about 10% relative. The fifth sample (62.0 mol % propylene) had no detectable nonalternating TFE structure, and therefore is not included in the plot.

The symmetry of the curves for percent nonalternating propylene and TFE sequences confirms the consistency of the analyses of both the ^1H and ^{19}F NMR spectra. Trial calculations of monomer reactivity ratios were made, based on the known monomer feed ratios, copolymer compositions, and percent nonalternating sequences. Quantitatively reliable estimates of reactivity ratios were not possible, because the observed percent nonalternating sequences could not be produced by Bernoulli trial statistics. The data available were insufficient for fitting higher order models.

Acknowledgment. The authors wish to thank Dr. G. S. Reddy, F. W. Barney, Jr., F. W. Yeager, and R. G. Hamilton for their help in obtaining the NMR spectra used in this work. Thanks are also due T. F. George for his help in the copolymerizations.

References and Notes

(1) (a) Elastomer Chemicals Department; (b) Central Research and Development Department; (c) contribution No. 317; (d) contribution no. 2267.
(2) F. A. Bovey, "High Resolution NMR of Macromolecules", Academic Press, New York, N.Y., 1972.
(3) R. C. Ferguson, *Kautsch. Gummi Kunstst.*, **18**, 723 (1965).
(4) K. Ishigure, Y. Tabata, and K. Oshima, *Macromolecules*, **3**, 27 (1970).
(5) K. Ishigure and Y. Tabata, *Macromolecules*, **3**, 450 (1970).
(6) K. Ishigure, Y. Tabata, and K. Oshima, *Macromolecules*, **6**, 584 (1973).
(7) Brasen and Cleaver, U.S. Patent 3,467,635
(8) R. Wickbold, *Angew. Chem.*, **64**, 133 (1952); **64**, 173 (1954).
(9) R. C. Ferguson, *Trans. N.Y. Acad. Sci.*, **29**, 495 (1967).
(10) D. D. Elleman, L. C. Brown, and D. Williams, *J. Mol. Spectrosc.*, **7**, 393 (1961).
(11) C. W. Haigh and R. B. Mallon, *J. Mol. Spectrosc.*, **29**, 478 (1969); C. W. Haigh, "Annual Review of NMR Spectroscopy", Vol. 4, E. F. Mooney, Ed., Academic Press, London, 1971.
(12) R. C. Ferguson, *J. Magn. Reson.*, **12**, 296 (1973).

Investigation of the Conformations of Four Tetrapeptides by Nuclear Magnetic Resonance and Circular Dichroism Spectroscopy, and Conformational Energy Calculations[1]

J. C. Howard,[2a,c] A. Ali,[2b,c] H. A. Scheraga,*[2c] and F. A. Momany[2d]

Departments of Chemistry, Cornell University, Ithaca, New York 14853, and Memphis State University, Memphis, Tennessee 38152. *Received March 24, 1975*

ABSTRACT: Proton nuclear magnetic resonance and circular dichroism studies were carried out on aqueous solutions of the tetrapeptide Asp-Lys-Thr-Gly (which appears as a bend at residues 35–38 of α-chymotrypsin) and its sequence variants Gly-Thr-Asp-Lys, Asp-Lys-Gly-Thr, and Lys-Thr-Gly-Asp; the N and C termini of all four tetrapeptides were blocked with CH_3CO and $NHCH_3$ groups, respectively. The spectroscopic data suggest that bend conformations may exist, to some extent, among the distributions of conformations in the first, third, and fourth, but not in the second, tetrapeptide. This result is consistent with empirical probabilities for the prediction of bend conformations in proteins. Conformational energy calculations on these four tetrapeptides support the indications from the experimental data. It thus appears that, because of short-range interactions, the tendency toward bend formation exists in short peptides, provided that both the composition and amino acid sequence are energetically favorable for bend formation.

In a previous paper,[3] a mechanism for the folding of proteins (involving the formation of β turns, or chain reversals in specific tetrapeptide sequences) was proposed. Subsequently, the conformational energies of several tetrapeptides were minimized[4] to determine the energetic factors responsible for bend and nonbend structures; similar calculations have been carried out on the central dipeptides of numerous tetrapeptides.[5,6] In order to examine the tendencies of specific tetrapeptide sequences to adopt a chain-reversing conformation, we have begun a combined experimental and computational investigation of several isolated tetrapeptides.[7] In this paper, we consider the sequence Asp-Lys-Thr-Gly, which appears as a bend conformation at residues 35–38 in α-chymotrypsin.[8]

Since the tendency toward bend formation depends on both the composition and sequence of the tetrapeptide,[3] we have examined four of the 24 possible sequence permutations of these four amino acid residues, viz., Gly-Thr-Asp-Lys, Asp-Lys-Thr-Gly, Asp-Lys-Gly-Thr, and Lys-Thr-Gly-Asp. The theoretical relative probabilities of occurrence of bends[3] in these four tetrapeptides are 1, 23, 41, and 41, respectively, and hence they are designated as low-probability (L), chymotrypsin-native (N), and high-probability (H_1 and H_2) bends, respectively. In order that these isolated tetrapeptides simulate similar-size segments of polypeptide chains, we have prepared them with blocked end groups, using CH_3CO and $NHCH_3$ at the N and C termini, respectively. Nuclear magnetic resonance and circular dichroism measurements, and conformational energy calculations,[9,10] were carried out on these four tetrapeptides to determine whether a detectable fraction of the population of conformations of each one can exist as a chain reversal and, if so, whether this tendency differs among L, N, H_1, and H_2.

I. Experimental and Computational Procedures

Materials. All solvents used were spectral grade and were used as purchased, except as indicated below. Isobutyl chloroformate, TFA,[11] and $BF_3O(Et)_2$ were purchased from Eastman, and THF and DMF were purchased from J. T. Baker. TFA was distilled over P_2O_5 before use, and $BF_3O(Et)_2$ was distilled as described by Zweifel and Brown.[12] Monomethylamine was purchased from Matheson Co., and *N*-methylmorpholine was purchased from Aldrich Chem. Co.

α-BOC (ϵ-Z) lysine was prepared by the procedure of Ali et al.,[13] and BOC glycine was prepared as described by Schnabel.[14] BOC threonine was prepared by the method of Hofmann et al.,[15] BOC

(β-OBz) aspartic acid by the method of Stewart and Young,[16] Z glycine by the method of Greenstein and Winitz,[17] and Z threonine by the procedure of Merrifield.[18] BOC glycine pentachlorophenyl ester and BOC (ϵ-Z) lysine pentachlorophenyl ester were prepared by the method of Johnson and Trask.[19] BOC (β-OBz) aspartate pentachlorophenyl ester was prepared as described by Johnson.[20] BOC threonine pentachlorophenyl ester and Z threonine pentachlorophenyl ester were prepared by the procedure of Visser et al.,[21] and (ϵ-Z) lysine *N*-carboxyanhydride was prepared as described by Fasman et al.[22]

Synthesis. The synthesis of the blocked tetrapeptides is described in the Appendix.

Chromatography. AG_1-X_2 ion exchange resin (−400 mesh, chloride form) was purchased from Bio-Rad Laboratories and converted to the acetate form by washing with 50% acetic acid. Carboxymethyl cellulose 52 was purchased from Whatman, and silica gel (60–200 mesh) was purchased from Davison Chemical.

The peptides were purified by liquid–solid chromatography, using ninhydrin to detect the peptide material in the fractions, and then the purity of the peptide material in each fraction containing peptide was determined by thin layer chromatography. The following columns were used. (A) AG_1-X_2 (-400 mesh) ion exchange resin in the acetate form was packed in a 2 × 60 cm column and equilibrated with 2% pyridine solution. The peptide was applied to the column in a small amount of 2% pyridine and eluted with 2% pyridine solution (200 ml) followed by 100 ml of 5% acetic acid solution. The flow rate was 8–10 ml/hr and fractions of 4–5 ml were collected. To use the column again, it was first washed with 50% acetic acid solution followed by water. (B) Carboxymethyl cellulose 52 was packed in a 2 × 120 cm column and equilibrated with 0.01 *N* acetic acid. The peptide was applied to the column in a small amount of 0.01 *N* acetic acid and eluted with a gradient of 0.01 *N* acetic acid (600 ml) and 3.5 *N* acetic acid (600 ml). The flow rate was 14–16 ml/hr and the fraction size was 7–8 ml. (C) AG_1-X_2 (-400 mesh) ion exchange resin in the acetate form was packed in a 2 × 60 cm column and equilibrated with 1% acetic acid solution. The peptide was applied to the column in a small amount of 1% acetic acid and eluted with 1% acetic acid solution. The flow rate and fraction size were the same as in A. (D) Silica gel (60–200 mesh), packed with acetone in a 4 × 45 cm column, was used for the protected amino acids and peptides. The compound was applied to the column as a solid or in a small amount of suitable solvent and eluted with an acetic acid–acetone (2:98) mixture. (E) Silica gel (60–200 mesh) was packed using chloroform in a 3 × 105 cm column. The compound was applied as above and eluted with increasing volume ratio of acetic acid–methanol and finally eluted with chloroform–methanol–acetic acid (91:6:3).

The purified end-blocked tetrapeptides, in the appropriate fractions eluted from columns A, B, and C, were recovered by lyophilization.

The purity of the final products and intermediate derivatives in the synthesis was checked by thin layer chromatography, using Merck's silica gel F-254 plates, layer thickness 0.25 mm, with the following systems: R_F^1, *n*-butyl alcohol–acetic acid–water, 60:20:20; R_F^2, *n*-butyl alcohol–acetic acid–water, 40:30:30; R_F^3, *n*-butyl alcohol–acetic acid–water, 30:40:40; R_F^4, *n*-butyl alcohol–pyridine–acetic acid–water, 15:10:3:12; R_F^5, acetic acid–chloroform–methanol, 5:85:10; R_F^6, methanol–chloroform, 50:50; R_F^7, acetic acid–acetone, 2:98. R_f^5 was found to be best for study of the active ester coupling reactions. The structures of the products and intermediate derivatives were verified by PMR spectroscopy.

Amino Acid Analysis and Test for Racemization. The composition of the peptides was confirmed by amino acid analysis (Technicon Amino Acid Analyzer). Acid hydrolyses were carried out in constant boiling HCl containing 10% by volume of a 0.5 *M* phenol solution at 105° for 20 hr in a sealed tube. Optical purity of the final peptides was checked by the procedure of Manning and Moore,[23] and the results are listed in Table I.

The optical rotation of the peptides at the sodium D line was determined at room temperature (23°) with a Perkin-Elmer Model 141 polarimeter, using a 1 dm cell. All melting points are uncorrected.

Circular Dichroism. Circular dichroism (CD) spectra were obtained with a Cary Model 60 spectropolarimeter equipped with a Model 6001 CD attachment and a 450 W Osram Xenon arc lamp or, alternatively, with a Jasco Model ORD/UV-5 spectropolarimeter with CD attachment and a 500 W Xenon high-pressure arc lamp. The sample temperature in the Cary instrument was ambient (26°) but the temperature in the Jasco apparatus (5°) was controlled (to within ±0.5°) with the aid of a water-jacketed quartz cell. The path length was 10.017 mm, and the polypeptide concentration (in H_2O) was 10^{-4} or 10^{-5} *M* in the 250–210 and 210–195 nm regions, respectively.

Table I
Optical Purity of Synthesized Peptides

Peptide	% D isomer found[a]		
	Asp	Lys	*a*thr
L	1.6	0.5	0
N	3.8	2.0	0
H_1	2.0	1.4	0
H_2	0.8	1.0	0

[a] Error no greater than ±0.2.

Proton Magnetic Resonance. Proton magnetic resonance (PMR) spectra were obtained in H_2O and D_2O with the following instruments: Varian HA-100 NMR spectrometer with a standard variable temperature controller and CAT capability; Varian HR-220 NMR spectrometer[24] with standard temperature controller; Bruker HX-90 multinuclei spectrometer equipped with a Bruker Typ B-ST 100/700 temperature controller unit and a NMR3 Digilab data system; and a 250 MHz NMR spectrometer.[25]

Measurements of the spin–lattice relaxation time, T_1, were made on the Bruker HX-90 using a standard 180°–τ–90°, *T* pulse sequence[26] on samples in 99.8% D_2O at 3° which had been degassed by several freeze–thaw cycles under vacuum (10^{-4} Torr). Nuclear Overhauser Effect (NOE) experiments were carried out on the above freeze–thaw degassed samples (at 5°) using the Bruker HX-90 in a continuous wave mode.

Because several instruments were used, a variety of referencing techniques were required and will be described as they arise in the figures and tables.

Conformational Energy Calculations. The nomenclature and conventions used here are those adopted by an IUPAC-IUB commission.[27] The bond angles, bond lengths, partial atomic changes, torsional potentials, and nonbonded and hydrogen-bond parameters for treating the tetrapeptides are given elsewhere.[10] All bond lengths and bond angles were maintained fixed, and only dihedral angles were allowed to vary; in addition, the terminal methyl groups of the tetrapeptide were kept in the trans (180°) conformation (i.e., with a methyl hydrogen trans to the nearest C′–N bond). Other dihedral angles which were held fixed at 180° are: χ^2, χ^3, and χ^4 of Lys, χ^5 being kept at 90°; χ^3 of Asp [i.e., the acid proton was kept in the extended (180°) conformation]; and the terminal peptides (ω_0 and ω_4 = 180°). These choices for Lys and Asp were based on previous experience,[9] and that for ω_0 and ω_4 was made because there does not appear to be any strain in these molecules which could cause the terminal peptide groups to depart from planarity. All other dihedral angles were allowed to vary.

The total energy of any tetrapeptide conformation is taken as a sum of the contributions[10] from all intramolecular torsional, nonbonded, hydrogen-bonding, and electrostatic energy terms. No solvent effects were included in these calculations, and the effective dielectric constant in the electrostatic term was taken as two.[10] The conformational energy of many starting sets of dihedral angles was minimized (as in ref 9) with respect to all the dihedral angles except those mentioned above.

