Functional and Structural Genomics and Medicine

The Protein Folding Puzzle is Solved by Viewing of it from Two Sides

Alexei V. Finkelstein*

Institute of Protein Research, Russian Academy of Sciences, Pushchino, Moscow Region, Russian Federation

Protein chain folding is a miracle. The protein chain is gene-encoded and initially has no structure (Figure 1, left panel). Its intricate structure, with every atom in its unique position, results from a spontaneous folding process (Figure 1, right panel).

This is as amazing as if a multicolored thread could produce a T-shirt itself!

The chain spontaneously finds its stable fold (Figure 1: from left to right) within minutes or faster (both in vitro and in living cells), although much more than the entire life-period of the Universe would be necessary to try all possible chain structures in search of the most stable one. This is called "the Levinthal's paradox" [1]. To solve it, various models of folding were proposed during a few decades (see [2] for a brief review).

In Figure 1 (adapted from [3]), the colored helix, β-strands (strips) and loops (bold lines) show the chain fixed in the folded (right) or semi-folded (in the middle) structures; the globular regions are dotted. The thin broken line shows the structureless chain (left) and unfolded parts of the intermediate. The amino acid residues (color beads) are gene-encoded. The most unstable semi-folded state acts as the free-energy barrier at the folding and unfolding pathways. Instability of this folding intermediate, which is typical of proteins, results from the additional (by natural or artificial selection) reinforcement of the correctly folded structure relatively to all its "incorrect" competitors.

However the models suggested before mid-90th fail to overcome the Levinthal's paradox for the most typical case when the globular structure stability is close to that of the unfolded chain (which is typical for proteins) and provide for no estimate of the folding rates (spanning, for folding of different proteins in vitro, over 11 orders of magnitude, see Figure 2).

The problem of protein folding rates has been first solved using the unfolding (not folding!) as the starting point, i.e., when the free-energy barrier between the globular and unfolded state (Figure 1, middle) was viewed "from the globule's side" (Figure 1: from right to left) [4].

The trick is that, firstly, the rates of the forward and reverse reactions coincide when the globular structure stability equals to stability of the unfolded chain (according to the "principle of detailed balance" well-known in physics). Secondly, it is much easier to imagine and investigate how the thread unfolds than how it obtains some certain fold among countless possibilities.

The resulting estimate of the folding time was TIME ~ 10ns × exp[(1 ± 0.5)$L^{2/3}$], where ≈ 10 ns is the experimentally known time of growth of a structure (e.g., an α-helix) by one residue, and L is the number of amino acid residues in the protein chain[4].

The validity of this theory (proposed two decades ago, when the experimental data were yet scarce) has been recently confirmed by all currently available experimental data

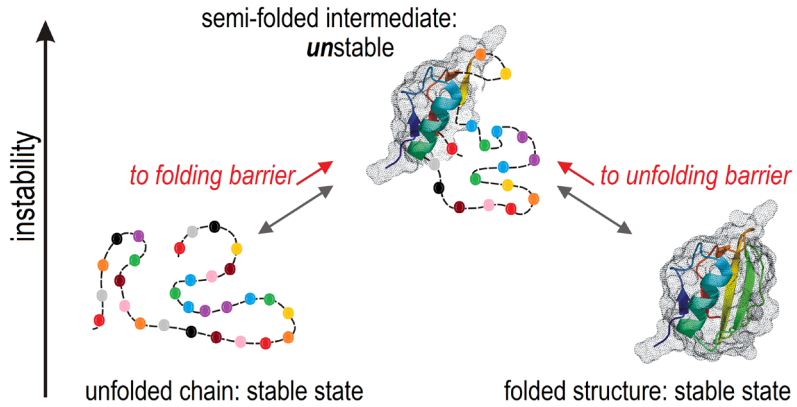

Figure 1: Protein folding and unfolding via an unstable semi-folded intermediate

Article Information

DOI: 10.31021/fsgm.20171101
Article Type: Letter to Editor
Journal Type: Open Access
Volume: 1 **Issue:** 1
Manuscript ID: FSGM-1-101
Publisher: Boffin Access Limited.

Received Date: 18 October 2017
Accepted Date: 03 December 2017
Published Date: 29 December 2017

***Corresponding author:**

Alexei V. Finkelstein
Institute of Protein Research
Russian Academy of Sciences
Pushchino
Moscow Region, Russian Federation
E-mail: afinkel@vega.protres.ru

Citation: Finkelstein AV. The Protein Folding Puzzle is Solved by Viewing of it from Two Sides. Funct Struct Genomics Med. 2017;1:101

[1,5] (Figure 2).

Yellow triangle in Figure 2 (adapted from [3]) shows the theoretically predicted (from consideration of unfolding!) range of these rates. The coincidence of experiment and theory is obvious.

However, a kind of dissatisfaction was felt, because in the folding problem has not been solved yet "from the viewpoint of the folding chain" (Figure 1: from left to right), and this initiated further efforts aimed to estimate a volume of conformations which must be tried by the folding protein chain in its search for its most stable fold [4,5].

Recently, this volume has been estimated at the level of formation and packing of the most strongly interacting protein structure elements (helices and β-strands, see Figures 1 and 3), as this has been suggested by Ptitsyn far ago [6].

