ER quality control: towards an understanding at the molecular level

Lars Ellgaard* and Ari Helenius†

The process of 'quality control' in the endoplasmic reticulum (ER) involves a variety of mechanisms that collectively ensure that only correctly folded, assembled and modified proteins are transported along the secretory pathway. In contrast, non-native proteins are retained and eventually targeted for degradation. Recent work provides the first structural insights into the process of glycoprotein folding in the ER involving the lectin chaperones calnexin and calreticulin. Underlying principles governing the choice of chaperone system engaged by different proteins have also been discovered.

Addresses
Institute of Biochemistry, ETH Zürich, Universitätstrasse 16, CH - 8092 Zürich, Switzerland
*e-mail: lars.ellgaard@bc.biol.ethz.ch
†e-mail: ari.helenius@bc.biol.ethz.ch

Current Opinion in Cell Biology 2001, 13:431–437

0955-0674/01/$ – see front matter
© 2001 Elsevier Science Ltd. All rights reserved.

Abbreviations

CRT	calreticulin
CNX	calnexin
ER	endoplasmic reticulum
ERAD	ER-associated degradation
PDI	protein disulfide isomerase
3D	three-dimensional
UPR	unfolded protein response
UGGT	UDP-glucose:glycoprotein glucosyltransferase

Introduction

The endoplasmic reticulum (ER) plays an essential role in the folding and maturation of newly synthesized proteins in the secretory pathway. It provides an environment optimized for folding, oxidation and oligomeric assembly of proteins translocated into the lumen or inserted into the membrane. Folding in the ER is assisted by a large variety of folding enzymes, molecular chaperones and folding sensors [1]. Many of these associate with growing nascent chains and continue to assist folding until a protein has acquired its native structure. To ensure the fidelity of the maturation process, exit from the ER is regulated by a stringent quality control system that inhibits the secretion of incompletely folded or misfolded proteins [2]. In addition to securing extended exposure of proteins to the folding machinery, quality control prevents deployment of potentially malfunctioning proteins that could be detrimental to the cell and the organism. For many proteins in the ER, proper folding and maturation depends on co- and post-translational modifications. Here we primarily focus on the ER quality control system that is in place for proteins containing N-linked glycans.

Quality control mechanisms

ER quality control operates at several levels and by multiple mechanisms [3]. At a general level, all proteins are subject to conformation-based screening by members of major molecular chaperone families. These chaperones have the capacity to recognize properties common to non-native proteins such as exposed hydrophobic areas. They selectively associate with proteins that display such features and in doing so promote folding and assembly. As long as they are engaged in interactions with the substrate proteins they also prevent export from the ER.

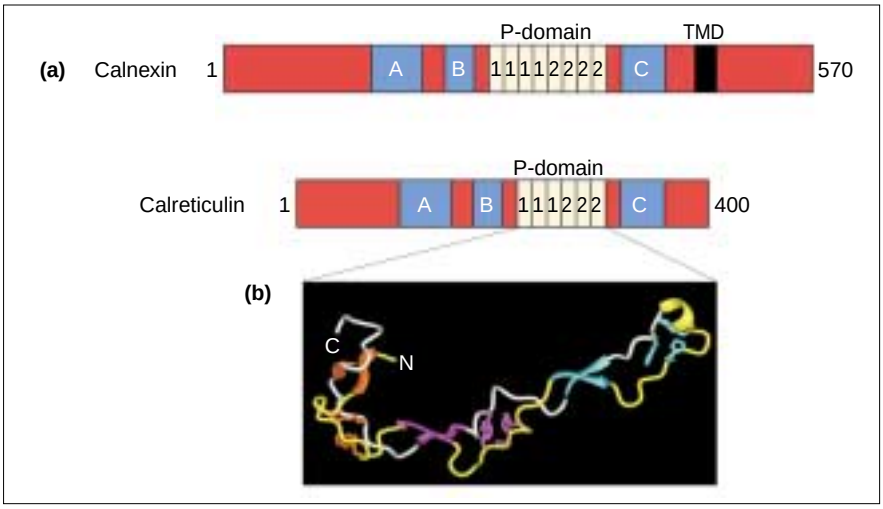
Another level of quality control is specific for individual protein species or protein families [3–5]. This level includes specialized folding factors such as HSP47, which only interacts with collagens [6]. It also includes special escort proteins, such as the receptor associated protein (RAP), which accompanies LDL (low density lipoprotein) receptor family members from the ER to the Golgi complex [7]. Moreover, there are also retention factors, such as TAP (transporter associated with antigen processing) and tapasin for MHC Class I antigens, that restrict transport to a limited set of conformers of specific substrate proteins [8].

The quality control process involves a complex sorting system that separates proteins according to their folding and maturation status. The folding and assembly process is thus functionally coupled to export by vesicular transport. For certain proteins, transport relies on the exposure of specific signal sequences. Such signals, like the DXE sequences in the cytosolic domains of many membrane proteins [9–11], may accelerate transport or, like RXR sequences in ion channels, they may prevent premature transport [12,13]. Selective retrieval of misfolded proteins bound to chaperones from the intermediate compartment or the *cis*-Golgi by retrograde transport to the ER has also been reported [14].

