(5) A. Bondi, *J. Phys. Chem.*, **68**, 441 (1964).
(6) E. F. Meyer, T. A. Renner, and K. S. Stec, *J. Phys. Chem.*, **75**, 642 (1971).
(7) E. F. Meyer, K. S. Stec, and R. D. Hotz, *J. Phys. Chem.*, **77**, 2140 (1973).
(8) D. E. Martire and P. Riedl, *J. Phys. Chem.*, **72**, 3478 (1968).
(9) M. J. Kamlet, private communication.
(10) S. Wasik, private communication.
(11) For the purposes of this plot, each Cl atom is counted as a carbon atom; see ref 6.

Conformational Analysis of Protected Norvaline Oligopeptides by High Resolution Proton Magnetic Resonance

E. S. Pysh*[1a] and C. Toniolo[1b]

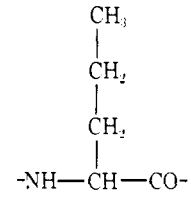
Contribution from the Department of Chemistry, Brown University, Providence, Rhode Island 02912 and the Biopolymer Research Center, C.N.R., Institute of Organic Chemistry, University of Padova, 35100 Padova, Italy. *Received February 15, 1977*

Abstract: Proton magnetic resonance spectra at 270 MHz are reported for Boc(L-Nva)$_n$OMe, n = 1-4, and for Ac-L-Nva-OMe, in $CDCl_3$ and $CDCl_3/Me_2SO$-d_6 mixtures. A method is reported for assigning the C$^\alpha$H resonance at the C-terminal residue of ester C-protected peptides, based on the downfield shift of that resonance in $CDCl_3$ with increasing temperature. No folded forms are found for the unaggregated dipeptide (0.009 M), but mixtures of several folded forms are found at low concentrations of the tripeptide (0.007 M) and tetrapeptide (0.005 M). At higher concentrations aggregates form, with increasing ease as the peptide chain lengthens. Whereas the dipeptide and tripeptide aggregates are best described as a network of extended "out-of-register" forms, the tetrapeptide associations are "in-register", i.e., primitive β sheets. These conclusions correlate with the results of previous circular dichroism and infrared studies.

Previous work[2-7] has shown that homo-oligoheptapeptides with alkyl, arylalkyl, and thioether-containing side chains take up the β conformation both in structure-supporting solvents and in the solid state. Of relevant interest has been the finding that the peptide chains generally align in an antiparallel fashion in forming the intermolecularly H-bonded β sheet when the amino acid side chains are small, but the presence of bulky, aromatic, or β-branched side chains leads to an increased tendency toward parallel alignment. In particular, in the case of the norvaline oligomers it was concluded that some of both the parallel and antiparallel arrangements are present in the β sheet which forms.[3-8] In solution, as the peptide concentration is lowered or the temperature is increased, nonaggregated forms become dominant.

High resolution ^{1}H NMR[9] has come to be a valuable and widely used technique in the area of conformational analysis of homo-oligopeptides in solution. Goodman and co-workers have studied by ^{1}H NMR several series of synthetic monodisperse protected peptides including those derived from γ-alkyl esters of glutamic acid,[10-12] alanine,[13-16] and isoleucine.[13,15] Bystrov et al.[17] examined in detail diastereomeric protected alanine dipeptides, mainly in CCl_4, and Dimicoli and Ptak[18] investigated a protected alanine dipeptide in $CDCl_3$.

This work is part of a study devoted to the investigation by high resolution ^{1}H NMR of the role of the side chains of saturated hydrocarbon amino acid residues on the formation of ordered secondary structures in protected homo-L-oligopeptides in solvents of low polarity. The number of carbon atoms, shape of the side chain (whether linear or branched), and position of branching will be considered. In this paper we wish to discuss the results obtained in the case of the synthetic monodisperse homologous series Boc(L-Nva)$_n$OMe where n = 1-4. The pentapeptide was not examined because its solubility in $CDCl_3$ is too low for a ^{1}H NMR study. Norvaline was investigated because it is the simplest amino acid residue having a δ carbon in the side chain. In this context, of particular interest is the question of the detailed structure of any nonassociated intramolecularly H-bonded folded forms which occur, and also the nature of the aggregated forms which develop at higher peptide concentrations.



norvaline

The CD of an internal Nva-Nva peptide chromophore within a randomly coiled chain in 1,1,1,3,3,3-hexafluoro-2-propanol has been determined by subtracting the total molar ellipticity values of the tripeptide from those of the tetrapeptide.[19] The CD properties of the dipeptide *in* aqueous solution[8] have represented the first experimental verification of Mattice's prediction[20] of the occurrence of a temperature and salt-sensitive dichroic band near 210-215 nm in aqueous solutions of linear randomly coiled peptides with saturated hydrocarbon side chains.

Experimental Section

The details of the synthesis and the chemical and optical characterizations of Boc(L-Nva)$_{2-4}$OMe are reported in ref 21.

Boc-L-Nva-OMe. This compound was prepared from *tert*-butyloxycarbonyl azide and HCl·H-L-Nva-OMe[22] in anhydrous ethyl acetate using *N*-methylmorpholine to deprotonate the ammonium group: yield 75% oil; $[\alpha]^{23}_D = -30.0°$ (c = 1; methanol).

Anal. for $C_{11}H_{21}NO_4$. Calcd: C, 57.1; H, 9.1; N, 6.1. Found: C, 56.5; H, 8.9; N, 6.0.

Ac-L-Nva-OMe.[23] This compound was prepared from acetyl chloride and HCl·H-L-Nva-OMe[22] in anhydrous chloroform using triethylamine to deprotonate the ammonium group: yield 82%; mp 53-54 °C (from ethyl acetate/petroleum ether); $[\alpha]^{23}_D = -57.4°$ (c = 2; water).

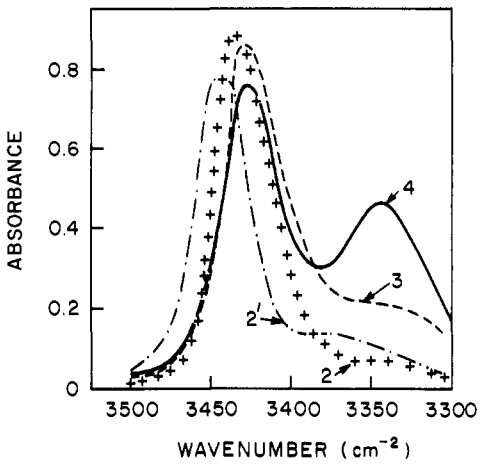


Figure 1. Infrared absorption spectra in deuteriochloroform solution of Boc(L-Nva)$_2$OMe (5.95×10^{-4} mol/L) (2), Boc(L-Nva)$_3$OMe (5.57×10^{-4} mol/L) (3), and Boc(L-Nva)$_4$OMe (3.41×10^{-4} mol/L) (4), and in carbon tetrachloride solution of Boc(L-Nva)$_2$OMe (6.03×10^{-4} mol/L) (2'). The band in the 3380-3340-cm^{-1} region corresponds to the intramolecularly H-bonded N-H vibration. The spectra of trimer and tetramer were taken from ref 4.

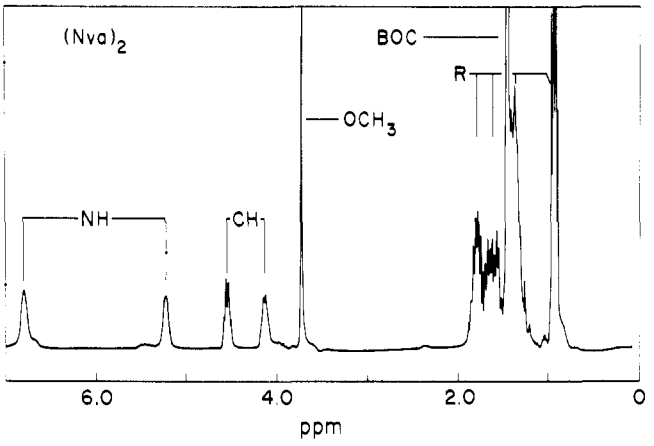


Figure 2. Complete ^{1}H NMR spectrum of Boc(L-Nva)$_2$OMe in $CDCl_3$ (0.48 M) at 270 MHz.

