^{15}N NMR Spectroscopy

30†—Structure/Shift Relationships of Oligopeptides and Copolypeptides, Including Gramicidin S

Hans R. Kricheldorf
Institut für makromolekulare Chemie der Universität, Stefan Meier Straße 31, D-7800 Freiburg, FRG

The ^{15}N NMR spectra of various oligopeptide derivatives of the Z—X—Y—Y—OMe structure, where X and Y are variable amino acids and Z is the benzyloxycarbonyl group, were measured in several protic and aprotic solvents. The shift difference of the ^{15}N of the Y—Y and X—Y bond (neighbouring residue effect) is discussed with respect to the nature of X and Y with respect to the solvent. Oligopeptides of the Z—X—Y—Y—OH and $\overset{\oplus}{N}H_3$—X—Y—Y—OMe structures were compared with the Z-peptide esters to investigate the spectroscopic influence of the protecting groups. The methyl ester hydrochlorides of the 25 most common amino acids were measured in water and DMSO to elucidate the solvent dependence of the substituent effects. Moreover, the methyl ester hydrochlorides were compared with Z-amino acids and *N*-acetyl-amino acid methyl esters in DMSO to establish whether the substituent effects depend on the nature of the amino acid derivatives. In this connection the assignments of the serine, threonine and glycine signals are discussed with respect to silk proteins. Furthermore, the assignments of the signals of copolypeptides by comparison with oligo- and homo-polypeptides are discussed. Finally, it was demonstrated that intramolecular H bonds cause downfield shifts of 7–10 ppm of the acceptor amide groups.

INTRODUCTION

^{15}N NMR spectroscopy has found growing interest for the characterization of primary and secondary structures of both naturally occurring and synthetic peptides.[1–12] The numerous investigations on amino acid derivatives and other organic compounds[13] have shown that the ^{15}N nucleus is highly sensitive to structural and conformational changes of the molecules, and also to solvent effects. Hence, ^{15}N NMR spectra of peptides and proteins can contain much information; the assignments of the signals, however, may present serious difficulties.

By comparing dipeptides in water solution, Roberts *et al.*[14] concluded that the chemical shift of the peptide nitrogen of dipeptides can vary with the nature of the *N*-substituent (neighbouring residue effect). However, because dipeptides have one or two charged groups close to the peptide nitrogen their application as models for larger peptide sequences is questionable. However, Randall *et al.*,[5] investigating peptides of glycine and alanine, came to the conclusion that neighbouring residue effects do not play an important role in the ^{15}N NMR spectra of peptides and proteins. Yet, the prediction of the Gramicidin S signals from the corresponding *N*-acetyl-amino acids was not satisfactory.[5] We have shown in previous papers that large neighbouring residue effects do exist in polypeptides[15,16] and that the acidity of the solvents has a strong influence on the chemical shifts of the polypeptides.[17] Hence, it is not clear to what extent the neighbouring residue effects depend on the polymer structure and to what extent on solvent effects. Furthermore, the substituent effects of amino acid esters are a potential basis for the prediction of peptide shifts.[18] However, on a quite limited basis of measurements we have demonstrated that substituent effects can depend on solvation[19] and on the structure of the amino acid derivatives under investigation.[20] Thus, this work was undertaken to shed more light on the relationship between neighbouring residue effects, substituent effects and solvent effects, and to clarify to what extent simple amino acid derivatives can serve as spectroscopic models for peptides and proteins.

† For Part 29, see H. R. Kricheldorf, *Polym. Bull.* in press.

EXPERIMENTAL

Materials

99% ^{15}N-enriched amino acids were purchased from Stohler Isotope Chemicals (Waltham, Massachusetts); they were diluted with normal amino acids (purchased from Merck & Co., Darmstadt, FRG) to 1% ^{15}N-enrichment. Gramicidin S (bishydrochloride) was obtained from Sigma Chemicals.

Peptide syntheses (general considerations)

All peptide derivatives listed in Tables 1–4 were prepared in a stepwise manner from Z-protected amino acids (see later for definition), α-amino acid methyl ester hydrochlorides and 2-ethoxy-1-ethoxycarbonyl-1,2-dihydroquinoline (EEDQ) as condensing reagent.[21] Crystalline end products were obtained in more than 90% of these syntheses. Their structures were checked by C-, H- and N-elemental analyses and ^{1}H NMR spectra and in some cases by ^{13}C NMR spectra. However, in most cases the products were not recrystallized or purified in any other way, because our

CCC-0030-4921/81/0015-0162$08.00

NMR measurements were not sensitive to impurities (in particular, quinoline). In the case of the compounds shown in Table 5 the *tert*-butoxy carbonyl group (BOC) was used for the peptide syntheses and finally removed by means of pure trifluoroacetic acid; all peptide ester trifluoroacetates crystallized well. Furthermore, *N,N'*-carbonyldiimidazole or isobutyl chloroformate were used as condensing reagents for the preparation of the oligopeptides listed in Tables 5 and 14. Examples of the three different synthetic procedures used are given below.

N-Acetyl-L-leucine-D-phenylalanine methyl ester. 16.3 g (0.1 mol) *N*-acetyl-L-leucine and 22 g (>0.1 mol) D-phenylalanine methyl ester hydrochloride were suspended in a mixture of 250 ml dry dimethylformamide and 50 ml methanol. 14 ml Triethylamine and 25 g (0.1 mol) 2-ethoxy-1-ethoxycarbonyl-1,2-dihydroquinoline[21] were added, and the thickening solution was vigorously shaken. After 24 h the initially formed paste of crystals had become a clear solution which was concentrated *in vacuo*. The residue was dissolved in 500 ml ethyl acetate and washed successively with two 100 ml portions of 1 N HCl, 100 ml 10% (by weight) K_2CO_3 solution and finally with 5% (by weight) citric acid solution. The organic phase was dried over Na_2SO_4 and concentrated *in vacuo*. The crystallization of the product was completed by portionwise addition of ligroin: yield 25.8 g (77%); m.p. 138–140 °C.

$$[\alpha]_D^{20} = -2.0 \ (c = 2 \text{ in DMSO})$$

Analyses calcd for $C_{18}H_{26}N_2O_4$ (334.4): C 64.64, H 7.83, N 8.37; found: C 64.99, H 7.57, N 8.17.

N-Methoxycarbonylvalyl-ornithine methyl ester N^δ-hydrochloride. 23.2 g (0.1 mol) N^δ-BOC-Ornithine and 40 g (0.25 mol) hexamethyldisilazane were refluxed in 100 ml dry toluene until the amino acid had dissolved. The reaction mixture was concentrated *in vacuo*, the residue diluted with 150 ml dry toluene and concentrated again.

17.6 (0.1 mol) *N*-methoxycarbonylvaline and 14.0 g (20.1 mol) isobutyl chloroformate were dissolved in 200 ml dry tetrahydrofuran and cooled to −10 °C. 14.5 ml Triethylamine were then added dropwise, and after 10 min the silylated N^δ-BOC ornithine. The reaction mixture was stirred at *c.* 0 °C for 1 h, diluted with 300 ml ethyl acetate and washed twice with 200 ml of a 10% (by weight) citric acid solution. The organic phase was dried over Na_2SO_4 and concentrated *in vacuo*.

The syrupy *N*-methoxycarbonylvalyl-N^δ-BOC-ornithine obtained was dissolved in 300 ml dry methanol and refluxed with 20 ml chlorotrimethylsilane until CO_2 evolution had ceased (*c.* 30 min). The reaction mixture was then concentrated *in vacuo* and the product crystallized by treatment with tetrahydrofuran.

Yield: 18.5 g (53%); m.p. 156–158 °C

$$[\alpha]_D^{20} - 14.5 \ (c = 2 \text{ in DMSO})$$

Analyses calcd for $C_{13}H_{26}N_3O_5Cl$ (339.8): C 45.94, H 7.71, N 12.36; $C_{13}H_{28}N_3O_6Cl$ (356.7): C 43.77, H 7.91, N 11.78; found: C 44.55, H 7.93, N 11.54.

The analyses suggest that the product contains *c.* 1 mol H_2O even after drying at 50 °C/12 mm. After recrystallization the same analyses were obtained as before. The ^{1}H NMR spectrum (in DMSO with TMS as internal standard) exhibits the two singlets of the methoxycarbonyl and methyl ester groups (equal intensities) at 3.57 and 3.65 ppm, whereas the signal of the BOC group was absent.

N-Benzyloxycarbonylglycyl-DL-threonine methyl ester. 10.5 g (50 mmol) Z-Glycine and 8.1 g (50 mmol) *N,N'*-carbonyldiimidazole were heated in 50 ml dry tetrahydrofuran until CO_2 evolution had ceased (solution A).

9.0 g (>50 mmol) D,L-Threonine methyl ester hydrochloride were suspended in 200 ml dry methylene chloride, 11.0 g (100 mmol) chlorotrimethylsilane were added immediately, and 15.2 g (150 mmol) triethylamine were added dropwise while stirring. The reaction mixture was refluxed for 30 min, cooled to room temperature and combined with solution A.

The reaction mixture was refluxed for 2 h, concentrated *in vacuo* and the residue diluted with 100 ml tetrahydrofuran and 300 ml ethyl acetate. This solution was stirred with 150 ml 1 N HCl for 2 h, and the organic layer was separated and washed with 100 ml of a 5% (by weight) citric acid solution, 100 ml of a 5% (by weight) K_2CO_3 solution and again with dilute citric acid. The organic phase was dried over sodium sulphate and concentrated *in vacuo*. The residual product was crystallized by treatment with diethyl ether and ligroin while cooling with ice.

Yield: 23.8 g (71%); m.p. 108–110 °C

Analyses calcd for $C_{15}H_{20}N_2O_6$ (323.2): C 55.72, H 5.92, N 8.66; found C 55.90, H 5.96, N 8.65.

All serine derivatives of Table 12 were prepared analogously and crystalline products with correct elemental analyses were obtained in all cases.

Measurements

The natural abundance ^{15}N NMR measurements shown in Tables 1–10 and 12–14 were carried out with a Bruker WH-90 FT spectrometer at 28–30 °C. 1.5 g Product were dissolved in 6–7 ml solvent and the spectra recorded in 20 mm diameter sample tubes with a coaxial 5 mm tube containing D_2O. The $NO_3^\ominus$ ion of a 25% (by weight) $^{15}NH_4{}^{15}NO_3$ solution in D_2O (5 mm tube) served for shift referencing. The frequency of the standard was insensitive to both the nature of the solvent and the solute. The shift difference between the $NO_3^\ominus$ and the $NH_4^\oplus$ ion was 356.5 ppm. The following acquisition parameters were used; pulse width 30 μs (*c.* 35°) for all low molecular weight amino acid derivatives but 40 μs (*c.* 45°) for the polypeptides (Table 7); 1 K data points for a spectral width of 500 Hz (or 2 K for 1000 Hz) were used for acquisition and 'zero-filled' to 2 K (or 4 K) for Fourier transform. The exponential line broadening was 1.5 Hz, or 0.8 Hz if required for better resolution. A delay time of 4.0 s was applied in the case of proline

and sarcosine derivatives. The number of accumulated transients was *c.* 1000–3000 for the measurements shown in Tables 8–10 and 13, 10 000–40 000 for the data of Tables 1–5, 12 and 14 and 40 000–50 000 for the copolypeptides of Table 6. In the case of the compounds in Tables 1–5 and 14 the molar concentrations were of the order of 0.4–0.8, while the amino acid derivatives of Tables 8–10 and 12 were measured at concentrations of 0.8–1.5 M. The measurements of Table 6 were carried out with 1.5 g polypeptide in 7 ml solvent.

