THE ROYAL SOCIETY

doi 10.1098/rstb.2000.0759

Quality control in the secretory assembly line

Ari Helenius

Swiss Federal Institute of Technology (ETH), Institute of Biochemistry, Universitätstrasse 16, CH-8092 Zurich, Switzerland (ari.helenius@bc.biol.ethz.ch)

As a rule, only proteins that have reached a native, folded and assembled structure are transported to their target organelles and compartments within the cell. In the secretory pathway of eukaryotic cells, this type of sorting is particularly important. A variety of molecular mechanisms are involved that distinguish between folded and unfolded proteins, modulate their intracellular transport, and induce degradation if they fail to fold. This phenomenon, called quality control, occurs at several levels and involves different types of folding sensors. The quality control system provides a stringent and versatile molecular sorting system that guaranties fidelity of protein expression in the secretory pathway.

Keywords: protein folding; glycosylation; chaperones; lectins; sorting

1. INTRODUCTION

The endoplasmic reticulum (ER) is the site of synthesis for the majority of secretory proteins, lysosomal proteins, and proteins of the plasma membrane and endocytic–exocytic membranes. In mammalian cells these proteins are first co-translationally translocated into the ER lumen or ER membrane. Their folding occurs co- and post-translationally while still in the ER, and many of them undergo oligomeric assembly before transport to the Golgi complex and their final locations within or outside the cell (Helenius *et al*. 1992).

Typically, the maturation of proteins also involves covalent processing such as proteolytic cleavage in the form of signal sequence removal, glycosylation by addition of N-linked core glycans, and cross-linking through the formation of intra- and intermolecular disulphide bonds. Some proteins undergo further modifications including proline and lysine hydroxylation (procollagens), addition of lipid (glycosylphosphatidyl inositol anchored proteins), and association of ligands such as haeme, retinol and lipids. These modifications are usually important for the folding and assembly process, and essential for export from the ER.

The folding, assembly and modification are supported by a large number of different ER-resident enzymes and molecular chaperones (Gething 1997). Some are present in the ER of all cells and often at high concentrations. Others are only present in cells that produce the corresponding protein substrates. Together with a precisely controlled ionic and redox milieu, they ensure that conditions in the ER are optimal for efficient protein folding and maturation. The ER is a highly specialized organelle largely devoted to protein synthesis and folding.

2. QUALITY CONTROL

Nevertheless, when one measures actual rates and efficiencies by which proteins fold in the ER of living cells, the processes often appear surprisingly slow and ineffective. For some proteins it takes hours to reach their mature folded form and, in many cases, less than half of a given polypeptide species actually reaches its native conformation (Hurtley & Helenius 1989). This means that cells continuously produce misfolded proteins, protein aggregates, unassembled subunits, and incorrectly assembled oligomers as side-products. Some of the 'junk' produced is caused by genuine folding problems, some to non-stoichiometric synthesis of subunits of hetero-oligomers.

To deal with this problem, cells have systems to sort folded from the incompletely folded proteins, and to deploy the folded and degrade the misfolded. This is called quality control (QC) (Brewer & Corley 1996; Ellgaard *et al*. 1999; Hammond & Helenius 1995; Hurtley & Helenius 1989; Nauseef 1999; Parodi 1999). It amounts to a highly stringent conformation-based (architectural), molecular sorting process that regulates which conformers are deployed (Hurtley & Helenius 1989; Klausner 1989). It differs from other types of sorting in that the conformation is the critical parameter. Once recognized as misfolded, proteins are, as a rule, retained in the ER. If they fail to fold, they are translocated from the ER to the cytosol, where they are degraded by a proteasome-dependent degradation system (Bonifacino & Weissman 1998; Plemper & Wolf 1999).

The main reason for the existence of QC is, no doubt, to prevent deployment of misfolded and incompletely oligomerized proteins that would be deleterious to the cell. The fact that a breakdown in QC would actually result in problems has recently been documented by studies on the conformational screening of adenosine triphosphate-sensitive K channels. When QC was relaxed, a variety of defective channels reached the plasma membrane, a situation dangerous for maintenance of the ion balance in the cell (Zerangue *et al*. 1999). However, QC also allows a protein to remain within the environment of the ER and thus increase its chances for folding.

