Biological applications of NMR spectroscopy

Tatiana Kutateladze

Department of Pharmacology
University of Colorado Denver
Anschutz Medical Campus

Suggested texts
- Protein NMR Spectroscopy: Principle and practice - J. Cavanagh, W. Fairbrother, A. Palmer III and N. Skelton
- NMR of Proteins and Nucleic Acids - K. Wüthrich
- Biomolecular NMR Spectroscopy - J. Evans
- NMR of Proteins - edited by G. Clore and A. Gronenborn
  - Introduction to Protein Structure - C. Branden and J. Tooze
  - Spin Dynamics - M. Levitt
  - Modern NMR Techniques for Chemistry Research - A. Derome
  - A complete introduction to modern NMR spectroscopy - R. Macomber

Structure-Based Drug Design is the major objective of structural biology

- the use of atomic resolution structural information about a macromolecule typically obtained by X-ray crystallography or NMR spectroscopy
- the 3D structure of a target protein bound to a lead compound allows the atom-by-atom modification and improvement of drug leads
- the process is iterative, with successive cycles of design, synthesis and testing that refine the properties of the drug lead

The structural biology methods
- NMR spectroscopy
- X-ray crystallography
- Electron microscopy
- Mass spectrometry
- CD spectroscopy
- IR
- UV
- EPR
- Molecular modeling and structure prediction

Protein Data Bank: www.rcsb.org/pdb/index.html
Research Collaboration for Structural Bioinformatics - Rutgers U, UC San Diego, NIST
Total 80,830 structures (17 May 2008)
41,013 structures (9 January 2007)
18,359 structures (1 August 2002)
3,298 structures in 2001
~90% - crystal; ~9% - solution structures

BioMagResBank: www.bmrb.wisc.edu/Welcome.html
University of Wisconsin-Madison
3,940 – proteins/peptides chemical shifts
123 – DNA coupling constants
81 - RNA NMR software library
pulse sequence library
reference information

The structure-based drug design cycle
NMR applications and limitations

- In solution - native state
- Kinetics
- Folding
- Intermolecular interactions
- Chemical characterization
- Detection of impurities
- Time consuming
- Complex
- Low sensitivity (requires 10 mg of a protein)
- Solubility (mM concentrations)
- Size limit ~ 30 kDa
- Isotope labeling (13C, 15N, 2H) is required for large proteins but can be difficult and expensive

NMR sample

- 550 µl of 0.1-2 mM unlabeled, 15N- or 13C-isotope labeled protein
- Deuterated or phosphate buffer
- pH 2-7
- At least 95% purity
- Not aggregated
- Stable for several days/weeks
- d20-DTT (Cys)
- NaH2PO4

NMR experiments

Identification of amino acid spin system and chemical shift assignment

- 2D heteronuclear TOCSY
- 2D homonuclear COSY
- 2D HBCGCDCEHE TOCSY
- 2D HBCGCDHD TOCSY
- 2D homonuclear NOESY
- 4D 1313C--edited NOESY--HSQC
- 1515N,N,1313C--edited NOESY--HSQC
- 1515N--edited HSQC
- 1313C--edited HSQC

Sequential assignment of amino acids

- 2D HBCGCDCEHE TOCSY
- 2D HBCGCDHD TOCSY
- 2D homonuclear NOESY
- 4D 1313C--edited NOESY--HSQC
- 1515N,N,1313C--edited NOESY--HSQC
- 1515N--edited HSQC
- 1313C--edited HSQC

Total experiment time 10³ hrs

What do we start with?

1. Instrument

900 MHz NMR spectrometer

Strategy to determine the protein structure by NMR

1. Protein purification and optimization of the conditions
   2. Amino acid analysis, secondary structure by CD and 1H NMR, integrity and purity by mass spectrometry
   3. Oligomeric state by NMR and sedimentation ultracentrifugation
   4. Collection of NMR spectra
   5. Analysis of the NMR data:
      - Sequence specific resonance assignment
      - NOEs (distance restraints)
      - Coupling constants (torsion angle restraints)
      - Exchange rate
      - Secondary structure
   6. Structure calculation and refinement
   7. Dynamics
   8. Functional assays

NMR experiments

COSY: correlation spectroscopy, provides through bond connections for protons separated by 2 or 3 bonds

TOCSY: total correlation spectroscopy, provides through bond connections for all protons in an amino acid side chain

NOESY: nuclear Overhauser effect spectroscopy, provides through space connections between protons that are close in the space

