IGOR ZHUKOV, ANDRZEJ EJCHART*)

Polish Academy of Sciences, Institute of Biochemistry and Biophysics
ul. Pawińskiego 5A, 02-106 Warszawa, Poland

NMR spectroscopy in structural proteomics. NMR-based protein structure determination

Summary — The pros and cons of NMR spectroscopy as a tool for the protein structure determination are discussed. Recently, the advance in the NMR equipment, spectral techniques and isotope labelling resulted in an enormous growth of NMR-determined protein structures. Modern approaches to the NMR-based protein structure determinations are based on several types of experimentally derived constraints. Short-range, distance and dihedral angle constraints are valuable, but cumulative errors can appear when successive constraints are used to determine spatial relationship of remote parts of a protein. Therefore, long-range constraints derived from residual dipolar couplings and nuclear relaxation data of anisotropically tumbling molecules are highly complementary to the short-range constraints.
Key words: NMR spectroscopy, protein structure in solution, NMR structural constraints.

Proteins play crucial roles in virtually all biological processes. They transmit the messages, repair a damage, provide the building blocks for tissues and carry out reactions essential for a life. Nowadays it is well documented that three-dimensional structure of proteins determines their function and interactions with other proteins, nuclei acids, small ligands and ions. Structural proteomics refers to the information about three-dimensional structures of a significant fraction of the proteins encoded by a given genome. Nowadays it is expected that the human genome alone encodes about 300 000 proteins. Together with proteins produced in pathogenic organisms or agriculturally important plants it makes an enormous number of structures to be determined in structural proteomics programs. Since the number of peptide and protein structures currently deposited in Protein Data Bank (PDB) approaches 16 000 a lot of effort should be placed in the field of structural proteomics [1].

NMR SPECTROSCOPY AND STRUCTURE DETERMINATION OF BIOPOLYMERS

Nuclear magnetic resonance spectroscopy (NMR) is well suited to play an important role in proteomics programs. This method provides structural information at the atomic resolution. Isotopes of biologically important elements (^{1}H, ^{13}C, ^{15}N, and ^{31}P) display narrow resonance lines despite the uselessness of the most abundant nitrogen (^{14}N) and carbon (^{12}C) isotopes in the NMR studies of macromolecules. Internuclear interactions, modulated by even small structural and conformational changes, influence line shape and intensity of signals in NMR spectra. Last but not least, NMR provides high resolution structures in solution allowing to study those proteins that fail to crystallize or to compare differences between their crystal and solution structures. Potential of NMR method, however, has not been reflected by the present number of deposited structures; only 14% of the PDB structures have been determined by NMR spectroscopy.

The following factors have hampered a broad use of NMR in the determination of three-dimensional structures of biopolymers:

— Low sensitivity of NMR spectroscopy. A typical amount of protein used in NMR studies is *ca.* 0.5 μM, several orders of magnitude more than in mass spectrometry or optical spectroscopy.

— Signals in spectra of individual isotopes of biopolymers are usually strongly superposed. For instance, in a protein built up of 200 amino acid residues one can expect *ca.* 1200 ^{1}H signals, 1000 ^{13}C signals, and more than 200 ^{15}N signals.

— In large proteins fast transverse nuclear relaxation brings about line broadening that aggravates superposition of signals and eliminates their fine structure.

— Strong solvent signal generates dynamic range problems in ^{1}H spectra ($S_{water}/S_{protein} \approx 10^{5}$) and obscures an important spectral region of H_{α} signals.

*) To whom all correspondence should be addressed.

— NMR derived data determine ambiguous structural constraints leaving a number of possible solutions.

— Since NMR measurements are long lasting, the protein stability in solution at room temperature is required for an extended period.

Recently, all these obstacles have been mostly overcome due to the progress in spectrometer designing (higher permanent magnetic fields generated in cryomagnets, pulse field gradient equipment, cryoprobes), isotope labelling (uniform, selective or segmental), and new spectral techniques (multidimensional and multinuclear spectra, TROSY and CRINEPT techniques)*).

The key aims of isotope labelling are to increase sensitivity, to eliminate signal overlap and to narrow linewidths [2, 3]. ^{15}N and ^{13}C labelling simplifies the spectral assignments and provides new types of structural constraints. On the other hand, deuteration allows to reduce linewidths, to eliminate partially or totally ^{1}H signals and to reduce parasitical spin diffusion effect. Uniform $^{15}N/^{13}C$ double labelling has been successfully used in the identification of signals in the spectra of ^{1}H, ^{13}C and ^{15}N isotopes. Selective $^{1}H/^{13}C$ labelling of methyl groups, in otherwise deuterated proteins, facilitates identification of side chain — side chain dipolar interactions crucial in the determination of the hydrophobic core conformation. Segmental isotope labelling allows to decrease the number of signals in NMR spectra thus reducing signal overlap. Such proteins are produced through the ligation of labelled and unlabelled polypeptide chains in a self-catalytic protein splicing process [4].

