Modern methods in protein research

The Integrative Structural Biology part

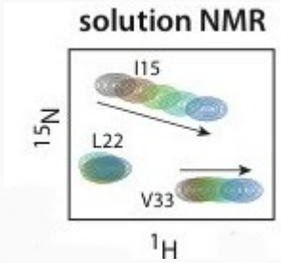
Josef Chmelík

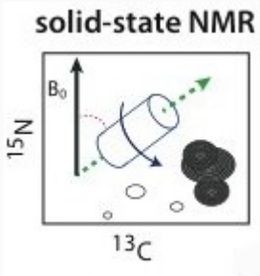
March 2020

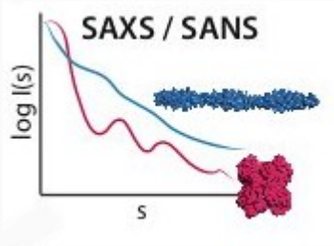
Integrative Structural Biology

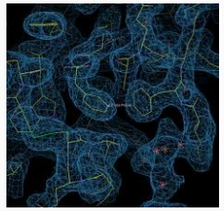
- combines data from multiple experimental techniques (X-ray, NMR, FRET, MS3D, CryoEM, ...)

  → model for the biological system of interest at atomic resolution
- Biomolecular Complexes – inaccessible by traditional methods like X-ray or NMR (or CryoEM)



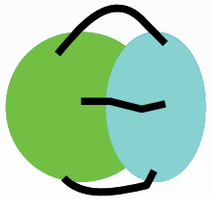




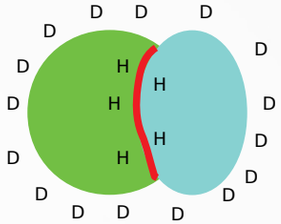


NMR

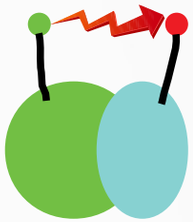
MS3D



HDX-MS



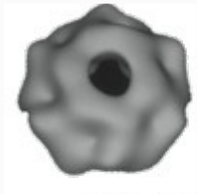
FRET



SAXS

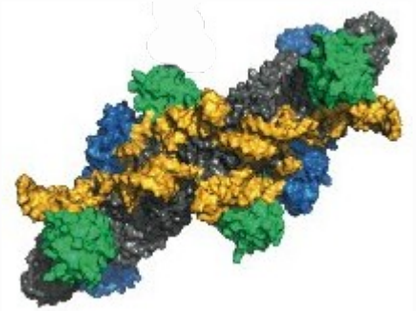
ISB

CryoEM

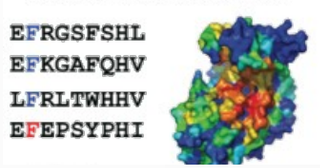


X-ray

3D structure



Bioinformatics, mutagenesis, …



Outline

1. Nuclear Magnetic Resonance
2. Small Angle X-ray Scattering
3. Förster Resonance Energy Transfer
4. Cryo Electron Microscopy

NMR

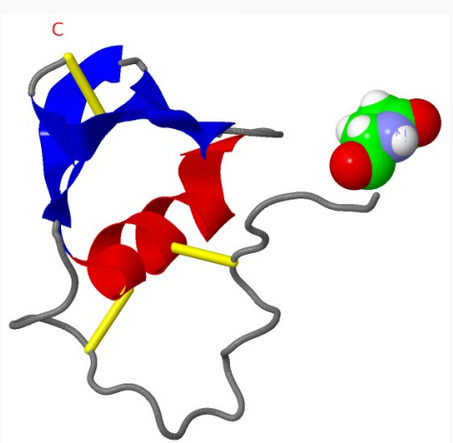
Nuclear Magnetic Resonance

History

The Nobel Prize in Chemistry:

1991: *Richard Ernst* – pulse NMR, FT tranformation

2002: *Kurt Wüthrich* – NMR methods for 3D structure determination in solution



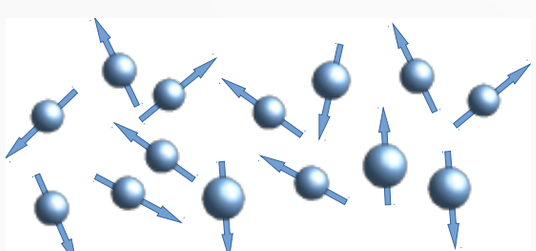
proteinase inhibitor IIA (1985)

Nuclear Magnetic Resonance

- Observe changes of local magnetic fields around atom nuclei
- Nuclear spin ≠ 0 (for biomolecules spin=½)

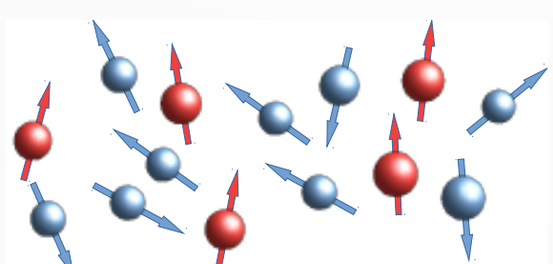
  → magnetic moment

$\vec{B}_0 = 0$



Random orientation:
Net magnetization:
$M_0 = 0$

$\vec{B}_0 \neq 0$

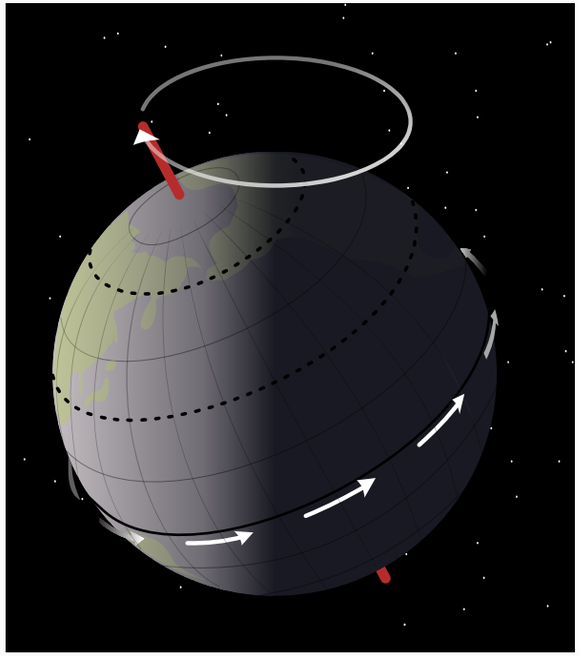


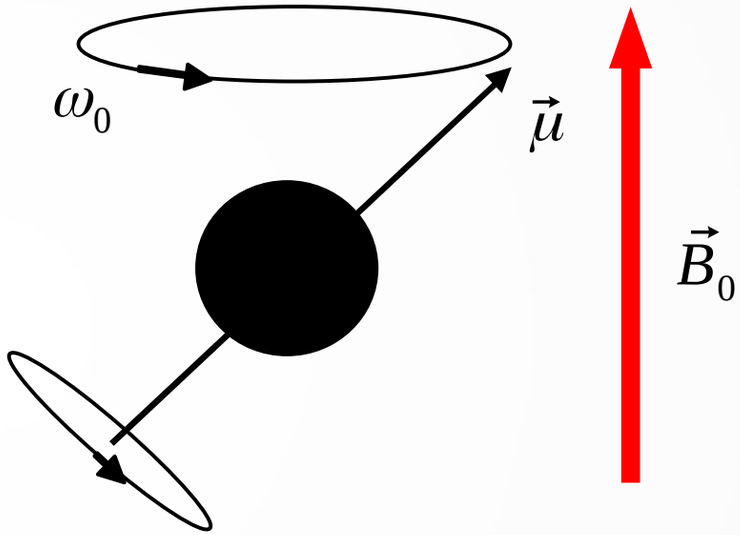
“preferred” orientation:
Net magnetization:
$M_0 \neq 0$

$$M_0 \approx \gamma^2 h^2 N_s B_0 / 4kT$$

γ - gyromagnetic ratio
h - Planck's constant
N_s - number of spins
B_0 - magnetic field strength
K - Boltzmann constant
T - absolute temperature

- Magnetic moment precession movement

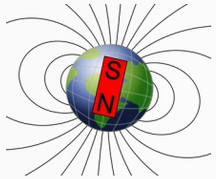




Larmor (resonance) frequency $\omega_0 = -\gamma B_0$

Nucleus	Spin	natural abundance [%]	γ [10^7rad $T^{-1}s^{-1}$]	NMR frequency [MHz] (11,74 T)	Relative sensitivity [%]
^{1}H	1/2	99,99	26,75	500,0	100
^{2}H	1	0,01	4,11	76,8	0,0001
^{3}H	1/2	0	28,54	533,3	0
^{12}C	0	98,93	0	0	0
^{13}C	1/2	1,07	6,73	125,7	0,02
^{14}N	1	99,63	1,93	36,1	0,1
^{15}N	1/2	0,37	-2,71	50,7	0,0004
^{16}O	0	99,96	0	0	0
^{19}F	1/2	100	25,18	470,4	83
^{31}P	1/2	100	10,84	202,4	6,6

Magnetic Field Strength



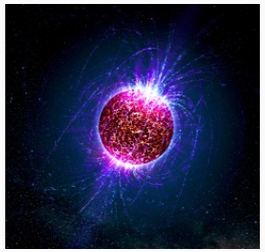
31 – 58 μT
earth

16 T
levitating frog

0.1 – 100 GT
Magnetar

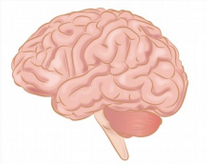
1.5 – 3 T
MRI

1 – 100 MT
Neutron star



$\vec{B}_0$

0.1 – 1 pT
brain



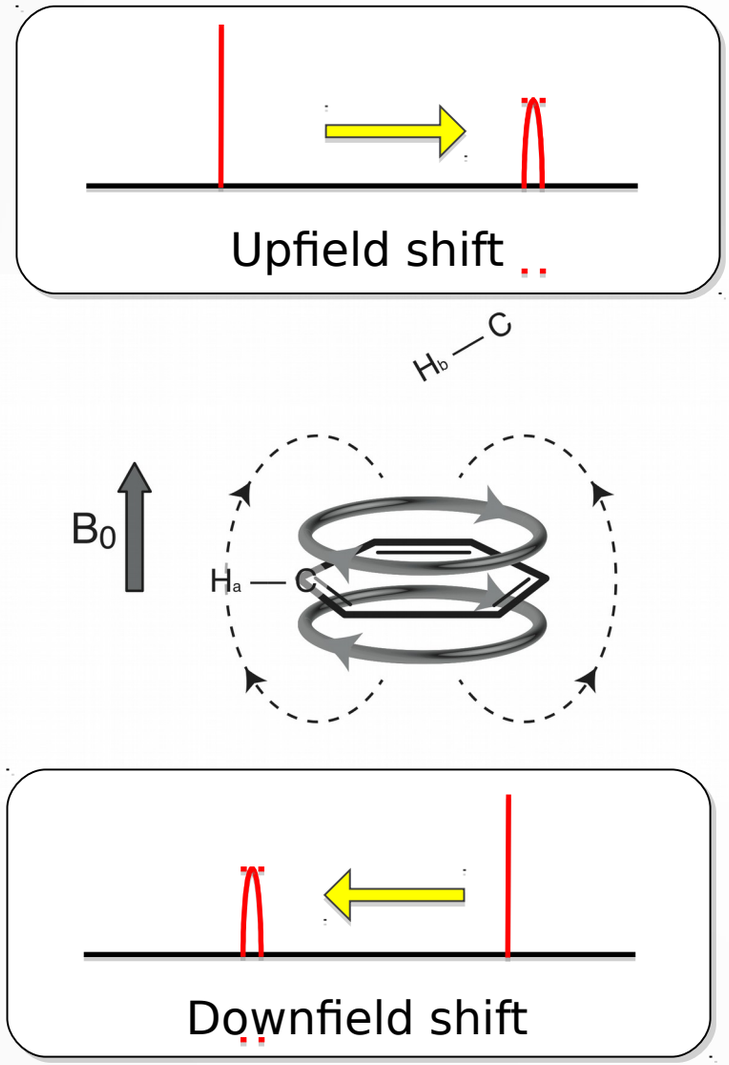
5 mT
fridge magnet

5 T
MAGLEV

7 – 26 T
NMR

0.5 mT
Safe zone for pacemaker

- Same nuclei same resonance frequency? NO!



Change of "effective" magnetic field by surrounding electrons

$$B = B_0 - B' = B_0(1 - \sigma)$$

Change of resonance (Larmor) frequency

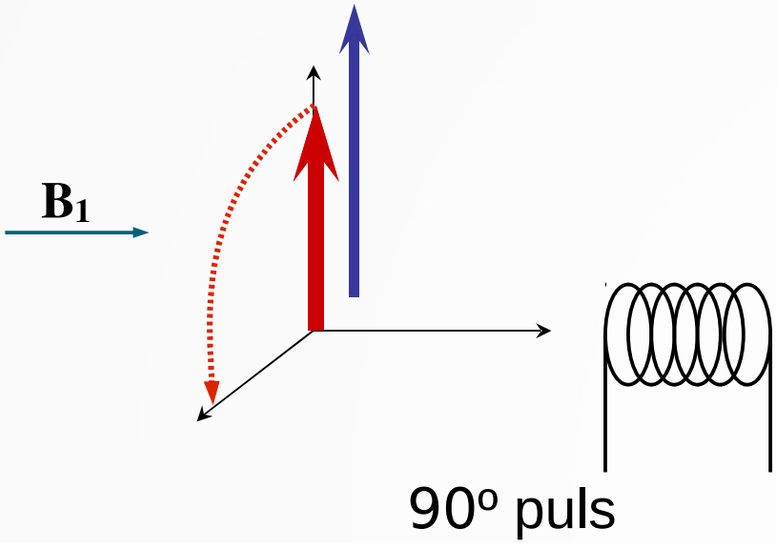
$$\delta = 10^6 \frac{\omega - \omega_{ref}}{\omega_0}$$

Chemical shift

"units" ppm (parts per million)

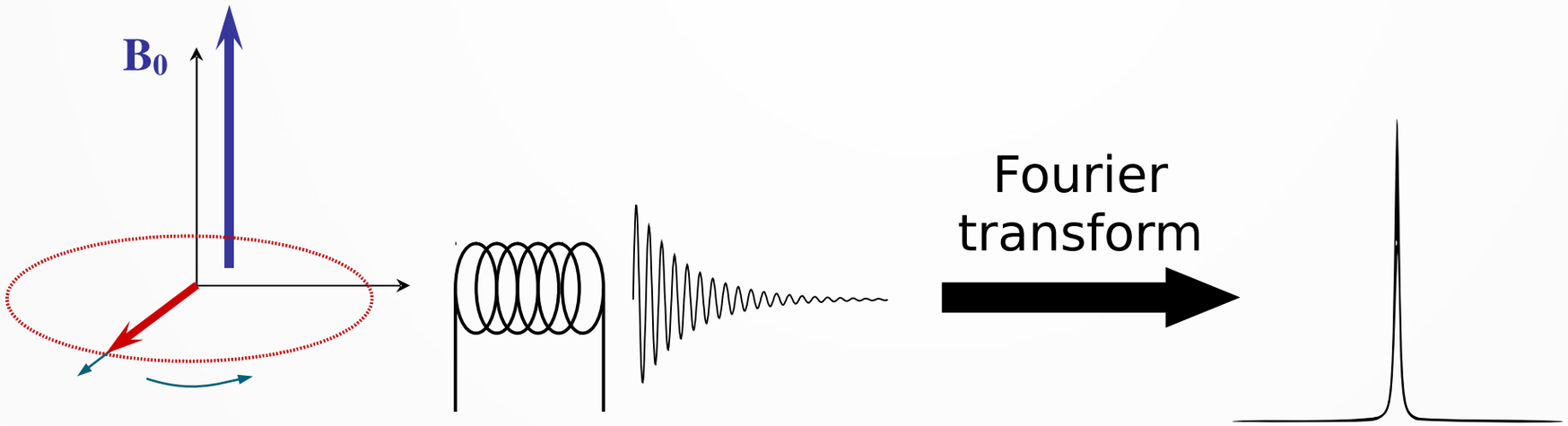
Same chemical shift at different B_0

- NMR signal

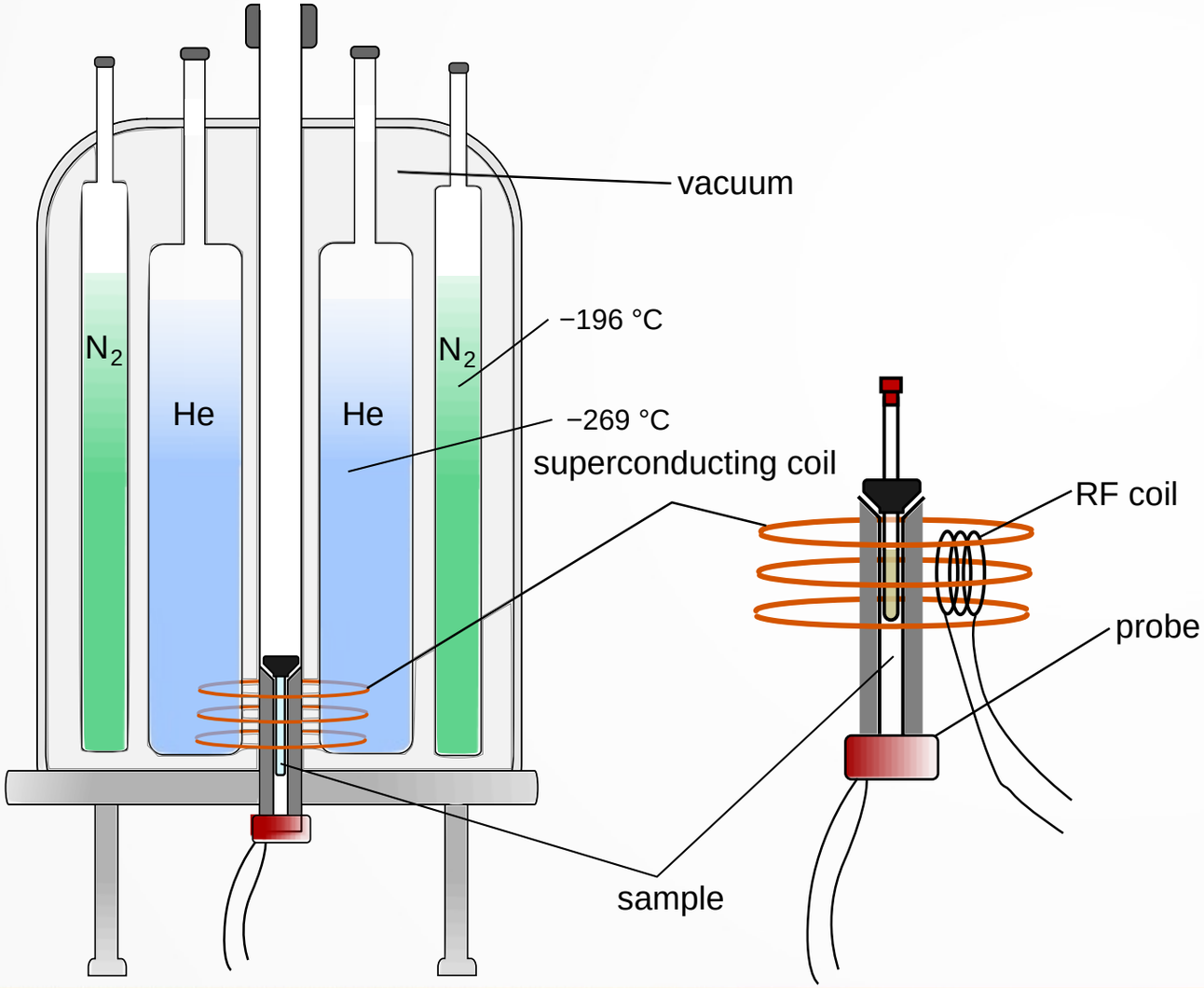


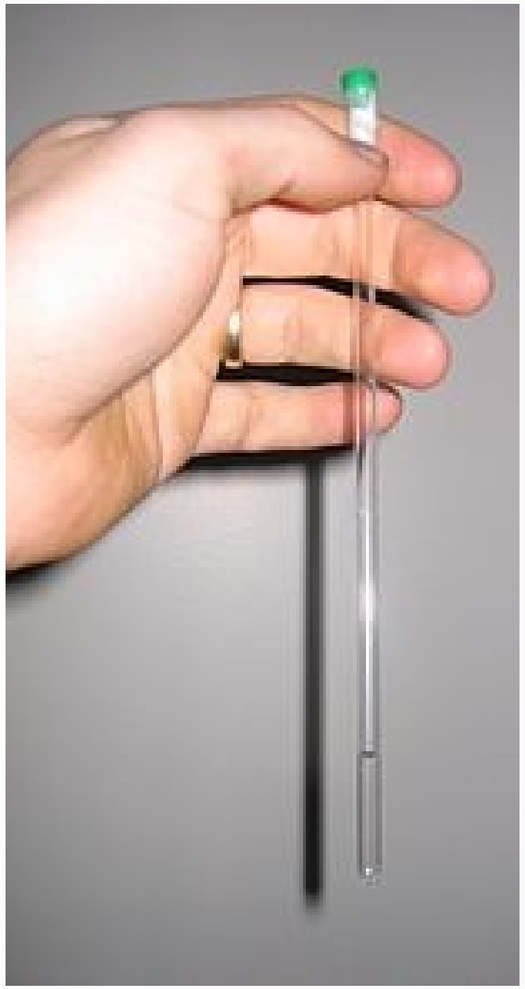
Manipulation with magnetization by radiofrequency filed B_1

