The endoplasmic reticulum (ER) is the subcellular site where proteins following the secretory pathway acquire their proper tertiary and, in certain cases, quaternary structures. Species that are not yet properly folded are prevented from exit to the Golgi apparatus and, if permanently misfolded, are transported to the cytosol, where they are degraded in the proteasomes. This review deals with a mechanism, applicable to proteins that are N-glycosylated in the ER, by which the quality control of folding is performed. Protein-linked monoglucosylated glycans, formed by glucosidase I- and glucosidase II-dependent partial deglucosylation of the oligosaccharides transferred from dolichol diphosphate derivatives in N-glycosylation (Glc$_3$Man$_2$GlcNAc$_3$), mediate glycoprotein recognition by two ER-resident lectins, membrane-bound calnexin (CNX) and its soluble homologue, calreticulin (CRT). A still not yet fully confirmed interaction between the lectins and the protein moieties of folding glycoproteins may occur after lectin recognition of monoglucosylated structures. Further deglucosylation of glycans by glucosidase II, and perhaps also by a change in CNX/CRT and/or in the substrate glycoprotein conformation, liberates the glycoproteins from their CNX/CRT anchors. Glycans may be then re-glucosylated by the UDP-Glc:glycoprotein glucosyltransferase (GT), and thus be recognized again by CNX/CRT, but only when linked to not yet properly folded protein moieties, as this enzyme behaves as a sensor of glycoprotein conformation. Deglucosylation/reglucosylation cycles catalysed by the opposing activities of glucosidase II and GT only stop when proper folding is achieved. The interaction between CNX/CRT and a monoglucosylated glycan is one of the alternative mechanisms by which cells retain not yet properly folded glycoproteins in the ER; in addition, it enhances folding efficiency by preventing protein aggregation and thus allowing intervention of classical chaperones and other folding-assisting proteins. There is evidence suggesting that both glycoprotein glycosylation and mannose removal, respectively mediated by GT and ER mannosidase I, might be involved in cell recognition of permanently misfolded glycoproteins bound for proteasome degradation.

Key words: calnexin/calreticulin, glucosyltransferase, glycoprotein degradation, quality control.

---

**PROTEIN GLYCOSYLATION AND GLYCAN PROCESSING IN THE ENDOPLASMIC RETICULUM (ER)**

The main features of the pathway leading to protein glycosylation and oligosaccharide processing in the lumen of the ER of mammalian, plant and yeast cells were already known at the beginning of the 1980s, but their full meaning was not clearly understood [1–3]. It had been established that a dolichol-P-P-linked oligosaccharide (Glc$_3$Man$_2$GlcNAc$_3$; Scheme 1) is transferred to Asn residues in nascent polypeptide chains. This glycan represents a well conserved structure in evolution, as the same compound is transferred in wild-type plant, fungal and mammalian cells [1–3].

The oligosaccharyltransferase is a very complex enzyme, formed by eight subunits located in the ER membrane. In cell-free assays, the enzyme transfers the triglucosylated oligosaccharide about 20-fold faster than the compound lacking Glc units. This property determines the occurrence of underglycosylated glycoproteins in cell mutants unable to elongate Man$_n$GlcNAc$_{2-n}$-P-P-dolichol [4]. The requirement of three Glc units in the oligosaccharide for an efficient transfer reaction did not initially seem fully rational, as these residues are removed immediately once the glycan is linked to protein.

Several ER enzymes are involved in oligosaccharide processing. Glucosidase I (GI; a membrane-bound enzyme) [3] removes the external Glc unit (residue $\mathbf{n}$, Scheme 1), whereas glucosidase II (GII; a soluble heterodimer composed of catalytic and ER-retaining subunits) [5,6] excises both remaining units (residues $\mathbf{i}$ and $\mathbf{m}$, Scheme 1). Two specific ER $\alpha$-mannosidases may remove up to two mannose units in the mammalian cell ER (residues $\mathbf{i}$ and $\mathbf{k}$, by $\alpha$-mannosidases I and II respectively; see Scheme 1) [7]. The single $\alpha$-mannosidase present in *Saccharomyces cerevisiae* ER removes residue $\mathbf{i}$ (Scheme 1) [8].

Further processing of the oligosaccharide may occur in the Golgi complex. The extreme structural diversity found in oligosaccharides linked to Asn residues in mature, fully processed glycoproteins of different organisms or of different cell types belonging to multicellular organisms originates by differential processing reactions in the Golgi apparatus. This structural diversity presents a certain logic, as the roles of protein-linked oligosaccharides in mature glycoproteins are related mainly to recognition phenomena [9]. In contrast, glycan processing in the...
ER is basically similar in all cells because, as will be described below, the role of these reactions is directly related to a property common to all glycoproteins, namely the acquisition of their correct tertiary structures.

**TRANSIENT GLUCOSYLATION OF GLYCOPROTEINS IN TRYPANOSOMATIDS**

Glucosylation of glucose-free high-mannose-type oligosaccharides was first detected in trypanosomatid cells (several trypanosomatid species are the causative agents of diseases endemic in developing countries). On isolation of dolichol-3-P-derivatives from live trypanosomatid cells incubated with [3H]Glc, it was found that trypanosomatid protozoa were (and so far still are) the only wild-type cells that transfer unglucosylated oligosaccharides in protein N-glycosylation in vivo (Man$_3$GlcNAc$_n$, Man$_6$GlcNAc$_n$, or Man$_9$GlcNAc$_n$, depending on the species; see Scheme 1) (for a review on protein N-glycosylation in trypanosomatid protozoa, see [10]). In contrast with oligosaccharide transfers from other eukaryotes, those from trypanosomatid protozoa transfer Man$_3$,GlcNAc$_n$ and Glc$_1$,Man$_3$,GlcNAc$_n$ at the same rate in cell-free assays [11]. Structural analysis revealed that oligosaccharides were transiently glucosylated once they were transferred to proteins. For instance, protein-linked Glc$_3$,Man$_6$,GlcNAc$_n$, Glc$_4$,Man$_9$,GlcNAc$_n$, and Glc$_5$,Man$_9$,GlcNAc$_n$ were characterized in Trypanosoma cruzi cells (this parasite is the causative agent of Chagas’ disease; it transfers Man$_3$,GlcNAc$_n$ in protein N-glycosylation) [12]. The precursor of the Glc units was probably UDP-Glc, because it was also found that these protozoa are able to synthesize dolichol-3-P-Glc, the donor of Glc units in Glc$_3$,Man$_6$,GlcNAc$_n$-3-P-dolichol [13]. Further work showed that Glc$_5$,Man$_9$,GlcNAc$_n$, and Glc$_6$,Man$_9$,GlcNAc$_n$, are formed by glucosylation of the unglucosylated species, and not by demannosylation of Glc$_3$,Man$_6$,GlcNAc$_n$. The monoglucosylated glycans had a transient existence, as they disappeared upon chasing cells with unlabelled Glc (glycan processing reactions occurring in T. cruzi ER are depicted in Scheme 2A).

