232. Preferred Spatial Arrangement of the Aromatic Side Chains in Linear Oligopeptides Containing Tyrosine

by **Kurt Wüthrich** and **Antonio de Marco**

Institut für Molekularbiologie und Biophysik
Eidgenössische Technische Hochschule, 8093 Zürich, Switzerland

(23. VI. 76)

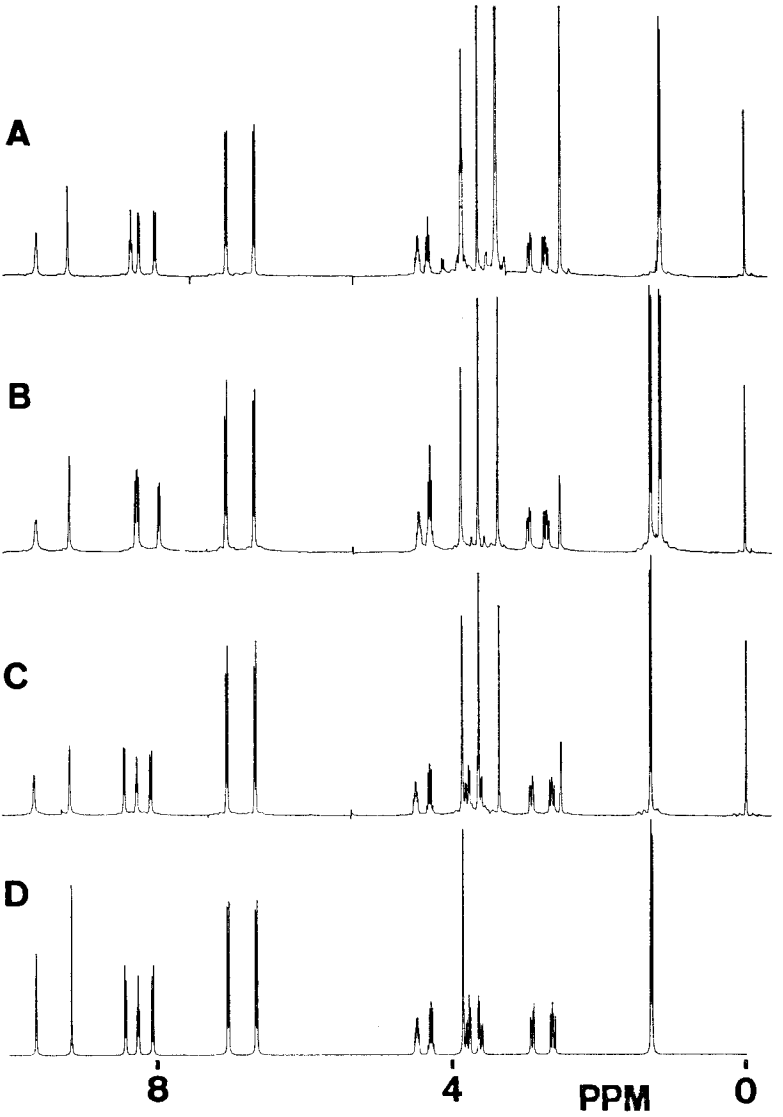
Summary. ^{1}H-NMR. studies of the protected linear tetrapeptides CF_3CO-Gly-Gly-L-Tyr-L-Ala-OCH_3, CF_3CO-Gly-L-Ala-L-Tyr-L-Ala-OCH_3 and CF_3CO-Gly-L-Ala-L-Tyr-Gly-OCH_3 showed that the side chain of the tyrosyl residue was in all three peptides preferentially oriented towards the amino terminus of the peptide chain. This preferred spatial arrangement of the aromatic side chain was manifested in the chemical shifts of the amino acid residue preceding tyrosine and in the vicinal spin-spin coupling constants $^3J_{H\alpha C\beta H}$ of tyrosine.

