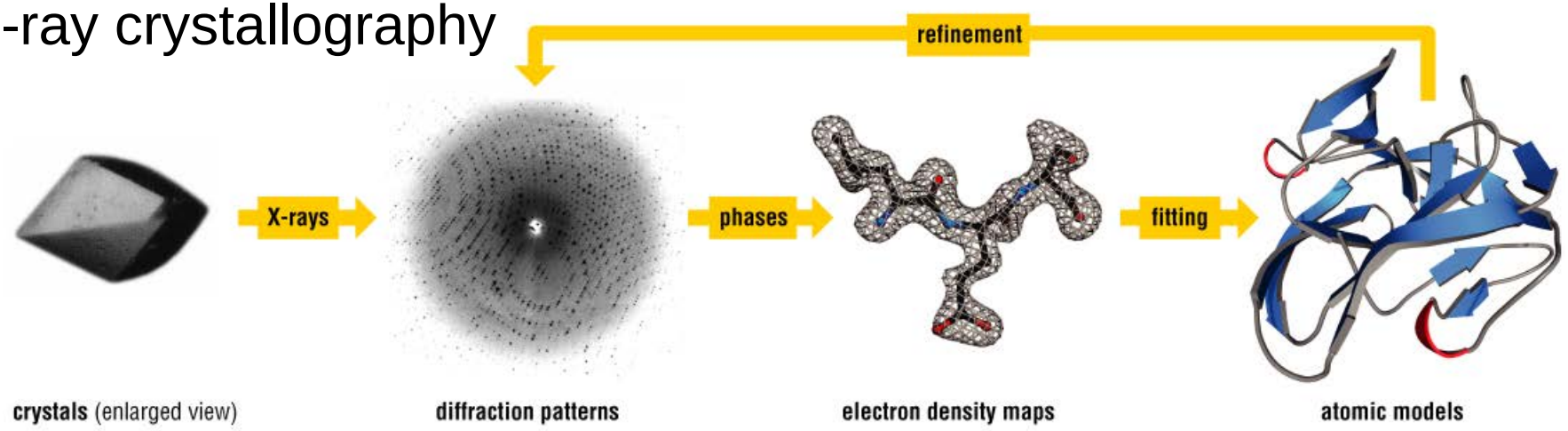
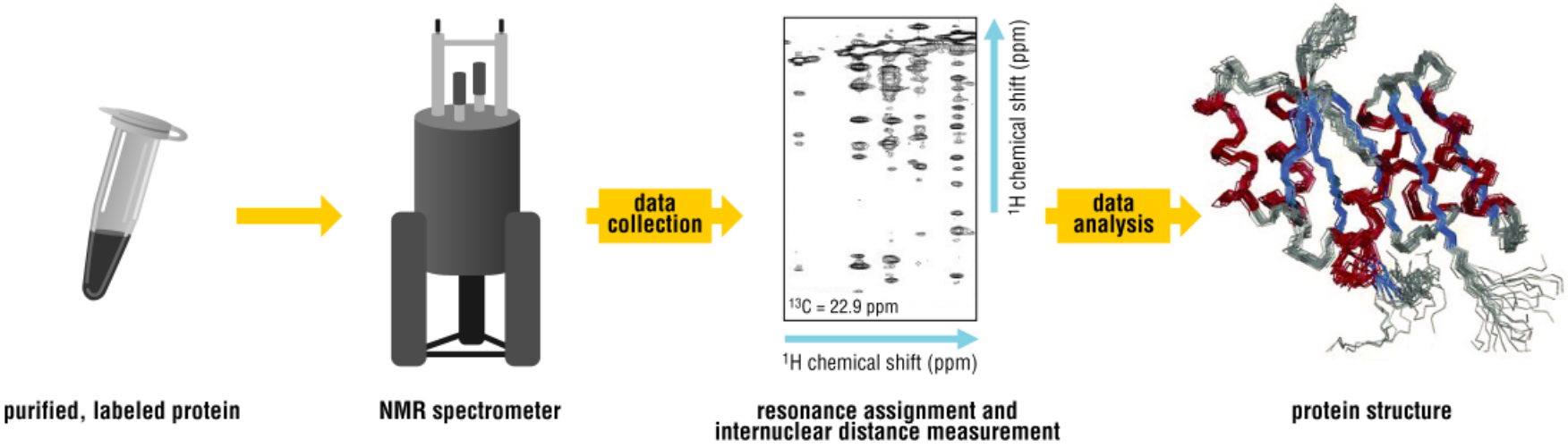
Methods to determine 3D structures of proteins

