Biochemical and Biophysical Research Communications 460 (2015) 434–438

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

The elastic free energy of a tandem modular protein under force

Jessica Valle-Orero [a, *], Edward C. Eckels [a], Guillaume Stirnemann [b], Ionel Popa [a], Ronen Berkovich [c], Julio M. Fernandez [a, *]

[a] Department of Biological Sciences, Columbia University, New York, NY 10027, USA
[b] Laboratoire de Biochimie Théorique, Institut de Biologie Physico-Chimique, 13, rue Pierre et Marie Curie, 75005 Paris, France
[c] Department of Chemical Engineering, Ben-Gurion University of the Negev, Beer-Sheva, Israel

ARTICLE INFO

Article history:
Received 4 March 2015
Available online 18 March 2015

Keywords:
Tandem modular protein
Free energy landscape
Force spectroscopy
Tissue elasticity
Protein folding
Langevin dynamics

ABSTRACT

Recent studies have provided a theoretical framework for including entropic elasticity in the free energy landscape of proteins under mechanical force. Accounting for entropic elasticity using polymer physics models has helped explain the hopping behavior seen in single molecule experiments in the low force regime. Here, we expand on the construction of the free energy of a single protein domain under force proposed by Berkovich et al. to provide a free energy landscape for *N* tandem domains along a continuous polypeptide. Calculation of the free energy of individual domains followed by their concatenation provides a continuous free energy landscape whose curvature is dominated by the worm-like chain at forces below 20 pN. We have validated our free energy model using Brownian dynamics and reproduce key features of protein folding. This free energy model can predict the effects of changes in the elastic properties of a multidomain protein as a consequence of biological modifications such as phosphorylation or the formation of disulfide bonds. This work lays the foundations for the modeling of tissue elasticity, which is largely determined by the properties of tandem polyproteins.

1. Introduction

Force spectroscopy studies have uncovered the basic elements of how proteins respond to a stretching force; domains unfold and refold in a time and force dependent manner [1–3]. The Bell model and the formalism of barrier crossing developed by Kramers were widely applied to understand the force dependency of these transitions. However, such models were developed to describe the rupture of bonds over length scales of only a few Angstroms, where changes in entropy do not play a significant role [4]. Using these models to explain protein unfolding–refolding reactions under force, where molecules extend and collapse over tens of nanometers, led to paradoxical results [5]. The departure of experimental findings from simple two-state behavior motivated the development of new theories that considered changes in entropy as a crucial component of the free energy [5,6]. These models included the laws of polymer physics in the unfolding of a single protein domain under a stretching force and demonstrated the force dependency of the unfolding rates and protein elongation. However, it was not clear how to extend these concepts to construct the elastic free energy of a multi-domain protein where many of these individual modules are arranged in tandem.

Models of tandem modular proteins are becoming increasingly important as these proteins are identified as determinant factors of tissue elasticity and modulators of cell signaling [7–9]. It remains poorly understood how the mechanical behavior of single molecules scales to an overall material property such as tissue elasticity. Here, we generalize the concepts described by Berkovich et al. [5] and show how to construct the elastic free energy of a tandem modular protein as a function of force. We perform Brownian dynamics simulations of the resulting free energy landscape and reproduce key experimental benchmarks of protein unfolding such as the Arrhenius dependency of the rates, the force dependency of the step size, and the force range at which domain refolding is favored over unfolding. Our model predicts that the effects of complex biological modifications, such as disulfide bond formation, can be incorporated into the free energy by simply changing the polymer properties. Thus, describing the elastic free energy of large tandem modular proteins is an important step towards understanding the origins and regulation of tissue elasticity and its function during animal motion [10–12].

* Corresponding authors.
E-mail addresses: jv2499@columbia.edu (J. Valle-Orero), jfernandez@columbia.edu (J.M. Fernandez).

http://dx.doi.org/10.1016/j.bbrc.2015.03.051
0006-291X/

2. Materials and methods

A free energy model for tandem modular proteins is developed here by extrapolating the procedures described in Berkovich et al. [5]. This model was defined for only one domain, but here we propose a numerical method to concatenate the free energy landscape to any number of identical protein domains *N*.

The free energy of a polyprotein can be described by the summation of three distinct components: the entropic elasticity of the polymer chain under force U^{WLC} (Eq. (S1)), a short range potential representing the hydrophobic interactions that drive folding U^{M} (Eq. (S2)), and an entropic barrier caused by removal of available polypeptide configurations between the collapsed and folded states U^{G} (Eq. (S3)). Here the worm-like chain (WLC) model is used to approximate the entropic elasticity of the polypeptide [6,13]. The entropic elasticity is controlled by two parameters: the increase in contour length ΔL_c and persistence length p. The free energy of a single domain along its pulling coordinate is defined by the summation of these three energy contributions in Eq. (1). A plot of the free energy of a single protein domain constructed from these three components (dashed lines) is shown in Fig. 1A (solid line).

$$U(x) = U^{WLC}(x, \Delta L_c) + U^{M}(x + R_C) + U^{G}(x) \tag{1}$$

The energy landscape is defined within the range $x_0 \leq x \leq x_1$. At short extensions, the free energy is dominated by the Morse potential and the protein lies in its native state at x_0. With application of enough force, mechanical unfolding drives the protein to an extended conformation with an energy located at the entropic minimum x_1 according to Eq. (S1).

