KRZYSZTOF KAZIMIERCZUK, MARIA MISIAK, ANNA ZAWADZKA, WIKTOR KOŹMIŃSKI*)

Warsaw University Chemistry Department
ul. Pasteura 1, 02-093 Warszawa, Poland

Progress in structural studies of proteins by NMR spectroscopy**)

Summary — In this review we discuss recent advances in the field of NMR structural studies of proteins. We mention the development of new spectrometer hardware, novel experimental techniques and sample preparation. The approaches improving obtained structures, like analysis of Residual Dipolar Couplings, and enabling studies of large molecules as for example TROSY and CRINEPT are compared. As the most important, achievements in the field of fast multidimensional NMR are discussed in details.
Key words: NMR structure determination, proteins, multidimensional NMR.

POSTĘP W BADANIACH STRUKTURALNYCH BIAŁEK ZA POMOCĄ SPEKTROSKOPII NMR
Streszczenie — Praca zawiera przegląd ostatnich osiągnięć dotyczących badań strukturalnych białek metodami spektroskopii NMR. Omówiono postęp w konstrukcji spektrometrów, nowe techniki eksperymentalne i sposoby przygotowania próbek. Wymieniono nowe procedury analizy wyników eksperymentalnych pozwalające uzyskać lepszą jakość obrazu badanych struktur (np. analiza resztkowych sprzężeń dipolowych), a także umożliwiające badania coraz większych molekuł (np. TROSY i CRINEPT).
Słowa kluczowe: badania strukturalne metodami NMR, białka, wielowymiarowy NMR.

For the last twenty years NMR spectroscopy has become one of the leading techniques for studying biomolecular dynamics and structure. Among the biopolymers, proteins are the most popular target for NMR experiments. The amount of structures in the largest structural database, PDB (Protein Data Bank) [1], is growing faster each year and by 2006 has reached total number of six thousand. It is about 17 % of all PDB structures. Although it is still much less than crystalographers can obtain, the NMR experiments allow observation of molecules in solution, their dynamics, interactions and chemical reactions, which is not possible by X-ray technique [2—5].

At first strategies for structure determination involved only two-dimensional homonuclear correlation experiments, based on measurement of proton signal modulated by frequencies of other hydrogen nuclei. Considering the type of interaction that causes correlation of proton frequencies, one can choose between techniques where correlation is caused by scalar coupling *via* chemical bonds, and experiments based on distance-dependent nuclear Overhauser effect (NOE).

However, the strategies mentioned above were sufficient only for very small proteins. The larger ones gave very crowded proton-proton spectra causing problems with interpretation. Addition of another dimension to homonuclear experiments reduced these problems, allowing structural studies of proteins of molecular weights up to 5000 Da [6, 7]. The further progress was obtained by $^{15}N/^{13}C$ isotope labeling technologies, which become well elaborated in 1980s. The isotopically enriched samples allowed for heteronuclear edited experiments, however the most important progress was achieved by triple-resonance NMR experiments based on heteronuclear magnetization transfer. This kind of experiments [8—11] enabled effective backbone and side chains nuclei assignment, and further increased molecular weight limit above 25 000 Da. Deuterium labeling was also used to improve protein NMR by increasing transverse relaxation times of amide protons and carbon nuclei, extending molecular weight limit of studied proteins well above 30 000 Da.

Another milestone in biomolecular NMR was employment of residual dipolar couplings (RDCs) as constraints for structure determination.

One of the main breakthroughs in the NMR pulse sequence improvement in last decade was development of transverse relaxation-optimized spectroscopy (TROSY) [12] and cross-relaxation enhanced polarization transfer (CRINEPT) [13]. Both techniques enable to study very large biomolecules.

Improvement in experimental techniques is not limited to pulse sequences. One of the goals of modern NMR is to shorten days-long experiment times required for well resolved spectra and wide spectral band. This can

*) Author for correspondence; e-mail: kozmin@chem.uw.edu.pl
**) Szkoła spektroskopii NMR.

be done in many ways, from making low dimensional projections of full-dimensional spectra to employing sophisticated statistical methods. They will be presented below, together with our method, based on one-step multidimensional Fourier transform.

Shortening of experiment time always decreases signal-to-noise ratio. The noise level can only be decreased by hardware improving. Nowadays, new superconducting magnets can generate magnetic fields of induction up to B_0 = 22.2 T (950 MHz of proton Larmor frequency). Improved electronic consoles contain high frequency analog-to-digital converters (allowing even "quadrature-less" measurements). Radio-frequency (RF) pulse synthesizers work with shorter unwanted "dead--times" and RF channels are better separated from each other. Introduction of cryoprobes containing helium--cooled detection coils and pre-amplifiers allowed the reduction of thermal noise even four-fold for solutions of small electric susceptibility. In the most of protein samples the electric susceptibility is quite high, and the effect is smaller. Negative high ionic strength effect can be neglected by using special, rectangular tubes, thinner than standard ones in the direction of electric part of RF field.

The process from NMR data to structure is independent on spectrometer hardware, however it still requires experience and is time consuming. This process could be simplified by the automatic methods which are still under development [15].

CRINEPT AND TROSY METHODS

Spectral lines in NMR are described by Lorentzian function of linewidth proportional to transverse relaxation rate. It means that rapidly decaying signal gives broad lines of low intensity, which may, in some cases, even disappear under the noise level. Relaxation occurs not only during signal measurement, but also during pulse sequence delay times. There are two main interactions causing nuclear relaxation in proteins: dipole-dipole interactions and chemical shift anisotropy (CSA). The effective relaxation is related to the field strength in both cases. For dipolar mechanism, outside extreme narrowing limit, it depends also on molecular weight, while the relaxation rates due to CSA increase proportionally to square of the B_0 field. In the case of larger proteins these effects cause very fast relaxation, making standard experiments impossible.

