James Diao
MB&B 452
4 April 2017

Section 8: Protein Simulation 2

The two readings build on last week's papers, seeking to better understand the molecular composition of proteins. The first, by Dill et al., is a review paper describing the task of understanding and predicting protein folding. The second, by Bowman et al., introduces MSMBUILDER, a python application and library for automating the construction of Markov state models (MSMs). Both papers describe recent advances in protein modeling and demonstrate the value of computational methods in understanding the nature of proteins.

Weikl et al. begin by describing the three main branches of the study of protein folding: (a) the folding code, (b) the folding mechanism, and (c) structure prediction. Recent advancements have enabled precise prediction of small protein structures, as well as applications in foldables and polymer design. This progress is driven by new methods and data, such as multiple sequence alignment, fast-homology methods, and the Protein Databank (PDB). On the mechanism side, Weikl et al. describe increasing evidence supporting a funnel-shaped protein-folding energy landscape, which provides an intuitive analogy for the conformation heterogeneity and chain entropy of proteins. The zipping and assembly (ZA) paradigm—where independent local optimization precedes joining into structures—helps explain the Levinthal paradox of fast folding. The newest challenges are to refine homology models, reduce errors to routinely better than 3 A, accommodate large, multi-domain proteins, and translating folding experiments and computer simulations into insights about molecular behavior.

The paper by Pande et al. introduce a tool for automated construction of MSMS, which provide a "kinetic clustering" of data. This means that conformations with rapid interconversions are grouped together, while conformations with slow interconversions are separated, creating a decomposition of the state space with the Markov property. Specifically, MSMBUILDER takes a bottom-up approach by lumping many Markov microstates (with high structural similarity) into Markov macrostates. This method was applied for villin headpiece, as an archetypal peptide, demonstrating applicability to full protein systems. In summary, MSMBUILDER is capable of precise structural prediction, with accurate thermodynamic and kinetic profiles. Future work will focus on estimation of transition probabilities and minimizing computational costs.

- Dill KA, Ozkan SB, Shell MS, Weikl TR. (2008) The Protein Folding Problem. Annu Rev Biophys,9, 37:289-316. PMID: 2443096.PDF
- Bowman GR, Beauchamp KA, Boxer G, Pande VS. "Progress and challenges in the automated construction of Markov state models for full protein systems," J. Chem. Phys. 131 (2009) 124101 PDF