Whereas folded proteins rapidly move via ER exit sites and the intermediate compartment to the *cis*-Golgi and beyond, persistently misfolded or unassembled proteins either aggregate or become degraded. Degradation is important because the folding process is far from quantitative even under normal cellular growth conditions. In most cases, ER-associated degradation (ERAD) of misfolded proteins involves their retrotranslocation to the cytosol, ubiquitination and degradation by proteasomes [15–17]. Not surprisingly, a number of diseases such as cystic fibrosis, α_1-antitrypsin deficiency and familial hypercholesterolaemia are associated with ER retention and degradation of folding-defective mutant proteins [4,18].

It is important to note that the accumulation of misfolded proteins in the ER, especially observed under conditions of stress, triggers activation of a wide range of genes encoding for proteins of the secretory pathway. This is the so-called unfolded protein response (UPR) [19–21].

Figure 1



Schematic representation of calnexin (CNX) and calreticulin (CRT). **(a)** CNX is a type 1 membrane protein of 570 residues, whereas CRT is a soluble, lumenal protein of 400 residues. The transmembrane domain (TMD) of CNX is depicted in black. Regions A, B and C show 50–55% sequence identity [63]. The central P-domain contains two sequence repeat types, designated 1 and 2, each repeated four times in CNX and three times in CRT. **(b)** A cartoon of the CRT P-domain NMR structure is shown (reproduced with permission from [36••]). 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.

Several of the UPR-induced proteins are involved in protein folding and glycosylation in the ER, in ERAD, in lipid metabolism and in vesicular transport [22••]. Thus, quality control, ERAD and UPR are tightly coordinated processes (recently reviewed in [23]).

The calnexin/calreticulin cycle

One of the most common modifications of proteins translocated into the ER is the addition of N-linked glycans. For glycoproteins, a particularly well studied ER quality control system is in place, involving two homologous lectins, calnexin (CNX) and calreticulin (CRT) (Figure 1). The process of N-linked glycosylation occurs through the transfer of a triglucosylated, branched core oligosaccharide ($Glc_3Man_9GlcNAc_2$) to the nascent polypeptide chain as it enters the ER lumen. Soon after transfer, trimming of the core oligosaccharide by the successive action of ER glucosidases I and II to the monoglucosylated form, $Glc_1Man_9GlcNAc_2$, allows the glycopolypeptide to interact with CRT and CNX. Being lectins they specifically interact with glycoproteins, but only if these have monoglucosylated N-glycans [24]. Interaction with CNX and CRT exposes the folding glycoprotein to the associated co-chaperone, ERp57, a thiol oxidoreductase of the protein disulfide isomerase (PDI) family [25].

The association between substrate glycoprotein and lectin is terminated by glucosidase II, which removes the remaining glucose from the glycan. If, at this point, the glycoprotein has reached its native conformation, it is no longer retained in the ER and can be transported to the Golgi complex. If not, re-addition of a glucose to the N-linked glycan occurs by the action of the UDP-glucose:glycoprotein glucosyltransferase (UGGT), a lumenal enzyme that acts as a folding sensor [26•]. The glycoprotein is thereby 'tagged' for renewed interaction with CRT and CNX. The possibility of multiple rounds of interaction with CRT and CNX is thus ensured. Overall, CRT and CNX together with glucosidase II and UGGT cooperate to increase the folding efficiency, to prevent premature oligomeric assembly and to prevent the export of misfolded glycoproteins from the ER. The main features of the CNX/CRT cycle are now well established, and a thorough review has recently been published [26•]. Below we address some detailed aspects of the cycle and discuss unresolved questions.

Chaperone selection in the endoplasmic reticulum

In addition to CNX and CRT, the ER contains a large collection of other molecular chaperones and folding factors with different properties and functions [27]. Each newly synthesized protein makes use of only a few of the available chaperones. What are the parameters that determine which chaperones a protein will engage and in which order?

Recent work shows that for glycoproteins the choice of chaperone depends, in part, on the position of the glycans in the sequence [28•]. Growing nascent chains that have N-linked glycans within the first ~50 residues from the amino terminus preferentially interact with CRT and CNX. In contrast, glycoproteins in which the glycans occur later in the sequence first interact with BiP, an abundant ER chaperone of the Hsp70 family that binds to hydrophobic peptide sequences, and later during post-translational folding with CRT and CNX.

