Introduction to NMR spectroscopy

Swiss Institute of Bioinformatics
I.Phan & J. Kopp

NMR: the background

- Complex technique. Requires knowledge in:
  - Mathematics
  - Physics
  - Chemistry
  - Biology
  - (Medicin)
- Involves a lot of computing

N.M.R.

- Nuclear Magnetic Resonance
  - spectroscopy
  - imaging
  - solid-state
  - ....and much more

NMR applications

dynamics

binding

structure

in-vivo

folding

kinetics

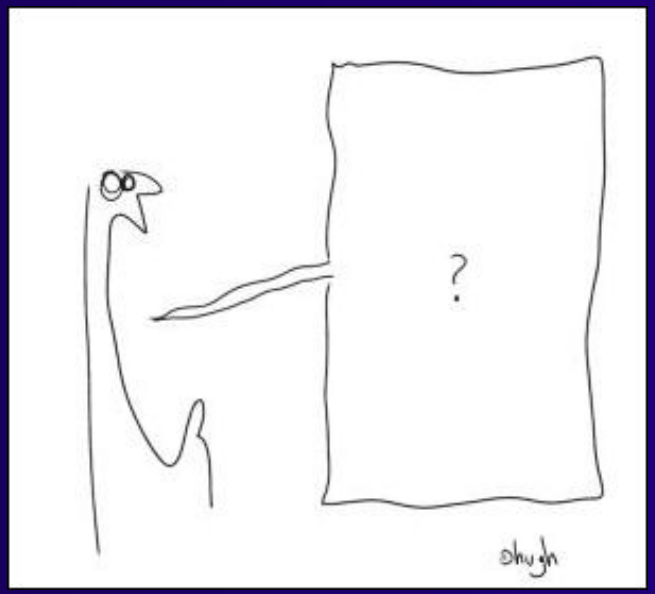
imaging

Can NMR be used for:

- Medical diagnosis?
- Drug-design?
- Computing?

NUCLEAR Magnetic Resonance

- NMR can detect atoms with a nuclear spin 1/2



Some nuclei with spin 1/2

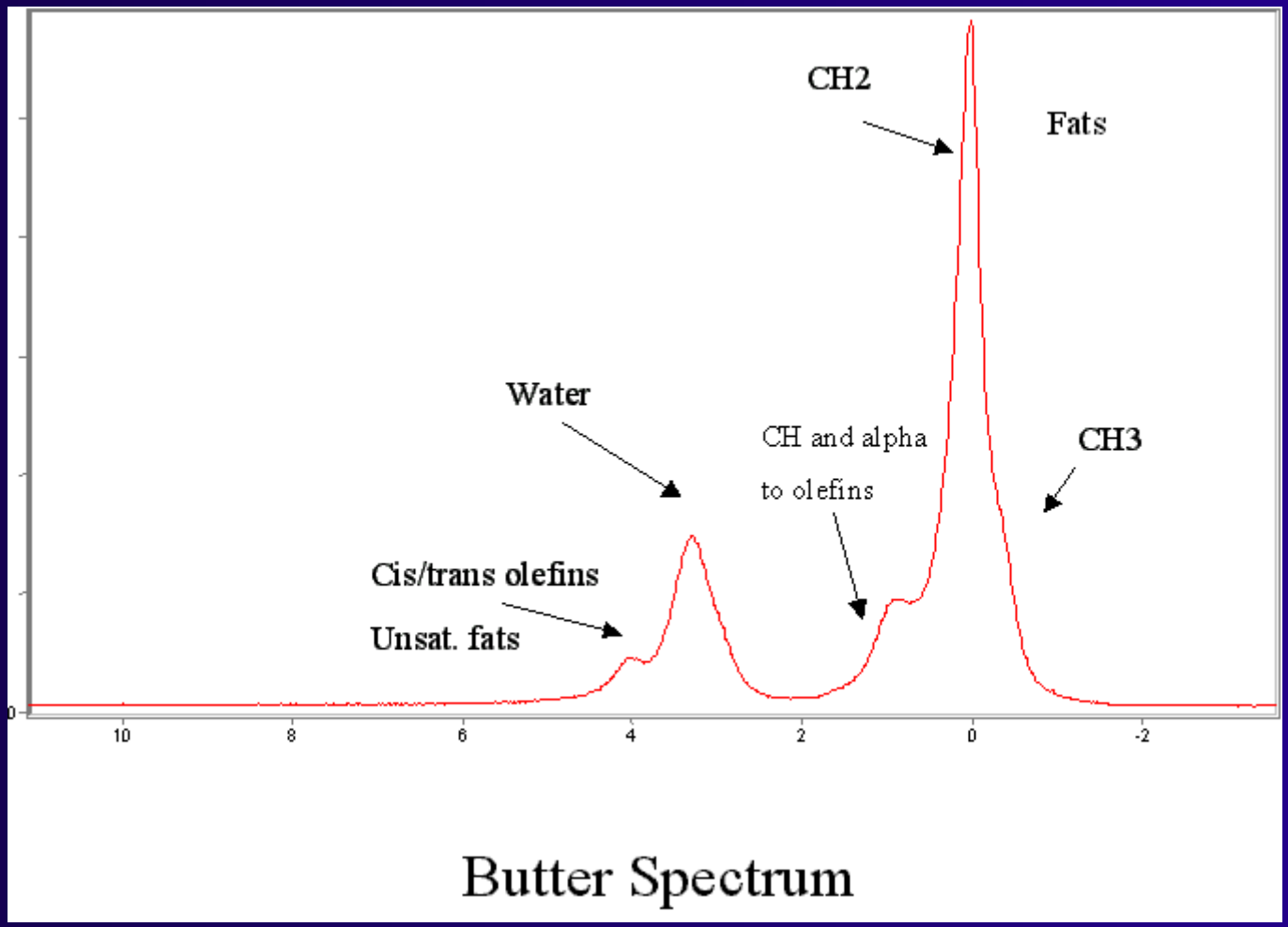
Isotope	Natural abundance (%)
^{1}H	99.98
^{13}C	1.11
^{15}N	0.37
^{19}F	100
^{31}P	100

^{12}C and ^{16}O do not have a spin
^{2}H, ^{14}N have spin 1

Is this good news for solving the structure of proteins by NMR?

NUCLEAR Magnetic Resonance

- Can NMR be used to detect:
  - Ozone?
  - Salt?
  - Butter?
  - A brain tumor?



Butter Spectrum

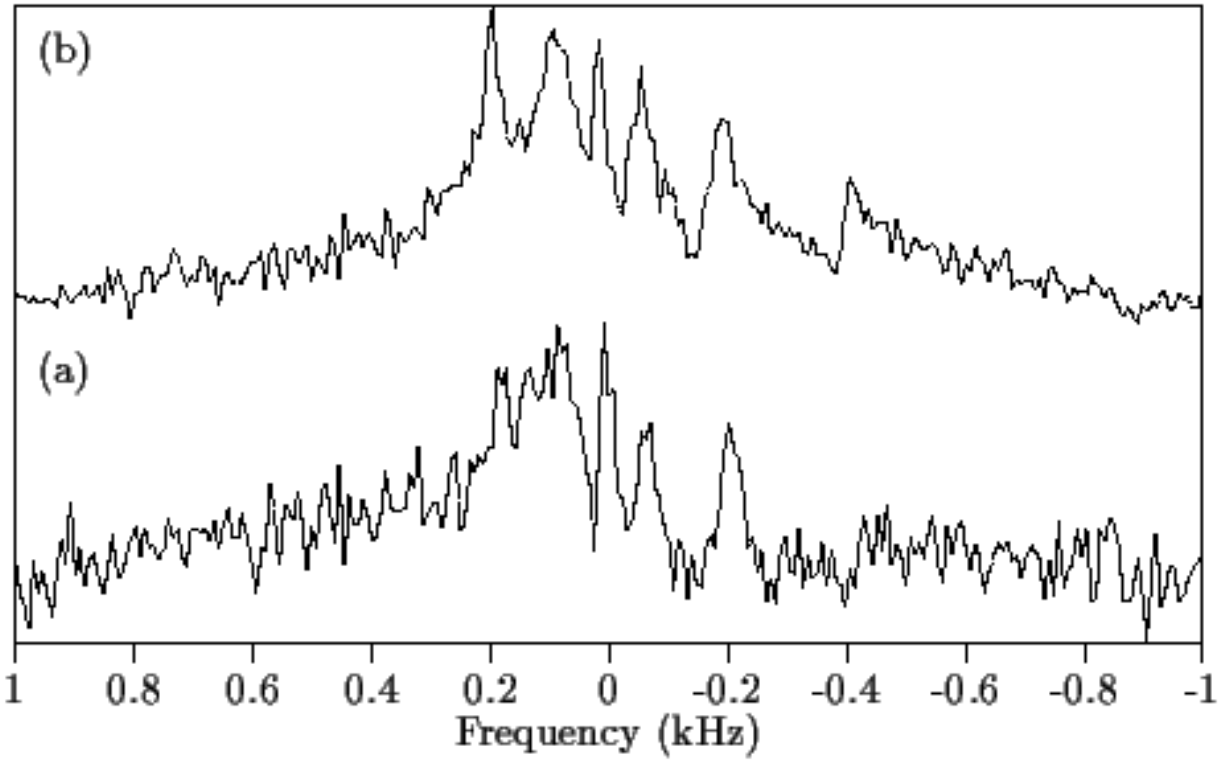
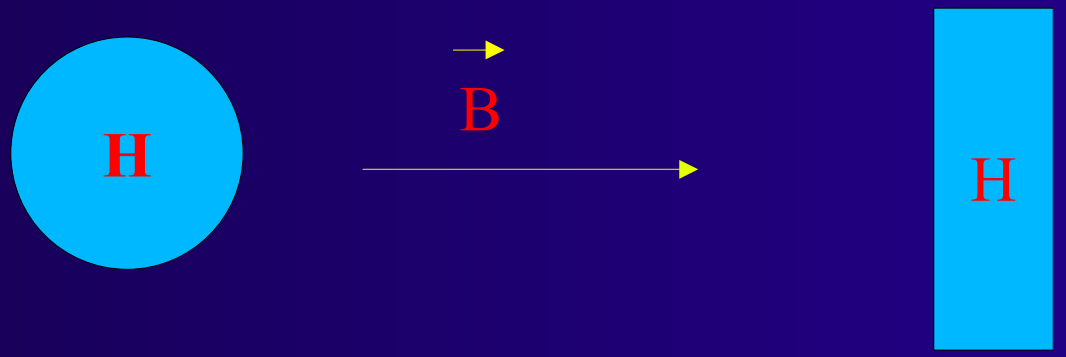


Figure 1.1: *In vivo* ^{31}P *NMR spectra of the human brain. (a) Healthy volunteer. (b) Patient with a tumor. Notice the much larger peak at 0.2 kHz.*

Nuclear MAGNETIC Resonance

- Nuclei with spin ½ behave like a magnet
- Placed in a constant magnetic field, they will align with that field



How do we create a constant magnetic field?

NMR instrumentation

- The spectrometer

sample tube goes in here

super-conducting magnet

NOTE:
the protein
is in solution
[1mM in water]

RF field generator

NMR is not sensitive

- Signals from the nuclei are measured in parts per million [ppm] of the static field strength
- NMR experiments require:
  - concentrated samples
  - strong fields = (very) big magnets

Biochemistry. 2001 Dec 25;40(51):15520-7.

High-resolution NMR structure of the chemically-synthesized melanocortin receptor binding domain AGRP(87-132) of the agouti-related protein.

McNulty JC, Thompson DA, Bolin KA, Wilken J, Barsh GS, Millhauser GL.

Department of Chemistry and Biochemistry, University of California, Santa Cruz, California 95064, USA.

The agouti-related protein (AGRP) is an endogenous antagonist of the melanocortin receptors MC3R and MC4R found in the hypothalamus and exhibits potent orexigenic (appetite-stimulating) activity. The cysteine-rich C-terminal domain of this protein, corresponding to AGRP(87-132), contains five disulfide bonds and exhibits receptor binding affinity and antagonism equivalent to that of the full-length protein. The three-dimensional structure of this domain has been determined by 1H NMR at 800 MHz. The first 34 residues of AGRP(87-132) are well-ordered and contain a three-stranded antiparallel beta sheet, where the last two strands form a beta hairpin. The relative spatial positioning of the disulfide cross-links demonstrates that the ordered region of AGRP(87-132) adopts the inhibitor cystine knot (ICK) fold previously identified for numerous invertebrate toxins. Interestingly, this may be the first example of a mammalian protein assigned to the ICK superfamily. The hairpin's turn region presents a triplet of residues (Arg-Phe-Phe) known to be essential for melanocortin receptor binding. The structure also suggests that AGRP possesses an additional melanocortin-receptor contact region within a loop formed by the first 16 residues of its C-terminal domain. This specific region shows little sequence homology to the corresponding region of the agouti protein, which is an MC1R antagonist involved in pigmentation. Consideration of these sequence differences, along with recent experiments on mutant and chimeric melanocortin receptors, allows us to postulate that this loop in the first 16 residues of its C-terminal domain confers AGRP's distinct selectivity for MC3R and MC4R.