Table 1. *Disease states involving ER folding and quality control*

disease	protein
secretory proteins and extra cellular matrix components	
emphysema–liver disease	α1-antitrypsin
α1-antichymotrypsin deficiency	α1-antichymotrypsinogen
scurvy	procollagen
micromelia (in chicken)	Aggrecan
osteogenesis imperfecta	procollagens I, II and IV
Marfan's syndrome	fibrillin
fibrinogen storage disease	fibrinogen
Von Willebrand's disease	vW factor
diabetes insipidus	vasopressin–vasopressin receptor
protein C deficiency	protein C
primary hypoparathyroidism	preproparathyroid hormone
type I hereditary angioderma	complement C1
factor H deficiency	155 kDa factor H subunit
plasma membrane proteins	
cystic fibrosis	cystic fibrosis transmembrane regulator
Glanzmann's thrombasthenia	integrin receptor
congenital sucrase–isomaltase deficiency	sucrase–isomaltase
hereditary haemochromatosis	transferrin receptor
familial hypercholesterolaemia	low-density lipoprotein receptor
type I chylomicronaemia	lipoprotein lipase
Charcot–Marie–Tooth syndrome	myelin protein 22
Perlizaeus–Merzbacher disease	proteolipoprotein
nephrogenic diabetes insipidus	Aquaporin
myeloperoxidase deficiency	myeloperoxidase
Laron dwarfism	growth hormone receptor
diabetes mellitus	insulin receptor
lysosomal enzymes and proteins	
hexosaminidase A deficiency	α-hexosaminidase
Sandhoff/Tay–Sachs	β-hexosaminidase
Hurler syndrome	α-L-iduronidase
aspartylglucoseaminuria	aspartylglucoseaminidase
Maroteaux–Lamy syndrome	lysosomal 4-sulphatase
GM2-gangliosidosis AB variant	GM2 activator protein
others	
retinitis pigmentosa	rhodopsin
congenital hypothyroid goitre	thyroglobulin
glycanosis CDG type 1	underglycosylation (OST?)
abetalipoproteinaemia	microsomal triglyceride transfer protein
melanoma	tyrosinase

3. A GLYCOPROTEIN-SPECIFIC FOLDING SENSOR

The QC process in the ER can be divided into sequential reactions that involve detection (recognition), retention, retro-translocation and degradation. The remainder of this review focuses on the first of these steps, i.e. on the system that decides whether a protein is folded or not.

The most common mechanism of recognition and retention of proteins involves association of a substrate protein with an ER-resident chaperone or folding enzyme. The resident proteins in question include major chaperones such as BiP, calnexin (CNX), calreticulin (CRT), Grp98, or the folding enzymes PDI and ERp57 (Gething 1997). As long as a newly synthesized protein interacts with an ER protein, even if the association is determined by an on–off equilibrium, it is 'defined' by the system as incompletely folded and unfit for transport. Chaperones thus function simultaneously as folding catalysts, folding sensors and retention anchors. Interaction between the substrate and these chaperones often begins at the level of the growing nascent chain (Chen & Helenius 2000; Chen *et al*. 1995; Molinari & Helenius 2000) and continues throughout the folding and assembly process (Hammond & Helenius 1994).

Most proteins made in the ER acquire N-linked glycans co-translationally, and are thus glycoproteins. This also means that they are likely to interact with a special glycoprotein-specific chaperone system often called the calnexin–calreticulin cycle (Helenius *et al*. 1997; Zapun *et al*. 1999). Mediated by two homologous lectins, CNX and CRT, this cycle plays a central role in folding and QC in the ER lumen.

It was first proposed in 1993 that folding, CNX binding and oligosaccharide trimming are closely linked functions (Hammond & Helenius 1993). Since then it has become quite clear that the glycans and their trimming in the ER determine their interaction with both CNX and CRT. They bind to proteins that have monoglucosylated core N-linked glycans, i.e. proteins that have lost

by trimming two of the three glucose residues of the original three in the core glycans. Together with ERp57, a thiol oxidase, with which CNX and CRT form complexes (Oliver *et al*. 1997), they promote proper folding of the glycoproteins through a binding and release cycle (Molinari & Helenius 1999; Zapun *et al*. 1998).

