Division of Molecular Structural Biology
Department of Medical Biochemistry and Biophysics
Karolinska Institutet, Stockholm, Sweden

Structural Studies of Protein-Protein Complexes in the Early Secretory Pathway

Maria Svärd

Karolinska Institutet

Stockholm 2013

Published by Karolinska Institutet. Printed by Åtta.45 Tryckeri AB.

ISBN 978-91-7549-378-7

ABSTRACT

In the eukaryotic cell, all proteins targeted to the extracellular medium, the plasma membrane, or specific organelles within the cell depend on transfer along the secretory pathway. Within this pathway newly synthesised proteins are folded, modified, and sorted, and an elaborate quality control machinery ensures that only proteins with correct conformations are transported to their final destinations. The functions of the secretory pathway are performed by a large number of proteins, such as chaperones, processing enzymes, and transport receptors. Although the roles of some of these proteins have been established, the functions of many remain to be elucidated.

In this thesis, components of three different protein-protein complexes in the early secretory pathway have been studied by NMR spectroscopy and X-ray crystallography, with the aim of giving detailed structural information on each protein and gaining a deeper insight into the molecular mechanisms by which each protein performs its function and interactions. The work has resulted in the three-dimensional structures of MCFD2, a protein involved in transport of coagulation factors V and VIII, Erv41p, a vesicle protein proposed to play a role in glycoprotein processing and transport, and the co-chaperone $P58^{IPK}$.

NMR spectroscopy showed MCFD2 to be disordered in the absence of Ca^{2+} ions, but to adopt a predominantly ordered conformation upon binding of Ca^{2+}. MCFD2 forms a Ca^{2+}-dependent complex with the glycoprotein transport receptor ERGIC-53, and our data suggest that the requirement of Ca^{2+} for folding of MCFD2 is the mechanism behind the Ca^{2+}-dependence of complex formation. In addition, we could explain the mechanism by which two missense mutations in MCFD2 cause the bleeding disorder combined deficiency of factor V and factor VIII.

The first structure of Erv41p or any of its homologues was determined by X-ray crystallography. The protein was shown to form a β-sandwich with only limited structural homology to other proteins and this structure, in combination with the suggested role of Erv41p in glycoprotein processing, provided a starting point for our further studies to gain a better understanding of its function in the cell.

X-ray crystallography was used to determine the structure of the co-chaperone $P58^{IPK}$. $P58^{IPK}$ had been shown to interact with the chaperone BiP and stimulate its ATPase activity, and a conserved HPD motif in the J domain of $P58^{IPK}$ is known to mediate interactions between these types of proteins. The structure provided some initial insights into how the interaction with the chaperone BiP could be mediated. Further investigations of this interaction, using surface plasmon resonance biosensor technology, showed that additional interactions, beyond the J domain and the HPD motif, are important for the stimulating effect of $P58^{IPK}$ on BiP.

LIST OF PUBLICATIONS

I. Guy, J.E., Wigren, E., **Svärd, M**., Härd, T., Lindqvist, Y., 2008. New Insights into Multiple Coagulation Factor Deficiency from the Solution Structure of Human MCFD2. Journal of Molecular Biology 381, 941-955.

II. Biterova, E.I., **Svärd, M**., Possner, D.D.D., Guy, J.E., 2013. Purification, crystallization and preliminary X-ray crystallographic analysis of the lumenal domain of the ER-vesicle protein Erv41p from Saccharomyces cerevisiae. Acta Crystallographica Section F - Structural Biology and Crystallization Communications 69, 544-546.

III. Biterova, E.I., **Svärd, M**., Possner, D.D.D., Guy, J.E., 2013. The Crystal Structure of the Lumenal Domain of Erv41p, a Protein Involved in Transport between the Endoplasmic Reticulum and Golgi Apparatus. Journal of Molecular Biology 425, 2208-2218.

IV. **Svärd, M**., Biterova, E.I., Bourhis, J.M., Guy, J.E., 2011. The Crystal Structure of the Human Co-Chaperone $P58^{IPK}$. PLoS ONE 6, e22337.

V. **Svärd, M**., Danielson, U.H., Guy, J.E., 2013. Studies of the Complex between the Chaperone BiP and the Co-Chaperone $P58^{IPK}$ Suggest a Dynamic Interaction. Manuscript.

TABLE OF CONTENTS

1 Introduction 1
1.1 General introduction 1
1.2 Overview of the early secretory pathway 2
1.2.1 Protein processing, folding and quality control in the ER 4
1.2.2 Protein transport in the secretory pathway 5
1.3 The ERGIC-53/MCFD2 transport receptor complex 9
1.4 The Erv41-Erv46 complex 12
1.5 The unfolded protein response 14
1.6 The Hsp70-Hsp40 chaperone system 16
1.6.1 BiP and the ERdj proteins 19
1.7 $P58^{IPK}$ (ERdj6) 20
2 Aim of the thesis 23
3 Results and discussion 24
3.1 NMR studies and biophysical characterisation of MCFD2 (Paper I) 24
3.1.1 NMR spectroscopy and structure determination 24
3.1.2 Overall structure of MCFD2 25
3.1.3 NMR and CD studies of F5F8D causing mutants 26
3.2 Structural studies of Erv41p (Papers II and III) 29
3.2.1 Protein production, crystallisation and structure determination 29
3.2.2 Overall structure of Erv41p 30
3.2.3 Potential interaction sites 31
3.2.4 Similarity to other structures 33
3.3 Functional studies of Erv41p and Erv46p 34
3.3.1 Expression and crystallisation of Erv46p 34
3.3.2 Thermofluor studies of the potential role of Erv41p in glycan-binding 34
3.3.3 Potential Ca^{2+}-binding of Erv41p 35
3.3.4 Glycan array screen with Erv41p and Erv46p 35
3.4 Structural studies of human $P58^{IPK}$ (Paper IV) 38
3.4.1 Protein production, crystallisation and structure determination 38
3.4.2 Overall structure of $P58^{IPK}$ 39
3.4.3 Comparison to the TPR domain of mouse $P58^{IPK}$ 40
3.4.4 Binding sites for interaction partners 40
3.5 Studies of the interaction between $P58^{IPK}$ and BiP (Paper V) 43
3.5.1 Surface plasmon resonance assays for studies of the $P58^{IPK}$-BiP interaction 43
3.5.2 Studies of $P58^{IPK}$'s stimulation of BiP's ATPase activity 46
3.5.3 Functional interactions between $P58^{IPK}$ and BiP 46
4 Conclusions 48
5 Acknowledgements 50
6 References 52

LIST OF ABBREVIATIONS

ATF6	activating transcription factor 6
BiP	immunoglobulin heavy-chain binding protein
CD	circular dichroism
COPI/II	coat protein complex I/II
CRD	carbohydrate-recognition domain
eIF2α	eukaryotic translation initiation factor 2 α
ER	endoplasmic reticulum
ERAD	ER-associated degradation
ERdj	ER-localised DnaJ-containing protein
ERGIC	ER-Golgi intermediate compartment
ERGIC-53	ER-Golgi intermediate compartment 53 kDa protein
Erv	ER vesicle protein
Erv41	41 kDa ER vesicle protein
Erv46	46 kDa ER vesicle protein
F5F8D	combined deficiency of factor V and factor VIII
HPD	histidine-proline-aspartate
Hsp	heat shock protein
Hsp70	70 kDa heat shock protein
Hsp40	40 kDa heat shock protein
HSQC	heteronuclear single quantum coherence
IRE1	inositol-requiring enzyme 1
MCFD2	multiple coagulation factor deficiency protein 2
NBD	nucleotide-binding domain
NEF	nucleotide exchange factor
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser enhancement
$P52^{rIPK}$	52 kDa repressor of the inhibitor of protein kinase
$P58^{IPK}$	58 kDa inhibitor of protein kinase
PDB	protein data bank
PDI	protein disulfide isomerase
PERK	protein kinase RNA-like ER kinase
PKR	protein kinase RNA-activated
RMSD	root mean square deviation
SAD	single-wavelength anomalous diffraction
SAXS	small-angle X-ray scattering
SBD	substrate-binding domain
SPR	surface plasmon resonance
TPR	tetratricopeptide repeat
UPR	unfolded protein response

1 INTRODUCTION

1.1 General introduction

At any given time point, a typical eukaryotic cell carries out thousands of reactions. Many of these reactions are mutually incompatible; for example build-up and break-down of the same metabolites occur simultaneously. Cells have evolved various approaches to avoid the chaos that would inevitably follow if these events took place side by side. One strategy, which is most developed in eukaryotic cells, is to separate the different metabolic processes into membrane-bound organelles. Although this solves one problem, it creates another, since the specific proteins constituting the machinery of each such organelle have to be selectively transported from where they are produced. In order to ensure that every protein arrives at the right compartment in its correct conformation, an elaborate system for processing and transport, known as the secretory pathway, has evolved.

It has been estimated that one third of the more than 20000 translated proteins in eukaryotic genomes enter the secretory pathway. All of these proteins have to be folded, modified, and sorted for transport to their correct localisation, and the fidelity with which these tasks are performed by chaperones, processing enzymes, and transport receptors is nothing short of amazing. The importance of the secretory pathway is clearly established, and malfunction of different components of the processing and transport machinery has been linked to a number of diseases. A few examples include diabetes, different types of cancer, and bleeding disorders, such as the combined blood coagulation factor V and VIII deficiency.

A proteomics study has identified more than 1400 proteins in the secretory pathway proteome, and although the roles of some of these are established, the functions of hundreds of the proteins remain unknown (Gilchrist *et al.*, 2006). This highlights the challenge of providing a complete picture of each process in the pathway, which would require identification of the key components of every step and determination of how they interact to perform their functions. It is difficult to predict the roles of these uncharacterised proteins based on their sequences, and in order to fully elucidate the molecular mechanisms of the secretory pathway, determination of the three-dimensional structure of each component in combination with studies of the many interactions between them will make valuable contributions.

This thesis describes the work on three protein-protein complexes with different roles in the early secretory pathway. X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy have been applied to determine the structures of the coagulation factor binding protein MCFD2, the vesicle protein Erv41p, and the co-chaperone $P58^{IPK}$. These studies have provided new insights into the proteins at a molecular level, and we also hope to have improved the general understanding of protein transport and processing in the early secretory pathway.

1.2 Overview of the early secretory pathway

In the eukaryotic cell, the correct folding, processing and transport of proteins destined for organelles of the exocytotic pathway or secretion is dependent on events in the secretory pathway. The journey along this pathway starts at the endoplasmic reticulum (ER), and continues with stops at the ER-Golgi intermediate compartment (ERGIC) and the Golgi apparatus before the final destination is reached. Proteins travel between the different organelles in transport vesicles, which bud off from the membrane of one compartment, pass through the cytosol, and then fuse with the destination compartment to release their cargo. Anterograde transport in the direction from the ER towards the Golgi is mediated by vesicles created by the coat protein complex COPII, while COPI vesicles perform retrograde transport from the Golgi or the ERGIC back to the ER (Figure 1).

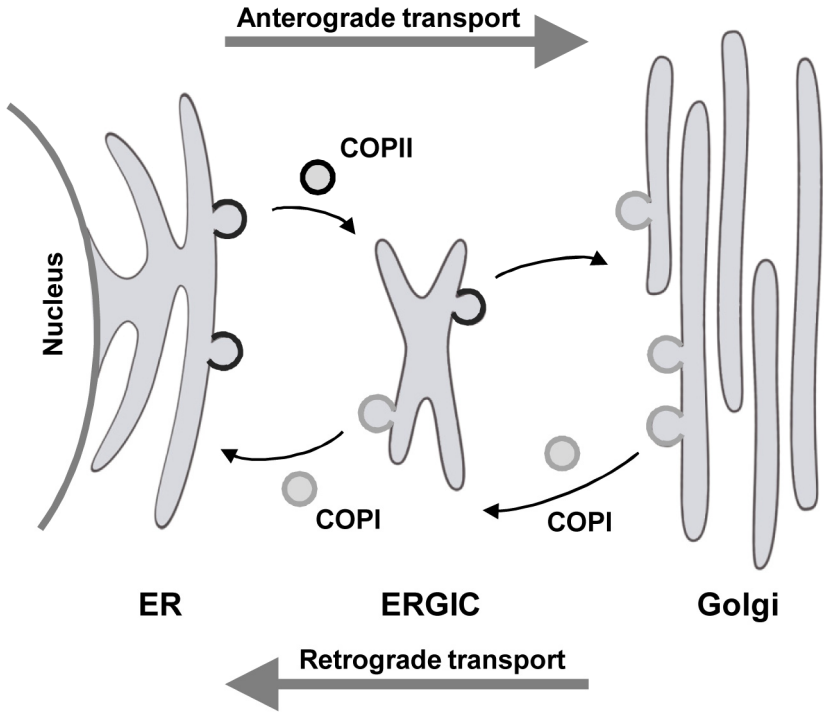


Figure 1. Schematic overview of the secretory pathway. Cargo proteins destined for secretion leave the ER in COPII vesicles (black) that bud off from the membrane and travel to the ERGIC. In the ERGIC, cargo proteins are sorted for further transport towards the Golgi apparatus and beyond. Retrograde transport is mediated by COPI vesicles (grey) that transport escaped ER proteins, or proteins that cycle between the compartments of the secretory pathway, from the Golgi and the ERGIC back to the ER.

Each organelle of the secretory pathway provides a unique environment that favours the gradual maturation of proteins. Chaperones assist folding and prevent unfolded proteins from aggregating, while other enzymes perform posttranslational modifications, such as glycosylation, alteration of the added glycan moieties, or formation of stabilising disulfide bonds. An elaborate quality control system is present at the early stages of the secretory pathway to ensure that only correctly folded proteins are transported to their final destinations. Proteins might undergo several rounds of unfolding and refolding before they are transported further, whereas terminally misfolded proteins are targeted for ER-associated degradation and retrotranslocated to the cytosol.

The endoplasmic reticulum

Already in 1902, the Italian scientist Emilio Veratti applied light microscopy to observe what was later named the endoplasmic reticulum (Veratti, 1961). However, the findings were only published in Italian, which could explain the fact that the knowledge of this structure was almost forgotten for a long time (Mazzarello et al., 2003). In 1945 however, Porter *et al.* noted the presence of a lace-like structure in tissue culture cells studied by electron microscopy. Based on its net-like appearance and location in the cytoplasm, the name endoplasmic reticulum, with the endoplasm being the innermost part of the cytoplasm and reticulum originating from the Latin word for small net, was chosen for the system (Porter & Thompson, 1948; Palade, 1956). Since then, the endoplasmic reticulum (ER) has been further studied and is now known to consist of a network of interconnected tubules and cisternae enclosed by a membrane. The ER membrane makes up more than half of the total membrane quantity in most cells, and the enclosed ER lumen can occupy more than 10% of the total cell volume (Alberts *et al.*, 2002). The ER is the first station of the secretory pathway, and the lumen provides an environment that differs from the cytosol in several ways. For example, the ER has a distinct redox environment and dedicated chaperones that promote formation and isomerisation of disulfide bonds and oxidative protein folding (Csala *et al.*, 2010). Furthermore, the total concentration of Ca^{2+} is approximately 1 mM in the ER (Prins & Michalak, 2011), which is significantly higher than the concentrations normally found in the cytosol (Pezzati *et al.*, 1997). Many chaperones involved in folding and quality control are dependent on Ca^{2+} for their function, and depletion of Ca^{2+} is known to affect the maturation of a number of proteins (Ashby & Tepikin, 2001).

The ER-Golgi intermediate compartment

The ER-Golgi intermediate compartment (ERGIC), also known as vesiculo-tubular clusters or pre-Golgi intermediates, is a part of the secretory pathway only present in higher eukaryotes. Since its discovery, the identity of the ERGIC has been controversial; while some studies have proposed it to be a specialised domain of the ER (Sitia & Meldolesi, 1992), others have assigned it as a part or the Golgi (Mellman & Simons, 1992), but it is now considered to be a distinct compartment between the ER and the Golgi with unique properties (Appenzeller-Herzog & Hauri, 2006). Whereas the total Ca^{2+} levels in the ER and the Golgi are high, they are significantly lower in the ERGIC (Pezzati et al., 1997), and the pH is thought to be somewhere in between that of the ER (pH 7.2-7.4) and that of the Golgi (pH 6.0-6.7) (Paroutis *et al.*, 2004; Appenzeller-Herzog & Hauri, 2006). These differences have been suggested to play a role in the function of the ERGIC as an important checkpoint for anterograde and retrograde transport, allowing secretory cargo to be concentrated and transported towards the Golgi, while ER resident proteins or proteins cycling between the compartments of the early secretory pathway are transported in the opposite direction. In addition to its function as a sorting station, the ERGIC has been suggested to function as a post-ER site for quality control and protein folding (Breuza *et al.*, 2004; Appenzeller-Herzog & Hauri, 2006).

The Golgi apparatus

The Golgi apparatus was first observed in 1898 by Camillo Golgi as a basket-like network surrounding the nucleus in stained Purkinje cells (Farquhar & Palade, 1998). For a long time, the functions of the Golgi apparatus were unknown. It had been seen

that the Golgi was highly developed in secretory cells, but it was not until the 1960s that the role of this organelle in secretion and glycosylation of proteins was established. The Golgi consists of disc-shaped cisternae organised into stacks (Nakamura *et al.*, 2012). Proteins arriving from the ERGIC enter at the *cis*-Golgi network, travel through the series of cisternae, and leave the organelle at the *trans*-Golgi network, where they are sorted into carriers for further transport to their final destinations. Each cisterna within a stack constitutes a separate reaction compartment with unique biochemical composition and specific reactions carried out, and as they pass through the Golgi, the secretory proteins undergo a number of sequential modifications (Kornfeld & Kornfeld, 1985; Nilsson *et al.*, 1993). Among the events to which proteins are exposed in the Golgi, the most common are various types of glycoprotein processing reactions, such as trimming and remodelling of N-linked or attachment of O-linked oligosaccharides (Hanisch, 2001; Wilson *et al.*, 2011). In contrast to the ER, there is no quality control mechanism in the Golgi to retain and reprocess or eliminate incorrectly processed proteins (Nakamura *et al.*, 2012). Instead, the arrangement of the Golgi into separate cisternae prolongs the exposure of cargo to the processing enzymes, thereby increasing the fidelity of the pathway.

1.2.1 Protein processing, folding and quality control in the ER

The first step in the maturation of proteins in the secretory pathway occurs already during translocation into the ER. Secretory proteins are translocated cotranslationally, directed by an N-terminal signal sequence (Blobel & Dobberstein, 1975). This signal sequence is removed by endoproteolytic cleavage of the nascent polypeptide chain during translocation through the Sec61 pore (Paetzel *et al.*, 2002). The multisubunit signal peptidase complex carrying out this function is thought to be located close to the Sec61 pore, with the active site near the lumenal surface of the ER membrane (Antonin *et al.*, 2000).

Also occurring concomitantly with protein translocation through the Sec61 pore is glycosylation of nascent proteins by attachment of asparagine-linked (N-linked) oligosaccharides. A 14-residue oligosaccharide core is transferred to proteins containing an Asn-X-Ser/Thr site by the oligosaccharyltransferase enzyme, which is composed of seven or eight transmembrane subunits and located close to the Sec61 pore (Kelleher & Gilmore, 2006). It has been proposed that N-linked glycosylation of secretory proteins plays a role in folding and quality control. Firstly, addition of a hydrophilic glycan moiety influences the thermodynamic stability and solubility of a protein. Secondly, the glycans allow the glycoproteins to interact with various glucosidases, glycosyltransferases, and lectins such as calnexin and calreticulin, which play roles in sorting or folding of maturing proteins, or targeting of terminally misfolded proteins for degradation (Helenius & Aebi, 2004). These processing steps occur in a sequential and protein conformation dependent manner, thereby contributing to the quality control.

Furthermore, mannose residues are added on serine or threonine side-chains of certain secretory proteins through O-linked glycosylation by protein O-mannosyltransferases. The specific role of O-linked glycosylation has not been established, but mutations in

genes encoding specific protein O-mannosyltransferases have been shown to affect degradation of misfolded proteins (Harty *et al.*, 2001; Hirayama *et al.*, 2008).

Although the folding and maturation processes in the ER are similar to those in the cytosol, there are some modifications, necessary for the processing of secretory proteins, which are unique to the ER. For example, the redox conditions differ, and only in the oxidising environment of the ER is it possible for disulfide bonds, important for folding and stabilisation of proteins, to be formed. Enzymes such as those of the PDI (protein disulfide isomerase) and Ero1 (ER oxidoreductin 1) families assist and regulate these processes of oxidation of cysteines to form new disulfide bonds and isomerisation of non-native disulfides (Bulleid & Ellgaard, 2011). Members of the PDI family have been shown to interact with other components of the ER quality control machinery, e.g. chaperones (Kimura *et al.*, 2005) or proteins involved in glycoprotein processing (Gauss *et al.*, 2011). It has been suggested that differences in domain organisation between the PDIs determine which client proteins they interact with, and whether they interact with chaperones to promote further folding of their client proteins or deliver their clients to the ER-associated degradation pathway (Vitu et al., 2008; Barlowe & Miller, 2013).

The existence of an ER quality control machinery was first suggested based on the discovery that misassembly of oligomeric viral proteins led to their inefficient export from the ER (Gething *et al.*, 1986; Kreis & Lodish, 1986). Since then it has been established that secretory proteins have to pass a number of quality control checkpoints and might undergo multiple rounds of unfolding and refolding before they are ready to leave the ER. To assist in these processes, a number of chaperones recognise and bind to proteins with deviations from their native conformations, thereby retaining them in the ER, preventing them from aggregating and assisting in their folding (Ellgaard & Helenius, 2003). Some of the chaperones in the ER belong to classical chaperone families, such as the families of heat shock proteins (Hsp), Hsp40, Hsp70, and Hsp90, that are represented also in the cytosol, whereas others, for example the glycoprotein-interacting chaperones calnexin and calreticulin, are unique to the secretory pathway.

Only once a protein is correctly folded and assembled and has passed all quality control checkpoints it is allowed to leave the ER for further processing in the down-stream compartments of the secretory pathway. Proteins that, despite several rounds of unfolding and refolding, are judged to be terminally misfolded or incompletely processed are instead retrotranslocated from the ER to the cytosol for ubiquitination and proteasomal degradation in a process known as ER-associated degradation (ERAD) (Brodsky, 2012).

1.2.2 Protein transport in the secretory pathway

Secretory proteins that have been properly assembled and modified in the ER are then separated from proteins of the quality control machinery and sorted for further transport along the secretory pathway. This anterograde transport is mediated by vesicles that bud off from one membrane and travel through the cytoplasm to fuse with a downstream compartment. Other vesicles take care of retrograde transport in the

opposite direction, allowing retrieval of cycling or ER resident proteins, thereby maintaining the balance in the secretory pathway.

Anterograde transport

Vesicles leaving the ER for anterograde transport towards the ERGIC are created by the COPII (coat protein complex II) coat (Barlowe *et al.*, 1994; Lee & Miller, 2007). This complex of proteins is responsible both for the formation of spherical vesicles from the ER membrane and for the recruitment of transport cargo to the vesicle. The budding of the COPII vesicles from the ER occurs at specific sites known as transitional ER or ER exit sites (Budnik & Stephens, 2009). Five proteins constitute the minimal COPII machinery: Sar1, Sec23, Sec24, Sec13, and Sec31 (Figure 2) (Barlowe *et al.*, 1994). The initial step in vesicle formation is the recruitment of the small G protein Sar1 to the cytosolic side of the ER membrane, and the exchange of its nucleotide from GDP to GTP by the ER membrane-bound nucleotide exchange factor Sec12 (Barlowe & Schekman, 1993). GTP-binding activates Sar1 and leads to exposure and membrane insertion of an amphipatic α-helix that is thought to induce initial curvature of the membrane (Lee et al., 2005). Activated Sar1 recruits the heterodimeric complex of Sec23 and Sec24. Sec23 is a GTPase-activating protein (Yoshihisa *et al.*, 1993), while Sec24 contains multiple independent domains that recognise specific sorting signals on transmembrane cargo proteins and transport receptors, and is responsible for the cargo-binding of COPII complexes (Miller *et al.*, 2002). After binding of cargo, the prebudding Sec23-24 complexes recruit heterotetrameric complexes of Sec13 and Sec31. These complexes assemble as an outer layer around the prebudding complexes and polymerise into a cage-like structure that further distorts and induces scission of the membrane, leading to formation of a vesicle that buds off from the ER (Stagg *et al.*, 2006). The cargo-containing vesicle travels towards the ERGIC, where it is captured through cooperation between Rab GTPases and tethering factors. Fusion of the vesicle membrane and the ERGIC membrane is then assisted by interactions between a SNARE (soluble N-ethylmaleimide-sensitive factor attachment protein receptor) protein on the transport vesicle and its SNARE-binding partner on the target membrane (Cai *et al.*, 2007; Brocker *et al.*, 2010).

In cases when efficiency of export is not critical, for example concerning abundant, nonessential proteins, cargo molecules can be packed into vesicles by nonspecific bulk flow (Pfeffer & Rothman, 1987; Wieland *et al.*, 1987; Thor *et al.*, 2009; Barlowe & Miller, 2013). However, this mode of transport cannot account for the efficiency and selectivity with which certain proteins are exported from the ER. Instead, some cargo proteins contain specific signal sequences, allowing concentrative sorting into COPII vesicles (Malkus *et al.*, 2002). Proteins transported by virtue of sorting signals can be divided into two classes: transmembrane cargo proteins that interact directly with the COPII coat and soluble cargo proteins that are linked to the coat through interaction with transmembrane cargo receptors (Barlowe, 2003).

