Volume 330, number 1, 1–4 FEBS 12918 September 1993

Hypothesis

Kin recognition

A model for the retention of Golgi enzymes

Tommy Nilsson*, Paul Slusarewicz, Mee H. Hoe and Graham Warren

Imperial Cancer Research Fund, PO Box 123, 44 Lincoln's Inn Fields, London WC2A 3PX, UK

(Received 16 July 1993)

The surprising result that the spanning domain causes retention of proteins in the Golgi stack poses the question as to the actual mechanism. Here we present a simple model that might have general applicability.

Golgi apparatus; Resident protein; Retention; Membrane traffic

1. INTRODUCTION

Transport of proteins to the cell surface occurs by default. Signals are not required for newly assembled proteins that leave the endoplasmic reticulum (ER) and are destined for the cell surface [1,2]. Signals are needed to divert proteins to the lysosomes [3] and secretory granules [4], and to retrieve proteins that are lost from the compartment in which they function [5]. In all cases the signals are present on the cytoplasmic or lumenal domains of the protein.

Golgi proteins also follow the default pathway but must stop at the correct cisterna and somehow be prevented from entering the vesicles budding from the dilated rims that carry other proteins further along the pathway [2]. Results from many laboratories show that the signal for retention lies not in the cytoplasmic or lumenal domains of the protein but in the spanning domain [6–15] . In many cases, retention does not appear to require the flanking, charged amino acids [6,10,12,14] so any model for retention must explain how an uncharged and hydrophobic stretch of amino acids can function to retain proteins in the Golgi stack.

2. THE MODEL

We would like to propose a model that explains most of the experimental observations. Golgi enzymes would form long hetero-oligomers by alternating interactions between the lumenal and spanning domains. By attaching the cytoplasmic domains to an underlying matrix, these oligomers would be prevented from entering the budding vesicles (Fig. 1). Each Golgi enzyme would be a homo-dimer in which the lumenal domains are bound to each other but the spanning domains are free to interact specifically with those of their neighbours (Fig. 1A). Different enzymes will interact with each other providing they share the same cisterna and each cisterna will contain a unique set of hetero-oligomers. Kin recognition would therefore explain the compartmental organisation of the Golgi stack [2].

This model is based on three assumptions. First, that Golgi enzymes are homodimers because their lumenal domains are bound together. The lumenal domain comprises both a catalytic domain and a stalk which connects it to the spanning domain (Fig. 1A). The catalytic domains can be separated from the rest of the protein by proteolytic cleavage both in vitro and in vivo and, in all cases so far analysed, have been shown to be tightly-bound dimers [16–18]. The stalks may also bind to each other but there is no direct evidence for this. The presence of acidic and basic amino acids in most stalks would certainly aid binding and there have been several studies implicating the stalks in the retention of some Golgi enzymes [8,11,14]. Of particular interest are those studies where the catalytic domains have been replaced by reporter molecules that are known to be monomers in their native state [6,8,9,13,15]. In most cases the stalk was preserved or replaced by sequences that could serve the same function. Interaction between the stalks could then generate the homodimers demanded by the proposed model. In other studies the reporter molecule was itself an oligomer which explains why the requirement for a lumenal interaction was not detected.

The second assumption is that the spanning domains of Golgi enzymes can specifically bind to those of their

