The molecular basis for ERp57/calreticulin complex formation

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Science, 2003.

THE MOLECULAR BASIS FOR ERP57/CALRETICULIN COMPLEX FORMATION.

Sarah Jayne Russell

Division of Biochemistry
School of Biological Sciences
University of Manchester

Contents

List of Figures

List of Tables

Abstract

In mammalian cells newly synthesised proteins are translocated across the ER membrane and their subsequent folding is facilitated by an array of folding factors present in the lumen. These include the lectins calreticulin and calnexin, which form complexes with ERp57 to generate glycoprotein specific molecular chaperones. ERp57 is a member of the protein disulphide isomerase (PDI) family and its binding to ER lectins can be reconstituted *in vitro*. I have exploited this approach to define the regions of ERp57 that are necessary and sufficient for its specific interaction with calreticulin and calnexin. Truncated forms of ERp57, chimeric proteins containing various domains of ERp57 and PDI (which does not interact with calreticulin) and ERp57 ***b'*** domain point mutants have been constructed. By analysing the interactions of ERp57 derivatives with calreticulin using both cross-linking and binding assays I have been able to provide detailed insights into the molecular basis for the specific assembly of these components within the ER lumen.

My results indicate that the ***b*** and ***b'*** domains of ERp57 are necessary, but not sufficient for binding to both calreticulin and calnexin. The more stringent binding assay revealed that the ***a'*** domain of ERp57 significantly enhanced binding to biotin-tagged calreticulin. The ERp57 C-terminal extension also increased binding to biotin-tagged calreticulin, perhaps by playing a role in the overall stability of the ERp57. In addition, the ERp57 ***b'*** domain point mutants show that certain amino acids in this domain, in particular residues F280, V283 and F299, may be crucial for binding to calreticulin, consistent with the principal lectin-binding site being located in the ***b'*** domain. However, the binding region clearly extends into other domains, in particular the ***b*** and ***a'*** domains.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Acknowledgements

I wish to thank all of the members of High and Woodman laboratories for their help and friendship which made the time pass so quickly.

I give special thanks to:

My supervisor, Professor Steve High, for enabling me to complete my postgraduate studies in his laboratory under his guidance. His encouragement and generosity with his time during my PhD has been greatly appreciated.

My research has been greatly aided by encouragement and advice from my advisor Professor Neil Bulleid and my friend and work colleague Fabienne Lecomte.

Many thanks to Alex Coe for his enduring support.

This work was supported by a BBSRC studentship. I would also like to thank the Division of Biochemistry in the University of Manchester for supporting my research and postgraduate training.

Abbreviations

AAT	α-Anti-Trypsin
AEBSF	4-(-2-Aminoethyl)-bezenesulphonylfluoride.HCL
ACTA	Aurin Tricarboxylic Acid
BiP	Heavy Chain Binding Protein
BMH	Bismaleimidohexane
BSA	Bovine Serum Albumin
DMSO	Dimethyl Sulfoxide
ER	Endoplasmic Reticulum
ERAD	ER-Associated Protein Degradation
FAD	Flavine Adenine Dinucleotide
GTP	Guanosine 5'-Triphosphate
MHC	Major Histocompatability Complex
PCR	Polymerase Chain Reaction
PDI	Protein Disulphide Isomerase
PMSF	phenylmethylsulfonyl fluoride
PPIase	Peptidyl Prolyl *cis-trans* Isomerase
SDS-PAGE	Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis
SMCC	Succinimidyl-4-(N-maleimidomethyl) Cyclohexane-1-Carboxylate
SPC	Signal Peptidase Complex
SPase	Signal Peptidase
SR	SRP Receptor
SRP	Signal Recognition Particle
TRAM	Translocating Chain-Associating Membrane Protein
UGGT	UDP-Glc:Glycoprotein Glucosyl-Transferase
UPR	Unfolded Protein Response

CHAPTER - 1

Introduction

1 Introduction

1.1 The Endoplasmic Reticulum (ER)

The endoplasmic reticulum (ER)* plays a central role in lipid and protein biosynthesis and it is the site of manufacture of transmembrane proteins and lipids for many of the cells' organelles, including the ER. Most proteins that are destined for the ER, the Golgi, lysosomes and the cell surface are initially delivered to the ER. Many ribosomes engaged in the synthesis of proteins bind to the ER membrane with the attached nascent chains being frequently co-translationally transported across the membrane, particularly in mammals (Johnson and van Waes, 1999). In practice, this means that the ribosome is directly attached to the ER membrane and nascent membrane or secretory proteins are never released into the cytosol. This mechanism is believed to prevent the synthesis of large, partially folded regions of polypeptide that may be unable to translocate across the ER membrane (Walter and Johnson, 1994).

Authentic translocation across, or insertion into, the membrane of the ER is a fundamental phase of biogenesis for most secretory and membrane proteins. In addition to transport across the membrane, several ER specific modifications are crucial for proper folding and maturation, and hence for the production of functional proteins. These modifications can include the cleavage of ER signal sequences, N-linked glycosylation, disulphide bond formation, intrachain folding and oligomerisation with other proteins. Many of these processes can only occur during protein translocation, and as this process provides a unique opportunity to gain access to all areas of the protein while it is at least partially unfolded (Hegde and Voigt, 1998).

---

* Abbreviations are listed on page 16

Protein folding is a complex event and many defects can arise during this process. In order for newly synthesised proteins to achieve their correct conformation a number of ER components are required, namely molecular chaperones and enzymes that catalyse protein folding (Zapun *et al.,*1999). Protein folding is an essential element of biology and of increasing relevance to medicine since the inability of specific proteins to form their native structures is the basis of a variety of human diseases (Thomas *et al.,*1995). A number of diseases are linked to changes in the sequence of a protein that result in alterations in the way it interacts with molecular chaperones. In the case of mutant proteins synthesised at the ER, these changes frequently result in the retention of the mutant protein within the cell and its subsequent degradation via a process usually known as 'quality control'. It has even been suggested that by fully understanding the mechanisms by which ER chaperones function, it may be possible to manipulate their actions and thereby alleviate some disease states (Thomas *et al.,*1995). Whilst protein folding at the ER has been studied in a number of different organisms, the principal focus of this review will be the role of molecular chaperones in the ER of mammalian cells.

1.2 ER targeting signals.

The first step in protein biosynthesis at the ER is the targeting of a nascent chain and its attached ribosome to a translocation site in the ER membrane. In mammalian cells, targeting usually occurs during translation of the nascent protein, and is achieved by a cytosolic targeting factor, called signal recognition particle (SRP). SRP recognises a signal sequence contained within the nascent chain. ER targeting signals can be described as a stretch of 6 – 20 hydrophobic amino acids that are usually located towards the amino terminus of a protein (von Heijne, 1985). Analysis of multiple signal sequences showed

that polar C-terminal and N-terminal regions normally flank its central hydrophobic region (Figure 1.1). This 'h-region' is however the most essential element and it is required for targeting and membrane insertion (von Heijne, 1985). The C terminal-region may contain small, uncharged residues at positions -1 and –3, which determine the site for the cleavage of the signal peptide by the signal peptidase (von Heijne, 1990).

Figure 1.1 Structure of ER signal sequences

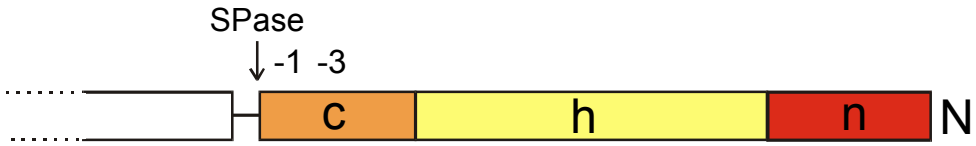


Figure 1.1 Structure of ER signal sequences (adapted from Martoglio and Dobberstein, 1998). The central hydrophobic h-region (yellow) is flanked by hydrophilic N- (red) and C-terminal (orange) regions. Arrow: cleavage site for signal peptidase (SPase).

ER signal sequences can be exchanged between different proteins and even between different organisms and still retain their function (Gierasch, 1989). However, subtleties in the signal sequence structures may endow them with additional functions (Martoglio and Dobberstein, 1998). The signal sequence determines whether a protein is translocated across the ER membrane or is inserted into it, and if inserted what transmembrane topology the protein has. Following targeting to the ER, the signal sequence may either be cleaved from the protein or may remain attached, forming an integral part of the mature protein.

1.3 SRP dependent targeting.

The SRP dependent targeting of most proteins to the mammalian ER involves the recognition of the signal sequence on a nascent polypeptide as it emerges from the ribosome, and the association of this complex with the ER membrane. The recognition of the signal sequence is achieved by the interaction of a cytosolic ribonucleoprotein complex, the signal recognition particle (SRP) with the signal sequence (Lutcke, 1995). SRP 'samples' nascent chains for the presence of signal sequences by interacting transiently with ribosomes producing nascent chains (Ogg and Walter, 1995). Interaction with the signal sequence causes SRP to bind the ribosome more tightly, causing either a slowing down in the rate of translation of the nascent protein (Wolin and Walter, 1989) or in some cases a complete arrest of protein synthesis (Walter and Blobel, 1981). The ternary complex of SRP, the ribosome and the nascent chain is then targeted to the ER membrane via a GTP-dependent interaction with its ER-located receptor, the SRP receptor (SR; Meyer and Dobberstein, 1980; Gilmore *et al.,* 1982a; Gilmore *et al.,* 1982b). The fundamental components and the molecular mechanism of SRP-dependent protein targeting have been conserved during evolution (for review see Keenan *et al.,* 2001)

1.3.1 The signal recognition particle (SRP).

Mammalian SRP is a ribonucleoprotein complex comprising a 7S RNA scaffold (SRP RNA) with six protein subunits named by their apparent molecular weight SRP9, SRP14, SRP19, SRP54, SRP68 and SRP72. The interaction of the signal sequence with SRP has been probed by photocross-linking experiments, which demonstrate that the 54-kDa subunit of the SRP is the site of signal sequence binding (High *et al.,* 1991; Kurzchalia *et*

al., 1986). This binding can only occur once the signal sequence has emerged from the ribosome, *i.e.*, once 70-80 residues of the nascent chain have been synthesised (Wiedmann *et al.,* 1987a). SRP54 is composed of three domains, termed M, G and N. The methionine-rich M domain binds both the SRP RNA, in the presence of the 19 kDa SRP subunit (Römisch *et al.,* 1990; Zopf *et al.,* 1990) and the signal sequence (High and Dobberstein, 1991). Structural features of the SRP54 M domain were revealed by determination of the crystal structure of the *Thermus aquaticus* homologue (Ffh) and include a large exposed hydrophobic groove, constructed from three alpha helices, and a long flexible loop (Keenan *et al.,* 1998). The dimensions of the M domain groove are compatible with binding a signal sequence in an α-helcial conformation (Keenan *et al.,* 1998). The G domain is a Ras-like GTPase with an additional subdomain unique to the SRP family of GTPases (Freymann *et al.,* 1997; Miller *et al.,* 1993). This domain is required for interaction with the SRP receptor when docking at the ER membrane (Zopf *et al.,* 1993). The amino-terminal N domain is a four-helix bundle closely associated with the G domain, and is thought to promote efficient recognition of the signal sequence (Newitt and Bernstein, 1997).

Other subunits of mammalian SRP are involved in the translational slowing that follows the binding of SRP to the ribosome (Siegel and Walter, 1985) and may also act to stimulate the dissociation of guanidine nucleotide from SRP54 (Althoff *et al.,* 1994).

1.3.2 The SRP receptor (SR).

The mammalian SRP receptor (SR) is composed of two GTPases, SRα and SRβ (Connolly and Gilmore, 1989; Miller *et al.,* 1995; Tajima *et al.,* 1986). The SRα subunit is a 69kDa peripheral membrane protein, with a GTPase domain identified at its C-terminus that is closely related to the GTPase domain of SRP54 (Freymann *et al.,* 1997; Montoya *et al.,* 1997). The SRα subunit associates with SRβ, a 30kDa integral membrane protein (Lauffer *et al.,* 1985; Miller *et al.,* 1995). It was shown that the role of SRβ in protein targeting and translocation was unlikely to be simply anchorage of SRα to the ER membrane, since deletion of the transmembrane domain of SRβ in *S.cerevisiae*, producing a soluble SR, does not significantly impair SR function (Ogg *et al.,* 1998). However, mutations in the GTPase domain of SRβ do disrupt SR function (Ogg *et al.,* 1998).

1.3.3 The SRP/SR GTPase Cycle

The SRP-dependent ER targeting process is intimately coupled to a GTPase cycle involving three GTPases; SRP54, SRα and SRβ. The GTPase properties of SRP and SR during the targeting reaction can be modulated by additional components. Firstly, GTP binding by SRP is thought to be stimulated by the ribosome (Bacher *et al.,* 1996). Therefore, SRP bound to the ribosome is likely to be in its GTP bound form before its interaction with SR. Secondly, in contrast to ribosome binding of SRP, signal sequence binding to SRP is thought to inhibit GTP binding and hydrolysis (Miller *et al.,* 1993; Miller *et al.,* 1994). This supports a model in which SRP binding to the nascent chain's signal sequence inhibits GTP hydrolysis (but not GTP binding) until the ribosome-nascent chain complex is transferred to the translocon in the ER membrane. Finally, GTP

hydrolysis by SRP is stimulated significantly by its interaction with SR (Powers and Walter, 1995).

As with SRP, the GTPase activity of SR is also modulated by other components. Firstly, GTP-bound SRβ interacts directly with ribosomes bearing a nascent chain and this stimulates the GTP hydrolysis of SRβ (Bacher *et al.,* 1999). Secondly, SRβ must be in its GTP-bound form in order to interact with SRα (Legate *et al.,* 2000). These data suggest that the localisation of SRα to the ER membrane might be indirectly regulated by the ribosome. In doing so, it allows SRα to scan the now membrane-bound ribosome for the presence of SRP. Interaction of SRα with SRP would then be the trigger for release of the signal sequence from SRP and its insertion into the translocation site, whilst the ribosome is released from its interaction with SRβ by ribosome-induced hydrolysis of its bound GTP (Bacher *et al.,* 1999). Upon transfer of the ribosome/nascent chain complex to the translocation site any pause in translation, imposed by SRP binding to the ribosome, is lifted and translation resumes as normal, with the nascent chain being threaded into the translocation channel (Walter and Blobel, 1981). Additionally, recent data suggests that the translocon may regulate the GTP hydrolysis cycle of the SRP-SR complex (Bacher *et al.,* 1999).

Current models of the SRP-dependent targeting process suggest that both the ribosome and translocon can regulate the enzymatic activities of SRP and SR, ensuring efficient and unidirectional targeting of nascent chains to the ER translocon (Figure 1.2). SRP interacts with ribosomes with exposed signal sequences in the cytosol, while SR associates with empty translocons in the ER membrane. These interactions would set both SRP and SR to the GTP-bound state and in this state SRP and SR display high affinity for each other. This in turn would facilitate transfer of the ternary complex of SRP, the ribosome and the nascent chain to the translocon and enable the transfer of the nascent chain

from SRP to the translocon as well as the formation of a functional ribosome-membrane junction.

Figure 1.2 Translocation at the Mammalian ER membrane.

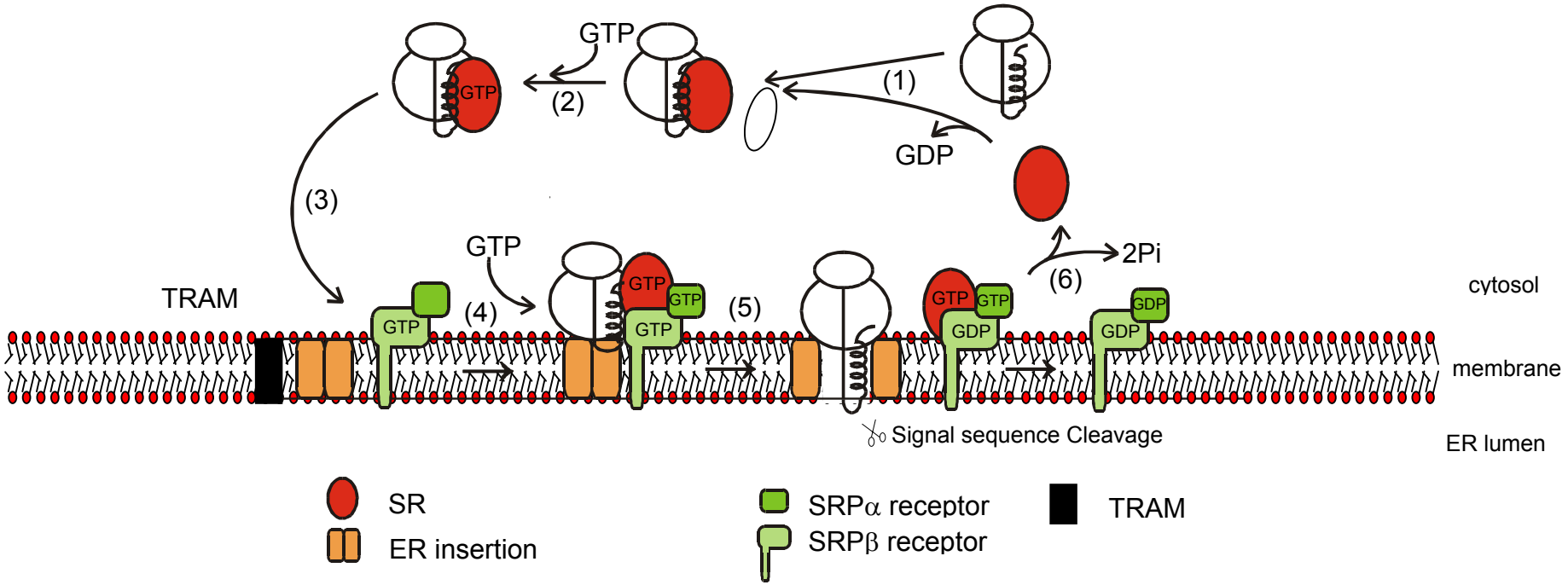


Figure 1.2 A recent model of SRP/SR mediated targeting of nascent chains to the ER membrane (adapted from Bacher *et al.,* 1999). The signal recognition particle (SRP) binds to the signal sequence (SS) of the nascent chain as it emerges from the ribosome and slows down the rate of translation (1). Due to a ribosome-induced increase in the affinity of SRP for GTP, SRP binds GTP (2). The ribosome / nascent chain / SRP(GTP) complex is targeted to the ER membrane by interaction of the ribosome with SRβ(GTP) (3). The proximity of SRα to the ribosome allows it to scan for the presence of SRP(GTP) and an interaction between SRP and SRα induces binding of GTP by SRα (4). As a result of GTP binding by SRα, the signal sequence is released by SRP(GTP) and inserts into the translocon (5), whilst GTP hydrolysis promotes the dissociation of the ribosome from SRβ, and of SRP from SRα (6). The TRAM protein enhances the early stages of protein translocation. The signal peptidase complex (scissors symbol) cleaves the signal sequence where a suitable site is present. The signal sequences of some membrane proteins are not cleaved and function as membrane anchors.

1.4 Translocation across the Mammalian ER

Proteins are translocated across the ER in an unfolded, or partially folded, state and via a process that is normally unidirectional (reviewed in Rapoport *et al.,* 1996a; Rapoport *et al.,* 1996b; Wilkinson *et al.,* 1997 Johnson and van Waes, 1999). The nascent chain is thought to insert in a loop-like conformation, and when the nascent chain is initially inserted into the translocon the chain is short and the lumenal side of the ER translocon channel is closed. As the polypeptide increases in length, the channel opens letting the nascent chain enter the ER lumen (Rapoport *et al.,*1996a; Rapoport *et al.,*1996b).

In vitro reconstitution and cross-linking studies have shown that the Sec61 complex is a core component of the ER translocon, and that it is necessary for translocation to occur (Görlich and Rapoport, 1993). High-resolution electron microscopy studies and image reconstruction of both the yeast and mammalian pore complexes suggest that the Sec61 complex assembles into a trimer or tetramer, forming a central pore that is likely to be the channel through which preproteins are translocated (Hanein *et al.,* 1996; see Figure 1.3). The interior of the pore functions as an aqueous channel (Gilmore and Blobel, 1985; Crowley *et al.,*1993) and measurements using fluorescent probes placed within a translocating polypeptide indicate that the diameter of the pore in an active translocon may be as large as 80 Å (Hamman *et al.,*1997). This value is significantly higher than that obtained by EM studies that indicate a size of ~15-35 Å (Hanein *et al*., 1996; Beckmann *et al.*, 1997; Beckmann *et al.,* 2001). Regardless of these apparent differences, the pore is clearly of sufficient diameter to accommodate polypeptides in an α-helical conformation, or containing some other form of secondary structure.

Figure 1.3 Translocon pore

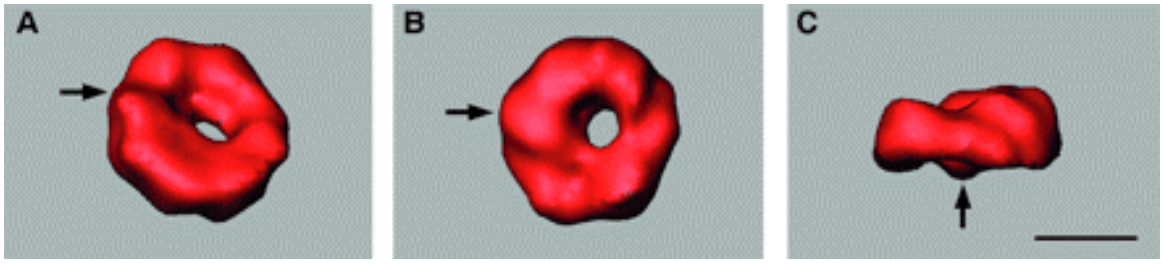


Figure 1.3 Close up views of the Sec61 oligomer (taken from Beckmann *et al.,* 1997). (**A**) Surface facing the ribosome. There is a vestibule (diameter of 35 Å) formed by the funnel-like structure and a pore (diameter of 15 Å). (**B**) Surface facing away from the ribosome. (**C**) View of the side opposite the attachment site. The ribosome would be situated underneath the channel. The wall opposite the attachment site is thinner and more regular. Arrows indicate the ribosome attachment site. Scale bar, 50 Å.

The minimal components required for the translocation of a number of proteins were found to consist of the signal recognition particle (SRP), the SRP receptor and the Sec61 complex. Many other proteins also required the translocating chain-associating membrane (TRAM) protein and even the transfer of so-called "TRAM-independent" proteins was found to be stimulated by TRAM (Hegde and Voigt, 1998). A number of other components are also considered to be part of the ER translocation machinery, and these may stimulate the efficiency of translocation or regulate the process in some way. It may be important to modulate the translocation process so that complex modifications, which are important for the maturation of certain proteins, can be carried out (Hedge and Voigt, 1998). For example, protein disulphide isomerase (PDI) can bind to nascent proteins as they emerge into the ER lumen, thus catalysing disulphide bond formation before the proteins are completely translocated (Freedman *et al.,*1994).

The same translocation complex allows proteins to either cross the ER membrane completely and enter the lumen or become integrated into the lipid bilayer (Do *et al*., 1996; Laird and High, 1997; Mothes *et al.*, 1997; Heinrich *et al.,* 2000). Photoreactive cross-

linking suggests that the nascent chain is transferred directly from the ribosome into the Sec61 translocon complex, of which Sec61α is the major constituent (Mothes *et al.,*1994). Cross-linking experiments also indicate that different parts of the nascent chain can interact with different components of the translocation machinery (High *et al.,*1993). Additionally, studies revealed that a single site within a nascent chain can interact with different components, specifically Sec61α and phospholipids, inferring that the nascent chain may be relatively mobile whilst in the translocation channel (Martogolio and Dobberstein, 1996). There are currently two main models describing the mechanism of lateral step-wise integration into the ER membrane (High and Laird, 1997), and it has been suggested that the transmembrane domains of proteins can either enter the lipid bilayer sequentially, as they are inserted into the translocon channel (Do *et al.,*1996), or only following the completion of translation and the release of the nascent chain from the ribosome (Borel, 1996).

Biochemical studies using reconstituted proteoliposomes depleted of lumenal components indicate that the driving force for cotranslational translation into the ER can be provided by the ribosome at least *in vitro*. Thus, the nascent chain can be 'pushed' into the lumen (Nicchitta *et al.,* 1990). However, the lumenal components in the ER can play a role in translocation by 'pulling' polypeptides into the ER. In the yeast *Saccharomyces cerevisiae*, numerous studies have led to the conclusion that BiP and Sec63p are responsible for driving polypeptide transport through the translocon channel and into the ER (reviewed in Brodsky *et al.,* 1996; see Section 1.5.1).

ER chaperones may also be required to gate the aqueous pore created by the ER translocon. Johnson and colleagues used an *in vitro* system in which lumenal access to a preprotein confined within the Sec61 complex could be assessed (Hamman *et al.,* 1998). They found that in mammalian microsomes it is BiP that seals the lumenal side of the pore

until the translocating polypeptide reaches a length of ~50 amino acids (Hamman *et al.,*1998). Recent results point to additional chaperone modulators required for protein translocation, and the *Yarrowia lipolytica* Sls1 protein, and its *S. cerevisiae* homologue (known as Per100p, Sil1p, or scSls1p), interact with the ATPase domain of BiP thereby regulating its activity (Kabani *et al.,*2000; Tyson and Stirling, 2000). Sls1p/Sil1p may act as a nucleotide exchange factor for BiP (Kabani *et al.,*2000), since the deletion of the SIL1 gene alone does not compromise translocation (Tyson and Stirling, 2000), but rather exacerbates the translocation defect obtained with a mutant form of BiP (Kabani *et al.,* 2000).

1.5 ER molecular Chaperones- an overview

As the nascent chain enters the ER it encounters an environment that is dramatically different from the cytoplasm, and in fact more similar to the environment present outside the cell. The ER is the primary regulator of cellular Ca^{++} levels (Stevens and Argon, 1999) and it possesses a higher concentration of free Ca^{++} than that found in the cytosol (~1mM in the ER compared with ~100nM in the cytosol (Chapman *et al.,*1998). The ER environment is also significantly more oxidising than the cytoplasm, with a redox potential of –230mV versus –150mV (Hwang *et al.,*1992). This favours disulphide bond formation in the ER, whereas disulphide bonds are virtually absent in the cytoplasm (Rietsch and Beckwith, 1998). The presence of disulphide bonds in many secretory proteins is often required for their correct folding and/or biological activity (Rietsch and Beckwith, 1998). Additionally, many secretory proteins are N-glycosylated in the ER, and this modification is also frequently required for proper folding and/or activity (Helenius and Aebi, 2001).

A key function of the ER is to serve as a protein-folding compartment and to ensure that proteins achieve their correct conformation (Hurtley and Helenius, 1989; Hammond and Helenius, 1995) . The ER lumen contains a number of distinct molecular chaperones and folding factors, which are involved in the modulation and assembly of newly synthesised proteins and protein complexes (reviewed in Ferrari and Söling, 1999; Zapun *et al.,* 1999). The basic primary sequence of a protein can often be sufficient to specify its ultimate tertiary structure. However, the folding of a secretory protein into a tertiary structure within a cell poses specific problems, since the nascent chain is gradually exposed on the trans-side of the membrane during translocation. Furthermore, many polypeptides must assemble into oligomeric complexes and must do so in the high effective protein concentration present within the ER, where there are many potential opportunities for inappropriate associations. Molecular chaperones and enzymes are employed within the ER lumen to prevent such mis-associations and to facilitate correct folding.

'Classical' molecular chaperones are proteins that associate transiently with the folding intermediates of a large number of different proteins via a cycle of binding and release. In this way they act to promote proper folding and assembly, and prevent protein aggregation (Zapun *et al.,*1999). The interactions of newly synthesised proteins with ER chaperones forms part of an elaborate mechanism, including ER retention and retrieval, regulation of protein expression, assisted protein folding and degradation of misfolded proteins, which ensures that only properly folded and oligomerised secretory proteins exit the ER (reviewed in Ellgaard *et al.,* 1999; Wickner *et al.,* 1999). The specific molecular mechanisms by which chaperones recognise their target proteins, and promote, inhibit or reverse folding and assembly, still remain to be elucidated. The major classes of ER-

resident chaperones (Table 1.1) are highly conserved in eukaryotes and will be discussed briefly below.

Table 1.1 Classes of chaperones found in the ER of mammalian cells.

Chaperone	Family	Protein name
Classical chaperones	GRP78/BiP GRP94	GRP78/BiP GRP94/gp96
Peptidyl propyl isomerases	FK506-binding proteins Cyclophilins	FKBP13, FKBP65 Cyclophilin B, C and 40
Protein disulphide isomerases	PDI PDI-like	PDI, ERp57, P5, ERp72, PDIR.
Lectin-like proteins	Calnexin/calreticulin	Calnexin Calreticulin
Glucosyltranferases	UGGT	UDP-glucose: glycoprotein glucosyltransferase

Table 1.1. Classes of chaperones found in the ER of mammalian cells (adapted from Zapun *et al.,* 1999).

1.5.1 BiP

BiP is an abundant molecular chaperone of the ER and a member of the Hsp70 family of chaperones (reviewed in Gething, 1999; Zapun *et al.,* 1999). BiP binds preferentially to extended stretches of hydrophobic polypeptides, suggesting that it may recognise unfolded proteins. Peptide binding to, and release from, BiP is coupled to the binding and hydrolysis of ATP (Flynn *et al.,*1989). BiP's role in protein folding has been well documented (Gething and Sambrook, 1992; Wei and Hendershot, 1996) and it probably aids protein folding by preventing 'off-pathway' folding intermediates from forming. BiP has been shown to interact transiently with a number of peptides and unfolded proteins, including unassembled immunoglobulin light and heavy chains (Melnick *et al.,*1994).

In the ER of yeast, BiP interacts with a 'DnaJ'-like domain of the Sec63p subunit of the ER translocation machinery. This interaction is required for the translocation of newly synthesised proteins, including the completion of translocation and the release from the translocon (Brodsky and Schekman, 1993; Lyman and Schekman, 1995; Lyman and Schekman, 1997). Recently data from studies in canine pancreatic microsomes have revealed a homologue of yeast protein Sec63p, which interacts with DnaK-like proteins, such as BiP (Tyedmers *et al.,* 2000). In this context, BiP may act as a molecular ratchet or a motor that either directs the nascent protein towards the lumen or pulls the nascent chain into it respectively (Rapoport *et al.,*1999).

1.5.2 GRP94

The glycoprotein GRP94 is abundant in the ER and belongs to the Hsp90 family of chaperones (reviewed in Nicchitta, 1998; Zapun *et al.,* 1999). Although it is abundant, its specific function(s) are unclear and the lack of an obvious homologue in *S. cerevisiae* has been taken to indicate that GRP94 may be unique to multicellular eukaryotes (Nicchitta, 1998). Analysis of GRP94 structure, indicates that the molecule is a tri-nodular rod consisting of two subunits (Wearsch and Nicchitta, 1996). The oligomerisation domain is in the C-terminus of the polypeptide and the amino-termini are orientated at opposite ends of the molecule (Wearsch and Nicchitta, 1996).

GRP94 is thought to have a similar function to BiP, since it has been shown to interact transiently with a number of peptides, including unassembled immunoglobulin chains (Melnick *et al.,*1994). However, it seems likely that GRP94 is mechanistically distinct from BiP, since it appears to possess only a weak ATPase activity (Wearsch and Nicchitta, 1997).

1.5.3 Peptidyl propyl isomerases

A distinct group of enzymes involved in protein folding in the ER are the peptidyl propyl isomerases, which catalyse the isomerisation between the *cis* and *trans* forms of the imino acid, proline (reviewed in Zapun *et al.,* 1999). This enzyme group is comprised of two main families: the immunophilins and the FK-binding proteins, which bind to the immunosuppressant cyclosporin and the FK506 compound respectively. FK-binding proteins have been identified in both the mammalian (Jin *et al.,*1991) and yeast (Partaledis and Berlin, 1993) ER implying that proline isomerisation is an important step during

protein folding in the ER lumen. PPIases are also induced by the unfolded protein response consistent with a role in promoting protein folding *in vivo* (UPR; Section 1.10.3; Travers et al., 2000).

1.5.4 Protein Disulphide Isomerases (PDIs).

In addition to the archetypal protein disulphide isomerase (PDI), there are several other 'PDI-like' proteins present in the ER lumen (see Figure 1.4; reviewed in Ferrari and Söling, 1999; Zapun *et al.,* 1999; Freedman *et al.,* 2002). Two of these PDI-like proteins, ERp57 and P5, each contain two thioredoxin-like domains and resemble PDI in both their size and modular organisation. Another two, ERp72 (also known as CaBP2 (Mazzarella *et al.,*1990)) and PDIR (PDI-related protein), contain three thioredoxin-like domain and are hence larger than PDI. A further PDI-like protein, PDIp, is expressed exclusively in the pancreas (Desilva and Lan, 1996). With the exception of ERp57, the specific function(s) of these various PDI homologues remains unclear at present.

Members of the PDI family have both enzymatic and chaperone functions to fulfil in the ER (Ferrari and Söling, 1999). Not only do they catalyse disulphide bond formation within newly synthesised proteins, but they also form transient disulphide bonded intermediates with their protein substrates (Molinari and Helenius, 1999b). In the ER lumen, the formation of disulphide bonds is an important protein modification and these bonds serve to stabilise the native conformation of many secretory proteins (see Section 1.9). Disulphide bridges are covalent bonds that are stable in oxidising conditions, including those normally found outside the cell. Hence, the formation of disulphide bonds enables secretory proteins to assume stable secondary structure folds that are difficult to unfold. *In vitro* the formation of disulphide bonds is often the rate-limiting step during

protein folding, and components that catalyse the reactions are considered to be 'true' folding enzymes (Freedman *et al.,*1994). The chemical steps involved in disulphide bond formation and rearrangement are intrinsically slow compared to conformational rearrangements within the polypeptide. Detailed studies have shown that some members of the PDI family have different substrate specificities, and that they interact with nascent and fully translocated proteins present within the ER (Molinari and Helenius, 1999b). The best-characterised members of the PDI family, PDI and ERp57, are discussed below in more detail (see sections 1.7 & 1.8).

Figure 1.4 Domain Structures of Mammalian PDI Family Members.

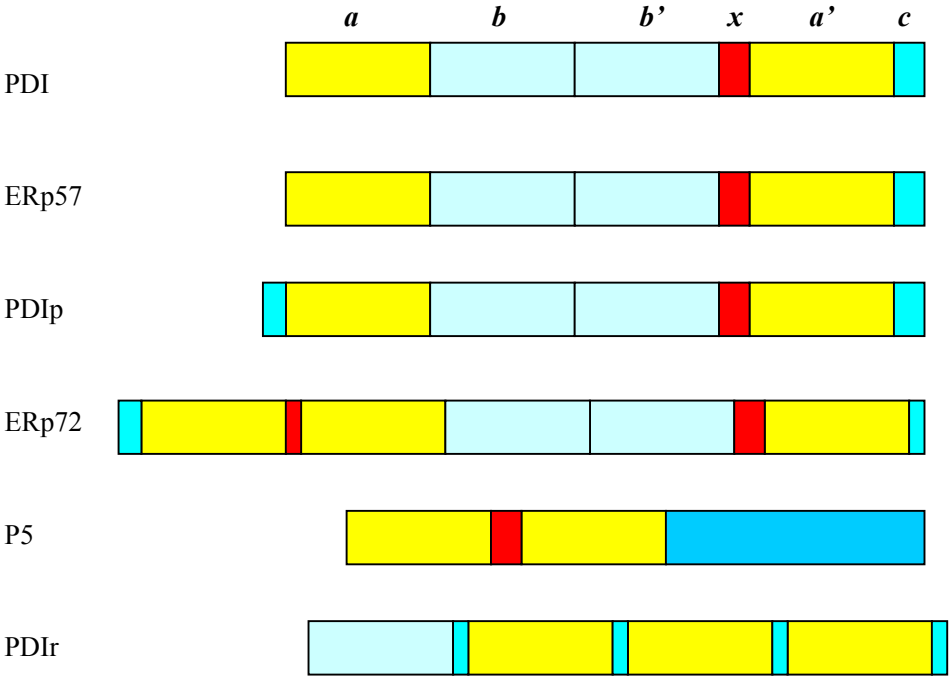


Figure 1.4 The overall domain structures of mammalian PDI family members are shown for comparison with that of human PDI (adapted from Freedman *et al.,* 2002). Homologous domains are coloured as for the corresponding domains in PDI. Regions of unknown function or structural homology are in shades of blue.

1.5.5 ER Lectins

Calnexin is a phosphorylated membrane protein of 65kDa with the bulk of its structure in the ER lumen, whilst calreticulin is its 46kDa soluble equivalent (discussed in more detail in section 1.6). Both proteins contain several Ca^{++} binding regions, the primary amino acid sequence of calreticulin and the lumenal domain of calnexin show extensive similarity and the proteins appear to be functionally homologous (Zapun *et al.,*1999).

Calnexin and calreticulin have been defined as ER chaperones and both function as lectins, binding transiently to many newly synthesised glycoproteins in the ER. It has been shown that calnexin (Kearse *et al.,*1994; Peterson *et al.,*1995) and calreticulin (Nauseef *et al.,*1995; Ora and Helenius, 1995) have a strong preference for monoglucosylated proteins ($Glc1Man_{5\text{-}9}GlcNAc_2$). In practice, the binding of calnexin and calreticulin retains partially folded polypeptides in the ER, preventing both premature degradation and the transport of misfolded proteins through the secretory pathway. Calnexin and calreticulin associate with a variety of proteins including soluble secretory proteins, complex membrane receptors and ions channels (for review see Trombetta & Helenius 1998). Both calnexin and calreticulin can form a protein-folding complex with another ER chaperone, ERp57 (discussed in sections 1.6.1, 1.6.3 & 1.8). The interaction of these two complexes (calnexin/ERp57 and calreticulin/ERp57) with glycoprotein folding intermediates is functionally relevant, since the presence of these complexes significantly increases the efficiency of glycoprotein folding (Zapun *et al.,* 1999).

1.5.6 UDP-glucose: glycoprotein glucosyltransferase

UDP-glucose: glycoprotein glucosyltransferase (UGGT) is a large, soluble enzyme present in the ER lumenal (reviewed in Parodi, 2000). UGGT has a conserved region of 300 amino acids in the carboxyl terminus of its catalytic domain, which is homologous to glycosyltransferases (Tessier *et al.,*2000). UGGT acts as a folding sensor in the calnexin/calreticulin cycle (see Figure 1.5). Thus, if a glycoprotein is recognised as misfolded by UGGT it functions to add back the innermost glucose residue onto the high mannose oligosaccharides, regenerating monoglucosylated glycans, see Figure 1.5 (Parodi *et al.,*1983; Fernandez *et al.,*1998) and thereby leading to the re-association of the glycoprotein with the ER lectins calnexin and calreticulin and re-entry in to the 'calnexin cycle' (Hammond *et al.,*1994a).

The exact mechanism by which UGGT distinguishes between folded and non-folded glycoproteins is not known. It is likely that it resembles a 'classical' molecular chaperone, as it recognises a large variety of unrelated glycoproteins and probably recognises features shared by incompletely folded proteins. It is known that UGGT does not recognise glycoproteins in a random coil conformation, but instead efficiently recognises and reglucosylates a variety of partially folded conformers (Trombetta and Helenius, 2000). Where glycoproteins have multiple domains, UGGT selectively recognises glycans in the misfolded domains (Ritter and Helenius, 2000).

1.6 Calnexin and Calreticulin

Two resident ER lectins, membrane-bound calnexin and its soluble homologue, calreticulin retain improperly folded glycoproteins and facilitate their correct folding. As discussed previously in Section 1.5.5 they recognise folding glycoproteins by their monoglucosylated oligosaccharides. The calnexin/calreticulin folding cycle (Section 1.6.1 below), the lectins structures (section 1.6.2) and their substrate binding are discussed below (Section 1.6.3).

1.6.1 The role of N-linked oligosaccharides and the calnexin/ calreticulin cycle

N-linked glycosylation is highly conserved from yeast to mammalian cells and the ER is the initiation site for this process. As proteins are transported through the ER translocon, N-linked glycans are added to the peptide chain by the oligosaccharyltransferase, which is associated with the translocation machinery (Nikonov *et al.,* 2002). The oligosaccharyl-transferase complex attaches a 14-residue oligosaccharide ($Glc3Man_{5-9}GlcNAc_2$) to the side chain of asparagine residues in the consensus sequence Asn-X-Ser/Thr. This modification allows newly synthesised glycoproteins to interact with the ER lectin-based chaperone system (Figure 1.5; reviewed in Trombetta & Helenius, 1998; Parodi, 2000). However, proteins that fold rapidly in the ER are less prone to this modification (Holst *et al.,*1996), suggesting that a steric block arising from secondary structure can occlude the consensus sequence. The position of the glycans in the sequence can influence which of the many chaperones in the ER interacts with a particular precursor (Molinari and Helenius, 1999a). If glycoproteins contain an N-linked glycan within the first ~50 residues from the amino terminus they interact co-translationally with calreticulin and calnexin. If

the glycan appears later in the sequence then the initial chaperone interaction occurs with BiP.

The core N-linked glycan originally attached to a newly synthesised protein contains a string of three glucose residues (see Figure 1.6). Glucosidase I removes the terminal α(1,2)-linked glucose and glucosidase II subsequently removes the two remaining α(1,3)-linked glucoses rapidly. A lumenal enzyme called UDP-Glc: glycoprotein glucosyl-transferase (UGGT; see Section 1.5.6), which is the folding sensor in the calnexin/calreticulin cycle can add back the innermost glucose residue onto the oligosaccharides. This regenerates monoglucosylated glycans 'tagging' unfolded proteins (Parodi *et al.,*1983; Fernandez *et al.,*1998) leading to their re-association with the ER lectins and prolonging their window of opportunity for achieving a native structure (Hammond *et al.,*1994a). If the glycoprotein is folded correctly, it is not reglucosylated by UGGT and hence it is released from the calnexin cycle (Trombetta and Helenius, 1998).

Calnexin and calreticulin have been found to associate with their substrates regardless of whether they are correctly folded or not, so long as their N-linked glycans remain monoglucosylated. Hence, when glucosidase II activity is inhibited, the release of substrates from calreticulin or calnexin is also inhibited, consistent with the need for glucose removal to enable the release of glycoprotein substrates (Hebert *et al.,*1995). The behaviour of ERp57 is identical under these conditions consistent with it functioning in concert with calnexin and calreticulin (Oliver *et al.,* 1997). The deglucosylation and reglucosylation cycle is also supported by studies in living cells, including those on the T cell receptor α and β subunits (Gardner and Kearse, 1999).

Whilst the ER is well accepted as a location for the quality control of newly synthesised membrane and secretory proteins, components such as calreticulin, glucosidase

II and UGGT have also been found in the ER-Golgi intermediate compartment suggesting this may also be a site at which quality control occurs (Zuber *et al.,* 2001).

Figure 1.5 Model for the lectin-assisted cycle of glycoprotein folding in the ER.

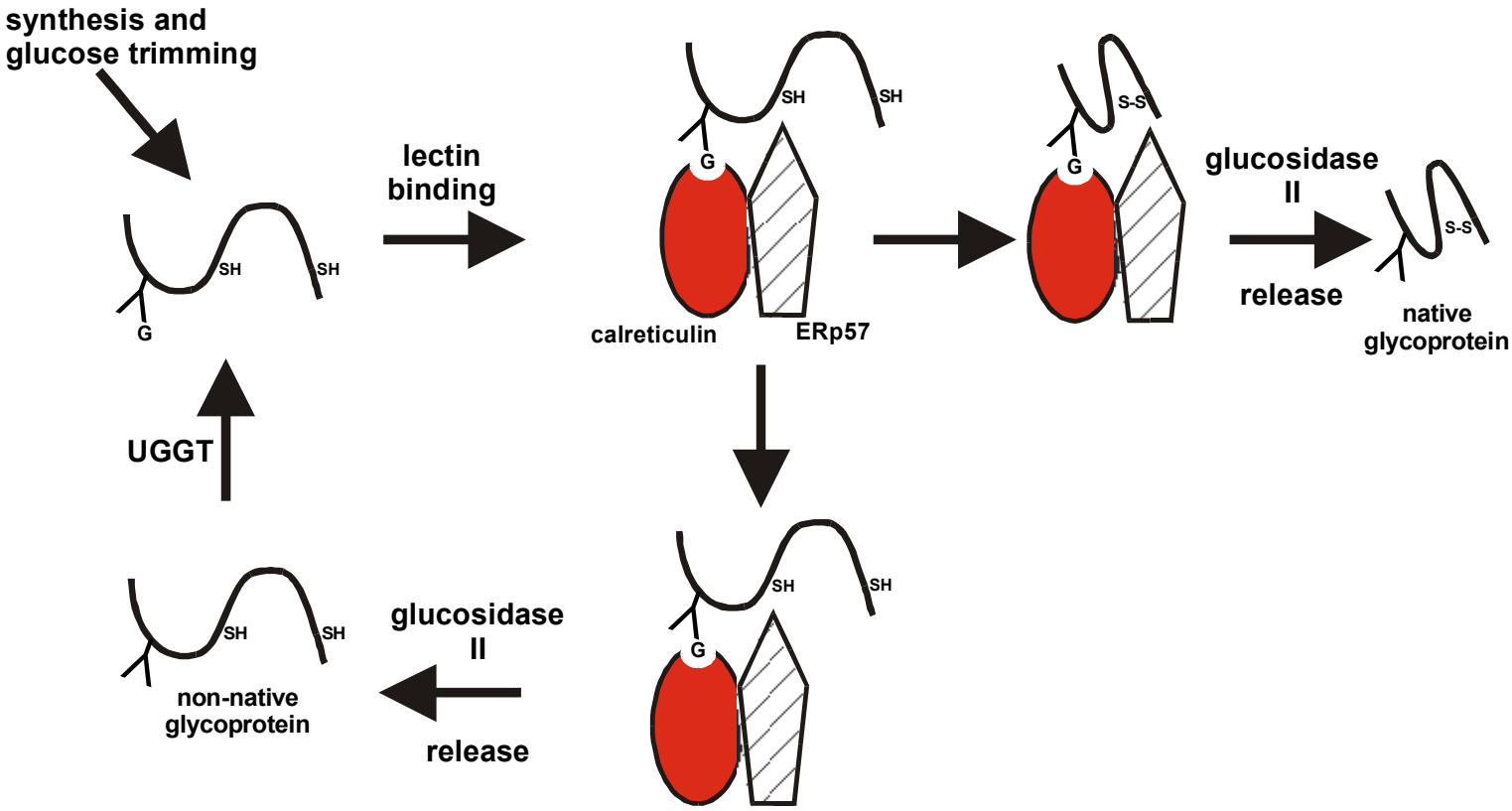


Figure 1.5 Model for glycoproteins folding in the ER (adapted from Oliver *et al.,* 1999). During synthesis and translocation into the ER lumen polypeptides are N-glycosylated. The oligosaccharide is then trimmed by glucosidase I and II. Calreticulin/ERp57 or calnexin/ERp57 complexes then preferentially bind monoglucosylated glycoproteins. The lectin interaction with the folding intermediates is prevented by the removal of the last glucose residue. UDP-glucose: glycoprotein glucosyltransferase reglucosylates misfolded glycoproteins, allowing the lectins to bind again. The deglucosylation and reglucosylation cycle continues until the glycoprotein is correctly folded and is no longer reglucosylated.

Figure 1.6 Structure of the N-linked oligosaccharide.

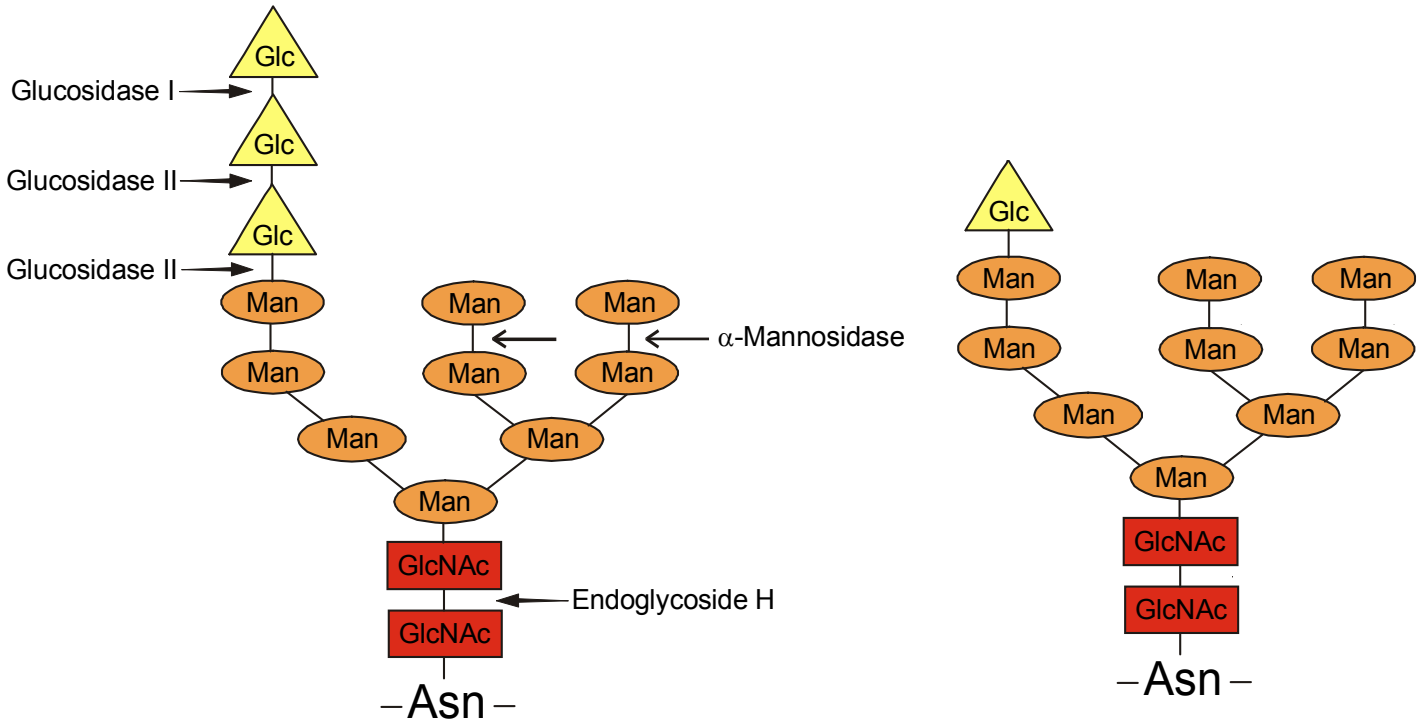


Figure 1.6 Structure of N-linked oligosaccharide (adapted from Trombetta & Helenius, 1998). a) Composition and structure of the oligosaccharide added to the proteins in the ER and structure of monoglucosylated form recognised by the ER lectins calnexin and calreticulin. b) Processing intermediate recognised by calnexin and calreticulin. Two glucoses have been removed and the single glucose residue is essential for recognition by calnexin and calreticulin. Structure

High resolution three-dimensional structures of the calreticulin P-domain, a region characterised by a series of short proline repeats (see Figure 1.7; Ellgaard *et al.,* 2001) and the calnexin ectodomain, including its P-domain (Schrag *et al.,* 2001; see Figure 1.8) have recently been solved. They show that the P-domain of both proteins comprises an unusual, extended hairpin fold. The crystal structure of the calnexin ectodomain shows the P-domain as a long arm protruding from a compact globular lectin domain (Figure 1.8). The P-domain forms a slightly curved arm of about 110Å in calreticulin (Ellgaard *et al.,*2001) and 140Å in calnexin (Schrag *et al.,*2001). The P domain of calreticulin comprises residues 189-288 and functions as an independently folding domain.

Given the close sequence similarity of calnexin and calreticulin, and the structural resemblance between their two P-domains, it is likely that the lectin binding domains of the two protein are also structurally similar. In calnexin, a single binding site for the oligosaccharide of the glycoprotein is located close to the site from which the P-domain emerges as an elongated loop (Schrag *et al.,* 2001; see Figure 1.10).

1.6.2 Substrate binding

Calreticulin and calnexin are functionally homologous and both bind monoglucosylated glycoproteins *in vitro* (Vassilakos *et al.,*1998). However, they only exhibit partially overlapping substrate specificities *in vivo* (Peterson *et al.,*1995; Van Leeuwen and Kearse, 1996). Additionally when calreticulin and calnexin bind to the same protein they have been shown to interact with different glycans in some cases (Hebert *et al.,*1995). It has been suggested that the difference in substrate specificity between calreticulin and calnexin is due in part to the fact calnexin is a membrane protein and calreticulin is a soluble protein (Peterson *et al.,*1995; Hebert *et al.,*1995).

Figure 1.7 Schematic representation of calnexin and calreticulin and the NMR structure of the P domain of calreticulin.

