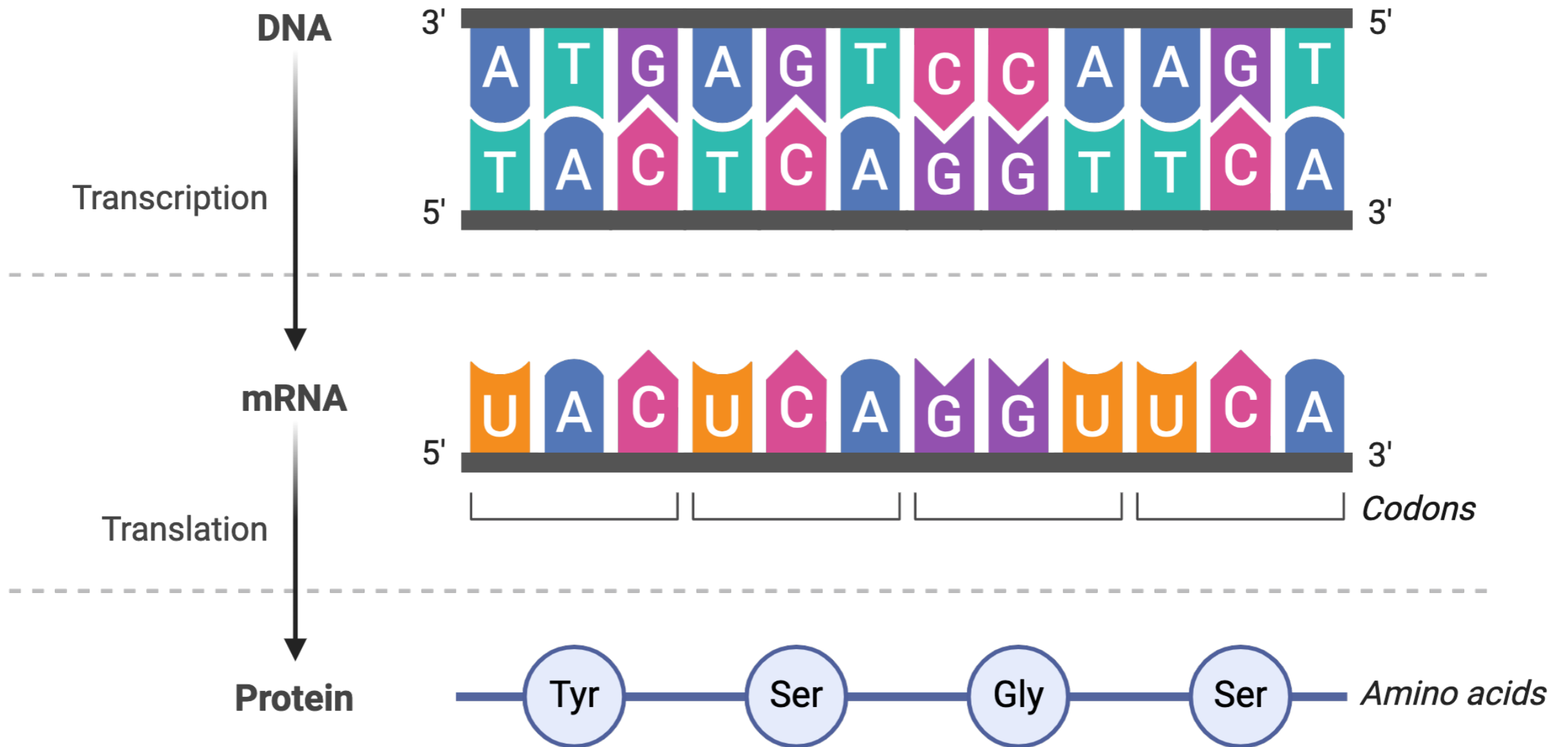


# **Structural Biology**

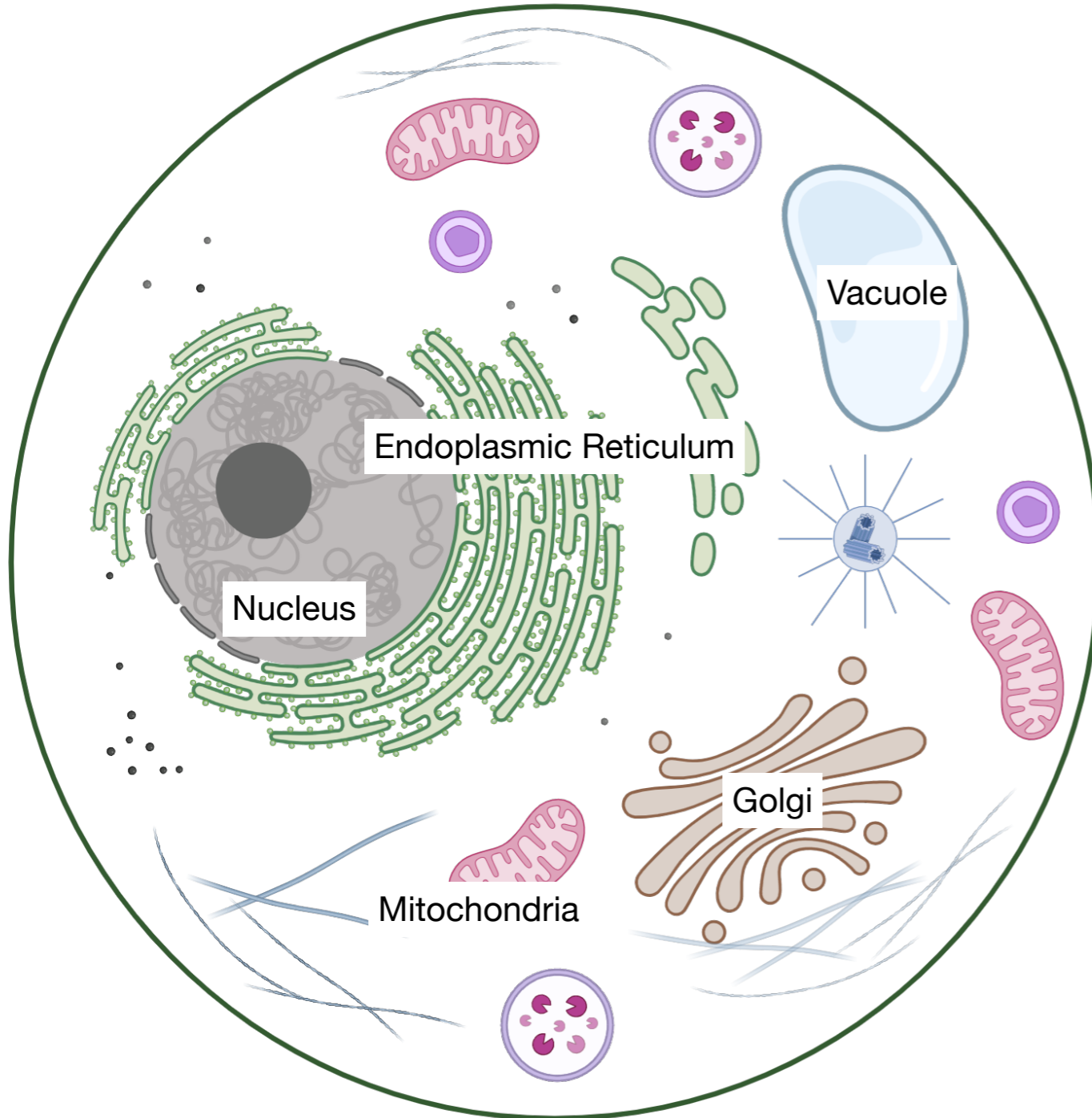
**PROPEL**

**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

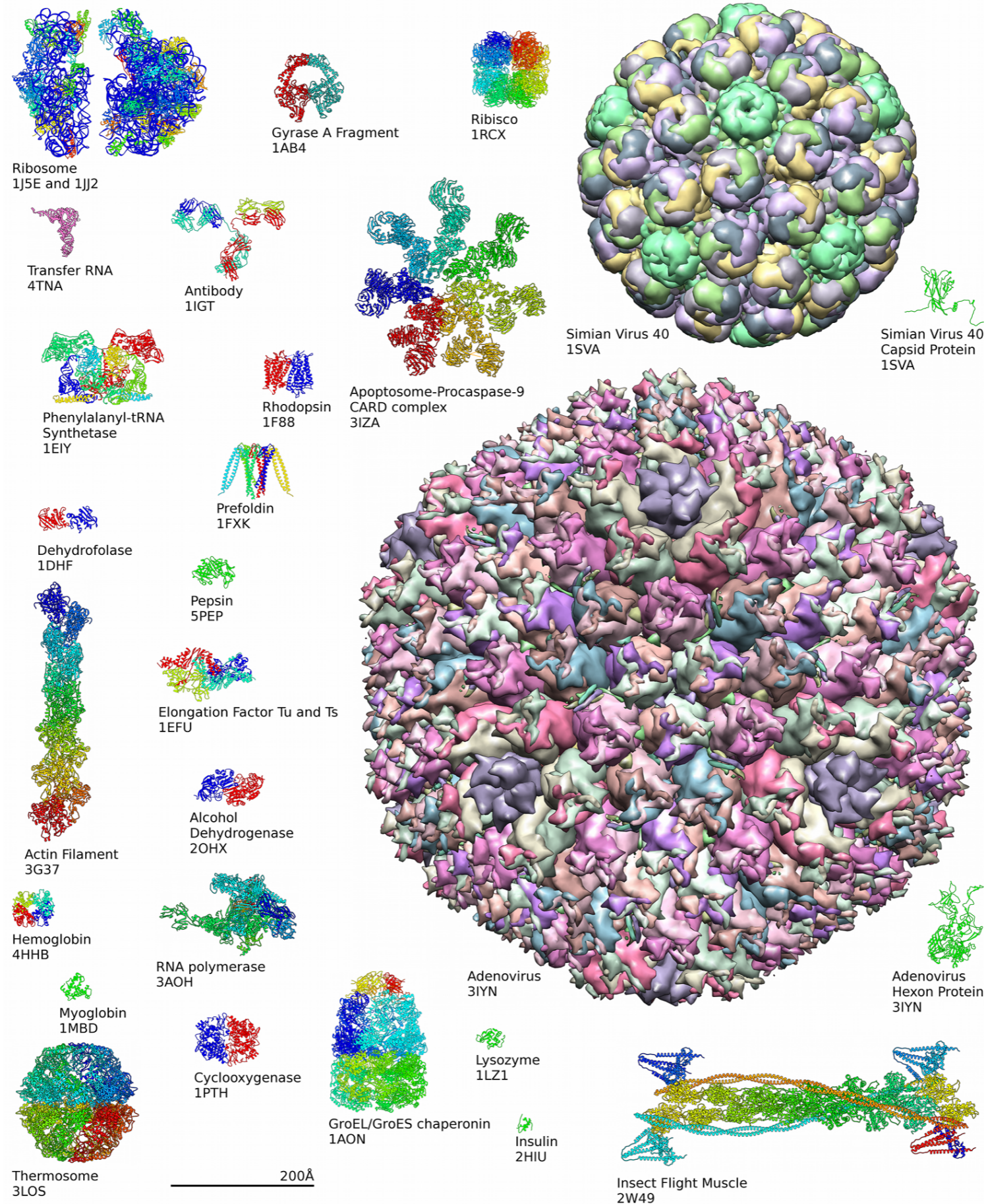
