Journal of Biomolecular NMR, **27:** 13–39, 2003.
KLUWER/ESCOM

Perspective

NMR studies of structure and function of biological macromolecules (Nobel Lecture)*

Kurt Wüthrich
Eidgenössische Technische Hochschule Zürich, CH-8093 Zürich, Switzerland, and The Scripps Research Institute, 10550 N. Torrey Pines Rd., La Jolla, CA 92037, U.S.A.

Received and accepted 20 March 2003

Introduction

Nuclear magnetic resonance (NMR) spectroscopy is unique among the methods available for three-dimensional structure determination of proteins and nucleic acids at atomic resolution, since the NMR data can be recorded in solution. Considering that body fluids such as blood, stomach liquid and saliva are protein solutions where these molecules perform their physiological functions, knowledge of the molecular structures in solution is highly relevant. In the NMR experiments, solution conditions such as the temperature, pH and salt concentration can be adjusted so as to closely mimic a given physiological fluid. Conversely, the solutions may also be changed to quite extreme non-physiological conditions, for example, for studies of protein denaturation. Furthermore, in addition to protein structure determination, NMR applications include investigations of dynamic features of the molecular structures, as well as studies of structural, thermodynamic and kinetic aspects of interactions between proteins and other solution components, which may either be other macromolecules or low molecular weight ligands. Again, for these supplementary data it is of keen interest that they can be measured directly in solution.

The NMR structure of the *Antennapedia* homeodomain (Figure 1) illustrates one of the exciting features of being able to perform structural studies in solution. The polypeptide chain in this protein is only partially folded, with both chain ends showing pronounced disorder (Qian et al., 1989). In the complex with its operator DNA (Figure 2) the N-terminal chain end is located in the minor groove of the DNA, where it adopts a well-defined structure (Otting et al., 1990). This mode of intermolecular recognition illustrates the functional importance of partially structured polypeptide chains. Mammalian prion proteins are an even more striking example of partial polypeptide folding, since the three-dimensional structure of the benign 'cellular' form (PrP^C) includes a flexibly disordered 100-residue tail linked to the N-terminal end of a globular domain (Figures 3 and 4) (Lopez Garcia et al., 2000). Considering that the mechanism of transformation of PrP^C into the aggregated, disease-related form of mammalian prion proteins is still subject to speculation, the observation of this flexible tail has been highly intriguing.

Partially folded polypeptide chains are usually difficult to crystallize. Furthermore, if crystals can be obtained, the chain segments that are disordered in solution will either be ordered by intermolecular contacts in the crystal lattice, or they will not be visible by diffraction methods. As a consequence, NMR has in many cases been the only method capable of providing structural information on partially folded polypeptides. Although a standard protocol for NMR structure determination provides a static picture of the unstructured chain segments (Figures 1 and 4) (Wüthrich, 1986), additional NMR experiments can provide information on the frequencies of the rate processes that mediate transitions between discrete states of the molecule within the conformation space spanned by the static bundle of NMR-conformers (Wüthrich, 1995a, b).

The ability of NMR techniques to characterize macromolecular structures and their intermolecular interactions with high spatial and temporal resolution has long attracted keen interest. This article reports on experience gained with NMR studies of proteins in my laboratory during the past 35 years (Wüthrich, 2001).

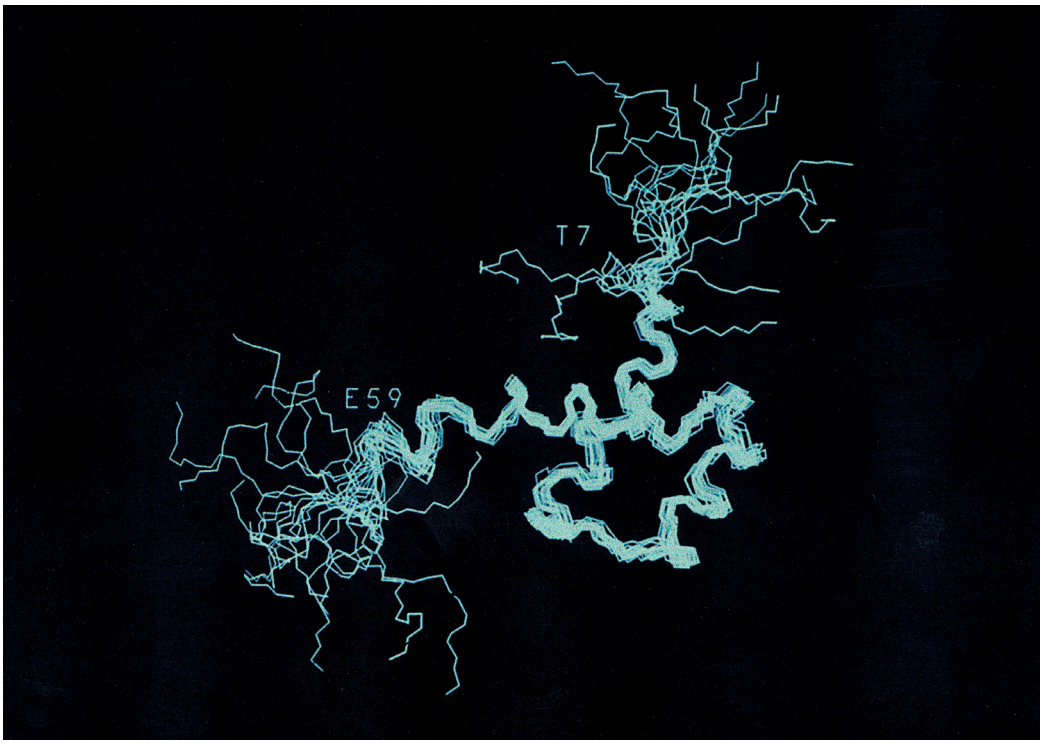


Figure 1. NMR structure of the *Antennapedia* homeodomain (Qian et al., 1989). A bundle of 20 superimposed conformers represents the polypeptide backbone. For the polypeptide segment 7–59 the tight fit of the bundle indicates that the structure is defined with high precision, whereas the two chain ends are disordered.

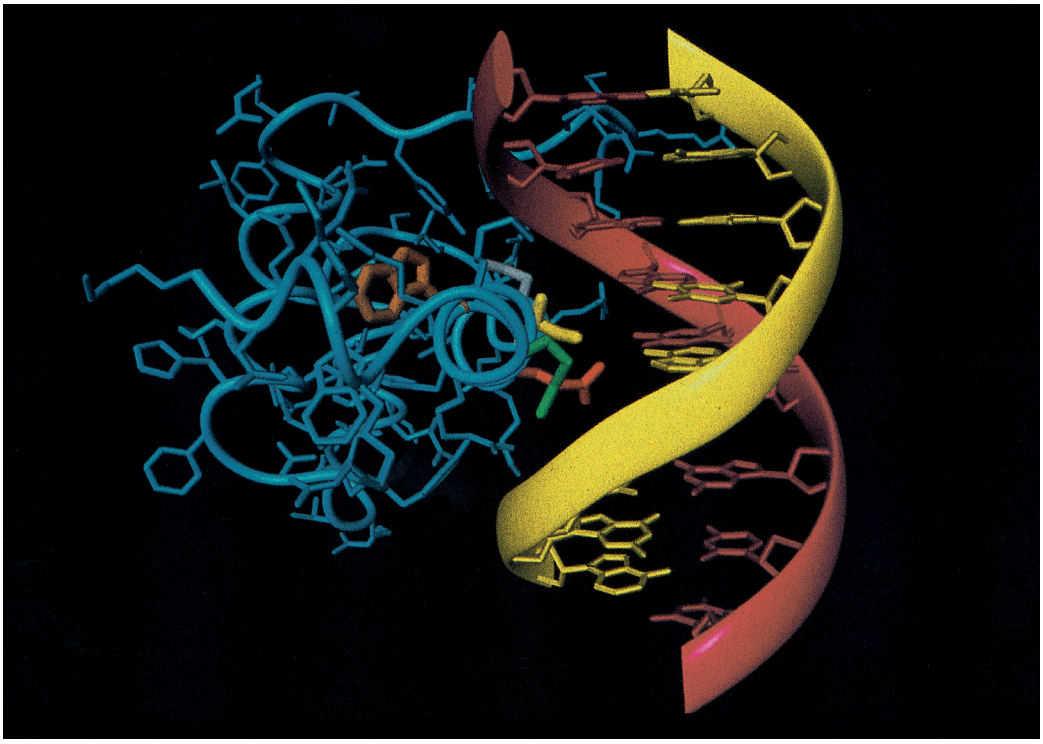


Figure 2. NMR structure of the complex formed between the *Antennapedia* homeodomain (blue, with functionally important residues in red, yellow, green and grey) and its operator DNA (yellow and red) (Otting et al., 1990).

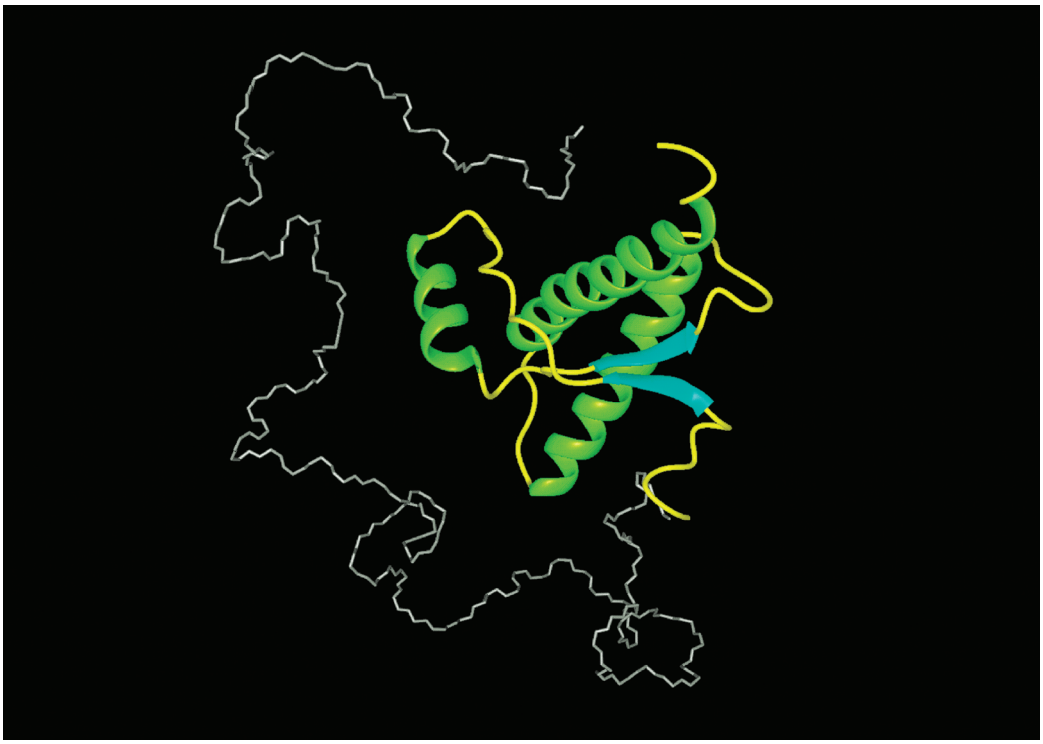


Figure 3. NMR structure of the bovine prion protein (Lopez Garcia et al., 2000). In the C-terminal globular domain of residues 126–230, α-helices are green, an antiparallel β-sheet is blue, and non-regular secondary structure is yellow; the 'unstructured' N-terminal tail of residues 23–125 is white.

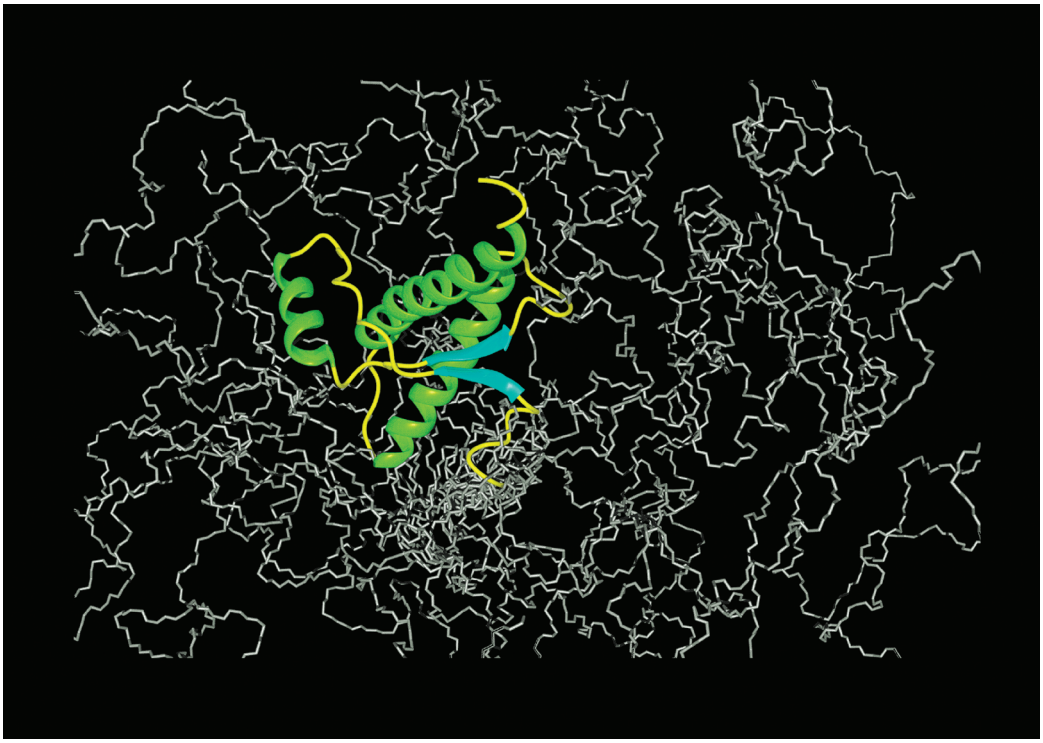


Figure 4. Visual impression of the variation of the bovine prion protein structure (Figure 3) during a time period of about 1 nanosecond. The superposition of 20 snapshots illustrates that the globular domain maintains its mean geometry, whereas the tail undergoes large-scale changes with time.

NMR spectra of proteins in solution

When I joined the field of biomacromolecular NMR spectroscopy in 1967, important qualitative NMR features of amino acids and proteins had already been noted and tentatively rationalized (Saunders and Wishnia, 1958; Jardetzky and Jardetzky, 1958; Kowalsky, 1962; Mandel, 1965; McDonald and Phillips, 1967). Most important, it had been well documented that the spectrum of a globular protein is more complex than the sum of the NMR lines from the constituent amino acid residues in the polypetide chain (Figure 5), and the differences between the ^{1}H NMR spectra of folded and unfolded ('random coil') forms of a polypeptide chain had been tentatively explained by different interactions with the solvent (Figure 6). The spectral analysis was primarily focused on the positions of the individual NMR lines in the ^{1}H NMR spectrum, as given by the 'chemical shift', δ, in parts per million (ppm) relative to a reference compound (Figure 5). Although the chemical shift is primarily determined by the covalent structure of the amino acid residue, it can also be significantly affected by the interactions with the solvent. Therefore, the exclusion of the solvent water from the interior of a globular protein (Figure 6) causes the chemical shifts of the core residues to be different from those of the water-exposed amino acid residues, so that even NMR lines originating from multiple residues of the same amino acid type can be distinguished. This 'conformation-dependent chemical shift dispersion' was found to be sufficiently large to enable ^{1}H NMR studies of protein denaturation (Figure 7) (McDonald and Phillips, 1967). This then indicated the exciting prospect of using NMR for detailed studies of protein folding, and in particular for distinguishing between two-state and multiple-state folding and unfolding transitions. For me, the observations and ideas illustrated by Figures 5–7 also showed that one would need the ability of resolving and analyzing highly complex NMR spectra (upper trace in Figure 5) in order to obtain information about physiologically active, folded forms of proteins. This was a clearly defined problem that appeared sufficiently challenging to be attractive for my own research.

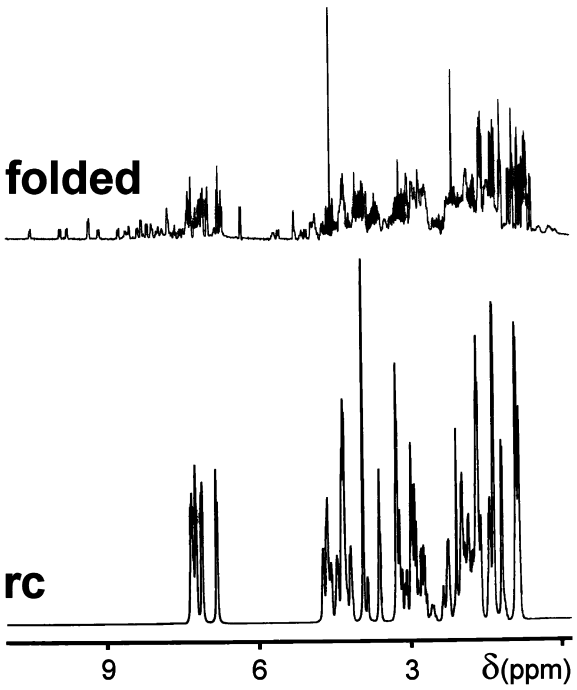


Figure 5. One-dimensional (1D) ^{1}H NMR spectra of the small protein bovine pancreatic trypsin inhibitor (BPTI, M $\approx$ 6000). Top: experimental spectrum of folded, active BPTI in a freshly prepared ^2H_2O-solution. Bottom: simulated spectrum for the unfolded, random coil form of the BPTI polypeptide chain.

In attempts at rationalizing in more detail the observations of Figures 5 and 7, it had been suggested that local magnetic fields generated by 'ring currents' in the aromatic amino acid side chains (Figure 8) should cause outstandingly large conformation-dependent chemical shift changes of hydrogen atoms located near the rings in the three-dimensional protein structure. This was qualitatively confirmed by comparison of the largest observed deviations from random coil chemical shifts (Wüthrich, 1986) in hen egg white lysozyme (Figure 7) with calculations of the 'ring current shifts' based on the then-available low-resolution crystal structure of this protein (McDonald and Phillips, 1967). It had also been pointed out that hemoproteins could be expected to have particularly large ring current shifts for hydrogen atoms located near the heme groups (Figure 9) (Kowalsky, 1962; McDonald and Phillips, 1967).

For my start in the field of NMR with proteins, the prediction of unique, large heme ring current shifts seemed to be an attractive feature of hemoproteins, and the presence of the heme iron (Figure 9) was appealing in view of my background in inorganic chemistry. As an additional advantage, low-resolution crystal structures for several hemoproteins were already available in 1967. It turned out that for several years nearly all my research projects were focused on the heme iron and its coordinatively linked ligands, and on the non-bonding interactions of the heme groups with their immediate environment in the hemoproteins. As

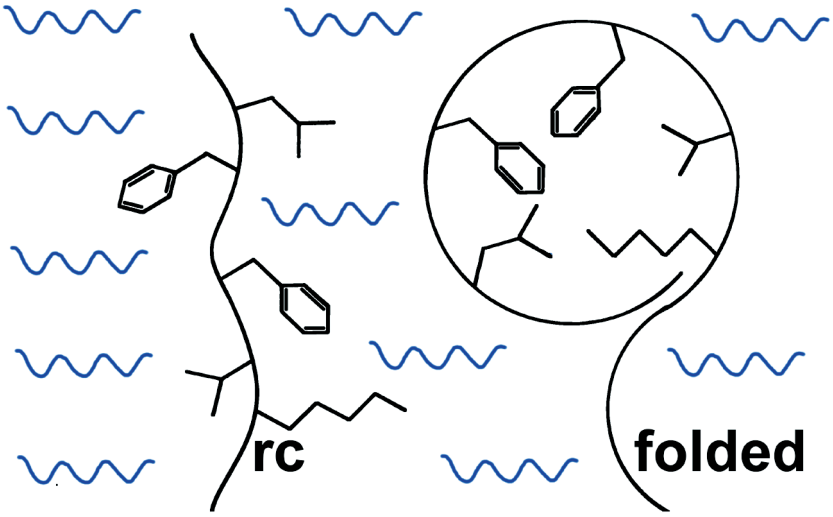


Figure 6. Scheme illustrating that the solvent (blue) has free access to all parts of a random coil polypeptide chain (rc), whereas it is excluded from the core of a folded globular protein.

anticipated, the large heme ring current field generates a small number of well-separated resonance lines in the spectra of folded, globular hemoproteins (Figure 10). For the paramagnetic states of hemoproteins, additional well-resolved lines result from interactions with the unpaired electrons of the heme iron. Although they represented less than 3% of all the hydrogen atoms in any given hemoprotein, the well-resolved resonances were of crucial importance, because these lines could be selectively irradiated in 1D ^{1}H NMR experiments. For example, performing some simple ligand exchange and redox reactions with cytochrome c (Figure 11) resulted in the identification of one of the axial ligands bound to the heme iron, which had not been seen in the low resolution X-ray crystal structure available in 1969.

The experiments in Figure 11 (Wüthrich, 1969) started with the oxidised form of cytochrome c. As was then learned through this study (see below), ferricytochrome c contains a methionine side chain of its polypeptide chain in one of the axial heme iron coordination sites. Upon addition of KCN to the protein solution, this methionine is replaced in the binding site on the heme iron by a cyanide ion, and therefore moves out of the heme ring current field. The completion of this reaction can readily be monitored, since different patterns of well-resolved resonance lines are present in the ^{1}H NMR spectra of ferricytochrome c and its cyanide complex, respectively. After reduction of the heme iron to the ferrous state, the cyanide complex is thermodynamically less stable than the native form of cytochrome c, and in a slow reaction the cyanide in the axial coordination site of the heme iron is again replaced with the natural methionine ligand. The return of the axial methionine into the heme ring current field was monitored in real time (Figure 12)

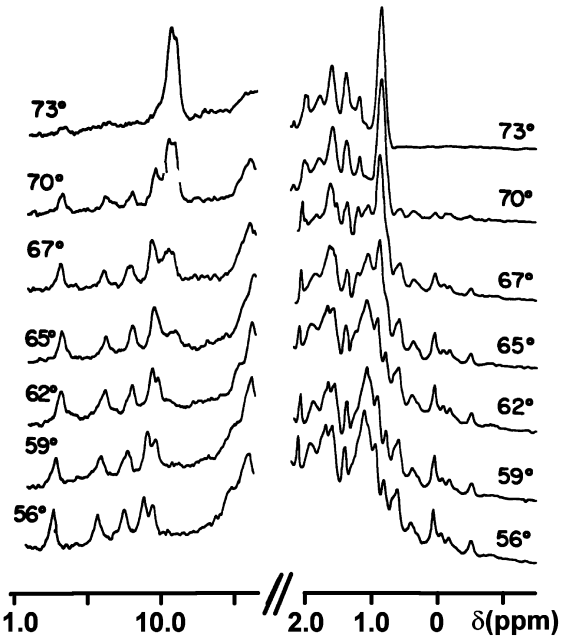


Figure 7. Low-field and high-field regions of 1D ^{1}H NMR spectra of the protein lysozyme at different temperatures, reflecting the transition from the folded form at 56 °C to the random coil form of the polypeptide chain at 73 °C (220 MHz, H_2O-solution) (adapted from McDonald and Phillips, 1967).

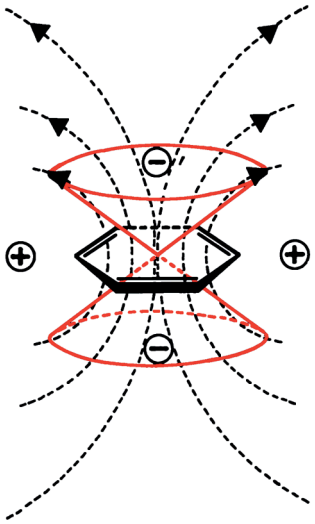


Figure 8. Local 'ring current field' around aromatic rings in solution induced by an external, static magnetic field. The shape of the ring current field is indicated by the red double-cone and by broken magnetic-field-lines. The minus-sign indicates that the NMR lines of hydrogen atoms located inside the cone in the three-dimensional protein structure are shifted 'upfield' (to the right in the spectra of Figure 5), whereas for atoms outside of the cone the shifts are 'downfield'.

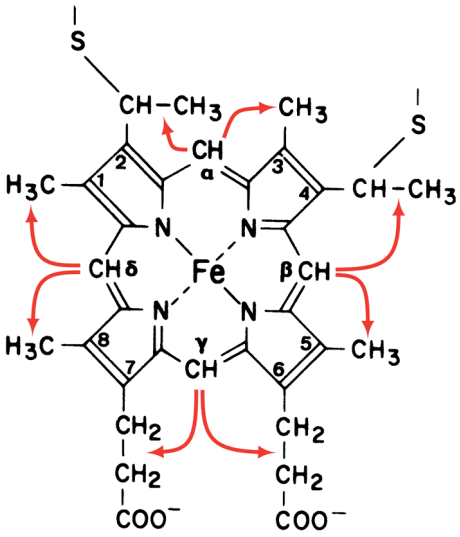


Figure 9. Chemical structure of heme c, which is the prosthetic group in cytochromes c. The red arrows connect groups of hydrogen atoms in the covalent structure of heme c that are separated by a sufficiently short distance to be connected by nuclear Overhauser effects (NOE).

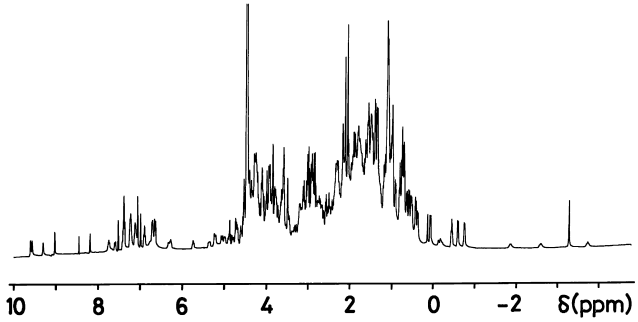


Figure 10. 1D ^{1}H NMR spectrum (360 MHz, 52 °C, solvent $^{2}H_2O$) of horse ferrocytochrome c (M ≈ 12000).

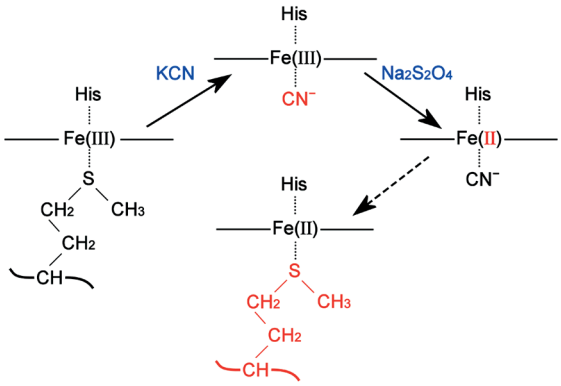


Figure 11. Three chemical reaction steps with cytochrome c that were used to identify methionine as one of the axial ligands of the heme iron (Wüthrich, 1969). The roman numerals indicate the oxidation state of the heme iron, and His stands for the amino acid residue histidine. The solid and broken arrows indicate rapid and slow reaction steps, respectively. Colour code: black, heme iron with ligands, where the horizontal line through Fe represents a side view of the heme c ring (Figure 9); blue, chemicals added to the protein solution; red, new structural features resulting from the preceding reaction.

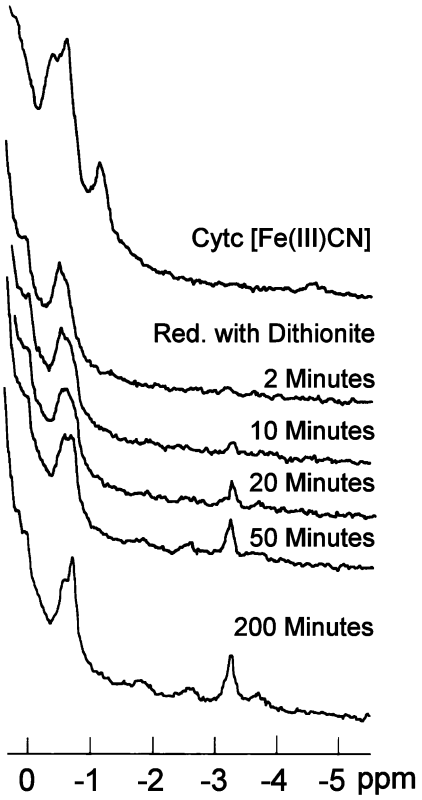


Figure 12. Changes with time in the 1D ^{1}H NMR spectrum of cytochrome c after reducing the ferric cyanide complex with dithionite (220 MHz, 9 °C, solvent ^2H_2O).

by the appearance of its typical high-field lines in the ^{1}H NMR spectrum (Figure 10). This experiment resulted in the identification of the nature of this axial ligand in the native, active form of cytochrome c, and further yielded information on the relative thermodynamic stabilities of the four different structures in Figure 11.

As a sideline, comparison of Figures 10 and 12 illustrates the influence of various experimental factors on the appearance of the NMR spectra. The higher resonance frequency and the high temperature used in the experiment of Figure 10, which was recorded around 1980 with Fourier transform spectroscopy, resulted in narrower and more clearly separated peaks in the region 0 to −5 ppm than the experiment in Figure 12, which was recorded in 1968 with continuous-wave NMR spectroscopy at low temperature.

Other early experiments based on the observation of well-resolved resonances in the ^{1}H NMR spectra of hemoproteins resulted in the characterization of conformation changes during the oxygenation of myoglobin and hemoglobin (Shulman et al., 1969; Davis et al., 1971), and to detailed characterization of the electronic structure of the heme groups in different classes of hemoproteins (Kurland et al., 1968; Wüthrich, 1970; Shulman et al., 1971).

The way to NMR structures of proteins

Around 1970 there was a lot of enthusiasm in my laboratory and elsewhere about the success of experiments of the types illustrated by Figures 5–12. During the following years, however, there was little further progress toward an NMR method for *de novo* structure determination of proteins. In hindsight this can readily be rationalized: Early successful structural interpretations of NMR data invariably supplemented a previously known low resolution X-ray crystal structure of the same protein. In spite of the high symmetry of the ring current fields (Figure 8) and other sources of local conformation-dependent chemical shifts, including natural paramagnetic centers and extrinsic 'paramagnetic shift reagents' (Wüthrich, 1970, 1976), the observed chemical shifts could thus in some instances yield exciting new information. In a *de novo* protein structure determination, however, the high symmetry of the local magnetic fields would lead to ambiguities in the structural interpretation of the resulting chemical shifts. Different approaches were therefore called for, and eventually a NMR method for protein structure determination could be based on the following four principal elements:

Measurement of NOE upper distance limits as conformational constraints

Nuclear Overhauser effects (NOE) are due to dipolar interactions between different nuclei. The intensity of the NOE is related to the product of the inverse sixth power of the internuclear distance and a correlation function, $f(\tau_c)$, which describes the modulation of the dipole–dipole coupling by stochastic rate processes, with an effective correlation time τ_c:

$$\mathrm{NOE} \alpha \frac{1}{\langle r \rangle^6} f\,(\tau_c) \qquad [1]$$

Although the NOE is a common phenomenon for all combinations of closely spaced nuclear spins, NOEs between pairs of hydrogen atoms are of prime interest for structural studies. A ^{1}H–^{1}H NOE is related to the 'through-space' distance between a pair of atoms that

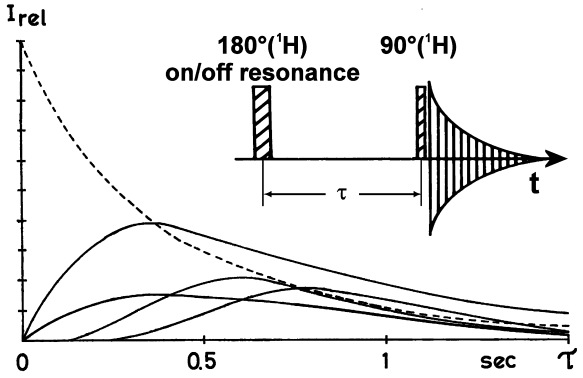


Figure 13. NOE build-up curves observed with 1D transient [^{1}H,^{1}H]-NOE measurements (Gordon and Wüthrich, 1978). The relative intensities, I_{rel}, of the pre-irradiated resonance line (broken curve) and of the lines experiencing NOEs (solid curves) are plotted *versus* the duration of the mixing period, τ. The experimental scheme used to record transient NOE spectra (inset in the upper right) includes a selective 180° radio-frequency pulse applied to the 'pre-irradiated' ^{1}H NMR line, and a non-selective 90°-pulse applied to the entire ^{1}H NMR spectrum. The two pulses are separated by the mixing period, τ, and followed by signal acquisition. When working with the crowded spectra of proteins, one records the difference between two spectra obtained with the pre-irradiation 180°(^{1}H)-pulse applied on- and off-resonance, respectively.

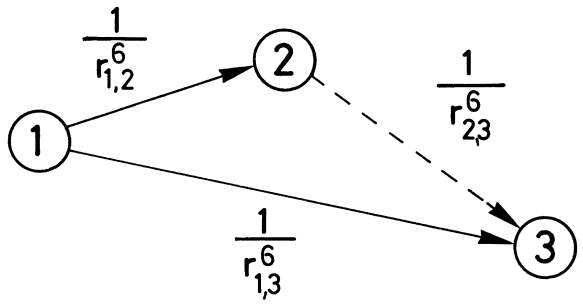


Figure 14. Spin diffusion. Transfer of magnetization between two hydrogen atoms 1 and 3 in the presence of additional hydrogen atoms goes through two competing paythways, i.e., the direct NOE across the distance $r_{1,3}$, and two- or multiple-step spin diffusion *via* intervening hydrogen atoms.

are either not at all linked by covalent bonds (intermolecular NOE), or that may be far apart in the amino acid sequence of a polypeptide chain. ^{1}H–^{1}H NOEs can be observed in double-irradiation 1D NMR experiments as the fractional change in intensity of a NMR line that results from preirradiation of another resonance (inset in Figure 13). Although NOE distance measurements had successfully been used for studies of small organic molecules (Noggle and Schirmer, 1971), including oligopeptides (Gibbons et al., 1972), and observation of NOEs in proteins had also been reported (Gupta and Redfield, 1970; Balaram et al., 1972), it was not clear whether the desired distance information could actually be obtained in proteins (Solomon, 1955; Noggle and Schirmer, 1971; Kalk and Berendsen, 1976). This uncertainty arose because the Brownian motions of large structures in solution are slow, with long effective correlation times, τ_c, for this stochastic process, and proteins contain a dense network of hydrogen atoms. With the combination of these two features, 'spin diffusion' could partially or fully deteriorate distance measurements based on ^{1}H–^{1}H NOE experiments (Solomon, 1955; Abragam, 1961; Kalk and Berendsen, 1976; Gordon and Wüthrich, 1978; Wagner and Wüthrich, 1978).

In the solution NMR experiments discussed here, spin diffusion arises as a consequence of the dependence of the NOE on the inverse sixth power of the internuclear distance, since magnetization transfer between two spins through multiple short steps may be more efficient than a one-step transfer over the longer, direct distance (Figure 14). We used 1D transient NOE difference experiments (inset in Figure 13) (Gordon and Wüthrich, 1978) and 1D truncated-driven NOE difference experiments (TOE) (Wagner and Wüthrich, 1979) to record 'NOE build-up curves' (Figure 13), which then provided a basis for ^{1}H–^{1}H distance measurements in macromolecules. These studies showed that even more favourable conditions for NOE distance measurements can be found in macromolecules, which have long effective correlation times for the modulation of dipole–dipole couplings, than in low-molecular-weight compounds, for which the condition of 'extreme motional narrowing' applies (Solomon, 1955; Abragam, 1961).

Analysis of ^{1}H–^{1}H NOE build-up curves recorded with 1D transient NOE experiments (Figure 13) is based on the consideration that during the early phase of the mixing period, direct magnetization transfer from spin 1 to spin 3 (Figure 14) will result in a linear increase of the magnetization on spin 3 with time. In contrast, transfer to atom 3 via an intermediate step through atom 2 will have a lag period, even though for long mixing periods, τ, the extent of the magnetization transfer by such 'spin diffusion' may for individual pairs of hydrogen atoms exceed that of the direct transfer (Figure 13). Corresponding considerations apply for the analysis of NOE build-up curves recorded with truncated-driven NOE experiments (Wagner and Wüthrich, 1979). With proper selection of the duration of the mixing period, one can thus measure highly selective ^{1}H–^{1}H NOEs between distinct pairs of hy-

drogen atoms, or groups of chemical shift-equivalent hydrogen atoms in proteins in solution.

There is a second criterion that needs to be satisfied for obtaining selective NOEs with 1D NMR experiments, i.e., at least one of the two lines that are connected by the NOE must be sufficiently well resolved (i.e., separated from all other lines in the spectrum) to enable selective radio-frequency pre-irradiation (inset in Figure 13). The previously discussed well-resolved lines in the ^{1}H NMR spectrum of ferrocytochrome c (Figure 10) thus had again an important role in enabling us to study the spin physics in the interior of this globular protein with 1D ^{1}H NMR experiments, as well as to obtain novel structural information. Figure 15 shows a series of highly selective 1D truncated-driven NOE measurements, which were used to determine ^{1}H NMR assignments for the heme group in a c-type cytochrome (Figure 9) (Keller and Wüthrich, 1978).

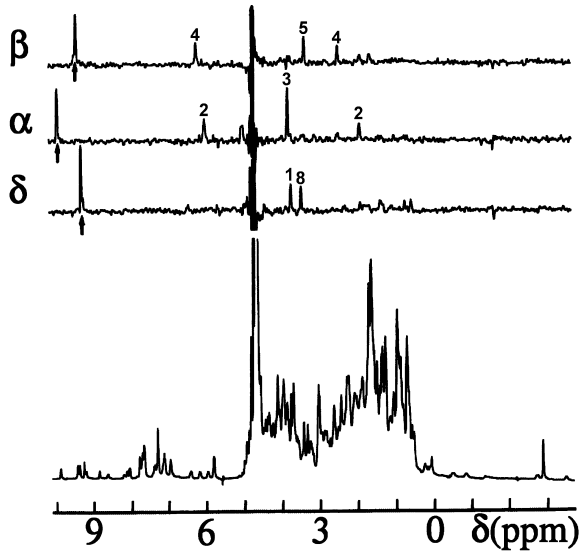


Figure 15. ^{1}H NMR assignments for the heme group of a c-type cytochrome using ^{1}H-^{1}H NOEs (see Figure 9). Bottom trace: 1D ^{1}H NMR spectrum of ferrocytochrome c-551 from *Pseudomonas aeruginosa* (M ≈ 12000). Upper traces: three 1D truncated-driven NOE difference spectra obtained with selective pre-irradiation (indicated by arrows) on the hydrogen atoms β, α and δ (see Figure 9), respectively (360 MHz, 35 °C, ^2H_2O-solution). The peaks in the NOE difference spectra are identified with numbers indicating the substituents attached to the corresponding porphyrin ring carbons (Figure 9).

Sequence-specific resonance assignments

Similar to the situation in heme groups (Figure 9), there are closely spaced pairs of hydrogen atoms in neighbouring residues of a polypeptide chain (Figure 16). These can be connected by the observation of 'sequential NOEs'. Figure 16 illustrates that NOE-based ^{1}H NMR assignments for a polypeptide chain can conceptually be considered as a two-step process. Each amino acid residue represents a 'spin system' (Wüthrich, 1986), i.e., it consists of an array of hydrogen atoms including an amide proton (H^N), an α-proton (H^α), and the side chain protons, which can be connected by steps over three or less covalent bonds through the observation of scalar spin–spin ('through-bond') couplings (see (Wüthrich, 1986, for the exceptions represented by proline and the aromatic side chains). In contrast, hydrogen atoms located in sequentially neighbouring amino acid residues are separated by at least four covalent bonds. Pairs of neighbouring residues in the sequence can therefore only be connected via NOEs manifesting short through-space distances, such as $d_{\alpha N}$ and d_{NN}. Suitable combinations of intraresidual ^{1}H–^{1}H connectivities established by scalar spin–spin couplings, and inter-residue connectivities established by sequential NOEs enable progressive resonance assignments while 'walking along the polypeptide backbone'. In other words, the spin systems of two neighbouring residues can be connected by the intervening H^α–H^N or H^N–H^N sequential NOE connectivities (Figure 16).

1D double-resonance experiments with selective irradiation of the well-resolved amide proton resonances between 8 and 10 ppm in folded BPTI (Figure 5) thus yielded assignments for most of the residues in the regular secondary structure elements (Figure 17) (Dubs et al., 1979). Further assignments were not possible, because the other regions of the polypeptide chain were not represented by well-separated ^{1}H NMR lines that could have been selectively irradiated in 1D NMR experiments.

Two-dimensional (2D) NMR

With the introduction of 2D NMR experiments, and subsequently 3D and 4D NMR experiments, NMR studies of biological macromolecules evolved from intellectually stimulating science to a practical approach for protein structure determination. The foundations of multi-dimensional NMR have been presented in the Nobel Lecture by Ernst (1992). Here, I only want to comment on two crucial consequences of multi-dimensional NMR for studies of proteins. First, 2D ^{1}H NMR enables the recording of selective interactions between pairs of hydrogen atoms, or groups of

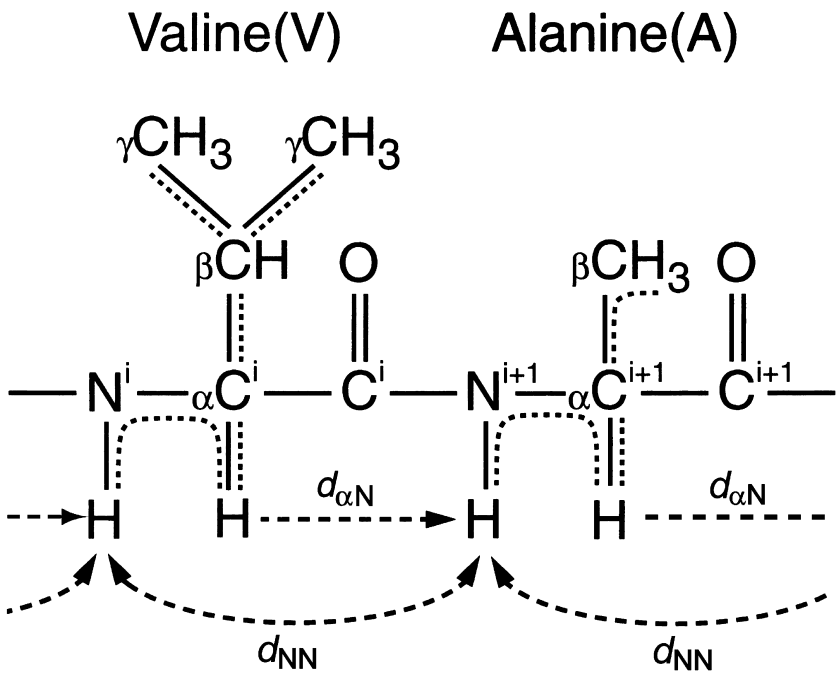


Figure 16. Sequential ^{1}H NMR assignment of proteins. The drawing shows the chemical structure of a – valine – alanine – dipeptide segment in a polypeptide chain. The dotted lines connect groups of hydrogen atoms that are separated by at most three chemical bonds and can therefore be connected using scalar spin–spin couplings. The broken arrows link pairs of hydrogen atoms in neighbouring amino acid residues that are separated by short through-space distances, $d_{\alpha N}$ and d_{NN}, and can therefore be connected by 'sequential NOEs'.

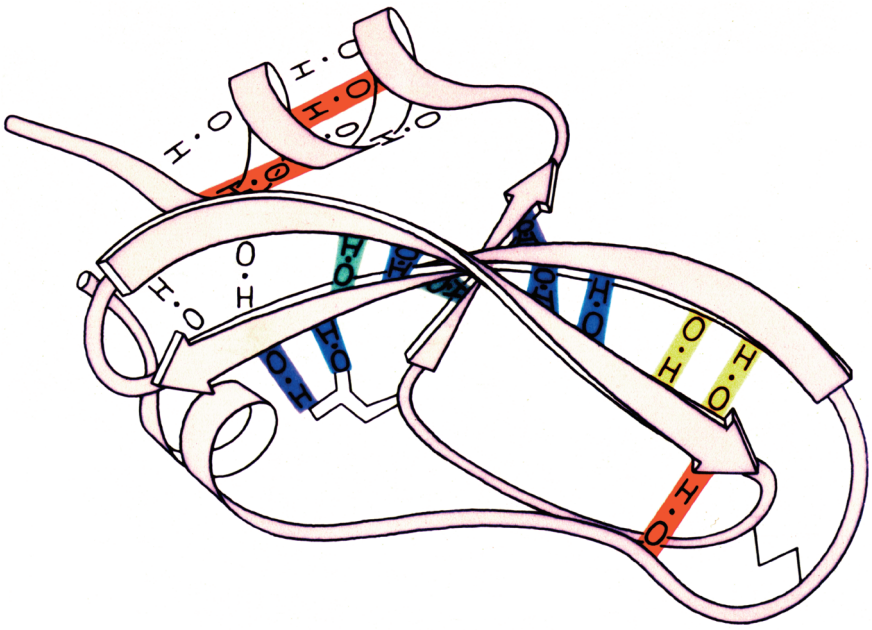


Figure 17. Sequence-specific resonance assignments for BPTI that were obtained using 1D ^{1}H NMR experiments (Dubs et al., 1979). The polypeptide backbone is shown, and the assigned residues are identified by indication of hydrogen bonds with their amide protons. (The colour code indicates variable exchange rates of the amide protons with the solvent; drawing by Jane Richardson, 1980.)

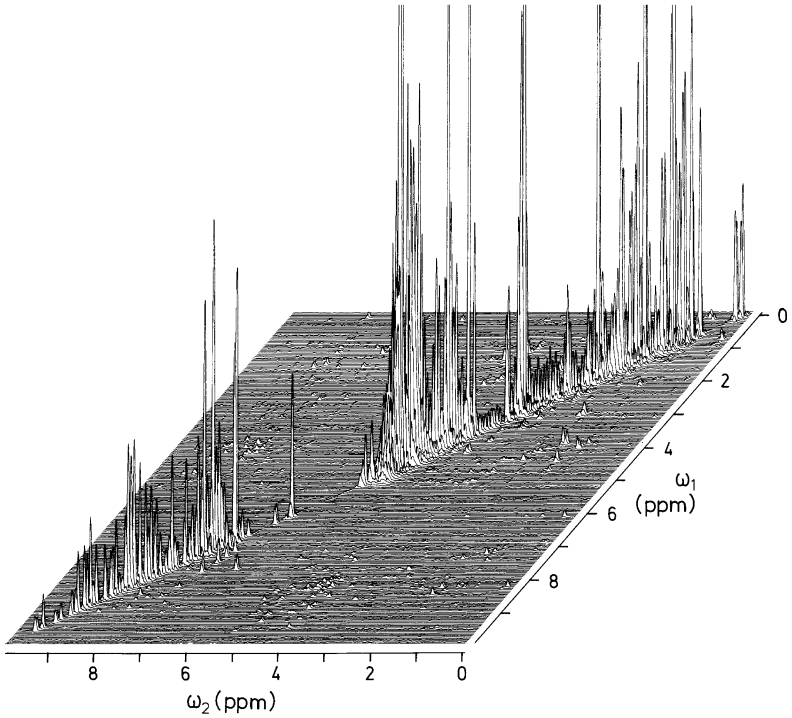


Figure 18. Two-dimensional (2D) [^{1}H,^{1}H]-NOE spectroscopy ([^{1}H,^{1}H]-NOESY). A stacked plot representation of a spectrum of the small protein bull seminal proteinase inhibitor IIA (BUSI IIA, M $\approx$ 6000) is shown (500 MHz, 45 °C, H_2O-solution).

chemical shift-equivalent hydrogen atoms, without selective irradiation of individual resonance lines. It thus enables a detailed analysis of the entire ^{1}H NMR spectrum of a protein, which contrasts with the 1D NMR situation of being limited to using only a small number of resolved lines at the periphery of the spectrum. Second, the dispersion of the resonances in a two-dimensional frequency plane affords greatly improved separation of the individual peaks. For example, in Figure 18 the intense lines on the diagonal from the lower left to the upper right correspond to the 1D NMR spectrum, and the weak 'cross peaks' in the plane outside of the diagonal manifest selective NOEs between pairs of hydrogen atoms, which are represented in the NMR spectrum by two distinct (but not necessarily 'resolved') chemical shift positions along the diagonal. The greatly improved separation of the individual cross peaks is best seen in contour plots of 2D NMR spectra (Figure 19), which is the presentation used for detailed analysis. The 1D decoupling experiments for the identification of scalar spin–spin couplings were thus replaced by 2D correlation experiments, such as COSY, SECSY and FOCSY, and the 1D NOE difference experiments were replaced by 2D

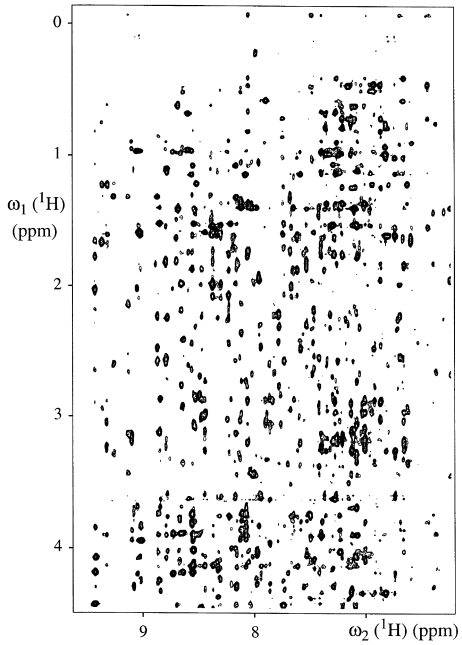


Figure 19. 2D [^{1}H,^{1}H]-NOESY spectrum of the plant pathogenesis-related protein P14A (M $\approx$ 15000). A contour plot of the spectral region [ω_1(^{1}H) = 0–4.3 ppm, ω_2(^{1}H) = 6.3–9.5 ppm] is shown (750 MHz, 30 °C, H_2O-solution).

[^{1}H,^{1}H]-NOESY, which yielded similar NOE build-up curves for each NOESY cross peak (Figure 20) as the 1D transient NOE technique (Figure 13) (Wider et al., 1984; Wüthrich, 1986; Ernst et al., 1987). Using these 2D ^{1}H NMR experiments, complete sequence-specific resonance assignments could be obtained for the protein BPTI (Figure 21) (Wagner and Wüthrich, 1982).

The 'NOESY-COSY connectivity diagram' in Figure 22 (Wagner et al., 1981; Wüthrich, 1986) makes use of the fact that standard 2D [^{1}H,^{1}H]-COSY and 2D [^{1}H,^{1}H]-NOESY spectra are symmetrical with respect to the diagonal peaks. Combining the upper left half of a NOESY spectrum with the lower right half of a COSY spectrum therefore enables a straightforward visualization of assignments by a succession of $d_{\alpha N}$ sequential NOEs and intraresidual H^N–H^α scalar coupling connectivities. The data were recorded in a freshly prepared solution of BPTI in ^2H_2O, where only the resonances of the slowly exchanging amide protons are seen (Wagner et al., 1981). The assignments start with the COSY cross peak identified by a black square and, as indicated by the arrows, go clockwise to the sequentially preceding isoleucine residue, and counter-clockwise to a sequence of four residues. It is customary that the data leading to the ^{1}H NMR assignments of a protein are collected in a plot *versus* the amino acid sequence (Figure 23). Figure 23 shows that most of the sequential connectivities are independently documented by two or three different sequential NOEs (Billeter et al., 1982; Wüthrich, 1986), and that possible remaining gaps in the assignment pathway are readily recognized in this presentation.

Structural interpretation of NOE distance constraints

A polypeptide chain with 100 amino acid residues has a length of about 400 Å, whereas NOE-observable distances are shorter than about 5 Å. Observation of a NOE between a pair of hydrogen atoms with assigned chemical shift positions therefore enforces the formation of a ring-like structure (Figure 24). A successful structure determination generates three-dimensional arrangements of the polypeptide chain that simultaneously contain all the small and large circular structures imposed by the ensemble of all NOESY cross peaks.

Partial structure determination has been obtained by an empirical approach for the identification of regular secondary structures in polypeptide chains, which relies on recognizing distinct patterns of NOEs (Wüthrich et al., 1984; Wüthrich, 1986). For example, in Figure 23 a succession of strong sequential d_{NN} NOEs in combination with the observation of medium-range NOEs in the same polypeptide segment identifies three α-helical structures, which are independently also indicated by successions of small values for the scalar spin–spin couplings $^3J_{HN\alpha}$ (Pardi et al., 1984; Wüthrich, 1986).

For the calculation of complete three-dimensional protein structures from NMR data (Wüthrich et al., 1982), it was quite clear from the outset that an input of quantitative NOE distance measurements would be difficult to obtain. The observed NOEs depend on the proton–proton distance, r, as well as on the effective rotational correlation times, τ_c (Equation [1]). Since for each pair of hydrogen atoms the effective τ_c-value is governed not only by the overall rotational molecular tumbling (Brownian motions), which depends on the size and shape of the protein as well as on the viscosity of the solvent, but can also be affected by intramolecular motions, $f(\tau_c)$ may vary for different pairs of hydrogen atoms in a protein molecule. Additional ambiguities could arise from partial quenching of individual ^{1}H–^{1}H NOEs by competitive spin relaxation processes, for example, spin diffusion (Figure 14), chemical or conformational exchange, and interactions with other nuclear or electronic spins. Furthermore, as a result of rapid intramolecular mobility, a given NOE may be the result of sampling over a range of distances between the two hydrogen atoms of interest (Braun et al., 1981). In view of these intrinsic limitations for efficient quantitative NOE distance measurements, we decided to use a constant value of the correlation function, $f(\tau_c)$ (Equation [1]), for all ^{1}H–^{1}H combinations in a protein, and to derive only upper limits on the ^{1}H–^{1}H distances from the NOE measurements. In practice, the input for a structure calculation then consists of allowed distance ranges, which are bounded by a NOE upper limit of 3.0 to 5.0 Å, depending on the intensity of the NOE, and a lower limit of 2.0 Å, which represents the sum of the van der Waals radii of the two NOE-connected hydrogen atoms. Although each individual entry in the input data thus has only limited precision, this procedure is robust and can conceptually account for the influence of intramolecular mobility in most of the situations that are commonly expected for the structured parts of globular proteins.

For the initial globular protein structure calculations from NMR data (Figure 25), we used a metric matrix distance geometry algorithm to search for molecular geometries that are consistent with the

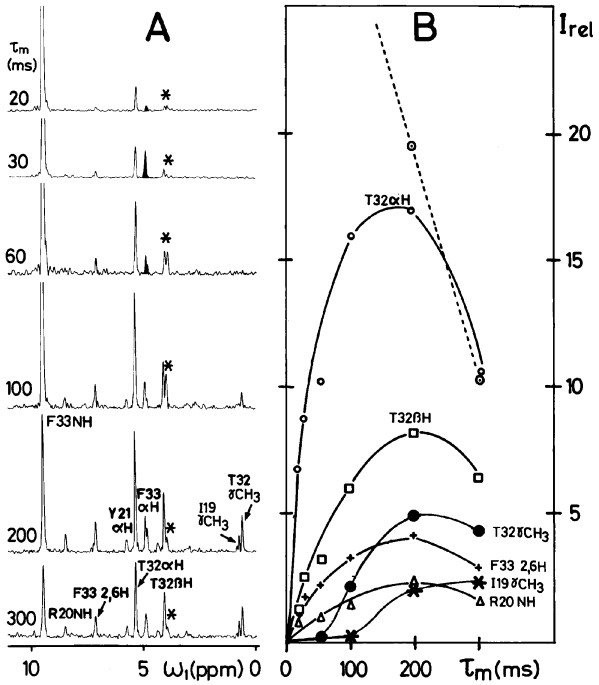


Figure 20. Measurement of NOE build-up curves using 2D [^{1}H,^{1}H]-NOESY experiments. (A) Six NOESY spectra of BPTI recorded with different mixing times, as indicated on the left. The same 1D cross-section along the ω_1 frequency axis through the diagonal peak of the Phe 33 amide proton is plotted for each spectrum. (B) NOE build-up curves obtained from the analysis of the data in (A); same presentation as in Figure 13, with the relative intensities of the NOESY cross peaks plotted versus the length of the mixing time, τ_m, and the broken line representing the decay of the magnetization on the diagonal peak of the Phe 33 amide proton.

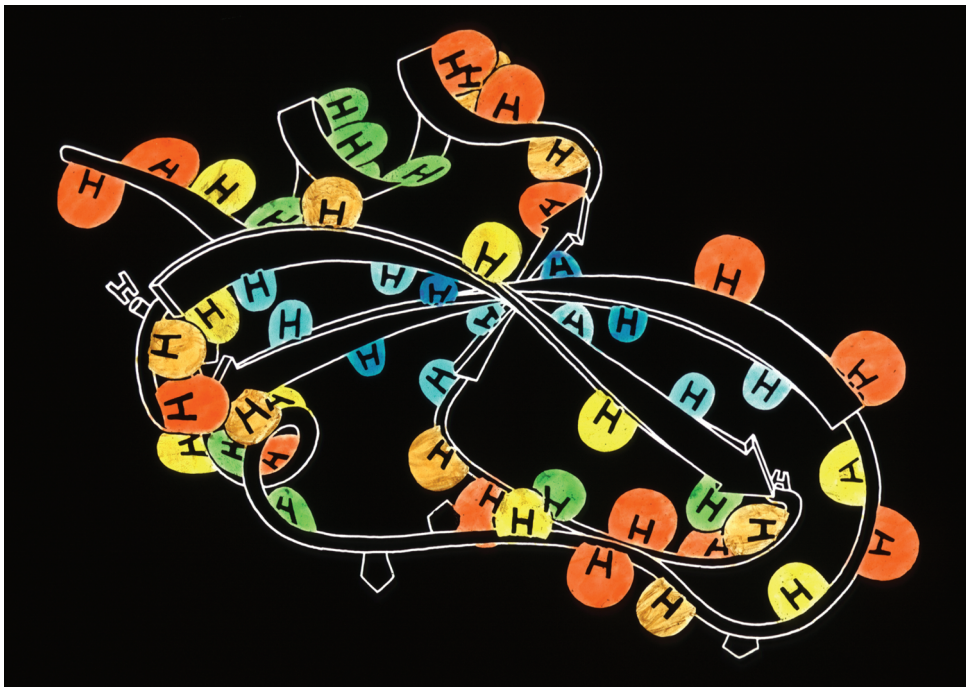


Figure 21. Complete sequence-specific resonance assignments for BPTI obtained using 2D ^{1}H NMR experiments (Wagner and Wüthrich, 1982). Assigned residues are identified by colored patches covering their amide protons. (The colour code indicates variable amide proton exchange rates; drawing by Jane Richardson, 1982.)

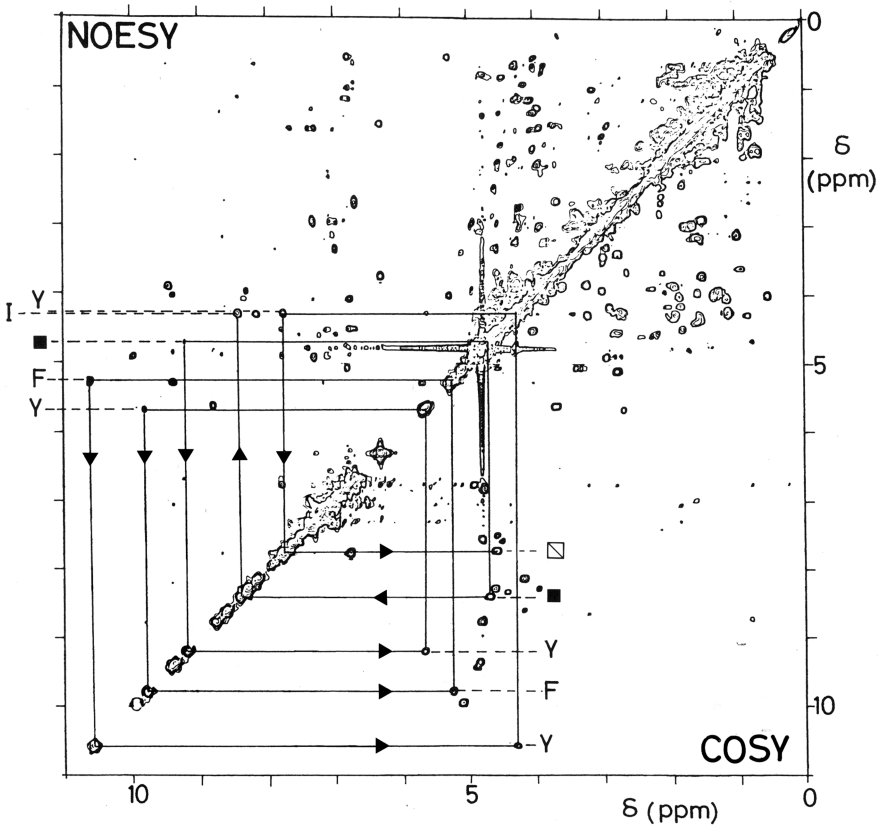


Figure 22. NOESY-COSY connectivity diagram for sequential ^{1}H NMR assignments, using the $d_{\alpha N}$ sequential NOEs. The data was recorded with BPTI in ^2H_2O solution. (see text).

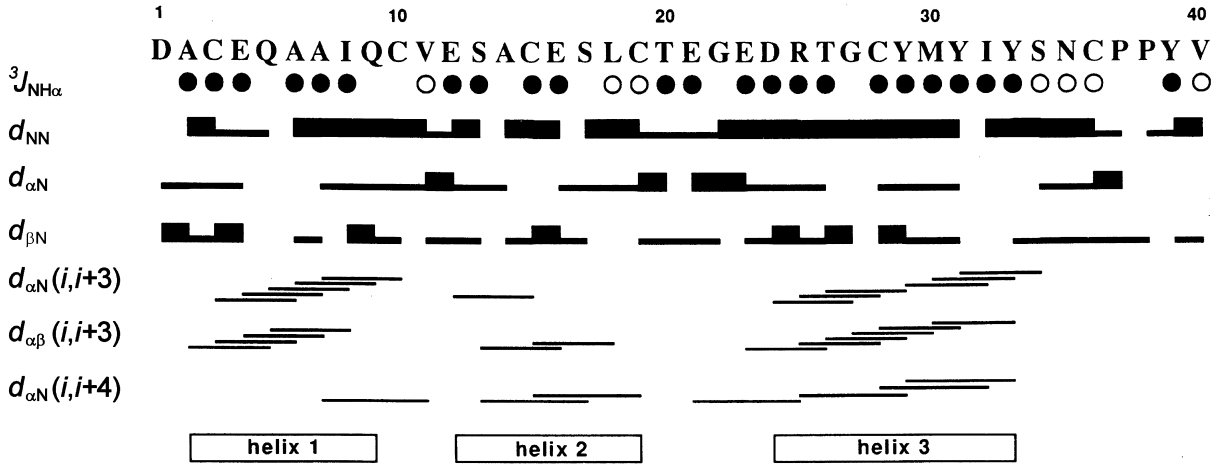


Figure 23. Standard presentation of ^{1}H NMR data leading to sequence-specific resonance assignments and the identification of regular secondary structures in proteins. Experimental data plotted *versus* the amino acid sequence are shown for the small pheromone protein E*r*-1 from *Euplotes raikovii* (M $\approx$ 4500). $J_{HN\alpha}$ are scalar spin–spin coupling constants, with small and big values indicated by filled and empty circles, respectively. d_{NN}, $d_{\alpha N}$ and $d_{\beta N}$ are distances manifested in sequential NOEs, where strong and weak sequential NOEs are indicated by a thick or a thin line, respectively. Small values of the distances $d_{\alpha N}(i,i+3)$, $d_{\alpha\beta}(i,i+3)$ and $d_{\alpha N}(i,i+4)$ are observed by medium-range NOEs linking the given atom types between residues spaced as indicated in the parentheses and by the short horizontal lines. The locations of three α-helices are indicated at the bottom.

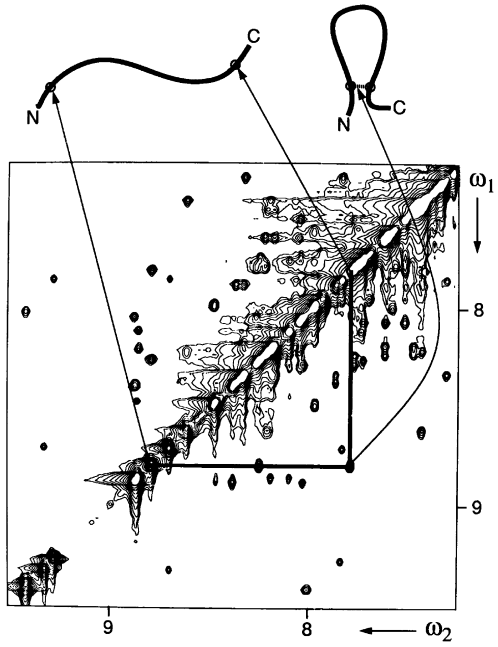
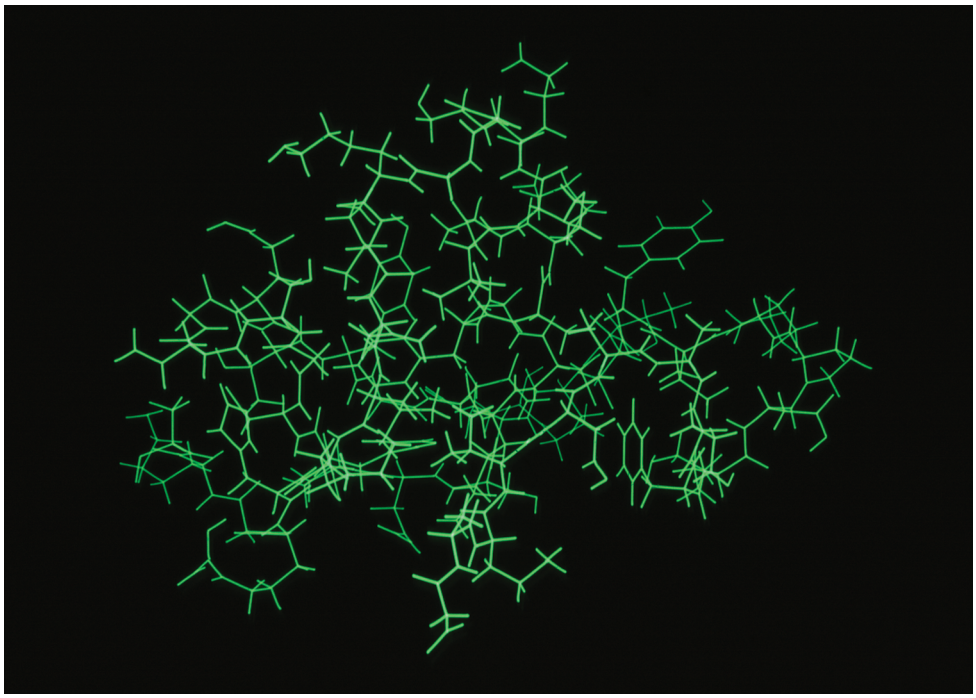


Figure 24. Scheme indicating the relations between an experimental 2D [^{1}H,^{1}H]-NOESY spectrum, a polypeptide with the chain ends indicated by N and C, sequence-specific assignments for two hydrogen atoms in the polypeptide chain indicated by circles, and the NOE upper distance constraint derived from the NOESY cross-peak connecting the chemical shift positions of the two assigned hydrogen atoms (see text).

ensemble of all experimentally determined NOE distance constraints (Havel and Wüthrich, 1984, 1985; Williamson et al., 1985). Each such calculation ends with the minimization of an error function, and the residual error function value represents a straightforward measure for the success of having found a molecular geometry that satisfies the experimental input data. In view of the aforementioned distance-range format of the input, it is further of keen interest to evaluate the uniqueness of the calculated structure. To this end, the structure calculation is repeated with identical input data but different boundary conditions, and the uniqueness of the resulting NMR structure is judged from the tightness of the fit among the resulting ensemble of conformers. Typically, about 100 conformers are generated, and a sub-group of the 20 conformers with the smallest residual error function values is selected to represent the NMR structure of the protein. The average of the pairwise root-mean-square distances (RMSD) calculated for this bundle of conformers (Figure 1) is then taken as a measure for the precision of the structure determination. Visually, a tight fit of the bundle of conformers indicates regions where the structure is defined with high precision by the NMR data, whereas structurally disordered polypeptide segments show a large dispersion among the members of the bundle, as exemplified by the two chain ends of the *Anten-*



Figuur 25. NMR structure of BUSI IIA (Williamson et al., 1985).

napedia homeodomain in Figure 1. In the absence of long-range NOE distance constraints, a properly functioning algorithm for the structure calculation will sample the entire conformation space that is accessible with the given length of the polypeptide chain, as exemplified by the 'unstructured' tail of the bovine prion protein in Figure 4.

NMR in structural biology

Standard protocol for NMR structure determination of biological macromolecules

The protocol for NMR structure determination includes the preparation of a homogeneous protein solution, the recording and handling of the NMR data sets, and the structural interpretation of the NMR data (Table 1). The techniques used in 1984 for the structure determination of bull seminal proteinase inhibitor IIA are listed in Table 1; the four steps of the structural interpretation (III in Table 1) were performed separately, although the result of the first round of constraint collection and structure calculation was subsequently used for additional checks on the sequence-specific resonance assignments as well as on the collection of conformational constraints. Since 1984, the protocol outlined in Table 1 has been used in over 3000 NMR structure determinations of proteins and nucleic acids, and greatly improved experimental techniques have been incorporated into this general scheme.

Major advances in the experimental techniques for NMR structure determination were spurned by the introduction of methods for the production of recombinant proteins labelled with stable isotopes, in particular carbon-13, nitrogen-15 and deuterium (Wagner, 1993; Kay and Gardner, 1997). For example, this opened the way for efficient use of heteronuclear NMR techniques with proteins, such as 3D [^{1}H,^{13}C,^{15}N]-triple resonance experiments, 3D ^{13}C- or ^{15}N-resolved [^{1}H,^{1}H]-NOESY (Figure 26) (Bax and Grzesiek, 1993; Wider, 1998), and the use of heteronuclear filters (Otting and Wüthrich, 1990). Important advances have also been made with the methods of structure calculation, where the c.p.u. time needed for the calculation of a small protein structure has been reduced from about one day in 1984 (Havel and Wüthrich, 1984, 1985) to a few seconds (Güntert, 1998). Currently, intense work is focused on the automation and combined execution of the individual steps in the structural interpretation of the NMR data (Moseley and Montelione, 1999; Herrmann et al., 2002). It is beyond the scope of this article to present in detail the wide range of beautiful novel experimental approaches developed by the community of macromolecular NMR spectroscopists during the past 15 years, which now enable studies with ever more complex systems (Figures 1–4).

Table 1. Standard protocol for NMR structure determination of proteins

Step[a]		BUSI IIA[b]
I	Sample preparation	Protein isolated from natural source; natural isotope distribution; 16 mM solutions in H_2O and in $^{2}H_2O$, respectively
II	NMR spectroscopy	2D ^{1}H NMR
IIIa	Resonance assignments	Sequential NOEs
IIIb	Conformational constraints	[^{1}H,^{1}H]-NOEs, $^{3}J_{HN\alpha}$, $^{3}J_{\alpha\beta}$
IIIc	Structure calculation	Metric matrix distance geometry
IIId	Structure refinement	Restrained energy minimization

[a]The structural interpretation of the NMR data, III, is somewhat arbitrarily divided up into four steps; in practice, one goes through multiple cycles of collection of conformational con straints (IIIb) and structure calculation (IIIc), and the completion of the sequence-specific assignments (IIIa) as well as the structure refinement (IIId) may also be part of this iterative approach.

[b]This column lists the techniques used in the first structure determination of a globular protein in 1984 (Williamson et al., 1985).

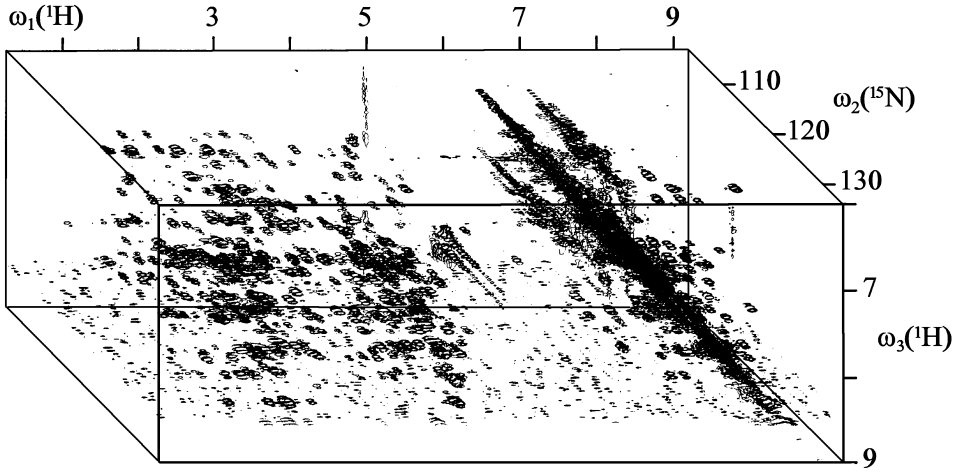


Figure 26. Three-dimensional (3D) ^{15}N-resolved [^{1}H,^{1}H]-NOESY spectrum (600 MHz, 28 °C, H_2O-solution) of the DNA-binding domain of the P22 c2 repressor (M ≈ 10000, uniformly ^{15}N-labeled).

Globular protein structures in solution

The static picture of a protein molecule obtained from the standard protocol for structure determination (Table 1) typically shows variable precision of the structure determination along the polypeptide chain, as manifested by the variations in the closeness of the fit among the bundle of conformers used to represent the NMR structure (Figure 1). Even in proteins where the entire polypeptide chain is part of the global fold, increased disorder is observed toward the periphery of the surface side chains (Figure 27). This pronounced surface disorder, which typically also involves the ends of the polypeptide chain, is in most instances the only significant difference between corresponding globular protein structures in single crystals and in solution (Wüthrich, 1995a). In part the increased surface disorder in solution arises because of a scarcity of packing constraints, when compared with the protein core, and a concomitant scarcity of NOE distance constraints. With the additional use of NMR spin relaxation measurements (Solomon, 1955; Abragam, 1961; Ernst et al., 1987; Luginbühl and Wüthrich, 2002), one can distinguish between static disorder, which would then presumably arise from the scarcity of constraints, and dynamic disorder, with intramolecular motions on the nanosecond to sub-nanosecond time scale. Overall, quite independent of the dynamics issue, the observation of partially folded polypeptide chains in solution (Figures 1, 4 and 27) is important complementary information to the data that can be obtained by studies in crystals. It is also the main reason why the quality of a NMR structure determination is not usually characterized by a single, global parameter (Wüthrich, 1986, 1995a, b).

An important extension of the characterization of proteins in solution resulted from high resolution NMR studies of protein hydration. Thereby the location of hydration waters is determined by the observation of NOEs between water protons and hydrogen atoms of the polypeptide chain (Otting et al., 1991). Because of the dependence of the NOE on the inverse sixth power of the ^{1}H–^{1}H distance, only one layer of hydration water molecules is observed (Figure 28). For the hydration studies, the dependence of the NOE intensity on the correlation function describing the stochastic modulation of the dipole–dipole coupling between the interacting protons (Equation [1]) has a key role. The value of $f(\tau_c)$ may be governed either by the Brownian rotational tumbling of the hydrated protein molecule, or by interruption of the dipolar interaction through translational diffusion of the water molecules relative to the protein surface, whichever is faster. On this basis it could be established that surface hydration of peptides and proteins is characterized by very short residence times of the water molecules in the hydration sites, in the range from about 20 to 300 picoseconds at 10 °C. This result presents an intuitive rationale for the generally observed dynamic disorder of the protein surface structure (Figure 27), and indicates that the dehydration of the polypeptide surface will hardly ever be a rate-limiting step either in protein folding or in intermolecular interactions with proteins.

