

Long-range ^{15}N - ^1H correlation as an aid to sequential proton resonance assignment of proteins

Application to the DNA-binding protein *ner* from phage Mu

G. Marius Clore, Ad Bax, Paul Wingfield* and Angela M. Gronenborn

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA and *Glaxo Institute for Molecular Biology SA, 46 Route des Acacias, CH-1211, Geneva, Switzerland

Received 13 July 1988; revised version received 25 July 1988

A method is described for sequential resonance assignment of protein ^1H -NMR spectra relying on the detection of long-range correlations between ^{15}N and C^2H atoms using ^1H -detected heteronuclear multiple-bond correlation spectroscopy. In particular, the observation of the two-bond $^{15}\text{N}(i)\text{-C}^2\text{H}(i)$ and three-bond $^{15}\text{N}(i)\text{-C}^2\text{H}(i-1)$ correlations enables one to connect one residue with the next. Because the magnitude of the long-range couplings is small (< 6 Hz), the sensitivity of this experiment is necessarily low and requires the use of ^{15}N -enriched protein samples. Further, because the size of the $^{15}\text{N}(i)\text{-C}^2\text{H}(i-1)$ coupling is very sensitive to the ψ backbone torsion angle, structural information can be derived.

The application of this experiment is illustrated with the 75-residue DNA-binding protein *ner* from phage Mu.

NMR; Sequential resonance assignment; ^{15}N labeling; Heteronuclear multiple bond correlation; *ner* protein; Phage Mu

1. INTRODUCTION

The determination of the three-dimensional structure of a protein in solution using NMR spectroscopy depends crucially on obtaining complete or virtually complete ^1H resonance assignments (reviews [1,2]): This can be achieved using a sequential method that relies on correlating through-bond and through-space (< 5 Å) connectivities in-

volving the backbone amide protons [1]. In particular, through-bond homonuclear correlation experiments are used to identify spin systems (i.e. intraresidue connectivities) while through-space nuclear Overhauser enhancement (NOE) experiments are used to identify interresidue connectivities. In this regard the most useful NOEs are of the type $\text{NH}(i)\text{-NH}(i+1)$, $\text{C}^\alpha\text{H}(i)\text{-NH}(i+1)$, $\text{C}^\beta\text{H}(i)\text{-NH}(i+1)$, $\text{C}^\alpha\text{H}(i)\text{-NH}(i+3)$ and $\text{C}^\alpha\text{H}(i)\text{-C}^\beta\text{H}(i+3)$. Difficulties arise, however, when there is a large number of overlapping NH and/or C^αH resonances. This is often the case in proteins containing more than ~ 100 residues and also in smaller helical proteins where the chemical shift dispersion of both the NH and C^αH resonances is small. One method of alleviating this problem is to correlate through-space and through-bond connectivities with the chemical shift of the directly bonded ^{15}N atom. This can be carried out by combining the ^1H -detected heteronuclear multiple quantum coherence (HMQC) experiment [3-10] with homo-

Correspondence address: G.M. Clore or A.M. Gronenborn, Laboratory of Chemical Physics, Building 2, Room 123, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA

Abbreviations: HMBC, ^1H -detected heteronuclear multiple-bond correlation spectroscopy; HMQC, ^1H -detected heteronuclear multiple quantum coherence spectroscopy; COSY, homonuclear correlated spectroscopy; NOESY, homonuclear nuclear Overhauser enhancement spectroscopy; HOHAHA, homonuclear Hartmann-Hahn spectroscopy