With respect to CRT and CNX, it is clear that despite their sequence similarity and identical oligosaccharide specificity *in vitro* [29–31], these two lectins exhibit only partially overlapping substrate specificities *in vivo* [32–34]. Moreover, when they bind to the same protein, they have in some cases been shown to interact with different glycans [35]. At the molecular level, the main differences are

that the extended P-domain arm in CRT (Figure 1) is shorter than the corresponding arm in CNX [36••] and that CRT is a soluble, lumenal protein, whereas CNX is a type I membrane protein of the ER membrane. As previously suggested [32,33,35], the latter difference clearly plays a role in substrate selection. Analyzing the substrates bound to a soluble, anchor-free mutant of CNX and a membrane-bound version of CRT, Danilczyk and coworkers [37] recently observed that the substrate specificities of the two chaperones were essentially inverted.

ERp57-catalyzed disulfide bond formation

The oxidizing environment and the presence of several different thiol oxidoreductases allow formation of disulfide bonds in the ER. One of the oxidoreductases, ERp57, functions as a 'co-chaperone' with CRT and CNX [25,38•]. It most probably forms one-to-one complexes with both CRT and CNX [38•] and has been shown to accelerate oxidative refolding of monoglucosylated RNaseB in the presence of CRT or CNX *in vitro* [39••]. The formation of transient intermolecular disulfide bonds between ERp57 and newly synthesized glycoproteins have, moreover, been observed in living cells [40••]. When the association of CRT and CNX with glycoproteins was blocked, formation of mixed disulfides with ERp57 was also inhibited, indicating that lectin binding is a prerequisite for substrate recognition by this thiol oxidoreductase.

Substrate recognition by UDP-glucose: glycoprotein glucosyltransferase

UGGT, the folding sensor in the CNX/CRT cycle, is a large, soluble, lumenal enzyme [26•]. Its catalytic domain displays a conserved 300 amino acid sequence at the carboxyl terminus of the protein with homology to glycosyltransferases of family 8 [41]. UDP-glucose, transported into the ER lumen from the cytosol [42], is the glucose donor, whereas the acceptors are glucose-free high mannose oligosaccharides attached to incompletely folded glycoproteins. UGGT is present throughout the ER including the transitional ER elements [43].

The exact mechanism by which UGGT distinguishes folded from non-native glycoproteins is not known. Given the large variety of unrelated glycoproteins that serve as substrates, it is likely that the enzyme resembles classical molecular chaperones in that it recognizes features shared by incompletely folded proteins. The enzyme is specific for glycoproteins as it uses neither glycans nor short glycopeptides as substrates [44]. Misfolded non-glycosylated proteins do not inhibit UGGT [44,45]; however, a misfolded glycoprotein containing only the innermost GlcNAc unit of the oligosaccharide can inhibit it [46]. Furthermore, although UGGT does not recognize glycoproteins in a random coil conformation, it efficiently reglucosylates a variety of partially folded conformers [47]. This is consistent with the finding that its function in cells coincides with later stages of folding [48]. In glycoproteins with multiple domains, UGGT selectively recognizes glycans in the misfolded domains [49].

Three main models can be proposed to explain how UGGT recognizes its substrates. First, exposed hydrophobic peptide elements in the glycoprotein substrate could be recognized by the enzyme. It has been shown that UGGT binds to immobilized hydrophobic peptides and that this interaction can be inhibited by denatured glycoproteins [46]. Second, recognition could involve the innermost GlcNAc unit of the oligosaccharide. In folded proteins this sugar interacts with neighboring amino acid residues, an interaction which may be lost upon denaturation [26•]. Finally, the enzyme may recognize the dynamic properties of the polypeptide moiety. In other words, it may be sensitive to the mobility or deformability of the protein to which the glycan is connected.

UGGT studies have been hampered by the lack of recombinant enzyme, the tendency of substrates to aggregate and the heterogeneity of substrate glycoforms. With the recent expression of UGGT in insect cells and the use of non-aggregating, homogenous substrates such as glycopeptides, yeast acid phosphatase and RNase B, more rigorous analysis of this interesting and important enzyme should be possible.

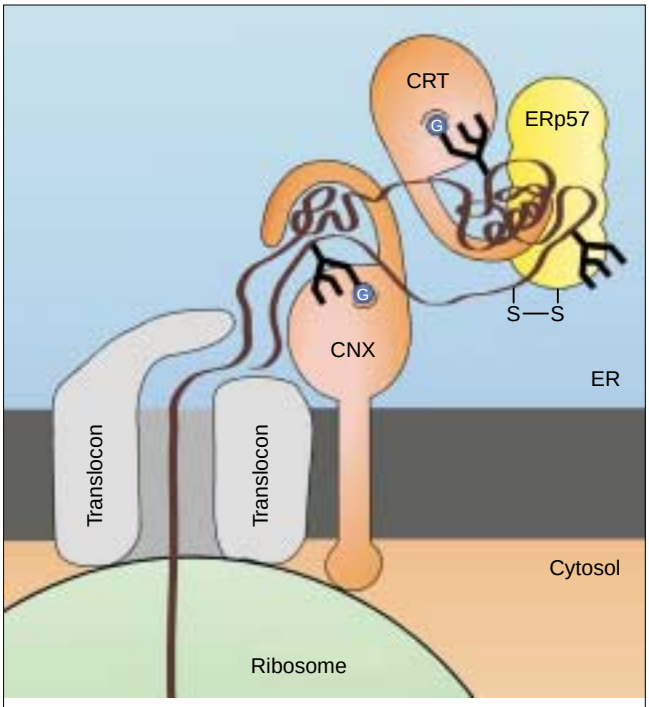
Three dimensional structures of calreticulin and calnexin