Protein structure file in the Protein Data Bank

```
REMARK 210 EXPERIMENTAL DETAILS
REMARK 210  EXPERIMENT TYPE                : NMR
REMARK 210  TEMPERATURE           (KELVIN) : 288.00
REMARK 210  PH                             : 5.00
REMARK 210  IONIC STRENGTH                 : 50MM PHOSPHATE
REMARK 210  PRESSURE                       : 1 ATM
REMARK 210  SAMPLE CONTENTS                : 1.9 MM AC-AGRP(87-132)
REMARK 210                                   (SAME AS 1.9 MM MARP OF
REMARK 210                                   1QU8)
REMARK 210
REMARK 210  NMR EXPERIMENTS CONDUCTED      : 2D-NOESY, DQF-COSY, E-
REMARK 210                                   COSY
REMARK 210  SPECTROMETER FIELD STRENGTH    : 800 MHZ
REMARK 210  SPECTROMETER MODEL             : UNITYPLUS
REMARK 210  SPECTROMETER MANUFACTURER      : VARIAN
REMARK 210
REMARK 210  STRUCTURE DETERMINATION.
REMARK 210   SOFTWARE USED                 : VNMR 5.2, MNMR 940501,
REMARK 210                                   XEASY 1.2, DYANA 1.5,
REMARK 210                                   PROCHECK 3.4.4
REMARK 210   METHOD USED                   : TORSION ANGLE DYNAMICS
REMARK 210
```

Structure data in the Protein Data Bank

* X-ray crystallography : over 80%
* NMR : about 16%

The cost of NMR

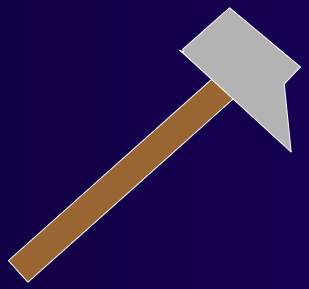
- Superconducting magnet : no resistance, no current loss
  - Requires cooling to almost absolute zero
    - Liquid helium
    - (+ Liquid nitrogen for insulation)
- Strong magnetic field
  - Requires special infrastructure
    - Minimize field disturbance
- NMR is a very expensive technique!

A typical NMR lab

radio frequency unit

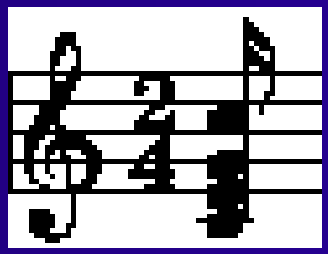
superconducting magnet

Nuclear Magnetic RESONANCE

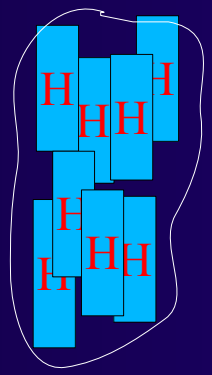


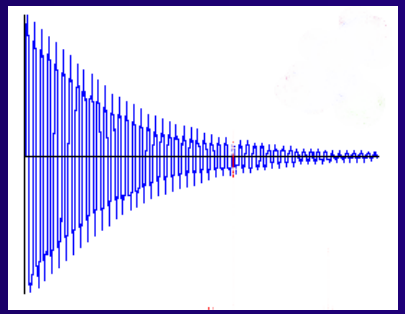
sound

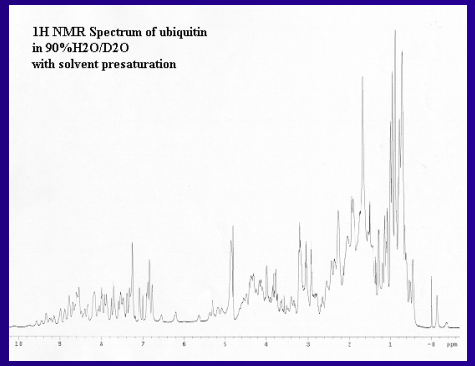
wave
to
frequency
transformation



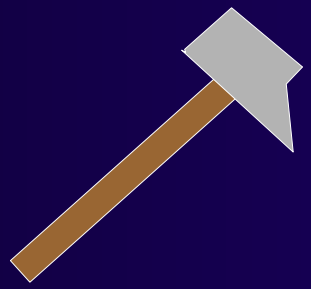
RF field

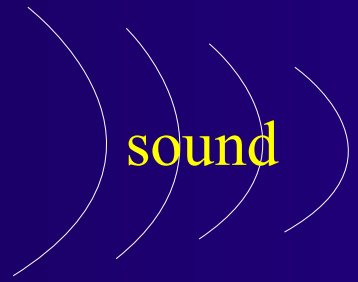




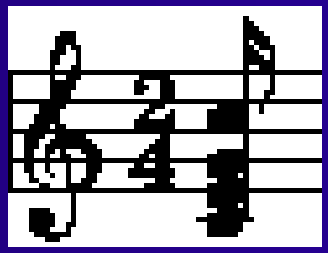


Nuclear Magnetic RESONANCE

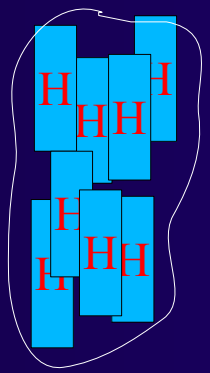


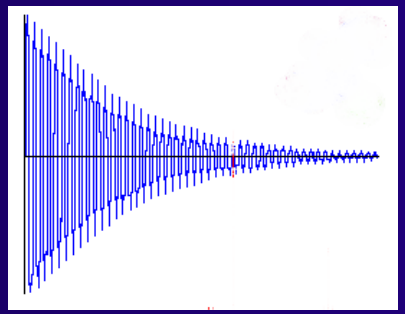


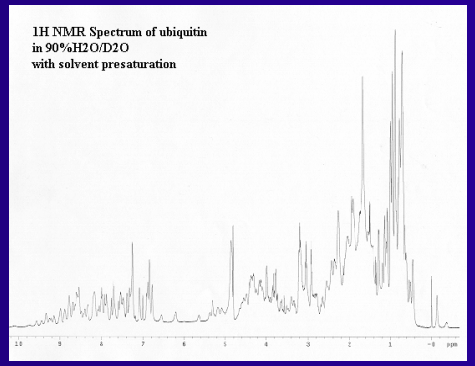
wave to frequency transformation



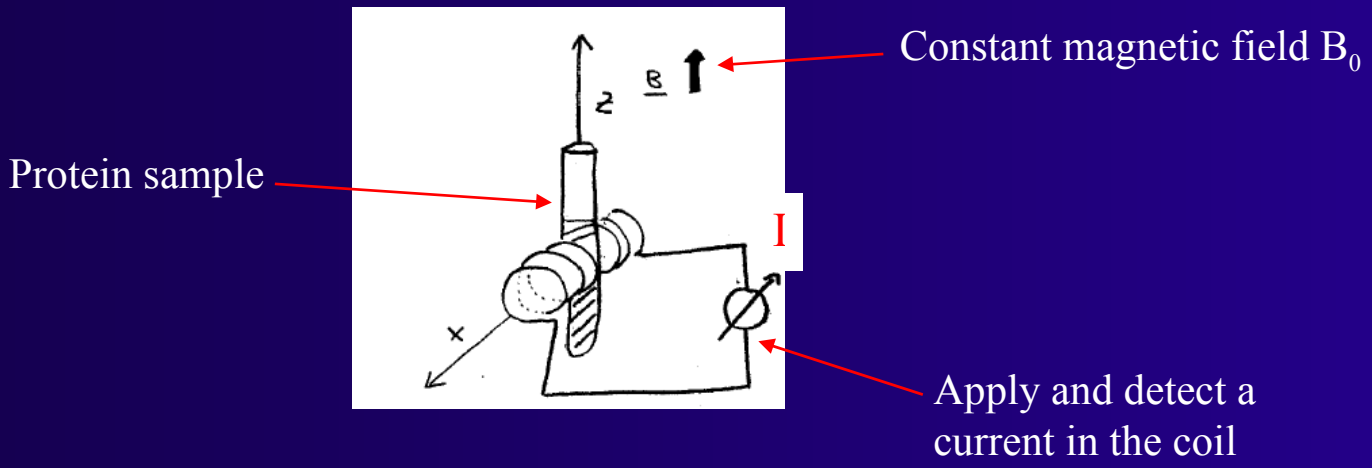
RF field

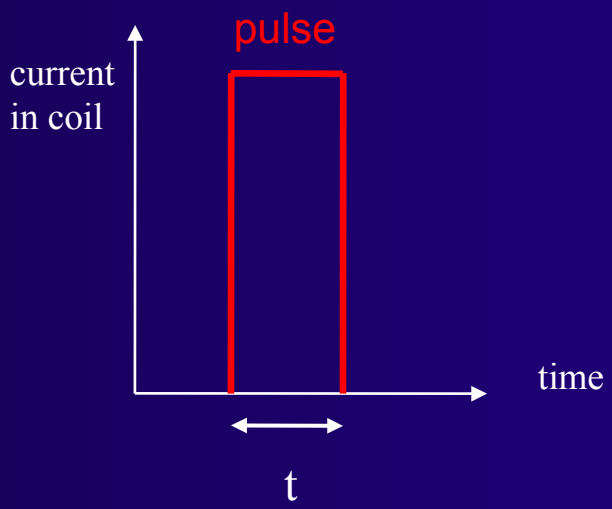




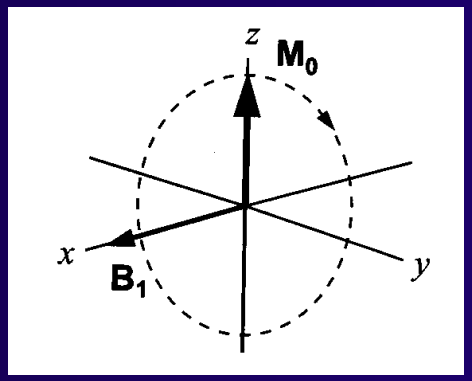


- In an NMR experiment, the energy input to make the nuclei resonate is produced by a Radio Frequency (RF) **PULSE**
- RF field: magnetic field B_1 perpendicular to the constant magnetic field B_0



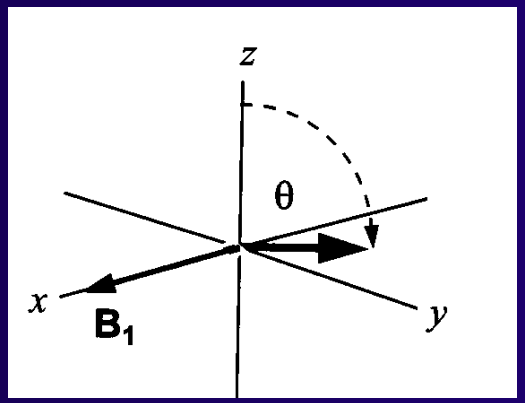


What is the effect of the pulse?



The **PULSE**
rotates the bulk magnetisation M_0
around x-axis

B_0



The angle of rotation
is proportional to the duration
of the RF pulse time t

NMR experiments

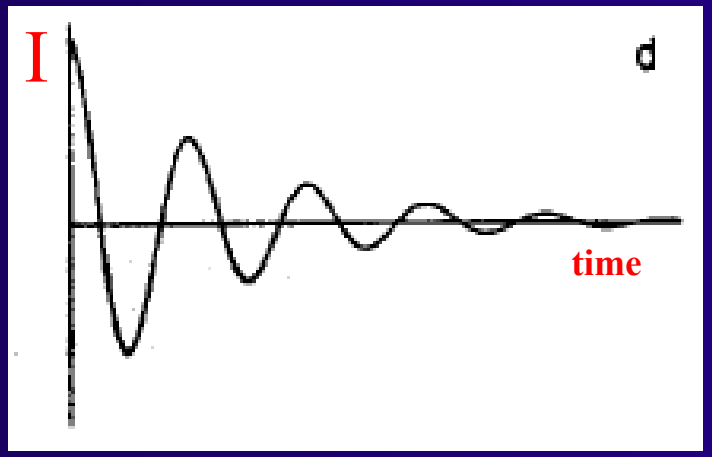
- Depending on the length of the pulses and delay between pulses, different effects are measured
  - Variety of NMR experiments: variety of spectra
- Depending on the frequency of the RF pulse, different nuclei can be detected
  - Proton 1H – NMR
  - 15N – NMR spectra
  - 13C – NMR spectra
  - ...etc...