Introduction. – Recently we investigated a series of twenty protected linear tetrapeptides CF_3CO-Gly-Gly-X-L-Ala-OCH_3, where X stands for one of the common amino acids [1] [2]. The NMR. data on these compounds indicated that the peptide backbone was predominantly in a flexible extended form, whereas the side chain of residue X was predominantly oriented towards the amino terminus of the peptide chain. This non-random side chain conformation was most clearly manifested for the aromatic amino acids, where the ring current field [3] [4] markedly affected the ^{1}H-NMR. chemical shifts of the glycyl residue preceding X [2]. Inspection of space filling molecular models indicated that the preferred orientation towards the amino terminus might be a consequence of steric interactions between the side chains of residues (3) and (4) in these tetrapeptides. To further investigate this hypothesis, the tetrapeptide ester CF_3CO-Gly-Gly-L-Tyr-L-Ala-OCH_3 was now compared with the two analogs CF_3CO-Gly-L-Ala-L-Tyr-L-Ala-OCH_3 and CF_3CO-Gly-L-Ala-L-Tyr-Gly-OCH_3.

Experimental Part. – The three peptides used in the present study were obtained from *Bachem* AG, Liestal, Switzerland. For the NMR. measurements, 0.1M solutions in d_6-DMSO were prepared.

^{1}H-NMR. spectra were recorded on a *Bruker* HXS 360 spectrometer. ^{13}C-NMR. spectra at 25.16 MHz and ^{19}F-NMR. spectra at 94 MHz were recorded on a *Varian* XL-100 spectrometer. In the context of the present study, the ^{13}C- and ^{19}F-NMR. data were mainly used to ascertain the covalent structure of the peptides [1] [2].

Results. – The ^{1}H-NMR. spectra of the three peptides studied are shown in Fig. 1 and the chemical shifts and spin-spin coupling constants are listed in Table 1. The components of the individual spin-systems were identified with double resonance experiments. The assignments of the spin systems to specific amino acid residues, in as far as they were not obvious from inspection of the covalent structure, were based on comparison of the spectra in Fig. 1 with those of CF_3CO-Gly-Gly-L-Ala-L-Ala-OCH_3 and CF_3CO-Gly-Gly-OH [2].



Fourier *transform* ^1H-*NMR. spectra at 360 MHz of 0.1* M *peptide solutions in* d_6-*DMSO at 25°*

A. CF_3CO-Gly-L-Ala-L-Tyr-Gly-OCH_3

B. CF_3CO-Gly-L-Ala-L-Tyr-L-Ala-OCH_3

C. CF_3CO-Gly-Gly-L-Tyr-L-Ala-OCH_3

D. Spectrum of peptide C computed with the parameters of Table 1 and a half width at half height of the lines of 1.1 Hz. The solvent resonances at 2.50 ppm (DMSO) and 3.38 ppm (H_2O), the methylester line at 3.62 ppm and the resonance of TMS at 0 ppm were omitted in the calculated spectrum.

The dependence on temperature of the ^{1}H-NMR. spectra in Fig. 1 was studied over the range 20–80°. The temperature coefficients for the amide protons are given in Table 2. Both the chemical shifts of non-labile protons and the spin-spin coupling constants given in Table 1 were found to be essentially independent of temperature.

The ^{13}C- and ^{19}F-NMR. spectra confirmed the covalent structures of the peptides given in Table 1. The ^{13}C chemical shifts were found to be essentially identical to those expected on the basis of previous studies of a series of analogous peptides [1].

Discussion. – The NMR. data of Tables 1 and 2 can provide information on the conformation of the peptide backbone and the spatial arrangement of the side chain of Tyr (3). These two aspects of the molecular conformation will now be discussed in more detail.

The peptide backbone conformation is most directly manifested in the vicinal coupling constants $J_{HN(\alpha)}$ between the amide protons and the C(α)-protons, and in the temperature dependence of the amide proton resonances. The observed coupling constants $J_{HN(\alpha)}$ (Table 1) correspond to the dynamic average of typical 'random coil' distributions of rotation states Φ_i about the $N_i - C_i(\alpha)$ bonds in the individual dipeptide fragments of the peptide chains [2] [5] [6][1]). The temperature coefficients for the chemical shifts of the amide protons (Table 2) are quite typical for solvent exposed peptide groups in DMSO [7] [8], which is certainly compatible with an extended flexible form of the polypeptide chain.