Chloroform (99.8% D) and Me_2SO (99.9% D) were purchased from the Wilmad Co., Buena, N.J. Weighed amounts of peptide were placed directly into NMR sample tubes and a known volume of solvent was then added. Measurements were made at the Southern New England High Field NMR Facility, New Haven, Conn. on a Brucker 270 HX spectrometer. At the most dilute concentrations (2.8 mg/mL; 0.28% w/v) 1100 scans were accumulated requiring 25 min. Chloroform was used as the internal standard. The IR absorption measurements were carried out with a Beckman Model IR-9 spectrophotometer using demountable 10 cm path length cells with calcium fluoride windows. The positions of the bands are accurate to ±1 cm^{-1}.

Results

To obtain preliminary information on the tendency toward intramolecularly H-bonded folded structure formation of Boc(L-Nva)$_{2-4}$OMe in solvents of low polarity we carried out an IR absorption study in the NH stretching region in very dilute $CDCl_3$ solutions.[4] The results indicate that the dipeptide contains only a small amount of intramolecularly H-bonded folded forms (less than 8%), the tripeptide a larger amount (approximately 21%), but that in the tetrapeptide a considerable percentage (approximately 38%) of the molecules adopt such forms (Figure 1). In an apolar solvent (CCl_4) the portion of intramolecularly H-bonded folded forms increases (approximately 16% for the dipeptide).

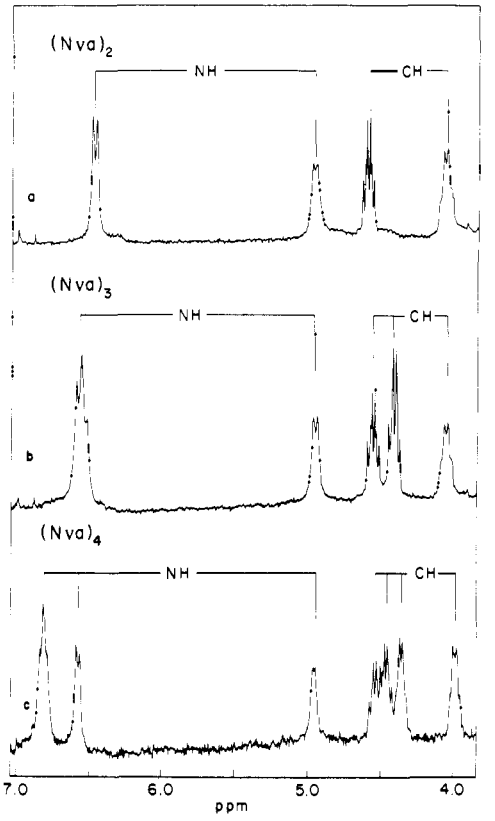


Figure 3. (a) Partial ^{1}H NMR spectrum of Boc(L-Nva)$_2$OMe in $CDCl_3$ (0.009 M) at 270 MHz. (b) Partial ^{1}H NMR spectrum of Boc(L-Nva)$_3$OMe in $CDCl_3$ (0.007 M) at 270 MHz. (c) Partial ^{1}H NMR spectrum of Boc(L-Nva)$_4$OMe in $CDCl_3$ (0.005 M) at 270 MHz.

A clearer picture of the conformational preferences of the norvaline oligopeptides was obtained using high resolution ^{1}H NMR. We number the amino acid residues from the amino end of the peptide chain, so that the two protons attached to the nitrogen and α-carbon atoms of the amino terminus residue are labeled N(1)H and C(1)H, and so on. In this study attention is focused upon the NH and CH resonances which are well separated from all other proton resonances. This feature is illustrated in Figure 2 which shows the complete ^{1}H NMR spectrum of the dipeptide at a concentration of 0.48 M (in $CDCl_3$). The protons of the blocking groups and of the side chain are always found at higher field than the NH and C$^\alpha$H protons.

Figure 3 displays representative spectra in the NH and C$^\alpha$H region in dilute $CDCl_3$ solutions for the dipeptide (0.009 M), tripeptide (0.007 M), and tetrapeptide (0.005 M). This figure illustrates that the broadest NH and C$^\alpha$H signals are those at highest field.

Assignment of the high field NH resonance to the Boc-protected amino terminus residue is based on the well-established observation that the presence of a urethane group, as in ROCONHR, shifts the NH signal to considerably higher fields compared with RCONHR.[18,24-26] Additional confirmation of this effect is indicated in the spectra of Ac-L-Nva-OMe and Boc-L-Nva-OMe (Figure 4) which also show the effect of the Boc protecting group. The resonance at 5.0 ppm in Figure 3a, for example, is therefore assigned to N(1)H.

Application of proton-proton double resonance experiments allows the C$^\alpha$H signals of the dipeptide also to be assigned; irradiation of the downfield C$^\alpha$H signal causes collapse of the N(2)H resonance and irradiation of the upfield C$^\alpha$H signal causes collapse of the N(1)H resonance. These results agree well with those of Dimicoli and Ptak on Boc(L-Ala)$_2$-OMe.[18]

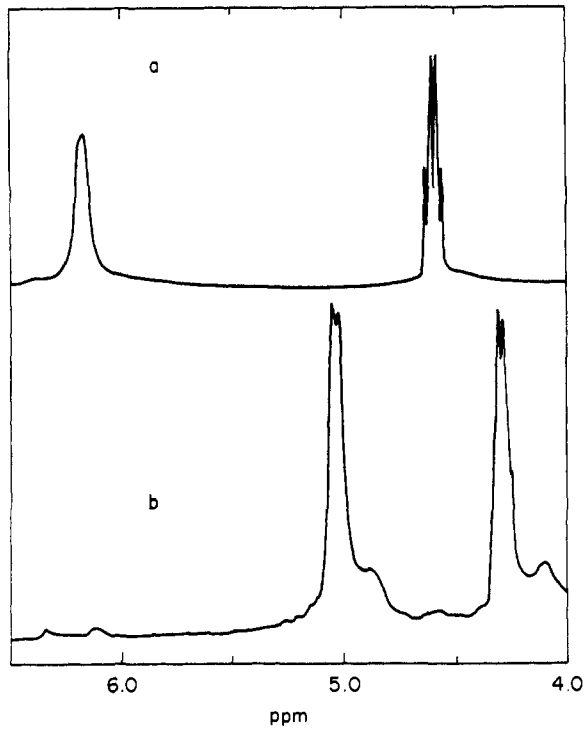


Figure 4. (a) Partial ^{1}H NMR spectrum of Ac-L-Nva-OMe in $CDCl_3$ (0.3 M). (b) Partial ^{1}H NMR spectrum of Boc-L-Nva-OMe in $CDCl_3$ (0.3 M).

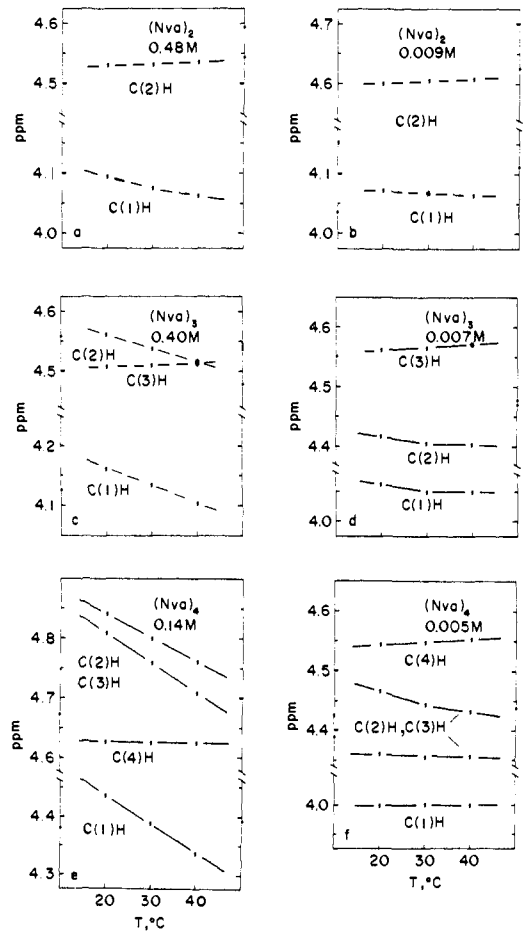


Figure 5. Temperature dependence of C$^\alpha$H resonances of Boc(L-Nva)$_2$OMe in $CDCl_3$: (a) 0.48 M; (b) 0.009 M. Temperature dependence of C$^\alpha$H resonances of Boc(L-Nva)$_3$OMe in $CDCl_3$: (c) 0.40 M; (d) 0.007 M. Temperature dependence of C$^\alpha$H resonances of Boc(L-Nva)$_4$OMe in $CDCl_3$: (e) 0.14 M; (f) 0.005 M.