**GLYCOPROTEIN GLUCOSYLATION IS NOT RESTRICTED TO TRYPANOSOMATID CELLS**

In contrast with what happens in trypanosomatids, Glc$_3$,Man$_6$,GlcNAc$_n$, Glc$_4$,Man$_9$,GlcNAc$_n$, and Glc$_5$,Man$_9$,GlcNAc$_n$ could theoretically be formed in mammalian cells by two different pathways: (1) partial deglucosylation of Glc$_3$,Man$_6$,GlcNAc$_n$ to Glc$_3$,Man$_6$,GlcNAc$_j$, followed by demannosylation of the latter compound, or (2) complete deglucosylation of the transferred compound, followed first by partial demannosylation of Glc$_3$,Man$_6$,GlcNAc$_n$ and then by direct glucosylation of unglucosylated glycans. The occurrence of direct glucosylation of protein-linked Man$_6$,GlcNAc$_n$, Man$_9$,GlcNAc$_n$, and Man$_9$,GlcNAc$_n$ in trypanosomatids suggested the design of several experiments, performed in live cells and in a subcellular fraction enriched in ER vesicles, that demonstrated that the same process occurred in mammalian, plant and fungal cells [14–16]. The ER was identified as the subcellular site of transient glycoprotein glucosylation (ER processing of protein-linked oligosaccharides in mammalian cells is shown in Scheme 2B). As in trypanosomatids, monoglucosylated glycans also had a transient existence in live mammalian cells. The structure of Glc$_3$,Man$_6$,GlcNAc$_j$ formed by direct glucosylation was identical with that formed by partial deglucosylation of Glc$_3$,Man$_6$,GlcNAc$_n$ (residues a–l in Scheme 1), thus indicating that GII was also responsible for deglucosylation of the former compound in vivo [14,17].

Although rat liver microsomes incubated with UDP-Glc glucosylated endogenous glycoproteins, no glucosylation of exogenously added high-mannose-type glycoproteins was observed. Further work showed that added glycoproteins had to be previously denatured in order to be glucosylated by the activity present in the ER, i.e. UDP-Glc:glycoprotein glucosyltransferase (GT) [17]. Glycopeptides such as Man$_3$,GlcNAc$_2$,Asn were not glucosylated, thus indicating that the effect of denaturation was not to render the oligosaccharide accessible to GT, but rather to expose or create protein domains whose recognition by the enzyme was apparently required for glycan
Oligosaccharide processing reactions occurring in the ER of \textit{T. cruzi} (A) and mammalian (B) cells are shown. GNA, GlcNAc; M, Man; G, Glc; D, dolichol; Pr, protein; OT, oligosaccharyltransferase; MI and MII, mannosidases I and II respectively.

Glucosylation [18]. A paper by Suh et al. [19] confirmed that, indeed, misfolded glycoproteins were glucosylated by GT \textit{in vivo}: glycoprotein G of the thermosensitive vesicular stomatitis virus (VSV) ts045 mutant was retained in the ER of infected cells at the non-permissive temperature (39 °C) in two interconvertible (monoglucosylated and unglucosylated) forms. Around 55% of all N-linked oligosaccharides in mature, properly folded glycoproteins were found to be monoglucosylated in \textit{T. cruzi} cells (in which monoglucosylated glycoproteins are formed exclusively by GT; see Schemes 2A and 2B) grown in the presence of a GII inhibitor (1-deoxynojirimycin) [20]. This indicated that glucosylation of glycoproteins was not restricted to permanently folded species, but that it also occurred during productive folding. Moreover, as glycoproteins generally have more than one N-oligosaccharide, the value obtained suggested that practically all glycoproteins were glucosylated in at least one of their oligosaccharides. Based on the unique properties of GT, it was suggested that the enzyme (and monoglucosylated oligosaccharides) could somehow be involved in the so-called ‘quality control’ of glycoprotein folding in the ER (see below) [17,18,20,21].

The availability of an assay for GT activity (incubation of Glc-labelled UDP-Glc, a denatured high-mannose-type glycoprotein and an enzyme source, followed by quantification of label in 10% trichloroacetic acid-insoluble material) allowed detection of the enzyme not only in mammalian and protozoan cell-derived microsomes, but also in vesicles isolated from fungi and plants [17]. GT was purified to homogeneity from rat liver and the fission yeast \textit{Schizosaccharomyces pombe} [21,22] and also, but for unexpected reasons, from \textit{Drosophila melanogaster} cells [23]. While studying extracellular-matrix formation in fly tissues, J. Fessler’s group purified to homogeneity a secreted protein [23]. Immunolocalization studies revealed that the protein was mainly localized in the ER. Its sequence, size and localization suggested that the protein could be the \textit{D. melanogaster} homologue of rat liver and \textit{Schiz. pombe} GT. Enzymic assays confirmed this idea. GT is a rather large (160 kDa) soluble ER protein that has an absolute requirement of millimolar Ca\textsuperscript{2+} concentrations for activity and uses UDP-Glc, but not TDP-Glc, ADP-Glc or UDP-Gal, as substrate donor [21,22].