The BPTI crystal structure contains four interior hydration water molecules. These are an integral part

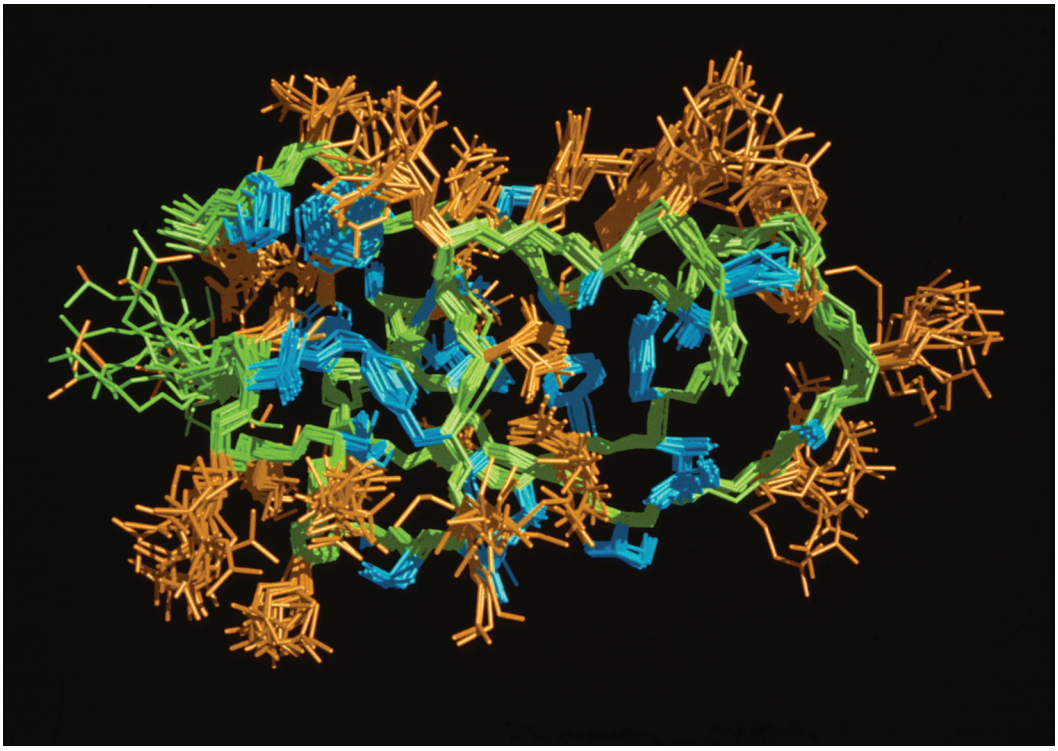


Figure 27. NMR structure of BPTI represented by a bundle of 20 conformers superimposed for best fit of the polypeptide backbone. The polypeptide backbone is green, core side-chains are blue, and solvent-accessible surface side-chains are red.

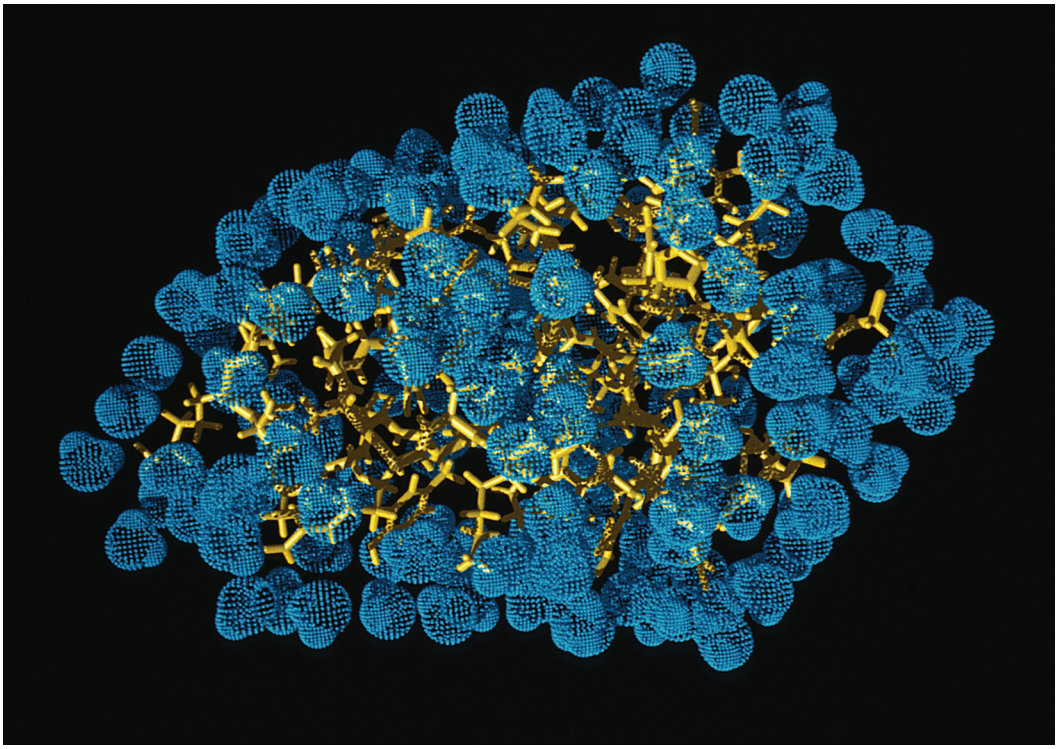


Figure 28. Molecular model of hydrated BPTI in H_2O solution. The drawing shows an all-heavy-atom-presentation of one of the conformers in Figure 27 (yellow) covered with a layer of hydration water molecules (dotted blue spheres).

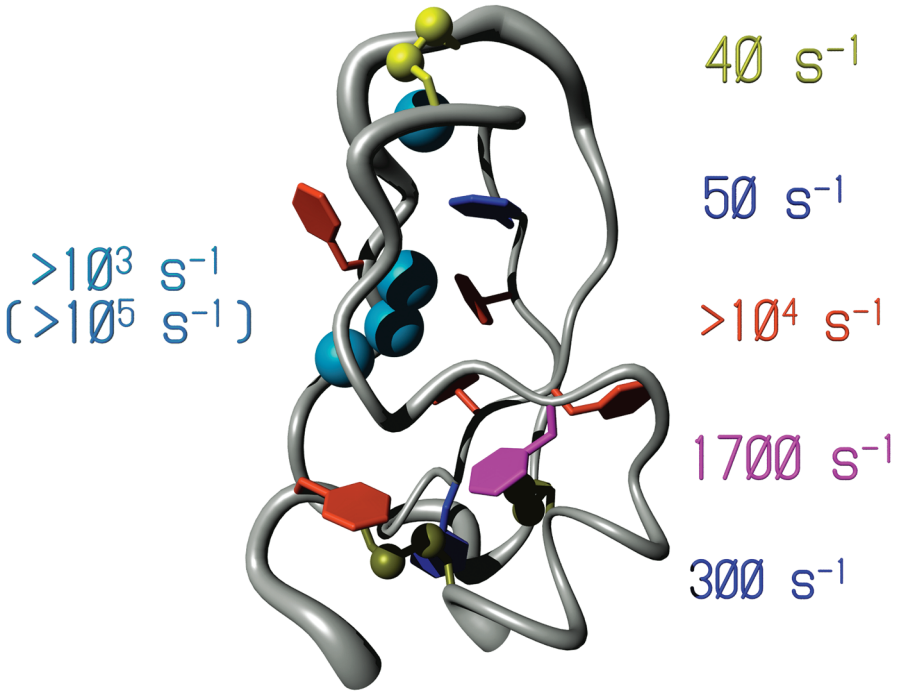


Figure 29. Intramolecular rate processes in BPTI. The polypeptide backbone is represented by a grey spline function through the α-carbon positions, with the thickness of the line representing the spread of the bundle of 20 conformers in Figure 27. Individual rate processes and their frequencies are indicated with the following colour code: dark blue, magenta and red: ring flips of phenylalanine and tyrosine; yellow, exchange between the R- and S-chiral forms of the disulfide bond at the top; cyan, interior hydration water molecules, with indication of the exchange rates with the bulk water.

of the molecular architecture and are inaccessible to the solvent in a rigid model of the three-dimensional structure (Figure 29). The chemical shift dispersion anticipated for these four water molecules on the basis of the considerations in Figures 5–8 was not observed. This degeneracy of the chemical shifts of bound water and bulk water was found to be due to rapid exchange of water molecules in and out of the protein molecule, with an upper limit on the life-time of about 1 millisecond (Otting et al., 1991; as indicated in Figure 29 by the number in parentheses, the actual life times for individual waters may be significantly shorter). Rapid exchange of interior hydration water molecules appears to be a general property of globular proteins, and was also observed for water molecules located in protein–DNA interfaces, for example, in the DNA complex with the *Antennapedia* homeodomain (Figure 2).

Another intriguing NMR observation of internal protein mobility are the 180°-ring flipping motions of phenylalanine and tyrosine (Wüthrich and Wagner, 1975). Observation of these 'ring flips' on the millisecond to microsecond time scale (Figure 29) was a genuine surprise for the following reasons: in the refined X-ray crystal structure of BPTI reported in 1975, the aromatic rings of phenylalanine and tyrosine are among the best-defined side chains, with the smallest temperature factors. In each ring the relative values of the temperature factors for the individual atoms increase toward the periphery, so that the largest positional uncertainty is indicated for the peripheral ζ-carbon atom on the symmetry axis through the C^{β}–C^{γ} bond, rather than for the four δ- and ε-ring carbon atoms, which undergo extensive movements during the ring flips. Theoretical studies then resolved this apparent contradiction: The crystallographic temperature factors sample multiple rotation states about the C^{α}–C^{β} bond, but they do not manifest the ring flips because the populations of all non-equilibrium rotational states about the C^{β}–C^{γ} bond are vanishingly small. Although the flipping motions about the C^{β}–C^{γ} bond have low frequencies (Figure 29), they are very rapid 180°-rotations connecting two indistinguishable equilibrium orientations of the ring. Similar to the exchange of internal hydration waters, the ring flip phenomenon is a general feature of globular proteins,

manifesting ubiquitous low-frequency internal motions that have activation energies of 60 to 100 kJM^{-1}, amplitudes larger than 1 Å, activation volumes of about 50 Å^3, and involve concerted displacement of numerous groups of atoms. Combined with sequence-specific NMR assignments, these experiments provide high spatio-temporal resolution for the description of rate processes in proteins. In Figure 29 this is illustrated by a mapping of the frequencies for ring flips and water exchange onto the NMR structure of BPTI.

In addition to the ring flips and the exchange of internal hydration water molecules, Figure 29 includes data on the exchange of a disulfide bond between the R and S chiral states. In contrast to the other two phenomena, this rate process connects two different molecular structures.

Although all the data collected in Figure 29 have been known for more than a decade, and some of them for nearly three decades, no widely accepted functional interpretation of these low-frequency motional processes has been advanced. The same holds for the conformational equilibria manifested by the 'protection factors' governing the amide proton exchange rates in folded proteins (Figures 17 and 21). Quite possibly these NMR measurements are ahead of their times, and represent a source for future novel insights into structure–function correlations in proteins.

Outlook to NMR applications in structural and functional proteomics

With the availability of a rapidly increasing number of completely sequenced genomes, new challenges arise for the methods used for three-dimensional structure determination. On the one hand, 'structural genomics' initiatives in several leading research centers focus on the development of technology for high-throughput structure determination to generate a comprehensive atlas of protein folds, so that remaining gaps could be filled by structure prediction methods. There is clearly a lot of room to further enhance the efficiency of each

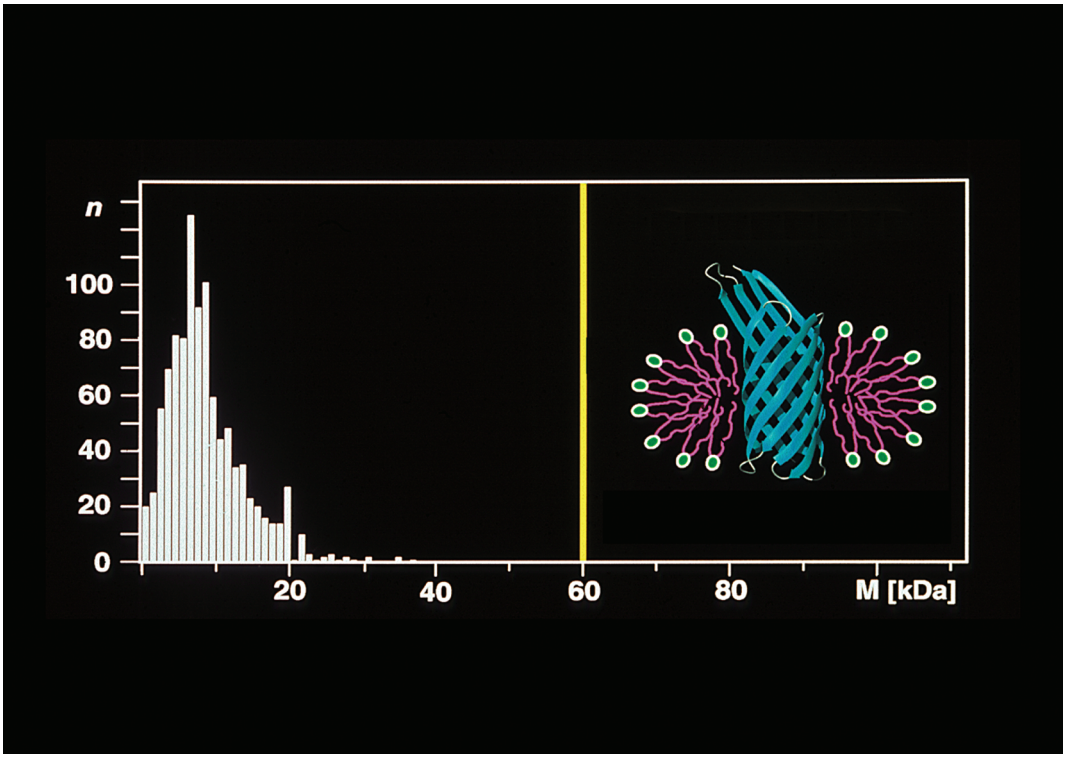


Figure 30. NMR structure determination and molecular weight. The horizontal axis covers the molecular weight-range 0–110 000. The left side shows the molecular weight distribution of the NMR structures in the protein databank (December 2000). On the right, the structure of the protein OmpX from *E. coli* (M ≈ 18 000) in DHPC micelles (M ≈ 70 000 for the mixed micelles) indicates along the horizontal axis the approximate molecular weight range for NMR structure determination of mixed micelles with membrane proteins that is presently accessible with the NMR techniques of Figures 31–33.

step of the NMR structure determination procedure (Table 1). On the other hand, we face the situation that newly determined protein structures should enable us to predict novel functions, whereas in classical structural biology one encounters more typically the challenge of rationalizing known functions on the basis of the three-dimensional structure. This section describes some recent work in our laboratory that may eventually contribute to future strategies for the discovery of new physiological functions from molecular structure data.

