Macromolecules **1985,** *18,* 1069-1073 **1069**

Helix-Coil Transition in Multicomponent Random Copolypeptides in Water. 2. Application to Random Copolymers of (Hydroxybutyl)-L-glutamine, L-Phenylalanine, and L-Lysine

T. Miki,[†] A. Kidera,[†‡] M. Oka,[†] T. Hayashi,[†] A. Nakajima,*[†] Y. C. Meinwald,[‡] T. W. Thannhauser,[‡] and H. A. Scheraga*[‡]

Department of Polymer Chemistry, Kyoto University, Kyoto 606, Japan, and Baker Laboratory of Chemistry, Cornell University, Ithaca, New York 14853.
Received October 9, 1984

ABSTRACT: Water-soluble three-component random copolymers of N^5-(4-hydroxybutyl)-L-glutamine, L-phenylalanine, and L-lysine were synthesized, fractionated, and characterized. The thermally induced helix-coil transitions of the fractions in water were determined by optical rotatory dispersion measurements, and the melting behavior was analyzed with the theory reported in the preceding paper. Fairly good agreement was obtained between the experimental melting curves and the theoretical ones computed directly from the Zimm-Bragg parameters obtained previously by the host-guest method for the corresponding binary random copolymers. The validity of the assumption that the statistical weights of a residue depend on the kind of residue but not on those of its neighbors has been confirmed even for this ternary copolypeptide containing a cationic residue (Lys) as one of its components.

I. Introduction

In the first paper[1] of this series, the nearest-neighbor Ising model theory for the helix-coil transition in a binary random copolymer was extended to multicomponent random copolymers consisting of any number of components. In formulating and using this multicomponent theory, it is implied that the conformational state of a polypeptide composed of any number of components can be predicted by using the helix-coil stability constants, viz., the Zimm-Bragg parameters[2] s and σ, obtained for the 20 naturally occurring amino acids in water,[3,4] on the basis of the assumption that the statistical weights depend on the kind of amino acid residue but are independent of the kinds of its neighboring residues. In order to demonstrate the validity of this assumption, the melting behavior of water-soluble multicomponent random copolymers is examined by comparing the experimental transition curves with the theoretical ones. In the preceding paper,[1] the validity of this assumption was confirmed for three-component copolypeptides consisting of the (hydroxypropyl)-L-glutamine (HPG), L-alanine, and glycine residues. In this paper, we examine this hypothesis further by using three-component random copolymers consisting of a host, N^5-(4-hydroxybutyl)-L-glutamine (HBG), and two guest residues, L-lysine (Lys) and L-phenylalanine (Phe). The ϵ-amino group of the lysine residue is positively charged under the conditions of the measurements, and the phenylalanine residue has a large nonpolar side chain with high hydrophobicity.[5,6] The lysine content is limited to less than 10% so as to avoid long-range electrostatic repulsions between the charged lysine side chains, and this copolypeptide can serve as a primitive model of a protein molecule in that it contains charged and hydrophobic residues, and is soluble in water.

The synthesis of water-soluble random copolymers of N^5-(4-hydroxybutyl)-L-glutamine, L-phenylalanine, and L-lysine is described in section II. The experimental results and associated computations, using the parameters obtained from the host-guest technique for binary copolymers, are presented in section III. Finally, the results are discussed in section IV.

Table I
Polymerization of Glu(OBzl) NCA, Phe NCA, and Lys(Boc) NCA

	composition, mol %					
	supernatant			precipitate		
time, min	Glu	Phe	Lys	Glu	Phe	Lys
0[a]	86.1	6.0	7.9		none	
10	83.0	7.2	9.4	88.0	9.0	3.0
20	82.3	7.5	10.2	87.6	8.7	3.7
30	82.6	6.4	10.8	88.7	7.3	4.0
40	82.0	5.5	12.2	89.2	7.4	3.4
50	82.1	4.6	13.3	89.2	5.8	5.0
60	80.0	3.9	16.1	89.4	4.6	6.0
90	78.3	3.5	18.2	88.2	5.1	6.6
120	84.0	5.3	10.6	88.1	4.2	7.7

[a] The composition of the original, protected NCA's was 85% Glu, 7.5% Phe, and 7.5% Lys (cf. to "original" polymer in Table II).