\textbf{Identification of structures recognized by GT in misfolded conformers}

Once proteins depart from their native conformation, they may adopt a variety of different structures that are extremely difficult to characterize. This fact has hindered identification of structural
features recognized by GT in misfolded conformers. Nevertheless, several experiments have provided hints as to the nature of these structures and the folding stages in which glycoproteins are glucosylated.

Although denatured non-glycosylated proteins do not affect GT activity, denatured glycoproteins from which the oligosaccharide had been removed by endo-β-N-acetylglucosaminidase H (Endo H) treatment (i.e. leaving a single GlcNAc linked to the Asn units) proved to be very efficient inhibitors [24]. For instance, Endo H-treated denatured RNase B inhibited glucosylation of denatured RNase B, whereas RNase A or Endo H-treated native RNase B had no effect (RNases A and B have exactly the same composition, except for the presence of a single N-oligosaccharide in the latter). On the other hand, although Man<sub>4</sub>GlcNAc<sub>2</sub>-Asn was not (or was very poorly) glucosylated by GT, even in the presence of denatured β-lactoglobulin (a non-glycosylated protein), cross-linking the two structures with glutaraldehyde resulted in a very efficient substrate [24]. It was concluded that, in an efficient GT acceptor substrate, the protein recognition elements and the oligosaccharide have to be co-valently linked and the innermost GlcNAc residue has to be accessible to GT. This last requirement probably contributes to the exclusive glucosylation of incompletely folded glycoproteins, as the innermost GlcNAc units interact, in native conformers, with neighbouring amino acids, and are often not accessible to macromolecular probes [25,26]. The requirement of a covalent linkage between the oligosaccharide and the protein indicates that the folding stage of a glycoprotein molecule does not influence the glucosylation of other molecules. GT recognizes not only the innermost GlcNAc but also the Man units, as the respective relative glucosylation rates of Man<sub>4</sub>GlcNAc<sub>2</sub>, Man<sub>4</sub>GlcNAc<sub>2</sub>, and Man<sub>4</sub>GlcNAc<sub>2</sub> are 100, 50 and 15 (this result was obtained considering only the isomers of the last two compounds that contain Man residue <i>g</i> in Scheme 1, i.e. the residue to which the Glc residue is attached). A decrease in the rate of deglucosylation by purified GII was also observed upon removal of Man units from Glc<sub>2</sub>Man<sub>2</sub>GlcNAc<sub>2</sub> [27].

With regard to the nature of protein structures recognized by GT in misfolded conformations, it was initially suggested that perhaps they could be patches of hydrophobic amino acid side chains, as the enzyme was found to bind hydrophobic amino acids under physiological conditions of pH and salt concentration, and exposure of hydrophobic amino acid side chains is a feature common to all misfolded conformers [18]. Nevertheless, no evidence indicating that such patches actually are the elements triggering glucosylation has been obtained.

Evidence gathered in both cell-free and in vivo assays suggested that GT might recognize structures exposed not in the initial, but in the final, folding stages. Thus, although a neoglycoprotein formed by chemical linkage between full-length staphylococcal nuclease (a non-glycosylated 149-amino-acid protein) and Man<sub>4</sub>GlcNAc<sub>2</sub>-Asn necessarily required previous denaturation in 8 M urea to be glucosylated by GT, truncated versions of the same molecule lacking 14 amino acids at the C-terminus were good GT substrates in the absence of any denaturing treatment [24]. The nuclease-specific activities of the truncated neoglycoproteins were about one-quarter to one-third of that of the full-length native neoglycoprotein, thus showing that GT substrates can display conformations closely resembling native ones. Cleavage of the bond between amino acids 20 and 21 in RNase B (a 124-amino-acid protein) yielded a peptide plus a protein core that had conserved full enzymic activity. The latter structure, which had the N-oligosaccharide, was not glucosylated by GT unless the peptide was removed, a procedure that also abolished RNase activity [27a]. The isolated 104-amino-acid core nevertheless retained much of the native structure, due to the presence of four disulphide bridges. On the other hand, complete reduction followed by S-alkylation of RNase B with iodoacetamide or iodoacetic acid produced fully soluble unfolded RNase B conformers that were poorly glucosylated by GT.

These results, suggesting that GT requires partially structured glycoproteins for glucosylation, agree with experiments performed in vivo indicating that induction of glycoprotein misfolding by the addition of dithiothreitol to <i>Schiz. pombe</i> cells does not result in a generalized increase in GT-mediated glycoprotein glucosylation [28]. Furthermore, the fact that formation of monoglucosylated oligosaccharides in <i>T. cruzi</i> depends exclusively on GT activity (Scheme 2A) allowed pinpointing, with some accuracy, of the folding stages at which cruzipain (a lysosomal proteinase having two or three N-oligosaccharides and six or seven disulphide bridges) was glucosylated. As will be described below, monoglucosylated oligosaccharides are recognized in vivo by ER-resident lectins, calnexin (CNX) and calreticulin (CRT). Interaction of cruzipain with CRT (trypansomatid protozoa lack CNX) occurred only when all or nearly all disulphide bridges had already been formed, thus indicating a requirement of at least a partially structured molecule for GT-mediated glucosylation [29].

GT has been sequenced from mammalian, insect, nematode and fungal sources ([23,30], and GenBank accession number U28735). Whereas the C-terminal portions (30% of the molecule) of GTs from different sources show a fairly high degree of similarity (65–70% identity), similarity is much lower at the N-terminal ends. This is particularly noticeable when comparing the sequence of <i>Schiz. pombe</i> GT with those from other sources. As there is a significant, although limited, identity between GT C-terminal portions and bacterial glycosyltransferases that utilize UDP-Glc or UDP-Gal as substrate donors, it may be speculated that part of the enzyme is responsible for sugar nucleotide (and probably also oligosaccharide) recognition. On the other hand, GT N-terminal portions show much lower sequence identity, as they probably have to sense the folding status of glycoproteins.

Recognition of a wide variety of different structures, exclusively mediated glucosylation [29], requires much less stringent sequence identity. All GTs sequenced so far display ER retrieval sequences for soluble proteins at their C-termini.

