'*Era of Integrative Structural Biology*'

'There is freedom waiting for you,
On the breezes of the sky,
And you ask "What if I fall?"
Oh but my darling,
What if you fly?'
© Erin Hanson

In 1976, the Protein Data Bank (PDB) was launched containing 13 experimentally determined structures. The number of structures has rapidly increased since then, reaching over 125,000 entries by the end of 2016 (www.pdb.org). The last 40 years can thus be described as the '*Era of Structural Biology'*.

Methods of structural biology have become very powerful, providing valuable mechanistic insights into basic biological processes, which could be translated into medical applications. Examples of such knowledge include the structural models of the DNA double helix, RNA polymerase, spliceosome, ribosome, and nuclear pore complex. The reconstitution of such and other assemblies have led to increased understanding of life processes.

The success of structural biology methods depends heavily on the technological advancements in instrumentation, which have made it possible to determine the structure of large molecular assemblies and to increase the quality of the structures determined for smaller molecules. An excellent example is the development in the field of cryo-electron microscopy (CryoEM), which took place over the course of the last few years[1].

Structural biology no longer represents '*dark magic*' for many scientists, which is a very important achievement. Nowadays more than ever, nuclear magnetic resonance (NMR), X-ray crystallography, and CryoEM are more user-friendly, automated, and accessible. Governments and private funding bodies established large high-throughput facilities that are accessible to a broad range of users based on grant application schemes. Therefore, from the technical and financial point of view, structure determination is easily accessible for scientists from areas outside structural biology. In fact, the inclusion and consideration of structural data in biological research has become an unspoken requirement for publication in highly ranked journals.

Structural genomics and other large-scale initiatives solved the structures of many proteins and protein domains from various genomes. It seems that all low-hanging fruits have been picked, and the discovery of novel folds is becoming a rare event. Because of technological advances, the level of complexity of the studied biological systems has increased significantly over the last years. In some cases, the level of complexity is so high that many phenomena cannot be addressed by conventional structural biology methods alone (NMR, X-ray crystallography, and CryoEM). While conventional methods work well for structured molecules, the presence of flexibility, conformational changes, and formation of transient complexes with weak affinities remain the greatest challenge in structural studies of macromolecules. An additional limitation of contemporary methods is related to the conditions under which macromolecules are studied, involving either concentrated solutions, crystals, or vitrified ice, which differ significantly from the conditions in the living cell. Current methods fail to take into account the complexity of the cell, the relevance of macromolecular crowding, and the co-existence of other phases such as aggregates, hydrogels, and glasses. Therefore, the determination of molecular structure directly inside the living cell remains a challenge.

The scientific community has always been aware of the above-mentioned limitations, which has boosted the

development of many complementary techniques. Consequently, the '*Era of Integrative Structural Biology*' became of age, and this is a beautiful era[2,3].

What is the integrative approach, and what does it offer us? For the purpose of this preface, I would divide integrative structural biology into two categories: first, determination of a structural model by integrating direct restraints from several high- and low-resolution biophysical methods [2]; second, follow-up study of the determined model, involving structure validation and interpretation, using a complementary or advanced variation of the original investigative method.

Methods of the first category include NMR, X-ray crystallography, CryoEM, small-angle scattering, chemical crosslinking followed by mass-spectrometry, and computational approaches, which are used to combine all the information. The individual constraints gathered using different methods provide restraints on the conformation, position, and orientation of the components in a macromolecular assembly. Combining all restraints improves the accuracy, precision, and completeness of a model, especially when limited high-resolution structural data on the entire assembly are available[2].

The second category includes experimental methods that are not directly used for three-dimensional model determination, but are invaluable for verifying the model and characterizing its behavior/dynamics. The verification of the model includes determination of binding affinity (*e.g.*, by fluorescence anisotropy, isothermal titration calorimetry, microscale thermophoresis, electromobility shift assay), confirmation of enzymatic activity by assays specific for the studied system, and validation of stoichiometry (*e.g.*, analytical ultracentrifugation, size exclusion chromatography, small-angle scattering, light scattering). Methods characterizing the dynamics of the system can be very diverse and usually include advanced biophysical methods such as solution NMR investigations (*e.g.*, NMR relaxation experiments, residual dipolar couplings), and single-molecule methods (*e.g.*, molecular tweezers, single-molecule fluorescence resonance energy transfer). Applying a combination of methods allows moving from the canonical, static understanding of the structural model to a more dynamic understanding of the process in action.

The availability of a wide range of established complementary methods indicates that we have reached an '*Era beyond the Technique*'. Developments are so fast and so broad that investigators are no longer bound to a technique, but become bound to the research question. In this context, it might be no longer worth to define ourselves by our training, *e.g.*, structural biologists, and it is rather necessary to train ourselves in how to ask good questions and how to communicate them to colleagues in order to find solutions to these questions. There are now powerful tools to aid us in our quest. The future is bright.

Acknowledgement

I would like to thank R. Stefl, K. Kubicek, J. Novacek, T. Shaikh, and C. M. Ionescu for the critical reading of the text, discussion and suggestions.

References

1. Callaway, E. The revolution will not be crystallized: a new method sweeps through structural biology. *Nat. News* **525,** 172 (2015).
2. Ward, A. B., Sali, A. & Wilson, I. A. Integrative Structural Biology. *Science* **339,** 913–915 (2013).
3. van den Bedem, H. & Fraser, J. S. Integrative, dynamic structural biology at atomic resolution—it's about time. *Nat. Methods* **12,** 307–318 (2015).