Considering the system of two interacting coupled spins I and S it may be proved that transverse relaxation rates of two possible spectral transitions for each of them are different due to the effects of interference between dipolar interaction and chemical shift anisotropy. In standard multidimensional experiments, like two-dimensional ^{15}N-^{1}H HSQC (heteronuclear single quantum correlation) decoupling is used in both dimensions, so that only one line of average relaxation rate is present (Fig. 1a). Without decoupling, four lines appear (Fig. 1b), each corresponding to one of transitions, and each having different linewidth because of difference in relaxation rates. In TROSY experiments the effects of CSA and dipole-dipole interactions are mutually cancelled for the narrowest line, while their effect is summed for broadest (anti-TROSY) line, two other multiplet components exhibit broadening in one and narrowing in the second dimension (Fig. 1c). The strongest effect of differentiation of individual relaxation rates of each multiplet component is predicted to have a maximum for the ^{1}H frequency range about 1.1 GHz for amide protons, but is approached and observed at available frequencies of 700—800 MHz.

The same effect is employed to avoid the loss of signal during long magnetization transfers. Appropriate method is called CRINEPT and is used instead of conventional INEPT sequence. Also in this case no decoupling is used. Combination of TROSY and CRINEPT allows one to study proteins of molecular weight above 100 000 Da, where the total effect makes sensitivity even two orders of magnitude better, but even for smaller proteins (with molecular weight *ca.* 25 000 Da) gain in sensitivity is two- or threefold.

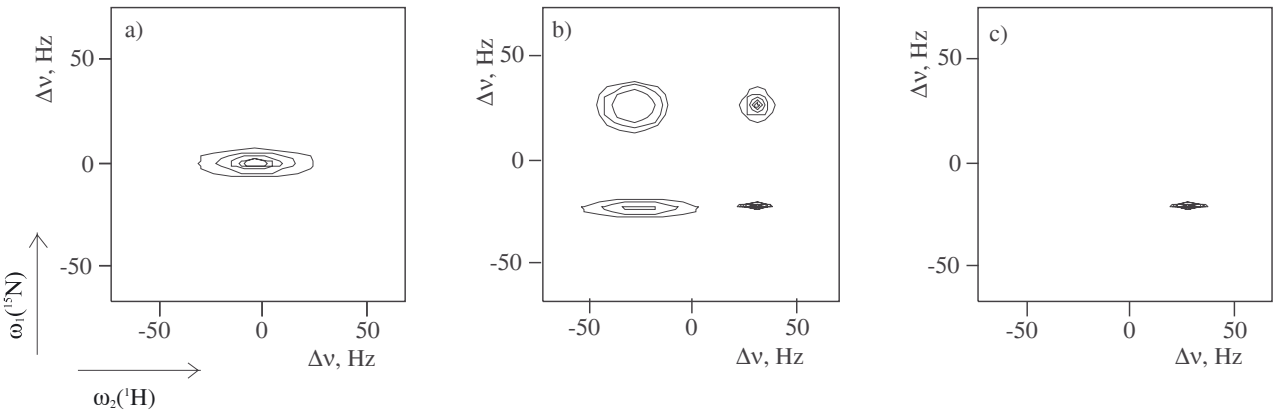

Fig. 1. Schematic illustration of the TROSY principle: a) ^{1}H-^{15}N HSQC with decoupling both time domains, b) the same without decoupling, four different multiplet components are observed, c) TROSY spectrum (narrowest line is selected)

ISOTOPE LABELING

As mentioned before, one of the greatest breakthroughs in protein NMR was development of structural determination strategies based on heteronuclear multidimensional experiments in late 1980s. Isotope labeling techniques employing transformed *E. Coli*, yeasts or mammalian cells become a routine. The ^{15}N, ^{13}C-labeled glucose, acetic acid, methanol, ammonium salts and amino acids are used as precursors, sometimes in mixtures with non-labeled compounds, if non-uniform labeling is needed. The best modern labeling method, however, is based on cell-free synthesis [16]. It may be used for site-specific labeling, valuable for larger proteins in solution and for protein solid state NMR. Site-specific labeling is used often to decrease a number of signals in a crowded spectra. Multidomain protein may be enriched sequentionally, domain after domain, and structure of each part may be determined separately, being the only observable part of the whole protein [17, 18]. This simple approach allows one to avoid overcrowding of spectra of larger molecules but will not solve the problem of high relaxation rates for them.

High relaxation rate, caused by dipole-dipole interaction may be avoided not only by TROSY/CRINEPT procedures described above, but also by deuteration. Deuteration leads to reduction of linewidths of signals from remaining ^{13}C and ^{1}H, ^{15}N nuclei. However, high level of deuteration means low proton concentration. In the case of completely deuterated proteins only amide and hydroxyl protons are observable, which is enough for NOE measurements but the information about residues is lost. On the other hand, medium deuteration and random fractional labeling lead to splitting of carbon signals because of strong isotope effect. Thus, more sophisticated methods, like stereoselective deuteration seem to be promising [19].

