2 fusions were biologically active (An 2015). Essentially, this finding suggests that published subcellular localisations of the ERD2 may not represent biologically meaningful constructs.

It is still unclear if ERD2 directly binds to cargo molecules and if there is signal specificity amongst different organisms. Previous binding studies (Townsley et al. 1993; Scheel et al. 1997) have not been carried out with purified components and for this reason, the ligand-binding and release conditions are far from established. In addition it is still unclear if the signal recognition system differentiates between related signals as shown in yeasts (Lewis & Pelham 1992a), or if signal specificity is much less discriminative as suggested in plants (Denecke et al. 1992).

The mechanism for receptor recycling between two compartments cannot be studied without considering the requirement for ligand-binding and ligand release. The recycling model proposed for the Man-6-P receptor has led to the popular belief that the ligand-release of all protein sorting receptors are influenced by the pH in the lumen of the organelle they occupy. However, potential changes in pH between the ER lumen and the cis-Golgi may be too small to mediate high fidelity ligand binding and release. In addition*, in vitro* pH binding curve of ERD2 reported in earlier studies (Wilson et al. 1993) revealed the highest ligand-binding affinity at a pH close to five, which does not correlate to the higher pH measured in the lumen of plant Golgi cisternae (Martinière et al.

2013). The exact mechanism for conditional ligand binding in vivo remains to be elucidated in a eukaryotic system.

One of the key-problems with most of the previous research approaches is that the more detailed studies of ERD2 function (i.e. Townsley et al. 1993) were not assessed directly by quantifying ligand transport in living cells. The main objectives of this thesis were thus built on a recently established ERD2 bioassay, which monitors its *in vivo* activity in reducing the secretion of HDEL-proteins (An 2015), and develop this assay further, to study how the receptor is targeted in the ER-Golgi system using biologically functional fusions. I next interrogated which sequence motifs are critical to ERD2 function and tested how COPI-mediated recycling affects ERD2 functionality. I also compared ERD2 with the closely related ERP gene family and finally, studied how conserved the receptor is amongst eukaryotes.

B **Results**

Chapter 1 **Establishing quantitative assays to monitor ERD2 activity for functional analysis in plants**

1.1 **Introduction**

Recent results in the host laboratory suggested that some of the generally accepted findings surrounding the functioning of ERD2 may have to be re-evaluated. These include the apparent inactivity of a widely used fusion protein in which the yellow fluorescent protein (YFP) or spectral variants were fused to the C-terminus of ERD2 (An 2015). If the chimeric protein is not biologically active, neither the localisation nor the specific interaction studies conducted may be reliable. In the same work, mutagenesis of the ERD2 C-terminus revealed a crucial role of the C-terminus of ERD2 which was also deemed unimportant for ERD2 function in earlier studies (Townsley et al., 1994). Since experiments were based on a gain-of-function activity of ectopically expressed ERD2 in reducing the partial secretion of AmyHDEL using *Arabidopsis thaliana* ERD2 in *Nicotiana tabacum* protoplasts, the scope for widening the experimental tools was substantial at the beginning of this project.

Nicotiana tabacum has been well established as a model plant to be used for quantitative protein transport studies in protoplasts (Denecke et al. 1989; Denecke & Vitale 1995). In addition, high quality live bio-imaging via confocal laser-scanning microscopy is often carried out in tobacco leaf epidermis cells after infiltration with Agrobacterium strains harbouring fluorescent constructs (Denecke et al. 2012). Nonetheless, the fact that its genome is tetraploid, its genome sequence is incomplete and not readily available for public research and the lack of mutant resources has stopped the plant from being a model as universal as *Arabidopsis thaliana*. In contrast, a very similar plant, *Nicotiana benthamiana*, has been increasingly used in recent years, it is diploid, and its genome has been fully sequenced (Goodin et al. 2008; Nakasugi et al. 2013). Since these features would permit gene targeting via homologous recombination and gene knockdowns by interfering RNA or antisense transcript inhibition, it could potentially be an attractive model plant for secretory pathway research.

In this Chapter 1 intended to explore if *N. benthamiana* may be used as an alternative model system to study ERD2 in both homologous and heterologous expression approaches and verify earlier gain-of-function results by complementing partial loss of function. The main aim was to combine the potential for genetic approaches with functional experiments without risking a loss of quantification.

1.2 Results

1.2.1 *Nicotiana benthamiana* leaves provide an excellent source of protoplasts with high recombinant protein expression.

To explore the potential of *N. benthamiana* for common working practices to studying protein trafficking in the plant secretory pathway, I directly compared it with *N. tabacum* side-by-side with regards to their protoplast production capacities and their transient expression capacity for recombinant proteins.

Figure 4A exemplifies an *in-vitro N. tabacum* plant and a protoplast sample visualised in bright-field microscopy. Figure 4B shows the same but for *N. benthamiana*. At first it became evident to me that *N. benthamiana* is a faster growing plant and it can be grown in a smaller tissue culture jar due to its reduced size. Although the leaves are smaller and slightly more folded and thus harder to handle via needle-bed perforation, first experiments quickly revealed that the leaves digested faster and with more diluted enzyme mixes compared to *N. tabacum*. As a consequence, the yield of protoplasts was also higher.

For the initial screen accessing the proteins synthesis of each species I decided to use the well stablished β-glucuronidase (GUS) marker due to its high stability and easiness to be measured, and because it has been previously shown to be an efficient expression marker (Gershlick et al. 2014). For that purpose, an expression vector which contains GUS coding region under the control of the TR2' mannopine synthase promoter of *Agrobacterium tumefaciens* (TR2) was used (plasmid pIKA9 - Adam 2013).

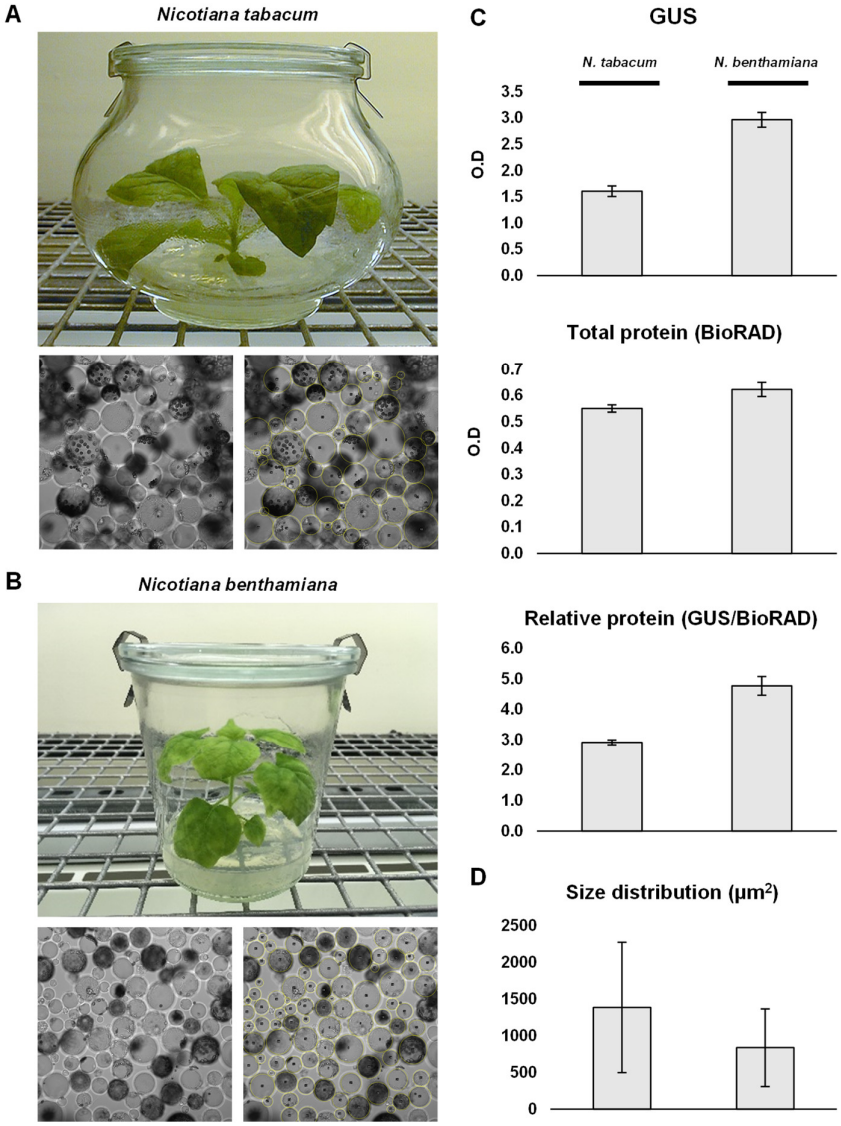


Figure 4 Comparison of protoplasts size and proteins synthesis from two plant species, *Nicotiana benthamiana* and *Nicotiana tabacum*.

Plant sterile tissue culture and brightfield microscopy picture of protoplast exemplifying size distribution from A) *Nicotiana tabacum* and B) *Nicotiana benthamiana*. C) Protoplast suspension of both species were transformed via electroporation with 50 µg of a GUS containing plasmid and incubated for 24 hours. Error bars are standard deviations of three independent transfections. Upper panel shows total GUS values, measured for whole cell suspensions by colorimetric assay and expressed as optical density (O.D). Second panel shows total protein values, measured via BioRAD assay for whole cell suspension and expressed as O.D. Bottom panel shows the relative protein content obtained by the ratio of GUS and BioRad measurement and expressed as arbitrary values. D) Protoplast suspension from both species were imaged in brightfield microscopy and posteriorly analysed in ImageJ software. At least 5 images from independent cultures per plant were screened until a total of 250 protoplast were measured. Total protoplast surface area was measured in µm, and only protoplast in focus were manually selected for measurement.

Protoplast suspensions from both species were used to count protoplasts and adjusted to contain equal cell densities, to be transfected with the same amount of plasmid DNA. Figure 4C shows a direct comparison illustrating that *N. benthamiana* protoplasts can produce almost twice as much GUS enzyme, whilst the measurement of the total protein content confirms that suspensions were adjusted adequately to yield comparable amounts of plant cells. Interestingly, *N. benthamiana* yielded smaller sized protoplasts compared to *N. tabacum* (Figure 4D). Protoplast size may influence the optimal electroporation conditions, and since the standard procedure was optimised for *N. tabacum*, it was possible that even higher transfection rates and thus recombinant protein yields could be expected if conditions were specifically optimised for *N. benthamiana*.

Since the first experiment was highly encouraging, I decided to invest in further experiments to establish a routine working protocol for *N. benthamiana* protoplasts, which forms the basis of this first results chapter.

1.2.2 Electroporation conditions for *N. benthamiana* require a lower voltage optimum compared to *N. tabacum* protoplasts

Having established that electroporated protoplasts from *N. benthamiana* can produce more GUS proteins than a comparable number of *N. tabacum* protoplasts, the next step was to establish the optimum electroporation conditions for *N. benthamiana*. I took advantage of the fact that plasmid pIKA9 is a dual expression plasmid harbouring two different reporter constructs, one for cytosolic expression and one for protein secretion via the secretory pathway. This would yield two independent measures to score transfection efficiency and transient expression performance and permit a solid comparison of the two species *N. benthamiana* and *N. tabacum*. The secreted Amylase reporter (Amy) is widely used for the purpose of studying protein secretion, ER retention and vacuolar transport (Phillipson et al. 2001; daSilva et al. 2005). GUS expression was driven by the TR2 promoter and Amy expression was expressed under the control of the CaMV 35S promoter. Both promoters control high levels of transcription in tobacco protoplasts (Denecke et al. 1990; Denecke et al. 1992) and were presumed to work efficiently in *N. benthamiana* as well.

To determine the optimal voltage for electroporation conditions, protoplast from each of the two species were prepared, adjusted for cell numbers and electroporated using the same concentration of pIKA9. Keeping the capacity of

the electric charge constant, the voltage was gradually increased forming a voltage-series ranging from 130mV to 220mV for both species. After a standard incubation of 24 hours, two identical samples of the protoplast suspensions were harvested and processed for either GUS or Amy analysis.

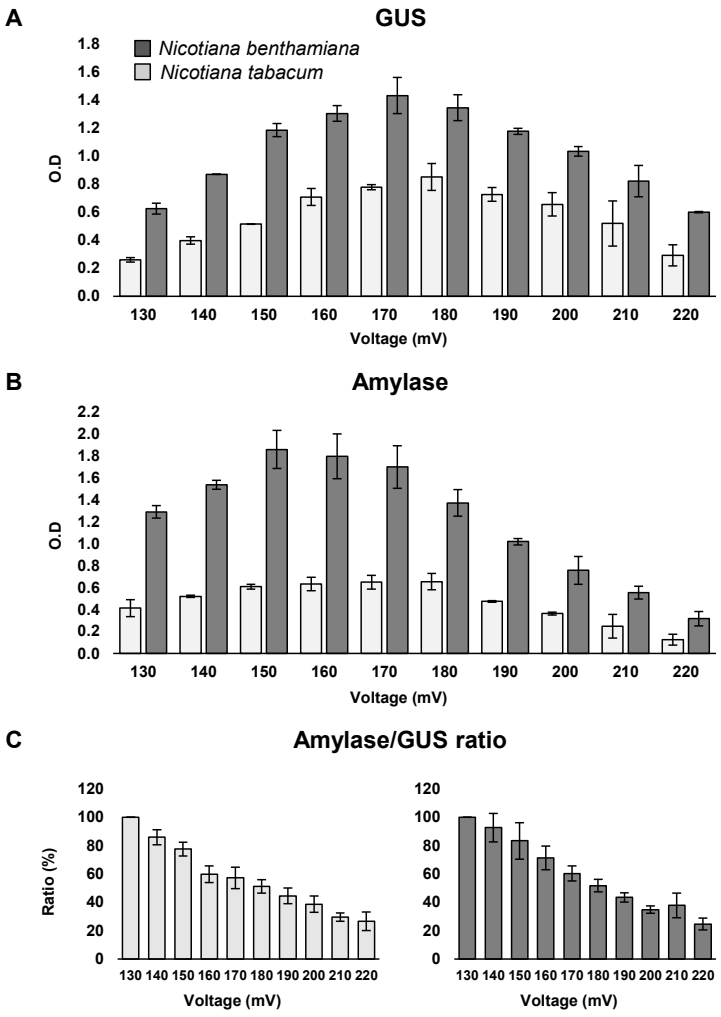


Figure 5 Optimum voltage to yield maximum synthesis of cytosolic and secretory molecules in transformed protoplast of *Nicotiana benthamiana* and *Nicotiana tabacum*.

Variable voltages ranging from 130 mV to 220mV were used to electroporate and transform protoplasts from *N. benthamiana* (dark grey) and *N. tabacum* (light grey) using a dual expression plasmid carrying GUS as a cytosolic marker and also Amylase as a secretory marker. Error bars are standard deviations of three independent transfections. A) Effect of voltage-series over GUS synthesis for both plant species plot together. 50 µg of cargo plasmid was electroporated and total GUS was measured for whole cell suspensions via colorimetric assay and expressed as O.D. B) Effect of voltage-series over Amylase synthesis for both plant species plot together. 50 µg of cargo plasmid was electroporated and total Amylase was measured for whole cell suspensions via colorimetric assay and expressed as O.D. C) Ratio between total Amylase O.D and total GUS O.D for each voltage step is shown in separate per plant species and colour coded as before.

Figure 5A shows that the GUS levels are consistently higher for *N. benthamiana* protoplasts when compared to *N. tabacum.* This difference was independent of the voltage used. The optimum voltage was remarkably similar for both species but appeared to be slightly lower for *N. benthamiana* (170 Volts) compared to *N. tabacum* (180 Volts). This was surprising as *N. benthamiana* protoplasts were seen to be smaller than those of *N. tabacum* (Figure 4D).

1.2.3 *N. benthamiana* protoplasts express particularly high levels of secretory proteins.

Measurement of the second reporter Amy was initially expected to yield a similar result, as it was encoded by the same plasmid as GUS and thus expected to be transfected with similar efficiency. Representative cell suspension samples from the same experiment as in Figure 5A were thus extracted using a different buffer and used to measure Amy activity.

Interestingly, Figure 5B shows that the optimum voltage for electroporation of both *N. benthamiana* and *N. tabacum* protoplasts was shifted towards lower voltages in both species when Amy was measured. The preference for lower voltage is clearly demonstrated by plotting the Amy-GUS ratio in function of voltage for each species (Figure 5C). In both cases, higher voltages appear to favour GUS over Amy, despite being encoded by the same transfected plasmid.

Another unexpected result was the fact that *N. benthamiana* protoplasts expressed much higher levels of Amy compared to *N. tabacum* protoplasts and that the difference, an approximate 3-fold increase (Figure 5B), was greater than for GUS which was only around 2-fold (Figure 5A). This was particularly obvious at low voltages.

In summary, *N. benthamiana* protoplasts performed better on all fronts compared to *N. tabacum* protoplasts and this was particularly true for the secretory marker. *N. benthamiana* could therefore be an ideal model system for the analysis of protein transport and receptor function, using both biochemical as well as genetic experiments.

1.2.4 *N. benthamiana* protoplasts show faster rates of protein secretion compared to *N. tabacum* protoplast.

Results so far suggest that *Nicotiana benthamiana* protoplasts appear to exhibit a higher capacity for recombinant protein production.

This was obvious for the cytosolic reporter GUS and in particular for the secreted reporter protein Amy (Figure 5). It is unlikely that this increase is simply due to a higher transfection efficiency leading to a higher percentage of transfected cells relative to the total number of cells, because otherwise the increase in *N. benthamiana* relative to *N. tabacum* would have been the same for the two reporters. To test if other physiological properties of *N. benthamiana* protoplasts show further differences compared to those of *N. tabacum*, I compared secreted Amy with engineered ER retained HDEL-tagged Amy (AmyHDEL), two proteins with very distinct transport properties. The latter was shown to first accumulate in the cells, until a threshold is reached after which secretion gradually increases (Phillipson et al. 2001).

Plasmids encoding either Amy or AmyHDEL were transfected in either N. benthamiana protoplasts or N. tabacum protoplasts via electroporation (160mV/1000µF) and incubated for 24 hours before harvesting. Rather than simply measuring the total O.D. for the amylase activity in the suspension as before (Figure 5), cells and medium were separated, and enzyme activity was measured in the two fractions to determine the “secretion index” (SI), the ratio between extracellular and intracellular fractions.

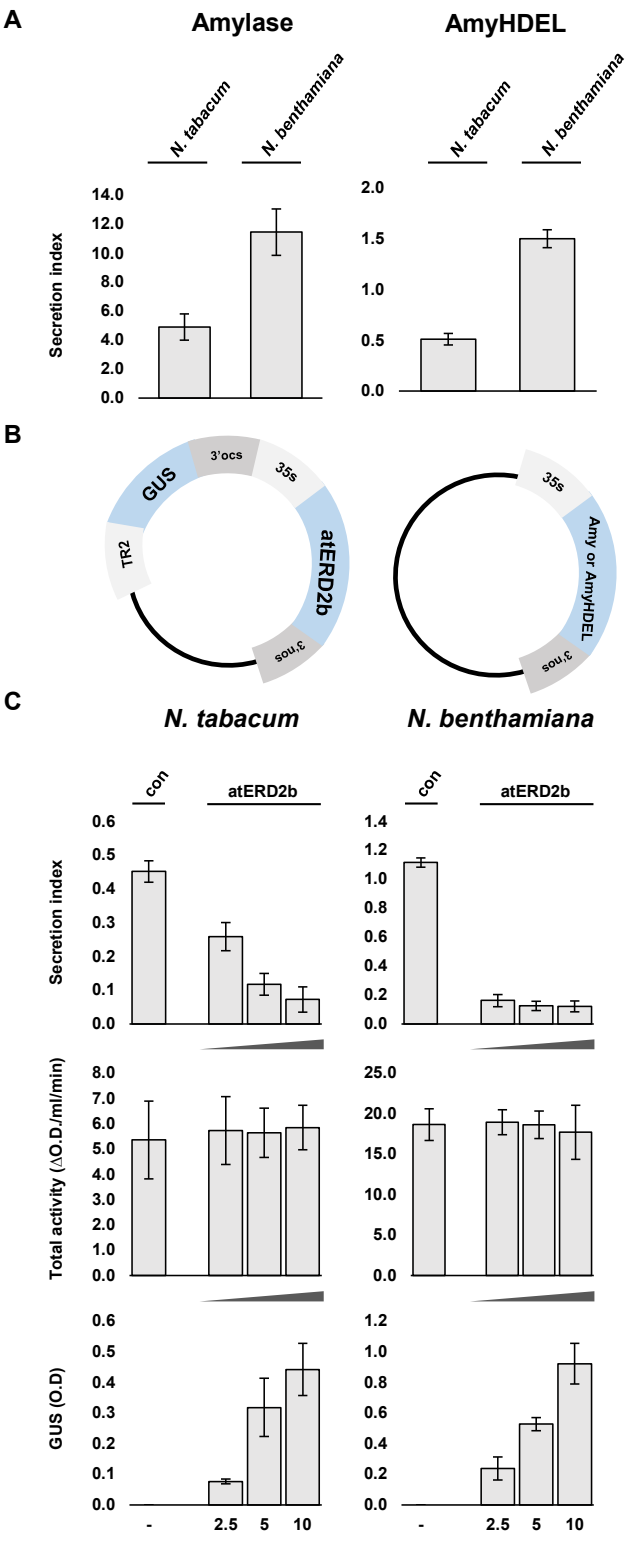


Figure 6 *N. benthamiana* higher performance of both the expression of two independent secretory cargo molecules and the receptor mediated retention of HDEL-ligands.

A) Transient expression experiment with both *N. tabacum* and *N. benthamiana* protoplasts co-expressing either Amylase or AmyHDEL molecules. 50 µg of cargo plasmid was electroporated. B) Schematic of dual expression system used for the assay based on the pGUSref plasmid (Gershlick et al., 2014) allowing normalisation of the transfection efficiency by the colorimetric GUS assay and the cargo plasmids C) Transient expression experiment with both *N. tabacum* and *N. benthamiana* protoplasts co-expressing AmyHDEL alone or combined with a atERD2b. 50 µg of cargo plasmid was electroporated together with increasing amounts of effector plasmid. Upper panel shows the secretion index obtained by the ration of medium and cell activities. Middle panel shows the total α-amylase activity obtained in each cell suspension given in arbitrary relative units (ΔO.D./ml/min). Bottom panel shows total GUS O.D used as internal control marker for transfection efficiency. Protoplast suspension of both plant species were adjusted by volume to be directly compared. Error bars are standard deviation of three independent transfections.

Figure 6A confirms that regardless of the species, Amy is secreted much faster than AmyHDEL, as expected (Phillipson et al. 2001). However, the data also show that SI values for Amy and AmyHDEL are significantly higher in *N. benthamiana* protoplast compared to *N. tabacum* protoplasts. The difference is slightly more noticeable for AmyHDEL which shows an almost 3-fold increase in the secretion index.

The results confirm that differences in protein production are not simply due to a higher percentage of transfected cells. Instead, *N. benthamiana* protoplasts synthesize recombinant proteins faster, and also appear to secrete faster than protoplasts from *N. tabacum*.

1.2.5 Ectopic expression of ERD2 drastically reduces secretion of HDEL proteins in *N. benthamiana*.

The fact that *N. benthamiana* protoplasts show a higher SI value for the AmyHDEL cargo could be explained by further saturation of the endogenous K/HDEL receptor ERD2, due to higher quantities of the cargo molecule AmyHDEL. A stronger ERD2 loss-of-function phenotype in this protoplast model could thus be highly suitable for an earlier established ERD2 bio-assay (An 2015). The assay was based on measuring the reduction of the AmyHDEL SI upon ectopic expression of *Arabidopsis thaliana* ERD2a or ERD2b in *N. tabacum* protoplasts (Silva-Alvim et al. 2018). To test if Arabidopsis ERD2 also functions in *N. benthamiana* protoplasts, a dose-response assay was performed in which constant levels of AmyHDEL plasmid were co-transfected with increasing quantities of Arabidopsis ERD2b plasmid. The two Nicotiana species were compared, and control conditions (con) without ERD2 plasmids provided the base-line for the dose-response assay. In addition, the ERD2 plasmid contained the internal marker GUS to ensure experimental conditions for a fair comparison. Cargo and effector plasmids used are schematically represented in Figure 6B.

Figure 6C shows a side by side comparison of the two plant species in the same experimental conditions. The results demonstrate the advantages of the *N. benthamiana* protoplast system. As observed in figure 6A, the SI value for AmyHDEL alone is much higher to start with in *N. benthamiana* protoplasts. Interestingly, even the lowest concentration of ERD2 plasmid yielded an almost maximum reduction in the SI. In comparison to the *N. tabacum* protoplasts, the ERD2 dose-response is much steeper. Total cargo activity is more than 3-fold

higher, and GUS levels from the ERD2 plasmid are approximately 2-fold higher in *N. benthamiana*. However, the lowest quantity of ERD2 plasmid (2.5µg) in *N. benthamiana* yielded the same reduction in the SI value as the highest quantity in *N. tabacum* (10µg).

Although the dose-response assay is an indirect measure for ERD2 activity, the results indicate that for multiple membrane-spanning proteins such as ERD2 the increased synthesis in *N. benthamiana* compared to *N. tabacum* could possibly be even higher than for GUS and Amy. Therefore, *N. benthamiana* provides a highly sensitive model system to monitor ERD2 activity. To obtain more gradual dose-response curves, ERD2 plasmids may have to be diluted further compared to *N. tabacum*.

1.2.6 Complementation of ERD2 function between *Arabidopsis thaliana* and *Nicotiana benthamiana*.

Results from the previous experiments depended on heterologous overexpression of *Arabidopsis thaliana* ERD2 in two separate Nicotiana species. Since *Nicotiana benthamiana* genome has been sequenced, homologous gene overexpression and knock-down was feasible. As a first step, a hybrid gene comprising the first half of *Nicotiana benthamiana* ERD2a (NbERD2a) followed by the second half of *Nicotiana benthamiana* ERD2b (NbERD2b) was designed and constructed. Figure 7A illustrates the exact position of the fusion point in a conserved region, resulting in a hybrid that shares the general overall structure of the two parent ERD2 genes.

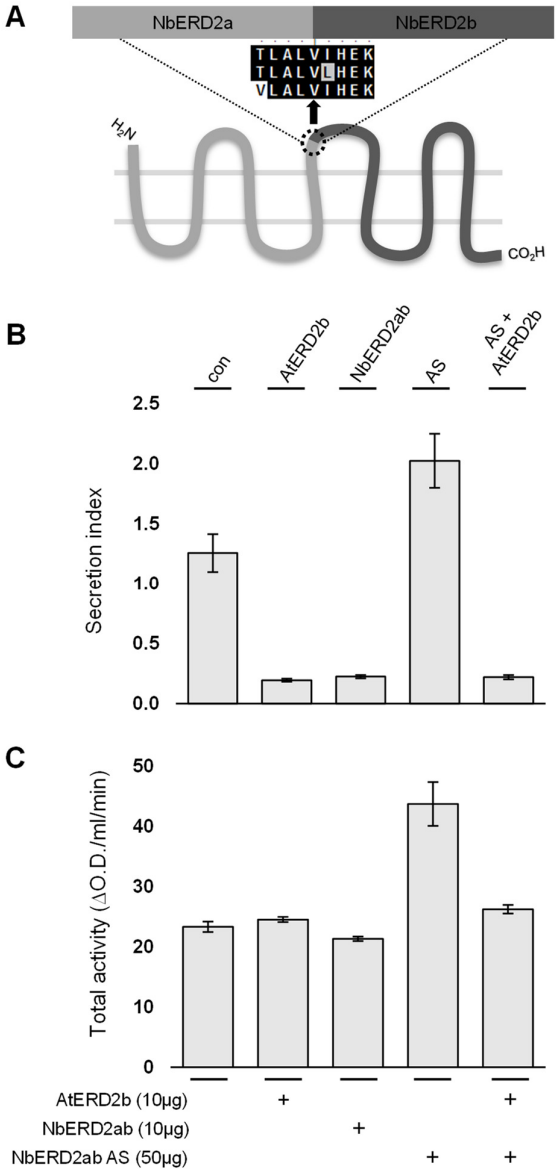


Figure 7 Evaluation of evolutionary conservation of ERD2 genes in *Arabidopsis thaliana* and *Nicotiana benthamiana.*

A) Illustration of the hybrid ERD2 transcript (NbERD2ab) which was generated as sense and as anti-sense constructs. The alignment shows the point where the fusion was made to generate a hybrid ERD2 coding region. B) Transient expression experiment with *Nicotiana benthamiana* protoplasts co-expressing AmyHDEL with either AtERD2b, sense NbERD2ab, antisense NbERD2ab (AS) or the combination of AS with AtERD2b and incubated for 48 hours to allow degradation of endogenous ERD2. 50µg of cargo plasmid was electroporated alone or co-electroporated together with sense or antisense ERD2 plasmids as indicated by "+". Error bars are standard deviations of three independent transfections. C) Total α-amylase activity obtained in each cell suspension given in arbitrary relative units (ΔO.D./ml/min). *Figure adapted from Silva-Alvim et al. (2018)

To test if the hybrid NbERD2ab is functional, its coding region was inserted in the GUS reference plasmid for direct comparison with AtERD2b in the same type of vector. To test if endogenous ERD2 genes can be down-regulated by anti-sense inhibition (Ecker and Davis, 1986), the hybrid NbERD2ab coding region was inserted in the opposite orientation in the GUS reference plasmid. After normalisation via the internal marker GUS, and using standard incubation times and effector plasmid concentrations, the impact of the antisense construct was not noticeable, whilst both sense constructs led to an equally strong reduction in the secretion index comparted to the control with cargo alone (data not shown). This result illustrates the high levels of functional conservation between ERD2a

and ERD2b, as well as conservation between different plant species, but also suggests that endogenous ERD2 turnover is too slow to compromise ER retention defects after just 24 hours. For this reason, the protocol was modified to pool two electroporations, include a washing step after 24 hours and incubate for a further 24 hours in fresh TEX medium. This was to eliminate the secreted AmyHDEL that accumulated whilst endogenous ERD2 knockdown was in progress. In addition, the quantity of the anti-sense plasmid was increased 5-fold.

Figure 7B shows that AtERD2b and NbERD2ab led to the same reduction of the AmyHDEL secretion, suppressing the loss-of-function effect caused by the saturation of endogenous receptor. In contrast, the anti-sense NbERD2ab led to increased cargo secretion. In addition, the total activity of AmyHDEL was increased almost two-fold by the anti-sense co-expression, indicating that higher secretion was not caused by cell mortality (Figure 7C). Increased yield of AmyHDEL was unexpected, but highly reproducible, and will not be discussed here because it was further explored in results Chapter 5.

Differences in the codon usage between *A. thaliana* and *N. benthamiana* render the AtERD2b coding region relatively insensitive to the effect of the antisense transcript as its transcript will not hybridise efficiently. Therefore, a complementation assay was performed and the last lane of Figure 7B shows that co-expressed AtERD2b abolished the effect of the antisense transcript and reduced the secretion of AmyHDEL to the same level as AtERD2b co-expression without anti-sense transcript. The results show that the effect of the anti-sense was caused by a reduction in the level of ERD2, which was fully compensated by ERD2 from a different plant species. The result also illustrates that low plasmid concentrations encoding sense ERD2 suffice to mediate strong addition retention of AmyHDEL, beyond the capabilities of endogenous ERD2.

1.2.7 The weak pNOS promoter can be used to study ERD2 due to high efficiency of the receptor

Previous experiments demonstrated that cytosolic, secreted and membrane proteins are highly expressed in *N. benthamiana* and that ERD2 seems particularly well expressed. The high efficiency in *N. benthamiana* makes it difficult to carry out dose-response assays, as seen in Figure 6C (*N. benthamiana* SI in function of increasing ERD2 levels). Although the dual

expression effector plasmid can be diluted further, this will also decrease the co-transfection efficiency with the AmyHDEL plasmid, which would give a false impression of reduced ERD2 activity. In order to keep plasmid concentrations and thus the co-transfection efficiency higher, and yet achieve a dose-response assay in *N. benthamiana*, I decided to explore the use of a weaker promoter to drive ERD2 expression. For that purpose, the CaMV 35S promoter present in the dual expression vector was replaced by the weaker nopaline synthase gene promoter (pNOS) from *Agrobacterium tumefaciens*. The resulting new construct was then compared with the original construct in *N. tabacum* aiming to compare two dose-response curves with gradual changes.

Results from Figure 8A show that when ERD2 is expressed under the control of pNOS promoter, the secretion of AmyHDEL is reduced in a shallower dose-dependent manner, compared to the CaMV 35S promoter driven construct. The results show that the pNOS promoter is indeed weaker than the CaMV 35S promoter in *N. tabacum* protoplasts. For instance, at the maximum concentration of the pNOS plasmid, the same reduced SI value was observed as for 4-fold lower levels of the CaMV 35S promoter plasmid (Figure 8A, compare lane 4 and lane 9 in upper panel). The internal marker GUS shows a comparable transfection efficiency for both plasmids (Figure 8B).

Taken together, these results suggest that the pNOS promoter is a good candidate to be used when lower levels of ERD2 are needed, particularly when working with *N. benthamiana* and for the purpose of mutant screens to identify ERD2 variants with partial loss of function that would be difficult to see if ERD2 levels were saturating.

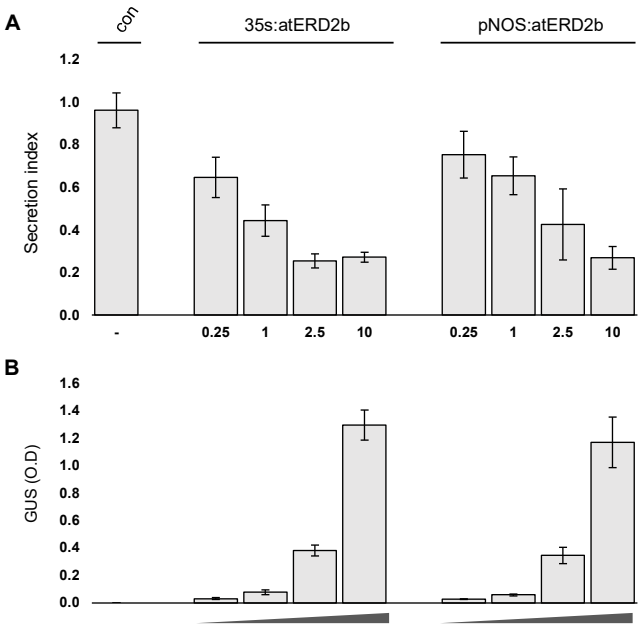


Figure 8 Effectiveness of different promoters whilst driving the expression of ERD2b in *Nicotiana tabacum* protoplast.

A) Transient expression experiment with *Nicotiana tabacum* protoplasts co-expressing AmyHDEL alone or with either CaMV35S driven atERD2b or pNOS driven atERD2b. Protoplast were incubated for 24 hours before harvesting. 50µg of cargo plasmid was electroporated alone or co-electroporated together with increasing amounts of effector plasmid as indicated below each lane in µg. Error bars are standard deviations of three independent transfections. B) Total GUS O.D used as internal control marker for transfection efficiency.

1.2.8 Identification of functional ERD2 mutants using a triple expression vector with three promoters

Taking together the outcomes from past experiments, and the fact that previous attempts to examine ERD2 function via site direct mutagenesis in the literature were done *in vitro*, I decided to analyse specific mutations of highly conserved amino acids using the bio-assay. Amino acids selected to be mutated are coloured and represented in Figure 9A, all have been replaced by alanine apart from leucine which have been replaced by glycine. Previous data from our laboratory (An 2015; Silva-Alvim et al. 2018) shows that the double replacement of C-terminally positioned leucine by glycine causes a complete knock-out of the receptor and for that reason this double mutant was used as negative control representing inactive ERD2.

To establish a reproducible comparison between mutants using a fast routine assay, a variant of the double-expression vectors (DV) was built to harbour three genes with three different promoters. The TR2’ promoter controls GUS expression to normalise transfection rates, the CaMV 35S promoter drives cargo AmyHDEL expression, and the pNOS promoter permits low expression levels of ERD2 and its variants. The newly created triple-expression vector (TV) is illustrated in Figure 9B.

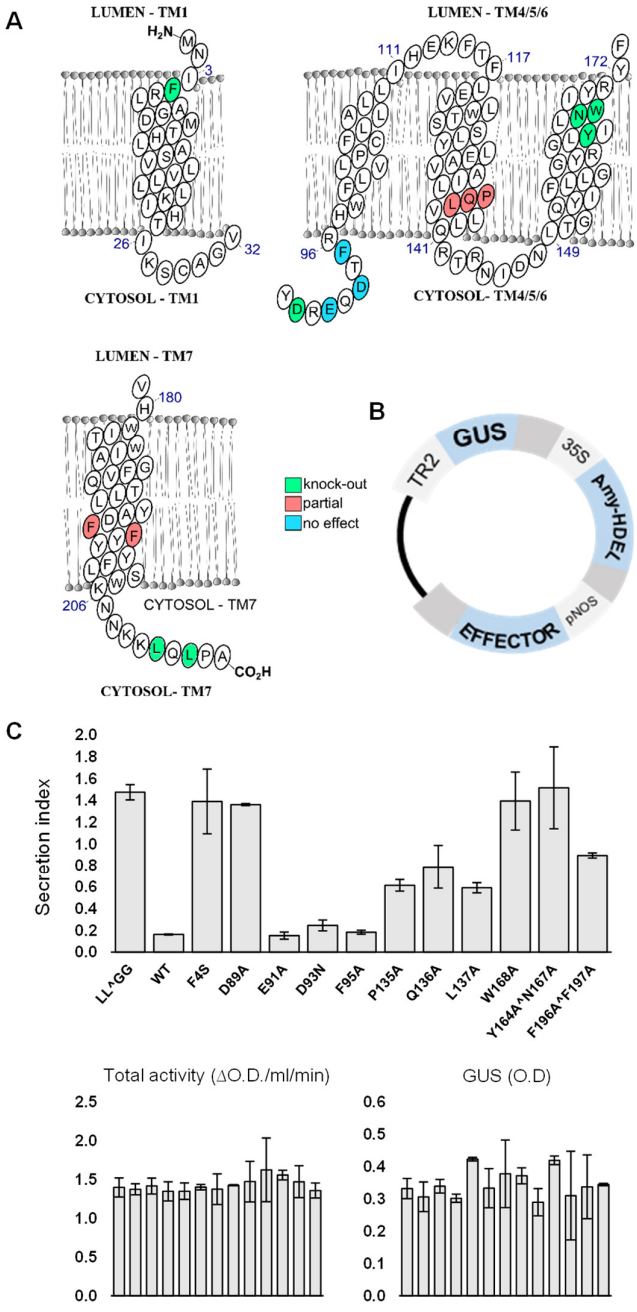


Figure 9 Mutagenesis analysis of ERD2 evaluating its biological activity using a triple expression vector system.

A) Illustration of amino acids selected for point mutagenesis analysis of atERD2b and the observed effects in the biological activity. B) Schematic of triple expression system used for the assay based on the pGUSref plasmid (Gershlick et al., 2014), but with the addition an extra gene cassette allowing for normalisation of the transfection efficiency and for equal cargo expression simultaneously C) Transient expression experiment with *Nicotiana benthamiana* protoplasts expressing triple vectors carrying ERD2 with desired mutations. Variable µg of effector plasmid was used and previously adjusted to yield equal GUS levels in pilot experiment. LL^GG mutation used as negative control based on pre-existent data (An Jing, 2015). Error bars are standard deviations of three independent transfections. Total α-amylase activity obtained in each cell suspension given in arbitrary relative units (ΔO.D./ml/min). Total GUS O.D used as internal control marker for transfection efficiency.

The results of the mutant screening using the TV are shown in Figure 9C which demonstrates that cargo levels as well as GUS levels were comparable. The generated mutants could be classified into three categories. Firstly, there are mutations that caused no effect (Figure 9A coloured in blue), secondly there are mutations causing a partial loss of function (Figure 9A, coloured in red) and lastly there are mutations causing a complete knock-out of the receptor-function (Figure 9A, coloured in green).

These findings provide insights for future research and for that reason the mutations leading to either partial or complete loss of function were selected to be further studied later in this thesis, aiming to find if there is a correlation between the localisation of the receptor and the lack of biological function

1.3 Discussion

The first chapter of this thesis compared side-by-side protein synthesis efficiency of two plant well know plant species. The first, *Nicotiana tabacum*, has been a classical model plant used in protein studies, but in contrast to *Nicotiana benthamiana* the genome sequence is not readily in the public domain and homologous gene expression is not straight-forward. The results revealed that *Nicotiana benthamiana* is an excellent candidate to be used in future studies because it exceeds the performance of *Nicotiana tabacum* protoplasts and yet offers the prospect of including genetic approaches to investigate protein synthesis in the secretory pathway. In addition, a number of unexpected results regarding the performance of secretory versus cytosolic expression were made which are worth discussing in more detail. Finally, a feasibility study to use *Nicotiana benthamiana* protoplasts to study ERD2 introduced new tools and findings, which are placed into perspective.

1.3.1 *Nicotiana benthamiana* is an attractive model plant to study the secretory pathway.

For the last two decades, *Arabidopsis thaliana* has been the dominant model for gene identification to study the secretory pathway in plants. For this reason, protein secretion essays and *in vivo* bioimaging using more suitable models such as tobacco was mostly dependent on heterologous expression of *Arabidopsis thaliana* genes. To enable experimental strategies based on homologous expression, I wanted to establish a working practice for *Nicotiana benthamiana* because its diploid genome is sequenced.

The data shown in the first half of this chapter illustrate that *N. benthamiana* is an excellent candidate to replace *Nicotiana tabacum* as model for quantitative transport assays using protoplasts. Only minor modifications of the electroporation procedure were required to identify optimal transfection conditions (Figure 5) and the main advantages are the high yield and quality of protoplast produced. Interestingly, the same number of *N. benthamiana*

protoplasts were capable of producing higher recombinant protein levels per total protein measured compared to *N. tabacum* protoplasts (Figure 4), and this was particularly obvious for proteins synthesized by the secretory pathway (Figure 6). A higher performance can either be explained by a higher transfection efficiency, or by a cytosol with a higher density of ribosomes and loaded tRNAs. In particular the difference between cytosolic and secretory protein expression may also indicate a more active secretory pathway (Figure 5).

1.3.2 Secretory versus cytosolic expression

Unexpected discrepancies between secretory protein yield versus cytosolic protein yield were observed during voltage optimisation of the electroporation conditions and whilst comparing protoplasts from *N. tabacum* and *N. benthamiana*. An interesting observation was that the Amy-GUS ratio is reduced at higher voltages (Figure 5C). Higher voltages are expected to increase the transfection efficiency and copy number of plasmids in electroporated cells, whilst also increasing mortality. A potential explanation for the reduction observed is the fact that GUS is synthesized by free ribosomes in the cytosol whilst Amy is synthesized on the ER membrane surface by membrane bound ribosomes associated with translocation pores. The latter may be a limiting factor in Amy biosynthesis. At higher plasmid copy number, the number of translocation pores may become increasingly saturated, causing the observed reduction of the Amy-GUS ratio. The hypothesis of translocation being the limiting factor for amylase synthesis could be tested via co-electroporation with BiP since it would help with the translocation, but this was beyond the scope of this thesis.

A consequence of the limitation caused by translocation pores for Amy synthesis can also be a plausible explanation for the 3-fold increase of Amy versus the 2-fold increase of GUS when comparing the two species directly (Figure 5A/B). This cannot be explained by differences in transfection efficiency because GUS and Amy were encoded by the same plasmid. Therefore, as hinted in the previous paragraph, it is possible that *N. benthamiana* protoplasts have a more active secretory pathway with either a larger ER membrane surface per cell or a higher density of ER translocation pores. An alternative explanation for this leap is that the 35S promoter functions differently in the two species. In that case it has a bigger advantage when controlling gene expression in *N. benthamiana*. Hereafter this argument could be confirmed/denied by swapping the promoters

controlling GUS and Amy and repeating the experiments from Figure 5A/B. Neither hypothesis could be tested within the scope of this thesis, but they may form the basis for future research.

1.3.3 Functional conservation of ERD2 gene between two plant species can be exploited to carry out gene knockdown as well as gain-of-function experiments

The documented advantage of the *N. benthamiana* protoplast system convinced me to attempt improving a recently established ERD2 activity assay, based on the suppression of HDEL-ligand receptor saturation (Phillipson et al. 2001) by overexpressing ERD2 (An 2015). In Figure 6C I have demonstrated that the overexpression of heterologous ERD2 gene in *N. benthamiana* protoplast is more efficient than in tabacum and in Figure 7B I show that a hybrid molecule generated from two *N. benthamiana* ERD2 genes is equally efficient. This indicates that the two copies of ERD2 found in all higher plants must be structurally highly related, despite a considerable sequence divergence (Figure 3). The hybrid transcript was successfully used to knock-down endogenous ERD2 expression, leading to increased HDEL-cargo secretion (Figure 7C). Nevertheless it required high plasmid concentrations and long incubation for 48 hours to visualise a mild but reproducible increased secretion of HDEL-cargo. This means that the endogenous ERD2 protein must be remarkably stable, suggesting that transport to post-Golgi compartments with lytic properties must be minimal. Anti-sense insensitive *Arabidopsis thaliana* ERD2 could fully suppress HDEL-cargo secretion, suggesting that ERD2 is highly conserved amongst plants and it is interchangeable between species. In addition it also shows via partial loss-of-function analysis that ERD2 is necessary for HDEL-retention, even though the gain-of-function assays already established that ERD2 must be the limiting factor in the retention mechanism.

1.3.4 ERD2 gene is extremely efficient but its biological function can be disrupted by specific point mutations.

The high efficiency of the receptor in *N. benthamiana* protoplast shown in Figures 6/7 point to the possible use of a weaker promoter being a viable option. That would be interesting allowing for the design of a triple-expression cassette. In Figure 8 I show that the weak pNOS promoter only lack behinds the strong 35S promoter at very low plasmid concentrations.

For that reason, it was used to drive the expression of ERD2 genes in the newly created triple expression vector and this way allowing me to quickly screen thought point-mutations of highly conserved amino acids in ERD2 backbone. The triple vector permits the equalising of the transfection efficiency via quantitative GUS assays. By adjusting plasmid dilutions to achieve equal GUS levels, it is likely that cargo synthesis is also comparable. The variable (effector) is the receptor construct with the mutations, and in principle it should have equal chances to be synthesized as well. Any differences observed in cargo transport may thus be fully attributed to the changes in the receptor constructs. This can be considered as an important methodology step forward to study ERD2 function in a quantitative manner.

I show in Figure 9 that the mutant screening lead to the identification of new amino acids that can completely knock-out the receptor function, similarly to the previous reported LLGG mutation. Lastly, the creation of the triple vector helped to also identify mutants 1) do not have effect on the biological function or 2) mutants causing a partial reduction of biological function.

1.4 **Conclusions**

Results presented here certainly suggest that *Nicotiana benthamiana* protoplasts are here to stay, in particular due to its apparent upregulation of the secretory pathway capacity. I also demonstrate that even though the H/KDEL receptor is extremely efficient, its biological function can be diminished by different approaches, such as overexpression of cargo molecules (loss-of-function phenotype observed in Figure 6C), anti-sense inhibition (Figure 7B) or by the point-mutagenesis of specific amino acids (Figure 9C). The identification of amino acids that can completely or partially compromises the receptor activity is exciting and form the basis for future research to be presented in this thesis aiming to show that at least in some cases, there is a link between the lack-of-function and a potential mis-localization of the receptor along the secretory pathway.

Chapter 2 **ERD2 C-terminus is crucial for its biological activity and Golgi residency**

2.1 **Introduction**

The results from the first chapter in this thesis provides an interesting experimental platform for a functional analysis of ERD2 in plants using *N. benthamiana* protoplasts. The secretion assay offers the tremendous advantage of highlighting defects at various stages of the ER retention process because the ability to retain HDEL cargo *in vivo* was directly measured with higher sensitivity compared to earlier models. In comparison, first functional results on ERD2 in eukaryotes were based on peptide-binding studies *in vitro*, together with a redistribution-assay for ERD2 itself, but did not attempt quantification of ligand-retention in the ER as a result of ERD2 expression *in vivo* (Townsley et al. 1993; Townsley et al. 1994; Scheel et al. 1997).

However, the secretion assay with protoplasts only permits a comparison of the severity of the phenotype, it does not allow to classify mutations into different functional categories. One of the key-observations of past studies was the subcellular localisation of ERD2, based on C-terminally tagged ERD2 using epitopes or fluorescent proteins. Having established that ERD2-YFP is biologically inactive (An 2015) despite an accepted dual Golgi-ER distribution, further fusion proteins were generated including N-terminally tagged YFP-ERD2 which was ER retained and also inactive. Supplementing the latter by an N-terminal signal peptide (SecYFP-ERD2) resulted in a Golgi-resident fusion protein, but which was also biologically inactive. Therefore, three different fluorescent proteins yielded three different subcellular localisations, but since none of these were biologically active, it is impossible to determine which of the observed subcellular localisations is the correct one.

In this chapter my aim was to link mutations in ERD2 identified in the secretion assay to potential changes in the subcellular localisation of ERD2. In order to do so successfully, it was important to establish a fluorescent fusion protein that could be observed *in vivo* and that retains biological activity in the secretion assay.

2.2 Results

2.2.1 ERD2-related proteins (ERPs) do not play a role in the retention of HDEL-cargo

Nearly two decades ago after the successful completion of the *Arabidopsis thaliana* genome sequencing project, an ERD2-related gene family was discovered *in silico*, termed ERPs (Hadlington & Denecke 2000). ERPs are uniquely found in plants and organism from the SAR-group and in contrast are not found in other eukaryotes (Silva-Alvim et al. 2018). Figure 10A shows a sequence alignment between one of the *Arabidopsis thaliana* ERP members ERP1 (AT4G38790) and ERD2b. ERP1 shows a significant sequence homology with ERD2, but it harbours an additional N-terminal region which contains a predicted transmembrane domain.

Very little is currently known about the biological function of ERPs and as part of a collaborative project with other laboratory members, I decided to investigate if ERP1 gene product plays a similar role to ERD2 in mediating ER retention of HDEL cargo. Due to the overall similarity between ERPs and ERD2, the first approach was to generate hybrids between the two proteins in highly conserved regions (Figure 10A, red arrows) and investigate the effect HDEL cargo retention. The overall structure of the hybrids is schematically represented in Figure 10B.

As it can be seen in Figure 10C, ERD2b can very efficiently inhibit the secretion of AmyHDEL (con) reducing its SI as shown in Chapter 1. However, the overexpression of ERP1 does not cause any reduction in the SI of AmyHDEL (compare first three lanes). Based on that observation it is clear that, although sharing a significant homology with ERD2, ERP1 does not play a role in mediating increased retention of HDEL cargo when ectopically expressed.

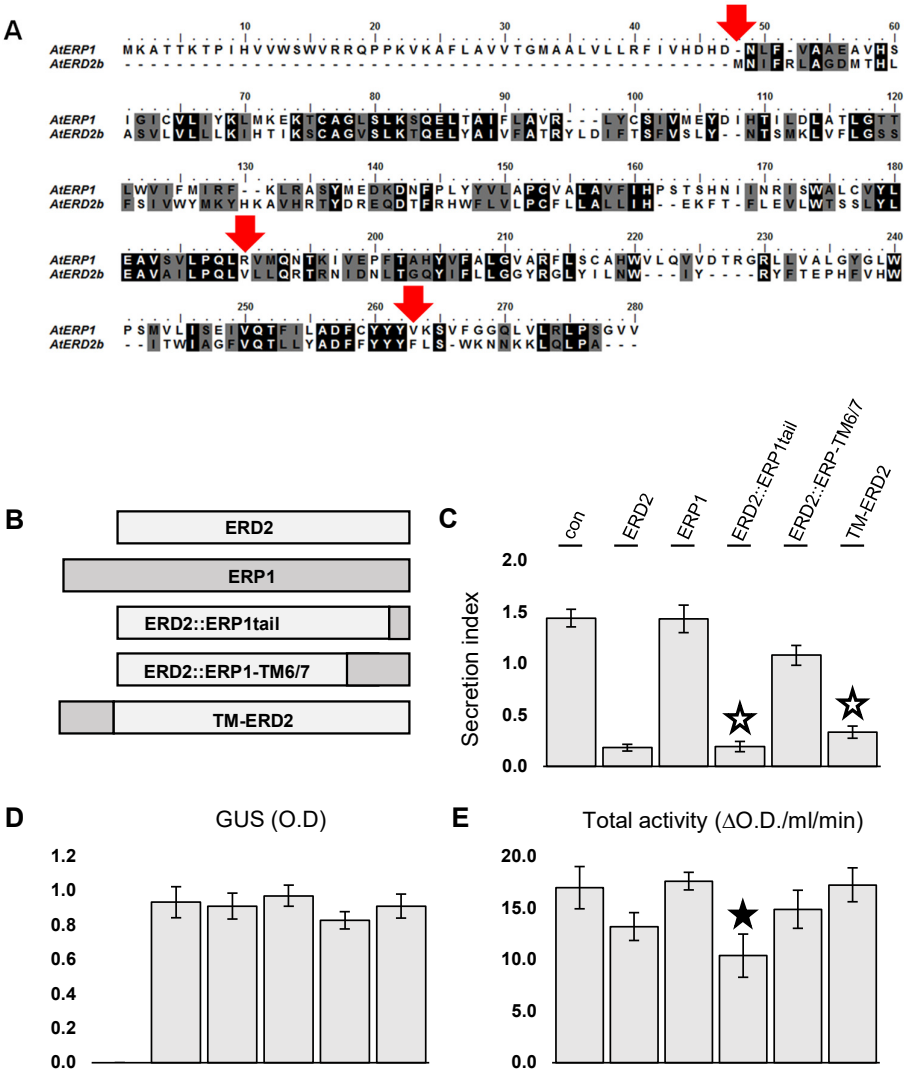


Figure 10 ERD2 and ERP1 do not play the same role mediating ER retention of HDEL cargo.