II. Experimental Section

A water-soluble random copolymer of HBG, Phe, and Lys was prepared from the N-carboxyanhydrides (NCA's) of γ-benzyl L-glutamate [Glu(OBzl)], L-phenylalanine, and N-ϵ-(*tert*-butyloxycarbonyl)-L-lysine [Lys(Boc)] using sodium methoxide as an initiator. The resulting copolymer was converted to the HBG derivative by treatment with 4-amino-1-butanol. The Boc protecting group was removed with 3.9 N HCl to give the final product.

A. Materials. L-Glutamic acid was purchased from Mann Research Laboratories, and L-phenylalanine from Aldrich. N-ϵ-(*tert*-butyloxycarbonyl)-L-lysine NCA was prepared previously in our laboratory[7] and was recrystallized before use. The synthesis of the NCA's was carried out as described in our earlier publications.[7,8]

B. Syntheses. 1. Poly[Glu(OBzl),Phe,Lys(Boc)]. A mixture of Glu(OBzl) NCA (2.24 g, 8.5 mmol), Phe NCA (0.14 g, 0.75 mmol), and Lys(Boc) NCA (0.20 g, 0.75 mmol) was dissolved with stirring in 80 mL of dioxane (freshly distilled from sodium). The reaction mixture was protected with a drying tube. Polymerization was initiated with 0.25 mL of 1 M sodium methoxide in benzene (A/I ratio = 40).

The polymerization was monitored for randomness by taking 0.5-mL aliquots every 10 min for the first hour and every 30 min thereafter. The aliquots were quenched in 4.5 mL of 0.1 M HCl in ethanol. The precipitated polymer was separated from the supernatant by centrifugation and washed thoroughly with ethanol. Samples of the supernatant were evaporated in vacuo, and the residues were hydrolyzed with 6 N HCl at 110 °C for 24 h.

[†] Kyoto University.
[‡] Cornell University.

0024-9297/85/2218-1069$01.50/0

Table II
Characterization of the Fractionated Copolymers

fraction no.	composition,[a] mol %			$\bar{M}_w \times 10^{-3}$ [b]	$\bar{M}_z/\bar{M}_w$ [c]	$\overline{DP}_w$
	Glu	Phe	Lys			
original	84.6	7.9	7.5			
F-2	84.9	8.6	6.5	68.4	1.79	358
F-3	84.8	8.1	7.1	55.2	1.87	289
F-4	84.3	7.9	7.8	45.5	1.43	239
F-5	84.1	7.5	8.4	35.8	1.46	188

[a] Determined by hydrolysis in 6 N HCl at 110 °C. [b] This value was obtained by conventional sedimentation equilibrium, with an extrapolation to zero concentration. [c] $\bar{M}_z/\bar{M}_w$ is reported for the sedimentation equilibrium run at the lowest concentration.

Samples of the precipitated polymer were hydrolyzed in the same manner. The results of the amino acid analysis (Table I) indicate that Phe is incorporated slightly faster than Lys in the copolymer. This small difference in the reactivities of the NCA's of the guests would not be expected to affect the melting behavior of the copolymers (see section IIIA).

After 2 h, when the NCA assay[9] indicated that the polymerization was 90% completed, the reaction mixture was poured into 800 mL of absolute ethanol with stirring. The copolymer that precipitated was collected, washed with ethanol, and dried to constant weight over P_2O_5 in vacuo. The yield was 1.68 g (78%).

The copolymer had an approximate $\overline{DP}$ of 1470, determined by the relationship of Fujita et al.[10] from viscosity measurements in dichloroacetic acid.

2. Poly[HBG,Phe,Lys]. Poly[Glu(OBzl),Phe,Lys(Boc)] (1.6 g) was treated with 4-amino-1-butanol (20 mL), and the Boc group was removed according to the procedure described previously.[7] The aminolysis reaction was completed (>99.5%) in 10 days, monitored by assaying for unexchanged benzyl groups as described earlier.[9] After dialysis against water, the polymer (1.36 g) was obtained by lyophilization. The complete removel of the Boc groups was checked by NMR spectroscopy.

C. Fractionation of Copoly(HBG,Phe,Lys). The copolymer was fractionated by adding ether to a methanol solution of the copolymer following the procedure described in a previous paper,[11] after which the fractions were dissolved in water, lyophilized, and dried in vacuo.

