
UDP-Glucose: Glycoprotein Glucosyltransferase in ER Glycoprotein Quality Control

110

Cecilia D'Alessio and Armando J. Parodi

Contents

Introduction	914
The Reaction-Substrates and Assays	915
The Enzyme Protein	916
Glucosidase II, the Enzyme Opposing UGGT Activity	917
Genes Encoding UGGT	917
UGGT Relevance for Cell Functions	918
References	919

Abstract

The *N*-glycan-dependent quality control of glycoprotein folding prevents endoplasmic reticulum (ER) to Golgi exit 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, a 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

C. D'Alessio (✉) • A.J. Parodi (✉)

Laboratory of Glycobiology, Fundación Instituto Leloir, Buenos Aires, Argentina

e-mail: cdalessio@leloir.org.ar; aparodi@leloir.org.ar

© Springer Japan 2015

N. Taniguchi et al. (eds.), *Glycoscience: Biology and Medicine*,

DOI 10.1007/978-4-431-54841-6_156

913

glycoprotein complexes. The glucosidase is a heterodimer whose activity diminishes as ER mannosidases remove mannose units from *N*-glycans. This mini review deals with our present knowledge on how the mechanism operates within the secretory pathway.

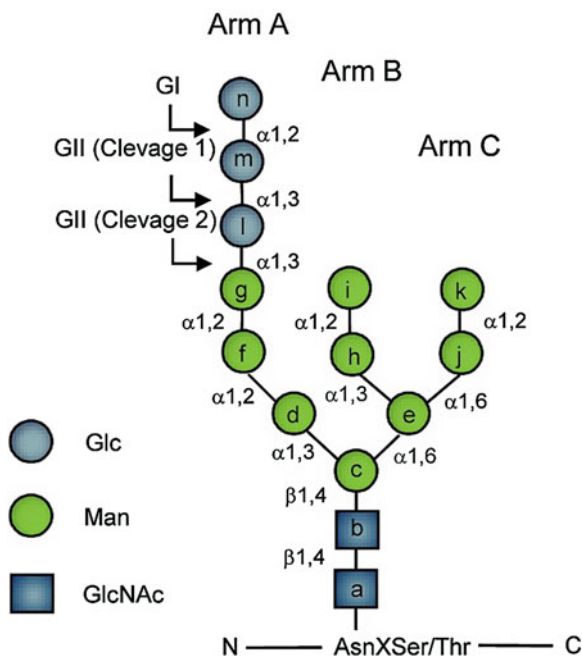
Keywords

Endoplasmic reticulum • Secretory pathway • Folding quality control • UDP-glucose/glycoprotein glucosyltransferase • Glucosidase II • Calnexin • Calreticulin • ERp57 • BiP • Protein tertiary and quaternary structures

Introduction

About one-third of proteins synthesized in a mammalian cell follow the secretory pathway and nearly 80 % of those entering this pathway are *N*-glycosylated in the endoplasmic reticulum (ER). In most cases a glycan (Glc₃Man₉GlcNAc₂, Fig. 1) is transferred either co-translationally or post-translationally from a lipid (dolichol)-P-P derivative to Asn residues in the sequon Asn-X-Ser/Thr, where X may be any amino acid except for Pro. Glycans transferred are then processed by removal and addition of monosaccharides. Folding of the protein moieties takes place simultaneously with glycan processing. Immediately after transfer, the more external glucose (residue *n*, Fig. 1) is removed by glucosidase I (GI), a membrane-bound enzyme. Successive removal of additional glucose units (first residue *m* and then *l*, Fig. 1) then proceeds by reactions catalyzed by glucosidase II (GII), a soluble enzyme. Removal of mannose units *i* and *k* (Fig. 1) may then occur in reactions catalyzed by ER mannosidases. It is worth mentioning that demannosylation is in most cells a process that is relatively slower than deglycosylation. Once the first two glucoses have been removed (residues *n* and *m*, Fig. 1) as described above, the Glc₁Man₉GlcNAc₂ structure thus created may be specifically recognized by two ER-resident lectins calnexin (CNX), a membrane-bound protein and its soluble homologue calreticulin (CRT). Interaction of monoglucosylated structures with the lectins has a dual effect: on one hand it prevents exit of folding intermediates and misfolded glycoproteins to the Golgi, and on the other, it enhances folding efficiency by preventing aggregation. Furthermore, the lectin-glycoprotein interaction facilitates formation of the correct disulfide bonds by a CNX-/CRT-associated oxidoreductase (ERp57). Upon removal of the innermost glucose unit by GII, the interaction between glycoproteins and lectins is lost. If the protein moieties have attained their native conformations, glycoproteins are secreted to the Golgi. If not, they may be reglycosylated by the UDP-Glc/glycoprotein glucosyltransferase (UGGT), thus recreating the monoglucosylated structures and thence the interactions with the lectins. The UGGT forms part, therefore, of the mechanism by which the structure of protein-linked glycans provides the cell information on the conformation of the protein moieties. Depending on the information received from protein-linked glycans, cells may then either allow further transit of glycoproteins through the Alactory secretory pathway, or retain them in the ER for further folding

