Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Editorial overview: Folding and binding: Dynamic conformational heterogeneity is pivotal to cell life

Ruth Nussinov and Jayant B Udgaonkar

Current Opinion in Structural Biology 2016, **36**:iv–vi

For a complete overview see the Issue

Available online 9th February 2016

http://dx.doi.org/10.1016/j.sbi.2016.01.012

Ruth Nussinov[1,2]

[1] Cancer and Inflammation Program, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, National Cancer Institute, Frederick, MD 21702, United States

[2] Sackler Inst. of Molecular Medicine, Department of Human Genetics and Molecular Medicine, Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

e-mail: NussinoR@helix.nih.gov

Ruth Nussinov received her B. Sc. from the University of Washington, and Ph.D. in 1977 from Rutgers University. She did post-doctoral work in the Structural Chemistry Department of the Weizmann Institute of Science. Subsequently she was at the Chemistry Department at Berkeley, the Biochemistry Department at Harvard, and a visiting scientist at the NIH. In 1984 she joined the Department of Human Genetics, at the Medical School at Tel Aviv University. In 1985, she accepted a concurrent position at the National Cancer Institute of the NIH, Leidos Biomedical Research, where she is a Senior Principal Scientist and Senior Principle Investigator heading the Computational Structural Biology Section at the NCI. She has authored over 500 scientific papers. She is the Editor-in-Chief in PLoS Computational Biology and Associate Editor and on the Editorial Boards of a number of journals. She is a frequent speaker in domestic and international meetings, symposia and academic institutions, won several award and elected fellow of several societies. Her National Cancer Institute website gives further details. https://ccr.cancer.gov/ruth-nussinov.

To function, a macromolecule needs to be in a distinct conformational state and to interact with its partner(s). Macromolecular folding and binding are decisive events in the life of a cell; aberrancy in either process can lead to dysfunction and occasionally to cell death. For decades, efforts have centered on the so-called protein folding problem. Even though understanding deepened and much progress has been made, the problem of ab initio folding still remains largely intractable. Important unsolved mechanistic problems include how the sequence of the protein chain dictates the sequence of structural events during folding; how the sequence and folding conditions modulate the cooperativity of the folding reaction; and how initial non-specific chain collapse facilitates subsequent native structure formation. These problems persist because of the continued lack of understanding of the chemical forces that govern folding. Major obstacles include the insufficiently accurate force fields, the incompletely understood folding pathways and the large conformational space that needs to be searched. This has led to the community expectation that eventually the problem will be solved by knowledge-based approaches: with a sufficiently large dataset of experimentally-determined structures of representatives of the diverse protein folds, homology modeling will overcome the challenge. Notwithstanding, how proteins fold in the living cell — rather than in solution or in silico — is still among the most profound questions. In vivo, conditions and processes differ, including effects of co-translational folding, the impact of cell crowding, the assistance lent by molecular chaperones, the stabilization/destabilization by post-translational modifications, small molecules and the variable environment. The disordered protein state further compounds the challenge, raising questions about its regulatory mechanisms and how it fulfills its specific roles in the various systems; its order-disorder transitions upon functional changes in the cellular environment as well as post-translational modifications; molecular recognition and more. Finally, even though ab initio folding is the holy grail of structural biology, relating the structure to function also presents a daunting challenge. This is where folding meets binding. In the living cell, no macromolecule works on its own; binding is essential for the protein to execute virtually all its functions.