**THE ER AS A CONTROL POINT**

Proteins enter the secretory pathway in the ER, where they are covalently modified (cleavage of the signal peptide, N-glycosylation, formation of disulphide bonds) and acquire their proper tertiary and, in some cases, also quaternary structures. Proteins that fail to fold properly are initially retained in the ER and eventually transported to the cytosol, where they are proteolytically degraded in the proteasomes [31,32]. Cells have, therefore, mechanisms for discriminating between different conformers. These mechanisms have been globally referred to as the ‘quality control’ of protein folding [33]. As will be described further below, there are exceptions in which misfolded glycoproteins are transported to their correct subcellular location.

Proper folding of proteins is facilitated in the ER by a battery of classical molecular chaperones [BIP (immunoglobulin heavy-chain binding protein)/glucose-regulated protein (GRP)68, GRP94, GRP170, etc.], by unconventional chaperones (CNX and CRT) and by proteins with activity to catalyse protein disulphide-bond formation [protein disulphide isomerase (PDI or ERP59), ERP72, ERP57, etc.] [34–39]. These proteins display thiol-protein disulphide oxidoreductase and isomerase activities.
The lumen of the ER is an oxidizing environment that facilitates the formation of disulphide bonds. On the other hand, the ER lumen is the main cellular Ca\(^{2+}\) reservoir. The function of several chaperones and folding-assisting proteins is dependent on the correct Ca\(^{2+}\) concentration [40].

CNX is a 572-amino-acid type I transmembrane protein (initially known also as p88 and IP90), whereas CRT (a 400-amino-acid protein) is its soluble homologue [41–43]. Both proteins have high-affinity/low-capacity and low-affinity/high-capacity Ca\(^{2+}\) binding sites. As will be described below, the middle (or P) domains of both proteins, which show high homology, have lectin properties.

**GLYCANS ARE REQUIRED FOR PRODUCTIVE FOLDING OF MANY GLYCOPROTEINS**

N-glycosylation may be prevented in vivo by inhibiting formation of the dolichol-P-P precursor (either by the addition of compounds inhibiting biosynthesis, such as tunicamycin, or by mutations affecting the enzymes involved) or by eliminating N-glycosylation consensus sequences by site-directed mutagenesis in glycoproteins under study. In addition, new N-glycosylation sites may also be created by site-directed mutagenesis. Numerous reports have shown that the presence of N-oligosaccharides is required for the correct folding of many glycoproteins [44–46]. Misfolded species lacking oligosaccharides are generally found in the ER, forming large aggregates that are non-covalently bound to BiP and other ER chaperones, and often are covalently linked to each other through aberrant interchain disulphide bonds [47,48]. In some glycoproteins all N-oligosaccharides are required for proper folding of all molecules, whereas in other cases a sizeable fraction of the molecule may fold correctly upon elimination of all N-glycans [49,50]. In some glycoproteins, some N-oligosaccharides are more important than others for proper folding; elimination of a particular glycosylation site may not affect folding, whereas elimination of a different one may result in complete misfolding of the same glycoprotein [51,52]. Rather surprisingly, creation of new N-glycosylation sites without elimination of the pre-existing ones resulted in a hyperglycosylated influenza virus haemagglutinin that showed temperature dependence in transport (i.e. probably in ER folding) [53]. In some cases, elimination of N-oligosaccharides may result in lower temperature-dependence for folding [54].

In other cases, however, the presence, but not the location, of N-oligosaccharides is important for proper folding. Thus elimination of both glycosylation sites from VSV G protein resulted in misfolding of the protein. Proper folding was restored on creation of new N-glycosylation sites at different locations [55,56]. Finally, there are many cases in which N-oligosaccharides are completely dispensable for proper folding and secretion [57–59].

That the effect of N-oligosaccharides on correct folding cannot be predicted may be highlighted by the observation that the G protein of only one of two VSV strains required N-oligosaccharides for proper folding. A single point mutation (Y172D) converted a form for which folding was N-oligosaccharide-dependent to an N-oligosaccharide-independent form [60]. Similarly, a chimaeric membrane protein formed by rat growth hormone (a non-glycosylated soluble protein) and a C-terminal extension containing VSV G protein transmembrane and cytosolic domains was not transported to the cell surface unless one N-glycosylation consensus sequence was created at a random position in the hormone portion of the molecule [61].

As will be seen below, not only do N-oligosaccharides provide bulky, highly hydrophilic groups that help to maintain glycoproteins in solution during the folding process or modulate protein conformation by forcing amino acids close to the linking Asn unit to be in the proximity of the water–glycoprotein interface, but the interaction of certain specific structures (monoglucosylated oligosaccharides) with ER lectins potentiates the folding-facilitating effect of the high-mannose core and provides a mechanism for retaining molecules that are not yet correctly folded in the ER.

**CNX AND CRT ARE LECTINS THAT SPECIFICALLY BIND MONOGLUCOSYLATED GLYCOPROTEINS**

CNX was initially described as a protein that transiently recognized class I MHC molecules, membrane-bound immunoglobulins and T-cell antigen receptor molecules [62,63]. Ou et al. [64] published a seminal article in which they showed that CNX was associated transiently with a series of glycoproteins synthesized by human hepatoma cells. What was remarkable was that tunicamycin (an inhibitor of protein N-glycosylation) prevented association. Moreover, the main non-glycosylated protein secreted by the above-mentioned cells (albumin) did not interact with CNX. The effect of tunicamycin was a serendipitous observation, because further reports showed that other glycoproteins interacted with CNX when synthesized either in the absence or in the presence of tunicamycin. As mentioned above, prevention of N-glycosylation leads, in many cases, to protein aggregation within the ER and, as will be discussed further below, such aggregates do indeed interact with CNX (as with other ER-resident proteins), but in a non-specific and non-productive way. Interaction of CNX with glycoproteins secreted by the human hepatoma cells was transient, and retention half-times for different glycoproteins were correlated with their respective half-times of secretion. Moreover, incorporation of a Pro analogue into the hepatoma cell glycoproteins prevented their proper folding and prolonged their interaction with CNX. These results suggested, therefore, that the exclusive interaction of glycoproteins with CNX could be somehow involved in the quality control of glycoprotein folding.

