Introduction to NMR Spectroscopy: Structure, Dynamics and Molecular Interactions

Trevor Rutherford
tjr@mrc-lmb.cam.ac.uk

NMR spectroscopy is a powerful method for resolving the structure, dynamics and interactions of proteins and nucleic acids in solution. The LMB has a state-of-the-art NMR facility with a mission to support programmes of research into the structure of proteins, characterisation of protein complexes, membrane and intrinsically disordered proteins and protein-protein interactions. We focus on understanding how signalling and ribosome maturation pathways work.

The LMB NMR facility was relocated to a purpose-built Magnetic Resonance Spectroscopy (MRS) building in 2001 (see: where to find us). The NMR facility contributes to projects across the full range of research activities at the LMB and is complementary to other structural and biophysical techniques. It is a biology-driven facility that makes NMR accessible to the non-specialist user. All relevant advanced NMR methods have been implemented and our support staff are able to adapt them to the specific needs of the LMB scientists.

Our website contains many useful resources including protein expression protocols, software pointers and in-house scripts for use in your NMR project. The contact for any enquiries about the use of the NMR facility and NMR support is Stefan Heinzel, ext 3074 or 3077.

Click on an image below to learn a little more about NMR.
Molecular Interactions

Wnt enhanceosome multiprotein complex

Characterisation of mechanistic principles:
• direct binding interactions
• contact surfaces
• stable or transient complexes

**Overview**

**Where NMR is good:**

<table>
<thead>
<tr>
<th>STRUCTURE</th>
<th>composition, conformation, complexation</th>
</tr>
</thead>
<tbody>
<tr>
<td>DYNAMICS</td>
<td>internal mobility, diffusion (MW), exchange</td>
</tr>
<tr>
<td>SITE SPECIFICITY</td>
<td>define the locus of changes &amp; binding interfaces</td>
</tr>
<tr>
<td>QUANTIFICATION</td>
<td>relative concentrations of materials in the same sample (kinetics, titration, $K_{eq}$)</td>
</tr>
</tbody>
</table>

**Where NMR is ‘less good’:**

<table>
<thead>
<tr>
<th>SYMMETRY</th>
<th>cannot resolve planes of symmetry (can the symmetry be broken ?)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HETEROGENEITY</td>
<td>requires &quot;pure&quot;, monodisperse, single conformers (?)</td>
</tr>
<tr>
<td>SAMPLE QUANTITY</td>
<td>requires minimum 0.28 ml of &gt;20 µM fixed volume, therefore there is a minimum concentration</td>
</tr>
<tr>
<td>HIGH MW</td>
<td>struggles with slowly rotating rigid molecules</td>
</tr>
<tr>
<td>LOW SPIN DENSITY</td>
<td>requires many ‘coupled’ NMR active atomic nuclei</td>
</tr>
<tr>
<td>SPEED</td>
<td>often time-consuming — data acquisition is getting faster but data analysis not yet reliably automated</td>
</tr>
</tbody>
</table>
DYNAMICS — NMR Sees All States in the Ensemble

Solid State methods report on individual frames; multiple states can be resolvable.

Solution State methods report all states in the ensemble — but structural data (NOEs) may be missing from flexible regions.

- NMR in combination with computational methods
- Quantify both the extent and rates of motion
- Interpreting NMR data in terms of a single structure can result in an erroneous virtual conformation

NMR Spectra are Sensitive to Motion Across a Wide Span of Timescales

- vibration
- side-chain rotation
- segmental motion
- ligand binding
- allosteric regulation
- protein folding
- relaxation ($T_1$, $T_2$, Het-NOE)
- $T_{1\rho}$
- ZZ exchange
- H-D exchange
- residual dipolar coupling
- relaxation dispersion (CPMG)
- EXSY
atomic nucleus spinning in a precessional orbit

\[ \Delta E = \frac{h}{\nu} = \frac{h \gamma B_0}{2\pi} \]

Nuclear magnetic dipole

accessible isotopes for biomolecular NMR

\[ ^1H \quad ^{13}C^* \quad ^{15}N^* \quad ^2H \quad ^{31}P \]

\[ ^{17}O \quad ^{33}S \quad ^{19}F \quad ^{23}Na \quad ^{7}Li \quad ^{67}Zn \quad ^{113}Cd \]

cryomagnet with cryoprobe

in the absence of other magnetic fields all nuclei of a given isotope resonate at identical frequency

\[ B_0 \]

\[ \text{spin } -\frac{1}{2} \]

\[ \text{spin } +\frac{1}{2} \]

NMR Spectrometer Field Strength Dependence

Advantages of increased \( B_0 \) field strengths:

