EDITORIAL

Integrative methods in structural biology

Rob Kaptein[1] · Gerhard Wagner[2]

During the last decade structural biologists have addressed ever larger protein complexes and biological machines. As these projects became more ambitious, it was necessary to use multiple experimental and computational methods, thereby taking advantage of the different strengths of each method. This is now known as using "integrative" (or "hybrid") approaches. Early examples are the combined use of NMR and SAXS in structure determination of proteins (Grishaev et al. 2005) and of solid-state NMR, cryo-EM, and Rosetta modeling in the study of the Type 3 secretion needle by Loquet et al. (2012). For large protein complexes the use of X-ray crystallography with cryo-EM has become popular (cf. Lander et al. 2012). It is noteworthy that mass spectrometry (MS) methods are now also part of the integrative toolkit. These methods include hydrogen–deuterium exchange MS (HDX-MS) and cross-linking MS (c.f. Liko et al. 2016). To deal with these widely different data from multiple sources, it is crucial to have proper modeling software to handle the various restraints and analyze the results in a statistically meaningful way (cf. Rieping et al. 2005; Alber et al. 2008).

It is remarkable that NMR often plays an important role in these hybrid approaches. Of course, high-resolution NMR is unique in that it provides atomic-level structural information for biomolecules *in solution*. In addition, dynamic behavior can be studied for an extremely wide range of timescales. For solid-state NMR there is no size limit, provided suitable isotope labeling can be used to beat the spectral complexity problem. For the present Special Issue, we have taken integrative methods in the sense of NMR plus anything else.

Luchinat and coworkers present a new version of REFMAC-NMR, a joint structure refinement program based on NMR and X-ray data (Carlon et al. 2019). For multidomain proteins they include constraints between alignment tensors for residual dipolar couplings (RDC's) and between susceptibility tensors for pseudo-contact shifts (PCS's). The authors show for several multidomain proteins and biomolecular complexes that by using joint X-ray/NMR refinement with these new features either the structural quality is improved, or the differences between crystal and solution structures can be defined more precisely.

Two papers deal with secretion pili, i.e. the bacterial fibers that are involved in transport of proteins to target cells or to the extracellular milieu. Lange and collaborators (Hwang et al. 2019) have used ^{1}H-detected solid-state NMR to study amide H/D exchange in pili from uropathogenic *Escherichia coli* bacteria. The H/D exchange rates were also modeled by MD simulations. Comparison of experimental and modeled data gave insight in the structure and packing of Type 1 secretion pili.

The Nilges group (Bardiaux et al. 2019) has studied a Type 2 pilus using a combination of several methods (solution NMR, cryo-EM, and HDX-MS). The structure of pilin, the monomeric subunit of the pilus, was solved by NMR and used to interpret the EM map of the assembled pilus. With dynamics information from normal-mode analysis this provided a detailed picture of the structure and assembly of Type 2 pili with implications for their function.

For the characterization of intrinsically disordered proteins (IDP's) or disordered regions of proteins NMR has been the method of choice, since information from X-ray crystallography or EM is often not meaningful for these proteins. Madl and coworkers (Hartlmüller et al. 2019) describe a method to study IDP's using solvent paramagnetic relaxation enhancement (sPRE). In this approach dissolved paramagnetic molecules are used rather than tags attached to the protein. A benchmark for sPRE's was introduced by calculating the average sPRE for the central residue of all 5-residue peptides of the protein, flanked on each side by 3 alanines. Lower values for the experimental sPRE's compared to the

✉ Rob Kaptein
  r.kaptein@uu.nl

[1] Bijvoet Centre, Utrecht University, 3584 CH Utrecht, The Netherlands

[2] Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA

back-calculated ones were interpreted in terms of regions of long-range interactions. The method was applied to the proteins α-synuclein, FOXO4, and p53.

The mechanism of T cell signaling has been addressed by Reinherz, Wagner, and coworkers (Mallis et al. 2019). Non-self peptides (p) are displayed on the cell surface by the major histocompatibility complex (MHC) for recognition by the αβT-cell antigen receptor (αβTCR). The αβTCR—pMHC recognition had been previously considered as an interaction between static structures. By using NMR and single-molecule techniques the authors bring our insight in this process now a step further. By measuring dynamics by NMR and binding forces using optical tweezers, they provide a more complete picture of the interaction process. This approach has led to a better understanding of T cell signaling and may form the basis of improved immunotherapies.

The groups of Polenova and Gronenborn combine solid state MAS NMR and cryo-EM for structural efforts of single- and multidomain proteins. They explore how many C–C distances derived from solid-state MAS NMR are needed to satisfactorily define protein structures when using 1,6-^{13}C glucose or 2-^{13}C glucose as the carbon source (Russell et al. 2019). They consider three known single-domain structures and pick randomly a certain percentage of distance constraints of up to either ~5.5 Å or ~7 Å. They conclude that for the 7 Å set, around 15 random restraints per residue are required to attain maximum accuracy. For multidomain proteins additional information is needed on domain orientations, quaternary structure, and protein shape, such as from cryo-EM.

A combination of experimental NMR data and computational studies is used by Meiler and coworkers to derive a unified structural model of the mammalian translocator protein (TSPO) from the outer mitochondrial membrane, both in the presence and absence of the TSPO ligand (Xia et al. 2019). The model is based on experimental structures from X-ray and NMR data of the mouse protein and of bacterial homologs in detergent micelles. Structural differences are suggested to be due to the use of detergents. Using the Rosetta approach, the authors derive a unified model of the translocator. Furthermore, they propose a binding site for a PET imaging ligand.