RESIDUAL DIPOLAR COUPLINGS

In the NMR spectrum recorded in isotropic phase all interactions are averaged due to fast random molecular reorientations. Therefore, most of useful information from NMR spectra of isotropically reoriented molecules are chemical shifts and scalar J-couplings. Other interactions like quadrupolar — for spins larger than 1/2 (*e.g.* ^{2}H), or dipolar couplings are described by traceless tenors and influence spectra only indirectly *e.g.* by well known NOE effect. In the presence of the weakly orienting media the probabilities of various molecular orientations are different. Thus, an averaging gives rise to modified spectrum. There are observed three effects: the chemical shift anisotropy which shifts the peak position, the dipole-dipole coupling and the quadrupolar splittings. In the case when weak alignment is introduced, the molecules adopt a preferred alignment direction in the medium, but, if the reorientation rate is still similar to observed in solution, the linewidths in the spectra remain almost unaffected. Therefore NMR spectra recorded in partially ordered media allow to provide more information about structure of investigated compounds than the spectra recorded in liquid state.

In isotropic phases the most frequently used is the NOE experiment based on dipole-dipole interaction. The 2D technique correlates nuclei, mostly ^{1}H-^{1}H which are not necessarily scalar coupled, but they are close in space (the maximum distance between them is *ca.* 5 Å) and are connected by dipolar interaction. The value of NOE is inversely proportional to the sixth power of the distance between dipolar interacting nuclei ($\sim r^{-6}$). Structural calculation requires a large amount of NOE data, which is time consuming and the number of NOE contacts is more important than precision of their intensities. Moreover, because the NOE effect yield the information only about the local structure of the molecule, there are difficulties with elucidation of three-dimensional structure of a large protein. In the case of non-globular proteins there are often structurally independent domains where there are no NOE contacts among them. When the NOE fails, the problem of structural investigations can be solved using RDC, which are directly related to the molecular structure and allow to determine the relative orientations of internuclear vectors *e.g.* amide NH pairs within a whole macromolecule [20]. The most frequent application of RDC is improvement and verification of protein structure quality.

The measurement of RDC plays an important role in the structure determination of proteins [21—24]. It enables studying of tertiary structure *i.e.* finding of relative orientations of protein domains. However, in order to obtain the structure based only on the RDC measurement, more than one orienting medium is required. The molecular alignment can be described using an alignment tensor, defined by the traceless and symmetric matrix (Saupe order matrix) with five independent elements. After transformation to the principal axis frame of the alignment tensor the dipolar coupling between nuclei i and j can be described as a function of polar coordinates (r_{ij}, θ, φ), which describe the orientation of the internuclear dipolar coupling vector:

$$D_{ij} = -\gamma_i \gamma_j \mu_0 h / 16\pi^3 r_{ij}^3 A_a \left\{ \left(3\cos^2\theta - 1\right) + 3/2 R \sin^2\theta \cos(2\varphi) \right\} \quad (1)$$

where: A_a, R — *axial and rhombic components of the alignment tensor, respectively;* γ_i, γ_j — *gyromagnetic ratios;* μ_0 — *magnetic susceptibility*.

Thus, RDCs are related to internuclear distances (r), angles between the pair of interacting spins and B_0 field (θ), and the degree of orientation. For the precise evaluation of RDCs the induced alignment should not be too strong, what causes the line broadening and not too weak due to limited accuracy of obtained RDCs. In the case of very large degree of orientation the NMR spectrum can be too difficult to analyze because of the presence of long range dipolar couplings.

Some biomolecules *e.g.* nucleic acids or α-helical proteins, which have a sufficiently large magnetic susceptibility anisotropy, orient themselves in strong external magnetic field without using orienting media to observe RDCs. In other case the introduction of orienting media is required. There are many orienting media, for example various liquid crystals [25] and anisotropic gels. The ordering can be achieved by steric or electrostatic forces, it depends on the charge of the media. Lipid bicells were first orienting medium used in biomolecular NMR [26]. It was formed by a mixture of dihexanoyl phosphatidylcholine (DHPC) with dimyristoyl phosphatidylcholine (DMPC) [27]. The disc shaped diamagnetic bicells are oriented with their normal perpendicular direction to the external magnetic field. DHPC/DMPC works in the temperature range of 30—40 °C in aqueous solutions. The disadvantage of using this medium is irreversible degradation at low pH due to hydrolysis. Another bicellular medium working at low pH and over large temperature range is a mixture of 1,2-di-*O*-dodecyl-*sn*-glycero-3-phospholine and 3-(chloroamidopropyl)dimethylammonio-2-hydroxyl-1-propane sulfonate [28]. However this liquid crystal is unstable at neutral pH.

Lamellar phases like cetylpyridium chloride/hexanol/NaCl [29] and cetylpyridium bromide/hexanol/NaBr [30] can be used also as the orienting media in wide temperature range. Other lamellar phases are formed by mixtures of *n*-alkyl-poly(ethylene glycol)/*n*-alkyl alcohol or glucopone/*n*-hexanol. There are many advantages of using these systems — first of all they are commercially available, cheaper than other media and they can be stored long at room temperature, in the case of damaging the restoration can be obtained by the addition of *n*-alkyl alcohol [25]. Glucopon micelles are negatively charged between pH 3 and 9, what causes the improvement in alignment because of the elactrostatic interactions. However there can be a problem with recovery of protein from these lamellar media, sometimes extensive dialysis and cationic exchange chromatography is necessary.

The alternative orienting media are filamentous viruses [31, 32] and purple membrane fragments [33, 34]. Rod shaped viruses are oriented spontaneously with their long axis parallel to the external magnetic field. The most commonly used filamentous phages are *tobacco mosaic virus* and *Pf*1, which grows on *Pseudomonas aeruginosa* or is commercially available. Purple Membrane (PM) is a bacterial membrane isolated from *Helobacterium salinarum*. PM, which has negatively charged, contains bacteriorhodopsin as the sole protein. The introduced orientation is based on the self-alignment caused by intrinsic magnetic susceptibility of the seventrans membrane alpha helices of bacteriorhodopsine, the direction of the membrane normaly is parallel to B_0 field.

