

MEETING REPORT

CECAM workshop on Intrinsically Disordered Proteins: Connecting Computation, Physics and Biology. ETH Zürich September 2nd to 5th 2013

Heike Rösner¹, Elena Papaleo¹, Gitte W. Haxholm¹, Robert B. Best^{2,†}, Birthe B. Kragelund^{1,†} and Kresten Lindorff-Larsen^{1,†}

¹ Structural Biology and NMR Laboratory, Department of Biology, University of Copenhagen, Copenhagen, Denmark

² Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland, United States

[†] To whom correspondence should be addressed. RBB: robertbe@helix.nih.gov; BBK: bbk@bio.ku.dk; KL-L: lindorff@bio.ku.dk

Abstract

With the increasing need to integrate different areas of science in the study of intrinsically disordered proteins we arranged a meeting entitled “Intrinsically Disordered Proteins: Connecting Computation, Physics and Biology” in Zürich in September 2013. The aim of the meeting was to bring together scientists from a range of disciplines to provide a snapshot of the field, as well as to promote future interdisciplinary studies that link the fundamental physical and chemical properties of intrinsically disordered proteins with their biological function. A range of important topics were covered at the meeting including studies linking structural studies of intrinsically disordered proteins with their function, the effect of post-translational modifications, studies of folding-upon-binding, as well as presentation of a number of systems in which intrinsically disordered proteins play a central role in important biological processes. A recurring theme was how computation, including various forms of molecular simulations, can be integrated with experimental and theoretical studies to help understand the complex properties of intrinsically disordered proteins. With this Meeting Report we hope to give a brief overview of the inspiration obtained from presentations, discussions and conversation held at the workshop and point out possible future directions within the field of intrinsically disordered proteins.

Background

With its offspring in the world of bioinformatics and accentuated by biophysics it is now confidently established that a new class of proteins exists that does not conform to the classical structure-function paradigm. These intrinsically disordered proteins (IDPs) or regions of disorder interspersed folded domains (IDRs) comprise large fractions of cellular proteomes. The nature of structural disorder in these proteins can range from proteins that are highly extended with little secondary structure content, to proteins that at most times adopt more compact, molten globule-like structures. The pervasive structural heterogeneity and complex dynamical properties that characterize IDPs and IDRs make them challenging to study both experimentally and through simulation studies. Thus, in addition to using the toolbox of methods that have been developed to study folded, globular proteins, a number of theoretical and biophysical approaches have been developed to characterize the structural, dynamical and biophysical properties of IDPs. To gain a more detailed view of IDPs it is typically necessary to combine multiple different techniques including, for example, NMR spectroscopy, ensemble spectroscopy (FRET, fluorescence, CD, etc.), SAXS, single molecule spectroscopy, molecular simulations and purely theoretical considerations. Although a number of systems have now been characterized in substantial detail using these methods, only a few studies have so far linked these structural studies to the biological function of the protein. Indeed, we believe that the field is now at a point where computation, physics and biology can be used together to provide a more comprehensive understanding of IDPs. With IDPs emerging as increasingly central to molecular and cellular biology, uniting these disciplines is likely to lead to discoveries of new functionalities and hitherto unknown mechanisms.

To highlight the importance of bringing together *in vivo*, *in vitro* and *in silico* approaches to study IDPs and to provide researchers from different fields with a forum in which they could learn from one another, an interdisciplinary workshop on IDPs for exchanging results, ideas and hypotheses was recently organized with the support of the Centre Européen de Calcul Atomique et Moléculaire (CECAM). The international workshop was held in Zürich from the 2nd to 5th of September 2013 in the unique location of the Swiss Federal Institute of Technology of Zürich (ETH Zürich). Overlooking Lake Zürich (Fig. 1), we enjoyed three-and-a-half days of intense and inspiring exchange of largely unpublished results, where a broad range of scientists in the field presented their work. The workshop brought together theoreticians and experimentalists, but also gathered different disciplines, notably physics, chemistry and biology. This interdisciplinary exchange of scientific competence and perspectives proved to be a very fruitful source of inspiration for all the participants. A total of 21 invited lectures and 10 selected oral presentations were given during the workshop. Furthermore, more than 25 poster presentations enriched the program through two poster sessions, covering various aspects of IDP characterisation, single-molecule studies, folding-upon-binding, simulations of IDPs and IDRs, bioinformatics analyses, and biological data. This report aims to summarize the essence of the workshop, the topics of greatest debate and the scientific challenges that were revealed. While this is not a detailed summary or a chronological rundown we hope this report will inspire those who did not attend and remind those of you who were present. The intention is to sum up the inspiration obtained from presentations, discussions and conversation held at the workshop and point out possible future directions within the field of IDPs.

