Biophysical Chemistry xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Biophysical Chemistry

journal homepage: http://www.elsevier.com/locate/biophyschem

The push-and-pull hypothesis in protein unfolding, misfolding and aggregation

Guilherme A.P. de Oliveira *, Jerson L. Silva *

Programa de Biologia Estrutural, Instituto de Bioquímica Médica Leopoldo de Meis, Instituto Nacional de Biologia Estrutural e Bioimagem, Centro Nacional de Ressonância Magnética Nuclear Jiri Jonas, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

ARTICLE INFO

Article history:
Received 14 January 2017
Received in revised form 25 March 2017
Accepted 27 March 2017
Available online xxxx

Keywords:
Protein misfolding
High-pressure spectroscopy
Molten globule
Urea
High hydrostatic pressure
Alpha-synuclein
p53

ABSTRACT

The combination of biophysical and structural techniques has allowed the visualization of species classified as dry molten-globule states. Further destabilization causes these structures to follow through a wet-globule stage to reach an unfolded chain. We have recently combined small angle X-ray scattering and nuclear magnetic resonance to observe these species, and we introduce a push-and-pull hypothesis to explain the dissimilar actions of urea and high pressure on proteins. The implications of these molten-globule states are further discussed in light of their potential physiological and pathological roles, especially in protein misfolding diseases.

1. Introduction

The trajectories that are adopted by polypeptide chains during folding reactions remain a notorious issue in the literature, and the difficult task of understanding protein folding phenomena and characterizing intermediates has puzzled researchers. From Anfinsen's dogma and the Levinthal's paradox to conformational diversity and functional promiscuity, science has proven its transformative capacity. The seminal work of Anfinsen on ribonuclease A folding revealed that amino acid composition is the determining factor for a small protein to achieve its native structure, and the native structure is a unique and stable conformation with a minimum free energy [1,2]. Cirus Levinthal posited that a tremendous amount of random searching would be required for a stretched polypeptide chain to find its correct native conformation, but considering a flat energy landscape there is an apparent paradox in which proteins require a fast timeframe to fold, in the range of micro to mili seconds. His conclusion was that there is a pathway for folding and that a well-defined sequence of interaction events might occur for proteins to fold properly and efficaciously [3]. However, a mathematical analysis imposing restraints at the order of thermal energy had locally unfavorable bond configurations and revealed a reduced time for proper folding that fits a realistic biological timeframe [4].

Intermediate conformers undoubtedly occur during the folding reactions of small and large proteins, but our current ability to characterize and interpret these species is limited by the timeframe under which human technologies are sensitive enough to detect them. High pressure is an important approach to help us visualize the plurality of intermediate states upon unfolding. This became evident in a 1976 work by Weber's group on the pressure-induced unfolding of lysozyme and chymotrypsinogen [5]. Additionally, the low energy barriers between the intermediate states (*e.g.*, $N \leftrightarrow I^1 \leftrightarrow I^2 \leftrightarrow I^n \leftrightarrow U$, where N is the native conformation; I^1, I^2 and I^n are multiple intermediate species; and U is the unfolded peptide), due to the high cooperativity among their rates of inter-conversion, provide a rate-limiting step to their visualization. The two-state mechanism of folding ($N \leftrightarrow U$, without populating appreciable intermediates) is the simplest and most attractive model to explain the folding reaction of small proteins (approximately 100 amino acids). In these circumstances, a nucleation-condensation mechanism would select distant and adjacent sites on polypeptide primary sequences to work as determinants for stability and should govern the rapid collapse of the native structure [6]. Here, the collapse is related to the idea that hydrophobic side-chains, when in contact with an aqueous solution, will have the tendency to move inward in the protein manifold to shield these groups from the solvent. Proteins with larger and more complex polypeptide chains *e.g.*, multi-domain proteins,