Another example of an ordered system is suspension of cellulose crystallites in water [35], especially dedicated to highly charged biological macromolecules that interact with other orienting media.

Nowadays crosslinked polyacrylamide gels are the most frequently used anisotropic phases. Application of these gels [36] introduces the steric alignment, which is independent on the magnetic field. Anisotropy can be induced by compressing or stretching the gel away from an initial isotropic state [36—38]. In the case of compressed phase the protein's long axis is oriented orthogonally to the magnetic field, in stretched gels the proteins are oriented parallel to the field. The protein penetrates the gel by diffusion from an external solution. The main advantage of the application of crosslinked gels is their inertness, therefore it is simple to recover the protein from this orienting phase. After replacement of 50 % of the acrylamide monomers by acrylic acid during the polymerization time there can be obtained a bigger degree of orientation due to electrostatic interaction. The alignment introduced in this way is stronger than in bicellular or inmodified polyacrylamide gels.

The RDCs are determined by comparison of spectra recorded in isotropic and anisotropic phases mostly for one bond interaction, because of the larger magnitude of dipolar coupling. The difference between the coupling constants observed in anisotropic phase and isotropic determines the values of RDCs, which reach tens of Hz. The most frequently measured RDCs, between ^{1}H-^{13}C and ^{1}H-^{15}N, are determined using ^{13}C-HSQC, ^{15}N-HSQC and HNCO experiments. The determination of RDCs between ^{1}H-^{1}H, which do not belong to the protein backbone, is also possible using COSY (correlated spectroscopy)-based techniques.

FAST MULTIDIMENSIONAL NMR

Here the main aims of NMR experiment and the demands for spectra, that are crucial for biomolecular NMR, are introduced. First of all resolution in spectra of large molecules is of great importance. When there are thousands of signals, the narrow linewidths make the spectrum possible to interpret. The lines in spectrum have their natural widths (caused by signal decay), but the signals are additionally broadened due to the measurement free induction decay (FID) function during the finite time. The linewidth is inversely proportional to the maximum time t_{max} for which the FID has been measured:

$$\Delta\nu \sim \frac{1}{t_{max}} \tag{2}$$

On the other hand, if the points in time domain are arranged regularly, then the Nyquist Theorem has to be fulfilled, *i.e.* the space between neighbouring points must be:

$$\Delta t \leq \frac{1}{2\nu_0} \tag{3}$$

where: ν_0 — *highest frequency in the spectrum*.

In the opposite case, the false (folded) signals will appear in the spectrum.

Conventional approach

The most popular scheme of processing of 2D NMR data is shown below (for two-dimensional experiment):

$$s(t_1,t_2) \xrightarrow{FT2} s(t_1,\omega_2) \xrightarrow{FT1} S(\omega_1,\omega_2) \quad (4)$$

In fact it is a sequence of two series of one-dimensional Fourier transforms, calculated using fast Fourier transform (FFT) algorithms. This kind of processing requires time domain data arranged in rows and columns, what has serious consequences as far as experiment time is concerned. Placing points on a Cartesian grid require very long measurement times to obtain desired resolution and fulfill Nyquist Theorem. Collecting of all spectra needed to solve the protein structure requires a few weeks of constant measurements.

Thus development of fast methods of multidimensional NMR is essential for biomolecular research and much effort is put into experiment time shortening. For last two decades the rapid grow in this domain has taken place. There is a wide variety of techniques. Here the most important ones will be introduced, with their advantages and limitations.

Radial sampling of evolution time space

The first attempt to shorten the experimental time by synchronous incrementing of two delays was the idea of accordion spectroscopy. It was based on simultaneous sampling of evolution and mixing periods and was applied for the exchange spectroscopy with exchange rates encoded in signal shapes [39, 40]. Analogous idea of saving experimental time was employed for simultaneous sampling of chemical shift and spin-spin coupling evolution [41, 42]. Application of this approach for sampling of two chemical shift evolution delays is based on the Fourier Transform Theorem [43, 44], claiming that the one-dimensional Fourier transform of the signal acquired for a set of linearly dependent times (t_1, $t_2 = at_1$) gives the projection of spectra in frequency domain on the plain ($\omega_2 = a\omega_1$). During the experiment two evolution times in pulse sequence are changed simultaneously, according to the a value. The achieved spectrum is one dimensional, however it contains twice more peaks. For example, when the signal is $s(t_1, t_2) = \cos(\Omega_1 t_1)\cos(\Omega_2 t_2) = \cos(\Omega_1 t_1)\cos(\Omega_1 a t_1)$ then we obtain peaks at $(\Omega_1 + a\Omega_2)$ and $(\Omega_1 - a\Omega_2)$. To get more information one needs to acquire spectra for more a values. The radial sampling of evolution time space is easily scalable for larger number of dimensions. There are several methods of treating such data, which will be described below.

Reduced dimensionality

Application of accordion spectroscopy principle for parallel sampling of two-chemical shift evolutions called reduced dimensionality (RD) [45—48] allows one to acquire N-dimensional spectrum encoded by radial sampling in two dimensions. Analysis of such projection spectra is based on calculation of the true frequencies by solving the set of equations for each signal. The applicability of this method was gained by an introduction of multiple quadrature detection to such experiments [49, 50]. The important drawbacks of these methods are the possible ambiguity that appears during the analysis and big number of signals; when the spectrum is too complicated solving becomes impossible.

