What is Archived in PDB

Amitava Roy

Some materials are borrowed from the PDB 101 resources
and from Dr. Shuchismita Dutta

Today’s Instructor

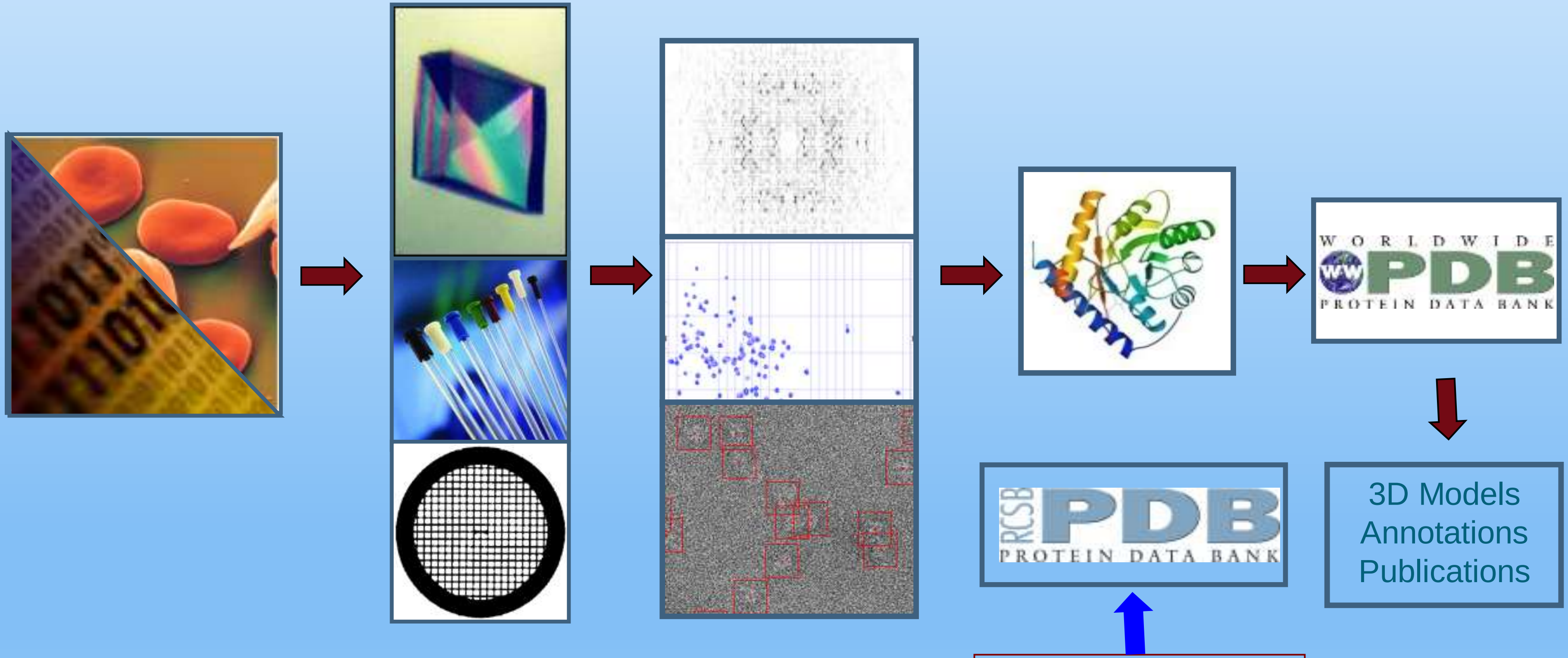
Amitava Roy, PhD

- Faculty affiliate at the Department Of Biomedical And Pharmaceutical Sciences, University of Montana
- Research interest: Computational Structural Biology and Drug Development
- Email: amitava.roy@umontana.edu

Objectives

- What is Archived in the PDB?

