The data universe of structural biology

Helen M. Berman,[a,b]* Brinda Vallat[a] and Catherine L. Lawson[a]

[a]Department of Chemistry and Chemical Biology, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA, [b]Department of Biological Sciences and Bridge Institute, University of Southern California, Los Angeles, CA 90089, USA, and [c]Institute for Quantitative Biomedicine, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA. *Correspondence e-mail: berman@rcsb.rutgers.edu

The Protein Data Bank (PDB) has grown from a small data resource for crystallographers to a worldwide resource serving structural biology. The history of the growth of the PDB and the role that the community has played in developing standards and policies are described. This article also illustrates how other biophysical communities are collaborating with the worldwide PDB to create a network of interoperating data resources. This network will expand the capabilities of structural biology and enable the determination and archiving of increasingly complex structures.

Keywords: Protein Data Bank; structural biology; X-ray crystallography; data resources; data standards

1. Introduction

2. The Protein Data Bank

2.1. Early history

Figure 1

2.2. Deposition guidelines

2.3. mmCIF standard

2.4. Validation

2.5. Current state of the PDB

Figure 2

Figure 3

Figure 4

3. Other structural biology databases

Figure 5

Figure 6

4. Perspectives

Acknowledgements

References