The four tetrapeptides, whose minimum-energy conformations were determined, are those listed in the introductory section, all residues being taken in the L configuration. Those side chains having ionizable groups were taken to be in their uncharged forms; the justification for taking these groups as neutral in water has been provided elsewhere.[28] It is clear that a completely ionized side chain would not hydrogen bond in the same way as the neutral groups used here. However, we have found previously[28] that, by using the uncharged groups, the effect of ion shielding is simulated, and the computed conformation reproduces experimental data more closely.[28] Conformers in which the charge effect may influence the choice of lowest-energy structure are discussed in the Experimental Results section. The selection of the various starting conformations for minimization of the total energy is discussed in the Results section. The contributions of the librational entropy term[9] to the stabilities of particular minimum-energy conformations are not calculated, but the energy surface around the mini-

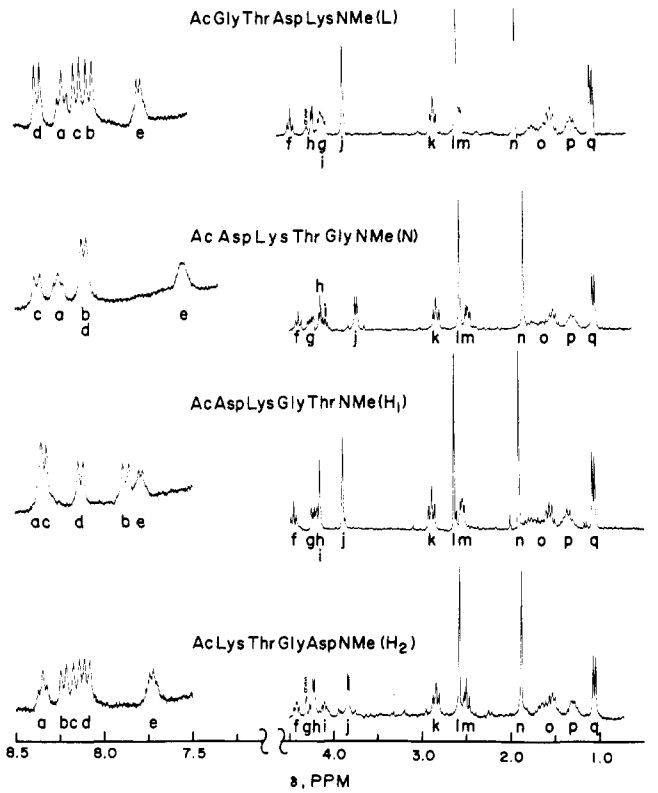

Figure 1. NMR spectra at 220 MHz of the blocked tetrapeptides L, N, H_1, and H_2 in H_2O and D_2O at 18° and pH_{app} ~5, with no salt or buffer added (peptide concentration = 4% (w/v)). TMS (Varian standard reference sample) was used as an external reference. Assignments: a, b, c, d, and e are the amide protons (in H_2O) of Gly, Thr, Lys, Asp, and $NHCH_3$, respectively. The aliphatic protons (in D_2O) are: f, Asp $C^{\alpha}H$; g, Lys $C^{\alpha}H$; h, Thr $C^{\alpha}H$; i, Thr $C^{\beta}H$; j, Gly $C^{\alpha}H_2$; k, Lys $C^{\epsilon}H_2$; l, NCH_3; m, Asp $C^{\beta}H_2$; n, CH_3CO; o, Lys $C^{\beta,\delta}H_2$; p, Lys $C^{\gamma}H_2$; q, Thr $C^{\gamma}H_3$.

mum-energy conformations is examined for each lowest-energy conformation in section III.

II. Experimental Results

PMR Spectra The PMR spectra at 220 MHz (at 18°) of the four blocked tetrapeptides, L, N, H_1, and H_2, are shown in Figure 1. The resonances of the aliphatic protons (shown on the right-hand side of Figure 1) were observed by taking the spectra in D_2O solution (pD_{ap} ~5). The amide proton resonances (shown on the left-hand side of Figure 1) were obtained in H_2O solution (pH ~5). The polypeptide concentrations were 4% w/v (~0.1 *M*). The temperature was 18°, and no salt or buffer was added. Fourier transform (FT) spectra were obtained using both dilute (0.1% w/v) and concentrated (5–20% w/v) solutions (90 MHz spectra in D_2O); no concentration dependence was observed when the different spectra were superimposed. The solubilities of the four tetrapeptides differed, with L and H_1 being soluble up to ~20% w/v, N to ~10% w/v, and H_2 to 5% w/v.

The assignments of the resonances are given in the legend of Figure 1 and were made by comparing the spectra obtained here to published spectra[29,30] for the same amino acid residues, and by use of coupling patterns, area ratios, and decoupling experiments. An example of a decoupling experiment is shown in Figure 2 for one of the tetrapeptides.

The coupling constants $J_{NH-C^{\alpha}H}$, obtained from the 5° data at 220 MHz, are given in Table II. From the values of $J_{NH-C^{\alpha}H}$ of Table II, the dihedral angle ϕ, for rotation about the N–C^{α} bond (in the current convention[27]), was obtained using the Bystrov relationship.[31] However, we have modified the relationship for ΣJ_{NH-CH_2} for glycyl residues because the value of the *projection* of the bond angle τ ($HC^{\alpha}H$ = 107°) of glycine on a plane that is perpendicular to the N–C^{α} bond is 116°,[10] rather than the 120° used by Bystrov et al.[31]

It should be realized that, as indicated in Table II, the derived values of ϕ are fourfold degenerate (and also have large uncertainties because of experimental error in $J_{NH-C^{\alpha}H}$ and because of the inexactness of the Bystrov relation); also, since a distribution of conformations probably exists for each tetrapeptide, the reported values of ϕ probably represent averages over the distributions. In addition, the values of ϕ computed for the Gly residue are less accurate than those for the other residues because of the uncertainty in the *projection* of the bond angle, τ ($HC^{\alpha}H$ = 107°), of glycine on a plane that is perpendicular to the N–C^{α} bond, mentioned above. Variations in the projection of τ (which influence $\Sigma J_{NH-C^{\alpha}H_2}$ of glycine) could lead to large errors in the computed value[31] of ϕ.

The temperature dependences of the chemical shifts of the NH resonances (measured between 5 and 35° at 220 MHz, with an external Varian TMS reference, and hence uncorrected for differences in susceptibility) are also presented in Table II. Since these values are all similar, and rather high, they provide no indication that any intramolecular hydrogen bonds are present with lifetimes longer than 10^{-3} sec,[32] over this temperature range. By the nature of such an experiment, a hydrogen bond present at 5° but not at, say, 25° would not be detected.

Conventional H–D exchange experiments were also carried out to obtain information about possible intramolecular hydrogen bonds. The NH region of the NMR spectrum (220 MHz) of each tetrapeptide was observed at 5°, and the sample tube was then removed from the probe. D_2O was added to a volume concentration of 50:50, and the sample was shaken and replaced in the probe. By this technique, it was not possible to control the temperature continuously. The NMR spectrum was observed as a function of time. About 1 min elapsed between the addition of D_2O and the recording of the first spectrum. The peak areas of all NH resonances were reduced to ~¼ of their original areas immediately because of the dilution by a factor of 2 and also because of rapid exchange.

The temperature coefficients were high (Table II), and comparable to values of 7.6×10^{-3} ppm/°C for glycyl NH and 5.1×10^{-3} ppm/°C for the methylamide NH, found here for a nonhydrogen-bonded model compound (Ac-Gly-$NHCH_3$) in H_2O at pH 6.2. Also, the H-D exchange was rapid. Thus, no evidence for strong hydrogen bonds in any of the tetrapeptides was obtained.

Since there is no definitive evidence for hydrogen bonding, and since the ranges deduced for ϕ are too large to be helpful, by themselves, in deducing an average conformation for each tetrapeptide in solution, two other types of NMR experiments were carried out, viz., observations of (1) the Nuclear Overhauser effect[33] (NOE) and (2) the spin–lattice relaxation time[26] T_1.

If a bend conformation were to exist to any significant extent in any tetrapeptide, model building indicates that the methyls of the CH_3CO and $NHCH_3$ blocking groups would be near each other. Such an intramolecular methyl–methyl contact might lead to an NOE.[33] For example, an NOE (10% enhancement) was observed between proton 3′ of the sugar moiety and proton 6 of the base moiety in a D_2O solution of cytidine (p 199 of ref 33). In camphene, saturation of the methyl signals at 1.01 and 1.05 ppm led to a 16% enhancement of the peak at 4.44 ppm but no change in the peak at 4.65 ppm, thus providing an unequivocal assignment of the syn- and anti-exo protons (p 171 of ref 33).

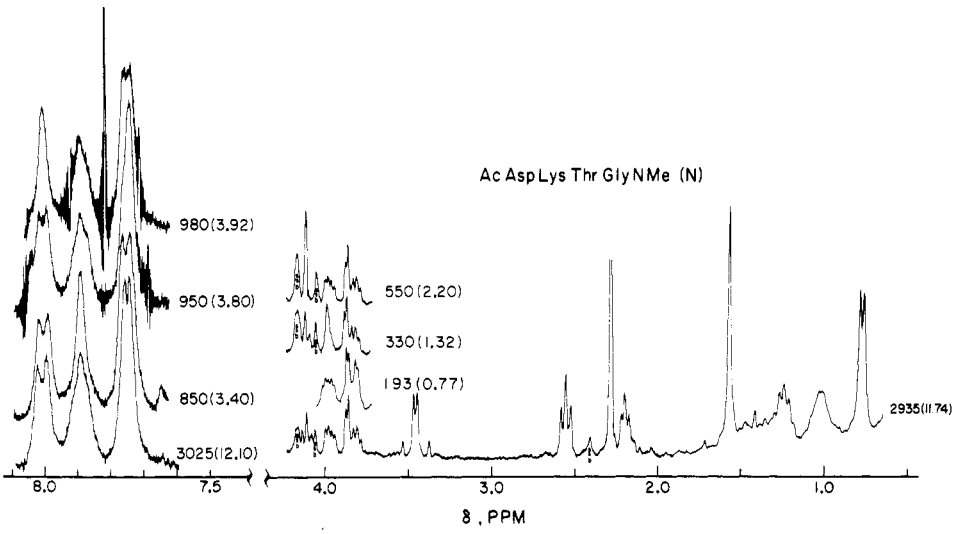

Figure 2. An example of the decoupling experiments, at 250 MHz at 27°, on peptide N at 4% (w/v) concentration in H_2O at pH_{app} 5 for $C^{\alpha}H$–NH, and in D_2O at pD_{app} 5 for $C^{\beta}H$–$C^{\alpha}H$, using 100% TMS capillary insert (Wilmad Glass Co. #520-2) for reference and lock signal. The numbers indicate the frequency in hertz of the decoupling radiofrequency transmitter relative to the lock frequency; the values in parentheses are the same frequencies in ppm.

Table II
Data for the Amide Protons of the Tetrapeptides

Peptide	Residue[a]	$J_{NH-C^{\alpha}H}$,[b] Hz	$d\delta/dt$[c] × 10^3, ppm/°C	ϕ,[d] deg
L	Asp	7.1	6.9	−160, −80, 35, 85
	Lys	8.2	9.9	−150, −90, 45, 75
	Thr	8.2	8.0	−150, −90, 45, 75
	Gly	Σ = 11.4	8.0	±60, ±135
	$NHCH_3$		7.6	
N	Asp	6.0–6.5	7.7–8.3	−165 to −160, −80 to −75, 25 to 30, 90 to 95
	Lys	7.6	9.3	−155, −85, 40, 80
	Thr	6.0–6.5	7.7–8.3	−165 to −160, −80 to −75, 25 to 30, 90 to 95
	Gly	Σ = 10.9	7.9	±55, ±140
	$NHCH_3$		6.1	
H_1	Asp	6.5	8.4	−160, −80, 30, 90
	Lys	7.1	9.2	−160, −80, 35, 85
	Thr	8.2	8.4	−150, −90, 45, 75
	Gly	Σ = 10.4	7.3	±50, ±145
	$NHCH_3$		7.8	
H_2	Asp	7.6	7.5	−155, −85, 40, 80
	Lys	8.2	10.1	−150, −90, 45, 75
	Thr	7.1	9.4	−160, −80, 35, 85
	Gly	Σ = 11.4	9.5	±60, ±135
	$NHCH_3$		9.4	

[a] The amino acids are listed in the same order (the sequence of peptide N) to facilitate comparisons within the table. [b] Obtained from J_{obsd} (220 MHz at 5°), using electronegativity correction suggested by Bystrov et al.[31] [c] 220 MHz data, using Varian standard reference sample (TMS) as external reference at each temperature between 5 and 35°. No correction was made for the variation of the bulk susceptibility of TMS with temperature because it is small, and only large changes in $d\delta/dt$ are of interest here. [d] Reference 31, rounded to nearest 5°. Some of the degeneracy in dihedral angle can be removed by energy considerations; see footnote *b* of Table VI.

Small negative NOE's involving a methyl group and a nearby proton were observed in β,β-dimethylacrylic acid and in dimethylformamide (p 168 of ref 33). A small negative (−2.1%) NOE was observed in trans crotonaldehyde while observing the methyl resonance.[34] Therefore, NOE experiments were carried out at 90 MHz by irradiating the CH_3CO methyl group and observing the change in area of the peak from the $NHCH_3$ methyl group. This experiment is feasible because the resonances of these methyl groups are separated by 60–65 Hz at 90 MHz. With 22 dB attenuation of radiofrequency power in the observing (f_1) channel being found to be necessary to detect saturation effects, the NOE experiments were performed using 42 dB attenuation of f_1, radiofrequency power. Thus, any radiofrequency in the irradiating channel (f_2) "leaking" over to the observed peak would be expected to cause an increase in peak area, rather than the decrease which was found (see Table III). The data of Table III, obtained at 5°, indicate that tetrapeptides N, H_1, and H_2 show a small (but detectable, with a large experimental error) decrease in area in the NOE experiment, and that no effect is observed for tetrapeptide L. The Gly $C^{\alpha}H_2$ peak showed no change in area, within the experimental error, for all four tetrapeptides. At 90 MHz, the $NHCH_3$ methyl peak overlaps that of the $C^{\epsilon}H_2$ of Lys

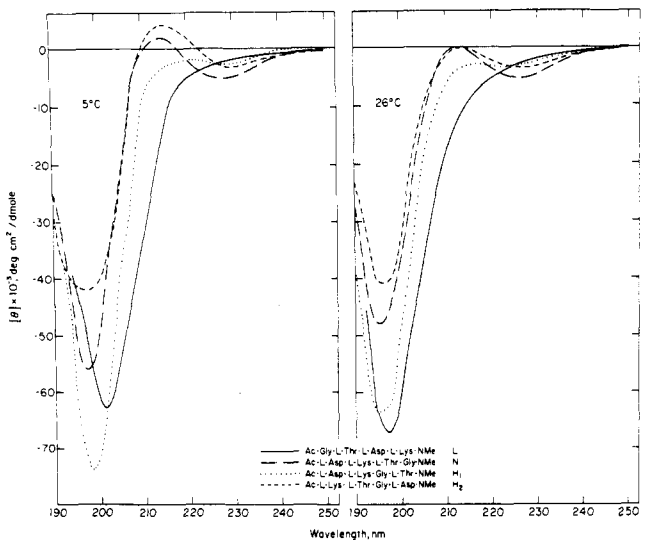

Figure 3. Circular dichroism spectra of the four tetrapeptides at 5 and 26° in H_2O at 10^{-4} to 10^{-5} M concentration of peptide. pH_{app} 5, with no salt or buffer added. The units of the CD data are (deg cm^2)/dmol of *tetrapeptide*.

