The ER Glycoprotein Quality Control System

Selma Dejgaard1, Johan Nicolay1, Maryam Taheri1, David Y. Thomas2 and John J.M. Bergeron1*

1Department of Anatomy and Cell Biology, McGill University, 3640 University Street, Montreal, Quebec, H3A 2B2, Canada
2Department of Biochemistry, McGill University, Montreal, Quebec, Canada

Abstract

The endoplasmic reticulum (ER) is the major site for folding and sorting of newly synthesized secretory cargo proteins. One central regulator of this process is the quality control machinery, which retains and ultimately disposes of misfolded secretory proteins before they can exit the ER. The ER quality control process is highly effective and mutations in cargo molecules are linked to a variety of diseases. In mammalian cells, a large number of secretory proteins, whether membrane bound or soluble, are asparagine (N)-glycosylated. Recent attention has focused on a sugar transferase, UDP-Glucose: glycoprotein glucosyl transferase (UGGT), which is now recognized as a constituent of the ER quality control machinery. UGGT is capable of sensing the folding state of glycoproteins and attaches a single glucose residue to the Man9GlcNAc2 glycan of incompletely folded or misfolded glycoproteins. This enables misfolded glycoproteins to rebind calnexin and reenter productive folding cycles. Prolonging the time of glucose addition on misfolded glycoproteins ultimately results in either the proper folding of the glycoprotein or its presentation to an ER associated degradation machinery.

Introduction

UDP-Glucose: glycoprotein glucosyl transferase (UGGT) is a lumenal endoplasmic reticulum (ER) enzyme that plays a sensor role in a molecular machine known as the calnexin/calreticulin (CNX/CRT) cycle (Figure 1) (for recent reviews, Helenius et al., 1997; Zapun et al., 1999; Jakob et al., 2001b; Parodi, 2000). The principal constituents of this molecular machine are two lectins, the ER transmembrane protein calnexin (Bergeron et al., 1994) and its soluble luminal parologue calreticulin (Michalak et al., 1999), and the enzymes glucosidase II (Brada and Dubach, 1984; Trombetta et al., 1996) and UGGT (Sousa et al., 1992; Trombetta and Parodi, 1992). In the ER lumen, the asparagine (N)-linked oligosaccharides of newly synthesized glycoproteins undergo trimming by glucosidases (Brada and Dubach, 1984; Hettkamp et al., 1984; Trombetta et al., 1996) immediately after transfer of the core glycan, Glc3Man9GlcNAc2, to the asparagine residue within the consensus sequence Asn-X-Ser/Thr (where X is any amino acid except proline) of the growing polypeptide chain (Kornfeld and Kornfeld, 1985; Burda and Aebi, 1999). Membrane bound glucosidase I (Hettkamp et al., 1984) and the soluble glucosidase II (Brada and Dubach, 1984; Trombetta et al., 1996) remove the alpha 1,2-glucose and alpha 1,3-glucose residues, respectively. The sequential action of glucosidase I and II generates glycoproteins having the monoglucosylated oligosaccharide Glc3Man9GlcNAc2, which is a substrate for binding to CNX/CRT (Hammappala et al., 1994). Calnexin and calreticulin also interact with the PDI orthologue Erp57. This interaction assists in disulfide interchange of calnexin associated glycoproteins (Oliver et al., 1997; Zapun et al., 1998; Molinari and Helenius, 1999). Glucosidase II, apparently irrespective of the protein conformation, trims the last glucose residue on the oligosaccharide side chain of glycoproteins (Pelletier et al., 2000; Schrag et al., 2001; Zapun et al., 1997; Rodan, 1996), thus eliminating their recognition by calnexin or calreticulin. Conversely, UGGT can add back a single glucose unit from UDP-Glc in an α(1-3) bond to the terminal mannose of the α(1-3)-α(1-2) branch of Manx9GlcNAc2, restoring the monoglucosylated oligosaccharide molecule as a substrate for CNX/CRT interaction (Parodi, 2000; Rodan et al., 1996; Sousa and Parodi, 1995; Trombetta and Parodi, 1992; Wada et al., 1997; Zapun et al., 1997). The sum of the action of the three components is to act as a molecular chaperone to detain incompletely folded proteins in the ER. That UGGT was a component of such a molecular chaperone system became apparent following the discovery by Parodi that the enzyme targets denatured substrates which may be taken as surrogates for incorrectly folded proteins (Parodi, 2000). UGGT can sense and select incompletely folded glycoproteins for a further cycle of folding (Parodi, 2000; Zapun et al., 1997; Rodan et al., 1996; Sousa and Parodi, 1995; Wada et al., 1997), as well as discriminate among different nonnative conformers (Parodi, 2000; Sousa et al., 1992; Rodan et al., 1996; Sousa and Parodi, 1995). If the protein is folded, it is not reglucosylated and escapes this cycle, thereby releasing the glycoprotein to exit from the ER, after ER mannosidase processing. If secretory proteins fail to fold correctly in the ER, they are targeted to the ER-associated degradation (ERAD) machinery (Fewell et al., 2001; Hampton, 2002; Jarosch et al., 2002), also after ER mannosidase processing (Weng and Spiro, 1993; Jelinek-Kelly and Herscovich, 1988) and the downstream involvement of a novel lectin (Mn11p/ Htm1p / EDEM; Hosokawa et al., 2001; Jakob et al., 2001a; Nakatsukasa et al., 2001; Figure 1). Collectively, these mechanisms have been referred to as the ER quality control of protein folding and degradation (for recent reviews, Helenius et al., 1997; Jakob et al., 2001a; Parodi, 2000; Zapun et al., 1999). UGGT is the sole known constituent of the calnexin cycle to read the polypeptide code for folding and thereby distinguish
between productively folded and misfolded glycoproteins (Zapun et al., 1999). However, the molecular mechanisms by which this process happens remain unclear. The identification of UGGT isoforms without glucosyl transferase activity further adds to the complexity of the sensors of glycoprotein folding. Understanding the mechanism of glycoprotein recognition of by the UGGT family may be of relevance to protein trafficking diseases.

