C h a p t e r 9

APPENDIX A

Principles of high-resolution NMR spectroscopy

The physical origin of NMR

It is well known from the diamagnetic and paramagnetic effect of matter that negatively charged electrons of an atom possess discrete magnetic spin moments which depend on the occupancy of outer shell electron orbitals. Likewise, the positively charged nuclei of atoms possess discrete magnetic spin moments dependent on the number of protons and neutrons present in the nucleus. Four of the five most abundant elements in biological macromolecules, namely hydrogen, carbon, nitrogen and phosphorus, exist as stable spin ½ nuclei (table 12). This means the magnetic moment of their atom nuclei can adopt exactly two discrete spin states: + ½ and – ½ allowing spectroscopic transitions from + ½ to – ½ or vice versa to take place.

While ^{1}H and ^{31}P are the most abundant natural isotopes, the spin ½ isotopes of carbon and nitrogen, ^{13}C and ^{15}N, only occur with a marginal natural abundance of 1.11 % and 0.37 %, respectively. This problem of extremely low abundance can be addressed by labeling biological macromolecules with ^{13}C- and ^{15}N- isotopes through biosynthetic routes.

table 12 **Spin ½ nuclei most useful for NMR of biological macromolecules**

Nucleus	$\gamma * 10^{-7}$ [rad $T^{-1}s^{-1}$]	frequency [MHz][1]	Natural abundance
^{1}H	26.75	750.000	99.98 %
^{13}C	6.73	188.580	1.11 %
^{15}N	-2.71	75.998	0.37 %
^{31}P	10.83	303.608	100.00 %

[1]Given are the resonance frequencies for a modern 750 MHz NMR spectrometer (B_o = 17.6 Tesla).

In the homogeneous magnetic field of a Nuclear Magnetic Resonance (NMR) spectrometer, the magnetic spin moment of such a nucleus precesses about the axis of the external magnetic field in one of two orientations: either parallel to the field in a low energy state or antiparallel to the field in a high energy state (fig. 54). By convention, the orientation of the external magnetic field, B_o, is always oriented along the Cartesian z axis. The energy gap, ΔE, between the two spin states is proportional to the strength of the magnetic field, B_o, and is given by:

$$\Delta E = (h/2\pi)\, \gamma\, B_o$$

in which h is Planck's constant. The gyromagnetic ratio, γ, is a constant unique to each nucleus under study reflecting a nucleus' sensitivity to the magnetic field. ^{1}H has the greatest gyromagnetic ratio, with ^{13}C and ^{15}N being only ¼ and 1/10 as sensitive as ^{1}H, respectively (table 12). Considering $\Delta E = h\nu$, this equation can be rewritten to give the widely known Larmor equation:

$$\omega_o = \gamma B_o$$

The angular velocity, ω_o, of precession is proportional to the strength of the magnetic field, B_o, and the gyromagnetic ratio, γ, of the nucleus under study. As $\omega_o = 2\pi\nu_o$, the same holds true for the resonance frequency, ν_o, of the transition between the two spin states, which is called the Larmor frequency.