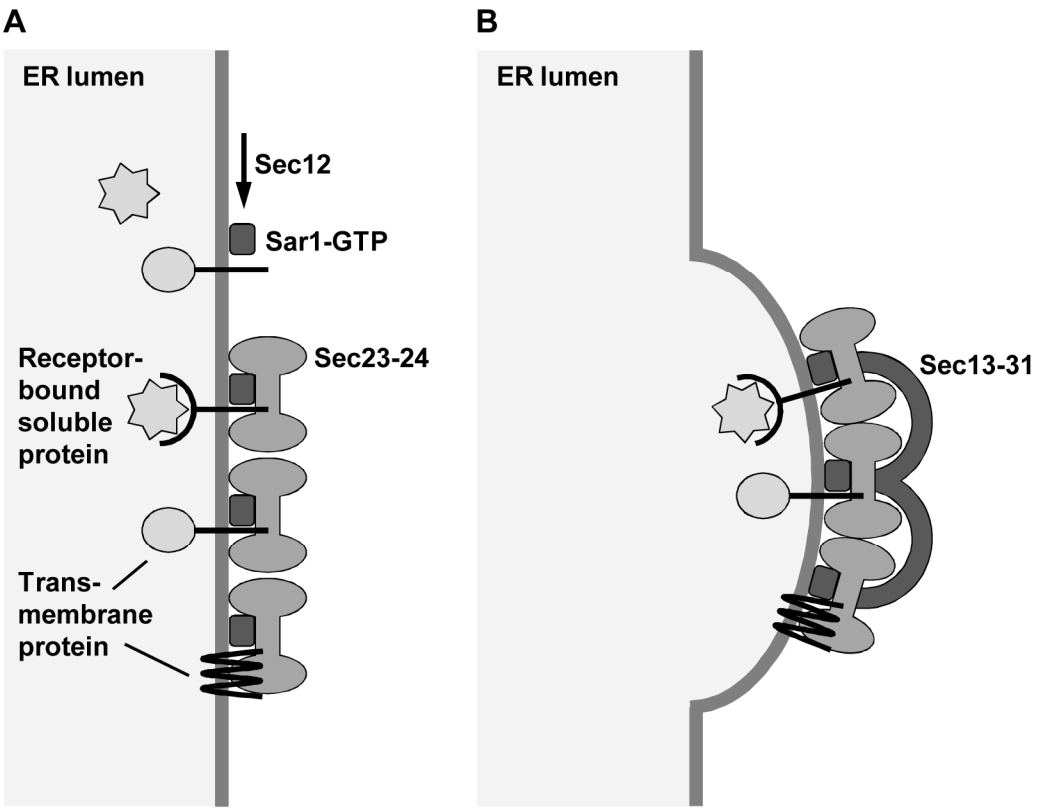


Figure 2. COPII vesicle formation and selective uptake of cargo proteins. (**A**) Vesicle formation is initiated when Sar1 is activated by the nucleotide exchange factor Sec12. Activated GTP-bound Sar1 recruits the Sec23-24 complex to the ER membrane. Transmembrane cargo proteins bind to Sec24 directly *via* their cytosolically exposed sorting signals, while the interaction of soluble proteins with the Sec23-24 complex is mediated by transmembrane cargo receptors. (**B**) The prebudding complex recruits the outer layer Sec13-31 complex which polymerises, leading to distortion and scission of the membrane and formation of a COPII-coated vesicle loaded with cargo. (Adapted from Dancourt & Barlowe, 2010.)

The first amino acid motif identified as a sorting signal of transmembrane cargo proteins was found in the cytoplasmic tail of the vesicular stomatitis virus glycoprotein (Nishimura & Balch, 1997). This diacidic Asp/Glu-X-Asp/Glu motif has been shown to interact directly with Sec23-24 of the COPII complex (Votsmeier and Gallwitz, 2001). Among other proteins depending on diacidic motifs for ER export are the angiotensin II receptor (Zhang *et al.*, 2011), the cystic fibrosis transmembrane conductance regulator (Wang *et al.*, 2004), and the potassium channels Kir1.1 and Kir2.1 (Ma *et al.*, 2001).

Another class of ER export signals comprises diaromatic or dihydrophobic motifs, such as Phe-Phe (FF), Tyr-Tyr (YY), Phe-Tyr (FY), or Leu-Leu (LL). Such a motif occurs for example in the human membrane lectin ERGIC-53, which contains two phenylalanines at the C-terminus required for exit from the ER (Kappeler *et al.*, 1997). In its yeast homologues Emp46p and Emp47p the corresponding signal consists of two leucines (Sato & Nakano, 2002). Dihydrophobic motifs are also found in both partners of the complex between the ER vesicle (Erv) proteins Erv41p and Erv46p; an isoleucine-leucine motif in Erv41p and a phenylalanine-tyrosine motif in Erv46p are necessary for sorting of the complex into COPII vesicles (Otte & Barlowe, 2002).

Many of the proteins with dihydrophobic export signals form oligomeric complexes, and it has been suggested that the multiple signals present in such complexes are required for their efficient ER exit (Barlowe, 2003). This has been proposed to be a part of the quality control of the ER, only allowing properly assembled oligomers to continue along the secretory pathway.

Recently, additional motifs for sorting into COPII vesicles have been discovered. Bovine anion exchanger 1 was shown to contain a new type of ER export motif for transmembrane proteins. This protein contains a Φ-X-Φ-X-Φ motif, where Φ is a hydrophobic amino acid and X is any amino acid, which has been shown to interact with Sec24 in the COPII coat (Otsu *et al.*, 2013). A conserved triple arginine motif was also found to function as an ER export signal for the transmembrane α_{2B}-adrenergic receptor, providing evidence of the association of G protein-coupled receptors with COPII vesicles (Dong *et al.*, 2012).

Soluble cargo proteins that are not transported by bulk flow depend on the association with transmembrane transport receptors for their packaging into COPII vesicles and export from the ER. It has been suggested that the transport receptors are important for the quality control and concentrative sorting in the ER, with only correctly folded and assembled cargo being recognised by their receptors and exported to downstream compartments of the secretory pathway (Barlowe, 2003). Although still not fully elucidated, a coordination between the ER folding machinery and sorting receptors has been indicated; a reduced binding of misfolded proteins to cargo receptors could help preventing competition between the two processes of folding and export (Dancourt & Barlowe, 2010).

Based on their abundance in membranes of different organelles and transport intermediates of the early secretory pathway, several different proteins, including ERGIC-53 (Schweizer *et al.*, 1988), were identified as potential cargo receptors. Their roles in sorting and transport were then confirmed by the observation that cells lacking a specific receptor displayed defective sorting and inefficient export of certain secretory proteins, whereas transport of other cargo proteins was unaffected (Dancourt & Barlowe, 2010).

When a COPII vesicle arrives at its final destination, soluble cargo proteins need to be released from their receptors. Cargo release has, in some cases, been suggested to be controlled by changes in pH and/or Ca^{2+} concentration between different organelles, which would affect the interaction between cargo and receptor (Barlowe, 2003).

Retrograde transport

Retrograde transport is a vital process of the secretory pathway; ER resident proteins that have escaped the ER retention system need to be retrieved, and components of the trafficking machinery need to be recycled. Such retrograde vesicular transport from the Golgi and the ERGIC is mediated by COPI vesicles. The COPI coat consists of seven coatomer subunits that are assembled at the ERGIC or Golgi membranes in a single step, initiated by the small GTPase Arf1 (Serafini *et al.*, 1991; Waters *et al.*, 1991).

The retrograde transport of soluble ER proteins is generally mediated by the transmembrane KDEL receptor, which binds proteins bearing a KDEL sequence motif (Lewis & Pelham, 1990). It has been proposed that binding of cargo to the receptor induces conformational changes that are transduced to the cytosolic side of the membrane, where they contribute to uptake into COPI vesicles (Dancourt & Barlowe, 2010). Proteins of the KDEL receptor family are highly conserved across species, although different versions of the C-terminal KDEL motif are recognised by different species (Lewis & Pelham, 1990; Lewis *et al.*, 1990). Also within a species the specificities of different isoforms of the KDEL receptor are known to vary; as an example, the three homologous human KDEL receptors (ERD21, ERD22, and ERD23) have distinct but overlapping substrate preferences and have been shown to recognise over 50 variants of the KDEL motif (Raykhel *et al.*, 2007). It has been shown that KDEL-receptor binding is strongly influenced by pH, and it has been proposed that the differences in pH between the compartments of the secretory pathway regulate the binding of cargo to the KDEL receptor in the Golgi and their release in the ER (Wilson *et al.*, 1993).

Another receptor mediating ER retrieval is the transmembrane protein Rer1p (Sato *et al.*, 1997). Unlike other known cargo receptors, Rer1p recognises and binds to sorting motifs in the transmembrane domain of its cargo (Sato *et al.*, 2001). Little is known about how the association of cargo with Rer1p is regulated to allow binding in the Golgi and the ERGIC and release in the ER.

Several different motifs that mediate direct interaction between membrane proteins and the COPI coat have been identified. One of the best characterised motifs for retrograde sorting is the dilysine Lys-Lys-X-X (KKXX) motif, which can be found at the cytoplasmic C-terminal end of some membrane proteins and interacts directly with subunits of the COPI coat (Jackson *et al.*, 1990). Another motif for ER retrieval is the arginine-based Arg-X-Arg motif, first identified in the ATP-sensitive K^+-channel (Zerangue *et al.*, 1999). Whereas the Lys-Lys-X-X motif has to be positioned close to the membrane surface for effective signal recognition, the Arg-X-Arg motif is functional only when spaced further from the membrane (Shikano & Li, 2003).

1.3 The ERGIC-53/MCFD2 transport receptor complex

ERGIC-53

One of the most extensively studied transport receptors is ERGIC-53, also known as LMAN1. This protein, named based on its localisation in the ER-Golgi intermediate compartment and its molecular mass of 53 kDa, was initially identified using a monoclonal antibody generated against Golgi membranes (Schweizer *et al.*, 1988). ERGIC-53 has since then been used as a marker for the ERGIC, where it is present at high concentrations, but it is also present in the ER and the *cis*-Golgi (Schweizer *et al.*, 1988; Hauri *et al.*, 2000), and in both COPII and COPI vesicles (Tisdale *et al.*, 1997; Wendeler *et al.*, 2007).

ERGIC-53 shares sequence and structural homology to leguminous lectins, a family of carbohydrate-binding proteins interacting specifically with mannose-rich

oligosaccharides in a Ca^{2+}-dependent manner (Itin *et al.*, 1996). This homology, in combination with sugar-binding studies, led to the suggestion of a role of ERGIC-53 in ER export of high-mannose glycoproteins. Since then, ERGIC-53 has been demonstrated to be involved in the transport of a number of glycoproteins, e.g. the lysosomal proteins cathepsin C (Vollenweider *et al.*, 1998) and cathepsin Z (Appenzeller *et al.*, 1999), and the glycosylated membrane protein nicastrin, a member of the γ-secretase complex (Morais *et al.*, 2006). Additional support for a role of ERGIC-53 as a glycoprotein receptor came from the finding that mutations in the *LMAN1* gene cause the bleeding disorder F5F8D (combined deficiency of factor V and factor VIII), which is characterised by impaired transport of the heavily glycosylated coagulation factors V and VIII (Nichols *et al.*, 1998). For cathepsin Z, it has been shown that both N-linked glycans and a surface-exposed β-hairpin loop are important for the association with ERGIC-53 (Appenzeller-Herzog *et al.*, 2005). A similar combination of a β-hairpin loop close to an N-linked oligosaccharide is also present in cathepsin C. In the interaction with coagulation factor VIII, a protein-protein interaction has also been shown to be a component of binding, in addition to the interaction with high-mannose glycans (Cunningham *et al.*, 2003).

ERGIC-53 contains a single transmembrane segment, a large N-terminal lumenal domain, and a short C-terminal cytosolic tail, which contains a diphenylalanine motif, governing interaction with COPII vesicles for anterograde transport (Kappeler *et al.*, 1997), and a dilysine motif, responsible for association with COPI vesicles for retrograde transport (Schindler *et al.*, 1993). The lumenal domain is composed of two subdomains, a carbohydrate-recognition domain (CRD) and a membrane-proximal stalk domain. Crystal structures of the CRD have revealed a large binding site capable of accommodating high-mannose glycans (Velloso *et al.*, 2002, 2003). The structures also support a Ca^{2+}-dependent mechanism for binding of carbohydrates; the presence of two Ca^{2+} ions stabilises and coordinates two loop regions in the vicinity of the glycan-binding site. The stalk domain is predicted to consist of four α-helices adopting a coiled-coil structure, and two cysteines in this domain form inter-molecular disulfide bonds that together with determinants in the transmembrane domain have been shown to mediate oligomerisation (Lahtinen *et al.*, 1999; Nufer *et al.*, 2003). ERGIC-53 has been demonstrated to exist as a homohexamer in the cell (Neve *et al.*, 2005), and it has been shown that the oligomerisation state of ERGIC-53 is important for its transport receptor function (Nufer *et al.*, 2003; Zheng *et al.*, 2010).

Although the mechanisms for some steps of ERGIC-53-mediated glycoprotein transport remain elusive, a model for its function as a transport receptor has been established (Figure 3). According to this model, hexameric ERGIC-53 binds folded glycoproteins in the ER and assists their transport with COPII vesicles to the ERGIC where dissociation of the cargo-receptor complex occurs as a consequence of the lower pH and Ca^{2+} concentration of this compartment. The glycoprotein is then further transported along the secretory pathway, while ERGIC-53 is recycled back to the ER in COPI vesicles (Hauri *et al.*, 2000).

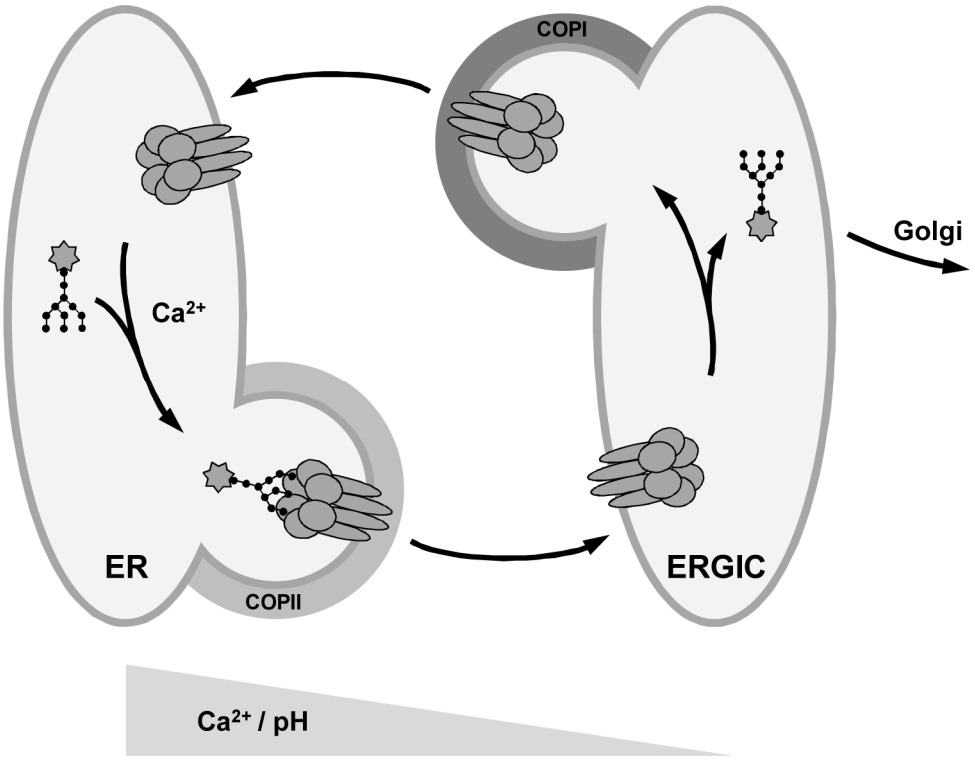


Figure 3. Schematic overview of the cargo receptor function and recycling of ERGIC-53. Hexameric ERGIC-53 binds to correctly folded glycoproteins in a Ca^{2+}-dependent manner in the ER. The diphenylalanine motif in the cytosolic tail of ERGIC-53 mediates interaction with the Sec23-24 complex and recruits the cargo-receptor complex to COPII vesicles. The vesicles then travel to and fuse with the ERGIC, where the cargo proteins are released from ERGIC-53 triggered by the lower pH and Ca^{2+} concentration in this compartment. The released cargo proteins are transported to the Golgi for further modifications, while the dilysine motif of ERGIC-53 interacts with the COPI components and mediates the packaging of the transport receptor into vesicles for recycling back to the ER. (Adapted from Hauri *et al.*, 2000.)

MCFD2

MCFD2 (multiple coagulation factor deficiency protein 2) was first linked to the secretory pathway when the gene encoding it was identified as the location of previously unexplained mutations causing F5F8D in patients with a wild-type *LMAN1* gene (Zhang *et al.*, 2003). MCFD2 is known to interact with ERGIC-53, forming a stable Ca^{2+}-dependent complex (Kawasaki *et al.*, 2008). It has also been demonstrated to interact with coagulation factor VIII independently of ERGIC-53, while interaction between ERGIC-53 and factor VIII requires complex formation with MCFD2 (Zhang *et al.*, 2005). Other cargo proteins are known to bind to ERGIC-53 in an MCFD2-independent way, and the requirement of MCFD2 for transport of coagulation factors V and VIII would suggest a role for MCFD2 as a specific recruitment factor for this subset of glycoproteins (Nyfeler *et al.*, 2006).

MCFD2 is a soluble 16 kDa protein with an N-terminal ER localisation sequence and two C-terminal Ca^{2+}-binding EF-hand motifs. MCFD2 has been shown to be localised

to the ER and the ERGIC, but in contrast to many other soluble ER proteins it lacks the KDEL sequence mediating ER retrieval, suggesting that localisation of MCFD2 is reliant on its interaction with ERGIC-53 (Zhang *et al.*, 2003).

Combined deficiency of factor V and factor VIII

Combined deficiency of factor V and factor VIII (F5F8D) is a rare bleeding disorder first described by Oeri *et al.* in 1954. The disease is characterised by reduced plasma concentrations of the blood coagulation factors V and VIII, and is manifested as a mild to moderate bleeding tendency with common symptoms such as nose bleeds, gum bleeding and easy bruising. The disease is known to be inherited and caused by mutations in either of two proteins: ERGIC-53 (Nichols *et al.*, 1998) or MCFD2 (Nichols *et al.*, 1998; Zhang *et al.*, 2003). To date, at least 36 different mutations in the *LMAN1* gene and 18 in the *MCFD2* gene have been reported (Zheng & Zhang, 2013). All of the mutations in *LMAN1* are null mutations except for two; one mutation causes a conformational change in the CRD and one breaks a disulfide bond important for oligomerisation (Yamada *et al.*, 2009; Zheng *et al.*, 2010). Some of the mutations in *MCFD2* are also null mutations, but in this gene missense mutations are more common than in *LMAN1*; eight of the reported mutations result in single amino acid substitutions (Zhang *et al.*, 2003; Zhang *et al.*, 2006; Jayandharan *et al.*, 2007; Ivaskevicius *et al.*, 2008; Zhang *et al.*, 2008; Abdallah *et al.*, 2010; Elmahmoudi *et al.*, 2011). Except for one mutation found in the binding surface to ERGIC-53, all of these mutations give rise to a disordered or destabilised protein, which is unable to form a complex with ERGIC-53, leading to an impaired transport of coagulation factors (Elmahmoudi *et al.*, 2011).

1.4 The Erv41-Erv46 complex

In an attempt to identify proteins involved in transport along the secretory pathway in yeast, Otte *et al.* (2001) performed a reconstituted COPII vesicle budding assay in combination with proteomics studies. They identified the previously uncharacterised ER vesicle proteins Erv41p and Erv46p, named based on their localisation and molecular mass. Erv41p and Erv46p are both integral membrane proteins sharing the same overall topology, with one single large lumenal domain, two membrane-spanning helices close to the termini of the protein, and short N- and C-terminal tails on the cytosolic side of the membrane. The expression levels of Erv41p and Erv46p have been shown to be co-dependent, and immunoprecipitation experiments have demonstrated that the two proteins form a stable complex, most likely composed of one subunit of each.

The C-terminal tails of both Erv41p and Erv46p contain a hydrophobic sequence that controls sorting into COPII vesicles for anterograde transport from the ER (Otte & Barlowe, 2002). The hydrophobic isoleucine-leucine sequence in the Erv41p tail is sufficient for binding of the complex to the COPII coat, but its combination with the phenylalanine-tyrosine sequence of Erv46p is required for export in COPII vesicles. Sucrose gradient fractionation has shown that the major fraction of the complex is located to the Golgi, with approximately 70% of all Erv46p, and a minor fraction is present in the ER. Removal of any of the COPII localisation signals leads to accumulation of the complex in the ER. Only Erv46p contains a signal for retrograde

transport; a conserved KKXX motif in its C-terminal tail binds to COPI vesicles (Otte & Barlowe, 2002; Orci *et al.*, 2003), indicating that retrieval of Erv41p to the ER is dependent on binding to Erv46p.

Erv41p and Erv46p are both conserved across eukaryotic species, and the human homologues, known as ERGIC2 or hErv41, and ERGIC3 or hErv46, share 30% and 41% sequence identity, respectively, with their yeast counterparts. Erv41p and Erv46p have been demonstrated to be parts of a large heterogeneous complex with a molecular mass between 200 and 400 kDa (Breuza *et al.*, 2004; Welsh *et al.*, 2006). In humans, ERGIC1, also known as ERGIC-32, has been shown to be a part of the ERGIC2-ERGIC3 complex through its interaction with ERGIC3 (Breuza *et al.*, 2004). ERGIC1 belongs to the same protein family as Erv41p and Erv46p, and shares their topology with a larger lumenal domain between two transmembrane segments. In contrast to Erv41p and Erv46p, ERGIC1 is mainly localised to the ERGIC. ERGIC1 is not believed to associate permanently with ERGIC2 and ERGIC3, but it has been suggested to stabilise monomeric ERGIC3 in the ER, thereby promoting correct assembly of the ERGIC2-ERGIC3 complex. It has also been demonstrated that the ER-localised glucosidase II, Rot2p, which trims the two inner glucose residues from N-linked oligosaccharides after glucosidase I has removed the most distal glucose, interacts with the Erv41p-Erv46p complex in yeast (Welsh *et al.*, 2006).

The exact function of the Erv41p-Erv46p complex in the early secretory pathway is not known. It has been proposed to play a role in sorting of cargo into transport vesicles, to help localising other components of the early secretory pathway, or to be involved in transport of lipids, but evidence now points towards a role in glycoprotein processing or transport (Otte *et al.*, 2001). An *in vitro* assay showed that absence of either or both of the two proteins led to a decreased glycoprotein transport between ER and Golgi, suggested to be caused by a defect at the stage of vesicle fusion (Otte *et al.*, 2001), and yeast strains lacking a cycling Erv41p-Erv46p complex display a mild glycoprotein processing defect (Welsh *et al.*, 2006).

Recently, both ERGIC1 and ERGIC3 have been linked to different types of cancer in human. The *ERGIC1* gene has been shown to be highly expressed in prostate cancer tissue, and silencing of *ERGIC1* induced antiproliferative effects in an ERG (v-ets erythroblastosis virus E26 oncogene homolog, avian) positive prostate cancer cell line (Vainio *et al.*, 2012). Results also indicated that silencing of *ERGIC1* could downregulate ERG mRNA expression, and it was suggested to be a potential drug target for ERG oncogene expressing tumours.

The *ERGIC3* gene has been demonstrated to be upregulated by the microRNA *miR-490-3p* in hepatocellular carcinoma tissues and cell lines (Zhang *et al.*, 2013). It was shown that *ERGIC3* enhanced cell viability and colony formation, as well as stimulated cell migration and invasion ability. Data also indicated that *ERGIC3* stimulated epithelial to mesenchymal transition, a process important for metastasis.

ERGIC3 was also found to be upregulated in lung cancer tissues and in lung cancer cell lines, and the ERGIC3 protein was overexpressed in both tumours and cell lines (Wu *et al.*, 2013). ERGIC3 was shown to have an effect on both cell proliferation and cellular

migration; reduced expression of ERGIC3 in a lung cancer cell line reduced the rate of proliferation and inhibited cellular migration, and overexpression of ERGIC3 in a normal cell line increased the rate of proliferation and promoted cellular migration. Based on the importance of the yeast homologue of ERGIC3, Erv46p, and its complex partner Erv41p in glycoprotein processing and transport (Welsh *et al.*, 2006), it was speculated that the mechanism by which decreased expression of ERGIC3 affects cancer cell proliferation and migration could involve disruption of glucosidase activity and intracellular protein transport.

1.5 The unfolded protein response

The quality control machinery of the ER has evolved to ensure that no misfolded proteins are forwarded from the organelle; proteins that are not correctly processed or assembled are either refolded or designated for degradation. However, in some physiological and pathological conditions the capacity of the quality control is exceeded, leading to accumulation of misfolded proteins, a state termed ER stress. In order to cope with the ER stress and restore the normal functions of the ER, pathways collectively known as the unfolded protein response (UPR) are activated.