Introduction of multidimensional spectroscopy exerted a strong impact on the biomolecular NMR [5, 6]. In the traditional one-dimensional spectra only small molecules with very limited number of nonequivalent nuclei do not show signal overlap. In biopolymers containing numerous repeating subunits, extensive superposition of spectral lines precludes their unequivocal identification and structural assignment. In multidimensional spectroscopy each nucleus is identified not only by its own resonance frequency but also by frequency or frequencies of nuclei interacting with it. Since usually resonance frequencies of interacting nuclei are weakly correlated, a probability of occurrence of identical sets of resonance frequencies characterizing a given nucleus is virtually eliminated. This feature is particularly well noticeable in the heteronuclear multidimensional spectra. For instance, the signal overlap in one-dimensional ^{1}H and ^{15}N spectra is usually removed in the two-dimensional $^{1}H/^{15}N$ correlation as shown in Fig. 1.

With increasing molecular size, rotational diffusion slows down resulting in faster and faster transverse nuclear relaxation and, thus, line broadening, which causes loss of sensitivity and spectral resolution. TROSY technique largely suppress these undesired effects selecting only the narrowest, most slowly relaxing component of scalar coupled multiplet [7].

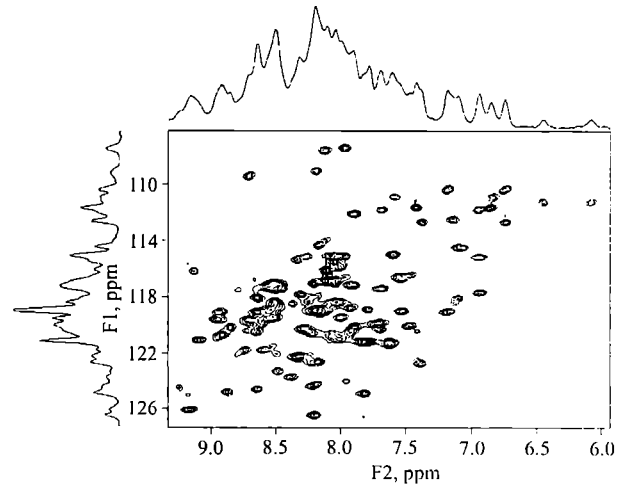
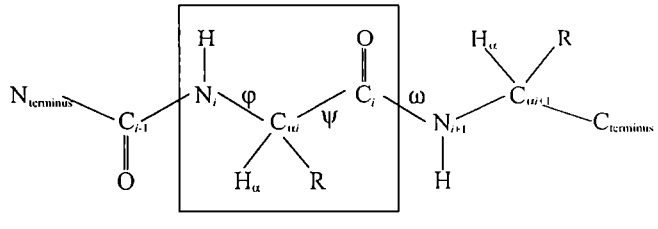


Fig. 1. Two-dimensional NMR spectrum of S100A1 protein correlating ^{1}H (horizontal axis) and ^{15}N (vertical axis) nuclei in amide groups shows very good spectral dispersion. Each of 93 cross peaks corresponds to the individual amide group. On the other hand, signals in both one-dimensional spectra (traces above and to the left of the spectrum) are strongly overlapped

In heteronuclear NMR techniques, magnetization is transferred between nuclei of isotopes *via* scalar couplings applying pulse sequence called INEPT [8]. Minimal duration of INEPT sequence is determined by scalar coupling values and fast transverse relaxation inherent in large molecules significantly deteriorates its efficiency. The CRINEPT technique [9] overcomes this limitation by combining INEPT with CRIPT sequence. The latter allows to compensate increasing transverse relaxation with shorter duration of transfer sequence.

