Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

The importance of dynamics in integrative modeling of supramolecular assemblies

Giorgio E Tamò[1], Luciano A Abriata[1] and Matteo Dal Peraro

Revealing the atomistic architecture of supramolecular complexes is a fundamental step toward a deeper understanding of cellular functioning. To date, this formidable task is facilitated by an emerging array of integrative modeling approaches that combine experimental data from different sources. One major challenge these methods have to face is the treatment of the dynamic rearrangements of the individual subunits upon assembly. While this flexibility can be sampled at different levels, integrating native dynamic determinants with available experimental inputs can provide an effective way to reveal the molecular recognition mechanisms at the basis of supramolecular assembly.

Address
Laboratory for Biomolecular Modeling, Institute of Bioengineering, School of Life Sciences, École Polytechnique Fédérale de Lausanne (EPFL) and Swiss Institute of Bioinformatics (SIB), Lausanne, Switzerland

Corresponding author: Dal Peraro, Matteo (matteo.dalperaro@epfl.ch)

Current Opinion in Structural Biology 2015, **31**:28–34

This review comes from a themed issue on **Theory and simulation**

Edited by **Claire Lesieur** and **Klaus Schulten**

For a complete overview see the Issue and the Editorial

Available online 18th March 2015

http://dx.doi.org/10.1016/j.sbi.2015.02.018

Introduction

Supramolecular complexes are the cornerstone of cellular architecture and function. Large assemblies of macromolecules are involved in DNA remodeling, translation and transcription, RNA processing, protein synthesis and degradation, import, export and injection of solutes through cell membranes and to different organelles, ATP synthesis, respiration and photosynthesis, packaging of viral nucleic acids, membrane reshaping, just to name some of the systems best characterized to date. Knowing the structure of these complexes at atomistic resolution is essential to understand how they work and to modulate their function in controlled ways. However, such characterization remains a daunting task for conventional techniques [1].

Whereas the atomic structure of single proteins and complexes of relatively low molecular weight can generally be obtained by NMR spectroscopy and/or X-ray crystallography, supramolecular assemblies are not routinely accessible to these techniques [1]. Although recent advances in cryo-electron microscopy (cryo-EM) are enabling near-atomistic resolution of large assemblies [2], a broad array of experimental methods can provide lower resolution data about overall shape, symmetry, composition, contact sites between constituent molecules, angular and distance restraints between domains, as extensively reviewed in [3]. In this context, *integrative modeling* attempts to consistently combine these heterogeneous, and sometime incomplete, data with the structures of the individual subunits that constitute a complex in order to generate models at near atomistic resolution [4]. However, as we highlight here, the structure of subunits as determined in isolation might differ from the conformation they adopt upon supramolecular assembly [5], further complicating the prediction of native architectures. Such scenario might arise either from artifacts induced by the conditions under which the structure of individual subunits is solved and/or because of actual structural and functional rearrangements taking place upon assembly. One way or the other, these call for the inclusion in integrative modeling protocols of the global and local dynamics of the individual subunits undergoing macromolecular assembly. Notably, the inclusion of dynamic features cannot only lead to more reliable models, but can also reveal fine details of the assembly mechanism and the existence of multiple functional states.

In this opinion we discuss the most recent integrative modeling approaches that attempt to determine supramolecular assembly considering the contribution of these dynamic determinants. We show how canonical integrative modeling approaches start to intersect with molecular simulation techniques leading to new powerful hybrid strategies to unveil the structure and function of large cellular complexes.

The quest for spatial restraints and dynamic determinants

Integrative modeling is open to all kinds of experimental restraints. In the most common flavor, high-resolution structures of subunits are complemented by coarser details of the assembly such as size, volume and shape

[1] These authors contributed equally to this work.

retrieved from small-angle X-ray or neutron scattering (SAXS/SANS) experiments, cryo-EM or tomography. Despite the tremendous progress of these techniques, these volumetric shapes might often not have the resolution required to unambiguously define position and orientation of constituent molecules. Thus, contact information such as residue–residue contacts unveiled through mutagenesis, chemical cross-linking and ion mobility mass spectrometry [6], or even from coevolution analysis [7–9], can help to define assembly rules. Complementarily, chemical footprinting can point at protein surfaces that must remain exposed [10].

