

From integrative structural biology to cell biology

A Sali¹

¹RCSB PDB/UCSF, San Francisco, CA

sali@salilab.org

Integrative modeling is an increasingly important tool in structural biology, providing structures by combining data from varied experimental methods and prior information. As a result, molecular architectures of large, heterogeneous, and dynamic systems, such as the ~52 MDa Nuclear Pore Complex, can be mapped with useful accuracy, precision, and completeness. Key challenges in improving integrative modeling include expanding model representations, increasing the variety of input data and prior information, quantifying a match between input information and a model in a Bayesian fashion, inventing more efficient structural sampling, as well as developing better model validation, analysis, and visualization. In addition, two community-level challenges in integrative modeling are being addressed under the auspices of the Worldwide Protein Data Bank (wwPDB). First, the impact of integrative structures is maximized by PDB-Dev, a prototype wwPDB repository for archiving, validating, visualizing, and disseminating integrative structures. Second, the scope of structural biology is expanded by linking the wwPDB resource for integrative structures with archives of data that have not been generally used for structure determination but are increasingly important for computing integrative structures, such as data from various types of mass spectrometry, spectroscopy, optical microscopy, proteomics, and genetics. To address the largest of modeling problems, a type of integrative modeling called metamodeling is being developed; metamodeling combines different types of input models as opposed to different types of data to compute an output model. Collectively, these developments will facilitate the structural biology mindset in cell biology and underpin spatiotemporal mapping of the entire cell.