Activation of the UPR is now known to have effects both on transcriptional and translational levels, and leads ultimately to reduced protein synthesis, increased capacity to fold nascent polypeptides in the ER, and increased retrotranslocation and degradation of misfolded proteins (Zhang & Kaufman, 2006). It is not only genes involved in protein folding, processing and degradation that are upregulated by the UPR, but also genes encoding proteins important for vesicle trafficking between the ER and the Golgi, or proteins with roles in lipid and inositol metabolism. This proves the importance of the unfolded protein response for integration of the functions of the secretory pathway and maintenance of ER homeostasis (Barlowe & Miller, 2013). In cases when ER homeostasis cannot be re-established, prolonged UPR signalling leads to apoptosis, thereby protecting the organism from accumulation of cells that are unable to ascertain the fidelity of protein production (Walter & Ron, 2011).

There are three major regulators of the metazoan UPR (Figure 4): IRE1 (inositol-requiring enzyme 1), ATF6 (activating transcription factor 6), and PERK (RNA-activated protein kinase-like ER kinase). These are all transmembrane proteins which sense abnormal conditions in the ER lumen and transmit signals across the membrane to the cytosol, from where they are further passed on to the nucleus (Gardner *et al.*, 2013).

IRE1 is the regulator of the most conserved branch of the UPR, and is the only ER stress sensor present in all eukaryotes (Mori, 2009). Apart from its N-terminal lumenal sensor domain, IRE1 contains cytosolic kinase and RNase domains at its C-terminus (Gardner *et al.*, 2013). Upon accumulation of misfolded proteins in the ER, IRE1 oligomerises and is phosphorylated *in trans* by other IRE1 molecules, leading to activation of the kinase and RNase domains (Shamu & Walter, 1996). The activated RNase domain specifically cleaves its mRNA substrate, yielding a spliced mRNA that is translated to the active form of the UPR-specific transcription factors XBP1 in

metazoans (Yoshida et al., 2001), or HAC1 in yeast (Cox & Walter, 1996). These transcription factors, in turn, directly mediate upregulation of UPR target genes, such as ER chaperone genes and genes encoding components of the ERAD machinery (Travers *et al.*, 2000; Yoshida *et al.*, 2001). In metazoans, IRE1 also mediates degradation of a subset of mRNAs encoding membrane and secreted proteins, thereby decreasing the uptake of proteins to the ER (Hollien & Weissman, 2006).

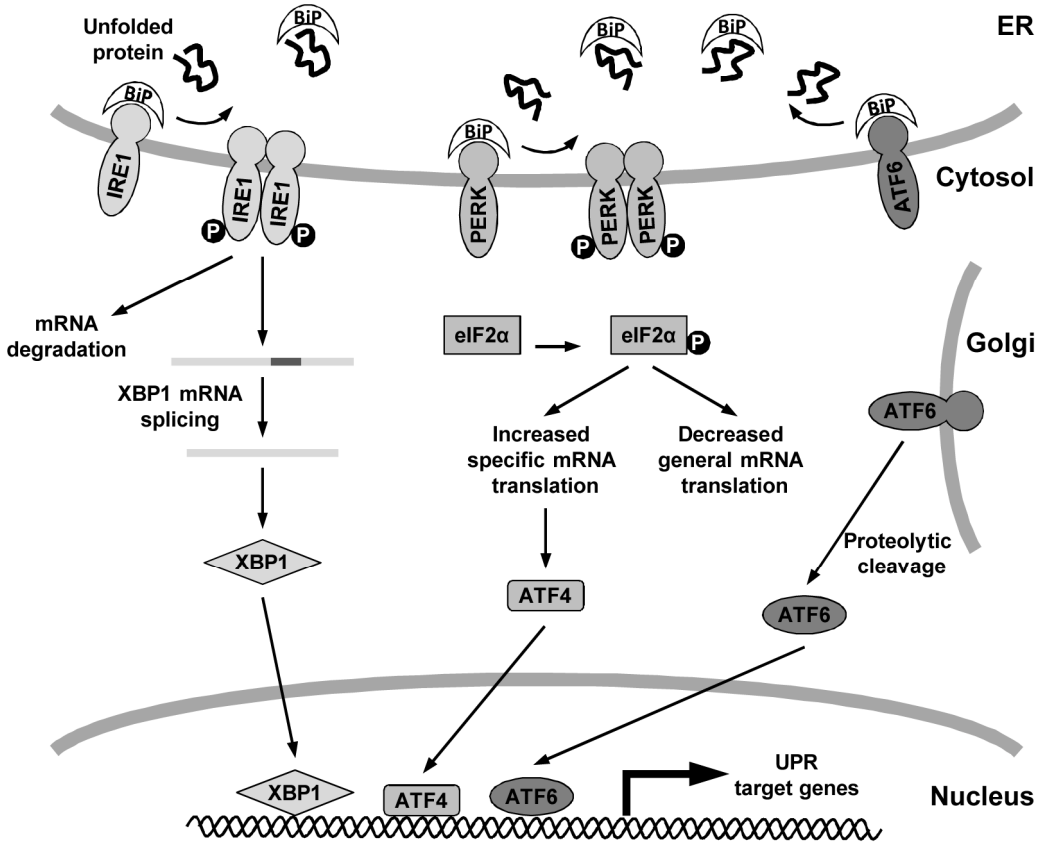


Figure 4. Schematic overview of the unfolded protein response. In unstressed cells BiP binds to the lumenal domains of the three stress sensors IRE1, PERK, and ATF6 and keep them in an inactive state. In the event of ER stress, BiP is sequestered from the stress sensors to bind to un- or misfolded proteins, and the released sensors are activated. Activation of IRE1 leads to splicing of the mRNA encoding the transcription factor XBP1, and the translated transcription factor induces expression of different UPR target genes. IRE1 also mediate degradation of specific mRNAs, leading to a decreased amount of proteins translocated into the ER. Activated PERK phosphorylates eIF2α, inhibiting global mRNA translation to reduce the protein load in the ER, and increasing the specific translation of mRNA encoding the transcription factor ATF4, which activates UPR target genes. Upon activation, ATF6 travels to the Golgi where it is cleaved, and its cytosolic transcription factor domain is released and free to enter the nucleus to activate transcription of UPR target genes. (Adapted from Cao & Kaufman, 2012.)

In response to ER stress, autophosphorylation and activation of the cytoplasmic kinase domain of PERK enable phosphorylation of eukaryotic translation initiation factor 2 α (eIF2α) (Harding *et al.*, 1999). This inhibits general mRNA translation, resulting in a decreased protein load in the ER, but also increases the translation of some specific mRNAs, such as that encoding the transcription factor ATF4, which activates

downstream UPR target genes (Harding *et al.*, 2000). One of these encodes the transcription factor CHOP, which controls genes involved in apoptosis (Wang *et al.*, 1998; Zinszner *et al.*, 1998).

The third UPR regulator, ATF6, is synthesised as a transmembrane protein with a lumenal ER stress-sensing domain and a cytosolic bZIP (basic leucine zipper) transcription factor domain. Upon exposure to high levels of misfolded proteins, it travels from the ER to the Golgi apparatus where it is proteolytically cleaved, and the transcription factor domain is released and enters the nucleus (Haze *et al.*, 1999). There, it binds to the ER stress response element and activates transcription of UPR target genes similar to those transcribed in response to IRE1 signalling (Yoshida *et al.*, 1998; Haze *et al.*, 1999; Adachi *et al.*, 2008).

An additional central player in the UPR is the molecular chaperone BiP. BiP has been shown to form stable complexes with the lumenal stress-sensing domains of IRE1, PERK and ATF6 in unstressed cells, keeping them in an inactive state, and to dissociate from the UPR regulators in the event of ER stress (Bertolotti *et al.*, 2000; Okamura *et al.*, 2000; Ma *et al.*, 2002; Shen *et al.*, 2002a). BiP is also known to bind directly to newly synthesised proteins in a transient manner, and more permanently to misfolded or unassembled proteins (Gething, 1999). According to the most supported view of how ER stress initiates the UPR, high concentrations of unfolded proteins compete with the ER stress sensors for BiP-binding (Kimata *et al.*, 2003; Carrara *et al.*, 2013). When BiP is sequestered from the lumenal sensor domains, they are activated and the stress signal can be propagated *via* their cytosolic domains. Moreover, BiP itself is one of the proteins that are upregulated upon accumulation of misfolded proteins and UPR activation (Kozutsumi *et al.*, 1988). It has been suggested that BiP expression is feed-back regulated on a translational level, helping the cell to fine-tune BiP levels in order to prevent ill-timed UPR (Gulow *et al.*, 2002).

1.6 The Hsp70-Hsp40 chaperone system

The heat shock response and the heat shock proteins were first discovered by Ritossa in the 1960s, when he discovered that elevated temperatures gave rise to a different pattern of RNA synthesis in *Drosophila melanogaster* (Ritossa, 1962, 1996). Since then, it has been shown that a number of environmental and pathological events, including those that activate the unfolded protein response, increase the expression of heat shock proteins. Although first discovered in connection with the response to cellular stress, members of the heat shock families are also expressed in non-stressed conditions, acting as chaperones that are important for the quality control machinery of the secretory pathway.

One of the most well-studied heat shock protein families is the Hsp70 (70 kDa heat shock protein) family. The chaperones within this family are highly conserved and are expressed in all cells and organelles (Awad *et al.*, 2008). In humans there are at least 13 identified members of the Hsp70 family (Vos *et al.*, 2008), and in most other organisms multiple Hsp70 chaperones are also expressed (Meimaridou *et al.*, 2009). The Hsp70 proteins are involved in many diverse processes including prevention of aggregation

and promotion of folding of nascent proteins, un- and refolding of misfolded or aggregated proteins, and translocation of proteins across membranes (Mayer & Bukau, 2005).

Hsp70 chaperones recognise proteins that are not natively folded *via* short degenerate motifs consisting of a core of four to five residues enriched in hydrophobic amino acids and flanked by regions enriched in basic residues (Rudiger *et al.*, 1997). Such motifs are common, and occur on average every 30-40 residues in most proteins. In the native folded protein, these motifs are usually buried in the hydrophobic core of the protein, but before and during folding, or upon denaturation, they are exposed and can be detected by the Hsp70 proteins, allowing the chaperones to distinguish between un- or misfolded proteins and their native counterparts (Mayer, 2013).

Un- or misfolded substrates cycle between the free and Hsp70-bound states until all molecules have obtained their native conformation, but the exact mechanism by which folding is assisted by the chaperone is still unclear. Two mechanisms have been proposed: Either the Hsp70 acts by stabilising unfolded intermediates, keeping the free substrate concentration sufficiently low to prevent aggregation and allow correct folding, or it actively helps partially unfolding the substrates, yielding productive folding intermediates that then fold spontaneously into their native states (Slepenkov & Witt, 2002).

The Hsp70 proteins contain an N-terminal nucleotide-binding domain (NBD) connected *via* a flexible linker to a C-terminal substrate-binding domain (SBD), which can be further subdivided into a β-sandwich domain, containing a substrate-binding pocket, and an α-helical lid (Mayer, 2013). The arrangement of the SBD changes depending on the nucleotide-state of the Hsp70. Binding of ATP to the NBD induces a conformational change in the SBD, which opens the lid from the substrate-binding pocket (Figure 5). In this state, substrates associate quickly with the chaperone, but the affinity is low. Binding of a substrate to the SBD stimulates the ATPase activity of the NBD, and hydrolysis of ATP induces a closing of the SBD lid over the substrate-binding pocket and entrapment of the substrate. The signal arising from ATP hydrolysis is transmitted within the NBD through a hydrogen bond network, starting at the catalytic center and converging onto a conserved surface-exposed arginine at the interface between the NBD and the SBD (Vogel *et al.*, 2006).

The activity of Hsp70 proteins is regulated by co-chaperones and nucleotide exchange factors (NEFs) (Liberek *et al.*, 1991). Co-chaperones from the Hsp40 (40 kDa heat shock protein) family, also known as the DnaJ family after the *Escherichia coli* Hsp40 homologue DnaJ, bind to their Hsp70 partners and stimulate their ATPase activity, thereby assisting in keeping the unfolded protein associated with the Hsp70 chaperone (Cyr *et al.*, 1992; Liberek *et al.*, 1991). Another important function of Hsp40 proteins is to bind and deliver specific client proteins to Hsp70 chaperones (Cheetham & Caplan, 1998). The interaction with Hsp70 is mediated by a J domain, containing a conserved histidine-proline-aspartate (HPD) motif known to be important for the co-chaperone function (Tsai & Douglas, 1996), but apart from this domain, the Hsp40 proteins exhibit large diversity. Three types of Hsp40 proteins have been distinguished (Cheetham & Caplan, 1998). Type I Hsp40 proteins show full domain conservation

with *E. coli* DnaJ; they contain, apart from the J domain, a Gly/Phe-rich region and a Cys-rich, zinc finger-like region. Type II Hsp40 proteins contain the J domain and the Gly/Phe-rich region, whereas type III Hsp40 proteins only share the J domain. There are more Hsp40 proteins than Hsp70 proteins, and the different properties of the various co-chaperones are believed to contribute to the versatile functions of the Hsp70 family (Mayer & Bukau, 2005; Kampinga & Craig, 2010). The nucleotide exchange factors are also important for the function of Hsp70 chaperones; they promote the release of ADP and the binding of ATP, and trigger thereby unloading of the bound substrate and recycling of the Hsp70 (Kampinga & Craig, 2010). Unlike the Hsp40 proteins, which all share the J domain responsible for their regulation of Hsp70, four different types of NEFs, without any sequence similarity, have been identified. Some well-known NEFs are GrpE, which releases ATP from DnaK, the *E. coli* Hsp70 homologue (Liberek *et al.*, 1991), and the BAG family of NEFs, which interact with cytosolic Hsp70 proteins in eukaryotes (Hohfeld & Jentsch, 1997).

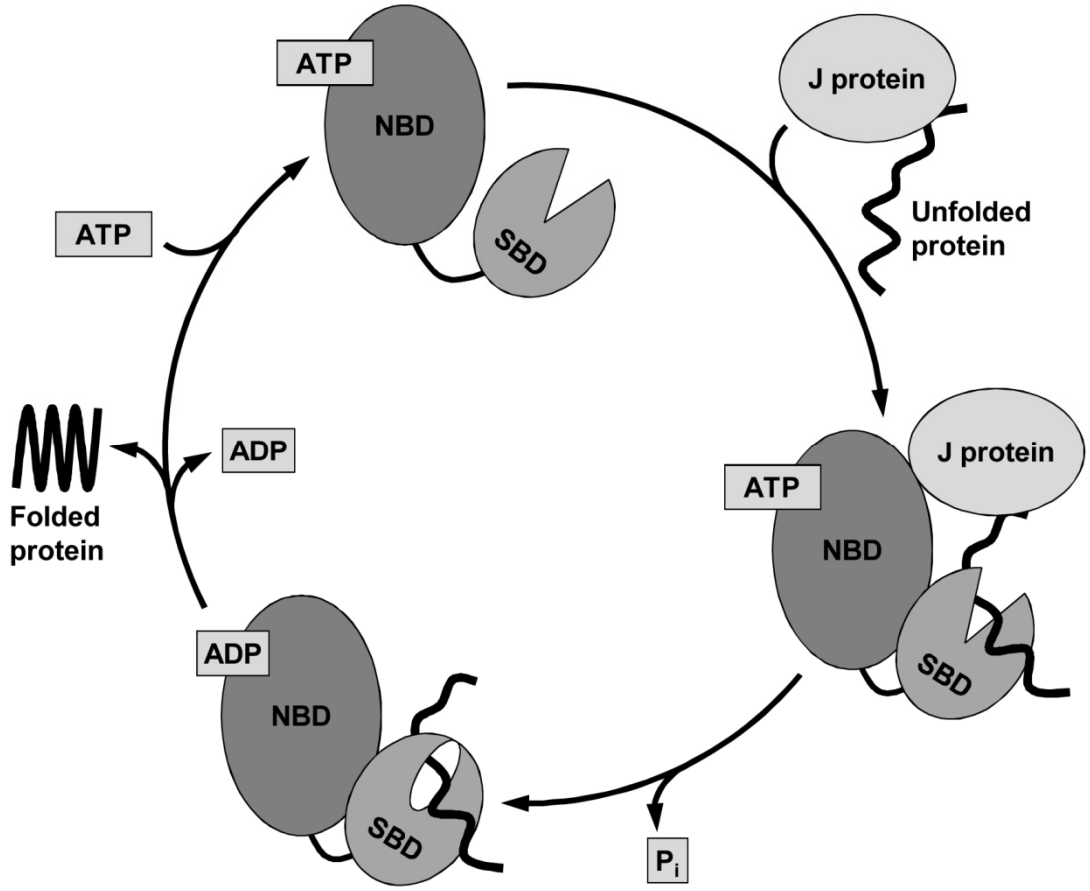


Figure 5. The folding cycle of Hsp70 proteins. Unfolded or misfolded proteins are delivered to the open ATP-bound Hsp70 by a J domain protein. The J domain stimulates the ATPase activity of Hsp70, leading to a conformational change that traps the client protein in the substrate-binding domain (SBD) where folding is promoted. Nucleotide exchange factors stimulate the exchange of ADP to ATP in the nucleotide-binding domain (NBD), leading to an opening of the SBD and release of the folded client protein. (Adapted from Meimaridou *et al.*, 2009.)

1.6.1 BiP and the ERdj proteins

The mammalian ER orthologue of Hsp70 is BiP (immunoglobulin heavy-chain binding protein), also known as GRP78 (glucose-regulated protein of 78 kDa) (Shiu *et al.*, 1977; Haas & Wabl, 1983; Munro & Pelham, 1986). BiP is involved in many of the processes in the ER, including folding of nascent proteins (Lee *et al.*, 1999), targeting misfolded proteins for degradation (Nishikawa *et al.*, 2001), gating the translocon (Hamman *et al.*, 1998), and regulating the unfolded protein response (Bertolotti *et al.*, 2000; Shen *et al.*, 2002a). The different functions of BiP are regulated by the Hsp40 homologues in the ER, the ER-localised DnaJ-containing (ERdj) proteins, of which at least seven have been identified to date (Figure 6) (Otero *et al.*, 2010; Vembar *et al.*, 2010).

The transmembrane protein ERdj1, also known as Mtj1, consists of a lumenal J domain and a cytosolic domain which interacts with translating ribosomes (Dudek *et al.*, 2005). When BiP is not bound to the J domain, protein synthesis is inhibited, ensuring that BiP is always present when newly synthesised proteins enter the ER.
ERdj2, or Sec63, is also a transmembrane protein with a lumenal J domain. It resembles yeast Sec63p, which plays a role in posttranslational translocation, but since mammalian translocation into the ER mainly occurs cotranslationally, its function is not clear (Meyer *et al.*, 2000).
The soluble ER protein ERdj3, also known as HEDJ, is upregulated during ER stress, and is suggested to assist the recruitment of BiP to nascent unfolded proteins or proteins unable to fold correctly (Shen & Hendershot, 2005).
ERdj4 (MDG-1), anchored in the ER membrane *via* its uncleaved signal sequence and with the main part of the protein in the ER lumen, is upregulated during ER stress and has been suggested to be involved in protein folding or ER-associated degradation (Shen *et al.*, 2002b).
ERdj5, or JPDI, is a soluble protein with four thioredoxin-like domains in addition to the J domain (Cunnea *et al.*, 2003; Hosoda *et al.*, 2003). It has been suggested to reduce non-native disulfide bonds in BiP substrates and assist in their refolding, as well as targeting terminally misfolded proteins for elimination *via* ERAD (Dong *et al.*, 2008; Oka *et al.*, 2013).
ERdj6, also known as $P58^{IPK}$ or DnaJC3, is mainly located in the ER lumen where it acts as a co-chaperone for BiP (Rutkowski *et al.*, 2007; Petrova *et al.*, 2008). It binds unfolded substrates *via* its tetratricopeptide repeat (TPR) domain, and is thought to assist in promoting protein folding in the ER lumen.
ERdj7 consists of two transmembrane segments and a lumenal J domain. Its function remains to be determined, but it is not upregulated by ER stress, and reduced levels of ERdj7 do not induce the UPR, implying that it is not likely to play a role in protein folding or degradation (Zahedi *et al.*, 2009).

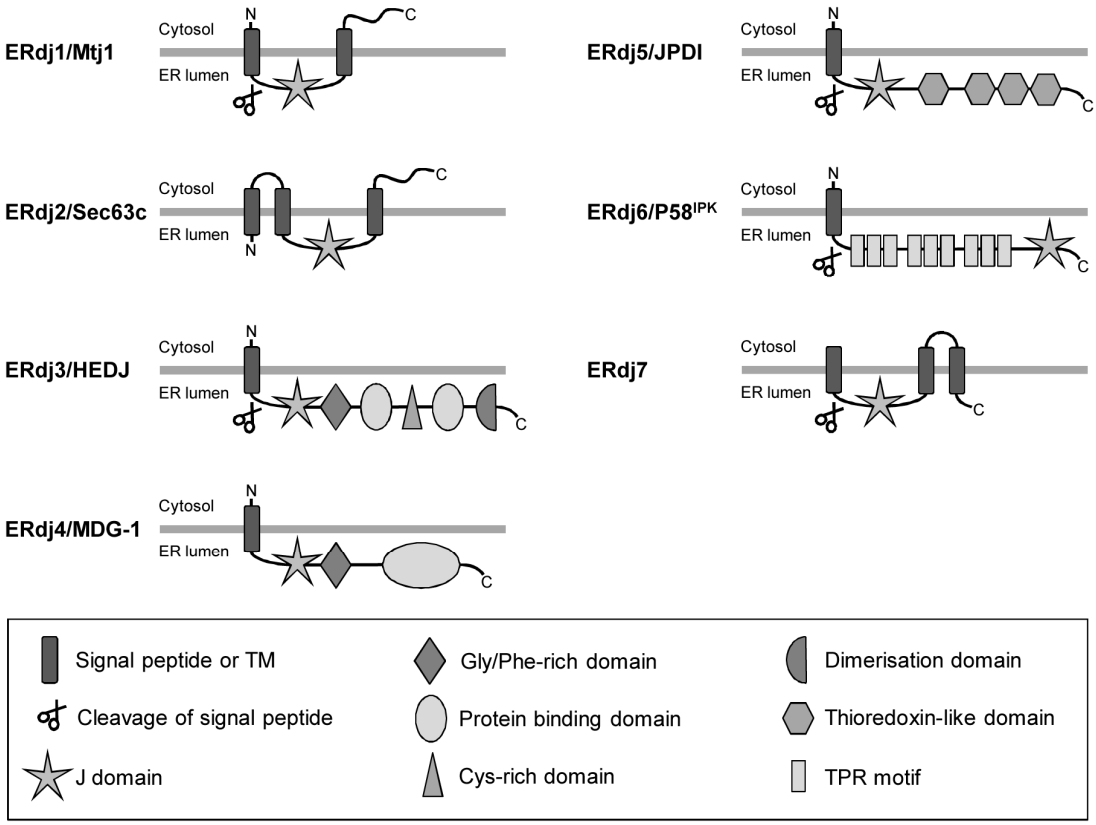


Figure 6. The ERdj proteins. The ERdj proteins all share the J domain, but are otherwise different. ERdj1, ERdj2, ERdj4, and ERdj7 contain transmembrane segments, whereas ERdj3, ERdj5, and ERdj6 are soluble. ERdj3 is a type I J domain protein with a Gly/Phe-rich region and a Cys-rich region, ERdj4 is a type II J domain proteins with a Gly/Phe-rich domain, and the remaining family members are type III J domain proteins. (Adapted from Otero *et al.*, 2010.)

1.7 P58^{IPK} (ERdj6)

P58^{IPK} (58 kDa inhibitor of protein kinase) was first identified as an inhibitor of the cytosolic protein PKR (protein kinase RNA-activated) in influenza infection (Lee *et al.*, 1990; Barber *et al.*, 1994). During viral infection, dsRNA-activated PKR phosphorylates the α-subunit of eIF2α, leading to protection of the cell through decreased synthesis of host and viral proteins. To prevent this, viruses recruit P58^{IPK} to inactivate PKR, thereby permitting continued synthesis of viral proteins (Lee *et al.*, 1990). The P58^{IPK}-mediated repression of PKR is, in turn, inhibited *via* binding of another protein, P52^{rIPK} (52 kDa repressor of the inhibitor of protein kinase), to P58^{IPK} (Gale *et al.*, 1998).

Later, the transcription of the *P58^{IPK}* gene was shown to be induced in response to ER stress (Yan *et al.*, 2002; van Huizen *et al.*, 2003). P58^{IPK} binds to the cytosolic part of PERK, one of the key regulators of the UPR, inhibiting PERK-mediated phosphorylation of eIF2α (Yan *et al.*, 2002). Thereby P58^{IPK} has been suggested to act as a component of a negative feedback loop, used to attenuate eIF2α signalling at the later stages of the UPR, when protein synthesis is needed.

The importance of P58^{IPK} for regulation of the stress response was demonstrated by knockout studies in mice (Ladiges *et al.*, 2005). After deletion of the *P58^{IPK}* gene, mice exhibited a diabetic phenotype, with glucosuria, hyperglycemia, and hypoinsulinemia, as well as increased apoptosis of pancreatic β-cells. These secretory cells are likely particularly susceptible to accumulation of unfolded proteins and ER stress, and the results are consistent with an important role for P58^{IPK} in the unfolded protein response.

Studies of mouse embryonic fibroblasts revealed a reduced ability to cope with misfolded proteins in the absence of P58^{IPK}, and based on these results a role for P58^{IPK} in protein processing and maturation was proposed (Rutkowski *et al.*, 2007). Since the major site of protein folding is the ER, this suggested that P58^{IPK} is located in the ER lumen rather than, as was previously thought, peripheral to the ER membrane. Reexamination of the localisation of P58^{IPK} then indeed showed that the protein is translocated into the ER directed by a cleavable N-terminal signal sequence, and that it is mainly located to the lumen, both in stressed and non-stressed cells. However, inefficiency of the signal sequence has been suggested to result in a small fraction of the protein not being translocated into the ER, but being able to perform its cytosolic interactions with PKR and PERK.