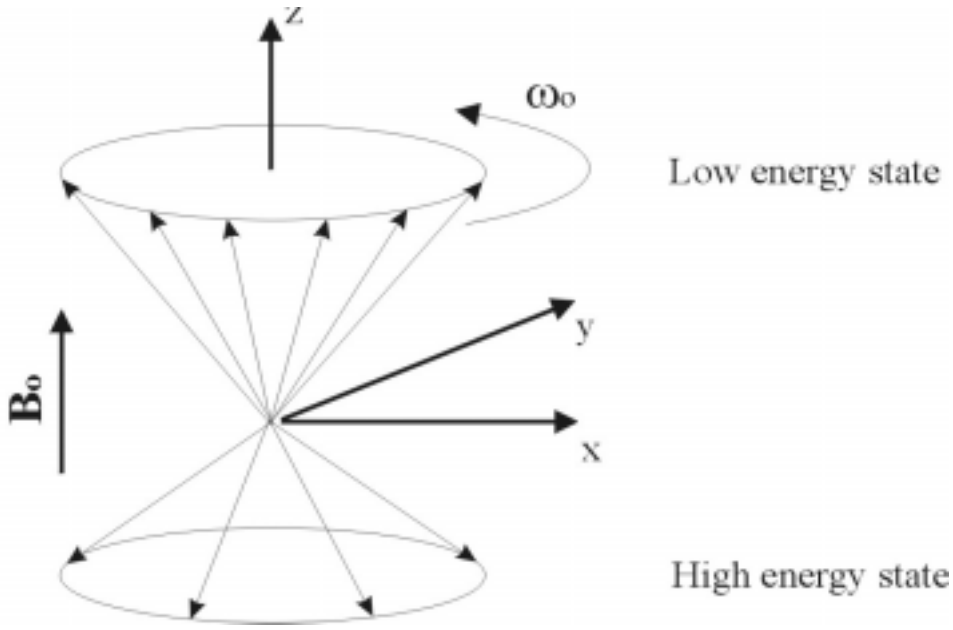

fig. 54 **Precession (ω_o) of nuclear magnetic spins with I=1/2 in an external magnetic field, B_o.** Note that more spins (arrows) populate the low energy state parallel to B_o than the high energy state antiparallel to B_o resulting in a net magnetic dipole moment along B_o.

The NMR signal

At thermal equilibrium, kT, the distribution of the nuclear magnetic spins in an external magnetic field is described by the Boltzmann equation,

$$n^+ = n^- \exp(-\Delta E/kT)$$

in which k represents Boltzmann's constant, and T is the temperature measured in Kelvin (K). As the number of spins in the lower energy level, n^-, is somewhat larger than the number of spins in the higher energy level, n^+, there is a net surplus, Δn, of magnetization oriented parallel to the external field. Considering that $\Delta E = (h/2\pi)\ \gamma\ B_o << kT$ at room temperature, this net magnetization along the z axis, M_z, can be approximated by:

$$M_z = h/(2\pi)\,\gamma\,\Delta n = h/(2\pi)\,\gamma\,(n^- - n^+) = (h/(2\pi))^2\,N/2\,\gamma^2\,B\,/\,(kT)$$

N represents the number of nuclei in the NMR sample, which equals the concentration of analyte or a multiple of it.

In the most basic 1D NMR experiment, equilibrium z-magnetization, M_z, which is also called longitudinal magnetization, is rotated into xy-magnetization by a brief radio frequency (rf) pulse and detected through a radio frequency receiver coil in the NMR spectrometer. In higher dimensional NMR experiments, a evolution period is inserted between the rf pulse and detection, during which correlations with adjacent nuclei can be recorded. Precessing with the Larmor frequency, $\omega_o/(2\pi)$, the xy-magnetization present at time t, $M_{xy}(t)$, induces an alternating voltage signal, U_{ind}, in the receiver coil given by:

$$U_{ind} \sim \omega_o\,M_{xy}(t) = \gamma\,B_o\,M_{xy}(t) = (h/(2\pi))^2\,N/2\,\gamma^3\;B_o^{\,2}\,/\,(kT)$$

The detected signal, which is often called the free induction decay (FID) (fig. 55), is composite of x-magnetization, $M_x(t)\cos(\omega_o t)$, and of y-magnetization, $M_y(t)\sin(\omega_o t)$. Both the x- and y-component, which are phase-shifted by 90°, are recorded by an electronic technique called 'quadrature detection', yielding simultaneously the amplitude as well as the phase of the signal. For 2D or higher dimensional spectra, complex signal detection is achieved by time-proportional phase incrementation (TPPI) [212] or by the States-Habekorn-Ruben method (States) [213] or by combinations of these methods (States-TPPI).

The complex time-domain signal, $\exp(-t/T_2)\exp(-i\omega_o t)$, (omitting the amplitude ξ_o) is converted into frequency-domain data by Fourier transformation (FT), resulting in a real absorptive frequency-domain NMR spectrum, $FT(\omega)_{real}$ (fig. 55), and an imaginary dispersive frequency-domain spectrum, $FT(\omega)_{imaginary}$, which is not further used.

$$FT(\omega) \sim {}_o\!\int^{\infty} \exp(-t/T_2)\exp(-i\omega_o t)\exp(i\omega t)\,dt$$

$$FT(\omega)_{complex} \sim 1/(T_2 + i(\omega - \omega_o))$$

$$FT(\omega)_{real} \sim T_2 / (1 + T_2^{\,2}(\omega - \omega_o)^2)$$

$$FT(\omega)_{imaginary} \sim i\,(\omega - \omega_o) / ((1/T_2)^2 + (\omega - \omega_o)^2)$$

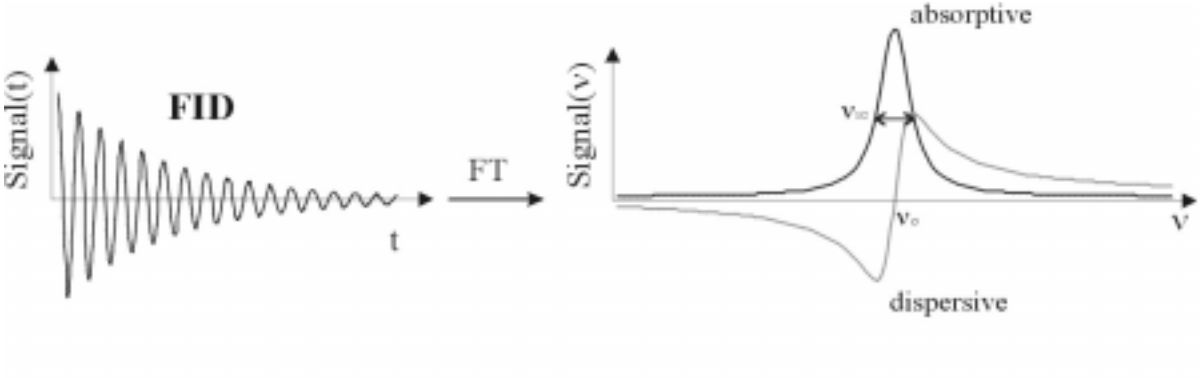

fig. 55 **The free induction decay (FID) and the resulting absorptive lineshape after Fourier transformation.** Note that the FID contains time-domain (t) data, whereas the absorptive NMR spectrum contains frequency-domain (ν)data.

The FID in fig. 55 is damped by an exponential decay, $\exp(-t/T_2)$, due to relaxation. This results in a Lorentzian lineshape given by

$$\text{signal}(\omega) \sim T_2 / (1 + T_2^2 (\omega_o - \omega)^2) \qquad \text{for } T_1 >> T_2$$

with a maximal signal height proportional to T_2 and a line width, $\nu_{1/2} = 1/(\pi T_2)$. This indicates that signal-to-noise and resolution, which correlates with small line width, of an absorptive NMR line is strongly influenced by the effective relaxation time T_2. The physical meaning of T_1 and T_2 are explained in the following description of relaxation in NMR spectroscopy.

Relaxation

During an NMR experiment, longitudinal and transverse relaxation lead to an exponential decrease of transverse magnetization. Longitudinal or spin lattice relaxation accounts for the return of xy-magnetization back to equilibrium z-magnetization at a rate $1/T_1$ by energy exchange processes between the elongated nuclear spins and the spins of the surrounding matter, generally called the lattice. The recovery of z-magnetization after an rf pulse is described by a first order exponential decay:

$$d M_z(t) / d t = - (1/T_1) M_z(t)$$

$$M_z(t) = M_z(\text{equilibrium}) (1 - \exp(-t/T_1))$$

Longitudinal relaxation is correlated with the overall rotational tumbling of the molecule in solution, because efficient spin-lattice energy transfer requires that the frequencies of rotational tumbling match those of the nuclear spin transitions (fig. 56). The longitudinal relaxation time, T_1, may be further affected by intramolecular mobility in flexible substructures.

Transverse relaxation or spin-spin relaxation accounts for the rapid dephasing of precessing xy-magnetization within the xy-plane at a rate $1/T_2$ by dynamic processes in the molecule. The transverse relaxation time, T_2, is shorter than or equal to T_1. In biological macromolecules with rotational correlation times, τ_c, in the nanosecond range, T_2 is considerably shorter than T_1 rendering T_2 the predominant relaxation pathway (fig. 56).

