

## Hypothesis

## Folding intermediates are involved in genetic diseases?

Valentina E. Bychkova<sup>a,\*</sup>, Oleg B. Ptitsyn<sup>a,b</sup><sup>a</sup>*Institute of Protein Research, Russian Academy of Sciences, 142292, Pushchino, Moscow Region, Russian Federation*<sup>b</sup>*Laboratory of Mathematical Biology, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892-5677, USA*

Received 22 December 1994

**Abstract** Recent experimental data show that some human genetic diseases are due to mutations in proteins which influence their trafficking and lead to retaining of proteins in the endoplasmic reticulum or their unproper processing. In this paper a hypothesis is proposed that these mutations are connected with an incomplete protein folding, blocking it at the stage of the kinetic molten globule or even earlier. If so, the specific drugs against these diseases may be ligands and other factors which facilitate the correct protein folding.

**Key words:** Protein folding; Genetic disease; Molten globule state; Mislocation of proteins

## 1. Introduction

It has been shown recently that some point mutations in a number of proteins lead to genetic diseases. These mutations mislocate proteins in a cell leading to the loss of their function (see [1] for review). What is the structural basis for this mislocation? The intriguing possibility is that the molten globule or similar state of protein molecules [2,3] may be involved in this process.

## 2. Some point mutations transform proteins into the molten globule state

There are in vitro experimental data showing that point mutations can hinder last stages of protein folding and arrest protein folding intermediates under physiological conditions.

The first evidence was obtained by King and his collaborators [4,5]. They have shown that some thermosensitive mutants of tailspike endorhamnosidase of phage P22 cannot fold correctly at non-permissive temperature (39–42°C). However, at lower temperature (30°C) some of these mutants can fold into the native structure and assemble correctly [6]. When folded, they remain stable at 42°C. It means that the lack of intracellular folding is due to the decreased stability of folding intermediates rather than of native proteins [7,8].

Craig et al. have reported [9] that two point mutations in  $\beta$ -lactamase (Thr-40→Ile and Asp-146→Asn) prevent the protein from the correct folding and trap it into an intermediate state.

Machamer et al. [10,11] have shown that thermosensitive

mutants of G protein of vesicular stomatitis virus with altered glycosylation sites also fold in a way different from that of the native protein. These mutants usually have non-native disulphide bonds and are retained in the endoplasmic reticulum (ER) in a complex with chaperones.

Lim et al. [12] have investigated point mutations in a non-polar core of the N-terminal domain of  $\lambda$ -repressor. Substitutions of non-polar groups by polar ones (Leu-18→Asn, Val-36→Asn and Val-36→Asp) drastically change the properties of mutant proteins. These proteins lose activity, do not bind antibodies against the native protein, have no cooperative temperature melting, their NMR resonances become broad and less dispersed than in the native protein. However, they retain a lot of their native far UV CD ellipticity, and their NMR spectra still have resonances well shifted from their random coil positions. The authors concluded that folding of these mutant proteins is blocked in the molten globule state.

Quite recently Horowitz and his group [13,14] have shown that replacements of Lys-249 or Arg-186 change specificity, stability and structural state of rhodanase. Mutant proteins have an increased hydrophobic surface, because they bind ANS, and quenching of their Trp fluorescence by acrylamide is increased. It may be suggested that these mutant proteins are in the molten globule state.

These examples show that even a single point mutation can transform the native protein into the molten globule.

## 3. Some point mutations associated with genetic diseases lead to a mislocation of proteins in a cell

### 3.1. Cystic fibrosis transmembrane conductance regulator

Cystic fibrosis is a human genetic disease caused by mutations in *cystic fibrosis transmembrane conductance regulator* (CFTR) [15–17]. In 70% of cases this disease is associated with a deletion of Phe-508 (mutant  $\Delta F$  508). Newly synthesized CFTR forms a transient complex with chaperone hsp70. The wild type protein dissociates from this complex and is transported to the Golgi with a subsequent traffic to a plasma membrane [15]. As a contrast, the mutant  $\Delta F$  508 remains in a complex with hsp70 in the ER and is rapidly degraded in a pre-Golgi non-lysosomal compartment [15]. It has been proposed that the deletion of Phe-508 perturbs protein folding, leads to its tighter binding to hsp70, and to retention of the mutant in the ER implicating the defect of its trafficking. The idea that this mutant has a defective folding is supported by a temperature sensitivity of this mutant [16,17].