The Data Pipeline



Genomic Based Target Selection

Isolation, Expression, Purification, Crystallization

Data Collection

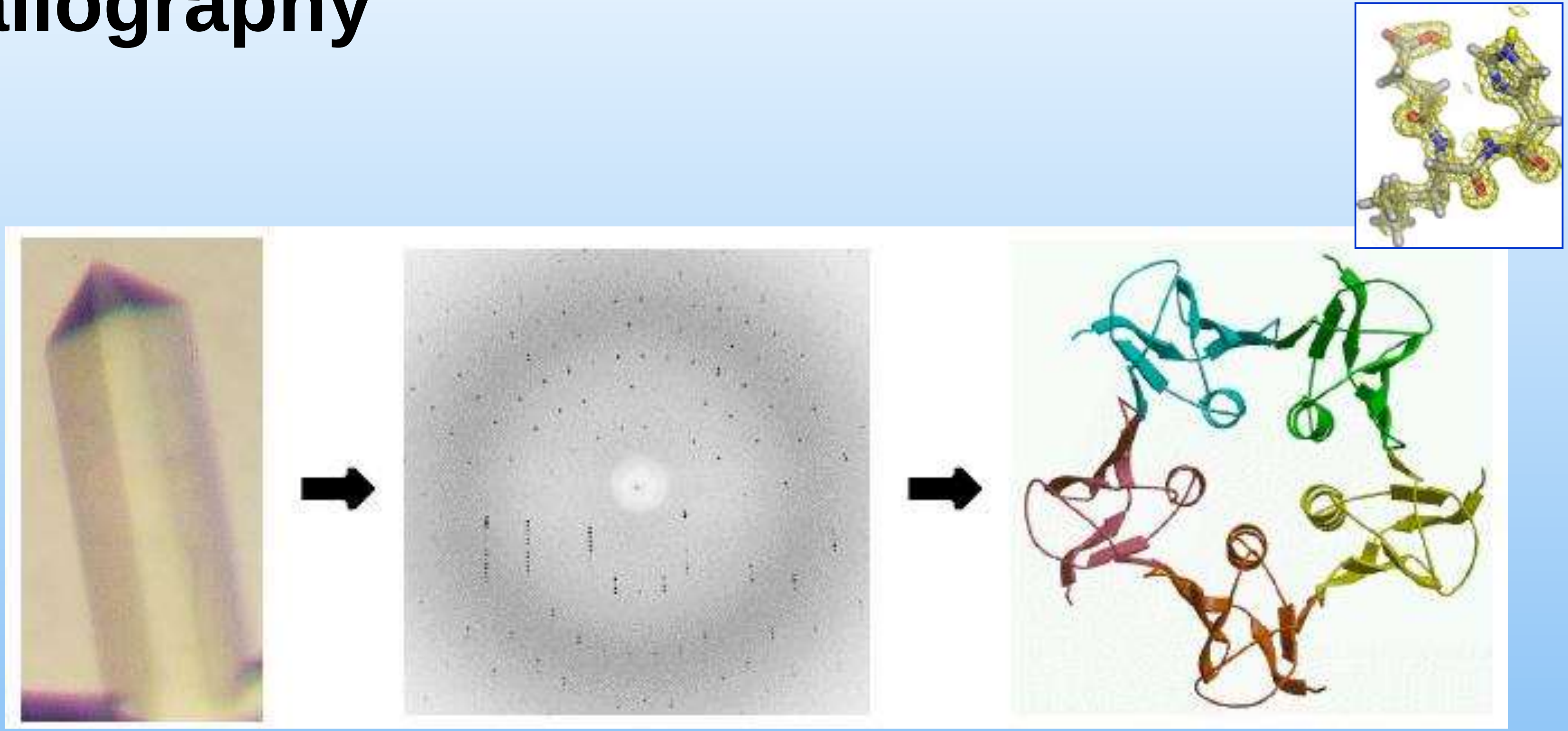
Structure Determination

PDB Deposition & Release

3D Models
Annotations
Publications