Fig. 1 Structure of *N*-glycans. The structure is that of the full-length glycan transferred to Asn residues from dolichol-P-P derivatives upon *N*-glycosylation in most wild-type eukaryotic cells. Lettering (*a, b, c, ...*) follows the order of addition of monosaccharides in the buildup of dolichol-P-P derivatives



attempts or, alternatively, retrotranslocate them to the cytosol for proteasomal degradation in a process known as ERAD (ER-associated degradation) (Hammond et al. 1994; Parodi 2000; Helenius and Aebi 2004; D'Alessio et al. 2010).

The Reaction-Substrates and Assays

UGGT transfers a single glucose unit from UDP-Glc to residue *g* (Fig. 1) in protein-linked high mannose type glycans. $\text{Man}_9\text{GlcNAc}_2$, $\text{Man}_8\text{GlcNAc}_2$, and $\text{Man}_7\text{GlcNAc}_2$ serve equally well as substrates provided they display the acceptor mannose unit (residue *g*, Fig. 1). UGGT recreates the same α(1,3) bond present in the partially deglycosylated glycan transferred from the lipid derivative. The reaction requires millimolar calcium concentrations for activity, a condition normally present in the ER lumen. K_m values of 18–40 μM for UDP-Glc have been reported for UGGTs from different species. UDP-Gal, TDP-Glc, and ADP-Glc were ineffective in the case of the *Schizosaccharomyces pombe* enzyme (Parodi 2000; D'Alessio et al. 2010). There are several different in vitro assays for UGGT. The most commonly used involves measuring label incorporated into trichloroacetic acid insoluble material formed upon incubation of UDP-Glc labeled in the glucose moiety, calcium ions, a denatured glycoprotein as thyroglobulin or soybean agglutinin as acceptor, a neutral buffer, and the enzymatic preparation. Alternatively, the reaction mixture may be submitted to SDS-PAGE followed by autoradiography. There is firm evidence obtained both in cell free assays and in

intact cells indicating that the protein moiety in the acceptor substrate has to display a collapsed, molten globule-like, non-native conformation thus exposing hydrophobic patches that are recognized by the enzyme (Caramelo et al. 2003). It has been shown that in intact cells, glycoproteins bearing several disulfide bonds are glucosylated by UGGT once most of such bonds have been already formed, that is, in the last folding stages. Relatively short glycopeptides may serve also as acceptors if rich in hydrophobic amino acids. Unstructured, random coil glycoproteins, that is, not displaying hydrophobic patches, are not recognized by UGGT. Further evidence gathered in intact cells showed that peptides entering the ER lumen are first recognized mainly by a chaperone of the Hsp70 family (BiP) while nonstructured, and it is only after collapsing that binding to this last chaperone is lost and UGGT-dependent CNX/CRT binding proceeds. UGGT also recognizes the innermost *N*-acetylglucosamine unit in the glycan (residue α , Fig. 1). As this residue is normally hidden in native but not in molten globule-like conformations, this feature provides an additional impediment for glucosylation of native structures. UGGT participates not only in the folding process of monomeric glycoproteins but also in the complete formation of multimeric complexes: incompletely formed complexes, even though composed of correctly folded monomers, may be glucosylated by the enzyme (and thence bound by CNX/CRT) upon UGGT recognition of hydrophobic protein interphases normally hidden in the complete complexes. As expected UGGT-mediated glucosylation ceases upon complete formation of the complexes (Parodi 2000; D'Alessio et al. 2010).

The Enzyme Protein

UGGT is an ER-soluble protein of about 150,000–160,000 kDa. It is present in most eukaryotic cells that transfer high mannose glycans upon *N*-glycosylation with the notable exception of *Saccharomyces cerevisiae* and organisms that transfer extremely short versions of the glycan as *Giardia lamblia* and *Plasmodium falciparum* (Parodi 2000; D'Alessio et al. 2010). The enzyme is composed by at least two distinct structural domains. The *N*-terminal one comprises about 80 % of the molecule and it is assumed that it is responsible for recognition of the folding status of the protein moiety although there is not a definitive evidence for this property. The C-terminal domain bears the catalytic site and displays a KDEL-like sequence at its C-terminus in most species, a feature characteristic of ER-ERGIC (ER-Golgi intermediate compartment) resident soluble proteins. A notable exception is the protozoon *Trypanosoma cruzi* UGGT, an enzyme that lacks that sequence in spite of being a soluble protein and having the correct subcellular location. The C-terminal domain shows a high degree of homology between UGGTs from different species (about 60–70 % identity), whereas no such homology is found in the N-terminal domains. For instance, the N-terminal domains from *S. pombe* and *Drosophila melanogaster* UGGTs only share a 16.5 % similarity. Nevertheless, switching both domains in chimeric enzymes expressed in the fission yeast was found to be active in vivo. This result suggests common features in the

