FEBS Letters xxx (2015) xxx–xxx

FEBS Letters

journal homepage: www.FEBSLetters.org

Review

A sweet code for glycoprotein folding

Julio J. Caramelo *, Armando J. Parodi

Fundación Instituto Leloir and Instituto de Investigaciones Bioquímicas de Buenos Aires (IIBBA-CONICET), Avda. Patricias Argentinas 435, Buenos Aires C1405BWE, Argentina

ARTICLE INFO

Article history:
Received 30 June 2015
Revised 15 July 2015
Accepted 15 July 2015
Available online xxxx

Edited by Wilhelm Just

Keywords:
Glycoprotein folding
Endoplasmic reticulum
Chaperones
Calnexin
Calreticulin
Glucosyltransferase
Proteasomal degradation

ABSTRACT

Glycoprotein synthesis is initiated in the endoplasmic reticulum (ER) lumen upon transfer of a glycan ($Glc_3Man_9GlcNAc_2$) from a lipid derivative to Asn residues (*N*-glycosylation). *N*-Glycan-dependent quality control of glycoprotein folding in the ER prevents exit to Golgi of folding intermediates, irreparably misfolded glycoproteins and incompletely assembled multimeric complexes. It also enhances folding efficiency by preventing aggregation and facilitating formation of proper disulfide bonds. The control mechanism essentially involves four components, resident lectin-chaperones (calnexin and calreticulin) that recognize monoglucosylated polymannose protein-linked glycans, lectin-associated oxidoreductase acting on monoglucosylated glycoproteins (ERp57), a glucosyltransferase that creates monoglucosylated epitopes in protein-linked glycans (UGGT) and a glucosidase (GII) that removes the glucose units added by UGGT. This last enzyme is the only mechanism component sensing glycoprotein conformations as it creates monoglucosylated glycans exclusively in not properly folded glycoproteins or in not completely assembled multimeric glycoprotein complexes. Glycoproteins that fail to properly fold are eventually driven to proteasomal degradation in the cytosol following the ER-associated degradation pathway, in which the extent of *N*-glycan demannosylation by ER mannosidases play a relevant role in the identification of irreparably misfolded glycoproteins.

1. Introduction – the ER as the initial site of protein secretion

Nearly one third of eukaryotic proteins belong to the secretory pathway, representing about 8000 proteins in humans. Most of them are synthetized by ribosomes attached to the endoplasmic reticulum (ER) and enter into the ER through the Sec61 αβγ translocon complex. Alternatively, some proteins may enter post-translationally. This pathway is more frequently employed in yeast [1]. Secretory pathway proteins fold and assemble in the ER before continuing their journey. Topologically equivalent to the cell exterior, the lumen of the ER is a highly crowded environment, with an oxidizing potential and high calcium concentration, which ranges in the order of millimolar. For this reason, secretory pathway proteins face unique challenges in order to fold properly in this potentially hostile environment. The ER molecular chaperones belong to protein families commonly found in other locations, such as HSP70 (BiP in the ER) and HSP90 (GRP94 in the ER), but lacks chaperonine-like proteins. This absence can be partially compensated by GRP94 which, by recognizing advanced folding intermediate, collaborates with BiP in assisting the folding pathway of selected substrates [2]. Concomitantly with their folding, most secretory pathway proteins acquire disulfide bridges and are *N*-glycosylated in the ER. Compared with proteins in other locations, secretory pathway proteins have a higher frequency of disulfide bonds, which stabilize their tertiary structure and oligomer association. A varied group of protein disulfide isomerases (PDIs), unique to the ER, guarantee the fidelity of the oxidation process. Equally important, the presence of *N*-glycans allows the operation of specialized mechanisms that assist the protein folding.

2. *N*-glycosylation

Approximately one quarter of the eukaryotic proteins are *N*-glycosylated at the lateral chain of Asn residues displayed within the consensus sequence Asn-Xxx-Ser/Thr, where Xxx cannot be Pro (in some cases Asn-Xxx-Cys, Asn-Gly or Asn-Xxx-Val sequences can also be used) [3]. This motif, named *N*-glycosylation sequon, is quite common, with a frequency of about 4–10 sequons every 1000 residues [4].

In most organisms the glycan $Glc_3Man_9GlcNAc_2$ (G3M9) (Figs. 1A, B and 2) is initially transferred *en bloc* from a dolichol-pyrophosphate-oligosaccharide, while some protozoans transfer a shorter version. For instance, *Trypanosome cruzi* uses an

* Corresponding author. Fax: +54 5238 7501.
E-mail addresses: jcaramelo@leloir.org.ar (J.J. Caramelo), aparodi@leloir.org.ar (A.J. Parodi).

http://dx.doi.org/10.1016/j.febslet.2015.07.021
0014-5793/

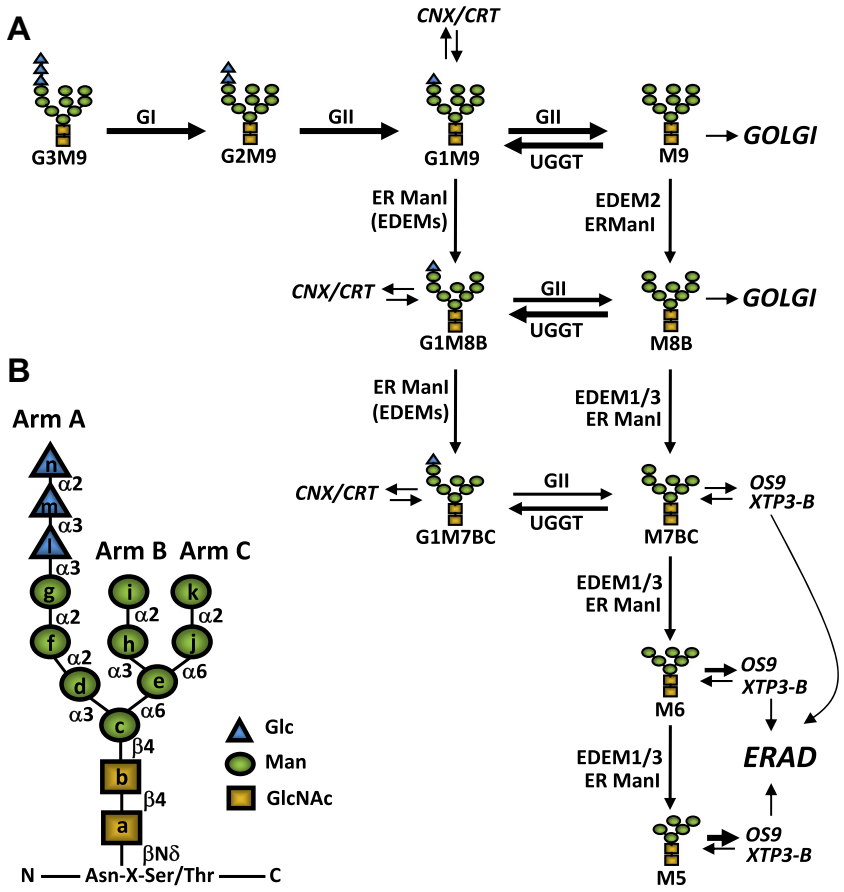

Fig. 1. (A) Processing and recognition of *N*-glycans in the early secretory pathway. (B) Structure of the $Glc_3Man_9GlcNAc_2$ glycan transferred by the oligosaccharyltransferase.

oligosaccharide that lacks glucose residues ($Man_9GlcNAc_2$, M9) [5]. The homogeneity of the *N*-glycan transferred among different species is in sharp contrast with the enormous variability of *N*-glycans displayed by mature glycoproteins at their final destinations. This diversity is acquired through the action of several glycosidases and glycosyltransferases located partially in the ER but mainly in the Golgi apparatus, following a complex and non-linear process. In mature complex glycoproteins the core glycan $Man_3GlcNAc_2$ (M3) is usually the only remnant of the transferred glycan. This observation points to the dual biological functions of *N*-glycans [6]. On one hand, the great variety of glycosidic linkages and monosaccharide units makes *N*-glycans perfectly suited to play a central role in many molecular recognition events at the cell surface. A diverse array of biological processes are modulated or mediated by *N*-glycans, such as differentiation, migration and proliferation. On the other hand, early high mannose *N*-glycans assist to the protein folding process of several glycoproteins. Due to their bulky and hydrophilic nature, *N*-glycans can be considered as covalently attached chemical chaperones. They prevent protein aggregation by hindering hydrophobic aggregation-prone regions, acting as a kind of hydrophilic shield. They can also facilitate the acquisition of secondary structure elements such as turns [7]. In addition, *N*-glycans are used as a "bar code" to display information regarding the folding status and age of glycoproteins. This code is generated and interpreted by several glycosidases, glycosyltransferases and lectins operating in the early secretory pathway. In brief, monoglucosylated *N*-glycans are a signature of glycoproteins displaying immature conformations, while progressive loss of mannose residues marks slowing folding or terminally misfolded proteins, eventually leading them to degradation. These intertwined processes, protein folding quality control (QC) and endoplasmic reticulum associated degradation (ERAD), are discussed in the following sections and are schematically outlined in Figs. 1A and 2).