Resonance frequency
$\omega_{RF} = \gamma B_0$

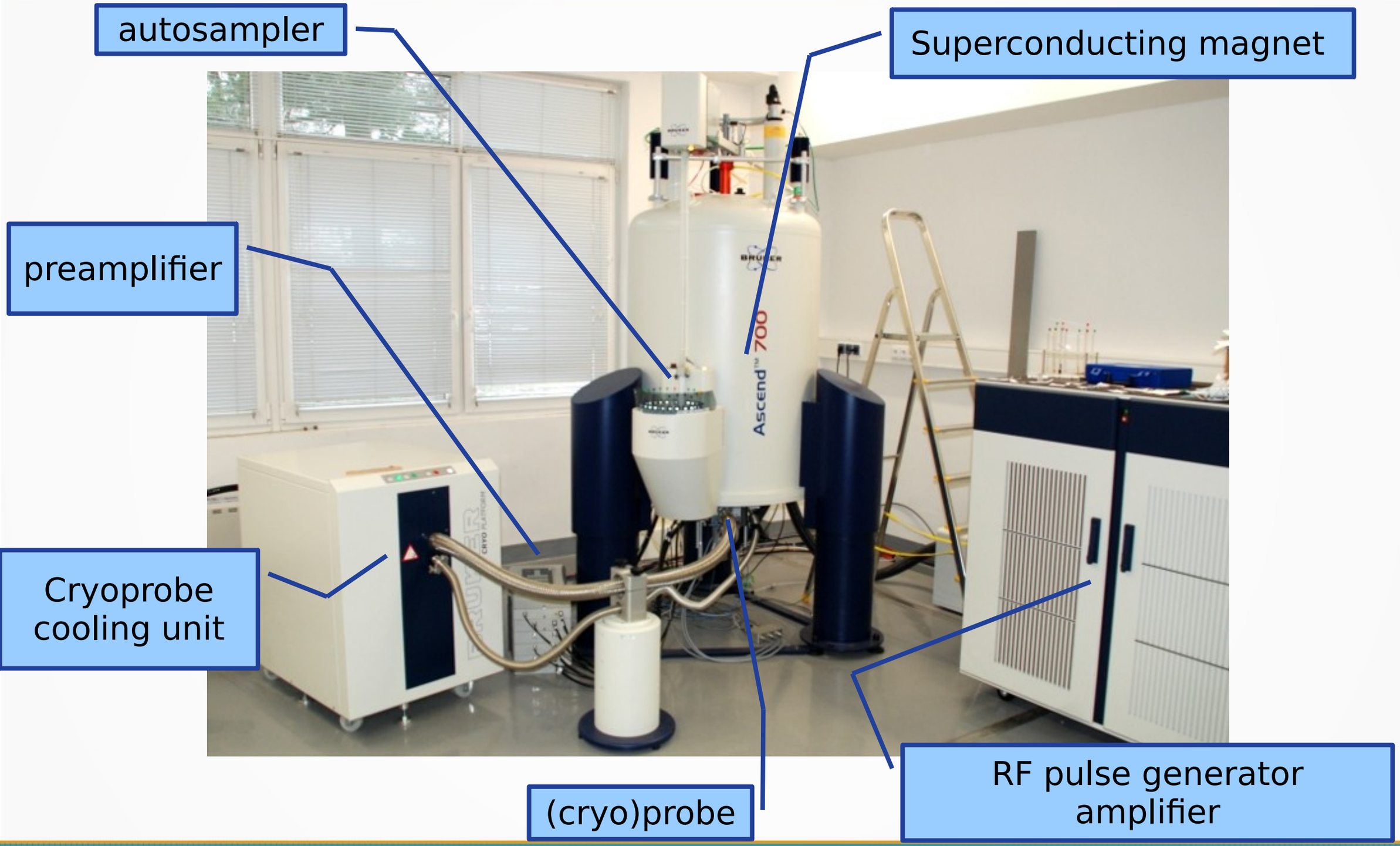


NMR spectrometer

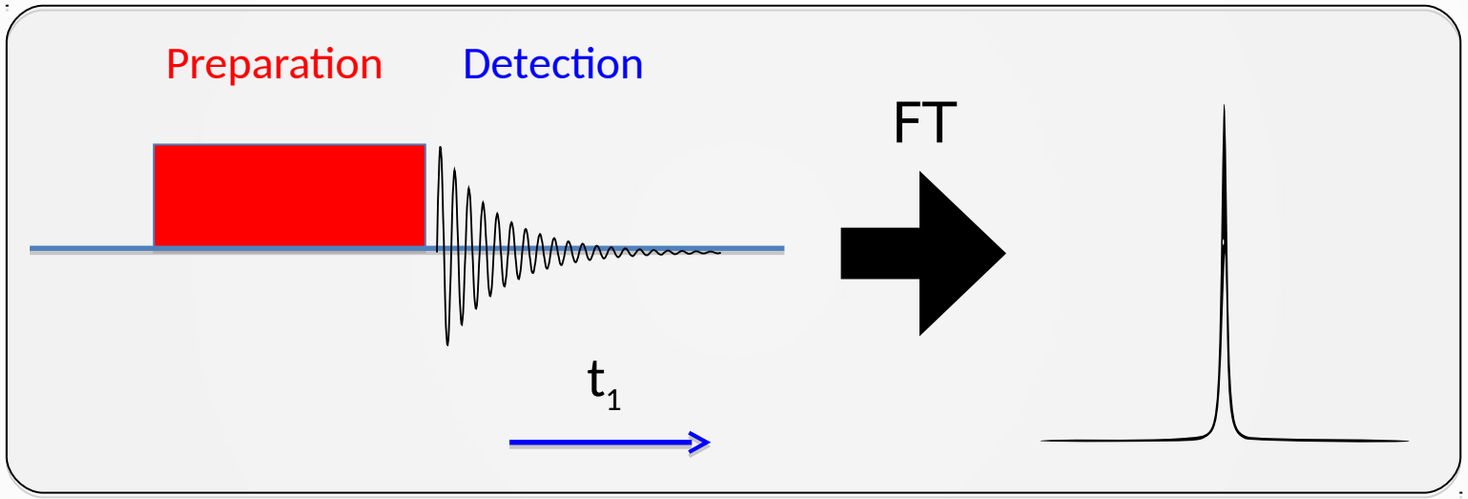




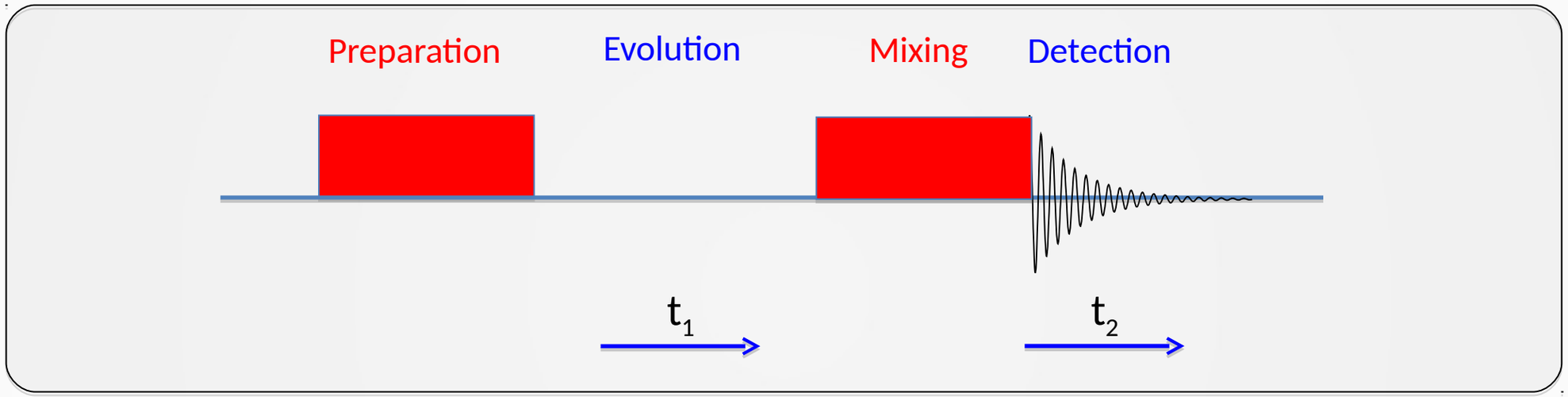
Sample in NMR tube


autosampler
Superconducting magnet
preamplifier
Cryoprobe cooling unit
(cryo)probe
RF pulse generator amplifier
Ascend™ 700

1D experiment



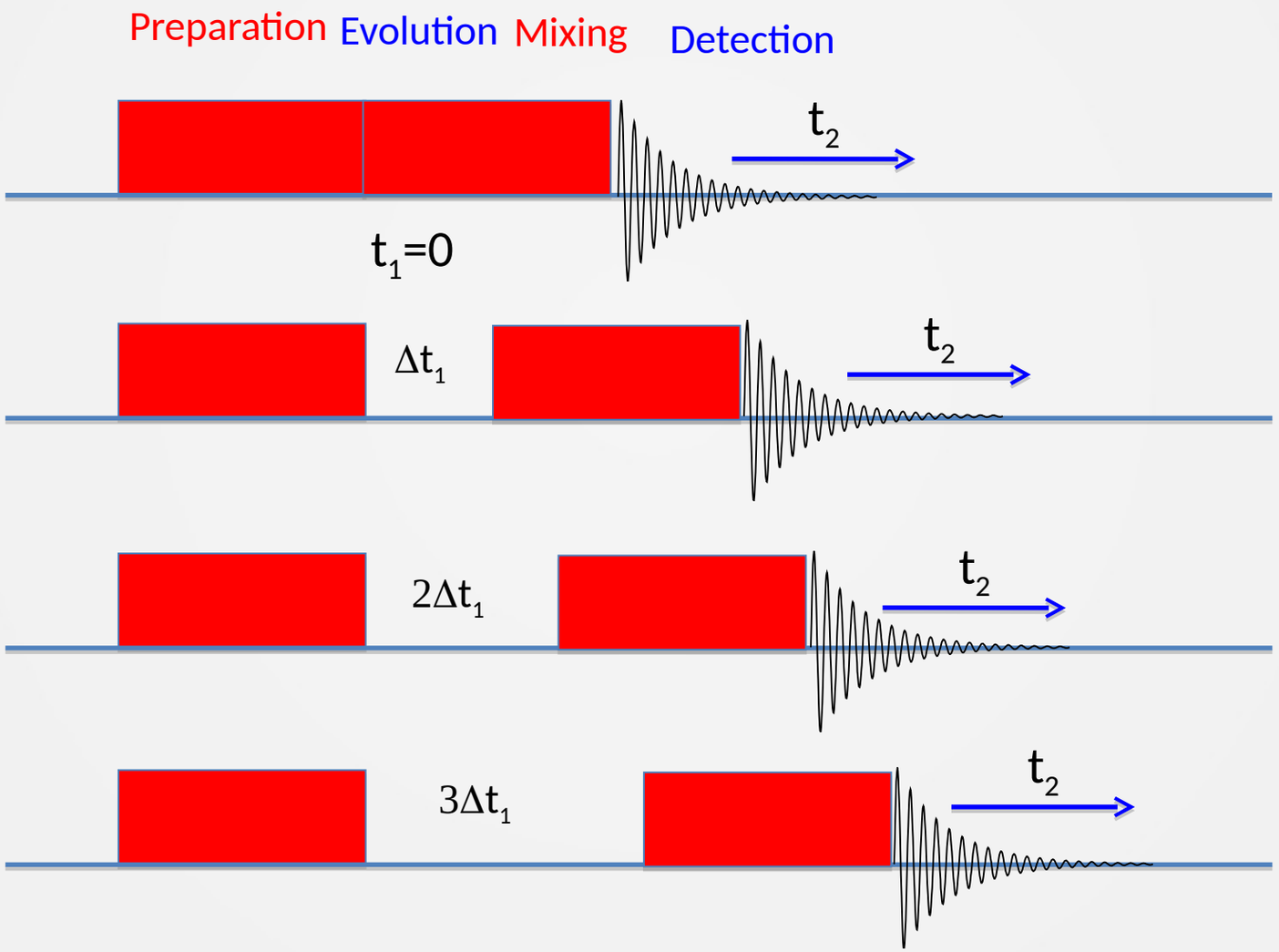
2D experiment



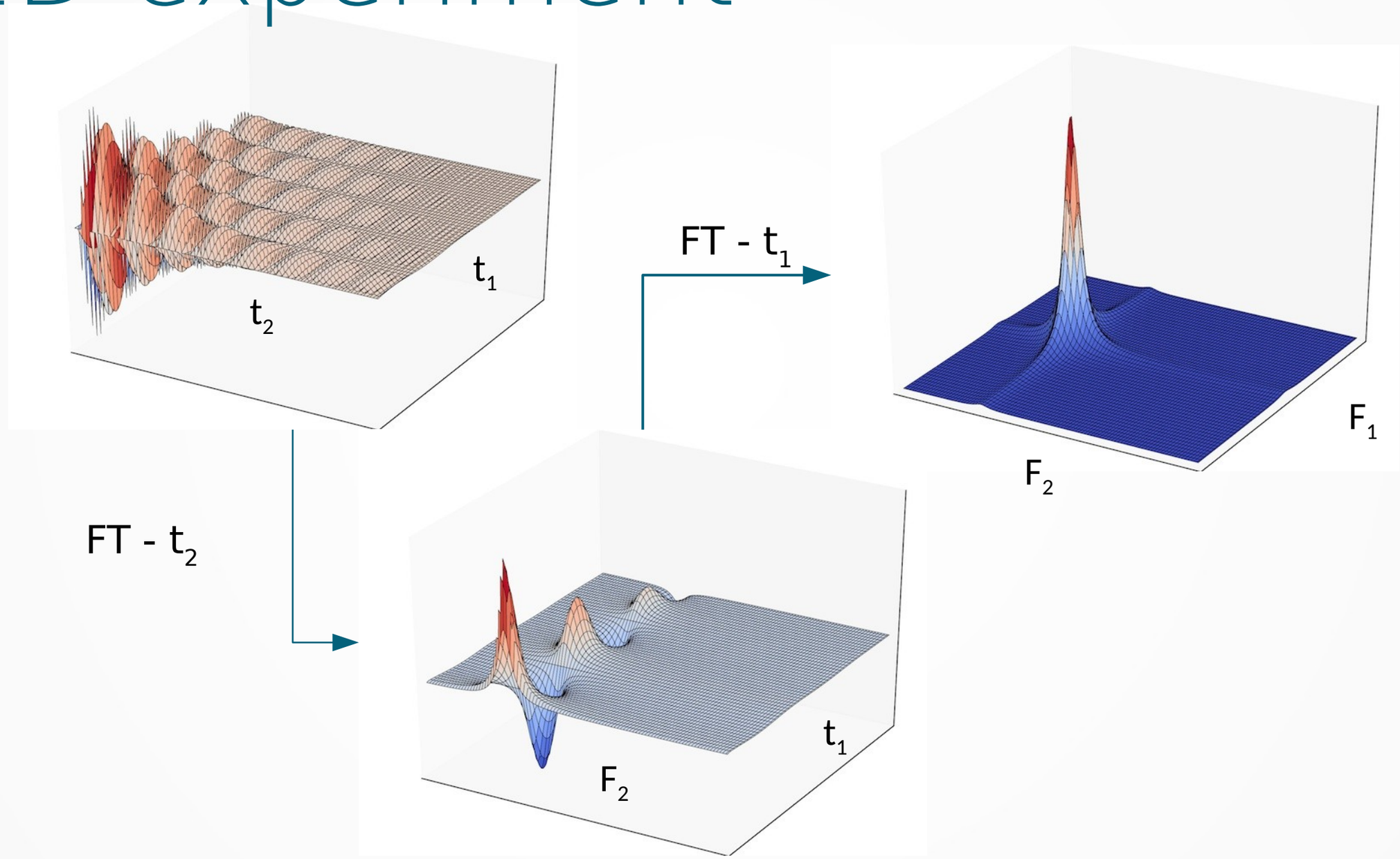
Preparation and Mixing do not change during experiment.

Jean Jeener, AMPERE Summer School in Basko Polje, Yugoslavia, September 1971

2D experiment



2D experiment



What can we do with NMR:

- 3D structure determination
- Protein – ligand interaction (small molecule, DNA/RNA, protein, atom) – even very weak interactions (K_d in mM)
- Monitoring of biochemical processes
- Dynamics (ps → days)
- intrinsically disordered proteins(IDPs)
- “In-cell” NMR

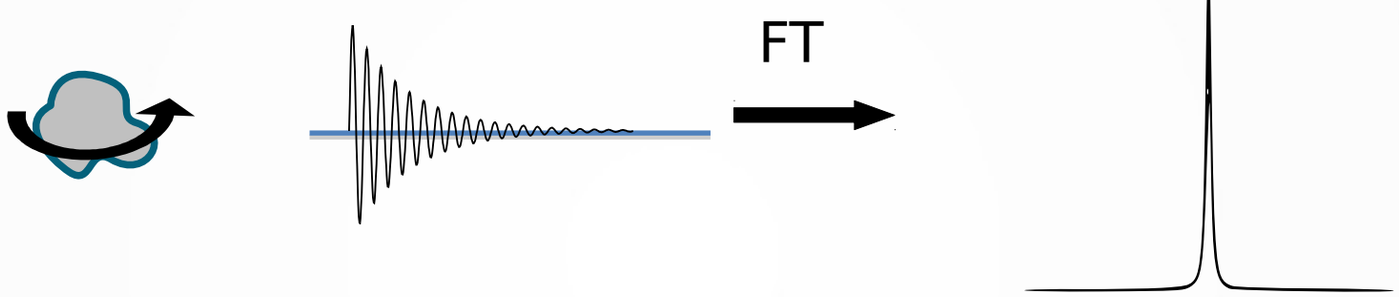
Limitations:

- Protein size
  - ✓ < 10 kDa – similar to small molecules (COSY, TOCSY, NOESY, …)
  - ✓ < 20 kDa – 100 % ^{13}C and ^{15}N isotopic enrichment, field ≥ 500 MHz
  - – up to ~100 kDa – 100 % ^{13}C, ^{15}N and ^{2}H isotopic enrichment, field ≥ 800 MHz
  - – big proteins – changes in structure/dynamics, selective labeling, field ≥ 900 MHz
- Sample volume and concentration
  - ✓ 300 – 600 µl, ≥ 0.2 mM, low salt concentration, (deuterated buffer)
- Sample stability
  - ✓ ideally weeks (at least ~ 3 days → fresh sample preparation)

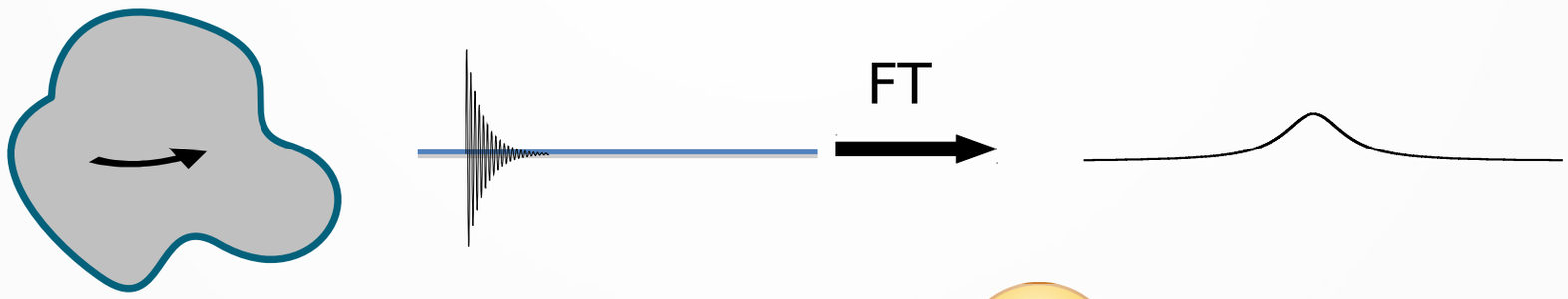
Limitations:

- Protein size!!!