Metabolic Diseases

**20,687 protein-coding genes**

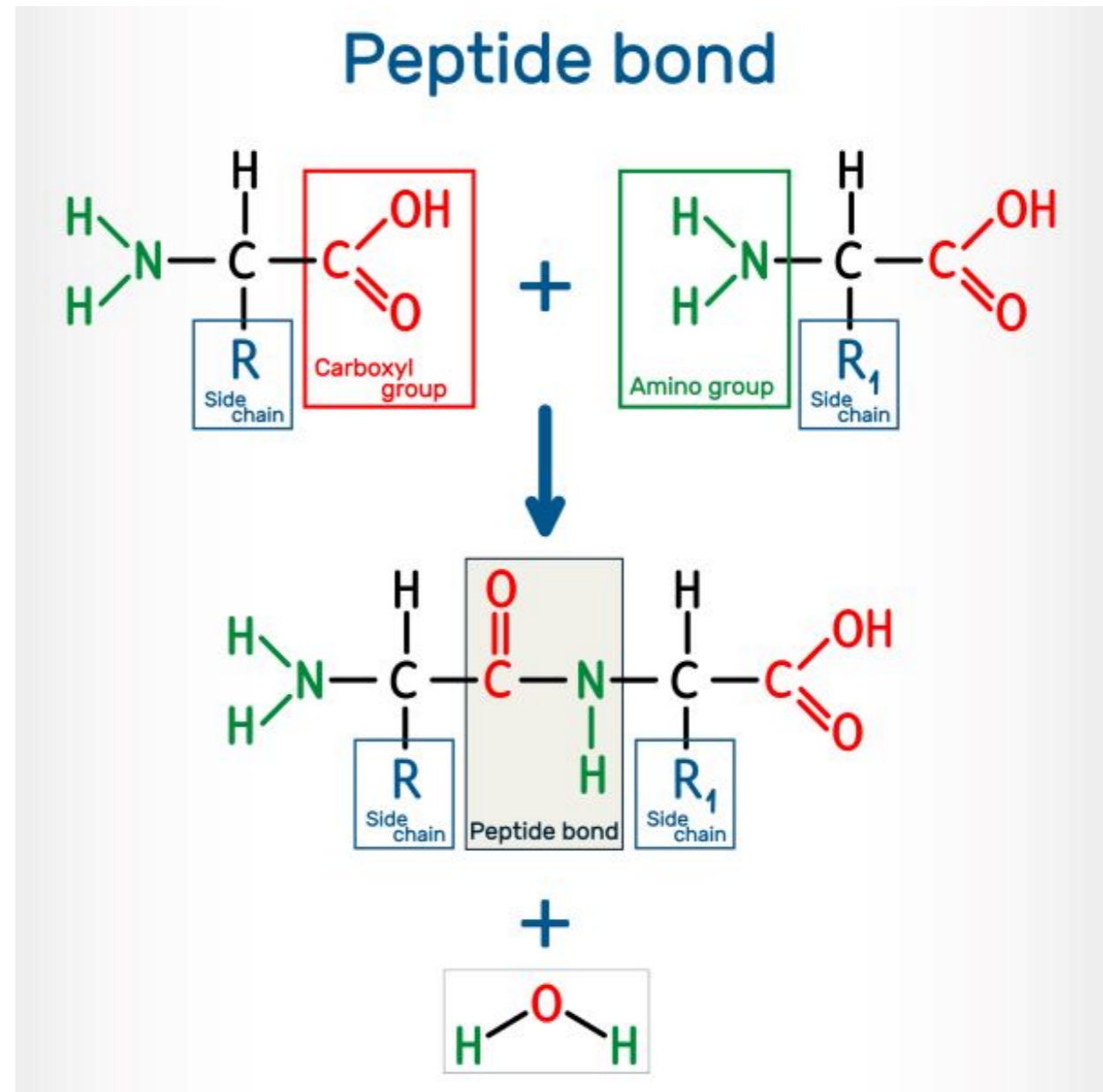
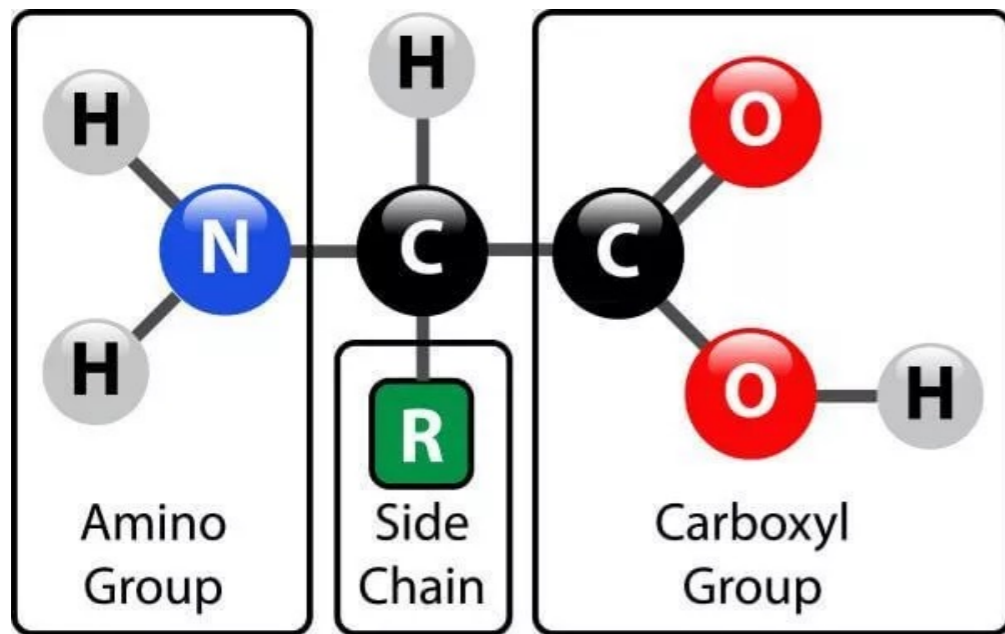
**42 million protein molecules per cell**