D. Determination of Composition. Each fraction (1.5 mg) was hydrolyzed twice for 24 h in 6 N HCl at 110 °C and 0.05 mmHg. The hydrolyzate was adjusted to pH 2.0 by addition of 6 N NaOH. Each hydrolyzate was analyzed twice, and the data are shown in Table II as the average values from two runs.

E. ORD and CD Measurements. Optical rotatory dispersion (ORD) and circular dichroism (CD) measurements were made with a Jasco J-20 recording spectropolarimeter using a 0.1-dm path length quartz cell, which was calibrated with testosterone for a wavelength check.[12] Solutions were prepared by dissolving the lyophilized material in twice-distilled water and filtering through a Millipore filter HA (0.45-μm pore size). All transition curves were obtained by both heating and cooling in order to check that the transitions were reversible. Temperature was kept constant to within ±0.2 °C throughout the range 5–60 °C. A Lorentz-type correction was applied to the data. The parameter b_0 of the Moffitt–Yang equation, obtained from the ORD data in the range 280–450 nm, was converted to the average fraction of helix states, θ_h, by means of the same equation used previously,[1] viz.

$$\theta_h = -b_0/750 \tag{1}$$

The choice of the values of 0 and −750 for b_0 for the complete coil and α-helix, respectively, was based on the results in previous papers.[7,8,11]

F. Determination of Molecular Weight. The molecular weights of the fractions were determined by conventional sedimentation equilibrium as reported earlier,[11] using interference optics with a MOM ultracentrifuge 3170-b (Hungarian Optical Works, Hungary). All measurements were made in 0.05 M phosphoric acid buffer solution at pH 7.0 and 25.0 °C in order to eliminate polyelectrolytic effects arising from the charged ϵ-amino groups of the lysine residues. The apparent weight-average molecular weight, $\bar{M}_w^{app}$, was obtained by using C_b (the concentration at the bottom) and C_m (the concentration at the meniscus), which were calculated by numerical integration, and the concentration dependence of the weight-average molecular weight, $\bar{M}_w$, was determined. The z-average molecular weight, $\bar{M}_z$, was computed from the data corresponding to the lowest concentration for each fraction. The partial specific volumes ($\bar{v}$) of the copolymer fractions, required for the calculation of molecular weights, were determined from the amino acid content as described by Cohn and Edsall.[13] The values of $\bar{v}$ = 0.816 for the HBG residue,[7,11] $\bar{v}$ = 0.77 for the Phe residue,[13] and $\bar{v}$ = 0.82 for the Lys residue[13] were used in the calculation of $\bar{v}$ for the copolymers. $\bar{v}$ = 0.813 mL/g was obtained for the copolymer fractions. The accuracy of the molecular weight is ±5%.

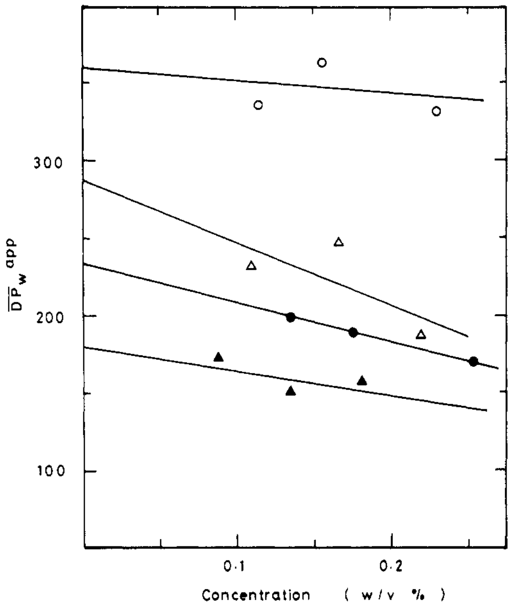

Figure 1. Determination of $\overline{DP}_w$ for the copolymer fractions from the concentration dependence of $\overline{DP}_w^{app}$. (○) F-2: $\overline{DP}_w$ = 358; (△) F-3: $\overline{DP}_w$ = 289; (●) F-4: $\overline{DP}_w$ = 239; (▲) F-5: $\overline{DP}_w$ = 188.

