Published in final edited form as:

Curr Opin Struct Biol. 2017 April ; 43: iv–vi. doi:10.1016/j.sbi.2017.04.008.

Editorial overview: Theory & computation

Carol Beth Post and Ronald M Levy

Three of the reviews appearing in COSB Volume 43 highlight innovative directions in atomistic simulations of proteins with the goal of understanding how their environments, both simple and complex, affect their molecular behavior, both structure and dynamics. The structure of a protein is strongly influenced by solvation. Statistical thermodynamics provides a framework for treating solvation effects and a key quantity is the solvation free energy. In this volume [Matubayashi](#) reviews a theory of solutions he has developed to compute the solvation free energy of proteins and shows how it can be applied to understand the thermodynamics of protein structural fluctuations in solution, the effects of solvent on protein-protein interfaces, and protein denaturation by urea. In a cellular environment, proteins are often present at high total concentrations, a situation known as molecular crowding. The paper by [Zhou](#) and [Qin](#) describes how atomistic and coarse grained simulations can be used to model cell like conditions of crowding and how crowding influences protein folding and misfolding. These authors describe a new cellular phase referred to as a “protein droplet”, they sketch a phase diagram and suggest a functional role for protein droplets. Voltage gated sodium channels (VGSCs) are membrane proteins that contain 24 transmembrane alpha-helices. These channels provide the conduit for the passage of ions across the plasma membrane, a process that results in the propagation of an electrical signal. The review by [Carnevale](#) and [Klein](#) describes how small molecules can bind to VGSCs and modulate their conductance. VGSCs open and close in response to a change in polarization of the lipid membrane. The pore domain is allosterically coupled to a voltage sensing domain; molecular simulations have helped to reveal the mechanism of the electro-mechanical coupling and shown how small molecule anesthetics like isoflurane can bind to specific locations that lead to a pore blocking mechanism of inhibition.

The ongoing challenge in the field of molecular simulations of modeling processes that occur on longer and longer timescales is the subject of three reviews in this volume. Simulating the long-time (e.g. microseconds to milliseconds) kinetics of biomolecules with atomistic molecular dynamics is challenging in part because of the many degrees of freedom of the problem and the difficulty identifying which coordinates govern the rate determining conformational transitions. The paper by [Noe](#) and [Clementi](#) reviews the concept of slow collective variables, and explains why they are so useful for studying high-dimensional dynamical systems. Recent developments of the theory are discussed, including the formulation of a variational principle to find optimal approximations to slow collective variables from simulation data of protein. Defining the conformational paths between energy wells on the free energy landscape is equally challenging. [Zuckerman](#), [Chong](#) and [Saglam](#) review methods using ensemble sampling to identify transition paths of ms-timescale processes such as protein conformational changes and molecular association. They outline the approaches based on complete paths (transition path sampling; dynamic importance sampling) and on trajectory segments (region-to-region; interface-to-interface), as well as

the advances that have been made in the efficiency of calculating transition trajectories and the ability to estimate rate constants. Applications to protein conformational transitions, protein folding and ligand binding are given. Another answer for making the simulation of long-time processes tractable is the use of adaptively biased potentials during the simulation. Dickson summarizes a number of adaptive biasing potentials that have emerged to enhance sampling, and highlights unique attributes in their behavior. The review explains a unified treatment of the biasing schemes into direct-histogram and log-histogram classes of bias. The various biasing potentials are compared by a visual illustration of their time evolution and by contrasting several features important for consideration in their implementation in simulations.

The continuing increase in computer resources and computational power greatly motivates the natural application of molecular simulation to explain the molecular basis of experimental data and provide insight into biological processes. Allison reviews the use of biased and unbiased ensemble approaches to match and interpret experimental data from NMR, SAXS and EPR. The use of maximum entropy and Bayesian inferences is described for the problems associated with the statistical accuracy in the simulation of rare events, as detected in relaxation dispersion NMR experiments. How simulated structural ensembles complement experimental measurements on intrinsically disordered proteins (IDP) is the focus of the second review in this area. The paper by Shea and Levine describes approaches to generate IDP ensembles, and issues related to force fields and modeling solvation. Limitations in water models become even more apparent in simulations of IDPs given the greater exposure to solvent of IDPs relative to globular proteins. The paper points out some of the functional insights into IDP behavior that have arisen from simulations, including mechanisms for IDPs binding proteins, aggregation and fibril formation, and separation into protein-rich phases as in forming membrane-less organelles in cells.

Increasing computational power also benefits the simulation of highly complex systems. Im, Patel and Qi outline how bacterial outer membranes and components of outer membrane proteins (OMPs) can now be modeled effectively to study the structural basis for transport of ions, substrates, antibiotics across the outer membrane. Such simulations require modeling complex lipopolysaccharides and setting up solvated membrane systems with a large number of diverse lipid molecules with defined composition. The paper cites applications to several OMPs including OmpF porin, TonB-dependent transporter FecA, and the BAM complex.

The paper by Abel, Friesner, and colleagues in this volume focuses on recent developments in the modeling of protein-ligand interactions and its role in drug discovery. It tells a story of free energy perturbation (FEP) simulations coming of age. FEP simulations of protein-ligand binding, first carried out in the 1980s, have become a very useful tool for computational drug design due to the improvement in force fields, the enormous increase in computer power, especially gpu computing, and the design of effective workflows to carry out FEP calculations on a large scale. The paper recounts how computational chemistry and structural biology served as the foundation for the successful design of an allosteric inhibitor of Acetyl-CoA carboxylase implicated in several disease pathologies including “fatty liver” disease, and of inhibitors of Tyk2 kinase implicated in autoimmune disease.

Potts models illustrate how strikingly biological physics intersects with structural biology and biology. Levy, Haldane and Flynn review recent developments in Potts models of sequence covariation in proteins and their potential promise for yielding principles needed to interpret the massive amounts of available sequence data. They explain how contacts inferred from evolutionary sequence patterns are used to predict structure and map out conformational free-energy landscapes. Additionally, the paper explains the experimental evidence to validate the premise that the statistical scores of the Potts Hamiltonian reflect protein fitness. The advantages of Potts models are also discussed for explorations of the mutational landscape concerned with epistasis, the phenomenon that the effect of a mutation depends on the background sequence.