Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Where the complex things are: single molecule and ensemble spectroscopic investigations of protein folding dynamics

Satoshi Takahashi, Kiyoto Kamagata and Hiroyuki Oikawa

Progress in our understanding of the simple folding dynamics of small proteins and the complex dynamics of large proteins is reviewed. Recent characterizations of the folding transition path of small proteins revealed a simple dynamics explainable by the native centric model. In contrast, the accumulated data showed the substates containing residual structures in the unfolded state and partially populated intermediates, causing complexity in the early folding dynamics of small proteins. The size of the unfolded proteins in the absence of denaturants is likely expanded but still controversial. The steady progress in the observation of folding of large proteins has clarified the rapid formation of long-range contacts that seem inconsistent with the native centric model, suggesting that the folding strategy of large proteins is distinct from that of small proteins.

Address
Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai, Miyagi 980-8577, Japan

Corresponding author: Takahashi, Satoshi (st@tagen.tohoku.ac.jp)

Current Opinion in Structural Biology 2016, **36**:1–9

This review comes from a themed issue on **Folding and binding**

Edited by **Ruth Nussinov** and **Jayant B Udgaonkar**

http://dx.doi.org/10.1016/j.sbi.2015.11.006

0959-440/Published by Elsevier Ltd.

Introduction

In the 1970s, Go and his collaborators enthusiastically investigated behaviors of a coarse-grained model of proteins using numerical calculations. When they placed a chain of 49 beads in a two-dimensional lattice and assumed attractive interactions only between pairs of beads forming native contacts, they observed cooperative transitions resembling heat denaturation curves of actual proteins (Figure 1a–c) [1]. The native centric model of proteins thereby originated with Go, but it was met with a lack of comprehension of many researchers in the 70s and 80s, leading him to propose the consistency principle of protein folding [2]. Now, the model is regarded as embodying the ''perfect funnel'' proposed in the energy landscape theory, which subsumes that the primary sequences of natural proteins are optimized for smooth folding transitions [3,4]. The native centric model incorporated into more sophisticated treatment of proteins reproduced not only the cooperative equilibrium transitions but also the tendency of folding rates and the structure of the folding transition state of small single domain proteins [5,6]. The energy landscape theory and the native centric model of proteins are milestones of current structural biology.

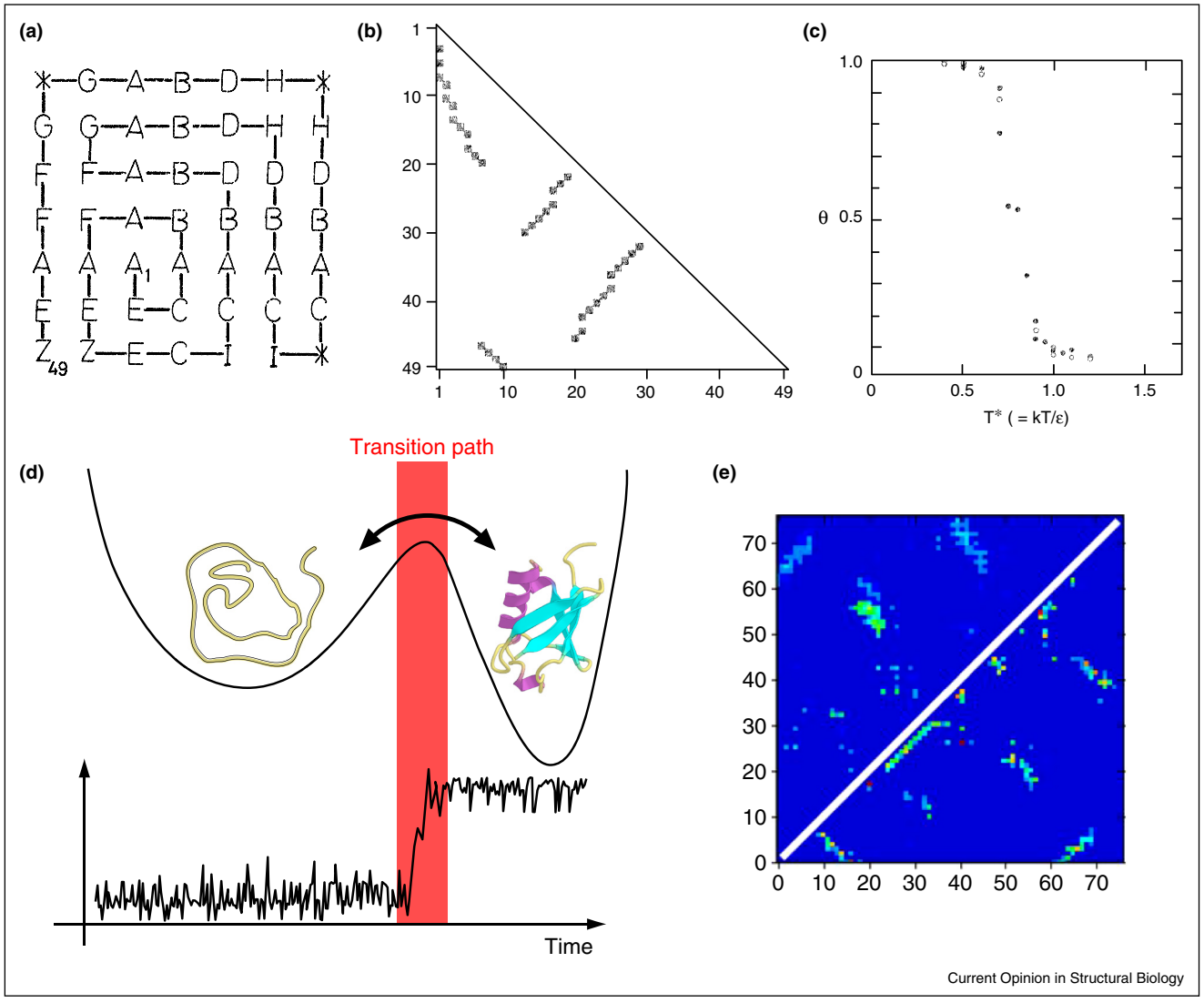
It is noteworthy, however, that important questions related to protein folding have remained unanswered [7,8]. The energy landscape theory and the native centric model are wonderfully simple, but neither includes consideration of details of intramolecular interactions arising from the primary sequence of amino acid residues and their interactions with water. Therefore, they do not explain how the primary sequences of natural proteins destabilize the misfolded conformations and minimize the roughness of the landscape. Partly because of the lack of such explanations, the *de-novo* design of artificial proteins and the prediction of large protein structures persist as difficult tasks. It is necessary to elucidate the mechanism that achieves apparent simplicity in the folding transition of small proteins and to reveal the complex folding of large proteins.