3. The ER glucosidases

The two outermost glucose residues of the initially transferred glycan G3M9 are rapidly removed by the sequential action of glucosidase I (GI) and glucosidase II (GII), thus generating the monoglucosylated intermediate $Glc_1Man_9GlcNAc_2$ (G1M9). This glycan is recognized by ER resident lectins calreticulin (CRT), which is a soluble protein and its type I integral membrane paralogue calnexin (CNX), that retain the glycoproteins in the ER until GII cleaves the last glucose residue. At this point, properly folded proteins leave the ER toward their final destination. By contrast, those proteins unable to attain a native conformation or incompletely assembled complexes are recognized by the enzyme UDP-glucose:glycoprotein glucosyltransferase (UGGT) that adds the last glucose removed by GII exclusively to glycoproteins not displaying their native conformations, thus re-creating the G1M9 structure. This process allows the re-association of glycoproteins and CRT/CNX. Cycles of deglucosylation by GII and reglucosylation by UGGT continue until the glycoprotein folds properly, that is, when they are no longer recognized by the last enzyme [8]. Alternatively, proteins unable to fold properly are eventually directed to degradation by the ERAD machinery [9].

The α1,2-exoglucosidase GI is a type II membrane protein with a short cytosolic tail that belongs to the glycosylhydrolase 63

Please cite this article in press as: Caramelo, J.J. and Parodi, A.J. A sweet code for glycoprotein folding. FEBS Lett. (2015), http://dx.doi.org/10.1016/j.febslet.2015.07.021

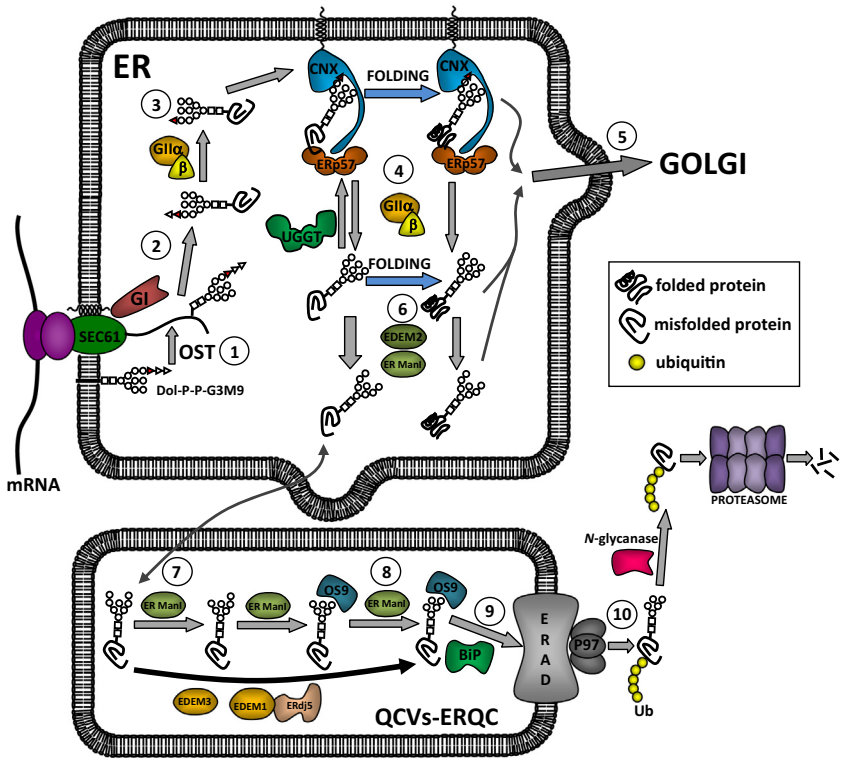

Fig. 2. Glycoprotein processing in the early secretory pathway. (1) G3M9 transfer to nascent proteins by the oligosaccharyltransferase complex (OST). (2) After trimming by glucosidase I (GI), (3) *N*-glycans are further processed by the heterodimeric glucosidase II (GIIα-GIIβ) to generate the monoglucosylated intermediates that are bound by the ER lectins calnexin (CNX) or calreticulin (not shown). These lectins retain the monoglucosylated intermediates in the ER, preventing their aggregation and facilitating their folding through the activity of the associated enzymes Erp57 and CypB (not shown). (4) Eventually, GII cleaves the remaining glucose residue. At this point properly folded species can leave de ER, while misfolded species are recognized by the folding sensor UDP-Glc:glycoprotein glycosyltransferase (UGGT). (5) High mannose glycoproteins leave the ER to the Golgi by bulk flow or associated with sorting receptors. (6) Mannose trimming initiates in the ER by cleavage of the mannose residue at branch B by EDEM-2 and ER α1,2-mannosidase I (ER ManI). (7) Further mannose trimming takes place in the ER or in the specialized quality control vesicles (QCVs) or ER quality control (ERQC) subcompartments. This process is mainly in charge of ER ManI and EDEMs 1 and 3. EDEM 1 associates with the PDI J-domain containing protein ERdj5, which reduces the disulfide bridges of misfolded species. (8) Uncovering of the α1,6-Man residue allows the interaction of the M6 and M5 glycans with the lectin OS-9 and XTP3-B (not shown), and (9) in conjunction with BiP deliver them to the ERAD retrotranslocation and ubiquitination machinery.

family. GI is associated to the Sec61 αβγ translocon complex, ensuring the removal of the outermost glucose almost immediately after transfer of the *N*-glycan [10]. The α1,3-exoglucosidase GII is an ER-resident soluble heterodimer composed by a catalytic subunit (GIIα) non-covalently bound to the GIIβ subunit [11–13]. GIIα displays the (G/F)(L/I/V/M)WXDMNE consensus sequence typical of glycosylhydrolase family 31. GIIα is retained in the ER through its association to GIIβ that has a KDEL-like ER retention/retrieval signal. GIIβ has a mannose receptor homolog (MRH) domain, which works as a lectin for high mannose glycans [14]. This domain presents the substrates to the GIIα subunit, thus dramatically increasing its activity toward high mannose *N*-glycans [15–17]. Interestingly, MRH domains also appear in other lectins of the secretory pathway as both the cation-dependent and independent Golgi receptors for lysosomal glycoprotein enzymes and in lectins involved in ERAD (see below), although with different specificities. Both *in vitro* and *in vivo* experiments showed that the activity of GII decreases as the mannose content on *N*-glycan arms B and C (Fig. 1A and B) decreases. The main residue influencing GII activity is the outermost residue in arm C (residue k, Fig. 1B). The molecular basis of this trend should be interpreted with precaution, since GII is inhibited differently by its end products. For instance, the decreased activity toward G1M7BC may reflect the greater inhibitory power of M7BC [18]. Regarding the relative speed of both glucose cleavages, current evidence suggest that the first cleavage is faster than the second [16,19], although this difference is not observed *in vitro* in the presence of crowding agents [20]. A slower cleavage of the second glucose would provide a better chance for glycoproteins to enter the CNX/CRT cycle, although this is matter of debate. Interestingly, GI and GII are potential targets for antiviral therapies against several viruses that depend strongly on the CNX/CRT cycle to mature [21].

4. The lectins (CNX and CRT)

CRT (about 400 residues, 46kDa) and CNX (about 572 residues, 65kDa) expression is induced by several types of stress such as heat shock, heavy metals, amino acid deprivation and calcium-mobilizing agents. CRT is a multifunctional protein found in several locations (ER lumen, cytosol, nucleus, secretory granules and plasma membrane). It has a central role in glycoprotein folding and calcium homeostasis, but it has also been implicated in trafficking of nuclear receptors, mRNA stability, complement activation and angiogenesis [22,23]. CRT and CNX display about 45% sequence similarity, with a similar domain organization and structure. Both proteins have an N-terminal signal peptide and share a β-sandwich N-terminal domain similar to that of leguminous lectins, which is responsible for the lectin activity, followed by a Pro-rich domain (P-domain) and an acidic C-terminal domain. An intervening transmembrane segment appears in CNX before the C-terminal domain, which faces the cytosolic side. CRT is retained in the ER through a KDEL-like C-terminal signal, while CNX displays an RKPRRE ER-localization signal on its cytosolic tail. The

Please cite this article in press as: Caramelo, J.J. and Parodi, A.J. A sweet code for glycoprotein folding. FEBS Lett. (2015), http://dx.doi.org/10.1016/j.febslet.2015.07.021