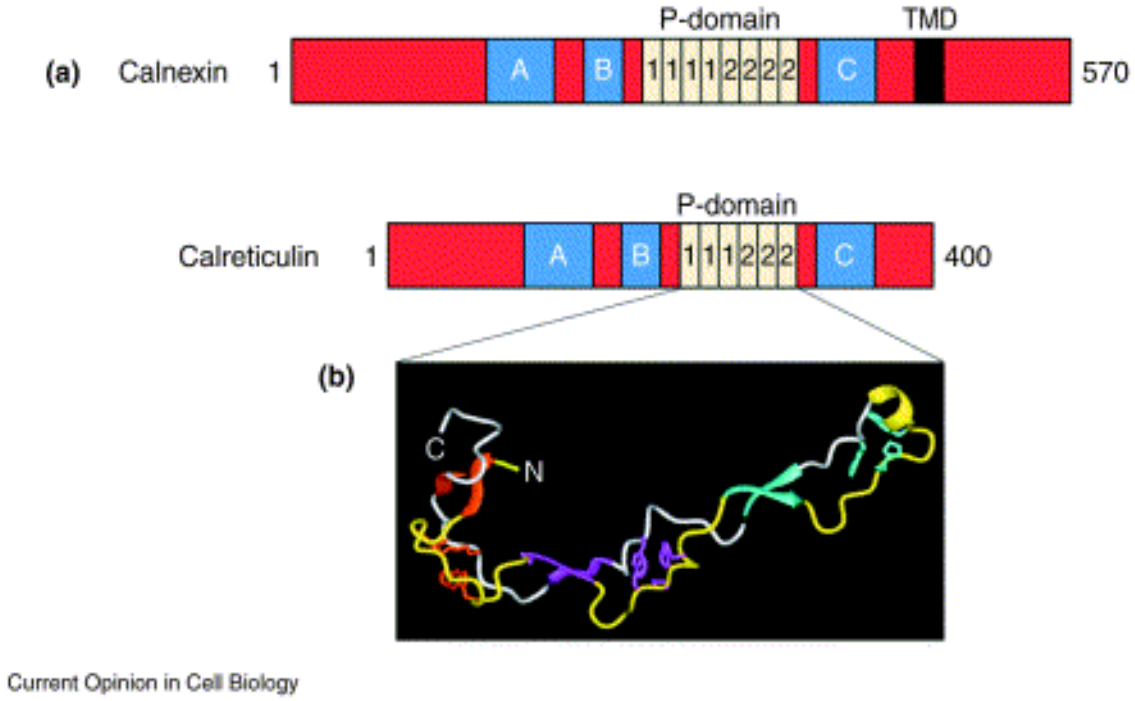


Figure 1.7 (a) Schematic representation of calnexin and calreticulin (adapted from Ellgaard and Helenius, 2001). Calnexin's transmembrane domain is depicted in black. Regions A, B and C show 50-55% sequence identity (David *et al.,*1993). The central P-domain contains two sequence repeat types, designated type 1 and 2, each repeated four times in calnexin and three times in calreticulin. (b) A cartoon of the calreticulin P-domain structure derived by NMR is shown (from Ellgaard *et al*., 2001). Type 1 repeats are indicated in yellow and type 2 repeats in white. The three β sheets and an α-helical turn are drawn as ribbons. Residues of three hydrophobic clusters are drawn as stick models.

Figure 1.8 Structure of the Lumenal Domain of Calnexin.

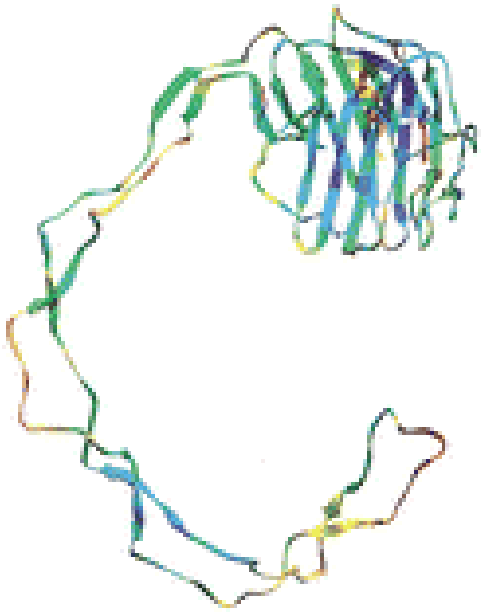


Figure 1.8 Ribbon diagram showing the concave β sheet of the lectin domain in front and the convex β sheet in back, which is a proteinase K-resistant fragment of calnexin (residues 61-458; taken from Schrag *et al.,* 2001). The molecule has two very distinctive regions: a compact globular domain (residues 61-262, 415-458) composed of a β sandwich of two anti-parallel β sheets, and a 145-residue long arm (residues 270-414) stretching 140 Å away from the globular domain.

Studies with various glycoproteins have revealed that calnexin and calreticulin interact with overlapping, but distinct sets of glycoprotein substrates. Studies in L cells, with calnexin being expressed in a soluble form revealed that its pattern of glycoprotein substrates resembled that of calreticulin (Danilczyk *et al.,*2000), whilst a membrane-bound version of calreticulin interacted with a similar set of glycoproteins as calnexin. Therefore, an important factor in the distinct substrate specificities of calnexin and calreticulin is their different topological environments (Danilczyk *et al.,* 2000).

There are currently two models for the mechanism by which calnexin and calreticulin enhance glycoprotein folding. One is the lectin only model (Hammond et al., 1994a; Hebert, 1995). In this model they function solely as lectins with cycles of binding and release regulated by availability of terminal glucose on monoglucosylated oligosaccharide. The lectin function is thought to be sufficient to prevent aggregation, and the ER lectins may also recruit other chaperones and folding enzymes to the unfolded glycoprotein, for example ERp57 (Oliver et al., 1997; Elliott et al., 1997; Zapun et al., 1997). *In vitro* studies showing complexes of RNase B and calnexin can be dissociated by oligosaccharide modification, and that calnexin does not appear to discriminate between reduced and native forms of the RNase B support the lectin only model (Rodan *et al.,*1996; Zapun *et al.,*1997). The alternative is the dual binding model (Ware, 1995; Williams, 1995). This model incorporates features of the lectin-only model, but also proposes that the ER lectins possess a second site that binds to polypeptide segments of unfolded glycoproteins. In this model, the release of the glycoprotein would require both a conformational change in the polypeptide-binding site and then the removal of the final glucose residue by glucosidase II. A key feature of the dual binding model is that both UGGT and calnexin function as folding sensors. Both calnexin and calreticulin have recently been found to bind to unfolded, non-glycosylated proteins but not to their native conformers *in vitro* (Ihara *et al.,* 1999; Saito *et al.,* 1999). More compellingly, it has been shown that calnexin can form complexes with diverse non-glycosylated proteins *in vivo* when mild isolation procedures are used (Danilczyk and Williams, 2001).

It has been reported that PDI can also interact with calreticulin, albeit at very low Ca^{++} concentrations (<100μM) (Corbett *et al.,*1999). However, the Ca^{++} concentrations used during this study are unlikely to be representative of those present in the ER, since the steady state Ca^{++} concentration is reported to be 500-800μM. Furthermore, *in vitro* cross-

linking studies revealed no association between calreticulin and PDI in intact endoplasmic reticulum-derived microsomes, although a strong association of calreticulin with ERp57 was observed (Oliver *et al.,*1999). The interaction between calreticulin and ERp57 has also been investigated using biochemical techniques, including ELISA, isothermal titration microcalorimetry (ITC), and NMR spectroscopy and it was found that ERp57 binds to the P-domain of calreticulin (Frickel *et al.,*2002). The binding of ERp57 to the distal end of the P-domain could generate a partially solvent-shielded cavity surrounded by the lectin domain, the P-domain, and ERp57 (Figure 1.9 below). The P-domain structure indicates a degree of plasticity (Ellgaard *et al.,* 2001) suggesting that the size of such a cavity could vary in order to accommodate substrates of different sizes.

Calnexin may also play a distinct glycan independent role in the quality control of membrane proteins and facilitate their assembly and/or quality control at the ER (Cannon and Cresswell, 2001; Swanton *et al.,* 2003, manuscript in press).

Figure 1.9 Scheme for a possible mode of cooperative action between Calreticulin and ERp57.

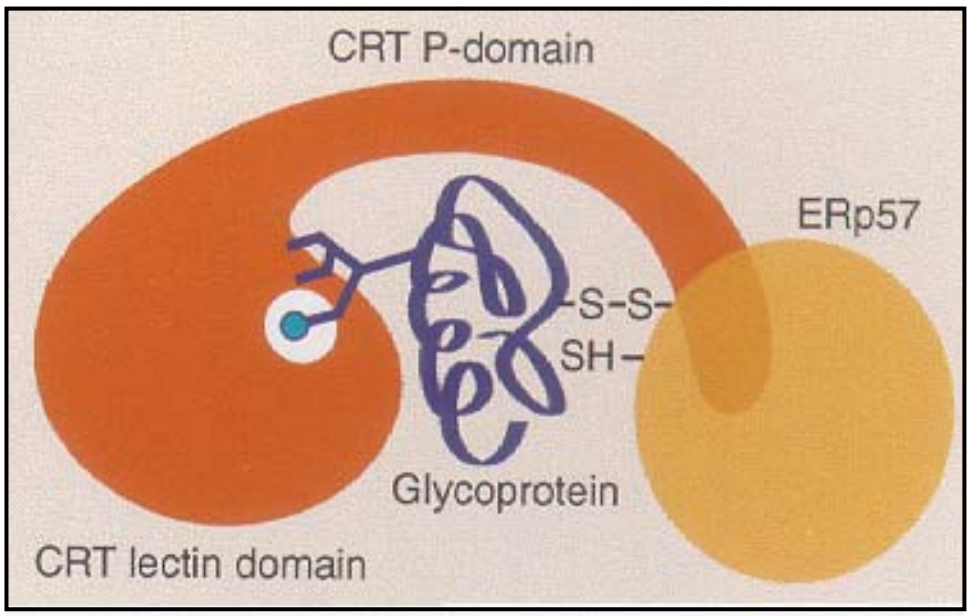


Figure 1.9 Scheme for a possible mode of cooperative action between calreticulin and ERp57 in assisting the folding of a glycoprotein (adapted from Frickel et al., 2002). The model is based on currently available structural, biochemical and cell-biological data (Parodi, 2000; Ellgaard *et al.,*2001; Schrag *et al.,*2001). The substrate glycoprotein (blue) binds to the calreticulin (CRT) lectin domain (red) by means of a branched oligosaccharide, which is monoglucosylated. The interaction of the terminal glucose (light blue circle) with the binding site (white circle) would place the glycoprotein polypeptide chain in the partially solvent-shielded cavity bounded by the calreticulin lectin domain, the P domain, and ERp57 (orange) bound to the distal end of the protruding P domain. It is further hypothesised that the CRT-ERp57 interaction places the thiol-disulphide oxidoreductase so that it favours the formation of intermolecular disulphide bonds (S-S) with the glycoprotein.

1.7 Protein disulphide isomerase (PDI)

PDI was the first folding enzyme to be identified and it catalyses the rate-limiting reactions of disulphide-bond formation, isomerisation and reduction within the ER (Goldberger *et al.,* 1963; Freedman *et al.,* 1994). In addition, PDI has two other well-established functions as a component of both prolyl-4-hydroxylase (P4H; Kivirikko *et al.,* 1989) and the triacyl glycerol transfer protein (MTP; Wetterau *et al.,* 1991). P4H catalyses the formation of 4-hydroxyprolyl residues in nascent collagen-like polypeptides and MTP facilitates the incorporation of lipids into newly synthesised core lipoproteins. The major features of PDI's primary structure are conserved across eukaryotes (for reviews see Freedman *et al.,* 1994; Ferrari and Söling, 1999; Kemmink *et al.,* 1997).

1.7.1 PDI and its structure

Since the first cDNA encoding PDI was sequenced, internal homologies within the protein sequence have been recognised and a multi-domain architecture proposed (Edman *et al.,* 1985; Darby *et al.,* 1998). Recent experimental data, particularly the proteolysis of native PDI and the characterisation of recombinant fragments, combined with bioinformatic analysis has lead to the current structural model of PDI (Kemmink et al., 1997; Freedman et al., 1998b). This model is comprised of four structural domains, ***a*, *b*, *b'*** and ***a'***, plus a linker region (*x*) between the ***b'*** and ***a'*** regions and a C-terminal acidic extension (Figure 1.10). The ***a*** and ***a'*** domains of PDI are structurally homologous to a thiol-disulphide oxidoreductase called thioredoxin (Martin, 1995). The two intervening segments, ***b*** and ***b'*,** are structurally homologous to each other and also contain a fold that is similar to that of thioredoxin, although there is no obvious sequence homology. Finally, the C-terminal

extension is a putative Ca^{2+}-binding region and is rich in acidic amino acids (Koivunen *et al.,*1999). The deletion of this '**c** domain' has no clear effect on the chaperone and disulphide isomerase functions of PDI (Koivunen *et al.,*1999).

Figure 1.10 The domain structure of human PDI.

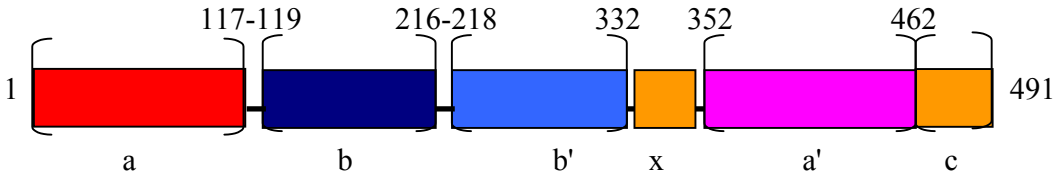


Figure 1.10 The domain structure of human PDI (adapted from Freedman *et al.,* 2002). The first and last residues of each known 3D structure are indicated. The boundaries presented are based on NMR-derived structural information for domains ***a***, ***b***, and ***a'*** (Kemmink *et al.,* 1996; Kemmink *et al.*, 1997; Dijkstra *et al.,* 1999), homology modelling of domain ***b'*** and data on limited proteolysis of intact PDI (Freedman *et al.,*1998b). The 'c domain' at the carboxyl terminus comprises a 24 residue acidic segment, rich in glutamate and aspartate. This is followed by the KDEL sequence (Lys-Asp-Glu-Leu) that is required for the retention of PDI within the ER. The ***x*** region is a 'linker' of undefined structure and function.

1.7.2 PDI domains

NMR studies on the recombinant ***a*** domain of human PDI have confirmed that it is a structural domain with a thioredoxin fold (Figure 1.11, panel A; Kemmink *et al.,* 1996). Thioredoxin is a small, ubiquitous protein of ~12kDa, involved in many cytoplasmic redox functions including that of a protein disulphide reductase (Martin, 1995). The single domain members of the thioredoxin superfamily (thioredoxins and glutaredoxins) all have the same characteristic α/β fold, with a βαβαβαββα structure. The five-stranded β-sheet surrounded by four α helixes forms the central core, with all of the strands parallel except

β4 (Figure 1.11). This core structure can be found in other members of the thioredoxin superfamily, including DsbA (a bacterial periplasmic oxidase) and members of the PDI family. The active site motif WCxxC is found between β2 and α2 in an exposed region of polypeptide chain. This site is redox active, it converts between the dithiol and disulphide forms and is involved in thiol-sulphide exchange reactions (Figure 1.14).

The ***a*** domain of PDI shares other common features with thioredoxin. Firstly, the active-site motif is located at the N-terminus of helix α2, which is distorted by a proline residue. Secondly, the peptide bond before the proline residue at the N-terminus of helix β4 is in the *cis* form. Thirdly, there is a buried acidic residue, Glu30, in an analogous position in the ***a*** domain of PDI to that of Asp26 in thioredoxin. This residue has been postulated to play a role in generating the redox properties of thioredoxin (Martin, 1995). The high-resolution structure of the ***a'*** domain of human PDI has not yet been derived, but preliminary NMR data and secondary structure assignment have confirmed that the overall fold of the ***a'*** domain is similar to that of the ***a*** domain (Dijkstra *et al.,*1999). The ***b*** and ***b'*** domains of PDI show significant sequence homology to each other, but not the ***a*** domain or thioredoxin. However, NMR analysis has revealed that the ***b*** domain also forms a domain with a thioredoxin fold (Kemmink *et al.,*1997). The ***b*** domain does not contain a thioredoxin-like active site, and other residues associated with redox properties have been replaced. As there is significant sequence similarity between the ***b*** and ***b'*** domains, this implies that the ***b'*** domain most likely also contains a thioredoxin fold.

Figure 1.11 Structures of the isolated ***a*** and ***b*** domains of human PDI.

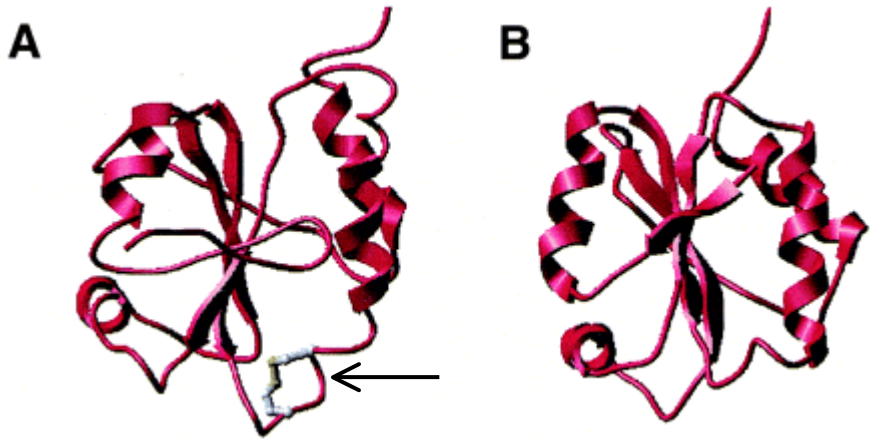


Figure 1.11 Structures of the isolated ***a*** (A) and ***b*** (B) domains of human PDI as determined by NMR (adapted from Kemmink *et al.,* 1996, Kemmink *et al.,* 1997). Both domains show a characteristic thioredoxin fold, despite a low degree of sequence similarity. The active-sites cysteines in the ***a*** domain are shown in grey and yellow (indicated by ←).

1.7.3 The function of PDI and its domains

Redox assays using simple peptide substrates show that the isolated ***a*** and ***a'*** domains of PDI have properties enabling them to act both as reductases or oxidases depending on the substrate and redox environment (Zapun *et al.,*1993; Darby and Creighton, 1995a; Darby and Creighton, 1995b). However, the activity of PDI is not just the sum of the activity of its two active domains (***a*** and ***a'***), suggesting that other parts of PDI are required for its full range of activities. For example, complex disulphide isomerisations in the folding pathway of bovine pancreatic trypsin inhibitor (BPTI) can be catalysed by full length PDI (Darby *et al.,*1998), but the isolated ***a*** and ***a'*** domains are not able to do so (Darby and Creighton, 1995c).

The contribution of individual domains to the activities of PDI was analysed using a variety of PDI constructs containing combinations of PDI domains (Darby and Creighton, 1995c; Darby *et al.,*1998). These constructs were used in assays studying oxidation, reduction and isomerisation. These experiments showed that simple thiol-

disulphide chemistry only requires the ***a*** or ***a'*** domains and simple isomerisation requires one these active domains in combination with the ***b'*** domain. However, complex isomerisation requires all four PDI domain (***abb'a'***), but not the C-terminal extension. The role of PDI domains was further explored using the chemical cross-linking of peptides and unfolded polypeptides to various PDI derivatives (Klappa *et al.,*1998). The ***b'*** domain was found to be essential for PDI-protein substrate interactions and sufficient for the binding of small peptides, although other domains contributed to the association of PDI with large peptides and partially unfolded proteins (Klappa *et al.,*1998). Hence, the ***b'*** domain of PDI appears to be the principle peptide binding site, but other domains are required for interactions with larger substrates. In contrast, both the ***b'*** and ***a'*** domains of PDI are the minimum requirement for PDI to function as a subunit of prolyl 4-hydroxylase (Pirneskoski *et al.,*2001).

1.8 ERp57

ERp57 (also known as GRP58, ERp60, ERp61, ER60, HIP-70, Q2 and P58) is an ER folding factor (reviewed in High *et al.,* 2000; see Figure 1.12). Several mammalian ERp57 cDNAs have been identified, and the encoding amino acid sequences all share significant homology with PDI (Bourdi *et al.,*1995). The overall amino acid sequence identity and similarity between the signal sequence processed forms of human ERp57 and PDI are 29% and 56% respectively (Koivunen *et al.,*1996). Most domains in the human and bovine ERp57 (Figure 1.13) or *Drosophila* ERp57 (Koivunen *et al.,* 1996; Table 1.2) are similar to the equivalent regions of PDI, the only exception being the C-terminal extension. ERp57 also contains two thioredoxin-like WCGHCK motifs that are identical to those found in PDI and ERp72 (Freedman *et al.,*1994).

Figure 1.12 The proposed domain structure of human ERp57.

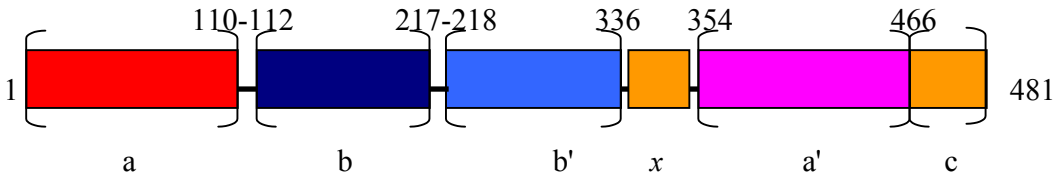


Figure 1.12 The domain structure of human ERp57. The first and last residues of each proposed structural domain are indicated. The boundaries are based on homology modelling with human PDI and data on limited proteolysis of intact ERp57 (personal communication, Dr Lloyd Ruddock). The '***c***-domain' at the carboxyl terminus comprises a basic segment. This is followed by the KDEL-like sequence (Lys-Asp-Glu-Leu), which is required for the retention of ERp57 in the ER. The ***x*** region is a linker of undefined structure and function.

Table 1.2 Homology comparison of human PDI and ERp57.

Domain(s)	Homology (% identity)
abb'a'	**33**
a	**49.4**
b	**23.1**
b'	**14.4**
a'	**53.6**
c	**-**

Table 1.2 Comparison of amino acid sequence identities between various domains of ERp57 and PDI (taken from Koivunen *et al.,* 1996).

Figure 1.13 Comparison of human and bovine ERp57 and PDI C-terminal sequences.

```
ER57BO SPYEVRGFPT IYFSPANKKQ NPKKYEGGRE LSDFISYLK. ..... ..REATNPPV
ER57HU SPYEVRGFPT IYFSPANKKL NPKKYEGGRE LSDFISYLQ. ..... ..REATNPPV
PDIHUM A.VKVHSFPT LKFFPASADR TVIDYNGERT LDGFKKFLES GGQDGAGDDD DLEDLEEAEE
PDIBOV A.VKVHSFPT LKFFPASADR TVIDYNGERT LDGFKKFLES GGQDGAGDDD DLEDLEEAEE

ER57BO IQEEKPKKKK KAQEDL
ER57HU IQEEKPKKKK KAQEDL
PDIHUM PDMEEDDDQK AVKDEL
PDIBOV PDLEEDDDQK AVKDEL
```

Figure 1.13 The QEDL and KDEL sequences at the extreme C-termini of the proteins are involved in their ER retention and retrieval. The basic residues that are unique to this region of ERp57 are underlined. The equivalent region of PDI is markedly acidic in character.

The first ERp57 cDNA was originally reported to encode for phosphatidylinositol-specific phospholipase-C-α (Bennet *et al.,*1988). The cDNA was later shown to encode for a glucose-regulated protein that had a PDI-like amino acid sequence and a thiol-dependent reductase activity *in vitro* (Bourdi *et al.,*1995; Hirano *et al.,*1995), indicating that it may influence protein folding. ERp57 has also been proposed to be a cysteine-dependent protease (Urade and Kito, 1992), a carnitine palmitoyltransferase (Murthy and Pande, 1993), and a hormone-induced protein of the brain (HIP-70; Mobbs *et al.,*1990a; Mobbs *et al.,*1990b).

More recent studies of ERp57 have shown that it is specifically involved in glycoprotein folding (reviewed in High et al., 2000). ERp57 is not a lectin *per se*, since it does not bind to oligosaccharides (Zapun *et al.,* 1998), rather it forms a discrete complex with one or another of the two ER lectins calnexin or calreticulin, to generate a glycoprotein specific assembly (Oliver *et al.,*1999). As would be expected for such a

complex, the interaction between ERp57 and newly synthesised glycoproteins shows the same requirement for glucose trimming of the N-linked glycan as the binding of calnexin and calreticulin (Oliver *et al.,*1997). The interaction of ERp57 with newly synthesised glycoproteins was further investigated in semi-permeabilised mammalian cells, a system in which the ER remains structurally 'intact' and therefore most closely resembles that present in a living cell (Van der Wal *et al.,* 1998). The association of ERp57 with glycoproteins was found to be transient, a characteristic of molecular chaperones, and the release of ERp57 from the glycoprotein was also found to require glucose trimming (Van der Wal *et al.,* 1998). Hence, both the binding and release of ERp57 are regulated by glucose trimming, and ERp57 was proposed to be a generic component of the glycoprotein specific folding machinery operating inside the ER (Oliver *et al.,*1997).

The interaction of ERp57 with either calnexin or calreticulin has been shown to greatly increase its disulphide-isomerase activity with respect to the refolding of a mono-glucosylated protein substrate (ribonuclease B). In contrast, the activity of PDI was reduced by the addition of these ER lectins in the same assay (Zapun *et al.,* 1998). This study provided direct evidence that ERp57, in combination with the ER lectins calreticulin and calnexin, can directly enhance the efficiency of glycoprotein folding and therefore act as a chaperone. Additionally, both ERp57 and PDI have been shown to interact directly with two viral glycoproteins in living cells and form transient disulphide bonded complexes with the newly synthesised proteins (Molinari and Helenius, 1999b). 'Chase' experiment revealed changes and heterogeneity in the complexes formed and these were consistent with the reshuffling of disulphide bonds in the folding glycoprotein (Molinari and Helenius, 1999b).

Studies into the assembly of the class I major histocompatability complex (MHC I) have shown that ERp57 also functions as part of the MHC class I peptide-loading complex

in the ER (Hughes and Cresswell, 1998). This complex contains the transporter associated with antigen processing (TAP), calnexin, calreticulin, tapasin, MHC class I, β_2m and ERp57 (Hughes and Cresswell, 1998; Lindquist *et al.,*1998; Morrice and Powis, 1998). ERp57 has been reported to interact at both an early (Morrice and Powis, 1998) and late stage of MHC class I biogenesis (Lindquist *et al.,*1998), and this has been confirmed using a synchronised system to exploit semi-permeabilised mammalian cells (Farmery *et al.,*2000). A role for ERp57 during the assembly of the MHC class I-peptide complex is entirely consistent with the well established roles for calnexin and calreticulin during this process (reviewed in Lehner and Trowsdale, 1998), and supports the view that ERp57 functions as a molecular chaperone. The specific role of ERp57 during MHC class I assembly remains to be determined, but recent studies have shown that it can form a mixed disulphide with another component of the loading complex, tapasin. Hence, one role for ERp57 may be the isomerisation of disulphide bonds within tapasin during the peptide loading of the MHC class I complex (Dick *et al.,*2002). Where MHC class I complexes are not peptide loaded, the complexes are normally retained at the ER and subsequently degraded via a process referred to as quality control (Section 1.10). In the case of the MHC class I complex, ERp57 may act to reduce intra-chain disulphide bonds to enable the unfolding of MHC class I molecules targeted for degradation (Antoniou *et al.,*2002). Taken together, the data outlined above indicate that ERp57 can function as a molecular chaperone that can promote glycoprotein folding, complex assembly and quality control.

1.9 Formation of disulphide bonds

Protein folding, as assayed by disulphide bond formation, occurs far more rapidly *in vivo* than *in vitro* suggesting that protein oxidation must be catalysed within living cells (Anfinsen, 1973). In eukaryotic cells, the formation of protein disulphide bonds in the ER lumen requires protein disulphide isomerases (reviewed in Freedman, 1994). Both ERp57 and PDI have been shown to interact directly with two viral glycoproteins in living cells and form transient disulphide bonded complexes with the newly synthesised proteins (Molinari and Helenius, 1999b).

The ratio of reduced and oxidised glutathione (GSH and GSSG, respectively) varies between the cytoplasm and the compartments of the secretory pathway, with the cytoplasm having a ratio of GSH: GSSG of ~100:1 and the secretory pathway having a ratio of ~3:1 (Hwang *et al.,*1992). Hence, the ER lumen is relatively oxidising, compared to the predominantly reducing environment of the cytosol, and this favours the formation of disulphide bonds. Since glutathione ratios differ, the presence of this gradient would in principle be sufficient to maintain the ER lumen as an environment favourable for protein oxidation. However, every disulphide bond formed within a nascent chain introduces reducing equivalents into the ER, and these must be disposed of in order to maintain the environment in an oxidised state in the ER. It was first proposed that oxidised glutathione may serve as the initial electron acceptor from PDI during the process of disulphide bond formation *in vivo* (Hwang *et al.,* 1992). However, it has now been shown that oxidised glutathione is not required for oxidative protein folding in the ER (Cuozzo and Kaiser, 1999).

Factors that are required for effective oxidative protein folding within the ER have been identified in *S. cerevisiae* (Frand and Kaiser, 1998; Pollard *et al.,*1998). One gene

product *ER* *O*xidoreductase 1 (ERO1; Figure 1.14) is specifically required for oxidative protein folding (for reviews see Freedman et al., 1998a; Frand *et al.,* 2000). Mutations in this gene result in a folding defect for three substrates containing disulphide bonds, while proteins that do not depend on disulphide formation for their folding are secreted normally (Frand & Kaiser, 1998, Pollard *et al.,* 1998). ERO1 activity also determines the overall redox balance within the cell, and its importance is underlined by the fact that homologues can be found in all eukaryotic organisms examined (Frand and Kaiser, 1998; Pollard *et al.,* 1998; Cabibbo *et al.*, 2000).

Figure 1.14 Proposed model for the formation of protein disulphide bonds in the ER Lumen.

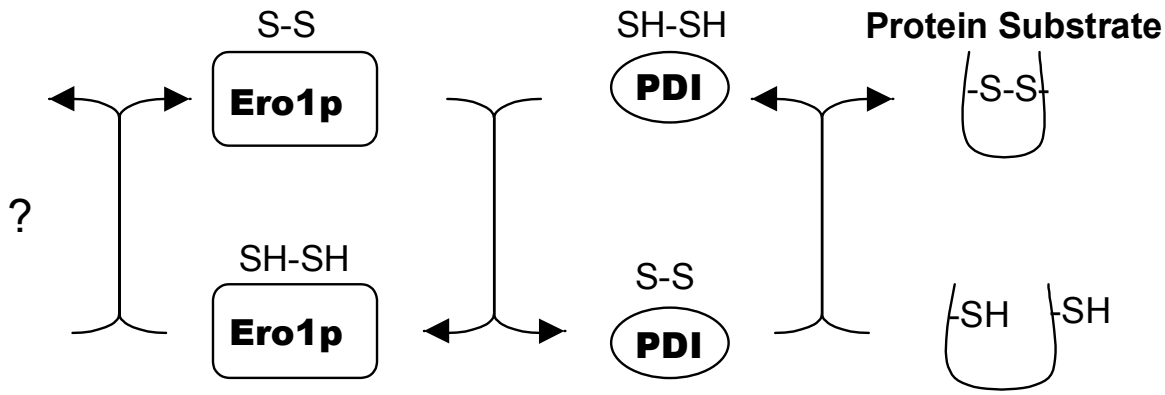


Figure 1.14 Model for protein disulphide bond formation (adapted from Frand and Kaiser, 1999). A dithiol-disulphide exchange occurs between Ero1p and PDI, resulting in PDI being oxidised. PDI then goes on to oxidise proteins in the ER, through dithiol-disulphide exchange, and create disulphide bonds. Only one active site of PDI is shown. The mechanism by which oxidised Ero1p is regenerated remains to be established.

PDI is normally found in an oxidised state *in vivo*, and mutations in *ERO1* result in the steady-state accumulation of PDI in its reduced form (Frand and Kaiser, 1999). Conversely, mutations in the active-site cysteines of PDI allow the isolation of stable mixed-disulphide complexes between Ero1p and PDI (Frand and Kaiser, 1999; Tu *et al.,*2000), and it is via this complex that oxidising equivalents are probably introduced to substrate proteins (Figure 1.14; Frand and Kaiser, 1999). Two human members of the ERO1 family, Ero1-Lα and Ero1-Lβ, have been identified. Only Ero1-Lβ is induced during the unfolded protein response (Pagani *et al.,* 2000), nevertheless, both genes complement the *ero1-1* yeast mutant (Cabibbo *et al.,*2000; Pagani *et al.,*2000). Ero1-Lα can be isolated in mixed-disulphides with PDI from mammalian cells (Benham *et al.,*2000), and both Ero1-Lα and Ero1-Lβ can facilitate disulphide bond formation in immunoglobulin subunits by selectively oxidising PDI (Mezghrani *et al.,*2001). An *in vitro* oxidative folding system has recently been developed in which PDI is a required for the Erop1-dependent oxidation of RNase A (Tu *et al.,* 2000). In contrast to PDI, the mechanism by which ERp57 is reoxidised is unknown, and no interaction between ERp57 and Ero1L-α or Ero1-Lβ has been detected to date (Mezghrani *et al.,*2001).

An important question is how oxidising equivalents are introduced into the ER. As all redox reactions involve the movement of electrons from one molecule to another, there must be a 'sink' for the excess of electrons generated within the ER. An oxidation system analogous to that in the ER of eukaryotic cells is found in the periplasm of *E. coli* (Rietsch and Beckwith, 1998). In *E. coli*, electrons pass through the respiratory chain (Kobayashi *et al.,*1997), ultimately being delivered to oxygen (Bader *et al.,*1999). Unlike this bacterial oxidation system, the one operating in *S. cerevisiae* is not affected by the inactivation of components of the respiratory chain (Tu *et al.,*2000), though oxygen is the preferred terminal acceptor (Tu *et al.,* 2002). Instead, oxidative folding appears to depend on

cellular levels of free flavine adenine dinucleotide (FAD; Tu *et al.,* 2000). A role for FAD was confirmed *in vitro,* since FAD was required for the oxidation of RNase A by Erop1 (Tu *et al.,* 2000). Screening in *S. cerevisiae* identified ERv2p, a flavoenzyme that can catalyse O_2-dependent formation of disulphide bonds, indicating that FAD does also play a role *in vivo* (Sevier *et al.,* 2001).

FAD may act as a co-factor for Erop1, in which case another protein could bind to Erop1 and catalyse the reaction $FADH_2 \rightarrow FAD$ ($FADH_2$ being generated during the oxidation of PDI). Alternatively, $FADH_2$ may be released and Erop1 may bind an oxidised molecule of FAD. In either case something must reoxide $FADH_2$ to FAD, even under anaerobic conditions, suggesting that another oxidising source must be present in the ER. Such a system has been found in the oxidising environment of the bacterial periplasm where fumarate acts as the terminal electron acceptor under anaerobic conditions (Bader *et al.,* 1999).

1.10 Quality Control at the ER

Protein folding in living cells is a complex, and potentially error-prone process. Inside the ER, proteins destined for secretion or transport to the cell surface are folded and glycosylated. Many of the processes mediated by the ER are essential for cell viability, and any disturbance of ER function presents significant stress to the cell. Furthermore, the cell employs numerous mechanisms to ensure that only correctly folded proteins are produced (reviewed in Ellgaard *et al.,* 1999; Wickner *et al.,* 1999; Cabral *et al.,* 2001; Ellgaard and Helenius, 2001). In the context of the ER, this process is normally referred to as 'quality control' and results in the retention of misfolded or misassembled proteins at the ER. Retention is most commonly achieved via the association of the misfolded proteins with components that are themselves normally resident in the ER (reviewed in Ellgaard *et al.,*1999). Chaperones are particularly good candidates for such retention factors, since they recognise residues/epitopes that are exposed on misfolded or misassembled proteins. Hence, the calnexin/calreticulin system (Zhang *et al.,*1997), BiP (Hurtley and Helenius, 1989; Hammond and Helenius, 1994b), and PDI (Puig and Gilbert, 1994) have all been implicated in several aspects of quality control at the ER.

When proteins fail to attain their correct conformation, a degradation mechanism can act to remove them from the ER (McCracken and Brodsky, 1996). Should both the chaperone and protease systems fail, aggregation of the misfolded proteins is the likely outcome (Wickner *et al.,*1999). Finally, signal transduction mechanisms are specifically activated when the cell is under stress and these can upregulate the production of factors involved in the control and regulation of protein folding (Chapman *et al.,*1998). These processes are generally known as ER Associated Degradation (ERAD) and the Unfolded

Protein Response (UPR) and they are discussed in more detail in sections 1.10.2 and 1.10.3.

1.10.1 Molecular Chaperones: Co-operation and Redundancy

A number of studies have revealed co-operation between different ER components giving rise to additional functions. For instance, BiP interacts with Sec63p during protein translocation (for review see Brodsky, 1996), and calreticulin recruits ERp57 to promote glycoprotein folding and disulphide bond formation (Oliver *et al.,*1997). It has been hypothesised that all resident ER proteins may form a dynamic matrix and interact loosely with each other (Liu *et al.,*1999). This network has been speculated to function like a chromatographic matrix adsorbing newly synthesised proteins, until they become correctly folded and thereby lose their affinity. Transmembrane protein, such as calnexin, could keep such a matrix associated with the translocation site at the rough ER membrane (Liu *et al.,*1999).

As previously discussed, most glycoproteins transiently interact with the lectins, calreticulin and calnexin during their maturation. However, it is clear that many glycoproteins do not require these associations to attain their correct conformation (Helenius and Aebi, 2001). This may result from several factors, firstly, glycans may be essential to the folding of some glycoproteins, but they may not be required for many others. Secondly, the ER contains many chaperones that assist protein folding. This provides a versatile system where if one chaperone is inactive, another one may substitute for its function. For example, when ERp57 binding is inhibited, the association of N-glycosylated protein precursors with both PDI and PDIp is enhanced (Elliott *et al.,*1998). This overlapping specificity and functional redundancy would facilitate the role of the ER in producing a diverse and complex population of proteins.

1.10.2 ERAD (ER-Associated Protein Degradation)

The existence of an intracellular protein degradation process, referred to as ER-associated protein degradation (ERAD), that removes aberrant polypeptides and unassembled subunits of multimeric protein complexes from the endoplasmic reticulum is now well established (for reviews see McCracken *et al.,* 1996, Ivessa *et al.,* 1999 and Fewell *et al.,* 2001). ERAD acts as an essential step in the quality control of newly synthesised proteins and is involved in the degradation of both soluble and integral membrane proteins present in the ER, thus ensuring that only functional molecules are deployed to the distal sites of the secretory pathway. ERAD involves three key steps i) recognition of the aberrant polypeptide; ii) export of soluble proteins back through the translocation pore ('retrotranslocation') back to the cytoplasm; and iii) degradation by the 26S proteosome.

Many of the same molecular chaperones involved in folding proteins in the ER are probably involved in the recognition and export of the aberrant polypeptides. Aberrant secretory proteins may have exposed structural motifs (Schmitz *et al.,*1995; Skowronek *et al.,*1998) or hydrophobic patches, which are buried when the protein is correctly folded and assembled. These structural motifs could prolong their interactions with chaperones and trigger their export from the ER and destruction. Aggregation of misfolded proteins prior to their export would preclude/prevent their retrotranslocation through the ER translocon. In yeast the chaperones BiP (Plemper *et al.,*1997; Chapman *et al.,*1998), calnexin (McCracken and Brodsky, 1996), and PDI (Gillece *et al.,*1999) are required for the degradation of some aberrant proteins.

A role for BiP in degradation is suggested by the observation that the enhanced binding of BiP to Ig light chain chimeras is associated with their accelerated degradation (Brodsky *et al.,*1999). BiP has also been shown to bind a mutant prion protein and mediate

its degradation (Jin *et al.,*2000). Additionally, defects in calnexin can compromise the degradation process (McCracken and Brodsky, 1996) and mammalian calnexin has been shown to prevent the aggregation of unfolded glycoproteins *in vitro* (Ihara *et al.,*1999). In mammalian cells the proteosomal degradation of α_1-anti-trypsin (AAT) requires a physical interaction with calnexin (Liu *et al.,* 1999). PDI also plays some role in ERAD, and the ERAD activity in yeast lacking PDI was restored by the expression of any one of four other PDI homologues (Niwa *et al.,*1999).

One issue that has been raised with respect to ERAD is whether there is a signal that distinguishes between proteins that are folding slowly and those that are terminally misfolded? A scheme proposing the removal of mannose from the oligosaccharide side-chain together with the association of calnexin as directing the misfolded AAT for degradation by the proteosome in mammalian cells, has been put forward (Liu *et al.,* 1999). This agrees with previous work showing that processing by ER mannosidases decreases the extent of glycoprotein glucosylation (Sousa *et al.,*1992). Further studies also suggest that extensive mannose trimming and/or the retention of glucose residues, due to the lack of glucose trimming, are signals for ER/proteosomal degradation. Some studies indicate this effect is independent of an interaction with calnexin (Ayalon-Soffer *et al.,* 1999; Chillaron *et al.,* 2000; Chung *et al.,* 2000; Tokunaga *et al.,* 2000) , though others studies indicate further associations with calnexin may be required (Liu *et al.,*1999; de Virgilio *et al.,*1999; Wilson *et al.,*2000). However, since not all proteins are N-glycosylated, there must presumably be other mechanisms present in the ER to detect terminally misfolded proteins.

The presence of a lectin, which targets proteins with an attached $Glc_{0-3}Man_8GlcNAc_2$ for degradation has been identified in *S.cerevisiae* (Jakob *et al.,* 2001, Nakatsukasa *et al.,* 2001) and in mouse cells (Hosokawa *et al.,* 2001). The two proteins

are α1,2-mannosidase-likeprotein, homologous to one another, and they lack enzymatic mannosidase activity, which suggests they are mannose-binding lectins (Figure 1.15). Recently it has been shown that overexpression of EDEM in mammalian cells results in quicker release of misfolded glycoproteins from calnexin (Molinari *et al.,* 2003; Oda *et al.,* 2003). On the other hand, down-regulation of EDEM leads to a delay in glycoprotein degradation (Molinari *et al.,* 2003). Additionally, proteosome inhibitors only mildly inhibit degradation induced by EDEM overexpression, suggesting the existence of additional degradation pathways (Hosokawa *et al.*, 2001).

The export of soluble ERAD substrates from the ER to the cytosol occurs via the Sec61p translocon, and since Sec61p is required for both ER import and ER export it seems likely that these two processes are related (Brodsky *et al.,*1999). However, the precise roles of BiP and Ssa1p (a yeast Hsp70 cytosolic protein) during ER import and export differ, indicating that these two processes are mechanistically distinct (Brodsky *et al.,*1999). Several studies suggest that BiP may deliver misfolded proteins to the Sec61 pore (Schmitz *et al.,* 1995; Knittler *et al.,* 1995; Skowronek *et al.,* 1998; Brodsky *et al.,* 1999), whilst PDI has also been proposed to play a role in targeting an ERAD substrate to BiP at the translocon (Gillece *et al.,*1999).

Like soluble proteins, misfolded membrane proteins may also exit the ER via the Sec61 pore, and they can be co-precipitated with a component of the mammalian ER translocon (Wiertz *et al.,*1996; Bebök *et al.,*1998; de Virgilio *et al.,*1998). Other transmembrane proteins may be directly extracted by the proteosome or attacked by different proteases (reviewed in Fewell *et al.,* 2001). It has recently been reported that the proteosome may be able to clip polypeptide loops of the cytoplasmic portions of integral ER membrane proteins (Hoppe *et al.,*2000). The resulting transmembrane domains might

be unstable and spontaneously dissociate to enter the cytoplasm or could be further cleaved (Weihofen *et al.,*2000).

Figure 1.15 Proposed Pathway of Quality Control

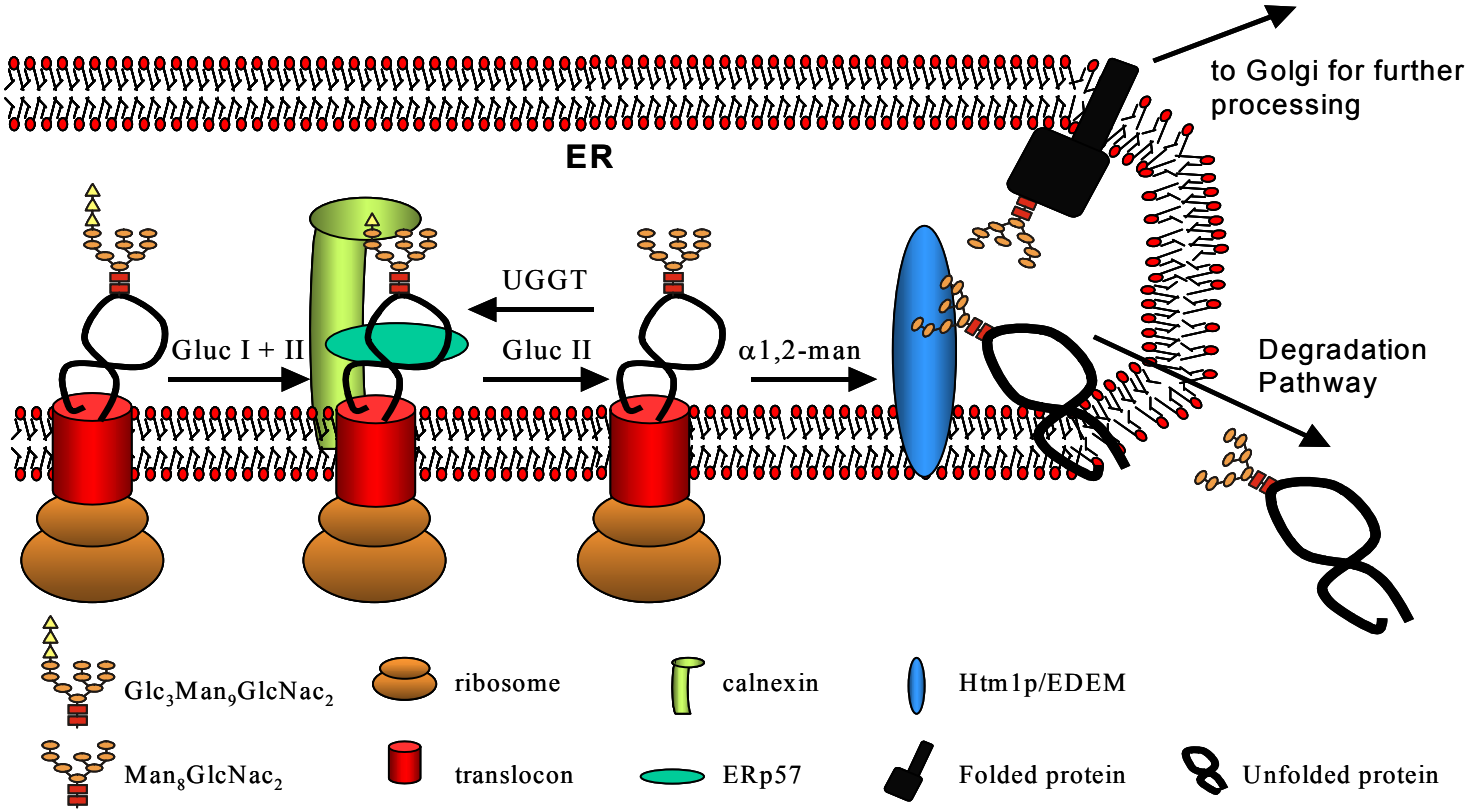


Figure 1.15 Proposed Pathway of Quality Control (adapted from Braakman 2001). Following translocation into the ER lumen N-linked glycans are added to asparagine residues on proteins. The first glucose is removed by glucosidase I, the second and third by glucosidase II. Monoglucosylated protein enter the calnexin/calreticulin folding cycle, which retains the immature glycoproteins in the ER allowing correct folding to occur. The third glucose can be reattached by glucosyltransferase, which recognises misfolded/unfolded proteins. The ER α1,2-mannosidase removes a single mannose and the Man_8 glycan, containing a yet unknown number of glucose residues, is recognised by Htm1p/EDEM and targeted for degradation.

Multiple studies indicate that ERAD substrates are degraded in the cytoplasm by the proteosome (Nishikawa *et al.,*2001). The proteosome consists of a catalytic 20S cylindrical core particle and two copies of the 19S regulatory particle that 'caps' the 20S subunit. Ubiquitination is necessary for proteosomal processing of most, but not all ERAD substrates. ERAD substrates must be de-glycosylated, de-ubiquitinated and unfolded before they can interact with the small aperture at the tip of the catalytic core of the proteosome (Wiertz *et al.,*1996; Voges *et al.,*1999; Suzuki *et al.,*2000). In mammalian cells the delivery of misfolded proteins from cytosolic chaperones to the proteosome may be mediated by a putative E3 ubiquitin ligase, CHIP (Ballinger *et al.,*1999; Connell *et al.,*2001; Meacham *et al.,*2001). It has also been suggested that Hsc70-bound substrates can be targeted to the proteosome by the mammalian nucleotide exchange factor BAG-1 (Höhfeld and Jentsch, 1997; Luders *et al.,*2000).

1.10.3 The Unfolded Protein Response (UPR)

Signal transduction pathways to the nucleus can be activated in response to the stress caused by disturbances in protein processing in the ER. These pathways include a well defined unfolded protein response or UPR (reviewed in Sidrauski *et al.,* 1998; Pahl, 1999; Patil and Walter, 2001). The signalling results in the activation of specific sets of genes, which increase the transcription of a variety of folding factors including BiP, PDI and ERp57 (Kaufman, 1999). The response is however complex, and in *S. cervisiae* nearly 400 genes have been identified as targets of the UPR (Travers *et al.,* 2000). The UPR is activated by the presence of mis- or unfolded proteins in the ER and is the primary mechanism by which eukaryotic cells counteract the accumulation of misfolded proteins. The UPR was initially identified in mammalian cells following glucose starvation, which

results in protein misfolding through the under glycosylation of nascent polypeptides and an accompanying induction of a specific set of proteins (Chapman *et al.,*1998). The induced proteins, which are mainly molecular chaperones, were referred to as glucose regulated proteins (GRPs) and hence GRP78 was the original name given to BiP. A detailed description of the eukaryote UPR pathways is beyond the scope of this thesis, and several recent reviews describing this process are available (Chapman *et al.,* 1998; Foti *et al.,* 1999; Fewell *et al.,* 2001; Patil and Walter, 2001).

1.11 In vitro analysis of protein-protein interactions

The assembly of the PDI-like protein ERp57 with the ER lectins calnexin and calreticulin is central to the function of these components in modulating glycoprotein folding (Zapun *et al.,* 1998). Furthermore, the association of these components into specific complexes can be reproduced *in vitro* using radiolabelled precursor proteins that have been synthesised using a cell free system (Oliver *et al.,* 1999). My goal was to exploit such *in vitro* systems to identity the basis for the specific association of ERp57 with the ER lectins.

1.11.1 Cross-linking

A variety of cross-linking approaches can be used to detect protein-protein interactions that have been reconstituted using cell free systems (reviewed in Martoglio and Dobberstein, 1996). All cross-linking techniques identify the nearest neighbours of the protein of interest and, if a radioactive protein is cross-linked to an unlabelled component, the increase in the apparent size of the resulting adduct can be used to estimate the size of adjacent components. During the course of this study, bi-functional cross-linking reagents

were used to cross-link ER chaperones (ERp57, PDI, calreticulin etc.) to endogenous ER components with which they were associated following the import of the radiolabelled ER chaperone precursor into dog pancreatic microsomes or the ER of semi-permeabilised mammalian cells.

The cross-linking reagents used in this study were a hetero-bifunctional cross-linking reagent specific for thiol and amino groups, succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate (SMCC), and a thiol-specific homo-bifunctional reagent bismaleimidohexane (BMH; Table 1.3). The succinimide moiety reacts with suitable amino groups whilst the maleimide group reacts with thiol groups. In practice maleimides will modify cysteine side chains and succinimides will modify lysine side chains and free amino-termini (Figure 1.16). Most importantly, the cross-linking reagents used during this study are membrane permeable and can readily cross the lipid bilayer that surrounds the ER and gain access to the lumen where the formation of the ER chaperone complexes occurs (cf. Oliver *et al.,* 1999).

Table 1.3 Cross-linking with bi-functional reagents

Reagent	Reactivity	Spacer Arm Length
BMH	-SH to –SH	~16.1 Å
SMCC	-SH to $–NH_2$	~11.6 Å

Table 1.3 Reactivity and solubility of the bifunctional chemical cross-linking reagents used in this study.

Figure 1.16 Structure and reactivity of chemical cross-linking reagents.

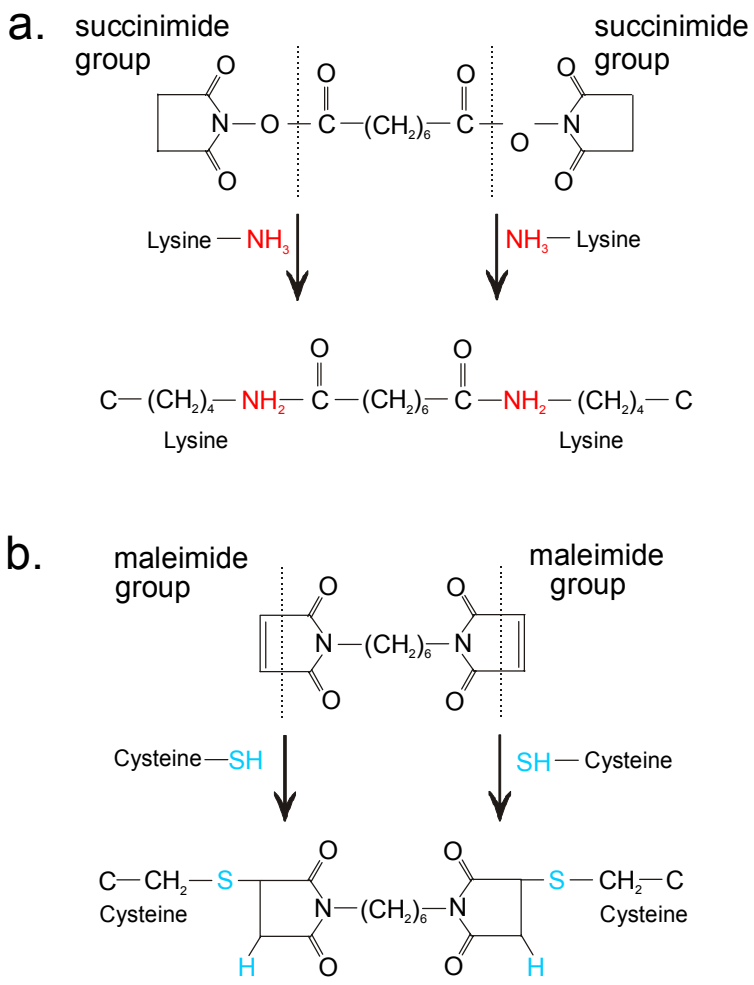


Figure 1.16 Structure and reactivity of, **a**, succinimide and **b**, maleimide chemical cross-linking reagents.

1.11.2 ERp57 Chimeras and point mutants

The modular structure of the PDI family (Figures 1.4, 1.10 and 1.12) has been used in the past to create a series of chimeras between PDI and ERp57 that have been used to identify the region of PDI that enables it to form a complex with the α subunit of prolyl 4 hydroxylase (Pirneskoski et al., 2001). During the course of this study, these chimeras plus additional ERp57 derivatives were used in comparison to wild-type ERp57 and PDI in order to identify the region of ERp57 that is responsible for its ability to form specific complexes with the ER lectins calnexin and calreticulin. Once a domain(s) of ERp57 was implicated in its specific binding to calreticulin and calnexin, site-directed mutagenesis was used to more closely define the molecular basis for this specific interaction.