Multi-way decomposition [51, 52]

It is the method of obtaining new projections of spectra by combining those that have been measured. It is based on the assumption that if the signal of reduced experiment (for example three-dimensional reduced to two dimensions) is:

$$f(t,t_3) = \sum_k f_1^k(t,c_{1m}) \cdot f_2^k(t,c_{2m}) \otimes f_3^k(t_3) \quad (5)$$

where: $\otimes$ — *symbolizes the relation between one-dimensional FID functions and three-dimensional signal*.

Than the peaks in two-dimensional projection spectra are convolutions of Lorentzian functions describing peaks:

$$F(\omega,\omega_3) = \sum_k \left[F_1^k(c_{1m}) \cdot F_2^k(c_{2m})\right](\omega) \otimes F_3^k(\omega_3) \quad (6)$$

Spectra from unmeasured projection, whose frequencies are linear combinations of frequencies of measured projections, can be obtained by convolution of appropriate spectra. The advantage of this method is its numerical efficiency, however the result is not a full-dimensional spectrum what requires further analysis, as in RD techniques.

Projection — reconstruction [53—55]

Another way of analyzing projection data is the reconstruction of the full-dimensional spectrum from a set of projections (as shown in Fig. 2), which is easier to interpret and so appears more friendly for users. This kind of methods is widely employed in image reconstruction not only obtained by NMR. Here one can choose between deterministic and statistical methods. In the first case the value of each point of frequency is calculated using the values of corresponding points in all projections. The latter method is searching for the model of spectrum which is compatible with experimental data, but it does not directly use the values of projection spectra.

Most popular deterministic method is additive back projection in which the intensity of each point is simply calculated as a sum corresponding frequency points in all projections. The main disadvantage is that in reconstructed spectrum the false signals (so called artifacts) appear. To cope with this problem the iterative algorithm may be used. It chooses the highest signal and removes all signals in projections, that relevant peak consists of. A

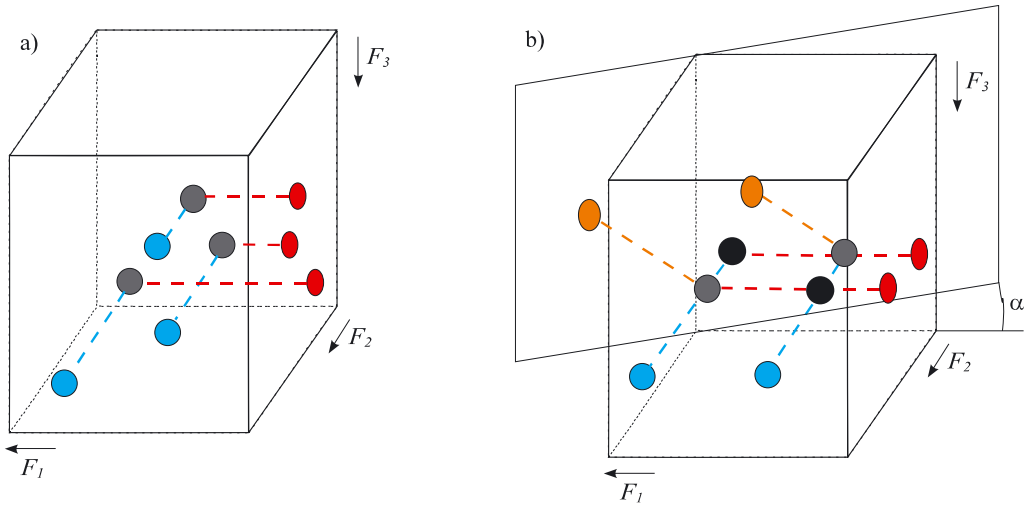

Fig. 2. The principle of 3D projection spectroscopy: a) example of unambiguous reconstruction for two projections, b) degenerated F_3 chemical shifts cause ambiguity which can be solved by third projection

method that deals with the problem of artifacts is lowest-value algorithm, where the intensity is assumed to be the intensity of corresponding point of that projection, in which it has the lowest value. Non-linearity of this algorithm changes the character of noise and makes the sensitivity weaker, which is the price one has to pay for reducing the number of artifacts. None of these methods deals with spectra with wide range of signal intensities, such as NOESY.

In statistical methods the starting point is a spectrum with arbitrary assumed intensities. Then the intensities are varied and the aim is to find the model that best matches the projection data. There are some different ways to choose the best possible model. One is iterative least-squares fitting, where the sum of square deviations is calculated and minimized. Problems with this technique are long computations and finding the starting point.

Arbitrary sparse sampling

A large time benefit can be achieved if conventional sampling of time domain is replaced by some kind of irregular sampling. Then the Nyquist theorem does not have to be fulfilled and higher maximum times of evolution (and so the better resolution) can be obtained without the effect of folding of signals. The question is how to get the spectrum; the conventional sequence Fourier transform is not applicable to such data.

Maximum entropy method [56, 57]

One possible scheme of processing is statistical technique called maximum entropy method (MEM). It is widely used in many areas of analyzing data. It may be also used as a projection reconstruction method, based on radial sampling of evolution time space. The essence of MEM is to consider many models of spectra and choose one with the highest value of entropy, defined as:

$$S = -\sum_{i=1}^{M} \frac{|f_i|}{A} \ln \frac{|f_i|}{A} \tag{7}$$

where: $A = \sum_{k=1}^{M} f_k$ — *normalizing factor*, f_i — *value of the model spectrum in its i point (the spectrum consists of M points)*.