STRUCTURE OF PROTEINS

Amino acids are the basic structural units of proteins. Twenty different amino acids are commonly found in proteins. They are linked by peptide bonds to form polypeptide chains (Fig. 2). The sequence of amino acids



i-th amino acid residue

Fig. 2. Repeating unit of the polypeptide chain; amino acids differ in the structure of the side chain R (φ, ψ, ω — see text)

*) TROSY = transverse relaxation-optimized spectroscopy; CRINEPT = CRIPT + INEPT; CRIPT = cross relaxation-induced polarization transfer, INEPT = insensitive nuclei enhanced by polarization transfer.

determines primary protein structure. Secondary structure refers to the conformational arrangement of the backbone segments of a polypeptide chain without regard to the conformation of the side chains or relationship to other segments. It is defined by triads of backbone dihedral angles (φ, ψ, ω). Among secondary structure elements one can distinguish periodic structures stabilized by hydrogen bonds, as α-helices or β-strands. Tertiary structure describes the spatial organization of an entire protein molecule consisting of a single chain. Proteins that contain more than one polypeptide chain, so-called multimeric proteins, additionally display quaternary structure. It describes the spatial organization of two or more chains with tertiary structure held together by hydrogen bonds, van der Waals, and electrostatic forces. All these levels of protein structure are shown in Fig. 3.

GSELETAMETLINVFIIAHSGKEGDKYKLSKKELKELLQ
TELSGFLDAQKDADAVDKVMKELDEDGDGEVDFQEY
VVLVAALTVACNNFFWENS

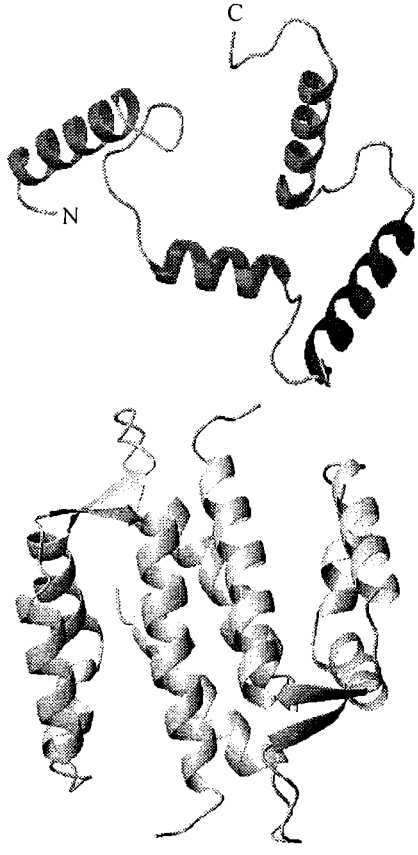


Fig. 3. Different levels of the protein structure are shown for the homodimeric S100A1 protein. Its polypeptide chain is built up of 93 amino acid residues listed in the upper part (primary structure). Basic secondary structure elements, four long α-helices, without regard to their relative orientation are shown in the central part (secondary structure). In the lower part relative orientation of secondary structure elements within subunit (tertiary structure) as well as relative orientation of two subunits (quaternary structure) are shown

STRATEGIES TO THE NMR-BASED PROTEIN STRUCTURE DETERMINATION

NMR-based procedure of structure determination comprises three stages: assignment of as many signals as possible in the spectra of NMR-active isotopes, identification of structural constraints, and calculation of a family of three-dimensional structures fulfilling experimental constraints. Approaches used at first two stages depend on the size of protein studied whereas approach applied at the third stage depends on the type and number of identified constraints.

Small proteins with molecular weight $MW < 10$ kDa are usually studied using two-dimensional (2D) ^{1}H NMR spectra solely [10, 11]. In the first step spin systems of similar topology are identified from the correlations utilizing ^{1}H-^{1}H scalar couplings. Sequential assignment of spin systems is based on the short range dipolar interactions identified in 2D nuclear Overhauser effect (NOESY) spectra*). Next, constraints characterizing secondary structure elements are elucidated from medium range NOEs and vicinal scalar couplings, $^3J(H_NH_\alpha)$. Finally, long range NOEs and hydrogen bonds are used for the tertiary structure determination.

For proteins with $MW > 10$ kDa two simultaneous limitations occur; progressive overlap of signals and their broadening owing to fast transverse nuclear relaxation [5, 11]. Usually proteins with $MW < 30$ kDa are uniformly ^{15}N/^{13}C double labelled in order to take advantage of heteronuclear, multidimensional techniques. They allow to replace NOE based sequential assignment with correlations transmitted through heteronuclear scalar couplings, which are more sensitive and display better signal dispersion. Moreover, correlations in NOESY spectra can be spread out due to ^{15}N and/or ^{13}C editing [11, 12]. Additionally, combined information on ^{1}H, ^{13}C and ^{15}N resonance frequencies of individual backbone nuclei can be used in such statistical methods of secondary structure determination as chemical shift index [13] or TALOS software [14]. These methods become insufficient for proteins with $MW > 30$ kDa. In very large proteins and protein assemblies, signal overlap can be further diminished by selective or segmental isotope labelling. On the other hand, relaxation based line broadening and loss of sensitivity can be limited by deuteration and by application of TROSY and CRINEPT techniques [15].

