A Protein-RNA Docking Benchmark

Amita Barik, Nithin C, Manasa P and Ranjit Prasad Bahadur*

Department of Biotechnology, Indian Institute of Technology, Kharagpur 721302, India

Address for correspondence:
E-mail: ranjitp_bahadur@yahoo.com

RNA plays a critical role in biological processes but due to limited knowledge of its structural model, the progress in protein-RNA docking is not highly paced. In this article, we present an overview of protein-RNA benchmark which will provide a guideline for the protein RNA docking methods. Our non-redundant protein-RNA docking benchmark dataset is derived from the available bound and unbound structures in the PDB involving polypeptide and nucleic acid chains. It comprises of 45 complexes out of which 9 are unbound-unbound cases where both the protein and RNA are available in the free form. The rest 36 cases are of unbound-bound type where only the protein is available in the free from. The conformation change upon complex formation is calculated by distance matrix alignment method, and based on that they are classified into rigid, semi-flexible and full-flexible. While in the rigid body category, complex formation does not accompany significant conformational change, the full flexible test cases show considerable conformational changes involving domain movement or RNA base modification. The benchmark covers four major groups of RNA, viz., t-RNA, Ribosomal RNA, Duplex RNA and Single-stranded RNA. In general RNA is more flexible compared to its partner protein, and the interface region between them is as flexible as their solvent exposed surfaces. The structural diversity of the complexes in the benchmark set should provide a common ground for comparison and development of protein-RNA docking methods which will help in the structure prediction of the large versatile RNAs. This benchmark is available for free on the internet.