For several years, 3D structure determination of CRT and CNX has been pursued by several groups. Now, as a first step towards a more detailed understanding of their function at the molecular level, the NMR structure of the CRT P-domain has been solved [36••,50]. In addition, the crystal structure of a CNX ectodomain fragment, for which crystallization conditions have previously been reported [51], has recently been solved but not yet published [23].

The NMR structure of the CRT P-domain (residues 189–288) shows an extended hairpin fold comprising the entire polypeptide chain with amino and carboxyl termini in close spacial proximity ([36••]; Figure 1). This unusual structure constitutes a new fold. It is stabilized by three short anti-parallel β-sheets as well as by three small hydrophobic clusters each involving two highly conserved tryptophyl residues, one from each strand of the hairpin. The three-fold repetition of both the β-sheets and the hydrophobic clusters reflects the repetitive nature of the P-domain sequence, which contains two sets of amino acid sequences each repeated three times (Figure 1). The topology and the elongated shape of the P-domain suggest that it constitutes an extended, somewhat curved protrusion from the CRT core domain [36••]. This is in agreement with the recent finding by gel filtration and sedimentation analysis that full-length CRT is an elongated molecule [52•].

A brief mention of the unpublished ectodomain crystal structure by Chevet and coworkers [23] describes the CNX ectodomain as containing 'a lectin domain, as well

Figure 2



A model of CNX, CRT and ERp57 interacting with a growing nascent chain (brown) of a glycoprotein in the ER. The growing nascent chain of the glycoprotein is cotranslationally translocated into the lumen of the ER through the translocon complex which contains a large, lumenal protrusion [64]. Phosphorylation of the cytosolic tail of CNX leads to increased association with the ribosome [65]. CNX, CRT and ERp57 are known to interact co- and post-translationally with glycoprotein chains in the ER of live cells. Early association with CNX and CRT is possible because, after addition of the core glycans by the oligosaccharyl tranferase enzyme, two of the glucoses are rapidly trimmed to generate the monoglucosylated form of the glycans (the glucose is represented by the blue circles labeled 'G'). The glycans bind to the lectin domains of CNX and CRT. The P-domain, which forms a long, slightly curved arm extending from the lectin domain, is likely to generate a partially closed space within which folding of the glycopolypeptide can occur in a protected environment. It is also possible, as shown here for CRT, that the P-domains of CNX and CRT interact with ERp57 allowing this thiol oxidoreductase to interact optimally with cysteines in the glycoprotein substrate and thereby promote proper formation of disulfide bonds.

as a distinct loop', indicating a structure similar to CRT. Taken together, a uniform picture of the 3D structure of CRT and CNX is emerging, where the P-domain constitutes a finger-like extension from the globular core structure that is responsible for the interaction with the oligosaccharide of the substrate glycoprotein. That the P-domain does not contain the carbohydrate binding site is consistent with the lack of structural homology with known lectins and the experimental observation that the P-domain alone does not bind to glycoproteins [53].

At present, we can only speculate about the structure of the lectin domain. However, outside the P-domain, CNX and CRT show weak sequence similarity to legume lectins such as pea lectin, which in turn shows structural homology to galectins and pentraxins [54]. These proteins are characterized by a β-sandwich structure containing two opposing β-sheets each of six or seven β-strands. Secondary structure prediction of CRT and CNX using the PHDsec algorithm [55] shows that both proteins are, in fact, likely to contain 10 or 11 β-strands outside the P-domain region. Therefore, the possibility exists that the CRT/CNX lectin domain is characterized by a similar fold to legume lectins. However, in contrast to the well characterized plant lectins, which are multivalent, CRT binding to monoglucosylated IgG is monovalent [56•]. Interestingly, ERGIC-53 and VIP36, two mannose specific lectins, shuttling between the ER and the Golgi complex also both show sequence similarity to the leguminous lectins [57,58].

Structural insights into glycoprotein folding

With the oligosaccharide binding function of CRT and CNX mapped to a distinct lectin domain, the functions of the P-domain become all the more intriguing. *A priori*, the P-domain could be a site for direct interaction with unfolded proteins. Two recent papers describe *in vitro* experiments that suggest such a function for CRT and CNX [59•,60]. Both proteins were found to bind to unfolded, non-glycosylated proteins but not to native conformers. In addition, they suppressed thermal denaturation and aggregation and kept substrates in a folding-competent state. Generally, interactions between classical molecular chaperones such as those belonging to the Hsp60, Hsp70 and Hsp90 families and their non-native substrates are mediated by hydrophobic contacts. However, the highly charged surface of the CRT P-domain [36••] provides no obvious sites for protein–protein interactions of this sort. Whether the results described in [59•,60] affect protein folding *in vivo* remains to be seen.

