

Trends in macromolecular structure data across 50 years of the PDB

Yuhe Liang¹

¹*RCSB Protein Data Bank*

yuhe.liang@rcsb.org

In 2021, we celebrate 50 years of the Protein Data Bank (PDB) archive - one of the longest-running open access scientific databases - managed collaboratively by the wwPDB (RCSB PDB, PDBe, PDBj, and BMRB). Throughout these 50 years, there has been significant evolution of the data in the PDB archive. This evolution has been driven by a number of factors including development in structural determination techniques, adaptation of biocuration practices, and increase in data capture via updated file formats. Here we will discuss some of the key trends in data across the PDB archive, highlighting how structural biology data has changed over time and how wwPDB biocuration practices have adapted to handle these changes. In addition to long term trends, we will also identify more recent trends and challenges and demonstrate how wwPDB biocuration practices are evolving to meet these challenges in the future.