The 40.55 MHz spectra of Fig. 2 and Table 7 were run at 30 °C on a Bruker WH 400 spectrometer for 1 g polypeptide in 4.5 ml TFA. A 15 mm sample tube with a 5 mm coaxial tube containing a $^{15}NH_4{}^{15}NO_3$ solution in D_2O was used for lock and shift referencing. The following acquisition parameters were applied: pulse width 45 μs (70°); 2 K data points for a spectral width of 2400 Hz (4 K for Fourier transform); delay time for inverse-gated ^1H decoupling 2.0 s; exponential line broadening 3.0 Hz; *c.* 20 000–40 000 transients.

RESULTS AND DISCUSSION

Neighbouring residue effects in oligopeptides

In a previous paper dealing with the ^{15}N NMR spectra of sequence polypeptides we have defined neighbouring residue effects (NREs) as the shift difference between peptide group X—Y and Y—Y.[16] This shift difference measures the spectroscopic influence of the amino acid unit X on Y with the homologous peptide group Y—Y as the internal standard. To study neighbouring residue effects in oligopeptides according to the above definition, and to avoid the disturbing influences of charged end groups, we have synthesized and measured in this work a series of tripeptides with the structure Z—X—Y—Y—OMe, where Z is the *N*-benzyloxycarbonyl protecting group. Some higher oligopeptide esters containing the X—Y—Y segment were also included. Z-protected tripeptides with free carboxyl groups (Z—X—Y—Y—OH) were compared with the corresponding methyl esters to elucidate the shift effect of the ester group (Tables 1–4). Furthermore, tripeptide esters with an ammonium end group ($\overset{\oplus}{N}H_3$—X—Y—Y—O—R) were recorded to establish the spectroscopic influence of a positively charged end group (Table 5). Those Z-protected oligopeptides having an X—Gly—Gly sequence are described in Table 1 and those with an X—Ala—Ala sequence in Table 2; those with an X—Val—Val, X—Leu—Leu or X—Met—Met sequence are described in Table 3 and those with an X—Phe—Phe block in Table 4. The syntheses of all these peptide derivatives were carried out in such a way that one of the Y-units was ^{15}N enriched (denoted by an asterisk) to allow an unambiguous assignment of all signals (Fig. 1). All oligopeptides were measured in five different solvents (as far as possible) in order to investigate the influence of solvation on the neighbouring residue effects (for discussion see below).

Comparison of the data in Tables 1–4 allows the derivation of one spectroscopic rule which seems to hold without exception: 'The sign (or sense) of the neighbouring residue effect (NRE) which X causes on Y is always opposite to the NRE which Y causes on X'.

For example, in Z—Leu—Gly—*Gly—OEt the Leu—Gly signal appears downfield of the Gly—Gly signal, while the Gly—Leu signal of Z—Gly—Leu—*Leu—OMe is upfield of the Leu—*Leu signal. This rule also holds for the peptide esters with a charged end group (Table 5). In order to test this rule, pairs of tripeptide esters with complementary sequences, e.g. X—Y—Y/Y—X—X, were measured and listed in Table 5. A more detailed study of the data in Tables 1–5 reveals further rules with a more limited scope. All amino acids (except alanine in acidic solvents) cause downfield shifts of Gly signals, i.e. a positive sign of the NRE in sequences of the type X—Gly—Gly. Glycine itself causes upfield shifts when attached to the nitrogen of other amino acids (except alanine in aprotic solvents, Table 2), i.e. the sign of the NRE is negative for Gly—Y—Y. When amino acids with saturated aliphatic side chains are compared with each other in both types of sequences, X—Gly—Gly and Gly—Y—Y, the NREs become larger with increasing bulkiness of the side chain, i.e. in the following order:

$$\text{Ala} < \text{Leu} < \text{Val} < \text{Ile}$$

This observation allows the conclusion, on the one hand, that NREs have a partially steric origin. On the other hand, the data of β-Ala—Y—Y sequences demonstrate that β-Ala causes strong downfield shifts of the attached Y nitrogen (positive NRE) by analogy with Val or Ile, even though it does not have a bulky side chain. Other unbranched ω-amino acids behave like β-Ala, as shown for Z-γ-Abu—Gly—*Gly—OEt in Table 1 and for various $(X—Gly—Gly)_n$ polypeptides in previous papers.[15,18] Hence, it must be concluded that another important contribution to the NRE results from an inductive, electronic influence of the *N*-acyl residue. Another tendency becomes detectable when the following pairs of *N*-protected tripeptide esters are compared: Gly—Gly—*Gly/Ala—Gly—*Gly, Gly—Leu—*Leu/Ala—Leu—*Leu, Gly—Val—*Val/Ala—Val—*Val, Gly—Phe—*Phe/Ala—Phe—*Phe. In all these cases the Ala—Y bond absorbs upfield of the Gly—Y bond; i.e. the NRE of Ala is more negative. Of course, a rule based only on four measurements may not be free of exceptions.

The data of Tables 1–5 allow, furthermore, the influence of the protecting groups on the chemical shifts and NREs of the tripeptide derivatives to be considered. When Z-tripeptides are compared with the corresponding esters in an aprotic solvent one finds that the *Y nitrogen in Z—X—Y—*Y—OH absorbs, in most cases, downfield of that in Z—X—Y—*Y—OAlkyl. *N*-protected dipeptides and their esters behave similarly, and a potential explanation of this spectroscopic behaviour has been presented in a previous paper.[22] The NREs of *N*-protected tripeptides and their esters have the same sign, but the value can differ up to 0.8 ppm. However, the pair Z—Gly—Ala—*Ala—OH/Z—Gly—Ala—*Ala—OMe is an ex-

Table 1. ^{15}N NMR chemical shifts[a] (ppm) and neighbouring residue effects Δδ (ppm)[b] of Z-protected oligopeptide esters with Gly—*Gly units

Gly—Gly Derivative	Formic acid[c] δ	Δδ	Chloroethanol δ	Δδ	DMSO δ	Δδ	CH_2Cl_2 δ	Δδ	Pyridine δ	Δδ
Tos—Gly—Gly—*Gly—OEt	−285.0		−284.6		−284.7				−284.3	
	−267.2	+0.6	−268.9	+1.3	−270.8	+1.8	Insoluble		−271.0	+2.6
	−267.8*		−270.2*		−272.6*				−273.6*	
Z—β—Ala—Gly—*Gly—OH	−292.0				−293.6					
	−262.9	+5.0	Insoluble		−266.6	+4.7	Insoluble		Insoluble	
	−267.9*				−271.3*					
Z—β—Ala—Gly—*Gly—OEt	−292.0		−293.1		−293.5				−294.5	
	−263.2	+4.8	−263.1	+7.1	−267.0	+5.5	Insoluble		−267.4	+6.1
	−268.0*		−270.2*		−272.5*				−273.5*	
Z—Ala—Gly—*Gly—OEt	−284.5		−284.8		−284.9				−285.2	
	−269.2	−1.3	−270.4	−0.4	−272.9	±0	—	—	−272.9	+0.9
	−267.9*		−270.0*		−272.9*				−273.8*	
Z—α—n—Abu—Gly—*Gly—OEt	−287.8		−287.9		−288.2				−288.0	
	−267.4	+0.4	−268.6	+1.6	−271.1	+1.9	Insoluble		−270.6	+3.2
	−267.8*		−270.2*		−273.0*				−273.8*	
Z—Val—Gly—*Gly—OEt	−290.7		−290.5		−290.7		−292.2		−290.5	
	−265.7	+2.2	−266.8	+3.4	−269.1	+3.8	−268.9	+4.4	−268.5	+5.3
	−267.9*		−270.2*		−272.9*		−273.3*		−273.8*	
Z—Ile—Gly—*Gly—OEt	−288.2				−288.9					
	−264.2	+3.5	—	—	−268.4	+4.4	Insoluble		—	—
	−267.7*				−272.8*					
Z—Leu—Gly—*Gly—OEt	−287.1				−287.8		−288.0		−287.8	
	−267.5	+0.5	—	—	−271.8	+1.2	−271.4	+1.4	−272.3	+1.4
	−268.0*				−273.0*		−272.9*		−273.7*	
Z^{α}—Z^{ε}—Lys—Gly—*Gly—OEt	—		−288.1		−287.6				−288.5	
	−289.5		−291.4		−292.4		—	—	−293.7	
	−267.9	±0	−267.8	+0.7	−271.4	+1.4			−272.3	+1.9
	−267.9*		−268.5*		−272.8*				−274.2*	
Z—Gly—Met—*Gly—OH	−299.3				−299.5				−300.3	
	−257.4		Insoluble		−261.0		Insoluble		−260.5	
	−266.6*				−270.3*				−269.8*	
Z—Gly—Met—*Gly—Gly—OEt	−299.7		−300.0		−299.5				−300.4	
	−257.3		−258.7		−260.4		Insoluble		−259.5	
	−267.2*	+0.6	−269.0*	+1.2	−270.7*	+2.1			−270.2*	+3.3
	−267.8		−270.2		−272.8				−273.5	
Z—Phg—Gly—*Gly—OEt	−290.4				−287.8					
	−266.3	+1.8	—	—	−269.5	+3.0	Insoluble		Insoluble	
	−268.1*				−272.5*					
Z—Tyr—Gly—*Gly—OEt	−286.9		−288.0		−287.9				−288.8	
	−266.7	+1.1	−267.9	+2.4	−271.0	+1.7	Insoluble		−269.7	+4.2
	−267.8*		−270.3		−272.7*				−273.9*	
Z—Trp—Gly—*Gly—OEt	−286.6		−286.8		−288.6				−288.2	
	−250.4		−248.4		−245.5		Insoluble		−246.7	
	−266.7	+1.4	−268.5	+1.5	−271.6	+1.5			−271.3	+2.3
	−268.1*		−270.0		−273.1*				−273.6*	
Z—Gly—Pro—Gly—*Gly—OEt	−299.8		−300.7		−300.7					
	−267.7[d]	+0.3	−269.2	+0.6	−271.2[d]	+2.1	Insoluble		Insoluble	
	−268.0*		−269.8*		−273.3*					
Z—γ—Abu—Gly—*Gly—OEt	−290.8				−293.0				−293.7	
	−264.0	+4.2	—	—	−267.7	+5.0	Insoluble		−268.3	+5.1
	−268.2*				−272.7*				−273.4*	

[a] The order of chemical shifts matches that of the monomer units in the formulae.
[b] $\delta_{Y-Y} - \delta_{X-Y}$ (Y = Gly; X = variable amino acid).
[c] Most urethane signals are broad due to proton exchange (see Fig. 1).
[d] The Pro signal was not detected; all other signals represent the *trans* isomer.