How is an NMR signal detected?

- After the pulse, the nuclei return to their ground energy state
- The nuclei precess back to their start position
- Precessing induces a current that is detected by a coil in the NMR spectrometer
- As the nuclei return to equilibrium, the induced current decreases back to zero = Free Induction Decay (FID)

The NMR signal: a Free Induction Decay (FID)

The FID signal



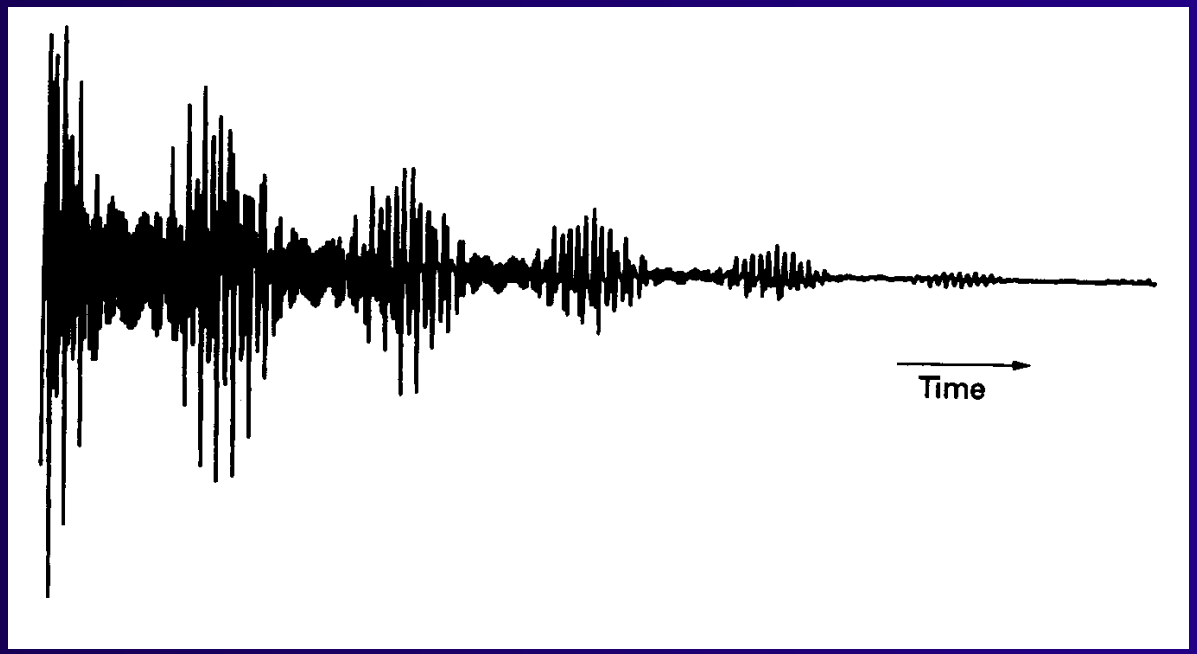
Current induced by one precessing spin decays after RF pulse

Nuclear Magnetic RESONANCE

- One spin: one bar magnet
- Many spins: bulk magnetisation
- Depending on the length of the RF pulse, the bulk magnetisation of an ensemble of spins will flip at a different angle with respect to the static field (B_0)
- After the pulse, each spin precesses individually and gives rise to an FID

?

FID for an ensemble of spins

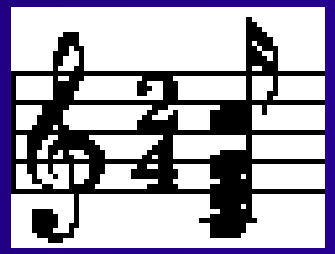


Current induced by precessing spins decays after RF pulse

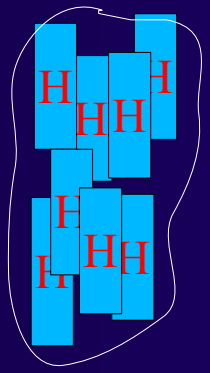
Nuclear Magnetic RESONANCE

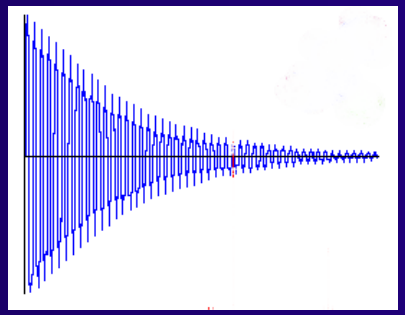
sound

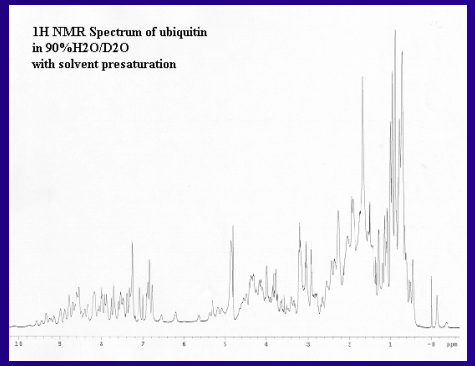
wave
to
frequency
transformation



RF field



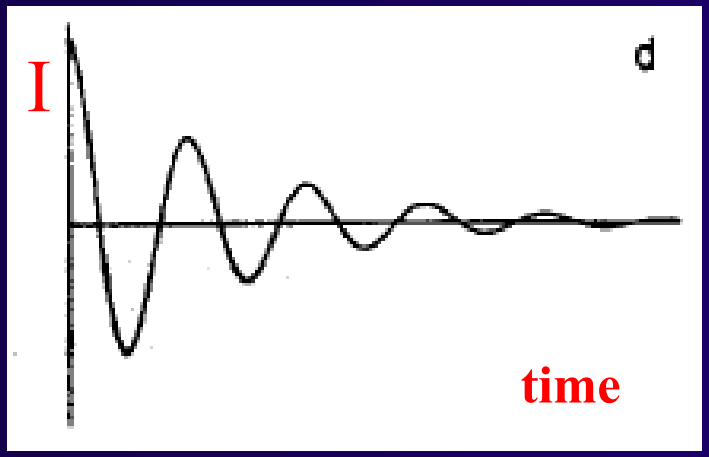




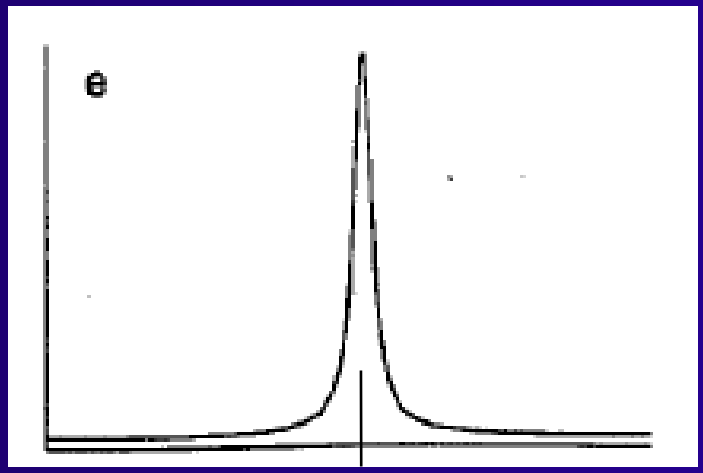
NMR signal processing

- NMR spectrum: a superposition of signals
- One signal: FID of one nucleus
- Interpretation is made easier by a simple mathematical formula that transforms of the FID from the time domain to the frequency domain
  - Fourier transformation

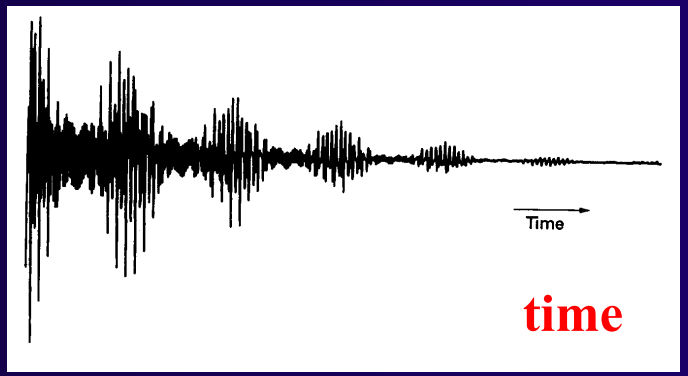
Fourier Transformation



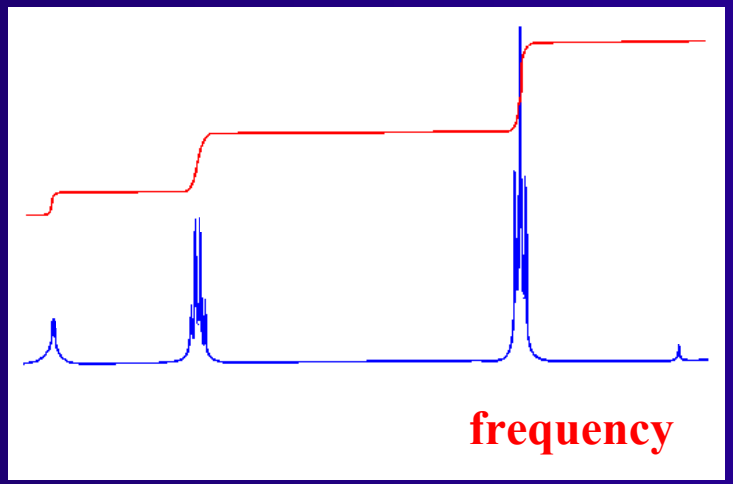
FT



resonance
frequency

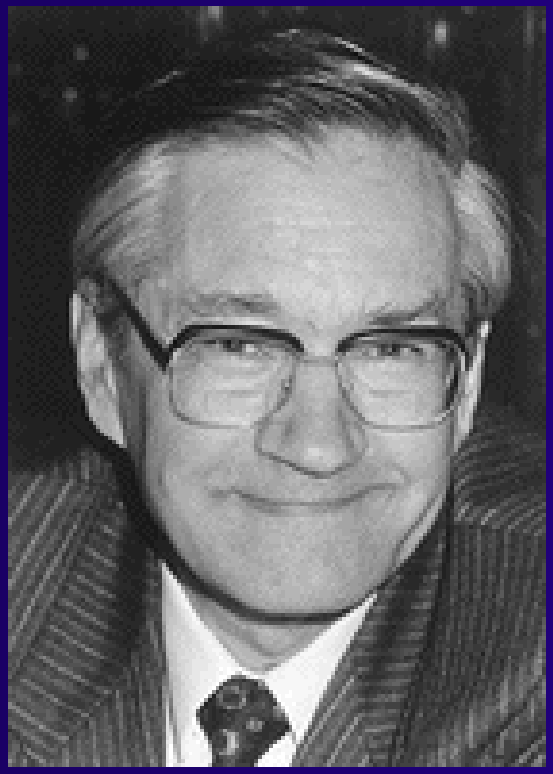


FT



Richard Ernst (ETHZ)

- Nobel 1991

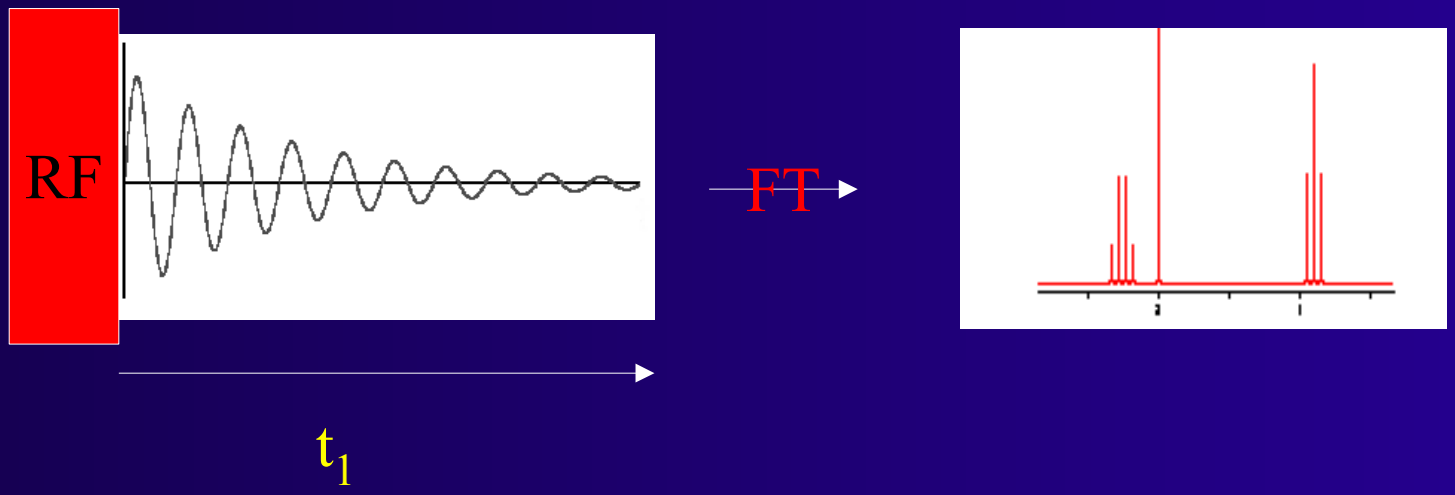


NMR structures

- How do we get a protein 3D structure with NMR?
  - NMR experiments
  - Data collection
  - Spectrum assignment
  - Structure calculation
  - Method assessment: is NMR really worth the effort?