The region of small rotational correlation times in fig. 56, where the curves of T1 and T2 overlap, is referred to as the extreme narrowing limit. The rotational correlation time, τ_c, of globular approximately spherical biomacromolecules can be derived from the Debye equation:

$$\tau_c = \eta V_{hyd} / (kT)$$

Thus, the relaxation time and hence spectral resolution also depend on the hydrodynamic volume, V_{hyd}, which reflects the size of the molecule under study, and the viscosity, η, of the solution, which decreases exponentially with increasing temperature. It follows from these parameters that it is advantageous to collect NMR spectra on a biomacromolecule, which has been engineered to be as small as possible, in a low-viscosity (aqueous) solution at a temperature as high as durable for sensitive biomolecules.

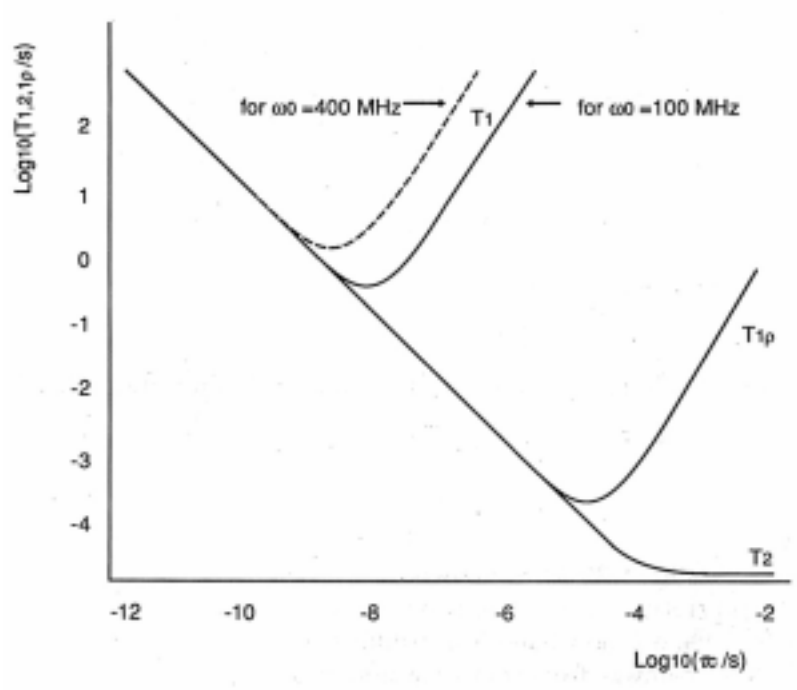

fig. 56 **Dependence of T_1 and T_2 on the rotational correlation times, τ_c.** At $\tau_c < 10^{-9}$ s, common to low molecular weight (< 1 kD) molecules, the magnitude of T_1 and T_2 are identical. However, at approximately $\tau_c = 10^{-9}$ s the T_1 curve inverts to become larger than T_2 for an operating frequency, w_o = 600 MHz common for protein NMR. Since proteins possess $\tau_c > 10^{-9}$ s, the decay of magnetization is dominated by T_2 relaxation in proteins.

The chemical shift

The chemical shift allows one to index each NMR-sensitive nucleus in the molecule under study with a certain number rendering it the most useful parameter for discriminating individual atoms. The chemical shift is defined as δ in parts per million (ppm):

$$\delta = (\nu - \nu_{ref}) / \nu_{ref} * 10^6$$

where ν_{ref} is the resonant frequency of a reference nucleus in Hz, and ν is the resonant frequency of the nucleus of interest in Hz. The chemical shift stems from moving electric charges of the electron cloud around the nucleus of interest. These moving charges induce a local magnetic field which opposes the applied external field. As a result, the external magnetic field B_o is reduced by a chemical shift tensor, σ, to an effective field B_{eff} at the position of the nucleus:

$$B_{eff} = B_o (1 - \sigma)$$

where B_{eff} relates to the resonant frequency ω via the Larmor equation:

$$\nu = \gamma B_{eff} / 2\pi$$

The chemical shift tensor is a 3x3 matrix containing the electronic shielding constants for each of the three dimensions in space. In solution fast motional averaging gives raise to an isotropic chemical shift tensor, σ_{iso}, which is the average of the x, y and z components of the trace of the matrix:

$$\sigma_{iso} = 1/3 (\sigma_{11} + \sigma_{22} + \sigma_{33})$$

The ^{1}H chemical shifts of proteins are commonly spread between –1 ppm and 10 ppm. Aliphatic ^{1}H atoms, e.g. of methyl and methylene groups, generally resonate in the upfield region (1 – 2 ppm), while aromatic ^{1}H atoms resonate in the downfield region (6 – 8 ppm) of the spectrum. In this manner, an empirical list of standard chemical shifts has been set up that is

fundamental to resonance assignment. Chemically equivalent ^{1}H atoms, such as the three protons of a methyl group, cannot be distinguished because they possess identical chemical shifts. Such groups are generally treated as a single virtual pseudo-atom located at the center of gravity of the real ^{1}H atoms.