1.11.3 'Pull-down' assays

During the course of this work, cross-linking was initially used to assay the ability of ERp57 derivatives to form complexes with calnexin and calreticulin. In order to study complex formation using an assay that would be unaffected by differences in the availability of amino acid side chains, and other factors that could reduce cross-linking efficiency, a pull-down assay for ERp57 binding was also developed. Pull-down assays have been widely used in the past (Updyke and Nicholson, 1984; Buckie and Cook, 1986; Gretch et al., 1987) and can be made both selective and stringent. Furthermore, in contrast to a cross-linking assay, it is generally accepted that a pull-down assay provides a true reflection of the strength of the interaction being studied. By analysing the interactions of ERp57 derivatives with calreticulin using both cross-linking and pull-down based assays I

have been able to provide detailed insights into the molecular basis for the specific assembly of these components within the ER lumen.

CHAPTER - 2

Materials & Methods

2 Materials & Methods

2.1 Materials

Epicurian Coli® sub-cloning grade *E coli*, the PCR-Script™ Amp Cloning kit, QuikChange™ Site-Directed Mutagenesis kit and cloned *Pfu* DNA polymerase were purchased from Stratagene Ltd (Cambridge). The ABI Prism BigDye Terminator™ Cycle Sequencing Ready Reaction Kit for DNA sequencing was supplied by PE Applied Biosystems (Warrington, Chesire). Restriction endonucleases were purchased from New England Biolabs, (Hitchin, Herts). T7 RNA polymerase, transcription buffers, rNTPs, RNAsin ribonuclease inhibitor, amino acids, RNAse A, Flexi® rabbit reticulocyte lysate lysate and the TNT® Coupled Reticulocyte System were supplied by Promega (Southampton, Hants). DC BioRad Protein Assay was purchased from BioRad laboratories, (Hemel Hempsted, Herts). Aurintricarboxylic acid (ATCA), cycloheximide, phenyl-methylsulfonyl fluoride (PMSF), Protease Inhibitor Cocktail (for bacterial cells) and TX-100 were purchased from Sigma (Poole, Dorset). Easytag L-[^{35}S]methionine was purchased from NEN Du Pont (Stevenage, Herts). All reagents for cell culture were obtained from GibcoBRL (Life Technologies) and digitonin was obtained from Calbiochem. The cross-linking reagents bismaleimidohexane (BMH) and succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate (SMCC) was supplied by Pierce & Warriner (Chester, Cheshire). Canine pancreatic microsomes were prepared essentially as described by Walter and Blobel (1983). All other chemicals were purchased from BDH/Merck (Poole, Dorset) and Sigma.

Rabbit anti-sera, raised against the C-terminal regions of mammalian PDI and ERp57 (High laboratory) were generated by Eurogentec. The re-immune serum used was from

one of the rabbits used to raise anti-sera against the C-terminus of mammalian PDI. A rabbit anti-ERp57 sera raised against the purified baboon protein was gifted by Dr Claude Bonfils (CNRS, Montpellier, France). The anti-calnexin serum raised against the C-terminus of the protein was a gift from Professor Ari Helenius (Swiss Federal Institute of Technology, Zürich, Switzerland). The polyclonal rabbit anti-calreticulin serum was obtained from Affinity Bioreagents (Cambridge Bioscience, Cambridge).

2.2 Constructs.

A full-length cDNA encoding human ERp57 in the vector pVL1392 and five PDI/ERp57 chimeras (constructs 2-6) were kindly provided by Professor Kari Kivirikko (University of Oulu, Finland). The ERp57 construct had previously been removed from the pVL1392 vector (designed for baculovirus expression) by *Eag*I and *Eco*RI digestion and ligated into the pBluescript II KS vector under the control of the T7 promoter (Oliver, J. et. al., 1999). PDI/ERp57 chimeric constructs 2-6 were similarly removed from the pVL1392 vector by restriction digestion and then ligated into either pBluescript KS II or pBluescript SK II under the control of the T7 promoter. PDI/ERp57 chimeric construct 7 was generated by two-step PCR and ligated into pPCRScript under control of the SP6 promoter (see Section 2.4 below). The ERp57Δc construct was generated by site directed mutagenesis (see Section 2.3 below). Human ERp57 sub-domain constructs ***abb'***, ***b'a'c***, ***bb'x*** and ***b'x***, and all thirteen ERp57 ***b'*** domain point mutants (V267A, V267W, Y269A, Y269W, E270A, K274A, R280A, N281A, V283A, V283L, V283W, F299A and F299W) in pBluescript II KS were provided by Dr Lloyd Ruddock (University of Oulu, Finland). Human PDI in pBluescript II SK(-) was kindly provided by Professor Neil Bulleid (University of Manchester, UK).

2.3 Site-Directed mutagenesis to construct ERp57∆c and ERp57∆ss.

All site-directed mutagenesis was performed using the Stratagene PCR-based Quikchange mutagenesis kit (Cat. No. 200518) according to the manufacturers' directions. This method requires two complementary primers that each incorporate the desired base change(s). The primers were annealed to the DNA template simultaneously in a single PCR reaction, and the entire plasmid amplified by *Pfu* DNA polymerase generating mutated copies that were also distinct in being non-methylated. The methylated parental template was then digested by the restriction enzyme *DpnI*, which is specific for the site -GA↓TC- where the adenine residue is methylated**,** leaving only non-methylated plasmid DNA which carried the mutation.

The QuikChange™ Site-Directed Mutagenesis kit (Stratagene) was used to remove the C-terminal region of ERp57. To construct ERp57Δc complementary primers were designed to introduce a stop codon in-frame after codon 465 following the manufacturers' recommendations. For ERp57Δss, complementary primers containing a ATG overhang on the forward primer were designed to introduce a start codon in-frame at the beginning of the ERp57 ***a*** domain before codon 25. A 5µl sample of the PCR reaction mix was then run on a 1% agarose TAE gel and a very high molecular weight product (above the 1Kb markers) could be seen in the mutagenised samples (data not shown). For transformation, 1µl of each of the *Dpn*I treated mutagenised DNA samples, or of a pUC18 plasmid 10ng/µl (a positive control for transformation) was added to 50µl of Epicurian Coli® sub-cloning grade cells according to the manufacturers' instructions.

All transformants were selected for by growth on LB-agar plates supplemented with ampicillin (1%w/v tryptone, 0.5% w/v yeast extract, 1% w/v NaCl, 1.5% w/v agar, 100μg/ml ampicillin). Single colonies were picked and grown in 3ml LB-ampicillin (100μg/ml) broth, and the plasmid DNA isolated using a Qiagen Spin column miniprep kit (Dorking, Surrey. Cat. No. 27104). The resulting plasmids were analysed by restriction digestion to remove the various ERp57/PDI encoding inserts. Clones containing the correct size of insert and vector were grown overnight in 250ml cultures and ultrapure pBluescript plasmid DNA containing the plasmid of interest was then isolated using a Qiagen Maxiprep column kit (Cat. No. 12123).

2.4 Two-Step PCR to construct PDIab ERp57b'a'c mutant.

The ERp57/PDI chimeric construct 7 PDI abERp57b'a'c was generated by a two-step PCR strategy. This was achieved by setting up two initial PCR reactions with *Pfu* polymerase, one to amplify the PDI***ab'*** fragment including the PDI signal sequence and the first 15 bases of the PDI ***b'*** domain. A second, separate PCR reaction to amplify the ERp57***b'a'c*** fragment, including the last 15 bases encoding the ERp57 ***b*** domain, was carried out. These fragments were then purified using Wizard PCR purification columns (Promega, Southampton). The purified fragments were used as a template in a second PCR reaction, where the PDI***ab*** forward primer and the ERp57***b'a'c*** reverse primers were used to amplify the PDI***ab***ERp57***b'a'c*** chimeric construct. The resulting blunt-ended product was then inserted into pPCRScript.

2.5 Cell culture and preparation of semi-intact cells.

Human HT-1080 fibrosarcoma cells (A.T.C.C.CCL121, American Type Culture Collection), were grown to about 90% confluence in 75cm^2 flasks with minimal essential medium (MEM) with Earle's salts, supplemented with 1/100 volume 200mM L-glutamine, 1/100 volume 100mM sodium pyruvate, 10% heat-inactivated foetal calf serum, 1/100 volume MEM non-essential amino acids and 7/500 volume MEM vitamins solution. Digitonin permeabilisation was essentially as described by Wilson *et al*., (1995), which is briefly as follows: cells were washed twice with phosphate-buffered saline (PBS) and detached from the flask by incubation with 2ml Trypsin-EDTA (0.5mg/ml trypsin; 0.2mg/ml EDTA) for 5 min. Further degradation by trypsin was inhibited by addition of 4ml KHM buffer (110mM KOAc, 2mM MgOAc, 20mM HEPES pH 7.2) containing 100μg/ml Soybean Trypsin Inhibitor, and the cells were transferred to a 15ml Falcon tube for collection by centrifugation at 250 x *g* for 3 min at 4˚C. The pellet was resuspended in 4ml KHM buffer and the cells permeabilised for 5 min on ice by addition of digitonin (40mg/ml stock in DMSO) to a final concentration of 40μg/ml. The digitonin was diluted by addition of 10ml ice-cold KHM buffer, and the cells pelleted immediately by centrifugation at 250 x *g* for 3 min at 4˚C. Following resuspension in 5ml of HEPES buffer (50mM KOAc, 90mM HEPES pH 7.2), the cells were incubated on ice for 10 min. After centrifugation again at 250 x *g* for 3 min at 4˚C cells were resuspended in 100μl KHM, transferred to a microfuge tube and spun for 10 seconds in a benchtop microcentrifuge. The cell pellet was resuspended in 100μl KHM buffer, $CaCl_2$ was added to 1mM and 1μl calcium-dependent micrococcal nuclease (1mg/ml; ~70 Units/mg) added to degrade endogenous RNA. After incubation for 12 min on ice, the nuclease was inactivated by chelation of the Ca^{2+} using EGTA at a final concentration of 4mM. Cells were isolated by spinning for 10 seconds in a microfuge, and finally resuspended in 100μl KHM buffer.

2.6 Transcription and Translation

2.6.1 In vitro Transcription

Purified human ERp57/PDI pBluescript II KS/SK plasmid DNA was linearised with either *Eco*RI (Construct 1) or *Bam*HI (Constructs 2-6). Construct 7, in pPCRScript was linerasied with *Eco*RI. Human ERp57, ΔERp57, ERp57 sub-domain constructs ***abb'***, ***b'a'c***, ***bb'x*** and ***b'x***, and all the ERp57 ***b'*** domain point mutants in pBluescript II KS were digested with *Eco*RI. All thirteen ERp57 ***b'*** domain point mutants (V267A, V267W, Y269A, Y269W, E270A, K274A, R280A, N281A, V283A, V283L, V283W, F299A and F299W), in pBluescript II KS were also digested with *Eco*RI. Human PDI in pBluescript II SK was linearised with *Hin*dIII.

Transcription reactions were in a 100μl total volume containing 10μg of linearised plasmid DNA as a template, 40mM Tris.HCl pH7.5, 10mM NaCl, 6mM $MgCl_2$, 2mM spermidine, 1mM ribonucleotide triphosphates, 10mM DTT, 1mM 7mGG cap analogue, 50units of RNAsin and 40 units of T7 or SP6 RNA polymerase. Transcription was allowed to proceed for 2 hours at 37°C before purification of the RNA using RNeasy®Mini Kit (Cat. No. 74104) and analysed on a 1% agarose TAE gel. The resulting mRNA was snap frozen in liquid nitrogen and stored at -80°C for use with Flexi® rabbit reticulocyte and nuclease treated lysates during *in vitro* translations (see section 2.6.2 below).

2.6.2 In vitro Translation

Translation was in a Flexi® rabbit reticulocyte lysate system (Promega). mRNA was heated 70°C for 3 minutes and then immediately cooled on ice to decrease secondary RNA structure. A standard reaction comprised 14µl lysate supplemented with 2µl RNA transcript, 15µCi L-[^{35}S]methionine (specific activity 1175 Ci/mmol), 20µM each amino acid except methionine and either 1.5-2µl of canine pancreatic microsomes at ~50 OD_{280}/ml, 4µl of semi-intact HT-1080 cells (equivalent to approx. 10^5 cells) or water where no membranes were used.

All translation mixes were incubated at 30°C for 1 hour. ACTA was added to a final concentration of 100µM to inhibit further initiation of translation and chain elongation was allowed to continue for 10 min at 30°C. Any incomplete nascent chains were detached from the ribosome prior to cross-linking, this was achieved by a 10 min incubation at 30°C with 2mM puromycin, a peptidyl-tRNA analogue which causes the termination of translation and release of the nascent chain from the ribosome.

2.6.3 In vitro Eukaryotic Coupled Transcription/Translation

Maxiprep plasmid DNA encoding ERp57/PDI Constructs 1 to 6 was used in a TNT coupled transcription and translation kit (Promega) based on the rabbit reticulocyte lysate system and T7 RNA polymerase. The reaction were carried out in the presence of L-[^{35}S]-methionine and canine pancreatic microsomes. A standard 25µl reaction comprised 12.5µl lysate supplemented with 250ng closed circular plasmid DNA, 15µCi L-[^{35}S]methionine (specific activity 1175 Ci/mmol), 20µM each amino acid except methionine, 40units T7 PNA polymerase, 50 units RNAsin and 1.5-2µl of canine pancreatic microsomes at ~50 OD_{280}/ml. ACTA was added to a final concentration of 100µM to inhibit further initiation of translation and chain elongation was allowed to continue for 10 min at 30°C. Nascent chains were detached from the ribosome prior to cross-linking, this was achieved by a 10 min incubation at 30°C with 2mM puromycin, a peptidyl-tRNA analogue which causes termination of translation and release of the nascent chain from the ribosome.

2.7 Cross-linking optimisation and Sample Analysis.

Following translation, the membrane-associated protein fraction was isolated by centrifugation through a high salt/sucrose cushion (250mM sucrose, 500mM potassium acetate, 5mM magnesium acetate, 50mM Hepes.KOH, pH7.9) at ~100,000 x g for 10 minutes at 4^{o}C. For this purpose, a maximum of 200µl reaction mix was layered onto the 800µl cushion in a pre-chilled TLA 100.2 tube. Following centrifugation the supernatant and cushion were removed and the membrane pellet resuspended in the equivalent starting translation volume of low salt/sucrose buffer (250mM sucrose, 100mM potassium acetate, 5mM magnesium acetate, 50mM Hepes.KOH, pH7.9). Semi-intact cells were purified

from the translation mix by washing twice in KHM buffer. A wash comprised collecting the cells by centrifugation for 10 seconds at full speed (15000 *g*) and resuspension in one volume ice-cold KHM buffer. Cells were finally resuspended into 20µl KHM buffer per 25µl initial translation reaction.

The cross-linking reagents were prepared as 20mM stock solutions in DMSO and final concentrations of 10µM to 1mM were tested. Following optimisation tests a final concentration of 0.5mM BMH and 250µM SMCC were used for cross-linking. For cross-linking, the reagent was added to the resuspended membranes or semi-intact cells and incubated at 30°C for 10 min before being quenched by addition of either 5mM β-mercaptoethanol (to BMH) or 5mM β-mercaptoethanol and 100mM glycine (to SMCC), for 10 minutes at 30^{o}C. Control reactions with no cross-linking reagent were incubated with the cross-linker solvent, DMSO before quenching.

2.8 Immunoprecipitation of cross-linking products.

Cross-linking products were further analysed by immunoprecipitation experiments that were carried out in parallel on the DMSO treated control reactions. Following cross-linking or an appropriate control, samples were denatured by adding 1% SDS and four volumes of Triton immunoprecipitation (IP) buffer (10mM Tris-HCl, pH7.6, 140mM NaCl, 1mM EDTA, 1% Triton X-100) and heating for 10 minutes at 70^{o}C. The diluted mixture was then left on ice for 15 minutes with intermittent vortexing. Once 'solubilisation' was complete the sample was spun in a microfuge for 1 minute ~12,000rpm to remove any debris and the resulting supernatant removed. To reduce any background from unincorporated ^{35}S methionine, methionine (1mM) was added to the rest of the sample. The protease inhibitor phenylmethylsulphonyl fluride (PMSF; 0.2mg/ml)

was also included. Following centrifugation, supernatant aliquots of 100µl or 200µl were incubated with a variety of anti-sera and the samples were mixed overnight at 4^{o}C. The next day 30µl of Protein A or a 50:50 mix of Protein A and G Sepharose bead slurry (Zymed Laboratories, San Francisco, CA) was added to each aliquot and the sample incubated for a further 2 hours. The Protein A and G Sepharose beads had been pre-incubated with 20% bovine serum albumin and washed 5 times in IP buffer. Following the 2 hour incubation, the Protein A Sepharose beads were pelleted by centrifugation in a microfuge and the beads were washed five times using 1ml of cold Triton IP buffer. Residual IP buffer was removed with a Hamilton syringe and SDS PAGE samples buffer added (see Section 2.11).

2.9 Biotin-tagged Calreticulin Production Binding Assay

The human calreticulin cDNA coding for the mature part of the protein had previously been sub-cloned into the Nru I site of the PinPiont Xa I vector (Promega) downstream of the biotin purification tag. *E.*coli containing the biotin-tagged calreticulin plasmid were grown on LB-agar plates supplemented with ampicillin and glucose (1%w/v tryptone, 0.5% w/v yeast extract, 1% w/v NaCl, 1.5% w/v agar, 100µg/ml ampicillin, 2% glucose). A 50 ml culture was set up containing 1%w/v tryptone, 0.5% w/v yeast extract, 1% w/v NaCl, 100µg/ml ampicillin, 2% glucose and 2mM biotin). At an O.D. of 0.4-0.6 fusion protein expression was induced with IPTG (0.1M) for 2 hours. Cells were harvested and 50ml Column Buffer buffer (20mM HEPES, 1mM EDTA, 50mM NaCl, pH7.5) and 1ml of protease inhibitor cocktail (AEBSF 0.7mM, EDTA 3.4mM, Bestatin 0.07mM, Pepstatin A 0.1mM and E-64 0.009mM in DMSO) was added per 20ml of pelleted cells, as as described by the manufacturer (Sigma, Poole, Dorset). Samples were then frozen

overnight at -20^{o}C to aid cell lysis and broken with a French press. The resulting cell lysate was captured using SoftLink avidin resin as described by the manufacturers (Promega). Biotin-tagged recombinant protein was eluted from this resin with elution buffer (50mM Tris Base, 50mM NaCL, 5% glycerol, 1mM EDTA and 5mM or 20mMBiotin) and exchanged into column buffer (20mM HEPES, 1mM EDTA, 50mM NaCl, pH7.5) by dialysis.

Samples of eluted fractions were tested by Western blotting with an anti-Strepavidin HRP antibody (1:5000 dilution). Fractions containing significant amounts of biotin-tagged calreticulin were pooled and the concentration of biotin-tagged calreticulin calculated using the DC BioRad Protein Assay as described by the manufacturer (BioRad laboratories, Hemel Hempsted, Herts).

2.10 Biotin-tagged Calreticulin Binding Assay

NeutrAvdin beads were pre-incubated with 20% bovine serum albumin and washed 5 times in TBS (50mM Tris-HCL, pH 7.5, 150mM NaCl). Biotin-tagged calreticulin (5μg) or biotin-tagged insulin (~0.6μg which is an equivalent molar quantity) were subsequently immobilised on UltraLink NeutrAvidin beads (Pierce) by mixing overnight at 4°C in the presence in the of 100μM PMSF and 0.25% TX-100 in TBS (total sample volume 150μl/20μl beads). Control reactions included a ten fold molar excess of 'native' or scrambled' RNase A (~1.4μg).

Following the overnight incubation, the beads were pelleted by centrifugation for 2 minutes at 3,000 rpm in a microfuge and the beads were washed 3 times using 1ml of 0.25% TX-100 in TBS. mRNAs were translated in the absence of microsomes as described above (Section 2.6.2) and briefly spun in the bench microfuge (1000xg) for 2

minutes. 10µl of each translation product was then mixed with either 20µl NeutrAvidin beads alone or 20µl NeurAvidin beads with biotin-tagged calreticulin or biotin-tagged insulin bound. Competition experiments with ^{35}S radiolabelled ERp57 were also carried out in the presence of a 10 fold molar excess of his-tagged calreticulin (25µg) or PDI (~30µg) with biotin-calreticulin loaded beads. Samples were incubated for 2 hours, rolling at room temperature, in presence of 100µM PMSF and 0.25% TX-100 in TBS (50µl/20µl beads). Following a 2 hour mixing at room temperature the beads were washed 4 times in 0.25% TX-100 in TBS and then once with TBS.

2.11 Sample Analysis

Samples were denatured and solubilised in SDS sample buffer by heating to 70°C for 10 minutes before being resolved on SDS-PAGE gels of between 8 and 15% acrylamide as indicated in the figure legend. Gels were fixed in acetic acid:methanol:water solution (1:2:7 v/v), dried and exposed for 3 days to a phosphorimaging plate for visualisation on a Fujix BAS-2000 bioimaging system. Quantification of the phosphorimages was then carried out using AIDA software supplied by Fuji.

CHAPTER - 3

Results

Cross-linking analysis of ERp57/PDI chimeras

3 Cross-linking analysis of ERp57/PDI chimeras

The formation of disulphide bonds in the ER lumen is an important post-translational modification and these bonds serve to stabilise the native conformation of many secretory proteins. It is well known that PDI plays a critical role in the correct formation of disulphide bonds in the ER, yet the ER also contains several other PDI-like proteins that are less well characterised (Ferrari and Söling, 1999). One such protein is ERp57, also referred to in the literature as GRP58, ERp60, ERp61, ER60, HIP-70, Q2 and P58 (Elliott *et al.,* 1997).

It has previously been demonstrated that ERp57 interacts specifically with both secretory and integral membrane glycoproteins in association with the ER lectins calnexin and calreticulin (Oliver *et al.,* 1997; Elliott *et al.,* 1997). However, ERp57 is not a lectin *per se*, since it does not bind to oligosaccharides (Zapun *et al.,* 1998). Rather, it forms discrete complexes with the ER lectins calnexin and calreticulin, and it is these complexes that interact specifically with newly synthesised glycoproteins (Oliver *et al.,* 1999). ERp57 therefore functions as a generic subunit of a glycoprotein specific folding machinery operating in the ER (Oliver *et al.,* 1997).

During the course of this study, the interaction of ERp57 with calnexin and calreticulin has been investigated at the molecular level by using a number of ERp57/PDI chimeric proteins and other ERp57 derivatives. The initial cross-linking studies that are described in this chapter used both homo- and hetero-bifunctional reagents and were aimed at identifying the domain(s) of ERp57 required for its specific interaction with the ER lectins calreticulin and calnexin.

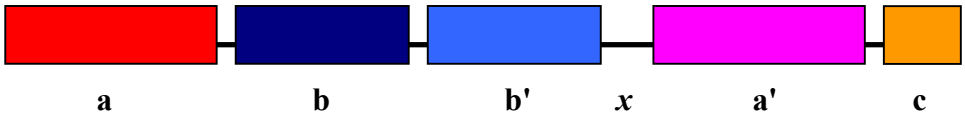
3.1 The Constructs

Six chimeric constructs (Constructs 2-7) containing various domains of ERp57 and PDI were used (Figure 3.1). Five of these constructs had previously been exploited to identify the region of PDI that is required to form an active $\alpha_2\beta_2$ prolyl 4 hydroxylase complex (Pirneskoski *et al.*, 2001). The amino acids derived from each protein that contribute to the various constructs are indicated in Figure 3.1. Endonuclease restriction digestion confirmed the cDNA inserts were of the predicted size for each construct (data not shown), and their coding regions were all sub-cloned into pBluescript II under the control of the T7 promoter for *in vitro* transcription. The coding regions for all of the constructs were sequenced and found to be correct (see Figure 3.1 and 3.2 for domain structure and approximate position of cysteine residues, and Appendix 1 for the amino acid sequences of the constructs). I created another chimeric construct (Construct 7; PDI ***ab*** ERp57 ***b'a'c***) in order to allow a complete comparison of the role of all of the ERp57 domains (see Section 2.4).

Amino acid sequence comparison of ERp57 and PDI reveals that their C-termini differ significantly, the C-terminus of ERp57 being essentially basic, whereas the equivalent region of PDI is markedly acidic (Koivunen *et al.,* 1996; Figure 1.13). I therefore hypothesised that the characteristic C-terminus of ERp57 may contribute to its specific association with calreticulin and calnexin. In order to test this hypothesis a premature stop codon was introduced in the ERp57 sequence at residue 466 just after the ***a'*** domain (ERp57Δc, see Figure 3.2).

The ERp57Δc derivative and ERp57/PDI chimeras were used in comparison with wild-type ERp57 and PDI in order to identify the region or regions of ERp57 responsible for its ability to form specific complexes with the ER lectins calnexin and calreticulin.

Figure 3.1 The domain structure of the ERp57/PDI constructs.



Constructs and amino acids present:

ERp57Δc ***abb'a'*** (1-466)

Construct 2 - ERp57 ***ab*** PDI ***b'a'c*** (1-218 and 219-491)

Construct 3 - ERp57 ***abb'x*** PDI ***a'c*** (1-351 and 351-491)

Construct 4 - PDI ***a*** ERp57 ***bb'a'c*** (1-119 and 111-481)

Construct 5 - PDI ***abb'x*** ERp57 ***a'*** (1-354 and 355-463)

Construct 6 - PDI ***abb'x*** ERp57 ***a'*** + AVKDEL (1-354 and 355-463)

Construct 7 - PDI ***ab*** ERp57 ***b'a'c*** (1-218 and 217-482)

Figure 3.1 The simplified domain structure of ERp57 and PDI is shown above along with the 7 constructs and the domains that they contain. The amino acids contributed by ERp57 and PDI to each particular construct are noted in brackets.

3.2 ERp57/PDI chimeras to map site of ER lectin binding on ERp57.

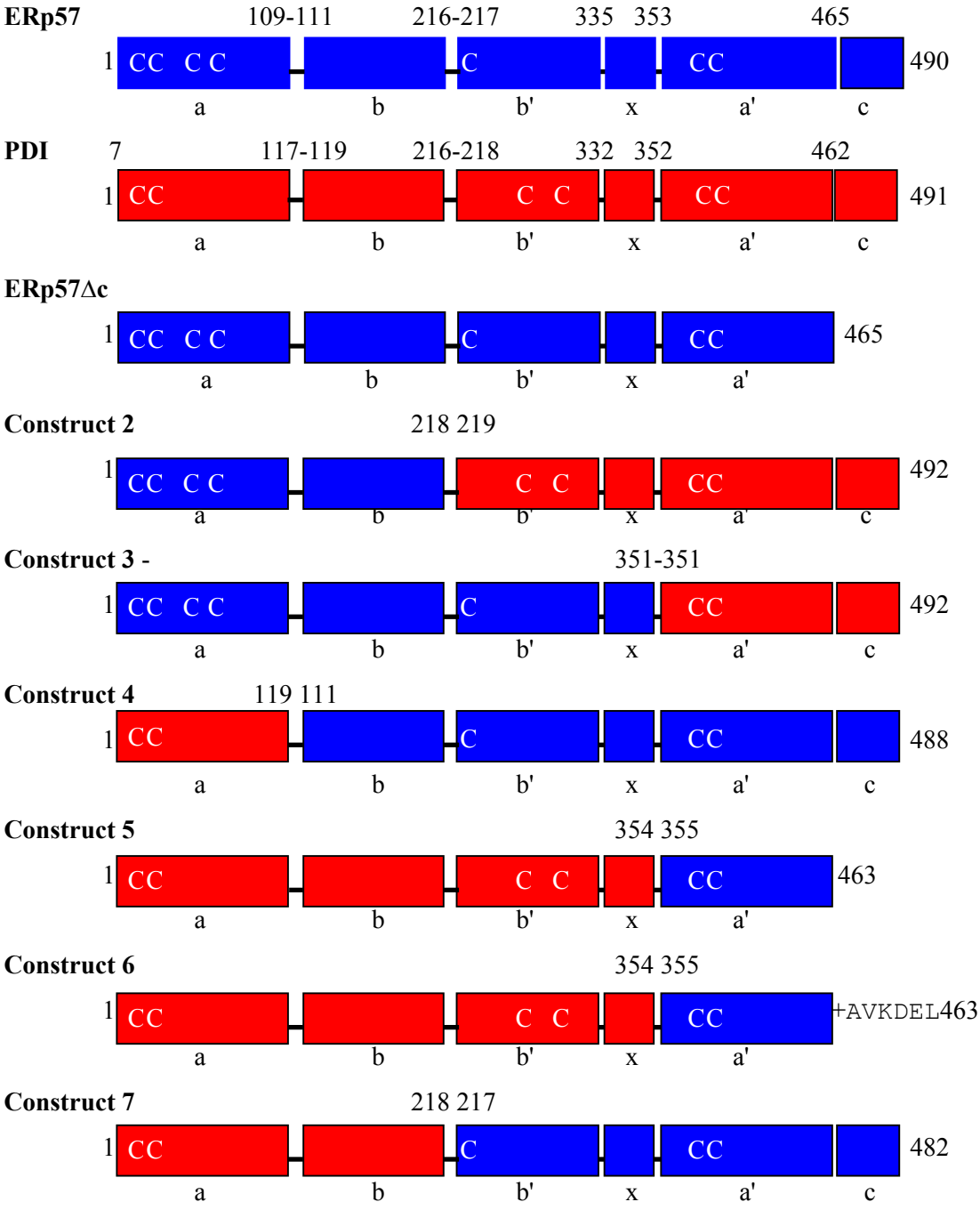


Figure 3.2 ERp57/PDI chimeras used to map the site of ER lectin binding within ERp57. PDI domains are shown in red with ERp57 domains highlighted in blue. The approximate positions of the cysteines present in each construct are indicated (white C).

3.2 Analysis of BMH-dependent cross-linking of ERp57/ER lectin complexes in microsomes

3.2.1 Analysis of total cross-linking products

The formation of ERp57/ER lectin complexes had previously been shown *in vitro* by importing a radiolabelled version of one of the partners into ER derived microsomes and analysing its interaction with the endogenous components of the ER lumen by cross-linking. Using this approach, a specific interaction of ERp57, but not PDI, with calnexin and calreticulin was demonstrated (Oliver *et al.,* 1999). The success of this *in vitro* analysis in reproducing the specific assembly of ER lectins with ERp57 suggested that I could exploit the same system to analyse the ERp57/PDI chimeras. Since PDI contains no information that allows it to bind calnexin and calreticulin (Oliver *et al.,* 1999), if any of the chimeras did assemble with the ER lectins, the domain or domains of ERp57 responsible should be apparent.

A preliminary analysis of the various chimeras was carried out using BMH-dependent (i.e. cysteine specific) cross-linking using wild-type ERp57 as a positive control. In this case, following *in vitro* synthesis of the radiolabelled precursors and their import into microsomes, total products obtained in the presence and absence of cross-linking reagent were analysed simply to establish whether any high molecular weight adducts similar to those previously obtained with ERp57 were formed (cf. Oliver *et al.,* 1999). This was seen as a prerequisite to any subsequent detailed analysis of cross-linking products by immunoprecipitation and the results are summarised in Table 3.1.

Table 3.1 BMH-dependent formation of ERp57-like adducts with ER luminal components.

Construct	ERp57-like BMH-dependent adducts detected in total products
ERp57	✓
ERp57Δc	✓
Construct 2	✖
Construct 3	✓
Construct 4	✓
Construct 5	✖
Construct 6	✖
Construct 7	✖

Since my preliminary study indicated at least some of the ERp57/PDI chimeras might generate complexes with ER lumenal components, I went on to analyse the nature of these complexes by immunoprecipitation.

3.2.2 Identification of BMH dependent cross-linking products.

The ERp57Δc and ERp57/PDI chimeric proteins are all non-natural polypeptides that may misfold in the ER lumen and thereby be bound by generic chaperones such as PDI or BiP. It was therefore essential for me to establish whether the adducts visible by analysing total products represented cross-linking to calreticulin and calnexin and hence reflected specific complex formation. To this end, ERp57, ERp57Δc, constructs 3 and 4 and PDI were translated *in vitro* in the presence of microsomal membranes. BMH and DMSO treated membrane fractions were prepared and the resulting products analysed by immunoprecipitation (Section 2.8). Each sample was immunoprecipitated using anti-sera specific for calreticulin, calnexin, PDI (C-terminal peptide), BiP (N-terminal peptide) and ERp57 (whole protein). All of the samples, except for construct 3, were also immunoprecipitated with a second anti-ERp57 serum (C-terminal peptide). Construct 3 was omitted from this analysis since it does not contain the ERp57 C-terminal epitope (see Figure 3.2) and is not recognised by the anti-ERp57 C-terminal peptide specific serum (data not shown).

In all instances BMH-dependent cross-linking products were observed in these experiments (Figures 3.3-3.7). In the case of wild-type ERp57, strong cross-linking products could be immunoprecipitated using anti-sera raised against a C-terminal ERp57 derived peptide (Figure 3.3; lane 5), purified ERp57 protein (Figure 3.3, lane 6) and recombinant calreticulin (Figure 3.3; lane 7). Three distinct cross-linking products between calreticulin and ERp57 were observed (Figure 3.3, lanes 5-7), most probably due to cross-linking between different cysteine residues within the two proteins. As previously observed (Oliver *et al.,* 1999), a distinct adduct of radiolabelled ERp57 with endogenous calnexin was immunoprecipitated by both anti-ERp57 sera (Figure 3.3, lanes 5 and 6*) and

an anti-calnexin serum (Figure 3.3; lane 10*), but not by the anti-calreticulin serum (Figure 3.3, lane 7). No higher molecular weight products were immunoprecipitated in the absence of BMH treatment (Figure 3.3, lanes 3 and 4), and no adducts were detected with anti-PDI, anti-BiP or a pre-immune serum in the presence of BMH treatment (Figure 3.3; lanes 8, 11 and 9 respectively).

Two distinct ERp57 derived products were detected when the protein was synthesised in the presence of membranes (Figure 3.3, lanes 1-6, arrow) and these are assumed to be the precursor before (upper product) and after (lower product) cleavage of the ER targeting signal (cf. Oliver *et al.,* 1999). The small amount of lower molecular weight products are most likely due to premature chain termination and/or alternative initiation at downstream methionine residues.

Figure 3.3 Interactions of *in vitro* synthesised ERp57 with the endogenous ER proteins of microsomes

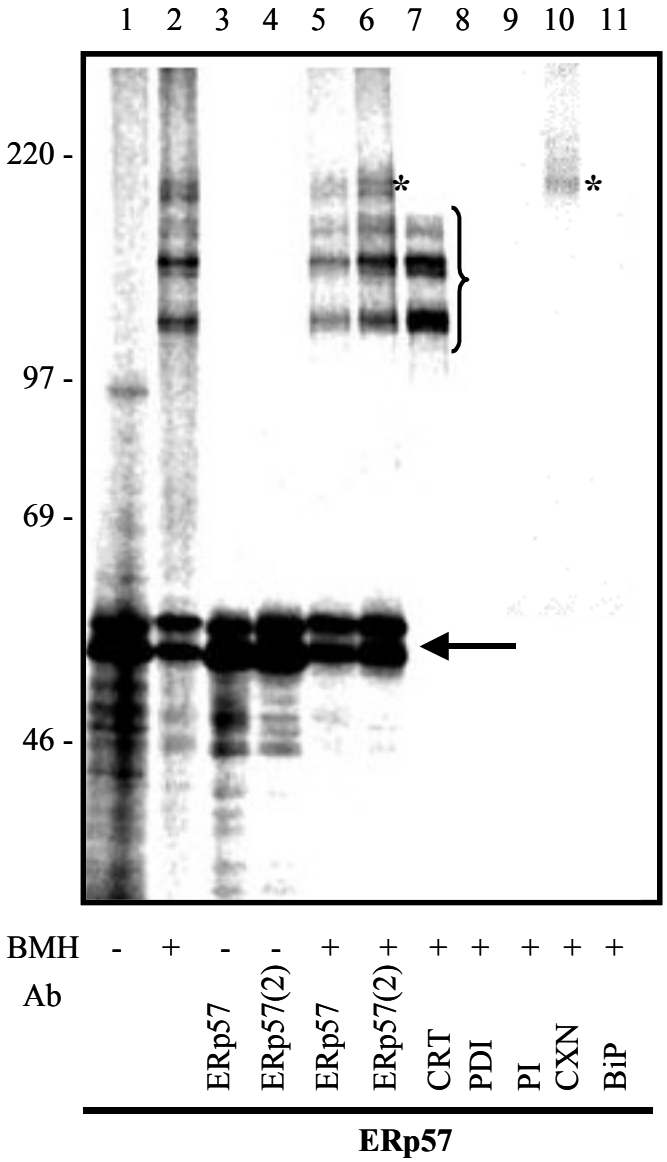


Figure 3.3 Interactions of *in vitro* synthesised ERp57 with the endogenous ER proteins of microsomes. ERp57 mRNA was translated in a rabbit reticulocyte lysate translation system in the presence of canine pancreatic microsomes. After termination of translation membrane-associated material was isolated by centrifugation, and incubated with either 1mM BMH (+) or mock-treated by incubation with DMSO (-). The reaction was quenched with 5mM β-mercaptoethanol and samples were denatured with 1% SDS diluted in 4 volumes of IP buffer and incubated overnight with the anti-sera indicated: ERp57, anti-ERp57 (C-terminal); ERp57 (2), anti-ERp57 (whole protein), CRT, anti-calreticulin; PDI, anti-PDI (C-terminal); PI, pre-immune; CXN, anti-calnexin; BiP, anti-BiP (N terminal). Antibody complexes were isolated by incubating samples for 2 hours, rolling at 4ºC, with protein A sepharose or a mix of protein A and G sepharose pre-blocked with BSA. The samples were analysed by electrophoresis on an 8% SDS-polyacrylamide gel. The major ERp57 derived translation products are indicated by a black arrow (←). Specific adducts with calreticulin (ϡ) and calnexin (*) are indicated.

Figure 3.4 Interactions of *in vitro* synthesised ERp57∆c with the endogenous ER proteins of microsomes.

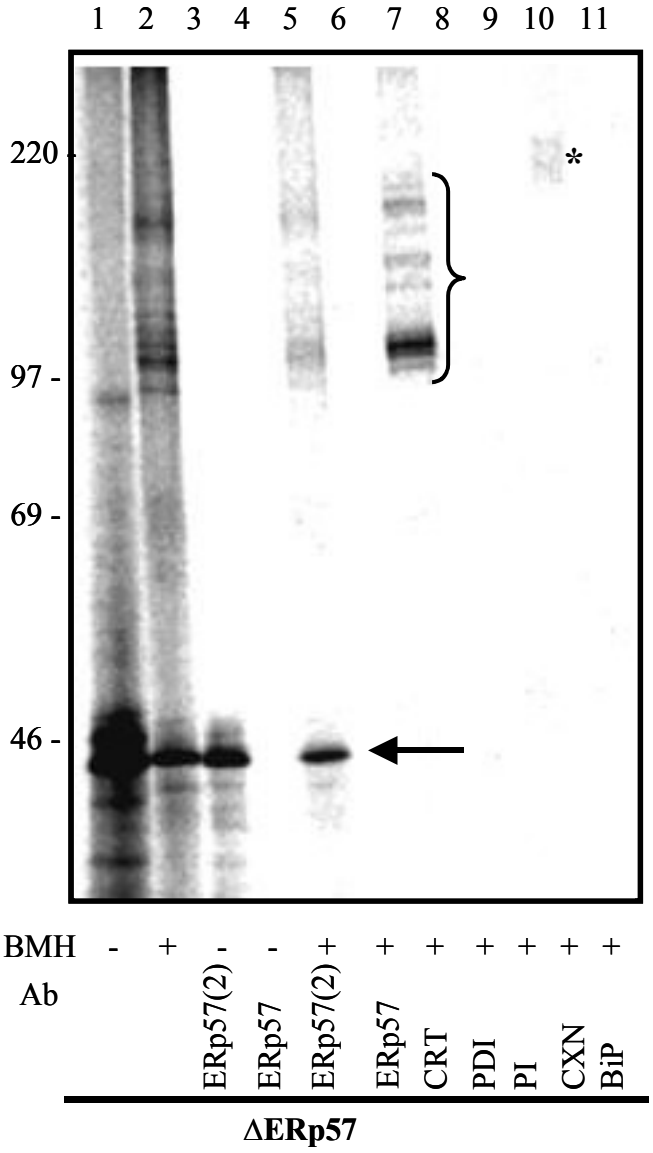


Figure 3.4 Interactions of *in vitro* synthesised ERp57Δc with the endogenous ER proteins of microsomes. ERp57Δc mRNA was translated in a rabbit reticulocyte lysate translation system as described for Figure 3.3 and immunoprecipitated. Anti-sera indicated: ERp57, anti-ERp57 (C-terminal); ERp57 (2), anti-ERp57 (whole protein), CRT, anti-calreticulin; PDI, anti-PDI (C-terminal); PI, pre-immune; CXN, anti-calnexin; BiP, anti-BiP (N terminal). Uncross-linked ERp57Δc imported polypeptides are indicated by arrows (←). Cross-links to calreticulin are denoted by (⊦) and to calnexin by a black asterisk (*).

The pattern of adducts observed after the cross-linking of ERp57Δc to ER lumenal components was revealing, and discrete cross-linking products could be immunoprecipitated by both an anti-ERp57 serum (raised against the purified protein; Figure 3.4, lane 5) and the anti-calreticulin serum (Figure 3.4, lane 7}). A weak, but distinct adduct between ERp57 and calnexin could also be detected with a calnexin specific antiserum (Figure 3.4, lane 10*). In contrast no ERp57Δc adducts were immunoprecipitated with anti-PDI, anti-BiP or a control pre-immune serum (Figure 3.4, lanes 8,11 and 9 respectively). Likewise no high molecular weight adducts were seen in the absence of cross-linking (Figure 3.4, lanes 3 and 4). Hence, ERp57Δc exhibited a specific interaction with the ER lectins calreticulin and calnexin, despite the loss of its C-terminus. Furthermore, no evidence of any interaction with PDI or BiP could be detected suggesting that the truncated polypeptide was not recognised as a misfolded protein.

Consistent with its mixed domain structure, the chimeric ERp57/PDI construct 3 was recognised by the anti-PDI (C-terminus) serum and the anti-ERp57 (purified protein) serum, (Figure 3.5; lanes 3 and 4). Both of these sera immunoprecipitated three cross-linking products (Figure 3.5; lanes 5 and 6}) that were clearly visible in the total products (Figure 3.5; lane 2). Most significantly, an identical pattern of cross-linking products was also immunoprecipitated by the anti-calreticulin serum (Figure 3.5; lane 7}). Two higher molecular weight adducts of construct 3 with calnexin could also be specifically identified by immunoprecipitation (Figure 3.5; lane 9*). In contrast, control experiments showed no products were isolated non-specifically (Figure 3.5; lanes 8, pre-immune serum) and that the adducts were seen only after cross-linking (Figure 3.5, lanes 3 and 4). Hence, the ERp57/PDI chimera, construct 3 appeared to be specifically associated with the ER lectins calreticulin and calnexin. Moreover, like ERp57Δc, no evidence of BMH-dependent cross-

linking to BiP could be detected (Figure 3.5, lane 10) indicating the chimera was not misfolded.

I found that a proportion of the construct 3 polypeptide displayed a decreased mobility upon BMH treatment (Figure 3.5, lanes 2, 5 and 6, <). This is most likely due to the modification of multiple cysteine residues within the construct 3 polypeptide by BMH molecules that do not cross-link to other proteins, but are quenched by β-mercaptoethanol.

Figure 3.5 Interactions of *in vitro* synthesised Construct 3 (ERp57 ***abb'*** PDI ***a'c***) with the endogenous ER proteins in microsomes.

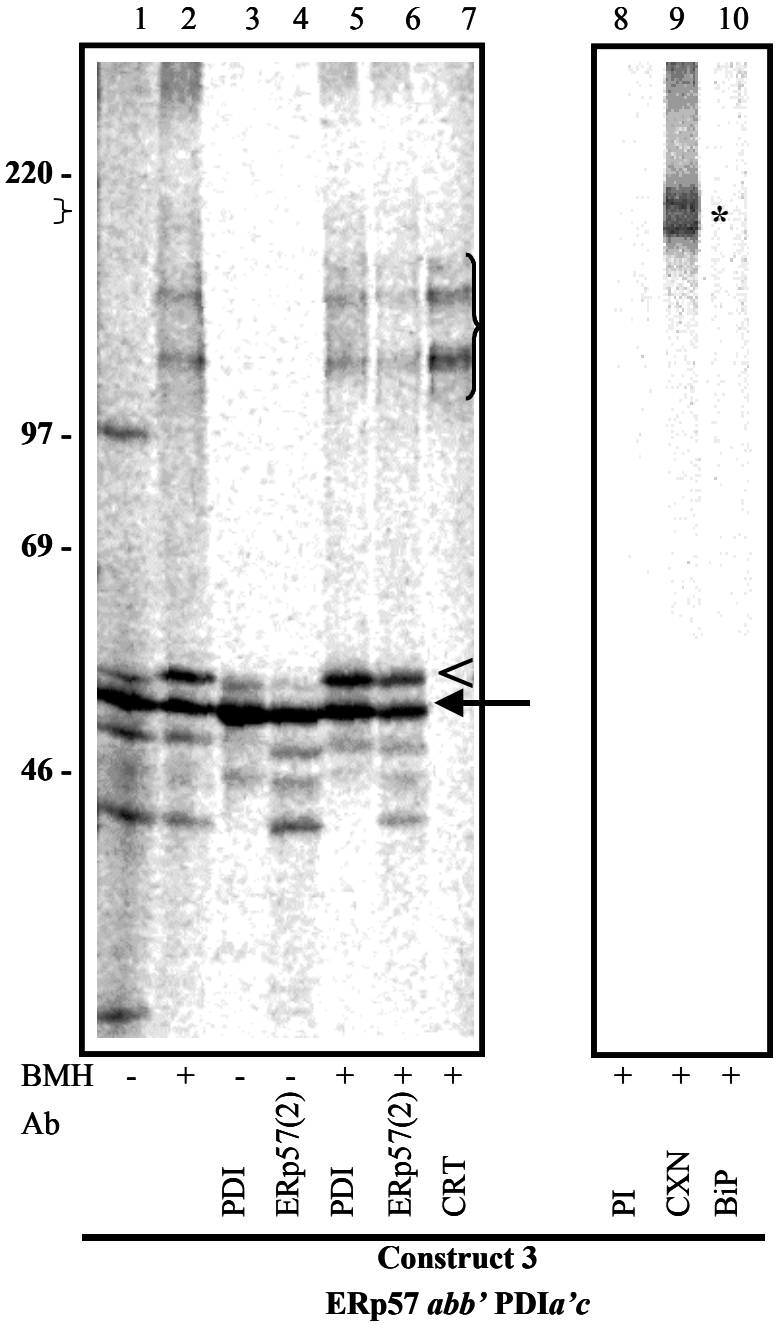


Figure 3.5 Interactions of *in vitro* synthesised Construct 3 (ERp57 ***abb'*** PDI ***a'c***) with the endogenous ER proteins in microsomes. Construct 3 mRNA was translated in a rabbit reticulocyte lysate translation system as described in Figure 3.3 and immunoprecipitated. Anti-sera indicated: ERp57 (2), anti-ERp57 (whole protein), CRT, anti-calreticulin; PDI, anti-PDI (C-terminal); PI, pre-immune; CXN, anti-calnexin; BiP, anti-BiP (N terminal). Principal construct 3 derived polypeptides are indicated by a black arrow (◄—). Adducts with calreticulin are denoted by (}) and to calnexin by a black asterisk (*). Polypeptide showing BMH-dependent alteration in mobility indicated by arrowhead (<).

The ERp57/PDI chimeric construct 4 was recognised by the anti-ERp57, C-terminal, serum (Figure 3.6; lane 2) and the BMH dependent cross-linking products were immunoprecipitated with this reagent (Figure 3.6; lane 3}). Most significantly, an identical pattern of cross-linking products was immunoprecipitated by the anti-calreticulin serum (Figure 3.6; lane 4}). Higher molecular weight adducts were also immunoprecipitated with the anti-calnexin serum (Figure 3.6; lane 5*). No construct 4 adducts were immunoprecipitated with anti-BiP, anti-PDI or a control pre-immune serum (Figure 3.6; lanes 6, 7 and 8 respectively). Likewise no products were seen in the absence of cross-linking (Figure 3.6, lane 2). Hence, the ERp57/PDI chimera, construct 4 appeared to be specifically associated with the ER lectins calreticulin and calnexin.

In contrast to the data presented for ERp57/PDI chimera 3 and 4, no evidence for any interaction between chimeras 2, 5, 6 or 7 and the ER lectins was obtained (data not shown, see also Figures 3.8 and 3.9).

Figure 3.6 Interactions of *in vitro* synthesised Construct 4 (PDI ***a*** ERp57 ***bb'a'c***) with the endogenous ER proteins of microsomes.

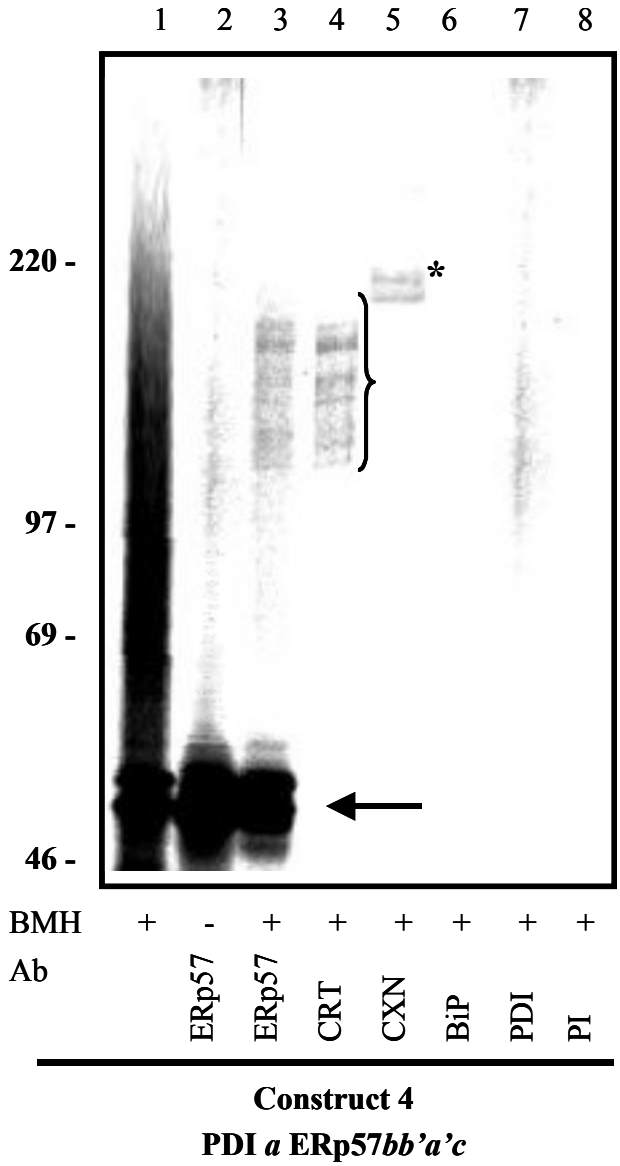


Figure 3.6 Interactions of *in vitro* synthesised Construct 4 (PDI ***a*** ERp57 ***bb'a'c***) with the endogenous ER proteins of microsomes. Construct 4 mRNA was translated in a rabbit reticulocyte lysate translation system as described in Figure 3.3 and immunoprecipitated. Anti-sera indicated: ERp57, anti-ERp57 (C-terminal); CRT, anti-calreticulin; PDI, anti-PDI (C-terminal); PI, pre-immune; CXN, anti-calnexin; BiP, anti-BiP (N terminal). Uncross-linked Construct 4 imported polypeptides are indicated by a black arrow (◄—). Cross-links to calreticulin are denoted by (⊃) and to calnexin by a black asterisk (*).

Archetypal PDI was used as a 'negative control', to demonstrate that the BMH dependent cross-linking of ERp57 and its derivates to endogenous calreticulin and calnexin was specific and representative of complex formation rather than some other phenomenon (cf. Oliver *et al.,* 1999). I found that whilst BMH dependent adducts between radiolabelled PDI and endogenous components of the ER lumen were observed (Figure 3.7, cf. lanes 1 and 2), none of these adducts were recognised by sera specific for calreticulin or calnexin (Figure 3.7, lanes 7 and 9). The higher molecular weight adducts were immunoprecipitated by the anti-PDI serum (Figure 3.7, lane 5}) but not by the anti-ERp57 or anti-BiP serum (Figure 3.7, lanes 6 and 10). These products were clearly dependent upon the addition of BMH (Figure 3.7, cf. lanes 3 and 5) although a distinct high molecular weight PDI product was seen in the absence of cross-linking after immunoprecipitation with anti-PDI and the anti-ERp57 serum (Figure 3.7, lanes 3 and 4, filled circle). This product appears to be non-specific since it could also be seen with a control pre-immune serum following BMH treatment (Figure 3.7, lane 8, filled circle).

The BMH dependent PDI adducts observed in Figure 3.7 are most likely PDI cross-linked to one or more isoforms of the canine Ero1 protein present in the microsomes used for this experiment. The Ero1 proteins act to facilitate the function of PDI, but not ERp57, and have been found to form mixed disulphides with PDI (Mezghrani *et al.,* 2001). The size and nature of the PDI cross-linking product I observe suggest they are most likely to reflect the interaction of PDI with Ero1 proteins during the normal cycle of PDI function (Frand *et al.,* 2000). Most importantly, the lack of cross-linking between PDI and the ER lectins calnexin and calreticulin reinforces the authenticity of the adducts observed with specific ERp57 derivatives and is consistent with particular regions of ERp57 being responsible for this interaction.

Figure 3.7 Interactions of *in vitro* synthesised PDI with the endogenous ER proteins of microsomes.