ception, because the signs of the NREs are opposite in pyridine (Table 2). This demonstrates, along with other observations discussed below, that Gly/Ala sequences behave abnormally with respect to NREs. Furthermore, the spectroscopic behaviour of the series Z—Gly—Phe—*Phe—OR where R = H, Me, Gly—OEt and Gly—*Gly—OEt is noteworthy (Table 4). When this series is measured in formic acid the Gly—Phe signal is not affected by the structural variation at the carboxyl end of the Z-tripeptide; yet, the signal of the *Phe nitrogen shifts downfield, so that the negative NRE increases. Because the basicity of the *Phe carbonyl group increases in the order R = H, Me, Gly—OEt, Gly—*Gly—OEt, the observed shift effect probably reflects an interaction of this CO group with formic acid. This conclusion is supported by the obser-

Table 2. ^{15}N NMR chemical shifts δ[a] (ppm, relative to external $NO_3^{\ominus}$) and neighbouring residue effects Δδ (ppm)[b] of Ala—Ala containing Z-tripeptide esters

Peptide derivative	Formic acid δ	Formic acid Δδ	2-Chloroethanol δ	2-Chloroethanol Δδ	DMSO δ	DMSO Δδ	Pyridine δ	Pyridine Δδ
Z—Gly—Ala—*Ala—OH	−299.3		−300.1		−299.5		−300.4	
	−253.5	±0	−255.0	−0.8	−257.3	−0.4	−257.2	−0.3
	−253.5*		−255.8*		−256.9*		−256.9*	
Z—Gly—Ala—*Ala—OMe	−299.3		−300.0		−299.6		−300.5	
	−253.5	±0	−255.7	±0	−257.7	±0	−257.4	+0.8
	−253.5*		−255.7*		−257.7*		−258.2*	
Z—β—Ala—Ala—*Ala—OMe	−292.1		−293.3		−294.0			
	−247.2	+6.4	−247.7	+8.2	−251.6	+6.4	Insoluble	
	−253.6*		−255.9*		−258.0*			
Z—Val—Ala—*Ala—OMe	−290.7				−291.1		−290.9	
	−250.5	+2.7	Insoluble		−254.8	+3.2	−253.1	+5.1
	−253.2*				−258.0*		−258.2*	
Z—Ile—Ala—*Ala—OMe	−289.7				−289.9			
	−249.9	+3.3	Insoluble		−254.3	+3.7	Insoluble	
	−253.2*				−258.0*			
Z—Leu—Ala—*Ala—OMe	−287.0		−287.2		−287.6		−297.6	
	−252.8	+0.6	−254.5	+1.4	−257.6	+0.5	−255.9	+2.6
	−253.4*		−255.9*		−258.1*		−258.5*	
Z—Met—Ala—*Ala—OMe					−290.5			
	—	—	—	—	−256.2	+1.9	—	—
					−258.1*			
Z—Phe—Ala—*Ala—OH	−288.8		−288.5		−289.0			
	−251.6	+1.7	−253.5	+1.6	−257.3	±0	Insoluble	
	−253.3*		−255.1*		−257.3*			
Z—Phe—Ala—*Ala—OMe	−288.7				−288.7		−289.0	
	−251.8	+1.5	Insoluble		−257.2	+0.8	−255.4	+2.9
	−253.3*				−258.0*		−258.3*	
Z—Gly—Pro—Ala—*Ala—OMe	−299.5[d]		−300.1[d]		−300.2[d]		−300.9	
	−253.5	±0	−255.9	±0	−258.6	−0.5	−247.7	
	−253.5*		−255.9*		−258.1*		−259.1	−1.1
							258.0*	

[a] The order of chemical shifts matches that of the monomer units in the formulae.
[b] $\delta_{Y-Y} - \delta_{X-Y}$ (Y = Ala; X = variable amino acid).
[c] Most urethane signals are broad due to proton exchange (see Fig. 1).
[d] The Pro signal was not detected; all other signals represent the *trans* isomer.

vation that no such regular shift effects exist in the less acidic solvents.

The comparison of tripeptide methyl and ethyl ester trifluoroacetates (Table 5) with the corresponding Z-protected tripeptide esters shows, on the one hand, that the NREs have identical signs, except in the case of Ala—Gly—*Gly—OEt in 2-chloroethanol. On the other hand, the values of the NREs differ by up to 3.8 ppm, and these large deviations apply to NREs with both positive and negative signs. Thus, oligopeptides or their esters with ammonium end groups are obviously poor spectroscopic models for peptide or protein sequences lacking a positively charged neighbourhood. However, oligopeptide and peptide ester salts have the advantage of allowing measurements in water, which is the only biologically relevant solvent, and they are also soluble and stable in TFA, the best solvent for synthetic polypeptides.

Solvent effects

Previous studies of amides, oligopeptides and polypeptides[18] have shown that the ^{15}N NMR chemical shifts of peptide groups are highly sensitive to the nature of the solvent. Hence, it was expected that the NREs also depend on solvation. The solvents used for the measurements of Tables 1–4 were selected for the following three reasons: first, structure and acidity should vary as largely as possible (the order of the solvents in Tables 1–5 matches the order of acidity); second, NH groups should not be present, and, third, at least 0.5 M solutions of as many oligopeptide derivatives as possible should be accessible, because natural abundance ^{15}N NMR measurements with a 2.1 T magnet require such high concentrations. 2-Chloroethanol was chosen as a non-acidic protic solvent which is, on the one hand, a much better solvent for peptide derivatives than methanol and, on the other hand, much cheaper than trifluoroethanol. In the case of serine and threonine derivatives the influence of all three alcohols on the chemical shifts of the peptide signals could be compared (Table 12). The chemical shifts found in methanol and trifluoroethanol (TFE) agree better with each other than with those of 2-chloroethanol (CE) solutions. However, the chemical shifts found in 2-chloroethanol are further downfield and, thus, agree better with those in water solution, as demonstrated by the measurements of Table 5. The data of Table 5 show, furthermore, that the NREs obtained from 2-chloroethanol solutions agree fairly well with those found in water, whereas the chemical shifts in water

Table 3. ^{15}N NMR chemical shifts δ[a] (ppm, upfield of external $NO_3^{\ominus}$) and neighbouring residue effects δΔ (ppm)[b] of protected tripeptide esters with Val—Val; Leu—Leu and Met—Met units

Z—peptide ester	Formic acid[c] δ	Δδ	Chloroethanol δ	Δδ	DMSO δ	Δδ	CH_2Cl_2 δ	Δδ	Pyridine δ	Δδ
Z—Gly—Val—*Val—OMe	−300.5		−300.2		−299.5		−302.0		−300.7	
	−257.8	−2.9	−259.8	−2.6	−263.6	−3.9	−260.4	−1.3	−261.6	−2.5
	−254.9*		−257.2*		−259.7*		−259.1*		−259.1*	
Z—β—Ala—Val—*Val—OMe	−292.4		−293.7		−293.8		−295.3		−294.7	
	−251.6	+3.9	−252.2	+5.4	−256.8	+3.1	−255.7	+4.0	−255.3	+3.7
	−255.5*		−257.6*		−259.9*		−259.7*		−259.0*	
Z—Ala—Val—*Val—OMe	−284.4				−285.0				−285.7	
	−260.3	−5.2	—	—	−264.5	−4.1	—	—	−259.7	−0.6
	−255.1*				−260.4*				−259.1*	
Z—Phe—Val—*Val—OMe	−287.5		−288.7		−288.7				−289.0	
	−254.6	+0.5	−256.8	±0.4	−259.9	±0	Insoluble		−258.8	±0
	−255.1*		−257.2*		−259.9*				−258.8*	
Z—Met—Val—*Val—OMe	−288.3				−288.6		−289.4		−289.2	
	−256.5	−1.5			−263.1	−3.0	−259.3	−0.6	−259.7	−0.7
	−255.0*				−260.1*		−259.3*		−259.0*	
Z—Gly—Pro—Val—*Val—OMe	−299.2[d]		−300.7[d]		−300.3[d]		−302.9[d]		−301.7[d]	
	−256.5	−1.5	−258.9	−1.3	−262.5	−2.0	−260.2	+1.1	−260.2	±0
	−255.0*		−257.6*		−260.5*		−261.3*		−260.2*	
Z—Gly—Leu—*Leu—OMe	−299.8		−300.0		−299.5		−301.7		−300.7	
	−255.5	−1.6	−256.7	−0.8	−259.5	−1.1	−257.8	−0.5	−258.9	−0.5
	−253.9*		−255.9*		−258.9*		−257.3*		−258.4*	
Z—β—Ala—Leu—*Leu—OMe	−292.3		−293.7		−293.8		−295.6		−294.7	
	−249.0	+5.1	−249.2	+7.0	−253.6	+5.5	−252.9	+5.2	−253.2	+5.0
	−254.1*		−256.2*		−259.1*		−258.1*		−258.2*	
Z—Ala—Leu—*Leu—OMe	−284.1		−285.3		−285.2		−286.6		−285.8	
	−256.0	−2.5	−257.4	−1.5	−260.0	−1.1	−256.8	±0	−258.6	−0.4
	−253.5*		−255.9*		−258.9*		−256.8*		−258.2*	
Z—Phe—Leu—*Leu—OMe	−288.0				−288.7		−299.0		−289.1	
	−253.5	±0	—		−258.9	+1.6	−255.5	+1.6	−256.6	+1.5
	−253.5*				−258.9*		−257.1		−258.1*	
Z—Gly—Pro—Leu—*Leu—OMe	−300.0[d]				−300.3		−301.7[d]			
	−255.0				−259.9		−256.7			
	−253.9*	−1.1	—		−259.6*	−0.3	−257.3*	+0.6	—	
Z—Gly—Met—*Met—OMe	−299.9				−299.5				−301.0	
	−257.1	−1.7	—		−261.1	−1.2	Insoluble		−260.4	−0.9
	−255.4*				−259.9*				−259.5*	

[a] The order of chemical shifts matches that of the monomer units in the formulae.
[b] $\delta_{Y-Y} - \delta_{X-Y}$ (Y = Val, Leu, Met; X = variable amino acid).
[c] Most urethane signals are broad due to proton exchange.
[d] The Pro signal was not detected; all other signals represent the *trans* isomer.

solutions are best approximated by measurements in formic acid. Hence, we conclude that in the case of the compounds in Tables 1–4, 2-chloroethanol is the best substitute for water with respect to the NREs while formic acid is the better substitute with respect to the chemical shifts.

Formic acid was used instead of TFA, which is the better solvent for polypeptides and peptide ester salts, because the Z group is not stable in TFA. The Z group was preferred to more stable protecting groups, such as tosyl, phthaloyl or trifluoroacetyl, because its electron withdrawing (–I) effect resembles more that of an aminoacyl unit. Furthermore, the Z group is, in addition to the *tert*-butoxycarbonyl (BOC) group, the most widely used protecting group in peptide chemistry. The spectroscopic effect of the Z group is almost identical with that of the BOC group,[23] yet the BOC group is unstable in formic acid.