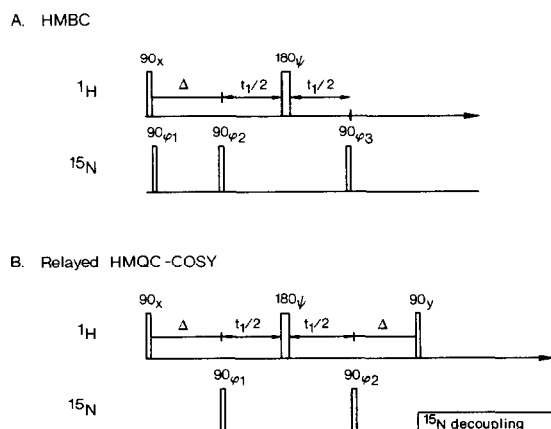


Fig.1. Pulse schemes for the ^{15}N - ^1H HMBC (A) and ^{15}N - ^1H relayed HMQC-COSY (B) experiments. The phase cycling for the HMBC experiment is as follows: $\psi = 8(x), 8(y), 8(-x), 8(-y)$; $\phi_1 = 2(x), 2(-x)$; $\phi_2 = x, -x$; $\phi_3 = 32(x), 32(-x)$; receiver = $8(x, -x), 8(-x, x), 8(x, -x), 16(-x, x), 8(x, -x), 8(-x, x), 8(x, -x)$. The phase of ϕ_2 is incremented by 90° for successive t_1 increments in order to obtain pure phase absorption in the F_1 dimension using the time-proportional incrementation (TPPI) method [28]. To compromise for the effects of the short ^1H transverse relaxation time, the delay Δ is set to a value significantly shorter than $1/(2^{15}J_{\text{NH}})$, namely 40 ms. No delay is used between the 90° (^1H) and the $90^\circ_{\phi_i}$ (^{15}N) pulses. The presence of such a delay normally serves to suppress one-bond correlations [29], a procedure not required in the present study. The phase cycling for the ^{15}N - ^1H relayed HMQC-COSY experiment is as follows: $\psi = 2(x), 2(-x)$; $\phi_1 = x, -x$; $\phi_2 = 4(x), 4(-x)$; receiver = $2(x, -x), 2(-x, x)$. To obtain a pure phase absorption spectrum using TPPI, the phase of ϕ_1 is incremented by 90° for every successive t_1 value. The delay Δ in this experiment is set to 4 ms, slightly shorter than $1/(2J_{\text{NH}})$. ^{15}N decoupling during the acquisition time (t_2) is achieved using the WALTZ-16 decoupling sequence [30]. Suppression of the water resonance in the relayed HMQC-COSY experiment is carried out by pre-saturation during the relaxation delay.

nuclear NOE [11], correlated (COSY) [12] or Hartmann-Hahn (HOHAHA) experiments [13,14]. Results from such experiments have recently been reported for the 75-residue helical protein *ner* from phage Mu [15]. Here, we direct attention to a complementary approach that involves correlating ^{15}N

chemical shifts with those of C^αH using ^1H -detected heteronuclear multiple-bond correlation (HMBC) spectroscopy [16,17]. A similar approach involving the correlation of the ^{13}C chemical shift of the carbonyl carbon atom with C^αH chemical shifts has been applied to small peptides [18–20]. In addition, because the magnitude of the $\text{C}^\alpha\text{H}(i-1)$ - $^{15}\text{N}(i)$ coupling is very sensitive to the ψ backbone torsion angle, qualitative structural information can be derived.