P-domain is a unique feature of these proteins. In CNX it consists of four copies of motif 1 (IxDP(D/E)(A/D)xKP(D/E)DWD(D/E)) followed by four copies of motif 2 (GxWxxPxIxNPxY) in a 11112222 pattern. These motifs are arranged in four modules with a head to tail disposition. The P-domain of CRT is shorter, having three copies of similar motifs. The P-domain of CNX protrudes about 140 Å from the globular domain. It is very flexible, embracing bound glycoproteins and preventing their aggregation. Glycan recognition by these lectins is strictly dependent on the glucose residue, while additional mannoses also contribute to the binding energy ($Kb = 2.2 \times 10^4\ M^{-1}$, $56 \times 10^4\ M^{-1}$ and $102 \times 10^4\ M^{-1}$ for Glcα1,3-Man, Glcα1–3Man α1,2-Man and Glcα1,3-Manα1–2Manα1,2-Man, respectively) [24]. The equatorially oriented 2-hydroxyl group of the glucose is necessary for binding, since 2-deoxyGlcα1,2-Man does not bind to CRT. Interestingly, glycan binging to CRT is enhanced in the presence of crowding agents, suggesting a stronger interaction *in vivo* with its client proteins [25]. Even though CRT and CNX glycan binding profiles are identical, they bind *in vivo* to a partially overlapping set of glycoproteins, the former protein preferentially associating to glycans located distally to the ER inner membrane and the latter binding to glycans close to it [26,27]. Glycoprotein binding to CNX or CRT can take place cotranslationally. A competition between BiP and CRT/CNX regulates the initial selection of chaperones, where *N*-glycosylation sequons located near the N-terminus of the substrates favors the interaction with the lectins [28]. In general, BiP and CRT/CNX cooperate to assist in the complete folding pathway of glycoproteins, where early folding intermediates bind to BiP and more advanced intermediates are bound by the lectins [29–31]. Interestingly, by delaying the oxidation of some key cysteine residues, early interaction with CNX ensures the proper disulfide bond formation of influenza membrane glycoproteins hemagglutinin and neuraminidase [32,33]. Of note, some microorganisms only have one of these lectins. For instance, trypanosomatid protozoa only have CRT, while *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* genomes only code for CNX. Interestingly, CRT and CNX can also interact through polypeptide-based contacts, similarly to conventional chaperones [34]. Indeed, the crystal structure of the globular domain of CRT revealed a putative peptide binding site located at the edge of the lectin site [35]. This ability is modulated in CNX by environmental stressors, such as calcium depletion or high temperature [36]. *In vitro* studies have shown that CRT can interact with glycosylated and non-glycosylated substrates with similar affinities but with different kinetics, and the conformation displayed by the P-domain depends on the nature of the bound substrate [37].

CRT is also one of the main calcium buffers in the ER, accounting for about half of the calcium content in the organelle. This activity is located in the acidic C-terminal domain of the protein, which can bind about 25 calcium ions with low affinity (Kd about 1 mM). The lectin and calcium buffering activities of CRT are mutually independent [38]. The C-terminal domain has natively unfolded structure, and upon calcium binding it adopts a more rigid and compact conformation [39] and increments its solvent exposition in the full length protein [40]. Interestingly, this domain seems to regulate the cellular localization of CRT [41], and upon depletion of ER calcium it could promote the partial retrotranslocation of CRT into the cytosol [42]. The presence of CRT in the cytosol and its biological relevance is a matter of debate. Cytosolic CRT can be arginylated at the N-terminal end after the signal peptide cleavage [43] and it has been recently shown that this modification increases the half-life of the protein [44]. Since the arginylation machinery is located in the cytosol, the discovery of arginylated CRT lacking its signal peptide is a good evidence of its alternative cytosolic localization. Cytosolic CRT may participate in many functions, such as regulation of nuclear hormone receptors localization [45] and mRNA stability [46]. CRT KO in mice is embryonically lethal at E14.5 as a consequence of an impaired cardiac development [47]. This effect can be prevented by expressing a constitutively active calcineurin in the heart [48], thus showing that CNX can compensate the absence of CRT in the QC system and pointing to an important role of CRT in regulating the cellular calcium homeostasis. A reciprocal situation occurs with CNX. Contrary to early reports [49], it was shown that CNX KO mice are viable [50]. These animals display a dysmyelination phenotype in the peripheral and central nervous systems, but their immune system is not affected, they are fertile and they have a normal lifespan.

5. The interaction of lectins and glycoproteins

Glycoprotein-CNX/CRT association results in increased folding efficiency, decreased aggregation and facilitation of disulfide bridge isomerization. This last effect is mediated by the activity of ERp57, a member of the PDI family. The domain architecture ERp57 (also known as PDIA3) is similar to PDI. Both proteins are composed of four thioredoxin domains (a, b, b′ and a′), where domains a and a′ display the redox active motif CGHC. While the b′ domain of PDI has a hydrophobic patch involved in client protein binding, a cluster of positively residues in the b′ domain of ERp57 mediates its interaction with the negatively charged tip of the CRT/CNX P-domain, which in turn presents most known substrates to ERp57 [51,52]. The interaction between CNX and ERp57 is modulated by a conserved disulfide bridge located at the P-domain, which is absent in CRT [53]. Deletion of this disulfide bridge leads to a 5-fold increase in the binding constant. The moderate affinities between CRT/CNX and ERp57 (Kd about 7 μM and 6 μM for CRT and CNX, respectively) and its fast off-rate (Koff > 1000 s-1) may imply that ERp57 can rapidly assist several glycoprotein-lectin complexes. In addition, the flexibility of the P-domain probably allows ERp57 to scan for disulfide bridges at distant positions. Heavily glycosylated and disulfide bond-rich substrates, such as low density lipoprotein receptor (LDLR), rely on this interaction for efficient folding and trafficking.

The tip of the P-domain also serves as a platform for interaction with cyclophilin B (CypB), an ER member of the peptidyl prolyl *cis–trans*-isomerase family, thus expanding the capabilities fulfilled by the CRT/CNX cycle. A patch of positively charged residues in CypB mediates this interaction, with a dissociation constant of about 10 μM [54]. By presenting the substrates to CypB, the interaction with CNX may improve the intrinsic low catalytic efficiency of the isomerase. Both CypB and ERp57 are very abundant proteins and the factors regulating the balance between ERp57-CRT/CNX and CypB-CRT/CNX complexes are unknown.

In addition, the TM domain of CNX displays one or two conserved CYS residue near the cytosolic side. Under non-stress conditions, palmitoylation of these residues directs CNX to the mitochondria-associated membrane (MAM) and facilitates its interaction with the sarcoendoplasmic reticulum calcium transport ATPase (SERCA) 2b [55]. This interaction modulates the ER Ca^{2+} content and the ER–mitochondria Ca^{2+} crosstalk. Upon ER stress, CNX is less palmitoylated and migrates from MAM to the ER and the pericentriolar ER-derived control compartment (ERQC) where it associates with ERp57 (and likely CypB), thus focusing its activity in glycoprotein folding. This exemplifies the dynamic nature of the ER, which can adapt and re localize its component according to the particular demands of the cell. Interestingly, the cytosolic domain of CNX has been recently shown to upregulate the activity of the transcription factor STAT3 [56]. Treatment with epidermal growth factor induces the release of this domain by caspase 8, which in turn inhibits the activity of PIAS3, an inhibitor of STAT3. This pathway depends on the palmitoylation of CNX, and it is inhibited under ER stress conditions, thus providing a potential link between ER stress and cell proliferation.

Please cite this article in press as: Caramelo, J.J. and Parodi, A.J. A sweet code for glycoprotein folding. FEBS Lett. (2015), http://dx.doi.org/10.1016/j.febslet.2015.07.021