NMR DERIVED STRUCTURAL CONSTRAINTS

Three dimensional structure of any molecule built up of n atoms is unequivocally determined by $3n$-6 internal coordinates, interatomic distances, valence angles, and dihedral angles. It is usually assumed that distances be-

*) NOESY = nuclear Overhauser effect spectroscopy.

tween directly bound atoms are well represented by bond lengths and the valence angles by their standard values. On the other hand, experimentally derived structural constraints are required for the determination of dihedral angles. Experimental constraints are also important for the verification of other internal coordinates. NMR spectroscopy can be a source of several types of structural constraints.

1. Interproton distances can be determined quantitatively or semiquantitatively from the nuclear Overhauser effect (NOE).

2. Hydrogen bond donor-acceptor distances between amino acid residues remote in the sequence can be qualitatively evaluated when scalar couplings *via* hydrogen bonds are detected.

3. Dihedral angles can be evaluated from vicinal scalar couplings.

4. Relative orientations of internuclear vectors can be calculated from the interference of nuclear relaxation mechanisms.

5. Orientations of a given type of internuclear vectors in the molecular reference frame can be derived from residual dipolar couplings or nuclear relaxation data of anisotropically tumbling macromolecules.

The constraints 1—4 are local, short-range ones and cumulative errors can appear when successive constraints are used to determine spatial relationship of remote parts of a macromolecule. On the other hand, type 5 constraints provide long-range orientation that is highly complementary to short-range constraints. Accuracy of structure determination strongly depends on the available number of constraints and, therefore, the most important objective in any NMR-based structure determination is to obtain the maximum number of NMR constraints [16].

Ad 1. The NOE arises due to cross relaxation taking place during appropriately designed and performed experiment, which results in the transfer of magnetization between protons close together in space [17]. In order to disperse superposed signals in the crowded NMR spectra of large proteins, multidimensional NOE spectroscopy is routinely used. NOESY spectra of medium and large proteins usually show strong signal overlap. This aspect can be visualized when corresponding parts of NOESY spectra (two-dimensional NOE) of proteins differing in size are compared (Fig. 4). To overcome this problem ^{15}N and/or ^{13}C edited 3D/4D NOESY spectra of respectively labelled proteins should be measured [6]. The most important dependence of the cross relaxation rate σ_{ij} between two protons i and j from a structural standpoint is that on the inverse sixth power of the internuclear distance, $\sigma_{ij} \sim r_{ij}^{-6}$. Therefore, the relative intensities of NOE cross peaks in NOE spectra can be used to quantify internuclear distances providing the cross peak(s) between protons of known separation were identified and used for the distance calibration purpose. In more conservative approach NOE cross peaks are divided according to intensity into three groups, namely strong (0.18 nm < r < 0.25 nm), medium (0.18 nm < r < 0.35 nm), and weak (0.18 nm < r < 0.50 nm). In practice, the maximum distance so available is *ca.* 0.5 nm [16]. Larger distances are usually influenced by spin diffusion, a multistep magnetization transfer [18], which can lead to incorrect internuclear distances and hence to imposition of tighter interproton distance constraints than is justified. These problems can be circumvented when complete relaxation matrix methods allowing for the spin diffusion are used [16]. One should add that a single internuclear distance determines a sphere and as many as four distances are required to remove the spatial ambiguity.

