Tetrahedron Vol. 46, No. 9, pp. 3287-3296, 1990
Printed in Great Britain

0040-4020/90 $3.00+.00

^{1}H AND ^{13}C NMR CHEMICAL SHIFTS OF THE DIASTEREOTOPIC METHYL GROUPS OF VALYL AND LEUCYL RESIDUES IN PEPTIDES AND PROTEINS

Dedicated to Professor Yu Wang on the occasion of his 80th birthday.

D. NERI, G. OTTING, K. WÜTHRICH*
Institut für Molekularbiologie und Biophysik
Eidgenossische Technische Hochschule - Hönggerberg
CH-8093 Zürich, Switzerland

(*Received in Japan* 21 *December* 1989)

Abstract - Stereospecific ^{1}H and ^{13}C NMR assignments for the diastereotopic methyl groups of the valyl and leucyl residues in a cyclic peptide and in two globular proteins were obtained using biosynthetically directed fractional isotope labeling with ^{13}C. The methyl chemical shifts are analyzed with respect to possible correlations with the stereochemistry of these amino acid side chains.

The results presented in this paper are an offspring of studies on the nuclear magnetic resonance (NMR)[1] method for the determination of protein three-dimensional structures.[2] NMR in solution is presently the only method besides X-ray diffraction in protein single crystals that can be used to determine protein structures at atomic resolution. Although the first complete solution structure of a globular protein was reported only in 1985[3] the method is by now quite well established,[4] and there is keen interest in additional refinements of the structures obtained by NMR. One avenue toward this goal is by obtaining, in addition to the sequence-specific ^{1}H NMR assignments,[2,5] stereospecific assignments for the diastereotopic groups of protons in the individual amino acid residues.[2,6-9] In this context we introduced the technique of biosynthetically directed fractional ^{13}C labeling for the individual assignment of the diastereotopic methyl groups of Val and Leu.[10] Using this method, stereospecific assignments for the methyl groups of Val and Leu have so far been determined in cyclosporin A[10] and the DNA-binding domain of the 434 repressor protein.[11] In the present paper corresponding assignments are reported for the DNA-binding domain of the P22 c2 repressor protein, and for L-valine and L-leucine in the hydrolysate of this protein. Although the prime interest in these results is with regard to the refinement of the protein structure determinations, which will be reported elsewhere, they also represent an important collection of ^{1}H and ^{13}C chemical shift data on the diastereotopic methyl groups. These chemical shifts are compiled in Table 1 and analyzed in the last section of this paper.

Stereospecific NMR assignments for the methyl groups of valine and leucine by biosynthetically directed fractional ^{13}C-labeling make use of the fact that the

biosynthesis of the amino acids valine and leucine from glucose is known to be stereoselective.[12] Thereby, as is shown in Fig. 1, the isopropyl group is made up of a two-carbon fragment from one pyruvate unit, while the second methyl group is transferred from another pyruvate unit. This methyl migration has been shown to be stereoselective, and the migrating methyl group is found in the pro-S position in both valine and leucine, i.e., it corresponds to γ^2CH_3, or δ^2CH_3, respectively, in the standard IUB/IUPAC nomenclature for amino acids.

Biosynthetically directed fractional ^{13}C labeling of proteins can be achieved by growing microorganisms on minimal media containing a mixture of roughly 10% [$^{13}C_6$]-