You come here

X-ray Crystallography



Target selection
Protein production
Crystallization

→

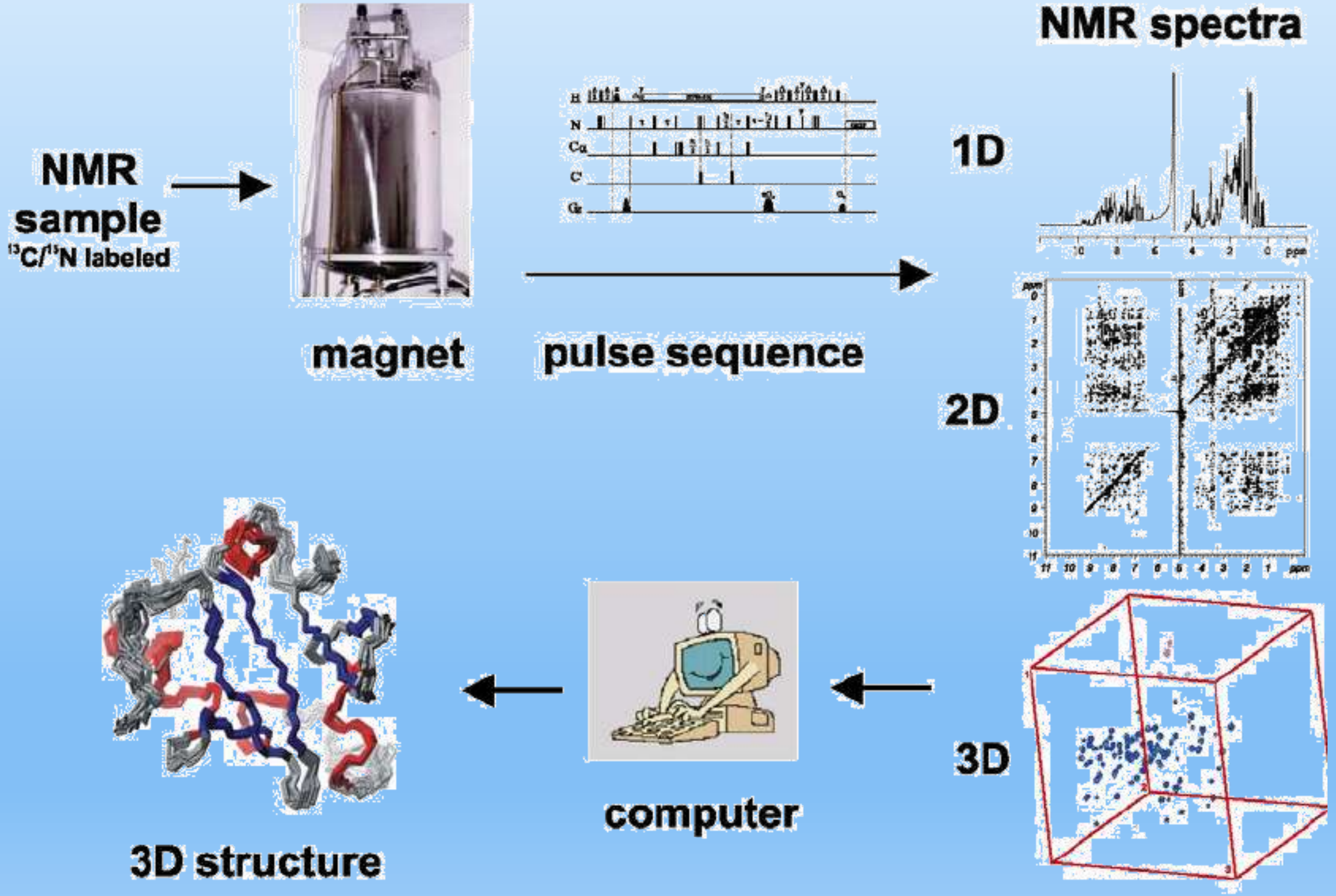
Data collection
Phasing

→

Structure calculation
Refinement
Validation
Deposition to PDB
Annotation

