Macromolecules 1980, 13, 1428–1440

(22) Chavez, L. G.; Scheraga, H. A. *Biochemistry* **1980**, *19*, 1005.

(23) Momany, F. A.; McGuire, R. F.; Burgess, A. W.; Scheraga, H. A. *J. Phys. Chem.* **1975**, *79*, 2361.

(24) The *side-chain* dihedral angles were the *final* values computed by Swenson et al.[25] from the experimental coordinates of Deisenhofer and Steigemann.[26] These coordinates[26] are referred to as X-ray 2 by Swenson et al.[25]

(25) Swenson, M. K.; Burgess, A. W.; Scheraga, H. A. "Frontiers in Physicochemical Biology"; Pullman, B., Ed.; Academic Press: New York, 1978; p 115.

(26) Deisenhofer, J.; Steigemann, W. *Acta Crystallogr., Sect. B* **1975**, *31*, 238.

(27) The values of the *backbone* dihedral angles were again the *final* X-ray 2 computed ones of Swenson et al.[25]

(28) Howard, J. C.; Momany, F. A.; Andreatta, R. H.; Scheraga, H. A. *Macromolecules* **1973**, *6*, 535.

(29) Hodes, Z. I.; Némethy, G.; Scheraga, H. A. *Biopolymers* **1979**, *18*, 1565, 1611.

(30) Némethy, G.; Hodes, Z. I.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 5760.

(31) Paterson, Y.; Leach, S. J. *Macromolecules* **1978**, *11*, 409 (see supplementary material).

(32) Zimmerman, S. S.; Scheraga, H. A. *Biopolymers* **1977**, *16*, 811.

(33) Crow, E. L.; Davis, F. A.; Maxfield, M. W. "Statistics Manual"; Dover: New York, 1960; p 100.

(34) Such an omission can also be justified on energetic grounds. For residues *not* in the helical conformation in the native molecule (see footnote *d* of Table II), the calculated (Ising model) free energy of the helical conformation was always several kcal/mol (and sometimes more) above that at the (nonhelical) global minimum. On the other hand, for residues that *are* in the helical conformation in the native molecule, the calculated (Ising model) free energy (at the minimum corresponding to the helical conformation) was always only 1–1.5 kcal/mol above that calculated for the global minimum; hence, the helical conformation was (incorrectly) not assigned. This small difference in free energy (and the accompanying entropy loss), however, would be more than offset if the stabilizing hydrogen bonds (and the dipole–dipole interactions between the first and fourth residues) were included.

(35) The definition of the regions used in ref 32 is slightly different from the definition used in ref 18, but the method described for the "perfect" prediction leads to the same dihedral angles.

(36) "Ala-type" pertains to all types of residues except Gly and Pro.

(37) Flory, P. J. "Principles of Polymer Chemistry"; Cornell University Press: Ithaca, N.Y., 1953; Chapter 14.

(38) Actually, we find that the native structure lies within 1 kcal/mol of the short-range model.

(39) Ueda, Y.; Scheraga, H. A., work in progress.

(40) Reiss, H. *J. Chem. Phys.* **1967**, *47*, 186.

(41) Allegra, G.; Calligaris, M.; Randaccio, L. *Macromolecules* **1973**, *6*, 390.

Helix–Coil Stability Constants for the Naturally Occurring Amino Acids in Water. 18. Tryptophan Parameters from Random Poly[(hydroxypropyl)glutamine-*co*-L-tryptophan][1]

J. A. Nagy,[2a] S. P. Powers, B. O. Zweifel,[2b] and H. A. Scheraga*[2c]

Baker Laboratory of Chemistry, Cornell University, Ithaca, New York 14853.

Received May 5, 1980

ABSTRACT: Water-soluble, random copolymers containing L-tryptophan and N^5-(3-hydroxypropyl)-L-glutamine have been synthesized, fractionated, and characterized, and the thermally induced helix–coil transitions of these copolymers in water have been investigated. The incorporation of L-tryptophan was found to increase the helix content of the polymers at all temperatures in the range 0–60 °C. The Zimm–Bragg parameters σ and s for the helix–coil transition in poly(L-tryptophan) in water were deduced from an analysis of the melting curves of the copolymers in the manner described in earlier papers. The computed values of s indicate that L-tryptophan enhances helix growth at low temperatures and reduces it at high temperatures; the large value of σ indicates that, in water, this residue has a tendency to promote helix–coil boundaries at all temperatures.

I. Introduction

This paper is concerned with the determination of the helix–coil stability constants of L-tryptophan in water and is a continuation of the series of papers[3-19] in which the conformational preferences of the naturally occurring amino acids in water have been investigated by use of the "host–guest" technique. In this technique, a water-soluble, α-helical host homopolymer with nonionizable side chains is selected, and various amounts of a guest residue are incorporated into it to form random copolymers. By studying the thermally induced helix–coil transitions in these copolymers, it is possible to calculate the Zimm–Bragg[20] helix–coil parameters σ and s for the *guest* residues from an examination of their influence on the helix–coil transition properties of the *host* homopolymer. L-Tryptophan residues are incorporated into copolymers with an N^5-(3-hydroxypropyl)-L-glutamine host, and the thermally induced helix–coil transitions in these copolymers in water are studied.

Although no experimental determination of the helix–coil stability constants for L-tryptophan in aqueous solution has been reported, several studies on the relative preference of L-tryptophan residues for helix and coil conformations have been carried out. Poly(L-tryptophan) has been investigated in organic solvents,[21-28] and both block and random copolymers of tryptophan and amino acids with ionized as well as nonionized side chains have been studied in water[29-32] and in organic solvents.[22,23,26-28,32] The results of the present study indicate that, in water, L-tryptophan has a pronounced ability to promote helix–coil boundaries at all temperatures and that it can either enhance or reduce helix growth, depending on the temperature.

The synthesis of water-soluble random copolymers of L-tryptophan with N^5-(3-hydroxypropyl)-L-glutamine is described in section II, and the experimental characterization of these copolymers and their melting behavior in aqueous solution are presented in section III. Finally, in section IV, the data are analyzed by means of an appropriate form of the theory[3] to determine the helix–coil stability parameters of L-tryptophan in water. The theory is based on evidence[33,34] that short-range interactions dominate in determining the local conformation of a polypeptide or protein. The parameters for L-tryptophan are compared with empirical observations on the behavior of this residue in proteins and with a theoretical analysis of these quantities.

II. Experimental Section

A. Preparation and Characterization of the Copolymers. The synthesis of the copolymers was achieved by first co-

0024-9297/80/2213-1428$01.00/0

Table I
Composition and Chain Lengths of Unfractionated Poly[Glu(OBzl), Trp]

polym no.	L-Trp content of reacn mixt, mol %	A/I[a]	reacn time, h	$\overline{DP}$[b]
A	10	40	2.75	588
B	15[c]	80	19.0	382
C	15	26	4.5	750

[a] Ratio of anhydride to initiator. [b] By viscometry in DMF. [c] Under the reaction conditions used here, only ~6-7% Trp was incorporated in the *final* polymer (see Table II).

Table II
Characterization of Fractionated Copolymers

polym no.	fraction	L-Trp content,[a] mol %	$10^{-3}\overline{M}_w$[b]	$\overline{M}_z/\overline{M}_w$[c]	$\overline{DP}_w$[b]
I	1	10.6	65.4	1.39	351
	2	10.5	46.7	1.12	251
	3	10.0	36.1	1.09	194
	4	10.7			
	5	10.8			
II	1	6.5			
	2	6.6			
	3	6.5	46.6	1.09	250
	4	6.5	30.0	1.01	161
	5	6.7			
III	1	14.5			
	2	13.0			
	3	13.1			
	4	13.5	26.6	1.28	143
	5	14.1	10.8	1.10	58

[a] Determined by hydrolysis in 3 N mercaptoethanesulfonic acid. [b] By conventional sedimentation equilibrium, with an extrapolation to zero concentration. The mean residue weight for all fractions was 186.2. [c] $\overline{M}_z/\overline{M}_w$ is reported for the run at the lowest concentration.

polymerizing the N-carboxyanhydrides (NCA's) of L-tryptophan and γ-benzyl L-glutamate in dioxane, using sodium methoxide as an initiator. The γ-benzyl ester groups were then substituted by 3-hydroxypropylamide groups to yield the desired poly[N^5-(3-hydroxypropyl)-L-glutamine, L-tryptophan].

B. Materials. L-Tryptophan was purchased from Aldrich and recrystallized at least twice from 50% aqueous ethanol immediately before use. 3-Aminopropanol (Aldrich) was dried over Linde molecular sieves (4 Å), distilled under reduced pressure in a nitrogen atmosphere, and stored over molecular sieves. Guanidine hydrochloride (Ultrapure) was obtained from Schwarz/Mann. Mercaptoethanesulfonic acid (3 N; sequanal grade) was purchased from Pierce. All other reagents and solvents were identical in quality and preparation with those used in paper 9 of this series.[11]

Poly[N^5-(3-hydroxypropyl)-L-glutamine], P(HPG), samples of weight-average degree of polymerization $\overline{DP}_w$ = 120, 220, and 360 were fractions IIC, IIB, and IIID, respectively, of paper 2 of this series.[4]

C. Synthesis. *N*-Carboxyanhydrides. γ-Benzyl L-glutamate N-carboxyanhydride was prepared by treatment of γ-benzyl L-glutamate (synthesized according to the procedure of Prestige et al.[35]) with phosgene, using tetrahydrofuran[36] or dioxane as a solvent, as described by Hirschmann et al.[37]

L-Tryptophan N-carboxyanhydride was synthesized by using L-tryptophan which had been recrystallized from 50% aqueous ethanol immediately before use. Since L-tryptophan is light sensitive, all reactions were protected from light as far as possible. The L-tryptophan was treated with a 10-20% excess of phosgene in dioxane. The mixture was stirred at 40 °C until all of the tryptophan had dissolved, and the greenish solution was concentrated and then diluted with ethyl acetate. Activated charcoal was added, and the solution was filtered over Celite and then treated with one volume of hexanes. A brownish green oil separated on cooling at −70 °C. The oil was removed and the supernatant was treated with more hexanes. This procedure was repeated until all of the oil had been removed. The colorless supernatant was decanted and treated with more hexanes, and the white crystals were isolated. L-Tryptophan N-carboxyanhydride was prepared in yields of about 8-10%; mp 144-145 °C [lit. mp: 135 °C,[38] 145-146 °C[22]]. Anal. Calcd for $C_{12}H_{10}N_2O_3$: C, 62.61; H, 4.38; N, 12.17. Found: C, 62.45; H, 4.44; N, 12.12. Only trace or no Cl was found.

Poly[γ-benzyl L-glutamate-*co*-L-tryptophan], Poly[Glu(OBzl), Trp], Copolymers A-C. Random copolymers of γ-benzyl L-glutamate with up to 14% L-tryptophan were synthesized by polymerization of the N-carboxyanhydrides in dioxane with sodium methoxide as initiator.[11] The progress of the polymerization reaction was monitored by assaying for unreacted N-carboxyanhydrides as described in paper 10 of this series.[12] Yields from the polymerization were generally in the range of 80-85%. The polymers were isolated by pouring the reaction mixture into absolute ethanol as usual.[11] The chain lengths of these polymers were determined roughly with the viscosity-molecular weight relationship of Fujita et al.,[39] using dimethylformamide as a solvent (see section IID4), and are given in Table I.

Poly[N^5-(3-hydroxypropyl)-L-glutamine-*co*-L-tryptophan], Poly(HPG, Trp), Copolymers I-III. The copolymers A-C were dissolved in dioxane and treated with 3-amino-1-propanol at 50 °C under nitrogen, in the dark, to yield water-soluble copolymers, poly(HPG, Trp) polymers I-III, respectively, as described previously.[6] The course of the aminolysis reaction was monitored as described in paper 10[12] and terminated when no fewer than 99.5% of the benzyl groups had been exchanged. The reaction mixture was then poured into an excess of 1 N acetic acid and then dialyzed against water at 20 °C, in the dark, until amines could no longer be detected by a ninhydrin test on 0.1 mL of the dialysate.[40] The water-soluble poly(HPG, Trp) was recovered by lyophilization in 50-80% yield, based on the number of moles of poly[Glu(OBzl), Trp]. The crude polymer was fractionated by the procedure described in paper 2 of this series.[4] After fractionation, the polymers were dissolved in water, lyophilized, and dried in vacuo over P_2O_5. The lyophilized fractions were stored over P_2O_5/KOH in the dark. The composition and chain lengths of the fractions analyzed for their thermally induced helix-coil transitions are summarized in Table II.