tertiary structures of N-terminal domains displaying highly divergent primary sequences. The enzyme is relatively easy to purify with relatively high yields from several sources as rat liver, *S. pombe*, and others. A standard procedure involves extraction of the soluble content from a microsomal preparation followed by DEAE-cellulose, ConA-Sepharose, monoQ, gel filtration, and finally phenyl-Superose chromatographies (D'Alessio et al. 2010).

Glucosidase II, the Enzyme Opposing UGGT Activity

It is necessary to describe certain features of GII, the enzyme whose activity opposes that of UGGT to get a full information on the influence of the last enzyme on glycoprotein folding. GII is a soluble heterodimeric protein. GII α bears the catalytic activity, whereas GII β has a dual function: on one hand, it displays a KDEL-like sequence at its C-terminus, thus ensuring a similar subcellular location as UGGT, and on the other, it expresses an MRH (mannose receptor homologous) domain. This domain presents significant homology with those of the Golgi receptors that recognize the Man 6-P epitopes present in mammalian lysosomal enzymes. The presence of the MRH domain in GII is required for an efficient catalytic reaction. GII β MRH domain showed a decreasing affinity for *N*-glycans as mannose units were removed from them by ER mannosidases (first residue *i* followed by *k*, Fig. 1) and thence GII activity also diminished. As mentioned above, demannosylation is a relatively slow process when compared to deglycosylation, and glycoproteins that are thus processed are mainly those that show longer permanences in the ER due to folding difficulties. Demannosylation then results in longer half-lives of monoglucosylated structures, as this process does not affect UGGT activity. Therefore, demannosylation provides glycoproteins with additional chances for proper folding mediated by UGGT-dependent interactions with CNX/CRT (D'Alessio et al. 2010; Stigliano et al. 2011).

Genes Encoding UGGT

Whereas in *S. pombe*, *D. melanogaster*, *T. cruzi*, and plants an enzymatically active UGGT is encoded by a single gene, there are two homologous genes coding for UGGT-like proteins in Euteleostomi, which is a successful clade that includes more than 90 % of the living species of vertebrates and at least in some species of nematodes belonging to the genus *Caenorhabditis*. It has been reported that disruption of genes coding for either UGGT, GII α , or GII β in the fission yeast triggers the unfolded protein response. This is to be expected as the proteins coded by those genes are involved in ER glycoprotein folding facilitation. The yeast *S. cerevisiae* genome codes for a single protein with UGGT homology (Kre5p) that lacks canonical enzymatic activity (Parodi 2000; D'Alessio et al. 2010). As mentioned above, humans have two homologous genes coding for UGGT-like proteins that share a 55 % identity, hUGGT1 and hUGGT2, but only the former appeared to be

enzymatically active. ER stress triggered by tunicamycin or the ionophore A23187 induced the expression of hUGGT1 but not hUGGT2. The biological function of hUGGT2 is unknown, but a chimeric protein consisting of the N-terminal domain of hUGGT1 and the C-terminal domain of UGGT2 is active, suggesting that UGGT2 has evolved to glycosylate a restricted set of glycoproteins (Arnold and Kaufman 2003). It may be speculated that hUGGT2 recognizes folding intermediates displaying structures not present in the denatured glycoprotein substrates commonly used in the standard in vitro assays because, as mentioned above, the N-terminal domain putatively recognizes the folding status of protein moieties, whereas the C-terminal one bears the catalytic site. Indeed, it has been shown recently that hUGGT2 does display a canonical glucosyltransferase activity when assayed using synthetic substrates as fluorescently labeled glycans as acceptors (Takeda et al. 2014). Both genes coding for UGGT homologues were found to be expressed in *C. elegans* (CeUGGT1 and CeUGGT2), but only one of them (CeUGGT1) showed enzymatic activity when expressed in *S. pombe* UGGT minus cells and further assayed with denatured thyroglobulin as acceptor. RNAi-mediated depletion of CeUGGT1 but not of CeUGGT2 resulted in a reduced life span, although depletion of both proteins resulted in a developmental delay. Both CeUGGT1 and CeUGGT2 played a protective role under ER stress conditions. Furthermore, CeUGGT2 alleviated endogenous ER stress during development. The fact that *cnx-1 crt-1* double null *C. elegans* mutants (i.e., mutants lacking both ER-resident lectins and thus unable to recognize monoglucosylated glycans) are viable, while expression of CeUGGT2 is required for viability indicated that both *C. elegans* have not identical biological functions (Buzzi et al. 2011).