2. EXPERIMENTAL

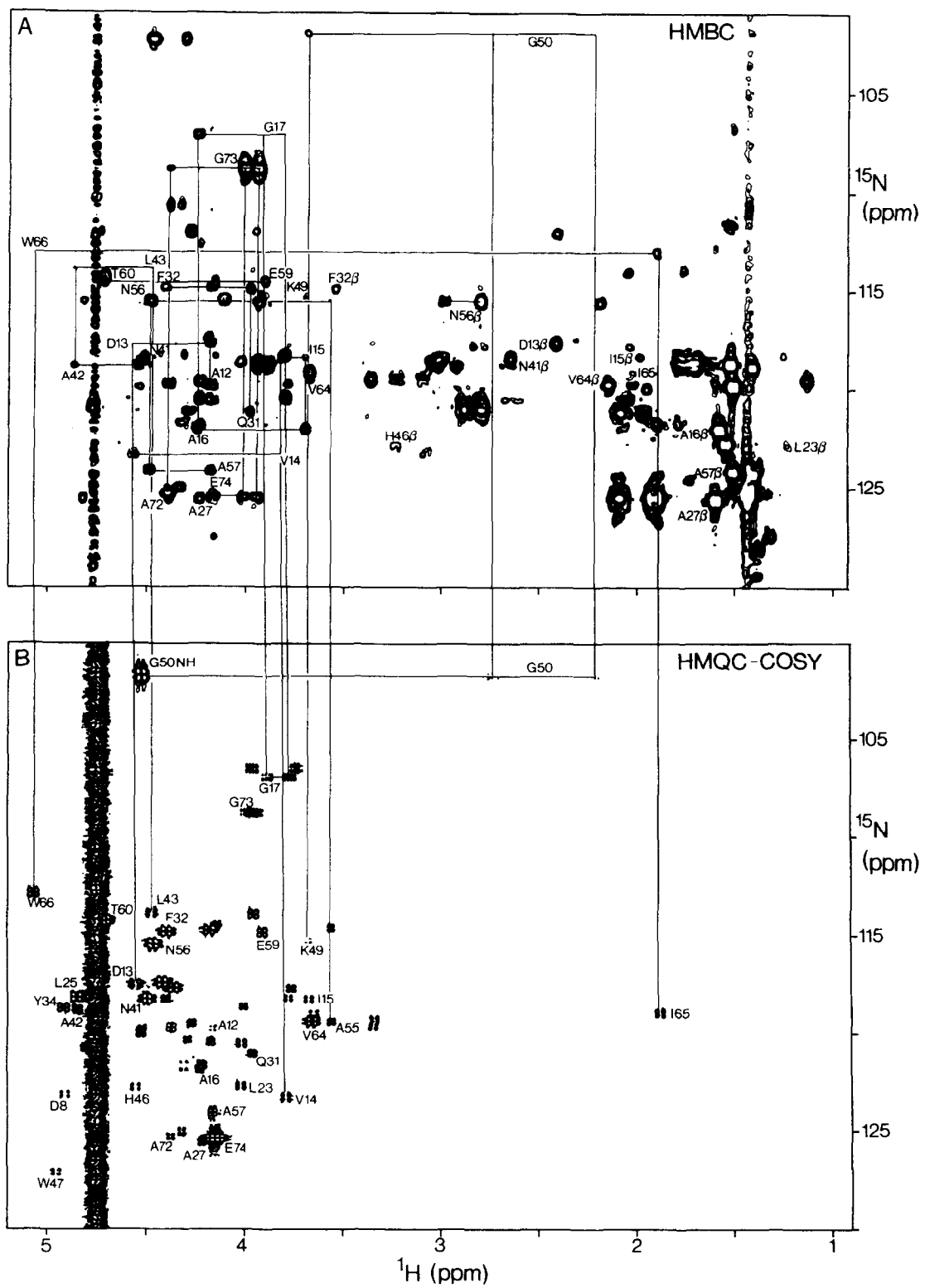
The protein *ner* from phage Mu was purified from *Escherichia coli* B containing the inducible plasmid pL-*ner* which directs high level production of the protein [21]. Complete ^{15}N labelling ($\sim 93\%$) was achieved by growing the bacteria in minimal medium using $^{15}\text{NH}_4\text{Cl}$ as the sole nitrogen source. The samples for NMR comprised 2 mM protein in either 99.996% D_2O or 90% $\text{H}_2\text{O}/10\%$ D_2O containing 150 mM sodium phosphate buffer, pH 7.0.

All NMR spectra were recorded on a Bruker AM-600 spectrometer at 27°C . The pulse schemes used for the HMBC [17] and relayed HMQC-COSY [15] experiments are given in fig.1. The HMBC experiment was recorded on the D_2O sample and 512 t_1 increments of 1 K data points were collected for spectral widths of 4166.67 and 5494 Hz in the ^{15}N and ^1H dimensions, respectively. The final digital resolution, after appropriate zero-filling, was 4.07 and 5.36 Hz/point in the ^{15}N and ^1H dimensions, respectively. The relaxation delay between scans was 1.5 s and 192 transients were recorded per increment, giving a total measurement time of ~ 48 h. The ^{15}N - ^1H relayed HMQC-COSY spectrum was recorded on the H_2O sample and 700 increments of 2 K data points were collected for spectral widths of 4166.67 and 10000 Hz in the ^{15}N and ^1H dimensions, respectively. The relaxation delay between scans was 1.5 s and 64 transients were recorded per increment, giving a total measurement time of ~ 22 h. The final digital resolution after zero-filling, was 4.07 and 2.88 Hz/point in the ^{15}N and ^1H dimensions, respectively.