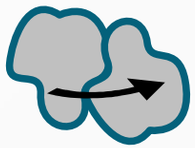
Small molecule – fast tumbling



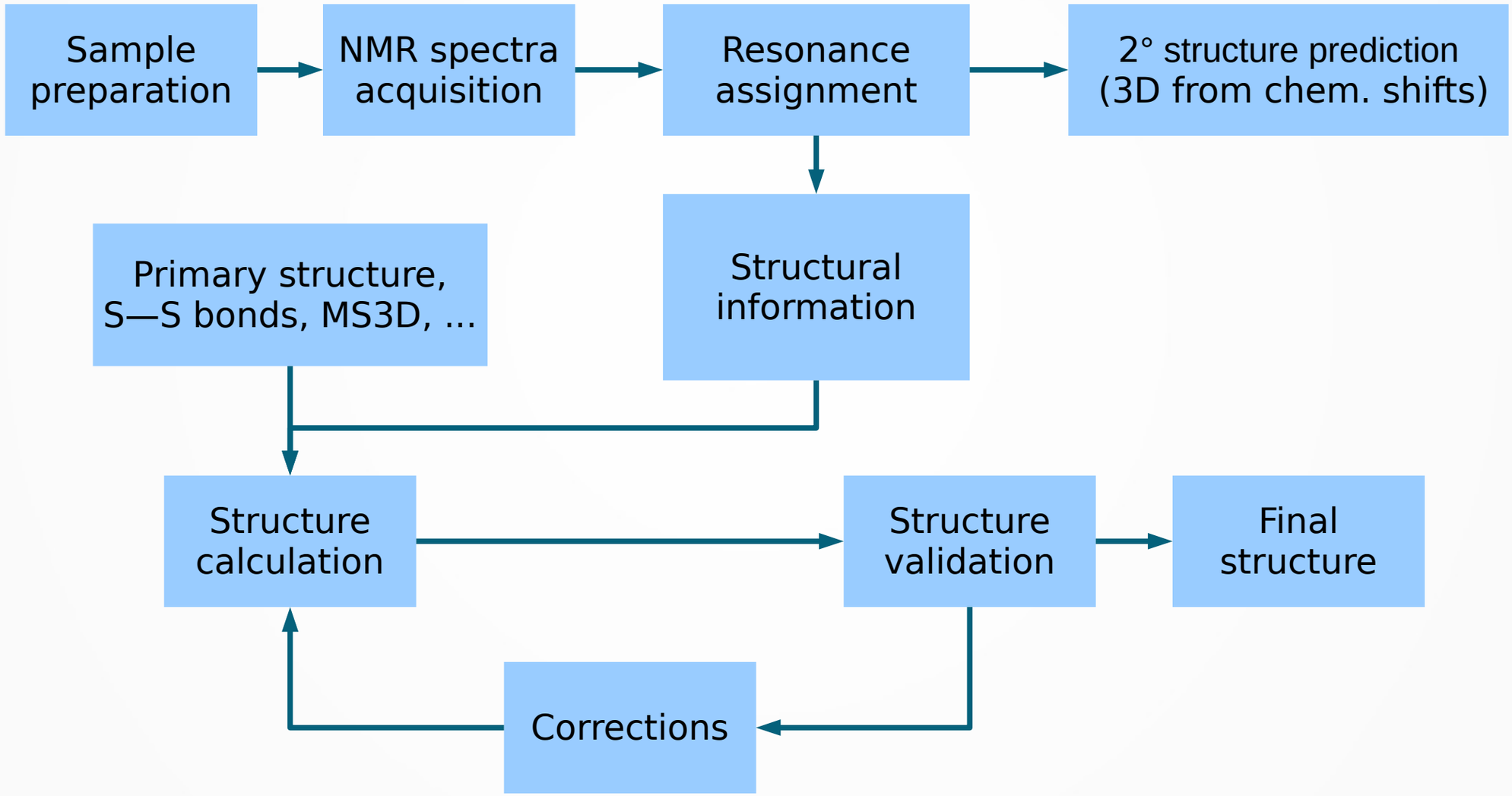
Big molecule – slow tumbling



oligomers



Procedure


Sample preparation
NMR spectra acquisition
Resonance assignment
2° structure prediction (3D from chem. shifts)
Structural information
Primary structure, S—S bonds, MS3D, ...
Structure calculation
Structure validation
Final structure
Corrections

Sample

- Original organism
  - ✔ Natural form (posttranslational modification)
  - ✗ Small amount, difficult isolation, no isotopic enhancement
- Protein expression in organism (bacteria, yeast, insect or mammalian cells)
  - ✔ easy uniform isotopic labeling ($^{15}N_4Cl$, [U-^{13}C] glucose), cheap, high yields, ^{2}H labeling, selective labeling
  - ✗ no posttranslational modification (*E. Coli*)
- “Cell Free” protein expression
  - ✔ Toxic proteins, selective labeling
  - ✗ Price, (posttranslational modifications)

NMR spectra acquisition

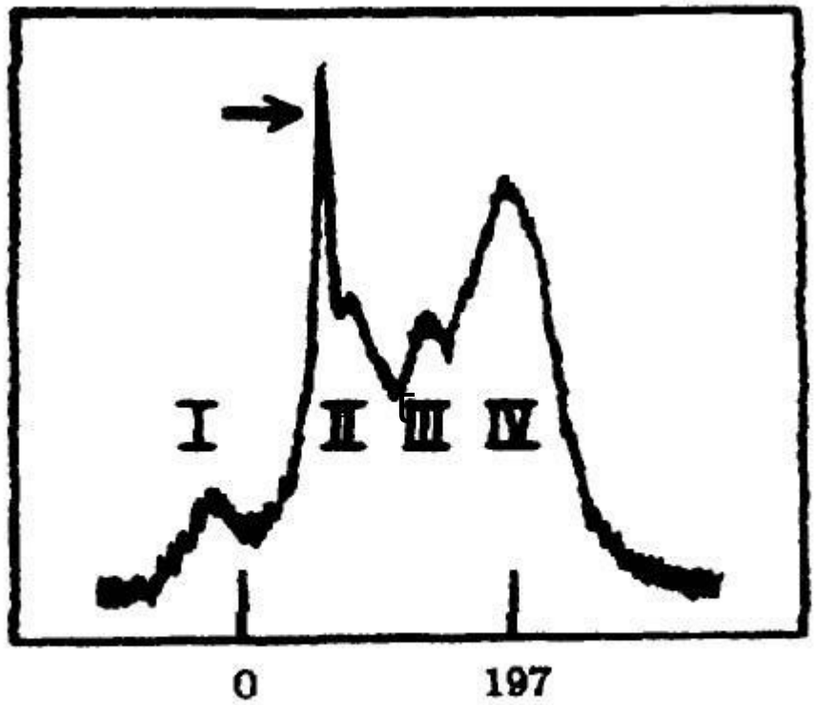
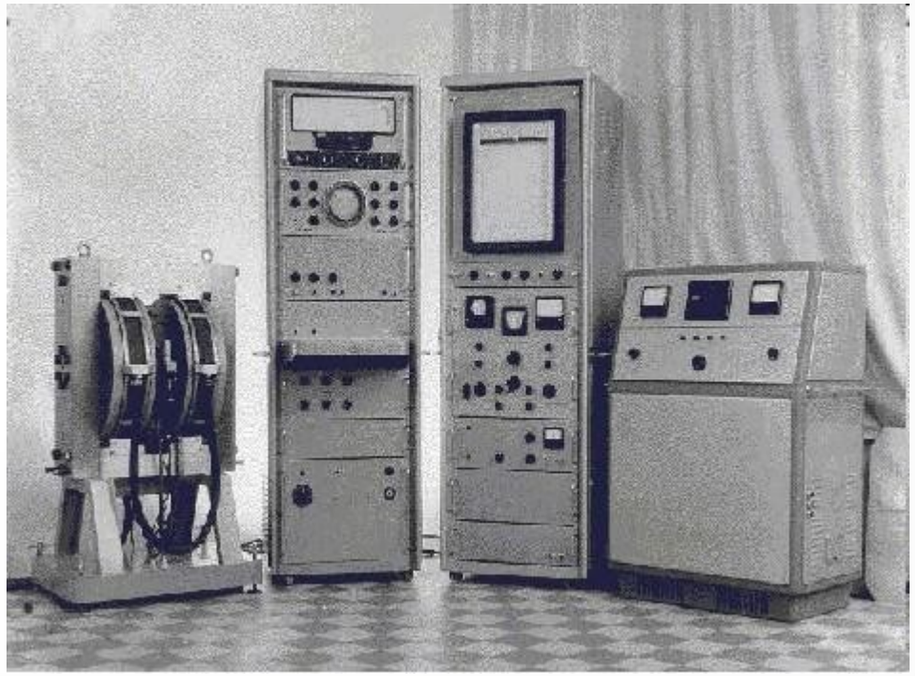


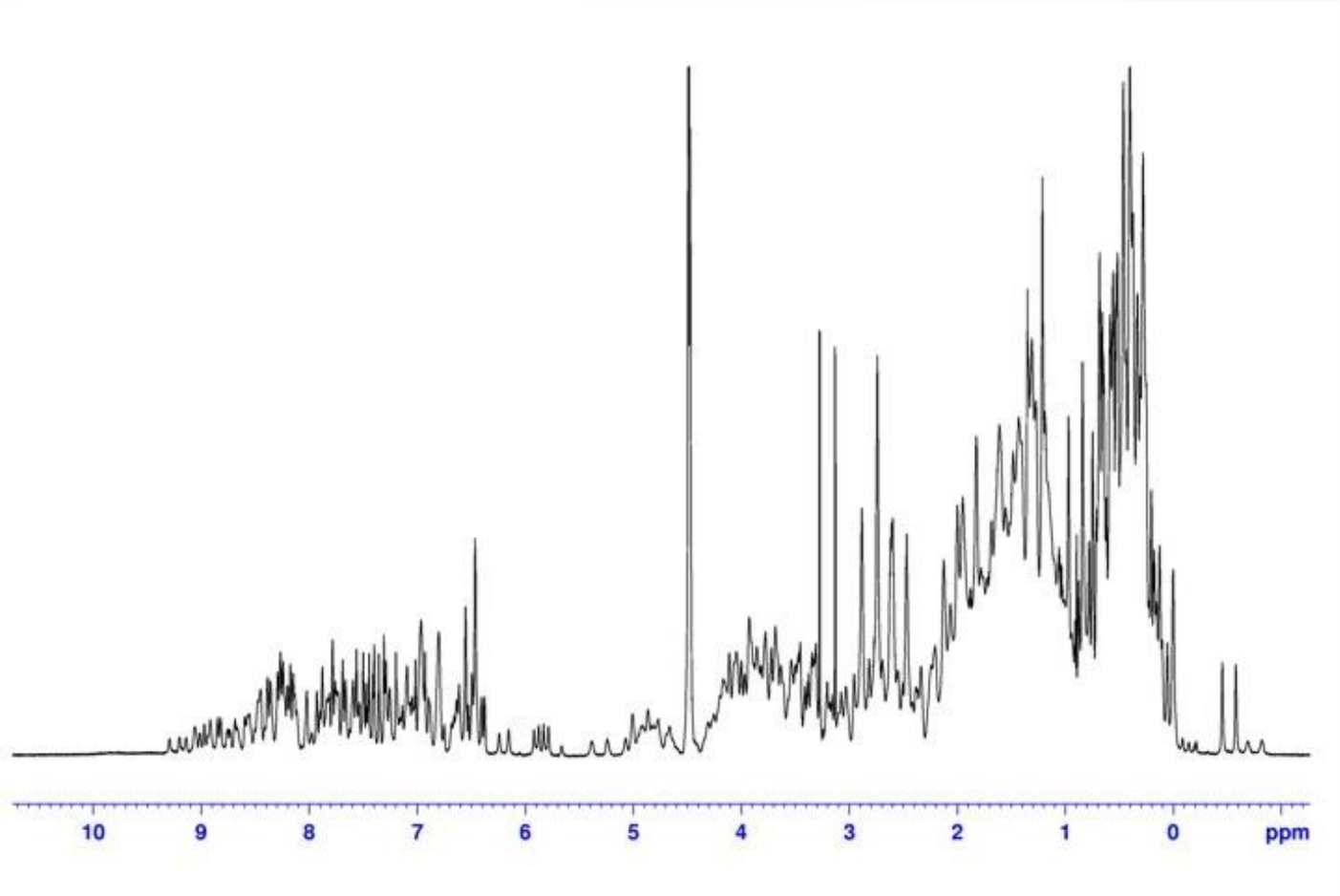
Fig. 1. The first NMR spectrum of a protein (RNase A; 20% (w/v) in D_2O) published, a single slow scan at 40 MHz [14].



40 MHz CW spectrometer

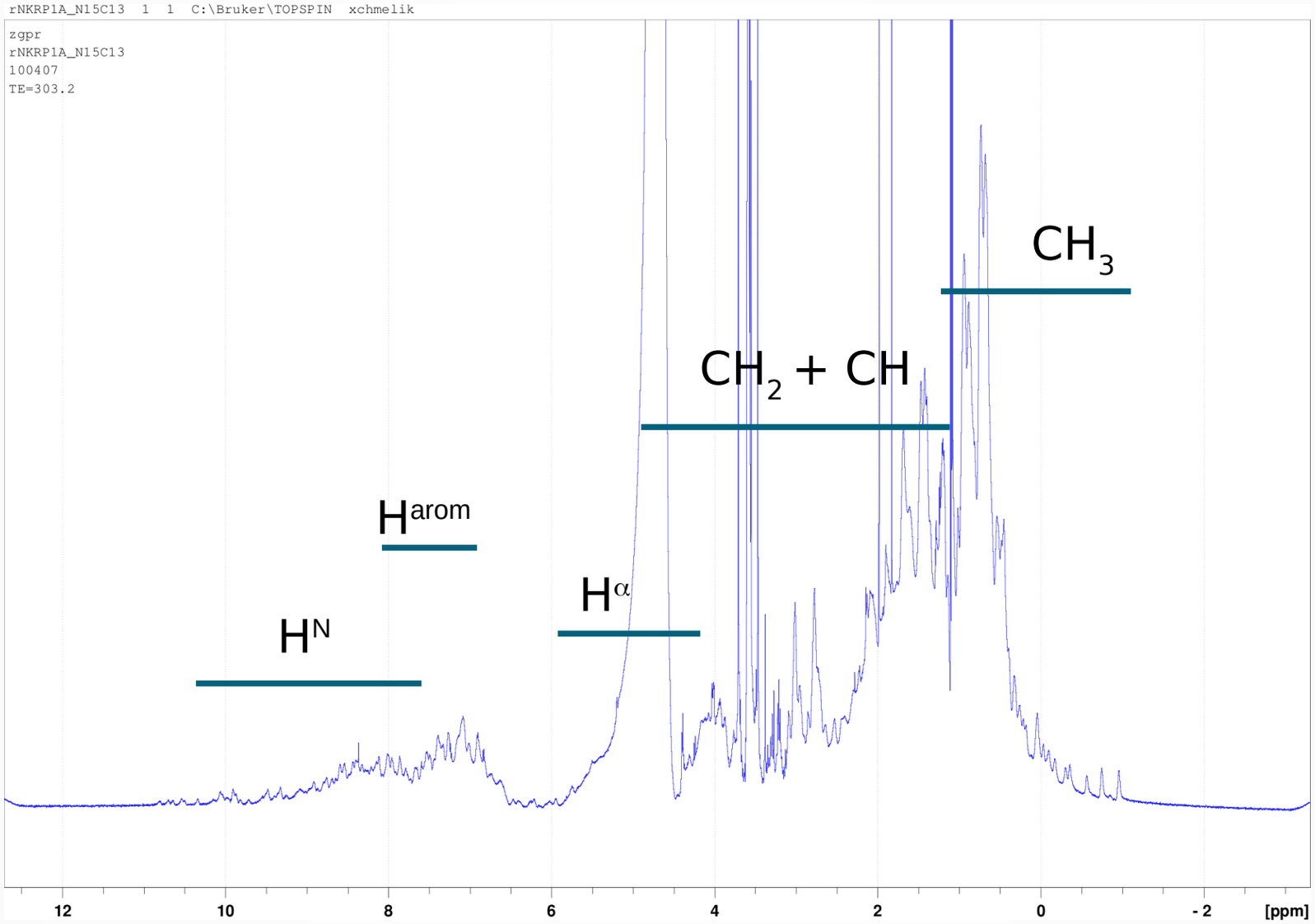
The first published protein NMR spectrum (1957)

1 GHz NMR spectrometer (23.5 T)



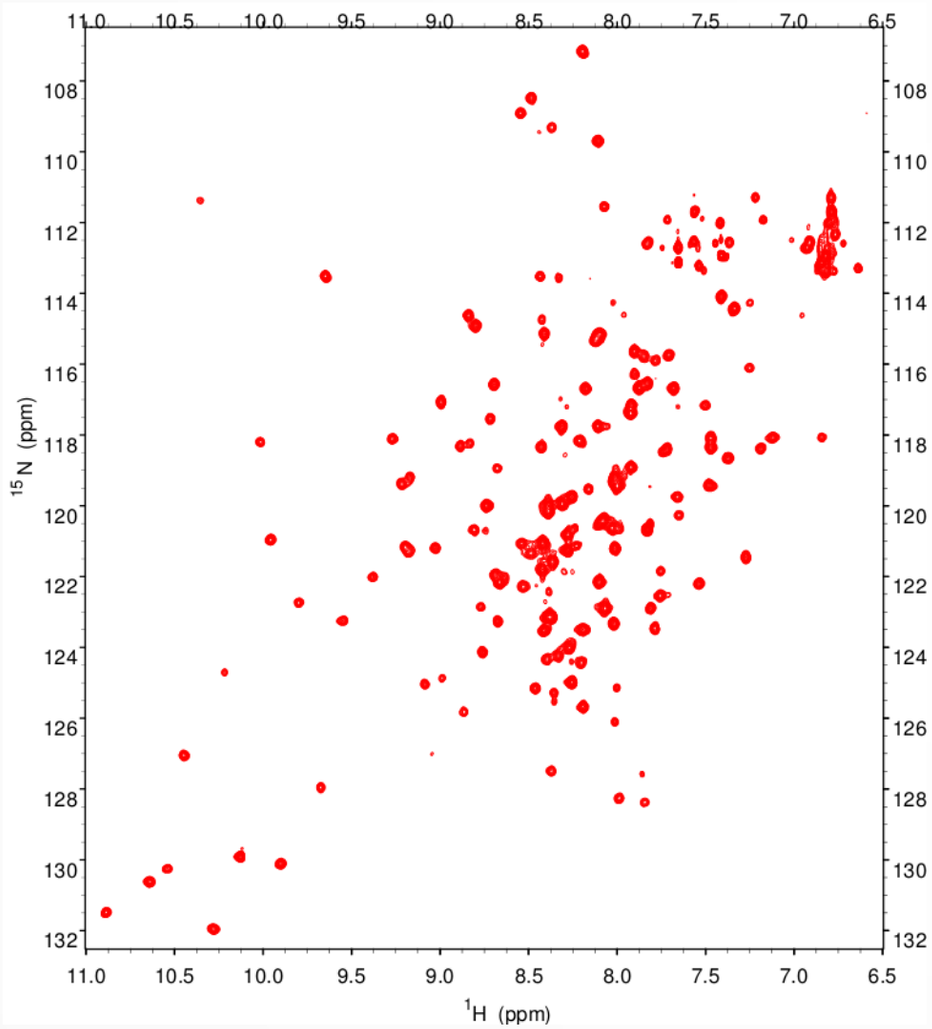
2.0 mM [^{13}C,^{15}N]-Ubiquitin in 90% H_2O/10% D_2O

Protein 1D proton NMR spectrum

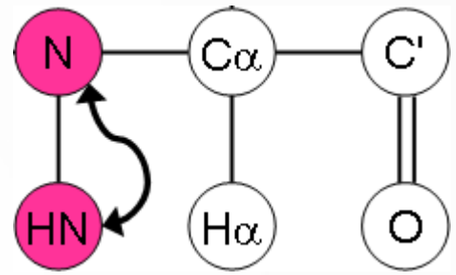


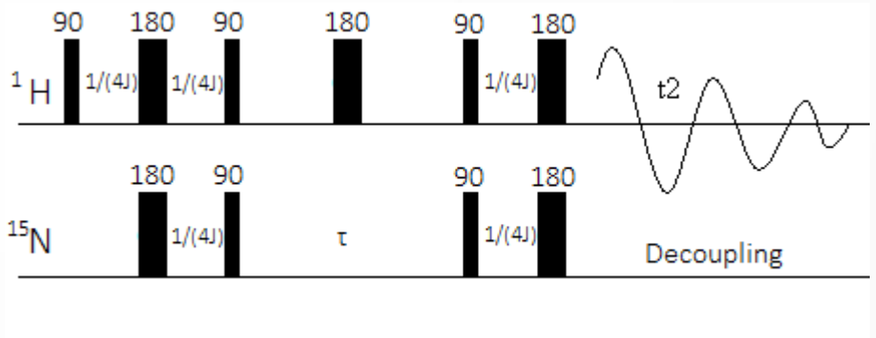
139 amino acid residues → 839 visible proton signals