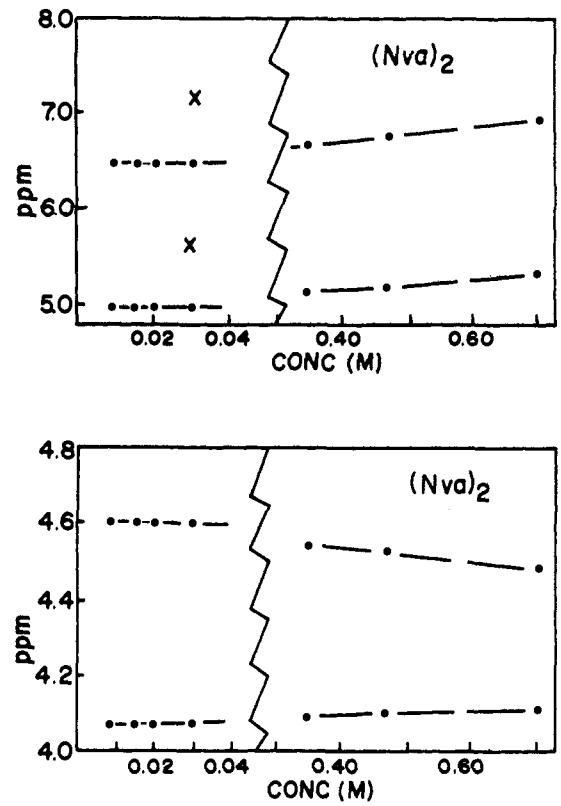


Figure 6. Concentration dependence of NH (above) and C$^\alpha$H (below) chemical shifts of Boc(L-Nva)$_2$OMe in $CDCl_3$: (×) $CDCl_3$/Me_2SO (50:50 vol).

We proceeded in the assignment of resonances in the tripeptide and tetrapeptide using the characteristic and unusual temperature dependence of the resonance of the C$^\alpha$H at the methyl ester end of the chain. In both concentrated and dilute $CDCl_3$ solutions the C(2)H resonance of the dipeptide shifts *downfield* as the temperature is increased (Figure 5). The C(1)H resonance shifts upfield as do both NH resonances (see below). The temperature dependence of the C(2)H resonance in the dipeptide is, in fact, the same as that of the C$^\alpha$H resonance in Ac-L-Nva-OMe, which we measured to be 0.0004 ppm/deg downfield with increasing temperature.

One of the three C$^\alpha$H resonances of the tripeptide shifts downfield with increasing temperature (Figure 5) and this resonance we assign to C(3)H by analogy with the behavior of C(2)H of the dipeptide and the C$^\alpha$H of Ac-L-Nva-OMe. Likewise only one C$^\alpha$H resonance of the tetrapeptide shifts downfield and we assign it to C(4)H (Figure 5). Decoupling experiments allow the resonance of the NH at the methyl ester end of the chain to be assigned in each case. N(2)H and C(2)H of the tripeptide are then assigned by elimination.

We cannot assign the N(2)H and N(3)H resonances of the tetrapeptide unambiguously, nor the C(2)H and C(3)H resonances. In Figure 3c, for example, we can only tell that the resonance at 6.6 ppm is either N(2)H or N(3)H. We can also tell from double resonance experiments that it is coupled to the C$^\alpha$H resonance at 4.45 ppm. The effect of this uncertainty in the assignments on our conclusions is marginal (see below); the main conclusions are not affected at all.

The concentration dependence of the NH signals and the C$^\alpha$H signals is shown in Figures 6-8 for the dipeptide, tripeptide, and tetrapeptide, respectively. Most of our conclusions are based on the behavior of the NH resonances; this behavior is summarized in Table I. The concentration dependence of the resonance of the C$^\alpha$H at the methyl ester end of the chain is unique, as is its temperature dependence, in that it shifts in

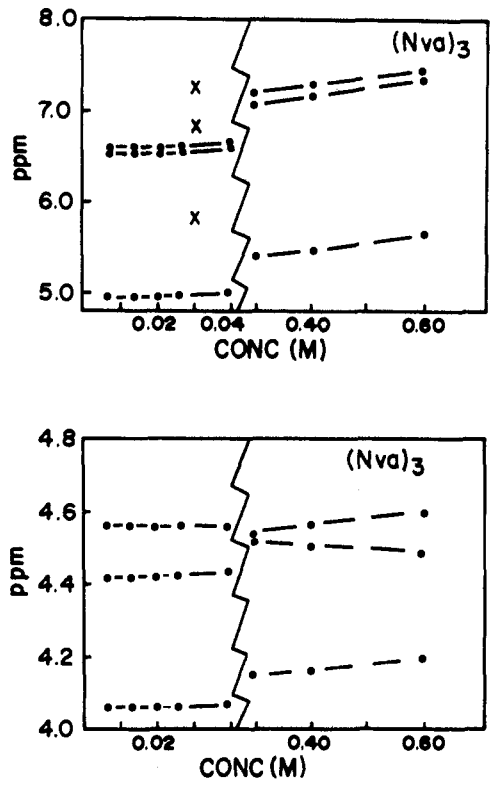


Figure 7. Concentration dependence of NH (above) and C$^{\alpha}$H (below) chemical shifts of Boc(L-Nva)$_3$OMe in $CDCl_3$: (×) $CDCl_3$/Me_2SO (50:50 vol).

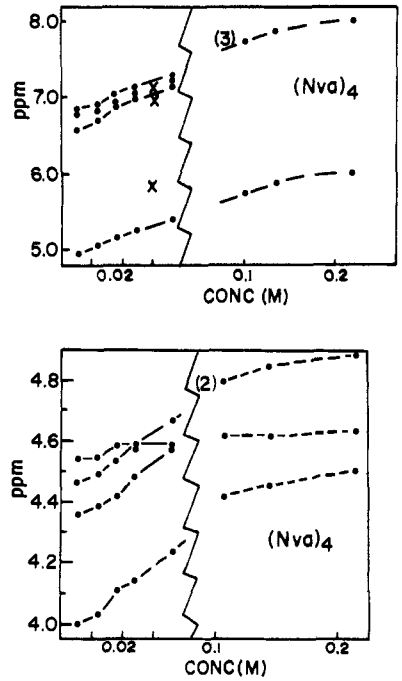


Figure 8. Concentration dependence of NH (above) and C$^{\alpha}$H (below) chemical shifts of Boc(L-Nva)$_4$OMe in $CDCl_3$: (×) $CDCl_3$/Me_2SO (50:50 vol).

the opposite direction compared with all other resonances, thus confirming our assignment of it.

The temperature dependence of the NH resonances is illustrated in Figure 9, and the slopes of the graphs, in units of ppm/deg, are shown in Table II.

