

¹H NMR Studies on the Preferred Interactions of Guanidinium and C-Terminal Carboxylate Groups in Arginine-containing Peptides

Vassilios Tsikaris,^a Mahn Thong Cung,^b Eugenia Panou-Pomonis,^a Constantinos Sakarellos^a and Maria Sakarellos-Daitsiotis^{*,a}

^a Department of Chemistry, University of Ioannina, Box 1186, GR-45110 Ioannina, Greece

^b LCPM-ENSIC-INPL, URA-CNRS-494, BP 451, F-54001 Nancy Cedex, France

Interactions between positively charged side-chain residues and C-terminal carboxylate groups are of great importance in stabilizing secondary structures of many biologically active peptide sequences. The conformational properties of the hexapeptide sequence L-Arg-L-Val-Gly-L-Arg-L-Val-Gly with an Arg residue located at the third position from the carboxy terminus and a second Arg moiety at the amino terminus of the molecule, have been investigated in DMSO solution at pH 6.5, by means of 1D and 2D ¹H NMR spectroscopy. The considerable downfield shift of the Arg⁴N⁴H resonance, when the pH of the solution is raised from 2 to 6.5, and the magnetically non-equivalent C^βH and C^γH methylene protons of the same residue at pH 6.5 provide evidence for a non-covalent hydrogen-bonding interaction between Arg⁴ guanidinium and C-terminal carboxylate groups. The intense backbone-backbone NOE connectivities observed for the L-Arg⁴-L-Val⁵-Gly⁶ C-terminal sequence, the NOE connectivities of Arg⁴ methylene protons and the retained ABX system of Gly⁶ C-terminal residue at high temperature are evidence for some mobility restrictions around this part of the molecule, and provide additional confirmation for a particular propensity of the arginine side chain, when occupying a position near the carboxy terminus, to interact with the C-terminal carboxylate group at neutral pH. Absence of ROE connectivities between protons of the Arg¹ side chain, and the fact that almost all methylene protons of the same residue appear magnetically equivalent, point out that this Arg moiety does not participate in any intramolecular interaction. Thus, differences in the behaviour of guanidinium side chains of the reported peptide sequence can be used as a general probe for the study of Arg side-chain mobility and for its ability to generate hydrogen-bonding interactions in other peptides and proteins.

Non-covalent hydrogen-bonding interactions or salt bridge formation are critical for the establishment of secondary structure features in peptides and proteins, and are intimately linked to many fundamental selective recognition processes inside cells.¹

For example, it has been postulated that stabilization of β-turns in the sequences Asp-X-Y-Arg and Arg-X-Y-Glu is modulated by the formation of a salt bridge between the side chains of Arg and Asp or Glu residues respectively, and the importance of these conformations in the expressed biological activities has been well documented.²⁻⁵ Investigations on the binding capacity of synthetic decapeptide-analogues, corresponding to the main immunogenic region (MIR, Trp⁶⁷-Asn-Pro-Ala-Asp⁷¹-Tyr-Gly-Gly-Ile-Lys⁷⁶) of the acetylcholine receptor (AChR), to anti-AChR monoclonal antibodies, revealed that a possible electrostatic interaction between Asp⁷¹ and Lys⁷⁶ side chains should restrain a close molecular recognition of MIR by the mAbs⁶.

The apparent importance of interactions between charged side-chain residues and N-, C-terminal groups was supported by conformational studies on the helical C-peptide (residues 1-13) of RNase A. It was found that Glu² and His¹², which are close to the positive and negative pole of the helical macrodipole respectively, can make favourable charge-dipole interactions that stabilize the helix.^{7,8} Likewise recent crystallographic studies on the Arg-Gly-Asp sequence, which was suggested as the key recognition element for cell adhesion, have shown that the guanidinium group of Arg forms, through a water molecule, a pair of H-bonds with the C-terminal carboxylate moiety, leaving the β-carboxylate group of Asp available for other interactions.⁹⁻¹¹

