FEBS Letters 476 (2000) 38–41 FEBS 23790

Minireview

Glycoprotein folding in the endoplasmic reticulum: a tale of three chaperones?

Stephen High*, Fabienne J.L. Lecomte, Sarah J. Russell, Benjamin M. Abell, Jason D. Oliver

School of Biological Sciences, University of Manchester, 2.205 Stopford Building, Oxford Road, Manchester M13 9PT, UK

Received 5 May 2000

Edited by Gunnar von Heijne

Abstract The endoplasmic reticulum (ER) is a major site of protein synthesis and its inside, or lumen, is a major site of protein folding. The lumen of the ER contains many folding factors and molecular chaperones, which facilitate protein folding by increasing both the rate and the efficiency of this process. Amongst the many ER folding factors, there are three components that specifically modulate the folding glycoproteins bearing *N*-linked carbohydrate side chains. These components are calnexin, calreticulin and ERp57, and this review focuses on the molecular basis for their capacity to influence glycoprotein folding.

Key words: Endoplasmic reticulum; Glycoprotein; Protein folding; Molecular chaperone; Calnexin; ERp57

1. Introduction

In eukaryotic cells, the endoplasmic reticulum (ER) acts as the entry point for newly synthesised proteins to enter the well-defined secretory pathway. Hence, the ER is a major site of protein synthesis with many nascent polypeptides being co-translationally translocated into and across its membrane. These polypeptides pass through the translocation sites of the ER membrane as partially unfolded polypeptide chains. It should therefore be no surprise that the inside, or lumen, of the ER functions as a specialised folding environment and that it contains a number of molecular chaperones and folding factors [1]. This ensures that as newly synthesised polypeptides enter the ER lumen, the nascent chains begin to fold rapidly into their native structures.

From amongst a complex palette of ER chaperones and folding factors, three components have been experimentally defined as interacting specifically with newly synthesised glycoproteins, that is proteins that have been covalently modified by the attachment of *N*-linked oligosaccharides. This specificity for glycoproteins is particularly apparent, since none of the other known ER folding factors select their substrates purely on the basis of whether or not they are *N*-glycosylated. This article will focus specifically upon these three glycoprotein-specific ER components, drawing particularly upon the more recent developments in this area.

2. Glycoproteins are special

Historically, calnexin was the first glycoprotein-specific ER component to be identified [2]. Calnexin is an integral membrane protein, and it soon transpired that its soluble homologue, the resident ER luminal protein calreticulin, also interacted specifically with newly synthesised glycoproteins [3]. Calnexin and calreticulin share regions of high amino acid sequence identity and have similar glycoprotein binding activities. These two ER components were initially described as molecular chaperones on the basis that they both associated transiently with many different newly synthesised glycoproteins as they entered the secretory pathway at the ER [2–4]. However, it was later shown that a principal factor in the binding of calnexin and calreticulin to newly synthesised glycoproteins is the structure of the carbohydrate side chain [5,6]. In contrast to what one would expect of a classical chaperone, the conformation of the polypeptide to which the carbohydrate is attached appeared to have little or no influence upon the binding of calnexin and calreticulin [5–7]. In practice this means that calnexin and calreticulin can either be described as ER lectins that recognise a very specific carbohydrate structure, or that one has to broaden one's definition of a molecular chaperone [4,8,9].