Scalar (through bond) coupling: Coherence

Two nuclei, designated I and S, that are connected by one to four chemical bonds are coupled through the magnetic moments of the electrons in the bonding orbitals, as expressed by the coupling energy, E_{coup}, for a weakly coupled system:

$$E_{coup} = h\, J\, I_z\, S_z$$

where J represents the coupling constant. This is also called scalar or through-bond coupling. The strength of the coupling constant is a function of the number of chemical bonds, the strength of the chemical bond and the angle between the chemical bonds. For instance, the magnitude of the three-bond (vicinal ^{1}H-C-C-^{1}H) coupling constant, $^{3}J_{HH}$, is given by the Karplus equation [214]:

$$^{3}J_{HH}(\theta) = A + B\cos(\theta) + C\cos^{2}(\theta)$$

where θ is the dihedral angle. A, B and C are coefficients that depend on the substituents of the carbon atoms. The Karplus equation is often used in protein NMR for the determination of dihedral angles, such as the phi angle in the protein backbone which is spanned by the atoms H-N-Cα-Hα.

The utilization of through-bond coupling is fundamental to the design of most two and higher dimensional NMR experiments (except for 2D NOESY and similar experiments). Through-bond coupling is used to transfer magnetization in a controlled manner from one NMR-sensitive nucleus to another. The 'insensitive nucleus enhancement by polarization transfer' (INEPT) experiment demonstrates the general utilization of through-bond coupling in NMR pulse sequences.

The gyromagnetic ratios of ^{13}C and ^{15}N are considerable smaller than that of ^{1}H (table 12). Therefore, it is always advantageous to start a pulse sequence with the equilibrium magnetization residing on ^{1}H nuclei. For instance, if the chemical shifts of ^{15}N nuclei in the protein backbone are to be acquired, the NMR experiment begins with the preparation of transverse magnetization from the ^{1}H amide proton (labeled I), which is attached to the ^{15}N nucleus (labeled S). The INEPT pulse sequence (fig. 57) is described best by the product operator formalism. Concise and thorough introductions to the use of product operators for the mathematical description of magnetization transfer in NMR experiments can be read elsewhere, [215] and [18], respectively.

The pulse sequence begins with a $90°_{y}$ rf pulse on the ^{1}H equilibrium magnetization, I_z, resulting in I_x magnetization that evolves under the coupling Hamiltonian operator with the coupling constant, J_{IS}, and under the free-precession (chemical shift) Hamiltonian operator. The latter does not need to be considered in this calculation because it is refocused by the $180°_{x}$ rf pulse on I in the center of the pulse sequence. After scalar coupling for a period of time, t, the magnetization is given by:

$= I_x \cos(\pi J_{IS}t) + 2 I_y S_z \sin(\pi J_{IS}t)$

The $180°_x$ rf pulse on I reverses the sign of the single-quantum antiphase term $2 I_y S_z$:

$= I_x \cos(\pi J_{IS}t) - 2 I_y S_z \sin(\pi J_{IS}t)$

and the $180°_x$ rf pulse on S reverses the sign of this term again:

$= I_x \cos(\pi J_{IS}t) + 2 I_y S_z \sin(\pi J_{IS}t)$

For the remaining time t, I evolves under the coupling Hamiltonian:

$= I_x \cos(\pi J_{IS}t) \cos(\pi J_{IS}t) + 2 I_y S_z \cos(\pi J_{IS}t) \sin(\pi J_{IS}t)$

$+ 2 I_y S_z \sin(\pi J_{IS}t) \cos(\pi J_{IS}t) - I_x \sin(\pi J_{IS}t) \sin(\pi J_{IS}t)$