In a previous study¹² we have demonstrated that the

guanidinium side chain of arginine in the L-Arg-X-Gly peptide sequence (X = L-Ala, L-Val, L-Leu) participates in an intramolecular interaction with the C-terminal carboxylate group in DMSO solutions at neutral pH. The present work was designed to further investigate, using 1D and 2D ¹H-NMR spectroscopy, whether the side chain of arginine shows a similar propensity to interact with the C-terminal carboxylate group in a peptide sequence longer than three residues, where various possibilities for secondary structures may occur. Therefore, we have synthesized the L-Arg-L-Val-Gly-L-Arg-L-Val-Gly hexapeptide with an Arg residue localized at the third position from the carboxy terminus, as in the L-Arg-L-Val-Gly tripeptide, and a second Arg moiety at the amino terminus of the peptide sequence. It is also apparent that differences in the behaviour of the two guanidinium side chains of the examined hexapeptide might serve as a general probe for the study of Arg side-chain mobility and accessibility for hydrogen bonding interactions in various peptides and proteins.

Experimental

Materials and Methods.—Preparation of the *N*^ε-*tert*-butyloxycarbonyl-*N*^ω-toluene-*p*-sulfonyl-L-arginyl-L-valylglycine methyl ester and the *N*^ε-*tert*-butyloxycarbonyl-*N*^ω-toluene-*p*-sulfonyl-L-arginyl-L-valylglycine have been reported previously.¹³ The *tert*-butyloxycarbonyl group of the protected tripeptide methyl ester was selectively removed by 100% anhydrous trifluoroacetic acid. Synthesis of the protected hexapeptide was realized by the mixed anhydride procedure as follows.

A solution of *N*^ε-*tert*-butyloxycarbonyl-*N*^ω-toluene-*p*-sulfonyl-L-arginyl-L-valylglycine (0.32 g, 0.55 mmol) and *N*-methylmorpholine (0.05 cm³, 0.55 mmol) in 2 cm³ tetrahydrofuran

Table 1 Proton magnetic resonance data (ppm) of the L-Arg¹-L-Val²-Gly³-L-Arg⁴-L-Val⁵-Gly⁶ hexapeptide (8×10^{-3} mol dm⁻³) in [2H₆]DMSO at 298 K from tetramethylsilane

pH	Residue	⁺ NH ₃	NH ^a	N ^ε H	N ^ε H ₂	C ^α H	C ^β H	C ^γ H	C ^δ H
2	Arg ¹	8.30		7.77	7.19	3.96	1.72	1.51	3.15
	Val ²		8.55			4.22	2.00	0.92	
	Gly ³		8.32			3.86		0.90	
	Arg ⁴		8.10	7.72	7.19	4.37	1.72	1.51	3.08
	Val ⁵		7.92			4.17	2.00	0.87	
	Gly ^{6,c}		8.32			3.81		0.85	
						3.71			
6.5	Arg ¹	n.d. ^b		n.d. ^b	7.13	3.19	1.62	1.52	3.08
	Val ²		7.99			4.18	1.97	0.86	
	Gly ³		8.31 (-3.88)			3.77		0.83	
						3.71			
	Arg ⁴		7.86 (-5.81)	9.46	7.13	4.41	1.89	1.52	3.11
	Val ⁵		8.27 (-3.23)		7.60	4.05	1.58	0.84	2.96
	Gly ⁶		7.37 (-3.87)			3.37			
						3.34			

^a The temperature coefficients (10^{-3} ppm K⁻¹) are specified in parentheses for the NH proton signals. ^b n.d. = not detected. ^c Gly⁶CO₂H at 12.45 ppm.