A) Alignment of AtERP1 (ERP1) with AtERD2b (ERD2) and the specific amino acids selected for fusions to generate hybrid coding regions between the two proteins B) Illustration of chimeric constructs created, ERD2 in light grey and ERP1 in dark grey C) Co-expression of AmyHDEL cargo alone or combined with either ERD2, ERP1 or fusions in *Nicotiana benthamiana* protoplasts. 50 µg of AmyHDEL was co-transfected with variable µg of effector plasmid previously adjusted to yield equal GUS levels in pilot experiment. White starts indicate fusions capable to mediate ER retention. D) Total GUS O.D used as internal control marker for transfection efficiency. E) Total α-amylase activity obtained in each cell suspension given in arbitrary relative units (ΔO.D./ml/min). Black star indicate fusion that reduced total α-amylase activity. Error bars are standard deviations of three independent transfections.*Figure adapted from Silva-Alvim et al. (2018)

2.2.2 Extending the ERD2 N-terminus with the additional transmembrane domain of ERP1 (TM-ERD2) does not compromise HDEL retention

In order to start identifying critical residues in distinguishing ERD2 and ERP1 function, three hybrids were created (Figure 10B). Results from transient expression illustrated that only constructs in which the majority of the ERD2 coding region is maintained retained the ability to mediate increased HDEL retention. Figure 10C shows that replacing the last 14 amino acids of ERD2 by the last 17 amino acids of ERP1 just downstream of the highly conserved tyrosine repeat (ERD2::ERP1tail) did not impair the ability of this ERD2-derivative to retain HDEL cargo (Figure 10C fourth lane). However, replacing a larger portion by ERP1 almost completely abolished its biological activity (Fifth lane). Interestingly, maintaining the entire ERD2 coding region but fusing the additional N-domain of ERP1 resulted in a fusion protein that maintains a strong HDEL-cargo retention activity (Last lane).

It is important to highlight that, for the same GUS levels (Figure 10D), ERD2::ERP1tail reduction of the AmyHDEL SI is accompanied by a drastic reduction in the total AmyHDEL activity (TA) (Figure 10E, black star). Although the two critical Leucine residues in the C-terminus of ERD2 appear to be conserved in ERP1, other residues may influence the biological function. However, the same is not true for TM-ERD2 which can efficiently reduce the SI without compromising amylase activity. It is possible that the N-domain is therefore not responsible for the differences between ERP1 and ERD2 with respect to its role in HDEL-cargo retention.

2.2.3 ERP1 is ER resident and does not reach the Golgi apparatus.

Having established that ERP1 does not play a role in the retention of HDEL cargo, it was interesting to determine its subcellular localisation as an important step forward to unveil the unknown function of the protein. Since the localisation data of any plant ERP's is yet to be reported, ERP1 was selected as initial test subject and a YFP tag was fused to its N-terminus (Figure 11A). Under the control of the weak TR2 promoter ERP1 was well expressed and localized to the ER network, Figure 11B (first row). Co-expression with a red fluorescent protein fusion with the classical Golgi marker Sialyltransferase (ST, Wee 1998) revealed

no co-localisation at all. Since it was shown that fluorescent tagging of ERD2 provides drastically different results depending on how the protein is tagged (An 2015), the observed localisation of YFP-ERP1 in the ER should not be interpreted as a conclusive result for the ERP1 family and ongoing research by other members of the host laboratory aims at establishing both the biological function of ERPs as well as their subcellular localisation.

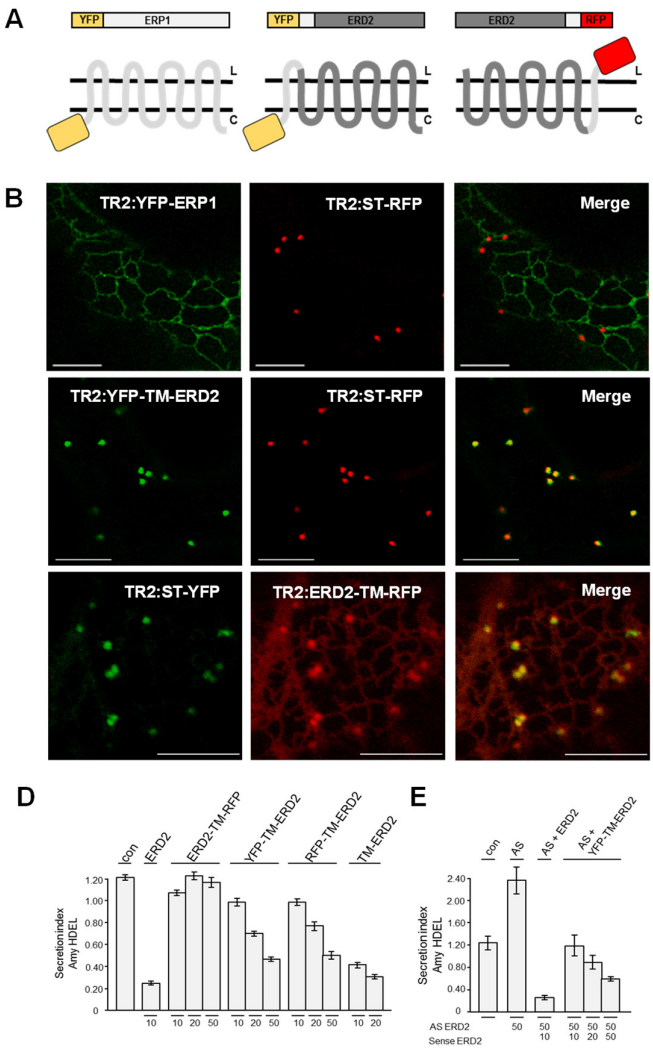


Figure 11 Addition of a transmembrane domain to either the C-terminus or the N-terminus of ERD2 to generate a viable and biological active fluorescent fusion.

A) Illustration of chimeric constructs create via the addition of ERP1 to ERD2 core B) Confocal laser scanning microscopy in *N. tabacum* leaf epidermis cells showing that YFP-ERP1 is ER resident and does not localise to the Golgi apparatus labelled with ST-RFP. The hybrid YFP-TM-ERD2 co-localises with the Golgi-marker ST-RFP and shows no evidence of ER staining. The hybrid ERD2-TM-RFP labels the ER and the Golgi apparatus. All scale bars are 10 µm D) Co-expression of the AmyHDEL with ERD2 and fusions containing an additional transmembrane domain at the N-terminus (YFP-TM-ERD2, RFP-TM-ERD2 and TM-ERD2) or the C-terminus (ERD2-TM-RFP) in *Nicotiana benthamiana* protoplasts. 50 µg of AmyHDEL was co-transfected with increasing amounts of effector plasmids given below each lane in µg. E) Knocking-down the endogenous ERD2 using the antisense (AS) NbERD2ab and complementation of the activity either by the sense wild type ERD2 (AtERD2b) or by the biologically active fusion YFP-TM-ERD2. Experimental conditions are as in Figure 7. *Figure adapted from Silva-Alvim et al. (2018)

2.2.4 N-terminally fluorescently labelled TM-ERD2 resides at the Golgi apparatus.

The high expression of YFP-ERP1 (the work above) compared to YFP-ERD2 (An 2015) suggests that in contrast to the latter, YFP-ERP1 may not be terminally misfolded. The high homology level between ERD2 and ERP1 core suggests that these proteins could have evolved from a common ancestral or from each other. Since addition of the additional N-domain to the coding region of ERD2 (TM-ERD2, figure 10C) had no negative effect on the ability to retain HDEL ligands in plant protoplasts, it is likely that the TM-ERD2 fusion maintains the structural integrity of the ERD2 core. For these two reasons, an attempt was made to create a fluorescent ERD2 fusion that retains biological activity by fusing YFP to the N-terminus of TM-ERD2 (YFP-TM-ERD2, Figure 11A). A second fusion contained the additional TM domain at the C-terminus of ERD2 followed by the red fluorescent protein (ERD2-TM-RFP).

Figure 11B (second row) reveals that YFP-TM-ERD2 labels exclusively the Golgi bodies together with the known Golgi marker ST-RFP. There was no hint of ER stain which was reported for C-terminally tagged ERD2 using either epitopes (Hsu et al. 1992; Lewis & Pelham 1992b) or fluorescent proteins (Boevink et al., 1998; daSilva et al., 2004; Xu and Liu, 2012; Montesinos et al., 2014). In fact, the fusion protein was very well expressed, in contrast to the N-terminal fusion without the additional TM domain (YFP-ERD2, An 2015), but highly similar to YFP-ERD2 supplemented with an N-terminal signal peptide (Sec-YFP-ERD2) which was also exclusively Golgi-localised (An 2015).

Placing the additional TM-domain at the C-terminus followed by RFP resulted in a dual ER-Golgi localisation pattern (Figure 11B, third row), similar to earlier observations with ERD2-YFP and related C-terminal constructs which were biologically inactive (An 2015).

2.2.5 Fluorescently tagged ERD2 retains biological activity if the lumenal side and the cytosolic tail remain unobstructed

Having observed two different subcellular localisation patterns for YFP-TM-ERD2 and ERD2-TM-RFP, it is possible that only one of these is biologically active or none at all as observed earlier with direct fusions to fluorescent proteins without TM domains as separator. Therefore, the gain-of-function assay using *N.*

benthamiana protoplasts was used to test biological activity of the new fusion proteins.

In Figure 11D it is demonstrated that the C-terminal fusion construct (ERD2-TM-RFP) lacked biological activity since even at high dosages this molecule is unable to reduce the SI of AmyHDEL compared to the control (con, first lane). However, the N-terminal fusion YFP-TM-ERD2 mediated a gradual reduction of the AmyHDEL secretion index with increasing plasmid concentrations, leading to a reproducible dose-response, albeit weaker than wild type ERD2. The same result was seen when YFP was replaced by RFP (RFP-TM-ERD2), which shows that the nature of the fluorescent protein was not critical. An additional control was the earlier studied ERP1-ERD2 hybrid (TM-ERD2, Figure 10) which shows higher activity than the two constructs harbouring fluorescent proteins but slightly lower than untagged native ERD2 (Figure 11D, compare second lane with last two lanes).

In addition to these results, and as a further control, the ability of YFP-TM-ERD2 to complement the induced-secretion effect caused by the partial gene knock-down in the presence of NbERD2ab anti-sense (AS) was assessed. Figure 11E shows that YFP-TM-ERD2 can negate the anti-sense effect and progressively cause increased retention of AmyHDEL with increasing plasmid concentrations. Again, the effect is weaker than for untagged ERD2, but the biological activity is clearly measurable.

Taking together all the results show that extending the ERD2 N-terminus by an additional transmembrane domain has only minor negative effects on its biological activity, and enables N-terminal ERD2 tagging via fluorescent proteins without obstructing the lumenal side and without compromising the C-terminus. The resulting construct resides in the Golgi and cannot be detected in the ER.

2.2.6 The new fluorescent fusion (TM-ERD2) remains Golgi resident even after over-expression of HDEL cargo

Since ERD2 discovery, almost three decades ago, only few studies have examined and shown a redistribution of ERD2 from the Golgi to the ER in response to co-expression of ligands in different organism (Lewis & Pelham 1992a; Montesinos et al. 2014). Having established that YFP-TM-ERD2 is a fluorescent tagged ERD2 molecule which retains biological activity the obvious

next step was to assess the possible effect of overexpression of cargo molecules on the localisation of the receptor.

As part of an earlier PhD project in the host laboratory, an innovative method was introduced to visualize ERD2-mediated cargo accumulation in the ER *in situ*. Upon the overexpression of a plasmid carrying two independent HDEL cargoes, the recombinant marker ST-YFP-HDEL and AmyHDEL, the retention system can be saturated leading to a leakage of ST-YFP-HDEL from the ER back to the Golgi, creating a dual localization of the marker (ER-Golgi). However, when co-expressed with untagged ERD2 the Golgi marker is completely retained in the ER once again (An 2015). This gain-of-function approach was used with the newly generated tagged ERD2, RFP-TM-ERD2.

Figure 12 shows the subcellular distribution of RFP-TM-ERD2 in the presence or absence of HDEL overdose. When ST-YFP-HDEL was co-expressed with control cargo Amy and RFP-TM-ERD2, the model cargo was exclusively found in the ER whilst the receptor fusion was exclusively found in the Golgi apparatus (Figure 12A).

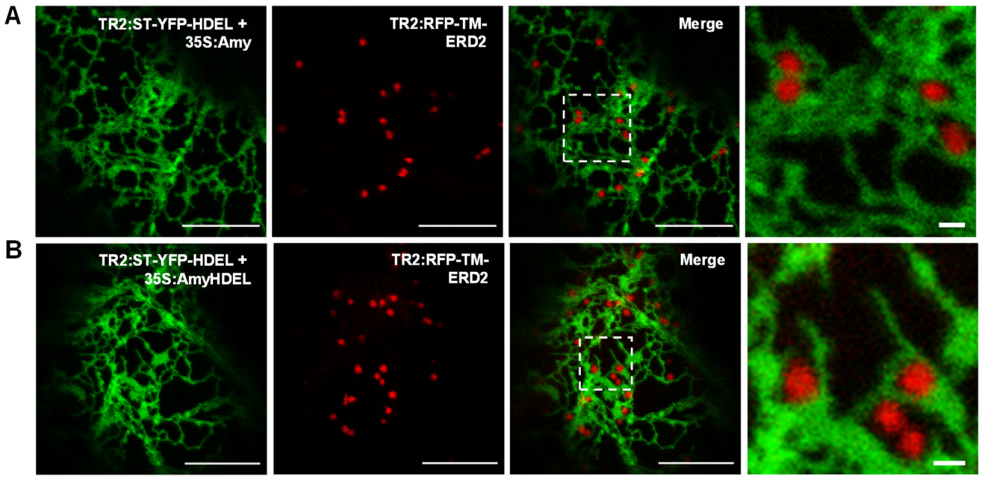


Figure 12 Evidence that ERD2 localisation is restricted to early Golgi cisternae with overexpressed ligands.

A) CLSM showing the distribution of RFP-TM-ERD2 in the absence of ligand over-expression by co-expression with the control construct (TR2:ST-YFP-HDEL + 35S:Amy). B) CLSM demonstrating in situ biological function of RFP-TM-EDR2 co-expressed with the HDEL overdose test construct (TR2:ST-YFP-HDEL + 35S:AmyHDEL). Scale bars are 10µm. Close-ups of the enlarged dashed rectangle in A) and B) show that RFP-TM-ERD2 punctate are well separated from the ER. Scale bars in the close-ups are 1µm. *Figure adapted from Silva-Alvim et al. (2018)

When Amy was replaced by AmyHDEL (Figure 12B), saturation of the HDEL retrieval was never achieved, no ST-YFP-HDEL molecules were detected in the Golgi apparatus, confirming the biological activity of RFP-TM-ERD2 in the *in situ* assay. Secondly, the Golgi-localisation of RFP-TM-ERD2 was never shifted partially to the ER in spite of two separate HDEL molecules being present, one of which being overexpressed.

The observations from Figure 12 argue against the most accepted model by which ERD2 is an active traveller moving in between the ER-Golgi system (Semenza et al. 1990; Lewis & Pelham 1990; Hugh & Pelham 1991). However, it should be noted that those earlier studies made use of C-terminally tagged ERD2.

2.2.7 Golgi-residency of TM-ERD2 fusions is independent of the fluorescent tag and it partially segregates from a trans-Golgi marker.

A real argument to be risen is that the choice of the tag to be used is of high importance since it could influence the localization of the receptor. With that in mind, both previously used ERD2 tagged versions, YFP and RFP, were co-infiltrated and their localization correlation was analysed.

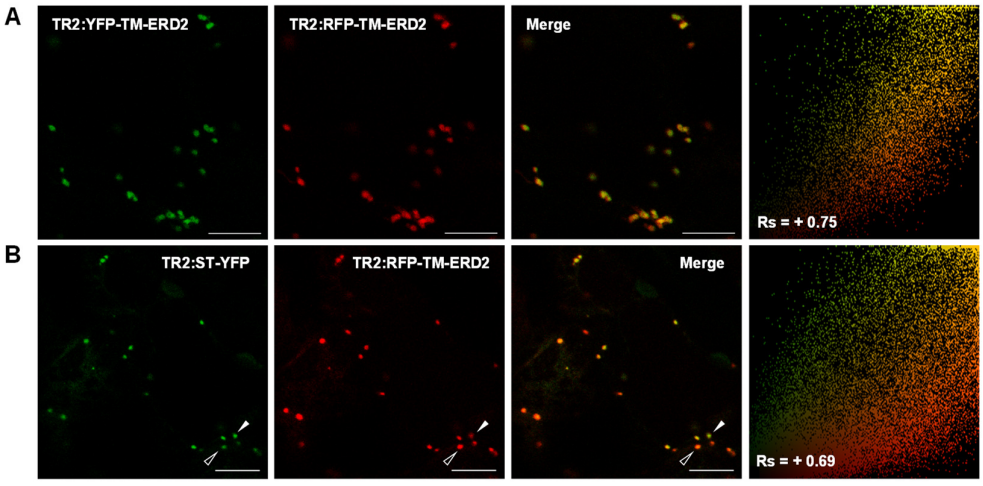


Figure 13 Testing the co-localization of biologically active ERD2 fusions.

A) CLSM image showing YFP-TM-ERD2 co-expressed with RFP-TM-ERD2 showing high level of co-localisation, illustrated by a single yellow pixel population in the scatterplot and a high positive Rs. B) CLSM image of RFP-TM-ERD2 co-expressed with the Golgi-marker ST-YFP showing consistent co-labelling of the same Golgi bodies, but with less correlation between green and red signals, showing a range between mostly red (open arrow heads) or mostly green (white arrow heads) structures, reflected by a broader scatterplot and a lower Rs. All scale bars are 10µm. *Figure adapted from Silva-Alvim et al. (2018)

As it can be seen in Figure 13A, YFP-TM-ERD2 and RFP-TM-ERD2 exclusively label Golgi punctate structures and the two fusions co-localized to a high level.

Figure 13B also shows that the co-expression of RFP-TM-ERD2 with the Golgi marker ST-YFP revealed the same degree of co-localization in the same structures as seen before for the combination YFP-TM-ERD2 and ST-RFP (Figure 11B, second row).

It is interesting to notice that the correlation between ST-YFP and RFP-TM-ERD2 is slightly lower (Figure 13B, last panel) than the co-localisation of YFP-TM-ERD2 and RFP-TM-ERD2. The partial stratification reveals predominantly red (open arrow head) or predominantly green (white arrow head) structures, suggesting that ERD2 may partially segregate from the common Golgi marker, possibly through a different cisternal distribution within the Golgi stack.

2.2.8 High-resolution Airyscan microscopy reveals further details of intra-organelle segregation

To further characterize TM-ERD2, as well as to better dissect the potential intra-organelle segregation, the resolution of the microscopy needed to be enhanced by the use of the Airyscan function in conjunction with higher magnification and narrower pinhole.

Firstly, both RFP and YFP forms of TM-ERD2 were co-expressed and their localization analysed under the new conditions to establish a baseline. Figure 14A once again reveals the strong co-localization resulting in a main diagonal yellow scatterplot and high positive correlation coefficient. This really confirms that the nature of the fluorescent tag affects neither localisation nor function.

When YFP-TM-ERD2 was co-expressed with ST-RFP, the segregation observed in Figure 13 became more evident by the appearance of red-only structures (arrow heads), the presence of a red-only population in the scatterplot and a lower correlation coefficient (Figure 14B). Interestingly all the structures containing YFP-TM-ERD2 also contain ST-RFP. This could indicate that ST-RFP may progress further in the Golgi apparatus compared to ERD2, and it is noteworthy that ST-XFP markers have been known as trans-Golgi markers (Wee 1998; Boevink et al. 1998; Neumann et al. 2003).

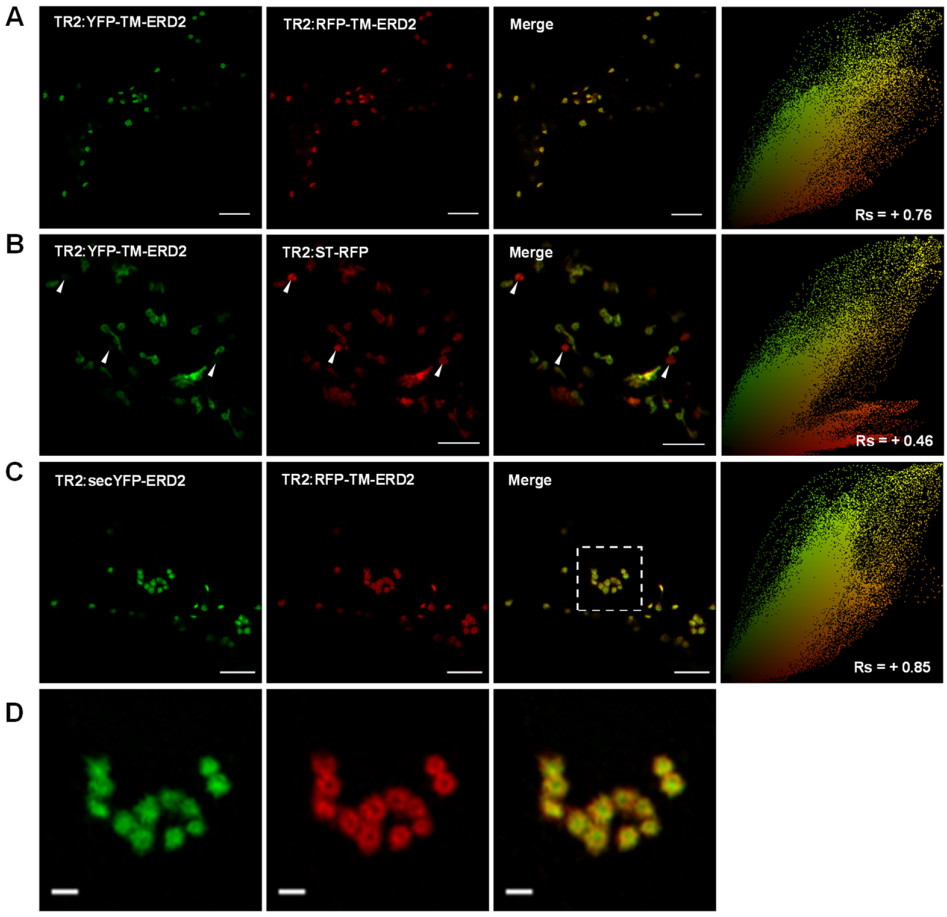


Figure 14 High-resolution Airy scan technology applied to the co-localization studies of ERD2 fusions.

A) CLSM using higher resolution Airy scan detector showing strong co-localisation of YFP-TM-ERD2 and RFP-TM-ERD2. Scatterplot and Spearman correlation coefficient were similar to data from conventional CLSM (Figure 13), confirming that both fusions can substitute for each other. B) CLSM using higher resolution Airy scan detector showing of YFP-TM-ERD2 co-expressed with the Golgi-marker ST-RFP shows a clear segregation of structures labelled solely by ST-RFP (white arrow heads) as can be noticed by the distinct red population on the scatter plot and a significantly lower correlation coefficient. C) CLSM using higher resolution Airy scan detector of non-functional secYFP-ERD2 and functional RFP-TM-ERD2, revealing a very strong co-localisation. D) Close-up of the enlarged dashed rectangle in C) shows that RFP-TM-ERD2 have the doughnut-like appearance labelling the periphery of Golgi bodies and secYFP-ERD2 appears to be more evenly distributed within a Golgi body without a central halo. Scale bars on panels A), B) and C) are 5µm, scale bar on close-up D) is 1µm. *Figure adapted from Silva-Alvim et al. (2018)

Lastly, a previously used tagged version of ERD2 (secYFP-ERD2: An 2015, Silva-Alvim et al. 2018) which highlights exclusively Golgi bodies but does not retain biological activity was co-expressed in conjunction with the newly created RFP-TM-ERD2. This was important because a function in ligand-sorting may yet

reveal a difference that may have been below the detection limit in my previous experiments exploring ligand-induced redistribution (Figure 12).

Figure 14C shows that, although lacking biological activity, secYFP-ERD2 strongly co-localizes with RFP-TM-ERD2 similar to the co-localisation of YFP-TM-ERD2 with RFP-TM-ERD2. The only discernible difference observed by the high-resolution provided by the Airyscan is an “architectural” discrepancy between the two proteins (Figure 14D). Whereas RFP-TM-ERD2 have the doughnut-like appearance labelling the periphery of Golgi bodies, secYFP-ERD2 appears to be more evenly distributed within a Golgi body without a central halo. This difference could be explained by the fact that RFP-TM-ERD2 displays the fluorescent protein on the cytosolic side of the Golgi membrane whilst secYFP-ERD2 is confined to the Golgi lumen. The topology was confirmed by a different team member in the host laboratory (Silva-Alvim et al. 2018).

All considered, the results could indicate that fluorescent TM-ERD2 constructs mainly localizes to the cis-cisternae of the Golgi apparatus, differently from ST-XFP constructs which can proceed to the trans-cisternae (Boevink et al. 1998; Ito et al. 2012). Additionally, the high-correlation between RFP-TM-ERD2 and secYFP-ERD2 argues that the former lack of biological activity is not due to a sorting defect but probably due to other issues, possibly ligand-binding defects caused by obstruction of the N-terminus or steric hindrance with the lumenal portion of ERD2.

2.2.9 A conserved di-leucine motif in the C-terminus of ERD2 is crucial for its Golgi residency

Results so far, in conjunction with earlier observations in the host laboratory, suggest that the alteration of the ERD2 C-terminus by either point-mutagenesis or via C-terminal fusions is highly detrimental to its biological function (An 2015; Silva-Alvim et al, 2108). In addition to that, contrasting results can be found in the literature. A broad examination of ERD2 function via site-directed mutagenesis done in the past revealed no specific residue critically present on that region of the protein (Townsley et al., 1993) but the phosphorylation of a C-terminal positioned serine was suggested to be of high importance to ERD2 transport in mammals (Cabrera et al., 2003).

The establishment that fluorescently tagged TM-ERD2 retains biological function opened the doors to the investigation aiming to link any lack-of-function mutant to an alteration of ERD2 localization.

A previous dissection of ERD2-tail via site-directed mutagenesis (An 2015, schematically shown in Figure 15A) revealed that the double mutant of two highly conserved leucines by glycines (LLGG) causes a strong knock-out inhibiting receptor activity *in vivo* (Silva-Alvim et al. 2018) and used as a negative control for ERD2 function in this thesis (Chapter 1, Figure 9).

Figure 15B shows that the introduction of LLGG mutation in the active fluorescent ERD2 fusion (YFP-TM-ERD2 LLGG) causes a redistribution of the receptor, similarly to the one caused by C-terminal fusions of ERD2 (i.e. Figure 11B). As it can be seen ERD2 still reaches the Golgi, since when co-expressed its partially localizes to ST-RFP, but a strong ER retention is also observed.

It is important to notice that the redistribution was specific to the mutated protein. The co-expressed wild-type fusion RFP-TM-ERD2 remained firmly Golgi-resident (Figure 15C), suggesting that introduced mutant ERD2 does not compromise the behaviour of wild type ERD2.

As a further control to emphasise the importance of ERD2-tail, a truncated version of YFP-TM-ERD2 was created that lacks last 5 amino acids (LQLPA) at the C-terminus (YFP-TM-ERD2ΔC5). Figure 15D shows that once again the receptor localization was shifted to a dual ER-Golgi distribution, showing a clear ER network as well as punctate structures that co-localised with ST-RFP. Finally, Figure 15E shows that the truncated ERD2 does not influence the localization of RFP-TM-ERD2 which remains Golgi-resident.

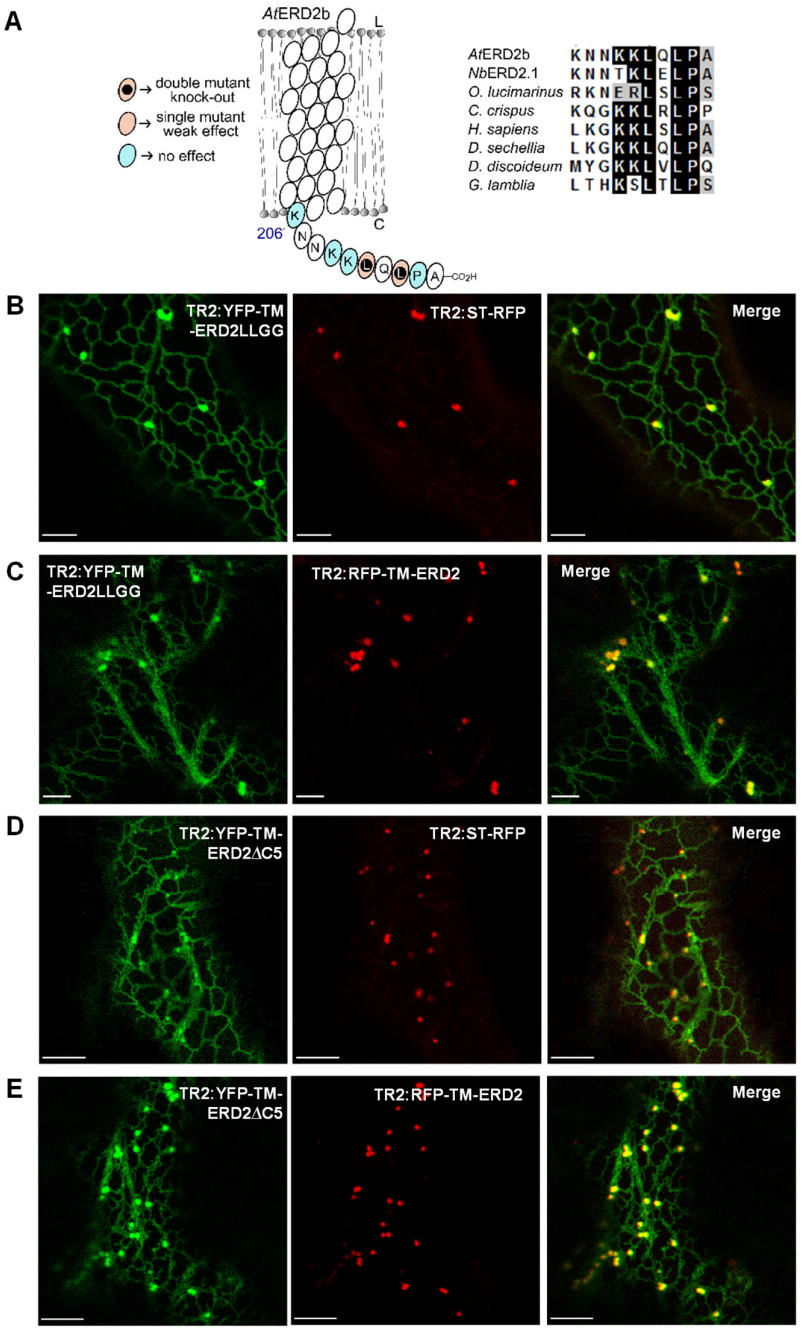


Figure 15 The C-terminus of ERD2 controls efficient ER export and is essential for its biological activity.

A) Illustration of point mutagenesis of the C-terminus and the observed effects in the biological activity followed by an alignment of ERD2 C-termini from different eukaryotes as indicated. B) CLSM showing the distribution of YFP-TM-ERD2-LLGG in comparison with RFP-TM-ERD2. Scale bars are 5µm. C) YFP-TM-ERD2-LLGG in comparison with the Golgi marker ST-RFP. Scale bars are 5µm. D) CLSM showing the distribution of YFP-TM-ERD2-ΔC5 in comparison with RFP-TM-ERD2. Scale bars are 10µm. E) YFP-TM-ERD2-ΔC5 in comparison with the Golgi marker ST-RFP. Scale bars are 10µm. Notice that the non-functional LLGG mutant as well as the C-terminal deletion are still capable to reach the Golgi apparatus but are now also retained in the ER. *Figure adapted from Silva-Alvim et al. (2018)

2.2.10 A different class of ERD2 Loss-of-function mutants maintains Golgi residency

The first mutational analysis of human ERD2 revealed that specific amino acids are required to a proper localization and also for the receptor *in vitro* binding activity be maintained (Townsley et al., 1993). Previous research in the host laboratory has re-evaluated some of these mutants in plants via the bio-assay (An 2015), including 5 original point-mutations with strong phenotypes from the first study (Townsley et al. 1993). Two of these mutants exhibited interesting results in the bio-assay and for that reason I decided to test their localization using the newly established YFP-TM-ERD2 reagent.

The arginine located at ERD2 N-terminus, fifth amino acid (Figure 2), is well conserved and its mutation to alanine (R5A) lead to a weak *in vitro* peptide binding activity (Townsley et al. 1993) and disrupted its *in vivo* biological function (An 2015). Figure 16A shows that the addition of the same mutation to fluorescently tagged ERD2 (YFP-TM-ERD2 R5A) causes no change in the receptor localization and it perfectly co-localizes to the Golgi-marker ST-RFP when co-expressed in tobacco leaf epidermal cells.

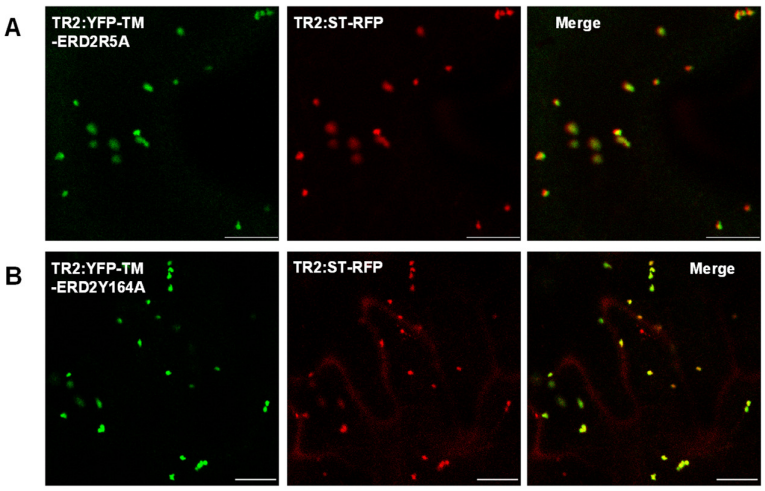


Figure 16 Co-localization studies of ERD2 mutants previously shown to disrupt its biological activity.

A) Confocal laser scanning microscopy in *N. tabacum* leaf epidermis cells showing the distribution of YFP-TM-ERD2R5A in comparison to ST-RFP. B) Confocal laser scanning microscopy in *N. tabacum* leaf epidermis cells showing the distribution of YFP-TM-ERD2Y164A in comparison to ST-RFP. Scale bars are 10µm. Both mutations were previously shown to disrupt ERD2 biological activity (An Jing, 2015) but notice that the receptor is not ER retained.

Figure 16B shows that the mutation of tyrosine 164 to alanine (Y164A) in the fluorescently tagged ERD2 (YFP-TM-ERD2 Y164A) does not compromise the receptor trafficking and still perfectly co-localizes with the Golgi marker. This mutation was particularly interesting as it was previously reported to cause induced secretion of HDEL cargo in *N. tabacum* protoplasts (An 2015) and showed no activity in *N. benthamiana* protoplasts (Chapter 1, Figure 9).

In addition to these two mutations, all the newly designed and analysed point-mutations causing reduction of biological activity in the bio-assay (Figure 9C) where inserted in RFP-TM-ERD2 core and their localization was tested.

Figure 17 summarizes the findings from the initial screen of the new mutants. Surprisingly, the vast majority of mutations remain in punctate structures, despite exhibiting clear reduced or abolished biological activity in the secretion assay.

To clear any doubts regarding the co-localization of the mutants, all were co-expressed with the Golgi marker ST-RFP in tobacco leaf epidermal cells. Figure 18 illustrates that indeed most of the mutations caused no effect on ERD2 trafficking and a perfect Golgi-only co-localization can be observed.

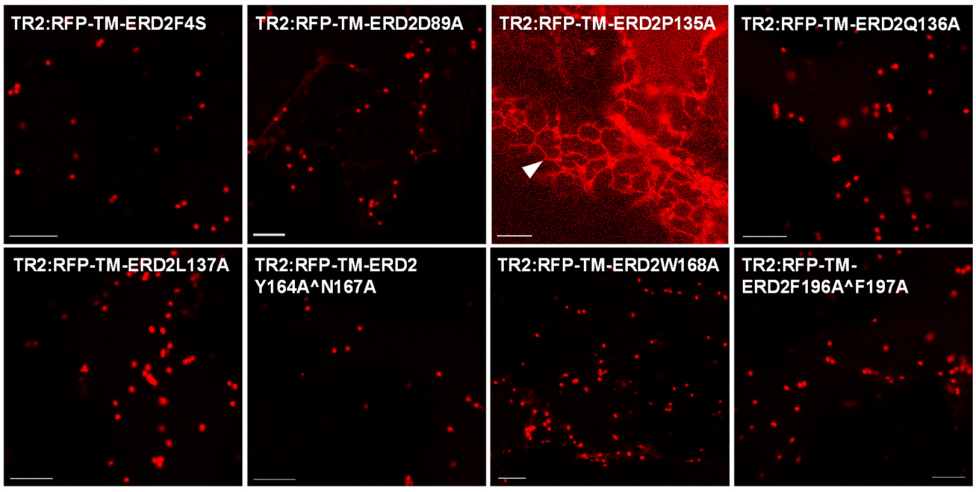


Figure 17 Localization studies of newly created ERD2 mutants that disrupt its biological activity.

Confocal laser scanning microscopy in *N. tabacum* leaf epidermis cells showing that most of the proteins highlights q Golgi-like punctate structure. Notice that RFP-TM-ERDP135A shows a dual distribution highlighting the ER (white arrow) and punctate structures. Scale bars are 10µm.

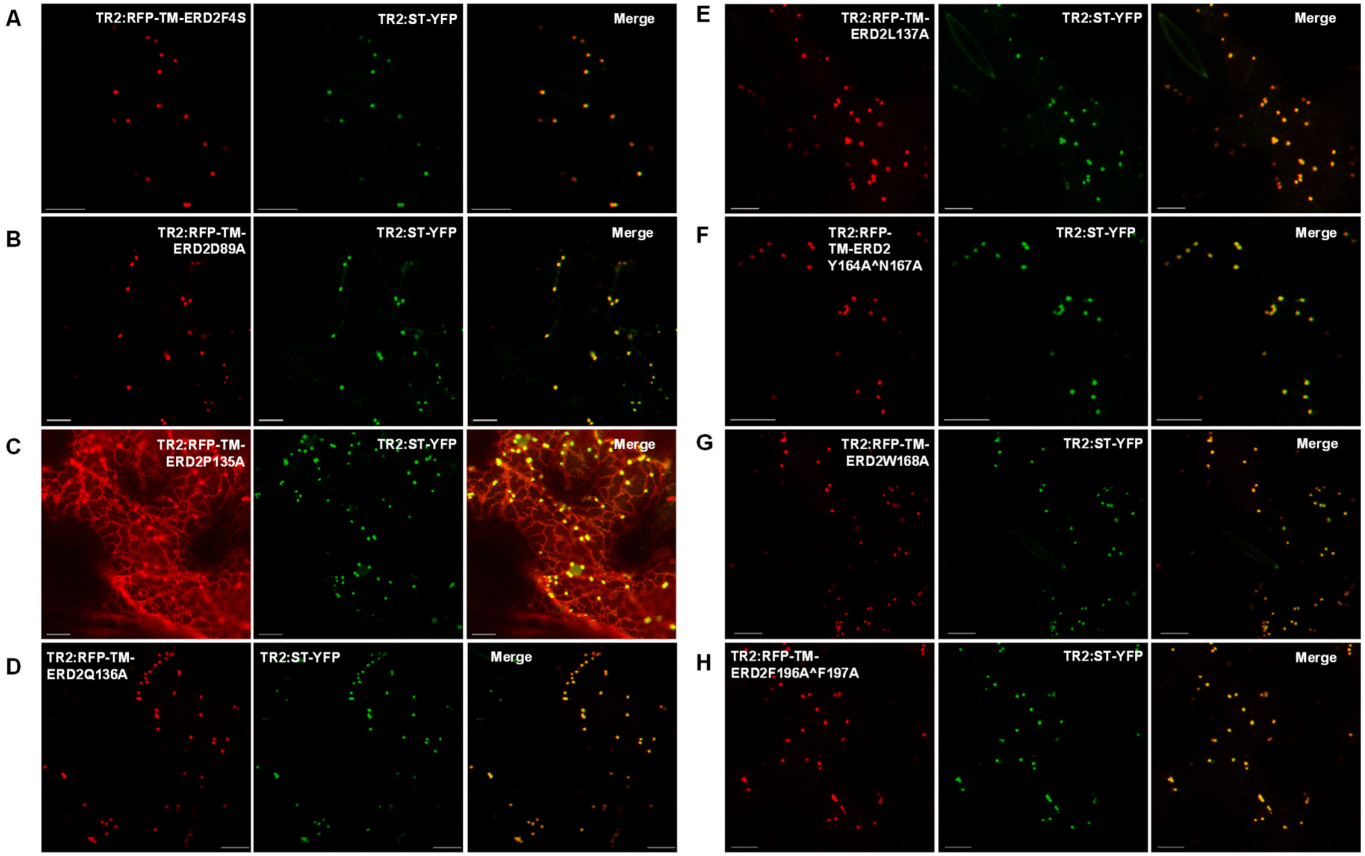


Figure 18 Co-localization studies of newly created ERD2 mutants to identify the nature of punctate structures.

A) to H) Confocal laser scanning microscopy in *N. tabacum* leaf epidermis cells showing the distribution of YFP-TM-ERD2 mutants in comparison to the Golgi marker ST-RFP. Scale bars are 10µm. Notice that with exception of P135A construct all mutants perfectly co-localises to the Golgi marker.

2.2.11 Partial ER retention of further ERD2 mutants reveals the importance of a conserved PQL region typical for PQ-loop proteins

In recent years the human ERD2 was used as initial template protein and a broad analysis looking for homologues has found and established a consensus region, consequently leading to the proposal of a family of proteins termed PQ-loop. All members share a similar structure, consisting of 7-TM domains and they all have at least one conserved PQ-motif (Saudek 2012). The importance of the conserved PQ-motif of ERD2 and how it could potentially affect the protein trafficking or function has not been demonstrated before.

Whilst analysing the co-localization of the newly created mutants the one exception causing miss-sorting of ERD2 was the change of proline located at position 135 for alanine (RFP-TM-ERD2P135A). This mutation is shown to causes a strong ER distribution (Figure 17, white arrow head), but the receptor can still reach the Golgi (Figure 18C). This proline is part of the PQL high conserved region of ERD2 classifying it as part of the PQ-loop family.

Thus, this study reports for the first time that the mutation of this conserved motif can directly influence the localization of ERD2. It is unlikely to be a misfolding mutation because the P135A mutant does not completely abolish the biological activity of the receptor in the secretion assay (Figure 9C), in contrast to the results obtained for the LLGG mutation which totally abolishes activity.

2.3 Discussion

In this chapter I have explored the properties of an ERD2-related protein (ERP1) to study the functioning of ERD2 in further detail. The results have paved the way ahead to further characterise ERPs and led to a strategy to build a new fusion protein of ERD2 that can be observed in situ and that retains biological activity. The work has led to several fundamental advances in our understanding of ERD2 and ERP1 and will be discussed here.

2.3.1 Differences in the N- and C-termini of ERP1 compared to ERD2 cannot explain the lack of an HDEL-retention function in this protein.

ERP1 and related family members have a long N-terminal extension (i.e. 47 amino acids) compared to ERD2, including a hydrophobic domain flanked by charged residues that is predicted to be a transmembrane domain (Figure 10A). ERP1 also contains a short but noticeable C-terminal extension of three amino acids. Having established that ERP1 overexpression has no effect on HDEL cargo in contrast to ERD2, I created ERP1/ERD2 hybrids to investigate potential key-elements of the coding region. Interestingly, adding the ERP1-specific 47 amino acid N-terminal extension to ERD2 had hardly any negative effect on the functioning of the ERD2 core (Figure 10B), thus suggesting that the ERD2 portion was correctly folded and active. Protease-protection experiments by another team member revealed that the native N-terminus of ERD2 is lumenal (Silva-Alvim et al. 2018). It is noteworthy that the additional transmembrane domain in ERP1 is preceded by positively charged amino acids. According to the positive-inside rule, the N-terminus of ERP1 is cytosolic, and fusion of the first 47 amino acids of ERP1 to ERD2 is therefore not expected to change the membrane topology of ERD2.

Interestingly, substituting the C-terminal cytosolic portion in ERD2 by the equivalent portion of ERP1, resulted in a fusion protein still capable of mediating increased HDEL cargo retention, despite the longer C-terminus. The combined results suggest that functional differences between ERP1 and ERD2 must be caused by changes in the core region (after the first 47 amino acids of ERP1) where both type of proteins show overall sequence similarity and a similar transmembrane domain structure. This was confirmed by a hybrid in which a larger portion of ERP1 was inserted to replace the corresponding ERD2 region (Figure 10B). It should be pointed out that the results only suggest that either an essential ERD2 element was lost in the fusion protein, or that the hybrid is not folded correctly. Since only the ability to retain HDEL cargo was scored, no conclusions can be drawn about a potential ERP1 function which is still unknown. In the absence of a biochemical activity assay for ERP1 function, further work on ERP1 was beyond the scope of this thesis.

2.3.2 Golgi-residency of a biological active ERD2 fusion.

The results shown in this chapter demonstrate the cruciality of having the fluorescent tag well separated from the ERD2 N-terminus and the necessity of maintaining ERD2 C-terminus unaltered so that the receptor can preserve its proper localization and biological function.

In contrast to previous published data I have evaluated the capacity of the R/YFP-TM-ERD2 to *in vivo* and *in situ* mediate the retention of HDEL ligands via two different assays (Figure 11 and 12). Both methods, bio-assay and *in situ* assay, are highly sensitive and in association to each other are strong tools to understand how the receptor behaves when altered.

For the first time it was shown here that ERD2 is a Golgi-resident molecule, independently of the presence of extra-stoichiometric levels of cargo molecules, and also that the receptor does not progress further to the trans-Golgi. The stratification observed in Figures 13/14B could be an indication of a possible cis-trans segregation due to be confirmed. This could also explain the high stability of endogenous ERD2 as suggested from the anti-sense inhibition assay in Chapter 1.

The Golgi apparatus operates as the main sorting station of the cells and handles both secreted and endocytosed proteins. A cis-trans segregation makes perfect sense to maintain the adequate balance and avoid mis-binding of receptors and cargoes. Lastly this segregation is in accordance to the most recent studies of the Golgi cisternal maturation model in mammalian cells, describing the Golgi as a tripartite organelle (Papanikou & Glick 2014; Day et al. 2013).

Finally, a difference was observed when localizing two ERD2 molecules tagged via different strategies (Figure 14D). This difference could be explained by the nature of the position of the fluorescence tag considering that for RFP-TM-ERD2 it is on the cytosolic side of the Golgi and for secYFP-ERD2 it is on the lumenal side. An alternative explanation is that the non-functional nature of secYFP-ERD2 causes an altered distribution compared to the functional RFP-TM-ERD2. Further work is needed to distinguish between these possibilities.

2.3.3 ERD2 mutations can be classified into different categories, revealing important amino acids controlling ERD2 localization.

Having established a biologically functional fluorescent ERD2 fusion was an important step forward and permitted the analysis of ERD2 mutants further. This chapter reveals that, so far, the mutants created and analysed can be classified in three categories 1) a double-mutant that causes lack-of-function and altered localization, 2) mutants that causes lack-of-function but do not alter the localization and finally 3) a mutant causing a partial reduction of function and altering the localization of the receptor.

It is clear that further analysis of mutants is needed to map functional domains of ERD2 and potential interactions with ligands or with other membrane proteins that could contribute to ERD2 function. Although the ERD2 C-terminus was earlier considered unimportant for function (Townsley et al. 1993), the current findings illustrate that both the Golgi-residency and its function in promoting HDEL-cargo retention are dependent on the C-terminus. Mass-alignment of ERD2 from many different species confirmed that the two identified C-terminally located leucine residues, positioned at -3 and -5, are highly-conserved. The pair of leucine is very often flanked by other conserved amino acids and a LXLP motif is most commonly found. It would be interesting to test if the LLGG mutation causes changes in protein-protein interactions involving the ERD2 C-terminus, but so-far co-IP experiments with ERD2 as bait have failed to reveal any binding partners (An and Denecke, unpublished results). Further essays beyond HDEL retention and subcellular ERD2 localisation would however be extremely useful in classifying the mutations into further sub-categories. One of these could be the potential oligomerisation of ERD2, which was proposed earlier but using C-terminal fluorescent ERD2 fusions (Xu & Liu 2012). Since these have been shown to be non-functional (An 2015; Silva-Alvim et al. 2018), the potential oligomerisation should be tested again using functional ERD2 fusions.

Chapter 3 **Further understanding the role of ERD2 C-terminus in the localization and efficient transport of the receptor.**

3.1 **Introduction**

The idea that protein sorting receptors would continuously shuttle in vesicles between donor and acceptor compartments was originally inspired by the mechanism of mannose 6-phosphate (Man-6P) dependent lysosomal protein targeting. Man-6-P receptor is a type I membrane spanning protein with a large lumenal domain that binds ligands in the Golgi, travels in clathrin coated vesicles to the early endosome where it releases them. The receptor then recycles back to the Golgi for new rounds of cargo selection (Duncan & Kornfeld 1988; Seaman 2005). This principle was quickly applied to the mechanism of K/HDEL-mediated ER retention, because it was demonstrated that the diffusion of soluble proteins in the ER was not influenced by the presence or absence of the KDEL signal (Ceriotti & Colman 1988), ruling out retention by association to a membrane receptor in the ER. In addition, fusing the KDEL signal to the lysosomal protein cathepsin D resulted in its accumulation in the ER, but it continued to be modified by phosphorylation, which suggests that the recombinant protein is exposed to Golgi-resident phosphotransferases (Pelham 1988). When it was finally shown that the K/HDEL receptor was found in both the ER and the Golgi and appeared to redistribute strongly to the ER when its ligands were overproduced (Lewis & Pelham 1992a), the field generally accepted that the K/HDEL receptor (ERD2) shuttles between the ER and the Golgi to mediate cargo transport.

Results from Chapter 2 show that a biological active receptor is Golgi-only localized (Figure 11) and does not redistribute to the ER even when HDEL-ligands are overproduced (Figure 12). To reconcile these results with the generally accepted model for the recycling of the K/HDEL receptor is difficult, but could be explained by a very fast anterograde transport of the receptor, rendering its steady state levels in the ER beyond the detection limit (Silva-Alvim et al. 2018).

Masking of the ERD2 C-terminus by fluorescent tags can lead to the well-described dual ER-Golgi distribution, but it is accompanied by a complete loss of function (An 2015). In the same work, a conserved di-leucine motif at the ERD2 C-terminus was identified that is essential for the biological activity of ERD2. In

Chapter 2 I show that this mutation causes a dual ER-Golgi localisation of the biologically active fusion protein YFP-TM-ERD2 (Figure 15B). This change in steady state distribution can either be explained by defective ER export, or by faster Golgi to ER transport, but this will require further analysis of the nature of the C-terminal sorting signal of ERD2.

Membrane spanning proteins often display sorting signals at their C-termini. A good example of these is the vacuolar sorting receptor (VSR) gene family, which encodes type I membrane spanning receptors which mediate the traffic of specific soluble cargo to the plant vacuoles. Although the exact nature of the donor and acceptor membranes is currently under debate, the transport does involve Golgi and post-Golgi organelles (daSilva et al. 2006; Foresti et al. 2010; Künzl et al. 2016; Früholz et al. 2018). The short cytosolic tail of VSRs is responsible for the targeting of the receptor and includes the well-known tyrosine motif (YXXϕ) for transport to and from the prevacuolar compartment (PVC), signals for ER export and signals for endocytosis to retrieve mis-targeted receptors from the plasma membrane (daSilva et al. 2006; Foresti et al. 2010; Gershlick et al. 2014). The C-terminus appears to function independently from the lumenal domain and has been studied via deletion and transplantation experiments.

Another example of a C-terminal sorting signal is the so-called KKXX motif for COPI-mediated retrieval of membrane proteins from the Golgi apparatus, displayed by members of the P24 protein family. These type I membrane spanning proteins are highly conserved amongst eukaryotes and have been shown to be components of both COPI and COPII vesicles (Kaiser 2000). p24δ5 (p24a) steady-state was shown to be the ER even though these proteins are thought to be actively travelling in between the ER-Golgi interface (Contreras et al. 2004; Langhans et al. 2008). Transplantation and deletion experiments have been used to describe these and other signal motifs involved in ER retention (Munro & Pelham 1987; Jackson 1993), post-Golgi sorting (daSilva et al. 2006; Foresti et al. 2010; Gershlick et al. 2014) and Golgi retention (Gao et al. 2012).

In this chapter my aim was to characterise the C-terminus of ERD2 via further transport studies to understand the nature of ERD2 trafficking. In addition, I have used a synthetic biology approach to replace the native LQLPA signal by different

types of sorting signals and tested how re-direction of the receptor influences its biological function.

