

Refined NMR solution structures of proteins using homo- and heteronuclear couplings, relaxation time measurements and relaxation matrix analysis

Heinz Rüterjans, Frank Löhr, Markus Blümel, Stefania Pfeiffer, Jan Engelke,
Norman Spitzner and Jürgen M. Schmidt

Institute of Biophysical Chemistry, J.W.Goethe-University Frankfurt/Main, Germany

Abstract

In order to compare high resolution crystal structures of proteins with the corresponding solution structures a detailed analysis of NMR parameters obtained for various proteins was carried out. As many NOE values as possible were transformed into distances using a relaxation matrix analysis. In addition homo- and heteronuclear 3J couplings from ^{13}C and ^{15}N enriched protein species were determined. From these couplings the dihedral angles ϕ , ψ and χ_1 were evaluated. It was possible to interpret the various 3J values in terms of either distinct dihedral angles or with a certain variance of angles or with an equilibrium of different rotameric states.

The refined solution structures were obtained using the distance constraints together with the dihedral angle constraints in a distance geometry algorithm (DIANA program package). The resulting DG structure was the starting conformation of a subsequent molecular dynamics simulation. From a determination of relaxation times T_1 , T_2 and NOE build-up rates of ^{15}N and ^{13}C nuclei order parameters were obtained to describe the dynamic behaviour of protein molecular parts.

Refined solution structures were obtained for ribonuclease T_1 , a flavodoxin from *D.vulgaris*, fatty acid binding protein from bovine heart and for a heat shock transcription factor from tomato. In most cases the high resolution crystal structures differ only slightly from the refined solution structures.

INTRODUCTION

With recent advances in multidimensional NMR spectroscopy together with isotopic enrichment techniques an almost complete assignment of proton- and heteronuclei (^{15}N and ^{13}C) in proteins has become feasible for molecules with a molecular weight of up to 30 kDa. NOE distance constraints together with homonuclear and heteronuclear couplings have permitted the first detailed studies of solution structures of such proteins. However, since protein structures are dynamic structures more information about the dynamic properties is required. Beside exchange processes NMR is able to provide motional parameters from relaxation time measurements. Using theoretical approximations these relaxation times are transferred to order parameters describing the internal motion of proteins. Using special NMR techniques it was possible to detect interactions of solvent molecules with backbone CO or NH groups via hydrogen bonding. Recently also residence times of some water molecules at distinct sites of the protein have been determined. In the following some of these techniques were used for the refinement of solution structures of various protein systems.

RESULTS AND DISCUSSION

Number of NOE Values, Relaxation Matrix Analysis

The most important NMR data for determining solution structures of proteins are scalar couplings and NOEs or NOE build-up rates which are translated into restraints on dihedral angles and interproton distances. An ensemble of structures is calculated with a distance geometry algorithm (DIANA program package) in order to fulfil as much of these constraints as possible. The quality of the resulting solution structure depends mostly on the number and to a less extent on precision and accuracy of the distance and

dihedral angle restraints. Only a limited amount of structure relevant constraints is necessary for the determination of secondary structure elements (1) or a fairly reasonable tertiary structure (2). Fig. 1A presents 20 DIANA structures of flavodoxin from *Desulfovibrio vulgaris* derived from 1415 distance and 112 ϕ -angle constraints. Using different NOESY and ROESY techniques the number and precision of distance restraints could be increased. Stereospecific assignments of prochiral protons and methyl groups improve the quality of the data, as pseudoatom corrections can be neglected. Fig. 1B presents the resulting refined solution structure of flavodoxin based on approximately 15 restraints per residue. The precision of the resulting DIANA structure is very high. Additional qualitative distance restraints will only lead to an increase of redundant information (3). In order to assess the accuracy of the structure, quantitative distance restraints are presently derived from NOE build-up rates and relaxation matrix analysis (4). In the future attempts will be necessary to consider the dynamic nature of proteins in the structure calculation.