Glycoprotein binding to CNX and CRT is regulated by two resident enzymes, glucosidase II (a releasing enzyme that removes glucose) and uridine diphosphate (UDP)-glucose:glycoprotein glucosyltransferase (GT, an association promoting enzyme that adds glucose). By removing and adding glucose residues (Parodi *et al*. 1983; Trombetta & Parodi 1992), respectively, they drive the cycle of binding and release.

It is evident from studies *in vitro* and *in vivo* that GT serves in this cycle as a folding sensor. It selectively re-glucosylates glycoproteins that have not undergone proper folding (Sousa *et al*. 1992; Sousa & Parodi 1995; Trombetta & Helenius 2000; Ritter & Helenius 2000). Glycoproteins therefore exit from the cycle when they have reached their native conformation.

How does GT, a 170 kDa soluble, resident ER enzyme, recognize the difference between folded and unfolded glycoproteins? Work by Parodi and colleagues (see Parodi 1996) has shown that GT 'sees' the oligosaccharide and protein moieties together. Thus it does not recognize oligosaccharides or small glycopeptides alone as substrates, nor are unfolded proteins without glycans inhibitory. A single protein-linked *N*-acetyl glucosamine is, however, sufficient to make a misfolded glycoprotein visible to GT (Sousa & Parodi 1995). Since the first glycan is known to be immobilized in folded proteins via interactions with the polypeptide chain, its dynamic properties may serve as an indicator for the local folding status in the recognition process. An alternative model suggests that GT recognizes hydrophobic surface features on misfolded proteins in the same way as many classical chaperones (Parodi 1996).

Recent experiments in our laboratory with RNase B, a simple glycoprotein, indicate that GT does not re-glucosylate random coils with glycans as efficiently as the same proteins in a partially folded form (Trombetta & Helenius 2000). Moreover, we have shown that if a glycoprotein has two domains, one folded and the other unfolded, GT only re-glucosylates the latter (Ritter & Helenius 2000). This means that it recognizes local folding defects and does not indiscriminately modify sugars in the whole molecule.

4. THE ROLE OF ENDOPLASMIC RETICULUM FOLDING AND QUALITY CONTROL IN DISEASE

A large number of disease states are known in which the underlying cause can be found in folding, maturation and QC at the ER level. Table 1 shows some of the disease states in which evidence exists for such 'ER storage disease' aetiology. Typically these diseases are hereditary and caused by mutations in specific proteins synthesized in the ER (Amara *et al*. 1992; Aridor & Balch 1999; Thomas *et al*. 1995). Quite often this disease aetiology is only seen for some of the alleles in each syndrome. Although synthesized and translocated into the ER, one or more proteins fail to fold and are therefore not efficiently transported to its proper site of residence in the cell or outside the cell. In some cases, the accumulation of the protein in the ER of the cells leads to additional symptoms.

The mutated or improperly modified proteins are often truncated, they lack important disulphide bonds or they tend to aggregate. In rare cases they seem, however, to reach a functional conformation but are defective enough structurally to be trapped by the QC system. Since many of them are glycoproteins, they may be unable to escape detection by GT.

A better understanding of protein folding and QC systems in the secretory pathway will be essential for understanding these diseases better, and to develop therapies. In general, our current understanding of protein folding in the living cell is quite incomplete. However, methods have now been developed that allow detailed analysis of early interactions in the ER and these can be applied to new proteins. Major progress can be expected as studies begin to focus on the folding process in its physiological milieu.

I thank members of my laboratory for discussions and advice, and the Swiss National Science Foundation and the Roche Foundation for support.

REFERENCES

Amara, J. F., Cheng, S. H. & Smith, A. E. 1992 Intracellular protein trafficking defects in human disease. *Trends Cell Biol.* **2**, 145–149.

Aridor, M. & Balch, W. E. 1999 Integration of endoplasmic reticulum signaling in health and disease. *Nat. Med.* **5**, 745–751.