When Me_2SO is added to the $CDCl_3$ (50/50% vol) there are changes in chemical shift, temperature dependence, and

Table I. Chemical Shifts in $CDCl_3$ (in ppm) of NH Peaks Extrapolated to Zero Concentration at 20 °C[a]

	N(1)H	N(2)H	N(3)H	N(4)H
Boc(Nva)$_2$OMe	5.0(0.2$_5$)	6.5(0.3)		
Boc(Nva)$_3$OMe	5.0(0.4$_5$)	6.5(0.6)	6.6(0.6$_5$)	
Boc(Nva)$_4$OMe	5.0(1.0)	6.8(1.2)[b]	6.6(1.4)[b]	6.8(1.2)

[a] In parentheses, the increase in chemical shift of NH peaks as concentration is increased to 0.48 M in the dimer, 0.30 M in the trimer, or 0.22 M in the tetramer; i.e., to an approximately uniform monomer concentration of 1 M. [b] In the tetrapeptide N(2)H and N(3)H cannot be distinguished unambiguously; see text.

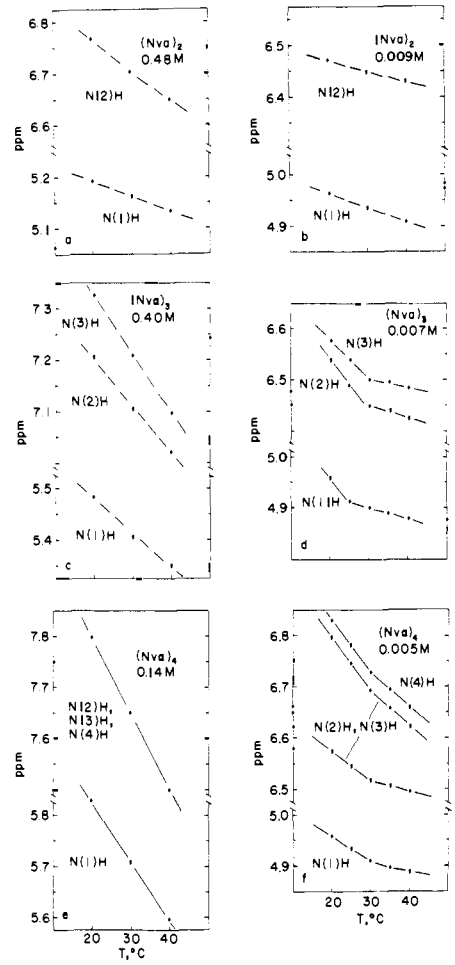


Figure 9. Temperature dependence of NH chemical shifts of Boc(L-Nva)$_2$OMe in $CDCl_3$: (a) 0.48 M; (b) 0.009 M. Temperature dependence of NH chemical shifts of Boc(L-Nva)$_3$OMe in $CDCl_3$: (c) 0.40 M; (d) 0.007 M. Temperature dependence of NH chemical shifts of Boc(L-Nva)$_4$OMe in $CDCl_3$: (e) 0.14 M; (f) 0.005 M.

spectral resolution. These changes were measured at a peptide concentration of 0.03 M. The × symbols in Figures 6-8 show the downfield shift in the NH resonances relative to their positions in $CDCl_3$. Decoupling experiments were repeated in the mixed solvent system as necessary to confirm assignments. In the presence of Me_2SO there is a general narrowing of the signals; it is especially marked in the N(1)H doublet of the dipeptide as illustrated in Figure 10. The temperature dependence of the NH resonances in the presence of Me_2SO is summarized in Table III.

The temperature dependencies of the NH and C$^{\alpha}$H chemical shifts in the model compounds Ac-L-Nva-OMe and Boc-L-Nva-OMe are shown in Figures 11 ($CDCl_3$/Me_2SO) and 12 ($CDCl_3$).

Table II. Temperature Dependence of Chemical Shift (in ppm/deg) of NH Resonances in $CDCl_3$[a]

	N(1)H	N(2)H	N(3)H	N(4)H
Boc(Nva)$_2$OMe				
0.009 M	0.002	0.002		
0.48 M	0.003	0.006		
Boc(Nva)$_3$OMe				
0.007 M	0.006	0.009	0.008	
	(0.002)	(0.002)	(0.002)	
0.40 M	0.007	0.009	0.011	
Boc(Nva)$_4$OMe				
0.005 M	0.005	0.005	0.011	0.011
	(0.002)	(0.003)[b]	(0.005)[b]	(0.005)
0.14 M	0.012	0.015	0.015	0.015

[a] Where two figures appear, the first refers to the temperature range 20–30 °C and the figure in parentheses refers to the range 30–40 °C. [b] In the tetrapeptide N(2)H and N(3)H cannot be distinguished unambiguously; see text.

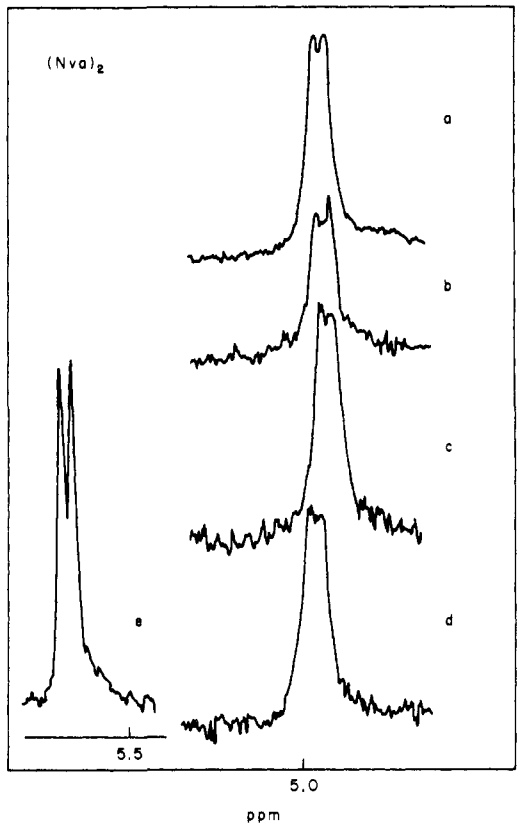


Figure 10. N(1)H resonance of Boc(L-Nva)$_2$OMe in $CDCl_3$ at 0.009 M (a), 0.015 M (b), 0.02 M (c), and 0.03 M (d); and in $CDCl_3/Me_2SO$ (50:50 vol) at 0.03 M.

Discussion

Assignment of the high field NH resonance to the N-terminal amino acid residue of urethane N-protected peptides is now well established in the literature.[17,18] Our method of assigning the resonance of the C$^\alpha$H at the methyl ester end of the chain in the tripeptide and tetrapeptide has, however, apparently not been used before. Since the downfield shift of that resonance with increasing temperature appears to be unique in $CDCl_3$ solutions, we propose that the resonance can be identified on the basis of such a shift and perhaps even in different solvent systems. The approach used here is useful in that, together with standard decoupling experiments, it allows an unambiguous assignment of the NH and C$^\alpha$H resonances of the amino acid residues at both ends of urethane N-protected and ester C-protected oligopeptides. Once the proper assignments of all resonances are at hand, a picture of the peptide conformational states can be constructed on the basis of the concentration dependence, temperature dependence, and solvent dependence of the chemical shifts of the NH and C$^\alpha$H resonances.

The downfield shift of the NH resonances upon increasing

Table III. Chemical Shift in $CDCl_3/Me_2SO$[a] (in ppm) of NH and C$^\alpha$H resonances[b]

	N(1)H	N(2)H	N(3)H	N(4)H
Boc(Nva)$_2$OMe	5.7	7.2		
	(0.004)	(0.0035)		
Boc(Nva)$_3$OMe	5.8	6.9	7.3	
	(0.0055)	(0.001)	(0.004)	
Boc(Nva)$_4$OMe	5.8	7.0	7.0	7.1
	(0.0055)	(0.002)	(0.002)	(0.003)
	C(1)	C(2)H	C(3)H	C(4)H
Boc(Nva)$_2$OMe	3.3	3.6		
Boc(Nva)$_3$OMe	3.2	3.5	3.6	
Boc(Nva)$_4$OMe	3.2	3.5	3.5	3.5

[a] 50:50 by volume. [b] Peptide concentration, 0.03 M. The temperature dependence of the NH resonances is shown in parentheses.