Turning now to the side chain conformation of Tyr(3) in the three peptides of Table 1, we have that the spin-spin coupling constants $^3J_{\alpha\beta}$ are indicative of non-random populations of the rotamers about the C(α)–C(β) bond. Following *Pachler* [9], one has the relations (1) to (3) between the populations of the three staggered rotamers (Table 3), n_1, n_2 and n_3, and the vicinal spin-spin coupling constants.

$$n_2 = \frac{J_g - J_{BX}}{J_g - J_t} \tag{1}$$

$$n_3 = \frac{J_g - J_{AX}}{J_g - J_t} \tag{2}$$

$$n_1 = 1 - (n_2 + n_3) \tag{3}$$

where J_{AX} and J_{BX} are the observed parameters $^3J_{\alpha\beta}$ (Table 1), $J_g = 2.6$ Hz is the value of $^3J_{\alpha\beta}$ for two protons in gauche configuration and $J_t = 13.6$ Hz the value of $^3J_{\alpha\beta}$ for two protons in *trans* configuration. Using the spin-spin coupling constants $^3J_{\alpha\beta}$ of Table 1 in equations (1) to (3), the rotamer populations in Table 3 were

[1]) A similar result was previously obtained from studies of a series of linear tetrapeptides CF_3CO-Gly-Gly-X-L-Ala-OCH_3, where X stands for one of the common amino acid residues [2]. In [2] it was also pointed out that in certain samples of the peptides with X = Ser, Thr or Tyr, a small apparent coupling constant $^3J_{HN(\alpha)}$ of Gly(1) had been observed. An apparent value of approximately 3 Hz for $^3J_{HN(\alpha)}$ of Gly(1) was now also observed in certain samples of the peptides of Table 1. By systematic variation of the water content in the d_6-DMSO solutions and by saturation transfer studies, it could be ascertained that $^3J_{HN(\alpha)} = 5.7$ Hz in all cases, and that the small residual line separations in certain samples resulted from rapid exchange of the amide proton of Gly(1).

obtained. Table 3 shows that for all three peptides studied, one of the rotamers (2) and (3) is preferentially populated. Since the C(β) methylene protons of Tyr(3) had not been individually assigned, the preferentially populated states in the individual peptides could not be further identified on the basis of the $^3J_{\alpha\beta}$ analysis alone.

It had been shown in previous studies of CF_3CO-Gly-Gly-L-Tyr-L-Ala-OCH_3 that the non-random orientation of the aromatic ring was manifested in high field shifts of the C(α)-protons of Gly(2) [2]. Inspection of the chemical shifts in Table 1 now revealed similar high field shifts for the ^{1}H-resonances of Ala(2) in the other two peptides. The chemical shifts of Ala(2) (Table 1) were compared with the corresponding chemical shifts of Ala(3) in CF_3CO-Gly-Gly-L-Ala-L-Ala-OCH_3, *i.e.* δ(C(α)H) = 4.34 ppm and δ(C(β)H) = 1.22 ppm [2]. It is seen that the resonances of the alanyl residue preceding Tyr are shifted upfield by approximately 0.07 ppm, which can readily be accounted for by interactions of Ala(2) with the ring current field [3] [4] of Tyr(3). On the other hand, the chemical shifts of Ala(4) in CF_3CO-Gly-L-Ala-L-Tyr-L-Ala-OCH_3 (Table 1) fall within the ranges observed for numerous analogous peptides, *i.e.* δ(C(α)H) = 4.28 $\pm$ 0.04 ppm and δ(C(β)H) = 1.30 $\pm$ 0.02 ppm. The chemical shifts of Gly(4) in CF_3CO-Gly-L-Ala-L-Tyr-Gly-OCH_3 are also very near the range of chemical shifts observed for Gly(2) in peptides CF_3CO-Gly-Gly-X-L-Ala-OCH_3, with a non-aromatic residue X, *i.e.* δ = 3.80 $\pm$ 0.05 ppm [2].

In conclusion, the combination of $^3J_{\alpha\beta}$ of Tyr(3) and the chemical shift data shows that the rotamer population about the C(α)–C(β) bond of Tyr(3) in the three peptides of Tables 1–3 is non-random, with preferred orientation of the aromatic ring towards the amino terminus of the peptide chain. In the notation of Table 3 this corresponds to preferred population of the rotamer(2). It thus appears that the spatial arrangement of the side chain of residue 3 in these linear flexible tetrapeptides is not greatly affected by substituting Gly with Ala, or *v.v.*, in the neighboring positions (2) and (4). It may be recalled from the earlier paper [2] that the behavior described here for tyrosine is typical for all four common aromatic amino acid residues and probably also for other bulky amino acid side chains.