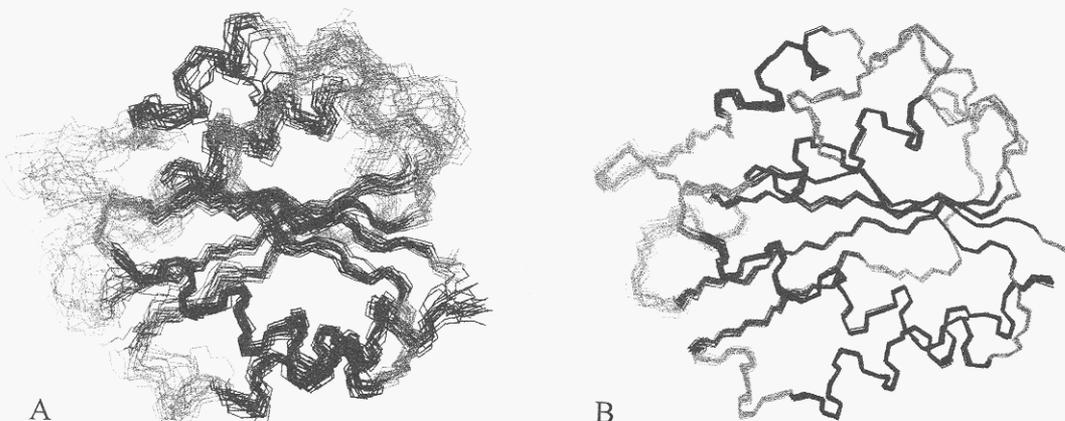


Fig. 1: Comparison of 20 DIANA structures of flavodoxin from *D. vulgaris* calculated on the basis of (A) 1415 distance and 112 ϕ -angle constraints or (B) on 2112 distance and 201 ϕ - and χ_1 -constraints together with 141 stereospecific assignments. The mean global backbone RMSD value drops from 1.79 (A) to 0.35 (B).

Vicinal Spin-Spin Coupling Constants

Homo- and heteronuclear 3J couplings have become available with a variety of heteronuclear NMR techniques based on the E.COSY principle (5,6,7). Sensitivity problems in the determination of heteronuclear couplings were overcome using isotopically labeled samples. Karplus parameters (8, 9) which relate 3J coupling constants to dihedral angles are known for most of the spin pairs relevant to the conformational analysis of proteins. Since Karplus relations are not single valued functions delivering up to four different dihedral angle values for a given 3J value, a set of homo- and heteronuclear coupling constants must be determined to unambiguously characterize the dihedral angle. However, the interpretation of vicinal coupling constants in terms of a unique dihedral angle is hampered by conformational mobility since coupling constants may be time averages over multiple conformations. Also for protein amino acid sidechains the staggered rotamer model might be inappropriate for the following reasons:

1. Non-staggered rotamers may occur due to a shifted torsional potential minimum which arises from structural interactions in the protein matrix. It was derived from x-ray protein structures that crystallographic χ_1 angles in the sidechains often deviate from the ideal staggered conformations.
2. A limitation to discrete rotamers may not be correct considering the local mobility in the backbone as well as in the side chains of a protein. In addition, distributions of dihedral angles may occur, especially for χ_1 angles of side chains on the surface of a protein.

We have determined four possible vicinal coupling constants related to the dihedral angle ϕ (10) $^3J_{\text{HNH}\alpha}$, $^3J_{\text{HNC}\beta}$, $^3J_{\text{HNC}\beta}$ and $^3J_{\text{H}\alpha\text{C}\beta-1}$ and most of the vicinal coupling constants related to χ_1 (11). The 3J data are analysed with respect to different models for the conformational dynamics of the backbone and of the side

