Review

Structure Determination of Biological Macromolecules in Solution Using Nuclear Magnetic Resonance Spectroscopy

Gerhard Wider
Eidgenössische Technische Hochschule Hönggerberg, Zürich, Switzerland

ABSTRACT

A detailed understanding of the function of a biological macromolecule requires knowledge of its three-dimensional structure. Most atomic-resolution structures of biological macromolecules have been solved either by X-ray diffraction in single crystals or by nuclear magnetic resonance (NMR) in solution. This review surveys the method of NMR structure determination. First, a brief introduction to NMR and its basic concepts is presented. The main part of the article deals with the individual steps necessary for an NMR structure determination. At the end, the discussion turns to considerations on the influence of the molecular size of the macromolecules on the structure determination by NMR. New techniques are discussed that greatly enhance the possibilities of applying NMR to large molecular systems.

INTRODUCTION

With the near completion of the human genome project, an even greater challenge awaits biologists: proteomics, the study of all the proteins coded by the genes. In many cases, it will be necessary to know the three-dimensional (3-D) structure of a protein to understand its function. The feasibility of such a structural proteomics project was recently demonstrated (10), and it was shown that two techniques will play a dominant part: X-ray crystallography and nuclear magnetic resonance (NMR). These are the two main techniques that can provide structures of macromolecules at atomic resolution. Both techniques are well established and already play a key role in biological and pharmaceutical research as a basis for a detailed understanding of molecular functions and as a foundation for protein design. Whereas X-ray crystallography requires single crystals, NMR measurements are carried out in solution under near physiological conditions. The important role that NMR plays in structural biology is illustrated by the far more than 1000 NMR solution structures deposited in the protein databank (3). NMR measurements not only provide structural data but also reach much further and can supply information on dynamics, conformational equilibria, folding and intra- as well as intermolecular interactions (14,16,28, 36,57). This review focuses on the application of NMR for structure determination of proteins in solution, but the methodology is applicable to other biomolecules, especially nucleic acids (53).

NMR does not directly produce an image of a protein. Rather, it is able to yield a wealth of indirect structural information from which the 3-D structure can only be revealed by extensive data analysis and computer calculations. The principles of an NMR structure determination can be summarized as follows (Figure 1): preparation of the protein solution, the NMR measurements, the assignment of NMR signals to individual atoms in the molecule, identification of conformational constraints (e.g., distances between hydrogen atoms) and the calculation of the 3-D structure on the basis of the experimental constraints. These different steps will be discussed in individual sections in this review. However, the NMR technique (8,9,15,48) is first introduced in the sections "Basic NMR", "Correlations in NMR" and "Multidimensional NMR".

Structure determination by NMR is an established technique and is routinely used to determine 3-D structures of biological macromolecules in solution with molecular weights up to 30 kDa (3,13,54,55). The power of NMR over other spectroscopic techniques results from the fact that every NMR-active nucleus gives rise to an individual signal (resonance line) in the spectrum that can be resolved by multidimensional NMR techniques. This becomes more difficult for larger molecular structures and puts a practical limit to the molecular size that can be studied in detail by NMR (11,29,56). Up to now, only very few NMR structures were determined of molecules larger than 30

kDa (18,32). In the section "NMR with Very Large Molecules" at the end of this article, new techniques will be discussed that allow detailed studies of macromolecules with a molecular weight far beyond 100 kDa (42,49).

BASIC NMR

This section presents some basic aspects of NMR that are important for the understanding of the methods used for structure determination by NMR (9,15, 48,53). The basis of all NMR experiments is a quantum mechanical property of the nucleus: the spin. The nuclei of interest in NMR of biological macromolecules carry a spin of 1/2 that allows only two different spin states, often referred to as spin up and spin down. Associated with the spin is a magnetic moment that, for a spin 1/2, can be interpreted as a magnetic dipole. These very small atomic dipoles can only orient parallel or antiparallel to an external magnetic field. The two possible orientations correspond to slightly different energies, and spins are allowed to jump from one orientation to the other, absorbing or emitting the energy difference in the form of electromagnetic radiation. It is remarkable that the difference between the number of spins in the two states is extremely small, even if the spins are in a very strong magnetic field. For example, in 1 million hydrogen (^{1}H) nuclei, the numbers of spins orienting parallel or antiparallel differ only by about 60 at room temperature and in the highest magnetic field strengths available for NMR. Only the very small difference between the number of parallel and antiparallel spins contributes to the NMR signal. This fact explains why NMR is an insensitive technique compared, for example, with optical absorption spectroscopy.

NMR experiments are performed on an NMR spectrometer consisting of two components: a high field superconductive magnet that produces an extremely homogeneous, strong static magnetic field and a console that can generate electromagnetic waves in any desired combination. For an NMR experiment, a glass tube containing the solution of the molecule of interest is placed in the room temperature bore of

the superconductive magnet. The small imbalance of nuclear magnetic moments oriented parallel and antiparallel gives rise to a small polarization of the nuclear spins in the sample, resulting in a net macroscopic magnetization. With suitable irradiation of electromagnetic waves, this magnetization can be manipulated as the NMR spectroscopist chooses (2,15,44,48). The frequency of the irradiation must match the energy difference between the two possible states of the spins and typically lies in the radio frequency (*rf*) range between 50 and 800 MHz. The irradiation is usually applied only for a few microseconds as *rf* pulses.

Most *rf* pulses used in NMR experiments belong to one of two specific classes. One class, the 90° pulses, equalizes populations of spins up and spins down when applied to the equilibrium state. In a pictorial view, a 90° pulse turns the macroscopic magnetization by 90°. The other class, the 180° pulses, inverts the populations of the spin states; in a pictorial view, a 180° pulse turns the

Figure 1. Strategy of structure determination by NMR. Outline of the general strategy used to solve the 3-D structure of biological macromolecules in solution by NMR.

magnetization by 180°. After a perturbation of the equilibrium state of the spins by one or several *rf* pulses, a NMR signal can be observed. The signal consists of *rf* waves with frequencies that match the energy difference between the spin-up and spin-down states of the individual nuclei involved. Since the system tends to return to its equilibrium state, the signal will decay (relax) and typically vanish within 100 ms. During this time, the NMR signal, the free induction decay (FID), is recorded. The FID consists of a sum of decaying cosine waves whose frequencies represent the resonance frequencies of the nuclei in the sample. From this data, the NMR frequency spectrum is then obtained by a Fourier transformation.

In an NMR spectrum, the nuclei are represented by their characteristic resonance frequencies. For example, hydrogen atoms (protons) are represented in a proton resonance spectrum or carbon atoms in a carbon spectrum. The resonance frequencies of different types of nuclei are widely different. For example, protons (^{1}H) resonate at a ten times higher frequency than nitrogen nuclei (^{15}N) and four times higher than carbon nuclei (^{13}C). The resonance frequencies of different nuclei of the same type lie in a much narrower frequency range. For example, the resonance lines for different protons in a molecule vary in a range of a few parts per million (ppm) around the average proton resonance frequency. This variation is due to the interaction between the nuclei and surrounding electrons, which affect the local magnetic field experienced by a particular nucleus and thus influence its resonance frequency. For example, the resonance frequencies of amide protons, α-protons, or methyls are quite distinct within the proton spectrum and allow a tentative assignment to the different classes of protons. The chemical shift is very sensitive to a multitude of structural, electronic, magnetic and dynamic variables and in principle contains a wealth of information on the state of the system under investigation. However, a direct interpretation of the chemical shift is usually not possible because the dependence on these parameters is either not fully understood or too intricate to allow the derivation of detailed information.
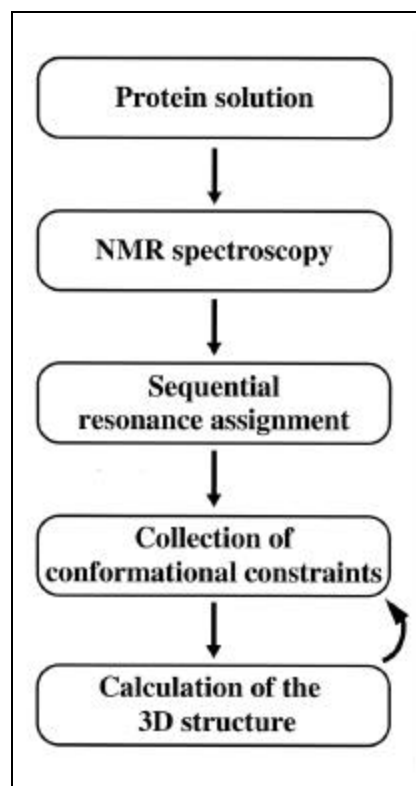
Review

CORRELATIONS IN NMR

A key feature exploited in NMR spectroscopy is the fact that the magnetic moments of the individual nuclei interact with the small magnetic fields created by the spins of nearby nuclei. This spin-spin interaction can be used to correlate different nuclei in a molecule with one another. Nuclei interact either "direct", through-space, or "indirect", through-bond. The latter are transmitted via polarization of bonding electrons and known as spin-spin coupling or J coupling; the former are the basis for the nuclear Overhauser effect (NOE), which permits distance measurements between hydrogen nuclei. The detailed analysis of protein spectra is based on through-bond and through-space correlations.

Through-bond correlations group individual spins into spin systems that are used for the analysis of the spectra. In proteins, J couplings over more than three chemical bonds are usually too small to be observed. Consequently, only spin systems within individual amino acids can be obtained in proton spectra. In proteins that are isotope-labeled with ^{15}N and ^{13}C, J couplings between ^{1}H, ^{15}N and ^{13}C allow through-bond correlations across the peptide bond. NMR experiments that correlate different nuclei via J coupling are often referred to as correlation spectroscopy (COSY)-type experiments (1,48,51). An important feature of COSY-type experiments is that they can transfer magnetization between different types of nuclei. This property makes it possible to start an experiment with one type of nucleus and to transfer the magnetization to another or several other types of nuclei in the course of the experiment. Such magnetization transfers are of great practical importance because widely different sensitivities are obtained with different types of nuclei. Hydrogen nuclei (protons) are the most sensitive type of nuclei. Thus, many NMR experiments with biological macromolecules start with the large proton magnetization and transfer the signal via heteronuclei (e.g., carbon and/or nitrogen) back to protons for the recording of the FID with maximal sensitivity.

Through-space correlations provide the basis for geometric information required to determine the structure of a macromolecule and are measured via the NOE (45,48). The NMR method for protein structure determination relies on a dense network of distance constraints derived from NOEs between nearby hydrogen atoms in the protein (53). NOEs connect pairs of hydrogen atoms separated by less than 0.5 nm. In contrast to COSY-type experiments, the nuclei involved in the NOE correlation can belong to amino acid residues that may be far apart along the protein sequence but close in space. The NOE reflects the transfer of magnetization between spins interacting via their associated dipoles. For molecules with a molecular weight of more than 5 kDa, the intensity of an NOE is approximately proportional to r^{-6} and to the molecular weight, where r is the distance between the two interacting spins. Due to the dependence on the inverse sixth power of the distance between the nuclei, the NOE intensity falls off rapidly with increasing distance. This fact explains why NOEs between protons in a globular protein separated by more than 0.5 nm are usually not observed. NMR experiments that measure the NOE are often referred to as NOE spectroscopy (NOESY) experiments (30,51).

MULTIDIMENSIONAL NMR

NMR spectra of biological macromolecules contain hundreds or even thousands of resonance lines that cannot be resolved in a conventional 1-D NMR experiment. However, for a detailed analysis, resolved lines are a prerequisite. Further, the interpretation of NMR data requires correlations between different nuclei that are implicitly contained in 1-D spectra but are often difficult to extract. Multidimensional NMR spectra provide both increased resolution and correlations that are easy to analyze. The price to pay for this are more complex NMR experiments that consist of a series of rf pulses separated by short time periods (delays) during which no external rf is applied. This pulse sequence is then followed by the recording of the resulting magnetization.

The crucial step in increasing the dimensionality of NMR experiments lies in the extension from one to two dimensions. A higher dimensional NMR experiment consists of a straightforward combination of 2-D experiments. All 2-D NMR experiments use the same basic scheme, which consists of the four following, consecutive time periods: excitation - evolution - mixing - detection.

During the excitation period, the spins are prepared in the desired state from which the chemical shifts of the individual nuclei are observed during the evolution period, t_1. In the mixing period, the spins are correlated with each other, and the information on the chemical shift of one nucleus ends up on an other nucleus of which the frequency is measured during the detection period, t_2. Technically, the individual periods consist of delays and rf pulses that together form the pulse sequence. To obtain a 2-D data set, a number of experiments are recorded with successively incremented lengths of the evolution period t_1 to generate a data matrix $S(t_1, t_2)$. A 2-D Fourier transformation of $S(t_1, t_2)$ yields the desired 2-D frequency spectrum $s(\omega_1, \omega_2)$. A resonance in the 2-D spectrum, a peak, represents a pair of nuclei that suitably interact during the mixing time. Figure 2 presents an example of a 2-D NMR spectrum. A 2-D $[^{15}N, ^{1}H]$ correlation spectrum (4) of a ^{15}N-labeled protein in water solution is shown. The spectrum exhibits correlations between amide nitrogens and the attached protons. Correspondingly, there are two different chemical shift axes, ω_1 and ω_2: one for nitrogens (^{15}N) and the other one for protons (^{1}H). Each cross peak in the spectrum of Figure 2 represents one NH-group in the backbone of the polypeptide chain. Since all amino acids except proline have such an amide group, the spectrum of Figure 2 represents a characteristic fingerprint for a protein.

The extension from a 2-D to an n-dimensional (n-D) NMR experiment consists of the combination of (n-1) 2-D experiments, which contains only one excitation and one detection period but repeats the evolution and mixing times (n-1) times. This results in a pulse sequence comprising (n-1) independently incremented evolution periods and (n-1)

Vis Genetics
4 C

Circle Reader Service No. 189

Review

mixing periods and one detection period. A typical n-D NMR experiment thus follows the scheme: excitation - (evolution - mixing)$_{n-1}$ - detection, where the bracket repeats (n-1) times. Only during the detection period the signal is physically measured, and this period is thus often referred to as the direct dimension in contrast to the evolution periods, which are referred to as indirect dimensions. To obtain an n-D data matrix, $S(t_1, t_2, ..., t_n)$, the pulse sequence is repeated many times with individually incremented lengths of the (n-1) evolution periods.

An n-D Fourier transform of $S(t_1, t_2, ..., t_n)$ provides an n-D spectrum that depends on n frequency variables, $s(\omega_1, \omega_2, ..., \omega_n)$. If two nuclei suitably interact with each other in a mixing time, then this interaction will be manifested by a resonance in the spectrum, a cross peak, at a position characterized by the resonance frequencies of the interacting nuclei. For example, in the schematic

drawing of Figure 3, the nuclei ^{1}H and ^{15}N interact in one mixing time and ^{15}N and ^{13}C in the other mixing time of a 3-D experiment. Correspondingly, one peak in the 3-D spectrum appears at the position indicated by the chemical shifts of the three nuclei.

Figure 3 illustrates the increasing resolution obtained when going from 2-D to 3-D spectra. Figure 3A shows a schematic [^{1}H,^{15}N]-correlation spectrum (Figure 2). A 3-D spectrum can, for example, be obtained by correlating the amide groups represented in the spectrum of Figure 3A with the α-carbon nuclei attached to [15]N. The chemical shifts of these carbon nuclei are used to spread the resonances from the 2-D plane into a third dimension. In Figure 3B, the 3-D spectrum is represented by a stack of 2-D planes that accords with the actual digitization of the data in the ^{13}C dimension. The same representation is used for the analysis

of 3-D spectra. Two frequency axes are chosen to define 2-D planes that are stacked in the third dimension.

PROTEIN SOLUTION

In the preceding three sections, theoretical and technical aspects of NMR were briefly introduced to the extent necessary to follow the discussion of the general strategy used to solve the 3-D structure of biological macromolecules in solution by NMR (Figure 1). The first step in this strategy is the preparation of the protein solution. An efficient structure determination by NMR requires a highly purified protein preparation. An inhomogeneous preparation and/or aggregation of the protein as well as low molecular weight protonated impurities may severely impair the structure determination. The first step in every protein NMR study therefore involves optimization of the measurement conditions. The pH, ionic strength and temperature can often be adjusted to mimic physiological conditions. The macromolecule under study should be stable in the chosen conditions for many weeks. Any buffers, co-solvents and additives (e.g., detergent molecules) used should be hydrogen-free or deuterated. For the NMR measurements, the protein is dissolved in 0.25–0.5 mL water that contains about 5% D_2O, which is necessary for the stabilization of the NMR instrument during the measurement. The inherent low sensitivity of the technique requires protein concentrations of about 1 mM, that is, for a 0.5-mL solution of a protein with a molecular weight of 10 kDa, about 5 mg protein are required.

Proteins with a molecular weight larger than 10 kDa must be isotope-enriched in ^{15}N and ^{13}C for an efficient structure determination. [15]N and ^{13}C are used because the most abundant carbon isotope (^{12}C) does not give an NMR signal and the most abundant nitrogen isotope (^{14}N) has undesired NMR properties. The natural abundance of ^{15}N and ^{13}C is only 0.37% and 1.1%, respectively, whereas levels close to 100% are required for efficient NMR experiments with macromolecules. For the isotope enrichment, the protein is usually overexpressed in a
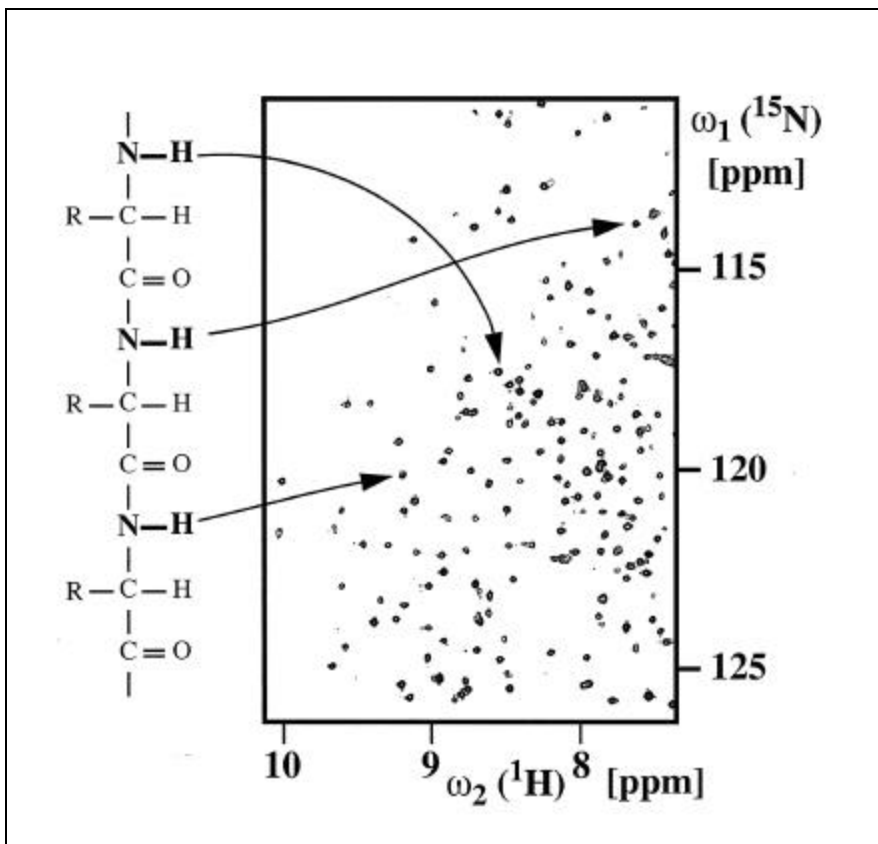
Figure 2. 2-D NMR spectrum. Through-bond correlation spectrum between the hydrogen (^{1}H) and the nitrogen (^{15}N) nuclei of amide groups in a protein. The location of the amide groups in the polypeptide backbone are sketched in the formulae on the left. The arrows indicate that each peak in the NMR spectrum corresponds to one NH-moiety. The axes of the NMR spectrum indicate the chemical shift of the ^{1}H and the ^{15}N nuclei.

USB
4 C
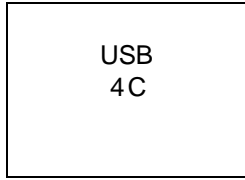
Circle Reader Service No. 190

Review

bacterial system. Fully isotope-labeled proteins can be produced by growing the bacteria on minimal medium containing $^{15}NH_4Cl$ and $^{13}C_6$-glucose as the sole nitrogen and carbon sources. In principle, all necessary NMR measurements could be done with one $[^{15}N,^{13}C]$-labeled sample. Since labeling with ^{15}N is only slightly more expensive than the preparation of an unlabeled sample, a ^{15}N-labeled molecule should be produced during optimization of the overexpression system. The ^{15}N-labeled protein can be used for 3-D ^{15}N-resolved NOESY experiments and for the study of the dynamics of the molecule (28,36). For systems larger than 30 kDa, deuteration and/or specific isotope labeling may be necessary (17). Partial deuteration can be achieved by exchanging the water by D_2O in the minimal medium used to grow the bacteria. With this procedure, the α-protons are replaced by deuterons but the side-chain protons much less and the aromatic rings stay mostly protonated. For an efficient production of a completely deuterated protein, the protonated glucose in the medium must be replaced by the expensive, fully deuterated glucose.

## NMR SPECTROSCOPY WITH BIOLOGICAL MACROMOLECULES

The foundations of NMR structural studies are high-quality NMR spectra recorded with good sensitivity and spectral resolution. Most NMR experiments with biological macromolecules are measured with two or three dimensions. The measurements almost always use protons (^1H) and depending on the isotope labeling ^{13}C and/or ^{15}N nuclei. In studies of nucleic acids, the natural isotope ^{31}P can also be used. The sensitivity obtainable with these types of nuclei greatly varies even if the sample is fully isotope-labeled with ^{13}C or ^{15}N. The proton offers the best sensitivity and hence constitutes the preferred nucleus for detection of the NMR signal. The other nuclei are usually measured during evolution periods of multidimensional NMR experiments, and their information is transferred to protons for detection.

The chemical shift plays a very important though mainly indirect role in a structure determination by NMR. The chemical shift is very sensitive to the microenvironment of a particular nucleus. Thus, in a folded protein, multiple copies of the same amino acid can be distinguished because of the conformation-dependent chemical shift. As mentioned before, a direct interpretation of the chemical shift is usually not (yet) possible. However, the large amount of data represented by the ^1H, ^{13}C and ^{15}N shifts in known 3-D NMR structures of proteins form the basis for a multitude of empirical and semi-empirical correlations with structural parameters (6,31,46,52).

For the necessary detailed analysis of NMR spectra, hundreds or even thousands of resonance lines must be observed as separate peaks. For protein spectra, this is only possible in multidimensional NMR spectra that are measured on NMR instruments operating with strong magnetic fields corresponding to a proton resonance frequency of 600 MHz or higher. Many NMR measurements with biological macromolecules rely on the presence of exchangeable protons in the molecules and must therefore be performed in water solution. Measuring the ^1H spectrum of a protein at a typical concentration of 1 mM dissolved in water requires special NMR techniques to massively reduce the enormous water resonance (20,48, 50) that is located in the center of the hydrogen spectrum.

Figure 4 shows a slightly simplified experimental scheme that will serve as an example to illustrate a typical pulse sequence applied with $[^{13}C,^{15}N]$-labeled proteins in solution. The experimental scheme correlates the chemical shifts of the amide proton and nitrogen (HN) with the α-carbon chemical shift (CA) of the preceding residue in the amino acid sequence via the intervening carbonyl carbon (CO). This 3-D experiment is called HN(CO)CA (19,43). A cross peak in the corresponding 3-D spectrum indicates a connection via the peptide bond of two sequential amino acids in a fully $[^{13}C,^{15}N]$-labeled protein. The experiment is one of a whole family of schemes that are used for the assignment of resonances to individual atoms in the protein. The pulse sequence (Figure 4) starts with the excita-
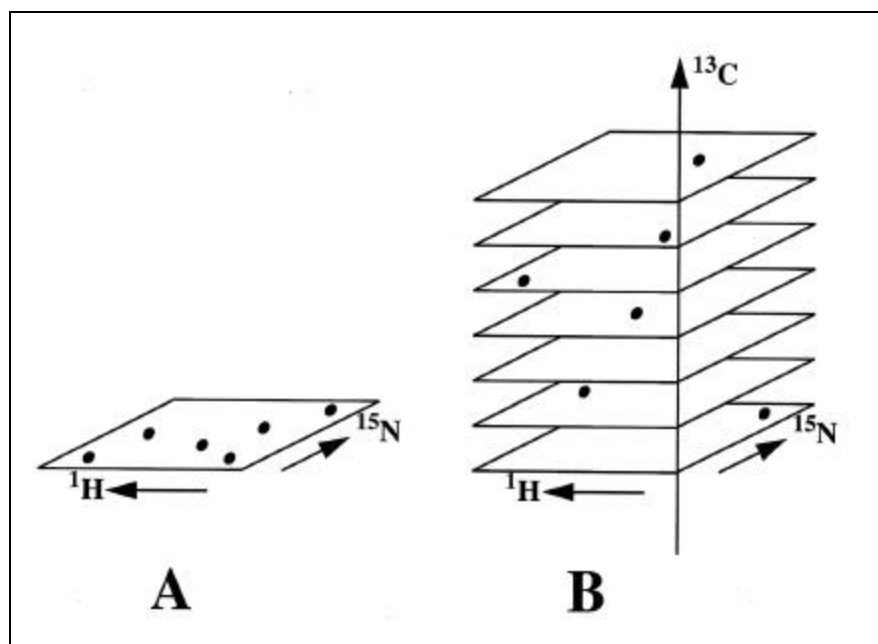
Figure 3. Schematic 2-D and 3-D NMR spectrum. The drawing illustrates the representation of a 3-D spectrum and the increased resolution obtained when going from 2-D to 3-D NMR spectra. The six peaks in the 2-D [^1H,^{15}N]-correlation spectrum (A) are separated in different planes of a 3-D spectrum (B) by an additional correlation with the α-carbon nuclei (^{13}C) attached to the nitrogen nuclei (^{15}N) in the same amino acid residue. The chemical shifts of the carbon nuclei are used to spread the resonances from the 2-D plane into a third dimension.

tion of amide protons (H in Figure 4) that are subsequently correlated to the attached ^{15}N nucleus (N) whose chemical shift is encoded during the evolution period t_1. Next, the ^{15}N nuclei are correlated via the carbonyl carbon (C′) with the α-carbon nuclei (C) of the preceding amino acid in the polypeptide chain. The chemical shift of the respective α-carbon is encoded during the second evolution period, t_2. Subsequently the information on the ^{15}N and ^{13}C chemical shift is transferred back to the amide proton (Figure 4). During the detection period t_3, the amide proton signal is measured. A 3-D Fourier transformation of the data provides a 3-D spectrum, as schematically shown in Figure 3B, with a ^{1}H, a ^{15}N and a ^{13}C dimension. A peak in this spectrum represents a correlation between the amide proton with the attached amide nitrogen of one amino acid and the α-carbon of the preceding amino acid residue.

SEQUENTIAL RESONANCE ASSIGNMENT

For a detailed analysis of the information content of NMR spectra, nearly complete assignments of signals in the spectra to individual atoms in the molecule are a prerequisite. This task cannot generally be solved on the basis of 1-D NMR spectra. Only the application of multidimensional NMR spectroscopy allowed the development of general strategies for the assignment of signals in proteins (53). All procedures use the known protein sequence to connect nuclei of amino acid residues that are neighbors in the sequence (i.e., the assignment procedure takes advantage of the sequential arrangement of individual amino acids in a polypeptide chain).

For unlabeled proteins smaller than 10 kDa, the combination of two 2-D spectra often allows the assignment of most proton NMR signals to individual

protons (51,53). The first experiment, the [^{1}H,^{1}H]-COSY, detects through-bond interactions between protons and correlates protons that are separated by up to three chemical bonds. With this experiment, the protons within an amino acid can be correlated; however, neighboring amino acids in the polypeptide sequence cannot be connected. The set of correlated proton nuclei is referred to as a spin system. For most amino acid residues, the ensemble of all ^{1}H atoms thus forms one spin system. For example, in Ala, the methyl protons are separated by three bonds from the α-proton, which in turn relates to the amide proton via three bonds, so that a continuous network of through-bond correlations can be established. Exceptions are the four aromatic side chains and some peripheral labile side-chain protons, where the connection to the aliphatic ^{1}H atoms of the same side chain must be established via through-space correlations.

Review

After complete analysis of the COSY spectrum, there are ideally spin systems for all amino acids in the protein. Each amino acid is represented by a particular spin system. However, as soon as an amino acid occurs more than once in a polypeptide chain, a direct assignment to a specific sequence position is not possible. For this purpose, the second experiment, the [^{1}H, ^{1}H]-NOESY (27, 30,51), is measured. Resonances in this spectrum manifest short distances between pairs of protons. Only short distances with backbone amide protons are extracted at this stage and used to establish correlations between protons of neighboring residues in the polypeptide sequence. With this information, the spin systems can be put in their proper sequential order. A comparison of the sequentially ordered spin systems with the known amino acid sequence allows one to assign the sequence position to every spin system, and thus complete sequence-specific resonance assignments are obtained (53).

For larger proteins, extensive signal overlap prevents complete assignments of all ^{1}H signals in proton spectra. This barrier can be overcome with 2-D and 3-D NMR techniques and uniformly ^{13}C- and ^{15}N-labeled proteins. With these methods, systems with molecular weights up to approximately 35 kDa can be studied. In [^{13}C, ^{15}N]-labeled proteins a sequential assignment strategy can be used that is based on through-bond correlations across the peptide bond between sequential amino acids. This procedure circumvents the use of NOESY spectra already in the assignment step. Most of these correlation experiments use the three types of nuclei, ^{1}H, ^{15}N, and ^{13}C, and are referred to as triple resonance experiments. Spectra of large molecules not only suffer from signal overlap but also from poor sensitivity due to faster relaxation (signal loss) with increasing molecular weight. Relaxation leads to broad signals (more overlap) and unreasonable long measuring times for decent spectra. The accessible molecular weight range can be extended to about 50 kDa by fully deuterating [^{13}C, ^{15}N]-labeled proteins to reduce signal loss by relaxation. Experiments with [^{2}H, ^{13}C, ^{15}N]-labeled proteins permit sequence-specific resonance assignment of the backbone atoms and the heavy atoms in the side chains. However, it is still essential to assign side-chain protons because the through-space correlations between protons provide the main source of geometric information used to calculate the structures (21,53). Thus, in addition to fully deuterated proteins used for the sequential assignment, a partially deuterated protein is required to collect structural information from through-space proton-proton correlations.

COLLECTION OF CONFORMATIONAL CONSTRAINTS

For use in structure calculation, geometric conformational information in the form of distances and/or torsion angles has to be derived from the NMR data. In general, NMR data alone would not be sufficient to determine the positions of all atoms in a biological macromolecule. It has to be supplemented by information about the covalent structure of the protein—the amino acid sequence, bond lengths, bond angles, chiralities and planar groups—as well as by steric repulsion between nonbonded atom pairs. Although a variety of NMR parameters contain structural information, the crucial information comes from NOE measurements that provide distance information between pairs of protons (Figure 5). Supplementary constraints can be derived from through-bond correlations in the form of dihedral angles (21,53). Further, chemical shift data, especially from ^{13}C, provides information on the type of secondary structure (31,46,52), and hydrogen bonds can be detected via through-bond interactions (12). When NOEs are scarce (e.g., in partially deuterated proteins), additional constraints can be obtained from residual dipolar couplings that are observable in weakly aligned molecules. These couplings are related directly to the orientation of N-H and C-H internuclear vectors relative to the molecular frame (32,40,47). Most proteins in solution give rise only to extremely small residual dipolar couplings, and to enhance the effect, they are brought into a liquid-crystalline environment by adding proper additives to the aqueous solution (24,40,47).

The quality of a solution structure increases with the number of consistent input constraints used in the structure calculation. For a high-resolution structure determination, the maximum possible number of NOE constraints must be collected as input for the calculation of the complete 3-D protein structure. The NOE input data have the format of allowed distance ranges. The lower limit is usually taken to correspond to the sum of two hydrogen atomic radii,
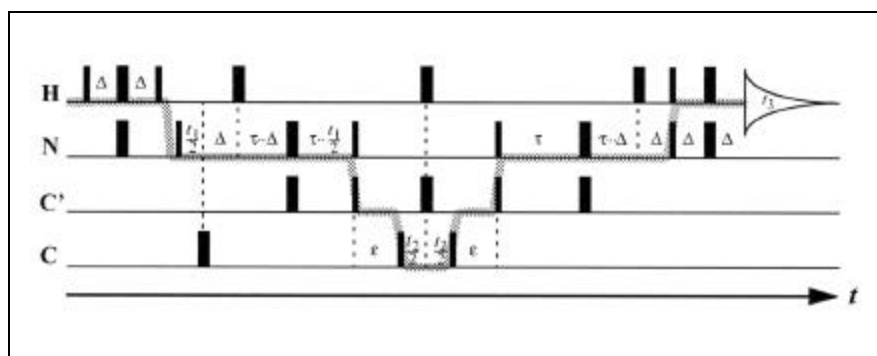
Figure 4. Simplified experimental scheme for an NMR measurement. The 3-D HN(CO)CA experiment (19,43) is used to illustrated an NMR pulse sequence. Narrow and wide black vertical bars represent 90° and 180° *rf* pulses, which are indicated on different horizontal lines for different *rf* frequencies: H, N, C′ and C stand for the resonance frequencies of the amide hydrogen, the amide nitrogen, the carbonyl carbon and the α-carbon nuclei, respectively. The time periods t_1, t_2 and t_3 denote the two chemical shift evolution periods and the detection period. The time periods labeled Δ, τ and ε have a fixed duration of 2.5, 17.6 and 8.5 ms, respectively. The gray line starting at the beginning of the line labeled H represents the flow of magnetization. The experiment starts with amide proton magnetization (H), which is transferred to the nitrogen nuclei attached to protons (N). In the next step, the magnetization is transferred from N to the bound carbonyl carbons C′ and further to the α-carbon C of the preceding amino acid residue. From the α-carbon, the magnetization is transferred back to the protons where it originated: the magnetization path goes back from C to C′ to N and finally to H.

0.2 nm. The upper limits are derived from the NOE intensities, typically in the range of 0.2–0.5 nm. In a folded protein, there are many proton pairs with an internuclear distance up to 0.5 nm; thus, NOESY spectra tend to be very crowded with peaks, even if they are measured as 3-D or even 4-D spectra to resolve the proton signals by the chemical shift of the attached ^{13}C and/or ^{15}N nucleus. There remain ambiguities that cannot be resolved based on the chemical shifts of the participating protons, and several possibilities may exist for the assignment of a distance to a pair of protons. In practice, as indicated in Figure 1, the determination of proton-proton distances is achieved in several cycles. Once a low-resolution structure has been calculated from an unambiguous subset of the NOE data, it is usually possible to discriminate between several possibilities, especially if some prospective proton pairs are not within 0.5 nm. In this way, successively more of the originally ambiguous distances can be assigned to specific pairs of protons, which leads to improved 3-D structures (Figure 6).

CALCULATION OF THE 3-D STRUCTURE

There are several computer programs available for the calculation of 3-D structures based on NMR constraints. Two approaches are mainly used (21). The first class of programs uses interatomic distances (25,34), and the second class works with torsion angles of the chemical bonds (22,23). In the end, all programs calculate the Cartesian coordinates of the spatial molecular structures, which are consistent with the set of constraints obtained from the NMR data (Figure 5). It must be kept in mind that the experimental constraints do not uniquely describe one exact 3-D structure because NMR-derived constraints typically describe a range of possible values and many distances cannot be determined. The structure calculation is thus repeated many times to determine an ensemble of structures consistent with the input data set. A good ensemble of structures minimizes violations of input constraints and samples the complete conformational space allowed by the constraints. For this reason, the NMR structures are usually represented by a bundle of structures (Figure 6). The quality of a structure is assessed by the calculation of the root-man-square deviation (RMSD) between the atoms of individual conformers in the bundle (Figure 6).

The two most common approaches for generation of structures use distance geometry (5,22) and restrained molecular dynamics (25,34). Distance geometry emphasizes the fact that the structure

Review

is derived using predominantly geometric criteria. Distance geometry algorithms determine ensembles of 3-D structures consistent with the distance constraints from NMR data and the chemical data (e.g., bond length, van der Waals' radii and planarity of aromatic rings). Two types of implementations exist that both do not operate in the normal 3-D space of atom positions. One class of algorithms works in the high dimensional space of all interatomic distances (metric matrix algorithm) where it is easy to find configurations that satisfy a given set of distance constraints. In the crucial step of these algorithms, the distance space must be converted into coordinates in the 3-D space, which represent one conformation that is compatible with the input data. The other class of distance geometry algorithms operates in torsional space, which retains only the essential degrees of freedom of a macromolecule, name-

ly the torsion angles of the chemical bonds. The covalent structural parameters are kept fixed, which reduces the number of degrees of freedom about tenfold compared to metric matrix algorithms. Determination of the structure starts from a polypeptide chain with the proper sequence in random conformation. The 3-D structure is then determined from the specified distance constraints by minimizing an error (target) function, which depends on the sum of all differences between the actual distance and the distance constraint (variable target function method).

The second approach for the generation of structures, restrained molecular dynamics (rMD), is based on classical mechanics and proceeds by numerically solving Newton's equation of motion to calculate the 3-D structure of a protein (25,34). The classical molecular dynamics force fields are supplemented (restrained) by square-well pseudo-en-

ergy terms based on the NMR-derived constraints. rMD has the advantage over distance geometry methods that the kinetic energy in the system allows energy barriers to be crossed, resulting in a sampling of a larger region of conformational space and producing structures with lower energy. This advantage can even be enhanced using simulated annealing (23,34). In this method, a simplified force field is used, and the energy barriers between states are reduced by setting the temperature to a high value (e.g., 1000 K). Subsequently the system is slowly cooled down (annealed) to a normal temperature of 300 K. Restrained molecular dynamics calculations can be performed in Cartesian as well as in torsion angle coordinates (torsion angle dynamics). However, currently, rMD using torsion angles can be performed about ten times faster than in Cartesian space (23).

In practice, a combination of

distance geometry and molecular dynamics is often used to calculate the structures (7,21). Initial conformations are generated by distance geometry that are used as starting structures for the molecular dynamics algorithms. For the highly efficient torsion angle dynamics calculation, only local NMR constraints (within four residues along the polypeptide chain) are usually considered, which results in an unfolded, extended starting conformer that can be obtained with minimal calculation time. For efficient rMD calculations in Cartesian coordinates, the starting structures should satisfy to a large extent the NMR-derived constraints before simulated annealing and restrained molecular dynamics calculations with proper force fields can be efficiently performed.

NMR WITH VERY LARGE MOLECULES

The foundation of successful NMR studies is high-quality spectra with good sensitivity and resolution. When studying larger molecules in solution, these basic requirements become harder to fulfill. Large molecules tumble slower, which results in faster relaxation and consequently broader lines in the NMR spectrum. Thus, the corresponding spectra show poor resolution and sensitivity. In practice, it becomes very hard to determine structures from proteins that have molecular weights above 40 kDa. Important advances in extending this size limit have recently been made with the introduction of novel NMR techniques and new biochemical approaches (49). Most importantly, relaxation in large molecules could be reduced with transverse relaxation-optimized spectroscopy (TROSY) (37), but also the potentially limiting spectral crowding is addressed by biochemical methods (61).

The TROSY technique reduces signal loss, which is the direct cause of the deterioration of the NMR spectra of large molecular structures. TROSY uses constructive interference between different relaxation mechanisms and
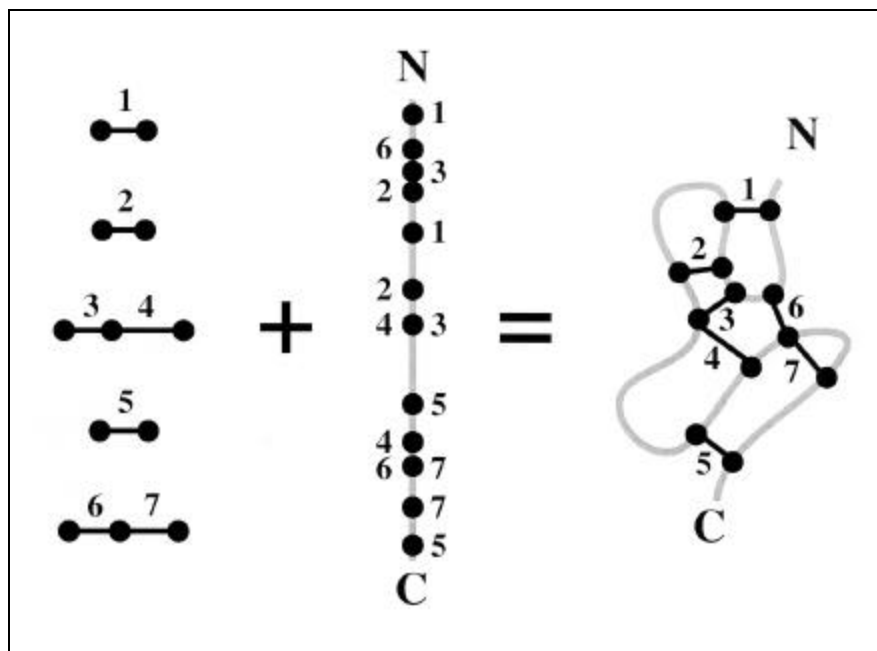
Figure 5. Principle of the structure determination by NMR. Schematic drawing of the use of distance constraints for the determination of the 3-D structure of biological macromolecules. On the left, distances obtained from NOESY spectra for seven pairs of protons are indicated and numbered from 1 to 7. In the middle, the gray vertical line represents the polypeptide chain, and the black dots represent the sequential positions of the individual protons that define the seven distances on the left-hand side of the figure. N and C mark the N and C terminus of the polypeptide chain. The protons are numbered according to the distances they define. Adding the information on the distance between pairs of protons and their position in the polypeptide sequence allows the construction of possible arrangements in which all distance constraints are fulfilled as indicated on the right-hand side of the figure.

Review

works best at the highest available magnetic field strengths in the range of 700–900 MHz proton resonance frequency (37,38). With TROSY, the molecular size of proteins accessible for detailed NMR investigations has been extended several fold. The TROSY experiments enable the recording of high-quality NMR spectra of macromolecules and supramolecular structures with molecular weights above 100 kDa (42). The TROSY technique benefits a variety of NMR experiments. Here, the HNCA experiments shall serve as an example (41,59). The HNCA experiment correlates the chemical shifts of the ^{1}H and the ^{15}N nuclei of an amide group with the intraresidual and the sequentially preceding α-carbon nuclei. The correlation with the two sequential α-carbons for each NH-group (Figure 7) allows sequential ordering of the amide groups, which finally results in sequence-specific assignments. Figure 7 shows a comparison of a 3-D TROSY-HNCA and a conventional 3-D HNCA spectrum that were measured with a 110-kDa protein under identical experimental conditions (41). The tremendous differences in signal intensities in favor of the TROSY-type spectrum are obvious. Only small parts, strips, of the complete 3-D spectrum are shown. In each strip, two peaks are expected that correspond to the two α-carbon nuclei adjacent to the amide

group represented by the strip. In Figure 7, the strip for Lys 121 is included as a control. As expected on the basis of the theory, the two spectra show similar signal intensities only for this residue, which is located at the flexible C-terminus of the polypeptide chain.

For larger molecules, there are more NMR-active nuclei and hence more resonance lines in the NMR spectra, which increases spectral overlap. New labeling techniques promise to alleviate the problem with spectral crowding (61). The methods use transsplicing or chemical ligation with inteins to reassemble proteins from two or more fragments (33,35,58,60). The complementing protein fragments are separately biosynthesized in bacteria with and without isotope labeling, respectively, and then combined to form a complete, segmentally labeled protein. The solution structure of only the labeled domain can then be determined by NMR. By labeling a different protein domain in each series of separate experiments, the structure of the entire large protein may be obtained using TROSY techniques, even for proteins that would otherwise produce overcrowded NMR spectra.

There are interesting biological systems that do not suffer from spectral crowding and are affected only by fast relaxation. These systems can readily be investigated by TROSY-type NMR ex-

periments. Such systems include symmetric oligomeric proteins or isotope-labeled proteins contained in unlabeled large particles, such as nucleic acid complexes, detergent micelles or lipid vesicles. Sequence-specific NMR assignments for such large structures have already been obtained (42), and TROSY-based NOESY experiments (39) for the collection of structural constraints are also available. It can be anticipated that, over the next few years, widespread use of the new techniques will result in many NMR structures for molecules and molecular complexes with a molecular weight exceeding 100 kDa.

CONCLUSION

Structure determination in solution using NMR is an established method that can routinely be applied to molecules with a molecular weight up to 30 kDa. There are some key requirements that must be fulfilled to permit a successful structure determination. The protein under investigation must be highly purified and soluble to a concentration of approximately 1 mM without aggregation. For molecules larger than 7–10 kDa, isotope enrichment in ^{13}C and ^{15}N and above 25 kDa possibly in ^{2}H is required. The NMR measurements have to be performed on a spectrometer operating at a proton reso-
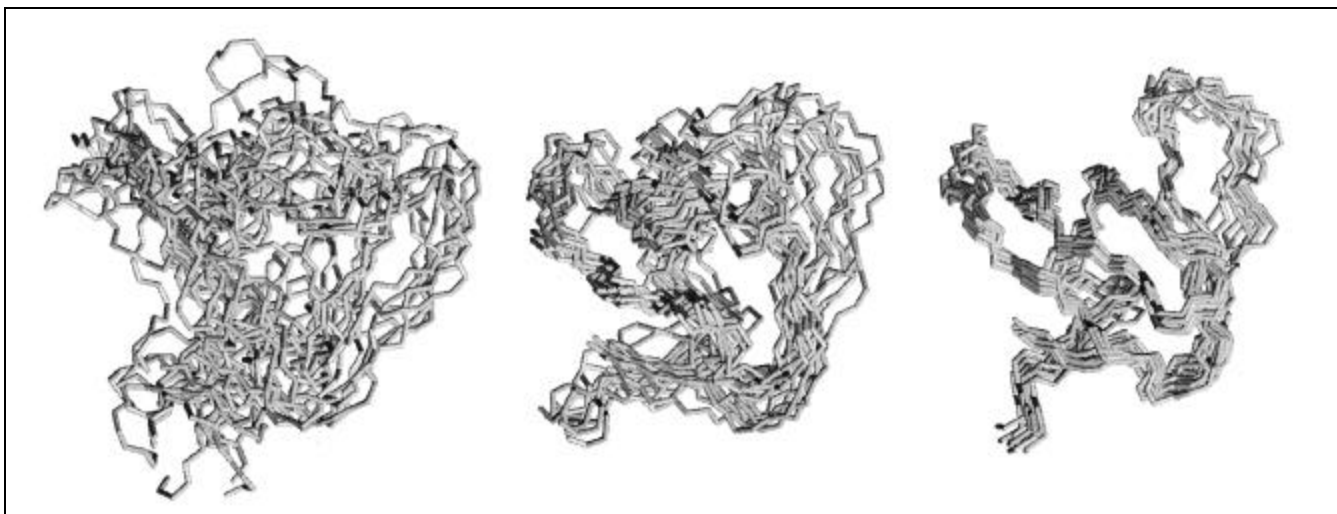
Figure 6. Protein backbone structures calculated with different numbers of NMR constraints. The structures show the SH$_3$ domain of human p56 Lck tyrosine kinase (26) at various stages of the assignment of additional distance constraints on the basis of preliminary structures (courtesy M. Salzmann). Backbone superpositions of ten conformers are shown with 1113 constraints (left), 1336 (middle) and 1687 constraints (right). (The RMSD between the positions of the polypeptide backbone atoms in the different conformers of the three bundles of structures are 4.2, 1.9 and 1.1 Å, respectively).

Labnet
4 C

Circle Reader Service No. 195

Review
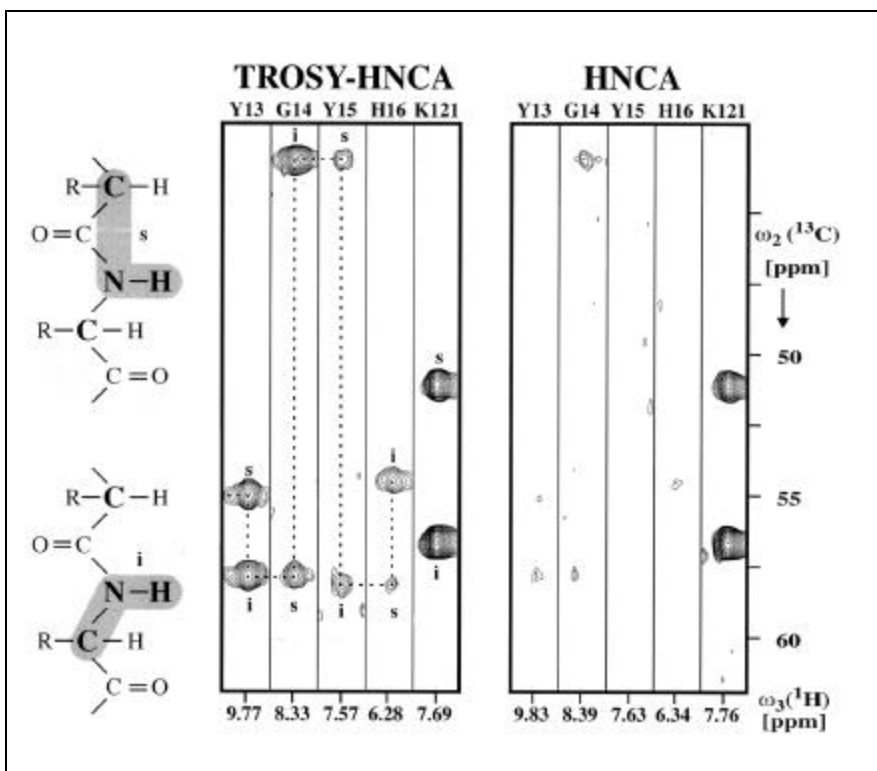
Figure 7. Comparison of a 3-D TROSY-HNCA and a conventional 3-D HNCA spectrum. Both spectra were measured with identical experimental conditions with a protein of 110 kDa molecular weight (41). The HNCA experiment correlates the chemical shifts of the ^{1}H and the ^{15}N of the amide group with the intraresidual and sequentially preceding α-carbon. This is indicated by the schematic drawing on the left. Only small parts, strips, of the complete 3-D spectrum are shown. For the presentation, the 3-D spectrum was separated in a stack of planes (Figure 3) along the ^{15}N dimension. From the [^{1}H, ^{13}C] planes, 130-Hz wide strips were cut out that run parallel to the ^{13}C axis. The strips are centered about the chemical shift of the amide proton attached to the ^{15}N represented by the plane. From both spectra, five strips are shown that represent the H-N-Cα correlations in the four sequential amino acids Tyr 13, Gly 14, Tyr 15, His 16 and in the C-terminal Lys 121 as indicated at the top of each strip by the one-letter amino acid symbol. The intraresidual correlation is marked with **i** and the sequential correlation with **s**. The broken line indicates the connections of the intraresidual and sequential HNCA correlations.

nance frequency of at least 600 MHz.

Although in this article little consideration was given to compounds other than peptides and proteins, the same methodology can be used with other biological macromolecules, in particular nucleic acids and their complexes with proteins and drugs. In addition, although not discussed in this review, NMR offers unique means of probing molecular motions on the picosecond-to-nanosecond and on the microsecond-to-millisecond time scales (28,36).

At the present time, only a few structures of proteins above 30 kDa molecular weight have been solved by NMR. One can anticipate, however, that in the not too distant future many more NMR structures of larger proteins and protein complexes will become available by the widespread use of TROSY, novel NMR experiments and creative isotope labeling schemes. In addition, the development of higher magnetic fields and improved spectrometer hardware will result in gains in resolution and sensitivity and will further increase the upper molecular weight limit for structural studies by NMR.

REFERENCES

1. **Aue, W.P., E. Bartholdi and R.R. Ernst.** 1976. Two-dimensional spectroscopy. Application to nuclear magnetic resonance. J. Chem. Phys. *64*:2229-2246.
2. **Bax, A. and S. Grzesiek.** 1993. Methodological advances in protein NMR. Acc. Chem. Res. *26*:131-138.
3. **Bernstein, F.C., T.F. Koetzle, G.J.B. Williams, E.F. Meyer, Jr., M.D. Brice, J.R. Rodgers, O. Kennard, T. Shimanouchi and M. Tasumi.** 1977. The protein data bank: a computer-based archival file for macromolecular structures. J. Mol. Biol. *112*:535-542.
4. **Bodenhausen, G. and D.J. Ruben.** 1980. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. Chem. Phys. Lett. *69*:185-189.
5. **Braun, W.** 1987. Distance geometry and related methods for protein structure determination from NMR data. Q. Rev. Biophys. *19*:115-157.
6. **Braun, D., G. Wider and K. Wüthrich.** 1994. Sequence-corrected ^{15}N "random coil" chemical shifts. J. Am. Chem Soc. *116*:8466-8469.
7. **Brünger, A.T., and M. Nilges.** 1993. Computational challenges for macromolecular structure determination by X-ray and solution NMR spectroscopy. Q. Rev. Biophys. *26*:49-125.
8. **Canet, D.** 1996. Nuclear Magnetic Resonance: Concepts and Methods. John Wiley & Sons, New York.
9. **Cavanagh, J., W.J. Fairbrother, A.G. Palmer III and N.J. Skelton.** 1996. Protein NMR Spectroscopy: Principles and Practice. Academic Press, San Diego.
10. **Christendat, D., A. Yee, A. Dharamsi, Y. Kluger, A. Savchenko, J.R. Cort, V. Booth, C.D. Mackereth et al.** 2000. Structural proteomics of an archaeon. Nat. Struct. Biol. *7*:903-909.
11. **Clore, G.M. and A.M. Gronenborn.** 1998. Determining the structures of large proteins and protein complexes by NMR. Trends Biotechnol. *16*:22-34.
12. **Cordier, F. and S. Grzesiek.** 1999. Direct observation of hydrogen bonds in proteins by interresidue $^{3h}J_{NC}$ scalar couplings. J. Am. Chem. Soc. *121*:1601-1602.
13. **Dötsch, V. and G. Wagner.** 1998. New approaches to structure determination by NMR spectroscopy. Curr. Opin. Struct. Biol. *8*:619-623.
14. **Dyson, H.J. and P.E. Wright.** 1996. Insights into protein folding from NMR. Annu. Rev. Phys. Chem. *47*:369-395.
15. **Ernst, R.R, G. Bodenhausen and A. Wokaun.** 1987 and 1994. Principles of Nuclear Magnetic Resonance in One and Two Dimensions. Clarendon Press, Oxford.
16. **Farrow, N.A., O. Zhang, J.D. Forman-Kay and L.E. Kay.** 1997. Characterization of the backbone dynamics of folded and denatured states of an SH3 domain. Biochemistry *36*:2390-2402.
17. **Gardner, K.H. and L.E. Kay.** 1998. The use of H-2, C-13, N-15 multidimensional NMR to study the structure and dynamics of proteins. Annu. Rev. Biophys. Biomol. Struct. *27*:357-406.
18. **Garrett, D.S., Y.-J. Seok, A. Peterkofsky, A.M. Gronenborn and G.M. Clore.** 1999. Solution structure of the 40,000 M_r phosphoryl transfer complex between the N-terminal domain of enzyme I and HPr. Nat. Struct. Biol. *6*:166-173.