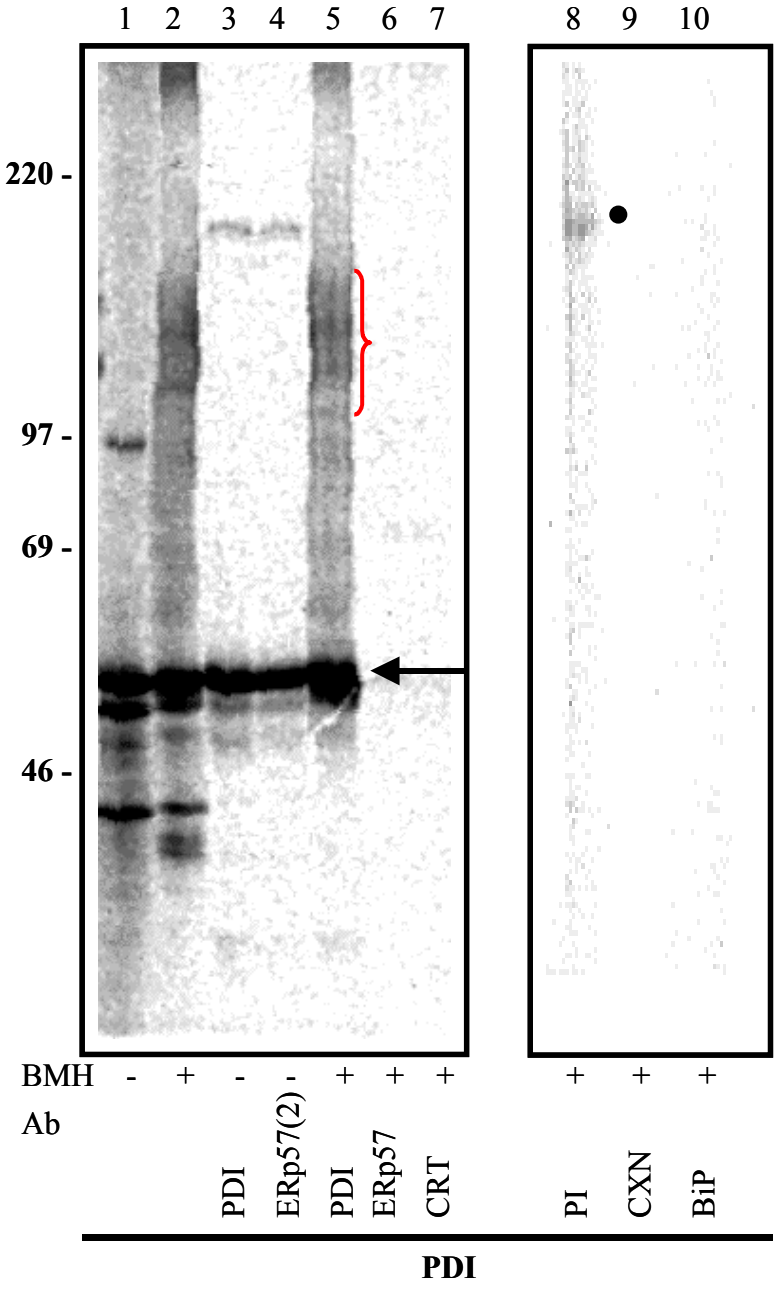


Figure 3.7 Interactions of *in vitro* synthesised PDI with the endogenous ER proteins of microsomes. PDI mRNA was translated in a rabbit reticulocyte lysate translation system as described in Figure 3.3 and immunoprecipitated. Anti-sera indicated: ERp57, anti-ERp57 (C-terminal); ERp57 (2), anti-ERp57 (whole protein), CRT, anti-calreticulin; PDI, anti-PDI (C-terminal); PI, pre-immune; CXN, anti-calnexin; BiP, anti-BiP (N terminal). samples were run on two separate gels. BMH-dependent cross-links recognised by anti-PDI sera are indicated by (}). Uncross-linked PDI imported polypeptides are indicated by a black arrows (←). A non-specific high molecular weight PDI product is indicated by a (black, filled circle)

3.3 Analysis of ERp57/ER lectin binding using SMCC dependent cross-linking

One limitation of using BMH to analyse the association of ERp57 with the ER lectins, calnexin and calreticulin, is that cross-linking can only occur between cysteine residues within the subunits of the complex. There are differences in both the number and location of cysteine residues in ERp57, PDI and the various chimeras of the two proteins (cf. Figure 3.2). It was therefore possible that some of the chimeras were not cross-linked to the ER lectins because they lacked suitable cysteine residues rather than being unable to interact. In order to address this issue, I repeated my cross-linking analysis of ERp57, PDI and the various ERp57 chimeras and derivatives using the cross-linking reagent succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate (SMCC). SMCC displays a broader specificity than BMH, being a hetero-bifunctional reagent that reacts with lysine and cysteine side chains. All of the proteins of interest contain a substantial number of lysine residues distributed throughout their sequence, in addition to the cysteine residues displayed in Figure 3.2.

A second parameter was also adjusted for this subsequent cross-linking study in that semi-permeabilised mammalian tissue culture cells were used in place of canine pancreatic microsomes as a source of ER during *in vitro* translation. Semi-permeabilised tissue culture cells can be prepared by treating the cells with low concentrations of digitonin in order to selectively permeabilise the plasma membrane (Wilson *et al.,* 1995). Previous studies had shown that the *in vitro* formation of ERp57/ER lectin complexes was efficient in both microsomes and semi-intact cells (Oliver *et al.,* 1999) and I confirmed this was the case using BMH dependent cross-linking to endogenous ER lectins with radiolabelled wild-type ERp57 (data not shown). The use of semi-intact cells also provides

the potential to exploit cell lines that are deficient in specific components of the ER glycoproteins folding machinery (Oliver *et al.,* 1999).

mRNAs encoding ERp57, chimeras 2-7 and PDI were translated in a rabbit reticulocyte lysate system, supplemented with, semi-permeabilised human HT1080 fibroblast derived tissue culture cells. The membrane-associated material was isolated and treated with SMCC at a final concentration of 250μM, which proved to give optimal cross-linking results (data not shown). The total products of the reaction before and after cross-linking were then examined and a number of specific products identified by immunoprecipitated as described in Section 3.2 above.

The use of SMCC as a cross-linking reagent generated much less clear adducts than those seen with BMH, particularly when the total products of the cross-linking reactions were analysed (Figures 3.8 A-D and 3.9 A-D, cf. lanes 1 and 2). However, by analysing the cross-linking products that were immunoprecipitated with sera specific for calreticulin and calnexin it was immediately apparent that only a subset of the ERp57 derivatives displayed SMCC dependent cross-linking to these ER lectins. Thus, multiple adducts with calreticulin were obtained with ERp57, construct 3 and construct 4 (Figures 3.8A, 3C and 3D, lane 7⊦) and a weak but detectable doublet with calnexin was also seen (Figures 3.8A, 3C and 3D, lane 8*). These data are very similar to those obtained with BMH dependent cross-linking (cf. Figures 3.3, 3.5 and 3.6). No such SMCC dependent adducts were detected with constructs 2, 5, 6 or 7 (Figures 3.8B, 3.9A, 3.9B, 3.9C) nor were they seen with PDI (Figure 3.9D). These data are entirely consistent with that obtained using BMH as a cross-linking reagent and indicate, that of the ERp57 chimeras analysed, only construct 3 and construct 4 can associate with the ER lectins calreticulin and calnexin.

In many of the experiments carried out using semi-permeabilised mammalian cells, radiolabelled products of ~55kDa, and in some cases ~90kDa, were immunoprecipitated by

the anti-calreticulin and anti-calnexin sera (Figures 3.8A-D and 3.9A-D, lanes 7 and 8* and * respectively). These products are almost certainly ^{35}S-methionine labelled calreticulin and calnexin resulting from the translation of endogenous mRNA associated with the ER of the semi-intact mammalian cells. This suggests that the treatment of the semi-permeabilised cells with micrococcal nuclease to remove the endogenous mRNAs was not completely efficient (cf. Wilson *et al.,* 1995).

As with BMH, archetypal PDI was used to control for the specificity of SMCC dependent cross-linking to ER lumenal components, whilst SMCC dependent adducts were clearly formed between radiolabelled PDI and endogenous ER lumenal proteins (Figure 3.9D, cf. lanes 4 and 5, }-), no adducts with calreticulin or calnexin were present (Figure 3.9D, lanes 7 and 8). These adducts of PDI are most likely with ER lumenal Ero1 proteins as discussed above. Hence, the SMCC dependent cross-linking of a specific subset of ERp57 chimeras to the ER lectins appears to indicate an authentic interaction between these components.

Figure 3.8 Interactions of *in vitro* synthesised ERp57, Construct 2-4 with the endogenous ER proteins of semi-permeabilised HT1080 cells.

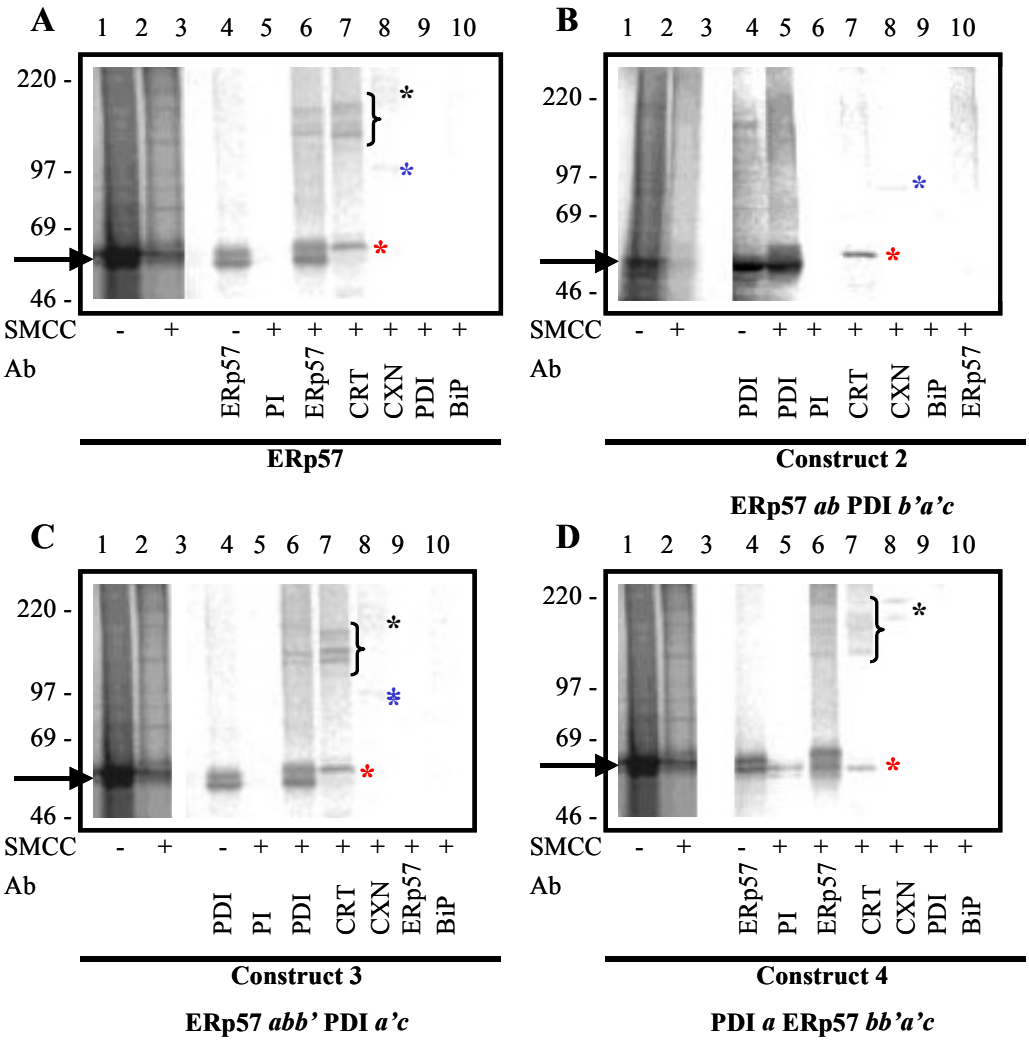


Figure 3.8 Interactions of *in vitro* synthesised ERp57 (Panel A), Construct 2 (Panel B; ERp57 ***ab*** PDI ***b'a'c***), Construct 3 (Panel C; ERp57 ***abb'*** PDI ***a'c***) and Construct 4 (Panel D; PDI ***a*** ERp57 ***bb'a'c***) with the endogenous ER proteins of semi-permeabilised HT1080 cells. mRNA was translated in a rabbit reticulocyte lysate translation system in the presence of semi-permeabilised HT1080 cells. After termination of translation membrane-integrated material was isolated by centrifugation, and incubated with either 0.25mM SMCC (+) or mock-treated by incubation with DMSO (-). The reaction was quenched with 5mM β-mercaptoethanol and 100mM glycine, the samples were denatured with 1% SDS, diluted in 4 volumes of IP buffer and incubated overnight with the antisera indicated: ERp57, anti-ERp57 (C-terminal); CRT, anti-calreticulin; PDI, anti-PDI (C-terminal); PI, pre-immune; CXN, anti-calnexin; BiP, anti-BiP (N terminal). Antibody complexes were isolated using protein A sepharose or a mix of protein A and G sepharose. The samples were analysed on an 8% SDS-polyacrylamide gel. Uncross-linked ERp57 and chimeric construct imported polypeptides are indicated by arrows (→). Cross-links to calreticulin are denoted by (⊣) and to calnexin by a black asterisk (*). Endogenous calreticulin and calnexin immunoprecipitated by calreticulin and calnexin antiserum are indicated by (* and* respectively).

Figure 3.9 Interactions of *in vitro* synthesised Constructs 5-7 and PDI with the endogenous ER proteins of semi-permeabilised HT1080 cells.

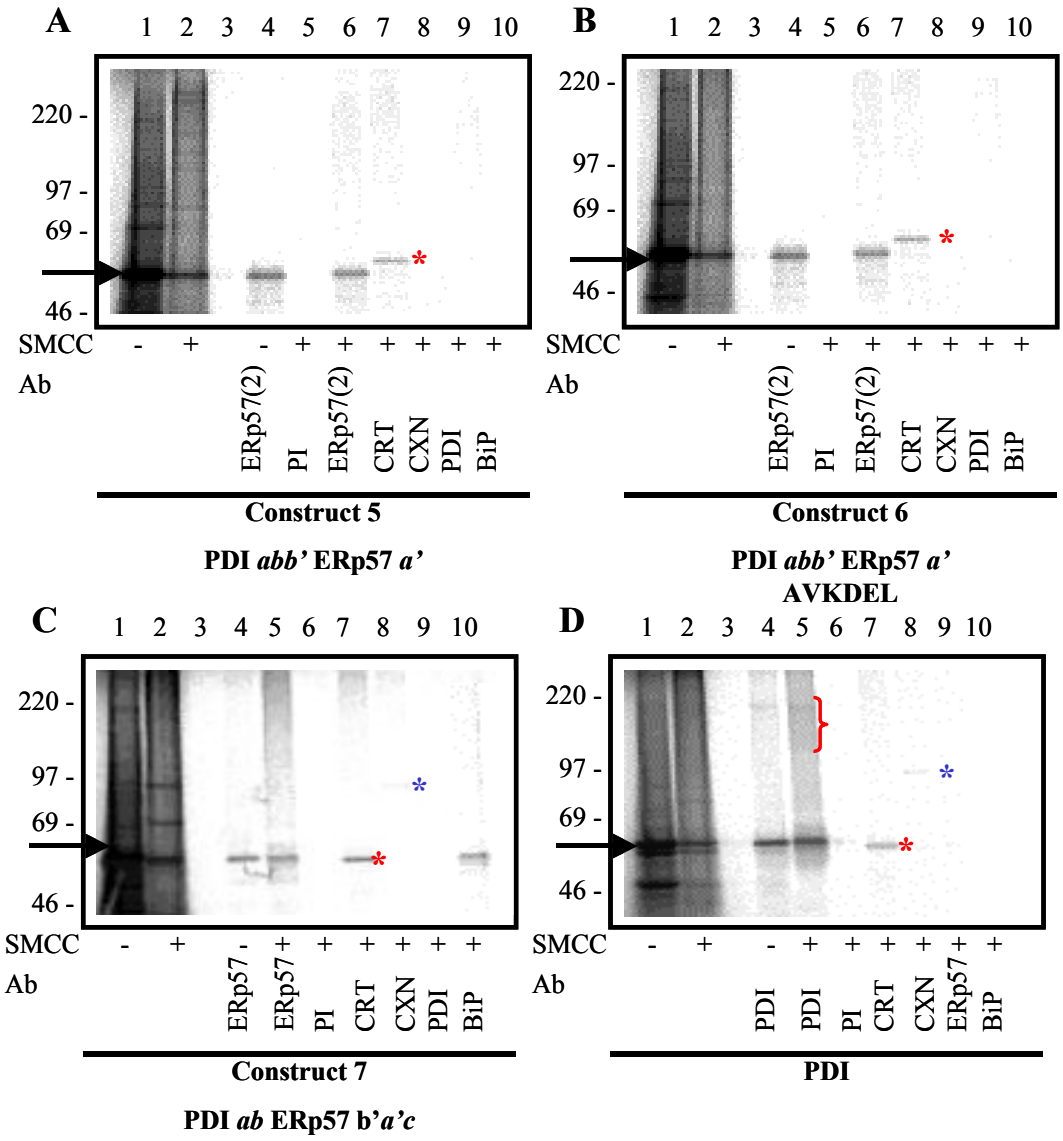


Figure 3.9 Interactions of *in vitro* synthesised Construct 5 (Panel A PDI ***abb'*** ERp57 ***a'***), Construct 6 (Panel B; PDI ***abb'*** ERp57 ***a'***+ AVKDEL), Construct 7 (Panel C; PDI ***ab*** ERp57 ***b'a'c***) and PDI (Panel D) with the endogenous ER proteins of semi-permeabilised HT1080 cells. mRNA was translated in a rabbit reticulocyte lysate translation system in the presence of semi-permeabilised HT1080 cells, as described in Figure 3.9. Samples were denatured and incubated overnight with the antisera indicated: ERp57, anti-ERp57 (C-terminal); ERp57 (2), anti-ERp57 (whole protein), CRT, anti-calreticulin; PDI, anti-PDI (C-terminal); PI, pre-immune; CXN, anti-calnexin; BiP, anti-BiP (N terminal). Uncross-linked ERp57 and chimeric construct imported polypeptides are indicated by arrows (→). Cross-links recognised by anti-PDI sera are denoted by (}). Non-specific bands immunoprecipitated by calreticulin and calnexin antiserum are indicated by (* and* respectively).

3.4 Summary

By analysing the interaction of ERp57 derivatives with the ER lectins present in microsomes and semi-intact cells, I hoped to define regions of the ERp57 protein that enabled it to specifically associate with these ER lectins. The formation of protein complexes in the ER lumen was assessed by cross-linking, and I found that wild-type ERp57 formed specific cross-linking products with calreticulin and calnexin whilst PDI did not. This was the case with two different cross-linking reagents and exploiting two different sources of ER derived membranes.

The C-terminal extension of ERp57 could be deleted with no obvious defect in the ability of the polypeptide to bind to calreticulin and calnexin. Likewise, two chimeras containing specific portions of ERp57, construct 3 and construct 4, also formed clear adducts with calreticulin and calnexin in both variations of the experimental system utilised. When the domain structures of the various chimeras used during this study are compared in detail (Figure 3.2)it can be seen that the common factor between the chimeras that bind to calreticulin and calnexin is the presence of both the ***b*** and ***b'*** domains from ERp57. Thus, these regions of ERp57 may be important for its specific interaction with the ER lectins.

3.5 Conclusions

The results presented in this chapter suggest that both the ***b*** and ***b'*** domains of ERp57 are necessary for its specific interaction with the ER lectins calreticulin and calnexin.

CHAPTER - 4

Results

Cross-linking analysis of ERp57 derivatives and point mutants

4 Cross-linking analysis of ERp57 derivatives and point mutants.

In Chapter 3, the molecular basis for the interaction of ERp57 with the ER lectins calreticulin and calnexin was investigated by importing wild-type ERp57, chimeras of ERp57/PDI and PDI into the ER and analysing the ability of these polypeptides to be cross-linked to endogenous calreticulin and calnexin. When the amino acid sequences of the various domains in ERp57 are analysed, the ***b*** and ***b'*** domains and the C terminal extension are found to show little or no homology with those of PDI (see Table 4.1 below). However, I found that the removal of the short C terminal extension of ERp57 had no effect on its association with the ER lectins, as detected by BMH cross-linking (see Chapter 3). Conversely both the ***b*** and ***b'*** domain of ERp57 were found to be necessary for the interaction of ERp57 derived chimeras with the ER lectins. Hence, following immunoprecipitation with anti-sera specific for calreticulin and calnexin, only constructs containing both the ***b*** and ***b'*** domains of ERp57 showed a similar cross-linking pattern to that of wild-type ERp57 (Chapter 3). Studies with the hetero-bifunctional cross-linking reagent SMCC and semi-permeabilised mammalian cells confirmed the initial cross-linking data obtained using the homo-bifunctional reagent BMH and canine pancreatic microsomes.

My data can be viewed in the light of the suggestion that for PDI the principal peptide binding site is the ***b'*** domain, with other domains contributing to the binding of larger substrates (Klappa *et al.,* 1998). Given the sequence divergence between the ***b'*** domains of PDI and ERp57 (Table 4.1), and the role of this domain as a peptide binding site in PDI, I hypothesised that the ***b'*** domain of ERp57 may function as a binding site for calreticulin and calnexin.

Table 4.1 Homology between domains of ERp57 and PDI

Domain	Homology (% identity)
abb'a'	**33**
a	**49.4**
b	**23.1**
b'	**14.4**
a'	**53.6**
c	**-**

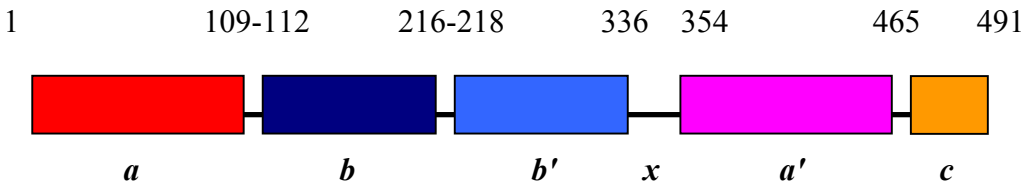
Table 4.1 Comparison of amino acid sequence identities between the various domains of ERp57 and PDI (taken from Koivunen *et al.,* 1996, see also Figure 4.1).

The aim of the work presented in this chapter was to test this hypothesis in two ways. Firstly, I wished to determine whether the ***b*** and/or ***b'*** domains of ERp57 are sufficient for its specific interaction with the ER lectins, and secondly, I wanted to establish whether specific point mutations, in the ***b'*** domain, based on mutants causing the loss of peptide binding in other members of the PDI family, could prevent ER lectin binding. To this end, ERp57 derivatives ***abb'***, ***b'a'c, bb'x*** and ***b'x*** were generated (see Figure 4.1), and a number of point mutants were constructed based on a ***b'*** domain peptide binding site model of PDI and PDIp (Figure 4.2; binding site).

4.1 The ERp57 derivatives studied

The ERp57 derivatives studied in this chapter include ***abb'***, ***bb'x***, ***b'a'c*** and ***b'x*** sub-domains (Figure 4.1), and the amino acids that contribute to the various constructs are indicated overleaf. For those constructs lacking the ***a*** domain, the codons for the first three amino acids of the mature chain were included to ensure authentic cleavage of the N-terminal signal sequence after import into the ER. Following signal sequence cleavage the polypeptides generated will therefore contain the sequence Ser-His-Met followed by the domains indicated.

Figure 4.1 ERp57 derivatives to map site of ER lectin binding on ERp57



Constructs and amino acids present

ERp57 ***abb'*** (1-336)

ERp57 ***b'a'c*** (1-3 and 218-482)

ERp57 ***bb'x*** (1-3 and 112-354)

ERp57 ***b'x*** (1-3 and 218-354)

Figure 4.1 The simplified domain structure of ERp57 is shown above with a list of ERp57 fragments studied. The numbering begins at the residue after the point of signal sequence cleavage and the amino acids present in each mature ERp57 derived protein are shown in brackets.

The ERp57 ***b'*** domain point mutants used in this study were based on a hypothetical model of the peptide binding site in the ***b'*** domain of PDI and PDIp (Figure 4.2), and each mutant was designed to alter a particular amino acid predicted to form part of the equivalent binding site in ERp57 (Figure 4.3).

The ERp57 derivatives and ERp57 point mutants detailed above were used in conjunction with wild-type ERp57 in a series of cross-linking experiments designed to determine their capacity to associate with ER lectins. In this way I hoped to more closely define the region or regions in ERp57 necessary and sufficient to form specific complexes with calnexin and calreticulin.

Figure 4.2 Hypothetical model of PDI structure.

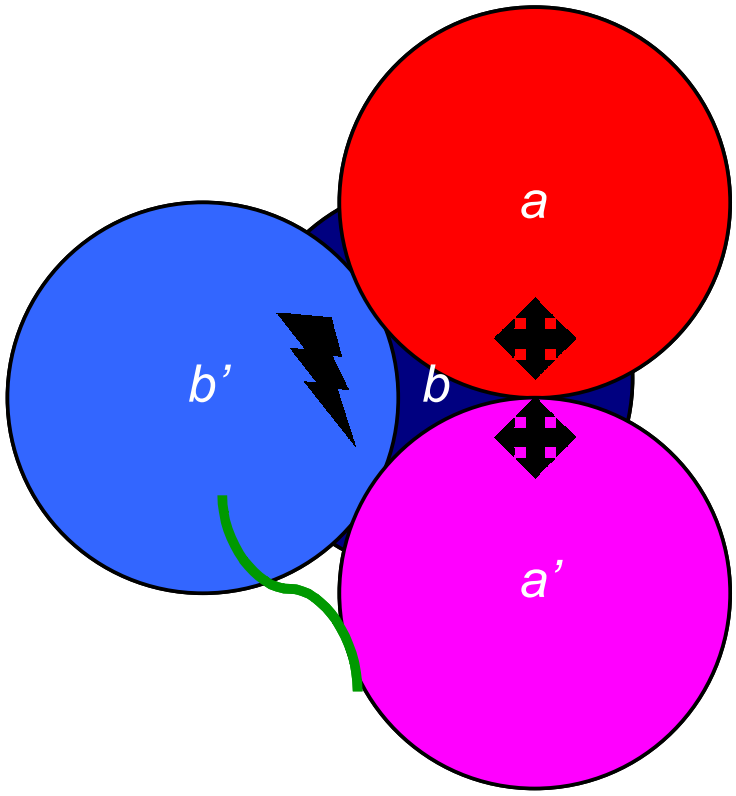
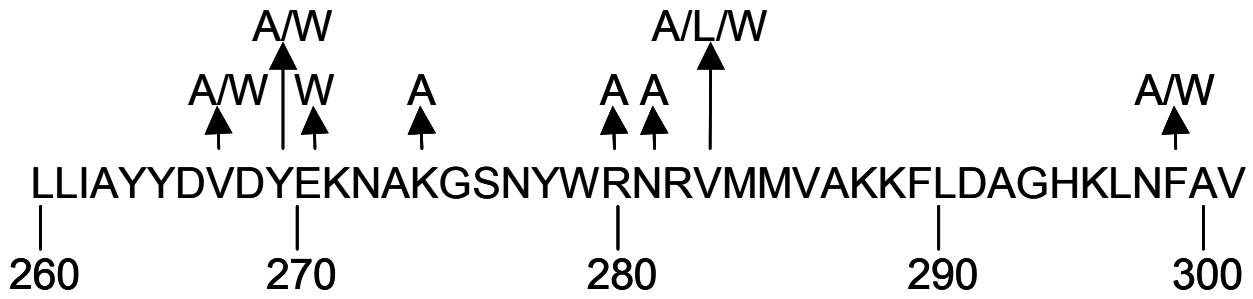


Figure 4.2 A hypothetical model of PDI structure used to identify amino acid residues within the ERp57 polypeptide that may be important for ER lectin binding. The CXXC active sites in the ***a*** and ***a'*** domains are marked (✥)and the putative substrate-binding site in the ***b'*** domain (↯). The x-linker region is depicted as a green line and would allow the ***a'*** domain to align close to the ***a*** domain bringing the two active sites into close proximity (Dr Lloyd Ruddock, University of Oulu, personal communication).

Figure 4.3 Schematic representation of ERp57 ***b'*** domain point mutants



ERp57 ***b'*** domain point mutants		
Point mutant 1 - ERp57 V267A	Point mutant 5 - ERp57 E270A	Point mutant 9 – ERp57 V283A
Point mutant 2 - ERp57 V267W	Point mutant 6 - ERp57 K274A	Point mutant 10 – ERp57 V283L
Point mutant 3 - ERp57 Y269A	Point mutant 7 - ERp57 R280A	Point mutant 11 – ERp57 V283W
Point mutant 4 - ERp57 Y269W	Point mutant 8 – ERp57 N281A	Point mutant 12 – ERp57 F299A
		Point mutant 13 – ERp57 F299W

Figure 4.3 Schematic representation of the ERp57 ***b'*** domain. The amino acid sequence is shown in one letter code with the point mutations shown above the wild-type amino acid sequence. A list of all 13 ERp57 point mutants, with a note of the amino acid change and residue number is also shown.

4.2 Analysis of SMCC-dependent cross-linking of ERp57 derivatives to ER lectins in semi-permeabilised cells

mRNAs encoding ERp57 and various ERp57 derivatives (ERp57***abb'***, ***b'a'c, bb'x*** and ***b'x***) were transcribed *in vitro* and then translated using the rabbit reticulocyte lysate system, supplemented with digitonin-permeabilised, semi-intact human HT1080 fibroblast cells. The cross-linking of the radiolabelled polypetides to endogenous ER lectins was carried out with the hetero-bifunctional reagent, succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate (SMCC), as described above (Section 3.3). SMCC was specifically used for the cross-linking of these derivatives so that adducts could be formed from both $-NH_2$ and –SH groups in the radiolabelled nascent chains. It was hoped that this would minimise any potential variation in cross-linking efficiency that may arise due to differences in the number of cysteine residues present in each derivative. Since the ***abb'***, ***b'a'c, bb'x*** and ***b'x*** derivatives of ERp57 are all non-natural polypeptides it was possible that they may misfold in the ER lumen and hence be bound by generic lumenal chaperones. For this reason, following cross-linking the samples were immunoprecipitated using anti-sera recognising the ER lectins calreticulin and calnexin, and the general chaperones/ folding factors BiP and PDI.

Following SMCC mediated cross-linking and immunoprecipitation of the samples with anti-calreticulin sera multiple adducts were revealed with wild-type ERp57 (Figure 4.4A, lane 7}) and ERp57 derivative ***abb'*** (Figure 4.4B, lane 6). Weak, but just detectable, cross-linking products with calnexin could also be seen with ERp57 (Figure 4.4A, lane 8*) and ERp57 ***abb'*** (Figure 4.4B, lane 7*). No adducts were detected in the absence of SMCC treatment of the wild-type ERp57 or ERp57 ***abb'*** (Figures 4.4A and 4.4B, lane 4). Likewise, after SMCC treatment no high molecular weight products were seen with a

control pre-immune serum or sera specific for PDI or BiP (Figure 4.4A lanes 5, 9 and 10; Figure 4.4B lanes 10, 8 and 9 respectively).

ERp57 derivatives ***b'a'c***, ***bb'x*** and ***b'x*** revealed an *in vitro* cross-linking pattern that was more difficult to resolve. In the case of the ERp57 ***bb'x*** derivative no clear adducts with the ER lectins could be detected (Figure 4.4C, lane 6 and 7), and only one faint adduct could be immunoprecipitated by the anti-calreticulin sera (Figure 4.4C, lane 6}). A number of weak cross-linking products were recognised by anti-sera specific for PDI (Figure 4.4C, lane 8}), and similar products were also found when ***b'a'c*** and ***b'x*** derivatives of ERp57 were analysed after SMCC treatment (data not shown). On the basis of this experiment and similar investigations that are not presented, I concluded that many of the ERp57 sub-domains could not be synthesised and correctly folded *in vitro.*

As I had observed in previous experiments carried out using semi-intact mammalian cells (see Chapter 3.3), radiolabelled products of ~55kDa, and in some cases ~90kDa, were immunoprecipitated by the anti-calreticulin (Figure 4.4A, lane 8 *; Figure 4.4B and 4.4C, lane 6 *) and anti-calnexin sera (Figures 4.4A, lane 8 * and 4.4C lane 6 *). This again indicates that the treatment of the semi-permeabilised cells with micrococcal nuclease to remove the endogenous mRNAs was not completely efficient and low levels of endogenous calreticulin and calnexin were becoming labelled with ^{35}S-methionine (cf. Wilson *et al.,* 1995).

Whilst my studies of the ERp57/PDI chimeras suggested that the ***b*** and ***b'*** domains of ERp57 are necessary for its interaction with the ER lectins (Chapter 3), my attempts to show that this region of ERp57 were sufficient for binding were less successful. I did find that the ***abb'*** region of ERP57 could specifically associate with calreticulin and calnexin, consistent with such a role. However, my attempts at analysing any smaller regions of

ERp57 than this were complicated by the apparent misfolding/instability of such ‘minimal’ domains.

Figure 4.4 Interactions of *in vitro* synthesised ERp57 (A), ERp57 ***abb'*** (B), ERp57 ***bb'x*** (C) with the endogenous ER proteins of semi-permeabilised HT1080 cells.

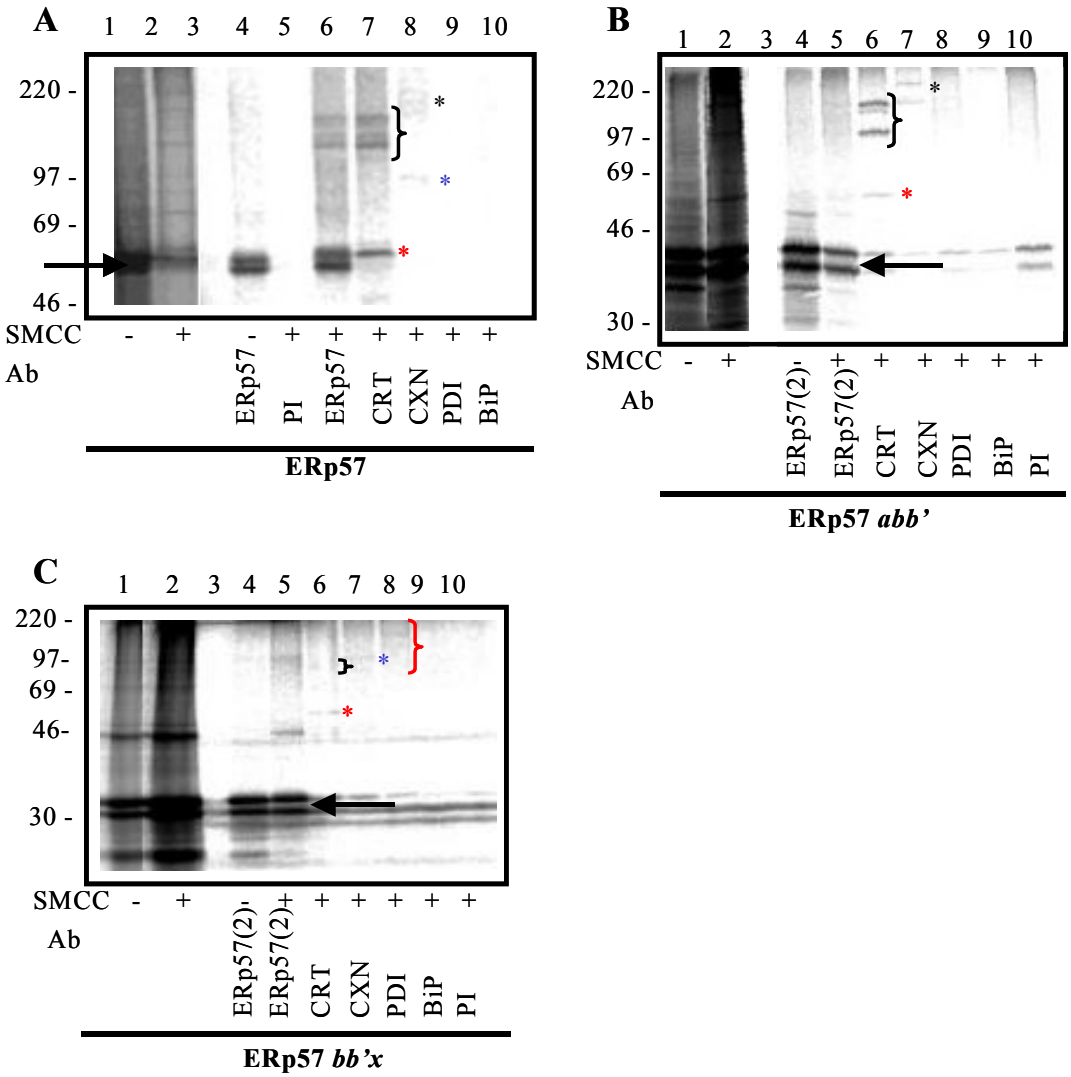


Figure 4.4 Interactions of *in vitro* synthesised ERp57 (A), ERp57 ***abb'*** (B), ERp57 ***bb'x*** (C) with the endogenous ER proteins of semi-permeabilised HT1080 cells. The mRNAs were translated in a rabbit reticulocyte lysate translation system in the presence of semi-permeabilised HT1080 cells. After termination of translation membrane-integrated material was isolated by centrifugation, and incubated with either 250mM SMCC (+) or mock-treated by incubation with DMSO (-). The reaction was quenched with 5mM β-mercaptoethanol and 100mM glycine. Samples were denatured with 1% SDS diluted in 4 volumes of IP buffer and incubated overnight with the anti-sera indicated: ERp57, anti-ERp57 (C-terminal); ERp57(2), anti-ERp57 (whole protein); PI, pre-immune of PDI; CRT, anti-calreticulin; CXN, anti-calnexin; BiP, anti-BiP (N terminal); PDI, anti-PDI (C-terminal). Antibody complexes were isolated by incubating samples for 2 hours, rolling at 4ºC, with protein A sepharose or a mix of protein A and G sepharose pre-blocked with BSA. The samples were analysed on SDS-polyacrylamide gels (8%, 12% and 10% respectively). Imported polypeptides that are not cross-linked are indicated by black arrows (→). Cross-linking products with calreticulin are denoted by (}) and with calnexin by a black asterix (*). Adducts with PDI are also indicated (}). Endogenous radiolabelled calreticulin and calnexin immunoprecipitated by the calreticulin and calnexin anti-sera are indicated (* and* respectively).

4.3 Cross-linking analysis of point mutants of the ERp57 b' domain

In order to investigate the ERp57 binding site for the ER lectins in more detail a series of point mutants located within the ***b'*** domain of the polypeptide were tested using the established *in vitro* cross-linking assay already described (see Figure 4.1). The ***b'*** domain was chosen as the site for selective mutagenesis because all of the ERp57 derivatives that interacted with the ER lectins contain the ***b'*** domain. Further more the ***b'*** domain is known to be the principle peptide-binding site of PDI (Klappa *et al.,* 1998). It was also hoped that the use of point mutants would enable the disruption of the putative ER lectin-binding site whilst allowing the ERp57 polypeptide to fold correctly into a wild type like structure. A number of single amino-acid substitutions within the ***b'*** domain of ERp57 were generated in order to establish whether this acted as the principal ER lectin binding site and, if possible, to determine which residues may be involved in any specific interaction (Figure 4.3). mRNAs encoding these ERp57 point mutants were then translated in the presence of ^{35}S-methionine and canine pancreatic microsomes, treated with the bifunctional reagent BMH, and immunoprecipitated with anti-ERp57 and anti-calreticulin sera.

In all cases the ERp57 point mutants were efficiently synthesised and imported into microsomes and BMH-dependent cross-linking products were observed, albeit that the intensity of the cross-linking products were variable (cf. Figures 4.5, Panels A-F, 4.6 Panels A-D and 4.7 Panels A-D, lanes 1-3). Since none of the mutations altered the number of cysteine residues present in ERp57, these variations in cross-linking efficiency were unlikely to result from differences in the numbers of thiol groups available for modification by BMH. In the case of wild-type ERp57 obvious BMH-dependent cross-linking products could be immunoprecipitated with anti-ERp57 sera (Figure 4.5A, cf. lanes

2 and 3). These were confirmed as ERp57/calreticulin adducts by immunoprecipitation with anti-calreticulin sera (Figure 4.5A, lane 4,).

For eleven of the thirteen ERp57 point mutants analysed, clear evidence of BMH dependent adducts with calreticulin could be observed (Figures 4.5 B-F, 4.6 A-D and 4.7 A and B, see lane 4 in each case). Two of the point mutants showed no obvious adducts with calreticulin in this preliminary analysis (Figures 4.7 C and D, lane 4), suggesting that the binding of ERp57 mutants 12 (F299A) and 13 (F299W) to the ER lectin calreticulin was significantly compromised. Careful inspection of the data presented in Figures 4.5-4.7 suggested that some of the other ERp57 point mutants may also cross-link less efficiently to calreticulin than others. However, the variation in the translation efficiency of the different point mutants made it difficult to be certain about the significance of such apparent differences. For this reason, the phosphorimaging data obtained for each ERp57 point mutant was analysed so as to quantify the relative proportion of that mutant that was cross-linked to calreticulin.

Figure 4.5 Interactions of wild-type and ERp57 *b'* domain point mutants (1-5) with calreticulin.

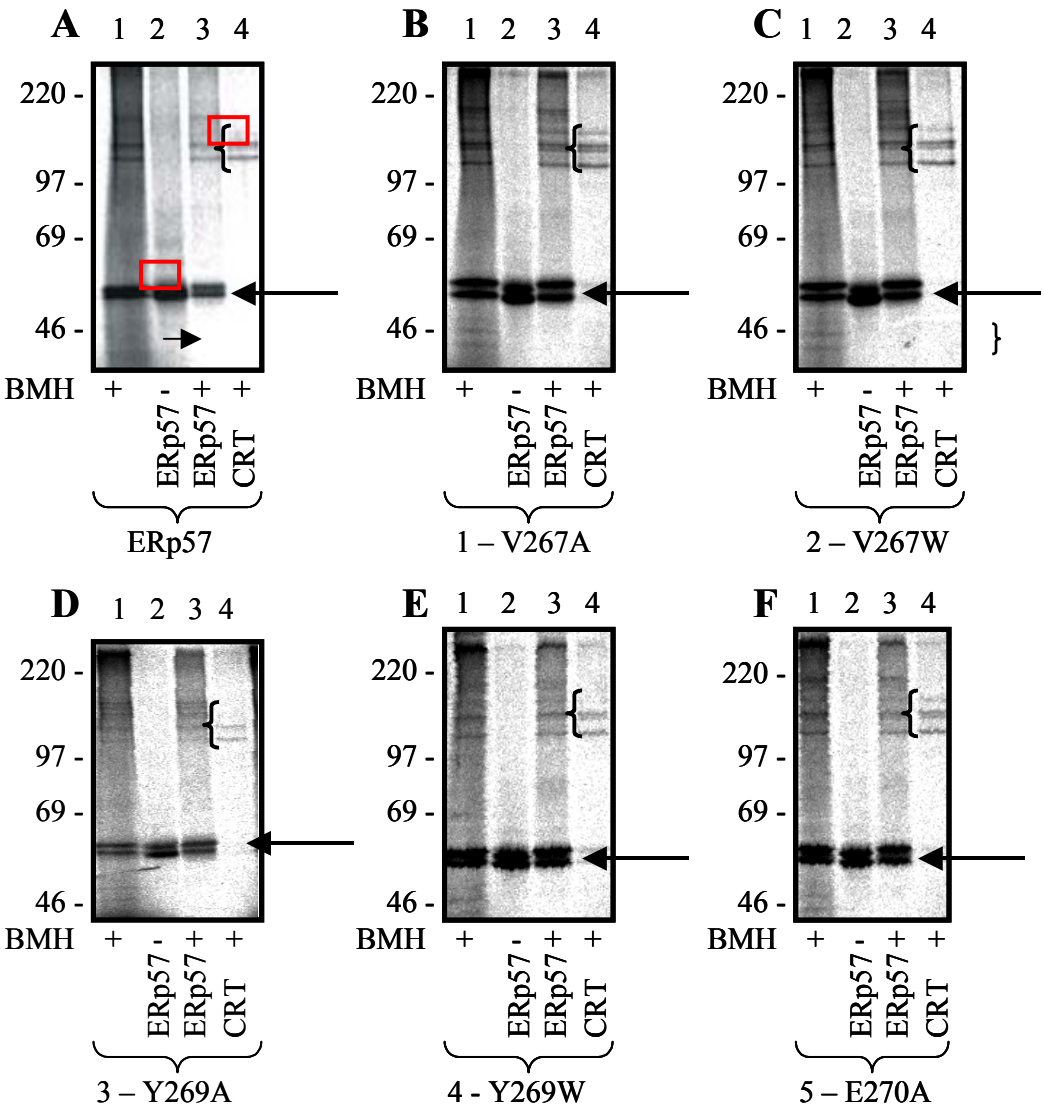


Figure 4.5 Interactions of wild-type and ERp57 *b'* domain point mutants (1-5) with calreticulin. The various mRNAs were translated in a rabbit reticulocyte lysate translation system in the presence of canine pancreatic microsomes. After termination of translation membrane-integrated material was isolated by centrifugation, and incubated with either 1mM BMH (+) or mock-treated by incubation with DMSO (-). The reaction was quenched with 5mM β-mercaptoethanol and samples were denatured with 1% SDS, diluted in 4 volumes of IP buffer, and incubated overnight with the anti-sera indicated: ERp57, anti-ERp57 (C-terminal); CRT, anti-calreticulin. Antibody complexes were isolated by incubating samples for 2 hours, rolling at 4ºC, with protein A sepharose pre-blocked with BSA. These samples together with an aliquot of the total products (lane 1) were analysed on 8% SDS-polyacrylamide gels. ERp57 polypeptides that have been imported but not cross-linked to endogenous components are denoted by arrows (→), whilst cross-linking products to calreticulin are shown by a (}). In the case of the wild-type protein, the 'total' ERp57 immunoprecipitated by the anti-ERp57 serum and the BMH-dependent ERp57/calreticulin adduct immunoprecipitated by the anti-calreticulin serum are indicated by red boxes. These products and their equivalents for the various ***b'*** mutants were quantified by phosphorimaging (see Figure 4.8).

Figure 4.6 Interactions of *in vitro* synthesised ERp57 ***b'*** domain point mutants (6-9) with calreticulin.

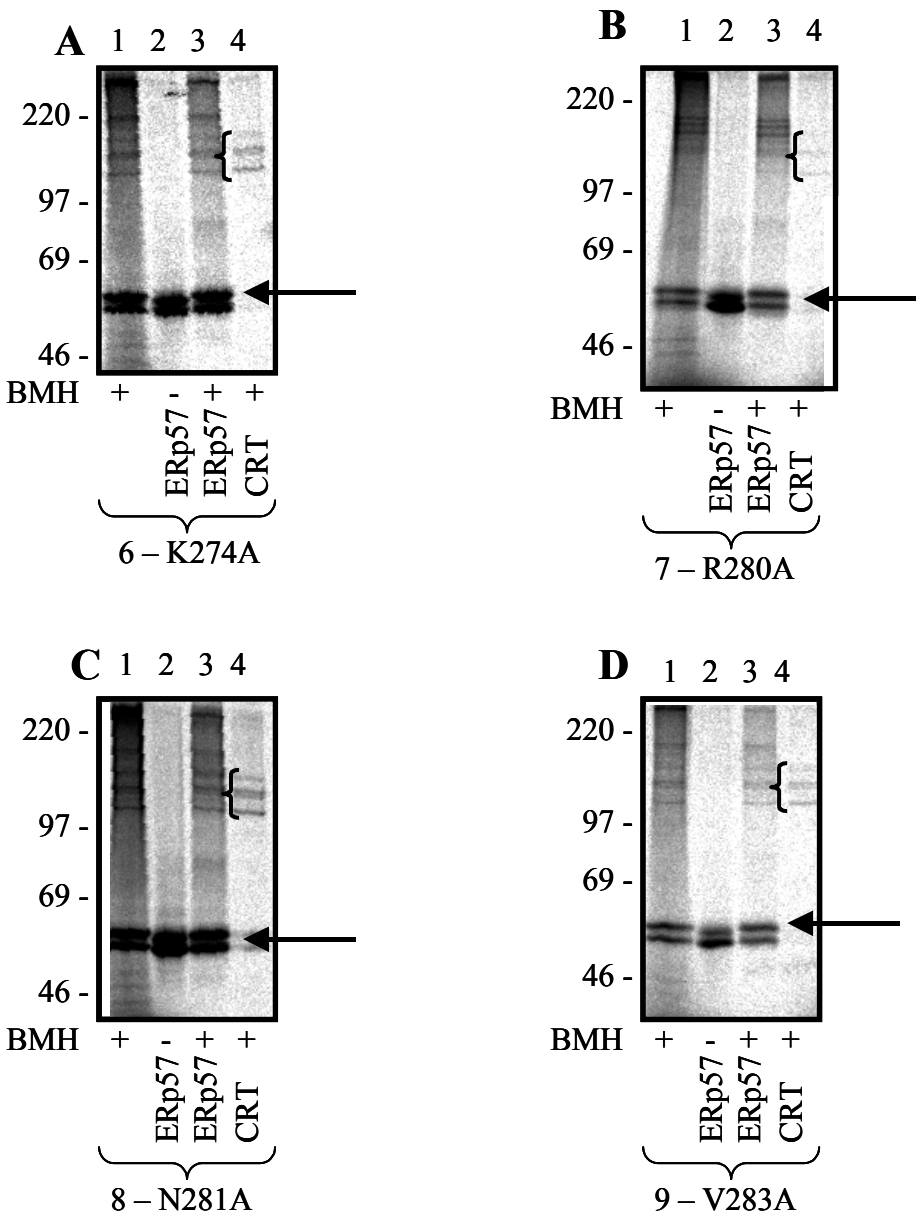


Figure 4.6 Interactions of *in vitro* synthesised ERp57 ***b'*** domain point mutants (6-9) with calreticulin. Experimental details, sample analysis and symbols are all as detailed in the legend to Figure 4.5.

Figure 4.7 Interactions of *in vitro* synthesised ERp57 ***b'*** domain point mutants (10-13) with calreticulin.

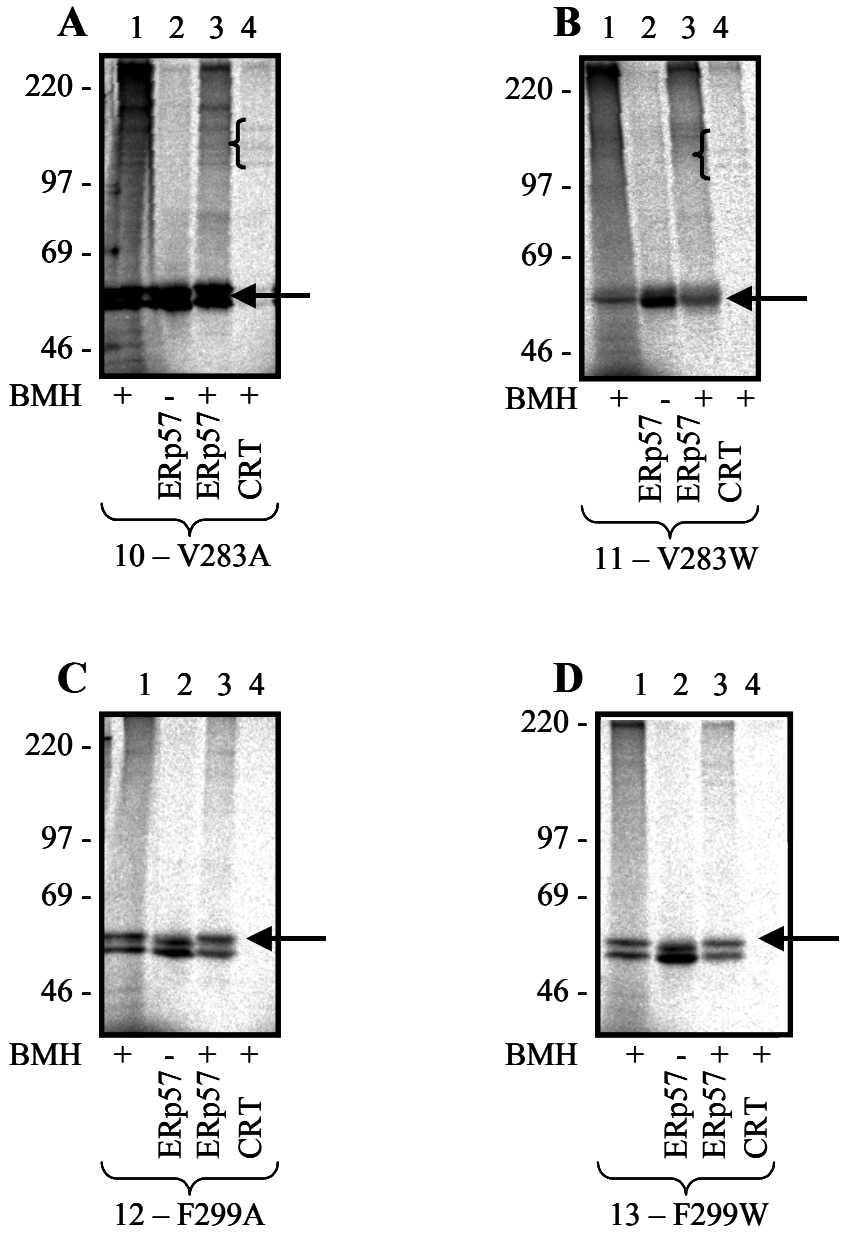


Figure 4.7 Interactions of *in vitro* synthesised ERp57 ***b'*** domain point mutants (10-13) with calreticulin. Experimental details, sample analysis and symbols are all as detailed in the legend to

Hence, the amount of total ERp57 immunoprecipitated by the anti-ERp57 serum and the amount of BMH-dependent ERp57/calreticulin adduct immunoprecipitated by the anti-calreticulin serum were both measured (see Figure 4.5, lanes 2 and 4, indicated by red boxes). This allowed the percentage of each ERp57 mutant that was cross-linked to calreticulin to be calculated and these figure could then be directly compared with one another in order to identify mutations that compromised binding (see Figure 4.8).

The use of quantitative phosphorimaging confirmed that there was a wide variation in the efficiency of cross-linking of ERp57 point mutants to calreticulin with the F299W mutant generating only approximately $^{1}/_{10}^{th}$ the amount of adduct seen with wild-type ERp57 (Figure 4.8). Several other point mutants showed adduct formation that was <50% of that seen with the wild-type protein whilst others were indistinguishable from the wild-type protein (Figure 4.8).

