Structural Bioinformatics

Elodie Laine

Master BIM-BMC Semestre 3, 2014-2015

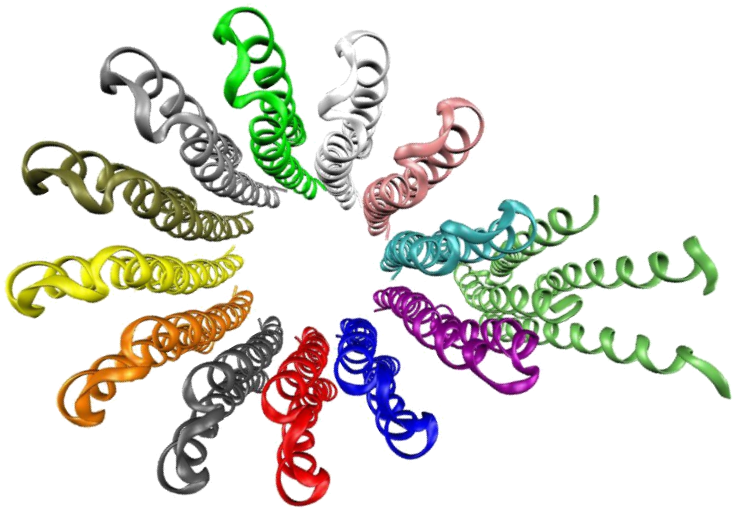
Laboratoire de Biologie Computationnelle et Quantitative (LCQB)
e-documents: http://www.lcqb.upmc.fr/laine/STRUCT
e-mail: elodie.laine@upmc.fr

Lecture 4 – Tertiary Structure Determination

Context

How can we determine a protein 3-dimensional coordinates ?

Experimental techniques
(X-ray crystallography, NMR, cryo-EM…)



In silico prediction
(bioinformatics methods)

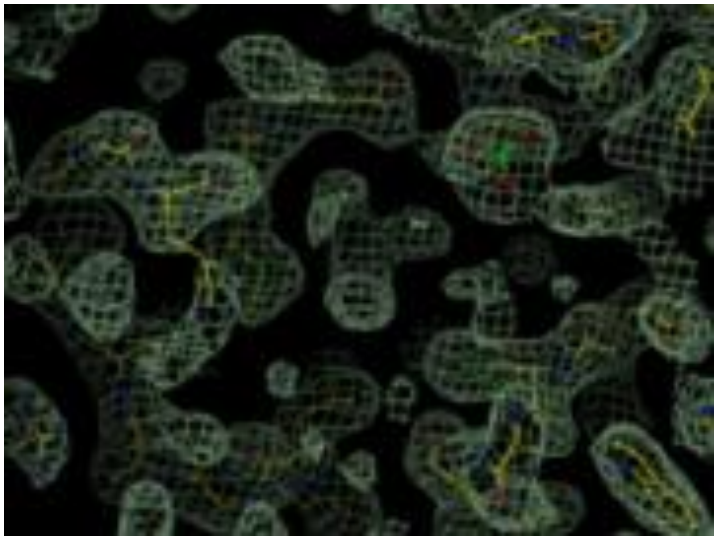
- ✓ *Ab initio* modelling
- ✓ Comparative modelling
- ✓ Fold recognition (threading)

Experimental techniques

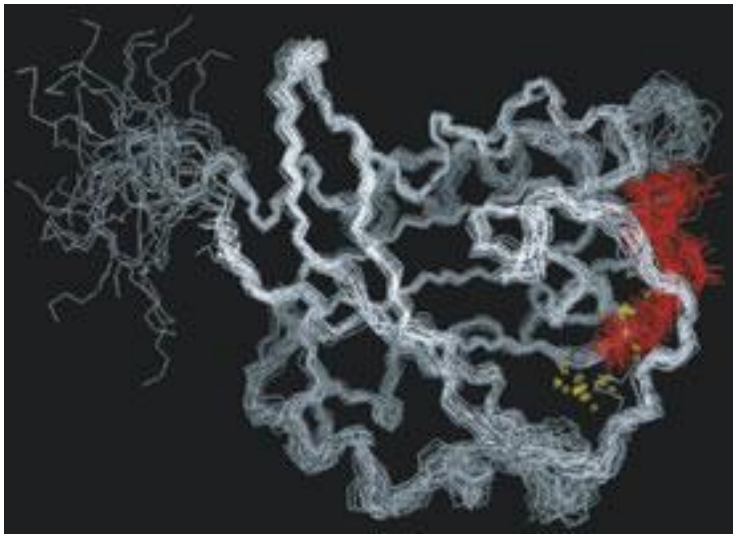
- X-ray crystallography
  - interactions of X-rays with electrons in the molecules in a crystal
- Nuclear magnetic resonance
  - interactions of an external magnetic field with the intrinsic magnetic properties of atomic nuclei which possess a spin angular momentum
- Cryo-electron microscopy
- Small angle x-ray scattering

Experimental techniques

Elec tronic density map from X-ray crystallography



Multiple models from nuclear magnetic resonance (NMR)



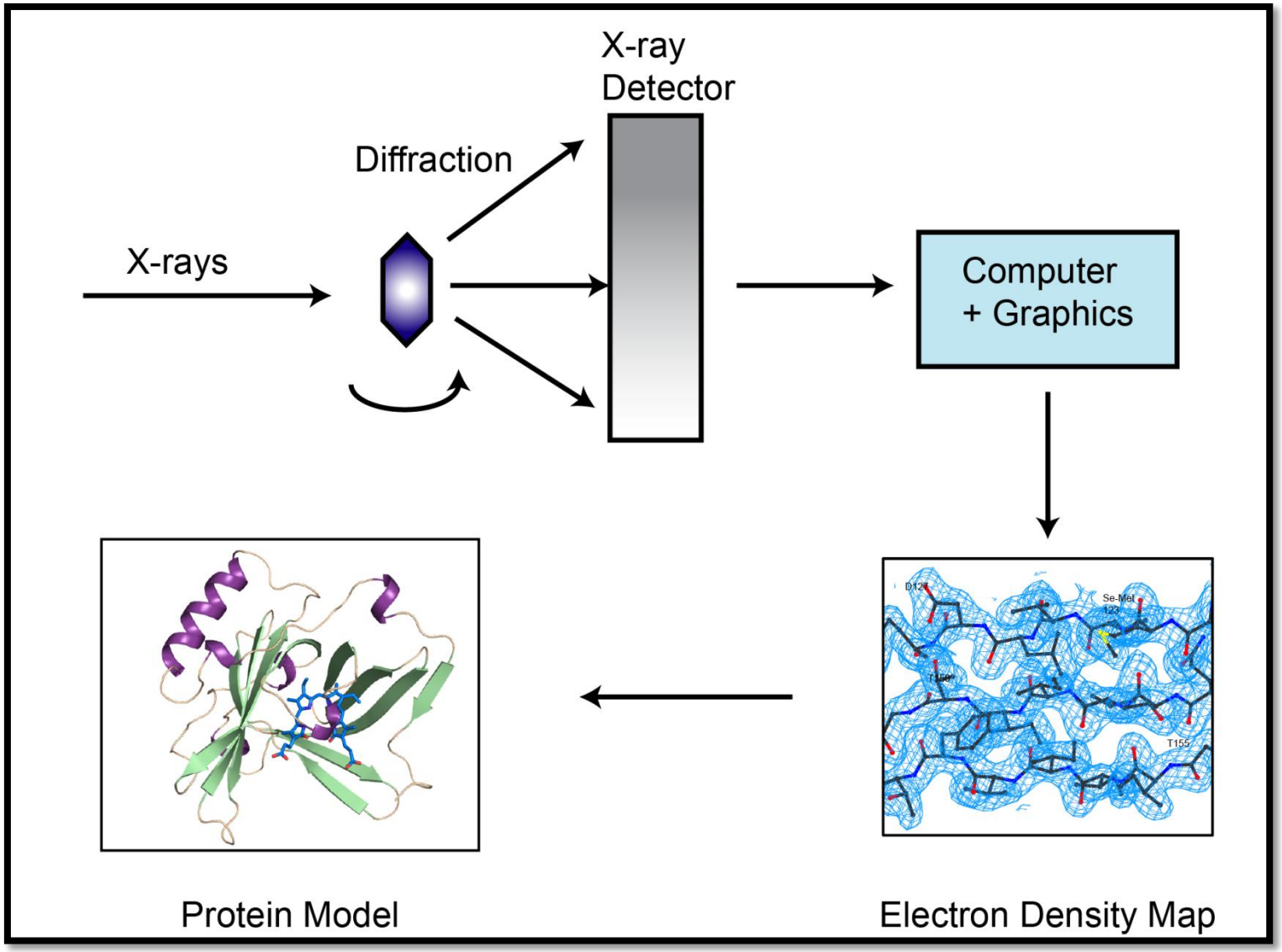
Both techniques need:

- purified protein
- amino acid sequence
- recombinant DNA technology
- computational methods…
  - molecular mechanics
  - molecular dynamics
  - molecular graphics

Experimental techniques

- **X-ray crystallography**
  - **interactions of X-rays with electrons in the molecules in a crystal**
- Nuclear magnetic resonance
  - interactions of an external magnetic field with the intrinsic magnetic properties of atomic nuclei which possess a spin angular momentum
- Cryo-electron microscopy
- Small angle x-ray scattering

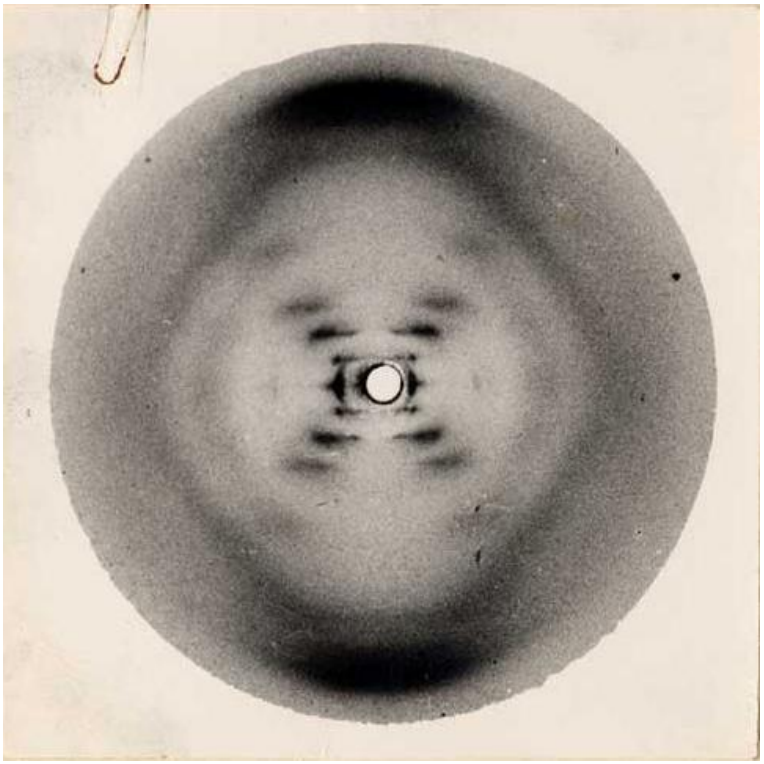
X-ray crystallography workflow



X-ray crystallography: a brief history

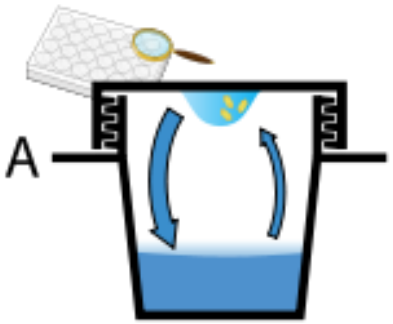
- **1895**: **Discovery of X-rays** by W. Röntgenand, a german Physicist (Nobel Prize in1901)
- **1912**: **Crystals were found to diffract X-rays** by M. von Laue (Nobel Prize in 1914)
- **1913**: **The structure of table salt** is the first to be determined by X-ray by W. H. Bragg & W. L. Bragg (Nobel Prize in 1915)
- **1946**: **Proteins were found to crystallize** by J. B. Sumner (Nobel Prize)
- **1962**: **First protein structures**, those of myoglobin & hemoglobin, to be solved by x-ray by J. Kendrew & M. Perutz (Nobel Prize)
- **1962**: **Discovery of the DNA double helix** by F. Crick, J. Watson & M. Wilkins

Rosalind Franklin's X-ray image of DNA

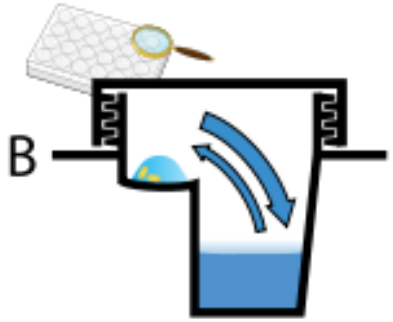


Crystal growth

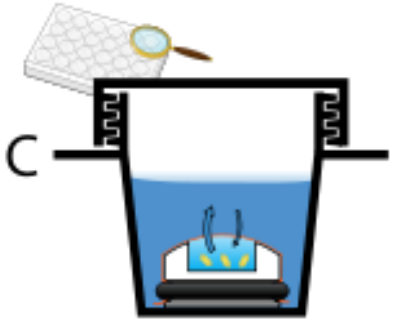
hanging drop



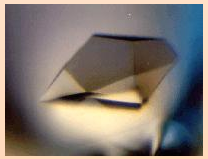
sitting drop



micro-dialysis



Protein crystals are almost always grown in solution. The most common approach is to lower the solubility of its component molecules very gradually.



There are two main steps:

- **nucleation** of a microscopic crystallite (possibly having only 100 molecules)
- **growth** of that crystallite, ideally to a diffraction-quality crystal

The crystal should be sufficiently **large** (>0.1 mm in all dimensions), **pure** in composition and **regular** in structure.