The structure–function paradigm which dominates molecular biology was inspired by the notion that even living things must conform to the laws of quantum mechanics and physical chemistry. The powerful idea of the energy landscape — which views molecules as dynamical ensembles of conformational states — was able to explain how function is executed, thereby deepening and shedding light on this paradigm. This profound concept led to the realization that biomolecules are not static sculptures;

Jayant B Udgaonkar
National Centre for Biological Sciences, Tata Institute of Fundamental Research, GKVK Campus, Bengaluru 560065, India
e-mail: jayant@ncbs.res.in

Jayant Udgaonkar obtained his BSc degree from the University of Mumbai and his MSc degree from the Indian Institute of technology, Madras, majoring in chemistry. Subsequently, he obtained his PhD degree in biochemistry with a minor in applied physics from Cornell University, Ithaca in 1986, carrying out chemical kinetic and patch-clamp studies on the acetylcholine receptor. He then carried out postdoctoral research in protein folding at Stanford University, Stanford, where he developed the HX-NMR methodology. Since 1990, he has been at the National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bengaluru, where is now a Senior Professor. He is on the editorial board of Protein Engineering, Design and Selection, and is an Associate Editor of Biochemistry.

instead, they are dynamical objects that continuously interconvert between different structures with varying energies, and that this interconversion between distinct conformational states is how they fulfill their tasks in cell life. This indicated that biomolecules should not be described in terms of their static structures but their dynamical statistics, which changes with the physical environment and cell state. The number of possible conformations that biological macromolecules can adopt for life is immense, and evolution has exploited and tuned them to optimize existing functions and gain new ones. This understanding has led to key questions including how one-dimensionally connected biomolecules can organize themselves into functional three-dimensional states, and how a single protein chain can assume multiple functions, each of which requiring distinct conformation, primarily through multiple distinct interactions. Distinct states are the way through which regulation is executed. Such questions are particularly challenging since the strong bonding constraints are relatively few in number. Structural biologists aim to comprehend how specific binding events take place at specific times and spatial locations in the cell; which events determine these and how; and how binding events of distinct protein conformations lead to distinct functions. They recognize that to function, biomolecules must remain incompletely organized.

The free energy landscape theory has had a profound impact on the field of protein folding. It also encapsulates the essence of binding and function. The importance of perceiving function in terms of conformational heterogeneity has increasingly shifted the interest in the community from folding to function. From the standpoint of the free energy landscape, the principles are unaltered: rather than considering the entire protein conformational landscape — which is the case in ab initio folding — the focus has shifted to the ensemble around the bottom of the folding funnel. For enzymes, the protein can be viewed as populating a single active or an ensemble of inactive and active states. The basins of the active and inactive states are separated by a surmountable barrier, which allows the conformations to switch between the states. Unless it is a repressor, under physiological conditions it typically populates an inactive state. Binding (or post-translational modification) triggers a switch to the active state. Allosteric mutations act by shifting the ensemble from the inactive to the active state (or vice versa) which can take place by either destabilizing the inactive state, stabilizing the active state, or both. A scaffolding protein which binds multiple partners similarly exploits (some of) its multiple states, each of which recognizes a distinct ligand with a subsequent shift of the conformational ensemble. The distribution of the states is often tuned by post-translational modifications.

Binding events are neither binary nor stationary. Architecture and multi-scale organization is what distinguishes a living cell from random assemblage in solution. Cellular architecture is important for the cell's properties, including morphology, motility, metabolism, chromatin organization and more; all temporal and dynamic, linked to external and internal signaling cues and the changing cellular environment. Binding events are crucial for signaling, and signals propagate through interactions involving dynamic reorganization of multiprotein complexes. Dynamic association implies not merely interactions forming and dissociating; it speaks of cooperativity. A key challenge is to understand this interplay, link it to the physicochemical basis of the conformational behavior of single molecules, and ultimately relate it to global cellular function. Cell signaling can be thought of as allostery-driven forming and reforming binding events taking place within dynamic, loosely preorganized assemblies. Cellular processes are temporal, and can be understood only in terms of dynamics within, and among

multimolecular complexes. And within this framework, interactions and coordination are governed by a conformational biasing mechanism, that is, population shift of the ensembles as described by the free energy landscape. Multivalent proteins with multiple partners are typically engaged in dynamic membrane-anchored and cytoskeleton-attached network, bestowing signaling with gel-like properties.