- increased sensitivity \( I \propto B_0^{3/2} \)
- increased signal dispersion (resolution)
- altered relaxation (TROSY, NOESY)

Disadvantages:

- cost & housing
- radiofrequency power handling
- some lines broaden at high field (particularly carbonyl \(^{13}C\))

Put in the most simple terms, the stronger the magnet the better
Linewidth, $\lambda$, is inversely proportional to $T_2$

$\lambda = \text{linewidth at half height; } T_2 = \text{exponential time constant for signal decay; } A = \text{signal amplitude; } \omega = \text{frequency in rad/s}$
Relationship Between Relaxation Rates and Linewidth

\[ \lambda = \frac{1}{\pi T_2} \]

- FID for macromolecule, \( T_2 = 20 \) ms
  - Fourier Transformation
  - \( \lambda \approx 0.2 \) Hz

- FID for small molecule, \( T_2 = 1.6 \) s
  - Fourier Transformation
  - \( \lambda \approx \frac{1}{\pi} \) Hz

Fourier Transformation

- Time domain: "free induction decay" (FID)
- Frequency domain
**Fourier Transformation**

- Frequency domain
- Time domain
- "Free induction decay" (FID)

---

**Multi-Dimensional NMR: Reducing Peak Overlap**
Magnetic Interactions Between a Nucleus and its Environment

\[ H = \sum_j \gamma B_0 (1 - \sigma_j) I_j^z + \sum_{jk} 2\pi J_{jk} I_j^z I_k^z + \sum_{jk} 2\pi D_{jk} I_j^z I_k^z \]

1. Shielding

- Shielding (chemical shift)
- J-coupling
- Dipolar coupling

\[ \text{chemical shift axis} \]
Magnetic Interactions Between a Nucleus and its Environment

\[ H = \sum_j \gamma_H B_0 (1 - \sigma) I_z^j + \sum_{j<k} 2 \pi J_{jk} I_z^j I_z^k + \sum_{j<k} 2 \pi D_{jk} I_z^j I_z^k \]

2. **J-Coupling** (scalar coupling through covalent bonds)
Magnetic Interactions Between a Nucleus and its Environment

\[ H = \sum_j \gamma_j B_0 (1 - \sigma_j) I_j + \sum_{jk} 2 \pi J_{jk} I^j I^k + \sum_{jk} 2 \pi D_{jk} I^j I^k \]

3. Dipole-Dipole Coupling (through space, up to 7Å)

Dipolar Coupling in Structure Determination

\[ D = D_{\text{max}} \left( \frac{1}{2} (3 \cos^2 \theta - 1) \right) \]

- **Solid state NMR**: resolving all dipolar couplings in all possible orientations loses the fine detail in the spectrum
- **NOE \( \propto \frac{1}{r^6} \)**: nuclear Overhauser effect — measure of inter-nuclear distance
- **rapid averaging**: 'residual' dipolar coupling — measure of bond orientation
- **incomplete averaging**: no averaging over time:
NOE-based Structural Analysis of Proteins
(Kurt Wüthrich, 1980’s; Nobel Prize 2002)

- structure determined from short-range H-H interactions (<6Å — i.e. within 4 carbon-carbon bond lengths)
- apply typically 25 restraints per residue in combined distance-geometry & simulated annealing

NOESY: a complex jigsaw puzzle

- Main structural information
  Intensity $\propto \frac{1}{r^6}$ (< 6.0 Å)
- Very complex to analyse
  Requires proton assignments
  Highly overlapped
  (requires 3D or 4D NMR)
  Spin diffusion
- Quality degrades as MW increases

Realistic MW limit for high resolution structure $\approx$ 40 kDa
Residual Dipolar Coupling
(N.Tjandra & A.Bax, Science, 278, 1111, 1997)

Sample with 1 part in \(-10^3\) excess population of its major diffusion axis aligned with \(B_0\) recovers a few Hertz of ‘residual’ dipolar coupling; the magnitude is sensitive to bond orientation.

RDC for Intrinsically Disordered Protein Segments

\[ E = \frac{1}{2} k (x - x_0)^2 \]

Add experimental restraints from NMR data

**Experimentally Derived Structural Restraints**

- **NOE (Nuclear Overhauser effect)**
  - structure determined from short-range H-H interactions (<6 Å)

- **Residual Dipolar Coupling**
  - bond orientation (WRT global coordinate frame)

- **Chemical Shift Index (TALOS)**
  - bond dihedral angles

- **J-coupling constants**
  - bond dihedral angles

- **Hydrogen bonds**
  - from hydrogen exchange protection factors

Energy Coordinate

NMR Structure PDB Submissions
superposition of 20-100 models depicting the consensus structure