Figure 4.8 Percentage of radiolabelled ERp57 cross-linked to calreticulin

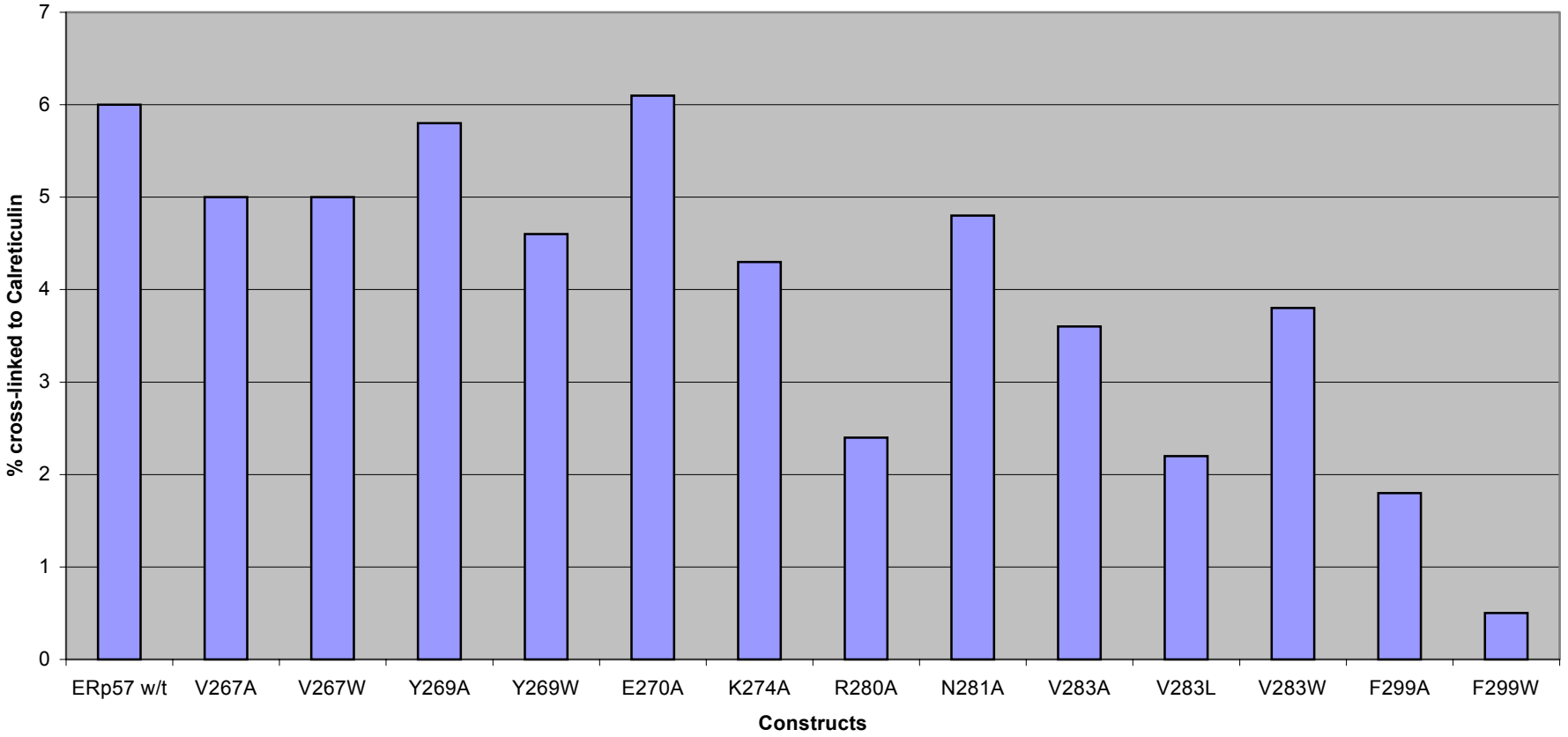


Figure 4.8 The data shown in Figures 4.5-4.7 has been analysed by quantitative phosphorimaging to determine the percentage of each ERp57 derivative that was cross-linked to calreticulin. The results are presented in the form of a bar graph allowing for a comparison of the various point mutants studied and the wild-type protein (ERp57 w/t). For details of the various point mutants refer to Figure 4.3.

4.4 Detailed analysis of selected ERp57 point mutants

The ERp57 point mutants are obviously non-natural polypeptides and the apparent effects of some of the mutations on ER lectin binding may be direct or indirect. Hence, the point mutants that showed a significant decrease in the detectable levels of association between ERp57 and calreticulin may be chronically misfolded. In order to address this issue, point mutants 7 (R280A), 12 (F299A) and 13 (F299W) that gave pronounced reductions in cross-linking to calreticulin were selected for a more detailed analysis in combination with the wild type protein.

The ERp57 ***b'*** domain point mutants 7 (R280A), 12 (F299A) and 13 (F299W) were imported into microsomes and subjected to BMH mediated cross-linking in parallel with wild type ERp57; cross-linking products were observed in all cases but with clear variation in efficiency (Figure 4.9A-D). In the case of wild-type ERp57, strong cross-linking products could be immunoprecipitated with anti-sera specific for ERp57 (Figure 4.9A; lane 3) and calreticulin (Figure 4.9A; lane 4}). As in previous experiments, multiple cross-linking products between calreticulin and ERp57 were observed (cf. Figure 4.9A, lanes 3 and 4}). A larger adduct of radiolabelled ERp57 cross-linked to endogenous calnexin was also immunoprecipitated by both the anti-ERp57 and anti-calnexin sera (cf. Figure 4.9A; lanes 3 and 5*), but not by the anti-calreticulin sera (Figure 4.9A, lane 4). No higher molecular weight products were immunoprecipitated in the absence of BMH treatment (Figure 4.9A, lane 2).

As in previous experiments (see Figure 4.4), I found no evidence of cross-linking between the newly synthesised wild type ERp57 and the endogenous ER chaperones/folding factors BiP and PDI (Figure 4.9A, lanes 6 and 7). This again suggested

that the wild type protein was folding correctly after import into the ER-derived microsomes and attaining a native structure.

A qualitative analysis of point mutants 7 (R280A), 12 (F299A) and 13 (F299W) suggested that the amount of cross-linking products formed with both calreticulin and calnexin were significantly less than those seen with the wild type protein (cf. Figure 4.9A-D lanes 3, } and 4*). This was entirely consistent with my previous analysis of adduct formation with calreticulin alone (see Chapter 4, Section 3 above). Furthermore, whilst all of the mutants were efficiently radiolabelled, no evidence of adducts with either BiP or PDI was obtained (Figures 4.9 A-D, lanes 6 and 7) suggesting that the point mutants were assuming a native like tertiary structure.

As previously, quantitative phosphorimaging was used to compare the amount of the radiolabelled ERp57 point mutants cross-linked to calreticulin and calnexin with that of the wild-type ERp57 protein. This quantitative analysis of the three ERp57 ***b'*** domain point mutants, 7 (R280A), 12 (F299A) and 13 (F299W) confirmed that the level of adduct formation with both calreticulin and calnexin was substantially reduced (Table 4.2).

From the data presented there is some indication that a single point mutation may have differential effects on the association of ERp57 with calreticulin and calnexin. However, it should be noted that the overall efficiency of cross-linking to calnexin was low, and hence the potential for experimental error proportionally higher. To draw firmer conclusions about such apparent differences in cross-linking efficiencies would require multiple repetitions of these experiments and subsequent statistical analysis. Unfortunately insufficient time was available to complete such a statically valid analysis.

Figure 4.9 Interaction of wild type ERp57 and selected ***b'*** domain point mutants with endogenous ER components.

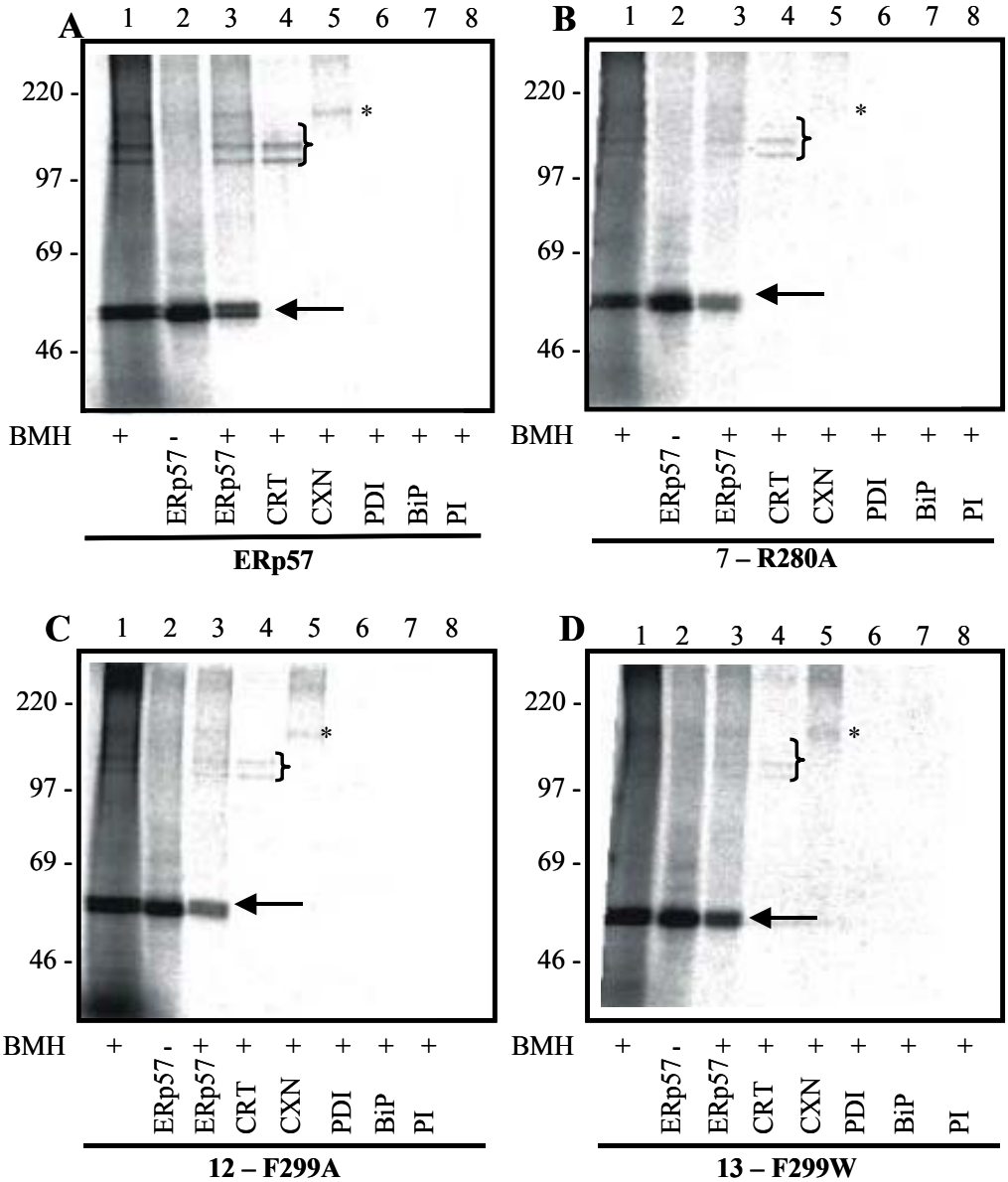


Figure 4.9 Interaction of wild type ERp57 and selected ***b'*** domain point mutants with endogenous ER components. The various mRNAs were translated in a rabbit reticulocyte lysate translation system in the presence of canine pancreatic microsomes. After termination of translation membrane-associated material was isolated by centrifugation, and incubated with either 1mM BMH (+) or mock-treated by incubation with DMSO(-). The reaction was quenched with 5mM β-mercaptoethanol and samples were denatured with 1% SDS, diluted in 4 volumes of IP buffer and incubated overnight with the anti-sera indicated: ERp57, anti-ERp57(C-terminal) ; CRT, anti-calreticulin. Antibody complexes were isolated by incubating samples for 2 hours, rolling at 4ºC, with protein A sepharose or a mix of protein A and G sepharose pre-blocked with BSA. The samples, including an aliquot of the total cross-linking reaction (lane 1), were analysed on an 8% SDS-polyacrylamide gel. Imported ERp57 polypetides that remain uncross-linked are shown by arrows (→), whilst adducts with calreticulin and calnexin are indicated by a bracket (}) and an asterix (*) respectively.

Table 4.2 Percentage of ERp57 point mutants cross-linked to calreticulin in comparison to wild-type ERp57.

Construct	% Cross-linked to calreticulin	Comparison with w/t ERp57 (%)	% Cross-linked to calnexin	Comparison with w/t ERp57 (%)
ERp57	5	100	1.4	100
Point mutant 7 (R280A)	2.5	50	0.5	36
Point mutant 12 (F299A)	1.8	36	1	71
Point mutant 13 (F299W)	0.5	10	0.9	64

The percentage of wild type and mutant versions of ERp57 cross-linked to the ER lectins calreticulin and calnexin. The figures were determined using quantitative phosphorimaging of the data presented in Figure 4.9.

4.5 Summary

In order to determine whether the ***b*** and ***b'*** domains of ERp57 are sufficient, as well as necessary, for the specific interaction of ERp57 with the ER lectins, calreticulin and calnexin, a number of ERp57 sub-domains were constructed. The ability of these ERp57 derivatives to interact with the ER lectins present in semi-intact cells was analysed by cross-linking. Only one of the four ERp57 sub-domains, ERp57 ***abb'***, showed a cross-linking pattern similar to wild-type ERp57. The other ERp57 fragments ***bb'x***, ***b'a'c***, ***b'x*** revealed a more complex cross-linking pattern, including adducts with PDI, and I concluded that these sub-domains were unable to fold correctly. Therefore, I was only able to conclude that the ***a***, ***b*** and ***b'*** fragment of ERp57 could function as a minimal element for ER lectin binding.

I found that the ***b'*** domain of ERp57 was present in all of the polypeptides that interacted with the ER lectins. Furthermore, the ***b'*** domain of PDI has been shown to be its principle peptide binding site (Klappa *et al.,* 1998) making its equivalent in ERp57 an obvious candidate for the site of ER lectin binding. With this in mind the interaction of ERp57 and the ER lectins was studied in more detail by using a number of ERp57 ***b'*** domain point mutants, based on the residues that create the peptide binding sites of PDI and PDIp. The ERp57 point mutants were imported into canine pancreatic microsomes and BMH-dependent adducts were identified by immunoprecipitation. Following immunoprecipitation the relative amount of ERp57 cross-linked to calreticulin was calculated by quantitative phosphorimaging. The amino acids, R280, V283 and F299, in the ***b'*** domain of ERp57 appear to be important for ER lectin binding site since the site-specific mutation of these amino acids resulted in a substantial decrease in the amount of ERp57 cross-linked to calreticulin.

Further cross-linking and immunoprecipitation analysis of the ERp57 point mutants R280A, F299A and F299W revealed that their cross-linking to calnexin was also decreased. Thus, amino acids F280 and F299 may form part of a specific calreticulin/calnexin-binding site present in ERp57.

4.6 Conclusion

The further analysis of the interaction between ERp57 and the ER lectins presented in this chapter shows that the ***a, b*** and ***b'*** domains of ERp57 are together sufficient for the specific interaction of the polypeptide with calreticulin and calnexin. My analysis of ERp57 ***b'*** domain point mutants, indicates that residues R280, V283, F299 may be important for the specific interaction of ERp57 with ER lectins and that they may form part of its calreticulin/calnexin-binding site.

CHAPTER - 5

Results

Direct analysis of ERp57 binding using a ‘pull-down’ assay

5 Direct analysis of ERp57 binding using a 'pull-down' assay

The data presented so far relied on cross-linking agents to stabilise the specific interactions between various ERp57 derivatives and the ER lectins calnexin and calreticulin. Using this cross-linking approach, I found that both the ***b*** and ***b'*** domains of ERp57 are necessary for its specific interaction with the ER lectins, since only polypeptides containing both of these domains are able to form cross-linking adducts with calreticulin and calnexin (see Chapter 3). However, these two domains alone may not be sufficient for the association with the ER lectins, and I found that the minimal requirement for the interaction was the ***a***, ***b*** and ***b'*** of ERp57 (see Chapter 4). Cross-linking analysis of a number of ERp57 ***b'*** domain point mutants, revealed that while several mutants exhibited adduct formation similar to that seen with the wild type protein, four mutants showed a marked decrease in the efficiency of adduct formation (see Chapter 4).

It was important to address the possibility that the variability in complex formation I observed may result from variations in cross-linking efficiency, rather than molecular differences in complex formation. For example, it is conceivable that the availability of amino acid side chains, or the conformation of different chimeras or point mutants, could reduce cross-linking efficiency without affecting complex formation. Therefore, in this chapter, the molecular basis for the interaction between ERp57 and the ER lectins was investigated using a distinct biochemical approach. For this reason, a 'pull-down' assay for ERp57 binding to immobilised calreticulin was fully established and optimised, and then used to investigate the binding capacity of the various ERp57 derivatives already described in the previous two chapters.

5.1 Components of the pull-down assay

A recombinant version of calreticulin had been previously constructed containing a naturally biotinylated purification tag, in place of the ER targeting signal normally present at the N-terminus of calreticulin. Recombinant calreticulin containing the N-terminal biotinylated tag was expressed in *E. coli*, isolated by binding to a monomeric avidin column, and eluted using free biotin. The purified biotin-tagged calreticulin was then immobilised on tetrameric NeutrAvidin coated beads and used in binding assays as the ligand for a variety of ERp57 derivatives that were generated by *in vitro* translation.

All of the ERp57 derivatives previously analysed by cross-linking, including ERp57/PDI chimeric constructs, ERp57 sub-domains and ERp57 ***b'*** domain point mutants (see Figures 3.1, 4.1 and 4.3), were used in conjunction with wild type ERp57 and PDI in the pull-down assay. For the purpose of this assay, the mRNA for each construct was translated in a flexi lysate system and the entire translation mixture added directly to samples of biotinylated calreticulin immobilised on NeutrAvidin beads. In contrast to the cross-linking experiments carried out in the presence of ER membrane, no signal sequence cleavage would occur in this system. Previous experiments had indicated that the presence of a signal sequence did not prevent ERp57 binding to calreticulin in solution (Oliver *et al.,* 1999). In order to establish whether the presence of a signal sequence influenced the ability of ERp57 to interact with the biotinylated calreticulin, I made an additional ERp57 derivative containing no signal sequence (see Chapter 2, Section 2.3). When translated this polypeptide contained an initiator Methionine residue followed by the mature region of ERp57.

5.2 Analysis of ERp57 binding to biotinylated-calreticulin

It has been suggested that calnexin and calreticulin may be able to function as 'general chaperones', and as such may be able to bind to partially folded proteins independently of any requirement for N-glycosylation (Ihara *et al.,* 1999; Saito *et al.,* 1999). This *in vitro* evidence is supported by data showing that calnexin can form complexes with non-glycosylated proteins *in vivo* when suitably 'mild' isolation procedures are used (Danilczyk and Williams, 2001). In the pull-down assay described here, radiolabelled ERp57 derivatives were synthesised in solution in the absence of ER lumenal folding factors and without signal sequence cleavage. Therefore, it was possible that a proportion of the ERp57 derived polypeptides might be unable to attain their native structure and could interact with biotinylated calreticulin as misfolded polypeptide substrates rather than as a consequence of authentic complex formation.

In order to address this issue I carried out a number of control experiments to ensure that the optimised pull-down assay conditions were both stringent and able to reflect specific interactions between ERp57 derivatives and biotin-tagged calreticulin. Firstly, native and 'scrambled' RNase A were used to compete with ERp57 for binding to the biotin-tagged-calreticulin/NeutrAvidin bead mixture. If the biotin-tagged calreticulin was acting as a 'general chaperone' under the conditions of the assay then the 'scrambled' RNase A would bind to the calreticulin and compete with the radiolabelled ERp57 whilst the native RNase A would not (cf. Klappa *et al.,* 1998). Secondly, to determine whether the presence of a signal sequence would affect the ability of ERp57 derivatives to interact with biotin-tagged calreticulin, wild type ERp57 lacking its signal sequence (ERp57Δss) was used. Finally, archetypal PDI was used in the same assay as a negative control. Since previous studies (Oliver *et al.,* 1999), and my own cross-linking analysis (see Chapters 3

and 4), indicated no interaction of PDI with calreticulin, its use would allow me to determine what level of non-specific binding between an *in vitro* translated protein and immobilised calreticulin could occur under the conditions of my assay.

Aliquots of NeutrAvidin beads were incubated in buffer, with biotin-tagged insulin or with biotin-tagged calreticulin (Figure 5.1A lanes 2-4 respectively) and washed prior to use in the pull-down assay. Additional samples were prepared using immobilised calreticulin that had been generated in the presence of a molar excess of either 'native' RNase A or 'scrambled' RNase A (Figure 5.1A, lanes 5 and 6 respectively). ^{35}S-labelled ERp57 was then added to the various samples and incubated for two hours. The beads were isolated and washed and the amount of ERp57 remaining bound at the end of the procedure estimated by SDS PAGE following the elution of bound material by its denaturation in the presence of SDS and DTT. The amount of ERp57 recovered, relative to the input (Figure 5.1A, lane 1) was estimated by quantitative phosphorimaging. As an additional control for binding specificity, the effects of adding a molar excess of polyhistidine tagged recombinant calreticulin and polyhistidine tagged recombinant PDI during the ERp57 binding step was also determined (Figure 5.1A, lanes 7 and 8).

A substantial fraction of ^{35}S-labelled ERp57 was bound to the beads loaded with biotin-tagged calreticulin but not to the control beads loaded with biotin-tagged insulin or unloaded beads (Figure 5.1A, cf. lanes 1-4). Quantification showed that ~33% of the total ^{35}S-labelled ERp57 present during the binding step remained bound to the calreticulin beads after isolation and extensive washing (Figure 5.2). In the absence of immobilised calreticulin, the amount of ERp57 recovered was ~3% of the total input for both the beads alone and the beads with biotinylated insulin bound (Figure 5.2). The pre-incubation of the beads in the presence of both biotinylated calreticulin and either scrambled or native RNase A did not have any substantial effect on ERp57 (~38% and 30% respectively, see

Figure 5.2). Thus, the binding of ERp57 appears to reflect the formation of a specific complex with immobilised calreticulin rather than the binding of a misfolded protein to a molecular chaperone. When an excess of unlabelled, recombinant his-tagged calreticulin was included during the incubation, this was able to efficiently compete for ERp57 binding reducing the amount of material recovered with the immobilised, biotinylated calreticulin to near background levels (~5%, see Figure 5.1A, lane 7 and Figure 5.2). When a similar molar excess of recombinant his-tagged PDI was added no competition for binding occurred and ~30% of the radiolabelled ERp57 was recovered (Figure 5.1 A, lane 8 and Figure 5.2).

Radiolabelled ERp57Δss and PDI were also synthesised *in vitro* and analysed for specific binding to immobilised calreticulin as well as levels of non-specific binding to immobilised insulin and the NeutrAvidin beads alone (Figure 5.1B and C). The ERp57Δss polypeptide showed efficient binding to the immobilised calreticulin but not to the beads alone or beads with insulin bound (Figure 5.1B, cf. lanes 2-4). Quantification showed that ~31% of ERp57Δss was recovered with the immobilised calreticulin fraction (Figure 5.2), and thus ERp57 binding in the presence and absence of the signal sequence was largely unaffected. The specificity of the interaction between radiolabelled ERp57 and immobilised, biotinylated calreticulin was further underlined by my observation that only ~2% of radiolabelled PDI was recovered in this pull-down assay under all three of the conditions tested (Figure 5.1C, Figure 5.2).

On the basis of this preliminary analysis I concluded that I could use the pull-down assay outlined above to detect the formation of specific complexes between immobilised calreticulin and radiolabelled ERp57 derivatives. Furthermore, this assay seemed well suited to a quantitative analysis of the efficiency of complex formation and was independent of any requirement for cross-linking. The specificity of the interaction was

essentially unaffected by the presence of the signal sequence (see also Oliver *et al.,* 1999) and the remainder of the experiments used full-length proteins with their signal sequences.

Figure 5.1 Binding of ERp57, ERp57Δss and PDI to immobilised calreticulin.

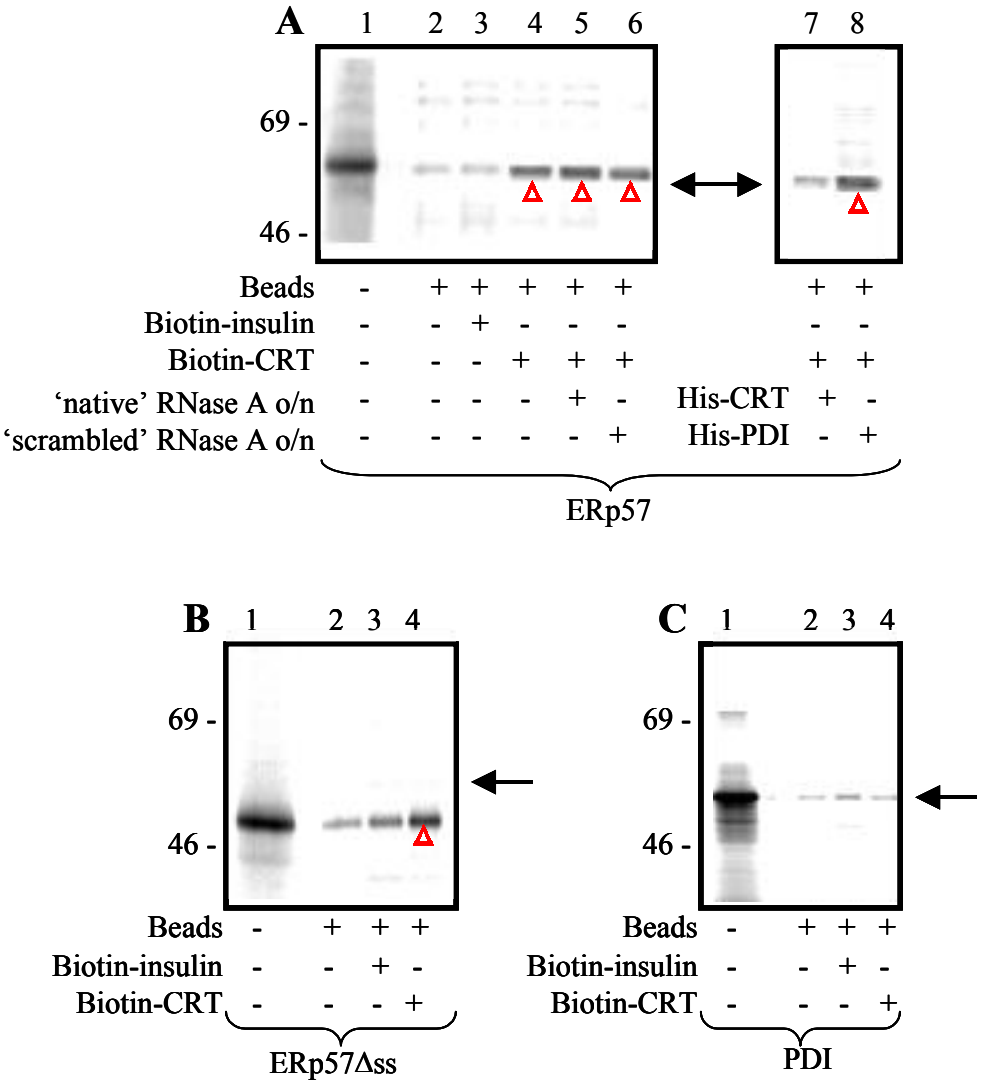


Figure 5.1 Binding of ERp57, ERp57Δss and PDI to immobilised calreticulin. ERp57 (Panel A), ERp57Δss (Panel B) and PDI (Panel C) were translated in the presence of reticulocyte lysate. ^{35}S-labelled ERp57, ERp57Δss and PDI were then mixed with NeutrAvidin beads alone (Panel A-C, lane 2), NeutrAvidin beads plus biotin-tagged insulin (Panel A-C, lane 3), NeutrAvidin beads plus biotin-tagged calreticulin (Panel A lanes 4, 7 and 8; Panels B-C lane 4) or NeutrAvidin beads plus biotin-tagged calreticulin pre-blocked with either 'native' or 'scrambled' RNase A (Panel A, lanes 5 and 6 respectively). His-tagged calreticulin and PDI were added in conjunction with radiolabelled ERp57 (samples 7 and 8) and the percentage recovery of ERp57 bound to biotin-tagged calreticulin was calculated. The beads were washed in 0.25% Triton X-100 in TBS and the bound material analysed on an 8 % SDS polyacrylamide gel followed by quantitative phophorimaging. An equal fraction of ERp57, ERp57Δss and PDI translation products was retained prior to incubation with NeutrAvidin beads and these were run in conjunction with the other samples (Panel A-C, lane 1). The 'specific' binding of radiolabelled translation products to biotin-tagged calreticulin is indicated by Δ.

Figure 5.2 ERp57, ERp57Δss and PDI binding to biotinylated-calreticulin in pull-down assays.

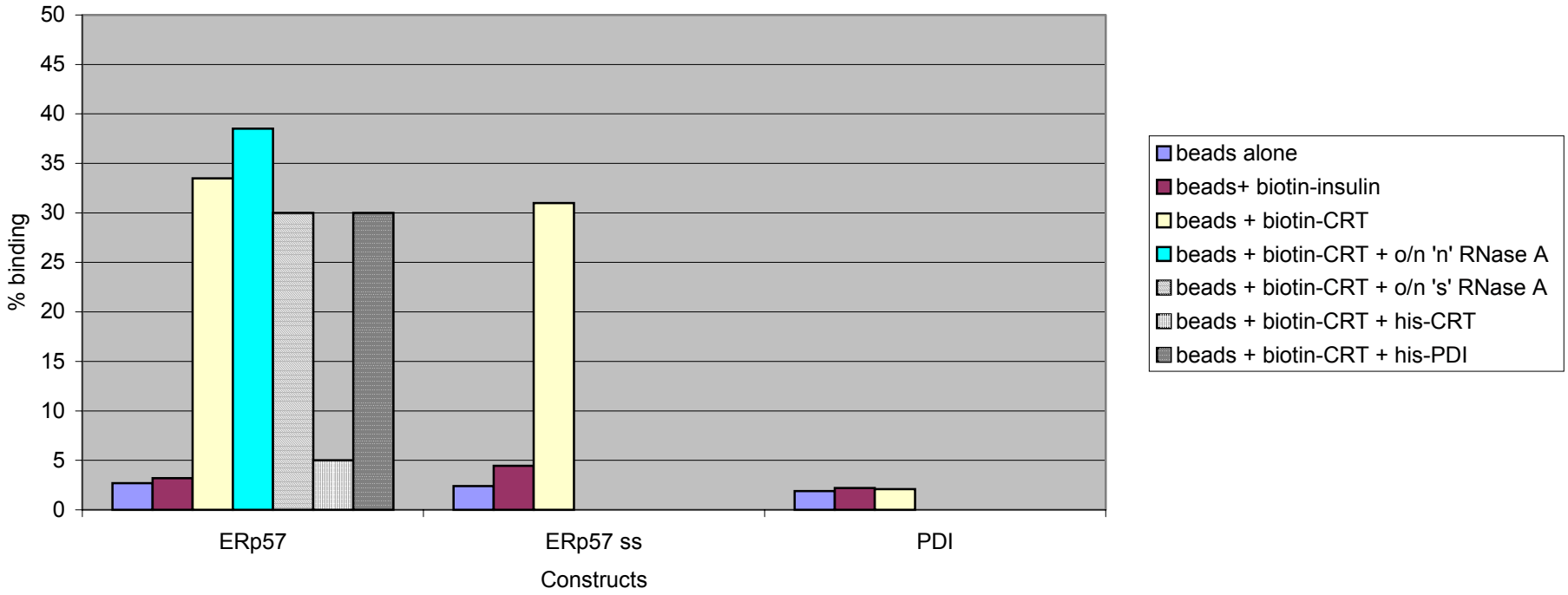


Figure 5.2 The data shown in Figure 5.1 was analysed by quantitative phosphorimaging to determine the percentage of radiolabelled ERp57 that was bound to biotin-tagged calreticulin under a number of experimental conditions (see legend and refer to Sections 2.10 and 5.2 for details of the experiment). The percentage of ERp57Δss and PDI that was bound to beads alone, biotinylated-insulin or biotinylated-calreticulin loaded beads is shown. The results presented are an average of 2 independent experiments one set of which are shown in Figure 5.1.

5.3 Binding of ERp57 chimeras to biotinylated-calreticulin

In order to investigate the molecular basis for the interaction of ERp57 and the ER lectins calnexin and calreticulin, a number of ERp57/PDI chimeric constructs and an ERp57 deletion construct (ERp57***abb'a'***) had been generated. When these various ERp57 derivatives were analysed by cross-linking, only ERp57Δc and chimeric constructs 3 (ERp57abb'PDIa'c) and 4 (PDI aERp57bb'a'c) were found to interact with both ER lectins (Chapter 3). This data indicated that both the ***b*** and ***b'*** domain of ERp57 were necessary for its specific interaction with calnexin and calreticulin and this hypothesis was now tested independently using a pull-down assay.

mRNAs encoding ERp57Δc (ERp57***abb'a'***) and various ERp57/PDI chimeras (constructs 2-7) were synthesised *in vitro* and translated using a rabbit reticulocyte lysate system. The radiolabelled polypeptides were then added to aliquots of NeutrAvidin beads alone, beads with biotin-tagged insulin or beads with biotin-tagged calreticulin bound. The beads were then washed and the samples analysed as described in Section 5.2 above.

As before, a substantial fraction of ERp57 was bound to the beads loaded with biotin-tagged calreticulin but not to the control beads loaded with biotin-tagged insulin or beads alone (cf. Figure 5.3A, lanes 2-4). Quantification showed that ~37% of the total ^{35}S-labelled ERp57 present during the binding step remained bound to the calreticulin beads after substantial washing (Figure 5.6). When the ERp57Δc polypeptide and the various ERp57/PDI chimeras were tested using the pull-down assay it was immediately apparent that only a subset of these proteins displayed a strong signal with the immobilised calreticulin compared to a low level background signal with the control baits (Figures 5.3B-D and 5.4A-D, cf. lanes 2-4). Thus, only ERp57Δc and chimeras 3 and 4 showed levels of specific binding that appeared comparable with wild type ERp57 (see Figures

5.3A, 5.3B, 5.3D and 5.4A, lane 4, Δ). No such binding was detected with constructs 2, 5, 6 or 7 (Figures 5.3C and 5.4B-D respectively). These results are consistent with the data obtained using the cross-linking reagents BMH and SMCC (Chapter 3), and show that of the ERp57 chimeras analysed, only construct 3 and construct 4 can associate strongly with the ER lectin calreticulin. This indicates that both the ***b*** and ***b'*** domains of ERp57 are required for efficient binding to calreticulin.

Upon quantification of the data presented in the Figure 5.3 and 5.4 it became clear that whilst wild-type ERp57, ERp57Δc and chimeras 3 and 4 all showed specific binding to calreticulin, there was a clear and reproducible difference in their efficiencies of binding as measured by the pull-down assay. In fact, only construct 4 showed binding equivalent to wild-type ERp57 (~37% and 40% respectively), while the binding of ERp57Δc was clearly reduced (~29%) and that of construct 3 significantly lower (~23%; see Figure 5.7). The clear variation in the binding ability of chimeric constructs 3 and 4 suggests that the ***a'*** domain of ERp57 significantly enhances its binding to calreticulin, whilst the reduced efficiency of the ERp57Δc protein indicates that the C-terminal extension of ERp57 may also facilitate the binding of ERp57 to calreticulin.

Figure 5.3 Binding of ERp57, ΔERp57 and ERp57/PDI chimeras 2 and 3 to immobilised calreticulin

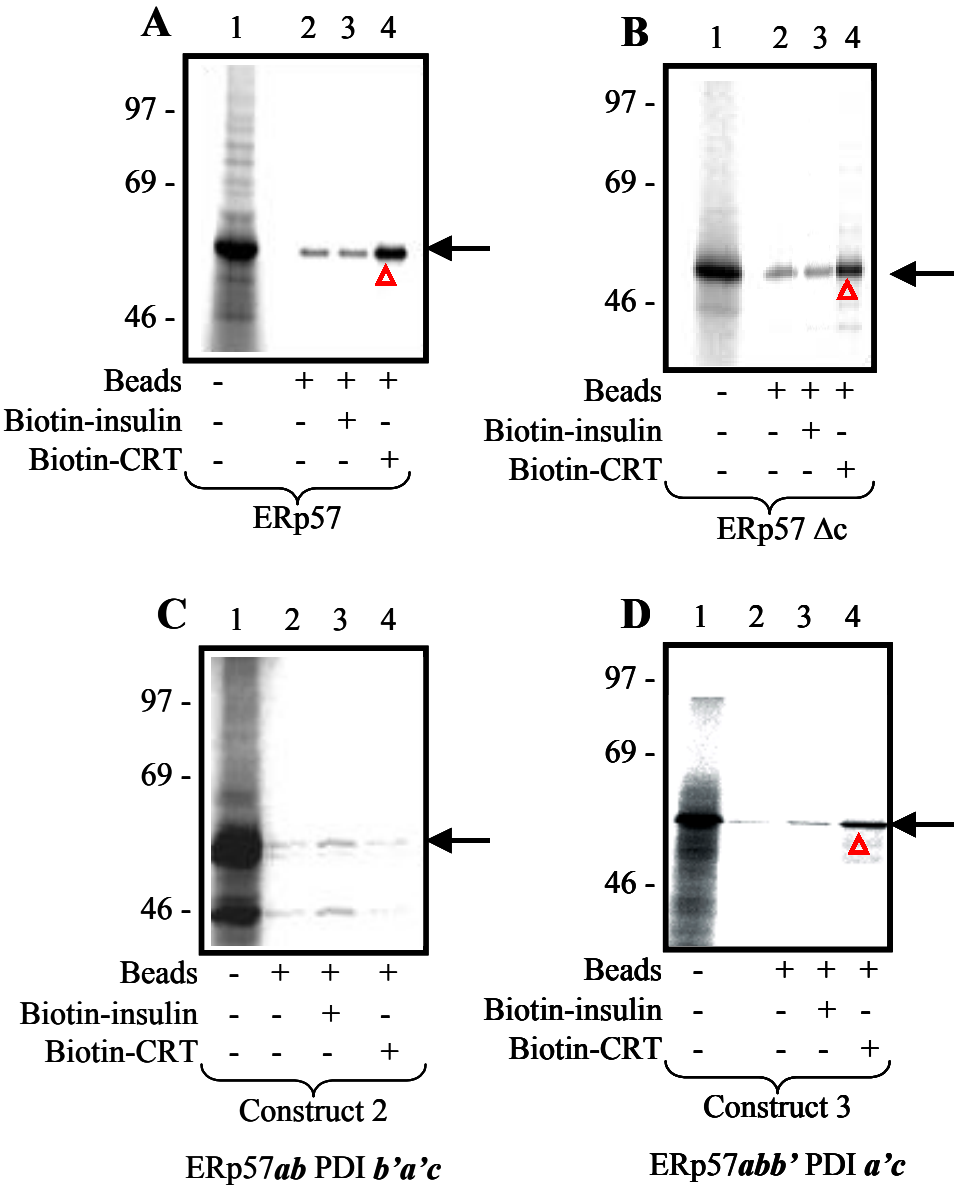


Figure 5.3 Binding of ERp57, ΔERp57 and ERp57/PDI chimeras to immobilised calreticulin. ERp57 (Panel A), ΔERp57 (Panel B), construct 2 (ERp57***ab***PDI***b'a'c***; Panel C) and construct 3 (ERp57***abb'***PDI***a'c***; Panel D) were translated in the presence of reticulocyte lysate. ^{35}S-labelled ERp57, ΔERp57, construct 3 and 4 were then mixed with NeutrAvidin beads alone (Panel A-D, lane 2), NeutrAvidin beads plus biotin-tagged insulin (Panel A-D, lane 3), NeutrAvidin beads plus biotin-tagged calreticulin (Panel A-D, lane 4). The beads were washed in 0.25% Triton X-100 in TBS and the bound material analysed on an 8% SDS polyacrylamide gel followed by quantitative phosphorimaging. An equal fraction of ERp57, ΔERp57, construct 3 and 4 translation products was retained prior to incubation with NeutrAvidin beads and these were run in conjunction with the other samples (Panel A-D, lane 1). The 'specific' binding of radiolabelled translation products to biotin-tagged calreticulin is indicated by Δ.

Figure 5.4 Binding of ERp57/PDI chimeras (4-7) to immobilised calreticulin

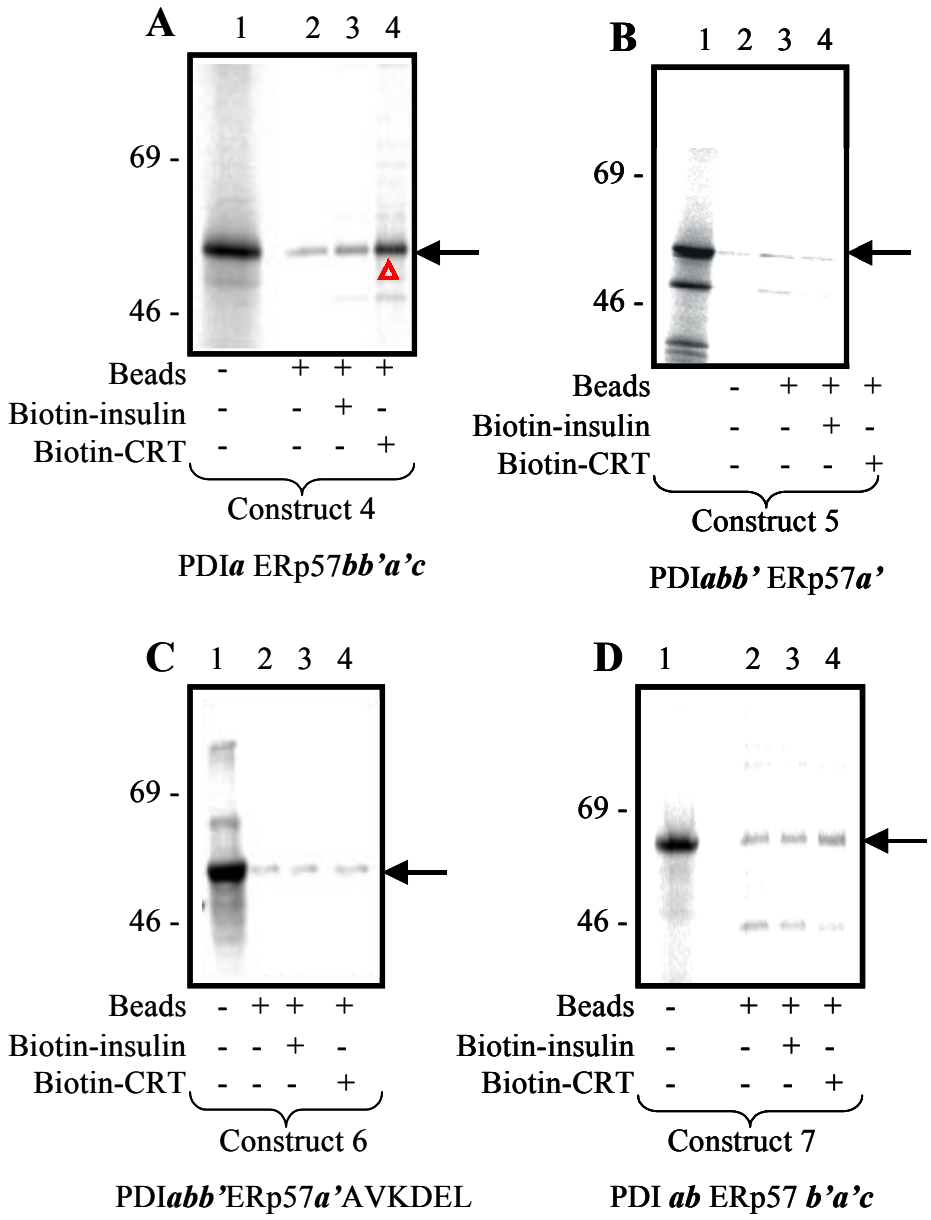


Figure 5.4 Binding of ERp57/PDI chimeras to immobilised calreticulin. Construct 4 (PDI***a***ERp57***bb'a'c***; Panel A), construct 5 (PDI***abb'***ERp57***a'***; Panel B), construct 6 (PDI***abb'***ERp57***a'***AVKDEL; Panel C) and construct 7 (PDI***ab***ERp57***b'a'c***; Panel D) were translated in the presence of reticulocyte lysate. ^{35}S-labelled construct 4, 5, 6 and 7 were then mixed with NeutrAvidin beads alone (Panel A-D, lane 2), NeutrAvidin beads plus biotin-tagged insulin (Panel A-D, lane 3), NeutrAvidin beads plus biotin-tagged calreticulin (Panel A-D, lane 4). The beads were washed in 0.25% Triton X-100 in TBS and the bound material analysed on an 8% SDS polyacrylamide gel followed by quantitative phosphorimaging. An equal fraction of construct 4, construct 5, construct 6 and construct 7 translation products was retained prior to incubation with NeutrAvidin beads and these were run in conjunction with the other samples (Panel A-D, lane 1). The 'specific' binding of radiolabelled translation products to biotin-tagged calreticulin is indicated by Δ.

5.4 Binding of ERp57 sub-domains to biotinylated-calreticulin

In order to identify the region or regions of ERp57 that are necessary and sufficient to form specific complexes with calnexin and calreticulin several ERp57 sub-domains were generated (Chapter 4, Section 4.2). I had previously found that, of these ERp57 sub-domain constructs, only the ERp57 ***abb'*** region produced a cross-linking pattern similar to the wild-type protein (Chapter 4). In order to analyse these ERp57 sub-domain polypeptides further, their ability to bind immobilised calreticulin was analysed using the pull-down assay as already described (see Sections 5.2 and 5.3).

It was immediately clear that none of the ^{35}S-labelled ERp57 sub-domains were able to bind efficiently to biotin-tagged calreticulin, and levels of binding appeared similar to control baits (Figure 5.5A-D cf. lanes 2-4). This qualitative impression was confirmed by a quantitative analysis of the ERp57 sub-domain recovery in the pull-down assay, and in no case was the amount of radiolabelled polypeptide recovered above the level of background (Figure 5.6). Thus, the pull-down assay confirmed that the ERp57***b'a'c***, ***bb'x***, ***b'x*** domains do not interact with calreticulin when present alone (cf. cross-linking data presented in Chapter 4). This pull-down analysis also challenges the cross-linking data obtained with the ***abb'*** sub-domain (Chapter 4), and indicates that this region does not stably associate with immobilised calreticulin unless the interaction is stabilised by covalent cross-linking.

Figure 5.5 Binding of ERp57 sub-domain derivatives to immobilised calreticulin

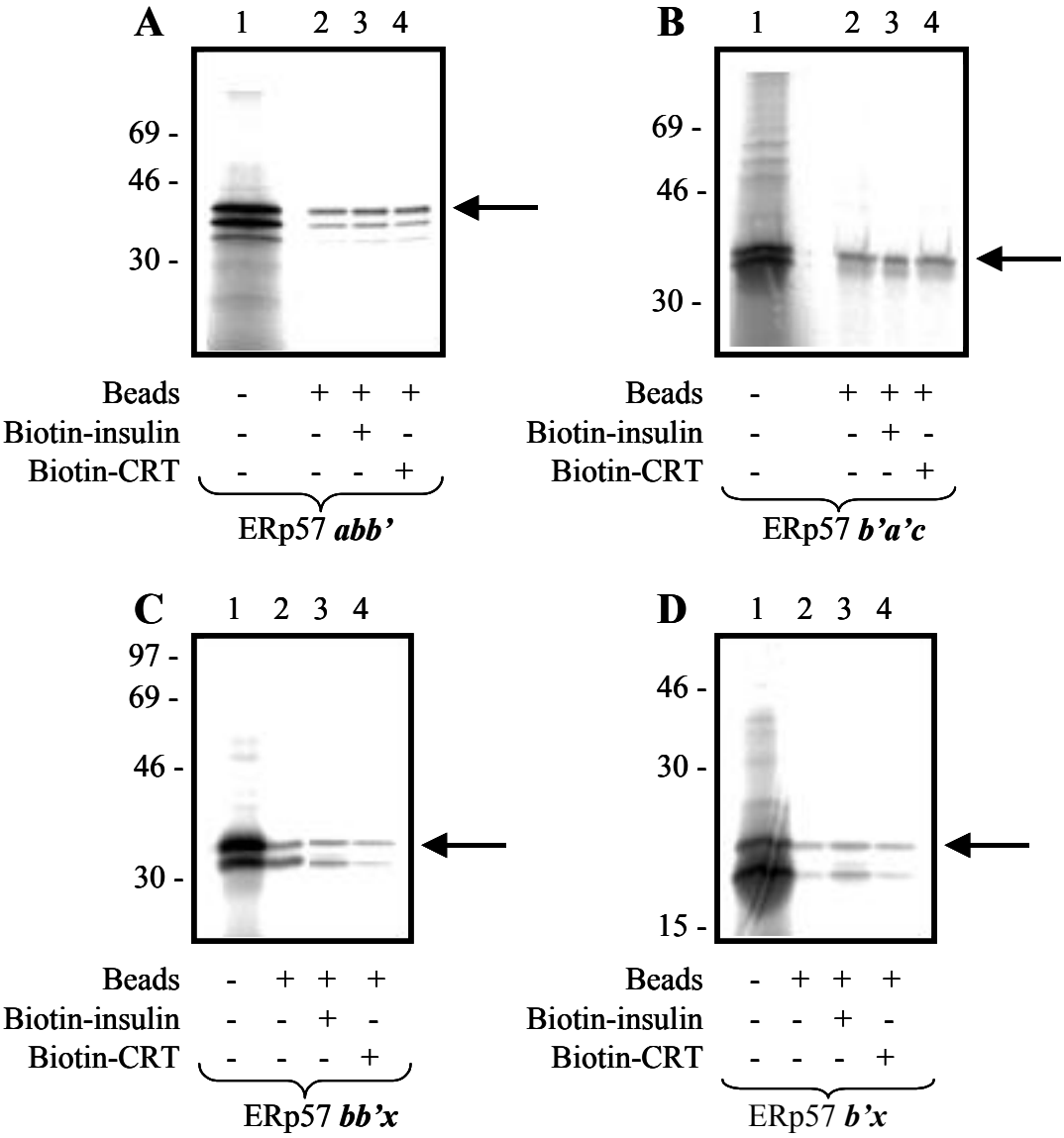


Figure 5.5 Binding of ERp57 sub-domain derivatives to immobilised calreticulin. ERp57***abb'*** (Panel A), ERp57***b'a'c*** (Panel B), ERp57***bb'x*** (Panel C) and ERp57***b'x*** (Panel D) were translated in the presence of reticulocyte lysate. ^{35}S-labelled ERp57***abb'***, ERp57***b'a'c***, ERp57***bb'x*** and ERp57***b'x*** were then mixed with NeutrAvidin beads alone (Panel A-D, lane 2), NeutrAvidin beads plus biotin-tagged insulin (Panel A-D, lane 3), NeutrAvidin beads plus biotin-tagged calreticulin (Panel A-D, lane 4). The beads were washed in 0.25% Triton X-100 in TBS and the bound material analysed on a 10-15% SDS polyacrylamide gel followed by quantitative phosphorimaging. An equal fraction of ERp57***abb'***, ERp57***b'a'c***, ERp57***bb'x*** and ERp57***b'x*** translation products was retained prior to incubation with NeutrAvidin beads and these were run in conjunction with the other samples (Panel A-D, lane 1).

Figure 5.6 Summary of ERp57 chimeras and sub-domain derivatives binding data in biotinylated-calreticulin pull-down assays.

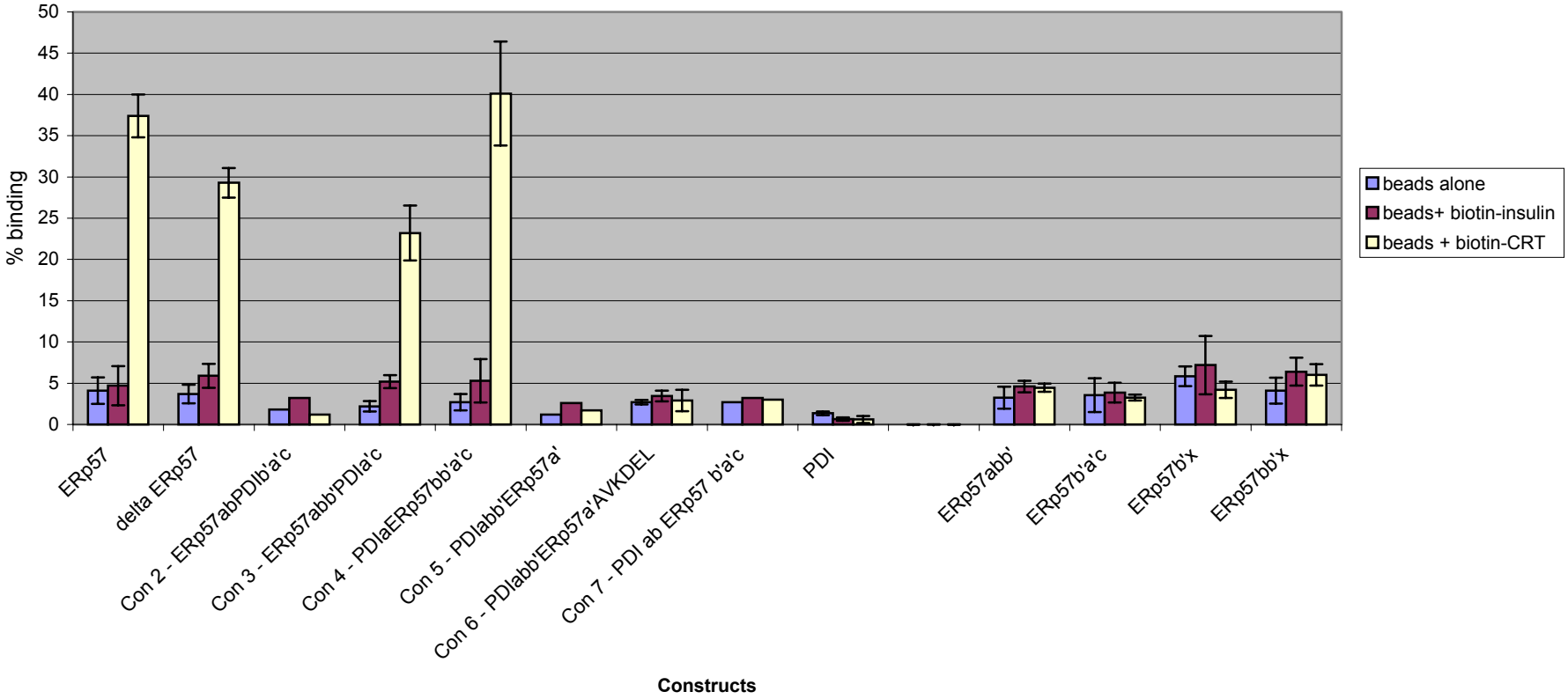


Figure 5.6 Data shown in Figures 5.3-5.5 has been analysed by quantitative phosphorimaging to determine the percentage of wild-type ERp57, ΔERp57, ERp57/PDI chimeras and ERp57 sub-domain derivatives bound to biotinylated calreticulin. The results are presented as a bar graph showing an average of at least 3 experiments with standard deviation. The specific experimental conditions are detailed in the appropriate figure legends (see Figures 5.3-5.5).