P58^{IPK} belongs to the DnaJ family or Hsp40 family of proteins that interact with chaperones of the Hsp70 type. This, in combination with its localisation to the ER lumen, gave rise to the hypothesis that P58^{IPK} could interact with the Hsp70 homologue of the ER, BiP (Rutkowski *et al.*, 2007). The interaction between P58^{IPK} and BiP was proven by immunoprecipitation; BiP from liver microsomes co-purified P58^{IPK}, and *vice versa*. This has led to the suggestion that P58^{IPK} acts as a co-chaperone for BiP. Further, it has also been demonstrated that P58^{IPK} can make direct interactions with unfolded protein substrates (Tao *et al.*, 2010). Although the mechanism of this co-chaperone function is still not completely clear, P58^{IPK} has been proposed to bind an unfolded protein substrate, stabilise it, and then, after binding to BiP, deliver the substrate to the chaperone.

In P58^{IPK}, the N-terminal ER translocation signal sequence is followed by nine tetratricopeptide repeats and a C-terminal J domain. Tetratricopeptide repeats are motifs of 34 amino acids with a degenerate consensus sequence defined by a pattern of small and large hydrophobic residues (Zeytuni & Zarivach, 2012). When comparing TPR motifs from functionally unrelated proteins, conservation is limited to eight consensus residues at positions 4, 7, 8, 11, 20, 24, 27, and 32, whereas functionally equivalent TPR motifs also can exhibit conservation outside of these residues (Blatch & Lassle, 1999). The TPR motif adopts a helix-turn-helix fold. Often, proteins contain tandem arrays of up to 16 TPR motifs, and when multiple TPR motifs are connected, they form a series of antiparallel α-helices. The angle between two adjacent helices is then such that a curvature is created, resulting in one concave and one convex surface (Zeytuni & Zarivach, 2012). TPR motifs are known to participate in protein-protein interactions in a variety of cellular processes. The ligand bound by the TPR can adopt an α-helical or extended coil conformation, or a combination of both.

The J domain, conserved between members of the Hsp40 family, is composed of two shorter and two longer α-helices. In the turn between the two longer helices is located

the HPD motif; a motif known to be important for the Hsp40-mediated stimulation of the ATPase activity of Hsp70 proteins. The positioning of the J domain at the C-terminus of $P58^{IPK}$ is unusual; more commonly it is found at the N-terminus of proteins (Cyr *et al.*, 1994).

To date, no structural information is available on $P58^{IPK}$ in complex with any of its interaction partners. However, various studies have localised the different interactions to specific parts of the protein. With respect to the BiP interaction, it has been shown by co-immunoprecipitation experiments that the J domain is important; a $P58^{IPK}$ construct without the J domain did not interact with BiP (Rutkowski *et al.*, 2007). Yeast two-hybrid experiments have suggested that the interaction between $P58^{IPK}$ and PKR involves TPR6 (Gale *et al.*, 1996; Tang *et al.*, 1996), and TPR7 has been proposed to be the binding site for $P52^{rIPK}$ (Gale *et al.*, 2002).

2 AIM OF THE THESIS

The overall aim of this thesis project was to gain a deeper understanding of different aspects of the early secretory pathway. Members of three different protein-protein complexes have been structurally characterised by X-ray crystallography and NMR spectroscopy, and further investigations have been performed.

The specific aims of this project have been:

- To determine the structure of MCFD2 and to give insights into its Ca^{2+}-dependent interaction with ERGIC-53.

- To clarify the mechanism by which mutations in MCFD2 cause the blood coagulation factor deficiency disorder F5F8D.

- To determine the structure of Erv41 and to increase the understanding of its role in the cell.

- To determine the structure of $P58^{IPK}$ and to further investigate its functions.

- To investigate the interaction between $P58^{IPK}$ and BiP in order to give insights into how co-chaperones regulate the different functions of the chaperone.

3 RESULTS AND DISCUSSION

3.1 NMR studies and biophysical characterisation of MCFD2 (Paper I)

When the work described in paper I was initiated, the importance of MCFD2 for efficient transport of the blood coagulation factors V and VIII had been clearly established, but its precise role in these processes still remained elusive. Ca^{2+} had been shown to be important for complex-formation with ERGIC-53, but little structural information was available to explain the mechanism behind this observation. In order to increase the understanding of the function of MCFD2, its interactions with ERGIC-53, and the mechanism by which mutations in the protein cause disease, the initial aim was to determine the structure of human MCFD2 by X-ray crystallography. However, despite extensive crystallisation screening, no crystals could be produced. Instead, NMR studies to determine the structure of the wild-type protein and to characterise several mutations that are found in patients with the bleeding disorder F5F8D, were performed in collaboration with Professor Torleif Härd at the Swedish NMR Centre at the University of Gothenburg.

3.1.1 NMR spectroscopy and structure determination

An MCFD2 construct encoding the full mature protein (residues 27-146), excluding the ER-localisation signal sequence, was used to produce native and isotope-labelled (^{15}N and $^{15}N/^{13}C$) protein in *E. coli* for the NMR experiments. Initial measurements were performed in Ca^{2+}-free phosphate buffer, resulting in spectra with the characteristics of an unfolded protein, with no evidence of secondary structure or of a hydrophobic core. In the ^{15}N-HSQC spectrum, this is seen as poorly resolved amide resonances in the random coil region, around 8.0 ppm (Figure 7A). Exchange into MES buffer and addition of Ca^{2+} ions resulted in a markedly different spectrum, showing a number of the well-dispersed peaks associated with presence of secondary and tertiary structure, although some disorder still remains (Figure 7B). Binding of Ca^{2+} causes functionally important conformational changes in many other EF-hand proteins (Yap *et al.*, 1999), but few are entirely dependent on Ca^{2+} for folding, and the complete loss of secondary structure observed for MCFD2 in the absence of Ca^{2+} is unusual.

Because of the apparent Ca^{2+}-dependent folding of MCFD2, all NMR spectra used for the structure calculations were recorded in the presence of $CaCl_2$. Several multidimensional heteronuclear NMR spectra were recorded to obtain the resonance assignments of ^{1}H, ^{13}C, and ^{15}N nuclei for the structure calculation. Despite the observed partial disorder in MCFD2, nearly complete backbone assignments were achieved, with the exception of a small number of residues in the unstructured region. Side-chain resonance assignment was almost complete in the ordered region of the structure, whereas over-crowding of the random coil regions made the assignment of several side-chains in the disordered regions impossible. A total of 2142 distance restraints were identified by nuclear Overhauser enhancement (NOE) assignments. Very few medium- or long-range restraints could be localised to the N-terminal

residues 27-66 and residues 102-112, localising the disordered parts of the protein to these regions. Based on the assignments, secondary structure elements were defined and the final ensemble of MCFD2 structures was calculated by automated iterative NOE assignment and structure calculation by simulated annealing.

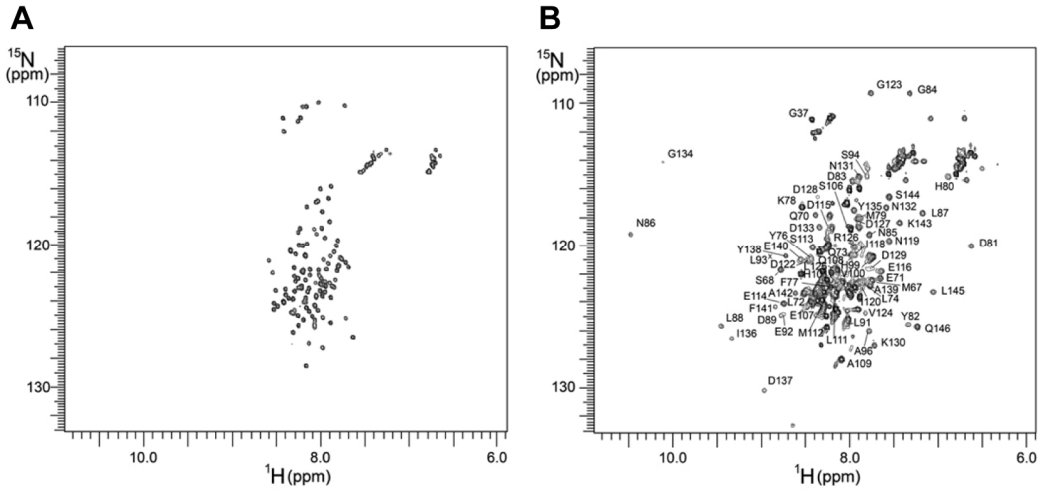


Figure 7. NMR spectra of MCFD2. (**A**) The ^{15}N-HSQC spectrum recorded in the absence of Ca^{2+}. (**B**) The ^{15}N-HSQC spectrum recorded after addition of 10 mM $CaCl_2$, with the assignments indicated for all backbone NH resonances in the ordered region.

3.1.2 Overall structure of MCFD2

The disordered N-terminus of MCFD2, extending from residue 27 to residue 69, is followed by four α-helices and two short β-strands forming two EF-hand motifs (Figure 8). These EF hands are packed against each other to form one single domain, with a fold similar to that of other EF-hand proteins such as parvalbumin and S100 proteins (Lewit-Bentley & Rety, 2000). As in other EF hands, one Ca^{2+} ion is coordinated in the loop connecting the two helices of each motif.

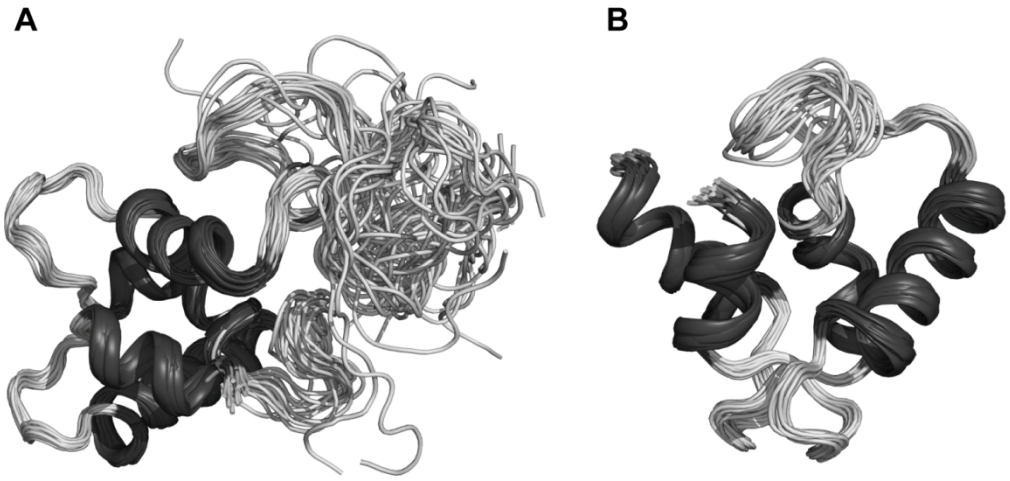


Figure 8. Overall NMR structure of MCFD2. (**A**) The ensemble of the 20 lowest-energy structures, showing the complete structure including the disordered N-terminal part. (**B**) The ordered region of MCFD2, residues 67-146 (PDB ID 2VRG).

Two features that distinguish MCFD2 from its structural homologues, the closest of which is the C-terminal domain of calmodulin with an RMSD (Root Mean Square Deviation) of 1.9 Å over 57 residues, are the significantly longer loop connecting the two EF-hand motifs, and the second helix, which is more bent and packed more tightly into the structure than in homologous structures. Backbone amide resonance broadening, indicative of conformational interconversions on the microsecond to millisecond time scales, was observed for residues in this region, suggesting that it might be able to adopt different conformations. This was later confirmed by the crystal structure of the complex between MCFD2 and ERGIC-53, which revealed conformational changes in this part of MCFD2 upon complex formation (Wigren *et al.*, 2010). In the complex, the second helix, which is bent in the NMR structure, is straightened, and the subsequent helix is tilted by approximately 20°.

Another interesting feature of MCFD2 is the disordered N-terminus. This region shows very little sequence identity to most other proteins, but is highly conserved in orthologues of MCFD2, implying a potential role in the function of the protein. One hypothesis is that this unstructured part is involved in binding of ERGIC-53 or the cargo proteins factor V and factor VIII, and that it could adopt a more ordered conformation upon interaction.

3.1.3 NMR and CD studies of F5F8D causing mutants

Several mutations in MCFD2 have been linked to disease in patients with the bleeding disorder F5F8D. At the time of this study, it had been demonstrated that two of these mutations, resulting in single amino acid substitutions in the second EF-hand motif, abolished interaction between MCFD2 and ERGIC-53 *in vivo* (Zhang *et al.*, 2003). This was suggested to imply Ca^{2+}-binding of MCFD2 to be required for formation of the transporter complex, but no structural information was available to support this theory.

The disease-causing mutations D129E and I136T, as well as the I136V mutation which causes reduced, but not lost, interaction with ERGIC-53 (Zhang *et al.*, 2003), were introduced into MCFD2 by site-directed mutagenesis. ^{15}N-labelled mutated proteins were expressed, purified, and used for circular dichroism (CD) spectroscopy and collection of ^{15}N-HSQC spectra. NMR data revealed that the two disease-causing mutations D129E and I136T prevent correct folding of the protein. ^{15}N-HSQC spectra recorded for these mutants are markedly different than the wild-type spectrum, showing the clustered peaks characteristic of an unfolded protein (Figure 9A and B). Despite being recorded in the presence of Ca^{2+}, these mutant spectra resemble that of the Ca^{2+}-free unfolded form of MCFD2. The I136V mutant, on the other hand, retains its ability to fold; the ^{15}N-HSQC spectrum recorded for this mutant is similar to that of the wild-type protein, but shows changes in the resonances of several residues, particularly Gly134 and Leu88 (Figure 9C). CD spectroscopy analysis confirmed the lower secondary structure content of the disease-causing mutants, again giving spectra similar to that obtained for the unfolded apo-form of MCFD2, while the spectrum of the I136V mutant is closer to that of the wild-type protein (Figure 9D).

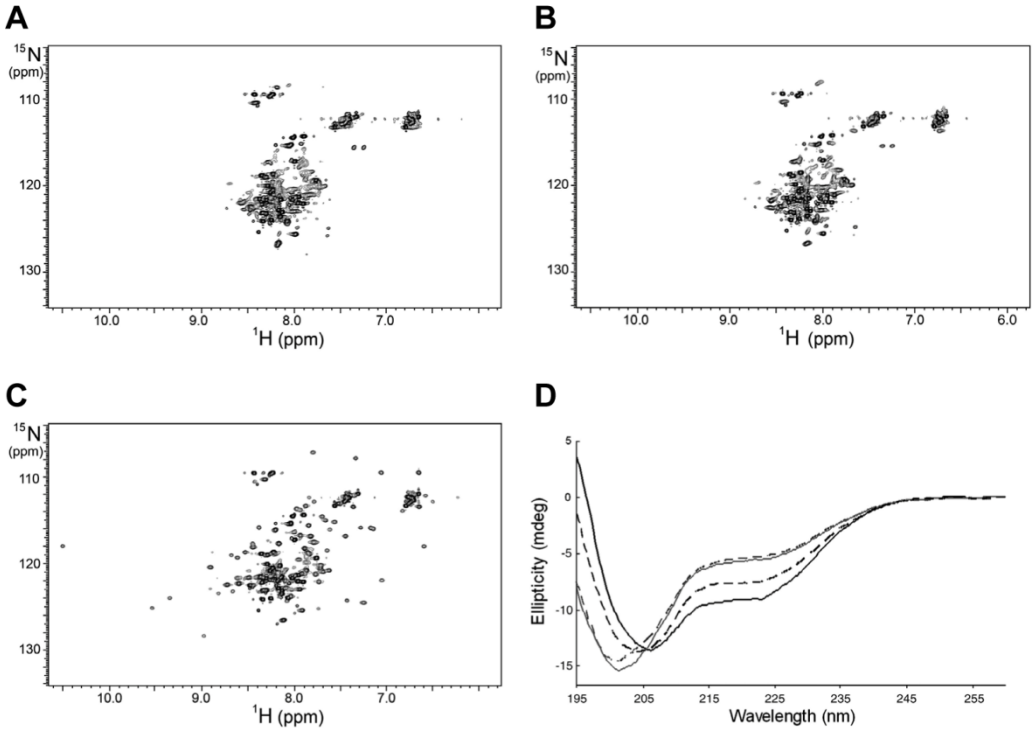


Figure 9. ^{15}N-HSQC spectra of the MCFD2 mutants, recorded in the presence of Ca^{2+}. (**A**) D129E (**B**) I136T (**C**) I136V (**D**) CD spectra of the D129E (broken grey line), I136T (continuous grey line), I136V (broken black line), together with wild-type MCFD2 (continuous black line). All spectra were recorded in the presence of Ca^{2+}.

Both mutated residues are positioned in the second EF-hand motif of MCFD2 (Figure 10). Asp129 is the first residue of the Ca^{2+}-binding consensus sequence of the EF-hand loop. The aspartate at this position is highly conserved; it binds Ca^{2+} and contributes to the structure of the loop by forming several intraloop hydrogen bonds (Gifford *et al.*, 2007). Substitution at this position prevents the correct formation of the EF hand, and disrupts thereby the core structure of the protein. Ile136 is also located to the second EF hand, but it is not involved in binding of Ca^{2+}. There is no complete conservation of the residue at this position, but it is always hydrophobic. In wild-type MCFD2, the side-chain of this residue faces away from the Ca^{2+}-binding site and is packed into the hydrophobic core of the protein. Introducing a polar residue, such as threonine, at this position would most likely disturb the packing of the hydrophobic core, resulting in a disordered protein, whereas the effect of mutation into another hydrophobic residue of similar size would be less dramatic. Our results indicated that mutation of the studied residues causes inefficient coagulation factor transport and F5F8D *via* complete disruption of the structure of MCFD2 resulting in its inability to interact with ERGIC-53. This suggested that the mutations are not necessarily located at the site for interaction with ERGIC-53, a suggestion that was later confirmed by the ERGIC-53/MCFD2 complex structure (Wigren *et al.*, 2010).

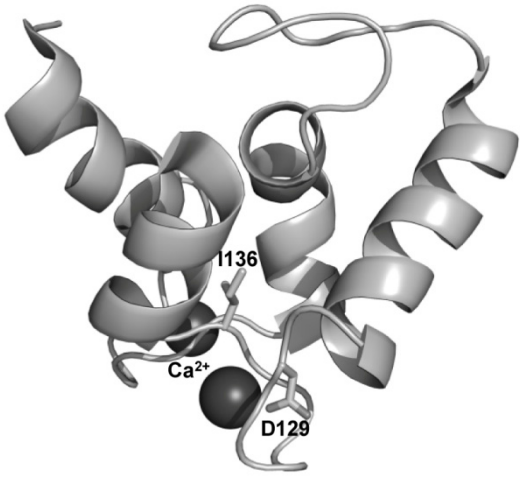


Figure 10. The positions of the mutated residues in the second EF-hand loop of MCFD2. Asp129 is involved in coordination of Ca^{2+} and Ile136 is packed into the hydrophobic core of the protein.

In summary, the first structure of human MCFD2 was determined and a detailed description of the protein was provided. The results demonstrated that the protein is unfolded in the absence of Ca^{2+} ions and that, even in the presence of Ca^{2+}, the N-terminal part of the protein remains disordered. This Ca^{2+}-dependent folding of MCFD2 gave a first insight into the function of Ca^{2+} as an allosteric activator of the protein and its interaction with ERGIC-53. The studies of disease-causing mutations enabled us to explain for the first time how these alterations of the protein cause inefficient coagulation factor transport in patients. Our results emphasise the importance of the intact EF-hand motifs for the structural stability and cargo-transporting function of MCFD2, and suggest that disruption of the ERGIC-53/MCFD2 interaction is the mechanism by which the studied mutants cause F5F8D.

3.2 Structural studies of Erv41p (Papers II and III)

Erv41p and its complex partner Erv46p are highly conserved throughout eukaryotes, implying that they play an important role in the cell. They have been linked to transport and processing of glycoproteins, and the human proteins have recently been implicated in the progression of three different types of cancers, but when the study described in papers II and III was initiated, little was known about the structure of any of these proteins or their precise functions. In order to provide a better understanding of the biological role of these proteins and enable future structure-function experiments and interaction studies, we have determined the structure of yeast Erv41p and initiated a detailed analysis of the properties of the protein.

3.2.1 Protein production, crystallisation and structure determination

Based on the previously reported domain topology and secondary structure prediction, multiple constructs of ERGIC2 (human Erv41) were designed and cloned for expression in *E. coli*. However, despite extensive screening for suitable expression and lysis conditions, only minute amounts of soluble protein could be obtained. In order to increase the chances of achieving expression of soluble protein, we decided to expand our study to include the *Saccharomyces cerevisiae* homologue Erv41p, which, coming from a lower eukaryote, could be expected to be more easily expressed in a prokaryotic host. Although several of the designed constructs of the yeast protein could be overexpressed, solubility was still a major problem. We went on to test different expression systems; expression of the human protein in mammalian cells yielded low amounts of protein, but only when the yeast protein was expressed in insect cells and secreted to the medium were good results obtained. With this setup, we were able to express the longest designed construct, Erv41p_LD, spanning the full lumenal domain of Erv41p (Figure 11), and we could purify protein of sufficient amounts and quality for crystallisation trials, with a final yield of 1-5 mg of protein per litre culture medium.

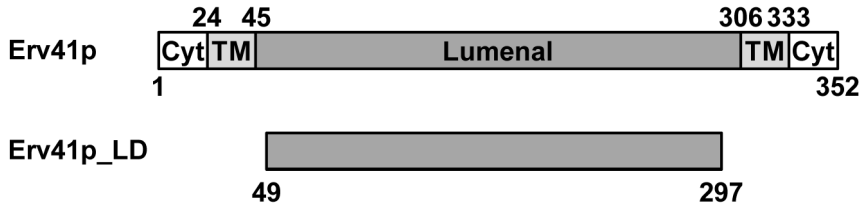


Figure 11. Graphical representation of full-length *S. cerevisiae* Erv41p with the membrane topology indicated. The lumenal domain is shown in dark grey, the transmembrane regions in light grey and the cytosolic tails in white. The construct of the lumenal domain used for crystallisation, Erv41p_LD, is shown below.

Initial crystallisation hits were established using commercial sparse-matrix screens and after optimisation of the conditions, crystals that diffracted to 2.0 Å resolution were obtained. Since no structure of any sufficiently homologous protein was available, we needed to determine phases for structure solution experimentally. A large number of heavy atom compounds were tested for soaking, and both soaking time and

concentration were optimised in order to find a condition giving an anomalous signal without destroying the crystal. The diffraction of the native crystals was already unreliable, and this problem was only amplified by soaking in heavy atoms; a large number of derivative crystals were screened on the beamline before we could collect a useful dataset. Finally, the structure was solved by single-wavelength anomalous diffraction (SAD) phasing, using an Yb^{3+} derivative with one heavy atom site per asymmetric unit. The Yb^{3+}-derivatised crystals were not isomorphous with the native crystals, and the inclusion of isomorphous replacement to improve phasing was therefore not possible. A model of the core of the protein could be built based on the maps from SAD phasing, and was then used for molecular replacement into the higher-resolution native dataset, allowing refinement of the structure to 2.0 Å resolution.

3.2.2 Overall structure of Erv41p

The lumenal part of Erv41p comprises one single domain composed of 14 β-strands and 5 short α-helical segments (Figure 12). The main body of the domain is made up of two twisted β-sheets of mixed character, one with eight strands and one with six, packed against each other to form a β-sandwich with both convex and concave regions.

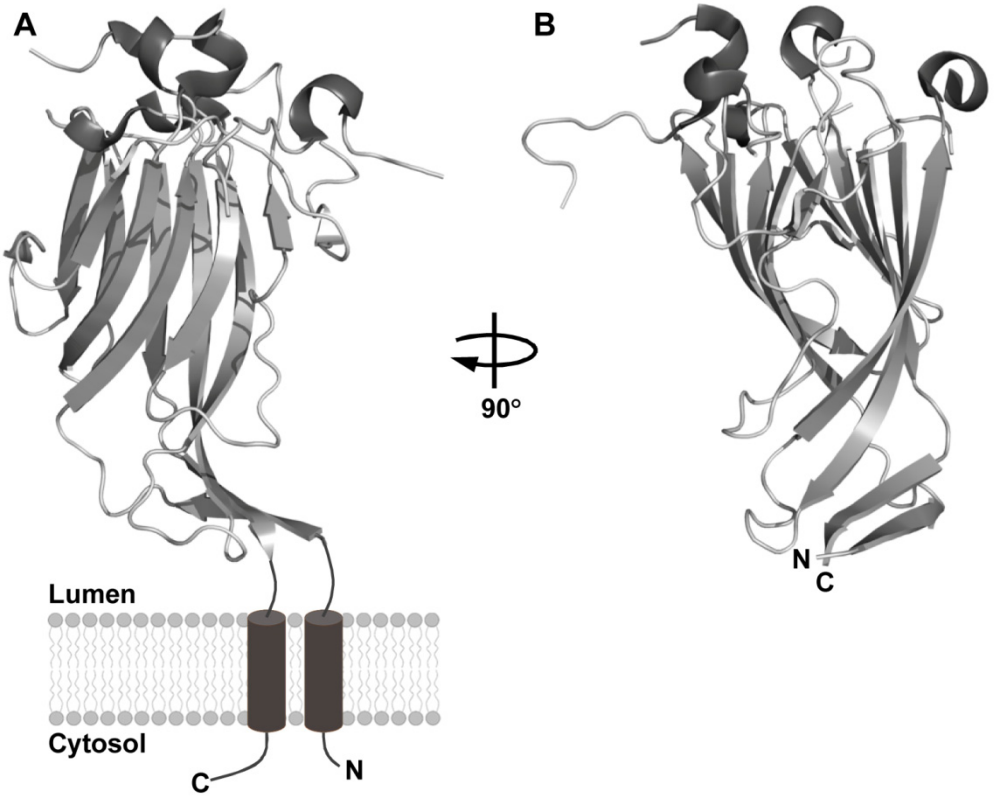


Figure 12. Overall structure of the lumenal domain of Erv41p. Cartoon representation of the structure shaded according to secondary structure, with β-strands in light grey and α-helices in dark grey (PDB ID 3ZLC). (**A**) The crystal structure combined with a schematic to show the position of the lumenal domain in the full protein. (**B**) Side view of the β-sandwich, related to the view in **A** by 90° rotation.