X-ray crystallography

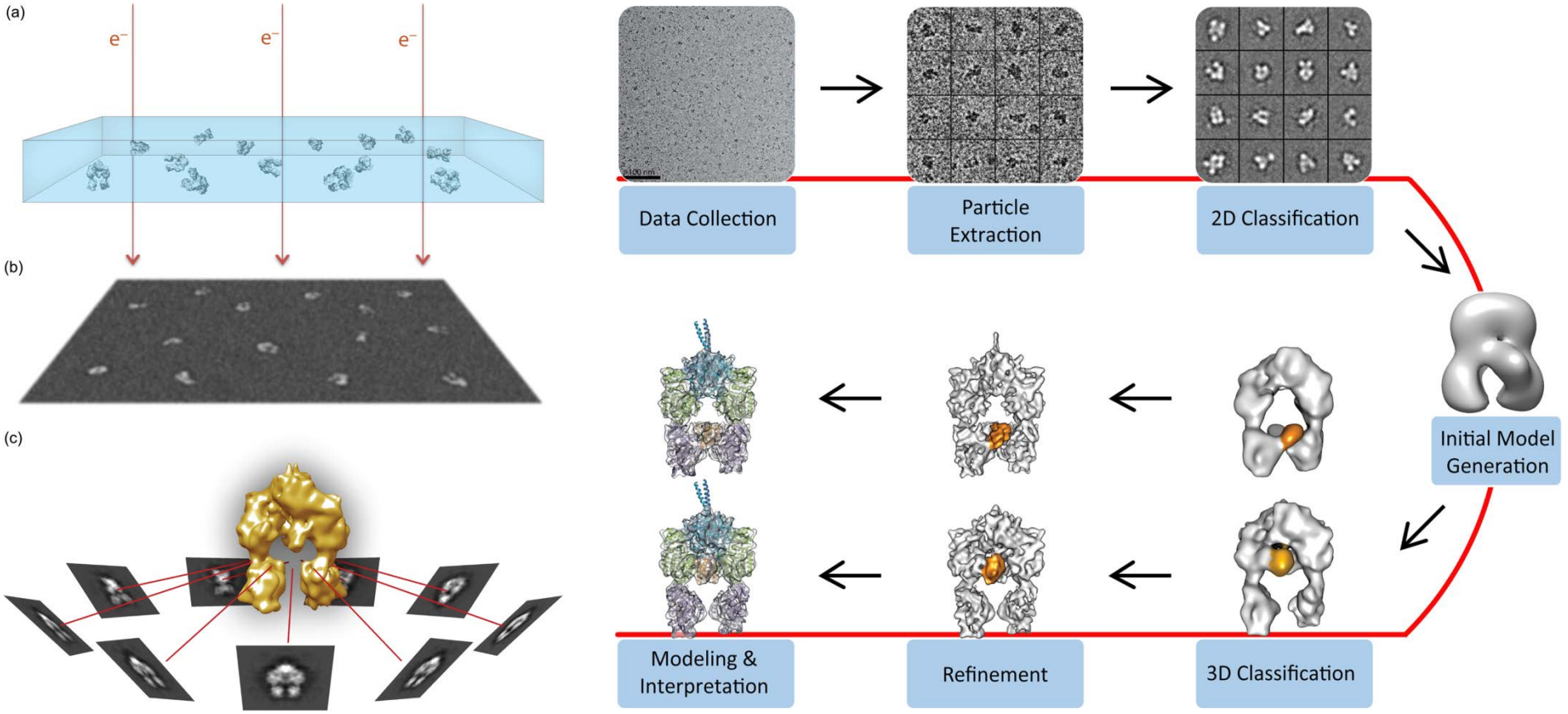


Nuclear magnetic resonance (NMR) spectroscopy



Petsko, G. and Ringe, D. Protein Structure and Function, OUP, 2003

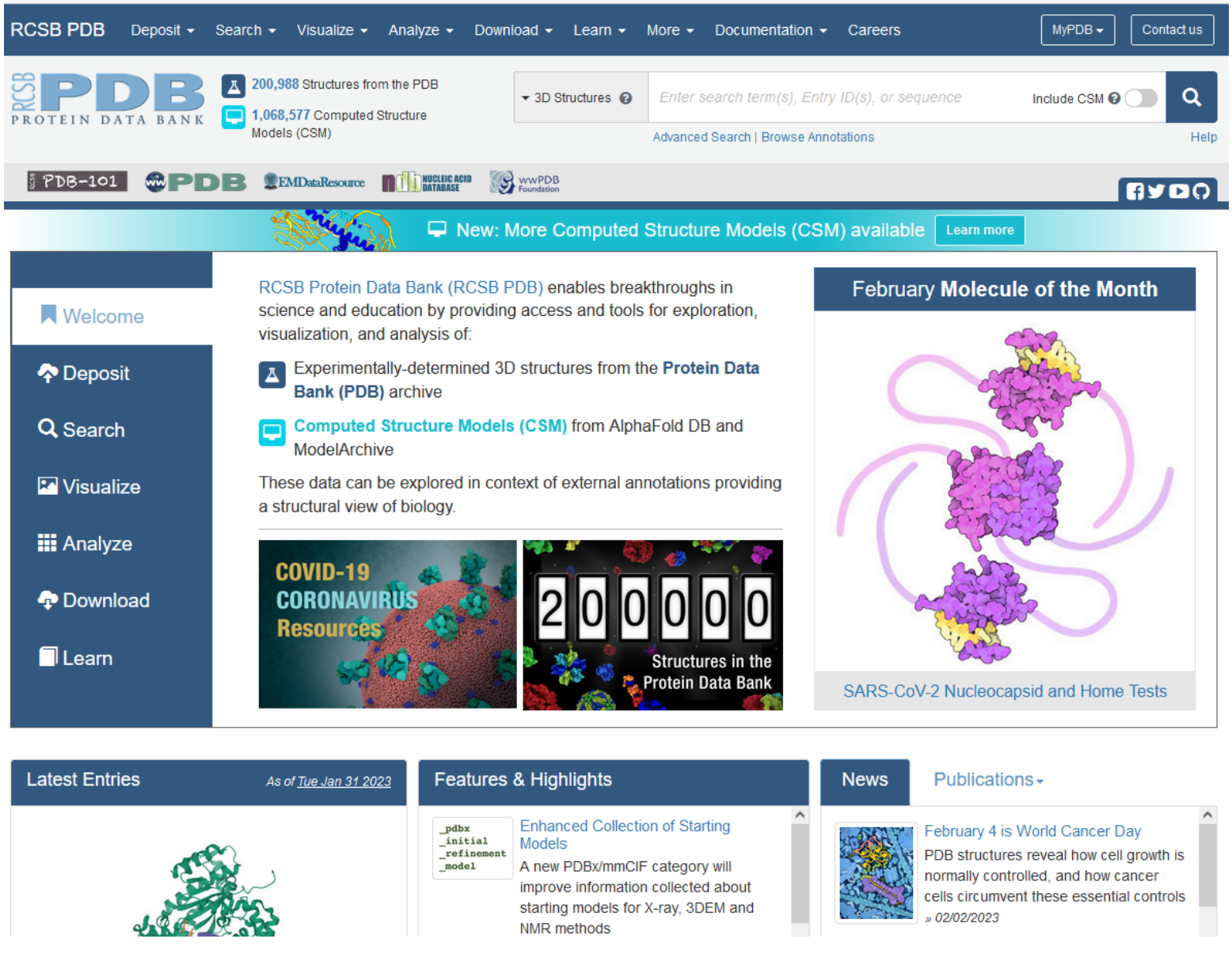
Single-particle cryo-electron microscopy



Skiniotis G, Southworth DR. Single-particle cryo-electron microscopy of macromolecular complexes. Microscopy (Oxf). 2016 Feb;65(1):9-22.

Protein Data Bank: depository of biomolecular structures

https://www.rcsb.org/



AlphaFold

- predicts protein structure from sequence using AI/deep learning
- most accurate structure prediction currently available
- trained using structures in the PDB
- uses multiple sequence alignment

AlphaFold protein structure database

- structure predictions for human proteins and proteins from model organisms
- https://alphafold.ebi.ac.uk/
- structures predicted by AlphaFold can also be accessed via UniProt and PDB

AlphaFold2 Colab

- google alphafold2 colab or AphaFold2.ipynb
- use for sequences not in Alphafold database and for complexes/multimers

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. Highly accurate protein structure prediction with AlphaFold. Nature. 2021 Aug;596(7873):583-589. https://www.nature.com/articles/s41586-021-03819-2

Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, Yuan D, Stroe O, Wood G, Laydon A, Žídek A, Green T, Tunyasuvunakool K, Petersen S, Jumper J, Clancy E, Green R, Vora A, Lutfi M, Figurnov M, Cowie A, Hobbs N, Kohli P, Kleywegt G, Birney E, Hassabis D, Velankar S. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. 2022 Jan 7;50(D1):D439-D444. https://doi.org/10.1093/nar/gkab1061

AlphaFold limitations

- Only predicts monomers

Complexes/multimers can be predicted with AlphaFold2 Colab (using AlphaFold2-Multimer) – good but not as good as AlphaFold2 for monomers

- Only predicts protein structure

Predictions will not contain bound metals, cofactors, post-translational modifications (glycosylation, phosphorylation…), small ligands, nucleic acids, bound peptides or proteins (unless provided in AlphaFold2-Multimer) etc. – but the structure is consistent with bound metals, cofactors, PTMs, bound ligands etc.

