Sci. Bull. (2015) 60(1):76–85
DOI 10.1007/s11434-014-0718-3

www.scibull.com
www.springer.com/scp

Article

Chemistry

The principle of compromise in competition: exploring stability condition of protein folding

Ji Xu · Mengzhi Han · Ying Ren · Jinghai Li

Received: 16 December 2014 / Accepted: 19 December 2014 / Published online: 8 January 2015

Abstract Thermodynamic hypothesis and kinetic stability are currently used to understand protein folding. The former assumes that free energy minimum is the exclusive dominant mechanism in most cases, while the latter shows that some proteins have even lower free energy in intermediate states and their native states are kinetically trapped in the higher free energy region. This article explores the stability condition of protein structures on the basis of our study of complex chemical systems. We believe that separating one from another is not reasonable since they should be coupled, and protein structures should be dominated by at least two mechanisms resulting in different characteristic states. It is concluded that: (1) Structures of proteins are dynamic, showing multiple characteristic states emerging alternately and each dominated by a respective mechanism. (2) Compromise in competition of multiple dominant mechanisms might be the key to understanding the stability of protein structures. (3) The dynamic process of protein folding should be depicted through the time series of both its energetic and structural changes, which is much meaningful and applicable than the free energy landscape.

Keywords Protein folding · Dynamic structure · Multiple mechanisms · Compromise in competition · Mesoscale · Stability

1 Introduction

Proteins are organic macromolecules consisting of twenty different kinds of amino acids and fold into specific spatial conformations, showing dynamic structures that are important for their biological active functions. Four distinct levels of protein structures have been identified: Primary structure is the sequence of amino acids, secondary structure refers to regular structures formed by local residues, tertiary structure is the three-dimensional (3D) structure of a single protein molecule, and quaternary structure refers to the 3D structure of a multi-subunit protein. Although the structures of the 20 amino acids are already known in detail, and protein sequence can be identified through experimental techniques like mass spectrometry [1], the resolution of dynamic structure remains a challenge because of the influence of environmental factors such as temperature [2], pH [3] and macromolecular concentration [4]. To date, the method used most widely to probe the 3D structure of a protein is to determine its crystal structure through X-ray crystallography and to determine the stable structures for small proteins through NMR, which is, however, unable to show the dynamic change of the protein.

Predicting the dynamic structures of protein using computers is still a challenge with traditional protein folding theories [5]. The dynamic nature and heterogeneity of proteins remain even in the crystalline form because of the high content of solvent in protein crystals [6]. In fact, it is logical that dynamic change is required for proteins to perform their biological functions. It is acknowledged that

Electronic supplementary material The online version of this article (doi:10.1007/s11434-014-0718-3) contains supplementary material, which is available to authorized users.

J. Xu · M. Han · Y. Ren (✉) · J. Li
State Key Laboratory of Multiphase Complex Systems, Institute of Process Engineering, Chinese Academy of Sciences, Beijing 100190, China
e-mail: yren@ipe.ac.cn

M. Han
University of Chinese Academy of Sciences, Beijing 100039, China

proteins exhibit anisotropic motion of individual atoms and collective, large-scale motion over a range of time scales, and that by such dynamic behaviors, proteins can assume a number of almost isoenergetic structures with different conformations [7]. Theoretically, a protein molecule can be considered stationary only when there is no thermal motion of any of the atoms. Therefore, understanding dynamic structures of proteins is a key to many challenging issues in life science. Although some techniques have been developed to provide valuable information about the folding process of proteins, including single-molecule techniques like fluorescence resonance energy transfer [8] and force spectroscopy [9], the laser-induced temperature jump [10] and hydrogen exchange [11] methods, it remains difficult to investigate their dynamic behavior at atomic level. Rapid protein folding events are difficult to capture by current measurement techniques because of their limited spatio-temporal resolution. Therefore, computer simulation, especially molecular dynamics (MD) simulation, has become an important tool to investigate the dynamic structures of proteins at atomic resolution [12].

It has been half a century since the problem of protein folding was first recognized. Great theoretical, experimental and computational efforts have been expended in attempts to understand the underlying mechanism of protein folding, but a definitive answer still remains elusive. Historically, two prevalent theories, the thermodynamic control hypothesis and the kinetic control hypothesis, have been proposed. The thermodynamic hypothesis was first proposed by Anfinsen [13] in 1972, and asserts that a native protein in its normal physiological environment is the system with the lowest Gibbs free energy. The experimental and computational evidence used to develop this hypothesis was the observation that the folding/unfolding of many small proteins are reversible [13, 14]. During the succeeding 40 years, the thermodynamic hypothesis has been developed and led to the concept of a funnel-shaped energy landscape [15], but the funnel hypothesis has yet to be validated experimentally. The thermodynamic hypothesis has been widely used in theoretical predictions of the native structures of proteins [16] and the study of folding dynamics [17]. Theoretically, the thermodynamic hypothesis could be tested by comparing all of the possible states through an exhaustive computer survey of conformational space. However, such an endeavor is unachievable because of the extremely huge amount of conformational space required. On the contrary, an alternative view assumes that the native structure of a protein is kinetically but not thermodynamically stable, and the observed properties should be statistically calculated from a large number of independent folding trajectories [18, 19]. This viewpoint holds that the native states are kinetically trapped in deep valley on the free energy landscape, which guarantees that the native states could exist long enough to perform functions even if they do not correspond to the minimum free energy [18, 20]. For example, the free energies of the native structures of α-lytic protease [21] and serpins [22] are higher than their respective unfolded structures.

Furthermore, protein folding is influenced by environmental factors. The folding of the influenza virus hemagglutinin is induced by low pH [23]. The β-hairpin folding of the β-switch region of glycoprotein Ibα is induced under flow [24]. The folding of some intrinsically disordered proteins [25] is induced by binding with other molecules like ions, small organic molecules and large biomolecules [26]. These findings, in addition to the requirement of dynamic changes for active biological functions, also indicate that the thermodynamic hypothesis is not universal and very limited. In fact, both kinetic and thermodynamic effects, as currently defined, should be simultaneously involved, that is, coupled, because the protein systems are open systems rather than isolated closed systems.

According to the interdisciplinary knowledge in complex chemical systems, in particular, from our understanding of mesoscale phenomena [27, 28], we believe that there might be something missing that is preventing us from understanding the stability condition of proteins. This is because protein molecules show multiscale spatio-temporal structures, and the dynamic changes of protein structures are believed to be a mesoscale phenomenon with dissipative and non-equilibrium features [12, 29]. From the study of complex systems, we believe that a dissipative process is usually dominated by at least two mechanisms [30, 31]. This provides a clue that separating thermodynamic hypothesis from kinetic stability may not be reasonable, and they should be coupled with each other to define stability conditions of protein structures. Protein folding may be dominated by multiple mechanisms, of which free energy minimum is only one of them. Correspondingly, a protein shows multiple characteristic states dynamically, and minimum free energy represents just one of these states. We also think that the concepts of a funnel-shaped energy landscape and the rugged free energy landscape with free energy barrier between native state and intermediate state may not reflect the real change in dynamic process, except the idealized smooth downhill folding landscape with no barriers. The switch from one minimum of a landscape to another may be induced by a totally different mechanism rather than the energy barrier. Surprisingly, this is the same principle as that behind the dynamic structure of gas–solid flow, which is characterized by alternating existence of a gas-rich dilute state dominated by minimization of energy consumption for transporting and suspending particles and a solid-rich dense state dominated by minimization of gravitational potential with respect to time and space [31, 32].