The N- and C-terminal ends of the structure, which would be very close to the membrane in the full-length protein, form two β-strands protruding from the main β-sandwich, which would function to separate it from the membrane. The two termini are positioned close together, suggesting that the two transmembrane helices attached to these ends in the full-length protein would also be situated next to each other in the membrane. On the opposite end of the β-sandwich, the helical segments and several long loops are clustered. This region is likely flexible, considering that several of the loops could not be completely modelled and that B-factors were higher here than in other parts of the structure.

3.2.3 Potential interaction sites

Erv41p is known to be part of a large complex in the cell, and its interactions with other proteins of this complex are most likely of importance to its function. In order to find out more about the character of such interactions, and to find potential interaction sites, we analysed the surface of the protein.

Analysis of the hydrophobic surface of the lumenal domain of Erv41p shows that a number of hydrophobic patches are distributed all over the surface, but none of these stands out as a clear site for protein-protein interactions. In contrast, the electrostatic surface potential reveals a dominant negatively charged area covering one entire face of the β-sandwich (Figure 13A and B). Due to the curved shape of the β-sandwich, this negative patch coincides with a broad shallow groove stretching across the protein. Furthermore, this region harbours a number of residues that are highly conserved across eukaryotic species (Figure 13C and D). Together, these structural features suggest that this could indeed be a region of physiological importance, most likely as a binding site for an interaction partner or ligand.

An additional cluster of conserved residues can be identified in the membrane-proximal region of the lumenal domain. The same region is also identified by the meta-PPISP server (Qin & Zhou, 2007) as the most likely site for protein-protein interactions in the structure (Figure 13E, F, and G). This area is involved in crystal-packing interactions between the two monomers seen in the asymmetric unit of our crystals. Although this interaction is unlikely to be physiological, since both size-exclusion chromatography and dynamic light scattering show the protein to be a monomer, it strengthens the hypothesis that this area is capable of participating in protein-protein interactions.

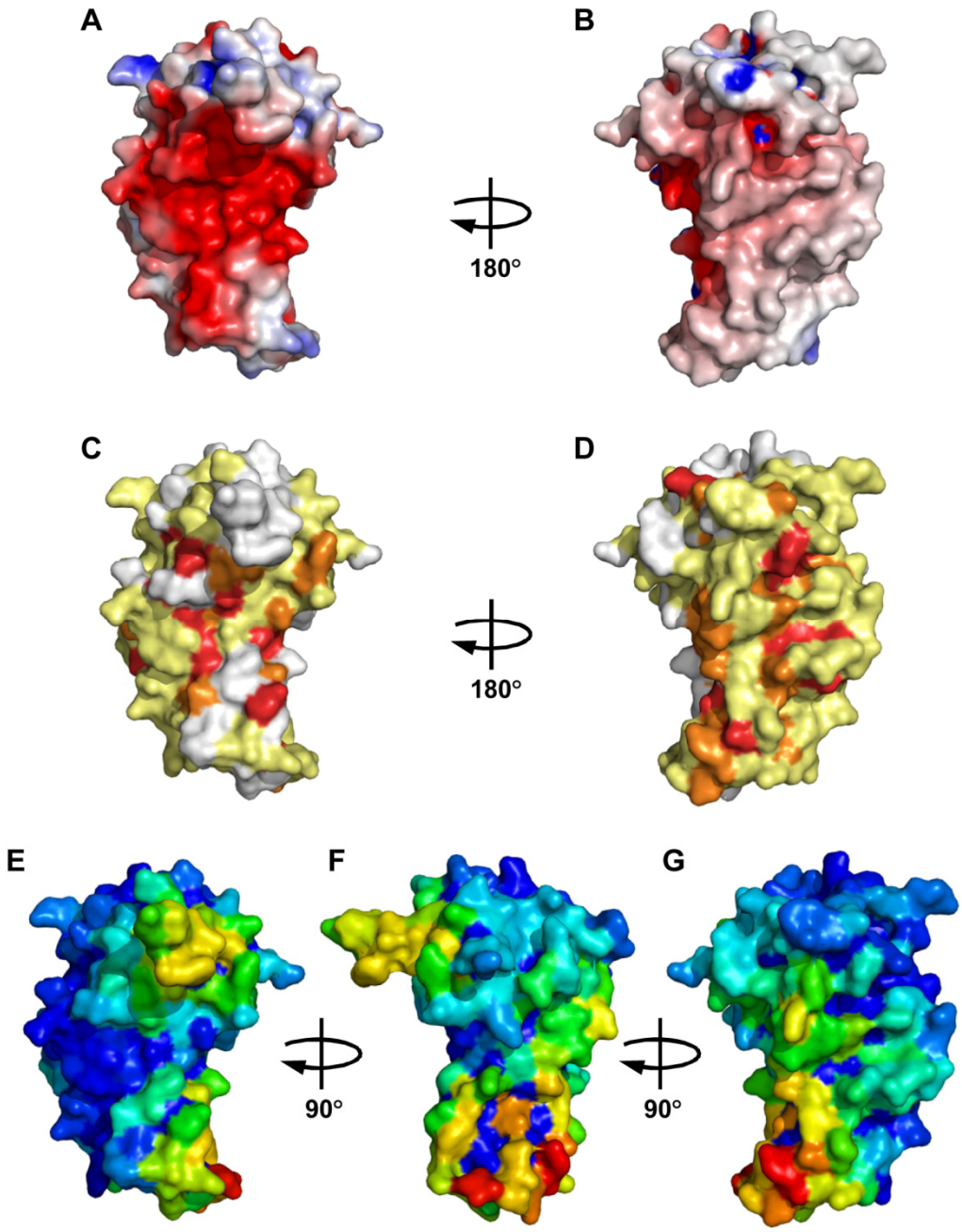


Figure 13. Surface properties of Erv41p. (**A** and **B**) Electrostatic representation of Erv41p showing negative surface potential in red and positive surface potential in blue. The two views are related by 180° rotation around a vertical axis and correspond to the two faces of the β-sandwich. (**C** and **D**) Representation of the evolutionary conservation of Erv41. The sequence of *S. cerevisiae* Erv41p was aligned with homologue sequences from *Homo sapiens*, *Bos taurus, Mus musculus, Danio rerio, and Saccharomyces pombe*. The surface of Erv41p is coloured according to residue conservation among these six organisms, with invariant residues in shown red, residues that are identical in five organisms shown in orange, and residues conserved in four organisms shown in yellow. The two views of the protein are the same as those in **A** and **B**, respectively. (**E**, **F**, and **G**) Predicted protein-protein interaction sites. The surface residues of Erv41p are coloured according to their relative probability of participating in protein-protein interactions, as calculated by the meta-PPISP server, from red indicating the highest probability to blue indicating the lowest probability. The three views are related by sequential 90° rotations, with **E** and **G** corresponding to the views in **A** and **B**, respectively.

3.2.4 Similarity to other structures

Since the role of Erv41p in the cell is not fully understood, we were hoping that structural similarity to proteins of known function could give some hints towards its function. However, the sequence homology of Erv41p to other proteins with known structure is low, and neither are there any close structural homologues. Two servers were used to identify structural homologues in the Protein Data Bank. Only very weak matches were found, and different results were obtained with each server; the Dali server (Holm & Rosenström, 2010) identified acid-sensing ion channels as the top hits, whereas the PDBeFold server (Krissinel & Henrick, 2004) found a carbohydrate-binding module as the closest structural homologue.

The complex between Erv41p and Erv46p has been suggested to play a role in glycoprotein processing and transport; it has been reported to interact with the ER glucosidase II, Rot2p, and strains lacking a cycling Erv41p-Erv46p complex display a mild glycoprotein processing defect (Welsh *et al.*, 2006). In the light of this, the β-sandwich structure of Erv41p is interesting, since a number of β-sandwich proteins have carbohydrate-binding properties. However, the structural homology to these proteins is limited and no conserved carbohydrate-binding motifs can be identified in Erv41p.

In conclusion, we determined the first structure of Erv41p or any of its homologues, and provided a detailed analysis of the structural features of the protein. The protein adopts a β-sandwich structure with only limited structural homology to other proteins. Two potential binding sites for interaction partners or ligands could be identified by analysis of the surface properties of the protein. The structure of Erv41p has provided a new level of knowledge about the protein, and will facilitate further, more detailed investigations into its function.

3.3 Functional studies of Erv41p and Erv46p

With the structure of Erv41p as a starting point, our aim was to further investigate the role of the protein in the cell, and its interactions with binding partners and ligands. In order to do this we have also expressed *S. cerevisiae* Erv46p for structure determination on its own, functional studies, studies of the interaction with Erv41p, and ultimately complex structure determination. With the suggested role of the Erv41p-Erv46p complex in glycoprotein processing in mind, we have investigated whether any of the proteins interacts directly with glycans. Furthermore, we have looked into the possibility that Ca^{2+} could stabilise the Erv41p structure.

3.3.1 Expression and crystallisation of Erv46p

ERGIC3 (human Erv46) was initially cloned for expression in *E. coli.* Several constructs could be overexpressed, but most of them not in soluble form, and only one very short construct could be purified. In analogy with Erv41p, however, the lumenal domain of yeast Erv46p could be expressed in insect cells and secreted to the medium in soluble form, and could be purified to homogeneity. The purified protein, both with and without the secretion signal sequence and His-tag included in the expression construct, has been used for crystallisation screening. An initial crystallisation condition has been established, and the very small crystals have been tested at a synchrotron beamline. No protein diffraction was observed, most likely due to the limited size of the crystals, but the crystals appeared not to be salt and optimisation of this crystallisation condition, as well as continued screening, is ongoing.

3.3.2 Thermofluor studies of the potential role of Erv41p in glycan-binding

The Erv41p-Erv46p complex has been strongly suggested to play a role in glycoprotein processing or transport, and the β-sandwich structure of Erv41p shows some similarity to other glycan-binding proteins in the early secretory pathway, e.g. ERGIC-53. No conserved sugar-binding site could be identified from the structure of Erv41p, but these sites are often difficult to identify since they are shallow and very variable. An obvious starting point for our functional studies was therefore to determine whether Erv41p or Erv46p bind to sugars, and if so, with what specificity.

In order to further investigate the proposed glycoprotein-interacting roles of Erv41p and Erv46p we initially performed Thermofluor ligand screens using a limited number of low-complexity carbohydrates. With Thermofluor, binding of a ligand and concomitant stabilisation of a protein can be detected as a shift in the melting temperature of the protein. A fluorescent dye that binds to hydrophobic patches is added to the protein and the temperature of the sample is gradually increased. As the protein is denatured, the fluorescence signal increases and a melting temperature can be determined. If a ligand that stabilises the protein is added, the melting temperature will increase and a thermal shift can be determined. No clear indication of interaction could be detected in these experiments, perhaps not surprisingly, since the glycan moieties of glycoproteins in a eukaryotic cell are substantially more complex than those tested.

3.3.3 Potential Ca^{2+}-binding of Erv41p

Although no carbohydrate ligands for Erv41p could be detected in our Thermofluor screen, we did make another interesting observation. Screening was performed both in the presence and absence of Ca^{2+}, and addition of Ca^{2+} was observed to result in a thermal shift of approximately 8 °C, suggesting that the ion had a stabilising effect on the protein. In the ER and the Golgi the Ca^{2+} concentration is known to be much higher than in the cytosol, whereas Ca^{2+} levels are below the detection limit in the ERGIC. These variations are believed to be of importance for the functions of the secretory pathway, and many of the proteins localised to these compartments are binding Ca^{2+}. This could be the case also for Erv41p; Ca^{2+} could for example be necessary to form a ligand-binding site.

To further investigate how Ca^{2+} affects the structure of Erv41p, soaking experiments with different concentrations of Ca^{2+} and different soaking times, as well as co-crystallisation using the condition established for the native protein, were performed. In case the packing in the current crystal form would prevent binding of Ca^{2+}, we also tried screening for new crystallisation conditions in presence of Ca^{2+}. Data could be collected to 2-3 Å resolution from both soaked crystals and crystals from co-crystallisation, but no density for Ca^{2+} ions could be observed, and no conformational changes compared to the native protein indicated any Ca^{2+}-induced effects on the structure.

3.3.4 Glycan array screen with Erv41p and Erv46p

As the range of glycans tested in our Thermofluor screens was very limited, a glycan array screen was performed in collaboration with the Consortium for Functional Glycomics with the aim to significantly expand the carbohydrate-space investigated. In this screen (Mammalian Printed Array version 5.1), more than 600 natural and synthetic mammalian glycans printed onto a microarray were tested for binding to the lumenal domains of both Erv41p and Erv46p. An anti-His antibody was used to detect His-tagged protein bound to glycans. Since many carbohydrate-binding β-sandwich proteins are dependent on Ca^{2+} for their function, and we had some evidence that Erv41p might be binding Ca^{2+} (see above), screening was performed both in the presence and absence of Ca^{2+}.

The results of this assay were classed by the Consortium for Functional Glycomics as negative; no carbohydrates were identified as potential ligands for either Erv41p (Figure 14A and B) or Erv46p (Figure 14C and D). The signal was very low for all tested glycans. In the case of Erv46p, some glycans did seem to give rise to a signal higher than the background, and they reproducibly did so both in the presence and absence of Ca^{2+}, but also for these the signals were much lower than could be expected from a true binder. Signals in this range have been classed as positive for other proteins, but then there have been structural correlations between the hit glycans, supporting their binding to the protein. This was not the case for Erv46p; the top glycans were varied (Figure 14E), and there were related structures on the array that did not bind.

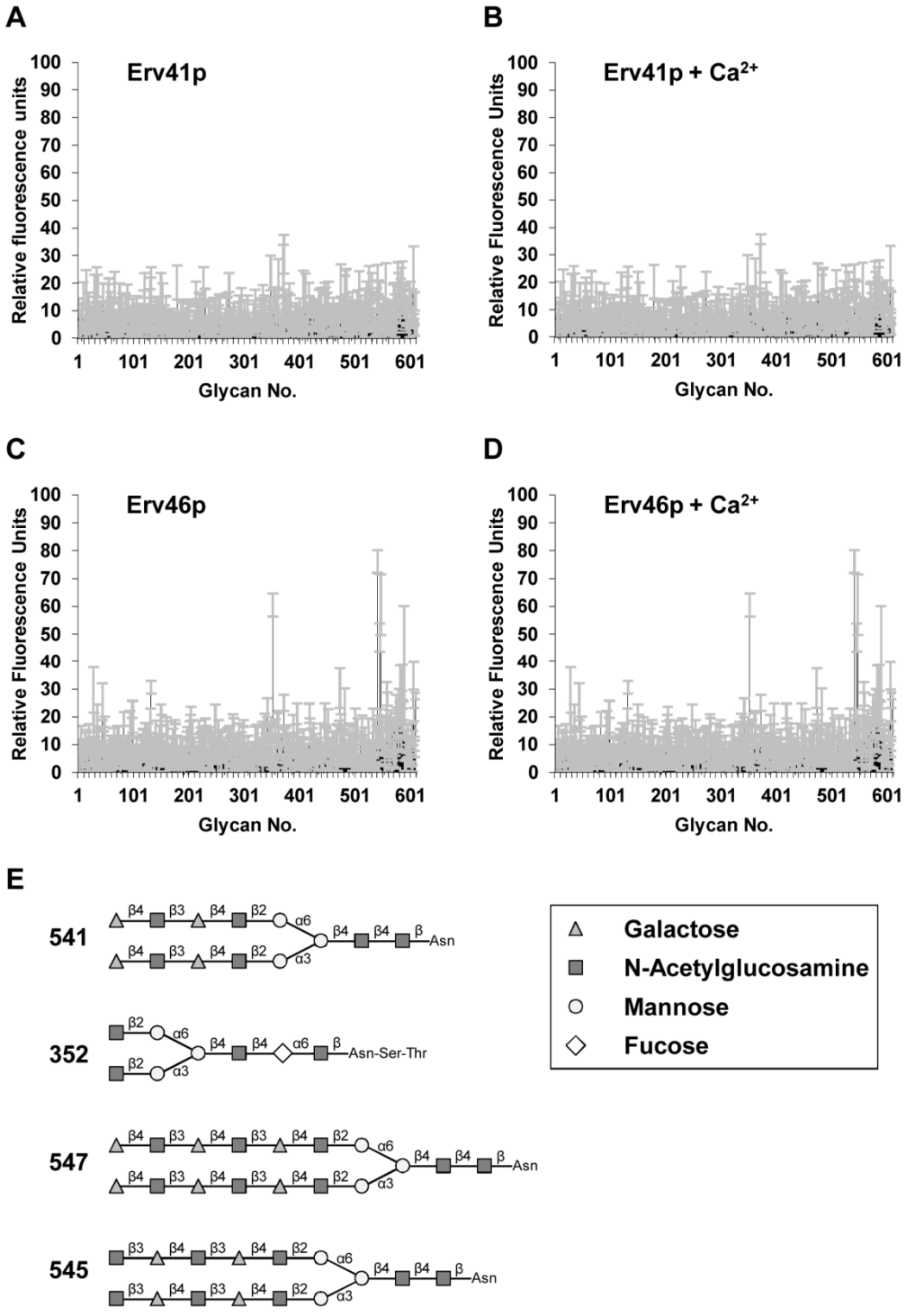


Figure 14. Glycan array screen results for Erv41p and Erv46p. Glycan binding was tested using the Mammalian Printed Array version 5.1 according to the standard procedure of the Consortium for Functional Glycomics. Error bars represent standard error of the mean from four measurements. (**A**) Erv41p analysed in the absence of Ca^{2+}. (**B**) Erv41p analysed in the presence of Ca^{2+}. (**C**) Erv46p analysed in the absence of Ca^{2+}. (**D**) Erv46p analysed in the presence of Ca^{2+}. (**E**) The top hits from the Erv46p screens.

Despite the suggested roles of these proteins in glycoprotein processing or transport it is of course possible that they do not in fact interact with carbohydrates. However, even if the proteins indeed bind glycan moieties in the cell, there could be several reasons for the negative results obtained in this *in vitro* assay. Negative results could be due to low binding affinity or inactive protein. In addition, technical problems, such as problems of the anti-His antibody to detect our proteins, could be a reason for the absence of binding. One potential problem with this assay was that our yeast proteins, which were the only homologues that we could express and purify in sufficient amounts, were screened against mammalian glycans, since no equivalent yeast array was available. Already the human glycome is not fully covered by the screen, and is possible that screening against a set of yeast glycans would give a different result. However, many of the samples for which positive hits have been found from this screen have been non-mammalian. Furthermore, the Erv41p-Erv46p complex has been suggested to interact with glycoproteins, and not necessarily isolated sugars, and it is possible that a combination of a glycan moiety and a protein determinant is necessary for interaction. In addition, it is conceivable that only the complex of Erv41p and Erv46p, and not the isolated proteins, binds to glycans, and further screening of carbohydrate-binding to the complex would therefore be interesting.

In summary, Thermofluor measurements suggest that Ca^{2+} ions stabilise Erv41p. No obvious binding site for Ca^{2+} can be identified based on our structure of Erv41p, but Ca^{2+}-binding sites often include a combination of acidic side-chains and main-chain atoms (Kretsinger, 1976), and can be difficult to identify or predict from sequences or metal-free structures. If Ca^{2+} would indeed be binding to the protein, it would not be surprising, considering that many other glycoprotein-interacting β-sandwich proteins of the secretory pathway depend on Ca^{2+} for their function. Based on the structural similarity of Erv41p to such carbohydrate-binding proteins, we performed a screen for binding to a large number of glycans, but none of the carbohydrates tested showed binding to either Erv41p or Erv46p.

3.4 Structural studies of human P58^{IPK} (Paper IV)

P58^{IPK} has been shown to play multiple roles in the cell. It is involved in influenza infection *via* its inhibition of PKR, it inhibits PERK at the later stages of the unfolded protein response and it interacts with BiP, supposedly as a co-chaperone delivering unfolded proteins. Although these interactions have been mapped to different regions of P58^{IPK}, the lack of structural knowledge has made it difficult to fully understand the mechanisms behind them. The aim of the work described in paper IV was to determine the structure of human P58^{IPK} in order to gain further insight into its functions in the cell, and to provide a basis for studies of the interactions with its binding partners.

3.4.1 Protein production, crystallisation and structure determination

Multiple constructs of human P58^{IPK} were cloned for expression in *E. coli*. The longest soluble construct comprised the mature protein up to the end of the J domain, without the N-terminal signal sequence and the 43 C-terminal residues, which were predicted to be unstructured. The expression construct included an N-terminal His-tag, and protein both with and without this tag was used for screening for crystallisation conditions. Extensive screening revealed only one condition yielding crystals, and this only when protein with the intact His-tag was used. Despite optimisation efforts and screening of a large number of crystals at the synchrotron, no resolution higher than 4-5 Å could be obtained from any of these native crystals.

With the hope that a homologous protein would give better crystals, P58^{IPK} from different organisms (*Danio rerio, Drosophila melanogaster,* and *Caenorhabditis elegans*) were cloned, expressed and purified and used for crystallisation screening. However, none of the constructs produced crystals in any of the crystallisation screens. Some spherulites were obtained, but before these were optimised we were able to solve the structure of the human protein, and the work on the homologues was thus not pursued.

At the time, no structure of homology high enough to be used as a model for molecular replacement was available. Therefore, selenomethionine-substituted human P58^{IPK} was produced for phasing purposes. This derivatised protein turned out to yield significantly better crystals than those previously obtained from the native protein, and data was collected at the Se-edge to 3.2 Å resolution in space group P312. At approximately the same time, a partial structure of mouse P58^{IPK} was published by another group (Tao *et al.*, 2010), allowing us to use a combination of molecular replacement and single-wavelength anomalous diffraction to solve the structure. Using the optimised condition identified for the selenomethionine-containing crystals, we finally found that native protein, without the His-tag, yielded crystals that were similar, but that diffracted in space group H32. One of these crystals finally allowed us to collect a dataset used to refine the structure to 3.0 Å. The structure is very similar in the two crystal forms, with an RMSD of 0.8 Å.

3.4.2 Overall structure of P58^{IPK}

P58^{IPK} adopts an elongated shape; the protein is approximately 120 Å long but only 20-50 Å wide (Figure 15). The N-terminal TPR domain consists of 19 α-helices arranged into nine TPR motifs. These are further divided into three groups of three TPR motifs each. Tandem arrays of TPR motifs often adopt a super-helix structure of antiparallel α-helices (Zeytuni & Zarivach, 2012), but in P58^{IPK} two longer helices create the three subdomains, each with one concave and one convex surface.

The first secondary structure element of the J domain is a turn of a helix (helix I) separated from the TPR domain by three amino acid residues adopting an extended conformation. The short helical segment is followed by a long loop, and together they act as a linker between the TPR domain and the remainder of the J domain, adding flexibility to the protein. The main body of the J domain is composed of two antiparallel α-helices (helix II and helix III) separated by a short loop region harbouring the conserved histidine-proline-aspartate (HPD) motif. Another turn of a helix (helix IV) could be built, but the last residues of the crystallised construct were not visible in the electron density, indicating that the C-terminus of P58^{IPK} is unstructured.

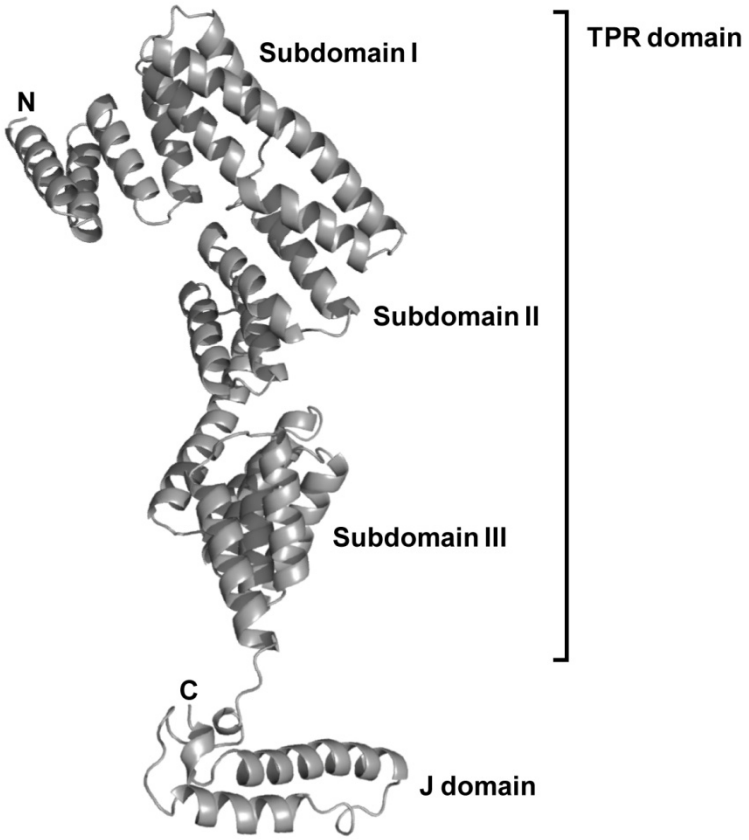


Figure 15. Overall structure of human P58^{IPK}. (PDB ID 2Y4T).