3.2 Results

3.2.1 The conserved di-leucine motif of the ERD2 C-terminus is not an ER export signal but may promote Golgi retention

To understand the nature of the putative Golgi localisation signal at the ERD2 C-terminus, the reasons for the partial ER localisation of the LLGG mutant was investigated further. One possibility is that the mutation inhibits rapid ER export of the receptor, leading to its increased retention in the ER. Alternatively, the mutant could exhibit faster Golgi-to-ER transport, thus resulting in a higher steady state level in the ER.

To distinguish between the two possibilities, I have measured fluorescence recovery after photobleaching (FRAP) using either the wild type construct YFP-TM-ERD2 or the mutant YFP-TM-ERD2LLGG and compared it with the trans-Golgi marker ST-YFP. Inhibited ER export would lead to a slower recovery. Faster Golgi to ER transport may result in the opposite scenario, a faster recovery since more receptors recycle. Tobacco leaf epidermal cells were thus infiltrated with the various constructs together with ST-RFP to track Golgi bodies. 48 hours after infiltration, small sections of the infiltrated leaves were removed and kept in a solution containing latrunculin B solution to promote disruption of the cytoskeleton and stop Golgi movement. Samples were then analysed via laser scanning confocal microscopy (CSLM).

To establish a baseline recovery ratio to future comparison, I firstly examined the movement of ST-YFP. Golgi bodies with high expression where determined as region of interest (ROI) and either selected to be bleached (black arrow-head) or not as a control (white arrow-head) Figure 19A shows that 50% recovery of the bleached areas was observed after approximately 400 seconds. YFP-TM-ERD2 behaved in almost the same manner (Figure 19B). In contrast, the mutant YFP-TM-ERD2LLGG recovered significantly faster, crossing the 50% threshold after just 240 seconds, and reaching almost 85% of recovery after 570 seconds.

The results obtained strongly rule out that increased steady state levels of the LLGG mutant of ERD2 in the ER are caused by a defect in ER export. Instead,

the faster recovery points at an increased proportion of ERD2 recycling between the ER and the Golgi, which can only be explained by a less efficient Golgi-retention.

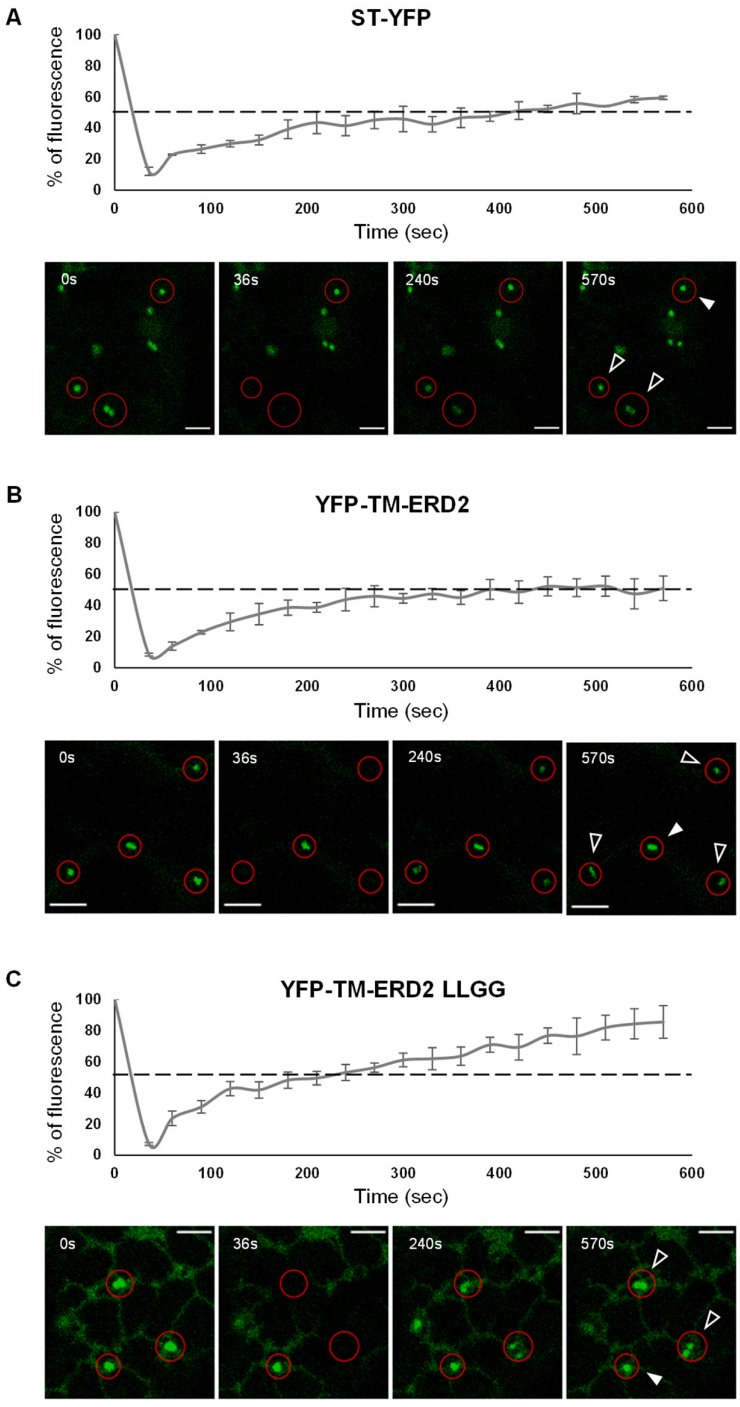


Figure 19 Fluorescence recovery after photobleaching (FRAP) to study the role of the LXLP motif.

CLSM analysis of transformed tobacco leaf epidermal cells using ST-YFP, YFP-TM-ERD2 or YFP-TM-ERD2 LLGG in combination with ST-RFP as a control to track Golgi bodies. Samples were treated with latrunculin B solution to stop Golgi movement via cytoskeleton disruption. A) Percentage of fluorescence recovered over the period of 570 seconds in bleached regions of interest (ROI, black arrow-head) and control ROI (white arrow-head) of leaves expressing ST-YFP. B) Percentage of fluorescence recovered over the period of 570 seconds in bleached regions of interest (ROI, black arrow-head) and control ROI (white arrow-head) of leaves expressing YFP-TM-ERD2. C) Percentage of fluorescence recovered over the period of 570 seconds in bleached regions of interest (ROI, black arrow-head) and control ROI (white arrow-head) of leaves expressing YFP-TM-ERD2 LLGG. Notice that the mutant has a much faster recovery compared to control samples. Scale bars are 5µm.

3.2.2 The cytosolic ERD2 C-terminus is necessary but not sufficient for Golgi residency

Experiments so-far suggest that the di-leucine motif is necessary for Golgi retention of ERD2, but it remains to be shown if the signal is sufficient for Golgi-residency. In addition, I wanted to test if the potential sorting signal can be used to titrate ERD2-binding partners and inhibit ER retention of HDEL ligands. The idea was inspired by previous experiments in which the cytosolic sorting motif of the plant vacuolar sorting receptor (VSR) was overproduced in an artificial construct lacking the ligand-binding motif, resulting in a strong competitor for endogenous receptors, leading to inhibition of vacuolar sorting (daSilva et al. 2005; daSilva et al. 2006). For this reason, the approach taken was to remove all but the last TM domain and the cytosolic tail of ERD2 and replace the deleted portion by a signal-peptide YFP fusion (secYFP). If the last TM domain of ERD2 remains embedded in the ER membrane when expressed as a single TM domain, the resulting construct would mature into a type I membrane spanning protein with the YFP-portion in the lumen and the ERD2 tail exposed in the cytosol.

The last 42 amino acids of ERD2b N-terminus were thus directly fused to the secYFP coding region to generate secYFP-ERD2b-TM7 (Figure 20A). Subsequently tobacco leaves were infiltrated and analysed via CLSM. Interestingly, in contrast to similar fusions with the TM domain and cytosolic tail of plant VSRs (daSilva et al. 2005; daSilva et al. 2006), secYFP-ERD2b-TM7 was very poorly expressed and when detected in very few cells, it was seen exclusively in the ER and the nuclear envelope (Figure 20A).

The results show that the fusion protein did not contain a dominant Golgi-localisation signal. In addition, a very low expression pattern with eventually week staining of the ER is commonly seen for unfolded proteins. It is possible that the 7th TM domain of ERD2 cannot mediate membrane spanning properties and the fusion protein may be misfolded in the lumen of the ER.

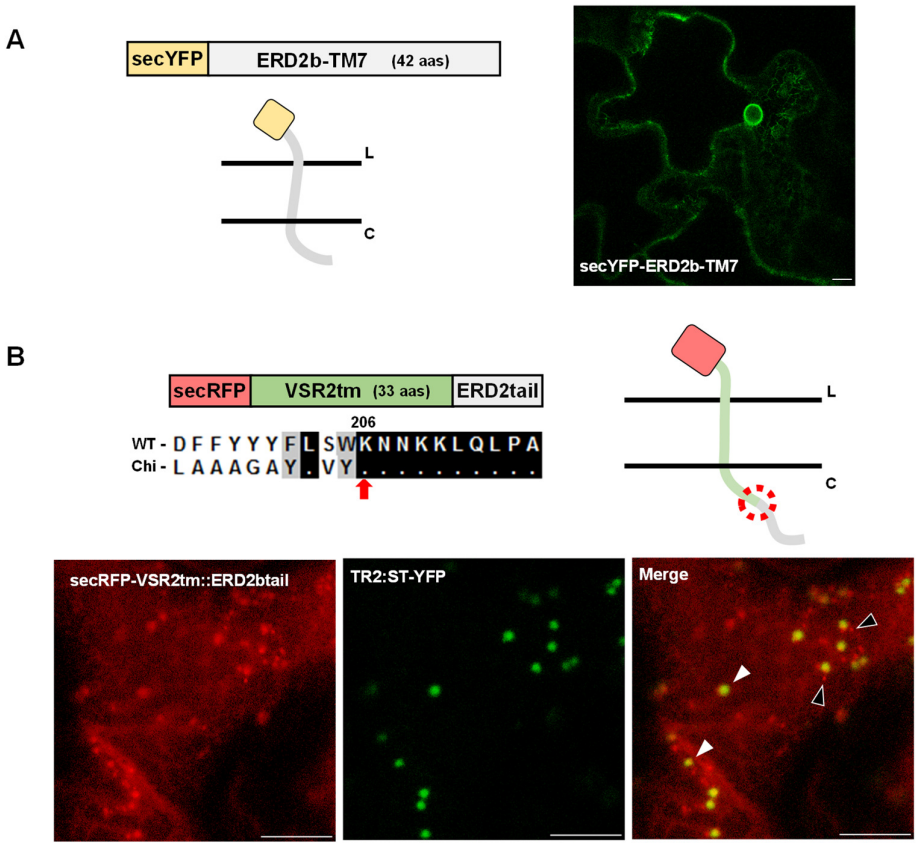


Figure 20 Studying ERD2 C-terminus Golgi-targeting sufficiency.

CLSM analysis of tobacco leaf epidermis expressing two proteins created using different strategies to isolate ERD2 C-terminus. A) First panel shows the illustration of truncated ERD2 construct where secYFP was fused directly fused to the last TM domain and its expected topology. The pattern observed with very low expression and eventual week staining of the ER is illustrated in the second panel. B) Upper panel illustrates the chimeric construct where the last 10 amino acids of ERD2b replaced the cytosolic tail of VSR and its expected topology. Bottom panel shows strong expression of secRFP-VSR2tm-ERD2tail, which can reach the Golgi since it co-localized with the marker ST-YFP but does not promote retention. Scale bars are 10µm.

An alternative approach was pursued to display the ERD2 C-terminus at the C-terminus of a confirmed type I membrane spanning protein marker. Earlier research on plant VSRs resulted in a well-characterised red fluorescent protein (RFP) construct that is transported from the ER-Golgi system to the PVC and recycles back to the ER-Golgi system (Foresti et al. 2010). The construct is composed of a secRFP coding region followed by the TM domain and cytosolic tail of VSR. The cytosolic tail of VSR was thus replaced by the last 10 amino acids of ERD2b (ERD2-tail) to achieve cytosolic display on the surface of the ER-Golgi membrane. The newly generated fusion protein, secRFP-VSR2tm::ERD2tail, is schematically represented in Figure 20B.

To test the subcellular localisation of secRFP-VSR2tm::ERD2tail, the construct was co-expressed with the Golgi marker ST-YFP in tobacco leaf epidermal cells and analysed via CSLM. Figure 20B shows that ERD2tail is again not sufficient by itself to promote Golgi retention since the protein is found not only in the Golgi (white arrow head) but also in structures that resemble the ER, vacuole and post-Golgi organelles (black arrow head), a characteristic localization achieved by the VSR2 TM domain alone (Gershlick et al. 2014).

In conclusion, the di-leucine motif in the cytosolic tail of ERD2 is necessary for Golgi localisation but it appears to be dependent on the structural context of the ERD2 protein to facilitate Golgi-retention.

3.2.3 The VSR C-terminus cannot mediate post-Golgi targeting to the PVC when displayed at the ERD2 C-terminus

In contrast to the ERD2 C-terminus, the VSR C-terminus contains sorting determinants that appear to be independent of the lumenal domain of VSR and mediate ER export, Golgi to PVC transport, recycling from the PVC and endocytosis from the plasma membrane (DaSilva et al., 2005, 2006, Foresti et al., 2010, Gershlick et al., 2014). To test if ERD2 can be re-directed to and from post Golgi organelles, I decided to test the reciprocal experiment and replace the last 10 amino acids from the ERD2 coding region by the entire 36 amino acid long VSR2 C-terminus. This strategy could potentially lead to a gain-of-function or altered-function phenotype, because mutation of the di-leucine motif or deletion of the last 5 amino acids of ERD2 did not impair arrival in the Golgi apparatus but resulted in faster retrograde transport to the ER. The VSR C-terminus could therefore potentially introduce post-Golgi trafficking of the ERD2-VSR hybrid. This experiment would also tackle the question of the Golgi-retention mechanism of ERD2 and what prevents it from reaching the plasma membrane or the vacuolar membrane, mechanisms which remain elusive to date (Pfeffer 2007).

For the purpose of co-localizing the synthetic ERD2 I took advantage of the recently created YFP-TM-ERD2 and replaced the most C-terminally located 9 amino acids of ERD2 by the entire 36 amino acid long VSR2 C-terminus, to create YFP-TM-ERD2::VSR2tail. Figure 21A schematically represents the new fluorescent protein and shows the specific amino acid where the change took place. Figure 21B shows that YFP-TM-ERD2::VSR2tail is partially confined to

the ER but also reaches round Golgi-like structures, but no evident post-Golgi organelle staining was detectable. Essentially, the subcellular localisation was similar to the C-terminal fusion ERD2-YFP (An 2015) or the mutation/deletion of the di-leucine motif in YFP-TM-ERD2 (Chapter 2, Figure 11B).

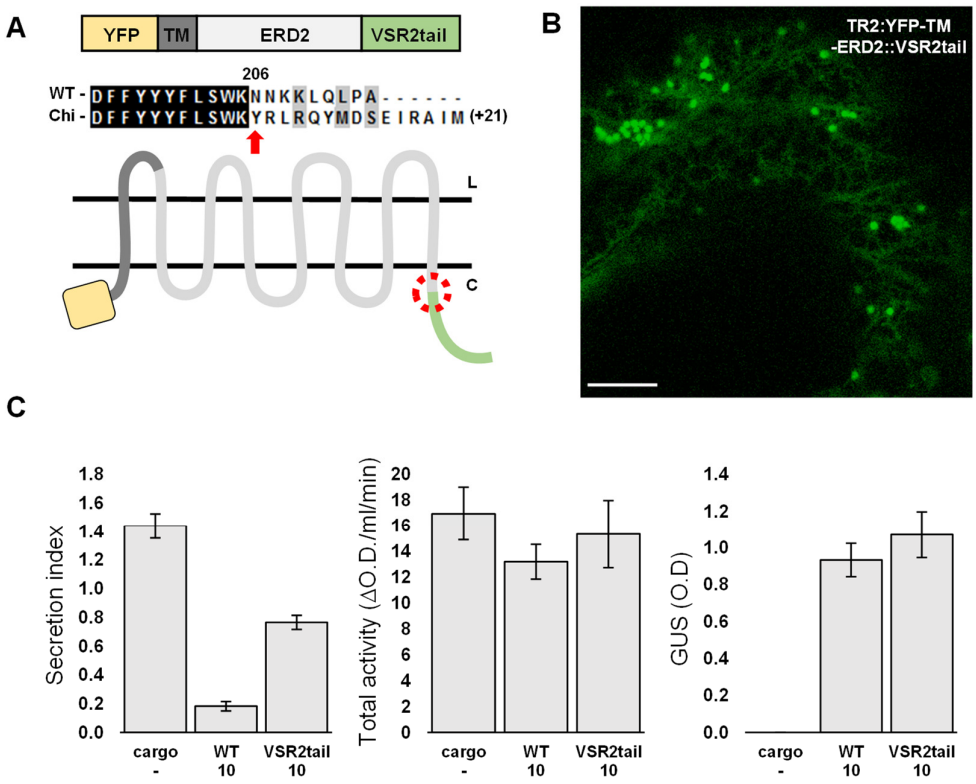


Figure 21 Synthetic biology approach to redirect ERD2 to post-Golgi intracellular compartments.

A) Illustration of the hybrid protein where the last 9 amino acids of ERD2 C-terminus were replaced by 36 amino acids from VSR2 C-terminus. The exact position where the replacement was made and the expected topology are also represented. B) Confocal laser scanning microscopy in tobacco leaf epidermis cells showing the distribution of YFP-TM-ERD2::VSR2tail. Scale bar is 10µm. C) Co-expression of AmyHDEL cargo alone or combined with atERD2b (WT) and the hybrid protein (VSR2tail) in *Nicotiana benthamiana* protoplasts. 50 µg of AmyHDEL was co-transfected with variable µg of effector plasmid adjusted to yield equal GUS levels in pilot experiment. Total α-amylase activity is given arbitrary relative units (ΔO.D./ml/min). Total GUS O.D used as internal control marker for transfection efficiency. Error bars are standard deviations of three independent transfections.

To carry out a biochemical transport assay on the cargo molecule AmyHDEL, the same C-terminal replacement construct was imposed on untagged ERD2 within the dual expression vector carrying the GUS expression marker (Chapter 1, Figure 6B). Figure 21C shows that for similar GUS levels compared to the construct expressing *wild type* ERD2 (WT), the newly generated synthetic

receptor has a reduced ability to convey increased AmyHDEL cell retention, but in contrast to the non-functional LLGG mutant, it is still capable of mediating a reproducible reduction of the AmyHDEL secretion index compared to the cargo-only control lane. This result was surprising because the di-leucine motif was completely replaced. It should, however, be noted that a reduction in the secretion index does not mean that AmyHDEL was re-distributed back to the ER. It is possible that the VSR tail mediates weak cycling through the PVC, but this was beyond the detection limit of confocal laser scanning microscopy (Figure 21B).

3.2.4 HDEL-overdose promotes ER retention of ERD2::VSR2tail

Unexpected results from Figure 21, the dual-localization of the synthetic receptor and the partial reduction of AmyHDEL secretion, convinced me that further analysis of ERD2::VSR2tail was necessary. A trustworthy co-localization, using Golgi markers, was needed to be able to determine the nature of the Golgi-like structures observed before (Figure 21B). YFP-TM-ERD2::VSR2tail was co-expressed with the Golgi marker ST-RFP in tobacco leaves and their co-localization was checked via CLSM. Figure 22A reveals that both markers perfectly co-localize to the typical Golgi-like structures, and that the synthetic receptor does not reach post-Golgi organelles in detectable quantities. The dual ER-Golgi localisation (Figure 21B) is therefore similar to earlier observations with mutated, deleted or masked di-leucine motifs (Chapter 2 and An 2015).

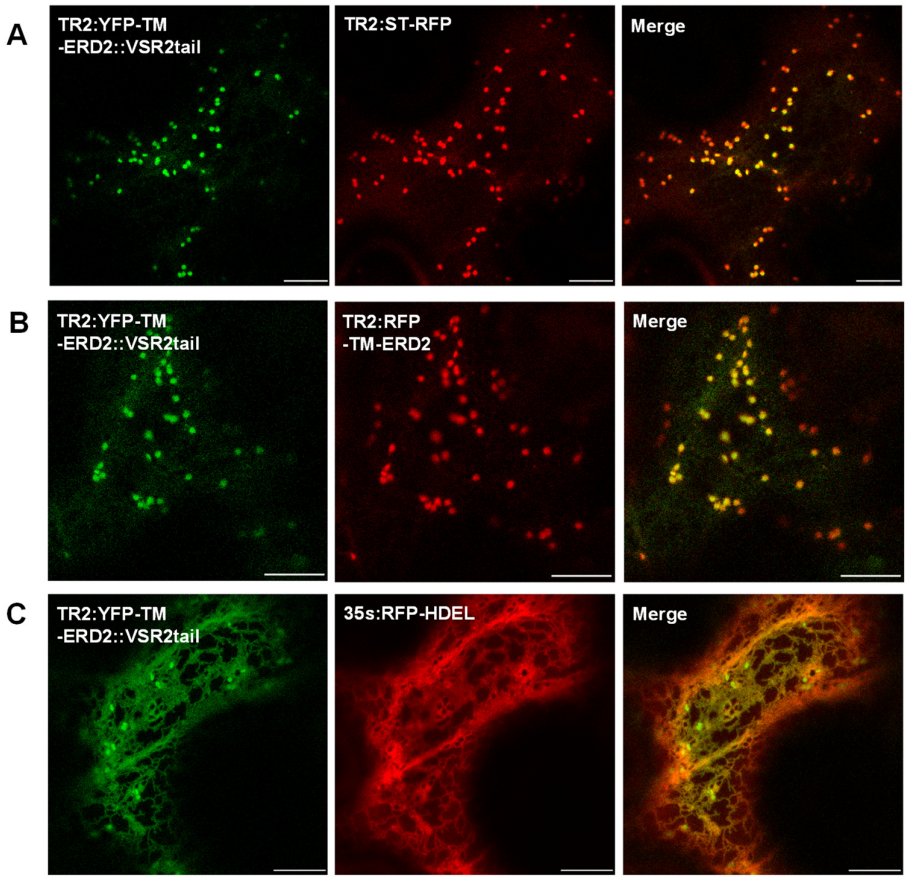


Figure 22 Co-localization studies of ERD2-VSR2 hybrid.

A) Confocal laser scanning microscopy in tobacco leaf epidermis cells comparing the intracellular distribution of YFP-TM-ERD2::VSR2tail and the Golgi marker ST-RFP showing that the proteins co-localise. B) CLSM comparing the intracellular distribution of YFP-TM-ERD2::VSR2tail and the Golgi marker RFP-TM-ERD2 showing that the proteins co-localise and do not alter the distribution of each other. C) CLSM comparing the intracellular distribution of YFP-TM-ERD2::VSR2tail and the ER marker RFP-HDEL showing that the presence of a HDEL cargo strongly re-distribute the hybrid to the ER. Scale bars are 10µm.

To determine if the synthetic receptor can interfere with the localization of endogenous receptors, YFP-TM-ERD2::VSR2tail was co-expressed with RFP-TM-ERD2. Figure 22B shows that both markers co-localise with each other in punctate Golgi structures but RFP-TM-ERD2 was not seen in the ER. Therefore, there is no detectable evidence for interference and/or miss-localization by ERD2-ERD2 interactions.

To conclude the localization studies of YFP-TM-ERD2::VSR2tail I wanted to verify the identity of the weak ER stain observed. I thus co-infiltrated tobacco leaves with *Agrobacterium* harbouring YFP-TM-ERD2::VSR2tail together with a

strain that carries the classical ER maker RFP-HDEL (Gershlick et al. 2014). Figure 22C shows that upon co-expression with RFP-HDEL, the synthetic receptor still reached punctate structures but was much more retained in the ER network, where it co-localised with RFP-HDEL. This is sharp contrast to XFP-TM-ERD2 fusions with a native C-terminus, which did not re-distribute upon HDEL-cargo co-expression (Chapter 2).

3.2.5 A canonical KKXX motif can cause complete redistribution of chimeric ERD2 to the ER

To test if active Golgi to ER recycling is compatible with the function of the receptor in mediating accumulation of H/KDEL cargo in the ER lumen, I decided to provide the ERD2 C-terminus with a well characterised signal for dual transport between the ER and the Golgi apparatus. It was shown previously that a member of the P24 family of type I membrane spanning proteins found in COPI and COPII vesicles depends on two C-terminal cytosolic signals, a di-lysine -3/-4 (KKXX) motif for COPI-mediated recycling and a di-hydrophobic -7/-8 sequence potentially for COPII mediated ER export (Contreras et al. 2004). Due to the nature of the p24 C-terminus I replaced the last 9 C-terminally located amino acids of ERD2b by the 9 most C-terminal amino acids of p24a (AT1G21900). The new synthetic ERD2 construct (ERD2::p24tail) was cloned in the double expression vector, carrying a GUS expression marker, to be analysed via the bioassay. The same replacement was done on RFP-TM-ERD2 to be able to pursue microscopy studies, generating RFP-TM-ERD2::p24tail.

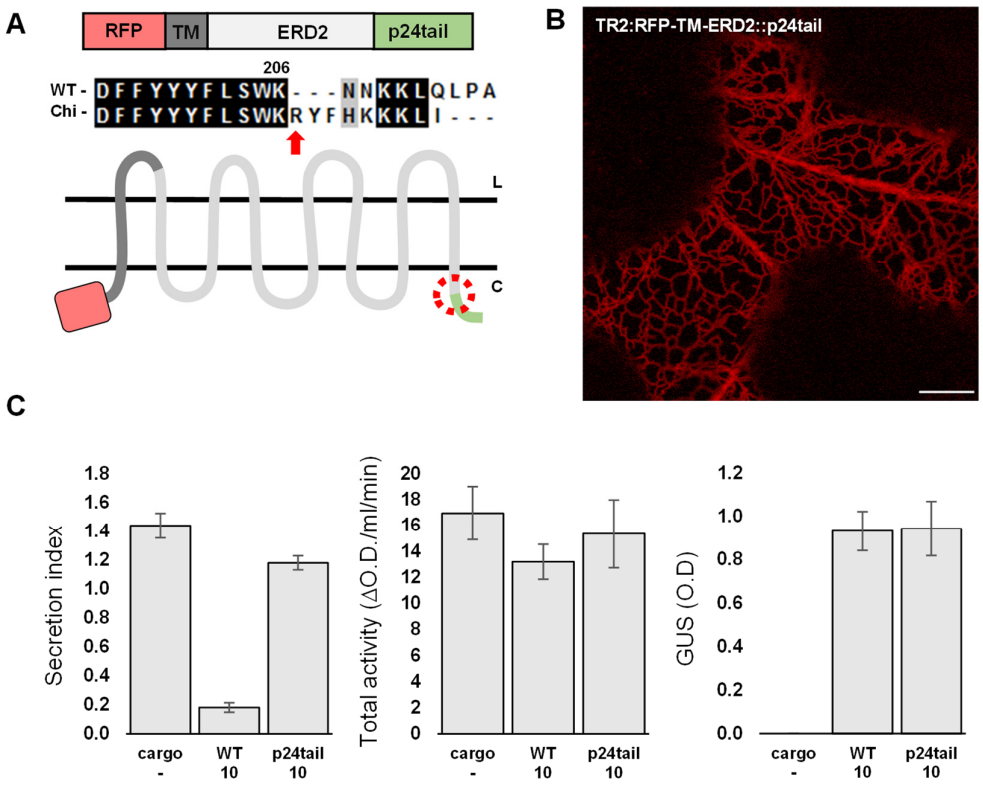


Figure 23 Synthetic biology approach adding a KKXX motif belonging to p24A to ERD2.

A) Illustration of the hybrid protein where the last 9 amino acids of ERD2 C-terminus were replaced by last 9 amino acids of p24a. The exact position where the replacement was made and the expected topology are also represented B) Confocal laser scanning microscopy in tobacco leaf epidermis cells showing the distribution of RFP-TM-ERD2::p24tail. Scale bar is 10µm. C) Co-expression of AmyHDEL cargo alone or combined with atERD2b (WT) and the hybrid protein (p24tail) in *Nicotiana benthamiana* protoplasts. 50 µg of AmyHDEL was co-transfected with variable µg of effector plasmid adjusted to yield equal GUS levels in pilot experiment. Total α-amylase activity is given arbitrary relative units (ΔO.D./ml/min). Total GUS O.D used as internal control marker for transfection efficiency. Error bars are standard deviations of three independent transfections.

Figure 23A schematically represents the new fluorescent protein and shows the exact position where the replacement was made. Figure 23B illustrates the localization pattern for RFP-TM-ERD2::P24tail. Interestingly, the fusion protein shows a complete ER staining with no evidence of Golgi bodies, which is in accordance to the published steady state localization of plant p24a (Langhans et al. 2008). This suggests that the signals embedded in the p24 tail were dominant over potential signals of the ERD2 core to achieve Golgi localisation. This could mean that the KKXX retrieval motif acts prior to mechanisms that allow arrival in the Golgi apparatus when the di-leucine signal is absent either by mutagenesis (Figure 11), deletion (Figure 15) or replacement (Figure 21).

Increased ERD2 recycling to the ER has been proposed to be a p24-mediated event caused to enhance retention of HDEL-protein in the ER (Montesinos et al. 2014). Therefore, I tested the ability of ERD2::p24tail to promote retention of HDEL cargo using the GUS-normalised biochemical assay. Figure 23C shows that the strong inhibition of AmyHDEL secretion seen for wild type ERD2 co-expression was all but abolished for ERD2::p24tail. Measurement of total AmyHDEL activity and GUS expression levels (Figure 23C) confirmed that the experimental conditions allowed a fair comparison with wild type ERD2.

In conclusion, re-designing ERD2 to become a typical COPI cargo and increasing the steady state levels in the ER did not enhance its ability to increase the cell retention of HDEL cargo as suggested before (Montesinos et al. 2014), but instead was found to abolish its biological function.

3.2.6 ER residency of ERD2::p24tail is not influenced by overexpression of TM-ERD2 or HDEL-proteins

Even though the overexpression of ERD2b::p24tail did not cause any detectable increase in HDEL-cargo retention (Figure 23C) the drastic change in the localization of the receptor (Figure 23B) justifies a careful analysis of the co-localization by co-infiltration in tobacco leaf with a Golgi marker, endogenous receptors and also to determine if the putative ER steady-state could be changed upon overexpression of ligands.

Results from Figure 24A clearly show that RFP-TM-ERD2::p24tail is ER retained and that there is absolutely no co-localization with the Golgi marker ST-YFP when co-expressed. Figure 24B shows that the ER-only distribution of RFP-TM-ERD2::p24tail and the Golgi-only distribution of YFP-TM-ERD2 are not influenced by co-expression of these two constructs and were unable to cause meaningful shifts in steady state levels. Lastly, Figure 24C shows that the co-infiltration of RFP-TM-ERD2::p24tail with YFP-HDEL does not mis-localize either of the two proteins and a perfect co-localization can be observed.

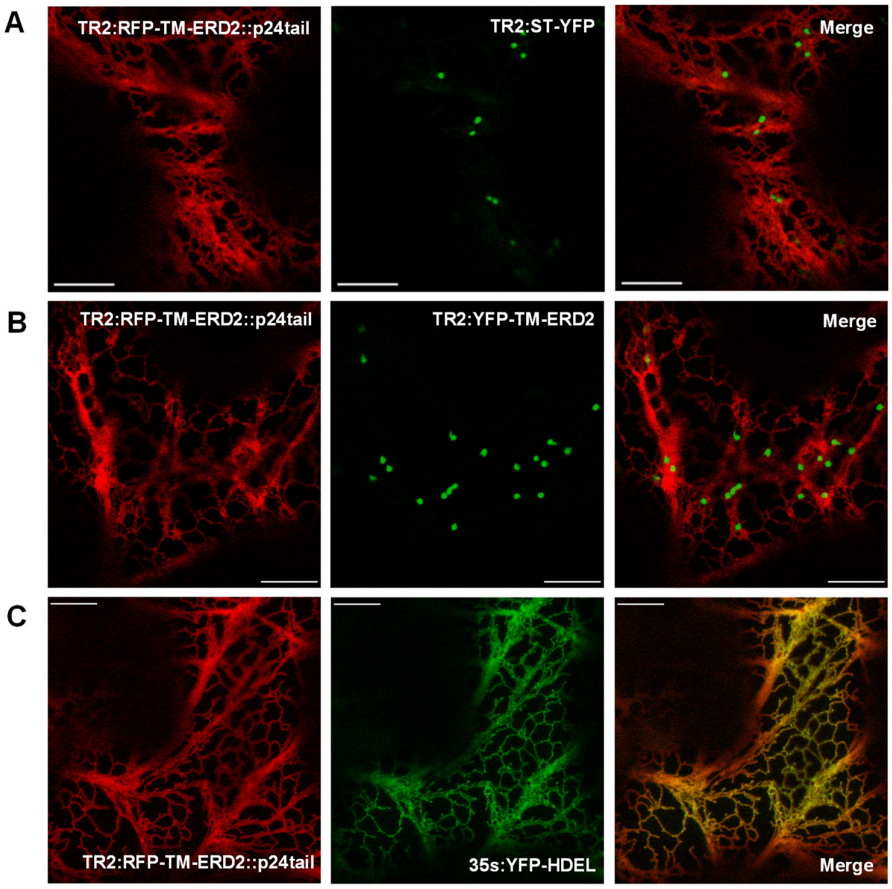


Figure 24 Co-localization studies of ERD2-p24 hybrid.

A) Confocal laser scanning microscopy in tobacco leaf epidermis cells comparing the intracellular distribution of RFP-TM-ERD2::p24tail and the Golgi marker ST-YFP showing that the proteins do not co-localise. B) Confocal laser scanning microscopy comparing the intracellular distribution of RFP-TM-ERD2::p24tail and the Golgi marker YFP-TM-ERD2 showing that the proteins do not co-localise and do not alter the distribution of each other. C) Confocal laser scanning microscopy comparing the intracellular distribution of RFP-TM-ERD2::p24tail and the ER marker YFP-HDEL showing that the proteins co-localise. Scale bars are 10µm.

3.2.7 Mutational analysis of the p24 KKXX motif reveals an unexpected function of the two aliphatic residues at the C-terminus and the bi-partite structure of this signal

The signalling of p24 proteins has been investigated in the past and, as mentioned before, two separate signals work collectively permitting the interaction with COPI and COPII coats, consequently allowing these proteins to participate in retrograde and anterograde transport (for review, Strating & Martens 2009). p24δ5 mutational and deletion analysis has shown that the C-

terminal di-lysine motif, KKXX, is necessary for its ER steady-state (Montesinos et al. 2012).

To further understand how the addition of a p24 tail to ERD2 core could be interfering and controlling the protein localisation I decided to verify previous observations associated with the properties of the canonical KKXX motif using infiltrated tobacco leaves in combination with CLSM. For that purpose, I created secRFP-p24aTM by isolating p24δ5 transmembrane domain (TMD) and its C-terminus and directly fusing it to a fluorophore which bears a signal peptide (secRFP), to allow the correct insertion of the new protein in the ER membrane in the same way as published before (Langhans et al. 2008).

Figure 25A shows that there is no co-localization when secRFP-p24aTM, under the control of the 35S promoter, is co-infiltrated with ST-YFP in tobacco leaves and that secRFP-p24aTM is exclusively ER resident, as reported before (Langhans et al. 2008). Figure 25B shows that the double-mutation of lysines into serines (secRFP-p24aTM KKSS) causes, as expected, the total redistribution of the receptor which now co-localizes with ST-YFP and also stains extra punctate structures (black arrow heads). So far, results obtained when analysing secRFP-p24aTM are in accordance with the literature.

Earlier experiments on the di-hydrophobic -7/-8 sequence in p24 did not suggest a strong influence on the targeting of a secRFP-p24aTM reporter (Langhans et al. 2008). However, I noticed that many members of the p24 family carry two aliphatic residues subsequently to the di-lysine, for instance KKLI in p24δ5. For that reason, secRFP-p24aTM was once again mutated and both aliphatic residues were replaced by glycine (secRFP-p24aTM-LIGG). Surprisingly, Figure 25C shows that the modification of these amino acids compromised the ER-only localisation of secRFP-p24aTM and now it is present in a dual ER-Golgi configuration, as demonstrated by the co-localization with ST-YFP. This suggests that the common annotation of a KKXX motif may be erroneous because at least in the case of p24δ5 the two amino acids following the di-lysine motif are not totally random, and thus the annotation XX is misleading.

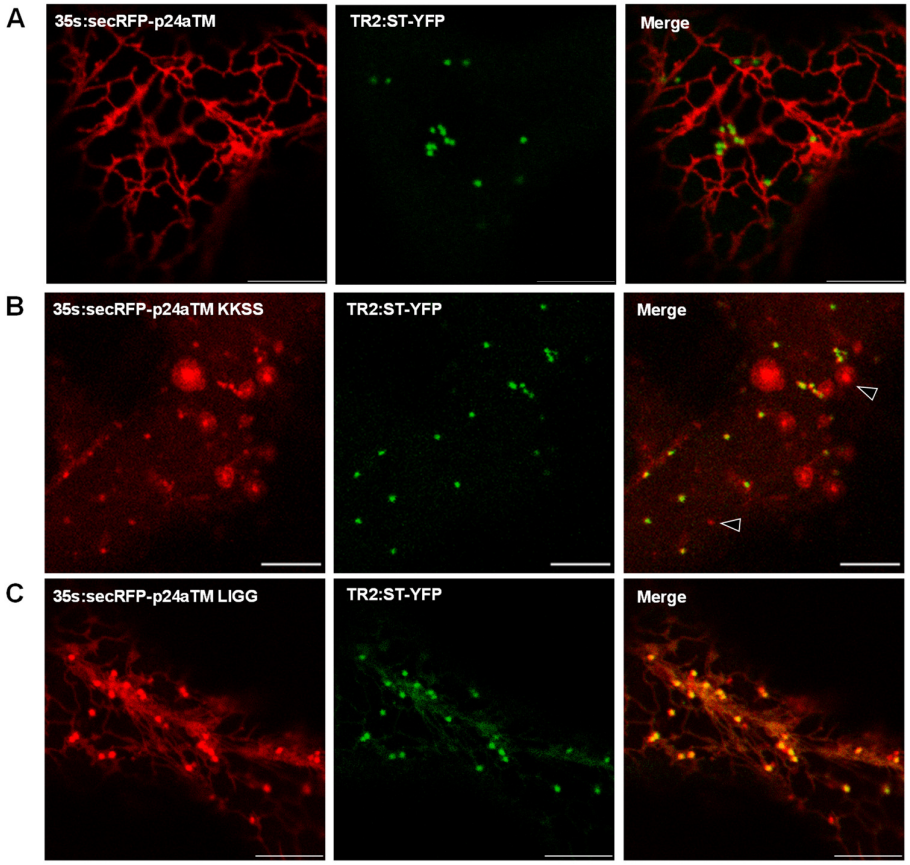


Figure 25 Studying the canonical KKXX motif of plant p24 proteins via point-mutagenesis and co-localization.

A) Confocal laser scanning microscopy in tobacco leaf epidermis cells comparing the intracellular distribution of the isolated p24δ5 transmembrane domain (TMD) and its C-terminus (secRFP-p24aTM) and the Golgi marker ST-YFP showing that the proteins do not co-localise and the secRFP-p24aTM is in the ER. B) CSLM comparing the intracellular distribution of secRFP-p24aTM KKSS and the Golgi marker ST-YFP showing that the ER localization was lost and now the proteins partially co-localise but secRFP-p24aTM KKSS is also seen in other organelles (black arrow-head). C) CSLM comparing the intracellular distribution of secRFP-p24aTM LIGG and the Golgi marker ST-YFP showing that the ER localization was compromised and now the proteins partially co-localise. Scale bars are 10µm.

3.2.8 The ERD2 core itself could carry sorting information

The addition of p24a tail to ERD2 core was shown to overrule the original receptor targeting (Figure 23) and I have also shown that the mutation of the p24a C-terminus can alter its signalling (Figure 25). These results together form the basis to explore if ERD2 core has any control on the protein localization. Thereby, the double-lysine modification (KKSS) and double-aliphatic modification (LIGG) were imposed on to the RFP-TM-ERD2::p24tail construct

and the receptor localization was once again verified by co-infiltration with the Golgi marker ST in tobacco leaves via CLSM.

Figure 26A shows that without the di-lysine motif the receptor is prone to leave the ER, similarly to what was observed in Figure 25B, but RFP-TM-ERD2::p24tail KKSS is perfectly localized to the Golgi apparatus and there is no evidence of either ER staining nor post-Golgi organelles. Thus, this result suggests that the structural core of ERD2 alone can contribute to a Golgi localisation, even though its C-terminus has been completely replaced.

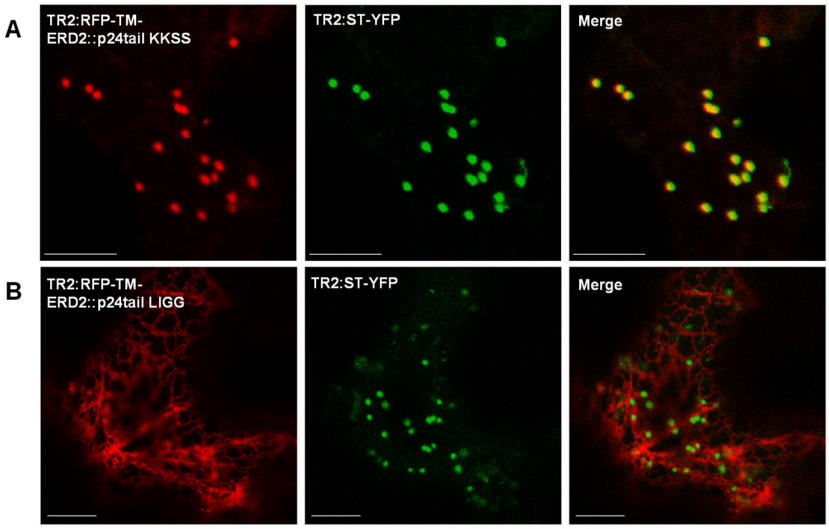


Figure 26 The presence of ERD2 core alters the effect of mutations on p24 KKXX motif.

A) Confocal laser scanning microscopy in tobacco leaf epidermis cells comparing the intracellular distribution of the double serine mutant (KKSS) version of RFP-TM-ERD2::p24tail and the Golgi marker ST-YFP showing that the proteins co-localise B) CSLM comparing the intracellular distribution of the double glycine mutant (LIGG) version of RFP-TM-ERD2::p24tail and the Golgi marker ST-YFP showing that the proteins do not co-localise. Scale bars are 10µm.

On the other hand, Figure 26B surprisingly shows that repeating the previous double-aliphatic replacement (LIGG) and generating RFP-TM-ERD2::p24tail LIGG does not cause any change in the localization of the fluorescent molecule. RFP-TM-ERD2::p24tail LIGG behaves exactly as the non-mutated one (Figure 24) and does not follow the pattern observed in Figure 25C. So here it appears as if the KKXX annotation is applicable and the two aliphatic resides can be replaced by glycines.

In summary, the functioning of the KKLI signal of p24 differs depending on how it is displayed, either as part of a single TM span type I membrane protein with a

cytosolic C-terminus, or as part of a 7 TM multiple membrane span protein with a cytosolic C-terminus. A further analysis of the differential roles of the two aliphatic residues was beyond the scope of this thesis, but the dramatic influence of the di-lysine motif on mediating ER retention strongly confirms that the C-terminus of ERD2 is cytosolic, as recently demonstrated (Silva-Alvim et al. 2018), and that a di-lysine motif takes precedence over the Golgi localizing function of the ERD2 core.

3.3 Discussion

Previous results indicated that a conserved di-leucine found at ERD2 C-terminus was involved with the receptor function and Golgi localisation. The objective of this chapter was to further characterize this potential signal motif and assess if was sufficient to mediate Golgi retention of the receptor. I also wanted to use synthetic biology to study to what extend the trafficking of the receptor could be manipulated via the addition of other classical signal motifs from membrane spanning proteins. The results provided further insight into the biological significance of ERD2's Golgi residency and how it is maintained in this position. The work also provided further insights into well-established COPI-dependent sorting signals which are discussed below.

3.3.1 Golgi retention of ERD2 is not caused by one specific motif but instead by a combination of factors.

Faster recovery of mutant YFP-TM-ERD2 LLGG fluorescence in the Golgi (Figure 19) is an important key-result because it establishes the LXLP signal as a Golgi retention signal, rather than an ER export signal. I believe that the most reasonable explanation is that the LLGG mutation leads to faster leakage of ERD2 from the Golgi apparatus back to the ER, which therefore creates a larger pool of ERD2 proteins that can replenish the levels in the Golgi apparatus via anterograde transport. In the case of the wild type, recovery is most likely dependent on *de novo* synthesis, which is probably very slow for a stable protein with low turnover. I propose that the di-leucine motif is therefore a Golgi retention signal that actively prevents retrograde transport back to the ER.

Further experiments (Figure 20) have shown that the transplantation of the motif to a type I membrane spanning protein capable of reaching post-Golgi compartments (Jiang & Rogers 1998; Foresti et al. 2010) does not promote Golgi

retention by itself, and it is therefore not sufficient for its function. The approach used in Figure 20B is similar to a strategy previously used in the literature, in which the C-terminal domain of the mammalian KDEL receptor replaced the last 41 amino acids of a chemokine receptor (CXCR4), a plasma membrane protein. This was done in order to investigate the targeting properties of this domain and has shown that the chimera was fully capable to reach the plasma membrane, concluding that the C-terminus of the KDEL receptor alone does not convey Golgi residency (Cabrera et al. 2003), but may depend on the rest of the ERD2 core structure.

3.3.2 Re-engineering ERD2 C-terminus reveals that the receptor cannot be easily targeted to the late compartments of the secretory pathway.

The addition of a PVC/LPVC targeting motif was not able to remove the chimeric receptor from the Golgi apparatus as shown in Figure 22. This observation is interesting and could indicate that ERD2 does not reach the later stacks of the Golgi, where VSRs targeting has been proposed to happen. Perhaps ERD2-core plays a unveiled role mediating the *cis-trans* Golgi partitioning observed before (Figure 13B and Silva-Alvim et al. 2018).

The apparent ligand-induced re-distribution of the ERD2 hybrid YFP-TM-ERD2::VSR2tail was very surprising because it was not observed with the wild type construct YFP-TM-ERD2 (Chapter 2, Figure 12). One possible explanation is that the hybrid is capable of binding ligands but can no longer release the ligands, thus assuming a higher ER steady state level. In Figure 22C is noticeable that in the presence of high levels of HDEL cargo YFP-TM-ERD2::VSR2tail had its partial ER localisation enhanced.

Fundamentally it is important to consider that this can be an indirect effect caused by the saturation of the trafficking machinery and consequently increased retrograde transport or decrease of ER export, not necessarily this is a real re-distribution of the synthetic receptor caused by, for instance, binding to ligands. Further experiments are necessary to distinguish between real re-distribution combined to poorly ligand-release from other possible causes. Apart from that, it would be interesting to check if the presence of vacuolar cargo could influence the localization of the receptor, even though the VSR2 C-terminus should not bear any binding properties.

3.3.3 A potential new KKΨΨ bi-partite signal is capable of mediating ER-containment of ERD2.

The observation that the two aliphatic (Ψ, Aasland et al. 2002) residues can influence the localization of p24 proteins (Figure 25) is innovative and was not observed before. This result shows that the signal motif in the C-terminus region of p24s is potentially bi-partite and that it possibly does not operate as the classical KKXX motif from other proteins, where the last two amino acids were shown to be irrelevant (Jackson et al. 1990). In the case of p24, XX cannot be GG, but further mutagenesis studies should help to elucidate permissive amino acids in the -1-2 position.

Combined results from Figures 25 and 26 are an indicative that the KKΨΨ motif found in family members of plant p24s could, differentially, control the exit ratio of these proteins from the ER. As consequence KKSS mutants could be exiting the ER much faster than LIGG mutants, leading to a post-Golgi localization combined with stronger ER retention observed when comparing Figure 25B/C and Figure 26A/B, respectively. To confirm this hypothesis, it would be interesting in the future to compare the recovery ratio of these proteins using FRAP and also consider analysing the modifications acquired by these proteins if an alternative tag was used, but these experiments were beyond the scope of this thesis.

Finally, It has been shown in the literature that both *N. benthamiana* ERD2 isoforms (ERD2a and ERD2b) have potential homotypic interactions and also that these molecules interact with other proteins related to COPI formation (Xu & Liu 2012). Results shown in Figures 22 and 24 suggests that both YFP-TM-ERD2::VSR2tail and RFP-TM-ERD2::p24tail do not have the capacity to interfere with the subcellular localization of XFP-TM-ERD2 fusions. These observations show that either the ERD2 C-terminus is mandatory for a putative interaction or that ERD2 homo/hetero-oligomerization is only achieved in specific circumstances. If the former is true, and an unobstructed C-terminus is necessary for a meaningful interaction, I believe that any experiments conducted to show any synergy need to be executed using biologically active ERD2 molecules, which is in discordance to what has been done previously using non-functional C-terminal fusions (Xu & Liu 2012; Aoe et al. 1997; Majoul et al. 2001).

3.4 **Conclusions**

Combined the data presented in this chapter confirm that ERD2 must reside in the Golgi apparatus in order to carry out efficient HDEL-cargo retention and that COPI-mediated recycling to the ER is probably not involved in ERD2 functions. This is an important discovery, as it questions the central dogma surrounding the mechanism of ER retention of soluble proteins.

Chapter 4 **The ERD2 gene product is remarkably conserved amongst eukaryotic organisms.**

4.1 **Introduction**

Phylogenomic studies debating the parenthood and relationship of Eukaryotes are constantly been published and the tree of life is regularly changing. Nevertheless, publications from the last few years have firmly established the existence of at least five, potentially eight, major supergroups. Figure 27A shows a simplified representation of the tree of life based on the most recent information available in the literature (Klinger et al. 2016; Burki et al. 2016; Burki 2014; Cavalier-Smith et al. 2014).

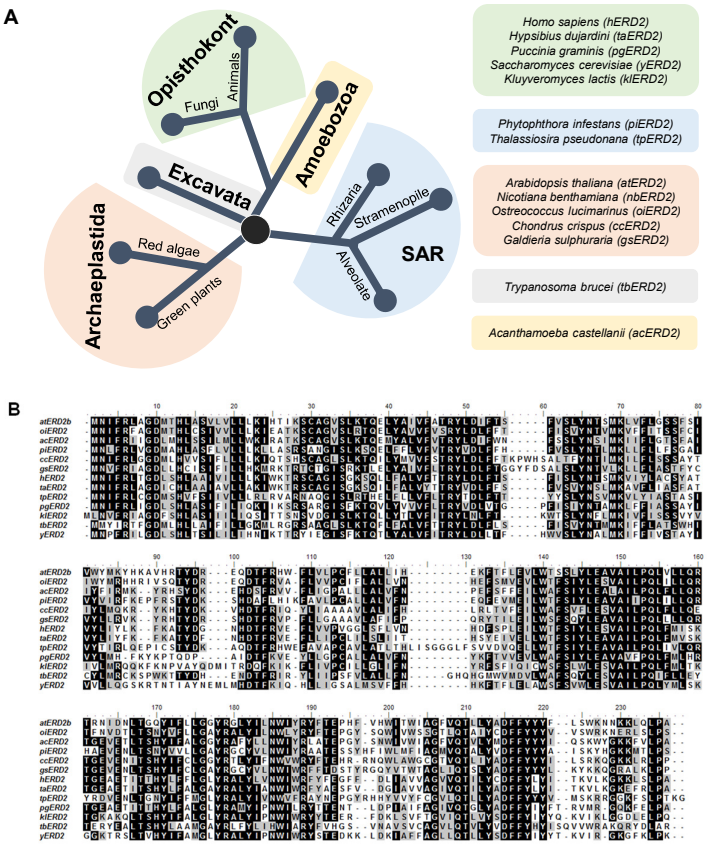


Figure 27 ERD2 gene product is highly conserved amongst eukaryotes.

A) Left panel shows a simplified three of life reconstruction based on recent studies (Klinger et al. 2016; Burki et al. 2016; Burki 2014; Cavalier-Smith et al. 2014). All major eukaryotic groups are represented. At least one representative organism was selected per major branch and are listed in right panel, those were used posteriorly for ERD2 protein homology analysis B) ClustalW2 multiple sequences alignment of ERD2 gene product from 13 organisms using BioEdit software. Identical amino acids are highlight in black and similar amino acids are highlighted in grey.

The secretory pathway, particularly the ER-Golgi network, is an ancient feature of Eukaryotes with conserved fundamental steps and common core of proteins, such as SNAREs and coat protein complexes (Dacks & Field 2007; Dacks & Doolittle 2001). In contrast to mitochondria and chloroplasts which have been acquired by primary or secondary endosymbiosis, the secretory pathway is hardwired in the eukaryotic cells and as such more suitable to study phylogeny of eukaryotes (Doolittle 1981). The machinery controlling the necessary steps for protein maturation, transport and secretion has been dissected by many studies since one central investigation conducted by George Palade (Palade 1975).