Linking experimental and computational studies on the binding of IDPs to their targets

The mechanisms by which IDPs and IDRs bind to their partner proteins remain incompletely understood. Certain systems may gain structure before binding whereas others would only develop structure after or during complex formation (Jane Clarke, Thomas Kiefhaber, Kaare Teilum). Notable differences between IDP binding and protein folding were observed – firstly the importance of non-

native interactions in accelerating binding rates, and secondly energetic “frustration” caused by local barriers in the energy landscape for binding reactions. Throughout the workshop it was highlighted that no single, general mechanism will likely represent all cases and that the various systems under study could utilize different routes. Rather, it was proposed that the nature of the systems, the means of their manipulation and the differential use of experimental conditions would determine which route is the more prominent for a particular system. It seems obvious that a combined folding/binding event cannot easily be described by a two-state mechanism and although the process is complex, e.g. involving at least an encounter complex, it is often only possible to observe a single phase in kinetic experiments. Thus, as a field we are still challenged by the necessity to develop experimental procedures and methods for analyses that will allow us to understand the mechanisms underlying recognition of IDPs.

One issue addressed at this workshop therefore centred around which experimental approaches would be suitable for obtaining insight into the determinants of the rate-limiting step in IDP/partner-interactions. Methods designed for classical protein folding studies as well as the need for new “out-of-the-box” approaches were discussed. Indeed, mutant design will respectfully have to take the special nature of IDPs into account. What seems to be a small effect on binding may have large effects both locally and globally on the conformational ensemble including effects on transient secondary structure and compactness (see below). Thus, mutagenesis studies applied to IDPs run the risk of causing large “ground state” effects complicating the interpretation of protein engineering experiments. Care must thus be taken when performing and analysing mutational studies of the binding of IDPs to their targets.

In general, the number of protein engineering studies of IDP binding kinetics is still relatively small and increased efforts toward collecting more data are desirable. Computational studies might be an ideal *in silico* pre-screening routine for optimal mutant design (Jeetain Mittal, Athi Narayanan Naganathan, Lennart Nilsson), addressing the possible effects induced by a mutation on the conformational ensembles. The complementarity between computation and biophysics was a strong and continuous element during the workshop. In the future, a closer interplay between experiments and simulations will provide an essential bridge to the development of new approaches in IDP investigations.

Defining the physical-chemical determinants for the function and structure of IDPs was a second widely discussed topic at the workshop. It crystallised that in addition to the role of transient structure, there could be several alternative deterministic traits to binding and behaviour of IDPs. Especially, the distribution of charges seems crucial. Examples of computational and theoretical studies (Rohit Pappu), *in vitro* single-molecule experiments (Ben Schuler) and NMR experiments (Frans Mulder) were presented showing individually how the number and distribution of charged residues would define the structure, compactness, dynamics and reactivity of IDPs. However, the functional importance of pre-formed transient structure, charge distribution and hydrophobic interactions as well as any possible interrelation between these remains an unsolved and open question. Again, due to the scarcity of the number of systems studied it is too early to be establishing general conclusions, which would call for an expansion of systems under study.

Structural biology of IDPs coming of age

Due to their substantial differences from well-folded proteins, IDPs have challenged the current methods in structural biology. At the workshop, structural characterisations of IDPs were presented, employing different biophysical techniques including NMR spectroscopy, SAXS and FRET analyses (e.g. Jane Dyson, Pau Bernado, Peter Tompa, Gitte Haxholm, Peter Wright, Richard Kriwacki, David Elzizer). Also here, simulations are needed to integrate those data to help study the structural ensemble of IDPs. Nevertheless, it was pointed out that atomistic molecular simulations alone are still not able to

routinely describe the conformational ensemble of IDPs. This is most likely still due to intrinsic limitations in classical force fields that may cause over-compactation or give rise to erroneous secondary structure content. A solution to overcome this risk can be to synergistically integrate classical force fields with experimental data such as that available from NMR, SAXS or FRET experiments (Jane Allison, Elena Papaleo), and in the long run to use these data to improve the force fields themselves (Jeetain Mittal). Such efforts will be aided by the availability of a new database for structural ensembles and experimental data (Peter Tompa).

While *in silico* studies and *in vitro* experiments that aim to reproduce biologically relevant conditions largely remain the only way to investigate the detailed intrinsic properties of proteins, such approaches often run short on a critical factor: in their natural milieu, proteins are surrounded by a wide range of other molecules. These molecules display a wide range of chemical properties and the resulting “crowding effects” are thus a complex combination of chemical and physical effects (David Eliezer, Philipp Selenko, He Fang). High-resolution in-cell NMR may provide a powerful tool to study structural properties of IDPs in their native cellular environments (Philipp Selenko).

