

Experimental methods for the elucidation of molecular structures

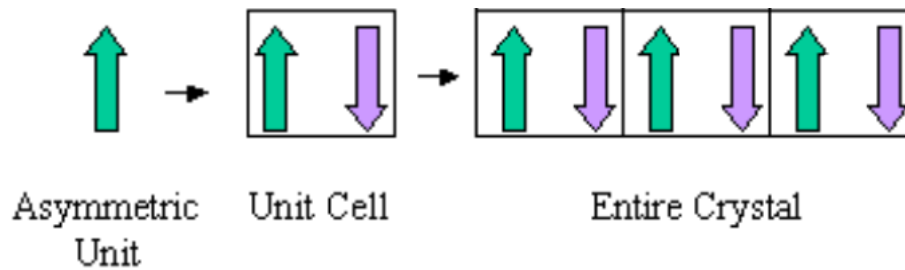
They can describe different aspect of the molecule: - **Atomic resolution**: position of atoms in the molecule. X-ray and NMR yield this info. - Outline of the molecule structure: SAXS, Cryo (but advances make it give high resolution...) - **Secondary structural elements** (CD, infra-red spectroscopy).

1 X-RAY CRYSTALLOGRAPHY

- X-ray source → molecule of interest scatters the beam and creates **diffraction pattern**** with. From pattern we can infer position of the atoms. (****Electron density map**** is the resulting pattern).
- Interpretation of electron density map was one of the first uses of computers in biology).
- X-Ray has the wavelength in the same scale as atoms. (1.54 Å. 10Å = 1nm).
- for long it damages the structure due to high energy.
- We use crystals because the regular pattern of trillions of molecules in the same conformation makes the results interpretable.
- Atoms cannot superpose, but waves can.
- X-Ray sources are for example **synchrotrons**.

1.1 CRYSTALLIZATION

- Not guaranteed to work. There are some proteins very difficult to crystallize
- **Hanging drop method** is used a lot. With a drop in a solution slowly evaporating.
- Parameters:
 - Concentrations of protein PH, Type and concentration of precipitant. It leads to a multifactorial optimization problem.
- Different crystal forms of the same protein can be reach.
- Membrane proteins are more difficult to predict.
- **Crystal**: periodic arrangement of proteins. Proteins in living systems don't have such conformations.
- Crystals have structures consisting of several molecules. **Unit cell** is the elementary unit you can use to replicate the crystal using only spatial translation. **Asymmetric unit** needs additional spatial rearrangement to produce the whole structure.



- Now there are automated programs that can take automatically of the different units and can help avoid mistakes.

1.2 SCATTERING/DIFFRACTION

- Huygens-Fresnel principle. Atoms absorb energy and creates circular emissions creating interferences, that can be constructive or desctructive, due to phase shifts. (like waves in water after rock).

Braggs law: Condition for constructive interference.

From the product of those physics law, diffraction patters are created and recorded.

Map is interpreted. Map is result of amplitude, angle of reflection and phase.

Large angles contain information about atoms close to eachother.

1.2.1 Solving the phase problem

Direct method: Brute force, trying all wavelenghts and observe which fits better.

MIR: Insertion of heavy atoms in the protein structure, difference between scattering patterns are used for finding the phase of the original protein.

Molecular replacement: Using homology model.

Result is electron density map.

From this information combined with the known primary structure we can infer the exact atomic 3d disposition.

Kendrew and Perutz solved the first structure in 1958 (myoglobin and hemoglobin).

1.3 QUALITY ASSESSMENT

1.3.1 R factor

Also called **residual or reliability factor**. Errors, missing data incorrect phase inference... Computed diffraction pattern is compared with the theoretical diffraction pattern of the created model.

$$R = \frac{\sum F_{obs} - F_{cal}}{\sum F_{obs}}$$

0 means total agreement.

0.15-0.2 is a good agreement.

0.6 is bad. Rfree is an alternative. We check our model with a test set.

1.3.2 Resolution

Determined by quality of the crystal (purity, regularity...)

Braggs law: $d \propto 1/\sin(\theta)$

Small angles: large d (close to the center).

Resolution gives information about the purity of the crystal, regularity, water content...

It's typically given in armstrong.

Is estimated from the diffraction pattern.

Lower armstrong means higher resolution.

Resolution doesn't mean standard deviation.

Resolution [Å]	Resolvable Structural Detail	Quality
4.0	General structural feature, domains	Poor
3.5	Secondary structural elements visible	
3.0	Amino acid residue side chains	typical
2.5	Defined side chains and torsion angles, Water molecules visible	
1.5	Heavy atoms visible	Very good
1.0	H-Atoms visible	Excellent

1.3.3 B-factor

For every atom

Also called Debye-Waller factor or temperature factor.

Atoms that are in a noisy/not pure crystalline form have a higher temperature (lower quality).

1.3.4 Crystalline form properties

Enzymes are still active in crystalline form

water content is similar to cellular environment

1.4 Websites

pdb: Contains information: Resolution, FASTA sequence, PDB format, mmCIF

2 NUCLEAR MAGNETIC RESONANCE

Properties of the nuclei:

Every nucleus with an odd number of protons and/or neutrons have magnetic moment.

Under a static magnetic field up and down spin are separated into two states. Delta-E measures the difference of difference between two states.

Boltzmann distribution: States of low energy are more frequent.

Electromagnetic radiation of energy $h \cdot \nu$, changes the spin of the atoms.

Emission depends on the electronic environment of a atom (solvation, aromatic systems...).

Absortion is compared with reference substance (TMS, Tetramethyl silane).

Chemical shift is the resulting measure, in ppm.

2.1 Instrumentation

Use of strong electromagnetic fields at very low temperature.

Different energy states are scanned. ### 2.2 NMR principle Different electronic environment: - Electrostatic charges: Electron cloud acts as an “electromagnetic shield”. Electrostatic forces attracts the electron cloud and causes the nuclei to be less shielded, changing the emission. - Solvent effects also changes the spectrum. - Different structures in different environments can be “quantified”. - Spin-Spin coupling: Spin emissions are coupled to the spins of protons bounded to adjacent carbons. Can be through bonds or through space. - Also 2D NMR #### 2.2.1 Cosy Correlation spectroscopy: Captures bond-mediated spin-spin coupling and allows to know the spatial distribution of aminoacids. #### 2.2.2 Nuclear-Overhauser-Effect: Dipol-dipol spin emission changes.

Captures the effect of molecules not close on the covalent structure, but close in space due to protein secondary structure.

2.2.3 Distance constraints are used to derive possible structures of a protein molecule.

Is called embedding. We only have distances up to 5 Amstrongs.

There's no unique solution.

Problems are iteratively solved, then is refined.

X-ray	NMR
+ works for large molecules	+ in solution
+ high resolution	+ provides information on flexibility of molecular regions
+ simple quality assesement (R-factor, Resolution)	- only for small molecules (< 100 or so amino acids)
- requires crystallization	- No quality metrics available
- Generally no H-atoms visible	

3 SMALL ANGLE X-RAY SCATTERING

We don't get atom level information, but it's of biological relevance.

We don't need a crystal but we have only low resolution information.

The higher the angle, the more information we have.

We try to place in space the "scatterers".

Then we perform iterative refinement against the theoretical scattering profile of our model

4 Circular Dichroism Spectroscopy

Molecules absorb circularly polarized light differentially.

Percentage of alpha helix, beta-fold and random is determined.

(Useful when certain molecules are difficult to crystallize, prions or unstructured proteins).

5. Cryo- Electron microscopy

Low resolution, but significant advances in recent years