Although a role for the P-domain in binding to unfolded protein cannot be ruled out, it seems likely that it participates in other protein–protein interactions. The topology of the P-domain places the tip of the hairpin loop at a discrete distance from the lectin domain. Likewise, the protein moiety of a bound glycoprotein substrate would be placed at a distance from the lectin domain due to the presence of the glycan. Thus, the tip of the hairpin loop could constitute a protein-binding site, with the most obvious ligand being the co-chaperone ERp57. The NMR data show indications of slow conformational exchange in the central region of the P-domain, suggesting a certain degree of plasticity [36••]. This plasticity could endow the bound ERp57 with some freedom of movement that would allow it to adopt different positions in respect to the glycoprotein substrates bound to the glycan-binding site and thus allow access to cysteines at different positions. The intrinsic flexibility of CRT was recently deduced from its hydrodynamic properties and was suggested to be of potential importance for the protein's function as a molecular chaperone [52•].

Although speculative at present, an additional function of the P-domain could be to act as a 'diffusion barrier' for glycoproteins upon dissociation from the glycan-binding site. As pointed out recently, the relatively low affinity of CRT for monoglucosylated IgG of ~10^5 M^{-1} implies rapid rounds of association and dissociation [56•]. A steric constraint on the diffusion of the glycoprotein away from the chaperone would ensure a higher local glycan concentration in the vicinity of the lectin domain and thereby facilitate renewed association. A similar effect would be achieved with multiple glycan chains, which are known to stabilize the complexes [61]. These ideas have been incorporated into the model shown in Figure 2 where CNX and CRT bind to the N-linked glycans of the growing nascent chain through the lectin domain. The P-domain of CRT, containing a bound molecule of ERp57, wraps around the chain to create a protective barrier around the folding chain and to position ERp57 for disulfide oxidation. As already mentioned, CNX and CRT are known to interact concurrently with influenza hemagglutinin and other proteins cotranslationally and binding occurs to distinct sugars [35,62].

Conclusions

The quality control mechanisms in the ER ensure the structural integrity of proteins delivered to the organelles of the secretory and endocytic pathways and the extracellular space. Recent progress has provided a better understanding of oxidative folding of glycoproteins through the cooperation of ERp57 and lectin chaperones CRT or CNX, of the basis for recognition of unfolded substrate glycoprotein by the UGGT, as well as of the structure of CRT and CNX. Some of the 'rules' underlying chaperone selection have also been deduced. In this context, it has to be kept in mind, however, that the chaperone system is quite flexible, with chaperones able to substitute for each other, cooperate in different ways and respond differently to physiological changes and cellular stress.

With regards to glycoprotein folding in the ER, future work is likely to focus on the interplay between the different components of the CNX/CRT cycle at the molecular level. Important unresolved questions include if and how conformational changes occur upon binding of substrate glycoprotein by CRT and CNX, how the co-chaperone ERp57 is recruited by CRT and CNX, how it manages to interact productively with the large number of different substrates, and finally to what extent CRT and CNX are involved in protein–protein interactions with unfolded proteins?

Acknowledgements

We are thankful to Rolf Moser for help with the figures and to Anna Mezzacasa and Christiane Ritter for critical reading of the manuscript. The work was supported by the Swiss National Science Foundation.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Gething MJ, Sambrook J: Protein folding in the cell. *Nature* 1992, 355:33-45.

2. Hurtley SM, Helenius A: Protein oligomerization in the endoplasmic reticulum. *Annu Rev Cell Biol* 1989, 5:277-307.

3. Ellgaard L, Molinari M, Helenius A: Setting the standards: quality control in the secretory pathway. *Science* 1999, 286:1882-1888.

4. Aridor M, Balch WE: Integration of endoplasmic reticulum signaling in health and disease. *Nat Med* 1999, 5:745-751.

5. Herrmann JM, Malkus P, Schekman R: Out of the ER — outfitters, escorts and guides. *Trends Cell Biol* 1999, 9:5-7.

6. Nagata K: Hsp47: a collagen specific molecular chaperone. *Trends Biochem Sci* 1996, 21:22-26.

7. Bu G, Geuze HJ, Strous GJ, Schwartz AL: 39kDa receptor-associated protein is an ER resident proteins and molecular chaperone for LDL receptor related proteins. *EMBO J* 1995, 14:2269-2280.

8. Ortmann B, Copeman J, Lehner PJ, Sadasivan B, Herberg JA, Grandea AG, Riddell SR, Tampe R, Spies T, Trowsdale J *et al.*: A critical role for tapasin in the assembly and function of multimeric MHC class I-TAP complexes. *Science* 1997, 277:1306-1309.

