Aim: Protein structure retrieval from PDB database and its visualization using any visualization software.

Theory:

PDB Database

The Protein Data Bank (PDB) is a database for the three-dimensional structural data of large biological molecules, such as proteins and nucleic acids. The data, typically obtained by X-ray crystallography, NMR spectroscopy, or, increasingly, cryo-electron microscopy, and submitted by biologists and biochemists from around the world, are freely accessible on the Internet. The PDB is overseen by an organization called the Worldwide Protein Data Bank, wwPDB. The PDB is a key in areas of structural biology, such as structural genomics. Most major scientific journals, and some funding agencies, now require scientists to submit their structure data to the PDB. Many other databases use protein structures deposited in the PDB. For example, SCOP and CATH classify protein structures, while PDBsum provides a graphic overview of PDB entries using information from other sources, such as Gene ontology.

The PDB database is updated weekly, along with its holdings list. Most structures are determined by X-ray diffraction, but about 10% of structures are determined by protein NMR. When using X-ray diffraction, approximations of the coordinates of the atoms of the protein are obtained, whereas using NMR, the distance between pairs of atoms of the protein is estimated. The final conformation of the protein is obtained from NMR by solving a distance geometry problem. A few proteins are determined by cryo-electron microscopy. Clicking on the numbers in the linked external table displays examples of structures determined by that method. For PDB structures determined by X-ray diffraction that have a structure factor file, their electron density map may be viewed. The data of such structures is stored on the "electron density server".

JMOL database:

Jmol is computer software for molecular modelling chemical structures in 3-dimensions. Jmol returns a 3D representation of a molecule that may be used as a teaching tool, or for research e.g., in chemistry and biochemistry. It is written in the programming language Java, so it can run on the operating systems Windows, macOS, Linux, and Unix, if Java is installed. It is free and open-source software released under a GNU Lesser General Public License (LGPL) version 2.0. A standalone application and a software development kit (SDK) exist that can be integrated into other Java applications, such as Bioclipse and Taverna. A popular feature is an applet that can be integrated into web pages to display molecules in a variety of ways. For example, molecules can be displayed as ball-and-stick models, space-filling models, ribbon diagrams, etc. Jmol supports a wide range of chemical

file formats, including Protein Data Bank (pdb), Crystallographic Information File (cif), MDL Molfile (mol), and Chemical Markup Language (CML). There is also a JavaScript-only (HTML5) version, JSmol, that can be used on computers with no Java.

Suggested readings: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC102472/

https://onlinelibrary.wiley.com/iucr/doi/10.1107/S0021889810030256