Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Protein–protein interactions

Editorial Overview

Joël Janin and Alexandre MJJ Bonvin

Current Opinion in Structural Biology 2013, **23**:859–861

For a complete overview see the Issue

Available online 12th November 2013

http://dx.doi.org/10.1016/j.sbi.2013.10.003

Joël Janin
IBBMC (Institut de Biochimie et Biologie Moléculaire et Cellulaire), Université Paris-Sud, 91405 Orsay, France
e-mail: joel.janin@u-psud.fr

Joël Janin did a PhD in Biochemistry (Paris, 1969), and a postdoc in Cambridge, UK, before starting one of the first French groups in protein crystallography at the Pasteur Institute in Paris. In 1981, he was appointed Professor of Biophysics at Université Paris-Sud, Orsay. He was Director of the Laboratoire de Biochimie et Enzymologie Structurales in Gif-sur-Yvette from 1992 to 2006, and now is Professor Emeritus at Institut de Biochimie et Biophysique Moléculaire et Cellulaire in Orsay. His research activity has long concerned the catalytic mechanism of phosphokinases and other enzymes, but in recent years, the main subject has been modeling protein–protein recognition, a field that he initiated in the 1970s and includes managing the CAPRI community-wide experiment.

Alexandre MJJ Bonvin
Faculty of Science – Chemistry, Bijvoet Center, Utrecht University, Padualaan 8, 3584 CH 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 postdoc periods at Yale University (Prof. AT Brunger) and at the ETHZ (Prof. WF 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 modeling of biomolecular complexes (http://haddock.science.uu.nl).

We are proud to present the first edition of the *Protein–protein interactions* Section of *Current Opinion in Structural Biology*. The Section is new, but the topic has been present in the journal from the very start. Volume 1, Issue 1, dated February 1991, had a review by Janin entitled *Protein–protein interactions and assembly*, and others by Bode and Huber on *Proteinase–inhibitor interaction*, and by Chothia on *Antigen recognition*. The Editorial Overview, signed by TE Creighton and PS Kim, noted that 'several new X-ray structures of protein complexes had been determined', and stated that, for such complexes, the 'overall goal was to understand the specificities of the binding reactions [and the] differential affinities'. At the time, most of the known structures displaying protein–protein interactions were oligomeric protein complexes, which are permanent assemblies. Protease–inhibitor and antigen–antibody complexes were the only examples of transient, although highly affine, interactions for which the questions of specificity and affinity could be raised, and therefore, the three reviews of Issue 1 pretty much covered the field. Soon afterwards, new X-ray structures illustrated the role of protein–protein interactions in several major cellular processes. A landmark was the structure of the transducin G_α–$G_{\beta\gamma}$ complex by Paul Sigler and collaborators (Lambright *et al.*, 1996, Nature 379:311). Transducin is the G protein that is coupled with rhodopsin in the retina, a key step of the visual signal transduction. The G_α–$G_{\beta\gamma}$ complex has many features that had not been seen in earlier structures. It is short-lived, whereas protease–inhibitor and antigen–antibody complexes are stable for hours or days; it displays large conformation changes, all essential to function. The December 1997 issue of *Current Opinion in Structural Biology* reviewed that structure and those of other complexes involved in signal transduction. Many more were determined afterwards, and they have been reviewed on a regular basis in several Sections of the journal, first of all *Folding and Binding*, *Catalysis and Regulation*, and *Macromolecular Assemblages*.

The new *Protein–protein interactions* Section of *Current Opinion in Structural Biology* will cover structural aspects as other Sections do, but it will pay particular attention to topics such as the specificity and affinity that Creighton and Kim mentioned in 1991. Binding affinities determine whether and when an interaction is functional and biologically relevant. In the present issue, we chose to make the relation between structure and affinity the theme of several reviews. The binding affinity of a protein for a ligand is quantified by the equilibrium dissociation constant of the complex K_d, and the related Gibbs free energy of dissociation ΔG_d. K_d is also related to the kinetic rate constants k_a and k_d of the association and dissociation reactions. Given the concentration of the components of a complex, the three constants determine whether the interaction actually takes place, and whether it is short-lived as in transducin, or stable as in protease–inhibitor complexes. Moreover, all these parameters depend on the environment,

especially the pH and ionic strength, and sometimes also on the presence of other molecules such as GDP and GTP as in the case of transducin. As a result, modeling the thermodynamics and kinetics of protein–protein interactions is a challenge, and the progress in the field is well worth following.