3.4.3 Comparison to the TPR domain of mouse P58^{IPK}

The structure of the TPR domain of mouse P58^{IPK} (PDB ID 3IEG) (Tao *et al.*, 2010) was determined shortly before we could solve the structure of the full-length human protein. Despite a high sequence identity between the two homologues, the structures of the TPR domains differ (Figure 16). The long helices connecting subdomain I and II, and II and III, are bent and rotated in the human structure. This results in a more curved shape of the human protein, with the subdomains closer together. Whether this structural difference is a true physiological feature is not easy to say. Considering the high sequence identity, it is perhaps less likely that the structure of the two proteins are indeed different. The observed differences could instead be explained by the absence of the J domain in the mouse structure; it is possible that the conformation of the TPR domain is altered when a J domain is attached. Yet another explanation for the differences could be that they arise from differences in crystal packing. The structural differences between the two TPR domains do however tell us that there is an inherent flexibility in the structure, a flexibility that is likely important for its function in binding different interaction partners.

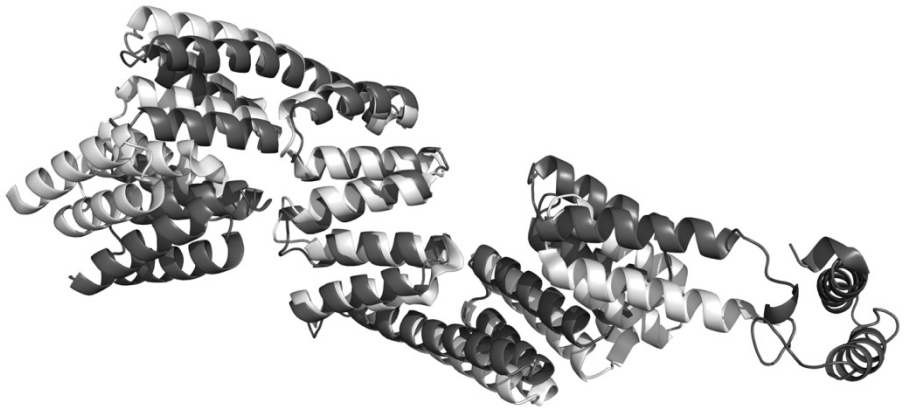


Figure 16. Comparison of human and mouse P58^{IPK} TPR domains. Superposition of the human P58^{IPK} structure (dark grey) and the structure of the TPR domain of mouse P58^{IPK} (light grey). The structural alignment was based on TPR subdomain II.

3.4.4 Binding sites for interaction partners

The HPD motif in the J domain is conserved among Hsp40 proteins and is commonly shown to be involved in interactions with Hsp70 chaperones, as has previously been shown to be the case with P58^{IPK} and BiP (Petrova *et al.*, 2008). In human P58^{IPK}, as in other J domain proteins, the HPD motif is positioned in the loop between the two long helices II and III of the J domain (Figure 17A). The motif is positioned at the edge of the elongated protein, readily accessible for interaction with BiP. For other Hsp40-Hsp70 pairs, it has been suggested that, in addition to the HPD motif, helix II, and to a lesser extent helix III and the loop between these two helices, are important for interaction (Greene *et al.*, 1998; Genevaux *et al.*, 2002), and especially positively charged residues in this region (Jiang *et al.*, 2003). In human P58^{IPK}, there is indeed a patch of positive charge on the surface of helix II (Figure 17B), but it is located on the

farther end from the HPD motif and is smaller than in most other J domain structures. Instead, there is a higher than usual proportion of hydrophobic residues immediately surrounding the HPD motif (Figure 17C).

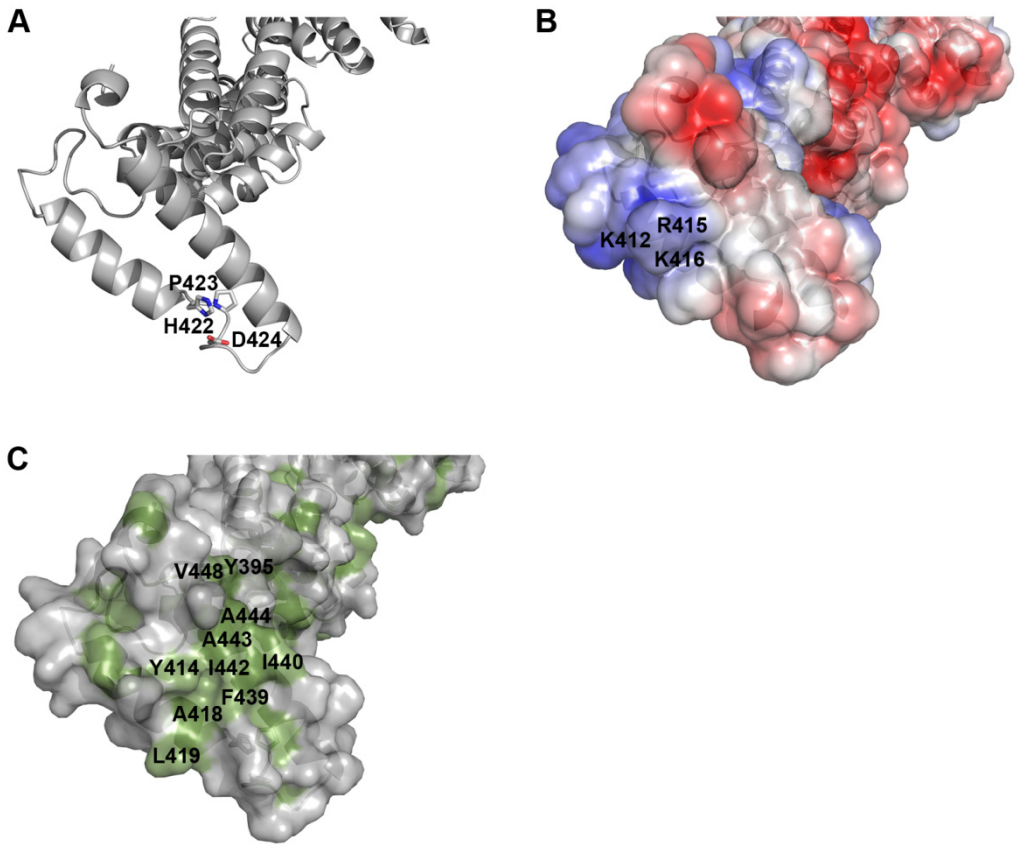


Figure 17. The J domain of P58^{IPK}. (**A**) Cartoon representation of the J domain of P58^{IPK} with the residues of the HPD motif shown as sticks. (**B**) Surface representation of the J domain, showing the electrostatic surface potential with positively charged residues in blue and negatively charged residues in red. Positively charged residues on helix II are labelled. (**C**) Surface representation of the J domain, with hydrophobic residues in green. Residues contributing to the hydrophobic patch adjacent to the HPD motif are labelled.

The residues surrounding the HPD motif do not coincide with those that are the most conserved among the BiP-interacting ERdj proteins. In contrast, the residues in this region, as well as the entire J domain, exhibit a higher degree of conservation between different P58^{IPK} homologues from various organisms. This could suggest differences in the interactions of the ERdj proteins with BiP, and give an insight into how the different family members are able to regulate the functions of BiP. Although the structure of J domains is generally well conserved, it is possible that properties such as the proportions of charged and hydrophobic surface could play a role in the regulation of which ERdj protein binds to BiP at a certain time point. Moreover, the significantly higher degree of J domain conservation among P58^{IPK} homologues from different species would indicate that the properties of the J domain have a large impact on the functions or interactions of the protein. However, although the ERdj proteins all share the J domain, there is a high degree of variation between the additional domains, and

their functions and effects on BiP are also distinct. Therefore it is perhaps not unlikely that the additional domains in each ERdj protein influence the interaction and that the standard J domain-mediated binding of BiP is only part of the picture. In the case of $P58^{IPK}$, this is very easy to imagine; it would be surprising if a large and elongated protein like this would interact with BiP only *via* the very tip of the protein.

The exact binding sites for the other interaction partners of $P58^{IPK}$ are not known, and are difficult to predict based on the structure. The interaction with PKR has been mapped to TPR6 (Gale *et al.*, 1996; Tang *et al.*, 1996), but the exact residues involved remain to be determined. Analysis of the surface of TPR6 in human $P58^{IPK}$ does not show any obvious binding site. $P52^{rIPK}$ has been shown to bind to $P58^{IPK}$ at TPR7 (Gale *et al.*, 2002). It has been suggested to bind *via* a charged domain, and the charged surface of TPR7 makes it plausible that this interaction is of an electrostatic nature. In the structure of the TPR domain of mouse $P58^{IPK}$, a hydrophobic patch in the first subdomain was suggested to bind unfolded proteins. A patch at the same position and with the same shape is present also in human $P58^{IPK}$, consistent with an important role for binding of unfolded proteins.

In summary, we determined the structure of human $P58^{IPK}$ and analysed its surface with respect to potential binding sites for different interaction partners. The interaction site for BiP is particularly interesting, since it could provide some initial insights into how the different BiP-ERdj interactions could be regulated. The structure reveals that $P58^{IPK}$ is a flexible protein; most likely the conformation of the TPR domain can vary depending on what it is binding, and the linker between the TPR domain and the J domain certainly would allow adaptation of the relationship between the two domains.

3.5 Studies of the interaction between P58^{IPK} and BiP (Paper V)

Based on the structure of P58^{IPK}, we could identify patches on the J domain of potential importance for interaction with BiP. The residues of these patches are not those that are most conserved among the ERdj proteins, and understanding how they are involved in interaction could explain how the ERdj proteins are able to regulate the different functions of BiP. In order to investigate these regions in more detail and to determine exactly which residues had a direct effect on BiP-binding, we created a series of mutants replacing residues in the surroundings of the HPD motif. Our aim was then to apply surface plasmon resonance (SPR) biosensor technology in order to quantify the interaction between native P58^{IPK} and BiP, and compare the results to the interactions of each mutant. However, the interaction turned out to be far more complex than we had expected, and constructs we had not initially expected to interact with BiP in fact did. Due to the complexity of the SPR data, interpretation of our results was performed in collaboration with Professor Helena Danielson at the Department of Chemistry at Uppsala University, a specialist within the field of biomolecular interactions and the surface plasmon resonance technique. Instead of determining which parts of the J domain, in addition to the HPD motif, interacted with BiP, we analysed the interaction between the two proteins on a domain level, in order to try to better understand this complex interaction and to determine whether it includes not only the well-established J domain-component, but also the TPR domain.

Since the interaction between P58^{IPK} and BiP proved to be so complex, we also tried to employ other methods to help us understand it better. In particular, an enzyme assay measuring the ATPase activity of BiP in the presence of the P58^{IPK} constructs used in the SPR measurements allowed us to determine whether the interactions we saw also had an effect on the ATPase activity. In order to fully understand the interaction between P58^{IPK} and BiP, an ultimate aim would be to determine the structure of the complex of the two proteins. We have made attempts to crystallise the complex of P58^{IPK} and BiP, but no crystals have been obtained, and co-crystallisation of Hsp70-Hsp40 pairs is known to be notoriously difficult. In addition, we have tried to obtain a low-resolution structure of the complex using small-angle X-ray scattering (SAXS), but this was not successful.

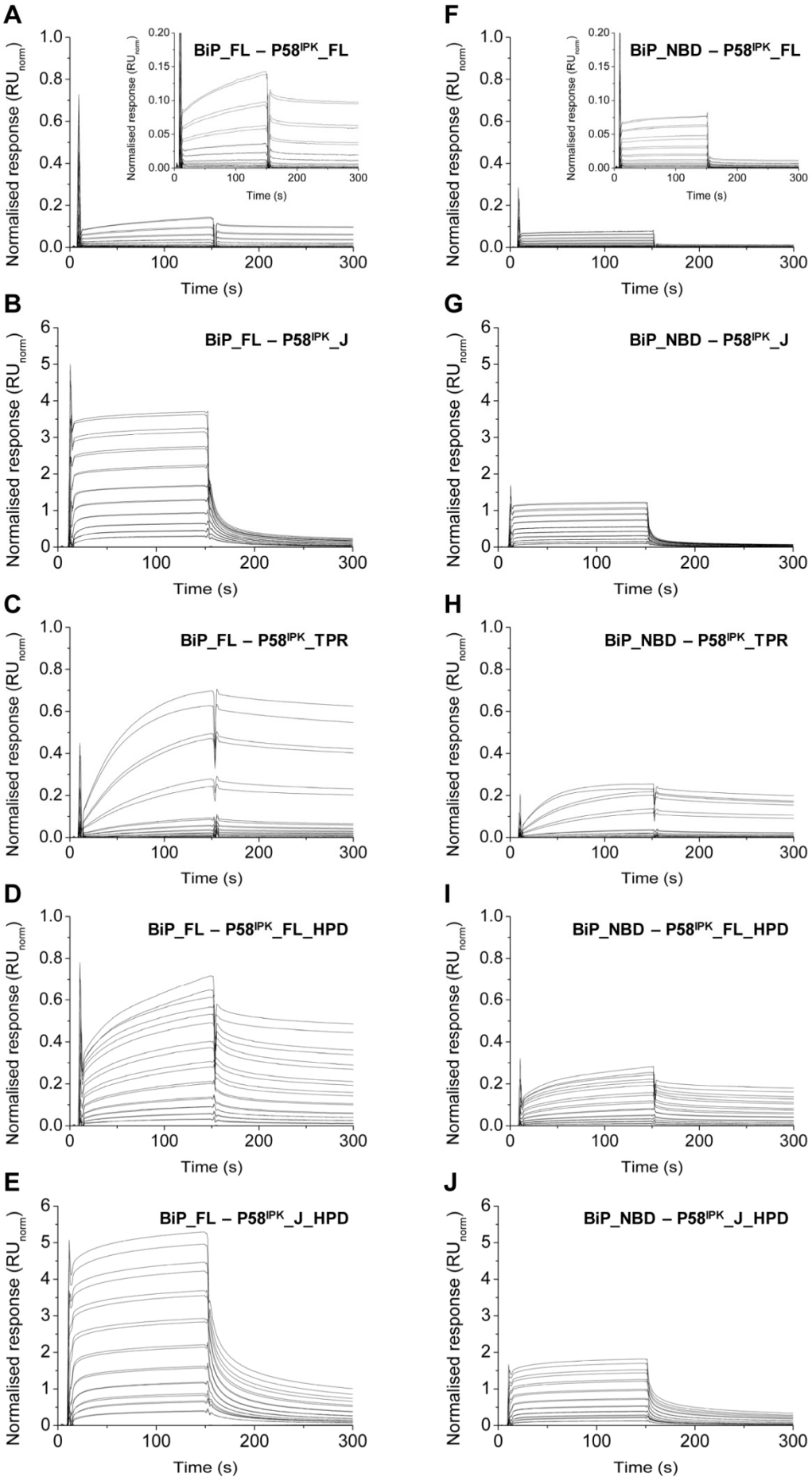
3.5.1 Surface plasmon resonance assays for studies of the P58^{IPK}-BiP interaction

In all SPR experiments, two different constructs of BiP were used as ligands; one construct spanning both the nucleotide-binding domain and the substrate-binding domain, and one comprising only the nucleotide-binding domain and the linker connecting the two domains. Different P58^{IPK} constructs were used as analytes; the full-length construct as used for solving the structure, and constructs of the separate TPR and J domains, as well as HPD-to-AAA mutated variants of the full-length and the J domain constructs. Based on previously published results, stating that the isolated TPR domain did not co-immunoprecipitate BiP, our intention was to employ this construct as a negative control.

We started out by measuring the interaction between the different constructs of native and HPD-mutated P58^{IPK} and BiP. Interaction was clearly detected in all cases, but the sensorgrams were not well modelled by a simple 1:1 single-step interaction model, suggesting that the interactions were more complex. The shapes of the sensorgrams suggested that there could be several components to the interactions, and an induced-fit model gave reasonable curve fits, although not sufficient to explain the mode of interaction or to calculate the kinetic parameters. None of the alternative models explored gave better results or allowed us to determine the nature of the complexity, and we therefore performed a qualitative analysis of the sensorgrams (Figure 18).

All tested constructs of P58^{IPK} interacted with BiP, although the interactions differed qualitatively. The full-length construct and the isolated J domain both interacted, in agreement with previous results. More surprisingly, also the TPR domain of P58^{IPK} interacted with both BiP constructs. However, the sensorgrams indicated that the interaction was of a different type than for constructs comprising the J domain. The interaction of the TPR domain with BiP was also different depending on whether full-length BiP or the NBD alone was used. This suggested that there is an interaction between P58^{IPK} and the SBD of BiP, and that the TPR domain, when connected to the J domain, could mediate this. Interactions outside the J domain have been suggested also for other ERdj-BiP pairs (Marcinowski *et al.*, 2011), and it is plausible that the interaction with BiP is not mediated by the relatively small J domain alone, but that also the larger TPR domain binds to BiP. However, TPR motifs are known to mediate a wide range of protein-protein interactions, and it is possible that the TPR domain of P58^{IPK} binds to BiP in a more unspecific way than the J domain. Based on our SPR data we cannot determine if this is the case, and although several of our normalised curves suggested a superstoichiometric interaction, equilibrium analysis showed that even these data are not consistent with a purely unspecific interaction.

Figure 18. (Opposite page) Sensorgrams of P58^{IPK} interacting with BiP. Full-length BiP (BiP_FL, residues 26-639) or the nucleotide-binding domain alone (BiP_NBD, residues 26-419) were immobilised on the chip. Five different constructs of P58^{IPK} were used as analytes; full-length P58^{IPK} (P58^{IPK}_FL, residues 35-461), the J domain (P58^{IPK}_J, residues 394-466), the TPR domain (P58^{IPK}_TPR, residues 35-401), the full-length protein with the HPD motif mutated (P58^{IPK}_FL_HPD, residues 35-461, HPD422,423,424AAA), and the J domain with the HPD motif mutated (P58^{IPK}_J_HPD, residues 394-466, HPD422,423,424AAA). The raw data were double referenced and normalised with respect to the immobilisation levels and the different molecular weights of the immobilised protein and the analyte. (**A**) BiP_FL and P58^{IPK}_FL. (**B**) BiP_FL and P58^{IPK}_J. (**C**) BiP_FL and P58^{IPK}_TPR. (**D**) BiP_FL and P58^{IPK}_FL_HPD. (**E**) BiP_FL and P58^{IPK}_J_HPD. (**F**) BiP_NBD and P58^{IPK}_FL. (**G**) BiP_NBD and P58^{IPK}_J. (**H**) BiP_NBD and P58^{IPK}_TPR. (**I**) BiP_NBD and P58^{IPK}_FL_HPD. (**J**) BiP_NBD and P58^{IPK}_J_HPD. For panels **A** and **F**, for which the signal is weaker, a second copy of the sensorgrams with larger scale is added to the right of the original to facilitate interpretation.


A
BiP_FL – P58IPK_FL
B
BiP_FL – P58IPK_J
C
BiP_FL – P58IPK_TPR
D
BiP_FL – P58IPK_FL_HPD
E
BiP_FL – P58IPK_J_HPD
F
BiP_NBD – P58IPK_FL
G
BiP_NBD – P58IPK_J
H
BiP_NBD – P58IPK_TPR
I
BiP_NBD – P58IPK_FL_HPD
J
BiP_NBD – P58IPK_J_HPD
Normalised response (RUnorm)
Time (s)

3.5.2 Studies of P58^{IPK}'s stimulation of BiP's ATPase activity

An important function of J domain proteins is to stimulate the ATPase activity of Hsp70 chaperones. With the aim of further investigating the functional interaction between P58^{IPK} and BiP, and to support the interpretation of our SPR results, we measured the ATPase activity of BiP using a malachite green based colorimetric assay detecting the release of free phosphate. Figure 19 shows the results from measurements of the ATPase activity of BiP alone and in the presence of different constructs of P58^{IPK}. Addition of full-length P58^{IPK} led to an approximately eightfold stimulation of the ATPase activity of BiP. Also the isolated J domain increased the ATPase rate, but not to the same extent as the full-length protein. Interestingly, the full-length protein with a mutated HPD motif retained some capability to stimulate the ATPase activity, whereas the mutation in the J domain alone completely abolished the stimulation. No clear stimulating effect was observed upon addition of the TPR domain.

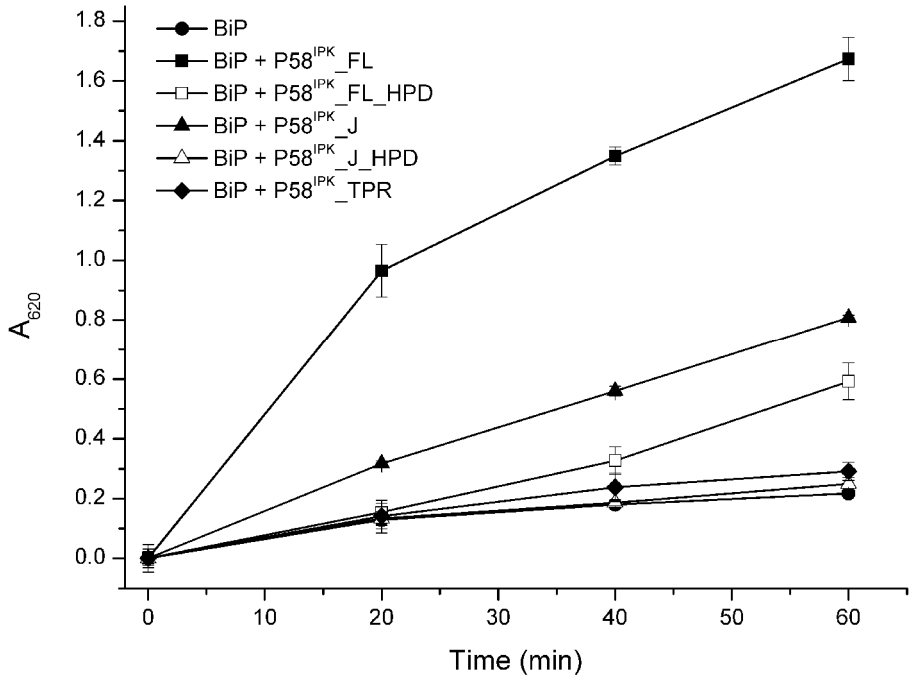


Figure 19. ATPase activity of BiP. BiP was incubated alone or with different constructs of P58^{IPK}. Reactions were started by addition of ATP, and the release of free phosphate was detected at the indicated time points by measurement of absorbance at 620 nm using a malachite green based colorimetric assay. The signal from ATP in buffer without proteins was subtracted to compensate for intrinsic hydrolysis. Error bars show standard deviation from three independent measurements.

3.5.3 Functional interactions between P58^{IPK} and BiP

Combining the results from our SPR measurements and the ATPase activity assay, we see that the TPR domain of P58^{IPK} is clearly involved in the BiP-stimulating effect, although only when attached to a J domain. On the other hand, the isolated TPR domain interacts with BiP, but the interaction seems to be different than that of the full-length protein. Also between full-length P58^{IPK} and the isolated J domain there is a

difference in interaction with BiP. These results all suggest that the interaction between the two proteins is dynamic, with the presence of one domain of $P58^{IPK}$ affecting the binding and function of the other, in particular the presence of the TPR domain enhancing the ATPase stimulating effect of the J domain, either by a second interaction or by allosteric effects.

The HPD motif of $P58^{IPK}$ is clearly involved in the interaction with BiP, but it is certainly not the only site of interaction between the two large proteins; the TPR domain and the remainder of the J domain are also important for binding. The fact that the full-length mutated protein retains some capability to stimulate BiP's ATPase activity confirms these results.

In conclusion, we have studied the interaction of $P58^{IPK}$ with BiP and its ability to stimulate ATPase activity. Our finding that not only the conserved J domain, but also the TPR domain, interacts with BiP and participates in the stimulation of ATPase activity has taken us one step further on the way to understanding the role of $P58^{IPK}$ as a co-chaperone and how the different functions of BiP are regulated by the ERdj proteins. To fully understand the interaction between BiP and $P58^{IPK}$ would require detailed structural information on the complex between the two proteins, but attempts to achieve that using X-ray crystallography and SAXS have proven difficult. In order to expand the knowledge on this chaperone-co-chaperone pair, further interaction studies could instead be undertaken. For example, repeating the performed experiments using a non-hydrolysable ATP analogue might help to separate the initial binding from a subsequent conformational change.

4 CONCLUSIONS

Within this thesis project, members of three complexes with different functions in the early secretory pathway have been studied. NMR spectroscopy and X-ray crystallography have been applied to determine the three-dimensional structures of MCFD2, Erv41p and $P58^{IPK}$, and additional methods, such as CD spectroscopy and SPR biosensor technology, have been used to further characterise each protein.

NMR studies of human MCFD2 resulted in the first structure of MCFD2 and demonstrated that the protein is unstructured in the absence of Ca^{2+} but adopts a predominantly ordered conformation upon binding of Ca^{2+}. These findings gave an insight into the function of Ca^{2+} as an allosteric regulator of the interaction between MCFD2 and ERGIC-53. Characterisation of mutant variants of MCFD2 allowed us to explain the mechanisms by which they cause F5F8D, and also further emphasise the importance of the intact EF-hand motifs for the structural stability of MCFD2.