UDP-Glucose:glycoprotein glucosyl transferase

The enzyme activities have been partially characterized from Trypanosoma cruzi (Parodi and Cazzullo, 1982; Parodi, 2000), Schizosaccharomyces pombe (Fernandez et al., 1994), Drosophila melanogaster (Parker et al., 1995) and rat liver (Trombetta and Parodi, 1992). UGGT is ubiquitously expressed in the ER of most eukaryotic species (Trombetta et al., 1989) and has been sequenced from different sources (Arnold et al., 2000; Fernandez et al., 1994; Parker et al., 1995; Tessier et al., 2000). Recombinant rat (Tessier et al., 2000) and human (Arnold et al., 2000) UGGTs have been expressed in insect and mammalian cells, respectively. The rat enzyme is a large, soluble glycoprotein of 170 kD with an ER localization signal (variants of the His-Asp-Glu-Leu (HDEL) retention signal) at its C-terminus. It is present in the ER lumen (Parodi, 2000; Trombetta and Parodi, 1992) and in pre-GoGi intermediates (Zuber et al., 2001). The optimal enzyme activity is at neutral pH and is Ca^{2+}- or Mn^{2+}-dependent (for a recent review, Parodi, 2000).

Classification of known glycosyltransferases has been based on sequence homologies (Breton et al., 1998; Campbell et al., 1997). Recently, the D. melanogaster and S. pombe UGGT homologues (which are grouped with C. elegans 2, and the sequence related killer toxin-resistance protein Kre5p from S. cerevisiae (Meaden et al., 1990) into family 24 in ref. Campbell et al., 1997) were placed in the galactosyltransferase Family B by their structural homology to bacterial enzymes (the glycosyltransferases of family 8 in ref. Campbell et al., 1997) involved in lipopolysaccharide core biosynthesis (Breton et al., 1998). The full-length sequence alignments of known and putative UGGTs reveal a highly conserved 300 amino acid sequence (30% of the molecule, 60-70% identity) in the C-terminal domain (Breton et al., 1998; Tessier et al., 2000) (Figure 2). Limited but significant similarity exists between this highly conserved domain and several bacterial transferases that utilize UDP-Glc or UDP-Gal as a substrate donor (Breton et al., 1998). Therefore, the C-terminal domain of UGGT is responsible for recognition of the donor nucleotide-sugar and likely contains the catalytic domain: In this C-terminal region, two conserved motifs have been detected. DxD (x is any amino acid) is the most conserved motif observed in the galactosyltransferase Family B (Figure 2) which probably provides the binding site for the UDP-sugar (Tessier et al., 2000; Arnold et al., 2000) and DQDxxN which is probably involved in recognition of the N-acetylgalactosamine residue linked to the Asn residue of the glycoprotein substrate to which the sugar is transferred (Tessier et al., 2000; Arnold et al., 2000). The N-terminal domains of UGGTs reveal a lower degree of sequence similarity and have been proposed to be responsible for the recognition of protein conformations (Parodi, 2000; Guerin and Parodi, 2003).

Substrate recognition by UGGT

UGGT can recognize both the glycan and the protein moiety of incompletely folded glycoproteins, preferentially in molten globule-like conformers (Parodi, 2000; Caramelo et al., 2003; Sousa et al., 1992; Sousa and Parodi, 1995). The innermost GlcNAc unit of the glycoprotein's oligosaccharide is proposed to be required for UGGT recognition. Denatured nonglycosylated proteins do not affect UGGT activity, whereas denatured glycoproteins from which oligosaccharides have been removed by endo-β-N-acetylglucosaminidase H (Endo H) treatment (i.e. leaving a single GlcNAC-Asn) are efficient inhibitors of UGGT activity (Sousa and Parodi, 1995), but Endo H digested native glycoproteins with a remaining GlcNAC residue are not inhibitors of UGGT activity (Parodi, 2000; Sousa and Parodi, 1995). Thus, this innermost GlcNAC moiety must be covalently linked to a denatured glycoprotein. UGGT then recognize the covalently-linked Man2,GlcNAC2 denatured protein as substrate (Sousa and Parodi, 1995).

The exposed hydrophobic patches in incompletely folded glycoproteins, that would otherwise be hidden domains in native conformers, are recognized by the enzyme (Sousa et al., 1992; Sousa and Parodi, 1995). This is supported by observations demonstrating that interaction with immobilized hydrophobic stretches, but not with hydrophilic peptides, is inhibited by denatured but not by native glycoproteins (Sousa et al., 1992; Sousa and Parodi, 1995). Consequently, it has been proposed that UGGT senses the exposed hydrophobic residues in a way similar to many classical molecular chaperones (Sousa and Parodi, 1995).