When the data of Tables 1–4 are compared with each other with respect to solvent effects, the following observations are noteworthy. All peptide signals shift several ppm downfield when the aprotic solvents are replaced by the more acidic solvents 2-chloroethanol and formic acid. Thus, the largest downfield shifts listed in Tables 1–4 are found in formic acid, those of Table 5 in TFA. However, the data of Table 5 show that the downfield shifts do not exactly parallel the solvent acidity, because the peptide signals in water are downfield of those obtained in formic acid. No systematic behaviour is, furthermore, detectable for the chemical shifts in aprotic solvents. Hence, we conclude that the solvent effects have at least two origins: first, direct interaction of solvent and amide group, i.e. dipolar forces, H bonds and protonation; and, second, conformational changes. In protic solvents the direct interactions are predominant and, thus, increasing acidity leads in most cases to downfield shifts, while the solvent effects of aprotic solvents are more complex and unpredictable, as already pointed out in a previous paper.[23]

In the following discussion we aim to distinguish between solvent effects on chemical shifts and those

Table 4. ^{15}N NMR chemical shifts δ[a] (ppm, relative to external $NO_3^{\ominus}$) and neighbouring residue effects Δδ (ppm)[b] of Phe—Phe-containing Z-protected oligopeptides

Z—Oligopeptide	Formic acid δ	Δδ	2-Chloroethanol δ	Δδ	DMSO δ	Δδ	CH_2Cl_2 δ	Δδ	Pyridine δ	Δδ
Z—Gly—Phe—*Phe—OH	−299.8		−300.2		−299.9				−300.8	
	−257.1	−1.3	−259.1	−1.8	−261.2	−1.7	Insoluble		−260.6	−1.8
	−255.8*		−257.3*		−259.5*				−258.8*	
Z—Gly—Phe—*Phe—OMe	−300.0				−299.8				−300.8	
	−257.0	−1.6	—	—	−261.2	−1.5	Insoluble		−260.2	−1.2
	−255.4*				−259.7*				−259.0*	
Z—Gly—Phe—*Phe—Gly—OEt	−299.8		−300.1		−299.8				−300.7	
	−257.1	−1.9	−258.2	−0.9	−260.8	−1.0	Insoluble		−259.8	−1.9
	−255.2*		−257.3		−259.8*				−257.9*	
	−266.6		−269.6		−271.6				−271.7	
Z—Gly—Phe—*Phe—Gly—*Gly—OEt	−299.4		−300.0		−299.8		−303.8		−300.4	
	−257.0	−2.1	−258.5	−1.2	−260.7	−1.7	−260.3	−3.6	−259.6	−2.5
	−254.9*		−257.3*		−259.0*		−256.7*		−257.1*	
	−266.7	+1.4	−268.9	+2.1	−270.7	+2.1	−272.2	+2.1	−270.1	+3.4
	−268.1*		−270.5*		−272.8*		−274.3*		−273.5*	
Z—Ala—Phe—*Phe—OMe	−284.5				−285.4				−285.8	
	−258.0	−2.6	Insoluble		−262.5	−2.6	Insoluble		−260.9	−1.5
	−255.4*				−259.9*				−259.4*	
Z—Ala—Phe—*Phe—Ala—OMe	−284.2									
	−258.1	−3.3	Insoluble		Insoluble		Insoluble		Insoluble	
	−254.8*									
	−251.0									
Z—β—Ala—Phe—*Phe—OH	−292.2		−293.2		−293.9				−294.6	
	−250.8	+5.0	−251.9	+5.3	−255.4	+4.9	Insoluble		−254.9	+4.1
	−255.8*		−257.2*		−260.3*				−259.0*	
Z—β—Ala—Phe—*Phe—OMe	−292.2				−293.9				−294.6	
	−250.8	+4.6	Insoluble		−255.0	+4.9	Insoluble		−254.4	+4.4
	−255.4*				−259.9*				−258.8*	
Z—Gly—Pro—Phe—*Phe—OMe					−300.1[d]				−302.3[d]	
	—	—	—	—	−262.0	−2.1	—	—	−260.0	+1.1
					−259.9*				−261.1*	

[a] The order of chemical shifts matches that of the monomer units in the formulae.
[b] $\delta_{Y-Y} - \delta_{X-Y}$ (Y = Phe; X = variable amino acid).
[c] Most urethane signals are broad due to proton exchange.
[d] The Pro signal was not detected; all other signals represent the *trans* isomer.

on NREs. Since the NREs are, according to our definition, the difference between two chemical shifts (X—Y and Y—Y), a systematic shift of one or both peptide signals should also lead to a systematic change of the NREs. Yet, the data of Tables 1–5 show that even in the case of non-systematic shifts, regular trends of the NREs are observable. For example, the NREs of the Gly—Gly derivatives (Table 1) increase with decreasing acidity of the solvent, the only exception being Z—β—Ala—Gly—*Gly—OEt, because all Z—β—Ala—Y—Y—OR peptides show the maximum NRE in 2-chloroethanol. In addition, the NREs shown in Table 5 increase with decreasing acidity of the solvents, except those of Gly/Ala sequences. Yet, in several cases of fully protected Val—Val and Leu—Leu peptides (Table 3) the largest NREs are observed in formic acid. Thus, we find that, first, solvents have a strong influence on the NREs and chemical shifts of peptides; second, the acidity is the predominant property of the solvents with respect to solvent effects; and, third, the manner in which the NREs vary with increasing acidity of the solvents depends on the nature of the peptide. Hence, the strong solvent effects have a negative and a positive aspect. On the one hand, NREs and chemical shifts are not easily predictable, and assignments based on comparisons with model compounds require measurements in the same solvent. On the other hand, changing the solvent is a simple measure to avoid overlapping of signals; this positive aspect is noteworthy in as much as the common shift reagents (lanthanide complexes) are rather useless in ^{15}N NMR spectroscopy of peptides, proteins and polyamides.[24]

Neighbouring residue effects in copolypeptides

In previous papers[25,26] we have shown that ^{15}N NMR spectroscopy is currently the only method which allows sequence analyses of most synthetic copolypeptides. ^{15}N NMR sequence analyses of copolypeptides are worthwhile for two reasons. The peptide sequences contain information on the kinetic course of copolymerization and, thus, on the hitherto not yet fully elucidated polymerization mechanisms. Copolypeptides are, furthermore, interesting models for proteins, yet a meaningful investigation of physical or biological properties requires that the sequences are known. A copolypeptide of two amino acids (binary system) may contain four different peptide groups (X—X, X—Y; Y—X and Y—Y) and a ternary copolypeptide (terpolypeptide) nine different peptide

Table 5. ^{15}N NMR chemical shifts δ[a] (ppm, relative to external $NO_3^{\ominus}$) and neighbouring residue effects Δδ (ppm) of tripeptide ester trifluoroacetates

Trifluoroacetate of	TFA δ	TFA Δδ	Formic acid δ	Formic acid Δδ	Water δ	Water Δδ	2-Chloroethanol δ	2-Chloroethanol Δδ	DMSO δ	DMSO Δδ
H—Gly—Ala—*Ala—OH	−350.1	[b]	−349.7	[b]						
	−251.7	+1.0	−253.4	±0	—		—		—	
	−252.7*		−253.4*							
H—Gly—Ala—*Ala—OMe	−350.8		−351.0		−349.6				−349.1	
	−251.9	+0.4	−253.5	±0	−252.4	+0.4	Insoluble		−257.4	±0
	−252.3*		−253.5*		−252.8*				−257.4*	
H—Ala—Gly—*Gly—OEt	−336.8		−336.4		−335.8		−335.6		−335.6	
	−268.3	−1.4	−269.2	−1.2	−268.1	−0.5	−269.6	+0.4	−273.0	−0.7
	−266.9*		−268.0*		−267.6*		−270.0*		−272.3*	
H—Phe—Gly—*Gly—OEt	−338.8		−338.8		−338.0		−337.5		−337.6	
	−264.9	+2.1	−266.0	+2.0	−263.5	+4.5	−265.8	+4.1	−267.0	+5.0
	−267.0*		−268.0*		−268.0*		−269.9*		−272.0*	
H—Val—Gly—*Gly—OEt	−342.0		−341.7		−340.9		−340.5		−340.3	
	−265.2	+2.7	−265.2	+2.7	−263.5	+4.2	−265.0	+4.9	−265.7	+6.5
	−267.9*		−267.9*		−267.7*		−269.9*		−272.2*	
H—Gly—Val—*Val—OMe	−350.7		−350.1		−350.1		−349.6		−349.1	
	−255.8	−2.2	−258.2	−3.0	−257.3	−3.5	−261.2	−3.6	−265.8	−5.7
	−253.6*		−255.2*		−253.8*		−257.4*		−260.1*	
H—Ala—Val—*Val—OMe	−336.8		−350.1		−350.1		−349.6		−349.1	
	−255.6	−2.1	−258.2	−3.0	−257.3	−3.5	−261.2	−3.6	−265.8	−5.7
	−253.5*		−255.2*		−253.8*		−257.4*		−260.1*	
H—Val—Ala—*Ala—OMe	−343.2		−342.5		−341.4		−341.4		−340.5	
	−248.5	+3.3	−249.7	+3.8	−248.5	+4.2	−250.5	+5.2	−250.8	+7.0
	−251.8*		−253.1*		−252.7*		−255.7*		−257.8*	

[a] The order of chemical shifts matches that of the monomer units in the formulae.
[b] The free tripeptide was dissolved in TFA.

groups.[27] A successful sequence analysis requires on the one hand, that the NREs are large enough (≥0.2 ppm) to allow good resolution of all different peptide groups. In this connection it should be mentioned that the signals of copolypeptides possess a greater line width than homopolypeptides or oligopeptides because of long-range neighbouring residue effects.[27] On the other hand, unambiguous assignments of all the signals are necessary and, thus, the question arises as to what extent oligopeptides can serve as models for copolypeptides.

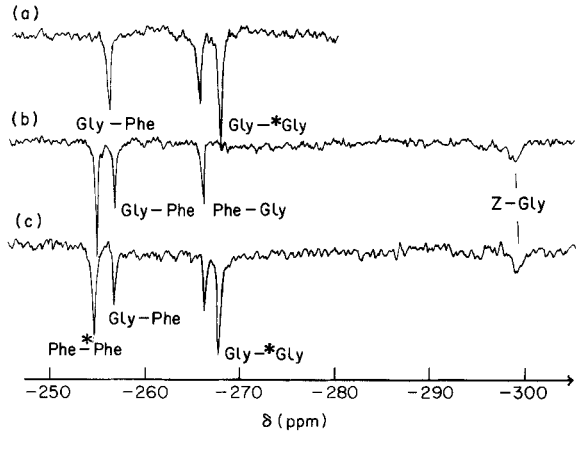


Figure 1. Natural abundance 9.12 MHz ^{15}N NMR spectra of (a) (Phe—Gly—*Gly)$_n$; (b) Z—Gly—Phe—*Phe—Gly—OEt; and (c) of Z—Gly—Phe—*Phe—Gly—*Gly—OEt in formic acid. Pulse width: 30 μs (*c.* 45°); 1 K data points for a spectral width of 600 Hz; exponential line broadening 1.5 Hz; 10 000–50 000 transients.