Thus, hsp70 can discriminate between normal and mutant forms of CFTR. These [15] and other [11,18] data suggest that

\*Corresponding author. Fax: (7) (095) 924-0493.  
E-mail: protres@ipr.serpukhov.su

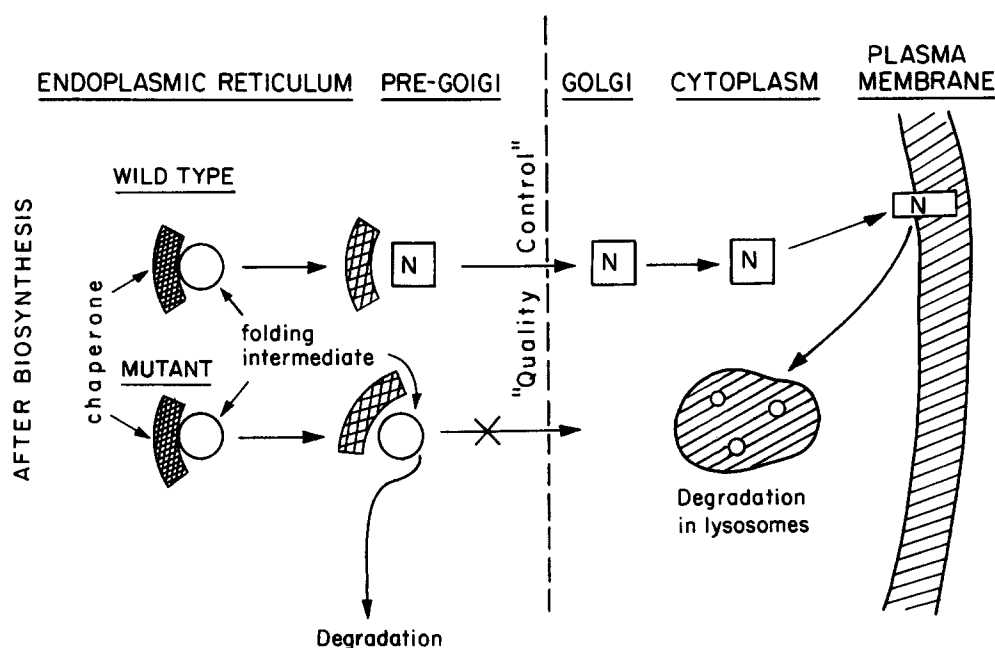


Fig. 1. Cell mislocation of mutant proteins associated with some genetic diseases.

a chaperone performs quality control during the biosynthesis of plasma-membrane and secreted proteins. Mutants which cannot be released from chaperones remain in the ER where they are degraded [11,15,18] instead of being transported from the ER to reach their destination in a cell.

### 3.2. $\alpha$ 1-antitrypsin

Another example is *emphysema* and *chronic liver diseases*, caused by the insufficiency of  $\alpha$ 1-antitrypsin (AAT), the serine protease inhibitor (see [1,19] and references therein). Under the mild denaturing conditions (1 M GdmCl and 37–41°C) AAT has a tendency to aggregate presumably due to the moving of 14-residues loop which contains the reaction inhibiting site [20].

The Z mutant of  $\alpha$ 1-antitrypsin has Glu-342 substituted by Lys [20]. Glu-342 is located at the beginning of the reactive loop, which is stabilized by salt bridge between Glu-342 and Lys-290 in the wild type protein. The mutation Glu-342→Lys leads to disruption of the salt bridge and makes this loop more flexible and  $\beta$ -sheet more 'open'. As a result, the reactive loop can interact with other AAT molecules being in similar 'open' conformation. This promotes AAT aggregation in the ER even under physiological conditions [20] and probably also its association with chaperones [19,21]. Thus, in this case we again face with a mislocation of protein in a cell.

### 3.3. Other proteins

A similar mechanism may operate also in the pathogenesis of some other diseases, including for example *hypercholesterolemia*, caused by mutation on the *low density lipoprotein* (LDL) receptor, responsible for the serum cholesterol level. Class 2 mutations of LDL receptor (50% of all cases) results in a protein which cannot be secreted from the ER [1,22,23]. Another example is the *Tay-Sachs disease* which is connected with mutations in *lysosomal enzymes*, which dimeric forms are finally processed in lysosomes. At least one of these mutations results in defective protein trafficking: the mutant of  $\beta$ -hexosaminidase is not secreted from the ER, not processed in lysosomes, and is degraded in the ER [1,24].

The list of mutant proteins which mislocation leads to different types of genetic diseases is given in Table 1.

