Parameters of Helix–Coil Transition Theory for Alanine-Based Peptides of Varying Chain Lengths in Water*

J. MARTIN SCHOLTZ,[1] HONG QIAN,[2] EUNICE J. YORK,[3] JOHN M. STEWART,[3] and ROBERT L. BALDWIN[1,†]

[1]Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305; [2]Institute of Molecular Biology, University of Oregon, Eugene, Oregon 97403; and [3]Department of Biochemistry, University of Colorado Health Science Center, Denver, Colorado 80262

SYNOPSIS

Thermal unfolding curves have been measured for a series of short alanine-based peptides that contain repeating sequences and varying chain lengths. Standard helix–coil theory successfully fits the observed transition curves, even for these short peptides. The results provide values for σ, the helix nucleation constant, ΔH°, the enthalpy change on helix formation, and for $s(0^\circ C)$, the average helix propagation parameter at $0^\circ C$. The enthalpy change agrees with the value determined calorimetrically. The success of helix–coil theory in describing the unfolding transitions of short peptides in water indicates that helical propensities, or s values, can be determined from substitution experiments in short alanine-based peptides.

INTRODUCTION

Substitution experiments made with short alanine-based[1-4] and other[5] monomeric peptides or dimeric coiled-coils[6] are providing new and sometimes surprising insights into the helical propensities of the amino acids. In particular, most of the results are quite different from what was expected from estimates of helical propensities obtained by the host–guest method.[7] Here we analyze, by helix–coil theory, the thermal unfolding transitions of peptides with the generic formula Ac-Y(AEAAKA)$_k$F-NH$_2$ (A: L-alanine; E: L-glutamic acid; K: L-lysine), which have varying numbers of repeats (k) of the same unit sequence (AEAAKA).

There are two basic motivations in fitting thermal unfolding curves for short peptides to helix–coil transition theory, as regards the problem of determining accurate helical propensities. One is to find out how well standard helix–coil transition theory fits the data for thermal unfolding transitions of short peptides in water. The other is to determine the helix nucleation constant σ and to check on the calorimetrically determined value[8] of ΔH°, the enthalpy change per mole residue for helix formation. Both questions are important in determining values for the helical propensity of each amino acid, which is identified here with its value of s, the propagation parameter of helix–coil theory.

To analyze the transition curves for thermal unfolding, we use the single helical sequence approximation either of the Zimm–Bragg (ZB) theory[9] or of the Lifson–Roig (LR) theory,[10] which can be used interchangeably by means of transformations (H. Qian and J. A. Schellman, *J. Phys. Chem.*, submitted) that relate the s and σ parameters of the ZB theory to w and v of the LR theory. In the peptide size range studied here, the single-sequence approximation of each theory is a very good approximation of the more general theory (see Discussion).

MATERIALS AND METHODS

Peptide Synthesis and Purification

Peptide synthesis was performed on a Biosearch 9500 automatic synthesizer with stepwise solid phase procedures[11] using Boc/benzyl strategy and HF

cleavage. p-Methylbenzhydrylamine (MBHA; polystyrene/1% divinylbenzene) resin was used to give the C-terminal amide. Double couplings and capping by acylation with acetyl imidazole were employed routinely. A third coupling using the active ester procedure[12] was used when monitoring by the qualitative Kaiser test showed the coupling to be incomplete. The syntheses were performed with 0.4–0.8 mmole of Boc-Phe-MBHA resin. As units of (AEAAKA) were added, aliquots of the resin (0.05–0.2 mmol) were removed for the addition of tyrosine and N-terminal acylation with acetic anhydride. The crude peptides were purified first by gel filtration on G-50 Sephadex in 0.1M acetic acid or on G-15 Sephadex in 50% acetic acid, then by reverse-phase high performance liquid chromatography (HPLC) on Vydac large-pore (300 Å) C_4 resin with gradients of acetonitrile containing 0.1% trifluoroacetic acid.