NMR experiments

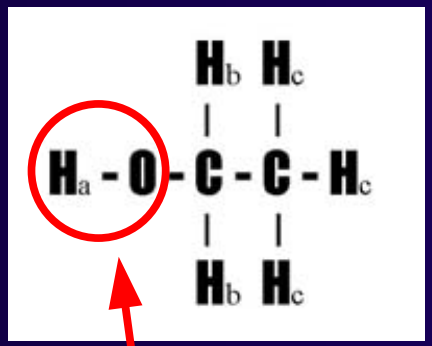
- 1D NMR
  - 'hit, measure' (1D pulse sequence)

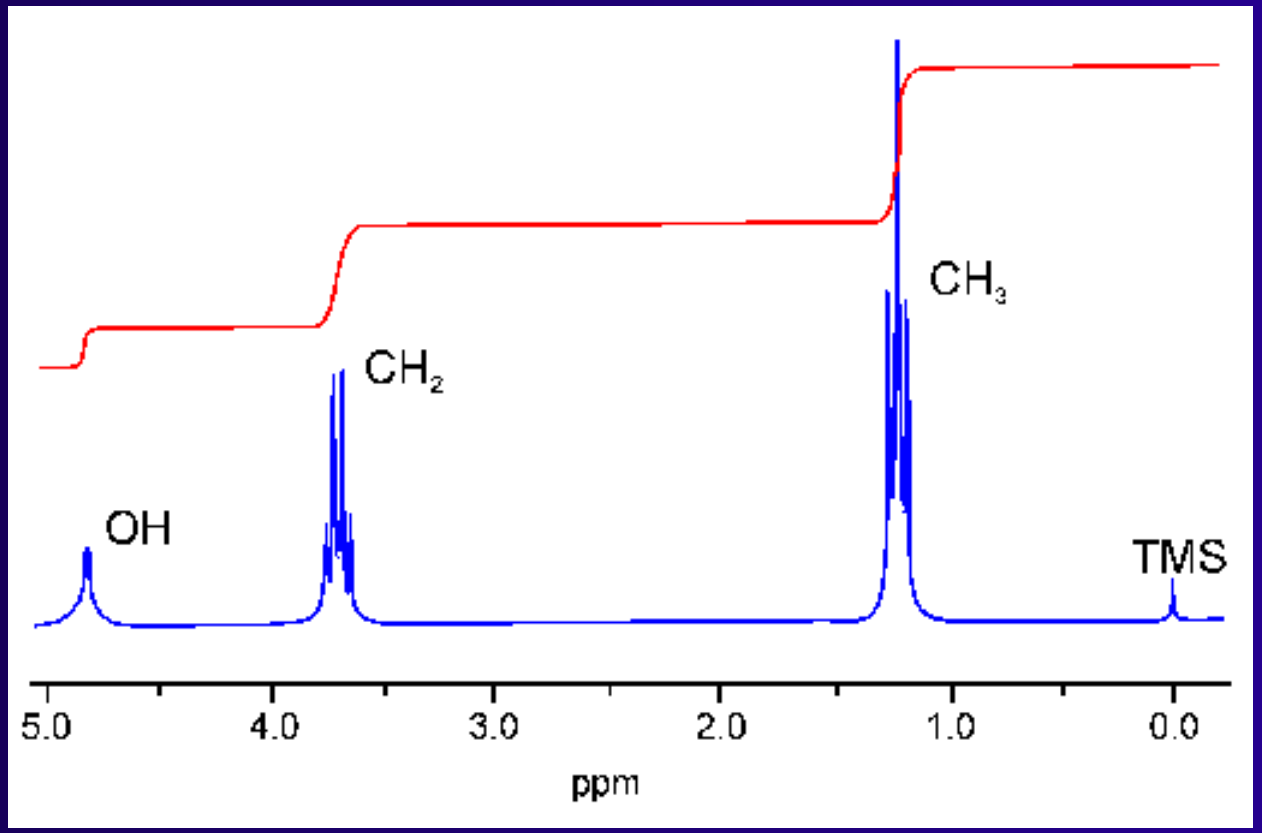


1H NMR experiment

- 1D NMR
  - 1 peak for each proton in a distinct environment within the protein
  - height $\propto$ number of structurally identical H (-CH3)
  - position (shift) $\propto$ electronegativity of surrounding
    - Minute differences in shifts: measured in **P**art **P**er **M**illion of the field
  - width $\propto$ protein size
    - size expressed in terms of tumbling (correlation) time
    - -> there is an experimental limit to the size of the proteins one can determine by NMR ! we need tricks, more tricks....

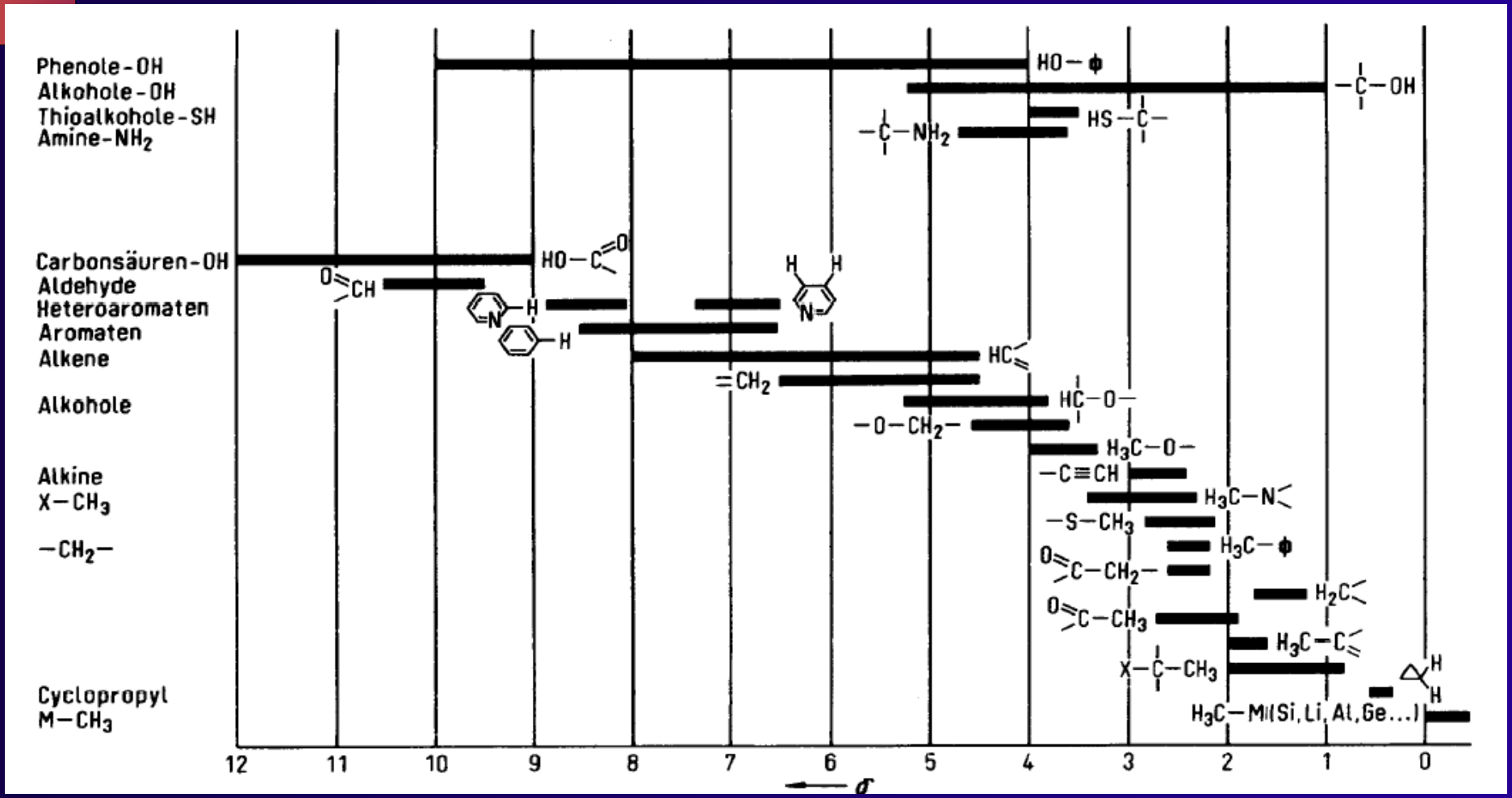
^{1}H NMR spectrum of ethanol





Electronegative group: proton signal is shifted compared to the reference

^{1}H NMR shifts in different molecular environment



What does it mean for NMR spectra of: lipids? Sugars? DNA? Proteins?

1H NMR spectra of proteins

- Problem Number 1: overlap. The larger the protein, the more protons, the worse the overlap

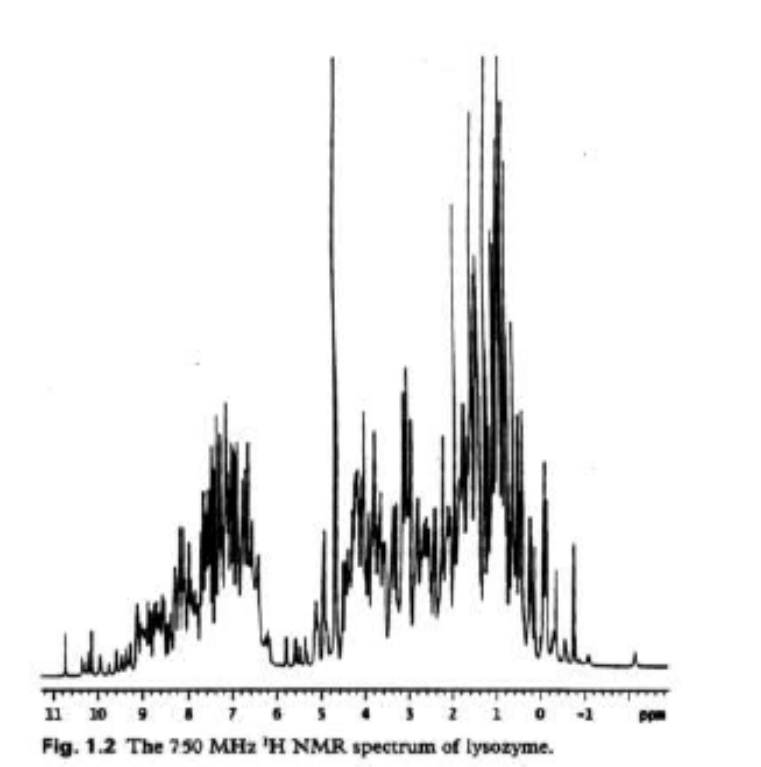


Fig. 1.2 The 750 MHz ^{1}H NMR spectrum of lysozyme.

