Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

# Editorial overview: Folding and binding: *In silico*, *in vitro* and *in cellula*

Amedeo Caflisch and Stefano Gianni

**Current Opinion in Structural Biology** 2018, **48**:iv–vii

For a complete overview see the Issue

https://doi.org/10.1016/j.sbi.2018.01.014



## Amedeo Caflisch

Department of Biochemistry, University of Zurich, CH-8057 Zurich, Switzerland
e-mail: caflisch@bioc.uzh.ch

Amedeo Caflisch studied physics at the ETH in Zurich. During 1992–1994 he was a postdoctoral fellow in the group of Martin Karplus at Harvard University. In 1996 he was offered an assistant professorship position at the Department of Biochemistry of the University of Zurich, and promoted to full professor in 2001. His research team is known for the development of methods for enhanced sampling molecular dynamics and data-driven analysis of simulations of protein folding, aggregation, and ligand binding. The group has also developed fragment-based docking protocols, which have been applied to proteases, kinases, and protein-protein recognition domains. In 2013, at the age of 50, he has decided to start experimental activities in his group, which have resulted in the crystal structures of 10 tyrosine kinases/inhibitor complexes and nearly 100 bromodomain/ligand complexes. Most of the ligands in these complexes originate from docking.

## Stefano Gianni

Istituto Pasteur – Fondazione Cenci Bolognetti, Dipartimento di Scienze Biochimiche "A. Rossi Fanelli" and Istituto di Biologia e Patologia Molecolari del CNR, Sapienza Università di Roma, 00185 Rome, Italy
e-mail: stefano.gianni@uniroma1.it

Stefano Gianni is Professor of Molecular Biology in the Department of Biochemistry of the Sapienza, University of Rome (Italy), where he is also Director of the PhD Program in Biochemistry. He completed his PhD in Rome and, after a post doc with Alan Fersht in Cambridge (UK), he started his independent group in 2005. His lab is known for the study of protein folding and protein–

The essence of any biological processes relies on the conformational states of macromolecules and their interactions. It comes therefore with no surprises that the study of folding and binding has been centre stage since the birth of structural biology. In this context, the collaborative efforts of experimentalists and theoreticians have tremendously increased our current knowledge on macromolecular structure and recognition. Nevertheless, several challenges and open questions are still present and a multidisciplinary approach would appear the most appropriate means to shed light onto the mechanisms of folding and binding to the highest level of detail. This thematic issue brings together a collection of reviews describing our current understanding of folding and binding, looking at these fundamental problems from a wide perspective ranging from the single molecule to the complexity of the living cell, drawing on approaches that span from computational (*in silico*), to the test tube (*in vitro*) and cell cultures (*in cellula*).

Our current interpretation of chemical reactions is founded on the model of Arrhenius, generally known as the transition state theory. Following Arrhenius, chemical kinetics has a probabilistic nature and reactants interconvert quasi instantaneously into products with a probability over time that is proportional to the free energy of the intervening barrier. The time required for a molecule to jump between states is called the 'transition path time' and, until very recently, has escaped experimental detection for any molecular reaction, until single molecule FRET folding experiments have been developed. Chung and Eaton provide a comprehensive review on the key findings arising from these studies. In particular, on the one hand, the authors highlight how the study of transition paths have provided an experimental validation of predictions arising from all-atom molecular dynamics and the energy landscape theory and, on the other hand, they depict the challenges ahead.

The experimental study of protein folding and binding has been classically carried out *in vitro* and in diluted solutions. Therefore much has still to be learned about how folding proceeds in the complex and crowded cellular environment. Davis, Gruebele and Sukenik, describe how recent innovative experimental techniques allow measuring folding in the living cell. Importantly, they show how these studies have demonstrated that weak but correlated interactions may lead to large physiological effects on both folding and binding. The authors describe in detail the role of molecular crowding and quinary interactions in affecting protein folding as well as in perturbing binding affinities to promote the formation of higher order complexes. These intermolecular interactions appear to play a key role in regulating protein structure and function during the cell cycle or in different cellular environments.

protein interactions, investigated through a combination of protein engineering and biophysical methods. He has also worked for several years on the characterization of PDZ domains, with particular emphasis on their allosteric behaviour. More recently, he directed his scientific attention towards the mechanisms of binding-induced folding of intrinsically disordered proteins.

