

MS12-P05 | PDBe KNOWLEDGE BASE (PDBe-KB) – INFRASTRUCTURE FOR FAIR

STRUCTURAL AND FUNCTIONAL ANNOTATIONS

Velankar, Sameer (Protein Data Bank in Europe (PDBe), Cambridge, GBR)

Growth in the number and complexity of macromolecular structures deposited to the Protein Data Bank (pdb.org) has continually increased, due to rapid advances in structure determination techniques. The PDB archive currently contains ~150,000 structures, referencing more than 45,000 unique UniProtKB entries, which facilitate greater understanding of protein function and mechanism. In addition to the PDB, over 400 further data resources are able to derive additional annotations related to structures. This both enhances the inherent value of macromolecular structures and provides their biological context. PDBe-KB (Protein Data Bank in Europe - Knowledge Base; pdbe-kb.org), established in 2018, is an international, community-driven resource with the primary goal of collating and making available structural and functional annotations contributed by partner resources. It is managed by the Protein Data Bank in Europe team at EMBL-EBI.

PDBe-KB partner data resources provide manually curated or computationally predicted annotations for ligand binding sites, catalytic sites, protein-protein interfaces and post-translational modification sites, as well as selected physico-chemical parameters (solvent accessibility, residue depth). These enriched data are combined with manual curations of functional sites and quaternary structure assemblies and other annotations available in PDBe. PDBe-KB facilitates data exchange and allows insights into the biological function by providing a comprehensive view of the functional context of the protein structure. Functional annotations are made available on PDBe-KB web pages, programmatically via an API, and as a distributed Neo4J graph database. As a first example, PDBe-KB has launched aggregated views of structure data for UniProtKB accessions, referenced by PDB structures.