HSQC: heteronuclear single quantum coherence, provides connections between a proton and a directly attached 15N or 13C nuclei

4.5-5.0 cm
NMR parameters

Chemical shift: \( \delta \) (corresponds to the resonance frequency of a nucleus; related to the chemical environment of the nucleus; affected by aromatic ring currents, solvent exchange and electronegativity of nearby nuclei; units: ppm parts per million)

Spin-spin coupling constant: J (characterizes through-bond (scalar) interactions; related to the bond angles; causes NMR lines of coupled nuclei to split; units: Hz)

Nuclear Overhauser effect: NOE (caused by through-space (dipolar) interactions between nuclei separated in space by less than 5 Å; detected by changes in NMR line intensity of one nucleus when another is irradiated; correlated by the inverse sixth power of the distance between the nuclei)

Relaxation parameters: linewidth, \( T_1 \) and \( T_2 \) (longitudinal and transverse relaxation times)

Scalar spin-spin coupling connectivities in TOCSY (total correlation spectroscopy)

Identification of \(^1H\) spin systems in aromatic residues

Protein structure determination using 3D and 4D spectra of \(^15N/\(^13C\) labeled proteins

Sequential assignment of amino acids using triple-resonance HNCA, HNCO/CO/NN = H/\(^15N\) HSQC

Identification of spin systems and complete chemical shift assignment intraresidue: 3D \(^15N\)- and \(^13C\)-edited TOCSY, Interresidue: 3D HNC and CCC TOCSY, HNCO

NOEs: 3D \(^15N\)- and \(^13C\)-edited NOESY

\( ^3J_{HH} \) coupling constants: 3D HNHA, HMOC-J

Structure calculation
Sequential resonance assignment

![Image of sequential resonance assignment](image1.png)

The secondary structure determination and collection of the $d_{\text{AX}}$ restraints in 3D $^{15}$N-edited NOESY

![Image of secondary structure determination](image2.png)

Computational approach to determine the structure of a protein

Input:
1. known sequence of the amino acid residues
2. known molecular structure of the residues, including bond distances and angles
3. known planar structure of the peptide group
4. NMR-derived restraints:
   - Interproton distances (NOEs) weak, medium and strong intensity
   - Intra- and interresidue bond angles $\Phi$, $\Psi$, $\chi$
   - Hydrogen bonds N-O and H-O distances
   - Heteroatom distance and angle restraints Ca$^2+$, Mg$^2+$, Zn$^2+$, etc.

Calculation and refinement of the 3D structure

Input
- Distance geometry
- Restrained molecular dynamics simulated annealing

Minimization of system energy: bond length and angles, van der Waals
Minimization of violations: NOEs, dihedral angles
Several cycles - calculation-analysis of violated restraints

Output: Structural statistics

A set of 20 lowest energy structures Average structure
Root Mean Square Deviation (RMSD) – a measure of the quality of the NMR structure

\[ \text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (r_i - r'_i)^2} \]

The number of restraints affects both precision and accuracy

8 – 15 restraints/res

Beyond the structure

- qualitative and quantitative binding assays: changes in NMR parameters allow us to determine a ligand's binding site and to estimate binding affinities
- rapid screening of lead compounds (SHAPES and other libraries)
- dynamic regions: motions of the protein backbone
- conformational changes
- pre-screening of structural targets for X-ray and NMR
- structures of proteins that do not crystallize
- structures of multi-domain proteins in solution
- determination of small protein structures (cryogenic probes, high field magnets, residual dipolar couplings, TROSY, micelles, automatic resonance assignment and NOE peaks picking)

NMR-based methods to monitor protein-ligand interactions

- chemical shift perturbations
- intermolecular NOEs
- \( T_1/T_2 \) relaxation

Mapping the binding site residues based on the chemical shift perturbations

Mapping the binding site of a ligand

The HSQC titration experiments provide quick information on the affinity

advantages: fast, low protein concentration, even weak binding, no false positive
NMR-based methods to monitor protein-ligand interactions

- chemical shift perturbations
- intermolecular NOEs
- $T_1/T_2$ relaxation changes

Labeling with stable isotopes, such as $^{15}$N, $^{13}$C and/or $^2$H together with 'isotope editing' experiments allows us to select signals.

Intermolecular NOEs – interproton distances between two molecules

Structure → Molecular mechanism → Function

Structural basis of signaling by F-binding domains