3. Carbohydrate processing is important

The carbohydrate specificity of calnexin and calreticulin has been well defined, and both components recognise proteins carrying one or more monoglucosylated oligosaccharide side chains [5,6]. The monoglucosylated carbohydrate structure, $GlcNAc_2Man_{7-9}Glc_1$, is initially generated by the action of glucosidases I and II on the *N*-linked glycan precursor, $GlcNAc_2Man_9Glc_3$ [8]. The abundance of monoglucosylated oligosaccharide present on a glycoprotein precursor will be determined by two ER-resident enzymes. Glucosidase II removes the final glucose residue generating a deglucosylated side chain and hence removing monoglucosylated glycoprotein substrates, whilst UDP glucose:glycoprotein glucosyl transferase (UGGT) can add a single glucose residue back onto this structure, thus regenerating the monoglucosylated side chain [10]. UGGT recognises only incompletely folded glycoproteins as its substrate and so it is UGGT that acts as a 'folding sensor' within the glycoprotein folding cycle (see Fig. 1). Hence, UGGT detects non-native glycoproteins

*Corresponding author. Fax: (44)-161-275 5082.
E-mail: shigh@fs1.scg.man.ac.uk

PII: S0014-5793(00)01666-5

and reglucosylates them, thereby allowing further rounds of binding to calnexin/calreticulin [8,9].

4. ERp57 – the missing link?

The crucial question to pose at this point is whether calnexin and calreticulin can actually exert any direct influence over glycoprotein folding? This is a very moot point [8,9,11] and one that we will return to later. What is now well established is that there is a third ER component, ERp57, which plays a key role in glycoprotein folding [12]. We believe that it is the combination of ERp57 with both calnexin and calreticulin (see Fig. 1) that forms the lynchpin of chaperone-mediated glycoprotein folding in the ER and will explore this hypothesis further below.

ERp57 has a host of pseudonyms and a 'chequered' past [13,14]. What is very clear is that ERp57 is a member of the protein disulphide isomerase (PDI) family [15]. Like archetypal PDI, ERp57 contains two 'thioredoxin motifs' which constitute the thiol/disulphide oxidoreductase active sites of PDI [16]. On the basis of its sequence similarity to PDI (29% identity, 56% similarity [22]), ERp57 was analysed for its functional homology several years ago. It was established that ERp57 does exhibit a thiol-dependent reductase activity, albeit that it was significantly less effective than PDI [17–19]. These data did show that ERp57 could, at least in principle, influence protein folding.

We first came across ERp57 when we were using a cross-linking approach to analyse the interactions of newly synthesised secretory and membrane proteins with ER components. Our crucial discovery was the finding that ERp57 interacts specifically with *N*-glycosylated polypeptides [12,13,20,21]. It was also clearly significant that both the binding of ERp57 to newly synthesised glycoproteins, and its subsequent release, required the trimming of glucose residues from the *N*-linked glycan [12,13,20,21]. Such a requirement for the glucose trimming of *N*-linked glycans had previously been identified as a hallmark for the binding of calnexin and calreticulin. Since there was no evidence that ERp57 had any intrinsic lectin-like properties, a prediction that was subsequently validated experimentally [27], we proposed that ERp57 functioned in concert with calnexin and calreticulin to specifically modulate glycoprotein folding [12].

The most obvious molecular basis for this concerted action was that ERp57 functioned as a specific subunit of discrete complexes formed with calnexin and calreticulin, and we therefore set out to test this hypothesis directly. Using a variety of approaches we showed that ERp57 forms distinct complexes with both calnexin and calreticulin, both within the lumen of the ER and when the proteins are mixed in solution [14]. We also clearly established that the binding of ERp57 to calnexin and calreticulin is direct and does not require the presence of the complex's substrate (i.e. glucose-trimmed glycoprotein, cf. Fig. 1). The association of ERp57 with calnexin and calreticulin is clearly specific. Hence, despite its sequence similarity, PDI does not associate with calnexin or calreticulin in any of our assays [14], although calreticulin/PDI complexes can be observed when the purified proteins are studied in vitro [23].