- For proteins with multiple conformations, it cannot be controlled which conformation it will predict

- Does not reliably predict the effect of mutations

- Does not reliably predict antibody:antigen interactions

What can you obtain from NMR spectroscopy?

- Observable signals from individual atoms: ^{1}H, ^{13}C, ^{15}N, ^{19}F, ^{31}P, ^{2}H...
- Structural information: bonds, dihedral angles, distances, bond orientations, solvent accessibility
- Interactions with other molecules: binding site, structure of the complex, affinity (K_d), binding kinetics
- Dynamics (ps-days): conformational exchange, rotational and translational diffusion

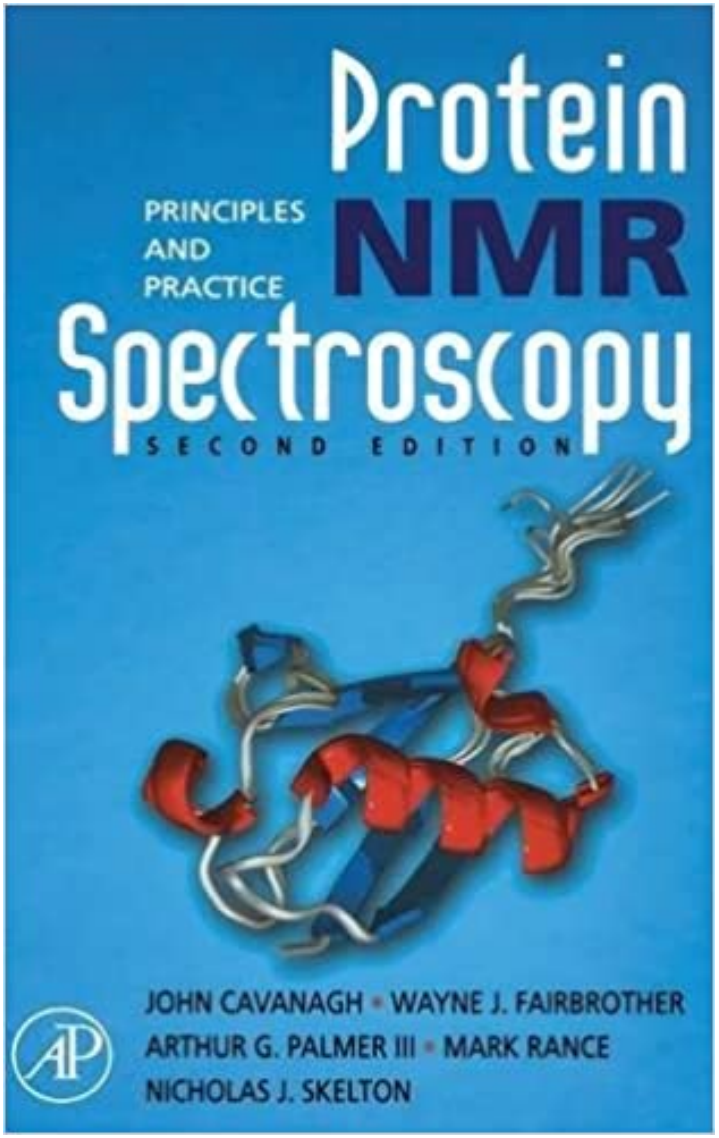
Brief history of NMR

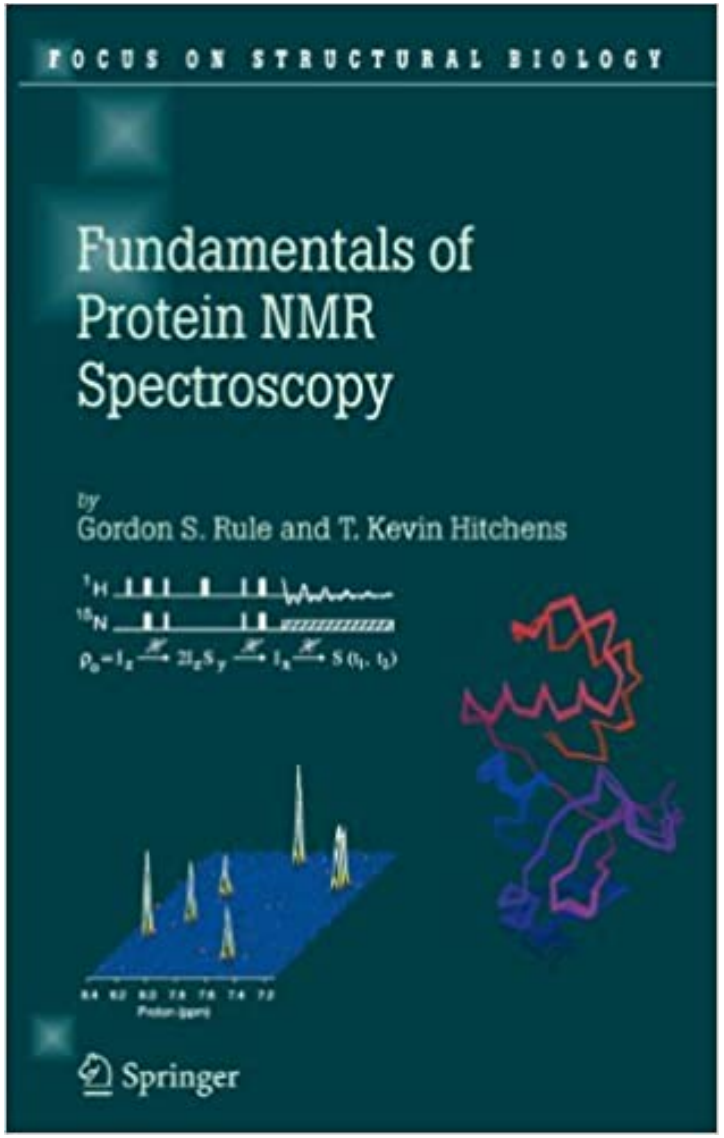
1937	Rabi	NMR in a molecular beam
1946	Bloch, Purcell	NMR in condensed phase
1955	Solomon	Nuclear Overhauser effect (NOE)
1966	Ernst, Anderson	Fourier transform NMR
1975	Jeener, Ernst	2D NMR
1985	Wüthrich	Solution structure of a small protein (BPTI) from NMR experimental data
1987		3D NMR ^{13}C and ^{15}N labeling
1990		Pulsed field gradients
1996/1997		Residual dipolar couplings TROSY