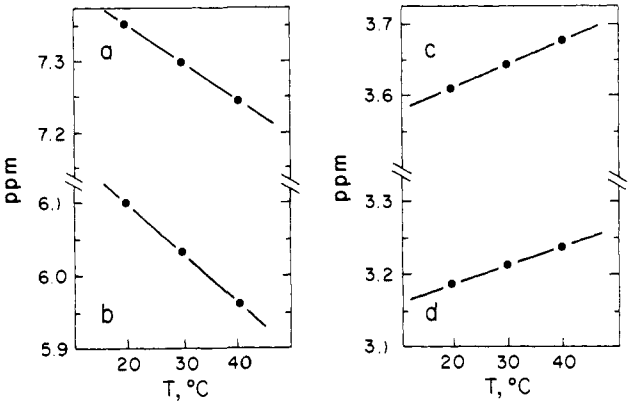


Figure 11. Temperature dependence of chemical shifts in $CDCl_3/Me_2SO$-d_6 (50:50 vol): (a) NH resonance of Ac-L-Nva-OMe, 0.3 M; (b) NH resonance of Boc-L-Nva-OMe, 0.3 M; (c) C$^\alpha$H resonance of Ac-L-Nva-OMe, 0.3 M; (d) C$^\alpha$H resonance of Boc-L-Nva-OMe, 0.3 M.

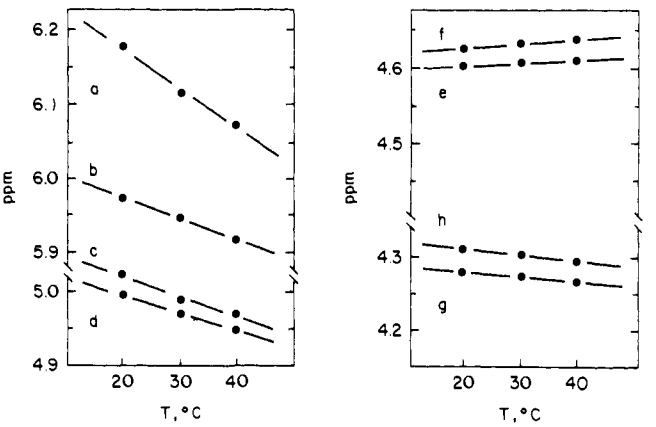


Figure 12. Temperature dependence of chemical shifts in $CDCl_3$: (a) NH resonance of Ac-L-Nva-OMe, 0.3 M; (b) NH resonance of Ac-L-Nva-OMe, 0.03 M; (c) NH resonance of Boc-L-Nva-OMe, 0.3 M; (d) NH resonance of Boc-L-Nva-OMe, 0.03 M; (e) C$^\alpha$H resonance of Ac-L-Nva-OMe, 0.3 M; (f) C$^\alpha$H resonance of Ac-L-Nva-OMe, 0.03 M; (g) C$^\alpha$H resonance of Boc-L-Nva-OMe, 0.3 M; (h) C$^\alpha$H resonance of Boc-L-Nva-OMe, 0.03 M.

the concentration is evidence of intermolecular H-bond formation.[27] As one measure of the strength of concentration dependence, the chemical shift at the lowest concentration measured can be compared with the chemical shift at some other specified concentration, which we define in Table I as 1.0 M on a monomer basis. A common concentration on a monomer basis is useful for purposes of comparing one peptide with another. The concentration dependence is greatest in the tetrapeptide and smallest in the dipeptide. Also indicating the strong aggregation tendency of the tetrapeptide is the fact that its NH chemical shifts remain concentration dependent down to a concentration of 0.005 M, whereas in the dipeptide and tripeptide there is no further dependence below approximately 0.03 M (Figures 6-8).

The temperature dependence results are summarized in Table II. The NH chemical shifts of the dipeptide at low concentration (0.009 M) vary only slightly with temperature (0.002 ppm/deg), as is expected in the absence of association (Figure 6) and in the presence of only a very small amount of folded forms (Figure 1). At higher dipeptide concentrations (0.48 M) there is an increase in the temperature dependence of the NH chemical shifts, reflecting more extensive association which melts out as the temperature is increased.

In the tripeptide, there is a larger temperature dependence in the chemical shifts at 0.40 M concentration than in those of the dipeptide at 0.48 M. This is a further indication that ease of association increases with chain length.

At low concentrations of tripeptide (0.007 M) however, the temperature dependence of NH shifts is more complex. Between 30 and 40 °C the temperature dependence is very small, like that of the unassociated dipeptide; but between 20 and 30 °C the change in chemical shift is substantial, and almost as large as in the concentrated solution. Since the lack of a concentration dependence in NH chemical shift in the 0.007 M range indicates the absence of aggregates, this large temperature dependence must reflect some process other than dissociation. We interpret the large temperature dependence as indicating the presence of intramolecularly H-bonded folded forms at 20 °C. The decreased temperature dependence at higher temperatures we take to indicate that these intramolecular H bonds are, at that point, completely thermally disrupted. Such a conclusion is consistent with the results of the infrared absorption study (Figure 1) in which we ascertained that at room temperature approximately 21% of the tripeptide is intramolecularly H bonded in very dilute $CDCl_3$ solutions.

In Figure 9d, straight lines having sharp breaks are shown, rather than gradually decreasing curves. The expected precision of the data justifies the presentation shown in Figure 9, but for the interpretative model we present here no sharp breaks in the data would be expected. Therefore, either the breaks are real and the temperature dependence in the tripeptide requires a more complex model than the melting of folded forms, or the model is correct and the chemical shifts decrease smoothly and gradually, implying that the actual precision of the data is slightly less than that expected.

The tetrapeptide at a concentration of 0.005 M similarly shows a small temperature dependence between 30 and 40 °C, but a larger dependence between 20 and 30 °C. This result partially reflects the unfolding of intramolecularly H-bonded forms which may be nearly complete around 30 °C. The concentration dependence curve for the tetrapeptide in the 0.005 M range is not flat (Figure 8), indicating that there is still some association at that low concentration. However, the temperature dependence is so large (Table II) that any residual association at that concentration can probably not account for the entire temperature dependence, and at least some of the temperature dependence likely comes from the presence of folded forms at 20 °C. Nonassociated intramolecularly H-bonded folded forms (38%) were found in this oligopeptide at room temperature in a 10^{-4} M $CDCl_3$ solution using infrared absorption (Figure 1).

The extremely large temperature dependence in all four NH chemical shifts of the tetrapeptide at 0.14 M concentration again suggests increasing association as the chain length increases.

Thus, from a combination of the concentration dependence and temperature dependence comes the picture: (a) of intermolecularly H-bonded structures forming with increasing ease as the peptide chain lengthens, (b) of these associated forms being disrupted as the temperature is increased or the concentration decreased, and (c) of intramolecularly H-bonded folded forms of the tripeptide and tetrapeptide (but not the dipeptide) being stable at 20 °C. We emphasize that these conclusions apply only to chloroform solutions.

The results obtained in the mixed solvent system $CDCl_3$/Me_2SO-D_6 (Table III) are also consistent with this picture. In the cases of the dipeptide and tripeptide, in which there is no significant association in $CDCl_3$ at 0.03 M, the strong H-bonding acceptor properties of Me_2SO lead to marked downfield shifts of all NH resonances. In the tetrapeptide at a concentration of 0.03 M the structure breaking properties of Me_2SO result in replacing a number of intermolecular peptide-peptide H bonds existing in $CDCl_3$ solutions with strong peptide-solvent H bonds, the net effect being only a rather small change in chemical shift of the NH resonances (Figure 8). The final conformational states in 50% Me_2SO are the same for all three peptides as evidenced by the similarity in chemical shifts of the NH resonances: 7.1-7.3 ppm for the NH group at the methyl ester end of the chain, 5.7-5.8 ppm for the NH group at the urethane end of the chain, and 6.9-7.0 ppm for the internal NH groups. The $C^{\alpha}H$ resonances all move upfield in 50% Me_2SO and there are characteristic chemical shifts for each position in the chain (Table III).

It was of special interest to attempt to determine the detailed structure of both the intermolecularly H-bonded forms and, for the tripeptide and tetrapeptide, the nonassociated intramolecularly H-bonded folded forms. In both cases we found the results of the temperature dependence experiments most useful. In 50% Me_2SO (Table III) the NH resonance for each position of the peptide chain has its characteristic temperature dependence, which reflects the stability of the peptide-solvent H bond in which that NH group participates. Thus, the NH group at the urethane N-protected end displays the greatest temperature dependence (0.004-0.0055 ppm/deg), the NH group at the carboxyl end an intermediate value (0.003-0.004 ppm/deg), and the internal NH groups the smallest values (0.001-0.003 ppm/deg).

In pure chloroform, however, all NH resonances show the same temperature dependence (0.002 ppm/deg) when the peptide is both unassociated and unfolded, e.g., the dipeptide at 0.009 M, and the tripeptide at 0.007 M at 30-40 °C (Table II). The slightly larger value for the tetrapeptide at 0.005 M concentration indicates that there is some residual amount of associated and/or folded forms at 30-40 °C. This finding, that the NH resonances at all positions of the peptide chain show the same temperature dependence when the peptide is neither aggregated nor folded, is very useful since it allows the hypothesis to be entertained that the *differences* which appear in the associated and folded forms are related to the degree of participation by each NH group in the H-bonds stabilizing those forms. We cannot, however, be sure that this is the case because we cannot determine what the absolute chemical shifts are of the completely (e.g., intermolecularly) hydrogen bonded NH groups. For example, consider the difference between the chemical shift for a totally (e.g., intermolecularly) hydrogen bonded NH group and that for the same group totally solvated. Only if that difference is known to be the same for two NH

groups could any difference in temperature dependence rigorously be attributed to different degrees of hydrogen bond participation.

In the aggregated dipeptide (e.g., at 0.48 M) the associations are probably not associations of folded forms since folded forms of the dipeptide were not found at dilute concentrations. The associations are therefore likely to be made up of extended forms, either "in-register" or "out-of-register" forms. The temperature dependence of N(2)H is twice that of N(1)H, which itself is only slightly larger than the value for the unassociated dipeptide (Figure 9 and Table II). This result likely reflects the production of unassociated dipeptides, and the small temperature dependence of the N(1)H chemical shift reflects a low degree of participation in the hydrogen bond network which stabilizes the aggregates.

Figure 13 illustrates the four possible "in-register" structures for associated dipeptides. Figures 13a and 13b show antiparallel chains, the other two parallel. The structure 13a has only N(2)H hydrogen bonds, 13b only N(1)H bonds. Our experimental results could be explained if the structure of Figure 13a is substantially more probable than 13b. Dimeric association of extended forms of the type shown in Figure 13a was found in CCl_4 solutions by Cung et al.[28,29] in the case of *N*-acylnorvaline monoalkyl amides. The temperature dependence of N(1)H is only slightly greater than in very dilute solution, which in itself is evidence that the group could be practically free. Its chemical shift, however, is 0.2 ppm greater at 0.48 M than at 0.009 M, indicating some participation in H-bond formation. This participation might be in the form of structures shown in Figures 13b-d.

The dipeptide aggregates might also have "out-of-register" structures, in which two dipeptides have only one H bond in common. If such a network is formed, our data indicate that N(2)H groups hold the network together to a greater extent than do N(1)H groups. Earlier work on the CD and IR properties of these oligopeptides in the solid state[3,4] gave no evidence of β structure formation in the dipeptide. This in itself is not strong evidence against the existence of the "in-register" structures of Figure 13 because end effects could be expected to distort the CD and IR properties of those structures relative to the properties characteristic of a large β structure. On the other hand, the structures of Figure 13 do contain two (antiparallel) or four (parallel) unit cells, and the optical properties, especially the IR, might well have been expected to differ more than they do from the properties characteristic of random chains.

The temperature dependence of the ^1H NMR spectra of Ac-L-Nva-OMe and Boc-L-Nva-OMe is reported (Figures 11 and 12) as corroborative evidence that the difference in temperature dependence between N(1)H and N(2)H in the concentrated dipeptide reflects different extents of participation in H bonding of the two groups rather than intrinsic differences stemming from different electronic environments within the molecule. Figure 11 illustrates the similarity in the ^1H NMR behavior of the two model compounds in $CDCl_3/Me_2SO$, the NH temperature dependence of both compounds being 0.005-0.007 ppm/deg. Figure 12 shows that in $CDCl_3$, the temperature dependence in the NH resonance of Boc-L-Nva-OMe (0.002 ppm/deg) is much lower than in $CDCl_3/Me_2SO$, both at 0.3 and 0.03 M concentration, indicating the absence of solvent-solute H bonds. The temperature dependence of the NH resonance in Ac-L-Nva-OMe, however, remains as high at 0.3 M in $CDCl_3$ as in $CDCl_3/Me_2SO$ (0.005 ppm/deg) showing that in $CDCl_3$, intermolecular peptide-peptide H bonds replace the solvent-solute H bonds that occur in Me_2SO. Thus the ease of H-bond formation by the NH group at the methyl ester end of the chain, relative to the NH group at the Boc N-protected end, is revealed even at the monomer level.

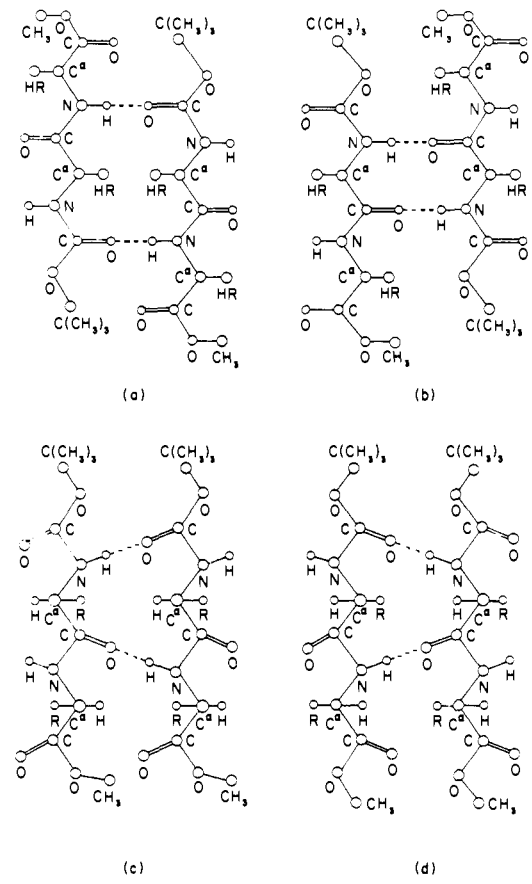


Figure 13. The four possible "in-register" structures for associated dipeptides: (a) and (b) antiparallel chains, (c) and (d) parallel chains. (a) includes only N(2)H H bonds; (b) includes only N(1)H H bonds.

The best picture that can be extracted from the combination of ^1H NMR, CD, and IR properties of the dipeptide aggregates, therefore, is that they consist largely of an "out-of-register" network of extended chains held together largely by N(2)H H bonds, and some portion of "in-register" associated forms of the type shown in Figure 13a.

With respect to the conformation of the dipeptide in dilute $CDCl_3$ solutions, our evidence is (Table II) that the solvent molecules interact sufficiently with the NH groups of the nonaggregated dipeptide to preclude almost completely formation of intramolecularly H-bonded folded structures. The ^1H NMR results are in accord with the IR evidence (Figure 1) by which it is concluded that less than 8% of the dipeptides have intramolecular H bonds. Bystrov et al.[17] found, in the case of a protected alanine dipeptide, that a folded form is stable in the more inert solvent CCl_4 at high dilution. The folded form which they found was a seven-membered ring involving an N(2)H H bonded to the Boc-carbonyl group. Similar results were reported by Cung[29] for *N*-acetyl-L-norvalinemethylamide in CCl_4 at high dilution.

In the case of the tripeptide, the lack of a concentration dependence in chemical shifts in the 0.007 M range (Figure 7) indicates unaggregated forms. The strong temperature dependence in NH chemical shifts from 20 to 30 °C (Figure 9, Table II) provides evidence that these unaggregated forms are folded. A number of folded forms can be taken up by the tripeptide. The N(3)H group can form a 10-membered ring, a β turn, by H bonding to the carbonyl group of the Boc moiety (Figure 14a). This ring is likely to be relatively stable. The N(3)H group can also participate in a seven-membered ring by forming an H bond with the carbonyl group on the residue next to the Boc moiety. This ring structure[17,28,29] can be further stabilized by the formation within the molecule of a second H

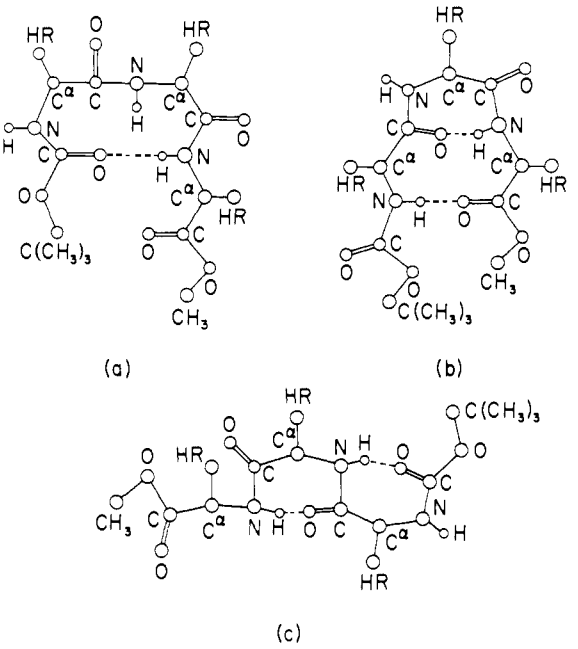


Figure 14. Folded structures of Boc(L-Nva)$_3$OMe: (a) one type of ten-membered ring; (b) seven-membered ring and 11-membered ring; (c) two seven-membered rings.

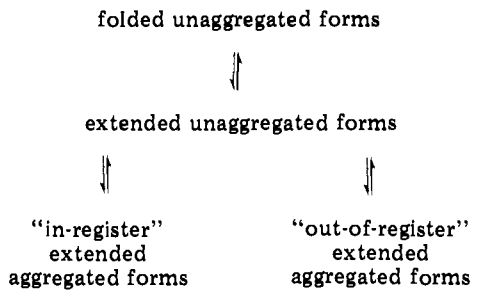
bond, either between N(1)H and the carbonyl of the residue at the methyl ester end of the chain to form an 11-membered ring, or between the N(2)H group and the carbonyl group of the Boc moiety to form a second seven-membered ring within the same structure (Figures 14b,c). The former ring structure, a γ turn, was proposed in 1972 by Némethy and Printz.[30a] The latter structure was found to exist in the solid state in horse heart ferricytochrome *c*.[30b] The N(1)H and N(2)H groups can also form eight-membered rings which are likely to be rather unstable.[31] Since all three NH resonances of the tripeptide show increased temperature dependence between 20 and 30 °C at 0.007 M concentration (Figure 9), there is likely an equilibrium among several of the folded structures just described. The two doubly H-bonded structures shown in Figures 14b and 14c may well compete with the 10-membered ring structure (Figure 14a) which has only one H bond. The presence of those two structures would account for the large temperature dependence observed in all three NH resonances.

These considerations shed light on the *absence* of folded forms in the dipeptide. In the dipeptide only two folded forms are possible. These are a seven-membered ring structure formed by N(2)H H bonding to the carbonyl group of the Boc moiety and an eight-membered ring structure formed by N(1)H H bonding to the carbonyl group at the methyl ester end of the chain. The eight-membered N(1)H ring is not found in the dipeptide just as it is not found in the tripeptide, and no other N(1)H stabilized folded forms are possible in the dipeptide. The N(2)H stabilized seven-membered ring structure is found in the tripeptide probably only because an 11-membered ring or a second seven-membered ring can be formed simultaneously. A seven-membered ring structure by itself, as would be required in the dipeptide, is not very stable in $CDCl_3$. Thus the absence of folded forms in the dipeptide is quite consistent with the presence of folded forms in the tripeptide.

Above 30 °C the temperature dependence of the NH groups of the tripeptide in dilute $CDCl_3$ solutions resembles that of the NH groups of the dipeptide which, according to the discussion of the dipeptide in dilute $CDCl_3$ solutions presented above, indicates the melting out of the folded forms.

At higher concentrations aggregates of the tripeptide are formed. These aggregates might be formed by association of extended tripeptides, or folded tripeptides, or both. The evidence concerning which if any of these predominates is mixed. There are two indications that there are significant amounts of associated folded tripeptides. First, the two downfield NH resonances, although having similar chemical shifts at high concentration, are still separately identifiable (Figure 7). This is contrary to the downfield NH resonances of the tetrapeptide at high concentrations (Figure 8). Models of associated extended tripeptides "in-register", i.e., primitive β sheets, show the protons of the N(2)H and N(3)H groups to be in virtually identical environments, and likely to have identical shifts. Secondly, each of the three NH resonances at 0.40 M concentration shows a different temperature dependence (Figure 9, Table II). Models of the extended "in-register" forms show that the amide proton of each residue participates in H bonding to the same extent in "in-register" associated forms, contrary to the case of the dipeptide described above (Figure 13). Thus, these two pieces of evidence, involving the observed differences in NH chemical shift and temperature dependence of chemical shift, point to the existence of aggregated folded forms.

On the other hand, it is also possible that there are "out-of-register" extended aggregated forms, with the following series of equilibria:



This scheme, too, could account for the differences in NH chemical shift and the differences in temperature dependence.

Table II and Figure 9 indicate that the temperature dependence of the N(2)H resonance is the same at low concentration from 20 to 30 °C (folded forms) as at high concentration (associations) which might indicate that no new N(2)H H bonds are formed during association. The temperature dependence of the N(3)H resonance, on the other hand, does increase as the associations are formed, so that at 0.40 M concentration its temperature dependence is the largest of the three NH resonances. This is evidence that the N(3)H groups are especially important in forming the aggregates. Such a result supports the picture that the tripeptide aggregates, like the dipeptide aggregates, consist largely of an "out-of-register" network of extended chains held together largely by N(3)H hydrogen bonds. Earlier work on the CD and IR properties of the tripeptide in the solid state[3,4] gave evidence that there was very little, if any, β structure present. The optical properties therefore support the picture just described and indicate that the amount of "in-register" extended chain aggregates is relatively small.

The temperature dependence data for the tetrapeptide in dilute $CDCl_3$ solutions at 20–30 °C (Figure 9, Table II) indicate the presence of folded forms. It should be noted that in this case not only ten-membered ring structures involving the N(3)H and N(4)H groups but also a 13-membered ring structure, i.e., one turn of α helix, involving the N(4)H group are possible. The α-helix forming character of the norvaline residue is well established.[32-37] Seven-membered ring structures are also possible, analogous to the folded tripeptide structures shown in Figure 14.

At higher concentrations the tetrapeptide aggregates, and the fact that the three low field NH resonances have the same chemical shift and display the same temperature dependence (Figure 9, Table II) is indication that the chains unfold as the associations are formed. In these respects the data are quite different from those described above for the associated tripeptide, in which case neither the chemical shifts nor the temperature dependencies of the chemical shifts are the same for the downfield resonances. Whereas the tripeptide associations are best described as a network of extended "out-of-register" forms, the tetrapeptide associations are better described as primitive β sheets. The earlier CD and IR studies[3,4] substantiate our conclusion here that β-like structures begin to appear at the level of the tetrapeptide.

We are currently employing high resolution ^{1}H NMR to establish in detail the conformations of other homo-oligopeptide series in $CDCl_3$ solution in order to shed more light on the structural dependence of folding, already noted by Shields and co-workers[38] and Palumbo et al.[4] using IR absorption.