Characteristics of the $^1H$-NMR Spectrum of Proteins

$^1H$ chemical shift (ppm)
**HSQC: the Protein ‘Fingerprint’**


15N-labelled wild-type p53 core domain, 219 residues

probe the structural environment at >200 individual sites

**Mapping Protein Binding Interfaces**

High Ambiguity Driven Docking “HADDOCK”

35 $^1$H($^{12}$C)-$^1$H($^{13}$C) intermolecular NOEs
perturbed chemical shift *and* >50% solvent exposed

Half-filtered NOESY
800 MHz $^1$H, 150 ms NOE mix
$\tau_1$ $^1$H($^{12}$C) $\tau_2$ $^1$H($^{13}$C)

Quantifying Binding Affinities

\[ \Delta_{\text{obs}} = \frac{[L_T] + [E_T] + K_d}{2} \]

\[ \Delta = \Delta_{\text{max}} \frac{[L_T] + [E_T] + K_d}{2} - 4 \frac{[L_T][E_T]}{[E_T]^2} \]

\( \Delta \) is the chemical shift perturbation, \([L_T]\) & \([E_T]\) are the total concentration of ligand & protein

Molecular Weight Limit for NMR?

\( ^{15} \text{N}-\text{labelled p53 (393 aa homotetramer)} \)

175 kDa complex

\( ^{15} \text{N}-\text{labelled p53 + unlabelled MDMX} \)

255 kDa complex

Molecular Weight Discrimination Using Relaxation Filters

(I.D. Campbell, C.M. Dobson, R.J.P. Williams & P.E. Wright, FEBS Lett., 57, 96, 1975)

FID for macromolecule, $T_2 = 20$ ms

FID for small molecule, $T_2 = 1.6$ s

$^1$H NMR of intact human erythrocytes

John O’Neill, LMB Cell Biology Division
**Molecular Weight Discrimination Using Relaxation Filters**

(J.D. Campbell, C.M. Dobson, R.J.F. Williams & F.E. Wright, *FEBS Lett.*, 57, 96, 1975)

- FID for macromolecule, $T_2 = 20$ ms
- FID for small molecule, $T_2 = 1.6$ s

- $^1$H NMR of intact human erythrocytes
- Relaxation filtered (CPMG)

---

**Effective Molecular Weight**

(J. Anglister, S. Grzesiek, H. Ren, C.B. Klee & A. Bax, *J. Biomol. NMR*, 3, 121, 1993)

- $T_2$ is obtained from the NH region peak intensity ratio between two spectra acquired with different values of the delay time $\Delta$.

\[
T_2 = \frac{2(\Delta_A - \Delta_B)}{\ln(I_A / I_B)}
\]

- $\tau_c \approx \frac{1}{5T_2} \approx \frac{1}{2} \text{MW} \ @ 20^\circ C$

---

For rotational correlation time $\tau$, in ps, relaxation time constant $T_2$, in $s$.

---

John O’Neill, LMB Cell Biology Division
Is My Protein Folded or Unfolded?

1H chemical shift (ppm)

N chemical shift (ppm)

Is My Protein Folded or Unfolded?
Optimising Domain Boundaries for Crystallisation or NMR

Quantifying Individual pKₐ Values

High-Throughput NMR Applications

Design of biochemical probes & measurement of binding affinities for:
- dissecting function of proteins
- target validation of protein drug targets
- (hit-to-lead development)

- chemical fragment libraries (1° screening)
- hits from HTS / in silico (2° screening)
- systematic exploration of conditions for synthesis

NMR as a Screening Tool in Drug Discovery & Design

Ligand-Observed NMR Assays
- Yes/No binding
- can assay as competition binding to a known site
- requires ~1 - 10 µM unlabelled protein
- no MW restriction for the protein

Basis
- change in diffusion rate (CPMG)
- transfer of energy from an irradiated protein (STD)
- nuclear Overhauser effect (waterLOGSY)
- chemical shift perturbation of 19F resonances

Protein-Observed NMR Assays
- Yes/No & characterise the binding site
- "gold standard" proof of binding
- requires 15N or 13C-labelled protein, >50 µM
- MW <50 kDa

Basis
- chemical shift perturbation (HSQC, methyl-TROSY)
Ligand Observed Methods for Fragment Based Screening

Summary

- **STRUCTURE**: NOE, RDC, chemical shifts
- **DYNAMICS**: internal mobility, diffusion & exchange
- **INTERACTIONS**: protein-protein, protein-DNA, protein-small molecule
- **QUANTIFICATION**: of reaction kinetics and $K_{eq}$, approx. MW, exchange rates
## Recommended NMR Texts

|--------------------------------------------------|----------------------------------------------------------------------------------|