Further work by Hammond et al. [65] showed that not only tunicamycin, but also inhibitors of GI and GII (but not of ER α-mannosidase I), prevented the interaction of glycoproteins with CNX. The G protein of the ts045 VSV mutant that, as mentioned above, was present at 39 °C in deglucosylated/monoglucosylated forms, interacted with CNX at that temperature. As the form that was precipitated with anti-CNX serum migrated in SDS-PAGE after jack-bean α-mannosidase treatment as having a larger size than that of the fully deglucosylated molecule, it was concluded that CNX specifically recognized monoglucosylated glycoproteins [the presence of glucose residue l](see Scheme 1) prevents full removal of α-linked Man units by α-mannosidase]. Further work confirmed this conclusion [66]. CRT, in addition to CNX, was also observed to interact with monoglucosylated glycoproteins [67–69]. From the long list of glycoproteins reported to date to associate with CNX/CRT, it may be concluded that all glycoproteins, irrespective of their final destination or their soluble or membrane-bound status, interact transiently with the ER lectins in mammalian cells. In agreement with the oligosaccharide processing pathway occurring in mammalian cells (Scheme 2B), in all cases tested the addition of GI/GII inhibitors (castanospermine, 1-deoxynojirimycin or its N-methyl or N-butyl derivatives) prevented lectin–glycoprotein association in vivo [67–80]. Accordingly, no CNX/CRT–glycoprotein interaction was observed in GI- or GII-deficient cell lines [73,81]. In addition, GI/GII inhibitors prevented the dissociation of already formed CNX/CRT–glycoprotein complexes [65,82,83].

© 2000 Biochemical Society
In contrast with what happens in mammalian systems, addition of 1-deoxynojirimycin to T. cruzi cells actually prolonged the CRT–cruzipain interaction [29,84]. This result agrees with known differences in N-oligosaccharide processing reactions occurring in mammalian and trypanosomatid cell ER (Schemes 2A and 2B).

**CNX AND CRT MAY RECOGNIZE DIFFERENT GLYCOPROTEINS IN VIVO**

Purified (or recombinant) immobilized CNX and CRT interact specifically with free monoglucosylated oligosaccharides when challenged with a mixture of tri-, di-, mono- and un-glucosylated glycans [29,85,86]. The strongest interaction was observed with Glc\(_2\)Man\(_\beta\)GlcNac\(_\beta\), and removal of mannose units from this species with a \(\alpha\)-mannosidase gradually reduced the binding capacity [86,87]. Nevertheless, Glc\(_3\)Man\(_\alpha\)GlcNac\(_\alpha\) (residues a-g and i in Scheme 1) still had 65\% of the binding capacity of the complete oligosaccharide. On the other hand, further removal of residue e from Glc\(_3\)Man\(_\alpha\)GlcNac\(_\alpha\) completely abolished glycan recognition by the lectins. No difference in glycan binding was observed between CNX and CRT [87]. The binding properties of CNX and CRT provide a rationale for the observation that the presence of monoglucosylated glycan is a sufficient condition for lectin interaction with isolated, properly folded glycoproteins [29,88,89]. Nevertheless, the presence of two glycans in the same glycoprotein appeared to be required for an efficient interaction, as judged by the co-immunoprecipitation technique commonly employed [88,90,91]. Milder, more sensitive techniques would be required to compare the relative lectin-association efficiencies of glycoproteins having different number of monoglucosylated glycans.

In spite of similar binding of CNX and CRT to monoglucosylated oligosaccharide observed in cell-free assays, the two lectins do not behave identically in vivo. Thus the patterns of glycoproteins precipitated with anti-CNX or anti-CRT antisera from lysed mammalian cells only partially overlapped [67]. For instance, CNX interacts with \(\alpha\), \(\beta\) and CD3-\(\gamma\),\(\delta\) and \(\epsilon\) T cell receptor subunits, but CRT only recognizes the first two polypeptides [92]. This difference might be related to the respective soluble and membrane-bound status of CRT and CNX, and to the relative positions of glycans in membrane glycoproteins. It may be speculated that oligosaccharides located in the proximity of the membrane would interact more easily with CNX, whereas those lumenerally oriented would preferentially associate with CRT. In fact, similar patterns of glycoproteins were found to interact with CRT and a truncated, soluble CNX fragment, or with the full-length version of CNX and CRT artificially anchored to the ER membrane by fusion with CNX or with an adenovirus glycoprotein transmembrane domain [70,92a]. Moreover, oligosaccharides located in the top/hinge domain of influenza virus haemagglutinin, i.e. in the more lumenerally oriented portion of the molecule, associated preferentially with CRT, whereas CNX was less discriminating, but mainly bound glycans close to the ER membrane [93].

An interesting case is that of the assembly of the human class I MHC. The heavy chain (a membrane glycoprotein) interacts first with CNX, but this interaction ceases upon association of the former protein with \(\beta\)_\(\alpha\)-microglobulin (a non-glycosylated protein). CRT then associates with the heavy chain, and this interaction persists during the rest of the assembly process, which involves transient interactions with other proteins, such as the transporters associated with antigen processing (TAP), and permanent association with a short, 8–10-amino-acid peptide [71,72,94–97]. The association with first CNX and then CRT would imply that the heavy-chain single oligosaccharide is first located close to the ER membrane, but that a change in the heavy-chain conformation resulting from \(\beta\)_\(\alpha\)-microglobulin binding makes it more accessible to soluble ER proteins such as CRT.

**MODELS PROPOSED FOR THE QUALITY CONTROL OF GLYCOPROTEIN FOLDING**

The following basic model for the quality control of glycoprotein folding was initially proposed by A. Helenius and co-workers [65]. Protein-linked monoglucosylated glycans formed by partial deglucosylation of the transferred oligosaccharide (Glc\(_3\)Man\(_\alpha\)GlcNac\(_\alpha\)) interact with CNX or CRT. The interaction would be disrupted by further GII-mediated deglucosylation, but GT would recreate monoglucosylated glycans in species that are not yet properly folded. A shuttle between glucosylated and un-glycosylated forms (i.e. a shuttle between CNX/CRT-bound and -unbound forms) would continue until the molecules adopt correctly folded structures (Figure 1A). It follows from this model that interactions between lectins and glycoproteins would exclusively retain conformers that are not correctly folded.