6. The misfolded glycoprotein sensor (UGGT)

Many proteins only need one round of association with CRT/CNX in order to attain their native conformations [57]. This is particularly true for those few organisms that lack UGGT such as *S. cerevisiae*. Alternatively, glycoproteins still displaying non-native structures or partially assembled complexes are reglucosylated by UGGT. By triggering the re-association of glycoproteins and CRT/CNX, UGGTs activity increases the solubility of partially folded species and improves their chances to fold successfully [58]. UGGT is a unique enzyme of about 170kDa (1555 residues in humans) that mixes the activity of a glycosyltransferase with the specificity of a classical chaperone [59]. UGGT KO in mouse is embryonically lethal at E13 [60,61] and viability of *S. pombe* lacking UGGT is only affected under extreme ER stress conditions [62]. UGGT is an ER soluble protein that displays a hydrophobic N-terminal signal peptide and a C-terminal KDEL-like ER retention/retrieval signal. The catalytic activity of the enzyme resides in the highly conserved C-terminal domain (about 20% of the protein), that belongs to the glycosyltransferase family 24. The N-terminal domain (80% of the protein) is believed to be responsible for misfolded acceptor recognition. In most organisms, monoglucosylated high mannose glycans can be formed by two different processes, either by the first GII cleavage or by UGGT activity. For this reason the occurrence of the enzyme was first demonstrated in *T. cruzi* [63]: as this parasite transfers M9 to nascent glycoproteins, detection of G1M9 necessarily implied the existence of a protein glucosylating activity. Early studies clearly showed a marked preference of UGGT for misfolded glycoproteins, and the first substrates employed in *in vitro* assays were 8 M urea-denatured bovine thyroglobulin, RNAse B or soybean agglutinin (SBA) [64]. The original aim of using denatured substrates was to improve the exposition of the *N*-glycans for increasing substrate glucosylation but an unexpected fortunate collateral consequence was the exposure of hydrophobic surface in the acceptors. *In vitro* assays using well defined substrates showed that UGGT uses a bipartite signal, recognizing on one hand the *N*-glycan innermost GlcNAc residue that is generally occluded in native conformations and exposed hydrophobic patches on high mannose-bearing glycoproteins. This last recognition process shows higher efficiency on advanced molten globule-like folding intermediates and UGGT catalytic efficiency correlates with the ANS-binding capacity of the acceptors (this drug associates preferentially to proteins exposing hydrophobic patches) [65–67]. The ability of UGGT can also be used to control the proper assembly of oligomeric structures. For example, it is capable of glucosylating well folded subunits of incompletely assembled SBA (a tetramer in the native state), a likely consequence of the presence hydrophobic patches exposed by this substrate [68]. In addition, glycopeptides can also work as glucose acceptors, provided they are long enough and that they have hydrophobic segments [69]. Surprisingly, UGGT can also recognize hydrophobic non-proteinaceous epitopes attached to high mannose glycans [70]. The distance required between hydrophobic patches and the *N*-glycan is still a matter of debate. While in some substrates it seems to be restricted to a short range [71], in others it can be extended to a few nanometers [72]. Recognition of substrate protein elements besides the acceptor glycan is not an exclusive feature of UGGT. For example, UDP-*N*-acetylglucosamine:lysosomal enzyme *N*-acetylglucosamine phosphotransferase selectively initiates the formation of the mannose 6-P signal in lysosomal glycoproteins by recognizing not only of a high mannose *N*-glycan but also a patch of positively charged amino acid residues, while the UDP-Gal:glycoprotein hormone N-acetylgalactosamyniltransferase selectively adds a GalNAc residue to certain but not all secreted hormones. There are two homologs of the enzyme in mammals, UGGT1 and UGGT2 that display a 55% sequence identity (75% similarity). It was initially believed that UGGT2 was inactive but more recent results showed that *C. elegans* knock down of each protein displayed different phenotypes [73] and that both proteins are active in a mouse hybridome [74]. Indeed, it has been recently shown *in vitro* that UGGT2 has a *bona fide* glucosyltransferase activity, with specificity similar to UGGT1 [60].

Even though it is expected that *in vivo* a wide range of glycoproteins are UGGT substrates, very few endogenous substrates have been identified so far. Two known examples are lysosomal proteases rich in disulfide bridges, prosaposin in mammals [75] and cruzipain (CZ) in *T. cruzi* [76]. UGGT KO in *T. cruzi* precludes the association of CZ and CRT and results in an enzyme impaired oxidative maturation [30]. In those parasites most CZ molecules are retained in the ER bound to BiP, with an important fraction forming covalently bound aggregates. A similar situation was observed for prosaposin in mouse uggt −/− cells. The exquisite selectivity of UGGT can be exploited in subtle ways. For instance, UGGT optimizes the loading to MHC class I (MHC-I) with high affinity peptides [77]. The peptide loading complex is composed by tapasin, ERp57, the heavy chain of MHC-I, β2-microglobulin and CRT. Tapasin forms a stable disulfide bridge with ERp57 domain a, while CRT stabilizes their association with MHC-I by binding ERp57 and a conserved *N*-glycan attached to MCH-I. Interestingly, UGGT preferentially reglucosylates the *N*-glycan of those complexes containing suboptimal peptides, thus precluding their exit from the ER.

UGGT activity is expected to facilitate binding of glycoproteins to CRT/CNX, thus slowing their passage through the ER. Nevertheless, experiments using mouse uggt −/− cells revealed a more complex scenario [57]. Three type of glycoproteins appeared in this system. On one hand there were glycoproteins whose transit was not affected or was accelerated upon UGGT deletion, the latter being a logical consequence of a shorter interaction with CRT/CNX. Surprisingly, some proteins showed a slower transit through the ER. This observation pointed to a more active role of UGGT in assisting the folding maturation of some glycoproteins. In this respect, the selenium containing thioredoxin-like protein Sep15 may modulate the activity of UGGT [78]. This protein displays an unusual CGU motif (where U stands for selenocysteine). Sep15 associates with UGGT in a 1:1 complex with very high affinity (Kd ~20 nM), thus retaining Sep15 in the ER [79]. Sep15 expression is enhanced upon ER stress [80], and it has been proposed that may facilitate the oxidative maturation of some selected glycoproteins while being recognized by UGGT. Sep15 deficient mice develop early cataracts, probably due to a misfolding defect of lens proteins [81]. In addition, *in vitro* assays showed that Sep15 increases the activity of UGGT1 and UGGT2, being this effect more dramatic for the later protein [74]. This activation is independent of any redox effect exerted by Sep15 on the substrates, pointing to a modulation of the activity of the glucosyltransferases.

The balance between GII and UGGT activities is pivotal in determining glycoprotein binding to CRT/CNX. *In vivo* observations suggest that as mannose residues are cleaved, UGGT activity diminishes much less sharply than that of GII [19,82]. This would ensure the retention of immature species in the ER. Nevertheless, this may also negatively affect the passage of terminally misfolded proteins to the ERAD machinery. Here lies one major puzzle, since the QC and ERAD machinery should discriminate folding intermediates within a productive pathway from terminally misfolded proteins. A mistake in any other way may have dangerous consequences. Interestingly, even though UGGT delays the secretion of immature glycoproteins, the enzyme does not affect the processing of misfolded species by ERAD [83]. This implies that

Please cite this article in press as: Caramelo, J.J. and Parodi, A.J. A sweet code for glycoprotein folding. FEBS Lett. (2015), http://dx.doi.org/10.1016/j.febslet.2015.07.021

somehow the ERAD machinery can extract very efficiently the substrates from the CRT/CNX cycle.

Recent structural studies on UGGT from the thermophilic fungus *Chaetomium thermophilum* (1509 residues long, 51% similarity with human UGGT1) suggest a plausible mechanism for substrate recognition [84]. A bioinformatics analysis predicted that the N-terminal portion of the protein contains three thioredoxin (Trx)-like domains that lack a CXYC motif, followed by a β-rich domain. This prediction was confirmed after solving the structure of the third Trx-like domain, showing a five-stranded β-sheet with a β1–β3–β2–β4–β5 topology surrounded by six α-helices, very similar to bacterial protein disulfide isomerase DsbA/C. Interestingly, the third domain could be crystallized in two forms, with and without a bound detergent molecule. A hydrophobic patch composed by several conserved residues from strands β2-β5 and helix α2 interacted alternatively with the detergent molecule or with four hydrophobic residues of helix α6 (Phe820, Phe825, Phe828 and Leu829). Homology modelling predicted similar, or even larger, hydrophobic patches in the other two Trx-like domains. Even though more biochemical and structural analysis are needed, it can be speculated that this hydrophobic patch is involved in the recognition of misfolded acceptors, while the plasticity of helix α6 may serve to regulate its exposition in the absence of substrates. The presence of three similar domains would provide the ability to scan for structural defects at various distances from the *N*-glycan. Given the high degree of similarity between *Chaetomium thermophilum* and others UGGTs, an analogous domain organization and structure is expected. In addition, these findings suggest an evolutionary mechanism for the origin of this enzyme, in which a portion of a PDI family protein may have fused with a glycosyltransferase domain.

7. ERAD or a time for oblivion

Progressive loss of α1,2-mannose residues marks glycoproteins for degradation. There are seven α1,2-mannosidades in the mammalian cell secretory pathway, all belonging to the glycosylhydrolase family 47: ER α1,2-mannosidase I (ER Man I), three ER degradation-enhancing -mannosidase-like proteins (EDEM 1, 2 and 3) and three Golgi α1,2-mannosidases (Golgi Man IA, IB and IC). The Golgi enzymes are responsible for trimming *N*-glycans from M9-M8 to $Man_5GlcNAc_2$ (M5) of proteins continuing their journey through the secretory pathway. This last trimming is the first step in the formation of hybrid- and complex-type glycans displayed by most mature glycoproteins. Subsequent trimming in the Golgi until the core glycan (M3) is reached is performed by α-mannosidase II, that cleaves the branching α1,3 and α1,6 bonds of the B and C arms. Interestingly, some mature glycoproteins display high mannose structures. It is assumed that glycans located distally to the Golgi inner membrane during secretion, may not be fully accessible to the membrane-bound Golgi processing mannosidases.

