*Russian Chemical Reviews* **71** (12) 1005 – 1040 (2002) © 2002 Russian Academy of Sciences and Turpion Ltd

DOI 10.1070/RC2002v071n12ABEH000757

# Determination of the three-dimensional structure for weakly aligned biomolecules by NMR spectroscopy

A A Shahkhatuni, A G Shahkhatuni

## Contents


| | | |
|---|---|---|
| I. | Introduction | 1005 |
| II. | Methods of investigation of the structure of non-oriented biomolecules in solution | 1007 |
| III. | Study of membrane proteins by solid-state NMR spectroscopy | 1009 |
| IV. | Methods for alignment of biomolecules in solutions | 1010 |
| V. | Measurement of orientation-dependent parameters | 1016 |
| VI. | Strategy of structure determination for weakly aligned biomolecules by NMR spectroscopy | 1020 |
| VII. | The use of NMR spectroscopy for the study of the structures of biomolecules | 1025 |
| VIII. | Conclusion | 1030 |
| IX. | Supplement | 1030 |


**Abstract. The key achievements and the potential of NMR spectroscopy for weakly aligned biomolecules are considered. Due to weak alignment, it becomes possible to determine a number of NMR parameters dependent on the orientation of biomolecules, which are averaged to zero in usual isotropic media. The addition of new orientational constraints to standard procedures of 3D structure determination markedly increases the achievable accuracy. The possibility of structure determination for biomolecules using only orientation-dependent parameters without invoking other NMR data is discussed. The methods of orientation, experimental techniques, and calculation methods are systematised. The main results obtained and the prospects of using NMR spectroscopy of weakly aligned systems to study different classes of biomolecules and to solve various problems of molecular biology are analysed. Examples of biomolecules whose structures have been determined using orientation-dependent parameters are given. The bibliography includes 508 references.**

## I. Introduction

The advances achieved in implementation of the programme of the human genome decoding marked a new stage in investigations of the structure and functioning of living organisms. Study of the human genome entailed the research into the genomes and gene products of many other organisms. The primary structures of many biomolecules have already been determined; however, for understanding of the functioning mechanisms, one should know the spatial structure of these molecules.

The key experimental methods used to determine the three-dimensional structures of biomolecules include X-ray crystallography (X-ray) and nuclear magnetic resonance spectroscopy (NMR).

**A A Shahkhatuni, A G Shahkhatuni** Molecular Structure Research Centre, National Academy of Sciences of the Republic of Armenia, prosp. Azatutyan 26, 375014 Yerevan, Republic of Armenia.
Fax (374-1) 28 13 80. E-mail: astra@msrc.am (A A Shahkhatuni)
Tel. (374-1) 28 74 23. E-mail: alexsh@msrc.am (A G Shahkhatuni)

Received 26 July 2002
*Uspekhi Khimii* **71** (12) 1132 – 1172 (2002); translated by Z P Bobkova

Currently, The Protein Data Bank (PDB) [1] stores the results of structural studies for ~ 19 000 proteins and their complexes; however, only ~ 6000 structures are different and of these, ~ 1000 structures have essentially different homologies. The 3D structures for an enormous number of proteins are still to be determined.

By the end of 2002, ~ 83% of the structures of known biomolecules have been determined by X-ray, ~ 16%, by NMR, and ~ 1%, by other methods (mainly, by neutron diffraction). The proportion of structures determined using NMR steadily grows. Indeed, by the end of 1997, 1008 structures of various biomolecules had been determined using NMR; during the following four years, this number doubled, and by the end of 2002, it trebled. Moreover, for ~ 75% of these biomolecules, no data on the crystal structure was included in the PDB.[2] Nuclear magnetic resonance proved especially efficient for investigation of nucleic acids. It was used to determine more than one-third of all DNA structures and 60% of RNA structures.

Comparison of numerous structural data obtained by NMR and X-ray diffraction methods showed a high degree of compliance, indicating that NMR is a worthy companion for X-ray, although the latter remains the predominant method used to determine the spatial structures of biomolecules. However, many proteins (for example, most of membrane proteins) cannot be crystallised; therefore, X-ray diffraction is inapplicable to them. In addition, to elucidate the role of proteins in various processes, one should know not only their structures but also the dynamic behaviour, furthermore, under conditions resembling, as close as possible, the environment [either liquid or liquid-crystalline (LC)] in which these proteins exist as well as the structures and dynamic behaviour of their partners and complexes. First of all, this refers to processes such as enzymic catalysis, immune response, signal transduction, regulation of protein synthesis, *etc*. These problems can be solved by using NMR spectroscopy, while X-ray diffraction is not always suitable for this purpose.[3, 4]

Despite the fact that NMR and X-ray techniques are based on different kinds of physical phenomena, they have much in common, first, investigation objects; second, computation software that allows processing of both X-ray and NMR data; and third, databases which accumulate, archive and annotate the obtained structural information of biomolecules. Therefore, it is not a surprise that most of the new NMR research centres are established at large functioning sources of synchrotron radiation,

which is used most often in structural studies. The resolution achieved in both methods is 1.0 – 1.5 Å and the amount of specimen required is 30 – 50 mg. Currently, NMR techniques are still much inferior to X-ray as regards the time and cost of determination of one structure. However, owing to the fast development of research using NMR, these parameters may soon become comparable for both methods.

The progress made due to the use of synchrotron radiation, automated systems for sample crystallisation and handling, effective array detectors for data collection and processing, and development of the multiwavelength anomalous diffraction method has transformed X-ray into a fast and convenient method for determining the structures of biomacromolecules. The NMR spectroscopy methods are used predominantly only at a few research centres. In particular, NMR spectroscopy is becoming the major component in investigations of the 3D structures of biomolecules in at least several large research centres.[2,5] For example, twenty high-frequency NMR spectrometers were installed last year at the RIKEN Genomic Sciences Centre in Japan. It is evident that NMR spectroscopy will be the main tool of investigating the structure – biological functions relationships for proteins.

The modern structural NMR spectroscopy has formed in two phases. The first phase (1985 – 1995) was characterised by vigorous development and active use of multidimensional Fourier Transform spectroscopy methods † for the study of structures of biomolecules.[6–21] Initially, the measurements were carried out in liquid isotropic solutions using two- or three-dimensional homonuclear NMR methods based on the nuclear Overhauser effect (NOE), and the weights of macromolecules did not exceed 10 kDa.

Further progress of the research was held up by two restrictions of principle. First, the widths of resonance lines in the NMR spectrum in solution are inversely proportional to the spin – spin relaxation time $T_2$, which decreases as the size of the molecule increases, and to the rotational correlation time. This diminishes the sensitivity of the pulse sequences used, as long delays are required for coherence transfer.

Second, larger molecules give rise to more complicated spectra, because the numbers of NMR-active nuclei and interactions between them increase.

In the early 1990s, methods of selective $^{15}N/^{13}C$-enrichment of biomolecules as well as multidimensional heteronuclear spectroscopy and triple resonance spectroscopy with pulse gradients were developed. Thus the mass of eligible proteins was increased to 25 kDa.[22] With the progress in a NMR hardware (the appearance of high-frequency spectrometers operating at 700, 800 and 900 MHz, probes with magnetic field pulse gradients, *etc*.) and with the advent of procedures for partial and selective deuteration of biomolecules, it became possible to study molecules with higher masses, up to 35 – 40 kDa.[23–25] Combination of triple resonance spectroscopy with $^{15}N/^{13}C$-enrichment and perdeuteration allows investigation of proteins weighing up to 50 – 60 kDa. The up-to-date TROSY (Transverse Relaxation Optimized Spectroscopy) and CRINEPT (Cross-Correlated Relaxation-Enhanced Polarisation Transfer) methods[26–33] have increased this value to more than 100 kDa.[34–36] For instance, the structure of an octameric protein with a 110 kDa mass has been determined.[37] Certainly, this is not the limiting value. For example, the segment-by-segment isotope labelling procedure reported recently,[38,39] the decrease in the rotational correlation time by using low-viscosity solvents,[40] and the use of cryogenic probes and high-frequency NMR spectrometers (with an operating frequency of 1 GHz or higher) provide the possibility of increasing substantially the size of molecules that can be studied, especially resorting to new, in principle, ideas and approaches.[3,41]

The second phase in the study of the structures of biomolecules by NMR spectroscopy, which started in 1995, was marked by the advent of a new approach. It was proposed to use the orientation of biomolecules on exposure to the spectrometer magnetic field.[42,43] For aligned molecules, one can determine additional parameters of NMR signals that bear structural information, first of all, the residual dipolar coupling constants (RDC).[44] However, even in strong magnetic fields, the orientation of biomolecules was very weak; therefore, it was possible to measure only the RDC for directly linked nuclei in biomolecules with highly anisotropic magnetic susceptibility. The accuracy of structure determination in these experiments was not inferior to that in NOE-based experiments but still left much to be desired.

Tjandra and Bax[45] proposed using micellar solutions and dilute LC phases for the orientation of biomolecules. In these media, the degree of alignment of biomolecules can be easily controlled by changing experimental conditions, for example, component concentrations and temperature. In this case, one can select the optimal values of the order parameters and to simplify the spectra.

By using NMR methods, one can determine a number of physical characteristics bearing information on the interatomic distances and angles in the molecule. Most of these values are orientation dependent, *i.e*., they depend on the orientation of biomolecules relative to the direction of the magnetic field; in usual isotropic solutions, they are averaged to zero due to the fast rotation of molecules. Nevertheless, there exist some parameters containing indirect information on the distances between the magnetic nuclei and on dihedral angles which can be elucidated in experiments with isotropic solutions. For example, these are nuclear spin relaxation times in NOE experiments depending on the dipole – dipole couplings between them; the spin – spin coupling constants through three covalent bonds ($^3J$) or cross-correlated relaxation rates depending on the dihedral angle between these bonds. In addition, the chemical shifts of some protons and carbon atoms depend crucially on the environment and can be useful for determining the tertiary structure of proteins.

Despite the low accuracy of the results, the complexity and high cost, NOE experiments are widely used to determine the spatial structures of biomolecules in solutions. Recently, several other parameters characterising specific properties of various types of relaxation were added to the known parameters, $^3J$ and chemical shifts. However, by using isotropic parameters, one can establish only the local structure, whereas the use of orientation-dependent values allows determination of both the local and long-range structures.

Direct quantitative measurement of orientation-dependent parameters, for example, RDC, from the splitting observed in the spectra of oriented molecules is certainly preferable to rough evaluation of these parameters in NOE experiments using qualitative analysis of cross-peak intensities for isotropic solutions. In the former case, we are dealing with an effect that can be interpreted using only the first-order perturbation theory; in the latter case, experimental results can be interpreted only by resorting to higher-order perturbation theory (*i.e*., a much rougher approximation); moreover, they can be masked by other effects. In addition, RDC can be measured faster and more accurately (and between a much larger number of different nuclei) in the spectra of oriented specimens than only between protons in NOE experiments. This means that RDC are more suitable for structural studies, as was brilliantly demonstrated in a publication[46] in which the 3D structure of a protein backbone was determined using only residual dipolar couplings without invoking the data of NOE experiments. This publication marks a change in the priority in the structural studies of biomolecules by NMR spectroscopy: cumbersome NOE-based experiments in isotropic media with approximate estimation of the distances between some close

† In 1991, Ernst (Switzerland) was awarded the Nobel Prize in chemistry for these studies.

protons are replaced by fast and efficient experiments with measurement of more informative and diverse parameters for oriented molecules directly related to their geometry.

A number of methods for the measurement of orientation-dependent NMR parameters have now been developed and development of new methods is in progress. They are based on investigations of biomolecules, either completely aligned in the solid phase or partially aligned in a liquid or LC medium. These methods are equally good for investigating both the structures of biomolecules and conformational changes, for determining the relative orientation of the domains, the formation of intermolecular complexes and for investigation of dynamic processes.[47,48] Experiments with weak alignment of biomolecules have been most successful so far, although it is expected that solid-state NMR techniques for highly oriented biomolecules (in the solid state) would prove to be more efficient in the study of structures and functions of membrane proteins.[49]

In modern approaches, the new methods mentioned above are combined with 'classical' methods based on the NOE, the use of vicinal $^3J$ constants, chemical shifts and relaxation times. This permits reliable determination of the structures of proteins with a mass of up to 100 kDa and even more. The research of biomolecules by NMR spectroscopy can be classified into three groups of studies carried out in parallel:[18–21]

— in liquid isotropic media;
— in weakly aligned media;
— solid-state experiments with magic angle spinning.

This review describes the scope of NMR spectroscopy of weakly aligned biomolecules in solutions. The procedures used in NMR studies of non-aligned molecules are also outlined. The ways of orienting biomolecules, the types of orientation-dependent parameters, experimental procedures, and calculation methods are described in detail; the main results and prospects for the use of NMR spectroscopy of weakly aligned molecules in the research of proteins, nucleic acids and oligosaccharides are analysed.

## II. Methods of investigation of the structure of non-oriented biomolecules in solution

The structure of a biomolecule can be represented by either a set of Cartesian coordinates ($x$, $y$, $z$) of each atom or by a set of intrinsic coordinates [for proteins, these are bond lengths and angles $\phi$, $\psi$, $\omega$, $\chi_1$, and $\chi_2$ for each amino-acid residue (Fig. 1)].

### 1. Nuclear Overhauser effect and its use in structural studies

Nuclear magnetic resonance techniques based on NOE have been used successfully for more than 20 years to elucidate the three-dimensional structures of biomolecules in solutions. NOE parameters are determined in 2D and 3D experiments and are used subsequently to calculate the interatomic distances. The studies of small proteins, transport RNA, oligonucleotides and interactions of proteins with ligands and nucleic acids are described in detail in many publications.[7–25,50]

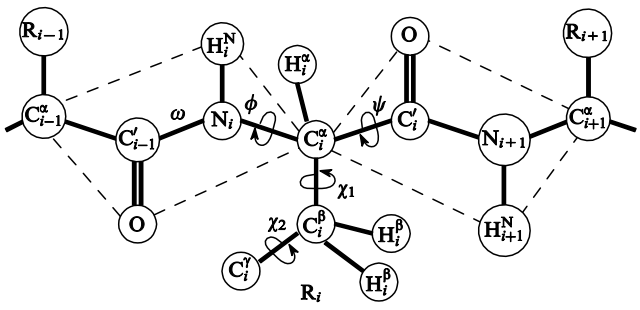



**Figure 1.** Fragment of the protein polypeptide chain.
Dashed lines designate the peptide planes. The backbone conformation is determined by the angles $\phi$ and $\psi$ for each amino acid residue R and the side chain conformations are determined by the angles $\chi_1$ and $\chi_2$.

The NOE is underlain by the fact that the rate constant for the dipole cross-relaxation is inversely proportional to the sixth power of the distance between the interacting nuclei. The intensities of NOE cross-peaks, which are proportional to the rate constants for cross-relaxation are determined in experiments. Thus, if one distance between protons ($R_0$) is known (for example, from the trivial covalent geometry of the molecule), the other, unknown $R_x$ can be determined from the relation

$$R_x = \left(\frac{I_o}{I_x}\right)^{1/6},$$

where $I_0$ and $I_x$ are the integral intensities of the corresponding cross-peaks.

A 2D (or 3D) NMR spectrum of a protein contains numerous NOE cross-peaks corresponding to different pairs of protons. Depending on the intensity, they are classified into strong, medium-intensity and weak. Strong cross-peaks correspond to protons neighbouring *via* covalent bonds (for example, geminal protons in the methylene groups or vicinal protons in aromatic rings) and to neighbouring protons in a regular secondary structure of the protein (for example, $H^N-H^\alpha$ along the chain, $H^\alpha-H^\alpha$ and $H^N-H^\alpha$ in β-sheets). The medium-intensity cross-peaks usually correspond to $H^N-H^\alpha$ pairs in helices. Even for proteins with a mass below 10 kDa, the number of these responses is about 100; this is sufficient for determining the spatial structure of the protein globule. The structures of many proteins with masses of up to 100 kDa have now been determined using several hundred (in some cases, above several thousand) cross-peaks.

The major advantage of these studies over X-ray studies is that the structure of a biomolecule is determined in a solution rather than in a single crystal. The structures of some biomolecules are known to change upon crystallisation; therefore, structure determination in solution is preferred, in particular, for physicians and biologists. In addition, the molecules in the crystal are exposed to anisotropic crystal fields; therefore, discrepancy between the structures determined in the crystal state and in solution does not mean by itself that either of these structures is erroneous.

However, the NOE arises only between closely spaced (not more than 5–6 Å) magnetic nuclei (mainly, protons), whereas X-ray diffraction analysis allows determination of the position of any atom. Yet another important difference between these methods is manifested in investigations of mobile molecules. The crystal structure determined by X-ray is a simple superposition of all molecules present in the crystal, whereas NOE experiments provide certain averaged $(\langle r_{ij}^{-6}\rangle)^{-1/6}$ values rather than the distances themselves. This implies that in the former case, the contributions of all fluctuations to the distances between the nuclei in the crystal structure are identical, while in the latter case, the contributions of states with closely located protons to the measured distances are much greater than the contributions of states with remote protons, resulting in apparent shortening of the averaged distance.

Nuclear Overhauser effect experiments successfully supplement X-ray studies, especially in those cases where crystal data are not readily available, in a comparative study of oligonucleotides with different sequences, and in a study of transitions between various DNA conformations.

Until recently, data on the geometry of biomolecules in solutions have been mainly gained in NOE experiments. However, the knowledge of distances between protons found in these experiments does not suffice for structure determination. Therefore, the use of any additional information is useful both for structure refinement and for facilitating the structure determination procedure.

As the mass of the biomolecule increases, the NOE signals are broadened, which may result in line coalescence in the spectra.

This phenomenon can be controlled by using partial deuteration; however, this markedly decreases the number of NOE signals and increases the cost of structure determination. This makes the additional information obtained in other experiments even more significant.

## 2. Spin – spin coupling constants

The use of spin – spin coupling constants through three bonds $^3J$ in structural studies is based on the fact that they obey an empirical dependence on the dihedral angle $\theta$ described by Karplus:[10, 13]

$$^3J(\theta) = A\cos^2\theta + B\cos\theta + C.$$

The coefficients $A$, $B$ and $C$ depend on the types of particular nuclei forming the dihedral angle; they are determined during optimisation of experimental $^3J$ constants and the angles found from high-resolution X-ray diffraction data for a multitude of molecules. There are several of these constants for protein molecules. The constant $^3J(H^N - H^\alpha)$ directly related to the torsion angle $\phi$ of the peptide chain is the most convenient for structural studies, as it can be determined rather simply and accurately. Data on the other torsion angle of the peptide chain $\psi$ are contained in the $^3J(N - H^\alpha)$ and $^3J(N - N)$ constants, but, unfortunately, they are both small and, moreover, change the sign. In return, data on the conformations of side chains, in particular on the torsion angle $\chi_1$, can be derived from the spin – spin coupling constants for different pairs: $H^\alpha - H^\beta$, $N - H^\beta$, $C' - H^\beta$, $H^\alpha - C^\gamma$, $N - C^\gamma$, $C' - C^\gamma$.

The possibility of using the spin – spin coupling constants through one or two bonds in structural research has been barely studied until recently. However, the advent of NMR spectroscopy of weakly aligned molecules brought about the necessity of separating the contributions of the coupling constants and RDC through one bond to the observed splitting because a researcher measures a spin – spin coupling constant every time he determines a RDC. At first glance, these constants provide little information and are measured with a too low accuracy, but they can still be used. For example, the $C^\alpha - H^\alpha$ and $C^\alpha - C^\beta$ constants proved to be 'sensitive' to the torsion angle $\psi$ of the peptide chain.[51] The number of spin – spin coupling constants suitable for conformational studies of nucleic acids and carbohydrates is greater and their information content is higher; therefore, they have been better studied.[52] In addition, the knowledge of spin – spin coupling constants is important due to the fact that they allow one to plan NMR experiments with coherence transfer *via* the corresponding bonds.

## 3. Chemical shifts

The chemical shifts of nuclei depend on their local environment and, actually, they bear all information needed to determine the 3D structure of the molecule. However, calculation of a 3D structure using chemical shifts alone has not been performed so far, although some progress along this line has been achieved.[53, 54]

The chemical shifts of the $^1H$, $^{15}N$, $^{13}C$ and $^{31}P$ nuclei are widely used as constraints in determining the structure of a biomolecule. Empirical correlations between the $\phi$ and $\psi$ angles, which determine the conformation of a section of the peptide chain, and the $H^N$, $H^\alpha$, $^{15}N$, $^{13}C'$, $^{13}C^\alpha$ and $^{13}C^\beta$ chemical shifts are known.[55] For example, the $^{13}C^\alpha$ and $^{13}C^\beta$ chemical shifts in the α-helices and β-sheets are markedly different, which allows reliable identification of these elements of the protein secondary structure. Similar correlations have been found for nucleic acids.[52]

The structural information contained in these chemical shifts duplicates to some extent the information obtained using $^3J$ constants (for example, data on $\phi$ angles). However, not all the $^3J$ constants can be measured (because they are too small or because of broadening or overlap of spectral lines), whereas chemical shifts can be determined for any atom.

The signal assignment and chemical shift determination procedures are almost entirely automated; they are performed at an initial stage of structural studies. When these values are known, one can establish and locate rather reliably the elements pertaining to the secondary structure of the biomolecule,[56] which is important for further analysis.

The chemical shifts of nuclei are determined not only to obtain tentative data concerning the structure; their changes during functioning of the biomolecule can serve as a fast-response tool suitable for investigating biochemical processes, screening low-molecular-mass ligands, and determining the sites of their binding to the biomolecule. The results of these studies are required, first of all, for the development of drugs.

## 4. Determination of constants through one hydrogen bond

Hydrogen bonds are the key stabilising factor which determines the secondary and tertiary structures of biomolecules. Until recently, the existence of hydrogen bonds could be inferred only indirectly from the NMR parameters, for example, from the rates of proton exchange, $^1H$ and $^{15}N$ chemical shifts, isotope shifts or the spatial proximity of partners.

However, it was found that the constants through one hydrogen bond can be determined rather easily for the molecules of nucleic acids or proteins. In particular, the $^2J_{N\cdots HN}$ constants between the $^{15}N$ atoms of the purine and pyrimidine bases in the Watson – Crick pairs have been measured for both RNA[57] and DNA.[27] The spin – spin coupling constants found are rather high ($^2J_{NN}$ is about 6 – 7 Hz); they depend on the length of the hydrogen bond.[58] Therefore, measurement of the through-hydrogen-bond constants is not only indicative of the presence of this bond (this can also be inferred indirectly from the chemical shifts) but also allows one to determine the bond length.[59, 60] Smaller constants ($^1J_{HN}$ of about 2 – 4 Hz) have been found between the imine group proton and its $^{15}N$ acceptor in the Watson – Crick pairs.

For proteins, the $^3J_{CO\cdots HN}$ constants have also been found.[61 – 63] They are equal to 0.2 – 1.0 Hz and also depend on the length of the hydrogen bond.[63]

## 5. Relaxation parameters

One relaxation parameter, the intensity of NOE peaks, has been mentioned above.

The times of heteronuclear relaxation contain information on the intramolecular dynamics. For example, the spin – spin relaxation time $T_2$ for a $^{15}N$ nucleus is a function of both frequency-dependent and frequency-independent constituents of the spectral density, and the spin – lattice relaxation time $T_1$ is a function of only frequency-dependent components. Due to anisotropy of rotational diffusion in the absence of large-amplitude motions in the molecule (and, hence, without broadening of NMR lines), the $T_1/T_2$ ratio for the $^{15}N$ nuclei in a protein molecule with an axially symmetrical diffusion tensor depends only on three variables: the angle between the vector of the N – H bond and the principal axis of the diffusion tensor and the parallel and perpendicular components of the diffusion tensor. This fact can be used to determine the orientation of the N – H bond vector, and the $T_1/T_2$ ratio can be included as an independent variable into the group of varied parameters for determining the structure of a biomolecule.[64, 65]

In recent years, methods based on cross-correlation (or interference) of the relaxation mechanisms related to the dipole – dipole interaction (DD) and chemical shift anisotropy (CSA) have been used for gaining structural and dynamic information on biomolecules.[66] These methods are based on elucidation and use of DD – DD,[67 – 69] DDI – CSA[70 – 73] or CSA – CSA[74] type cross-correlations. One more type of cross-correlation is found in paramagnetic proteins, namely, that between the DD and the Curie spin, which is the net magnetic moment of the system of interacting nuclei and the unpaired electron.[75] These tensor correlations are very convenient for studying the dynamics[76] and determining various torsion angles in biomolecules.[77, 78]

For example, to find the $\psi$ angle in proteins, one can measure the following relaxation rates: $DD(C^\alpha - H^\alpha) - DD(N - H^N)$,[68, 69] $DD(C^\alpha - H^\alpha) - CSA(^{15}N)$,[68] $DD(C^\alpha - H^\alpha) - CSA(^{13}C')$[70, 71] and

CSA($^{13}C'$) – CSA($^{13}C'$).[72] The use of cross-correlated DD(C – H) – CSA($^{31}P$) relaxation has been proposed for determining the angles in the phosphodiester chain of RNA.[73]

The cross-correlation of relaxation parameters has also been used to study the glycoside bond in polysaccharides,[69] to explore the intramolecular dynamics in proteins [76, 79] and to determine the conformations of the sugar rings in RNA [80] and the geometry of metal binding in zinc-containing protein molecules.[81] The research making use of relaxation parameters is only at the beginning but encouraging results have already been obtained.[82, 83] For instance, TROSY spectroscopy has been embodied owing to these studies.

## III. Study of membrane proteins by solid-state NMR spectroscopy

Studies into the structure of biomolecules by solid-state NMR spectroscopy carried out by several research groups are currently in progress.[84–91]

Biomolecules are oriented to the greatest extent in the solid phase; however, due to the absence of rotational or translational degrees of freedom, the NMR spectra exhibit lines corresponding to interactions of magnetic nuclei not only within the molecule but also with the magnetic nuclei of neighbouring molecules. This makes the spectrum more complicated; it contains a large number of overlapping lines. In addition, the protein molecules forming a crystal unit cell are oriented in different ways relative to the crystal axes, and, hence, the RDC for the same pairs of nuclei are different in the magnetic field of the spectrometer.

In order to simplify the NMR spectra, one can use the LC phases of the studied substances. In an ideal case, the spectra should exhibit only intramolecular interactions, while intermolecular interactions should be averaged to zero. However, the mere use of LC phases would not ensure a full success due to the same problem: there are too many magnetic nuclei in a biomolecule. Nevertheless, the fact that intermolecular interactions are not manifested in the spectra is a step forward. By using methods for simplification of NMR spectra with allowance for the fact that the principal axes of molecules are unidirectional, it becomes possible to interpret the results obtained much more accurately than with crystals. This new line of research is currently vigorously developing, especially in investigations of insoluble membrane proteins.

Membrane proteins participate in important processes in cells (intercellular bonds, signal transduction, toxin secretion, transport of ions and nutrients); therefore, they are of interest as potential targets for drugs. It is difficult to determine the spatial structure of a membrane protein by X-ray or liquid-state NMR. These proteins virtually do not crystallise and are almost insoluble; even if a crystal or a solution is obtained, the protein structure may no longer correspond to the native structure. In addition, for understanding of the mechanism of functioning of a membrane protein, it is important to know the positions of elements of the secondary structure of the protein itself relative to the membrane.

Only several dozens of structures for membrane proteins have now been determined by X-ray with a resolution of 0.3 nm, while for soluble proteins, this number is more than 3000 for the same resolution. The structures of several more membrane proteins have been established using electron microscopy of two-dimensional crystals obtained by crystallisation of protein-containing lipid bilayers. However, this does not provide a satisfactory resolution. Thus, NMR spectroscopy appears to be the most promising technique for large-scale determinations of the structures of membrane proteins.

Solid-state NMR is convenient for investigations of membrane proteins in lipid bilayers, because both the tertiary structure of the membrane-embedded protein and the protein orientation relative to the membrane surface can be determined simultaneously.

Membrane proteins are oriented in the same way with respect to the membrane but are arranged chaotically on its surface. Actually, both the membrane and the protein are in the LC state, the degree of alignment of biomolecules being very high. This accounts for the multiple dipolar splittings of lines, which makes the resultant spectrum unresolved, like the spectra of usual organic crystals. However, in the case of membrane proteins, in view of the fact that the proteins are similarly oriented on the membrane, there is a unique opportunity to prepare specimens in which the membranes themselves are also aligned. Such a specimen would be suitable for investigation by solid-state NMR techniques.

In most of these techniques, magic-angle spinning of the oriented specimen is combined with other NMR tools (multidimensional spectroscopy, multi-quantum experiments, *etc.*). In this case, the orientation-dependent information is lost (due to averaging), but it can be restored by analysing, for example, the spinning sidebands.[92–94]

In order to decrease the RDC value for highly aligned molecules and to simplify the NMR spectrum, magic-angle spinning in the spin space is used; this is attained by using special pulse sequences.

Magic-angle spinning (or spinning at a very similar angle) results in virtually complete averaging of relatively small RDC from remote nuclei, leaving only the contributions from non-averaged large RDC, which are usually due to directly linked or very close nuclei. Since the distances between such nuclei are often known with high accuracy (the distances between the nuclei at the sites of contact of different chain sections are an exception), the experimental RDC values can be used to find the orientation of each pair of nuclei relative to the magnetic field. The set of these parameters called orientational constraints determines the conformation of the biomolecule in space. This method developed in a number of studies [95–99] can be successfully used to determine the structure and polymorphism of the membrane proteins.

Up till now, solid-state experiments have been carried out using the RDC of a covalently bonded $^{15}N-^{1}H$ pair which lies in the plane of the peptide bond. Potentially suitable are the RDC for the directly linked $^{13}C-^{1}H$, $^{17}O-^{13}C$, and $^{15}N-^{2}H$ heteronuclear pairs,[96, 100] the $^{13}C-^{13}C$ and $^{15}N-^{15}N$ homonuclear pairs, and the $^{2}H$-quadrupole couplings.[97] The first attempts were undertaken using the $^{1}H-^{1}H$ coupling constants for exchanging protons. They were determined by the homonuclear dipole – dipole decoupling technique in the MAS (Magic Angle Spinning) spectra of highly deuterated proteins and peptides.[101, 102]

The $^{15}N-^{1}H$ dipolar coupling constants are mainly determined in SLF (Separated Local Field) [103] and PISEMA (Polarisation Inversion with Spin Exchange at the Magic Angle) experiments.[104] The procedures of these experiments make it possible to exclude the homonuclear proton – proton constants and to increase sharply the resolution of spectra. These procedures are vigorously developing now.[85, 105, 106] The known REDOR (Rotational Echo Double Resonance) procedure has started to be used in the solid-state investigations of biomolecules.[103, 107, 108]

The PISEMA spectra as well as the usual heterocorrelated $^{1}H/^{15}N$ NMR spectra exhibit readily diagnosable cross-peaks, which form resonance patterns called PISA-Wheel, because they correspond to α-helices or β-sheets that are somewhat tilted in the membrane body. These resonance patterns serve as excellent and clearly visible indicators for the presence of secondary structure elements (α-helix or β-sheet) and their orientation with respect to the membrane.[109–112] For visualisation and reliable identification of this type of spectral images and for numerical simulation of the spectra of proteins, a very convenient SIMMOL program,[113] integrated into the SIMPSON program,[114] used for the analysis of the solid-state protein spectra, has been developed.