X-ray diffraction

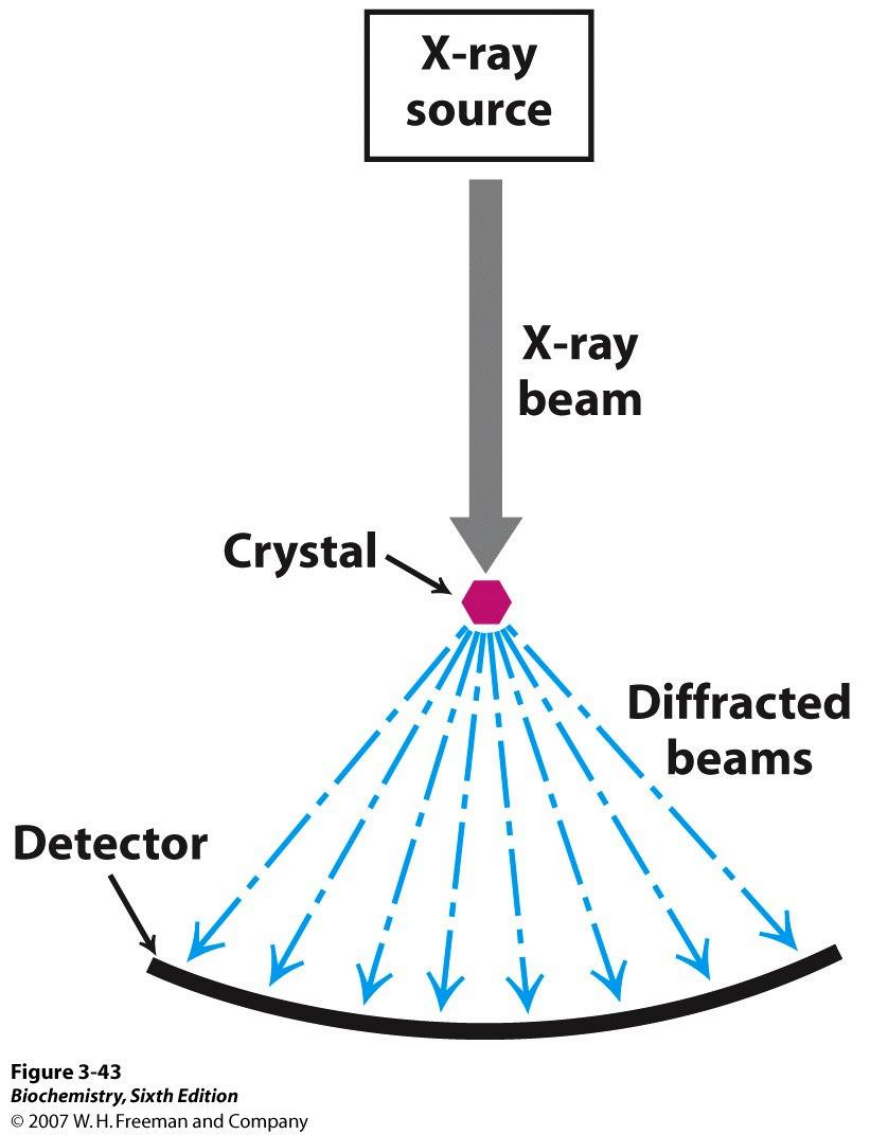


Figure 3-43
Biochemistry, Sixth Edition

Crystals

The repeating unit of a crystal, corresponding ~ to the volume occupied by a single molecule is called a **unit cell**
A crystal is built by billions of identical unit cell

X-Rays

Electromagnetic radiation of wavelength 1.54 Å
They are produced by a beam of accelerating e- on a copper anode target

X-ray diffraction

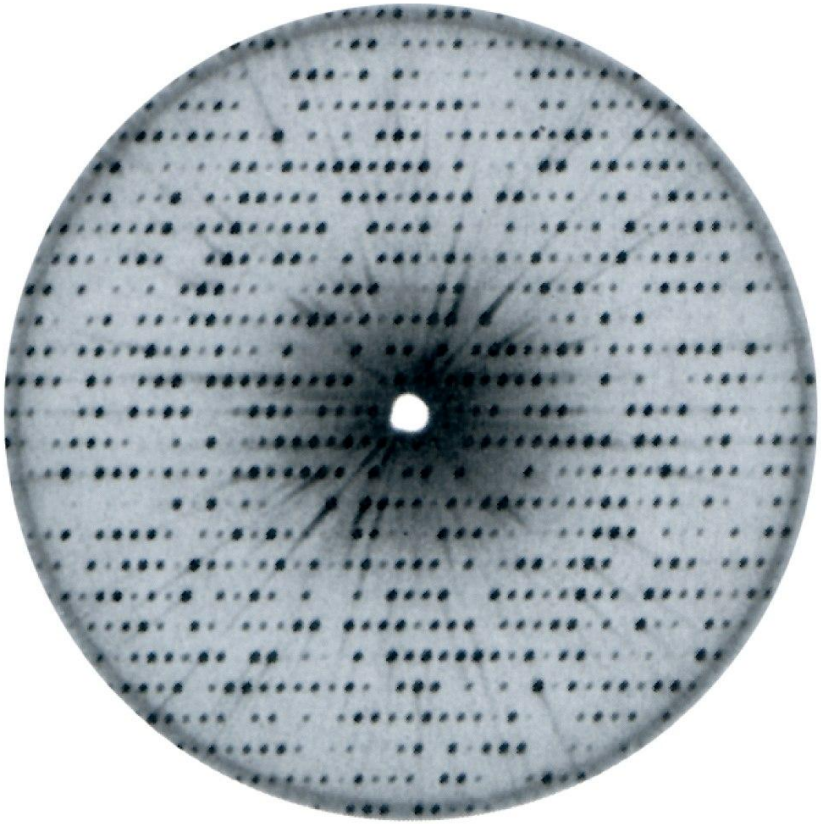
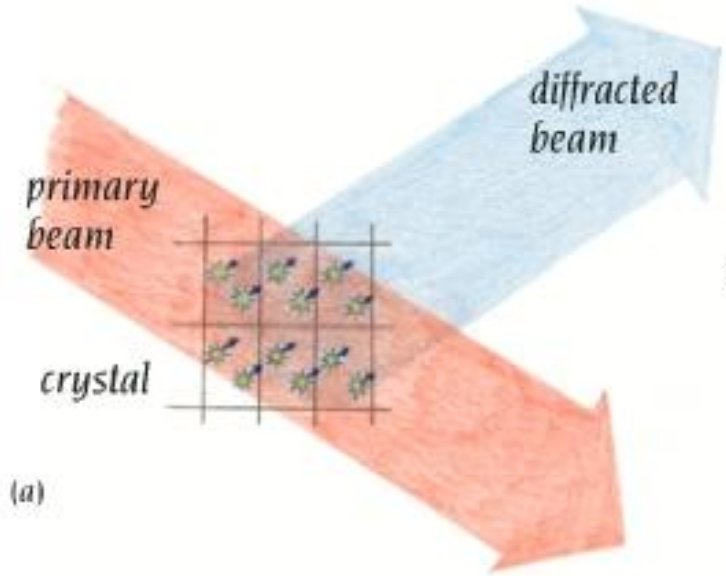
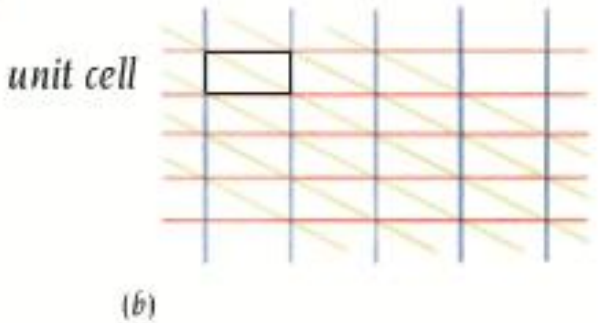


