Volume 249, number 1, 113-118 FEB 07168 May 1989

Stereospecific assignment of the methyl ^{1}H NMR lines of valine and leucine in polypeptides by nonrandom ^{13}C labelling

H. Senn*, B. Werner, B.A. Messerle, C. Weber, R. Traber* and K. Wüthrich

Institut für Molekularbiologie und Biophysik, ETH-Hönggerberg, CH-8093 Zürich and **Präklinische Forschung, Sandoz Ltd, CH-4002 Basel, Switzerland*

Received 5 April 1989

An alternative method for the stereospecific assignment of the methyl groups of valine and leucine in the nuclear magnetic resonance (NMR) spectra of peptides and proteins is proposed, and its practical application is demonstrated with the assignment of all valine and leucine methyl groups in cyclosporin A. The method is based on the use of a mixture of fully ^{13}C-labelled and unlabelled glucose as the sole carbon source for the biosynthetic production of the polypeptide studied, knowledge of the independently established stereoselectivity of the reaction pathways of valine and leucine biosynthesis, and analysis of the distribution of ^{13}C labels in the valyl and leucyl residues of the product by two-dimensional heteronuclear NMR correlation experiments.

NMR; Protein structure; Stereoselective isotope labeling; Metabolic pathway; Resonance assignment; Methyl group stereospecific assignment

1. INTRODUCTION

Nuclear magnetic resonance (NMR) may be used to determine the three-dimensional structure of noncrystalline peptides and proteins [1]. A structure determination by this method includes the assignment of the ^{1}H NMR spectrum, the collection of conformational constraints using NMR measurements, and the calculation of the structure from these experimental input data. The precision of the protein structures thus obtained can be significantly improved if stereospecific assignments are available for the prochiral groups of protons [2]. The present paper describes a novel avenue for obtaining stereospecific ^{1}H NMR assignments for the methyl groups of valine and leucine in polypeptides.

For the assignment of protein ^{1}H NMR spectra the sequential assignment technique [1,3,4] is an efficient procedure, but it does not usually yield stereospecific assignments. As a consequence, the conformational constraints with prochiral groups of protons can only be defined relative to centrally located pseudoatoms [1,5]. This entails a considerable loss of structural information, which is particularly significant for NOE distance constraints with the methyl protons of Val and Leu. Stereospecific assignments for these isopropyl groups have in favorable cases been derived from distance constraints in low precision three-dimensional structures calculated from pseudoatom data (e.g. [6,7]), and for Val they may also be derived from strictly local conformational constraints [2]. The assignments described in this paper use ^{13}C isotope labelling and rely on knowledge about the stereospecificity of Val and Leu biosynthesis; they are thus independent of the molecular conformation and can be completed prior to the start of the structure calculations.

Labelling of amino acid residues selectively with deuterium has been used previously to determine unambiguously stereospecific assignments [8,9], and heteronuclear couplings between ^{1}H and ^{13}C have also been used for this purpose (e.g.[10,11]).

Correspondence address: H. Senn, Institut für Molekularbiologie und Biophysik, ETH-Hönggerberg, CH-8093 Zürich, Switzerland

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

However, on a much larger scale labelling with stable, NMR-active isotopes has made NMR a powerful tool for gaining insight into metabolic processes [12]. Standard labelling experiments in metabolic studies involve the observation of changes in NMR signal intensities during or following the administration of labelled precursors to an organism. They provide information about the metabolic processes occuring and the reaction rates of the chemical processes involved. Probably even more important, though not fully exploited, are nonrandom labelling techniques where, for instance, two adjacent labelled nuclei in a molecule effectively label the intervening chemical bond. This makes it possible to follow the fate of chemical bonds during metabolic precesses, and hence to obtain additional information about the reaction pathways involved. Nonramdom labelling can also result in patterns involving sites that are not directly adjacent. This paper demonstrates that based on the known stereochemistry of the substrates and the reaction pathways, nonrandom labelling patterns observed in enzymatically produced polypeptides can yield stereospecific assignments in the NMR spectra.