Note that usual methods of NMR spectroscopy in the liquid state are also applicable to membrane proteins provided they are placed in, for example, micelles.[115, 116] The NMR spectra of solutions of these encapsulated membrane proteins resemble the

spectra of soluble proteins. For increasing the relaxation times, one can use low-viscosity solvents or reversed micelles.[40] This procedure was used to establish the structure of the transmembrane domain of membrane protein A from *E. coli* embedded in dodecylphosphocholine micelles [117] (the structure proved to be identical to the structure determined by X-ray diffraction analysis in n-octyltetraoxyethylene micelles) and the OmpX membrane protein from *E. coli* in dihexanoyl phosphatidylcholine micelles.[33]

Gramicidin A is the first membrane protein whose structure has been determined by solid-state NMR using isotope labels.[96, 118–120] Subsequently PISEMA experiments were used to sequence the structure of the transmembrane M2 segment of the nicotine acetylcholine receptor,[121, 122] which is a single helix arranged at an angle of 12° to the membrane normal. A membrane pore designed in molecular modelling accommodated five of these helices. It was established in similar experiments that the M2 protein of the influenza virus A forms an ion channel out of four helices located at a 35° angle to the membrane normal,[123, 124] while the Vpu protein forms a similar channel in the HIV-1 membrane at an angle of 15°.[125] Structural and topological studies of some other membrane proteins have been reported.[126–128]

The results of numerical modelling of the high-resolution solid-state NMR spectra of uniformly isotope-enriched membrane proteins, rhodopsin and porin, in oriented lipid layers made it possible to increase the efficiency of the experimental procedures used.[129]

The structures of only two membrane proteins have now been deposited with the PDB, namely, gramicidin A and M2 of the acetylcholine receptor. The intensive development of solid-state NMR spectroscopy provides every reason for hoping that the number of these structures will increase within the next few years.

Encouraging results were also obtained in the development of procedures for determination of the structures of single-crystalline specimens of $^{13}C$- and $^{15}N$-enriched peptides.[88, 130]

The high degree of alignment of molecules in solid specimens brings about specific problems; the procedures used to solve these problems allow one to diminish artificially the effects of the high degree of alignment down to acceptable values which normally correspond to weakly aligned molecules. Weak alignment of molecules can also be attained in the liquid phase.

## IV. Methods for alignment of biomolecules in solutions

The idea of the orientation of molecules for gaining structural information sprang up in the early 1960s. Later, this was implemented in practice with the advent of the method for the orientation of molecules by placing them in an LC medium; this method has been used to determine the structures of small organic molecules.[131] Other ways of orientation appeared non-competitive at that time due to low degrees of alignment achieved and the limitation on the number of magnetic nuclei in the molecule, which had to be not more than 10–12; otherwise, interpretation of the NMR spectrum became impossible due to overlap of lines.

At that time, the possibility of studying biomolecules in this way seemed impracticable. However, the idea to use weakly aligned molecules for structural purposes, paradoxical at first glance, made it possible to simplify the NMR spectrum and to eliminate the limitation on the number of magnetic nuclei in the molecule. In addition, the lines in the NMR spectra of weakly aligned biomolecules, unlike the lines of highly aligned molecules, are not virtually broadened due to faster rerientation.

Weak alignment of biomolecules can be attained by a much greater number of methods than high alignment. The type of alignment depends on the specific interactions of the biomacromolecule with the orienting forces or particles; therefore, the orientation mechanisms in different media can differ in principle; this is important for selecting the optimal experimental conditions, more facile data interpretation, and for automation of experiments and calculations.

The simplest methods used for orientation of biomolecules include exposure to magnetic or electric fields or pressure, stretching or straining of microporous matrices, flow and gradient of velocity in the flow. Lyotropic liquid crystals, polymer gels, zeolites and so on can be used as media for the orientation. Virtually all of these methods have been covered in a review.[132] Orientation by particles is accomplished due to stereospecific interactions of the biomolecules with highly anisotropic particles shaped like discs (bicelles, membrane fragments) or rods (phages, viruses) which are readily aligned under the action of a magnetic field. The molecules of the orienting media can be either oriented (additionally oriented, reoriented) under the action of the magnetic field of a spectrometer or retain their initial orientation.

The media that can be oriented due to steric interactions are preferred because they have no influence on the rotational correlation time and, hence, the line widths in the NMR spectrum remain constant. The orientation on complex formation often affects appreciably the mechanisms of rotational correlation, and, hence, $T_2$. The orienting properties of various media have been described in detail in reviews.[133–135] Currently, intensive search for new methods for the orientation of biomolecules is in progress. Figure 2 sketches different types of LC media used for weak alignment of biomolecules in NMR spectroscopy.

### 1. Alignment induced by an external magnetic field

Molecules with a non-zero anisotropy of magnetic susceptibility can be aligned on exposure to an external magnetic field. The orientation of molecules in solution obeys the Boltzmann law, the degree of alignment being proportional to the magnetic field strength squared.[136] In practice, the degree of alignment of molecules is slight and cannot be observed under usual conditions.

The partial alignment of molecules induced by a magnetic field, resulting in new or additional line splittings in high-resolution NMR spectra, was first observed by Bothner-By and coworkers.[137, 138] They studied the paramagnetic complex bis[(*p*-tolyl)-tris(pyrazolyl)borato]cobalt(II). The orientation under the action of a magnetic field was very weak and took place only for some of the molecules, mainly complexes with paramagnetic metals. With the progress in NMR hardware, the magnetic fields in the spectrometers have become stronger; therefore the orientation effects are observed more and more frequently [139] and they should be taken into account for precise determination of spin–spin coupling constants.

It was shown [42] that biological molecules, for example, DNA duplexes and quadruplexes can also be aligned in the magnetic field of the spectrometer. In particular, the researchers observed rather substantial changes in the $^1J_{CH}$ constants in DNA dodecamers upon an increase in the magnetic field strength. It was proposed to use this effect in structural studies.

The research group headed by Prestegard [43] demonstrated the orientation of molecules of a paramagnetic protein, cyanomethmyoglobine, on exposure to a magnetic field. The measured $^1J_{NH}$ values differed from the known ones. The difference between the constants (−1 to 5 Hz) was attributed to the residual $^{15}N-^1H$ dipolar coupling.

Tjandra *et al.* [140] observed magnetic field-induced orientation even for the diamagnetic $^{15}N$-enriched human ubiquitin protein. The anisotropy of magnetic susceptibility of this protein is an order of magnitude lower than those for cyanomethmyoglobine or DNA dodecamers; correspondingly, the residual constants are also lower ($< 0.3$ Hz). The studies were carried out using three NMR spectrometers with different magnetic fields. The constants $^1J_{NH}$ were measured using two procedures. A constant was found to comprise two components: an orientation-independent component caused by the dynamic frequency shift and orientation-dependent component caused by the $^{15}N-^1H$ RDC. The residual $^{13}C-^1H$ dipolar couplings for the same protein enriched in $^{13}C$ have been measured.[141] The measurements were carried out using

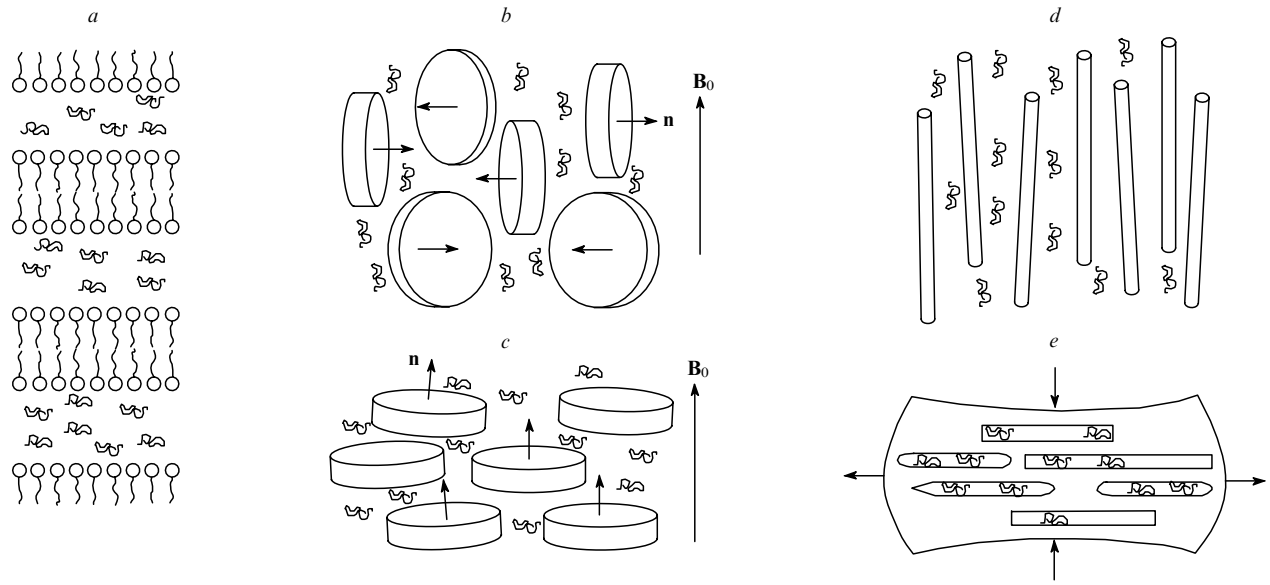



**Figure 2.** Schematic view of the media for weak alignment of biomolecules.
(*a*) Helfrich lamellar phases with bilayers remote from one another; (*b*) aqueous solution of disc-like particles with a negative anisotropy of diamagnetic susceptibility orienting with their normal **n** perpendicular to the magnetic field $\mathbf{B}_0$; (*c*) aqueous solution of the same particles with a positive anisotropy of diamagnetic susceptibility orienting parallel to the magnetic field $\mathbf{B}_0$; (*d*) aqueous solution of rod-like particles (suspensions of filamentous phages and viruses, cellulose crystallites), which align with their long axis parallel to the field; (*e*) stretched or compressed porous polymer gels.

two spectrometers operating at 360 and 750 MHz. However, the residual dipolar couplings were relatively weak (due to the weak degree of alignment of the biomolecule) and inapplicable for structural studies.

The complexes of proteins with DNA or with paramagnetic metals are aligned much better on exposure to a magnetic field.

## 2. Alignment induced by paramagnetic ions

To enhance the alignment of biomolecules in a magnetic field, it is necessary to increase the anisotropy of magnetic susceptibility of these molecules or the orienting system (membranes, media, and so on). This can be done, for example, by adding paramagnetic metal ions (mainly, lanthanide ions). These ions form a paramagnetic complex with a studied molecule, which is aligned more efficiently in the magnetic field of the spectrometer. This was shown for the first time using the paramagnetic complex of myoglobine.[43] The degree of alignment thus attained is not very large and is inferior to that attained in dilute LC media but it is quite sufficient for structural investigations.

A paramagnetic ion can be linked to the studied biomolecule in several ways. Some proteins initially contain paramagnetic ions, for example, the myoglobine mentioned above and heme-containing proteins. In other metal-containing biomolecules, diamagnetic metal ions can be replaced by paramagnetic ions. The proteins devoid of metal can be built up by attaching a protein fragment containing metal atoms to one of its ends. The orientation of the biomolecule would be different depending on whether the paramagnetic label is attached to the C- or N-end. This provides additional structural information and proves fairly useful for removing the degeneracy of orientation data in determining the three-dimensional structure of a biomolecule. By successive replacement of several different ions in a biomolecule, one can attain its multiple orientations.

The position of the $Eu^{3+}$ in the Eu–DNA quadruplex system was determined by using magnetic orientation induced by the paramagnetic ion.[142]

As a result of orientation induced by replacement of the $Ca^{2+}$ ion in calcium-containing biomolecules by lanthanide ions [143–147] [in particular, by $Tb^{3+}$ (see Ref. 144) or by $Ce^{3+}$ (see Ref. 146)], it became possible to observe the dipole splitting and to determine the structures of a number of proteins. The same replacement made it possible to measure the RDC for pairs of nuclei in the $NH_2$ groups and to study the conformational mobility of the side chains in the asparagine and glutamine residues in the molecule of the $Ca_2$-calbindin protein.[145]

A situation where two aligning factors, for example, a paramagnetic ion and an LC medium, compete with each other is favourable for structure determination. In these cases, the resulting RDC is equal to the sum of the RDC for the initial diamagnetic protein molecule induced by the LC phase and that induced by the paramagnetic ion. When the diamagnetic ion in the protein is replaced by different paramagnetic ions, the alignment tensor can be determined unambiguously.[148]

The method in which a standard paramagnetic label is attached to the N- or C-end of the protein molecule is promising. Various small proteins and protein tags have been proposed as labels, for example, the Zn-finger protein in which the diamagnetic $Zn^{2+}$ ion can be replaced by paramagnetic $Co^{2+}$ and $Mn^{2+}$ ions,[149] or the terminal fragment of the EF-hand oncomodulin protein consisting of 12 amino acid residues in which $Co^{2+}$ can be replaced by paramagnetic lanthanide ions.[150] The range of these labels is broad.[151] Donaldson *et al.*[152] proposed the shortest sequence, consisting of only three amino acid residues, binding the $Cu^{2+}$ ion for adding a paramagnetic label to the N-end of a studied protein. The protein molecules containing a paramagnetic label exhibit a substantial anisotropy of paramagnetic susceptibility and are efficiently aligned in a magnetic field, whereas the orientation of a diamagnetic protein is very weak. By recording the NMR spectra before and after the substitution, one can determine the dipolar constants and the dipolar shifts. If in paramagnetic relaxation experiments, the $Mn^{2+}$ ion is replaced by a paramagnetic ion, it is possible to determine the distance between the unpaired electron and the proton of the amide group. When these distances and the dipolar shifts of the amide protons are known, one can calculate the $z$ coordinates for these protons.[149]

In addition to the orientation (or reorientation) effect in the magnetic field and the possibility of measuring RDC, the presence of a paramagnetic ion in a biomolecule allows one to measure other parameters containing structural information also. In particular, these are the paramagnetic shifts of nuclei, the constants of cross-correlated relaxation between the electron and

nuclear spins [153, 154] and between nuclear spins and the Curie spin.[75, 155, 156]

The pseudocontact shifts caused by the dipolar coupling between the magnetic moments of the electron and the nucleus also carry structural information and can be used as additional parameters for structure determination of a protein.[157] The paramagnetic heme-containing protein of the rat cytochrome $b_5$ in the paramagnetic (oxidised) and diamagnetic (reduced) states was studied in this way.[158] The structures determined from the data obtained in NOE-based experiments, by means of RDC, and using information on the pseudocontact shifts are in good agreement.

Using data on dipolar coupling, one can determine the anisotropy of the magnetic susceptibility of the whole protein. In particular, comparison of the dipolar couplings and the anisotropies of the magnetic susceptibility of cytochrome-c′ from *Rhodobacter capsulatus* in the paramagnetic and diamagnetic states found by different methods with one another and with published results [159, 160] showed good agreement of these values. It was also shown that the protein exists in solution in the monomer form.

Taking into account the self-orientation of cytochrome $b_{592}$ molecules from *E. coli* containing heme in the oxidised state, Arnesano *et al.*[161] were able to refine the structure of this protein and to separate the diamagnetic and paramagnetic contributions from the total magnetic anisotropy of the molecule.

## 3. Alignment induced by an electric field

An electric field can be used as an orientation factor; however, this restricts the possible electrochemical and electroconvective processes involving biomolecules in aqueous and salt-containing aqueous solutions.[162, 163]

An electric field was first used to align the *p*-nitrotoluene molecules.[164] This study was undertaken to determine the sign of the spin – spin coupling constant.

The orientation of molecules on exposure to an electric field has an obvious advantage over the magnetic field-induced orientation: there is no need to carry out measurements using at least three NMR spectrometers to disinguish between the orientation effects and those caused by dynamic frequency shifts; it is sufficient to employ only one spectrometer but three different electric field strengths.

It has been proposed [165, 166] to use an alternating pulse electric field synchronised with a sequence of radiofrequency pulses. In this case, multidimensional NMR experiments can be carried out with the electric field being switched on and off at particular time instants. The researchers cited developed two 2D experimental procedures for determination of the RDC. In one experiment, the electric field was applied only in the second dimension to observe the dipolar splitting, while the isotropic chemical shifts were observed in the direct dimension. In the other experiment, the electric field was switched on only during mixing; as a consequence, cross-peaks were observed only for the pairs connected by dipolar coupling. In these experiments, alternating electric fields with a frequency of 28.5 kHz and a strength of $40-70$ kV cm$^{-1}$ were used.

A similar 2D $^2$H NMR experiment for observing the correlation between the isotropic chemical shift and the residual anisotropic couplings (in this case, quadrupole coupling induced by a pulse-modulated constant electric field) has been proposed.[167] For decreasing the influence of the molecular motion on the measured signal, multiquantum filtration is used.[168, 169] This technique is equally suitable for dipole-coupled systems with a spin of 1/2.

In these experiments, nitrobenzene thoroughly purified from ionic impurities was used as an orienting compound. The molecules of some proteins have higher dipole moments than nitrobenzene; however, it is still unknown to what extent the method involving orientation in an electric field is applicable to biomolecules. These studies are only at the beginning.

## 4. Orientation in liquid-crystalline media

The self-orientation of biomolecules (except for the molecules containing paramagnetic centres) even on exposure to strong magnetic and electric fields is very low. A higher degree of alignment can be attained by placing a biomolecule in the LC medium, which is oriented in a magnetic field and thus aligns the biomolecules it contains. This medium should meet a number of requirements: it should be miscible with water, not interact chemically with the studied molecule, accommodate high concentrations of the biomolecules under study and the required buffer components and also to be stable for a long time over a broad temperature range. However, every LC possesses a high degree of alignment; therefore, the main difficulty in searching for these systems is the requirement of weak alignment of the biomolecules, *i.e.*, several orders of magnitude weaker than the degree of alignment of the LC molecules.

Systematic search resulted in a number of LC systems suitable for structural studies of biomolecules. For weak alignment of biomolecules, first, traditional lyotropic LC systems were tested, for example, solutions in quarternary ammonium, pyridinium,[170, 171] and other salts in aqueous alcohol. Mere dilution by water allows one to decrease somewhat the degree of alignment in these systems but only to certain critical concentrations below which the system is no longer liquid-crystalline and, hence, aligned. Liquid-crystalline media based on cetylpyridinium chloride – hexanol [170] and cetylpyridinium bromide – hexanol proved to be the systems of choice.[171] These systems form lamellar phases of the Helfrich type consisting of double layers, which can be moved apart by adding a solvent, so that the distance between them will markedly exceed the thickness of the bilayer itself (see Fig. 2 *a*). These lyotropic phases are very convenient and simple systems, relatively stable over a broad range of temperatures and concentrations. They withstand the concentrations of salts optimal for the existence of biomolecules and are stable for a long period of time. The magnetic anisotropy of the pyridinium rings determines the orientation of lamellae in a magnetic field: the symmetry axes of the lamellae are arranged along the field direction.

A new non-ionic lyotropic LC medium based on alkyl poly(ethylene glycols) and alkanols, which has certain advantages, has been proposed.[172] In particular, the solution can be prepared outside the magnetic field of the spectrometer. The specimen has a low viscosity and can be transferred into the ampoule just by a pipette. In a magnetic field, the system is rapidly oriented. The components of this system are commercially available and inexpensive compared to the components of other systems used to study biomolecules by NMR spectroscopy techniques. They can be used with equal efficiency to study both proteins and nucleic acids. Advantages of these systems include low sensitivity to the pH and to high concentrations of salts and the protein.

## 5. Orientation by disc-like particles

### a. Bicelles

Disc-like particles (bicelles, membrane fragments, *etc.*) can also be oriented in a magnetic field, form ordered LC media and align biomolecules that occur together with them in a solution.

Ram and Prestegard [173] were the first to propose disc-like mixed micelles as a model membrane system for NMR investigations. Bicelles, which are formed in aqueous solutions of lipids and surfactants and represent discrete bilayer disc-shaped structures, proved to be most convenient for this purpose.[174] They are equally applicable to studies of membrane proteins, which are incorporated and fixed in bicelles, by solid-state or liquid-state NMR and to NMR studies of water-soluble proteins in the presence of bicelles which are aligned in a magnetic field and align the protein molecules due to specific interactions with them.[175 – 179]

The researchers' attention to bicelles as a medium for the study of biomolecules was enhanced when it was found that enzyme activity is retained in this medium.[167] Tjandra and Bax [45, 177] proposed using a dilute solution of bicelles as an orienting solvent

for biomolecules. Even the first results demonstrated that the distances and angles determined using $^{1}H-^{15}N$, $^{1}H-^{13}C$ and $^{13}C-^{13}C$ RDC in the molecules of human ubiquitin in dilute bicellar LC media were in agreement with the corresponding parameters determined in their crystal structure.[178]

Bicelles (especially, those containing phospholipids) are much more similar to native membranes than small micelles. Therefore, the structure determined for membrane proteins encapsulated into bicelles is more plausible.[180] (However, data on micelle-induced bending of biomolecules have been reported.[181]) If the bicelle discs are small, they are freely and rapidly reoriented in solution and are not oriented under the action of a magnetic field. Such bicelles can form an ideal system for the investigation of encapsulated membrane proteins by conventional high-resolution NMR techniques used to study water-soluble proteins.[182] In addition, a bicelle is oriented much more easily by a magnetic field than a spherical micelle with an embedded protein. This allows one to use procedures based on residual dipolar couplings in the studies of membrane proteins encapsulated in bicelles.

An anisotropic bicelle with a high ratio of the radius to the thickness can be oriented on exposure to an external magnetic field. By varying the composition and component ratio and their concentration in an aqueous solution, one can change gradually the degree of bicelle orientation and, hence, the degree of alignment of the studied biomolecules (which occur in the aqueous phase or inside the bicelles).

Studies of membrane proteins and peptides encapsulated in bicelles are usually carried out with high concentrations of phospholipids. This method was used to determine the orientation of the N-terminal fragment of the ARF1 protein relative to the surface of a bicelle oriented in a magnetic field [183] and to compare the orientation of the same fragment relative to micelle or bicelle surfaces.[184]

Below we consider the use of bicelles for the studies of biomolecules only in solutions. By varying the component concentrations, one can change the degree of alignment of the dissolved biomolecules in these systems over broad ranges in order to attain the desired values of residual dipolar splitting. Various additives can be used to improve the properties of the system.

Bicelles consist of a mixture of at least two phospholipids with short and long hydrophobic tails; long-chain phospholipids form the flat surface of the disc, while the short-chain ones form the rounded edges. A typical thickness of a lipid bilayer is 40 – 50 Å. The dimensions of the bicelles increase with an increase in the molar ratio of the long-chain to short-chain lipids.[185] An important feature of such systems is the ability to pass from the nematic LC state into an isotropic state on changing the temperature; as this takes place, one can record the spectra of the dissolved molecules in both states of the system, which is necessary for determining the dipole contributions to the observed splitting. In particular, the transition of the basic DMPC – DHPC system (DMPC is dimyristoylphosphatidylcholine, DHPC is dihexanoylphosphatidylcholine) from the nematic to isotropic state takes place at 35 °C. This system is used most widely. It forms an alignable LC phase at very low bicelle concentrations ($\approx 3\% - 5\%$). The distances between separate bicelles in solution reach 300 Å; this is quite sufficient for biomolecules to fit between them. The disc radius is directly proportional to the [DMPC]/[DHPC] ratio and is inversely proportional to the total concentration of phospholipids in water with their ratio remaining constant.[186]

The procedures for sample preparation for NMR studies and the properties of the basic DMPC – DHPC bicellar system have been described in detail in the literature. The dependences of the stability and the orienting properties of this medium on the temperature, pH, salt composition, the relative concentrations of components and additives have been studied.[187 – 189]

The properties of bicelles (size, shape, surface charge, the anisotropy of diamagnetic susceptibility or the anisotropy of dielectric constant) can be varied over broad ranges. Usual bicelles are stable only in narrow ranges of temperatures and pH. Charged lipids [190] or aromatic surfactants [191] can be added to bicellar systems in order to extend the temperature range, increase the stability, enhance the orienting ability and change the type of orientation. The additives counterbalance the destructive influence of high salt concentrations on the system and minimise the interactions between the macromolecule under study and the bicelle. This effect shows itself with both positively and negatively charged additives. The addition of cationic lipids imparts to the system a low positive charge, which has an influence on the alignment of proteins caused by electrostatic and steric forces.[192]

A large number of compositions able to form aqueous bicellar solutions under certain conditions have now been found.[179] The most promising among them are DLPC – CHAPSO [193] [DLPC is dilaurylphosphatidylcholine, CHAPSO is 3-(chloroamidopropyl)dimethylammonio-2-hydroxy-1-propane sulfonate] for studies over a broad temperature range and DIODPC – CHAPSO [194] (DIODPC is 1,2-di-*O*-dodecyl-*sn*-glycero-3-phosphocholine) for studies of macromolecules in highly acidic media. Struppe *et al.*[195] have modified the DMPC – DHPC system by adding 25 mass % of dimyristoylphosphatidylserine (DMPS) or dimyristoylphosphatidylglycerol (DMPG). These 'acidic' bicelles are stable at pHs from 5.5 to 7 and are suitable for use in NMR studies of membrane proteins, because natural membranes are often charged, as they contain negatively charged lipids.

Many bicellar systems are susceptible for hydrolysis, which hampers long-term experiments (taking many days). Ottiger and Bax [196] proposed replacing diacyl phospholipids by dialkyl analogues in order to make the bicelles stable against hydrolysis. New bicelle systems are similar in orienting properties to DMPC – DHPC type systems but they are suitable for operation in both acidic and alkaline media. The DIODPC – CHAPSO system is also stable against hydrolysis.[194]

The mechanism of orientation of biomolecules by bicelles is complicated and includes, apart from steric repulsion, electrostatic interactions and adsorption on the bicelle surface. The steric interactions are usually rather weak and can be easily modelled.[197]

Bicelles can be oriented in different ways: by attaching an additional peptide tail to the C-end of the protein molecule, for example, so-called peptide Hi-tag; by changing pH; by adding cationic lipids, which would make the system slightly positively charged; by adding particles with highly anisotropic diamagnetic susceptibility or paramagnetic ions.[192] The orientation of the director of the bicellar LC phase can also be varied by rotating the sample.[198]

Usually, the degree of alignment of bicelles and transition of a bicellar LC system into the isotropic state are detected by measuring the quadrupole splittings in the spectra of heavy water, which is the major component of the system. Another method for measuring the degree of alignment of phospholipid bicelles in a magnetic field based on $^{31}P$ NMR spectroscopy has been proposed.[199]

The morphology of three lyotropic LC phases — bicellar phase, a mixture of pentaethylene glycol monododecyl ether with hexanol and the cetylpyridinium bromide – hexanol system — was studied in detail in translational diffusion anisotropy experiments.[200] Bicelles remain stable in the studies under elevated pressures. The interest in these studies is increasing at present.[201]

Thus, lyotropic bicellar systems are very convenient media for investigating aligned proteins. However, there exist a number of proteins that destroy bicelles when interacting with them; therefore, study of the structures of these proteins using bicelles is problematic.[178, 202]

**b. Orientation of biomolecules in bicelles doped with lanthanide ions**

Bicelles are usually oriented in a magnetic field in such a way that their normal is perpendicular to the field direction,[203] while the planes of the discs are located chaotically. This orientation can be changed by adding paramagnetic ions or by incorporating into the

bicelles an aromatic compound or macromolecules with a great anisotropy of diamagnetic susceptibility.

The orienting properties of bilayer membranes doped with lanthanide ions [204–207] are rather attractive not only for investigating membrane proteins but also for investigating proteins encapsulated in them in solutions.[150, 208, 209] This system combines two mechanisms for the ordering of biomolecules, namely, the influence of the membrane and the effect of paramagnetic ions. In the presence of $Er^{3+}$, $Eu^{3+}$, $Tm^{3+}$, or $Yb^{3+}$ ions, the negative diamagnetic anisotropy of bicelles increases and their symmetry axis becomes parallel to the magnetic field. The biomolecules in the bicelles are also reoriented.[204, 205] The transition accompanied by the change in the orientation of the bicellar system with respect to the magnetic field from the perpendicular to the parallel upon the addition of paramagnetic ions, as well as the formation of the LC phase upon increasing the bicelle concentration are cooperative phenomena depending on the interactions between the bicelles and on the temperature. In some cases, two types of alignment have been observed, the equilibrium between them being dependent on the concentrations of the paramagnetic ions and bicelles.[208]

The systematic research of mixtures of phospholipids with long and short chains over a broad range of concentrations and temperatures by small-angle neutron scattering showed the presence of phases with different morphologies such as lamellae, multilamellae, monolamellar vesicles and bicellar discs.[210]

However, the large concentration of paramagnetic ions in the system results in undesirable line broadening and line shifts in the NMR spectra, which hampers their analysis. When chelating phospholipids are added in an equimolar amount with respect to paramagnetic ions,[206] the ions form chelates incorporated in the membrane, and the paramagnetic effects are levelled. In addition, due to the alignment of the lanthanide chelates in the membranes, the total anisotropy of the diamagnetic susceptibility increases to such an extent that the lanthanide ion concentration needed for the reorientation of bicelles can be decreased. The possibility of inducing substantial changes in the type of bicelle orientation and, hence, the orientation of biomolecules by adding paramagnetic ions is also attractive due to the fact that a second set of RDC can be obtained rather easily. Owing to the new set, the number of RDC is actually doubled; on the one hand, this markedly increases the accuracy of the determined structure and, on the other hand, eliminates the ambiguity of the order parameters (see below).

If phospholipids containing, for example, a biphenyl group [203] in one of the long chains, are added instead of nonlipid amphiphilic aromatic compounds or paramagnetic ions, it is possible to create bicellar systems with positive anisotropy of magnetic susceptibility ($\Delta\chi$).[211, 212]

Study of a number of hydrophobic membrane proteins in solution became possible due to the effect of binding of lanthanide ions to protein-containing micelles.[213] The use of bicelles, membrane fragments and viruses for this purpose is problematic because membrane proteins change appreciably the structures of these particles when react with their lipids. This influences the quality of NMR spectra.

#### c. Purple membrane fragments

A method for aligning water-soluble biomolecules in the presence of purple membrane fragments (PM) was proposed almost simultaneously in two publications.[214, 215] Purple membrane fragments are prepared from halobacteria, which form wide disc-like particles consisting of bacteriorhodopsin and some residual lipids when occuring in an aqueous solution with a concentration of $1-10$ mg ml$^{-1}$. These particles are stable on sharp changes in the temperature, pH and ionic strength of the solution. They have great negative charges and, therefore, they can induce orientation of biomolecules (due to electrostatic interaction), which can be controlled by changing the concentrations of the PM fragments or by adding cations for reducing the electrostatic interaction between the macromolecule and the membrane.[215] These systems are highly stable against temperature or pH variations and represent an excellent alternative to bicelles.

The orientation of biomolecules induced by PM fragments was demonstrated using two proteins, namely, human ubiquitin and p53. The maximum values of the $^1H-^{15}N$ RDC were $\sim 20$ Hz. It is worth mentioning that in this case, too, the axial and rhombic components and the direction of the principal axis of the alignment tensor differ appreciably from the corresponding values in bicelles and other orienting media, which makes the membrane fragments a convenient medium for determining the additional set of RDC. The presence of several sets of RDC in which the principal axes of the corresponding diagonalised tensors are arranged at substantially different angles allows one to construct the complete 'triangulation' system of biological macromolecules in solution (see below).

A solution of PM fragments is not a liquid-crystalline system. Typical PM fragments are several micrometers in diameter and are 45 Å thick. Their alignment in a magnetic field is due to seven $\alpha$-helices of bacteriorhodopsin. In a magnetic field, PM fragments are oriented in such a way that the normal to the membrane is parallel to the magnetic field direction and, due to high concentration of bacteriorhodopsin, the degree of alignment is extremely high.[214]

The concentration of liquid crystals in solution cannot be lower than the value where the LC phase is formed; conversely, the concentration of the aligning particles in a suspension can be arbitrarily low. The alignment can be monitored based on quadrupole splitting of deuterium in the spectrum of heavy water.

The use of lanthanide ions, as in the case of bicelles, allows one to change the direction of the predominant orientation of biomolecules in the presence of PM fragments.[216] In a study dealing with the structure of a membrane protein,[217] an oxygen molecule has been used as a paramagnetic probe.

## 6. Orientation by rod-like particles

It should be borne in mind that bicelles are stable only in narrow temperature and pH ranges (beyond this range, hydrolysis of the ester bond can take place) and that many proteins interact with bicelle membranes to destroy them. Therefore, it is more preferable to use solutions of more stable particles that can be oriented and to orient the dissolved biomolecules without destroying their structure. This role can be played by any rigid, charged, rod-like macromolecular ensembles consisting, for example, of rod-shaped viruses and bacteriophages of the filamentous phage family, inorganic filamentary crystallites, polymeric crystallites and other relatively rigid filamentary polymeric structures (for example, polytetrafluoroethylene whiskers).

#### a. Liquid-crystalline colloid suspensions of rod-shaped viruses

Widely known suspensions of bacteriophage fd and the tobacco mosaic virus,[202] bacteriophage Pf1 or M13 [218] were among the first systems to be proposed as alternatives to bicelles. These systems are very convenient for attaining medium degrees of alignment of the dissolved biomolecules; moreover, it is possible to control the orientation parameters by adding the bacteriophage Pf1 subject to magnetic alignment.[218, 219] This procedure was used to determine the $^1H-^1H$, $^1H-^{13}C$ and $^1H-^{15}N$ RDC of various proteins and oligosaccharides. The orientation of the molecular alignment tensor in these systems differs substantially from the orientation in the medium of usual bicelles; this allows the measurements to be taken simultaneously in both media.

The procedure of preparation of the bacteriophage Pf1 specimen for the partial alignment of biomacromolecules has been described in detail elsewhere.[220–222] The bacteriophage Pf1 is a DNA phage composed of 7349 nucleotides and packed in a protein shell with a nucleotide : protein ratio of 1 : 1. This is a rod-like system with a length of 20 000 Å and 60 Å in diameter, which is spontaneously oriented in a magnetic field; Pf1 has a negative surface charge. The orienting properties of the phage under various experimental conditions were considered in detail

using DNA and RNA as examples. Based on quadrupole splitting of deuterium, one can judge the degree of system alignment, which can be changed by varying the phage concentration and the temperature.[220] The researchers recommended a medium pH of 6.5 – 8.0 and [NaCl] < 100 mmol litre$^{-1}$.

Most biomolecules and orienting particles are charged and this should be taken into account when selecting the appropriate medium. Bacteriophages represent ideal media for the orientation of RNA and DNA molecules: negatively charged molecules of nucleic acids are aligned due to steric interactions with negatively charged phages Pf1 , which produce a stable orientation over a broad range of conditions in a solution.

The use of negatively charged particles to align positively charged biomolecules is problematic, although in some cases, such media prove very convenient for investigations of proteins.[220] Even when the bacteriophage concentration is abnormally low, these systems are effective.[223] Presumably, in this case, a different orientation mechanism is involved associated with weak binding of the protein under study to the phage. The observed orientation is due to fast exchange between the bound and free states of the protein.

In the cholesteric phase of the bacteriophage fd, the degree of alignment of the dissolved protein and, what is especially important, the rhombicity, depend on the pH.[224] The pH value can be varied over a broad range, from 2 to 8 (an erroneous opinion that working with phages is possible only at high pH values has been established).

The alignment of a phage-based LC system can be fixed and made independent of an external magnetic field. For this purpose, the system is placed in a polyacrylamide matrix. The polymer-stabilised system fully retains its specific orientation characteristics.[225]

#### b. Cellulose crystallites

Suspensions of cellulose crystallites are, probably, the cheapest, the most stable and the simplest media for the orientation of biomolecules.[226] They are produced from natural cellulose fibres (for example, from filter paper) by careful hydrolysis by sulfuric acid, concentration and extensive dialysis. After some period, the system separates into an upper isotropic phase and a lower anisotropic phase. The latter is isolated, concentrated by evaporation and then used.

The cellulose crystallites are rods with dimensions $\sim 200 \times 10$ nm. A separate crystallite has a large negative diamagnetic anisotropy; therefore it is oriented in a magnetic field with its long axis normal to the direction of the applied field. The orientation process is slow; to attain complete orientation, a freshly prepared specimen should be kept for several hours in a magnetic field. The surface negative charge of the crystallite caused by sulfate groups is smaller than the surface charge of bacteriophages; therefore, they orient biomolecules by a steric mechanism rather than an electrostatic one.

## 7. Stressed polymer gels

Polymer gels form promising media for the orientation of biomolecules. The effect of orientation of small molecules incorporated in rubber-like polymers has been known for a long time.[227] Two groups of researchers working independently of each other proposed using this phenomenon for the orientation of biomolecules in polyacrylamide gels. The orientation is attained either upon mechanical treatment (stretching or straining)[228, 229] or by means of aligned PM fragments or other particles embedded in the gel.[229]

The gel is prepared by polymerisation of an aqueous solution of a mixture of acrylamide and methylenebisacrylamide in a special tube whose diameter differs slightly from the diameter of an NMR ampoule. The polymerisation is initiated by a mixture of ammonium persulfate with tetramehylethylenediamine. After polymerisation, the gel is thoroughly washed, dried and placed in an NMR ampoule, and an aqueous solution of the studied protein is added. The protein and water slowly diffuse into the gel, which swells. Subsequently the gel is subjected to mechanical treatment and, as a result, the macromolecules having penetrated into the gel are aligned. If the diameter of the tube used initially to prepare the gel is smaller than the diameter of the ampoule used in NMR experiments, the gel is compressed, while in the opposite case, the gel is stretched, and the axis of the alignment tensor of the biomolecules incorporated in the gel is oriented across or along the ampoule, respectively.

The pore sizes and the diffusion properties of the polyacrylamide gels can be changed by controlling the concentration of acrylamide and methylenebisacrylamide. Instead of the simple protein – water system, one can use protein – surfactant – water micellar systems.

Yet another option is to incorporate other particles into the gel during polymerisation, for example, PM fragments, phages and viruses able to orient protein molecules. Polymerisation is carried out in a magnetic field and, hence, the above particles are incorporated in the gel in the oriented state. The orientation of molecules of the protein introduced into the gel is due to the interaction with the oriented (orienting) particles.

The orientation of particles immobilised in the gel during polymerisation can be attained using other types of external macroscopic treatment, for example, electric or mechanical treatment.

A distinctive feature of polymer gels is that the direction of their orientation is determined by the preparation procedure and does not depend on a magnetic field. This allows one to change smoothly the gel orientation with respect to the direction of the magnetic field. This feature can also be used to produce a new, biaxial orientation of the system after completion of NMR measurements. In this axially asymmetrical system, one can determine the set of orientation-dependent parameters, which cannot be found in usual uniaxial LC media.[228] The orientation in these systems is not violated when the specimen is rotated. This allows one to measure the spin – spin coupling constants during magic angle spinning of the specimen (to set the RDC to zero).[225]

By varying the pressure, polyacrylamide concentration or cross-linking density in the gel, one can change over wide limits the degree of alignment of biomolecules, whereas the direction and asymmetry of the alignment tensor remain virtually constant.[230] Relatively large protein molecules can also be studied in gels because their relaxation times in the gel are rather long.

A simple funnel-like apparatus has been proposed for uniform squeezing of the polyacrylamide hydrogel.[231] The degree of alignment depends on the ratio of the diameters of a cylindrical hydrogel specimen before and after squeezing.

## 8. Mineral liquid crystals

The media for the orientation of biomolecules considered above are organic materials; therefore, the solvent signals predominate in the NMR spectra, because the number of molecules of the medium is much greater than the number of the biomolecules under study. This creates certain difficulties, especially in the studies of oligosaccharides or oligonucleotides for which the synthesis of isotopically enriched analogues is very complex and expensive.

Mineral liquid crystals devoid of $^{1}H$ or $^{13}C$ nuclei are very convenient in this respect. Systematic research into these crystals started not long ago. Thy are colloid systems consisting of highly anisotropic ribbon-, plate- and disc-shaped or column nanoparticles, which tend to form various types of LC phases. Systems oriented by a magnetic field are of special interest for NMR spectroscopy.

One system of this type is an aqueous suspension of vanadium pentoxide $V_2O_5$, in which a nematic phase is formed within a certain range of concentrations and pH.[232] The particles of this phase are long (several hundreds of nanometers) 1 nm thick ribbons with a width of 25 nm. Due to the high negative charge of the particles, the mechanism of orientation of biomolecules is

similar to that effective in suspensions of viruses and phages. Yet another class of mineral liquid crystals well oriented in a magnetic field has also been proposed, namely, a series of swelling gels of salts of the solid acids generally described as $H_nM_nZ_2O_{3n+5}$ (M = Sb, Nb, Ta; Z = P, As; $n$ = 1, 3), whose particles are plate-like.[233]

The $^{13}C-^{1}H$ dipolar coupling constants of pentasaccharide Lewis X measured in these systems with a natural abundance of the $^{13}C$ isotope proved to be comparable in the order of magnitude with the RDC measured in organic LC, in particular, in a bicellar medium.

Advantages of inorganic systems over organic systems include high stability and inertness over a broad temperature range, the possibility of recovery of the substance under study by flocculation and the fact that at lower concentrations, they are oriented in the magnetic field much better and faster.

Unfortunately, none of the systems considered above is universal for structural research.[172] The DMPC–DHPC bicellar system is susceptible to irreversible degradation due to hydrolysis at low pH values.[188] The DIODPC–CHAPSO system is more stable at low pH values but does not form an LC phase in a neutral medium.[194] The cetylpyridinium chloride–hexanol system is destabilised at high protein concentrations and, to increase its stability, high salt concentrations are required.[170] Conversely, the cetylpyridiunium bromide–hexanol system is stable at low ionic strength of the solution.[171] Phages are effective in neutral media, but at pH < 6, they tend to aggregate.[218] Purple membrane fragments and bicelles often bind proteins too strongly, which is manifested as line broadening in the NMR spectra and in perturbation of the alignment tensor; as a consequence, it is impossible to determine the real structure of the molecule in solution.[214, 215]

## V. Measurement of orientation-dependent parameters

### 1. Orientation-dependent parameters

Today, a number of NMR parameters containing diverse structural information concerning the molecules under study can be measured in experiments. Table 1 presents some of these NMR parameters. The most valuable and diverse structural information can be derived from orientation-dependent parameters such as RDC, the quadrupole coupling constants (QCC), and the chemical shift anisotropy.

The dipolar coupling constants proved to be more useful at any stage of structure determination for a biomolecule than the traditional parameters obtained mainly in NOE-based experiments. Comparison of the distances found in these experiments with those found using RDC showed that the use of the latter is preferred. The distances longer than 5–6 Å are inaccessible for NOE procedures, whereas the use of RDC allows determination of even longer distances.

**Table 1**. NMR parameters carrying structural information.

| NMR parameter | Structural information |
|---|---|
| Chemical shifts | signal assignment, secondary structure, |
| $^{1}H$, $^{13}C$, $^{15}N$, $^{31}P$ | torsion angles |
| Constants $^{3}J$ | torsion angles |
| NOE signals | proton–proton distances (< 5 Å) |
| Relaxation parameters | torsion angles, bond orientation, dynamics |
| RDC | bond orientation, internuclear distances, dynamics |
| CSA | bond or functional group orientation |
| QCC | distances, angles, bond orientation, dynamics |

The fundamental difference between these methods is as follows: NOE signals permit one to estimate the length of the vector connecting the interacting nuclei, while RDC, in addition, provides information on the orientation of this vector with respect to the direction of the magnetic field; this allows one to determine the relative spatial orientation of the remote fragments of the molecule.

Another important difference is that the measured parameters depend in different ways on the distance. Dipolar couplings depend on the distance as $1/r^3$ and the NOE signals, as $1/r^6$. With a smoother dependence, one can distinguish sections corresponding to interactions between more remote nuclei.

Even when the number of NOE restraints is great, the use of RDC markedly increases the accuracy of structure determination. In those cases where NOE signals provide little information, as in the studies of nucleic acids or large, partially deuterated (to simplify the spectrum) proteins, RDC can play a crucial role.[234]

In order to determine orientation-dependent parameters, one must measure NMR spectra in both isotropic and oriented states. Two different patterns of these measurements are possible, each of them having advantages and shortcomings:

— separate measurements (isotropic and anisotropic parameters are determined in different experiments corresponding to isotropic and oriented states);

— simultaneous measurements (isotropic and anisotropic parameters are determined in one correlation experiment).

Up to the present day, separate measurements for the isotropic and oriented states have been carried out most often. In these experiments, it is possible to select the optimal conditions for recording the spectra for each of the states, in particular, the solvent, the concentration, the temperature, *etc*. However, the structure of a biomolecule may change depending on the experimental conditions. In addition, this type of measurement often requires large amounts of substances because a non-destructive transfer of biomolecules from one medium to another is very complicated. Yet another problem is the difficulty of signal assignment due to the dependence of chemical shifts and their anisotropies on the experimental conditions.

The techniques based on synchronous measurements of isotropic and anisotropic parameters are devoid of the above drawbacks. The measurements are carried out in multidimensional correlation experiments of various types, in one stage of which (evolution/recording), the orientation-dependent parameters are not averaged to zero.

This can be implemented in, for example, a COSY experiment in two ways. If the initial sample has anisotropic interactions, they can be temporarily removed at one stage of the correlation experiment by magic angle spinning (in the real or spin space). If the initial sample is isotropic, anisotropic interactions can be induced by some orienting factor.

This idea was proposed rather long ago; however, only experiments using magic angle spinning have been developed. Currently, experiments of a different type are quite topical, namely, those associated with fast controlled orientation/disorientation of the molecule. One of these experiments which uses orientation in an electric field [165–167] was mentioned above.

It is noteworthy that NMR studies of oriented biomolecules were often preceded by traditional experiments in isotropic media. In many cases, this is a necessary phase because well-developed NMR procedures making use of orienting media are often missing. This is especially true for identification of the resonance signals of a primary sequence. However, another existing tendency is to make NMR spectroscopy of oriented molecules all-sufficient, *i.e*., to develop a method which would allow determination of the structure of a biomolecule relying only on the measurements in an aligned phase.[46] However, we rather believe that both types of experiments, those in oriented and isotropic states, should supplement each other.

Below we discuss the experiments in which NMR spectra of weakly aligned molecules have been recorded; attention is focused

on methods specific to these molecules. In terms of the value to be measured, the experiments can be classified into those determining RDC, CSA and QCC.

## 2. Determination of the residual dipolar coupling constants

The possibility of measuring the RDC for biomolecules accounts for the interest in the NMR spectroscopy of weakly aligned systems.[47, 235]

The weak orientational contributions to the measured spin–spin coupling constants can be distinguished most easily for directly linked nuclei, which are located at the shortest distances and, hence, have the largest dipolar contribution. The pairs of nuclei encountered in biomolecules are H–C, H–N, C–C and C–N. However, due to low natural abundance of the required magnetic isotopes, one must use isotope-enriched specimens.

Almost any type of procedure for precise determination of spin–spin coupling constants developed to date is suitable for RDC determination. First of all, these are experiments based on E COSY (Exclusive COSY) or *J*-correlations and their modifications.

As a rule, RDC are determined in 2D or pseudo-2D experiments from the difference between the splittings of the scalar constant *J* in isotropic and aligned phases. In the case of large protein molecules where analysis of the results of 2D experiments is difficult (especially, if the protein contains predominantly α-helices), 3D NMR experiments are carried out.

In principle, dipolar constants can be measured between any pair of magnetic nuclei with a significant interaction but in practice, only several pairs are taken.

The first RDC between directly linked nuclei to be used were the $^{1}H-^{13}C$ and $^{1}H-^{15}N$ constants.[45, 141] Currently, the RDC between these nuclei constitute the major array of values measured in the NMR spectroscopy of weakly aligned biomolecules. Their interpretation in structural studies is simple because the distances between directly linked nuclei in biomolecules are well known and are virtually constant. This allows unambiguous determination of the orientation of the bond vector relative to the direction of the magnetic field. In $^{15}N/^{13}C$-enriched proteins, one can determine the bond vector for any pair of interacting nuclei (except for oxygen). This information, which can be supplemented by the data from conventional isotropic experiments, is often sufficient for a considerable increase in the accuracy of structure calculations.

Although each of these RDC determines the orientation of the bond vector for one pair of nuclei, the total structure of the biomolecule still remains uncertain, no matter how many constants are known, because they do not contain translational information.

The next step was to use the $^{1}H-^{1}H$ RDC.[218, 219] Unlike the RDC for heteronuclear pairs, $^{1}H-^{13}C$ and $^{1}H-^{15}N$, which determine the orientation of vectors connecting these nuclei, the constants for the $^{1}H-^{1}H$ pair simultaneously bear information on the distances between the nuclei (up to 7.4 Å).[219]

In the presence of scalar and residual dipolar couplings, the efficiency of transfer in different types of homonuclear mixing (isotropic or dipolar) depends appreciably on the relative magnitudes of the constants.[236–238] In the general case, efficient transfer for all possible relative values of the indirect and dipolar constants requires two different mixing sequences (Hartmann–Hahn and COSY).[239]

Today, in a typical study of a biomolecule by NMR spectroscopy, numerous experiments of different types are carried out to determine the RDC for both directly linked and remote nuclei. The procedures for RDC measurement are being further developed by adding new nuclei, for example, $^{19}F$ (see Ref. 240). The TROSY technique developed recently for increasing the resolution and sensitivity was immediately used in experiments for the determination of RDC,[241] in particular, in HSQC (Heteronuclear Single Quantum Coherence),[242] HNCO (3D $H^N$, N, C′ Heteronuclear Correlation Experiment),[243, 244] HN(CO)CA (3D $H^N$, C′, $C^\alpha$ Heteronuclear Correlation Experiment)[245] and other experiments.

Procedures for determining small scalar couplings and dipolar constants in purine or pyrimidine bases in nucleic acids[246, 247] and for detection of internucleotide dipolar couplings[248] have been proposed.

### a. Dipolar coupling constants between the nuclei separated by one bond

Numerous effective procedures for the measurement of RDC between nuclei separated by one bond have now been developed and used successfully for model proteins, most often, ubiquitin. Specific features of the application of these procedures to particular pairs of nuclei in biomolecules have been considered in detail in a number of publications. In proteins, these are RDC for $N-H^N$ (see studies[140, 241–243, 249–257]), $C^\alpha-H^\alpha$ (see studies[177, 251, 253, 258–261]), $C^\alpha-C'$ (see studies[241–245, 253, 259, 261, 262]), $C^\alpha-N$ (see studies[241, 244, 252, 253, 255, 263]) and $C'-N$ (see studies[241, 243, 244, 257]) pairs in the backbone and the $C^\beta-H^\beta$ (see studies[259, 260, 264]), $C^\alpha-C^\beta$ (see studies[244, 245, 259]), $C^{Me}-C$ and $C^{Me}-H$ pairs in the side chains.[265, 266] Many procedures are applicable to the determination of various types of constants and some allow one to measure simultaneously a number of constants.[252, 253] Special procedures for measuring the constants characterising particular groups, for example the side-chain $NH_2$ groups in asparagine and glutamine are being developed.[267]

The direct RDC in proteins are interpreted as values determining the direction of the vector of the corresponding bond, the bond length being considered constant. In particular, the C–H length is 1.095 Å, N–H is 1.02 Å, $C^\alpha-C'$ is 1.525 Å, C–C is 1.517 Å, $N-C^\alpha$ is 1.329 Å and N–C′ is 1.45 Å.

In the case of $^{15}N$-enriched RNA and DNA, one can determine only a limited number of $^{1}H-^{15}N$ RDC, because imine-group protons usually undergo fast exchange. Therefore, it is more convenient to use $^{1}H-^{13}C$ RDC for usual or $^{13}C$-enriched specimens.[268] Some C–H resonance signals of nucleic acids overlap appreciably even in 2D spectra, and special 3D experiments have been developed to resolve the signals.[269]

An elegant E COSY procedure of the $S^3E$ (Spin State Selective Excitation) type, which was called IPAP (In-Phase/Anti-Phase) procedure, has been proposed.[251] This procedure was successfully used in $^{1}H$-coupled [$^{15}N$,$^{1}H$]-HSQC and HN(CO) experiments, which made it possible to record in-phase and anti-phase *J*-coupled NMR spectra that supplement each other. The sum of and the difference between these spectra contain either only the high-field or only the low-field component of the doublet; therefore when they overlap, *J*-splittings can still be measured and the resolution does not deteriorate compared to that in the decoupled spectrum.

In most studies, different constants are found in different experiments. However, many specimens are unstable; moreover, the orientation of molecules in them can change. Therefore, 3D procedures have been developed; they allow one to measure several RDC of different types within a single experiment.[270]

Ottiger and Bax[253] measured simultaneously four RDC between directly bonded nuclei in the backbone of the $^{13}C/^{15}N$-enriched ubiquitin molecule. The results were in good agreement, especially with the $N-C'$ and $C^\alpha-C'$ RDC. The distances found using these constants coincided with the corresponding X-ray diffraction data, while the $N-H^N$ and $C^\alpha-H^\alpha$ distances proved to be somewhat longer. The researchers attributed this to vibrations whose amplitudes are much greater than the $N-C'$ and $C^\alpha-C'$ vibration amplitudes.

The $^{1}H-^{1}H$ coupling, resulting in line broadening in the spectra of oriented specimens, can be eliminated by using either adiabatic $^{1}H$-selective homonuclear decoupling[254] or partial deuteration.

The measurement of the $C'-N$ RDC is faced with some difficulties due to low gyromagnetic ratios of the nitrogen and carbon nuclei and to the longer $C'-N$ distance (1.33 Å). The

$D_{C'N}$ value is 8.3 times lower than $D_{NH}$. Therefore, the accuracy of determination of $D_{C'N}$ by this method can be considered to be critically low.[253]

The $D_{C'N}$ constant is usually determined from the difference between the $^1J_{C'N}$ splitting in isotropic and oriented phases, or from the in-phase splittings in HSQC (or TROSY) $^{13}C'$-coupled spectra,[252, 253] or from the anti-phase splittings in TROSY – HNCO $^{13}C'$-coupled spectra.[255] However, due to the insufficient resolution of many of the $^{15}N-\{^{13}C'\}$ doublets, their splitting cannot be measured with the required accuracy. To solve this problem, Permi and Annila[241] proposed that the doublet components should be displayed in two different subspectra (this is done in a similar way in the IPAP and $S^3E$ procedures described above).[251]

A different, more facile method for solving the problem of precise measurement of splittings in poorly resolved multiplets based on *J*-correlation spectroscopy principles has been proposed.[263] According to this method, the value of interaction is determined from the ratio of the amplitudes of two resonances rather from the difference between the resonance frequencies. To this end, relative intensities of the signals in two 3D TROSY – HNCO spectra with different $^1J_{C'N}$ dephasing gaps, equal to $1/2\,^1J_{C'N}$ and $1/^1J_{C'N}$, are found. The researchers also proposed a different, perhaps, more facile, way of recording two CT TROSY (Constant Time TROSY) $^{15}N-^1H$ spectra, namely, with $^{13}C'$-decoupling or with a 180-degree pulse at the $^{13}C'$ frequency applied simultaneously with the $^{15}N$-pulse shifted by 180°.

Yet another modification of the HNCA – TROSY experiment allows the assignment of the signals of atoms along the biomolecular chain, together with the possibility of measuring the scalar and dipolar constants.[271]

A procedure of spinning the specimen at different angles has been developed for fast determination of the sign of the $^1H-^{13}C$ RDC.[272]

**b. Dipolar coupling constants between the nuclei separated by several bonds**

The above described experiments are suitable for determination of only RDC for directly linked atoms. The measurement of homonuclear and heteronuclear RDC through several bonds is hampered not only due to their lower values but also due to the lower values of spin – spin coupling constants.[273] Nevertheless, in some studies, this is done using the procedures designed for the measurement of RDC of directly bonded nuclei. In particular, procedures for determination of $C'-H^N$ (see studies [241, 243, 252, 257, 274]), $C^\alpha-H^N$ (see studies [241, 243, 244, 274]), $N-C^\alpha$ (see studies [241, 244, 274]), $N-C'$ [243, 252, 274] and $C^\alpha-C'$ [243, 274] RDC, the $H^\beta-H^\beta$ and $H^N-H^N$ RDC for side chains [264, 274, 275] and H – H RDC for methyl groups [265] have been developed.

Two-dimensional two- and zero-quantum experiments for the measurement of geminal spin – spin coupling constants and RDC in the $NH_2$ and $CH_2$ groups of proteins and for the determination of the signs of these constants have been described.[274]

Vicinal RDC in proteins were measured most often for the $H^\alpha-H^N$, $C^\alpha-H^N$, $C^\gamma-H^N$ and $C^\delta-H^N$ pairs.[244, 276] Special *J*-correlated 2D spectroscopy techniques [277] and a new CT COSY (Constant Time COSY) type experiment [278] have been elaborated to measure the vicinal $^1H3'-^{31}P$ RDC for nucleic acids.

Below, we describe in more detail the works devoted to measurement of vicinal $^1H-^1H$ RDC.

The procedures described above designed for measuring the RDC of directly bonded atoms are poorly applicable to the $^1H-^1H$ coupling due to low values of the corresponding *J*-constants. Nevertheless, numerous methods for determining these RDC in overlapping multiplets have been developed, because these constants allow one to determine longer distances than the methods based on NOE.[219] A special 2D experiment has been proposed on the basis of the TOCSY (Total Correlation Spectroscopy) pulse sequence and called DCOSY (Dipolar Coupling Spectroscopy). The technique is based on the dependence of the direct cross-peak intensity on the RDC value. In this procedure, usual cross-peaks displayed in the spectra can be distinguished from the peaks of the relayed coherence transfer because they have different signs. A problem arises in the case of superposition of cross-peaks of these two types because signal deconvolution is impossible; this restricts the scope of the experiment.

The multiple relay cross-peaks are eliminated in another simple method for estimation of the $^1H-^1H$ RDC based on analysis of the diagonal and cross-peak intensities in CT COSY experiments.[279] However, since the cross-peaks in this experiment are anti-phase or out-of-phase, their integration is often faced with difficulties. Therefore, a 3D variant with additional $^{15}N$-editing has been developed;[280] the sensitivity of this method is not inferior to that of the NOE-based method. Another method for determining the $^1H-^1H$ RDC in CT COSY experiments is based on the time dependence of the zeroing conditions for proton – proton cross-peaks.[281]

Apart from methods based on estimation of the cross-peak intensities, special methods for direct determination of the sign and the magnitudes of the $^1H-^1H$ RDC from the splittings measured in the spectra have also been developed.[274, 282]

An HNCA type experiment with observation of E COSY between the $H^N$ and $H^\alpha$ protons has been developed.[282] Unfortunately, this method is applicable to the precise determination of the values and signs of RDC only between $H^N$ and $H^\alpha$.

Peti and Griesinger [283] have further developed this method into a three-dimensional one by adding a NOESY (Nuclear Overhauser Enhancement Spectroscopy) stage. Their method, called *J*HH NOESY ($J_{HH}$-based NOESY), requires a minimum isotope enrichment and can be combined with elements of the $S^3E$ experiment for further spectral resolution enhancement.

An advantage of another 3D experiment (HNHA) developed for the measurement of unresolved RDC between the protons of amide groups and other protons in $^{15}N$-enriched protein molecules is the possibility of using homonuclear decoupling and the TROSY procedure, which allows one to study large oriented proteins.[284] Yet another simple and reliable method for the measurement of proton – proton RDC in the spectra with unresolved but not very large multiplets is based on iterative analysis of the shapes of multiplets in phase-sensitive COSY spectra.[285]

The signs of $^1H-^1H$ RDC depend on the orientation of the proton – proton vector with respect to the external magnetic field, and their determination is often difficult, especially in the absence of scalar coupling between the protons or if the protons do not have a common partner through the bonds. Therefore, a very convenient *J*HH TOCSY ($J_{HH}$-based TOCSY) type experiment has been developed and called 'signed COSY'.[286] This method determines the sign of the $^1H-^1H$ RDC with respect to the sign of the $^1H-X$ direct coupling constants in linear $X-^1H-^1H$ three-spin systems (X = $^{13}C$ or $^{15}N$). The advantages of this method include high sensitivity and the possibility of diagonal peak suppression, cross-peaks being observed only between the directly interacting protons.

Recently, it was found unexpectedly that the signs and magnitudes of the $H^N-H^\alpha$ RDC can be determined even in a usual homonuclear COSY experiment.[287] It was found that only RDC with positive signs are manifested in the spectrum, while cross-peaks corresponding to negative RDC are either missing or greatly broadened and can be easily identified.

## 3. Measurement of the chemical shift anisotropy

The chemical shift is a tensor value. It is anisotropic, only the average value being observed for an isotropic liquid. In a study of a single crystal, all the six components of the symmetrical tensor can be found experimentally by measuring the values along the principal axes of the crystal. Only diagonal elements can be determined for a powdered specimen. However, due to the pronounced line broadening in the NMR spectra of solids, the accuracy of these measurements is low. Liquid crystals have some

advantage in this respect, although only a linear combination of diagonal elements can be determined.

The chemical shift anisotropy $\Delta\delta$ for each nucleus on passing from isotropic into an oriented state is determined using the relation

$$\Delta\delta = \sum\sum A_{jj}\delta_{ii}\cos^2\theta_{ij},$$

where $\theta_{ij}$ is the angle between the principal axis of the chemical shift tensor $\delta_{ii}$ and the principal axis of the diagonalised molecular alignment tensor $A_{jj}$; summation involves all $i$ and $j$.

The $\Delta\delta$ value is a key parameter in investigations of biomolecular structures by solid-state NMR and in relaxation measurements to study chain dynamics. The correlation between the CSA and dipolar coupling provides important information about the angles [67] and is a key value in TROSY experiments.[26]

The $\Delta\delta$ value for aligned biomolecules can be directly measured. In this case, the change in the chemical shift induced by transition of a biomolecule from the isotropic phase to the aligned phase provides information on the orientation of the CSA tensor of this nucleus with respect to the molecular alignment tensor.[175] In order to use this information, one has initially to find the components of the order parameter tensor of a biomolecule (which can be determined, for example, from RDC) and the principal components of the CSA tensor for the given nucleus (which can be determined from solid-state experiments or calculated if the structure of the protein under study is known). The measured differences between the chemical shifts of the isotropic and oriented molecule are converted into values describing the orientation of the vectors of a pair of bonds relative to the order parameter tensor. For example, for $^{13}C'$ nuclei of the carbonyl groups in the proteins, the vectors of the $C'-O$ and $C'-N$ bonds can serve as this pair;[288] for phosphate $^{31}P$ atoms in nucleic acids, these are $P-O(1)$ and $P-O(2)$ bond vectors.[289]

The possibility of using the CSA parameter, together with RDC, in the structural studies of biomolecules was first indicated by Tjandra and Bax;[45] the same researchers confirmed this experimentally.[141]

According to Ottiger *et al.*,[290] the average $\Delta\delta$ value for the $^{15}N$ nucleus in the amide group in protein molecules is $-168 \pm 20$ ppm and the angle $\alpha$ between the $N-H$ bond vector and the principal axis of the $^{15}N$ CSA is $13 \pm 5^\circ$. A more thorough determination of the $^{15}N$ $\Delta\delta$ for the lysozyme molecule confirmed the legitimacy of using the average $\Delta\delta$ and $\alpha$ values in proteins:[291] values of $-171.5 \pm 2.2$ and $18.3 \pm 0.8^\circ$, respectively, were found.

A correlation between the $^{13}C^\alpha$ CSA and elements of the secondary structure of the biomolecule has been established.[292] The $\Delta\delta$ value for $C^\alpha$ atoms in $\alpha$-helices was found to be $6.1 \pm 4.9$, that for $\beta$-sheets is $27.1 \pm 4.3$ ppm.

The procedure for involvement of CSA values in structure determination has been substantiated and developed taking the CSA of carbonyl groups (which can be easily measured) as examples.[288, 293] The increase in the structure accuracy attained thereupon was estimated.

In a series of studies, Tjandra and Bax and co-workers justified the possibility of inclusion of this type of information, namely, the CSA of the amide nitrogen atoms,[290] carbonyl $C'$ atoms [293] as well as $C^\alpha$ (see Ref. 292) and $^1H^N$ (see Ref. 294) atoms into standard procedures for structure determination of biomolecules.

Particular examples of using CSA can be found in subsequent publications. The $\Delta\delta$ value for the carbonyl-group $C'$ atom was used as a restraint in the well-known XPLOR programme.[295]

In metal-containing proteins, so-called pseudocontact paramagnetic shifts appear due to the presence of a paramagnetic centre;[296] these can be used as additional structure constraints. These constraints were taken into account in determining the structures of a number of protein biomolecules. Owing to the pseudocontact shifts, one can exactly determine the position of the metal ion in a protein and to determine the components of the magnetic susceptibility tensor of a protein.[158, 297]

## 4. Measurement of the quadrupole coupling constants

Nuclei with a spin of $> 1/2$ possess a nuclear electrical quadrupole moment and interact with an electric field gradient. This interaction is described by the second-range quadrupole coupling tensor **q**. Its $z$-component, defined as

$$q_{izz} = \frac{\varepsilon Q_i V_{izz}}{h},$$

where $\varepsilon Q_i$ is the quadrupole moment of an $i$th nucleus, $V_{izz}$ is the element of the electric field gradient tensor at the $i$th nucleus, can be manifested in the NMR spectra.

The quadrupole coupling constant can be found from line splitting in the NMR spectrum and order parameters of the molecule in the nematic phase. Since a QCC depends only on the electronic properties of the ground state and the nuclear electric quadrupole moment, it bears information on the electronic structure around the nucleus under study. The QCC values found from the quadrupole splitting observed in the nematic phase can be used to determine the orientational and geometric parameters of the molecule.

The tensor **q** has a zero trace and, therefore, it has almost no influence on the NMR spectra in isotropic solutions, except for line broadening in some cases. Line broadening is also observed in anisotropic solutions, but simultaneously, additional splittings appear due to the non-zero average $q_{izz}$.

In practice, this splitting is not observed for most of the quadrupole nuclei due to fast relaxation. The iodine, bromine and chlorine nuclei, having very high $\varepsilon Q_i$ values, cannot furnish resolved NMR spectra unless the symmetry or alignment are such that they give relatively small $V_{izz}$. In the case of $^2H$ and $^{14}N$ nuclei, the $\varepsilon Q_i$ value is much smaller; therefore, quadrupole splitting can be observed in anisotropic solvents (however, the relaxation rate for $^{14}N$ is most often too great for observing a high-resolution spectrum).

The most convenient nucleus is deuterium with a much smaller nuclear quadrupole moment and moderate line broadening. Therefore, well resolved dipole–dipole interactions between protons and deuterons can be observed in a proton spectra of partially deuterated compounds. It should be noted that $^2H$ NMR spectra often do not display resolved dipole splittings, mainly due to the fact that deuteron–deuteron interactions are 36 times weaker than the corresponding proton–proton interactions.

Quadrupole couplings are much stronger than dipolar couplings under the same experimental conditions; therefore, they could be an alternative to the latter in the structural studies of weakly aligned biomolecules. In this case, deuterium is also the nucleus of choice. However, low sensitivity (100 times lower than the sensitivity of a proton), together with low natural abundance in biomolecules (667 times less than that of hydrogen) restrict its practical use.

The dependence of quadrupole couplings on the degree of alignment of molecules shows itself in the quadrupole splittings of signals. In NMR experiments with oriented molecules, these splittings are used only to confirm and measure the ordering of the lyotropic LC phase.

The quadrupole coupling constants of deuterium in $^{13}C/^{15}N$-enriched proteins dissolved in $D_2O$, in which all the protons of amide groups are subject to exchange with solvent deuterons, can be used to describe $N-{}^2H\cdots O=C$ hydrogen bonds, moreover, in non-oriented specimens.[298] This is done relying on the dependence of the relaxation time $T_1$ for deuterium on QCC, which, in turn, depends on the length of the hydrogen bond and the angle between the amide and carbonyl bonds.

The use of QCC for investigations of weakly aligned biomolecules became possible owing to the recently developed procedures for partial deuteration. However, line broadening caused by

quadrupole relaxation and non-resolved dipole splittings in the NMR spectra make QCC less promising than RDC. Nevertheless, using the procedure for measurement and interpretation of $^2$H quadrupole splittings in the spectra of partially deuterated and oriented proteins reported by Mittermaier and Kay,[299] one can observe the $^2$H quadrupole splittings of the side-chain methyl groups in a protein molecule.

The possibility of using QCC for determining the structure and dynamics of weakly aligned deuterated carbohydrates has been demonstrated.[300] A special pulse scheme of $^1$H–$^2$H 2D experiment was developed, and good prospects for this procedure for investigation of small carbohydrate molecules were shown. This requires, however, partial deuteration and $^{13}$C-enrichment of the molecules. The applicability of this scheme to large biomolecules is also restricted by the $^2$H relaxation times.

## VI. Strategy of structure determination for weakly aligned biomolecules by NMR spectroscopy

### 1. Algorithms of structure calculation

Different types of information are used to determine the structure of a biomolecule. An algorithm for structure solution[301] implies minimisation of the system total energy function $E_t$

$$E_t = E_c + E_{vW} + E_{NMR},$$

where $E_c$ determines the covalent geometry (bond lengths, angles, planarity, chirality); $E_{vW}$ takes into account contacts without covalent bonding (for example, repulsion or some empirical functions including the Lennard–Jones, electrostatic and hydrogen bond potentials); $E_{NMR}$ describes the experimental NMR restraints which are established during different NMR experiments carried out in different media.

The last term $E_{NMR}$ consists (arbitrarily) of three terms

$$E_{NMR} = E_i + E_a + E_{DB},$$

where $E_i$ includes the NMR parameters which do not depend on the orientation; $E_a$ includes the orientation-dependent NMR parameters; $E_{DB}$ corresponds to empirical information found in structural databases.

The term $E_i$ involves the NMR parameters determined from isotropic spectra. In particular,

$$E_i = E_O + E_J + E_{J_{CH}} + E_{J_{CC}} + E_{C(shift)} + E_{H(shift)} + \\ + E_{relax} + E_{N(relax)} + E_{CCR} + \ldots$$

Here, $E_O$ takes into account the NOE cross-peaks; $E_J$ is for the vicinal $^3J$ constants; $E_{J_{CH}}$ corresponds to $^1J_{CH}$, $^2J_{CH}$ and $^3J_{CH}$ constants; $E_{J_{CC}}$ includes the $^1J_{CC}$, $^2J_{CC}$ and $^3J_{CC}$ constants; $E_{C(shift)}$ includes the $C^\alpha$ and $C^\beta$ chemical shifts; $E_{H(shift)}$ is the term for proton chemical shifts; $E_{relax}$ is the ratio of the $^{13}$C and $^2$H relaxation times; $E_{N(relax)}$ is the ratio of the $T_1/T_2$ relaxation times for the $^{15}$N atom; $E_{CCR}$ takes into account cross-correlated relaxation data, *etc*.

For each term in the function, a potential is introduced. Usually, it has the form

$$E = k(\delta_{calc} - \delta_{obs})^2,$$

where $k$ is the force constant, and $\delta_{calc}$ and $\delta_{obs}$ are the calculated and measured values of the corresponding energy, respectively.[141, 302]

Now we consider the $E_a$ term in more detail

$$E_a = E_d + E_{CSA} + E_{QCC} + \ldots,$$

where $E_d$ represents the RDC for the $^{15}$N–$^1$H, $^{13}$C–$^1$H, $^1$H–$^1$H and $^{15}$N–$^{13}$C pairs; $E_{CSA}$ is the chemical shift anisotropy; $E_{QCC}$ are the quadrupole couplings.

A set of so-called distance and angle restraints for the molecule under study is formed from the different-type data obtained in NMR experiments; these restraints are used in the iteration procedure for minimisation of the total energy of the biomolecule in order to find the global energy minimum. This procedure, called adjustment, may use various methods developed in molecular dynamics. The 'simulated annealing' technique is the most popular.

The result of adjustment is evaluated using numerical performance criterion.[303] In crystallography, the *R*-factor — a quantitative estimate of the degree of agreement between the measured and calculated reflections — is a key criterion of the quality of structure determination. It is defined as

$$R = \sum \frac{|\, A_{exp} - A_{calc} \,|}{|\, A_{exp} \,|},$$

where $A_{exp}$ and $A_{calc}$ are the experimental and calculated values, respectively. A similar criterion can also be defined for the set of RDC. However, introduction of the *R*-factor in the same form is impossible for RDC, because the dipolar coupling tensor is unitary; hence, the average $|A_{exp}|$ value will be zero.

It has been proposed[293] to define the *Q*-factor for CSA in the form

$$Q = \frac{\mathrm{rms}(\Delta\delta_1 - \Delta\delta_2)}{\mathrm{rms}(\Delta\delta_1)},$$

where the numerator is the root-mean-square difference between the measured ($\Delta\delta_1$) and predicted ($\Delta\delta_2$) values.

This criterion is also applicable for RDC and for other orientation-dependent values. However, for these, it is not quite convenient for interpretation. Although in the case of full fitting, its value is 0, the value for completely random structure is $\sqrt{2}$. Therefore, the following *R*-factor was proposed for RDC:[303]

$$R_{dip} = \left[\frac{\langle (D_1 - D_2)^2 \rangle}{2\langle D_1 \rangle^2}\right]^{1/2},$$

where $D_1$ and $D_2$ are the predicted and calculated RDC values.

This fitting criterion is smaller than the previous one by a factor of $\sqrt{2}$. Theoretically, this quality criterion corresponds to the case where the distribution of bond vectors is completely random, while their number is infinite. In reality, the vector distribution is not entirely random, the number of measured values is small and the $(2\langle D_1^2 \rangle)^{1/2}$ value determined experimentally is not equal to the expected root-mean-square difference between the measured and predicted values. Therefore, in the same publication,[303] a different quality criterion was proposed to take into account the pattern of distribution of the vectors

$$R_{dip} = \left[\frac{5\langle (D_1 - D_2)^2 \rangle}{2A_a^2(4 + 3\eta^2)}\right]^{1/2},$$

where $A_a$ is the axial component of the order parameter tensor and $\eta = A_r/A_a$ is its rhombicity.

The success of using NMR spectroscopy techniques for determination of the structures of biomolecules is largely determined by the possibility of automation. With the development of multidimensional methods and triple resonance experiments, signal assignment procedures and 3D structure determination were consecutively automated. This process is currently in progress and there are grounds to believe that in the near future, NMR would be competitive with X-ray in the determination of the 3D structure of biomolecules regarding not only the accuracy but also the duration and cost.[304]

Now we consider a typical procedure used to analyse the spectra taking protein structure determination as an example.

First, automated processing of the multidimensional spectrum is carried out in order to identify, describe and catalogue all the observed signals and also to verify them (distinguish real signals from possible artifacts).

The next stage is to assign the identified signals to particular members of the primary sequence; this is accompanied by identification of secondary structure elements in the biomolecule. This is

done using information for homologous molecules with a known structure. The secondary structure elements are recognised or confirmed using known chemical shifts, *J*-constants, specific NOE signals and hydrogen bonds. The use of $^1H-{}^1H$ RDC in standard calculation programmes is faced with some problems, because the corresponding distances and angles with the magnetic field are unknown for most protons, unlike the constants through one or two bonds. Therefore, the procedure for the calculation of a biomolecular structure should envisage the possibility of varying both the length and orientation of the internuclear vector. In particular, it was shown for ubiquitin that inclusion of the $^1H-{}^1H$ RDC in the calculation markedly increases the accuracy of the structure determination.[305]

Finally, the last stage is determination of the 3D structure of the protein. Of prime importance in this stage is the problem of folding. Homologous structures for folding are found by searching through databases (see below).

An algorithm for fast packing of the already known elements of the secondary structure of the protein molecule to form a protein fold using the residual RDC has been developed:[306] first, the relative orientations of the secondary structure elements with respect to one another are determined, then, using data on the elements located at short distances, the packing of secondary structure elements is completed.

The inclusion of the option of RDC analysis into the Rosetta *ab initio* programme allows one to solve the protein folding problem relying only on these constants, without using additional information.[307]

Tian *et al.*[308] developed a strategy of simultaneous resonance assignment and folding of the protein molecule based on RDC measurements. This strategy allows rather quick and reliable identification of most signals in the NMR spectrum and provides a good initial approximation for the structure of a biomolecule on the basis of brief information on the chemical shifts and vicinal constants. This sharply accelerates the NMR structure determination procedure for proteins. Another similar strategy[309] is based on a single 3D triple-resonance experiment (HA)CANH.

In the general case, information of different types can be used as the initial data:

— distance restraints found in the analysis of multidimensional NOESY spectra;

— torsion angle restraints determined from experimental and/or statistical data including data from relaxation experiments, NOESY, chemical shifts, and spin–spin coupling constants;

— residual dipolar couplings, CSA and QCC;

— distance restraints from the data on disulfide bridges and hydrogen bonds.

There exist several algorithms for structure determination using this set of data embodied in different computer programmes.[141, 301] Most of the conventional programmes were developed before the experiments with oriented biomolecules were carried out. As spectroscopic information on oriented molecules was accumulated, these programmes were upgraded. In particular, the XPLOR programme was modified,[302, 305] the algorithms for inclusion of RDC information into the CNS programme were developed,[310] and the DYANA programme was supplemented by the PSEUDYANA module designed to take into account the pseudo-contact shifts.[154, 158, 297] New versions of these programmes contain all the parameters determined from the NMR spectra of oriented biomolecules.

An important component of the structure determination for a biomolecule is graphical visualisation of molecular structures. This is done, most often, using the graphical viewers MOLMOL, GRASP, RASMOL, WHATIF and MIDAS. For the computation and visualisation of the structures of biomolecules from different-type NMR data, the combined VMD–XPLOR programme package was developed;[311] The special IVM module, which allows use of the intrinsic coordinates instead of Cartesian ones, was designed to accelerate the adjustment procedure.[312]

The interactive MODULE programme,[313] which permits one to determine order parameters for macromolecules and for multidomain systems and complexes on the basis of RDC, is an excellent add-in to the existing programme packages used in the molecular simulation.

The achievements and prospects for the development of research into the theory and calculation of the structures of biomacromolecules using NMR data have been covered in detail in reviews.[304, 314]

## 2. Dipolar coupling and the order parameter tensor

Dipolar coupling is a unique feature of matter which provides information on molecular structure and dynamics. The dipolar coupling between the spins of nuclei *i* and *j* situated in a magnetic field is characterised by the dependence on the distance $r_{ij}$ between them and on the angle $\theta$ between the internuclear vector $\mathbf{r}_{ij}$ and the magnetic field (Fig. 3)

$$D_{ij} = \frac{K(3\cos^2\theta - 1)}{2r_{ij}^3},$$

where $K = (h\gamma_i\gamma_j/4\pi^2)$ is the constant depending on the gyromagnetic ratios $\gamma_i$, $\gamma_j$ of nuclei *i* and *j*, respectively.

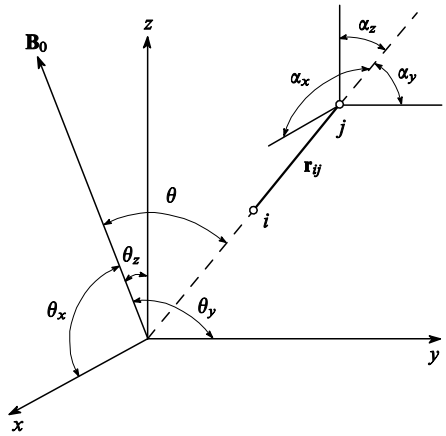



**Figure 3.** Scheme representation of the parameters used for determination of the dipolar coupling and elements of the Saupe order matrix; *i* and *j* are nuclei located in magnetic field $\mathbf{B}_0$; $\theta$ is the angle between the internuclear vector $\mathbf{r}_{ij}$ and the magnetic field direction; (*x*,*y*,*z*) is a molecular-fixed system of coordinates; $\theta_x$, $\theta_y$, $\theta_z$ are the angles between the corresponding molecular axes and the magnetic field direction; $\alpha_x$, $\alpha_y$, $\alpha_z$ are the angles between the vector connecting the nuclei *i* and *j* and the molecular axes *x*, *y*, *z*.

Dipolar couplings occur both between the nuclei of one molecule and between the nuclei of neighbouring molecules. In an isotropic medium, rigid molecules are arranged chaotically and execute chaotic translational and reorientation motion; therefore, dipolar coupling is averaged over the whole molecular ensemble and over the observation time, *i.e.*,

$$D_{ij} = K\left\langle \frac{3\cos^2\theta - 1}{2r_{ij}^3} \right\rangle$$

(the angled brackets mean averaging over time and over all the possible orientations and distances).

Apart from participation in the molecular motion as a whole, the nuclei are involved in the intramolecular motion; the RDC are averaged over these motions, too. Since the characteristic frequencies of these types of motion are substantially different, they are supposed to be not correlated, and these two averaging operations are considered to be independent, *i.e.*,

$$D_{ij} = K\left\langle \frac{3\cos^2\theta - 1}{2} \right\rangle \langle r_{ij}^3 \rangle^{-1}.$$

If the magnetic field is uniform (as is normally the case for NMR spectrometers), the translational motions do not influence the averaged constant and it depends only on the reorientation motion and the instantaneous distribution of the molecular axes in the magnetic field.

In an isotropic medium in the absence of aligning factors, the spatial distribution of molecular axes is spherically symmetrical and the times of molecular reorientation are markedly shorter than the characteristic times of an NMR experiment; therefore, in this medium, the dipolar couplings are completely averaged and are not manifested in the spectra because the average $\langle(3\cos^2\theta - 1)/2\rangle$ is zero.

Under the action of aligning factors or when the molecules are placed in an aligned medium, the molecular motion in the specimen becomes anisotropic, the direction along the orientation axis, which is called the director, being the predominant one. The molecules of the compound under study are compelled to move anisotropically even if they do not tend to be oriented along the director. This anisotropic motion, together with the partial alignment along the director is responsible for the fact that magnetic dipole–dipole interactions are observable, because in this case, the average $\langle(3\cos^2\theta - 1)/2\rangle$ value is no longer equal to zero.

Saupe[315] proposed describing the degree of ordering of a molecule by the value $\langle(3\cos^2\theta - 1/2\rangle$, which he called the order parameter tensor. This value can range from $-0.5$ to 1. Zero corresponds to the absence of orientation and the values 1 and $-0.5$ correspond to the maximum degrees of alignment along and normal to the direction of the orienting factor, respectively.

Since the magnetic field in a spectrometer has spatially fixed axes, it is convenient to express the components of coupling tensors in this laboratory system of coordinates, which can be converted into a molecular-fixed system of coordinates in different ways, namely, using the order matrix, the formalism of the motion constants or averaged Wigner rotation matrix. In the Saupe scheme, the alignment of molecules is described by a symmetrical order matrix $S$ having a zero trace, whose elements are determined by the expression

$$S = \frac{1}{2}(3\cos\theta_p \cos\theta_q - \delta_{pq}),$$

where $p, q$ mean one of the Cartesian axes ($p, q = x, y, z$) of a molecular-fixed system of coordinates; $\theta_x$, $\theta_y$ and $\theta_z$ are the angles between the corresponding molecular axes and the direction of the magnetic field (see Fig. 3); $\delta_{pq}$ is the Kronecker symbol ($\delta_{pq} = 1$ for $p = q$ and $\delta_{pq} = 0$ for other cases).

The degree of order along the axis containing two nuclei $i$ and $j$ and having the angles $\alpha_x$, $\alpha_y$ and $\alpha_z$ with molecular axes $x$, $y$ and $z$ is determined by the relation

$$S_{ij} = \sum \cos\alpha_p \cos\alpha_q S_{pq}.$$

The $3 \times 3$ order matrix is characterised by only five independent elements (because it is symmetrical and has a zero trace, *i.e.*, $S_{xx} + S_{yy} + S_{zz} = 0$). With appropriately chosen molecular axes, the matrix can be reduced to a form in which its elements are independent and their number varies from 1 to 5, depending on the symmetry of the molecule or, more precisely, the symmetry of the system of interacting nuclei. For example, the degree of alignment of highly symmetrical molecules having a threefold (or higher-fold) axis of symmetry can be described by only one independent element of the matrix. The degrees of alignment of molecules having two perpendicular symmetry planes, can be described by two independent elements. When the molecule has only one plane of symmetry, knowledge of the three independent elements of the order parameter matrix is required. In the general case, where the molecule is asymmetric, three elements determine the direction of the principal axis of the tensor with respect to the magnetic field and the other two describe the degree of orientation along this direction and the degree of orientation in the plane normal to the principal axis. In the system of $x$, $y$, $z$ principal axes, the order parameter tensor is diagonal and only two independent values determine the above degrees of orientation: $S_{zz}$ and the asymmetry parameter

$$\eta = \frac{S_{xx} - S_{yy}}{S_{zz}}.$$

Since the symmetry of biomacromolecules is low, as a rule, all five elements of the order parameter tensor need to be known to describe their orientation.[316] In these molecules, the distances between the directly bonded nuclei virtually do not change and only two independent values determine the orientation, namely, the order parameter and the angle between the bond vector and the magnetic field direction; therefore, these parameters are often called orientational constraints.

For clarity, the orientation of biomolecules is commonly described, instead of $S_{ij}$ tensor components of the order matrix, as is usually done in NMR spectroscopy of small molecules,[131] by the axial $A_a$ and rhombic $A_r$ components of the diagonalised molecular orientation tensor **A**, *i.e.*, in the system of the principal axes of the tensor of the magnetic susceptibility anisotropy of the molecule and three Euler angles determining the position of the principal axis of this tensor relative to the direction of the magnetic field. In this system, the expression for the RDC between the nuclei P and Q has the form

$$D_{PQ}(\theta,\varphi) = S\frac{\mu_0}{4\pi}\gamma_P\gamma_Q h \frac{A_a(3\cos^2\theta - 1) + (3/2)A_r \sin^2\theta\cos 2\varphi}{4\pi^2 r_{PQ}^3},$$

where $\theta$ and $\varphi$ are the angles for the vectors of the nuclei P and Q in the coordinate system of tensor **A**, $h$ is the Planck constant, $\gamma_P$ and $\gamma_Q$ are the gyromagnetic ratios of these nuclei, $r_{PQ}$ is the distance between them, $S$ is the main order parameter, $\mu_0$ is the magnetic susceptibility of a vacuum.

If the magnetic susceptibility tensor is axially symmetric, then $\eta$ (or $A_r$) is equal to 0, and only one parameter $S_{zz}$ (or $A_a$) determines the RDC, which, in turn, determine the direction of vectors connecting the pairs of nuclei relative to the principal orientation axis. This allows easy interpretation of the dipolar constants, for example, for the DNA molecule or some DNA–protein complexes.[141] For the case where the aligned molecules have a nearly axial symmetry, a simple algorithm for including the information on dipolar couplings has been developed using the CNS computation programme as an example.[317]

However, most often, the molecule is not axially symmetrical; therefore, for data interpretation, one should also determine the second component of the tensor, which measures the degree of its deviation from the axial symmetry, *i.e.*, its rhombicity and asymmetry. This is a difficult task for molecules with a small number of magnetic nuclei.

For biomolecules containing a large number of identical pairs of nuclei whose vectors are approximately uniformly distributed in space, this task is markedly easier.[318] In this case, by analysing the column diagram for the distribution of dipolar couplings (corresponding to the spectrum of a powdered specimen), one can determine unambiguously the axial and rhombic components of the diagonalised tensor (Fig. 4). No preliminary structural information on the biomolecule is required; in this stage, there is even no need to determine the main direction of the tensor.

The same principle underlies a method used to determine components of the diffusion tensor from the $T_1/T_2$ distribution for $^{15}$N nuclei of the amide groups in the protein molecules.[319] It should be noted that diffusion and orientation tensors for proteins in bicelles are unidirectional,[320] which can also be used in determining the orientation of a biomolecule.

To evaluate the possibility of determining the tensor components from the set of measurable orientation-dependent parame-

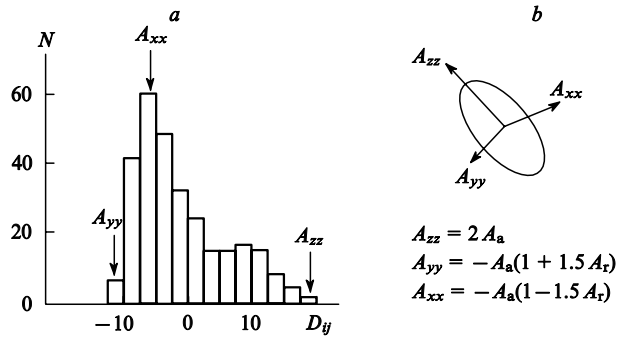



**Figure 4.** Column diagram for the distribution of dipolar couplings (*a*) and the component of the diagonalised molecular alignment tensor **A** (*b*).[318]

ters, a method for analysing the pattern of filling of the orientational space by internuclear vectors has been proposed.[321]

If only few RDC have been determined for a biomolecule, or they are non-uniformly distributed in space, or they contain several domains with different orientations, the column diagram analysis mentioned above is inapplicable. Therefore, a more advanced procedure for the analysis of the whole set of RDC and determination of the axial and rhombic components of the order parameter has been developed on the basis of the maximum likelihood method.[322]

Determining the orientation of the order parameter tensor relative to the molecule-fixed system of coordinates is a rather complicated problem, which can be solved unambiguously if the structure of the biomolecule is already known. This results in multiple possible orientations of the vectors of each pair of nuclei, corresponding to the mirror reflections relative to each axis of the order tensor (altogether 8). This appreciably complicates the energy hypersurface in searching for the minimum. Finally, the convergence of the iteration is very slow.

Attempts at estimating the order parameter tensor of the biomolecule on the basis of calculation results or specific experiments have been undertaken in a number of studies. The most trivial approach implies a network search among the possible pairs of the parameters $A_a$ and $A_r$.[302]

If data on the biomolecule structure have been derived preliminarily from other experiments, they can be used for more precise determination of the axial and rhombic components of the tensor.[323, 324] Finally, one can do without the preliminary determination of the values of this tensor by including them as variable parameters into programmes for calculation of the biomolecule geometry.[325] In particular, such a modification of the XPLOR algorithm has substantially facilitated the calculation and accelerated its convergence; the obtained results are very close to those found with the conventional algorithm.[326]

To simplify the analysis of RDC, a method for predicting the orientational order has been proposed, based on a model taking into account the steric repulsion between the protein molecule and the aligning particles of the LC medium.[197]

A more advanced method for predicting the order parameter taking into account only steric effects has been reported;[327] in this method, the protein molecule under study is modelled by an ellipsoid. The principal axes of the orientation tensor and the moment of inertia tensor of the molecule coincide. The other components of the orientation tensor are determined using the standard algorithm for minimisation of the difference between the predicted and experimental RDC. The order parameter tensor differs from the moment of inertia tensor because it has a zero trace. Therefore, when information about the other two axes of the moment of inertia tensor is used to predict the other elements of the order parameter tensor, the eigenvectors of these tensors are taken to be parallel if there are no electrostatic effects.[328] An analytical solution to the problem of alignment of ellipsoidal particles by planar obstacles has been obtained. If the 3D structure of the protein molecule is known, the order parameters of both the protein molecules incorporated in membranes and the molecules aligned under the action of a magnetic field or an LC medium can be predicted using simple expressions.

In a more convenient algorithm for predicting the order parameter tensor from the moment of inertia tensor, the eigenvalues of the order tensor are proportional to the molecular asymmetry in the planes determined by the corresponding pair of eigenvectors of the moment of inertia.[329] This method, which the authors called TRAMITE (Tracking Alignment from Moment of Inertia Tensors), proved to be very convenient in investigating not only protein molecules but also non-rigid or small molecules.

The relative orientation of the order parameter tensors can be determined without knowing the protein structure on the basis of correlation analysis of two sets of RDC by conducting experiments in two media that orient in different directions.[330] The graphical representation of the correlation map of such sets depends appreciably on the relative orientation of the tensors; therefore, it allows one to obtain a good initial approximation for further analysis of RDC.

A new strategy for the calculation of the 3D structure based on the transformation of a set of RDC into a set of angles between vectors has been developed.[331, 332] The values comprised in this set no longer depend on the direction of the orientation tensor with respect to the molecular axes; they can be used in calculations instead of RDC.

Skrynnikov and Kay[333] also proposed that a set of pairwise angular restraints independent of the chosen system of coordinates be used instead of RDC. This set proved to be very convenient for evaluating the compliance of the set of RDC with the next structure generated in adjustment programmes. Using $n$ vectors of dipole couplings, $n(n-1)/2$ new values determining the angles between each pair of vectors are composed. In addition, the maximum and minimum values are determined for each angle between the vectors proceeding from the current generated structure. If the vector angle does not fall in the range between these values, it is considered as a mismatch. The total number of these mismatches can serve as a quality criterion for a given structure and can be used in calculations. The compliance of RDC with the structure is visualised by the 'RDC compliance map'. Using this map, it is convenient to evaluate the quality of each generated structural element or domain and its compliance with experimental data.