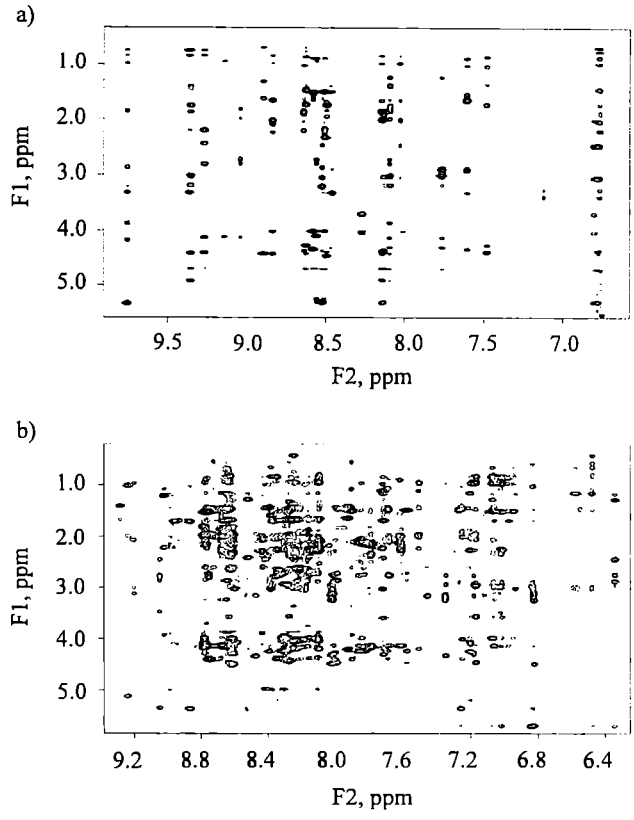


Fig. 4. Cross peaks reflecting dipolar interactions among H_N protons (horizontal axis) and H_α and aliphatic side chain protons (vertical axis) in NOESY spectra are shown for two proteins. In the upper part (a) a fragment of NOESY spectrum of small, 29 amino acid residue protein CMTI-I (M8L), shows no cross peak overlap. The lower part (b) displaying corresponding fragment of NOESY spectrum of larger, 93 amino acid residue protein S100A1, shows extensive cross peak overlap precluding full identification of interacting protons

Ad 2. Hydrogen bonds are of key importance for protein structures stabilizing. The presence of hydrogen bonds indicates the spatial proximity and relative arrangement of the atoms involved. Direct evidence for the existence of hydrogen bonds can be established in proteins by the observation of scalar couplings between an amide ^{15}N and carbonyl ^{13}C nuclei of two residues,

^{15}N-H···O=^{13}C across a hydrogen bond [19, 20]. Thus, the detection of scalar coupling through hydrogen bond unambiguously imposes a valuable distance constraint at the stage of the backbone assignment in a protein: 0.18 nm < $d(H_N···O)$ < 0.20 nm and 0.27 nm < $d(N···O)$ < 0.30 nm. This method requires the use of ^{15}N/^{13}C double labelled proteins.

Ad 3. In conformational studies, the Karplus relation between vicinal (through three bonds) scalar couplings, ^3J, and dihedral angles, φ, is of great importance [21]. This relation can be represented by the general formula:

$$^3J = A \cdot cos^2\varphi + B \cdot cos\varphi + C \quad (1)$$

Coefficients *A*, *B*, and *C* depend on the variety of molecular parameters. Among them, the type of elements forming the central bond as well as electronegativity and relative position of their substituents are most important [22]. Therefore, for a given class of molecules, empirical calibration of the coefficients derived from measurements of model compounds has so far used. In structural studies of proteins, scalar couplings determining the φ backbone angle have been the most widely used [23]. Six homo- and heteronuclear scalar couplings are related to this dihedral angle and corresponding experimental φ-dependent Karplus curves are shown in Fig. 5. Owing

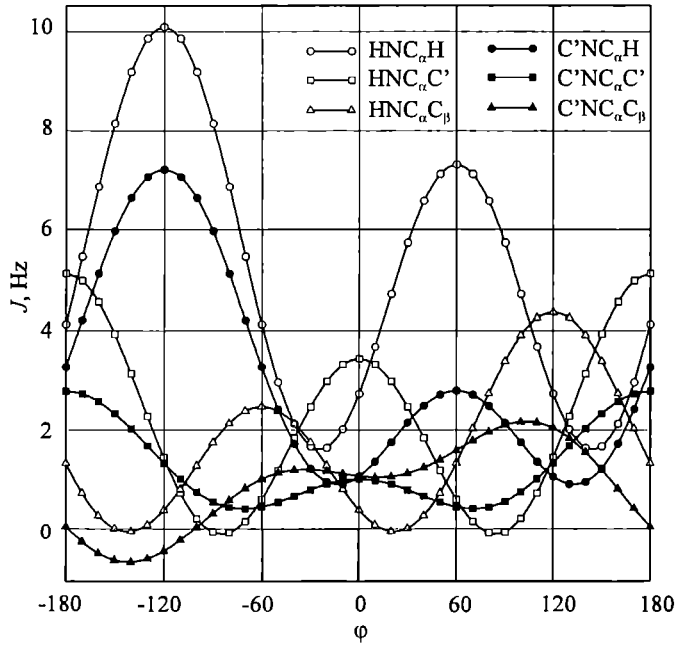


Fig. 5. Experimental Karplus dependencies for all six vicinal scalar couplings across backbone N—Cα bond determining dihedral backbone φ angle

to the periodicity of Karplus equation a single value of vicinal coupling constant can correspond to as many as four different dihedral angles introducing an ambiguity to the scalar coupling based constraints. However, an