* Corresponding authors at: Programa de Biologia Estrutural, Instituto de Bioquímica Médica Leopoldo de Meis, Instituto Nacional de Biologia Estrutural e Bioimagem, Centro Nacional de Ressonância Magnética Nuclear Jiri Jonas, Universidade Federal do Rio de Janeiro, 21941-902 Rio de Janeiro, RJ, Brazil.
E-mail addresses: gaugusto@bioqmed.ufrj.br (G.A.P. de Oliveira), jerson@bioqmed.ufrj.br (J.L. Silva).

http://dx.doi.org/10.1016/j.bpc.2017.03.007
0301-4622/

would fit timeframes at which we can measure these intermediate conformers by different biophysical techniques. The presence of intermediate conformers in small proteins, in contrast, has already been detected with several different biochemical and biophysical efforts, such as urea-gradient electrophoresis near 0 °C, folding kinetics using amide hydrogen exchange and fast circular dichroism measurements [7 and refs therein]. Notably, a combination of computational and experimental restraints was able to represent an ensemble of structures in the unfolded state of a *Drosophila* SH3 domain that preserved a reasonable amount of native-like contacts and a compact state [8].

The classical views of substrate binding to enzymes, *i.e.*, the lock-and-key and induced-fit models, support the idea that rigid structures exist for binding or that small local changes around the active site occur to accommodate the substrate. These seminal models are in opposition to the most recent visualization of functionality, which indicates the presence of higher energetic states that form the basis for the conformational diversity of proteins. In this new model, a pre-existing ensemble of conformers under equilibrium with small energy barriers populates the native structure; upon ligand binding, a shift to the high-affinity conformer is favored. The diversity of conformational motions ranges from side-chain fluctuations, loop and secondary structure motions and even tertiary rearrangement of domains in complex multi-domain proteins. There are several examples of conformational diversity in proteins, such as the inter-conversion of cellular prion (PrP^{C}) to Scrapie prion (PrP^{Sc}), *i.e.*, all α-helix to all β-sheet [9], and the monoclonal antibody SPE7 that samples different binding-site conformations and thus binds to unrelated antigens [10]. Additionally, the elements that contribute most to conformational diversity are intrinsic disordered proteins like the alpha-synuclein involved in synaptic vesicle release and trafficking, modification of neurotransmitter release, neuronal survival and plasticity. The fact that a protein is able to sample different conformations suggests that distinct functionalities could arise from the same polypeptide chain; therefore, the paradigm that one function comes from one structure fails. Indeed, antibody multi-specificity, in which an ensemble of conformers exists with similar free energies, was proposed by Linus Pauling in the 1930s [11]. Recently, we and others characterized subtle changes in the tumor suppressor transcription factor p53 due to cancer-related mutations, which convert the protein to misfolded conformers, such as molten-like or native-like structures, that culminate in aggregation with severe consequences for cancer [12–18]. These examples of conformational diversity have intriguing implications for understanding the phenomena of protein folding and the investigation of hidden intermediates in physiological and pathological states.

Here we present a brief discussion of the formation of dry and wet molten-globule intermediates and some molecular features of these intermediates that are marginally observable with current techniques. If the universality of these dry/wet molten-globule states is proven for different classes of protein structures, it may represent a broader comprehension of the protein folding phenomena that could shed light on the different states proteins can adopt during pathological conversion or their physiological roles. We connect this discussion to our recent push-and-pull hypothesis that explains the dissimilar actions of urea and high hydrostatic pressure to populate these challenging species.

2. The basics of the push-and-pull hypothesis

The basic idea of the push-and-pull model comes from the dissimilar actions of chemical and physical forces in unfolding biological systems. The name "push-and-pull" comes from the contrast between preferential hydration under high hydrostatic pressure and the preferential binding of chemical denaturants, such as urea or guanidinium. We have recently provided experimental evidence combining local (nuclear magnetic resonance spectroscopy) and global techniques (small angle X-ray scattering and fluorescence) to assess the push-and-pull effects during protein unfolding [19].