As mentioned above, the ERAD and QC machineries must discriminate productive folding intermediates from irreparably misfolded species, a particularly difficult task for proteins with slow folding pathways. Mannose trimming in the ER is carried out by ER Man I and the EDEMs, which are not particularly efficient proteins. For this reason, even though the three EDEMs display all the residues needed for catalysis and can bind the inhibitor kifunensin, their mannosidase activity was difficult to detect. It is believed that the low efficiency of those enzymes enables the occurrence of a mechanism known as "mannose timer", that allows the conformational maturation of slow folding proteins (see below). In basal situations ER Man I is mainly located in ER-derived quality control vesicles (QCVs) and upon ER-stress it migrates to ER sites harboring the main elements of the QC (ERQCs, see below) [85–88]. This initial physical separation would protect folding glycoproteins from ER Man I. Upon ER stress, misfolded glycoproteins would migrate to the ERQCs where they encounter a high concentration of ER Man I, thus sharply increasing the mannose trimming pace. In addition, other important components of the ERAD and QC systems such as CRT, CNX, EDEM1, OS-9 and XTP3-B also localize to the ERQC. Probably, CRT and CNX may commute their client proteins between the ER and the ERQCs. ER Man I and EDEM1 are single-pass type II membrane proteins, while EDEM2 and EDEM3 are soluble proteins. EDEM1 has five *N*-glycosylation sequons, where the most C-terminal site is partially occupied. In addition, EDEM1 signal peptide is processed very slowly, thus rendering a mix of soluble and membrane bound species [89]. Membrane-bound EDEM1 associates with the ERAD scaffold protein SEL1L, while the soluble form binds to ERdj5 [90], a unique J-domain containing PDI with a strong reductase activity. This complex is reminiscent of the CRT/CNX-ERp57 complexes. ERdj5 reduces the disulfide bridges of glycosylated ERAD substrates, and handles them to BiP through its J-domain in an ATP-dependent manner; BiP in turn delivers the substrates to the ERAD machinery [91]. Non-glycoprotein ERAD substrates may be delivered by BiP to ERdj5, and after reduction of their disulfide bridges they are also transferred by BiP to the SEL1L-containing complex [92]. Interestingly, glycoproteins use this last pathway upon inhibition of glucose trimming, meaning that under stress condition the ERAD pathways for glycosylated and non-glycosylated substrates can be interchangeable.

The exact contribution of each α1,2-mannosidade to the trimming process is still under investigation. It was originally believed that generation of the isomer M8B, which lacks the more external mannose residue of B arm (residue i, Fig. 1B), was the signal for degradation. It was recently shown that EDEM2 is the main responsible for this step, with some contribution of ER Man I. Nevertheless, M8B is also displayed by properly folded proteins that reach the Golgi. It is accepted now that further mannose trimming is needed for diversion to ERAD and that most ERAD substrates end displaying $Man_6GlcNAc_2$ (M6) and M5 [93]. Apparently, EDEM 1 and 3, and to some extent ER Man I, are in charge of the subsequent loss of mannose that generate M7BC and M7AB isomers (M7s lacking residues i and k or g and i, respectively, Fig. 1B) [94]. Those enzymes are also believed to be responsible for the acquisition of the final M6 and M5 glycoforms. A crucial event is the cleavage the outer α1,2-mannose of A arm (residue g, Fig. 1B), the acceptor site for UGGT activity, since this cleavage definitively extracts the glycoprotein from the CRT/CNX cycle. Removal of residue g may be performed also by a *cis* Golgi endomannosidase (EM) that yields the disaccharide Glc1Man1 from G1M9 or G1M8B. As mentioned above, forcing the occurrence of a monoglucosylated glycoform does not seem to affect the degradation kinetics of some glycoproteins [83]. In this sense, ER Man I can use glucosylated substrates, thus indicating that it may be operative while they are still in the CRT/CNX cycle [95].

A very important step is the uncovering of the α1,6-mannose in arm C (residue j, Fig. 1B) [78,79]. This residue is the main epitope recognized by lectins OS-9 [96,97] and XTP3-B [98,99], which through their MRH domains (one in OS-9 and two in XTP3-B) escort the substrate to the dislocon and the ubiquitination machinery of the ERAD. Interestingly, it has been found that XTP3-B may have a stabilizing effect on some misfolded substrates, probably by delaying its premature degradation [100]. Both mannose removal steps, conversion of M9 to M8B and of this last glycan to M7BC are slow processes in mammalians cells when compared to the deglucosylation reactions. This fact affords two checkpoints ensuring that only terminally misfolded glycoproteins, and not folding intermediates, are derived to proteasomal degradation after a

Please cite this article in press as: Caramelo, J.J. and Parodi, A.J. A sweet code for glycoprotein folding. FEBS Lett. (2015), http://dx.doi.org/10.1016/j.febslet.2015.07.021

relatively long ER residence. A somewhat different situation is present in the budding yeast, in which an ER α-mannosidase I and only one EDEM-like protein called Htm1p are present. The former transforms M9 to M8B very rapidly in practically all glycoproteins, whereas the latter converts M8B to M7BC in a slow fashion, thus affording only one checkpoint for the exclusive degradation of terminally misfolded glycoproteins. Htm1p was found to be forming a complex with protein disulfide isomerase. The oxidoreductase enhanced the mannosidase activity of Htm1p and participated in the recognition of ERAD substrates. In the fission yeast *S. pombe* there are, similarly to *S. cerevisiae* only one ER α-mannosidase and Htm1p-like proteins, but in this case both demannosylation reactions are slow processes [101].

Retrotranslocation of misfolded glycoproteins from the ER lumen to the cytosol to be eventually degraded by proteasomes ultimately depends of protein complexes some of whose components are integral ER membrane proteins [102]. Depending on the location of the misfolded domain, proteins are delivered to proteasomal degradation through different complexes. Thus, ER membrane proteins with folding defects in the cytosolic portion of the molecule (ERAD-C) are extracted from the ER membrane by the so called Doa10 complex, whereas those in which the folding defect is present in the lumenal (ERAD-L) or membrane (ERAD-M) portions of the molecules utilize the Hrd1 complex. Both Doa10p and Hrd1p are integral membrane proteins showing E3 ligase activity in their cytosolic portions. In addition, other proteins both lumenal and cytosolic, form part of the complexes. Some of them are chaperones, others display E2 ubiquitin-conjugating activity or participate in misfolded protein recognition. For instance, glycoproteins exposing α(1,6)-linked mannosyl units bound to OS9 are driven to degradation only if present in unstructured polypeptides, a feature recognized by Hrd3p, a protein found in the Hrd1 complex. Finally, there are protein components common to both complexes, as for instance p97 in mammals, responsible for the membrane extraction of misfolded proteins in an ATP-dependent manner. It should be mentioned that the pore by which ERAD-L substrates are actually transported to the cytosol has not been unequivocally identified yet but it has been established that to be able to be translocated to the cytosol, misfolded lumenal proteins must be previously unfolded in the ER lumen, a process involving conventional chaperones and protein disulfide isomerases.

N-glycans are removed from the protein moieties prior to proteasomal degradation. It appears that a cytoplasmic peptide:*N*-glycanase (PNGase) plays an important role in both removing the glycan and constructing an efficient predegradation complex. PNGase recognizes only misfolded or denatured glycoproteins and the enzyme is bound to the proteasome by subunits S4 and HR23B as a complex with cytoplasmic protein Cdc48, a component of both ERAD complexes [103,104]. Degradation of the released glycan occurs in two stages. First, partial cleavage occurs between the chitobiose core via a cytoplasmic endo-β-N-acetylglucosaminidase or possibly a neutral-pH cytoplasmic chitobiase. A cytoplasmic α-mannosidase cleaves up to four mannose residues to generate $Man_5GlcNAc$ (residues b, c, d, e, f and g, Fig. 1B) [105]. This glycan is then taken into the lysosome for final degradation to monosaccharides via an ATP-dependent lysosomal membrane transporter [106].

8. Perspectives

Over the last years the field of protein folding and degradation in the ER has become more complex and a myriad of new players have been discovered. The ER has revealed as a very rich compartment, with a dynamic spatial organization that responds to the different demands of the cell. Its proper function relies not only on the molecular specificities of its components, but also on their timely and spatial coordination. Subtle changes in calcium concentration, pH and membrane composition among the different subcompartments of the early secretory pathway may regulate the activity of various cargo receptors. The fact that many components of the glycoprotein recognition and processing systems are sensitive toward the folding status of their ligands and substrates could explain the exquisite discrimination ability of the ERAD. Interestingly, during evolution a limited repertoire of structural domains (Trx, MRH, J, etc) was combined to create those marvelous systems. How the ER responds to widely different requirements and how coordination between the QC and ERAD machineries is achieved is still under active research. A detailed comprehension of these issues could provide new therapeutic strategies to treat a wide spectrum of human diseases, mainly the so called "conformational diseases", that is diseases in which key proteins are unable to properly fold and are initially retained in the ER and eventually sent to degradation by the ERAD machinery.