To date, it has been unclear how these exposed hydrophobic stretches influence UGGT's ability to recognize and reglucosylate its substrates. However, hydrophobic residues on substrate glycoproteins may directly and/or with other molecular chaperones (BIP [binding protein, a member of the heat shock protein 70 family]) enhance the preferential selectivity of UGGT for incompletely folded glycoproteins (Caramelo et al., 2003; Taylor et al., 2003). In contrast to previous reports, it has recently been shown that short glycopeptides with different amino acid sequences are recognized by UGGT. Furthermore, hydrophobic residues close to N-linked glycan seem to be the main determinant for recognition by UGGT (Taylor et al., 2003). We speculate that these peptide-recognition elements may be in close proximity to glycans in folding intermediates (Taylor et al., 2003). Although both proteins recognize hydrophobic patches exposed during folding process, UGGT glucosylates glycoproteins preferentially in molten globule-like conformations, whereas BIP recognizes heptapeptides with large hydrophobic residues (Blond-Eiguindi et al., 1993) in an extended structure (Caramelo et al., 2003). It has consequently been speculated that BIP would interact with a relatively extended structure in early stage of glycoprotein folding rather than UGGT. This is consistent with other observations that
UGGT functions at later stages of glycoprotein folding (Parodi, 2000) and efficiently recognizes a variety of partially folded conformers (Trombetta and Helenius, 2000). Indeed, it has been shown that UGGT glucosylates the endogenous trypanosome substrate of UGGT, cruzipain (a lysosomal cysteine proteinase with two or three N-linked oligosaccharides and six or seven disulfide bridges) after it has obtained a tertiary structure that closely resembles the native conformer (Parodi, 2000). Since hydrophobic patches on glycoprotein substrates are introduced into a cleft in the bacterial homolog of BiP (DNaK) structure, it is then suggested that UGGT might require a larger surface interaction with its substrate glycoproteins for reglucosylation than BiP (Caramelo et al., 2003).

In addition, in glycoproteins with multiple independently folding domains, UGGT recognizes folding defects at the level of individual domains and only reglucosylates glycans in the misfolded domains (Ritter and Helenius, 2000). This may allow the CNX/CRT cycle to interact with only unfolded parts of the glycoproteins depending on distribution of...
Dejgaard et al. reported that glycosylation sites. Such recognition may be important for large molecules that have multiple independently folding domains (Ritter and Helenius, 2000). This then raises the more fundamental question that concerns the evolution of N-linked sites of glycosylation in mammalian glycoproteins. Have these sites evolved to identify locations in a polypeptide that are critical to UGGT-mediated presentation to the CNX/CRT cycle? The degree of conservation of most N-linked glycans in glycoproteins is high (Dwek et al., 2002; Rudd et al., 2001a; Rudd et al., 2001b; Rudd et al., 2001c). For viral envelope glycoproteins, all of which use the CNX/CRT cycle for productive folding, these conserved sites of N-linked glycosylation have evolved as essential for viral biogenesis (Rudd et al., 2001a; Rudd et al., 2001b; Rudd et al., 2001c).

Figure 2. Alignment of protein sequences from UGGTs in the C-terminal region of the protein. Although the cDNA sequence encoding UGGT protein from different sources shows a high degree of sequence homology extending over the entire sequence of the protein, it is particularly high in the C-terminal region of the protein. The color key is: red box, white letter for strict identity, red character for similarity in a group, blue frame filled in yellow shows similarity across groups. The boundary of the proposed catalytic domain (is shown with ▲) identified for H. sapiens1 (Arnold et al., 2000) is about 300 residues from the C-terminus. Highly conserved catalytic domain motifs (in H. sapiens UGGT 1 sequence, residues 1357-1360 and 1452-1457 respectively) in all UGGTs except S. cerevisiae Kre5p are shown with ▲, in green box, white letter. Conserved residues in all homologues but both H. sapiens UGGT 2 (Arnold et al., 2000) and S. cerevisiae (except the first residue) are denoted with blue ▼ under the sequences (in HUGGT1 sequence, I 1256, K 1331, G 1409, H 1461). Other candidate residues that differ from the known active UGGTs (H. sapiens1, R. norvegicus, D. melanogaster, S. pombe UGGT proteins) for the inactive UGGT sequences (H. sapiens2 and S. cerevisiae proteins) are donated with red ▼ under the sequences (in HUGGT1 sequence, N 1311, T 1365).
et al., 2001c). Hence, evolution may have selected the polypeptide domains within which a subset of N-glycans is found for presentation to UGGT as part of the protein folding code required for correct glycoprotein folding in the calnexin cycle.

UGGT in quality control, degradation and sorting

Quality control
Calnexin and/or calreticulin bind transiently to almost all soluble and membrane bound glycoproteins during folding or oligomeric assembly of the glycoproteins in the ER (Hammond et al., 1994; Helenius et al., 1997; Ou et al., 1993). Trypanosomatid protozoa cells express calreticulin but lack calnexin and in these cells Man$_{6}$GlcNAc$_{2}$ or Man$_{9}$GlcNAc$_{2}$ is transferred to the nascent polypeptide chain by oligosaccharyltransferase. Thus, in these cells creation of a Glc$_{3}$Man$_{9}$GlcNAc$_{2}$ glycoprotein that can bind calreticulin is only via the action of UGGT (Labriola et al., 1999; Parodi and Cazzulo, 1982). In contrast, all mammalian species transfer triligosylated (i.e. Glc$_{1}$Man$_{3}$GlcNAc$_{2}$) structures to nascent polypeptide chains and also possess UGGT and thus there is a dual route of presentation to CNX/CRT either via the sequential activity of glucosidase I and II or by reglucosylation by UGGT (Parodi, 2000). This difference between cells from various trypanosome species and mammalian cells may predict regulation of entry into the CNX/CRT cycle at the level of substrates presented to the oligosaccharyltransferase. Such a regulation of dolichol-based intermediates in the biosynthetic pathway of Glc$_{3}$Man$_{9}$GlcNAc$_{2}$ has recently been observed in normal primary cells undergoing a stress response (Doerrler and Lehrman, 1999), although its precise physiological significance remains to be established. Misfolded glycoproteins enter a cycle of binding to and release from CNX/CRT mediated by the sequential actions of UGGT and glucosidase II as originally proposed by Hammond et al (Hammond et al., 1994; Parodi, 2000; Trombetta and Parodi, 1992). This ultimately results in either the proper folding of the glycoprotein or its disposal (Fewell et al., 2001; Helenius et al., 1997; Jarosch et al., 2002; Ou et al., 1993; Parodi, 2000). In this context, the CNX/CRT cycle acts as a kinetic trap, retaining conformers that are not correctly folded, and allowing only the native conformers to move further along the secretory pathway (Labriola et al., 1999; Le et al., 1994; Ou et al., 1993; Pind et al., 1994). In support of this suggestion is the observation that the ER retention half-times of glycoproteins correlate with half times of their secretion, as well as their rank order of calnexin binding (Lodish et al., 1983; Ou et al., 1993). The demonstration in mammalian cells that UGGT activity can transform intermolecular oxidized aggregates of misfolded transferrin, a secretory glycoprotein of the liver, into monomeric productively folded transferrin is good evidence that reglucosylation by UGGT has a central role in quality control in vivo (Wada et al., 1997).