9. Balch WE, McCaffery JM, Plutner H, Farquhar MG: Vesicular stomatitis virus is sorted and concentrated during export from the endoplasmic reticulum. *Cell* 1994, 76:841-852.

10. Nishimura N, Balch WE: A di-acidic signal required for selective export from the endoplasmic reticulum. *Science* 1997, 277:556-558.

11. Ma D, Zerangue N, Lin YF, Collins A, Yu M, Jan YN, Jan LY: Role of ER export signals in controlling surface potassium channel numbers. *Science* 2001, 291:316-319.

12. Zerangue N, Schwappach B, Jan YN, Jan LY: A new ER trafficking signal regulates the subunit stoichiometry of plasma membrane K(ATP) channels. *Neuron* 1999, 22:537-548.

13. Zerangue N, Malan MJ, Fried SR, Dazin PF, Jan YN, Jan LY, Schwappach B: Analysis of endoplasmic reticulum trafficking signals by combinatorial screening in mammalian cells. *Proc Natl Acad Sci USA* 2001, 98:2431-2436.

14. Hammond C, Helenius A: 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* 1994, 126:41-52.

15. Werner ED, Brodsky JL, McCracken AA: Proteasome-dependent endoplasmic reticulum-associated protein degradation: an unconventional route to a familiar fate. *Proc Natl Acad Sci USA* 1996, 93:13797-13801.

16. Hiller MM, Finger A, Schweiger M, Wolf DH: ER degradation of a misfolded luminal protein by the cytosolic ubiquitin-proteasome pathway. *Science* 1996, 273:1725-1728.

17. Bonifacino JS, Weissman AM: Ubiquitin and the control of protein fate in the secretory and endocytic pathways. *Annu Rev Cell Dev Biol* 1998, 14:19-57.

18. Kopito RR, Ron D: Conformational disease. *Nat Cell Biol* 2000, 2:E207-209.

19. Mori K, Sant A, Kohno K, Normington K, Gething MJ, Sambrook JF: A 22 bp cis-acting element is necessary and sufficient for the induction of the yeast KAR2 (BiP) gene by unfolded proteins. *EMBO J* 1992, 11:2583-2593.

20. Chapman R, Sidrauski C, Walter P: Intracellular signaling from the endoplasmic reticulum to the nucleus. *Annu Rev Cell Dev Biol* 1998, 14:459-485.

21. Kaufman RJ: Stress signaling from the lumen of the endoplasmic reticulum: coordination of gene transcriptional and translational controls. *Genes Dev* 1999, 13:1211-1233.

22. •• Travers KJ, Patil CK, Wodicka L, Lockhart DJ, Weissman JS, Walter P: Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. *Cell* 2000, 101:249-258.

Using DNA microarrays, the authors investigate which genes are induced by the UPR in *S. cerevisiae*. Approximately 400 UPR targets were identified representing a wide range of proteins of the secretory pathway. The demonstration of coordinated responses of UPR and ERAD for the disposal of misfolded ER proteins links the two processes mechanistically.

23. Chevet E, Cameron PH, Pelletier MF, Thomas DY, Bergeron JJ: The endoplasmic reticulum: integration of protein folding, quality control, signaling and degradation. *Curr Opin Struct Biol* 2001, 11:120-124.

24. Hammond C, Braakman I, Helenius A: Role of N-linked oligosaccharides, glucose trimming and calnexin during glycoprotein folding in the endoplasmic reticulum. *Proc Natl Acad Sci USA* 1994, 91:913-917.

25. Oliver JD, van der Wal FJ, Bulleid NJ, High S: Interaction of the thiol-dependent reductase ERp57 with nascent glycoproteins. *Science* 1997, 275:86-88.

26. • Parodi A: Protein glucosylation and its role in protein folding. *Annu Rev Biochem* 2000, 69:69-93.
An outstanding review on the role of N-linked glycosylation in protein folding.

27. Ellgaard L, Helenius A: Folding of influenza hemagglutinin in the living cell. In *Mechanisms of Protein Folding: Frontiers in Molecular Biology,* edn 2. Edited by Pain R. Oxford University Press; 2000:352-363.

28. • Molinari M, Helenius A: Chaperone selection during glycoprotein translocation into the endoplasmic reticulum. *Science* 2000, 288:331-333.
Using virus-infected cells, chaperone selection for viral glycoproteins during translocation into the ER was investigated. The results show that if glycoproteins contain an N-linked glycan within the first ~50 residues from the amino terminus they interact cotranslationally with CRT and CNX. If the glycan appears later in the sequence, the initial chaperone interaction occurs through BiP.

29. Ware FE, Vassilakos A, Peterson PA, Jackson MR, Lehrman MA, Williams DB: The molecular chaperone calnexin binds $Glc_1Man_9GlcNAc_2$ oligosaccharides as an initial step in recognizing unfolded glycoproteins. *J Biol Chem* 1995, 270:4697-4704.