was cooled to -17 °C with stirring. Isobutyl chloroformate (0.07 cm³, 0.55 mmol) was added, followed after 5 min by a solution of *N*^ω-toluene-*p*-sulfonyl-L-arginyl-L-valylglycine methyl ester trifluoroacetate (0.30 g, 0.5 mmol) and triethylamine (0.07 cm³, 0.5 mmol) in 1.5 cm³ cold tetrahydrofuran. The reaction mixture was maintained for 1 h at 0 °C and, after 24 h at room temperature, evaporated to dryness. The product was precipitated with anhydrous diethyl ether and triturated several times with water, 1 mol dm⁻³ NaHSO₄, water, 1 mol dm⁻³ NaHCO₃ and water. The peptide was found homogenous by thin-layer chromatography in chloroform-carbon tetrachloride-methanol (6:3:1) ($R_f = 0.27$) and in acetone-acetic acid (98:2) ($R_f = 0.77$). The yield was 0.39 g (73%).

Saponification of the *N*^α-*tert*-butyloxycarbonyl-*N*^ω-toluene-*p*-sulfonyl-L-arginyl-L-valylglycyl-*N*^ω-toluene-*p*-sulfonyl-L-arginyl-L-valylglycine methyl ester (0.75 g, 0.7 mmol) was carried out in a solution of 7 cm³ dioxane containing 2.7 cm³, 1 mol dm⁻³ sodium hydroxide. The solution was stirred for 1.5 h at room temperature, neutralized with 1 mol dm⁻³ NaHSO₄, evaporated to remove dioxane, and acidified at 0 °C. The precipitate was filtered and washed with cold water. The peptide was found homogeneous by thin-layer chromatography in butanol-acetic acid-water (4:1:5) ($R_f = 0.67$) and the yield was 0.67 g (91%).

The unprotected hexapeptide was obtained after treatment of the *N*^α-*tert*-butyloxycarbonyl-*N*^ω-toluene-*p*-sulfonyl-L-arginyl-L-valylglycyl-*N*^ω-toluene-*p*-sulfonyl-L-arginyl-L-valylglycine (0.50 g, 0.7 mmol) with 15 cm³ liquid hydrogen fluoride (HF) in the presence of 2 cm³ anisole, for 1 h at 0 °C. The product was precipitated with anhydrous ether, and the precipitate was dissolved in water and lyophilized. Purification was carried out by gel filtration on Sephadex G-25 in 2 mol dm⁻³ acetic acid and the overall yield was 0.25 g (55%).

The NMR samples were prepared by dissolving the solid material in water and adjusting the pH to the desired value with hydrochloric acid or sodium hydroxide. The obtained aqueous solutions were lyophilized and weighed amounts were dissolved in [2H₆]DMSO at concentrations *ca.* 8×10^{-3} mol dm⁻³. At this concentration intermolecular associations were excluded, as was proved by dilution studies.

1D and 2D ¹H NMR experiments were performed on Bruker AC 200 and AM 400 apparatus. ROESY experiments were carried out by changing the frequency offset of the carrier in order to be sure that the NOEs observed represented true proximity between protons and were not the result of Hartman-Hahn transfer¹⁴ coupled with incoherent magnetization transfer. A spin-locking mixing time of 350 ms was used with a radiofrequency field strength of 2.4 KHz. The MLEV 17 HOHAHA experiment¹⁵ was scanned with a spin locking mixing time of 100 ms and a radiofrequency field strength of 9.6 kHz.

Results and Discussion

¹H Resonance Assignments.—The spectral analysis of the L-Arg¹-L-Val-Gly-L-Arg⁴-L-Val-Gly⁶ hexapeptide was based on the combined use of COSY, HOHAHA and ROESY experiments. The complete assignment of all proton resonances is summarized in Table 1 and is in agreement with previously reported ¹H NMR studies on arginine-containing peptides.^{4,12,16-18}

The chemical shifts of the N^εHs of both Arg¹ and Arg⁴ residues were found almost identical (*ca.* 7.7 ppm) at pH 2. When the pH of the hexapeptide solution was raised from 2 to 6.5, the Arg⁴N^εH signal shifted to low field (from 7.72 to 9.46 ppm), and the Arg¹N^εH signal disappeared. Coincidence of the N^εH signals (7.19 ppm) of both Arg¹ and Arg⁴ guanidinium groups was observed at pH 2, while splitting of this signal (at pH 6.5) provided two unsymmetric resonance peaks at 7.13 and 7.60 ppm, suggesting that one of the ArgN^εH₄ groups was split at pH 6.5 (Fig. 1).