D. Analytical Methods. 1. Determination of Composition. Despite the availability of convenient and accurate methods for the analysis of most amino acids, the determination of tryptophan has remained difficult. Largely due to its acid lability, methods for its analysis suggest hydrolysis under carefully controlled acid[41,42] or alkaline conditions in the presence of reducing agents[43] or by enzymatic means.[44,45] Alternatively, spectrophotometric procedures employ reagents which produce colored products on reaction with tryptophan.[46-57] The near-ultraviolet absorption spectra of tryptophan in alkali,[58] in guanidine hydrochloride,[59] or in H_2O[60] have also been used for quantitative analysis of tryptophan. Fluorescence[61-64] and magnetic circular dichroism[65-68] measurements have recently been applied as methods to determine the tryptophan content of peptides and proteins.

Three methods of tryptophan analysis were employed to determine the composition of the copolymers: (i) amino acid analysis following hydrolysis with 3 N mercaptoethanesulfonic acid,[69] (ii) spectrophotometric measurement of the tryptophan content in polymer solutions in 6 M guanidine hydrochloride,[59] and (iii) evaluation of tryptophan content, using magnetic circular dichroism measurements.[67]

a. Amino Acid Analysis. The copolymer fractions were hydrolyzed in evacuated, degassed, sealed ampules, according to the procedure of Penke et al.,[69] using 3 N mercaptoethanesulfonic acid for 22 h at 110 °C. A Technicon TSM amino acid analyzer was used for the analyses. For those fractions which were used to obtain melting curve data, two separate samples were hydrolyzed. The average experimental error in the determination of the composition for these water-soluble fractions was ±5%. One to three milligrams of tryptophan was reacted with 1 mL of 3 N mercaptoethanesulfonic acid in the described way for 22 h to check for loss of tryptophan during hydrolysis. Free tryptophan gave recoveries of about 95% after 22 h. Standard mixtures of glutamic acid and tryptophan were also subjected to the same hydrolytic conditions. The same Glu:Trp ratios were achieved when the samples were analyzed before and after hydrolysis, showing no loss of glutamic acid and no loss of tryptophan due to the presence of glutamic acid. An advantage of the mercaptoethanesulfonic acid method of hydrolysis was that it was then not necessary to evaporate the hydrolysates which might cause side reactions such as the loss of glutamic acid through ester formation with 3-amino-1-propanol.[15]

b. Spectrophotometric Determination of Composition. The method of Edelhoch[59] was used to determine the tryptophan content of the various fractions of copolymers I–III. In a typical determination, a 1-mL aliquot of a copolymer solution was dissolved in 5 mL of a 7.2 M guanidine hydrochloride solution to make a 6 M guanidine hydrochloride solution (pH 6.5, 0.02 M phosphate). Absorbance measurements at 280 nm were made on a Cary 14 spectrophotometer; ϵ_{280} = 5690[59] determined for *N*-acetyl-L-tryptophanamide in 6 M guanidine hydrochloride (pH 6.5, 0.02 M phosphate) was used for calculating the concentration of tryptophan. Another aliquot of the original copolymer solution was used to determine total concentration of copolymer by nitrogen analysis (see below). A combination of these results yielded the composition of tryptophan in each of the copolymer fractions studied.

c. Magnetic Circular Dichroism (MCD) Measurements. Studies of tryptophan-containing peptides and proteins and of indole derivatives have demonstrated that the discrete and prominent magnetically induced dichroic band between 290 and 300 nm, characteristic of tryptophan, can serve as a convenient and nondestructive means for quantitative analysis of the residue.[65–68] The magnitude of the positive MCD band is a linear function of tryptophan content. Using the molar magnetic ellipticity per tryptophanyl residue, 777 deg cm^2 dmol^{-1} kG^{-1},[67] the quantitative measurement of the tryptophan was based on the measurement of the positive band and a comparison of its intensity with that of a known concentration of *N*-acetyltryptophanamide (solutions made up gravimetrically from water-recrystallized and vacuum-dried crystals). Assessment of the total polymer concentration, by nitrogen analysis (see below), in conjunction with the molar concentration of tryptophan from MCD measurements gave the tryptophan content. A value of ϵ_{280} for tryptophan in these copolymers in H_2O was determined from a combination of measurements of MCD and A_{280} made on aqueous solutions of fraction 2 of copolymer I (see Results).

Magnetic circular dichroism measurements were made with a Cary 61 circular dichroic spectrophotometer equipped with a Varian Model V4145 superconducting magnet, energized by a Varian V4106 power supply. A magnetic field strength of 45.5 kG was used. Scans were made at a speed of 0.2 nm/s with a spectral bandwidth of 2 nm. CD measurements were made with a Cary 61 spectropolarimeter equipped with a Cary 6001 CD attachment. A cylindrical neckless cuvette constructed of Suprasil of 1-cm path length was used. Absorbances of all samples were maintained below 1.0 absorbance. Absorbance spectra were obtained with a Cary 14 spectrophotometer equipped with a Cary tungsten–iodine high-intensity light source. All measurements were made at ambient temperatures of 21–23 °C.

In a typical experiment, a solution of the sample to be analyzed was placed in a cuvette and its absorption spectrum was measured with a Cary 14 spectrophotometer using air as the reference. The cuvette was transferred immediately to the Cary 61 to record the corresponding MCD spectrum. The CD spectrum of the sample was also measured. Base-line spectra were measured, using deionized, distilled water by the same three techniques.

CD data are reported as molar ellipticities, $[\Theta] = 100(\Theta)/bM$ in units of deg cm^2/dmol, where Θ is the ellipticity in degrees as obtained directly from instrumental readings, b is the path length, and M is the molar concentration of the sample. Values of $[\Theta]_M$, the molar magnetically induced ellipticity, were determined from the difference between the ellipticity in the presence of the magnetic field and the natural CD and are normalized to a field of 1 kG with units of deg cm^2 dmol^{-1} kG^{-1}.

The ratio of ellipticity of the positive MCD peak near 290 nm of *N*-acetyl-L-tryptophanamide to its absorption at 280 nm measured on the Cary 14 (0.0621° at 43.3 kG) served as a control standard to calibrate the magnitude of the magnetic field. Concentrations of the standard *N*-acetyl-L-tryptophanamide solutions were based on absorption: ϵ_{280} = 5520.[67] An estimated overall error based on the summation of operational factors involved in MCD (i.e., signal/noise, magnetic field setting) is ~±5%. The UV and MCD spectra for various concentrations of copolymer I, fraction 2 are shown in Figure 1.

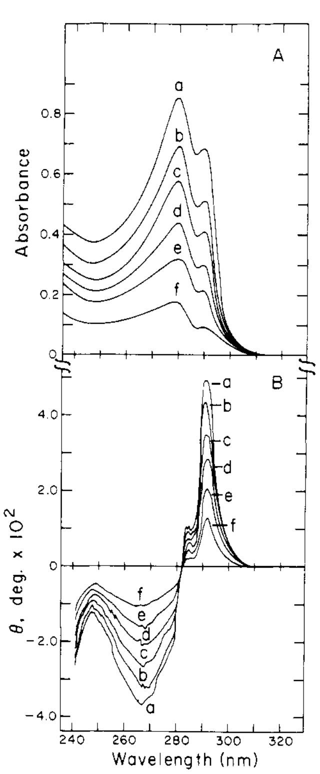

Figure 1. Observed absorbance (A) and ellipticity in the presence of a 45.5-kG magnetic field (B) for aqueous solutions of fraction I-2 of various copolymer concentrations (w/v) %: (a) 0.0259%; (b) 0.0225%; (c) 0.0182%; (d) 0.0148%; (e) 0.0109%; (f) 0.0068%. In the absence of the magnetic field, there was no observable ellipticity in this wavelength region.

2. Determination of Concentration. The concentrations of the copolymers in the solutions used in the various physical measurements were determined by micro-Kjeldahl nitrogen analysis, using Lang's method[70] for digestion and the semiautomated method of detection of ammonia of Noel and Hambleton.[71] In order to ensure complete digestion of the indole moiety, a time-dependent study of concentration vs. digestion time was performed on a standard L-tryptophan solution, whose concentration was determined by absorption: L-tryptophan, ϵ_{280} = 5600.[60] For digestion times ≥3 h, the agreement in concentration by absorption and by nitrogen analysis was ±3%. Therefore, all copolymer samples were digested for 3 h. The error in these concentration determinations, made by nitrogen analyses, was estimated to be ±3%.

3. Assay for D-Amino Acids. The starting amino acids as well as the final copolymers, poly(HPG, Trp), were checked for the presence of D residues by the L-leucyl dipeptide or L-glutamyl dipeptide methods of Manning and Moore[72] for L-glutamic acid or L-tryptophan, respectively. The starting amino acids were found to contain less than 0.1% of their D isomers.

Tryptophan and glutamic acid were isolated separately from acid hydrolysates (3 N mercaptoethanesulfonic acid) of the copolymers prior to derivatization. Complete separation of these two amino acids from the hydrolysate of 5–10 mg of polymer was achieved by column chromatography on a 0.8 × 20 cm column of Dowex AG1-X8 (200–400 mesh) in the acetate form, using a modification of the method of Hirs et al.[73] The column was eluted first with distilled water to separate 3-amino-1-propanol and tryptophan, sequentially, and then with 0.5 N acetic acid to remove the glutamic acid (flow rate 15 mL/h). The mercaptoethanesulfonic acid was retained on the column resin. For best results each column was used only once. The L-Glu-L-Trp and L-Glu-D-Trp dipeptides were separated on a Technicon Auto Analyzer

Table III
Determination of L-Tryptophan Content

copolym fraction	mol % of tryptophan		
	hydrolysis with MES[a]	A_{280} in 6 M Gdn·HCl[b]	A_{280} in H_2O using ϵ_{MCD}[c]
I-1	10.6	10.3	10.5
I-2	10.5	10.2	10.0
I-3	10.0	10.1	10.6
II-3	6.5	6.6	6.8
II-4	6.5	6.5	7.0
III-4	13.5	13.0	13.6
III-5	14.1	13.7	13.1

[a] See ref 69; MES = mercaptoethanesulfonic acid. [b] See ref 59; Gdn·HCl = guanidine hydrochloride. [c] See ref 67.

with 0.2 N sodium, 0.1 M citrate elution buffer at pH 4.68 on a 120-cm column of Chromobeads B at 62 °C. The L-Leu-L-Glu and L-Leu-D-Glu dipeptides were separated on the same system with 0.4 N sodium, 0.05 M citrate buffer at pH 2.9. By this technique, the tryptophan and glutamic acid residues in the polymers were found to contain less than 3% of their D isomers, after correction for racemization during hydrolysis, as determined by carrying a mixture of L-tryptophan, L-glutamic acid, and 3-amino-1-propanol through the same procedure.

4. Viscometry. The viscosities of the tryptophan copolymers could not be determined in dichloroacetic acid (DCA), the usual solvent,[4] because of the immediate destruction of the indole moiety by this solvent, as indicated by gross discoloration of the solutions. Therefore the intrinsic viscosity of each of the copolymers A–C was determined in dimethylformamide at 25.0 ± 0.1 °C in a Cannon-Ubbelohde semimicro dilution viscometer. A rough estimate of the molecular weight was determined for these polymers by using poly(γ-benzyl L-glutamate) as a model and the relationship determined by Fujita et al.[39] for this polymer in dimethylformamide.

5. Optical Rotatory Dispersion and Circular Dichroism Measurements. The optical rotatory dispersion (ORD) and the circular dichroism (CD) measurements were made on a Cary Model 60 spectropolarimeter equipped with a Model 6001 CD attachment. Temperature control was maintained to within ±0.2 °C with water-jacketed quartz cells. The ORD studies used in estimating the helix content of the copolymers were performed at two or more concentrations for each fraction in the range 0.05–0.35% (w/v) over the wavelength range 450–320 nm. The solutions were filtered through 0.45-μm Millipore filters before use. The optical rotation of solutions of tryptophan copolymers could be determined only down to 320 nm because of interference by the absorbance of the indole moiety at lower wavelengths.