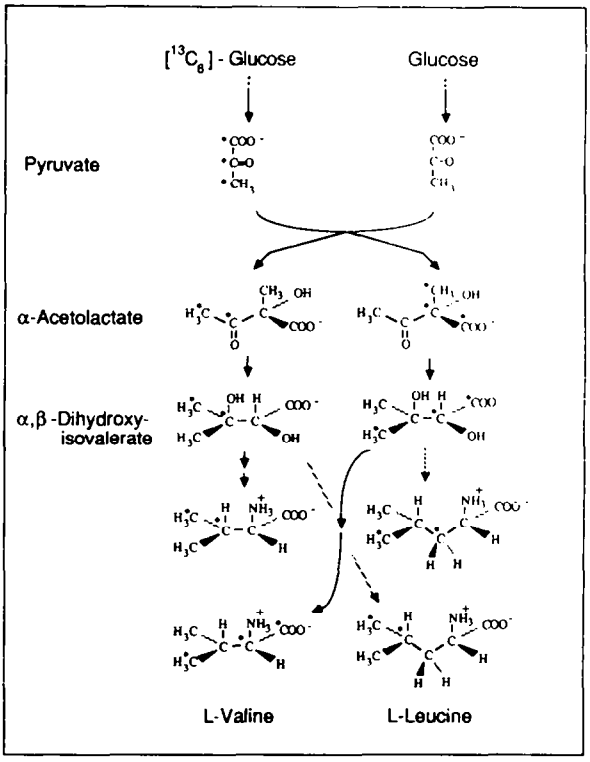

Fig. 1 Reaction pathways for the biosynthesis of valine and leucine from a mixture of fully ^{13}C labeled and unlabeled glucose, showing the stereochemistry and the principal labeling patterns (A star indicates a carbon position containing 10% ^{13}C, the absence of the star indicates that the ^{13}C abundance is 1.2% (see text)).

glucose and 90% unlabeled glucose as the sole carbon source. The carbon positions in such preparations are ^{13}C labeled to an extent of about 10%. Assuming a random distribution of the ^{13}C labels and taking into account the natural ^{13}C abundance of 1.1% in the unlabeled glucose, the probability that two adjacent carbon positions are labeled in the same molecule is then 1.2%. The only exceptions to this arise if two adjacent carbon atoms originate from the same source molecule, whence this probability becomes 10%. These two situations are both present in the isopropyl group in valine and leucine (Fig. 1): The pro-R methyl group (γ^1, or δ^1, respectively) and the adjacent >CH- group originate from the same pyruvate molecule and are, in the absence of isotope scrambling, labeled with ^{13}C in the same amino acid molecules. On the other hand, the pro-S methyl group and the adjacent carbon atom originate from two different pyruvate molecules (Fig. 1).

The stereospecific distinction between the pairs of isopropyl methyl groups in a fractionally ^{13}C labeled protein is clearly evidenced in ^1H-decoupled ^{13}C NMR spectra, where the ^{13}C resonance of the pro-R methyl group is a doublet with a splitting of about 33 Hz due to the one-bond ^{13}C-^{13}C coupling with the neighboring ^{13}C spin, while

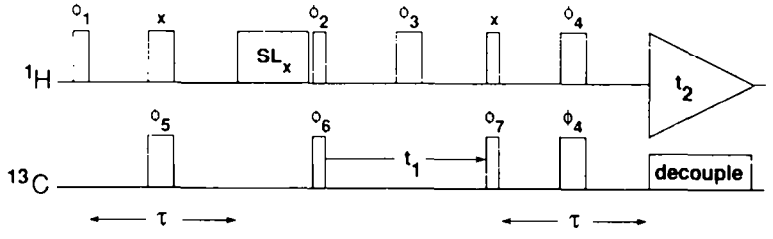

Fig. 2 Experimental scheme of [^{13}C,^1H]-COSY[14] which provides decoupling of the ^1H-^{13}C scalar couplings in both dimensions, and of the ^1H-^1H couplings in ω_1. 90-degree pulses are indicated by thin bars, and 180-degree pulses by broad bars. SL_x denotes a spin lock pulse of 2 ms duration with phase x, which purges the magnetization from ^{12}C-bound protons.[15] The delay τ is set to 1/2 $^1J(^1H,^{13}C)$ (in the experiment of Fig. 4, τ = 3.6 ms). The following 64-step phase cycle was used: $\Phi_1=(x,x,-x,-x)_{16}$, $\Phi_2=[(y)_4(-y)_4]_8$, $\Phi_3=[(x)_8(-x)_8]_4$, $\Phi_4=[(x)_{16}(-x)_{16}]_2$, $\Phi_5=(x)_{64}$, $\Phi_6=[(x)_4(-x)_4]_8$, $\Phi_7=(x,-x)_{16}(-x,x)_{16}$, receiver-$(x,-x,-x,x)_8(-x,x,x,x)_8$. The phases Φ_5 and Φ_6 are subjected to TPPI for quadrature detection in ω_1.[16] ^{13}C broadband decoupling during acquisition was obtained using WALTZ-16 with cycled sidebands.[17]

the ^{13}C NMR signal of the pro-S methyl group is a singlet. For work with proteins, 2D NMR experiments are usually employed to ensure a workable spectral resolution.[2,13]