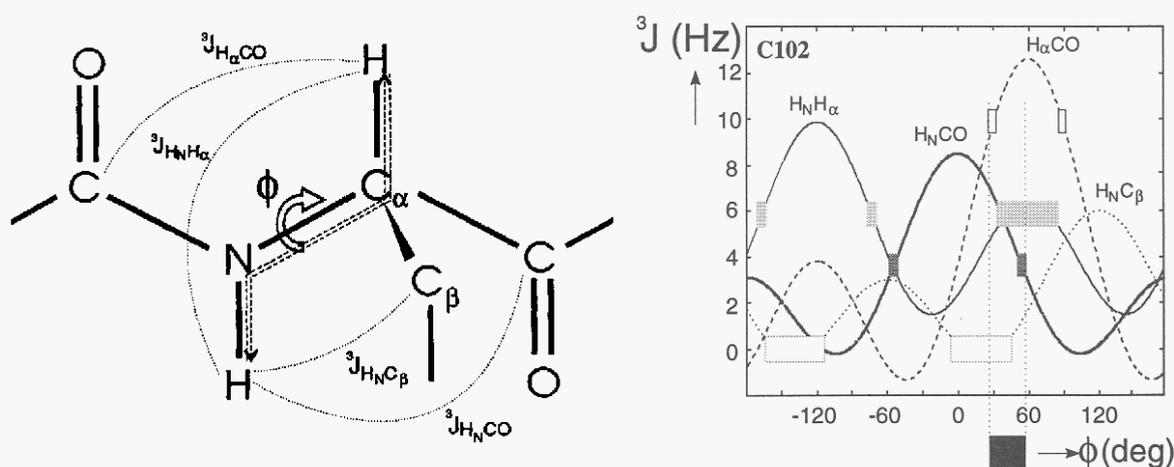


Fig. 2: The diagram on the right side displays the Karplus relations for the ϕ -angle determination. The ϕ -angle should correspond to all measured 3J values. The experimental 3J couplings are indicated with open and hatched rectangles with an uncertainty of ± 0.5 Hz. The black rectangle below the abscissa denotes the most probable ϕ -angle from the coincidence of the experimental 3J couplings for the *D. vulgaris* flavodoxin Cys 102 residue.

chains. The conformations obtained from J coupling data are examined with respect to consistency with NOE data. The results are also compared with crystal structures. Dihedral angles were obtained by fitting one or multiple ϕ or χ_1 rotamers to the set of experimental coupling constants. The dihedral angle dependence of the 3J coupling constants used for the ϕ -angle determinations is indicated in Fig. 2. A scheme for the possible 3J couplings for determining ϕ is also indicated in Fig. 2. In general the backbones or the side chains are not necessarily restricted to a single conformation. Internal motion may lead to an averaging of the observed 3J coupling constants, therefore three different models of internal dynamics were applied to describe the most probable dihedral angle θ ($= \phi$ or χ_1) or the corresponding angle distributions. In model A a fixed angle was assumed in a single parameter fit to minimize the difference between the calculated and experimental coupling constants according to:

$$J_k^{\text{expt}} \equiv J_k^{\text{calc}}(\theta) \quad (1)$$

In model B the dihedral angle was allowed to exhibit local mobility according to a unimodal Gaussian distribution leading to a two-parameter fit of θ (in average) and σ_θ as given by:

$$J_k^{\text{expt}} \equiv J_k^{\text{calc}}(\bar{\theta}, \sigma_\theta) = \frac{1}{\sigma_\theta \sqrt{2\pi}} \int_{-\pi}^{+\pi} J_k^{\text{calc}}(\theta) \cdot \exp\left\{-\frac{(\theta - \bar{\theta})^2}{2\sigma_\theta^2}\right\} d\theta \quad (2)$$

In the three-site jump model C comprising the three staggered rotamers, e.g. with $\chi_1 = 180^\circ$, -60° and $+60^\circ$ the respective populations p_I , p_{II} and p_{III} were allowed to vary leading again to a two-parameter fit of p_I and p_{II} according to:

$$J_k^{\text{expt}} \equiv J_k^{\text{calc}}(p_I, p_{II}) = p_I J_k^{\text{calc}}(-60^\circ) + p_{II} J_k^{\text{calc}}(\pm 180^\circ) + (1 - p_I - p_{II}) \cdot J_k^{\text{calc}}(+60^\circ) \quad (3)$$