Satisfactory amino acid composition for each peptide was determined by analysis on a Beckman 6300 amino acid analyzer after hydrolysis for 22 h at 110°C in 6 N HCl. Peptide purity was ascertained by reversed-phase HPLC on C_4, C_{18}, or diphenyl resins to be greater than 95%. Molecular weights were confirmed by fast-atom bombardment mass spectroscopy.

CD Measurements

CD spectra were taken on an Aviv 60DS spectropolarimeter equipped with a Hewlett-Packard 89100A temperature control unit. Cuvettes with 10- or 1-mm path lengths were employed. Ellipticity is reported as mean molar residue ellipticity $[\theta]$ (deg cm^2 dmol^{-1}), and was calibrated with (+)-10-camphorsulfonic acid.[13] CD samples were prepared by diluting aqueous stock solutions of peptide with either a buffer consisting of 1 mM sodium citrate, 1 mM sodium phosphate, 1 mM sodium borate, and 0.10M sodium chloride, or with 1 mM potassium phosphate containing 0.10M potassium fluoride. In either case the pH was adjusted with HCl and KOH or NaOH to pH 7.0 at room temperature. Stock peptide concentration was determined by measuring tyrosine absorbance in phosphate-buffered 6M guanidine hydrochloride, pH 6.0, as described.[14,15]

Application of Helix–Coil Models

The one-helical sequence form of either the ZB model[9] or the LR model[10] for the helix–coil transition of a homopolymer was used in all fitting attempts. The exact form of the expression for the ZB model is given as equation 3b in Ref. 9. The temperature dependence of s (or w) was generated using the van't Hoff relationship and solving for $s(0°C)$ and $\Delta H°$. In all cases, $\Delta H°$ and σ were assumed to be temperature independent. A nonlinear function minimization program based on standard Gauss–Newton iteration, originally developed by Michael Johnson[16] and modified for use on the Macintosh (R. Brenstein and D. W. Bolen) or IBM-PC (D. Whitman), was used to fit the experimental data to the helix–coil models.

Each formalism, the ZB model or the LR model, describes the fractional helicity at any temperature $[f_H(T)]$ in terms of four parameters: the chain length (n), a propagation parameter (s in ZB notation, w in LR), a nucleation parameter (σ in ZB notation, v in LR), and an enthalpy change $(\Delta H°)$ associated with the propagation parameter (s or w). The models differ slightly in their definition of reference states, and so the propagation parameters s and w are not numerically equivalent. The reference state in the ZB model is a coil residue whose amide group is not involved in hydrogen bonding, whereas the LR model defines states in terms of helical and nonhelical (ϕ, ψ) space. Nonetheless, a simple transformation allows the parameters from the two models to be compared directly.

The LR parameters (w and v) can be converted to the ZB parameters (s and σ) using the following expressions (H. Qian and J. A. Schellman, *J. Phys. Chem.*, submitted):

$$s = w/(1 + v) \tag{1}$$

$$\sigma = v^2/(1 + v)^4 \tag{2}$$

The two formalisms also differ in their definition of chain length and helix nucleation. In the LR model, the chain length (n) is defined as the number of residues that have peptide bonds on both sides, whereas the ZB model defines n as the number of amide units in the chain. For the peptides studied in this report, with blocked C- and N-termini, the number of amide units is one more than the number of residues. This difference in the definition of n also leads to a different physical interpretation of helix nucleation, which is reflected in the parameter σ or v. In the ZB model, helix initiation involves the formation of the first amide hydrogen bond, from residue i to residue $i + 4$. In contrast, the LR definition of helix initiation requires that three contiguous residues occupy helical (ϕ, ψ) space. These differences necessitate the use of the transformations in Eqs. (1) and (2) in order to compare the results of the two models. We report all parameters using

the ZB nomenclature, regardless of the model employed.