5. The role of ERp57 in glycoprotein folding

Our current model for the role of ERp57 is that it functions in the ER lumen as a subunit of discrete complexes formed

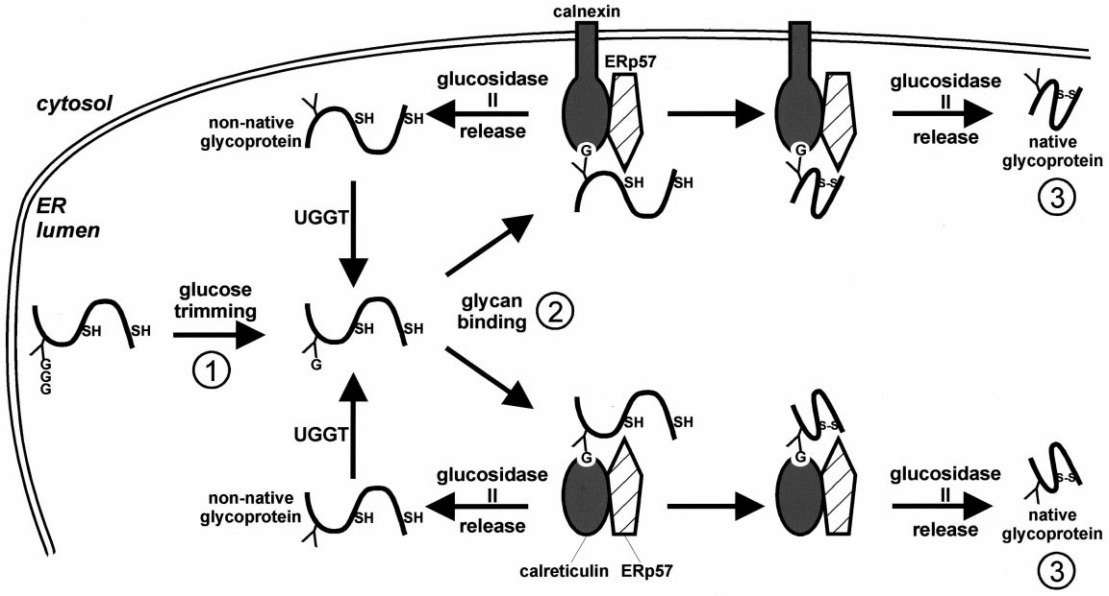

Fig. 1. Glycoprotein folding in the ER lumen. The *N*-linked oligosaccharide side chains present on newly synthesised glycoproteins are rapidly trimmed by glucosidases I and II to generate monoglucosylated forms of the oligosaccharides (step 1). These monoglucosylated glycoproteins are specifically bound by two resident ER lectins, calnexin, an integral membrane protein, and calreticulin, its soluble homologue. By virtue of its stable association with calnexin and calreticulin, ERp57 is brought into contact with the newly made glycoproteins (step 2) and is able to modulate their folding. In this model we have shown the role of ERp57 as that of promoting native disulphide bond formation and this is supported by experimental evidence [27,28]. However, ERp57 may also play a more general role as a chaperone or folding factor (see text). The glycoprotein substrate is released from calnexin or calreticulin, and thereby its specific association with ERp57, after the removal of the last glucose residue from the glycan by glucosidase II. If its folding has been successful, the native glycoprotein is free to be transported along the secretory pathway (step 3). If the glycoprotein is incorrectly or incompletely folded the folding sensor UGGT adds back a single glucose residue and the glycoprotein can undergo another round of lectin binding and ERp57 action (re-entry at step 2).