*Corresponding author

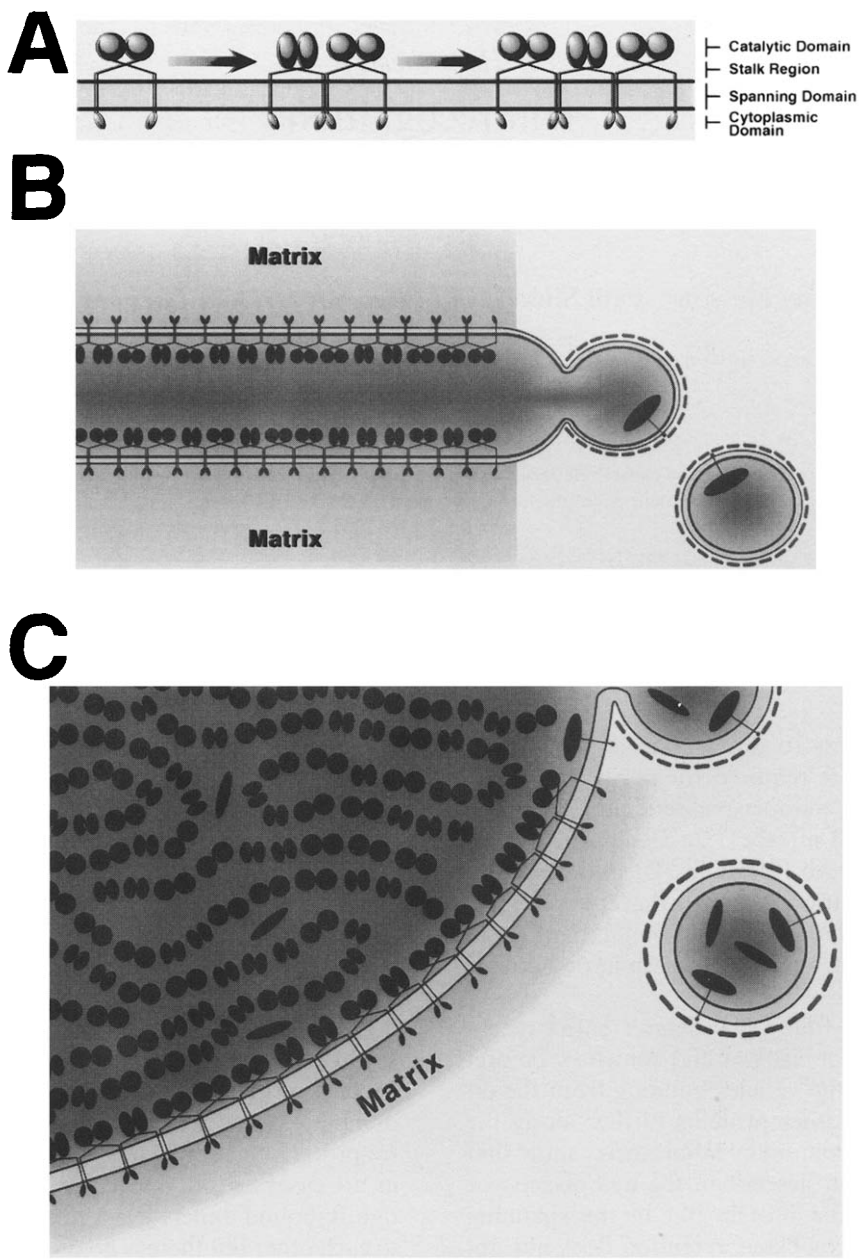

Fig. 1. Model for kin recognition of Golgi enzymes. (A) Each Golgi enzyme is assumed to be a homodimer [35] in which the lumenal domains are bound together by interactions between the catalytic domains and/or stalk regions. The spanning domains and perhaps the stalk regions of these dimers are free to bind to those of their kin at one [25] or more sites [7,11,12]. Cut-away side view (B) and top view (C) of a Golgi cisterna and transport vesicle showing the long, linear hetero-oligomers formed by kin recognition. Binding to the matrix prevents entry into the COP-coated vesicles budding from the dilated cisternal rims, whereas proteins undergoing transport are free to move

kin. The *medial* enzymes, *N*-acetylglucosaminyltransferase I (NAGT I) and mannosidase II (Mann II), should bind to each other but not to the *trans* enzyme, β1,4-galactosyltransferase (GalT). This has recently been tested by replacing the cytoplasmic domains of each of these enzymes by an ER retention signal [19]. Tagged NAGT I causes redistribution of Mann II from the Golgi to the ER and tagged Mann II has a similar effect on NAGT I. Neither have any effect on GalT and tagged GalT has no effect on either NAGT I or Mann II. These studies provide good evidence for kin recognition though they do not show directly that the interaction occurs via the spanning domains. Such an interaction has only been shown for a Golgi protein in one instance [6] and the problem is compounded by the lack of any obvious sequence homology that might be used to predict such an interaction [20]. Fortunately, such interactions are a well-established feature of plasma

membrane proteins. Glycophorin A will only bind synthetic peptides encompassing the spanning domain if they have the sequence of glycophorin A but not C [21] Spanning domains of glycophorin A dimerise in a highly specific manner [22] and modelling suggests that the best structure is a right handed supercoil [23]. The ζ chain of the T cell receptor (TCR) forms homodimers dependent on a cysteine and an aspartic acid in the spanning domain [24]. The oncogenic form of the neu receptor dimerises in the absence of ligand because of a single amino acid substitution in the spanning domain [25]. Furthermore, interaction can occur between the spanning domains of different proteins. The binding of the α chain of the TCR and the CD3 δ chain is stabilised by charged amino acids in the spanning domain [26]. Lastly, a truncated form of the E5 oncoprotein of papillomavirus containing only the spanning domain will bind to both the pore-forming protein of vacuolar ATPases and the PDGF receptor [27]. Together these data show that specific binding between different spanning domains can occur as required by the proposed model.