It is obvious that this is not the sole quality control mechanism of protein folding occurring in the ER. Other mechanisms, applicable to both glycoproteins and non-glycosylated proteins, must be operative in order to prevent secretion of all misfolded conformers. One of these mechanisms is based on the folding polypeptide forming reversible disulphide bonds with matrix proteins of the ER [98]. Additional mechanisms for retention in the ER of misfolded conformers were revealed by preventing the formation of monoglucosylated glycans: although the addition of GII/GII inhibitors resulted in an increase in the secretion of misfolded glycoproteins in CHO cells, the majority of misfolded conformers were still retained in the ER by associating with BiP, and probably also with other classical chaperones [99].

Another consequence of CNX/CRT–glycoprotein associations is a decrease in the folding rate, but an increase in folding efficiency and in the correct formation of both oligomeric structures and disulphide bonds. Prevention of the above-mentioned association by addition of GII/GII inhibitors to a rabbit reticulocyte/dog pancreas microsome system expressing influenza virus haemagglutinin doubled the rate of disulphide bond formation and protein oligomerization, but the overall efficiency of maturation decreased due to aggregation and degradation [100]. Similar results were obtained in CHO cells expressing the human insulin receptor to which GII/GII inhibitors were added [78]. Co-expression of CNX and class I MHC heavy-chain molecules in D. melanogaster cells enhanced the folding efficiency of the latter molecules, as judged by increased reactivity with conformation-dependent monoclonal antibodies [71].

The main obstacle to productive folding is aggregation; by binding oligosaccharides, CNX and CRT maintain glycoproteins in solution, thus allowing a functional interaction between the protein moieties of glycoproteins and classical ER chaperones, such as BiP/GRP68 [79,101], or with other folding-assisting proteins. Cross-linking experiments performed in intact cells showed that ERp57, an ER thiol:protein oxide reductase and isomerase, interacts exclusively with monoglucosylated forms of both membrane and soluble glycoproteins. An CNX/CRT–ERp57 association mediated this interaction [102–104]. The effect of addition of ERp57 on the proper folding of isolated bovine pancreas monoglucosylated RNase B was enhanced if CNX or CRT was added to the cell-free system. ERp57 associated
Figure 1  Models proposed for the quality control of glycoprotein folding

(A) ‘Lectin only’ model. Protein-linked Glc3Man9GlcNAc2 is partially deglucosylated to the monoglucosylated derivative by GI and GII, and this structure is recognized by CNX/CRT. Man9GlcNAc2 is glucosylated by GT if complete deglucosylation occurs before lectin binding. The glycoprotein is liberated from the CNX/CRT anchor by GII and reglucosylated by GT only if not properly folded. This process allows rebinding of the glycoprotein to the lectins. Upon adoption of the native tertiary structure, the glycoprotein is released from CNX/CRT by GII and not reglucosylated by GT.

(B) ‘Lectin first, followed by protein–protein interaction’ model. Binding of protein-linked Glc1Man9GlcNAc2 to the lectins is a necessary step for a subsequent protein–protein interaction between the glycoprotein and CNX/CRT. Release of the glycoprotein from the CNX/CRT anchor is a consequence of GII activity and a change in conformation in the glycoprotein, in CNX/CRT or in both. In both models CNX/CRT recruits ERp57 for proper disulphide-bond formation in folding glycoproteins.

An issue not yet settled is whether CNX and CRT behave exclusively as lectins, as proposed in the model of quality control of glycoprotein folding described above, or if the interaction of monoglucosylated glycans with CNX/CRT is the first and necessary step for a further interaction of the protein moieties of glycoproteins and the ER lectins. According to this second model, both CNX and CRT would behave first as lectins and then as classical chaperones (Figure 1B). Although in both models GT represents the element that senses glycoprotein conformation, in the second model the disruption of the CNX/
CRT–glycoprotein complex would result not only from GII activity, but also from a change in the protein conformation of one or both elements of the complex. Evidence supporting both models has been derived from experiments that have used a similar experimental approach: CNX/CRT–glycoprotein complexes were treated with glycosidases (GII, Endo H, N-Glycanase) which removed the whole or part of the oligosaccharides from glycoproteins. Disruption or persistence of the CNX/CRT–glycoprotein association after the enzymic treatment, as judged by the absence or presence of glycoproteins in immunocomplex precipitates obtained upon addition of anti-CNXCRT antisera, was taken as evidence for the ‘lectin only’ or the ‘lectin first and then chaperone’ models respectively. The main drawback of using this experimental approach for supporting the second model (Figure 1B) is that the presence of glycoproteins in the immunocomplex precipitates after partial or total removal of saccharide could be due, in cases in which CNX–membrane-glycoprotein associations were studied, to the persistence of the lectin and the glycoproteins in the same detergent micelles after glycosidase treatment [72,85,106,107]. In the case of a glycosidase-modified soluble glycoprotein, its presence in the immunoprecipitates may be ascribed to the poor solubility often observed for not yet properly folded glycoproteins [85]. No such drawbacks are encountered in experiments supporting the ‘lectin only’ model (Figure 1A) [29,88,89]. It has become evident that a different experimental approach must be used to support the ‘lectin first and then chaperone’ model. The possibility cannot be excluded, however, that the mechanism applicable may depend on the glycoprotein under study.

It has been reported that CRT and a truncated form of CNX behaved as classical chaperones in cell-free folding assays, with respect not only to glycoproteins devoid of monoglucosylated epitopes, but also to non-glycosylated proteins [108,109]. A similar role for CNX and CRT in vitro is doubtful, as it would contradict the well documented requirement for monoglucosylated oligosaccharides for interaction with CNX/CRT (see above). On the other hand, experiments performed in intact cells indicated that CNX co-immunoprecipitated with proteins in which N-glycosylation had been prevented by mutation of consensus sequences or by addition of tunicamycin [106,110]. A detailed study using VSV G glycoprotein having two, one or no N-glycosylation sites, expressed in a rabbit reticuloocyte/dog pancreas microsome system, strongly suggested that co-immunoprecipitations of CNX and non-glycosylated glycoproteins could represent non-specific associations between the lectins and the aggregates formed by many glycoproteins upon prevention of N-glycosylation [91]. Non-specific co-precipitation of misfolded glycoprotein aggregates with classical ER chaperones has been observed previously [111,112].