2. MATERIALS AND METHODS

Cyclosporin A (Sandimmun^R) was produced by growing the fungus *Tolypocladium inflatum* on a minimal medium with gluclose as the sole carbon source. The production of partially labelled cyclosporin A was started by replacing the natural [$^{12}C_6$]glucose with a mixture of roughly 90% [$^{12}C_6$]glucose and 10% uniformly labelled [$^{13}C_6$]glucose obtained from CIL, Cambridge. After 21 days the mycelium was harvested and the cyclosporin was extracted and purified [13]. The NMR sample was prepared by dissolving 20 mg of the ^{13}C-labelled cyclosporin in 0.6 ml of C^2HCl_3. This solution was studied in a sealed NMR tube. The NMR measurements were done at 20°C.

All NMR measurements were performed on a Bruker AM-360 spectrometer. Two-dimensional NMR spectra were acquired in the phase-sensitive mode, with time-proportional phase incrementation along ω_1 [14]. The data were processed using standard Bruker software. A homonuclear ^{13}C 2QF-COSY spectrum was acquired using the standard pulse sequence and standard phase-cycling [15]. WALTZ decoupling was used to remove the ^{1}H–^{13}C couplings in both dimensions of the spectrum. 860 t_1 values with 4096 data points each were acquired, and the sweep width in both dimensions was 18 000 Hz. The total experimental time was 60 h. A TOCSY-relayed (^{13}C,^{1}H) COSY experiment with a mixing time of 100 ms was measured using the pulse sequence and phase cycling of Otting and Wüthrich [16]. 4096 t_1 values with 4096 data points each were acquired over 48 h. The ^{1}H sweep width was 3500 Hz and the ^{13}C sweep width was 12 000 Hz.

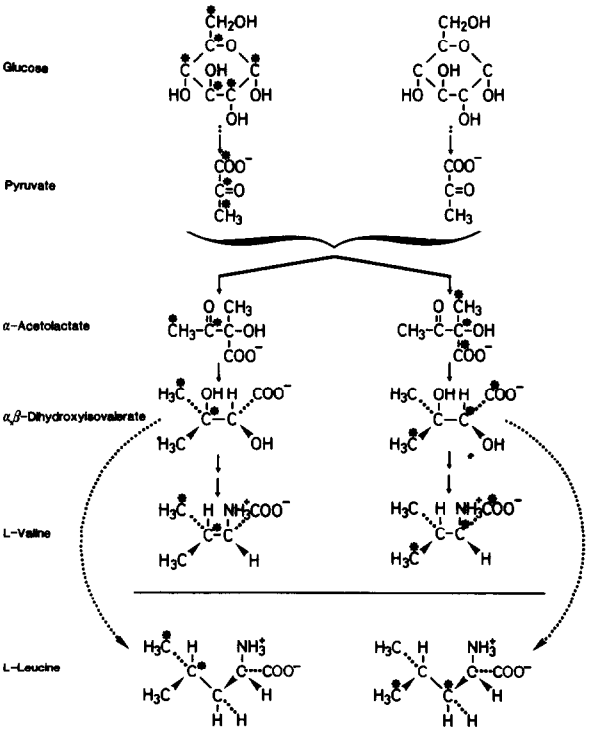

Fig.1. Reaction pathways for the biosynthesis of valine and leucine from a mixture of ^{13}C-labelled and unlabelled glucose, showing the stereochemistry and possible labelling patterns (* indicates a ^{13}C label, all other positions contain the natural ^{13}C abundance of 1.1%).

The approach used here for obtaining stereospecific NMR assignments relies on the fact that the biosynthesis of the amino acids valine and leucine from glucose is known to be stereoselective [8,9,17–19]. 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 (fig.1). This methyl migration has been shown to be stereoselective, and the migrating methyl group is *pro-S* in both valine and leucine, i.e. it is γ^2CH_3 or δ^2CH_3, respectively. Direct proof for this stereoselective biosynthetic pathway was obtained only for *E. Coli* and *Salmonella typhimurium*, but the wide occurrence of these reaction pathways among different species [20] suggests that similar stereoselectivity is present also in other species capable of synthesizing branched side chain amino acids.