The molecular conformations proposed on the basis of $^3J_{\alpha\beta}$ of Tyr(3) and the chemical shifts of the non-labile protons seem to provide an explanation for other experimental observations. One is that the amide proton of Gly(1) exchanges more rapidly with protons of H_2O than the amide protons of the other three residues (footnote 1). Saturation transfer studies showed that the hydroxyl proton of Tyr(3) was also involved in this rapid proton exchange, suggesting that the increased rate of exchange of amide proton(1) is a consequence of its close proximity to the hydroxyl group of Tyr (3) in the spatial structures of these tetrapeptides. Second, in all three peptides the temperature coefficient of the amido proton of residue(2) is slightly smaller than those for the other three residues (Table 2). In the proposed molecular species this is quite likely a consequence of partial shielding from the solvent of amide proton(2) by the nearby aromatic ring of Tyr(3). For the characterization of the peptide conformations it is further interesting that the NMR. parameters of the non-labile protons, in particular $^3J_{\alpha\beta}$ of Tyr(3), are within the accuracy of the present experiments essentially independent of temperature over the range 20–80°. This indicates that the NMR. parameters of Table 1 correspond to an average over an ensemble of rapidly interconverting molecular species which occupy a region of the conformational energy

Table 1. *^{1}H-NMR. parameters at $T = 25°$ of the three protected linear tetrapeptides CF_3CO-Gly-Gly-L-Tyr-L-Ala-OCH_3, CF_3CO-Gly-L-Ala-L-Tyr-L-Ala-OCH_3 and CF_3CO-Gly-L-Ala-L-Tyr-Gly-OCH_3.* The chemical shifts δ are in ppm from TMS, the spin-spin coupling constants J in Hz ($\pm$ 0.3 Hz). The parameters were verified by spectrum simulation, as illustrated in Fig.1 C and D.

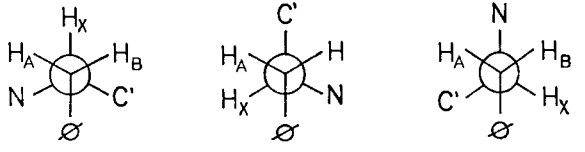
Amino Acid Residue	NMR. Parameter	CF_3CO-Gly-Gly-Tyr-Ala-OCH_3	CF_3CO-Gly-Ala-Tyr-Ala-OCH_3	CF_3CO-Gly-Ala-Tyr-Gly-OCH_3
Gly(1)	δ(NH)	9.625	9.587	9.587
	δ(C(α)H)	3.840	3.844	3.838
		3.840	3.844	3.838
	$^3J_{HN(\alpha)}$	5.7	5.7	5.7
		5.7	5.7	5.7
Gly(2) or Ala(2)	δ(NH)	8.242	8.225	8.200
	δ(C(α)H)	3.765	4.259	4.286
		3.607		
	δ(C(β)H)		1.146	1.153
	$^3J_{HN(\alpha)}$	5.8	6.8	7.3
		5.8		
	$^2J_{\alpha\alpha}$	16.6		
	$^3J_{\alpha\beta}$		6.8	6.8
Tyr(3)	δ(NH)	8.052	7.927	7.981
	δ(C(α)H)	4.464	4.406	4.427

	δ(C(β)H)	2.903	2.926	2.914
		2.627	2.682	2.699
	δ(C(δ)H)	6.641	6.644	6.640
	δ(C(ε)H)	7.027	7.027	7.018
	δ(OH)	9.147	9.144	9.166
	$^3J_{HN(\alpha)}$	8.3	8.3	8.7
	$^3J_{\alpha\beta}$	4.4	4.4	4.8
		9.7	9.7	9.2
	$^2J_{\beta\beta}$	13.6	13.6	13.6
	$^3J_{\delta\varepsilon}$	8.3	8.3	8.3
Gly(4) or Ala(4)	δ(NH)	8.415	8.234	8.313
	δ(C(α)H)	4.274	4.259	3.860
				3.818
	δ(C(β)H)	1.290	1.278	
	$^3J_{HN(\alpha)}$	7.3	7.3	5.8
				5.8
	$^2J_{\alpha\alpha}$			17.5
	$^3J_{\alpha\beta}$	7.3	7.3	
Methylester	δ(OCH_3)	3.623	3.613	3.622

Table 2. *Temperature coefficients* (in 10^{-3} ppm deg^{-1}) *of the amide proton resonances in* d_6*-DMSO solutions of the three linear protected tetrapeptides of Table 1*

Amino Acid Residue	CF_3CO-Gly-Gly-Tyr-Ala-OCH_3	CF_3CO-Gly-Ala-Tyr-Ala-OCH_3	CF_3CO-Gly-Ala-Tyr-Gly-OCH_3
Gly(1)	−4.58	−6.86	−5.53
Gly(2) or Ala(2)	−3.56	−4.79	−4.31
Tyr(3)	−5.50	−6.45	−5.76
Gly(4) or Ala(4)	−5.74	−6.79	−5.98

Table 3. *Populations of the three staggered rotamers about the* χ_1 *angle of Tyr(3) computed from the spin-spin coupling constants* $^3J_{\alpha\beta}$ (see text)



(1) (2) (3)

Rotamer Populations	CF_3CO-Gly-Gly-Tyr-Ala-OCH_3	CF_3CO-Gly-Ala-Tyr-Ala-OCH_3	CF_3CO-Gly-Ala-Tyr-Gly-OCH_3
n_1	0.19	0.19	0.18
n_2 (or n_3)	0.16	0.16	0.21
n_3 (or n_2)	0.65	0.65	0.61

map bounded by rather steep energy gradients, so that only relatively few additional conformations are admixed when the temperature is raised from 20° to 80°. In other words the preferred non-random conformations observed in these overall rather flexible oligopeptides are maintained over the temperature range where thermal denaturation is observed for most globular proteins!

Recent theoretical work on protein folding dynamics [10] [11] emphasizes the importance of local non-random conformations in short segments of linear flexible polypeptide chains as primary nucleation sites in the formation of globular proteins. Conformational studies of linear oligopeptides might produce experimental support for these hypotheses. However, these model investigations are in general rather difficult, mainly because the observed parameters correspond to an average over a manifold of different conformations in rapid equilibrium, which makes variations in the relative populations of certain conformational states difficult to detect [12]. It is therefore quite gratifying that the experiments described in this and a previous paper [2] appear to point towards a conformational phenomenon which might be quite common for aromatic amino acid side chains in flexible polypeptide chains, and possibly for other bulky side chains as well.

Financial support by the *Schweizerischer Nationalfonds zur Förderung der wissenschaftlichen Forschung* (Project 3.1510.73) and the *Italian N.R.C.* (fellowship to *A. De M.*) is gratefully acknowledged.

REFERENCES

[1] *C. Grathwohl & K. Wüthrich*, J. magn. Res. *13*, 217 (1974).
[2] *A. Bundi, C. Grathwohl, J. Hochmann, R. M. Keller, G. Wagner & K. Wüthrich*, J. magn. Res. *18*, 191 (1975).
[3] *J. A. Pople*, J. chem. Physics *24*, 1111 (1956).
[4] *C. E. Johnson & F. A. Bovey*, J. chem. Physics *29*, 1012 (1958).
[5] *G. N. Ramachandran & V. Sasisekharan*, Advan. Protein Chemistry *23*, 283 (1968).
[6] *G. N. Ramachandran, R. Chandrasekharan & K. D. Kopple*, Biopolymers *10*, 2113 (1971).
[7] *M. Ohnishi & D. W. Urry*, Biochem. biophys. Res. Commun. *36*, 194 (1969).
[8] *K. Wüthrich*, NMR. in Biological Research: Peptides and Proteins (North Holland, Amsterdam 1976).
[9] *K. G. R. Pachler*, Spectrochim. Acta *20*, 581 (1964).
[10] *B. Gelin & M. Karplus*, Proc. Nat. Acad. Sci. *72*, 2002 (1975).
[11] *M. Karplus & D. L. Weaver*, Nature *260*, 404 (1976).
[12] *K. Wüthrich & C. Grathwohl*, F.E.B.S. Letters *43*, 337 (1974).