RESULTS

Peptide Design and Synthesis

The sequence repeat of the series of peptides used in these studies is based upon the $(i, i + 3)$E,K peptide described earlier,[17] which contains Glu and Lys separated by two alanine residues. The repeating blocks of AEAAKA are flanked by an N-terminal acetyltyrosine and a C-terminal phenylalanyl carboxamide. The blocking groups on the termini eliminate unfavorable charge–helix dipole interactions, while the tyrosine and phenylalanine residues are included to facilitate accurate determination of peptide concentration and for an internal control in amino acid analysis, respectively. The $(i, i + 3)$ spacing of the Glu,Lys residues was selected because side-chain interactions are minimal when compared to the $(i, i + 4)$ arrangement of the Glu and Lys residues, which stabilize the helix by forming intrahelical ion pairs.[17] These stabilizing interactions have not been demonstrated with the $(i, i + 3)$E,K peptides. Since we wish to investigate the properties of the helix–coil transition associated with the polypeptide backbone, we desire a peptide that contains minimal side-chain interactions. This peptide appears to be well suited for this purpose.

CD Measurements

There are three criteria that must be satisfied in order to apply the helix–coil transition theory to these peptides: each of the peptides must form an α-helix and no other organized structure in the conditions studied, helix formation must be monomolecular and not the result of aggregation or oligomerization, and the thermally induced helix to coil transition must be reversible. Figure 1 shows the CD spectrum of each of the six peptides under op-
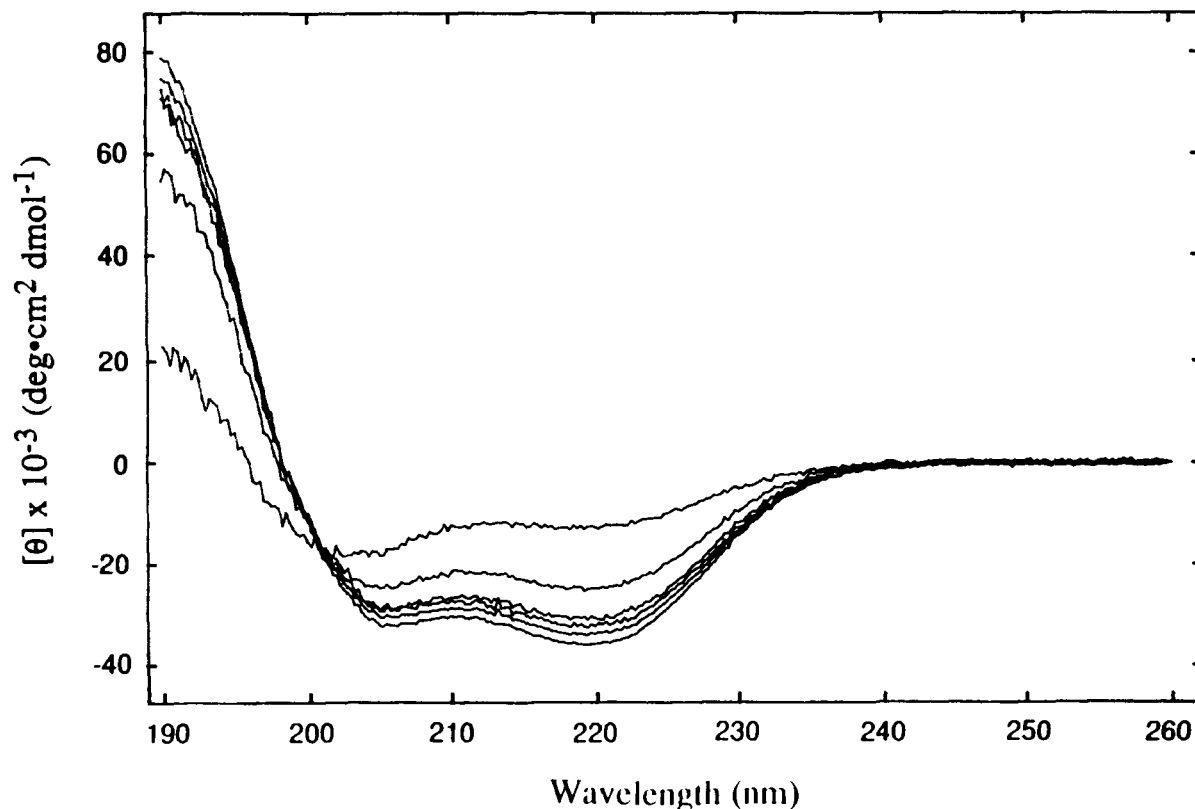