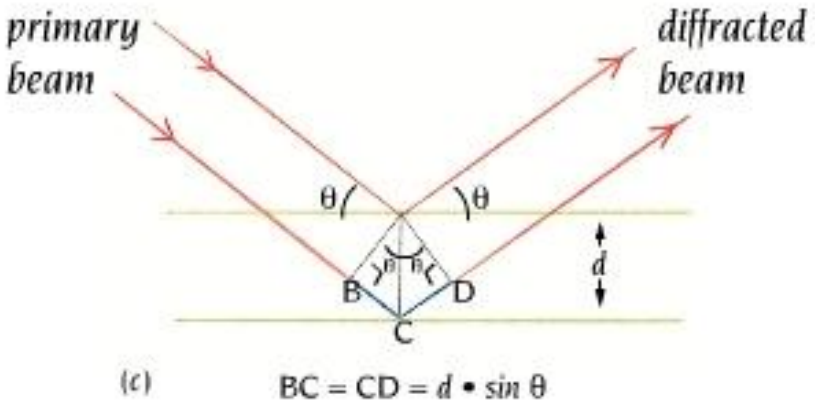
Figure 3-44
Biochemistry, Sixth Edition

When a crystal is mounted and exposed to an intense beam of X-rays, it scatters the X-rays into a **pattern of spots or reflections**. The relative intensities of these spots provide the information to determine the arrangement of molecules within the crystal in atomic detail. One image of spots is insufficient to reconstruct the whole crystal : to collect all the necessary information, the crystal must be rotated step-by-step through 180°, with an image recorded at every step.

X-ray diffraction





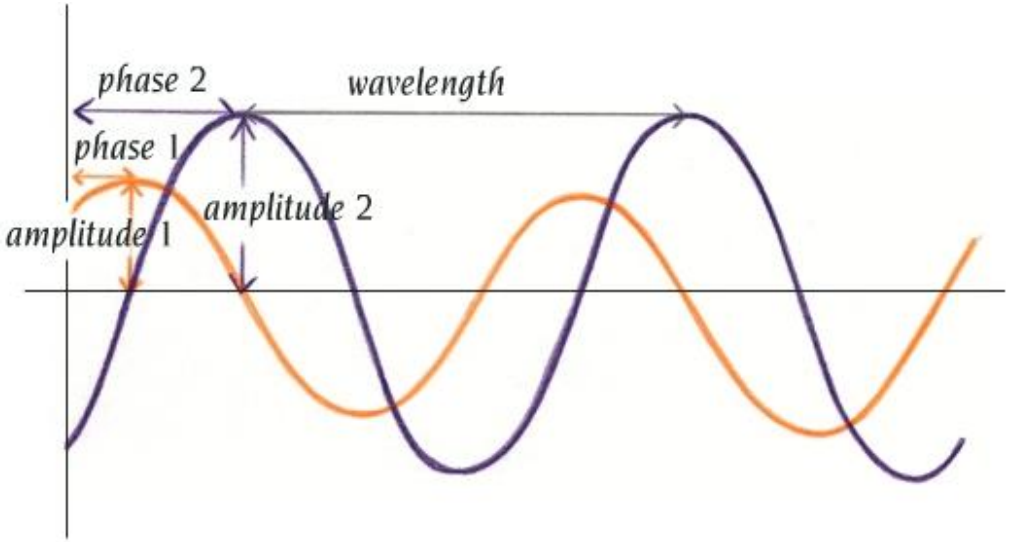


Braggs law:

$$2d\sin\theta = \lambda$$

X-ray data processing

The conversion from 2D diffraction pattern to 3D dimensional model is performed using Fast Fourier Transform.



Three main steps

1/ **Indexing** the reflections, *i.e.* identifying the dimensions of the unit cell and which image peak corresponds to which position in reciprocal space

2/ **Merging** and **scaling** of the different images to determine the relative strengths of the spots

3/ **Phasing**, i. e. determine the phases of the diffracted waves, which are lost in the experiment (the phase problem)

Initial phasing

- ✓ ***Ab initio* phasing or direct methods**: for small proteins(<1000 non-hydrogen atoms)
- ✓ **Molecular replacement**: if a related structure is known, it can be used as a search model in molecular replacement to determine the orientation and position of the molecules within the unit cell
- ✓ **Anomalous X-ray scattering**: by recording full sets of reflections at three different wavelengths for anomalously diffracting atoms (ex. selenium atoms in seleno-methionines)
- ✓ **Heavy atom methods (multiple isomorphous replacement)**: using electron-dense metal atoms

Model building and phase refinement

Having obtained initial phases, an initial model can be built. This model can be used to refine the phases, leading to an improved model, and so on. Refinement is achieved by adjusting the Debye-Waller factors (or B-factors) to fit the observed diffraction data.

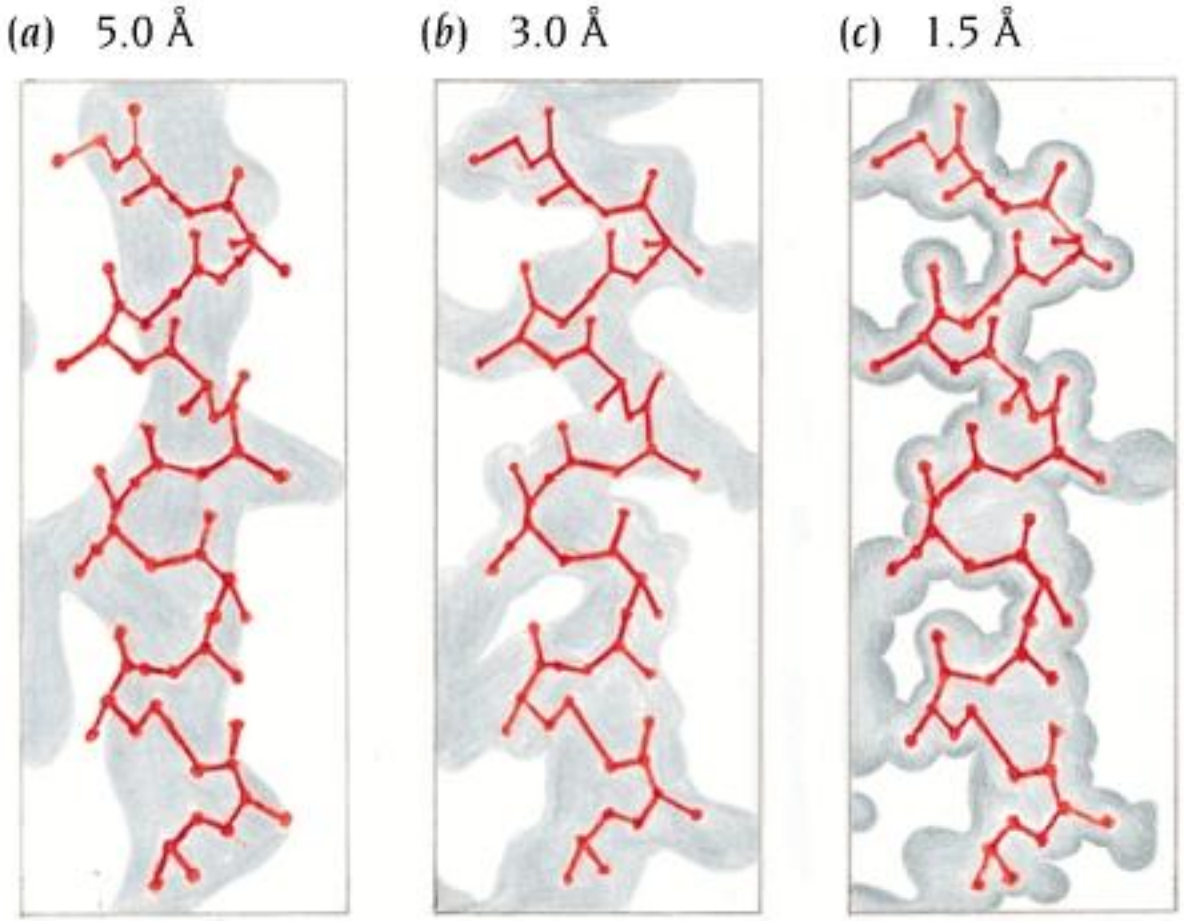
The agreement between the model and the diffraction data is measured by an R-factor:

$$R = \frac{\sum_{\text{all reflections}} \left| F_o - F_c \right|}{\sum_{\text{all reflections}} \left| F_o \right|}$$

where F is the **structure factor**.