The ^{15}N NMR spectra of glycine containing copolypeptides are summarized in Table 6 together with the NREs, while the data of several alanine containing co- and ter-polypeptides are listed in Table 7. Both series of copolypeptides, that of glycine (X = Gly, Y = variable amino acid) and that of alanine (X = Ala), show good qualitative agreement with the spectroscopic behaviour of the corresponding oligopeptides. All amino acids shift the Gly signal (Y—X bond) downfield relative to that of the Gly—Gly (X—X) bond, i.e. the NREs are positive, whereas glycine causes upfield shifts in the signals of other amino acids (negative NREs). Again Gly/Ala copolypeptides behave differently, as found for the Gly/Ala oligopeptides in Tables 1–5. Additionally, in the case of alanine containing copolypeptides alanine shifts the signals of other amino acids upfield (negative NREs of X—Y bonds) while the Ala signals are shifted downfield by all other amino acids (positive NREs of Y—X bonds). Furthermore, in both series (Gly and Ala containing copolypeptides) the largest NREs (*c.* 4–6 ppm) are found for β-Ala or Val containing sequences. However, the quantitative agreement of the NREs of oligopeptides and copolypeptides can be poor, even if determined in the same solvent. For example, the pair H—Ala—Val—*Val—OMe/H—Val—Ala—*Ala—OMe gives NREs of −2.1 and +3.3 in TFA (Table 5) whereas the copolypeptide (Ala/Val)$_n$ shows NREs of −5.9 and +4.3 (Table 7). Furthermore, the NRE of HPhe—Gly—*Gly—OEt in TFA is +2.1, but is +3.1 in the case of (Phe—Gly—Gly)$_n$, and the NRE of H—Ala—Gly—*Gly—OEt in TFA is −1.4 while it is −0.4 for (Gly/Ala)$_n$. Better

Table 6. ^{15}N NMR chemical shifts δ[a] (ppm, relative to external $NO_3^{\ominus}$) and neighbouring residue effects Δδ (ppm) of glycine-containing sequence—peptides and copolypeptides

Polypeptides	Solvent / Bond	Trifluoroacetic acid A—A	B—A	B—B	A—B	Formic acid A—A	B—A	B—B	A—B
(Gly—Gly—Ala)$_n$	δ	−267.6	−268.0	—	−251.4	−268.3	−269.0	—	−253.0
	Δδ	−0.4		—		−0.7		—	
(Gly—Ala—Ala)$_n$	δ	—	−268.1	−252.2	−251.4	—	−268.9	−253.6	−253.0
	Δδ	—		+0.8		—		+0.6	
(Gly/Val)$_n$	δ	−267.4	−264.1	−250.0	−255.5	Insoluble		Insoluble	
	Δδ	+3.3		−5.5					
(Gly—Gly—Val)$_n$	δ	−267.1	−263.9	—	−255.5	−268.1	−265.2	—	−258.3
	Δδ	+3.2		—		+2.9		—	
(Gly/Leu)$_n$	δ	−267.1	−266.6	−250.7	−253.2	Insoluble		Insoluble	
	Δδ	+0.5		−2.5					
(Gly—Gly—Pro)$_n$	δ	−267.1	−266.2	—	−237.9	—		—	
	Δδ	+0.9		—					
(Gly/Z$^{\varepsilon}$—Lys)$_n$	δ	Not stable		Not stable		−268.3	−267.3	−253.8	−255.8
	Δδ					+1.0		−2.0	
(Gly/D,L—α-Abu)$_n$	δ	−267.3	−266.1	−251.6	−253.8	Insoluble		Insoluble	
	Δδ	+1.2		−2.2					
(Gly—/γ-Me—Glu)$_n$	δ	−267.3	−266.1	−254.1	−255.8	−268.3	−267.3	−257.1	−255.8
	Δδ	+1.2		−1.7		+1.0		−1.3	
(Gly/Met)$_n$	δ	−267.2	−265.7	−253.2	−255.1	Insoluble		Insoluble	
	Δδ	+1.5		−1.8					
(Gly/Bzl—Cys)$_n$	δ	−267.2	−264.7	−2[illegible]	−255.4	Insoluble		Insoluble	
	Δδ	+2.5		3					
(Gly—Gly—Phe)$_n$	δ	−267.6	−254.5	—	−254.5	−268.4	−266.5	—	−256.7
	Δδ	+3.1				+1.9			
(Gly/Phg)$_n$	δ	−267.4	−265.8	(−250.9)	(−253.2)	Insoluble		Insoluble	
	Δδ	+1.6		(−2.3)[b]					
(Gly—Gly—β-Ala)$_n$	δ	−268.1	−261.6	—	−255.0	−268.2	−263.4	—	−258.7
	Δδ	+6.5		—		+4.8			
(Gly—β—Ala—β-Ala)$_n$	δ	—	−262.5	−245.3	−255.6	—	—	—	—
	Δδ			−10.3		—		—	
(Gly—Gly—γ-Abu)$_n$	δ	−267.0	−261.1	—	−252.6	−267.8	−263.7	—	−256.5
	Δδ	+5.9				+4.1[c]			[c]

[a] Measured at 9.12 MHz.
[b] Broad signals.
[c] Interpolated from the data in Ref. 18.

agreement is found when Z—Val—Gly—*Gly—OEt, Z—β—Ala—Gly—*Gly—OEt and Z—γ—Abu—Gly—*Gly—OEt are compared with the corresponding copolypeptides in formic acid. Probably, fully protected oligopeptides are the better models, if the same solvent can be used for the comparison. We have already compared free tripeptides with copolypeptides in two previous papers,[15,28] and the conclusion drawn from those comparisons is now confirmed by measurements in this work. Oligopeptides and their derivatives

Table 7. ^{15}N NMR chemical shifts δ[a] (ppm, relative to external $NO_3^{\ominus}$) and neighbouring residue effects Δδ[b] (ppm) of alanine containing copolypeptides in TFA

Copolypeptide	A—A	B—A	C—A	B—B	Bond[c] A—B	C—B	C—C	A—C	B—C
(Ala/Gly)$_n$	−250.6	−250.2	—	−266.1	−266.8				
		+0.4	—		−0.7				
(Ala/Val)$_n$	−250.7	−246.4	—	−248.9	−254.8				
		+4.3	—		−5.9				
(Ala/Leu)$_n$	−250.7	−248.5	—	−249.6	251.5				
		+2.5	—		−1.9				
(Ala/Phe)$_n$	−250.5	−247.6	—	−250.5	−253.9				
		+3.0	—		−3.5				
(Ala/Gly/Val)$_n$	−250.5	−250.0	−246.3	−265.9	−266.8	−262.8	−248.8	−254.3	−254.6
					−0.8	+3.1		−5.5	−5.8
(Ala/Gly/Phe)$_n$	−250.6	−250.1	−247.2	−265.9	−266.7	−263.6	−250.1	−253.3	−253.7
		+0.5	+3.4		−0.8	+2.3		−3.2	−3.6

[a] Measured at 40.55 MHz; these chemical shifts are 1.1±0.1 ppm downfield of those measured at 9.12 MHz with an iron magnet[22] (Tables 1–6 and 8–14).
[b] $\delta_{A-A} - \delta_{B-A}$ etc.
[c] A, B, C represent the aminoacyl units according to the sequence given in the first column.

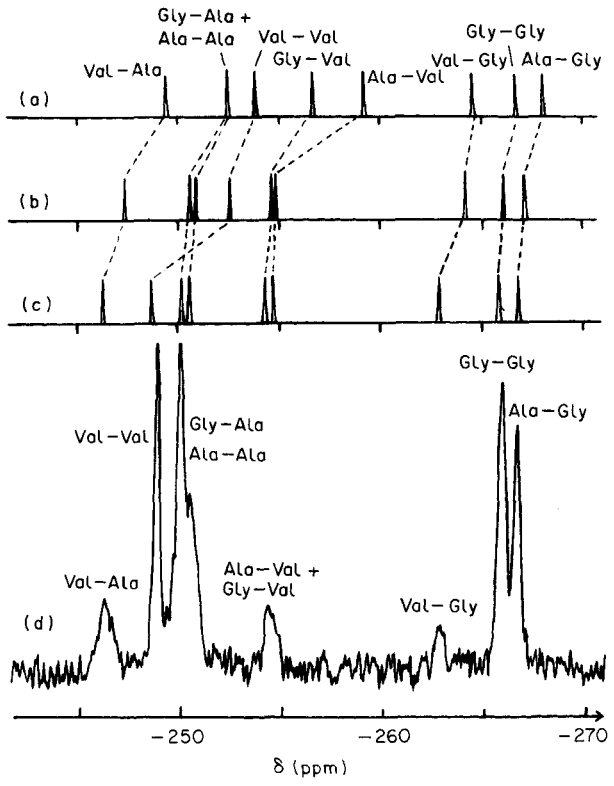


Figure 2. Natural abundance 40.55 MHz ^{15}N NMR spectrum of the terpolypeptide $(Gly/Ala/Val)_n$ in trifluoroacetic acid (d). Signal pattern predicted from Z-tripeptide methyl esters in formic acid (Tables 1–4) (a); signal pattern predicted from tripeptide methyl esters in TFA (Table 5) (b); signal pattern predicted from polypeptides in TFA (Tables 6, 7) (c).

allow us to predict reliably the sign (or sense) of the NREs of polypeptides (and proteins), yet the chemical shifts and NREs cannot be predicted quantitatively.

To test the predictive power of oligopeptides in a concrete case, we have synthesized and measured two terpolypeptides. One terpolypeptide was prepared by benzylamine initiated terpolymerization of glycine-*N*-carboxyanhydride (Gly—NCA), Ala—NCA and Val—NCA, and the other was obtained from Gly—NCA, Ala—NCA and Phe—NCA. Although 40.5 MHz spectra of these terpolypeptides were measured (Table 7), only eight of the nine possible signals were resolved in the case of $(Gly/Ala/Val)_n$ and seven signals for $(Gly/Ala/Phe)_n$. Figure 2(d) displays the spectrum of $(Gly/Ala/Val)_n$; Fig. 2(a) shows the predicted signal pattern based on the Z-oligopeptide esters measured in formic acid (Tables 1–3); Fig. 2(b) shows the pattern predicted from the tripeptide esters in TFA (Table 5) and Fig. 2(c) is the signal pattern predicted from homo- and co-polypeptides (Tables 6 and 7). This comparison allows the following conclusions: first, the signal pattern of the X—Gly bonds is well predictable from all three model systems while the Val—Val signal cannot be predicted from oligopeptides; second, the predictive power increases from Fig. 2(a) to 2(c); and, third, only the polypeptides allow the precise prediction of the chemical shifts of the terpolypeptide. The fact that the predictive powers of the Z-oligopeptide esters is worse than that of the tripeptide esters emphasizes the above conclusion that model systems must be measured in the same solvent to ensure acceptable assignments. The finding that Val—Val signals of polypeptides are several ppm downfield from those of oligopeptides agrees with our previous observation that the substituent effect of a Val unit in polyvaline is much larger than that of valine and low molecular weight derivatives of valine (compare Table 3 in Ref. 20 and Table 11 in this work). The chemical shift of $(Ile)_n$,[29] the substituent effects of isoleucine derivatives (Table 11) and the NREs of Z–Ile peptides (Tables 1 and 2) suggest that isoleucine behaves like valine. The final conclusion drawn from these measurements is that the crowding of amino acyl units with bulky side chains in a peptide causes unusual downfield shifts. Hence, the chemical shifts of Val and Ile containing sequences are most difficult to predict from low molecular weight model compounds, and one may speculate that other amino acids with bulky side chains, such as tryptophane, can also cause unpredictable shift effects when neighboured with valine or isoleucine. Finally, it should be mentioned that the reasons for the different intensities and line widths of the signals of glycine containing terpolypeptides were discussed in a previous paper.[27]

Substituent effects

The above discussed neighbouring residue effects have shown that the ^{15}N NMR chemical shifts of amino acid derivatives depend on the nature of the N-bound residue. Hence, the calculation and assignment of ^{15}N NMR chemical shifts of peptides require the knowledge of the neighbouring residue and solvent effects. Instead of this complication, such calculations and assignments would be relatively easy if the substituent effects were additive, i.e. if the spectroscopic effect of a substituent attached to the α-carbon of a glycine derivative could be added to the chemical shift of the glycine derivative, regardless of the nature of the derivative and the solvent. Such an additivity would allow one to calculate all possible peptide shifts from a standard set of substituent effects, and the chemical shifts of numerous glycine derivatives measured in various solvents.

In a previous paper we have already compared the chemical shifts of several amino acids in various acids.[19] A slight solvent dependence of the substituent effects was found; however, the strong acids used for this investigation[19] might behave irregularly and are not useful for the investigation of proteins or naturally occurring peptides. On comparison of several amino acids, cyclodipeptides and polypeptides in TFA we have, furthermore, observed that the substituent effects depend strongly on the nature of the amino acid derivatives. However, this investigation[20] also represents an extreme case, because only one strong acid was used as solvent and because amino acid derivatives with extremely differing structures were compared. Thus, in this work we have enlarged our study on substituent effects using water and/or DMSO as solvents and using amino acid methyl ester hydrochlorides (Table 8), Z-protected amino acids (Table 9) and *N*-acetyl-amino acid methyl esters (Table 10) as substrates. Substituent effects of amino acid methyl

Table 8. ^{15}N NMR chemical shifts δ (ppm, relative to external $NO_3^{\ominus}$) of amino acid methyl ester hydrochlorides

Methyl ester hydrochloride	In H_2O(pH 3–4)	In DMSO
Gly	−347.9 (347.7)[a]	−343.4 (343.6)[a]
Ala	−334.8	−331.1
β-Ala	−343.8	−339.9
α-Aibu	−323.5	−320.3
γ-Abu	−342.3	−338.8
Val	−340.5	−334.8
Leu	−336.9	−332.4
Ile	−339.3	−334.7
Phg	−332.6	−329.6
Phe	−337.9	−332.8
Tyr	−338.0	−333.9
Ser	−341.7	−338.0
Thr	—	−340.3
Met	−337.3	−333.3
Cys	−338.2	−334.7
Cys_2	−338.4	−333.6
γ-OMe—Glu	−336.6	−333.4
β-OMe—Asp	−337.7	−333.7
Arg HCl	−336.8 −303.7 −291.6	Insoluble
Lys(N^{ε}·HCl)	−337.0 −341.9	−332.6 −337.8
His(N^{im}·HCl)	−336.6 −200.2 −202.2	Insoluble
Trp	−338.0 −245.8	−333.4 −241.9
Pro	−322.3	−318.5
Hypro	−326.0	−322.7
Sar	−346.0	−343.2
4-Abe	−324.3	−316.6

[a] Ethyl ester.

ester hydrochlorides were first calculated by Randall *et al.* for highly concentrated (5–9 M) solutions at pH 0.5–2.0.[30] A comparison with amino acid solutions in aqueous HCl showed good agreement,[5,7] suggesting that substituent effects are useful parameters for the prediction of chemical shifts. Our measurements in water and DMSO (Table 8) demonstrate that the substituent effects depend on the solvation (up to *c.* 1.4 ppm) even in non-acidic solvents (Table 11).

Table 10. ^{15}N NMR chemical shifts δ (ppm, relative to external $NO_3^{\ominus}$) of N-acetyl-amino acid methyl esters in dimethyl sulphoxide

N-Acetyl derivative of	δ	N-Acetyl derivative of	δ (other nitrogens)
Gly—OMe	−267.3	Tyr—OMe	−254.6
Ala—OMe	−251.8	Met—OMe	−255.9
β-Ala—OMe	−259.6	γ-OMe—Glu—OMe	−255.9
α-Aibu—OMe	−241.3	β-OMe—Asp—OMe	−253.5
γ-Abu—OMe	−258.1	Trp—OMe	−253.8 −245.0
Val—OMe	−257.9	4-Abe—OMe[a]	−240.2
Leu—OMe	−254.3	Sar—OMe	−273.6
Ile—OMe	−256.2		−273.3
Phg—OMe	−252.5		
Phe—OMe	−255.2		

[a] Derivative of 4-aminobenzoic acid.

When various amino acid derivatives are compared in DMSO the substituent effects of individual amino acids show differences up to *c.* 4 ppm (e.g. β-alanine, Table 11). Thus, it is clear that an accurate prediction of ^{15}N NMR chemical shifts by means of substituent effects is not possible. This conclusion marks a characteristic difference between ^{15}N and ^{13}C NMR spectroscopy. It is the lone pair of the nitrogen which renders the ^{15}N NMR chemical shift so sensitive to slight changes in the neighbourhood, regardless of whether molecular structure, conformation or solvation is concerned.

Nonetheless, crude assignments of ^{15}N NMR signals are possible on the basis of substituent effects which differ by more than 3 ppm. For example, in a peptide containing only Gly, Ala, Ser and Pro units all signals can be attributed unambiguously to the individual amino acids, simply by means of substituent effects.

The fact that glycine absorbs upfield of all other amino acids makes it easy to assign Gly signals. However, up to now, threonine and its derivatives have never been measured. The data in Tables 8 and 9 and the substituent effects in Table 11 demonstrate that the chemical shifts of threonine are closest to those of glycine. The hydroxy group and the methyl group in γ-position relative to the nitrogen are responsible for

Table 9. ^{15}N NMR chemical shifts δ (ppm, relative to external $NO_3^{\ominus}$) of *N*-benzyloxycarbonyl α-amino acids in dimethyl sulphoxide

Amino acid derivative	δ	Amino acid derivative	δ	δ (other nitrogens)
Z—Glycine	−299.5	Z—Methionine	−288.9	—
Z—Alanine	−285.1	Z—Aspartic acid	−288.8	—
Z—β-Alanine	−294.2	Z—Asparagine	−288.0	−266.5
Z—α-Aminoisobutyric acid	−275.6	Z—Glutamic Acid	−288.7	—
Z—γ-Aminobutyric acid	−293.3	Z—Glutamine	−288.2	−267.5
		Z—Tryptophane	−287.1	−245.5
Z—Valine	−291.2	Z—Arginine·HCl	−288.0	−290.7 −300.9
Z—Leucine	−287.7	Z—Lysine (N^{ε}·HCl)	−287.8	−339.2
Z—Isoleucine	−290.1	$^{\alpha}$Z, $^{\varepsilon}$BOC—Ornithine	−287.9	−291.4
		Z—Proline[a]	−278.2	—
Z—Phenylglycine	−285.7	Z—4-Hydroxyproline[a]	−281.8	—
Z—Phenylalanine	−288.5	Z—Pyroglutamic acid	−226.4	—
Z—Tyrosine	−288.6	Z—4-Aminobenzoic acid	−268.2	—
Z—Serine	−291.5			
Z—Threonine	−295.7			
Z—Cystine	Not soluble			

[a] Signal of the *cis* isomer.

Table 11. ^{15}N NMR substituent effects $\Delta\delta^a$ (ppm) of various amino acid derivatives

Amino acid	Δδ of methyl ester (HCl) in H_2O	Δδ of methyl ester (HCl) in DMSO	Δδ of Z-derivatives in DMSO	Δδ of *N*-acetyl methyl ester in DMSO
Ala	+13.1	+12.3	+14.4	+15.5
β-Ala	+4.1	+3.5	+5.3	+7.7
α-Aibu	+24.4	+23.1	+23.9	+26.0
γ-Abu	+5.6	+4.6	+6.2	+9.2
Val	+7.4	+8.6	+8.3	+9.4
Leu	+11.0	+11.0	+11.7	+13.0
Ile	+8.6	+8.7	+9.4	+11.1
Phg	+15.3	+13.7	+13.8	+14.8
Phe	+10.0	+10.6	+11.0	+12.1
Tyr	+9.9	+9.5	+10.9	+13.7
Ser	+6.2	+5.4	+8.0	—
Thr	—	+3.1	+3.8	—
Met	+10.6	+10.1	+10.6	+11.4
Cys	+9.7	+8.7	—	—
β-Me Asp	+10.2	+9.7	—	+13.8
Lys	+10.9	+10.8	+11.7	—
Trp	+9.9	+10.0	+12.4	+13.5
Pro	+25.6	+24.9	+21.3 (*cis*)	—
4-Hypro	+21.9	+20.7	+17.7 (*cis*)	—

[a] $\delta_{\text{Gly deriv.}} - \delta_{\text{amino acid deriv.}}$

the absorption at high field. A comparison of the substituent effects of serine and alanine, or of 4-hydroxyproline and proline, shows that the upfield shift effect of a hydroxy group in γ-position is of general validity. We have measured the chemical shifts of several serine- and threonine-containing dipeptides in various solvents to clarify whether the signals of these two amino acids can appear upfield of the glycine signal, depending on the neighbouring residue effects (Table 12). This question is of interest because serine, threonine, glycine, alanine and valine are among the most abundant amino acid units in silk proteins. The data of Table 12 show that with respect to neighbouring residue effects serine and threonine behave similarly to glycine (Table 1): first, the Val—Ser bond absorbs downfield of the Phe—Ser, Ala—Ser and Gly—Ser bonds, just as the Val—Gly signal appears downfield of the Phe—Gly, Ala—Gly and Gly—Gly signals and, second, the Ala—Ser and Ala—Thr bonds absorb slightly upfield of the Gly—Ser and Gly—Thr bonds in most solvents, analogous with the spectroscopic behaviour of Z—Ala—Gly—Gly—OEt in formic acid and in 2-chloroethanol. However, even in the most favourable case, namely upon comparison of Ala—Ser and Val—Gly, the Gly signal absorbs upfield of the Ser signal. In other words, the order of chemical shifts in the Ser/Gly system is dominated by the substituent effect of serine. The opposite is true for threonine, because both Ala—Thr and Gly—Thr bonds absorb upfield of the Val—Gly group. Hence, the high field signals in a silk protein spectrum cannot be exclusively assigned to X—Gly bonds. In this connection it is also noteworthy that sarcosine derivatives can absorb upfield of glycine derivatives (see Table 10), although Sar—OMe·HCl absorbs downfield of Gly—OMe·HCl.

Finally, it should be pointed out that the substituent effects we have measured for amino acid methyl ester hydrochlorides in water are not all identical with those in Refs 5 and 7. Three reasons may account for this: first, the different concentration, second, the different pH values and, third, partial saponification during the measurements. Three measures were taken to ensure that our data are not affected by saponification. First, the methyl ester hydrochlorides were measured immediately after they had dissolved in cold water (in most cases the measurements required only 10–15 min); second, the pH was only in the range of 3–4

Table 12. ^{15}N NMR chemical shifts δ (ppm relative to external $NO_3^{\ominus}$) of serine and threonine containing dipeptide derivatives

Compound	TFE	CE	Methanol	DMSO	Pyridine
Z—Gly—Ser—OMe	−301.2	−300.1	−301.4	−299.9	−300.8
	−265.2	−263.8	−265.2	−265.2	−265.9
Z—Ala—Ser—OMe	−285.8	−285.2	−286.5	−285.4	Insoluble
	−265.3	−264.1	−265.2	−265.5	
Z—Val—Ser—OMe	−291.3	Insoluble	Insoluble	−291.3	Insoluble
	−261.0			−261.4	
Z—Phe—Ser—OMe	−289.1	−288.5	−289.6	−288.9	−289.0
	−262.8	−262.5	−263.2	−264.0	−263.3
Z—Gly—Thr—OMe	−301.2	−300.1	−301.5	−299.8	−300.6
	−268.3	−267.3	−268.7	−269.1	−269.7
Z—Ala—Thr—OMe	−286.0	—	—	−285.4	−285.8
	−268.8			−269.8	−270.0

(0.5–2.0 in Refs 5 and 30), and, third, the stability of the most reactive methyl ester hydrochlorides, e.g. Gly—OMe·HCl and Ser—OMe·HCl, was monitored in D_2O by means of ^1H NMR spectroscopy.