It has been widely recognized that supplementing the determination of new protein folds with data on intermolecular interactions may provide a key for the identification of unknown gene functions. Since efficient use of conventional NMR spectroscopy in solution had been limited to particle sizes with molecular weights up to about 30 000 (Figure 30, left side), a new challenge for solution NMR techniques then arose from the fact that the supramolecular structures resulting from interactions of two or several proteins, or of other macromolecular components tend to have high molecular weights. Although a 30 000-Dalton size limit allowed to work with a large pool of physiologically interesting proteins, it was hardly compatible with extensive use of NMR for studies of such supramolecular structures. For example, this size limit would severely narrow down the range of potential receptor systems accessible to NMR in drug discovery projects (Shuker et al., 1996; Pellechia et al., 2002), restrict studies of protein–nucleic acid complexes (Figure 2) to a small number of systems with modest size, and prevent the use of solution NMR for studies of membrane proteins, since these have to be reconstituted and solubilized in mixed micelles with detergents or lipids (Figure 30, right side). A few years ago this limitation was successfully challenged, since the size-range for applications of solution NMR techniques could be significantly extended through the introduction of transverse relaxation-optimized spectroscopy (TROSY) (Pervushin et al., 1997). As an illustration, Figure 31 shows a [^{15}N,^{1}H]-TROSY correlation spectrum of a membrane protein reconstituted in detergent micelles. Sharp, well-separated peaks are obtained in spite of the large size of the mixed micelles (Figure 30, right side). With the use of

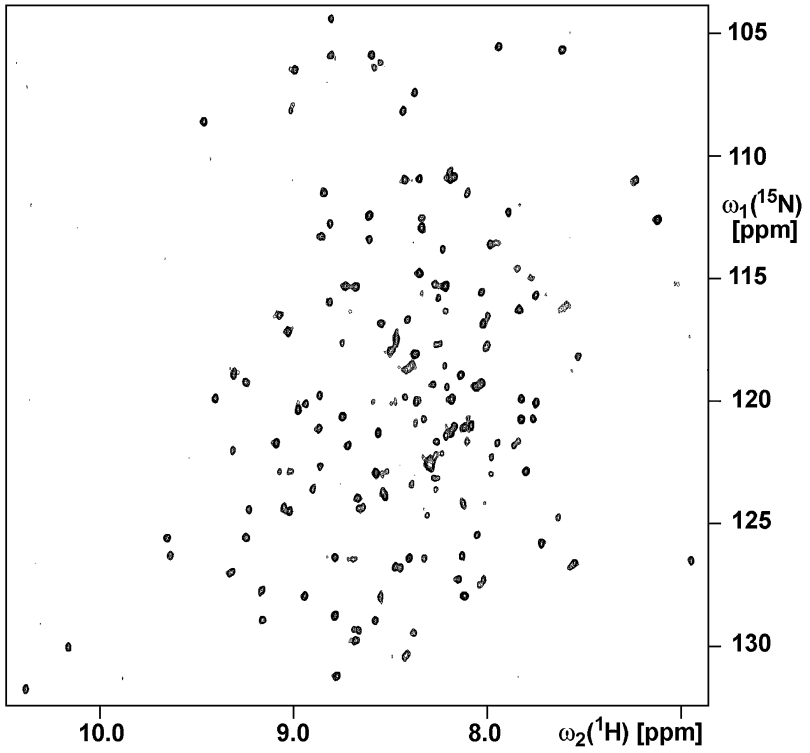


Figure 31. Transverse relaxation-optimized [^{15}N,^{1}H]-correlation spectroscopy (TROSY). A spectrum of the uniformly ^{2}H- and ^{15}N-labeled membrane protein OmpX reconstituted in DHPC micelles is shown (750 MHz, 30 °C, H_2O-solution).

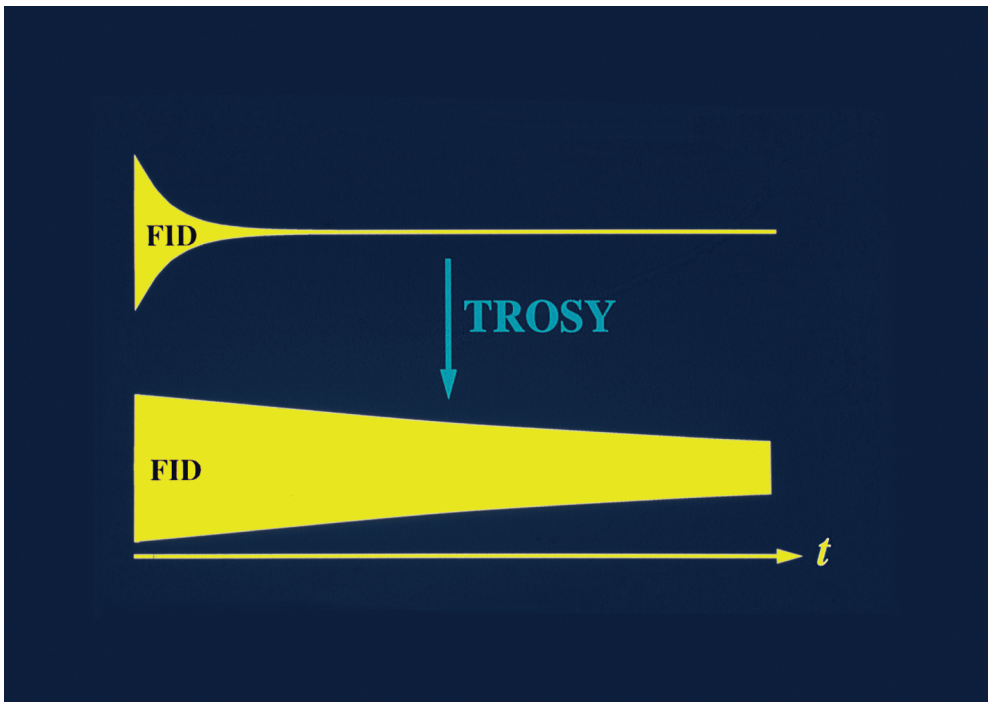


Figure 32. Schematic drawing of the free induction decay (FID) for a large protein with and without the use of TROSY. The period covered along the time axis is of the order of 100 milliseconds. The FID represents the primary recording of the NMR data in the 'time domain', from which the frequency domain spectrum (for example, Figures 18, 19 and 31) is obtained by Fourier transformation (Ernst et al., 1987; Ernst, 1992).

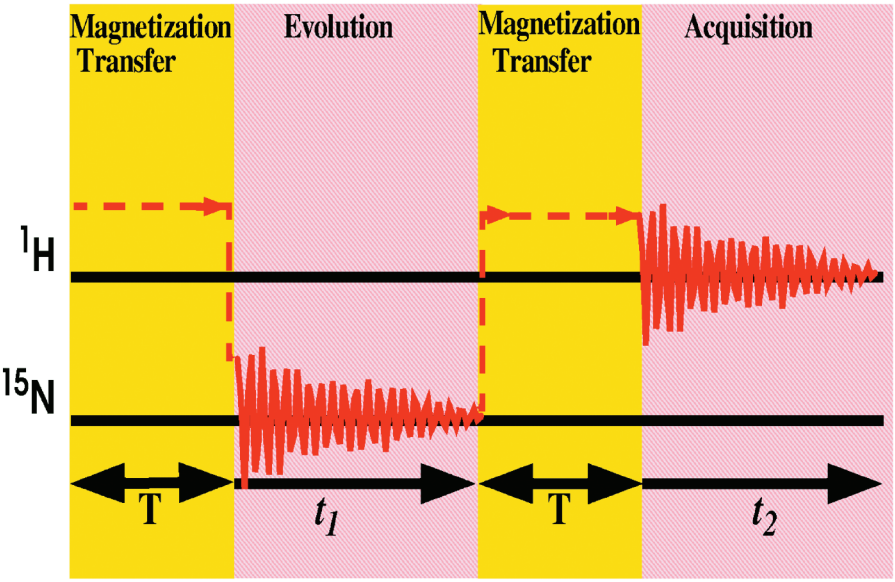


Figure 33. Basic features of a 2D [^{15}N,^{1}H]-correlation experiment. On the left, ^{1}H and ^{15}N indicate radio-frequency channels for irradiation of these isotopes. Yellow shading identifies the ^{1}H→^{15}N and ^{15}N→^{1}H magnetization transfer periods, T, and pink shading the evolution and acquisition periods, t_1 and t_2. During t_1 and t_2 the decay of the magnetization (FID) is schematically indicated. The overall duration of the experiment, $2T+t_1+t_2$, can be of the order of 10 to 1000 milliseconds, depending on the size of the structure studied and the intended purpose of the measurement.

the TROSY-principle, the more complex NMR experiments needed for a structure determination with [^{2}H,^{13}C,^{15}N]-labelled proteins can also be obtained with high quality (Salzmann et al., 1998), so that solution NMR can now be used for *de novo* membrane protein structure determination (Fernández et al., 2001).

The improved quality of the NMR spectra of large structures with the use of TROSY can be qualitatively rationalized by the following considerations. As has previously been indicated in the discussion on NOEs (Equation [1]), the appearance of solution NMR spectra is intimately related to the effective correlation time, τ_c, which characterizes the thermal motions of the molecule considered. The increase of the τ_c-values in larger structures results in line broadening due to rapid transverse spin-relaxation. For example, if the spectrum of Figure 31 had been recorded with conventional NMR techniques, most of the resonance lines would not be individually resolved due to severe line broadening, and one would have experienced a severe loss of sensitivity for detection of the NMR signals. The reduced sensitivity can be readily appreciated by examination of the 'free induction decays' (FID) in the time domain data. Figure 32 shows that rapid loss of magnetization in a conventional NMR experiment with a large structure can be slowed down by the use of TROSY, which corresponds to the reduction of the line width in the frequency domain spectrum. The impact on the sensitivity is visualized in Figure 33 with a simplified scheme for a 2D correlation experiment, which includes two magnetization transfer periods, and the evolution and acquisition periods. During evolution and acquisition, the system is not subjected to external perturbations, and loss of magnetization occurs at a rate determined by the 'transverse relaxation time'. It is readily apparent that rapid loss of magnetization in a conventional NMR experiment (upper trace in Figure 32) leads to weak or even vanishing signal intensity at the end of the evolution period, and accordingly the sensitivity for detection of the signal during the acquisition period is very low. In contrast, one can obtain much improved sensitivity with the use of TROSY (lower trace in Figure 32), since plenty of signal intensity will be preserved at the end of the evolution period, and the signal can be recorded with high signal-to-noise ratio during a large portion of the acquisition period. Following such considerations for minimizing the loss of magnetization during all four time periods indicated in Figure 33, and with additional optimization of the magnetization transfer techniques (Riek et al., 1999), solution NMR spectra have by now been recorded for structures with molecular weights up to 870000. In the spectrum of the co-chaperonin GroES bound to the chaperonin GroEL (Figure 34), resonance lines that provided novel information on structural and dynamic features of the GroES–GroEL interface are colored in red and blue (Fiaux et al., 2002).

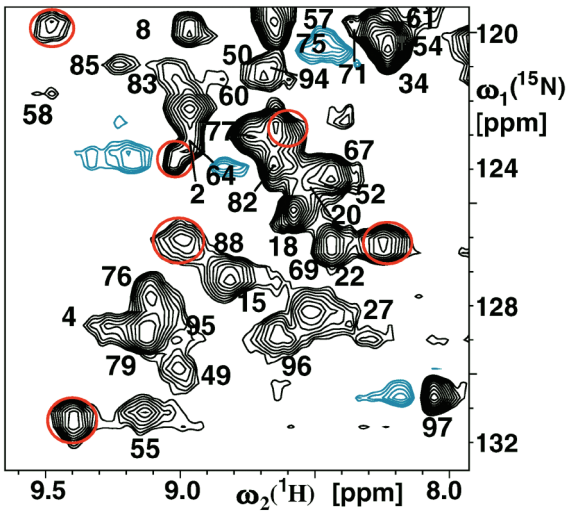


Figure 34. [^{15}N,^{1}H]-correlation spectrum of the co-chaperonin GroES (M $\approx$ 70000) bound to the chaperonin GroEL (M $\approx$ 800 000). The spectrum was recorded with a [^{15}N,^{1}H]-CRIPT-TROSY experiment (Rick et al., 1999; Fiaux et al., 2002). The red and blue colouring is explained in the text.

The principal contributions to the rate of transverse spin-relaxation can be traced to two different types of interactions, i.e., dipole–dipole coupling of the observed spin with other, nearby spins, and chemical shift anisotropy (CSA) (Abragam, 1961). These interactions are modulated by the stochastic rotational motions in solution, and as a consequence the rate of transverse relaxation increases for larger structures with slower Brownian motions. TROSY exploits constructive interference between the two relaxation mechanisms, and actually uses CSA-relaxation at high fields to cancel the dipolar relaxation (Pervushin et al., 1997). In this way the appearance of the NMR spectrum is effectively uncoupled from the Brownian motions, which then enables the recording of solution NMR spectra with large structures.

It remains to be seen how these new NMR techniques will be employed most profitably in the future. Intriguing possibilities include that NMR can now be

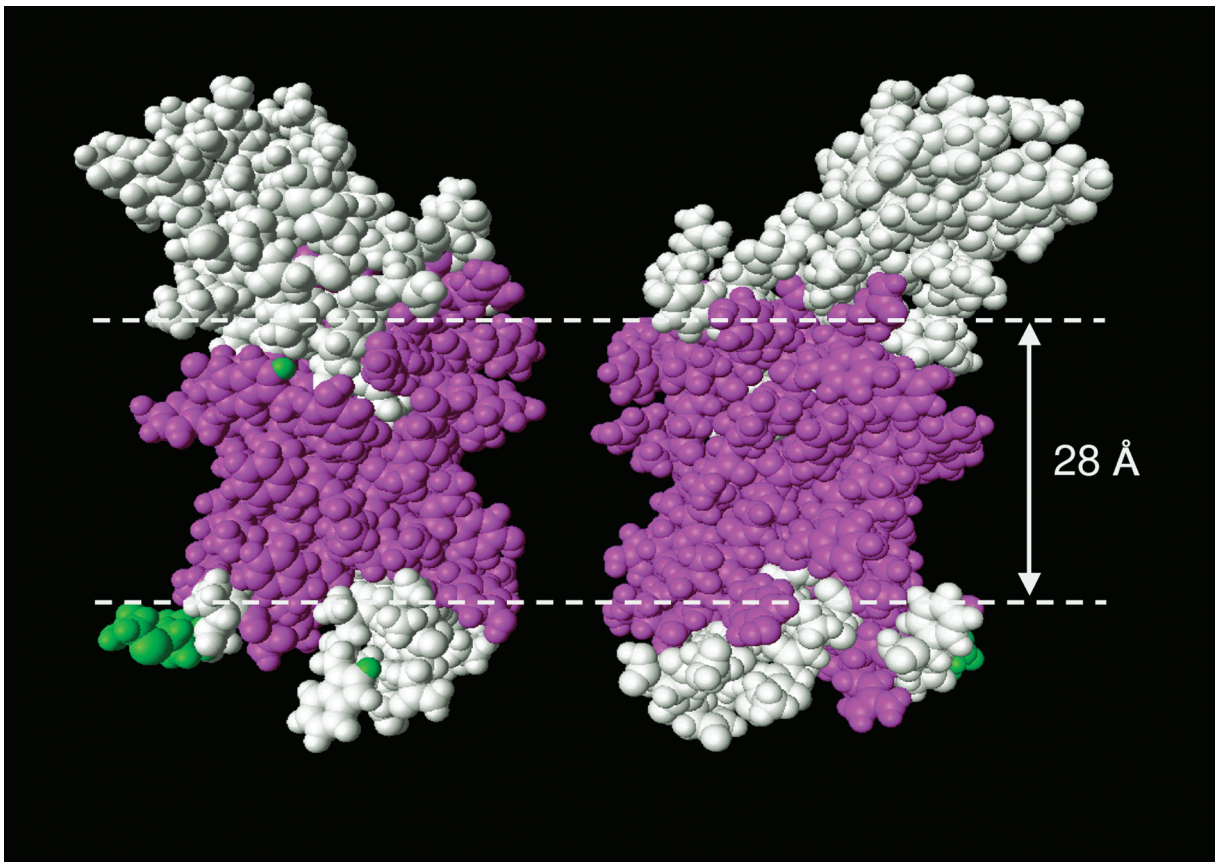


Figure 35. NMR structure of the membrane protein OmpX reconstituted in DHPC micelles. Front and back views of a space-filling model of OmpX are shown, with pink and green coloring identifying the surface area of the protein that is in contact with DHPC molecules in the mixed micelles. The two broken horizontal lines indicate the thickness of the lipid phase in the *E. coli* cell membrane (adapted from Fernandez et al., 2002).

employed in drug discovery projects with very large receptors (Shuker et al., 1996; Pellecchia et al., 2002). Combined with suitable isotope-labelling strategies, TROSY-based NMR techniques have also been shown to provide a powerful approach for investigations of intermolecular interactions in supramolecular structures with two or several macromolecular components (Pellecchia et al., 1999; Frickel et al., 2002). In these applications, a detailed structural interpretation of the NMR data will in most instances be dependent on the availability of an independently determined atomic-resolution structure for one or multiple components, which may have been obtained either by NMR in solution or by diffraction methods in single crystals. Applications of the new NMR techniques for *de novo* determination of large structures appears to be particularly attractive for, but not limited to, nucleic acid–protein complexes and small membrane proteins reconstituted in soluble detergent or lipid micelles. The characterization of a membrane protein structure has also been extended to include the detergents in the mixed micelles (Figure 35) (Fernández et al., 2002), since the areas of the protein surface in contact with the detergent molecules could be delineated by the observation of intermolecular NOEs between hydrogen atoms of the protein and the detergent. In these particular mixed micelles this surface area is virtually identical to the protein surface in contact with the lipid phase in the biological membrane. Since this leaves both ends of the β-barrel of OmpX freely accessible to the aqueous solvent, this reconstitution system should also be suitable for functional studies of the membrane protein with NMR.