The lumen of the ER is filled with soluble proteins that ensure the correct folding of nascent polypeptides, such as the many *hsps* (De Maio 1999). These ER lumenal proteins can escape via *bulk-flow* to the Golgi apparatus, some of which are thought to be quickly recognised and recovered upon arrival (Wieland et al. 1987; Pelham et al. 1988). Retrograde transport of ER lumenal proteins is widely believed to be mediated by an active receptor called ERD2 which recognizes a conserved K/HDEL motif (Semenza et al. 1990). However, results in Chapters 2 and 3 question the suggested recycling of ERD2 back to the ER. ERD2 from animals, plants and fungi have been cloned, expressed and studied in the past (Semenza et al. 1990; Tang et al. 1993; Lewis & Pelham 1990). Due to the highly conserved nature of the Golgi-ER interface major players, such ERD2, are plausibly conserved as well. However, it is entirely possible that predominant Golgi-residency is a plant-specific phenomenon and ERD2 molecules in other organisms, such as humans, may operate in a different manner.

To verify how conserved ERD2 is amongst eukaryotes at least one organism of each supergroup was selected, as it can be seen in Figure 27A (second panel), using sequences from online databases (NCBI BLAST - Altschul et al. 1997; SOL Genomics - Mueller et al. 1992) and by a multiple alignment analysis using BioEdit (Hall 1999), I found that at least one gene encoding for an ERD2 orthologue protein is found per organism. Figure 27B shows a multiple alignment of these sequences and reveals the high homology level, with clear highly conserved regions. This observation is another indication that the biological function of this protein is likely to be conserved.

Sequential studies after the original identification and isolation of ERD2 revealed the high homology of the gene in different organisms when compared to *S. cerevisiae* ERD2, 50% identity for the human ERD2 (Lewis & Pelham 1990) and 59% for *K. lactis* (Lewis et al. 1990). Interestingly, researchers have unsuccessfully tried to complement the lethal phenotype caused by yeast ERD2 deletions using the human ERD2 (Lewis & Pelham 1990). However, after isolation and expression of *A. thaliana* ERD2 (atERD2) it was shown that the plant ERD2 was capable of complementing the same yeast deletions (Lee et al. 1993).

These results show that even though ERD2 is extremely conserved amongst eukaryotes some level of divergence may be present. Similarly to the strategy used before in Figure 10A (Chapter 2), the high structural conservation of ERD2 (Figure 27B) makes domain-swap an attractive approach to identify crucial regions and possibly important amino acids contained in these regions. For this reason, I aimed to identify ERD2 orthologs that would potentially fail to function in the gain of function assay described in Chapter 1.

The efficiency and accuracy of the bioassay previously used in this thesis (Chapter 1, Figure 6) allows for it to be used with the purpose to investigate ERD2 evolutionary conservation and signal-specific amongst eukaryotes via ectopic overexpression of heterologous ERD2s in plant protoplast. In this chapter I present functional data on ERD2 orthologues from 13 different organisms including all major eukaryotic supergroups, suggesting that ERD2 is a remarkably conserved protein possibly controlling the most ancient of transport steps in the secretory pathway

4.2 Results

4.2.1 Cloning of *Saccharomyces cerevisiae* ERD2 in *E. coli* plasmids is troublesome

The first step to start ERD2 evolutionary conservation analysis was to clone the coding region of gene synthesised ERD2 orthologs into the previously used double-expression vector containing GUS as internal marker for direct comparison. For that purpose, two specific restriction sites (ClaI and BamHI) were selected to be flanking all sequences, and the coding regions for the various ERD2 orthologues were obtained by custom gene synthesis. This included

Ostreococcus lucimarinus (oiERD2), *Acanthamoeba castellanii* (acERD2), *Phytophthora infestans* (piERD2), *Chondrus crispus* (ccERD2), *Galdieria sulphuraria* (gsERD2), *Homo sapiens* (hERD2), *Hypsibius dujardini (*Tardigrade - taERD2), *Thalassiosira pseudonana* (diatom - tpERD2), *Puccinia graminis* (pgERD2), *Kluyveromyces lactis* (klERD2), *Trypanosoma brucei* (tbERD2) and *Saccharomyces cerevisiae* (yERD2). If necessary, gene synthesis involved the removal of internal restriction sites that would interfere with the planned sub cloning strategy and further subsequent cloning steps. Engineered coding sequences are shown in appendix 1.

All orthologue sequences were removed from the gene synthesis vector as ClaI-BamHI fragments and with only one exception they were easy to handle and did not lead to problems during sub-cloning and subsequent production of concentrated plasmid DNA preparations. The exception was the *Saccharomyces cerevisiae* ERD2 (yERD2) coding sequence, which resulted in mixed population of colonies growing after transforming the ligation mixture, including distinctly small (white arrow head) amongst fewer normal sized (black arrow head) colonies (Figure 28A).

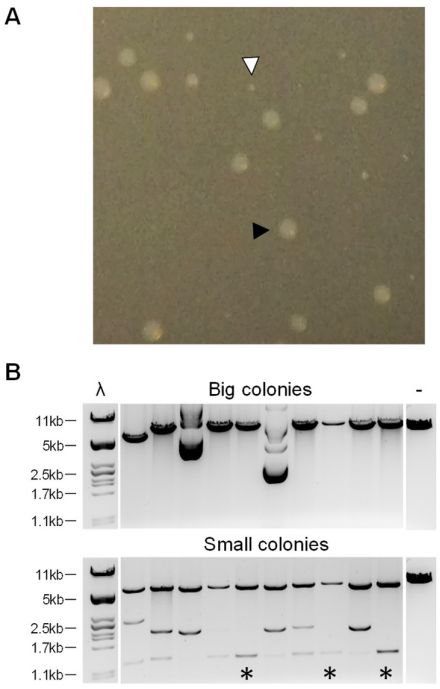


Figure 28 Screening for positive insertion of *S. cerevisiae* ERD2 coding region into an expression vector.

A) Example of dual-sized population of *E. coli* colonies observed in plates, large (black arrow-head) and small (white arrow-head). B) Upper panel is a qualitative digestion of 10 DNA preparations originated from big colonies using EcoRI and NcoI restriction enzymes and comparing to parent plasmid (negative control, last lane). In bottom panel the same qualitative digest was repeated with 10 DNA preparations originated from small colonies and compared to parent plasmid (negative control, last lane). Plasmids containing yERD2 were expected to yield two separate fragments (1.3kb and 5.9kb, represented by an asterisk), whilst the parent plasmid would yield a single fragment (7.2kb) for the same digest.

A dual population was unusual but experienced before in the host laboratory when weak expression of recombinant genes in *E. coli* had a detrimental effect on clone stability and bacterial growth (An, personal communication). For that

reason, 10 DNA preparations were cultured for both classes of colony and recombinants were tested for the presence of the desired insert. Plasmids containing yERD2 were expected to yield two separate fragments (1.3kb and 5.9kb), whilst the parent plasmid would yield a single fragment (7.2kb) for the same digest. Figure 28B (upper panel) shows that DNA preparations originated from large colonies did not contain the desired insertion. Only a minor set of preparations from small colonies yielded the expected fragment (asterisks in bottom panel). However, upon sequencing of those positive DNA preparations from small colonies it became evident that, without exception, ERD2 coding region always contained random point mutations. After different attempts using different *E. coli* strains, it was ultimately found that if *E. coli* was grown at lower temperatures (22^{o}C) for a prolonged period of 3 days, colonies of the same size were obtained and clones with the expected insert size were obtained that showed the expected wild type yERD2 sequence. Thus, for the purpose of this thesis all cultures of *E. coli* transformed with a plasmid containing the coding region of an ERD2 orthologue derived from yeasts was incubated at 22^{o}C for 3 days, including all liquid cultures for maxiprep. In addition, sequencing was carried out each time when a new DNA preparation was made to verify that the sequence was not mutated.

4.2.2 Functional conservation of ERD2 between plants and other eukaryotes except yeasts

After successful gene synthesis and sub cloning of ERD2 orthologs from 13 different organism into the GUS reference vector, protein transport assays using AmyHDEL as cargo were carried out in either *Nicotiana tabacum* or *Nicotiana benthamiana* protoplasts to assess the ability to mediate increased cell retention of the reporter enzyme *in vivo*.

Nicotiana benthamiana protoplast were transformed via electroporation and incubated for 24 hours for later harvesting of medium and cell fractions, as before. After GUS normalization, Figure 29A shows that a minimum degree of homology is critically important for ERD2 to be able to promote increased HDEL-retention in the cells compared to the partial secretion of cargo alone (con, first lane). As it can be seen, the vast majority of ERD2s were efficiently capable of reducing the SI to levels comparable to those obtained with *Arabidopsis thaliana* ERD2 (atERD2 - second lane). Nevertheless, after crossing a threshold of 50%

homology with atERD2, overexpression of *K. lactis* and *T. brucei* ERD2 were not efficiently capable to suppress the loss-of-function effect (lanes 12 and 13). Most significant was the effect observed by the overexpression of *S. cerevisiae* ERD2 (yERD2 – last lane), which was found to cause an opposite effect leading to the induced secretion of AmyHDEL molecules.

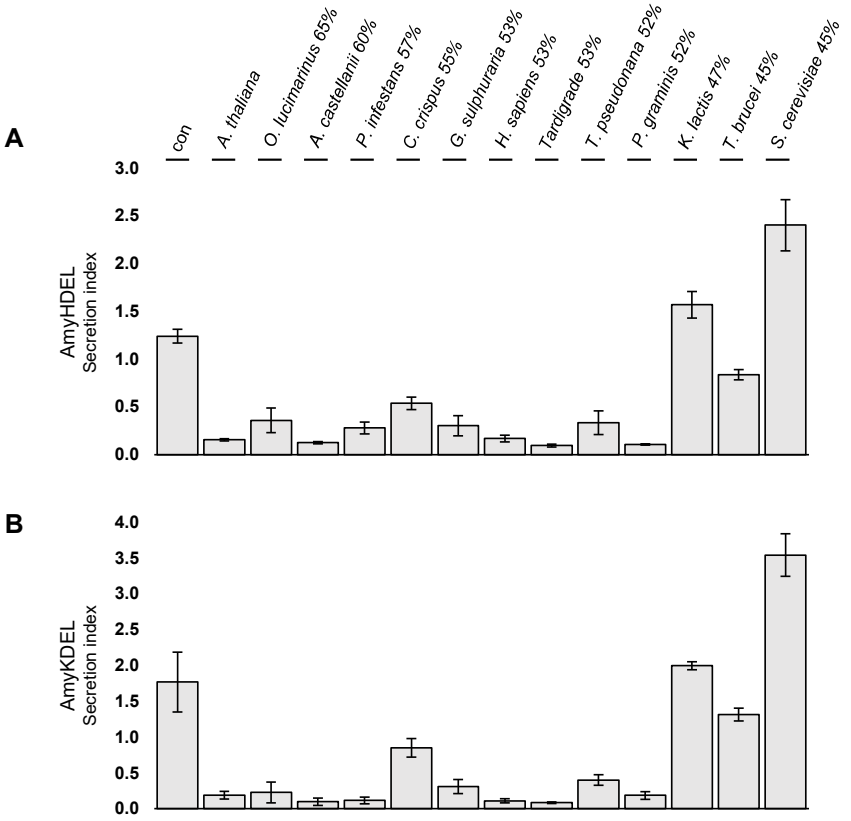


Figure 29 ERD2 biological function is highly conserved amongst eukaryotes and it is not signal-specific.

A) Co-expression of AmyHDEL cargo alone or combined with ERD2 from 13 organisms, including *Arabidopsis thaliana* (atERD2b) *Ostreococcus lucimarinus* (oiERD2), *Acanthamoeba castellanii* (acERD2), *Phytophthora infestans* (piERD2), *Chondrus crispus* (ccERD2), *Galdieria sulphuraria* (gsERD2), *Homo sapiens* (hERD2), *Hypsibius dujardini* (Tardigrade - taERD2), *Thalassiosira pseudonana* (tpERD2), *Puccinia graminis* (pgERD2), *Kluyveromyces lactis* (klERD2), *Trypanosoma brucei* (tbERD2) and *Saccharomyces cerevisiae* (yERD2) in *Nicotiana benthamiana* protoplasts. The homology level of the proteins compared to the plant ERD2 is shown in percentage on top of each lane. 50 µg of AmyHDEL was co-transfected with variable µg of effector plasmid, adjusted to yield equal GUS levels in pilot experiment.. B) Same experimental condition as panel A, but AmyHDEL was replaced by AmyKDEL. Error bars are standard deviations of three independent transfections

4.2.3 KDEL and HDEL are interchangeable in eukaryotes

ERD2 signal-specificity was shown by the observation that the receptor from *S. cerevisiae* does not efficiently reduce the secretion of invertase bearing different signal motifs, such as the one found into *K. lactis* BiP, DDEL (Lewis et al. 1990).

Nonetheless, ERD2 signal-specificity was challenged when researchers tested the retention capacities of multiple signal motifs in plants (Denecke et al. 1992).

To verify how conserved the use of different signal motifs is amongst eukaryotes four different ER lumenal proteins were selected and their respective orthologues sequences were blasted. Proteins selected were Binding immunoglobulin protein (BiP), Endoplasmin, Calreticulin and Protein disulfide isomerase (PDI). Table I shows the remarkable signal diversity amongst eukaryotes, whereas at least nine different signals can be found. Table I also shows the low diversity in specific organisms, such as *S. cerevisiae*.

Protein	Organism	Motif	Accession #
BiP	*A. thaliana*	DESHDEL	AT5G28540.1
	H. sapiens	TAEKDEL	BAI45446.1
	S. cerevisiae	YFEHDEL	UniProtKB A6ZPU2
	O. lucimarinus	VDEHDEL	XP_001416181.1
	A. castellanii	DYVADEL	XP_004334973.1
	P. infestans	DDEHDEL	XP_002909118.1
	G. sulphuraria	DEGHDEL	XP_005707456.1
	C. crispus	DDSHDEL	XP_005718499.1
	C. crispus	YQGKDEL	XP_005714152.1
	T. pseudonana	VGFDDEL	XP_002295019.1
	T. pseudonana	DFGDDEL	XP_002288567.1
	P. graminis	LGSHDEL	XP_003323708.1
	k. lactis	DYFDDEL	X54709.1
	T. brucei	PQPMDDL	XP_011780014.1
Endoplasmin	*A. thaliana*	ENTKDEL	AT4G24190.1
	H. sapiens	TAEKDEL	AAH66656.1
	O. lucimarinus	EAPKDEL	XP_001421809.1
	O. tauri	EPPKDEL	UniProtKB A0A096P903
	A. castellanii	DDEHDEL	XP_004339231.1
	P. infestans	VDEKDEL	XP_002999135.1
	G. sulphuraria	TVDHEEL	XP_005704163.1
	C. crispus	SEDHDEL	XP_005715704.1
Calreticulin	*A. thaliana*	DAAHDEL	AT1G56340.1
	H. sapiens	GQAKDEL	BAG70222.1
	O. tauri	VPVRDEL	UniProtKB A0A090LYV3
	O. tauri	IPNRDEL	UniProtKB Q01BQ5
	A. castellanii	GHDHEDL	XP_004349632.1
	P. infestans	KTEKDEL	XP_002997040.1
	T. pseudonana	QKKDREE	XP_002291609.1
	Trypanosoma brucei	KEDKSDL	XP_847570.1
PDI	*A. thaliana*	TAAKDEL	AT1G77510.1
	H. sapiens	DLGKDEL	UniProtKB Q15084
	H. sapiens	SRTKEEL	UniProtKB A0A090N8Y2
	S. cerevisiae	DAIHDEL	YCL043C
	O. tauri	RTPHMEL	UniProtKB Q018Z4
	A. castellanii	DPEHDEL	XP_004334355.1
	P. infestans	QKEHEEL	XP_002895135.1
	G. sulphuraria	HIQKEEL	XP_005704356.1
	C. crispus	EADKEEL	XP_005711820.1
	C. crispus	EEEKEEL	XP_005712420.1
	k. lactis	ELEQDEL	XP_452244.1
	T. brucei	NVDKQDL	XP_011778178.1

Organism	KDEL	HDEL	Other
A. thaliana	2	2	
H. sapiens	4		1
S. cerevisiae		2	
Ostreococcus sp.	1	1	3
A. castellanii		3	1
P. infestans	2	1	1
G. sulphuraria		1	2
C. crispus	1	2	2
T. pseudonana			3
P. graminis	1		
K. lactis			2
T. brucei			3

Table 1 Signal diversity of ER residents amongst eukaryotes.

Sequence of four conserved ER lumenal proteins were blasted and used for this analysis, it includes Binding immunoglobulin protein (BiP), Endoplasmin, Calreticulin and Protein disulfide isomerase (PDI).

The low diversity of ER retention signal motifs in *S. cerevisiae*, mostly KDEL (Hadlington & Denecke 2000), If other organisms have high preference for certain sequence-variants, a single cargo molecule may not be sufficient to study functional conservation of ERD2. Nevertheless the results with *S. cerevisiae* were unexpected because plant ERD2 was reported to complement yeast ERD2 (Lee et al. 1993). Therefore, an alternative cargo molecule bearing a KDEL signal (AmyKDEL) was used to repeat the comparison of ERD2 orthologs.

Figure 29B shows that, independently of the signal motif used, the suppression of amylase secretion previously observed follows the same trend and only minor changes are observed, for instance *O. lucimarinus* and *P. infestans* show a slightly higher affinity for AmyKDEL. The most important observation is that the overexpression of yERD2 again did not supress the secretion of AmyKDEL and, as before, it caused the induced secretion of the amylase reporter.

Results presented in Figure 29 on the overexpression of exogenous ERD2 in plant systems shows that the receptors can be classified into functional receptors, partially functional receptor, non-functional receptor and an unexpected gain-of-function phenotype characterized by the induced secretion of K/HDEL cargo.

4.2.4 *Saccharomyces cerevisiae* ERD2 induces increased secretion of HDEL proteins in *Nicotiana tabacum*

To better characterize the induced secretion effect caused by the overexpression of *S. cerevisiae* ERD2 I decided to shift back to *Nicotiana tabacum* protoplasts due to the higher chances of seeing a dose-response curve, as discussed before in Chapter 1 (Figure 6C).

After GUS normalization in a pilot assay, a constant amount of plasmid containing AmyHDEL was delivered in combination with increasing amount of effector plasmids (1µl, 3 µl, 5µl or 10µl) containing either *A. thaliana* (atERD2) or *S. cerevisiae* (yERD2) ERD2 via electroporation to *N. tabacum* protoplasts and incubated for 24 hours.

Figure 30 shows that the induced-secretion effect is dose dependant and increasing delivery of yERD2 plasmids has the exact opposite effect of the increasing suppression of AmyHDEL secretion characteristically of atERD2. This observation indicates that the presence of yERD2 interferes with the endogenous ER retention mechanism. The results suggest that yERD2 cannot be totally non-functional in plants and may retain some ability to interact with at least part of the sorting machinery, resulting in the semi-dominant effect which is dosage dependent.

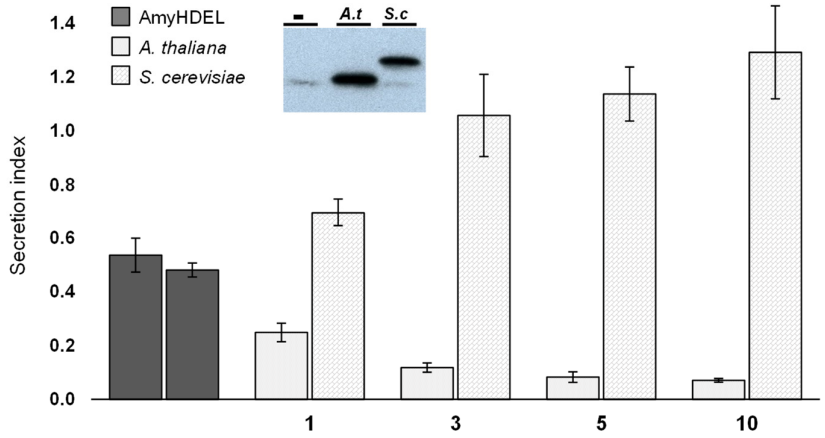


Figure 30 Induced-secretion caused by overexpression of yERD2 is dose-dependent.

Co-expression of the AmyHDEL cargo alone (dark grey) or with either atERD2b (dotted pattern) or yERD2 (brick pattern) in *Nicotiana tabacum* protoplast. 50 µg of AmyHDEL was co-transfected with increasing amounts of effector plasmids given below each lane in µg. Total GUS O.D used as internal control marker for transfection efficiency in all experiments (data not shown). Error bars are standard deviations of three independent transfections. Insert shows immune blot against HA-tag (YPYDVPDYA) for same GUS levels.

As an alternative control to the GUS internal marker an HA-tag (YPYDVPDYA) was added to the C-terminus of both atERD2 and yERD2. Expression levels of both ERD2s was checked via immunoblotting, probing with anti-HA. Figure 30 (insert) shows that for the same GUS levels loaded on the gel, yERD2 and atERD2 show the same expression levels in *N. tabacum* protoplasts.

4.2.5 Induced secretion of HDEL cargo caused by yeast ERD2 can be suppressed by extra levels of plant ERD2

ERD2 self-interaction and oligomerization was shown before, both in animals and plants, as an important feature mediating the interaction of the receptor with other proteins (Aoe et al. 1997; Majoul et al. 2001; Xu & Liu 2012). Such important observation should be taken into consideration when evaluating the surprising effect caused by the overexpression of yERD2.

To verify if the induced secretion effect was a consequence of yERD2 interaction with atERD2 I have decided to ectopically express incremental levels of atERD2 in a yERD2 background.

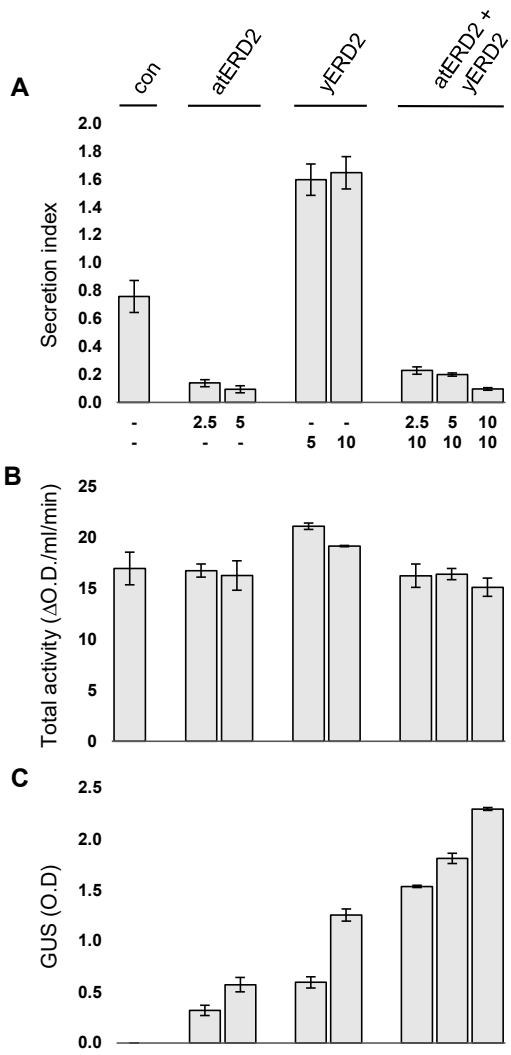


Figure 31 Induced-secretion caused by overexpression of yERD2 is suppressible.

A) Co-expression of the AmyHDEL cargo alone or with effector plasmids in *Nicotiana benthamiana* protoplast. Cargo was co-expressed with either atERD2b alone, yERD2 alone or with both. 50 µg of AmyHDEL was co-transfected with increasing amounts of effector plasmids given below each lane in µg and in the combination with fixed amount of yERD2 (10 µg). Error bars are standard deviations of three independent transfections. B) Total α-amylase activity is given arbitrary relative units (ΔO.D./ml/min) C) Total GUS O.D used as internal control marker for transfection efficiency in all experiments.

Figure 31A shows that the overexpression of atERD2 can quickly recover the induced-secretion effect caused by yERD2 in a dose-response manner (last three lanes) as if yERD2 was never present in the system. Very importantly, the suppression of induced secretion is not accompanied by any reduction of total amylase activity (Figure 31B), which shows that any observed effects were not caused cell death. Figure 31C also shows that increment in GUS levels is in accordance to the addition of more plasmid. These results are a first step to demonstrate that the induced secretion effect is unlikely to be caused by direct ERD2 oligomerization and dominant re-direction of endogenous ERD2.

4.2.6 ERD2 domain-swap analysis reveals that the yeast ERD2 C-terminus is not functional in plants

The main purpose of studying the evolutionary conservation of ERD2 amongst eukaryotes in plant was to identify variants of the receptor with lack-of-function or partial-function to further pursue a domain-swap analysis to narrow down regions of interest and avoid a point mutagenesis screen of the entire ERD2 coding region. Given the observed functional differences between yeast ERD2 and other ERD2 variants tested, in particular the potential altered function

(induced secretion) phenotype, and considering the earlier complementation results (Lee et al. 1993), a domain-swap using yERD2 and atERD2 appeared to be an ideal approach to gain further insight into functional domains of ERD2.

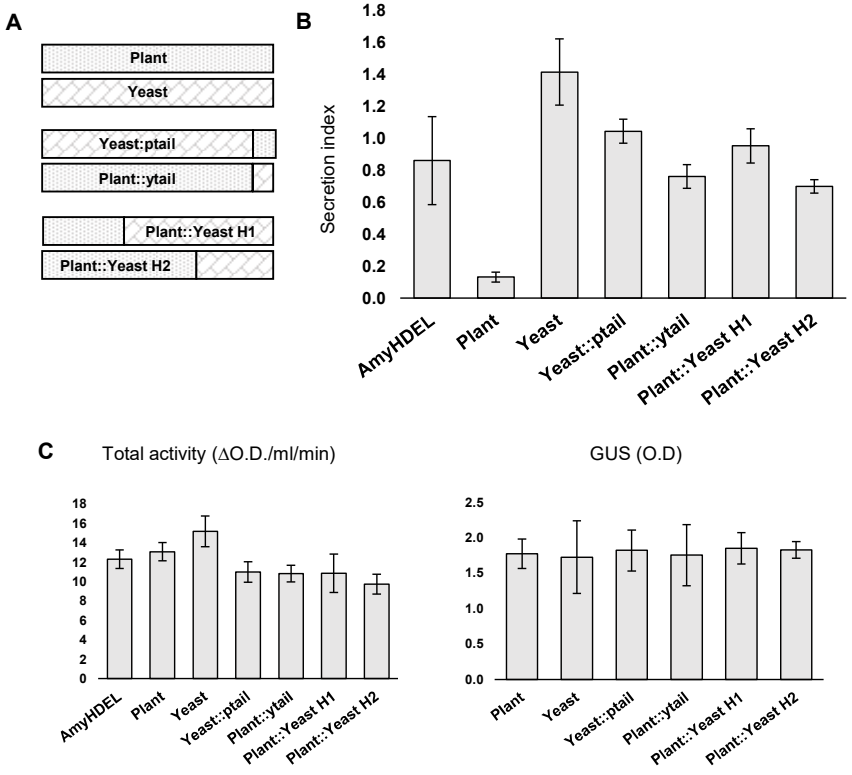


Figure 32 Domain-swap analysis to dissect the induced-secretion effect caused by yERD2.

A) Illustration of chimeric constructs created by swapping atERD2 (dotted pattern) and yERD2 (brick pattern) TM domains. B) Co-expression of AmyHDEL cargo alone or combined with hybrid ERD2 fusions in *Nicotiana benthamiana* protoplasts. 50 µg of AmyHDEL was co-transfected with variable µg of effector plasmid adjusted to yield equal GUS levels in pilot experiment. C) Total α-amylase activity obtained in each cell suspension given in arbitrary relative units (ΔO.D./ml/min). D) Total GUS O.D used as internal control marker for transfection efficiency.

Two hybrid molecules were initially created by the permutation of the last 16 C-terminally located amino acids in both orthologs just after the conserved DFFY region, creating Yeast::ptail and Plant::ytail ERD2s. The next step was to progressively replace the plant ERD2 core by longer parts of the yeast protein, thus creating Plant::Yeast H1 by the replacement of 158 amino acids and Plant::Yeast H2 by the replacement of 78 amino acids. In all cases, conserved regions were selected for the junctions between the hybrids. Figure 32A shows a schematic representation of the new hybrids.

The replacement of the atERD2 C-terminus within the yeast core was not sufficient to generate a functional receptor, but it was sufficient to stop the induced-secretion effect (Figure 32B, compare fourth lane to third and second). Nonetheless, the reciprocal domain swap in which the yERD2 C-terminus replaced the corresponding region in the plant ERD2 core completely abolished the biological function of the hybrid. Finally, the longer domain replacements of yERD2 within atERD2 did not yield further insight as the hybrid molecules failed to show any effect on the cargo molecule. Unfortunately a specific region causing the gain-of-function characteristic to induce AmyHDEL secretion could not be identified with this approach. Figure 32C demonstrates that the experiments were done with equal GUS loading and that amylase total activity was never compromised. Thereby, the main conclusion from the domain-swap analysis is that the yeast C-terminus, which is 4 amino acids longer than the *Arabidopsis thaliana* ERD2 C-terminus and does not contain the conserved LXLP motif identified earlier (An 2015; Silva-Alvim et al. 2018, Chapter 2), is unable to act when displayed at the atERD2 C-terminus.

4.2.7 The dual ER-Golgi localisation of yeast ERD2 is not affected by HDEL-cargo nor the presence of the plant ERD2

The induced-secretion effect caused by the overexpression of yERD2 in plant protoplasts observed before (Figure 30) could be caused by different reasons, for instance ERD2 self-interactions previously discussed and partially challenged (Figure 31), as well as by major disturbances on the membrane balance between the ER-Golgi interface. To further test the former, I decided to apply the same strategy previously used to generate a biologically active fluorescent fusion of ERD2 (Silva-Alvim et al. 2018).

The ERP1 first transmembrane domain bearing an RFP on its N-terminus was once again used (Chapter 2, Figure 14) and directly fused to the coding region of yERD2, as represented in Figure 33A, to create RFP-TM-yERD2. The new fluorescent protein was infiltrated in Tobacco leaf epidermal cells and its localization tested via CLSM. Figure 33A reveals a dual distribution of the protein, which highlights both the ER network as well as Golgi-like punctate structures. Co-infiltration with the Golgi marker ST-YFP reveals that the punctate structures were indeed Golgi bodies, Figure 33B (upper panel). A stratification previously

discussed can be noticed and mostly red dots (white arrow heads) and mostly green dots (black arrow heads) are in evidence.

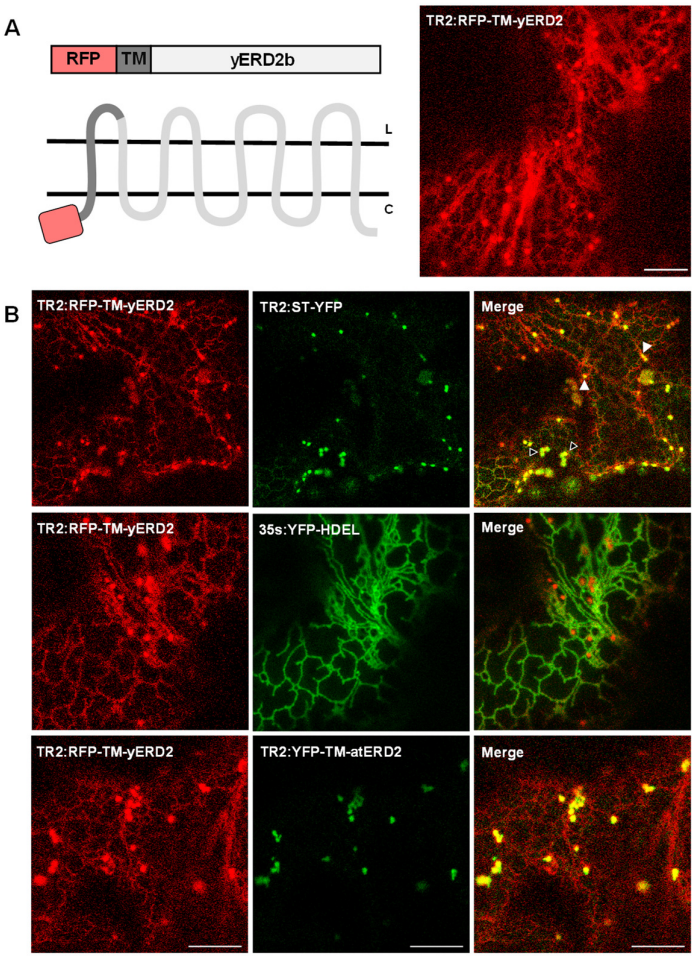


Figure 33 Co-localization studies of yERD2 in plants

A) Illustration of the fluorescent fusion where yERD2 was fused to ERP1 TM bearing RFP. The expected topology is also represented B) Confocal laser scanning microscopy in tobacco leaf epidermis cells showing the distribution of RFP-TM-yERD2 in the ER and punctate structures. Scale bar is 10µm. C) Comparison of intracellular distribution of RFP-TM-yERD2 and the Golgi marker ST-YFP shows that the proteins do co-localise, but an uneven pattern is present, mostly green (black arrow-head) and mostly red (white arrow-head) Golgi. Middle panel shows that the distribution of RFP-TM-yERD2 do not change in the presence of the ER marker YFP-HDEL. Bottom panel shows that co-expression of YFP-TM-yERD2 and RFP-TM-ERD2 do not alter the localization of either proteins. Scale bars are 10µm.

Since the overexpression of yERD2 caused the secretion of AmyHDEL in the bioassay I decided to test the fate of the classical ER maker YFP-HDEL by co-infiltration with RFP-TM-yERD2. Figure 33B (middle panel) shows that there were no changes in the co-localization of either protein.

Finally, I also wanted to challenge once more ERD2 self-interaction and oligomerization concept *in situ*. Figure 33B (bottom panel) shows that the co-infiltration of RFP-TM-yERD2 and the equivalent fusion using the *Arabidopsis thaliana* ERD2 (YFP-TM-ERD2) did not alter the distribution of either proteins, which maintain their typical localization, ER-Golgi for RFP-TM-yERD2 and Golgi-only for YFP-TM-ERD2. This result is an indicative that the induced-secretion

effect caused by the overexpression of yERD2 is unlikely to be caused by interaction with endogenous ERD2.

4.2.8 *Homo sapiens* ERD2 activity in plant cells is not dependent on S209 phosphorylation

A study carried in mammalian cells proposed that a serine located in position 209 (S209) in the C-terminus of the human ERD2 (Figure 34C) is crucial for its interaction with ARF-GAP due to this region being PKA phosphorylated (Cabrera et al. 2003). Such important feature should be highly conserved, however Figure 27B shows that S209 is not conserved. Due to the fact that the *Homo sapiens* ERD2 (hERD2) is fully functional in Nicotiana protoplasts and can efficiently suppress the loss-of-function effect caused by overexpression of K/HDEL ligands in plant protoplasts (Figure 29) I have decided to inspect S209 phosphorylation *in vivo*.

To carry out a similar analysis as the one done by Cabrera et al. (2003) I decided to repeat point mutations in hERD2 used in their work and access the capabilities of these mutants to suppress AmyHDEL secretion (cargo). I introduced an alanine replacement (S209A) that cannot be phosphorylated (permanently inactive) and also tested the substitution of serine by aspartic-acid (S209D), which is a phosphomimetic change (permanently active).

The mutations were directly compared the wild type *Homo sapiens* ERD2 via electroporation and *N. benthamiana* protoplasts using the AmyHDEL bioassay. As it can be seen in Figure 34A, neither S209 modifications had any noticeable effects on the activity of hERD2 in plants. Combined with the observation that this amino acid is not conserved, it is unlikely that PKA phosphorylation plays a role in ERD2 function and interactions with other proteins. Figure 34B shows that experiments were conducted under same GUS loadings and that amylase total activity was stable.

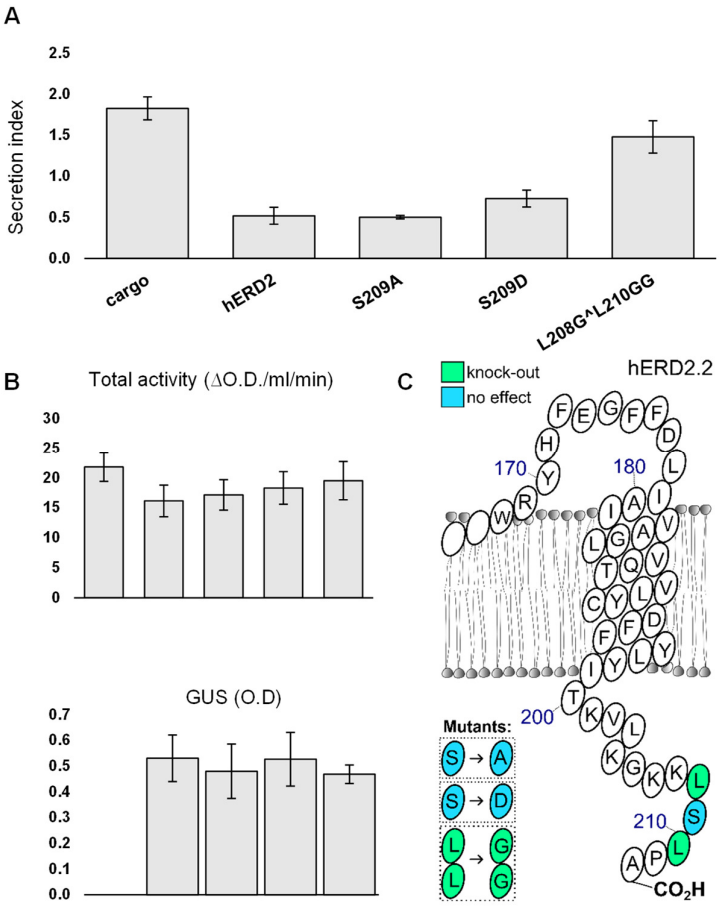


Figure 34 Point-mutagenesis of hERD2 confirms the importance of the conserved LXLP motif.

A) Co-expression of AmyHDEL cargo alone or combined with wild-type hERD2 and mutants in *Nicotiana benthamiana* protoplasts showing that S209 mutations have no effect but that LLGG mutations disrupts hERD2 activity. 50 µg of AmyHDEL was co-transfected with variable µg of effector plasmid adjusted to yield equal GUS levels in pilot experiment. B) Total α-amylase activity obtained in each cell suspension given in arbitrary relative units (ΔO.D./ml/min). Total GUS O.D used as internal control marker for transfection efficiency. C) Schematic representation of hERD2 C-terminus, mutations and the results obtained on the bioassay.

4.2.9 Mutation of the conserved di-leucine motif caused the reduction of human ERD2 activity in plants

In contrast to Cabrera et al. (2003) observations results from our laboratory show that modifying two conserved leucine residues in the ERD2 C-terminus can drastically change the biological function and localisation of the receptor (An 2015; Silva-Alvim et al. 2018). These two residues are conserved also in the human ERD2 C-terminus. For that reason, I decided to reproduce the double-leucine replacement within the context of the human ERD2 C-terminus (L208G^L210G) and test if this double mutant could influence its biological function.

Figure 34A shows that the LxLP motif is equally important for human ERD2 to mediate increased retention of AmyHDEL cargo in plant protoplasts. The double-mutant repeatedly caused the complete loss-of-function of the receptor. As before, Figure 34B shows same GUS loadings and stable amylase total activity

for all plasmids that have been electroporated. Finally, Figure 34C condenses the findings in a schematic representation of hERD2 last transmembrane domain and the point-mutations tested.

4.2.10 YFP-TM-ERD2 expressed in HEK293 cells is more confined to perinuclear regions compared to ERD2-YFP which also labels the cell periphery

Most studies encircling ERD2 function were, and still are, done in either animal cells lines or in yeast, with only few studies been conducted using plant systems. Recent findings linking a disruption of ERD2 biological activity by C-terminally positioned tags and mutations (Chapters 2 and 3; Silva-Alvim et al. 2018) were done in plant systems. Figure 29 shows that human ERD2 can efficiently supress the loss-of-function effect caused by overexpression of K/HDEL ligands in *N. benthamiana* protoplasts.

As consequence of the above mentioned, I have decided to explore the possibility of endorsing these observations in an animal system. I aimed to be able to verify the localisation and biological function of two independent fluorescent fusions of *Arabidopsis thaliana* ERD2 in mammalian cell lines (HEK293). Therefore, to reproduce the amylase secretion assay and to co-localise the protein fusions four plasmids were constructed and are schematically represented on Figure 35A. In all cases the genes were flanked by the strong cytomegalovirus (CMV) immediate-early enhancer and promoter and Bovine Growth Hormone Polyadenylation Signal (BGHpA). The genes chosen to be used are the non-functional and mis-localised C-terminal fusion of ERD2 (ERD2-YFP) the recently published TM-ERD2 (YFP-TM-ERD2, Silva-Alvim et al. 2018), amylase and the engineered amylase bearing a HDEL signal (Amy-HDEL).

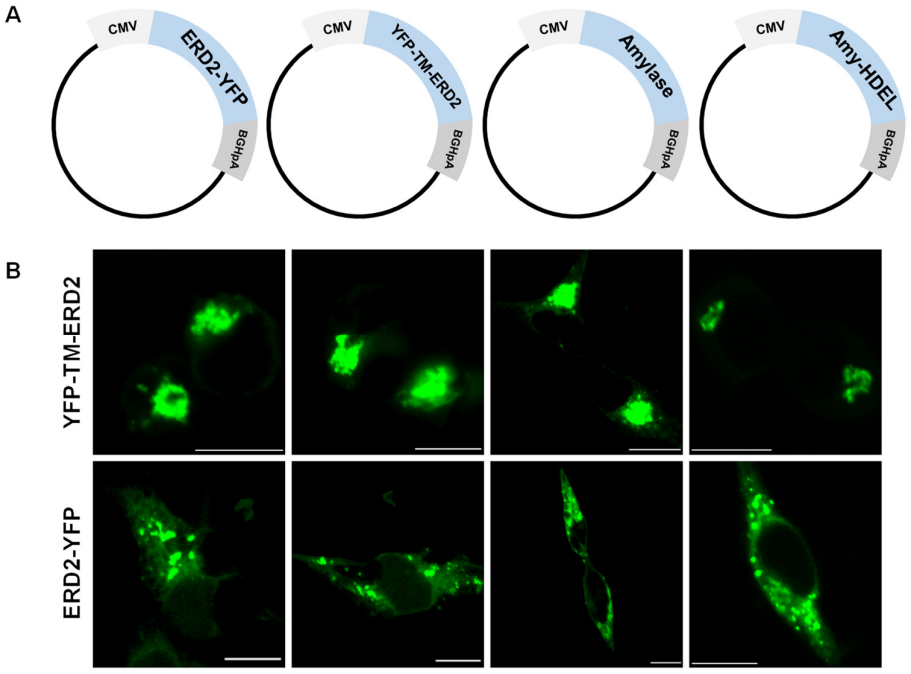


Figure 35 Two ERD2 fusions have different intracellular distribution in HEK293 cells

A) Schematic of expression vectors created to reproduce amylase bioassay and CLSM of mammalian cell lines. B) CSLM of fixed HEK293 cells overexpressing YFP-TM-ERD2b (upper panel) or ERD2b-YFP (bottom panel) shows the conflicting distribution of the proteins. Scale bar is 10µm.

Amylase secretion assay (bioassay) to test ERD2 *in vivo* biological function was preliminary tested in HEK293 cells using the newly created plasmids. However, from the initial test amylase levels were too low to be detected using the standard methods and need to be optimized. Optimization of the method was beyond the scope of this thesis.

However, detection of the two fluorescent plant ERD2 fusions was straightforward and revealed interesting differences. Similarly to what has been observed before in plants (Silva-Alvim et al. 2018), YFP-TM-ERD2 and ERD2-YFP have different distribution and occupy distinct cellular compartments when compared to each other (Figure 35B). YFP-TM-ERD2, a biological active receptor, is clustered close to the nucleus whilst ERD2-YFP seems to be further spread to the periphery of the cells. In addition, the cellular morphology was abnormal when YFP-TM-ERD2 was overexpressed, which could have resulted from the earlier described Brefeldin-A like effect (Hsu et al. 1992). Since ERD2-

YFP is non-functional (An 2015), a more normal morphology is observed, but this requires further experimentation.

4.2.11 YFP-TM-ERD2 appears to partially co-localise with the ERGIC marker ERGIC53

To be able to properly identify the localization of each fusion a co-expression with organelle markers was also necessary. The ER-Golgi interface has many similarities amongst eukaryotes, but, differently from plants (Brandizzi & Barlowe 2013a; Robinson et al. 2015), in animals an intermediate compartment called ERGIC is found and is well characterized by the presence a protein called ERGIC-53 (Hauri et al. 2000). Calnexin is a protein that interacts with a variety of molecules at the ER and for that reason is constantly used as an ER marker (Leach et al. 2002). Preliminary results from Figure 35B indicated that YFP-TM-ERD2 and ERD2-YFP were not in the same compartment.

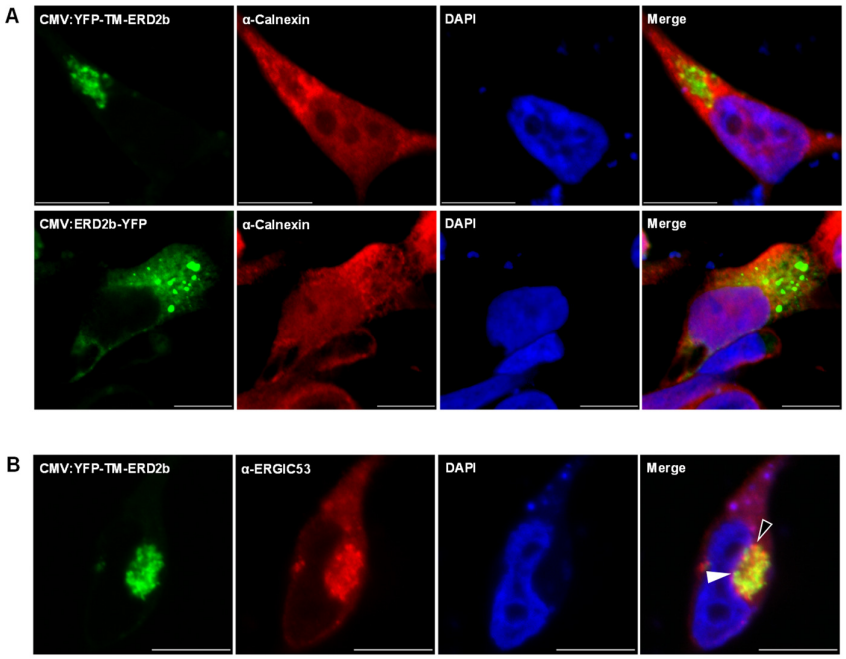


Figure 36 Co-localization studies of two different fluorescent ERD2 fusions in HEK293 cells

A) Immunocytochemistry (ICC) and CSLM of fixed tissue using RFP conjugated anti-calnexin and DAPI in HEK293 cells overexpressing YFP-TM-ERD2b (upper panel) or ERD2b-YFP (bottom panel). Reveals the partial co-localisation with the ER marker for the N-terminal fusion. B) Immunocytochemistry (ICC) and CSLM of fixed tissue using RFP conjugated anti-ERGIC53 and DAPI in HEK293 cells overexpressing YFP-TM-ERD2b. Revealing a stratified partial co-localisation, green only structures (white arrow head) and red only structure (black arrow head).. Scale bar is 10µm.

HEK293 cells were transformed with either YFP-TM-ERD2 or ERD2-YFP in addition to specific markers. After fixation, direct immunofluorescence (dIF) was carried out using conjugated antibodies against Calnexin in conjunction to anti-YFP and DAPI. Figure 36A shows that YFP-TM-ERD2 does not co-localize with the ER marker, but that ERD2-YFP has partial localization with the same marker. This observation agrees with the results obtained in plants (An 2015; Silva-Alvim et al. 2018). dIF of HEK293 cells expressing YFP-TM-ERD2 combined with conjugated antibody against the ERGIC compartment marker (ERGIC53) in Figure 36B reveals a partial co-localization, indicating that plant receptor has the capacity to possibly be trafficking in between the Golgi and the intermediate compartment. Figure 36B also brings to evidence a stratification, green only structures (white arrow head) and red only structure (black arrow head).

In conclusion, expression of fluorescent plant ERD2 fusions in mammalian cells confirms that the position of the fluorescent tag can influence localisation.

4.3 Discussion

In Chapter 1 of this thesis I demonstrate that ERD2 from plants species were interchangeable. In this chapter, ERD2 genes from 13 different organisms were expressed in *N. benthamiana* protoplasts and revealed a remarkable degree of functional conservation even in distantly related eukaryotes from different kingdoms. Potential functional divergence and conservation are discussed below, in the light of experiments conducted in plant cells and in mammalian cells.

4.3.1 Conservation of ERD2 gene is not accompanied by signal-specificity

In Figure 28 difficulties found whilst cloning and expressing *S. cerevisiae* ERD2 were discussed. The reasons behind that are unclear and were not the scope of this project. However, it is important to highlight that in the past the expression of the same yeast ERD2 was problematic and unsuccessful in animal cells (Lewis & Pelham 1990).

Upon surpassing any cloning difficulties and followed by the overexpression of all ERD2 orthologs in plants it became evident that unless a homology divergence threshold is surpassed the receptors from other organisms are

mostly functional in plants (Figure 29A). This observation reveals how well conserved the receptor is amongst eukaryotes since the ERD2 from organisms as different as mammals and plants are fully interchangeable. Results from Figure 29B confirms the argument risen in the past against the signal-divergence of ERD2 since no meaningful difference was observed when the HDEL motif was exchanged for KDEL.

Nevertheless, the fact that yERD2 overexpression caused an unexpected induced-secretion which is dose dependent effect (Figure 30) points once again towards a high conservation level because yERD2 must at least interact with some of the sorting machinery in order to interfere with or titrate some essential component of the retention machinery. However, results indicated that yERD2 is clearly not capable of mediating increased HDEL-cargo retention in plant cells. Interestingly, the induced-secretion effect caused by overexpression of yERD2 can be fully complemented by the ectopic addition of atERD2, which argues to a problem not caused by hetero-oligomerization of the receptors.

Finally, the same strategy used to create a functional and fluorescently tagged plant ERD2 (Chapter 2) was applied to yERD2 and it brought to evidence that this protein is mis-targeted to the ER in addition to the expected Golgi localization. The mis-targeting of ERD2 to the ER is usually accompanied to loss-of-function of the receptor (Chapter 2) and the fact that yERD2 causes an induced-secretion effect cannot be explained simply by the localization of the receptor. It would be interesting to test the hypothesis that this protein can titrate important components of the recycling machinery, for instance via co-localization studies with COPI and COPII components, but it was beyond this thesis.

In matter of fact the few receptors to be found non-functional, klERD2 and tbERD2, might be functional at some extent and maybe the expression levels were too low compared to others originating the weak effect observed. It might be interesting to add an HA-tag to these proteins, similarly to the strategy used in Figure 30, and check the expression levels, but it was beyond the time frame for this thesis.

4.3.2 Mutational analysis of *Homo sapiens* ERD2 contradicts previous observations and confirms the importance of the conserved LXLP motif

Evidence shown in this chapter strongly advocates against any functional divergence of ERD2 in plants and animals. Unlike what was observed when overexpressing yERD2 in plants, the overexpression of hERD2 caused a strong suppression of AmyHDEL and AmyKDEL secretion (Figure 29), confirming the higher tolerance for HDEL-related signals reported before (Denecke et al. 1992).

For that reason, I have decided to use our bioassay to test the hypothesis that S209 of hERD2 may be linked to PKA phosphorylation and functionality of the receptor (Cabrera et al. 2003). Mutations in hERD2 revealed that S209 does not play a role in the biological function of hERD2 in plants. In contrast, the conserved di-leucine motif proved to be equally crucial for the functioning of hERD2 in plants (Figure 34). Furthermore, the overexpression of the two fluorescently tagged plant ERD2 in animal cells confirmed that a biological active N-terminally tagged ERD2 (YFP-TM-ERD2, Silva-Alvim et al. 2018) is Golgi/ERGIC resident and does not co-localize to the ER, in contrast to the C-terminally tagged non-functional receptor which shows a more disperse ER-like localization (Figure 36).

These observations are of extreme importance since they show that results obtained in plant systems are truthful and probably representative of how a functional receptor localizes in different organisms, which is in contrast to previous observations made using C-terminally tagged receptors, claiming that the receptor steady-state is Golgi localized and that it can redistribute to the ER upon overexpression of ligands (Lewis & Pelham 1992a; Li et al. 2009; Montesinos et al. 2014). It will be interesting to establish similar quantitative KDEL or HDEL cargo transport assays using mammalian cells, and to try other mammalian cell lines (i.e. Hela, Cos) which may prove more suitable for confocal laser scanning microscopy.

Chapter 5 **Interesting stand-alone data that may form the basis for future research**

5.1 **Introduction**

One of the key-features of the secretory pathway is the integrated way in which the individual compartments are interconnected via sequential transport steps through budding and fusion of vesicular and/or tubular transport carriers (Bonifacino & Glick 2004). As a result, some organelles can exhibit a relatively fixed identity whereas others mature over time, giving rise to different biochemical compositions. Moreover, individual compartments can be held together by filaments of the cytoskeleton (Sparkes et al. 2009) on which they also readily move within the cell.

Due to its complexity, certain steps controlling the secretory pathway cannot always be studied via genetics alone. This is because the most obvious phenotypes may be secondary consequences, rather than primary causes, because many processes are interconnected. For instance, anterograde and retrograde transport in the pathway depend on each other, and inhibition of a single step may bring an entire recycling pathway down, as shown before by the use of Brefeldin A (BFA) to prevent the activation of ARF1p and consequently the formation of COPI (Pimpl et al. 2000). Some organelles form branching points in the secretory pathway where biosynthetic transport from the ER as starting point meets endocytic transport from the plasma membrane. Some pathways are designed for continuous recycling, whilst other pathways lead to degradation in the lytic compartments. In all these cases, loss-of-function genetics alone is often insufficient to conclusively attribute a specific gene function.