with calreticulin and calnexin. We propose that it is these complexes which directly modulate glycoprotein folding and are therefore acting as molecular chaperones (Fig. 1). Whilst our own work has not directly addressed the role of ERp57 during glycoprotein folding, several other recent studies do indeed indicate that ERp57 functions to influence protein folding, and confirm that it does so in combination with calnexin and calreticulin. Firstly, several groups have demonstrated the presence of ERp57, in addition to calnexin and calreticulin, during the assembly of the MHC class I complex [24–26]. The heavy chain of the MHC class I complex is a glycoprotein, and its assembly with β2-microglobulin in the ER appears to be mediated by a specific set of ER folding factors which includes calnexin, calreticulin and ERp57. Secondly, the PDI activity of ERp57 has been studied directly by analysing the refolding of an authentic monoglucosylated precursor. It was found that the disulphide isomerase activity of ERp57 is greatly enhanced by the presence of calreticulin or calnexin [27]. This result provides direct evidence that it is the combination of ERp57 with calnexin or calreticulin that specifically modulates glycoprotein folding, in this case by promoting the formation of native disulphide bonds [27]. Thirdly, a direct role for ERp57 in the specific catalysis of native disulphide bond formation in glycoproteins has recently been experimentally demonstrated [28]. Hence, mixed disulphide species formed in vivo between ERp57 and specific viral glycoprotein substrates could be identified [28]. These mixed disulphides represent transient intermediates during the catalysis of disulphide bond formation and exchange [29], and they were only obtained with ERp57 when the precursors were both *N*-glycosylated and suitably glucose-trimmed [28].

The precise molecular basis for the dramatic increase in *disulphide isomerase activity* that is observed by combining *ERp57* with *calnexin or calreticulin* remains to be established. The ER lectins calnexin and calreticulin may simply act to bring a chaperone (ERp57) and its substrate (a glucose-trimmed glycoprotein) together efficiently [27]. This role could be more refined, since the relative positions of the glycans and cysteine residues within a polypeptide might specifically facilitate the formation of particular disulphide bonds. In contrast to the stimulation of ERp57 activity, the binding of glycoprotein substrates to calnexin or calreticulin inhibits the isomerase activity of PDI in vitro [27]. Whether this inhibition of PDI function is in any way a crucial facet of the glycoprotein specific folding pathway of the ER lumen is also at present unclear.

6. Where to next?

It seems clear that ERp57 functions as a glycoprotein-specific PDI, but does ERp57 have any other functions? It is surely significant that ERp57 can associate with glycoproteins that lack any cysteine residues, and that do not therefore require any intramolecular disulphide exchange [12,13]. This observation suggests that the actions of ERp57 may not be limited to the catalysis of disulphide bond formation and exchange. Rather, as appears to be the case for PDI, ERp57 may have a wider role as a molecular chaperone [15]. Following the identification of a role for calnexin and calreticulin in glycoprotein folding, a recurring question has been what do these ER lectins actually do, and can they exert any direct influence over protein folding. Whilst our own analysis of glycoprotein folding has highlighted the contribution of ERp57, others have now shown that both calnexin and calreticulin can influence the in vitro refolding of proteins in the absence of ERp57 [30,31]. In these assays, an effect on both glycosylated and non-glycosylated proteins was observed, and calnexin and calreticulin both behaved as classical molecular chaperones [30,31].

Since the initial identification of calnexin as a glycoprotein-specific ER component [2] we have come a long way in our understanding of how newly synthesised glycoproteins are folded in the ER lumen. A combination of recent publications [30–32] and our own better judgement tell us that there is still much to understand.

Acknowledgements: Our contribution to the work described has been supported by funding from the BBSRC and the Wellcome Trust. S.H. would like dedicate this review to Sarah, Dan, Joe, Zac and Eddie for their continuing support.

