4190260

Advanced searches of the Protein Data Bank in Jupyter notebooks Board #176

Date March 24, 2025

OVERVIEW COMMENTS

The Protein Data Bank (PDB) holds an extensive amount of information, and can be a vital tool when performing background research for biochemical work. In an attempt to make the information in the PDB more accessible, the RCSB Search API was employed within Jupyter Notebooks to create more customizable and user-friendly tools with simple Python code. Areas of focus include structure motif searches used to predict the function of proteins based on the 3-dimensional shape of their active sites, searches for FDA Approved Drugs, as well as searches targeting ligands with specific characteristics. This code has been built into Jupyter Notebook templates that include both examples of these searches as well as annotated code that users can customize to more efficiently run advanced searches on the PDB and download structure and small molecule files returned by the search. Future plans include increasing the amount and type of information available from a search, as well as expanding the scope of our notebooks to cover more types of searches.

Presenter

PC Paul Craig

Co-Authors

AR Anna Rack

MP Murphy Pick

VG Victoria Greever