X-ray crystallography was used to determine the first structure of Erv41p or any of its homologues. Detailed analysis of the surface properties of the protein enabled us to identify two potential binding sites for ligands or interaction partners. Although the structural homology between Erv41p and other proteins is limited, the β-sandwich structure in combination with the suggested role of Erv41p in glycoprotein processing is interesting, and provided a starting point for our further investigations to elucidate its function in the cell. Our results from thermal shift assays suggest that Ca^{2+} stabilises Erv41p. Considering the proposed role of Erv41p in glycoprotein processing or transport, this would not be surprising, since binding of Ca^{2+} is required for the function of other glycoprotein-interacting β-sandwich proteins in the secretory pathway, and this could be further investigated. Future studies of Erv41p should also include its complex partner Erv46p. We have made initial attempts to crystallise this protein alone, and although the structure of Erv46p would be interesting on its own, it is possible that we will only fully understand the roles of these proteins when we know how they interact, and a main goal would be to determine the three-dimensional structure of the complex between the two proteins.

The structure of human $P58^{IPK}$ was determined by X-ray crystallography. Analysis of the surface of the J domain provided some initial insights into how interaction with BiP could be mediated, and how the functions of BiP could be regulated by the different members of the ERdj family. These findings led us to further investigate the interaction between $P58^{IPK}$ and BiP using SPR biosensor technology. The interaction between the two proteins turned out to be more complex than we initially expected, with both domains of $P58^{IPK}$ being involved in the interaction. These results were confirmed by measurements of the ability of different domains of $P58^{IPK}$ to stimulate the ATPase activity of BiP. We showed that, while the J domain alone is able to increase BiP's ATPase activity, the presence of the TPR domain in the full-length protein significantly improves the stimulation. Although our findings are interesting, they do not provide the full story of the interaction between the two proteins. In order to find additional pieces

of the puzzle, many future experiments could be suggested, with the determination of the complex structure as a challenging ultimate aim.

In conclusion, the structures determined within this thesis work have given new insights into different aspects of the early secretory pathway. Our studies of their interactions with complex partners or ligands have provided increased knowledge of their roles in the cell, and future studies will hopefully yield a full understanding of the functions of each protein and how these functions are facilitated by their interactions.

5 ACKNOWLEDGEMENTS

Many people have contributed to making my years as a PhD student a great experience and now I would like to express my gratitude.

First of all I would like to thank my main supervisor Jodie Guy for all the help during the years. Thank you for giving me freedom and support, for your patience and eye for detail. I think this has very much been a journey we made together.

I would also like to thank Professors Ylva Lindqvist and Gunter Schneider for leading the group with an inspiring dedication to science, for sharing your experience and knowledge, and for creating a friendly atmosphere. Thank you Ylva for first accepting me as a PhD student and for your continued support as a co-supervisor throughout my time in the group.

Jean-Marie Bourhis, my other co-supervisor, thank you for suggesting a challenge like $P58^{IPK}$ as a target for my studies, for the practical help during my first year in the lab, and for the long-distance encouragement during the remaining time. I have never forgotten about the dessert task.

I would like to acknowledge the people we have collaborated with. Helena Danielson, thank you for convincing us about the usefulness of our Biacore data. Anne Tuukkanen, thank you for help with SAXS measurements and data analysis, and thank you for the guided tour in Hamburg.

I also want to thank the students that have contributed to my project. So Yeon, thank you for your enthusiasm and for the interesting food choices. Isabel, without you I would never have tried to speak about chromatography in Spanish. Lance, you were so productive that I still have some of your samples left.

Thanks to all the past and present members of the Molecular Structural Biology group: The PhD students who finished long before me, Catrine, Stina, Daniel and Hanna. To see you reach the goal inspired me to carry on with my project. The post-docs that left a long time ago, Laura, Mikko and Leif, and more recently, Jola, Agatha and Atsushi. A special thanks to my former office mates. Cyprian for your calm sense of humour and Jason for your ability to always cheer me up and our shared fascination with growing birches and oaks. Rajesh, thank you for showing me how to run the Biacore machine and for being an excellent host in Lisbon. Katja, for sharing the burden of the ER proteins and for the long lunches out. Edvard, thank you for introducing me to the lab work, for provision of bike tools, and for always being friendly and helpful, not least during my thesis writing. Magnus, thank you for your good mood, for being such a sensible person, and for shared teaching experiences. Katharina, I am so happy that you joined the group. Thank you for listening to me when I needed it the most, and for letting me listen to you. And thanks for the company on trips in all weathers. Dominik, thanks for thoughtful discussions about life and science and for good advice concerning insect cells. Sorry for hiding my proteins in the freezer, soon you can inherit them all. Ming Wei, for interesting discussions about weight lifting and nutrition. I hope you will

soon be back in the gym again. Dominic, I wish we had got the opportunity to sing and play in the same concert. Eva Maria, good luck with the skiing, I would be really tempted to go for another Vasalopp. Ömer, thanks for new perspectives on Sweden and the Swedes, for your imaginative associations, and for teaching me the importance of salad dressing. Robert, for your long and fascinating stories and for the way you always seem happy to share your knowledge. Bernie, for always being ready to give helpful scientific advice and for not hesitating to take the Swedish last piece of cake. Doreen, thanks for quietly reorganising the seminar list when I baked semlor instead of preparing my talk.

Thanks Victoria for your quick responses to all practical questions, and thank you Eva for your help and care during my early years in the group. Ahmad, thank you for running the lab and for company at coffee time.

I would also like to thank the teaching unit at MBB: Åke Rökaeus, Susanne Rothstein, Susie Björkholm, Agneta Nordberg and Christine Jansson. You have helped making teaching fun and interesting, and I think that I have often learnt just as much as the students. And all the other people I have met at MBB, thanks for practical help and nice company.

Tack till alla vänner utanför arbetet. Tack för att ni har lyssnat på mina historier om vedermödorna i labbet, och för att ni har fått mig att tänka på helt andra saker.

Slutligen vill jag tacka min familj. Tack Henrik för squashmatcher, våffelluncher och diskussioner om doktorandlivet. Tack Lovisa för strapatser i kajak och på skidor. Mamma och Pappa, tack för er omtanke, er hjälpsamhet och ert oändliga stöd i allt jag tar mig för. Ni betyder så mycket.

6 REFERENCES

Abdallah, H.E., Gouider, E., Amor, M.B., Jlizi, A., Meddeb, B., Elgaaied, A., 2010. Molecular analysis in two Tunisian families with combined factor V and factor VIII deficiency. Haemophilia 16, 801-804.

Adachi, Y., Yamamoto, K., Okada, T., Yoshida, H., Harada, A., Mori, K., 2008. ATF6 Is a Transcription Factor Specializing in the Regulation of Quality Control Proteins in the Endoplasmic Reticulum. Cell Structure and Function 33, 75-89.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P., 2002. Molecular Biology of the Cell, 4th ed. Garland Science, New York.

Antonin, W., Meyer, H.A., Hartmann, E., 2000. Interactions between Spc2p and other components of the endoplasmic reticulum translocation sites of the yeast Saccharomyces cerevisiae. Journal of Biological Chemistry 275, 34068-34072.

Appenzeller, C., Andersson, H., Kappeler, F., Hauri, H.P., 1999. The lectin ERGIC-53 is a cargo transport receptor for glycoproteins. Nature Cell Biology 1, 330-334.

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, 2173-2183.

Appenzeller-Herzog, C., Nyfeler, B., Burkhard, P., Santamaria, I., Lopez-Otin, C., Hauri, H.P., 2005. Carbohydrate- and conformation-dependent cargo capture for ER-exit. Molecular Biology of the Cell 16, 1258-1267.

Ashby, M.C., Tepikin, A.V., 2001. ER calcium and the functions of intracellular organelles. Seminars in Cell & Developmental Biology 12, 11-17.

Awad, W., Estrada, I., Shen, Y., Hendershot, L.M., 2008. BiP mutants that are unable to interact with endoplasmic reticulum DnaJ proteins provide insights into interdomain interactions in BiP. Proceedings of the National Academy of Sciences of the United States of America 105, 1164-1169.

Barber, G.N., Thompson, S., Lee, T.G., Strom, T., Jagus, R., Darveau, A., Katze, M.G., 1994. The 58-kilodalton inhibitor of the interferon-induced double-stranded RNA-activated protein-kinase is a tetratricopeptide repeat protein with oncogenic properties. Proceedings of the National Academy of Sciences of the United States of America 91, 4278-4282.

Barlowe, C., 2003. Signals for COPII-dependent export from the ER: what's the ticket out? Trends in Cell Biology 13, 295-300.

Barlowe, C., Schekman, R., 1993. Sec12 encodes a guanine-nucleotide-exchange factor essential for transport vesicle budding from the ER. Nature 365, 347-349.

Barlowe, C., Orci, L., Yeung, T., Hosobuchi, M., Hamamoto, S., Salama, N., Rexach, M.F., Ravazzola, M., Amherdt, M., Schekman, R., 1994. COPII - a membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic-reticulum. Cell 77, 895-907.

Barlowe, C.K., Miller, E.A., 2013. Secretory Protein Biogenesis and Traffic in the Early Secretory Pathway. Genetics 193, 383-410.

Bertolotti, A., Zhang, Y.H., Hendershot, L.M., Harding, H.P., Ron, D., 2000. Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response. Nature Cell Biology 2, 326-332.

Blatch, G.L., Lassle, M., 1999. The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. Bioessays 21, 932-939.

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. Journal of Cell Biology 67, 835-851.

Breuza, L., Halbeisen, R., Jeno, P., Otte, S., Barlowe, C., Hong, W.J., Hauri, H.P., 2004. Proteomics of endoplasmic reticulum-golgi intermediate compartment (ERGIC) membranes from brefeldin A-treated HepG2 cells identifies ERGIC-32, a new cycling protein that interacts with human Erv46. Journal of Biological Chemistry 279, 47242-47253.

Brocker, C., Engelbrecht-Vandre, S., Ungermann, C., 2010. Multisubunit Tethering Complexes and Their Role in Membrane Fusion. Current Biology 20, R943-R952.

Brodsky, J.L., 2012. Cleaning Up: ER-Associated Degradation to the Rescue. Cell 151, 1163-1167.

Budnik, A., Stephens, D.J., 2009. ER exit sites - Localization and control of COPII vesicle formation. Febs Letters 583, 3796-3803.

Bulleid, N.J., Ellgaard, L., 2011. Multiple ways to make disulfides. Trends in Biochemical Sciences 36, 485-492.

Cai, H.Q., Reinisch, K., Ferro-Novick, S., 2007. Coats, tethers, Rabs, and SNAREs work together to mediate the intracellular destination of a transport vesicle. Developmental Cell 12, 671-682.

Cao, S.S., Kaufman, R.J., 2012. Unfolded protein response. Current Biology 22, R622-R626.

Carrara, M., Prischi, F., Ali, M.M.U., 2013. UPR Signal Activation by Luminal Sensor Domains. International Journal of Molecular Sciences 14, 6454-6466.

Cheetham, M.E., Caplan, A.J., 1998. Structure, function and evolution of DnaJ: conservation and adaptation of chaperone function. Cell Stress & Chaperones 3, 28-36.

Cox, J.S., Walter, P., 1996. A novel mechanism for regulating activity of a transcription factor that controls the unfolded protein response. Cell 87, 391-404.

Csala, M., Margittai, E., Banhegyi, G., 2010. Redox Control of Endoplasmic Reticulum Function. Antioxidants & Redox Signaling 13, 77-108.

Cunnea, P.M., Miranda-Vizuete, A., Bertoli, G., Simmen, T., Damdimopoulos, A.E., Hermann, S., Leinonen, S., Huikko, M.P., Gustafsson, J.A., Sitia, R., Spyrou, G., 2003. ERdJ5, an endoplasmic reticulum (ER)-resident protein containing DnaJ and thioredoxin domains, is expressed in secretory cells or following ER stress. Journal of Biological Chemistry 278, 1059-1066.

Cunningham, M.A., Pipe, S.W., Zhang, B., Hauri, H.P., Ginsburg, D., Kaufman, R.J., 2003. LMAN1 is a molecular chaperone for the secretion of coagulation factor VIII. Journal of Thrombosis and Haemostasis 1, 2360-2367.

Cyr, D.M., Lu, X.Y., Douglas, M.G., 1992. Regulation of Hsp70 Function by a Eukaryotic DnaJ Homolog. Journal of Biological Chemistry 267, 20927-20931.

Cyr, D.M., Langer, T., Douglas, M.G., 1994. DnaJ-like proteins: molecular chaperones and specific regulators of Hsp70. Trends in Biochemical Sciences 19, 176-181.

Dancourt, J., Barlowe, C., 2010. Protein Sorting Receptors in the Early Secretory Pathway. Annual Review of Biochemistry 79, 777-802.

Dong, C.M., Nichols, C.D., Guo, J.H., Huang, W., Lambert, N.A., Wu, G.Y., 2012. A Triple Arg Motif Mediates a2B-Adrenergic Receptor Interaction with Sec24C/D and Export. Traffic 13, 857-868.

Dong, M., Bridges, J.P., Apsley, K., Xu, Y., Weaver, T.E., 2008. ERdj4 and ERdj5 are required for endoplasmic reticulum-associated protein degradation of misfolded surfactant protein C. Molecular Biology of the Cell 19, 2620-2630.

Dudek, J., Greiner, M., Muller, A., Hendershot, L.M., Kopsch, K., Nastainczyk, W., Zimmermann, R., 2005. ERj1p has a basic role in protein biogenesis at the endoplasmic reticulum. Nature Structural & Molecular Biology 12, 1008-1014.

Ellgaard, L., Helenius, A., 2003. Quality control in the endoplasmic reticulum. Nature Reviews Molecular Cell Biology 4, 181-191.

Elmahmoudi, H., Wigren, E., Laatiri, A., Jlizi, A., Elgaaied, A., Gouider, E., Lindqvist, Y., 2011. Analysis of newly detected mutations in the MCFD2 gene giving rise to combined deficiency of coagulation factors V and VIII. Haemophilia 17, E923-E927.

Farquhar, M.G., Palade, G.E., 1998. The Golgi apparatus: 100 years of progress and controversy. Trends in Cell Biology 8, 2-10.

Gale, M., Tan, S.L., Wambach, M., Katze, M.G., 1996. Interaction of the interferon-induced PKR protein kinase with inhibitory proteins p58(IPK) and vaccinia virus K3L is mediated by unique domains: Implications for kinase regulation. Molecular and Cellular Biology 16, 4172-4181.

Gale, M., Blakely, C.M., Darveau, A., Romano, P.R., Korth, M.J., Katze, M.G., 2002. P52(rIPK) regulates the molecular cochaperone P58(IPK) to mediate control of the RNA-dependent protein kinase in response to cytoplasmic stress. Biochemistry 41, 11878-11887.

Gale, M., Blakely, C.M., Hopkins, D.A., Melville, M.W., Wambach, M., Romano, P.R., Katze, M.G., 1998. Regulation of interferon-induced protein kinase PKR: Modulation of p58(IPK) inhibitory function by a novel protein, P52(rIPK). Molecular and Cellular Biology 18, 859-871.

Gardner, B.M., Pincus, D., Gotthardt, K., Gallagher, C.M., Walter, P., 2013. Endoplasmic Reticulum Stress Sensing in the Unfolded Protein Response. Cold Spring Harbor Perspectives in Biology 5, 15.

Gauss, R., Kanehara, K., Carvalho, P., Ng, D.T.W., Aebi, M., 2011. A Complex of Pdi1p and the Mannosidase Htm1p Initiates Clearance of Unfolded Glycoproteins from the Endoplasmic Reticulum. Molecular Cell 42, 782-793.

Genevaux, P., Schwager, F., Georgopoulos, C., Kelley, W.L., 2002. Scanning mutagenesis identifies amino acid residues essential for the in vivo activity of the Escherichia coli DnaJ (Hsp40) J-domain. Genetics 162, 1045-1053.

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

Gething, M.J., McCammon, K., Sambrook, J., 1986. Expression of wild-type and mutant forms of influenza hemagglutinin - the role of folding in intracellular-transport. Cell 46, 939-950.

Gifford, J.L., Walsh, M.P., Vogel, H.J., 2007. Structures and metal-ion-binding properties of the Ca2+-binding helix-loop-helix EF-hand motifs. Biochemical Journal 405, 199-221.

Gilchrist, A., Au, C.E., Hiding, J., Bell, A.W., Fernandez-Rodriguez, J., Lesimple, S., Nagaya, H., Roy, L., Gosline, S.J.C., Hallett, M., Paiement, J., Kearney, R.E., Nilsson, T., Bergeron, J.J.M., 2006. Quantitative proteomics analysis of the secretory pathway. Cell 127, 1265-1281.

Greene, M.K., Maskos, K., Landry, S.J., 1998. Role of the J-domain in the cooperation of Hsp40 with Hsp70. Proceedings of the National Academy of Sciences of the United States of America 95, 6108-6113.

Gulow, K., Bienert, D., Haas, I.G., 2002. BiP is feed-back regulated by control of protein translation efficiency. Journal of Cell Science 115, 2443-2452.

Haas, I.G., Wabl, M., 1983. Immunoglobulin heavy-chain binding-protein. Nature 306, 387-389.

Hamman, B.D., Hendershot, L.M., Johnson, A.E., 1998. BiP maintains the permeability barrier of the ER membrane by sealing the lumenal end of the translocon pore before and early in translocation. Cell 92, 747-758.

Hanisch, F.A., 2001. O-glycosylation of the mucin type. Biological Chemistry 382, 143-149.

Harding, H.P., Zhang, Y.H., Ron, D., 1999. Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. Nature 397, 271-274.

Harding, H.P., Novoa, I., Zhang, Y.H., Zeng, H.Q., Wek, R., Schapira, M., Ron, D., 2000. Regulated translation initiation controls stress-induced gene expression in mammalian cells. Molecular Cell 6, 1099-1108.

Harty, C., Strahl, S., Romisch, K., 2001. O-mannosylation protects mutant alpha-factor precursor from endoplasmic reticulum-associated degradation. Molecular Biology of the Cell 12, 1093-1101.

Hauri, H.P., Kappeler, F., Andersson, H., Appenzeller, C., 2000. ERGIC-53 and traffic in the secretory pathway. Journal of Cell Science 113, 587-596.

Haze, K., Yoshida, H., Yanagi, H., Yura, T., Mori, K., 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, 3787-3799.

Helenius, A., Aebi, M., 2004. Roles of N-linked glycans in the endoplasmic reticulum. Annual Review of Biochemistry 73, 1019-1049.

Hirayama, H., Fujita, M., Yoko-O, T., Jigami, Y., 2008. O-mannosylation is required for degradation of the endoplasmic reticulum-associated degradation substrate Gas1*p via the ubiquitin/proteasome pathway in Saccharomyces cerevisiae. Journal of Biochemistry 143, 555-567.

Hohfeld, J., Jentsch, S., 1997. GrpE-like regulation of the Hsc70 chaperone by the anti-apoptotic protein BAG-1. EMBO Journal 16, 6209-6216.

Hollien, J., Weissman, J.S., 2006. Decay of endoplasmic reticulum-localized mRNAs during the unfolded protein response. Science 313, 104-107.

Holm, L., Rosenström, P., 2010. Dali server: conservation mapping in 3D. Nucleic Acids Research 38, W545-W549.

Hosoda, A., Kimata, Y., Tsuru, A., Kohno, K., 2003. JPDI, a novel endoplasmic reticulum-resident protein containing both a BiP-interacting J-domain and thioredoxin-like motifs. Journal of Biological Chemistry 278, 2669-2676.

Itin, C., Roche, A.C., Monsigny, M., Hauri, H.P., 1996. ERGIC-53 is a functional mannose-selective and calcium-dependent human homologue of leguminous lectins. Molecular Biology of the Cell 7, 483-493.

Ivaskevicius, V., Windyga, J., Baran, B., Bykowska, K., Daugela, L., Watzka, M., Seifried, E., Olenburg, J., 2008. The first case of combined coagulation factor V and coagulation factor VIII deficiency in Poland due to a novel p.Tyr135Asn missense mutation in the MCFD2 gene. Blood Coagulation & Fibrinolysis 19, 531-534.

Jackson, M.R., Nilsson, T., Peterson, P.A., 1990. Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic-reticulum. EMBO Journal 9, 3153-3162.

Jayandharan, G., Spreafico, M., Viswabandya, A., Chandy, M., Srivastava, A., Peyvandi, F., 2007. Mutations in the MCFD2 gene are predominant among patients with hereditary combined FV and FVIII deficiency (F5F8D) in India. Haemophilia 13, 413-419.

Jiang, J., Taylor, A.B., Prasad, K., Ishikawa-Brush, Y., Hart, P.J., Lafer, E.M., Sousa, R., 2003. Structure-function analysis of the auxilin J-domain reveals an extended Hsc70 interaction interface. Biochemistry 42, 5748-5753.

Kampinga, H.H., Craig, E.A., 2010. The HSP70 chaperone machinery: J proteins as drivers of functional specificity. Nature Reviews Molecular Cell Biology 11, 579-592.

Kappeler, F., Klopfenstein, D.R.C., Foguet, M., Paccaud, J.P., Hauri, H.P., 1997. The recycling of ERGIC-53 in the early secretory pathway - ERGIC-53 carries a cytosolic endoplasmic reticulum exit determinant interacting with COPII. Journal of Biological Chemistry 272, 31801-31808.

Kawasaki, N., Ichikawa, Y., Matsuo, I., Totani, K., Matsumoto, N., Ito, Y., Yamamoto, K., 2008. The sugar-binding ability of ERGIC-53 is enhanced by its interaction with MCFD2. Blood 111, 1972-1979.

Kelleher, D.J., Gilmore, R., 2006. An evolving view of the eukaryotic oligosaccharyltransferase. Glycobiology 16, 47R-62R.

Kimata, Y., Kimata, Y.L., Shimizu, Y., Abe, H., Farcasanu, R.C., Takeuchi, M., Rose, M.D., Kohno, K., 2003. Genetic evidence for a role of BiP/Kar2 that regulates Ire1 in response to accumulation of unfolded proteins. Molecular Biology of the Cell 14, 2559-2569.

Kimura, T., Hosoda, Y., Sato, Y., Kitamura, Y., Ikeda, T., Horibe, T., Kikuchi, M., 2005. Interactions among yeast protein-disulfide isomerase proteins and endoplasmic reticulum chaperone proteins influence their activities. Journal of Biological Chemistry 280, 31438-31441.

Kornfeld, R., Kornfeld, S., 1985. Assembly of asparagine-linked oligosaccharides. Annual Review of Biochemistry 54, 631-664.

Kozutsumi, Y., Segal, M., Normington, K., Gething, M.J., Sambrook, J., 1988. The presence of malfolded proteins in the endoplasmic-reticulum signals the induction of glucose-regulated proteins. Nature 332, 462-464.

Kreis, T.E., Lodish, H.F., 1986. Oligomerization is essential for transport of vesicular stomatitis viral glycoprotein to the cell-surface. Cell 46, 929-937.

Kretsinger, R.H., 1976. Calcium-binding proteins. Annual Review of Biochemistry 45, 239-266.

Krissinel, E., Henrick, K., 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallographica Section D - Biological Crystallography 60, 2256-2268.

Ladiges, W.C., Knoblaugh, S.E., Morton, J.F., Korth, M.J., Sopher, B.L., Baskin, C.R., MacAuley, A., Goodman, A.G., LeBoeuf, R.C., Katze, M.G., 2005. Pancreatic beta-cell failure and diabetes in mice with a deletion mutation of the endoplasmic reticulum molecular chaperone gene P58(IPK). Diabetes 54, 1074-1081.

Lahtinen, U., Svensson, K., Pettersson, R.F., 1999. Mapping of structural determinants for the oligomerization of p58, a lectin-like protein of the intermediate compartment and cis-Golgi. European Journal of Biochemistry 260, 392-397.

Lee, M.C.S., Miller, E.A., 2007. Molecular mechanisms of COPII vesicle formation. Seminars in Cell & Developmental Biology 18, 424-434.

Lee, M.C.S., Orci, L., Hamamoto, S., Futai, E., Ravazzola, M., Schekman, R., 2005. Sar1p N-terminal helix initiates membrane curvature and completes the fission of a COPII vesicle. Cell 122, 605-617.

Lee, T.G., Tomita, J., Hovanessian, A.G., Katze, M.G., 1990. Purification and partial characterization of a cellular inhibitor of the interferon-induced protein kinase of Mr 68,000 from influenza virus-infected cells. Proceedings of the National Academy of Sciences of the United States of America 87, 6208-6212.

Lee, Y.K., Brewer, J.W., Hellman, R., Hendershot, L.M., 1999. BiP and immunoglobulin light chain cooperate to control the folding of heavy chain and ensure the fidelity of immunoglobulin assembly. Molecular Biology of the Cell 10, 2209-2219.

Lewis, M.J., Pelham, H.R.B., 1990. A human homolog of the yeast HDEL receptor. Nature 348, 162-163.

Lewis, M.J., Sweet, D.J., Pelham, H.R.B., 1990. The ERD2 Gene Determines the Specificity of the Luminal ER Protein Retention System. Cell 61, 1359-1363.

Lewit-Bentley, A., Rety, S., 2000. EF-hand calcium-binding proteins. Current Opinion in Structural Biology 10, 637-643.

Liberek, K., Marszalek, J., Ang, D., Georgopoulos, C., Zylicz, M., 1991. Escherichia coli DnaJ and GrpE heat shock proteins jointly stimulate ATPase activity of DnaK. Proceedings of the National Academy of Sciences of the United States of America 88, 2874-2878.