$= I_x \cos(\pi J_{IS}\, 2t) + 2 I_y S_z \sin(\pi J_{IS}\, 2t)$

and for $2t = 1/(2 J_{IS})$, the cosine and sine terms become 0 and 1, respectively:

$= - 2 I_y S_z$

The two $90°_x$ pulses on I and S convert the antiphase magnetization on I into antiphase magnetization on S,

$= 2 S_y I_z$

which can be converted into observable inphase S_x magnetization by another coupling period of 2t, while chemical shift evolution is decoupled. In summary, the INEPT pulse sequence has transferred longitudinal magnetization from a ^{1}H amide proton, I_z, to transverse single-quantum magnetization on a ^{15}N atom, $2 S_y I_z$, thereby increasing the signal strength by a factor of $\gamma(^1H) / \gamma(^{15}N) \sim 10$ provided that relaxation is disregarded.

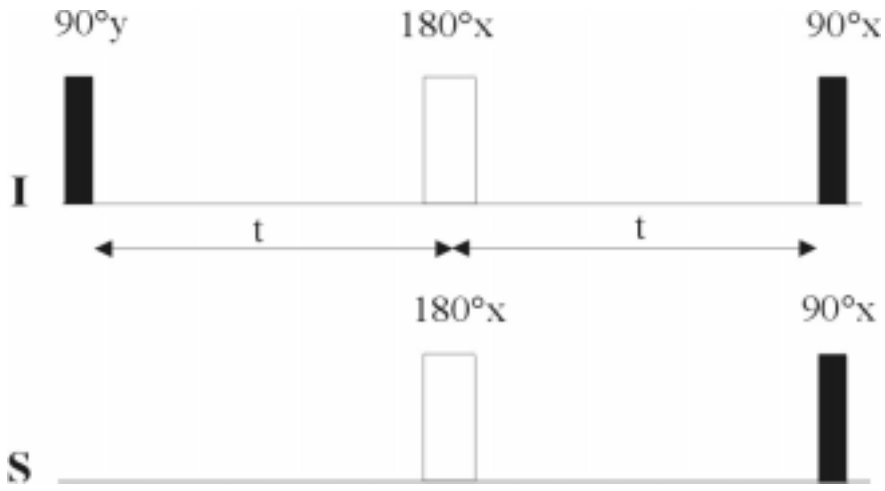

fig. 57 **Coherence transfer in the INEPT pulse sequence.** Magnetization from spin I (a ^{1}H atom) is transferred to single-quantum transverse magnetization on spin S (a ^{15}N or ^{13}C atom). 90° and 180° radio frequency (rf) pulses are illustrated by filled and open rectangles, respectively. Rf pulses on the upper and lower line are selective for spin I (^{1}H) and S (^{15}N or ^{13}C), respectively.

The frequently used HSQC (heteronuclear single quantum coherence) building block combines INEPT type of magnetization transfer with chemical shift labeling of the insensitive heteronucleus. The product operator treatment of the HSQC pulse sequence (fig. 58) begins with an INEPT transfer from a ^{1}H atom (spin I) to the attached heteronucleus (spin S, e.g. ^{15}N or ^{13}C):

$= 2 S_y I_z$

During the period t_1, the single quantum 2 S_yI_z magnetization evolves under the chemical shift Hamiltonian, $S_y\Omega_S t_1$, of spin S. The coupling Hamiltonian is refocused by the central 180° pulse on spin I after $t_1/2$:

$$= 2\ S_yI_z \cos(\Omega_S t_1)$$

After t_1 a reverse INEPT pulse sequence transfers the magnetization back to the ^{1}H atom,

$$= I_y \cos(\Omega_S t_1)$$

which may be detected during t_2 to give a 2D HSQC spectrum with cross peaks at (F1, F2) = (Ω_I, Ω_S). Alternatively, the $I_y \cos(\Omega_S t_1)$ magnetization is fed into further pulse sequence blocks, for example to generate 3D HSQC-NOESY or 3D HSQC-TOCSY experiments.