3. RESULTS AND DISCUSSION

Sequential assignment using HMBC spectroscopy relies on detecting multiple-bond scalar correlations between the $^{15}\text{N}(i)$ atom and the $\text{C}^\alpha\text{H}(i)$

Fig.2. ^{15}N - ^1H HMBC spectrum in D_2O (A) and ^{15}N - ^1H relayed HMQC-COSY spectrum in H_2O (B) of 2 mM uniformly labelled ^{15}N Mu *ner* at 27°C . Labels are at the positions of the intraresidue correlations and a number of sequential connectivities are indicated. The peaks in the HMBC spectrum (A) arise from long range two- and three-bond correlations of the type $\text{C}^\alpha\text{H}(i)$ - $^{15}\text{N}(i)$, $\text{C}^\alpha\text{H}(i-1)$ - $^{15}\text{N}(i)$, $\text{C}^\alpha\text{H}(i)$ - $^{15}\text{N}(i)$, while those in the relayed HMQC-COSY spectrum (B) arise only from $\text{C}^\alpha\text{H}(i)$ - $^{15}\text{N}(i)$ connectivities via an indirect pathway involving two successive one-bond scalar connectivities, namely $^{15}\text{N}(i)$ - $\text{NH}(i)$ and $\text{NH}(i)$ - $\text{C}^\alpha\text{H}(i)$.



and $C^\alpha H(i-1)$ atoms. The size of the three-bond $C^\alpha H(i-1)-^{15}N(i)$ couplings follows the well-known Karplus behaviour [22], varying from -6 to 1 Hz, depending on the ψ peptide backbone torsion angle [23]. Very few data are available to predict the size of two-bond $C^\alpha H(i)-^{15}N(i)$ couplings in peptides and proteins. The fact, however, that in our experiments some of the corresponding intraresidue correlations are relatively intense indicates that this coupling can be as large as $4-5$ Hz. As the size of the multiple-bond couplings is smaller than the typical natural linewidth of the $C^\alpha H$ resonances, the sensitivity of this experiment is necessarily low and requires isotopic labelling with ^{15}N . The HMBC experiment is the most sensitive NMR method for observing such correlations and its pulse scheme is depicted in fig.1A. As it is impossible to record HMBC spectra in the pure absorption mode [16], the best compromise for optimizing both sensitivity and resolution is to record the experiment in a mixed mode, with pure phase absorption in the ^{15}N chemical shift dimensions (F_1) and absolute value in the 1H dimension (F_2) [17].

The HMBC experiment cannot distinguish between a $C^\alpha H(i)-^{15}N(i)$ and a $C^\alpha H(i-1)-^{15}N(i)$ correlation. The intraresidue $C^\alpha H(i)-^{15}N(i)$ connectivities, however, can be identified unambiguously by the relayed $^{15}N-^1H$ HMQC-COSY experiment [15] using an indirect pathway that involves two successive one-bond scalar connectivities, namely $^{15}N(i)-NH(i)$ and $NH(i)-C^\alpha H(i)$. The pulse scheme for this experiment is shown in fig.1B. We note that in favourable cases (e.g. small peptides) where there is little overlap of ^{15}N , NH and $C^\alpha H$ resonances, it should also be possible to discriminate the $C^\alpha H(i)$ from the $C^\alpha H(i-1)$ coupling to the $^{15}N(i)$ atom from a combination of $^{15}N-^1H$ HMQC and $^1H-^1H$ COSY experiments.

The application of the HMBC and relayed $^{15}N-^1H$ HMQC-COSY experiments is illustrated using the DNA-binding protein *ner* from phage Mu. This protein has been cloned and overexpressed in *E. coli* [21] and the determination of its solution structure is currently under way in our laboratory. The *ner* protein is mainly α -helical and consists of 75 residues. The HMBC spectrum of 2 mM uniformly ^{15}N -labelled *ner* in D_2O is shown in fig.2A. A reasonable number of correlations can be observed and stretches of sequential connectivities involving residues 12-17, 31-32, 41-43,

49-50, 55-57, 59-60 and 72-74 are indicated. The identification of the two-bond correlations in the HMBC spectrum is easily made by reference to the $^{15}N-^1H$ relayed HMQC-COSY spectrum recorded in H_2O (fig.2B). These assignments were aided by the simultaneous identification of spin systems using homonuclear HOHAHA spectroscopy (not shown) and confirmed by analysis of homonuclear NOESY spectra (not shown).

