Abstract

The Protein–RNA Interface Database (PRIDB) is a comprehensive database of protein–RNA interfaces extracted from complexes in the Protein Data Bank (PDB). It is designed to facilitate detailed analyses of individual protein–RNA complexes and their interfaces, in addition to automated generation of user-defined data sets of protein–RNA interfaces for statistical analyses and machine learning applications. For any chosen PDB complex or list of complexes, PRIDB rapidly displays interfacial amino acids and ribonucleotides within the primary sequences of the interacting protein and RNA chains. PRIDB also identifies ProSite motifs in protein chains and FR3D motifs in RNA chains and provides links to these external databases, as well as to structure files in the PDB. An integrated Jmol applet is provided for visualization of interacting atoms and residues in the context of the 3D complex structures. The current version of PRIDB contains structural information regarding 926 protein–RNA complexes available in the PDB (as of 10 October 2010). Atomic- and residue-level contact information for the entire data set can be downloaded in a simple machine-readable format. Also, several non-redundant benchmark data sets of protein–RNA complexes are provided.