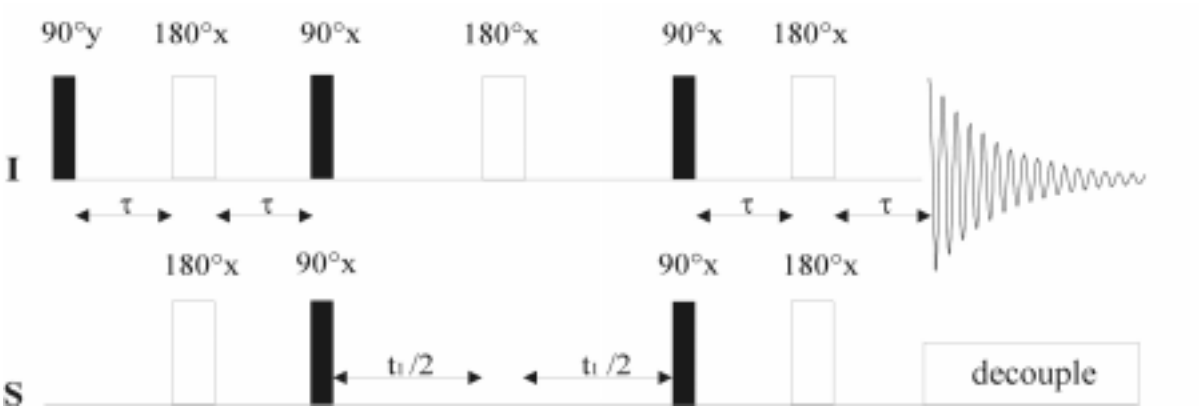

fig. 58 **Coherence transfer and chemical shift labeling in the HSQC pulse sequence.** The damped cosine wave represents the acquisition of the NMR signal (FID). Labeling is identical to fig. 57.

Dipolar (through space) interaction: The NOE

In addition to through-bond coupling, nuclear spins interact through space, because the magnetic moment of a spin I senses the magnetic field originating from the magnetic moment of another spin S in the vicinity. This through-space interaction between neighboring nuclear magnetic dipoles gives raise to the Nuclear Overhauser Effect (NOE), η_I, which describes the alteration of the intensity of a spin I when the transition of another spin S close in space is perturbed:

$$\eta_I = (\ \text{Intensity(I)} - \text{Intensity}(I_{eq.})\ ) / \text{Intensity}(I_{eq.})$$

in which $I_{eq.}$ is the equilibrium intensity.

The NOE intensity is proportional to the rotational correlation time, τ_c, and to the inverse sixth power of the distance, r^{-6}, between the two interacting spins under extreme narrowing conditions ($1/\tau_c >> \omega_o$). This r^{-6} dependence makes the NOE the most important parameter for the determination of molecular structures by NMR spectroscopy because it delivers the information about the fold of the molecule. For macromolecules with $\tau_c > 1$ ns, the theoretical maximum Nuclear Overhauser Enhancement is greatest for ^{1}H nuclei. NOE

measurements between ^{1}H atoms are also particularly well suited for structure determination because ^{1}H atoms always reside on the surface of the carbon/nitrogen skeleton of biological macromolecules. Therefore, ^{1}H-^{1}H distances between discontinuous parts of a macromolecule are minimal. Hence the corresponding NOE intensities are maximal, as compared with ^{13}C-^{13}C or ^{15}N-^{15}N distances. In practice, NOE intensities are converted into ^{1}H-^{1}H distances by calibration with a known intramolecular ^{1}H-^{1}H distance, r_{ref}, using the following equation:

$$r_{is} = r_{ref} (I_{ref} / I_{is})^{1/6}$$

The unknown distance, r_{is}, between the two ^{1}H atoms i and s is obtained from the sixth root of the ratio of the measured NOE intensities, I_{ref} and I_{is}, multiplied with the reference distance, r_{ref}. Suitable references include the distance between two methylene ^{1}H atoms (1.78 Å) and regular distances in the backbone of α-helices and β-sheets.