Solutions:
- higher field, stronger magnet?
- 2D experiments
- 15N, 13C labelling

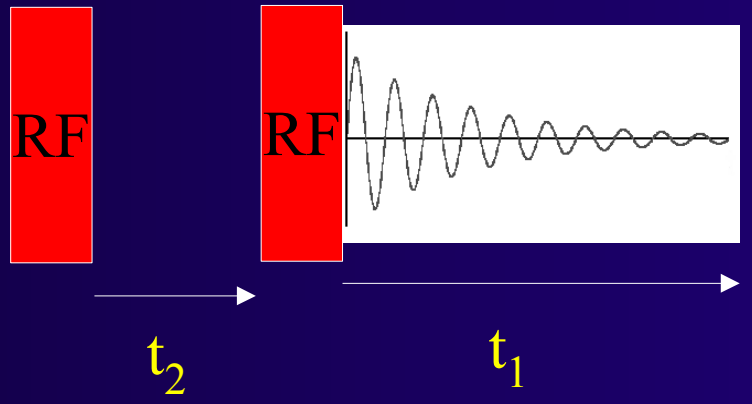
1H NMR spectra of proteins

- Problem number 2: in NMR terms, a large protein [> 250 residues] means also besides overlapping peaks:
  - faster decaying FID
  - broader lines
  - poorer sensitivity

No practical solution for this problem (?).

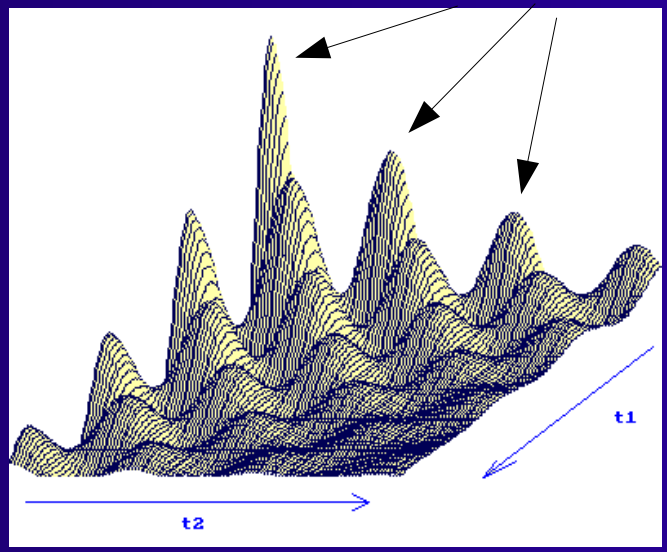
NMR experiments

- 2D NMR
  - 'hit, wait, hit, measure' (2D pulse sequence)

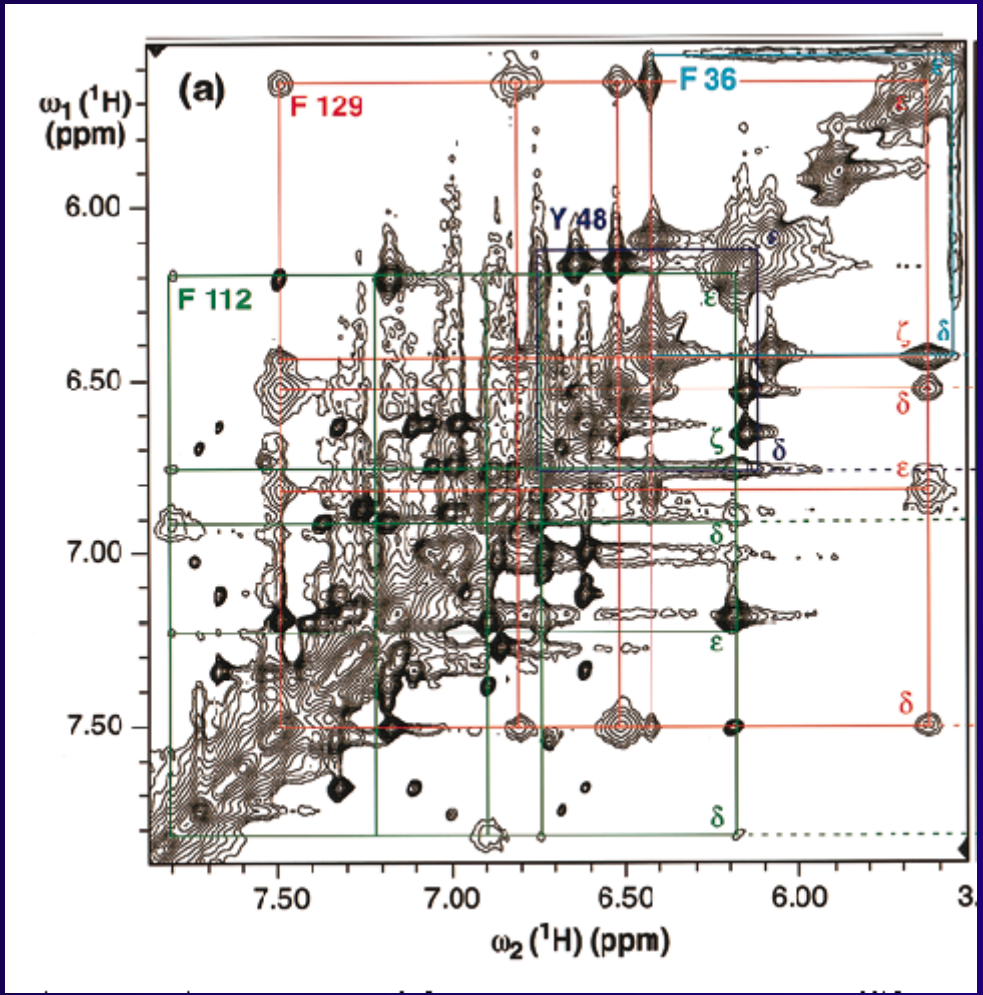


FT

crosspeaks



A slice from above gives a map of crosspeaks.

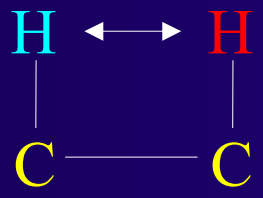


The diagonal contains the 1-dimensional spectrum

Off diagonal cross peaks contain **new information**

NMR experiments

- Useful 2D NMR techniques
  - H-COSY: through-bond connections visible if protons are at most 3 bonds apart



peak on 2D map with position f(H),f(H)

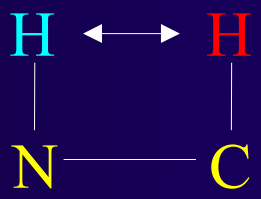
  - H-NOESY: through-space connections visible **if d < 6Å**

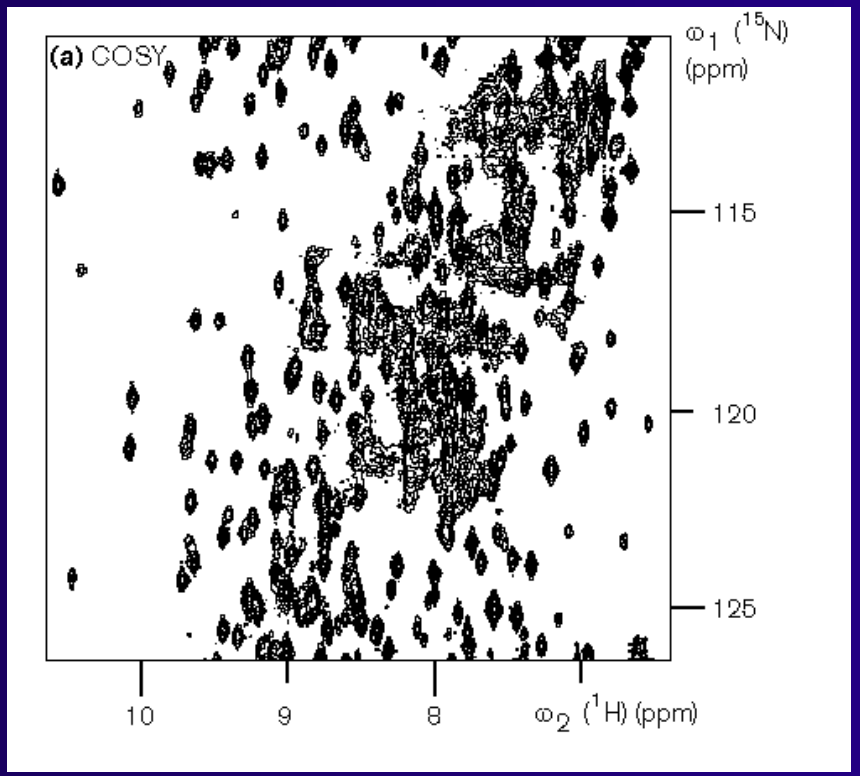
Heteronuclear experiments

- When overlap is too bad for solving a structure by 1H NMR alone
- Label protein with 15N and 13C
  - Expensive, time-consuming, bad yields
  - ...but 15N and 13C resonate at completely different frequencies from 1H
    - Multi-dimensional experiments
    - ... and much more...

Heteronuclear experiments

15N-1H COSY





^{15}N

^{1}H

NMR assignment

- Assignment of spectra
  - method developped by K. Wuethrich
  - first protein 3D-structure solved by NMR in 1983
  - many protons, even in small proteins
    - complex problem
    - relatively simple solution
    - arguably the most fastidious stage of protein structure determination by NMR

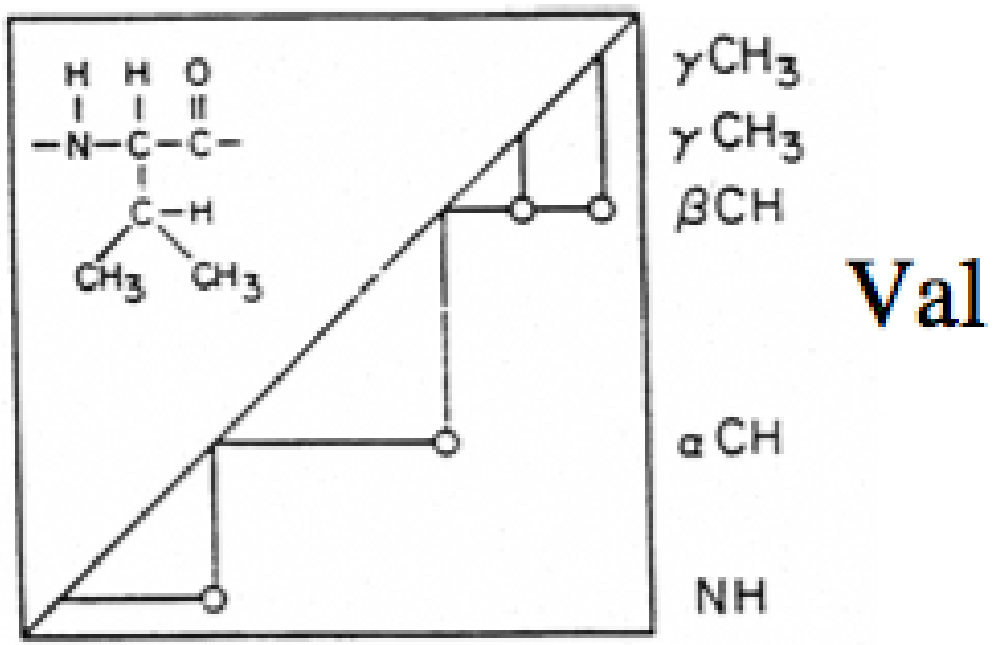
Kurt Wuethrich (ETHZ)

Nobel 2002

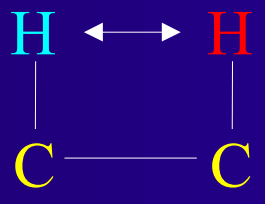
NMR assignment

- Assignment of spectra (K. Wuethrich)
  - map individual amino acids using COSY spectrum
    - set of 2D peaks particular for each side-chain 'spin-system', or relative arrangement of protons