The higher entropy has the model, the closer to the true spectrum it is. Main problem of MEM is very long computation time when modeling spectra of full dimensionality.

Polynomial interpolation [58]

The idea of this method is to obtain data regularly placed in time domain from points arranged in arbitrary way by polynomial interpolation. Then the data may be transformed conventionally, using FFT algorithms. The starting data may be irregular only in one dimension, so the method is used for two dimensional experiments (with one irregularly sampled evolution time) or for reduced dimensionality experiments. In the spectra obtained with this technique the additional noise occurs, because of non-matching of sampled function (sum of decaying periodic functions) and polynomials.

Multidimensional Fourier transformation (MFT) [59]

This is general method which enables direct calculation of the spectrum from arbitrary chosen time domain points. Contrary to conventional sequence of 1D Fourier transforms it can be applied to any (also randomly chosen) set of time points. Here each point of spectrum (in frequency domain) is calculated in one step, using all points from time domain. For example for an assuming of the two-dimensional signal as:

$$s(t_1,t_2) = \cos(\Omega_1 t_1) \cdot \cos(\Omega_2 t_2) \tag{8}$$

the transform is calculated as:

$$S(\omega_1,\omega_2) = \sum_{t_1=0}^{t_{1\,\max}} \sum_{t_2=0}^{t_{2\,\max}} \cos(\Omega_1 t_1) \cdot \cos(\Omega_2 t_2) \cdot \cos(\omega_1 t_1) \cdot \cos(\omega_2 t_2) \cdot w(t_1,t_2) \tag{9}$$

where: $w(t_1, t_2)$ — *optional weights associated with time domain data points*.

The quadrature is realized, for collecting four sets of data with different modulations and adding four transforms:

$$s_{cc}(t_1,t_2) = \cos(\Omega_1 t_1)\cdot\cos(\Omega_2 t_2) \rightarrow S_{cc} = \sum_{t_1=0}^{t_{1\max}}\sum_{t_2=0}^{t_{2\max}} s_{cc}\cdot\cos(\omega_1 t_1)\cdot\cos(\omega_2 t_2)\cdot w(t_1,t_2) \tag{10a}$$

$$s_{cs}(t_1,t_2) = \cos(\Omega_1 t_1)\cdot\sin(\Omega_2 t_2) \rightarrow S_{cs} = \sum_{t_1=0}^{t_{1\max}}\sum_{t_2=0}^{t_{2\max}} s_{cs}\cdot\cos(\omega_1 t_1)\cdot\sin(\omega_2 t_2)\cdot w(t_1,t_2) \tag{10b}$$

$$s_{sc}(t_1,t_2) = \sin(\Omega_1 t_1)\cdot\cos(\Omega_2 t_2) \rightarrow S_{sc} = \sum_{t_1=0}^{t_{1\max}}\sum_{t_2=0}^{t_{2\max}} s_{sc}\cdot\sin(\omega_1 t_1)\cdot\cos(\omega_2 t_2)\cdot w(t_1,t_2) \tag{10c}$$

$$s_{ss}(t_1,t_2) = \sin(\Omega_1 t_1)\cdot\sin(\Omega_2 t_2) \rightarrow S_{ss} = \sum_{t_1=0}^{t_{1\max}}\sum_{t_2=0}^{t_{2\max}} s_{ss}\cdot\sin(\omega_1 t_1)\cdot\sin(\omega_2 t_2)\cdot w(t_1,t_2) \tag{10d}$$

$$S(\omega_1,\omega_2) = S_{cc}(\omega_1,\omega_2) + S_{cs}(\omega_1,\omega_2) + S_{sc}(\omega_1,\omega_2) + S_{ss}(\omega_1,\omega_2) \tag{11}$$

This method allows reaching the same resolution as in conventional case in much shorter time what radically improves the resolution if the experiment last the same time (comparison of conventional and MFT spectra is shown in Fig. 3). It also allows analyzing spectra with wide range of signal intensities. The great advantage is lack of assumption at the beginning, which makes the method general. The price we have to pay is decrease in signal to noise ratio, as the sampling noise (with the amplitude of the same order as the thermal noise) appears, but in the case of random sampling there are no coherent artifacts.