References

[1] Zimmermann, R., Eyrisch, S., Ahmad, M. and Helms, V. (2011) Protein translocation across the ER membrane. Biochim. Biophys. Acta 1808, 912–924.
[2] Marzec, M., Eletto, D. and Argon, Y. (2012) GRP94: An HSP90-like protein specialized for protein folding and quality control in the endoplasmic reticulum. Biochim. Biophys. Acta 1823, 774–787.
[3] Zielinska, D.F., Gnad, F., Wisniewski, J.R. and Mann, M. (2010) Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. Cell 141, 897–907.
[4] Cui, J., Smith, T., Robbins, P.W. and Samuelson, J. (2009) Darwinian selection for sites of Asn-linked glycosylation in phylogenetically disparate eukaryotes and viruses. Proc. Natl. Acad. Sci. U.S.A. 106, 13421–13426.
[5] Parodi, A.J. (1993) N-glycosylation in trypanosomatid protozoa. Glycobiology 3, 193–199.
[6] Larkin, A. and Imperiali, B. (2011) The expanding horizons of asparagine-linked glycosylation. Biochemistry 50, 4411–4426.
[7] O'Connor, S.E. and Imperiali, B. (1996) Modulation of protein structure and function by asparagine-linked glycosylation. Chem. Biol. 3, 803–812.
[8] Hebert, D.N., Foellmer, B. and Helenius, A. (1995) Glucose trimming and reglucosylation determine glycoprotein association with calnexin in the endoplasmic reticulum. Cell 81, 425–433.
[9] Vembar, S.S. and Brodsky, J.L. (2008) One step at a time: endoplasmic reticulum-associated degradation. Nat. Rev. Mol. Cell Biol. 9, 944–957.
[10] Dejgaard, K., Theberge, J.F., Heath-Engel, H., Chevet, E., Tremblay, M.L. and Thomas, D.Y. (2010) Organization of the Sec61 translocon, studied by high resolution native electrophoresis. J. Proteome Res. 9, 1763–1771.
[11] D'Alessio, C. and Dahms, N.M. (2015) Glucosidase II and MRH-domain containing proteins in the secretory pathway. Curr. Protein Pept. Sci. 16, 31–48.
[12] Trombetta, E.S., Simons, J.F. and Helenius, A. (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.
[13] Trombetta, E.S., Fleming, K.G. and Helenius, A. (2001) Quaternary and domain structure of glycoprotein processing glucosidase II. Biochemistry 40, 10717–10722.
[14] Olson, L.J., Orsi, R., Peterson, F.C., Parodi, A.J., Kim, J.J., D'Alessio, C. and Dahms, N.M. (2015) Crystal Structure and Functional Analyses of the Lectin Domain of Glucosidase II: Insights into Oligomannose Recognition. Biochemistry.
[15] Hu, D. et al. (2009) Sugar-binding activity of the MRH domain in the ER alpha-glucosidase II beta subunit is important for efficient glucose trimming. Glycobiology 19, 1127–1135.
[16] Stigliano, I.D., Caramelo, J.J., Labriola, C.A., Parodi, A.J. and D'Alessio, C. (2009) Glucosidase II beta subunit modulates N-glycan trimming in fission yeasts and mammals. Mol. Biol. Cell 20, 3974–3984.
[17] Watanabe, T., Totani, K., Matsuo, I., Maruyama, J., Kitamoto, K. and Ito, Y. (2009) Genetic analysis of glucosidase II beta-subunit in trimming of high-mannose-type glycans. Glycobiology 19, 834–840.
[18] Bosis, E., Nachliel, E., Cohen, T., Takeda, Y., Ito, Y., Bar-Nun, S. and Gutman, M. (2008) Endoplasmic reticulum glucosidase II is inhibited by its end products. Biochemistry 47, 10970–10980.
[19] Totani, K., Ihara, Y., Matsuo, I. and Ito, Y. (2006) Substrate specificity analysis of endoplasmic reticulum glucosidase II using synthetic high mannose-type glycans. J. Biol. Chem. 281, 31502–31508.
[20] Totani, K., Ihara, Y., Matsuo, I. and Ito, Y. (2008) Effects of macromolecular crowding on glycoprotein processing enzymes. J. Am. Chem. Soc. 130, 2101–2107.
[21] Chang, J., Block, T.M. and Guo, J.T. (2013) Antiviral therapies targeting host ER alpha-glucosidases: current status and future directions. Antiviral Res. 99, 251–260.

Please cite this article in press as: Caramelo, J.J. and Parodi, A.J. A sweet code for glycoprotein folding. FEBS Lett. (2015), http://dx.doi.org/10.1016/j.febslet.2015.07.021