G. Concentration Determination. Concentrations were determined by dry weight, and experience indicates that these polymers contain from 0% to 5% moisture after "drying" in vacuo.[1] Helix contents reported here were calculated for two cases, in which the assumed moisture contents were 0% and 5%, respectively.

III. Results

A. Characterization of Copolymers. The average compositions and the average degrees of polymerization of the fractionated copolymers are shown in Table II. The average compositions of the fractionated copolymers agreed with each other and with that of the unfractionated copolymer (see also Table I). This result indicates that no large departures from random copolymerization occurred. In any case, it has been amply demonstrated theoretically in a previous paper[14] that small deviations from randomness do not influence the melting behavior of binary copolymers (see also section IV of the present paper).

The apparent degrees of polymerization $\overline{DP}_w^{app}$ were found to be concentration dependent, and extrapolation to infinite dilution was required to obtain the values of $\overline{DP}_w$ (shown in Figure 1). The $\bar{M}_z/\bar{M}_w$ ratios given in Table II indicate that the fractionation procedure yielded relatively homogeneous material.

B. ORD and CD Data. The CD spectrum for fraction F-3 at 24.0 °C is shown in Figure 2. This spectrum is ascribed to a mixture of right-handed α-helix and random coil[11] and shows no evidence of β or any other structure.[15]

The experimental values of θ_h in Figure 3 are represented by rectangles whose upper and lower sides represent the

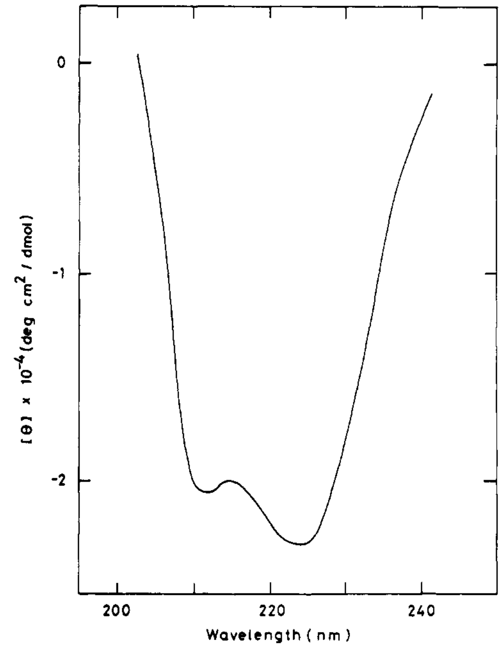

Figure 2. CD data for fraction F-3 in water at 24.0 °C.

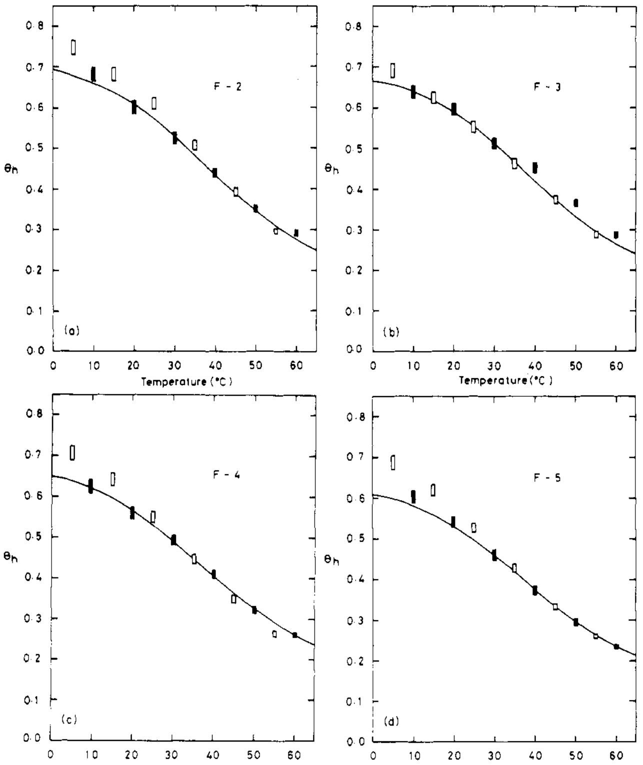

Figure 3. Comparison of the experimental data for θ_h with the calculated melting curves obtained with the exact theory,[1] using previously determined values of σ and s for HBG, Phe, and Lys. (□) Experimental points measured upon cooling; (■) experimental points measured upon heating. The upper and lower sides of the rectangles represent the experimental results obtained by assuming a moisture content of 5% and 0%, respectively. The solid curves are theoretical.

results based on assumed moisture contents of 5% and 0%, respectively.