A comparison of conformations in solution with x-ray data revealed that in most cases the dihedral angles are in agreement. Some χ_1 angles especially from side chains on the surface of the protein differed in that sense that equilibria of two or three different rotameric structures occur in solution while one rigid conformation was found in the crystal. In some cases the fit between the experimentally determined coupling constants and a dihedral angle was only possible allowing for a local mobility according to a unimodal Gaussian distribution.

pH-dependent NMR Studies

From pH-dependent studies of side chain resonances in particular of ionizable side chains pK values were derived. In Fig. 3 the pH-dependence of ^{13}C carboxyl resonances of some glutamic and aspartic acid residues is shown. It may be readily recognized that one of the carboxylic groups is not ionized because this residue is buried in the protein core. Because of a strong interaction between the Glu58 and His40 in ribonuclease T_1 , the pK-value of this residue is very low, whereas the pK-value of glutamic acid 28 is very high due to the fact that the carboxyl group of this side chain is located on top of the C-terminal end of the α -helix in ribonuclease T_1 . The strong polarisation of the α -helix prevents the carboxyl group to ionize at low pH-values. In fact a deviation of pK-values in proteins from those in peptides provides information about the extent of interactions between side chains (12).

^{15}N and ^{13}C Relaxation Times

NMR relaxation experiments can provide a detailed description of protein dynamics. In general dynamic parameters are obtained from a determination of T_1 , T_2 and NOE values. ^1H detected sensitive one- and two-dimensional heteronuclear NMR experiments have been reported for studying the relaxation properties of ^{15}N and ^{13}C nuclei. While the ^{15}N relaxation experiments only provide information about the backbone dynamics ^{13}C relaxation times can be used to study the motion of both, the backbone and the side chains. The analysis of the relaxation data using a suitable model of motion provide dynamical parameters of the motion of the backbone. Since a typical property of protein structures is the planarity of the peptide plane, between the carbonyl and the amide group the motion of the NH bond should be highly correlated with the motion of the CO bond. Therefore it should be possible to study the dynamics of the peptide plane even with the ^{15}N amide or with the ^{13}C carbonyl relaxation. In contrast to ^{15}N relaxation time measurements the ^{13}C relaxation has three major advantages. No artifacts will occur due to the H^{N} exchange with water protons. In addition, proline residues are detectable and finally the relation between the measured relaxation times and the spectral density function is much easier to interpret because the ^{13}CO relaxation follows almost entirely the chemical shift anisotropy mechanism rather than the dipolar mechanism. Relaxation times and NOE build-up rates were analyzed using the model free approach according to Lipari and Szabo (13). According to this model the motion of a protein is divided into two parts: the overall rotational diffusion and the internal motion. A so-called order parameter S^2 describes essentially the internal motion. If this quantity is equal to one the protein is a rigid body without internal motion. If this quantity is equal to zero its isotropic motion without any restrictions takes place. In Fig. 4 order parameters for each of the residues of ribonuclease T_1 is shown using ^{13}CO (A), ^{15}N (B) and $^{13}\text{C}\alpha$ (C) relaxation data. The correlation of motions within a peptide plane is indicated for the pairs Asp49 and Phe50 and for Ser51 and Val52. The analysis of this data indicates that within the secondary structure elements motions are restricted whereas in loop regions the backbone is rather mobile. Further motional restrictions are induced with ligand binding (14).