CNX/CRT-mediated facilitation of folding is not absolutely dependent on glucosylation/deglucosylation cycles catalysed by the opposing activities of GT and GII. Mutation of the Schiz. pombe gene coding for the dolichol-P-Glc-dependent glucosyltransferase that adds the first Glc unit to Man,GlcNAc,P-P-dolichol resulted in a mutant that synthesized underglycosylated proteins (as it transferred Man,GlcNAc in protein N-glycosylation) and which required GT-mediated glucosylation, but not GII-dependent deglucosylation, for survival under conditions of severe ER stress, such as high temperature [113]. In a different experimental approach, CHO cells were infected with the ts045 VSV mutant at the non-permissive temperature, thus yielding CNX-G protein complexes [82]. On lowering the temperature in the absence or presence of GII inhibitors (thus allowing or not the occurrence of glucosylation/deglucosylation cycles), G protein folded properly with the same efficiency under both conditions, although the folding rate was somewhat higher in the absence of the inhibitor.

**IMPORTANT OF FOLDING FACILITATION MEDIATED BY CNX/CRT–GLYCOPROTEIN INTERACTIONS FOR CELL VIABILITY**

The interaction of CNX/CRT with folding glycoproteins is not essential for cell viability under normal growth conditions. Mutation of GI/GII-encoding genes yielded mammalian or Schiz. pombe cells in which monoglucosylated oligosaccharide could not be formed by either of the possible pathways, which nevertheless not only were able to grow, but also did not show any discernable phenotype [6,114,115]. Moreover, a drastic decrease in the formation of monoglucosylated oligosaccharides caused by disruption of the Schiz. pombe GT-encoding gene did not affect cell growth [30]. As folding is an error-prone but essential process, cells have alternative ways for helping proteins to acquire the correct tertiary structures. When one folding facilitation system is absent, an alternative one carries out the task. For instance, about half of influenza virus haemagglutinin folded correctly when translated in a rabbit reticuloocyte/dog pancreas microsome system in the presence of GI/GII inhibitors [100]. Prevention of CNX/CRT–glycoprotein interaction leads to accumulation of misfolded glycoprotein in the ER, but this triggers the up-regulation of classical chaperones and folding-facilitating proteins (unfolded protein response) [6,73,116].

It has not yet been clearly established whether monoglucosylated oligosaccharides do indeed participate in facilitation of glycoprotein folding in *Saccharomyces cerevisiae*, as (a) this yeast is the only organism known so far that is devoid of GT [22,117]; (b) in contrast with what happens in mammalian and Schiz. pombe cells, disruption of the GII-encoding gene does not lead to the accumulation of misfolded glycoproteins in the ER [117,118]; and (c) *Sacch. cerevisiae* CNX presents significant structural variations when compared with its mammalian and Schiz. pombe counterparts [119] (neither *Sacch. cerevisiae* nor Schiz. pombe present CRT-encoding genes). It has been reported that *Sacch. cerevisiae* mutants in which Glc,GlcNAc is transferred in protein N-glycosylation and lack GII accumulated lower amounts of misfolded glycoproteins in the ER when subjected to an exogenoe stress (addition of dithiothreitol) than mutants that transfer Man,GlcNAc or Glc,GlcNAc and also lack GII, as judged by levels of BiP mRNA synthesized [117]. This would suggest that monoglucosylated proteins are somehow involved in the facilitation of glycoprotein folding.

On the other hand, as mentioned above, CNX–glycoprotein interactions were found to be essential for Schiz. pombe viability under conditions of severe ER stress, such as underglycosylation of glycoproteins caused by a mutation that determined Man,GlcNAc transfer in N-glycosylation, and high temperature [113]. The affected glycoprotein(s) was apparently involved in cell wall formation, as viability at high temperature could be regained not only upon transfection of double mutants lacking GT and transferring Man,GlcNAc, with a GT-encoding expression vector, but also by growing cells in a hyperosmotic medium (1 M sorbitol). Homozygous-null mice for calmein, a testis-specific CNX homologue, were nearly sterile, although they showed normal spermatogenesis and mating [120]. Their sperm did not adhere, however, to the extracellular matrix (zona pellucida), probably due to the total or partial absence of a docking glycoprotein at the plasma membrane. In addition, production of infectious HIV-1 and hepatitis virus B was prevented by addition of GI/GII inhibitors to the growth medium [121–123]. These results show, therefore, that in certain cases CNX/CRT-mediated facilitation of glycoprotein folding is indeed required for several cellular processes.
ARE OLIGOSACCHARIDE PROCESSING REACTIONS INVOLVED IN RECOGNITION OF PERMANENTLY MISFOLDED GLYCOPROTEINS?

As mentioned above, it has become evident in recent years that in both Sacch. cerevisiae and mammalian cells the 26 S cytotoxic proteasome is the main site of degradation of soluble and transmembrane misfolded glycoproteins retained in the ER [31,32]. Retrograde transport from the lumen of the ER to the cytosol and degradation involves the translocon heterotrimERIC Sec61p complex and the ubiquitin/proteasome system. Several proteasome inhibitors prevented ER-retained glycoprotein degradation and induced synthesis of mRNAs coding for ER chaperones, such as BiP, GRP94 and ERp72, a consequence of the accumulation of misfolded protein in the ER. This response was not triggered by inhibitors of non-proteasome proteases [124]. Glycoproteins are apparently deglycosylated in both yeast and mammalian cells by a neutral cytosolic N-glycanase activity prior to proteasome degradation [125–127]. This activity generates an N,N'-diacetylchitobiose structure at the reducing end of the oligosaccharides and a conversion of the Asn residue into Asp in the protein moiety. Free high-mannose-type oligosaccharides are then successively degraded in the mammalian cell cytosol by an Endo H-like activity and an α-mannosidase that yield ManGlcNAc2. This oligosaccharide is then degraded in the lysosomes [128].