Affinity relies on non-covalent interactions; those that hold a protein–protein complex together are the same as those that define the fold in protein folding. As a consequence, many of the computational methods and force fields used to model protein–protein interactions derive from those developed for folding. An alternative approach is protein–protein docking, which has in many cases its own methods and scoring functions. In this issue, Moal, Moretti, Baker and Fernandez-Recio review the application of force fields and scoring functions to docking, binding affinity estimation and the prediction of mutation effects. Having performed an extensive benchmark of the methods on a set of affinity and kinetic data, Kastritis and Bonvin look for an 'Archimedean point' that offers a wider view of the question, and draw the conclusion that the description of binding affinity must incorporate effects of conformation changes and long-range interactions. The procedures and scoring functions tested by Moal *et al.* and by Kastritis and Bonvin include atomic-level and coarse-grain (residue-level) force fields, and empirical potentials designed for docking. Coarse-grain potentials and their application to model association processes are further described by Baaden and Marrink. They highlight the parallel developments in studying protein association in both solution and membrane environments, with coarse-grain models being combined with knowledge-based approaches predict and refine structural models of protein–protein complexes, applied to their *de novo* prediction and to the modeling of large scale protein aggregates.

The mechanism of the association reaction is the theme of the review by Zhou and Bates, who show that kinetic data on k_a can be accurately modeled when conformational changes are not rate-limiting, which is the case of about half of the systems that have been studied to date. Binding rate predictions are less reliable in the presence of conformational changes, but promising methods are being developed, and they may eventually deal with cases where the interaction involves flexible or disordered protein segments. This is the rule in protein–peptide interactions, represented in the Section by a review of London, Ravey and Schueler-Forman on peptide docking. The review also covers structure-based methods to model the binding affinity and specificity of peptides and linear motifs, which are implicated in many protein–protein interactions (up to 40% in higher eukaryotes).

In structural biology, modeling is a companion to experimental methods, and an important application is the computational design of novel interactions, which has met with remarkable success in recent years. Schreiber and Fleishman review experiments in which protein–protein complexes have been designed *ab initio*, and the designs shown to bind in the predicted way. In other experiments, binding rates have been manipulated by introducing mutations, also in conformity with the prediction. The link between modeling and experiment is particularly tight when it comes to studying very transient complexes or characterizing encounter complexes on the pathway to the native complex as reviewed by Schilder and Ubbink. Such complexes are 'invisible' to classical NMR and X-ray structural studies, and it is the advent of paramagnetic NMR that has made possible to finally reveal them. Schilder and Ubbink discuss recent developments in mapping the free energy landscapes of these complexes (in major part redox complexes). They describe the various experimental approaches (mainly kinetic and NMR experiments), which, combined with modeling, can give us insight into their properties. They highlight the role of hydrophobic interactions during encounter complex formation and the discovery of futile encounter complexes. Another class of challenging associations to study experimentally involves membrane and membrane associated complexes. Solid-state NMR offers insights into their structure and association in the membrane, as reviewed by Miao and Cross. The membrane environment plays an important role in defining and modulating protein structures and their associations. Solid-state NMR provides a unique means to study those associations in their native-like environment. While mostly oligomeric assemblies have been characterized to date, the field is developing rapidly. Gaining the ability to routinely study protein association in and on the membrane has a huge potential impact considering that about 50% of the current pharmaceutical targets are membrane proteins.

Genome sequencing and the subsequent advent of post-genomics in the last decade have marked the field of protein–protein interactions as it has the rest of biology. High-throughput genome-wide studies have relied on genetic (yeast two-hybrid or related) and analytic tools (affinity purification/mass spectroscopy) to identify hundreds of thousands of interactions between the proteins (or rather, the gene products) of yeast, human, and other organisms. The results are presented in the form of networks aiming to describe all the interactions that occur in a cell or an organism. They are stored in specialized databases, reviewed by Mosca, Pons, Céol, Valencia and Aloy. Mosca *et al.* also survey computational methods designed to exploit the data in them, and mention the role of the HUPO Protein Standard Initiative — Molecular Interaction. This consortium has defined a controlled vocabulary that applies to a wide variety of data on macromolecules and their interactions. Wodak, Vlasbom,

Turinsky and Pu review studies of the consistency and the quality of the data issued from high-throughput genome-wide studies. Different methods yield different sets of interactions, and the overlap is generally poor. Many of the discrepancies can be traced back to artifacts and biases proper to each method, but Wodak *et al.* suggest that the main culprit may be specificity: in any large dataset, some non-specific interactions must be present along with those that are biologically relevant. Returning to Creighton and Kim, we should add that affinity is also of concern. If the question is 'do two proteins interact?', the answer should depend on k_d and concentration, yet high-throughput data are interpreted in a yes or no manner, and this must introduce errors.

The last review in the Section deals with evolution. Protein–protein interactions govern the structure and the function of multi-component assemblies involved in all sorts of biological processes. They must be subject to strong Darwinian constraints. The interplay between the opposite requirements of conservation and diversification can be analyzed from many points of view. The one chosen by Zhang, Perica and Teichmann is the phylogeny of residue contact networks, a concept that is common to the interior of proteins and to the interfaces of protein–protein complexes. Evolution is a natural guideline through the complexity of life. We expect it to be a recurrent theme of a Section on *Protein–protein interactions*, as this subject is relevant to nearly all aspects of biology, and only a few are covered in this first issue.