2D H—N correlation

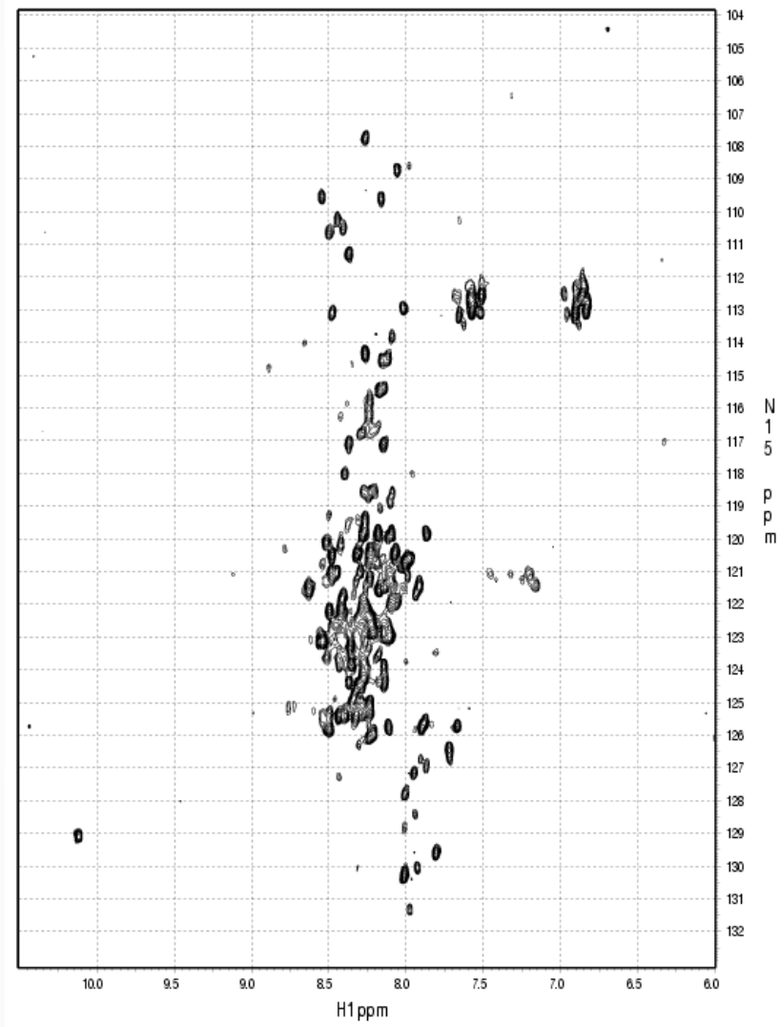


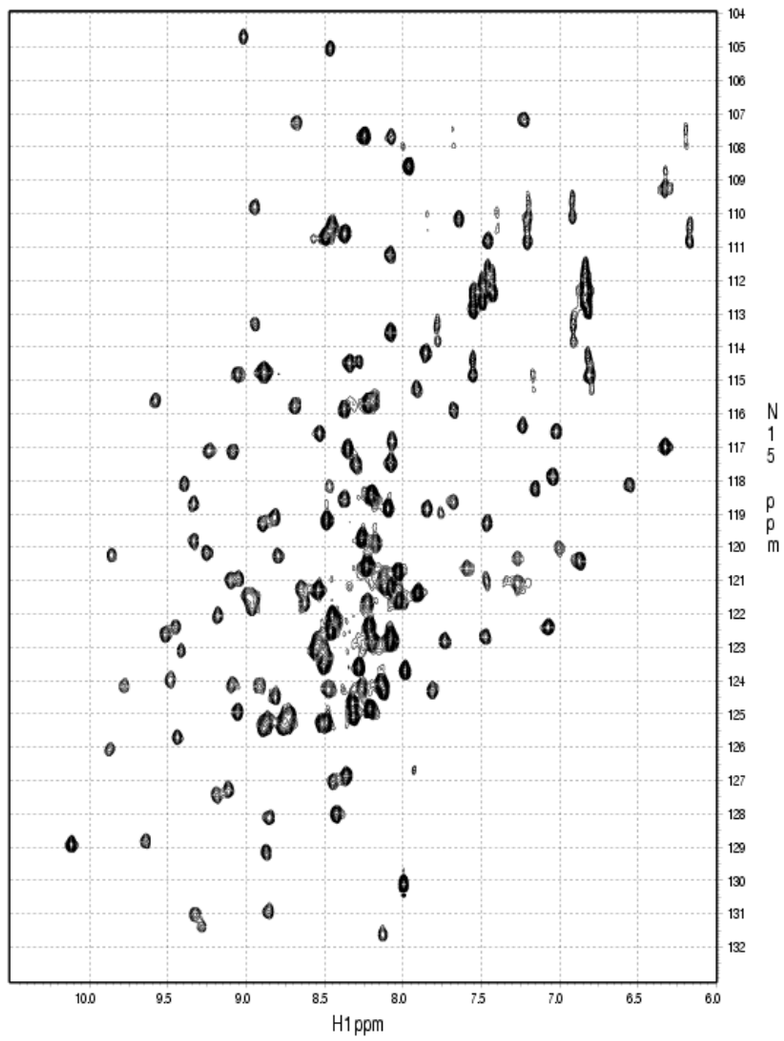
Protein “fingerprint”



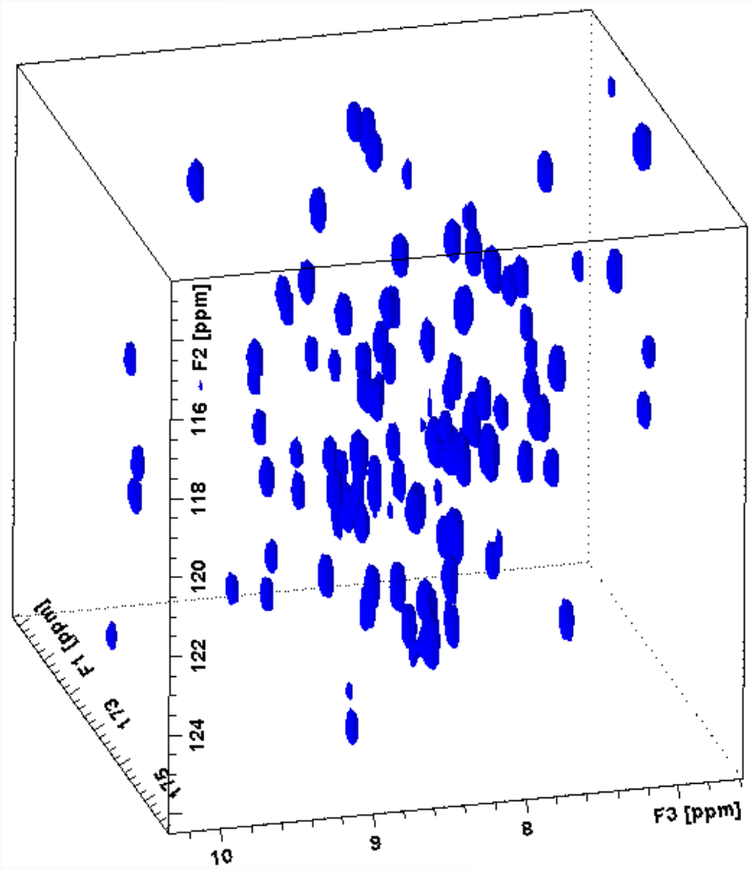


Folded vs unfolded protein

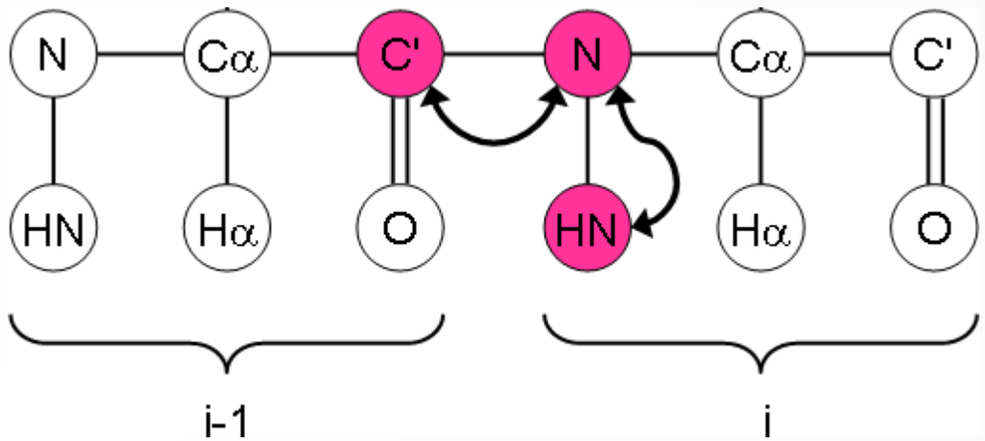


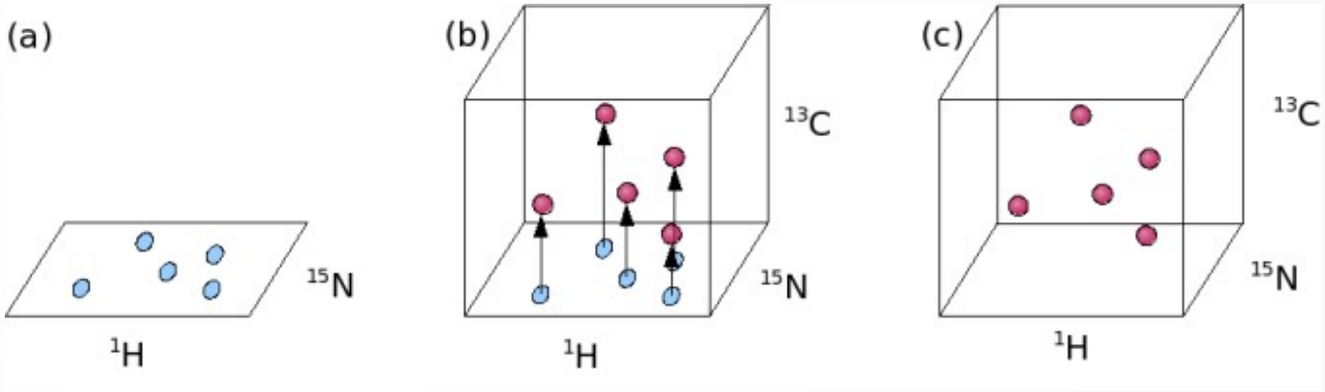


3D correlation



3D HNCO spectrum





Spectra for 3D structure

- Protein back-bone assignment
  - 1 2D experiment + 2–6 3D spectra
- Side-chain assignment
  - 2 2D + 2–5 3D spectra
- Inter-proton distances
  - 3 3D spectra

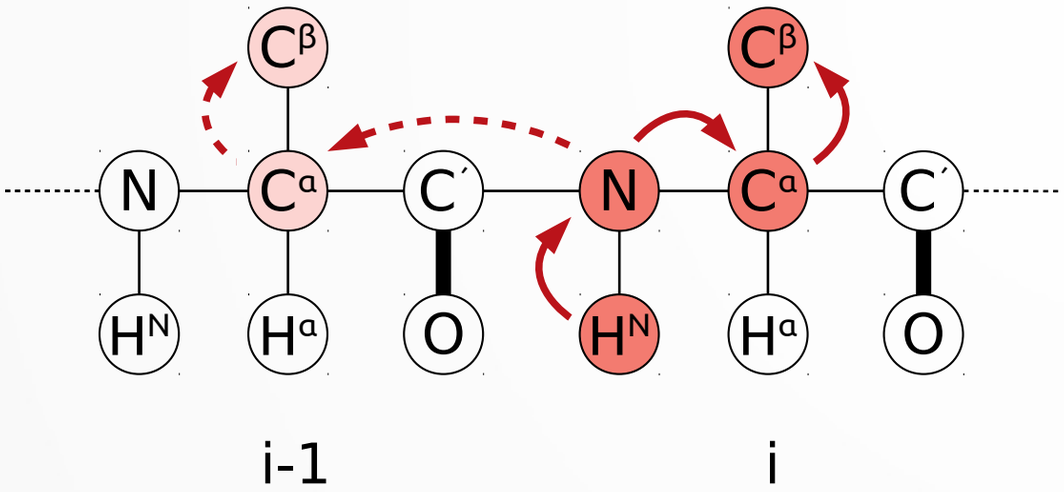
→ 2–3 weeks of NMR time

Intrinsically disordered protein → series of 4D, 5D, 6D experiments

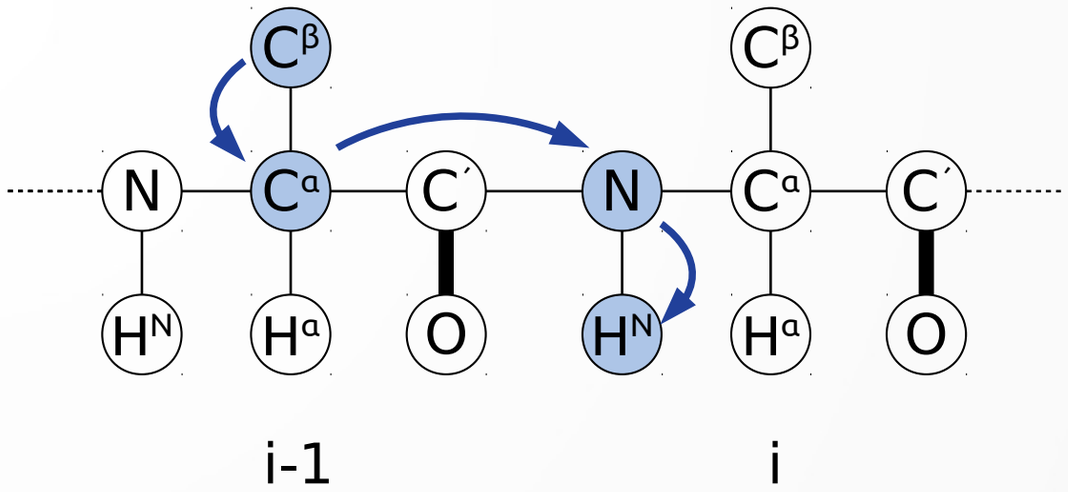
Atom resonance frequency assignment

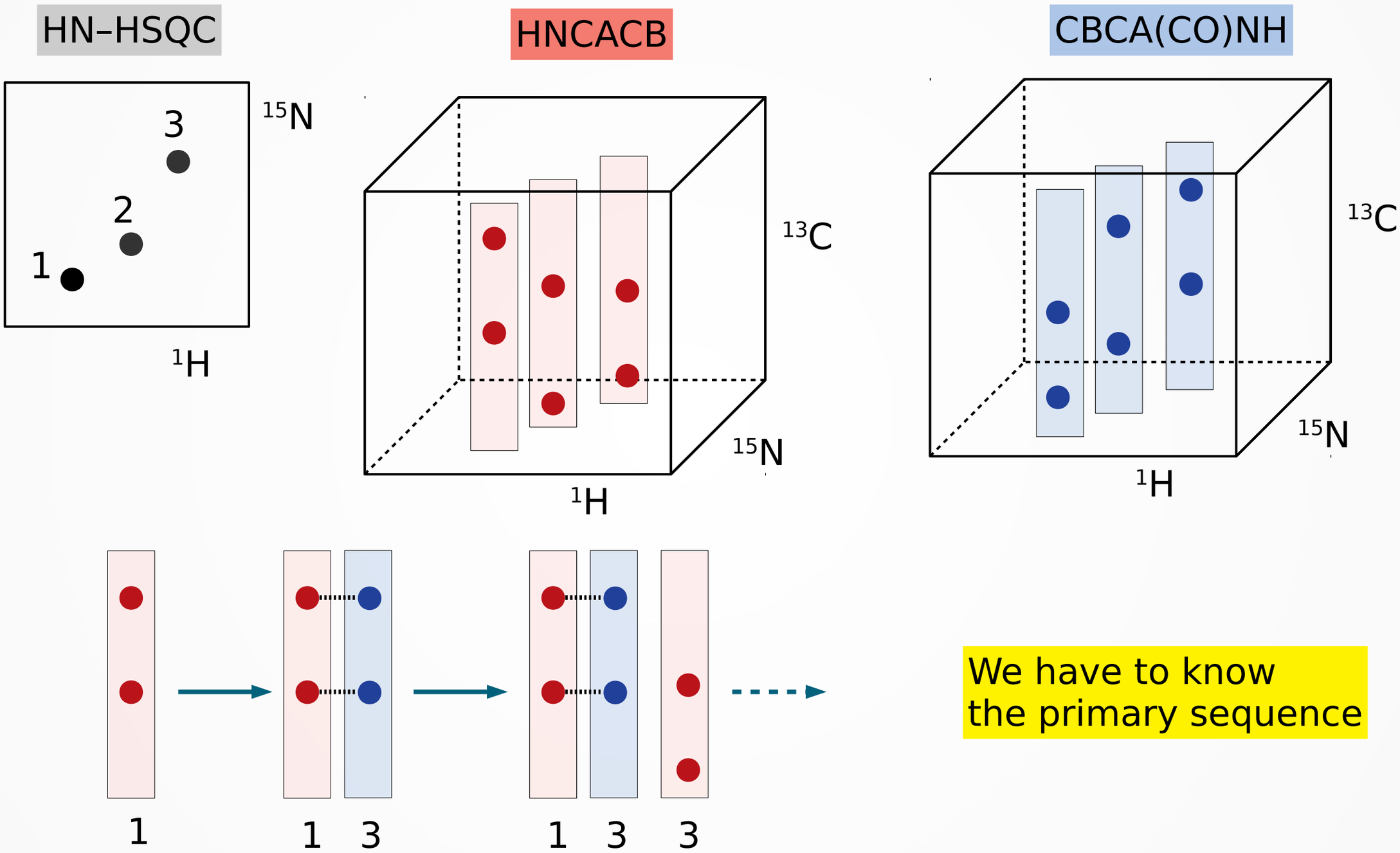
- 2 complementary experiments

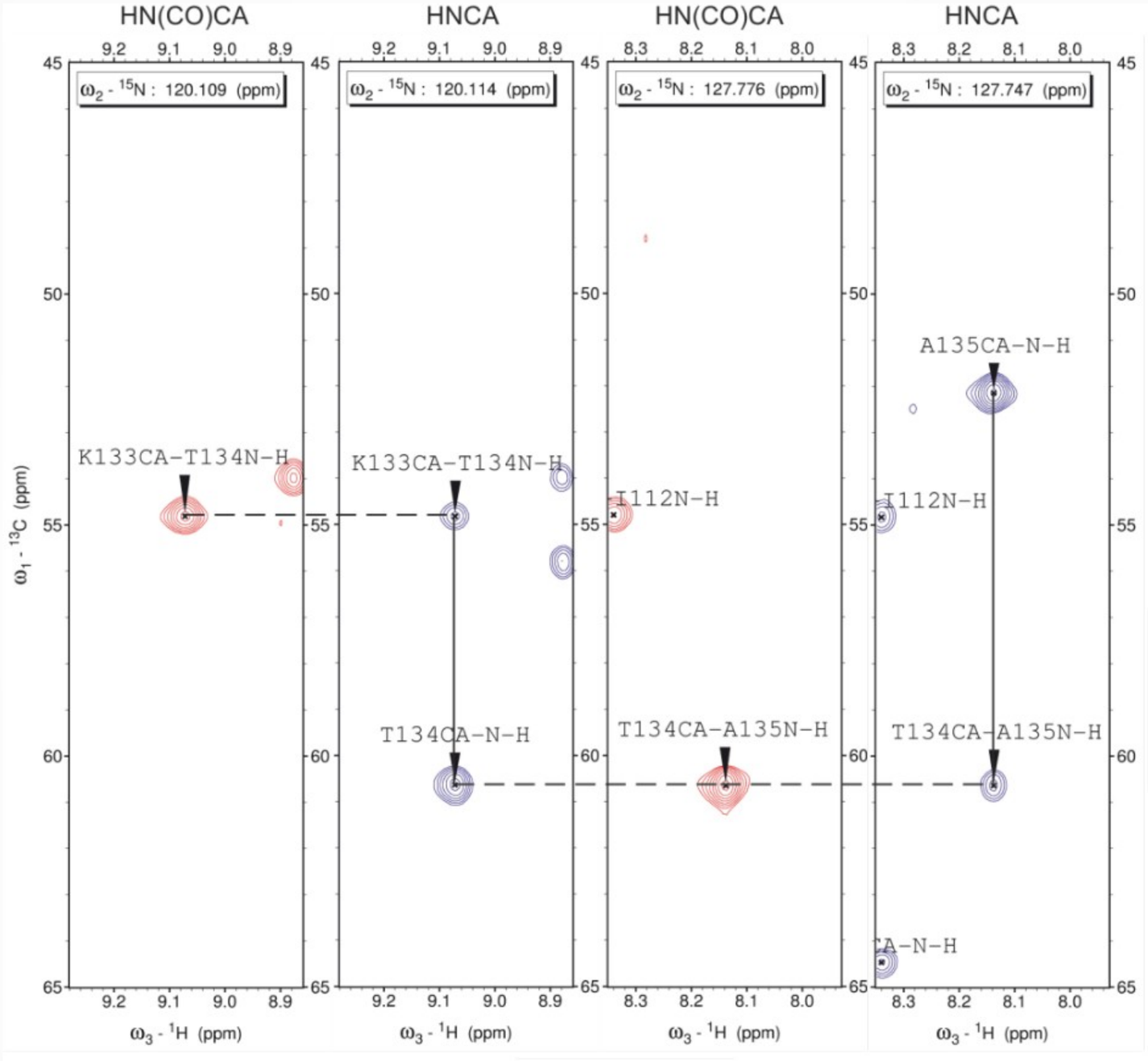
HNCACB



CBCA(CO)NH




HN–HSQC
HNCACB
CBCA(CO)NH
3
2
1
15N
1H
13C
15N
1H
13C
15N
1H
1
1 3
1 3 3
We have to know
the primary sequence


HN(CO)CA
HNCA
HN(CO)CA
HNCA
ω2 - 15N : 120.109 (ppm)
ω2 - 15N : 120.114 (ppm)
ω2 - 15N : 127.776 (ppm)
ω2 - 15N : 127.747 (ppm)
K133CA-T134N-H
K133CA-T134N-H
I112N-H
I112N-H
A135CA-N-H
T134CA-N-H
T134CA-A135N-H
T134CA-A135N-H
ω1 - 13C (ppm)
ω3 - 1H (ppm)

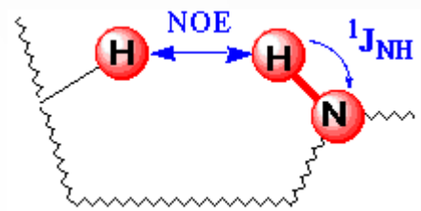
Structural information from NMR

- Inter-proton distances – NOE
- Torsion angles – scalar interaction constant, chemicals shifts
- Orientation of bonds in space – residual dipolar couplings
- …

