Biophysical *Journal*
New and Notable

Biophysical Society

Detailing Protein Landscapes under Pressure

Rocío Espada,[1] Ignacio E. Sánchez,[1] and Diego U. Ferreiro[1,*]
[1]Protein Physiology Lab, Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, UBA-CONICET-IQUIBICEN, Buenos Aires, Argentina

Natural protein molecules are remarkable physical objects. Despite the astronomical number of competing structural forms, these systems self-organize into beautiful structural ensembles in biologically short timescales, puzzling out the feat of specifically bringing together thousands of atoms interacting by a myriad of weak forces. Moreover, most protein domains appear to fold with ease in a deceptively simple two-state manner, populating either the fully folded or fully unfolded ensembles. This is certainly not a property of random amino acid chains…how do proteins do it?

Today, protein folding is conceptually understood within the framework of the energy landscape theory, which describes the diversity of structural possibilities in statistical mechanical terms, setting the groundwork for the interpretation and design of novel experiments that detail our understanding of the phenomenon (1). A new and notable example of such interaction is provided by Fossat et al. (2) in this issue of the *Biophysical Journal*.

To map the energy landscape, Fossat et al. (2) brought together three key experimental aspects: the use of hydrostatic pressure to favor the population of folding intermediates, the power of nuclear magnetic resonance (NMR) to detect the signals with residue-specific resolution, and the use of a model protein that has high internal symmetry. Proteins that are composed of tandem arrangements of repeating motifs are expected to navigate landscapes where structure consolidates in somewhat discrete folding steps (3). Once an initial nucleation takes place, folding proceeds via a propagating structure to neighboring units, much like a front of droplets in a phase transition (4). The local stability of the units and their interactions define the preferred routes, as well as the appearance of partially folded species along the trajectories (5). Then, it is simple: just get the local energies and compute the population of states. Well, maybe not that simple.

The folded leucine-rich repeat domain of the pp32 protein consists of five leucine-rich repeats capped on both the N- and C-termini by short helices. Chemical unfolding of pp32 can be well described by a two-state transition with high cooperativity, but kinetic experiments revealed the transient population of an intermediate structured around the C-cap and repeat-5 (6). Although it has five repeating motifs, three folding units were postulated (6). As for other repeat-protein architectures, the structural repeats do not necessarily correspond with the cooperative units, a fact often overlooked by typical sequence-based analysis (7). When monitored with single-residue resolution by NMR and perturbing the structure with pressure (8), each individual signal can also be portrayed as populating only two states, and an apparent stability (ΔG^0_i) and cooperativity (ΔV_i) can describe each transition. The main new finding is that these parameters do not coincide for every residue in pp32, but show a very large spread (2). How can the residues that constitute the protein appear to occupy a multiplicity of states while the whole protein only two? Maybe the reconciliation lies in recognizing what each residue signal is reporting.

Hydrostatic pressure denatures proteins because the folded ensemble has a higher partial molar volume than the unfolded ensemble (8). Picture that surface corrugation is as extensive as cavity formation, meaning that folded proteins are not true three-dimensional objects but rather a "Swiss cheese" of atoms (9). As frustration from imperfect packing builds up with molecular size, it allows larger voids to form in larger molecules (9). Because these packing imperfections disappear when water molecules solvate the polypeptide chain, pressure stabilizes unfolded conformations. This is illustrated in Fig. 1. The free energy difference between the states is proportional to the applied pressure. If the whole protein occupies mainly the fully folded and fully unfolded states, then every single residue must report the same thermodynamic transition (U-F), and the ΔG^0_i and ΔV_i values coincide

Submitted September 8, 2016, and accepted for publication October 27, 2016.

*Correspondence: ferreiro@qb.fcen.uba.ar

Editor: James Shorter.
http://dx.doi.org/10.1016/j.bpj.2016.10.038

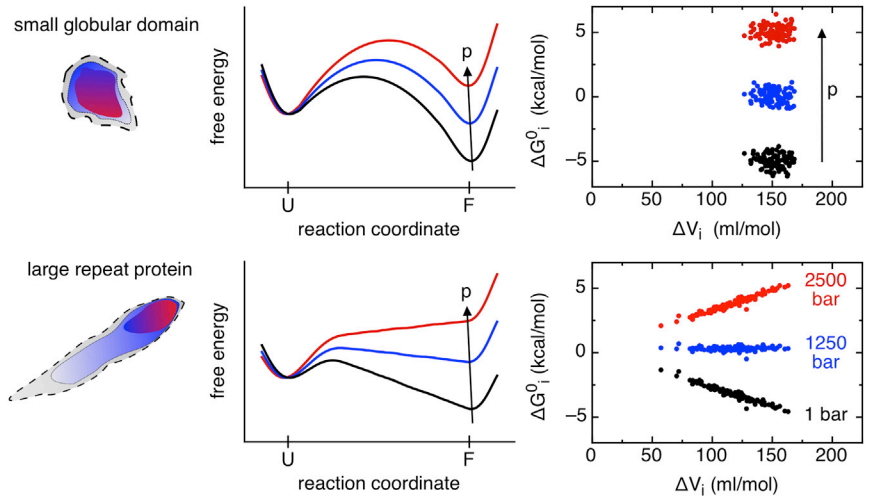

FIGURE 1 Putting pressure on protein folding landscapes. (*Left*) Sketches of proteins folding at different pressures, where different phases of the protein are possible: folded core (*red*), interfacial region (*blue*), and unfolded halo (*gray*). Having a gradual progression of phases can lower the activation energy and broaden the separation between folded and unfolded regions. (*Top*) Free energy diagram of a hypothetical small two-state folding protein. A large barrier separates the fully folded (F) and unfolded (U) states; folding appears highly cooperative at low (*black line*), intermediate (*blue line*), and high pressure (*red line*). (*Right*) Hypothetical experimental results as mapped by each residue in high-resolution NMR at low (*black points*), intermediate (*blue points*), and high pressure (*red points*). (*Bottom*) Hypothetical free energy diagram of a large repeat-protein at low (*black line*), intermediate (*blue line*), and high pressure (*red line*). (*Right*) Experimental results for pp32 mapped at single-residue resolution extrapolated at low (*black points*), intermediate (*blue points*), and high pressure (*red points*). The experimental data for the ΔG^0_i and ΔV_i values of pp32 were extracted from Table S1 of Fossat el al. (2). To see this figure in color, go online.