There are also indications that more than one type of hetero-oligomer can occupy the same cisterna. Approximately one third of the NAGT I and half the GalT in HeLa cells share the *trans* cisterna [28] yet they do not interact [19,29]. This means that the specificity of the interaction must depend on the spanning sequences themselves and not on the lipid composition or environment of the cisterna.

The third and last assumption is that the hetero-oligomers are anchored. Arranging the enzymes as long, linear structures would ensure that they had complete access to their substrates, the oligosaccharides attached to the proteins being transported through the stack. However, it is difficult to see how such oligomers, even if they were very long, could be prevented from entering the budding vesicles. The simplest solution is to assume that they are anchored to an underlying matrix by their cytoplasmic domains (Fig. 1B and C). This binding would be weaker than that between the spanning domains since the cytoplasmic domains alone cannot retain hybrid proteins [7]. We have recently isolated a Triton- and salt- resistant matrix from Golgi stacks that binds almost all of the NAGT I and Mann II and is present in the intercisternal space [29]. Such matrices, by binding to the enzymes of adjacent cisternae, would generate the characteristic Golgi stack.

3. EXPERIMENTAL OBSERVATIONS

The model explains several experimental observations made after over-expression of hybrid proteins. In all cases, the Golgi retention system could not be saturated even at expression levels several hundred-fold higher than normal [7–9,11,13]. In our model, the excess would simply add on to the end of pre-existing oligomers causing an increase in their length. Over-expression frequently causes the hybrid proteins to back up into the ER [7–9,11,13]. This could be explained by premature oligomerisation.

There are some experimental observations that are not explained by the model. The first are those that cast doubt on the specificity of interaction between spanning domains. Those parts of the domain responsible for the retention of GalT have been mapped by mutagenesis in two studies, with contradictory results [7,12]. The entire spanning domain of α2,6 sialyltransferase (SialylT) has been replaced by a stretch of 17 leucines and this does not affect the ability of the hybrid protein to reach the *trans* Golgi [8]. Increasing the length to 23 leucines causes the protein to appear on the cell surface suggesting that length can play an important role in retention. These discrepancies will only be resolved once the structure of these interacting domains is solved by the application of biophysical methods.

The second set of observations are those using hybrid proteins where the lumenal domain is a reporter molecule that is a monomer in the native state [6,8,9,13,15]. These are, nevertheless, retained in the Golgi, even in stable cell lines [15]. One explanation, mentioned above, is that the stalks alone could generate the needed homodimers but in some hybrid proteins even the stalks have been replaced and the substituted sequences do not have any features that would suggest they might dimerise. If the hybrid proteins are monomers one could imagine that they add on to the ends of existing oligomers, effectively capping them. A similar consequence would be predicted from an experiment where the lumenal domain alone of SialylT was expressed [11]. The stalk region caused retention perhaps by binding to an endogenous, full-length monomer in the ER, forming a truncated dimer. These dimers could still bind to the ends of oligomers but the absence of one of the spanning domains would effectively cap them. The problem is that for both types of hybrid protein the number of molecules needed to cap all of the oligomers would be very small. More would break up the oligomers leading to a massive loss of Golgi enzymes. One solution to this problem is to suggest that these lost Golgi enzymes would be rapidly retrieved, as are soluble ER proteins. Several lines of evidence suggest that Golgi enzymes are retrieved [28], but the signal is not known. Soluble ER proteins are retrieved by a receptor which recognises the C-terminal tetra-peptide, -KDEL [5]. For membrane proteins, the signals for moving them from one membrane to another are almost always found in the cytoplasmic domain [30]. At first sight these domains do not appear to be suitable candidates. They improve the retention efficiency of hybrid proteins in the Golgi but this role can be served by the cytoplasmic domains of the reporter proteins [7]. Of course, these reporter proteins may themselves, on occasion, need to be retrieved so it is possible to imagine a mechanism which recognises the

cytoplasmic domains on proteins that have moved beyond their correct location and return them to the ER. From here they could move back to their correct position in the cell determined by the spanning region.

Such an idea would physically separate the retrieval and retention signals and this idea is borne out by work on ER proteins. When the -KDEL sequence is removed from the immunoglobulin binding protein (BiP), the protein is secreted, but very slowly [5], suggesting that there is also a retention signal. Some ER membrane proteins have a double-lysine motif which was thought to act as a retention signal [31,32]. However, when this motif was grafted onto a plasma membrane protein, it stayed in the ER but acquired oligosaccharide modifications showing that it had passed through the Golgi apparatus [33]. The motif is therefore a retrieval and not a retention signal which must lie elsewhere in the protein. Recent work on other ER proteins shows that this retention signal probably resides in the spanning domain [34]. We would, therefore, suggest that all proteins along the secretory pathway have both retention and retrieval signals and that kin recognition, operating through the spanning domains, will prove to be the primary mechanism for keeping proteins in the correct location.