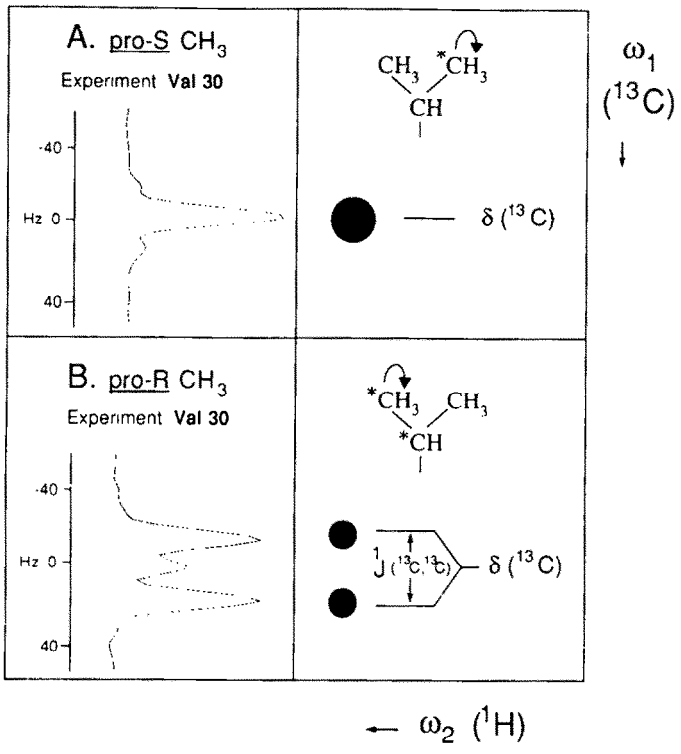

Fig. 3 The right side shows schematic drawings of the dominant multiplet fine structures expected for the methyl ^{13}C-^{1}H cross peaks of Val and Leu in [^{13}C,^{1}H]-COSY spectra recorded using the experiment of Fig. 2 with a protein preparation containing biosynthetically directed fractional ^{13}C labeling in the extent of ca. 10% (Fig.1). The arrows in the chemical structures indicate the relevant coherence transfer. The left side shows cross sections through the two methyl cross peaks of Val 30 in the spectrum of the P22 c2 repressor(1-76) in Fig. 4B. (A) For γ^2CH_3 of Val and δ^2CH_3 of Leu the cross peak consists of a single component. This pattern is the same as that seen for all methyl groups in a protein with natural abundance of ^{13}C (Fig. 4A). (B) For γ^1CH_3 of Val and δ^1CH_3 of Leu the cross peak consists of two components separated along ω_1 by the $^1J(^{13}C,^{13}C)$ coupling with the methine carbon. The weak signal in the center comes from the natural abundance of ^{13}C.

For the studies with the P22 c2 repressor(1-76) we used the experimental scheme of Fig. 2. Compared to the NMR experiments employed previously for the same purpose,[10,11] this experiment has the advantage to afford ^{13}C broadband decoupling during the ^{1}H observation period t_2, which results in improved sensitivity and spectral resolution. The pro-S and pro-R methyl resonances can readily be distinguished from the different multiplicities along the ω_1 axis, as is illustrated in Fig. 3 by schematic drawings as well as by an experimental example.

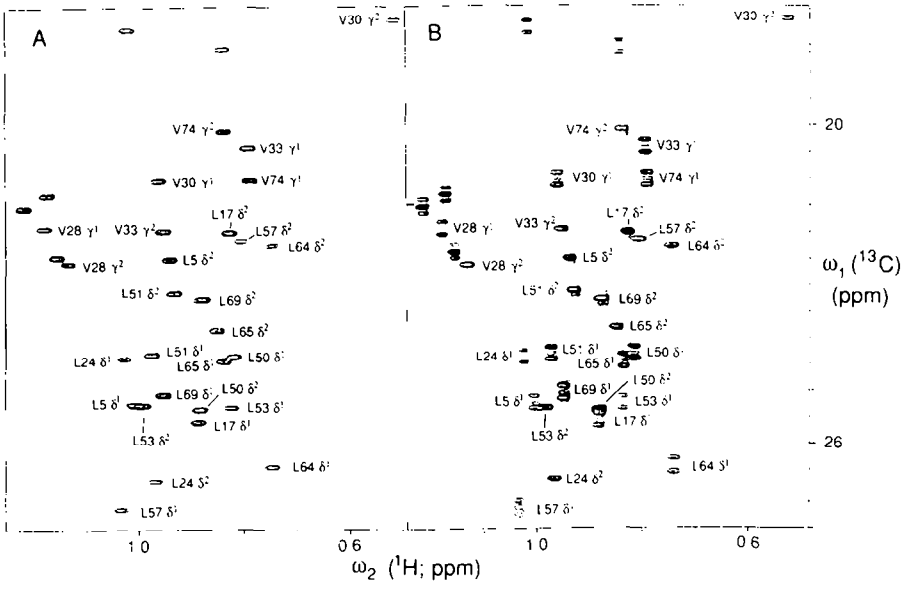

Fig. 4 [^{13}C,^{1}H]-COSY spectra of 5 mM solutions of P22 c2 repressor(1-76) in $^{2}H_2O$, p^{2}H = 4.8, T = 28°C, recorded with the experiment of Fig. 2. The spectral regions shown contain the cross peaks between the directly bonded ^{13}C and ^{1}H atoms of the methyl groups of valine and leucine. The resonance assignments are given by the one-letter code for the amino acids, the sequence position, and individual identifications of the methyl groups according to the standard IUB-IUPAC nomenclature, where γ^1 and δ^1 are pro-R, and γ^2 and δ^2 are pro-S. (A) Natural abundance of ^{13}C. Spectral parameters: ^{1}H-frequency 500 MHz, t_{1max} = 97 ms, t_{2max} = 204 ms, total recording time about 58 h. (B) Biosynthetically directed fractional ^{13}C labeling in the extent of 8%. Spectral parameters: ^{1}H-frequency 600 MHz, t_{1max} = 90 ms, t_{2max} = 188 ms, total recording time about 27 h.

In the presently used preparation of P22 c2 repressor(1-76), comparison of the ^{13}C satellites with the main peaks in a ^{1}H NMR spectrum showed that the extent of ^{13}C labeling was ca. 8%. The methyl resonances of the 4 Val and 10 Leu residues in this protein are all in the spectral region shown in Fig. 4. At natural abundance of ^{13}C (Fig. 4A), all methyl cross peaks have the fine structure of Fig. 3A, as expected. In the labeled protein (Fig. 4B) half of the methyl cross peaks of Val and Leu have the two-component fine structure of Fig. 3B, which identifies these methyls as γ^1 of Val, or δ^1 of Leu, respectively. Fig. 4 shows that all 28 ^{13}C-^{1}H methyl cross peaks of Val and Leu in the P22 c2 repressor(1-76) were resolved in the decoupled [^{13}C,^{1}H]-COSY spectrum, and complete stereospecific assignments were obtained (Table 1). The sequence-specific resonance assignments given in Fig. 4 are based on conventional sequential ^{1}H NMR assignments.[18]

Table 1 Methyl Chemical Shifts of L-Valine, L-Leucine and the Valyl and Leucyl Residues in Cyclosporin A, 434 Repressor(1-69) and P22 c2 Repressor(1-76)

Polypeptide	Residue	$\delta(^{1}H)$[ppm]$^{+}$			$\delta(^{13}C)$[ppm]$^{+}$		
		pro-R	pro-S	$\Delta\delta(^{1}H)^{*}$	pro-R	pro-S	$\Delta\delta(^{13}C)^{*}$
L-Valine		1.04	1.00	0.04	20.8	19.5	1.3
L-Leucine		0.96	0.94	0.02	24.9	23.8	1.1
Cyclosporin A[+]	Val 5	0.90	1.06	-0.16	18.5	19.8	-1.3
	MeVal 11	1.01	0.86	0.15	18.7	20.3	-1.6
	MeLeu 4	0.95	0.88	0.07	23.5	21.2	2.3
	MeLeu 6	0.85	0.94	-0.09	21.9	23.9	-2.0
	MeLeu 9	0.97	0.89	0.08	23.7	21.9	1.8
	MeLeu 10	0.98	0.98	0.00	23.8	23.4	0.4
434 Repressor[+]	Val 6	1.15	1.16	-0.01	21.3	25.9	-4.6
	Val 24	0.75	0.88	-0.13	22.6	24.1	-1.5
	Val 54	0.46	1.04	-0.58	21.3	21.2	0.1
	Val 56	1.02	1.09	-0.07	22.1	23.6	-1.5
	Val 68	0.97	0.97	0.00	21.4	21.7	-0.3
	Leu 13	1.02	0.96	0.06	26.2	22.9	3.3
	Leu 15	0.85	0.83	0.02	26.6	23.3	3.3
	Leu 20	0.89	0.95	-0.06	23.5	26.9	-3.4
	Leu 34	0.94	1.07	-0.13	24.8	28.3	-3.5
	Leu 45	0.98	0.92	0.06	25.7	25.4	0.3
	Leu 48	1.05	1.05	0.00	25.1	27.7	-2.6

Table 1 (continued)

	Leu 52	1.05	0.85	0.20	27.3	22.6	4.7
	Leu 59	0.94	1.00	-0.06	26.4	24.6	1.8
	Leu 60	0.89	0.87	0.02	25.6	23.3	2.3
P22 c2 Repressor	Val 28	1.17	1.13	0.04	21.9	22.5	-0.6
	Val 30	0.95	0.51	0.44	21.0	18.0	3.0
	Val 33	0.78	0.94	-0.16	20.4	21.9	-1.5
	Val 74	0.76	0.83	-0.07	21.0	20.0	1.0
	Leu 5	0.99	0.92	0.07	25.3	22.5	2.8
	Leu 17	0.88	0.83	0.05	25.6	22.0	3.6
	Leu 24	1.02	0.96	0.06	24.4	26.6	-2.2
	Leu 50	0.81	0.86	-0.05	24.3	25.4	-1.1
	Leu 51	0.96	0.92	0.04	24.3	23.1	1.2
	Leu 53	0.83	0.97	-0.14	25.2	25.3	-0.1
	Leu 57	1.02	0.80	0.22	27.1	22.1	5.0
	Leu 64	0.73	0.73	0.00	26.4	22.3	4.1
	Leu 65	0.83	0.85	-0.02	24.5	23.8	0.7
	Leu 69	0.94	0.87	0.07	25.0	23.3	1.7

+ The ^1H and ^{13}C chemical shifts are listed for the methyl groups of all Val and Leu residues in the compounds studied. Pro-R is γ^1 in Val and δ^1 in Leu; pro-S is γ^2 in Val and δ^2 in Leu. Chemical shifts were measured at the following conditions: L-valine and L-leucine, solvent ^2H_2O at p^2H = 1.0, T = 20°C, internal reference TSP; Cyclosporin A, solvent C^2HCl_3, T = 20°C, internal reference TMS; 434 repressor, solvent ^2H_2O at p^2H = 6.0, T = 28°C, internal reference TSP; P22 c2 repressor, solvent ^2H_2O at p^2H = 4.8, T = 28°C, internal reference TSP.

* $\Delta\delta = \delta$(pro-R) - δ(pro-S)

† From ref. 10

‡ From ref. 11

In addition to the data for the P22 c2 repressor(1-76) Table 1 lists the ^1H and ^{13}C chemical shifts of L-valine and L-leucine, which were measured in a hydrolysate of fractionally labeled P22 c2 repressor, and of cyclosporin A[10] and 434 repressor(1-69).[11] The chemical shift differences between the resonances of the pro-R and the pro-S methyls in each isopropyl group, $\Delta\delta(^1H)$ and $\Delta\delta(^{13}C)$, are also listed. Comparison of these chemical shift differences in different compounds may be more significant than comparison of the individual shifts, since these may be affected by the referencing used. In the free amino acids, for example, the chemical shift differences agree closely with those reported in the literature,[12] even though the

absolute values of the individual shifts are somewhat different. The chemical shift differences observed in Val and Leu were of course expected from the systematic studies by Mislow and Raban[19] of the inequivalence of the chemical shifts of diastereotopic substituents.

In the cyclic peptide cyclosporin A the ^1H chemical shifts among the individual valyl residues, or leucyl residues, respectively vary only over a narrow range of values, but the variation of the resulting chemical shift differences $\Delta\delta(^1H)$ is nonetheless about tenfold larger than the intrinsic value in the free amino acids. The same general statement applies for the ^1H shifts in the two proteins. The only exceptions are Val 54 in 434 repressor(1-69), and Val 30 in P22 c2 repressor(1-76), which can be attributed to ring current shifts.[2] This small dispersion of the ^1H chemical shifts is rather unusual for globular proteins,[22] but it is clearly a consequence of the fact that 434 repressor(1-69) and P22 c2 repressor(1-76) contain, respectively, only 2 and 3 aromatic residues. In spite of the scarcity of aromatic residues in the two proteins considered here, the nonbonding, conformation-dependent influences on the ^1H chemical shifts are clearly dominant when compared to the intrinsic shift difference between the diastereotopic substituents.

Table 2 Comparison of the methyl ^{13}C and ^1H chemical shifts of the homologous valyl and leucyl residues in 434 repressor(1-69) and P22 c2 repressor(1-76).

Homologous residues of Val and Leu*		Lower field methyl chemical shift+			
		$\delta(^1H)$		$\delta(^{13}C)$	
434	P22 c2	434	P22 c2	434	P22 c2
Leu 13	Leu 17	δ^1	δ^1	δ^1	δ^1
Leu 20	Leu 24	δ^2	δ^1	δ^2	δ^2
Val 24	Val 28	γ^2	γ^1	γ^2	γ^2
Leu 45	Leu 50	δ^1	δ^2	δ^1	δ^2
Leu 48	Leu 53	(δ^2)	δ^2	δ^2	(δ^2)
Leu 52	Leu 57	δ^1	δ^1	δ^1	δ^1
Leu 59	Leu 64	δ^2	(δ^2)	δ^1	δ^1
Leu 60	Leu 65	δ^1	δ^2	δ^1	δ^1
Val 68	Val 74	(γ^2)	γ^2	γ^1	γ^1

* Only those pairs of residues are listed that were reported in ref.21 to be in homologous positions in the two proteins.

\+ Each column lists the methyl group with the lower field chemical shift at fixed radio-frequency. Parentheses are used for those residues where the chemical shifts of the two diastereotopic methyls are nearly degenerate. γ^1 in Val and δ^1 in Leu are pro-R, γ^2 and δ^2 are pro-S.

Qualitatively similar observations as for the ^1H shifts can be made for the ^{13}C chemical shifts. Again the conformation-dependent shifts are severalfold larger than the intrinsic difference between the two diastereotopic carbon atoms. Furthermore, it is worth noting that there is no apparent correlation between the conformation-dependent effects on the ^1H and ^{13}C chemical shifts. This statement applies to the extent and sign of the individual chemical shifts as well as to the chemical shift differences within the individual isopropyl groups, indicating that the conformation-dependent shifts of the two spin types are caused by different mechanisms. Experimental evidence leading to similar conclusions had already been reported previously, based on sequence-specific ^1H and ^{13}C NMR assignments for methyl resonances in a protein.[20]

For continued studies with peptides and proteins there is little hope that the intrinsic chemical shifts of the diastereotopic methyls could be widely used for the purpose of facile resonance assignments. From various analyses of the data in Table 1 there resulted nonetheless an indication that the sign of the chemical shift difference $\Delta\delta(^{13}C)$ might be of some use in studies of homologous proteins. As is seen from Table 2, which compares the Val and Leu residues that occupy homologous positions in 434 repressor(1-69) and P22 c2 repressor(1-76), there is only a single exception to the rule that $\Delta\delta(^{13}C)$ has the same sign for homologous residues. In contrast, one observes a random distribution of the signs for $\Delta\delta(^1H)$ among the same homologous residues (Table 2).

Acknowledgments

We thank the Schweizerische Nationalfonds (project 31.25174.88) for financial support and Mr. R. Marani for the careful processing of the manuscript.