Most of the glycoproteins are reglucosylated during their maturation in the ER, including influenza HA (Hebert et al., 1995), vesicular somatitis virus G protein (Suh et al., 1989), transferrin (Wada et al., 1997), T cell receptor subunits (Van Leeuwen and Kearse, 1997) and cruzipain (Labriola et al., 1999). Reglucosylation may mediate the selection of the chaperones in vivo (Trombetta and Helenius, 2000). For example, vesicular somatitis virus glycoprotein binds to BiP during its early stage of folding, and later it is possible that reglucosylation mediates its binding to calnexin (Hammond et al., 1994). Recently, it has been speculated (Caramelo et al., 2003) that the preferential recognition of the glycoprotein substrates by UGGT and BiP might provide a molecular rationale for sequential interaction between BiP and CNX/CRT with folding glycoproteins (Hammond et al., 1994; Kim and Arvan, 1995; Molinari and Helenius, 2000).

In vitro, both UGGT and glucosidase II showed the highest relative rate of glucosylation/deglucosylation for glycoproteins containing Man$_{5}$GlcNAc$_{2}$ (Grinnia and Robbins, 1980; Sousa et al., 1992). Relative glucosylation rates by UGGT were respectively 100, 50, and 15 for Man$_{9}$, Man$_{8}$, and Man$_{7}$ (Parodi, 2000; Sousa and Parodi, 1995). Similar to UGGT, glucosidase II also revealed reduced deglucosylation rates for substrate glycoproteins upon the removal of mannose residues (Grinnia and Robbins, 1980). Hence these data led to a proposal that quality control in the ER is also regulated by demannosylation (Cabral et al., 2001).

ER-associated degradation

ER-associated degradation has a central clearance function in the cell (Fewell et al., 2001; Hampton, 2002; Jarosch et al., 2002). In this pathway, terminally misfolded proteins are subjected to trimming by ER α1,2-mannosidase(s) (Jelinek-Kelly and Herscovics, 1988; Weng and Spiro, 1993) and are transported to the cytosol via the Sec61p translocon complex (Pilon et al., 1997; Wiertz et al., 1996). They are then polyubiquitinated and proteolytically degraded by the cytosolic 26S proteasome. Ubiquitination appears to be required for both retrotranslocation to the cytosol and proteasomal degradation (see in reviews, Fewell et al., 2001; Hampton, 2002; Jarosch et al., 2002).

The relatively slow-acting ER α1,2-mannosidase is responsible for mannose trimming of the oligosaccharide side chain, irrespective of the protein conformation, generating primarily Man$_{6}$GlcNAc$_{2}$ isomer B in mammalian cells (Weng and Spiro, 1993), and only Man$_{6}$GlcNAc$_{2}$ isomer B in S. cerevisiae (Jelinek-Kelly and Herscovics, 1988). The resulting Man$_{6}$ structure can be recognized by an enzymatically inactive mannosidase I, called Mm11p (mannosidase-like protein; Nakatsukasa et al., 2001) or Htm1p (homologous to mannosidase I; Jakob et al., 2001a) in yeast and EDEM (ER Degradation Enhancing α-Mannosidase-like protein; Hosokawa et al., 2001) in mammalian cells. EDEM seems to be up-regulated by the unfolded protein response through the XBP1 pathway (Yoshida et al., 2003) and to target the misfolded protein for retrotranslocation and degradation by promoting release from the calnexin (Molinari et al., 2003; Oda et al., 2003). Calnexin, BiP and PDI have also been proposed to work in the recognition phase of misfolded glycoproteins for subsequent retrotranslocation and degradation (Molinari et al., 2002). Other factors may participate in the targeting
### Table 1. List of diseases and ER quality control implications*