A similar $\mathbf{R_{free}}$ measure can be evaluated from a subset (~10%) of reflections not included in the structure refinement. R_{free} should be approximately the resolution in angstroms divided by 10.

Electron density map



Experimental techniques

- X-ray crystallography
  - interactions of X-rays with electrons in the molecules in a crystal
- **Nuclear magnetic resonance**
  - **interactions of an external magnetic field with the intrinsic magnetic properties of atomic nuclei which possess a spin angular momentum**
- Cryo-electron microscopy
- Small angle x-ray scattering

NMR principle

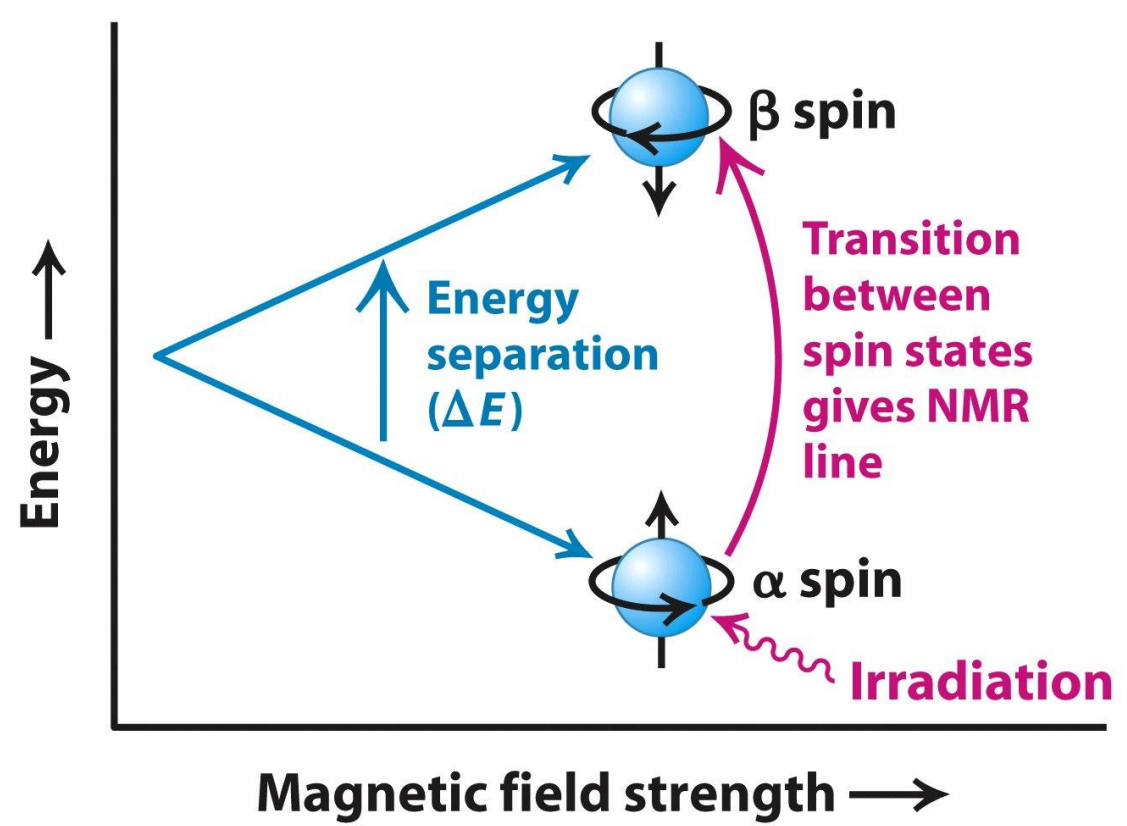


Figure 3-48
Biochemistry, Sixth Edition

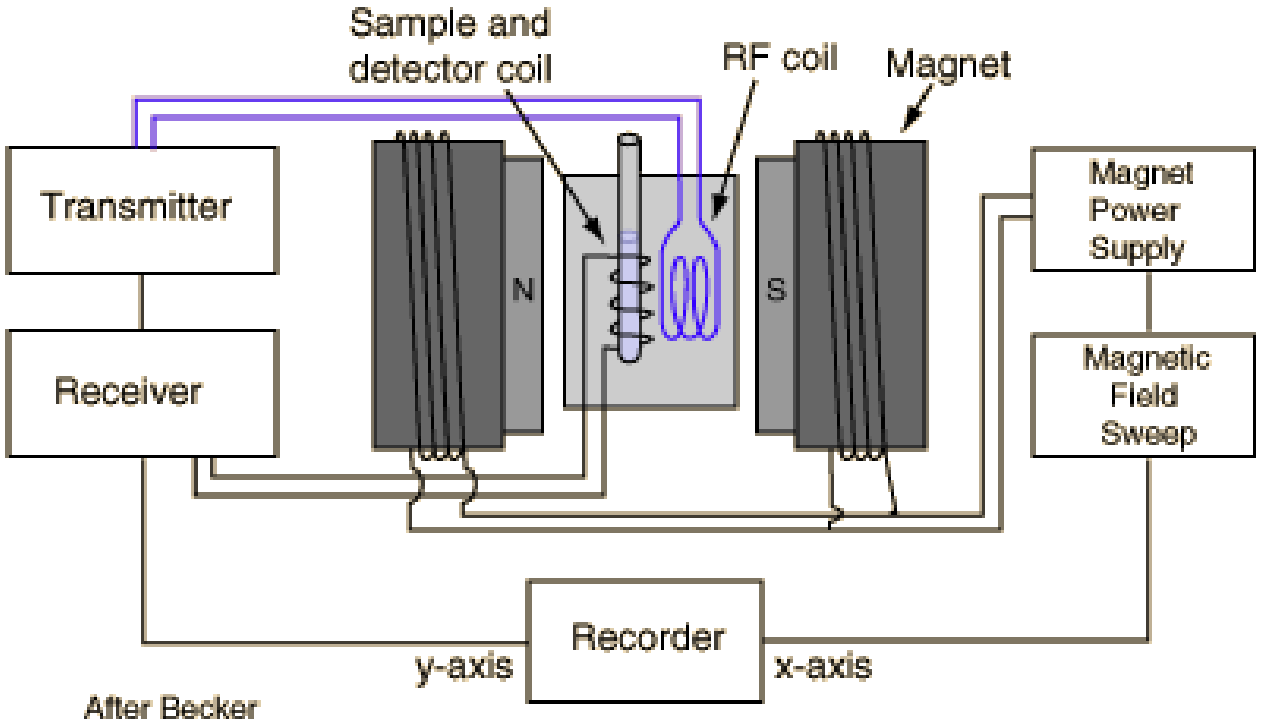
Atomic nuclei (^{1}H or ^{15}N) possess an intrinsic spin angular momentum, resulting in a magnetic momentum that can interact with an externally applied magnetic field B.