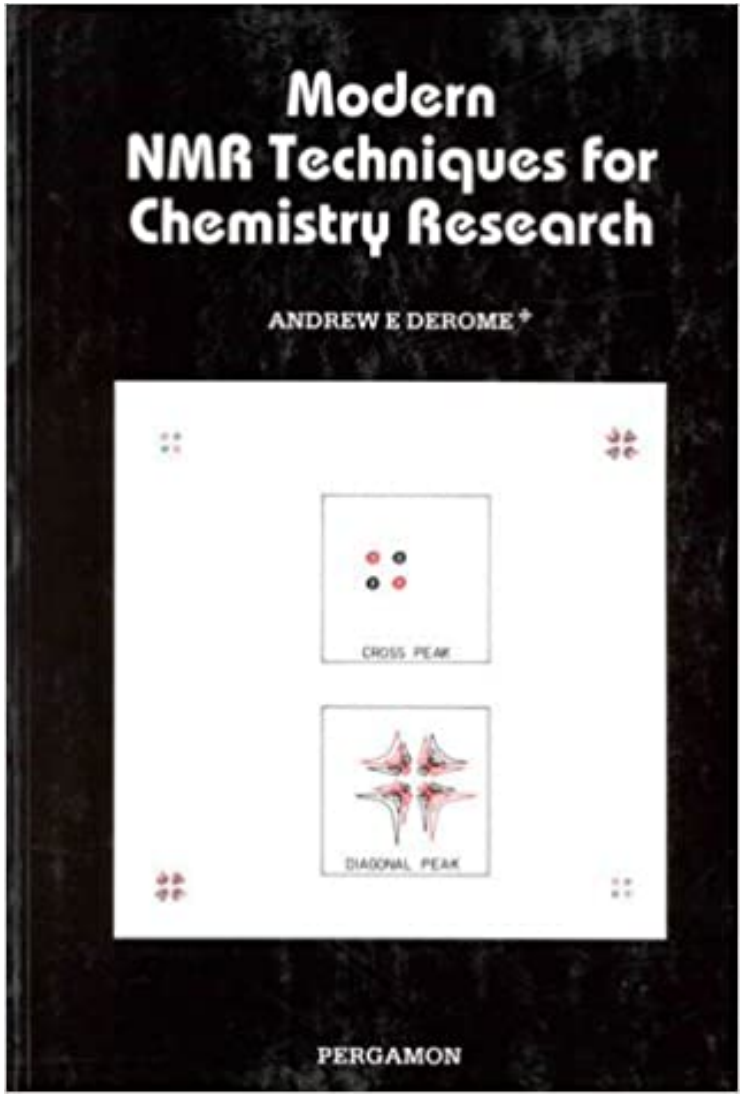
Nobel Prizes for NMR

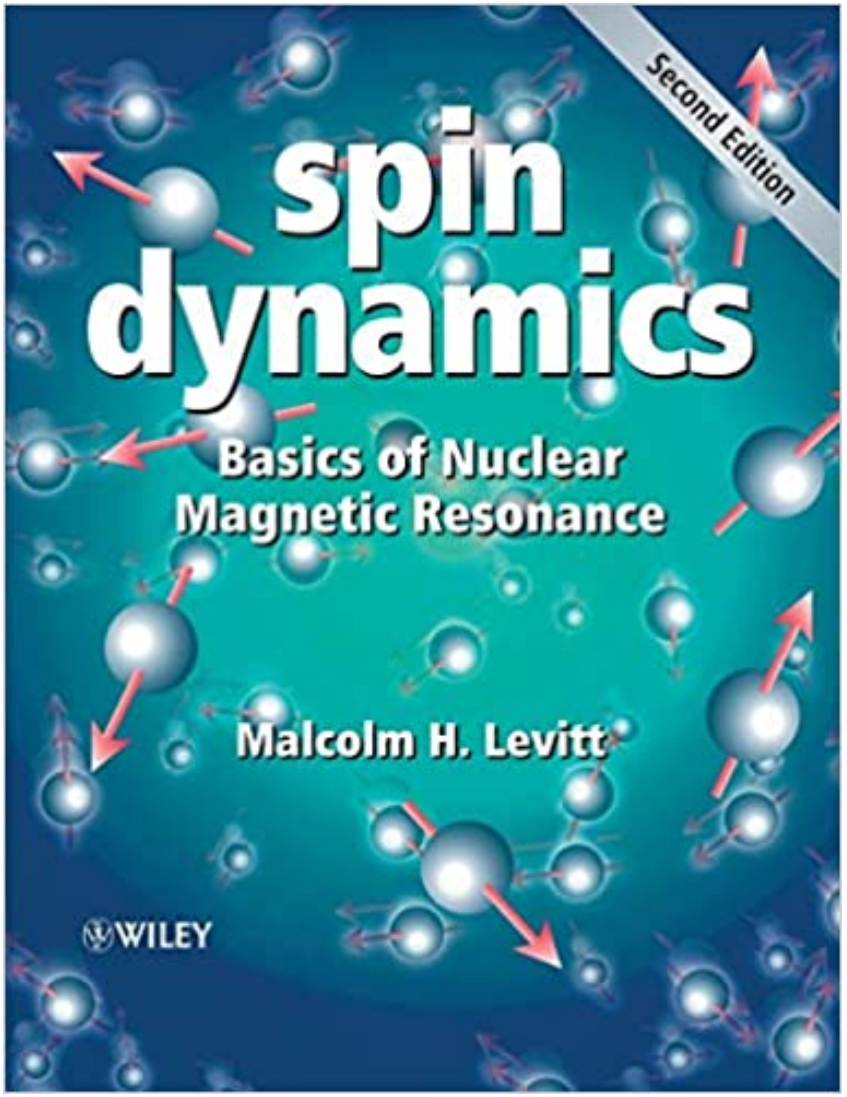
1944	Physics	Rabi (Columbia)	Measurements of nuclear magnetic moment using a molecular beam in an oscillating magnetic field (the foundation of NMR)
1952	Physics	Bloch (Stanford) Purcell (Harvard)	NMR in condensed phase (water & paraffin)
1991	Chemistry	Ernst (ETH Zurich)	2D NMR
2002	Chemistry	Wüthrich (ETH Zurich)	NMR methods for determining the 3D structure of biological macromolecules in solution
2003	Medicine	Lauterbur (Urbana) Mansfield (Nottingham)	MRI