Figure 1. CD spectra of all six peptides recorded at 0°C. The spectra correspond to chain lengths of 50, 38, 32, 26, 20, and 14 residues, respectively, reading from the lower curve at 222 nm to the upper curve. The spectra were recorded at peptide concentrations of 8.5–50 μM in 1 mM potassium phosphate (pH 7.0) containing 0.10M potassium fluoride.

timal helix-forming conditions. The spectra share features that are characteristic of an α-helical structure: minima at 222 and 208 nm, and a maximum around 190 nm.[18] The presence of an isodichroic point shows that, at the resolution provided by CD spectra, each residue exists in only one of two conformations—helix or coil—regardless of the length of the peptide.

The thermally induced helix–coil transition for each peptide, as monitored by CD at 222 nm, is shown in Figure 2. Identical curves are obtained for samples that contain different peptide concentrations as well as for samples that either are heated from 0 to 80°C or cooled from 80 to 0°C. The data presented in Figures 1 and 2, along with other data,[8] suggest that the three criteria listed above are satisfied for these peptides.

Application of the Helix–Coil Models

In order to apply the models for the helix–coil transition to the data in Figure 2, the mean residue ellipticity at 222 nm, $[\theta]_{222}$, must be converted into fractional helicity. This conversion requires a knowledge of $[\theta]_{222}$ for the completely helical and completely coiled forms of each peptide at every temperature. We employ the following expressions for $[\theta]_{222}$ corresponding to the complete helix (θ_H) and the complete coil (θ_C):