In order to generalize this result for tandem modular proteins, we consider the effect of unfolding a second domain in the protein. At constant force *F*, unfolding a second domain increases the total length of the polymer by ΔL_c such that the total contour length of the polymer is $2 \cdot \Delta L_c$. The entropic minimum of this polymer chain is now located at a new extension x_2. Thus, extending a polymer with N unfolded domains is well described by Eq. (S1) with a contour length of $n \cdot \Delta L_c$ where $1 \leq n \leq N$, and an entropic minimum located at extension x_n (pink curves; Fig. 1B).

Fig. 1B is a graphical representation of how to construct the free energy landscape of a tandem polyprotein from the segments comprising the free energy of each individual domain. Prior to unfolding, the second domain lies in the minimum of a Morse potential located at x_1 and must cross a transition state barrier at $x_1 + x_b$. After unfolding, the second domain lies at its entropic minimum x_2. Thus, we sum the three energetic components as described in Eq. (1) over the range $[x_1, x_2]$ for a polymer with contour length $2 \cdot \Delta L_c$. For a protein with *N* folded domains, the free energy is divided into several segments, defined by the evaluation of Eq. (1) on $[x_{n-1}, x_n]$ for $n = 1, 2, \ldots N$ (blue curves, Fig. 1B). A general expression for the free energy of any segment *n* is provided here:

$$\begin{aligned} U_n(x) = & U^{WLC}(x, n\Delta L_c) + U^{M}(x - x_{n-1} + R_C) \\ & + U^{G}(x - x_{n-1}), \quad x_{n-1} \leq x \leq x_n \end{aligned} \tag{2}$$

Furthermore, we assume that the free energy must be continuous at the boundaries of each segment, so that the entropic minimum of the extended domain coincides with the minimum Morse energy of the subsequent folded domain. This is numerically achieved by concatenating all the segments to satisfy the following boundary condition (black curve; Fig. 1B).

$$U_{n-1}(x_{n-1}) = U_n(x_{n-1}), \quad 1 \leq n \leq N \tag{3}$$

As shown in Fig. 1B, a polypeptide with *N* structured domains (black curve) has a different energy compared to an unstructured polypeptide with an equivalent contour length (pink curve, $n = N$) due to Eq. (3) [14]. Finally, we append to the free energy a stiff segment representing *N* tandem folded domains plus any polymer linkers used for attachment chemistry to the probe (red curve, Fig. 1A and B). The free energy for this segment is calculated using only the WLC (Eq. (S1)) with a high persistence length (Table S1) on the range $[0, x_0]$.

The resulting continuous free energy landscape for a tandem modular protein with $N = 8$ domains is shown in Fig. 2. The free energy was calculated at several forces: 4, 7, 12, 15 and 18 pN. As the force applied to the polypeptide increases, the location of the entropic minima x_n increases according to the WLC model (Fig. 2A, dashed lines). The trajectories of the minima as a function of force will be referred to here as E-curves. For a protein with $N = 8$ domains, there are $N + 1$ E-curves representing the entropic minima

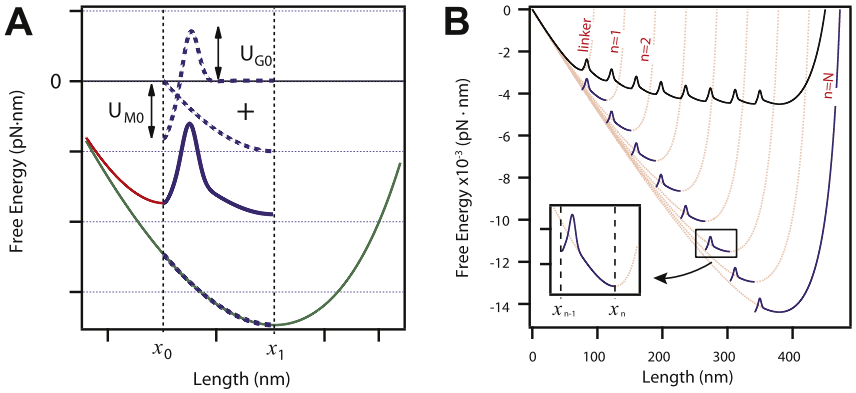

Fig. 1. Method of the construction of the elastic free energy of a tandem modular protein. A) Construction of the free energy for a single domain unfolding and extending. The free energy of unfolding and extending a domain is constructed from three elements (dotted lines): a Morse well of U_{M0} depth, a Gaussian barrier of U_{G0} height, and a WLC potential of contour length ΔL_c (green line) from which only the section between x_0 and x_1 is considered. The red curve represents the initial extension to x_0 calculated from the WLC with a contour length L_0. B) Expansion of the free energy model to a polyprotein with *N* number of domains. The pink lines correspond to WLC curves calculated for an octamer polyprotein at a force $F = 20$ pN, and for contour length $n \cdot \Delta L_c$, where *n* is the number of unfolded domains, $\Delta L_c = 19$ nm, and $p = 0.4$ nm. At each contour length the WLC has a minimum that serves as the obligatory starting point for the next segment. Segments (thick blue lines) are constructed by adding a Morse well and a Gaussian barrier to the entropic elasticity as described in the text and in Fig. 1A within the range x_{n-1} to x_n (insert). The WLC for the linker is also shown in the plot ($L_0 = 42$ nm and $p_f = 10$ nm, extended at $F = 20$ pN). The final free energy profile is constructed by concatenating all segments at their boundaries (thick black line). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)