5.5 Binding of ERp57 point mutants to biotinylated-calreticulin

In the Chapter 4, the potential role of the ***b'*** domain of ERp57 in mediating its specific association with the ER lectins calnexin and calreticulin was investigated. This was achieved by using a series of specific point mutants (see Table 4.1 and Figure 4.2 and 4.3) and analysing their association with the ER lectins by cross-linking. Given the successful analysis of the various ERp57 chimeras in the pull-down assay already described above, I next investigated the association of each of the ERp57 ***b'*** domain point mutants with immobilised calreticulin using exactly the same pull-down assay.

Significantly, the data obtained using the pull-down assay revealed clear discrepancies with the previous cross-linking analysis of the ERp57 ***b'*** domain point mutants. Thus, whilst several of the point mutants showed cross-linking to calreticulin of a similar intensity to the wild type protein (see Figures 4.8 and 4.9), most of the mutants appeared to be severely compromised in the pull-down assay (Figure 5.7-5.10). When the relative efficiencies for the binding of the various ERp57 ***b'*** domain point mutants were quantified as previously described, it was apparent that only one of the thirteen point mutants analysed, 'E270A', bound to immobilised calreticulin with the same efficiency as the wild type protein (~35% and ~37% respectively, see Figure 5.11). A second mutant, 'V283A', exhibited an intermediate level of binding (~20%, Figure 5.11). Whilst three other mutants, V267A', 'Y269W' and 'V283L', all showed levels of binding that were significantly above their matched controls, but that were substantially reduced in comparison to the wild type protein (~12%, ~15% and ~12% respectively, Figure 5.11). All of the other point mutants displayed levels of binding to immobilised calreticulin that were not significantly above the background level (Figure 5.11). Taken together, these data suggest that the ***b'*** domain of ERp57 may be required for its stable association with

the ER lectins calreticulin and calnexin, and that specific perturbations in this region can effectively abolish ER lectin binding.

Figure 5.7 Binding of wild-type ERp57 and ERp57 ***b'*** domain point mutants (1-3) to immobilised calreticulin

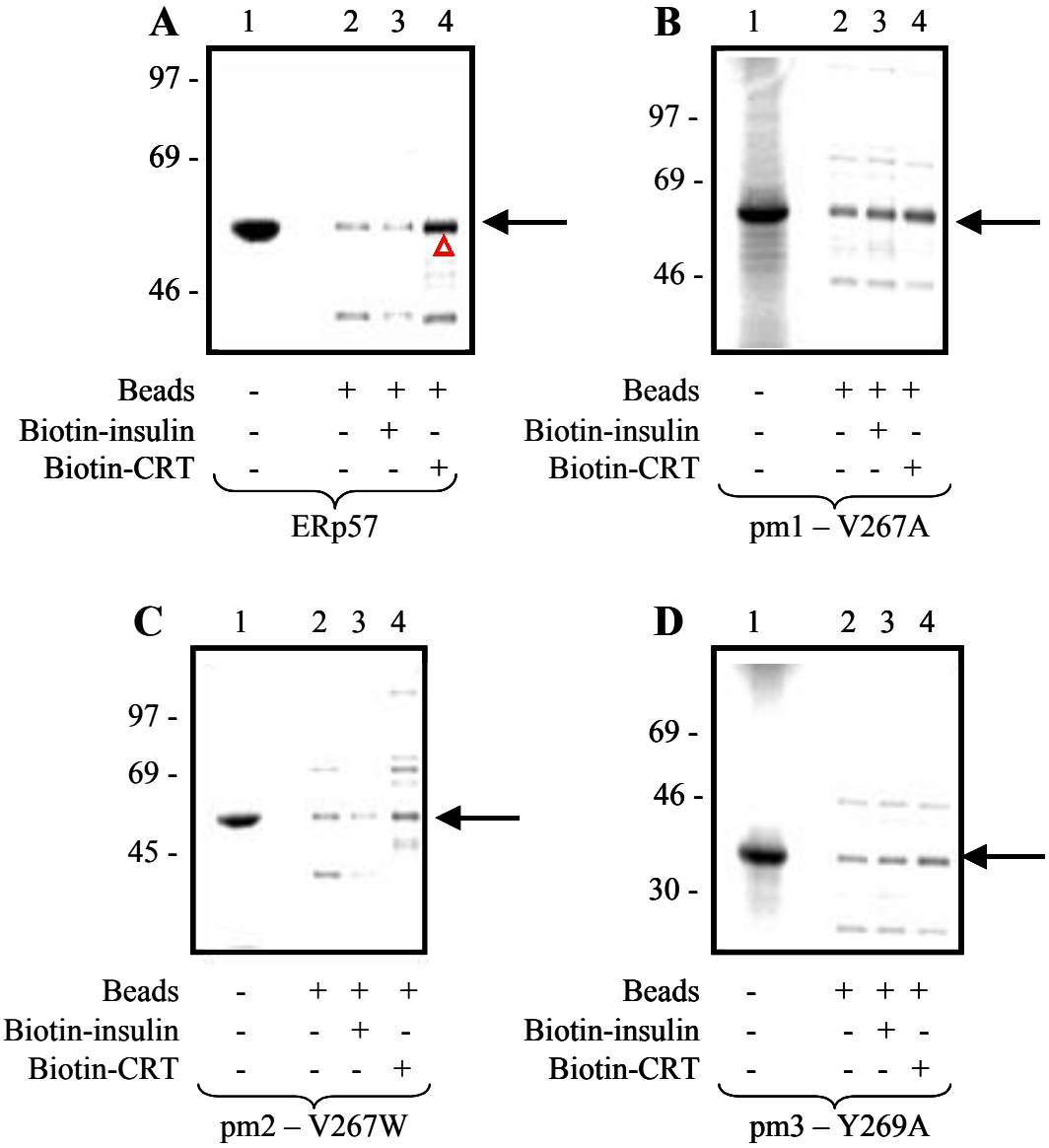


Figure 5.7 Binding of wild-type ERp57 and ERp57 ***b'*** domain point mutants to immobilised calreticulin. ERp57 (Panel A), point mutant 1–V267A (Panel B), point mutant 2-V267W (Panel C) and point mutant 3-Y269A (Panel D) were translated in the presence of reticulocyte lysate. ^{35}S-labelled wild-type ERp57, point mutant 1, 2 and 3 were then mixed with NeutrAvidin beads alone (Panel A-D, lane 2), NeutrAvidin beads plus biotin-tagged insulin (Panel A-D, lane 3), NeutrAvidin beads plus biotin-tagged calreticulin (Panel A-D, lane 4). The beads were washed in 0.25% Triton X-100 in TBS and the bound material analysed on an 8% SDS polyacrylamide gel followed by quantitative phosphorimaging. An equal fraction of each translation products was retained prior to incubation with NeutrAvidin beads and these were run in conjunction with the other samples (Panel A-D, lane 1). The efficient binding of radiolabelled translation products to biotin-tagged calreticulin is indicated by Δ.

Figure 5.8 Binding of ERp57 ***b'*** domain point mutants (4-7) to immobilised calreticulin

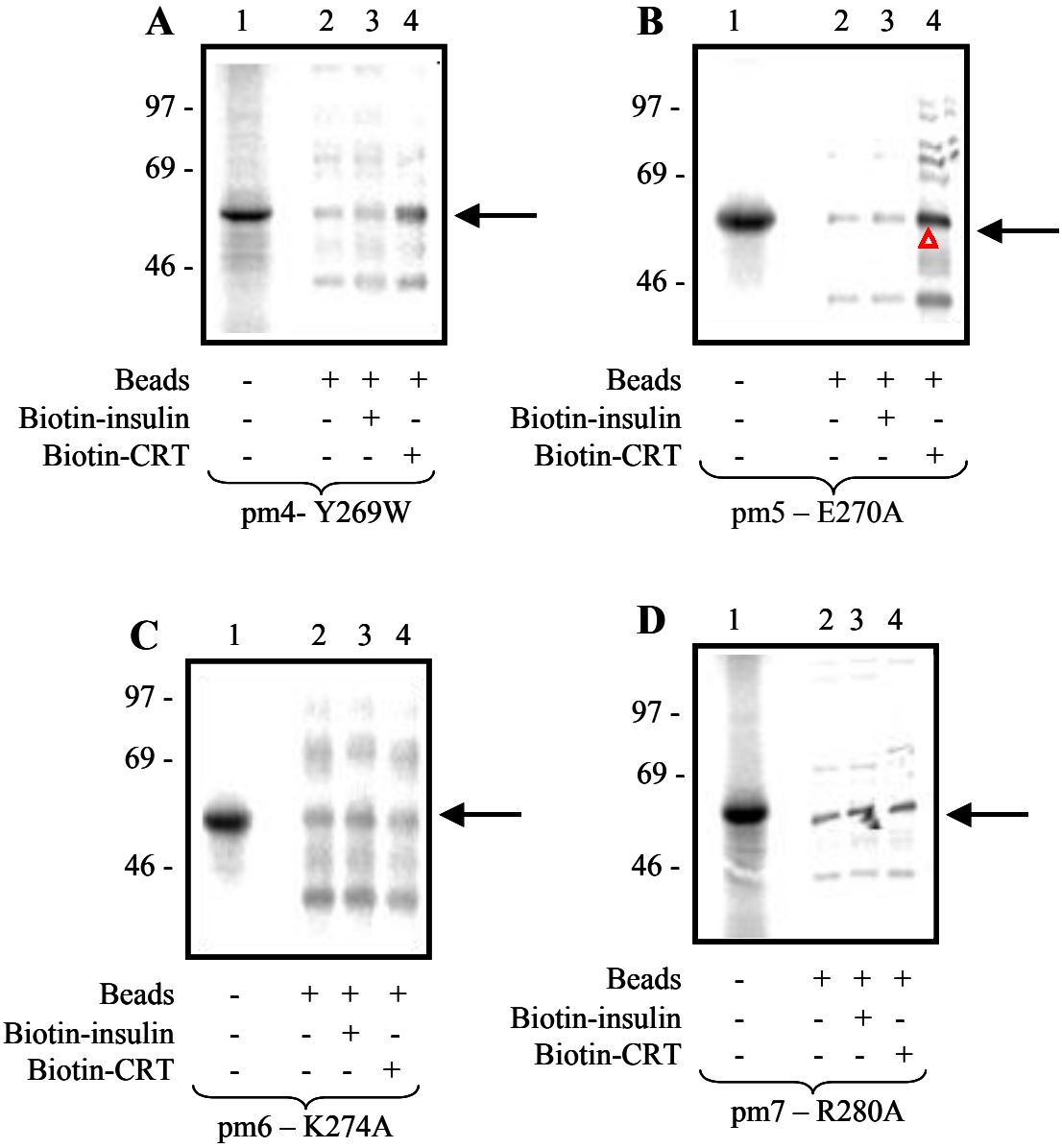


Figure 5.8 Binding of ERp57 ***b'*** domain point mutants to immobilised calreticulin. Point mutant 4-Y269W (Panel A), point mutant 5–E270A (Panel B), point mutant 6-K274A (Panel C) and point mutant 7-R280A (Panel D) were translated in the presence of reticulocyte lysate. ^{35}S-labelled point mutants 4-7 were then mixed with NeutrAvidin beads alone (Panel A-D, lane 2), NeutrAvidin beads plus biotin-tagged insulin (Panel A-D, lane 3), NeutrAvidin beads plus biotin-tagged calreticulin (Panel A-D, lane 4). The beads were washed in 0.25% Triton X-100 in TBS and the bound material analysed on an 8% SDS polyacrylamide gel followed by quantitative phosphorimaging. An equal fraction of each translation products was retained prior to incubation with NeutrAvidin beads and these were run in conjunction with the other samples (Panel A-D, lane 1). The efficient binding of radiolabelled translation products to biotin-tagged calreticulin is indicated by Δ.

Figure 5.9 Binding of ERp57 ***b'*** domain point mutants (8-11) to immobilised calreticulin

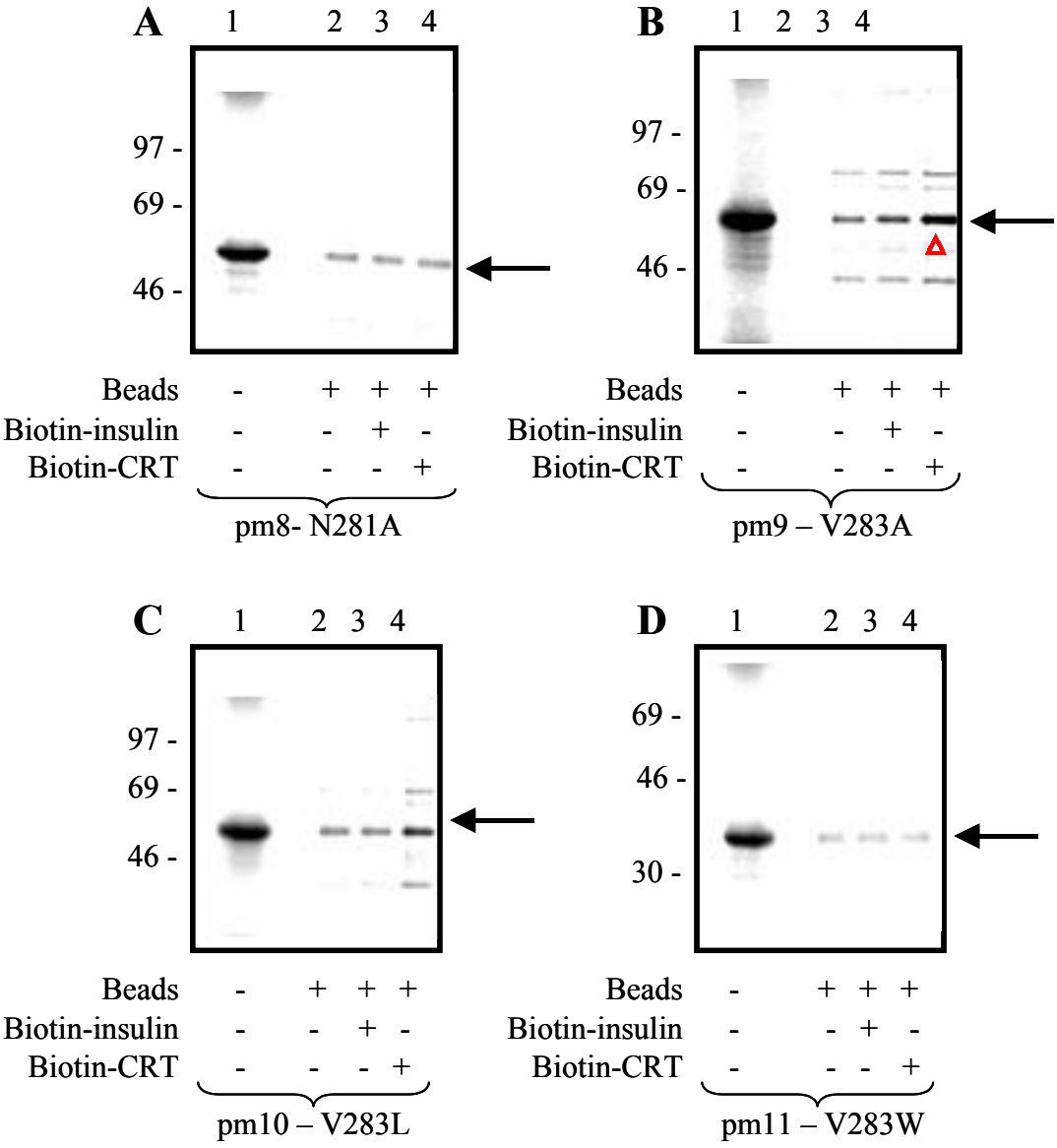


Figure 5.9 Binding of ERp57 ***b'*** domain point mutants to immobilised calreticulin. Point mutant 8-N281A (Panel A), point mutant 9-V283A (Panel B), point mutant 10-V283L (Panel C) and point mutant 11-V283W (Panel D) were translated in the presence of reticulocyte lysate. ^{35}S-labelled point mutants 8-11 were then mixed with NeutrAvidin beads alone (Panel A-D, lane 2), NeutrAvidin beads plus biotin-tagged insulin (Panel A-D, lane 3), NeutrAvidin beads plus biotin-tagged calreticulin (Panel A-D, lane 4). The beads were washed in 0.25% Triton X-100 in TBS and the bound material analysed on an 8% SDS polyacrylamide gel followed by quantitative phosphorimaging. An equal fraction of each translation products was retained prior to incubation with NeutrAvidin beads and these were run in conjunction with the other samples (Panel A-D, lane 1). The efficient binding of radiolabelled translation products to biotin-tagged calreticulin is indicated by Δ.

Figure 5.10 Binding of ERp57 ***b'*** domain point mutants 12 and 13 to immobilised calreticulin

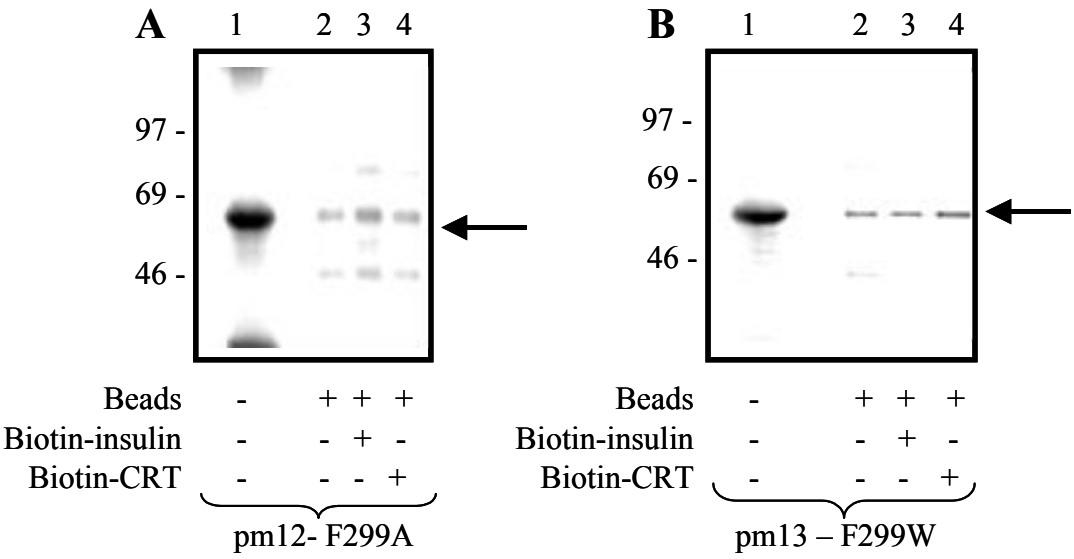


Figure 5.10 Binding of ERp57 ***b'*** domain point mutants to immobilised calreticulin. Point mutant 12-F299A (Panel A), point mutant 13-F299W (Panel B) were translated in the presence of reticulocyte lysate. ^{35}S-labelled point mutants 12 and 13 were then mixed with NeutrAvidin beads alone (Panel A-D, lane 2), NeutrAvidin beads plus biotin-tagged insulin (Panel A-D, lane 3), NeutrAvidin beads plus biotin-tagged calreticulin (Panel A-D, lane 4). The beads were washed in 0.25% Triton X-100 in TBS and the bound material analysed on an 8% SDS polyacrylamide gel followed by quantitative phosphorimaging. An equal fraction of each translation products was retained prior to incubation with NeutrAvidin beads and these were run in conjunction with the other samples (Panel A-D, lane 1).

Figure 5.11 Summary of ERp57 ***b'*** domain point mutant binding data in biotinylated-calreticulin pull-down assay

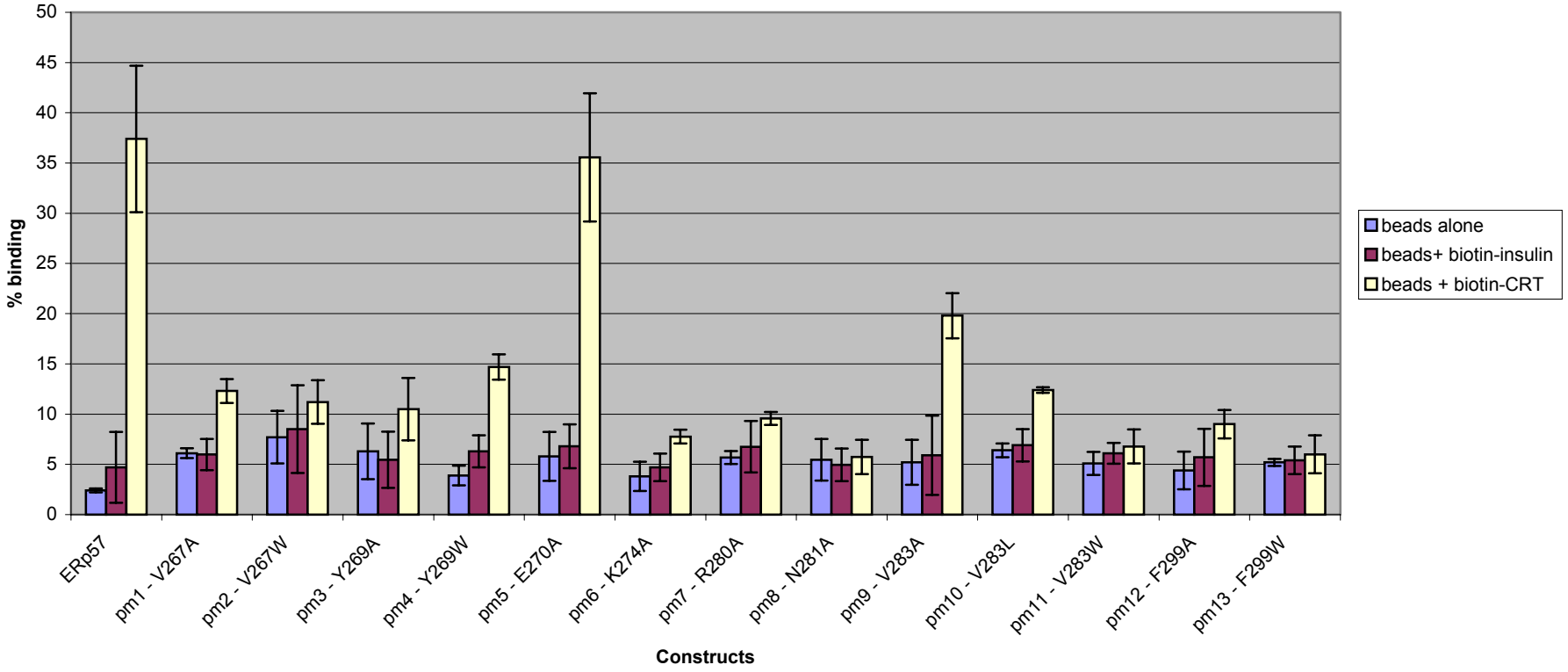


Figure 5.11 Data shown in Figures 5.7-5.10 has been analysed by quantitative phosphorimaging to determine the percentage of wild-type ERp57 and ERp57 ***b'*** domain point mutant derivatives bound to biotinylated calreticulin. The results are presented as a bar graph showing an average of at least 3 experiments with standard deviation. The specific experimental conditions are detailed in the appropriate figure legends (see Figures 5.7-5.10).

5.6 Summary

It was possible that differences in the assembly between various ERp57 derivatives and ER lectins that I could detect by cross-linking, did not fully reflect the ability of all of the ERp57 derivatives to interact with calnexin and calreticulin. In particular, the absence of a cross-linking product alone may not reflect the absence of a stable interaction and the cross-linking approach might have 'missed' ERp57 derivatives that bound to the ER lectins in a way that could not be detected by this approach. For this reason, I established a second, independent assay for ERp57 binding to the ER lectins in the form of a pull-down assay. The data I obtained with the pull-down assay showed that in every case, the ER derivatives that did not cross-link to the ER lectins did not bind to immobilised calreticulin. Furthermore, some of the ERp57 derivatives that did yield ER lectin cross-linking products did not bind to immobilised calreticulin in the pull-down assay, and those that did showed substantial differences in the efficiency of the interaction. I conclude that the pull-down assay is likely to be more stringent than the cross-linking assay and hence better reflects the formation of stable complexes between the two groups of proteins (see Discussion).

A number of control experiments confirmed the authenticity of the ERp57-calreticulin association, as detected by the pull-down assay, and showed that the presence of the ER targeting signal on the ERp57 derived proteins had little effect on their binding. When the ERp57/PDI chimeras were tested using the pull-down assay only ERp57Δc and chimeras 3 and 4 were found to bind to the immobilised calreticulin at a level approaching that found with the wild type protein. All of these ERp57 derivatives had also shown clear cross-linking to the ER lectins. Thus, the pull-down assay confirmed a central role for the ERp57 ***b*** and ***b'*** domains in ER lectin binding. However, the quantitative nature of the

assay was also able to reveal that both the ***a'*** domain and C-terminal extension contribute to the efficiency of binding and thus only chimera 4 bound as efficiently as the wild type protein.

The pull-down assay was particularly useful in dissecting out useful information from the ERp57 sub-domain analysis, and although the ERp57***abb'*** region showed clear cross-linking to the ER lectins no binding of it, or any of the other ERp57 sub-domains, was detectable in the pull-down system. Thus, the ERp57***abb'*** region is incapable of the robust/stable interaction necessary to obtain a signal in the pull-down assay using immobilised calreticulin as bait.

The last set of ERp57 derivatives to be analysed were the ***b'*** point mutants. Once again the pull-down assay was clearly able to discriminate between mutants that behaved identically in the cross-linking assay (see Discussion). In most cases point mutants in the ERp57 ***b'*** domain caused a marked reduction in the efficiency of binding to calreticulin, supporting a model where this region of ERp57 makes an important contribution to the ER lectin binding site.

5.7 Conclusion

In this chapter a pull-down assay was used to analyse the molecular basis for the interaction between ERp57 and immobilised calreticulin revealing that both the ***b*** and ***b'*** domains of ERp57 are required for the specific interaction. In addition the ***a'*** domain significantly enhances binding and the C-terminal extension may also facilitate complex formation with calreticulin. None of the ERp57 sub-domain derivatives remained bound to biotin-tagged calreticulin in this assay, and only one of the ERp57 ***b'*** domain point mutants (E270A) showed binding similar to wild-type protein. Hence, certain amino acids in the ***b'*** domain of ERp57 may contribute to its specific interaction with the ER lectins.

CHAPTER - 6

Discussion

6 Discussion

6.1 Introduction

The ER lumen contains a number of distinct molecular chaperones and folding factors, that are involved in modulating the folding and assembly of newly synthesised proteins and protein complexes (Ferrari and Söling, 1999; Zapun *et al.,* 1999). These include the members of the PDI family, which appear to fill both enzymatic and chaperone functions within the ER (Ferrari and Söling, 1999). In the ER lumen, the formation of disulphide bonds is an important protein modification, and these bonds serve to stabilise the native conformation of many secretory proteins. Apart from PDI, there are several additional PDI-like proteins (Ferrari and Söling, 1999). One such protein is ERp57, also known as GRP58, ERp60, ERp61, ER60, HIP-70, Q2 and P58 (Elliott *et al.,* 1997, Freedman *et al.,* 2002)

ERp57 has been studied in detail and found to be specifically involved in glycoprotein folding (Oliver *et al.,* 1997; Elliott *et al.,* 1997; Van der Wal *et al.,* 1998; Elliott *et al.,* 1998; Zapun *et al.,* 1998; Oliver *et al.,* 1999; Molinari and Helenius, 1999). However, ERp57 itself is not a lectin *per se*, since it does not bind to oligosaccharides (Zapun *et al.,* 1998). Rather it forms discrete complexes with the ER lectins calnexin and calreticulin, and it is these complexes that interact specifically with newly synthesised glycoproteins (Oliver *et al.,* 1999). ERp57 is not glycosylated, so these complexes are not the result of 'typical' calnexin/calreticulin-monoglucosylated glycoprotein substrate interactions. A member of the same family of proteins as ERp57, archetypal PDI, shares 29% sequence identity and 56% sequence similarity with ERp57 (Koivunen *et al.,* 1996). Although PDI is structurally and functionally related to ERp57, *in vitro* cross-linking

studies revealed no association between calreticulin and PDI under conditions where a strong association of calreticulin with ERp57 was observed (Oliver *et al.,* 1999). ERp57 therefore functions as a generic subunit of a glycoprotein specific folding machinery operating in the ER (Oliver *et al.,* 1997, High *et al.,* 2000).

The aim of this project was to investigate the molecular basis for the specific interaction between ERp57 and the two ER lectins, focussing primarily on the soluble lectin, calreticulin. A number of ERp57/PDI chimeras, ERp57 sub-domain derivatives and ERp57 ***b'*** domain point mutants were synthesised *in vitro* in the presence of canine pancreatic microsomes and semi-permeabilised cells. The ability of the imported proteins to interact with the endogenous ER lectins was then analysed by cross-linking and immunoprecipitation. The ability of the various ERp57 derivatives to bind specifically to biotin-tagged calreticulin was also analysed using an alternative pull-down assay that could be used quantitatively.

6.2 The b and b' domains of ERp57 are both essential for its cross-linking to the ER lectins

Radiolabelled ERp57, ERp57Δc and ERp57/PDI chimeric polypeptides were imported into microsomes and semi-permeabilised cells and treated with hetero- and homo-bifunctional cross-linking reagents. The products were then immunoprecipitated with a variety of anti-sera specific for calreticulin, calnexin, PDI, BiP and ERp57 (both the whole protein and a C-terminal peptide). In this way, adducts of calreticulin with ERp57, ERp57Δc, construct 3 (ERp57***abb'***PDI***a'c***) and construct 4 (PDI***a***ERp57***bb'a'c***) were identified (Figures 3.3-3.6, 3.8 and 3.9, indicated by}). In addition, a high molecular weight adduct with calnexin

was identified (Figures 3.3-3.6, 3.8 and 3.9, indicated by *) and these data are summarised in Table 6.1.

Table 6.1 Comparison of ERp57 derivative pull-down and cross-linking data with published prolyl 4-hydroxylase activity data

Polypeptide	**% bound in pull down**	**ER lectin cross-links**	***% Prolyl 4-hydroxylase activity***	
			α(I)subunit	***α(II)subunit***
ERp57	38	Yes	*<4.2*	*<4.2*
ERp57Δ*c*	29	Yes	-	-
ERp57*abb'*	0	Yes	-	-
Construct 2	0	No	*7.3*	*17.4*
Construct 3	23	Yes	*<4.2*	*<4.2*
Construct 4	40	Yes	*<4.2*	*<4.2*
Construct 5	0	No	*<4.2*	*<4.2*
Construct 6	0	No	-	-
Construct 7	0	No	-	-
PDI	0	No	*100*	*100*

Table 6.1 Comparison of ERp57 derivatives in the pull-down assay (see Section 6.6 for discussion), cross-linking assay and published prolyl 4-hydroxylase data (Pirneskoski *et al.,* 2001).

The results outlined above suggested that the C-terminal extension, which is markedly different between ERp57 and PDI, was not necessary for the specific interaction of ERp57 with calreticulin and calnexin, since the removal of this domain did not abolish the cross-linking of ERp57 to these two proteins. Conversely, both the ***b*** and ***b'*** domains of ERp57 were found to be necessary, since constructs 2 (ERp57***ab***PDI***b'a'c***), 5 (PDI***abb'x***ERp57 ***a'***), 6 (PDI***abb'x***ERp57***a'*** + AVKDEL) and 7 (PDI***ab***ERp57***b'a'c***) did not form adducts with the ER lectins or any other obvious ER component (Figure 3.8,

Panel B and 3.9, Panels A-C). In contrast, constructs 3 (ERp57 ***abb'*** PDI ***a'c***) and 4 (PDI***a***ERp57***bb'a'c***) did form specific adducts with calreticulin and calnexin (Figure3.5, 3.6, and 3.8C and D, indicated by } and *).

The chimeric constructs 2-6 have also been used in a published study of prolyl 4-hydroxylase complex formation, for which PDI functions as the β subunit. This analysis indicates that these chimeras are able to attain a native-like conformation, and some can assemble into an active prolyl 4-hydroxylase complex (summarised in Table 6.1, Pirneskoski *et al.,* 2001). In addition, cross-linking of the chimeras to 'general' ER chaperones such as PDI or BiP was not detected, again suggesting that although 'non-native' these proteins are not grossly misfolded and do not act as misfolded protein substrates for a range of ER lumenal chaperones including the ER lectins (see Chapter 3).

Viewed in isolation, the cross-linking data indicate that the ***a*** and ***a'*** domains, and the C-terminal extension of ERp57 are not essential for its association with the ER lectins. However, subsequent analysis using a more stringent pull down assay indicates that this is an over simplification of the actual situation as discussed below (Section 6.6).

6.3 ERp57abb' is sufficient for adduct formation with the ER lectins

In order to determine whether the ***b*** and/or ***b'*** domains of ERp57 were sufficient for its specific interaction with the ER lectins the ERp57 sub-domain derivatives ***abb'***, ***b'a'c***, ***bb'x*** and ***b'x*** were studied. In these *in vitro* cross-linking experiments the ERp57 ***bb'x*** or ***b'x*** fragments alone were not sufficient for adduct formation with the ER lectins (Figure 4.4C and ERp57***b'x*** data not shown). In fact, the minimal element of ERp57 that was found to give detectable adducts with calreticulin and calnexin was the ERp57***abb'*** fragment (Figure 4.4B, lane 6 } and lane 7 * respectively).

The analysis of ERp57 sub-domains in isolation was hampered by their tendency to generate complex cross-linking data and/or yield a very high background in the pull down assay (e.g. Figures 4.4A-C and Figure 5.6 showing summary of pull-down data). Taken together, these results suggest that, at least in the form used in this study, these sub-domains of ERp57 are unable to attain a native-like structure and may therefore be intrinsically unsuitable for such studies. When present in the form of chimeras, these regions are presumably able to 'pack' with the complementary regions from PDI and reach a more stable/native-like structure.

6.4 Point mutations in the ERp57 b' domain can severely disrupt adduct formation with calreticulin

A detailed study of PDI had shown that its ***b'*** domain provided the principal peptide binding site, although other domains contribute to its association with larger misfolded proteins (Klappa *et al.,* 1998). For this reason, a number of ERp57 mutants with single amino acid changes in the equivalent ***b'*** domain were analysed for their ability to bind to calreticulin by cross-linking and the relative efficiency of adduct formation was quantified by phosphorimaging. Four mutants, F280A, V283L, F299A and F299W, all showed a significant reduction in their association with calreticulin as judged by cross-linking (Table 6.2). Further analysis of F280A, F299A and F299W (Figure 4.9) showed no cross-linking to the generic ER chaperones PDI and BiP indicating that the ERp57 polypeptides were not recognised as being chronically misfolded. Whilst further analysis of these ERp57 point mutants will be required to confirm they can attain a near native structure, taken together these data support a model where the ***b'*** domain forms a major element of the ER lectin binding site.

6.5 The ERp57/biotin-tagged calreticulin interaction is specific and not dependent upon signal sequence cleavage

A major caveat with the cross-linking data outlined above is that this assay only determines whether two proteins are in close molecular proximity, and it cannot be assumed that they are interacting directly. To show that ERp57 binds to calreticulin, and to establish whether the formation of ERp57 cross-linking products reflected an authentic non-covalent interaction, a binding assay that was not dependent upon cross-linking was established.

Previous studies had indicated that ERp57 was capable of binding to calreticulin in solution (Corbett *et al.,* 1999, Oliver *et al.,* 1999) and that this interaction was important for ERp57 function (Zapun *et al.,* 1998). More recently, the dissociation constant for the complex was estimated as $(9.1\pm3.0)\text{x}10^{-6}$M using the P-domain of calreticulin and wild-type ERp57 (Frickel *et al.,* 2002), suggesting that binding would be strong enough to survive a pull-down assay. This was indeed found to be the case, and a substantial proportion of wild type ERp57 could be specifically recovered with immobilised calreticulin in the pull-down assay I developed.

Previous studies had indicated that both calnexin and calreticulin can exhibit associations with non-glycosylated proteins both *in vitro* and *in vivo* (Ihara *et al.,* 1999; Danilczyk and Williams, 2001), and this may reflect a role as general chaperones or folding factors, that is additional to their lectin function. It was therefore important to establish that the binding of ERp57 to calreticulin I detected in the pull-down assay reflected the formation of a biologically relevant complex, and not the binding of misfolded ERp57 to calreticulin acting as a broad specificity chaperone. For this reason I investigated the effects of 'native' and 'scrambled' RNase A upon the capacity of immobilised calreticulin to bind to ERp57. These substrates are well established and, in the case of PDI, the scrambled form of RNase A efficiently competes with peptides for

binding to the site of substrate interaction whilst 'native' RNase A does not (Klappa *et al.,* 19998). These control experiments confirmed that the interaction detected in the pull-down assay reflected authentic complex formation between ERp57 and calreticulin. No specific interaction of PDI with immobilised calreticulin was detected, confirming the specificity of the interaction.

A second potential drawback of the binding assay was that because the polypeptides were added directly to immobilised calreticulin, in contrast to precursors imported into microsomes or semi-permeabilised cells, the N-terminal signal sequences were not removed from the proteins. It has been shown that the mature region of a protein can often attain a conformation indistinguishable from that observed with the signal sequence present. In the case of maltose binding protein, the presence of the N-terminal signal sequence does not prevent the protein from binding to its natural ligand, indicating that even with the signal sequence it is able to fold into a native structure (Park *et al.,* 1998). Furthermore, when calreticulin was synthesised *in vitro* in the absence of signal sequence cleavage, the protein was found to form an intrachain disulphide bond characteristic of the functional protein (Oliver *et al.,* 1999). When the effect of the ERp57 signal sequence was analysed directly using the pull down assay the presence of the signal sequence had no substantial effect upon binding to immobilised calreticulin (Figures 5.1 and 5.2). Thus, the pull down assay could be used to analyse the various ERp57 derivatives that I had previously investigated by a cross-linking assay.

6.6 Pull down analysis reveals that the a' domain and C-terminal extension of ERp57 enhance its interaction with calreticulin.

The nature of a pull-down assay requires that the binding of the soluble partner to the immobilised bait occurs, and the assay includes several washing steps to remove any non-specific and/or low affinity interactions. It is therefore reasonable to expect that such a pull-down assay can provide more rigorous and discriminating measure of true binding than a cross-linking assay.

The chimeric ERp57/PDI constructs tested in the pull-down assay mirrored the results previously obtained by cross-linking (cf. Figures 3.3-3.9 and 5.6), thus, only ERp57Δc (ERp57***abb'a'***) and chimeric constructs 3 (ERp57***abb'***PDI***a'c***) and 4 (PDI***a***ERp57***bb'a'c***) remained bound at significant levels to biotin-tagged calreticulin after washing. However, the pull-down assay also revealed differences in the ability of these three constructs to interact with biotin-tagged calreticulin, as only construct 4 displayed binding equivalent to wild-type ERp57. In contrast, construct 3 and ERp57Δc displayed only about half and three quarters of the wild type level of binding (Figure 5.6 and Table 6.1).

Furthermore, the pull down analysis showed that the ERp57***abb'*** region was not strongly bound to immobilised calreticulin. Thus, ERp57Δc was the minimal element that showed substantial binding in the pull down assay. There was insufficient time to generate additional sub-regions of ERp57 for specific analysis using the pull down assay, however, in future studies it would be interesting to determine whether the ***bb'a'*** region of ERp57 could bind calreticulin, and if so with what efficiency. Since the binding of chimeric construct 4 is at wild type levels, the ***a*** domain may either be completely indispensable, or may enhance binding without being essential for the association. Furthermore, it was clear

that even the binding of the ERp57***abb'a'*** region was enhanced by the presence of the C-terminal extension, cf. ERp57Δc and wild type ERp57 or chimeric construct 3 (Table 6.1).

The data presented in this thesis indicating that the ***bb'*** and ***a'*** domains of ERp57 are involved in calreticulin binding (see Figure 6.1) are consistent with a previous study of PDI chimeras showing that the ***b'*** and ***a'*** domains of PDI are crucial for its efficient assembly into a functional prolyl 4-hydroxylase complex (Pirneskoski *et al.,* 2001). The pull-down assay also indicates a role for the C-terminal extension of ERp57 in its binding to calreticulin. This is perhaps not completely unexpected, given the unique nature of this region when compared to other members of the PDI family. However, in the case of PDI, the removal of the equivalent C-terminal extension had no effect on its known functions (Koivunen *et al.,* 1999) suggesting that C-terminal extension of ERp57 may play a specialised role in its assembly with calreticulin (see Figure 6.1).

In future experiments it would be useful to place the ERp57***bb'*** and ERp57***bb'a'*** regions together with the PDI***a--a'c*** and ***a---c*** regions respectively. This may facilitate the assembly of the minimal ERp57 sub-domains and establish whether the reduction in binding efficiency of the ERp57Δc construct is directly due to a contribution to binding by the C-terminal extension, or reflects a degree of misfolding or misassembly of the rest of the polypeptide in its absence.

6.7 Most ERp57 b' domain point mutants are severely compromised in the pull down assay

The most striking result of the analysis of the ERp57 ***b'*** domain point mutants using the pull-down assay was that all except one of the mutants showed a large reduction in calreticulin binding. In fact, only two of the thirteen mutants displayed binding levels that were significantly above background, namely E270A and V283A. For the rest of the mutants no substantial calreticulin binding was detected independent of whether cross-linking to the ER lectins had been seen (Table 6.2).

The variation between cross-linking and pull-down results is probably due to assay conditions. In cross-linking experiments any interactions are stabilised by adduct formation, however, in the case of the pull-down assay the interaction between biotin-tagged calreticulin and ERp57 derivatives is not stabilised by additional reagents. Samples then undergo repeated washes, which should select for 'high-affinity' interactions. The folding factors, that are normally present in the ER, and hence in microsomes and semi-intact cells, are also absent in the pull-down assay. Therefore, the ERp57/PDI chimeras and other ERp57 derivatives, which are all engineered proteins, may have more difficultly in attaining their correct conformation in the conditions of the pull-down assay as compared to the cross-linking assay.

A pivotal role of the ***b'*** domain in generating a binding site for calreticulin is suggested by the disruptive nature of most of the point mutants that were targeted to this area (Table 6.2, see also Figure 6.1). These data are consistent with studies of the substrate binding domains of PDI that indicate peptides interact primarily with its ***b'*** domain whilst larger proteins also require the ***b*** and ***a'*** domains for efficient binding (Klappa *et al.,* 1998; Klappa *et al.,* 2000). Thus, we can view ERp57 as a specialised version of PDI that binds

specifically to two related ER lectins, calnexin and calreticulin, in place of the broad range of substrates that can associate with PDI (Figure 6.1, cf. Freedman *et al.,* 2002).

There are several caveats to this conclusion and further studies will be necessary to confirm some of the details. Thus, whilst the pull-down assay allowed quantitative data to be obtained, other assays for analysing binding at higher resolution would be extremely valuable. The use of microcalorimetry or surface plasmon resonance ('BiaCore') would allow true dissociation constants for the chimeras and point mutants to be obtained and compared with the wild-type protein (cf. Frickel *et al.,* 2002). Such studies would require up to milligram quantities of purified recombinant proteins for each of the ERp57 derivatives in addition to purified calreticulin, as used in the pull down assay. Even more importantly, it will be crucial to confirm that the ERp57 ***b'*** domain point mutants are correctly folded into a native or near native structure since, in contrast to the ERp57/PDI chimeras which have been used in other studies (Pirneskoski *et al.,* 2001), the ERp57 ***b'*** point mutants are not fully characterised. Preliminary studies using limited proteolysis and CD spectroscopy suggest that the structure of the F299W ERp57 ***b'*** mutant is similar to that of the wild-type protein (Dr Lloyd Ruddock, University of Oulu, personal communication). If so then its complete inability to associate with the ER lectins would support a key role for the ***b'*** domain in this interaction. The status of other ***b'*** domain mutants, in comparison to the wild type protein, remains to be established. Whilst the cross-linking analysis detailed in this thesis analysed interactions of ERp57 with both calnexin and calreticulin, the pull-down assay focussed on calreticulin for technical reasons, i.e. because it is a soluble protein. Whilst the working model for the ERp57/ER lectin complex presented in Figure 6.1 describes calreticulin, it seems likely that the binding of calnexin occurs in an identical manner.

Table 6.2 Comparison of ERp57 *b'* domain point mutants

ERp57 derivative	**% cross-linked to calreticulin**	**% bound in pull-down assay (less background binding)**
ERp57	6	35
Point mutant 1 V267A	5	6
Point mutant 2 V267W	5	3.5
Point mutant 3 Y269A	5.8	4
Point mutant 4 Y269W	4.6	11
Point mutant 5 E270A	6.1	30
Point mutant 6 K274A	4.3	4
Point mutant 7 R280A	2.4	4
Point mutant 8 N281A	4.8	<1
Point mutant 9 V283A	3.6	15
Point mutant 10 V283L	2.2	6
Point mutant 11 V283W	3.8	<1
Point mutant 12 F299A	1.8	4
Point mutant 13 F299W	0.5	<1

Table 6.2 Comparison of ERp57 ***b'*** domain point mutant pull-down and cross-linking data (Figures 4.8 and 5.11).

6.8 Working model for ERp57 binding to calreticulin

On the basis of the data presented in this thesis, I propose that ERp57 binds to calreticulin across several of its domains, with only the ***a*** domain making no contribution at all (Figure 6.1). In essence, my data suggest that the binding of calreticulin to ERp57 mirrors the binding of a misfolded protein substrate to PDI (Figure 6.1). If this is viewed in terms of 'molecular evolution', then the specialisation of the ***b*** and ***b'*** domains of ERp57 for binding to calreticulin is reflected by their considerable divergence from the equivalent domains of PDI (23% and 14% sequence identity respectively). This divergence is much greater then that seen in the flanking ***a*** and ***a'*** domains (49% and 54% sequence identity respectively).

In the case of ERp57 my data suggest an additional role for the C-terminal extension (Figure 6.1). In terms of ERp57 function, it appears that the ER lectin binds specifically to the monoglucosylated glycoprotein substrate, and specifically recruits them to the ERp57/ER lectin complex (Figure 6.1). This model is entirely consistent with all of the functional data regarding the action of ERp57 on glycoproteins (see High *et al.,* 2000). The principal site of ER lectin binding to ERp57 is the so-called 'P-loop' which is distinct from the region that acts as a lectin (Ellgaard *et al.,* 2001; Frickel *et al.,* 2002; Schrag *et al.,* 2001). Furthermore, this model would leave the ***a*** domain entirely free of any constraints imposed by binding to calreticulin, and therefore readily available as a flexible 'redox active' domain capable of catalysing disulphide bond exchange within glycoprotein substrates bound to the ER lectin part of the complex.

In the long term, the nature of the ERp57/ER lectin-binding site will require the high resolution structure of one of the complexes to be solved. To date, no crystal structure for a member of the PDI family has been solved and our information comes from

NMR studies of functional sub-domains (see Introduction, Freedman *et al.,* 2002). Since the P-domains of calreticulin and calnexin appear to be the sole region of ERp57 binding, this ~10kDa fragment would be suitable for NMR studies (Frickel *et al.,* 2002). However, I find that only the ***a*** domain of ERp57 is dispensible for calreticulin binding and the remaining ***bb'a'c*** region of ERp57 is ~45kDa. At present, the putative minimal complex of the calreticulin P-loop and the ERp57***bb'a'c*** region is too large to obtain a high-resolution structure by normal NMR techniques.

Figure 6.1 Substrate binding of ERp57 versus PDI

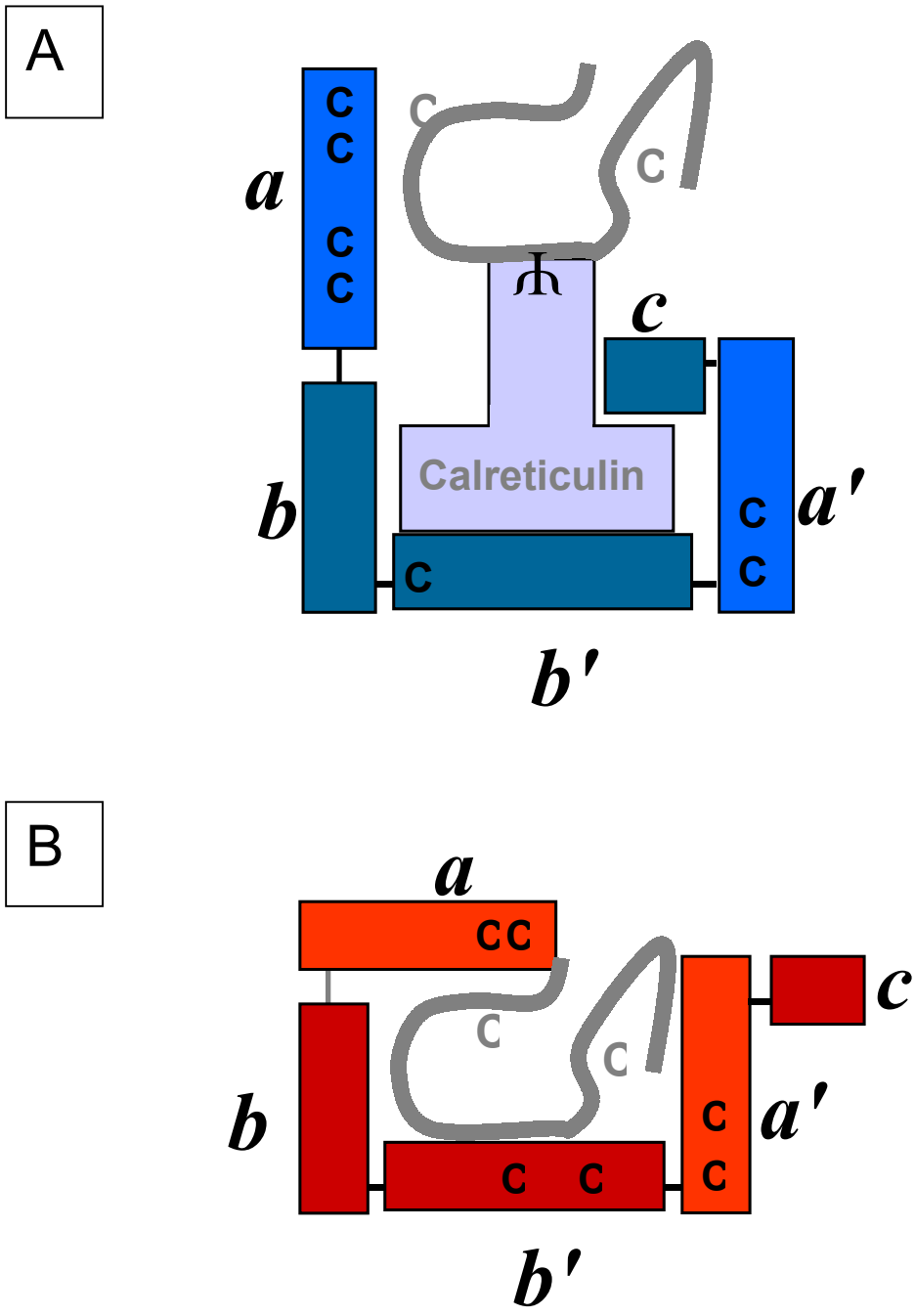


Figure 6.1 Substrate binding to ERp57 (A) versus PDI (B). Models indicating the regions of ERp57 and PDI that interact with calreticulin and large substrates respectively are shown. High resolution structures of domains of calnexin (Schrag *et al.,* 2001) and calreticulin (Ellgaard *et al.,* 2001) indicate that ERp57 binds to the conserved P-loop of these two lectins and recent NMR studies confirm this model (Frickel *et al.,* 2002). ERp57 shows an indirect interaction with its substrate via calreticulin with the ***b*** and ***b'*** domains of ERp57 essential for ER lectin binding, whilst the ***a'*** and C-terminal extension can enhance binding. In comparison PDI shows a direct interaction with its substrate, principally through its ***b'*** domain, but the ***a'*** domain/C-terminal extension are required for the binding of larger substrates.

6.9 Conclusions

By analysing the interactions of ERp57 derivatives with calreticulin using both cross-linking and pull-down assays I have been able to provide detailed insights into the molecular basis for the specific assembly of these components within the ER lumen. Bearing in mind the reservations outlined above, the ERp57/PDI chimeric constructs and ERp57 sub-domain derivatives have revealed that the ***b*** and ***b'*** domains of ERp57 are necessary, but not sufficient for its binding to the ER lectins. Whilst ERp57 ***b'*** domain point mutants suggest that certain amino acids in this region may be crucial for complex formation, the binding region extends to the ***b*** and ***a'*** domains with the C-terminal extension either contributing directly to the association of the two proteins or acting to stabilise ERp57 in a conformation that favours complex formation. Further experiments, including structural studies, will be needed to confirm these results and to determine whether the molecular requirements for the binding of calnexin to ERp57 are identical to those I have defined for calreticulin.

Bibliography

Althoff, S., Selinger, D., and Wise, J. A. (1994). Molecular evolution of SRP cycle components: functional implications. *Nucleic Acids Res.* **22**:1933-1947.

Anfinsen, C.B. (1973). Principles that Govern the Folding of Protein Chains. *Science* **181**:223-230.

Antoniou, A.N., Ford, S., Alphey, M., Osborne, A., Elliott, T., and Powis, S.J. (2002). The Oxidoreductase ERp57 Efficiently Reduces Partially Folded in Preference to Fully Folded MHC Class I Molecules. *EMBO J*. **21**:2655-2663.