Bonifacino, J. S. & Weissman, A. M. 1998 Ubiquitin and the control of protein fate in the secretory and endocytic pathways. *A. Rev. Cell Dev. Biol.* **14**, 19–57.

Brewer, J. W. & Corley, R. B. 1996 Quality control in protein biogenesis: thiol-mediated retention monitors the redox state of proteins in the endoplasmic reticulum. *J. Cell Sci.* **109**, 2383–2392.

Chen, W. & Helenius, A. 2000 Role of ribosome and translocon complex during folding of influenza hemagglutinin in the endoplasmic reticulum of living cells. *Mol. Biol. Cell* **11**, 765–772.

Chen, W., Helenius, J., Braakman, I. & Helenius, A. 1995 Cotranslational folding and calnexin binding during glycoprotein synthesis. *Proc. Natl Acad. Sci. USA* **92**, 6229–6233.

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

Gething, M.-J. (ed.) 1997 *Guidebook to molecular chaperones and protein-folding catalysts*. Oxford University Press.

Hammond, C. & Helenius, A. 1993 A chaperone with a sweet tooth. *Curr. Biol.* **3**, 884–885.

Hammond, C. & Helenius, A. 1994 Folding of VSV G protein: sequential interaction with BiP and calnexin. *Science* **266**, 456–458.

Hammond, C. & Helenius, A. 1995 Quality control in the secretory pathway. *Curr. Opin. Cell Biol.* **7**, 523–529.

Helenius, A., Tatu, U., Marquardt, T. & Braakman, I. 1992 Protein folding in the endoplasmic reticulum. In *Cell biology and biotechnology* (ed. R. G. Rupp & M. S. Oka), pp. 125–136. Berlin and Heidelberg, Germany: Springer.

Helenius, A., Trombetta, E. S., Hebert, D. N. & Simons, J. F. 1997 Calnexin, calreticulin and the folding of glycoproteins. *Trends Cell Biol.* **7**, 193–200.

Hurtley, S. M. & Helenius, A. 1989 Protein oligomerization in the endoplasmic reticulum. *A. Rev. Cell Biol.* **5**, 277–307.
Klausner, R. D. 1989 Architectural editing: determining the fate of newly synthesized membrane proteins. *New Biol.* **1**, 3–8.
Molinari, M. & Helenius, A. 1999 Glycoproteins form mixed disulfides with oxidoreductases during folding in living cells. *Nature* **402**, 90–93.
Molinari, M. & Helenius, A. 2000 Chaperone selection during glycoprotein translocation and folding in the ER. *Science* **288**, 331–333.
Nauseef, W. M. 1999 Quality control in the endoplasmic reticulum: lessons from hereditary myeloperoxidase deficiency. *J. Lab. Clin. Med.* **134**, 215–221.
Oliver, J. D., Van der Wal, F. J., Bulleid, N. J. & High, S. 1997 Interaction of the thiol-dependent reductase ERp57 with nascent glycoproteins. *Science* **275**, 86–88.
Parodi, A. J. 1996 The UDP-Glc:glycoprotein glycosyl transferase and the quality control of glycoprotein folding in the endoplasmic reticulum. *Trends Glycosci. Glyc.* **8**, 1–12.
Parodi, A. J. 1999 Reglucosylation of glycoproteins and quality control of glycoprotein folding in the endoplasmic reticulum of yeast cells. *Biochim. Biophys. Acta* **1426**, 287–295.
Parodi, A. J., Mendelzon, D. H. & Lederkremer, G. H. 1983 Transient glucosylation of protein bound Man9GlcNac2, Man8GlcNac2 and Man7GlcNac2 in calf thyroid cells. A possible recognition signal in the processing of glycoproteins. *J. Biol. Chem.* **258**, 8260–8265.
Plemper, R. K. & Wolf, D. H. 1999 Retrograde protein translocation: ERADication of secretory proteins in health and disease. *Trends Biochem. Sci.* **24**, 266–270.
Ritter, C. & Helenius, A. 2000 Recognition of local misfolding by the ER folding sensor UDP-glucose:glycoprotein glucosyltransferase. *Nature Struct. Biol.* **7**, 278–280.
Sousa, M. & Parodi, A. J. 1995 The molecular basis for the recognition of misfolded glycoproteins by the UDP-Glc:glycoprotein glucosyltransferase. *EMBO J.* **14**, 4196–4203.
Sousa, M. C., Ferrero-Garcia, M. A. & Parodi, A. J. 1992 Recognition of the oligosaccharide and protein moieties of glycoproteins by the UDP-Glc:glycoprotein glucosyltransferase. *Biochemistry* **31**, 97–105.
Thomas, P. J., Qu, B. & Pederson, P. L. 1995 Defective protein folding as a basis of human disease. *Trends Biochem. Sci.* **20**, 456–459.
Trombetta, E. S. & Helenius, A. 2000 Conformational requirements for glycoprotein reglucosylation in the endoplasmic reticulum. *J. Cell. Biol.* **148**, 1123–1130.
Trombetta, S. E. & Parodi, A. J. 1992 Purification to apparent homogeneity and partial characterization of rat liver UDP-glucose:glycoprotein glucosyltransferase. *J. Biol. Chem.* **267**, 9236–9240.
Zapun, A., Darby, N. J., Tessier, D. C., Michalak, M., Bergeron, J. J. & Thomas, D. Y. 1998 Enhanced catalysis of ribonuclease B folding by the interaction of calnexin or calreticulin with ERp57. *J. Biol. Chem.* **273**, 6009–6012.
Zapun, A., Jakob, C. A., Thomas, D. Y. & Bergeron, J. J. 1999 Protein folding in a specialized compartment: the endoplasmic reticulum. *Struct. Fold Des.* **7**, R173–R182.
Zerangue, N., Schwappach, B., Jan, Y. N. & Jan, L. Y. 1999 A new ER trafficking signal regulates the subunit stoichiometry of plasma membrane K(ATP) channels. *Neuron* **22**, 537–548.