6. Molecular Weight Determination. The weight- and z-average molecular weights of the poly(HPG, Trp) fractions were determined on aqueous solutions by the conventional sedimentation equilibrium method, using a Spinco Model E ultracentrifuge. The sedimentation equilibrium data were analyzed by the procedure described by Chervenka.[74] Determination of the initial concentration was made by calibrating the fringe shift against polymer concentration as in paper 7.[9] The concentration dependence of the weight-average molecular weight, $\bar{M}_w$, was determined for each sample, and $\bar{M}_z$ was computed from the run at the lowest concentration for each fraction. The estimated precision in the values of $\overline{DP}_w$ was ±5%.

The partial specific volumes ($\bar{v}$) of the several fractions, required for the calculation of the molecular weights, were determined from the amino acid content as described by Cohn and Edsall.[75] A value of $\bar{v}$ = 0.79 for P(HPG) was used in the calculation of $\bar{v}$ for the copolymers.[4]

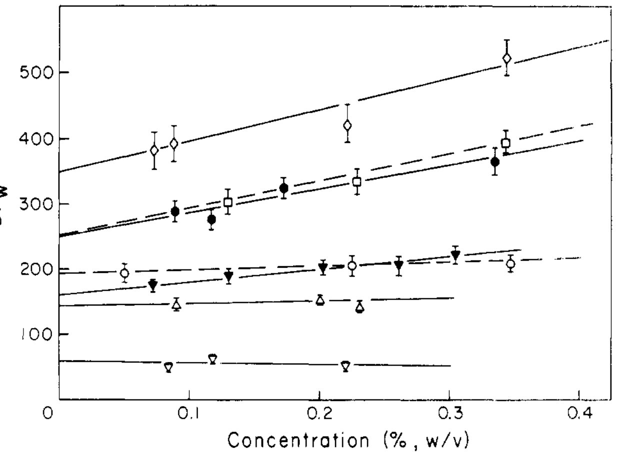

Figure 2. Concentration dependence of molecular weights for fractions used in analysis to obtain σ and s: (◇) 10.6% Trp, $\overline{DP}_w$ = 351 (fraction I-1); (□) 10.5% Trp, $\overline{DP}_w$ = 251 (fraction I-2); (○) 10.5% Trp, $\overline{DP}_w$ = 194 (fraction I-3); (●) 6.5% Trp, $\overline{DP}_w$ = 250 (fraction II-3); (▼) 6.5% Trp, $\overline{DP}_w$ = 161 (fraction II-4); (△) 13.5% Trp, $\overline{DP}_w$ = 143 (fraction III-4); (▽) 14.1% Trp, $\overline{DP}_w$ = 58 (fraction III-5). The error symbols represent the experimental error in each measurement.

III. Results

A. Characterization of the Copolymers. Table I summarizes the composition of the reaction mixtures and the composition and average degree of polymerization ($\overline{DP}$) of the unfractionated poly[Glu(OBzl), Trp] copolymers. A comparison of the various methods used to determine the tryptophan composition of the copolymers is presented in Table III. The results for a particular fraction are in agreement to within ±5%. Table II summarizes the composition, $\overline{DP}_w$, and $\bar{M}_z/\bar{M}_w$ data for those fractions of the corresponding exchanged copolymers poly(HPG, Trp) which were used to determine the Zimm–Bragg parameters σ and s. The usual decrease in $\overline{DP}$ attributed to transamidation during aminolysis[76] is apparent by a comparison of Table I and Table II.

The data in Table II indicate that the amino acid composition is independent of the chain length for a series of fractions from a given copolymer parent. While this criterion has been used in previous papers in this series to indicate that the polymers were sufficiently random for the purpose required here, a more rigorous proof was provided for copolymers containing methionine[77] as a guest residue, by analyzing the distribution of methionine in the fragments which result from treatment of the copolymers with CNBr. In any case, the assumption in the theory that the copolymers are truly random has been shown to be not too critical. The presence of short blocks of guest residues (number of such residues ~10–15% of $\sigma^{-1/2}$) has been shown to have little effect on the melting behavior of a random copolymer.[3]

The concentration dependencies of the apparent molecular weights of the fractions studied are shown in Figure 2. The data have been extrapolated to infinite dilution to obtain $\overline{DP}_w$. The values of $\bar{M}_z/\bar{M}_w$ for the fractions do not depart significantly from unity, indicating that the fractions used in the determination of σ and s were relatively homogeneous. This is important, since the molecular weights are fairly small and, with shorter chains, the theory used to calculate σ and s makes more stringent demands on the uniformity of chain length of the samples than with chains approaching infinite length, effectively $\overline{DP}_w$ ~ 600 or greater.

As a means of ensuring internal consistency in the various physical measurements (molecular weight, ORD, CD, etc.) which would be subject to error introduced by spurious results from a particular series of micro-Kjeldahl analyses, a calibration curve was prepared from careful

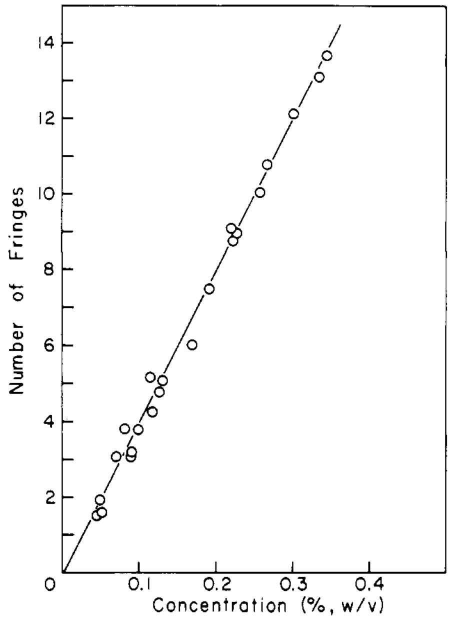

Figure 3. Plot of the number of fringes obtained from a synthetic boundary run in the ultracentrifuge against the concentration of the sample obtained from nitrogen analysis.

measurements of concentration, c_0 in fringe units, from synthetic boundary runs in the ultracentrifuge against polymer concentration assayed by nitrogen analyses. The calibration curve so obtained is shown in Figure 3. All the solutions studied were checked against this curve for consistency in the concentration measurements.

Several copolymer solutions were checked for any possible destruction of tryptophanyl residues due to photooxidation which might have occurred during preparation and handling. N'-Formylkynurenine, the major degradation product,[78-81] can be recognized by its absorption spectrum, λ_{max} at 321 nm, ϵ 3750, and 260 nm, ϵ 10 980[82] at pH 7.0; therefore, if tryptophanyl side chains were destroyed, the absorption at 280 nm would fall and that over 300 nm and near 260 nm would rise. UV absorption spectra of the copolymer solutions were recorded before and after the solutions were used for determinations of b_0. No decrease in the ratio A_{280}/A_{321} or in A_{280}/A_{260} was observed for any of the copolymer fractions, indicating that less than 3×10^{-6} mol of N'-formylkynurenine was formed (using ϵ = 2750 at 321 nm[82]).

The presence in the copolymers of less than 3% of the tryptophan content as D-tryptophan (and similarly for glutamic acid) is not considered to be large enough to affect the computed values of σ and s significantly.

A value of ϵ_{280} = 5700 (±5%) was determined for tryptophan from the results of MCD and A_{280} measurements (Figure 1). The mol % tryptophan for all other copolymer fractions could be determined by using this value, A_{280}, and nitrogen analysis. This value of ϵ may be compared with ϵ_{280} = 5690[59] and 5520[67] for N-acetyl-L-tryptophanamide in 6 M guanidine hydrochloride and H_2O, respectively, and with 5600 for L-tryptophan in H_2O.[60]

B. ORD and CD Data for the Copolymers. The ORD and CD data in the region 190–240 nm for representative fractions of poly(HPG, Trp) in water are shown in Figure 4. Both the ORD and the CD data are interpretable in terms of combinations of right-handed α-helix and random-coil states,[83-85] the relative contribution of each conformation being a function of both temperature and composition.

One assumption that is made when the types and amounts of backbone structure in a protein or a polypeptide are estimated from a given rotatory parameter, i.e., from b_0, $[m']_{233\ \text{nm}}$, or $[\Theta]_{280\ \text{nm}}$, is that the peptide bonds of the backbone chain are the only source of optical activity contributing to the rotatory parameter.[86] The presence of Cotton effects due to the aromatic transitions of that, may interfere with the visible or far-UV ORD and CD measurements. Contributions to the ORD and CD spectra of the copolymers arising from the tryptophan side-chain transitions are not evident, however, in the spectra of Figure 4, when compared with spectra of poly(L-tryptophan) or copolymers of tryptophan in a variety of solvents[21-28] and amount to at most a very small perturbation on the large amide background. This result is consistent with available experimental evidence, e.g., that of Peggion and co-workers,[27] who reported that the CD spectrum of a 16% copolymer of L-tryptophan and γ-ethyl L-glutamate studied in ethylene glycol monomethyl ether is that of a pure right-handed α helix, with negative bands at 208 and 222 nm; the latter values are in quantitative agreement with those previously reported for polypeptides in the α-helix conformation.[87,88] The CD spectrum of an α helix is distorted by incorporation of increasing amounts (>16%) of L-tryptophan, with the 208- and 222-nm bands of the α helix being replaced by a negative band at 220 nm and with a positive band (attributed to exciton splitting[27]) appearing at 230 nm.[27] This change is gradual but not linear with the mole fraction of tryptophan.[27] In the copolymers studied in this paper, however, the tryptophan content is below those reported[27] to show CD bands attributable to side-chain chromophores. Our spectra

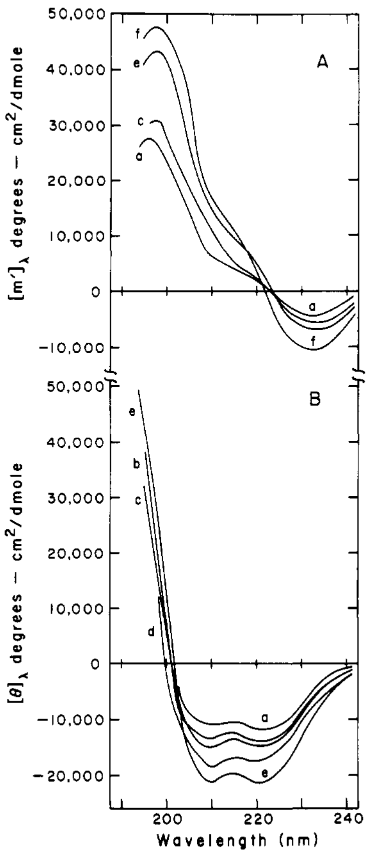

Figure 4. (A) ORD and (B) CD data in water for representative fractions of tryptophan copolymers: (a) 6.5% Trp, $\overline{DP}_w$ = 161 at 20.2 °C; (b) 6.5% Trp, $\overline{DP}_w$ = 250 at 20.2 °C; (c) 13.5% Trp, $\overline{DP}_w$ = 143 at 20.2 °C; (d) 10.0% Trp, $\overline{DP}_w$ = 194 at 20.2 °C; (e) 10.6% Trp, $\overline{DP}_w$ = 351 at 20.2 °C; (f) 10.6% Trp, $\overline{DP}_w$ = 351 at 5.3 °C. The first and last curves are labeled in the 208–240-nm range. Others follow in alphabetical order.

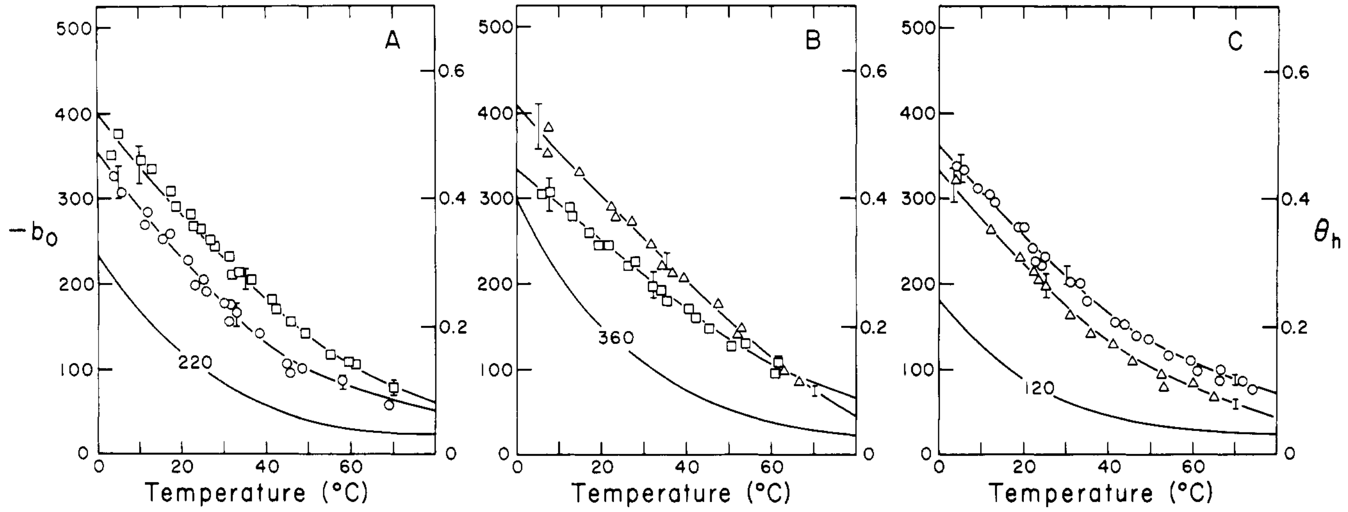

Figure 5. Temperature dependence of $-b_0$ for poly(HPG, Trp) copolymers in water. The melting curves of poly(HPG) (lines without experimental points) of $\overline{DP}_w = 120$ (fraction IIC of paper 2), $\overline{DP}_w = 220$ (fraction IIB of paper 2), and $\overline{DP}_w = 360$ (fraction IIID of paper 2) are included for comparison: (A) (□) 10.5% Trp, $\overline{DP}_w = 251$, (○) 6.5% Trp, $\overline{DP}_w = 250$; (B) (△) 13.5% Trp, $\overline{DP}_w = 143$, (□) 10.6% Trp, $\overline{DP}_w = 351$; (C) (△) 6.5% Trp, $\overline{DP}_w = 161$, (○) 10.0% Trp, $\overline{DP}_w = 194$. The points are the experimental ones and the lines represent the smoothed data used in subsequent calculations. The error symbols reflect the errors in b_0 arising from errors in the determination of solution concentration and in the slope of the Moffitt–Yang plot.

showed no evidence of a positive CD band at 230 nm. Because of the low mole fractions of tryptophan in our copolymers and because of the nature of our host amino acid, any specific orienting of tryptophan side chains (that might have led to exciton splitting) would be expected to be minimal and would result in negligibly small side-chain rotational strengths.

The information in Figure 4 clearly demonstrates that incorporation of L-tryptophan increases the helix content of the copolymer (compare curve D of Figure 3 of paper 2[4] with curve d of Figure 4B). The ORD curves labeled e and f in Figure 4A show that, for a particular polymer fraction, the relative amount of helix decreases as the temperature increases, thus indicating that the copolymers undergo a thermally induced helix-to-coil transition in water.

The ORD spectra of aqueous solutions of the polymers in the range 450–320 nm were used to calculate the value of the Moffitt–Yang parameter b_0 at various temperatures. Data for six representative fractions are shown in Figure 5, together with data obtained earlier on P(HPG).[4] ORD spectra were extended only down to 320 nm in order to minimize contributions to b_0 from the long-wavelength tryptophanyl transition. The Moffitt–Yang plots used to determine b_0 show no departure from linearity in the range 450–320 nm. Although there are no discernible contributions of the tryptophan chromophore in the far-UV ORD and CD spectra of the copolymers, there may, however, be contributions of the aromatic chromophores to the visible ORD since side-chain chromophores may interact with one another and with the peptide chromophore and the various Cotton effects would contribute to the tail of the ORD in the visible region. An upper limit of possible tryptophan side-chain contributions to b_0 and their possible effect on the computed conformational parameters of tryptophan are discussed in section IIIC. The procedure used to obtain the curves of Figure 5 was otherwise the same as shown in paper 7.[9] No concentration dependence was observed, and all curves were reproducible. The error symbols in Figure 5 reflect two standard deviations in b_0 (or in helix content θ_h) calculated[8,9] from standard deviations in the concentration ($\pm 0.025 b_0$) and in the slope of the Moffitt–Yang plot (±3, in b_0 units). The reversibility of the transition was checked for each melting curve. In Figure 6, the reversibility of one of the transitions studied is shown in detail.

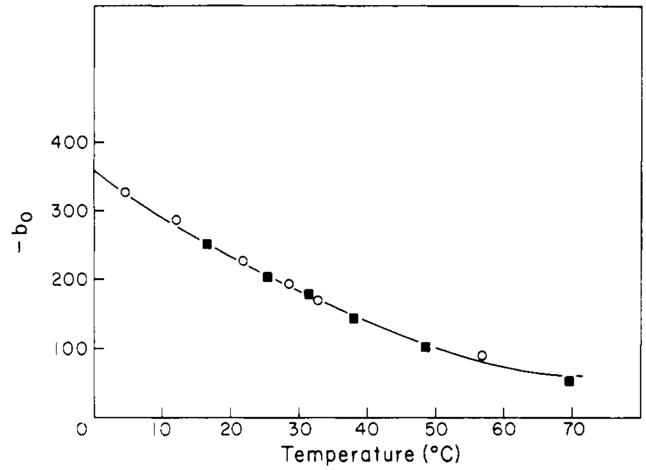

Figure 6. Plot of $-b_0$ vs. temperature to demonstrate the reversibility of the helix–coil transition. The filled symbols were determined while heating the solution and the open symbols were determined while cooling the solution (6.5% Trp copolymer, $\overline{DP}_w$ = 250).

Several observations can be made from these thermally induced melting curves. First, at low temperature, incorporation of L-tryptophan increases the helix content of a copolymer relative to that of the corresponding homopolymer (i.e., the one of similar chain length). Only at high temperatures do the helix contents of copolymer and homopolymer approach each other. Second, for similar chain-length copolymers at a given temperature, the ability of tryptophan to increase helix content is not linear in the mole fraction of tryptophan. This was also observed for copolymers containing L-tyrosine[12] and L-methionine.[16] These aspects of the copolymer transition behavior will be delineated with the aid of the helix–coil transition parameters obtained for tryptophan in section IV.

C. b_0 for Complete Helix and Complete Coil. For the homopolymer P(HPG) studied in paper 2,[4] the magnitude of b_0 for the complete helix has been taken as −750

Table IV
Values of b_0 for Copolymers in Trifluoroethanol and Guanidine Hydrochloride

copolym fraction	L-Trp content, mol %	solvent	T, °C	b_0 ($\overline{\mathrm{DP}}_\mathrm{w}$)	b_0 ($\overline{\mathrm{DP}}_\mathrm{w}$)[a]
I-2	10.5	TFE	2.8	−727 (250)	−720 (300)
II-4	6.5	TFE	2.2	−669 (161)	−690 (210)
III-4	13.5	TFE	2.3	−688 (143)	−660 (150)
III-5	14.1	TFE	2.3	−638 (58)	−650 (100)
III-4	13.5	6 M Gdn·HCl	70.2	+35	

[a] Maximum values of b_0 at ~5 °C for samples of P(HPG) in methanol, with the indicated $\overline{\mathrm{DP}}_\mathrm{w}$'s from Figure 3 of Okita et al.[92]

and that for the complete coil as zero. Because these values may vary with the nature of the side chain[90] and especially because there may possibly be contributions to b_0 from the tryptophanyl side-chain chromophore, in contrast to their apparent absence in the ORD and CD curves in the 190–240-nm range (Figure 4), several fractions were examined, first, in trifluoroethanol (TFE) at ~2.2 °C to determine b_0 for the complete helix and, second, in 6 M guanidine hydrochloride at 70.2 °C to determine b_0 for the complete coil. In each case, b_0 was corrected for the dispersion of the refractive index of the solvent.[91] The results are shown in Table IV, together with values of b_0 for homopolymers of P(HPG) of corresponding chain length.[92]

Values of $-b_0$ less than 750 for the *homopolymers* indicate that the infinite-chain behavior (100% helix content) is attained only when the chain length approaches 1000 residues. Small chain lengths have a correspondingly smaller fraction of helical residues.

The values of $-b_0$ obtained for the *copolymers* in TFE indicate several things. First there is the expected decrease in the values of $-b_0$ with decreasing chain length, for a given composition. Second, the values of $-b_0$ increase with increasing tryptophan content, for a given chain length. These increases in $-b_0$ reflect either (i) a gain of helix content (from the action of tryptophan as a possible helix maker), (ii) a positive contribution to $-b_0$ from the tryptophanyl side-chain chromophore without change in conformation, or (iii) a gain in helix content, compensated by negative contributions to $-b_0$ by tryptophanyl side-chain chromophores. Third, whatever the cause of the high values of $-b_0$, several of these values are the same (within an experimental error of ±15–20 b_0 units) as those of the host homopolymer of the corresponding chain length; e.g., the values of $-b_0$ = 727 and 669 for I-2 and II-4, respectively, are similar to the values 720 and 660 for homopolymers of approximately the same chain length; thus, the effect of tryptophan incorporation on $-b_0$ is indeed small. Even though we cannot specify the reason for this effect, we suspect that it arises from a gain of helix content (see earlier discussion of Figure 4).

Nevertheless, to exclude the possibility that the value of b_0 might be erroneous because of rotational contributions from side-chain tryptophanyl Cotton effects to the tail of the helical ORD curve measured from 320 to 450 nm, the CD spectra of several polymers were examined in the *near*-UV region from 260–320 nm in H_2O and TFE as was done in the case of the tyrosine copolymers.[12] In no case were any dichroic bands distinguishable from the base lines in the 260–320-nm region of the CD spectrum of the tryptophan copolymers. This is in agreement with further results of Peggion et al.,[27] i.e., that observable deviations from base-line CD at 268 nm appear only at tryptophan content >16% (and our copolymers contain <16% Trp). Only low-temperature CD studies on model compounds such as *N*-acetyl-L-tryptophanamide[93–95] and H-Trp-Trp-OH[96–102] have shown ellipticities resulting from the L_a and L_b *near*-UV bands of tryptophan, i.e., at 77 and 203 K, respectively, but not at 298 K.

A possible rationalization is that the bands appearing in the near-UV region at low temperature reflect the conformational rigidity of the aromatic residue (i.e., a restriction of the torsion angles χ_1 and χ_2).[100] As with these model compounds, multiple conformations may exist for the tryptophan side chains of the copolymers at 298 K, which are interconvertible by rotations about single bonds. The circular dichroism spectrum at 298 K appears to represent the weighted average of all these conformations. For some, the circular dichroism bands of a given vibronic transition may be negative; for others, the same circular dichroism band may be positive.[93] Thus, at 298 K the circular dichroism strength for some conformations would cancel that of others. When the temperature is lowered to 77 K, the lower energy conformations would become more heavily populated in accordance with the Boltzmann factor. The average circular dichroism strength would increase because a much larger fraction of the tryptophan molecules would have the same circular dichroism spectra. Presumably, this is why the near-UV bands of tryptophan are seen only at low temperature. The tyrosine copolymer, on the other hand, did exhibit a near-UV CD band at room temperature.[12] Because of the symmetry of the tyrosine side chain (whereby a 180° variation of χ^2 does not change the position of the aromatic ring), in contrast to the tryptophan side chain, a much larger fraction of the tyrosine molecules would have the same CD spectra even at room temperature, resulting in a higher circular dichroism strength as evidenced by the prominent near-UV CD band. Nevertheless, the contribution of the tyrosyl side-chain chromophore, calibrated from this observed ellipticity, was shown[12] to have a negligible effect on the calculated values of σ and s.

While the above observations are reassuring that there are no discernible contributions from side-chain tryptophan chromophores to the value of $-b_0$, there are still both positive and negative Cotton effects which, in principle, could contribute to $-b_0$. Negative Cotton effects have been observed at 225 nm in compounds such as H-Gly-Trp-(Gly)$_n$-Trp-Gly-OH, where n = 0, and in H-Trp-Trp-OH and H-Trp-Trp-Gly-OH, provided that the N-terminal α-amino group is not protonated.[96] This characteristic behavior in the 220–230-nm range has been attributed to a conformation in which the two aromatic side chains are rigidly positioned (but not stacked) relative to each other; stacking interactions do not appear important in stabilizing this conformationally rigid structure.[100] Rather, intramolecular electrostatic interactions (e.g., hydrogen bonding or polar interactions between the aromatic side chain and the peptide backbone) are suggested to be the salient forces.[100]

Positive Cotton effects have been observed in compounds such as H-Gly-Trp-(Gly)$_n$-Trp-Gly-OH, with n = 1, 2,[96] and in synthetic aromatic-containing polypeptides in ordered conformations.[103–105] In the case of poly(L-tryptophan) a value of +570 for the Moffitt–Yang param-

Table V
Comparison of the Values of θ_h, Calculated with the Approximate and Exact Theories[a] for Finite Chains

L-Trp content, mol fraction	$\overline{DP}_w$	T, °C	$(\theta_h)_{exptl}$	$(\theta_h)_{theor}$		
				Lifson[b]	Allegra[c]	Lehman-McTague[c]
0.135	143	0	0.447	0.481	0.483	0.481
		30	0.280	0.296	0.297	0.296
		60	0.144	0.145	0.145	0.145
0.106	351	0	0.567	0.532	0.532	0.532
		30	0.319	0.307	0.307	0.307
		60	0.143	0.144	0.144	0.145
0.065	250	0	0.455	0.476	0.475	0.473
		30	0.233	0.244	0.243	0.242
		60	0.100	0.113	0.114	0.114

[a] The parameters used for P(HPG) were those of Table II in paper 2.[4] [b] The parameters used for L-tryptophan were obtained by fitting the data by the Lifson theory with $\sigma = 6.8 \times 10^{-3}$. [c] The parameters used for L-tryptophan were obtained by fitting the data by the Allegra theory with $\sigma = 7.7 \times 10^{-3}$.

eter b_0 in DMF has been reported,[22] even though poly(L-tryptophan) has been shown to be a right-handed α helix.[22,27] The observed $+b_0$ value for poly(L-tryptophan) may arise from interactions between indole residues which are not simply additive (i.e., a conformational-dependent stacking in a helical array).[23] Positive side-chain Cotton effects may partially cancel the negative peptide Cotton effect in the range 200–240 nm and therefore may result in too low an estimate for the α-helical content. However, the far-UV ORD spectrum for helical polypeptides containing 10% of tyrosine was calculated and this curve was the same within experimental error as that for helical polypeptides lacking aromatic side chains.[106] Fasman et al.[107] showed that aromatic groups would not be expected to display significant Cotton effects if these groups comprise less than 10% of the total amino acid residues. Hence, for tryptophan content less than 10%, the effect on b_0 would be of the same order of magnitude as the experimental error in evaluating it.

In order to show that these effects do not influence the computed values of σ and s, we will first assume that the increase in $-b_0$ in TFE results only from a negative contribution of the tryptophanyl side-chain chomophore to the rotational strength [i.e., the $(-b_0)_{obsd}$ would be smaller if the guest residue did not possess an optically active side chain]. Thus we will assume that $-b_0 = 750$ for the complete helix in the copolymer and correct the observed values $(-b_0)_{obsd}$ by adding a correction $b_0'(T)$ to obtain the corrected value $(-b_0)_{corr}$ which, when divided by 750 gives θ_h [which will be lower than the value of θ_h computed by omitting the correction term $b_0'(T)$].

To exaggerate the tryptophanyl side-chain contribution all values of $(-b_0)_{obsd}$ will be decreased by a certain percentage, equal to the percent tryptophan in the particular fraction being studied [i.e., values of $(-b_0)_{obsd}$ for fractions I-1, I-2, and I-3 will undergo a 10% reduction, II-3 and II-4 a 6.5% reduction, and III-4 a 13.5% reduction], at all temperatures. The values of σ and s will then be recomputed with these new values $(-b_0)_{corr}$.

It is also possible, however, that the larger values of $-b_0$ in TFE for the copolymers as compared to the homopolymers reflect a large increase in the helix content of the copolymers which has been decreased by positive contributions to the rotational strength from the tryptophan side chains [i.e., the $(-b_0)_{obsd}$ would be larger if the guest residue had the same helix-promoting ability of tryptophan but did not possess an optically active side chain].

To test this assumption and show its effect on σ and s, the values of $(-b_0)_{obsd}$ will be increased to compensate for any positive contribution from the tryptophan side chains. Fractions I-1, I-2, and I-3 will have their values of $(-b_0)_{obsd}$ increased by 10%, II-3 and II-4 increased by 6.5%, and III-4 increased by 13.5%.

Thus, all values of $(-b_0)_{obsd}$ were corrected to $(-b_0)_{corr}$ by an addition or subtraction of $b_0'(T)$ (the magnitude of which depends on the tryptophan content of the particular fraction studied) and converted to θ_h by dividing by 750. The values of σ and s were computed from these "corrected" melting data by the same procedure (to be presented in section IV) used to analyze the "uncorrected" melting data.

Despite the intentionally exaggerated values (designed to overestimate any tryptophanyl side-chain contribution to b_0), the correction had only a moderate effect on the calculated thermodynamic quantities [an increase in $(-b_0)_{obsd}$ led to an increase in s and a reduction in σ, while a decrease in $(-b_0)_{obsd}$ led to a decrease in s and an increase in σ], their variation being within the originally determined limits of experimental error. This correction of the experimental helix contents also did not lead to any better agreement with the recomputed theoretical melting curves, either collectively (as measured by the parameter τ; see below) or in groups (e.g., for groups of polymers of specific tryptophan content). For these reasons, the Discussion which follows will focus on the results obtained by using values of $-b_0 = 750$ and 0 for the complete helix and coil, respectively, together with the experimentally measured values of $(-b_0)$, viz., $(-b_0)_{obsd}$.

IV. Discussion

Helix–Coil Parameters for Poly(L-tryptophan). The melting curves described in section III were analyzed according to the LAPS (Lifson–Allegra–Poland–Scheraga) hierarchy of approximations to obtain the Zimm–Bragg parameters σ and s for poly(L-tryptophan) in water. This procedure has been discussed extensively in previous papers in this series.[3-5] First, the approximation corresponding to the theory of Lifson[108] was used to obtain estimates of the parameters; then, better values were obtained with the second approximation, corresponding to the theory of Allegra.[109] Finally, the data from the Allegra approximation were checked with the exact theory of Lehman and McTague[110] in three representative cases. (All computer programs used in these calculations can be obtained as directed in footnotes 26 and 27 of paper 1.[3]) The results of these calculations are shown in Table V, along with the original experimental data for comparison. Both the second-order (Allegra) and the first-order (Lifson) approximations give results which agree well with those obtained from the Lehman–McTague method. The higher order Allegra approximation will be used in all subsequent discussion of the tryptophan parameters.

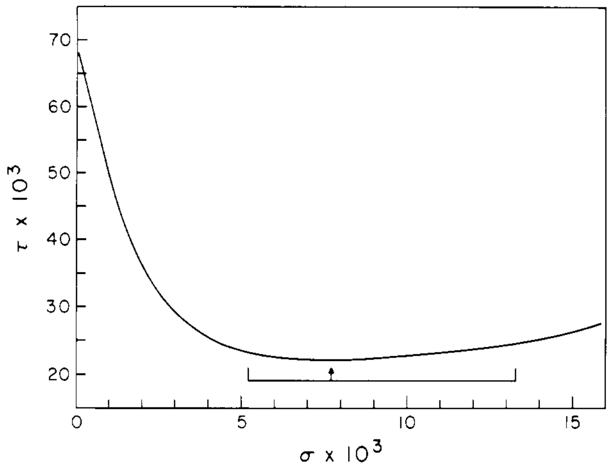

Figure 7. Determination of the best temperature-independent value of σ as the one which corresponds to the lowest value of τ for the tryptophan copolymers, using the Allegra approximation. The arrow represents this best value (7.7×10^{-3}), while the brackets represent the limits of the error in σ (discussed in section IV). The Lifson theory yielded a slightly lower value of σ, 6.8 $\times 10^{-3}$.

The close agreement between the approximate theories and the exact theory arises here because the values of s for HPG and Trp are both close to unity over the range of study and because the difference between the values of σ for P(HPG) and poly(Trp) is small and the respective values of σ are large. As was shown in paper 1,[3] this behavior causes the approximations to converge very rapidly to the exact results. In fact, as we have seen, even the first-order theory is adequate for fitting the data.

The copolymer melting data were analyzed by finding the best value of σ by application of the "goodness of fit" criterion, expressed in terms of the parameter τ defined in paper 2.[14] With σ taken as independent of temperature, the best fit of the copolymer data was obtained by minimizing τ. Figure 7 shows a plot of τ vs. σ for the approximation corresponding to the Allegra theory. The best value of σ was taken as 7.7×10^{-3}. (The value of σ obtained from the Lifson theory was 6.8×10^{-3}.) The uncertainty in the best value of σ was determined by calculating the value of this parameter for each fraction at each temperature individually, using in the calculation the values of s corresponding to the "best" σ. To represent accurately the skewness of the resulting distribution of σ values, an arithmetic average of all values of σ greater than the best value ($\sigma = 7.7 \times 10^{-3}$) was computed. Similarly, an arithmetic average was computed for all values of σ less than the best value. These arithmetic averages are shown as error limits in Figure 7. The temperature dependence of σ was computed by finding the "best" value of σ at each temperature for *all* fractions. The resulting variation of σ with temperature was found to be well within the limits of error of a temperature-*in*dependent σ. Therefore, σ was taken to be independent of temperature in the determination of $s(T)$ for L-tryptophan. The melting data for fraction III-5 have been omitted from the analysis because their inclusion increased the value of τ considerably. Since this fraction has a low value of $\overline{\mathrm{DP}}_{\mathrm{w}}$ and a high content of Trp, it may involve specific Trp–Trp interactions presumably not present in the other fractions.

The values of s at each temperature were found from both the Lifson and Allegra theories, using the appropriate "best" value of σ, and are given in Table VI. It can be seen that the agreement between the two approximate theories is very good, particularly at the higher temperatures.

Table VI
Values of the Zimm–Bragg Parameter s for Poly(L-tryptophan) in Water from 0 to 60 °C

T, °C	s		T, °C	s	
	Lifson[a]	Allegra[b]		Lifson[a]	Allegra[b]
0	1.10_4	1.12_3	35	1.03_6	1.04_0
5	1.10_5	1.12_5	40	1.02_1	1.02_2
10	1.10_2	1.12_2	45	0.99_9	0.99_6
15	1.09_4	1.11_2	50	0.97_6	0.96_9
20	1.08_9	1.10_7	55	0.95_4	0.94_5
25	1.07_0	1.08_3	60	0.93_3	0.92_1
30	1.05_6	1.06_5			

[a] Computed with the Lifson theory, using $\sigma = 6.8 \times 10^{-3}$. [b] Computed with the Allegra theory, using $\sigma = 7.7 \times 10^{-3}$.

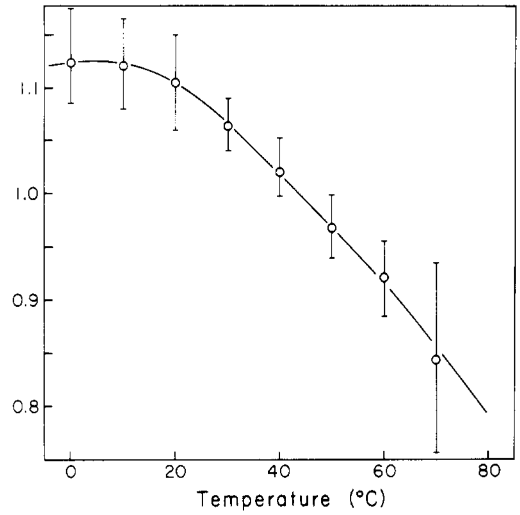

Figure 8. Plot of s vs. T for poly(L-tryptophan) in water. The error symbols are described in section IV. The solid line is drawn to pass through all the points.