Table III
NOE Results at 5° for the Four Tetrapeptides

Tetrapeptide	% change in area[a]	Tetrapeptide	% change in area[a]
L	0	H_1	−6 ± 6
N	−4 ± 4	H_2	−5 ± 5

[a] This is the percent change in the observed area of $NHCH_3$ methyl protons when the CH_3CO methyl protons are irradiated at 90 MHz. The data for each peak were obtained from ten spectra, and the area of each peak was measured five times; i.e., each area is an average over 50 values. The signal-to-noise ratio ranged between 40:1 and 70:1 for the methyl peaks.

and $C^\beta H_2$ of Asp, to some extent, in all the tetrapeptides. [Since the $C^\epsilon H_2$ of the lysyl residue is far from the CH_3CO methyl group in all four tetrapeptides, no NOE involving the $C^\epsilon H_2$ group is possible. However, the $C^\beta H_2$ of the aspartyl residue is about 5 Å from the CH_3CO methyl group in tetrapeptides H_1 and N, and this proximity might be the source of the observed change in area. To exclude this possibility, a similar NOE experiment was carried out on the model compound *N*-acetyl-L-aspartyl-*N'*-methylamide at pH 8. The total area of the $C^\beta H_2$ resonances and the $NHCH_3$ methyl resonances was found to be 103.8 ± 1.8 (arbitrary units) when the irradiating frequency, f_2, was offset to high field, and was 104.0 ± 2.0 when f_2 was adjusted to saturate the CH_3CO methyl resonance. This absence of an NOE is reasonable since the $C^\beta H_2$ protons are strongly spin–spin coupled to the $C^\alpha H$ proton; this coupling would be expected to decrease or eliminate any NOE involving the $C^\beta H_2$ protons with other groups.] Therefore, the change in *total* area of the Lys $C^\epsilon H_2$, Asp $C^\beta H_2$, and $NHCH_3$ methyl resonances in the tetrapeptide spectra (when f_2 was first adjusted to saturate the CH_3CO methyl resonance and then offset to high field, i.e., about 1 ppm higher than the resonance due to the $C^\gamma H_3$ of Thr) was attributed to an NOE at the $NHCH_3$ methyl protons. The results in Table III indicate that peptides N, H_1, and H_2 differ from peptide L, and that, among the distributions of populations of peptides N, H_1, and H_2, there may exist a not insignificant fraction of chain-reversal conformations which bring the methyls of the blocking end groups near each other to give rise to the small but detectable NOE.

Table IV
NMR Spin-Lattice Relaxation Times for the CH_3CO and $NHCH_3$ Methyl Groups of the Four Tetrapeptides at 3°

	T_1, sec	
Tetrapeptide	CH_3CO methyl	$NHCH_3$ methyl
L	1.24 ± 18%	1.35 ± 22%
N	1.06 ± 30%	0.86 ± 23%
H_1	0.48 ± 10%	0.50 ± 15%
H_2	1.06 ± 26%	0.98 ± 22%

The second additional type of NMR experiment involved the measurement of T_1 at 3° for the hydrogen resonances of the methyls of the CH_3CO and $NHCH_3$ groups to try to detect interactions between these methyls, i.e., restricted motion of these groups (which might arise from chain-reversal conformations). Measurements of T_1 were not made at higher temperatures, since the other NMR data (and the CD data described below) indicated that, if chain-reversal conformations exist at all, they are more likely to occur at low temperature. The spin–lattice relaxation times were measured by following the change in peak height in the Fourier transformed spectrum (90 MHz) as a function of the delay, τ, in a standard 180°–τ–90° T pulse sequence[26] relative to the peak height at $T \geq 5T_1$. The results of these experiments are given in Table IV. The error cited in Table IV is the standard deviation in the slope of the best straight line obtained from a least-squares fit of the logarithm of the difference in peak height (compared to $T \geq 5T_1$) plotted against τ, and is not an average of several experiments. The error is so large because this plot is not linear; hence, it is concluded that the methyl resonances of the end groups do not follow a first-order relaxation mechanism (assumed by the analysis) and thus that no statement can be made about the relative correlation times. The complex motion of the end groups, manifesting itself as a departure from linearity, may be the reason that the observed NOE was negative. Thus, the T_1 measurements may rationalize the sign of the observed NOE, but in themselves provide no information about the distances between methyl groups.

Circular Dichroism Spectra. The circular dichroism spectra of the four tetrapeptides at 5 and 26° are shown in Figure 3. Each curve is an average of two independent runs made on two separate samples of tetrapeptide. The data at 5° were obtained on the Jasco instrument, while those at 26° were obtained on the Cary instrument. To compare the results for the two instruments, tetrapeptide N was run at 26° on both instruments, with no significant difference in the two spectra.

The spectra of peptide L are the same at 5 and at 26°, whereas those of the other three peptides change with temperature in this range. Further, the spectra of peptides N, H_1, and H_2 differ from that of peptide L. The trough observed in the range of 220–240 nm for peptides N, H_1, and H_2 [with $[\theta]$ in the range of ~−6000], but not for peptide L, is displaced to longer wavelengths from that observed by Crippen and Yang[35] for *N*-acetyl-L-alanine-*N'*-methyl amide in 1,2-dichloroethane, with $[\theta] \sim -12{,}000$ (see their Figure 2). While Crippen and Yang interpreted their CD data (together with accompanying infrared data) in terms of the presence of a signifcant fraction of a seven-membered hydrogen-bonded ring conformation, a similar conclusion cannot be drawn here for tetrapeptides N, H_1, and H_2 without additional data (but see section IV for further discussion of the interpretation of the CD data).

III. Results of Conformational Energy Calculations

The conformational energy calculations were carried out without any initial bias from the experimental results. The resulting minimum-energy conformations were then considered in light of the NMR and CD results.

Selection of Starting Conformations. From the point of view of the feasibility of an extensive search of its conformational space, a tetrapeptide is a rather large molecule.[36] In order to search this space adequately by energy minimization, we have used the low-energy conformations of single residues[9] as initial ones for each of the four residues of the tetrapeptides. The initial conformations of the single residues were actual minimum-energy values[9] which lie in the following ranges, and the minima are identified by this nomenclature: (1) extended (β_1) ϕ = −150 to −180°, ψ = 130 to 180°; (2) extended puckered (2_1 and β_2) ϕ = −140 to −150°, ψ = 60 to 80° and ϕ = −60 to −80°, ψ = 120 to 150°, respectively, or ($2_1'$ for glycine), ϕ = 140 to 150°, ψ = −70 to −80°; (3) equatorial seven-membered ring (C_7^{eq}) ϕ = −70 to −80°, ψ = 75 to 85°; (4) right-handed α helix (α_R) ϕ = −60 to −80°, ψ = −50 to −70° and (α'_R), ϕ = −150 to −170°, ψ = −55 to −65°; (5) left-handed α helix (α_L) ϕ = 60 to 80°, ψ = 50 to 70°; (6) axial seven-membered ring (C_7^{ax}) ϕ = 70 to 80°, ψ = −75 to −85°; (7) types I through III′ bends (see ref 4).

For each minimum-energy set of values of ϕ and ψ[9] for Gly, Asp, Lys, and Thr, which lies in the above ranges, one or more best side-chain conformations[9] were selected as initial ones. All initial values of ω were taken as 180°. By starting from this selected set of conformations, and minimizing the conformational energy with respect to ϕ, ψ, ω, and the χ_i's, it is possible to reach a minimum in the tetrapeptide energy that is not a combination of the single-residue minima. While it is not feasible to search the complete conformational space of each tetrapeptide, the choice of single-residue minima (within 2 to 3 kcal/mol of the global minimum), which are also those found in proteins, is a reasonable compromise which should provide a high probability of finding the lowest-energy tetrapeptide conformations by energy minimization from these starting points.

Peptide L (Ac-Gly-Thr-Asp-Lys-NMe). The starting conformations for peptide L are listed in Table V. In most cases, only one cycle of minimization[37,38] was carried out to relieve steric overlaps and to search for a path to a lower-energy conformation. The energy listed in the last column of Table V is that for the lowest-energy conformer in each group (corresponding to the last symbol under Gly), after one cycle of minimization.

Two observations result from an examination of Table V. First, if the ith and (i + 1)th (or, correspondingly, the (i + 2)th and (i + 3)th) are locked in a local minimum- (but moderately high-) energy conformation (i.e., 2–3 kcal/mol above the energy of other conformers after the same number of cycles of minimization), no variation in the conformation of the other two residues will lower the total energy. For example, with Asp-Lys in the C_7^{eq}–α_R conformation, no alteration of the Gly-Thr backbone could reduce the total energy below −16 kcal/mol. Similarly, with Gly-Thr in the C_7^{ax}(or C_7^{eq})–β_1 conformation, no alteration of the Asp-Lys backbone could reduce the total energy below −15 kcal/mol. Thus, it is possible to discard many combinations of conformations (such as those described above), whose energy cannot be lowered further by additional cycles of minimization. Second, if residues (i + 1) and (i + 2) are placed initially in β_1 or β_2 conformations, the initial energy of the tetrapeptide is generally raised by several kilocalories per mole, and (upon further minimization) no low-energy structure (with the i + 1 or i + 2 residue remaining in the β_1 or β_2 conformations) could be obtained. The reason for this is not that β_1 or β_2 are high-energy conformations of a single residue, but rather that the resulting structure is not as compact as those with the (i + 1)th or (i + 2)th residues in a different conformation. Generally, compact structures of a tetrapeptide are found here to have the lowest energies.

The observations cited above were tested by selecting combinations of residue conformations that apparently were not compatible, in the sense that further minimization would not be expected to lower the energy of such combinations (even though such combinations had moderately low energy, e.g., −15 to −16 kcal/mol), and then carrying out ten cycles of energy minimization. In nearly every case (see those indicated by superscript *d* in Table V for examples), the energy was reduced only if at least one residue altered from its starting conformation to a different conformation (see arrows in Table V). In particular, it was obvious that Thr preferred the α_R and C_7^{eq} conformations by several kilocalories per mole over the β_1 or β_2 conformations, in agreement with the observations after only one cycle of minimization. Each conformer of Table V with an energy lower than −16 kcal/mol was carried through further cycles of energy minimization, and the five lowest-energy conformations are shown in Table VI. In each case, minimization was carried out until the total energy did not change by more than 0.01 kcal/mol; this usually required ~30 cycles of energy minimization.

A partial test for convergence and flexibility was carried out for all low-energy conformations. For this purpose, each variable dihedral angle was incremented by a few degrees around the minimum-energy position, and the total energy was recalculated. In no case was a lower energy conformation found by this procedure. Although this does not guarantee that the conformation is at a minimum in the energy surface, it does show that no change in a *single* variable will lower the energy. At the same time, this procedure provides information about the flexibility (or conformational entropy) of each low-energy conformation, i.e., about the energy required to produce a small deviation from the low-energy position. These results will be described for the lowest-energy conformers of each tetrapeptide.

Each minimum-energy conformer of each tetrapeptide was also examined to determine the magnitudes of various contact distances. In particular, the $C^{\alpha}_i \cdots C^{\alpha}_{i+3}$ distance was computed, since a chain reversal or bend corresponds to a value of less than ~7 Å for this distance. Further, the $C^{\alpha}_{i-1} \cdots C^{\alpha}_{i+4}$ distance (i.e., the distance between the two terminal methyl carbon atoms) was computed, since a close van der Waals contact of these two groups (at ~4 Å) would be expected to give rise to a significant NOE. Of course, it is possible for a tetrapeptide to exist in a bend conformation, and yet *not* exhibit an NOE. We will discuss each tetrapeptide separately, and consider whether the molecular flexibility is sufficient to lead to some small amount of $C^{\alpha}_{i-1} \cdots C^{\alpha}_{i+4}$ contact. For peptide L, it appears that no one minimum-energy conformation is of sufficiently low energy to dominate in the population of conformers in solution. Thus, it appears that several conformations of this peptide probably exist in solution.

Structure A of Table VI, which has the lowest energy, is shown in Figure 4. This puckered form [which is a bend, since the $C^{\alpha}_i \cdots C^{\alpha}_{i+3}$ distance is <7 Å (i.e., 6.3 Å, as shown in Table VI)] *does not* correspond to a combination of single-residue minima. In particular, ϕ_i and ψ_i of glycine fall in the "bridge" region of a ϕ–ψ map. This is a very important result, showing that "bridge" region conformations, which are frequently observed in proteins,[39] can result from the potential energy functions used here, even though

Table V
Starting Conformations for Ac-Gly-Thr-Asp-Lys-NMe (L)

No. of different starting conformations[a]	Residue				Energy,[c] kcal/mol
	Gly ϕ_i, ψ_i	Thr ϕ_{i+1}, ψ_{i+1}	Asp ϕ_{i+2}, ψ_{i+2}	Lys ϕ_{i+3}, ψ_{i+3}	
7	β_1, 2_1, $2_1'$, α_R, α_L, C_7^{eq}, C_7^{ax}	C_7^{eq}	C_7^{eq}	β_1	−15
7[b]	β_1, 2_1, $2_1'$, α_R, C_7^{eq}, C_7^{ax}	C_7^{eq}	C_7^{eq}	C_7^{eq}	−16
4	β_1, α_R, $2_1'$, C_7^{eq}	C_7^{eq}	C_7^{eq}	α_R	−15
11[b]	α_R, C_7^{eq}, C_7^{ax}, β_1, 2_1	α_R	α_R	C_7^{eq}	−17
4	2_1, β_1, α_R, C_7^{eq}	α_R	α_R	α_R	−15
5[b]	$2_1'$, α_R, C_7^{eq}	α_R	α_R	β_1	−16
4[b]	C_7^{eq}, β_1, α_R	α_R	C_7^{eq}	α_R	−16
5[b]	α_R, β_1, C_7^{ax}	α_R	C_7^{eq}	β_1	−16
7[b]	β_1, C_7^{ax}, C_7^{eq}, α_R	α_R	C_7^{eq}	C_7^{eq}	−16
3	C_7^{ax}, 2_1, α_R	α_R	β_1	C_7^{eq}	−17
2	2_1, α_R	α_R	β_1	α_R	−16
4	α_R, β_1, 2_1, C_7^{ax}	α_R	β_1	β_1	−15
1	2_1	α_R	β_1	C_7^{eq}	−14
1	α_R	α_R	C_7^{eq}	α_R	−14
1	C_7^{eq}	β_1	α_R	α_R	−5
2	α_R, C_7^{ax}	β_1	α_R	C_7^{eq}	−15
1	C_7^{ax}	β_1	α_R	α_R	−5
2	α_R, C_7^{ax}	β_1	C_7^{eq}	α_R	−15
2	C_7^{eq}, C_7^{ax}	β_1	C_7^{eq}	C_7^{eq}	−15
1	C_7^{ax}	β_1	β_1	α_R	−14
2	C_7^{ax}, α_R	β_1	β_1	C_7^{eq}	−14
1	β_1	$\beta_1 \rightarrow \beta_2$	β_2	$2_1 \rightarrow C_7^{eq}$	−13[d]
1	C_7^{ax}	α_R	C_7^{eq}	C_7^{eq}	−18[d]
1[e]	α_R'	α_R	$\beta_2 \rightarrow 2_1$	α_R'	−21[d]
1[e]	α_R	α_R	2_1	2_1	−19[d]
1	$\beta_2' \rightarrow C_7^{ax}$	β_2	$\beta_2' \rightarrow \alpha_R$	$C_7^{eq} \rightarrow 2_1$	+8[d,f]
1[e]	$C_7^{ax} \rightarrow \beta_2'$	$2_1 \rightarrow C_7^{eq}$	$2_1 \rightarrow \beta_1$	C_7^{eq}	−19[d]
1[e]	β_2'	$\beta_1 \rightarrow C_7^{eq}$	$2_1 \rightarrow \beta_1$	β_1	−19[d]
1	C_7^{ax}	β_2	α_L	$\beta_1' \rightarrow \alpha_L$	−15[d]
1	$\alpha_R \rightarrow \alpha_R'$	C_7^{eq}	β_1	$\beta_1 \rightarrow \beta_2$	−17[d]
1[e]	α_R	C_7^{eq}	$\beta_1 \rightarrow \alpha_R'$	α_R	−21[d]

[a] A total of ~80 starting conformations was used. In some cases,[b] the initial backbone conformations were the same, but the initial side-chain conformations differed. [b] More than one initial side-chain conformation was considered for these backbone conformations. [c] This is the conformational energy of the whole tetrapeptide after one cycle of minimization, starting from an initial conformation. The energy reported corresponds to the last conformation listed for Gly, and is the lowest energy for that set of conformers. In many cases, the starting energy was positive. [d] This is the conformational energy of the whole tetrapeptide after ten cycles of minimization. These conformers were chosen for ten cycles of minimization because the results from one cycle of minimization indicated that their energies *might* be lowered by further minimization. In many cases, at least one residue changed from one single-residue starting position to another, upon minimization through ten cycles as indicated by the arrows. [e] These led to conformations B, E, C, D, A (in the order given) of Table VI, after a total of ~30 cycles of minimization. [f] This was a test case to determine whether bad steric interactions could be relieved by further cycles of minimization. As can be seen, the energy was not lowered significantly.

they are not low-energy conformations of end-blocked *single* residues. The values of ϕ_{i+1} and ψ_{i+1} are typical ones for C_7^{eq}, but those for ϕ_{i+2} and ψ_{i+2} are α'_R values, which again are not those found as the lowest-energy conformations of an end-blocked *single* residue (although a moderately high-energy minimum does exist in this region[9]). ϕ_{i+3} and ψ_{i+3} are near α_R, but somewhat displaced from the single-residue α_R conformation (viz., $\phi = -72°$ and $\psi = -55°$).[9] The hydrogen-bonding network, shown by dotted lines in Figure 4, is extensive. The fact that this low-energy conformation (obtained by energy minimization) differs from single-residue minimum-energy conformations served as a guide in the examination of the three other tetrapeptides.

Structure B of Table VI, which is next in energy, is shown in Figure 5. In this structure, which is a type I bend,[4] the threonine and aspartic acid side chains form a strong hydrogen bond which helps lock this structure in its calculated conformation. This conformation might be the more stable one if the side-chain carboxyl group of Asp were in the COO^- form. Ionization of the Asp side-chain carboxyl group could still occur in the conformation of Figure 5. This would not be true of structure A (Figure 4), where the OH of the carboxyl group is involved in the hydrogen bond. These differences must be kept in mind when these results are compared with the experimental ones, which were obtained in aqueous solution at a pH where the carboxyl group is ionized.

Structure C is close to a fully C_7^{eq} conformation, except for Gly which is partially extended. Structure D, which is ~3 kcal/mol less stable than structure A, is a more extended structure. Structure E is a type III bend,[4] and has a hydrogen bond between the NH of Gly and the C=O of Lys, which helps stabilize this conformation.

Table VI
Minimum-Energy Conformations of Peptide L (Ac-Gly-Thr-Asp-Lys-NMe)

Dihedral angle, deg	Conformation[a]					Low-energy values[b] from Table II
	A	B	C	D	E	
ϕ_i	−86	−158	71	80	80	±60, ±135
ψ_i	−11	−65	169	−179	−80	
ω_i	−179	179	178	−179	177	
ϕ_{i+1}	−79	−66	−87	−98	−77	−90, −150
ψ_{i+1}	78	−42	90	81	−41	
ω_{i+1}	180	−175	−172	−176	−179	
χ^1_{i+1}	56	46	58	61	38	
$\chi^{2,1}_{i+1}$	66	88	62	65	163	
$\chi^{2,2}_{i+1}$	83	74	81	79	62	
ϕ_{i+2}	−173	−106	−96	−90	−75	−80, −160
ψ_{i+2}	−51	47	75	135	−37	
ω_{i+2}	173	−173	−174	180	−176	
χ^1_{i+2}	27	−50	−54	−50	−59	
χ^2_{i+2}	61	59	−69	−70	84	
ϕ_{i+3}	−86	−158	−88	−150	−153	−90, −150
ψ_{i+3}	−45	−57	77	134	98	
χ^1_{i+3}	−68	−173	−67	−165	−170	
			Energy, kcal/mol			
Electrostatic	−3.1	−1.3	−2.1	−1.4	−0.9	
Nonbonded	−20.5	−20.8	−19.3	−19.2	−19.6	
Torsional	0.4	0.7	0.6	0.9	0.3	
Total	−23.2	−21.4	−20.8	−20.5	−20.2	
			Contact distances, Å			
$C^\alpha_i \cdots C^\alpha_{i+3}$	6.3	5.0	9.0	9.5	4.9	
$C^\alpha_{i-1} \cdots C^\alpha_{i+4}$	9.4	7.7	12.8	13.9	6.3	

[a] Conformations A–E were obtained by further minimization from all starting conformations of Table V for which ten cycles of minimization had been carried out. [b] By "low-energy values", we mean those dihedral angles obtained from the NMR coupling constants which are compatible with low-energy conformations of a single residue.

None of the low-energy conformers of this tetrapeptide had a close contact between the side-chain COOH group of Asp and the side-chain NH_2 group of Lys; i.e., backbone conformations, which would bring these groups together in *this* tetrapeptide, are not of low energy. Thus, in *this particular* tetrapeptide sequence, a salt bridge between the Asp and Lys side chains would not be likely to occur.

The energies of many other conformations were also minimized extensively, but those listed were the only ones which moved toward low-energy conformations upon minimization. All other conformations remained ~5 kcal/mol or more higher in energy than conformation A.

Figure 6 shows how the energy increases as the conformation of structure A (Table VI) of peptide L departs from its minimum-energy dihedral angles. It is apparent from this figure that the dihedral angles ω_i, ω_{i+1}, ϕ_i, ϕ_{i+2}, ψ_{i+1}, and χ^1_{i+2} are relatively stiff, while all the other dihedral angles are moderately soft, or easily varied by several degrees about their minimum-energy positions. In conformation A, ω_i and ϕ_i are directly involved in the formation of a hydrogen bond between the terminal acetyl carbonyl and the Asp side-chain carboxyl O–H group, and any variation in these dihedral angles breaks this hydrogen bond. The dihedral angles ϕ_{i+1} and ψ_{i+1} are involved in a C_7^{eq} structure, and can also vary to break the acetyl–Asp hydrogen bond, as would a variation of χ^2 of Asp. It appears from Figure 6 that this conformation is quite flexible, as will be seen in Figures 8, 10, and 12 when compared with the other three tetrapeptides.

Further, from Figures 4 and 5, and Table VI, it is clear that the terminal methyl groups are not in contact; even if some brief contact were made, no one conformation is sufficiently stable to keep these methyls in contact long enough for an NOE to be observed. This is in agreement with the experimental data which gave no indication of an NOE for this tetrapeptide.

In structure B, which may be favored in the ionized state, the bend conformation brings the terminal methyl groups closer together than in structure A. However, from model building and from flexibility criteria, there appears to be no way for the terminal methyls to come in contact with each other. Thus, we would not expect to observe an NOE for structures A or B (similarly for structures C, D, and E), in all of which the contact between the end methyl groups is hindered.

The *low-energy* values of ϕ, deduced from the NMR coupling constants, are also given in Table VI for comparison. From these results it appears that structure A most closely agrees with the NMR data, with ϕ_i (of glycine) deviating the most from the NMR values. The three other ϕ's are very close (within about ±10°) to the NMR values. The problem of assigning the value of ϕ of glycine from the NMR data was discussed in section II.

No other lower energy minima were found for this tetrapeptide when ten cycles of minimization were carried out, starting from all the bend types[4] I through III′. Molecular models were used in these cases to obtain optimum initial side-chain conformations, i.e., those without overlaps and with optimal hydrogen bonding between side chains.

Peptide N (Ac-Asp-Lys-Thr-Gly-NMe). The tetrapeptide N, having the sequence of residues 35–38 of α-chymotrypsin, was examined briefly in our earlier paper.[4] Using the procedure described for peptide L, we again selected single-residue minimum-energy conformations as

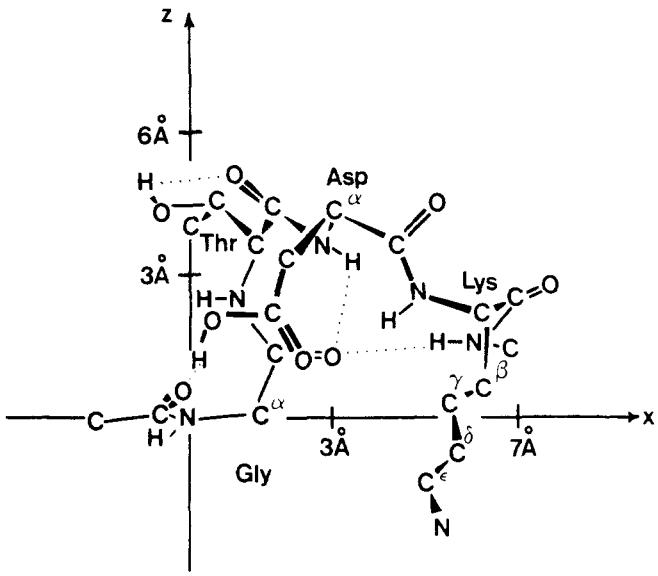

Figure 4. Conformation A (Table VI) of peptide L. Hydrogen bonds are indicated by dotted lines. Backbone and side-chain aliphatic hydrogen atoms have been omitted for clarity.

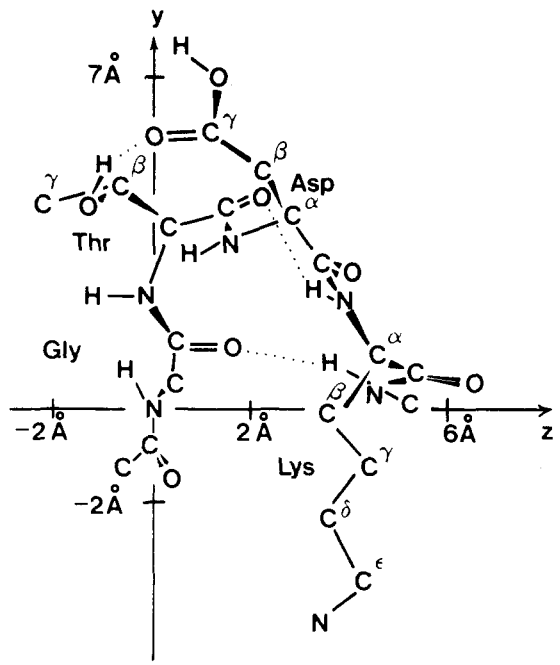

Figure 5. Conformation B (Table VI) of peptide L. Hydrogen bonds are indicated by dotted lines. Backbone and side-chain aliphatic hydrogen atoms have been omitted for clarity.

initial ones for minimization. However, instead of using so many initial conformations for the first cycle of energy minimization, a more limited set (~35 conformations) was selected, and these were each subjected immediately to ten cycles of minimization. These were selected on the basis of the results for peptide L; i.e., lower-energy puckered structures rather than higher-energy extended ones were taken; also, low-energy conformations of Asp-Lys (obtained from an analysis of Table V) were used. It is valid to use this more limited set of starting conformations if the molecule is sufficiently flexible to move from one region of conformation space to another in the minimization procedure. For peptides N, H_1, and H_2, changes of up to 50° in dihedral angles were obtained during minimization and, in some cases, the conformations of single residues did move from one region of conformation space to another. With

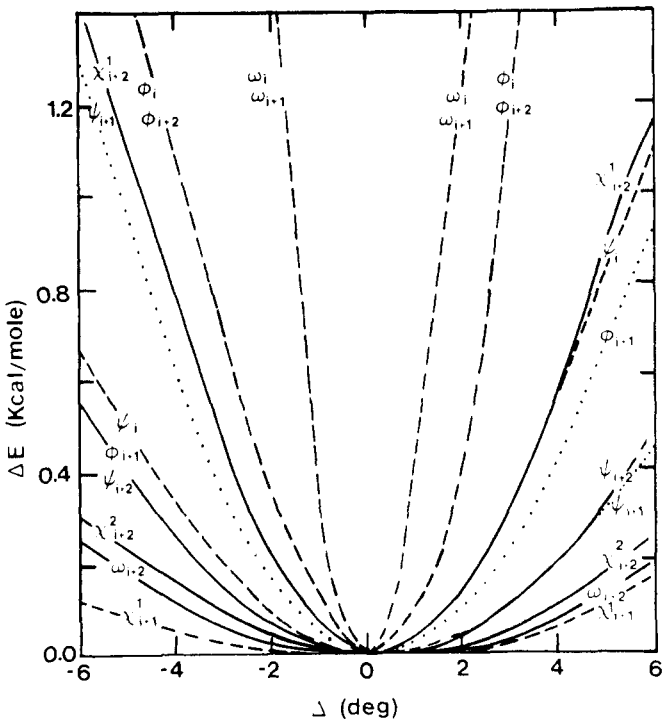

Figure 6. Variation in energy as the conformation of structure A (Table VI) of peptide L departs from its minimum-energy dihedral angles. The energy at the lowest-energy minimum was −23.2 kcal/mol. Those dihedral angles not shown (viz., $\chi^{2,1}{}_{i+1}$, $\chi^{2,2}{}_{i+1}$, ϕ_{i+3}, ψ_{i+3}, $\chi^1{}_{i+3}$) exhibit variations in energy similar to that shown for $\chi^1{}_{i+1}$. The curve for ϕ_{i+1} on the left falls under that for ψ_{i+2}.

this movement, fast convergence to low-energy structures was obtained. Thus, although only ~35 starting conformations were taken, a considerable portion of conformation space was covered in the subsequent minimization.