This review presents a summary of the exciting progress and notable discussions related to the simplicity and complexity of protein-folding dynamics that have appeared since 2013. We first introduce results demonstrating applicability of the native centric model for the dynamics of proteins in the transition path. We next introduce results showing structural and dynamical complexity in the unfolded state of even small proteins. Finally, we introduce progress in the investigations of complex folding mechanisms of large proteins.

Native centric model can explain the dynamics of proteins in the transition path

Small proteins with fewer than roughly 100 residues usually demonstrate two-state folding. They provided ample data supporting the energy landscape theory. The recent examination of the folding transition path of small proteins further demonstrated that the theory is almost perfectly satisfied. Chung and Eaton estimated the duration necessary for the folding transitions that are observed as discrete jumps in the single molecule fluorescence data, and termed the duration as the transition

Figure 1



Native centric potential can reproduce various properties of protein folding transitions. **(a)** Two-dimensional protein model originally proposed by Go *et al*. [1]. Each character represents one residue. **(b)** Contact map corresponding to the folded structure in the panel a. **(c)** Temperature denaturation transition calculated for the model protein in panel A assuming the native centric potential. The figures in panels (a)–(c) were adapted with permission from [1]. **(d)** One-dimensional reaction coordinate of protein folding. The two-state folding transition observed for small proteins can be described by one-dimensional potential diagram. The unfolded protein in the left can convert to the native state in one step. Thus, in the single molecule time series measurements, the transition can be detected as the jumps in the fluorescence signals as shown in the putative time series shown in the bottom. Time duration necessary for the folding transition is designated as the transition path time. In the native basin, the folded structure of ubiquitin is presented. **(e)** The contact map in the lower right half represents the native structure of ubiquitin. The upper left half represents the probability of forming inter-residue contacts in the transition path of the all atom MD calculation for folding of ubiquitin. The data were adapted from [15•].

path time (Figure 1d) [9,10••,11•]. The duration was in the range of one to several tens of microseconds, which is consistent with the speed limit of protein folding. All-atom molecular dynamics (MD) calculations of protein folding roughly reproduced the duration [12,13•,14]. The transition path time can be analyzed using Kramer's theory assuming a diffusive barrier crossing over a single reaction coordinate. Furthermore, the native centric potential can explain the behaviors of proteins in the transition path of the all-atom MD calculations (Figure 1e) [15•,16•]. Consequently, the folding transition occurs as a simple diffusive process over a one-dimensional reaction coordinate with a minimum contribution from the non-native contacts.