PyMine - A PyMol plugin to integrate and visualize chemical and biological data for drug discovery

Zhijun Li and Rajan Chaudhari
University of the Sciences, USA

Tremendous amount of chemical and biological data are being generated by various high-throughput biological techniques which could facilitate modern drug discovery. These data include protein sequences, protein three-dimensional structures, single amino acid variations (SAV), known ligands and substrates of proteins, and pathways. Currently, these data are classified and deposited into various chemical-informatics and bio-informatics databases based on the data type. However, the lack of integration makes it very challenging for individual scientist to access and understand all the data related to a specific protein of interest. We developed a bioinformatics tool, PyMine, which extracts various chemical and biological data from a variety of high-quality databases and presents them in a graphic and uniformed way. The primary databases accessible by PyMine include protein UniProt database, KEGG pathway database, PDB macromolecule structure database, ChEMBL database of small ligands, IBIS database of protein binding sites and HUMSAVAR database of SAVs. Users input the PDB ID of the protein of interest; all related data will then be extracted and whenever possible, mapped to the three-dimensional structure of the protein. Thus, PyMine can be used as a central data-hub to visualize and access various types of data related to the protein of interest as well as to develop new idea for structure-based molecule design. For convenience, PyMine was developed as a plugin of PyMol, a popular and free graphic modeling tool.

Biography
Zhijun Li received his PhD degree from Vanderbilt University in 2000 and underwent Postdoctoral training at Massachusetts Institute of Technology and Vanderbilt University. He is currently the Associate Professor of Bioinformatics at University of the Sciences, formerly Philadelphia College of Pharmacy. He has published more than 28 papers in peer-reviewed journals and has been awarded a number of grants by various funding agencies for his research on protein structure and function prediction and computer-aided molecular design. He has been a reviewer for a variety of scientific journals and funding agencies including NSF.

Notes: