Oct 2002

Introduction to biomolecular NMR spectroscopy

Michael Sattler, Structural & Computational Biology EMBL Heidelberg

Contents

Introduction

Nuclear magnetic resonance (NMR) spectroscopy is next to X-ray crystallography the only biophysical method which can provide **high-resolution structures of biological molecules** such as proteins and nucleic acids and their complexes at atomic resolution. The first NMR derived three-dimensional **solution structure** of a small protein was determined in 1985, which means that NMR is about 25 years younger than biomolecular X-ray crystallography. Since then major improvements in NMR hardware (magnetic field strength, cryoprobes) and NMR methodology, combined with the availability of molecular biology and biochemical methods for preparation and isotope labeling of recombinant proteins have dramatically increased the use of NMR for the characterization of structure and dynamics of biological molecules in solution. These improvements are on-going and are designed to overcome the two main problems with NMR of biomacromolecules, namely signal-to-noise and spectral overlap. Importantly, biomolecular NMR spectrocopy can provide information about **conformational dynamics** and **exchange processes** of biomolecules at timescales ranging from picoseconds to seconds, and is very efficient in determining **ligand binding** and **mapping interaction surfaces** of protein/ligand complexes.

The nuclei of naturally occuring atomic isotopes that constitute biological molecules either have a nuclear spin at natural abundance (^{1}H, ^{31}P) or naturally less-abundant isotopes with a nuclear spin (^{13}C, ^{15}N) can be incorporated into biomolecules through **isotope labeling**. The nuclear spin is associated with a magnetic moment which is required for obtaining nuclear magnetic resonance and defines the basic resonance frequency, i.e. 600/150 MHz for ^{1}H/^{13}C at 14.1 Tesla. Each spin in a molecule gives rise to a nuclear magnetic resonance line. The exact resonance frequency depends on the chemical environment of each spin, such that for example the NMR spectrum of a protein will show NMR signals with slightly different frequencies. These differences are called **chemical shifts**. The first step of a structure determination by NMR consists of **assigning** the **chemical shifts** of all the atoms/spins of the molecule which are observed in an NMR spectrum. Once the NMR signals are assigned, experimental parameters which define the three-dimensional structure are measured. The most important structural information derived from NMR spectra is based on the **nuclear Overhauser effect (NOE)**, which is a result of cross-talk between different spins (normally protons) in a molecule and depends on the *through-space* distance between these spins. NOEs are typically only observed between protons which are separated by less than 5-6 Å. **J-coupling constants**, which are mediated through chemical bonds provide information about dihedral angles, and thus can define the peptide backbone and side chain conformations. Recently, new NMR parameters, such as **residual dipolar couplings (RDCs)** and **cross-correlated relaxation effects**

(CCRs) have been shown to provide distance independent projection angles for bond vectors, e.g. N–H and C^{α}–H^{α} bonds in proteins. Residual dipolar couplings are measured in anisotropic solution, e.g. dilute (3%) liquid crysalline media. Especially, RDCs are extremely useful for defining the relative orientation of two domains of a protein (NOEs between separate domains are often not observed due to the 5 Å upper distance limit for observation of an NOE).

Nowadays, three-dimensional structures can be obtained for proteins up to 50 kDa molecular weight, and NMR spectra can be recorded for molecules well above 100 kDa. In the following, the use and applications of biomolecular NMR in structural biology are introduced, basic principles and observables of biomolecular liquid-state NMR are described and the practical aspects are discussed to provide an overview about the utility but also limitations of NMR spectroscopy for structure/function studies of biomolecules.

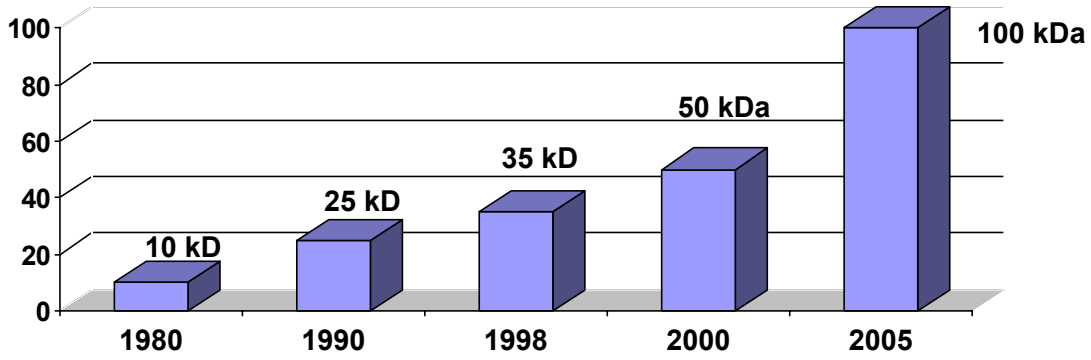
History

1946	**Bloch, Purcell**	first nuclear magnetic resonance
1955	**Solomon**	NOE (nuclear Overhauser effect)
1966	**Ernst, Anderson**	Fourier transform NMR
1975	**Jeener, Ernst**	2D NMR
1985	**Wüthrich**	first solution structure of a small protein (BPTI) from NOE derived distance restraints
1987	3D NMR + ^{13}C, ^{15}N isotope labeling of recombinant proteins (resolution)	
1990	pulsed field gradients (artifact suppression)	
1996/7	new *long range* structural parameters:	
	- **residual dipolar couplings** from partial alignment in liquid crystalline media	
	- projection angle restraints from **cross-correlated relaxation**	
	TROSY (molecular weight > 100 kDa)	

Nobel prizes

1944	*Physics*	Rabi (Columbia)
1952	*Physics*	Bloch (Stanford), Purcell (Harvard)
1991	*Chemistry*	Ernst (ETH)
2002	*Chemistry*	Wüthrich (ETH)

Methodological developments for structure determination by NMR



□

magnets	500 MHz	600 MHz	800 MHz	900 MHz	1 GHz
cryoprobes				500,600	700,800
NMR	2D	3D/4D	TROSY		
		triple resonance			
		pulsed field gradients			
			RDCs, CCR		
Labeling	no labeling	$^{13}C/^{15}N$	$^{2}H/^{13}C/^{15}N$	segmental labeling	

Important improvements:

1. Magnets — S/N > 2400:1 @ 900 MHz
   spectral resolution
   shielded magnets (up to 700 MHz) => no extra building needed
2. Cryoprobes — S/N (> 4000:1) @ 600 MHz
3. Methods — TROSY
   Cross-correlated relaxation (CCR)
   Residual dipolar couplings (RDCs)
4. Labeling — selective ^{1}H-labeling (uniformly $^{13}C/^{15}N/^{2}H$ & $^{1}H/^{13}C$-methyl)
   segmental labeling (domain 1: $^{13}C/^{15}N$, domain 2:unlabeled)

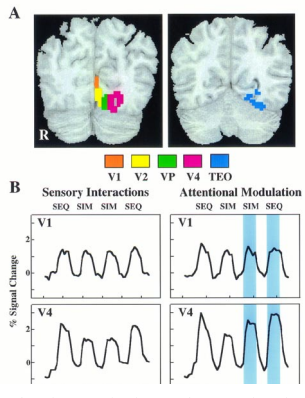
NMR in structural biology

What information can be obtained from NMR?

NMR has a wide range of applications, for example NMR of solids (crystalline or powders) is applied in inorganic chemistry and material sciences to characterize polymers. **Solid state NMR methods** are available for studying membrane proteins immobilized in lipid bilayers and other membrane mimics. New methods are being developed to determine 3D structures using similar methods as for structure determination in the liquid state. These developments are very important, since it is still difficult to crystallize membrane proteins, while no crystals are required for NMR.

Another important application is **Magnetic Resonance Imaging (MRI)** which allows *in vivo* imaging of human tissue. MRI based computer tomography is used to obtain slice images of a human body, for example in cancer diagnostics. An extension of this is **localized spectroscopy**, which allows to study metabolism in different tissues or organs *in vivo*. However, the predominant application of nuclear magnetic resonance in chemical and biological research is **liquid-state NMR**:

- *constitution* and *conformation* of organic molecules
- *3D structure of proteins, nucleic acids, sugars and their complexes in solution*
- *conformational dynamics and mobility of biomacromolecules in solution*
  - rotational correlation time τ_c, internal motion, order parameter S^2
  - mobility of peptide backbone in proteins
  - populations of amino acid side chain conformations
  - proline cis/trans isomerization
  - aromatic ring flips
  - sugar puckering, base pair dynamics in DNA, RNA
- *chemical exchange, conformational exchange*
  - hydrogen exchange rates → hydrogen bonding
  - base pair opening dynamics in DNA, RNA
  - k_D of acidic/basic side chains from pH titrations
- *enzyme mechanism and chemical reactions*
- *dynamics and lifetimes of bound water molecules*
- *partially structured biomacromolecules → folding studies*
- *rational drug design: SAR by NMR*



Mechanisms of Directed Attention in the Human Extrastriate Cortex as Revealed by Functional MRI.
Science (1998) ***282****, 108-111.*

NMR and X-ray crystallography are complementary

Apart from the high-resolution structure determination of biological molecules, biomolecular NMR provides unique information which can be complementary to crystallographic studies. Molecules are studied **in solution**. Therefore, crystallization is not required, and crystal packing affects may not influence the structure especially on the surface of a protein and even more so in nucleic acids where interaction with other molecules may occur. Solution studies should be closer to native-like conditions found in the cell. Since crystals are not needed, **protein folding** studies can be done by monitoring NMR spectra upon folding or denaturing of a protein in *real time.* More importantly, denatured states of a biomolecule, folding intermediates and even transition states can be characterized using NMR methods. NMR provides information about **conformational or chemical exchange, internal mobility and dynamics at** timescales ranging from picoseconds to seconds. For example, flexible loops which may not be visible in a crystal structure due to spatial averaging of electron density, are easily seen in NMR spectra. Even though they may not adopt a single conformation, and the mobility of such a flexible linker with respect to the rest of the molecule can be characterized. Averaging of side chain conformations can be measured and the populations of different conformations can be determined. NMR is very efficient in **mapping interactions with other molecules**, e.g. protein/protein, protein/nucleic acid, protein/ligand or nucleic acid/ligand interactions. Titration of a ligand that interacts with a molecule induces changes in its NMR spectrum for signals of atoms near the binding site. This **chemical shift mapping** therefore localizes the binding site with ligand and can be used to determine dissociation constants (k_D), especially in the μM and mM range. Chemical shift mapping requires very simple and fast (< 1h) NMR experiments and has numerous applications.

On the other hand, currently, the **upper weight limit for NMR structure determination** is ~ 50 kDa. Above this molecular weight, X-ray crystallography is currently the only method for high-resolution structure determination. Nevertheless, since NMR spectra can be obtained for molecules well above 100 kDa , molecular interactions can be mapped for complexes of that size using NMR in combination with a crystal structure. Also dynamical properties of such a molecule can be determined by NMR in conjunction with the crystal structure.

Principles of nuclear magnetic resonance

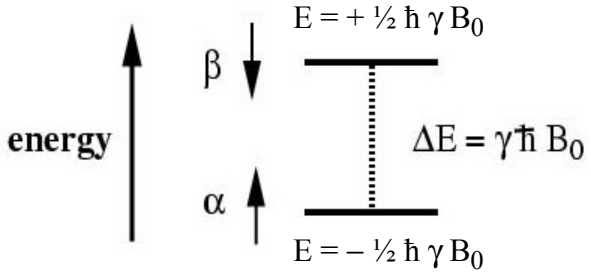
Most atoms which constitute biomolecules occur in isotopes which have a nuclear spin $\boldsymbol{I} > 0$. The nuclear spin gives rise to a magnetic dipole moment $\mu = \gamma \boldsymbol{I}$, which is depends on the atom type.

Nucleus	I	γ [T∗s]$^{-1}$	$\gamma_{relative}$	Natural abundance [%]
^{1}H	½	$2.675 * 10^{8}$	1	99.98
^{2}H	1	$4.107 * 10^{7}$	0.15	0.02
^{13}C	½	$6.728 * 10^{7}$	0.25	1.11
^{15}N	½	$-2.712 * 10^{7}$	0.1	0.36
^{31}P	½	$1.084 * 10^{8}$	0.41	100

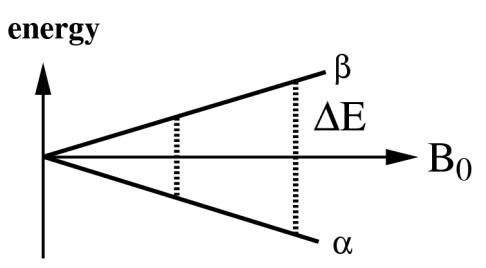
The energy associated with a magnetic dipole moment in a magnetic field of strength B_0 is given by $E = -\mu B_0 = -\gamma \boldsymbol{I} B_0$. The nuclear spin moments are quantitized, thus, only one of the three Cartesian coordinates, I_z can be specified simultaneously with the magnitude of the spin, $\boldsymbol{I}^2$. For a spin ½ nucleus, two $(2I + 1)$ orientations are possible ($I_z = \pm ½$), and the corresponding energy levels are populated according to a Boltzmann distribution.

Energy levels, populations & signal-to-noise

After inserting our NMR sample into the static magnetic field B_0 the spins in the ensemble of molecules in our NMR tube are either in the α- or β-state. That is the spins precess around the axis of the static magnetic field with a *z*-axis projection being parallel or antiparallel to B_0, respectively. The population of the α- and β-states according to a Boltzmann distribution is almost equal (~10^{-4} difference) since the energies involved are fairly small. Nevertheless the small population difference produces an effective magnetization along the *z*-axis. Note, however, that there is no magnetization observed in the *x,y*-plane since the individual precession frequencies are uncorrelated, e.g. out of phase such that the *x*- or *y*- components average to zero. In order to measure the precession frequencies the system is disturbed and brought into a non-equilibrium state which allows to monitor transverse (*x* or *y*) magnetization (see below).

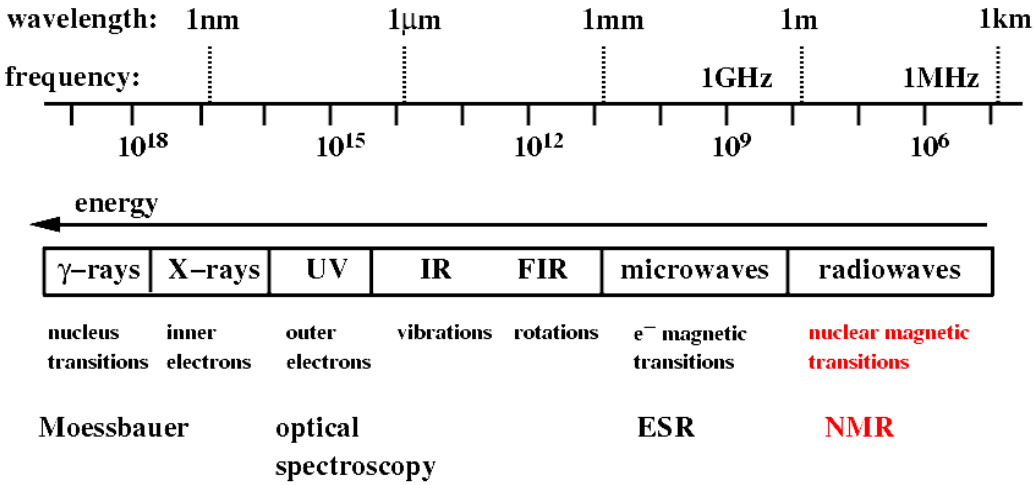


Boltzmann distribution: $\frac{N(\alpha)}{N(\beta)} \sim \frac{1.0001}{1}$



$N(\alpha/\beta) = \exp\{ \mu B_0/(k_B T) \} \approx 1 + \mu B_0/(k_B T)$

The Electromagnetic Spectrum



NMR involves resonance frequencies that correspond to radiofrequency wavelengths and is thus located at the low frequency end of the electromagnetic spectrum. The separation of the energy levels defines the population difference (according to Boltzmann statistics) which provides the magnetization that is measured in an NMR experiment. Since ΔE is very small NMR is a rather insensitive spectroscopic method, and optimizing the signal-to-noise ratio is always a critical issue for NMR. The signal-to-noise (S/N) can be increased by using more spins (higher sample concentration), but also by using higher strengths for the static magnetic field (B_0), and optimizing the design of NMR experiments, e.g. using spins with the highest γ (such ^1H) for excitation and detection.

$$S/N \sim N\, \gamma_{exc}\, \gamma_{det}^{3/2}\, B_0^{3/2}\, NS\, T_2^{1/2}$$

S/N	signal-to-noise
N	number of spins (sample concentration)
γ_{exc}	gyromagnetic ratio of excited spins
γ_{det}	gyromagnetic ratio of detected spins
B_0	static magnetic field strength (i.e. 600 MHz ^1H)
NS	number of scans
T_2	transverse relaxation time

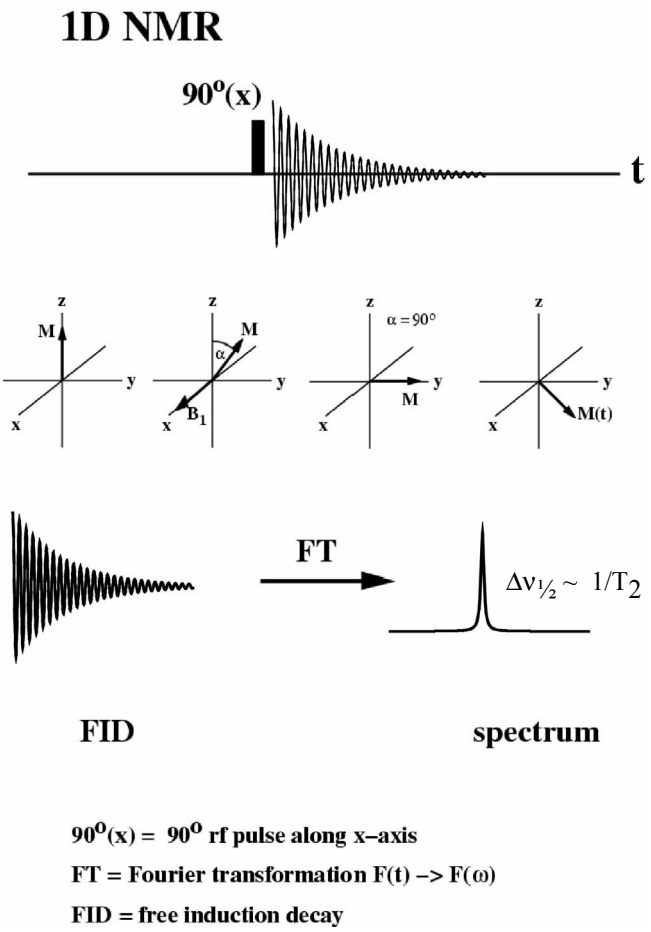
Nuclear magnetic resonance

In order to induce nuclear magnetic resonance, a oscillatory magnetic field has to be applied at the frequency which corresponds to the separation (ΔE) of the two spin energy levels. Since in a protein the frequencies vary slightly due to chemical shifts (see below), a number of frequencies have to be applied. There are two principal ways to achieve this. Either, the radiofrequency is varied within the expected frequency range, in the same way as a radiostation is tuned in (**continuous wave**, **CW NMR**), or a very short radiofrequency pulse is applied which inherently encodes a range of

frequencies allowing to induce resonance for the whole frequency spectrum in one experiment (**Fourier transform, FT NMR**). Since it takes time to step-wise tune over a range of frequencies by CW NMR while by FT NMR all signals can be excited at once, FT NMR is by far more sensitive. The response obtained from a FT NMR experiment is a superposition of the frequencies of all spins in the molecule as a function of time, F(t). In order to obtain the corresponding spectrum F(ω) (intensity as a function of frequency) a Fourier transformation is performed. Fourier transformation is a simple mathematical operation which translates a function in the time domain into the frequency domain and *vice versa*:

$$F(t) \longrightarrow F(\omega): \qquad F(\omega) = \int F(t)\, \exp^{-i\omega t}\, dt$$

$$F(\omega) \xrightarrow{IFT} F(t): \qquad F(t) = \int F(\omega)\, \exp^{i\omega t}\, d\omega$$



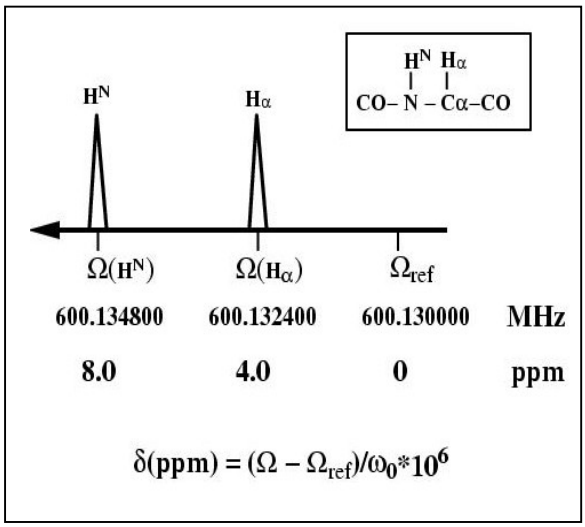
Relaxation and NMR linewidths

After a 90° pulse has been applied to the equilibrium *z*-magnetization, transverse magnetization is created in which the phases of the individual spin precession frequencies are correlated. Such a non-random superposition of states is called **coherence**. However, the population difference between the α- and β-states (which corresponds to *z*-magnetization) has disappeared. Thus, the system is in a

non-equilibrium state and will therefore relax back into the equilibrium state. This is described by two relaxation processes. The loss of phase-coherence in the transverse plane due to spin-spin interactions defines the T_2 or transverse relaxation time. In addition, the population difference along *z* is restored by interactions with the surrounding. This spin-lattice relaxation is described by the T_1 or longitudinal relaxation time. The linewidth of an NMR signal is described by the exponential decay of the FID, and reciprocal to the transverse relaxation time T_2. It scales with the molecular tumbling rate in solution and therefore increases with higher molecular weight. This is the main reason for the molecular weight limitation of NMR.

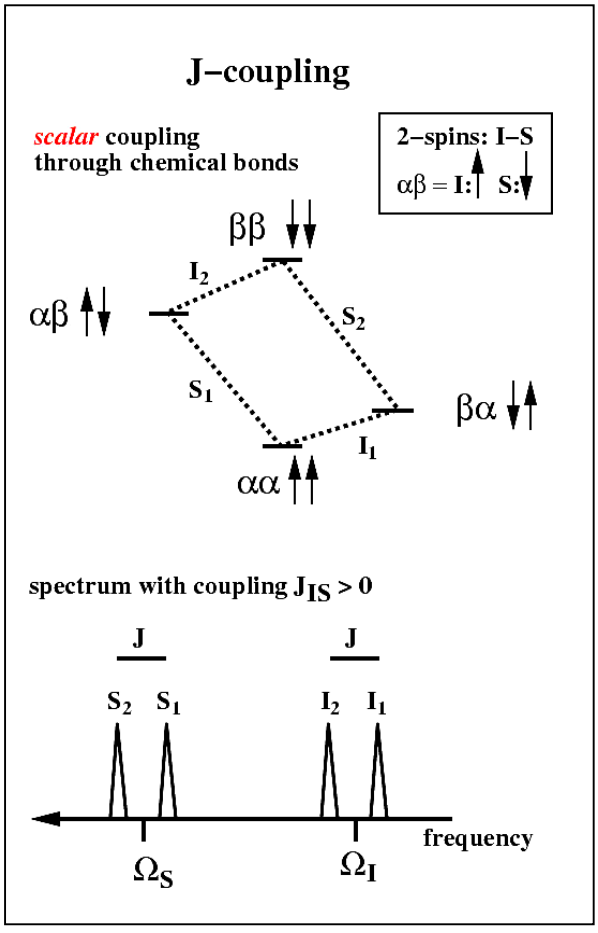
Chemical Shifts

The static magnetic field B_0 is shielded by the electrons in the local environment of a spin. Therefore, the individual resonance frequencies are slightly different reflecting the different chemical surrounding. The resonance frequencies are called chemical shifts and are measured in *parts per million (ppm)* in order to have chemical shift values independent of the static magnetic field strength. For example, backbone amide protons H^N in a protein resonate around 8 ppm, while H^α spins have resonance frequencies between 3.5-5.5 ppm.



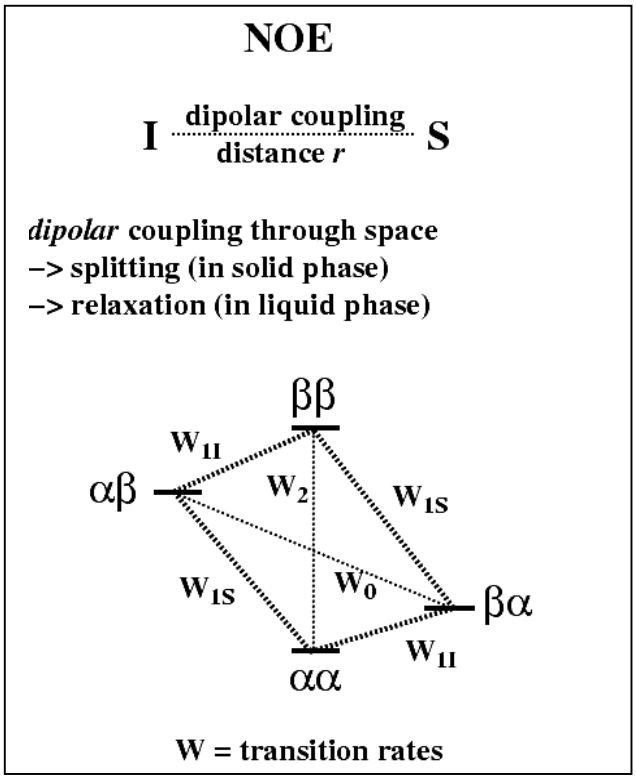
J-coupling

Scalar or J-couplings are mediated through *chemical bonds* connecting two spins. The energy levels of each spin are slightly altered depending on the spin state of a scalar coupled spin (α or β). This gives rise to a splitting of the resonance lines. Typical coupling constants are: $^1J(H^N,N) \sim 92$ Hz, $^3J(H^N,H^\alpha) \sim 2$-10 Hz, $^{>3}J < 1$. ^{3}J-couplings are well-correlated with the central dihedral angle by an empirical correlation, the Karplus curve. E.g. $^3J(H^N,H^\alpha)$ defines the backbone angle ϕ in proteins. . Scalar couplings are used in multidimensional (2D, 3D, 4D) **correlation experiments** to transfer magnetization from one spin to another in order to identify **spin systems**, e.g. spins which are connected by not more than three chemical bonds. For example, the amino acid ^{1}H spin system of Alanine consists of H^N, H^α and H^β.



Nuclear Overhauser Effect (NOE)

The nuclear Overhauser effect is a result of cross-relaxation between *dipolar coupled* spins as a result of spin/spin interactions *through space*. Dipolar couplings are usually in the kHz range, and depend on the distance between the two spins and the orientation of the internuclear vector with respect to the static magnetic field B_0. Fortunately, due to the fast overall tumbling of molecules in solution, the dipolar couplings are averaged to zero. Nevertheless, the dipolar couplings give rise to spin/spin and spin/lattice relaxation. The NOE is a result of cross-relaxation between spins and is defined by the transition rates W_0 and W_2 which involve spin flips of both spins. The NOE allows to transfer magnetization from one spin to another *through space*, and scales with the distance r between the two spins (NOE ~ $1/r^6$), e.g. two protons in a protein. Therefore the NOE is related to the three-dimensional structure of a molecule. For interproton distances > 5 Å, the NOE is too small and not observable.



Residual Dipolar couplings

Recently, it was found, that it is possible to measure residual dipolar couplings in solution by *weakly aligning a molecule in the liquid state*. If only a fraction of 1 out of 1000 molecules is aligned, the dipolar couplings, which in the solid state are in the kHz range, are scaled down to a few Hz. This allows to utilize the distance and the angle dependence of dipolar couplings in order to obtain **orientational restraints**, while still obtaining high-resolution NMR spectra:

$$D = \text{const} * 1/r^3 * < (3\cos^2\theta - 1)/2 >$$

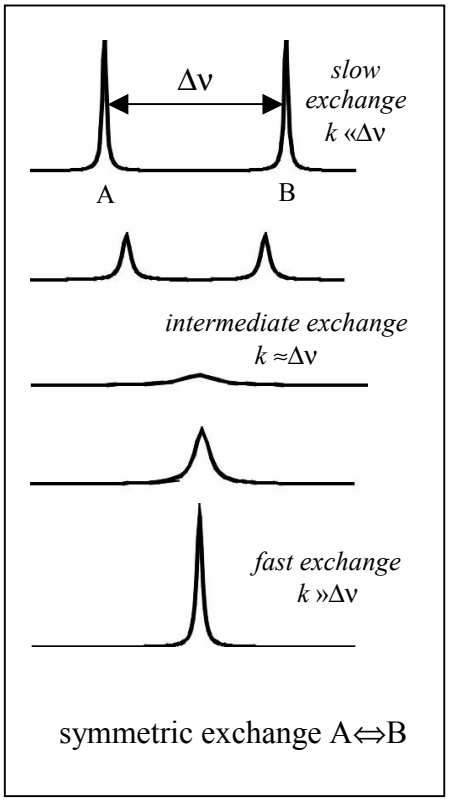
Here, r is the distance between the two spins which are dipolar coupled and θ is the angle of the distance vector with respect to the principal axis of the alignment tensor.