Concentration effects

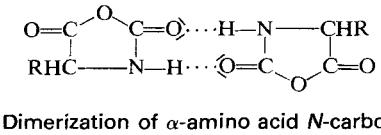
Since the ^{15}N NMR chemical shift of amide groups is highly sensitive to solvent effects, and because amides possessing an N-proton can associate in less polar solvents (e.g. to form dimers), it is not unlikely that concentration and temperature will influence the chemical shift. An investigation of this problem has two interesting aspects; first, the detection or elucidation of association equilibria and, second, the assignment of signals in spectra of ^{15}N-enriched naturally occuring proteins by means of synthetic model compounds. Routine natural abundance ^{15}N NMR measurements of synthetic amino acid derivatives require relatively high concentrations (0.2–2 M) because of the low sensitivity of the ^{15}N nucleus. On the other hand, it is possible to produce *c.* 95% ^{15}N-enriched proteins by feeding microorganisms with salts of $^{15}NH_3$. These ^{15}N-enriched proteins can be measured at much lower concentrations (*c.* 1–10 mM). Thus, the question arises whether the concentration difference affects the usefulness of synthetic model compounds for the assignments of protein signals.

We have measured ^{15}N-enriched polyglycine, poly-L-leucine, *N*-acetylglycine anilide and phenylalanine-*N*-carboxyanhydride in various solvents (Table 13). The latter compound was chosen because it was available with 45% ^{15}N-enrichment from polymerization experiments, and because it is much more soluble in various organic solvents than *N*-acetylglycine anilide. The data of Table 13 do not show any concentration effect for DMSO solutions of *N*-acetylglycyl anilide. Since DMSO is a good H-bond acceptor, and a highly polar solvent, association of amide groups is unlikely in this solvent. Thus, the solvation shell remains unaffected by a change of concentration and the chemical shift likewise. In TFA, decreasing concentration leads to a weak downfield shift. Because the chemical shift in the TFA solutions is 4.5 ppm downfield to that in the DMSO solutions, we believe that *N*-acetylglycine anilide is partially protonated in TFA and that the extent of protonation increases slightly with decreasing concentration. Phe—NCA does not show concentration effects, either in DMSO or in TFA. Because the basicity of the *N*-carboxyanhydride ring is lower than that of amide groups, protonation cannot play a major role for the solvation in TFA. However, the good solvation via H bonds with DMSO as acceptor and TFA as donor prevents association of NCA rings and, thus, concentration effects are not expected. Fitting well into this interpretation is that in the less polar solvents, dioxane and methylene chloride, concentration effects of up to *c.* 3 ppm are observable. Since neither solvent is an H-bond donor or a good acceptor, and because increasing concentration leads to downfield shifts (opposite to the influence of protonation!) we conclude that dimerization takes place according to Scheme 1. The dimerization favours the

Table 13. Influence of concentration and temperature on the ^{15}N NMR chemical shifts δ (ppm, upfield of external $NO_3^{\ominus}$) of various amino acid derivatives

Compound	Solvent	Concentration (M)	Temperature (°C)	δ
L-Phe—NCA (40% ^{15}N)	Methylene chloride	1.00	29–30	−288.66
		0.30	29–30	−289.31
		0.10	29–30	−290.11
		0.03	29–30	−290.54
		0.01	29–30	−290.80
	Dioxane	1.00	29–30	−290.16
		0.30	29–30	−290.54
		0.10	29–30	−290.75
		0.03	29–30	−290.79
		0.01	29–30	−290.84
		1.00	95–96	−290.16
		0.10	95–96	−290.59
	DMSO	1.00	29–30	−285.88
		0.10	29–30	−285.84
		0.01	29–30	−285.84
	TFA	1.00	29–30	−284.72
		0.30	29–30	−284.72
		0.10	29–30	−284.72
		0.01	29–30	−284.72
N-Acetyl-glycine anilide (10% ^{15}N)	DMSO	1.00	129–130	−247.03
		1.00	79–80	−246.98
		1.00	29–30	−246.96
		0.30	29–30	−247.01
		0.10	29–30	−247.01
		0.01	29–30	−247.01
	TFA	1.00	29–30	−243.05
		0.30	29–30	−242.57
		0.10	29–30	−242.35
		0.03	29–30	−242.21
		0.01	29–30	−242.11
$(Gly)_n$ (10% ^{15}N)	TFA	3.00	29–30	−267.34
		1.00	29–30	−267.30
		0.30	29–30	−267.30
		0.10	29–30	−267.25
		0.03	29–30	−267.25
	TFA+5% MSA	3.00	29–30	−264.8
		1.00	29–30	−262.5
		0.10	29–30	−261.4
		0.03	29–30	−261.2
$(Leu)_n$ (10% ^{15}N)	TFA+5% MSA	2.00	29–30	−248.8
		1.00	29–30	−248.2
		0.50	29–30	−247.9
		0.03	29–30	−247.7

delocalization of the nitrogen lone pair, which is responsible for the observed downfield shift. There are two reasons why the association of Phe—NCA involves the C-2 and not the C-5 carboxyl group of the oxazolidine-2,5-dione ring. First, various chemical reactions and IR and ^{13}C NMR spectra have demonstrated that the carbonyl group (C-2) is less electrophilic and more basic than the C-5 carbonyl group. Second, association involving C-5 would lead to a polymerization via H bonds, a process which is entropically less favourable than dimerization. Furthermore,



Scheme 1. Dimerization of α-amino acid *N*-carboxyanhydrides via H bonds in less polar solvents.

we have investigated whether the association of Phe—NCA affects the ^{15}N—^{1}H coupling constants; these results will be published elsewhere.[31]

Interesting spectroscopic behaviour is also observed when polyglycine and poly-L-leucine are measured at various concentrations in acidic solvents. Polyglycine does not show any concentration effect in pure TFA. This finding agrees with our conclusion that polypeptides are not protonated in TFA or, at most, of the order of only 1–3%, from a study of ^{15}N NMR chemical shifts of various polypeptides and polyamides dissolved in solvents of different acidities.[18] In the absence of a protonation/deprotonation equilibrium the solvation of peptide groups does not change upon variation of the concentration, so that the chemical shift must remain constant. However, when methanesulphonic acid (MSA) is added to the TFA solution of polyglycine or poly-L-leucine a concentration effect is clearly detectable (Table 13). Methanesulphonic acid is known to be several orders of magnitude more acidic than TFA.[32] This fact, and the observation that a polyglycine solution in MSA absorbs *c.* 9 ppm downfield of a polyglycine solution in TFA, establishes that polypeptides are to a large extent protonated in pure MSA. Hence, a partial protonation, probably <50%, is expected for peptide solutions in TFA containing 5% (by volume) MSA. A variation of the polyglycine concentration from 3 M to 3×10^{-2} M means that the MSA/NH ratio varies from 1.7 to 1.7×10^{2}. It is obvious that such a strong variation of the acid/base ratio has a considerable influence on the degree of protonation and, thus, on the chemical shifts. However, the comparison of polyglycine and poly-L-leucine demonstrates that the sensitivity of various polypeptides to concentration effects may vary greatly. Finally, we conclude that concentration effects on the ^{15}N NMR chemical shifts of amides and carbamic acid derivatives can reach 4 ppm, but they are only expected if concentration dependent association or protonation equilibria are present.

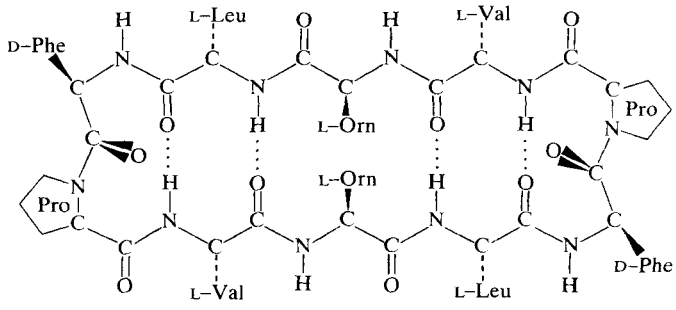


Figure 3. Two-dimensional structure of Gramicidin S with four intramolecular H bonds.[37]

Gramicidin S

To complete this study on structure/shift relationships of peptides we have included a comparison of Gramicidin S and non-cyclic model peptides to investigate the spectroscopic effect of intramolecular H bonds. In a previous paper we have shown that the Pro-nitrogens of the cyclopeptide (Gly^{1}—Pro^{1}—Gly^{2}—Pro^{2}—Ala) absorb *c.* 7–10 ppm downfield of those of non-cyclic model peptides in aprotic solvents.[23] Intramolecular H bonds directed to the carbonyl groups of the two Gly units were postulated to be responsible for these extraordinary downfield shifts. To find more experimental evidence for this conclusion we have now compared Gramicidin S, which possesses four intramolecular H bonds (Fig. 3), with various non-cyclic model peptides in DMSO/methanol and in formic acid (Table 14 and Fig. 4). The assignments of the Gramicidin S signals used for this comparison are those reported recently by Hawkes and Randall.[33]

Table 14. ^{15}N NMR chemical shifts δ (ppm, upfield of external NO_3^-) of Gramicidin S and related oligopeptides

Peptide	Solvent	δ[a]
Ac—D—Phe—L—Pro—OMe	DMSO/MeOH[b]	−245.6 (*c*) −253.0 (*t*)
		−246.6 (*t*) −253.5 (*c*)
N-Acetyl—Leu—D-Phe—OMe	DMSO/MeOH[b]	−253.5; −260.4
	Formic acid	−248.9; −256.2
N-Z—Orn(δN·HCl)Leu—OMe	DMSO/MeOH[b]	−289.1; −258.8; −340.6
	Formic acid	−287.8; −255.9; −343.0
N-Methoxycarbonyl—Val—Orn(δN·HCl)—OMe	DMSO/MeOH[b]	−292.0; −257.5; −340.8
	Formic acid	[c] −254.5; −343.0
N-Methoxycarbonyl—Val—Orn(δN·HCl)—Leu—OMe	DMSO/MeOH[b]	−292.2; −256.8; −258.5; −340.7
Z—Gly—Pro—Val—*Val—OMe[d]	DMSO/MeOH[b]	−300.3; —; −262.5; −260.5*
	Formic acid	−299.2; —; −255.5; −253.9*
(—Pro—Val—Orn(δN·HCl)Leu—D—Phe—)$_2$	DMSO/MeOH[b]	−241.8; −262.4; −250.9; −340.6
		−252.6; −248.7
	Formic acid	−238.2; −255.5; −251.8; −343.2
		−248.3; −251.8

[a] The order of δ values matches that of the amino acid units of the formulae in the first column. Those signals compared with Gramicidin S are underlined.
[b] 4 : 1 by volume.
[c] Broad, because of proton exchange with the solvent.
[d] Only the signals of the *trans* isomer were measured.