The results for the lowest-energy conformations of this tetrapeptide are given in Table VII together with the dihedral angles from the X-ray structure and the values of ϕ from the NMR coupling constants.

Structure A was found to be the one of lowest energy, and no other conformation was found with an energy less than 4 kcal/mol above this one. This result differs from that observed for peptide L, where several conformations in the range of −20 kcal/mol (but none as low as −30 kcal/mol) were found. Figure 7 shows the intricate network of hydrogen bonds formed in this conformation, and accounts for the very low nonbonded energy (which includes the hydrogen-bonding contribution[10]). Variation of χ^1 of Lys to the −60° region raises the energy by only ~0.5 kcal/mol, showing that the conformation of the Lys side chain is not important for the stability of this conformation. Structure A is similar to a type I bend,[4] except that ψ_{i+2} is 82° instead of the standard value[4] of 0°; i.e., it is really a type IV bend.[4] The X-ray structure is a type III bend. The difference between structure A and the minimized X-ray structure (see Table VII) is not unexpected since the chain observed by X-ray diffraction is not broken but continuous, and the residues further along the chain, in both directions, would influence the conformation. The values of ϕ of structure A agree quite well with those obtained from the NMR coupling constants, except for ϕ of Gly, as found for peptide L.

Figure 8 shows how the energy of conformation A of peptide N increases when the structure departs from its minimum-energy conformation. It is apparent that this conformation is considerably stiffer (with respect to changes in dihedral angle) than that found for peptide L, the reason being the more extensive network of hydrogen bonds. In

Table VII
Minimum-Energy Conformations of Peptide N (Ac-Asp-Lys-Thr-Gly-NMe)

Dihedral angle, deg	Conformation A	B	C	X-Ray[a]	X-Ray (min.)[b]	Low-energy values[c] from Table II
ϕ_i	−161	−78	−78	−37	−77	−80, −160
ψ_i	107	95	99	−167	93	
ω_i	178	176	−179	176	179	
$\chi^1{}_i$	172	172	−179	−92	−64	
$\chi^2{}_i$	47	−126	−85	23	88	
ϕ_{i+1}	−72	−71	−94	−113	−77	−85, −155
ψ_{i+1}	−38	−46	82	−11	−30	
ω_{i+1}	178	−179	172	−178	−177	
$\chi^1{}_{i+1}$	−168	−170	−168	−52	−70	
ϕ_{i+2}	−76	−84	−79	−92	−70	−80, −160
ψ_{i+2}	82	94	87	−11	−37	
ω_{i+2}	177	−179	−179	171	−178	
$\chi^1{}_{i+2}$	34	56	−67	57	35	
$\chi^{2,1}{}_{i+2}$	66	64	87	166	164	
$\chi^{2,2}{}_{i+2}$	56	80	41	63	60	
ϕ_{i+3}	114	140	84	116	−151	±55, ±140
ψ_{i+3}	−75	−70	−74	7	83	
			Energy, kcal/mol			
Electrostatic	−3.3	−3.1	−1.9		−1.1	
Nonbonded	−26.8	−22.7	−20.9		−14.8	
Torsional	0.1	0.1	0.6		0.2	
Total	−30.0	−25.7	−22.2	200+	−15.7	
			Contact distances, Å			
$C^\alpha{}_i \cdots C^\alpha{}_{i+3}$	6.3	6.3	9.0		5.4	
$C^\alpha{}_{i-1} \cdots C^\alpha{}_{i+4}$	5.8	8.6	11.0		7.5	

[a] Reference 8. [b] This structure was obtained by starting from the X-ray structure[8] and minimizing the conformational energy. There are small differences in most dihedral angles [viz., ±3°, except for ψ_{i-1} which was +91° in ref 4 and −30° here, $\chi_{i+1}{}^1$ which was −176° in ref 4 and −70° here, ψ_{i-2} which was −56° in ref 4 and −37° here, and ϕ_{i-3} which was 75° in ref 4 and −151° here] between the energy-minimized structure given here and that reported in ref. 4. These differences arise from some changes in the potential energy function.[10] [c] See footnote *b* of Table VI.

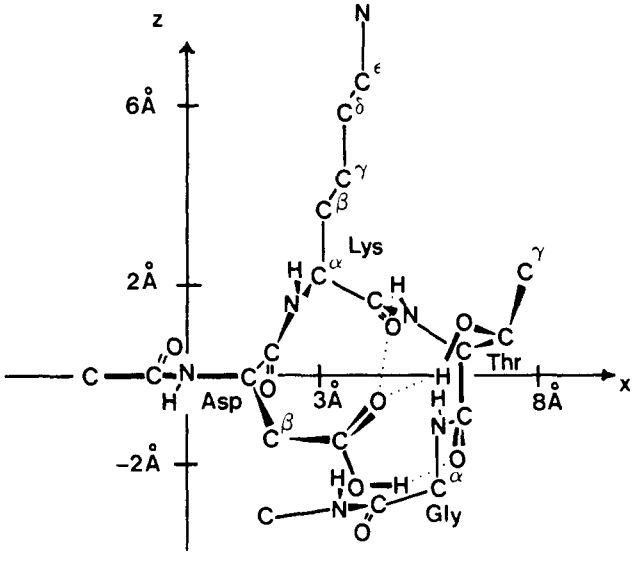

Figure 7. Conformation A (Table VII) of peptide N. Hydrogen bonds are indicated by dotted lines. Backbone and side-chain aliphatic hydrogen atoms have been omitted for clarity.

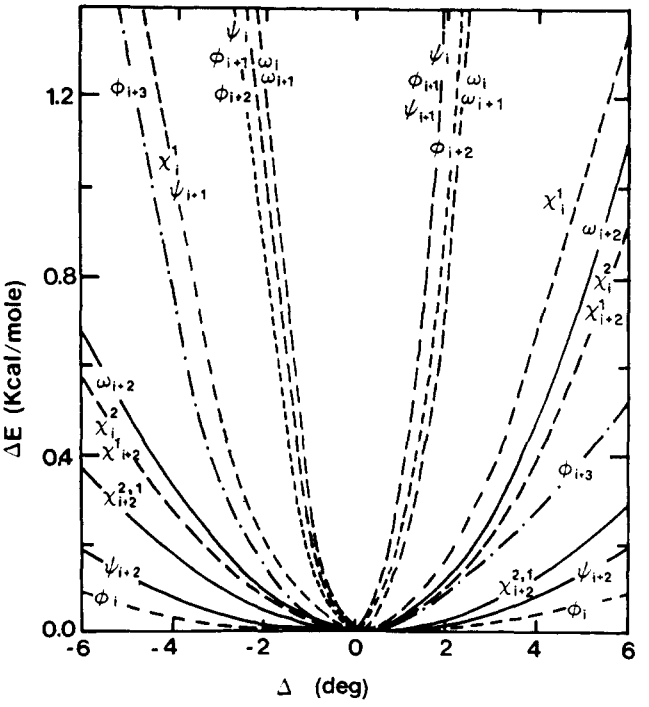

Figure 8. Variation in energy as the conformation of structure A (Table VII) of peptide N departs from its minimum-energy dihedral angles. The energy at the lowest-energy minimum was −30.0 kcal/mol. Those dihedral angles not shown exhibit the following variations in energy: $\chi^1{}_{i+1} \simeq \chi^{2,2}{}_{i+2} \simeq \phi_i$ and $\psi_{i+3} \simeq \psi_{i+2}$.

particular, the interaction between the side chains of Asp and Thr is found to be very favorable, helping to lock the complete conformation into its low-energy form. It is also of interest that, in this tetrapeptide, the ends of the molecule are in general much easier to rotate through a few degrees (see ϕ_i and ψ_{i+3} in Figures 8, 10, and 12), while the

Table VIII
Minimum-Energy Conformations of Peptide H_1 (Ac-Asp-Lys-Gly-Thr-NMe)

Dihedral angle, deg	Conformation A	B	C	D	Low-energy values[a] from Table II
ϕ_i	−162	−73	−159	−158	−80, −160
ψ_i	136	119	130	138	
ω_i	178	178	−175	−178	
$\chi^1{}_i$	169	174	−167	−165	
$\chi^2{}_i$	71	53	−87	−89	
ϕ_{i+1}	−147	−92	−79	−79	−80, −160
ψ_{i+1}	131	70	−46	−53	
ω_{i+1}	180	−177	−177	−179	
$\chi^1{}_{i+1}$	−165	−68	−170	−171	
ϕ_{i+2}	88	75	163	164	±50, ±145
ψ_{i+2}	−65	−89	61	59	
ω_{i+2}	179	176	179	−177	
ϕ_{i+3}	−65	−120	−168	−95	−90, −150
ψ_{i+3}	152	76	131	82	
$\chi^1{}_{i+3}$	34	64	−95	62	
$\chi^{2,1}{}_{i+3}$	78	65	66	64	
$\chi^{2,2}{}_{i+3}$	55	77	47	81	
		Energy, kcal/mol			
Electrostatic	−2.4	−3.6	−2.2	−1.6	
Nonbonded	−26.0	−22.6	−22.9	−19.2	
Torsional	0.2	0.2	0.2	0.1	
Total	−28.2	−26.0	−24.9	−20.7	
		Contact distances, Å			
$C^\alpha{}_i \cdots C^\alpha{}_{i+3}$	6.6	6.6	6.8	6.8	
$C^\alpha{}_{i-1} \cdots C^\alpha{}_{i+1}$	11.0	8.3	11.5	13.2	

[a] See footnote *b* of Table VI.

dihedral angles of the $(i + 1)$th and $(i + 2)$th residues are generally much more rigidly fixed. This flexibility of the end groups is most likely responsible for the low values found in the NOE experiments, the middle part of the molecule being fairly rigidly held in a puckered conformation, thus allowing the end methyl groups to come together and give a small, but significant NOE. The short contact distance between the end methyl groups of conformation A (Table VII) suggests the possibility of observing an NOE.

Structure B, which involves primarily a 180° variation in χ^2 of Asp, illustrates the decrease in stability when the side-chain carboxylic acid group is not hydrogen bonded (ϕ_i and ψ_i also differ). Structure B, which also agrees with the NMR coupling constant data, could be the preferred conformation in solution at the pH of the experimental study (i.e., if the carboxyl group were in the COO^- form). However, even structure B is still of considerably lower energy (by ~3.5 kcal) than the next best conformation (C), but is not as compatible with the NOE as is conformation A.

Structure C is made up of a series of C_7^{eq} conformations, stabilized again by a favorable carboxyl side-chain interaction of Asp with the hydroxyl group of Thr. In this conformation, the end methyl groups are very far apart, and would not contribute to the NOE.

The minimum-energy structure obtained from the starting X-ray dihedral angles is not of low energy, and deviates in many respects from the sterically hindered X-ray structure. This result is not surprising since longer segments (up to nine in length) are required[40] in order that an oligopeptide have the same conformation as it does in a native protein.

Peptide H_1 (Ac-Asp-Lys-Gly-Thr-NMe). Peptide H_1, like H_2, has Gly in position $i + 2$, which gives it a high probability of forming a bend conformation in a protein.[3] Also, like peptide N, it has the Asp-Lys sequence in positions i and $(i + 1)$. We have used low-energy conformations of Asp-Lys, as found in the calculations on peptides L and N, as starting points for all these residues, as well as single-residue values of ϕ and ψ (for all residues), as described for peptide L; i.e., ~30 starting conformations were taken through ten cycles of minimization, and the ten lowest-energy structures were then minimized further through 10–30 cycles, or until the energy changed by less than 0.01 kcal/mol. Table VIII gives the low-energy conformations computed; no other comformations in the energy range shown were found, even when bends of types I–III were included.

Structure A (Table VIII) of peptide H_1 is shown in Figure 9. The hydrogen-bonding network is somewhat different from that found for peptides L and N in that the side chain of Asp, interacting from the ith position (as in peptide N), does not interact with the Gly residue in position $i + 2$, but now interacts with the hydroxyl group of Thr [in position $(i + 3)$]. In order for this to occur, the bend or pucker conformation must be somewhat more open than that found for peptide N. This is shown by the β conformation of the Asp and Lys residues. In this case, the carboxylic acid group comes back to hydrogen bond with the Thr backbone carbonyl as well as with the side-chain hydroxyl group. The C_7^{ax} conformation of the Gly residue allows this to occur, so that this conformation can be a low-energy one. However, a residue with a C^β side chain in position $(i + 2)$ would not allow $i \rightarrow i + 3$ interactions of this type to occur.

