Hierarchical Approach to Coarse-Grain Modeling of Protein Structures which Accounts for Fast Local Fluctuations

Maria Kurnikova, Dept. of Chem. Carniegie-Mellon University

Strength, as well as spatial and temporal distribution of individual hydrophobic contacts and hydrogen bonds formed within a protein structure and with surrounding water affect protein fluctuations and its flexibility/rigidity on both short (ps-ns) and long (ns-ms) time-scales and ultimately affect its function. Characterizing protein long time-scale dynamics is needed in order to understand functionally important protein dynamics and conformational rearrangements, e.g. in allosteric regulation, domain rearrangements upon ligand binding, or in signal propagation in protein receptors. It seems thus logical to coarse-grain a protein model to achieve simulations on longer time-scales, as it is expected that on longer time-scales collections of atoms should exhibit consorted motions or ensemble behavior. Several schemes of modeling proteins as collections of mutually interacting blocks of atoms had been proposed all of which require an initial "preprocessing" of an atomistic structure to define the sub-collectives of atoms as blocks. Since both structure and short-time dynamics affect long-time dynamics in proteins and protein complexes we both should be taken into account in an algorithm for coarse-graining. However, extrapolating short-time behavior as well as small fluctuations beyond linear regime is a non-trivial task. In order to approach this goal we studied the interplay between short-time protein fluctuation properties and their conformational space using 10-20 ns molecular dynamics (MD) simulations of several proteins of various sizes solvated in water at room temperature. The duty cycles of the non-covalent pair-wise contacts (hydrogen bonds and hydrophobic contacts) were defined as the percentage of time a given contact is present, which characterize stability and fluctuations of individual contacts. These are then used to construct a topological network of persistent bonds corresponding to a protein structure with its fast fluctuations "integrated out". Thus constructed topological network is then analyzed using graph theoretical algorithm FIRST (*Jacobs et al. Proteins* 44, 150–165) to infer "rigid" blocks in the structure, which may further be used in long time-scale simulations. We will also compare dynamically smoothed algorithm and the original structure/energy only based algorithm outcomes for block decomposition.