**3.9 million new proteins synthesized per minute**

# Proteins have different shapes and dimensions



# Proteins are made up on amino acids



**TWENTY-ONE  
PROTEINOGENIC  
 $\alpha$ -AMINO ACIDS**

Side chain charge  
at physiological  
pH 7.4

$pK_a$  values shown  
italicized

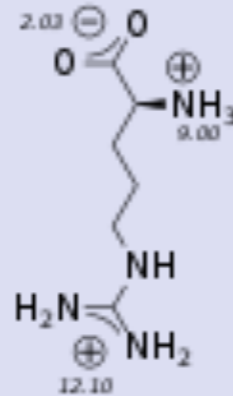
⊕ Positive  
⊖ Negative

*A. Amino Acids with Electrically Charged Side Chains*

Positive

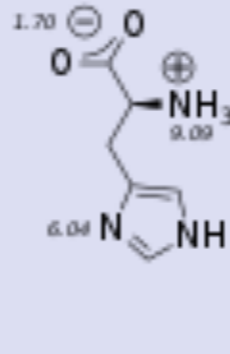
Arginine

Arg R



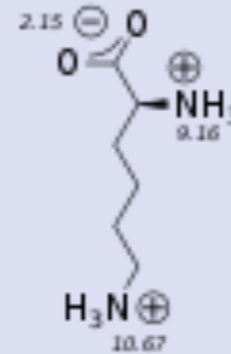
Histidine

His H



Lysine

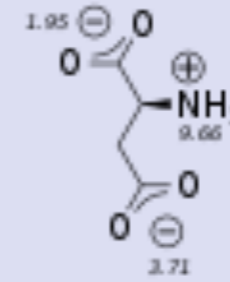
Lys K



Negative

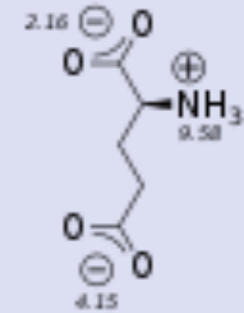
Aspartic Acid

Asp D



Glutamic Acid

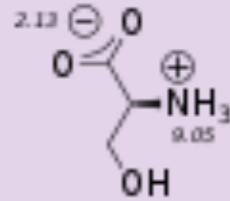
Glu E



*B. Amino Acids with Polar Uncharged Side Chains*

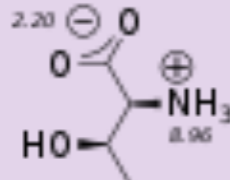
Serine

Ser S



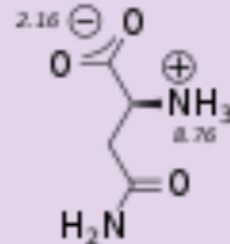
Threonine

Thr T



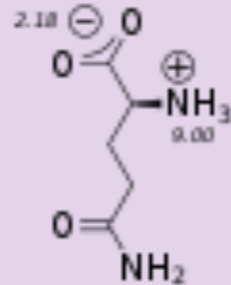
Asparagine

Asn N



Glutamine

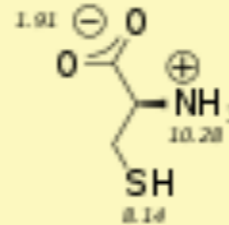
Gln Q



*C. Special Cases*

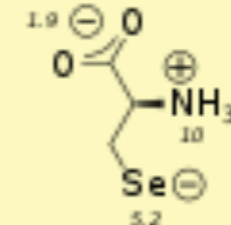
Cysteine

Cys C



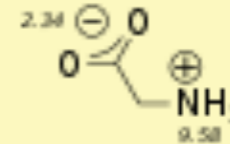
Selenocysteine

Sec U