REFERENCES

[1] Pfeffer, S.R. and Rothman, J.E. (1987) Annu. Rev. Biochem. 829–852.
[2] Rothman, J.E and Orci, L. (1992) Nature 355, 409–416.
[3] Kornfeld, S. and Mellman, I (1989) Annu. Rev Cell Biol. 5, 483–525
[4] Burgess, T L. and Kelly, R.B. (1987) Annu. Rev. Cell Biol. 3, 243–293.
[5] Pelham, H.R. (1989) Annu. Rev. Cell. Biol. 5, 1–23.
[6] Swift, A.M. and Machamer, C.E (1991) J. Cell Biol. 115, 19–30.
[7] Nilsson, T., Lucocq, J.M., Mackay, D. and Warren, G (1991) EMBO J. 10, 3567–3575
[8] Munro, S (1991) EMBO J. 10, 3577–3588.
[9] Teasdale, R.D., D'Agostaro, G. and Gleeson, P.A. (1992) J Biol. Chem. 267, 4084–4096.
[10] Wong, S.H., Low, S.H. and Hong, W. (1992) J. Cell Biol. 117, 245–258.
[11] Colley, K.J., Lee, E.U and Paulson, J.C. (1992) J. Biol Chem. 267, 7784–7793.
[12] Aoki, D., Lee, N., Yamaguchi, N., Dubois, C and Fukuda, M.N. (1992) Proc Natl. Acad. Sci. USA 89, 4319–4323.
[13] Russo, R.N., Shaper, N.L , Taatjes, D.J and Shaper, J H (1992) J. Biol. Chem. 267, 9241–9247.
[14] Tang, B.L., Wong, S.H., Low, S.H. and Hong, W. (1992) J. Biol. Chem. 267, 10122- 1–126.
[15] Burke, J., Pettitt, J.M., Schachter, H., Sarkar, M and Gleeson, P.A. (1992) J. Biol Chem. 267, 24433–24440.
[16] Navaratnam, N., Ward, S., Fisher, C., Kuhn, N.J., Keen, J.N. and Findlay, J.B.C. (1988) Eur. J. Biochem. 171, 623–629.
[17] Khatra, B.S., Herries, D.G. and Brew, K. (1974) Eur. J. Biochem 44, 537–560.
[18] Moremen, K.W., Touster, O. and Robbins, P.W. (1991) J. Biol. Chem. 266, 16876–16885.
[19] Nilsson, T., Hoe, M., Slusarewicz, P., Watzele, G., Berger, E.G. and Warren, G. (1993) submitted.
[20] Machamer, C.E. (1991) Trends Cell Biol. 1, 141–144.
[21] Bormann, B.-J , Knowles, W.J. and Marchesi, V.T (1989) J. Biol. Chem. 264, 4033–4037.
[22] Lemmon, M.A., Flanagan, J.M., Treutlein, H.R., Zhang, J. and Engelman, D.M. (1992) Biochemistry 31, 12719–12725.
[23] Treutlein, H.R., Lemmon, M.A., Engelman, D.M. and Brünger, A T (1992) Biochemistry 31, 12726–12733.
[24] Rutledge, T., Cosson, P., Manolios, N , Bonifacino, J.S. and Klausner, R.D (1992) EMBO J. 11, 3245–3254.
[25] Sternberg, M.J.E. and Gullick, W.J. (1989) Nature 339, 587.
[26] Cosson, P., Lankford, S.P., Bonifacino, J.S. and Klausner, R.D. (1991) Nature 351, 414–416.
[27] Goldstein, D., Andresson, T , Sparkowski, J.J. and Schlegel, R. (1992) EMBO J 11, 4851–4859.
[28] Nilsson, T., Pypaert, M., Hoe, M.H., Slusarewicz, P , Berger, E.G and Warren, G (1993) J. Cell Biol. 120, 5–14
[29] Slusarewicz, P., Nilsson, T., Hui, N., Watson, R. and Warren, G. (1993) submitted
[30] Hopkins, C R. (1992) Trends Biochem. Sci. 17, 27–31
[31] Nilsson, T., Jackson, M.R. and Peterson, P.A. (1989) Cell 58, 707–718.
[32] Jackson, M R., Nilsson, T. and Peterson, P.A. (1991) EMBO J. 9, 3153–3162
[33] Jackson, M.R., Nilsson, T. and Peterson, P.A. (1993) J Cell Biol. 121, 317–333
[34] Smith, S. and Blobel, G. (1993) J. Cell Biol 120, 631–637.
[35] Fleischer, B., McIntyre, J.O. and Kempner, E.S (1993) Biochemistry 32, 2076–2081.