Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Editorial overview: Protein–protein interactions

Alexandre MJJ Bonvin and Özlem Keskin

Current Opinion in Structural Biology 2015, **35**:vii–ix

For a complete overview see the Issue

Available online 26th November 2015

http://dx.doi.org/10.1016/j.sbi.2015.11.004

Alexandre MJJ Bonvin
Bijvoet Center for Biomolecular Research, Science Faculty/Chemistry, Utrecht University, Utrecht, The Netherlands
e-mail: a.m.j.j.bonvin@uu.nl

Alexandre Bonvin (1964) studied Chemistry at Lausanne University and obtained his PhD at Utrecht University (1993). After two post-doc periods at Yale University (USA) (Prof. A.T. Brunger) and at the ETHZ (Prof. W.F. van Gunsteren) he joined Utrecht University in 1998 where he was appointed full professor of computational structural biology in 2009. His research focuses on the development of reliable computational approaches to predict, model and dissect biomolecular interactions at atomic level, integrating various experimental information sources in order to obtain a comprehensive description of the structural and dynamic landscape of complex biomolecular machines. His group develops the widely used HADDOCK software for the modelling of biomolecular complexes (http://bonvinlab.org/software).

Özlem Keskin
Chemical and Biological Engineering, College of Engineering, and Center for Computational Biology and Bioinformatics, Koc University, 34450 Istanbul, Turkey
e-mail: okeskin@ku.edu.tr

Özlem Keskin is currently a professor of Chemical and Biological Engineering Department at Koc University, Istanbul. Before, she was a postdoctoral fellow at the National Cancer Institute-National Institutes of Health, U.S.A., during 1999–2001. She is a member of the Science Academy, Turkey. Her work focuses on understanding the principles of protein–protein interactions (PPIs), the molecular mechanisms, physical principles and dynamics of macromolecular systems. The Computational Systems Biology (COSBI) group aims to construct protein interactomes and integrate atomistic details of protein-protein interfaces. Various tools/methods are developed in the group:

Proteins and their intricate network of interactions are the mainstay of any cellular process. Dissecting these networks at atomic detail is invaluable, as this paves the route to a mechanistic understanding of biological function and provides the first essential step toward the development of new drugs. Adding the structural dimension to those networks is a challenging task that requires structural biologists to resort more and more to a combination of various experimental and computational methods. In this section about Protein–Protein Interactions, experts in the field review a number of recent developments.

Next to obtaining structures, or structural models, of biomolecular complexes, it is crucial to get proper information on the dynamics, stoichiometry, thermodynamics and kinetic parameters of the biologically relevant assemblies. There have been indeed cases, for example in CAPRI (Critical Assessment of Predicted Interactions), where the crystal structure was misleading and not pointing to the functional biological assembly. Similarly, high quality thermodynamics data on biomolecular interactions are key to pushing forward our understanding of the principles that govern binding affinity as current methods are still very much limited in their predictive power. The review by Byron and Vestergaard covers those aspects, together with the various experimental methods that can provide us with pieces of the supra-structural biomolecular complex puzzle. They describe the complexity caused by dynamics, allostery, heterogeneity and aggregation, which can only be properly tackled by using trans-disciplinary biophysical approaches.

From an overview of experimental approaches we then move into the modelling field with a first contribution by Muratcioglu *et al.* who describe how the current structural knowledge available in the Protein Data Bank (PDB) can be exploited for template-based modelling of interactions, both by global structural alignment or by focusing only on the interfaces, as they remark: ''Similar protein–protein interfaces are observed between different proteins''. Despite the wealth of interface information in the PDB (estimated to be over 27,000), there are still many novel interfaces to be explored, which calls for more structural templates and thus experimental structures to be solved. While template-based modelling has the potential to significantly increase the structural coverage of the interactome, there will always be specific cases where it might be misleading and thus should be applied with caution. A single mutation, post-translational modifications, environmental changes/factors or conformational dynamics might all play governing roles in the selectivity of interactions.

From template-based, we then move into *ab initio* modelling with Park *et al.* who review the status of protein-protein docking. Their focus is on

PRISM to model 3D structures of protein complexes, HotPoint to predict hotspots at protein–protein interfaces and Piface to query the clusters of protein–protein interfaces (http://prism.ccbb.ku.edu.tr).

high-resolution methods where interactions are described at the atomic level, comparing global search methods with more targeted searches, and the various way of accounting for flexibility in the modelling process. Despite recent advances, often catalyzed by the CAPRI experiment, flexibility and conformational changes remain very much an open challenge, especially when those changes exceed 2–3 Å interface root-mean-squared-deviations, not to mention complexes involving the binding of an intrinsically disordered peptide (discussed in a contribution by Tompa *et al.* later in this section). Park *et al.* also discuss the challenges of accounting for interfacial solvent molecules. The progresses in that area indicate that the field is maturing, also driven by CAPRI.

Next to the classical experimental techniques like NMR (see review by Wiesner and Spranger), X-ray crystallography and cryo-electron microscopy (not reviewed in details in this issue), mass spectrometry (MS), reviewed by Liu and Heck, is establishing itself as one of the experimental methods that will contribute significantly to large scale studies of protein-protein interactions, even in a cellular context. Liu and Heck focus on cross-link MS (XL-MS), which, in combination with affinity purification MS and advances in analysis software, allows the identification very large and intricate biomolecular assemblies from very low sample concentrations and even mixtures such as cell lysates. This allows for large scale, proteome-wide analysis of interactions. The authors even recently demonstrated the feasibility of carrying such studies in a cellular environment. Clearly, the information delivered by XL-MS is perfectly suited for combination with a variety of modelling approaches (such as docking reviewed here by Park *et al.*). It will become a standard method in the integrative toolbox of structural biologists.