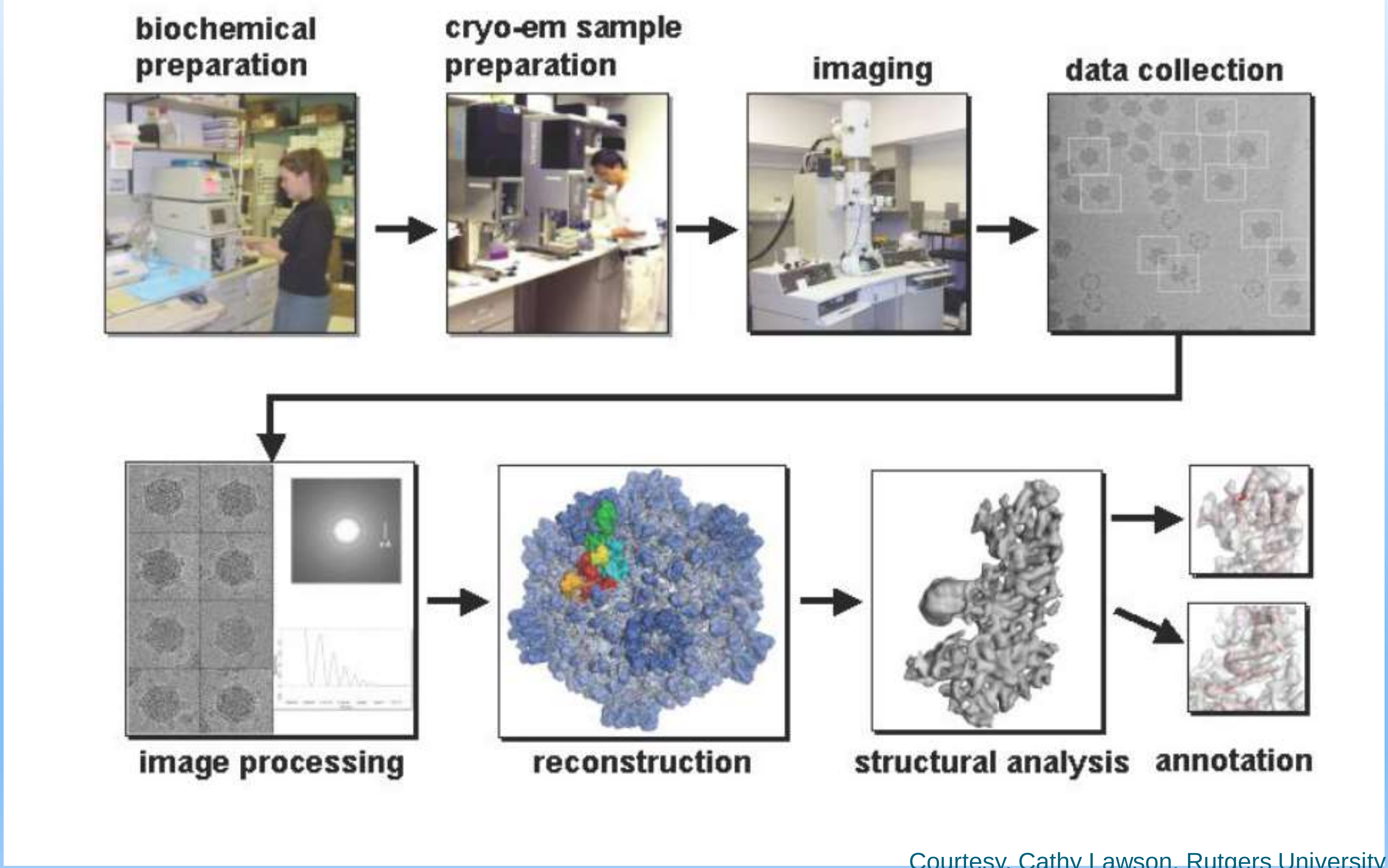
http://www-structmed.cimr.cam.ac.uk/Course/Overview/Overview.html#methods

Nuclear Magnetic Resonance (NMR)



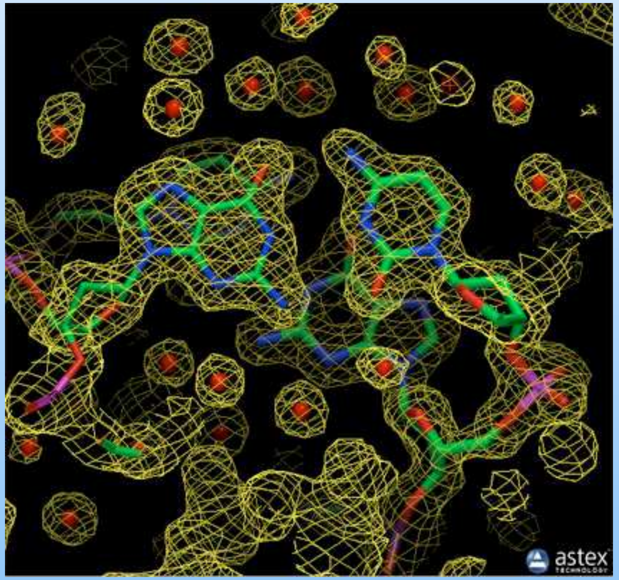
http://www.embl.de/nmr/sattler/teaching

Electron Microscopy (EM)

