Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Editorial overview: Folding and binding

James C A Bardwell and Gideon Schreiber

Current Opinion in Structural Biology 2014, **24**:viii–x

For a complete overview see the Issue

Available online 6th February 2014

0959-440X/$ – see front matter,

http://dx.doi.org/10.1016/j.sbi.2014.01.012

James C Bardwell

Howard Hughes Medical Institute, University of Michigan, Ann Arbor, MI 48109-1048, USA
e-mail: jbardwel@umich.edu

James C Bardwell is a Howard Hughes Investigator at the University of Michigan. His lab is involved in the discovery and characterization of chaperones. We have developed biosensors that can monitor stability *in vivo* and select for improved folding. By linking the folding of proteins to cell growth, we force organisms to find solutions for a variety of different folding problems. By examining these solutions, we hope to gain insight into the *in vivo* folding process. Our folding biosensors have allowed us to address questions about the relationship between stability and function, to explore the *in vivo* folding environment, and to discover previously unknown chaperones. We also use biophysical approaches to study the mechanism of action of these new chaperones as well as previously discovered chaperones. The knowledge gained by using these approaches should allow us to better understand and manipulate *in vivo* protein folding.

Gideon Schreiber

Department of Biological Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel
e-mail: gideon.schreiber@weizmann.ac.il

Gideon Schreiber from the Department of Biological Chemistry at the Weismann Institute of Science is studying various aspects of protein-protein interactions. Specific protein–protein interactions form a major part of the basic organization of living cells. My laboratory studies structure/function relations of protein–protein interactions and how these relate to biological signaling. For this purpose, we adopted a multidisciplinary approach including biophysical and biological bench work, protein-engineering, bioinformatics, and algorithm development and applied the gained knowledge towards solving biological questions.

Studying proteins has shifted in recent years from the reductionist approach of studying single proteins and their interactions *in vitro* to the study of dynamic systems in their native environment. This is now possible, thanks to the many new tools developed, allowing the measurements of quantitative, thermodynamic and kinetic data in complex environments that contain many components, at single particle resolution. Recent advances in these methodologies and some of the exciting biological understandings that have been revealed by these methods are reviewed in the articles of this section.

A longstanding dream of many in the protein-folding field is to understand folding in sufficient detail in order to be able to model the process. Piana *et al.* estimate that is now possible to model with all atom high-speed MD simulations the folding of about 10% of the single chain proteins in the PDB, namely those that are expected to have folding times of a few milliseconds or less. He then discusses how limitations in the current accuracy of the force fields used in these simulations differentially affects the various predictions, native-state structures and folding rates appear to be better predicted than are the detailed kinetics or the structural properties of the unfolded state.

While protein folding *in vitro* is a spontaneous well-studied process, folding in the cell is much less studied and complicated by the complex cellular environment. Folding *in vivo* occurs in a co-translational manner in a very crowded environment and is positively influenced by the action of chaperones and catalysts, and negatively influenced by aggregation processes. Hingorani and Gierasch describe recent efforts to bridge the gap between our knowledge on folding *in vitro* and *in vivo*. They describe how crowding may positively influence folding and how co-translational folding and the presence of chaperones can alter the folding landscape and the evolutionary trade-offs that affect *in vivo* folding. While detailed understanding of how proteins fold in the cell is no longer a receding aim, it still is a distant goal.

The concept of co-translational folding is examined in more detail by Bukau and coworkers. They introduce the concept of translational machinery as a hub coordinating protein maturation events. The active role of the ribosome itself in providing an early folding environment and in slowing overall folding while increasing its accuracy is described, as are the roles of ribosome associated nascent chain-interacting factors. Though it is becoming clear that the ribosome helps to detect and coordinate structural signals on the emerging nascent chain that signal the need for membrane targeting, co-translational folding or enzymatic processing. However, exactly how it detects and processes these structural signals will keep workers busy for some time to come.

Controlling protein-degradation is not less important than protein production. An overview of progress in understanding the proteasome structure

and its regulation is presented by Inobe and Matouschek. A flurry of papers have appeared recently that describe the structure of the 19S component of the proteasome, the component that is involved in substrate recognition and the regulation of the proteasome's activity. The substantial conformational changes that allow the proteasome to cycle between a open mouthed channel that can digest its substrates and a presumably inactive closed mouth channel have now been visualized clearly. It has long been known that ubiquitination is one way of targeting proteins for proteolytic degradation but about half of ubiquinated proteins are not degraded, a second signal, in the form of a disordered region may also be required. In some cases disordered regions can suffice to initiate degradation. Regulation of degradation is both richer and more subtle than previously thought.