The procedures considered are based on the strategy of a generalised, model-free order parameter, developed previously for analysing the diffusion tensor in the interpretation of relaxation data.[334, 335]

Yet another method which allows one to solve the problem of ambiguity and to obtain an exact solution for the set of equations with RDC is to convert these equations to elliptical equations in the principal alignment frame.[336]

The effects caused by molecular motions of different types that influence the resultant RDC values of a globular protein have been studied by molecular dynamics methods.[337] The structures of biomolecules calculated by virtue of these constants for different models of globule reorientation proved to be model-dependent. It was found, however, that the application of the developed self-coordination procedure to the sets of RDC measured for several (in the ideal case, five) different oriented LC media allows one to determine the structure without making suggestions concerning the pattern of the reorientation motion of the globule. This procedure does not require the simulation of these motions.

Taking rubredoxin as an example, it was shown that two or more different alignment media can be used to remove the degeneracy of orientation parameters when determining the structure of protein molecules (in order to obtain a larger number of parameters).[338, 339]

Actually, the dependence of dipolar couplings on the degree of alignment of the protein molecules and the globule orientation in

the spectrometer magnetic field can be used for obtaining numerous sets of parameters for different directions and magnitudes of the order parameter tensor, for example, in experiments with various aligning factors and also with a change in the conditions affecting the orientation (temperature, relative concentrations of components and additives). However, the sets of parameters determined are often linearly related and do not provide new information.

The precision of NMR determination of the structure of a biomolecule is still inferior to that of X-ray determination. One of the reasons for this is the lack of perfection of the computation algorithms used, which mainly copy the X-ray algorithms. With the progress in NMR techniques, specific computation algorithms that reflect the specific character of NMR experimental data are developed.[340]

## 3. The effect of vibrational motions

When the set of RDC is interpreted, the bond lengths are usually assumed to be constant and known for each type of bond. Therefore, the measured RDC describe either the variation in the angle orientation of the mean position of the bond or fluctuations around this mean position. However, RDC-based studies of the structures of small molecules have shown that the effect of the molecular vibrations may be appreciable in some cases.[341]

The influence of vibrational motions on the averaging of RDC has been studied by quantum chemistry methods for direct $N-H^N$, $C^\alpha-H^\alpha$, $C'-N$ and $C^\alpha-C'$ RDC and also the RDC between $C'-H$ nuclei separated by two bonds.[342] The vibrational averaging was shown to influence the constants and to entail a resultant elongation of the effective bond lengths by 0.5% – 3%, the effect of bending or torsion vibrations being more pronounced than the effect of stretching vibrations. The hydrogen bond formed by an amide group also increases the effective N – H bond length. The changes in the effective bond length are fairly small but, nevertheless, they can be used, especially in those cases where RDC from different types of bonds (nuclear pairs) are used to determine the structure of a biomolecule.

## 4. Databases for energy minimisation and for the search for homology and motifs

The use of homologous sequences with a known structure to generate structural models for a biomolecule is helpful for structural studies.

The most labour-consuming and prolonged stage in NMR investigations of the protein structure is the assignment of signals in the spectra to particular amino acid residues of a known primary sequence. This is especially true for NOESY spectra.

In protein databases, information on the spatial structure of a large number of proteins has been accumulated and the results of processing of numerous NMR spectra are presented. This information can be used for recognising separate structural elements and motifs often repeated in proteins. This is especially useful at the preliminary stage.

Apparently, RDC are the most convenient values that can be measured and used to search for similar images in protein databases. A trivial method is based on comparing the whole set of the N – H RDC measured for a biomolecule with the sets of these constants calculated from known structures deposited with protein databases.[343] Thus in the search for homologies in the PDB, Meiler *et al.*[344] created the DipoCoup programme taking into account all types of RDC and even pseudocontact shifts.

Delaglio *et al.*[345] have developed a method for the construction of a 3D structure of a protein molecule by means of the search for similar fragments in a specially composed sample from the PDB. The fragment should contain 7 amino acid residues. Comparison is carried out against not only RDC but also chemical shifts taken from a special database for each fragment.

Instead of comparing the whole set of N – H RDC, one can search through databases for structural motifs (fragments) containing, for example, $\sim 20$ amino acid residues using image recognition techniques. Andrec *et al.*[346] intend to compose a library of structural motifs on the basis of available protein databases for use in RDC-based studies. By applying special fitting criteria (for example, self-consistency of overlapping fragments) for each set of experimental data, one can restrict consecutively the fragment retrieval.[347]

It is possible to compose a large number of such protocols for the design of a protein structure based on homologies and RDC. However, the information contained in protein and nucleic acid databases can also be used in a different procedure for determining the structures of biomolecules.

The crystal structures of many proteins and nucleic acids have now been established with high accuracy. Analysis has shown that only a few dihedral angles determining the molecular conformation are possible in these structures for energy reasons. This conformational information accumulated as a database can be used for determining the structures of new biomolecules.

In order to select the conformations that are physically realised and described in databases from amongst the possible conformations proposed by the procedure of structure adjustment in experimental results, an additional pseudo-potential is introduced into the energy function in the following form:

$$E_{DB}(i) = -k_{DB}(\ln P_i),$$

where $k_{DB}$ is a scale coefficient or a force constant and $P_i$ is the partial probability that this parameter falls in the allowed range of values.

The introduction of this pseudo-potential into the energy function restricts the possible range of dihedral angles and accelerates the convergence during the iteration procedure.

Similar potentials with one-, two-, three- and four-dimensional potential surfaces describing the degree of likelihood of various combinations of torsion angles have been reported.[348 – 355]

The use of such potentials for minimisation is especially efficient for the adjustment of parameters determined from the spectra of oriented specimens, for example, dipolar couplings. In particular, the introduction of this type of potential for describing the positional coupling between pairs of bases in nucleic acids has sharply increased the accuracy of DNA structures determined in NOE-based experiments and using RDC.[356]

## 5. *De novo* structure determination using only residual dipolar couplings

At present, the 3D structure of a biomolecule is commonly determined using NOE. The new NMR techniques developed in recent years are usually considered as auxiliary and are used in those cases where the number of NOE signals is too small. However, the new methods are gradually becoming self-sufficient (without using the NOE data) methods for structure determination.

The method for determining the structure of a biomolecule with the use of only RDC may have been all-sufficient if it was not for the ambiguity of RDC interpretation caused by the fourfold degeneracy and the ambiguity of determination of the order tensor orientation.

An attempt at using only orientation parameters in structural studies was undertaken by Tjandra.[357] Subsequently, in a study of cytochrome c′ (although using several methods), Hus *et al.*[358] developed an algorithm that allows the structures of any paramagnetic proteins and complexes to be determined without resorting to NOE data. The algorithm employs the geometric information contained in paramagnetic pseudocontact shifts [158] and cross-correlated Cuire spin – dipole electron – nuclear relaxation [153, 359] as well as the information contained in RDC (only N – H RDC were used). This algorithm underlies the SCULPTOR programme (Structure Calculation using Long-Range, Paramagnetic, Tensorial and Orientational Restraints) meant for determining the structures of biomolecules using only orientation-

dependent data. This algorithm is also applicable to many metal-containing molecules.[143]

An algorithm for determining protein backbone structure on the basis of only RDC has been proposed.[46] The authors have defined it as MECCANO (Molecular Engineering Calculations using Coherent Association of Nonaveraged Orientations). This is a key study which demonstrates the real possibility of determining the structure of a biomolecule using the set of RDC alone. To solve the ambiguity problem, two sets of RDC were determined for the molecule in the LC media with different orientations and the self-coordination procedure was carried out for all RDC and each peptide plane.

The ambiguity of interpretation of the RDC determined in one experiment has been eliminated by invoking additional information.[345] In particular, a new algorithm for the protein folding was proposed, which can be used to attain self-consistency of all sorts of protein fragments during the search and screening of the structures of protein fragments in the database.[347]

Yet another algorithm using only RDC and the changes in the chemical shifts of carbonyl carbon atoms has been proposed.[288]

The all-sufficiency of RDC for the structural studies of biological macromolecules by NMR spectroscopy has been proven.[360]

## VII. The use of NMR spectroscopy for the study of the structures of biomolecules

As has been noted in the Introduction, the X-ray and NMR techniques are based on different kinds of physical phenomena and the structural data obtained by these methods supplement each other. X-Ray diffraction analysis produces a three-dimensional map of the electron density of the molecule, which is used to identify the relative positions of atoms (each experimental point on the map bears information on all atoms). An NMR study gives the distances and the angles between some pairs of atoms, which are used to reproduce a set of non-contradictory three-dimensional structures, and each experimental point bears information on the distance between some pair of atoms, or the size of some angle, or on the orientation of some bond vector with respect to a specified direction. In X-ray diffraction studies, hydrogen atoms almost cannot be seen, whereas in the case of NMR, they provide the main bulk of information. In turn, NMR spectroscopy virtually does not determine oxygen and sulfur. Whereas solution of the phase problem by X-ray requires the presence of a heavy atom in the molecule, in the case of NMR, isotope enrichment is needed to replace the non-magnetic $^{12}C$ and quadrupole $^{14}N$ nuclei by magnetic isotopes, $^{13}C$ and $^{15}N$, respectively. An X-ray study gives rise to a static (frozen) structure of a biomolecule containing no mobile groups, while NMR allows one to study biochemical processes and the behaviour of flexible groups.

The main achievements and the potential of NMR spectroscopy of weakly aligned molecules in the research into various classes of biomolecules are briefly considered below.

### 1. Proteins

An advantage of NMR procedures for aligned specimens is the fact that information on the long-range order can be derived from dipolar coupling parameters. This provides a substantial increase in the accuracy of determination of the molecular structures of proteins and protein complexes determined by other NMR techniques.

The number of measured RDC characterising the polypeptide chain is great. It can be markedly increased by replacing non-magnetic nuclei by magnetic isotopes. The efficiency and simplicity of including RDC into the standard set of parameters for determining the 3D structures of not only biomolecules but also their complexes have been convincingly demonstrated in many publications.[361–367] For example, Gronwald *et al.*[368] could determine the structure of the protein kinase Byr2 only after adding RDC to NOE data.

The structure of a very important anti-HIV protein, cyanovirin-N, which binds to the glucoproteins of the viral shell and inhibits their activity, has been studied.[369] This protein can become a key point in the development of a new anti-AIDS strategy. The $^{1}H$, $^{13}C$, $^{15}N$ NMR signals were assigned and a twisted folding was confirmed using RDC analysis;[370] this is significant for further research into the structure–activity relationship for this protein.

The structure of the signalling human Gα interacting protein (GAIP) in solution was determined in two different LC media.[371] Comparison with the crystal structure of the homologous rat RGS4 protein complexed to the α-subunit of the G protein showed the presence of equally aligned fragments which do not change after binding to Gα, indicating that these fragments are not involved in the Gα interaction. Comparative analysis of the structures of the bound and free proteins allowed the researchers to represent the binding geometry and the GTPase activation mechanism.

The structure of the DinI protein inhibiting the activity of the RecA protein from *Escherichia coli*, which initiates the repair or mutagenesis of DNA, was also determined in two different LC media (bicelles and phages).[366] This made it possible to identify the role of the C-terminal α-helix, which forms a negatively charged ridge on the globule; this mimics electrostatically the DNA phosphodiester backbone and thus it can compete with DNA for binding to RecA.

The structure of hen lysozyme was studied in particular detail in two bicellar LC phases with different orientations;[372] comparative analysis with and without RCD data was performed and the importance of RDC restraints has been demonstrated. The protein structure was determined more precisely in this study than this had been done by X-ray crystallography.

Using RDC, it was possible to study the thermal stability of the protein aporubredoxin in the native (metal-containing) and metal-free forms.[373] The native form of the protein proved to be fairly stable over a broad temperature range, while the metal-free form is susceptible to reversible thermal denaturation.

Analysis of the RDC for molecules of denatured proteins in a strained polyacrylamide gel showed that the topology of the native state of some structure elements is retained [374,375] and is insensitive to the replacement of hydrophobic residues in the amino acid sequence by polar residues of a similar size.[376]

### 2. Domain orientation in proteins

The molecules of many proteins consist of two or more packed independent domains. The type of biological process executed by multidomain proteins, for example, enzymic catalysis, often depends on the relative orientation of the domains. Therefore, one of the goals of structural studies is to determine the relative orientation of the domains. The crystal structures of these proteins may differ from the structures occurring in liquids, because the forces acting during crystal packing may change the relative orientation of the domains. By investigating multidomain proteins by liquid-state NMR, one can determine the domain orientation and its changes induced by different factors (*e.g.*, ligand binding or covalent interactions). However, usually, the number of observable domain–domain contacts is small. Therefore, only measurement of the orientation-dependent NMR parameters offers new opportunities for determining the orientation of domains or secondary structure elements for the molecules of non-rigid proteins.

The foregoing can be illustrated by the results of the investigation of a two-domain protein barley lectin.[377] First, it was demonstrated that the $^{15}N-^{1}H$ RDC are applicable for determination of the relative domain orientation in the fragment of this protein and the use of the SVD method (Singular Value Decomposition) was theoretically substantiated.[378] The orientation of each domain can be determined with a few RDC (five is plenty). The domain orientation and dynamics were studied in different LC media.[379] The domain orientation was found to differ from

that determined for a highly homologous protein in the crystal. Furthermore, differences in the order tensor values suggest that the two domains have different reorientational dynamics and can be considered as independently reoriented.

In a study[380] dealing with determination of the relative orientation of Zn-finger domains of the TFIIIA protein in the protein–DNA complex, RDC were also shown to be highly sensitive and reliable parameters. After inclusion of RDC, the spread in interdomain angles (thoroughly studied by X-ray, NMR and computer simulation) is reduced from 24 to 9°. However, in this case, too, the local structure of the domains should be known before; moreover its accuracy influences also the accuracy of RDC-based determination of the orientation.

The relative orientation of the monomer subunits in the cyanovirin-N domain-swapped dimer[381] and the full structure of its monomeric mutant[382] were established. This was done using a combination of SVD techniques[378] and calculations based on the molecular shape;[197] less than one-third of the total number of RDC was involved. A combination of 'simulated annealing', torsion angle dynamics and rigid-body minimisation proved to be even more successful for determining the full structure of this protein.[383] The combinations of these methods open up good prospects for fast determination of the full structure of homologous or chimeric proteins and structural analysis of mutations if the structure of the main protein is known.

The different orientation of domains in the ribosomal protein S4 Δ41 in the crystalline and liquid states has been demonstrated[384] using RDC measured in two different uniaxial LC phases formed by DMPC/DHPC bicelles and bacteriophage Pf1. It was shown that domain orientations determined in the isotropic medium using NOE and in two anisotropic media are identical and the helical subdomain is twisted by 45° around the long axis of the molecule with respect to the subdomain in the crystalline phase.

Orientation of paramagnetic protein metal complexes in the magnetic field has also been used to establish the relative orientation of domains.[144] For this purpose, it is possible, for example, to replace the diamagnetic $Ca^{2+}$ ion in calcium-containing proteins by a paramagnetic $Tb^{3+}$ ion. This replacement virtually does not influence the protein structure as a whole. It was shown that $Tb^{3+}$ is bound in only one of the domains and induces a substantial protein orientation. It is possible to measure not only domain–domain long-range pseudocontact shifts but also RDC.

If the crystal structure of a multidomain protein is known, the relative orientation of its domains in a liquid can be easily determined by measuring the RDC for only some easily identifiable residues incorporated in different domains.[385] In particular, this procedure was used to determine the conformation of the complex formed by the maltodextrin-binding MBP protein with β-cyclodextrin containing 370 residues; it was also demonstrated that the relative orientation of its domains in the crystal structure differs substantially (by $\sim 11°$) from that in the liquid, although the structure of each domain corresponds to that in the crystal. Full structure determination for this protein using RDC, NOE and vicinal constants[310] confirmed the domain orientation determined by a fast assay.[385] Measurements of the diffusion tensor anisotropy[386] also confirmed the relative orientation of the domains established using RDC and showed substantial deviations from X-ray data. Finally, the RDC procedure was used to carry out a systematic study of the relative orientation of the domains of free and β-cyclodextrin- and maltotriose-bound MBP proteins.[387] The domain orientations for apo-MBP and the MBP complex with maltotriose were found to coincide with the orientation in the crystalline phase, whereas in the case of the MBP–β-cyclodextrin complex, they were found to be different. The authors attributed this to the crystal packing effect.

The discrepancy of the conformations of multidomain proteins determined in the crystal phase and in solution appears to be the rule rather than an exception. Therefore, NMR spectroscopy of oriented biomolecules can be considered to be a unique method for investigating their conformations in solution. This is especially important in the studies of enzymes and in the search for new drugs.

The $^{13}C^{\alpha}-^{13}C^{\beta}$ RDC in $^{15}N$-, $^{13}C$- and $^{2}H$-enriched proteins can also be used to determine domain orientation.[245] It should be noted that both intra- and interdomain structures of one of the studied proteins, namely, the maltose-bound protein in the liquid and crystalline states were found to be similar.

The RDC method was used to demonstrate the unique possibility of fast identification of medium- and large-scale interdomain motions in modular proteins.[388] Actually, RDC provide information on the relative motions of domains. By using these data, together with structural and relaxation data, one can calculate the amplitudes and time scales of these motions.

In a study of relative orientation of Bacteriophage lysozyme T4 domains using RDC, the averaged conformation of this protein in solution was determined.[389] The lysozyme T4 hydrolyses the β(1-4)-glycoside bond in the peptideglycan substrate, which is bound to the protein in a gap between two domains. The crystal structures of various mutants of this enzyme are characterised by different opening angles, which implies a substantial mobility of the domains. In the native enzyme structure in solution, the opening angle was found to be 17° greater than that in the crystal phase.

A significant deviation (25°) from the crystal structure was found for $Ca^{2+}$-calmodulin.[390]

The applicability of RDC to investigations of the dependence of functional properties on minor structural differences in homologous series of proteins was demonstrated in relation to the variation of the calmodulin-binding geometry for five different calmodulin-regulated kinases.[391]

Yet another example is determination of the relative orientation of the subdomains of the heparin-binding domain of the vascular endothelium growth factor $VEGF_{165}$; this became possible only after RDC had been included in the analysis.[392]

Using RDC, it is also possible to study the dynamics of flexible fragments in non-rigid proteins; the RDC values averaged over various modes of motion are collated with the experimental values. For example, the mutual arrangement and the mobility of three helices in the $p8^{MTCP1}$ protein molecule were determined in this way.[393]

The MODULE graphical interactive programme[313] mentioned above is convenient for the simulation of multidomain macromolecules. This programme allows each domain to be described by its own set of RDC. By varying domain orientations, one finds the best agreement of the sets of RDC with experimental data.

The methods developed for investigating the relative orientation and the dynamics of separate protein domains can also be used to study the structure and dynamics of individual rigid fragments of various molecular systems, for example, in non-rigid proteins, oligosaccharides and molecular complexes. Evidently, they would also be useful in the study of non-rigid organic molecules and liquid crystals.[394]

## 3. Protein–ligand complexes

Determination of the geometry of protein binding to the ligand is important as regards elucidation of the structure–biological activity relationships for proteins.[395] The spatial structure of the protein–ligand complex needs to be known for the development of new pharmaceuticals, because this allows the search for a new drug to be target-directed, *i.e.*, aimed at providing the structural and charge complementarity of the required ligand to the active cavity of the protein.

The NMR spectrum of the ligand is usually rather simple and the signals can be readily identified even without isotope enrichment. This is taken into account in the most widely used scheme of investigation of protein complexes which implies the use of $^{15}N/^{13}C$-enriched proteins and ligands with natural isotope abundance. Hence, it is possible to use special heteronuclear filtration

and editing methods in experiments for resonance correlations of only the protein, only the ligand, or only through-space protein–ligand couplings.[24, 396]

Study of the change in the ligand conformation upon binding in protein complexes is often difficult, especially in the complexes with low affinity. In this case, the ligand cannot be observed separately in the free and bound states in the spectrum due to fast exchange. Therefore, it was proposed to use the dipolar couplings of the ligand in the oriented complex. Using the protein complex of a trisaccharide as an example, the possibility of measuring the orientation parameters of a free and bound ligand and determining its conformation has been demonstrated.[397]

The isotope enrichment of proteins, unlike that of small organic molecules, is currently a relatively easy task. In the development of new drugs, the use of traditional NMR procedures is complicated by the fact that the ligand becomes a kind of invisible. The structure of a non-enriched ligand complexed with an enriched protein can be determined by measuring the RDC of the ligand in an oriented complex.[398] The structure of the whole complex can be determined with the highest accuracy if RDC are also used in determining the protein structure.

Algorithms for determination of the geometry of complexes making use of orientation parameters of both the ligand and the protein and taking into account the molecular symmetry have been developed.[399] The geometry of the α-methyl-D-mannoside and its complex with a 53 kDa protein has been found using this type of algorithm.[400]

The orientation of a ligand (a transducin peptide fragment) weakly bound to light-activated membrane rhodopsin has been studied.[401, 402] The rhodopsin-containing discs are oriented similarly to membrane fragments. The possibility of investigating the binding dynamics and ligand orientation was shown.

The measurements of the RDC for a $^{13}$C-enriched $Gb_3$-oligosaccaride[403] made it possible to locate its binding site to the protein verotoxin 1. The exchange between the bound and free states of the trisaccharide was fast on the NMR time scale and the measured RDC values were weight-averaged over these states. One of the three possible binding sites was reliably identified on the basis of simulation results (X-ray). The structural changes induced by the ligand have been observed for MBP binding to β-cyclodextrin.[387]

## 4. Protein–protein complexes

In recent years, NMR studies of multiprotein complexes and protein–protein interactions, like the studies of multidomain proteins, have become nearly routine. These studies make use of intermolecular NOE, cross-saturation, paramagnetic methods, analysis of chemical shift perturbations, analysis of relaxation parameters and so on. It is obvious that combination of RDC-based procedures with TROSY is the most promising for relative positioning of the proteins in the complex.[404]

Most of the above methods for determining the relative orientation of domains are also applicable to studies of various types of complexes (protein–protein, protein–ligand, protein–nucleic acid).

If the structure of each protein forming the complex is known, one can determine rapidly and accurately the structure of the whole complex using RDC and intermolecular NOE signals. This was demonstrated[405] in relation to a protein–protein 40 kDa complex. In the algorithm described, dockings of the proteins incorporated in the complex are considered rigid. The use of this method markedly accelerates determination of the structures of complexes by NMR. The same method was used to determine the structures of two protein–protein complexes from the phosphotransferase system of *E. coli* sugars.[406, 407]

McCoy and Wyss[408] were able to solve the problem of protein docking using only the information contained in the RDC and chemical shift perturbations.

To determine the relative orientation of domains or molecules in calcium-containing proteins and complexes, one can use a facile procedure which does not require aligning solvents,[143] orientation being induced by replacement of the $Ca^{2+}$ ions by lanthanide ions having a similar ionic radius.

The presence of a paramagnetic label in one monomer in symmetric complexes, for example, in homodimers, eliminates the symmetry-related degeneracy and the overlap in the spectra and allows one to determine the RDC and pseudocontact shifts specific for each of the monomers.[409]

## 5. Protein side chain conformations

For the present moment, the attention of researchers focuses on the 3D structure of the protein backbone. However, significant items are determined by the side chain conformation; in some cases, they play a crucial role in the functioning of a biomolecule.

The assignment of signals corresponding to the side chain nuclei to particular members of the protein sequence is the most difficult and lengthy phase in determination of the protein structures by traditional methods based on NOE.

The dipolar coupling constants proved to be very convenient for investigating the side chain conformations and dynamic behaviour. The RDC values for side chains can also be useful for determining the backbone structure as additional structure constraints. A lot of various RDC can be measured (homonuclear or heteronuclear, between either directly linked or remote nuclei). The $C^{\beta}-H^{\beta}$ (see studies[259, 260, 264]) and $C^{\alpha}-C^{\beta}$ (see studies[244, 245, 259]) RDC as well as C–C,[265, 410] C–H[266, 410] and H–H[264, 265] RDC for methyl groups are used most often.

Even the earliest measurements of RDC in side chains showed good agreement with the values predicted on the basis of the 3D structure of the protein backbone.[411] The subsequent studies confirmed the high information content of these constants not only in protein structure determination but also in a study of side chain dynamics.

In particular, the torsion angles $\chi_1$ and their dynamic behaviour in $^{13}$C-enriched partially deuterated proteins were determined with high accuracy using the direct $^1H^{\beta}-^{13}C^{\beta}$ RDC in CH and $CH_2$ groups[260] (which can be interpreted rather easily, especially using the force potential of torsion angle databases).[355]

The possibility of precise determination of the orientation of functional groups in asparagine and glutamine in the protein side chains was shown.[276] Determination of the arrangement of these groups is especially important because they are involved in specific hydrogen bonds.

The information concerning the positions of the methylene and methyl groups is important in determining the positions of the hydrophobic side chains in the protein.

A simple and efficient method for the measurement of RDC in methyl and methylene groups has been developed based on the 3D version of the well-known [$^1$H–$^{13}$C] CT HSQC experiment.[411] The only drawback of this method is the fact that the sum of only two C–H RDC in the methylene group is determined.

Another method specially developed on the basis of the $S^3E$ type HSQC experiment is more convenient for investigating the methylene groups of $^{13}$C-enriched proteins.[264] A new pulse sequence, SPITZE–HSQC (Spin State Selective Zero Overlap HSQC), made it possible to measure simultaneously the C–H and H–H RDC. The glycine methylene groups in ubiquitin were used as an example to demonstrate the unique possibilities of this method for fast determination of the magnitudes and signs for all three RDC in the methylene group, *D*(H1,H2), *D*(C,H1) and *D*(C,H2), in the spectra with minimal spectral overlap and a better signal-to-noise ratio compared to that in the HSQC experiment.

The first RDC-based systematic study of the methyl groups in the proteins showed a similarity in their geometry; the average value for the C–C–H angle (110.9°) proved to be greater than 109.5° (see Ref. 410).

Study of the methyl groups is complicated by two circumstances. First, a fourfold increase in the number of resonance signals after switching-off the $^1$H decoupling may result in peak superposition and deteriorate the resolution. The specially devel-

oped IPAP–CT HSQC type protocol permits one to separate the multiplet components and markedly facilitates the measurement of RDC in the methyl groups.[266] Second, the rotation of the methyl groups around a threefold symmetry axis decreases RDC by a factor of $(3\cos^2\theta-1)$ ($\theta$ is the angle between the symmetry axis of the group and the C–H bond), *i.e.*, by almost 70%. The geminal H–H constants for which both the magnitude and the sign can be measured in a single experiment are more advantageous in this respect.[265] The measurements of these constants are especially reasonable if deuterium substitution is used, *i.e.*, the protein molecule is perdeuterated except for the methyl groups.

The measurement of the quadrupole splittings in the spectra of aligned proteins with partially deuterated and $^{13}$C-enriched methyl groups ($^{13}CH_2D$) is yet another promising line of investigations of these groups. The high correlation observed between quadrupole splittings and the $^{13}C^{Me}-^{13}C$ RDC has been interpreted[299] as being indicative of the similarity of all the methyl groups in proteins. The researchers cited calculated the average QCC for the methyl groups in the free rotation approximation taking the angle between the rotation axis and the axis of the electric field gradient tensor to be 109.5°. They proposed that the resulting value ($167\pm1$ kHz) should be used subsequently in interpreting the $^2$H relaxation data in the protein methyl groups.

An isotope effect was observed on deuterium substitution at the methyl groups. The spin–spin coupling constants for directly linked C–H and C–C nuclei decreased with an increase in the number of deuterium atoms in the group.[412] The very weak isotope effect also had an influence on the methyl group geometry: the C–C–H angle slightly changed (by fractions of a degree).

## 6. Carbohydrates

Oligosaccharides are important compounds participating in many processes that take place in a cell. They occur in biological systems either in the free state in solution or as parts of proteins (glycoproteins) or lipids (glycolipids) on the cell surface. They are used more and more often in the development of drugs.

The cell functions of oligosaccharides are directly related to their structures, conformations and dynamic behaviour. Knowing these structures both in the free state in solution and in the bound state is an important prerequisite for understanding of the major factors that control the recognition of oligosaccharides by specific cell receptors.

At present, multidimensional high-resolution NMR spectroscopy in combination with molecular simulation is the major method used to determine the structures of carbohydrates. The advent of the procedures based on the measurement of RDC gave a renewed impetus to these studies.

The use of RDC for structure determination and conformational analysis of oligosaccharides was started in 1998 simultaneously by three research groups, headed by Prestegard,[413] Widmalm[414] and Homans[415,416] and, later, by a group headed by Bush.[417] Actually, by that time, the methods based on the measurements of RDC and CSA have already been used to determine the structures of oriented glycosides, glycolipids and other carbohydrates,[418–427] but this line of research has been generally recognised only after successful use of these methods in the studies of biomolecules.

Bolon and Prestegard[413] proposed a simple COSY experiment for measurement of the $^1$H–$^1$H RDC in oriented oligosaccharides and demonstrated its efficiency in relation to a trisaccharide. The high information content of these new data for the conformational analysis of oligosaccharides was used by Kiddle and Homans[415] to refine the conformation of another trisaccharide, studied previously using NOE and $^{13}$C–$^1$H and $^1$H–$^1$H vicinal constants. The $^{13}$C–$^1$H RDC were measured for a $^{13}$C-enriched trisaccharide dissolved in a bicellar LC medium.

Rundlof *et al.*[414] used C–H dipolar coupling to study the conformation of a tetrasaccharide, lacto-N-neotetrose (LNnT), from human milk with natural $^{13}$C abundance. By using energy minimisation, experimental data were adjusted to the molecular model generated by the molecular dynamics method. Good correspondence was observed for a conformation in which the linear tetrasaccharide was described as a uniaxial cylinder. To confirm this model, LNnT was also studied in two other lyotropic LC media with conformation averaging.[428] Using the Monte Carlo and molecular dynamics methods, various molecular structures were generated, the orientation parameters and RDC were calculated for them and compared with experimental values. It was shown that, apart from the expected difference between the orientational characteristics of the tetrasaccharide LNnT in two different media, a slight difference between the conformations was found.

Martin-Pastor and Bush[417] studied the conformations of six other natural oligosaccharides isolated from human milk; they used molecular dynamics and NMR spectroscopy based on NOE and RDC measurements for molecules oriented in bicelles. The inclusion of information contained in RDC markedly increased the reliability and the accuracy of results. The conformational study of these oligosaccharides and the pentasaccharides LNF-2 and LNF-3 (LNF is lacto-N-fucopentose) was continued on the basis of analysis of C–H RDC.[418,429] (It should be noted that the conformations of these complex oligosaccharides were studied without isotope enrichment.) To this end, a strategy was developed for fast comparison and calculation of the difference between the experimental and calculated RDC for a large number of structures generated during computer simulation.[429] Two different consecutive algorithms were used to analyse thousands of generated structures in the space of dihedral angles corresponding to the high and low values of the energy to be minimised. This two-stage strategy proved to be very efficient for a conformational study of the carbohydrate antigenes of blood groups Lewis X and Lewis A trisaccharide[429] and two synthetic trisaccharides of blood groups A and B and tetrasaccharide alditol of blood group A.[430]

These studies demonstrated the possibility of measuring the heteronuclear C–H RDC in oligosaccharides with a natural $^{13}$C abundance. The 2D *J*-modulation HSQC procedures developed recently allow one to measure them with a higher accuracy.[431]

The paramount importance of the $^1$H–$^1$H RDC for determination of the 3D structures of oligosaccharides was noted in another study.[432] The authors developed a new procedure for the measurement of RDC and determined the structure of a partially deuterated decasaccharide also using traditional NMR methods. The decasaccharide structure could be determined only when direct ($^1D_{CH}$) and long-range ($D_{HH}$) dipolar couplings have been included in the analysis. A method for determination of the conformation distribution function for disaccharides in the LC medium has been proposed.[433]

A new strategy for conformational analysis of flexible oligosaccharides based on the use of RDC has been reported.[434] This strategy implies analysis of order matrices for each of the sugar rings. The resultant RDC values averaged over internal motions are used to estimate the amplitude of these motions. These data are well combined with the results of conventional relaxation methods.

The $^{13}$C–$^1$H and $^1$H–$^1$H dipolar coupling constants were used successfully to determine the structure and dynamics of simple sugars in solution.[435] The structures of saccharose (disaccharide) and raffinose (trisaccharide) in solution were found to differ substantially from their structures in the crystal phase. For example, the disaccharide can exist with four different orientations of monosaccharides with respect to each other (for example, like protein domains[399]). However, only one orientation was observed, possibly due to the short bridging unit. Meanwhile, in the case of the galactose fragment of raffinose, clear-cut dynamic behaviour was found, which markedly decreased the degree of orientation of this fragment compared to the other two fragments.

The quadrupole couplings in deuterated carbohydrates were used to determine the structure and dynamic behaviour of $^{13}$C/$^2$H-

enriched methyl-β-D-glucopyranoside.[300] The quadrupole constants of endocyclic deuterons proved to be very uniform; therefore, a single average value of this constant can be used in relaxation studies of glycan fragments in glycoproteins and glycopeptides.

Comparison of the results of structure calculations of the pentasaccharide constructed from D-mannose residues by molecular dynamics methods with the structures obtained using relaxation data and RDC has shown[328] that, although the results obtained using RDC are in good agreement with both the model of averaging over the whole possible ensemble of conformations and with one structure of this ensemble, it is still more legitimate to take into account all the conformational motions of the molecule. The relaxation data can be interpreted only in the ensemble model. The researchers also carried out quantitative conformational analysis of the core region of N-glycans in water using RDC and molecular dynamics methods.[436]

In determining the structure of a flexible heptasaccharide,[437] success was achieved owing to the combination of traditional NMR experiments with the measurement of direct ($^1D_{CH}$) and long-range ($D_{HH}$) dipolar couplings.

The 3D structure of the enterobacterial cyclic antigen containing four repeating trisaccharide units was also determined by combined use of RDC, data of NOE experiments and vicinal $^3J_{CH}$ constants.[438]

Using only RDC, Freedberg was able to study rather easily the deformation of the fructofuranosyl ring of sucrose.[439]

Thus, the NMR spectroscopy techniques for aligned molecules open up new, unique opportunities for research into the structures and for conformational analysis of both simple sugars and complex oligosaccharides and their protein complexes in solution.[440]

## 7. Nucleic acids

Study of the structure of nucleic acids by conventional NMR spectroscopy methods is much more difficult than the study of protein structures, due to a number of specific features of biomolecules of this class. First, there is a relatively low density of protons in nucleic acids, and, correspondingly, small number of NOE constraints. Second, due to the non-globular structure of the DNA molecule, the spectra contain no NOE signals between different sections of the chain but contain only signals due to nucleotides that adjoin each other in the sequence. Even in the RNA molecule, which is less linear and has a clear-cut tertiary structure, the number of internucleotide contacts is very little. Third, the dispersion of nuclear chemical shifts is insignificant; therefore, only a relatively small number of resonance signals can be resolved. Finally, it is necessary to determine five (rather than two as in proteins) backbone angles per structural unit.

The foregoing implies that global characteristics of nucleic acid structures such as helical twisting, helix step and torsion angles in the chain are difficult to determine, although a number of local conformation parameters (for example, the conformations of carbohydrate chains and the torsion angles of the glycoside bonds) can be established rather accurately. The RDC and CSA measurements in weakly aligned nucleic acids provide new, in principle, opportunities for structural studies.[218,441–444]

Almost any of the alignment methods mentioned above is applicable to nucleic acids; the use of filamentous phages is considered to be the method of choice.[220]

The attempts to measure the RDC for nucleic acids have been undertaken even in the first experiments with weakly aligned biomolecules in bicellar systems[45] and in LC suspensions of phages;[219] however, Bayer *et al.*,[445] who used the $^1H-^{13}C$ and $^1H-^{15}N$ RDC in the protein–RNA complex, were able to considerably improve the resolution in the determination of the structure of a nucleic acid.

The structure of the DNA dodecamer $d(CGCGAATTCGCG)_2$ has been studied in a bicellar medium.[446] In addition to the 162 NOE responses, 198 $^{13}C-^1H$, 10 $^{15}N-^1H$ and 200 $^1H-^1H$ RDC were determined. It was shown that the dodecamer structure could in principle be determined using only RDC.

The same dodecamer was used to demonstrate a procedure of direct determination of the parameters of nucleic acid helices using only RDC[447] based on the analysis of their order matrix. An advantage of this method is that it does not require the knowledge of the refined molecular structure.

A method for fast determination of a stoichiometry in homomultimeric nucleic acid complexes by analysis of RDC induced by a magnetic field has been proposed.[448] Depending on the number of monomer units in the multimer, the components of its diamagnetic susceptibility vary. Having found or calculated the diamagnetic susceptibility of the monomer, one can determine the diamagnetic susceptibility of the multimer from the RDC values and, hence, the number of monomer units in the multimer. This does not require the knowledge of any structural information or even assignment of the resonance signals.

Using RDC, it is possible to determine the relative lengths of hydrogen bonds[248] and their orientation in DNA.[449] A systematic study of two DNA duplexes has shown[450] that inclusion of RDC into the set of NMR parameters used in structural studies substantially increases the accuracy of determination of both the local and global structures of duplexes.

The use of RDC allows fast determination of the relative positions of known elements of RNA secondary structure.[441] For example, a few additional $^{15}N-^1H$ RDC proved sufficient for determining the global structure of transport RNA[451] and a hammerhead ribozyme,[452] in which the local structures of helical regions had already been known.

The spatial packing of nucleic acid homodimers with $C_2$ symmetry has been established unambiguously using two sets of RDC measured independently.[453] In this case, two independently oriented states of DNA quadruplexes were obtained using a bacteriophage solution as an orienting medium and on exposure to a strong magnetic field of the NMR spectrometer (800 MHz). The developed procedure was tested in relation to a $^{15}N/^{13}C$-enriched deoxynucleotide sequence d(GGGTTCAGG), which dimerises in solution to give a quadruplex and forms a pair of G(CA) triads and GGGG tetrads.

The measurement of RDC is a rather fast and facile method for detecting the specific conformational changes taking place in RNA upon binding to drugs. This was demonstrated using the conformational changes induced by binding of an aminoglycoside antibiotic to the RNA chain.[454]

Measurement of the $^1H-^{31}P$ RDC has been proposed for determining the geometry of binding of nucleic acids to the phosphodiester backbone.[455,456] The efficiency of measurement of the vicinal H3′–P RDC (NOESY version of the CT COSY experiment) was demonstrated in the study of the $d(CGCGAATTCGCG)_2$ dodecamer mentioned above.[456] The scalar and dipolar H3′–P constants were measured in isotropic and LC states both without and with $^{31}P$ decoupling. Yet another variant of using through-space $^1H-^{31}P$ RDC has been developed on the basis of the TOCSY experiment.[455] In this case, one should measure the $^{31}P-^1H$ RDC with base protons H6 and H8. The efficiency of this procedure was demonstrated by determining the structure of HIV-2 TAR 30-mer RNA in a solution of bacteriophage Pf1.

Wu *et al.*[289] demonstrated that the great $^{31}P$ CSA values observed in aligned samples of nucleic acids can be efficiently used to determine the orientation of the phosphodiester groups relative to the molecular alignment tensor.

When in the NMR spectrum many proton resonances overlap, and, moreover, protons have many coupling partners (as, for example, the protons of the RNA ribose), it can be difficult to measure the $^1H-^1H$ constants. In these cases, one can use selectively fluorinated compounds. In particular, an E COSY-based procedure with X-filtration has been developed for these measurements. The resulting $^1H-^{19}F$ RDC were used successfully

to determine the structure of 2′-fluorolabelled 21-mer RNA hairpin.[457]

The use of RDC simplifies the study of the structures of protein–DNA [380, 458–460] and protein–RNA complexes.[445] In particular, combined use of the measured $J_{NN}$ constants through a hydrogen bond and the RDC induced due to alignment of the RNA molecule by the magnetic field made it possible to determine the geometry of binding of the 16-mer HTLV-I Rex peptide to the 30-mer RNA aptamer.[461]

A new algorithm for fast assignment of NMR signals and simultaneous determination of the tertiary RNA structure has been developed;[462] in this case, the time-consuming NOE-based sequential assignment step is completely excluded.

Determination of components of the alignment tensor is much more complicated in the case of nucleic acids because the column diagram approach (see Ref. 318) based on the rather uniform distribution in space of the directions of internuclear vectors in proteins is inapplicable in this case. The mapping strategy developed by Warren and Moore [322] may be a good substitute for nucleic acids. The authors used this strategy to determine the structure of the sarcin-ricin RNA loop.[463]

Thus, the use of RDC markedly extends the scope of modern NMR techniques as applied to structure determination for any type of nucleic acids and their complexes.[463–466]

### 8. Internal dynamics

The study of dynamic processes in macromolecules represents yet another promising application of NMR spectroscopy of aligned biomolecules. Unlike the spin relaxation or chemical exchange methods used conventionally, this method provides information on the dynamic behaviour on a time scale typical of biologically significant motions (picoseconds to seconds).[467, 468]

We have already mentioned studies in which RDC have been used to examine the dynamics of protein side chains,[260] domain orientation [379, 385] and the conformational mobility in carbohydrates.[328, 434, 436] It is usually assumed that internal motions in the biomolecules are either missing or negligible or that they are similar and have an axial symmetry.

However, in reality, internal motions of various types always occur in biomolecules and neglecting them should undoubtedly influence the accuracy of structure determination based on RDC. Indeed, when the biomolecule contains non-rigid fragments, RDC are determined by averaging not only over reorientations of a biomolecule as a rigid particle but also over internal motion. The additional averaging gives rise to deviation of the RDC values of flexible fragments from the values expected for the given degree of orientation of the biomolecule as a whole and deteriorates the accuracy of the determined structure. These deviations can serve as a measure of mobility of the given fragment and a criterion for choosing a particular model of motion.

For separating the structural information in RDC from dynamic information, a special strategy based on analysis of the order matrix was developed; this strategy allows one to distinguish the most flexible fragments of a biomolecule. The internal motions are especially pronounced in carbohydrates; naturally, RDC were first used to study their dynamics.[434, 435] The potential of this method was also demonstrated for proteins in analysis of the internal dynamics of ubiquitin [469] and the functional dynamics of the active state of ribonucleotide binase.[470] The most extensive prospects are opened for investigations of the RNA internal dynamics and the dynamics of RNA binding to proteins.[471]

Another strategy for extracting dynamic information from RDC [337, 472] is based on a large number (at least five) of independent measurements in different orienting media.

## VIII. Conclusion

The information considered in this review demonstrates the vigorous development of the application of NMR spectroscopy to biomolecules.‡

An advantage of NMR spectroscopy for weakly aligned biomolecules is the fact that, along with structure determination, it allows one to determine rather promptly the relative arrangement of individual groups during functioning of the biomolecule without the necessity of repeated labour-consuming determination of the structures of all current states of the dynamic system.§ Therefore, this method can be regarded as most promising in the studies of the structure–biological function relationship for biomolecules.

The development of NMR spectroscopy of weakly aligned biomolecules is especially topical for biology and medicine. Apart from the possibility of determining the tertiary and quaternary structures and the dynamics of biomolecules and their complexes, this technique is suitable for: quick diagnostics of molecules in solutions and investigation of their functions; screening of drugs; fast elucidation of structure–property relationships for mutant or modified biomolecules; spectroscopic studies of the biomolecules of various tissues and organs *in vivo* either in the oriented state or in exchange between oriented and non-oriented states; visualisation of the structures of various biological objects using the degree of alignment of biomolecules or their fragments as the contrast factor.[473–477]

The combination of this method with other methods (for example, with small-angle X-ray scattering in solution,[478] diffraction methods and advanced structural analysis using synchrotron radiation [479]) opens up new opportunities for research into biomolecules.

The Tables presented in the Supplement summarise characteristics of the systems used for weak alignment of biomolecules and the results of studies using NMR of weakly aligned molecules for compounds deposited with the PDB and the BMRB (Biological Magnetic Resonance Bank).[480]

The authors are grateful to Ph.Ds S S Mamyan and G A Panosyan for stimulation and support of the work, and their discussion and valuable remarks. This review was written with financial support from the U.S. Civilian Research and Development Foundation (CRDF, Grant CGP 2002 AC1-2304-YE-02).

## IX. Supplement

Table. S.1 presents brief characteristics of the systems used for weak alignment of biomolecules. Table S.2 lists the structures of biomolecules deposited with the PDB and the BMRB which were determined using dipolar coupling parameters. Note that the term 'structure' is used to mean both a set (ensemble) of structures with close values of minimisation criteria in the adjustment programmes used and a single (the most probable) of these structures. Some identical biomolecules are recorded in the databases under different numbers. The structures of these biomolecules were determined by different groups of researchers or by the same group but based on different assumptions.

‡ K Wütrich (Switzerland) was awarded a Nobel Prize in 2002 for a series of studies dealing with NMR spectroscopy of biomolecules.

§ For the series of fundamental research dealing with the development of the procedures for the measurement of residual dipolar couplings of weakly aligned molecules, three outstanding scientists, A Bax, A Bothner-By and J H Prestegard, were awarded the Laukien Prize at the 43rd Experimental NMR Conference, Pacific Grove, CA, 2002.

**Table S.1**. Media used for weak alignment of biomolecules (the alignment direction is indicated relative to the magnetic field direction).

| System | | | | Biomolecule | Aligning factor | Alignment direction | Ref. |
|---|---|---|---|---|---|---|---|
| composition | description | advantages | drawbacks | | | | |
| Without orienting medium | isotropic solution of biomolecules | no special orienting media are required | very weak alignment | nucleic acids, paramagnetic proteins | magnetic field of the spectrometer | ∥ | 42, 43, 140, 142 – 147, 149, 150, 152, 161 |
| DMPC/DHPC, DMPC/CHAPSO | isotropic solution of discs with encapsulated protein | the same | – | membrane proteins | the same | ⊥ | 183, 184 |
| DMPC/DHPC | bicellar LC, flat discs, 4 – 5 nm thick | easily modifiable | hydrolysis at low pH | proteins, nucleic acids, carbohydrates | bicelles with negative $\Delta\chi$ | ⊥ | 174, 175, 179, 180, 182, 185 – 190, 196, 200 |
| DMPC/DHPC/CTAB | the same | stable, positively charged | – | proteins | the same | ⊥ | 192, 196, 200 |
| DMPC/DHPC/SDS | " | stable, negatively charged | – | proteins, nucleic acids | " | ⊥ | 192, 196 |
| Dialkyl phospholipids | " | broad pH range, stable against hydrolysis | – | proteins, nucleic acids, carbohydrates | " | ⊥ | 196 |
| DMPC/CHAPSO | " | – | – | proteins, glycolipids | " | ⊥ | 175, 179, 203 |
| DLPC/CHAPSO | " | stable over a broad temperature range | | proteins | " | ⊥ | 193, 203 |
| DIODPC/CHAPSO | " | stable against hydrolysis, stable at low pH | | " | " | ⊥ | 194, 203 |
| Acid bicelles DMPG/DHPC, DMPS/DHPC, DMPC/DHPC/DMPX | " | negatively charged, analogue of natural membranes | – | membrane proteins | | ⊥ | 182, 195 |
| Bicelles doped with lanthanide ions DMPC/DHPC + $Ln^{3+}$ | " | – | – | proteins | bicelles with positive $\Delta\chi$ | ∥ | 204 – 207 |
| DMPC/DHPC/DMPG + DMPE – DTPA (doped with $Yb^{3+}$) | " | – | – | " | the same | ∥ | 206 |
| DBPC/DHPC | bicellar LC, one lipid chain contains a biphenyl group | – | – | carbohydrates | " | ∥ | 211, 212 |
| DBBPC/DHPC | the same | – | – | " | " | ∥ | 212 |
| DMPC/CHAPSO + aromatic amphiphiles | " | – | – | proteins | " | ∥ | 191 |
| Bacteriophage Pf1 | LC medium, negatively charged, 2 μm long cylindrical particles 6.5 nm in diameter | broad range of concentrations and temperatures; easy sample recovery by centrifugation | aggregated at low pH, binds to some proteins | proteins, nucleic acids, carbohydrates | filamentous phages, viruses | ∥ | 218, 220 – 222, 441 |

**Table S.1** (continued).

| System | | | | Biomolecule | Aligning factor | Alignment direction | Ref. |
|---|---|---|---|---|---|---|---|
| composition | description | advantages | drawbacks | | | | |
| Bacteriophage fd | negatively charged, 2 μm long cylindrical particles 6.5 nm in diameter | broad range of concentrations and temperatures; easy sample recovery by centrifugation | aggregated at low pH, binds to some proteins | proteins, nucleic acids, carbohydrates | filamentous phages, viruses | ‖ | 202, 224 |
| Tobacco mosaic virus | the same | the same | the same | the same | the same | ‖ | 202, 234 |
| CTAB – pentanol, disc-like micelles | forms a nematic LC phase | stable | strong orienting effect | proteins | lamellae | ‖ | 160 |
| CTAB – rod-like particles | the same | " | unstable | " | " | ‖ | 160 |
| Cetylpyridinium chloride – hexanol | positively charged, forms a lamellar LC phase | " | sensitive to salts | proteins, carbohydrates | " | ‖ | 170 |
| Cetylpyridinium bromide – hexanol | the same | " | – | the same | " | ‖ | 171, 200 |
| C8E5 – n-octanol, C12E6 – n-hexanol | non-charged, forms a lamellar LC phase | inert, stable; easy to make, easy sample recovery by dialysis | – | proteins, nucleic acids | coaxial cylindrical lamellae | ‖ | 172, 200 |
| Polyacrylamide gel: acrylamide – methyne-bis(acrylamide) copolymer | hydrogels containing anisotropic cavities, non-charged, oriented under mechanical impact (compression, stretching) | inert, orientation can be changed, biaxial orientation is possible, easy sample recovery by elution | non-homogeneous, low rates of rotational diffusion | proteins | deformed pores of the gel | depends on the type of mechanical treatment | 228 – 231 |
| Natural cellulose fibres | LC medium, negatively charged rods, 200 × 10 nm | stable, inert | slow alignment | " | cellulose crystallites | ⊥ | 226 |
| Purple membrane fragments of halobacteria containing bacteriorhodopsin | non-liquid crystal, negatively charged, particles, aligned by magnetic field | low concentrations, stabilised by acrylamide gel | interacts with proteins, aggregated at high ion strength | proteins, peptides | electrostatic interaction with aligned membrane fragments | ‖ | 214, 215, 229 |
| The same doped with lanthanide ions | the same | – | the same | the same | the same | ⊥ | 216 |
| Vanadium pentoxide | aqueous LC suspension, negatively charged units up to 600 nm long | stable, inert, absence of $^{1}H$ and $^{13}C$ nuclei, easy sample recovery by flocculation | low pH | carbohydrates | LC medium of ribbons | ‖ | 232 |

**Note**. The following abbreviations are used: C8E5, C12E6 are n-alkylpoly(ethylene glycols) C*m*E*n*, where *m* is the number of carbon atoms in the alkyl group, *n* is the number of ethylene glycol units; DBPC is 1-dodecanoyl-2-(biphenyl-4-acetyl)-*sn-glycero*-3-phosphocholine; DBBPC is 1-dodecanoyl-2-[4-(4-biphenyl)butanol]-*sn-glycero*-3-phosphocholine; DMPE – DTPA is 1,2-dimyristoyl-*sn-glycero*-3-phosphoethanolammonium diethylenetriaminepentaacetate.

**Table S.2**. Examples of biomolecules whose structures were determined using RDC.

| PDB code | BMRB number | Description | The number of calculated structures | The number of residues | The number of RDC | Ref. |
|---|---|---|---|---|---|---|
| 1B4C | | rat apo-S100B (ββ) metal binding protein | 21 | 184 | 521 | 362 |
| 1C05 | 4577 | ribosomal protein S4 Δ41 | see [a] | 159 | 101 | 384 |
| 1C06 | 4577 | the same | 16 | 159 | 101 | 384 |
| 1C7U | | DNA binding core domain of the transcription factor Mef2A with a 20-mer oligonucleotide | – | 210 | 368 | 458 |
| 1CMZ | 4407 | human Gα interacting protein GAIP | 20 | 152 | 291 | 371 |
| 1D2B | 4327 | N-terminal domain of human tissue inhibitor of metalloproteinases-1 | 29 | 126 | 159 | 365 |
| 1D3Z | | ubiquitin | 10 | 76 | 945 | 293 |
| 1D8V | 4522 | antitumour protein MAP30 | see [a] | 263 | 387 | 364 |
| 1DUF | | DNA dodecamer $d(CGCGAATTCGCG)_2$ | 5 | 24 | 408 | 446 |
| 1E8L | 4831 | Hen lysozyme | 50 | 129 | 209 | 372 |
| 1EKY | | cytochrome c′ from *Rhodobacter capsulatus* | 32 | 129 | 68 | 358 |
| 1EZP | | maltodextrin binding protein complexed with β-cyclodextrin | 10 | 370 | 970 | 310 |
| 1F0A | | DNA-damage-inducible protein I (DinI) | 20 | 81 | 534 | 366 |
| 1F16 | 4632 | pro-apoptosis protein Bax | 20 | 192 | 575 | 481 |
| 1F3Y | | nudix enzyme diadenosine tetraphosphate hydrolase from *Lupinus angustifolius L.* | 25 | 165 | 94 | 482 |
| 1F70 | | calmodulin N-terminal domain | 10 | 76 | 72 | 483 |
| 1F71 | | calmodulin C-terminal domain | 10 | 67 | 65 | 483 |
| 1FH1 | | lipid binding nodulation protein Nodf | see [a] | 92 | 41 | 306 |
| 1FZX | | DNA dodecamer d(GGCAAAAAACGG)/d(CCGTTTTTTGCC) | 10 | 24 | 85 | 484 |
| 1G03 | 4649 | N-terminal domain of human T-cell leukemia virus type I HTLV-I $CA_{1\text{-}134}$ | 20 | 134 | 193 | 485 |
| 1G14 | | DNA dodecamer d(GGCAAGAAACGG)/d(CCGTTTCTTGCC) | 10 | 24 | 80 | 484 |
| IGE9 | | ribosome recycling factor | 15 | 184 | 108 | |
| 1GGR | | complex of enzyme $IIA^{Glc}$ and the histidine-containing phosphocarrier protein HPr from *E. coli* | see [a] | 253 | 77/118 | 407 |
| IGHH | | DNA-damage-inducible protein I (DinI) | 20 | 81 | 534 | 366 |
| IGIP | | DNA dodecamer $d(CGCGAATTCGCG)_2$ | 2 | 24 | 408 | 356 |
| 1GJJ | | N-terminal constant region of the nuclear envelope protein LAP2 | see [a] | 50 | 123 | 486 |
| 1HJ7 | | LDL receptor EGF-AB $Ca^{2+}$-binding epidermal growth factor-like domains | 20 | 80 | 45 | 487 |
| 1H9C | 4955 | cysteinyl-phosphorylated enzyme IIB of the *N,N′*-diacetylchitobiose specific phosphoenol-pyruvate-dependent phosphotransferase system of *E. coli* | 9 | 106 | 73 | 488 |
| 1H92 | | SH3 domain of human Lck tyrosine kinase | 30 | 63 | 44 | 489 |
| 1I35 | 4463 | Ras-binding domain of the protein kinase Byr2 | 10 | 95 | 28 | 368, 490 |
| 1I42 | | Ubx domain from P47 | 10 | 89 | 56 | |
| 1IB8 | | conserved protein SP14.3 encoded by an essential *Streptococcus Pneumoniae* gene | see [a] | 164 | 153 | 491 |
| 1IUF | | two DNA-binding domains in *Schizosaccharomices pombe* Abp1 protein | " | 144 | 47 | 459 |
| 1J46 | | wild type HMG-Box domain of the human male sex determining factor SRY complexed to DNA | " | 113 | 274/197, 101/96 | 492 |
| 1J47 | | M9I mutant of HMG-Box domain of the human male sex determining factor SRY complexed to DNA | " | 113 | 274/197, 101/96 | 492 |
| 1J4V | | domain-swapped dimer of the HIV-inactivating protein cyanovirin-N | " | 202 | 136 | 383 |
| 1J4W | | complex of the KH3 and KH4 domains of human FBP with single-stranded 29-mer DNA | " | 203 | 140/154 | 460 |
| 1J5K | | complex of the KH3 domain of hnRNP K with a single-stranded 10-mer DNA | " | 99 | 147/140 | 493 |
| 1J6T | | complex of the enzyme $IIA^{Mtl}$ and histidine-containing phosphocarrier protein HPr from *E. coli* | " | 233 | 528 | 494 |
| 1J7O | | calcium calmodulin N-terminal domain | 3 | 71 | 323 | 390 |
| 1J7P | | calcium calmodulin C-terminal domain | 3 | 67 | 305 | 390 |
| 1JO7 | | influenza A virus promoter RNA | 32 | 31 | 22 | 464 |
| 1JOX | | LP5.1 hairpin from *Bacillus* RNase P RNA | 24 | 21 | 56 | 495 |

**Table S.2** (continued).

| PDB code | BMRB number | Description | The number of calculated structures | The number of residues | The number of RDC | Ref. |
|---|---|---|---|---|---|---|
| 1JWE | 4297 | N-terminal domain of *E. coli* DnaB helicase | 20 | 114 | 84 | 496 |
| 1K31, 1KQV | | Ca Ln calbindin $D_{9K}$ | 30 | 79 | 64 | 497 |
| 1KSM | | the same | see [a] | 79 | 64 | 497 |
| 1KAT | 5185, 5198 | phage-derived peptide antagonist complexed with vascular endothelial growth factor | 24 | 118 | 146 | 498 |
| 1KHM | 4405 | C-terminal KH domain of hnRNP K (KH3) | 20 | 89 | 209 | 361 |
| 1KMX | 5238 | heparin-binding domain of vascular endothelial growth factor | 20 | 55 | 38 | 392 |
| 1KQQ | | dead ringer ARID-DNA complex | 20 | 169 | 94 | 499 |
| 1KR8 | | d(GCGAAGC) hairpin | 14 | 7 | 38 | 466 |
| 1KVN | 4935 | protein SRP19 of *Archaeoglobus fulgidus* | 10 | 104 | 66 | 500 |
| 1KVV | 4935 | the same | see [a] | 104 | 66 | 500 |
| 1L4S | 5315 | ribosome associated factor Y | 20 | 112 | 80 | 501 |
| 1L5E | | the domain-swapped dimer of cyanovirin-N in a metastable folded state | see [a] | 202 | 80 | 502 |
| 1LUD | 5396 | dihydrofolate reductase complexed with trimethoprim and NADPH | 24 | 162 | 229 | |
| 1LVZ | 5376 | metarhodopsin II bound C-terminal peptide of transducin α-subunit | 20 | 11 | 38 | 402 |
| 1M2Y | | mutant *Pyrococcus furiosus* rubredoxin | see [a] | 54 | 400 | 308 |
| 1MAG | | gramicidin A | " | 34 | 141 | 95 |
| 1MPE | | tetrameric mutant of the B1 domain of streptococcal protein G | 21 | 224 | 33 | |
| 1UD7 | 4493 | ubiquitin core mutant 1D7 | 20 | 76 | 63 | 504 |
| 2EZM | | HIV-inactivating protein cyanovirin-N | see [a] | 101 | 380 | 369 |
| 2EZN | | the same | 40 | 101 | 380 | 369 |
| 1QCK | | human barrier-to-autointegration factor Baf | 20 + see [a] | 178 | 259 | 503 |
| 2EZX | | the same | 20 | 178 | 496 | 505 |
| 2EZY, 2EZZ | | " | 20 | 178 | 496 | 505 |
| 2GAT | | C-terminal domain of chicken transcription factor GATA-1 bound to DNA | see [a] | 98 | 48 | 141 |
| 3GAT | | the same | 34 | 98 | 48 | 141 |
| 3EZA | | complex of the N-terminal domain of enzyme I and the histidine-containing phosphocarrier protein HPr from *E. coli* | see [a] | 334 | 238 | 406 |
| 3EZB | | the same | 40 | 334 | 238 | 406 |
| 3GB1 | | immunoglobulin binding protein B1 domain of streptococcal protein G | – | 56 | 300 | 503 |
| 4GAT | | wild type DNA-binding domain of area complexed to a 13-mer DNA containing a CGATA site | see [a] | 92 | 48 | 141 |
| 6GAT | | L22V mutant DNA-binding domain of area complexed to a 13-mer DNA containing a TGATA site | " | 92 | 49 | 141 |
| | 4319 | oxidised rubredoxin | – | 54 | 53 | 506 |
| | 4320 | reduced rubredoxin | – | 54 | 53 | 506 |
| | 4335 | calerythrin (EF-hand calcium-binding protein) | – | 176 | 172 | 363 |
| | 4822 | regulatory domain of human cardiac troponin C (apo form) | – | 89 | 53 | 507 |
| | 4823 | regulatory domain of human cardiac troponin C (holo-form) | – | 89 | 53 | 507 |
| | 4824 | regulatory domain of human cardiac troponin C (complex with peptide) | – | 89 | 52 | 507 |
| | 4927 | circular permuted variant of the potent HIV-inactivating protein cyanovirin-N | – | 202 | 75 | 370 |
| | 4986 | maltodextrin-binding protein (MBP) (ligand free) | – | 370 | 903 | 387 |
| | 4987 | maltodextrin-binding protein (MBP) (maltotriose bound state) | – | 370 | 908 | 387 |
| | 5368 | human ADP ribosylation factor 1 (ARF1) | – | 360 | – | 508 |

[a] Average structure.