Arnold, D., Wahl, C., Faath, S., Rammensee, H. G., and Schold, H. (1997). Influences of Transporter Associated with Antigen Processing (TAP) on the Repertoire of Peptides Associated with the Endoplasmic Reticulum-Resident Stress Protein gp96. *J. Exp. Med.* **186**:461-466.

Ayalon-Soffer, M., Shenkman, M., and Lederkremer, G.Z. (1999). Differential Role of Mannose and Glucose Trimming in the ER Degradation of Asialoglycoprotein Receptor Subunits. *J. Cell Sci*. **112**:3309-3318.

Bacher, G., Lutcke, H., Jungnickel, B., Rapoport, T.A., and Dobberstein, B. (1996) Regulation by the Ribosome of the GTPase of the Signal-Recognition Particle during Protein Targeting. *Nature* **381**:248-251.

Bacher, G., Pool, M., and Dobberstein, B. (1999). The Ribosome Regulates the GTPase of the β-Subunit of the Signal Recognition Particle Receptor. *J. Cell Biol.* **146**:723-730.

Bader, M., Muse, W., Ballo, D.P., Gassner, C., and Bardwell, J.C. (1999). Oxidative Protein Folding is Driven by the Electron Transport System. *Cell* **98**:217-227.

Ballinger, C.A., Connell, P., Wu, Y., Hu, Z., Thompson, L.J., Yin, L., and Patterson, C. (1999). Identification of CHIP, a Novel Tetratricopeptide Repeat-Containing Protein That Interacts with Heat Shock Proteins and Negatively Regulates Chaperone Functions. *Mol. Cell Biol*. **19**:4535-4545.

Bebök, Z., Mazzochi, C., King, S.A., Hong, J.S., and Sorscher, E.J. (1998). The Mechanism Underlying Cystic Fibrosis Transmembrane Conductance Regulator Transport from the Endoplasmic Reticulum to the Proteasome Includes Sec61β and a Cytosolic, Deglycosylated Intermediary. *J. Biol. Chem*. **273**:29873-29878.

Beckmann, R., Bubeck, D., Grassucci, R., Penczek, P., and Verschoor, A. (1997). Alignment of Conduits for the Nascent Polypeptide Chain in the Ribosome-Sec61 Complex. *Science* **278**:2123-2126.

Beckmann, R., Spahn, C.M.T., Eswar, N., Helmer, J., Penczek, P.A., Sali, A., Frank, J., and Blobel, G. (2001) Architecture of the Protein-Conducting Channel Associated with the Translating 80S Ribosome. *Cell* **107**:361-372

Benham, A.M., Cabibbo, A., Fassio, A., Bulleid, N.J., and Braakman, I. (2000). The CXXCXXC Motif Determines the Folding Structure and Stability of Human Ero1-Lα. *EMBO J*. **19**:4493-4502.

Bennet, C.F., Balcarek, J.M., Varrichio, A., and Crooke, S.T. (1988). Molecular Cloning of and Complete Amino-Acid Sequence of Form-I Phosphoinositide-Specific Phopholipase C. *Nature* **334**:268-270.

Borel, A.C.S.S.M. (1996). Biogenesis of Polytopic Membrane Proteins: Membrane Segments Assembled Within Translocon Channels Prior to Membrane Integration. *Cell* **85**:379-389.

Bourdi, M., Demady, D., Martin, J.L., Jabbour, S.K., Martin, B.M., George, J.W., and Pohl, L.R. (1995). cDNA Cloning and Baculovirus Expression of the Human Liver Endoplasmic Reticulum P58: Characterization as a Protein Disulphide Isomerase Isoform, But Not as a Protease or a Carnitine Acyltransferase. *Arch. Biochem. Biophys.* **323**:397-403.

Braakman, I. (2001). A Novel Lectin in the Secretory Pathway. An Elegant Mechanism for Glycoprotein Elimination. *EMBO Reports* **2**:666-668.

Brodsky, J.L. (1996) Post-Translational Protein Translocation: Not All Hsc70s are Created Equal. *Trends Biochem.Sci.* **21**:122-126.

Brodsky, J.L. and Schekman, R. (1993). A Sec63p-BiP Complex From Yeast is Required for Protein Translocation in a Reconstituted Proteoliposome. *J. Cell Biol.* **123**:1355-1363.

Brodsky, J.L., Werner, E.D., Dubas, M.E., Goeckeler, J.L., Kruse, K.B., and McCracken, A.A. (1999). The Requirement for Molecular Chaperones During Endoplasmic Reticulum-Associated Protein Degradation Demonstrates That Protein Export and Import are Mechanistically Distinct. *J. Biol.Chem*. **274**:3453-4360.

Buckie, J.W. and Cook, G.M.W. (1986). Specific Isolation of Surface Glycoproteins From Intact Cells by Biotinylated Concanavalin A and Immobilized Strepavidin. *Anal. Biochem.* **156**:463-472.

Cabibbo, A., Pagani, M., Fabbri, M., Rocchi, M., Farmery, M.R., Bulleid, N.J., and Sitia, R. (2000). ERO1-L, a Human Protein That Favours Disulphide Bond Formation in the Endoplasmic Reticulum. *J. Biol. Chem*. **275**:4827-4833.

Cabral, C.M., Liu, L., and Sifers, R.N. (2001). Dissecting Glycoprotein Quality Control in the Secretory Pathway. *Trends Biochem. Sci.* **26**:619-624.

Cannon, K.S. and Cresswell, P. (2001) Quality Control of Transmembrane Domain Assembly in the Tetraspanin CD82. *EMBO J.* **10**:2453.

Chapman, R., Sidrauski, C., and Walter, P. (1998). Intracellular Signalling from the Endoplasmic Reticulum to the Nucleus. *Annu.Rev.Cell Dev.Biol.* **14**:459-485.

Chillaron, J., Adan, C., and Haas, I.G. (2000). Mannosidase Action, Independent of Glucose Trimming, Is Essential for Proteasome-Mediated Degradation of Unassembled Glycosylated Ig Light Chains. *Biol. Chem*. **318**:1155-1164.

Chung, D.H., Ohashi, K., Watanabes, M., Miyasaka, N., and Hirosawa, S. (2000). Mannose Trimming Targets Mutant α_2-Plasmin Inhibitor for Degradation by the Proteosome. *J. Biol. Chem.* **275**:4981-4987.

Connell, P., Ballinger, C.A., Jiang, J., Wu, Y., and Thompson, L.J. (2001). The Co-Chaperone CHIP Regulates Protein Triage Decisions Mediated By Heat-Shock Proteins. *Nat. Cell Biol.* **3**:93-96.

Connolly, T., and Gilmore, R. (1989). The Signal Recognition Particle Receptor Mediates the GTP-dependent Displacement of SRP from the Signal Sequence of the Nascent Polypeptide. *Cell* **57**:599-610.

Corbett, E.F., Oikawa, K., Francois, P., Tessier, D.C., Kay, C., Bergeron, J.J.M., Thomas D.Y, Krauser, K., and Michalak, M. (1999). Ca^{2+} Regulation of Interactions Between Endoplasmic Reticulum Chaperones. *J. Biol. Chem.* **274**:6203-6211.

Crowley, K.S., Reinhart, G.D., and Johnson, A.E. (1993). The Signal Sequence Moves Through a Ribosomal Tunnel Into a Non-Cytoplasmic Aqueous Environment at the ER Membrane Early in Translocation. *Cell* **73**:1101-1116.

Cuozzo, J.W. and Kaiser, C.A. (1999). Competition Between Glutathione and Protein Thiols for Disulphide Bond Formation. *Nature Cell Biol*. **1**:130-135.

Danilczyk, U.G., Cohen-Doyle, M., and Williams, D.B. (2000). Functional Relationships Between Calreticulin, Calnexin, and the Endoplasmic Reticulum Luminal Domain of Calnexin. *J. Biol. Chem.* **275**:13089-13097.

Danilczyk, U.G. and Williams, D.B. (2001). The Lectin Chaperone Calnexin Utilizes Polypeptide-Based Interactions to Associate with Many of Its Substrates *In vivo*. *J. Biol. Chem.* **276**:25532-25540.

Darby, N.J. and Creighton, T.E. (1995a). Catalytic Mechanism of DsbA and its Comparison with that of Protein Disulphide-Isomerase. *Biochem.* **34**:3576-3587.

Darby, N.J. and Creighton, T.E. (1995b). Characterization of the Active-Site Cysteine Residues of the Thioredoxin-Like Domains of Protein Disulphide-Isomerase. *Biochem.* **34**:16770-16780.

Darby, N.J. and Creighton, T.E. (1995c). Functional-Properties of the Individual Thioredoxin-Like Domains of Protein Disulphide-Isomerase. *Biochem.* **34**:11725-11735.

Darby, N.J., Penka, E., and Vincentelli, R. (1998). The Multi-Domain Structure of Protein Disulphide Isomerase is Essential for High Catalytic Efficiency. *J. Mol. Biol.* **276**:239-247.

David, V., Hochstenbach, F., Rajagopalan, S., and Brenner, M.B. (1993). Interaction with Newly Synthesized and Retained Proteins in the Endoplasmic Reticulum Suggests a Chaperone Function for Human Integral Membrane Protein IP90 (Calnexin). *J. Biol. Chem.* **268**:9585-9592.

de Virgilio, M., Kitzmüller, C., Schwaiger, E., Klein, M., Kreibich, G., and Ivessa, N.E. (1999). Degradation of a Short-lived Glycoprotein from the Lumen of the Endoplasmic Reticulum: The Role of N-linked Glycans and the Unfolded Protein Response. *Mol. Biol. Cell* **10**:4059-4073.

de Virgilio, M., Weninger, H., and Ivessa, N.E. (1998). Ubiquitination is Required for the Retro-translocation of Short-lived Luminal Endoplasmic Reticulum Glycoprotein to the Cytosol for Degradation by the Proteasome. *J. Biol. Chem.* **273**:9734-9743.

Desilva, M.G. and Lan, M.S. (1996). Characterisation and Chromosomal Localisation of a New Protein Disulphide Isomerase, PDIp, Highly Expressed in Human Pancreas. *DNA Cell Biol.* **15**:9-16.

Dick, T.P., Bangia, N., Peaper, D.R., and Cresswell, P. (2002). Disulphide Bond Isomerization and the Assembly of MHC Class I-Peptide Complexes. *Immunity* **16**:87-98.

Dijkstra, K., Karvonen, P., Pirneskoski, A., Koivunen, P., Kivirikko, K.I., Darby, N.J., van Straaten, M., Scheek, R.M., and Kemmink, J. (1999). Assignment of ^{1}H, ^{13}C and ^{15}N Resonances of the *a′* Domain of Protein Disulphide Isomerase. *J. Biomol. NMR* **14**:195-196.

Do, H., Falcone, D., Lin, J., Andrews, DW., and Johnson, AE. (1996). The Cotranslational Integration of Membrane Proteins into the Phospholipid Bilayer is a Multistep Process. *Cell* **85**:369-378.

Edman, J.C., Ellis, L., Blacher. R.W., Roth, R.A., and Rutter, W.J., (1985). Sequence of Protein Disulphide Isomerase and Implications of Its Relationship to Thioredoxin. *Nature* **317**:267-270.

Ellgaard, L. and Helenius, A. (2001). ER Quality Control: Towards an Understanding at the Molecular Level. *Curr. Opin. Cell Biol* **13**:431-437.

Ellgaard, L., Molinari, M., and Helenius, A. (1999). Setting the Standards: Quality Control in the Secretory Pathway. *Science* **286**:1882-1888.

Ellgaard, L., Riek, R., Herrmann, T., üntert, P., Braun, D., Helenius, A., and Wüthrich, K. (2001). NMR Structure of the Calreticulin P-Domain. *Proc. Natl. Acad. Sci.* **98**:3133-3138.

Elliott, J.G., Oliver, J., and High, S. (1997). The Thiol-Dependent Reductase ERp57 Interacts Specifically with N-Glycosylated Integral Membrane Proteins. *J. Biol. Chem.* **272**:13849-13855.

Elliott, J.G., Oliver, J., Volkmer, J., Zimmerman, R., and High, S. (1998). *In vitro* Characterisation of the Interaction Between Newly Synthesised Proteins and a Pancreatic Isoform of Protein Disulphide Isomerase. *Eur. J. Biochem.* **252**:372-377.

Farmery, M.R., Allen, S., Allen, A.J., and Bulleid, N.J. (2000). The Role of ERp57 in Disulphide Bond Formation during the Assembly of Major Histocompatability Complex Class I in a Synchronized Semipermeabilized Cell Translation System. *J. Biol. Chem.* **275**:14933-14938.

Fernandez, F., D'Alessio, C., Fanchiotti, S., and Parodi, A.J. (1998). A Misfolded Protein Conformation is not a Sufficient Condition for *In vivo* Glycosylation by the UDP-Glc: Glycoprotein Glucosyltransferase. *EMBO J.* **17**:5877-5886.

Ferrari, D.M. and Söling, H. (1999). The Protein Disulphide-Isomerase Family: Unravelling a String of Folds. *Biochem. J.* **339**:1-10.

Fewell, S.W., Travers, K.J., Weissman, J.S., and Brodsky, J.L. (2001). The Action of Molecular Chaperones in the Early Secretory Pathway. *Annu. Rev. Genet.* **35**:149-191.

Flynn, G.C., Chappell, T.G., and Rothman, J.E. (1989). Peptide Binding and Release by Proteins Implicated as Catalysts of Protein Assembly. *Science* **245**:385-390.

Foti, D.M., Welihinda, A., Kaufman, R.J., and Lee, A.S. (1999). Conservation and Divergence of the Yeast and Mammalian Unfolded Protein Response. *J. Biol. Chem.* **274**:30402-30409.

Frand, A.R., and Cuozzo, J.W., and Kaiser, C.A. (2000). Pathways for Protein Disulphide Bond Formation. *Trends Cell Biol.* **10**:203-210.

Frand, A.R. and Kaiser, C.A. (1998). The ERO1 Gene of Yeast Is Required for the Oxidation of Protein Dithiols in the Endoplasmic Reticulum. *Mol. Cell* **1**:161-170.

Frand, A.R. and Kaiser, C.A. (1999). Ero1p Oxidizes Protein Disulphide Isomerase in a Pathway for Disulphide Bond Formation in the Endoplasmic Reticulum. *Mol. Cell* **4**:477

Freedman, R. B. (2002). Protein Disulphide Isomerases Exploit Synergy Between Catalytic and Specific Binding Domains. *EMBO reports* **3**:136-140.

Freedman, R.B., Dunn, A.D., and Ruddock, L.W. (1998a). Protein Folding: A Missing Redox Link in the Endoplasmic Reticulum. *Curr. Biol.* **8**:R468-470.

Freedman, R.B., Gane, P.J., Hawkins, H.C., Holdan, R., McLaughlin, S.H., and Parry, J.W.L. (1998b). Experimental and Theoretical Analyses of the Domain Architecture of Mammalian Protein Disulphide Isomerase. *Biol. Chem.* **379**:321-328.

Freedman, R.B., Hirst, T.M., and Tuite, M.F. (1994). Protein Disulphide Isomerase: Building Bridges in Protein Folding. *Trends Biotech*. **19**:331-336.

Freymann, D. M., Keenan, R. J., Stroud, R. M., and Walter, P. (1997). Structure of the Conserved GTPase Domain of the Signal Recognition Particle. *Nature* **385:**361-364.

Frickel, E., Riek, R., Jelesarov, I., Helenius, A., Wüthrich, K., and Ellgaard, L. (2002). TROSY-NMR Reveals Interaction Between ERp57 and the Tip of the Calreticulin P-domain. *Proc. Natl. Acad. Sci.* **99**:1954-1959.

Gardner, T.G. and Kearse, K.P. (1999). Modification of the T Cell Antigen Receptor (TCR) Complex by UDP-glucose:Glycoprotein Glucosyltransferase. *J. Biol. Chem.* **274**:14094-14099.

Gething, M. (1999). Role and Regulation of the ER Chaperone BiP. *Sem. Cell & Dev. Biol.* **10**:465-472.

Gething, M. and Sambrook, J. (1992). Protein Folding in the Cell. Nature **355**:33-45.

Gierasch, L. M. (1989). Signal sequences. *Biochemistry* **28**:923-930.

Gillece, P., Luz, J.M., Lennarz, W.J., and Javier de la Cruz, F. (1999). Export of a Cysteine-free Misfolded Secretory Protein from the Endoplasmic Reticulum for Degradation Requires Interaction with Protein Disulphide Isomerase. *J. Cell Biol.* **147**:1443-1456.

Gilmore, R., Blobel, G., and. Walter, P. (1982a) Protein Translocation Across the Endoplasmic Reticulum. I. Detection in the Microsomal Membrane of a Receptor for the Signal Recognition Particle. *J. Cell Biol.* **95**:463-469.

Gilmore, R., Walter, P., and Blobel, G. (1982b) Protein Translocation Across the Endoplasmic Reticulum. II. Isolation and Characterisation of the Signal Recognition Particle Receptor. *J. Cell Biol.* **95**:470-477.

Gilmore, R. and Blobel, G. (1985). Tranlsocation of Secretory Proteins Across the Microsomal Membrane Occurs Through an Environment Accessible to Aqueous Perturbants. Cell *42*, 495-497

Goldberger, R.F., Epstein, C.J., and Anfinsen, C.B. (1963). Acceleration of Reactivation of Reduced Bovine Pancreatic Ribonuclease by a Microsomal System from a Rat Liver. *J. Biol. Chem.* **238**:628-635.

Görlich, D. and Rapoport, TA. (1993). Protein Translocation into Proteoliposomes Reconstituted from Purified Components of the Endoplasmic Reticulum Membrane. *Cell* **75**:615-630.

Gretch, D.R., Suter, M., and Stinski, M.F. (1987). The Use of Biotinylated Monoclonal Antibodies and Streptavidin-Affinity Chromotography to Isolate Herpesvirus Hydrophobic Proteins or Glycoproteins. *Anal. Biochem*. **163**:270-277.

Hamman, B.D., Chen, J.C., Johnson, E.E., and Johnson, A.E. (1997). The Aqueous Pore Through the Translocon Has a Diameter of 40-60 Å During Cotranslational Protein Translocation at the ER Membrane. *Cell* **89**:535-544.

Hamman, B.D., Hendershot, L.M., and Johnson, A.E. (1998). BiP Maintains the Permeability Barrier of the ER Membrane by Sealing the Lumenal End of the Translocation Pore Before and Early in Translocation. *Cell* **92**:747-758.

Hammond, C., Braakman, I., and Helenius, A. (1994a). Role of N-Linked Oligosaccharide Recognition, Glucose Trimming, and Calnexin in Glycoprotein Folding and Quality Control. *Proc. Natl. Acad. Sci.* **91**:913-917.

Hammond, C. and Helenius, A. (1994b). Quality Control in the Secretory Pathway: Retention of a Misfolded Viral Membrane Glycoprotein Involves Cycling Between the ER, Intermediate Compartment, and Golgi Apparatus. *J. Cell Biol.* **126**:41-52.

Hammond, C. and Helenius, A. (1995). Quality control in the Secretory Pathway. Current *Opin. Cell Biol*. **7**:523-529.

Hanein, D., Matlack, K.E., Jungnickel, B., Plath, K., and Kalies, K.U. (1996). Oligomeric Rings of the Sec61p Complex Induced by Ligands Required for Protein Translocation. *Cell* **87**:721-732.

Hebert, D.N., Foellmer, B., and Helenius, A. (1995). Glucose Trimming and Reglycosylation Determine Glycoprotein Association with Calnexin in the Endoplasmic Reticulum. *Cell* **81**:425-433.

Hegde, RS. and Voigt, S. (1998). TRAM Regulates the Exposure of Nascent Secretory Proteins to the Cytosol during Translocation into the Endoplasmic Reticulum. *Cell* **92**:621-631.

Heinrich, S.U., Mothes, W., Brunner, J., and Rapoport, T.A. (2000). The Sec61p Complex Mediates the Integration of a Membrane Protein by Allowing Lipid Partitioning of the Transmembrane Domain. *Cell* **102**:233-244.

Helenius, A. and Aebi, M. (2001). Intracellular Functions of N-linked Glycans. *Science* **291**:2364-2369.

High, S. and Dobberstein, B. (1991) The Signal Sequence Interacts with the Methioine-rich Domain of the 54-kD Protein of Signal Recognition Particle. *J. Cell. Biol.* **113**:229-233.

High, S., Gorlich, D., Wiedmann, M., Rapoport, T. A., and Dobberstein, B. (1991). The identification of proteins in the proximity of signal-anchor sequences during their targeting to and insertion into the membrane of the ER. *J. Cell Biol.* **113,** 35-44.

High, S., Lecomte, F.L., Russell, S.J., Abell, B.M., and Oliver, J.D. (2000). Glycoprotein Folding in the Endoplasmic Reticulum: a Tale of Three Chaperones. *FEBS Letters* **476**:38-41.

High, S. and Laird, V. (1997). Membrane protein biosynthesis - all sewn up? *Trends Cell Biol.* **7**:206-210.

High, S., Martogolio, B., Görlich, D., Andersen, S., Ashford, AJ., Giner, A., Hartmann, E., Prehn, S., Rapoport, T., Dobberstein, B., and Brunner, J. (1993). Site-specific Photocross-linking Reveals That Sec61p and TRAM Contact Different Regions of a Membrane-inserted Signal Sequence. *J. Biol. Chem.* **268**:26745-26751.

Hirano, N., Shibasaki, F., Sakai, R., Tanaka, T., Nishida, J., Yazaki, Y., Takenawa, T., and Hirai, H. (1995). Molecular Cloning of the Human Glucose-Regulated Protein ERp57/GRP58 a Thiol-Dependent Reductase. *Eur. J. Biochem.* **234**:336-342.

Holst, B., Bruun, A.W., Kielland-Brandt, M.C., and Winther, J.R. (1996). Competition Between Folding and Glycosylation in the Endoplasmic Reticulum. *EMBO J.* **15**:3538-3546.

Hoppe, T., Matuschewski, K., Rape, M., Schlenker, S., Ulrich, H.D., and Jentsch, S. (2000). Activation of a Membrane-Bound Transcription Factor by Regulated Ubiquitin/Proteasome-Dependent Processing. *Cell* **102**:577-586.

Höhfeld, J. and Jentsch, S. (1997). GrpE-like Regulation of the Hsc70 Chaperone by the Anti-Apoptotic Protein BAG-1. *EMBO J.* **16**:6209-6216.

Hosokawa, N., Wada, I., Hasegawa, K., Yorihuzi, T., Tremblay, L.O., Herscovics, A., and Nagata, K. (2001). A Novel ER α-mannosidase-like Protein Accelerates ER-Associated Degradation. *EMBO Reports* **2**:415-422.

Hughes, E.A. and Cresswell, P. (1998). The Thiol Oxidoreductase ERp57 is a Component of the MHC I Peptide-Loading Complex. *Current Biol.* **8**:709-712.

Hurtley, S.M. and Helenius, A. (1989). Protein Oligomerization in the Endoplasmic Reticulum. *Annu. Rev. Cell Biol*. **5**:277-307.

Hwang, C., Sinskey, A.J., and Lodish, H.F. (1992). Oxidized State of Glutathione in the Endoplasmic Reticulum. *Science* **257**:1496-1502.

Ihara, Y., Cohen-Doyle, M., Saito, A., and Williams, D.B. (1999). Calnexin Discriminates Between Protein Conformational States and Functions as a Molecular Chaperone *In vitro*. *Molecular Cell* **4**:331-341.

Ivessa, N.E., Kitzmüller, C., and de Virgilio, M. (1999) Endoplasmic-Reticulum-Associated Protein Degradation Inside and Outside of the Endoplasmic Reticulum. *Protoplasm* **207**:16-23.

Jakob, C.A., Bodmer, D., Spirig, U., Bättig, P., Marcil, A., Dignard, D., Bergeron, J.J.M., Thomas, D.Y., and Aebi, M. (2001). Htm1p, a Mannosidase-like Protein, Is Involved in Glycoprotein Degradation in Yeast. *EMBO Reports* **2**:423-430.

Jin, T., Gu, Y., Zanusso, G., Sy, M., Kumar, A., Cohen, M., Gambetti, P., and Singh, N. (2000). The Chaperone BiP Binds to a Mutant Prion Protein and Mediates its Degradation by the Proteasome. *J. Biol. Chem.* **275**:38699-38704.

Jin, Y.J., Albers, M.W., Lane, W.S., Bierer, B.E., Schreiber, S.L., and Burakoff, S.J. (1991). Molecular Cloning of a Membrane-Associated Human FK506- and Rapamysin-Binding Protein, FKBP-13. *Proc. Natl. Acad. Sci.* **88**:6677-6681.

Johnson, A.E and van Waes, M.A., (1999). The translocon: A Dynamic Gateway at the ER Membrane. *Annu. Rev. Cell Dev. Biol.* **15**:799-842.

Kabani, M., Beckerich, J., and Gaillardin, C. (2000). Sls1p Stimulates Sec61p-Mediated Activation of Kar2p in a Conformation-Dependent Manner in the Yeast Endoplasmic Reticulum. *Mol. Cell Biol.* **20**:6923-6934.

Kaufman, R.J. (1999). Stress Signalling from the Lumen of the Endoplasmic Reticulum: Coordination of Gene Transcriptional and Translational Controls. *Genes & Develop.* **13**:1211-1233.

Kearse, K.P., Williams, D., and Singer, A. (1994). Persistence of Glucose Residues on Core Oligosaccharides Prevents Association of TCR Alpha and TCR Beta Proteins with

Calnexin and Results Specifically in Accelerated Degradation of Nascent TCR Alpha Proteins Within the Endoplasmic Reticulum. *EMBO J*. **13**:3678-3686.

Keenan, R. J., Freymann, D. M., Walter, P., and Stroud, R. M. (1998). Crystal Structure of the Signal Sequence Binding Subunit of the Signal Recognition Particle. *Cell* **94:**181-191.

Keenan, R.J., Freymann, D.M., Stroud, R.M., and Walter, P. (2001) The Signal Recognition Particle. *Annu. Rev. Biochem.* **70**:755-75.

Kemmink, J., Darby, N.J.; Dijkstra, K.; Nilges, M.; Creighton, T.E. (1996). Structure Determination of the N-terminal Thioredoxin-like Domain of Protein Disulphide Isomerase Using Multidimensional Heteronuclear $^{13}C/^{15}N$ NMR Spectroscopy. *Biochem.* **35**:7684-7691.

Kemmink, J., Darby, N.J., Dijkstra, K., Nilges, M., and Creighton, T.E. (1997). The Folding Catalyst Protein Disulphide Isomerase is Constructed of Active and Inactive Thioredoxin Molecules. *Current Biol.* **7**:239-245.

Kivirikko, K.M., Myllyla, R., and Pihlajaniemi, T. (1989). Protein Hydroxylation: Prolyl 4-hydroxylase, an Enzyme with Four Cosubstrates and a Multifunctional Subunit. *FASEB J* **3**:1609-1617.

Klappa, P., Ruddock, L.W., Darby, N.J., and Freedman, R.B. (1998). The b' Domain Provides the Principal Peptide-Binding Site of Protein Disulphide Isomerase But All Domains Contribute to Binding of Misfolded Proteins. *EMBO J*. **17**:927-935.

Knittler, M., Dirks, R.S., and Haas, I.G. (1995). Molecular Chaperones Involved in Protein Degradation in the Endoplasmic Reticulum: Quantitative Interaction of the Heat Shock Cognate Protein BiP with Partially Folded Immunoglobulin Light Chains that are Degraded in the Endoplasmic Reticulum. *Proc. Natl. Acad. Sci.* **92**:764-768.

Kobayashi, T., Kishigami, S., Sone, M., Inokuchi, H., Mogi, T., and Ito, K. (1997). Respiratory Chain is Required to Maintain Oxidized States of the DsbA-DsbB Disulphide Bond Formation System in Aerobically Growing *Escherichia coli* Cells. *Proc. Natl. Acad. Sci.* **94**:11857-11862.

Koivunen, P., Helaakoski, T., Annunen, P., Veijola, J., Räisänen, S., Pihlajaniemi, T., and Kivirikko, K.I. (1996). ERp60 Does Not Substitute for Protein Disulphide Isomerase as the β Subunit of Prolyl 4-Hydroxylase. *Biochem. J.* **316**:599-605.

Koivunen, P., Pirneskosko, A., Karvonen, P., Ljung, J., Helaakoski, T., Notbohm, H., and Kivirikko, K.I. (1999). The Acidic C-Terminal Domain of Protein Disulphide Isomerase is not Critical for the Enzyme Subunit Function or for the Chaperone or Disulphide Isomerase Activities of the Polypeptide. *EMBO J.* **18**:65-74.

Kurzchalia, T. V., Wiedmann, M., Girshovich, A. S., Bochkareva, E. S., Bielka, H., and Rapoport, T. A. (1986). The Signal Sequence of Nascent Preprolactin Interacts with the 54k Polypeptide of the Signal Recognition Particle. *Nature* **320:**634-636.

Lauffer, L., Garcia, P. D., Harkins, R. N., Coussens, L., Ullrich, A., and Walter, P. (1985). Topology of SRP Receptor in ER Membrane. *Nature* **318:**334-338.

Legate, K.R., Falcone, D., and Andrews, D.A. (2000). Nucleotide-Dependent Binding of the GTPase Domain of the Signal Recognition Particle Receptor β-Subunit to the α-Subunit. *J. Biol. Chem.***275**:27439-27446

Lehner, P.J., and Trowsdale, J. (1998). Antigen Presentation: Coming out Gracefully. *Curr. Biol.* **8**:R605-608.

Lindquist, J.A., Jensen, O.N., Mann, M., and Hämmerling, G.J. (1998). ER-60, a Chaperone With Thiol-Dependent Reductase Activity Involved in MHC Class I Assembly. *EMBO J.* **17**:2186-2195.

Liu, Y., Choudhury, P., Cabral, C.M., and Sifers, R.N. (1999). Oligosaccharide Modification in the Early Secretory Pathway Directs the Selection of a Misfolded Glycoprotein for Degradation by the Proteasome. *J. Biol. Chem.* **274**:5861-5867.

Luders, J., Demand, J., and Höhfeld, J. (2000). The Ubiquitin-Related BAG-1 Provides a Link Between the Molecular Chaperones Hsc70/Hsp70 and the Proteasome. *J. Biol. Chem*. **275**:4613-4617.

Lutcke, H. (1995). SRP, a Ubiquitous Inhibitor of Protein Translocation. *Eur. J. Biochem.* **228:**531-550.

Lyman, S.K. and Schekman, R. (1995). Interaction between BiP and Sec63p is Required for the Completion of Protein Translocation Into the ER of *Saccharomyces cerevisiae.* *J. Cell Biol.* **131**:1163-1171.

Lyman, S.K. and Schekman, R. (1997). Binding of Secretory Precursor Polypeptides to a Translocon Subcomplex is Regulated by BiP. *Cell* **88**:85-96.

Martin, J.L. (1995). Thioredoxin - a Fold for all Reasons. *Structure* **3**:245-250.

Martogolio, B. and Dobberstein, B. (1996). Snapshots of Membrane-Translocating Proteins. *Trends Cell Biol.* **6**:142-147.

Martoglio, B., and Dobberstein, B. (1998). Signal sequences: more than just greasy peptides. *Trends Cell Biol.* **8:**410-415.

Mazzarella, R.A., Crinivasan, M., Haugejorden, S.M., and Green, M. (1990). ERp72, an Abundant Luminal Endoplasmic Reticulum Protein, Contains Three Copies of the Active Site Sequences of Protein Disulphide Isomerase. *J. Biol. Chem.* **265**:1094-1101.

McCracken, A.A. and Brodsky, J.L. (1996). Assembly of ER-Associated Protein Degradation *in vitro*: Dependence on Cytosol, Calnexin and ATP. *J. Cell Biol.* **132**:291-298.

Meacham, G.C., Patterson, C., Zhang, W., Younger, M., and Cyr, D.M. (2001). The Hsc70 Co-Chaperone CHIP Targets Immature CFTR for Proteasomal Degradation. *Nat. Cell Biol.* **3**:100-105.

Melnick, J., Dul, J.L., and Argon, Y. (1994). Sequential Interaction of the Chaperones BiP and GRP94 with Immunoglobulin Chains in the Endoplasmic Reticulum. *Nature* **370**:373-375.

Meyer, D.I. and Dobberstein, B. (1980) Identification and Characterisation of a Membrane Component Essential for the Translocation of Nascent Proteins Across the Membrane of the Endoplasmic Reticulum. *J. Cell Biol.* **87**:503-508.

Mezghrani, A., Fassio, A., Benham, A., Simmen, T., Braakman, I., and Sitia, R. (2001). Manipulation of Oxidative Protein Folding and PDI Redox State in Mammalian Cells. *EMBO J.* **20**:6288-6296.

Miller, J.D., Wilhelm, H., Gierasch, L., Gilmore, R., and Walter, P. (1993) GTP Binding and Hydrolysis by the Signal Recognition Particle During Initiation of Protein Translocation. *Nature* **366**:351-354.

Miller, J.D., Bernstein, H.D., and Walter, P. (1994) Interaction of *E.coli* Ffh/4.5S Ribonucleoprotein and FtsY Mimics that of Mammalian Signal Recognition Particle and its Receptor. *Nature* **367**:657-659.

Miller, J. D., Tajima, S., Lauffer, L., and Walter, P. (1995). The β-subunit of the SRP receptor is a transmembrane GTPase that anchors the α subunit, a peripheral membrane GTPase, to the ER membrane. *J. Cell Biol.* **128:**273-282.

Mobbs, C.V., Fink, G., and Pfaff, D.W. (1990a). HIP-70: a Protein Induced by Estrogen in the Brain and LH-RH in the Pituitary. *Science* **247**:566-567.

Mobbs, C.V., Fink, G., and Pfaff, D.W. (1990b). HIP-70: an Isoform of Phosphoinositol-Specific Phospholipase C-alpha. *Science* **249**:566-567.

Molinari, M. and Helenius, A. (1999a). Chaperone Selection During Glycoprotein Translocation into the Endoplasmic Reticulum. *Science* **288**:331-333.

Molinari, M. and Helenius, A. (1999b). Glycoproteins Form Mixed Disulphides With Oxidoreductases During Folding in Living Cells. *Nature* **402**:90-93.

Molinari, M., Calanca, V., Galli, C., Lucca, P., and Paganetti, P. (2003). Role of EDEM in the Release of Misfolded Glycoproteins from the Calnexin cycle. *Science* **299**:1397-1400.

Montoya, G., Svensson, C., Luirink, J., and Sinning, I. (1997). Crystal structure of the NG domain from the signal-recognition particle receptor FtsY. *Nature* **385:**365-368.

Morrice, N.A. and Powis, S.J. (1998). A Role for the Thiol-Dependent Reductase ERp57 in the Assembly of MHC Class I Molecules. *Current Biol*. **8**:713-716.

Mothes, M., Prehn, S., and Rapoport, TA. (1994). Systematic Probing of the Environment of a Translocating Secretory Protein during Translocation Through the ER Membrane. *EMBO J*. **13**:3973-3982.

Mothes, W., Heinrich, SU., Graf, R., Nilsson, I., Von Heijne, G., Brunner, J., and Rapoport, TA. (1997). Molecular Mechanism of Membrane Protein Integration into the Endoplasmic Reticulum. *Cell* **89**:523-533.

Murthy, M.S., and Pande, S.V. (1993). Carnitine Medium/Long Chain Acyltransferase of Microsomes Seems to be the Previously Cloned Approximately 54kDa Protein of Unknown Function. *Mol. Cell Biochem.* **122**:133-138.

Nakatsukasa, K., Nishikawa, S., Hosokawa, N., Nagata, K., and Endo, T. (2001). Mn1lp, an α-Mannosidase-like Protein in Yeast *Saccharomyces cerevisiae*, Is Required for Endoplasmic Reticulum-Associated Degradation of Glycoproteins. *J. Biol. Chem.* **276**:8635-8638.

Nauseef, W.M., McCormick, S.J., and Clark, R.A. (1995). Calreticulin Functions as a Molecular Chaperone in the Biosynthesis of Myeloperoxidase. *J. Cell Biol*. **270**:4741-4747.

Newitt, J.A., and Bernstein, H.D. (1997). The N-domain of the Signal Recognition Particle 54-kDa Subunit Promotes Efficient Signal Sequence Binding. *Eur. J. Biochem.* **245:**720-729.

Nicchitta, C.V. (1998). Biochemical, Cell Biological and Immunological Issues Surrounding the Endoplasmic Reticulum Chaperone GRP94/gp96. *Curr. Opin. Immunol.* **10**:103-109.

Nicchitta, C.V., and Blobel, G. (1990). Assembly of Translocation-Competant Proteoliposomes from Detergent-Solubilized Rough Microsomes. *Cell* **60**:259-269.

Nikonov, A. V., Snapp, E., Lippincott-Schwartz, J., and Kreibich, G. (2002). Active Translocon Complexes Labeled with GFP-Dad1 Diffuse Slowly as Large Polysome Arrays in the Endoplasmic Reticulum. *J. Cell Biol.* **158**:497-506.

Nishikawa, S., Fewell, S.W., Kato, Y., Brodsky, J.L., and Endo, T. (2001). Molecular Chaperones in the Yeast ER Maintain the Solubility of Proteins for Retrotranslocation and Degradation. *J. Cell Biol.* **153**:1061-1069.

Niwa, M., Sidrauski, C., Kaufman, R.J., and Walter, P. (1999). A Role for Presenilin-1 in Nuclear Accumulation of Ire1 Fragments and Induction of the Mammalian Unfolded Protein Response. *Cell* **99**:691-702.

Oda,Y., Hosokawa, N., Wada, I., Nagata, K. (2003). EDEM As an Acceptor of Terminally Misfolded Glycoproteins Released from Calnexin. *Science* **299**:1394-1397.

Ogg, S. C., and Walter, P. (1995). SRP Samples Nascent Chains for the Presence of Signal Sequences by Interacting with Ribosomes at a Discrete Step During Translation Elongation. *Cell* **81:**1075-1084.

Ogg, S. C., Barz, W. P., and Walter, P. (1998). A Functional GTPase Domain, but not its Transmembrane Domain, is Required for Function of the SRP Receptor beta-subunit. *J. Cell Biol.* **142:**341-354.

Oliver, J., Roderick, H.L., Llewellyn, D.H., and High, S. (1999). ERp57 Functions as a Subunit of Specific Complexes Formed With the ER Lectins Calreticulin and Calnexin. *Mol. Biol.Cell* **10**:2573-2582.

Oliver, J., Van der Wal, F., Bulleid, N.J., and High, S. (1997). Interaction of the Thiol-Dependent Reductase ERp57 with Nascent Glycoproteins. *Science* **275**:86-88.

Ora, A. and Helenius, A. (1995). Calnexin Fails to Associate with Substrate Proteins in Glucosidase-deficient Cell Lines. *J. Biol. Chem.* **270**:26060-26062.

Pagani, M., Fabbri, M., Benedetti, C., Fassio, A., Pilati, S., Bulleid, N.J., Cabibbo, A., and Sitia, R. (2000). Endoplasmic Reticulum Oxidoreduction 1-Lβ (ERO1-Lβ), a Human Gene Induced in the Course of the Unfolded Protein Response. *J. Biol. Chem.* **275**:23685-23692.

Pahl, H.L. (1999). Signal Transduction From the Endoplasmic Reticulum to the Cell Nucleus. *Physiol. Rev.* **79**:683-697.

Parodi, A.J. (2000). Protein Glucosylation and Its Role in Protein Folding. *Annu. Rev. Biochem.* **69**:69-93.

Parodi, A.J., Lederkremer, G.Z., and Mendelzon, D.H. (1983). Protein Glycosylation in *Trypanosoma cruzi*. The Mechanism of Glycosylation and Structure of Protein-Bound Oligosaccharides. *J. Biol. Chem*. **258**:5589-5595.

Partaledis, J.A. and Berlin, V. (1993). The FKB2 Gene of *Saccharomyces cerevisiae*, Encoding the Immunosuppressant-binding Protein FKBP-13, is Regulated in Response to Accumulation of Unfolded Proteins in the Endoplasmic Reticulum. *Proc. Natl. Acad. Sci*. **90**:5450-5454.

Patil, C. and Walter, P. (2001). Intracellular Signaling from the Endoplasmic Reticulum to the Nucleus: the Unfolded Protein Response in Yeast and Mammals. *Curr. Opin. Cell Biol* **13**:349-356.

Peterson, J.R., Ora, A., Van, P.N., and Helenius, A. (1995). Transient, Lectin-Like Association of Calreticulin With Folding Intermediates of Cellular and Viral Glycoproteins. *Mol. Biol. Cell* **6**:1173-1184.

Pirneskoski, A., Ruddock, L.W., Klappa, P., Freedman, R.B., Kivirikko, K.I., and Koivunen, P. (2001). Domains *b′* and *a′* of Protein Disulphide Isomerase Fulfill the Minimum Requirement for Function as a Subunit of Prolyl 4-Hydroylase. J. Biol. Chem. **276**:11287-11293.

Plemper, R.K., and Wolf, D.H. (1999). Retrograde Protein Translocation: ERADication of Secretory Proteins in Health and Disease. *Trends Biochem. Sci.* **24**:266-270.

Plemper, R.K., Wolf, D.H., Bordallo, J., Sommer, T., and Wo (1997). Mutant Analysis Links to Translocon and BiP to Retrograde Protein Transport for ER Degradation. *Nature* **388**:891-895.

Pollard, M.G., Travers, K.J., and Weissman, J.S. (1998). Ero1p: A Novel and Ubiquitous Protein with an Essential Role in Oxidative Protein Folding in the Endoplasmic Reticulum. *Mol. Cell* **1**:171-182.

Powers, T., and Walter, P. (1995) Reciprocal Stimulation of GTP Hydrolysis by Two Directly Interacting GTPases. *Science* **269**:1422-1424.

Puig, A., and Gilbert, H.F. (1994). Protein Disulphide Isomerase Exhibits Chaperone and Anti-Chaperone Activity In the Oxidative Folding of Lysozyme. *J. Biol. Chem.* **269:**7764-7771.

Rapoport, T.A., Matlack, K.E.S., Plath, K., Misselwitz, B., and Staeck, O. (1999). Posttranslational Protein Translocation Across the Membrane of the Endoplasmic Reticulum. *Biol. Chem.* **380**:1143-1150.

Rapoport, TA., Jungickel, B., and Ulrike, K. (1996a). Protein Transport across the Eukaryotic Endoplasmic Reticulum and Bacterial Inner Membranes. *Annu. Rev. Biochem* **65**:271-303.

Rapoport, TA., Rolls, MM., and Jungickel, B. (1996b). Approaching the Mechanism of Protein Transport across the ER Membrane. *Curr. Opin. Cell Biol*. **8**:499-504.

Rietsch, A. and Beckwith, J. (1998). The Genetics of Disulphide Bond Metabolism. *Annu. Rev. Genet.* **32**:163-184.

Ritter, C. and Helenius, A. (2000). Recognition of Local Misfolding by the ER Folding Sensor UDP-Glucose: Glycoprotein Glucosyltransferase. *Nat. Struct. Biol*. **7**:278-280.

Rodan, A.R., Simons, J.F., Trombetta, E.S., and Helenius, A. (1996). N-Linked Oligosaccharides Are Necessary and Sufficient for Association of Glycosylated Forms of Bovine RNase With Calnexin and Calreticulin. *EMBO J.* **15**:6921-6930.

Römisch, K., Webb, J., Lingelbach, K., Gausepohl, H., and Dobberstein, B. (1990). The 54-kDa Protein of Signal Recognition Particle Contains a Methionine-rich RNA Binding Domain. *J. Cell Biol.* **111:**1793-1802.

Saito, Y., Ihara, Y., Leach, M.R., Cohen-Doyle, M.F., and Williams, D.B. (1999). Calreticulin Functions *In vitro* as a Molecular Chaperone for Both Glycosylated and Non-Glycosylated Proteins. *EMBO J.* **18**:6718-6729.

Schmitz, A., Maintz, M., Kehle, T., and Herzog, V. (1995). *In vivo* Iodination of a Misfolded Proinsulin Reveals Co-localized Signals for BiP Binding and for Degradation in the ER. *EMBO J.* **14**:1091-1098.

Schrag, J.D., Bergeron, J.J.M., Li, Y., Borisova, S., Hahn, M., Thomas, D.Y., and Cygler, M. (2001). The Structure of Calnexin, an ER Chaperone Involved in Quality Control of Protein Folding. *Mol. Cell* **8**:633-644.

Sevier, C.S., Cuozzo, J.W., Vala, A., Åslund, F., and Kaiser, C.A. (2001). A Flavoprotein Oxidase Defines a New Endoplasmic Reticulum Pathway for Biosynthetic Disulphide Bond Formation. *Nat. Cell Biol.* **3**:874-882.

Sidrauski, C., Chapman, R., and Walter, P. (1998). The Unfolded Protein Response: an Intracellular Signalling Pathway with Many Surprising Features. *Trends Biochem. Sci.* **8**:245-249.

Siegel, V., and Walter, P. (1985). Elongation Arrest is not a Prerequisite for Secretory Protein Translocation Across the Microsomal Membrane. *J. Cell Biol.* **100:**1913-1921.

Skowronek, M.H., Hendershot, L.M., and Haas, I.G. (1998). The Variable Domain of Nonassembled Ig Light Chains Determines Both Their Half-Life and Binding to the Chaperone BiP. *Proc. Natl. Acad. Sci*. **95**:1574-1578.

Song, W., Raden, D., Mandon, E.C., and Gilmore, R. (2000). Role of Sec61α in the Regulated Transfer of the Ribosome-Nascent Chain Complex From the Signal Recognition Particle to the Translocation Channel. *Cell* **100**:333-343.

Sousa, M., Ferrero-Garcia, M.A., and Parodi, A.J. (1992). Recognition of the Oligosaccharide and Protein Moieties of Glycoproteins by the UDP-Glc:Glycoprotein Glucosyltranserase. *Biochemistry* **31**:97-105.

Spee, P., and Neefjes, J. (1997). TAP-Translocated Peptides Specifically Bind Proteins in the Endoplasmic Reticulum, Including gp96, Protein Disulphide Isomerase and Calreticulin. *Eur. J. Immunol.* **27**:2441-2449.

Stevens, F.J. and Argon, Y. (1999). Protein Folding in the ER. *Seminars Cell & Dev. Biol.* **10**:443-454.

Suzuki, T., Park, H., Hollingworth, N.M., Sternglanz, R., and Lennarz, W.J. (2000). *PNG1*, a Yeast Gene Encoding a Highly Conserved Peptide: N-Glycanase. *J. Cell Biol.* **149**:1039-1051.

Tajima, S., Lauffer, L., Rath, V. L., and Walter, P. (1986). The Signal Recognition Particle Receptor is a Complex that Contains Two Distinct Polypeptide Chains. *J. Cell Biol.* **103:**1167-1178.

Tessier, D.C., Dignard, D., Zapun, A., Radominska-Pandya, A., Parodi, A.J., Bergeron, J.J.M., and Thomas, D.Y. (2000). Cloning and Characterisation of Mammalian UDP-Glucose Glycoprotein: Glucosyltransferase and the Development of a Specific Substrate for this Enzyme. *Glycobiology* **10**:403-412.

Thomas, P.J., Qu, B., and Pedersen, P.L. (1995). Defective Protein Folding as a Basis of Human Disease. *Trends Biochem. Sci* **20**:456-459.

Tokunaga, F., Brostrom, C., Koide, T., and Arvan, P. (2000). Endoplasmic Reticulum (ER)-associated Degradation of Misfolded N-linked Glycoproteins is Suppressed upon Inhibition of ER Mannosidase I. *J. Biol. Chem.* **275**:40757-40764.

Travers, K.J., Patil, C.K., Wodicka, L., Lockhart, D.J., Weissman, J.S., and Walter, P. (2000). Functional and Genomic Analyses Reveal an Essential Coordination Between the Unfolded Protein Response and ER-Associated Degradation. *Cell* **101**:249-258.

Trombetta, E.S. and Helenius, A. (1998). Lectins as Chaperones in Glycoprotein Folding. *Curr. Biol.* **8**:587-592.

Trombetta, E.S. and Helenius, A. (2000). Conformational Requirements for Glycoprotein Reglucosylation in the Endoplasmic Reticulum. *J. Cell Biol.* **148**:1123-1129.

Tu, B., Ho-Schleyer, S.C., Travers, K.J., and Weissman, J.S. (2000). Biochemical Basis of Oxidative Folding in the Endoplasmic Reticulum. *Science* **290**:1571-1574.

Tu, B.P., and Weissman, J.S. (2002). The FAD- and O_2-Dependent Reaction Cycle of Ero1-Mediated Oxidative Protein Folding in the Endoplasmic Reticulum. *Mol. Cell* **10**:983-994.

Tyedmers, J., Lerner, M., Bies, C., Dudek, J., Skowronek, M. H., Haas, I. G., Nastainczyk, W., Volkmer, J., Zimmermann, R. (2000). Homologs of the Yeast Sec Complex Subunits Sec62p and Sec63p are Abundant Proteins in Dog Pancreas Microsomes. *Proc. Nat. Acad. Sci.* **97**:7214-7219

Tyson, J.R. and Stirling, C.J. (2000). LHS1 and SIL1 Provide Lumenal Function that is Essential for Protein Translocation into the Endoplasmic Reticulum. *EMBO J*. **30**:794-797.

Updyke, T.V. and Nicholson, G.L. (1984). Immunoaffinity Isolation of Membrane Antigens with Biotinylated Monoclonal Antibodies and Immobilized Streptavidin Matrices. *J. Immunol. Meths*. **73**:83-95.

Urade, R. and Kito, M. (1992). Inhibition by Acidic Phospholipids of Protein Degradation by ER-60 Protease, a Novel Cysteine Protease, of Endoplasmic Reticulum. *FEBS Letters* **312**:83-86.

Van der Wal, F., Oliver, J., and High, S. (1998). The Transient Association of ERp57 With N-Glycosylated Proteins is Regulated by Glucose Trimming. *Eur. J .Biochem*. **256**:51-59.

Van Leeuwen, J.E. and Kearse, K.P. (1996). Deglucosylation of N-Linked Glycans is an Important Step in the Dissociation of Calreticulin-Class I-TAP Complexes. *Proc. Natl. Acad. Sci.* **93**:13997-14001.

Vassilakos, A., Michalak, M., Lehrman, M., and Williams, D. (1998). Oligosaccharide Binding Characteristics of the Molecular Chaperones Calnexin and Calreticulin. *Biochem.* **37**:3480-3490.

Voges, D., Zwickel, P., and Baumeister, W. (1999). The 26S Proteasome: a Molecular Machine Designed for Controlled Proteolysis. *Annu. Rev. Biochem* **68**:1015-1068.

von Heijne, G.(1985). Protein Sorting Signals: Simple Peptides with Complex Functions. *EXS* **73**:67-76.

von Heijne, G. (1990). The signal peptide. *J. Membr. Biol.* **115**:195-201.

Walter, P., and Blobel, G. (1981). Translocation of Proteins Across the Endoplasmic Reticulum. III. Signal Recognition Protein (SRP) Causes Signal-sequence Dependent and

Site-specific Arrest of Chain Elongation that is Released by Microsomal Membranes. *J. Cell Biol.* **91**:557-561.

Ware, F.E, Vassilakos, A., Peterson, P. A., Jackson, M. R., Lehrman, M A., and Williams, D. B. (1995). The Molecular Chaperone Calnexin Binds $Glc_1Man_9GlcNAc_2$ Oligosaccharide as an Initial Step in Recognizing Unfolded Glycoproteins. *J. Biol. Chem.* **270**:4697-4704.

Wearsch, P. A., Nicchitta, C. V. (1996). Endoplasmic Reticulum Chaperone GRP94 Subunit Assembly is Regulated Through a Defined Oligomerization Domain. *Biochem.* **35**:16760-16769.

Wearsch, P. A., Nicchitta, C. V. (1997). Interaction of Endoplasmic Reticulum Chaperone GRP94 with Peptide Substrates is Adenine Nucleotide-Independent. *J. Biol. Chem.* **272**:5152-5156.

Wei, J. and Hendershot, L.M. (1996). Protein Folding and Assembly in the Endoplasmic Reticulum. *EXS* **77**:41-55.

Weihofen, A., Lemberg, M.K., Ploegh, H.L., Bogyo, M., and Martoglio, B. (2000). Release of Signal Peptide Fragments into the Cytosol Requires Cleavage in the Transmembrane Region by a Novel Cysteine Protease Inhibitor. *J. Biol. Chem.* **275**:30951-30956.

Wetterau, J.R., Combs, K.A., McLean, L.R., Spinner, S.N., and Aggerbeck, L.P. (1991). Protein Disulphide Isomerase Appears Necessary to Maintain the Catalytically Active Structure of the Microsomal Triglyceride Transfer Protein. *Biochem*. **30**:9728-9735.

Wickner, S., Maurizi, M.R., and Gottesman, S. (1999). Posttranslational Quality Control: Folding, Refolding and Degrading Proteins. *Science* **286**:1888-1893.

Wiedmann, M., Kurzchalia, T. V., Bielka, H., and Rapoport, T. A. (1987a). Direct Probing of the Interaction Between the Signal Sequence of Nascent Preprolactin and the Signal Recognition Particle by Specific Cross-linking. *J. Cell Biol.* **104**:201-209.

Wiertz, E.J., Tortorella, D., Bogyo, M., Yu, J., and Mothes, W. (1996). Sec61p-mediated Transfer of a Membrane Protein from the Endoplasmic Reticulum to the Proteasome for Destruction. *Nature* **384**:432-438.

Wilkinson, BM., Regnacq, M., and Stirling, CJ. (1997) Protein Translocation Across the Membrane of the Endoplasmic Reticulum. *J. Memb. Biol.* **155**:189-197.

Williams, D.B. (1995). Calnexin: A Molecular Chaperone with a Taste for Carbohydrate. *Biochem. Cell Biol.* **73**:123-132.

Wilson, R., Allen, A.J., Oliver, J., Brookman, J.L., High, S., and Bulleid, N.J. (1995). The Translocation, Folding, Assembly and Redox-Dependent Degradation of Secretory and Membrane Proteins in Semi-Permeabilised Mammalian Cells. *Biochem. J.* **307**:679-687.

Wilson, C.M., Farmery, M.R., and Bulleid, N.J. (2000). Pivotal Role of Calnexin and Mannose Trimming in Regulation the Endoplasmic Reticulum-associated Degradation of Major Histocompatability Complex Class I Heavy Chain. *J. Biol. Chem.* **275**:21224-21232.