## 4. Hypothesis

It was predicted [25] and confirmed experimentally [26–30] that chaperones recognize proteins in the molten globule (or a little more expanded [29,30]) state. It supports the idea that chaperones bind newly synthesized protein chains trapping them in the molten globule state or even in the earlier kinetic intermediate. On the other hand, there is also some direct evi-

Table 1  
Some point mutations in proteins and related genetic diseases

Protein	Mutations	Disease	Ref.
Cystic fibrosis transmembrane conductance regulator	Deletion of F508	Cystic fibrosis	[15]
$\alpha$ 1-antitrypsin	Glu342→Lys	Emphysema, chronic liver diseases	[20]
Low density lipoprotein receptor	Class 2	Familial hypercholesterolemia	[22,23]
$\beta$ -hexosaminidase ( $\alpha$ -subunit)	Deletion of cytosine at position 1510 of the mRNA coding sequence	Tay-Sachs disease	[24]

dence that a protein should be transformed into the molten globule state before its degradation *in vivo* (see [3] for a review). Thus mutations which block protein folding in the molten globule or in the earlier kinetic intermediate can lead to dislocation of a protein in a cell and facilitate its degradation.

We suggest that mutations which promote some genetic diseases connected with an intracellular misrouting of proteins belong to the mutations which inhibit the last stage of protein folding. As a result, mutant proteins are trapped in the molten globule or in another kinetic intermediate. This does not allow mutant proteins to use a normal intracellular pathway and leads to their mislocation and subsequent degradation.

Thus we propose the hypothesis that *incomplete protein folding blocked at the kinetic molten globule or another intermediate state leads to mislocation of proteins in a cell and may provoke genetic diseases*. Fig. 1 illustrates this hypothesis. It shows that a wild type protein is bound to a chaperone just after its biosynthesis, then is released in the ER, and is transported to a cell membrane or to another cell compartment where it normally functions. On the contrary, a mutant protein which is in the molten globule or another denatured state cannot be released from chaperone and therefore is retained in the ER where it is later degraded.

Quite recently experimental data on tumor suppressor protein p53 appeared which seem to be consistent with our hypothesis. Among many different functions of this protein in a cell there is the preventing from cancer. Mutations of p53 which lead to inactivation of its tumor suppressor activity contribute to development of many human cancers (see [31–33] for references). One class of these mutations inhibits binding p53 to DNA, while another class leads to denaturation of this protein [32]. In fact, mutants of the second class bind antibodies specific to denatured rather than to native p53 and are highly sensitive to proteolytic enzymes.

The mutants of the second class (for example, replacements of Arg-175 and Val-143) are bound by heat shock proteins hsp70 [34,35], and it was also shown that some of these mutants are mislocated in a cell [34].

Thus, at least in one case mutations of a protein lead to its denaturation and to its mislocation in a cell, which is consistent with our hypothesis. Of course, a lot of work should be performed to check this hypothesis by direct structural data.

If this hypothesis is correct, a study of the influence of ligands and other factors on protein folding *in vitro* may lead to a rational design of new drugs against some genetic diseases. An example of these drugs can be antibodies against the native structure as well as other templates which might facilitate protein folding. In another case, these may be small specific peptides which can compete with flexible parts of protein molecules for intermolecular interactions [20].

This hypothesis emphasizes how important it is to study the non-native states of proteins in a cell for different biological and even medical purposes including the treatment of some genetic diseases.

**Acknowledgements:** We are grateful to the referee of our paper who attracted our attention to mutations of p53 tumor suppressor. This research was supported in part by grant from the International Human Frontier Science program Organization (Grant RG-331/93).