NOE intensities between individual ^{1}H atoms are obtained from NOESY-type of NMR experiments. The magnetization transfer pathway in a 2D NOESY (Nuclear Overhauser Enhancement Spectroscopy) experiment (fig. 59) is described by product operators as follows:

The initial 90°$_y$ pulse prepares I_x magnetization, which is chemical shift labeled during the evolution period t_1 under the free precession Hamiltonian:

$= I_x \cos(\Omega_I t_1) + I_y \sin(\Omega_I t_1)$

The second 90°$_y$ pulse creates inverted longitudinal magnetization, $- I_z$,

$= - I_z \cos(\Omega_I t_1) + I_y \sin(\Omega_I t_1)$

The I_y term is removed by pulsed field gradients or phase cycling. The $- I_z$ magnetization interacts with the cross-relaxation rate constants, a_{IS}, a_{IT} and a_{IU}, through space with the nuclei S, T, U, etc. during the mixing period τ_m:

$= - I_z a_{II} \cos(\Omega_I t_1) - S_z a_{IS} \cos(\Omega_I t_1) - T_z a_{IT} \cos(\Omega_I t_1) - U_z a_{IU} \cos(\Omega_I t_1)$

The final 90°$_y$ pulse converts these terms into observable inphase transverse magnetization,

$= I_x a_{II} \cos(\Omega_I t_1) + S_x a_{IS} \cos(\Omega_I t_1) + T_x a_{IT} \cos(\Omega_I t_1) + U_x a_{IU} \cos(\Omega_I t_1)$

which is detected during the FID multiplied by $\cos(\Omega_I t_2)$, $\cos(\Omega_S t_2)$, $\cos(\Omega_T t_2)$ and $\cos(\Omega_U t_2)$ for each of the above terms, respectively. The first term, $I_x a_{II} \cos(\Omega_I t_1)$, yields the diagonal peak at the position (F1, F2) = (Ω_I, Ω_I) in the 2D spectrum. The other terms give rise to NOE cross-peaks at (F1, F2) = (Ω_I, Ω_S), (Ω_I, Ω_T), (Ω_I, Ω_U), etc. Identical through-bond interactions are observed for the magnetization transfer from $-S_z$, $-T_z$, $-U_z$, etc. to $- I_z$ resulting in symmetry-related NOE cross-peaks at (F1, F2) = (Ω_S, Ω_I), (Ω_T, Ω_I), (Ω_U, Ω_I) in the 2D NOESY spectrum.

The cross-relaxation rate constants a_{IS}, a_{IT}, a_{IU}, etc grow linear with increasing τ_m, reach a maximum and drop at long τ_m when other relaxation pathways become dominant. This τ_m dependence of the NOE intensity necessitates that NOESY spectra are measured at various mixing times because the r^{-6} distance relationship is only valid in the initial linear part of the NOE build-up curve.

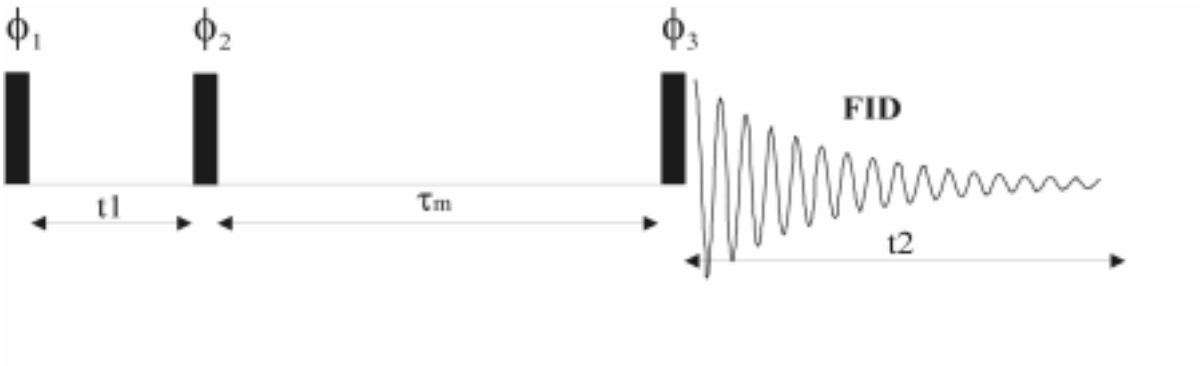

fig. 59 **The basic pulse sequence of a 2D NOESY experiment.** The evolution time, t1, is incremented giving rise to the second dimension of the spectrum. The three 90° pulses are labeled Φ_1, Φ_2 and Φ_3 to indicate that an appropriate phase cycling scheme has to be applied. The mixing time, τ_m, is to be optimized empirically for a certain molecule under study to reach maximal signal-to-noise and minimal spin diffusion.