UGGT Relevance for Cell Functions

UGGT is not essential for viability under physiological conditions in unicellular organisms. As mentioned above, knocking out or down expression of the enzyme results in the upregulation of ER chaperones and other folding assisting proteins that compensate for the absence of UGGT. Nevertheless, it was reported that in *S. pombe*, UGGT is indeed essential under extreme stress conditions as a generalized protein underglycosylation and high temperature (Parodi 2000; D'Alessio et al. 2010). This result points to the existence of a glycoprotein(s) essential for viability that necessarily requires UGGT-mediated interaction with CNX/CRT for proper folding under those conditions. Knocking out UGGT in the protozoan parasite *T. cruzi* resulted in a reduced infectivity (D'Alessio et al. 2010). Disruption of the UGGT encoding gene in *Arabidopsis thaliana* did not lead to an obvious phenotype but, as expected, several ER chaperones and folding assisting enzymes such as BiP and PDI were upregulated. However, UGGT knock out is embryonically lethal in mice but mouse embryo fibroblasts derived from those embryos are viable. Two experimental results will be mentioned here, as they point to the rather subtle influence of UGGT in cell functions: first, the ability to recognize minor folding defects was illustrated in *A. thaliana* in which the UGGT gene was isolated

using a complementation genetic screen aiming to restore the normal growth of a plant expressing a mutant brassinosteroid receptor (*bri1-9*). The phenotype due to the *bri1-9* mutation arises from receptor retention in the ER by CNX. Interestingly, upon disruption of the UGGT encoding gene, the mutant receptor reached the plasma membrane in a functional form, thus showing that the structural perturbation that triggered UGGT recognition did not involve the receptor activity (Jin et al. 2007). This observation suggests that some disease resulting in the ER retention of defective glycoproteins could be ameliorated by inhibitors of UGGT activity. Finally, it was shown that major histocompatibility complexes I (MHC I) loaded with suboptimal peptides were reglucosylated by UGGT, a reaction that resulted in their CNX-/CRT-mediated ER retention. On the other hand, complexes formed with high-affinity peptides were poor UGGT substrates and were transported to the cell surface (Zhang et al. 2011).

References

- Arnold SM, Kaufman RJ (2003) The noncatalytic portion of human UDP-glucose:glycoprotein glucosyltransferase I confers UDP-glucose binding and transferase function to the catalytic domain. *J Biol Chem* 278:43320–43328
- Buzzi LI, Simonetta SH, Parodi AJ et al (2011) The two *Caenorhabditis elegans* UDP-glucose:glycoprotein glucosyltransferase homologues have distinct biological functions. *PLoS One* 6: e27025
- Caramelo JJ, Castro OA, Alonso L et al (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
- D'Alessio C, Caramelo JJ, Parodi AJ (2010) UDP-Glc:glycoprotein glucosyltransferase-glucosidase II, the ying-yang of the ER quality control. *Semin Cell Dev Biol* 21:491–499
- Hammond C, Braakman I, Helenius A (1994) Role of *N*-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control. *Proc Natl Acad Sci U S A* 91:913–917
- Helenius A, Aebi M (2004) Roles of *N*-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem* 73:1019–1049
- Jin H, Yan Z, Nam KH et al (2007) Allele-specific suppression of a defective brassinosteroid receptor reveals a physiological role of UGGT in ER quality control. *Mol Cell* 26:821–830
- Parodi AJ (2000) Protein glucosylation and its role in protein folding. *Annu Rev Biochem* 69:69–93
- Stigliano ID, Alculumbre SG, Labriola CA et al (2011) Glucosidase II and *N*-glycan mannose content regulate the half lives of monoglucosylated species in vivo. *Mol Biol Cell* 22:1810–1823
- Takeda Y, Seko A, Hachisu M et al (2014) Both isoforms of human UDP-Glucose:glycoprotein glucosyltransferase are enzymatically active. *Glycobiology* 24:344–350
- Zhang W, Wearsch PA, Zhu Y et al (2011) A role for UDP-glucose: glycoprotein glucosyltransferase in expression and quality control of MHC class I molecules. *Proc Natl Acad Sci U S A* 108:4956–4961