30. Spiro RG, Zhu Q, Bhoyroo V, Söling HD: Definition of the lectin-like properties of the molecular chaperone, calreticulin, and demonstration of its copurification with endomannosidase from rat liver Golgi. *J Biol Chem* 1996, 271:11588-11594.

31. Vassilakos A, Michalak M, Lehrman MA, Williams DB: Oligosaccharide binding characteristics of the molecular chaperones calnexin and calreticulin. *Biochemistry* 1998, 37:3480-3490.

32. Peterson JR, Ora A, Nguyen Van P, Helenius A: Transient, lectin-like association of calreticulin with folding intermediates of cellular and viral glycoproteins. *Mol Biol Cell* 1995, 6:1173-1184.

33. Wada I, Imai S, Kai M, Sakane F, Kanoh H: Chaperone function of calreticulin when expressed in the endoplasmic reticulum as the membrane-anchored and soluble forms. *J Biol Chem* 1995, 270:20298-20304.

34. Van Leeuwen JEM, Kearse KP: The related molecular chaperones calnexin and calreticulin differentially associate with nascent T cell antigen receptor proteins within the endoplasmic reticulum. *J Biol Chem* 1996, 271:25345-25349.

35. Hebert DN, Zhang JX, Chen W, Foellmer B, Helenius A: The number and location of glycans on influenza hemagglutinin determine folding and association with calnexin and calreticulin. *J Cell Biol* 1997, 139:613-623.

36. •• Ellgaard L, Riek R, Herrmann T, Guntert P, Braun D, Helenius A, Wuthrich K: NMR structure of the calreticulin P-domain. *Proc Natl Acad Sci USA* 2001, 98:3133-3138.
The three-dimensional structure of the CRT P-domain solved by NMR spectroscopy reveals a novel hairpin fold formed by the entire polypeptide chain. The elongated structure is stabilized by three short anti-parallel β-sheets and three small hydrophobic clusters. The structure provides new clues for the function of CRT as a molecular chaperone.

37. Danilczyk UG, Cohen-Doyle MF, Williams DB: Functional relationship between calreticulin, calnexin, and the endoplasmic reticulum luminal domain of calnexin. *J Biol Chem* 2000, 275:13089-13097.

38. • Oliver JD, Roderick HL, Llewellyn DH, High S: ERp57 functions as a subunit of specific complexes formed with the ER lectins calreticulin and calnexin. *Mol Biol Cell* 1999, 10:2573-2582.
Demonstration that both CRT and CNX can be isolated in discrete complexes with ERp57, but not PDI, from, for instance, semi-permeabilized cells and microsomes. The interaction was shown to be independent of the presence of substrate glycoprotein.

39. •• Zapun A, Darby NJ, Tessier DC, Michalak M, Bergeron JJ, Thomas DY: Enhanced catalysis of ribonuclease B folding by the interaction of calnexin or calreticulin with ERp57. *J Biol Chem* 1998, 273:6009-6012.
Using purified components, it was clearly demonstrated that ERp57 accelerates the *in vitro*, oxidative refolding of monoglucosylated RNaseB only in the presence of CRT or CNX. The paper provides the first direct evidence that the disulfide isomerase activity of ERp57 towards glycoproteins is enhanced through association with CRT and CNX.

40. •• Molinari M, Helenius A: Glycoproteins form mixed disulphides with oxidoreductases during folding in living cells. *Nature* 1999, 402:90-93.
The first demonstration that transient mixed disulfides are formed *in vivo* between thiol oxidoreductases PDI and ERp57 and their respective substrate proteins. Furthermore, ERp57 was found to be the oxidoreductase of choice for CRT- and CNX-bound substrate glycoproteins.

41. Tessier DC, Dignard D, Zapun A, Radominska-Pandya A, Parodi AJ, Bergeron JJ, Thomas DY: Cloning and characterization of mammalian UDP-glucose glycoprotein: glucosyltransferase and the development of a specific substrate for this enzyme. *Glycobiology* 2000, 10:403-412.

42. Vanstapel F, Blanckaert N: Carrier-mediated translocation of uridine diphosphate glucose into the lumen of endoplasmic reticulum-derived vesicles from rat liver. *J Clin Invest* 1988, 82:1113-1122.

43. Cannon KSl, Helenius A: Trimming and readdition of glucose to N-Linked oligosaccharides determines calnexin association of a substrate glycoprotein in living cells. *J Biol Chem* 1999, 274:7537-7544.

44. Sousa MC, Ferrero-Garcia MA, Parodi AJ: Recognition of the oligosaccharide and protein moieties of glycoproteins by the UDP-Glc: glycoprotein glucosyltransferase. *Biochemistry* 1992, 31:97-105.

45. Trombetta S, Bosch M, Parodi AJ: Glucosylation of glycoproteins by mammlian, plant, fungal and trypanosomatid protozoa microsomal membranes. *Biochemistry* 1989, 28:8108-8116.

46. Sousa M, Parodi AJ: The molecular basis for the recognition of misfolded glycoproteins by the UDP-Glc:glycoprotein glucosyltransferase. *EMBO J* 1995, 14:4196-4203.