In Figure 8 the temperature dependence of s is shown for the Allegra theory. The error symbols on the computed values of s are standard deviations in s at a given temperature as calculated from the values of s found, with σ held fixed ($=7.7 \times 10^{-3}$), when each fraction was fit individually at that particular temperature. The values of s calculated in this way were not found to vary in any regular way with composition, implying that there are no discernible tryptophan–tryptophan interactions in these copolymers. Such interactions should indeed be minimal for "random" copolymers of the compositions employed in these studies.

Figure 9 shows the computed melting curves, using the best-fit Allegra values, along with the experimental points. The error symbols on the computed curves arise from the errors in composition (±5%) and molecular weight (±5%); no allowance has been made for possible errors in the Zimm–Bragg parameters for P(HPG). In all cases, the agreement between the calculated and experimental values of θ_{h} is reasonably good.

The values of the thermodynamic parameters $\Delta G°$ (the free energy), $\Delta H°$ (the enthalpy), and $\Delta S°$ (the entropy) for the conversion of a coil residue of L-tryptophan to a helical one at the end of a long helical sequence can be obtained from the values of s and its temperature dependence. Figure 10 shows a plot of $\Delta G°$ ($=-RT \ln s$) vs. temperature with error symbols calculated from the standard deviation in s. The enthalpy $\Delta H°_{20}$ was calculated from this plot, and the error in this quantity is derived from the procedure used to find the slope of the $\Delta G°$ vs. temperature curve at 20 °C. The entropy at 20 °C was

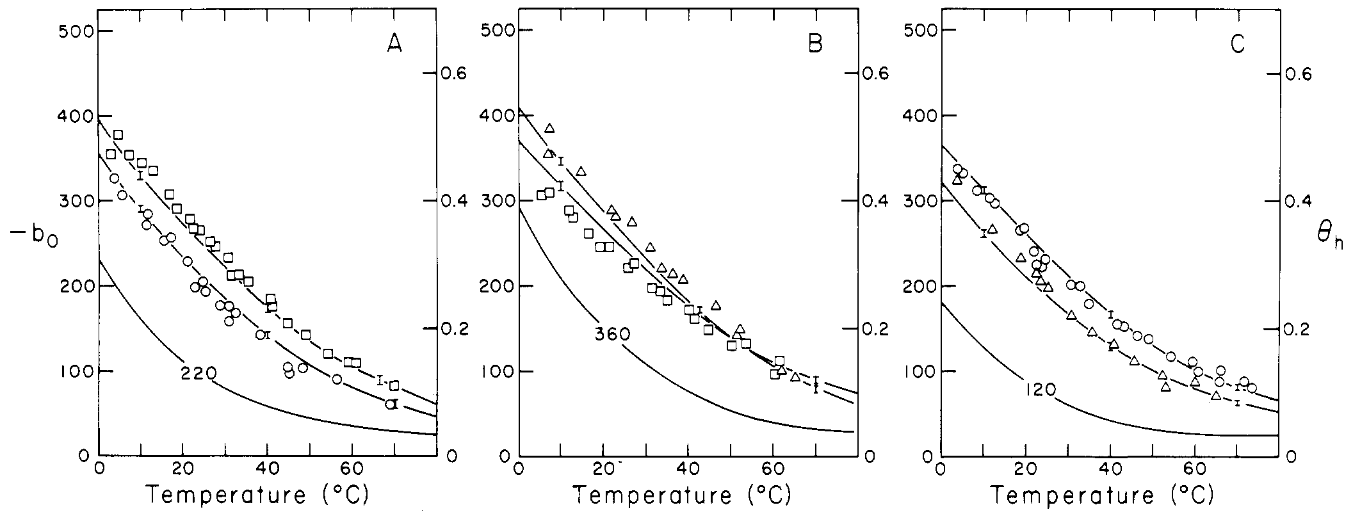

Figure 9. Comparison of the calculated melting curves obtained from the parameters of the Allegra theory (with $\sigma = 7.7 \times 10^{-3}$) for L-tryptophan given in Table VII and those of poly(HPG) of Table II in ref 4, with the experimental points for the copolymers in water. The curves for poly(HPG) $\overline{DP}_w = 120$, $\overline{DP}_w = 220$, and $\overline{DP}_w = 360$ (see ref 4) are included for comparison: (A) (□) 10.5% Trp, $\overline{DP}_w = 251$, (○) 6.5% Trp, $\overline{DP}_w = 250$; (B) (△) 13.5% Trp, $\overline{DP}_w = 143$, (□) 10.6% Trp, $\overline{DP}_w = 351$; (C) (△) 6.5% Trp, $\overline{DP}_w = 161$, (○) 10.0% Trp, $\overline{DP}_w = 194$. The error symbols indicate errors in the calculated values of θ_h arising from the errors in composition and chain length. See Figure 5 for additional errors in experimental points.

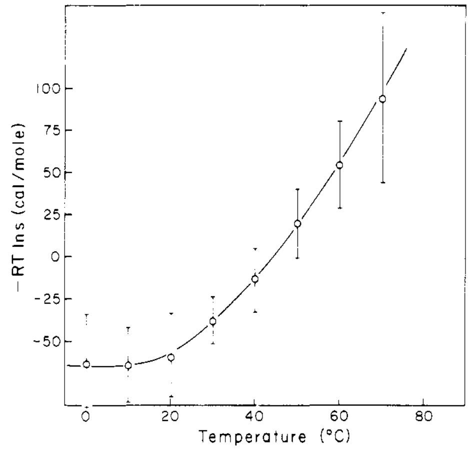

Figure 10. Plot of $-RT \ln s$ (i.e., $\Delta G°$) vs. T for poly(L-tryptophan) in water. The solid line has been drawn through the points obtained from the Allegra approximation (with $\sigma = 7.7 \times 10^{-3}$). The error symbols were calculated as described in section IV.

Table VII
Thermodynamic Parameters for L-Tryptophan

$\Delta G°_{20}$, cal/mol	-59 ± 24	$\Delta S°_{20}$, eu	-1.34 ± 0.61
$\Delta H°_{20}$, cal/mol	-452 ± 156	σ	7.7×10^{-3}

then calculated by using $\Delta G°$ and $\Delta H°$ at 20 °C. The thermodynamic parameters at 20 °C are summarized in Table VII.

The $\Delta G°$ vs. temperature curve has a minimum in the temperature range 5–30 °C. In this region, where the slope is very small, any small variation in the value of $\Delta G°$ (or s) has a large effect on the computed value of $\Delta H°$. Therefore, the values of $\Delta H°_{20}$ and $\Delta S°_{20}$ presented in Table VII should be considered only as rough estimates of these quantities.

Comparison with Other Results. Although no previous determinations of the helix–coil stability constants for L-tryptophan in water have been made, some qualitative results have been reported. Adler et al.,[31] working with random copolymers of N^5-(2-hydroxyethyl)-L-glutamine and L-tryptophan in water (temperature not reported), found that incorporation of L-tryptophan promoted helix formation in the copolymers, in qualitative agreement with our result that tryptophan is a helix former (large value of σ and $s > 1$ for $T < 45$ °C).

A quantitative measure of the tendency toward helix formation can also be obtained from X-ray data in proteins of known three-dimensional structure. Statistical analyses of the conformation of tryptophan in proteins has led to the designation of tryptophan as a helix maker.[111-124] Unfortunately, in the case of tryptophan, a representative analysis of this type (the frequency of occurrence of the α-helical conformation of tryptophan) is made difficult by the low frequency of occurrence of this residue in proteins.

Therefore, the data of these references[118-124] were recalculated as relative frequencies, using the procedure of Isogai et al.[125] The normalized relative frequency[125] of finding tryptophan in the helix was found to be 1.462, based on 8 proteins,[118] 1.139, based on 15 proteins,[119] 1.368, based on 16 proteins,[120] 1.069, based on 20 proteins,[121] 1.08, based on 29 proteins,[122] 1.105, based on 16 proteins,[123] 0.957, based on 26 proteins,[124] and 0.986, based on 23 proteins (unpublished data in connection with ref 125). If this normalized relative frequency is equal to the value of 1, then the relative frequency with which tryptophan is found in the helix is the same as the relative frequency with which the helix conformational state is found in the entire protein sample (i.e., tryptophan is indifferent with respect to being in the helical conformational state). The magnitude of the normalized relative frequency depends on the sample size, on the choice of proteins in the sample, and on the limits of the dihedral angles (ϕ, ψ) used to assign a particular residue to the α-helical state, and these factors may account for the variation in the values of the normalized relative frequencies reported above.

Relative frequencies of occurrence of interior helical states compared to all other states in proteins can be interpreted in terms of the statistical weight s.[34] It should be recognized that computations of values of s from frequencies of occurrence imply the assumption that $\sigma = 1$; i.e., no account is taken of the influence of neighboring residues when computing the frequency of occurrence of any type of residue in a helical state. Tanaka and Scheraga also point out[120,124] that any quantitative comparison between results obtained from X-ray data and those obtained from analyses of experimental data on polypeptides is premature because of the present incomplete understanding of the effects of solvents on these parameters.

From a theoretical analysis, Finkelstein et al.[126] assigned a value of $s = 1.1$ for L-tryptophan (which we interpreted, from their context, as pertaining to 20 °C). This is in excellent agreement with our experimental value of $s = 1.10_7$ at the same temperature.

Implications. Tryptophan is a widely distributed naturally occurring amino acid, essential to man and many living organisms. It contains the largest aromatic side chain of all of the amino acids that occur in proteins and contributes significantly to the optical, stereooptical, and fluorescent properties of proteins.[127] Tryptophanyl residues have been postulated to be involved in enzyme catalysis[128,129] and substrate binding[130,131] and their spectral characteristics are considered to be important gauges of protein conformation.[132-136] These residues may assume a structural role in the stabilization of the three-dimensional structure of a protein by interaction with other nonpolar side chains or, as indicated by its large value of σ, may participate in the determination of or initiation of α-helical regions in proteins.

Copolymers of poly(HPG, Trp) have been shown to be more helical than the corresponding homopolymer P(HPG) of the same chain length. The values of s for L-tryptophan, deduced from the application of the theory, indicate that L-tryptophan forms a stable helix, and the value of σ, the largest of any of the amino acids studied thus far, except for uncharged glutamic acid, indicates that, relative to other amino acids, L-tryptophan is a stronger promoter of the formation of boundaries between helix and coil states.

The effect of bulky nonpolar side chains on the stability of the α helix has been studied by Auer and Doty[137] in block and random copolymers of poly(L-leucine) and poly(L-phenylalanine). L-Leucine residues were found to form more stable α-helical systems than L-phenylalanine residues. The relative instability of L-phenylalanine residues was attributed to steric interference between the bulky side chain of L-phenylalanine and the α-helical backbone in α-helical poly(L-phenylalanine). Auer and Doty proposed that L-tryptophan could be expected to provide at least as much steric interference with the polypeptide backbone as that of L-phenylalanine and would be expected to form a less stable α helix than that of leucine. At 25 °C, L-leucine, L-phenylalanine, and L-tryptophan were found to have s values equal to 1.14,[10] 1.07,[9] and 1.08, respectively.

In aqueous solutions, bulky nonpolar side chains might also participate in favorable noncovalent interactions (i.e., hydrophobic interactions) between themselves, which are capable of enhancing the overall conformational stability of α-helical polypeptides. Auer and Doty also considered the relative helical stability of L-alanine and L-leucine.[137] In aqueous medium, the L-leucine residues form more stable helical sequences than do L-alanine residues. Molecular models suggest that, in the α helix, side-chain interactions between the isobutyl side chains of L-leucine residues may occur, whereas none are found between the methyl side chains of L-alanine residues. These associations form the basis for the stabilization of L-leucine residues in water.

Since the side chain of L-tryptophan is large enough to permit pairwise interactions with near neighbors in the α helix, L-tryptophan residues may be expected to form more stable helical sequences than do L-alanine residues. A comparison of the values of s obtained by the "host–guest" technique, at 25 °C, is 1.14[10] for L-leucine, 1.08 (this work) for L-tryptophan, and 1.06 for L-alanine.