Wolin, S. L., and Walter, P. (1989). Signal Recognition Particle Mediates a Transient Elongation Arrest of Preprolactin in Reticulocyte Lysate. *J. Cell Biol.* **109**:2617-2622.

Zapun A, Darby, N.J., Tessier, D.C., Michalak, M., Bergeron, J.J.M., and Thomas D.Y (1998). Enhanced Catalysis of Ribonuclease B Folding by the Interaction of Calnexin or Calreticulin with ERp57. *J. Biol. Chem.* **273**:6009-6012.

Zapun, A., Bardwell, J.C., and Creighton, T.E. (1993). The Reactive and Destabilizing Disulphide Bond of DsbA, a Protein Required for Protein Disulphide Bond Formation *in vivo*. *Biochem*. **32**:5083-5092.

Zapun, A., Jakob, C., and Thomas, D.Y. (1999). Protein Folding in a Specialized Compartment: the Endoplasmic Reticulum. *Structure* **7**:R173-R182

Zapun, A., Petrescu, S.M., Rudd, P.M., Dwek, R.A., Thomas, D.Y., and Bergeron, J.J.M. (1997). Conformation-Independent Binding of Monoglucosylated Ribonuclease B to Calnexin. *Cell* **88**:29-38.

Zhang, J.X., Braakman, I., Matlack, K.E., and Helenius, A. (1997). Quality Control in the Secretory Pathway: the Role of Calreticulin, Calnexin and BiP in the Retention of Glycoproteins with C-terminal Truncations. *Mol. Biol. Cell* **8**:1943-1954.

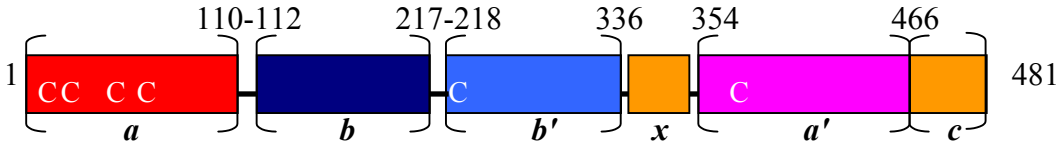
Zopf, D., Bernstein, H. D., Johnson, A. E., and Walter, P. (1990). The Methionine-Rich Domain of the 54kD Protein Subunit of the Signal Recognition Particle Contains an RNA Binding Site and can be Cross-linked to a Signal Sequence. *EMBO J.* **9**:4511-4517.

Zopf, D., Bernstein, H. D., and Walter, P. (1993). GTPase domain of the 54kDa Subunit of the Mammalian signal Recognition Particle is Required for Protein Translocation but not for Signal Sequence Binding. *J. Cell Biol.* **120**:1113-1121.

Zuber, C., Fan, J., Guhl, B., Parodi, A., Fessler, J.H., Parker, C., and Roth, J. (2001). Immunolocalization of UDP-Glucose:Glycoprotein Glucosyltransferase Indicates Involvement of Pre-Golgi Intermediates in Protein Quality Control. *Proc. Natl. Acad. Sci.* **98**:10710-10715.

Appendix

1.1 ERp57



Contains 7 cysteines (approximate placement shown above in white and underlined below) and 48 lysines. The ***a*** domain and ***a'*** domains are shown in red and pink respectively, while the ***b*** and ***b'*** domains are shaded navy and light blue respectively. Both the ***x***-linker region and the C-terminal extension are shown in orange.

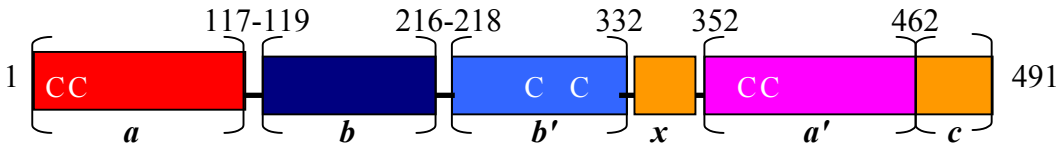
Amino Acid Sequence

```
  1 MRLRRLALFP GVALLLAAAR LAAASDVLEL TDDNFESRIS DTGSAGLMLV EFFAPWCGHC   60
 61 KRLAPEYEAA ATRLKGIVPL AKVDCTANTN TCNKYGVSGY PTLKIFRDGE EAGAYDGPRT  120
121 ADGIVSHLKK QAGPASVPLR TEEEFKKFIS DKDASIVGFF DDSFSEAHSE FLKAASNLRD  180
181 NYRFAHTNVE SLVNEYDDNG EGIILFRPSH LTNKFEDKTV AYTEQKMTSG KIKKFIQENI  240
241 FGICPHMTED NKDLIQGKDL LIAYYDVDYE KNAKGSNYWR NRVMMVAKKF LDAGHKLNFA  300
301 VASRKTFSHE LSDFGLESTA GEIPVVAIRT AKGEKFVMQE EFSRDGKALE RFLQDYFDGN  360
361 LKRYLKSEPI PESNDGPVKV VVAENFDEIV NNENKDVLIE FYAPWCGHCK NLEPKYKELG  420
421 EKLSKDPNIV IAKMDATAND VPSPYEVRGF PTIYFSPANK KLNPKKYEGG RELSDFISYL  480
481 QREATNPPVI QEEKPKKKKK AQEDL
```

DNA Sequence

```
atgcgcctcc gccgcctagc gctgttcccg ggtgtggcgc tgcttcttgc cgcggcccgc   60
ctcgccgctg cctccgacgt gctagaactc acggacgaca acttcgagag tcgcatctcc  120
gacacgggct ctgcgggcct catgctcgtc gagttcttcg ccccctggtg tggacactgc  180
aagagacttg cacctgagta tgaagctgca gctaccagat taaaagggat agtcccatta  240
gcaaaggttg attgcactgc caacactaac acctgtaata aatatggagt cagtggatat  300
ccaaccctga agatatttag agatggtgaa gaagcaggtg cttatgatgg acctaggact  360
gctgatggaa ttgtcagcca cttgaagaag caggcaggac cagcttcagt gcctctcagg  420
actgaggaag aatttaagaa attcattagt gataaagatg cctctatagt aggtttttc   480
gatgattcat tcagtgaggc tcactccgag ttcctaaaag cagccagcaa cttgagggat  540
aactaccgat ttgcacatac gaatgttgag tctctggtga acgagtatga tgataatgga  600
gagggtatca tcttatttcg tccttcacat ctcactaaca agtttgagga caagactgtg  660
gcatatacag agcaaaaaat gaccagtggc aaaattaaaa agtttatcca ggaaaacatt  720
ttggtatct gccctcacat gacagaagac aataaagatt tgatacaggg caaggactta  780
cttattgctt actatgatgt ggactatgaa aagaacgcta aaggttccaa ctactggaga  840
aacagggtaa tgatggtggc aaagaaattc ctggatgctg ggcacaaact caactttgct  900
gtagctagcc gcaaaacctt tagccatgaa ctttctgatt ttggcttgga gagcactgct  960
ggagagattc ctgttgttgc tatcagaact gctaaaggag agaagtttgt catgcaggag 1020
gagttctcgc gtgatgggaa ggctctggag aggttcctgc aggattactt tgatggcaat 1080
ctgaagagat acctgaagtc tgaacctatc ccagagagca atgatgggcc tgtgaaggta 1140
gtggtagcag agaattttga tgaaatagtg aataatgaaa ataaagatgt gctgattgaa 1200
ttttatgccc cttggtgtgg tcattgtaag aacctggagc ccaagtataa agaacttggc 1260
gagaagctca gcaaagaccc aaatatcgtc atagccaaga tggatgccac agccaatgat 1320
gtgccttctc catatgaagt cagaggtttt cctaccatat acttctctcc agccaacaag 1380
aagctaaatc caaagaaata tgaaggtggc cgtgaattaa gtgattttat tagctatcta 1440
caaagagaag ctacaaaccc ccctgtaatt caagaagaaa aacccaagaa gaagaagaag 1500
gcacaggagg atctctaa
```

1.2 PDI



Contains 6 cysteines (approximate placement shown above in white and underlined below) and 47 lysines. The ***a*** domain and ***a'*** domains are shown in red and pink respectively, while the ***b*** and ***b'*** domains are shaded navy and light blue respectively. Both the ***x***-linker region and the C-terminal extension are shown in orange.

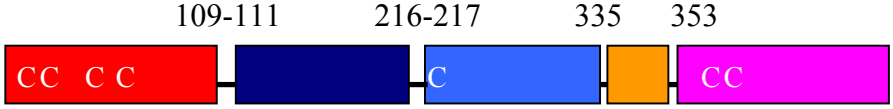
Amino Acid Sequence

```
  1 MLRRALLCLA VAALVRADAP EEEDHVLVLR KSNFAEALAA HKYLLVEFYA PWCGHCKALA   60
 61 PEYAKAAGKL KAEGSEIRLA KVDATEESDL AQQYGVRGYP TIKFFRNGDT ASPKEYTAGR  120
121 EADDIVNWLK KRTGPAATTL PDGAAAESLV ESSEVAVIGF FKDVESDSAK QFLQAAEAID  180
181 DIPFGITSNS DVFSKYQLDK DGVVLFKKFD EGRNNFEGEV TKENLLDFIK HNQLPLVIEF  240
241 TEQTAPKIFG GEIKTHILLF LPKSVSDYDG KLSNFKTAAE SFKGKILFIF IDSDHTDNQR  300
301 ILEFFGLKKE ECPAVRLITL EEEMTKYKPE SEELTAERIT EFCHRFLEGK IKPHLMSQEL  360
361 PEDWDKQPVK VLVGKNFEDV AFDEKKNVFV EFYAPWCGHC KQLAPIWDKL GETYKDHENI  420
421 VIAKMDSTAN EVEAVKVHSF PTLKFFPASA DRTVIDYNGE RTLDGFKKFL ESGGQDGAGD  480
481 DDDLEDLEEA EEPDMEEDDD QKAVKDEL
```

DNA sequence

```
atgctgcgcc gcgctctgct gtgcctggcc gtggccgccc tggtgcgcgc cgacgccccc    60
gaggaggagg accacgtcct ggtgctgcgg aaaagcaact tcgcggaggc gctggcggcc   120
cacaagtacc tgctggtgga gttctatgcc ccttggtgtg gccactgcaa ggctctggcc   180
cctgagtatg ccaaagccgc tgggaagctg aaggcagaag gttccgagat caggttggcc   240
aaggtggacg ccacggagga gtctgacctg gcccagcagt acggcgtgcg cggctatccc   300
accatcaagt tcttcaggaa tggagacacg gcttccccca aggaatatac agctggcaga   360
gaggctgatg acatcgtgaa ctggctgaag aagcgcacgg gccggctgtc caccaccctc   420
cgtgacggcg cagctgcaga gtccttggtg gagtccagcg aggtggctgt catcggcttc   480
ttcaaggacg tggagtcgga ctctgccaag cagtttttgc aggcagcaga ggccatcgat   540
gacataccat ttgggatcac ttccaacagt gacgtgttct ccaaatacca gctcgacaaa   600
gatggggttg tcctctttaa gaagtttgat gaaggccgga acaactttga aggggaggtc   660
accaaggaga acctgctgga ctttatcaaa cacaaccagc tgcccttgt catcgagttc   720
accgagcaga cagccccgaa gatttttgga ggtgaaatca agactcacat cctgctgttc   780
ttgcccaaga gtgtgtctga ctatgacggc aaactgagca acttcaaaac agcagccgag   840
agcttcaagg gcaagatcct gttcatcttc atcgacagcg accacaccga caaccagcgc   900
atcctcgagt tctttggcct gaagaaggaa gagtgcccgg ccgtgcgcct catcaccctg   960
gaggaggaga tgaccaagta caagcccgaa tcggaggagc tgacggcaga gaggatcaca  1020
gagttctgcc accgcttcct ggaggccaaa atcaagcccc acctgatgag ccaggagcgt  1080
gccggagact gggacaagca gcctgtcaag gtgcctgttg ggaagaactt tgaagacgtg  1140
gcttttgatg agaaaaaaaa cgtctttgtg gagttctatg ccccatggtg tggtcactgc  1200
aaacagttgg ctcccatttg ggataaactg ggagagacgt acaaggacca tgagaacatc  1260
gtcatcgcca agatggactc gactgccaac gaggtggagg ccgtcaaagt gcacagcttc  1320
cccacactca agttctttcc tgccagtgcc gacaggacgg tcattgatta caacggggaa  1380
cgcacgctgg atggttttaa gaaattcctg gagagcggtg gccaggatgg ggcaggggat  1440
gatgacgatc tcgaggacct ggaagaagca gaggagccag acatggagga agacgatgat  1500
cagaaagctg tgaaagatga actgtaa
```

1.3 ΔERp57 (ERp57 ***abb'a'***)



1 465

a ***b*** ***b'*** ***x*** ***a'***

Contains 7 cysteines (approximate placement shown above in white and underlined below) and 45 lysines. The ***a*** domain and ***a'*** domains are shown in red and pink respectively, while the ***b*** and ***b'*** domains are shaded navy and light blue respectively. Both the ***x***-linker region is shown in orange.

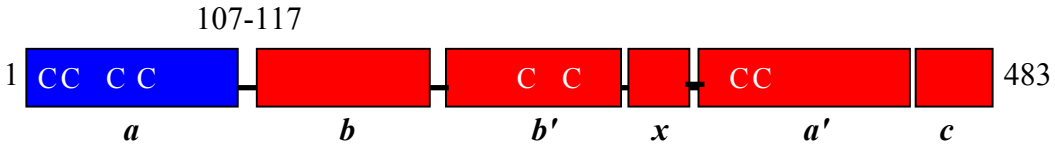
Amino Acid Sequence

```
  1 MRLRRLALFP GVALLLAAAR LAAASDVLEL TDDNFESRIS DTGSAGLMLV EFFAPWCGHC   60
 61 KRLAPEYEAA ATRLKGIVPL AKVDCTANTN TCNKYGVSGY PTLKIFRDGE EAGAYDGPRT  120
121 ADGIVSHLKK QAGPASVPLR TEEEFKKFIS DKDASIVGFF DDSFSEAHSE FLKAASNLRD  180
181 NYRFAHTNVE SLVNEYDDNG EGIILFRPSH LTNKFEDKTV AYTEQKMTSG KIKKFIQENI  240
241 FGICPHMTED NKDLIQGKDL LIAYYDVDYE KNAKGSNYWR NRVMMVAKKF LDAGHKLNFA  300
301 VASRKTFSHE LSDFGLESTA GEIPVVAIRT AKGEKFVMQE EFSRDGKALE RFLQDYFDGN  360
361 LKRYLKSEPI PESNDGPVKV VVAENFDEIV NNENKDVLIE FYAPWCGHCK NLEPKYKELG  420
421 EKLSKDPNIV IAKMDATAND VPSPYEVRGF PTIYFSPANK KLNPKKYEGG RELSDFISYL  480
481 QREATNPPVI
```

DNA Sequence

```
atgcgcctcc gccgcctagc gctgttcccg ggtgtggcgc tgcttcttgc cgcggcccgc   60
ctcgccgctg cctccgacgt gctagaactc acggacgaca acttcgagag tcgcatctcc  120
gacacgggct ctgcgggcct catgctcgtc gagttcttcg ccccctggtg tggacactgc  180
aagagacttg cacctgagta tgaagctgca gctaccagat taaaagggat agtcccatta  240
gcaaaggttg attgcactgc caacactaac acctgtaata aatatggagt cagtggatat  300
ccaaccctga agatatttag agatggtgaa gaagcaggtg cttatgatgg acctaggact  360
gctgatggaa ttgtcagcca cttgaagaag caggcaggac cagcttcagt gcctctcagg  420
actgaggaag aatttaagaa attcattagt gataaagatg cctctatagt aggtttttc   480
gatgattcat tcagtgaggc tcactccgag ttcctaaaag cagccagcaa cttgagggat  540
aactaccgat ttgcacatac gaatgttgag tctctggtga acgagtatga tgataatgga  600
gagggtatca tcttatttcg tccttcacat ctcactaaca agtttgagga caagactgtg  660
gcatatacag agcaaaaaat gaccagtggc aaaattaaaa agtttatcca ggaaaacatt  720
tttggtatct gccctcacat gacagaagac aataaagatt tgatacaggg caaggactta  780
cttattgctt actatgatgt ggactatgaa aagaacgcta aaggttccaa ctactggaga  840
aacagggtaa tgatggtggc aaagaaattc ctggatgctg ggcacaaact caactttgct  900
gtagctagcc gcaaaacctt tagccatgaa ctttctgatt ttggcttgga gagcactgct  960
ggagagattc ctgttgttgc tatcagaact gctaaaggag agaagtttgt catgcaggag 1020
gagttctcgc gtgatgggaa ggctctggag aggttcctgc aggattactt tgatggcaat 1080
ctgaagagat acctgaagtc tgaacctatc ccagagagca atgatgggcc tgtgaaggta 1140
gtggtagcag agaattttga tgaaatagtg aataatgaaa ataaagatgt gctgattgaa 1200
ttttatgccc cttggtgtgg tcattgtaag aacctggagc ccaagtataa agaacttggc 1260
gagaagctca gcaaagaccc aaatatcgtc atagccaaga tggatgccac agccaatgat 1320
gtgccttctc catatgaagt cagaggtttt cctaccatat acttctctcc agccaacaag 1380
aagctaaatc caaagaaata tgaaggtggc cgtgaattaa gtgattttat tagctatcta 1440
caaagagaag ctacaaaccc ccctgtaatt
```

1.4 Construct 1 (ERp57 ***a*** PDI ***bb'a'c***)



Contains 8 cysteines (approximate placement shown above in white and underlined below) and 43 lysines. ERp57 and PDI domains shown in blue and red respectively.

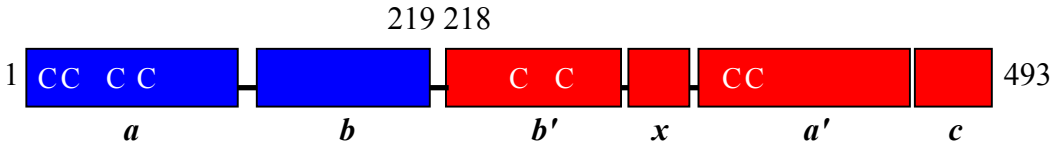
Amino Acid Sequence

```
  1 MRLRRLALFP GVALLLAAAR LAAASDVLEL TDDNFESRIS DTGSAGLMLV EFFAPWCGHC    60
 61 KRLAPEYEAA ATRLKGIVPL AKVDCTANTN TCNKYGVSGY PTLKIFRDGE EAGAYDGPRT   120
121 ADGIVSHLKK QAGPAATTL PDGAAAESLV ESSEVAVIGF FKDVESDSAK QFLQAAEAID   180
181 DIPFGITSNS DVFSKYQLDK DGVVLFKKFD EGRNNFEGEV TKENLLDFIK HNQLPLVIEF   240
241 TEQTAPKIFG GEIKTHILLF LPKSVSDYDG KLSNFKTAAE SFKGKILFIF IDSDHTDNQR   300
301 ILEFFGLKKE ECPAVRLITL EEEMTKYKPE SEELTAERIT EFCHRFLEGK IKPHLMSQEL   360
361 PEDWDKQPVK VLVGKNFEDV AFDEKKNVFV EFYAPWCGHC KQLAPIWDKL GETYKDHENI   420
421 VIAKMDSTAN EVEAVKVHSF PTLKFFPASA DRTVIDYNGE RTLDGFKKFL ESGGQDGAGD   480
481 DDDLEDLEEA EEPDMEEDDD QKAVKDEL
```

DNA Sequence

```
atgcgcctcc gccgcctagc gctgttcccg ggtgtggcgc tgcttcttgc cgcggcccgc        60
ctcgccgctg cctccgacgt gctagaactc acggacgaca acttcgagag tcgcatctcc       120
gacacgggct ctgcgggcct catgctcgtc gagttcttcg ccccctggtg tggacactgc       180
aagagacttg cacctgagta tgaagctgca gctaccagat taaaagggat agtcccatta       240
gcaaaggttg attgcactgc caacactaac acctgtaata aatatggagt cagtggatat       300
ccaaccctga agatatttag agatggtgaa gaagcaggtg cttatgatgg acctaggact       360
gctgatggaa ttgtcagcca cttgaagaag caggcgggcc cgccggctgc caccaccctc       420
cgtgacggcg cagctgcaga gtccttggtg gagtccagcg aggtggctgt catcggcttc       480
ttcaaggacg tggagtcgga ctctgccaag cagtttttgc aggcagcaga ggccatcgat       540
gacataccat ttgggatcac ttccaacagt gacgtgttct ccaaatacca gctcgacaaa       600
gatggggttg tcctctttaa gaagtttgat gaaggccgga acaactttga aggggaggtc       660
accaaggaga acctgctgga ctttatcaaa cacaaccagc tgccccttgt catcgagttc       720
accgagcaga cagccccgaa gatttttgga ggtgaaatca agactcacat cctgctgttc       780
ttgcccaaga gtgtgtctga ctatgacggc aaactgagca acttcaaaac agcagccgag       840
agcttcaagg gcaagatcct gttcatcttc atcgacagcg accacaccga caaccagcgc       900
atcctcgagt tctttggcct gaagaaggaa gagtgcccgg ccgtgcgcct catcaccctg       960
gaggaggaga tgaccaagta caagcccgaa tcggaggagc tgacggcaga gaggatcaca      1020
gagttctgcc accgcttcct ggagggcaaa atcaagcccc acctgatgag ccaggagcgt      1080
gccggagact gggacaagca gcctgtcaag gtgcctgttg ggaagaactt tgaagacgtg      1140
gcttttgatg agaaaaaaaa cgtctttgtg gagttctatg ccccatggtg tggtcactgc      1200
aaacagttgg ctcccatttg ggataaactg ggagagacgt acaaggacca tgagaacatc      1260
gtcatcgcca agatggactc gactgccaac gaggtggagg ccgtcaaagt gcacagcttc      1320
cccacactca agttctttcc tgccagtgcc gacaggacgg tcattgatta caacggggaa      1380
cgcacgctgg atggttttaa gaaattcctg gagagcggtg gccaggatgg ggcaggggat      1440
gatgacgatc tcgaggacct ggaagaagca gaggagccag acatggagga agacgatgat      1500
cagaaagctg tgaaagatga actgtaa
```

1.5 Construct 2 (ERp57 ***ab*** PDI ***b'a'c***)



Contains 8 cysteines (approximate placement shown above in white and underlined below) and 45 lysines. ERp57 and PDI domains shown in blue and red respectively.

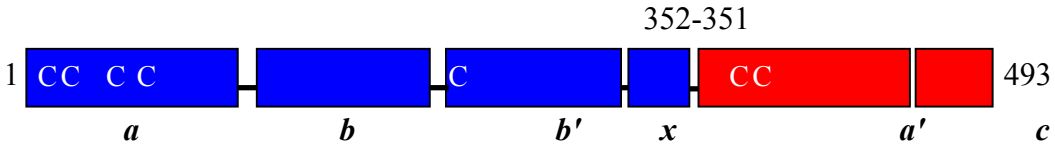
Amino Acid Sequence

```
  1 MRLRRLALFP GVALLLAAAR LAAASDVLEL TDDNFESRIS DTGSAGLMLV EFFAPWCGHC   60
 61 KRLAPEYEAA ATRLKGIVPL AKVDCTANTN TCNKYGVSGY PTLKIFRDGE EAGAYDGPRT  120
121 ADGIVSHLKK QAGPASVPLR TEEEFKKFIS DKDASIVGFF DDSFSEAHSE FLKAASNLRD  180
181 NYRFAHTNVE SLVNEYDDNG EGIILFRPSH LTNKFEDKTV AYTEQKMTSG KIKKFIQENI  240
241 FGIPLVIEFT EQTAPKIFGG EIKTHILLFL PKSVSDYDGK LSNFKTAAES FKGKILFIFI  300
301 DSDHTDNQRI LEFFGLKKEE CPAVRLITLE EEMTKYKPES EELTAERITE FCHRFLEGKI  360
361 KPHLMSQELP EDWDKQPVKV LVGKNFEDVA FDEKKNVFVE FYAPWCGHCK QLAPIWDKLG  420
421 ETYKDHENIV IAKMDSTANE VEAVKVHSFP TLKFFPASAD RTVIDYNGER TLDGFKKFLE  480
481 SGGQDGAGDD DDLEDLEEAE EPDMEEDDDQ KAVKDEL
```

DNA Sequence

```
atgcgcctcc gccgcctagc gctgttcccg ggtgtggcgc tgcttcttgc cgcggcccgc   60
ctcgccgctg cctccgacgt gctagaactc acggacgaca acttcgagag tcgcatctcc  120
gacacgggct ctgcgggcct catgctcgtc gagttcttcg ccccctggtg tggacactgc  180
aagagacttg cacctgagta tgaagctgca gctaccagat taaaagggat agtcccatta  240
gcaaaggttg attgcactgc caacactaac acctgtaata aatatggagt cagtggatat  300
ccaaccctga agatatttag agatggtgaa gaagcaggtg cttatgatgg acctaggact  360
gctgatggaa ttgtcagcca cttgaagaag caggcaggac cagcttcagt gcctctcagg  420
actgaggaag aatttaagaa attcattagt gataaagatg cctctatagt aggttttttc  480
gatgattcat tcagtgaggc tcactccgag ttcctaaaag cagccagcaa cttgagggat  540
aactaccgat ttgcacatac gaatgttgag tctctggtga acgagtatga tgataatgga  600
gagggtatca tcttatttcg tccttcacat ctcactaaca agtttgagga caagactgtg  660
gcatatacag agcaaaaaat gaccagtggc aaaattaaaa agtttatcca ggaaaacatt  720
tttggtatcc ttgtcatcga gttcaccgag cagacagccc cgaagatttt tggaggtgaa  780
atcaagactc acatcctgct gttcttgccc aagagtgtgt ctgactatga cggcaaactg  840
agcaagcaac ttcaaaacag cagccgagag cttcaagggc aagatcctgt tcatcttcat  900
cgacagcgac cacaccgaca accagcgcat cctcgagttc tttggcctga agaaggaaga  960
gtgcccggcc gtgcgcctca tcaccctgga ggaggagatg accaagtaca agcccgaatc 1020
ggaggagctg acggcagaga ggatcacaga gttctgccac cgcttcctgg agggcaaaat 1080
caagccccac ctgatgagcc aggagcgtgc cggagactgg gacaagcagc ctgtcaaggt 1140
gcctgttggg aagaactttg aagacgtggc ttttgatgag aaaaaaaacg tctttgtgga 1200
gttctatgcc ccatggtgtg gtcactgcaa acagttggct cccatttggg ataaactggg 1260
agagacgtac aaggaccatg agaacatcgt catcgccaag atggactcga ctgccaacga 1320
ggtggaggcc gtcaaagtgc acagcttccc cacactcaag ttctttcctg ccagtgccga 1380
caggacggtc attgattaca acggggaacg cacgctggat ggttttaaga aattcctgga 1440
gagcggtggc caggatgggg caggggatga tgacgatctc gaggacctgg aagaagcaga 1500
ggagccagac atggaggaag acgatgatca gaaagctgtg aaagatgaac tgtaa
```

1.6 Construct 3 - ERp57 ***abb'x*** PDI ***a'c***



Contains 7 cysteines (approximate placement shown above in white and underlined below) and 44 lysines. ERp57 and PDI domains shown in blue and red respectively.

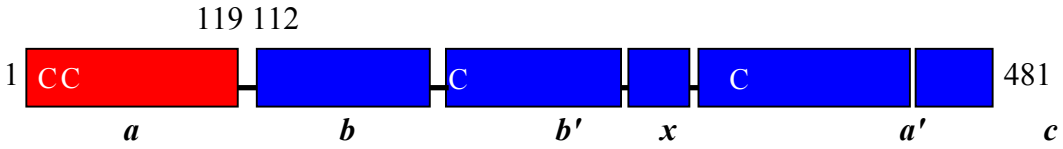
Amino Acid Sequence

```
  1 MRLRRLALFP GVALLLAAAR LAAASDVLEL TDDNFESRIS DTGSAGLMLV EFFAPWCGHC   60
 61 KRLAPEYEAA ATRLKGIVPL AKVDCTANTN TCNKYGVSGY PTLKIFRDGE EAGAYDGPRT  120
121 ADGIVSHLKK QAGPASVPLR TEEEFKKFIS DKDASIVGFF DDSFSEAHSE FLKAASNLRD  180
181 NYRFAHTNVE SLVNEYDDNG EGIILFRPSH LTNKFEDKTV AYTEQKMTSG KIKKFIQENI  240
241 FGICPHMTED NKDLIQGKDL LIAYYDVDYE KNAKGSNYWR NRVMMVAKKF LDAGHKLNFA  300
301 VASRKTFSHE LSDFGLESTA GEIPVVAIRT AKGEKFVMQE EFSRDGKALE RFLQDYFDGN  360
361 LKRYLKSEPI PESNDGPVKV LVGKNFEDVA FDEKKNVFVE FYAPWCGHCK QLAPIWDKLG  420
421 ETYKDHENIV IAKMDSTANE VEAVKVHSFP TLKFFPASAD RTVIDYNGER TLDGFKKFLE  480
481 SGGQDGAGDD DDLEDLEEAE EPDMEEDDDQ KAVKDEL
```

DNA Sequence

```
atgcgcctcc gccgcctagc gctgttcccg ggtgtggcgc tgcttcttgc cgcggcccgc    60
ctcgccgctg cctccgacgt gctagaactc acggacgaca acttcgagag tcgcatctcc   120
gacacgggct ctgcgggcct catgctcgtc gagttcttcg ccccctggtg tggacactgc   180
aagagacttg cacctgagta tgaagctgca gctaccagat taaaagggat agtcccatta   240
gcaaaggttg attgcactgc caacactaac acctgtaata aatatggagt cagtggatat   300
ccaaccctga agatatttag agatggtgaa gaagcaggtg cttatgatgg acctaggact   360
gctgatggaa ttgtcagcca cttgaagaag caggcaggac cagcttcagt gcctctcagg   420
actgaggaag aatttaagaa attcattagt gataaagatg cctctatagt aggttttttc   480
gatgattcat tcagtgaggc tcactccgag ttcctaaaag cagccagcaa cttgagggat   540
aactaccgat ttgcacatac gaatgttgag tctctggtga acgagtatga tgataatgga   600
gagggtatca tcttatttcg tccttcacat ctcactaaca agtttgagga caagactgtg   660
gcatatacag agcaaaaaat gaccagtggc aaaattaaaa agtttatcca ggaaaacatt   720
tttggtatct gccctcacat gacagaagac aataaagatt tgatacaggg caaggactta   780
cttattgctt actatgatgt ggactatgaa aagaacgcta aaggttccaa ctactggaga   840
aacagggtaa tgatggtggc aaagaaattc ctggatgctg ggcacaaact caactttgct   900
gtagctagcc gcaaaacctt tagccatgaa ctttctgatt ttggcttgga gagcactgct   960
ggagagattc ctgttgttgc tatcagaact gctaaaggag agaagtttgt catgcaggag  1020
gagttctcgc gtgatgggaa ggctctggag aggttcctgc aggattactt tgatggcaat  1080
ctgaagagat acctgaagtc tgaacctatc ccagagagca atgatcctgt caaggtgcct  1140
gttgggaaga actttgaaga cgtggctttt gatgagaaaa aaaacgtctt tgtggagttc  1200
ctatgcccca tggtgtggtc actgcaaaca gttggctccc atttgggata aactgggaga  1260
gacgtacaag gaccatgaga acatcgtcat cgccaagatg gactcgactg ccaacgaggt  1320
ggaggccgtc aaagtgcaca gcttccccac actcaagttc tttcctgcca gtgccgacag  1380
gacggtcatt gattacaacg gggaacgcac gctggatggt tttaagaaat tcctggagag  1440
cggtggccag gatggggcag gggatgatga cgatctcgag gacctggaag aagcagagag  1500
ccagacatgg aggaagacga tgatcagaaa gctgtgaaag atgaactgtaa
```

1.7 Construct 4 - PDI ***a*** ERp57 ***bb'a'c***



Contains 5 cysteines (approximate placement shown above in white and underlined below) and 52 lysines. ERp57 and PDI domains shown in blue and red respectively.

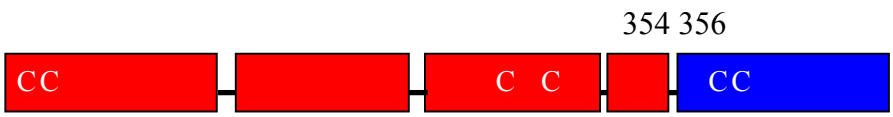
Amino Acid Sequence

```
  1 MRLRRLALFP GVALLLAAAR LAAASDVLEL TDDNFESRIS DTGSAGLMLV EFFAPWCGHC    60
 61 KRLAPEYEAA ATRLKGIVPL AKVDCTANTN TCNKYGVSGY PTLKIFRDGE EAGAYDGPRT   120
121 ADGIVSHLKK QAGPASVPLR TEEEFKKFIS DKDASIVGFF DDSFSEAHSE FLKAASNLRD   180
181 NYRFAHTNVE SLVNEYDDNG EGIILFRPSH LTNKFEDKTV AYTEQKMTSG KIKKFIQENI   240
241 FGICPHMTED NKDLIQGKDL LIAYYDVDYE KNAKGSNYWR NRVMMVAKKF LDAGHKLNFA   300
301 VASRKTFSHE LSDFGLESTA GEIPVVAIRT AKGEKFVMQE EFSRDGKALE RFLQDYFDGN   360
361 LKRYLKSEPI PESNDGPVKV VVAENFDEIV NNENKDVLIE FYAPWCGHCK NLEPKYKELG   420
421 EKLSKDPNIV IAKMDATAND VPSPYEVRGF PTIYFSPANK KLNPKKYEGG RELSDFISYL   480
481 QREATNPPVI QEEKPKKKKK AQEDL
```

DNA sequence

```
atgctgcgcc gcgctctgct gtgcctggcc gtggccgccc tggtgcgcgc cgacgccccc        60
gaggaggagg accacgtcct ggtgctgcgg aaaagcaact tcgcggaggc gctggcggcc       120
cacaagtacc tgctggtgga gttctatgcc ccttggtgtg gccactgcaa ggctctggcc       180
cctgagtatg ccaaagccgc tgggaagctg aaggcagaag gttccgagat caggttggcc       240
aaggtggacg ccacggagga gtctgacctg gcccagcagt acggcgtgcg cggctatccc       300
accatcaagt tcttcaggaa tggagacacg gcttccccca aggaatatac agctggcaga       360
gaggctgatg acatcgtgaa ctggctgaag aagcgcacgg gcccggcttc agtgcctctc       420
aggactgagg aagaatttaa gaaattcatt agtgataaag atgcctctat agtaggtttt       480
ttcgatgatt cattcagtga ggctcactcc gagttcctaa aagcagccag caacttgagg       540
gataactacc gatttgcaca tacgaatgtt gagtctctgg tgaacgagta tgatgataat       600
ggagagggta tcatcttatt tcgtccttca catctcacta acaagtttga ggacaagact       660
gtggcatata cagagcaaaa aatgaccagt ggcaaaatta aaaagtttat ccaggaaaac       720
atttttggta tctgccctca catgacagaa gacaataaag atttgataca gggcaaggac       780
ttacttattg cttactatga tgtggactat gaaaagaacg ctaaaggttc caactactgg       840
agaaacaggg taatgatggt ggcaaagaaa ttcctggatg ctgggcacaa actcaacttt       900
gctgtagcta gccgcaaaac ctttagccat gaactttctg attttggctt ggagagcact       960
gctggagaga ttcctgttgt tgctatcaga actgctaaag gagagaagtt tgtcatgcag      1020
gaggagttct cgcgtgatgg gaaggctctg gagaggttcc tgcaggatta ctttgatggc      1080
aatctgaaga gatacctgaa gtctgaacct atcccagaga gcaatgatgg gcctgtgaag      1140
gtagtggtag cagagaattt tgatgaaata gtgaataatg aaaataaaga tgtgctgatt      1200
gaattttatg ccccttggtg tggtcattgt aagaacctgg agcccaagta taaagaactt      1260
ggcgagaagc tcagcaaaga cccaaatatc gtcatagcca agatggatgc cacagccaat      1320
gatgtgcctt ctccatatga agtcagaggt tttcctacca tatacttctc tccagccaac      1380
aagaagctaa atccaaagaa atatgaaggt ggccgtgaat taagtgattt tattagctat      1440
ctacaaagag aagctacaaa ccccccctgta attcaagaag aaaaacccaa gaagaagaag      1500
aaggcacagg aggatctctaa
```

1.8 Construct 5 - PDI ***abb'x*** ERp57 ***a'***



1 465

a ***b*** ***b′*** ***x*** ***a′***

Contains 6 cysteines (approximate placement shown above in white and underlined below) and 45 lysines. ERp57 and PDI domains shown in blue and red respectively.

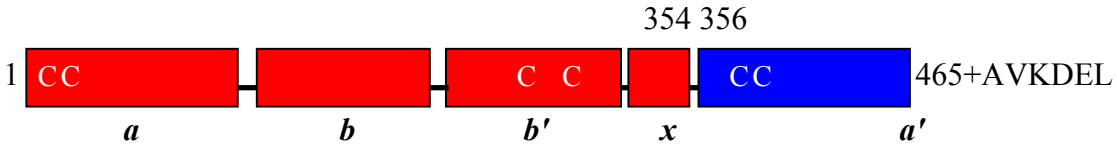
Amino Acid Sequence

```
  1 MLRRALLCLA VAALVRADAP EEEDHVLVLR KSNFAEALAA HKYLLVEFYA PWCGHCKALA   60
 61 PEYAKAAGKL KAEGSEIRLA KVDATEESDL AQQYGVRGYP TIKFFRNGDT ASPKEYTAGR  120
121 EADDIVNWLK KRTGPAATTL PDGAAAESLV ESSEVAVIGF FKDVESDSAK QFLQAAEAID  180
181 DIPFGITSNS DVFSKYQLDK DGVVLFKKFD EGRNNFEGEV TKENLLDFIK HNQLPLVIEF  240
241 TEQTAPKIFG GEIKTHILLF LPKSVSDYDG KLSNFKTAAE SFKGKILFIF IDSDHTDNQR  300
301 ILEFFGLKKE ECPAVRLITL EEEMTKYKPE SEELTAERIT EFCHRFLEGK IKPHLMSQEL  360
361 PEDWDKQPVK VVVAENFDEI VNNENKDVLI EFYAPWCGHC KNLEPKYKEL GEKLSKDPNI  420
421 VIAKMDATAN DVPSPYEVRG FPTIYFSPAN KKLNPKKYEG GRELSDFISY LQREATNPPV  480
```

DNA sequence

```
atgctgcgcc gcgctctgct gtgcctggcc gtggccgccc tggtgcgcgc cgacgccccc   60
gaggaggagg accacgtcct ggtgctgcgg aaaagcaact tcgcggaggc gctggcggcc  120
cacaagtacc tgctggtgga gttctatgcc ccttggtgtg gccactgcaa ggctctggcc  180
cctgagtatg ccaaagccgc tgggaagctg aaggcagaag gttccgagat caggttggcc  240
aaggtggacg ccacggagga gtctgacctg gcccagcagt acggcgtgcg cggctatccc  300
accatcaagt tcttcaggaa tggagacacg gcttccccca aggaatatac agctggcaga  360
gaggctgatg acatcgtgaa ctggctgaag aagcgcacgg gcccggctgc caccaccctc  420
cgtgacggcg cagctgcaga gtccttggtg gagtccagcg aggtggctgt catcggcttc  480
ttcaaggacg tggagtcgga ctctgccaag cagtttttgc aggcagcaga ggccatcgat  540
gacataccat ttgggatcac ttccaacagt gacgtgttct ccaaatacca gctcgacaaa  600
gatggggttg tcctctttaa gaagtttgat gaaggccgga acaactttga aggggaggtc  660
accaaggaga acctgctgga ctttatcaaa cacaaccagc tgccccttgt catcgagttc  720
accgagcaga cagccccgaa gatttttgga ggtgaaatca agactcacat cctgctgttc  780
ttgcccaaga gtgtgtctga ctatgacggc aaactgagca acttcaaaac agcagccgag  840
agcttcaagg gcaagatcct gttcatcttc atcgacagcg accacaccga caaccagcgc  900
atcctcgagt tctttggcct gaagaaggaa gagtgcccgg ccgtgcgcct catcaccctg  960
gaggaggaga tgaccaagta caagcccgaa tcggaggagc tgacggcaga gaggatcaca 1020
gagttctgcc accgcttcct ggagggcaaa atcaagcccc acctgatgag ccaggagcgt 1080
gccggagact gggacaagca gcctgtcaag gtagtggtag cagagaattt tgatgaaata 1140
gtgaataatg aaaataaaga tgtgctgatt gaattttatg ccccttggtg tggtcattgt 1200
aagaacctgg agcccaagta taaagaactt ggcgagaagc tcagcaaaga cccaaatatc 1260
gtcatagcca agatggatgc cacagccaat gatgtgcctt ctccatatga agtcagaggt 1320
ttcctacca tatacttctc tccagccaac aagaagctaa atccaaagaa atatgaaggt 1380
ggccgtgaat taagtgattt tattagctat ctacaaagag aagctacaaa ccccctgta  1440
```

1.9 Construct 6 - PDI ***abb'x*** ERp57 ***a'*** +AVKDEL



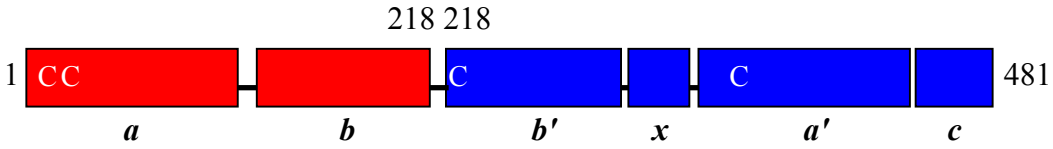
Contains 6 cysteines (approximate placement shown above in white and underlined below) and 46 lysines. ERp57 and PDI domains shown in blue and red respectively.

Amino Acid Sequence

```
   1 MLRRALLCLA VAALVRADAP EEEDHVLVLR KSNFAEALAA HKYLLVEFYA PWCGHCKALA   60
  61 PEYAKAAGKL KAEGSEIRLA KVDATEESDL AQQYGVRGYP TIKFFRNGDT ASPKEYTAGR  120
 121 EADDIVNWLK KRTGPAATTL PDGAAAESLV ESSEVAVIGF FKDVESDSAK QFLQAAEAID  180
 181 DIPFGITSNS DVFSKYQLDK DGVVLFKKFD EGRNNFEGEV TKENLLDFIK HNQLPLVIEF  240
 241 TEQTAPKIFG GEIKTHILLF LPKSVSDYDG KLSNFKTAAE SFKGKILFIF IDSDHTDNQR  300
 301 ILEFFGLKKE ECPAVRLITL EEEMTKYKPE SEELTAERIT EFCHRFLEGK IKPHLMSQEL  360
 361 PEDWDKQPVK VVVAENFDEI VNNENKDVLI EFYAPWCGHC KNLEPKYKEL GEKLSKDPNI  420
 421 VIAKMDATAN DVPSPYEVRG FPTIYFSPAN KKLNPKKYEG GRELSDFISY LQREATNPPV  480
..481 AVKDEL
```

As Construct 5 (PDI ***abb'x*** ERp57 ***a'***), but with additional PDI retention sequence (AVKDEL).

1.10 Construct 7 - PDI ***ab*** ERp57 ***b'a'c***



Contains 5 cysteines (approximate placement shown above in white and underlined below) and 51 lysines. ERp57 and PDI domains shown in blue and red respectively.

Amino Acid Sequence

```
  1 MLRRALLCLA VAALVRADAP EEEDHVLVLR KSNFAEALAA HKYLLVEFYA PWCGHCKALA   60
 61 PEYAKAAGKL KAEGSEIRLA KVDATEESDL AQQYGVRGYP TIKFFRNGDT ASPKEYTAGR  120
121 EADDIVNWLK KRTGPAATTL PDGAAAESLV ESSEVAVIGF FKDVESDSAK QFLQAAEAID  180
181 DIPFGITSNS DVFSKYQLDK DGVVLFKKFD EGRNNFEGEV TKENLLDFIK HNQLPGICPH  240
241 MTEDNKDLIQ GKDLLIAYYD VDYEKNAKGS NYWRNRVMMV AKKFLDAGHK LNFAVASRKT  300
301 FSHELSDFGL ESTAGEIPVV AIRTAKGEKF VMQEEFSRDG KALERFLQDY FDGNLKRYLK  360
361 SEPIPESNDG PVKVVVAENF DEIVNNENKD VLIEFYAPWC GHCKNLEPKY KELGEKLSKD  420
421 PNIVIAKMDA TANDVPSPYE VRGFPTIYFS PANKKLNPKK YEGGRELSDF ISYLQREATN  480
481 PPVIQEEKPK KKKKAQEDL
```

DNA sequence

```
atgctgcgcc gcgctctgct gtgcctggcc gtggccgccc tggtgcgcgc cgacgccccc   60
gaggaggagg accacgtcct ggtgctgcgg aaaagcaact tcgcggaggc gctggcggcc  120
cacaagtacc tgctggtgga gttctatgcc ccttggtgtg gccactgcaa ggctctggcc  180
cctgagtatg ccaaagccgc tgggaagctg aaggcagaag gttccgagat caggttggcc  240
aaggtggacg ccacggagga gtctgacctg gcccagcagt acggcgtgcg cggctatccc  300
accatcaagt tcttcaggaa tggagacacg gcttccccca aggaatatac agctggcaga  360
gaggctgatg acatcgtgaa ctggctgaag aagcgcacgg gcccggctgc caccaccctc  420
cgtgacggcg cagctgcaga gtccttggtg gagtccagcg aggtggctgt catcggcttc  480
ttcaaggacg tggagtcgga ctctgccaag cagtttttgc aggcagcaga ggccatcgat  540
gacataccat ttgggatcac ttccaacagt gacgtgttct ccaaatacca gctcgacaaa  600
gatggggttg tcctctttaa gaagtttgat gaaggccgga acaactttga aggggaggtc  660
accaaggaga acctgctgga ctttatcaaa cacaaccagc tgccccttgg tatctgccct  780
cacatgacag aagacaataa agatttgata cagggcaagg acttacttat tgcttactat  840
gatgtggact atgaaaagaa cgctaaaggt tccaactact ggagaaacag ggtaatgatg  900
gtggcaaaga aattcctgga tgctgggcac aaactcaact ttgctgtagc tagccgcaaa  960
acctttagcc atgaactttc tgattttggc ttggagagca ctgctggaga gattcctgtt 1020
gttgctatca gaactgctaa aggagagaag tttgtcatgc aggaggagtt ctcgcgtgat 1080
gggaaggctc tggagaggtt cctgcaggat tactttgatg gcaatctgaa gagatacctg 1140
aagtctgaac ctatcccaga gagcaatgat gggcctgtga aggtagtggt agcagagaat 1200
tttgatgaaa tagtgaataa tgaaaataaa gatgtgctga ttgaatttta tgccccttgg 1260
tgtggtcatt gtaagaacct ggagcccaag tataaagaac ttggcgagaa gctcagcaaa 1320
gacccaaata tcgtcatagc caagatggat gccacagcca atgatgtgcc ttctccatat 1380
gaagtcagag gttttcctac catatacttc tctccagcca acaagaagct aaatccaaag 1440
aaatatgaag gtggccgtga attaagtgat tttattagct atctacaaag agaagctaca 1500
aacccccctg taattcaaga agaaaaaccc aagaagaaga agaaggcaca ggaggatctc 1560
taa
```

1.11 Calreticulin

Amino Acid Sequence

MLLSVPLLLG LLGLAVAEPA VYFKEQFLDG DGWTSRWIES KHKSDFGKFV LSSGKFYGDE 60
EKDKGLQTSQ DARFYALSAS FEPFSNKGQT LVVQFTVKHE QNIDCGGGYV KLFPNSLDQT 120
DMHGDSEYNI MFGPDICGPG TKKVHVIFNY KGKNVLINKD IRCKDDEFTH LYTLIVRPDN 180
TYEVKIDNSQ VESGSLEDDW DFLPPKKIKD PDASKPEDWD ERAKIDDPTD SKPEDWDKPE 240
HIPDPDAKKP EDWDEEMDGE WEPPVIQNPE YKGEWKPRQI DNPDYKGTWI HPEIDNPEYS 300
HIPDPDAKKP EDWDEEMDGE WEPPVIQNPE YKGEWKPRQI DNPDYKGTWI HPEIDNPEYS 360
QDEEQRLKEE EEDKKRKEEE EAEDKEDDED KDEDEEDEED KEEDEEEDVP GQAKDEL 417

N domain, P domain and C domain are shown in blue, red and green respectively. The three cysteine residues are marked in bold, italic and underlined.

1.12 Calnexin

Amino Acid Sequence

MEGKWLLCML LVLGTAIVEA HDGHDDDVID IEDDLDDVIE EVEDSKPDTT APPSSPKVTY 60
KAPVPTGEVY FADSFDRGTL SGWILSKAKK DDTDDEIAKY DGKWEVEEMK ESKLPGDKGL 120
VLMSRAKHHA ISAKLNKPFL FDTKPLIVQY EVNFQNGIEC GGAYVKLLSK TPELNLDQFH 180
DKTPYTIMFG PDKCGEDYKL HFIFRHKNPK TGIYEEKHAK RPDADLKTYF TDKKTHLYTL 240
ILNPDNSFEI LVDQSVVNSG NLLNDMTPPV NPSREIEDPE DRKPEDWDER PKIPDPEAVK 300
PDDWDEDAPA KIPDEEATKP EGWLDDEPEY VPDPDAEKPE DWDEDMDGEW EAPQIANPRC 360
ESAPGCGVWQ RPVIDNPNYK GKWKPPMIDN PSYQGIWKPR KIPNPDFFED LEPFRMTPFS 420
AIGLELWSMT SDIFFDNFII CADRRIVDDW ANDGWGLKKA ADGAAEPGVV GQMIEAAEER 480
PWLWVVYILT VALPVFLVIL FCCSGKKQTS GMEYKKTDAP QPDVKEEEEE KEEEKDKGDE 540
EEEGEEKLEE KQKSDAEEDG GTVSQEEEDR KPKAEEDEIL NRSPRNRKPR RE 592

The potential lumenal domain is marked in red with the potential transmembrane and cytoplasmic domains marked in purple and blue respectively. The 6 cysteine residues in the mature protein are marked in bold, italic and underlined.

1.13 Pull down data

1.13.1 Binding of ERp57, ERp57∆ss and PDI to immobilised biotin-calreticulin (Figure 5.1)

Construct	Beads alone (%)	Beads + biotin-insulin (%)	Beads + biotin-CRT (%)	Beads + biotin-CRT + o/n native RNase A (%)	Beads + biotin-CRT + o/n scrambled RNase A (%)	Beads + biotin-CRT + Xs his-CRT (%)	Beads + biotin-CRT + Xs his-PDI (%)
ERp57	2.1	2	32	37	28	37	28
	3.3	4.4	35	40	32	28	30
ERp57∆ss	2.1	4.7	30	-	-	-	-
	2.7	4.2	32	-	-	-	-
PDI	1.8	3.2	1.2	-	-	-	-
	2	1.2	3	-	-	-	-

1.13.2 Binding of ERp57, ERp57∆c, ERp57/PDI chimeric and ERp57 sub-domain polypeptides to immobilised biotin-calreticulin (Figure 5.6)

Construct	Beads alone	Beads + biotin-insulin	Beads + biotin-CRT
ERp57	3.6	7.7	39.5
	4.9	7.2	34.5
	6	4.3	38.2
ERp57∆c	2.4	5.2	21.3
	4.3	7.6	18.6
	4.4	5	17.9
Construct 2	1.8	2.2	2.2
Construct 3	2.8	2.8	22.8
	1.5	1.5	19.4
	1.8	1.8	27.5
	2.6	2.6	23.2
Construct 4	3.3	3.3	47.2
	1.4	1.4	32.2
	2.3	2.3	42
	2.4	2.4	42.5
Construct 5	1.2	1.2	1.7
Construct 6	2.9	2.9	2.9
	2.5	2.5	2.9
Construct 7	2.9	3.5	4
	2.5	2.9	2
ERp57 *abb'*	2.3	1.5	0.3
	4.2	1.2	0.9
ERp57 *b'a'c*	2.1	2.3	4.1
	5	4.2	4.8
ERp57 *bb'x*	6.7	2.1	3
	5	5	3.5
ERp57 *b'x*	5.2	6.7	4.9
	3	5	3.5
PDI	1.5	5.2	6
	1.2	3	6

1.13.2 Binding of wild-type ERp57and ERp57 *b'* domain point mutants to immobilised biotin-calreticulin (Figure 5.11)

Construct	Beads alone	Beads + biotin-insulin	Beads + biotin-CRT
ERp57	2.3	3.2	26.5
	2.6	8.8	44.8
	2.25	2.1	39.2
pm1-V267A	6	5.9	12.7
	5.65	4.5	11
	6.62	7.55	13.2
pm2-V267W	9.5	8.2	11.85
	5.65	4.3	8.8
	8.9	13	13
pm3-Y269A	7	7.8	10
	3.2	2.4	7.6
	8.6	6.2	13.8
pm4-Y269W	5	7.6	14.55
	3.3	6.7	16
	3.3	4.5	13.5
pm5-E270A	8	8.1	41.7
	3.2	4.3	28.9
	6.2	8	35.9
pm6-K274A	3.1	3.3	7.15
	2.9	4.9	7.6
	5.5	6	8.5
pm7-R280A	5.15	6.2	10.3
	6.4	4.5	9.3
	5.5	9.55	9.1
pm8-N281A	3.2	3.7	4
	7.1	4.45	7.2
	6.05	6.8	6
pm9-V283A	3.7	2.6	20
	7.8	10.3	22
	4.2	4.9	17.5
pm10-V283L	6.3	6.3	12.15
	7.1	5.6	12.4
	5.75	8.7	12.7
pm11-V283W	6.4	7.2	8.7
	4.4	5.15	6.35
	4.4	5.95	5.3
pm12-F299A	6.5	8.2	8.7
	3.1	2.6	7.7
	3.45	6.2	10.5
pm13-F299W	4.8	5.6	3.8
	5.5	4	7
	5.2	6.7	7.1