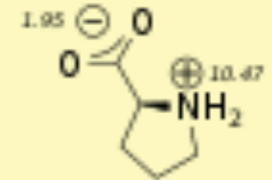
Glycine

Gly G



Proline

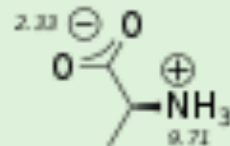
Pro P



*D. Amino Acids with Hydrophobic Side Chains*

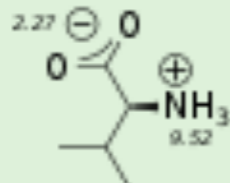
Alanine

Ala A



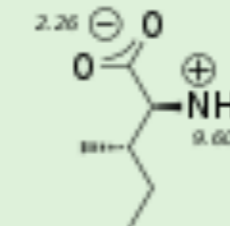
Valine

Val V



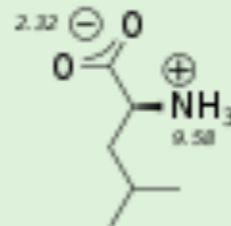
Isoleucine

Ile I



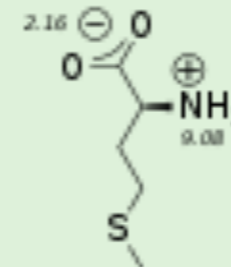
Leucine

Leu L



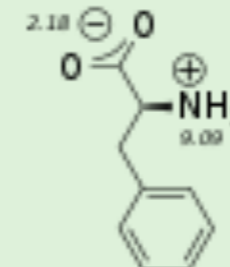
Methionine

Met M



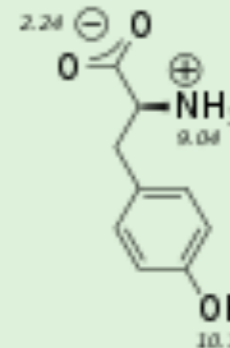
Phenylalanine

Phe F



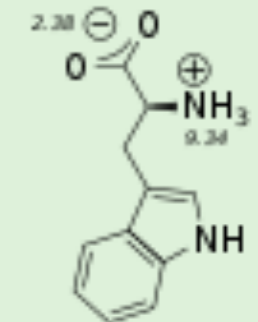
Tyrosine

Tyr Y



Tryptophan

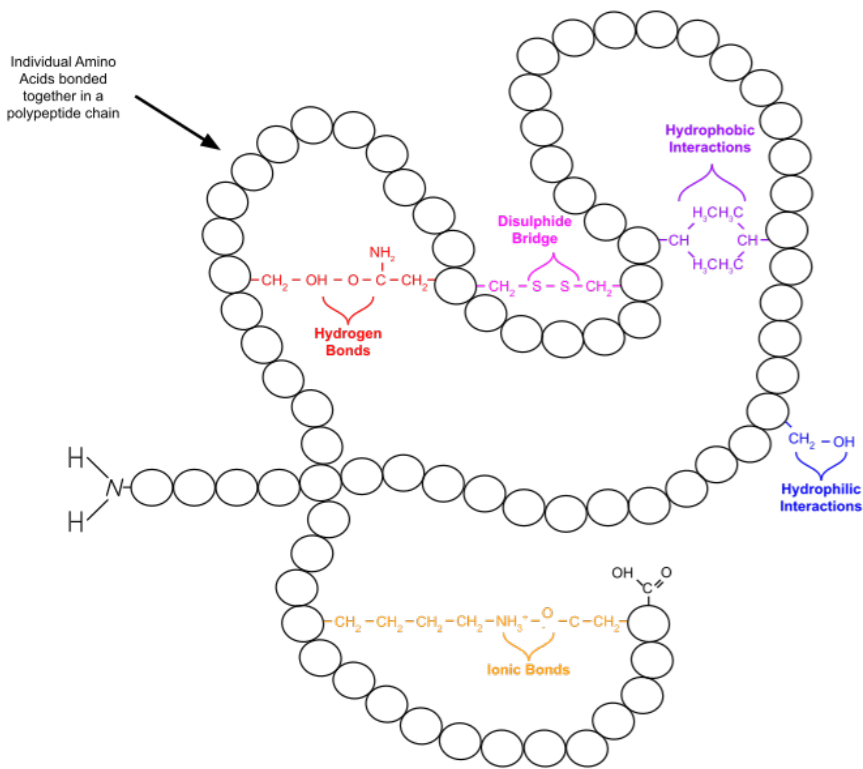
Trp W



# Overview of Protein Structure:

## Main Interactions

## Different Levels of Organization:



## Bond Types

### Hydrophobic Interactions:

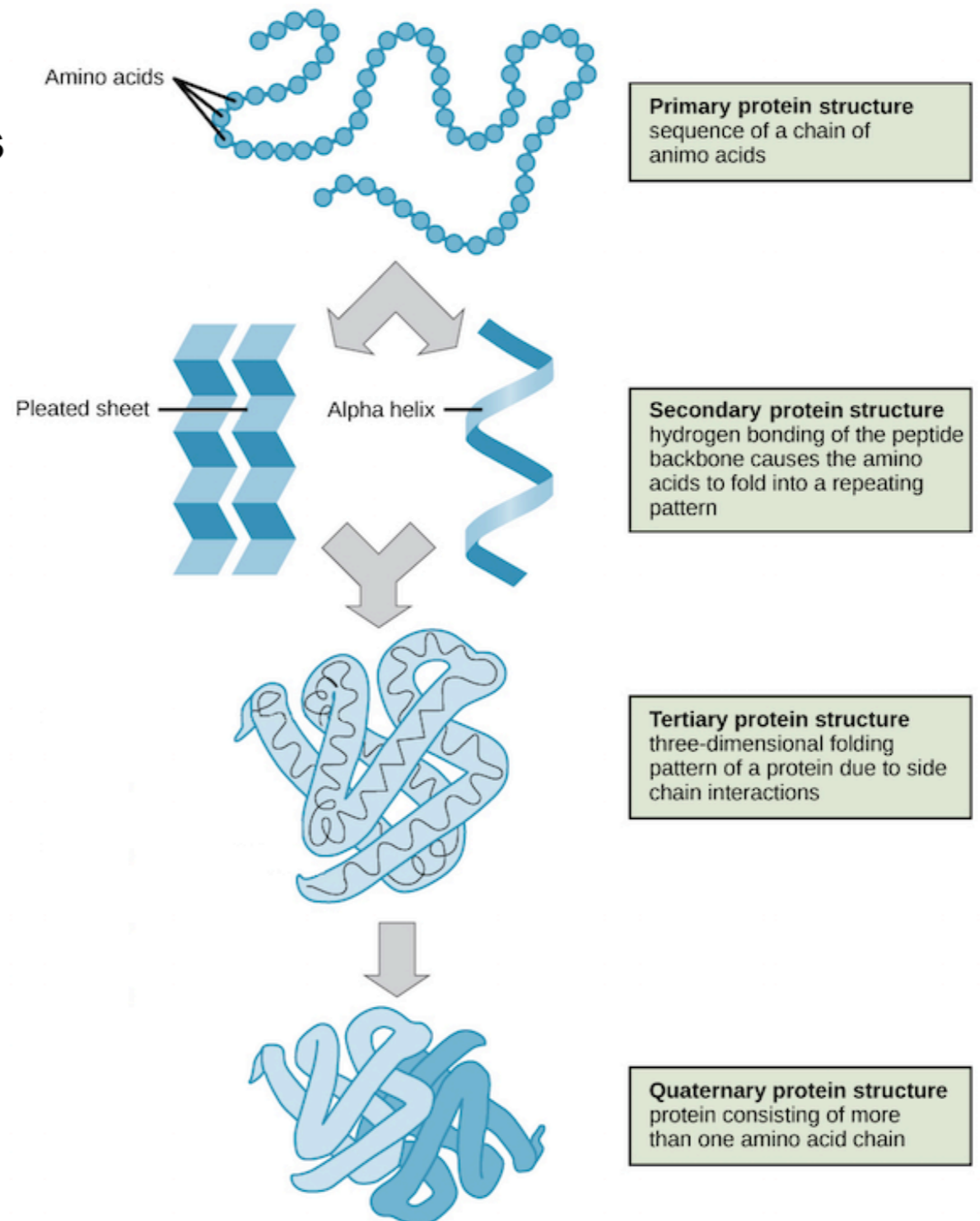
These amino acids orient themselves towards the center of the polypeptide to avoid the water

**Disulphide 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



# What can you do with protein structural information:

Understand their function



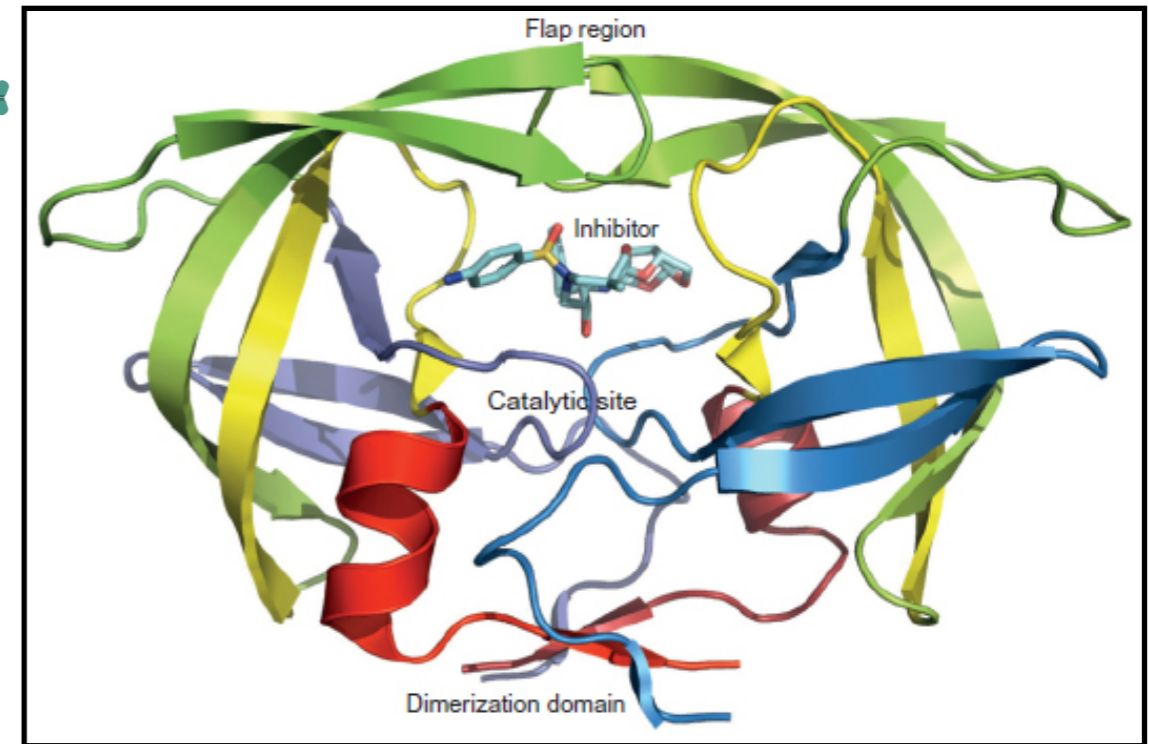
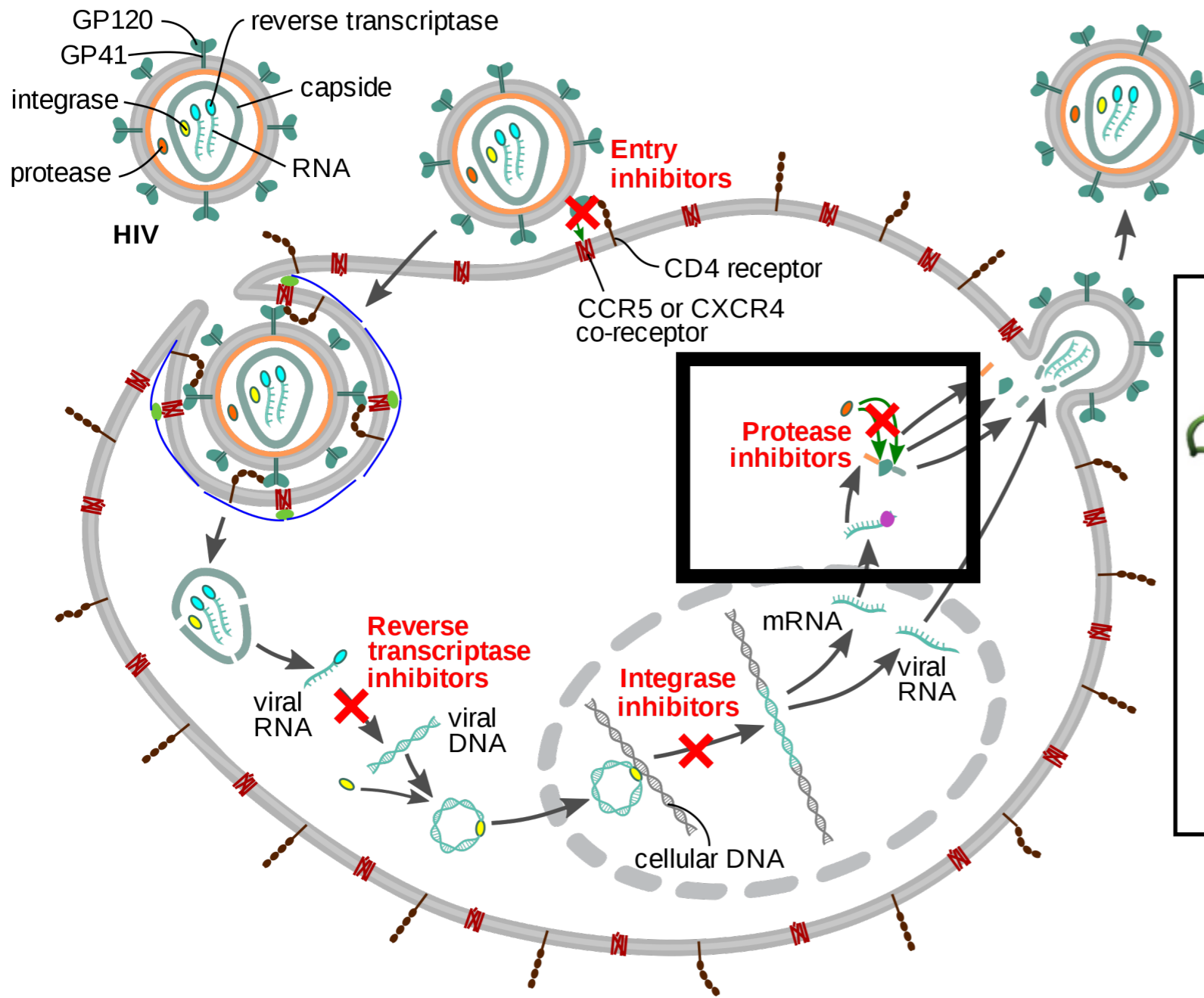
Design mutations to test their function