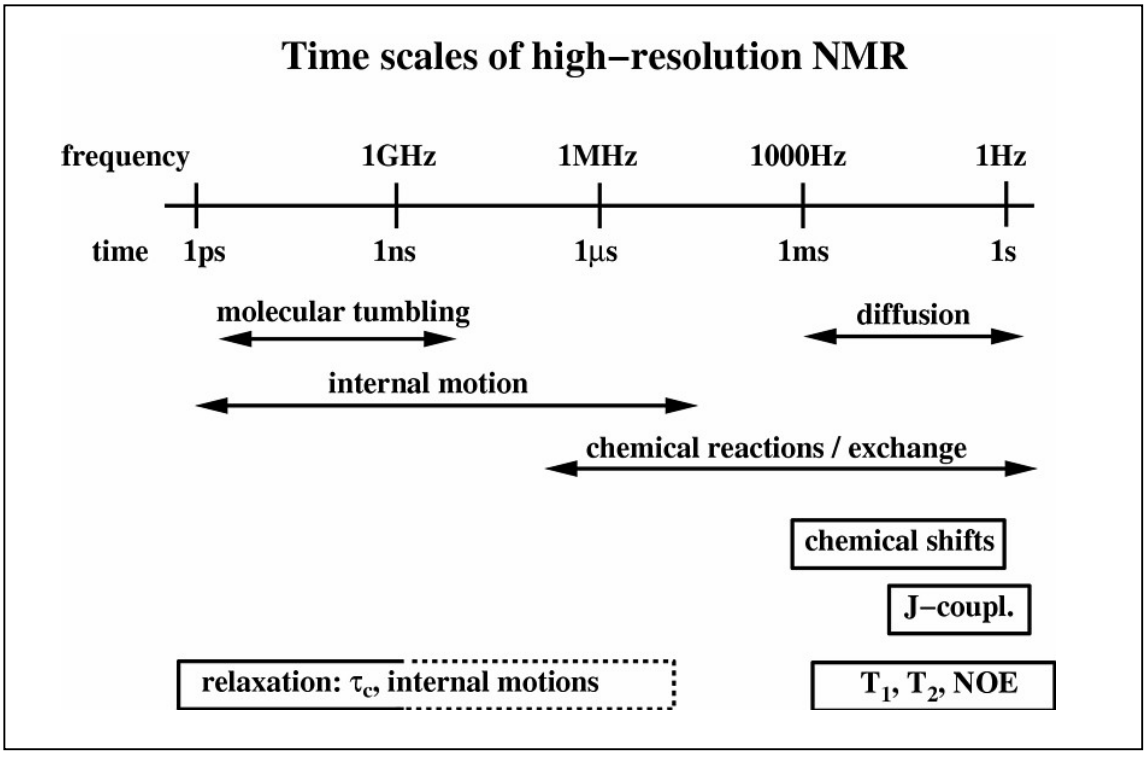
In order to achieve this weak alignment, NMR samples are prepared in slightly anisotropic solutions by adding e.g. a 5% mixture of DHPC/DMPC which form bicelles in aqueous solution, or

filamentous phages. Both, bicelles and phages align in the static magnetic field, and indirectly provide the weak alignment required for the observation of RDCs. Residual dipolar couplings provide not only additional structural information, but are crucial to define *long range* interactions, e.g. the relative orientation of two protein domains.

Exchange

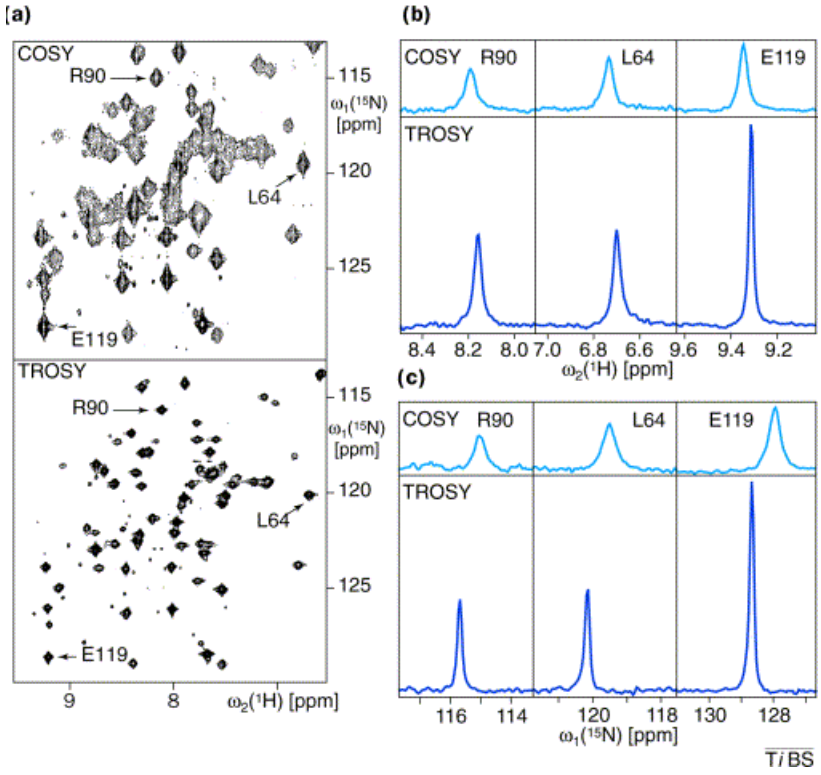
The exchange between two conformations, e.g. free and ligand bound forms of a protein, but also chemical exchange usually gives rise to two distinct NMR signals for a given spin due to different chemical environments in the two exchanging forms. If the exchange rate k is slow on the chemical shift time scale, two sets of signals are observed, if the exchange rate is fast on the chemical shift time scale only one signal is observed at an average frequency corresponding to the populations of the two conformations. If $k \sim \Delta\Omega$ intermediate exchange gives rise to very large linewidths. From lineshape analysis the exchange rate k and the populations of the conformations can be obtained. Since the frequency difference in Hz between the two signals is proportional to the magnetic field strength B_0, the exchange regime may change between different B_0 field strengths.