Interaction of Water Molecules with the Protein Surface

Juranic *et al.* (15) found that the coupling between the ^{15}N and the ^{13}C nucleus of a peptide bond is depending on the extent of hydrogen bonding of either the amide proton or the carbonyl oxygen. Usually the $^1J_{\text{NC}}$ coupling constant has a value of about 15 Hz. The value of the coupling constant is modulated by the N-C' bond length which varies in case of hydrogen bonding within the protein or with water molecules. Hydrogen bonds within the protein are weaker than hydrogen bonds with the solvent. In case the amide group is involved in a strong hydrogen bond the value of the $^1J_{\text{NC}}$ coupling constant drops below 15 Hz. In case the carbonyl group is involved in a strong hydrogen bond the value of the $^1J_{\text{NC}}$ coupling constant is increased to more than 16 Hz. If there is no structural indication for a hydrogen bond within the protein, the extreme value of the coupling constant arises from a hydrogen bond with a water molecule. In Fig. 5 the amide protons and the carbonyl oxygens of peptide groups with strong hydrogen bonds to water molecules are indicated with dark shaded balls in the secondary structure of the protein.

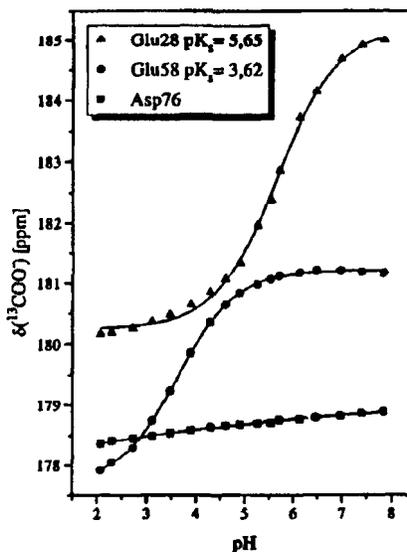


Fig. 3: pH-dependence of ^{13}C carboxyl resonances of three side chains of ribonuclease T_1

Although the exchange of water molecules from the surface of the protein to the bulk water is rapid on the NMR time scale, it is possible to detect magnetisation transfer due to NOE or ROE from water protons to protein protons (16). The magnetisation transfer depends on the distance between the involved protons and the correlation time of the interproton vector. This correlation time corresponds to the average residence time τ_R of the water molecules in distinct protein sites. While the ROE is positive over the entire range of correlation times the NOE changes to a negative sign at $\omega_0 \tau_R = 1.12$, corresponding to $\tau_R = 300$ ps at a Larmor frequency of 600 MHz (17). Hence from the comparison of the ROE and NOE values average