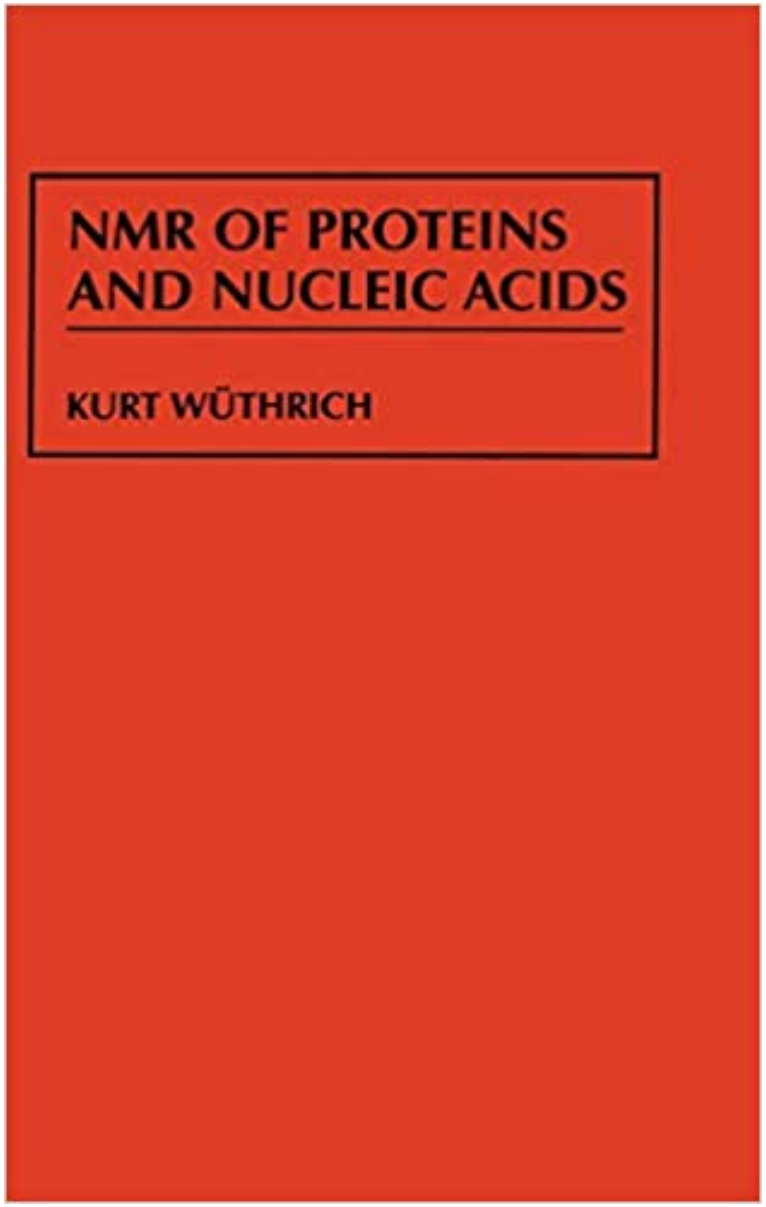
http://mriquestions.com/who-discovered-nmr.html


Protein
NMR
Spectroscopy
PRINCIPLES
AND
PRACTICE
SECOND EDITION
JOHN CAVANAGH • WAYNE J. FAIRBROTHER
ARTHUR G. PALMER III • MARK RANCE
NICHOLAS J. SKELTON
AP


FOCUS ON STRUCTURAL BIOLOGY
Fundamentals of
Protein NMR
Spectroscopy
by
Gordon S. Rule and T. Kevin Hitchens
Springer


Modern
NMR Techniques for
Chemistry Research
ANDREW E DEROME*
CROSS PEAK
DIAGONAL PEAK
PERGAMON


Second Edition
spin
dynamics
Basics of Nuclear
Magnetic Resonance
Malcolm H. Levitt
WILEY



Miscellaneous web resources:

- Understanding NMR spectroscopy course (James Keeler) notes: http://www-keeler.ch.cam.ac.uk/
- ANZMAG lectures: https://www.youtube.com/user/ANZMAG/featured
- ICMRBS (International Conference on Magnetic Resonance in Biological Systems) YouTube channel: https://www.youtube.com/channel/UCsxup-QiNEeBrfo-4d5w33Q/videos

Principles of NMR spectroscopy

Atomic nuclei have spin angular momentum (spin) $\boldsymbol{I}$ characterized by spin quantum number I

$$|\boldsymbol{I}| = \hbar\,[I(I+1)]^{\frac{1}{2}}$$

$\hbar = \frac{h}{2\pi} = 1.055 * 10^{-34}\ Js$ — h: Planck constant

- Odd mass number: half-integer spins
- Even mass number, even atomic number: spin=0
- Even mass number, odd atomic number: integer spin

Nuclei with spin have magnetic dipole moment: $\boldsymbol{\mu} = \gamma \boldsymbol{I}$

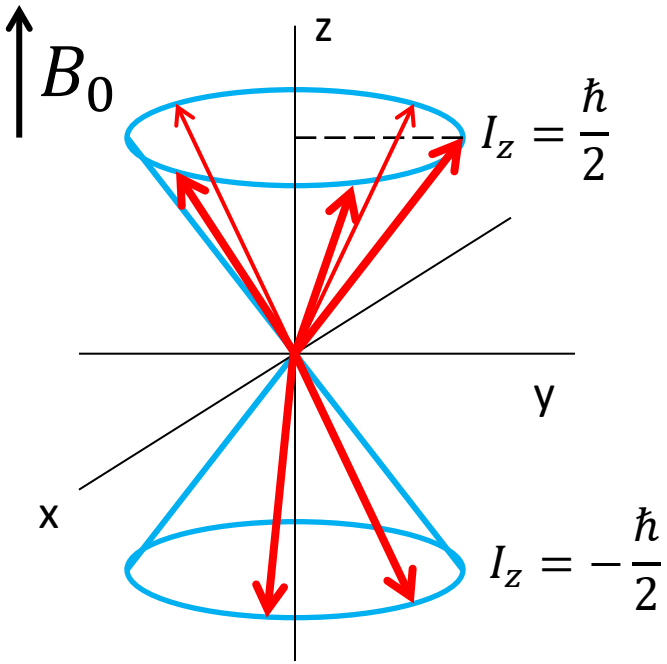
γ: gyromagnetic ratio

Nuclei with spin > ½ also have electric quadrupole moment: faster relaxation, broad signals, not so good for NMR