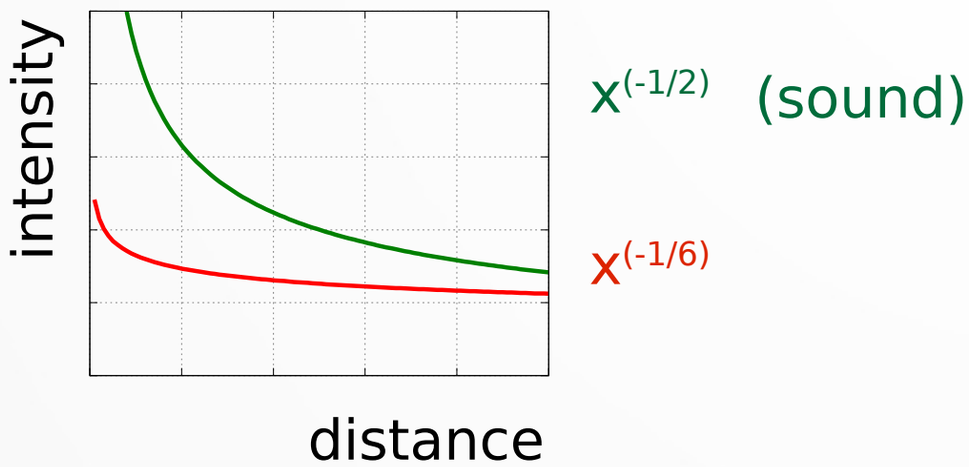
Inter-proton distances

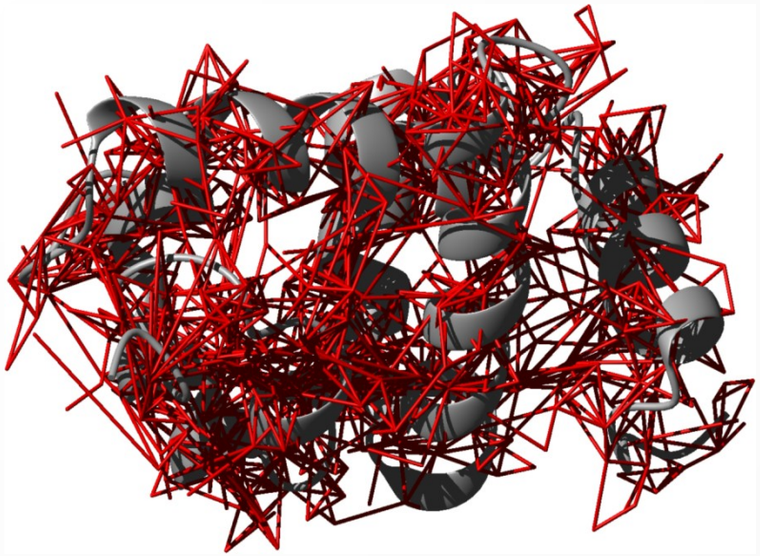
- Main source of structural information
- Short-range interaction:

  distances up to 5–6 Å

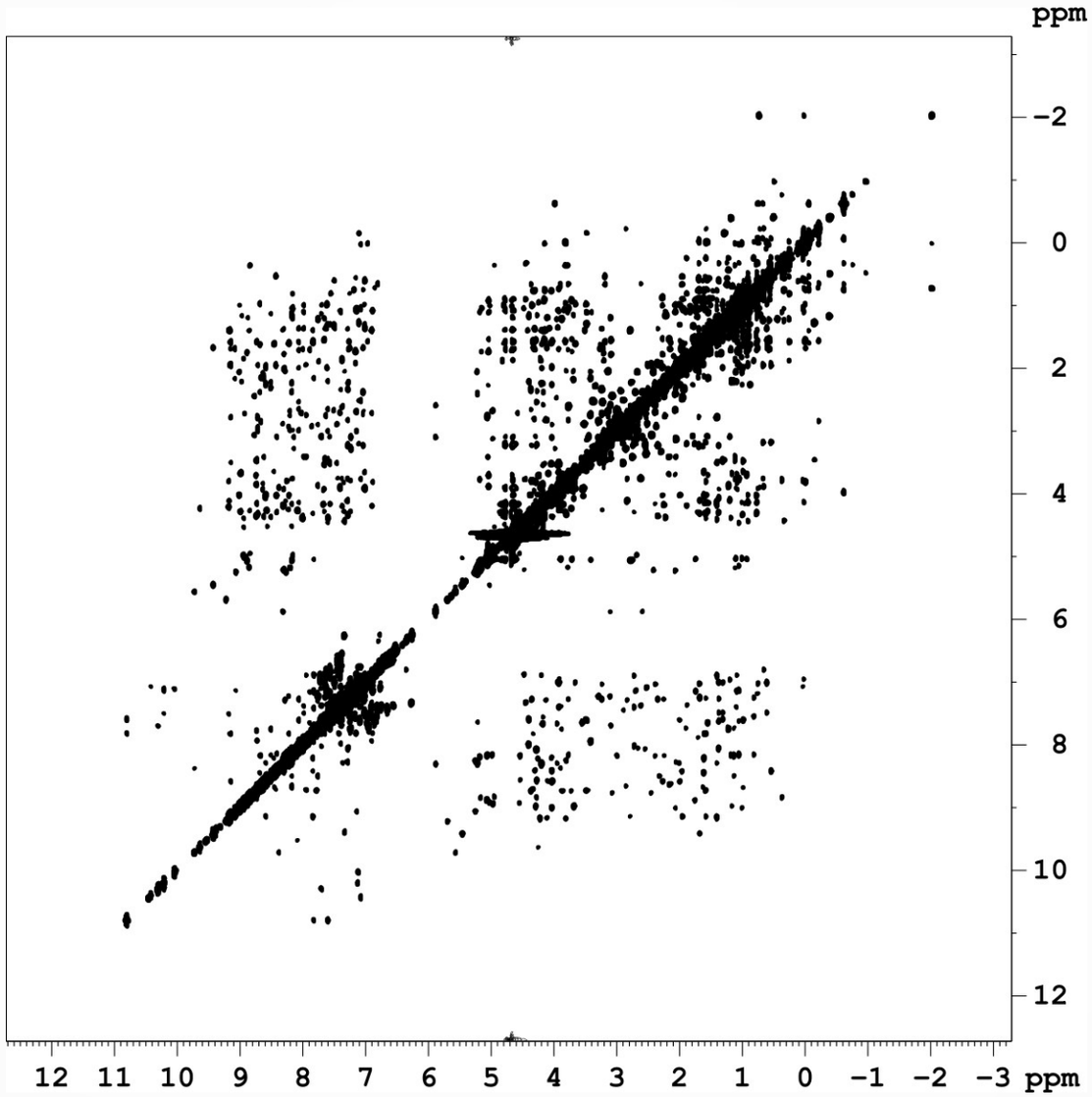


$$r \sim V^{-\frac{1}{6}}$$

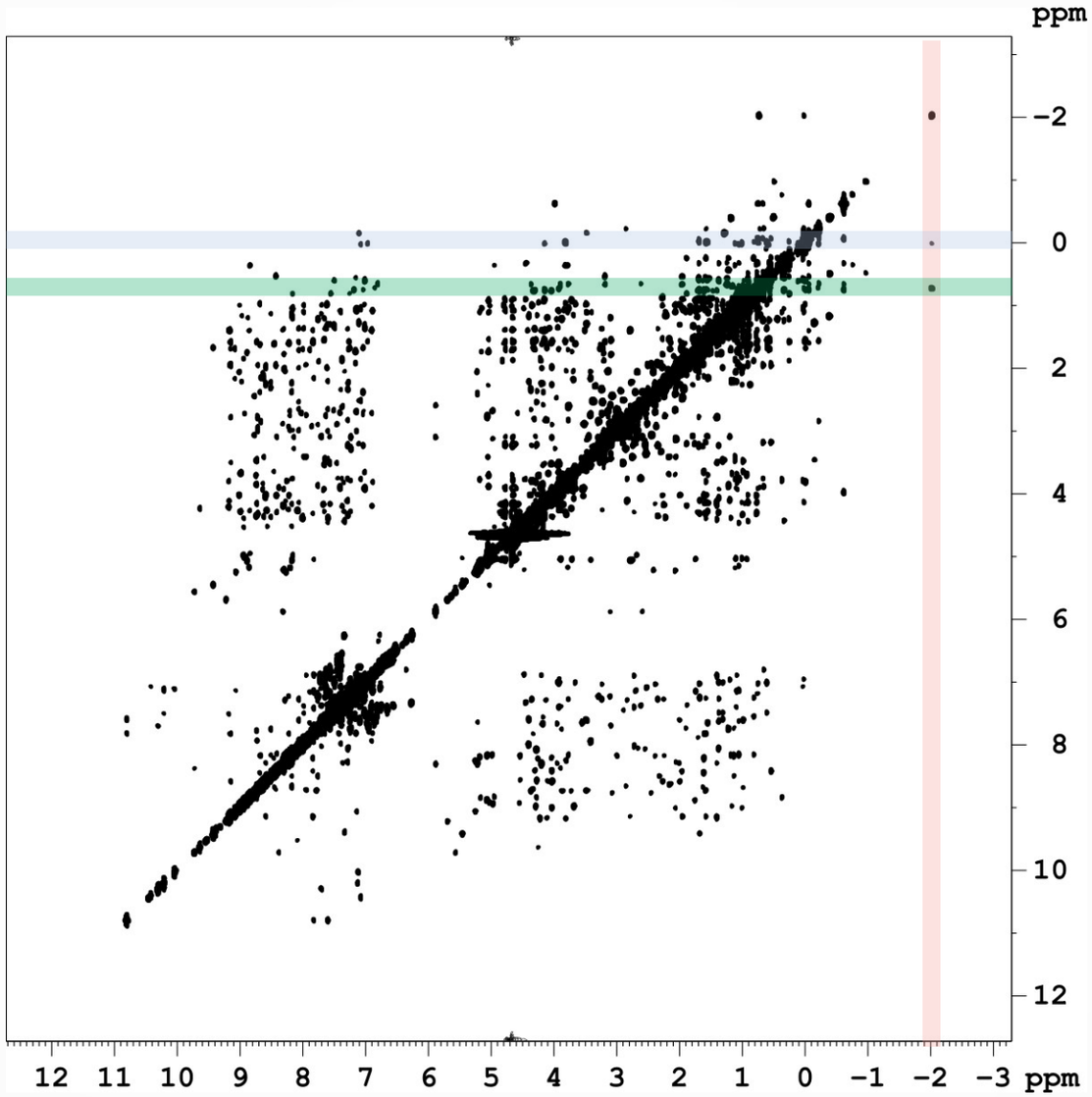


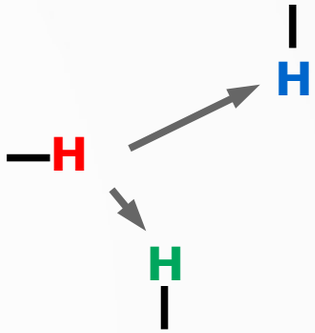


2D NOESY

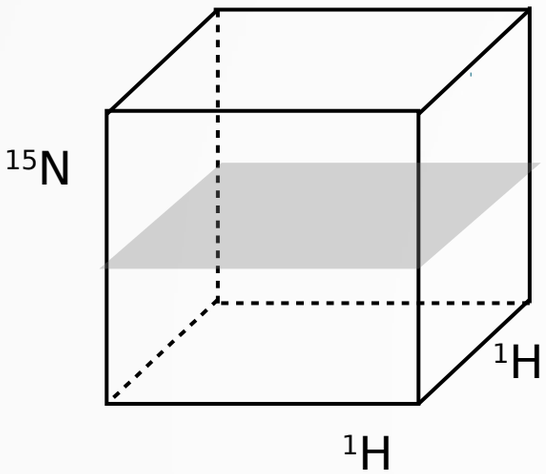


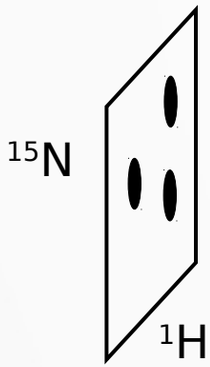
2D NOESY

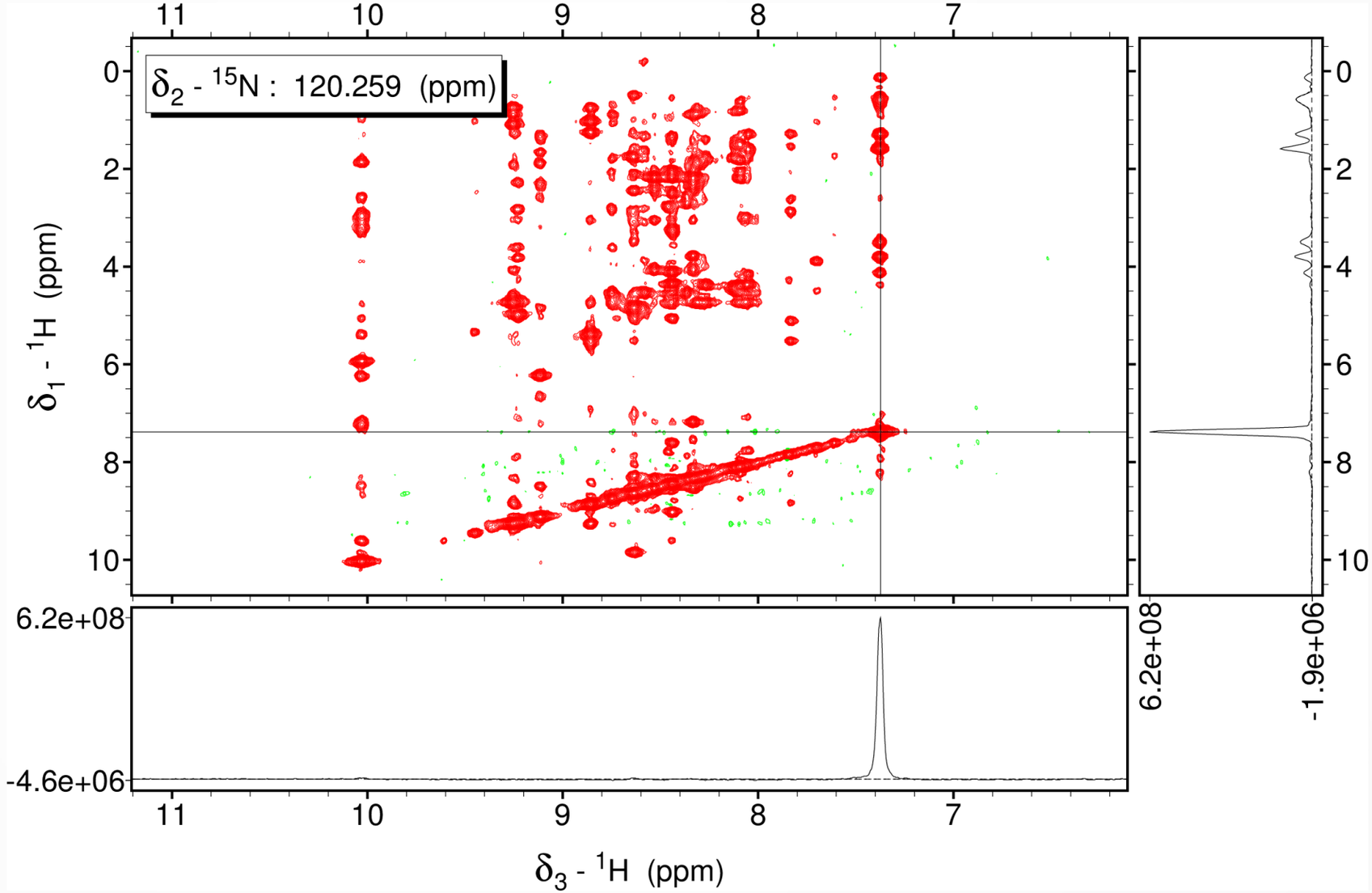




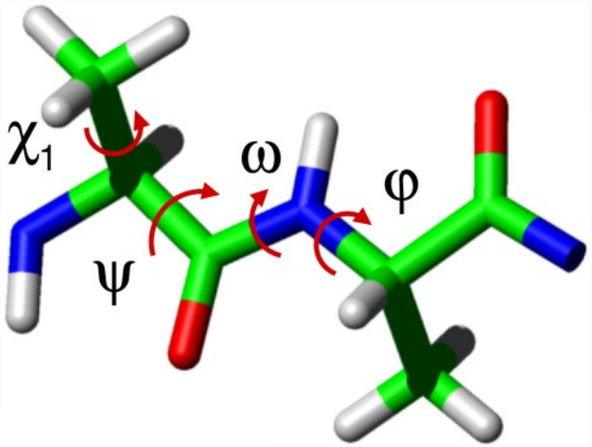
3D NOESY – HSQC







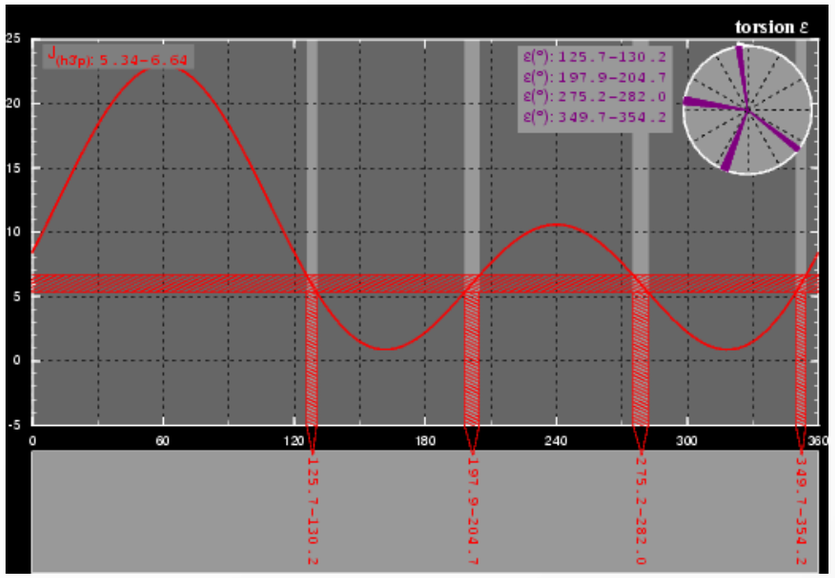
Torsion angles



- Angles φ, Ψ, χ_1, ω (*cis/trans*)
- Karplus equation
  - ^3J scalar interaction constant depends on torsion angle

    $^3J = A\cos^2(\varphi) - B\cos(\varphi) + C$
- 4 possible solutions
- Difficult to measure

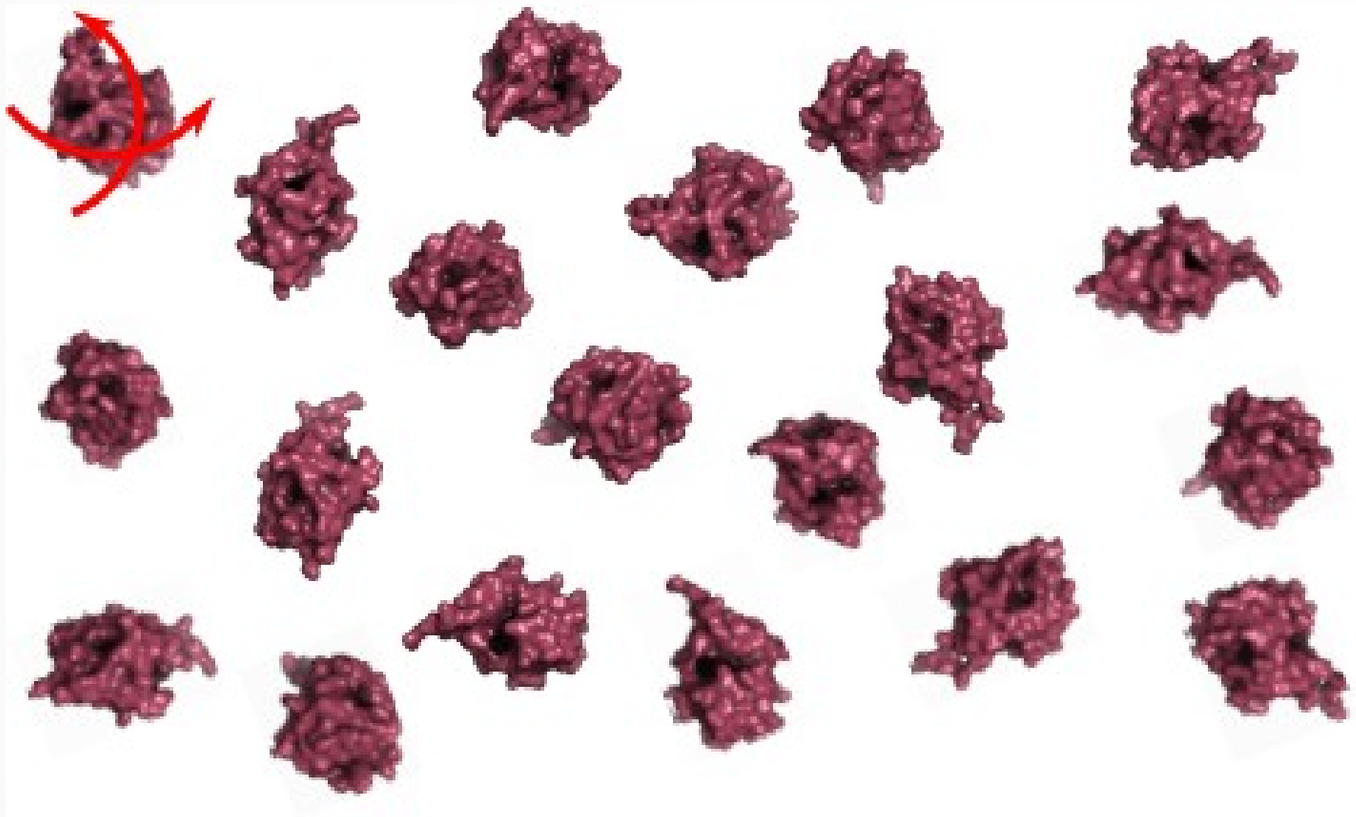
Can be predicted from the back-bone chemical shifts

