

# NMR with Proteins and Nucleic Acids

Kurt Wüthrich, Zurich

(Institute for Molecular Biology and Biophysics, ETH)

Initial observations of the physical phenomenon of nuclear magnetic resonance (NMR) date back to the late 1940's, but the use of this technique for investigations of complex biological systems has become attractive only during the last decade. Such applications include studies of biomacromolecular structure and function, work on biological membranes, *in vivo* studies of biochemical processes, and imaging of macroscopic objects. Because of imminent practical applications in medical diagnosis and thanks to extensive coverage by the popular news media, interest in some of these developments has spread far beyond the scientific community, making biological NMR a widely popular field.

This article presents a survey of recent progress in NMR investigations of proteins and nucleic acids. Somewhat shielded from publicity by the excitement over applications in medicine, NMR spectroscopy of these macromolecules has recently also experienced decisive progress. With the use of two-dimensional (2D) experiments, NMR has become the first technique applicable to the determination of three-dimensional biopolymer structures in solution and other non-crystalline states.

## NMR and the Structure and Function of Biopolymers

Present NMR studies of macromolecular structures are based primarily on three recent developments. First, the use of two-dimensional spectroscopy allows us to improve the resolution of highly complex NMR spectra, and to delineate efficiently networks of scalar couplings and dipole-dipole couplings which may extend over the entire molecule. Second, new strategies for studies of proteins and nucleic acids have enabled us to assign individual NMR lines to distinct nuclear spins in the macromolecules. Third, new concepts have been developed and implemented for the determination of the spatial structure of polymer chains from data on scalar and dipolar couplings between nuclear spins.

Overall, the success of an NMR project depends critically on obtaining

numerous resonance assignments for distinct, individual spins. These represent the basis for spatial structure determination, as well as for studies of the molecular dynamics and of intermolecular interactions. The latter may relate to topical problems such as nucleic acid-drug interactions in cancer chemotherapy, or the regulation of protein and nucleic acid functions by specific interactions with low molecular weight effectors or with other macromolecules.

## 2D NMR Spectra of Proteins

Fig. 1 shows a 2D correlated (COSY) spectrum of the protons in a small protein. The lines on the diagonal from the upper right to the lower left correspond to the resonances of the approximately 500 individual hydrogen atoms in this molecule. Off-diagonal "cross peaks" manifest interactions between different nuclear spins. This is indicated in the contour plot of a 2D nuclear Overhauser enhancement (NOESY) spectrum in Fig. 2A by lines connecting two diagonal peaks via a cross peak. In COSY spectra the cross peaks are caused by scalar, "through-bond" interactions (spin-spin couplings) between protons located in the same chemical structure and sepa-

rated by three or less covalent bonds. COSY connectivities therefore allow us to divide the resonance lines into groups arising from chemical entities, for example the individual amino acids in a protein or the individual building elements in nucleic acids. In NOESY spectra which are in appearance similar to COSY, the cross peaks are a consequence of dipole-dipole, "through-space" couplings between protons which are located near each other in the folded molecule, at a distance of less than *ca.* 5.0 Å. They thus allow us to observe parameters which are directly related to the three-dimensional molecular structure. The origin of the NOESY cross peaks lies in the dipole interaction effect on the populations of the spin states. In conventional, one-dimensional experiments, the NOE's are observed as variations in the intensity of one NMR line when the resonance of another, nearby spin is selectively irradiated with an RF field.

For high resolution NMR studies the samples are placed in a highly homogeneous polarizing magnetic field of typically 11 T, which is produced by a superconducting coil. A 2D NMR experiment is recorded in a two-dimensional time space. In the simplest case, COSY, the following experimental scheme is used:  $90^\circ t_1$ ,  $90^\circ t_2$ , where  $90^\circ$  indicates a non-selective RF pulse,  $t_1$  is the "evolution period" and  $t_2$  is the "observation period". In COSY, the polarized nuclear spins are thus perturbed by a succession of two RF pulses. After the second pulse, the return of the magnetization to equilibrium is recorded versus time (" $t_2$ -axis"), and a second time dimension is

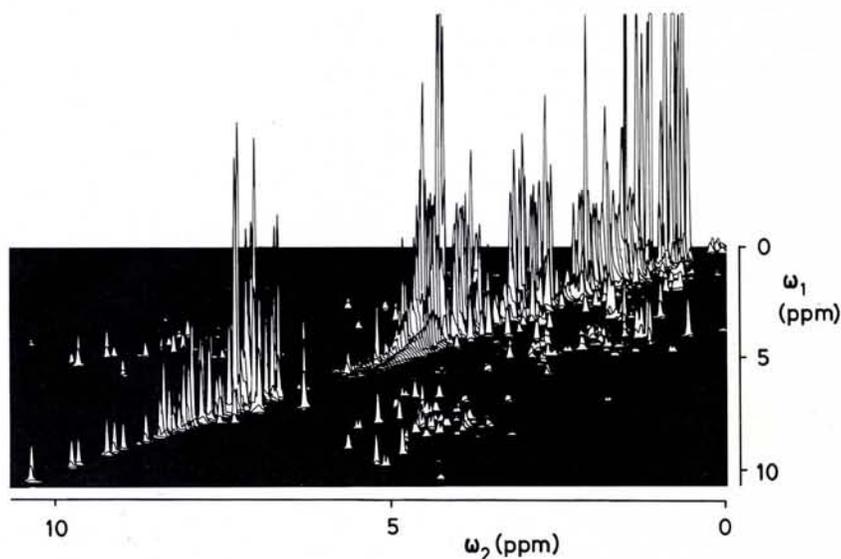


Fig. 1 —  $^1\text{H}$  COSY spectrum of the basic pancreatic trypsin inhibitor (BPTI), a small protein with molecular weight 6500, in  $\text{H}_2\text{O}$  solution. Larmor frequency 500 MHz. The two frequency axes,  $\omega_1$  and  $\omega_2$ , indicate difference frequencies in ppm from the Larmor frequency. Differences between the resonance frequencies of different spins of the same type are commonly referred to as "chemical shifts".

obtained by repetition of the same experiment with different time separation of the two RF pulses ("t<sub>1</sub>-axis"; typically 200 to 2000 increments of t<sub>1</sub> are recorded). For biomacromolecules, the data matrix  $s(t_1, t_2)$  consists typically of between 4 and 64 million data points. A two-dimensional Fourier transformation produces the desired frequency spectrum  $S(\omega_1, \omega_2)$  (Figs. 1 and 2A). The NOESY experimental scheme comprises three RF pulses:  $90^\circ t_1$ ,  $90^\circ \tau_m$ ,  $90^\circ t_2$ . The "mixing time",  $\tau_m$ , is kept constant throughout the measurements needed to record the data matrix  $s(t_1, t_2)$ . Employing 2D NMR in work with macromolecules brings the advantages that scalar or dipolar interactions between distinct, individual spins are observed with the use of non-selective RF pulses, and a complete set of all spin-spin interactions of a given kind in the entire molecule is manifested in a single experiment.

### Sequence-Specific Resonance Assignments

Fig. 2 illustrates NMR assignments in a polymer chain, which provide the basis for structure determination and studies of functional properties by NMR. First, the resonance lines on the diagonal must be attributed to spins in distinct locations of the chain (Fig. 2B). Once these "sequence-specific" assignments are available, the individual cross peaks can then be correlated with interactions

between distinct spins along the polymer chain. In the example of Fig. 2C, a NOESY cross peak indicates close approach between two spins located near the chain ends.

Obtaining sequence-specific resonance assignments is not straightforward because the polymer chains of proteins and nucleic acids contain multiple copies of the individual building blocks (amino acids or mono-nucleotides). Therefore, identification of the groups of scalar-coupled spins belonging to individual building blocks is in general not sufficient for the determination of a unique sequence location. The problem can be solved either by isotope labelling of specified positions in the molecular structures, or by employing NMR techniques capable of identifying the NMR lines of sequentially neighbouring amino acids or nucleotides. Prior to the recent introduction of this latter approach, lack of resonance assignments was a major stumbling block opposing detailed analysis of the NMR spectra of biopolymers.

### Spatial Structure Determination

The major source of information on the spatial molecular structures are the <sup>1</sup>H NOESY spectra. Typically, a large number of NOESY cross peaks are observed in globular proteins (Fig. 2A) or in double-helical nucleic acids, indicating that these molecular structures contain numerous pairs of closely spaced hydrogen atoms. Combined with sequence-

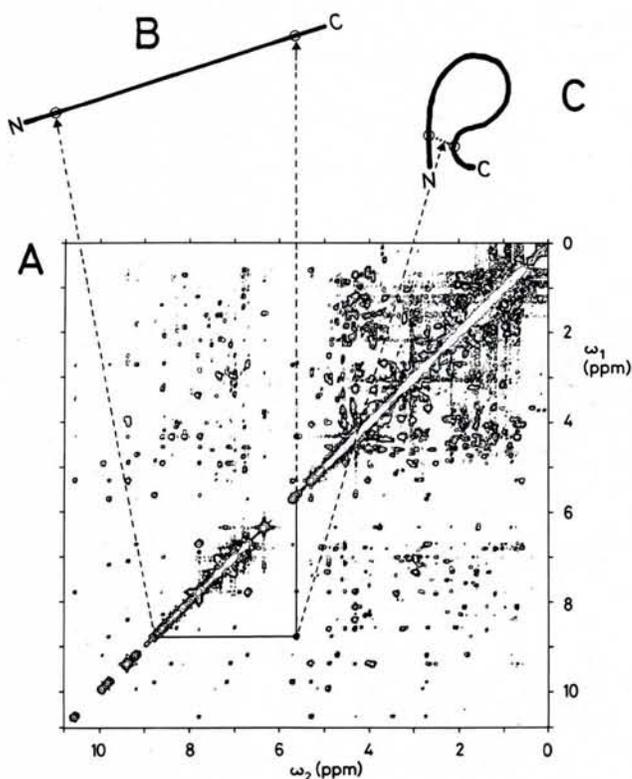


Fig. 2 — A: <sup>1</sup>H NOESY spectrum of BPTI in D<sub>2</sub>O solution; Larmor frequency 500 MHz. B: The straight line represents a polypeptide chain. The circles are protons along the chain. C: Same as B, with formation of a loop manifested by a NOESY cross peak between the two protons.

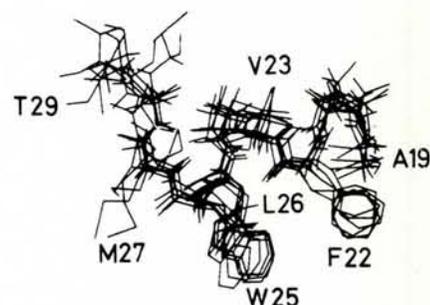


Fig. 3 — Superposition of five conformers of a segment of micelle-bound glucagon.

specific resonance assignments these "distance constraints" define the formation of loops by chain segments of variable lengths, as is illustrated for a single NOESY cross peak in Fig. 2C. With model building or the use of suitable mathematical procedures, notably distance geometry, the conformation space can then be searched for spatial arrangements of the polymer chain which are compatible with the experimental distance constraints.

In practice so far, the NOESY cross peaks in proteins were interpreted in terms of upper bounds on the distances between hydrogen atoms, rather than as exact distances (e.g. observation of a "strong" NOE between two protons would correspond to an upper distance limit of  $\leq 2.5$  Å, or a "weak" NOE would be interpreted as a distance-bound of  $\leq 4.0$  Å). Since in the linear polypeptide chain the protons in question may be separated by as much as 100 Å, or more, such distance constraints constitute nonetheless a drastic confinement of the accessible conformation space. However, the result of a structure determination based on qualitative distance constraints is not a unique structure, but a group of related conformers which all are compatible with the experiments. In Fig. 3 this is illustrated with a segment of micelle-bound glucagon. Five conformers have been superimposed for minimal root mean square distances, which were obtained in five independent distance geometry calculations using the same experimental distance constraints as the input.

Structure determinations by NMR are of interest because the biological macromolecules can be studied in environments which may closely resemble the physiological milieu. Examples are aqueous solutions, or the lipid-water interphase of lipid micelles, which was used in the case of glucagon (Fig. 3) to mimic the environment of the polypeptide hormone on the surface of a biological membrane. The NMR approach is limited to relatively small macromole-

cules ( $M \leq 20000$ ), but the conformational data can usually also include intricate details of the molecular dynamics.

### Molecular Dynamics

The use of NMR for studies of molecular dynamics relies on the observation of spectral properties in distinct NMR lines, which can be correlated with intramolecular motional processes. Once the NMR lines used for this purpose have independently been assigned, the intramolecular motions can be attributed to specified locations in the molecular structure, thus providing a map of the internal motility across the molecule. For both, proteins and double-helical nucleic acids, thermal fluctuations of distinct groups of atoms can thus be outlined in the interior, "structured" core regions of the molecules. Furthermore, in proteins

the amino acid side chains on the surface may either represent loose ends sticking out into the solvent, or they may be locked into distinct spatial arrangements by mutual interactions. The intriguing implication is that characterization of a protein in solution must include both a precise description of the molecular core, and a statistical account of the more pronouncedly flexible surface areas.

The NMR approach to studies of biological macromolecules outlined in this article is of very recent origin. It has already opened avenues to new insights into structure-function correlations with biopolymers. In addition to the use in basic research, technical applications can also be foreseen, for example for the evaluation of the action of drugs, or for product control in the production of pro-

teins by genetic engineering or by chemical methods.

### BIBLIOGRAPHY

— Wüthrich K., Wider G., Wagner G. and Braun W., 'Sequential Resonance Assignments as a Basis for Determination of Spatial Protein Structures by High Resolution Proton Nuclear Magnetic Resonance', *J. Mol. Biol.* **155** (1982) 311-319.

— Wider G., Macura S., Kumar Anil, Ernst R.R. and Wüthrich K., 'Homocyclic Two-Dimensional  $^1\text{H}$  NMR of Proteins. Experimental Procedures', *J. Magn. Reson.* **56** (1984) 207-234.

— Havel T.F. and Wüthrich K., 'An Evaluation of the Combined Use of Nuclear Magnetic Resonance and Distance Geometry for the Determination of Protein Conformations in Solution', *J. Mol. Biol.* **182** (1985) 281-294.

— Wüthrich K., *NMR of Proteins and Nucleic Acids* (Wiley, New York) 1986.

## Important Publications in Physics

### Nevill Mott Festschrift

*in celebration of his eightieth birthday*

**A special issue of Philosophical Magazine B**  
**Volume 52 No 3**

Professor Sir Nevill Mott, Nobel Laureate in Physics and Emeritus Professor of Physics, University of Cambridge has exerted an immense influence on generations of physicists whom he taught and collaborated with. The Editors and Publishers of the Philosophical Magazine have devoted this issue to a collection of appreciations and papers by friends and colleagues of Professor Sir Nevill Mott.

**0 85066 972 3 819 pp Sept 1985 £40.00**

### A Life in Science

**Sir Nevill Mott, Nobel Laureate in Physics,**  
*Emeritus Professor of Physics,*  
*University of Cambridge, UK*

This autobiography describes Sir Nevill's life in science, which covered a period of far-reaching developments in physics, including the beginning of quantum mechanics and its subsequent applications to solid-state physics.

**0 85066 333 4 220 pp approx. Spring 1986 £15.00 approx.**

### The Metallic and Nonmetallic States of Matter

Edited by **P P Edwards, University of Cambridge,**  
**& C N R Rao, Indian Institute of Science,**  
*Bangalore*

This book addresses itself to the fundamental question: what makes some materials metallic? It puts forward criteria for the occurrence of metallic character and shows systems for transition from the metallic to nonmetallic state. This is the first book to bring together all aspects of the metallic and nonmetallic states of matter - and transitions between the two.

**0 85066 321 0 400 pp Hardback Sept 1985 £40.00**

### An Introduction to Polymer Mechanics

**R G C Arridge, University of Bristol, UK**

This monograph is concerned with the mechanical properties of polymers, largely their elastic and viscoelastic properties. The book is divided into seven chapters, each designed to provide the serious student with a critical review of its subject matter as a basis for further research.

**0 85066 284 2 244 pp May 1985 £22.00**

### Potential Energy Surfaces

**D M Hirst, University of Warwick, UK**

The potential energy surfaces of molecules play a central role in our understanding of many chemical phenomena. This book considers several different approaches to the potential energy surface.

**0 85066 275 3 248 pp Hardback Feb 1985 £19.00**

### Fundamental Particles

**an introduction to quarks and leptons**

**Brian G. Duff, University College, London**

This book gives a simple introduction to the subject using the current view of particle physics which is embodied in the standard model, it emphasizes the basic physical models of elementary particles in terms of quarks and leptons and discusses in a simple way the forces which act between them.

**0 85066 332 6 130 pp Softback Nov 1985 £12.00**

### Introduction to Percolation Theory

**D Stauffer, University of Cologne, FRG**

This book provides a basic introduction to a modern interdisciplinary field of research, with emphasis on computer simulation. Percolation is used as an entrance to a more complicated phase transition theory, renormalization group, and fractals.

**0 85066 315 6 132 pp Softback July 1985 £12.00**

Send for our free physics catalogue



**Taylor & Francis**

Rankine Rd, Basingstoke, Hants RG24 0PR, UK  
242 Cherry St, Philadelphia, PA 19106-1906, USA