All methylene protons of Arg¹ and Arg⁴ residues gave single signals, indicating that these protons were magnetically equivalent at pH 2. In contrast, the Arg⁴C^βH₂, C^δH₂ protons became magnetically non-equivalent ($\Delta\delta$ values were 0.31 and 0.15 ppm respectively), when the pH was raised to 6.5. The same was also true only for the Arg¹C^βH₂ protons, which however showed a smaller $\Delta\delta$ value (0.24 ppm) compared to the Arg⁴C^βH₂ protons at the same pH (Fig. 1).

The Gly⁶CO₂H resonance at 12.45 ppm vanished, as expected, when the pH of the peptide solution was raised to 6.5, whereas the ArgN⁺H₃ signal at 8.30 ppm of the acidic peptide

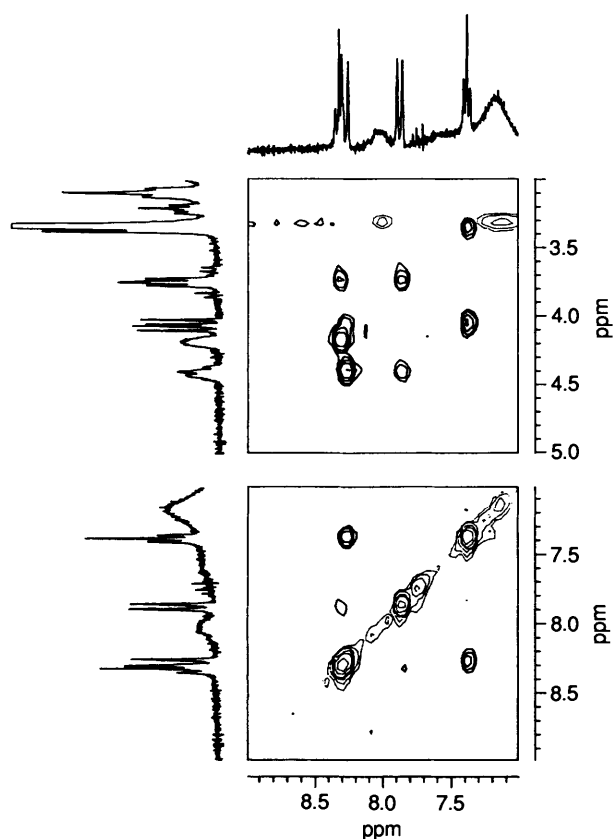


Fig. 3 Portions of ROESY map of L-Arg-L-Val-Gly-L-Arg-L-Val-Gly hexapeptide in $[^2\text{H}_6]$ DMSO at pH 6.5

and 3.71 ppm, coupling constant values $^3J_{\text{NH-C}^\alpha\text{H}} = 5.9$ Hz, $^3J_{\text{NH-C}^\alpha\text{H}} = 5.8$ Hz and $^2J_{\alpha\text{H},\alpha'\text{H}} = -17.5$ Hz) were considerably upfield shifted, upon deprotonation of the adjacent Gly⁶CO₂H, as already mentioned (Fig. 1). At neutral pH and 298 K, the Gly⁶ methylene protons were almost magnetically equivalent with resonance peaks at 3.37 and 3.34 ppm (chemical shift difference equal to 0.03 ppm) and the ABX system of Gly⁶ residue was estimated ($^3J_{\text{NH-C}^\alpha\text{H}} = 5.1$ Hz, $^3J_{\text{NH-C}^\alpha\text{H}} = 4.0$ Hz and $^2J_{\alpha\text{H},\alpha'\text{H}} = -16.6$ Hz). When the temperature of the neutral peptide solution was raised from 298 to 328 K, the chemical shift difference of the Gly⁶C^αH₂ protons increased from 0.03 to 0.05 ppm, whereas the coupling constant values remained unchanged.