Since ERD2 is controlling a central position in the secretory pathway, it is not surprising that over the course of this project (Chapters 1 to 4) I came across a number of interesting observations that suggest an indirect role of ERD2 in processes that seem unrelated to the primary function in promoting retention of soluble proteins in the ER lumen. These fell outside the scope of this work but nevertheless were interesting and solid enough to form a basis for future research. As this type of additional findings is very often lost and erroneously classified as “unfinished or preliminary” data, I have decided to make these available to the scientific community in the form of a final results chapter.

Chapter 5 of this thesis will describe individual experiments which will be introduced and discussed separately in self-contained subheadings, followed by a brief perspective section to explore how the potential conclusions can be tested by further experiments.

5.2 Results and discussion

5.2.1 Anti-sense inhibition of ERD2 promotes specific accumulation of HDEL-cargo but not secretory cargo

Results from Chapter 1 (Figure 7C) show that anti-sense NbERD2b transcript-induced AmyHDEL secretion was accompanied by an approximately 2-fold increase in the total activity (TA) of AmyHDEL. This increase was reproducible and could thus not be dismissed. Since reduced ERD2 activity and a partial defect in ER retention may lead to decreased ER chaperone levels and associated ER stress, an increase in total cargo reporter activity was totally unexpected. It was therefore decided to investigate this phenomenon in more detail.

Figure 37A confirms that the presence of the anti-sense NbERD2b transcript led to induced AmyHDEL secretion, accompanied by an increase in the total activity of AmyHDEL. When comparing the distribution of cargo in the medium with cargo in the cells separately, it is evident that increased presence of AmyHDEL in the medium cannot originate from the secretion of intracellular AmyHDEL alone.

Intracellular AmyHDEL levels remain the same or are slightly higher than that of the cargo-only control, which means that cells either synthesize more AmyHDEL or degrade less AmyHDEL when ERD2 synthesis is reduced. This is in contrast to other effectors that inhibit ER retention in a different way, for instance the effect of *Saccharomyces cerevisiae* ERD2 which causes induced secretion of AmyHDEL at the expense of intracellular AmyHDEL, resulting in a more constant total activity (Figure 30 and 31, data not shown).

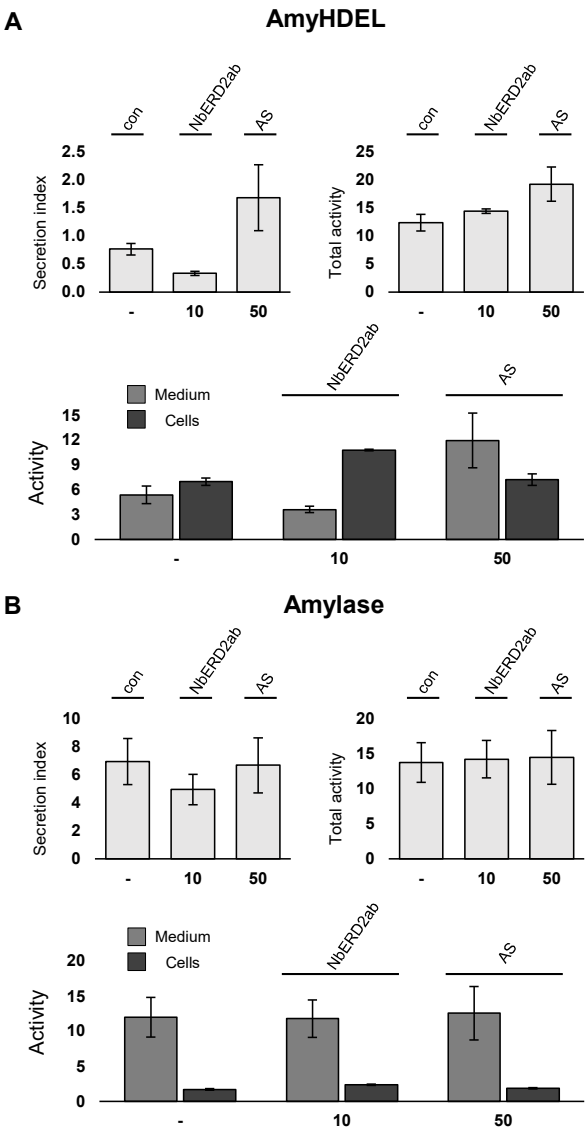


Figure 37 Anti-sense inhibition of ERD2 has a specific effect on the total synthesis of HDEL tagged α-amylase.

A) Transient expression experiment with *Nicotiana benthamiana* protoplasts co-expressing AmyHDEL with either sense NbERD2ab or antisense NbERD2ab. Protoplast were incubated for 48 hours to allow degradation of endogenous ERD2. 50µg of cargo plasmid was electroporated alone or co-electroporated together with sense or antisense ERD2 plasmids. B) Same experimental conditions as before, but using wild-type Amylase. Total α-amylase activity obtained in each cell suspension given in arbitrary relative units (ΔO.D./ml/min). Error bars are standard deviations of three independent transfections.

To investigate this result further, in addition to AmyHDEL, the control cargo Amy was used as independent reporter for secretory protein synthesis by the ER membrane. This was to test if modifying ERD2 levels affects the overall protein synthesis capacity of the secretory pathway, and since Amy and AmyHDEL only differ by 4 amino acids, the polypeptides were not expected to exhibit any major differences in the protein synthesis rate or protein translocation rate.

Figure 37B illustrates that the transport properties of constitutive secretory cargo are unaffected by ERD2 overexpression or anti-sense inhibition. This is particularly clear when comparing individual medium and cell values which change dramatically for AmyHDEL but not at all for Amy. These results were expected because Amy does not have an ER retention signal. However, the figure also shows that in contrast to AmyHDEL, the total activity levels of the control cargo Amy were unchanged under all conditions. This result was certainly

unexpected, and the clear increase in AmyHDEL yield upon ERD2 anti-sense inhibition may not be caused by changes in protein translocation and folding but may result from specific post-translational events affecting HDEL cargo alone. Equal amount of GUS levels is a control (data not shown) indicating that the transfection of sense and anti-sense plasmids was comparable for Amy and AmyHDEL series and could not account for the discrepancy in the total activity of AmyHDEL.

Results from these experiments support the idea that in the absence of the receptor cells may compensate by increasing *de novo* synthesis of chaperones or activating other mechanisms to promote secretory protein synthesis. This is in contrast to earlier work linking ER stress by the drug tunicamycin to reduction of secretory versus cytosolic protein synthesis (Leborgne-Castel et al. 1999), although experiments at the time were solely done with Amy, not AmyHDEL.

It has been proposed before that ERD2 can play an important role in ER quality control (Yamamoto et al. 2003) and results presented here could indeed indicate a link between a functional receptor and the activation/control of ER stress response. Previous observations were done in different organisms and I believe that ERD2 possibly plays an active role in the ER quality control which has not been observed before in plants. These observations made during my PhD and hereby presented make the basis for a future research project that could lead to the characterization of a previously unknown ERD2 function. Further work is required to substantiate these findings, such as a measurement of chaperone levels upon ERD2-antisense expression, or a repetition of Amy/GUS ratios upon tunicamycin treatment with AmyHDEL instead of Amy, but this was beyond the scope of this thesis.

5.2.1.1 Development of transgenic lines to push-forward the research on the effects of ERD2 anti-sense inhibition

Due to the lethality of ERD2 knock-out an anti-sense inhibition strategy was conceived and applied *in vivo*, as discussed in Chapter 1. To further access the outcomes of ERD2 inhibition and complementation *in planta* using YFP-TM-ERD2 I have started the development of *N. benthamiana* transgenic lines. Transgenic plants were generated using the standard leaf-disc transformation technique (Gallois & Marinho 1995). Plasmids chosen to transform discs via

Agrobacterium tumefaciens incubation, as well as, the transformation efficiency in percentage can be seen in Figure 38A.

The different combinations were chosen so that the anti-sense inhibition effect could be distinguishable by the presence or absence of HDEL cargo in the Golgi apparatus, similarly to the *in situ* assay presented in Chapter 2.

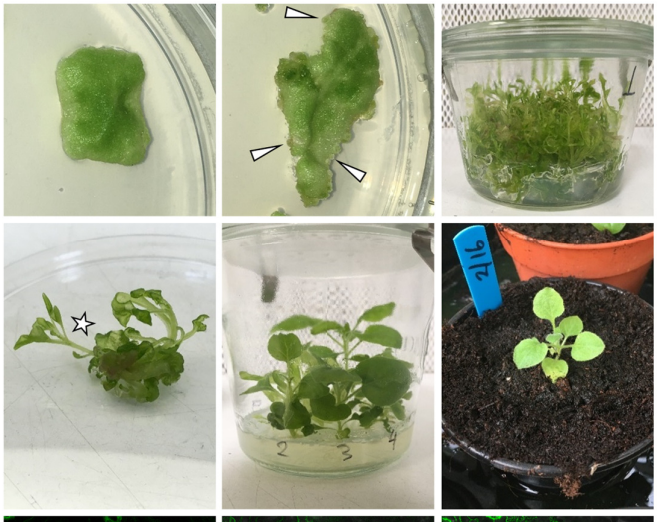
A step-by-step procedure is shown in Figure 38B. Small leaf-discs were excised from *Nicotiana benthamiana* plants and exposed via incubation to the desired plasmids carried by *Agrobacterium tumefaciens*. Leaf-discs where then Incubated at constant temperature in plates containing plant medium for 4 weeks. As soon as small calli were evident hey were individually removed (Figure 38B, white arrow heads) and transferred to growing jars. At least 50 calli were isolated per combination and grown for a period of 4 to 6 weeks. Overgrown calli were individually separated and one shoot per calli was selected for regeneration in MS medium (Figure 38B, white star).

A

Plasmid 1	Plasmid 2	% of transf.
TR2:ST-YFP-HDEL / 35S:NbERD2ab antisense	-	22
TR2:YFP-1TM-ERD2b / 35S:NbERD2ab antisense	-	55
TR2:RFP-1TM-ERD2	-	47.8
TR2:ST-YFP-HDEL / 35s:Amy	-	100
TR2:ST-YFP-HDEL / 35s:AmyHDEL	-	90
TR2:ST-YFP-HDEL / 35S:NbERD2ab antisense	TR2:RFP-1TM-ERD2	69 / 14 *
TR2:YFP-1TM-ERD2b / 35S:NbERD2ab antisense	TR2:RFP-1TM-ERD2	33 / 0 *
TR2:ST-YFP-HDEL / 35s:Amy	TR2:RFP-1TM-ERD2	55 / 0 *
TR2:ST-YFP-HDEL / 35s:AmyHDEL	TR2:RFP-1TM-ERD2	68 / 16 *

* 1st number represent single transformants and 2nd number double transformants, n ≥ 9.

B



C

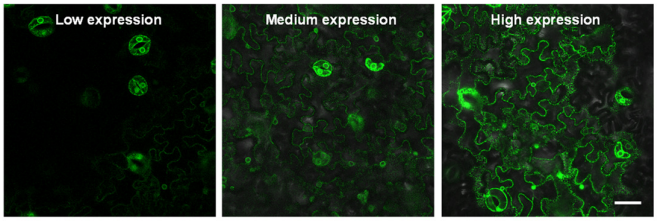


Figure 38 Generation of *Nicotiana benthamiana* transgenic lines to explore the effects of ERD2 inhibition and YFP-TM-ERD2 complementation *in planta*.

A) Demonstrative table showing the different combination of plasmids used for transformation. Last lane shows the success levels of single and double transformants in percental B) Step-by-step procedure to generate *in vitro* transgenic *N. benthamiana* plants by the use of leaf-disk transformation followed by calli regeneration. C) Example of CLSM analysis of regenerated plants screening for high expression and also for double-transformants. Scale bar 50μm.

Regenerated plants were analysed and scored accordingly to their expression levels using confocal laser scanning microscopy, as shown in Figure 38C, the same microscope settings maintained constant so that all new plants were equally evaluated.

Transgenic lines obtaining during the last months of my PhD are important and will allow for future research aiming to shown that the antisense inhibition technique was sufficiently capable of *knock-down* both endogenous copies of the ERD2 gene found in *Nicotiana benthamiana* and that also the newly generate fluorescent tagged ERD2 (YFP-TM-ERD2) is capable to rescue any phenotype caused by the absence of the endogenous genes. For the future, in complementation to the scoring of transgenic plants via CLSM, it would be interesting to check how low the endogenous ERD2 levels can get in comparison to wild type plants using, for instance, quantitative polymerase Chain Reaction (qPCR).

The increasing advances of CRISPR-CAS9 technology, particularly in plants (Bortesi & Fischer 2015; Liu et al. 2017; Ma et al. 2016), is also an interesting idea to be considered for the future since it could allow for the *knock-in* of YFP-TM-ERD2 in combination to a double *knock-out* of nbERD2s.

5.2.2 Overexpression of ERD2 inhibits vacuolar sorting

Chapter 4 of this thesis was focused on exploring the evolutionary conservation of ERD2 gene amongst eukaryotes. For that purpose, ERD2 coding region of 13 different species was overexpressed in *N. benthamiana* protoplast and their capacity to supress ERD2 loss-of-function effect caused by overexpression of ligands was accessed.

The secretory pathway is a very well-regulated system due to the importance to maintain organelles size and composition balanced via the membrane flux caused by vesicle mediated transports (for review, Lippincott-Schwartz 2011). When analysing ERD2 evolutionary conservation I have decided to use two extra control experiments to ensure that ERD2 overexpression was not causing an imbalance of the secretory pathway and that any observed differences were due to ERD2 partial or complete loss of function. As discussed before, vacuolar sorting and retrograde transport from the Golgi to the ER are mediated by different receptor in plants, VSRs and ERD2 respectively (Kirsch et al. 1994; Lee

et al. 1993), which are believed to bind to different targeting motifs (Pelham 1989; Johnson et al. 1987) thus AmySPO was used a negative control. For that purpose the experiment from Figure 29 was repeated using either Amylase or a engineered amylase bearing the vacuolar signal of sporamin (AmySPO, Pimpl et al. 2003)

Figure 39A shows the surprising result, obtained after GUS normalization, that the overexpression of most ERD2 variants inhibits vacuolar sorting, leading to induced-secretion of AmySPO.

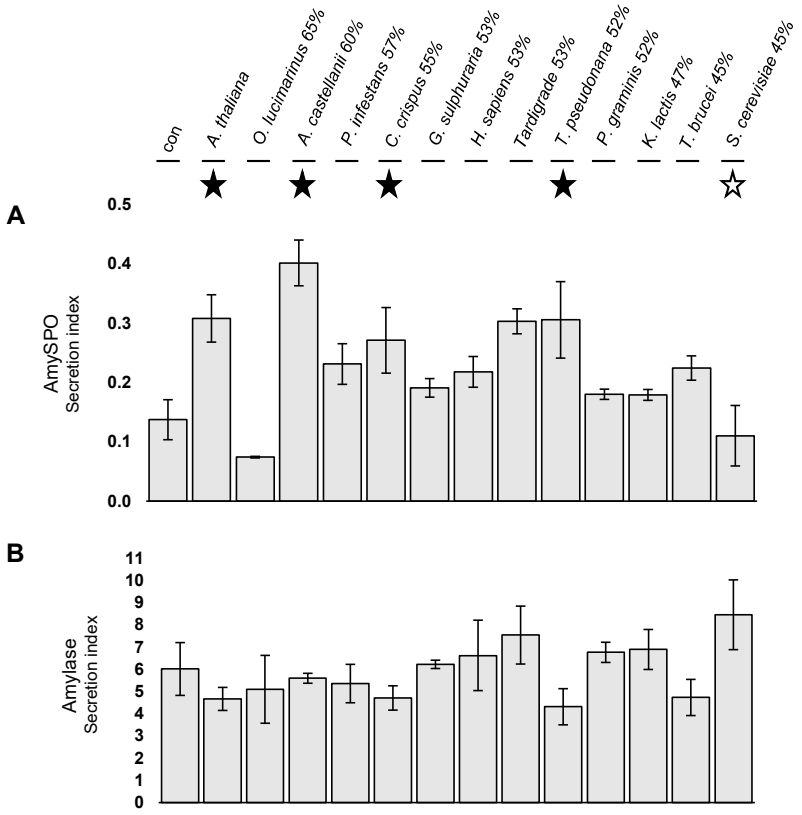


Figure 39 Overexpression of ERD2 can affect vacuolar sorting

A) Co-expression of AmySPO cargo alone or combined with ERD2 from 13 organisms, including *Arabidopsis thaliana* (atERD2b) *Ostreococcus lucimarinus* (oiERD2), *Acanthamoeba castellanii* (acERD2), *Phytophthora infestans* (piERD2), *Chondrus crispus* (ccERD2), *Galdieria sulphuraria* (gsERD2), *Homo sapiens* (hERD2), *Hypsibius dujardini* (Tardigrade - taERD2), *Thalassiosira pseudonana* (tpERD2), *Puccinia graminis* (pgERD2), *Kluyveromyces lactis* (klERD2), *Trypanosoma brucei* (tbERD2) and *Saccharomyces cerevisiae* (yERD2) in *Nicotiana benthamiana* protoplasts. The homology level of the proteins compared to the plant ERD2 is shown in percentage on top of each lane. 50 µg of AmySPO was co-transfected with variable µg of effector plasmid, adjusted to yield equal GUS levels in pilot experiment. Black star mark variants with strong effect and white star denotes no effect on AmySPO secretion. B) Same experimental condition as panel A, but using wild-type Amylase. Error bars are standard deviations of three independent transfections.

The negative effect is particularly strong for ERD2 variants that can efficiently suppress AmyHDEL secretion (black stars). It is also shown that some ERD2 variants have no negative effect on vacuolar sorting, for instance *Ostreococcus lucimarinus* ERD2 and also *S. cerevisiae* ERD2 which causes induced-secretion of AmyHDEL (white star). Figure 39B show that the same experiment repeated using Amylase as a reporter does not have drastic variances amongst ERD2 from different eukaryotic organisms.

Vacuolar sorting interference is a very interesting result because, even-though both receptors are believed to operate in separate routes, they were proposed to compete with each other in the *cis*-Golgi cisternae for a dual signal cargo carrying both type of sorting signal (Gershlick et al. 2014). Vacuolar sorting and Golgi-ER retrograde transport uses different vesicles, but the formation of these vesicles maybe be controlled by the same GTPase, ARF1 (Pimpl et al. 2003).

Therefore, the initial observation that overexpression of ERD2 can compromise vacuolar traffic could form the basis for research aiming to show that the retention of ER soluble proteins is process that may be further integrated in downstream sorting events. One possible hypothesis is that ERD2 overexpression interferes with efficient VSR trafficking, which can be analysed using the tools generated in this thesis.

5.2.3 ERD2 overexpression causes ARF1-dissociation from the Golgi apparatus

Currently, there are two established assays to measure and characterize the biological activity of ERD2, an *in vivo* bioassay and a *in situ* assay (Silva-Alvim et al. 2018). Aiming to further characterize and possibly track the biological activity of ectopically expressed ERD2 *in vivo*, I wanted to explore if ERD2 overdose could influence the tight machinery involved in anterograde and retrograde transport.

ADP-ribosylation factor 1 (ARF1) is a low molecular weight GTPase belonging to and extensive family and that mediates the formation of COPI and clathrin coated vesicles (Vernoud 2003). ARF1 localization has been controversial in the past (Bohlenius et al. 2010; Robinson et al. 2011).

Since N-termini of ARF1 GTPases is required for membrane association, in collaboration with laboratory members, a C-terminal fluorescent fusion of ARF1

(ARF1-RFP) was generated. The strong CaMV35S promoter fusions yielded excessive cytoplasmic background labelling for this construct (data not shown), and to avoid that issue the coding region was replaced under the transcriptional control of weak the TR2 promoter (Bottanelli et al. 2012). ARF1-RFP was co-expressed in tobacco leaf cells with dual the expression constructs harbouring a fluorescent Golgi marker (ST-YFP) under the control of TR2 promoter together with a mock effector PAT or ERD2 under the control of the CaMV35S promoter.

Figure 40A shows that ARF1-RFP strongly labelled the Golgi bodies when co-expressed with the double vector carrying ST-YFP and the mock effector. Occasional extra-Golgi structures were seen as well (white triangles), giving rise to the low correlation and extra red population seen in the scatterplot. These extra red punctate structures were co-localised with the TGN marker RFP-SYP61 (data not shown).

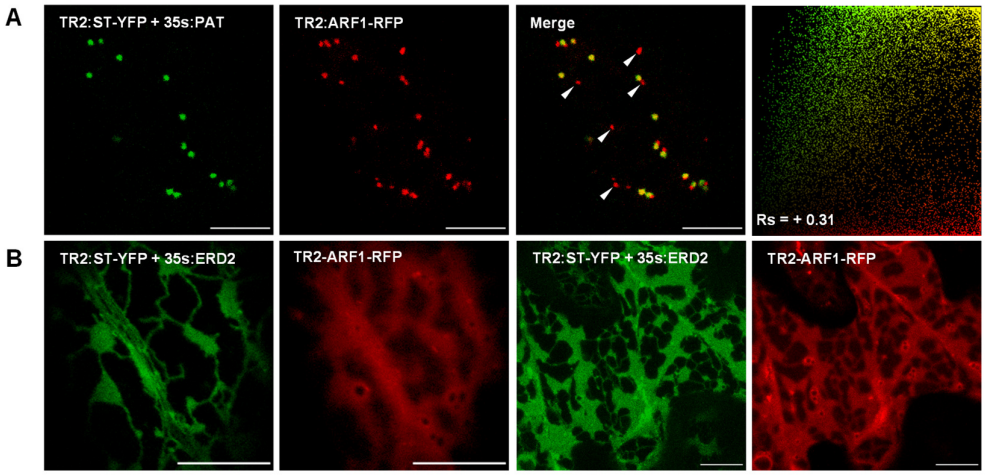


Figure 40 Overexpression of ERD2 drives ARF1 from the Golgi to the cytosol

A) Co-expression of a dual expression vector (ST-YFP + PAT) combined with ARF1-RFP in tobacco leaf epidermal cells analysed by CLSM show partial co-localisation and a red-only population (white arrow-head). Scatterplot and Spearman correlation coefficient (Rs), confirming the observation. B) Co-expression of a dual expression vector (ST-YFP + ERD2) combined with ARF1-RFP analysed by CLSM shows the redistribution of both proteins to the ER and cytosol, respectively. Scale bars are 10µm.

However, when the Golgi-marker ST-YFP was co-expressed with ERD2 and ARF1-RFP, the Golgi marker ST-YFP was completely redistributed to the ER and ARF1-RFP was completely cytosolic as it can be seen in Figure 40B. Since the drastic re-distribution of the Golgi marker was not seen by ERD2 overexpression alone and not by ARF1-RFP expression alone either, it was the

simultaneous ectopic expression of ARF1 and ERD2 that caused a strong effect on the maintenance of the Golgi marker. An earlier reported BFA-like effect (Hsu et al. 1992) may thus be caused by a more complicated interplay between ARF1 and ERD2.

The mechanism by which ST-YFP maintains high steady state levels at the Golgi are still unknown but it may depend on active ARF1 residing on Golgi membranes. Membrane dissociation of ARF1 caused by ERD2 overexpression can be explained by the observed membrane recruitment of ARF1-GAP (Aoe et al. 1997), which would lead to GTP hydrolysis and inactivation of ARF1. However, this alone would not explain how ERD2 is transported in the cell, and results from this thesis suggest that ERD2 does not recycle back to the ER but is a permanent Golgi-resident (Chapter 3, Figure 19).

The dramatic change of ARF1 localization in the presence of ERD2 is an interesting find that can possibly in the future yield an alternative method to verify ERD2 biological activity. Based on these observations the step forward would be investigating if ARF1 cytosolic redistribution is achieved with non-functional ERD2, such as the di-leucine mutant. It should also be realised that ERD2 mediated redistribution of ARF1 from the TGN-membranes as well, because no punctate structures remain. This result may help to explain why ERD2 overexpression caused induced secretion of vacuolar cargo because ARF1 has been implicated in vacuolar sorting as well (Pimpl et al. 2003).

5.2.4 ERD2 overexpression partially recruits Sar1 from the ER surface to the Golgi vicinity

Regardless of its Golgi-residency, ERD2 is synthesized on the rough ER and needs to be transported to the Golgi apparatus. Earlier results suggest that ERD2 anterograde ER to Golgi transport is relatively slow (Chapter 3, Figure 19), but this could be due to low levels of *de novo* synthesis. Furthermore, strong Golgi localisation is secured by a combination of Golgi-retention mechanisms dependent on, for instance, the C-terminus to prevent retrograde transport back to the ER (FRAP results on mutant) and on the entire ERD2 core to prevent escape to post-Golgi compartments (VSR tail construct does not go beyond Golgi).

ER export is mediated by COPII vesicles (Barlowe et al. 1994) and depends on an ARF1-related GTPase known as secretion-associated and ras-superfamily related gene (Sar1 - Nakańo & Muramatsu 1989). Sar1 homologues have been identified in plants and their localization has been shown to be both ER associated and also cytosolic (Bar-Peled & Raikhel 1997) . Very similarly to the strategy used in Figure 40 in a collaborative effort Sar1-Venus was created under the control of TR2 promoter. This strategy was chosen because earlier work with stronger promoter constructs yielded a significant level of cytosolic Sar1 (daSilva et al. 2004), which could have been overexpression artefacts. Co-expression analysis with the double vector carrying a Golgi marker ST-RFP and either a mock effector PAT or ERD2 under the control of the CaMV35S promoter was carried-out.

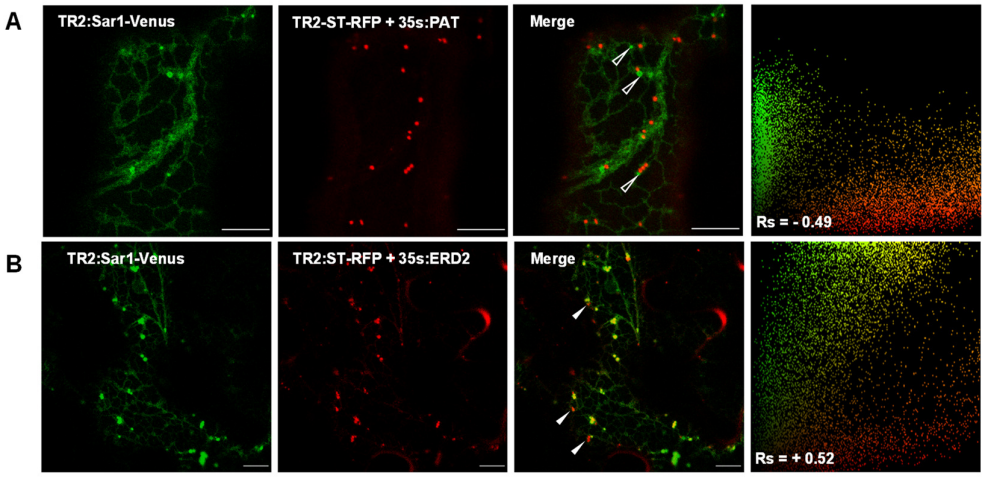


Figure 41 Overexpression of ERD2 drives Sar1 from the ER to the Golgi periphery

A) Co-expression of a dual expression vector (ST-RFP + PAT) combined with Sar1-Venus in tobacco leaf epidermal cells analysed by CLSM show no co-localisation and two distinct populations. Scatterplot and Spearman correlation coefficient (Rs), confirming the observation. B) Co-expression of a dual expression vector (ST-RFP + ERD2) combined with Sar1-Venus analysed by CLSM shows the redistribution of Sar1–Venus from the ER to the Golgi in the presence of ERD2, and the proteins now show partial co-localisation. Scatterplot and Spearman correlation coefficient (Rs), confirming the observation. Scale bars are 10µm.

Using the weaker promoter construct, Sar1-Venus clearly labelled the ER with occasional punctate clusters of more intense fluorescence (Figure 41A). These punctate structures did not co-localise with the Golgi marker when co-expressed in the presence of a mock effector. However, in the presence of ERD2, Sar1-Venus was strongly redistributed from the ER network to bright punctate structures, Figure 41B. All punctuate structures now co-localised with the Golgi

marker ST-RFP and a notable difference can be seen when comparing the scatterplots for Figure 41A/B. A minority of ST-RFP punctate was not labelled by Sar1-Venus (white arrow heads), but the vast majority of ER fluorescence was re-distributed to Golgi bodies. The results show that ERD2 overexpression mediates strong recruitment of the GTPase Sar1 from the ER network to the Golgi vicinity. These data confirm that ERD2 expression promotes ER export site formation (daSilva et al. 2004). The data could also indicate the presence of a cascade effect between GTPases controlling COPII and COPI transport because they behave in an opposite manner in response to ERD2 overexpression when considering membrane recruitment. This strong effect caused by the presence of ERD2 can be explored further with the various loss-of-function mutants that have been generated in the course of this thesis and help to classify them into different categories.

5.2.5 An artificial single chain ERD2 heterodimer molecule can mediate retention of HDEL-cargo and localizes exclusively to the Golgi apparatus

It has been proposed that the KDEL receptor can self-oligomerize and that this feature will enhance the recruitment of GAP (Aoe et al. 1997). Based on that observation I postulated that the overexpression of a chimeric protein, which simulates ERD2 heterodimerization, could shed light on this theory. For the construction of this chimera I selected the coding region of ERD2 from two different species, *P. graminis and A. thaliana* ERD2, to avoid direct repeats and potential clone instability by deletion. Since the ERP1 TM domain can be fused to ERD2 N-terminus without affecting its properties (Chapter 2 and Silva-Alvim et al. 2018) I decided to use that TM as a linker for the heterodimer. The newly created untagged chimeric protein, pg::atERD2, was tested for its biological function using the bioassay and Figure 42A demonstrate that it was efficient in suppress the secretion of HDEL ligands. It is likely that the biological activity originates from the C-terminal ERD2 because it was shown in Chapter 2 that C-terminal extension of ERD2 abolishes its activity.

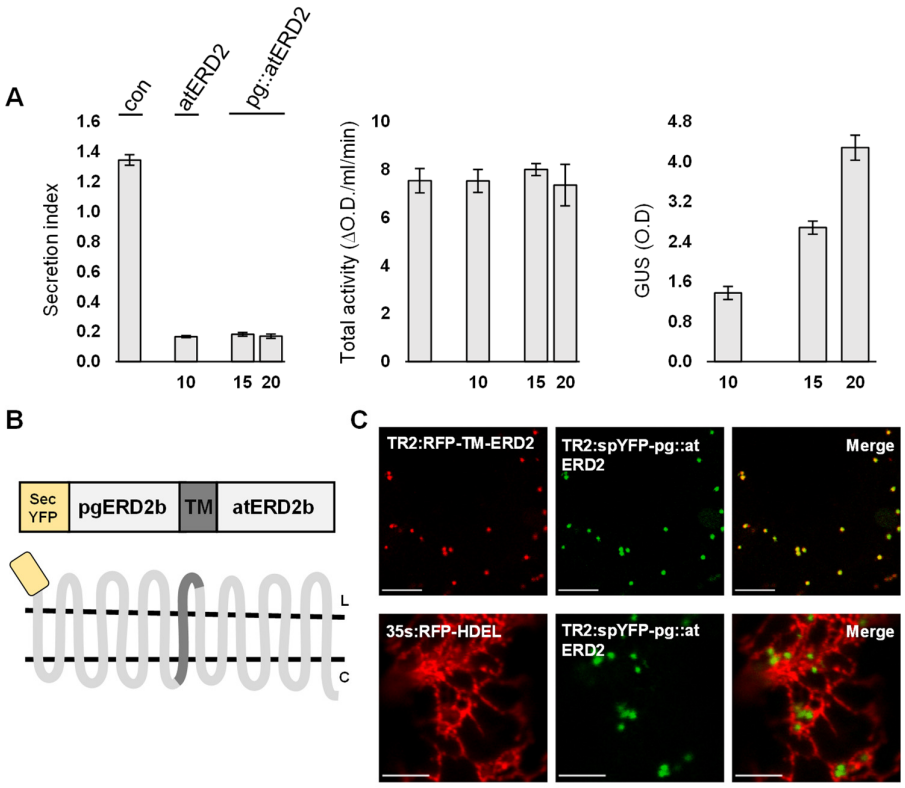


Figure 42 Artificial ERD2 heterodimer can mediate retention of HDEL cargo

A) Co-expression of AmyHDEL cargo alone or combined with atERD2b and the chimeric protein pg::atERD2 in *Nicotiana benthamiana* protoplasts. 50 µg of AmyHDEL was co-transfected with amounts of effector plasmid in µg shown below each lane. Error bars are standard deviations of three independent transfections. Total α-amylase activity is given arbitrary relative units (ΔO.D./ml/min). Total GUS O.D used as internal control marker for transfection efficiency B) Illustration of the fluorescent tagged hybrid protein and the expected topology are shown C) Confocal laser scanning microscopy in tobacco leaf epidermis cells showing the intracellular distribution of spYFP-pg::atERD2 in the presence of either RFP-TM-ERD2 or RFP-HDEL. Scale bar is 10µm.

In a second step I wanted to investigate the sub-cellular localisation of the new chimera. To avoid extending the molecule even more I fused a signal-peptide YFP (secYFP) fluorophore directly to pg::atERD2 N-terminus. A schematic representation of the fluorescently tagged forced heterodimer (secYFP-pg::atERD2) can be seen in Figure 42B.

Figure 42C shows that secYFP-pg::atERD2 perfectly co-localizes with co-expressed RFP-TM-ERD2 in tobacco, indicating that the chimera is Golgi localized. Co-infiltration with the ER marker RFP-HDEL shows that the chimera does not reach the ER. Based on that observation I concluded that, the chimeric protein and the biological active fluorescent receptor do not seem to mis-localize

each other. Furthermore, the chimeric protein does not redistribute in the presence of HDEL ligands. The results suggest that ERD2 can tolerate quite extensive fusion proteins without losing its subcellular sorting fidelity and biological function. Whether the fusion protein reflects biologically meaningful dimers is yet to be tested.

5.2.6 ERD2 heterodimer seems to enhance tubular emanations connecting Golgi bodies

As discussed before the chimeric protein simulating a heterodimer form of ERD2, linked by an artificial TM domain, is Golgi resident and is capable of mediating the increased retention of HDEL ligands (Figure 42). Interestingly, one of the features observed by YFP-TM-ERD2 is the formation of tubular emanations coming from and linking different Golgi bodies, which has been briefly mentioned by Silva-Alvim et al. (2018). Figure 43A shows that the chimera is still capable of forming these tubules without compromising the formation of tubules by the non-dimerized receptor (white arrow head and black arrow head, respectively). Interestingly, during data collection I observed that in the presence of secYFP-pg::atERD2 the formation of tubular emanations is drastically enhanced and that these tubules can either be connecting static Golgi clusters (Figure 43B, black star) and/or in transit mobile Golgi bodies (Figure 43B, white star).

Based on these results, it would be important to further analyse potential ERD2 dimerization and to see if the secYFP-pg::atERD2 would interfere with the localization of possible interaction partners, such as ARF-GAP. It would also be interesting to create a forced homodimer using two identical proteins, but with different codon usage to avoid recombination, and compare its localization and biological function to the heterodimer.

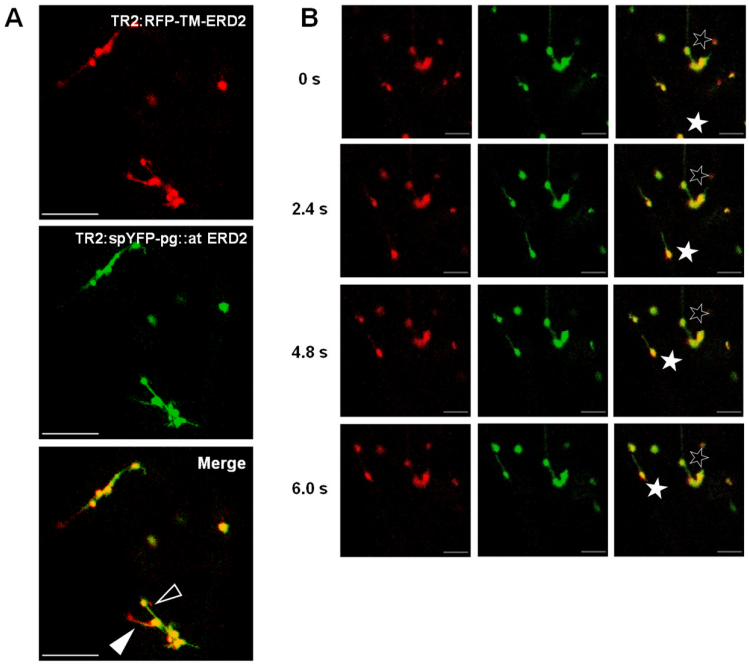


Figure 43 Artificial ERD2 heterodimer caused enhanced tubular emanations connecting adjacent Golgi

A) Confocal laser scanning microscopy in tobacco leaf epidermis cells showing the intracellular distribution of spYFP-pg::atERD2 in the presence of RFP-TM-ERD2 and demonstrating the enhanced effect with tubules connecting multiple Golgi. Also showing that tubules are formed by the WT molecule (black arrow-head) and by the chimera (white arrow-head). B) Time-series using same experimental conditions as panel A, showing that the tubular emanations can connect static Golgi clusters (black star) and/or in transit mobile Golgi bodies (white star). Scale bar is 10µm.

It is important to mention that the accounted results could purely be coming from the second fused ERD2, which has an un-obstructed C-terminus, and for that reason future experiments should test that possibility. However, due to the enhanced tubular formation phenotype, seen in Figure 43B, it is possible that the chimeric protein as a whole is biologically active.

Due to the enhanced tubulation observed by the overexpression of secYFP-pg::atERD2 I consider it as being an important reagent, which makes it a good candidate to be used in the future to further characterize these structures. Since research in this thesis has suggested that ERD2 may not recycle back to the ER, it is possible that its function is entirely different from the classical receptor model that suggests ligand binding in one compartment, vesicle transport, ligand-release in the acceptor compartment. One possibility is that ERD2 forms an affinity matrix in the early Golgi cisternae that allows secreted and vacuolar proteins to pass through but excludes H/KDEL proteins. Whilst it is not clear how such an exclusion mechanism could operate, the proposed structure of ERD2 does not include large lumenal domains for ligand-binding. Instead, it probably mostly occupies the membrane itself, similar to reticulons (Tolley et al. 2008; Sparkes et al. 2009; Sparkes et al. 2010) and Golgi-resident TM9 proteins (Gao

et al. 2012) and could contribute to tightly controlled membrane structures that exclude access of K/HDEL cargo to the Golgi-lumen.

5.3 **Conclusions**

Albeit unrelated to each other, the diverse additional observations all suggest that ERD2 may play a central role in various steps in the secretory pathway, ranging from protein synthesis and turnover to vacuolar sorting and Golgi morphology. The established use of a transmembrane linker to create extensive additions to the N-terminus of ERD2 without compromising its biological activity opens up new strategies to overproduce ERD2 for potential crystallisation and structural studies, which would help to shed light on how one of the first sorting steps in the secretory pathway functions.

C General discussion and future consideration

At the start of my project, the general dogma surrounding protein transport at the ER-Golgi interface conveyed the view that ER soluble proteins are lost by bulk-flow and that a receptor, called ERD2 (Semenza et al. 1990), retrieves those proteins via recognition of a C-terminal motif from the Golgi back to the ER (Pelham 1988; Pelham et al. 1988; Dean & Pelham 1990). However, signals controlling ERD2 traffic between these two compartments, in particular the manner in which retrograde ERD2 transport takes place and the mechanisms that prevent post-Golgi trafficking of the receptor remained enigmatic (Pfeffer 2007).

Results obtained from my work spearheaded further refinement of quantitative bioassays to monitor ERD2 function and suggest that many of the earlier findings surrounding the subcellular localisation of ERD2 may have been erroneous due to masking of critical amino acid residues at the ERD2 C-terminus, either by epitope tags or fluorescent proteins. Furthermore, my work strongly suggests that Golgi-residency is essential for the rate-limiting function of ERD2 in mediating K/HDEL-cargo accumulation in the plant ER, and that the findings may not be specific for the plant field alone, but could apply to most if not all of the eukaryotic kingdoms. The following discussion places the most important advances and findings within the context of the field and explores how further work can shed more light on the actual mechanism that mediates accumulation of soluble proteins in the ER lumen.

C.1 *Nicotiana benthamiana* an unexpected new player coming from the bench

The initial reasoning for the use *Nicotiana benthamiana* was solely based on the advantages from a readily available genome sequence and the opportunity of heterologous expression of ERD2. Thus, I set to optimize a routinely ERD2 bioassay to be used with *Nicotiana benthamiana* protoplasts. In addition to that I was able to establish a gene silencing protocol to simultaneously knock-down the expression of both ERD2 isoforms found in benthamiana. A delightful surprise, not anticipated, were the results indicating that *Nicotiana benthamiana* possibly have a more active secretory pathway.

This plant species is a rising star, becoming a standard model in different plant biology topics, a summary of these was well pointed by Goodin and colaborators (2008). One of the main advantages of *N. benthamiana* is the high amount of biomass produced combined with fast seeding (Leuzinger et al. 2013). *Nicotiana benthamiana* has also been used as a platform for the biosynthesis of pharmaceuticals (van Herpen et al. 2010) and antibodies (Giritch et al. 2006)

The achievement of food security, particularly by a boost in quality and distribution, is an increasing concern worldwide. The manipulation of the secretory pathway to allow for accumulation of high protein levels in specific plant tissues is an interesting concept to pursue in the future aiming to solve this problem. The fact that ERD2 can efficiently mediate the retention of artificially tagged HDEL proteins in the ER, summed to the overall elevated protein synthesis in *N. benthamiana,* are good arguments for future research tailored considering this species as a platform for the production and accumulation of elevated levels of nutritionally relevant proteins. Thereafter, based on my findings I strongly advocate for the use of *N. benthamiana* not just for studying proteins synthesis and transport but also to explore biotechnology research.

C.2 **Recycling the recycling principle, bouncing the ligands.**

The recycling mechanism, widely applied to protein receptors, is mostly based on the *modus operandi* of the MPRs, whereas the ligand binding and release occurs in separate compartments and once released the ligands do not return to their point of origin. The established model explaining ERD2 function postulates that it binds to ligands at the Golgi, undergoes retrograde transport, releases ligands in the ER and finally returns to the Golgi for a new cycle. A crucial problem is that, in contrast to other protein receptors, ERD2 releases its ligands in their point of origin where they are joined by *de novo* synthesised cargo. Thus, after each transport cycle the situation would be worse, and ERD2 would be always in a constant and increasing demand. This issue has not be discussed in the literature. Furthermore, radioactive labelling and stoichiometric experiments analysing the proportion between redistributed HDEL cargo and ectopically introduced ERD2 established a minimum ratio of 200:1 (An 2015). The classical recycling principle combined with cargo offload at the same site of *de novo* synthesis cannot explain this very high ratio. The putative ERD2 binding-site has

been debated many times and no common ground is established, as discussed before.

Data presented in this thesis indicate that an active receptor steady-state is Golgi-only, independent of the presence of ligands (Figure 12), and that partial ER localisation caused by mutagenesis of a conserved di-leucine motif (LLGG) at the ERD2 C-terminus is accompanied by loss-of-function (Figure 9 and Figure 15). I show that an unobstructed and unaltered C-terminus is crucial for ERD2 function and for that reason I would suggest that the use of C-terminally tagged ERD2 in most of previous reports could be the reason for the discrepancy in subcellular localisation. Interestingly, the LLGG mutation did not cause inhibition of ERD2 export from the ER to the Golgi apparatus, instead it appeared to accelerate retrograde transport (Figure 19). Therefore, the very transport step that earlier was associated with receptor-function (Lewis & Pelham 1992a; Montesinos et al. 2014) is now shown to be detrimental to receptor function. Finally, slow turnover of wild type ERD2 in the Golgi was not only seen from photo-bleaching results, but also indicated by the mild effect of the anti-sense inhibition even after 48 hours of severe overexpression. The results suggest a very high intrinsic stability of ERD2, which would suit a protein that is not expected to leave the Golgi once it has arrived. It will be interesting to evaluate ERD2 promoter strength in the future and compare the strength with that of other promoters, for instance the plant VSRs.

When comparing the behaviour of ERD2 with plant VSRs, the difference could not be more dramatic. VSR is mainly detected in the compartment where it is expected to release its ligands (Foresti et al. 2010), unless mutation of anterograde or retrograde transport signals shifts the balance to an earlier or later compartment. In contrast, ERD2 is mainly found in the Golgi, the compartment thought to be the place of ligand-binding. In addition, redirection of VSRs to the cell surface has been documented to lead to secretion of its vacuolar ligands (Gershlick et al. 2014), a strong indication of *in vivo* ligand binding, confirming the *in vitro* binding activity with purified VSRs (Kirsch et al. 1996). However, the biologically active fluorescent ERD2 fusion used in this study was never co-localised with its HDEL ligands despite overexpression of the latter (Figure 12), and if ERD2 would bind to its ligands one would assume at least

some level of co-localisation, particularly if binding occurs in the punctate, and thus intense fluorescent, Golgi bodies.

I believe that the use of the classical recycling mechanism to explain ERD2 function can be erroneous at some extent and with the new tools generated in this thesis it should be re-visited and re-analysed. ERD2 is definitely a rate-limiting factor in the retention of soluble proteins, otherwise ectopic expression would not have any effect. It would be interesting to see if pull-down experiments using YFP-TM-ERD2 can show the direct interaction to HDEL cargo. Perhaps ERD2 does not participate directly in the recycling of cargoes, but instead acts similarly to a bouncer at the entrance of a night-club, ready at the early stacks of the Golgi to identify and prevent the entrance of ER-lumenal residents. How such a mechanism would work, and how it would result in retrograde transport back to the ER is of course totally unclear. If the “bouncer” theory is correct, however, it would explain why it is essential that ERD2 remains at the Golgi entrance, rather than to accompany inappropriate Golgi “guests” back to the ER. Perhaps the arguments and data produced in the past proposing ERD2 oligomerization could be linked to the formation of a matrix or “net” by the receptor-receptor interactions to avoid cargo offload at the *cis*-Golgi, maybe preventing vesicle fusion.

C.3 Synthetic biology, an attempt to reengineer the secretory pathway

The use of signal motifs in conjunction to protein receptors, to induce the retention of attractive exogenous proteins in plant’s natural compartmentalization, is a process quite well understood and a well-used strategy. The re-direction of plant VSRs to the cell surface to mediate vacuolar protein secretion is a good example of this (Gershlick et al. 2014). Likewise Di Sansebastiano et al. (2014) also re-engineered the normal pathway of cargo molecules to trap them in a new intracellular compartment, which is very interesting because it shows how malleable plant secretory pathway can be and how fascinating is the use of synthetic biology. It is also very important to realise that the entire pathway is thought to be reversible (Pelham et al. 1992).

As an initial attempt to use synthetic biology to re-engineer the secretory pathway I tried to redirect the H/KDEL receptor to different intra-cellular compartments.

Surprisingly, the addition of a well-known vacuolar targeting sequence to ERD2-core was not sufficient to causes the re-distribution from the Golgi to later sub-cellular compartments, at least not at high enough levels to allow for visualisation using CLSM. In contrast to that, the observed distribution was a dual ER-Golgi and for that reason I expected to cause complete loss-of-function. However, for the first time a mis-localised receptor was able to mediate the reduction of AmyHDEL secretion (Figure 21), the reasons behind that effect are not clear and need to be further explored to determine if the reduction was caused by real ER retention or perhaps targeting to the vacuole for degradation at very low levels.

In addition to that strategy, my second attempt to re-engineer ERD2 pathway was by the using the C-terminus of a P24 protein that contained signals for COPII mediated ER export as well as a canonical KKXX motif for COPI-mediated recycling from the Golgi back to the ER. If ERD2 behaves like a recycling receptor than this tail should have been fit for purpose. However, as a result, the ERD2 fusion activity was drastically compromised (Figure23) and totally ER retained.

Nevertheless, the most exciting find coming form that strategy lays on the observation that the KKXX motif present at the C-terminus of plant P24 proteins could be functioning in a different way when compared to the canonical KKXX motif of animal proteins, whereas the last two amino acids where shown to be irrelevant. In my study, the two aliphatic residues in the XX position were important in the targeting of a type I membrane protein (Figure 25) even though at the ERD2-C-terminus they seemed irrelevant (Figure 26). Further experiments are necessary to characterize better the nature of this motif, for instance by the mutation of the hypothetical KK$\Psi\Psi$ motif in other P24 members. It is also fascinating that the ERD2 core itself appears to be important in mediating ERD2 retention in the Golgi, in conjunction with its C-terminal di-leucine motif, and further work on potential kin-recognition or protein-protein interactions with other Golgi residents will be extremely interesting in the future.

C.4 **Protein turnover, vacuolar sorting and Golgi tubules**

As mentioned before increased proteins misfolding leads to UPR (Hetz et al. 2011) and synthesis of chaperones (Smith et al. 2011). Excessive synthesis of chaperones could possibly lead to a higher demand for ERD2 retrograde

transport and for that reason it is justifiable to consider a link between ERD2 trafficking/function and the UPR. Co-silencing of ERD2 caused increased levels of both BiP mRNA and the protein itself (Xu et al. 2012). More recently ERD2 has also been shown to play a role in stress responses in T cells via the regulation of a phosphatase, PP1 (Kamimura et al. 2015). In summary more and more evidences are emerging demonstrating that ERD2 may play multiple roles in addition to the retrieval of ER lumenal residents.

Unexpectedly, anti-sense inhibition of ERD2 expression caused a very specific increase in the yield of Amy-HDEL but not Amy, despite only 4 additional amino acids in the former (Figure 37). This could not be explained by a general UPR effect and may point towards a specific mechanism to degrade excess HDEL proteins, possibly to deal with overload of chaperone-misfolded protein complexes. Further work will be needed to establish the pathway by which HDEL-protein turnover could be achieved.

One candidate for protein turnover is post-Golgi trafficking to the lytic vacuole which is well stablished and has been extensively studied. Nevertheless, new information is constantly emerging, for instance, recently it has been proposed that VSRs bind to their ligands in a much earlier stage than what was previously though, at the cis-Golgi (Früholz et al. 2018). The retrieval of ER proteins has been shown to occur as far as the TGN (Miesenböck & Rothman 1995), even though the receptor does not localize to the plant TGN marker SYP61 (Silva-Alvim et al. 2018). If these receptors can be interacting with ligands in the same compartments there is a possibility that competition could cause problems, as suggested earlier using cargo molecules bearing both vacuolar sorting and ER retention signals (Gershlick et al. 2014). The fact that ERD2 overexpression can induce the secretion of vacuolar cargo (Figure 39) could match the induced secretion of HDEL cargo via anti-sense inhibition, if a portion of HDEL proteins are targeted to the vacuoles. Also the fact that ERD2 overexpression can drive ARF1 way from the Golgi to the cytosol is another evidence indicating a dependency of ER-Golgi-ER and post-Golgi trafficking. I strongly believe that the observations from Figures 39 and 40 are linked and most likely due to the fact that ARF1 controls the formation of both COPI and clathrin coated vesicles, retrograde transport and vacuolar transport respectively (Cevher-Keskin 2013). In summary, perhaps ERD2 disruption of vacuolar targeting is an indirect effect

caused by the reduced levels of ARF1, which could lead to a reduction of Clathrin mediated transport from the Golgi to the vacuole.

Another interesting, but not much understood aspect of the Golgi, are the tubular formations emanating from it, which have been initially observed two decades ago (Sciaky et al. 1997). Yet, what triggers the formation of these tubular emanations as well as their functions is still up to debate, particularly in plants (Martínez-Menárguez 2013; Martínez-Alonso et al. 2013; Weidman 1995). The findings from Figures 42 and 43 point that the chimeric ERD2-dimer can be an interesting tool to further understand the properties of Golgi tubules due to its enhanced tabulation properties.

C.5 **Concluding remarks**

The main result of this dissertation is the fact that ERD2 contains a Golgi retention motif which is crucial for its function, and that various findings argue against the canonical receptor recycling model. One of the key questions the field should ask is if ERD2 really acts as a receptor that binds and releases ligands, and that shuttles between donor and acceptor compartments. The tools generated in this thesis may help to provide answers to these questions.