Design specific drugs



# Structure-based drug design for HIV-AIDS

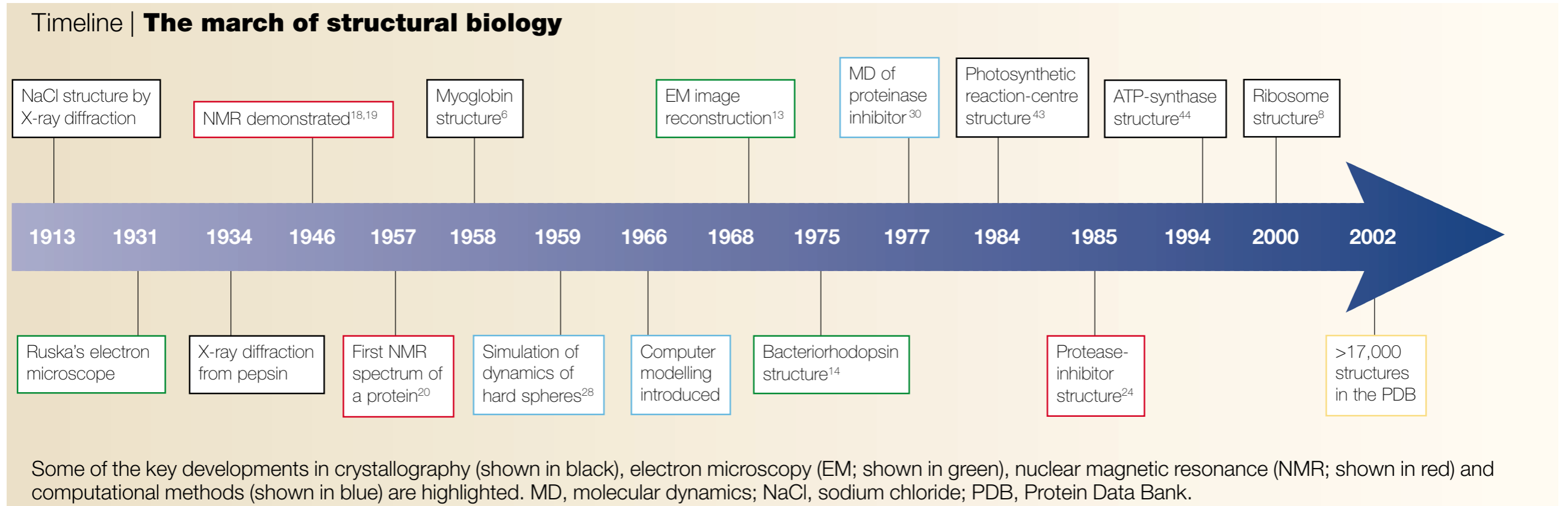


# Methods to determine protein structures:

X-ray crystallography

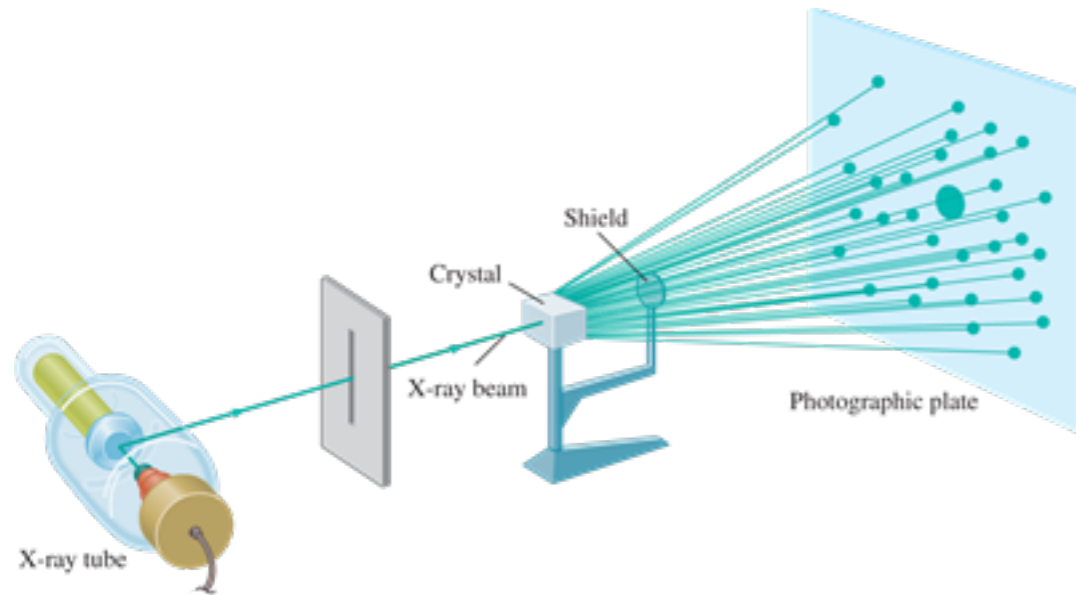
Nuclear Magnetic Resonance (NMR)

Cryogenic Electron Microscopy (CryoEM)