Nucleus	I	γ $(\mathrm{T\,s})^{-1}$	Natural abundance (%)
^{1}H	1/2	2.6752×10^{8}	99.99
^{2}H	1	4.107×10^{7}	0.012
^{13}C	1/2	6.728×10^{7}	1.07
^{14}N	1	1.934×10^{7}	99.63
^{15}N	1/2	-2.713×10^{7}	0.37
^{17}O	5/2	-3.628×10^{7}	0.038
^{19}F	1/2	2.518×10^{8}	100.00
^{23}Na	3/2	7.081×10^{7}	100.00
^{31}P	1/2	1.0839×10^{8}	100.00
^{113}Cd	1/2	-5.961×10^{7}	12.22

Cavanagh, J., Fairbrother, W.J., Palmer, A.G.III, Rance, M., Skelton, N.J. Protein NMR Spectroscopy: Principles and Practice, 2nd edition, 2007, Academic Press

Nucleus with spin ½ in external magnetic field B_0 along z



Two possible states:

$$I_z = \hbar m$$

magnetic quantum number

$m = -I, -I+1, \ldots, I$

for $I = \frac{1}{2}$: $m = -\frac{1}{2}, +\frac{1}{2}$

Energy of a magnetic dipole in external magnetic field

$$\boldsymbol{\mu} = \gamma \boldsymbol{I}$$

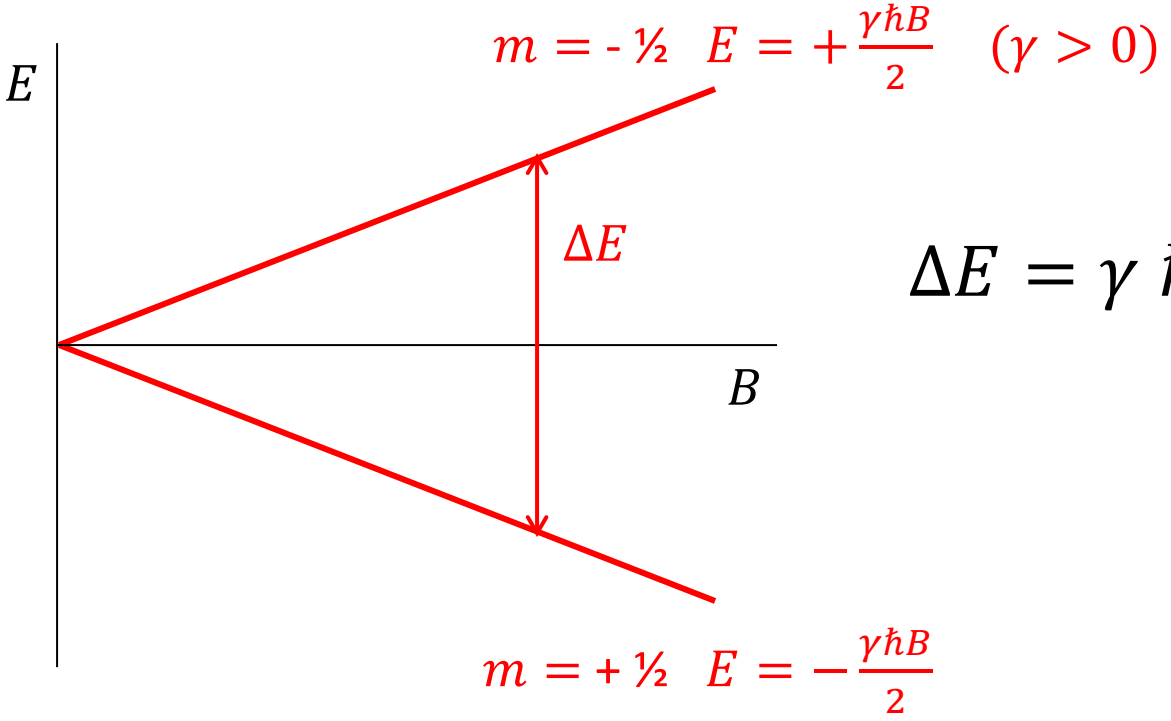
$$\mu_z = \gamma I_z = \gamma \hbar m$$

$$E = -\boldsymbol{\mu} \cdot \boldsymbol{B} = -\mu_z B = -\gamma \hbar m B$$

Energy of a magnetic dipole in external magnetic field

$$\boldsymbol{\mu} = \gamma \boldsymbol{I}$$
$$\mu_z = \gamma I_z = \gamma \hbar m$$

$$E = -\boldsymbol{\mu} \cdot \boldsymbol{B} = -\mu_z B = -\gamma \hbar m B$$



$$\Delta E = \gamma \hbar \Delta m B = \gamma \hbar B$$

$$\Delta E = \gamma\, \hbar\, \Delta m\, B = \gamma\, \hbar\, B$$ Energy difference between spin states

$$E = h\nu = \frac{h}{2\pi} 2\pi\nu = \hbar\omega$$ Energy of a photon

frequency in s^{-1}=Hz

$\omega = 2\pi\nu$
angular velocity (angular frequency)
in radians/s

$$\hbar\, \omega = \gamma\, \hbar\, B$$

$$\boxed{\omega = \gamma B}$$ Larmor frequency

$$\nu = \frac{\omega}{2\pi} = \frac{\gamma B}{2\pi}$$

ΔE (and ω) can be tuned as a function of B
ΔE is rather small $\rightarrow$ ω is radiofrequency

$\nu(^1\text{H})$=500 MHz: $B = \frac{2\pi\nu}{\gamma} = \frac{(2\pi)(500*10^6 s^{-1})}{2.6752*10^8 T^{-1}s^{-1}} = 11.74\ T$

Earth magnetic field ~ 30-60 µT

Populations (Boltzmann distribution) of nuclei with $m = \pm\frac{1}{2}$

Taylor expansion

$$\frac{N_m}{N} = \frac{e^{-\frac{E_m}{k_B T}}}{\sum_m e^{-\frac{E_m}{k_B T}}} = \frac{e^{\frac{\gamma\hbar mB}{k_B T}}}{\sum_m e^{\frac{\gamma\hbar mB}{k_B T}}} \approx \frac{1+\frac{\gamma\hbar mB}{k_B T}}{\sum_m\left(1+\frac{\gamma\hbar mB}{k_B T}\right)} = \frac{1+\frac{\gamma\hbar mB}{k_B T}}{2} = \frac{1}{2} \pm \frac{\gamma\hbar B}{4k_B T}$$