It has been shown that this "Ptitsyn's volume" is ~ L^N (where N is the number of structural elements, which is at least by an order of magnitude smaller than L), i.e., this volume is by many orders of magnitude smaller than the "Levintal's volume", which, considered at the level of individual amino acid residues (beads in Figure 1) is ~ 3^L or even ~ 100^L [1,2,8]. The rate of search in the "Ptitsyn's volume", at folding, become physically and biologically reasonable (and more or less close to the maximal unfolding-derived estimates, ~ 10ns × exp [$1.5L^{2/3}$]; see Figure 2).

The netted shading in Figure 2 shows a recent theoretical estimate of the lower boundary of the rate of an exhaustive search,

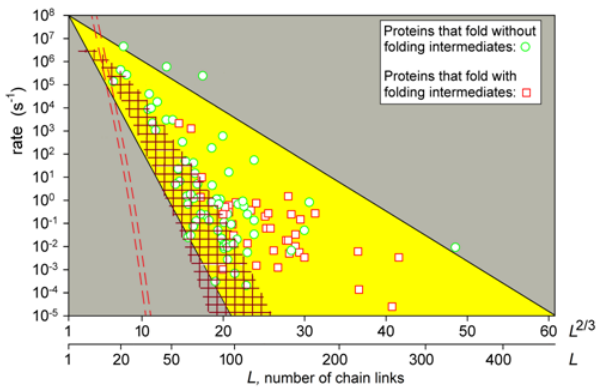

Figure 2: Folding rates (circles and squares) of proteins, experimentally studied at equal stability of their folded and unfolded states

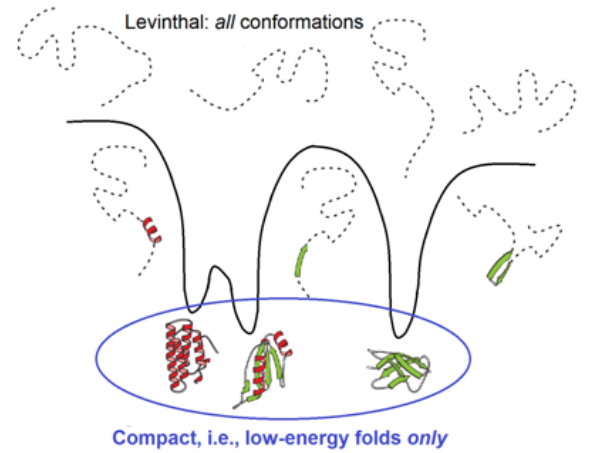

Figure 3: Comparison of a huge search among all, mostly disordered chain conformations, and a much less voluminous search only among compact and well-structured globules (corresponding to the deep energy minima). Adapted from [7]

at folding, of all possible packings of the protein secondary elements (helices and strands) [2,8]. The upper limit of the "Levinthal's search rate" is shown (in the same Figure) by the double dashed line.

Thus, the protein folding puzzle is solved by viewing of it from two sides: the side of unfolding and the side of folding.

The first part of this work has been partially supported by the Howard Hughes Medical Institute Awards and the Russian Academy of Sciences Program "Molecular and Cell Biology" (Grants No. 01200957492, 01201358029); the second part has been partially supported by the Russian Science Foundation Grant No. 14-24-00157.

References

1. Levinthal C (1969) How to fold graciously. In: Mössbauer Spectroscopy in Biological Systems: Proceedings of a meeting held at Allerton House, Monticello, Illinois (P. Debrunner, J.C.M. Tsibris, E. Munck, eds.). University of Illinois Press: Urbana-Champaign, IL 22–24.
2. Finkelstein AV, Badretdin AJ, Galzitskaya OV, Ivankov DN, Bogatyreva NS, et al. (2017) There and back again: Two views on the protein folding puzzle. Phys Life Rev 21: 56-71. doi: 10.1016/j.plrev. 2017;01:025
3. Finkelstein AV (2015) Two views on the protein folding puzzle. Atlas of Science, http://atlasofscience.org/two-views-on-the-protein-folding-puzzle/
4. Finkelstein AV, Badretdinov AY (1997) Rate of Protein Folding Near the Point of Thermodynamic Equilibrium between the Coil and the Most Stable Chain Fold. Structure. 2: 115-121.
5. Garbuzynskiy SO, Ivankov DN, Bogatyreva NS, Finkelstein AV (2013) Golden Triangle for Folding Rates of Globular Proteins. Proc Natl Acad Sci USA. 110: 147-150.
6. Ptitsyn OB (1973) Stages in the Mechanism of Self-organization of Protein Molecules. Dokl Akad Nauk SSSR 210: 1213-1215.
7. Finkelstein AV (2017) Some additional remarks to the solution of the protein folding puzzle: Reply to comments on "There and back again: Two views on the protein folding puzzle". Phys Life Rev 21: 77-79.
8. Finkelstein AV, Garbuzynskiy SO (2015) Reduction of the Search Space for the Folding of Proteins at the Level of Formation and Assembly of Secondary Structures: A New View on Solution of Levinthal's Paradox. Chem Phys Chem, 16: 3375-3378.