

PDBe: Bringing structure to biology and beyond

Matthew J Conroy¹, Sameer Velankar¹

¹Protein Data Bank In Europe EMBL-EBI, Cambridge, United Kingdom

E-mail: conroy@ebi.ac.uk

Protein Data Bank in Europe (PDBe) [1] a founder member of the wwPDB, has a mission to bring structure to biology. Data in the PDB archive is utilised by a diverse range of users including biochemists, geneticists, medicinal chemists, physicians, and even artists and school children. The complex nature of 3D structural data means that those with a limited background or training in structural biology do not always find it easy to exploit the rich information content of the structural archives to help them answer their research questions. It is for archive keepers such as PDBe to address this challenge and develop new ways of making structural information more easily accessible, relevant, and up-to-date.

Recently, PDBe implemented several new features to facilitate easy access to macromolecular structure data. These are inherently visual and, in collaboration with CEITEC in the Czech Republic, we have incorporated an easy to use, in-browser 3D visualizer on all structure webpages. With minimal data transfer needed, this works on all devices, including smartphones. In addition to models, electron density (for crystal structures) and electric potential maps (for cryoEM models) can be viewed where possible. 3D visualisation is interactively linked to sequence and topology views, a helpful teaching tool to link sequence and structure. All views can be overlaid with annotations of domain architecture and validation.

Structure data is rendered a much richer and powerful pedagogical aid if linked to accurate up-to-date annotations. The SIFTS resource (pdbe.org/sifts) [2] provides updated residue-level mappings to UniProt and annotations from a variety of other databases on a weekly basis. In conjunction with Europe PMC [3], PDBe is able to display figures and legends from open-access publications describing PDB entries and datamine literature for instances of PDB codes.

PDBe pages show the availability of, and indicate the results from, other resources useful to understanding macromolecular structure using a collection of PDB components. These include data from the PDB redo team at NKI, Amsterdam; Rajini from Dr. Madan Babu's lab at LMB, Cambridge; and a component providing biological relevance of assembly annotation in the PDB from Dr. Emanuel Levy's lab.

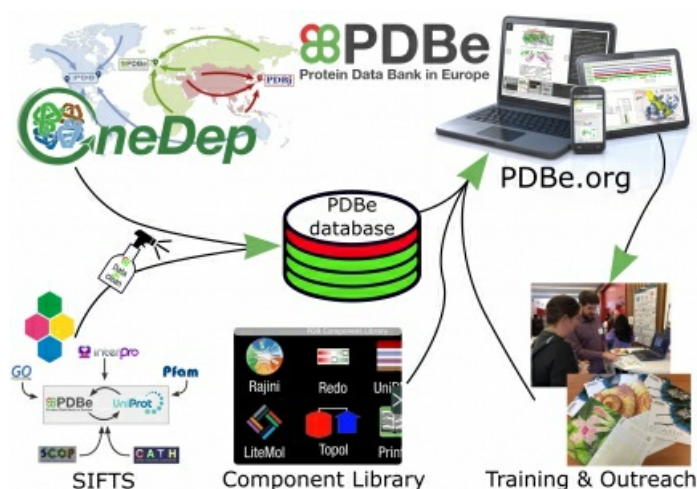
Via a RESTful API, data can be served programmatically and this is used in educational services such as Proteopedia, where a 'preferred' assembly of a structure is selected, and JalView where structure selection is used to select an appropriate structure and annotation is displayed on sequence data.

Resources are being developed to support the school science curriculum in collaboration with Newcastle University. We are also introducing structures to school students studying art, and this venture bringing together art and science, has already produced some stunning artworks. We hope that this collaboration will inspire the next generation of scientists and molecular graphic artists, and also enable the current generation of scientists to view their data through fresh eyes.

[1] Velankar et al. (2016) Nucleic Acids Res. 44, D385-395

[2] Velankar et al. (2013) Nucleic Acids Res. 41, D483-489

[3] McEntyre et al. (2015) Nucleic Acids Res. 43, D1042-1048



Keywords: [PDB](#), [Protein Data Bank](#), [data visualization](#)