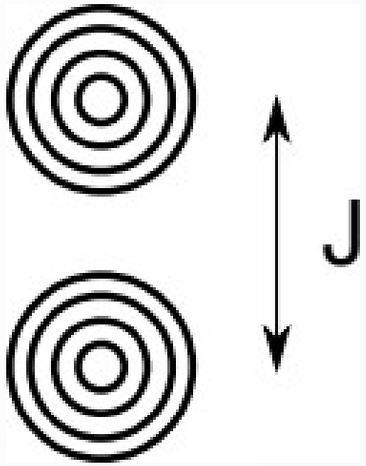


Torsion angle

Residual dipolar interaction

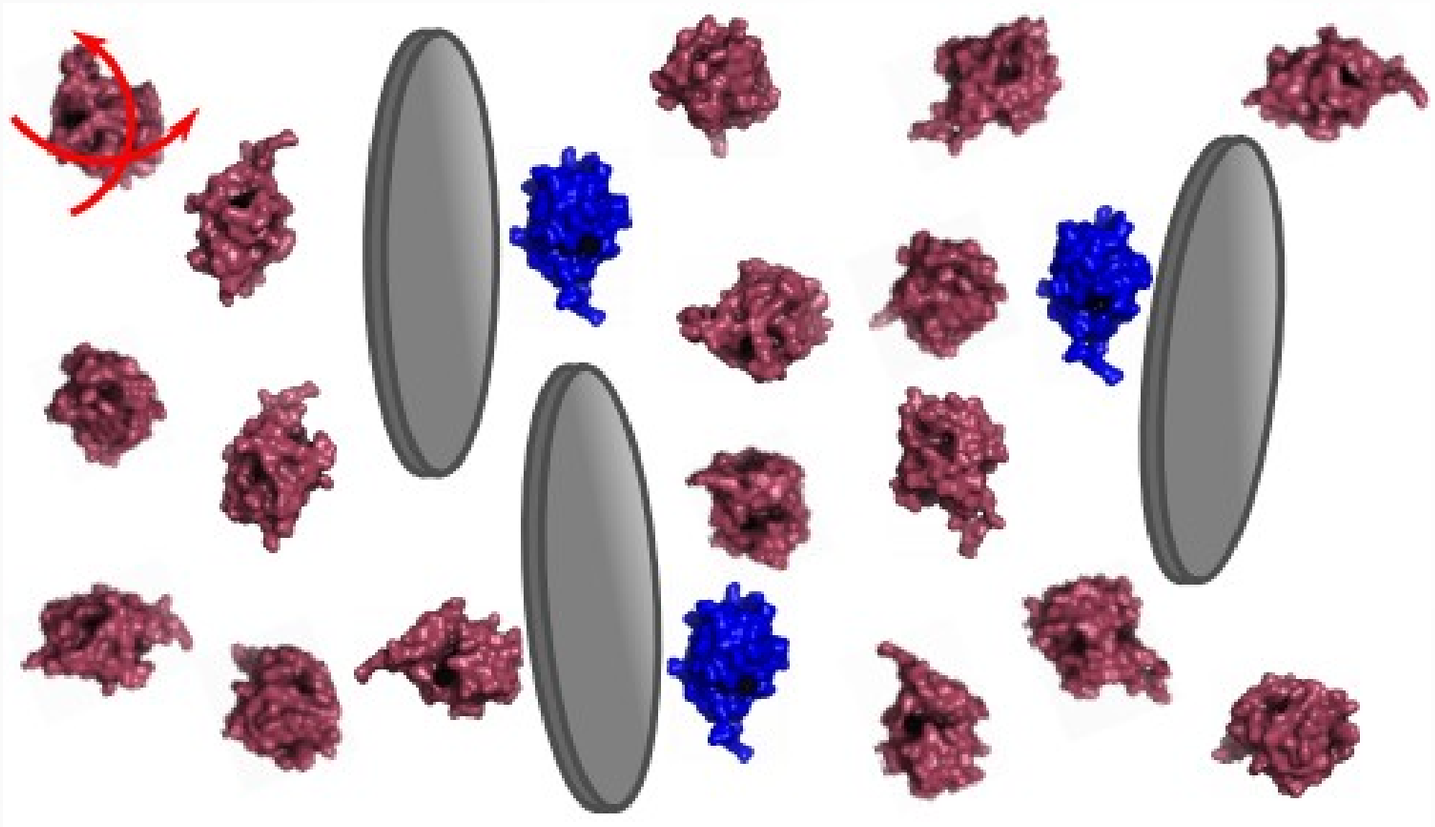
- Direct dipol—dipol interaction
- Undetectable in isotropic solution → averaged to 0

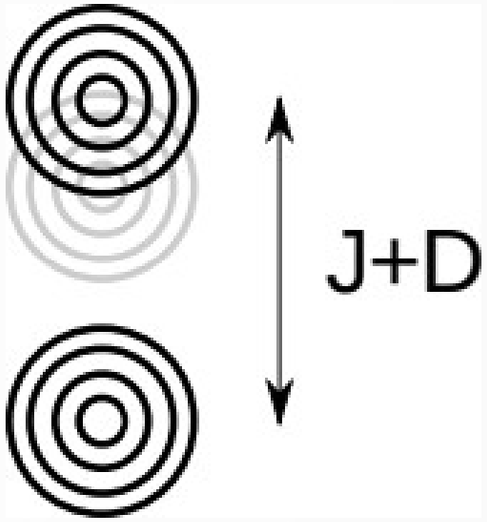




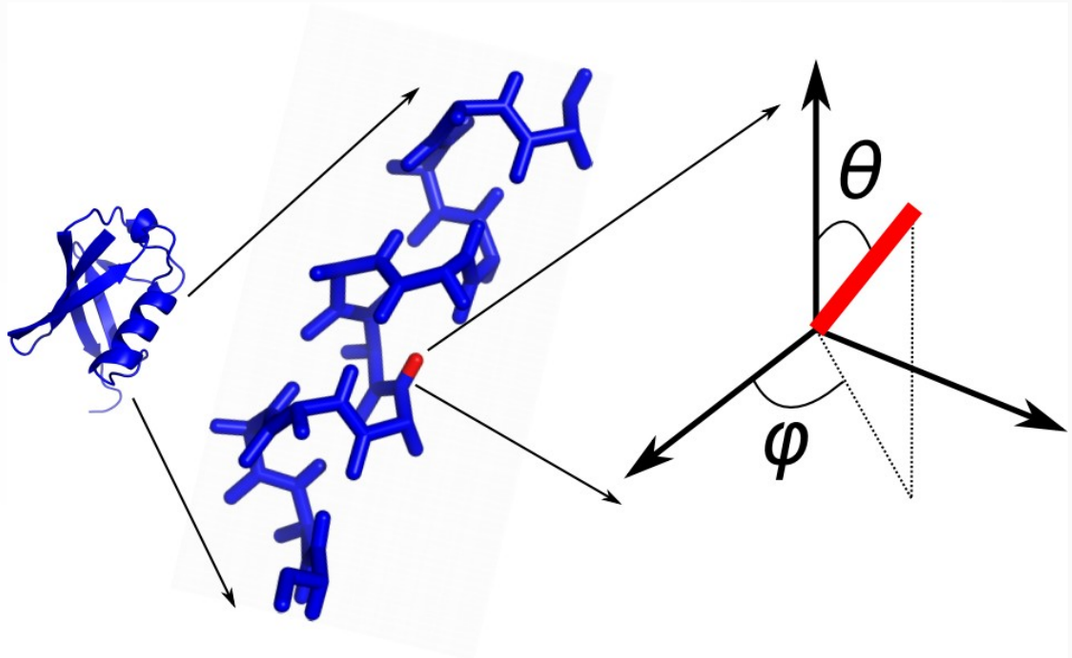
- Observable in anisotropic solution

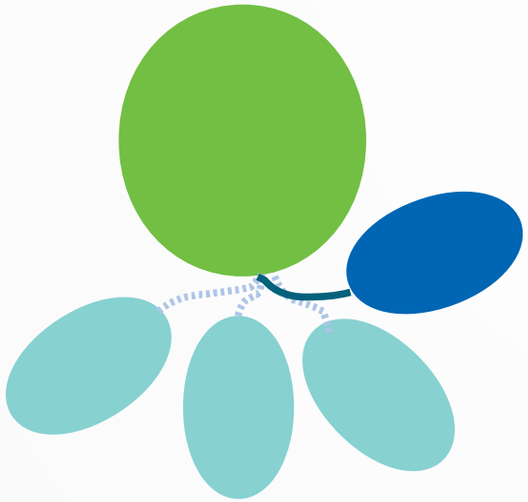
  liquid crystals – PEG, pressed PAG, virus particles (Bacteriophage)





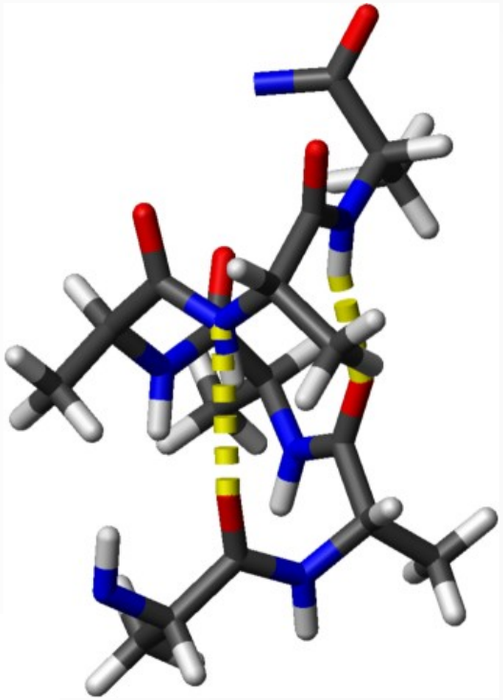
- Long-range interaction
- Size of *D* depends on the bond orientation with respect to the external magnetic field
- Orientation of domains, molecules in complexes

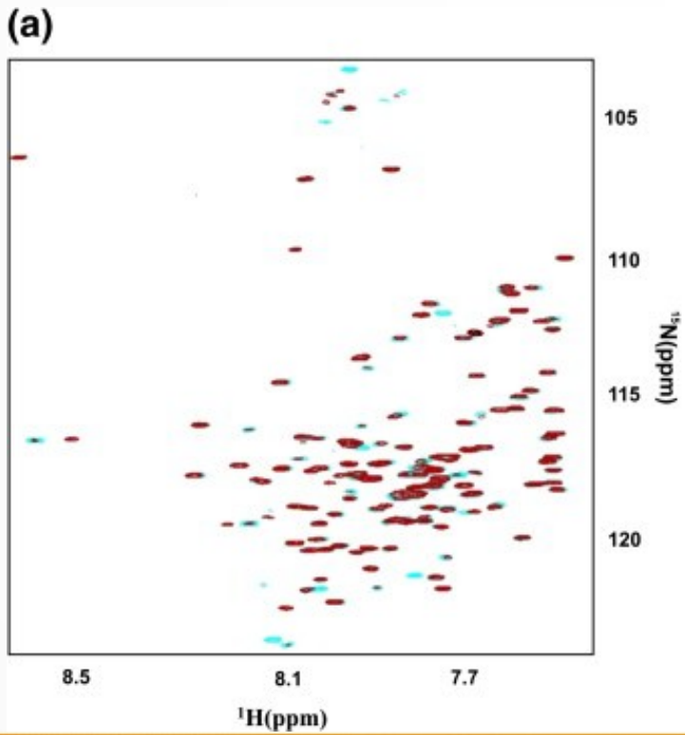


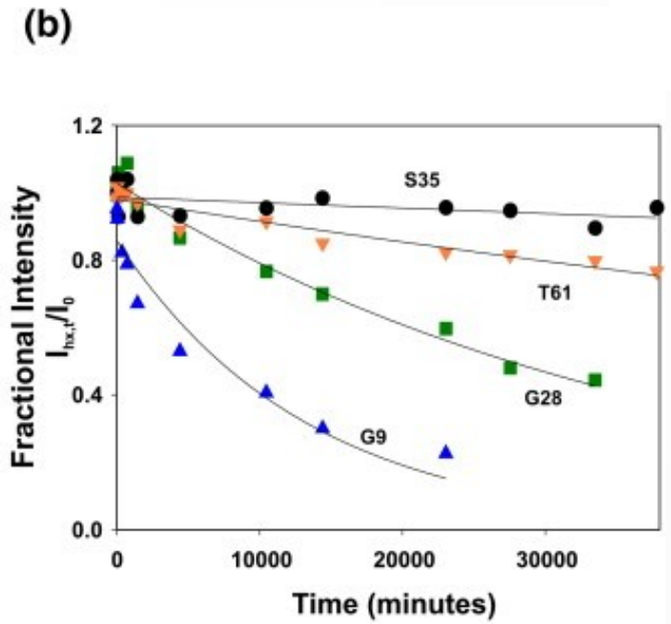


Hydrogen bonds

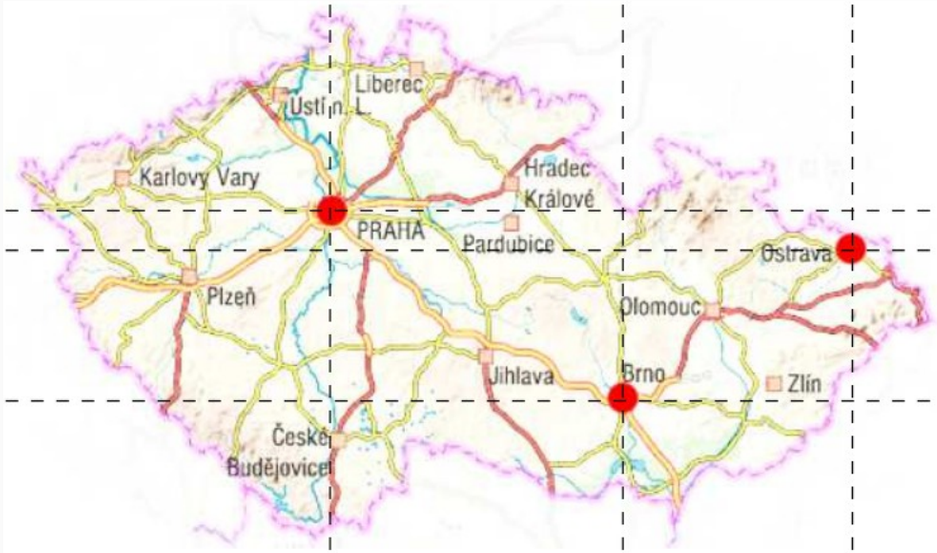
- H/D exchange measurement
- Predicted on 2° structure
- Scalar interaction trough hydrogen bond

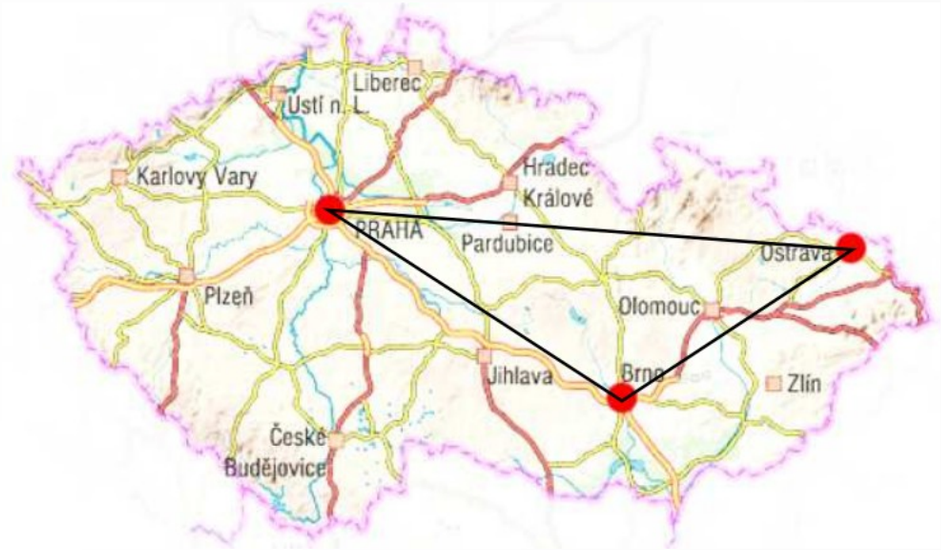






Structure calculation
X-ray vs NMR





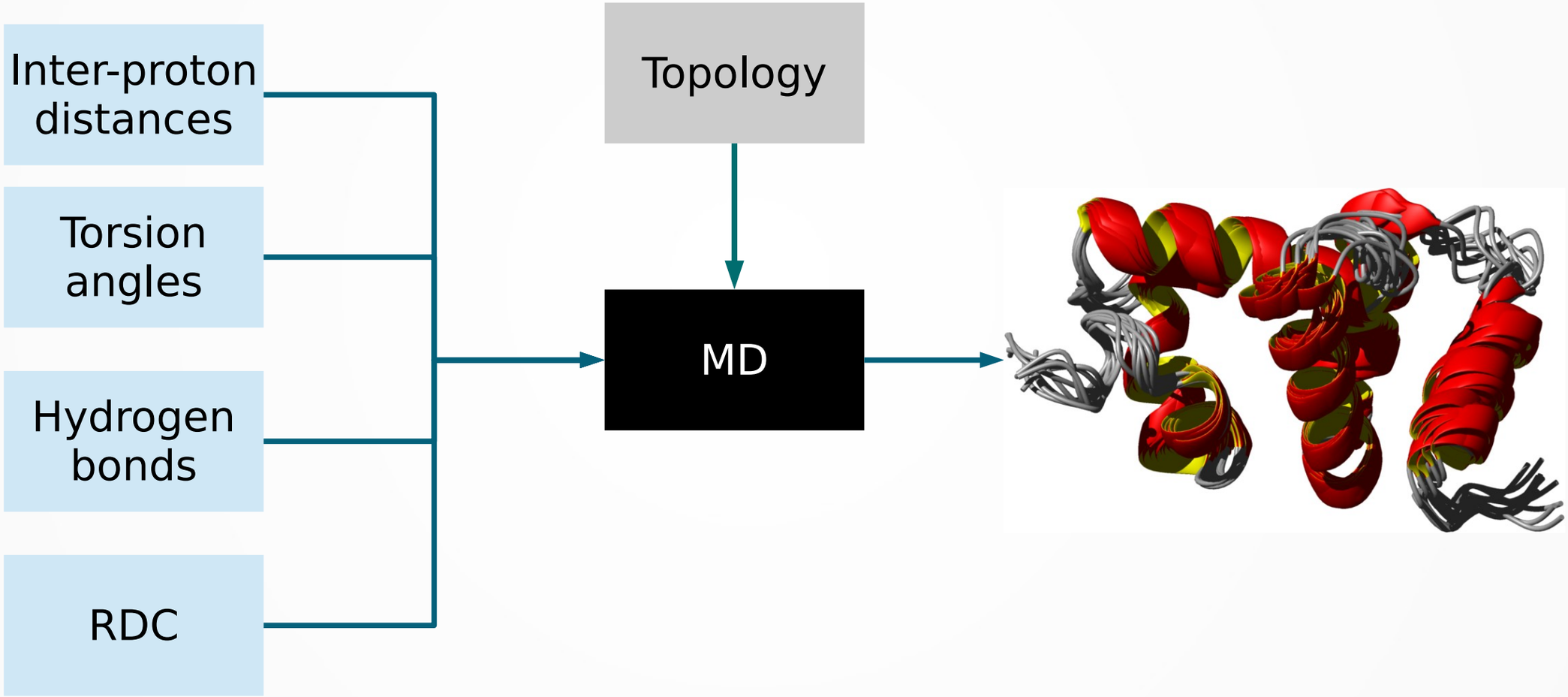
X-ray: GPS coordinates

Praha:	50°04’N	14°27’E
Brno:	49°11’N	16°37’E
Ostrava:	49°48’N	18°15’E

NMR: distances, angles, orientation

distacne(Praha-Brno) ~ 200 km
angle(Praha-Brno-Ostrava) ~ 100°
orientation(Praha-Brno) ~ SE

Molecular dynamics

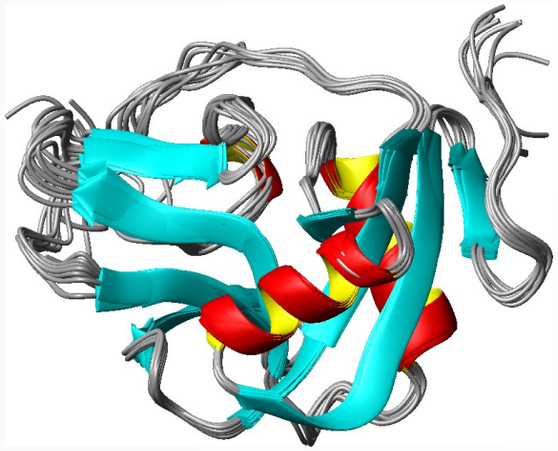


Calculated structures

- Set of similar structures with minimal $E_{potential}$
- Not only 1 structure (like in X-ray)

  → ensemble of structures which fulfill all experimental data
- Experimental data used as interval

  (*distance ± error)*



$RMSD_{back\text{-}bone}$=0.912 Å

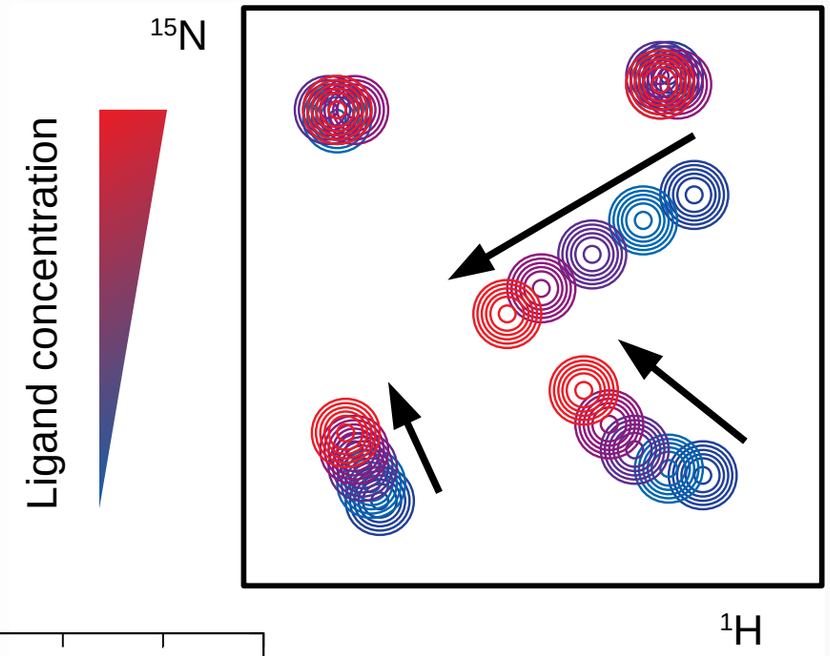
- root-mean-square deviation (RMSD)

