

# Expedience of Protein Folding Modeling during Progressive Elongation of Polypeptide Chain

M.A. Basharov<sup>\*,1</sup> and S.I. Allakhverdiev<sup>\*,2</sup>

<sup>1</sup>Institute of Theoretical and Experimental Biophysics, <sup>2</sup>Institute of Basic Biological Problems, Russian Academy of Sciences, Pushchino, Moscow Region 142290, Russia

**Abstract:** The problem of protein folding, i.e. how does a polypeptide chain fold to native protein following its synthesis on the ribosome, is recognized as a major unsolved problem of biology at molecular level. To solve the problem experimental studies of denaturation and renaturation of native proteins and a variety of theoretical-computational simulations of full-length polypeptide chains have usually been used as relevant *in vitro* models. However, at present these conventional approaches evidently seem to be hypothetical and have been hardly found. Nevertheless, there are a lot of convincing evidences that proteins fold progressively during the residue-by-residue elongation on the ribosome from the N- to the C-terminus. On this basis, therefore, simulations of the folding and formation of the native spatial structure of the proteins will be expedient. These points are briefly highlighted in the current minireview.

The problem of protein folding is formulated simply as either how a protein molecule acquires the native spatial structure following its synthesis on the ribosome, or how does the primary amino acid sequence determine protein native tertiary structure, or what is the mechanism of folding the proteins. To solve the problem the traditional *in vitro* approaches present protein folding as a spontaneous process during which the formation of the native spatial structure of proteins occurs from an unfolded random coil state both *in vivo* and *in vitro* [1-5]. For example, unfolded full-length polypeptide chains are usually considered as the objects under study to explore the mechanism of folding the proteins following their synthesis on the ribosome. The consideration has been developed in the 1970's on the basis of the results of the classical experiments on denaturation and renaturation of the small single-domain proteins [6], herewith of special importance were the results of the experiments on bovine pancreatic ribonuclease [7-10]. Such experiments demonstrated: firstly, proteins unfold and adopt a random coil state under strongly denaturing conditions with no residual ordered structure being present, and secondly, some of the denatured proteins restore their initial native state relatively quickly and spontaneously after removal of denaturation [6-10].

At present, it is evidently clear that neither denatured nor newly synthesized protein is randomly coiled. A slight perceptible indication of residual regular structures was detected in some proteins under strongly denaturing conditions yet in the early classical denaturation experiments [6, 11, 12].

Since then the evidence had been accumulated in a great body of new generation of precise kinetic experiments that the denatured state of proteins was far from the random chain. Undoubtedly, at present almost all the proteins, including those examined in the classical denaturation experiments, do not unfold fully and attain a random coil state under any, even strongly denaturing conditions (such as 6M guanidine-hydrochloride, 8M urea, or high temperature), however they contain a considerable amount of residual both native and non-native folded structures [13]. This fact causes doubt on the basis of evidence of utilization of unfolded random coil polypeptides in the studies of the protein folding process. If the same fact is taken into account, the phenomenon of spontaneous renaturation of some denatured proteins has plain explanation as well, since manifestation of the physical properties of the proteins *per se*, was caused by the residual folded structure available, because a native protein is used initially in the denaturation-renaturation experiments [14]. The residual structural elements can really be the initiating centers of renaturation as soon as regeneration of the initial native spatial structure of the protein starts.

Relatively fast and spontaneous restoration of the initial native state of the denatured protein after removal of the denaturant is not a common phenomenon in the experiments on renaturation. Usually, only small proteins without the disulfide bond renature quickly provided they are denatured by sparing denaturing condition. For a disulfide bond containing proteins with reduced disulfide bonds regeneration of the native spatial structure in the renaturation experiments is well known to be more problematic. For any protein, the time and the quality of renaturation almost without exception strongly depends on the type and power of denaturing action, on the renaturation condition used, as well as on the physical properties of protein the *per se*. Depending both on the flexibility (or stiffness) of the native structure and the character and the strength of denaturing

\*Address correspondence to these authors at the <sup>1</sup>Institute of Theoretical and Experimental Biophysics, <sup>2</sup>Institute of Basic Biological Problems, Russian Academy of Sciences, Pushchino, Moscow Region 142290, Russia; Tel: 7(496) 773-1837; 7-926-853-321; Fax: 7 (496)-7330532; E-mails: m-basharov@rambler.ru (Basharov); suleyman@issp.serpukhov.su, suleyman20042000@yahoo.com (Allakhverdiev)

action, the protein may exhibit such properties as stability, or elasticity (most likely the property of rubber-like elasticity), or irreversible denaturation (probably, the property of glassiness) after denaturation [14].

