Structural Biology

PROPEL

Smriti Sangwan, PhD  22nd February 2024
Proteins are the building blocks of life.
Proteins are the building blocks of life

Protein Misfolding is linked to numerous diseases:
- Neurodegeneration
- Cancer
- Metabolic Diseases

20,687 protein-coding genes

42 million protein molecules per cell

3.9 million new proteins synthesized per minute
Proteins have different shapes and dimensions
Proteins are made up of amino acids.
TWENTY-ONE PROTEINOGENIC α-AMINO ACIDS

Side chain charge at physiological pH 7.4

pKₐ values shown italicized

Positive

Negative

A. Amino Acids with Electrically Charged Side Chains

Arginine [Arg R]
Histidine [His H]
Lysine [Lys K]
Aspartic Acid [Asp D]
Glutamic Acid [Glu E]

B. Amino Acids with Polar Uncharged Side Chains

Serine [Ser S]
Threonine [Thr T]
Asparagine [Asn N]
Glutamine [Gln Q]
Cysteine [Cys C]
Selenocysteine [Sec U]
Glycine [Gly G]
Proline [Pro P]

C. Special Cases

D. Amino Acids with Hydrophobic Side Chains

Alanine [Ala A]
Valine [Val V]
Isoleucine [Ile I]
Leucine [Leu L]
Methionine [Met M]
Phenylalanine [Phe F]
Tyrosine [Tyr Y]
Tryptophan [Trp W]
Overview of Protein Structure:

Main Interactions:

Different Levels of Organization:

Bond Types

**Hydrophobic Interactions:**
These amino acids orient themselves towards the center of the polypeptide to avoid the water

**Disulfide Bridge:** The amino acid cysteine forms a bond with another cysteine through its R group

**Hydrogen Bonds:** Polar “R” groups on the amino acids form bonds with other Polar R groups

**Hydrophilic Interactions:**
These amino acids orient themselves outward to be close to the water

**Ionic Bonds:** Positively charged R groups bond together

Primary protein structure
sequence of a chain of amino acids

Secondary protein structure
hydrogen bonding of the peptide backbone causes the amino acids to fold into a repeating pattern

Tertiary protein structure
three-dimensional folding pattern of a protein due to side chain interactions

Quaternary protein structure
protein consisting of more than one amino acid chain
What can you do with protein structural information:

- Understand their function
- Design mutations to test their function
- Design specific drugs
Structure-based drug design for HIV-AIDS
Methods to determine protein structures:

X-ray crystallography

Nuclear Magnetic Resonance (NMR)

Cryogenic Electron Microscopy (CryoEM)
X-ray crystallography

Basic Principle:

Experiment:
NMR

Basic Principle:

- No external magnetic field
- Apply external magnetic field $B_0$

Experiment:

1. Purified protein
2. NMR sample preparation
3. Data acquisition
4. Spectral processing
5. Structural analysis
Nobel Prizes for the applications of NMR

1991
Richard R. Ernst  
b. 1933  
ETH, Switzerland

2002
Kurt Wüthrich  
b. 1938  
ETH, Switzerland

2003
Paul C. Lauterbur  
1929-2007  
U. Of Illinois, IL, US

2003
Peter Mansfield  
b. 1933  
U. Of Nottingham

Multidimensional NMR

MRI

3D structure of biomolecules
CryoEM

Basic Principle:

CRYO-ELECTRON MICROSCOPY
A beam of electron is fired at a frozen protein solution. The emerging scattered electrons pass through a lens to create a magnified image on the detector, from which their structure can be worked out.

Experiment:

Purified protein  Freezing / Negative staining  EM data collection  Particle picking  Particle alignment and classification  3D model reconstruction  Model refinement
2017 Nobel Laureates in Chemistry

Jacques Dubochet (University of Lausanne, Switzerland)
Joachim Frank (Columbia University, New York)
Richard Henderson (MRC Laboratory of Molecular Biology, Cambridge, U.K.)
<table>
<thead>
<tr>
<th></th>
<th>Pros</th>
<th>Cons</th>
<th>Sample Types</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>X-ray crystallography</strong></td>
<td>Established High Resolution Broad Molecular Weight Range Easy for model building</td>
<td>Static Crystalline State Protein needs to crystallize Not suitable for dynamic interactions Challenging for protein complexes</td>
<td>Soluble Proteins, Membrane Proteins, Small Molecules DNA/RNA and Protein Complexes</td>
<td>High</td>
</tr>
<tr>
<td><strong>NMR</strong></td>
<td>High Resolution 3D structure in solution Protein dynamics</td>
<td>High Sample Purity Difficult Sample Preparation Size limit</td>
<td>MW below 40-50 kDA Water Soluble Samples</td>
<td>High</td>
</tr>
<tr>
<td><strong>CryoEM</strong></td>
<td>Low amounts of sample Structure in native state</td>
<td>Size Limitation Costly EM equipment (6M $) Resolution limited</td>
<td>&gt;100kDa Membrane Proteins Large assemblies like virions, ribosomes</td>
<td>&gt;3.5 Angstrom</td>
</tr>
</tbody>
</table>
Structural Data Repositories:

Protein Data Bank: https://www.rcsb.org/

EM Data Bank: https://www.ebi.ac.uk/emdb/

For each PDB deposition: 4 letter code
EMDB: 4 or 5 letter code

Can you find the code for the structure in the paper and look it up?
3FBV

Crystal structure of the oligomer formed by the kinase-ribonuclease domain of Ire1

PDB DOI: https://doi.org/10.2210/pdb3FBV/pdb

Classification: TRANSFERASE, HYDROLASE

Organism(s): Saccharomyces cerevisiae S288C

Expression System: Escherichia coli

Mutation(s): None

Deposited: 2008-11-19  Released: 2008-12-16


Experimental Data Snapshot

Method: X-RAY DIFFRACTION

Resolution: 3.20 Å

R-Value Free: 0.283

R-Value Work: 0.235

R-Value Observed: 0.237

Ligand Structure Quality Assessment

Worse 0 Better 1

This is version 1.3 of the entry. See complete history.
Questions?