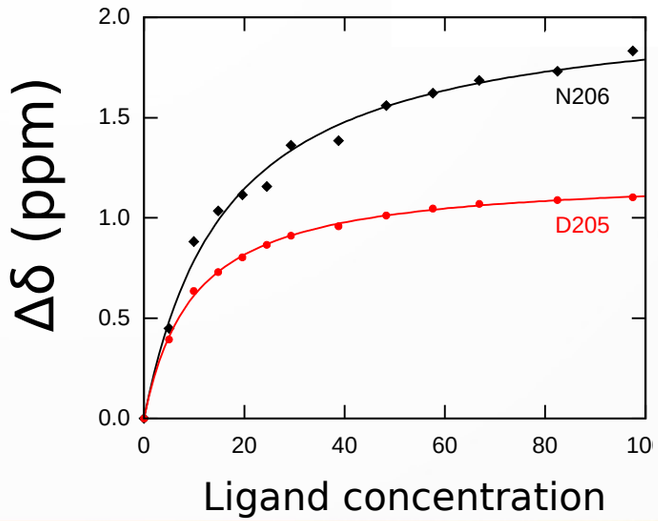
$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N} distance\left(C^{\alpha}\right)^2}$$

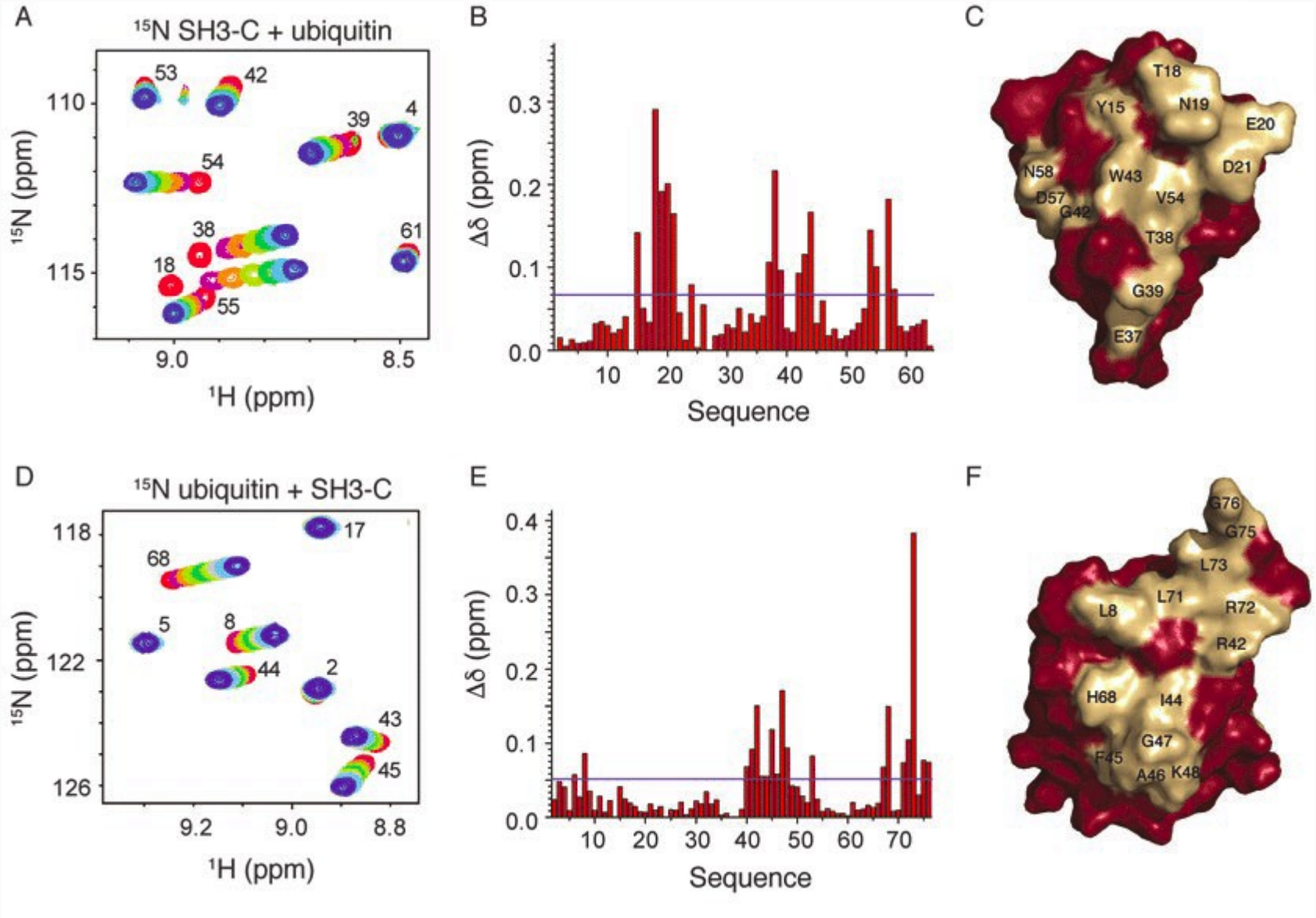
Protein—ligand interaction

- protein—protein, protein—nucleic acid, protein—small molecule, oligomeration
  - Identification of binding/interaction site
  - Determination of dissociation constant

- Titration experiment:
  - Small additions of ligand
  - Chemical shifts are very sensitive → changes in NMR spectra
  - Dissociation constant
  - Experimental data for docking





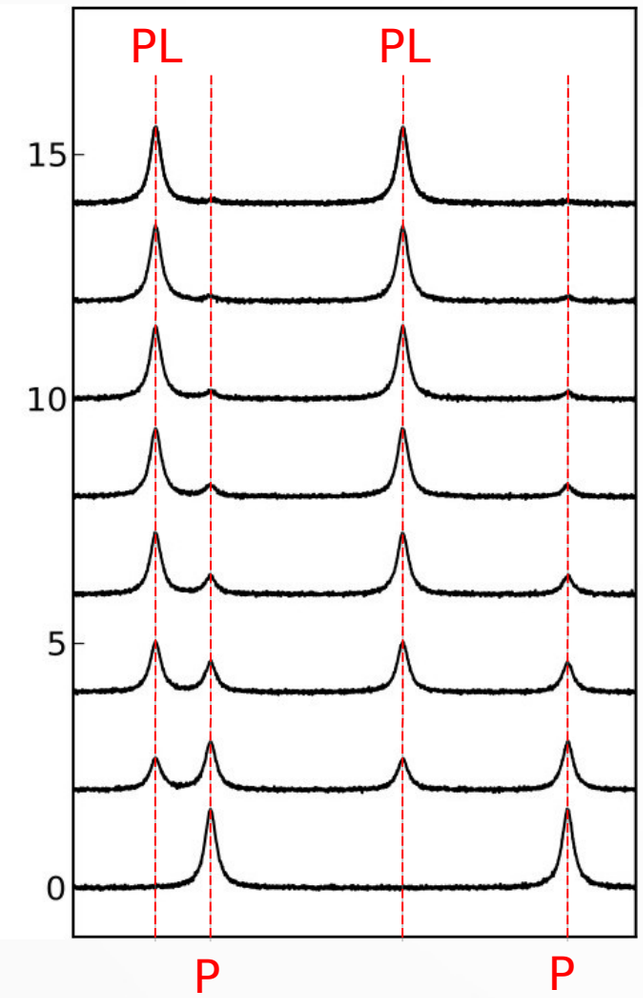


→ experimental data for docking calculations

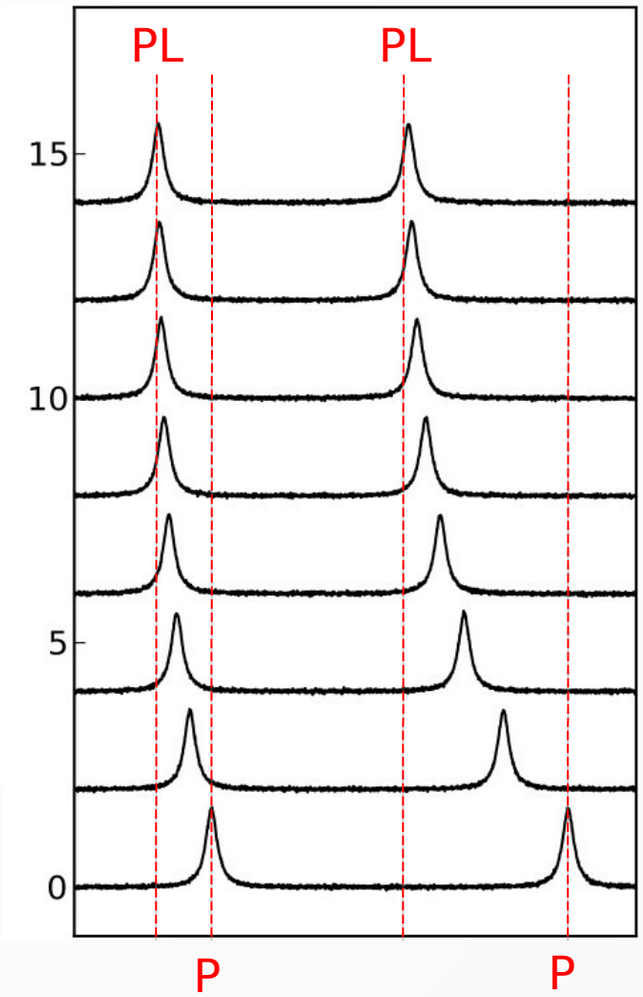
Joseph A. Marsh (ed.), Protein Complex Assembly: Methods and Protocols, Methods in Molecular Biology, vol. 1764

Slow exchange = strong interaction

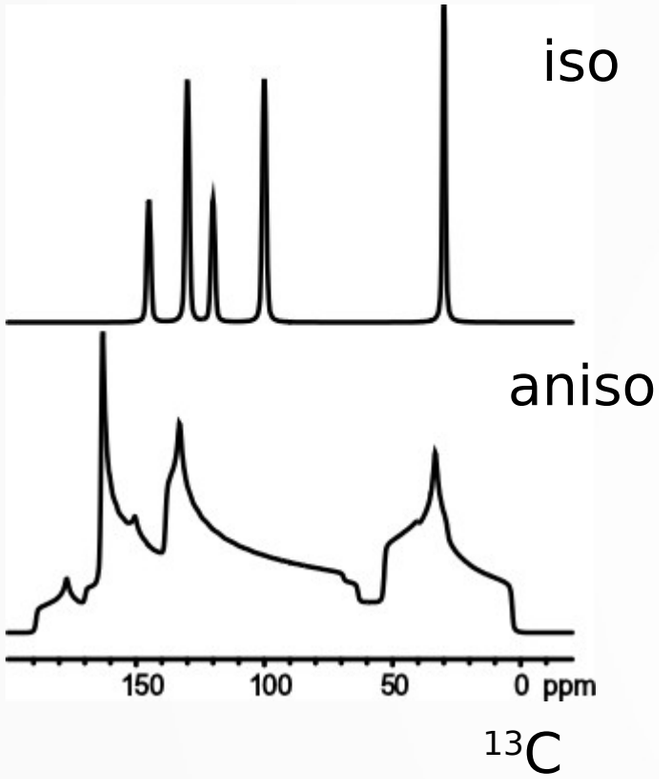
Fast exchange = weak interaction

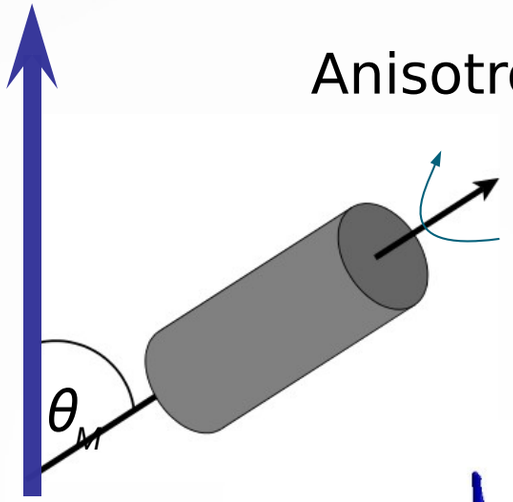


$$P + L \rightleftharpoons PL$$



Solid-state NMR (Magic Angle Spinning NMR)

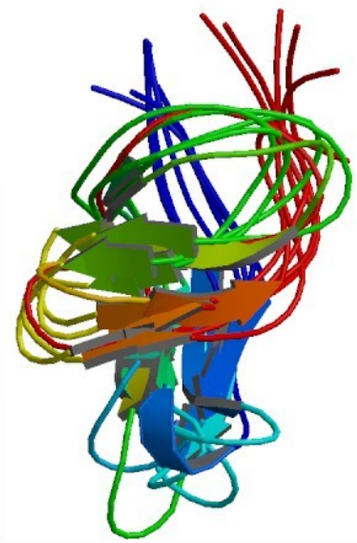




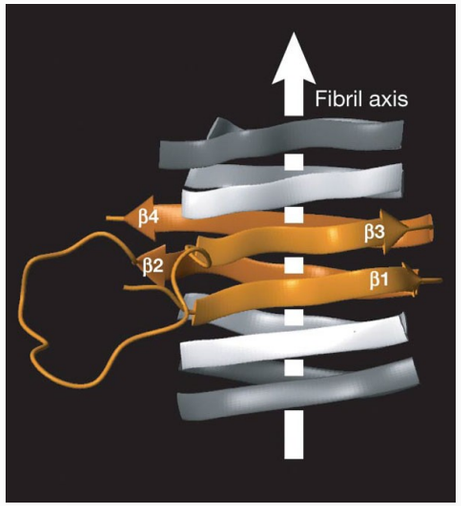
Anisotropic interactions = $3 \cos^2 \theta - 1 = 0$

$\theta_M = 54.74°$

Spinning: 1 – 110 kHz



2002 – SH3 domain of α-spectrin
1. ssNMR structure



amyloid fibril

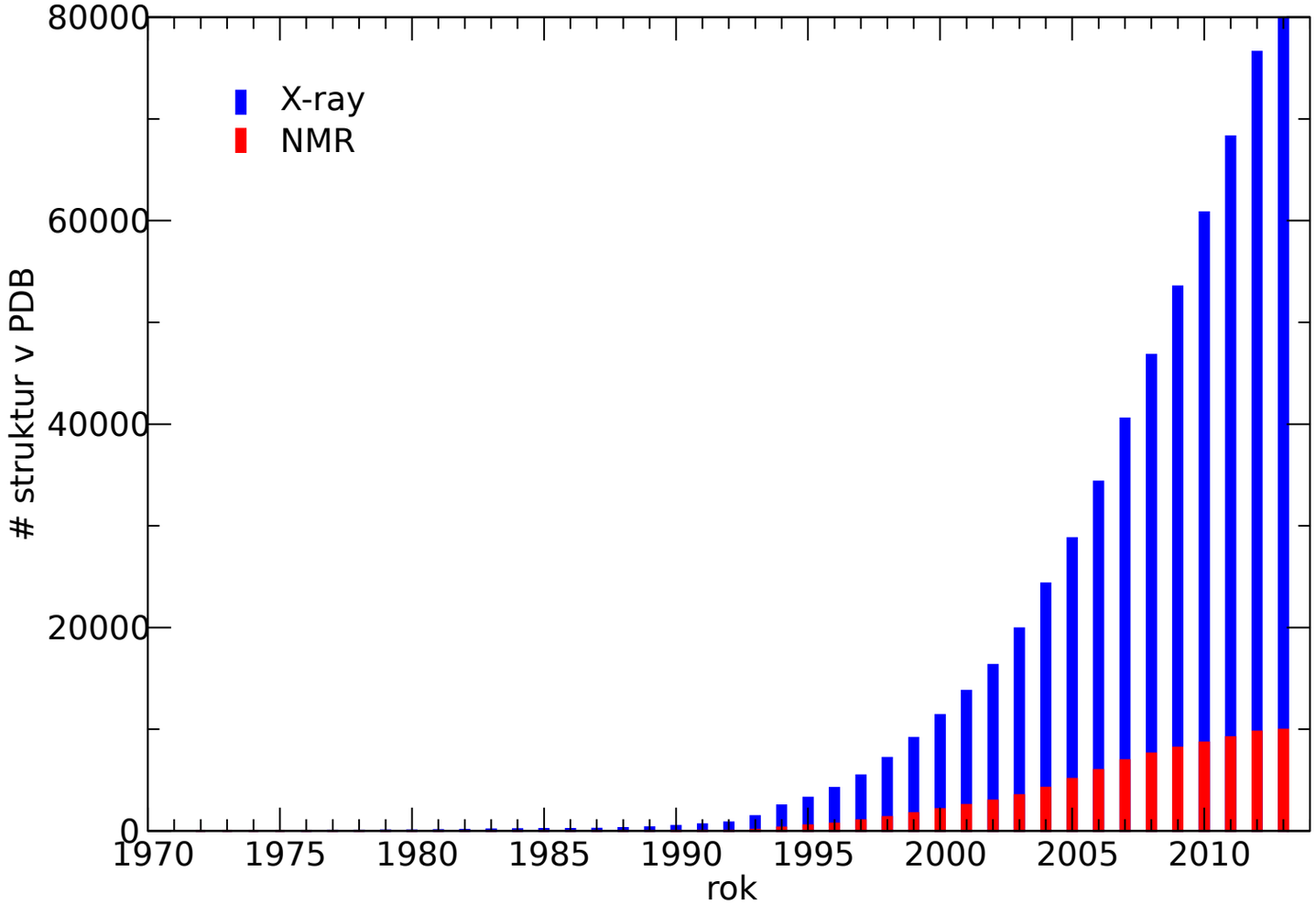
Sample:
- Microcrystalline form
- Selective labeling

Applications:
- Membrane proteins
- Insoluble proteins (prions,...)

NMR vs X-ray

	NMR	X-ray
sample	Liquid (solid) 😀	Crystal ☹️
Isotopic labeling	☹️	😀
Structure quality (resolution, correctness)	😐	😀
Protein size	☹️	😀
Unstructured proteins	😀	☹️
Dynamics	😀	😐
Time consumption	😐 ☹️	😐 😀

Structures in PDB

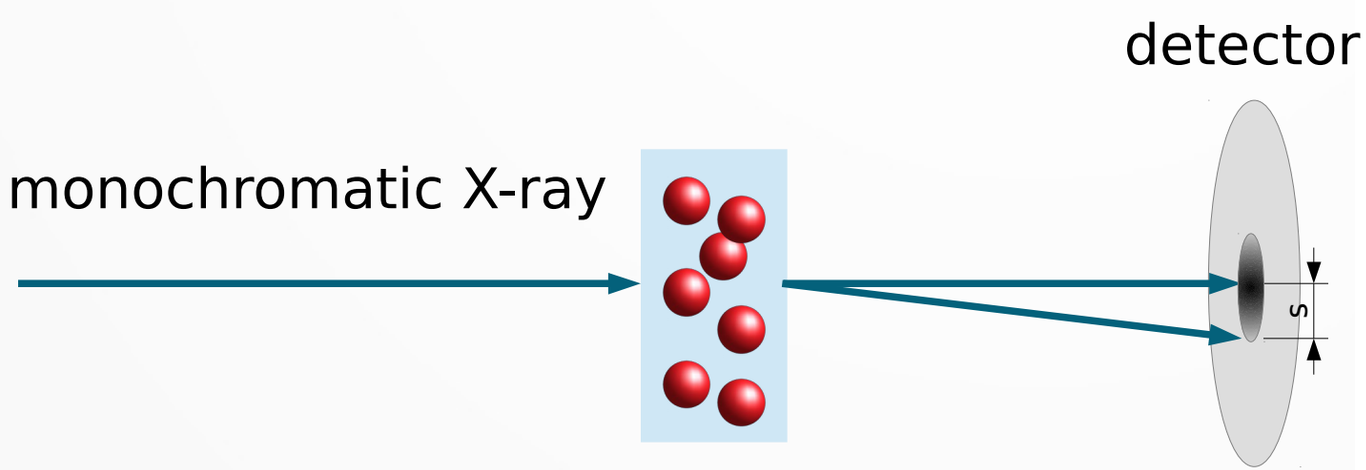


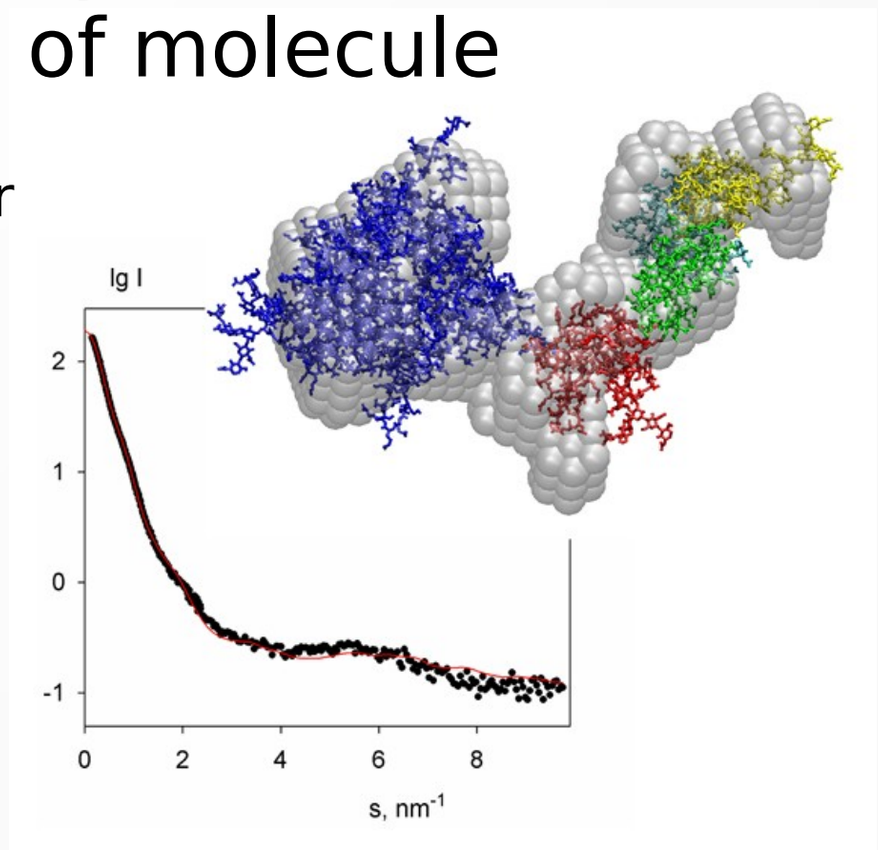
SAXS

Small-Angle X-ray Scattering

SAXS

- Scattering of monochromatic X-ray beam in protein *solution* → overall shape of molecule