References

1. Abbreviations: NMR, nuclear magnetic resonance; [^{13}C,^1H]-COSY, two-dimensional ^{13}C-^1H correlated spectroscopy; TMS, tetramethylsilane; TSP, 3-(trimethylsilyl) [2,2,3,3-$^2H_4$]-propionate, sodium salt; P22 c2 repressor(1-76), N-terminal DNA-binding domain containing residues 1-76 of P22 c2 repressor; 434 repressor(1-69), N-terminal DNA-binding domain containing residues 1-69 of 434 repressor.
2. Wuthrich, K. NMR of Proteins and Nucleic Acids; Wiley; New York, **1986**.
3. Williamson, M.P.; Havel, T.F.; Wuthrich, K. J. Mol. Biol. **1985**, 182, 295-315.
4. (a) Bax, A.; Annu. Rev. Biochem. **1989**, 58, 223-256. (b) Clore, G.M.; Gronenborn, A.M. in Crit. Rev. Biochem. Mol. Biol. **1989**, 24, 479-564. (c) Kaptein, R.; Boelens, R.; Scheek, R.M.; van Gunsteren, W.F. Biochemistry **1988**, 27, 5389-5395. (d) Wright, P.E. Trend Biochem. Sciences **1989**, 14, 255-260. (e) Wuthrich, K. Acc. Chem. Res. **1989**, 22, 36-44. (f) Wuthrich, K. Science **1989**, 243, 45-50.
5. Wüthrich, K.; Wider, G.; Wagner, G.; Braun, W. J. Mol. Biol. **1982**, 155, 311-319.
6. Wuthrich, K.; Billeter, M.; Braun, W. J. Mol. Biol. **1983**, 169, 949-961.
7. Kline, A.D.; Braun, W.; Wuthrich, K. J. Mol. Biol. **1988**, 204, 675-724.
8. Guntert, P.; Braun, W.; Billeter, M.; Wuthrich, K. J. Am. Chem. Soc. **1989**, 111, 3997-4004.
9. Driscoll, P.C.; Gronenborn, A.M.; Clore, G.M. FEBS Lett. **1989**, 243, 223-233.
10. Senn, H.; Werner, B.; Messerle, B.; Weber, C.; Traber, R.; Wuthrich, K. FEBS Lett. **1989**, 249, 113-118.
11. Neri, D.; Szyperski, T.; Otting, G.; Senn, H.; Wüthrich, K. Biochemistry **1989**, 28, 7510-7516.
12. (a) Crout, D.H.G.; Hedgecock, J.R., Lipscomb, E.L.; Armstrong, F.B. Chem. Commun. **1980**, 304-305. (b) Sylvester, S.R.; Stevens, C.M. Biochemistry **1979**, 18, 4529-4531. (c) Hill, R.K.; Yan, S.; Arfin, S.M. J. Am. Chem. Soc. **1973**, 95, 7857-7859. (d) Hill, R.K.; Sawada, S.; Arfin, S.M. Bioorg. Chem. **1979**, 8, 175-189.

13. Ernst, R.R.; Bodenhausen, G.; Wokaun, A. Principles of Nuclear Magnetic Resonance in One and Two Dimensions; Clarendon Press; Oxford, **1987**.
14. Bodenhausen, G.; Ruben, D. Chem. Phys. Lett. **1980**, 69, 185-188.
15. Otting, G.; Wüthrich, K. J. Magn. Reson. **1988**, 204, 569-574.
16. Marion, D.; Wüthrich, K. Biochem. Biophys. Res. Comm. **1983**, 113, 967-974.
17. Shaka, A.J.; Barker, P.B.; Bauer, C.J.; Freeman, R. J. Magn. Reson. **1986**, 67, 396-401.
18. Otting, G. Strukturermittlung an kleinen Proteinen mit NMR: neue Methoden und Anwendungen am Beispiel der Cardiotoxine von Naja Mossambica Mossambica und des P22 c2 Repressors. Ph.D. Thesis Nr.8314, ETH Zurich, **1987**.
19. Mislow, K.; Raban, M. Topics Stereochem. **1967**, 1, 1-38.
20. Richarz, R.; Wüthrich, K. FEBS Lett. **1977**, 79, 64-68. Richarz, R.; Wüthrich, K. Biochemistry **1978**, 17, 2263-2269.
21. Sauer, R.T.; Yocum, R.R.; Doolittle, R.F.; Lewis, M.; Pabo, C.O. Nature **1982**, 298, 447-451.
22. Wüthrich, K. NMR in Biological Research: Peptides and Proteins; North Holland; Amsterdam, **1976**.