While NMR is a well-established structural method, it has suffered from size limitations when it comes to characterizing large macromolecular assemblies. Here Wiesner and Spranger review how the combination of methyl group labeling approaches and relaxation-optimized NMR experiments has been lifting the size limitations usually associated with NMR. They illustrate this with various application examples covering the range from structural studies (e.g. the characterization from chemical shift perturbations and paramagnetic relaxation enhancement effect of the 650 kDa ClpB-DnaK chaperone system) to the characterization of dynamics and allosteric changes in large and complex assemblies (e.g. the identification of allosteric conformational changes linked to ligand binding in GPCRs), or even the characterization of 'invisible' states (low populated states typically invisible to X-ray crystallography). They end by describing the recently introduced LEGO NMR approach that allows the preparation of highly asymmetric complexes, often present in eukaryotes, which are NMR active in only a subset of the subunits, thereby simplifying the NMR analysis. Again, all this NMR information, complemented by for example MS and SAXS data, can be used together with structures or models to build models of large macromolecular assemblies following an integrative modelling approach.

We then move to an experimental study bringing together classical microbial genetics, biochemistry, biophysics, structural biology and imaging. Kleanthous *et al.* demonstrate how spatiotemporal dynamics of bacterial outer membrane proteins (OMPs) govern their turnover in the membrane. They show with the novel OMP labeling strategies coupled with imaging techniques what happens to OMPs once they inserted in the OM of bacteria, principally for the model organism *Escherichia coli*. They demonstrate that OMPs have restricted mobility compared to inner membrane proteins. They

also discuss how promiscuous interactions between outer membrane proteins are utilized in the formation of large OMP clusters as the basis for their characteristic diffusion behavior in the membrane. The review not only organizes and discusses contradictory existing data but also provides a historical aspect and progress on the methods. The authors nicely speculate on the areas of bacterial cell biology that could be influenced by these latest advances.

While, when speaking about protein-protein interactions, we are often thinking about interactions between structurally well-defined molecules, intrinsically disordered proteins (IDPs) or disordered regions within one protein (IDRs) are actually playing an important role in protein-protein interactions, accounting for a very significant portion of the interactome. This is reviewed in the contribution by Tompa *et al.* who discuss various examples and aspects of those interactions. These systems are challenging to both model and characterize experimentally, major challenges being the lack of well-defined structure and folding upon binding events. Rather small motifs are typically involved in such interactions, with post-translational modifications controlling their binding behavior. As for very large and complex assemblies, only a combination of various experimental and modelling methods can shed light on those 'fuzzy' complexes.

Haliloglu and Bahar introduce structure-encoded dynamics (also called intrinsic dynamics) as one of the major determinants of protein-protein and protein-ligand interaction mechanisms. Intrinsic dynamics represents the conformational motions, or the spectrum of modes, uniquely defined by the 3-dimensional structure. Conformational dynamics and flexibility appears to be essential for protein-substrate interactions, protein-protein interactions and allosteric control. They discuss how soft modes define pre-existing pathways of structural change on the energy landscape. In their review, they discuss how evolution of sequences and structures make use of intrinsic dynamics to determine functional interactions and to adapt multispecificity. The authors conclude that "Methods that exploit the intrinsic dynamics of proteins are likely to open the way to new strategies that for design, discovery and therapeutics."

Sethi *et al.* provide an overview of the growth of genomic and structural data over years and how data integration at sequence-, structure- and network-levels could enhance our understanding of the effect of rare and disease-causing mutations in humans. The authors provide a comprehensive review on existing and ongoing human genome sequencing projects. They first explain how genomic information is used to identify pathological disease associated variants. Then, they show how structural information is utilized to understand the harmful effects of different variants. They finally corroborate that while structural and sequence data are critical to understand the deleterious effects of certain disease-causing and rare variations, it is often difficult to interpret the phenotypic effects of an individual variant without considering the broader cellular context, that is at network level.

The last review combines and discusses protein-protein interactions, dynamics, evolution and genome-wide detected non-synonymous SNVs. Following the concepts introduced in Sethi *et al.*, Kumar *et al.* discuss recent advances in methods and approaches to predict the disease-related impacts of nsSNVs. Methods that predict the disease-associated (phenotypic) effects of mutations and those that assess ΔG of binding are overviewed. As also discussed in Haliloglu and Bahar's review, this article highlights that changes in conformational dynamics through allosteric regulation lead to new functions in proteins. They introduce dynamic flexibility index (DFI), which is a position-specific metric that quantifies the resilience of each residue to a perturbation occurring at another part of the chain, thus identifying the flexible and rigid parts of a protein. They indicate that disease-associated nsSNVs occur predominately at low DFI sites.

In summary, this issue aims to bring together different points of views from experimental and computational studies. It is now clear that we need integration of computations with data obtained from different experiments at different scales: from atomic details to abstract networks of interactions. These types of data will surely become more available in the near future with the advancement of experimental methods. Personalized medicine will make use of sequencing technologies that will provide individual human's genetic variations. However, making sense out of it will only be possible provided simultaneous progresses take place in the analysis of these data.