In a B field the spins of H align.

Equilibrium alignment can be perturbed by pulses of radiofrequency (RF).

NMR experimental setup

A sample containing protons (hydrogen nuclei) is placed in a strong magnetic field to produce partial polarization of the protons. A strong RF field is also imposed on the sample to excite some of the nuclear spins into their higher energy state.



When this strong RF signal is switched off, the spins tend to return to their lower state, producing a small amount of radiation at the Larmor frequency associated with that field.

1D-NMR

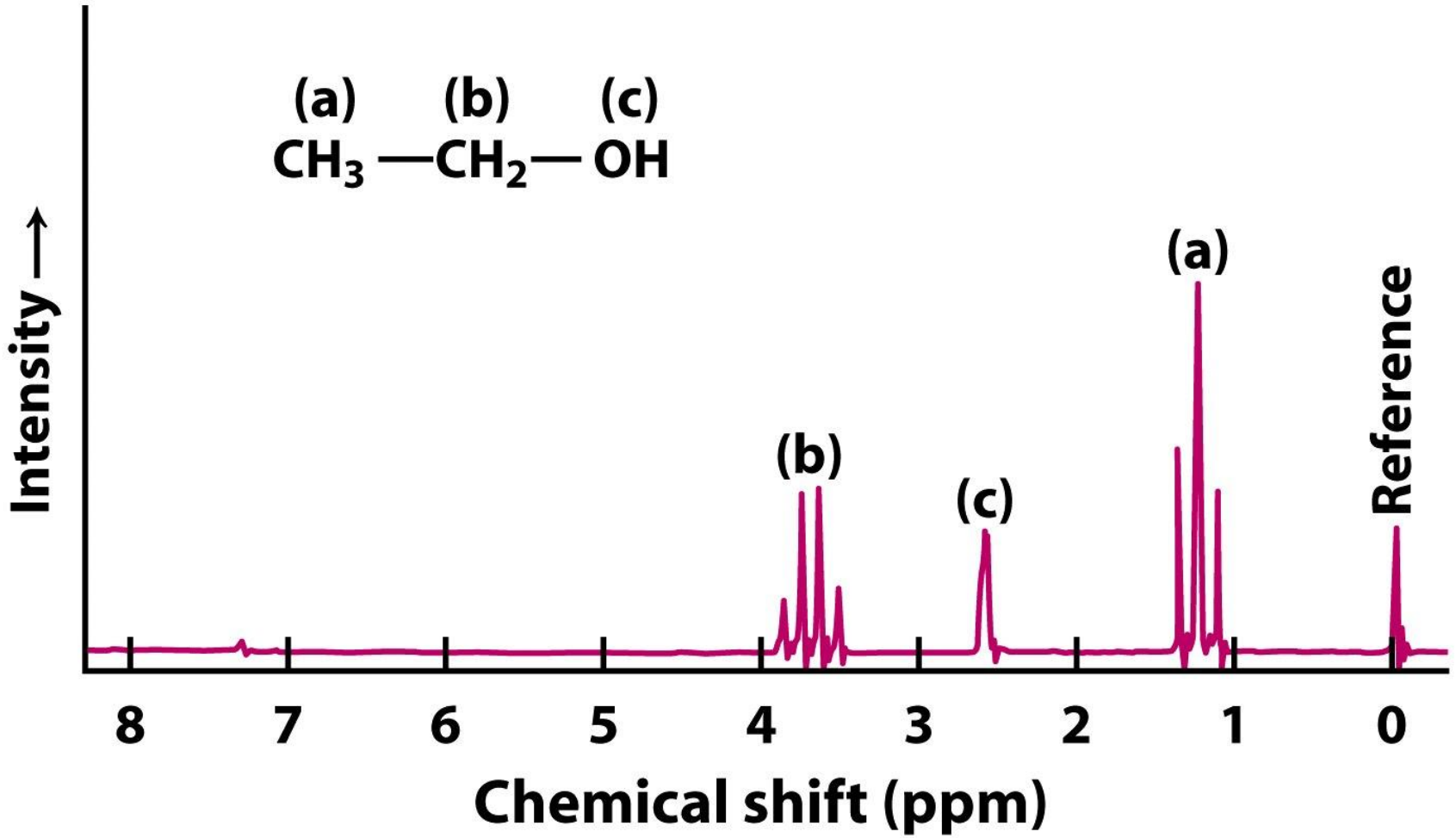
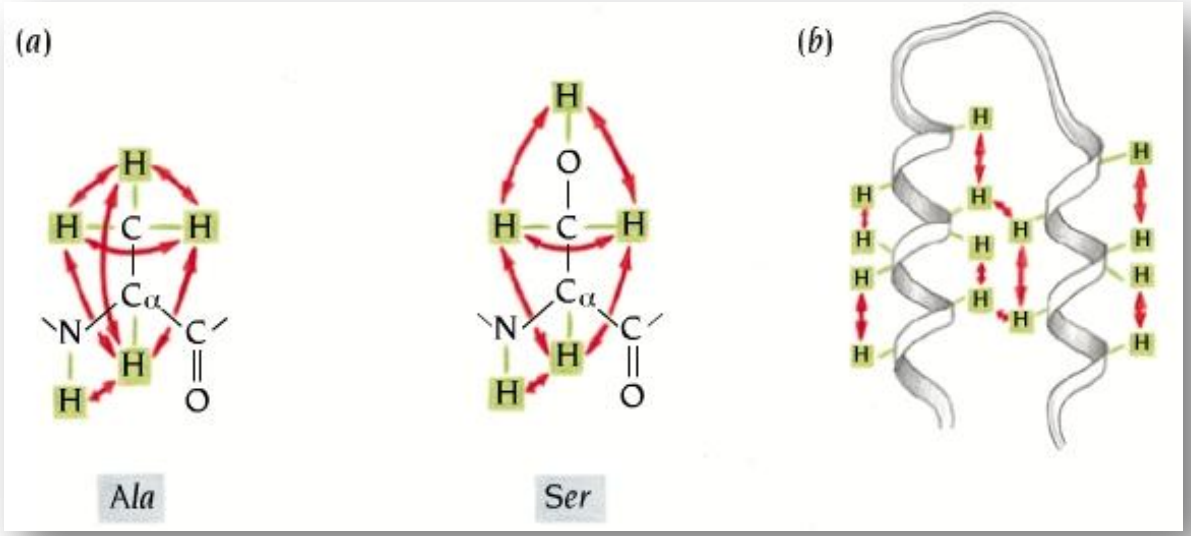


