PDBe: Quick tour

Sameer Velankar [1]

- Proteins
- Structures
- Beginner
- 0.5 hour

This quick tour provides a brief introduction to the Protein Data Bank in Europe (PDBe), EMBL-EBI's resource for the collection, organisation and dissemination of data on biological macromolecular structures.

Learning objectives:

- Basic understanding of PDBe and what it can do
- Knowing where to find out more about PDBe

What is PDBe?

EMBL-EBI's Protein Data Bank in Europe [2] (PDBe) is the European resource for the collection, organisation and dissemination of data about biological macromolecular structures. PDBe is one of four partners in the Worldwide Protein Data Bank [3] (wwPDB, see information box below), the consortium entrusted with the collation, maintenance and distribution of the single global repository of macromolecular structure data. The members of the wwPDB exchange data on a daily basis, ensuring that users access the most comprehensive and up to date collection of structures.

PDBe’s mission is ‘bringing structure to biology’, and the team provide a comprehensive range of tools and services that make the complex and rich structural and functional information in the PDB [4] more accessible and useful to the wider biomedical community.

The Worldwide Protein Data Bank [3] (wwPDB) consists of four organisations that act as deposition, data-processing and distribution centres for the PDB. The founding members are the Research Collaboratory for Structural Bioinformatics [5] (RCSB, USA), the PDBe team [2] at EMBL-EBI (Europe) and PDBj [6] (Japan). In 2006, the BioMagResBank [7] (BMRB) in the USA joined wwPDB. The mission of the wwPDB is to maintain a single Protein Data Bank archive of macromolecular structural data that is freely and publicly available to the global scientific community.
community.

What can I do with PDBe

Why do we need PDBe?

Structural biology continues to have an enormous impact on our understanding of biology and medicine. Three-dimensional structures give us insight into how macromolecules work, and help to explain how their functions are disrupted by mutation or by interaction with small molecules [8]. Biologists are now in a position to combine structural biology with genomics approaches to determine every protein fold encoded by a genome. As these efforts gain momentum, the demand for efficient access to sequence, structure and protein-family information is rising; biologists need standardised ways of viewing and describing proteins so that they can share and use data effectively.
Figure 1 The PDBe homepage provides access to tools and services for depositing, searching, exploring and analysing macromolecular structural data.

PDBe services

Macromolecular structures are complex and therefore advanced tools are needed for their analysis. PDBe’s suite of web-based services allows you to make the most of the global collection of publicly available structural data.

<table>
<thead>
<tr>
<th>Service</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access to PDB entries</td>
<td>Get more information about PDB entry 1xyz.</td>
</tr>
<tr>
<td>Download PDB entry</td>
<td>Download the PDB file of entry 1xyz (access from home page).</td>
</tr>
<tr>
<td>PDB Deposition Service</td>
<td>Deposit data to the PDB using AutoDep.</td>
</tr>
<tr>
<td>PDBeView</td>
<td>PDB search tool.</td>
</tr>
<tr>
<td>PDBeChem</td>
<td>Dictionary of chemical components (ligands, small molecules and monomers) in the PDB. Search and visualise ligands using keywords, PDB codes, SMILES strings, fragments or fingerprints.</td>
</tr>
<tr>
<td>PDBeExplore</td>
<td>Browse, list and analyse the structural knowledge embodied in the PDB based on classification systems that are familiar and intuitive to molecular biologists, biochemists and other life scientists.</td>
</tr>
<tr>
<td>PDBeMotif</td>
<td>Query and analysis of ligands, sequence and structure motifs, their relative position and the neighbouring environment.</td>
</tr>
<tr>
<td>PDBePisa</td>
<td>Search and analysis of protein interfaces, surfaces and assemblies. Can be used to predict probable quaternary structures (assemblies).</td>
</tr>
<tr>
<td>PDBeFold</td>
<td>Secondary Structure Matching (SSM) service for comparing and aligning 3D protein structures; search for structures that are similar to that of a reference protein, and visualise superimposed structures.</td>
</tr>
<tr>
<td>SIFTS</td>
<td>The Structure Integration with Function, Taxonomy and</td>
</tr>
</tbody>
</table>

SIFTS [21]
Sequence (SIFTS) resource provides up-to-date mappings between macromolecular structures and other major bioinformatics resources.

**NMR at PDBe** [22]
Access the NMR project pages at PDBe.

**EM Data Bank** [23]
Access the EM Data Bank pages at PDBe.

**BioBar** [24]
A downloadable toolbar for Mozilla browsers, providing instant access to PDBe and many other freely available bioinformatics services.

**PDBe** [9] Services: at a glance

**Analysing data in PDBe**
Searching and visualising data from PDBe

**Basic searches:** With PDBView you can perform simple searches based on a PDB [25] code, keyword, author name(s), ligand [26] names or code, or any piece of text. The query can be restricted to specific types of experiments (crystallography, NMR or electron microscopy [27], for example), species, or to a particular range of resolutions. You can also use sequence information to search for structures with closely matching sequences.

**Structure comparisons:** PDBFold [28] allows you to identify structures that are similar to that of a reference protein. Pairwise or multiple comparisons as well as 3D alignments of structures can be performed.

**Ligands and their binding sites:** PDBChem [29] allows you to search for and visualise any molecule in the PDB's ligand dictionary using keywords, PDB codes, SMILES [15] strings (stereo or non-stereo), fragments or fingerprints. You can use the PDBMotif search interface to query ligand-binding sites and their geometry or the statistics of ligand-binding sites.