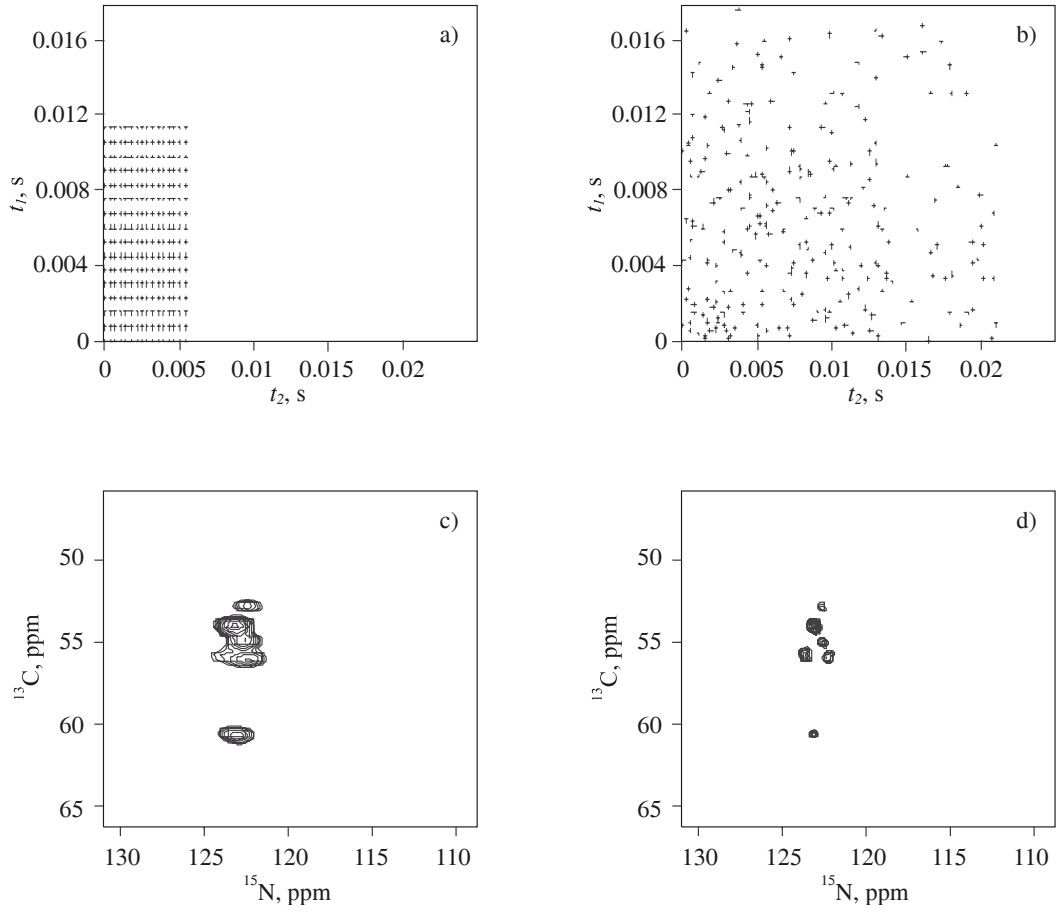

Fig. 3. Part of 3D HNCA spectrum of ubiquitin distribution of t_1/t_2 domain data points conventional (a) random (b) and cross sections showing resolution enhacement in spectrum measured using random distribution of time domain data points (d) in comparison to conventional spectrum (c) (according to [59])

T a b l e 1. Comparison of key features of various new techniques of fast multidimensional NMR

	Conventional	RD	Multi-way decomp.	PR-additive back projection	PR-lowest value	PR-least squares	Maximum entropy	Interpolation	MFT
short experiment	—	+	+	+	+	+	+	+	+
no additional noise	+	—	—	—	+	+	+	—	—
no start assumptions	+	+	+	+	+	—	—	—	+
fast calculations	+	—	—	+	+	—	—	+	+
high dynamic range	+	—	—	—	—	—	—	—	+
comments	very long exp. time for higher dimensional spectra	fails when too many signals	requires similar calculations to RD	false signals	signals disappear changed line shapes	very long calculation times	numerically expensive required start model	applicable to 2D spectra (or to RD)	avoidable sampling noise

CONCLUSIONS

The most important features of above mentioned methods are compared in Table 1.

The fast progress in NMR methods and hardware development enable determination and improvement in growing number of protein structures. Although the ability of obtaining of high quality protein spectra does not mean that the structure could be solved, the continuous advances in this field are observed. We expect that in addition to the achievements discussed above, the new methods of preparation of isotopically enriched samples, new approaches of automatic data analysis and structure conclusions would bring the further progress in the near future.

REFERENCES

[1] Berman H. M., Westbrook J., Feng Z., Gilliland G., Bhat T. N., Weissig H., Shindyalov I. N., Bourne P. E.: *Nucleic Acids Research* 2000, **28**, 235. [2] Sattler M., Schleucher J., Griesinger C.: *Prog. Nucl. Magn. Reson. Spectrosc.* 1999, **34**, 93. [3] Fiaux J., Bertelsen E. B., Horwich A. L., Wüthrich K.: *Nature* 2002, **418**, 207. [4] Fernandez C., Wider G.: *Curr. Opin. Struct. Biol.* 2003, **13**, 570. [5] Kay L.: *J. Magn. Reson.* 2005, **173**, 193. [6] Oschkinat H., Griesinger C., Kraulis P. J., Sorensen O. W., Ernst R. R., Gronenborn A. M., Clore G. M.: *Nature* 1988, **332**, 374. [7] Vuister G. W., Boelens R., Kaptein R.: *J. Magn. Reson.* 1988, **80**, 176. [8] Marion D., Kay L. E., Sparks S. W., Torchia D. A., Bax A.: *J. Am. Chem. Soc.* 1989, **111**, 1515. [9] Ikura M., Kay L. E., Bax A.: *Biochemistry* 1990, **29**, 4659. [10] Bax A., Grzesiek S.: *Acc. Chem. Res.* 1993, **26**, 131.

[11] Yamazaki T., Lee W., Arrowsmith C. H., Muhandiram D. R., Kay L. E.: *J. Am. Chem. Soc.* 1994, **116**, 11 655. [12] Pervushin K., Riek R., Wider G., Wüthrich K.: *Proc. Natl. Acad. Sci. USA* 1997, **94**, 12 366. [13] Riek R., Wider G., Pervushin K., Wüthrich K.: *Proc. Natl. Acad. Sci. USA* 1999, **96**, 4918. [14] Morris G. A., Freeman R.: *J. Am. Chem. Soc.* 1979, **101**, 760. [15] Güntert P.: *Prog. Nucl. Magn. Reson. Spectrosc.* 2003, **43**, 105. [16] Kigawa T., Nuto Y., Yokoyama S.: *J. Biomol. NMR* 1995, **6**, 129. [17] Xu R., Ayers B., Cowburn D., Muir T. W.: *Proc. Natl. Acad. Sci. USA* 1999, **96**, 388. [18] Lian L.-Y., Middleton D. A.: *Prog. Nucl. Magn. Reson. Spectrosc.* 2001, **39**, 90. [19] Otomo T., Teruya K., Uegaki K., Yamazaki T., Kyogoku Y.: *J. Biomol. NMR* 1999, **14**, 105. [20] Simon B., Sattler M.: *Angew. Chem. Int. Ed.* 2002, **41/3**, 437.