<table>
<thead>
<tr>
<th>Disease</th>
<th>Protein</th>
<th>Glycoprotein</th>
<th>CNX/CRT Assoc.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>I: Loss of coupling to ER export leading to degradation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cystic fibrosis (Find et al., 1994)</td>
<td>Cystic fibrosis transmembrane regulator</td>
<td>+</td>
<td>CNX, Prolonged Assoc.</td>
</tr>
<tr>
<td>Hereditary emphysema (Le et al., 1994; Spiro et al., 1996)</td>
<td>cf1-Antitrypsin (non PiZ variants)</td>
<td>+</td>
<td>CNX, Prolonged Assoc.</td>
</tr>
<tr>
<td>Hereditary hemochromatosis (Kühn, 1999)</td>
<td>HFE. Loss of binding transferrin receptor</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Protein C deficiency (Tokunaga et al., 2000)</td>
<td>Protein C</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Type I hereditary angioedema (Verpy et al., 1993)</td>
<td>Complement C1 inhibitor</td>
<td>+</td>
<td>CNX</td>
</tr>
<tr>
<td>Tay-Sachs (Kaback and Desnick, 2001)</td>
<td>β-Hexosaminidase</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Congenital sucrase-isomaltase deficiency (Naim et al., 1988)</td>
<td>Sucrase-isomaltase</td>
<td>+</td>
<td>CNX</td>
</tr>
<tr>
<td>Crigler-Najjar type II (Sampietro and Iolascon, 1999)</td>
<td>Bilirubin-UDP-glucuronosyltransferase 1</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Polyendocrinopathy/Hypertinsulinaemia (Remik and Fricker, 2001)</td>
<td>Carboxypeptidase E</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Diabetes mellitus (Ibs et al., 1998)</td>
<td>Insulin receptor</td>
<td>+</td>
<td>CNX/CRT, Prolonged Assoc.</td>
</tr>
<tr>
<td>Linon syndrome (Ansems et al., 1991)</td>
<td>Growth hormone receptor</td>
<td>+</td>
<td>CNX</td>
</tr>
<tr>
<td>Hereditary molycoexpeptide (Nauseef, 1999; Nauseef et al., 1998)</td>
<td>Myeloperoxidase</td>
<td>+</td>
<td>CNX/CRT, Prolonged Assoc.</td>
</tr>
<tr>
<td>Primary hyperparathyroidism (Garfield and Karagulis, 2001)</td>
<td>Preparathyroid hormone</td>
<td>+</td>
<td>CNX</td>
</tr>
<tr>
<td>Oculudentaneous albinism (Halaban et al., 1997; Halaban et al., 2002;</td>
<td>Tyrosinase</td>
<td>+</td>
<td>CNX, Prolonged Assoc.</td>
</tr>
<tr>
<td>Vainagamousorthy and Rajakumar, 1996)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fabry disease (ishi et al., 2000)</td>
<td>α-D-galactosidase</td>
<td>+</td>
<td>CNX transient assoc.</td>
</tr>
<tr>
<td>Congenital long QT syndrome (Deutsch, 2002; Furutani et al., 1999)</td>
<td>Voltage gated potassium channel (HERG)</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Autosomal dominant retinitis pigmentosa (Frederick et al., 2001;</td>
<td>Rhodopsin</td>
<td>+</td>
<td>CNX transient assoc.</td>
</tr>
<tr>
<td>Salita et al., 2002; Illing et al., 2002)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>II: Loss of coupling to ER export leading to accumulation in the ER</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liver diseases/Hereditary emphysema (Qu et al., 1997;</td>
<td>cf1-Antitrypsin (PiZ variants)</td>
<td>+</td>
<td>CNX</td>
</tr>
<tr>
<td>Coakley et al., 2001)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congenital hypothyroidism/related disorders (Kim and Arvan, 1998)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thyroglobulin deficiency (Kim and Arvan, 1995)</td>
<td>Thyroglobulin</td>
<td>+</td>
<td>CNX/CRT, Prolonged Assoc.</td>
</tr>
<tr>
<td>Thyroid peroxidase deficiency (de Carvalho et al., 1994;</td>
<td>Thyroid peroxidase</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Kim and Arvan, 1995; Fayedat et al., 2000)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thyroxin-binding globulin deficiency</td>
<td>Thyroxin-binding globulin</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>(Miura et al., 1994; Relfoff et al., 1996)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Osteogenesis imperfecta (Lamande and Bateman, 1999)</td>
<td>Type I procollagen</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Hereditary hypofibrinogenemia (Roy et al., 1996)</td>
<td>Fibrinogen</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>α1-Antichymotrypsin (ACT) deficiency (Callea et al., 1992)</td>
<td>α1-Antichymotrypsin</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Neurophyseal diseases insipidus (Morello et al., 2001)</td>
<td>Vasopressin precursor protein</td>
<td>+</td>
<td>CNX, Prolonged Assoc.</td>
</tr>
<tr>
<td>Nephrogenic diabetes insipidus (Tamarappoo et al., 1999)</td>
<td>Aquaporin II</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Charcot-Marie-Tooth disease (Thomas, 1999; Mendell, 1998)</td>
<td>Peripheral myelin protein 22</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Pelizaeus-Merzbacher disease (Yool et al., 2000; Swanton et al., 2003)</td>
<td>Perioendrinoprotein</td>
<td>-</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Althemer disease (Shastry, 2001)</td>
<td>Presenilin</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Strassler-Scheinker syndrome (Rudd et al., 2001b; Collins et al., 2001)</td>
<td>Prion protein processing defect</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Hereditary Creutzfeld-Jacob disease (Rudd et al., 2001b;</td>
<td>Prion protein processing defect</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Collins et al., 2001)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>von Willebrand Disease</td>
<td>von Willebrand factor (VWF)</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Type IIA (Englender et al., 1996; Lyons et al., 1992)</td>
<td>VWF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Types I and III (Allen et al., 2001)</td>
<td>Types I and III associated VWF variant</td>
<td>+</td>
<td>CNX/CRT, CNX, Prolonged Assoc.</td>
</tr>
<tr>
<td><strong>III: Defective transport machinery</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined factorsV and VIII deficiency (Nichols et al., 1998)</td>
<td>Ergic-53</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Spondyloepiphyseal dysplasia tarda (Gedeon et al., 1999)</td>
<td>SELD (sedlin)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Viral Infections: Selected examples that are known to be associated with CNX/CRT quality control</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIDS (Land and Braakman, 2001)</td>
<td>160/120</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Herpes simplex-I (Yamashita et al., 1996b)</td>
<td>B, C and D</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Cytomegalovirus disease (Yamashita et al., 1996a)</td>
<td>B</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Fluenceza (Hebert et al., 1996)</td>
<td>Haemagglutinin</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Hepatitis B (Prange et al., 1999)</td>
<td>M</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Hepatitis C (Chouki et al., 1998)</td>
<td>E1 and E2</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Rubella (Nakhashi et al., 2003)</td>
<td>E1 and E2</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Measles (Boit, 2001)</td>
<td>Haemagglutinin/Fusion</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Newcastle disease (McGimmies and Morrison, 1998)</td>
<td>Haemagglutinin-neuraminidase</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Dengue hemorrhagic fever (Wu et al., 2002)</td>
<td>M, E and NS1</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Japanese encephalitis (Wu et al., 2002)</td>
<td>M, E and NS1</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Unkiniemi virus infection (Veijola and Pettersson, 1999)</td>
<td>G</td>
<td>+</td>
<td>CNX/CRT</td>
</tr>
<tr>
<td>Vesicular somatitis (Cannon et al., 1996)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Unknown consequences</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polycystic liver disease (PCLD, OMIM 174050) (Dreht et al., 2003)</td>
<td>Hepatocystin (also identifiented as the β-subunit of glucosidase II)</td>
<td>+</td>
<td>CNX/CRT, CYCLE constituent</td>
</tr>
<tr>
<td>Congenital disorders of N-glycosylation (CDG)-II b (De Praeter et al., 2000)</td>
<td>Glucosidase I and II</td>
<td>+</td>
<td>CNX/CRT, CYCLE constituent</td>
</tr>
</tbody>
</table>