Long-distance information can be retrieved from electron paramagnetic resonance (EPR) experiments on samples with localized radical labels [11], as well as from FRET experiments using native or engineered fluorophores [12,13]. The effect of a localized paramagnetic tag on the nuclear relaxation properties of another molecule can also be used to infer spatial proximity through NMR experiments [14]. All these constraints arising from NMR, EPR and FRET usually reflect ensemble distributions and thus internal dynamics. On the more static, but still very informative end, direct interactions between monomers can be inferred from NOEs and from chemical shift perturbations upon binding [15], whereas residual dipolar couplings provide information about relative angular orientations [16,17]. H/D exchange experiments followed by NMR or by mass spectrometry can in turn provide information about solvent exposure, complementing the information about contacts [18,19].

NMR ensembles of individual subunits can moreover explicitly provide information about internal dynamics and conformational heterogeneity that could be important upon assembly. The same holds for solid-state NMR, with the advantage that it is less limited by molecular size therefore increasingly helping in structure determination for large assemblies, as recently shown for the needle structure of a type III secretion system [20].

Integrating structures, restraints and dynamics into models of supramolecular complexes

In the best-case scenario, structures of the individual subunits that constitute the complex are available from experiment or can be modeled with high confidence. However, these structures are often static and representative of single states which conformations might differ from those that fit in a supramolecular complex. This discrepancy can be due to the particular conditions in which the structure of the monomer was determined or may arise from the native dynamics underlying the molecular recognition pathways that lead to assembly through mechanisms as diverse as conformational selection or induced fit. By reviewing key works and strategies in the field, we posit that accounting for the flexible and dynamic nature of subunits is a key ingredient of the modeling process, which can not only facilitate supramolecular structure prediction, but can also lead to biologically relevant outcomes.

On one hand, some integrative modeling approaches use experimental restraints to drive deformations of the subunits upon assembly, gradually improving the fit to the experimental restraints. This typically involves biasing with the given restraints the algorithms that perform backbone/sidechain refinement, normal modes analysis [21] or molecular dynamics (MD) simulations of variable granularity [5,22] (Figure 1a). The other main avenue consists in using the experimental restraints to select conformations from existing ensembles, namely compiled from X-ray and NMR ensembles, homology models, and/or MD trajectories (Figure 1b). Whereas the first set of strategies resembles and can potentially describe induced-fit mechanisms underlying protein recognition, the second class of approaches seems well positioned to capture structural assembly driven by conformational selection.

Using experimental restraints to drive deformation of subunits into a supramolecular complex

A number of methods use experimental spatial restraints, as obtained from low-resolution experiments, to directly drive the physical deformation of starting structures into consistent conformations (Figure 1a). With a very simple and coarse way to represent protein flexibility, Situs [24] performs very well at fitting assemblies to volumetric maps derived from EM experiments within a broad range of resolution (e.g., up to 30 Å [25]). Within this approach flexibility is considered by converting starting structures into a skeleton that is allowed to sample conformations following distributions observed in the Protein Data Bank. In this way the fitting protocol uses a knowledge-based force field on which restraints are implemented to penalize shape differences between the assembled model and the experimental volume [24]. Although the dynamics of individual subunits is not completely sampled, this is a simple way to enlarge the conformational space accessible for assembly, which has been successfully applied to several systems [26–29], most notably myosin fibers [25] and full muscle filaments [30].

Similar ways to adjust individual structures into assemblies make use of normal modes computed from a deformable elastic network, as done by DireX [31] and iMODfit [32,33]. Recently, it has been also shown that the combination of diverse flexible fitting protocols of this kind can improve pseudo-atomistic models based on intermediate-resolution EM maps, providing in turn a general way to assess the fits [34]. ATTRACT is another software that makes explicit use of flexibility by exploiting normal mode analysis. Specifically, it performs systematic energy