References

[1] Gething, M.-J. and Sambrook, J. (1992) Nature 355, 33–45.
[2] Ou, W.-J., Cameron, P.H., Thomas, D.Y. and Bergeron, J.J.M. (1993) Nature 364, 771–776.
[3] Peterson, J.R., Ora, A., Van, P.N. and Helenius, A. (1995) Mol. Biol. Cell 6, 1173–1184.
[4] Helenius, A., Trombetta, E.S., Hebert, D.N. and Simons, J.F. (1997) Trends Cell Biol. 7, 193–200.
[5] Rodan, A.R., Simons, J.F., Trombetta, E.S. and Helenius, A. (1996) EMBO J. 15, 6921–6930.
[6] Zapun, A., Petrescu, S.M., Rudd, P.M., Dwek, R.A., Thomas, D.Y. and Bergeron, J.J.M. (1997) Cell 88, 29–38.
[7] Allen, S. and Bulleid, N.J. (1997) Biochem. J. 328, 113–119.
[8] Trombetta, E.S. and Helenius, A. (1998) Curr. Opin. Struct. Biol. 8, 587–592.
[9] Zapun, A., Jakob, C.A., Thomas, D.Y. and Bergeron, J.J. (1999) Structure 7, R173–R182.
[10] Trombetta, S.E., Bosch, M. and Parodi, A.J. (1989) Biochemistry 28, 8108–8116.
[11] Ellgaard, L., Molinari, M. and Helenius, A. (1999) Science 286, 1882–1888.
[12] Oliver, J.D., van der Wal, F.J., Bulleid, N.J. and High, S. (1997) Science 275, 86–88.
[13] Elliott, J.G., Oliver, J.D. and High, S. (1997) J. Biol. Chem. 272, 13849–13855.
[14] Oliver, J.D., Roderick, H.L., Llewellyn, D.H. and High, S. (1999) Mol. Biol. Cell 10, 2573–2582.
[15] Ferrari, D.M. and Soling, H.D. (1999) Biochem. J. 339, 1–10.
[16] Freedman, R.B., Hirst, T.R. and Tuite, M.F. (1994) Trends Biochem. Sci. 19, 331–336.
[17] Bourdi, M., Demady, D., Martin, J.L., Jabbour, S.K., Martin, B.M., George, J.W. and Pohl, L.R. (1995) Arch. Biochem. Biophys. 323, 397–403.
[18] Hirano, N., Shibasaki, F., Sakai, R., Tanaka, T., Nishida, J., Yazaki, Y., Takenawa, T. and Hirai, H. (1995) Eur. J. Biochem. 234, 336–342.
[19] Srivastava, S.P., Fuchs, J.A. and Holtzman, J.L. (1993) Biochem. Biophys. Res. Commun. 192, 971–978.
[20] Oliver, J.D., Hresko, R.C., Mueckler, M. and High, S. (1996) J. Biol. Chem. 271, 13691–13696.
[21] Van der Wal, F.J., Oliver, J.D. and High, S. (1998) Eur. J. Biochem. 256, 51–59.
[22] Koivunen, P., Helaakoski, T., Annunen, P., Veijola, J., Räisänen, S., Pihlajaniemi, T. and Kivirikko, K. (1996) Biochem. J. 316, 599–605.
[23] Corbett, E.F. et al. (1999) J. Biol. Chem. 274, 6203–6211.
[24] Hughes, E.A. and Cresswell, P. (1998) Curr. Biol. 8, 709–712.
[25] Lindquist, J.A., Jensen, O.N., Mann, M. and Hammerling, G.J. (1998) EMBO J. 17, 2186–2195.
[26] Morrice, N.A. and Powis, S.J. (1998) Curr. Biol. 8, 713–716.

[27] Zapun, A., Darby, N.J., Tessier, D.C., Michalak, M., Bergeron, J.J. and Thomas, D.Y. (1998) J. Biol. Chem. 273, 6009–6012.
[28] Molinari, M. and Helenius, A. (1999) Nature 402, 90–93.
[29] Freedman, R. (1999) Nature 402, 27–29.
[30] Ihara, Y., Cohen-Doyle, M.F., Saito, Y. and Williams, D.B. (1999) Mol. Cell 4, 331–341.
[31] Saito, Y., Ihara, Y., Leach, M.R., Cohen-Doyle, M.F. and Williams, D.B. (1999) EMBO J. 18, 6718–6729.
[32] Gillece, P., Luz, J.M., Lennarz, W.J., de La Cruz, F.J. and Romisch, K. (1999) J. Cell Biol. 147, 1443–1456.