3. RESULTS AND DISCUSSION

Nonramdomly labelled cyclosporin A was obtained from *T. inflatum* grown on a minimal medium with a mixture of roughly 90% [$^{12}C_6$]-glucose and 10% uniformly labelled [$^{13}C_6$]glucose as the sole carbon source. By integration of the ^{13}C satellites in the ^{1}H NMR spectrum, the peptide was

found to be uniformly labelled with ^{13}C to the extent of 14%. The relative peak intensities in the ^{13}C NMR spectrum were interpreted on the basis of the following statistical considerations. On supplying a mixture of unlabelled glucose and uniformly labelled [$^{13}C_6$]glucose for the biosynthetic production of the labelled polypeptide, there are two ways in which a pair of ^{13}C-labelled nuclei may be incorporated into adjacent positions. Either this may happen by chance, with a probability of ~2% in the present experiment ($2\% \approx 14\% \times 14\%$), or an intact pair may be incorporated from the labelled glucose, which would give a yield of 14%. The one-dimensional ^1H decoupled ^{13}C spectrum of cyclosporin A showed that the major products of the biosynthesis contained nonrandomly labelled sites, with additional small amounts of randomly labelled compounds indicating that metabolic scrambling processes are also occuring.

In the homonuclear ^{13}C 2QF-COSY spectrum (fig.2) it is clearly seen that for the side chains of both Val-5 and MeVal-11 only one strong cross peak connects the β-carbon and a γ-methyl carbon resonance. There is no cross peak correlating the resonance of the β-carbon with the second γ-methyl group. Following fig.1 the observed correlation must come from the intact βC-$\gamma^1 C$ (*pro-R*) fragment originating from the labelled glucose.

To benefit from the improved spectral resolution compared to homonuclear ^{13}C COSY, additional data were collected using TOCSY-relayed (^1H,^{13}C) COSY [16]. During the application of this pulse sequence, proton magnetization is first transferred to the directly bound carbon nuclei, where it is labelled with the ^{13}C frequency. After transfer back to the protons, the ^1H TOCSY relay distributes the magnetization among all spins of the proton spin system, as is indicated by the arrows in fig.3. In this way the magnetization derived from the methyl carbons was observed on the well resolved α-proton resonances. The four possible cross peak patterns that may result in this experiment from

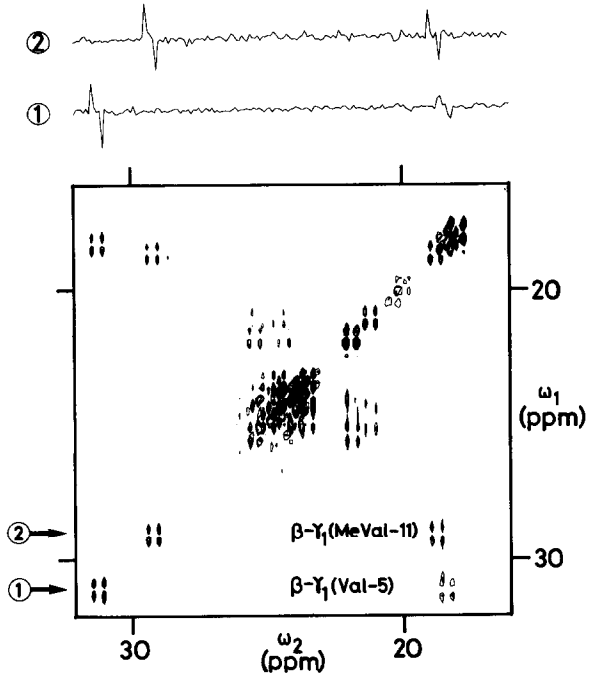

Fig.2. Part of the phase-sensitive homonuclear 90 MHz ^{13}C 2QF-COSY spectrum of nonrandomly ^{13}C-labelled cyclosporin A (20 mg in 0.6 ml $CDCl_3$, $T = 20°C$). The cross peaks due to the correlation between the β-carbon and the γ^1-methyl carbon (*pro-R*) of Val-5 and MeVal-11 are identified in the spectrum. At the top there are two cross sections along ω_2, which were taken at the ω_1 positions indicated by arrows.