For L-tryptophan, a combination of these two effects, (i) steric interference of the aromatic side chain with the polypeptide backbone and (ii) noncovalent interactions between the side chains themselves (hydrophobic stabilization),[138,139] results in a favorable contribution toward the overall stability of the α-helical sequences in water, as is evidenced by the value of the helix–coil stability constant s (=1.08) at 25 °C.

Acknowledgment. We are indebted to Dr. B. L. Vallee, Biophysics Research Laboratory, Department of Biological Chemistry, Harvard Medical School and the Division of Medical Biology, Peter Bent Brigham Hospital, Boston, Mass., for the use of the equipment to perform the magnetic circular dichroism measurements and to Drs. B. Holmquist, Y. Konishi, R. R. Matheson, Jr., F. R. Maxfield, G. Nêmethy, R. K. Scheule, E. R. Stimson, H. E. Van Wart, and R. W. Woody for helpful discussions. We thank Mr. G. Davenport and Mr. T. Thannhauser for technical assistance.

References and Notes

(1) This work was supported by research grants from the National Institute of Arthritis and Metabolic Diseases, U.S. Public Health Service (AM-08465), and from the National Science Foundation (PCM79-20279).

(2) (a) NIH Predoctoral Trainee. (b) Chemical Production, Pharmaceutical Division, Sandoz Ltd., Basel, Switzerland. (c) Author to whom requests for reprints should be addressed.

(3) Von Dreele, P. H.; Poland, D.; Scheraga, H. A. *Macromolecules* **1971**, *4*, 396.

(4) Von Dreele, P. H.; Lotan, N.; Ananthanarayanan, V. S.; Andreatta, R. H.; Poland, D.; Scheraga, H. A. *Macromolecules* **1971**, *4*, 408.

(5) Ananthanarayanan, V. S.; Andreatta, R. H.; Poland, D.; Scheraga, H. A. *Macromolecules* **1971**, *4*, 417.

(6) Platzer, K. E. B.; Ananthanarayanan, V. S.; Andreatta, R. H.; Scheraga, H. A. *Macromolecules* **1972**, *5*, 177.

(7) Hughes, L. J.; Andreatta, R. H.; Scheraga, H. A. *Macromolecules* **1972**, *5*, 187.

(8) Alter, J. E.; Taylor, G. T.; Scheraga, H. A. *Macromolecules* **1972**, *5*, 739.

(9) Van Wart, H. E.; Taylor, G. T.; Scheraga, H. A. *Macromolecules* **1973**, *6*, 266.

(10) Alter, J. E.; Andreatta, R. H.; Taylor, G. T.; Scheraga, H. A. *Macromolecules* **1973**, *6*, 564.

(11) Maxfield, F. R.; Alter, J. E.; Taylor, G. T.; Scheraga, H. A. *Macromolecules* **1975**, *8*, 479.

(12) Scheule, R. K.; Cardinaux, F.; Taylor, G. T.; Scheraga, H. A. *Macromolecules* **1976**, *9*, 23.

(13) Dygert, M. K.; Taylor, G. T.; Cardinaux, F.; Scheraga, H. A. *Macromolecules* **1976**, *9*, 794.

(14) Matheson, R. R., Jr.; Nemenoff, R. A.; Cardinaux, F.; Scheraga, H. A. *Biopolymers* **1977**, *16*, 1567.

(15) van Nispen, J. W.; Hill, D. J.; Scheraga, H. A. *Biopolymers* **1977**, *16*, 1587.

(16) Hill, D. J. T.; Cardinaux, F.; Scheraga, H. A. *Biopolymers* **1977**, *16*, 2447.

(17) Konishi, Y.; van Nispen, J. W.; Davenport, G.; Scheraga, H. A. *Macromolecules* **1977**, *10*, 1264.

(18) Kobayashi, Y.; Cardinaux, F.; Zweifel, B. O.; Scheraga, H. A. *Macromolecules* **1977**, *10*, 1271.

(19) Hecht, M. H.; Zweifel, B. O.; Scheraga, H. A. *Macromolecules* **1978**, *11*, 545.

(20) Zimm, B. H.; Bragg, J. K. *J. Chem. Phys.* **1959**, *31*, 526.

(21) Sela, M.; Steinberg, I. Z.; Daniel, E. *Biochim. Biophys. Acta* **1961**, *46*, 433.

(22) Fasman, G. D.; Landsberg, M.; Buchwald, M. *Can. J. Chem.* **1965**, *43*, 1588.

(23) Cosani, A.; Peggion, E.; Terbojevich, M.; Portolan, A. *Chem. Commun.* **1967**, 930.

(24) Stevens, L.; Townend, R.; Timasheff, S. N.; Fasman, G. D.; Potter, J. *Biochemistry* **1968**, *7*, 3717.

(25) Goodman, M.; Davis, G. W.; Benedetti, E. *Acc. Chem. Res.* **1968**, *1*, 275.

(26) Cosani, A.; Peggion, E.; Verdini, A. S.; Terbojevich, M. *Biopolymers* **1968**, *6*, 963.

(27) Peggion, E.; Cosani, A.; Verdini, A. S.; DelPra, A.; Mammi, M. *Biopolymers* **1968**, *6*, 1477.

(28) Peggion, E.; Fontana, A.; Cosani, A. *Biopolymers* **1969**, *7*, 517.

(29) Fasman, G. D.; Bodenheimer, E.; Pesce, A. *J. Biol. Chem.* **1966**, *241*, 916.

(30) Ohnishi, A.; Hayashi, K.; Noguchi, J. *Bull. Chem. Soc. Jpn.* **1969**, *42*, 1113.

(31) Adler, A. J.; Greenfield, N. J.; Fasman, G. D. In "Methods of Enzymology"; Hirs, C. H. W., Timasheff, S. N., Eds.; Academic Press: New York, 1973; Vol. 27, pp 712–27.

(32) Arfmann, H. A.; Labitzke, R.; Lawaczeck, R.; Wagner, K. G. *Biochemie* **1974**, *56*, 53.

(33) Scheraga, H. A. *Pure Appl. Chem.* **1973**, *36*, 1.

(34) Scheraga, H. A. *Pure Appl. Chem.* **1978**, *50*, 315.

(35) Prestige, R. L.; Harding, D. R. K.; Battersby, J. E.; Hancock, W. S. *J. Org. Chem.* **1975**, *40*, 3287.

(36) Blout, E. R.; Karlson, R. H. *J. Am. Chem. Soc.* **1956**, *78*, 941.

(37) Hirschmann, R.; Schwam, H.; Strachan, R. G.; Schoenewaldt, E. F.; Barkemeyer, H.; Miller, S. M.; Conn, J. B.; Garsky, V.; Veber, D. F.; Denkewalter, R. G. *J. Am. Chem. Soc.* **1971**, *93*, 2746.

(38) Patchornik, A.; Sela, M.; Katchalski, E. *J. Am. Chem. Soc.* **1954**, *76*, 299.

(39) Fujita, H.; Teramoto, A.; Yamashita, T.; Okita, K.; Ikeda, S. *Biopolymers* **1966**, *4*, 781.

(40) Ferger, M. F.; Jones, W. C.; Dyckes, D. F.; du Vigneaud, V. *J. Am. Chem. Soc.* **1972**, *94*, 982.

(41) Liu, T. Y.; Chang, Y. H. *J. Biol. Chem.* **1971**, *246*, 2842.

(42) Gruen, L. C.; Nicholls, P. W. *Anal. Biochem.* **1972**, *47*, 348.

(43) Hugli, T. E.; Moore, S. *J. Biol. Chem.* **1972**, *247*, 2828.

(44) Ikeda, S.; Fukui, S. *FEBS Lett.* **1974**, *41*, 216.

(45) Bennett, H. P. J.; Elliot, D. F.; Evans, B. E.; Lowry, P. J.; McMartin, C. *Biochem. J.* **1972**, *129*, 695.

(46) Spies, J. R.; Chambers, D. C. *Anal. Chem.* **1949**, *21*, 1249.

(47) Bencze, W. L.; Schmid, K. *Anal. Chem.* **1957**, *29*, 1193.

(48) Opiénska-Blauth, J.; Charęziński, M.; Berbéc, H. *Anal. Biochem.* **1963**, *6*, 69.

(49) Spande, T. F.; Witkop, B. In "Methods of Enzymology"; Hirs, C. H. W., Ed.; Academic Press: New York, 1967; Vol. 11, pp 498–506.

(50) Barman, T. E.; Koshland, D. E., Jr. *J. Biol. Chem.* **1967**, *242*, 5771.

(51) Scoffone, F.; Fontana, A.; Rocchi, R. *Biochemistry* **1968**, *7*, 971.

(52) Buccu, E.; Veronese, F. M.; Fontana, A.; Benassi, C. A. *Eur. J. Biochem.* **1970**, *13*, 188.

(53) Gaitonde, M. K.; Dovey, T. *Biochem. J.* **1970**, *117*, 907.

(54) Fridkin, M.; Patchornik, A. *Annu. Rev. Biochem.* **1974**, *43*, 419.

(55) Dalby, A.; Tsai, C. Y. *Anal. Biochem.* **1975**, *63*, 283.

(56) Hassan, S. S. M. *Anal. Chem.* **1975**, *47*, 1429.

(57) Karkhanis, Y. D.; Carlo, D. J.; Zeltner, J. *Anal. Biochem.* **1975**, *69*, 55.

(58) Goodwin, T. W.; Morton, R. A. *Biochem. J.* **1946**, *40*, 628.

(59) Edelhoch, H. *Biochemistry* **1967**, *6*, 1948.

(60) Bailey, J. E.; Beaven, G. H.; Chignell, D. A.; Gratzer, W. B. *Eur. J. Biochem.* **1968**, *7*, 5.

(61) Shelton, K. R.; Rogers, K. S. *Anal. Biochem.* **1971**, *44*, 134.

(62) Gaitonde, M. K. *Biochem. J.* **1974**, *139*, 625.

(63) Sasaki, T.; Abrams, B.; Horecker, B. L. *Anal. Biochem.* **1975**, *65*, 396.

(64) Pajot, P. *Eur. J. Biochem.* **1976**, *63*, 263.

(65) Barth, G.; Records, R.; Bunnenberg, E.; Djerassi, C.; Voelter, W. *J. Am. Chem. Soc.* **1971**, *93*, 2545.

(66) Barth, G.; Bunnenberg, E.; Djerassi, C. *Anal. Biochem.* **1972**, *48*, 471.

(67) Holmquist, B.; Vallee, B. L. *Biochemistry* **1976**, *12*, 4409.

(68) Holmquist, B.; Vallee, B. L. In "Methods of Enzymology"; Hirs, C. H. W., Timasheff, S. N., Eds.; Academic Press: New York, 1978; Vol. 49, pp 149–79.

(69) Penke, B.; Ferenczi, R.; Kovács, K. *Anal. Biochem.* **1974**, *60*, 45.

(70) Lang, C. A. *Anal. Chem.* **1958**, *30*, 1692.

(71) Noel, R. J.; Hambleton, L. G. *J. Assoc. Off. Anal. Chem.* **1976**, *59*, 134; *Chem. Abstr.* **1976**, *84*, 149347.

(72) Manning, J. M.; Moore, S. *J. Biol. Chem.* **1968**, *243*, 5591.

(73) Hirs, C. H. W.; Moore, S.; Stein, W. H. *J. Am. Chem. Soc.* **1954**, *76*, 6063.

(74) Chervenka, C. H. "A Manual of Methods for the Analytical Ultracentrifuge"; Beckman Instruments: Palo Alto, Calif., 1969; pp 47–9.

(75) Cohn, E. J.; Edsall, J. T. "Proteins, Amino Acids, and Peptides"; Reinhold: New York, 1943; p 371.

(76) Lupu-Lotan, N.; Yaron, A.; Berger, A.; Sela, M. *Biopolymers* **1965**, *3*, 625.

(77) Hill, D. J. T.; Cardinaux, F.; Scheraga, H. A. *Biopolymers* **1977**, *16*, 2469.