Single-molecule spectroscopy also holds great promises for characterisation of the structural and dynamical properties of IDPs. At the workshop, Ben Schuler and Hagen Hofmann presented applications of single molecule FRET integrated with nanosecond correlation spectroscopy and microfluidic mixing. Using these approaches, the structure and dynamics of IDPs can be studied on different timescales and even in heterogeneous environments.

In summary, we now have available a substantial set of experimental and computational tools that can be used to study the structural and dynamical properties of IDPs, both in their free states and in their complexes. Together with kinetic experiments that probe the dynamical properties and mechanisms of binding, we are beginning to build a consistent and generally applicable toolbox to link IDP structure, dynamics and biological function.

Tuning affinity through ensemble redistribution

It was noted by several speakers that IDPs could be compacted, perhaps down to the dimensions of folded proteins, without formation of any substantial amount of secondary or specific tertiary structure. Based on simulation, a phase diagram can be constructed where the amino acid composition decides whether a protein adopts a globular, coil-like or swollen coil-like conformation, and the distribution of charges appears to be an important determinant of the behaviour of the chain. In lively discussion, it was speculated whether the cell actively use post-translational modifications (PTMs), such as phosphorylation or acetylation, to influence the pattern of charged residues and thereby cause phase transitions within the ensembles (Rohit Pappu, Yaakov Levy). Several examples from experimental studies gave a first idea about the regulatory power of PTMs in increasing or blocking target interactions and thereby tuning e.g. the affinity. Such affinity modulation would affect e.g. membrane binding of IDPs (David Eliezer, Gitte Haxholm) or the sensitivity for allosteric modulations (Richard Kriwacki). What in particular remains to be addressed is whether it is the PTM per se, its precise sequential location, a changed charge-pattern, changes in formation of transient structure, or a combination of these factors that are responsible for the altered target interaction.

Another interesting issue addressed was the allosteric effect observed for some IDPs (Vincent Hilser, Xavier Salvatella). Disordered proteins are often identified as hubs in signalling networks and exhibit switch-like behaviour. It was suggested that in the energy landscape, some IDPs reside close to the transition from disorder to order, which would make them ideally suited to respond allosterically to small changes either chemically (e.g. PTMs) or functionally (e.g. by binding). Disordered regions in signalling proteins might allosterically modulate the function of folded domains through a broad array

of regulatory strategies to propagate the signal over long distances. Thus, IDPs in hubs may act both agonistically and antagonistically depending on the partners interacting with the hub. Remarkably, folded proteins similarly can also undergo a regulated unfolding as a mechanism of signalling (Richard Kriwacki).

Generally, the relationships between structure, conformational dynamics, PTMs and function in IDPs have only been brought to the test *in vivo* in a few systems. Furthermore, the full effects of modulations by PTMs as well as allosteric modulations in hub scenarios need to be elevated in more complex systems such that insight into the global behaviour of the system can be obtained.

Broad mechanistic properties: new meeting points with biology

As mentioned above, one of the prime functions ascribed to IDPs and IDRs is their involvement in signal transduction and regulation e.g. acting as hubs in signalling networks and in communication cascades. Here, specific linear peptide motifs are of great importance for directing protein-protein interactions. Several speakers argued, however, that the functional portfolio of IDPs is not only restricted to the more passive role of transmitting signals via acting as binding partners with sequence specific motifs, but can also be expanded to include IDPs that more directly perform active functions for a cell.

IDRs are highly abundant in DNA-binding proteins. Through clusters of charged residues they can modulate the affinity for DNA binding, ease the movement along DNA strands, and promote inter-strand crossings. IDRs can modulate both one-dimensional sliding along DNA and help optimize the search in three dimensions. The speed of motif search along a DNA molecule by transcription factors, for example, is highly influenced by its disordered domains (Yaakov Levy). Instead of the protein being restricted to reside on one DNA molecule, the IDR further facilitates jumping between two DNA strands in close proximity. The extent to which the distribution of charges and the modulation of this by PTMs, contribute to these more mechanical processes is an area for further studies.

IDPs have also been identified in a wide range of organisms as having chaperone-like activities e.g. in the plant LEA proteins. The chaperone mechanism of an IDP differs from those of traditional chaperones by e.g. acting as molecular shields against dysfunctional protein-protein associations by means of weak and transient interactions (Alan Tunnacliffe).

IDRs can also act as modifiers for protein turnover rates by affecting proteasomal degradation. The variations within IDRs in proteins can tune protein *in vivo* half-life, with a direct relationship between the half-life of the protein and the presence and characteristics of IDRs (Robin van der Lee).

Together, these and other examples presented at the workshop showed that IDPs can take an active part in essential cellular tasks like protein turnover, stress tolerance and protein mobility. As such, in the quest for further linking the structural and dynamical properties of IDPs with their functional properties, the field needs to take a wide range of potential functional mechanisms into account.

IDPs and promiscuity: a friend in evolution but a foe for drugability?