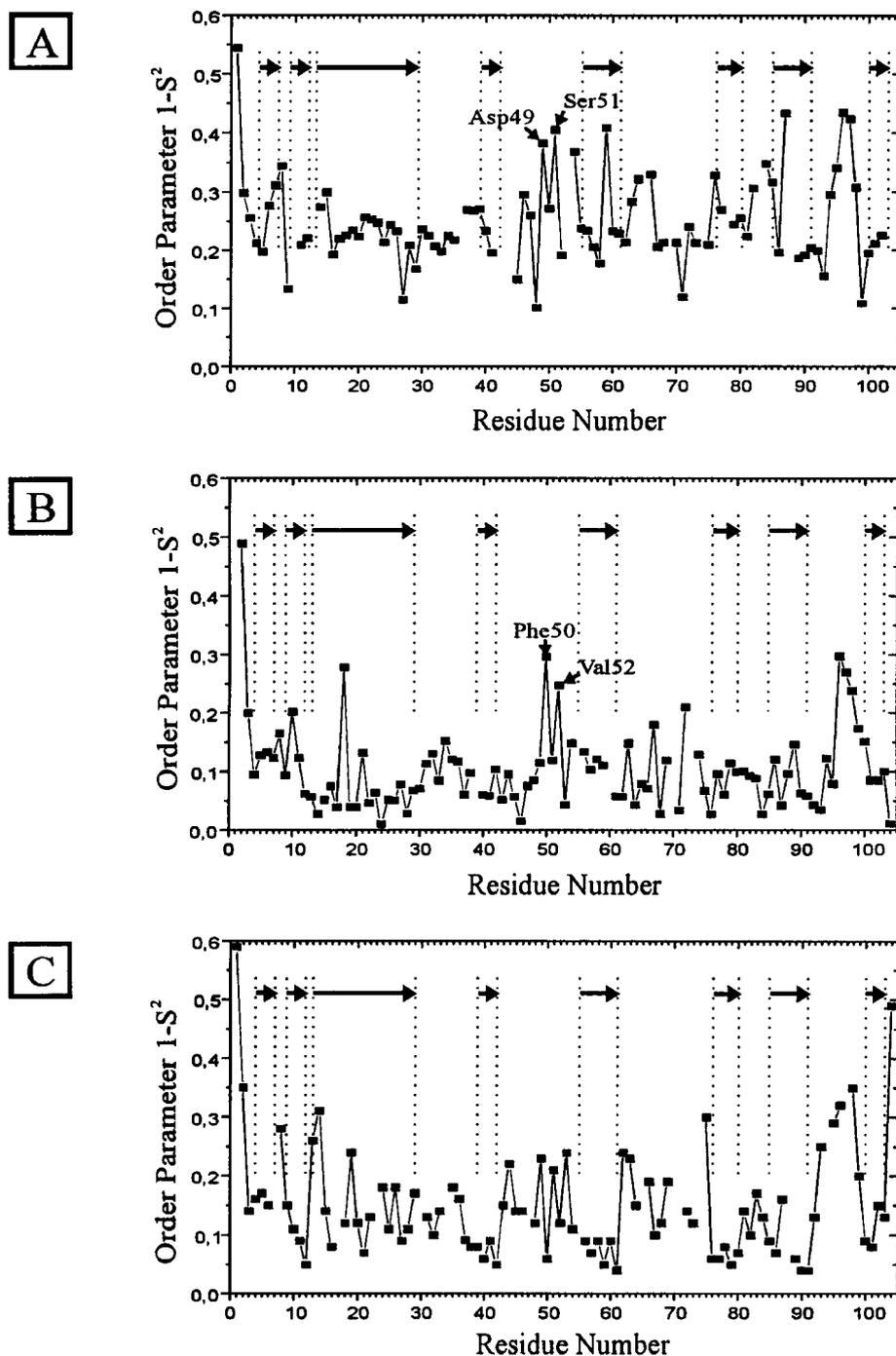


Fig. 4: Order parameters for backbone atoms as derived from ^{15}N and ^{13}C relaxation data for amino acid residues of ribonuclease T1.

A: order parameters from ^{13}C carbonyl relaxation data

B: order parameters from ^{15}N relaxation data

C: order parameters from $^{13}\text{C}_\alpha$ relaxation data

Arrows on the top indicate secondary structure elements in the sequence

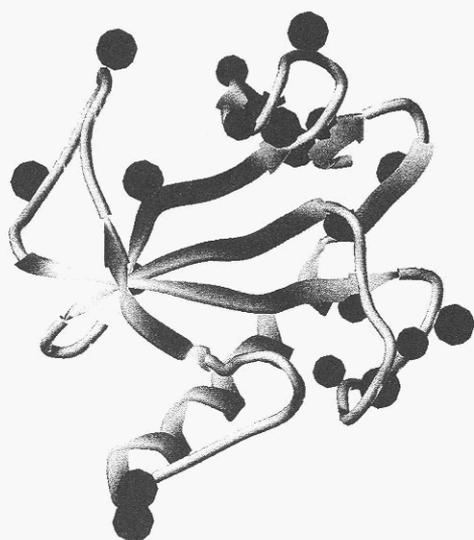


Fig. 5: Structure of ribonuclease T₁. Small spheres indicate NH and larger spheres indicate CO groups involved in H-bonds to water molecules.