Structure B is only ~2 kcal/mol less stable than A, and has a different network of hydrogen bonds. In this case, the Asp carboxylic acid group interacts strongly with the carbonyl of Gly and with the N-terminal amide. The Thr hydroxyl interacts only with the Thr carbonyl and not with

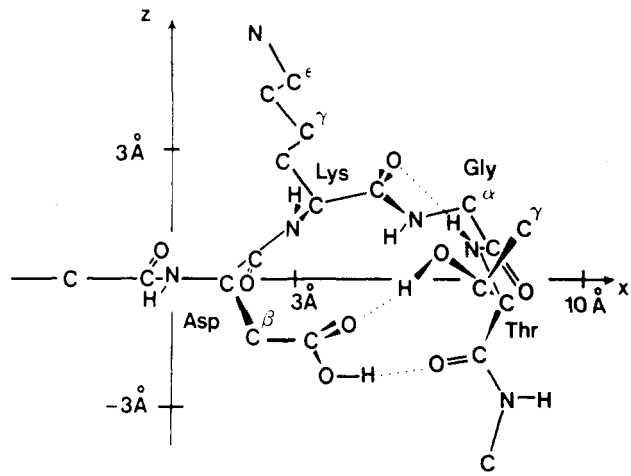

Figure 9. Conformation A (Table VIII) of peptide H_1. Hydrogen bonds are indicated by dotted lines. Backbone and side-chain aliphatic hydrogen atoms have been omitted for clarity.

the Asp side chain. The possibility that the terminal methyl groups interact is somewhat better for conformation B than A, as can be seen from the contact distances given in Table VIII.

Figure 10 shows the flexibility allowed around the minimum-energy conformation, A. As was found for peptide N, this molecule is quite rigid, with only the end residues being free to move.

Peptide H_2 (Ac-Lys-Thr-Gly-Asp-NMe). This tetrapeptide, like H_1, has Gly in position $(i + 2)$. The best Lys-Thr conformations of peptide N, as well as single-residue conformations (for all residues), were taken as starting conformations for energy minimization; i.e., 50 starting conformations were initially taken through one cycle of minimization, and the ~15 lowest-energy structures were then minimized further until the energy changed by less than 0.01 kcal/mol (i.e., ~30 cycles of minimization), as described for peptide L. The minimum-energy results are given in Table IX.

Structure A (Table IX) of peptide H_2 is seen to have an energy of −31 kcal/mol, close to that of structure A of peptide N. Also, as found for peptide N, no other conformation with an energy close to −31 kcal/mol was located (except for another conformation like A, but with χ^1 of Lys at −164°, and an increase in energy of ~0.2 kcal/mol). Structure A is similar to a type II bend[4] (but, more properly, a type IV bend[4]). The extensive hydrogen-bonding network is shown in Figure 11.

Similar to peptide N, Asp and Thr are separated by one residue (in this case, Gly), and this particular sequence appears to be most favorable for side-chain interactions between the Thr hydroxyl and Asp carboxylic acid groups. The end methyl groups are sufficiently close in this conformation to lead to a small NOE, in agreement with the experimental data. The good agreement with the ϕ values deduced from NMR coupling constants is seen in Table IX.

Structure B is closely similar to A, except for ψ_i. The result of this change is to break a hydrogen bond (—OH···O=C) from the carboxylic acid OH of Asp to the backbone carbonyl of the N-acetyl group. Also, this change destroys the C_7^{eq} conformation at the Lys residue. The close contact between the end methyls indicates that an NOE might be observed if this structure exists to an appreciable extent in solution. In structure C, the χ^1 value of Asp is −55°, which turns the carboxylic acid group away from the Thr OH and other hydrogen-bonding atoms. The loss of energy (~7 kcal/mol) is considerable, but this conformation may be a very probable one in water solution. The reason for this is

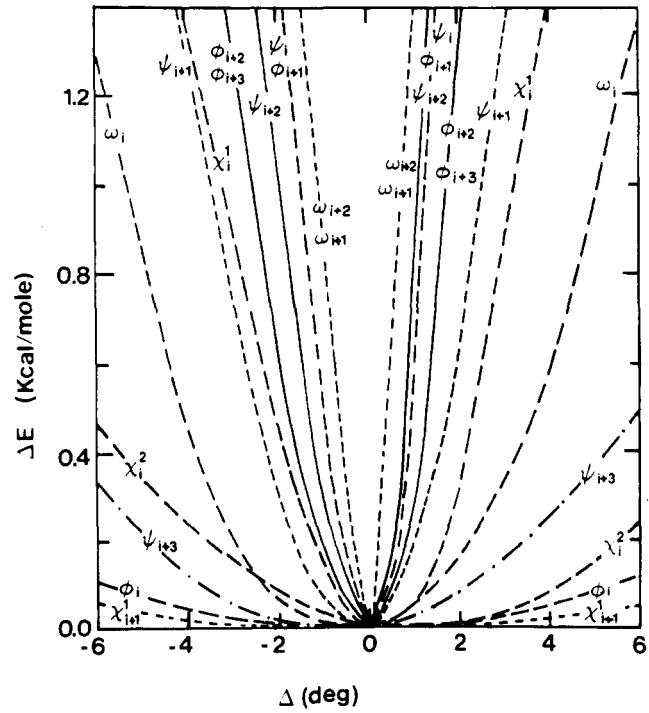

Figure 10. Variation in energy as the conformation of structure A (Table VIII) of peptide H_1 departs from its minimum-energy dihedral angles. The energy at the lowest-energy minimum was −28.2 kcal/mol. Those dihedral angles not shown exhibit the following variations in energy: $\chi^1{}_{i+3} \simeq \chi^2{}_i$ and $\chi^{2,1}{}_{i+3} \simeq \chi^{2,2}{}_{i+3} \simeq \chi^1{}_{i+1}$.

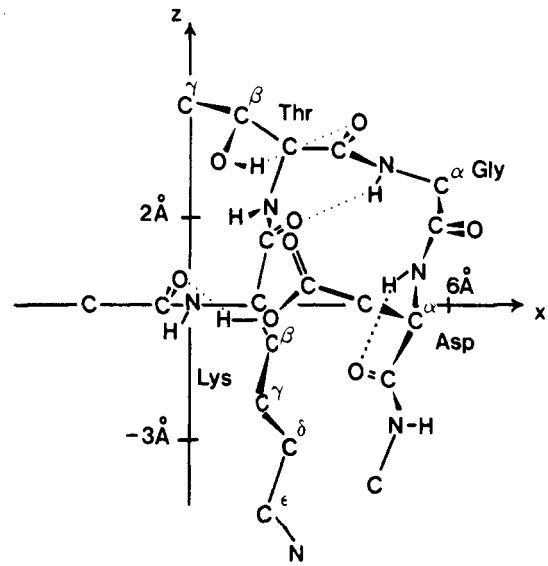

Figure 11. Conformation A (Table IX) of peptide H_2. Hydrogen bonds are indicated by dotted lines. Backbone and side-chain aliphatic hydrogen atoms have been omitted for clarity. There is also a hydrogen bond (not shown here) between the side-chain OH of Thr and the side-chain C=O of Asp.

the possible pH-induced ionization of the carboxylic acid group, which, upon hydration, could force it out into the solvent. Structure C is also similar to a type II bend,[4] with Gly in the α_L conformation, and its ϕ values agree with those deduced from NMR coupling constants.

Figure 12 shows the degree of flexibility in conformation A, the ends and some of the side chains being the least rigidly held.

IV. Discussion

Before discussing the experimental and computational

Table IX
Minimum-Energy Conformations of Peptide H_2 (Ac-Lys-Thr-Gly-Asp-NMe)

Dihedral angle, deg	Conformation A	B	C	D	Low-energy values[a] from Table II
ϕ_i	−89	−80	−85	−70	−90, −150
ψ_i	67	−54	80	−45	
ω_i	178	176	178	180	
$\chi^1{}_i$	−69	−166	−73	−170	
ϕ_{i+1}	−88	−133	−82	−73	−80, −160
ψ_{i+1}	68	53	69	−40	
ω_{i+1}	−172	−173	−171	−178	
$\chi^1{}_{i+1}$	39	45	40	34	
$\chi^{2,1}{}_{i+1}$	60	68	63	92	
$\chi^{2,2}{}_{i+1}$	61	−166	62	56	
ϕ_{i+2}	64	71	63	133	±60, ±135
ψ_{i+2}	44	31	48	73	
ω_{i+2}	−178	178	−174	−171	
ϕ_{i+3}	−162	−166	−144	−150	−85, −155
ψ_{i+3}	157	145	151	141	
$\chi^1{}_{i+3}$	53	58	−55	−58	
$\chi^2{}_{i+3}$	−103	−84	118	91	
		Energy, kcal/mol			
Electrostatic	−2.8	−0.6	−1.2	−0.2	
Nonbonded	−28.6	−24.2	−23.0	−22.0	
Torsional	0.4	0.4	0.7	0.8	
Total	−31.0	−24.4	−23.5	−21.4	
		Contact distances, Å			
$C^\alpha{}_i \cdots C^\alpha{}_{i+3}$	5.0	5.7	4.8	7.3	
$C^\alpha{}_{i-1} \cdots C^\alpha{}_{i+4}$	8.1	4.6	9.3	6.1	

[a] See footnote *b* of Table VI.

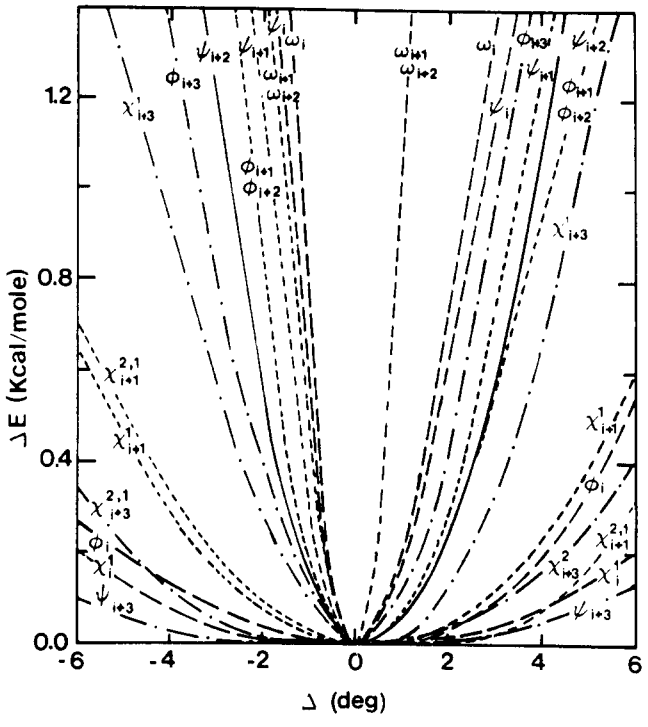

Figure 12. Variation in energy as the conformation of structure A (Table IX) of peptide H_2 departs from its minimum-energy dihedral angles. The energy at the lowest-energy minimum was −31.0 kcal/mol. One dihedral angle ($\chi^{2,2}{}_{i+1}$, not shown) exhibits a variation in energy similar to ψ_{i+3}.

results, we should anticipate that each tetrapeptide may be flexible, and may exist in a distribution of conformations in aqueous solution. Thus, the NMR and CD data would represent average properties over the distributions, while the computational results would provide information about the relative energies of various conformations within each distribution.

The deuterium–hydrogen exchange experiments and temperature dependence of the chemical shift provide no evidence for strong intramolecular hydrogen bonding, i.e., for hydrogen bonds in which the proton would be shielded from the solvent. This result is not unexpected since these (short) linear tetrapeptides are probably considerably more flexible than most cyclic peptides[41–44] and (longer) helical segments from proteins,[45] in which some protons are shielded. Because of this flexibility, the lifetime of a hydrogen bond (on the NMR time scale) may be too short to be observed.

Because of the error involved in measuring a small difference between two large numbers (each of which involves experimental error), the NOE experiments provide only a marginal indication that the terminal methyl groups may be near each other in some fraction of the total number of conformations for peptides N, H_1, and H_2, but not L. The proximity of these terminal methyl groups would indicate the existence of bend conformations in some of the conformations. Contact distances of the end methyl groups as short as 4.6 and 5.8 Å were computed for individual conformations of tetrapeptides H_2 and N, respectively, which may be compatible with the NOE experiments. However, the flexibility inherent in these tetrapeptides may also be important in determining whether the terminal methyl groups approach each other. The values of ϕ, computed for the low-energy structures, are, in general, in good agreement with those deduced from the NMR data.

The CD spectra indicate that peptide L differs from the other three, the latter exhibiting noticeable changes in

spectra between 5 and 26°. The spectra of N, H_1, and H_2 at 5°, showing a positive peak ~215 nm and negative peaks ~225 and ~198 nm, bear some resemblance to those computed by Woody[46] for various types of bends (see, e.g., his Figure 3b). Figure 3 of our paper thus provides some indication that a detectable fraction of the total number of conformations of peptides N, H_1, and H_2 may exist as bends at low temperature, and that the concentrations of such bend conformations are lower at 26°. The spectrum of peptide L provides no indication of a detectable amount of bend conformations, even at 5°.

The CD spectra of peptides N and H_2 (Figure 3) resemble those obtained by Mattice and Harrison[47] for Ac-L-Ala-L-Ala-L-Ala-L-Ala-OMe in water at 15 and 75°, respectively. Several years ago, Poland and Scheraga[48] discussed the stabilization of bend structures in oligo-alanines by hydrophobic bonds. In the absence of water, a type I bend in tetraalanine was computed[4] to be 1.0 and 0.8 kcal higher in energy than the C_7^{eq} and α_R structures, respectively.[49] However, following Poland and Scheraga,[48] two hydrophobic bonds can exist in the type I bend conformation of Lewis et al.,[4] viz., one between the side-chain methyls of Ala-1 and Ala-4, and another between the side-chain methyls of Ala-2 and Ala-3. Since *each* of these hydrophobic bonds can provide a stabilization *free* energy[50] of 0.7 kcal, it is reasonable to expect that a significant fraction of the tetraalanine molecules can exist in the type I bend conformation[4] *in water*, stabilized by hydrophobic bonds.[48] If so, then the CD spectra obtained by Mattice and Harrison[47] for tetraalanine may provide additional evidence that peptides N and H_2, which exhibit similar CD spectra, may exist to some extent in a bend conformation at low temperature.

Thus, taken together, the NMR and CD data provide circumstantial evidence that a detectable amount of bend conformations exists in peptides N, H_1, and H_2, but not in peptide L. To this extent, the experimental results are consistent with the relative probabilities of occurrence of β bends in these peptides, cited in the introductory section. These results then suggest that the *tendency* toward bend formation exists, not only in the sequence Asp-Lys-Thr-Gly in α-chymotrypsin, but also in the oligomeric fragments having the same sequence and high-probability variants thereof. This would constitute an experimental verification of the concept[3] that the *tendency* toward bend formation arises primarily because of short-range interactions, and depends on both the amino acid composition and sequence.