Another longstanding goal is to be able to determine not just which proteins interact with each other, as presented by two dimensional 'interactome' maps but to be able to visualize these interactions in 3 dimensions at high resolution. Szilagyi and Zhang discuss how the expanding structural database of protein–protein complexes can be utilized through template based modeling to predict the structure of complexes that are not yet in the database. High speed identification of templates and improvements in the ability to construct models starting from sequences alone have allowed template based modeling approaches to achieve impressive results. These seem now to be comparable in accuracy to high throughput efforts such as the yeast two hybrid approach, at least at the level of predicting if two proteins interact or not.

Protein–protein interactions are investigated at different spatial and temporal resolutions. X-ray crystallography provides mainly high-resolution spatial data, while kinetic reaction measurements provide temporal data. NMR is unique in its ability to provide both. Recent advances in methodologies, as discussed in the review by Hass and Ubbink make it possible to investigate dynamic processes at molecular resolution also for large systems. They present practical considerations for structure determination of protein–protein interactions using paramagnetic NMR restraints. The restraints are applied in a similar manner as NOEs have been used in the past. While paramagnetic NMR restraints are generally fewer in number, they provide information over much larger distances and thus can be used on larger protein-complexes and provide valuable information on the dynamics of interactions investigated. A number of interesting examples of protein complexes, with both their structure and dynamics being characterized by paramagnetic NMR data are presented.

The development of deep sequencing methods has revolutionized many aspects of biology. In this issue, Tripathi and Varadagajan review another use of deep sequencing data that gained much attention in recent years: retrieving residue specific contributions to protein and protein-complex stability from saturation mutagenesis and deep sequencing. The use of deep sequencing has the potential to provide complete data on all point mutations at the binding surface extracted from a natural pool of binders, or following random or focused mutagenesis and *in vitro* selection for binding. This allows for nonbiased mapping of mutational effects on binding and stability. The basic principle behind this method is that selection will enrich for amino-acids that contribute to a specific function. Deep sequencing then provides the amino-acid preference at any randomized position, which is related (with a Boltzmann like distribution) to the contribution towards the selected trait (for example binding). The first part of their review describes methods for creating the libraries and their selection. This is followed by potential applications, including guiding protein-design, protein-model discrimination, refining computational predictions and epistatic effects. Importantly, the review also addresses potential pitfalls of the methodology, including the extent that the amino-acid preferences revealed by this approach indeed reflect differences in binding affinity and protein stability.

Antibody binding to protein–antigens is a classical, and well-studied example of protein–protein interactions. In addition to its natural role in identifying foreign antigens in our body and directly and indirectly eliminating them, antibodies have become a cornerstone for drug development, synthetic biology, and molecular biology. To be able to rapidly produce specific, high affinity antibodies needed for these different applications, synthetic antibody technologies have been developed. Sidhu and Adams present in their review the current state of synthetic antibody technologies and their varied applications. Particularly, they discuss display technologies that enable amplification, selection and manipulation of antibodies *in vitro*. The controlled nature of the procedure allows the development of tools frameworks and position specific selections that would not be naturally possible. In addition to reviewing the various selection methodologies, they also provide an exiting update on the many uses of synthetic antibodies, including as crystallographic chaperons, modulators of protein function, and detection of posttranslational modifications.

An overview on how quantitative protein-interaction analysis can be done in living cells is presented by Piehler. Real time spectroscopic methods have significantly improved in recent years due to developments in super resolution microscopy and quality of fluorescence labels, allowing for single molecule tracking and binding measurements to be performed in quantitative ways. These, in turn have significantly enhanced our understanding of the biophysical basis of biological processes. However, this review also points to the many obstacles

that have to be overcome, and that in addition to specialized experimental equipment also sophisticated data evaluation is required to achieve optimal results.

Post-translational modifications (PTMs) often dictate the potential of proteins to bind. Here, Woodsmith and Stelzl present a collection of recently published large protein–protein interaction datasets with emphasize towards establishing the relationships between protein–protein interaction, PTMs and signaling. System wide analysis of PTMs has advanced thanks to the development of proteomic mass spectrometry tools. However, much less is known about the relation between the modifying enzyme (for example a kinase) and its substrate protein due to the weak, transient interactions between them. This is particularly important for defining modifying enzyme to substrate relationships and in deciphering the PTM conditional interactions, which are a basic feature of controlling signaling cascades. Overall, Woodsmith and Stelzl make a strong case that analysis of PTMs coupled to protein interaction information will promote a better understanding of enzyme–substrate relationships, the dynamics of PTM-mediated signal flow and the consequences of PTM-mediated recognition events, that is the rewiring of molecular networks as a signaling response.

Overall, this exiting issue provides us with a clear view on the changes occurring in studying folding and binding during the last few years. Advances in technology have allowed us to study these systems at a much higher level of complexity, whether by using high throughput methods to achieve system wide resolution, studying the processes within the complex cellular environment or by retrieving high resolution kinetic and structural data of large systems. We are clearly on our way to understand cellular structural biology at a molecular level, with exciting implications to future medicine.