Acknowledgments. This work was supported by NATO Grant No. 1099 and by National Institutes of Health Research Grant No. 1-P07-PR00798 from the Division of Research Resources.

References and Notes

(1) (a) Brown University; present address: Department of Chemistry, State University of New York at Binghamton, Binghamton, N.Y. 13901. (b) University of Padova.
(2) This work is part 39 of that series; for part 38 see M. M. Kelly, E. S. Pysh, G. M. Bonora, and C. Toniolo, *J. Am. Chem. Soc.*, **99**, 3264 (1977).
(3) J. S. Balcerski, E. S. Pysh, G. M. Bonora, and C. Toniolo, *J. Am. Chem. Soc.*, **98**, 3470 (1976).
(4) M. Palumbo, S. Da Rin, G. M. Bonora, and C. Toniolo, *Makromol. Chem.*, **177**, 1477 (1976).
(5) C. Toniolo and M. Palumbo, *Biopolymers*, in press.
(6) C. Toniolo, G. M. Bonora, M. Palumbo, and E. S. Pysh, in "Peptides 1976", A. Loffet, Ed., Presses University de Bruxelles, Bruxelles, 1976, p 597.
(7) C. Toniolo and G. M. Bonora in "Peptides: Chemistry, Structure and Biology", R. Walter and J. Melenhofer, Ed., Ann Arbor Science, Ann Arbor, Mich., 1975, p 145.
(8) C. Toniolo, G. M. Bonora, and A. Fontana, *Bull. Soc. Chim. Belg.*, **84**, 305 (1975).
(9) The following abbreviations are used: ^{1}H NMR (proton magnetic resonance), CD (circular dichroism), IR (infrared), Ac (acetyl), Boc (*tert*-butyloxycarbonyl), Nva (norvaline), Ala (alanine), Val (valine), OMe (methoxy), Me_2SO (dimethyl sulfoxide).
(10) M. Goodman, A. S. Verdini, C. Toniolo, W. D. Phillips, and F. A. Bovey, *Proc. Natl. Acad. Sci.*, **64**, 444 (1969).
(11) M. Goodman, C. Toniolo, and A. S. Verdini in "Peptides 1969", E. Scoffone, Ed., North Holland, Amsterdam, 1971, p. 207.
(12) P. A. Temussi and M. Goodman, *Proc. Natl. Acad. Sci.*, **68**, 1767 (1971).
(13) M. Goodman, C. Toniolo, and F. Nalder in "Peptides, Polypeptides and Proteins", E. R. Blout, F. A. Bovey, M. Goodman, and N. Lotan, Ed., Wiley, New York, N.Y., 1974, p. 308.
(14) M. Goodman, N. Ueyama, and F. Nalder, *Biopolymers*, **14**, 901 (1975).
(15) M. Goodman, N. Ueyama, and F. Nalder, *Biopolymers*, **14**, 915 (1975).
(16) M. Goodman, F. Toda, and N. Ueyama, *Proc. Natl. Acad. Sci.*, **70**, 331 (1973).
(17) V. F. Bystrov, S. L. Portnova, V. I. Tsetlin, V. T. Ivanov, and Yu. A. Ovchinnikov, *Tetrahedron*, **25**, 493 (1965).
(18) J. L. Dimicoli and M. Ptak, *Tetrahedron Lett.*, 2013 (1970).
(19) C. Toniolo and G. M. Bonora, *Can. J. Chem.*, **54**, 70 (1976).
(20) W. L. Mattice, *Biopolymers*, **13**, 169 (1974).
(21) G. M. Bonora, A. Maglione, A. Fontana, and C. Toniolo, *Bull. Soc. Chim. Belg.*, **84**, 299 (1975).
(22) H. Nesvadba, J. Honzl, and J. Ridlnger, *Collect. Czech. Chem. Commun.*, **28**, 1691 (1963).
(23) J. B. Jones and C. Niemann, *Biochemistry*, **1**, 1093 (1962).
(24) N. S. Baccha, L. F. Johnson, and J. N. Shoolery, Ed., Varian NMR Spectra Catalog, Vol. 1, 1962, Spectra No. 85, 319.
(25) J. E. Shields, *Biochemistry*, **5**, 1041 (1966).
(26) T. A. Foglia and D. Swern, *J. Org. Chem.*, **31**, 3625 (1966).
(27) V. F. Bystrov and V. P. Lezina, *Opt. Spektrosc.*, **16**, 790, 1004 (1964).
(28) M. T. Cung, M. Marraud, and J. Néel in "Peptides 1976", A. Loffet, Ed., Presses University de Bruxelles, in press; *Biopolymers*, **15**, 2081 (1976).
(29) M. T. Cung, Ph.D. Thesis, Polytechnic de Lorraine, Nancy, France, 1976.
(30) (a) G. Némethy and M. P. Printz, *Macromolecules*, **5**, 755 (1972); (b) C. B. Anfinsen and H. A. Scheraga, *Adv. Protein Chem.*, **29**, 305 (1975), and references therein.
(31) C. Toniolo, *Bioorg. Chem.*, in press.
(32) E. R. Blout in "Polyamino Acids, Polypeptides and Proteins", M. A. Stahmann, Ed., The University of Wisconsin Press, Madison, Wisc., 1962, p 275.
(33) V. Sasisekharan and P. K. Ponnuswamy, *Biopolymers*, **10**, 583 (1971).
(34) P. K. Ponnuswamy and V. Sasisekharan, *Int. J. Peptide Protein Res.*, **3**, 9 (1971).
(35) K. Itoh, T. Shimanouchi, and M. Oya, *Biopolymers*, **7**, 649 (1969).
(36) M. Oya, M. Tomizawa, K. Uno, and Y. Ywakura, *Bull. Soc. Chem. Jpn.*, **43**, 3486 (1970).
(37) B. R. Malcolm, *J. Polym. Sci., Part C*, **34**, 87 (1971).
(38) J. E. Shields, S. T. McDowell, J. Pavlos, and G. R. Gray, *J. Am. Chem. Soc.*, **90**, 3549 (1968).

Nuclear Magnetic Resonance Studies of *p*-Fluorocinnamate–α-Chymotrypsin Complexes

J. T. Gerig,* B. A. Halley, and C. E. Ortiz

Contribution from the Department of Chemistry, University of California, Santa Barbara, California 93106. *Received December 27, 1976*

Abstract: Fluorine magnetic resonance spectroscopy has been used to characterize the complexes formed between *p*-fluorocinnamate anion and the proteolytic enzyme, α-chymotrypsin. Equilibrium binding constants determined by competitive inhibition kinetics or a dye-displacement technique agree with the value found by NMR. Consideration of protein-induced chemical shifts and spin-lattice and transverse relaxation rates strongly indicates that *p*-fluorocinnamate binds equally well to the monomeric, dimeric, and trimeric forms of the enzyme with the properties of the binding sites so revealed being indistinguishable. A substantial part (about 62%) of the fluorine-proton dipolar relaxation derives from interaction with protons of the enzyme.

Bender and his co-workers have nicely elucidated the mechanism of chymotrypsin catalysis especially as regards the hydrolysis of cinnamoyl esters and amides.[1] A minimal mechanism for this process is represented by

$$E + AL \rightleftharpoons [E\cdot\cdot AL] \underset{+L}{\overset{-L}{\rightleftharpoons}} \hat{E}A \underset{-H_2O}{\overset{H_2O}{\rightleftharpoons}} [E\cdot\cdot A] \underset{k_{-4}}{\overset{k_4}{\rightleftharpoons}} E + A \quad (1)$$

This sequence portrays as the first stage of the reaction formation of an enzyme-substrate complex ([E··AL]) within which the acyl group of the substrate is transferred to the serine-195 residue of the protein, giving an acylated enzyme, $\hat{E}A$. Hydrolysis of this enzyme ester affords an enzyme-product complex ([E··A]) which then can reversibly dissociate to regenerate the free enzyme. When the acyl group A is cin-