In the case of pressure, the mechanism affecting protein structure was heavily explored and uncovered by several groups [19–25] and recently revised [26]. The basis for the idea is that pressure triggers the hydration of nonpolar amino acid residues located in the interior of the protein manifold that are normally oriented to the bulk of non-exposed cavities formed due to packing defects during folding. Computational studies in proteins and compound models support the solvation of nonpolar groups [27,28]. Volume manipulation of internal cavities by direct mutagenesis indicated the relevance of these "empty regions" to pressure-induced unfolding [23]. In light of our push-and-pull model and in contrast to the action of denaturants, pressure affects proteins in a non-homogeneous way, leading to the formation of cooperative units, *i.e.*, states that preserve a few secondary characteristics but lack major tertiary contacts. Our data revealed that pressure disturbs the protein structure due to its compressibility effects, culminating in the shortening of hydrogen bonds and hydration of hydrophobic motifs with the formation of these cooperative units (push effect) [19]. In contrast, the direct mechanism of urea binding to the protein surface seems to govern hydrogen bond shifts and the expulsion of water molecules from the solvation shell, causing instability in the overall packing that is maintained up to a limit concentration of denaturant (pull effects) [19].

The molecular mechanisms by which urea and pressure lead to protein unfolding have been broadly explored in the literature [19 and refs therein]. In the case of urea, there are two proposed mechanisms. The first hypothesizes an indirect mechanism where the denaturant changes the water structure, leading to the solvation of hydrophobic groups, and the second relates to a direct interaction of urea with the protein surface, withdrawing water molecules from around its surface. We recently showed that the direct mechanism is the driving force for urea-induced unfolding [19]. Several computational studies using all-atom replica exchange MD simulations have supported this direct mechanism [29–30]. The water expulsion effect due to denaturant binding at low concentrations seems to generate a swollen protein architecture (higher ΔV_u values as measured by fluorescence spectroscopy) that likely abolishes crucial protein-solvent and intramolecular interactions, resulting in protein destabilization without affecting the conformational packing. This swollen and more fragile state that exists at low concentrations of denaturant becomes evident when low levels of denaturant are combined with high-pressure titration (Figs. 1 and 2). Thermodynamic parameters including the $p_{1/2}$ and the ΔV_u values were measured to provide information about these states. The $p_{1/2}$ values were calculated from the fits obtained by the fluorescence data (Fig. 2a) and express the pressure at the midpoint of the curves. In contrast, ΔV_u values represent the volume changes from the native to the unfolded state and were calculated globally from the fluorescence fits (Fig. 2a) and locally from the NMR intensities of the H-N fits (Fig. 2b) at each urea concentration (Fig. 2c). We observed an appreciable reduction in $p_{1/2}$ values upon increasing denaturant, as measured by fluorescence spectroscopy (Fig. 2a) and an increase in local ΔV_u values with a decrease in the heterogeneity of the pressure effects on ΔV_u [19] (Fig. 2b and c). These swollen, highly unstable conformers promoted by low doses of denaturant are consistent with a dry globule state in which a few protein-solvent and intramolecular tertiary contacts were lost, but not in a sufficient way to disrupt its tertiary architecture packing. Upon increasing the denaturant to a limiting concentration, or when pressure is applied, most of the interactions are lost in a manner sufficient to disrupt the tertiary fold and promote wet molten-globule states, *i.e.*, states with native-like secondary structures but with disrupted native folding patterns due to the liquid-like interactions between nonpolar groups.

3. The challenging species between the N ↔ U

Groundbreaking evidence for these challenging intermediate species dates back the 1970s, when perturbing agents, such as acid or temperature, showed different cooperative transitions than those induced by urea or guanidinium [31]. The molten-globule (MG) term was coined

Please cite this article as: G.A.P. de Oliveira, J.L. Silva, The push-and-pull hypothesis in protein unfolding, misfolding and aggregation, Biophys. Chem. (2017), http://dx.doi.org/10.1016/j.bpc.2017.03.007