Editorial Overview

Integrative Biophysics: Protein Interaction and Disorder

Over the last decades, individual biophysical techniques and their complementary data integrations have been pivotal to advance functional and disease-driven understanding of biomolecules and their dynamics. On one hand, high-resolution biophysical methods such as X-ray crystallography, nuclear magnetic resonance (NMR, and cryo-electron microscopy (cryo-EM) have contributed enormously to our ability to decipher the atomistic details of biomolecules and their assemblies. On the other hand, when such high-resolution information is unattainable, sparse structural data often come from alternative approaches of small-angle scattering (SAXS/SANS) and Förster resonance energy transfer (FRET) as well as mass spectrometry based techniques such as hydrogen/deuterium exchange, crosslinking, and footprinting [1]. In both cases, computations and molecular modeling are critical to either interpret high-resolution raw data for structural elucidation or join different sources of data to better understand protein dynamics and function in an integrative fashion.

Thanks to genome-sequencing initiatives and large-scale interactome studies, it is becoming clear that all organisms demonstrate the intricate protein "sociology" required for biological function [2]. Therefore, it is necessary to obtain the three-dimensional arrangements of these critical biomolecular assemblies to understand function. A key advantage of integrative approaches is the enhanced capacity to elucidate details of these assemblies that perform every essential biological process in a concerted way. Another advantage is to go beyond static molecular pictures and establish a connection between structural dynamics and function. It is well established that biomolecular dynamics offers an additional dimension to available three-dimensional structures, and that it contributes (either *via* spatial flexibility or temporal coupling of motions at different scales) to most biological functions [3]. In this context, intrinsically disordered proteins (IDPs) represent a very challenging category [4], because their structure and dynamics are nontrivial to be disentangled by any single technique *per se*. As such, integrative approaches, which synergistically combine information from different sources, are becoming a powerful strategy to derive structure–dynamics–function relationships [5,6,7].

This Special Issue contains nine review articles and seven original articles that discuss some of the technical advances in integrative biophysics and recent applications for protein structure, dynamics, and interaction. These studies can be loosely divided into four categories. The first concerns the assembling of high-resolution structures of individual components that are obtained from crystallography or NMR so they fit the cryo-EM density or overarching structural data of their complexes, with the goal of creating more accurate pictures of biomolecular systems. Reviews by Srivastava *et al*. [8] and Koukos and Bonvin [9] provide a comprehensive survey on a variety of experimental information that can be used for integrative/hybrid modeling as well as computational approaches available for data interpretation, including the use of low-resolution cryo-EM and X-ray free-electron laser. Chance *et al*. [10] provide a detailed description on the capacity of hydroxyl radical protein footprinting to deliver residue-specific information of solvent exposure for these amino acids on protein-interaction surfaces and on its combination with SAXS for studying protein assemblies and IDPs. The original research article by Pond *et al*. [11] demonstrates how a combination of solution NMR, neutron reflectometry, and molecular dynamics simulations can provide a state-of-the-art picture about how different domains are spatially arranged within a multi-domain Src-family kinase Hck in the physiological environment of lipid bilayers. These are just a few examples to illustrate how such data integration enable the assembly of individual sub-components into their functional complexes.

The second category addresses the synergistic integration of different sources of sparse or low-resolution structural information that can be individually and routinely acquired. The nature of the information is diverse, including intramolecular distances, overall shape, solvent accessibility, and large-scale motions. The original research article by Tsytlonok *et al*. [12] examines the binding interactions between an IDP p27 and cyclin-dependent kinase Cdk2/Cyclin A, by combining time-resolved single-molecule (smFRET) and replica exchange simulations to describe the ensemble dynamics and conformational switching between ensembles. The review by Clark *et al*. [13] describes an up-to-date view about the role of water as a good solvent in unfolded states of folded globular proteins, summarizing a series of theoretical and experimental studies including SAXS and FRET. In addition to resolving biomolecular structures at high-resolution,

NMR is a very rich source of observables reporting on dynamics and interaction. This is exemplified by two original research articles. Capraro *et al*. [14] use various NMR chemical shifts to assess conformational and allosteric changes in a knotted methyltransferase, while Alderson *et al*. [15] combine backbone chemical shifts, ^{15}N relaxation rates, and ^{1}H–^{15}N residual dipolar couplings to investigate the connection between dissociation and partial unfolding of a chaperone. Combination of structural and dynamic NMR data with complementary biophysical techniques is also discussed in the following two reviews. Delhommel *et al*. [16] compile recent examples of the integration of local NMR data (reporting on local information) with SAXS/SANS techniques, with a special focus on the use of neutron scattering and contrast variation to study RNA complexes. Tang and Gong [17] provide a critical review on how to simultaneously interpret distance measurements of both smFRET and NMR data from the perspective of methodological integration.