A continuous challenge in anti-viral therapy is the development of drug resistance. The groups of Schiffer and Ishima use a combination of NMR relaxation studies and crystal structure-based molecular dynamics simulations to elucidate mechanisms of how HIV-1 protease develops drug resistance by mutations (Ishima et al. 2019). The combined data from NMR, MD simulations, and pulsed EPR methods exhibit a wide variety of global and local changes in dynamics, stability and structure of the resistance mutants. Some mutations abrogate binding to particular inhibitors while others maintain similar binding free energies through enthalpy/entropy compensation. While the combination of several techniques provides a better insight into biophysical changes of drug-resistance HIV-1 protease, guidelines to the design of more robust inhibitors are still not available and will need more advanced technologies.

The structure and function of the integrase BamA from the outer membrane of Gram-negative bacteria has been found to adopt different β-barrel conformations based on several X-ray and cryo-EM structures with open and closed side-way gates. Hiller and coworkers used a selection of conformation-selective nanobodies that stabilize BamA in specific conformations. They followed up with solution-NMR based screening to select complexes of open and closed states of BamA. Three crystal structures were solved that contained open and closed states, which differ in the association of the N- and C-terminal β strands (Kaur et al. 2019).

The impact of integrative structural models depends strongly on archiving and dissemination of the results. The groups of Berman and Sali describe the needs for and efforts to integrate multiple structural tools from the computational side (Vallat et al. 2019). Integrative modeling is cast as a computational optimization problem and strategies to consider various experimental data are discussed. The authors describe in detail the open-source integrative modeling platform (IMP) software they developed over a decade (https://integratitivemodeling.org). They follow up by discussing the requirements for archiving integrative models. They include how the Protein Data Bank handles multi-method structural data and provide standards for archiving integrative models and PDB-provided tools. An impressive set of examples is given ranging from small systems up to a model of the nuclear pore.

The introduction of integrative approaches has considerably changed the mode of operation in structural biology. From one lab, one technique we have seen a shift to more collaborative projects involving a combination of methods. It is clear from the papers in this Special Issue and other recent work that NMR spectroscopy plays a crucial role in this area.

References

Alber F, Förster F, Korkin D, Topf M, Sali A (2008) Integrating diverse data for structure determination of macromolecular assemblies. Annu Rev Biochem 77:443–477

Bardiaux B, Cordier F, Brier S, Lopez-Castilla A, Izadi-Pruneyre N, Nilges M (2019) Dynamics of a type 2 secretion system pseudopilus unraveled by complementary approaches. J Biomol NMR. https://doi.org/10.1007/s10858-019-00246-4

Carlon A, Ravera E, Parigi G, Murshudov GN, Luchinat C (2019) Joint X-ray/NMR structure refinement of multidomain/multisubunit systems. J Biomol NMR. https://doi.org/10.1007/s10858-018-0212-3

Grishaev A, Guo L, Bax A (2005) Refinement of multidomain protein structures by combination of solution small-angle X-ray scattering and NMR data. J Am Chem Soc 127:16621–16628

Hartlmüller C, Spreitzer E, Göbl C, Falsone F, Madl T (2019) NMR characterization of solvent accessibility and transient structure in intrinsically disordered proteins. J Biomol NMR. https://doi.org/10.1007/s10858-019-00248-2

Hwang S, Öster C, Chevelkov V, Giller K, Lange S, Becker S, Lange A (2019) Characterization of H/D-exchange in type 1 pili by proton-detected solid-state nmr and molecular dynamics simulations. J Biomol NMR. https://doi.org/10.1007/s10858-019-00247-3

Ishima R, Kurt-Yilmaz N, Schiffer CA (2019) NMR and MD studies combined to elucidate inhibitor and water interactions of HIV-1 protease and their modulations with resistance mutations. J Biomol NMR. https://doi.org/10.1007/s10858-019-00260-6

Kaur H, Hartmann JB, Jakob RP, Zahn M, Zimmermann I, Maier T, Seeger MA, Hiller S (2019) Identification of conformation-selective nanobodies against the membrane protein insertase BamA by an integrated structural biology approach. J Biomol NMR. https://doi.org/10.1007/s10858-019-00250-8

Lander GC, Estrin E, Matyskiela ME, Bashore C, Nogales E, Martin A (2012) Complete subunit architecture of the proteasome regulatory particle. Nature 482:186–193

Liko I, Allison TM, Hopper JTS, Robinson CV (2016) Mass spectrometry guided structural biology. Curr Opt Struct Biol 40:136–144

Loquet A et al (2012) Atomic model of the type III secretion system needle. Nature 486:276–279

Mallis RJ, Brazin KN, Duke-Cohan JS, Hwang W, Wang J, Wagner G, Arthanari H, Lang MJ, Reinherz EL (2019) NMR: an essential structural tool for integrative studies of T cell development, pMHC ligand recognition and TCR mechanobiology. J Biomol NMR. https://doi.org/10.1007/s10858-019-00234-8

Rieping W, Habeck M, Nilges M (2005) Inferential structure determination. Science 309:303–306

Russell RW, Fritz MP, Kraus J, Quinn CM, Polenova T, Gronenborn AM (2019) Accuracy and precision of protein structures determined by magic angle spinning NMR spectroscopy: for some 'with a little help from a friend'. J Biomol NMR. https://doi.org/10.1007/s10858-019-00233-9

Vallat B, Webb B, Westbrook JD, Sali A, Berman HM (2019) Archiving and disseminating integrative structure models. J Biomol NMR. https://doi.org/10.1007/s10858-019-00264-2

Xia Y, Ledwitch K, Kuenze G, Duran A, Li J, Sanders CR, Manning C, Meiler J (2019) A unified structural model of the mammalian translocator protein (TSPO). J Biomol NMR 5:4. https://doi.org/10.1007/s10858-019-00257-1