

# Lecture 2

## Protein and NMR Structural Biology I

In this lecture we mainly give a brief introduction of the nuclear magnetic resonance (NMR) and a survey of using NMR for studies of structures and dynamics of proteins and nucleic acids.

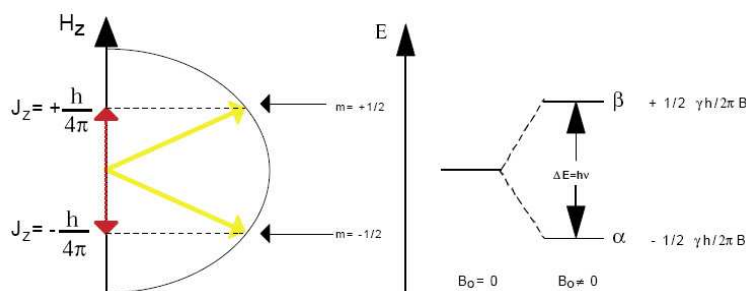
### 1 Overview of NMR

Nuclear magnetic resonance (NMR) is a physical phenomenon that measures the magnetic properties of each nucleus in a magnetic field based upon the magnetic properties of an atom's nucleus. All nuclei contain odd numbers of nucleons and some that contain even numbers of nucleons but odd charge have an intrinsic magnetic moment. NMR works by applying a strong magnetic field (the stronger the field the higher the resolution) to a sample and then measuring how the system responds to radio frequency waves. The radio frequency wave is electromagnetic radiation with an appropriate frequency which can be absorbed by a nucleus and switch its spin. Currently, NMR spectroscopy and X-ray crystallography are the only techniques capable of determining the three-dimensional structures of macromolecules at atomic resolution. In contrast to X-ray, the NMR approach is complimentary in different ways [1]:

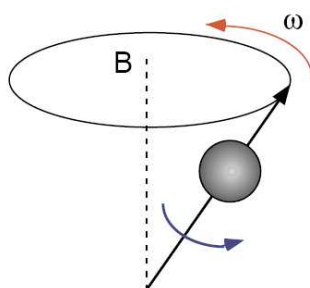
1. NMR directly measures protons,  $^{13}\text{C}$  and  $^{15}\text{N}$ .
2. NMR can be applied to molecules for which no single crystals are available. Thus new structures can be obtained, which are not available from X-ray studies.
3. NMR spectroscopy is a powerful technique for investigating intramolecular dynamics on many timescales, e.g. picoseconds to seconds. In addition, for a characterization of internal dynamics of biomacromolecular structures, NMR provides direct, quantitative measurements of the frequencies of certain high activation energy motional processes and at least semi-quantitative information on additional high frequency processes. In comparison, X-ray structure determinations may include an outline of the conformation space covered by high frequency structural fluctuations.

### 2 The Physical Basis of NMR Spectroscopy [2]

**Preliminary: Where does the magnetization come from?**  $^1\text{H}$  nuclei (or  $I = 1/2$  nuclei) possess a spin  $1/2$  which is associated with an angular momentum and a magnetic dipole moment. Each  $^1\text{H}$  nucleus thus behaves like a little magnet that spins around its dipole axis. A single spin  $1/2$  has the quantum mechanical property that, in an external magnetic field  $B_0$  (measured in Tesla, abbreviated T), only two orientations (called “parallel”



**Figure 1:** Energy levels of the  $\alpha$ - and  $\beta$ - states of  $^1\text{H}$  (and  $I = 1/2$  nuclei).



**Figure 2:** Rotation of nuclear momentum about its own axis (blue) and about the magnetic field axis (red).

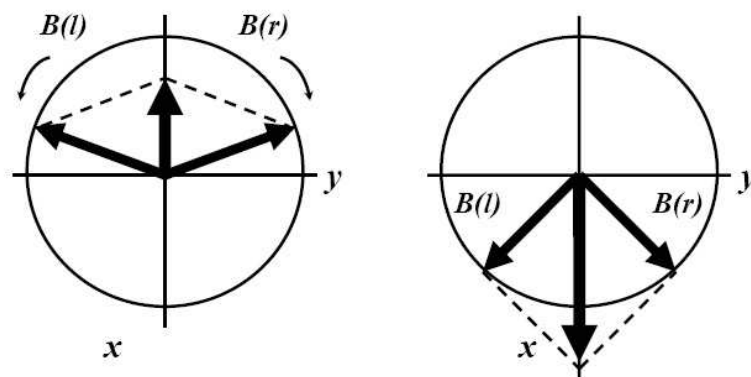
and “antiparallel” or “ $\alpha$ ” and “ $\beta$ ”, see Fig 1) with respect to  $B_0$  have a defined energy. The energy difference between these two states,  $\Delta E$ , is given by

$$\Delta E = \hbar\nu = -\frac{\hbar\gamma B_0}{2\pi}$$

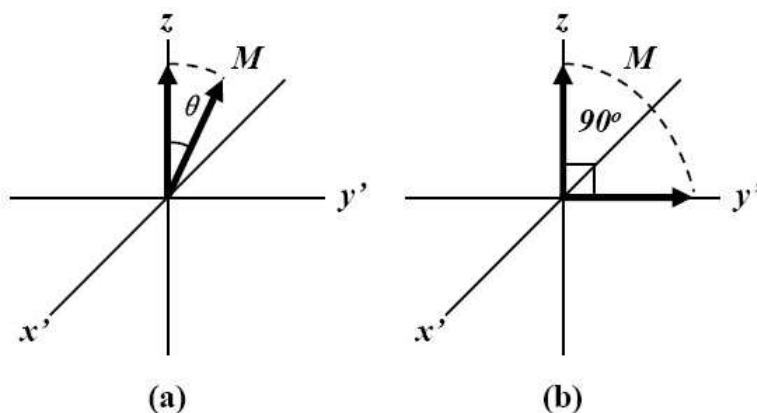
where  $\hbar$  is Planck’s constant,  $\gamma$  is the magnetogyric ratio ( $2.67 \times 10^8 T^{-1} s^{-1}$  for  $^1\text{H}$ ). Because of the angular momentum associated with spins, spins don’t align with the magnetic field  $B_0$ , but precess around an axis parallel to  $B_0$  with a frequency  $\nu_0$  (the so-called Larmor frequency,  $\omega_0 = 2\pi\nu_0 = -\gamma B_0$ , see Fig 2).

In an NMR experiment, many spins are present. As the energy difference between the states  $\alpha$  and  $\beta$  is small compared to  $kT$  ( $k$  Boltzmann constant,  $T$  temperature), there is, at room temperature, only a small excess of spins in the energetically lower state, even for the highest available magnetic fields. In thermal equilibrium, the average magnetization from all molecules constitutes a macroscopic magnetization  $M$  parallel to the magnetic field.

**The NMR experiment in a nutshell:** In a high resolution NMR experiment a glass tube containing a solution of the molecule of interest is placed in a *static magnetic field*  $B_0$  and then subjected to irradiation by one (could be several) radio-frequency (rf) field.  $B'$  is applied along the direction of the  $x$ -axis, which can be represented by two vectors with



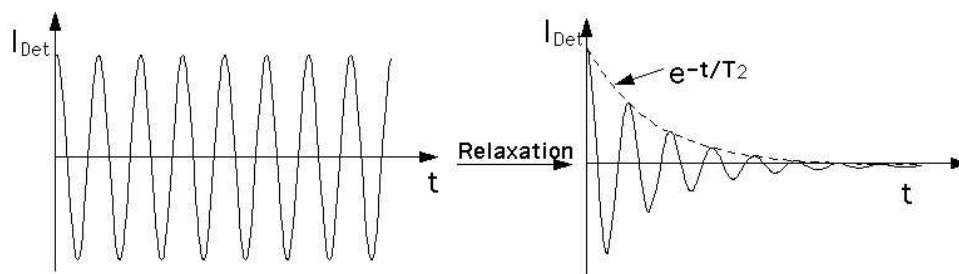
**Figure 3:** Representation of a linear alternating field (max  $2B$ ) as the sum of two rotating fields,  $B(r)$  (clockwise) and  $B(l)$  (anticlockwise).



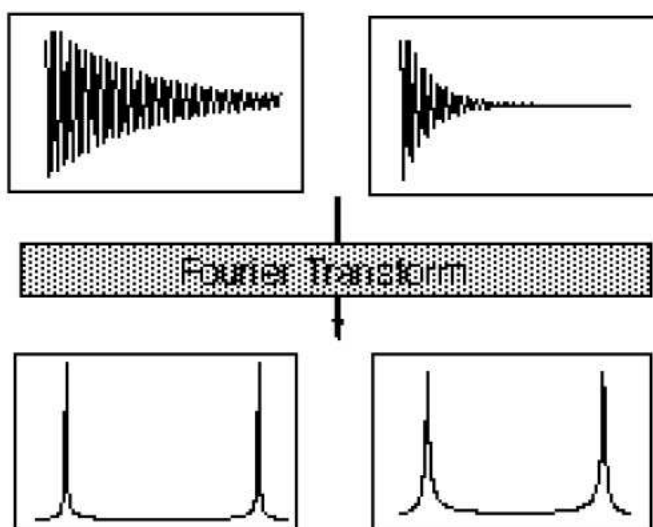
**Figure 4:** Direction of the macroscopic magnetization vector  $M$  in the rotating coordinate system: (a) after a pulse of arbitrary angle  $\theta x'$ ; (b) after a  $90^\circ x'$  pulse.

the same magnitude  $B_1$ , rotating in the  $x, y$ -plane with the same frequency  $\nu_L$ . One of the vectors  $B(r)$ , rotates clockwise, and the other,  $B(l)$ , rotates anticlockwise (Fig 3). Only the component precessing around  $B_0$  with the same frequency and the same direction as the precessing nuclear dipoles can interact with  $M$  (the macroscopic magnetization), denoted as  $B_1$ . If instead of the fixed coordinate system  $x, y, z$ , we use a rotating coordinate system  $x', y', z$ , which rotates with the same frequency as  $B_1$ , the effect of  $B_1$  is to turn the vector  $M$  about the  $x'$ -axis, *i.e.* in the  $y', z$ -plane. The angle  $\theta$  through which  $M$  is tipped increases with the amplitude  $B_1$  and the duration of the pulse (Fig 4a). A  $90^\circ x$  pulse is an rf pulse with the exact duration which takes  $M$  to become transverse (*i.e.* perpendicular to  $B_0$ ) on its way to the southern hemisphere (Fig 4b).

If we place a detector (*i.e.* a coil) along the  $y$ -axis, after a  $90^\circ x$  pulse, the system will induce a current in the detector, which is called free induction decay (FID) (Fig 5). The FID



**Figure 5:** Left: signal in absence of transverse relaxation; right: real FID.



**Figure 6:** Slowly decaying FIDs lead to narrow lines (left), rapidly decaying ones to broad lines (right).

decays exponentially with a decay constant  $T$ . The decay constant  $T$  (also called "relaxation time") is generated by the fact that the magnetization vector  $M$  turns back to equilibrium state when the rf pulse is switched off. There are two different time-constants describe this behavior:

1. The re-establishment of the equilibrium  $\alpha/\beta$  state distribution ( $T_1$ )
2. Dephasing of the transverse component (destruction of the coherent state,  $T_2$ ).

The NMR spectrum is obtained from Fourier transformation of the FID by plotting the amplitude of the sine and cosine functions versus their frequencies (Fig 6), where Fourier transformation is a mathematical procedure to represent any signal as a sum of sine and cosine functions with different frequencies.

### 3 Chemical Shifts

As mentioned in the former section, an atomic nucleus can have a magnetic moment (nuclear spin), which gives rise to different energy levels and resonance frequencies in a magnetic field. The total magnetic field experienced by a nucleus includes local magnetic fields induced by currents of electrons in the molecular orbitals. The electron distribution of the same type of nucleus (e.g.  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ) usually varies according to the atom's local electronic environment (including electrons from nearby atoms, *i.e.* binding partners, bond lengths, angles between bonds, ...). Therefore, the electronic environment of an atom shields or de-shields it from the external magnetic field and the local magnetic field at each nucleus (different atoms with different local electronic Environments “feel” a different external field and resonate at different frequencies). This variation of nuclear magnetic resonance frequencies of the same kind of nucleus, due to variations in the electron distribution, is called the chemical shift. The size of the chemical shift is given with respect to a reference frequency or reference sample, usually a molecule with a barely distorted electron distribution. Since an atom's chemical shift is very sensitive to the local electronic environment (and thus structure), the chemical shift can be considered as a sort of fingerprint for each atom, which plays an important role in NMR spectroscopy to identify the atoms and reveal the aspects of their local geometry. (see Chapter 2 of [3] for details.)

**Note:** Chemical shift is usually expressed in parts per million (*ppm*) by frequency, because it is calculated from

$$\delta = \frac{\omega_{\text{reference}} - \omega_{\text{observed}}}{\omega_{\text{reference}}} \times 10^6.$$

### 4 Introduction of NMR Experiments

**NMR spectrometer:** NMR spectrometers have a certain operating frequency, which is dependent on the strength of the applied magnetic field. Generally speaking, the higher the operating frequency, the better the resolution of the spectrum, e.g., a 500MHz spectrometer has a much stronger magnet than a 60MHz spectrometer. In structural biology, 500, 600, 700, 800, 900 and 950MHz instruments are commonly used. Duke has some 600MHz (around \$800K) and a very nice 800MHz NMR spectrometers (Fig 7), and a 950MHz spectrometer is being installed at UNC Charlotte, which will be operated by Duke.

**Protein NMR:** Protein NMR is performed on aqueous samples of highly purified protein. The source of the protein can be either natural or produced in an expression system using recombinant DNA techniques through genetic engineering. Recombinantly expressed proteins are usually easier to produce in sufficient quantity, and makes isotopic labelling possible.

The most abundant isotopes of carbon and oxygen,  $^{12}\text{C}$  and  $^{16}\text{O}$ , have no net nuclear spin, and thereby NMR spectroscopy cannot be directly exploited to identify them. And the most abundant isotopes of nitrogen,  $^{14}\text{N}$ , although has a net nuclear spin, also has a



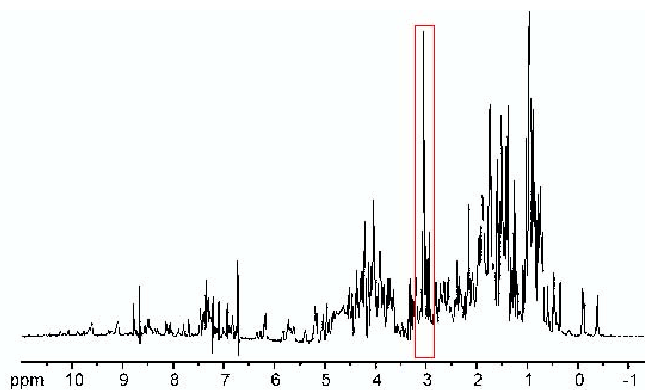
**Figure 7:** Duke has a world class set of high field NMR spectrometers, including 500MHz, 600MHz, and 800MHz instruments. Shown here is the 800MHz spectrometer, which is currently the state of the art for macromolecular NMR structural determinations.

large quadrupolar moment, a property of the atomic nuclei which prevents high-resolution information to be obtained from this isotope. In fact, NMR of proteins from natural sources are restricted to utilizing nuclear magnetic resonance based solely on proton. Nevertheless, the less common isotopes,  $^{13}\text{C}$  and  $^{15}\text{N}$ , are suitable for NMR experiments, and therefore labelling the proteins with these compounds open up possibilities for doing more advanced experiments which also detect or use these nuclei. Isotopical labelling is generally done by growing the expression host in a growth media enriched with the desired isotopes.