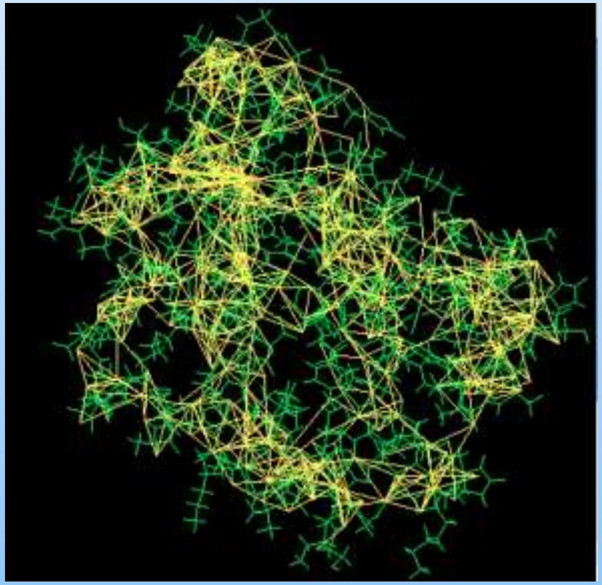


Courtesy, Cathy Lawson, Rutgers University

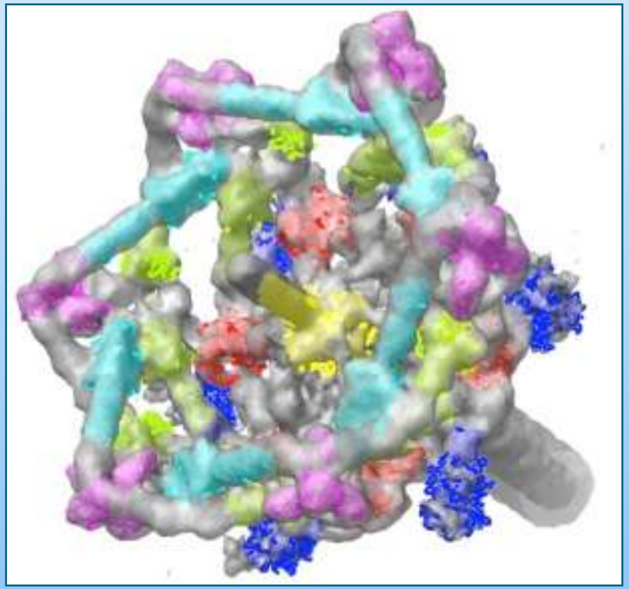
Experimental Data from Different Methods



X-ray Crystallography: Electron density from a structure of DNA is shown here (PDB entry 196d), along with the atomic model



NMR: Restraints used to solve the structure of a small monomeric hemoglobin. The protein (1vre and 1vrf) is shown in green, and restraints are shown in yellow.



EM: Tail of the T4 bacteriophage. Surface rendering of the EM data (emd-1048) with atomic coordinates from PDB entries 1pdf, 1pdi, 1pdl, 1pdm, 1pdp, and 2fl8.

Expanding the limits of PDB

- There are many millions of proteins whose structures have not yet been solved.
- In 2020, two different projects [AlphaFold2 (Jumper, J. et al., 2021, Varadi M. et al., 2022) and RoseTTAFold (Baek et al., 2021)] used artificial intelligence (AI) and machine learning (ML) to predict protein structures from their sequences successfully.
- To make it easier for users to query, organize, visualize, analyze and compare experimental and predicted structures alongside each other, RCSB PDB has integrated Computed structure models (CSMs) from a few specific resources.

What is in the PDB?

