Abstract

Specificity and affinity of protein-RNA recognition rely on the three-dimensional conformational plasticity of the binding partners. In the event of the recognition, both the binding partners often undergo essential conformational changes to facilitate the event. A significant understanding of conformational adaptations requires a robust dataset of protein-RNA complexes along with their unbound forms. To address this, a protein-RNA docking benchmark integrated with binding affinity data has been developed. Bound and unbound structures of RNA-binding proteins (RBPs) and RNAs from this dataset are used to evaluate conformational changes upon binding. To characterize the conformational plasticity of RBPs, a detailed analysis of side-chain conformations of amino acid residues in bound and unbound forms is carried out. Moreover, rotamer libraries have been developed to identify the side-chain transitions of amino acid residues upon binding RNA. Analyses performed in this dissertation shows that side-chain conformational transitions of the interface residues are more frequent compared to those observed for non-interface residues. Furthermore, it has been also observed that the side-chain rotamers vary depending on the backbone conformations. Majority of the conformational transitions are associated with the interface residues, which often undergo from higher-to-lower energy states. These analyses lead to the development of six secondary structure-dependent and six backbone-dependent rotamer libraries corresponding to interface, non-interface and overall protein surface of both bound and unbound forms of RBPs. Similarly, the conformations of dinucleotide RNAs at the interface and the non-interface regions are investigated. An RNA dinucleotide conformer library is developed to aid in identifying specific RNA conformers at the interface that facilitate the protein-RNA recognition. To further understand the effect of conformational dynamics in RNA recognition, binding between U1 SL-4 snRNA and splicing Factor 3A1 are analysed using molecular dynamics simulations along with the effects of mutations in binding. A comparison of rotamers across unbound, bound and mutant complexes highlights the role of the RGG motif and UUCG tetraloop of RNA in binding specificity. Knowledge generated in this dissertation will further contribute to the development of flexible docking methods, algorithms that aim to predict binding affinity and machine learning-based structure modeling platforms for studying protein-RNA recognition.

Keywords: protein-RNA recognition, protein-RNA docking benchmark, binding affinity, rotamer libraries, RNA conformers library, conformational changes, machine learning, molecular dynamics simulations.