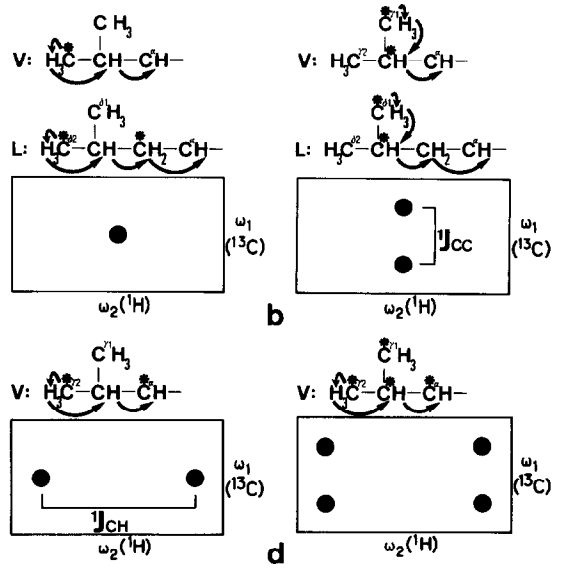

Fig.3. Schematic presentation of the methyl ^{13}C-α proton cross peak patterns expected for TOCSY-relayed (^1H,^{13}C) COSY with differently ^{13}C-labelled valine and leucine. (a) Neither the methyl ^{13}C nor the α proton has a one-bond coupling with ^{13}C. (b) The methyl ^{13}C is directly coupled to another ^{13}C spin. (c) The observed α proton is coupled directly to $\alpha^{13}C$. (d) Both the methyl ^{13}C and α proton are directly coupled to ^{13}C. The arrows indicate the coherence transfer pathway in TOCSY-relayed (^{13}C,^1H) COSY. In (b,c) $^1J_{CC}$ and $^1J_{CH}$ indicate one-bond spin-spin couplings. The molecular structures represent different ^{13}C-labelled species of Val and Leu resulting from the biosynthetic pathway of fig.1.