[22] Raghavan, M., Wijeyesakere, S.J., Peters, L.R. and Del Cid, N. (2013) Calreticulin in the immune system: ins and outs. Trends Immunol. 34, 13–21.
[23] Gold, L.I., Eggleton, P., Sweetwyne, M.T., Van Duyn, L.B., Greives, M.R., Naylor, S.M., Michalak, M. and Murphy-Ullrich, J.E. (2010) Calreticulin: non-endoplasmic reticulum functions in physiology and disease. FASEB J. 24, 665–683.
[24] Kapoor, M., Srinivas, H., Kandiah, E., Gemma, E., Ellgaard, L., Oscarson, S., Helenius, A. and Surolia, A. (2003) Interactions of substrate with calreticulin, an endoplasmic reticulum chaperone. J. Biol. Chem. 278, 6194–6200.
[25] Matsushima, H., Hirano, M., Ito, Y. and Totani, K. (2013) Diverse effects of macromolecular crowding on the sequential glycan-processing pathway involved in glycoprotein quality control. ChemBioChem 14, 753–758.
[26] Wada, I., Imai, S., Kai, M., Sakane, F. and Kanoh, H. (1995) Chaperone function of calreticulin when expressed in the endoplasmic reticulum as the membrane-anchored and soluble forms. J. Biol. Chem. 270, 20298–20304.
[27] Hebert, D.N., Zhang, J.X., Chen, W., Foellmer, B. and Helenius, A. (1997) The number and location of glycans on influenza hemagglutinin determine folding and association with calnexin and calreticulin. J. Cell Biol. 139, 613–623.
[28] Molinari, M. and Helenius, A. (2000) Chaperone selection during glycoprotein translocation into the endoplasmic reticulum. Science 288, 331–333.
[29] Wang, N., Daniels, R. and Hebert, D.N. (2005) The cotranslational maturation of the type I membrane glycoprotein tyrosinase: the heat shock protein 70 system hands off to the lectin-based chaperone system. Mol. Biol. Cell 16, 3740–3752.
[30] Labriola, C.A., Giraldo, A.M., Parodi, A.J. and Caramelo, J.J. (2011) Functional cooperation between BiP and calreticulin in the folding maturation of a glycoprotein in Trypanosoma cruzi. Mol. Biochem. Parasitol. 175, 112–117.
[31] Kim, P.S. and Arvan, P. (1995) Calnexin and BiP act as sequential molecular chaperones during thyroglobulin folding in the endoplasmic reticulum. J. Cell Biol. 128, 29–38.
[32] Daniels, R., Kurowski, B., Johnson, A.E. and Hebert, D.N. (2003) N-linked glycans direct the cotranslational folding pathway of influenza hemagglutinin. Mol. Cell 11, 79–90.
[33] Wang, N., Glidden, E.J., Murphy, S.R., Pearse, B.R. and Hebert, D.N. (2008) The cotranslational maturation program for the type II membrane glycoprotein influenza neuraminidase. J. Biol. Chem. 283, 33826–33837.
[34] Williams, D.B. (2006) Beyond lectins: the calnexin/calreticulin chaperone system of the endoplasmic reticulum. J. Cell Sci. 119, 615–623.
[35] Chouquet, A., Paidassi, H., Ling, W.L., Frachet, P., Houen, G., Arlaud, G.J. and Gaboriaud, C. (2011) X-ray structure of the human calreticulin globular domain reveals a peptide-binding area and suggests a multi-molecular mechanism. PLoS ONE 6, e17886.
[36] Thammavongsa, V., Mancino, L. and Raghavan, M. (2005) Polypeptide substrate recognition by calnexin requires specific conformations of the calnexin protein. J. Biol. Chem. 280, 33497–33505.
[37] Wijeyesakere, S.J., Rizvi, S.M. and Raghavan, M. (2013) Glycan-dependent and -independent interactions contribute to cellular substrate recruitment by calreticulin. J. Biol. Chem. 288, 35104–35116.
[38] Conte, I.L., Keith, N., Gutierrez-Gonzalez, C., Parodi, A.J. and Caramelo, J.J. (2007) The interplay between calcium and the in vitro lectin and chaperone activities of calreticulin. Biochemistry 46, 4671–4680.
[39] Villamil Giraldo, A.M. et al. (2010) The structure of calreticulin C-terminal domain is modulated by physiological variations of calcium concentration. J. Biol. Chem. 285, 4544–4553.
[40] Wijeyesakere, S.J., Gafni, A.A. and Raghavan, M. (2011) Calreticulin is a thermostable protein with distinct structural responses to different divalent cation environments. J. Biol. Chem. 286, 8771–8785.
[41] Afshar, N., Black, B.E. and Paschal, B.M. (2005) Retrotranslocation of the chaperone calreticulin from the endoplasmic reticulum lumen to the cytosol. Mol. Cell. Biol. 25, 8844–8853.
[42] Labriola, C.A., Conte, I.L., LopezMedus, M., Parodi, A.J. and Caramelo, J.J. (2010) Endoplasmic reticulum calcium regulates the retrotranslocation of Trypanosoma cruzi calreticulin to the cytosol. PLoS ONE.
[43] Decca, M.B., Carpio, M.A., Bosc, C., Galiano, M.R., Job, D., Andrieux, A. and Hallak, M.E. (2007) Post-translational arginylation of calreticulin: a new isospecies of calreticulin component of stress granules. J. Biol. Chem. 282, 8237–8245.
[44] Goitea, V.E. and Hallak, M.E. (2015) Calreticulin and arginylated calreticulin have different susceptibility to proteasomal degradation. J. Biol. Chem..
[45] Grespin, M.E., Bonamy, G.M., Roggero, V.R., Cameron, N.G., Adam, L.E., Atchison, A.P., Fratto, V.M. and Allison, L.A. (2008) Thyroid hormone receptor alpha1 follows a cooperative CRM1/calreticulin-mediated nuclear export pathway. J. Biol. Chem. 283, 25576–25588.
[46] Mueller, C.F., Wassmann, K., Berger, A., Holz, S., Wassmann, S. and Nickenig, G. (2008) Differential phosphorylation of calreticulin affects AT1 receptor mRNA stability in VSMC. Biochem. Biophys. Res. Commun. 370, 669–674.
[47] Mesaeli, N. et al. (1999) Calreticulin is essential for cardiac development. J. Cell Biol. 144, 857–868.
[48] Guo, L., Nakamura, K., Lynch, J., Opas, M., Olson, E.N., Agellon, L.B. and Michalak, M. (2002) Cardiac-specific expression of calcineurin reverses embryonic lethality in calreticulin-deficient mouse. J. Biol. Chem. 277, 50776–50779.
[49] Denzel, A., Molinari, M., Trigueros, C., Martin, J.E., Velmurgan, S., Brown, S., Stamp, G. and Owen, M.J. (2002) Early postnatal death and motor disorders in mice congenitally deficient in calnexin expression. Mol. Cell. Biol. 22, 7398–7404.
[50] Kraus, A. et al. (2010) Calnexin deficiency leads to dysmyelination. J. Biol. Chem. 285, 18928–18938.
[51] Maattanen, P., Kozlov, G., Gehring, K. and Thomas, D.Y. (2006) ERp57 and PDI: multifunctional protein disulfide isomerases with similar domain architectures but differing substrate-partner associations. Biochem. Cell Biol. 84, 881–889.
[52] Jessop, C.E., Tavender, T.J., Watkins, R.H., Chambers, J.E. and Bulleid, N.J. (2009) Substrate specificity of the oxidoreductase ERp57 is determined primarily by its interaction with calnexin and calreticulin. J. Biol. Chem. 284, 2194–2202.
[53] Coe, H., Schneider, J.D., Dabrowska, M., Groenendyk, J., Jung, J. and Michalak, M. (2012) Role of cysteine amino acid residues in calnexin. Mol. Cell. Biochem. 359, 271–281.
[54] Kozlov, G. et al. (2010) Structural basis of cyclophilin B binding by the calnexin/calreticulin P-domain. J. Biol. Chem. 285, 35551–35557.
[55] Lynes, E.M. et al. (2013) Palmitoylation is the switch that assigns calnexin to quality control or ER Ca2+ signaling. J. Cell Sci. 126, 3893–3903.
[56] Lakkaraju, A.K. and van der Goot, F.G. (2013) Calnexin controls the STAT3-mediated transcriptional response to EGF. Mol. Cell 51, 386–396.
[57] Solda, T., Galli, C., Kaufman, R.J. and Molinari, M. (2007) Substrate-specific requirements for UGT1-dependent release from calnexin. Mol. Cell 27, 238–249.
[58] Ferris, S.P., Jaber, N.S., Molinari, M., Arvan, P. and Kaufman, R.J. (2013) UDP-glucose:glycoprotein glucosyltransferase (UGGT1) promotes substrate solubility in the endoplasmic reticulum. Mol. Biol. Cell 24, 2597–2608.
[59] Sousa, M.C., Ferrero-Garcia, M.A. and Parodi, A.J. (1992) Recognition of the oligosaccharide and protein moieties of glycoproteins by the UDP-Glc:glycoprotein glucosyltransferase. Biochemistry 31, 97–105.
[60] Molinari, M., Galli, C., Vanoni, O., Arnold, S.M. and Kaufman, R.J. (2005) Persistent glycoprotein misfolding activates the glucosidase II/UGT1-driven calnexin cycle to delay aggregation and loss of folding competence. Mol. Cell 20, 503–512.
[61] Mitchell, K.J. et al. (2001) Functional analysis of secreted and transmembrane proteins critical to mouse development. Nat. Genet. 28, 241–249.
[62] Fanchiotti, S., Fernandez, F., D'Alessio, C. and Parodi, A.J. (1998) The UDP-Glc: Glycoprotein glucosyltransferase is essential for Schizosaccharomyces pombe viability under conditions of extreme endoplasmic reticulum stress. J. Cell Biol. 143, 625–635.
[63] Parodi, A.J. and Cazzulo, J.J. (1982) Protein glycosylation in Trypanosoma cruzi. II. Partial characterization of protein-bound oligosaccharides labeled "in vivo". J. Biol. Chem. 257, 7641–7645.
[64] Sousa, M. and Parodi, A.J. (1995) The molecular basis for the recognition of misfolded glycoproteins by the UDP-Glc:glycoprotein glucosyltransferase. EMBO J. 14, 4196–4203.
[65] Caramelo, J.J., Castro, O.A., Alonso, L.G., De Prat-Gay, G. and Parodi, A.J. (2003) UDP-Glc:glycoprotein glucosyltransferase recognizes structured and solvent accessible hydrophobic patches in molten globule-like folding intermediates. Proc. Natl. Acad. Sci. U.S.A 100, 86–91.
[66] Caramelo, J.J., Castro, O.A., de Prat-Gay, G. and Parodi, A.J. (2004) The endoplasmic reticulum glucosyltransferase recognizes nearly native glycoprotein folding intermediates. J. Biol. Chem. 279, 46280–46285.
[67] Izumi, M., Makimura, Y., Dedola, S., Seko, A., Kanamori, A., Sakono, M., Ito, Y. and Kajihara, Y. (2012) Chemical synthesis of intentionally misfolded homogeneous glycoprotein: a unique approach for the study of glycoprotein quality control. J. Am. Chem. Soc. 134, 7238–7241.
[68] Keith, N., Parodi, A.J. and Caramelo, J.J. (2005) Glycoprotein tertiary and quaternary structures are monitored by the same quality control mechanism. J. Biol. Chem. 280, 18138–18141.
[69] Taylor, S.C., Thibault, P., Tessier, D.C., Bergeron, J.J. and Thomas, D.Y. (2003) Glycopeptide specificity of the secretory protein folding sensor UDP-glucose glycoprotein:glucosyltransferase. EMBO Rep. 4, 405–411.
[70] Totani, K., Ihara, Y., Tsujimoto, T., Matsuo, I. and Ito, Y. (2009) The recognition motif of the glycoprotein-folding sensor enzyme UDP-Glc:glycoprotein glucosyltransferase. Biochemistry 48, 2933–2940.
[71] Ritter, C. and Helenius, A. (2000) Recognition of local glycoprotein misfolding by the ER folding sensor UDP-glucose:glycoprotein glucosyltransferase. Nat. Struct. Biol. 7, 278–280.
[72] Taylor, S.C., Ferguson, A.D., Bergeron, J.J. and Thomas, D.Y. (2004) The ER protein folding sensor UDP-glucose glycoprotein-glucosyltransferase modifies substrates distant to local changes in glycoprotein conformation. Nat. Struct. Mol. Biol. 11, 128–134.
[73] Buzzi, L.I., Simonetta, S.H., Parodi, A.J. and Castro, O.A. (2011) The two Caenorhabditis elegans UDP-glucose:glycoprotein glucosyltransferase homologues have distinct biological functions. PLoS ONE 6, e27025.
[74] Takeda, Y. et al. (2014) Both isoforms of human UDP-glucose:glycoprotein glucosyltransferase are enzymatically active. Glycobiology 24, 344–350.
[75] Pearse, B.R., Tamura, T., Sunryd, J.C., Grabowski, G.A., Kaufman, R.J. and Hebert, D.N. (2010) The role of UDP-Glc:glycoprotein glucosyltransferase 1 in the maturation of an obligate substrate prosaposin. J. Cell Biol. 189, 829–841.
[76] Labriola, C., Cazzulo, J.J. and Parodi, A.J. (1995) Retention of glucose units added by the UDP-GLC:glycoprotein glucosyltransferase delays exit of glycoproteins from the endoplasmic reticulum. J. Cell Biol. 130, 771–779.
[77] Wearsch, P.A., Peaper, D.R. and Cresswell, P. (2011) Essential glycan-dependent interactions optimize MHC class I peptide loading. Proc. Natl. Acad. Sci. U.S.A. 108, 4950–4955.