47. Trombetta ES, Helenius A: Conformational requirements for glycoprotein reglucosylation in the endoplasmic reticulum. *J Cell Biol* 2000, 148:1123-1130.

48. Labriola C, Cazzulo JJ, Parodi AJ: ***Trypanosoma cruzi*** calreticulin is a lectin that binds monoglucosylated oligosaccharides but not protein moieties of glycoproteins. *Mol Biol Cell* 1999, 10:1381-1394.

49. Ritter C, Helenius A: Recognition of local misfolding by the ER folding sensor UDP-glucose:glycoprotein glucosyltransferase. *Nat Struct Biol* 2000, 7:278-280.

50. Ellgaard L, Riek R, Braun D, Herrmann T, Helenius A, Wuthrich K: Three-dimensional structure topology of the calreticulin P-domain based on NMR assignment. *FEBS Lett* 2001, 488:69-73.

51. Hahn M, Borisova S, Schrag JD, Tessier DC, Zapun A, Tom R, Kamen AA, Bergeron JJM, Thomas DY, Cygler M: Identification and crystallization of a protease-resistant core of calnexin that retains biological activity. *J Struct Biol* 1998, 123:260-264.

52. • Bouvier M, Stafford WF: Probing the three-dimensional structure of human calreticulin. *Biochemistry* 2000, 39:14950-14959.
Structural properties of CRT were characterized by thorough biochemical and biophysical analysis. Analytical ultracentrifugation and gel filtration indicate that the protein is a highly assymmetric molecule. Circular dichroism measurements shows a relatively low thermal denaturation transition midpoint of 42.5°C, whereas limited proteolysis identifies stable core domains, with one corresponding to the P-domain.

53. Peterson JR, Helenius A: ***In vitro*** reconstitution of calreticulin-substrate interactions. *J Cell Sci* 1999, 112:2775-2784.

54. Vijayan M, Chandra N: Lectins. *Curr Opin Struct Biol* 1999, 9:707-714.

55. Rost B, Sander C: Prediction of protein secondary structure at better than 70% accuracy. *J Mol Biol* 1993, 232:584-599.

56. Patil AR, Thomas CJ, Surolia A: Kinetics and the mechanism of
• interaction of the endoplasmic reticulum chaperone, calreticulin, with monoglucosylated ($Glc_1Man_9GlcNAc_2$) substrate. *J Biol Chem* 2000, 275:24348-24356.

This paper contains a convincing demonstration that CRT interacts only with the monoglucosylated ($Glc_1Man_9GlcNAc_2$) form of chicken IgG *in vitro*, that the binding is monovalent and that it does not involve any protein–protein interaction. A moderate affinity of CRT for its substrate in the range of ~10^5 M^{-1} was determined by surface plasmon resonance.

57. Hauri H, Appenzeller C, Kuhn F, Nufer O: Lectins and traffic in the secretory pathway. *FEBS Lett* 2000, 476:32-37.

58. Fiedler K, Simons K: A putative novel class of animal lectins in the secretory pathway homologous to leguminous lectins. *Cell* 1994, 77:625-626.

59. Saito Y, Ihara Y, Leach MR, Cohen-Doyle MF, Williams DB:
• Calreticulin functions ***in vitro*** as a molecular chaperone for both glycosylated and non-glycosylated proteins. *EMBO J* 1999, 18:6718-6729.

Using *in vitro* assays, CRT was shown to bind unfolded, glycosylated as well as non-glycosylated proteins, prevent their aggregation and keep them in a folding competent state. The paper indicates a potential role for CRT as a classical molecular chaperone and advocates the idea of a 'dual-binding' mode of CRT to substrate glycoproteins.

60. Ihara Y, Cohen-Doyle MF, Saito Y, Williams DB: Calnexin discriminates between protein conformational states and functions as a molecular chaperone ***in vitro***. *Mol Cell* 1999, 4:331-341.

61. Rodan AR, Simons JF, Trombetta ES, Helenius A: N-linked oligosaccharides are necessary and sufficient for association of RNase B with calnexin and calreticulin. *EMBO J* 1996, 15:6921-6930.

62. Hebert DN, Zhang JX, Helenius A: Protein folding and maturation in a cell-free system. *Biochem Cell Biol* 1998, 76:867-873.

63. David V, Hochstenbach F, Rajagopalan S, Brenner MB: 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* 1993, 268:9585-9592.

64. Menetret J, Neuhof A, Morgan DG, Plath K, Radermacher M, Rapoport TA, Akey CW: The structure of ribosome-channel complexes engaged in protein translocation. *Mol Cell* 2000, 6:1219-1232.

65. Chevet E, Wong HN, Gerber D, Cochet C, Fazel A, Cameron PH, Gushue JN, Thomas DY, Bergeron JJ: Phosphorylation by CK2 and MAPK enhances calnexin association with ribosomes. *EMBO J* 1999, 18:3655-3666.