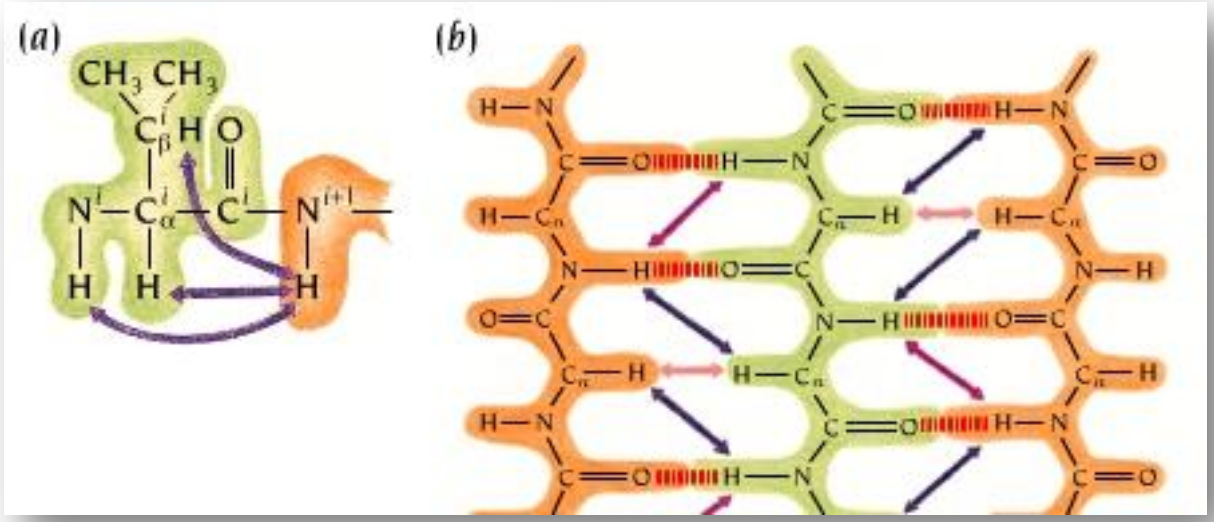
Figure 3-49a
Biochemistry, Sixth Edition

H-H interactions in proteins



α-helix

β-sheet



2D-NMR

2D-NMR spectra provide more information about a molecule than one-dimensional NMR spectra and are especially useful in determining the structure of a molecule.

The diagonal is similar to a 1D spectrum. Peaks that are off-diagonal (cross-peaks) correspond to interactions of H atoms that are close to each other in space.

COSY

(Correlation) → Fingerprint of a.a.
Distance between BONDED H atoms (≤ three chemical bonds, *i.e.* within the same a.a.)

NOESY

(NOE, Nuclear Overhausser Effect)
Distance between H atoms close together in space (≤ 5 Å)

NMR 2D spectra assignment

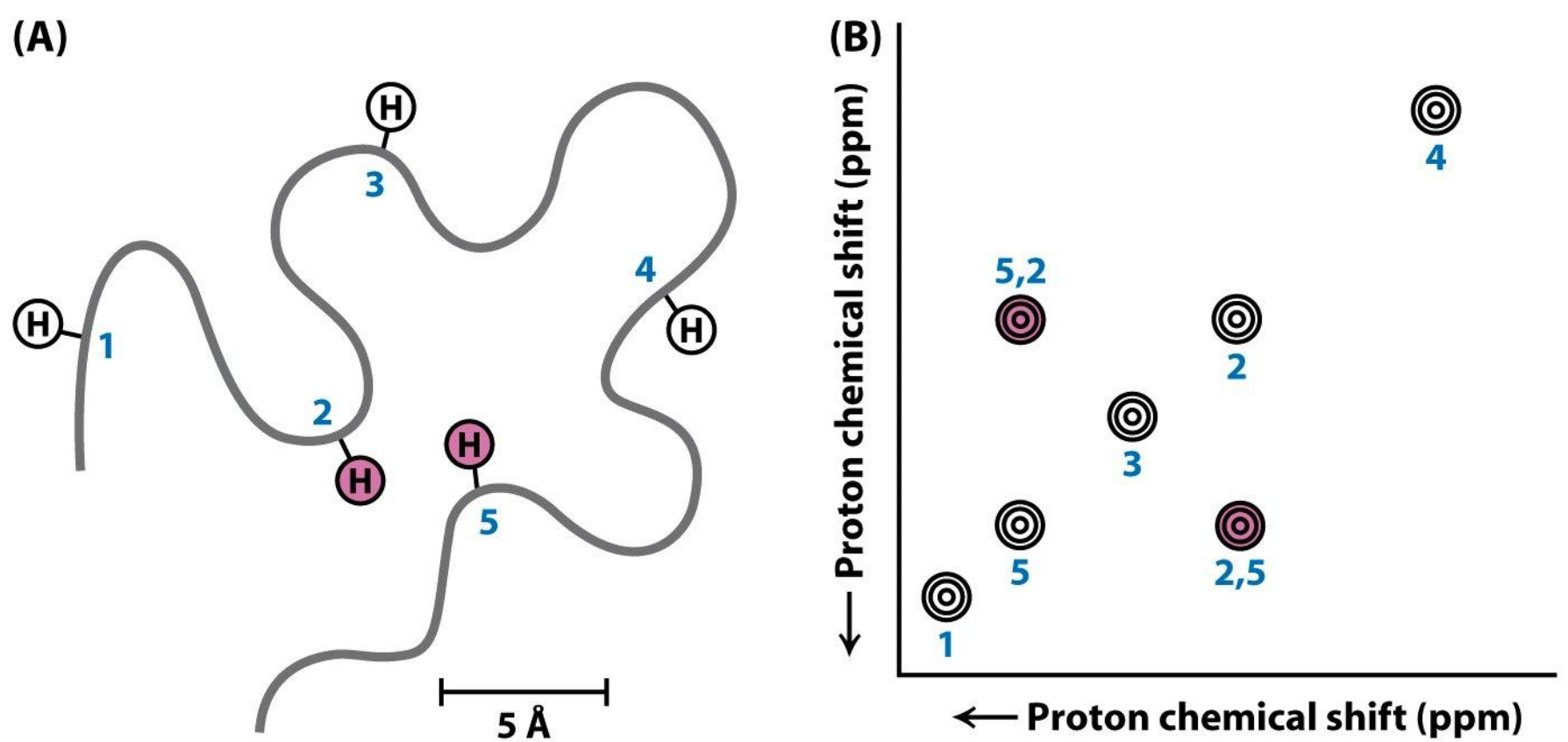
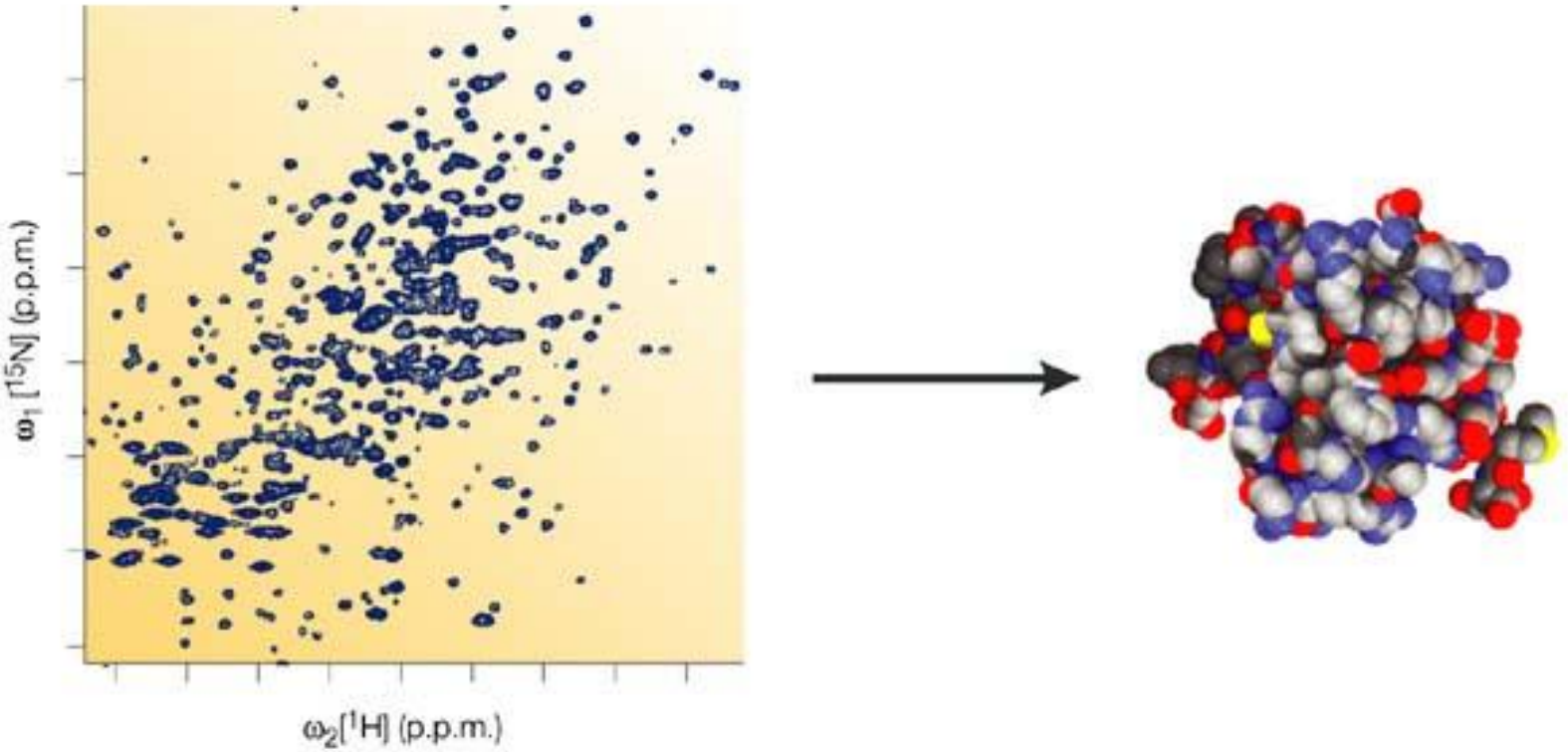