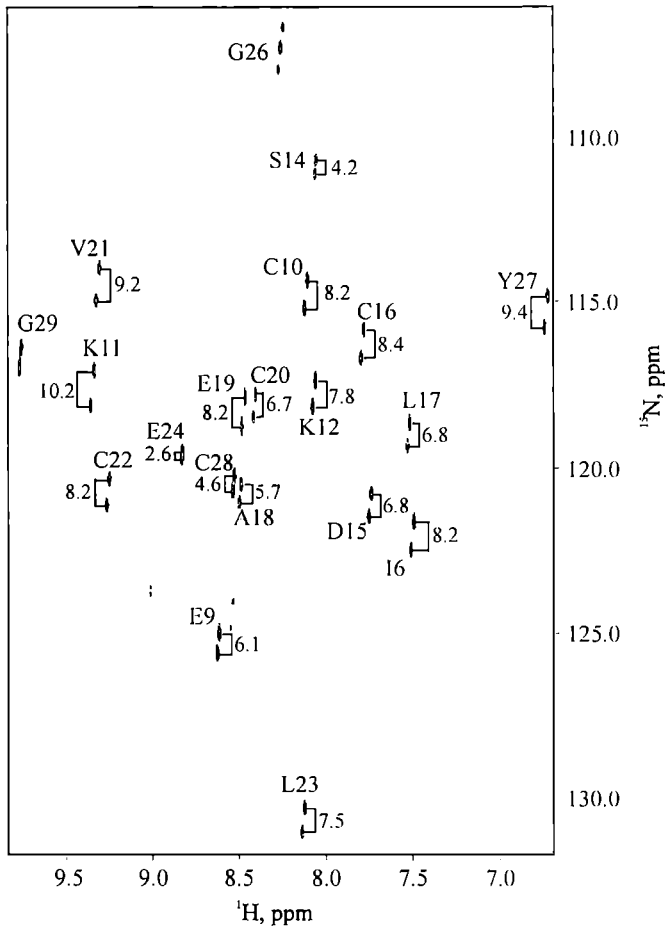


Fig. 6. PPJ-HMQC spectrum [24] measured for the ^{15}N labelled CMTI-I(M8L) protein (PPJ-HMQC = pure phase homonuclear J-modulated heteronuclear multiple quantum coherence). Homonuclear $^3J(H_NH_\alpha)$ scalar couplings result in the splitting of correlation signals along vertical axis. Sequential assignments and scalar coupling values are given in spectrum

appropriate dihedral angle can be derived from a combination of several scalar couplings. Isotopic labelling, which is required when heteronuclear scalar couplings are measured, facilitates the determination of homonuclear scalar couplings as well (Fig. 6) [24].

Ad 4. Two mechanisms dominating the relaxation of heteronuclei in proteins, namely dipol-dipol mechanism (DD) and chemical shift anisotropy mechanism (CSA), can interfere one another [25]. Arising cross correlation terms together with auto-correlation terms contribute to the total nuclear magnetic relaxation. Recently interference effects of different DD mechanisms or DD and CSA mechanisms or different CSA mechanisms in ^{15}N/^{13}C double labelled proteins have been used to determine angles between vectors characterizing cross correlated mechanisms [26, 27]. In turn, these angles can be related to the backbone dihedral angles. For instance, determination of the backbone dihedral angle φ_i was obtained from the interference of DD($C_{\alpha i}H_\alpha$) and DD($N_{i+1}H$) or