Discussion

R. Sitia (*Molecular Immunology, Instituto Scientifico San Raffaele, Milan, Italy*). You have presented convincing evidence that the ER quality control machinery patrols local folding in a polypeptide via GT. How can we reconcile this with the fact that GT does not recognize glycopeptides?

A. Helenius. There are two obvious possibilities. Although it is difficult to imagine, GT may function in such a way that there is a size limit below which substrates are not recognized. On the other hand, our results show that to be a good substrate a glycoprotein must be partially structured. A random coil is not a substrate. It is therefore possible that glycopeptides are poor substrates, not because they are small, but because they are unstructured. How large a partially unfolded region in a larger protein needs to be, to make the protein a substrate, is not clear.

B. E. P. Swoboda (*Department of Biological Sciences, University of Warwick, UK*). Does the addition of a glucose residue to a branched mannose glycoprotein by the UDP-glucose transfer enzyme (GT) programme a disordered protein for destruction?

A. Helenius. Yes, it probably does. The presence of the single glucose residue tells the ER system that the glycoprotein is still incompletely folded. While enough for retention, this is not sufficient for degradation. There must be another signal, a signal that tells the system that a protein is permanently misfolded and not a folding intermediate. The nature of the second signal is not known, but there is evidence that it involves the loss of mannose residues from the N-linked core glycan through the action of α-mannosidases in the ER lumen. For example, if mannosidase I is inhibited, degradation of misfolded glycoproteins generally comes to a halt.

B. E. P. Swoboda. Does the degree of glucosylation of a protein and the progressive removal of sugar residues define the lifetime of a glycoprotein?

A. Helenius. Referring to the answer to the previous questions, it seems that the presence of glucose and the loss of mannose determines in some way how long a glycoprotein survives in the ER lumen before it is degraded. The actual half-life depends no doubt in addition on the activity of the various enzymes (which in the case of the mannosidase is relatively slow), on the number of glycans, the amount of misfolded protein, etc. To allow slowly folding proteins time to mature, degradation should not be triggered too quickly, yet it is not desirable to let the ER fill up with incompletely folded protein.