Progress in Biophysics and Molecular Biology xxx (2017) 1–23

Contents lists available at ScienceDirect

Progress in Biophysics and Molecular Biology

journal homepage: www.elsevier.com/locate/pbiomolbio

On a generalized Levinthal's paradox: The role of long- and short range interactions in complex bio-molecular reactions, including protein and DNA folding

Alexey V. Melkikh [a, *], Dirk K.F. Meijer [b]

[a] Ural Federal University, Yekaterinburg, 620002, Mira str. 19, Russia
[b] University of Groningen, The Netherlands

ARTICLE INFO

Article history:
Received 18 April 2017
Received in revised form
27 August 2017
Accepted 17 September 2017
Available online xxx

Keywords:
Levinthal's paradox
Protein folding and misfolding
Molecular docking
NP-Completeness
Long-range interactions
Drug design

ABSTRACT

The current protein folding literature is reviewed. Two main approaches to the problem of folding were selected for this review: geometrical and biophysical. The geometrical approach allows the formulation of topological restrictions on folding, that are usually not taken into account in the construction of physical models. In particular, the topological constraints do not allow the known funnel-like energy landscape modeling, although most common methods of resolving the paradox are based on this method. The very paradox is based on the fact that complex molecules must reach their native conformations (complexes that result from reactions) in an exponentially long time, which clearly contradicts the observed experimental data. In this respect we considered the complexity of the reactions between ligands and proteins. On this general basis, the folding-reaction paradox was reformulated and generalized. We conclude that prospects for solving the paradox should be associated with incorporating a topology aspect in biophysical models of protein folding, through the construction of hybrid models. However, such models should explicitly include long-range force fields and local cell biological conditions, such as structured water complexes and photon/phonon/soliton waves, ordered in discrete frequency bands. In this framework, collective and coherent oscillations in, and between, macromolecules are instrumental in inducing intra- and intercellular resonance, serving as an integral guiding network of life communication: the electrome aspect of the cell. Yet, to identify the actual mechanisms underlying the bonds between molecules (atoms), it will be necessary to perform dedicated experiments to more definitely solve the particular time paradox.

Contents

* Corresponding author.
E-mail addresses: melkikh2008@rambler.ru (A.V. Melkikh), meij6076@planet.nl (D.K.F. Meijer).

https://doi.org/10.1016/j.pbiomolbio.2017.09.018
0079-6107/

1. Introduction

The problem of protein folding is one of the most important problems of molecular biology. A central problem (the so called Levinthal's paradox) is that the protein is first synthesized as a linear molecule that must reach its native conformation in a short time (on the order of seconds or less). The protein can only perform its functions in this (often single) conformation. The problem, however, is that the number of possible conformational states is exponentially large for a long protein molecule. Despite almost 30 years of attempts to resolve this paradox, a solution has not yet been found. A number of authors (see, e.g., Ben-Naim, 2013; Onuchic and Wolynes, 2004; Finkelstein et al., 2017) believe that there is a solution, but they disagree on the reasons. Other scientists (see, e.g., Berger and Leighton, 1998; Davies, 2004) believe that the paradox is not yet resolved.

The issue of folding is typically considered using two fundamentally different approaches that can be called *"biophysical"* and *"geometrical"*. Researchers that use either one of these approaches mostly do not refer to the work based on the other type of approach. The "biophysical" approach uses concepts such as free energy, entropy, and temperature to study protein folding. Simulations of folding are based on statistical physics. The "geometrical" approach does not focus on these parameters; folding is considered geometrically as a part of the broader context of the folding of figures of different topologies. In particular, computational biology has shown that the problem of folding that is based on H-P (hydrophobic-hydrophilic) model belongs to *NP* complexity class (i.e., generally requires an exponential number of steps).

Another major problem, that is essentially ignored in the literature, is the folding and the function of folded DNA because of its much greater length and thus, its much larger number of possible conformational states.

To solve the problem of folding it seems necessary to somehow unite these areas. We must at least discuss the results obtained, using the different approaches and attempt to develop of a single view on the folding problem.

However, the complex interaction of biologically important molecules is connected not only to their folding but also to the possible *reactions* between these molecules. These reactions are in fact the fundamental basis of all the processes that occur in a living system. The scientific field of "molecular docking" can also be regarded as a specialty that is not directly connected to folding processes in life practice. Within this framework, algorithms that can be used to calculate the interaction of the ligand and protein are considered. Complexity also poses a problem to these studies and requires a solution.

In this regard, it seems urgent to develop a more rigorous formulation of the problem of folding and biochemical reactions in general and to discuss the possible solutions in a broader biological context.

2. Levinthal's paradox and its possible solutions (biophysical approach)

The process of protein folding is one of the most important problems of molecular biology. According to the first estimates of Levinthal (1968), the average folding time for a long protein molecule is exponentially large because of the large number of conformational degrees of freedom. Levinthal concluded that a random search can, for this reason, not being performed. In that case, what is the folding mechanism? This problem (the so called Levinthal's paradox) has been considered repeatedly (see, e.g., Anfinsen, 1973, Dill, 1985; Shakhnovich and Gutin, 1989, Zwanzig et al., 1992; Berezovsky and Trifonov, 2002 Trifonov and Berezovsky, 2003; Finkelstein and Ptitsyn, 2002; Bai, 2003, 2006; Grosberg, 2002; Grosberg and Khokhlov, 2010; Finkelstein et al., 2017).

Anfinsen (1973) has proposed the hypothesis that the native protein conformation corresponds to attaining the minimum of its Gibbs energy. Yet, from the point of view of thermodynamics and statistical physics, the problem is to understand how such a complex system reaches equilibrium.

However, despite the large number of publications on this issue, researchers disagree not only about the solution to the paradox but also about whether the problem even exists.

In the following we show the calculations underlying the paradox.

First, let us estimate the chain length for which the enumeration problem does not occur, as previously described (Melkikh, 2015). The total number of states of a protein chain can be estimated as (see, e.g., Berezovsky and Trifonov, 2002)

$$3^N.$$

Here, it was assumed that each domain of a protein has 3 different conformations. If we take the maximal possible populations of such molecules as 1050, then we obtain the following:

$$3^N = 10^{50},$$

Thus, $N \approx 10^2$. Moreover, we can consider the fact that each domain in a protein contains several amino acids to obtain the following rough estimate:

$$N \approx 10^3.$$

Longer chains with more information would not be able to find their native conformation through a random search, at least, during the lifetime of the biosphere. If, however $N << 10^3$, such molecules might find their native conformation by a simple enumeration of variants, and this time must be small (~ 1c) for intracellular processes. As a consequence, N will also be relatively small.

Zwanzig et al., 1992, developed a statistical model of protein

Please cite this article in press as: Melkikh, A.V., Meijer, D.K.F., On a generalized Levinthal's paradox: The role of long- and short range interactions in complex bio-molecular reactions, including protein and DNA folding, Progress in Biophysics and Molecular Biology (2017), https://doi.org/10.1016/j.pbiomolbio.2017.09.018