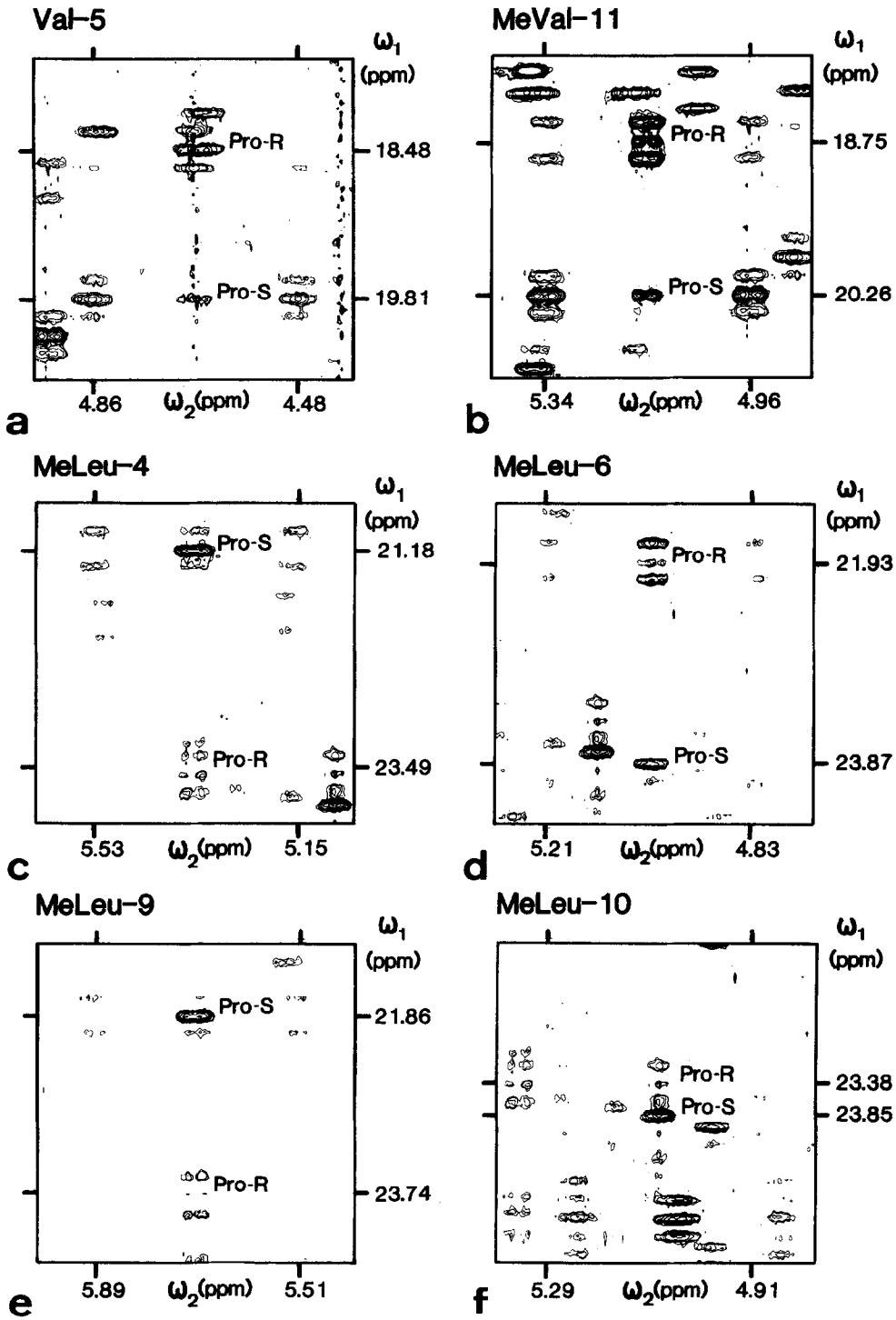

Fig.4. Cross peaks correlating the methyl ^{13}C spins of Val and Leu with the α protons in TOCSY-relayed (^{13}C,^{1}H) COSY experiments with nonrandomly ^{13}C-labelled cyclosporin A (20 mg in 0.6 ml $CDCl_3$, $T=20°C$, proton frequency 360 MHz). The methyl ^{13}C chemical shifts are indicated along ω_1 and the α proton chemical shifts along ω_2. The two Val (a,b) and four Leu (c–f) residues are identified at the top of each spectrum, and stereospecific assignments are given in the spectra.

different ^{13}C labelling are depicted in fig.3. A cross peak connecting an isolated ^{13}C label with a proton bound to ^{12}C consists of a single component (fig.3a). If the observed proton signal arises from a ^{13}C-bound proton, a splitting due to this one-bond coupling will result in the proton dimension of the spectrum (fig.3c). Similarly, in fig.3b a one-bond ^{13}C-^{13}C coupling is manifested along ω_1, and if both correlated spins are directly bound to ^{13}C one has the multiplet splittings along both frequency axes (fig.3d). Fig.3 further shows those ^{13}C-labelled species of Val and Leu resulting from the biosynthetic pathway of fig.1, which contribute the dominant components to the experimental spectra (fig.4). For valine (fig.4a,b) these are the structures of fig.3b for γ^1CH_3 (*pro-R*), and of fig.3c for γ^2CH_3 (*pro-S*), respectively, which confirms the result obtained with the experiment of fig.2. The additional Val species of fig.3a,d result from the synthetic pathway of fig.1 with approx. 10-fold lower yield, but the corresponding spectral components clearly dominate the background in fig.4a,b. For leucine, fig.1 predicts a dominant abundance of approx. 14% of the species of fig.3a,b, for δ^2CH_3 (*pro-S*) and δ^1CH_3 (*pro-R*), respectively. These dominant spectral components can readily be recognized in fig.4c–f, where the stereospecific assignments are indicated. The background is less clearly apparent than in the Val cross peaks, but it can qualitatively be rationalized with the basic patterns of fig.3.

The stereospecific assignments obtained are given in table 1, where they are compared with those determined by Loosli et al. [21] based on evaluation of ^1H-^1H NOE intensities and ^1H–^1H coupling constants. The presently used labelling technique provided also the stereospecific assignment of the ^1H and ^{13}C resonances of the methyl groups of Leu-10, which could not be determined with ^1H NMR alone. The agreement between the two sets of data in table 1 demonstrates that non-random labelling is a valid technique for obtaining stereoselectivity assignments of the isopropyl groups of Val and Leu. The observations made here in *T. inflatum* also support the previously made assumption that the stereoselectivity of Val and Leu biosynthesis in *E. coli* and *S. typhimurium* prevails also in other organisms. The biosynthesis of valine and leucine is clearly dominated by the metabolic pathways outlined in fig.1, with a direct flux of the glucose fragments into the peptide via the amino acid building blocks, and without much scrambling of the labels by alternative metabolic processes.