Among the cellular processes regulating protein folding, the role of molecular chaperones is one the most prominent. Bardwell and co-workers present a review of our current knowledge of an important class of molecular chaperons that assist protein folding by continuously interacting with the client. In these cases, folding occurs while binding to the chaperone in the absence of ATP hydrolysis, a mechanism which has been demonstrated in the case of Spy, GroEL and SecB. While recapitulating the main findings, the key experiments and the physiological roles of these chaperones, the authors contribute a general description of such a mechanism and speculate that it could be evolutionary ancient, with ATP hydrolysis appearing later to modulate chaperone regulation.

The review by Nitsche and Otting focuses on nuclear magnetic resonance (NMR) methods developed for the analysis of ligands that bind tightly to the protein target, that is, when ligand dissociation is not fast enough to allow magnetization transfer between bound and free ligand. The authors compile first a meta-review of seven review articles on NMR studies of ligand binding published in 2016–2017. They then discuss recent advances in synthetic chemical tags, in particular paramagnetic lanthanide tags. Paramagnetism is due to the presence of unpaired electrons in metal ions and a few non-metal compounds such as nitroxides. Paramagnetic tags influence the NMR spectrum of the entire protein because unpaired electrons have a strong magnetic moment. It is important to note that the chemical tags report on structural information independent of stable isotope labelling, which is very expensive. The authors provide evidence that NMR spectroscopy methods are useful for both the initial and advanced phases of drug discovery, in particular ligand-observed NMR and synthetic chemical tags, respectively.

Frustration is a condition arising from the perceived inability to achieve a given goal. The funneled energy landscape theory is based on the assumption that proteins are minimally frustrated, such that their sequences are evolved to avoid mis-folded traps. In these conditions, therefore, the energy landscape appears strongly biased towards the native conformation being by and large funneled. But, because proteins are not only evolved to fold, but also to function, it is reasonable to assume that they may locally display some frustration patterns, which may be associated to their biological role. Ferreiro, Komives and Wolynes provide an insightful review describing the critical importance of localizing frustration in proteins. In particular, after briefly introducing their methods to localize energetic frustration from the structure and sequence of proteins, they highlight the role and significance of such frustration arising from topology, function and naturally occurring mutagenesis. In doing so they provide an interesting view that poses frustration as a central concept in structural biology.

Huang and MacKerell Jr review recent advances in empirical force field development for molecular dynamics simulations of intrinsically disordered proteins (IDPs). The most commonly used experimental methods for the analysis of structure and function of proteins, for example, protein X-ray crystallography and cryo-electron microscopy, are not directly applicable to IDPs. In contrast, thanks to the continuous improvement of hardware performances, for example, clusters of specialized processors and graphics processing units (GPUs), and vigorous development of enhanced sampling protocols, atomistic simulations can reach time scales (tens of microseconds) that allow for the analysis of equilibrium properties even for IDPs of up to about 100 residues. Recent force field improvements for IDPs have focused on the propensity for regular secondary structure (α-helices and

β-sheets) and the fine-tuning of the relative contributions of protein–protein and protein–solvent interactions. The authors conclude their review with an optimistic forecast on the further improvement not only of parameter sets but also potential energy functions with the inclusion of electronic polarizability.

The identification of specific binders to proteins is a very powerful tool in drug discovery. For this purpose, computational methods have tremendously improved in the last decade. Sledz and Caflisch contribute a comprehensive review describing the current status of computer-based approach in drug selection. In particular, they first describe the major advances of molecular docking, which taking advantage of the lock and key concept is frequently used for the high-throughput screening of molecular libraries, and then address the role of explicit solvent molecular dynamics simulations protocol to describe in greater detail the interaction between a protein and its drug taking into account also the dynamics and the mutual adaptability. The authors succeed in providing a wide view on the current status of *in silico* drug discovery, comprising both the description of the methodological approaches as well as the general features of ligand binding.