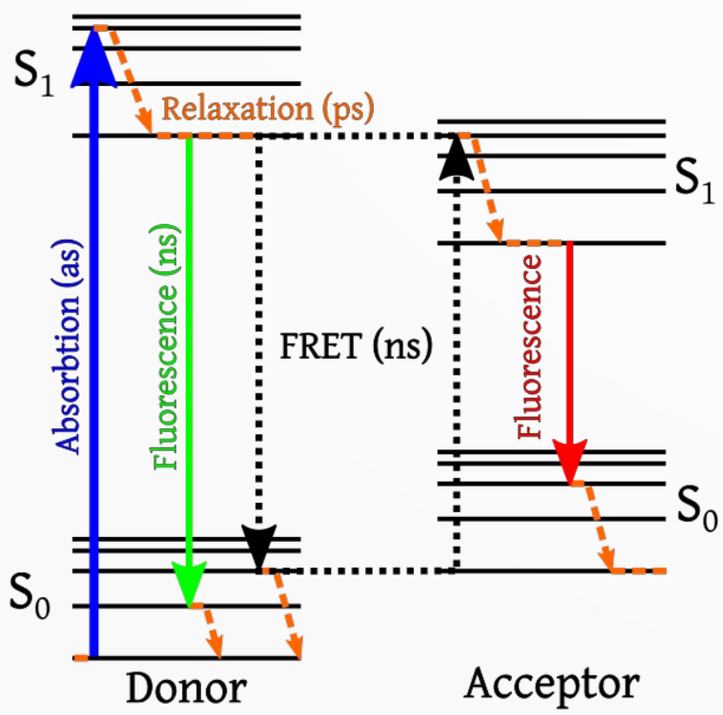
- 1D curve → 3D shape → more than 1 solution
- Validation – scattering curve can be easily back-calculated from protein 3D structure
- Applications:
  - Complex formation
  - Additional method for molecular modeling
- Variants:
  - Wide Angle X-ray Scattering (WAXS)
  - Small Angle Neutron Scattering (SANS)

FRET

Förster resonance energy transfer

FRET

- (Fluorescence resonance energy transfer)
- Energy transfer between two chromophores through nonradiative dipole—dipole interaction



$$E \sim \frac{1}{r_{DA}^{6}}$$

E – FRET efficiency (quantum yield of the energy-transfer transition)

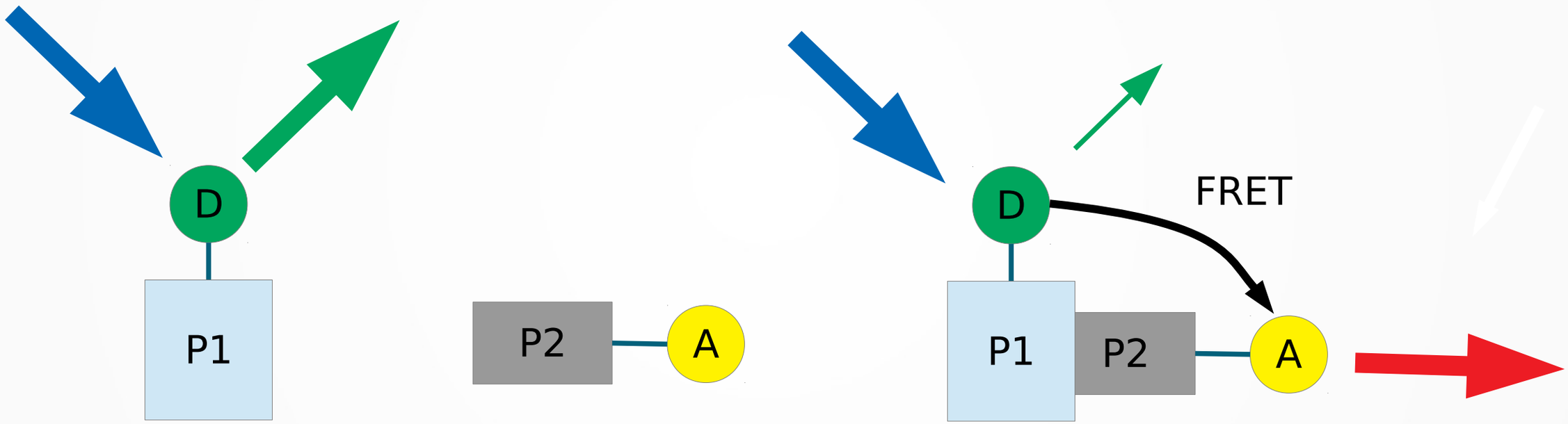
r_{DA} – donor—acceptor distance

Distances 10–100 Å

E depends on:

- distance (r_{AD}) between the donor and the acceptor
- spectral overlap of the donor emission spectrum and the acceptor absorption spectrum
- relative orientation of the donor emission dipole moment and the acceptor absorption dipole moment

- Application:
  - Protein conformation (intra-domain distances)
  - Complex formation/breakup



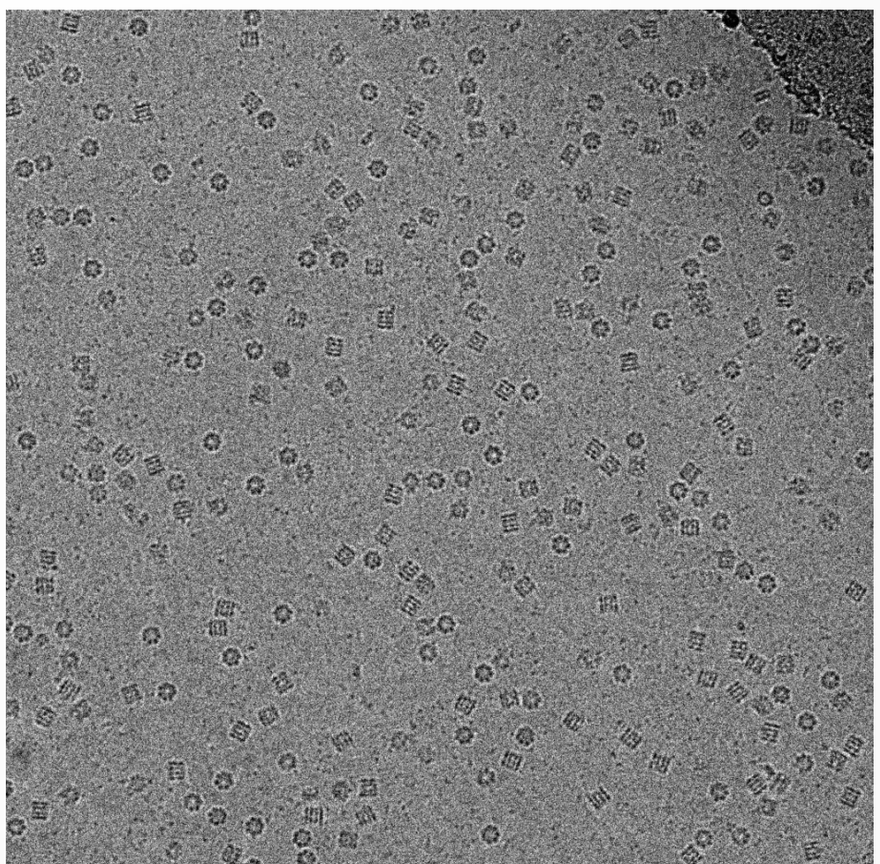
CryoEM

Cryo Electron Microscopy

CryoEM

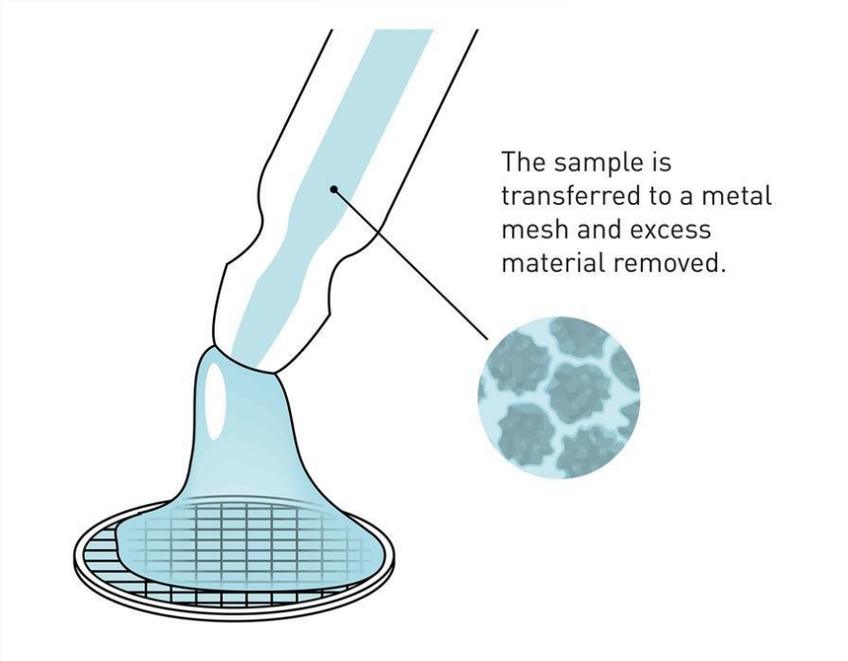
- Transmission electron microscopy
- Sample is frozen in amorphous ice
- Cryo-electron tomography

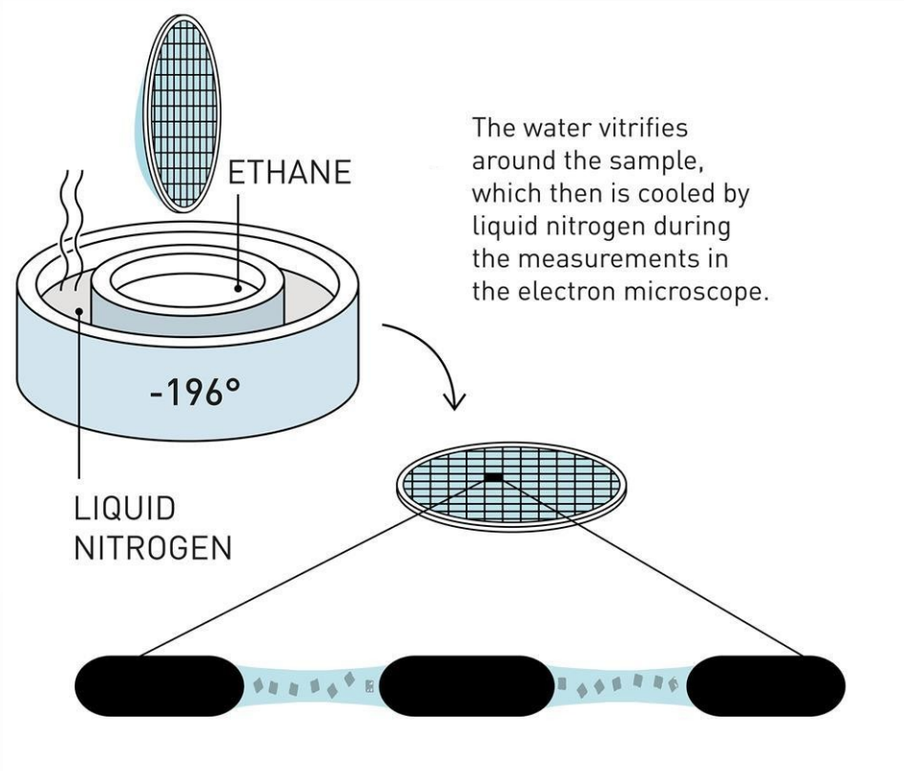
  → 3D reconstructuion
- 2017 – Nobel prize in Chemistry



GroEL

Sample preparation

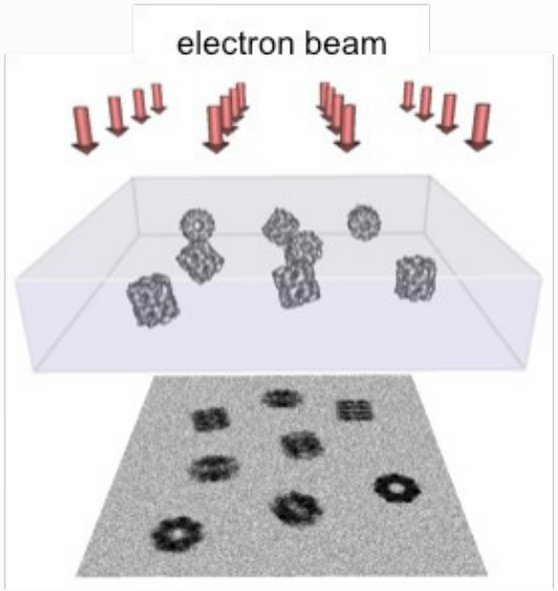


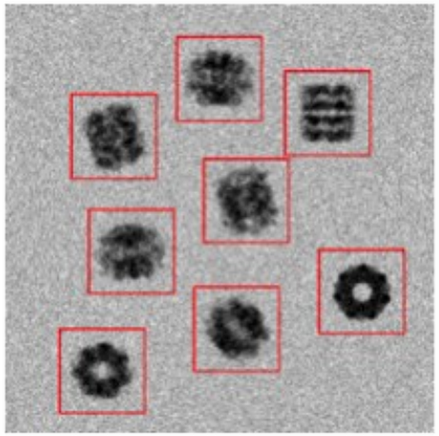


Vitrification

3D reconstruction

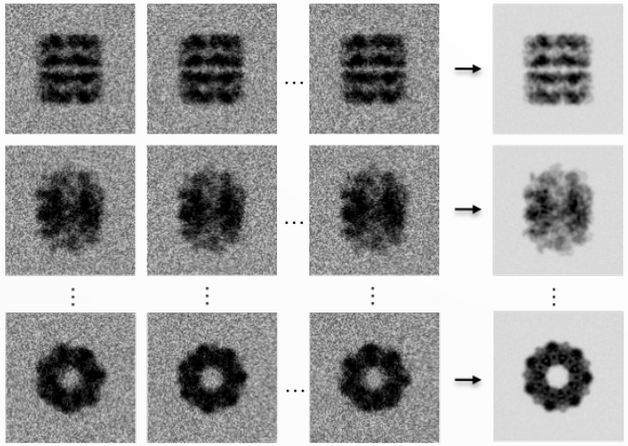
Data acquisition

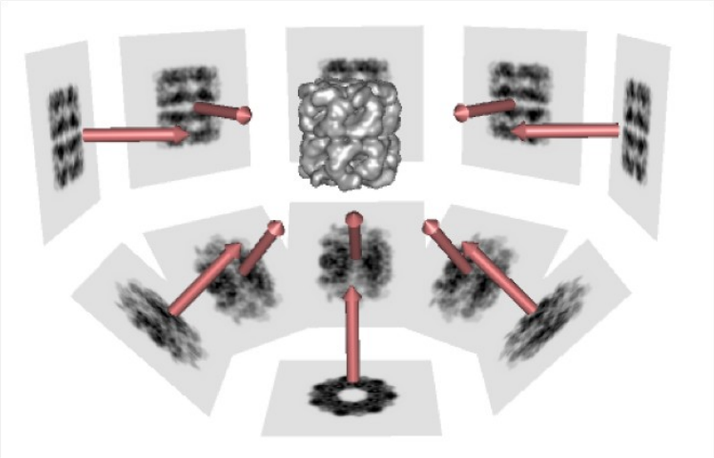




Boxing

Averaging
→ increasing S/N

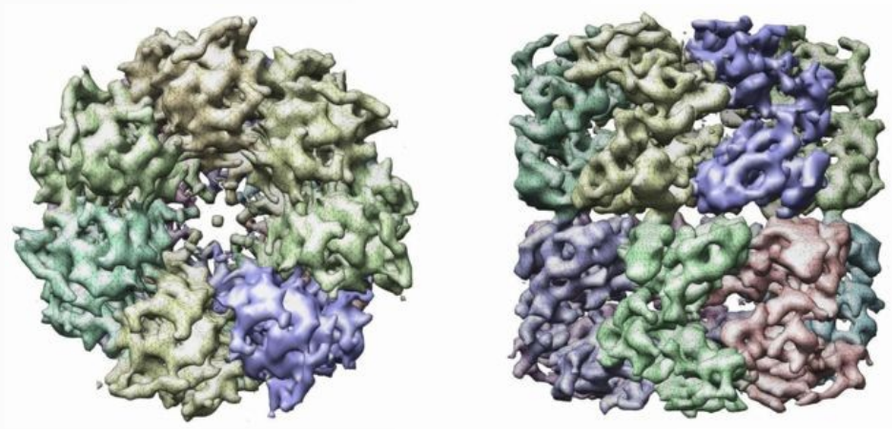




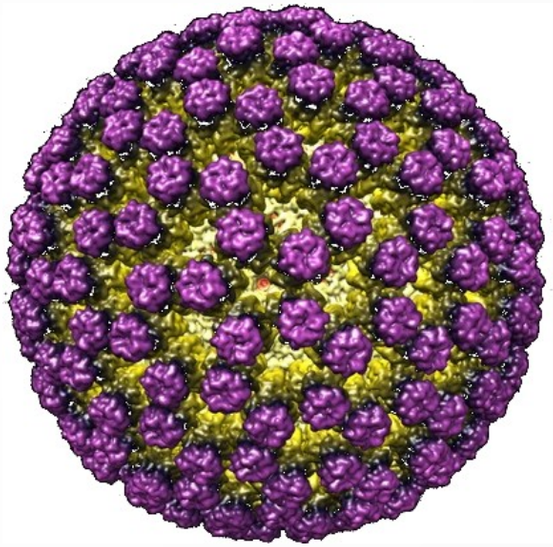
3D image reconstruction from different projections

Application

- structure of high molecular weight complexes

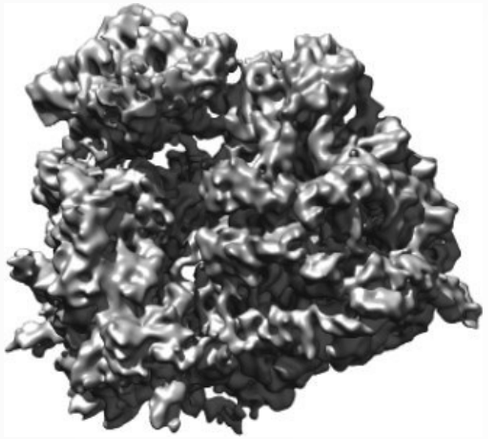


GroEL

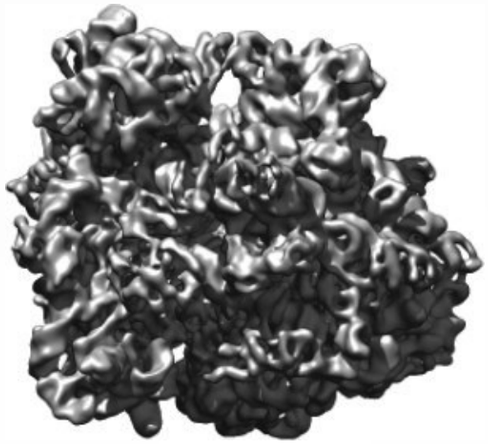


rotavirus

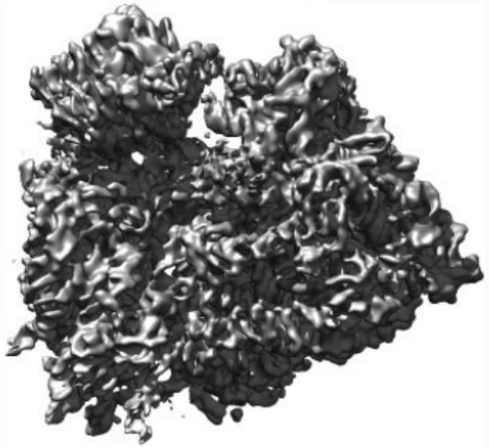
Ribosome CryoEM structure



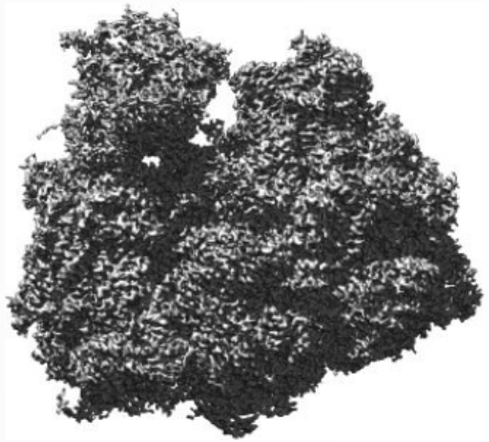
2006: 7.3 Å



2008: 6.7 Å



2010: 5.5 Å



2014: 3.2 Å

2016: 2.5 Å (60S) a 3.9 Å (40S)

CryoEM facility in CEITEC