No significant concentration dependence in the range 0.1–0.25 wt % was observed, and the curves were reproducible. No influence of long-range electrostatic repulsions between the charged lysine side chains could be detected for the melting curves of these copolymers (with less than 10% lysine content), obtained in water. These results are consistent with the previous results[7] for poly[(hydroxybutyl)glutamine-*co*-L-lysine].

Table III
Values of the Parameters s and σ Used for Calculating the Curves of Figure 3

temperature °C	values of s: HBG[a] (σ = 0.00067)	Phe[b] (σ = 0.0018)	Lys[c] (σ = 0.0001)
0	1.045	1.061	0.857
10	1.032	1.084	0.909
20	1.019	1.086	0.939
30	1.008	1.069	0.947
40	0.998	1.047	0.939
50	0.988	1.016	0.926
60	0.979	1.003	0.911

[a] Computed from the values of ΔH and ΔS given in reference 11. [b] From reference 8. [c] From reference 7.

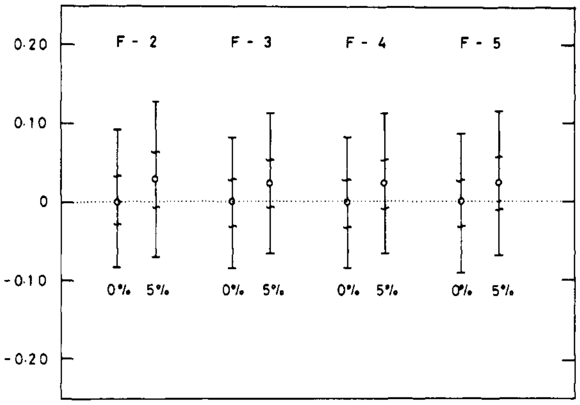

Figure 4. $\Delta\theta_h$ of eq 2 for each fraction, for assumed moisture contents in the "dried" sample of 0 and 5%, respectively. The symbols Ɨ and I represent the maximum errors in the experimental values only, and in both the experimental and theoretical values, respectively. The dashed line indicates the average value of $\Delta\theta_h$ for all the fractions (for an assumed moisture content of 0%). These values are shown in Table IV.

Figure 3 also shows the associated melting curves computed directly from the theory for finite chains; the values of s and σ for HBG, L-phenylalanine, and L-lysine (shown in Table III) were obtained from experiments on the homopolymer[11] (poly[(hydroxybutyl)glutamine] (for HBG)) and on the corresponding binary copolymers,[7,8] (for Phe and Lys). All of the computations in Figure 3 were carried out with the exact theory of Lehman and McTague[18] to avoid introducing any approximations.[1]

From an examination of these transition curves, it appears that θ_h increases with increasing chain length as expected.

The errors in the experimental points arise from errors in the determination of b_0 and from those inherent in the ORD measurements; errors in the absolute values of the optical rotation are ±3%, and those in the wavelength are ±0.5 nm. There are also errors in each of the theoretical curves; they arise from the standard deviations in the value of s as determined from the ORD data (±0.08 θ_h unit at 0 °C, and ±0.02 θ_h unit at 70 °C, at the most) and from experimental errors in the amino acid analyses and in the determination of $\overline{\mathrm{DP}}_w$ (±0.01 θ_h unit at 0 °C, and ±0.003 θ_h unit at 70 °C, at the most). The symbols Ɨ and I shown in Figure 4 represent the maximum errors in the experimental values only, and in both the experimental and theoretical values, respectively.