D **Material and methods**

D.1 **Buffers and solutions**

- LB (Luria Bertani) medium: 10 g/L Bacto-tryptone; 5 g/L Bacto-yeast extract; 10 g/L NaCl. Autoclave sterile. For solid medium 15 g/l Bacto-Agar was added prior autoclaving
- 2xYT medium: 16 g/L Bacto-tryptone; 10 g/L Bacto-yeast extract; 5 g/L NaCl. pH 7.0 adjusted with NaOH. Autoclave sterile
- TFBI solution: 30 mM $KC_2H_3O_2$; 100 mM RbCl; 10 mM $CaCl_2 \cdot 2H_2O$; 50 mM $MnCl_2 \cdot 4H_2O$; 15 % v/v glycerol. pH 5.8 using 0.2 M CH_3COOH. Filter sterilised and stored at +4°C
- TFBII solution: 10 mM MOPS; 10 mM RbCl; 75 mM $CaCl_2 \cdot 2H_2O$;15 % v/v glycerol. pH 6.6 using 5 M KOH. Filter sterilised and stored at +4°C
- TES solution: 10mM Tris HCl pH8.0, 5mM EDTA, 250mM sucrose; filter sterile
- TE buffer: 10mM Tris HCl pH8.0, 0.1mM EDTA
- TEX buffer: B5 salts, 500 mg/l MES, 750 mg/l $CaCl_2 \cdot 2H_2O$, 250 mg/l NH_4NO_3, and 0.4 M sucrose [13.7%], brought to pH 5.7 with KOH
- Electroporation buffer: 0.4 M sucrose [13.7%], 2.4 g/l HEPES, 6 g/l KCl, and 600 mg/l $CaCl_2$, brought to pH 7.2 with KOH
- α-amylase extraction buffer: 50 mM malic acid, 50 mM NaCl, 2 mM $CaCl_2$ and 0.02% sodium azide, 0.02% BSA
- To 10ml of GUS Extraction buffer: 5ml sodium phosphate buffer pH7.0, 1ml Na_2EDTA, 1ml 0.1% sodium lauryl sarcosine, 0.1ml 0.1% Triton X-100 and 7.8µl βMeEtOH added prior to use.
- To 5ml of GUS Reaction buffer: 2.5ml sodium phosphate buffer pH7.0, 0.1ml 0.1% Triton, 0.5ml PNPG, 3.9µl βMeEtOH added prior to use.
- Leaf extraction buffer: 100mM Tris HCl, pH 7.8, 200mMNaCl, 1mMEDTA, 0.2% Triton X-100, and 2% β-mercaptoethanol)
- ECL Solution 1: 1 ml 1 M Tris HCl (pH 8.5), 100 µL 250 mM Luminol, 44 µL 90 mM p-coumaric acid, 8.85 mL dH_2O
- ECL Solution 2: 6 µl 30 % H_2O_2, 1 mL 1 M Tris HCl pH 8.5, 9 mL dH_2O

D.2 **Molecular biology techniques**

DNA manipulations performed according to well established procedures for molecular biology, and unless stated all media and buffers were prepared according to Sambrook et al. (1989). Agarose gels were made up in 0.5x TBE (Tris, boric acid and EDTA) buffer and restriction digestion were carried out in TE buffer supplemented with restriction buffers compatible to the enzymes, normally *Cutsmart*, as recommended by the manufacturer. All the restriction enzymes were purchased from New England Biolabs. *E.coli* strain MC1061 was used for all construct amplifications (Casadaban & Cohen 1980).

PCR amplifications were set up by the KOD DNA polymerase protocol from Novagen, Darmstadt, Germany (Novagen, 2011). Oligonucleotides were purchased from Eurogentec (Liege, Belgium). DNA was amplified using a thermocycler (GeneCycler BioRad, Hercules, CA, USA) and conditions were adjusted according to the reaction. Quick-change protocol was used for all mutagenesis carried out during the thesis and were performed by using primer pairs designed as necessary and a typical PCR cycling as following:

1. Initial denaturing 95°C, 2 mins
2. Denaturing 95°C, 20sec
3. Annealing 55°C, 20sec
4. Extension 72°C, 60sec
5. Final extension 72°C, 10mins

Step 2-4 is then repeated for 15-25 cycles depending on the amplification product. Followed by a 2 hours digestion using DpNI restriction enzyme to eliminate parent plasmid and subsequent transformation of competent cells.

For the generation of chimeric proteins anti-sense primers specifically designed for the replacement of the C-terminal regions were used in combination with sense primers upstream the promoters. A typical PCR cycling condition for amplification is as following:

1. Initial denaturing 95°C, 2 mins
2. Denaturing 95°C, 20sec
3. Annealing 50°C, 20sec
4. Extension 72°C, 60sec
5. Final extension 72°C, 5mins

Step 2-4 is then repeated for 30 cycles depending on the amplification product. PCR products were digested using necessary restriction sites for further ligation into desired vectors followed by transformation of competent cells.

D.2.1 DNA preparations

To make mini-preparations (miniprep) fresh overnight liquid cultures, from a single colony that was inoculated in 3 ml of LB medium the night before, were used to fill approximately 1.5 ml of a labelled Eppendorf tube. Tubes were centrifuged for 1 minute at 14,000 rpm and the supernatant was removed. As soon as the pellets were re-suspended in 150 µl of TES, 20 µl of lysozyme solution was quickly added and tubes were incubated for 5 minutes at room temperature. Subsequently 300 µl of distilled water was added to the suspensions and tubes were incubated at 73°C for 15 minutes. Afterwards, the tubes were centrifuged at 14,000 rpm for 15 minutes and the supernatants were recovered in new labelled tubes. 50 µl of 5 M $NaClO_4$ was added to the supernatants, usually 500 µl and if not the volume was completed to with TE, and the tubes were vigorously shaken. 400 µl of isopropanol was then added and mixed followed by another 15 minutes centrifugation at 14,000 rpm. Supernatants were removed as previously and the “empty” tubes were further centrifuged for 2 minutes to remove the leftover liquid in the tube. Finally, tubes were dried at 37°C with open caps for 15 minutes and pellets were then re-suspended in 50 µl of TE.

To yield higher DNA concentrations and purity for sequencing the protocol was scaled-up to wizard-preparation (wizprep) from a single colony that was inoculated in 10 ml of LB medium the night before and Promega Wizard® *PLUS* SV was used for DNA extraction and purification.

To generate DNA maxi-preparations (maxiprep), to be used for transient expression in the bioassay, a single colony was used as pre-inoculum in 3 ml of LB, grown for 3 hours before final inoculation to be grown 24 hours in 500 ml of LB followed by purification as before.

D.2.2 Recombinant DNA plasmid

Constructs used in this thesis, both new constructs and pre-existing in the host laboratory, are listed in table 2. All plasmids were check by qualitative digest and sequencing by Source BioScience©. Constructs indicated by an asterisk (*) I have used gene synthesis services by Eurofins Genomics, sequences for genes

coding regions were selected from online databases (NCBI) and designed to be delivered in pUC57 vectors with specific restriction sites flanking both ends (ClaI and BamHI).

For the bioassay analysis all gene were sub-cloned into an existing pJA31 vector via classical cloning, substituting ERD2b gene; this is a double expression vector with a GUS internal marker as described in (Gershlick et al. 2014).

For transient expression via leaf infiltration a pre-existing plant vector was already available in the host laboratory and was used a backbone to receive or generate fluorescent fusions via either EcoRI + HindIII or ClaI + BamHI restriction sites.

Construction of the triple expression vector is explained in recent published work (Silva-Alvim et al. 2018) and was done via the modification of pGUSRef by the insertion of AmyHDEL coding region under the control of the CaMV35S promoter, elimination of unnecessary inconvenient restriction sites, gene synthesis of the *Arabidopsis thaliana* ADH 3'end (AT1G77120) carrying a polyadenylation signal and a polylinker as well as the modification of the nopaline synthase promoter from pDE1001 to exhibit the necessary restriction site to receive ERD2b and ERD2b mutants.

For the expression in mammalian cells lines YFP-TM-ERD2 and ERD2-YFP coding regions were sub-cloned using classical restriction digest into pCDNA3.1 vector, plasmids were amplified and purified as others in this thesis to later be used for transfection.

Table 2 List of constructs used in this project.

Pre-existing and new constructs are listed bellow

Description	Reference/Generated by
35S:α-amylase	(Crofts et al. 1999)
35S:α-amylase-HDEL	(Phillipson et al. 2001)
35S:α-amylase-SPORAMIN	(Pimpl et al. 2003)
35S:α-amylase-KDEL	(Phillipson et al. 2001)
TR2:GUS 35S: α-amylase	(Adam 2013)
TR2:GUS 35S:ERD2b	(An 2015)
TR2:GUS 35S:NbERD2ab sense	J. Alvim*
TR2:GUS 35S:NbERD2ab antisense	J. Alvim*
TR2:GUS pNOS:ERD2b	B. Lee
TR2:GUS 35S:α-amylase pNOS ERD2b	J. An

TR2:GUS 35S:α-amylase pNOS ERD2b L211G^L213G	J. An
TR2:GUS 35S:α-amylase pNOS ERD2b F4S	J. Alvim
TR2:GUS 35S:α-amylase pNOS ERD2b D89A	J. Alvim
TR2:GUS 35S:α-amylase pNOS ERD2b E91A	J. Alvim
TR2:GUS 35S:α-amylase pNOS ERD2b D93N	J. Alvim
TR2:GUS 35S:α-amylase pNOS ERD2b F95A	J. Alvim
TR2:GUS 35S:α-amylase pNOS ERD2b P135A	J. Alvim
TR2:GUS 35S:α-amylase pNOS ERD2b Q136A	J. Alvim
TR2:GUS 35S:α-amylase pNOS ERD2b L137A	J. Alvim
TR2:GUS 35S:α-amylase pNOS ERD2b W168A	J. Alvim
TR2:GUS 35S:α-amylase pNOS ERD2b Y164A^N167A	J. Alvim
TR2:GUS 35S:α-amylase pNOS ERD2b F196A^F197A	J. Alvim
TR2:GUS 35S:ERP1	J. An
TR2:GUS 35S:ERD2b::ERP1tail	J. Alvim
TR2:GUS 35S:ERD2::ERP1-TM6/7	J. Alvim
TR2:GUS 35S:TM-ERD2b	(Silva-Alvim et al. 2018)
TR2:YFP-ERP1	(Silva-Alvim et al. 2018)
TR2:ERD2-TM-RFP	(Silva-Alvim et al. 2018)
TR2:RFP-TM-ERD2	(Silva-Alvim et al. 2018)
TR2:YFP-TM-ERD2	(Silva-Alvim et al. 2018)
TR2:ST-YFP-HDEL + 35S:α-amylase	(An 2015)
TR2:ST-YFP-HDEL + 35S:α-amylase-HDEL	(An 2015)
TR2:secYFP-ERD2	(Silva-Alvim et al. 2018)
TR2:YFP-TM-ERD2 L211G^L213G	J. Alvim
TR2:YFP-TM-ERD2 ΔC5	J. Alvim
TR2:YFP-TM-ERD2 R5A	J. Alvim
TR2:YFP-TM-ERD2 Y164A	J. Alvim
TR2:RFP-TM-ERD2F4S	J. Alvim
TR2:RFP-TM-ERD2D89A	J. Alvim
TR2:RFP-TM-ERD2P135A	J. Alvim
TR2:RFP-TM-ERD2Q136A	J. Alvim
TR2:RFP-TM-ERD2L137A	J. Alvim
TR2:RFP-TM-ERD2 Y164A^N167A	J. Alvim
TR2:RFP-TM-ERD2W168A	J. Alvim
TR2:RFP-TM-ERD2F196A^F197A	J. Alvim
TR2:secYFP-ERD2b-TM7	J. Alvim
TR2:secRFP-VSR2tm::ERD2btail	J. Alvim
TR2:GUS 35S:ERD2b::VSR2tail	E. Talbot

TR2:YFP-TM-ERD2::VSR2tail	J. Alvim
TR2:GUS 35S:ERD2b::p24tail	J. Alvim
TR2:RFP-TM-ERD2::p24tail	J. Alvim
35S:secRFP-p24aTM	J. Alvim
35S:secRFP-p24aTM KKSS	J. Alvim
35S:secRFP-p24aTM LIGG	J. Alvim
TR2:RFP-TM-ERD2::p24tail KKSS	J. Alvim
TR2:RFP-TM-ERD2::p24tail LIGG	J. Alvim
TR2:ST-YFP	(Bottanelli et al. 2012)
TR2:ST-RFP	(An 2015)
35S:YFP-HDEL	(Gershlick et al. 2014)
35S:RFP-HDEL	(Gershlick et al. 2014)
TR2:GUS 35S:oiERD2	J. Alvim*
TR2:GUS 35S:acERD2	J. Alvim*
TR2:GUS 35S:piERD2	J. Alvim*
TR2:GUS 35S:ccERD2	J. Alvim*
TR2:GUS 35S:gsERD2	J. Alvim*
TR2:GUS 35S:hERD2	J. Alvim*
TR2:GUS 35S:taERD2	J. Alvim*
TR2:GUS 35S:tpERD2	J. Alvim*
TR2:GUS 35S:pgERD2	J. Alvim*
TR2:GUS 35S:klERD2	J. Alvim*
TR2:GUS 35S:tbERD2	J. Alvim*
TR2:GUS 35S:yERD2	J. Alvim*
TR2:GUS 35S:atERD2b-HA	(An 2015)
TR2:GUS 35S:yERD2-HA	J. Alvim
TR2:GUS 35S:ERD2 Yeast::ptail	J. Alvim
TR2:GUS 35S:ERD2 Plant::ytail	E. Talbot
TR2:GUS 35S:ERD2 Plant::Yeast H1	J. Alvim
TR2:GUS 35S:ERD2 Plant::Yeast H2	J. Alvim
TR2:RFP-TM-yERD2	J. Alvim
TR2:GUS 35S:hERD2 S209A	J. Ranger
TR2:GUS 35S:hERD2 S209D	J. Ranger
TR2:GUS 35S:hERD2 L208G^L210G	J. Ranger
CMV:YFP-TM-ERD2b	J. Alvim
CMV:ERD2b-YFP	J. Alvim
CMV: α-amylase	J. Alvim
CMV: α-amylase-HDEL	J. Alvim
TR2:YFP-1TM-ERD2b + 35S:NbERD2ab sense	J. Alvim
TR2:YFP-1TM-ERD2b + 35S:NbERD2ab antisense	J. Alvim
TR2:ST-RFP + 35S:ERD2b	(An 2015)

TR2:ST-RFP + 35S:PAT	(An 2015)
TR2:Sar1-Venus	P. Mansot
TR2:ARF1-RFP	P. Mansot
TR2:GUS 35S:pg::at ERD2	J. Alvim
TR2:spYFP-pg::at ERD2	N. Bhatia

Abbreviations: 35S: Cauliflower mosaic virus promoter; 3'nos: 3' untranslated end of nopaline synthase gene; TR2': TR-DNA derived mas 2'; 3'ocs: 3' untranslated end of the octopine synthase gene; PAT: phosphinothricin acetyl transferase; pNOS: nopaline-synthase gene promoter; 3'ADH: *Arabidopsis thaliana* ADH 3'end (AT1G77120). CMV: cytomegalovirus immediate-early enhancer and promoter.

D.2.3 Preparation of *E. coli* competent cells

A fresh aliquot of MC1061 *E.coli* strain was streaked out on a LB-agar plate and let to grow overnight at 37°C. 3 ml of 2xYT medium was inoculated with a fresh colony and incubated at 37°C with vigorous shaking (200 rpm). Turbidity of pre-inoculum was checked over time and when turbid inoculated into a final culture of 200 ml of pre-warmed 2xYT medium and incubated as before at 37°C. When the culture reached an O.D550 of approximately 0.400-0.450 was transferred into four sterile 50 ml conical tubes. These were placed on ice for 5 minutes to arrest cell division. In a cold room, at 4°C, culture aliquots were centrifuged at 3000g in a refrigerated swing-out rotor, for 20 minutes. The cell pellets were re-suspended in a total of 80 ml of ice-cold TFBI solution and placed on ice for 5 minutes. The cell suspension was centrifuged as before and then re-suspended in a total of 8 ml of TFBII, pooled and left on ice for 15 minutes. Using pre-chilled pipettes tips, 100 µl aliquots of homogenous cell suspension were transferred to pre-chilled microfuge tubes (placed on ice). Aliquots were then frozen in dry ice and stored at -80°C.

D.3 Plant material and transient expression experiments

Sterile grown *Nicotiana tabacum* cv., Petit Havana (Maliga et al. 1973) and *Nicotiana benthamiana* (Goodin et al. 2008) plants were grown from surface-sterilized seeds. Plants were used for *Agrobacterium tumefaciens* mediated leaf infiltration experiments and α-amylase assay as described in previously published protocol (Leborgne-Castel et al. 1999; Foresti et al. 2006). For *Nicotiana benthamiana* transient expression of transformed protoplasts the

methodology was adapted as described in Chapter 1, otherwise all protoplasts experiment were performed as previously published protocol (Denecke & Vitale 1995). For anti-sense inhibition and complementation analysis, protoplasts harvesting was modified as described further. Detailed descriptions of the procedures are as following.

D.3.1 Preparation of protoplasts

Tobacco leaf protoplasts were prepared with supplement of 1x digestion mix which was prepared from TEX buffer supplemented with 0.2% Macerozyme R10 and 0.4% Cellulase R10 (Yakult). Stocks with 10-fold concentrated digestion enzymes were prepared by dissolving the lyophilized powders in TEX buffer for 2 h, followed by centrifugation at 5000 *g* for 15 minutes and filter sterilization (0.2 µm) of the clear supernatant. The filtered supernatant was aliquot in 5 ml and kept at −80°C for routine use. The 1x digestion mix was always prepared freshly by addition of 45 ml of TEX buffer to these stocks.

Overnight digestions of floating leaves were prepared by using a needle bed. These digestions were then filtered through a 100-µm nylon mesh and brief washed with electroporation buffer to release further protoplasts from the tissue remnants. The protoplast suspensions were then centrifuged in Falcon tubes (50 ml) for 15 minutes at 100 g at room temperature in a swing-out rotor. Centrifugation was stopped without brake to prevent re-suspension of the floating protoplast band. The pellet and the underlying medium were removed and discarded using a peristaltic pump and a sterile Pasteur pipette until the band of floating living protoplasts reached the bottom. Then the cells were re-suspended in 25 ml of electroporation buffer and a further centrifugation at 100g for 10 minutes was initiated. The pellet and the underlying medium were removed again and this procedure was repeated twice.

D.3.2 Electroporation of protoplasts

After the final wash, protoplasts were re-suspended in electroporation buffer at an expected concentration of 5×10^6 protoplasts/ml. 500 µl of the obtained protoplasts mix was then pipetted into a disposable 1 ml plastic cuvette and was added to it a total volume of 100 µl, consisting of variable volumes of effector plasmid DNA (gene of interest) plus 50 µl of cargo molecule (1:10 dilution of maxiprep plasmid DNA of Amy, AmyKDEL, AmySpo or AmyHDEL) and the difference to 100 µl with electroporation buffer. The protoplast suspensions were

then incubated for 5 minutes and electroporated for 5 seconds with stainless steel electrodes at a distance of 3.5 mm. A complete exponential discharge of a 1000-µF capacitor charged at 160 V was connected to the electrodes. Electroporated protoplasts were rested for 30 minutes and were then removed from the cuvettes by washing in 1 ml of TEX buffer twice and transferred to 5 cm Petri dishes. All incubations were performed for at least 24h, otherwise specified. For *Nicotiana benthamiana* a optimization was necessary and it is explained in details in Chapter 1.

D.3.3 Harvesting of electroporated protoplast

After incubation, the protoplast where harvest for different experiments. For the GUS experiment 500 µl where recovered into a sterile Eppendorf tube for further analysis (see GUS-normalized assay). The difference in the volume, 2.0 ml of the cell suspension, was transferred into a small clear Falcon tube (10 ml) and used for either protein extraction or Alpha-amylase assay. Protoplast suspension was centrifuged at 100g for 5 minutes. Approximately 500 µl of the underlying medium was manually removed with a refined Pasteur pipette. This obtained medium was further cleared by centrifugation in a refrigerated microfuge (4°C, 14000 rpm, and 10 minutes) and was kept on ice for further analysis (see Alpha-amylase assay). The cells were diluted 10-fold with 250 µM NaCl in 10 ml Falcon tubes to recover the total cell population of the remaining suspension. Hence the suspension was centrifuged for 3 minutes at 150g. The supernatant was then removed with a peristaltic pump and the compact cell pellet was kept on ice for subsequent extraction and analysis (see Alpha-amylase assay).

All procedures were the same and constantly used thought the thesis. However, after initial tests is was necessary to modify the standard protocol for anti-sense inhibition assays and complementation assays. In those cases protoplast were incubated for an additional period of 24 hours, to a total of 48 hours. A washing step was added after the initial 24 hours and protoplast were re-suspended in fresh TEX media to allow for the accumulation of secreted cargo only after the knock-down of the receptor was in place.

D.3.4 Alpha-amylase assay

Alpha-amylase assay reagents were purchased from Megazyme (http://secure.megazyme.com). The protoplast samples of centrifuged medium suspensions from the harvesting procedure were extracted and diluted with α-

amylase extraction buffer to obtain suitable dilutions for the assay. In contrast, the cell pellet samples was re-suspended with the same α-amylase extraction buffer, to a total volume of 1ml, and subsequently sonicated for 5 seconds (130W, 20 KHz – 60% amplitude) and vortexed. The sonicated cell pellets were then centrifuged for 10 minutes at 14000 rpm at 4°C and the supernatants were recovered.

Sample extracts were on ice all the time between assays. The assays were carried out at 45°C using 30 µl of the medium or cell samples and the appropriate time. The reaction was hence initiated by addition of 30 µl of the substrate (R-CAAR4) consisting of blocked *p*-nitrophenyl maltoheptaoside (BPNPG7, 54.5 mg) and thermostable α-glucosidase (125 units at pH 6.0)) which was dissolved according to the manufacturer's instructions in 10 ml of autoclaved distilled water and stored at −80°C as 1 ml aliquots. The reaction was stopped by the addition of 150 µl of reaction stop buffer (1% (w/v) Tris). Finally the absorbance was measured at 405 nm and readings were recorded. Negative controls for correction of absorbance were obtained from mock-transformed protoplasts.

The α-amylase activity was calculated for medium and cells samples accordantly to the formulas 1 and 2, respectively:

1: (average ∆OD/sample µl)/time*dilution*1000

2: (average ∆OD/sample µl)/time/concentration*dilution*1000

*concentration refers to the re-suspension of cells pellet

Only readings within the linear range between ΔO.D 0.1 and 1.2 were used for calculations to avoid inaccuracy and substrate limitations. The assay was repeated at least three times for each extract including controls, and the average activity was calculated once appropriate dilutions and incubation times were established.

D.3.5 GUS-normalized effector Dose-response assay

The quantification of the GUS activity is used as an internal marker because it makes possible to measure the transfection efficiency of difference plasmids and equalize them. The assay was carried as described by Gershlick et al. (2014), with small modifications. All the reagents used were made accordantly to the previously mentioned paper. 500 µl specifically collected from the transient expression for this assay was immediately mixed with 500 µl of GUS extraction buffer and kept always on ice. The samples were then sonicated 5 seconds (130W, 20 KHz – 60% amplitude), vortexed and centrifuged (refrigerated

microfuge 4°C, 14000 rpm, and 10 minutes). 10 µl of the samples were individually aliquoted in a 96-well microtiter plate and mixed with 100 µl of extraction buffer and 90 µl of reaction buffer. The samples were incubated for at least 16 hours at 37°C. The reaction was stopped by the addition of 80 µl of stop buffer (2.5 M 2-amino-2-methyl propanediol). As a negative control the same sets of sample were prepared but the stop buffer was added to the mixed before the incubation starts.

D.3.6 Tobacco leaf infiltration and microscopy

Tobacco leaves were infiltrated with fresh overnight *Agrobacterium tumefaciens* cultures carrying desired plasmids. Before infiltration concentration of cultures were appropriated adjusted to a fixed optical density and infiltrated leaf areas were analysed after 2 days of further growth via confocal laser scanning microscopy (CLSM), unless otherwise stated in figure legends. Preparation of cultures for infiltration has strictly followed previously published protocol (Sparkes et al. 2006). Fresh colonies were inoculated in 3 ml of MGL and grown for 24 hours at 28ºC. 1 ml of the culture was centrifuged at 5000 rpm, 5 minutes at room temperature. Pellet was re-suspended in 1 ml of the infiltration buffer. Suspension was washed twice more as before. Final suspension was diluted to appropriate concentration to obtain an absorbance OD600 of approximately 0.1.

D.3.6.1 Organelle markers

All organelle markers used have been previously published as follow. Golgi-marker ST-RFP was based on *Agrobacterium tumefaciens* dual expression vector ST-YFP (Bottanelli et al. 2012), except that YFP was replaced by RFP (An 2015). ER markers used were Y/RFP-HDEL (Gershlick et al. 2014).

D.3.6.2 Fluorescence confocal microscope imaging and analysis

Infiltrated tobacco leaf squares (0.5 x 0.5 cm) were mounted in tap water with the lower epidermis facing the thin cover glass (22 x 50 mm; No. 0). Confocal imaging was performed using an upright Zeiss LSM 880 Laser Scanning Microscope (Zeiss) with a PMT or a high-resolution Airyscan detector, a Plan-Apochromat 40x/1.4 oil DIC M27 objective or Plan-Apochromat 63x/1.4 oil DIC M27 objective. When YFP-fusions were imaged alone, the excitation wavelength was 514 nm and fluorescence was detected with a bandpass filter 519-620 nm. When RFP-fusions were imaged alone, the excitation wavelength was 561 nm

and fluorescence was detected with a bandpass filter 585-650 nm. To image YFP-fusions together with RFP-fusions, samples were excited using an Argon ion laser at the wavelength of 488 nm for YFP and a HeNe ion laser at 561 nm for RFP. A 488/543 dichroic beam splitter was used to detect fluorescence, YFP fluorescence was detected with a bandpass filter 493-529 nm and RFP fluorescence was detected with a bandpass filter 585-650 nm. All dual colour imaging was performed by line switching to obtain adequate live bio-imaging data that are not distorted by organelle motion. Post-acquisition image processing was performed with the Zen 2.3 lite blue edition (Zeiss) and ImageJ (Collins 2007; Schneider et al. 2012; Rueden et al. 2017). Image analysis was undertaken using the ImageJ analysis program and the PSC co-localization plug-in (French et al. 2008) to calculate co-localization and to produce scatter plots as described before (Foresti et al. 2010).

Samples to be used in fluorescence recovery after photobleaching (FRAP) studies were pre-treated to promote Golgi lock-down before analysis. Confocal imaging was performed using an upright Zeiss LSM 880 Laser Scanning Microscope (Zeiss) with a PMT detector and a Plan-Apochromat 40x/1.4 oil DIC M27 objective. Zen 2.3 black edition (Zeiss) software was used to record pre and post-bleached signals and to modulate laser beam intensity. Signals were sampled before bleach treatment using standard confocal setting as described before. Bleaching was achieved by scanning with high-intensity illumination of selected regions of interest (ROI) and every 30 seconds after bleaching with low-intensity illumination following recommendation of previously published protocol (Brandizzi et al. 2002).

D.4 **Protein extraction and western blot**

To allow for the simultaneous measurement of GUS internal marker activity and to measure synthesis of ERD2-Ha fusions (Chapter 4) the standard harvesting of protoplast, previously described, was slightly modified. For the GUS experiment 500 µl of electroporated protoplast were recovered into a sterile Eppendorf and analysed as described before. The difference in the volume, 2 ml of the cell suspension, was pelleted by the addition of 250mM NaCl for a final volume of 10 ml. After centrifugation supernatant was removed and pellet resuspended in 250 µL of leaf extraction buffer, followed by brief sonication and

10 min centrifugation at 19,000 g at 4°C. The supernatant was used for protein gel blotting analysis. and probed with antibodies against HA-tag (Foresti et al. 2006; daSilva et al. 2006).

A pre-stained protein ladder ranging from 10 to 160 kDa (Fermentas Life Science) was used as a molecular weight marker. Protein extracts were loaded in equal volumes after two fold dilution with 2X SDS loading buffer and brief boiling (5 minutes, 95°C). Electrophoresis was performed in running buffer and at a limiting current of 40 mA, and a voltage of 200 V. Proteins were transferred onto nitrocellulose membranes via electroblotting, and transference was checked through Ponceau staining. Washing with 1x Phosphate buffered saline (PBS)+0.5% tween20 was performed several times and the membrane was subsequently incubated in the blocking solution (0.5% tween20, 5% milk powdered in PBS) at room temperature for 2 hours with slow agitation. PBS+0.5% tween20 and PBS washing was performed to removal of blocking solution. Primary antibody, diluted in a 1% BSA solution with 0.02% sodium azide in 1x PBS was added. Rabbit monoclonal antiserum rose against HA (1:5000 dilutions, Molecular Probes Inc.) was used. Incubation with anti-HA was performed overnight at 4°C and the membrane was washed once again with PBS+0.5% tween20 previously to addition of secondary antibody. Immunodetection was performed using enhanced chemiluminescence with freshly prepared ECL solutions 1 and 2. The solutions were mixed over the membranes followed by 5 minutes incubation and exposure to x-ray films.

D.4.1 Bio-rad assay

5µl of undiluted protein extracts from protoplast were diluted with 155 µl of autoclaved water and were then assayed at room temperature with 40 µl of Bio-rad reagent (contained phosphoric acid and methanol). The absorbance was then measured at 605 nm and readings were recorded.

D.5 **Drug treatment**

Infiltrated tobacco leaf squares (0.5 x 0.5 cm) were used for drug treatment preceding confocal imaging for photobleaching recovery studies, FRAP. To promote actin depolymerization and top Golgi movement treatment with latrunculin B followed previously published protocol modified as necessary

(Brandizzi et al. 2002). Samples were submerged in 12µM solution of the latrunculin B (Cayman Chemical Co.) in water for one hour and analysed soon after.

D.6 Mammalian expression and confocal laser scanning microscopy

All experimental procedures involving the expression of *Arabidopsis thaliana* ERD2 fluorescent fusions in mammalian cells were conducted in collaboration with the laboratory coordinated by professor Sreenivasan Ponnambalam at the University of Leeds.

Human HEK-293T (human embryo kidney cell line 293) cells were grown, transformed, cultured, fixed and the immunostaining was performed at his laboratory using standard procedures according to published work (Ponnambalam et al. 1996; Bruns et al. 2010; Jopling et al. 2011).

Cells are maintained using High Glucose DMEM (Invitrogen) with 10% FCS, glutamine, non-essential amino acids and penicillin/streptomycin. Cells are grown in 10 cm dishes and split at 1:10 every 2-3 days and re-suspended before use to 10-20% confluency for a final volume of 1 ml per transformation. Plasmids DNA were diluted with dH_2O to a final concentration of 5-25 µg and final volume of 439 µl, mixed by vortexing with 61 µl of 2M CaCl per sample. 500 µl of fresh thawed 2x HBS (16 g/l NaCl, 0.4 g/l $Na_2HPO_4.7H_2O$, 13 g/l Hepes) is added drop-wise to obtain a final volume of 1 ml, under constant vortex. DNA mix was added drop-wise and evenly to 6-well dishes containing sterile coverslips and 1ml of resuspended cells mixed with 8 ml of fresh media to obtain a final volume of 10 ml. After 20-24 hours of incubation medium was aspirate and 10 ml of fresh medium was added. 48 hours post-transfection cells were fixed and immunostained. Cells were fixed using methanol.

Monoclonal antibodies used for immunostaining were kindly provided and used in accordance to published work from Prof. Vas laboratory. As organelle markers primary conjugate anti-Calnexin and anti-ERGIC53 (Prescott et al. 2001; Towler et al. 2000) were used. Additional staining with Diamidino-2-phenylindole (DAPI) dye was used as control.

Immunofluorescence analysis was performed and images captured using a Zeiss LSM 880 Laser Scanning Microscope (Zeiss) with a PMT detector, a Plan-

Apochromat 40x/1.4 oil DIC M27 objective. 4′,6-Diamidino-2-phenylindole (DAPI) was excited using a 405-nm laser diode, and signal was collected through a bandpass (BP) 420- to 480-nm filter. To image YFP-fusions together with RFP-organelle markers, samples were excited using an Argon ion laser at the wavelength of 488 nm for YFP and a HeNe ion laser at 561 nm for RFP. A 488/543 dichroic beam splitter was used to detect fluorescence, YFP fluorescence was detected with a bandpass filter 493-529 nm and RFP fluorescence was detected with a bandpass filter 585-650 nm.

D.7 Generation of transgenic plants by leaf-disk transformation

Nicotiana benthamiana (Goodin et al. 2008) plants, were grown in Murashige and Skoog (MS) medium with 2% sucrose in a controlled room at 25°C with a 16-h daylength at a light irradiance of 200 μE $m^{-2}s^{-1}$. Stably transformed plants were obtained by *Agrobacterium infection* of leaf disks as shown in Chapter 5 and described by Denecke et al. (1990). Selection of transformants was accomplished in MS medium supplemented with 3% sucrose and containing 100 μg/mL kanamycin and 250 μg/mL cefotaxime. Regenerated plants were analysed and scored by CLSM.

E **Bibliography**

Aasland, R. et al., 2002. Normalization of nomenclature for peptide motifs as ligands of modular protein domains. *FEBS Letters*, 513(1), pp.141–144.

Adam, I.K., 2013. *Starch-based Bioethanol Process Innovation*. University of Leeds.

Alanen, H.I. et al., 2011. Beyond KDEL: the role of positions 5 and 6 in determining ER localization. *Journal of molecular biology*, 409(3), pp.291–7.

Altschul, S.F. et al., 1997. *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs*, Oxford University Press.

An, J., 2015. *Elucidating the recycling mechanism of ER resident proteins with ERD2*. University of Leeds.

Aoe, T. et al., 1997. The KDEL receptor, ERD2, regulates intracellular traffic by recruiting a GTPase-activating protein for ARF1. *EMBO Journal*, 16(24), pp.7305–7316.

Appenzeller-Herzog, C. & Hauri, H.-P., 2006. The ER-Golgi intermediate compartment (ERGIC): in search of its identity and function. *Journal of cell science*, 119(Pt 11), pp.2173–83.

Balch, W.E. et al., 2008. Adapting Proteostasis for Disease Intervention. *Science*, 319(5865), pp.916–919.

Balch, W.E. et al., 1984. Reconstitution of the transport of protein between successive compartments of the golgi measured by the coupled incorporation of N-acetylglucosamine. *Cell*, 39(2), pp.405–416.

Banfield, D.K., 2011. Mechanisms of Protein Retention in the Golgi. *Cold Spring Harbor Perspectives in Biology*, 3(8), pp.a005264–a005264.

Bar-Peled, M. & Raikhel, N. V, 1997. Characterization of AtSEC12 and AtSAR1 (Proteins Likely Involved in Endoplasmic Reticulum and Golgi Transport). *Plant Physiology*, 114(1), pp.315–324.

Barlowe, C. et al., 1994. COPII: A membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell*, 77(6), pp.895–907.

Baumann, O. & Walz, B., 2001. Endoplasmic reticulum of animal cells and its organization into structural and functional domains. *International review of cytology*, 205, pp.149–214.

Del Bem, L.E. V., 2011. The evolutionary history of calreticulin and calnexin genes in green plants. *Genetica*, 139(2), pp.255–259.

Bergeron, J.J.M. et al., 1994. Calnexin: a membrane-bound chaperone of the endoplasmic reticulum. *Trends in Biochemical Sciences*, 19(3), pp.124–128.

Blobel, G. & Dobberstein, B., 1975. TRANSFER OF PROTEINS ACROSS MEMBRANES I. Presence of Proteolytically Processed and Unprocessed Nascent Immunoglobulin Light Chains On Membrane-Bound Ribosomes of Murine Myeloma. *THE JOURNAL OF CELL BmLOCV*, 9, pp.835–851.

Boevink, P. et al., 1998. Stacks on tracks: The plant Golgi apparatus traffics on an actin/ER network. *Plant Journal*, 15(3), pp.441–447.

Bohlenius, H. et al., 2010. The Multivesicular Body-Localized GTPase ARFA1b/1c Is Important for Callose Deposition and ROR2 Syntaxin-Dependent Preinvasive Basal Defense in Barley. *THE PLANT CELL ONLINE*, 22(11), pp.3831–3844.

Bonifacino, J.S. & Glick, B.S., 2004. The Mechanisms of Vesicle Budding and Fusion. *Cell*, 116(2), pp.153–166.

Booth, C. & Koch, G.L., 1989. Perturbation of cellular calcium induces secretion of luminal ER proteins. *Cell*, 59(4), pp.729–737.

Bortesi, L. & Fischer, R., 2015. The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology Advances*, 33(1), pp.41–52.

Bottanelli, F., Gershlick, D.C. & Denecke, J., 2012. Evidence for Sequential Action of Rab5 and Rab7 GTPases in Prevacuolar Organelle Partitioning. *Traffic*, 13(2), pp.338–354.

Boulaflous, A. et al., 2009. Cytosolic N-terminal arginine-based signals together with a luminal signal target a type II membrane protein to the plant ER. *BMC Plant Biology*, 9(1), p.144.

Bourne, H.R., Sanders, D.A. & McCormick, F., 1990. The GTPase superfamily: a conserved switch for diverse cell functions. *Nature*, 348(6297), pp.125–32.

Bowers, K. & Stevens, T.H., 2005. Protein transport from the late Golgi to the vacuole in the yeast Saccharomyces cerevisiae. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1744(3), pp.438–454.

Brach, T. et al., 2009. Non-invasive topology analysis of membrane proteins in the secretory pathway. *Plant Journal*, 57(3), pp.534–541.

Brandizzi, F. et al., 2002. Membrane protein transport between the endoplasmic reticulum and the Golgi in tobacco leaves is energy dependent but cytoskeleton independent: evidence from selective photobleaching. *The Plant cell*, 14(6), pp.1293–309.

Brandizzi, F., 2002. The Destination for Single-Pass Membrane Proteins Is Influenced Markedly by the Length of the Hydrophobic Domain. *THE PLANT CELL ONLINE*, 14(5), pp.1077–1092.

Brandizzi, F. & Barlowe, C., 2013a. Organization of the ER-Golgi interface for membrane traffic control. *Nature reviews. Molecular cell biology*, 14(6), pp.382–92.

Brandizzi, F. & Barlowe, C., 2013b. Organization of the ER–Golgi interface for membrane traffic control. *Nature Reviews Molecular Cell Biology*, 14(6), pp.382–392.

Brodsky, J.L. & McCracken, A.A., 1999. ER protein quality control and proteasome-mediated protein degradation. *Seminars in Cell & Developmental Biology*, 10(5), pp.507–513.

Bruns, A.F. et al., 2010. Ligand-Stimulated VEGFR2 Signaling is Regulated by Co-Ordinated Trafficking and Proteolysis. *Traffic*, 11(1), pp.161–174.

Bubeck, J. et al., 2008. The syntaxins SYP31 and SYP81 control ER-Golgi trafficking in the plant secretory pathway. *Traffic*, 9(10), pp.1629–1652.

Burd, C. & Cullen, P.J., 2014. Retromer: A Master Conductor of Endosome Sorting. *Cold Spring Harbor Perspectives in Biology*, 6(2), pp.a016774–a016774.

Burki, F., 2014. The Eukaryotic Tree of Life from a Global Phylogenomic Perspective. *Cold Spring Harbor Perspectives in Biology*, 6(5), pp.a016147–a016147.

Burki, F. et al., 2016. Untangling the early diversification of eukaryotes: a phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista. *Proceedings of the Royal Society B: Biological Sciences*, 283(1823), p.20152802.

Cabrera, M. et al., 2003. The Retrieval Function of the KDEL Receptor Requires PKA Phosphorylation of Its C-Terminus. *Molecular biology of the cell*, 14, pp.4114–4125.

Cancino, J. et al., 2014. Control Systems of Membrane Transport at the Interface between the Endoplasmic Reticulum and the Golgi. *Developmental Cell*, 30(3), pp.280–294.

Capitani, M. & Sallese, M., 2009. The KDEL receptor: New functions for an old protein. *FEBS Letters*, 583(23), pp.3863–3871.

Carlton, J. et al., 2005. Sorting Nexins - Unifying Trends and New Perspectives. *Traffic*, (6), pp.75–82.

Casadaban, M.J. & Cohen, S.N., 1980. Analysis of gene control signals by DNA fusion and cloning in Escherichia coli. *Journal of Molecular Biology*, 138(2), pp.179–207.

Cavalier-Smith, T. et al., 2014. Multigene eukaryote phylogeny reveals the likely protozoan ancestors of opisthokonts (animals, fungi, choanozoans) and Amoebozoa. *Molecular Phylogenetics and Evolution*, 81, pp.71–85.

Ceriotti, a & Colman, a, 1988. Binding to membrane proteins within the endoplasmic reticulum cannot explain the retention of the glucose-regulated protein GRP78 in Xenopus oocytes. *The EMBO journal*, 7(3), pp.633–638.

Cevher-Keskin, B., 2013. ARF1 and SAR1 GTPases in Endomembrane Trafficking in Plants. *International Journal of Molecular Sciences*, 14(9), pp.18181–18199.

Chen, J., Qi, X. & Zheng, H., 2012. Subclass-specific localization and trafficking of Arabidopsis p24 proteins in the ER-Golgi interface. *Traffic (Copenhagen, Denmark)*, 13(3), pp.400–15.

Chen, Y. & Brandizzi, F., 2013. IRE1: ER stress sensor and cell fate executor. *Trends in Cell Biology*, 23(11), pp.547–555.

Cole, C., Barber, J.D. & Barton, G.J., 2008. The Jpred 3 secondary structure prediction server. *Nucleic acids research*, 36(Web Server issue), pp.197–201.

Collins, T.J., 2007. ImageJ for microscopy. *BioTechniques*, 43(1S), pp.S25–S30.

Connolly, T. & Gilmore, R., 1989. The signal recognition particle receptor mediates the GTP-dependent displacement of SRP from the signal sequence of the nascent polypeptide. *Cell*, 57(4), pp.599–610.

Connolly, T., Rapiejko, P. & Gilmore, R., 1991. Requirement of GTP hydrolysis for dissociation of the signal recognition particle from its receptor. *Science*, 252(5009), pp.1171–1173.

Contreras, I. et al., 2000. Characterization of Cop I coat proteins in plant cells. *Biochemical and biophysical research communications*, 273(1), pp.176–182.

Contreras, I. et al., 2004. Sorting signals in the cytosolic tail of plant p24 proteins involved in the interaction with the COPII coat. *Plant & cell physiology*, 45(12), pp.1779–1786.

Cosson, P. & Letourneur, F., 1994. Coatomer interaction with di-lysine endoplasmic reticulum retention motifs. *Science*, 263(5153), pp.1629–1631.

Cottam, N.P. & Ungar, D., 2012. Retrograde vesicle transport in the Golgi. *Protoplasma*, 249(4), pp.943–955.

Cox, J.S., Shamu, C.E. & Walter, P., 1993. Transcriptional induction of genes encoding endoplasmic reticulum resident proteins requires a transmembrane protein kinase. *Cell*, 73(6), pp.1197–1206.

Crofts et al., 1999. Saturation of the endoplasmic reticulum retention machinery reveals anterograde bulk flow. *The Plant cell*, 11(11), pp.2233–48.

Crofts, A.J. & Denecke, J., 1998. Calreticulin and calnexin in plants. *Trends in Plant Science*,

3(10), pp.396–399.

Cullen, P., 2008. Endosomal sorting and signalling: an emerging role for the sorting nexins. *Nature perspectives*, 9(july).

Dacks, J.B. & Doolittle, W.F., 2001. Reconstructing/Deconstructing the Earliest Eukaryotes. *Cell*, 107(4), pp.419–425.

Dacks, J.B. & Field, M.C., 2007. Evolution of the eukaryotic membrane-trafficking system: origin, tempo and mode. *Journal of Cell Science*, 120(17), pp.2977–2985.

Dacks, J.B., Poon, P.P. & Field, M.C., 2008. Phylogeny of endocytic components yields insight into the process of nonendosymbiotic organelle evolution. *Proceedings of the National Academy of Sciences*, 105(2), pp.588–593.

Dancourt, J. & Barlowe, C., 2010. Protein sorting receptors in the early secretory pathway. *Annual review of biochemistry*, 79, pp.777–802.

daSilva, L.L.P. et al., 2005. cDNA cloning and functional expression of KM+, the mannose-binding lectin from Artocarpus integrifolia seeds. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1726(3), pp.251–260.

daSilva, L.L.P. et al., 2004. Endoplasmic Reticulum Export Sites and Golgi Bodies Behave as Single Mobile Secretory Units in Plant Cells. *Society*, 16(July), pp.1753–1771.

daSilva, L.L.P., Foresti, O. & Denecke, J., 2006. Targeting of the plant vacuolar sorting receptor BP80 is dependent on multiple sorting signals in the cytosolic tail. *The Plant cell*, 18(6), pp.1477–97.

Day, K.J., Staehelin, L.A. & Glick, B.S., 2013. A three-stage model of Golgi structure and function. *Histochemistry and Cell Biology*, 140(3), pp.239–249.

Dean, N. & Pelham, H., 1990. Recycling of proteins from the Golgi compartment to the ER in yeast. *The Journal of cell biology*, 111(August), pp.369–377.

Denecke, J. et al., 1989. Quantitative analysis of transiently expressed genes in plant cells. *Methods Mol. Cell. Biol*, 1, pp.19–27.

Denecke, J. et al., 2012. Secretory pathway research: the more experimental systems the better. *The Plant cell*, 24(4), pp.1316–26.

Denecke, J., Botterman, J. & Deblaere, R., 1990. Protein secretion in plant cells can occur via a default pathway. *The Plant Cell Online*, 2(January), pp.51–59.

Denecke, J., De Rycke, R. & Botterman, J., 1992. Plant and mammalian sorting signals for protein retention in the endoplasmic reticulum contain a conserved epitope. *The EMBO journal*, 11(6), pp.2345–55.

Denecke, J. & Vitale, A., 1995. The use of plant protoplasts to study protein synthesis, quality control, protein modification and transport through the plant endomembrane system. *Methods in Cell Biology*, 50, pp.335–348.

Dereeper, A. et al., 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Research*, 36(Web Server), pp.W465–W469.

Donohoe, B.S., Kang, B.-H. & Staehelin, L.A., 2007. Identification and characterization of COPIa- and COPIb-type vesicle classes associated with plant and algal Golgi. *Proceedings of the National Academy of Sciences*, 104(1), pp.163–168.

Doolittle, W.F., 1981. The endosymbiont hypothesis. *Science (New York, N.Y.)*, 213(4508), pp.640–1.

Duncan, J.R. & Kornfeld, S., 1988. Intracellular movement of two mannose 6-phosphate receptors: return to the Golgi apparatus. *The Journal of cell biology*, 106(3), pp.617–28.

Dunlop, M.H. et al., 2017. Land-locked mammalian Golgi reveals cargo transport between stable cisternae. *Nature Communications*, 8(1), p.432.

Duwi Fanata, W.I., Lee, S.Y. & Lee, K.O., 2013. The unfolded protein response in plants: a fundamental adaptive cellular response to internal and external stresses. *Journal of proteomics*, 93, pp.356–68.

Ellgaard, L. & Helenius, A., 2003. Quality control in the endoplasmic reticulum. *Nature Reviews Molecular Cell Biology*, 4(3), pp.181–191.

Emr, S. et al., 2009. Journeys through the Golgi—taking stock in a new era. *The Journal of Cell Biology*, 187(4), pp.449–453.

English, A.R. & Voeltz, G.K., 2013. Endoplasmic Reticulum Structure and Interconnections with Other Organelles. *Cold Spring Harbor Perspectives in Biology*, 5(4), pp.a013227–a013227.

Fasshauer, D. et al., 1998. Conserved structural features of the synaptic fusion complex: SNARE proteins reclassified as Q- and R-SNAREs. *Proceedings of the National Academy of Sciences*, 95(26), pp.15781–15786.

Fassio, A. & Sitia, R., 2002. Formation, isomerisation and reduction of disulphide bonds during protein quality control in the endoplasmic reticulum. *Histochemistry and Cell Biology*, 117(2), pp.151–157.

Figura, K. V & Hasilik, A., 1986. Lysosomal enzymes and their receptors. *Annual review of biochemistry*, 55(1), pp.167–193.

Foresti, O. et al., 2010. A recycling-defective vacuolar sorting receptor reveals an intermediate compartment situated between prevacuoles and vacuoles in tobacco. *The Plant cell*, 22(12), pp.3992–4008.

Foresti, O., daSilva, L.L.P. & Denecke, J., 2006. Overexpression of the Arabidopsis Syntaxin PEP12/SYP21 Inhibits Transport from the Prevacuolar Compartment to the Lytic Vacuole in Vivo. *THE PLANT CELL ONLINE*, 18(9), pp.2275–2293.

Foresti, O. & Denecke, J., 2008. Intermediate organelles of the plant secretory pathway: Identity and function. *Traffic*, 9(10), pp.1599–1612.

French, A.P. et al., 2008. Colocalization of fluorescent markers in confocal microscope images of plant cells. *Nature Protocols*, 3(4), pp.619–628.

Früholz, S. et al., 2018. Nanobody-triggered lockdown of VSRs reveals ligand reloading in the Golgi. *Nature Communications*, 9(1), p.643.

Gallois, P. & Marinho, P., 1995. Leaf Disk Transformation Using Agrobacterium tumefaciens-Expression of Heterologous Genes in Tobacco. In *Plant Gene Transfer and Expression Protocols*. New Jersey: Humana Press, pp. 39–48.

Gao, C. et al., 2014. Retention mechanisms for ER and Golgi membrane proteins. *Trends in Plant Science*, 19(8), pp.508–515.

Gao, C. et al., 2012. The Golgi-Localized Arabidopsis Endomembrane Protein12 Contains Both Endoplasmic Reticulum Export and Golgi Retention Signals at Its C Terminus. *The Plant Cell*, 24(5), pp.2086–2104.

Gardner, B.M. et al., 2013. Endoplasmic Reticulum Stress Sensing in the Unfolded Protein Response. *Cold Spring Harbor Perspectives in Biology*, 5(3), pp.a013169–a013169.

Gershlick, D.C. et al., 2014. Golgi-dependent transport of vacuolar sorting receptors is regulated by COPII, AP1, and AP4 protein complexes in tobacco. *The Plant cell*, 26(3), pp.1308–29.

Gething, M.-J., 1999. Role and regulation of the ER chaperone BiP. *Seminars in Cell & Developmental Biology*, 10(5), pp.465–472.

Gething, M.-J. & Sambrook, J., 1992. Protein folding in the cell. *Nature*, 355(6355), pp.33–45.

Giannotta, M. et al., 2012. The KDEL receptor couples to Gα q/11 to activate Src kinases and regulate transport through the Golgi. *The EMBO Journal*, 31(13), pp.2869–2881.

Gilmore, R., 1982. Protein translocation across the endoplasmic reticulum. I. Detection in the microsomal membrane of a receptor for the signal recognition particle. *The Journal of Cell Biology*, 95(2), pp.463–469.

Giritch, A. et al., 2006. Rapid high-yield expression of full-size IgG antibodies in plants coinfected with noncompeting viral vectors. *Proceedings of the National Academy of Sciences*, 103(40), pp.14701–14706.

Goodin, M.M. et al., 2008. Nicotiana benthamiana: Its History and Future as a Model for Plant-Pathogen Interactions. */ 1015 MPMI*, 21(8), pp.1015–1026.

Griffiths, G. et al., 1985. Exit of newly synthesized membrane proteins from the trans cisterna of the Golgi complex to the plasma membrane. *The Journal of cell biology*, 101(3), pp.949–64.

Griffiths, G. et al., 1994. Localization of the Lys, Asp, Glu, Leu tetrapeptide receptor to the Golgi complex and the intermediate compartment in mammalian cells. *The Journal of Cell Biology*, 127(6), p.1557 LP-1574.

Haas, I.G. & Wabl, M., 1983. Immunoglobulin heavy chain binding protein. *Nature*, 306(5941), pp.387–389.

Hadlington, J. & Denecke, J., 2000. Sorting of soluble proteins in the secretory pathway of plants. *Current opinion in plant biology*, pp.461–468.

Hall, T.A., 1999. *Bioedit: A User Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT*,

Hanton, S.L. et al., 2007. De novo formation of plant endoplasmic reticulum export sites is membrane cargo induced and signal mediated. *Plant physiology*, 143(4), pp.1640–1650.

Hanton, S.L. et al., 2005. Diacidic motifs influence the export of transmembrane proteins from the endoplasmic reticulum in plant cells. *The Plant cell*, 17(11), pp.3081–3093.

Harding, H.P., Zhang, Y. & Ron, D., 1999. Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. *Nature*, 397(6716), pp.271–274.

Hardwick, K. et al., 1992. Genes that allow yeast cells to grow in the absence of the HDEL receptor. *The EMBO …*, 11(11), pp.4187–4195.

Hardwick, K., Lewis, M. & Semenza, J., 1990. ERD1, a yeast gene required for the retention of luminal endoplasmic reticulum proteins, affects glycoprotein processing in the Golgi apparatus. *The EMBO …*, 9(3), pp.623–630.

Hardwick, K.G. & Pelham, H.R.B., 1992. SED5 encodes a 39-kD integral membrane protein required for vesicular transport between the ER and the Golgi complex. *Journal of Cell Biology*, 119(3), pp.513–521.

Harter, C. & Wieland, F., 1996. The secretory pathway: mechanisms of protein sorting and

transport. *Biochimica et biophysica acta*, 1286(2), pp.75–93.

Hartl, F.U., 1996. Molecular chaperones in cellular protein folding. *Nature*, 381(6583), pp.571–580.

Hartl, F.U., 2017. Protein Misfolding Diseases. *Annu. Rev. Biochem*, 86, pp.21–26.

Hartl, F.U. & Hayer-Hartl, M., 2009. Converging concepts of protein folding in vitro and in vivo. *Nature Structural & Molecular Biology*, 16(6), pp.574–581.

Hauri, H.P. et al., 2000. ERGIC-53 and traffic in the secretory pathway. *Journal of cell science*, 113 (Pt 4, pp.587–96.

Hawes, C., 2005. Cell biology of the plant Golgi apparatus. *The New phytologist*, 165(1), pp.29–44.

Hawes, C., Kiviniemi, P. & Kriechbaumer, V., 2015. The endoplasmic reticulum: A dynamic and well-connected organelle. *Journal of Integrative Plant Biology*, 57(1), pp.50–62.