Please cite this article in press as: Caramelo, J.J. and Parodi, A.J. A sweet code for glycoprotein folding. FEBS Lett. (2015), http://dx.doi.org/10.1016/j.febslet.2015.07.021

[78] Korotkov, K.V., Kumaraswamy, E., Zhou, Y., Hatfield, D.L. and Gladyshev, V.N. (2001) Association between the 15-kDa selenoprotein and UDP-glucose:glycoprotein glucosyltransferase in the endoplasmic reticulum of mammalian cells. J. Biol. Chem. 276, 15330–15336.
[79] Labunskyy, V.M., Ferguson, A.D., Fomenko, D.E., Chelliah, Y., Hatfield, D.L. and Gladyshev, V.N. (2005) A novel cysteine-rich domain of Sep15 mediates the interaction with UDP-glucose:glycoprotein glucosyltransferase. J. Biol. Chem. 280, 37839–37845.
[80] Labunskyy, V.M., Yoo, M.H., Hatfield, D.L. and Gladyshev, V.N. (2009) Sep15, a thioredoxin-like selenoprotein, is involved in the unfolded protein response and differentially regulated by adaptive and acute ER stresses. Biochemistry 48, 8458–8465.
[81] Kasaikina, M.V. et al. (2011) Roles of the 15-kDa selenoprotein (Sep15) in redox homeostasis and cataract development revealed by the analysis of Sep 15 knockout mice. J. Biol. Chem. 286, 33203–33212.
[82] Stigliano, I.D., Alculumbre, S.G., Labriola, C.A., Parodi, A.J. and D'Alessio, C. (2011) Glucosidase II and N-glycan mannose content regulate the half-lives of monoglucosylated species in vivo. Mol. Biol. Cell 22, 1810–1823.
[83] Tannous, A., Patel, N., Tamura, T. and Hebert, D.N. (2015) Reglucosylation by UDP-glucose:glycoprotein glucosyltransferase 1 delays glycoprotein secretion but not degradation. Mol. Biol. Cell 26, 390–405.
[84] Zhu, T., Satoh, T. and Kato, K. (2014) Structural insight into substrate recognition by the endoplasmic reticulum folding-sensor enzyme: crystal structure of third thioredoxin-like domain of UDP-glucose:glycoprotein glucosyltransferase. Sci. Rep. 4, 7322.
[85] Huyer, G., Longsworth, G.L., Mason, D.L., Mallampalli, M.P., McCaffery, J.M., Wright, R.L. and Michaelis, S. (2004) A striking quality control subcompartment in Saccharomyces cerevisiae: the endoplasmic reticulum-associated compartment. Mol. Biol. Cell 15, 908–921.
[86] Avezov, E., Frenkel, Z., Ehrlich, M., Herscovics, A. and Lederkremer, G.Z. (2008) Endoplasmic reticulum (ER) mannosidase I is compartmentalized and required for N-glycan trimming to Man5-6GlcNAc2 in glycoprotein ER-associated degradation. Mol. Biol. Cell 19, 216–225.
[87] Benyair, R., Ogen-Shtern, N., Mazkereth, N., Shai, B., Ehrlich, M. and Lederkremer, G.Z. (2015) Mammalian ER mannosidase I resides in quality control vesicles, where it encounters its glycoprotein substrates. Mol. Biol. Cell 26, 172–184.
[88] Benyair, R., Ogen-Shtern, N. and Lederkremer, G.Z. (2014) Glycan regulation of ER-associated degradation through compartmentalization. Semin. Cell Dev. Biol..
[89] Tamura, T., Cormier, J.H. and Hebert, D.N. (2011) Characterization of early EDEM1 protein maturation events and their functional implications. J. Biol. Chem. 286, 24906–24915.
[90] Ushioda, R., Hoseki, J., Araki, K., Jansen, G., Thomas, D.Y. and Nagata, K. (2008) ERdj5 is required as a disulfide reductase for degradation of misfolded proteins in the ER. Science 321, 569–572.
[91] Hagiwara, M. and Nagata, K. (2012) Redox-dependent protein quality control in the endoplasmic reticulum: folding to degradation. Antioxid. Redox Signal. 16, 1119–1128.
[92] Williams, J.M., Inoue, T., Banks, L. and Tsai, B. (2013) The ERdj5-Sel1L complex facilitates cholera toxin retrotranslocation. Mol. Biol. Cell 24, 785–795.
[93] Frenkel, Z., Gregory, W., Kornfeld, S. and Lederkremer, G.Z. (2003) Endoplasmic reticulum-associated degradation of mammalian glycoproteins involves sugar chain trimming to Man6-5GlcNAc2. J. Biol. Chem. 278, 34119–34124.
[94] Hosokawa, N., Tremblay, L.O., Sleno, B., Kamiya, Y., Wada, I., Nagata, K., Kato, K. and Herscovics, A. (2010) EDEM1 accelerates the trimming of alpha1,2-linked mannose on the C branch of N-glycans. Glycobiology 20, 567–575.
[95] Aikawa, J., Takeda, Y., Matsuo, I. and Ito, Y. (2014) Trimming of glucosylated N-glycans by human ER alpha1,2-mannosidase I. J. Biochem. 155, 375–384.
[96] Mikami, K. et al. (2010) The sugar-binding ability of human OS-9 and its involvement in ER-associated degradation. Glycobiology 20, 310–321.
[97] Hosokawa, N., Kamiya, Y., Kamiya, D., Kato, K. and Nagata, K. (2009) Human OS-9, a lectin required for glycoprotein endoplasmic reticulum-associated degradation, recognizes mannose-trimmed N-glycans. J. Biol. Chem. 284, 17061–17068.
[98] Groisman, B., Shenkman, M., Ron, E. and Lederkremer, G.Z. (2011) Mannose trimming is required for delivery of a glycoprotein from EDEM1 to XTP3-B and to late endoplasmic reticulum-associated degradation steps. J. Biol. Chem. 286, 1292–1300.
[99] Yamaguchi, D., Hu, D., Matsumoto, N. and Yamamoto, K. (2010) Human XTP3-B binds to alpha1-antitrypsin variant null(Hong Kong) via the C-terminal MRH domain in a glycan-dependent manner. Glycobiology 20, 348–355.
[100] Fujimori, T., Kamiya, Y., Nagata, K., Kato, K. and Hosokawa, N. (2013) Endoplasmic reticulum lectin XTP3-B inhibits endoplasmic reticulum-associated degradation of a misfolded alpha1-antitrypsin variant. FEBS J. 280, 1563–1575.
[101] Movsichoff, F., Castro, O.A. and Parodi, A.J. (2005) Characterization of *Schizosaccharomyces pombe* ER α-mannosidase. A re-evaluation of the role of the enzyme on ER-associated degradation. Mol. Biol. Cell 16, 4714–4724.
[102] Ruggiano, A., Foresti, O. and Carvalho, P. (2014) ER-asssociated degradation: protein quality control and beyond. J. Cell Biol. 204, 869–879.
[103] Li, G., Zhou, C.X., Zhao, G., Schindelin, H. and Lennarz, W.J. (2005) Multiple forms of interaction of the deglycosylation enzyme, mouse peptide:N-glycanase with the proteasome. Proc. Natl. Acad. Sci. U.S.A. 102, 15809–15814.
[104] Joshi, S., Katiyar, S. and Lennarz, W.J. (2005) Misfolding of glycoproteins is a prerequisite for peptide:*N*-glycanase mediated deglycosylation. FEBS Lett. 579, 823–826.
[105] Suzuki, T., Yano, K., Sugimoto, S., Kitajima, K., Lennarz, W.J., Inoue, S., Inoue, Y. and Emory, Y. (2002) Endo-beta-N-acetylglucosaminidase, an enzyme involved in processing of free oligosaccharides in the cytosol. Proc. Natl. Acad. Sci. U.S.A. 99, 9691–9696.
[106] Moore, S.E.H. (1999) Oligosaccharide transport: pumping waste from the ER into lysosomes. Trends Cell Biol. 9, 441–446.

Please cite this article in press as: Caramelo, J.J. and Parodi, A.J. A sweet code for glycoprotein folding. FEBS Lett. (2015), http://dx.doi.org/10.1016/j.febslet.2015.07.021