IV. Discussion

Previous results[1] for random copolymers of N^5-(3-hydroxypropyl)-L-glutamine, L-alanine, and glycine showed

Table IV
Values of $\Delta\theta_h$ for the Copolymer Fractions

fraction	$\Delta\theta_h(0\%)^a$	$\Delta\theta_h(5\%)^a$
F-2	0.00360	0.0298
F-3	−0.00135	0.0236
F-4	−0.00019	0.0240
F-5	0.00332	0.0262
$\Delta\theta_h^{\text{average}}$	0.00135	0.0259

[a] Assumed moisture content.

good agreement between the experimental melting curves and the theoretical ones computed with the values of σ and s obtained from experiments on the corresponding two-component random copolymers.[11,18,19] While one might have anticipated poorer agreement for such a comparison in the case of the three-component copolymer investigated here, because of the charge on the lysine residue and the hydrophobic character of the large phenylalanine side chain, the agreement is actually quite satisfactory, as indicated in Figure 3. Since the Lys and Phe contents are both less than 10%, and these residues are distributed randomly in the polymer chain, there are apparently no significant differences (between helix and coil) in long-range electrostatic interactions involving Lys side chains or stacking interactions involving Phe side chains. As a result, the theoretical curves (computed with the parameters of Table III) match the experimental data fairly well.

This agreement is demonstrated in another way in Figure 4 and Table IV in terms of the average deviations $\Delta\theta_h$ of the experimental values θ_h^{exptl} from the calculated values θ_h^{calcd}, where $\Delta\theta_h$ is defined for each fraction as

$$\Delta\theta_h = \sum_{i=1}^{12} [\theta_h^{\text{exptl}}(T_i) - \theta_h^{\text{calcd}}(T_i)]/12 \qquad (2)$$

where the average is taken over the 12 experimental points for each fraction.[21] It can be seen that $\Delta\theta_h$ is very small compared with the magnitude of the errors in this quantity, with no evidence of any trend in the data. These results support the assumption that the statistical weight of a residue is independent of the kind of its neighboring residues.

In previous papers, it was shown that no significant departures from random copolymerization occurred and that small deviations from randomness did not influence the melting behavior of the binary copolymers. Further, the results on the three-component random copolymers of N^5-(3-hydroxypropyl)-L-glutamine, L-alanine, and glycine showed that side chain–side chain interactions played a minor role in determining the transition curves.[1] To check the effect of deviations from randomness, i.e., of blocking, on the theoretical melting curves, two types of three-component copolymers of 400 residues each with a composition (mole percent) of HBG:Phe:Lys of 0.8:0.1:0.1 were generated on a computer. The first type was a random copolymer, and the second type had a random arrangement of blocks of k identical residues of each of the three components (where k was 4, 8, and 20, respectively); e.g., for k = 4, blocks of four residues each of HBG, Phe, and Lys were connected randomly, with an overall composition of 8:1:1. The deviation $\delta\theta_h$ defined by

$$\delta\theta_h(T_i) = (\theta_h)^{\text{calcd}}_{\text{block}}(T_i) - (\theta_h)^{\text{calcd}}_{\text{random}}(T_i) \qquad (3)$$

was calculated at each of several temperatures, using the parameters of Table III, and is shown in Figure 5. It can be seen from the small values of $\delta\theta_h$ (compared to those of $\Delta\theta_h$ in Figure 4) that, even for k = 20, there is little deviation from the theoretical melting curve obtained for the three-component random copolymer. Thus, the effect of deviations of this magnitude from randomness have a negligible effect on the results reported here, at least for copolymers of the composition used in this investigation.

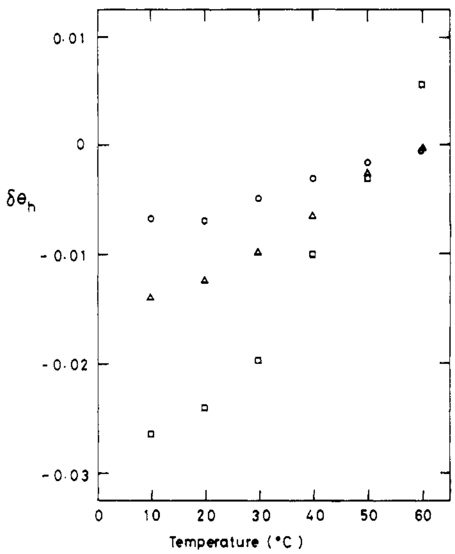

Figure 5. $\delta\theta_h$ of eq 3, calculated for copolymers with mole fractions f(HBG) = 0.8, f(Phe) = 0.1, and f(Lys) = 0.1. The chain lengths k of the block portions are k = 20 (□), k = 8 (△), and k = 4 (○).