Third is the integration of all this structural/dynamic information with different levels of resolution, which requires highly advanced and flexible computational platforms with robust statistical tests to exploit their complementarity or compare with other existing data for cross-validation. Danielsson *et al*. [18] describe how a combination of hydrogen/deuterium exchange, backbone NMR chemical shifts, and molecular dynamics simulations revealed residue-specific knowledge about protein dynamics spanning multiple time-scales in a pierced lasso protein leptin. Another example is the original research article by Girodat *et al*. [19] describing the multi-level computational studies for cryo-EM reconstructed structures to achieve an energy landscape mapping of a tRNA selection process, consistent with smFRET findings. Two excellent reviews are also included with regard to computational aspects of biophysical methods. Garcia de la Torre and Hernandez Cifre [20] review the power of hydrodynamic techniques and solution properties such as translational or rotational coefficients as well as computational modeling for their structural interpretation. Grawert and Svergun [21] describe the latest advances in experimental SAXS data acquisition and discuss related computational modeling for large biomolecular assemblies. Both reviews offer the use of such techniques for studies of IDPs through tailored data-driven computations.

The final category of the articles in this Special Issue covers the (re)-examination of structural studies of IDPs using state-of-the-art NMR approaches. The original article by Mateos *et al*. [22] characterizes the structural disorder of a 200-residue protein osteopontin and the structural role of its proline residues in particular, through the exquisite resolution of ^{13}C-detected NMR experiments developed by their group. An exemplary review by Spreitzer *et al*. [23] describes how solvent paramagnetic relaxation enhancement experiments can map solvent exposed regions in proteins and biomolecular complexes, which can be especially useful for studying IDPs' residue-specific solvent exposure to unveil partially structured elements or interaction surfaces in fuzzy complexes.

The field of *integrative biophysics* has emerged to tackle historically challenging topics such as multi-protein assemblies and IDPs, by harvesting the collective power of technological advances made in various biophysical techniques. While general data-integration strategies and specific data-acquisition protocols are being developed [24], the articles in this Special Issue represent a collection of individual and group efforts undertaken with the goal of answering important questions about protein dynamics, interaction, and disorder. As we are entering a new era of biophysics and structural biology, by placing a focus on how vast amounts of biophysical data can be integrated, this Special Issue is poised to proactively stir up the debate with regard to data-integration strategies and ultimately accelerate technological development for the benefit of the biophysics community at large.

References

[1] C.V. Robinson, A. Sali, W. Baumeister, The molecular sociology of the cell, Nature 450 (2007) 973–982.
[2] R.B. Russell, P. Aloy, Targeting and tinkering with interaction networks, Nat. Chem. Biol. 4 (2008) 666–673.
[3] H. Frauenfelder, S.G. Sligar, P.G. Wolynes, The energy landscapes and motions of proteins, Science 254 (1991) 1598–1603.
[4] V.N. Uversky, Natively unfolded proteins: a point where biology waits for physics, Protein Sci. 11 (2002) 739–756.
[5] T.N. Cordeiro, N. Sibille, P. Germain, P. Barthe, A. Boulahtouf, F. Allemand, R. Bailly, V. Vivat, et al., Interplay of protein disorder in retinoic acid receptor heterodimer and its corepressor regulates gene expression, Structure 27 (1270–1285) (2019), e6.
[6] M.A. Lietzow, M. Jamin, H.J. Dyson, P.E. Wright, Mapping long-range contacts in a highly unfolded protein, J. Mol. Biol. 322 (2002) 655–662.
[7] Y. Peng, S. Cao, J. Kiselar, X. Xiao, Z. Du, A. Hsieh, S. Ko, Y. Chen, et al., A metastable contact and structural disorder in the estrogen receptor transactivation domain, Structure 27 (229–240) (2019), e4.
[8] A. Srivastava, S.P. Tiwari, O. Miyashita, F. Tama, Integrative/hybrid modeling approaches for studying biomolecules, J. Mol. Biol. 432 (9) (2020) 2846–2860.
[9] P.I. Koukos, A. Bonvin, Integrative modelling of biomolecular complexes, J. Mol. Biol. 432 (9) (2020) 2861–2881.
[10] M.R. Chance, E.R. Farquhar, S. Yang, D.T. Lodowski, J. Kiselar, Protein footprinting: auxiliary engine to power the structural biology revolution, J. Mol. Biol. 432 (9) (2020) 2973–2984.
[11] M.P. Pond, R. Eells, B.W. Treece, F. Heinrich, M. Losche, B. Roux, Membrane anchoring of Hck kinase via the intrinsically disordered SH4-U and length scale