COSY amino acid patterns



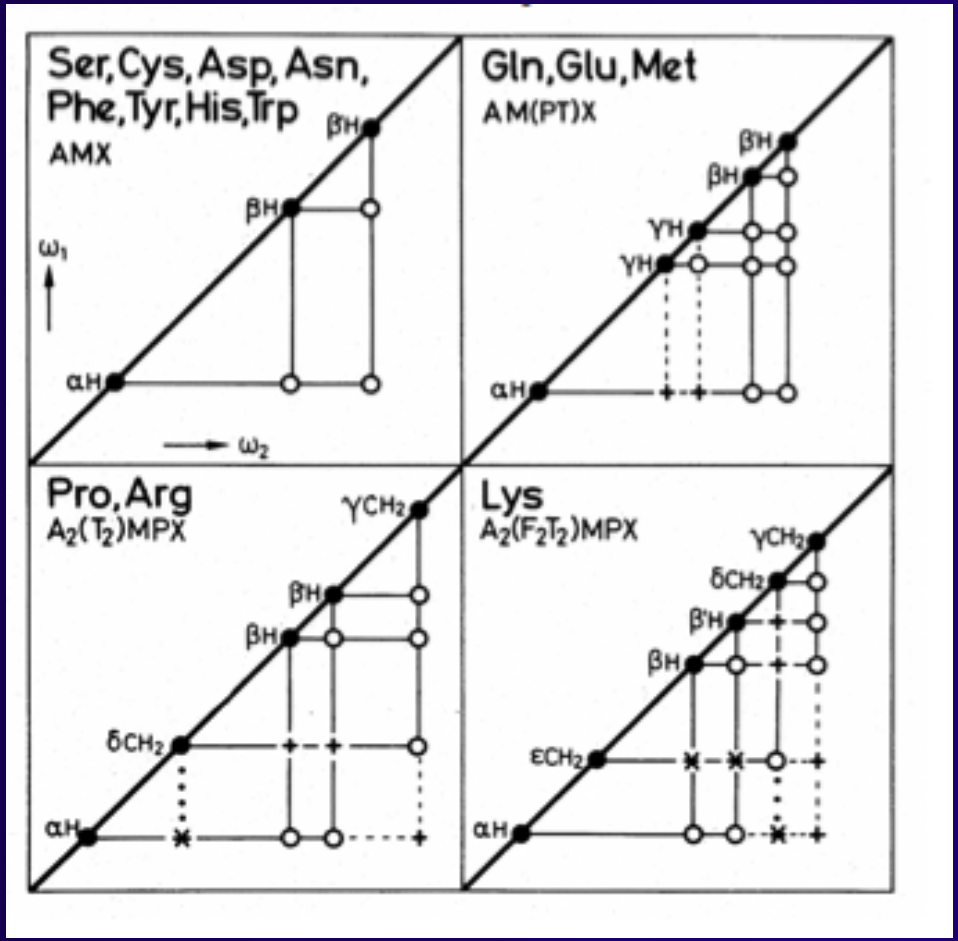
Val



peak on 2D map with position f(H),f(H)

Do you see potential problems for proteins?

Degeneracy of COSY patterns



Some amino acids share the same spin system

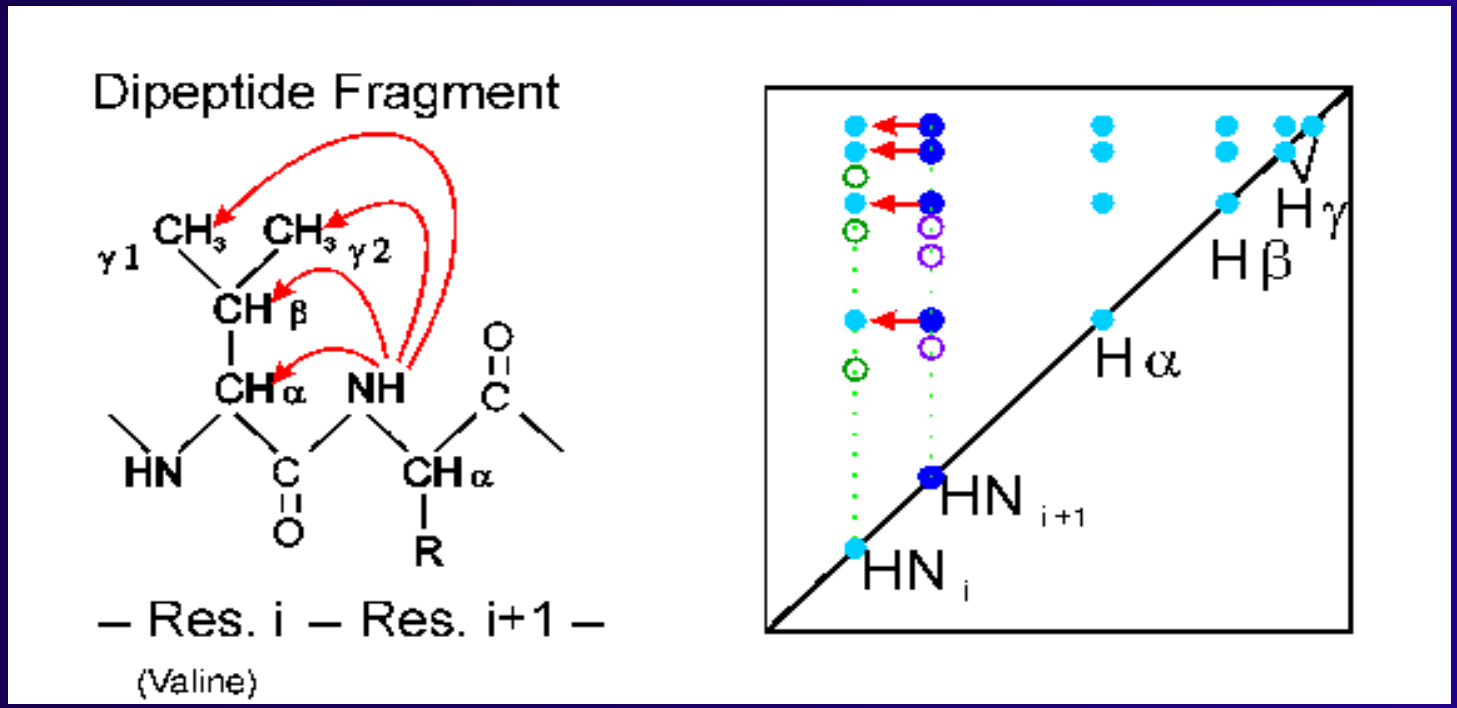
... for DNA it's even worse!

NMR assignment

- Assignment of spectra (K. Wuethrich)
  - map individual amino acids using COSY spectrum
    - set of 2D peaks particular for each side-chain 'spin-system', or relative arrangement of protons
  - locate individual amino acids within the sequence using NOESY spectrum (sequential assignment)
    - through-space connections from HA(i) to HN(i+1)

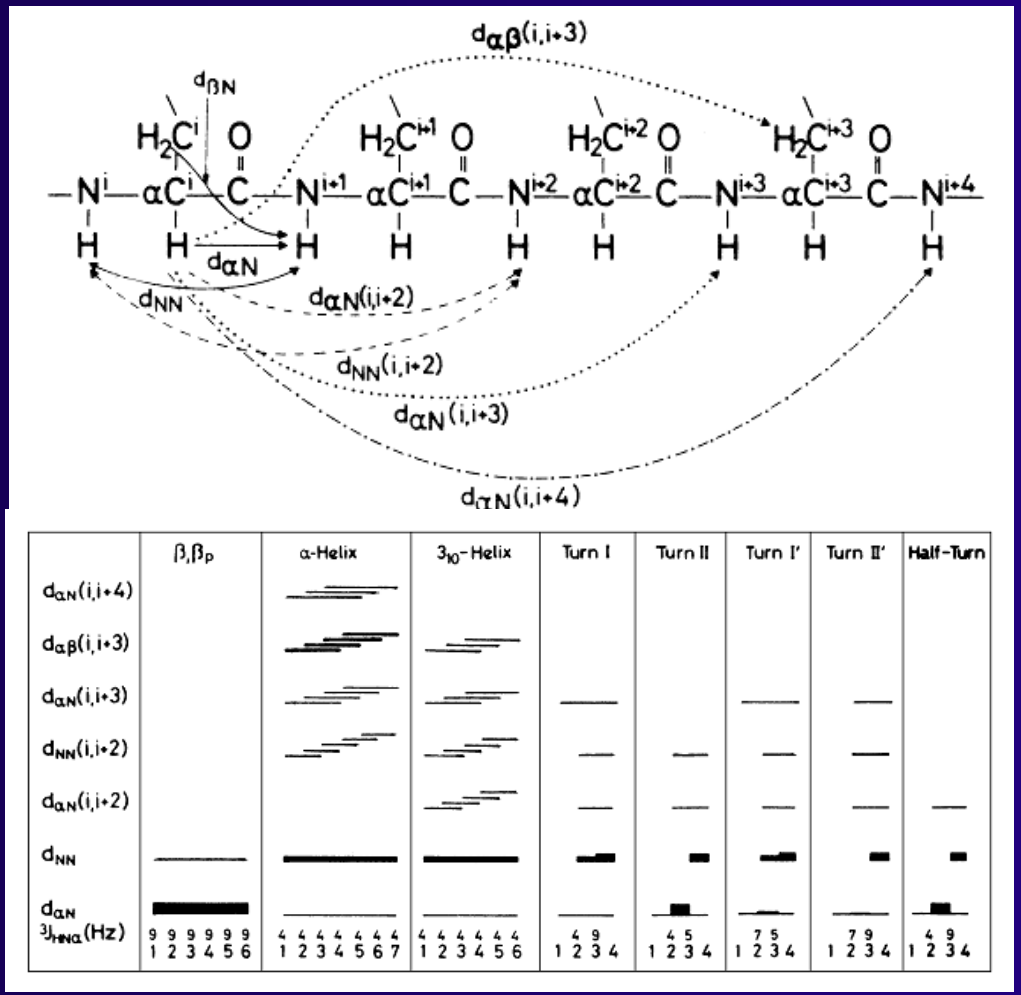
NOESY sequential assignment

NOESY gives through-space connections:
peak on 2D map with position f(H),f(H)
if distance **d < 6Å**