In summary, both of the three-component copolymers (HPG–Ala–Gly reported earlier[1] and HBG–Phe–Lys investigated here) have experimental helix–coil transition curves that can be matched theoretically with values of σ and s determined from the corresponding binary random copolymers. This agreement supports the validity of the assumption that the statistical weights of the residues are independent of the kinds of their neighbors, i.e., that short-range interactions are dominant in the helix–coil transition behavior of random copolypeptides.

Acknowledgment. This work was supported by research grants from the Japan Society for the Promotion of Science (Grant MPCR-035), from the National Science Foundation (Grants PCM79-20279 and R-MPC-0043), and from the National Institute of Arthritis and Metabolic Diseases, National Institutes of Health, U.S. Public Health Service (Grant AM-08465).

Registry No. Glu(OBzl) NCA, 3190-71-4; Phe NCA, 14825-82-2; Lys(Boc) NCA, 33043-60-6; poly[Glu(OBzl),Phe,Lys(Boc)], 96412-76-9; 4-amino-1-butanol, 13325-10-5.

References and Notes

(1) Kidera, A.; Mochizuki, M.; Hasegawa, R.; Hayashi, T.; Sato, H.; Nakajima, A.; Fredrickson, R. A.; Powers, S. P.; Lee, S.; Scheraga, H. A. *Macromolecules* **1983**, *16*, 162.
(2) Zimm, B. H.; Bragg, J. K. *J. Chem. Phys.* **1959**, *31*, 526.
(3) Experimental results of s and σ for 18 of the 20 naturally occurring amino acids have already been published, the most recent being those for His.[4] The parameters for Cys and Pro have not yet been determined.
(4) Sueki, M.; Lee, S.; Powers, S. P.; Denton, J. B.; Konishi, Y.; Scheraga, H. A. *Macromolecules* **1984**, *17*, 148.
(5) Wertz, D. H.; Scheraga, H. A. *Macromolecules* **1978**, *11*, 9.
(6) Meirovitch, H.; Rackovsky, S.; Scheraga, H. A. *Macromolecules* **1980**, *13*, 1398.
(7) Dygert, M. K.; Taylor, G. T.; Cardinaux, F.; Scheraga, H. A. *Macromolecules* **1976**, *9*, 794.
(8) Van Wart, H. E.; Taylor, G. T.; Scheraga, H. A. *Macromolecules* **1973**, *6*, 266.
(9) Scheule, R. K.; Cardinaux, F.; Taylor, G. T.; Scheraga, H. A. *Macromolecules* **1976**, *9*, 23.
(10) Fujita, H.; Teramoto, A.; Yamashita, T.; Okita, K.; Ikeda, S. *Biopolymers* **1966**, *4*, 781.
(11) von Dreele, P. H.; Lotan, N.; Ananthanarayanan, V. S.; Andreatta, R. H.; Poland, D.; Scheraga, H. A. *Macromolecules* **1971**, *4*, 408.

(12) Djerassi, C.; Riniker, R.; Riniker, B. *J. Am. Chem. Soc.* **1956**, *78*, 6377.

(13) Cohn, E. J.; Edsall, J. T. "Proteins, Amino Acids, and Peptides"; Reinhold Publishing Co.: New York, 1943; Chapter 16, p 371.

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

(15) Types I and III β-turn conformations in *dipeptides* exhibit similar CD spectra.[16] Since the backbone dihedral angles of such β-turns are similar to those of the α-helix,[17] there is no inconsistency in assigning the polymer CD spectra of Figure 2 to a mixture of α-helix and coil.

(16) Bandekar, J.; Evans, D. J.; Krimm, S.; Leach, S. J.; Lee, S.; McQuie, J. R.; Minasian, E.; Nêmethy, G.; Pottle, M. S.; Scheraga, H. A.; Stimson, E. R.; Woody, R. W. *Int. J. Pept. Protein Res.* **1982**, *19*, 187.