## References


1. H M Berman, J Westbrook, Z Feng, G Gilliland, T N Bhat, H Weissig, I N Shindyalov, P E Bourne *Nucleic Acids Res.* **28** 235 (2000)
2. G T Montelione, D Zheng, Y J Huang, K C Gunsalus, T Szyperski *Nat. Struct. Biol.* **7** 982 (2000)
3. J H Prestegard, H Valafar, J Glushka, F Tian *Biochemistry* **40** 8677 (2001)
4. U Heinemann, G Illing, H Oschkinat *Curr. Opin. Biotechnol.* **12** 348 (2001)
5. P R E Mittl, M G Grütter *Curr. Opin. Chem. Biol.* **5** 402 (2001)
6. R R Ernst, G Bodenhausen, W Wokaun *Principles of Nuclear Magnetic Resonances in One and Two Dimensions* (Oxford: Oxford University Press, 1986)
7. K Wüthrich *NMR of Proteins and Nucleic Acids* (New York: Wiley, 1986)
8. J C Hoch, F M Poulsen, C Redfield (Eds) *Computational Aspects of the Study of Biological Macromolecules by NMR Spectroscopy* (New York; London: Plenum, 1991)
9. G C K Roberts (Ed.) *NMR of Macromolecules. A Practical Approach* (Oxford: IRL Press, 1993)
10. J N S Evans *Biomolecular NMR Spectroscopy* (Oxford: Oxford University Press, 1995)
11. J L Markley, S J Opella (Eds) *Biological NMR Spectroscopy* (London: Oxford University Press, 1997)
12. N Beckman (Ed.) *Carbon-13 NMR Spectroscopy of Biological Systems* (San Diego: Academic Press, 1995)
13. J Cavanagh, W J Fairbrother, A G Palmer, N J Skelton *Protein NMR Spectroscopy: Principles and Practice* (San Diego: Academic Press, 1996)
14. D G Reid (Ed.) *Protein NMR Techniques. Methods in Molecular Biology*. Vol. 60 (Totowa, NJ: Humana Press, 1997)
15. K H Gardner, L E Kay *Annu. Rev. Biophys. Biomol. Struct.* **27** 357 (1998)
16. P Güntert *Q. Rev. Biophys.* **31** 145 (1998)
17. G Wider *Prog. Nucl. Magn. Reson. Spectrosc.* **32** 193 (1998)
18. N R Krishna, L J Berliner (Eds) *Modern Techniques in Protein. Series 'Biological Magnetic Resonance'* Vol. 16 (New York: Kluwer Academic – Plenum, 1999)
19. N R Krishna, L J Berliner (Eds) *NMR Structural Computation and Dynamics in Protein NMR. Series 'Biological Magnetic Resonance'* Vol. 17 (New York: Kluwer Academic – Plenum, 1999)
20. T L James, V Dotsch, U Schmitz (Eds) *NMR of Biological Macromolecules. Part A. Series 'Methods in Enzymology'* Vol. 338 (San Diego: Academic Press, 2001)
21. T L James, V Dotsch, U Schmitz (Eds) *NMR of Biological Macromolecules. Part B. Series 'Methods in Enzymology'* Vol. 339 (San Diego: Academic Press, 2001)
22. L E Kay , K H Gardner *Curr. Opin. Struct. Biol.* **7** 722 (1997)
23. V Dötsch, G Wagner *Curr. Opin. Struct. Biol.* **8** 619 (1998)
24. G M Clore, A M Gronenborn *Curr. Opin. Chem. Biol.* **2** 564 (1998)
25. M Sattler, J Schleucher, C Griesinger *Prog. Nucl. Magn. Reson. Spectrosc.* **34** 93 (1999)
26. K Pervushin, R Riek, G Wider, K Wüthrich *Proc. Natl. Acad. Sci. USA* **94** 12366 (1997)
27. K Pervushin, A Ono, C Fernández, T Szyperski, M Kainosho, K Wüthrich *Proc. Natl. Acad. Sci. USA* **95** 14147 (1998)
28. M Salzmann, K Pervushin, G Wider, H Senn, K Wüthrich *Proc. Natl. Acad. Sci. USA* **95** 13585 (1998)
29. R Riek, G Wider, K Pervushin, K Wüthrich *Proc. Natl. Acad. Sci. USA* **96** 4918 (1999)
30. M Salzmann, G Wider, K Pervushin, H Senn, K Wüthrich *J. Am. Chem. Soc.* **121** 844 (1999)
31. K Pervushin *Q. Rev. Biophys.* **33** 161 (2000)
32. R Riek, K Pervushin, K Wüthrich *Trends Biochem. Sci.* **25** 462 (2000)
33. C Fernández, K Adeishvili, K Wüthrich *Proc. Natl. Acad. Sci. USA* **98** 2358 (2001)
34. G Siegal, J van Duynhoven, M Baldus *Curr. Opin. Chem. Biol.* **3** 530 (1999)
35. G Wider, K Wüthrich *Curr. Opin. Struct. Biol.* **9** 594 (1999)
36. H Yu *Proc. Natl. Acad. Sci. USA* **96** 332 (1999)
37. M Salzmann, K Pervushin, G Wider, H Senn, K Wüthrich *J. Am. Chem. Soc.* **122** 7543 (2000)
38. R Xu, B Ayers, D Cowburn, T W Muir *Proc. Natl. Acad. Sci. USA* **96** 388 (1999)
39. N K Goto, L E Kay *Curr. Opin. Struct. Biol.* **10** 585 (2000)
40. A J Wand, M R Ehrhardt, P F Flynn *Proc. Natl. Acad. Sci. USA* **95** 15299 (1998)
41. R A Venters, R Thompson, J Cavanagh *J. Mol. Struct.* **602 – 603** 275 (2002)
42. H C Kung, K Y Wang, I Goljer, P H Bolton *J. Magn. Reson., Ser. B* **109** 323 (1995)
43. J R Tolman, J M Flanagan, M A Kennedy, J H Prestegard *Proc. Natl. Acad. Sci. USA* **92** 9279 (1995)
44. J H Prestegard *Nat. Struct. Biol., NMR Suppl.* **5** 517 (1998)
45. N Tjandra, A Bax *Science* **278** 1111 (1997)
46. J-C Hus, D Marion, M Blackledge *J. Am. Chem. Soc.* **123** 1541 (2001)
47. J R Tolman *Curr. Opin. Struct. Biol.* **11** 532 (2001)
48. H M Al-Hashimi, D J Patel *J. Biomol. NMR* **22** 1 (2002)
49. A Arora, L K Tamm *Curr. Opin. Struct. Biol.* **11** 540 (2001)
50. D Neuhaus, M P Williamson *The Nuclear Overhauser Effect in Stereochemical and Conformational Analysis* (New York: Wiley, 2000)
51. G Cornilescu, A Bax, D A Case *J. Am. Chem. Soc.* **122** 2168 (2000)
52. S S Wijmenga, B N M van Buuren *Prog. Nucl. Magn. Reson. Spectrosc.* **32** 287 (1998)
53. D A Case *Curr. Opin. Struct. Biol.* **8** 624 (1998)
54. D A Case *Curr. Opin. Struct. Biol.* **10** 197 (2000)
55. D S Wishart, A M Nip *Biochem. Cell Biol.* **76** 153 (1998)
56. G Cornilescu, F Delaglio, A Bax *J. Biomol. NMR* **13** 289 (1999)
57. A J Dingley, S Grzesiek *J. Am. Chem. Soc.* **120** 8293 (1998)
58. A J Dingley, J E Masse, R D Peterson, B Barfield, J Feigon, S Grzesiek *J. Am. Chem. Soc.* **121** 6019 (1999)
59. S Grzesiek, F Cordier, A J Dingley *Methods Enzymol.* **338** 111 (2001)
60. A J Dingley, F Cordier, S Grzesiek *Concepts Magn. Reson.* **13** 103 (2001)
61. F Cordier, S Grzesiek *J. Am. Chem. Soc.* **121** 1601 (1999)
62. G Cornilescu, J-S Hu, A Bax *J. Am. Chem. Soc.* **121** 2949 (1999)
63. G Cornilescu, B E Ramirez, M K Frank, G M Clore, A M Gronenborn, A Bax *J. Am. Chem. Soc.* **121** 6275 (1999)
64. N Tjandra, D S Garett, A M Gronenborn, A Bax, G M Clore *Nat. Struct. Biol.* **4** 443 (1997)
65. E de Alba, N Tjandra *J. Magn. Reson.* **144** 367 (2000)
66. B Brutscher *Concepts Magn. Reson.* **12** 207 (2000)
67. B Reif, M Hennig, C Griesinger *Science* **276** 1230 (1997)
68. D Yang, R Konrat, L E Kay *J. Am. Chem. Soc.* **119** 11938 (1997)
69. D Yang, K H Gardner, L E Kay *J. Biomol. NMR* **11** 213 (1998)
70. D Fushman, D Cowburn *J. Biomol. NMR* **13** 139 (1999)
71. C Richter, B Reif, C Griesinger, H Schwalbe *J. Am. Chem. Soc.* **122** 12728 (2000)
72. N R Skrynnikov, R Konrat, D R Muhandiram, L E Kay *J. Am. Chem. Soc.* **122** 7059 (2000)
73. I Bertini, J Kowalewski, C Luchinat, G Parigi *J. Magn. Reson.* **152** 103 (2001)
74. N Tjandra, A Szabo, A Bax *J. Am. Chem. Soc.* **118** 6986 (1996)
75. R Ghose, J H Prestegard *J. Magn. Reson.* **134** 308 (1998)
76. R Ghose, K Huang, J H Prestegard *J. Magn. Reson.* **135** 487 (1998)
77. D Yang, Y-K Mok, D R Muhandiram, J D Forman-Kay *J. Am. Chem. Soc.* **121** 3555 (1999)
78. B Reif, A Diener, M Hennig, M Maurer, C Griesinger *J. Magn. Reson.* **143** 45 (2000)
79. C Richter, C Griesinger, I Felli, P T Cole, G Varani, H Schwalbe *J. Biomol. NMR* **15** 241 (1999)
80. S J F Vincent, C Zwahlen *J. Am. Chem. Soc.* **122** 8307 (2000)
81. K Kloiber, W Schüler, R Konrat *J. Biomol. NMR* **19** 347 (2001)
82. M Rance, in *Encyclopedia of NMR* Vol. 9 (Eds D M Grant, R K Harris) (Chichester: Willey, 2002) p. 354
83. C Griesinger, J Meiler, W Peti, in *Biological Magnetic Resonance* (Eds N R Krishna, L J Berliner) (New York: Kluwer Academic – Plenum, 2002)

84. S J Opella, Y Kim, P McDonnell *Methods Enzymol.* **239** 536 (1994)
85. F M Marassi, S J Opella *Curr. Opin. Struct. Biol.* **8** 640 (1998)
86. R Q Fu, T A Cross *Annu. Rev. Biophys. Biomol. Struct.* **28** 235 (1999)
87. H J M De Groot *Curr. Opin. Struct. Biol.* **10** 593 (2000)
88. R Tycko *Annu. Rev. Phys. Chem.* **52** 575 (2001)
89. S J Opella, C Ma, F M Marassi *Methods Enzymol.* **339** 285 (2001)
90. S R Kiihne, H J M De Groot (Eds) *Perspectives on Solid State NMR in Biology* (Dordrecht: Kluwer Academic, 2001)
91. M J Duer (Ed.) *Solid-State NMR Spectroscopy. Principles and Applications* (Oxford: Blackwell, 2002)
92. C Glaubitz, A Watts *J. Magn. Reson.* **130** 305 (1998)
93. C Glaubitz, I J Burnett, G Gröbner, A J Mason, A Watts *J. Am. Chem. Soc.* **121** 5787 (1999)
94. C Glaubitz *Concepts Magn. Reson.* **12** 137 (2000)
95. R R Ketchem, K-C Lee, S Huo, T A Cross *J. Biomol. NMR* **8** 1 (1996)
96. R R Ketchem, B Roux, T A Cross *Structure* **5** 1655 (1997)
97. T A Cross, J R Quine *Concepts Magn. Reson.* **12** 55 (2000)
98. J R Quine, T A Cross *Concepts Magn. Reson.* **12** 71 (2000)
99. R Bertram, J R Quine, M S Chapman, T A Cross *J. Magn. Reson.* **147** 9 (2000)
100. M Cotten, R Fu, T A Cross *Biophys. J.* **76** 1179 (1999)
101. B Reif, C P Jaroniec, C M Rienstra, M Hohwy, R G Griffin *J. Magn. Reson.* **151** 320 (2001)
102. V Ladizhansky, M Veshtort, R G Griffin *J. Magn. Reson.* **154** 317 (2002)
103. C H Wu, A Ramamoorthy, S J Opella *J. Magn. Reson., Ser. A* **109** 270 (1994)
104. A Ramamoorthy, C H Wu, S J Opella *J. Magn. Reson., Ser. B* **107** 88 (1995)
105. F M Marassi, A Ramamoorthy, S J Opella *Proc. Natl. Acad. Sci. USA* **94** 8551 (1997)
106. R Fu, C Tian, T A Cross *J. Magn. Reson.* **154** 130 (2002)
107. J R Garbow, T Gullion, in *Carbon-13 NMR Spectroscopy of Biological Systems* (San Diego: Academic Press, 1995) p. 65
108. R G Griffin *Nat. Struct. Biol.* **5** 508 (1998)
109. F M Marassi, S J Opella *J. Magn. Reson.* **144** 150 (2000)
110. J Wang, J Denny, C Tian, S Kim, Y Mo, F Kovacs, Z Song, K Nishimura, Z Gan, R Fu, J R Quine, T A Cross *J. Magn. Reson.* **144** 162 (2000)
111. F M Marassi *Biophys. J.* **80** 994 (2001)
112. J K Denny, J Wang, T A Cross, J R Quine *J. Magn. Reson.* **152** 217 (2001)
113. M Bak, R Schultz, T Vosegaard, N C Nielsen *J. Magn. Reson.* **154** 28 (2002)
114. M Bak, J T Rasmussen, N C Nielsen *J. Magn. Reson.* **147** 296 (2000)
115. G D Henry, D B Sykes *Methods Enzymol.* **239** 515 (1994)
116. R R Vold, R S Prosser, A J Deese *J. Biomol. NMR.* **9** 329 (1997)
117. A Arora, F Abildgaard, J H Bushweller, L K Tamm *Nat. Struct. Biol.* **8** 334 (2001)
118. V Bystrov, A Arseniev, I Barsukov, A Lomize *Bull. Magn. Reson.* **8** 84 (1987)
119. R R Ketchem, W Hu, T A Cross *Science* **261** 1457 (1993)
120. F Kovacs, J Quine, T A Cross *Proc. Natl. Acad. Sci. USA* **96** 7910 (1999)
121. F M Marassi, J J Gesell, A P Valente, Y Kim, M Oblatt-Montal, M Montal, S J Opella *J. Biomol. NMR* **14** 141 (1999)
122. S J Opella, F M Marassi, J J Gesell, A P Valente, Y Kim, M Oblatt-Montal, M Montal *Nat. Struct. Biol.* **6** 374 (1999)
123. F A Kovacs, J K Denny, Z Song, J R Quine, T A Cross *J. Mol. Biol.* **295** 117 (2000)
124. Z Song, F A Kovacs, J Wang, J K Denny, S C Shekar, J R Quine, T A Cross *Biophys. J.* **79** 767 (2000)
125. F M Marassi, C Ma, H Gratkowski, S K Straus, K Strebel, M Oblatt-Montal, M Montal, S J Opella *Proc. Natl. Acad. Sci. USA* **96** 14336 (1999)
126. F M Marassi, S J Opella, P Juvvadi, R B Merrifield *Biophys. J.* **77** 3152 (1999)
127. K G Valentine, S-F Liu, F M Marassi, G Veglia, S J Opella, F-X Ding, S-H Wang, B Arshava, J M Becker, F Naider *Biopolymers* **59** 243 (2001)
128. S O Smith, D Song, S Shekar, M Groesbeek, M Ziliox, S Aimoto *Biochemistry* **40** 6553 (2001)
129. T Vosegaard, N C Nielsen *J. Biomol. NMR* **22** 225 (2002)
130. Y Ishii, R Tycko *J. Am. Chem. Soc.* **122** 1443 (2000)
131. J W Emsley, J C Lindon *NMR Spectroscopy Using Liquid Crystal Solvents* (Oxford: Pergamon Press, 1975)
132. A D Buckingham, K A McLauchlan, in *Progress in NMR Spectroscopy* (Oxford: Pergamon Press, 1967) p. 63
133. J H Prestegard, J R Tolman, H M Al-Hashimi, M Andrec, In *Biological Magnetic Resonance: Structure Computation and Dynamics in Protein NMR* Vol. 17 (New York: Kluwer Academic – Plenum, 1999) p. 311
134. J H Prestegard, H M Al-Hashimi, J R Tolman *Q. Rev. Biophys.* **33** 371 (2000)
135. J H Prestegard, A I Kishore *Curr. Opin. Chem. Biol.* **5** 584 (2001)
136. A A Bothner-By, in *Encyclopedia of NMR* (Chichester: Wiley, 1996) p. 2932
137. C Gayathri, A A Bothner-By, P C M van Zijl, C MacLean *Chem. Phys. Lett.* **87** 192 (1982)
138. A A Bothner-By, C Gayathri, P C M van Zijl, C MacLean, J-J Lai, K M Smith *Magn. Reson. Chem.* **23** 935 (1985)
139. E W Bastiaan, C MacLean, P C M van Zijl , A A Bothner-By *Annu. Rep. NMR Spectrosc.* **19** 35 (1987)
140. N Tjandra, S Grzesiek, A Bax *J. Am. Chem. Soc.* **118** 6264 (1996)
141. N Tjandra, J G Omichinski, A M Gronenborn, G M Clore, A Bax *Nat. Struct. Biol.* **4** 732 (1997)
142. R D Beger, V M Marathias, B F Volkman, P H Bolton *J. Magn. Reson.* **135** 256 (1998)
143. M A Contreras, J Ubach, O Millet, J Rizo, M Pons *J. Am. Chem. Soc.* **121** 8947 (1999)
144. R R Biekofsky, F W Muskett, J M Schmidt, S R Martin, J P Browne, P M Bayley, J Feeney *FEBS Lett.* **460** 519 (1999)
145. I Bertini, I C Felli, C Luchinat *J. Biomol. NMR* **18** 347 (2000)
146. I Bertini, M B L Janik, G Liu, C Luchinat, A Rosato *J. Magn. Reson.* **148** 23 (2001)
147. J Feeney, B Birdsall, A F Bradbury, R R Biekofsky, P M Bayley *J. Biomol. NMR* **21** 41 (2001)
148. R Barbieri, I Bertini, Y-M Lee, C Luchinat, A H Velders *J. Biomol. NMR* **22** 365 (2002)
149. V Gaponenko, A Dvoretsky, C Walsby, B M Hoffman, P R Rosevear *Biochemistry* **39** 15217 (2000)
150. C Ma, S J Opella *J. Magn. Reson.* **146** 381 (2000)
151. A Dvoretsky, V Gaponenko, P Rosevaar *FEBS Lett.* **528** 189 (2002)
152. L W Donaldson, N R Skrynnikov, W-Y Choy, D R Muhandiram, B Sarkar, J D Forman-Kay, L E Kay *J. Am. Chem. Soc.* **123** 9843 (2001)
153. J Boisbouvier, P Gans, M Blackledge, B Brutscher, D Marion *J. Am. Chem. Soc.* **121** 7700 (1999)
154. R Barbieri, I Bertini, G Cavallaro, Y-M Lee, C Luchinat, A Rosato *J. Am. Chem. Soc.* **124** 5581 (2002)
155. I Bertini, M Longinetti, C Luchinat, G Parigi, L Sgheri *J. Biomol. NMR* **22** 123 (2002)
156. I Bertini, G Cavallaro, M Cosenza, R Kümmerle, C Luchinat, M Piccioli, L Poggi *J. Biomol. NMR* **23** 115 (2002)
157. I Bertini, C Luchinat, G Parigi *Solution NMR of Paramagnetic Molecules. Applications to Metallobiomolecules and Models* (Amsterdam: Elsevier, 2001)
158. L Banci, I Bertini, J G Huber, C Luchinat, A Rosato *J. Am. Chem. Soc.* **120** 12903 (1998)
159. H Déméné, P Tsan, P Gans, D Marion *J. Phys. Chem. B* **104** 2559 (2000)
160. I Bertini, F Castellani, C Luchinat, G Martini, G Parigi, S Ristori *J. Phys. Chem. B* **104** 10653 (2000)
161. F Arnesano, L Banci, I Bertini, K van der Wetering, M Czisch, R Kaptein *J. Biomol. NMR* **17** 295 (2000)
162. C W Hilbers, C MacLean *NMR Basic Princ. Progr.* **7** 1 (1972)
163. T M Planteriga, B H Ruessink, C Maclean *Chem. Phys.* **48** 359 (1980)
164. A D Buckingham, K A McLauchlan *Proc. Chem. Soc., London* 144 (1963)

doi> 165. A Peshkovsky, A E McDermott *J. Phys. Chem. A* **103** 8604 (1999)
doi> 166. A Peshkovsky, A E McDermott *J. Magn. Reson.* **147** 104 (2000)
doi> 167. S A Riley, M P Augustine *J. Phys. Chem. A* **104** 3326 (2000)
doi> 168. S A Riley, M P Augustine *J. Magn. Reson.* **144** 288 (2000)
169. S A Riley, S Vyas, M P Augustine *Concepts Magn. Reson.* **13** 171 (2001)
doi> 170. R S Prosser, J A Losonczi, I V Shiyanovskaya *J. Am. Chem. Soc.* **120** 11010 (1998)
doi> 171. L G Barrientos, C Dolan, A M Gronenborn *J. Biomol. NMR* **16** 329 (2000)
doi> 172. M Rückert, G Otting *J. Am. Chem. Soc.* **122** 7793 (2000)
173. P Ram, J H Prestegard *Biochim. Biophys. Acta* **940** 289 (1988)
174. C R Sanders II, J P Schwonek *Biochemistry* **31** 8898 (1992)
doi> 175. C R Sanders II, B J Hare, K P Howard, J H Prestegard *Prog. Nucl. Magn. Reson. Spectrosc.* **26** 421 (1994)
176. C R Sanders II, G C Landis *Biochemistry* **34** 4030 (1995)
doi> 177. N Tjandra, A Bax *J. Magn. Reson* **124** 512 (1997)
178. A Bax, N Tjandra *J. Biomol. NMR* **10** 289 (1997)
179. C R Sanders, R S Prosser *Struct. Fold. Des.* **6** 1227 (1998)
doi> 180. K P Howard, S J Opella *J. Magn. Reson., Ser. B* **112** 91 (1996)
doi> 181. J J Chou, J D Kaufman, S J Stahl, P T Wingfield, A Bax *J. Am. Chem. Soc.* **124** 2450 (2002)
182. P A Luchette, T N Vetman, R S Prosser, R E W Hancock, M-P Nieh, C J Glinka, S Krueger, J Katsaras *Biochim. Biophys. Acta* **1513** 83 (2001)
doi> 183. J A Losonczi, J H Prestegard *Biochemistry* **37** 706 (1998)
doi> 184. J A Losonczi, F Tian, J H Prestegard *Biochemistry* **39** 3804 (2000)
doi> 185. R R Vold, R S Prosser *J. Magn. Reson., Ser. B* **113** 267 (1996)
186. K J Glover, J A Whiles, G Wu, N-J Yu, R Deems, J O Struppe, R E Stark, E A Komives, R R Vold *Biophys. J.* **81** 2163 (2001)
doi> 187. J Struppe, R R Vold *J. Magn. Reson.* **135** 541 (1998)
doi> 188. M Ottiger, A Bax *J. Biomol. NMR* **12** 361 (1998)
doi> 189. J Struppe, E A Komives, S S Taylor, R R Vold *Biochemistry* **37** 15523 (1998)
doi> 190. J A Losonczi, J H Prestegard *J. Biomol. NMR* **12** 447 (1998)
191. C R Sanders II, J E Schaff, J H Prestegard *Biophys. J.* **64** 1069 (1993)
doi> 192. B E Ramirez, A Bax *J. Am. Chem. Soc.* **120** 9106 (1998)
doi> 193. H Wang, M Eberstadt, E T Olejniczak, R P Meadows, S W Fesik *J. Biomol. NMR* **12** 443 (1998)
doi> 194. S Cavagnero, H J Dyson, P E Wright *J. Biomol. NMR* **13** 387 (1999)
195. J Struppe, J A Whiles, R R Vold *Biophys. J.* **78** 281 (2000)
doi> 196. M Ottiger, A Bax *J. Biomol. NMR* **13** 187 (1999)
doi> 197. M Zweckstetter, A Bax *J. Am. Chem. Soc.* **122** 3791 (2000)
doi> 198. G Zandomeneghi, M Tomaselli, J D van Beek, B H Meier *J. Am. Chem. Soc.* **123** 910 (2001)
199. F Picard, M-J Paquet, J Levesque, A Bélanger, M Auger *Biophys. J.* **77** 888 (1999)
doi> 200. S Gaemers, A Bax *J. Am. Chem. Soc.* **123** 12343 (2001)
doi> 201. E Brunner, M R Arnold, W Kremer, H R Kalbitzer *J. Biomol. NMR* **21** 173 (2001)
doi> 202. G M Clore, M R Starich, A M Gronenborn *J. Am. Chem. Soc.* **120** 10571 (1998)
203. C R Sanders II, J H Prestegard *Biophys. J.* **58** 447 (1990)
doi> 204. R S Prosser, S A Hunt, J A DiNatale, R R Vold *J. Am. Chem. Soc.* **118** 269 (1996)
205. R S Prosser, J S Hwang, R R Vold *Biophys. J.* **74** 2405 (1998)
206. R S Prosser, V B Volkov, I V Shiyanovskaya *Biophys. J.* **75** 2163 (1998)
doi> 207. R S Prosser, V B Volkov, I V Shiyanovskaya *Biochem. Cell Biol.* **76** 443 (1998)
208. K J Crowell, P M Macdonald *Biophys. J.* **81** 255 (2001)
doi> 209. R S Prosser, I V Shiyanovskaya *Concepts Magn. Reson.* **13** 19 (2001)
210. M-P Nieh, C J Glinka, S Krueger, R S Prosser, J Katsaras *Biophys. J.* **82** 2487 (2002)
doi> 211. G Cho, B M Fung, V B Reddy *J. Am. Chem. Soc.* **123** 1537 (2001)
212. C Tan, B M Fung, G Cho *J. Am. Chem. Soc.* **124** 11827 (2002)
doi> 213. G Veglia, S J Opella *J. Am. Chem. Soc.* **122** 11733 (2000)
doi> 214. B W Koenig, J-S Hu, M Ottiger, S Bose, R W Hendler, A Bax *J. Am. Chem. Soc.* **121** 1385 (1999)
doi> 215. J Sass, F Cordier, A Hoffmann, M Rogowski, A Cousin, J G Omichinski, H Löwen, S Grzesiek *J. Am. Chem. Soc.* **121** 2047 (1999)
doi> 216. P Bernado, R Barbieri, E Padros, C Luchinat, P Pons *J. Am. Chem. Soc.* **124** 374 (2002)
doi> 217. P A Luchette, R S Prosser, C R Sanders *J. Am. Chem. Soc.* **124** 1778 (2002)
doi> 218. M R Hansen, L Mueller, A Pardi *Nat. Struct. Biol.* **5** 1065 (1998)
doi> 219. M R Hansen, M Rance, A Pardi *J. Am. Chem. Soc.* **120** 11210 (1998)
220. M R Hansen, P Hanson, A Pardi *Methods Enzymol.* **317** 220 (2000)
221. M R Hansen, P Hanson, A Pardi *J. Biomol. Struct. Dyn.* **11** 365 (2000)
doi> 222. M Zweckstetter, A Bax *J. Biomol. NMR* **20** 365 (2001)
doi> 223. D D Ojennus, R M Mitton-Fry, D S Wuttke *J. Biomol. NMR* **14** 175 (1999)
doi> 224. L G Barrientos, J M Louis, A M Gronenborn *J. Magn. Reson.* **149** 154 (2001)
doi> 225. J-F Trempe, F G Morin, Z Xia, R H Marchessault, K Gehring *J. Biomol. NMR* **22** 83 (2002)
doi> 226. K Fleming, D Gray, S Prasannan, S Matthews *J. Am. Chem. Soc.* **122** 5224 (2000)
227. B Deloche, E T Samulski *Macromolecules* **14** 575 (1981)
doi> 228. R Tycko, F J Blanco, Y Ishii *J. Am. Chem. Soc.* **122** 9340 (2000)
229. H J Sass, G Musco, S J Stahl, P T Wingfield, S Grzesiek *J. Biomol. NMR* **18** 303 (2000)
doi> 230. Y Ishii, M A Markus, R Tycko *J. Biomol. NMR* **21** 141 (2001)
doi> 231. J J Chou, S Gaemers, B Howder, J M Louis, A Bax *J. Biomol. NMR* **21** 377 (2001)
doi> 232. H Desvaux, J-C P Gabriel, P Berthault, F Camerel *Angew. Chem., Int. Ed.* **40** 373 (2001)
doi> 233. J-C P Gabriel, F Camerel, B J Lemaire, H Desvaux, P Davidson, P Batail *Nature (London)* **413** 504 (2001)
doi> 234. G M Clore, M R Starich, C A Bewley, M Cai, J Kuszewski *J. Am. Chem. Soc.* **121** 6513 (1999)
235. A Bax, G Kontaxis, N Tjandra *Methods Enzymol.* **339** 127 (2001)
236. F Kramer, B Luy, S J Glaser *Appl. Magn. Reson.* **17** 173 (1999)
doi> 237. B Luy, S J Glaser *J. Magn. Reson.* **142** 280 (2000)
doi> 238. B Luy, S J Glaser *J. Magn. Reson.* **148** 169 (2001)
doi> 239. F Kramer, S J Glaser *J. Magn. Reson.* **155** 83 (2002)
doi> 240. B Luy, J J Barchi Jr, J P Marino *J. Magn. Reson.* **152** 179 (2001)
doi> 241. P Permi, A Annila *J. Biomol. NMR* **16** 221 (2000)
doi> 242. F Cordier, A J Dingley, S Grzesiek *J. Biomol. NMR* **13** 175 (1999)
doi> 243. D Yang, A V Venters, G A Mueller, W Y Choy, L E Kay *J. Biomol. NMR* **14** 333 (1999)
doi> 244. P Permi, P R Rosevear, A Annila *J. Biomol. NMR* **17** 43 (2000)
doi> 245. J Evenäs, A Mittermaier, D Yang, L E Kay *J. Am. Chem. Soc.* **123** 2858 (2001)
doi> 246. L Žídek, H Wu, J Feigon, V Sklenář *J. Biomol. NMR* **21** 153 (2001)
doi> 247. J Yan, T Corpora, P Pradhan, J H Bushweller *J. Biomol. NMR* **22** 9 (2002)
doi> 248. Z Wu, A Ono, M Kainosho, A Bax *J. Biomol. NMR* **19** 361 (2001)
doi> 249. J R Tolman, J H Prestegard *J. Magn. Reson., Ser. B* **112** 245 (1996)
doi> 250. J R Tolman, J H Prestegard *J. Magn. Reson., Ser. B* **112** 269 (1996)
doi> 251. M Ottiger, F Delaglio, A Bax *J. Magn. Reson.* **131** 373 (1998)
doi> 252. Y-X Wang, J L Marquardt, P Wingfield, S J Stahl, S Lee-Huang, D A Torchia, A Bax *J. Am. Chem. Soc.* **120** 7385 (1998)
doi> 253. M Ottiger, A Bax *J. Am. Chem. Soc.* **120** 12334 (1998)
doi> 254. C W Vander Kooi, E Kupče, E R P Zuiderweg, M Pellecchia *J. Biomol. NMR* **15** 335 (1999)
doi> 255. G Kontaxis, G M Clore, A Bax *J. Magn. Reson.* **143** 184 (2000)
256. B Cutting, J R Tolman, S Nanchen, G Bodenhausen *J. Biomol. NMR* **23** 195 (2002)
257. K Ding, A M Gronenborn *J. Magn. Reson.* **158** 173 (2002)