[21] Annila D., Permi P.: *Concepts Magn. Reson.* 2004, **23A**, 22. [22] de Alba E., Tjandra N.: *Prog. Nucl. Magn. Reson. Spectrosc.* 2002, **40**, 175. [23] Blackledge M.: *Prog. Nucl. Magn. Reson. Spectrosc.* 2005, **46**, 23. [24] Bax A.: *Protein Sci.* 2003, **12**, 1. [25] Rückert M., Otting G.: *J. Am. Chem. Soc.* 2000, **122**, 7793. [26] Ottiger M., Bax A.: *J. Biomol. NMR* 1998, **12**, 361. [27] Tjandra N., Bax A.: *Science* 1997, **278**, 1111. [28] Cavagnero S., Dyson H. J., Wright P. E.: *J. Biomol. NMR* 1999, **13**, 387. [29] Prosser R. S., Losonczi J. A., Shiyanovskaya I. V.: *J. Am. Chem. Soc.* 1998, **120**, 11 010. [30] Barrientos L. G., Dolan C., Gronenborn A. M.: *J. Biomol. NMR* 2000, **16**, 329.

[31] Clore G. M., Starich M. R., Gronenborn A. M.: *J. Am. Chem. Soc.* 1998, **120**, 10 571. [32] Hansen M. R., Mueller L., Pardi A.: *Nature Struct. Biol.* 1998, **5**, 1065. [33] Sass H. J., Cordier F., Hoffmann A., Rogowski M., Cousin A., Omichinski J. G., Lowen H., Grzesiek S.: *J. Am. Chem. Soc.* 1999, **121**, 2047. [34] Koenig B. W., Hu J. S., Ottiger M., Bose S., Hendler R. W., Bax A.: *J. Am. Chem. Soc.* 1999, **121**, 1385. [35] Fleming K., Gray D., Presannan S., Matthews S.: *J. Am. Chem. Soc.* 2000, **122**, 5224. [36] Tycko R., Blanco F. J., Ishii Y.: *J. Am. Chem. Soc.* 2000, **122**, 9340. [37] Sass H. J., Musco G., Stahl S. J., Wingfield P. T., Grzesiek S.: *J. Biomol. NMR* 2000, **18**, 303. [38] Chou J. J., Gaemers S., Howder B., Louis J. M., Bax A.: *J. Biomol. NMR* 2001, **21**, 377. [39] Bodenhausen G., Ernst R. R.: *J. Magn. Reson.* 1981, **45**, 367. [40] Bodenhausen G., Ernst R. R.: *J. Am. Chem. Soc.* 1982, **104**, 1304.

[41] Koźmiński W., Sperandio D., Nanz D.: *Magn. Res. Chem.* 1996, **34**, 311. [42] Koźmiński W., Nanz D.: *J. Magn. Res.* 1997, **124**, 383. [43] Bracewell R. N.: *Aust. J. Phys.* 1956, **9**, 198. [44] Nagayama K., Bachmann P., Wüthrich K., Ernst R. R.: *J. Magn. Reson.* 1978, **31**, 133. [45] Szyper-

ski T., Wider G., Buschweller J. H., Wüthrich K.: *J. Am. Chem. Soc.* 1993, **115**, 9307. [46] Szyperski T., Wider G., Buschweller J. H., Wüthrich K.: *J. Biomol. NMR* 1993, **3**, 127. [47] Brutscher B., Simorre J. P., Caffrey M. S., Marion D.: *J. Magn. Reson.* 1994, **B 105**, 77. [48] Löhr F., Rüterjans H.: *J. Biomol. NMR* 1995, **6**, 189. [49] Kim S., Szyperski T.: *J. Am. Chem. Soc.* 2003, **125**, 13 958. [50] Koźmiński W., Zhukov I.: *J. Biomol. NMR* 2003, **26**, 157.

[51] Malmodin D., Billeter M.: *J. Magn. Reson.* 2005, **176**, 47. [52] Malmodin D., Billeter M.: *J. Am. Chem. Soc.* 2005, **127**, 13 486. [53] Kupče Ē., Freeman R.: *J. Am. Chem. Soc.* 2003, **125**, 13 958. [54] Coggins B. E., Venters R. A., Zhou P.: *J. Am. Chem. Soc.* 2005, **127**, 11 562. [55] Yoon J., Godsill S., Kupče Ē., Freeman R.: *Magn. Res. Chem.* 2006, **44**, 197. [56] Leue E. D., Maygner M. R., Skilling J., Staunton J.: *J. Magn. Reson.* 1986, **68**, 14. [57] Rovnyak D., Frueh D. P., Sastry M., Sun Z. Y. J., Stern A. S., Hoch J. C., Wagner G.: *J. Magn. Reson.* 2004, **170**, 15. [58] Marion D.: *J. Biomol. NMR* 2005, **32**, 141. [59] Kazimierczuk K., Zawadzka A., Koźmiński W., Zhukov I.: *J. Biomol. NMR* 2006, **36**, 157.