Haze, K. et al., 1999. Mammalian Transcription Factor ATF6 Is Synthesized as a Transmembrane Protein and Activated by Proteolysis in Response to Endoplasmic Reticulum Stress. *Molecular Biology of the Cell*, 10, pp.3787–3799.

von Heijne, G., 1985. Signal sequences. *Journal of Molecular Biology*, 184(1), pp.99–105.

Helenius, A., Marquardt, T. & Braakman, I., 1992. The endoplasmic reticulum as a protein-folding compartment. *Trends in Cell Biology*, 2(8), pp.227–231.

Herman, E., Tague, B. & Hoffman, L., 1990. Retention of phytohemagglutinin with carboxyterminal tetrapeptide KDEL in the nuclear envelope and the endoplasmic reticulum. *Planta*, pp.305–312.

van Herpen, T.W.J.M. et al., 2010. Nicotiana benthamiana as a Production Platform for Artemisinin Precursors H. Yang, ed. *PLoS ONE*, 5(12), p.e14222.

Herrmann, J.M. & Riemer, J., 2014. Three Approaches to One Problem: Protein Folding in the Periplasm, the Endoplasmic Reticulum, and the Intermembrane Space. *Antioxidants & Redox Signaling*, 21(3), pp.438–456.

Hetz, C. et al., 2011. The Unfolded Protein Response: Integrating Stress Signals Through the Stress Sensor IRE1α. *Physiological Reviews*, 91(4), pp.1219–1243.

Houston, N.L. et al., 2005. Phylogenetic analyses identify 10 classes of the protein disulfide isomerase family in plants, including single-domain protein disulfide isomerase-related proteins. *Plant physiology*, 137(2), pp.762–78.

Hsu, V., Shah, N. & Klausner, R., 1992. A brefeldin A-like phenotype is induced by the overexpression of a human ERD-2-like protein, ELP-1. *Cell*, 69, pp.625–635.

Hsu, V.W. & Yang, J.-S., 2009. Mechanisms of COPI vesicle formation. *FEBS Letters*, 583(23), pp.3758–3763.

Hugh, R. & Pelham, B., 1991. Recycling of proteins between the endoplasmic and Colgic complex. *Current opinion in cell biology*, 3, pp.585–591.

Hulo, N. et al., 2006. The PROSITE database. *Nucleic acids research*, 34(Database issue), pp.D227-30.

Ito, Y. et al., 2012. cis -Golgi proteins accumulate near the ER exit sites and act as the scaffold for Golgi regeneration after brefeldin A treatment in tobacco BY-2 cells T. Yoshimori, ed. *Molecular Biology of the Cell*, 23(16), pp.3203–3214.

Iwata, Y. & Koizumi, N., 2012. Plant transducers of the endoplasmic reticulum unfolded protein response.

Jackson, M.R., 1993. Retrieval of transmembrane proteins to the endoplasmic reticulum. *The Journal of Cell Biology*, 121(2), pp.317–333.

Jackson, M.R., Nilsson, T. & Peterson, P.A., 1990. Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. *The EMBO journal*, 9(10), pp.3153–62.

Janson, I.M. et al., 1998. KDEL Motif Interacts with a Specific Sequence in Mammalianerd2Receptor. *Biochemical and Biophysical Research Communications*, 247(2), pp.447–451.

Jiang, L. et al., 2001. The protein storage vacuole. *The Journal of Cell Biology*, 155(6), pp.991–1002.

Jiang, L. & Rogers, J.C., 1998. Integral Membrane Protein Sorting to Vacuoles in Plant Cells: Evidence for Two Pathways. *The Journal of Cell Biology*, 143(5), pp.1183–1199.

Johnson, A.E. & van Waes, M.A., 1999. The Translocon: A Dynamic Gateway at the ER Membrane. *Annual Review of Cell and Developmental Biology*, 15(1), pp.799–842.

Johnson, L., Bankaitis, V. & Emr, S., 1987. Distinct sequence determinants direct intracellular sorting and modification of a yeast vacuolar protease. *Cell*, 48, pp.875–885.

Jopling, H.M. et al., 2011. The VEGFR2 receptor tyrosine kinase undergoes constitutive endosome-to-plasma membrane recycling. *Biochemical and Biophysical Research Communications*, 410(2), pp.170–176.

Kaiser, C., 2000. Thinking about p24 proteins and how transport vesicles select their cargo. *Proceedings of the National Academy of Sciences*, 97(8), pp.3783–3785.

Kaiser, C. a. & Schekman, R., 1990. Distinct sets of SEC genes govern transport vesicle formation and fusion early in the secretory pathway. *Cell*, 61(4), pp.723–733.

Kamimura, D. et al., 2015. KDEL receptor 1 regulates T-cell homeostasis via PP1 that is a key phosphatase for ISR. *Nature Communications*, 6(1), p.7474.

Kaufman, R., 2004. Regulation of mRNA translation by protein folding in the endoplasmic reticulum. *Trends in Biochemical Sciences*, 29(3), pp.152–158.

Keenan, R.J. et al., 2001. The Signal Recognition Particle. *Annual Review of Biochemistry*, 70(1), pp.755–775.

Kelly, E.E. et al., 2012. The Rab family of proteins: 25 years on. *Biochemical Society Transactions*, 40(6), pp.1337–1347.

Kim, Y.E. et al., 2013. Molecular Chaperone Functions in Protein Folding and Proteostasis. *Annual Review of Biochemistry*, 82(1), pp.323–355.

Kirsch, T. et al., 1996. Interaction of a Potential Vacuolar Targeting Receptor with Amino- and Carboxyl-Terminal Targeting Determinants. *Plant Physiology*, 111(2), pp.469–474.

Kirsch, T. et al., 1994. Purification and initial characterization of a potential plant vacuolar targeting receptor. *Cell Biology*, 91, pp.3403–3407.

Klinger, C.M. et al., 2016. Resolving the homology—function relationship through comparative genomics of membrane-trafficking machinery and parasite cell biology. *Molecular and Biochemical Parasitology*, 209(1–2), pp.88–103.

Koch, G.L., 1987. Reticuloplasmins: a novel group of proteins in the endoplasmic reticulum. *Journal of cell science*, 87 (Pt 4), pp.491–492.

Kojima, R., Okumura, M. & Inaba, K., 2013. Structural Basis of Disulfide Bond Formation in the Bacterial Periplasm and Mammalian ER. In *eLS*. Chichester, UK: John Wiley & Sons, Ltd.

Kornfeld, S., 1992. Structure and Function of the Mannose 6-Phosphate/Insulinlike Growth Factor II Receptors. *Annual Review of Biochemistry*, 61(1), pp.307–330.

Krause, K.-H. & Michalak, M., 1997. Calreticulin. *Cell*, 88(4), pp.439–443.

Künzl, F. et al., 2016. Receptor-mediated sorting of soluble vacuolar proteins ends at the trans-Golgi network/early endosome. *Nature Plants*, 2(4), p.16017.

Kurokawa, K., Okamoto, M. & Nakano, A., 2014. Contact of cis-Golgi with ER exit sites executes cargo capture and delivery from the ER. *Nature Communications*, 5(1), p.3653.

Laboissiere, M.C., Sturley, S.L. & Raines, R.T., 1995. The essential function of protein-disulfide isomerase is to unscramble non-native disulfide bonds. *The Journal of biological chemistry*, 270(47), pp.28006–9.

Langhans, M. et al., 2008. In vivo trafficking and localization of p24 proteins in plant cells. *Traffic*, 9(5), pp.770–785.

Lanoix, J. et al., 2001. Sorting of Golgi resident proteins into different subpopulations of COPI vesicles. *The Journal of Cell Biology*, 155(7), pp.1199–1212.

Lavoie, C. et al., 1999. Roles for 2p24 and COPI in Endoplasmic Reticulum Cargo Exit Site Formation. *The Journal of Cell Biology*, 146(2), pp.285–299.

Leach, M.R. et al., 2002. Localization of the Lectin, ERp57 Binding, and Polypeptide Binding Sites of Calnexin and Calreticulin. *Journal of Biological Chemistry*, 277(33), pp.29686–29697. A

Leborgne-Castel, N. et al., 1999. Overexpression of BiP in tobacco alleviates endoplasmic reticulum stress. *The Plant cell*, 11(3), pp.459–470.

Lee, H.-I.I. et al., 1993. The Arabidopsis endoplasmic reticulum retention receptor functions in yeast. *Proceedings of the National Academy of Sciences*, 90(23), pp.11433–11437.

Lee, M.C.S. et al., 2004. BI-DIRECTIONAL PROTEIN TRANSPORT BETWEEN THE ER AND GOLGI. *Annual Review of Cell and Developmental Biology*, 20(1), pp.87–123.

Lemmon, S.K. & Traub, L.M., 2000. Sorting in the endosomal system in yeast and animal cells. *Current Opinion in Cell Biology*, 12(4), pp.457–466.

Lerich, A. et al., 2012. ER Import Sites and Their Relationship to ER Exit Sites: A New Model for Bidirectional ER-Golgi Transport in Higher Plants. *Frontiers in Plant Science*, 3(July), pp.1–21.

Letourneur, F. et al., 1994. Coatomer is essential for retrieval of dilysine-tagged proteins to the endoplasmic reticulum. *Cell*, 79(7), pp.1199–1207.

Leuzinger, K. et al., 2013. Efficient Agroinfiltration of Plants for High-level Transient Expression of Recombinant Proteins. *Journal of Visualized Experiments*, (77), p.50521.

Lewis, M. & Pelham, H., 1990. A human homologue of the yeast HDEL receptor. *Nature*, 348, pp.162–163.

Lewis, M. & Pelham, H., 1992a. Ligand-induced redistribution of a human KDEL receptor from the Golgi complex to the endoplasmic reticulum. *Cell*, 66, pp.353–364.

Lewis, M. & Pelham, H., 1992b. Sequence of a second human KDEL receptor. *Journal of molecular biology*, pp.913–916.

Lewis, M., Sweet, D. & Pelham, H., 1990. The ERD2 gene determines the specificity of the luminal ER protein retention system. *Cell*, 61, pp.1359–1363.

Li, J. et al., 2009. Specific ER quality control components required for biogenesis of the plant innate immune receptor EFR. *Proceedings of the National Academy of Sciences*, 106(37), pp.15973–15978.

Lindquist, S., 1986. The Heat-Shock Response. *Annual Review of Biochemistry*, 55(1), pp.1151–1191.

Lippincott-Schwartz, J., 2011. An evolving paradigm for the secretory pathway? *Molecular Biology of the Cell*, 22(21), pp.3929–3932.

Liu, X. et al., 2017. Application of CRISPR/Cas9 in plant biology. *Acta Pharmaceutica Sinica B*, 7(3), pp.292–302.

Losev, E. et al., 2006. Golgi maturation visualized in living yeast. *Nature*, 441(7096), pp.1002–1006.

Ma, W., Goldberg, E. & Goldberg, J., 2017. ER retention is imposed by COPII protein sorting and attenuated by 4-phenylbutyrate. *eLife*, 6, p.e26624.

Ma, X. et al., 2016. CRISPR/Cas9 Platforms for Genome Editing in Plants: Developments and Applications. *Molecular Plant*, 9(7), pp.961–974.

Machamer, C. & Rose, J., 1987. A specific transmembrane domain of a coronavirus E1 glycoprotein is required for its retention in the Golgi region. *The Journal of cell biology*, 105(September), pp.1205–1214.

De Maio, A., 1999. HEAT SHOCK PROTEINS: FACTS, THOUGHTS, AND DREAMS. *Shock*, 11(1), pp.1–12.

Majoul, I. et al., 2001. KDEL-Cargo Regulates Interactions between Proteins Involved in COPI Vesicle Traffic: Measurements in Living Cells Using FRET. *Developmental Cell*, 1(1), pp.139–153.

Malhotra, V. et al., 1989. Purification of a novel class of coated vesicles mediating biosynthetic protein transport through the Golgi stack. *Cell*, 58(2), pp.329–336.

Maliga, P., Sz-Breznovits, A. & Márton, L., 1973. Streptomycin-resistant plants from callus culture of haploid tobacco. *Nature: New biology*, 244(131), pp.29–30.

Malsam, J. et al., 2005. Golgin tethers define subpopulations of COPI vesicles. *Science (New York, N.Y.)*, 307(5712), pp.1095–8.

Malsam, J. & Sollner, T.H., 2011. Organization of SNAREs within the Golgi Stack. *Cold Spring Harbor Perspectives in Biology*, 3(10), pp.a005249–a005249.

De Marcos Lousa, C., Gershlick, D.C. & Denecke, J., 2012. Mechanisms and Concepts Paving the Way towards a Complete Transport Cycle of Plant Vacuolar Sorting Receptors. *The Plant Cell*, 24(5), pp.1714–1732.

Marcusson, E.G. et al., 1994. The sorting receptor for yeast vacuolar carboxypeptidase Y is encoded by the VPS10 gene. *Cell*, 77(4), pp.579–86.

Margittai, É. et al., 2015. Composition of the redox environment of the endoplasmic reticulum and sources of hydrogen peroxide. *Free Radical Biology and Medicine*, 83, pp.331–340.

Martínez-Alonso, E., Tomás, M. & Martínez-Menárguez, J.A., 2013. Golgi tubules: their structure, formation and role in intra-Golgi transport. *Histochemistry and Cell Biology*, 140(3), pp.327–339.

Martínez-Menárguez, J.A., 2013. Intra-Golgi Transport: Roles for Vesicles, Tubules, and Cisternae. *ISRN Cell Biology*, 2013, pp.1–15.

Martinière, A. et al., 2013. In vivo intracellular pH measurements in tobacco and Arabidopsis reveal an unexpected pH gradient in the endomembrane system. *The Plant cell*, 25(10), pp.4028–43.

Matheson, L. a, Hanton, S.L. & Brandizzi, F., 2006. Traffic between the plant endoplasmic reticulum and Golgi apparatus: to the Golgi and beyond. *Current opinion in plant biology*, 9(6), pp.601–9.

Matlack, K.E.., Mothes, W. & Rapoport, T.A., 1998. Protein Translocation: Tunnel Vision. *Cell*, 92(3), pp.381–390.

Matsuoka, K. & Neuhaus, J.-M., 1999. Cis-elements of protein transport to the plant vacuoles. *Journal of Experimental Botany*, 50(331), pp.165–174.

McArthur, A.G. et al., 2001. The Evolutionary Origins of Eukaryotic Protein Disulfide Isomerase Domains: New Evidence from the Amitochondriate Protist Giardia lamblia. *Molecular Biology and Evolution*, 18(8), pp.1455–1463.

McCartney, A.W. et al., 2004. Membrane-bound fatty acid desaturases are inserted co-translationally into the ER and contain different ER retrieval motifs at their carboxy termini. *The Plant Journal*, 37(2), pp.156–173.

McKay, D.B., 1993. Structure and Mechanism of 70-kDa Heat-Shock-Related Proteins. In *Adv. Protein Chem.* pp. 67–98.

Mei, M. et al., 2017. Characterization of aromatic residue–controlled protein retention in the endoplasmic reticulum of Saccharomyces cerevisiae. *Journal of Biological Chemistry*, 292(50), pp.20707–20719.

Meldolesi, J. & Pozzan, T., 1998. The endoplasmic reticulum Ca2+ store: a view from the lumen. *Trends in Biochemical Sciences*, 23(1), pp.10–14.

Meyer, D.I. & Dobberstein, B., 1980. Identification and characterization of a membrane component essential for the translocation of nascent proteins across the membrane of the endoplasmic reticulum. *The Journal of cell biology*, 87(2 Pt 1), pp.503–8.

Miesenböck, G. & Rothman, J.E., 1995. The capacity to retrieve escaped ER proteins extends to the trans-most cisterna of the Golgi stack. *The Journal of cell biology*, 129(2), pp.309–19.

Mironov, A.A. et al., 2007. Evolution of the Endoplasmic Reticulum and the Golgi Complex. In *Eukaryotic Membranes and Cytoskeleton*. New York, NY: Springer New York, pp. 61–72.

Mironov, A.A. & Pavelka, M., 2008. *The Golgi Apparatus: State of the art 110 years after Camillo Golgi's discovery* A. A. Mironov & M. Pavelka, eds., Vienna: Springer Vienna.

Mitra, K. et al., 2004. Modulation of the bilayer thickness of exocytic pathway membranes by membrane proteins rather than cholesterol. *Proceedings of the National Academy of Sciences*, 101(12), pp.4083–4088.

Montesinos, J.C. et al., 2014. Arabidopsis p24δ5 and p24δ9 facilitate Coat Protein I-dependent transport of the K/HDEL receptor ERD2 from the Golgi to the endoplasmic reticulum. *The Plant journal : for cell and molecular biology*, 80(6), pp.1014–30.

Montesinos, J.C. et al., 2012. Coupled transport of Arabidopsis p24 proteins at the ER-Golgi interface. *Journal of experimental botany*, 63(11), pp.4243–61.

Montesinos, J.C. et al., 2013. Putative p24 complexes in Arabidopsis contain members of the delta and beta subfamilies and cycle in the early secretory pathway. *Journal of experimental botany*, 64(11), pp.3147–67.

Mowbrey, K. & Dacks, J.B., 2009. Evolution and diversity of the Golgi body. *FEBS Letters*, 583(23), pp.3738–3745.

Mueller, L.A., 2005. The SOL Genomics Network. A Comparative Resource for Solanaceae Biology and Beyond. *PLANT PHYSIOLOGY*, 138(3), pp.1310–1317.

Munro, S., 2011. What is the Golgi apparatus, and why are we asking? *BMC Biology*, 9(1), p.63.

Munro, S. & Pelham, H., 1987. A C-terminal signal prevents secretion of luminal ER proteins. *Cell*, 48, pp.899–907.

enes of Nicotiana benthamiana P. Meyer, ed. *PLoS ONE*, 8(3), p.e59534.

Neumann, U., Brandizzi, F. & Hawes, C., 2003. Protein transport in plant cells: In and out of the Golgi. *Annals of Botany*, 92(2), pp.167–180.

Nilsson, T. et al., 1993. Kin recognition. *FEBS Letters*, 330(1), pp.1–4.

Nilsson, T., Jackson, M. & Peterson, P.A., 1989. Short cytoplasmic sequences serve as retention signals for transmembrane proteins in the endoplasmic reticulum. *Cell*, 58(4), pp.707–18.

Nishimura, N. et al., 1999. A di-acidic (DXE) code directs concentration of cargo during export from the endoplasmic reticulum. *Journal of Biological Chemistry*, 274(22), pp.15937–15946.

Nishimura, N. & Balch, W.E., 1997. A di-acidic signal required for selective export from the endoplasmic reticulum. *Science (New York, N.Y.)*, 277(5325), pp.556–558.

Novick, P., Field, C. & Schekman, R., 1980. Identification of 23 complementation groups required for post-translational events in the yeast secretory pathway. *Cell*, 21(1), pp.205–215.

Nufer, O. et al., 2002. Role of cytoplasmic C-terminal amino acids of membrane proteins in ER export. *Journal of cell science*, 115(Pt 3), pp.619–28.

Orci, L. et al., 1993. Budding from Golgi membranes requires the coatomer complex of non-clathrin coat proteins. *Nature*, 362(6421), pp.648–652.

Orci, L., Glick, B.S. & Rothman, J.E., 1986. A new type of coated vesicular carrier that appears not to contain clathrin: Its possible role in protein transport within the Golgi stack. *Cell*, 46(2), pp.171–184.

Orci, L., Stamnes, M. & Ravazzola, M., 1997. Bidirectional Transport by Distinct Populations of COPI-Coated Vesicles. *Cell*, 90, pp.335–349.

Ostermeier, M., De Sutter, K. & Georgiou, G., 1996. Eukaryotic Protein Disulfide Isomerase Complements Escherichia coli dsbA Mutants and Increases the Yield of a Heterologous Secreted Protein with Disulfide Bonds. *Journal of Biological Chemistry*, 271(18), pp.10616–10622.

Otsu, W. et al., 2013. A New Class of Endoplasmic Reticulum Export Signal Φ X Φ X Φ for Transmembrane Proteins and Its Selective Interaction with Sec24C. *Journal of Biological*

Chemistry, 288(25), pp.18521–18532.

Paez Valencia, J., Goodman, K. & Otegui, M.S., 2016. Endocytosis and Endosomal Trafficking in Plants. *Annual Review of Plant Biology*, 67(1), pp.309–335.

Palade, G., 1975. Intracellular aspects of the process of protein synthesis. *Science (New York, N.Y.)*, 189(4206), p.867.

Papanikou, E. & Glick, B.S., 2014. Golgi compartmentation and identity. *Current Opinion in Cell Biology*, 29, pp.74–81.

Pastor-Cantizano, N. et al., 2018. Loss of Arabidopsis p24 function affects ERD2 trafficking and Golgi structure, and activates the unfolded protein response. *Journal of Cell Science*, 131(2), p.jcs203802.

Patil, C. & Walter, P., 2001. Intracellular signaling from the endoplasmic reticulum to the nucleus: The unfolded protein response in yeast and mammals. *Current Opinion in Cell Biology*.

Patterson, G.H. et al., 2008. Transport through the Golgi Apparatus by Rapid Partitioning within a Two-Phase Membrane System. *Cell*, 133(6), pp.1055–1067.

Paul, M.J. & Frigerio, L., 2007. Coated vesicles in plant cells. *Seminars in cell & developmental biology*, 18(4), pp.471–8.

Pearse, B. & Bretscher, M., 1981. Membrane recycling by coated vesicles. *Annual review of biochemistry*, 50, pp.85–101.

Pelham, H., 1989. Control of protein exit from the endoplasmic reticulum. *Annual review of cell biology*, 6.

Pelham, H., 1990. The retention signal for soluble proteins of the endoplasmic reticulum. *Trends in biochemical sciences*, (December), pp.483–486.

Pelham, H., Roberts, L. & Lord, J., 1992. Toxin entry: how reversible is the secretory pathway? *Trends in cell biology*, 2, pp.183–185.

Pelham, H.R., 1988. Evidence that luminal ER proteins are sorted from secreted proteins in a post-ER compartment. *The EMBO Journal*, 7(4), pp.913–918.

Pelham, H.R., Hardwick, K.G. & Lewis, M.J., 1988. Sorting of soluble ER proteins in yeast. *The EMBO journal*, 7(6), pp.1757–62.

Pelham, H.R.B., 2001. Traffic through the Golgi apparatus. *The Journal of Cell Biology*, 155(7), pp.1099–1102.

Pfeffer, S.R., 2007. Unsolved Mysteries in Membrane Traffic. *Annual Review of Biochemistry*, 76(1), pp.629–645.

Phillipson, B. a et al., 2001. Secretory bulk flow of soluble proteins is efficient and COPII dependent. *The Plant cell*, 13(9), pp.2005–20.

Pimpl, P. et al., 2000. In situ localization and in vitro induction of plant COPI-coated vesicles. *The Plant cell*, 12(11), pp.2219–36.

Pimpl, P. et al., 2003. The GTPase ARF1p Controls the Sequence-Specific Vacuolar Sorting Route to the Lytic Vacuole. *THE PLANT CELL ONLINE*, 15(5), pp.1242–1256.

Pimpl, P. & Denecke, J., 2000. ER Retention of Soluble Proteins: Retrieval, Retention, or Both? *THE PLANT CELL ONLINE*, 12(9), pp.1517–1519.

Pohlschröder, M. et al., 1997. Protein Translocation in the Three Domains of Life: Variations on a Theme. *Cell*, 91(5), pp.563–566.

Polishchuk, R.S., Capestrano, M. & Polishchuk, E. V, 2009. Shaping tubular carriers for intracellular membrane transport. *FEBS Letters*, 583(23), pp.3847–3856.

Ponnambalam, S. et al., 1996. Primate homologues of rat TGN38: primary structure, expression and functional implications. *Journal of cell science*, 109 (Pt 3, pp.675–85.

Prescott, A.R. et al., 2001. Evidence for Prebudding Arrest of ER Export in Animal Cell Mitosis and its Role in Generating Golgi Partitioning Intermediates. *Traffic*, 2(5), pp.321–335.

Pulvirenti, T. et al., 2008. A traffic-activated Golgi-based signalling circuit coordinates the secretory pathway. *Nature Cell Biology*, 10(8), pp.912–922.

Rabouille, C. & Klumperman, J., 2005. Opinion: The maturing role of COPI vesicles in intra-Golgi transport. *Nature reviews. Molecular cell biology*, 6(10), pp.812–817.

Rapoport, T.A., 2007. Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. *Nature*, 450(7170), pp.663–669.

Raykhel, I. et al., 2007. A molecular specificity code for the three mammalian KDEL receptors. *The Journal of cell biology*, 179(6), pp.1193–204.

Rexach, M.F. & Schekman, R.W., 1991. Distinct biochemical requirements for the budding, targeting, and fusion of ER-derived transport vesicles. *Journal of Cell Biology*, 114(2), pp.219–229.

Robinson, D.G. et al., 2011. ARF1 Localizes to the Golgi and the Trans-Golgi Network. *The Plant Cell*, 23(3), pp.846–849.

Robinson, D.G., 2003. The Golgi Apparatus and the Plant Secretory Pathway. *Annual Plant Reviews*, 9, p.260.

Robinson, D.G. et al., 2015. Vesicles versus Tubes: Is Endoplasmic Reticulum-Golgi Transport in Plants Fundamentally Different from Other Eukaryotes? *Plant Physiology*, 168(2), pp.393–406.

Rojo, E. & Denecke, J., 2008. What Is Moving in the Secretory Pathway of Plants? *PLANT PHYSIOLOGY*, 147(4), pp.1493–1503.

Ron, D. & Walter, P., 2007. Signal integration in the endoplasmic reticulum unfolded protein response. *Nature Reviews Molecular Cell Biology*, 8(7), pp.519–529.

Ronchi, P. et al., 2008. Transmembrane domain-dependent partitioning of membrane proteins within the endoplasmic reticulum. *Journal of Cell Biology*, 181(1), pp.105–118.

Rothman, J.E., Urbani, L.J. & Brands, R., 1984. Transport of protein between cytoplasmic membranes of fused cells: correspondence to processes reconstituted in a cell-free system. *The Journal of cell biology*, 99(1 Pt 1), pp.248–59.

Rothman, J.E. & Wieland, F.T., 1996. Protein Sorting by Transport Vesicles. *Science*, 272(5259), pp.227–234.

Rueden, C.T. et al., 2017. ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics*, 18(1), p.529.

Sahagian, G.G., Distler, J. & Jourdian, G.W., 1981. Characterization of a membrane-associated receptor from bovine liver that binds phosphomannosyl residues of bovine testicular beta-galactosidase. *Proceedings of the National Academy of Sciences of the United States of America*, 78(7), pp.4289–93.

Sambrook, J., Fritsch, E.F. & Maniatis, T., 1989. *Molecular cloning: a laboratory manual.*, Cold Spring Harbor: Cold Spring Harbor laboratory.

Di Sansebastiano, G. Pietro et al., 2014. Subcellular compartmentalization in protoplasts from Artemisia annua cell cultures: Engineering attempts using a modified SNARE protein. *Journal of biotechnology*.

Saudek, V., 2012. Cystinosin, MPDU1, SWEETs and KDELR Belong to a Well-Defined Protein Family with Putative Function of Cargo Receptors Involved in Vesicle Trafficking F. Fraternali, ed. *PLoS ONE*, 7(2), p.e30876.

Schaecher, S.R., Diamond, M.S. & Pekosz, A., 2008. The Transmembrane Domain of the Severe Acute Respiratory Syndrome Coronavirus ORF7b Protein Is Necessary and Sufficient for Its Retention in the Golgi Complex. *Journal of Virology*, 82(19), pp.9477–9491.

Scheel, A. a & Pelham, H.R.B., 1998. Identification of Amino Acids in the Binding Pocket of the Human KDEL Receptor. *Journal of Biological Chemistry*, 273(4), pp.2467–2472.

Scheel, J. et al., 1997. Dissociation of Coatomer from Membranes Is Required for Brefeldin A–induced Transfer of Golgi Enzymes to the Endoplasmic Reticulum. *The Journal of Cell Biology*, 137(2), pp.319–333.

Schmitt, H.D. et al., 1986. The ras-related YPT1 gene product in yeast: A GTP-binding protein that might be involved in microtubule organization. *Cell*, 47(3), pp.401–412. Available at: https://ac.els-cdn.com/0092867486905970/1-s2.0-0092867486905970-

Schneider, C.A., Rasband, W.S. & Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7), pp.671–675.

Schwarz, D.S. & Blower, M.D., 2016. The endoplasmic reticulum: structure, function and response to cellular signaling. *Cellular and Molecular Life Sciences*, 73(1), pp.79–94.

Sciaky, N. et al., 1997. *Golgi Tubule Traffic and the Effects of Brefeldin A Visualized in Living Cells*.

Seaman, M., 2005. Recycle your receptors with retromer. *Trends in Cell Biology*, 15(2), pp.68–75.

Seaman, M.N.J., Michael McCaffery, J. & Emr, S.D., 1998. A Membrane Coat Complex Essential for Endosome-to-Golgi Retrograde Transport in Yeast. *The Journal of Cell Biology*, 142(3), pp.665–681.

Segev, N., 2001. Ypt/Rab GTPases: Regulators of Protein Trafficking. *Science Signaling*, 2001(100), pp.re11-re11.

Segev, N., Mulholland, J. & Botstein, D., 1988. The yeast GTP-binding YPT1 protein and a mammalian counterpart are associated with the secretion machinery. *Cell*, 52(6), pp.915–924.

Semenza, J. et al., 1990. ERD2, a yeast gene required for the receptor-mediated retrieval of luminal ER proteins from the secretory pathway. *Cell*, 61, pp.1349–1357.

Sharpe, H.J., Stevens, T.J. & Munro, S., 2010. A Comprehensive Comparison of Transmembrane Domains Reveals Organelle-Specific Properties. *Cell*, 142(1), pp.158–169.

Shen, J. et al., 2013. Organelle pH in the arabidopsis endomembrane system. *Molecular Plant*, 6(5), pp.1419–1437.

Shibata, Y. et al., 2009. Mechanisms Shaping the Membranes of Cellular Organelles. *Annual*

Review of Cell and Developmental Biology, 25(1), pp.329–354.

Shinshi, H. et al., 1988. Evidence for N- and C-terminal processing of a plant defense-related enzyme: Primary structure of tobacco prepro- -1,3-glucanase. *Proceedings of the National Academy of Sciences*, 85(15), pp.5541–5545.

Silva-Alvim, F.A.L. et al., 2018. Predominant Golgi Residency of the Plant K/HDEL Receptor Is Essential for Its Function in Mediating ER Retention. *The Plant Cell*, 30(9), pp.2174–2196.

Singh, P. et al., 1993. Transmembrane topology of the mammalian KDEL receptor. *Molecular and cellular biology*, 13(10), pp.6435–41.

Sleat, D.E. & Lobel, P., 1997. Ligand Binding Specificities of the Two Mannose 6-Phosphate Receptors. *Journal of Biological Chemistry*, 272(2), pp.731–738.

Smith, M.H., Ploegh, H.L. & Weissman, J.S., 2011. Road to Ruin: Targeting Proteins for Degradation in the Endoplasmic Reticulum. *Science*, 334(6059), pp.1086–1090.

Sparkes, I. et al., 2010. Five Arabidopsis Reticulon Isoforms Share Endoplasmic Reticulum Location, Topology, and Membrane-Shaping Properties. *The Plant Cell*, 22(4), pp.1333–1343.

Sparkes, I. a et al., 2006. Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nature protocols*, 1(4), pp.2019–2025.

Sparkes, I.A. et al., 2009. The plant endoplasmic reticulum: a cell-wide web. *Biochemical Journal*, 423(2), pp.145–155.

Stefanic, S. et al., 2009. Neogenesis and maturation of transient Golgi-like cisternae in a simple eukaryote. *Journal of Cell Science*, 122(16), pp.2846–2856.

Stefano, G. et al., 2006. In tobacco leaf epidermal cells, the integrity of protein export from the endoplasmic reticulum and of ER export sites depends on active COPI machinery. *The Plant journal : for cell and molecular biology*, 46(1), pp.95–110.

Stenmark, H., 2009. Rab GTPases as coordinators of vesicle traffic. *Nature reviews. Molecular cell biology*, 10(8), pp.513–25.

Stornaiuolo, M. et al., 2003. KDEL and KKXX Retrieval Signals Appended to the Same Reporter Protein Determine Different Traffickin between Endoplasmic Reticulum, Intermediate Compartment, and Golgi Complex. *Molecular biology of the cell*, 14, pp.889–902.

Strating, J.R.P.M. et al., 2009. A Comprehensive Overview of the Vertebrate p24 Family: Identification of a Novel Tissue-Specifically Expressed Member. *Molecular Biology and Evolution*, 26(8), pp.1707–1714.

Strating, J.R.P.M. & Martens, G.J.M., 2009. The p24 family and selective transport processes at the ER-Golgi interface. *Biology of the Cell*, 101(9), pp.495–509.

Tabas, I. & Ron, D., 2011. Integrating the mechanisms of apoptosis induced by endoplasmic reticulum stress. *Nature Cell Biology*, 13(3), pp.184–190.

Tang, B.L. et al., 2005. COPII and exit from the endoplasmic reticulum. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1744(3), pp.293–303.

Tang, B.L. et al., 1993. Molecular cloning, characterization, subcellular localization and dynamics of p23, the mammalian KDEL receptor. *The Journal of cell biology*, 120(2), pp.325–38.

Tolley, N. et al., 2008. Overexpression of a plant reticulon remodels the lumen of the cortical

endoplasmic reticulum but does not perturb protein transport. *Traffic*, 9(1), pp.94–102.

Towler, M.C. et al., 2000. The Manganese Cation Disrupts Membrane Dynamics along the Secretory Pathway. *Experimental Cell Research*, 259(1), pp.167–179.

Townsley, F.M., Frigerio, G. & Pelham, H.R.B., 1994. Retrieval of HDEL proteins is required for growth of yeast cells. *Journal of Cell Biology*, 127(1), pp.21–28.

Townsley, F.M., Wilson, D.W. & Pelham, H.R., 1993. Mutational analysis of the human KDEL receptor: distinct structural requirements for Golgi retention, ligand binding and retrograde transport. *The EMBO journal*, 12(7), pp.2821–2829.

Traub, L.M. & Kornfeld, S., 1997. The trans-Golgi network: a late secretory sorting station. *Current Opinion in Cell Biology*, 9, pp.527–533.

Tu, L. et al., 2008. Signal-Mediated Dynamic Retention of Glycosyltransferases in the Golgi. *Science*, 321(5887), pp.404–407.

Tveten, K. et al., 2007. 4-Phenylbutyrate restores the functionality of a misfolded mutant low-density lipoprotein receptor. *The FEBS journal*, 274(8), pp.1881–93. A

Uemura, T., 2016. Physiological Roles of Plant Post-Golgi Transport Pathways in Membrane Trafficking. *Plant and Cell Physiology*, 57(10), pp.2013–2019.

Uemura, T. et al., 2004. Systematic Analysis of SNARE Molecules in *Arabidopsis*: Dissection of the post-Golgi Network in Plant Cells. *Cell Structure and Function*, 29(2), pp.49–65.

Valls, L. et al., 1987. Protein sorting in yeast: the localization determinant of yeast vacuolar carboxypeptidase Y resides in the propeptide. *Cell*, 48, pp.887–897.

Vernoud, V., 2003. Analysis of the Small GTPase Gene Superfamily of Arabidopsis. *PLANT PHYSIOLOGY*, 131(3), pp.1191–1208.

Vincent, M.J., Martin, A.S. & Compans, R.W., 1998. Function of the KKXX motif in endoplasmic reticulum retrieval of a transmembrane protein depends on the length and structure of the cytoplasmic domain. *Journal of Biological Chemistry*, 273(2), pp.950–956.

Vitale & Denecke, 1999. The endoplasmic reticulum-gateway of the secretory pathway. *The Plant cell*, 11(4), pp.615–28.

Voeltz, G.K., Rolls, M.M. & Rapoport, T.A., 2002. Structural organization of the endoplasmic reticulum. *EMBO reports*, 3(10), pp.944–950.

Vogel, J.P., Misra, L.M. & Rose, M.D., 1990. Loss of BiP/GRP78 function blocks translocation of secretory proteins in yeast. *The Journal of cell biology*, 110(6), pp.1885–95.

Walter, P. & Lingappa, V., 1986. Mechanism of protein translocation across the endoplasmic reticulum membrane. *Annual review of cell biology*, pp.499–516.

Wang, S. et al., 2018. Plant G proteins interact with endoplasmic reticulum luminal protein receptors to regulate endoplasmic reticulum retrieval. *Journal of Integrative Plant Biology*, 60(7), pp.541–561.

Waters, M.G., Serafini, T. & Rothman, J.E., 1991. “Coatomer”: a cytosolic protein complex containing subunits of non-clathrin-coated Golgi transport vesicles. *Nature*, 349(6306), pp.248–251.

Wee, E.G.-T., 1998. Targeting of Active Sialyltransferase to the Plant Golgi Apparatus. *THE PLANT CELL ONLINE*, 10(10), pp.1759–1768.

Weidman, P.J., 1995. Anterograde transport through the Golgi complex: do Golgi tubules hold

the key? *Trends in Cell Biology*, 5(8), pp.302–305.

Wieland, F. et al., 1987. The rate of bulk flow from the endoplasmic reticulum to the cell surface. *Cell*, 50, pp.289–300.

Wilkinson, B. & Gilbert, H.F., 2004. Protein disulfide isomerase. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1699(1–2), pp.35–44.

Wilson, D.W., Lewis, M.J. & Pelham, H.R.B., 1993. pH-dependent binding of KDEL to its receptor in vitro. *Journal of Biological Chemistry*, 268(10), pp.7465–7468.

Wu, M.M. et al., 2000. Organelle pH studies using targeted avidin and fluorescein–biotin. *Chemistry & Biology*, 7(3), pp.197–209.

Xu, G. et al., 2012. Plant ERD2-like proteins function as endoplasmic reticulum luminal protein receptors and participate in programmed cell death during innate immunity. *The Plant journal : for cell and molecular biology*, 72(1), pp.57–69.

Xu, G. & Liu, Y., 2012. Plant ERD2s self-interact and interact with GTPase-activating proteins and ADP-ribosylation factor 1. *Plant Signaling & Behavior*, 7(9), pp.1092–1094.

Yamamoto, K. et al., 2003. The KDEL Receptor Modulates the Endoplasmic Reticulum Stress Response through Mitogen-activated Protein Kinase Signaling Cascades. *Journal of Biological Chemistry*, 278(36), pp.34525–34532.

Zagouras, P. & Rose, J., 1989. Carboxy-terminal SEKDEL sequences retard but do not retain two secretory proteins in the endoplasmic reticulum. *The Journal of cell biology*, 109(6), pp.2633–2640.

Zerangue, N. et al., 1999. A New ER Trafficking Signal Regulates the Subunit Stoichiometry of Plasma Membrane KATP Channels. *Neuron*, 22(3), pp.537–548.

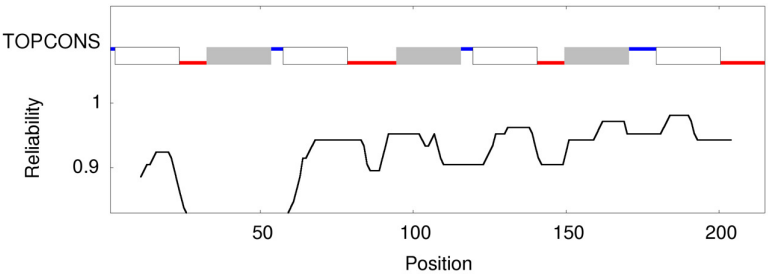
F **Appendix 1**

Protein topology of all ERD2's used.

Arabidopsis thaliana

MNIFRLAGDMTHLASVLVLLLKIHTIKSCAGVSLKTQELYAIVFATRYLDIFTSFVSLYNTSMKLVF
LGSSFSIVWYMKYHKAVHRTYDREQDTFRHWFLVLPCFLLALLIHEKFTFLEVLWTSSLYLEAV
AILPQLVLLQRTRNIDNLTGQYIFLLGGYRGLYILNWIYRYFTEPHFVHWITWIAGFVQTLLYADF
FYYYFLSWKNNKKLQLPA

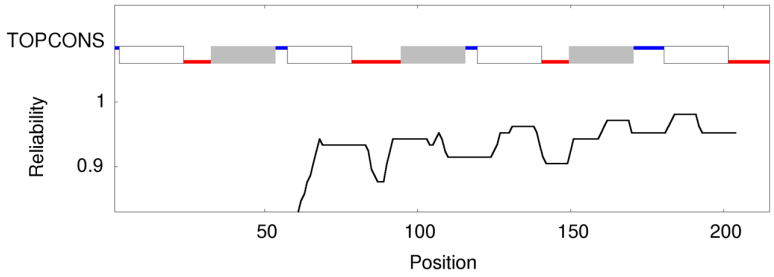
```
--HHHHHHHHHHHHHHHHHHHHHH-------HHHHHHHHHHHHHHHHHH-------
HHHHHHHHHHHHHHHHHHHHHH----------------HHHHHHHHHHHHHH-------
HHHHHHHHHHHHHHHHHHHHHHHHEE-----HHHHHHHHHHHHHHHHHHHHHHEE-------
HHHHHHHHHHHHHHHHHHHHHHHHH----E----
```



Ostreococcus lucimarinus

MNIFRFAGDMTHLCSIVVLLLKIEATKSCAGVSLRTQELYAVVFVSRYLDLFFTFISVYNTVMKV
FFITSSFCIIWYMRHHRIVSQTYDREQDTFRVAFLVVPCIFLALLVNHEFSMVEVLWTFSIYLESV
AILPQLILLQRTFNVDTLTSNYVFLLGAYRALYILNWLYRYFTEPGYSQWIVWSSGTLQTAIYCD
FFYYYVVSWRKNERLSLPS

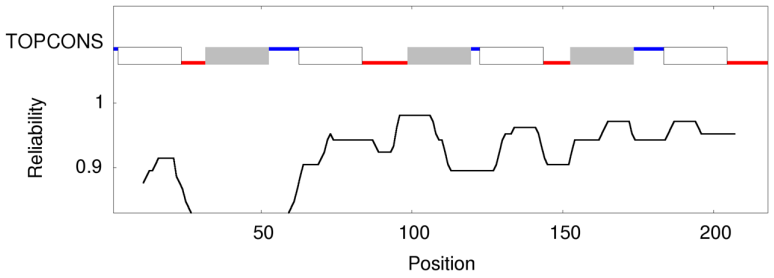
```
--HHHHHHHHHHHHHHHHHHHHHH-------HHHHHHHHHHHHHHHHH-------
HHHHHHHHHHHHHHHHHHHHHHHH---------------EEEHHHHHHHHHHH-------
HHHHHHHHHHHHHHHHHHHHHHHHE------HHHHHHHHHHHHHHHHHHHHHHHH-------
HHHHHHHHHHHHHHHHHHHHHHHHH---------
```



Chondrus crispus

MNIFRLGGDMLHVVSIFLLLLKIQTSHSCAGLSLKTQILYMVVFSTRYLDLFFTKPWHSALTFYN
TIMKILFLSSSAYTIYLMQKRYKHTYDKVHDTFRIQYLIAAAAVLALIFHLRLTVFEILWAFSVFLES
VAILPQLFLLQETGEVENITSHYIFCLGGYRTLYIFNWVWRYFTEHRRNQWLAWGCGTVQTLIY
ADFFYYYILSRKQGKKLRLPP

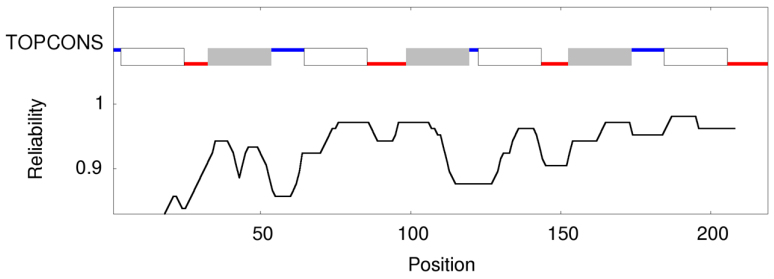
```
--HHHHHHHHHHHHHHHHHHHHHH--------HHHHHHHHHHHHHHHHH---------
HHHHHHHHHHHHHHHHHHHHHHHH----------------EEHHHHHHHHHHH------
HHHHHHHHHHHHHHHHHHHHHHHH-------HHHHHHHHHHHHHHHHHHHHHHEE--------
HHHHHHHHHHHHHHHHHHHHHHHH---------
```



Galdieria sulphuraria

MNVFRIAGDLLHCISIFILLHKMRKTRTCTGISRKTLELYAIVFLTRYLDLFTGGYFDSALSLYNTV
LKLLFLASTFYCVYLLRVKYRHTYDRSHDTFRVPFLLGAAAVLAFIFPQRYTILEILWSFSQYLEA
VAILPQLLLLQRTGEVENLTSHYIFCLGAYRGCYVLNWIWRFFTDSTYRGQYVTWTAGLIQTSL
YADFFYYYLKYKKQGRALKLPP

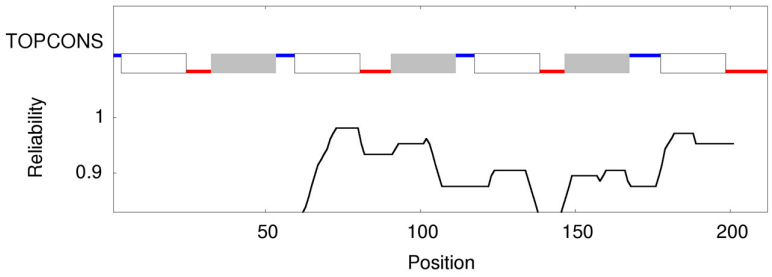
```
--HHHHHHHHHHHHHHHHHHHHHH-------HHHHHHHHHHHHHHHHHHH-------
HHHHHHHHHHHHHHHHHHHHHHHHHHEE--------------EEHHHHHHHHHHH-------
HHHHHHHHHHHHHHHHHHHHHHHHHH------HHHHHHHHHHHHHHHHHHHHHHEE---------
HHHHHHHHHHHHHHHHHHHHHHHHHH---------
```



Homo sapiens

MNIFRLTGDLSHLAAIVILLLKIWKTRSCAGISGKSQLLFALVFTTRYLDLFTSFISLYNTSMKVIYL
ACSYATVYLIYLKFKATYDGNHDTFRVEFLVVPVGGLSFLVNHDFSPLEILWTFSIYLESVAILPQ
LFMISKTGEAETITTHYLFFLGLYRALYLVNWIWRFYFEGFFDLIAVVAGVVQTILYCDFFYLYITK
VLKGKKLSLPA

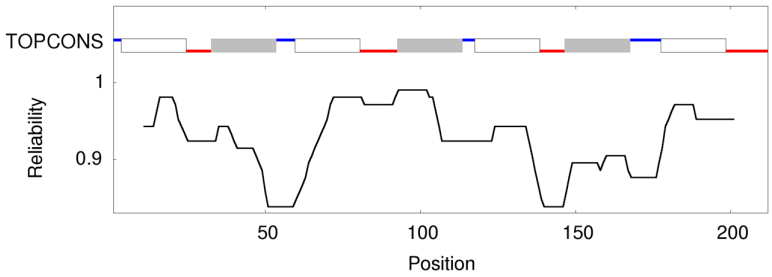
```
--HHHHHHHHHHHHHHHHHHHHHHH--------HHHHHHHHHHHHHHHH--------
HHHHHHHHHHHHHHHHHHHHHHHEE--------------EHHHHHHHHHHHH------
HHHHHHHHHHHHHHHHHHHHHHHHHE-------HHHHHHHHHHHHHHHHHHHHHHHHE-------
HHHHHHHHHHHHHHHHHHHHHHHHHH---EE----
```



Hypsibius dujardini (Tardigrade)

MNIFRLAGDICHLAAIAVLLAKIWKTRSCAGISGKSQILFALVYTTRYVDLFFSFVSVYNSLMKAV
FLIASFATVYLIYFKFKATYDFNHDTFRVEFLLIPCLILSLIITHSYEIVELLWTFSIYLEAVAILPQLF
MVSKTGEAETITSHYLFALGAYRALYIANWIWRFYAESFVDGIAVVAGIVQTILYADFFYLYITKV
LKGKEFRLPA

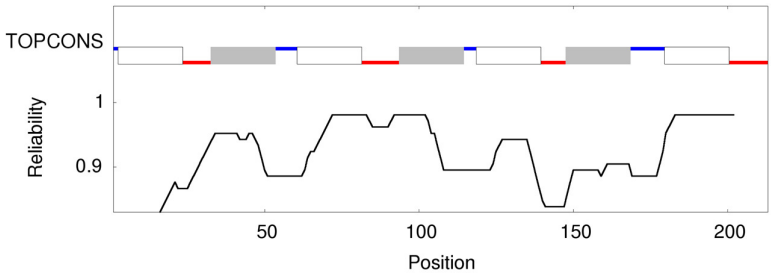
```
--HHHHHHHHHHHHHHHHHHHHHH---------HHHHHHHHHHHHHHHHHH-----
HHHHHHHHHHHHHHHHHHHHHH----------------HHHHHHHHHHHH------
HHHHHHHHHHHHHHHHHHHHHHHHHE-------HHHHHHHHHHHHHHHHHHHHHHH-------
HHHHHHHHHHHHHHHHHHHHHHHHHH---------
```



Puccinia graminis

MNIFRLIGDLSHLASIFILIQKIIKSRSARGISFKTQVLYVVVFLTRYVDLVTGPFISIYNTAMKLFFI
ASSAYIVYLMHFKYKPTQDPAIDTFKVEYLLGPCALLALVFNYKFTVVEVLWAFSIYLEAVAVFP
QLFMLHRTGEAETITTHYLFALGLYRAMYIPNWILRYTTENTLDPIAIFAGIVQTGLYADFFYIYFT
RVMRGQKFELPA

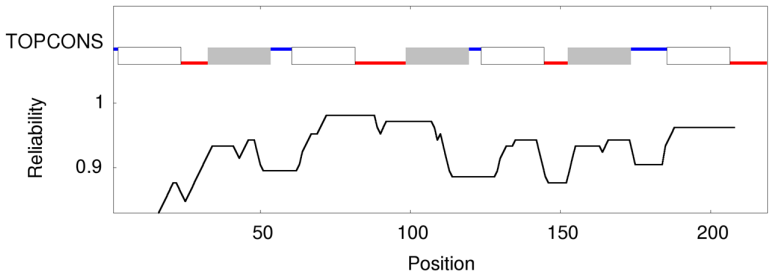
```
--HHHHHHHHHHHHHHHHHHHHHH-------HHHHHHHHHHHHHHHHHHHH------
HHHHHHHHHHHHHHHHHHHHHHEE----------------HHHHHHHHHHHHH-----
HHHHHHHHHHHHHHHHHHHHHHHHHE-------HHHHHHHHHHHHHHHHHHHHHHHHE------
HHHHHHHHHHHHHHHHHHHHHHHHHH---------
```



Saccharomyces cerevisiae

MNPFRILGDLSHLTSILILIHNIKTTRYIEGISFKTQTLYALVFITRYLDLLTFHWVSLYNALMKIFFI
VSTAYIVVLLQGSKRTNTIAYNEMLMHDTFKIQHLLIGSALMSVFFHHKFTFLELAWSFSVWLES
VAILPQLYMLSKGGKTRSLTVHYIFAMGLYRALYIPNWIWRYSTEDKKLDKIAFFAGLLQTLLYS
DFFYIYYTKVIRGKGFKLPK

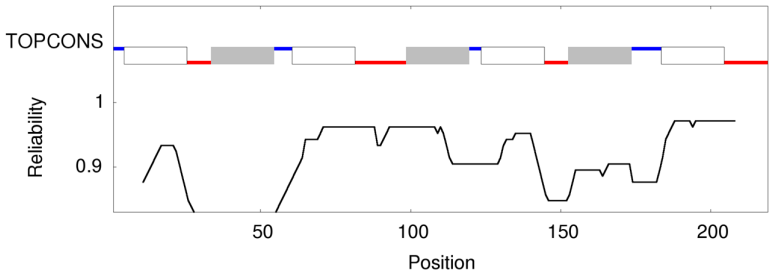
```
--HHHHHHHHHHHHHHHHHHHHHHH-------HHHHHHHHHHHHHHHHHHH------
HHHHHHHHHHHHHHHHHHHHHHHE------------------HHHHHHHHHHHHHH------
HHHHHHHHHHHHHHHHHHHHHHHHHH-------HHHHHHHHHHHHHHHHHHHHHHEE-------
EEHHHHHHHHHHHHHHHHHHHHHHHH---------
```



Kluyveromyces lactis

MLNVFRIAGDFSHLASIIILIQSITTSNSVDGISLKTQLLYTLVFITRYLNLFTKWTSLYNFLMKIVFI
SSSVYYVIVLMRQQKFKNPVAYQDMITRDQFKIKFLIVPCILLGLIFNYRFSFIQICWSFSLWLESV
AILPQLFMLTKTGKAKQLTSHYIFALGLYRALYIPNWIWRYYTEERFDKLSVFTGVIQTLVYSDFF
YIYYQKVIKLGGDLELPQ