258. D Yang, J R Tolman, N K Goto, L E Kay *J. Biomol. NMR* **12** 325 (1998)
259. J J Chou, A Bax *J. Am. Chem. Soc.* **123** 3844 (2001)
260. A Mittermaier, L E Kay *J. Am. Chem. Soc.* **123** 6892 (2001)
261. A W Giesen, L C Bae, C L Barrett, J A Chyba, M M Chaykovsky, M-C Cheng, J H Murray, E J Oliver, S M Sullivan, J M Brown, F W Dahlquist, S W Homans *J. Biomol. NMR* **19** 255 (2001)
262. A W Giesen, L C Bae, C L Barrett, J A Chyba, M M Chaykovsky, M-C Cheng, J H Murray, E J Oliver, S M Sullivan, J M Brown, S W Homans *J. Biomol. NMR* **22** 21 (2002)
263. J J Chou, F Delaglio, A Bax *J. Biomol. NMR* **18** 101 (2000)
264. T Carlomagno, W Peti, C Griesinger *J. Biomol. NMR* **17** 99 (2000)
265. A Kaikkonen, G Otting *J. Am. Chem. Soc.* **123** 1770 (2001)
266. G Kontaxis, A Bax *J. Biomol. NMR* **20** 77 (2001)
267. P Permi *J. Magn. Reson.* **153** 267 (2001)
268. B Brutscher, J Boisbouvier, A Pardi, D Marion, J-P Simorre *J. Am. Chem. Soc.* **120** 11845 (1998)
269. P Vallurupalli, P B Moore *J. Biomol. NMR* **24** 63 (2002)
270. E de Alba, M Suzuki, N Tjandra *J. Biomol. NMR* **19** 63 (2001)
271. P Permi, A Annila *J. Biomol. NMR* **20** 127 (2001)
272. F Tian, J A Losonczi, M W F Fischer, J H Prestegard *J. Biomol. NMR* **15** 145 (1999)
273. B Brutscher *J. Magn. Reson.* **151** 332 (2001)
274. P Permi, S Heikkinen, I Kilpeläinen, A Annila *J. Magn. Reson.* **139** 273 (1999)
275. P Permi *J. Biomol. NMR* **22** 27 (2002)
276. M Cai, Y Huang, G M Clore *J. Am. Chem. Soc.* **123** 8642 (2001)
277. G M Clore, E C Murphy, A M Gronenborn, A Bax *J. Magn. Reson.* **134** 164 (1998)
278. T Carlomagno, M Hennig, J R Williamson *J. Biomol. NMR* **22** 65 (2002)
279. F Tian, P J Bolon, J H Prestegard *J. Am. Chem. Soc.* **121** 7712 (1999)
280. F Tian, C A Fowler, E R Zartler, F A Jenney Jr, M W Adams, J H Prestegard *J. Biomol. NMR* **18** 23 (2000)
281. Z Wu, A Bax *J. Magn. Reson.* **151** 242 (2001)
282. M Cai, H Wang, E T Olejniczak, R P Meadows, A H Gunasekera, N Xu, S W Fesik *J. Magn. Reson.* **139** 451 (1999)
283. W Peti, C Griesinger *J. Am. Chem. Soc.* **122** 3975 (2000)
284. M Pellecchia, C W Vander Kooi, K Keliikuli, E R P Zuiderweg *J. Magn. Reson.* **143** 435 (2000)
285. F Delaglio, Z Wu, A Bax *J. Magn. Reson.* **149** 276 (2001)
286. G Otting, M Rückert, M H Levitt, A Moshref *J. Biomol. NMR* **16** 343 (2000)
287. E Brunner, J Ogle, M Wenzler, H R Kalbitzer *Biochem. Biophys. Res. Commun.* **272** 694 (2000)
288. W Y Choy, M Tollinger, G A Mueller, L E Kay *J. Biomol. NMR* **21** 31 (2001)
289. Z Wu, N Tjandra, A Bax *J. Am. Chem. Soc.* **123** 3617 (2001)
290. M Ottiger, N Tjandra, A Bax *J. Am. Chem. Soc.* **119** 9825 (1997)
291. J Boyd, C Redfield *J. Am. Chem. Soc.* **121** 7441 (1999)
292. N Tjandra, A Bax *J. Am. Chem. Soc.* **119** 9576 (1997)
293. G Cornilescu, J L Marquardt, M Ottiger, A Bax *J. Am. Chem. Soc.* **120** 6836 (1998)
294. G Cornilescu, A Bax *J. Am. Chem. Soc.* **122** 10 143 (2000)
295. R S Lipsitz, N Tjandra *J. Am. Chem. Soc.* **123** 11 065 (2001)
296. I Bertini, I C Felli, C Luchinat *J. Magn. Reson.* **134** 360 (1998)
297. L Banci, I Bertini, M A Cremonini, G Gori-Savellini, C Luchinat, K Wüthrich, P Güntert *J. Biomol. NMR* **12** 553 (1998)
298. A C LiWang, A Bax *J. Magn. Reson.* **127** 54 (1997)
299. A Mittermaier, L E Kay *J. Am. Chem. Soc.* **121** 10608 (1999)
300. A Kjellberg, S W Homans *J. Magn. Reson.* **151** 90 (2001)
301. G M Clore, A M Gronenborn *Proc. Natl. Acad. Sci. USA* **95** 5891 (1998)
302. G M Clore, A M Gronenborn, N Tjandra *J. Magn. Reson.* **131** 159 (1998)
303. G M Clore, D S Garett *J. Am. Chem. Soc.* **121** 9008 (1999)
304. H N B Moseley, G T Montelione *Curr. Opin. Struct. Biol.* **9** 635 (1999)
305. N Tjandra, J Marquardt, G M Clore *J. Magn. Reson.* **142** 393 (2000)
306. C A Fowler, F Tian, H M Al-Hashimi, J H Prestegard *J. Mol. Biol.* **304** 447 (2000)
307. C A Rohl, D Baker *J. Am. Chem. Soc.* **124** 2723 (2002)
308. F Tian, H Valafar, J H Prestegard *J. Am. Chem. Soc.* **123** 11 791 (2001)
309. M Zweckstetter, A Bax *J. Am. Chem. Soc.* **123** 9490 (2001)
310. G A Mueller, W Y Choy, D Yang, J D Forman-Kay, R A Venters, L E Kay *J. Mol. Biol.* **300** 197 (2000)
311. C D Schwieters, G M Clore *J. Magn. Reson.* **149** 239 (2001)
312. C D Schwieters, G M Clore *J. Magn. Reson.* **152** 288 (2001)
313. P Dosset, J-C Hus, D Marion, M Blackledge *J. Biomol. NMR* **20** 223 (2001)
314. G M Clore, C D Schwieters *Curr. Opin. Struct. Biol.* **12** 146 (2002)
315. A Saupe *Z. Naturforsch., A* **19** 161 (1964)
316. E Brunner *Concepts Magn. Reson.* **13** 238 (2001)
317. J A Mueller, W Y Choy, N R Skrynnikov, L E Kay *J. Biomol. NMR* **18** 183 (2000)
318. G M Clore, A M Gronenborn, A Bax *J. Magn. Reson.* **133** 216 (1998)
319. G M Clore, A M Gronenborn, A Szabo, N Tjandra *J. Am. Chem. Soc.* **120** 4889 (1998)
320. E de Alba, J L Baber, N Tjandra *J. Am. Chem. Soc.* **121** 4282 (1999)
321. D Fushman, R Ghose, D Cowburn *J. Am. Chem. Soc.* **122** 10 640 (2000)
322. J J Warren, P B Moore *J. Magn. Reson.* **149** 271 (2001)
323. J-C Hus, J J Prompers, R Brüschweiler *J. Magn. Reson.* **157** 119 (2002)
324. M Zweckstetter, A Bax *J. Biomol. NMR* **23** 127 (2002)
325. S Moltke, S Grzesiek *J. Biomol. NMR* **15** 77 (1999)
326. H-J Sass, G Musco, S J Stahl, P T Wingfield, S Grzesiek *J. Biomol. NMR* **21** 275 (2001)
327. M X Fernandes, P Bernadó, M Pons, J G de la Torre *J. Am. Chem. Soc.* **123** 12 037 (2001)
328. A Almond, J Bunkenborg, T Franch, C H Gotfredsen, J Ø Duus *J. Am. Chem. Soc.* **123** 4792 (2001)
329. H F Azurmendi, C A Bush *J. Am. Chem. Soc.* **124** 2426 (2002)
330. K Nomura, M Kainosho *J. Magn. Reson.* **154** 146 (2002)
331. J Meiler, N Blomberg, M Nilges, C Griesinger *J. Biomol. NMR* **16** 245 (2000)
332. W Peti, J Meiler, R Brüschweiler, C Griesinger *J. Am. Chem. Soc.* **124** 5822 (2002)
333. N R Skrynnikov, L E Kay *J. Biomol. NMR* **18** 239 (2000)
334. G Lipari, A Szabo *J. Am. Chem. Soc.* **104** 4546 (1982)
335. G Lipari, A Szabo *J. Am. Chem. Soc.* **104** 4559 (1982)
336. W J Wedemeyer, C A Rohl, H A Scheraga *J. Biomol. NMR* **22** 137 (2002)
337. J Meiler, J J Prompers, W Peti, C Griesinger, R Brüschweiler *J. Am. Chem. Soc.* **123** 6098 (2001)
338. H M Al-Hashimi, H Valafar, M Terrell, E R Zartler, M K Eidsness, J H Prestegard *J. Magn. Reson.* **143** 402 (2000)
339. J-C Hus, R Brüschweiler *J. Biomol. NMR* **24** 123 (2002)
340. C A E M Spronk, J P Linge, C W Hilbers, G W Vuister *J. Biomol. NMR* **22** 281 (2002)
341. P Diehl, in *Accurate Molecular Structures* (London: Oxford University Press, 1992) p. 299
342. D A Case *J. Biomol. NMR* **15** 95 (1999)
343. A Annila, H Aitio, E Thulin, T Drakenberg *J. Biomol. NMR* **14** 223 (1999)
344. J Meiler, W Peti, C Griesinger *J. Biomol. NMR* **17** 283 (2000)
345. F Delaglio, G Kontaxis, A Bax *J. Am. Chem. Soc.* **122** 2142 (2000)
346. M Andrec, P Du, R M Levy *J. Am. Chem. Soc.* **123** 1222 (2001)
347. M Andrec, P Du, R M Levy *J. Biomol. NMR* **21** 335 (2001)
348. D S Garrett, J Kuszewski, T J Hancock, P J Lodi, G W Vuister, A M Gronenborn, G M Clore *J. Magn. Reson., Ser. B* **104** 99 (1994)
349. J Kuszewski, A M Gronenborn, G M Clore *J. Magn. Reson., Ser. B* **107** 293 (1995)
350. J Kuszewski, J Qin, A M Gronenborn, G M Clore *J. Magn. Reson., Ser. B* **106** 92 (1995)
351. J Kuszewski, A M Gronenborn, G M Clore *J. Magn. Reson., Ser. B* **112** 79 (1996)

352. J Kuszewski, A M Gronenborn, G M Clore *Protein Sci.* **5** 1067 (1996)
353. J Kuszewski, A M Gronenborn, G M Clore *J. Magn. Reson.* **125** 171 (1997)
354. J Kuszewski, G M Clore *J. Magn. Reson.* **146** 249 (2000)
355. G M Clore, J Kuszewski *J. Am. Chem. Soc.* **124** 2866 (2002)
356. J Kuszewski, C Schwieters, G M Clore *J. Am. Chem. Soc.* **123** 3903 (2001)
357. N Tjandra *Struct. Fold. Des.* **7** R205 (1999)
358. J-C Hus, D Marion, M Blackledge *J. Mol. Biol.* **298** 927 (2000)
359. M Allegrozzi, I Bertini, M B L Janik, Y-M Lee, G Lin, C Luchinat *J. Am. Chem. Soc.* **122** 4154 (2000)
360. B Simon, M Sattler *Angew. Chem., Int. Ed.* **41** 437 (2002)
361. J L Baber, D Libutti, D Levens, N Tjandra *J. Mol. Biol.* **289** 949 (1999)
362. A C Drohat, N Tjandra, D M Baldisseri, D J Weber *Protein Sci.* **8** 800 (1999)
363. H Aitio, A Annila, S Heikkinen, E Thulin, T Drakenberg, I Kilpeläinen *Protein Sci.* **8** 2580 (1999)
364. Y-X Wang, N Neamati, J Jacob, I Palmer, S J Stahl, J D Kaufman, P L Huang, P L Huang, H E Winslow, Y Pommier, P T Wingfield, S Lee-Huang, A Bax, D A Torchia *Cell* **99** 433 (1999)
365. B Wu, S Arumugam, G Gao, G Lee, V Semenchenko, W Huang, K Brew, S R Van Doren *J. Mol. Biol.* **295** 257 (2000)
366. B E Ramirez, O N Voloshin, R D Camerini-Otero, A Bax *Protein Sci.* **9** 2161 (2000)
367. E Wasielewski, R A Atkinson, M A Abdallah, B Kieffer *Biochemistry* **41** 12488 (2002)
368. W Gronwald, E Brunner, F Huber, M Wenzler, C Herrmann, H R Kalbitzer *Protein Sci.* **10** 1260 (2001)
369. C A Bewley, K Gustafson, M Boyd, D G Covell, A Bax, G M Clore, A M Gronenborn *Nat. Struct. Biol.* **5** 571 (1998)
370. L G Barrientos, R Campos-Olivas, J M Louis, A Fiser, A Šali, A M Gronenborn *J. Biomol. NMR* **19** 289 (2001)
371. E de Alba, L de Vries, M G Farquhar, N Tjandra *J. Mol. Biol.* **291** 927 (1999)
372. H Schwalbe, S B Grimshaw, A Spencer, M Buck, J Boyd, C M Dobson, C Redfield, L J Smith *Protein Sci.* **10** 677 (2001)
373. E R Zartler, F E Jenney Jr , M Terrell, M K Eidsness, M W Adams, J H Prestegard *Biochemistry* **40** 7279 (2001)
374. D Shortle, M S Ackerman *Science* **293** 487 (2001)
375. M S Ackerman, D Shortle *Biochemistry* **41** 3089 (2002)
376. M S Ackerman, D Shortle *Biochemistry* **41** 13791 (2002)
377. J L Weaver, J H Prestegard *Biochemistry* **37** 116 (1998)
378. J A Losonczi, M Andrec, M W F Fischer, J H Prestegard *J. Magn. Reson.* **138** 334 (1999)
379. M W F Fisher, J A Losonczi, J L Weaver, J H Prestegard *Biochemistry* **38** 9013 (1999)
380. V Tsui, L Zhu, T H Huang, P E Wright, D A Case *J. Biomol. NMR* **16** 9 (2000)
381. C A Bewley, G M Clore *J. Am. Chem. Soc.* **122** 6009 (2000)
382. C A Bewley *J. Am. Chem. Soc.* **123** 1014 (2001)
383. G M Clore, C A Bewley *J. Magn. Reson.* **154** 329 (2002)
384. M A Markus, R B Gerstner, D E Draper, D A Torchia *J. Mol. Biol.* **292** 375 (1999)
385. N R Skrynnikov, N K Goto, D Yang, W-Y Choy, J R Tolman, G A Mueller, L E Kay *J. Mol. Biol.* **295** 1265 (2000)
386. P M Hwang, N R Skrynnikov, L E Kay *J. Biomol. NMR* **20** 83 (2001)
387. J Evenäs, V Tugarinov, N R Skrynnikov, N K Goto, R Muhandiram, L E Kay *J. Mol. Biol.* **309** 961 (2001)
388. D T Braddock, M Cai, J L Baber, Y Huang, G M Clore *J. Am. Chem. Soc.* **123** 8634 (2001)
389. N K Goto, N R Skrynnikov, F W Dahlquist, L E Kay *J. Mol. Biol.* **308** 745 (2001)
390. J J Chou, S Li, C B Klee, A Bax *Nat. Struct. Biol.* **8** 990 (2001)
391. T K Mal, N R Skrynnikov, K L Yap, L E Kay, M Ikura *Biochemistry* **41** 12899 (2002)
392. M E Stauffer, N J Skelton, W J Fairbrother *J. Biomol. NMR* **23** 57 (2002)
393. H Déméné, T Ducat, P Barthe, M-A Delsuc, C Roumestand *J. Biomol. NMR* **22** 47 (2002)
394. B Stevensson, A V Komolkin, D Sandström, A Maliniak *J. Chem. Phys.* **114** 2332 (2001)
395. D J Craik, J A Wilce *Methods Mol. Biol. (Tototwa, N.J.)* **60** 195 (1997)
396. B W Koenig *ChemBiochem.* **3** 975 (2002)
397. H Shimizu, A Donohue-Rolfe, S W Homans *J. Am. Chem. Soc.* **121** 5815 (1999)
398. E T Olejniczak, R P Meadows, H Wang, M Cai, D G Nettesheim, S W Fesik *J. Am. Chem. Soc.* **121** 9249 (1999)
399. H M Al-Hashimi, P J Bolon, J H Prestegard *J. Magn. Reson.* **142** 153 (2000)
400. P J Bolon, H M Al-Hashimi, J H Prestegard *J. Mol. Biol.* **293** 107 (1999)
401. B W Koenig, D C Mitchell, S König, S Grzesiek, B J Litman, A Bax *J. Biomol. NMR* **16** 121 (2000)
402. B W Koenig, G Kontaxis, D C Mitchell, J M Louis, B J Litman, A Bax *J. Mol. Biol.* **322** 441 (2002)
403. G S Thompson, H Shimizu, S W Homans, A Donohue-Rolfe *Biochemistry* **39** 13153 (2000)
404. E R P Zuiderweg *Biochemistry* **41** 1 (2002)
405. G M Clore *Proc. Natl. Acad. Sci. USA* **97** 9021 (2000)
406. D S Garrett, Y-J Seok, A Peterkofsky, A M Gronenborn, G M Clore *Nat. Struct. Biol.* **6** 166 (1999)
407. G Wang, J M Louis, M Sondej, Y-J Scok, A Peterkofsky, G M Clore *EMBO J.* **19** 5635 (2000)
408. M A McCoy, D F Wyss *J. Am. Chem. Soc.* **124** 2104 (2002)
409. V Gaponenko, A S Altieri, J Li, R A Byrd *J. Biomol. NMR* **24** 143 (2002)
410. M Ottiger, A Bax *J. Am. Chem. Soc.* **121** 4690 (1999)
411. M Ottiger, F Defaglio, J L Marquardt, N Tjandra, A Bax *J. Magn. Reson.* **134** 365 (1998)
412. A Mittermaier, L E Kay *J. Biomol. NMR* **23** 35 (2002)
413. P J Bolon, J H Prestegard *J. Am. Chem. Soc.* **120** 9366 (1998)
414. T Rundlöf, C Landersjö, K Lycknert, A Maliniak, G Widmalm *Magn. Reson. Chem.* **36** 773 (1998)
415. G R Kiddle, S W Homans *FEBS Lett.* **436** 128 (1998)
416. S W Homans *Biochem. Soc. Trans.* **26** 551 (1998)
417. M Martin-Pastor, C A Bush *Biochemistry* **39** 4674 (2000)
418. M Martin-Pastor, C A Bush *Carbohydr. Res.* **323** 147 (2000)
419. J H Prestegard, V W Miner, P M Tyrell *Proc. Natl. Acad. Sci. USA* **80** 7192 (1983)
420. P Ram, L Mazzola, J H Prestegard *J. Am. Chem. Soc.* **111** 3176 (1989)
421. C R Sanders II, J H Prestegard *J. Am. Chem. Soc.* **113** 1987 (1991)
422. C R Sanders II, J H Prestegard *J. Am. Chem. Soc.* **114** 7096 (1992)
423. B J Hare, K P Howard, J H Prestegard *Biophys. J.* **64** 392 (1993)
424. Y Aubin, J H Prestegard *Biochemistry* **32** 3422 (1993)
425. B A Salvatore, R Ghose, J H Prestegard *J. Am. Chem. Soc.* **118** 4001 (1996)
426. K P Howard, J H Prestegard *J. Am. Chem. Soc.* **118** 3345 (1996)
427. K P Howard, J H Prestegard *Biophys. J.* **71** 2573 (1996)
428. C Landersjö, C Höög, A Maliniak, G Widmalm *J. Phys. Chem. B* **104** 5618 (2000)
429. H F Azurmendi, M Martin-Pastor, C A Bush *Biopolymers* **63** 89 (2002)
430. H F Azurmendi, C A Bush *Carbohydr. Res.* **337** 905 (2002)
431. T N Pham, T Liptaj, K Bromek, D Uhrin *J. Magn. Reson.* **157** 200 (2002)
432. K Lycknert, A Maliniak, G Widmalm *J. Phys. Chem. A* **105** 5119 (2001)
433. B Stevensson, C Landersjö, G Widmalm, A Maliniak *J. Am. Chem. Soc.* **124** 5946 (2002)
434. F Tian, H M Al-Hashimi, J L Craighead, J H Prestegard *J. Am. Chem. Soc.* **123** 485 (2001)
435. H Neubauer, J Meiler, W Peti, C Griesinger *Helv. Chim. Acta* **84** 243 (2001)
436. A Almond, J Ø Duus *J. Biomol. NMR* **20** 351 (2001)
437. M Martin-Pastor, C A Bush *J. Biomol. NMR* **19** 125 (2001)
438. M Staaf, C Höög, B Stevensson, A Maliniak, G Widmalm *Biochemistry* **40** 3623 (2001)
439. D I Freedberg *J. Am. Chem. Soc.* **124** 2358 (2002)

440. J H Prestegard, J Glushka, in *NMR of Glyco-Conjugates*. (Eds J Jimenez-Barbero, T Peters) (New York: Wiley, 2002)
441. H Zhou, A Vermeulen, F M Jucker, A Pardi *Biopolymers* **52** 168 (2000)
442. D MacDonald, P Lu *Curr. Opin. Struct. Biol.* **12** 337 (2002)
443. E T Mollova, A Pardi *Curr. Opin. Struct. Biol.* **10** 298 (2000)
444. L Žídek, R Štefl, V Sklenář *Curr. Opin. Struct. Biol.* **11** 275 (2001)
445. P Bayer, L Varani, G Varani *J. Biomol. NMR* **14** 149 (1999)
446. N Tjandra, S Tate, A Ono, M Kainosho, A Bax *J. Am. Chem. Soc.* **122** 6190 (2000)
447. L Trantirek, M Urbášek, R Štefl, J Feigon, V Sklenář *J. Am. Chem. Soc.* **122** 10 454 (2000)
448. H M Al-Hashimi, J R Tolman, A Majumdar, A Gorin, D J Patel *J. Am. Chem. Soc.* **123** 5806 (2001)
449. A Majumdar, D J Patel *Acc. Chem. Res.* **35** 1 (2002)
450. A Vermeulen, H Zhou, A Pardi *J. Am. Chem. Soc.* **122** 9638 (2000)
451. E T Mollova, M R Hansen, A Pardi *J. Am. Chem. Soc.* **122** 11 561 (2000)
452. K Bondensgaard, E T Mollova, A Pardi *Biochemistry* **41** 11 532 (2002)
453. H M Al-Hashimi, A Majumdar, A Gorin, A Kettani, E Skripkin, D J Patel *J. Am. Chem. Soc.* **123** 633 (2001)
454. S R Lynch, J D Puglisi *J. Am. Chem. Soc.* **122** 7853 (2000)
455. M Hennig, T Carlomagno, J R Williamson *J. Am. Chem. Soc.* **123** 3395 (2001)
456. Z Wu, N Tjandra, A Bax *J. Biomol. NMR* **19** 367 (2001)
457. B Luy, J P Marino *J. Biomol. NMR* **20** 39 (2001)
458. K Huang, J M Louis, L Donaldson, F-L Lim, A D Sharrocks, G M Clore *EMBO J.* **19** 2615 (2000)
459. J Kikuchi, J Iwahara, T Kigawa, Y Murakami, T Okazaki, S Yokoyama *J. Biomol. NMR* **22** 333 (2002)
460. D T Braddock, J M Louis, J L Baber, D Levens, G M Clore *Nature (London)* **415** 1051 (2002)
461. H M Al-Hashimi, A Gorin, A Majumdar, D J Patel *J. Am. Chem. Soc.* **123** 3179 (2001)
462. H M Al-Hashimi, A Gorin, A Majumdar, Y Gosser, D J Patel *J. Mol. Biol.* **318** 637 (2002)
463. J J Warren, P B Moore *J. Biomol. NMR* **20** 311 (2001)
464. S-H Bae, H-K Cheong, J-H Lee, C Cheong, M Kainosho, B-S Choi *Proc. Natl. Acad. Sci. USA* **98** 10602 (2001)
465. N Sibille, A Pardi, J-P Simorre, M Blackledge *J. Am. Chem. Soc.* **123** 12 135 (2001)
466. P Padrta, R Štefl, L Králík, L Žídek, V Sklenář *J. Biomol. NMR* **24** 1 (2002)
467. J R Tolman, J M Flanagan, M A Kennedy, J H Prestegard *Nat. Struct. Biol.* **4** 292 (1997)
468. L E Kay *Nat. Struct. Biol.* **5** 513 (1998)
469. J R Tolman, H M Al-Hashimi, L E Kay, J H Prestegard *J. Am. Chem. Soc.* **123** 1416 (2001)
470. L Wang, Y Pang, T Holder, J R Brender, A V Kurochkin, E R P Zuiderweg *Proc. Natl. Acad. Sci. USA* **98** 7684 (2001)
471. H M Al-Hashimi, Y Gosser, A Gorin, W Hu, A Majumdar, D J Patel *J. Mol. Biol.* **315** 95 (2002)
472. J R Tolman *J. Am. Chem. Soc.* **124** 12 020 (2002)
473. A Bax, J J Chou, B E Ramirez, in *Encyclopedia of NMR* Vol. 9 (Eds D M Grant, R K Harris) (Chichester: Wiley, 2002) p. 401
474. J H Prestegard, N U Jain, S B Levery, in *Encyclopedia of NMR* Vol. 9 (Eds D M Grant, R K Harris) (Chichester: Wiley, 2002) p. 441
475. E de Alba, N Tjandra *Prog. Nucl. Magn. Reson. Spectrosc.* **40** 175 (2002)
476. M Schneider, D E Demco, B Blümich *Macromolecules* **34** 4019 (2001)
477. C Boesch, R Kreis *NMR Biomed.* **14** 140 (2001)
478. M-L Mattinen, K Pääkkönen, T Ikonen, J Craven, T Drakenberg, R Serimaa, J Waltho, A Annila *Biophys. J.* **83** 1177 (2002)
479. Ya V Zubavichus, Yu L Slovokhotov *Usp. Khim.* **70** 429 (2001) [*Russ. Chem. Rev.* **70** 373 (2001)]
480. E L Ulrich, I L Markley, Y Kyogoku *Protein Seq. Data Anal.* **2** 23 (1989)
481. M Suzuki, R J Youle, N Tjandra *Cell* **103** 645 (2000)
482. J D Swarbrick, T Bashtannyk, D Maksel, X-R Zhang, G M Blackburn, K R Gayler, P R Gooley *J. Mol. Biol.* **302** 1165 (2000)
483. J J Chou, S Li, A Bax *J. Biomol. NMR* **18** 217 (2000)
484. D MacDonald, K Herbert, X Zhang, T Polgruto, P Lu *J. Mol. Biol.* **306** 1081 (2001)
485. C C Cornilescu, F Bouamr, X Yao, C Carter, N Tjandra *J. Mol. Biol.* **306** 783 (2000)
486. M Cai, Y Huang, R Ghirlando, K L Wilson, R Craigie, G M Clore *EMBO J.* **20** 4399 (2001)
487. S Saha, J Boyd, J M Werner, V Knott, P A Handford, I D Campbell, A K Downing *Structure* **9** 451 (2001)
488. E AB, G K Schuurman-Wolters, D Nijlant, K Dijkstra, M H Saier, G T Robillard, R M Scheek *J. Mol. Biol.* **308** 993 (2001)
489. K Schweimer, S Hoffmann, F Bauer, U Friedrich, C Kardinal, S M Feller, B Biesinger, H Sticht *Biochemistry* **41** 5120 (2002)
490. W Gronwald, F Huber, P Grünewald, M Spörner, S Wohlgemuth, C Herrmann, H R Kalbitzer *Structure* **9** 1029 (2001)
491. L Yu, A H Gunasekera, J Mack, E T Olejniczak, L E Chovan, X Ruan, D L Towne, C G Lerner, S W Fesik *J. Mol. Biol.* **311** 593 (2001)
492. E C Murphy, V B Zhurkin, J M Louis, G Cornilescu, G M Clore *J. Mol. Biol.* **312** 481 (2001)
493. D T Braddock, J L Baber, D Levens, G M Clore *EMBO J.* **21** 3476 (2002)
494. G Cornilescu, B R Lee, C C Cornilescu, G Wang, A Peterkofsky, G M Clore *J. Biol. Chem.* **277** 42289 (2002)
495. T C Leeper, M B Martin, H Kim, S Cox, V Semenchenko, F J Schmidt, S R Van Doren *Nat. Struct. Biol.* **9** 397 (2002)
496. J Weigelt, S E Brown, C S Miles, N E Dixon, G Otting *Struct. Fold. Des.* **7** 681 (1999)
497. I Bertini, A Donaire, B Jiménez, C Luchinat, G Parigi, M Piccioli, L Poggi *J. Biomol. NMR* **21** 85 (2001)
498. B Pan, B Li, S J Russell, J Y K Tom, A G J Cochran, W J Fairbrother *J. Mol. Biol.* **316** 769 (2002)
499. J Iwahara, M Iwahara, G W Daughdrill, J Ford, R T Clubb *EMBO J.* **21** 1197 (2002)
500. O N Pakhomova, S Deep, Q Huang, C Zwieb, A P Hinck *J. Mol. Biol.* **317** 145 (2002)
501. K Ye, A Serganov, W Hu, M Garber, D J Patel *Eur. J. Biochem.* **269** 5182 (2002)
502. L G Barrientos, J M Louis, I Botos, T Mori, Z Han, B R O'Keefe, M R Boyd, A Wlodawer, A M Gronenborn *Structure* **10** 673 (2002)
503. J Kuszewski, A M Gronenborn, G M Clore *J. Am. Chem. Soc.* **121** 2337 (1999)
504. E C Johnson, G A Lazar, J R Desjarlais, T M Handel *Struct. Fold. Des.* **7** 967 (1999)
505. M Cai, Y Huang, R Zheng, S-Q Wei, R Ghirlando, M S Lee, R Craigie, A M Gronenborn, G M Clore *Nat. Struct. Biol.* **5** 903 (1998)
506. B F Volkman, S J Wilkens, W M Westler, A L Lee, R Beger, B Xia, J L Markley *J. Am. Chem. Soc.* **121** 4677 (1999)
507. K Pääkkönen, T Sorsa, T Drakenberg, P Pollesello, C Tilgmann, P Permi, S Heikkinen, I Kilpeläinen, A Annila *Eur. J. Biochem.* **267** 6665 (2000)
508. J C Amor, R D Seidel III, F Tian, R A Kahn, J H Prestegard *J. Biomol. NMR* **23** 253 (2002)