**Protein interfaces, surfaces and assemblies:** PDBFisa [30] allows you to interactively explore macromolecular interfaces and surfaces, predict probable quaternary structures (assemblies) and search the PDB for structurally similar interfaces and assemblies.

**Motif [31]-based searches:** PDBMotif [32] allows complex searches of the PDB based on small 3D motifs, sequence motifs in conjunction with ligand environment, secondary structure [20] patterns (such as helix–turn–helix motifs), Prosite [33] patterns and combinations of torsion angles and catalytic sites. PDBMotif also allows you to
search by sequence and perform multiple alignments of sequences or structures.

**Browser of structural knowledge**: PDBeXplore allows you to browse the structural knowledge embodied in the PDB based on classification systems that are familiar to molecular biologists, biochemists and other life scientists including Enzyme Class [34], CATH domains [35], Pfam families [36], sequence and chemical compounds [37].

## Submitting data to the PDB and EMDB

### Submitting to PDB [4]

Scientific journals require authors of structural biology papers to submit their molecular structure data to the PDB. **AutoDep** [38] is our web-based tool for the submission of X-ray crystallography [39] and Nuclear Magnetic Resonance [40] (NMR) spectroscopy structures and data to the PDB and BMRB. Following deposition, expert curators process the submission and send back the completed entry, normally within two working days.

### Submitting to EMDB [41]

**EMDep** [42] is a web-based tool that facilitates the submission of data to the Electron Microscopy Data Bank [43] (EMDB), an archive of high-resolution 3D cryo-electron microscopy data. EMDB contains 3D maps (volumes), masks, images and bibliographic citations, as well as processed primary data where appropriate. As with submissions to the PDB, you obtain a unique identifying code for each submission and you can choose whether to release the data immediately or upon publication in a journal. The deposition system allows you to deposit 3D maps to EMDB and associated coordinate data to the PDB.

## Your feedback

Please tell us what you thought about this course. Your feedback is invaluable and helps us to improve our courses and thus enhance your learning experience.

## Get help and support on PDBe

The EBI's PDBe Team [2] develops and maintains the EBI's Protein Data Bank in Europe [44].

### Support

- Try the PDBe tutorials [45] to find out more about macromolecular structures and how to navigate the PDBe website.
- For general enquiries about the PDB [4], email the pdbhelp [at] ebi.ac.uk (PDB help desk)
- For deposition enquiries, email the pdbdep [at] ebi.ac.uk (PDB deposition)
- For general enquiries about the Electron Microscopy Data Bank, email the emhelp [at] ebi.ac.uk (EMDB [41] help desk)
Reference

Collaborators
PDBe collaborates with the X-ray crystallography [39], Nuclear Magnetic Resonance [40] (NMR) spectroscopy and cryo-Electron Microscopy (EM) communities. To keep abreast of new developments in the NMR community, PDBe has participated in EU projects and continues to contribute to the Collaborative Computational Project for the NMR community [48] (CCPN). PDBe also operates EMDB [43], the international repository for density maps, which are created using high-resolution biological transmission electron microscopy in collaboration with RCSB [5] and Baylor College of Medicine [49]. EMDB contains both macromolecular images and structures reconstructed using the single-particle method and images of sub-cellular regions from electron tomography [50].

Funding
Core support for PDBe is provided by the European Molecular Biology Laboratory [51] and The Wellcome Trust [52]. Specific projects are funded by the European Union [53], the UK Biotechnology and Biosciences Research Council [54](BBSRC) and the National Institutes of Health [55] (NIH).

Contributors
Gary Battle [56]

EMBL-EBI
Project Leader Outreach - Kleywegt team: Protein Data Bank in Europe (EMBL-EBI alumni)

Gary was the outreach coordinator for the Protein Data Bank in Europe (PDBe). He was responsible for helping users with a broad range of backgrounds and interests to make the most of macromolecular structural data. Gary has a PhD is synthetic organic chemistry from the University of Warwick. Before joining the EMBL-EBI, he worked for over 10 years at the Cambridge Crystallographic Data Centre where he gained a wealth of experience in supporting software tools and data resources for pharmaceutical discovery, life science research and materials design.
Sameer Velankar [1]

EMBL-EBI
Team Leader, PDBe content and integration

Sameer [57] is a team leader at the EBI, responsible for content and integration of the Protein Data Bank in Europe (PDBe) resource. He earned his PhD in Structural Biology from the Indian Institute of Science in Bangalore, India in 1997, working on protein crystallographic studies of thymidylate synthase and triose phosphate isomerase. He then joined Dale Wigley's group in Oxford for post-doctoral research on the elucidation of the mechanism of DNA helicase. He has worked with and contributed to all parts of the PDBe team's operations, from annotation of newly deposited structures to the development of advanced PDBe services.

Source URL: https://www.ebi.ac.uk/training/online/course/pdbe-quick-tour

Links
[1] https://www.ebi.ac.uk/training/online/trainers/sameer
[4] https://www.ebi.ac.uk/training/online/glossary/pdb
[8] https://www.ebi.ac.uk/training/online/glossary/small-molecules
[9] https://www.ebi.ac.uk/training/online/glossary/pdbe
[15] https://www.ebi.ac.uk/training/online/glossary/smiles
[20] https://www.ebi.ac.uk/training/online/glossary/secondary-structure
[25] https://www.ebi.ac.uk/training/online/glossary/term/532
[26] https://www.ebi.ac.uk/training/online/glossary/term/429
[27] https://www.ebi.ac.uk/training/online/glossary/electron-microscopy
[29] http://www.ebi.ac.uk/pdbe-srv/pdbeckem/