* Please see Aridor and Hannan, 2002 for a more complete list of the ER quality control related diseases. § placed in the first group in Aridor and Hannan, 2000.  ¶ mutant protein may alter the processing of oligosaccharide chains of various glycoproteins. ¤ where the defects are in the trimming and modification of the core oligosaccharide which had already been transferred to the target proteins.
of misfolded glycoproteins to ERAD (Fewell et al., 2001; Hampton, 2002; Jarosch et al., 2002). For example, the AAA ATPase family members Cdc48 in yeast and p97 in mammals have been shown to be required for the transport of misfolded proteins to the cytosol (Braun et al., 2002; Ye et al., 2001). As well, ubiquitin ligases (the F-box protein Fbx2 of an ubiquitin ligase complex, SCF (Fbx2) [Yoshida et al., 2002], a chaperone-containing ligase, CHIP and E2 [Meacham et al., 2001]) have been shown to participate in ERAD of misfolded proteins, suggesting potential links between ubiquitination, ERAD and quality control of glycoproteins (Fewell et al., 2001; Hampton, 2002; Jarosch et al., 2002). Indeed Qu et al. (Qu et al., 1996) have proposed that calnexin itself is polyubiquitinated during ERAD, but this observation has not been confirmed.

Retention, retrieval and ER-associated degradation

In cells which overexpress a temperature sensitive mutant of vesicular somatitis virus G protein, the mutant protein escapes initial ER retention but is retrieved back to the ER bound to the molecular chaperone BiP from the intermediate compartment (IC) and the cis-Golgi network (Hammond et al., 1994; Hsu et al., 1991). However, the misfolded VSV G mutant protein at the ‘exit sites’ in the ER is reglucosylated by UGGT and returned to the ER instead of being transported to the Golgi complex (Mezzacasa and Helenius, 2002). Proteins that are localized in the ER possess retention and retrieval signals, including specific C-terminal motifs, such as Lys-Asp-Glu-Leu (KDEL) for soluble proteins (Pelham, 1996), or dilysine localized in the ER possess retention and retrieval signals, ER is reglucosylated by UGGT and returned to the ER.

The enzymatically “active” and “silent” UGGTs

Recently, UGGT was also identified as a part of the heavy chain-BiP complex including molecular chaperones and folding enzymes BiP, Glucose-regulated protein (GRP)94 (Endoplasmim), GRP170 (an ER heat shock protein 70 family member), an ER Hsp40 cochaperone (ERdj3), and several PDI family members (PDI, ERP72, CaBP1), cyclophilin B (an ER immunophilin protein) and the SDF2-L1 (an ER stress inducible protein; Meunier et al., 2002).

The existence of such a network(s) (Kim and Arvan, 1995; Kuznetsov et al., 1994; Kuznetsov et al., 1997; Tatu and Helenius, 1997) led to a proposal that the ER is organized into different networks containing distinct pool of the ER chaperones (Meunier et al., 2002). This might also explain the retention of some molecular chaperone without KDEL sequences in the ER (see in ref. Meunier et al., 2002).

Such a complex may also explain the sequential/simultaneous interactions of the molecular chaperones with misfolded proteins (as mentioned above) (Hammond et al., 1994). Furthermore, these data suggest that UGGT may well be a part of dynamic molecular chaperone complex, which may also determine the sorting and retrieval of secretory proteins. Such a dynamic network(s) of chaperones could prevent the forward movement of misfolded proteins by their retention (see in ref. Hendershot, 2000). This finding does not rule out the possibility that there are other pools of UGGT not present in the complex (Meunier et al., 2002).

UGGT has also been shown to associate with a misfolded variant of α1-antitrypsin (non PiZ; Choudhury et al., 1997), ER resident enzymes such as the folding enzyme protein disulfide isomerase (PDI), the chaperone BiP, and carboxylesterase (a specific quality control factor which limits ER export of C-reactive protein; Amouzadeh et al., 1997). The enzyme has also been observed in a complex with the selenoprotein, Sep15 (Korotkov et al., 2001), which is suggested to play a role in cancer etiology (see in ref. Korotkov et al., 2001). It is speculated that Sep15 may play a role in redox reactions in the complex, which would then have an affect on CNX/CRT-mediated folding. However, UGGT is detected in both selenoprotein-bound and selenoprotein-free forms (Korotkov et al., 2001). The physiological significance of UGGT in a complex with selenoproteins and/or other proteins (Amouzadeh et al., 1997; Choudhury et al., 1997; Korotkov et al., 2001; Meunier et al., 2002) is yet unclear and is fully functional as a glycosyltransferase in vitro in the absence of other proteins.