The collection of papers included in this Folding and Binding issue aim to capture some of the functional aspects of folding and binding as well as predictive and analytical methodologies. Rao and Gosavi use the folding landscapes of proteins to understand protein function. They point out that proteins fold in a biologically-relevant timescale because of their funnel-shaped energy landscape, which evolution has etched by selecting sequences that stabilize interactions present in the folded state and disfavoring stable non-native interactions. At the same time, evolution selects residues which are involved in function such as binding-site and active-site residues that do not optimize folding. However, they argue that the interactions of such functional residues with residues that promote folding modulate the shape of the energy landscape which would affect folding, and provide examples in support of this view. They also provide a scheme to detect such effects. Takahashi *et al.* review spectroscopic investigations of single molecule and ensemble protein folding dynamics. Data show that substates that contain residual structures in the unfolded and partially populated states, lead to a complex behavior in the early folding dynamics of small proteins, which can be explained by the native-centric model. In contrast, recent observations on large proteins point to the rapid formation of long-range contacts that appear inconsistent with the native centric model, suggesting that the folding strategy of large proteins differs from that of small proteins. Their review brings out the necessity of understanding the conformational fluctuations that occur in the unfolded state, which precede and prime structure formation. Muñoz *et al.* bring out the important point that protein folding reactions have limited cooperativity. Folding cooperativity has been difficult to address experimentally. However, new analytical procedures demonstrate a general scenario of limited cooperativity linking between how fast a protein folds and unfolds, and how cooperative is its equilibrium unfolding. They note that because it affects unfolding more than folding, reduced cooperativity also destabilizes the native structure. This leads them to define cooperativity scale that goes from the 'pliable' two-state of slow folders to the gradual unfolding of one-state downhill, and at the end of the spectrum to intrinsic disorder and they suggest a conformational rheostat mechanism for the allosteric effects of folding coupled to binding. So *et al.* discuss protein misfolding leading to amyoid fibril formation. They revisit supersaturation as a factor in amyloid fibrillation. Amyloid fibrils are involved in many diseases, and usually form by a nucleation-growth mechanism. They emphasize that as in crystallization of solutes, solubility and supersaturation are also two key factors that impact the formation of amyloid fibrils. They point out that the impact of these two factors on the partitioning between distinct types of aggregates (e.g. fibrils vs. amorphous aggregates) can be explained by kinetic and thermodynamic competition which suggests that supersaturation is a key component in phase transitions of denatured proteins and aggregation. Gianni *et al.* ask what we can learn from the kinetics of coupled binding and folding of intrinsically disordered proteins. They argue that to fully appreciate why protein disorder is advantageous for protein–protein interactions we need to understand the mechanism(s) of their interactions, which is challenging. Their review focuses on how kinetics in combination with protein engineering and structural information can be used to obtain details of protein–protein interactions involving intrinsically disordered proteins. Banerjee *et al.* discuss the many roles of the disordered hypervariable variable region (HVR) of KRAS4B, a splice variant of KRAS, a highly oncogenic Ras isoform. Classically, the role of the post-translationally-modified HVR is to navigate Ras, a GTPase, in the cell and to anchor it in localized plasma membrane regions. However, based on their recent work and supported by the literature the authors pointed out additional, overlooked, regulatory roles. These include autoinhibition by shielding the effector binding site in the GDP-bound state and release upon GTP binding and in the presence of certain oncogenic mutations. They proposed that the released positively charged and post-translationally-modified HVR can interact with calmodulin. They also propose that oncogenic mutations (G12V/G12D) may modulate the HVR-phospholipid binding specificity. Overall, the disordered state of the HVR exemplifies the critical role of the unstructured tail of K-Ras4B in cancer. Finally, Perez *et al.* overview advances in free-energy-based simulations of protein folding and ligand binding, pointing out that free-energy-based simulations are increasingly able to narrate protein structural dynamics and biological mechanisms. They focus on two recent successes, first that it is becoming practical to fold small proteins with free-energy methods without knowing substructures, and second to compute ligand-protein binding affinities — not just their binding poses. These successes are largely due to GPU-based computing, improved fast-solvation methods, and continued advances in force fields, and conformational sampling methods.