Not all glycoproteins that fail to fold properly in the ER are degraded in the proteasomes. The N-terminal domain of the α repressor fused to the secreted glycoprotein invertase was expressed in Sacch. cerevisiae. Whereas the wild-type fusion protein was secreted, hybrids containing repressor mutants that were thermodynamically unstable were targeted to and further degraded in the vacuole (the yeast equivalent of the lysosome). Evidence was presented indicating that targeting to the vacuole was mediated by a Golgi-located receptor (Vps10p) [129,130]. The significance of the dual pathway for disposing of improperly folded proteins in Sacch. cerevisiae (proteasome or vacuolar degradation) is obscure for the moment, but it may be speculated that the latter pathway could be designed for proteins not showing gross folding defects that are able to be transported from the ER to the Golgi. Finally, in some cases incorrectly folded glycoproteins may not be degraded at all, but sorted to the ER to the Golgi. In hepatoma cells, of the intrinsically unstable TCR CD3 ε subunit in thymocytes, of a soluble variant of ribophorin I expressed in HeLa cells, of two mutant α2-plasmin inhibitors and of the yeast prepro-α-factor expressed in GH3 rat pituitary cells [135–139]. However, no effect of 1-deoxymannojirimycin was observed on the proteasome-mediated degradation of the TCR-α subunit (also intrinsically unstable) or of MHC class I heavy chain molecules synthesized in an assembly-defective cell line [135,140]. The apparently contradictory effects of the addition of ER α-mannosidase inhibitors on the degradation of misfolded glycoproteins may reflect the fact that although, as described above, removal of Man residues from N-oligosaccharides yields poorer substrates for both GII and GT, the influence of the Man content on the two activities is not identical. The possibility exists, therefore, that, depending on the relative amounts of these opposing enzyme activities and their differential specificities with respect to saccharides of varying Man content, inhibition of Man removal might favour or not the production of monoglucosylated glycans. Only in cases in which production of these species was enhanced (even only slightly) by ER α-mannosidase I inhibitors would binding to CNX/CRT delay translocation of misfolded glycoproteins to the cytosol.

The CNX–glycoprotein complex in misfolded glycoproteins was also suggested to be an element that is recognized by the degradation machinery. Thus it was reported that a misfolded α2-antitrypsin mutant interacted with CNX and specifically induced the polyubiquitination of the cytoplasmic tail of the lectin [141]. In addition, ubiquitinated, concanavalin A-reactive and CNX-bound apolipoprotein B (an intrinsically unstable glycoprotein) was detected in hepatoma cells [142]. Moreover, microsomes derived from a Sacch. cerevisiae mutant defective in Cne1p (the yeast CNX homologue) had a decreased ability to degrade prepro-α-factor in vitro [143]. This result apparently contradicts the fact that degradation of this same protein expressed in mammalian cells was accelerated upon addition of GI/GII inhibitors, i.e. by preventing the interaction between prepro-α-factor and CNX [137]. Similarly, hindrance of the interaction of class I MHC heavy chain with CNX accelerated its degradation [140]. Accordingly, a recent report showed that the rate of proteasomal degradation of a soluble variant of ribophorin I increased upon preventing its interaction with CNX by addition of the GI/GII inhibitor castanospermine [139]. Thus the notion that the CNX/CRT–glycoprotein complex could be one of the elements recognized by the degradation machinery contradicts substantial available evidence indicating that the interaction between the lectin and the folding glycoprotein prevents premature degradation of the latter species.
CONCLUDING REMARKS

The CNX/CRT-mediated retention and folding facilitation of misfolded glycoproteins constitutes a novel cellular mechanism in which recognition of protein conformation is performed by an enzyme (GT) that covalently tags not yet properly folded conformers. Recognition of this tag by the retention elements allows facilitation of folding by classical chaperones and, in the alternative model proposed, also by the said elements. Although basic knowledge on this novel mechanism is already available, there are certain aspects that merit further studies, for example the structural features of misfolded conformers that determine exclusive GT glucosylation, the importance of GT in multicellular organisms, the possible interaction of CNX/CRT with the protein moieties of folding glycoproteins, and the roles of Man removal and/or lectin–saccharide complexes in the disposal of permanently misfolded species. The qualification of ‘novel’ to the mechanism refers to its relatively recent description and not to its appearance in evolution, as it also occurs in organisms that, like trypanosomatid protozoa, diverged from the mammalian evolutionary line a long before plants and fungi.

I thank Ari Helenius and David B. Williams for communicating results before publication. Work in the author's laboratory was mainly supported by grants from the United Nations Development Program/World Bank/World Health Organization Program for Research and Training in Tropical Diseases, the National Institutes of Health (U.S.A.) and the Howard Hughes Medical Institute.

REFERENCES

6 Weng, S. and Spiro, R. G. (1993) Demonstration that a kifunin-resistant α-mannosidase with a processing action on N-linked oligosaccharides occurs in rat liver endoplasmic reticulum and various cultured cells. J. Biol. Chem. 268, 25656–25663
A. J. Parodi

12 A. J. Parodi

1. Introduction

2. Methods

3. Results

4. Discussion

5. Conclusion

6. References

7. Acknowledgments

8. Appendix

9. Supplemental Information

10. Appendix

11. Figure Legends

12. Table Legends

13. Abbreviations

14. Appendix

15. Figure Legends

16. Table Legends
of degradation of T cell receptor subunits from the endoplasmic reticulum (ER) in T cells: importance of oligosaccharide processing, ubiquitination, and proteasome-dependent removal from ER membranes. J. Exp. Med. 187, 835–846


140 Moore, S.E. and Spiro, R. G. (1993) Inhibition of glucose trimming by castanospermine results in rapid degradation of unassembled major histocompatibility complex class I molecules. J. Biol. Chem. 268, 3800–3812


142 Liao, W., Yeung, S.-C. J. and Chan, L. (1998) Proteasome-mediated degradation of apolipoprotein B targets both nascent peptides cotranslationally before translocation and full-length apolipoprotein B after translocation into the endoplasmic reticulum. J. Biol. Chem. 273, 27225–27230