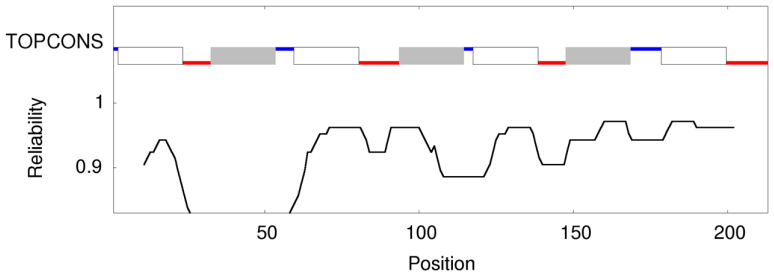
```
--HHHHHHHHHHHHHHHHHHHHHHHH----HHHHHHHHHHHHHHHHHH---------
HHHHHHHHHHHHHHHHHHHHHHHH--------------------EHHHHHHHHHHHHH-------
HHHHHHHHHHHHHHHHHHHHHHHHE-----HHHHHHHHHHHHHHHHHHHHHHHHHH-------
HHHHHHHHHHHHHHHHHHHHHHHH-----EE---
```



Acanthamoeba castellanii

MNIFRIIGDLMHLSSILMLLWKIRATKSCAGVSLKTQEMYALVFVTRYLDIFWNFSSLYNSIMKIIF
LGTSFAIIYFIRMKYRHSYDKEHDSFRVVFLIGPALLLALVFNPEFSFFEILWAFSIYLEALAILPQL
FLLQRTGEVETLTSHYIFALGGYRAFYLLNWIYRLATEPGYSNWIVWIAGFVQTVLYMDFFYYYI
QSKWYGKKFVLPA

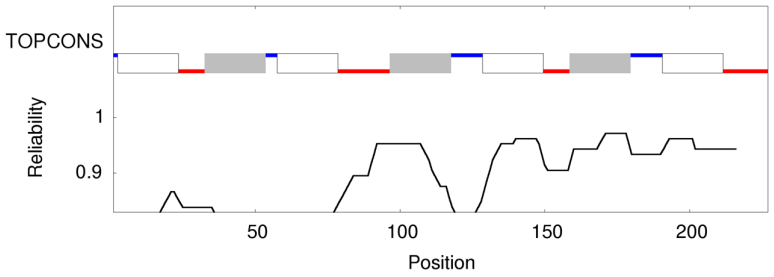
```
--HHHHHHHHHHHHHHHHHHHHHHH------HHHHHHHHHHHHHHHHH--------
HHHHHHHHHHHHHHHHHHHHHHH----------------HHHHHHHHHHH-------
HHHHHHHHHHHHHHHHHHHHHHHE-------HHHHHHHHHHHHHHHHHHHHHHEE-------
HHHHHHHHHHHHHHHHHHHHHHHH---------
```



Thalassiosira pseudonana

MNIFRLCGDMSHVFSIIVLLLRLRVARNAQGISLRTHELFLLVFLTRYTDLFTTYYSLYNSVMKVL
YIASTASIVYTIRLQEPICSTYDKAQDTFRHWEFAVAPCAVLATLTHLISGGGLFSVVDVQELLW
TFSIYLEAVAILPQLIVLQRYRDVENLTGNYIFFMGLYRALYIVNWVFRAYNEPGYRHHYVVYFC
GVLQTLLYADFFYYYVMSKRRGGKFSLPTKG

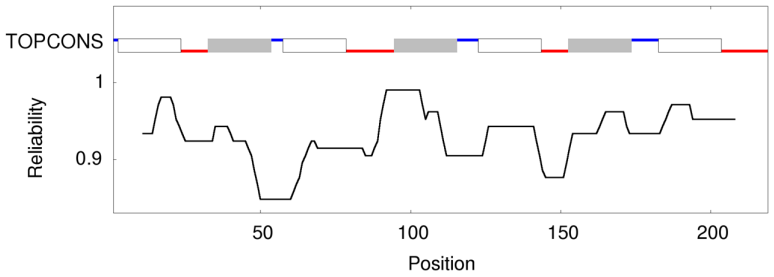
```
--HHHHHHHHHHHHHHHHHHHHHH-------HHHHHHHHHHHHHHHHHHH-----
HHHHHHHHHHHHHHHHHHHHHHHH----------------HHHHHHHHHHHHHH-------------
HHHHHHHHHHHHHHHHHHHHHHHHHHHEE----HHHHHHHHHHHHHHHHHHHHHHHHHE---------
EHHHHHHHHHHHHHHHHHHHHHHHH----EE-----
```



Trypanosoma cruzi

MNYLRTFGDMLHLLAIIILLGKMLRQRSAAGLSLKTQFLFALVFTTRYLDLFTSFLSLYNTLMKIF
FLTTSWHICYLMRNKSPWKATYDHENDTFRIRYLIVPCIVLALLFHGKPRGGWLMDFLWAFSQ
YLEAVAILPQIFLLEYTERYEALTSHYLAAMGAYRFFYLLHWIHRYLVLDRVNVVSVSAGVLQTV
LYVDFFYHYLTQVVRRAKQRYDLAR

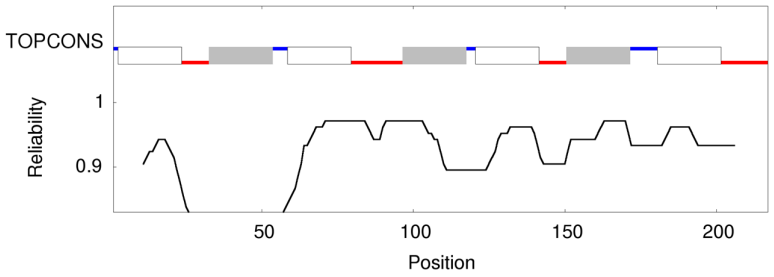
```
--HHHHHHHHHHHHHHHHHHHHHH------HHHHHHHHHHHHHHHHHH------
HHHHHHHHHHHHHHHHHHHHHHHH----------------HHHHHHHHHHHHHH--------
HHHHHHHHHHHHHHHHHHHHHHHHHHE------HHHHHHHHHHHHHHHHHHHHHHHHE------
HHHHHHHHHHHHHHHHHHHHHHHHHH-----------
```



Phytophthora infestans

MNLFRLVGDMAHLASFLVLLLKLLASRSANGISLKSQELFFLVFVTRYVDLFFHFVSLYNTLMKL
LFLLFSGAIVYVIRFKEPFRSTYDKSHDAFLHIKFAVLPCALLALVFNEQFEVMEILWTFSIYLEAV
AIIPQLILLQRHAEVENLTSNYVVLLGAYRGCYVLNWIYRAATESSYHFIWLMFIAGMVQTALYV
DFFYYYAISKYHGKKMTLPS

```
--HHHHHHHHHHHHHHHHHHHHHHH-------HHHHHHHHHHHHHHHHHHHH-----
HHHHHHHHHHHHHHHHHHHHHHHHH-----------------EEHHHHHHHHHHHH-------
HHHHHHHHHHHHHHHHHHHHHHHHHEE-------HHHHHHHHHHHHHHHHHHHHHHEE---------
HHHHHHHHHHHHHHHHHHHHHHHHH---------
```



Protein sequence of all ERD2's used.

pJCA59 TR2-GUS-35S-NbERD2ab sense

```
                              ClaI
3751 TCTATAAATC TATCTCTCTC TCTATATCGA TGAATATCTT CAGACTGGCC
                                 M   N  I  F   R  L  A    Frame 3

3801 GGAGACATGA CACATTTGAT TAGTGTCTTA GTTCTCCTCC TCAAAATCTA
     G  D  M  T   H  L  I   S  V  L   V  L  L  L   K  I  Y   Frame 3

3851 CGCTACTAAA TCATGCTCAG GAATATCATT GAAGACACAG GAGTTATATG
      A  T  K   S  C  S  G   I  S  L   K  T  Q   E  L  Y  A  Frame 3

3901 CTATCGTGTT CTTGGCTCGG TATCTGGATT TGTTCAGTGA CTTCATATCT
       I  V  F   L  A  R   Y  L  D  L   F  S  D   F  I  S   Frame 3

3951 CTTTACAACA CTGTGATGAA ACTGGTTTTC ATTGGAAGCT CTTTGGCAAT
     L  Y  N  T   V  M  K   L  V  F   I  G  S  S   L  A  I   Frame 3

                    KpnI
4001 TGTCTGGTGT ATGAGGTACC ACCGAGTTGT CAGGCGCTCA TATGACCGTG
      V  W  C   M  R  Y  H   R  V  V   R  R  S   Y  D  R  E  Frame 3

4051 AGCTAGATAC ATTTCGTTAC TGGATTCTTG TTGGAGCGTG TTTCACTTTG
       L  D  T   F  R  Y   W  I  L  V   G  A  C   F  T  L   Frame 3

4101 GCGCTTGTTA TACATGAGAA GTTTACCTTC AAGGAGATAA TGTGGACCTT
     A  L  V  I   H  E  K   F  T  F   K  E  I  M   W  T  F   Frame 3

                                          PvuII
4151 TTCCATATTC TTGGAAGCTG TTGCCATCCT TCCTCAGCTG GTCTTGTTGC
      S  I  F   L  E  A  V   A  I  L   P  Q  L   V  L  L  Q  Frame 3

4201 AGAGAACGAG AAATATAGAC AACTTGACTG GACAATACAT TTTACTCTTG
       R  T  R   N  I  D   N  L  T  G   Q  Y  I   L  L  L   Frame 3

4251 GGTGCATATC GGTCACTCTA CATCTTGAAC TGGGTATATC GCTACTTCAC
     G  A  Y  R   S  L  Y   I  L  N   W  V  Y  R   Y  F  T   Frame 3

4301 AGAACCCCAC TTTGTACATT GGATAACGTG GATTGCAGGA CTCGTGCAGA
      E  P  H   F  V  H  W   I  T  W   I  A  G   L  V  Q  T  Frame 3

4351 CAGCGGTTTA CGCTGATTTC TTTTATTACT ACTTCCAAAG CTGGAAGAAT
       A  V  Y   A  D  F   F  Y  Y  Y   F  Q  S   W  K  N   Frame 3

                                      XbaI   BamHI
4401 AACACCAAAC TCGAACTTCC TGCCTGAATC TAGAGGATCC GAAGCAGATC
     N  T  K  L   E  L  P   A  *   Frame 3
```

pJCA60 TR2-GUS-35S-NbERD2ab antisense

```
                                  NcoI   XbaI
3751 TCTATAAATC TATCTCTCTC TCTATAACCA TGGTCTAGAT TCAGGCAGGA
                                                 *  A  P   Frame 5

3801 AGTTCGAGTT TGGTGTTATT CTTCCAGCTT TGGAAGTAGT AATAAAAGAA
      L  E  L   K  T  N  N   K  W  S   Q  F  Y   Y  Y  F  F  Frame 5

3851 ATCAGCGTAA ACCGCTGTCT GCACGAGTCC TGCAATCCAC GTTATCCAAT
       D  A  Y   V  A  T   Q  V  L  G   A  I  W   T  I  W   Frame 5

3901 GTACAAAGTG GGGTTCTGTG AAGTAGCGAT ATACCCAGTT CAAGATGTAG
     H  V  F  H   P  E  T   F  Y  R   Y  V  W  N   L  I  Y   Frame 5

3951 AGTGACCGAT ATGCACCCAA GAGTAAAATG TATTGTCCAG TCAAGTTGTC
      L  S  R   Y  A  G  L   L  L  I   Y  Q  G   T  L  N  D  Frame 5

                                  PvuII
4001 TATATTTCTC GTTCTCTGCA ACAAGACCAG CTGAGGAAGG ATGGCAACAG
```

```
      I  N  R  T  R  Q  L  L  V  L  Q  P  L  I  A  V     Frame 5

4051 CTTCCAAGAA TATGGAAAAG GTCCACATTA TCTCCTTGAA GGTAAACTTC
     A  E  L  F  I  S  F  T  W  M  I  E  K  F  T  F  K   Frame 5

4101 TCATGTATAA CAAGCGCCAA AGTGAAACAC GCTCCAACAA GAATCCAGTA
      E  H  I  V  L  A  L  T  F  C  A  G  V  L  I  W  Y  Frame 5

                                                     KpnI
4151 ACGAAATGTA TCTAGCTCAC GGTCATATGA GCGCCTGACA ACTCGGTGGT
       R  F  T  D  L  E  R  D  Y  S  R  R  V  V  R  H    Frame 5

4201 ACCTCATACA CCAGACAATT GCCAAAGAGC TTCCAATGAA AACCAGTTTC
     Y  R  M  C  W  V  I  A  L  S  S  G  I  F  V  L  K   Frame 5

4251 ATCACAGTGT TGTAAAGAGA TATGAAGTCA CTGAACAAAT CCAGATACCG
      M  V  T  N  Y  L  S  I  F  D  S  F  L  D  L  Y  R  Frame 5

4301 AGCCAAGAAC ACGATAGCAT ATAACTCCTG TGTCTTCAAT GATATTCCTG
       A  L  F  V  I  A  Y  L  E  Q  T  K  L  S  I  G    Frame 5

4351 AGCATGATTT AGTAGCGTAG ATTTTGAGGA GGAGAACTAA GACACTAATC
     S  C  S  K  T  A  Y  I  K  L  L  L  V  L  V  S  I   Frame 5

                                              ClaI   BamHI
4401 AAATGTGTCA TGTCTCCGGC CAGTCTGAAG ATATTCATCG ATGGATCCGA
      L  H  T  M  D  G  A  L  R  F  I  N  M  S  P  D  S  Frame 5
```

pJCA2 TR2-GUS-35S-yERD2

```
        ClaI
3801 CTATATCGAT GAATCCGTTT AGAATCTTAG GTGATTTATC ACATCTAACC
           M  N  P  F  R  I  L  G  D  L  S  H  L  T     Frame 2

3851 AGTATACTGA TCCTGATTCA TAATATCAAG ACCACAAGGT ACATTGAAGG
     S  I  L  I  L  I  H  N  I  K  T  T  R  Y  I  E  G   Frame 2

3901 TATTTCTTTC AAGACCCAAA CGTTGTACGC TTTGGTTTTC ATAACACGAT
      I  S  F  K  T  Q  T  L  Y  A  L  V  F  I  T  R  Y  Frame 2

3951 ACTTGGATCT CTTGACTTTT CACTGGGTAT CCCTATACAA TGCTCTAATG
       L  D  L  L  T  F  H  W  V  S  L  Y  N  A  L  M    Frame 2

4001 AAAATATTTT TCATTGTATC TACCGCTTAC ATTGTAGTGC TATTACAAGG
     K  I  F  F  I  V  S  T  A  Y  I  V  V  L  L  Q  G   Frame 2

4051 GTCTAAAAGA ACCAACACCA TTGCGTATAA TGAAATGCTT ATGCATGATA
      S  K  R  T  N  T  I  A  Y  N  E  M  L  M  H  D  T  Frame 2

4101 CCTTTAAGAT CCAGCATTTA CTAATTGGGA GTGCTCTAAT GAGTGTTTTT
       F  K  I  Q  H  L  L  I  G  S  A  L  M  S  V  F    Frame 2

4151 TTCCATCACA AGTTCACTTT TCTTGAATTA GCATGGAGTT TTTCTGTATG
     F  H  H  K  F  T  F  L  E  L  A  W  S  F  S  V  W   Frame 2

4201 GTTGGAGAGT GTGGCTATTC TACCTCAATT GTACATGCTA TCTAAGGGAG
      L  E  S  V  A  I  L  P  Q  L  Y  M  L  S  K  G  G  Frame 2

                                                 NcoI
4251 GGAAGACTAG AAGTCTAACT GTTCATTATA TTTTTGCCAT GGGATTATAC
       K  T  R  S  L  T  V  H  Y  I  F  A  M  G  L  Y    Frame 2

4301 CGTGCATTGT ATATTCCTAA CTGGATTTGG AGGTACAGCA CGGAAGATAA
     R  A  L  Y  I  P  N  W  I  W  R  Y  S  T  E  D  K   Frame 2

4351 AAAATTGGAC AAGATTGCCT TCTTCGCGGG ACTTTTGCAA ACTCTGTTGT
```

```
      K  L  D   K  I  A  F   F  A  G   L  L  Q   T  L  L  Y   Frame 2

4401 ACTCTGATTT CTTTTACATT TACTACACTA AAGTCATCAG AGGAAAGGGT
       S  D  F   F  Y  I   Y  Y  T  K   V  I  R   G  K  G     Frame 2

                        XbaI   BamHI
4451 TTCAAACTGC CAAAATAATC TAGAGGATCC GAAGCAGATC GTTCAAACAT
     F  K  L  P   K  *   Frame 2
```

pJCA3 TR2-GUS-35S-olERD2

```
        ClaI
3801 CTATATCGAT GAACATTTTT CGGTTCGCCG GTGACATGAC GCACCTGTGT
             M   N  I  F   R  F  A  G   D  M  T   H  L  C    Frame 2

3851 AGCATAGTCG TTCTTCTGTT GAAAATAGAG GCGACGAAAT CGTGTGCTGG
     S  I  V  V   L  L  L   K  I  E   A  T  K  S   C  A  G   Frame 2

3901 AGTTTCGTTG AGGACTCAGG AACTATACGC CGTCGTTTTC GTGAGTAGAT
      V  S  L   R  T  Q  E   L  Y  A   V  V  F   V  S  R  Y   Frame 2

3951 ACTTGGATTT GTTTTTCACT TTCATCTCAG TGTACAACAC TGTTATGAAA
       L  D  L   F  F  T   F  I  S  V   Y  N  T   V  M  K     Frame 2

4001 GTCTTCTTTA TTACCAGCAG CTTCTGTATA ATATGGTACA TGCGGCATCA
     V  F  F  I   T  S  S   F  C  I   I  W  Y  M   R  H  H   Frame 2

4051 TAGGATAGTA TCACAGACGT ATGATCGTGA ACAGGACACA TTTCGTGTCG
      R  I  V   S  Q  T  Y   D  R  E   Q  D  T   F  R  V  A   Frame 2

                                                       EcoRI
4101 CATTTCTCGT TGTACCGTGC ATTTTCCTTG CACTGCTGGT CAACCACGAA
       F  L  V   V  P  C   I  F  L  A   L  L  V   N  H  E     Frame 2

         NcoI
4151 TTCTCCATGG TGGAAGTGTT GTGGACTTTT TCAATTTACT TGGAATCGGT
     F  S  M  V   E  V  L   W  T  F   S  I  Y  L   E  S  V   Frame 2

                                PstI
4201 CGCGATCTTA CCGCAGCTCA TCTTGCTGCA GCGGACATTT AATGTTGATA
      A  I  L   P  Q  L  I   L  L  Q   R  T  F   N  V  D  T   Frame 2

4251 CACTGACCAG CAACTATGTC TTCTTGTTGG GCGCATATCG CGCCTTGTAC
       L  T  S   N  Y  V   F  L  L  G   A  Y  R   A  L  Y     Frame 2

4301 ATATTAAACT GGCTGTATCG GTATTTTACT GAGCCGGGAT ATTCACAGTG
     I  L  N  W   L  Y  R   Y  F  T   E  P  G  Y   S  Q  W   Frame 2

4351 GATAGTCTGG AGCAGTGGAA CTTTGCAAAC TGCGATTTAT TGCGACTTTT
      I  V  W   S  S  G  T   L  Q  T   A  I  Y   C  D  F  F   Frame 2

4401 TCTACTACTA TGTCGTGAGT TGGCGAAAAA ATGAACGCCT GTCGCTACCT
       Y  Y  Y   V  V  S   W  R  K  N   E  R  L   S  L  P     Frame 2

           XbaI   BamHI
4451 AGTTGATCTA GAGGATCCGA AGCAGATCGT TCAAACATTT GGCAATAAAG
     S  *   Frame 2
```

pJCA4 TR2-GUS-35S-hERD2

```
       ClaI
3801 CTATATCGAT GAACATTTTC CGGCTGACTG GGGACCTGTC CCACCTGGCG
             M   N  I  F   R  L  T  G   D  L  S   H  L  A    Frame 2

                            BglII
3851 GCCATCGTCA TCCTGCTGCT GAAGATCTGG AAGACGCGCT CCTGCGCCGG
     A  I  V  I   L  L  L   K  I  W   K  T  R  S   C  A  G   Frame 2

3901 AATTTCTGGG AAAAGCCAGC TTCTGTTTGC ACTGGTCTTC ACAACTCGTT
      I  S  G   K  S  Q  L   L  F  A   L  V  F   T  T  R  Y   Frame 2
```

```
3951 ACCTGGATCT TTTTACTTCA TTTATTTCAT TGTATAACAC ATCTATGAAA
       L  D  L   F  T  S   F  I  S  L   Y  N  T   S  M  K     Frame 2

4001 GTTATCTACC TTGCCTGCTC CTATGCCACA GTGTACCTGA TCTACCTGAA
     V  I  Y  L   A  C  S   Y  A  T   V  Y  L  I   Y  L  K    Frame 2

4051 ATTTAAGGCA ACCTACGATG GAAATCATGA TACCTTCCGA GTGGAGTTTC
      F  K  A   T  Y  D  G   N  H  D   T  F  R   V  E  F  L   Frame 2

                   StuI
4101 TGGTGGTCCC TGTGGGAGGC CTCTCATTTT TAGTTAATCA CGATTTCTCT
       V  V  P   V  G  G   L  S  F  L   V  N  H   D  F  S     Frame 2

4151 CCTCTTGAGA TCCTCTGGAC CTTCTCCATC TACCTGGAGT CCGTGGCTAT
     P  L  E  I   L  W  T   F  S  I   Y  L  E  S   V  A  I    Frame 2

4201 CCTTCCGCAG CTATTTATGA TCAGCAAGAC TGGGGAGGCC GAGACCATCA
      L  P  Q   L  F  M  I   S  K  T   G  E  A   E  T  I  T   Frame 2

4251 CCACCCACTA CCTGTTCTTC CTGGGCCTCT ATCGTGCTTT GTATCTTGTC
       T  H  Y   L  F  F   L  G  L  Y   R  A  L   Y  L  V     Frame 2

4301 AACTGGATCT GGCGCTTCTA CTTTGAGGGC TTCTTTGACC TCATTGCTGT
     N  W  I  W   R  F  Y   F  E  G   F  F  D  L   I  A  V    Frame 2

4351 GGTGGCCGGC GTAGTCCAGA CCATCCTATA CTGTGACTTC TTCTACTTGT
      V  A  G   V  V  Q  T   I  L  Y   C  D  F   F  Y  L  Y   Frame 2

             ScaI                                        XbaI
4401 ACATTACAAA AGTACTCAAG GGAAAGAAGC TCAGTTTGCC AGCATAATCT
       I  T  K   V  L  K   G  K  K  L   S  L  P   A  *    Frame 2

        BamHI
```

pJCA18 TR2-GUS-35S-tpERD2

```
        ClaI
3801 CTATATCGAT GAACATCTTT CGTCTCTGCG GCGACATGTC CCACGTCTTC
            M   N  I  F   R  L  C  G   D  M  S   H  V  F     Frame 2

3851 TCCATCATCG TCCTCCTCCT CCGCCTTCGC GTCGCCCGCA ACGCCCAAGG
     S  I  I  V   L  L  L   R  L  R   V  A  R  N   A  Q  G    Frame 2

3901 AATCTCCCTC CGCACGCACG AACTCTTCCT CCTCGTCTTC CTCACGCGAT
      I  S  L   R  T  H  E   L  F  L   L  V  F   L  T  R  Y   Frame 2

3951 ACACCGATCT CTTCACCACG TATTACAGTT TGTACAACTC CGTTATGAAG
       T  D  L   F  T  T   Y  Y  S  L   Y  N  S   V  M  K     Frame 2

4001 GTGTTGTATA TTGCTTCTAC TGCGAGTATT GTGTATACCA TTCGGTTGCA
     V  L  Y  I   A  S  T   A  S  I   V  Y  T  I   R  L  Q    Frame 2

4051 GGAGCCGATT TGTTCAACGT ACGATAAGGC GCAGGATACG TTTAGGCATT
      E  P  I   C  S  T  Y   D  K  A   Q  D  T   F  R  H  W   Frame 2

4101 GGGAGTTCGC GGTGGCCCCG TGTGCGGTGT TGGCGACCTT GACTCATTTG
       E  F  A   V  A  P   C  A  V  L   A  T  L   T  H  L     Frame 2

4151 ATAAGTGGAG GGGGGCTGTT CTCGGTTGTG GACGTGCAGG AGCTGCTTTG
     I  S  G  G   G  L  F   S  V  V   D  V  Q  E   L  L  W    Frame 2

                   XhoI
4201 GACGTTTAGC ATATACCTCG AGGCGGTGGC GATTTTGCCT CAATTGATCG
      T  F  S   I  Y  L  E   A  V  A   I  L  P   Q  L  I  V   Frame 2

4251 TCCTTCAGCG TTATCGTGAT GTGGAGAATT TGACAGGGAA TTACATCTTC
       L  Q  R   Y  R  D   V  E  N  L   T  G  N   Y  I  F     Frame 2

4301 TTTATGGGTT TGTACCGTGC TCTATACATT GTCAATTGGG TCTTCAGGGC
     F  M  G  L   Y  R  A   L  Y  I   V  N  W  V   F  R  A    Frame 2
```

```
4351 TTACAATGAG CCGGGGTATC GTCATCACTA TGTGGTTTAC TTCTGCGGAG
      Y  N  E   P  G  Y  R   H  H  Y   V  V  Y   F  C  G  V   Frame 2

4401 TTTTGCAGAC GCTGCTTTAT GCGGACTTCT TCTACTACTA TGTCATGAGC
       L  Q  T   L  L  Y   A  D  F  F   Y  Y  Y   V  M  S    Frame 2

                                                    XbaI  BamHI
4451 AAACGCCGTG GAGGAAAGTT CAGTCTCCCA ACCAAGGGAT AATCTAGAGG
     K  R  R  G   G  K  F   S  L  P   T  K  G  *    Frame 2
```

pJCA21 TR2-GUS-35S-piERD2

```
                                ClaI
3751 TCTATAAATC TATCTCTCTC TCTATATCGA TGAATCTATT CCGCCTGGTG
                                    M   N  L  F   R  L  V    Frame 3

                          NheI
3801 GGCGATATGG CACACCTGGC TAGCTTCTTG GTGCTGCTAC TTAAGCTGCT
     G  D  M  A   H  L  A   S  F  L   V  L  L  L   K  L  L   Frame 3

3851 GGCCTCGCGC TCCGCGAATG GCATTTCACT TAAGTCGCAA GAACTTTTCT
      A  S  R   S  A  N  G   I  S  L   K  S  Q   E  L  F  F   Frame 3

3901 TCTTGGTGTT CGTAACTCGC TATGTGGATC TTTTCTTCCA CTTTGTGAGT
       L  V  F   V  T  R   Y  V  D  L   F  F  H   F  V  S    Frame 3

3951 TTGTACAACA CGCTCATGAA GCTGCTCTTC CTGCTCTTCT CGGGTGCCAT
     L  Y  N  T   L  M  K   L  L  F   L  L  F  S   G  A  I   Frame 3

4001 CGTCTACGTT ATTCGCTTCA AGGAACCCTT CCGCAGCACG TACGACAAGT
      V  Y  V   I  R  F  K   E  P  F   R  S  T   Y  D  K  S   Frame 3

4051 CACATGACGC CTTTTTGCAC ATCAAGTTCG CTGTCTTGCC CTGTGCGCTA
       H  D  A   F  L  H   I  K  F  A   V  L  P   C  A  L    Frame 3

4101 TTGGCTTTGG TCTTCAATGA GCAGTTCGAA GTTATGGAGA TCCTATGGAC
     L  A  L  V   F  N  E   Q  F  E   V  M  E  I   L  W  T   Frame 3

                  XhoI
4151 CTTTTCCATC TACCTCGAGG CGGTCGCCAT CATTCCGCAA CTCATTCTTC
      F  S  I   Y  L  E  A   V  A  I   I  P  Q   L  I  L  L   Frame 3

4201 TCCAACGCCA TGCAGAAGTG GAAAACCTTA CCAGCAACTA CGTCGTGCTA
       Q  R  H   A  E  V   E  N  L  T   S  N  Y   V  V  L    Frame 3

               SacII
4251 CTTGGAGCCT ACCGCGGCTG CTACGTGCTT AACTGGATCT ATCGAGCCGC
     L  G  A  Y   R  G  C   Y  V  L   N  W  I  Y   R  A  A   Frame 3

4301 CACCGAGTCG TCCTACCACT TTATCTGGCT CATGTTCATC GCCGGCATGG
      T  E  S   S  Y  H  F   I  W  L   M  F  I   A  G  M  V   Frame 3

                       SalI
4351 TGCAGACGGC GCTTTATGTC GACTTTTTCT ACTACTACGC TATCAGCAAA
       Q  T  A   L  Y  V   D  F  F  Y   Y  Y  A   I  S  K    Frame 3

                                          XbaI  BamHI
4401 TACCACGGCA AGAAGATGAC GTTGCCTTCT TAATCTAGAG GATCCCGTAG
     Y  H  G  K   K  M  T   L  P  S   *   Frame 3
```

pJCA22 TR2-GUS-35S-pgERD2

```
                                  ClaI
3751 TCTATAAATC TATCTCTCTC TCTATATCGA TGAATATCTT TCGATTGATC
                                    M   N  I  F   R  L  I    Frame 3

3801 GGTGATCTAT CCCATCTGGC CTCGATCTTC ATCCTCATTC AAAAGATCAT
     G  D  L  S   H  L  A   S  I  F   I  L  I  Q   K  I  I   Frame 3

                                                  ScaI
```

```
3851 CAAATCAAGA TCAGCCAGAG GAATCAGCTT CAAGACTCAA GTACTCTACG
      K  S  R   S  A  R  G   I  S  F   K  T  Q   V  L  Y  V   Frame 3

3901 TGGTGGTATT CTTGACGAGA TATGTTGACC TTGTCACCGG TCCTTTCATC
       V  V  F   L  T  R   Y  V  D  L   V  T  G   P  F  I    Frame 3

3951 TCTATCTACA ATACGGCGAT GAAACTGTTC TTCATTGCAT CCTCAGCTTA
     S  I  Y  N   T  A  M   K  L  F   F  I  A  S   S  A  Y   Frame 3

4001 CATAGTTTAC CTGATGCATT TCAAGTATAA ACCAACTCAA GACCCGGCAA
      I  V  Y   L  M  H  F   K  Y  K   P  T  Q   D  P  A  I   Frame 3

4051 TCGACACATT CAAAGTAGAA TACCTACTAG GACCATGCGC ATTATTAGCA
       D  T  F   K  V  E   Y  L  L  G   P  C  A   L  L  A    Frame 3

                                        ScaI
4101 TTAGTATTCA ACTATAAATT TACGGTAGTC GAAGTACTTT GGGCATTCAG
     L  V  F  N   Y  K  F   T  V  V   E  V  L  W   A  F  S   Frame 3

4151 TATCTACTTG GAAGCAGTCG CCGTTTTCCC CCAATTGTTC ATGCTCCACA
      I  Y  L   E  A  V  A   V  F  P   Q  L  F   M  L  H  R   Frame 3

4201 GGACTGGTGA AGCCGAGACG ATCACTACAC ATTACCTATT TGCCCTTGGC
       T  G  E   A  E  T   I  T  T  H   Y  L  F   A  L  G    Frame 3

4251 CTCTACCGAG CAATGTATAT CCCCAACTGG ATTCTCAGGT ACACAACCGA
     L  Y  R  A   M  Y  I   P  N  W   I  L  R  Y   T  T  E   Frame 3

4301 AAACACCCTT GATCCGATTG CGATTTTTGC CGGAATCGTT CAAACTGGAC
      N  T  L   D  P  I  A   I  F  A   G  I  V   Q  T  G  L   Frame 3

4351 TTTATGCAGA CTTTTTCTAC ATTTACTTCA CAAGAGTAAT GAGGGGTCAG
       Y  A  D   F  F  Y   I  Y  F  T   R  V  M   R  G  Q    Frame 3

                            XbaI  BamHI
4401 AAATTTGAAT TACCAGCATA ATCTAGAGGA TCCGAAGCAG ATCGTTCAAA
     K  F  E  L   P  A  *    Frame 3
```

pJCA63 TR2-GUS-35S-klERD2

```
                                  ClaI
3751 TCTATAAATC TATCTCTCTC TCTATATCGA TGTTGAACGT TTTCAGAATA
                                  M   L  N  V   F  R  I    Frame 3

3801 GCAGGTGATT TCTCTCATTT GGCTAGTATC ATCATTTTGA TACAATCAAT
     A  G  D  F   S  H  L   A  S  I   I  I  L  I   Q  S  I   Frame 3

3851 CACAACATCT AACTCAGTTG ATGGTATCTC ATTGAAAACT CAACTGCTAT
      T  T  S   N  S  V  D   G  I  S   L  K  T   Q  L  L  Y   Frame 3

3901 ACACCTTGGT CTTTATCACA CGTTATTTGA ACCTATTTAC CAAATGGACC
       T  L  V   F  I  T   R  Y  L  N   L  F  T   K  W  T    Frame 3

3951 TCCTTGTACA ACTTCTTAAT GAAAATTGTT TTCATTTCAT CTTCGGTTTA
     S  L  Y  N   F  L  M   K  I  V   F  I  S  S   S  V  Y   Frame 3

4001 CGTCATTGTG TTAATGCGCC AACAAAAATT TAAAAACCCT GTCGCATATC
      V  I  V   L  M  R  Q   Q  K  F   K  N  P   V  A  Y  Q   Frame 3

          BclI
4051 AAGACATGAT CACCAGAGAT CAATTTAAAA TCAAGTTTTT AATAGTACCA
       D  M  I   T  R  D   Q  F  K  I   K  F  L   I  V  P    Frame 3

4101 TGCATTCTCC TAGGATTAAT TTTCAATTAT CGTTTCAGTT TTATACAAAT
     C  I  L  L   G  L  I   F  N  Y   R  F  S  F   I  Q  I   Frame 3

4151 ATGCTGGTCC TTCTCTCTAT GGTTGGAAAG TGTTGCAATC CTTCCTCAAT
      C  W  S   F  S  L  W   L  E  S   V  A  I   L  P  Q  L   Frame 3

4201 TGTTTATGTT GACTAAAACA GGTAAAGCAA AACAATTGAC ATCTCATTAT
```

```
      F  M  L   T  K  T   G  K  A  K   Q  L  T   S  H  Y     Frame 3

4251 ATTTTTGCAT TGGGGTTATA CCGTGCCTTG TACATTCCAA ATTGGATATG
     I  F  A  L   G  L  Y   R  A  L   Y  I  P  N   W  I  W   Frame 3

4301 GAGGTATTAT ACCGAAGAAA GATTCGATAA ATTATCAGTT TTCACCGGTG
      R  Y  Y   T  E  E  R   F  D  K   L  S  V   F  T  G  V  Frame 3

4351 TGATTCAAAC TTTAGTGTAC TCTGATTTCT TCTACATCTA TTATCAGAAG
       I  Q  T   L  V  Y   S  D  F  F   Y  I  Y   Y  Q  K    Frame 3

                                                      BamHI
4401 GTTATCAAAC TCGGTGGAGA TTTAGAATTA CCACAATAGA GGATCCGAAG
     V  I  K  L   G  G  D   L  E  L   P  Q  *   Frame 3
```

pJCA66 TR2-GUS-35S-gsERD2

```
                                ClaI
3751 TCTATAAATC TATCTCTCTC TCTATATCGA TGAACGTTTT CCGTATTGCT
                                   M   N  V  F   R  I  A    Frame 3

3801 GGAGACCTTC TTCATTGCAT AAGTATTTTC ATTTTACTAC ATAAGATGAG
     G  D  L  L   H  C  I   S  I  F   I  L  L  H   K  M  R   Frame 3

3851 GAAAACTCGT ACTTGCACAG GAATTTCAAG AAAAACTTTA GAACTTTATG
      K  T  R   T  C  T  G   I  S  R   K  T  L   E  L  Y  A  Frame 3

                         EcoRV
3901 CGATTGTGTT TTTAACTCGA TATCTGGACT TGTTCACTGG AGGTTATTTT
       I  V  F   L  T  R   Y  L  D  L   F  T  G   G  Y  F    Frame 3

3951 GATAGTGCGC TATCTTTGTA CAACACAGTG TTGAAATTGC TATTTCTTGC
     D  S  A  L   S  L  Y   N  T  V   L  K  L  L   F  L  A   Frame 3

4001 GAGTACCTTT TATTGCGTCT ATCTATTGAG GGTGAAATAT AGACACACAT
      S  T  F   Y  C  V  Y   L  L  R   V  K  Y   R  H  T  Y  Frame 3

                                                        PstI
     PvuI                                              PvuII
4051 ACGATCGTTC ACACGACACT TTTCGAGTGC CGTTTCTCTT GGGAGCAGCT
       D  R  S   H  D  T   F  R  V  P   F  L  L   G  A  A    Frame 3

4101 GCAGTGTTGG CGTTCATTTT CCCTCAACGA TATACTATTT TGGAAATTCT
     A  V  L  A   F  I  F   P  Q  R   Y  T  I  L   E  I  L   Frame 3

                                                     PvuII
4151 GTGGTCCTTT TCACAATACC TGGAAGCAGT TGCAATTCTC CCTCAGCTGT
      W  S  F   S  Q  Y  L   E  A  V   A  I  L   P  Q  L  L  Frame 3

4201 TGTTGCTTCA ACGAACTGGT GAGGTCGAAA ATCTCACTTC ACACTATATA
       L  L  Q   R  T  G   E  V  E  N   L  T  S   H  Y  I    Frame 3

4251 TTTTGTCTTG GAGCATATCG TGGCTGTTAT GTTTTGAACT GGATATGGCG
     F  C  L  G   A  Y  R   G  C  Y   V  L  N  W   I  W  R   Frame 3

                                                       PstI
4301 TTTCTTTACT GACAGTACAT ATCGTGGACA GTATGTTACT TGGACTGCAG
      F  F  T   D  S  T  Y   R  G  Q   Y  V  T   W  T  A  G  Frame 3

4351 GCCTAATACA AACATCTCTT TATGCAGACT TTTTTTATTA TTATTTGAAA
       L  I  Q   T  S  L   Y  A  D  F   F  Y  Y   Y  L  K    Frame 3

                                                      BamHI
4401 TATAAGAAAC AGGGTCGCGC ATTGAAACTT CCTCCGTAAA GGATCCGAAG

     Y  K  K  Q   G  R  A   L  K  L   P  P  *   Frame 3
```

pJCA67 TR2-GUS-35S-ccERD2

```
                              ClaI
3751 TCTATAAATC TATCTCTCTC TCTATATCGA TGAACATATT CCGACTCGGC
```

```
                                    M   N   I   F   R   L   G     Frame 3

3801 GGCGACATGC TGCATGTTGT CTCCATCTTC CTGCTTCTTC TCAAGATCCA
     G   D   M   L   H   V   V   S   I   F   L   L   L   L   K   I   Q     Frame 3

3851 AACCTCTCAT TCCTGTGCAG GTCTCTCTCT CAAGACCCAA ATTCTCTACA
      T   S   H   S   C   A   G   L   S   L   K   T   Q   I   L   Y   M   Frame 3

3901 TGGTCGTCTT CAGCACCCGC TATCTCGATC TCTTCTTCAC AAAACCCTGG
       V   V   F   S   T   R   Y   L   D   L   F   F   T   K   P   W     Frame 3

3951 CATTCCGCTT TGACTTTCTA CAACACCATC ATGAAGATCC TCTTTCTCTC
     H   S   A   L   T   F   Y   N   T   I   M   K   I   L   F   L   S   Frame 3

4001 CTCGTCTGCA TACACCATTT ATCTCATGCA GAAGCGCTAC AAACACACTT
      S   S   A   Y   T   I   Y   L   M   Q   K   R   Y   K   H   T   Y   Frame 3

4051 ACGATAAGGT TCACGACACG TTCCGAATCC AATATCTCAT CGCCGCTGCC
       D   K   V   H   D   T   F   R   I   Q   Y   L   I   A   A   A     Frame 3

4101 GCCGTGCTTG CGCTCATCTT CCATCTTCGT CTCACTGTGT TTGAGATCCT
     A   V   L   A   L   I   F   H   L   R   L   T   V   F   E   I   L   Frame 3

4151 TTGGGCCTTT TCTGTATTCT TGGAATCTGT CGCCATCCTT CCGCAGCTTT
      W   A   F   S   V   F   L   E   S   V   A   I   L   P   Q   L   F   Frame 3

4201 TCTTGCTACA GGAAACCGGC GAGGTCGAGA ATATCACATC GCATTACATC
       L   L   Q   E   T   G   E   V   E   N   I   T   S   H   Y   I     Frame 3

4251 TTTTGTCTTG GCGGGTATCG GACACTCTAC ATTTTCAACT GGGTGTGGAG
     F   C   L   G   G   Y   R   T   L   Y   I   F   N   W   V   W   R   Frame 3

4301 GTACTTCACA GAGCATCGTA GAAACCAGTG GCTGGCCTGG GGTTGTGGCA
      Y   F   T   E   H   R   R   N   Q   W   L   A   W   G   C   G   T   Frame 3

4351 CTGTTCAGAC TCTCATTTAC GCCGACTTCT TCTATTATTA CATCTTAAGT
       V   Q   T   L   I   Y   A   D   F   F   Y   Y   Y   I   L   S     Frame 3

                                                  BamHI
4401 CGGAAGCAGG GAAAGAAGCT CCGTTTGCCT CCATAAAGGA TCCGAAGCAG
     R   K   Q   G   K   K   L   R   L   P   P   *   Frame 3
```

pJCA68 TR2-GUS-35S-acERD2

```
                                ClaI
3751 TCTATAAATC TATCTCTCTC TCTATATCGA TGAATATCTT CAGGATCATT
                                    M   N   I   F   R   I   I     Frame 3

3801 GGCGATCTTA TGCACCTGTC GAGCATCCTT ATGCTCCTCT GGAAGATCAG
     G   D   L   M   H   L   S   S   I   L   M   L   L   W   K   I   R   Frame 3

3851 GGCCACCAAA TCATGCGCGG GCGTCTCGCT CAAGACGCAG GAAATGTATG
      A   T   K   S   C   A   G   V   S   L   K   T   Q   E   M   Y   A   Frame 3
                    M   R   G   R   L   A   Q   D   A   G   N   V   C     Frame 2
                    *   A   R   A   D   R   E   L   R   L   F   H   I     Frame 6

3901 CGCTGGTGTT TGTGACACGC TACCTCGACA TCTTCTGGAA CTTCTCCTCC
       L   V   F   V   T   R   Y   L   D   I   F   W   N   F   S   S     Frame 3
      A   G   V   C   D   T   L   P   R   H   L   L   E   L   L   L   P   Frame 2
      R   Q   H   K   H   C   A   V   E   V   D   E   P   V   E   G   G   Frame 6

3951 CTCTACAACT CCATCATGAA GATCATTTTC CTCGGCACCT CCTTTGCCAT
     L   Y   N   S   I   M   K   I   I   F   L   G   T   S   F   A   I   Frame 3
       L   Q   L   H   H   E   D   H   F   P   R   H   L   L   C   H     Frame 2
       E   V   V   G   D   H   L   D   N   E   E   A   G   G   K   G     Frame 6

4001 CATCTACTTC ATTCGGATGA AGTACCGCCA CTCCTACGAC AAGGAGCACG
      I   Y   F   I   R   M   K   Y   R   H   S   Y   D   K   E   H   D   Frame 3
     H   L   L   H   S   D   E   V   P   P   L   L   R   Q   G   A   R   Frame 2
     D   D   V   E   N   P   H   L   V   A   V   G   V   V   L   L   V   Frame 6
```

```
4051 ATTCGTTCCG CGTGGTATTC CTCATCGGTC CGGCGCTCCT GCTGGCGCTG
      S  F  R   V  V  F   L  I  G  P   A  L  L   L  A  L     Frame 3
    F  V  P   R  G  I  P   H  R  S   G  A  P   A  G  A  G    Frame 2
    I  R  E   A  H  Y  E   E  D  T   R  R  E   Q  Q  R  Q    Frame 6

4101 GTGTTCAACC CCGAGTTCTC GTTCTTCGAG ATCCTGTGGG CCTTCTCCAT
    V  F  N  P   E  F  S   F  F  E   I  L  W  A   F  S  I    Frame 3
      V  Q  P   R  V  L   V  L  R  D   P  V  G   L  L  H     Frame 2
      H  E  V   G  L  E   R  E  E  L   D  Q  P   G  E  G     Frame 6

4151 CTACCTCGAG GCGCTGGCCA TCCTGCCGCA GCTCTTCCTC CTCCAGCGCA
     Y  L  E   A  L  A  I   L  P  Q   L  F  L   L  Q  R  T   Frame 3
    L  P  R  G   A  G  H   P  A  A   A  L  P  P   P  A  H    Frame 2
    D  V  E  L   R  Q  G   D  Q  R   L  E  E  E   E  L  A    Frame 6

4201 CCGGCGAGGT GGAGACCCTC ACCTCGCACT ACATCTTCGC CCTGGGCGGC
      G  E  V   E  T  L   T  S  H  Y   I  F  A   L  G  G     Frame 3
    R  R  G   G  D  P  H   L  A  L   H  L  R   P  G  R  L   Frame 2
    G  A  L   H  L  G  E   G  R  V   V  D  E   G  Q  A  A   Frame 6

4251 TACCGCGCGT TCTACCTCCT CAACTGGATC TACCGCCTCG CCACCGAGCC
    Y  R  A  F   Y  L  L   N  W  I   Y  R  L  A   T  E  P    Frame 3
      P  R  V   L  P  P   Q  L  D  L   P  P  R   H  R  A     Frame 2
      V  A  R   E  V  E   E  V  P  D   V  A  E   G  G  L     Frame 6

4301 AGGCTATTCT AACTGGATCG TTTGGATCGC TGGGTTCGTG CAGACCGTGT
     G  Y  S   N  W  I  V   W  I  A   G  F  V   Q  T  V  L   Frame 3
    R  L  F  *    Frame 2
    W  A  I  R   V  P  D   N  P  D   S  P  E  H   L  G  H    Frame 6

4351 TGTACATGGA CTTCTTCTAC TACTACATCC AGAGCAAGTG GTACGGCAAG
      Y  M  D   F  F  Y   Y  Y  I  Q   S  K  W   Y  G  K     Frame 3
    Q  V  H   V  E  E  V   V  V  D   L  A  L   P  V  A  L   Frame 6

                            BamHI
4401 AAGTTCGTCC TTCCCGCCTA AAGGATCCGA AGCAGATCGT TCAAACATTT

    K  F  V  L   P  A  *    Frame 3
```

pJCA69 TR2-GUS-35S-tbERD2

```
                               ClaI
3751 TCTATAAATC TATCTCTCTC TCTATATCGA TGATGTACAT TCGTACCTTC
                                  M   M  Y  I   R  T  F     Frame 3

3801 GGTGACATGC TTCACCTCCT GGCTATTTTT ATTCTACTAG GAAAGATGCT
    G  D  M  L   H  L  L   A  I  F   I  L  L  G   K  M  L   Frame 3

3851 GCGGGGGCGT TCTGCTGCGG GTCTTTCTCT TAAAACGCAA TTTCTTTTTG
     R  G  R   S  A  A  G   L  S  L   K  T  Q   F  L  F  A   Frame 3

3901 CTCTTGTCTT CACGACGCGG TACCTTGACC TGTTTTTGTC CTTTATTTCT
      L  V  F   T  T  R   Y  L  D  L   F  L  S   F  I  S     Frame 3

3951 GTGTACAACA CCATGATGAA GATATTTTTT CTCGCAACCT CGTGGCATAT
    V  Y  N  T   M  M  K   I  F  F   L  A  T  S   W  H  I   Frame 3

4001 TTGTTACCTA ATGCGTTGTA AAAGCCCGTG GAAGACAACC TATGACCACG
     C  Y  L   M  R  C  K   S  P  W   K  T  T   Y  D  H  E   Frame 3

4051 AGAACGACAC CTTCCGCATC CGTTACCTGA TCATTCCCTC GTTTGTTCTC
      N  D  T   F  R  I   R  Y  L  I   I  P  S   F  V  L     Frame 3

4101 GCCCTTCTTT TCAACGGTCA CCAGCATGGC ATGTGGGTGA TGGACGTGTT
    A  L  L  F   N  G  H   Q  H  G   M  W  V  M   D  V  L   Frame 3

4151 ATGGGCATTT TCGCAGTACC TGGAGTCTGT GGCTATTCTA CCACAAATCT
     W  A  F   S  Q  Y  L   E  S  V   A  I  L   P  Q  I  F   Frame 3
```

```
4201 TCTTACTGGA ATATACAGAA CGCTACGAAG CACTCACATC GCATTACCTC
       L  L  E   Y  T  E   R  Y  E  A   L  T  S   H  Y  L     Frame 3

4251 GCCGCTATGG GTGCCTACCG CCTCTTTTAC CTTATACATT GGATTGCTCG
     A  A  M  G   A  Y  R   L  F  Y   L  I  H  W   I  A  R    Frame 3

4301 CTATTTCGTG CACGGAAGCG TTAATGCTGT CTCTGTGTGC GCTGGCGTGC
      Y  F  V   H  G  S  V   N  A  V   S  V  C   A  G  V  L   Frame 3

4351 TTCAGACAGT GTTGTACGTT GACTTCTTTT ACCACTACAT CAGCCAGGTT
       Q  T  V   L  Y  V   D  F  F  Y   H  Y  I   S  Q  V    Frame 3

                                                      BamHI
4401 GTGTGGCGCG CTAAGCAGCG ATATGATTTG GCGCGATGAA GGATCCGAAG
     V  W  R  A   K  Q  R   Y  D  L   A  R  *   Frame 3
```

pJCA70 TR2-GUS-35S-taERD2

```
                              ClaI
3751 TCTATAAATC TATCTCTCTC TCTATATCGA TGAATATCTT CCGGCTTGCT
                                 M   N  I  F   R  L  A    Frame 3

3801 GGGGATATTT GCCACTTGGC GGCCATCGCT GTCTTGTTGG CAAAAATCTG
     G  D  I  C   H  L  A   A  I  A   V  L  L  A   K  I  W   Frame 3

3851 GAAGACACGT TCATGTGCAG GCATTTCTGG AAAGTCTCAA ATCCTCTTTG
      K  T  R   S  C  A  G   I  S  G   K  S  Q   I  L  F  A  Frame 3
                *  T  C   A  N  R   S  L  R  L   D  E  K     Frame 6

                              SalI
3901 CTCTGGTCTA CACCACCCGC TACGTCGACC TCTTCTTCTC TTTCGTTTCC
       L  V  Y   T  T  R   Y  V  D  L   F  F  S   F  V  S    Frame 3
      S  Q  D   V  G  G  A   V  D  V   E  E  E   R  E  N  G   Frame 6

3951 GTTTACAATT CGCTTATGAA AGCCGTCTTC TTGATTGCCA GCTTCGCCAC
     V  Y  N  S   L  M  K   A  V  F   L  I  A  S   F  A  T   Frame 3
       N  V  I   R  K  H   F  G  D  E   Q  N  G   A  E  G    Frame 6

                                          NdeI
4001 CGTCTACCTC ATCTACTTCA AATTCAAGGC CACATATGAT TTCAATCATG
      V  Y  L   I  Y  F  K   F  K  A   T  Y  D   F  N  H  D  Frame 3
     G  D  V  E   D  V  E   F  E  L   G  C  I  I   E  I  M   Frame 6

4051 ACACGTTCCG GGTGGAATTT CTTCTCATTC CTTGCCTCAT TCTGTCGCTC
       T  F  R   V  E  F   L  L  I  P   C  L  I   L  S  L    Frame 3
      V  R  E   P  H  F  K   K  E  N   R  A  E   N  Q  R  E   Frame 6

                                 SacI
4101 ATCATCACCC ACTCCTACGA GATTGTCGAG CTCCTGTGGA CCTTCTCCAT
     I  I  T  H   S  Y  E   I  V  E   L  L  W  T   F  S  I   Frame 3
       D  D  G   V  G  V   L  N  D  L   E  Q  P   G  E  G    Frame 6

4151 CTATCTGGAG GCAGTAGCTA TTCTGCCGCA GTTGTTCATG GTGAGCAAGA
      Y  L  E   A  V  A  I   L  P  Q   L  F  M   V  S  K  T  Frame 3
     D  I  Q  L   C  Y  S   N  Q  R   L  Q  E  H   H  A  L   Frame 6

4201 CAGGGGAAGC AGAAACCATC ACCAGCCACT ACCTCTTTGC ATTGGGAGCG
       G  E  A   E  T  I   T  S  H  Y   L  F  A   L  G  A    Frame 3
      C  P  F   C  F  G  D   G  A  V   V  E  K   C  Q  S  R   Frame 6

4251 TATCGGGCAC TTTACATCGC CAATTGGATC TGGCGCTTCT ATGCAGAGTC
     Y  R  A  L   Y  I  A   N  W  I   W  R  F  Y   A  E  S   Frame 3
       I  P  C   K  V  D   G  I  P  D   P  A  E   I  C  L    Frame 6

        SalI
4301 TTTCGTCGAC GGAATCGCCG TTGTCGCTGG CATTGTCCAA ACCATATTGT
      F  V  D   G  I  A  V   V  A  G   I  V  Q   T  I  L  Y   Frame 3
     R  E  D  V   S  D  G   N  D  S   A  N  D  L   G  Y  Q    Frame 6

4351 ATGCCGATTT CTTCTACCTT TACATCACCA AAGTGTTGAA GGGGAAAGAA
```

```
      A  D  F   F  Y  L   Y  I  T  K   V  L  K   G  K  E     Frame 3
    I  G  I   E  E  V  K   V  D  G   F  H  Q   L  P  F  F   Frame 6

                          BamHI
4401 TTTCGGCTGC CGGCTTAAAG GATCCGAAGC AGATCGTTCA AACATTTGGC
     F  R  L  P   A  *   Frame 3
```