Larger molecules, relaxation and TROSY

The main problem in studies of biomolecules with molecular weights above 30 kDa is the fast decay of the NMR signal due to relaxation. The line widths in the NMR spectra are inverse proportional to the relaxation rates. Therefore the signal-to-noise in NMR spectra of larger molecules is poor. The line width problem can be overcome (i) by the use of **^{2}H-labeling** to eliminate ^{1}H mediated relaxation pathways and (ii) by a new NMR method called **Transverse Relaxation Optimized Spectroscopy (TROSY)**. By the combined use of ^{2}H-labeling and TROSY, high-resolution NMR spectra of structures with molecular weights >100 kDa can be recorded. This, significantly extends the range of macromolecular systems that can be studied by NMR in solution. For example, spectra of molecules in the 250 kDa range have been reported. TROSY has already been used to map protein/protein interfaces, to conduct structural studies on membrane proteins and to study nucleic acid conformations in multimolecular assemblies. Thus, the conformational states of individual macromolecular components can be studied and *de novo* protein structure determination in large supramolecular structures will be feasible in the future.



Comparison of a "conventional" 2D H^{N},N correlation experiment (COSY) with TROSY applied to a 110 kDa protein.

Structure determination by solution NMR

NMR parameters

Observable	Information
Chemical shifts ^{1}H, ^{13}C, ^{15}N, ^{31}P	Assignments, secondary structure
J-couplings (*through bond*) $^{3}J(H^{N},H\alpha)$, $^{3}J(H\alpha, H\beta)$, …	Through-bond correlation experiments dihedral angles: ϕ, χ, Karplus curves
NOE *(through space)*	Interatomic distances (<5Å)
solvent exchange (H^{N})	Hydrogen bonds
relaxation / linewidths ^{1}H, ^{13}C, ^{15}N cross-correlated relaxation (CCR)	Mobility, dynamics, conformational/chemical exchange; projection angles
residual dipolar couplings (RDCs) ^{1}H-^{15}N, ^{1}H-^{13}C, ^{13}C-^{13}C, …	Projection angles Nat. Struct. Biol. (1997) **4**, 732 Science (1997) **278,** 1111

NMR sample preparation

Typically larger proteins consists of a number of globular domains of 100-150 amino acid size. Such domains often have modular functions, e.g. one domain is required for binding to one ligand, another may interact with a different ligand. Therefore, in order to define a certain molecular interaction it is often sufficient to determine the structure of the protein domain of interest. The first step is then to overexpress this protein domain in a bacterial host, normally *E. coli*, and to prepare isotopically labeled samples enriched with ^{13}C and ^{15}N isotopes, for larger molecules (> 25 kDa) in addition ^{2}H-labeling will be required. RNA molecules are obtained from *in vitro* transcription. Often different constructs of varying size and domain boundaries are tried, in order to obtain samples which give good NMR spectra. Based on a 1D NMR experiment which takes 5 minutes, one can easily determine if a protein is folded, i.e. adopts a tertiary structure, or not. Different buffer conditions (pH, salt concentration, buffer) have to tested in order to optimize the quality of the NMR spectra and to avoid aggregation. Currently, milligram amounts (e.g. 5-10 mg for a 20 kDa molecule) are required for an NMR sample which typically consists of 0.5 ml of a 1 mM solution. However, with the new cyroprobe technology, only ¼ of this is required.

NMR spectroscopy & structure calculations

The next step is acquisition of the NMR data. A set of **heteronuclear multidimensional NMR experiments** are recorded for the assignment of all chemical shifts (^{1}H, ^{15}N, ^{13}C). ^{13}C- and ^{15}N-edited 3D NOESY spectra are measured in order to obtain structural information based on interproton NOEs. The total time for data acquisition for a protein structure determination is about 4 weeks, but this can be significantly reduced using cryoprobes or at higher magnetic field strengths (> 600 MHz). The NMR spectra are acquired with the FT NMR method and therefore have to be processed using Fourier transformation to obtain frequency domain data. During the processing the raw data can be improved with respect to digital resolution and signal-to-noise.

The **chemical shift assignments** typically can be obtained within one week and can also be automated. The chemical shifts already allow to define the secondary structure elements of a protein. The assignment of interatomic distances based on **proton/proton NOEs observed in 3D NOESY** spectra is quite time consuming (~ 1500 NOEs for a 10 kDa protein). However, programs have been developed in order to automatically assign the cross peaks in a NOESY spectrum once the chemical shift assignments are known. In addition to NOE-derived distance restraints, **^{3}J-coupling constants** and other parameters, e.g. **residual dipolar couplings** are measured, in order to derive additional structural information as input for the structure calculation. The protection of amide protons against chemical exchange with D_2O can be determined by adding D_2O to a lyophylized protein sample and monitoring the amide proton NMR signals in a $^{1}H/^{15}N$ correlation experiment. Hydrogen bonded H^N protons are protected from **H/D exchange**, while the signals of the other amides disappear quickly. Recently it has been demonstrated that **scalar couplings exist across hydrogen bonds** in proteins and nucleic acids. Therefore, hydrogen bonds can also be measured directly.

All the experimentally derived structural parameters: distance restraints from NOEs and H/D exchange, dihedral angle restraints from ^{3}J-couplings, projection angle restraints from residual dipolar couplings are used in a **restrained molecular dynamics/simulated annealing (MD/SA)** structure calculation. Structure calculation and NOE assigment is an iterative process. The result of an NMR structure calculation is an ensemble of structures, all of which are consistent with the experimental NMR data. The calculated structures should converge to the same fold. Poor convergence may indicate problems the experimental restraints and/or the MD/SA protocol.

Structural statistics and structural quality

In a journal article, the table summarizing the structural statistics of an NMR structure may look like this:

Table 1

Structural statistics for the Sgs1p HRDC domain.