The transition from single-cell to multicellular organisms requires cellular differentiation and intercellular signalling and this transition was interestingly hypothesised by Keith Dunker to be linked to an increase in the IDP portfolio of an organism. The binding promiscuity of IDPs and the relatively few alterations needed to modify binding preferences allows for a comparatively fast evolution and adaption of signalling pathways involving IDPs compared to the evolutionary timescale of folded proteins (Keith Dunker, Anthony Wright). IDPs may be well described as being *anti-fragile*, a term put forward by Steven Metallo, meaning that the organisms benefit from the property of the IDP to be

acceptably mutated at a higher rate and adapt faster to environmental changes.

Low binding affinities, overlapping binding sites and thus an inherent multi-specificity, result in a certain biological dosage sensitivity of IDPs. When IDPs get — artificially or accidentally — overexpressed, they are prone to engage in more promiscuous interactions resulting in pathological conditions, perhaps more severe to the organism than the overexpression of a folded protein would usually be. Thus, besides linking mutation-prone regions in IDPs to disease states, also the level of the individual IDPs need to be taken into account when considering possible pathological conditions.

As a result of the above considerations, one price IDPs have to pay for their promiscuity is that they require tight regulation mechanisms. Periodically changing protein levels have been suggested to be a regulatory mechanism for a number of IDPs (Philipp Selenko, Yosef Shaul). Also, while IDPs alone are at high risk of degradation as soon as they are synthesized, so-called 'nanny proteins' of IDPs may assume a protective role in a biological context, and can thus establish and support oscillating levels of IDPs (Yosef Shaul).

Despite promiscuous binding sites and lower than average binding affinities (when compared to folded proteins), IDPs remain interesting drug targets (Steven Metallo). Thus, while certainly a difficult problem to tackle, IDPs seem to be drugable, and binding hot spots may possibly be identified similarly to in complexes of folded protein. Further, the tight biological regulation of IDPs suggests that relatively minor modulations by drug molecules might be sufficient to elicit a biological response. With a general low affinity there is room for optimization within a whole new drug-space, and the development of agonists and antagonist now seems possible at least on the exploratory level. Future studies, however, need to address how to obtain modulators with sufficient affinity, specificity and drug residence time. In addition to the traditional mechanisms utilized to modulate folded proteins, other mechanisms to interfere with IDP function may be possible targets. These may include controlling the protein levels, blocking or locking binding interactions or otherwise interfering with folding-upon-binding reactions.

Closing words

A great workshop always leaves open questions. Despite major progress in the field of structural biology of IDPs, there is still a strong need for improvements and development of special experimental and computational tools to achieve a comprehensive understanding of the general biophysical principles that underlie IDP function. In particular, the integration of *in silico* and experimental studies is greatly encouraged, as it would enable a better atomic-level understanding of protein disorder.

Functional disorder is clearly a fundamental concept, and its discovery has inspired theoretical physics, polymer chemistry, and various simplified models to be used to capture a wide range of properties of IDPs and IDRs. It is, however, also obvious that the complexity of biological molecules calls for descriptions of their action within a true biological context. How important is the role of the electrostatic patterns and hydrophobic interactions *in vivo*, where other molecules may bind and either screen out or enhance such interactions? How do macromolecular association and assembly influence the biophysical properties of IDPs? To what extent does the strong spatial inhomogeneity (protein complexes, cellular compartmentalization, etc.) have to be taken into account when understanding IDP function? What is a minimal set of properties that is necessary to describe IDP dynamics so that they closely relate to the actual *in vivo* conditions? Can we predict whether the protein acts through folding-upon-binding, fuzzy complexes or linkers from the compactness properties and charge patterns? How do we investigate molecular switches in a biological context, e.g. the function of PTMs, and not only in binary scenarios? These are all essential questions open to the IDP field, raised by the great results presented at the workshop, which will help to inspire future directions.

Acknowledgements and prizes

In conclusion, the workshop successfully brought together scientists from a range of disciplines to discuss the physics, biochemistry and biology of IDPs. Great insightful posters were displayed during the entire conference, and Cambridge University Press provided prizes for the six most outstanding posters. The winners of the poster prizes were Sarah Rauscher and Nathaniel Stanley for their posters in the category “Simulation” and Eva de Mol, Joseph Rogers, Andrea Soranno and Sigrid Milles in the category “Experimental”. The whole workshop could not have taken place without logistic help and funding from CECAM as well as additional funding received from BASF, Novartis, Roche, Merck Serono, Syngenta and Cortecnet. Their support in bringing together computational and experimental sciences was truly important.



Figure 1. A spectacular view of lake Zürich from the ETH as captured by Gaetano Invernizzi during the CECAM workshop on intrinsically disordered proteins in computation, physics and biology held in Zürich in September 2013.