for each residue signal. At different pressures, the populations of states change, but ΔV remains the same. Analogously, if the protein occupies an intermediate state (I) for which a portion of the chain remains in its native structure (F), each residue reports the transition in which it is involved, either U-I or I-F. Two clusters of signals should appear in the ΔG^0_i-versus-ΔV_i plot. This is not what is observed for pp32.

Putting pressure on pp32 reveals that each residue reports a transition for which ΔG^0_i and ΔV_i are coupled. These signals spread linearly on the ΔG^0_i-versus-ΔV_i plane (*lower-right panel* of Fig. 1), as predicted by the capillarity approximation of protein folding in funneled landscapes (4). The slope denotes the bulk free energy difference per unit volume, suggesting that each of the residues' signals come from the population of a variety of structurally related intermediates. These intermediates must be asymmetric, and appear as an almost continuous consolidation of structure around a common nucleus. Fossat et al. (2) go a step further and model the structural ensembles compatible with the NMR signals. They find that the folding of pp32 is indeed asymmetric with structure consolidating from C- to N-termini, in accordance with the local stabilities previously mapped (6). The main folding barrier is low, located close to the unfolded state, and corresponds to the formation of a ~40-residue nuclei, about the size needed for the energetic compensation of the entropy cost mapped in other repeat-protein systems (4). After the nucleation is established, folding proceeds downhill in free energy, and pressure modulates the steepness of the gradient (Fig. 1). The fact that a multiplicity of transitions is reported by the single residues is indicative that the landscape is effectively smooth, with low energetic roughness between the intermediates. This is reminiscent of the continuum of states that have been mapped by H/D exchange and chemical denaturation of globular proteins (10). The models constructed by Fossat et al. (2) confirm this expectation and anticipate that the roughness is in the order of kT.

The high-resolution mapping of the folding of pp32 qualitatively confirms the expectations of the energy landscape description, and calls for a quantitative dissection of the underlying factors governing the transitions. Is it simple or not? Energy landscape theory recognizes that natural protein molecules are fundamentally evolved polymers that must simultaneously satisfy a diversity of constraints such as folding, localization, interactions, and diverse functions. These biological requirements often come into conflict with physical necessities, thus the details of the landscapes of individual proteins are tuned by their evolutionary histories. Subtle energetic changes brought about by sequence or environmental modifications modulate the stabilization of structural states, and these can be coopted by biological functions (11). It is now clear that the population of these excited states in repeat proteins is intrinsically linked to their physiological behaviors, such that the collective influence of local interactions allows sites to thermodynamically modulate each other even at a considerable distance. The mechanistic complexity that remains on simple funneled landscapes may be ripe to be uncovered from first principles.

ACKNOWLEDGMENTS

R.E., I.E.S., and D.U.F. are supported by Universidad de Buenos Aires-Consejo Nacional de Investigaciones Científicas y Técnicas and Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT grant Nos. PICT2012-2550 and PICT2012-01647).

REFERENCES

1. Wolynes, P. G., W. A. Eaton, and A. R. Fersht. 2012. Chemical physics of

protein folding. *Proc. Natl. Acad. Sci. USA.* 109:17770–17771.

2. Fossat, M. J., T. P. Dao, …, C. A. Royer. 2016. High-resolution mapping of a repeat protein folding free energy landscape. *Biophys. J.* 111:2368–2376.

3. Ferreiro, D. U., A. M. Walczak, …, P. G. Wolynes. 2008. The energy landscapes of repeat-containing proteins: topology, cooperativity, and the folding funnels of one-dimensional architectures. *PLOS Comput. Biol.* 4:e1000070.

4. Ferreiro, D. U., and P. G. Wolynes. 2008. The capillarity picture and the kinetics of one-dimensional protein folding. *Proc. Natl. Acad. Sci. USA.* 105:9853–9854.

5. Gruebele, M. 2014. CARPe diem. *Biophys. J.* 107:3–4.

6. Dao, T. P., A. Majumdar, and D. Barrick. 2015. Highly polarized C-terminal transition state of the leucine-rich repeat domain of PP32 is governed by local stability. *Proc. Natl. Acad. Sci. USA.* 112:E2298–E2306.

7. Espada, R., R. G. Parra, …, D. U. Ferreiro. 2015. Repeat proteins challenge the concept of structural domains. *Biochem. Soc. Trans.* 43:844–849.

8. Roche, J., M. Dellarole, …, C. Roumestand. 2015. Exploring the protein folding pathway with high-pressure NMR: steady-state and kinetics studies. *In* High Pressure Bioscience. Springer, Dordrecht, Netherlands, pp. 261–278.

9. Chowdary, P. D., and M. Gruebele. 2009. Molecules: what kind of a bag of atoms? *J. Phys. Chem. A.* 113:13139–13143.

10. Parker, M. J., and S. Marqusee. 2000. A statistical appraisal of native state hydrogen exchange data: evidence for a burst phase continuum? *J. Mol. Biol.* 300:1361–1375.

11. Kauffman, S. A. 2014. Prolegomenon to patterns in evolution. *Biosystems.* 123:3–8.