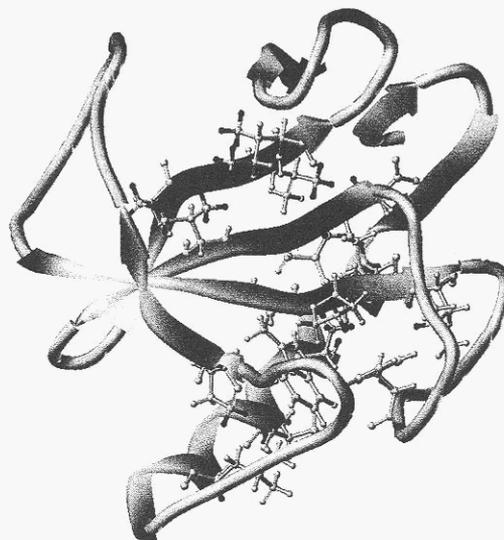


Fig. 6: Structure of ribonuclease T₁. Protons which are in close proximity to water molecules are indicated with black top balls within the side chains.

residence times of water molecules at distinct sites of the protein surface are possible. In Fig. 6 the protein protons which are in close proximity to water molecules are indicated with black balls in the ball-stick representation of the involved amino acid residues.

CONCLUSIONS

NMR spectroscopy can provide a detailed picture of the dynamic structure of proteins. In general crystal structures and solution structures are very similar. Albeit there may be differences in the conformations of side chains in particular on the surface of the protein structure. In both types of structure the secondary structure elements are more rigid while the loop regions are more mobile.

ACKNOWLEDGEMENT

We would like to thank the Deutsche Forschungsgemeinschaft for a grant (Ru 145/8-7, Ru 145/11-2).

REFERENCES

1. M.K. Knauf, F. Löhr, G.P. Curley, P. O'Farrell, S.G. Mayhew and H. Rüterjans, *Eur. J. Biochem.* **213**, 167-184, (1993).
2. M.K. Knauf, F. Löhr, M. Blümel, S.G. Mayhew, and H. Rüterjans, submitted to *Eur. J. Biochem.*
3. G.M. Clore, M.A. Robien and A.M. Gronenborn, *J. Mol. Biol.* **231**, 82-102, (1993).
4. M.J.J. Bonvin, J.A.C. Rullmann, R.M.J.N. Lamerichs, R. Boelens and R. Kaptein, *Proteins* **15** 385-400, (1993).
5. C. Griesinger, O.W. Sørensen, R.R. Ernst, *J. Am. Chem. Soc.* **107**, 6394 (1985).
6. C. Griesinger, O.W. Sørensen, R.R. Ernst, *J. Am. Phys.* **85**, 6837 (1986).
7. C. Griesinger, O.W. Sørensen, R.R. Ernst, *J. Magn. Reson.* **75**, 474 (1987).
8. M. Karplus, *J. Am. Chem. Soc.* **85**, 2870 (1963).
9. U.F. Bystrov, *Prog. NMR Spectrosc.* **10**, 41 (1976).
10. F. Löhr and H. Rüterjans, *J. Biomol. NMR* **5**, 25 (1995).
11. Y. Karimi-Nejad, J.M. Schmidt, H. Rüterjans, H. Schwalbe, C. Griesinger, *Biochemistry* **33**, 5481 (1994).
12. M. Laskowski Jr. and H. Scheraga, *J. Am. Chem. Soc.* **76**, 6305, (1954).
13. Lipari, G., Szabo, A., *J. Am. Chem. Soc.* **104**, 4546-4559, (1982).
14. J. Engelke and H. Rüterjans, *J. Biomol. NMR* **5**, 173-182 (1995).
15. N. Juranic, P.K. Ilich, S. Macura, *J. Am. Chem. Soc.* **117**, 405 - 410, (1995).
16. G. Otting *et al.*, *Science* **254**, 974, (1991).
17. S. Grezsiak, A. Bax, *J. Am. Chem. Soc.* **115**, 12593, (1993).