## References

- [1] Amara, J.F., Cheng, S.H. and Smith, A.E. (1992) *Trends Cell Biol.* 2, 145–149.
- [2] Ptitsyn, O.B. (1992) in: *Protein Folding*, (Creighton, T.E. ed.) pp. 243–300, W.H. Freeman and Co., New York.
- [3] Bychkova, V.E. and Ptitsyn, O.B. (1993) *Chemtracts-Biochem. Mol. Biol.* 4, 133–163.
- [4] Goldenberg, D.P. and King, J. (1981) *J. Mol. Biol.* 145, 633–651.
- [5] Smith, D.H. and King, J. (1981) *J. Mol. Biol.* 145, 653–676.
- [6] Goldenberg, D.P., Smith, D.H. and King, J. (1983) *Proc. Natl. Acad. Sci. USA* 80, 7060–7064.
- [7] Haase-Pettingell, C.A. and King, J. (1988) *J. Biol. Chem.* 263, 4977–4983.
- [8] Sturtevant, J.M., Yu, M., Haase-Pettingell, C.A. and King, J. (1989) *J. Biol. Chem.* 264, 10963–10968.
- [9] Creig, S., Hollecker, M., Creighton, T.E. and Pain, R.H. (1985) *J. Mol. Biol.* 185, 681–687.
- [10] Machamer, C.E. and Rose, J.K. (1988) *J. Biol. Chem.* 263, 5955–5960.
- [11] Machamer, C.E., Doms, R.W., Bole, D.G., Helenius, A. and Rose, J.K. (1990) *J. Biol. Chem.* 265, 6879–6883.
- [12] Lim, W.A., Farruggio, D.C. and Sauer, R.T. (1992) *Biochemistry* 31, 4324–4333.
- [13] Islam, T.A., Miller-Martini, D.M. and Horowitz, P.M. (1994) *J. Biol. Chem.* 269, 7908–7913.
- [14] Luo, G.X. and Horowitz, P.M. (1994) *J. Biol. Chem.* 269, 8220–8225.
- [15] Yang, Y., Janich, S., Cohn, J.A. and Wilson, J.M. (1993) *Proc. Natl. Acad. Sci. USA* 90, 9480–9484.
- [16] Denning, G.M., Anderson, M.P., Amara, J.F., Marshall, J., Smith, A.E. and Welsh, M.J. (1992) *Nature* 358, 761–764.
- [17] Thomas, P.J., Shenbagamurthi, P., Sonek, J., Hulihan, J.M. and Pedern, P.L. (1992) *J. Biol. Chem.* 267, 5727–5730.
- [18] Accili, D., Kadowaki, T., Kadowaki, H., Mosthaf, L., Ullrich, A. and Taylor, S.I. (1992) *J. Biol. Chem.* 267, 586–590.
- [19] Sifers, R.N. (1992) *Nature* 357, 541–542.
- [20] Lomas, D.A., Evans, D.L., Finch, J.T. and Carrell, R.W. (1992) *Nature* 357, 605–607.
- [21] Le, A., Ferrell, G.A., Dishon, D.S., Le, Q.-Q. and Sifers, R.N. (1992) *J. Biol. Chem.* 267, 1072–1080.
- [22] Hobbs, H.H., Russell, D.W., Brown, M.S. and Goldstein, J.L. (1990) *Annu. Rev. Genet.* 24, 133–170.
- [23] Yamamoto, T., Bishop, R.W., Brown, M.S., Goldstein, J.L. and Russell, D.W. (1986) *Science* 232, 1230–1237.
- [24] Lau, M.M.H. and Neufeld, E.F. (1989) *J. Biol. Chem.* 264, 21376–21380.
- [25] Bychkova, V.E., Pain, R.H. and Ptitsyn, O.B. (1988) *FEBS Lett.* 238, 231–234.
- [26] Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A.L. and Hartl, F.-U. (1991) *Nature* 352, 36–42.
- [27] Martin, J., Horwich, A.L. and Hartl, F.-U. (1992) *Science* 258, 995–998.
- [28] Van der Vies, S.M., Viitanen, P.V., Gatenby, A.A., Iorimer, G.H. and Janike, R. (1991) *Biochemistry* 31, 3635–3644.
- [29] Okazaki, A., Ikura, T., Nikaido, K. and Kuwajima, K. (1994) *Nature Struct. Biol.* 1, 439–445.
- [30] Robinson, C.V., Gros, M., Eyles, S.J., Ewbank, J.J., Mayhew, M., Hartl, F.-U., Dobson, C.M. and Radford, S.E. (1994) *Nature* 372, 646–651.
- [31] Friend, S. (1994) *Science* 265, 334–335.
- [32] Cho, Y., Gorina, S., Jeffrey, P.D. and Pavletich, N.P. (1994) *Science* 265, 346–355.
- [33] Clore, M.G., Omichiuski, J.G., Sakaguchi, K., Zambrano, N., Sakamoto, H., Appella, E. and Gronenborn, A.M. (1994) *Science* 265, 386–391.
- [34] Sturzbecher, H.-W., Chumakov, P., Welch, W.J. and Jenkins, J.R. (1987) *Oncogene* 1, 201–211.
- [35] Hinds, P.W., Finlay, C.A., Quartin, R.S., Baker, S.J., Fearon, E.R., Vogelstein, B. and Levina, A.J. (1990) *Cell Growth Diff.* 1, 571–580.