The conformational energy calculations support the conclusions derived from the NMR and CD data. Peptide L appears to exist in many different conformations with no distinct preference for a bend. However, peptides N and H_2 (and, to a lesser extent, H_1) show a distinct preference for a lower-energy bend conformation (approximately −30 kcal/mol compared with approximately −23 kcal/mol). The preference for bends in N, H_1, and H_2, but not in L, is qualitatively consistent, not only with the experimental data, but also with the empirical probabilities[3] for bend formation.

A comparison of the conformations of tetrapeptide H_2 with that of a fully extended minimum-energy structure (i.e., β_1-β_1-β_1-β_1), whose energy was computed here to be only −11.5 kcal/mol, suggests that the interactions involving side chains stabilize bend conformations (side chain to side chain, and side chain to backbone), and account for the differences in stabilization energy in the tetrapeptides studied here. Those energetic factors which favor bend structures in tetrapeptides N, H_1, and H_2, but not L, are related to the amino acid sequence. These conclusions seem reasonable, despite the omission of an explicit treatment of solvent effects and ionizable side chains, for the same reasons that a similar comparison[28] of experimental and calculated results (in sarcosyl-sarcosine) led to consistent results.

If there were no energetic preferences for bend conformations in tetrapeptides N, H_1, and H_2, i.e., if bend conformations occurred by a random selection from the conformational space of the peptides, then all four tetrapeptides would have similar physical properties. We can make a very rough estimate of the number of conformations accessible to an end-blocked tetrapeptide by assuming that each amino acid residue of the tetrapeptide can exist in five discrete backbone conformational states, each of which is accessible with equal probability. There are 5^4 (or 625) possible conformations of each tetrapeptide, of which a bend conformation is probably only a small number of these. The different CD and NOE properties of tetrapeptides N, H_1, and H_2, on the one hand, and L, on the other, could not arise if the bend conformations represented a small fraction of the total number of conformations. Thus, energetic factors, as illustrated by the conformational energy calculations reported here, play an important role in increasing the tendency of N, H_1, and H_2 to adopt bend conformations. Such energetic factors are probably also involved in determining the tendency toward folding, observed in other tetrapeptides.[51]

APPENDIX

SYNTHESIS OF TETRAPEPTIDES

Peptide L

Z

Lys-NHCH$_3$ (I)

ε-Z-Lysine N-carboxyanhydride (7.1 g; 0.023 mole) was dissolved in dry THF (80 ml), and this solution was added dropwise to a stirred and cooled (0°–5°) solution of monomethylamine in THF-water (40 + 5 ml). After the addition of lysine N-carboxyanhydride was completed, the mixture was stirred for one hour while slowly warming it to room temperature. The solvent was evaporated and the residue was triturated with hexane. The residue (I) was dried under vacuum to give a semi-solid material. Yield 6.5 g; 96%.

OBz Z

BOC-Asp-Lys-NHCH$_3$ (II)

BOC(β-OBz)Asp (7.11 g; 0.022 mole) was dissolved in THF (150 ml). The solution was cooled to −20°C and stirred vigorously. To the solution was added N-methylmorpholine (2.64 ml; 0.022 mole) and isobutylchloroformate (2.86 ml; 0.022 mole). After three minutes, I (6.5 g; 0.022 mole) in THF (100 ml) was added. The mixture was stirred overnight while slowly warming to room temperature. The solvent was evaporated and the residue was dissolved in chloroform (500 ml). The chloroform layer was washed with 10% citric acid solution (1 × 150 ml), water (1 × 150 ml), $NaHCO_3$ solution (1 × 150 ml), and water (2 × 150 ml), respectively, and finally dried over $MgSO_4$ and evaporated. The residue (II) was triturated with ether, and recrystallized from chloroform–hexane to give 11.1 g; 84%, m.p. 146–147°C.

OBz Z

BOC-Thr-Asp-Lys-NHCH$_3$ (III)

II (9.0 g; 0.015 mole) was dissolved in ethyl acetate (60 ml), and a saturated solution of HCl in ethyl acetate (30 ml) was added to remove the BOC group. After twenty minutes, the mixture was evaporated and the residue crystallized under ether. The dipeptide hydrochloride was precipitated from methanol–ether.

The dipeptide hydrochloride was dissolved in DMF–THF (125 + 25 ml), neutralized with triethylamine, and BOC threonine pentachlorophenyl ester (8.0 g; 0.017 mole) was added. After three days at room temperature, the solvent was evaporated and the residue was dissolved in a small amount of methanol. This solution was kept in the cold room overnight, and the unreacted BOC threonine pentachlorophenyl ester precipitated out and was filtered off. The filtrate was evaporated, and the residue dissolved in ethyl acetate (600 ml). The organic layer was washed with 10% citric acid solution (1 × 200 ml), water (1 × 200 ml), $NaHCO_3$ solution (1 × 200 ml), and water (2 × 200 ml) respectively, dried over $MgSO_4$, and evaporated. The residue (III) was recrystallized twice from methanol–ether to give 8.0 g; 75%, m.p. 171–172°C.

OBz Z

BOC-Gly-Thr-Asp-Lys-NHCH$_3$ (IV)

The BOC group of III (2.10 g; 0.003 mole) was removed by HCl/ethyl acetate treatment in the same manner as for II. The hydrochloride was precipitated from methanol–ethyl acetate and filtered off.

The hydrochloride was dissolved in DMF (30 ml), neutralized with triethylamine and BOC Gly pentachlorophenyl ester (1.7 g; 0.004 mole) was added. After three days, the solvent was evaporated, and the residue was triturated under ether. The residue (IV) was precipitated from methanol–water and methanol–ether, respectively, to give 1.36 g; 60%, m.p. 187–188°C.

OBz Z

Ac-Gly-Thr-Asp-Lys-NHCH$_3$ (V)

IV (1.51 g; 0.002 mole) was dissolved in acetic acid (6 ml), and $BF_3O(Et)_2$ (1.5 ml) was added to remove the BOC group. After 10 minutes, the product was precipitated with anhydrous ether, filtered, washed with more ether, and dried over KOH.

The residue was dissolved in THF (100 ml) – water (10 ml) and cooled to 0–5°C. To the cooled and stirred mixture was added triethylamine (0.56 ml; 0.004 mole) and acetic anhydride (0.38 ml; 0.004 mole). After stirring for one hour at the same temperature, the mixture was stirred another two hours while slowly warming to room temperature. The solvent was evaporated

4

and the residue was precipitated from methanol-water and methanol-ether respectively, to give 1.15 g, (82% yield), m.p. 209-210°C.

Ac-Gly-Thr-Asp-Lys-NHCH$_3$ (VI)

V (0.7 g; 0.001 mole) was dissolved in acetic acid (10 ml) and diluted with water (10 ml) and methanol (100 ml). After adding 10% palladium-charcoal (200 mg), hydrogenation was carried out for two hours. The catalyst was filtered off and the solvent evaporated. The residue was purified by fractionation on column A twice and column B once to give 0.332 g, 70%, m.p. slowly decomposes above 180°C. $[\alpha]_D^{23}$ = -54.7° (C = 1, H_2O); R_F^1 = 0.10, R_F^2 = 0.26, R_F^3 = 0.36, R_F^4 = 0.19, chlorine and ninhydrine positive spot. Amino acid analysis: Asp = 1.00, Lys = 1.06, Gly = 0.95, Thr = 0.99, NH_2CH_3 = 1.04.

PEPTIDE N

Z-Gly-NHCH$_3$ (VII)

Z-Gly (5.86 g; 0.028 mole) was dissolved in THF (100 ml). The solution was cooled to -20°C and stirred while N-methylmorpholine (3.36 ml; 0.028 mole) and isobutylchloroformate (3.64 ml; 0.028 mole) were added to the solution. After five minutes, a saturated solution of monomethylamine in THF (20 ml) was added to the mixture. This mixture was then stirred for 3 hours while slowly warming to room temperature. The solvent was evaporated and the residue was fractionated on column D to give 5.6 g; 90%, m.p. 105-106°C.

7

The solution was cooled to -20°C and stirred vigorously. To the solution was added N-methymorpholine (6.0 ml, 0.05 mole) and isobutylchloroformate (6.5 ml; 0.05 mole). After three minutes, a saturated solution of monomethylamine in water (15 ml) was added and the mixture was stirred for one hour while slowly warming to room temperature. The solvent was evaporated and the residue was dissolved in ethyl acetate (500 ml). The ethylacetate layer was washed with a saturated NaCl-saturated $NaHCO_3$ (150 ml : 50 ml) solution and then with saturated NaCl solution (2 x 150 ml). It was finally dried over $MgSO_4$ and evaporated. Purification was accomplished by fractionation on column D using acetone as eluent instead of acetone:acetic acid 98:2. The product was an oil (9.4 g, 81%).

BOC-Gly-Thr-NHCH$_3$ (XIII)

BOC-Thr-NHCH$_3$ (3.02 g; 0.013 mole) was dissolved in TFA (12 ml) and after 45 minutes at room temperature, the TFA was evaporated. The residue was triturated under dry ether and the ether decanted off. This residue was dried over KOH in a vacuum desiccator overnight. Then it was dissolved in THF (50 ml) and neutralized with triethylamine and evaporated. The resulting residue was dissolved in THF (60 ml) and BOC glycine pentachlorophenyl ester (7.62 g; 0.018 mole) added. After four days at room temperature, the product precipitated out. It was crystallized from methanol-ether to give 3.01 g, 80%, m.p. 172-173°C.

10

rated. The residue was dissolved in ethyl acetate and washed with water several times. The ethyl acetate layer was dried over $MgSO_4$ and evaporated. The residue was crystallized from ethyl acetate-hexane to give 9.0 g; 89%, m.p. 102-103°C.

OBz
BOC-Gly-Asp-NHCH$_3$ (XIX)

The BOC group of compound XVIII (6.1 g; 0.018 mole) was removed by HCl/ethyl acetate as described under III. The residue was dissolved in THF (60 ml) and neutralized with N-methylmorpholine.

BOC-Gly (3.5 g; 0.02 mole) was dissolved in THF (250 ml). The solution was cooled to -20°C and stirred vigorously. To this solution was added N-methylmorpholine (2.4 ml; 0.02 mole) and isobutylchloroformate (2.6 ml; 0.02 mole). After five minutes, the (ß-OBz)Asp-NHCH$_3$ solution was added to the mixture and stirred for three hours while slowly warming to room temperature. The solvent was evaporated and the residue was dissolved in ethyl acetate (400 ml). The ethyl acetate layer was washed with 10% citric acid solution (1 x 150 ml), water (1 x 150 ml), $NaHCO_3$ solution (1 x 150), and water (2 x 150), respectively, and finally dried over $MgSO_4$ and evaporated. The residue was crystallized twice from ethyl acetate-ether to give 6.0 g; 85%, m.p. 101-102°C.

5

Z-Thr-Gly-NHCH$_3$ (VIII)

VII (5.1 g; 0.023 mole) was hydrogenated in methanol (100 ml) using 10% palladium-charcoal (0.5 g) for 2 hours. The residue was coupled with Z-Thr-pentachlorophenyl ester (12.54 g; 0.025 mole) in DMF (100 ml) for four days. The solvent was evaporated and the residue was triturated under ether. The residue was fractionated on column D to give 6.7 g; 90%, m.p. 150-151°C.

Z
BOC-Lys-Thr-Gly-NHCH$_3$ (IX)

VIII (6.47 g; 0.02 mole) was hydrogenated in methanol (150 ml) as under VIII. The residue was coupled with BOC(ε-Z)Lys pentachlorophenyl ester (14.46 g; 0.023 mole) in DMF-THF (50 + 10 ml) as described under III. Purification was carried out by precipitating the residue from methanol-ether and fractionation on column D. The product was crystallized from methanol-ether to give 8.2 g; 74%, m.p. 100-101°C.

OBz Z
BOC-Asp-Lys-Thr-Gly-NHCH$_3$ (X)

The BOC group of IX (5.1 g; 0.0092 mole) was removed with acetic acid-$BF_3O(Et)_2$ as described under V, and the residue was dissolved in DMF-THF (20 + 40 ml). The solution was neutralized with N-methylmorpholine, and BOC(ß-OBz)Asp pentachlorophenyl ester (6.30 g; 0.011 mole) was added. After three

8

Z
BOC-Lys-Gly-Thr-NHCH$_3$ (XIV)

The BOC group of XIII (1.01 g; 0.0035 mole) was removed by dissolving XIII in TFA (8 ml) and keeping the solution for 25 minutes at room temperature and then evaporating off the TFA. The residue was dissolved in THF (20 ml), neutralized with triethylamine and evaporated. This residue was dissolved in THF (60 ml) and BOC(ε-Z)lysine pentachlorophenyl ester (2.83 g; 0.0045 mole) added. After four days, the mixture was concentrated and the product precipitated with ether. The product was purified by fractionation on column D and crystallized from an acetic acid-ether-hexane mixture to give 1.3 g; 67%, m.p. 162-163°C.

OBz Z
BOC-Asp-Lys-Gly-Thr-NHCH$_3$ (XV)

The BOC group of XIV (0.77 g; 0.0014 mole) was removed by dissolving XIV in TFA-water (8 + 2 ml) and holding the solution for 20 minutes at room temperature. The TFA-water was then evaporated off and the residue was neutralized in methanol (20 ml) with triethylamine and evaporated. This residue was dissolved in THF (60 ml) and BOC(ß-OBz)Asp-pentachlorophenyl ester (1.03 g; 0.0018 mole) added. After three days the solvent was evaporated and the residue was precipitated from methanol-ether. Finally, the product was fractionated on column E to give 0.70 g; 66% m.p. 141-142°C.

11

OBz
BOC-Thr-Gly-Asp-NHCH$_3$ (XX)

The BOC group of XIX (6.3 g; 0.016 mole) was removed by acetic acid-$BF_3O(Et)_2$ treatment as described under V. The residue was dissolved in DMF (150 ml), neutralized with N-methylmorpholine and coupled with BOC-Thr-pentachlorophenyl ester (9.35 g; 0.02 mole). After three days, the solvent was evaporated and the residue was triturated under ether. The residue was dissolved in ethyl acetate (500 ml) and the organic layer was washed with saturated NaCl solution-10% citric acid solution (80 + 40 ml), saturated NaCl solution (100 ml), saturated NaCl solution-$NaHCO_3$ solution (80 + 40 ml) and saturated NaCl solution (100 ml), respectively, and finally dried over $MgSO_4$ and evaporated. The product was fractionated using column E to give 5.7 g; 72%, m.p. 150-151°C.