(78) Savige, W. E. *Aust. J. Chem.* **1971**, *24*, 1285.

(79) Pirie, A. *Biochem. J.* **1971**, *125*, 203.

(80) Savige, W. E. *Aust. J. Chem.* **1975**, *28*, 2275.

(81) Borkman, R. F. *Photochem. Photobiol.* **1977**, *26*, 163.

(82) Mehler, A. H.; Knox, W. E. *J. Biol. Chem.* **1950**, *187*, 431.

(83) Blout, E. R.; Schmier, I.; Simmons, N. S. *J. Am. Chem. Soc.* **1962**, *84*, 3193.

(84) Greenfield, N.; Davidson, B.; Fasman, G. D. *Biochemistry* **1967**, *6*, 1630.

(85) Greenfield, N.; Fasman, G. D. *Biochemistry* **1969**, *8*, 4108.

(86) Carver, J. P.; Shechter, E.; Blout, E. R. *J. Am. Chem. Soc.* **1966**, *88*, 2562.

(87) Holzwarth, G.; Doty, P. *J. Am. Chem. Soc.* **1965**, *87*, 218.

(88) Velluz, L.; Legrand, M. *Angew. Chem., Int. Ed. Engl.* **1965**, *4*, 838.

(89) Bevington, P. R. "Data Reduction and Error Analysis for the Physical Scientist"; McGraw-Hill: New York, 1969.

(90) Vournakis, J. N.; Yan, J. F.; Scheraga, H. A. *Biopolymers* **1968**, *6*, 1531.

(91) Partington, J. R. "An Advanced Treatise on Physical Chemistry"; Longmans, Green and Co.: London, 1960; Vol. IV, pp 92, 99.

(92) Okita, K.; Teramoto, A.; Fujita, H. *Biopolymers* **1970**, *9*, 717.

(93) Strickland, E. H.; Horwitz, J.; Billups, C. *Biochemistry* **1969**, *8*, 3205.

(94) Auer, H. E. *J. Am. Chem. Soc.* **1973**, *95*, 3003.

(95) Myer, Y. P.; MacDonald, L. H. *J. Am. Chem. Soc.* **1967**, *89*, 7142.

(96) Luisi, P. L.; Rizzo, V.; Lorenzi, G. P.; Straub, B.; Suter, U.; Guarnaccia, R. *Biopolymers* **1975**, *14*, 2347.

(97) Guarnaccia, R.; Rizzo, V.; Gianola, P.; Luisi, P. L. *Biopolymers* **1976**, *15*, 1103.

(98) Goux, W. J.; Kadesch, T. R.; Hooker, T. M., Jr. *Biopolymers* **1976**, *15*, 977.

(99) Rizzo, V.; Luisi, P. L.; Straub, B.; Guarnaccia, R. *Biopolymers* **1977**, *16*, 449

(100) Rizzo, V.; Luisi, P. L. *Biopolymers* **1977**, *16*, 437.

(101) Wiget, P.; Luisi, P. L. *Biopolymers* **1978**, *17*, 167.

(102) Skrabal, P.; Rizzo, V.; Baici, A.; Bangerter, F.; Luisi, P. L. *Biopolymers* **1979**, *18*, 995.

(103) Imahori, K.; Nicola, N. A. In "Physical Principles and Techniques of Protein Chemistry, Part C"; Leach, S. J., Ed.; Academic Press: New York, 1973; pp 357–444.

(104) Sears, D. W.; Beychok, S. In ref 103, pp 445–593.

(105) Woody, R. W. *J. Polym. Sci., Macromol. Rev.* **1977**, *12*, 181.

(106) Shiraki, M. Ph.D. Thesis, University of Tokyo, 1968 (cited in ref 103).

(107) Fasman, G. D.; Bodenheimer, E.; Lindblow, C. *Biochemistry* **1964**, *3*, 1665.

(108) Lifson, S. *Biopolymers* **1963**, *1*, 25.

(109) Allegra, G. *J. Polym. Sci., Part C* **1967**, *16*, 2815.

(110) Lehman, G. W.; McTague, J. P. *J. Chem. Phys.* **1968**, *49*, 3170.

(111) Kotelchuck, D.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1969**, *62*, 14.

(112) Pain, R. H.; Robson, B. *Nature (London)* **1970**, *227*, 62.

(113) Lewis, P. N.; Gō, N.; Kotelchuck, D.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1970**, *65*, 810.

(114) Robson, B.; Pain, R. H. *J. Mol. Biol.* **1971**, *58*, 237.

(115) Gō, N.; Lewis, P. N.; Gō, M.; Scheraga, H. A. *Macromolecules* **1971**, *4*, 692.

(116) Lewis, P. N.; Scheraga, H. A. *Arch. Biochem. Biophys.* 1971, *144*, 576.

(117) Lewis, P. N.; Bradbury, E. M. *Biochim. Biophys. Acta* **1974**, *336*, 153.

(118) Burgess, A. W.; Ponnuswamy, P. K.; Scheraga, H. A. *Isr. J. Chem.* **174**, *12*, 239.

(119) Chou, P. Y.; Fasman, G. D. *Biochemistry* **1974**, *13*, 211.

(120) Tanaka, S.; Scheraga, H. A. *Macromolecules* **1976**, *9*, 142.

(121) Maxfield, F. R.; Scheraga, H. A. *Biochemistry* **1976**, *15*, 5138.

(122) Fasman, G. D.; Chou, P. Y.; Adler, A. J. *Biophys. J.* **1976**, *16*, 1201.

(123) Tanaka, S.; Scheraga, H. A. *Macromolecules* **1976**, *9*, 168.

(124) Tanaka, S.; Scheraga, H. A. *Macromolecules* **1977**, *10*, 305.

(125) Isogai, Y.; Němethy, G.; Rackovsky, G.; Leach, S. J.; Scheraga, H. A., *Biopolymers* **1980**, *19*, 1183.

(126) Finkelstein, A. V.; Ptitsyn, O. B.; Kozitsyn, S. A. *Biopolymers* **1977**, *16*, 497.

(127) Fontana, A.; Toniolo, C. In "Progress in the Chemistry of Organic Natural Products"; Herz, W., Grisebach, H., Kirby, G. W., Eds.; Springer-Verlag: Wien/New York, 1976; Vol. 33, pp 309–450.

(128) Glickson, J. D.; Phillips, W. D.; Rupley, J. A. *J. Am. Chem. Soc.* **1971**, *93*, 4031.

(129) Robbins, F. M.; Holmes, L. G. *J. Biol. Chem.* **1972**, *247*, 3062.

(130) Poulos, T. L.; Price, P. A. *J. Biol. Chem.* **1971**, *246*, 4041.
(131) Robinson, G. W. *J. Biol. Chem.* **1970**, *245*, 4832.
(132) Donovan, J. W. *J. Biol. Chem.* **1969**, *244*, 1961.
(133) Herskovits, T. T.; Sorensen, M. *Biochemistry* **1968**, *7*, 2523.
(134) Herskovits, T. T.; Sorensen, M. *Biochemistry* **1968**, *7*, 2533.
(135) Wetlaufer, D. B. *Adv. Protein Chem.* **1962**, *17*, 303.
(136) Kauzmann, W. *Adv. Protein Chem.* **1959**, *14*, 1.
(137) Auer, H. E.; Doty, P. *Biochemistry* **1966**, *5*, 1716.
(138) Peggion, E.; Verdini, A. S.; Cosani, A.; Scoffone, E. *Macromolecules* **1969**, *2*, 170.
(139) Arfmann, H. A.; Labitzke, R.; Wagner, K. G. *Biopolymers* **1975**, *14*, 1381.

Differential Geometry and Polymer Conformation. 2. Development of a Conformational Distance Function[1]

S. Rackovsky[2a] **and H. A. Scheraga***[2b]

Baker Laboratory of Chemistry, Cornell University, Ithaca, New York 14853.

Received March 21, 1980

ABSTRACT: In paper 1, we described a differential-geometric representation of protein backbone structure which reflects the influence of both short- and medium-range interactions. In this paper, we continue our development of this representation, particularly of its application to the comparison of protein structures. Initially, two mathematical features of the representation are discussed. They are (a) the dependence of the curvature (κ_i) and torsion (τ_i) on the sign of p_ip_{i+1} (where p_i is the scalar product of two nonperpendicular unit vectors at the ith and $(i-1)$th α-carbon atoms, respectively), and the conditions under which p_ip_{i+1} changes sign, and (b) the existence (and size and shape) of a gap in the (κ, τ) plane, centered on the τ axis. Both of these features lead to discontinuities in κ and τ, analogous to those which exist in systems with periodic boundary conditions. Application of these results to the distribution of residues in a large protein sample in (κ, τ) space reveals the existence of a *continuum of bend structures*, ranging from α_R-like bends through flat structures to α_L-like bends. A corrected treatment of chain handedness is then given, thereby greatly increasing the utility of the differential-geometric method for determining chain handedness. A discussion is given of the inversion of the (κ, τ) representation to recover C^α coordinates, and results of such an inversion are presented for bovine pancreatic trypsin inhibitor. We then develop a function to represent the conformational distance between structures represented by any two points in the (κ, τ) plane. The utility of this distance function is demonstrated by comparison with a new superposition method for backbone segments composed of four C^α atoms which requires no computational optimization of superposition. The differential geometric comparison method is applied to two specific cases: (a) the comparison of reduced and oxidized cytochrome *c*, treated preliminarily in paper 1, and (b) the comparison of oxidized cytochrome *c* and glyceraldehyde phosphate dehydrogenase. The results demonstrate the validity and utility of the differential-geometric comparison method. This method is then used to illustrate the detection of conformational similarity between two portions of a protein molecule, the test case being two domains of ferredoxin. Finally, a general discussion of the differential-geometric representation is given. It is pointed out that this representation is complementary to other representations of protein structure in current use, in that it operates on a length scale (that of four C^α atoms) not conveniently treated by these others. It is therefore capable of revealing structural features that are not transparently evident in other representations. In addition, it is noted that comparison of the (κ, τ) and (ϕ, ψ) representations reveals a type of degeneracy in protein folding in which a given type of structure on one length scale can be attained in several ways through combinations of structures on smaller length scales.

I. Introduction

In the first paper in this series[3] (hereafter referred to as paper 1), a differential-geometric (DG) representation of protein backbone structure was presented. The fundamental unit of folding treated by this representation is a length of backbone consisting of four C^α atoms (three virtual bonds in the virtual-bond representation). The conformation of such a unit is described in terms of the curvature (κ_i), the torsion (τ_i) and the two virtual bond angles (θ_i, θ_{i+1}). It was demonstrated in paper 1 that the various types of characteristic backbone structure fall into regions of varying degrees of localization in the (κ, τ) plane, independent of the virtual bond angles. It was also pointed out that the DG representation provides a method for comparing the local folding of two backbone conformations at every point along the chain. The preliminary method suggested in paper 1 involved *separate* comparisons of κ and of τ of the two conformations at each residue.

In this paper, we improve the comparison procedure by constructing a function which represents the conformational distance between two points in the (κ, τ) plane and by using this function to measure the structural similarity of two four-C^α units. For this purpose, we first develop a number of mathematical properties of the DG representation. We also correct an error in the discussion of chain parity (handedness) in paper 1, thereby greatly extending the usefulness of the DG $\mathbf{p}_i\mathbf{p}_{i+1}$ for determining local parity.

In section II, we summarize the DG notation briefly. In section III, we begin our treatment of mathematical questions. Two points will be covered: (1) the dependence of κ_i and τ_i on the sign (τ_i), p_ip_{i+1}, and the conditions under which p_ip_{i+1} changes sign, and (2) the existence, and the size and shape, of a discontinuity in the (κ, τ) plane, centered on the τ axis. In section IV, we give a corrected treatment of parity. In section V, we discuss the inversion of the (κ, τ) representation to recover C^α coordinates. Section VI deals with the construction of the distance function in (κ, τ) space and the comparison of conformations. Two methods are presented—a superposition method which allows the determination of actual atomic deviations between two four-C^α units without the necessity for computational optimization of superposition and the DG method. The former is used as a standard by which to judge the accuracy of the latter. The DG method is then used in section VII to discuss the comparisons of several

0024-9297/80/2213-1440\$01.00/0