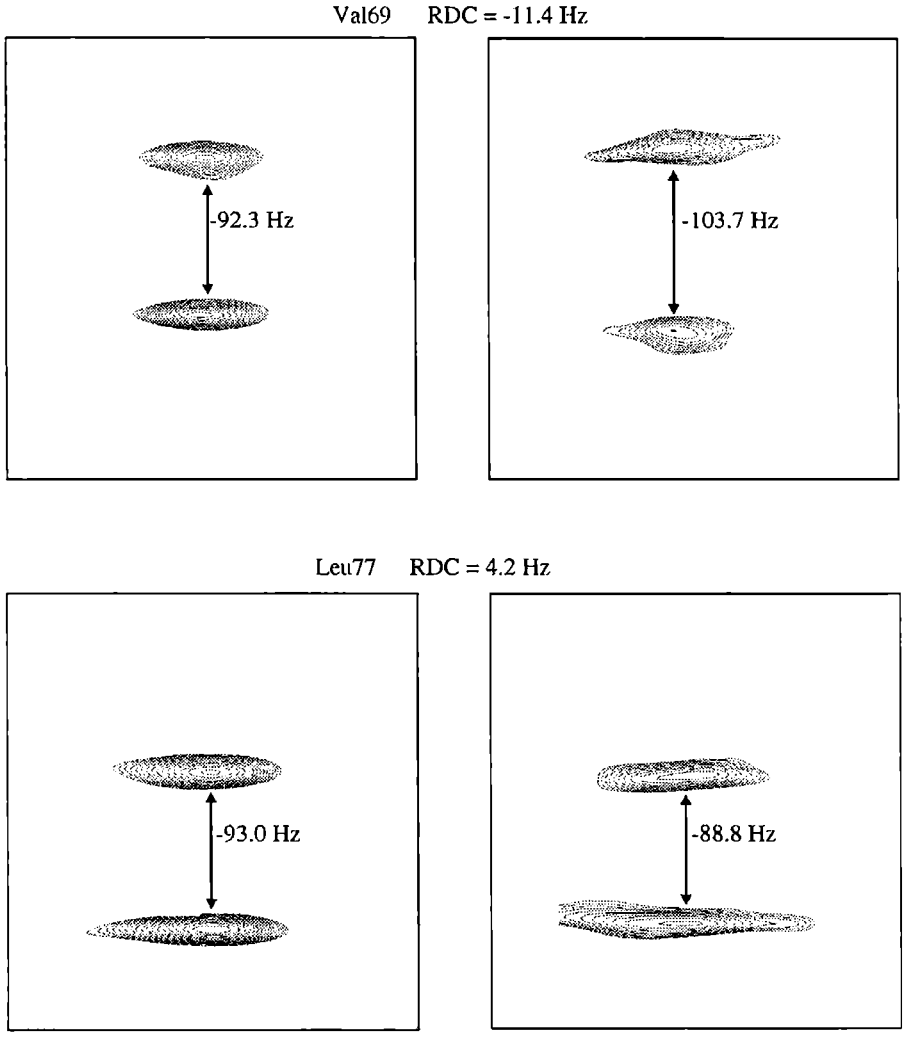


Fig. 7. Examples of the determination of residual dipolar couplings (RDC) in the $^{15}N/^{13}C$ double labelled S100A1 protein. The $^{1}H/^{15}N$ correlations for the backbone amide groups of Val69 [$\delta(^{1}H)$= 9.15 ppm; $\delta(^{15}N)$= 116.2 ppm] and Leu77 [$\delta(^{1}H)$= 7.42 ppm; $\delta(^{15}N)$= 122.8 ppm] were chosen. Left-side figures display fragments of $^{1}H/^{15}N$ correlation spectrum measured in isotropic solution: vertical splittings in the ^{15}N dimension correspond to $^{1}J(NH)$ scalar couplings. Right-side figures display the same fragments of spectrum obtained in the anisotropic bicelle solution. RDCs are obtained from the difference of two splittings

$DD(C_{\alpha i}H_{\alpha})$ and $CSA(C'_i)$ mechanisms. In general, as many as four dihedral angle values can correspond to a single interference relaxation rate. When more than one interference rate related to a given dihedral angle is available, the ambiguity can be reduced or removed [28].

Ad 5. Orientational information relative to a common molecular reference frame can be obtained from residual dipolar couplings (RDC) owing to partial molecular alignment [29—31] or from heteronuclear relaxation in anisotropically tumbling molecules [32, 33]. The partial alignment of proteins can be induced by solvation in dilute anisotropic media such as phospholipid bicelles [34], filamentous phage [35] or strained gels [36]. It prevents complete averaging of dipolar interactions as in isotropic solution. The direct measurement of the RDCs provides long range orientational information for internuclear vectors positioned throughout the studied macromolecule. Similar information is provided by heteronuclear relaxation parameters in anisotropically tumbling molecules. Geometric dependence of RDC values on the orientation of internuclear vectors determining dipolar interaction relative to the order matrix is similar to the dependence of relaxation parameters on the orientation of specific relaxation vectors relative to the diffusion tensor. Identical experimental values are distributed on two elliptic cone surfaces. Measurements of RDCs as well as relaxation parameters require labelled proteins.

An example of the determination of RDC values is shown in Fig. 7. It is noteworthy that multitude of different vectors in proteins whose orientations are available from RDCs allows to improve greatly both the precision and accuracy of solution structures of proteins and their complexes. Recently the determination of protein backbone conformation using only RDCs constraints has been reported [37].

CONCLUSIONS

Recent advance in the NMR equipment and novel multidimensional NMR techniques combined with new isotope labelling strategies of proteins greatly facilitate the study of a wide range of proteins and protein complexes. Three-dimensional structures, their dynamics, characterization of conformational changes as well as supramolecular interactions become accessible from NMR studies [38]. One can expect that high quality NMR spectra of proteins well beyond the present size limit of *ca.* 100 kDa will become a reality soon.