Table 1

The stereospecific assignments obtained for the ^{13}C and ^1H resonances of the methyl groups of valine and leucine in cyclosporin A

Amino acid	Methyl group chemical shifts[a]		Assignment from coherent labelling	Assignment from Loosli et al. [13]
	^{13}C	^1H		
Val-5	18.48	0.90	*pro-R*	*pro-R*
	19.81	1.06	*pro-S*	*pro-S*
MeVal-11	18.75	1.01	*pro-R*	*pro-R*
	20.26	0.86	*pro-S*	*pro-S*
MeLeu-4	23.49	0.95	*pro-R*	*pro-R*
	21.18	0.88	*pro-S*	*pro-S*
MeLeu-6	21.93	0.85	*pro-R*	*pro-R*
	23.87	0.94	*pro-S*	*pro-S*
MeLeu-9	23.74	0.97	*pro-R*	*pro-R*
	21.86	0.89	*pro-S*	*pro-S*
MeLeu-10	23.85	0.98	*pro-R*	not assigned
	23.38	0.98	*pro-S*	not assigned

[a] In ppm relative to internal tetramethylsilane (TMS)

The nonrandom labelling method proposed here requires no structural information for stereospecific assignments to be made. It therefore has the potential of yielding unambiguous assignments for all Val and Leu isopropyl groups, except in cases of complete resonance overlap. The observation of one clearly resolved cross peak is sufficient for the unambiguous stereospecific assignment of the respective methyl groups. It is possible to obtain nonrandomly labelled proteins by growing the bacteria for protein production on appropriately diluted [$^{13}C_6$]glucose or [$^{13}C_3$]glycerol. Alternatively, expression of recombinant proteins on amino acids supplemented with nonrandomly labelled valine and leucine may be used, where nonrandomly labelled valine and leucine can be obtained by growing bacteria or fungi which overproduce these amino acids on [$^{13}C_6$]glucose. As mentioned in section 1, obtaining complete stereospecific assignments for the methyl groups of valine and leucine may significantly improve the precision of a protein structure determination by NMR in solution [1,2].

Acknowledgements: Financial support by the Schweizerischer Nationalfonds (project 3.198.085) and the Ernst Hadorn-Stiftung (fellowship to B.W.) is gratefully acknowledged. We thank Mrs M. Schütz for the careful processing of the manuscript.

REFERENCES

[1] Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, Wiley, New York.
[2] Güntert, P., Braun, W., Billeter, M. and Wüthrich, K. (1989) J. Am. Chem. Soc., in press.
[3] Billeter, M., Braun, W. and Wüthrich, K. (1982) J. Mol. Biol. 155, 321–346.
[4] Wagner, G. and Wüthrich K. (1982) J. Mol. Biol. 155, 347–366.
[5] Wüthrich, K., Billeter, M. and Braun, W. (1983) J. Mol. Biol. 169, 949–961.
[6] Zuiderweg, E.R.P., Boelens, R. and Kaptein, R. (1985) Biopolymers 24, 601–611.
[7] Kline, A.D., Braun, W. and Wüthrich, K. (1988) J. Mol. Biol. 204, 675–724.
[8] Sylvester, S.R. and Stevens, C.M. (1979) Biochemistry 18, 4529–4531.
[9] Crout, D.H.G., Hedgecock, J.R., Lipscomb, E.L. and Armstrong, F.B. (1980) J. Chem. Soc. Chem. Comm. 7, 304–305.
[10] Feeney, J., Hansen, P.E. and Roberts, G.C.K. (1974) J. Chem. Soc. Chem. Commun., 465–466.
[11] Kessler, H., Griesinger, C. and Wagner, K. (1987) J. Am. Chem. Soc., 109, 6927–6933.
[12] London, R.E. (1988) Progress in NMR Spectroscopy 20, 337–383.
[13] Von Wartburg, A. and Traber, R. (1986) in: Cyclosporin, Progress in Allergy 38 (Borel, J.F. ed.) pp. 28–45, Karger, Basel.
[14] Marion, D. and Wüthrich, K. (1983) Biochem. Biophys. Res. Commun. 113, 967–974.
[15] Piantini, U., Sørensen, O.W. and Ernst, R.R. (1982) J. Am. Chem. Soc. 104, 6800–6801.
[16] Otting, G. and Wüthrich, K. (1988) J. Magn. Reson., 76, 569–574.
[17] Hill, R.K., Yan, S. and Arfin, S.M. (1973) J. Am. Chem. Soc., 95, 7857–7859.
[18] Hill, R.K. and Yan S. (1971) Bioorg. Chem. 1, 446–456.
[19] Hill, R.K., Sawada S. and Arfin, S.M. (1979) Bioorg. Chem. 8, 175–189.
[20] Stieglitz, B.I. and Calvo, J.M. (1974) J. Bacteriol. 118, 935–941.
[21] Loosli, H.R., Kessler, H., Oschkinat, H., Weber, H.P., Petcher, T.J. and Widmer, A. (1985) Helv. Chim. Acta 68, 682–704.