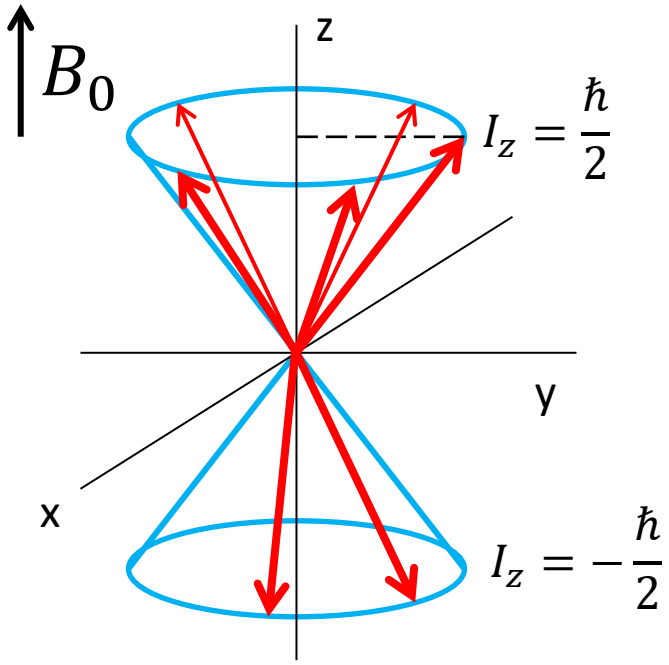
Boltzmann constant
$k_B = 1.381 * 10^{-23} JK^{-1}$

$\sim 2 * 10^{-5}$
for $\nu(^1\text{H})$=500 MHz and 25°C

^{1}H at 500 MHz and 25°C

m = - ½: population≈ 0.49998 (49.998 %)

m = ½: population≈ 0.50002 (50.002 %)



- Population difference → sample has bulk magnetic dipolar moment along z
- Population difference is very small → the magnetic moment is very small → low sensitivity
- Population difference $\approx \frac{\gamma \hbar B}{2k_B T}$ (proportional to γ and B)

Chemical shift

- Electrons surrounding the nucleus move under the influence of the applied magnetic field
- This movement of electrons generates additional magnetic field that usually opposes the externally applied magnetic field ("shielding")

nuclear shielding

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

external magnetic field

actual magnetic field perceived at the nucleus

$$\omega = \gamma B = \gamma(1 - \sigma)\, B_0$$

Larmor frequency (ω) of each nucleus is unique and very sensitive to local electronic environment

- Covalent bonds
- Conformation
- 3D structure, in particular ring currents, charges, hydrogen bonding...

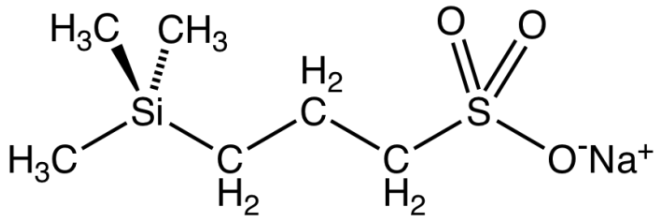
Chemical shift (continued)

$$\omega = \gamma(1-\sigma)\, B_0$$

$$\delta = \frac{\nu - \nu_0}{\nu_0} * 10^6$$

chemical shift
units: ppm (parts per million)

Reference frequency
= frequency of a reference compound:
tetramethylsilane (TMS) in organic solvents
trimethylsilylpropanesulfonate = 4,4-dimethyl-4-silapentane-1-sulfonate (DSS) in aqueous solutions



$\nu(^1\text{H})$=500 MHz (11.74T): 1 ppm $= \frac{500\ MHz}{10^6} = \frac{500 * 10^6\ Hz}{10^6} = 500\ Hz$

$\nu(^1\text{H})$=600 MHz (14.09T): 1 ppm $= \frac{600\ MHz}{10^6} = \frac{600 * 10^6\ Hz}{10^6} = 600\ Hz$

J-coupling (scalar coupling)

Interaction between nuclear spins mediated by electrons that form the bond(s) between the nuclei

- occurs over 1-4 bonds
- 3-bond J-coupling depends on conformation (dihedral angle)

$$E = \hbar\omega_I m_I + \hbar\omega_S m_S + 2\pi\hbar J_{IS} m_I m_S$$

J_{IS}: scalar coupling constant [Hz]

$$m_S = +\frac{1}{2}: \Delta E_I = \hbar\omega_I + 2\pi\hbar J_{IS}\left(\frac{1}{2}\right) = \hbar\omega_I + \pi\hbar J_{IS}$$

$$m_S = -\frac{1}{2}: \Delta E_I = \hbar\omega_I + 2\pi\hbar J_{IS}\left(-\frac{1}{2}\right) = \hbar\omega_I - \pi\hbar J_{IS}$$

ΔE_I depends on spin S: $\pm\pi\hbar J_{IS} = \pm\frac{1}{2} h J_{IS}$

J-coupling (scalar coupling)

$$E = \hbar\omega_I m_I + \hbar\omega_S m_S + 2\pi\hbar J_{IS} m_I m_S$$

scalar coupling constant

$$m_I = +\frac{1}{2}: \Delta E_S = \hbar\omega_S + 2\pi\hbar J_{IS}\left(\frac{1}{2}\right) = \hbar\omega_S + \pi\hbar J_{IS}$$

$$m_I = -\frac{1}{2}: \Delta E_S = \hbar\omega_S + 2\pi\hbar J_{IS}\left(-\frac{1}{2}\right) = \hbar\omega_S - \pi\hbar J_{IS}$$

ΔE_S depends on spin I: $\pm\pi\hbar J_{IS}$

Same splitting as for ΔE_I

Does not depend on magnetic field!