Two UGGT family members are apparently catalytically inactive: one in human, hUGGT2 (Arnold et al., 2000) and one in S. cerevisiae, Kre5p (Meaden et al., 1990; Figure 2). In S. cerevisiae the loss of enzyme activity might be due to the lack of conservation of critical D residues in the C-terminal catalytic domains of these enzymes (Tessier et al., 2000) (green boxes, shown with in Figure 2).
However, the conservation of these same motifs in the second catalytically inactive UGGT (hUGGT2) sequence suggests additional requirements for the enzyme activity. Comparisons of hUGGT1 and hUGGT2 protein sequences revealed differing residues in the catalytic domain (as identified by a blue ★ under the sequences in Figure 2); these residues may coincide with the loss of activity. Interestingly, three of these four residues also varied in the S. cerevisiae Kre5p sequence compared to all other homologues. Additionally, there is divergence within the N-terminus region of hUGGT1 and hUGGT2 which may affect their substrate specificity (Arnold et al., 2000). As well, except first residue, all of these residues differ from all other homologues proteins with S. cerevisiae Kre5p sequence. Furthermore, we compared the known inactive UGGT sequences (hUGGT2 and Kre5p) to the known active UGGT sequences (H. sapiens UGGT 1, R. norvegicus, D. melanogaster, and S. pombe UGGT proteins), revealing other candidate residues in the catalytic domain (as identified by a red ★ under the sequences in Figure 2).

Remarkably, the catalytically inactive S. cerevisiae gene is essential for cell viability (for a recent review, (Parodi, 2000) and appears to function early in the (1,6)-β-D-glucan synthesis pathway (Meaden et al., 1999). Higher eukaryotes (e.g. worms, rodents and humans) have evolved two UGGT genes with only one predicted to be catalytically active. An exciting possibility is that the catalytically inactive variants of UGGT including Kre5p are required for the disaggregation of misfolded proteins. At least four possible scenarios for these variants have also been proposed (Arnold et al., 2000) including different substrate specificity, a nucleotide sugar donor other than UDP-glucose, targeting a substrate for degradation, or binding of catalytically inactive hUGGT 2 to an unknown protein partner which would then elicit a gain of UGGT enzyme activity. Since, the enzymatically inactive EDEM (Hosokawa et al., 2001; Jakob et al., 2001a; Nakatsukasa et al., 2001) has been linked to quality control (Molinari et al., 2003; Oda et al., 2003) and ERP57 displays increased isomerase activity when associated with calnexin (Oliver et al., 1997; Zapun et al., 1998), this then raises the possibility of a role for the enzymatically "silent" UGGTs in quality control.

Quality control implications for diseases

Many human diseases can be classified as "protein trafficking diseases" where mutant secretory proteins are subjected to the ER quality control system and its associated ERAD (Amara et al., 1992; Aridor and Balch, 1999; Kim and Arvan, 1998; Kopito, 1999; Olkkonen and Ikonen, 2000; Thomas et al., 1995) (Table 1). These can be divided into three groups (I, II, and III, Table 1). The first group of diseases corresponds to loss of coupling to the ER export machinery leading to degradation of misfolded proteins. The second group of diseases correlates with ER accumulation of mutant proteins that are uncoupled from the ER export machinery and fail to be degraded and forms aggregates in the ER. The third group of diseases is due to defects in the machinery required for transport from the ER to the Golgi complex (Aridor and Balch, 1999; Aridor and Hannan, 2000). In addition, viral and bacterial pathogens manipulate ER function for their immunological survival (Land and Braakman, 2001; Ploegh, 1998; Rust et al., 2001; Yamashita et al., 1998a) or to deliver their toxic products to the cytosol (Lord and Roberts, 1998). Viral infections are often linked to massive production of viral proteins and their accumulation in the ER (Aridor and Balch, 1999; Ploegh, 1998). Mutant glycoproteins associated with protein trafficking diseases are also shown in Table 1 (Amara et al., 1992; Aridor and Balch, 1999; Kopito, 1999). The majorities of these proteins interact with calnexin and/or calreticulin and therefore are also potential substrates for UGGT. It is inferred that UGGT and its reglucosylation of mutant proteins is responsible for re-binding to CNX/CRT (for example, vesicular stomatitis virus G (Peterson and Helenius, 1999), hemagglutinin (Peterson and Helenius, 1999), α1-antitrypsin (Choudhury, 1997), thyroglobulin (Parker et al., 1995) (Table 1), subunits of the T cell receptor (Gardner and Kearse, 1999).

From a therapeutic perspective, the problem of inherited protein misfolding is being addressed by several related strategies: One approach is based on attempts to chaperone misfolded proteins into a native-like structure that can evade the ER quality control machinery using either biological or chemical chaperones (Chow et al., 2001; Loo and Clarke, 1997). In another approach, to the accumulation of a mutant variant (PiZ) that is mostly retained in the ER (Qu et al., 1997) osmotolytes such as trimethylamine-N-oxide and sarcosine significantly reduce the rate of α1-antitrypsin mutant polymerization with no effect on the normal inhibitory function of α1-antitrypsin for serine proteases (Chow et al., 2001).

The other approach focuses on circumventing the quality control machinery of the ER, allowing proteins to be secreted, irrespective of their structural abnormalities (Burrows et al., 2000; Choo-Kang and Zeiitlin, 2001; Rubenstein et al., 1997). The relevant example of this approach is the transmembrane conductance regulator protein (CFTR). The CFTR ΔF508 mutant is the most common cystic fibrosis allele and the mutated but otherwise functionally active protein is thus retained in the ER and eventually targeted for degradation, rather than being transported to the plasma membrane (Kopito and Ron, 2000; Kopito, 1999). The use of 4-phenylbutyrate (4PBA) increases the expression of the ΔF508-CFTR mutant to the plasma membrane (Rubenstein et al., 1997), possibly acting through a cytosolic molecular chaperone, Hsp70 (Choo-Kang and Zeiitlin, 2001). The use of specific mannannosidase inhibitors may also be relevant to overcome the increased degradation and mislocation of mutant α1-antitrypsin (Marcus and Perlmutter, 2000).