Ma, D.K., Zerangue, N., Lin, Y.F., Collins, A., Yu, M., Jan, Y.N., Jan, L.Y., 2001. Role of ER export signals in controlling surface potassium channel numbers. Science 291, 316-319.

Ma, K., Vattem, K.M., Wek, R.C., 2002. Dimerization and release of molecular chaperone inhibition facilitate activation of eukaryotic initiation factor-2 kinase in response to endoplasmic reticulum stress. Journal of Biological Chemistry 277, 18728-18735.

Malkus, P., Jiang, F., Schekman, R., 2002. Concentrative sorting of secretory cargo proteins into COPII-coated vesicles. Journal of Cell Biology 159, 915-921.

Marcinowski, M., Holler, M., Feige, M.J., Baerend, D., Lamb, D.C., Buchner, J., 2011. Substrate discrimination of the chaperone BiP by autonomous and cochaperone-regulated conformational transitions. Nature Structural & Molecular Biology 18, 150-158.

Mayer, M.P., 2013. Hsp70 chaperone dynamics and molecular mechanism. Trends in Biochemical Sciences 38, 507-514.

Mayer, M.P., Bukau, B., 2005. Hsp70 chaperones: Cellular functions and molecular mechanism. Cellular and Molecular Life Sciences 62, 670-684.

Mazzarello, P., Calligaro, A., Vannini, V., Muscatello, U., 2003. Timeline - The sarcoplasmic reticulum: its discovery and rediscovery. Nature Reviews Molecular Cell Biology 4, 69-74.

Meimaridou, E., Gooljar, S.B., Chapple, J.P., 2009. From hatching to dispatching: the multiple cellular roles of the Hsp70 molecular chaperone machinery. Journal of Molecular Endocrinology 42, 1-9.

Mellman, I., Simons, K., 1992. The Golgi Complex: In Vitro Veritas? Cell 68, 829-840.

Meyer, H.A., Grau, H., Kraft, R., Kostka, S., Prehn, S., Kalies, K.U., Hartmann, E., 2000. Mammalian Sec61 is associated with Sec62 and Sec63. Journal of Biological Chemistry 275, 14550-14557.

Miller, E., Antonny, B., Hamamoto, S., Schekman, R., 2002. Cargo selection into COPII vesicles is driven by the Sec24p subunit. EMBO Journal 21, 6105-6113.

Morais, V.A., Brito, C., Pijak, D.S., Crystal, A.S., Fortna, R.R., Li, T., Wong, P.C., Doms, R.W., Costa, J., 2006. N-glycosylation of human nicastrin is required for interaction with the lectins from the secretory pathway calnexin and ERGIC-53. Biochimica Et Biophysica Acta-Molecular Basis of Disease 1762, 802-810.

Mori, K., 2009. Signalling Pathways in the Unfolded Protein Response: Development from Yeast to Mammals. Journal of Biochemistry 146, 743-750.

Munro, S., Pelham, H.R.B., 1986. An Hsp70-like Protein in the ER: Identity with the 78 kd Glucose-Regulated Protein and Immunoglobulin Heavy Chain Binding Protein. Cell 46, 291-300.

Nakamura, N., Wei, J.H., Seemann, J., 2012. Modular organization of the mammalian Golgi apparatus. Current Opinion in Cell Biology 24, 467-474.

Neve, E.P.A., Lahtinen, U., Pettersson, R.F., 2005. Oligomerization and interacellular localization of the glycoprotein receptor ERGIC-53 is independent of disulfide bonds. Journal of Molecular Biology 354, 556-568.

Nichols, W.C., Seligsohn, U., Zivelin, A., Terry, V.H., Hertel, C.E., Wheatley, M.A., Moussalli, M.J., Hauri, H.P., Ciavarella, N., Kaufman, R.J., Ginsburg, D., 1998. Mutations in the ER-Golgi intermediate compartment protein ERGIC-53 cause combined deficiency of coagulation factors V and VIII. Cell 93, 61-70.

Nilsson, T., Pypaert, M., Hoe, M.H., Slusarewicz, P., Berger, E.G., Warren, G., 1993. Overlapping Distribution of Two Glycosyltransferases in the Golgi Apparatus of HeLa Cells. Journal of Cell Biology 120, 5-13.

Nishikawa, S., Fewell, S.W., Kato, Y., Brodsky, J.L., Endo, T., 2001. Molecular chaperones in the yeast endoplasmic reticulum maintain the solubility of proteins for retrotranslocation and degradation. Journal of Cell Biology 153, 1061-1069.

Nishimura, N., Balch, W.E., 1997. A di-acidic signal required for selective export from the endoplasmic reticulum. Science 277, 556-558.

Nufer, O., Kappeler, F., Guldbrandsen, S., Hauri, H.P., 2003. ER export of ERGIC-53 is controlled by cooperation of targeting determinants in all three of its domains. Journal of Cell Science 116, 4429-4440.

Nyfeler, B., Zhang, B., Ginsburg, D., Kaufman, R.J., Hauri, H.P., 2006. Cargo selectivity of the ERGIC-53/MCFD2 transport receptor complex. Traffic 7, 1473-1481.

Oeri, J., Matter, M., Isenschmid, H., Hauser, F., Koller, F. 1954. Congenital factor V deficiency (parahemophilia) with true hemophilia in two brothers, pp. 575-588, Vol. 58, Bibliotheca Paediatrica.

Oka, O.B.V., Pringle, M.A., Schopp, I.M., Braakman, I., Bulleid, N.J., 2013. ERdj5 Is the ER Reductase that Catalyzes the Removal of Non-Native Disulfides and Correct Folding of the LDL Receptor. Molecular Cell 50, 793-804.

Okamura, K., Kimata, Y., Higashio, H., Tsuru, A., Kohno, K., 2000. Dissociation of Kar2p/BiP from an ER sensory molecule, Ire1p, triggers the unfolded protein response in yeast. Biochemical and Biophysical Research Communications 279, 445-450.

Orci, L., Ravazzola, M., Mack, G.J., Barlowe, C., Otte, S., 2003. Mammalian Erv46 localizes to the endoplasmic reticulum-Golgi intermediate compartment and to cis-Golgi cisternae. Proceedings of the National Academy of Sciences of the United States of America 100, 4586-4591.

Otero, J.H., Lizák, B., Hendershot, L.M., 2010. Life and death of a BiP substrate. Seminars in Cell & Developmental Biology 21, 472-478.

Otsu, W., Kurooka, T., Otsuka, Y., Sato, K., Inaba, M., 2013. A New Class of Endoplasmic Reticulum Export Signal Phi X Phi X Phi for Transmembrane Proteins and Its Selective Interaction with Sec24C. Journal of Biological Chemistry 288, 18521-18532.

Otte, S., Barlowe, C., 2002. The Erv41p-Erv46p complex: multiple export signals are required in trans for COPII-dependent transport from the ER. EMBO Journal 21, 6095-6104.

Otte, S., Belden, W.J., Heidtman, M., Liu, J., Jensen, O.N., Barlowe, C., 2001. Erv41p and Erv46p: New components of COPII vesicles involved in transport between the ER and Golgi complex. Journal of Cell Biology 152, 503-517.

Paetzel, M., Karla, A., Strynadka, N.C.J., Dalbey, R.E., 2002. Signal peptidases. Chemical Reviews 102, 4549-4579.

Palade, G.E., 1956. The endoplasmic reticulum. Journal of Biophysical and Biochemical Cytology 2, 85-98.

Paroutis, P., Touret, N., Grinstein, S., 2004. The pH of the secretory pathway: Measurement, determinants, and regulation. Physiology 19, 207-215.

Petrova, K., Oyadomari, S., Hendershot, L.M., Ron, D., 2008. Regulated association of misfolded endoplasmic reticulum lumenal proteins with P58/DNAJc3. EMBO Journal 27, 2862-2872.

Pezzati, R., Bossi, M., Podini, P., Meldolesi, J., Grohovaz, F., 1997. High-resolution calcium mapping of the endoplasmic reticulum Golgi exocytic membrane system - Electron energy loss imaging analysis of quick frozen freeze dried PC12 cells. Molecular Biology of the Cell 8, 1501-1512.

Pfeffer, S.R., Rothman, J.E., 1987. Biosynthetic protein transport and sorting by the endoplasmic reticulum and Golgi. Annual Review of Biochemistry 56, 829-852.

Porter, K.R., Thompson, H.P., 1948. A particulate body associated with epithelial cells cultured from mammary carcinomas of mice of a milk-factor strain. Journal of Experimental Medicine 88, 15-24.

Porter, K.R., Claude, A., Fullam, E.F., 1945. A study of tissue culture cells by electron microscopy - methods and preliminary observations. Journal of Experimental Medicine 81, 233-246.

Prins, D., Michalak, M., 2011. Organellar Calcium Buffers. Cold Spring Harbor Perspectives in Biology 3, a004069.

Qin, S., Zhou, H.X., 2007. meta-PPISP: a meta web server for protein-protein interaction site prediction. Bioinformatics 23, 3386-3387.

Raykhel, I., Alanen, H., Salo, K., Jurvansuu, J., Nguyen, V.D., Latva-Ranta, M., Ruddock, L., 2007. A molecular specificity code for the three mammalian KDEL receptors. Journal of Cell Biology 179, 1193-1204.

Ritossa, F., 1962. New Puffing Pattern Induced by Temperature Shock and DNP in Drosophila. Experientia 18, 571-573.

Ritossa, F., 1996. Discovery of the heat shock response. Cell Stress & Chaperones 1, 97-98.

Rudiger, S., Germeroth, L., SchneiderMergener, J., Bukau, B., 1997. Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. EMBO Journal 16, 1501-1507.

Rutkowski, D.T., Kang, S.W., Goodman, A.G., Garrison, J.L., Taunton, J., Katze, M.G., Kaufman, R.J., Hegde, R.S., 2007. The role of p58IPK in protecting the stressed endoplasmic reticulum. Molecular Biology of the Cell 18, 3681-3691.

Sato, K., Nakano, A., 2002. Emp47p and its close homolog Emp46p have a tyrosine-containing endoplasmic reticulum exit signal and function in glycoprotein secretion in Saccharomyces cerevisiae. Molecular Biology of the Cell 13, 2518-2532.

Sato, K., Sato, M., Nakano, A., 1997. Rer1p as common machinery for the endoplasmic reticulum localization of membrane proteins. Proceedings of the National Academy of Sciences of the United States of America 94, 9693-9698.

Sato, K., Sato, M., Nakano, A., 2001. Rer1p, a retrieval receptor for endoplasmic reticulum membrane proteins, is dynamically localized to the Golgi apparatus by coatomer. Journal of Cell Biology 152, 935-944.

Schindler, R., Itin, C., Zerial, M., Lottspeich, F., Hauri, H.P., 1993. ERGIC-53, a membrane protein of the ER-Golgi intermediate compartment, carries an ER retention motif. European Journal of Cell Biology 61, 1-9.

Schweizer, A., Fransen, J.A.M., Bachi, T., Ginsel, L., Hauri, H.P., 1988. Identification, by a Monoclonal Antibody, of a 53-kD Protein Associated with a Tubulo-Vesicular Compartment at the cis-side of the Golgi Apparatus. Journal of Cell Biology 107, 1643-1653.

Serafini, T., Orci, L., Amherdt, M., Brunner, M., Kahn, R.A., Rothman, J.E., 1991. ADP-Ribosylation Factor Is a Subunit of the Coat of Golgi-derived COP-Coated Vesicles: A Novel Role for a GTP-Binding Protein. Cell 67, 239-253.

Shamu, C.E., Walter, P., 1996. Oligomerization and phosphorylation of the Ire1p kinase during intracellular signaling from the endoplasmic reticulum to the nucleus. EMBO Journal 15, 3028-3039.

Shen, J.S., Chen, X., Hendershot, L., Prywes, R., 2002a. ER stress regulation of ATF6 localization by dissociation of BiP/GRP78 binding and unmasking of golgi localization signals. Developmental Cell 3, 99-111.

Shen, Y., Hendershot, L.M., 2005. ERdj3, a stress-inducible endoplasmic reticulum DnaJ homologue, serves as a cofactor for BiP's interactions with unfolded substrates. Molecular Biology of the Cell 16, 40-50.

Shen, Y., Meunier, L., Hendershot, L.M., 2002b. Identification and characterization of a novel endoplasmic reticulum (ER) DnaJ homologue, which stimulates ATPase activity of BiP in vitro and is induced by ER stress. Journal of Biological Chemistry 277, 15947-15956.

Shikano, S., Li, M., 2003. Membrane receptor trafficking: Evidence of proximal and distal zones conferred by two independent endoplasmic reticulum localization signals. Proceedings of the National Academy of Sciences of the United States of America 100, 5783-5788.

Shiu, R.P.C., Pouyssegur, J., Pastan, I., 1977. Glucose depletion accounts for induction of two transformation-sensitive membrane proteins in Rous sarcoma virus-transformed chick embryo fibroblasts. Proceedings of the National Academy of Sciences of the United States of America 74, 3840-3844.

Sitia, R., Meldolesi, J., 1992. Endoplasmic-reticulum - a dynamic patchwork of specialized subregions. Molecular Biology of the Cell 3, 1067-1072.

Slepenkov, S.V., Witt, S.N., 2002. The unfolding story of the Escherichia coli Hsp70 DnaK: is DnaK a holdase or an unfoldase? Molecular Microbiology 45, 1197-1206.

Stagg, S.M., Gurkan, C., Fowler, D.M., LaPointe, P., Foss, T.R., Potter, C.S., Carragher, B., Balch, W.E., 2006. Structure of the Sec13/31 COPII coat cage. Nature 439, 234-238.

Tang, N.M., Ho, C.Y., Katze, M.G., 1996. The 58-kDa cellular inhibitor of the double stranded RNA-dependent protein kinase requires the tetratricopeptide repeat 6 and DnaJ motifs to stimulate protein synthesis in vivo. Journal of Biological Chemistry 271, 28660-28666.

Tao, J.H., Petrova, K., Ron, D., Sha, B.D., 2010. Crystal Structure of P58(IPK) TPR Fragment Reveals the Mechanism for its Molecular Chaperone Activity in UPR. Journal of Molecular Biology 397, 1307-1315.

Thor, F., Gautschi, M., Geiger, R., Helenius, A., 2009. Bulk Flow Revisited: Transport of a Soluble Protein in the Secretory Pathway. Traffic 10, 1819-1830.

Tisdale, E.J., Plutner, H., Matteson, J., Balch, W.E., 1997. p53/58 binds COPI and is required for selective transport through the early secretory pathway. Journal of Cell Biology 137, 581-593.

Travers, K.J., Patil, C.K., Wodicka, L., Lockhart, D.J., Weissman, J.S., Walter, P., 2000. Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. Cell 101, 249-258.

Tsai, J., Douglas, M.G., 1996. A conserved HPD sequence of the J-domain is necessary for YDJ1 stimulation of Hsp70 ATPase activity at a site distinct from substrate binding. Journal of Biological Chemistry 271, 9347-9354.

Vainio, P., Mpindi, J.P., Kohonen, P., Fey, V., Mirtti, T., Alanen, K.A., Perala, M., Kallioniemi, O., Iljin, K., 2012. High-Throughput Transcriptomic and RNAi Analysis Identifies AIM1, ERGIC1, TMED3 and TPX2 as Potential Drug Targets in Prostate Cancer. PLoS ONE 7, e39801.

van Huizen, R., Martindale, J.L., Gorospe, M., Holbrook, N.J., 2003. P58(IPK), a novel endoplasmic reticulum stress-inducible protein and potential negative regulator of eIF2 alpha signaling. Journal of Biological Chemistry 278, 15558-15564.

Velloso, L.M., Svensson, K., Pettersson, R.F., Lindqvist, Y., 2003. The crystal structure of the carbohydrate-recognition domain of the glycoprotein sorting receptor p58/ERGIC-53 reveals an unpredicted metal-binding site and conformational changes associated with calcium ion binding. Journal of Molecular Biology 334, 845-851.

Velloso, L.M., Svensson, K., Schneider, G., Pettersson, R.F., Lindqvist, Y., 2002. Crystal structure of the carbohydrate recognition domain of p58/ERGIC-53, a protein involved in glycoprotein export from the endoplasmic reticulum. Journal of Biological Chemistry 277, 15979-15984.

Vembar, S.S., Jonikas, M.C., Hendershot, L.M., Weissman, J.S., Brodsky, J.L., 2010. J Domain Co-chaperone Specificity Defines the Role of BiP during Protein Translocation. Journal of Biological Chemistry 285, 22484-22494.

Veratti, E., 1961. Investigations on fine structure of striated muscle fiber. Journal of Biophysical and Biochemical Cytology 10, 1-59.

Vitu, E., Gross, E., Greenblaft, H.M., Sevier, C.S., Kaiser, C.A., Fass, D., 2008. Yeast Mpd1p Reveals the Structural Diversity of the Protein Disulfide Isomerase Family. Journal of Molecular Biology 384, 631-640.

Vogel, M., Bukau, B., Mayer, M.P., 2006. Allosteric regulation of Hsp70 chaperones by a proline switch. Molecular Cell 21, 359-367.

Vollenweider, F., Kappeler, F., Itin, C., Hauri, H.P., 1998. Mistargeting of the lectin ERGIC-53 to the endoplasmic reticulum of HeLa cells impairs the secretion of a lysosomal enzyme. Journal of Cell Biology 142, 377-389.

Vos, M.J., Hageman, J., Carra, S., Kampinga, H.H., 2008. Structural and functional diversities between members of the human HSPB, HSPH, HSPA, and DNAJ chaperone families. Biochemistry 47, 7001-7011.

Votsmeier, C., Gallwitz, D., 2001. An acidic sequence of a putative yeast Golgi membrane protein binds COPII and facilitates ER export. EMBO Journal 20, 6742-6750.

Walter, P., Ron, D., 2011. The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation. Science 334, 1081-1086.

Wang, X.D., Matteson, J., An, Y., Moyer, B., Yoo, J.S., Bannykh, S., Wilson, I.A., Riordan, J.R., Balch, W.E., 2004. COPII-dependent export of cystic fibrosis transmembrane conductance regulator from the ER uses a di-acidic exit code. Journal of Cell Biology 167, 65-74.

Wang, X.Z., Kuroda, M., Sok, J., Batchvarova, N., Kimmel, R., Chung, P., Zinszner, H., Ron, D., 1998. Identification of novel stress-induced genes downstream of chop. EMBO Journal 17, 3619-3630.

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, 248-251.

Welsh, L.M., Tong, A.H.Y., Boone, C., Jensen, O.N., Otte, S., 2006. Genetic and molecular interactions of the Erv41p-Erv46p complex involved in transport between the endoplasmic reticulum and Golgi complex. Journal of Cell Science 119, 4730-4740.

Wendeler, M.W., Paccaud, J.P., Hauri, H.P., 2007. Role of Sec24 isoforms in selective export of membrane proteins from the endoplasmic reticulum. EMBO Reports 8, 258-264.

Wieland, F.T., Gleason, M.L., Serafini, T.A., Rothman, J.E., 1987. The rate of bulk flow from the endoplasmic-reticulum to the cell-surface. Cell 50, 289-300.

Wigren, E., Bourhis, J.M., Kursula, I., Guy, J.E., Lindqvist, Y., 2010. Crystal structure of the LMAN1-CRD/MCFD2 transport receptor complex provides insight into combined deficiency of factor V and factor VIII. Febs Letters 584, 878-882.

Wilson, C., Venditti, R., Rega, L.R., Colanzi, A., D'Angelo, G., De Matteis, M.A., 2011. The Golgi apparatus: an organelle with multiple complex functions. Biochemical Journal 433, 1-9.

Wilson, D.W., Lewis, M.J., Pelham, H.R.B., 1993. pH-dependent binding of KDEL to its receptor invitro. Journal of Biological Chemistry 268, 7465-7468.

Wu, M.S., Tu, T., Huang, Y.C., Cao, Y., 2013. Suppression subtractive hybridization identified differentially expressed genes in lung adenocarcinoma: ERGIC3 as a novel lung cancer-related gene. Bmc Cancer 13, 44.

Yamada, T., Fujimori, Y., Suzuki, A., Miyawaki, Y., Takagi, A., Murate, T., Sano, M., Matsushita, T., Saito, H., Kojima, T., 2009. A novel missense mutation causing abnormal LMAN1 in a Japanese patient with combined deficiency of factor V and factor VIII. American Journal of Hematology 84, 738-742.

Yan, W., Frank, C.L., Korth, M.J., Sopher, B.L., Novoa, I., Ron, D., Katze, M.G., 2002. Control of PERK eIF2 alpha kinase activity by the endoplasmic reticulum stress-induced molecular chaperone P58(IPK). Proceedings of the National Academy of Sciences of the United States of America 99, 15920-15925.

Yap, K.L., Ames, J.B., Swindells, M.B., Ikura, M., 1999. Diversity of conformational states and changes within the EF-hand protein superfamily. Proteins: Structure Function and Genetics 37, 499-507.

Yoshida, H., Haze, K., Yanagi, H., Yura, T., Mori, K., 1998. Identification of the cis-acting endoplasmic reticulum stress response element responsible for transcriptional induction of mammalian glucose-regulated proteins - Involvement of basic leucine zipper transcription factors. Journal of Biological Chemistry 273, 33741-33749.

Yoshida, H., Matsui, T., Yamamoto, A., Okada, T., Mori, K., 2001. XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. Cell 107, 881-891.

Yoshihisa, T., Barlowe, C., Schekman, R., 1993. Requirement for a GTPase-activating protein in vesicle budding from the endoplasmic-reticulum. Science 259, 1466-1468.

Zahedi, R.P., Volzing, C., Schmitt, A., Frien, M., Jung, M., Dudek, J., Wortelkamp, S., Sickmann, A., Zimmermann, R., 2009. Analysis of the membrane proteome of canine pancreatic rough microsomes identifies a novel Hsp40, termed ERj7. Proteomics 9, 3463-3473.

Zerangue, N., Schwappach, B., Jan, Y.N., Jan, L.Y., 1999. A new ER trafficking signal regulates the subunit stoichiometry of plasma membrane K-ATP channels. Neuron 22, 537-548.

Zeytuni, N., Zarivach, R., 2012. Structural and Functional Discussion of the Tetra-Trico-Peptide Repeat, a Protein Interaction Module. Structure 20, 397-405.

Zhang, B., Kaufman, R.J., Ginsburg, D., 2005. LMAN1 and MCFD2 form a cargo receptor complex and interact with coagulation factor VIII in the early secretory pathway. Journal of Biological Chemistry 280, 25881-25886.

Zhang, B., Cunningham, M.A., Nichols, W.C., Bernat, J.A., Seligsohn, U., Pipe, S.W., McVey, J.H., Schulte-Overberg, U., de Bosch, N.B., Ruiz-Saez, A., White, G.C., Tuddenham, E.G.D., Kaufman, R.J., Ginsburg, D., 2003. Bleeding due to disruption of a cargo-specific ER-to-Golgi transport complex. Nature Genetics 34, 220-225.

Zhang, B., Spreafico, M., Zheng, C., Yang, A., Platzer, P., Callaghan, M.U., Avci, Z., Ozbek, N., Mahlangu, J., Haw, T., Kaufman, R.J., Marchant, K., Tuddenham, E.G.D., Seligsohn, U., Peyvandi, F., Ginsburg, D., 2008. Genotype-phenotype correlation in combined deficiency of factor V and factor VIII. Blood 111, 5592-5600.

Zhang, B., McGee, B., Yamaoka, J.S., Guglielmone, H., Downes, K.A., Minoldo, S., Jarchum, G., Peyvandi, F., de Bosch, N.B., Ruiz-Saez, A., Chatelain, B., Olpinski, M., Bockenstedt, P., Sperl, W., Kaufman, R.J., Nichols, W.C., Tuddenham, E.G.D., Ginsburg, D., 2006. Combined deficiency of factor V and factor VIII is due to mutations in either LMAN1 or MCFD2. Blood 107, 1903-1907.

Zhang, K.Z., Kaufman, R.J., 2006. The unfolded protein response - A stress signaling pathway critical for health and disease. Neurology 66, S102-S109.

Zhang, L.Y., Liu, M., Li, X., Tang, H., 2013. miR-490-3p Modulates Cell Growth and Epithelial to Mesenchymal Transition of Hepatocellular Carcinoma Cells by Targeting Endoplasmic Reticulum-Golgi Intermediate Compartment Protein 3 (ERGIC3). Journal of Biological Chemistry 288, 4035-4047.

Zhang, X.P., Dong, C.M., Wu, Q.J., Balch, W.E., Wu, G.Y., 2011. Di-acidic Motifs in the Membrane-distal C Termini Modulate the Transport of Angiotensin II Receptors from the Endoplasmic Reticulum to the Cell Surface. Journal of Biological Chemistry 286, 20525-20535.

Zheng, C.L., Zhang, B., 2013. Combined Deficiency of Coagulation Factors V and VIII: An Update. Seminars in Thrombosis and Hemostasis 39, 613-620.

Zheng, C.L., Liu, H.H., Yuan, S.G., Zhou, J.H., Zhang, B., 2010. Molecular basis of LMAN1 in coordinating LMAN1-MCFD2 cargo receptor formation and ER-to-Golgi transport of FV/FVIII. Blood 116, 5698-5706.

Zinszner, H., Kuroda, M., Wang, X.Z., Batchvarova, N., Lightfoot, R.T., Remotti, H., Stevens, J.L., Ron, D., 1998. CHOP is implicated in programmed cell death in response to impaired function of the endoplasmic reticulum. Genes & Development 12, 982-995.