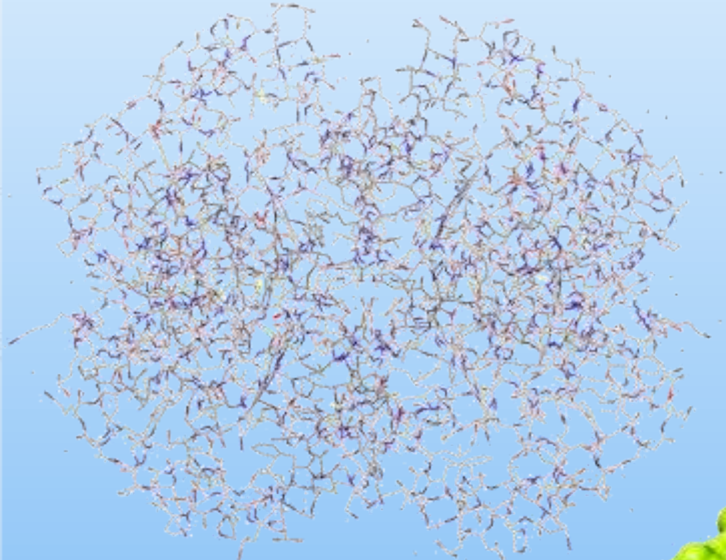
- Coordinate and experimental data files
- Details about sample preparation, data collection and structure solution
- Sequence(s) of polymers (proteins and nucleic acids) in the structure
- Information about ligands in the structure
- Coordinate from computational structure models

Download/Use directly

- Links to various resources that describe the sequence, function and other properties of the molecule.
- Classification of structures by sequence, structure, function and other criteria

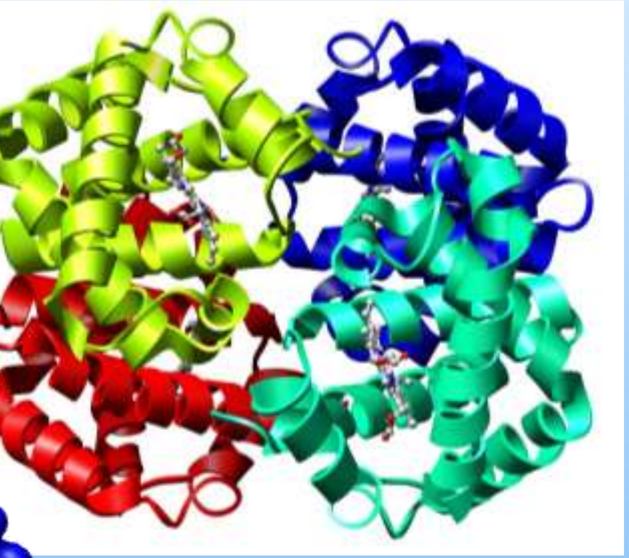
Link to/Find similar

Visualization Conventions



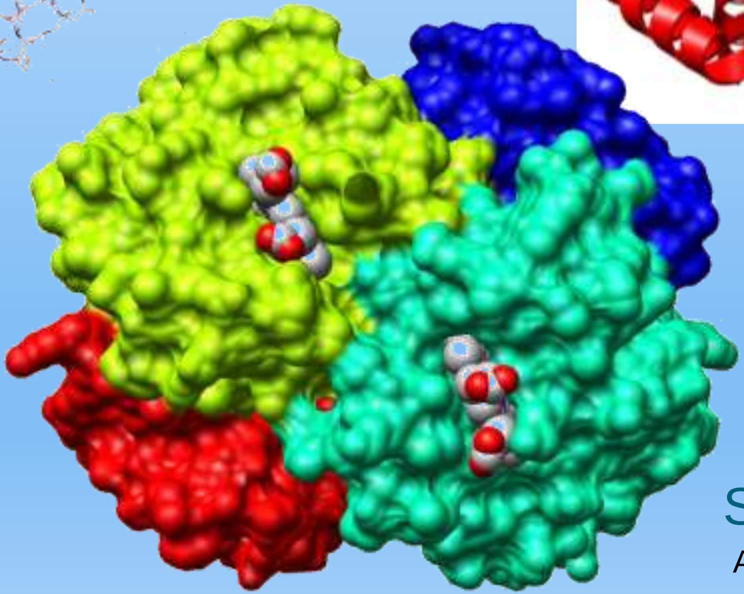
Wireframe

All atoms seen



Ribbons

Only backbone
C-alpha atoms
seen



Spacefill

All atoms seen