	<SA>	<SA>$^{water\text{-}refined}$
Rmsd (Å) from experimental distance restraints*		
Unambiguous (1807 total)	0.009 ± 0.003	0.019 ± 0.002
Intraresidue (959)	0.014 ± 0.007	0.020 ± 0.004
Sequential (|i – j|) = 1 (280)	0.002 ± 0.002	0.005 ± 0.003
Medium range 1 < (|i – j|) ≤ 4 (269)	0.007 ± 0.0008	0.019 ± 0.003
Long range (|i – j|) ≥ 5 (299)	0.008 ± 0.001	0.021 ± 0.002
Ambiguous (83)	0.012 ± 0.001	0.023 ± 0.004
Hydrogen bonds (34)	0.010 ± 0.002	0.015 ± 0.005
Rmsd (°) from experimental torsion restraints†		
Dihedral angles (44 ϕ, 7 χ_1)	0.23 ± 0.05	0.43 ± 0.17
Rmsd from experimental ψ restraints‡		
$\Gamma^{H\alpha C\alpha C'}$ [Hz] (39)	1.8 ± 0.08	3.3 ± 0.16
$^3\Delta C^{\alpha}_{(ND)}$ [ppb] (39)	8.2 ± 0.03	8.1 ± 0.13
Coordinate precision (Å)§		
N, C^{α}, C′ (residue 13–88)	0.43 ± 0.07	0.61 ± 0.09
All heavy atoms (residue 13–88)	0.91 ± 0.10	1.07 ± 0.11
Structural quality¶		
E_{L-J}¥	−303 ± 8	−444 ± 15
No. of bad contacts (PROCHECK)	0 ± 0	0.07 ± 0.24
Quality index (WHATIF)	−1.03 ± 0.08	−0.76 ± 0.09
Ramachandran plot		
% in most-favored region	88.0 ± 1.5	91.5 ± 0.9
% in additionally allowed region	10.5 ± 1.1	8.4 ± 0.8

NMR structures are characterized by the **coordinate precision** which is defined as the **root-mean-square deviation (rmsd) for the atomic coordinates** between the structures of the NMR ensemble. The **rmsd from experimental distance, dihedral angle and projection angle restraints** should be small. Distance restraints should not be violated by more than 0.5 Å. The structural quality is evaluated based on the **electrostatic potential** term ($E_{L.-J.}$, Lennard-Jones potential), which should be negative. Since electrostatic potentials are not used during the MD/SA protocol, good structural quality is reflected by a reasonable $E_{L.-J.}$. The number of **bad contacts**, e.g. too short distances between atoms in a molecule should be small, otherwise the experimental restraints might be applied too strong and cause van der Waals clashes. In addition, the **Ramachandran plot**, which specifies the fraction of backbone ϕ,ψ angles which correspond to favored, additionally allowed, generally allowed and disallowed conformations based on statistical analysis of high-resolution crystal structures, is a good quality measure. Most ϕ,ψ angles should be within the allowed regions, a large number of ϕ,ψ angles in forbidden regions of the Ramachandran plot indicate poor structural quality.

Analysis of the structure: structure/function

Once the structure is obtained questions can be addressed to further characterize the function of the protein or nucleic acid molecule. Ligand interactions or the importance of certain residues for the stability of a structure can be probed by site-directed mutagenesis based on the 3D structure. Here, a 1D NMR can be used to test whether a point mutation affects the tertiary fold of the protein or not. This allows to distinguish between structural vs. functional effects of the mutation for ligand binding.

- Novel fold? Known fold? → *function*
- Potential binding sites? → *molecular recognition*
- Structure-based mutational study → *specificity, selectivity*
- NMR relaxation → *mobility, dynamics*

Dynamics and ligand binding of biomolecules

NMR is a very powerful method for characterizing the **overall tumbling and internal dynamics** of a molecule. The decay or relaxation of the NMR signals is a result of mutual interactions between the spins and with the environment. The coupling with the environment is a result of random modulation of the local magnetic fields due to the tumbling of the molecule in solution. Therefore, relaxation measurements provide information about dynamical parameters such as the overall tumbling correlation time in solution (τ_c) but also internal motions. Internal mobility of a region in a protein structure may correlate with induced fit recognition of binding partners as often is observed in protein/nucleic acid interactions. It is therefore important to understand the dynamical properties of a protein in addition to its three-dimensional structure.

Molecular interactions can be very efficiently characterized using simple and very sensitive NMR experiments. Changes in the environment of a spin due to binding of a ligand give rise to chemical shift changes in the NMR spectrum. These changes are largest near the binding site. Therefore, the binding surface of a protein with a ligand can be easily mapped. In addition, from **NMR titration experiments** dissocation constants can be determined. Due to the relatively high sample concentration, even very weak interactions can be detected. This chemical shift mapping is also used in a method called **SAR by NMR** for the rational design of drug molecules.

Perspectives

- higher magnetic fields (900 MHz), cryogenic probe technology
- new isotope labeling schemes (segmental labeling) for proteins and nucleic acids
- membrane proteins by solid state NMR
- convergence of solid state and high resolution NMR

- more methodological developments
  - automation of NMR based structure determination
  - quality of structures (new structural parameters)
  - molecular weight limit ⇒ now 900 kDa GroEL/GroES for binding site mapping

dynamics of biomolecular complexes

molecular interactions in large molecular assemblies (> 100 kDa)

membrane proteins by liquid & solid state NMR

rational drug design: SAR by NMR

complementarity with X-ray crystallography and EM

Literature

NMR of Proteins and Nucleic Acids.

Kurt Wüthrich. Wiley (1986)

Protein NMR spectroscopy – Principles and Practice.

John Cavanagh, Wayne J. Fairbrother, Arthur G. PalmerIII, Nicholas J Skelton. Academic Press (1996)

Nat. Struct. Biol. (1997) **4**, 841-865. (NMR supplement I)

Nat. Struct. Biol. (1998) **5**, 492-522. (NMR supplement II)

NMR spectroscopy of large molecules and multimolecular assemblies in solution. Wider, G. & Wüthrich, K. *Curr. Op. Struct. Biol.* (1999) **9**, 594-601

TROSY and CRINEPT: NMR with large molecular and supramolecular structures in solution, Pages 462-468 *Trends Biochem Sci* (2000) **25**, 462

Establishing a degree of order: obtaining high-resolution NMR structures from molecular alignment. Tjandra, N. *Structure* (1999) **7**, R205-R211