Conformational State of the L-Arg¹-L-Val-Gly-L-Arg⁴-L-Val⁵-Gly⁶ Hexapeptide in $[^2\text{H}_6]$ DMSO Solution at pH 6.5.—When the pH of the hexapeptide solution was raised from 2 to 6.5, intense backbone-backbone NOE connectivities appeared in a ROESY spectrum (Fig. 3) between C^αH(*i*)/NH(*i* + 1) and NH(*i*)/C^αH(*i*) protons. However the unique intense NOE cross peak between successive amide protons was that of Val⁵NH and Gly⁶NH, similar to the NOE effect recorded for the consecutive Val and Gly residues in the L-Arg-Val-Gly tripeptide unit.^{1,2} Strong NOE correlations were also found between the geminal Arg⁴C^βH₂ protons, the Arg⁴C^αH and each of the Arg⁴C^βH/Arg⁴C^γH₂ protons, contrary to the Arg¹ methylene protons at pH 6.5. We therefore assume that the L-Arg⁴-L-Val⁵-Gly C-terminal sequence of the hexapeptide shows a rigid conformation with some mobility restrictions at the Arg⁴ side chain and resembles the L-Arg-L-Val-Gly tripeptide at neutral pH. It is likely that the ABX system already mentioned for the Gly⁶ moiety, though a carboxy-terminus residue, provides further evidence for some mobility restrictions at the C-terminal part of the molecule.

The above mentioned findings, the considerable downfield

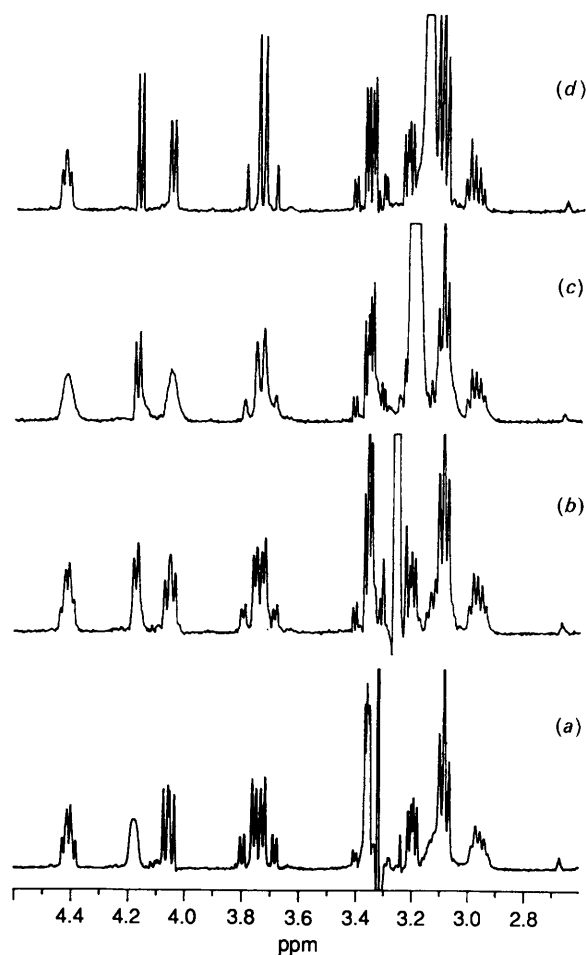


Fig. 4 The 400 MHz spectra of L-Arg-L-Val-Gly-L-Arg-L-Val-Gly hexapeptide in $[^2\text{H}_6]$ DMSO, pH 6.5, at 298 (a), 308 (b), 318 (c) and 328 K (d)

shift of the Arg⁴N^εH resonance upon neutralization of the peptide solution, and the fact that the Arg⁴N^εH₄ signal of the hexapeptide showed a similar NMR pattern [Fig. 1(b)] to the Arg⁴N^εH₄ of the L-Arg-L-Val-Gly tripeptide, are evidence for an intramolecular interaction between Arg⁴ guanidinium and Gly⁶ C-terminal carboxylate groups at neutral pH.