This last section of my presentation certainly indicates only a narrow range of potential further developments of solution NMR techniques and their applications in structural biology and structural proteomics. I look forward with great expectations to the

Students, Research associates, Technical staff 1970–2002

Werner Kindlimann
Jean-Paul Meraldi
André Masson
Regula M. Keller
Jan Götz
Rudolf Baumann
Bernard Donzel
Jiri Hochmann
Christoph Grathwohl
Kaspar Elber
Robert Hetzel
Gabriele Kurz
Jolan M. Targonski
Andreas Gerber
Antoinette Schumacher
Kurt Gabler
Gerhard Wagner
Arno Bundi
Maureen Dammbach
Peter Sehr
Aaron J. Gilboa
Miguel Llinás
Antonio De Marco
Hans-Ulrich Gremlich
Willi Meier
Larry R. Brown
René Richarz
Alexandra S. Frey
Christoph Bösch
Vincenza Viti
Lea Stassinopoulou
Jürgen Lauterwein
Peter Bachmann

Klaus Loth
Kuniaki Nagayama
Stephen J. Perkins
Sidney L. Gordon
Lotti Rutz
Andreas Dubs
Daniel Picot
Werner A. Braun
Heinrich Roder
Albert Eugster
Gerhard Wider
Christoph Briner
Chrit Moonen
Adela Chrzeszczyk
Markus Meier
Subbaraya Ramaprasad
Nobuhiro Gō
Mitiko Gō
Johannes Schaffner
Markus Schmidiger
Hans Senn
Wayne E. Steinmetz
Anil Kumar
Yvonne Hunziker
Sadato H. Yabuki
Edith Huber
Fredi Stucki
Olle Teleman
Martin Billeter
Slobodan Macura
Kong-Hung Lee
Ramakrishna Hosur
Alexandre Arseniev

Petr Štrop
Dominique Marion
Erik R. P. Zuiderweg
Françoise Guerlesquin
Arthur Pardi
Kenneth R. Metz
Timothy Havel
Michael P. Williamson
Paula Y. Watnick
Martin Hassler
David Neuhaus
Peter Schultze
Hans Widmer
Werner Leupin
Satoshi Ebina
Reto Marani
Walter J. Chazin
Michael H. Frey
Winfried Denk
Sven Hyberts
Gottfried Otting
Allen D. Kline
Alexander M. Labhardt
Gaetano Montelione
Chantal Wynants
Thomas Schaumann
Phuntsok Ernst
Qiwen Wang
Erich Wörgötter
Norbert Müller
Luzi Jenny
Bernardo Celda
Rolf Grütter

Daniel Brühwiler
Manuel Glauser
Oscar D. Redwine
Francesco Riva
Hydeyuki Haruyama
Daniel Baumann
Bennett Th. Farmer II
Christoph Weber
Hua Qingxin
Peter Güntert
Francesco Avilés
Tai-he Xia
Leonard P. M. Orbons
Barbara Messerle
Craig D. Eccles
Yan-qiu Qian
Dario Neri
Hervé Darbon
Hugo Amacker
Wolfgang Kaissl
Ivan Vranesic
Kandala V. R. Chary
Daniel Braun
Beat Werner
Patrick Sodano
Martin Bos
Maria Paz Sevilla
Agostino Fede
Josep Vendrell
Thomas A. Szyperski
Kurt Berndt
Claus Spitzfaden
Jonathan M. Moore

Berthold von Freyberg
Edvards Liepinsh
Johanna Blatter
Miguel S. Col
John F. O'Connell
Barbara Leiting
John H. Bushweller
Tadakazu Maeda
Martin Schick
Walfrido Antuch-Garcia
Peter Luginbühl
Christian Bartels
Volker Dötsch
Siggi Mronga
Jan Johansson
Gregg Siegal
Oliver Schott
Nikolaus Schaefer
Gabriel Hänggi
Reto Koradi
Helena Kovacs
Marcel Ottiger
Ralph Daumke
Matej Grilc
Serge Altmann
Maurizio Pellecchia
César Fernández-Estrabao
Christian Mumenthaler
Oliver Zerbe
Bénédicte Prêcheur
Hideo Iwai
Oliver Ohlenschläger
René Brunisholz

Gerhard Frank
Walter Sidler
Konstantin Pervushin
Rosmarie Hug
Cyrill Arnet
Roland Riek
Francisco López-García
Martin Wahl
Jihui Wu
Ralf Glaser
Aizhuo Liu
Michael Salzmann
Alexander C. Sobol
Fred F. Damberger
Chih-yen King
Bogdan Banecki
Yves-Laurent Viollier
Ralph Zahn
Michel Hochuli
Shin-ichi Tate
Eliane Ulrich
Jörg Pieper
Tian Y. Tsong
Thorsten Lührs
Christine von Schroetter
Roberto Fattorusso
Helmut Grubmüller
Reinhard Wimmer
Martha Geier
Sandra Zangger
Torsten Herrmann
Kai Schnee

Jocelyne Fiaux
Luigi Calzolai
Regitze Vold
Reto Horst
Donghan Lee
Koba Adeishvili
Dominikus A. Lysek
Sophie Bonjour
Christian Hilty
Morten Hohwy
Rochus L. J. Keller
Giorgio Pioda
Lars Dreier
Beat Vögeli
Touraj Etezady-Esfarjani
Christian Schorn
Francesco Fiorito
Pascal Bettendorff
Simone Hornemann
Vicent Esteve Moya
Valmir Fadel
Kent A. Baker
Hakim Tafer
Alvar Gossert
Lucas Nivon
Sebastian Hiller
Ida Arm
Yolanda Auchli
Daniel Pérez
Marianne Wüthrich
Isabelle Allen
Gabriela Arm

Figure 36. My collaborators at the ETH Zürich from 1970 to 2002 listed in the order of their starting dates.

future evolution of this awesome and beautiful technique, which has given me so many years of joy and excitement in studies of the molecules of life.

Acknowledgements

From 1970 through 2002, 229 students, postdoctoral research associates, and technical and administrative staff worked with me at the ETH Zürich. I am deeply indebted to all of them for their enthusiasm and dedication, which resulted in the work summarized in this article. Their names are collected in Figure 36, and individual contributions can be recognized in the reference list. Although human minds stand behind all progress in science, the success of our research projects also depended critically on financial resources. I would like to acknowledge the ETH Zürich, the Swiss National Science Foundation, the Kommission für Technologie und Innovation (KTI), Bruker-Biospin AG, and the Scripps Research Institute in La Jolla, CA, USA, for their support.

References

Abragam, A. (1961) *Principles of Nuclear Magnetism*, Clarendon Press, Oxford.

Balaram, P., Bothner-By, A.A. and Dadok, J. (1972) *J. Amer. Chem. Soc.*, **94**, 4015–4017.

Bax, A. and Grzesiek, S. (1993) *Accounts Chem. Res.*, **26**, 131–138.

Billeter, M., Braun, W. and Wüthrich, K. (1982) *J. Mol. Biol.*, **155**, 321–346.

Braun, W., Bösch, C., Brown, L.R., Gō, N. and Wüthrich, K. (1981) *Biochim. Biophys. Acta*, **667**, 377–396.

Davis, D.G., Lindstrom, T.R., Mock, N.H., Baldassare, J.J., Charache, S., Jones, R.T. and Ho, C. (1971) *J. Mol. Biol.*, **60**, 101–111.

Dubs, A., Wagner, G. and Wüthrich, K. (1979) *Biochim. Biophys. Acta*, **577**, 177–194.

Ernst, R.R. (1992) *Angew. Chemie – Int. Ed.*, **31**, 805–823.

Ernst, R.R., Bodenhausen, G. and Wokaun, A. (1987) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Oxford University Press, Oxford.

Fernández, C., Hilty, C., Wider, G. and Wüthrich, K. (2002) *Proc. Natl. Acad. Sci. USA*, **99**, 13533–13537.

Fernández, C., Hilty, C., Bonjour, S., Adeishvili, K., Pervushin, K. and Wüthrich, K. (2001) *FEBS Lett.*, **504**, 173–178.

Fiaux, J., Bertelsen, E., Horwich, A. and Wüthrich, K. (2002) *Nature*, **418**, 207–211.

Frickel, E.M., Riek, R., Jelesarov, I., Helenius, A., Wüthrich, K. and Ellgaard, L. (2002) *Proc. Natl. Acad. Sci. USA*, **99**, 1954–1959.

Gibbons, W.A., Alms, H., Bockman, R.S. and Wyssbrod, H.R. (1972) *Biochemistry*, **11**, 1721–1725.

Gordon, S.L. and Wüthrich, K. (1978) *J. Amer. Chem. Soc.*, **100**, 7094–7096.

Güntert, P. (1998) *Quart. Rev. Biophys.*, **31**, 145–237.

Gupta, R.K. and Redfield, A.G. (1970) *Science*, **169**, 1204–1206.

Havel, T.F. and Wüthrich, K. (1984) *Bull. Math. Biol.*, **46**, 673–698.

Havel, T.F. and Wüthrich, K. (1985) *J. Mol. Biol.*, **182**, 281–294.

Herrmann, T., Güntert, P. and Wüthrich, K. (2002) *J. Biomol. NMR*, **24**, 171–189.

Jardetzky, O. and Jardetzky, C.D. (1958) *J. Biol. Chem.*, **233**, 383–387.

Kalk, A. and Berendsen, H.J.C. (1976) *J. Magn. Reson.*, **24**, 343–366.

Kay, L.E. and Gardner, K.H. (1997) *Curr. Opin. Struct. Biol.*, **7**, 564–570.

Keller, R.M. and Wüthrich, K. (1978) *Biochem. Biophys. Res. Comm.*, **83**, 1132–1139.

Kowalsky, A. (1962) *J. Biol. Chem.*, **237**, 1807–1819.

Kurland, R.J., Davis, D.G. and Ho, C. (1968) *J. Amer. Chem. Soc.*, **90**, 2700–2701.

Lopez Garcia, F., Zahn, R., Riek, R. and Wüthrich, K. (2000) *Proc. Natl. Acad. Sci. USA*, **97**, 8334–8339.

Luginbühl, P. and Wüthrich, K. (2002) *Progr. Nucl. Magn. Reson. Spectr.*, **40**, 199–247.

Mandel, M. (1965) *J. Biol. Chem.*, **240**, 1586–1592.

McDonald, C.C. and Phillips, W.D. (1967) *J. Amer. Chem. Soc.*, **89**, 6332–6341.

Moseley, H.N.B. and Montelione, G.T. (1999) *Curr. Opin. Struct. Biol.*, **9**, 635–642.

Noggle, J.H. and Schirmer, R.E. (1971) *The Nuclear Overhauser Effect*, Academic Press, New York.

Otting, G. and Wüthrich, K. (1990) *Quart. Rev. Biophys.*, **23**, 39–96.

Otting, G., Liepinsh, E. and Wüthrich, K. (1991) *Science*, **254**, 974–980.

Otting, G., Qian, Y.Q., Billeter, M., Müller, M., Affolter, M., Gehring, W.J. and Wüthrich, K. (1990) *EMBO J.*, **9**, 3085–3092.

Pardi, A., Billeter, M. and Wüthrich, K. (1984) *J. Mol. Biol.*, **180**, 741–751.

Pellecchia, M., Sem, D.S. and Wüthrich, K. (2002) *Nature Rev. Drug Disc.*, **1**, 211–219.

Pellecchia, M., Sebbel, P., Hermanns, U., Wüthrich, K. and Glockshuber, R. (1999) *Nature Struct. Biol.*, **6**, 336–339.

Pervushin, K., Riek, R., Wider, G. and Wüthrich, K. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 12366–12371.

Qian, Y.Q., Billeter, M., Otting, G., Müller, M., Gehring, W.J. and Wüthrich, K. (1989) *Cell*, **59**, 573–580.

Riek, R., Wider, G., Pervushin, K. and Wüthrich, K. (1999) *Proc. Natl. Acad. Sci. USA*, **96**, 4918–4923.

Salzmann, M., Pervushin, K., Wider, G., Senn, H. and Wüthrich, K. (1998) *Proc. Natl. Acad. Sci. USA*, **95**, 13585–13590.

Saunders, M. and Wishnia, A. (1958) *Ann. N.Y. Acad. Sci.*, **70**, 870–874.

Shuker S.B., Hajduk P.J., Meadows R.P. and Fesik S.W. (1996) *Science*, **274**, 1531–1534.

Shulman, R.G., Glarum, S.H. and Karplus, M. (1971) *J. Mol. Biol.*, **57**, 93–115.

Shulman, R.G., Ogawa, S., Wüthrich, K., Yamane, T., Peisach, J. and Blumberg, W.E. (1969) *Science*, **165**, 251–257.

Solomon, I. (1955) *Phys. Rev.*, **99**, 559–565.

Wagner, G. (1993) *J. Biomol. NMR*, **3**, 375–385.

Wagner, G. and Wüthrich, K. (1979) *J. Magn. Reson.*, **33**, 675–680.

Wagner, G. and Wüthrich, K. (1982) *J. Mol. Biol.*, **155**, 347–366.

Wagner, G., Anil-Kumar and Wüthrich, K. (1981) *Eur. J. Biochem.*, **114**, 375–384.

Wider, G. (1998) *Prog. Nucl. Magn. Reson. Spectr.*, **32**, 193–275.

Wider, G., Macura, S., Anil Kumar, Ernst, R.R. and Wüthrich, K. (1984) *J. Magn. Reson.*, **56**, 207–234.

Williamson, M.P., Havel, T.F. and Wüthrich, K. (1985) *J. Mol. Biol.*, **182**, 295–315.

Wüthrich, K. (1969) *Proc. Natl. Acad. Sci. USA*, **63**, 1071–1078.
Wüthrich, K. (1970) *Structure and Bonding*, **8**, 53–121.
Wüthrich, K. (1976) *NMR in Biological Research: Peptides and Proteins*, North Holland, Amsterdam.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wüthrich, K. (1995a) *Acta Cryst.*, **D51**, 249–270.
Wüthrich, K. (1995b) *NMR in Structural Biology: A Collection of Papers by Kurt Wüthrich*, World Scientific, Singapore.
Wüthrich, K. (2001) *Nature Struct. Biol.*, **8**, 923–925.
Wüthrich, K. and Wagner, G. (1975) *FEBS Lett.*, **50**, 265–268.
Wüthrich, K., Billeter, M. and Braun, W. (1984) *J. Mol. Biol.*, **180**, 715–740.
Wüthrich, K., Wider, G., Wagner, G. and Braun, W. (1982) *J. Mol. Biol.*, **155**, 311–319.