Figure 3-50
Biochemistry, Sixth Edition

From NMR 2D spectra to 3D molecular model



Biologically important nuclei giving NMR signals

Nucleus	Natural abundance (% by weight of the element)
^{1}H	99.984
^{2}H	0.016
^{13}C	1.108
^{14}N	99.635
^{15}N	0.365
^{17}O	0.037
^{23}Na	100.0
^{25}Mg	10.05
^{31}P	100.0
^{35}Cl	75.4
^{39}K	93.1

Table 3-4
Biochemistry, Sixth Edition

3D-molecular model

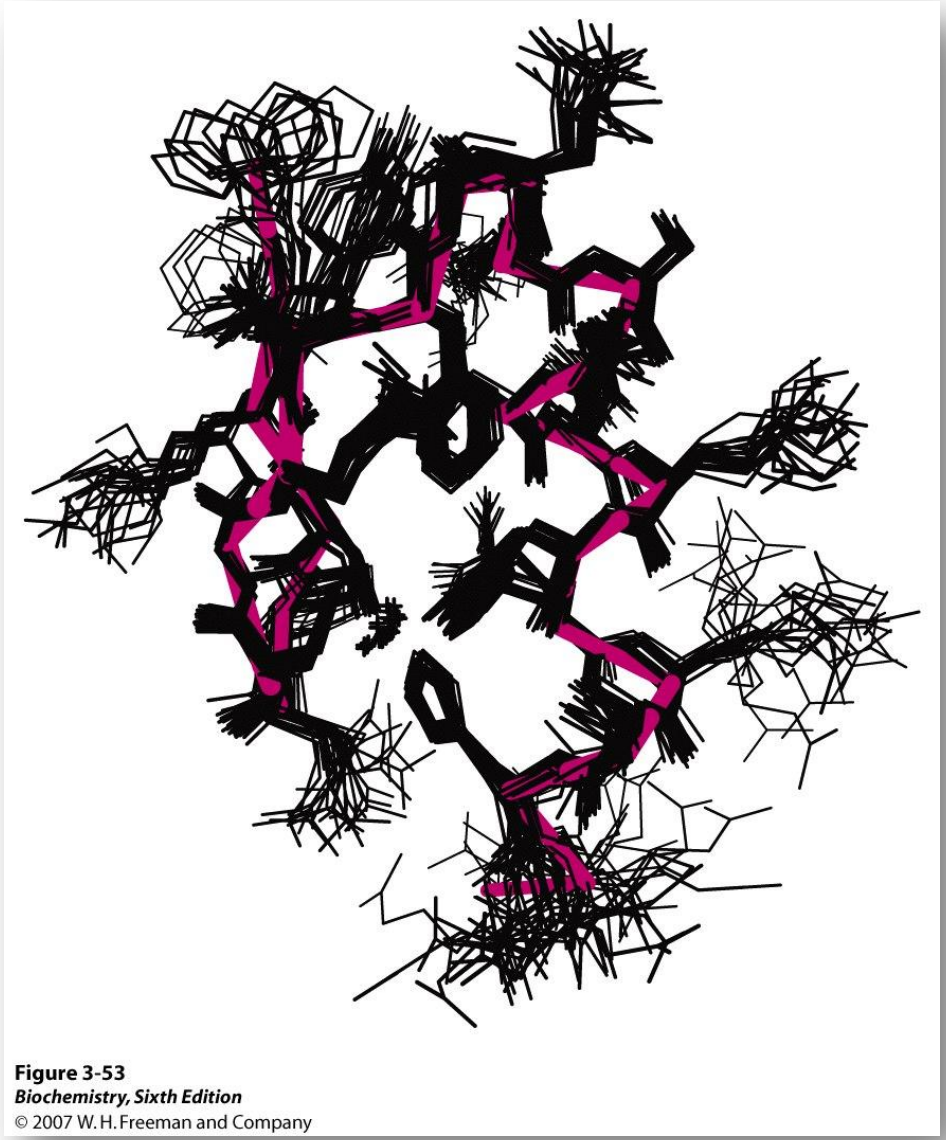


Figure 3-53
Biochemistry, Sixth Edition

The molecular model is refined using molecular modeling techniques. No unique structure is determined, but **different structural models** that are compatible with data (ambiguity).

Experimental techniques

X-ray crystallography

• Protein crystal of virtually any size • Atomic model inferred from the electron density map	• Few high quality crystals • Requires phase determination • Frozen static structure

NMR

• 3D-molecular model from distance contraints • Can probe the time resolved dynamics of a molecule	• Limited size (~30 kDa) • Requires protein in high concentration • requires isotopes (^{1}H, ^{13}C, ^{15}N)

Conclusion

- **Protein sructures deposited in the PDB are molecular models** that were determined using computational tools based on experimental data.
- **Protein structures determined by X-ray cristallography** are static and correspond to a frozen state of the molecule in non-native environment
- **Proteins structures determined by NMR** are flexible, to a certain extent.