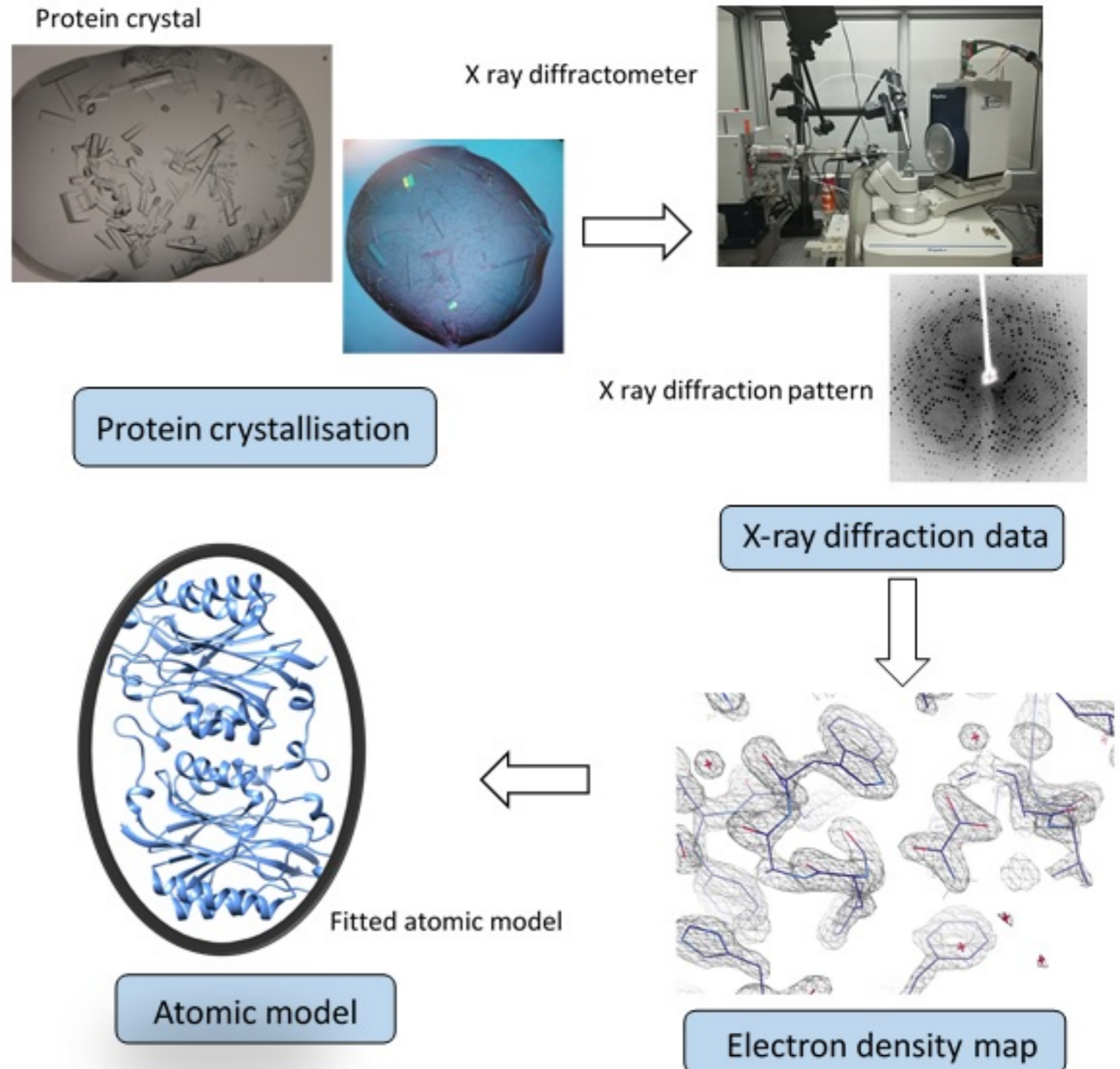


# X-ray crystallography

## Basic Principle:

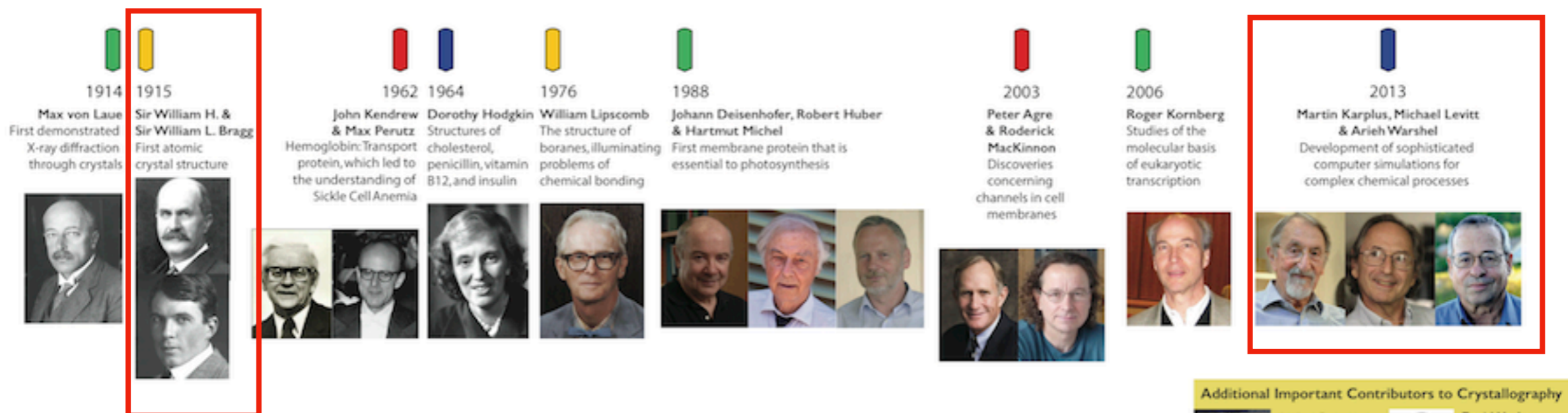
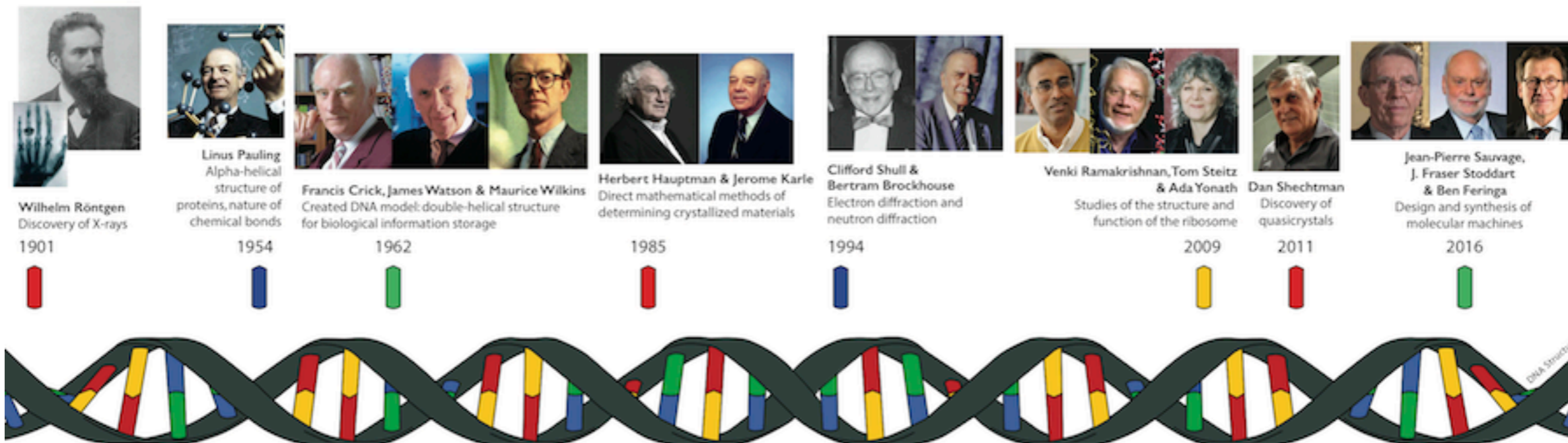


## Experiment:



# Highlights of the Many Nobel Prizes Awarded to Crystallographers

See a complete list of winners at [iucr.org/people/nobel-prize](http://iucr.org/people/nobel-prize)