Csizmok and Forman-Kay review the effects of multiple post-translational modification (PTMs) on the regulation of protein stability and function. A large variety of more than 300 PTMs have been found in eukaryotic proteins and mainly IDP segments. PTMs are essential for the regulation of many biochemical processes, for example, gene expression, enzyme activity, and protein–protein association. The physico-chemical properties of individual PTMs are well understood. As an example, lysine acetylation removes a positive charge on the tip of the side chain resulting in a reduced affinity for negatively charged groups, for example, aspartate or glutamate side chains of phosphate groups. In contrast, the interplay of multiple PTMs can result in very different effects depending on the number and positioning (along the sequence of the modified protein) of the PTMs. An important conclusion of this review is that new experimental methods have to be developed to shed light on the crosstalk of multiple PTMs playing a key role in several fundamental biochemical processes.

α-Synuclein (αS) is an intrinsically disordered protein primarily expressed in brain, whose aggregation is associated to Parkinson disease. Its biological role is mediated by its partitioning to a soluble and membrane-bound state, which appears to be tightly regulated *in vivo*. De Simone and co-workers provide a complete review of the structural and dynamic features of αS. Of particular interest, they describe in high detail the characteristics of αS when in its association with membrane as well as the influence of pathological mutations and PTMs in the membrane binding processes. The review presents three very nice pictures that allow to capture the structural features of αS in isolation, in its membrane bound state as well as explaining the promotion of vesicle–vesicle interaction via a so called 'membrane-anchoring mechanism'.

Wang *et al.* review recent studies in which molecular dynamics simulations of transmembrane receptors were used to help in the interpretation of experimental data, for example, to analyse the sequence of events during transitions between structures of different functional states. Examples from three classes of receptors are presented: G-protein coupled receptors, ligand-gated ion-channels, and single-pass receptors (tyrosine kinases and class I cytokines). These examples include receptors that vary in the number of transmembrane segments and show a broad range of conformational heterogeneity and types of motion with time scales ranging from ns to seconds. In many of these studies atomistic simulations on dedicated hardware, coarse-grained models, enhanced sampling techniques, and/or out-of-equilibrium initial conditions have resulted in sufficient sampling for the investigation of conformational transitions that require microseconds to milliseconds. Four beautiful and very clear figures elegantly illustrate the main points of this review.

Zhang and co-workers focus their review on membrane-associated guanylate kinases (MAGUKs). This family of proteins is characterized by a PDZ-SH3-GK tandem motif of their C-terminal end and is critical for the tissue development and homeostasis. Most remarkably, mutations of MAGUKs were associated many diseases in human ranging from cancer to psychological disorders. The review provides an insightful and complete review of MAGUKs structure and function highlighting both the different mechanism of recognition of MAGUKs, the allosteric features of the PDZ-SH3-GK supramodule, as well as the ligand-induced oligomerization of MAGUKs and the link phase transitions of synaptic signaling complexes.

The collection of reviews presented in this issue of COSB highlight the complexity of folding and binding from very heterogeneous methodological and biological perspectives. Despite the heterogeneity of approaches, two main points seem to emerge. First, reaching and maintaining the functional form of a protein depends mainly on the amino acid sequence but also on the presence of chaperones, limited crowding, multiple post-translational modifications, and/or a macromolecular binding partner. Thus protein folding and (macro) molecular binding are substantially more complex than a simple monomolecular and bimolecular process, respectively. Second, novel computational and experimental techniques are required to analyse individually each of the factors that influence protein (mis)folding and binding, and more importantly, to capture the interplay

between them. We look forward to the development of new computational models and experimental techniques for investigating the intricate aspects of protein folding and binding. New methods will improve our scarce understanding of protein misfolding, amyloid aggregation and related toxic species. A better understanding at the molecular level is essential for the discovery of curative treatments for Alzheimer's disease and other (neurological) disorders whose hallmarks are protein misfolding and aggregation.