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

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

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

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

(21) It should be noted that, in eq 2, we use average errors, which are more significant than mean-square deviations. Figure 4 and Table IV show that the data scatter around the line for $\Delta\theta_h = 0$ rather than lying all on one side of this line. The mean-square deviations would all be positive and mask this trend in the data.

Theory of the Helix–Coil Transition in Single-Chain Polypeptides with Interhelical Contacts. The Broken α-Helical Hairpin Model

Jeffrey Skolnick†

Department of Chemistry, Washington University, St. Louis, Missouri 63130.

Received September 12, 1984

ABSTRACT: Because of the possible role of the α-helical hairpin in the early stages of globular protein folding, the theory of the helix–coil transition in single-chain polypeptides containing interhelical contacts originally developed by Poland and Scheraga (*Biopolymers* **1965**, *3*, 305) has been extended to encompass a broader range of accessible conformational states. In the "broken α-helical hairpin" model developed here, the individual polypeptide chain may possess a single pair of interacting helices, joined together by an interior random coil loop of variable length greater than or equal to some minimum value and where each interacting helix may perhaps have an appended tail containing alternating stretches of noninteracting helices and random coils. Expressions for the partition function, the overall helix content, the helix probability profiles, and the fraction of random coils in free-end random coil sequences and in the interior random coil loop between interacting helices have been developed for both the perfect matching limit where the molecule is essentially folded in half as well as the imperfect matching, out-of-register case in which any non-nearest-neighbor pair of α-helical turns is allowed to interact. The broken α-helical hairpin model has been applied to a hypothetical homopolypeptide whose intrinsic helix content in the absence of interhelical contacts is small. The effect of the inclusion of out-of-register states on the breadth of the helix–coil transition of broken α-helical hairpins is investigated. Furthermore, a comparison of the helix–coil transition in a single-chain, broken α-helical hairpin and an analogous two-chain, coiled coil is made; the helix–coil transition in the former case is seen to be substantially sharper than in the latter case and reflects the enhanced cooperativity introduced by an interior random coil loop between interacting helices.

I. Introduction

About 20 years ago, Poland and Scheraga first examined the stabilization of the α-helical conformation in an individual polypeptide chain by the formation of interhelical hydrophobic bonds and specifically treated the helix–coil transition of a single chain capable of forming multiple α-helical, hairpinlike structures.[1–3] In view of recent conjectures about the possible role of α-helices and hairpin bends in the early stages of folding in globular proteins,[4–7] we believe it is worthwhile to reexamine and extend the original treatment of Poland and Scheraga to include a broader range of conformational states, thereby making the theory less restrictive and (hopefully) more realistic. In particular, in this paper we develop the theory of the "broken α-helical hairpin" model of the helix–coil transition in which each chain may possess a pair of interacting helices joined together by an interior random coil loop of arbitrary length (the "hairpin"), and each of the interacting helices may perhaps be preceded by alternating stretches of random coils and noninteracting helices.

The broken α-helical hairpin model differs from the model originally developed by Poland and Scheraga for chains of finite length[1] in several important respects. First of all, the interior random coil loop between interacting helices may be of arbitrary length and is not arbitrarily fixed at the minimum length bend that enables the two α-helices to be in contact, estimated on the basis of Courtauld's space-filling models to be one residue.[1] Unlike the short chain limit, for chains of moderate length we demonstrate that this assertion is unduly restrictive. Secondly, the present model is not isomorphic to the DNA-type helix. In the DNA isomorphic model developed by Poland and Scheraga,[3] the only allowed helical states are those that occur on both chains; hence, helical stretches add in pairs. Here, there may be noninteracting helices as well. On the basis of previous work on the effect of loop entropy in two-chain, coiled coils, we would expect the statistical weight of interacting, helical stretches punctuated by interior random coil loops between pairs of interacting helices and not in bends to be negligible;[8–10] we a priori set them equal to zero. Hence, we develop the "broken α-helical hairpin" model. Thirdly, for the sake of simplicity we restrict the treatment to a single pair of interacting helices, whereas the number of interacting helices is arbitrary in the Poland–Scheraga formalism. (We should point out that the formalism developed below can be extended to an arbitrary number of interacting helices; the possible contribution of multiple interacting helices is qualitatively examined in section IV.) Finally, while

---

† Alfred P. Sloan Foundation Fellow.

0024-9297/85/2218-1073$01.50/0