Fig.2A shows that not all of the expected two- and three-bond correlations are observed. The intensity of the long-range correlation is directly related to the magnitude of the time domain signal, detected during the acquisition time, t_2 . It is easily shown that the magnitude of this signal, $s(t_1, t_2)$, is described to a good approximation by the function

$$|s(t_1, t_2)| = \exp\{- (\Delta + t_1 + t_2)/T_2\} \sin(\pi J_{NA}\Delta) \times \sin(\pi J_{NA}t_2) \prod_k \cos[\pi J_{Ak}(\Delta + t_1 + t_2)]$$

where T_2 is the transverse 1H relaxation time, J_{NA} the $^1H-^{15}N$ long-range coupling, and J_{Ak} the coupling between protons A and k. It is clear from this expression that the intensity of the heteronuclear correlation is strongly dependent on (i) the size of the heteronuclear coupling, (ii) the transverse relaxation time T_2 , and (iii) the multiplicity of the 1H multiplet (i.e. the size of the couplings J_{Ak}). Considering that the 1H T_2 values in small proteins are relatively short (about 50 ms), couplings smaller than about 3 Hz are not expected to yield observable correlations.

The magnitude of the three-bond $C^\alpha H(i-1)-^{15}N(i)$ coupling is related to the ψ backbone torsion angle via a Karplus type relationship [22] and has a maximum value of ~ 6 Hz for $\psi = -60^\circ$. For $-120^\circ < \psi < 0^\circ$ the $C^\alpha H(i-1)-^{15}N(i)$ coupling constant is greater than 2 Hz, while for all other values of ψ it is less than 2 Hz [23]. α -Helices generally have ψ values between -40 and -80° while β -strands have ψ values between 40 and 180° [24], so that large three-bond couplings would be predicted for α -helices and small ones for β -sheets. Up till now, however, only peptide data are available for this coupling [23,25,26]. Nearly all the three-bond correlations that can be seen in the HMBC spectrum shown in fig.2A occur in α -helices, as ascertained by stretches of $NH(i)-NH(i+1,2)$, $C^\alpha H(i)-NH(i+3,4)$ and $C^\alpha H(i)-C^\beta H(i+3)$ connectivities in

the ^1H NOESY spectrum (unpublished), thus confirming the prediction of large three-bond couplings in α -helices. This nicely complements the homonuclear experiments as in helices the small size of the $^3J_{\text{NH}\alpha}$ couplings often makes it difficult to observe $\text{NH}(i)\text{-C}^\alpha\text{H}(i)$ correlations in either COSY or HOHAHA spectra and the $\text{C}^\alpha\text{H}(i-1)\text{-NH}(i)$ distance is relatively large such that the corresponding sequential NOE is not usually seen.

In addition to correlations involving the C^αH and ^{15}N atoms, some three-bond $^{15}\text{N}(i)\text{-C}^\beta\text{H}(i)$ correlations can be observed. In general, however, only *trans* couplings (~ 5 Hz) are sufficiently large [23] to yield significant correlations. Consequently, the $^{15}\text{N}(i)\text{-C}^\alpha\text{H}(i)$ correlations, supplemented by the relative magnitudes of intrasidue NOEs involving the C^βH protons on the one hand and the NH and C^αH protons on the other, can provide information on stereospecific assignments of β -methylene protons and on the conformation of the χ_1 side chain torsion angle. Such arguments have previously been used in studies of small peptides [25-27], and can now be applied, at least qualitatively, to proteins. For example, in the case of Asp 13, Leu 23, Asn 41, His 46 only one $^{15}\text{N}(i)\text{-C}^\beta\text{H}(i)$ correlation is observed. In a similar vein, the observation of a $^{15}\text{N}(i)\text{-C}^\beta\text{H}(i)$ correlation for Ile, Thr or Val residues is indicative of a χ_1 side chain torsion angle in the vicinity of $60 \pm 60^\circ$. This is the case, for example, for Ile 15, Ile 54 and Val 64. When both $^{15}\text{N}(i)\text{-C}^\beta\text{H}_2(i)$ correlations are observed, as in the case of Asn 56, it is likely that the side chain possesses conformational flexibility.

In conclusion, we believe that heteronuclear long-range correlation techniques will provide a powerful supplementary tool for the sequential resonance assignment of medium-sized proteins, complementing the existing methods relying on homonuclear experiments. Further, the presence or absence of a heteronuclear three-bond correlation reflects the size of the J coupling and therefore provides qualitative structural information.