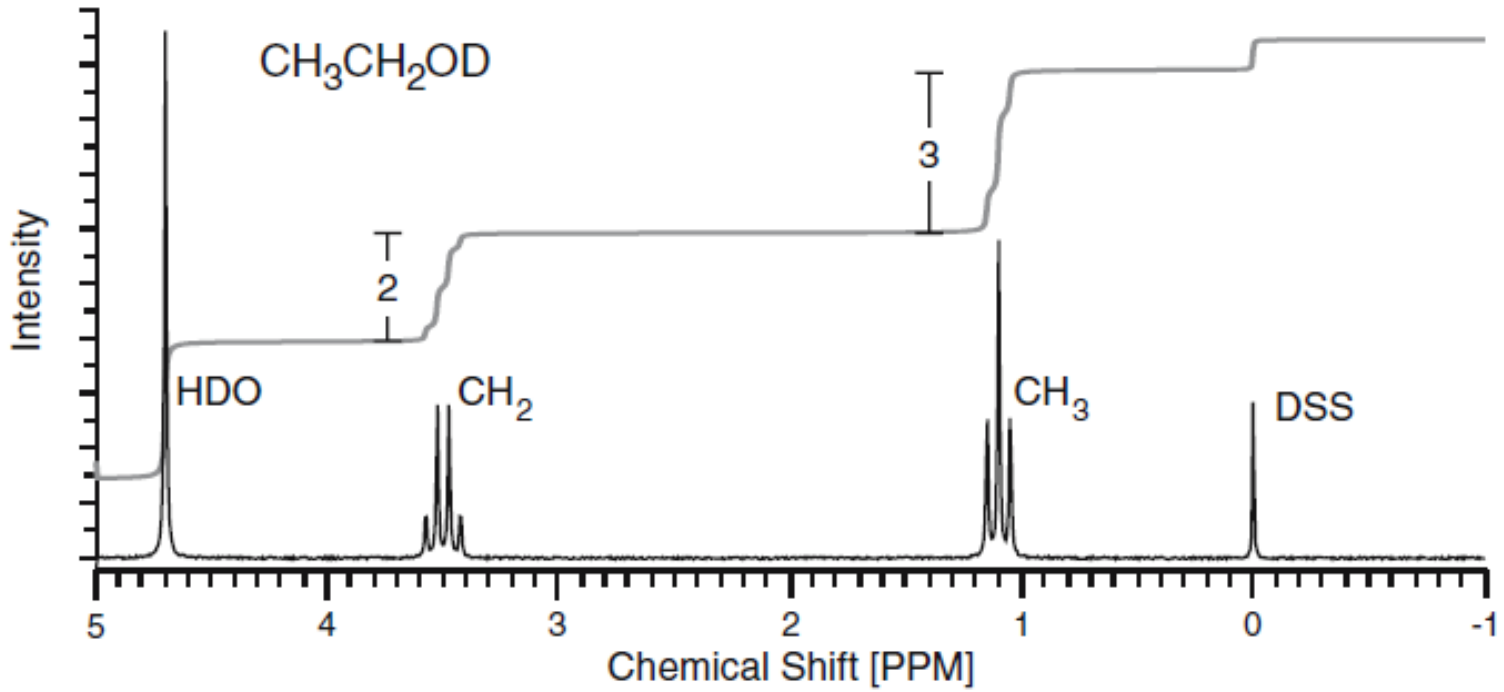
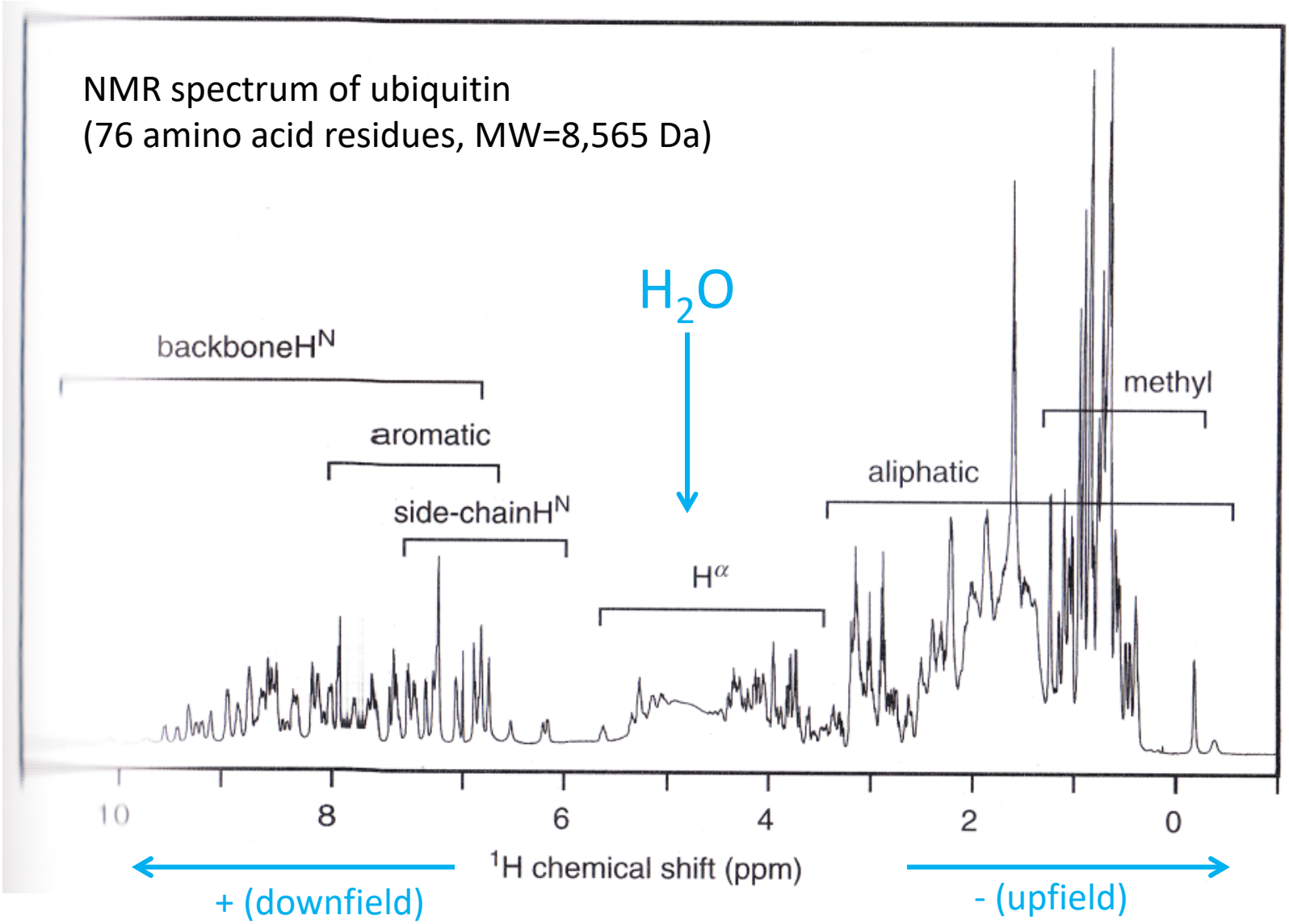


Figure 2.6. NMR spectrum of ethanol. The proton NMR spectrum of ethanol, in D_2O, is shown. The chemical shifts are approximate and are for illustrative purposes. The methyl protons give rise to the peak at 1.1 ppm and the methylene protons give rise to the peak at 3.5 ppm. The hydroxyl proton is absent in this example due to exchange with the solvent. The line at 4.7 ppm arises from residual protons in the solvent. The line at zero ppm is from the methyl groups of the reference compound (DSS, see Chapter 3). The gray line is the integral of the spectrum; the CH_2 and CH_3 peaks have a relative area of 2:3.

Rule, G.S. and Hitchens, T.K. Fundamentals of Protein NMR Spectroscopy, 2006, Springer



Cavanagh, J., Fairbrother, W.J., Palmer, A.G.III, Rance, M., Skelton, N.J. Protein NMR Spectroscopy: Principles and Practice, 2nd edition, 2007, Academic Press