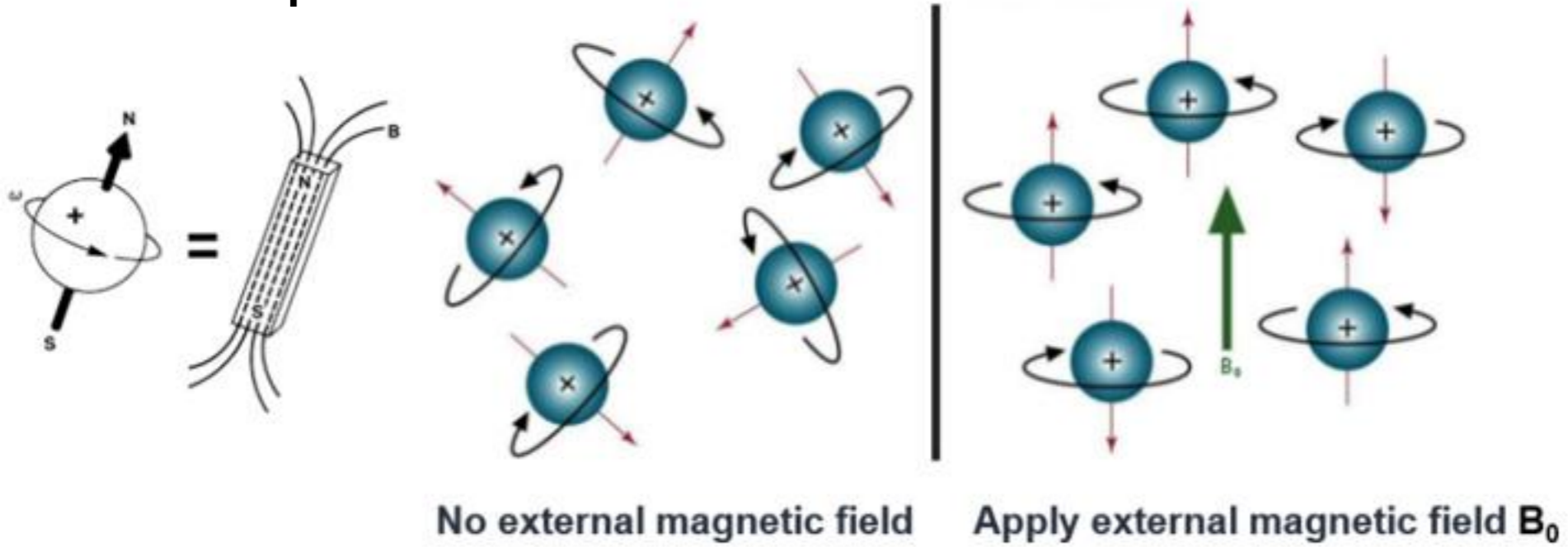


**Additional Important Contributors to Crystallography**

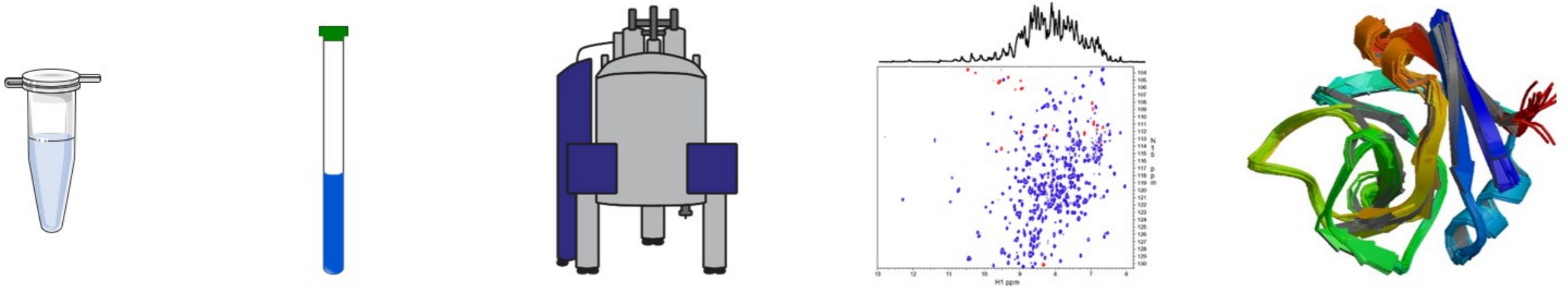
- Arthur Patterson** - The Patterson Function (equation) gives a map of the vectors between atoms
- David Harker** - Applied Patterson's map to identify planes and sections on different axes in molecular structures

# NMR

## Basic Principle:



## Experiment:



Purified protein

NMR sample preparation

Data acquisition

Spectral processing

Structural analysis

## Nobel Prizes for the applications of NMR

1991



Richard R. Ernst  
b. 1933  
ETH, Switzerland

2002



Kurt Würthrich  
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

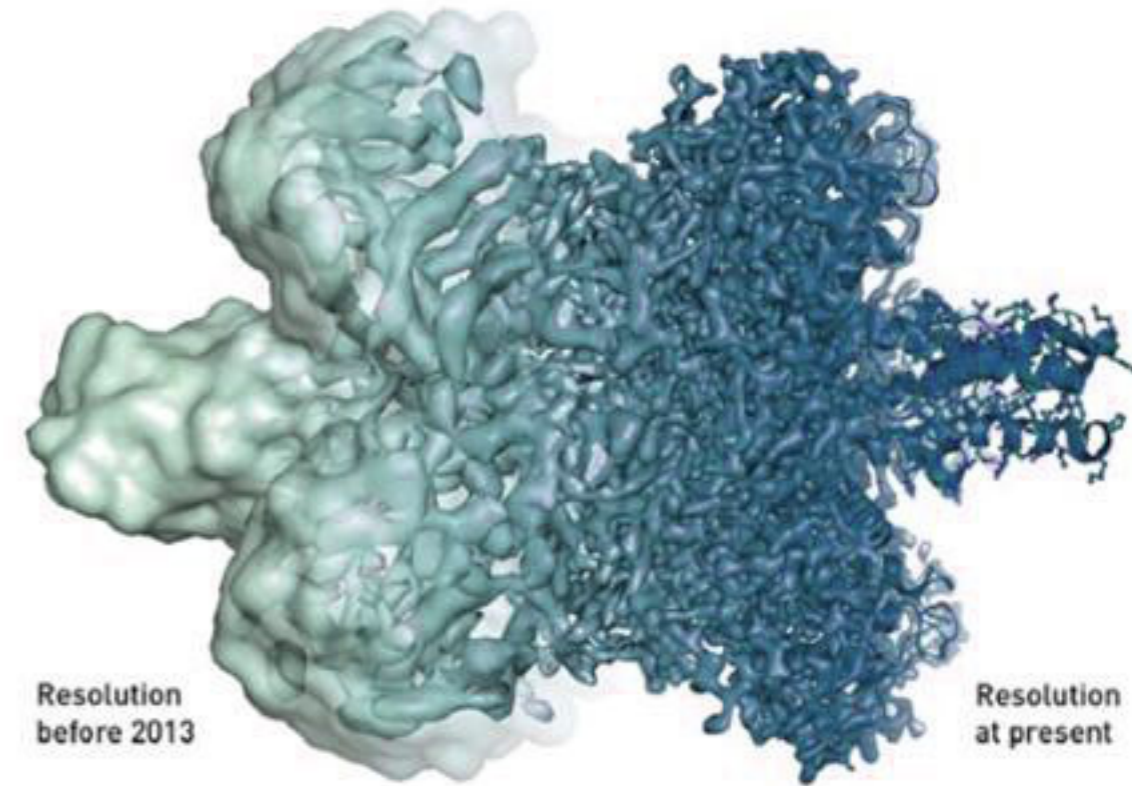
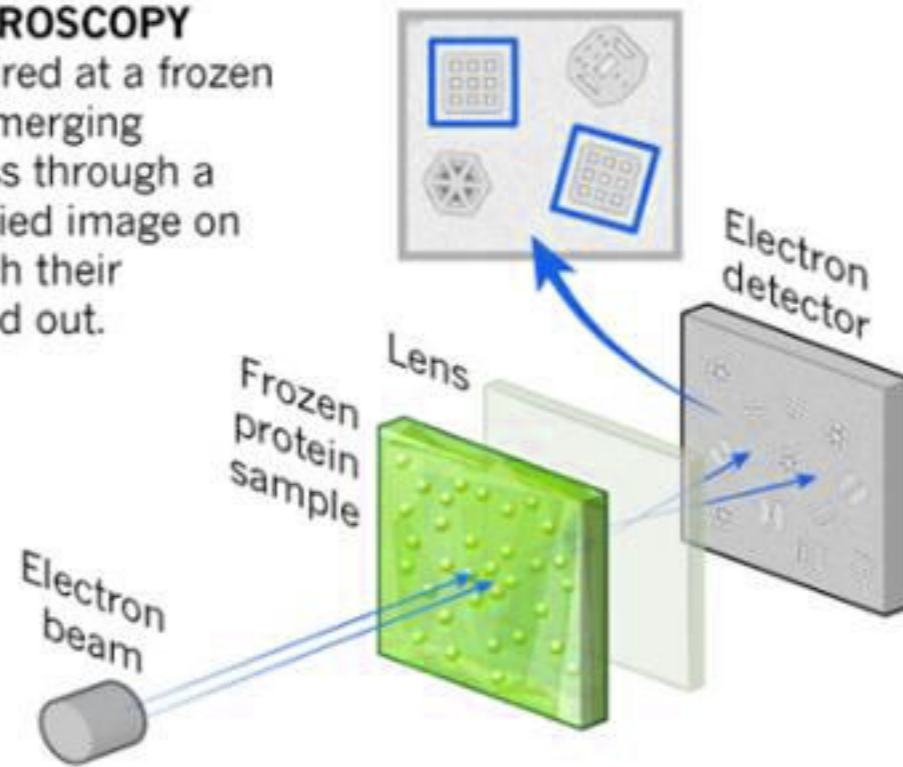
3D structure of biomolecules

# CryoEM

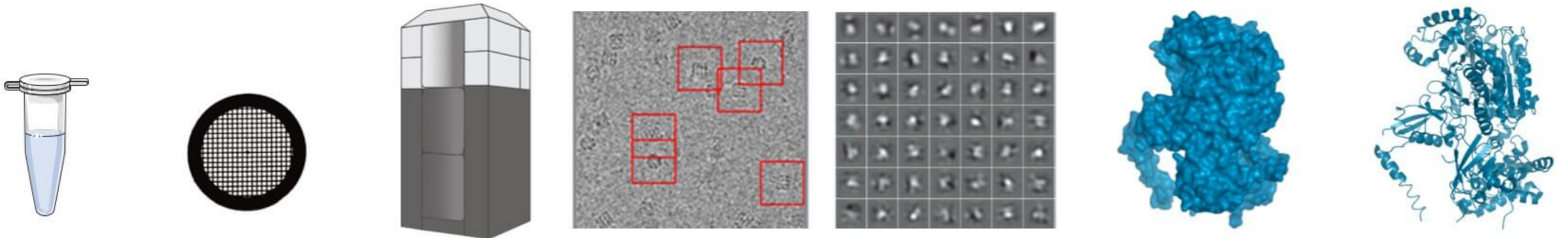
## Basic Principle:

### 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

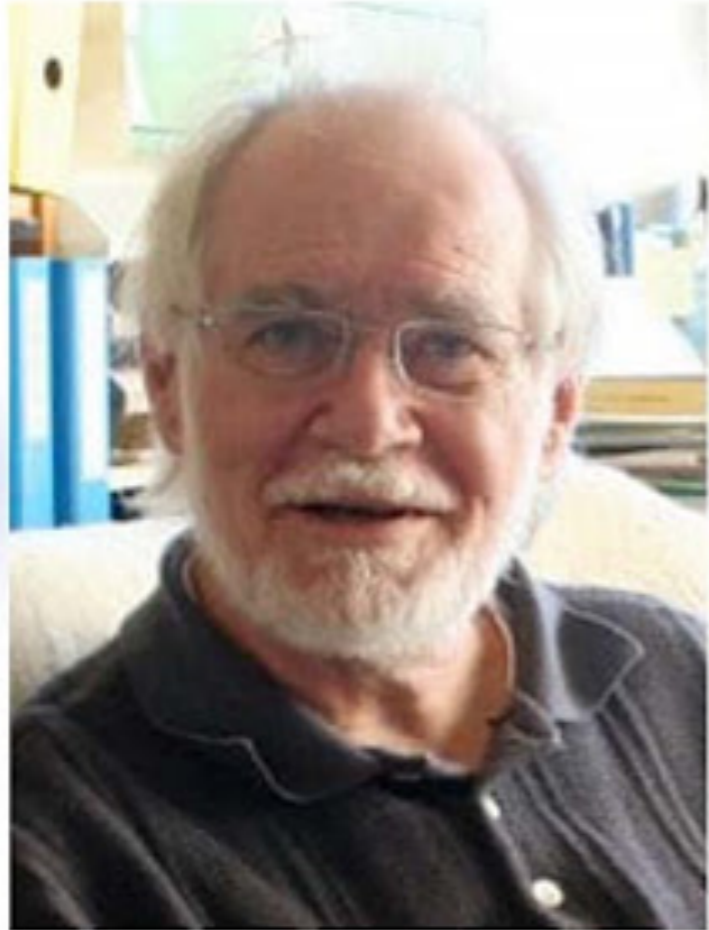
Particle picking

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.)



	<b>Pros</b>	<b>Cons</b>	<b>Sample Types</b>	<b>Resolution</b>
<b>X-ray crystallography</b>	<p>Established High Resolution Broad Molecular Weight Range Easy for model building</p>	<p>Static Crystalline State Protein needs to crystallize Not suitable for dynamic interactions Challenging for protein complexes</p>	<p>Soluble Proteins, Membrane Proteins, Small Molecules DNA/RNA and Protein Complexes</p>	High
<b>NMR</b>	<p>High Resolution 3D structure in solution Protein dynamics</p>	<p>High Sample Purity Difficult Sample Preparation Size limit</p>	<p>MW below 40-50 kDA Water Soluble Samples</p>	High
<b>CryoEM</b>	<p>Low amounts of sample Structure in native state</p>	<p>Size Limitation Costly EM equipment (6M \$) Resolution limited</p>	<p>&gt;100kDa Membrane Proteins Large assemblies like virions, ribosomes</p>	>3.5 Angstrom

# Structural Data Repositories:

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

The screenshot shows the RSCB PDB website homepage. At the top, there is a navigation bar with links for Deposit, Search, Visualize, Analyze, Download, Learn, About, Documentation, Careers, and COVID-19. Below this, the PDB logo is displayed along with statistics: 215,908 Structures from the PDB and 1,068,577 Computed Structure Models (CSM). A search bar is present with the placeholder text "Enter search term(s), Entry ID(s), or sequence". Below the search bar, there are links for "Advanced Search" and "Browse Annotations". A sidebar on the left contains navigation options: Welcome, Deposit, Search, Visualize, Analyze, Download, and Learn. The main content area features a "February Molecule of the Month" section with a 3D molecular model of Nanowires. Other sections include "Explore NEW Features" and "PDB-101 Training Resources".

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

The screenshot shows the EMDB website homepage. The header includes navigation links for Home, Deposition, Documentation, Resources, FTP Archive, REST API, About, Feedback, and Share. The main content area features the EMDB logo and a search bar with the placeholder text "Enter your search term(s) in the box below or build an advanced search query". Below the search bar, there are examples of search terms: "1001, Apoferritin, Tomography, Rossmann MG, 5A1A".

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?

EMDB (the Electron Microscopy Data Bank) is a public repository for electron cryo-microscopy maps and tomograms of macromolecular complexes and subcellular structures. It covers a variety of techniques, including single-particle analysis, electron tomography, sub-tomogram averaging, fibre diffraction and electron crystallography. [More...](#)  
As of 14 February 2024, EMDB contains 33009 entries ([latest entries](#), [trends](#)).

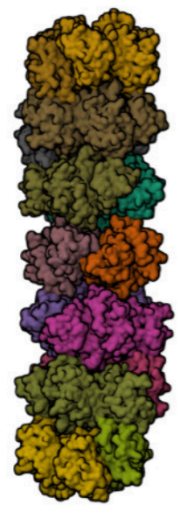
## EMDB News

- 06 January 2024: We are delighted to share that EMDB has officially been elevated to a Core Data Resource by ELIXIR, recognising it as a European data resource of fundamental importance to the wider life-science community and the long-term preservation of biological data.
- 22 November 2023: We are happy to share the most recent version of the EMDB Nucleic Acids Research paper. The paper describes recent developments and future plans of the EMDB in the context of the worldwide Protein Data Bank. Check it out now! <https://doi.org/10.1093/nar/gkad1019>

- ### Quick links
- EMDB Policies
  - Talks & Tutorials
  - Validation Analysis
  - Chart builder
  - EMICSS
  - Volume Browser
  - EMDB Citations
  - EMPIAR
  - PDBe
  - BiolImage Archive
  - EMDataResource
  - EM Navigator
  - 3D EM History

The footer navigation bar consists of four colored buttons: "Browse EMDB" (blue), "EMDB statistics" (yellow), "SARS-CoV-2 entries" (green), and "Deposit data" (red).

Biological Assembly 1



**Explore in 3D:** [Structure](#) | [Sequence Annotations](#) | [Electron Density](#) | [Validation Report](#) | [Ligand Interaction \(APJ\)](#)

**Global Symmetry:** Cyclic - C2 ([Explore in 3D](#))  
**Global Stoichiometry:** Homo 14-mer - A14

[Find Similar Assemblies](#)

Biological assembly 1 assigned by authors.

**Macromolecule Content**

- Total Structure Weight: 731.02 kDa
- Atom Count: 48,018
- Modelled Residue Count: 5,862
- Deposited Residue Count: 6,272
- Unique protein chains: 1

[Display Files](#) | [Download Files](#) | [Data API](#)

**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):** No

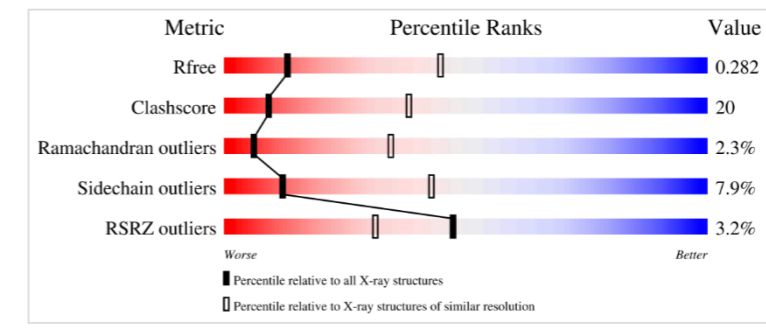
**Deposited:** 2008-11-19 **Released:** 2008-12-16  
**Deposition Author(s):** [Korennykh, A.V.](#), [Egea, P.F.](#), [Korostelev, A.A.](#), [Finer-Moore, J.](#), [Zhang, C.](#), [Shokat, K.M.](#), [Stroud, R.M.](#), [Walter, P.](#)

**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

**wwPDB Validation**

[3D Report](#) | [Full Report](#)



**Ligand Structure Quality Assessment**



This is version 1.3 of the entry. See complete [history](#).

Questions?