REFERENCES

1. Berman H. M.: *Nucleic Acid Res.* 2000, **28**, 235; http://www.rcsb.org/pdb/
2. Goto N. K., Kay L. E.: *Curr. Opin. Struct. Biol.* 2000, **10**, 585.
3. Lian L.-Y., Middleton D. A.: *Prog. Nucl. Magn. Reson. Spectrosc.* 2001, **39**, 171.
4. Yamazaki T. *et al.*: *J. Am. Chem. Soc.* 1998, **120**, 5591.
5. Bax A., Grzesiek S.: *Acc. Chem. Res.* 1993, **26**, 131.
6. Clore G. M., Gronenborn A. M.: *Protein Sci.* 1994, **3**, 372.
7. Pervushin K. *et al.*: *Proc. Natl. Acad. Sci. USA* 1997, **94**, 12 366.
8. Bodenhausen G., Ruben D. J.: *Chem. Phys. Letters* 1980, **69**, 185.
9. Riek R. *et al.*: *Proc. Natl. Acad. Sci. USA* 1999, **96**, 4918.
10. Wüthrich K.: "NMR of Proteins and Nucleic Acids", Wiley, New York 1986.
11. Cavanagh J. *et al.*: "Protein NMR Spectroscopy", Academic Press, New York 1996.
12. van de Ven F. J. M.: "Multidimensional NMR in Liquids", VHC, New York 1995.
13. Wishart D. S. *et al.*: *J. Biomol NMR* 1995, **5**, 67.
14. Cornilescu G. *et al.*: *J. Biomol NMR* 1999, **13**, 289.
15. Wider G., Wüthrich K.: *Curr. Opin. Struct. Biol.* 1999, **9**, 594.
16. Dyson H. J., Wright P. E.: "Protein Structure Calculation Using NMR Restraints" in "Two-Dimensional NMR Spectroscopy" (Eds. Croasmun W. R., Carlson R. M. K.), VCH, New York 1994.
17. Neuhaus D., Williamson M. P.: "The Nuclear Overhauser Effects in Structural and Conformational Analysis", VHC, New York 1989.
18. Kalk A., Berendsen H. J. C.: *J. Magn. Reson.* 1976, **24**, 343.
19. Cordier F., Grzesiek S.: *J. Am. Chem. Soc.* 1999, **121**, 1601.
20. Cornilescu G. *et al.*: *J. Am. Chem. Soc.* 1999, **121**, 2949.
21. Karplus M.: *J. Am. Chem. Soc.* 1963, **85**, 2870.
22. Imai K., Osawa E.: *Magn. Reson. Chem.* 1990, **28**, 668.
23. Ejchart A.: *Bull. Pol. Ac.: Chem.* 1999, **47**, 1.
24. Koźmiński W.: *J. Magn. Reson.* 1999, **141**, 185.
25. Engelke J., Rüterjans H.: "Recent Developments in Studying the Dynamics of Protein Structures from ^{15}N and ^{13}C Relaxation Time Measurements" in "Biological Magnetic Resonance" vol. 17 (Eds. Rama Krishna N., Berliner L. J.), Kluver/Plenum, New York 1999, p. 357.
26. Reif B. *et al.*: *Science* 1997, **276**, 1230.
27. Yang D. *et al.*: *J. Am. Chem. Soc.* 1997, **119**, 11 938.
28. Yang D., Kay L. E.: *J. Am. Chem. Soc.* 1998, **120**, 9880.
29. Prestegard J. H. *et al.*: "Protein Structure and Dynamics from Field-Induced Residual Dipolar Couplings" in "Biological Magnetic Resonance" vol. 17 (Eds. Rama Krishna N., Berliner L. J.), Kluver/Plenum, New York 1999, p. 311.
30. Tolman J. R.: *Curr. Opin. Struct. Biol.* 2001, **11**, 532.
31. Brunner E.: *Concepts Magn. Reson.* 2001, **13**, 238.
32. Tjandra N. *et al.*: *Nature Struct. Biol.* 1997, **4**, 443.
33. Clore G. M., Gronenborn A. M.: *Proc. Natl. Acad. Sci. USA* 1998, **95**, 5891.
34. Tjandra N., Bax A.: *Science* 1997, **278**, 1111.
35. Hansen M. R. *et al.*: *Nature Struct. Biol.* 1998, **5**, 1065.
36. Sass H.-J. *et al.*: *J. Biomol. NMR* 2000, **18**, 303.
37. Prestegard J. H. *et al.*: *Biochemistry* 2001, **40**, 8677.
38. Hus J.-C. *et al.*: *J. Am. Chem. Soc.* 2001, **123**, 1541.