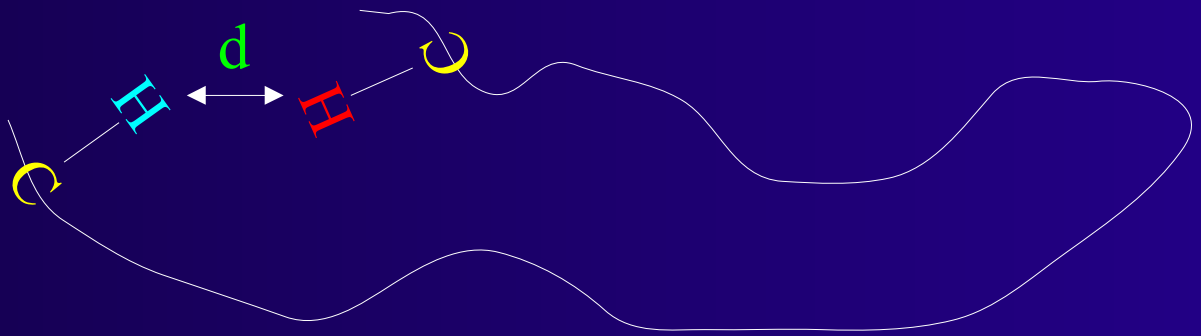
More from the NOESY spectrum

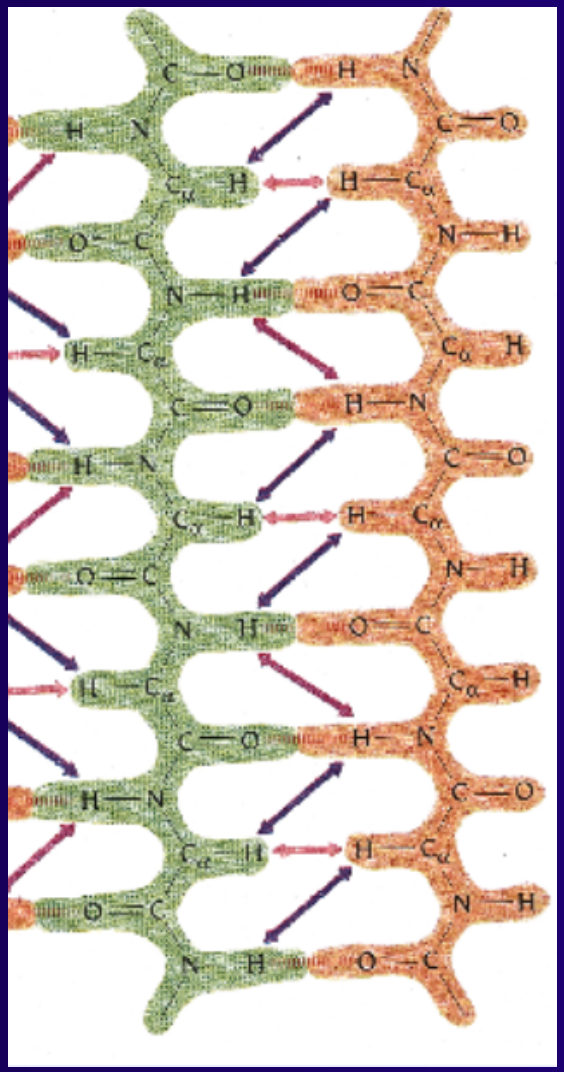
- Secondary structure patterns



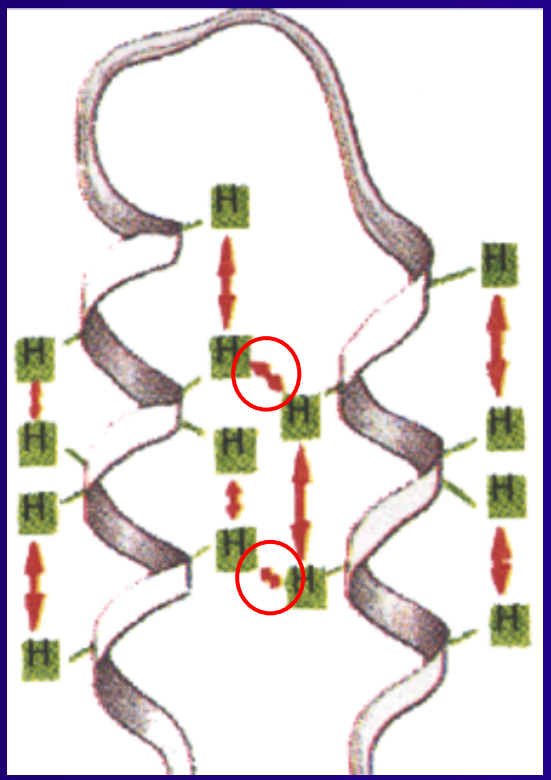
Nuclear Overhauser Effect

- Through-space connections, as given by the NOESY experiment, are the key to solving a protein structure by NMR
- Long-range interactions give the fold of the protein chain





adjacent beta-strands

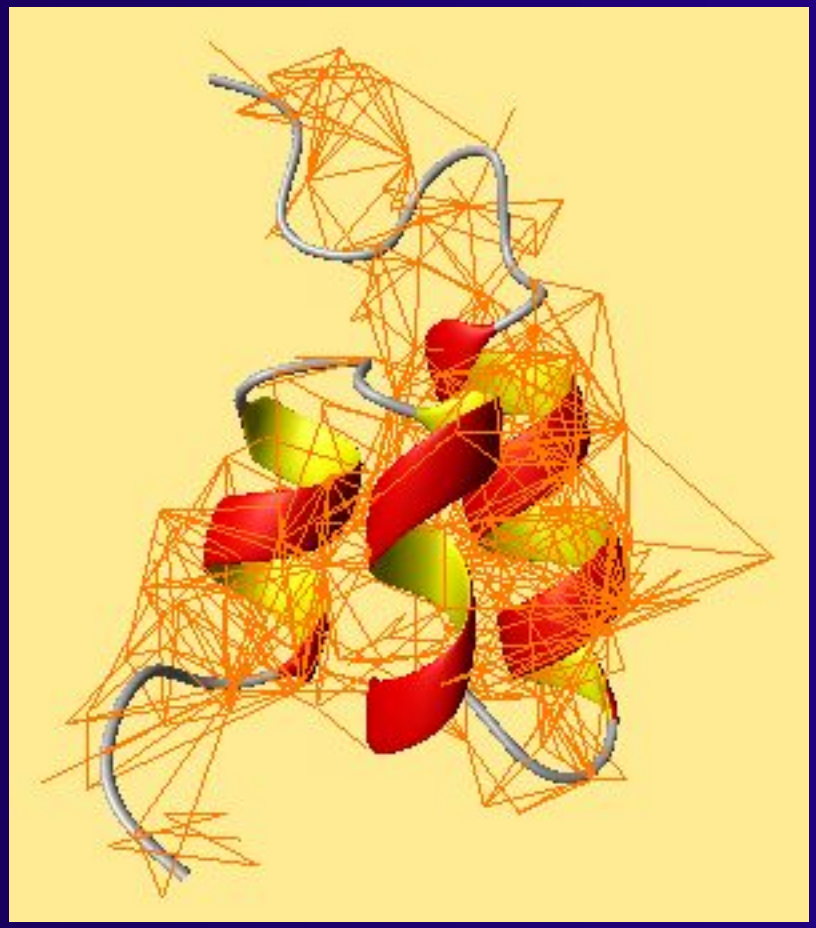


Positioning of two helices relative to each other

NMR assignment

- Assignment of spectra (K. Wuethrich)
  - map individual amino acids using COSY spectrum
    - set of 2D peaks particular for each side-chain 'spin-system', or relative arrangement of protons
  - locate individual amino acids within the sequence using NOESY spectrum (sequential assignment)
    - through-space connections from HA(i) to HN(i+1)
  - compile list of all other peaks arising from through-space connections (NOEs)
    - -> set of pairwise distance restraints

NMR distance restraints



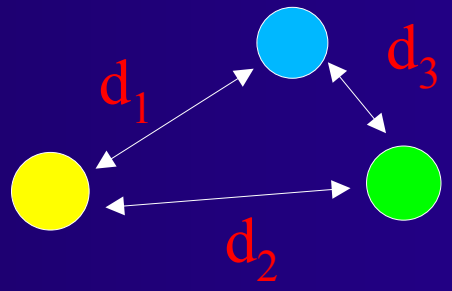
NMR structure calculation

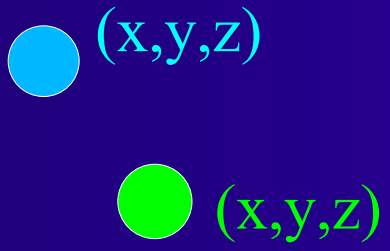
- How to get from the NOEs to the 3D model?

pairwise distances

?

atomic coordinates





NMR structure calculation

- Distance geometry
  - solves the triangular inequality
- Simulated annealing (Michael Nilges)
  - fancy monte-carlo simulation ("travelling salesman problem")
- Torsion angle dynamics (Peter Güntert)
  - hybrid method in dihedral angle space

Michael Nilges (EMBL)

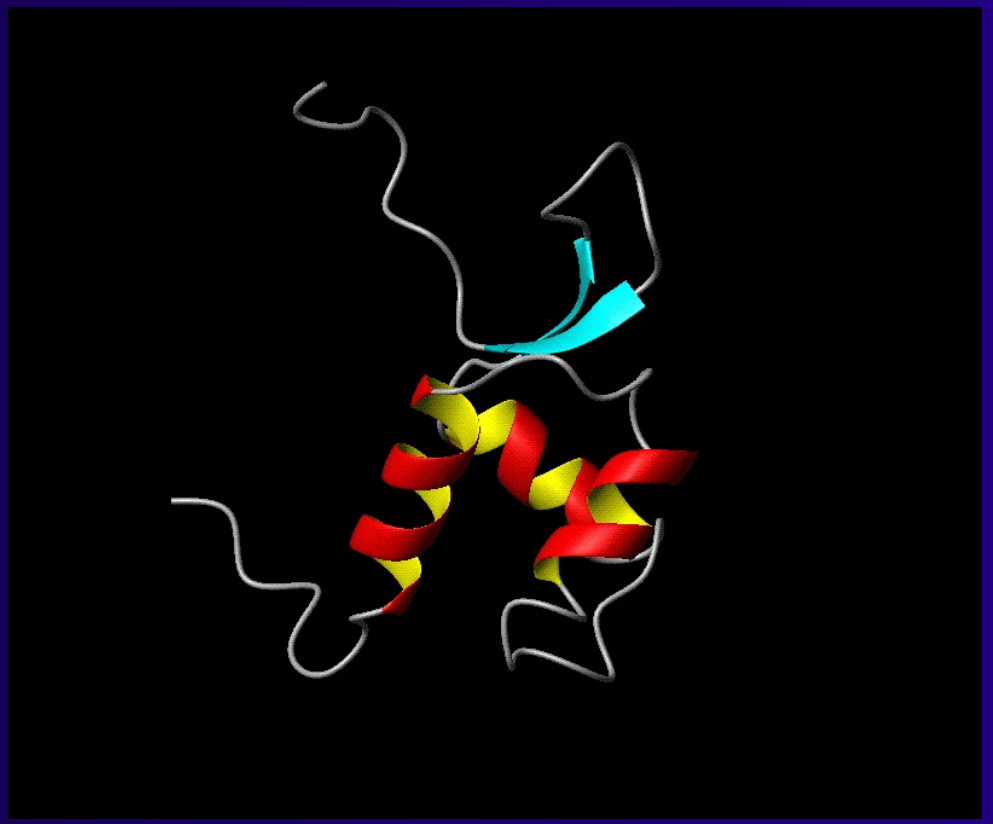
- XPLOR structure calculation and molecular dynamics algorithm

Peter Guentert (ETHZ)

- DYANA structure calculation programme with integrated automatic assignment protocol

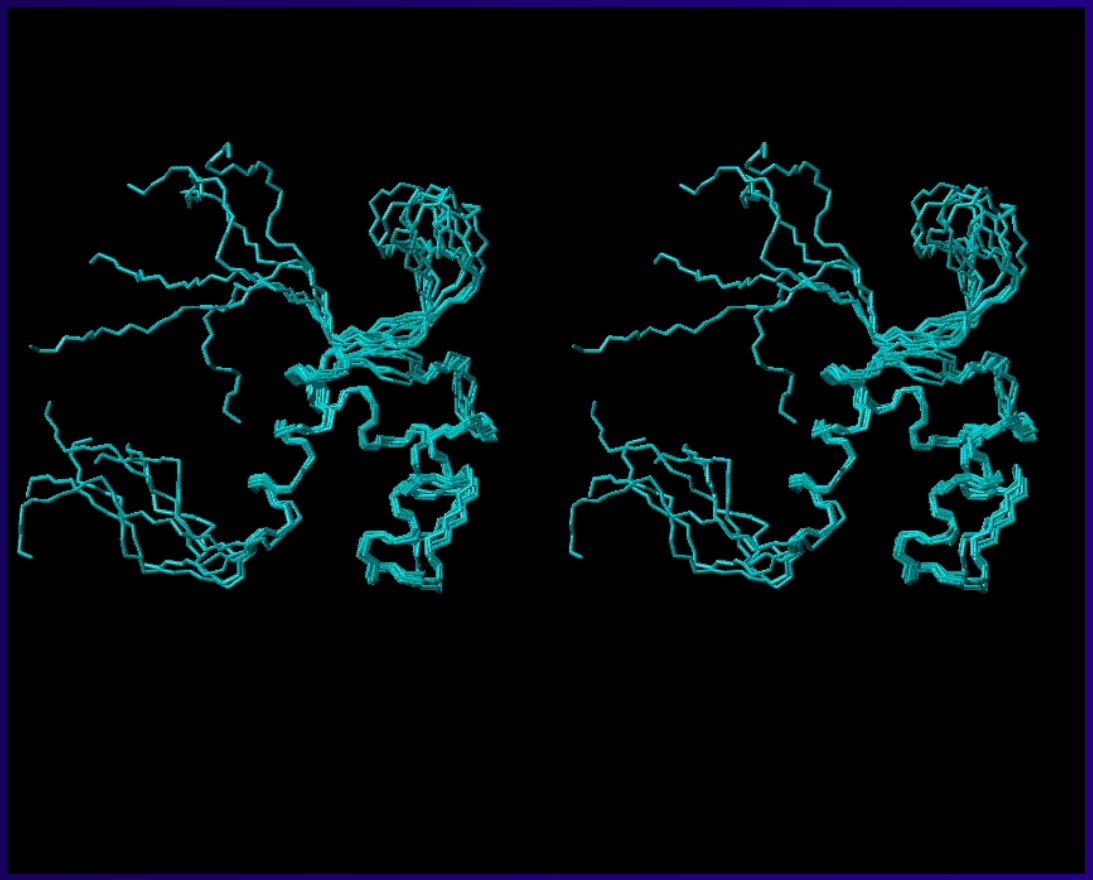
NMR structures

- What you see...



NMR structures

- ...is not what you get from NMR!