The temperature coefficient values of all the NHs of the hexapeptide (Table 1) showed that the backbone amide protons do not participate in any hydrogen bonds and they are exposed to the solvent. In particular, the Val²NH temperature coefficient was not determined due to significant broadening of the NH signal even in the range from 303 to 308 K. Further confirmation that the amide protons in the peptide sequence are exposed, was obtained from the following observation. The couplings with the amide protons decreased gradually, when the temperature of the peptide solution increased. Consequently, the C^αH signals detected at higher temperatures were of lower splitting (multiplicity) due to couplings with the C^βH protons. For instance, the quadruplet resonance peak of Val⁵C^αH at 298 K appeared as triplet at 308 K, broad peak at 318 K and doublet at 328 K, whereas the ABX system of Gly³ at 298 K was changed to an AB system (quadruplet signal) at 328 K (Fig. 4). A unique exception constitutes residue Gly⁶, which was detected as a well resolved ABX system even at 328 K. In fact the described loss of coupling between NH and C^αH protons due to the increase of temperature, could possibly result from a very fast exchange of the amide protons with residual water, whereas the Gly⁶NH was strongly solvated by DMSO and thus protected against the proton exchange with those of the water residual molecules.

One may also emphasize that the side chain of the Arg¹ residue in the hexapeptide sequence does not participate in any intramolecular interaction, as we can reasonably assume from the absence of ROE connectivities between protons of the Arg¹ side chain and the fact that almost all methylene protons appeared magnetically equivalent at neutral pH [Figs. 1(b), 2 and Table 1]. Thus, an obvious conclusion is that differences in the conformational properties of arginine moieties in a peptide sequence can prove advantageous in differentiating proton assignment of guanidyl side chains.

In the proposed guanidinium-carboxylate interaction, the Arg⁴N^εH and Arg⁴N^δH₂ groups may form two parallel (type B) or convergent (type D) hydrogen bonds with the carboxylate moiety, as was also postulated for the L-Arg-X-Gly tripeptide.^{12,20}

Our study on the arginine-containing hexapeptide pointed out that the guanidinium charged group of arginine at a position near to the carboxy-terminus of the peptide sequence, preferentially interacts with the C-terminal carboxylate group, of opposite sign, at neutral pH. It thus confirmed a particular tendency of the protonated side chain of arginine to interact with the C-terminal carboxylate group in sequences longer than three residues.

Conclusions

The NMR results presented here show that the guanidinium group (Arg⁴) at the third position from the carboxy-terminus of a peptide sequence with six residues (L-Arg-L-Val-Gly-L-Arg-L-Val-Gly) has a pronounced tendency to interact with the C-terminal carboxylate group in DMSO solution at pH 6.5. In fact, the considerable downfield shift of the Arg⁴N^εH signal upon neutralization and the magnetically non-equivalent methylene protons of the same residue at pH 6.5 further confirm our previous observations for a non-covalent hydrogen-bonding interaction between guanidinium and C-terminal carboxylate groups in the arginine-containing tripeptides L-Arg-X-Gly (X = L-Ala, Val, Leu) in DMSO solution at neutral pH. Besides that, the simultaneous presence of a second guanidinium group (Arg¹) at the amino terminus of the peptide sequence, which does not participate in any intramolecular interaction, allows the establishment of a general probe for the study of Arg side-chain mobility and accessibility in hydrogen bonding interactions in peptides and proteins.