Z OBz
BOC-Lys-Thr-Gly-Asp-NHCH$_3$ (XXI)

Acetic acid-$BF_3O(Et)_2$ was used to remove the BOC group of XX (4.95 g; 0.01 mole) and the residue was coupled in DMF-THF (30 + 100 ml) with BOC(ε-Z)Lys pentachlorophenyl ester (8.17 g; 0.013 mole) as described under X. The residue was purified by precipitation from methanol-water and by recrystallization (twice) from methanol-ethyl acetate to give 4.0 g; 53%, m.p. 133-134°C.

6

days, the solvent was evaporated and the residue triturated under ether. The product was precipitated from methanol-water and then methanol-ether to give 5.57 g; 80%, m.p. 152-153°C.

OBz Z
Ac-Asp-Lys-Thr-Gly-NHCH$_3$ (XI)

The removal of the BOC group from compound X (3.41 g; 0.0045 mole) with acetic acid-$BF_3O(Et)_2$ and acetylation of the residue was carried out as described under V. The product was purified by precipitating it from acetic acid-water and then acetic acid-methanol-ether to give 2.11 g; 67%, m.p. 200-201°C.

Ac-Asp-Lys-Thr-Gly-NHCH$_3$ (XII)

Compound XI (0.70 g; 0.001 mole) was hydrogenated in acetic acid (5 ml) : water (5 ml) : methanol (50 ml) as described under VI, and the residue was fractionated on columns A, B and then A again to give 0.370 g; 78%, m.p. decomposed slowly above 185°C. $[\alpha]_D^{23}$ = -52.8° (C = 1, H_2O). R_F^1 = 0.1, R_F^2 = 0.26, R_F^3 = 0.36, R_F^4 = 0.19, chlorine and ninhydrin positive spot. Amino acid analysis: Asp = 0.99, Lys = 1.00, Gly = 1.01, Thr = 1.00, NH_2CH_3 = 1.07.

PEPTIDE H$_3$:

BOC-Thr-NHCH$_3$

BOC-Thr (10.96 g; 0.05 mole) was dissolved in THF (150 ml).

9

OBz Z
Ac-Asp-Lys-Gly-Thr-NHCH$_3$ (XVI)

The BOC group of XV (0.280 g; 0.00037 mole) was removed by acetic acid (3 ml) and $BF_3O(Et)_2$ (0.6 ml) as in V. The acetylation of the residue and the purification of the product was also carried out as described under V. Yield 0.202 g, 78%, m.p. 217-219°C.

Ac-Asp-Lys-Gly-Thr-NHCH$_3$ (XVII)

XVI (0.196 g; 0.00028 mole) was hydrogenated as described under VI. The residue was purified by fractionation on column C to give 0.108 g; 81%, m.p. slowly decomposes above 178°C. $[\alpha]_D^{23}$ = -46.9 (C = 1, H_2O). R_F^1 = 0.10; R_F^2 = 0.26; R_F^3 = 0.36; R_F^4 = 0.19. Chlorine and ninhydrin positive spot. Amino acid analysis: Asp = 1.01, Lys = 1.01, Gly = 1.01, Thr = 0.97, NH_2CH_3 = 0.93.

PEPTIDE H$_2$

OBz
BOC-Asp-NHCH$_3$ (XVIII)

BOC(ß-OBz)Asp (9.70 g; 0.03 mole) was dissolved in THF (300 ml). The solution was cooled to -20°C and stirred vigorously while N-methylmorpholine (3.6 ml; 0.03 mole) and isobutylchloroformate (3.9 ml; 0.03 mole) added. After three minutes, monomethylamine in THF (50 ml) was added to the reaction mixture. The mixture was stirred at -20°C for 20 minutes and then evapo-

12

Z OBz
Ac-Lys-Thr-Gly-Asp-NHCH$_3$ (XXII)

The BOC group of XXI (2.2 g; 0.0029 mole) was removed by acetic acid-$BF_3O(Et)_2$, and acetylation was carried out as described under V. The product was purified by precipitation from methanol-water and then crystallization from methanol-ether to give 1.7 g; 84%, m.p. 216-218°C.

Ac-Lys-Thr-Gly-Asp-NHCH$_3$ (XXIII)

XXII (0.350 g; 0.0005 mole) was dissolved in acetic acid (5 ml), water (5 ml) and methanol (50 ml). After addition of the catalyst (10% palladium-charcoal, 0.100 g), the hydrogenation was carried out for two hours as described under VI. The residue was fractionated using column A twice and then column B twice to give 0.154 g; 65%, m.p. decomposes slowly above 200°C. $[\alpha]_D^{23}$ = -29.9 (C = 1, H_2O). R_F^1 = 0.10, R_F^2 = 0.26, R_F^3 = 0.36, R_F^4 = 0.19, chlorine and ninhydrin positive spot. Amino acid analysis: Asp* = 1.00, Thr = 0.96, Gly = 1.01, Lys = 1.02, NH_2CH_3* = 1.01.

*Corrected for the unhydrolyzed Asp-NHCH$_3$ which occurs as another peak after the Lys peak. This was shown to be the case by preparing standards of hydrolyzed (under exactly the same conditions used for compound XXIII) BOC-(ß-OBz)-Asp-NHCH$_3$ mixed with Gly and Lys with the same molar ratios and run on the amino acid analyzer.

Acknowledgment. We wish to thank G. T. Taylor for obtaining the data of Table I, E. R. Stimson for a gift of Ac-Gly-NMe and Ac-L-Asp-NMe, C. B. Anfinsen, S. J. Leach, J. H. Griffen and R. Bradley for the use of (and help with) the 220 MHz instrument at NIH, and similarly J. Dadok and R. F. Sprecher for the 250 MHz instrument at Carnegie-Mellon University, and H. S. Hair for the 90 MHz instrument, K. Nishikawa, E. R. Stimson and S. S. Zimmerman for helpful discussions, R. W. Woody for discussions of his calculations on the CD of bend conformations, and the Memphis State University Office of Computer Services for computer time.

Miniprint Material Available. Full-sized photocopies of the miniprint material from this paper only or microfiche (105 × 148 mm, 24× reduction, negatives) containing all of the miniprint and supplementary material for the papers in this issue may be obtained from the Business Office, Books and Journals Division, American Chemical Society, 1155 16th St., N.W., Washington, D.C. 20036. Remit check or money order for $4.50 for photocopy or $2.50 for microfiche, referring to code number MACRO-75-607.

References and Notes

(1) This work was supported by a research grant (AM-13743) from the National Institute of Arthritis and Metabolic Diseases of the National Institutes of Health, U.S. Public Health Service, and by a research grant (BMS71-00872 A04) from the National Science Foundation.
(2) (a) NSF and NIH predoctoral trainee; (b) NIH postdoctoral trainee; (c) Cornell University; (d) Memphis State University.
(3) P. N. Lewis, F. A. Momany, and H. A. Scheraga, *Proc. Natl. Acad. Sci. U.S.*, **68**, 2293 (1971).
(4) P. N. Lewis, F. A. Momany, and H. A. Scheraga, *Biochim. Biophys. Acta*, **303**, 211 (1973).
(5) K. Nishikawa, F. A. Momany, and H. A. Scheraga, *Macromolecules*, **7**, 797 (1974).
(6) S. S. Zimmerman and H. A. Scheraga, work in progress.
(7) H. A. Scheraga, P. N. Lewis, F. A. Momany, P. H. Von Dreele, A. W. Burgess, and J. C. Howard, *Fed. Proc., Fed. Am. Soc. Exp. Biol.*, **32**, 495 (1973).
(8) J. J. Birktoft, B. W. Matthews, and D. M. Blow, *Biochem. Biophys. Res. Commun.*, **36**, 131 (1969).
(9) P. N. Lewis, F. A. Momany and H. A. Scheraga, *Isr. J. Chem.*, **11**, 121 (1973).
(10) F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, *J. Phys. Chem.*, in press.
(11) Abbreviations used here are: TFA, trifluoroacetic acid; THF, tetrahydrofuran; DMF, dimethylformamide; KOH, potassium hydroxide; $BF_3O(Et)_2$, boron trifluoride diethyl etherate; BOC, *tert*-butyloxycarbonyl; Z, benzyloxycarbonyl; OBz, benzyl ester; Ac, acetyl; NMe, methyl amide end group; NMR, nuclear magnetic resonance; CD, circular dichroism.
(12) G. Zweifel and H. C. Brown, *Org. React.*, **13**, 28 (1963).
(13) A. Ali, F. Fahrenholz, and B. Weinstein, *Angew. Chem.*, **84**, 259 (1972).
(14) E. Schnabel, *Justus Liebigs Ann. Chem.*, **702**, 188 (1967).
(15) K. Hofmann, R. Schmiechen, R. D. Wells, Y. Wolman, and N. Yanaihara, *J. Am. Chem. Soc.*, **87**, 611 (1965).
(16) J. M. Stewart and J. D. Young, "Solid Phase Peptide Synthesis", W. H. Freeman, San Francisco, Calif., 1969, p 29.
(17) J. P. Greenstein and M. Winitz, "Chemistry of the Amino Acids", Vol. 2, Wiley, New York, N.Y., 1961, p 891.
(18) R. B. Merrifield, *J. Biol. Chem.*, **232**, 43 (1958).
(19) B. J. Johnson and E. G. Trask, *J. Org. Chem.*, **33**, 4521 (1968).
(20) B. J. Johnson, *J. Pharm. Sci.*, **62**, 1019 (1973).
(21) S. Visser, J. Roeloffs, K. E. T. Kerling, and E. Havinga, *Recl. Trav. Chim. Pays-Bas*, **87**, 559 (1968).
(22) G. D. Fasman, M. Idelson, and E. R. Blout, *J. Am. Chem. Soc.*, **83**, 709 (1961).
(23) J. M. Manning and S. Moore, *J. Biol. Chem.*, **243**, 5591 (1968).
(24) Made available to us through the courtesy of Dr. C. B. Anfinsen, National Institutes of Health, Bethesda, Md.
(25) Made available to us through the courtesy of Dr. J. Dadok, NMR Facility for Biomedical Research, Mellon Institute, Pittsburgh, Pa., sponsored by NIH Grant No. RR00292.
(26) T. C. Farrar and E. D. Becker, "Pulse and Fourier Transform NMR", Academic Press, New York, N.Y., 1971, p 20.
(27) IUPAC-IUB Commission on Biochemical Nomenclature, *Biochemistry*, **9**, 3471 (1970).
(28) J. C. Howard, F. A. Momany, R. H. Andreatta, and H. A. Scheraga, *Macromolecules*, **6**, 535 (1973).
(29) The SADTLER Standard Spectra, published by Sadtler Research Laboratories, Philadelphia, Pa. 19104.
(30) W. B. Smith and A. M. Ihrig, "The NMR Spectra of Eighteen Essential Amino Acids," TCU Research Foundation, Ft. Worth, Texas, 1969.
(31) V. F. Bystrov, V. T. Ivanov, S. L. Portnova, T. A. Balashova, and Yu. A. Ovchinnikov, *Tetrahedron*, **29**, 873 (1973).
(32) J. W. Emsley, J. Feeney, and L. H. Sutcliffe, "High-Resolution Nuclear Magnetic Resonance Spectroscopy", Pergamon Press, New York, N.Y., 1967, p 534.
(33) J. H. Noggle and R. E. Schirmer, "The Nuclear Overhauser Effect, Chemical Applications", Academic Press, New York, N.Y., 1971, Chapters 4 and 8; see especially pp 168, 171, and 199.
(34) R. Rowan, J. A. McCammon, and B. D. Sykes, *J. Am. Chem. Soc.*, **96**, 4773 (1974).
(35) G. M. Crippen and J. T. Yang, *J. Phys. Chem.*, **78**, 1127 (1974).
(36) G. M. Crippen and H. A. Scheraga, *Arch. Biochem. Biophys.*, **144**, 453 (1971); *J. Comput. Phys.*, **12**, 491 (1973).
(37) One cycle of minimization includes a complete search around each variable by a line-search technique[38] but does not necessarily lead to a minimum. This cycle usually involves the calculation of 40 or more (even up to 200) energy values, depending on the steric hindrance encountered.
(38) M. J. D. Powell, *Comput. J.*, **7**, 155 (1964).
(39) A. W. Burgess, P. K. Ponnuswamy, and H. A. Scheraga, *Isr. J. Chem.*, **12**, 239 (1974).
(40) P. K. Ponnuswamy, P. K. Warme, and H. A. Scheraga, *Proc. Natl. Acad. Sci. U.S.*, **70**, 830 (1973).
(41) K. D. Kopple, M. Ohnishi, and A. Go, *J. Am. Chem. Soc.*, **91**, 4264 (1969).
(42) S. L. Portnova, V. V. Shilin, T. A. Balashova, J. Biernat, V. F. Bystrov, V. T. Ivanov, and Yu. A. Ovchinnikov, *Tetrahedron Lett.*, 3085 (1971).
(43) D. A. Torchia, A. di Corato, S. C. K. Wong, C. M. Deber, and E. R. Blout, *J. Am. Chem. Soc.*, **94**, 609 (1972).
(44) D. A. Torchia, S. C. K. Wong, C. M. Deber, and E. R. Blout, *J. Am. Chem. Soc.*, **94**, 616 (1972).
(45) D. N. Silverman, D. Kotelchuck, G. T. Taylor, and H. A. Scheraga, *Arch. Biochem. Biophys.*, **150**, 757 (1972).
(46) R. W. Woody, "Peptides, Polypeptides and Proteins", E. R. Blout, F. A. Bovey, M. Goodman, and N. Lotan, Ed., Wiley, New York, N.Y., 1974, p 338.
(47) W. L. Mattice and W. H. Harrison, *Biopolymers*, in press.
(48) D. C. Poland and H. A. Scheraga, *Biopolymers*, **3**, 321 (1965).
(49) It should be noted that Lewis et al.[4] made only a limited examination of the conformational space of the tetraalanine oligopeptide, rather than the extensive search that was made here for peptides L, N, H_1, and H_2.
(50) G. Nemethy and H. A. Scheraga, *J. Phys. Chem.*, **66**, 1773 (1962).
(51) J. E. Shields, S. T. McDowell, J. Pavlos, and G. R. Gray, *J. Am. Chem. Soc.*, **90**, 3549 (1968).