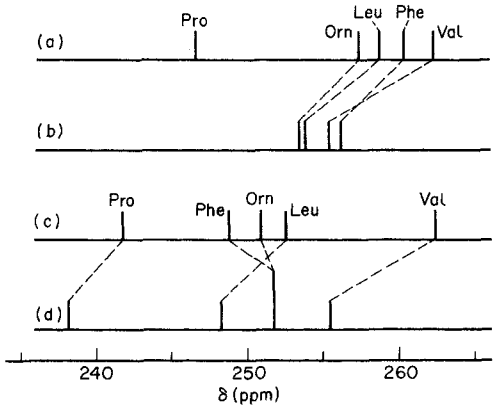
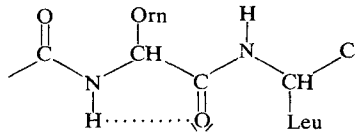


Figure 4. ^{15}N NMR chemical shifts of peptide groups in model compounds of Gramicidin S (Table 14), measured in (a) DMSO+20% MeOH; and (b) formic acid; and ^{15}N NMR signals of Gramicidin S in (c) DMSO+20% MeOH; and (d) formic acid.

Several papers dealing with ^{15}N NMR spectra of Gramicidin S have been published by various authors,[33–36] yet a systematic comparison of this cyclopeptide with model peptides simulating the NREs has not been made. In their second paper on Gramicidin S Hawkes *et al.* presented a comparison with various *N*-acetyl-amino acids;[34] however, the large shift differences were not explained. Urry *et al.*[36] found good agreement when Boc—Gly—Val—Gly—OMe was compared with the Val nitrogen of Gramicidin S in DMSO, yet models for other nitrogens of Gramicidin S were not studied. Our model peptides, systematically simulating the NREs of the Gramicidin S sequence (Table 14), demonstrate that the Val nitrogen of the cyclopeptide and that of the Pro—Val bond in Z—Gly—Pro—Val—Val—OMe do, indeed, have identical chemical shifts. However, all other nitrogens of Gramicidin S absorb several ppm downfield of the corresponding model peptides, and the largest shift differences (*c.* 7 and 12 ppm) are observed for the Phe and Orn units. We attribute these downfield shifts to the intramolecular H bonds which attack the carbonyl groups of Leu and Val. These H bonds favour the delocalization of the lone pairs of the attached Orn and Phe nitrogens, thus reducing their electron density and the activation energy of the $n-\pi^*$ transition. Since both effects increase the paramagnetic term of the screening constant (σ_p), a distinct downfield shift is the result. Our interpretation is confirmed by measurements in formic acid. Compared with DMSO solutions, downfield shifts are found for the Val, Leu and Pro nitrogens, obviously because strong H bonds are formed between solvent molecules and the carbonyl groups of Orn, Phe and Pro. This spectroscopic behaviour is in close analogy with that of the model peptides (Fig. 4), whereas the Phe and Orn nitrogens shift slightly upfield in the acidic solvent. Similar, but less pronounced, solvent effects were reported for Gramicidin S solutions in trifluoroethanol.[33,36]

Two observations still need discussion, namely (a) the downfield shifts of the Pro and Leu nitrogens of Gramicidin S relative to the corresponding model peptides; and (b) the fact that intramolecular H bonds cause stronger downfield shifts (10 ± 2 ppm) than intermolecular H bonds (5 ± 1.5 ppm).



Scheme 2. Intramolecular H bond of the Orn unit in Gramicidin S.

In the case of (a) we cannot offer a satisfactory explanation. Because the downfield shifts of the Pro and Leu nitrogen exist in DMSO and formic acid, they do not reflect solvent effects. Deviations of the Phe—Pro and Orn—Leu bond from planarity ($\omega \neq 0°, 180°$) are also unlikely, because they reduce the delocalization of the nitrogen lone pair and, thus, must result in upfield shifts. We assume a partial steric origin, such as compression of van der Waals radii, because our study of polyvaline and polyisoleucine demonstrates that the crowding of a bulky side chain in the close neighbourhood of peptide groups causes downfield shifts of several ppm (Fig. 2 and Ref. 20). The downfield shifts of L-L-cyclodipeptides relative to the L-D isomers[20] can be considered as further support for this hypothesis. In the case of the Leu nitrogen another potential explanation can be advanced through the existence of a weak H bond between the NH group of Orn and its own carbonyl group (Scheme 2). A relatively slow H/D exchange of the Orn N-proton also supports this hypothesis.[37] The strong shift effect due to intramolecular H bonds (b) is easier to explain. First, intramolecular H bonds exist permanently, in contrast to intermolecular H bonds. Second, rotations around the CO—N σ bond which reduce the delocalization of the nitrogen lone pair may occur in linear peptides, but not in cyclopeptides with a rigid conformation. Both effects together allow a satisfactory explanation why the ^{15}N NMR chemical shift of peptide nitrogens is a sensitive indicator of intramolecular H bonds.

CONCLUSIONS

Numerous investigations by various authors have shown the ^{15}N nucleus to be highly sensitive to changes of conformation and solvation. Since no method is known which allows the fast and accurate determination of the conformation of peptides in solution, the secondary structure of most peptides is still unknown. Hence, the ^{15}N NMR chemical shifts of peptides must be described in terms of easily accessible parameters such as neighbouring residue effects, substituent effects and solvent effects. These quasi-formal parameters allow one to establish a few crude structure/shift relationships and they have been found to be helpful for assignments of ^{15}N NMR signals. However, the present work also demonstrates clearly that reliable predictions of signal patterns, or accurate calculation of chemical shifts analogous to the increment method used for ^{13}C NMR spectra, are not feasible. Furthermore, the investigation of Gramicidin S spectra revealed, in agreement with previous result,[23] that intramolecular H bonds cause strong downfield

shifts of those peptide groups which are H-bond acceptors. Since cyclopeptides of polypeptides and noncyclic oligopeptides may adopt different conformations, even under otherwise identical conditions, signal assignments based exclusively on comparisons with model compounds are not reliable. Nevertheless, comparisons with model peptides may be worthwhile, because conspicuous shift differences allow the detection and interpretation of structural or conformational features of the peptide under investigation.

Acknowledgement

I wish to thank Dr W. E. Hull for the measurements of terpolypeptide spectra (Fig. 2) and the Deutsche Forschungsgemeinschaft for financial support.

REFERENCES

1. M. Llinas, W. J. Horseley and M. P. Klein, *J. Am. Chem. Soc.* **98,** 7554 (1976).
2. D. Gust, R. Moon and J. D. Roberts, *Proc. Nat. Acad. Sci. USA* **72,** 4696 (1975).
3. K. L. Williamson, L. G. Pease and J. D. Roberts, *J. Am. Chem. Soc.* **101,** 714 (1979).
4. G. E. Hawkes, W. M. Litchman and E. W. Randall, *J. Magn. Reson.* **19,** 255 (1975).
5. G. E. Hawkes, E. W. Randall and C. H. Bradley, *Nature London* **257,** 767 (1975).
6. G. E. Hawkes, E. W. Randall, W. E. Hull, D. Gattageno and F. Conti, *Biochemistry* **17,** 3986 (1978).
7. D. Gattageno, G. E. Hawkes and E. W. Randall, *J. Chem. Soc. Perkin Trans. 2* 1527 (1976).
8. V. Renugopalakrishnan, A Khaled, K. Okamoto and D. W. Urry, *Int. J. Quant. Chem. Quant. Biol. Sym.* **4,** 97 (1977).
9. H. R. Kricheldorf and W. E. Hull, *Biopolymers* **16,** 1609 (1977).
10. W. E. Hull and H. R. Kricheldorf, *Biopolymers* **17,** 2427 (1978).
11. H. R. Kricheldorf and W. E. Hull, *Org. Magn. Reson.* **12,** 607 (1979).
12. H. R. Kricheldorf and W. E. Hull; *Makromol. Chem.* **180,** 1715 (1979).
13. M. Witanowski, L. Stefaniak and H. Januszewski, in *Nitrogen NMR,* 1st Edn, ed. by M. Witanowski and G. A. Webb, Plenum Press, New York (1973).
14. T. B. Posner, V. Markowski, L. Loftus and J. D. Roberts, *Chem. Commun.* 769 (1975).
15. H. R. Kricheldorf and W. E. Hull, *J. Polym. Sci. Polym. Chem. Ed.* **16,** 583 (1978).
16. H. R. Kricheldorf and W. E. Hull, *Makromol. Chem.* **180,** 161 (1979).
17. H. R. Kricheldorf, *Makromol. Chem.* **179,** 2675 (1978).
18. J. A. Sogn, W. A. Gibbons and E. W. Randall, *Biochemistry* **12,** 2100 (1973).
19. H. R. Kricheldorf, *Org. Mag. Reson.* **12,** 414 (1979).
20. H. R. Kricheldorf, *Org. Magn. Reson.* **13,** 54 (1980).
21. J. Belleau and G. Malek, *J. Am. Chem. Soc.* **90,** 1651 (1968).
22. H. R. Kricheldorf and W. E. Hull, *Org. Magn. Reson.* **13,** 335 (1980).
23. H. R. Kricheldorf and W. E. Hull, *Biopolymers* **19,** 1103 (1980).
24. H. R. Kricheldorf, *Makromol. Chem.* **181,** 507 (1980).
25. H. R. Kricheldorf, *Makromol. Chem.* **180,** 147 (1979).
26. H. R. Kricheldorf and W. E. Hull, *Macromolecules* **13,** 87 (1980).
27. H. R. Kricheldorf and W. E. Hull, *Makromol. Chem.* **181,** 1949 (1980).
28. H. R. Kricheldorf and W. E. Hull, *J. Macromol. Sci. Chem.* **12,** 51 (1978).
29. H. R. Kricheldorf and W. E. Hull, *Makromol. Chem.* **178,** 253 (1977).
30. P. S. Pregosin, E. W. Randall and A. I. White, *Chem. Commun.* 1602 (1971).
31. H. R. Kricheldorf, *Org. Magn. Reson.* in preparation.
32. J. F. Cotzee and C. D. Ritchie, in *Solute–Solvent Interactions,* 1st Edn, p. 108. Marcel Dekker, New York (1969).
33. G. E. Hawkes, E. W. Randall and W. E. Hull, *Biopolymers* in press.
34. G. E. Hawkes, E. W. Randall and C. H. Bradley, *Nature London* **257,** 767 (1975).
35. G. E. Hawkes, W. M. Litchman and E. W. Randall, *J. Magn. Reson.* **19,** 255 (1975).
36. A. Khaled, D. W. Urry, H. Sugano, M. Miyoshi and N. Izumiya, *Biochemistry* **17,** 2490 (1978).
37. A. Stern, W. A. Gibbons and L. C. Craig, *Proc. Nat. Acad. Sci. USA* **61,** 734 (1968).

Received 6 May 1980; accepted (revised) 29 July 1980