The intense backbone-backbone NOE connectivities attested for the L-Arg⁴-L-Val⁵-Gly⁶ C-terminal sequence, the NOE

connectivities of Arg⁴ methylene protons and the ABX system of Gly⁶ moiety point to a local rigid conformation around this part of the molecule and provide further support for an intramolecular interaction between Arg⁴ guanidinium and Gly⁶ C-terminal carboxylate groups at neutral pH.

References

- 1 D. K. Eggleston and D. J. Hodgson, *Int. J. Peptide Protein Res.*, 1985, **25**, 242.
- 2 P. Y. Chou and G. D. Fasman, *J. Mol. Biol.*, 1977, **115**, 135.
- 3 H. C. Marsh, Y. C. Meinwald, T. W. Thannhauser and H. A. Scheraga, *Biochemistry*, 1983, **22**, 4174.
- 4 R. Mayer and G. Lancelot, *J. Am. Chem. Soc.*, 1981, **103**, 4738.
- 5 S. Loukas, D. Varoucha, C. Zioudrou, R. A. Steaty and W. A. Klee, *Biochemistry*, 1983, **22**, 4567.
- 6 M. T. Cung, P. Demange, M. Marraud, V. Tsikaris, C. Sakarellos, I. Papadouli, A. Kokla and S. J. Tzartos, *Biopolymers*, 1991, **31**, 769.
- 7 K. R. Shoemaker, P. S. Kim, D. N. Brems, S. Marqusee, E. J. York, I. M. Chaiken, J. M. Stewart and R. L. Baldwin, *Proc. Natl. Acad. Sci. USA*, 1985, **82**, 2349.
- 8 K. R. Shoemaker, P. S. Kim, E. J. York, J. M. Stewart and R. L. Baldwin, *Nature*, 1987, **326**, 563.
- 9 M. D. Pierschbacher and E. Ruoslahti, *Nature (London)*, 1984, **309**, 30.
- 10 R. Rytelka, M. D. Pierschbacher, M. H. Ginsberg, E. F. Plow and E. Ruoslahti, *Science*, 1986, **231**, 1559.
- 11 D. E. Eggleston and S. H. Feldman, *Int. J. Peptide Protein Res.*, 1990, **36**, 161.
- 12 V. Tsikaris, M. Sakarellos-Daitsiotis, E. Panou-Pomonis, E. Detsikas, C. Sakarellos, M. T. Cung and M. Marraud, *Peptide Research*, 1992, **5**, 110.
- 13 E. Panou-Pomonis, C. Sakarellos and M. Sakarellos-Daitsiotis, *Biopolymers*, 1986, **25**, 655.
- 14 D. Neuhaus and J. Keeler, *J. Magn. Reson.*, 1986, **68**, 568.
- 15 A. Bax and G. Davis, *J. Magn. Reson.*, 1985, **65**, 355.
- 16 V. Tsikaris, M. Sakarellos-Daitsiotis, N. Theophanidis, C. Sakarellos, M. T. Cung and M. Marraud, *J. Chem. Soc., Perkin Trans. 2*, 1991, 1353.
- 17 C. D. Bello, L. Gozzini, M. Tonellato, M. G. Corradini, G. D'Auria, L. Paolillo and E. Trivellone, *Bipolymers*, 1989, **28**, 421.
- 18 M. J. Bogusky, A. M. Naylor, S. M. Pitzengerger, R. F. Nutt, S. F. Brady, C. D. Colton, J. T. Sisko, P. S. Anderson and D. F. Veper, *Int. J. Pept. Protein Res.*, 1992, **39**, 63.
- 19 M. Sheinblatt, *J. Am. Chem. Soc.*, 1966, **88**, 2845.
- 20 D. M. Salunke and M. Vijayan, *Int. J. Pept. Protein Res.*, 1981, **18**, 348.

Paper 3/00623I

Received 1st February 1993

Accepted 23rd March 1993