associated with subcellular localization, J. Mol. Biol. 432 (9) (2020) 2985–2997.
[12] M. Tsytlonok, K. Hemmen, G. Hamilton, N. Kolimi, S. Felekyan, C.A.M. Seidel, P. Tompa, H. Sanabria, Specific conformational dynamics and expansion underpin a multi-step mechanism for specific binding of p27 with Cdk2/Cyclin A, J. Mol. Biol. 342 (9) (2020) 2998–3017.
[13] P.L. Clark, K.W. Plaxco, T.R. Sosnick, Water as a good solvent for unfolded proteins: folding and collapse are fundamentally different, J. Mol. Biol. 432 (9) (2020) 2882–2889.
[14] D.T. Capraro, D.J. Burban, P.A. Jennings, Unraveling Allostery in a knotted minimal methyltransferase by NMR spectroscopy, J. Mol. Biol. 432 (9) (2020) 3018–3032.
[15] T.R. Alderson, J. Ying, A. Bax, J.L.P. Benesch, A.J. Baldwin, Conditional disorder in small heat-shock proteins, J. Mol. Biol. 432 (9) (2020) 3033–3049.
[16] F. Delhommel, F. Gabel, M. Sattler, Current approaches for integrating solution NMR spectroscopy and small angle scattering to study the structure and dynamics of biomolecular complexes, J. Mol. Biol. 432 (9) (2020) 2890–2912.
[17] C. Tang, Z. Gong, Integrating non-NMR distance restraints to augment NMR depiction of protein structure and dynamics, J. Mol. Biol. 432 (9) (2020) 2913–2929.
[18] J. Danielsson, J.K. Noel, J.M. Simien, B.M. Duggan, M. Oliveberg, J.N. Onuchic, P.A. Jennings, E. Haglund, The pierced lasso topology leptin has a bolt on dynamic domain composed by the disordered loops I and III, J. Mol. Biol. 432 (9) (2020) 3050–3063.
[19] D. Girodat, S.C. Blanchard, H.J. Wieden, K.Y. Sanbonmatsu, Elongation factor Tu switch I element is a gate for aminoacyl-tRNA selection, J. Mol. Biol. 432 (2) (2020) 3064–3077.
[20] J. Garcia de la Torre, J.G. Hernandez Cifre, Hydrodynamic properties of biomacromolecules and macromolecular complexes: concepts and methods. A tutorial mini-review, J. Mol. Biol 432 (9) (2020) 2930–2948.
[21] T.W. Grawert, D.I. Svergun, Structural modeling using solution small-angle X-ray scattering (SAXS), J. Mol. Biol. 432 (9) (2020) 3078–3092.
[22] B. Mateos, C. Conrad-Billroth, M. Schiavina, A. Beier, G. Kontaxis, R. Konrat, I.C. Felli, R. Pierattelli, The ambivalent role of proline residues in an intrinsically disordered protein: from disorder promoters to compaction facilitators, J. Mol. Biol. 342 (9) (2020) 3093–3111.
[23] E. Spreitzer, S. Usluer, T. Madl, Probing surface accessibility in the dynamic meshwork of protein–protein interactions, J. Mol. Biol. (2020).
[24] H.M. Berman, P.D. Adams, A.A. Bonvin, S.K. Burley, B. Carragher, W. Chiu, F. DiMaio, T.E. Ferrin, et al., Federating structural models and data: outcomes from a workshop on archiving integrative structures, Structure 27 (2019) 1745–1759.

Sichun Yang
Center for Proteomics and Department of Nutrition, School of Medicine, Case Western Reserve University, Cleveland, OH 44106, USA
E-mail address: sichun.yang@case.edu

Pau Bernadó
Nuclear Magnetic Resonance (NMR), INSERM, CNRS, Université de Montpellier, 29, rue de Navacelles, 34090 Montpellier, France
E-mail address: pau.bernado@cbs.cnrs.fr