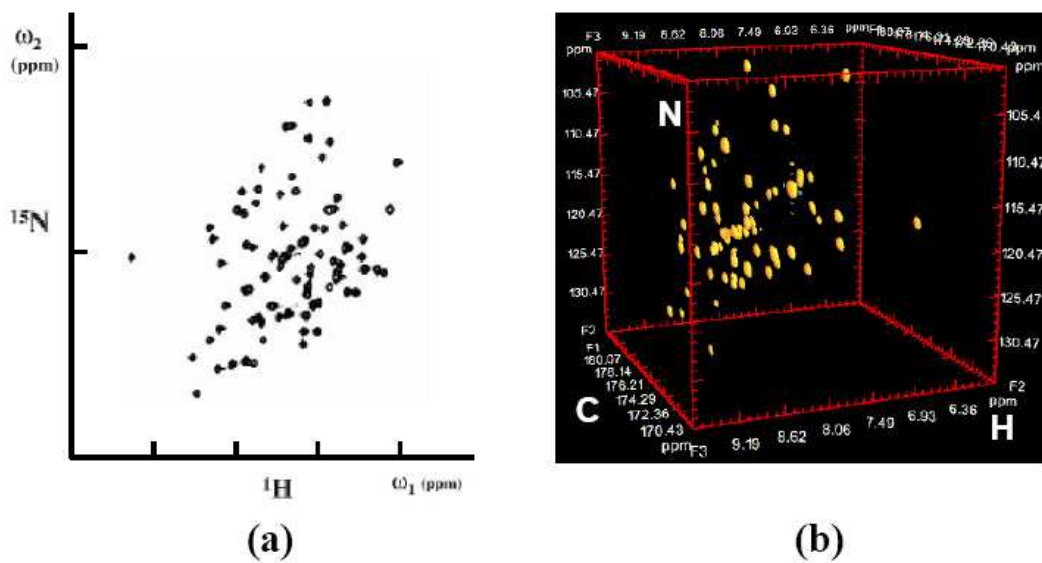
**Note:** There are many different kinds of NMR experiments, and each type reveals some differences about the structure or dynamics of the proteins. The different experiments also select for some specific kinds of atoms, e.g., all protons or all protons that are part of a methyl group, however, it is not possible to select for a single residue.

**Multi-dimensional spectra:** Although one-dimensional NMR spectral analysis, such as  $^1\text{H}$ -NMR spectral, is useful for elucidating the structures of small organic molecules, for more complex samples like biopolymers, the resulting spectral is far too complicated for interpretation as most of the signals overlap heavily (Fig 8). Therefore, we requires two- and three-dimensional analysis for these samples using  $^{13}\text{C}$  and  $^{15}\text{N}$  nuclei.

Two dimensional analysis basically measures the spectra of two different nuclei in a sample and plots them against each other. Figure 9a is an example of the two dimensional  $^1\text{H}$ - $^{15}\text{N}$  correlation spectra. These  $^1\text{H}$ - $^{15}\text{N}$  correlation experiments exploit the relatively strong shielding effects (J coupling) between the amide hydrogens and their corresponding amide nitrogens in the protein samples. However, complex protein samples may still have some degree of spectral overlap with 2D-NMR, thereby high dimensional experiments have been devised to deal with this problem. Currently, 3D & 4D spectra are available, and 5D & 6D experiments are possible in the laboratory.



**Figure 8:** In a 1D spectrum assignment is very difficult because of the crowding/overlap.



**Figure 9:** (a) Two dimensional  $^1\text{H}$ - $^{15}\text{N}$  correlation spectra. (b) Three dimensional  $^1\text{H}$ - $^{13}\text{C}$ - $^{15}\text{N}$  correlation spectra.

## **References**

- [1] Kurt Wuthrich. *NMR of Proteins and Nucleic Acids*. John Wiley & Sons. 1996.
- [2] P. J. Hore. *Nuclear Magnetic Resonance*. Oxford University Press, 1995.
- [3] Horst Friebolin. *Basic One- and Two-Dimensional NMR Spectroscopy*. VCH Pub. 1990.