Competitive or noncompetitive inhibition of UGGT, leading to inhibition of the CNX/CRT cycle, may be an alternative approach to development of therapies for protein misfolding diseases (see ref. Kopito and Ron, 2000). Although no specific inhibitors of this enzyme are known, related inhibitors may prove valuable to dissect a UGGT link to the diseases listed in Table 1 (Block and Jordan, 2001; Dwak et al., 2002). It is expected that mutation or
removal of UGGT genes as well as components of the CNX/CRT cycle in mouse models of the diseases indicated in Table 1 will provide further insight into new targets.

Acknowledgements

We thank Pamela Cameron, Eric Chevet, Kurt Deiglaard, Ali Fazel, Annalyn Gilchrist, Michael Jain, and Hugues Nicolay for their kind help. Special thanks to Inci Ozer for her support. We apologize to colleagues whose original work we have not cited. Supported by grants to JJMB and DYT from the Canadian Institutes of Health Research.

References


Trombetta, E.S., J.F. Simons, and A. Helenius. 1996. Endoplasmic reticulum glucosidase II is composed of a
catalytic subunit, conserved from yeast to mammals, and a tightly bound noncatalytic HDEL-containing subunit. J. Biol. Chem. 271:27509-27516.


Further Reading

Caister Academic Press is a leading academic publisher of advanced texts in microbiology, molecular biology and medical research. Full details of all our publications at caister.com

• MALDI-TOF Mass Spectrometry in Microbiology
  Edited by: M Kostrzewa, S Schubert (2016)
  www.caister.com/malditof

• Aspergillus and Penicillium in the Post-genomic Era
  Edited by: RP Vries, IB Gelber, MR Andersen (2016)
  www.caister.com/aspergillus2

• The Bacteriocins: Current Knowledge and Future Prospects
  Edited by: RL Dorfl, SM Roy, MA Riley (2016)
  www.caister.com/bacteriocins

• Omics in Plant Disease Resistance
  Edited by: V Bhadauria (2016)
  www.caister.com/opdr

• Acidophiles: Life in Extremely Acidic Environments
  Edited by: R Quatrini, DB Johnson (2016)
  www.caister.com/acidophiles

• Climate Change and Microbial Ecology: Current Research and Future Trends
  Edited by: J Marxsen (2016)
  www.caister.com/climate

• Biofilms in Bioremediation: Current Research and Emerging Technologies
  Edited by: G Lear (2016)
  www.caister.com/biorem

• Microalgae: Current Research and Applications
  Edited by: MN Tsaoiloglu (2016)
  www.caister.com/microalgae

• Gas Plasma Sterilization in Microbiology: Theory, Applications, Pitfalls and New Perspectives
  Edited by: H Shintani, A Sakudo (2016)
  www.caister.com/gasplasma

• Virus Evolution: Current Research and Future Directions
  Edited by: SC Weaver, M Denison, M Roossinck, et al. (2016)
  www.caister.com/viruserv

• Arboviruses: Molecular Biology, Evolution and Control
  Edited by: N Vasilakis, DJ Gubler (2016)
  www.caister.com/arbo

• Shigella: Molecular and Cellular Biology
  Edited by: WD Picking, WL Picking (2016)
  www.caister.com/shigella

• Aquatic Biofilms: Ecology, Water Quality and Wastewater Treatment
  Edited by: AM Romani, H Guasch, MD Balaguer (2016)
  www.caister.com/aquaticbiofilms

• Alphaviruses: Current Biology
  Edited by: S Mahalingam, L Herrero, B Herring (2016)
  www.caister.com/alpha

• Thermophilic Microorganisms
  Edited by: F Li (2015)
  www.caister.com/thermophile

• Flow Cytometry in Microbiology: Technology and Applications
  Edited by: MG Wilkinson (2015)
  www.caister.com/flow

• Probiotics and Prebiotics: Current Research and Future Trends
  Edited by: K Venema, AP Carmo (2015)
  www.caister.com/probiotics

• Epigenetics: Current Research and Emerging Trends
  Edited by: BP Chadwick (2015)
  www.caister.com/epigenetics2015

• Corynebacterium glutamicum: From Systems Biology to Biotechnological Applications
  Edited by: A Burkovski (2015)
  www.caister.com/cory2

• Advanced Vaccine Research Methods for the Decade of Vaccines
  Edited by: F Bagnoli, R Rappuoli (2015)
  www.caister.com/vaccines

• Antifungals: From Genomics to Resistance and the Development of Novel Agents
  Edited by: AT Coste, P Vandeputte (2015)
  www.caister.com/antifungals

• Bacteria-Plant Interactions: Advanced Research and Future Trends
  www.caister.com/bacteria-plant

• Aeromonas
  Edited by: J Graf (2015)
  www.caister.com/aeromonas

• Antibiotics: Current Innovations and Future Trends
  Edited by: S Sánchez, AL Demain (2015)
  www.caister.com/antibiotics

• Leishmania: Current Biology and Control
  Edited by: S Adak, R Datta (2015)
  www.caister.com/leish2

• Acanthamoeba: Biology and Pathogenesis (2nd edition)
  Author: NA Khan (2015)
  www.caister.com/acanthamoeba2

• Microarrays: Current Technology, Innovations and Applications
  Edited by: Z He (2014)
  www.caister.com/microarrays2

• Metagenomics of the Microbial Nitrogen Cycle: Theory, Methods and Applications
  Edited by: D Marco (2014)
  www.caister.com/n2

Order from caister.com/order