There is much convincing evidences accumulated since the early 1960's that the proteins fold during the synthesis on the ribosome or during translocation of the nascent polypeptides into the cellular compartments across the membranes. First of all, there is a great body of experiments, both *in vivo* and on the model systems, demonstrating that polypeptides synthesized on the ribosome exhibit enzymatic activity, bind conformational antibodies, cofactors or ligands especially for the corresponding mature proteins, and form the correct intra- and interchain disulfide bonds and protease-resistant compact structures ([14], for references). Besides, such processes in the cells as the binding of the signal recognition particles to the N-terminal signal sequences and of chaperones to the specific sequences of most of newly synthesizing polypeptides in the cytoplasm, as well as the splitting of the signal sequences of nascent polypeptides by the signal peptidases in the cellular compartments are observed occur when these sequences are properly folded during the synthesis and translocation from the ribosome or during the translocation into the compartments across the membranes [15].

In addition to the above, numerous data are available of cell biology on the rates of synthesis of polypeptide chains on the ribosome in the cells and of physical chemistry on the formation of secondary structure elements ( $\alpha$ -helices,  $\beta$ -sheets, and turns) and the compact states in the unfolded polypeptides ([14], for references). In comparison of these data it becomes clear that elongation of a polypeptide chain on the ribosome by every of its individual amino-acid residue occurs much more slowly (the minimal time required is no less than  $10^{-2}$ s) relative to the secondary structure elements, and the compact structure is formed in an unfolded polypeptide ( $10^{-7} - 10^{-4}$ s and  $10^{-5} - 10^{-3}$ s, respectively). Consequently, during elongation of a polypeptide chain by each of its residue on the ribosome, either the secondary or the compact structures should be formed of the N-terminal part of the chain that has emerged from the ribosome in the cytoplasm. This is sound deductive reason for sequential folding of the proteins during the residue-by-residue elongation on the ribosome from the N- to the C-terminus, which allows a tentative folding model to be proposed as well [14].

Undoubtedly, all the information required to determine the native secondary and spatial structures of the proteins is based on their genetically determined primary linear amino acid sequence [7, 8]. The problem of protein folding is to clarify how this information is realized for determination of the native structure of a protein molecule following its synthesis on the ribosome [1-3], or how does the primary

amino acid sequence determine protein native tertiary structure to predict the native structure from the known amino acid sequence. An inference of this minireview is that utilization of the full-length polypeptide chains and their unfolded states seems evidently hypothetical to solve the problem of protein folding. Meanwhile, there is a lot of irrefutable evidence that the information contained in the primary structure of a protein about its spatial structure is realized in the cells during synthesis on the ribosome, and numerous data available is a sound argument for sequential folding of the polypeptides, gradually during the residue-by-residue elongation from the N- to the C-terminus. Thus, there are a lot of reasons and prerequisites for initiation of the studies on the development of the approaches and the methods for simulations of the folding and formation of native conformation of the proteins based on progressive residue-by-residue elongation of their polypeptide chains from the N-terminus.

## ACKNOWLEDGEMENT

This work was funded by the Russian Academy of Sciences.

## REFERENCES

- [1] Radford SE, Dobson CM. From computer simulations to human disease: emerging themes in protein folding. *Cell* 1999; 97: 291-8.
- [2] Daggett V, Fersht AR. The present view of the mechanism of protein folding. *Nat Rev Mol Cell Biol* 2003; 4: 497-502.
- [3] Dobson CM. Protein folding and misfolding. *Nature* 2003; 426: 884-90.
- [4] Baldwin RL. Energetics of protein folding. *J Mol Biol* 2007; 371: 283-301.
- [5] Ferguson N, Sharpe TD, Johnson CM, Schartau PJ, Fersht AR. Structural Biology: Analysis of 'downhill' protein folding. *Nature* 2007; 445: E14 - E15.
- [6] Tanford C. Protein denaturation. *Adv Protein Chem* 1968; 23: 121-282.
- [7] White FH, Jr. Regeneration of native secondary and tertiary structures by air oxidation of reduced Ribonuclease. *J Biol Chem* 1961; 236: 1353-60.
- [8] Anfinsen CB, Haber E. Studies on the reduction and re-formation of protein disulfide bonds. *J Biol Chem* 1961; 236: 1361-63.
- [9] Anfinsen CB, Haber E, Sela M, White FH, Jr. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc Natl Acad Sci USA* 1961; 47: 1309-14.
- [10] Anfinsen CB. Principles that govern the folding of protein chains. *Science* 1973; 181: 223-30.
- [11] Aune KC, Salahuddin A, Zarlengo MH, Tanford C. Evidence for residual structure in acid- and heat-denatured proteins. *J Biol Chem* 1967; 242: 4486-89.
- [12] Grebenshikov YuB, Irjak VI, Kuzub LI, Kush PP, Enikolopyan NS. Study of conformations of polymer chains in concentrated solutions by paramagnetic labeling. *Dokl Acad Nauk USSR* 1973; 210: 1124-26.
- [13] Rose GD. Unfolded proteins. *Adv Protein Chem* 2002; 62: 1-389.
- [14] Basharov MA. Protein folding. *J Cell Mol Med* 2003; 7: 223-37.
- [15] Schatz G, Dobberstein B. Common principles of protein translocation across membranes. *Science* 1996; 271: 1519-26.