Robbins
4C

Circle Reader Service No. 196

Review

19. **Grzesiek, S. and A. Bax.** 1992. Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. J. Magn. Reson. *96*:432-440.
20. **Grzesiek, S. and A. Bax.** 1993. The importance of not saturating H₂O in protein NMR. Application to sensitivity enhancement and NOE measurements. J. Am. Chem. Soc. *115*:12593-12594.
21. **Güntert, P.** 1998. Structure calculation of biological macromolecules from NMR data. Q. Rev. Biophys. *31*:145-237.
22. **Güntert, P., W. Braun and K. Wüthrich.** 1991. Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. J. Mol. Biol. *217*:517-530.
23. **Güntert, P., C. Mumenthaler and K. Wüthrich.** 1997. Torsion angle dynamics for NMR structure calculation with the new program DYANA. J. Mol. Biol. *273*:283-298.
24. **Hansen, M.R., L. Mueller and A. Pardi.** Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. Nat. Struct. Biol. *5*:1065-1074.
25. **Havel, T.F.** 1991. An evaluation of computational strategies for use in the determination of protein structure from distance constraints obtained by nuclear magnetic resonance. Prog. Biophys. Mol. Biol. *56*:43-78.
26. **Hiroaki, H., W. Klaus and H. Senn.** 1996. Determination of the SH₂ domain of the human p56 Lck tyrosine kinase. J. Biomol. NMR *8*:105-122.
27. **Jeener, J., B.H. Meier, P. Bachmann and R.R. Ernst.** 1979. Investigation of exchange processes by two-dimensional NMR spectroscopy. J. Chem. Phys. *71*:4546-4553.
28. **Kay, L.E.** 1998. Protein dynamics from NMR. Biochem. Cell Biol. *76*:145-152.
29. **Kay, L.E. and K.H. Gardner.** 1997. Solution NMR spectroscopy beyond 25 kDa. Curr. Opin. Struct. Biol. *7*:722-731.
30. **Kumar, A., R.R. Ernst and K. Wüthrich.** 1980. A 2D NOE experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. Biochem. Biophys. Res. Commun. *95*:1-6.
31. **Luginbühl, P., T. Szyperski and K. Wüthrich.** 1995. Statistical basis for the use of $^{13}C^\alpha$ chemical shifts in protein structure determination. J. Magn. Reson. B *109*:229-233.
32. **Mueller, G.A., W.Y. Choy, D. Yang, J.D. Forman-Kay, R.A. Venters and L.E. Kay.** 2000. Global folds of proteins with low densities of NOEs using residual dipolar couplings: application to the 370-residue maltodextrin-binding protein. J. Mol. Biol. *300*:197-212.
33. **Muir T.W., D. Sondhi and P.A. Cole.** 1998. Expressed protein ligation: a general method for protein engineering. Proc. Natl. Acad. Sci. USA *95*:6705-6710.
34. **Nilges, M., G.M. Clore and A.M. Gronenborn.** 1988. Determination of three-dimensional structures of proteins from interproton distance data by hybrid distance geometry-dynamical simulated annealing calculations. FEBS Lett. *229*:317-324.
35. **Otomo, T., K. Teruya, K. Uegaki, T. Yamazaki and Y. Kyogoku.** 1999. Improved segmental isotope labeling of proteins and application to a larger protein. J. Biol. NMR *14*:105-114.
36. **Palmer, A.G.** 1997. Probing molecular motion by NMR. Curr. Opin. Struct. Biol. *7*:732-737.
37. **Pervushin, K., R. Riek, G. Wider and K. Wüthrich.** 1997. Attenuated T_2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc. Natl. Acad. Sci. USA *94*:12366-12371.
38. **Pervushin, K., R. Riek, G. Wider, and K. Wüthrich.** 1998. Transverse relaxation-optimized spectroscopy (TROSY) for NMR studies of aromatic spin systems in ¹³C-labeled proteins. J. Am. Chem. Soc. *120*:6394-6400.
39. **Pervushin, K., G. Wider, R. Riek and K. Wüthrich.** 1999. The 3D NOESY-[¹H,¹⁵N, ¹H]- ZQ-TROSY NMR experiment with diagonal peak suppression. Proc. Natl. Acad. Sci. USA *96*:9607-9612.
40. **Prestegard, J.H.** 1998. New techniques in structural NMR—anisotropic interactions. Nat. Struct. Biol. *5*:517-522.
41. **Salzmann, M., K. Pervushin, G. Wider, H. Senn and K. Wüthrich.** 1998. TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. Proc. Natl. Acad. Sci. USA *95*:13585-13590.
42. **Salzmann, M., K. Pervushin, G. Wider, H. Senn and K. Wüthrich.** NMR Assignment and secondary structure determination of an octameric 110 kDa protein using TROSY in triple resonance experiments. J. Am. Chem. Soc. *122*:7543-7548.
43. **Salzmann, M., G. Wider, K. Pervushin, H. Senn and K. Wüthrich.** 1999. TROSY-type triple-resonance experiments for sequential assignments of large proteins. J. Am. Chem. Soc. *121*:844-848.
44. **Sattler, M., J. Schleucher and C. Griesinger.** 1999. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Progr. NMR Spectrosc. *34*:93-158.
45. **Solomon, I.** 1955. Relaxation processes in a system of two spins. Phys. Rev. *99*:559-565.
46. **Spera, S. and A. Bax.** 1991. Empirical correlation between protein backbone conformation and C^α and C^β ¹³C nuclear-magnetic-resonance chemical-shifts. J. Am. Chem. Soc. *113*:5490-5492.
47. **Tjandra, N. and A. Bax.** 1997. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. Science *278*:1111-1114.
48. **Wider, G.** 1998. Technical aspects of NMR spectroscopy with biological macromolecules and studies of hydration in solution. Progr. NMR Spectrosc. *32*:193-275.
49. **Wider, G. and K. Wüthrich.** 1999. NMR spectroscopy of large molecules and multimolecular assemblies in solution. Curr. Opin. Struct. Biol. *9*:594-601.
50. **Wider, G., R.V. Hosur and K. Wüthrich.** 1983. Suppression of the solvent resonance in 2D NMR spectra of proteins in H₂O solution. J. Magn. Reson. *52*:130-135.
51. **Wider, G., S. Macura, Anil Kumar, R.R. Ernst and K. Wüthrich.** 1984. Homonuclear Two-dimensional ¹H NMR of proteins. Experimental Procedures. J. Magn. Reson. *56*:207-234.
52. **Wishart, D.S. and B.D. Sykes.** 1994. The ¹³C chemical-shift index—a simple method for the identification of protein secondary structure using ¹³C chemical-shift data. J. Biomol. NMR *4*:171-180.
53. **Wüthrich, K.** 1986. NMR of Proteins and Nucleic Acids. John Wiley & Sons, New York.
54. **Wüthrich, K.** 1996. Biological macromolecules: structure determination in solution, p. 932-939. *In* D.M. Grant, R.K. Harris (Eds.), Encyclopedia of Nuclear Magnetic Resonance, Vol. 2. John Wiley & Sons, New York.
55. **Wüthrich, K.** 1997. NMR—an alternative to x-ray cristallography for protein and nucleic acid structure determination, p. 242-257. *In* A. Haase, G. Landwehr and E. Umbach (Eds.), Röntgen Centennial: X-rays in Natural and Life Sciences. World Scientific, Singapore.
56. **Wüthrich, K.** 1998. The second decade—into the third millennium. Nat. Struct. Biol. *5*:492-495.
57. **Wüthrich, K., M. Billeter, P. Güntert, P. Luginbühl, R. Riek and G. Wider.** 1996. NMR studies of the hydration of biological macromolecules. Faraday Disc. *103*:245-253.
58. **Xu, R., B. Ayers, D. Cowburn and T.W. Muir.** 1999. Chemical ligation of folded recombinant proteins: segmental isotopic labeling of domains for NMR studies. Proc. Natl. Acad. Sci. USA *96*:388-393.
59. **Yamazaki, T., W. Lee, M. Revington, D.L. Mattiello, F.W. Dahlquist, C.H. Arrowsmith and L.E. Kay.** 1994. An HNCA pulse scheme for the backbone assignment of ¹⁵N, ¹³C, ²H-labeled proteins: application to a 37-kDa *Trp* repressor-DNA complex. J. Am. Chem. Soc. *116*:6464-6465.
60. **Yamazaki, T., T. Otomo, N. Oda, U.K. Kyogoku, K. Uegaki, N. Ito, Y. Ishino and H. Nakamura.** 1998. Segmental isotope labeling for protein NMR using peptide splicing. J. Am. Chem. Soc. *120*:5591-5592.
61. **Yu, H.** 1999. Extending the size limit of protein nuclear magnetic resonance. Proc. Natl. Acad. Sci. USA *96*:332-334.

Address correspondence to:

Dr. Gerhard Wider
*Institut für Molekularbiologie und Biophysik
Eidgenössische Technische Hochschule
 Hönggerberg*
CH - 8093 Zürich, Switzerland
e-mail: wider@mol.biol.ethz.ch