NMR structures

- NMR-derived distance restraints (NOEs) are upper-limits ("d < 6 Å")
  - transformation of distances to coordinates gives many solutions
  - NMR relies on cooperativity of distance restraints:
    - the more restraints per residue, the better defined the structure
  - one NOE set produces a family of structures:
    - loops: few experimental restraints -> bad definition -> "fuzzy"
    - core: lots of long-range restraints -> good definition -> "compact"

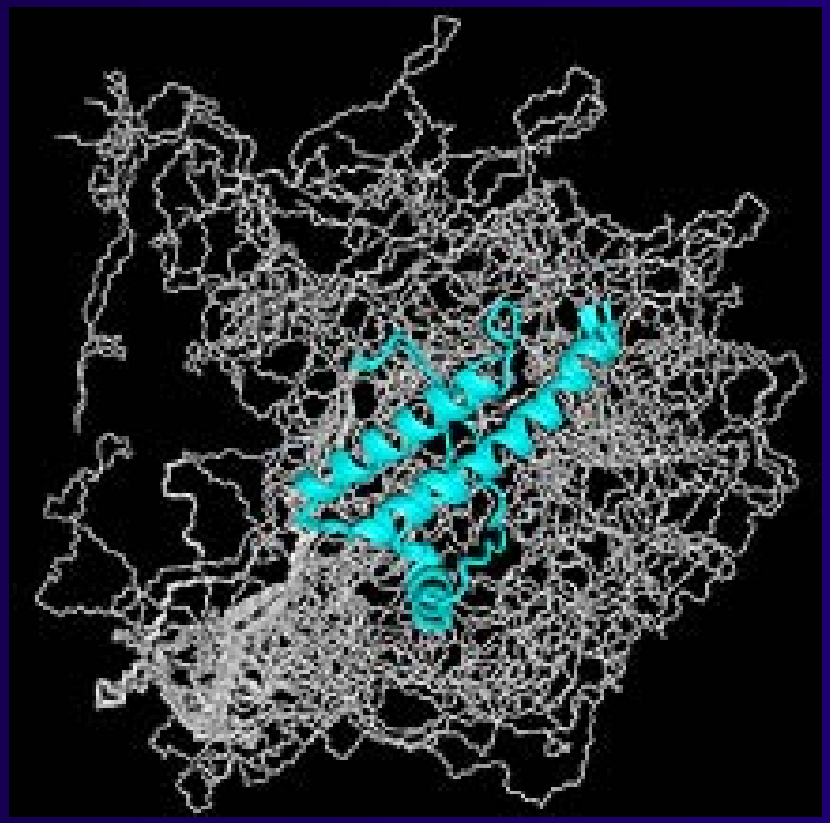
NMR structure determination

- Difficulties
  - protein in solution: protein has to be soluble
  - insensitive method: requires high concentrations of proteins
  - overlap: direct determination of 3D structures for small proteins only (150-200 residues)
- Advantages
  - no chemical modification necessary
  - protein in solution: no crystal packing artefacts, allows direct binding experiments, hydrodynamic and folding studies
  - assignment of labile regions possible: no gaps in structure

Detecting unstructured loops

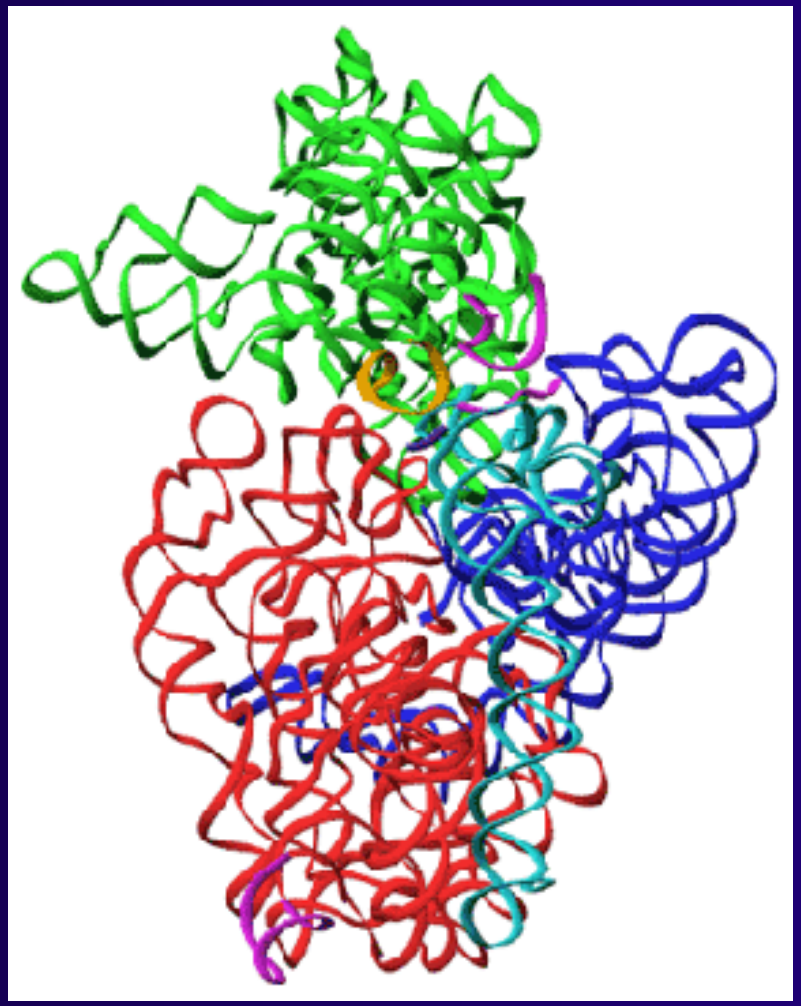
- NMR spectrum
  - spectrum shows no long-distance interactions but sequential assignment is possible
  - backbone is free to adopt a range of conformations: greater variation in structure coordinates for loop residues
- Crystallography
  - electron density map shows nothing at all
  - structures will have gaps for residues in mobile loops

Prion protein by NMR



Not explorable
by XRay

Ribosome structure by XRay



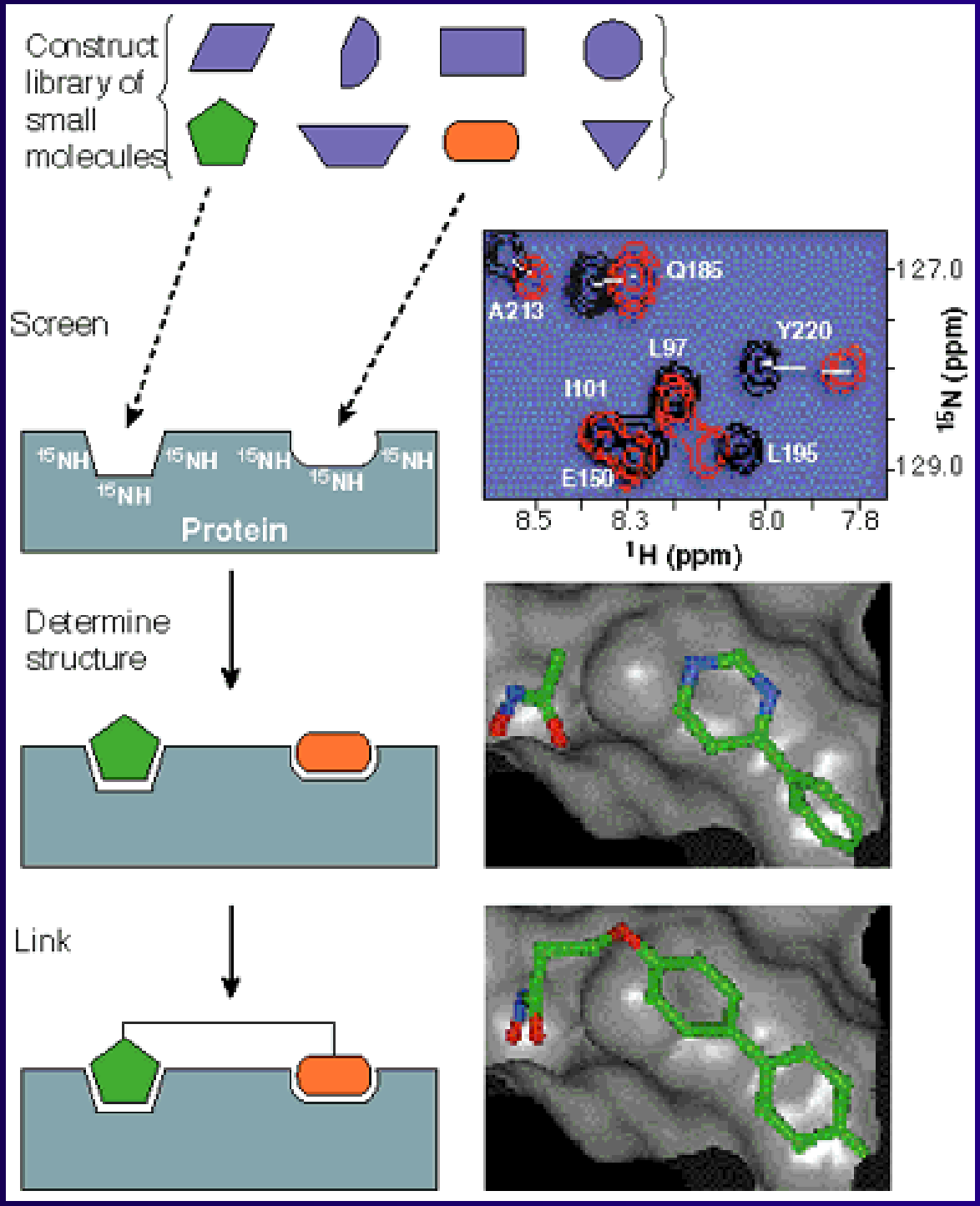
Impossible task
for NMR

NMR applications

- SAR: "shot-gun" approach to drug design
- Exploring Fibrils by solid-state NMR
- Protein folding mechanism
- NMR of proteins in bicelles (semi-crystalline state)
- TROESY

Structure Activity Relationship by NMR

- Drug design
  - method developed by Abbott Laboratories
- Aim:
  - discover high-affinity ligands for proteins
- Example application:
  - antiviral agent against the human papillomavirus

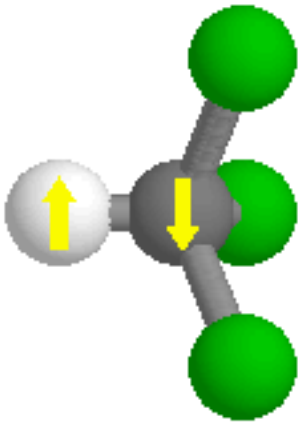

Construct library of small molecules
Screen
15NH
15NH
15NH
15NH
15NH
Protein
Determine structure
Link
Q185
A213
Y220
L97
I101
L195
E150
127.0
129.0
15N (ppm)
8.5
8.3
8.0
7.8
1H (ppm)

Quantum computing by NMR

- Computers
  - molecules
- Information
  - atomic nuclei state
- Programming
  - radio-frequency pulse

Quantum computing by NMR

- from bits: 0 1 to <u>qubits</u>: $|0,0\rangle$ $|0,1\rangle$ $|1,0\rangle$ $|1,1\rangle$



The ^{1}H and ^{13}C nuclei in isotopically labelled chloroform behave like small magnets, and interact with an external magnetic field. Nuclear spins aligned with the field correspond to qubits in state $|0\rangle$ while those aligned against the field correspond to state $|1\rangle$. The molecule depicted above thus represents a computer in the state $|0,1\rangle$. The three chlorine nuclei, shown in green, can be ignored.

Quantum computing by NMR

- WHY?
  - atoms change energy states very quickly
    - -- much more quickly than even the fastest computer processors.
  - each qubit can take the place of an entire processor
    - -- 1,000 ions of barium could take the place of a 1,000 processor computer.

Quantum computing by NMR

- WHAT sort of problem a quantum computer would be able to solve in principle?
  - Large-scale cryptography
  - modelling and indexing very large databases

Quantum computing by NMR

- WHO wants to build a quantum computer?
  - IBM/MIT/Berkeley/ Stanford
    - Isaac Chuang (IBM) Neil Gershenfeld (MIT)
    - "Enabling technology" for NMR-based quantum computing; scale up to 10-40 qubits
  - Harvard/ MIT/Los Alamos
    - David Cory (Harvard)
    - Quantum algorithms and NMR-based systems
  - Oxford University
    - David Deutsch, Jonathan Jones
    - Ion-trap and NMR implementations; quantum information theory

finally...the R.Ernst view of NMR