Acknowledgement: This work was supported by the Intramural AIDS Targeted Antiviral Program of the Office of the Director of the National Institutes of Health.

REFERENCES

- [1] Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
- [2] Clore, G.M. and Gronenborn, A.M. (1987) *Protein Eng.* 1, 275-288.
- [3] Mueller, L. and Ernst, R.R. (1979) *Mol. Phys.* 38, 963.
- [4] Mueller, L. (1979) *J. Am. Chem. Soc.* 101, 4481-4484.
- [5] Bendall, M.R., Pegg, D.T. and Dodrell, D.M. (1983) *J. Magn. Reson.* 52, 81-117.
- [6] Redfield, A.G. (1983) *Chem. Phys. Lett.* 96, 537-540.
- [7] Markley, J.L., Westler, W.M., Chan, T.M., Kojiro, C.L. and Ulrich, E.L. (1984) *Fed. Proc.* 43, 2648-2655.
- [8] Roy, S., Redfield, A.G., Papastavros, M.Z. and Sanchez, V. (1984) *Biochemistry* 23, 4395-4400.
- [9] Sklenar, V. and Bax, A. (1987) *J. Magn. Reson.* 74, 469-479.
- [10] Glushka, J. and Cowburn, D. (1987) *J. Am. Chem. Soc.* 109, 7879-7871.
- [11] Jeener, J., Meier, B.H., Bachmann, P. and Ernst, R.R. (1979) *J. Chem. Phys.* 71, 4546-4553.
- [12] Aue, W.P., Bartholdi, E. and Ernst, R.R. (1976) *J. Chem. Phys.* 64, 2229-2246.
- [13] Braunschweiler, L.R. and Ernst, R.R. (1983) *J. magn. Reson.* 53, 52-528.
- [14] Davis, D.G. and Bax, A. (1985) *J. Am. Chem. Soc.* 107, 2821-2822.
- [15] Gronenborn, A.M., Bax, A., Wingfield, P.T. and Clore, G.M. (1988) *J. Am. Chem. Soc.*, in press.
- [16] Bax, A. and Summers, M.F. (1986) *J. Am. Chem. Soc.* 108, 2093-2094.
- [17] Bax, A. and Marion, D. (1988) *J. Magn. Reson.* 78, 186-191.
- [18] Bystrov, V.F., Okhanov, V.V., Miroshnikov, A.I. and Ovchinnikov, Y.A. (1980) *FEBS Lett.* 119, 113-119.
- [19] Bystrov, V.F., Okhanov, V.V., Arseniev, A.S., Adanasiev, V.A. and Gurevich, A.Z. (1980) *Biorg. Khim.* 6, 383-392.
- [20] Wynants, C., Hallenga, K., van Binst, G., Michel, A. and Zanen, J. (1984) *J. Magn. Reson.* 57, 93-98.
- [21] Allet, B., Payton, M., Mattaliano, R.J., Gronenborn, A.M., Clore, G.M. and Wingfield, P.T. (1988) *Gene* 65, 259-268.
- [22] Karplus, M. (1963) *J. Am. Chem. Soc.* 85, 2870-2871.
- [23] Bystrov, V.F. (1976) *Prog. Nucl. Magn. Reson. Spectrosc.* 10, 41-81.
- [24] Schulz, G.E. and Schirmer, R.E. (1979) *Principles of Protein Structure*, Springer, Berlin.
- [25] Cowburn, D., Live, D.H. and Fischman, A.J. (1983) *J. Am. Chem. Soc.* 105, 7453-7442.
- [26] Fischman, A.J., Live, D.H., Wyssbrod, H.R., Agosta, W.C. and Cowburn, D.J. (1980) *J. Am. Chem. Soc.* 102, 2533-2539.
- [27] Kessler, H., Griesinger, C. and Wagner, K. (1987) *J. Am. Chem. Soc.* 109, 6927-6933.
- [28] Marion, D. and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* 113, 967-974.
- [29] Kogler, H., Sørensen, O.W., Bodenhausen, G. and Ernst, R.R. (1983) *J. Magn. Reson.* 55, 157-167.
- [30] Shaka, A.J., Keeler, J., Frenkiel, T.A. and Freeman, R. (1982) *J. Magn. Reson.* 52, 535-540.