$$\theta_H = -40,000 \cdot (1 - x/n) + 100 \cdot T \qquad (3)$$

$$\theta_C = +640 - 45 \cdot T \qquad (4)$$

with θ_H and θ_C expressed in deg cm^2 dmol^{-1}; T is in °C, n is the number of residues in the chain, and x
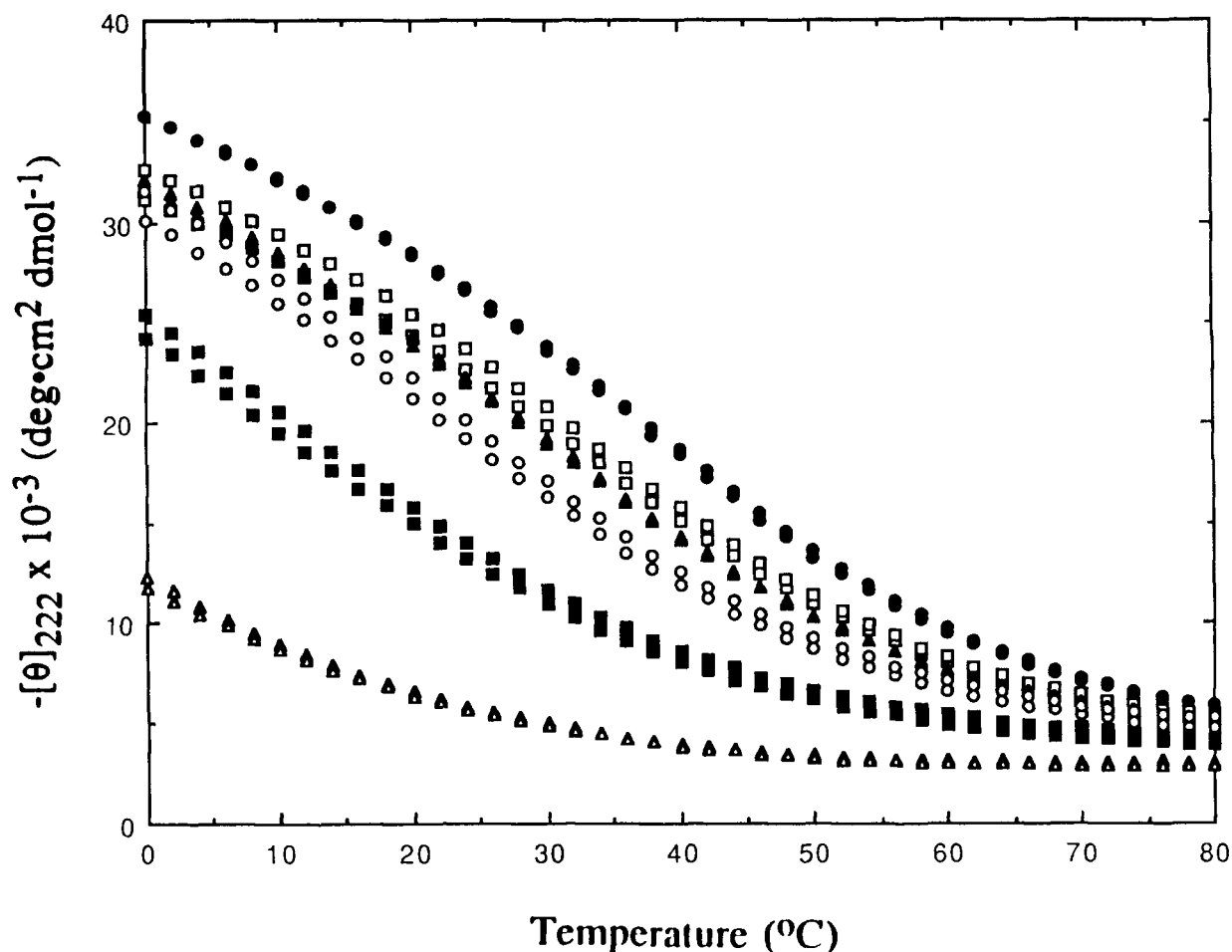

Figure 2. Thermal unfolding curves for the peptides monitored by CD. Two thermal unfolding curves for each peptide are depicted; the variation represents the uncertainty in the measurements of $-[\theta]_{222}$. Curves are shown for peptides with chain lengths of 50 (●), 38 (□), 32 (▲), 26 (○), 20 (■), and 14 (△) residues. The thermal scans were performed as described in Materials and Methods.

Table I   The Effect of the Chain-Length Dependence of θ_H on the Calculated Parameters of Helix–Coil Theory

x^a	$s(0°C)$	σ	ΔH^0 (cal/res)	Error^b
0	1.33	0.0029	−930	0.87
1	1.35	0.0028	−955	0.30
2.5	1.35	0.0033	−955	0.51
3	1.37	0.0027	−985	0.91
Mean	1.35 ± 0.02	0.0029 ± 0.0003	−960 ± 20	

^{a} The parameter x gives the dependence on chain length of θ_H, the value of $[\theta]_{222}$ for the complete helix, in Eq. (3).

^{b} The error in fitting the data to helix–coil theory, expressed as the sum of the squares of the residuals.

is a constant used to correct for nonhydrogen bonded carbonyls that do not contribute to θ_H.

The first term on the right-hand side of expression (3) or (4) is the value of $[\theta]_{222}$ for the complete helix or coil at 0°C, and the expression also gives the temperature dependence of $[\theta]_{222}$ for that structure. The expression for the complete coil [Eq. (4)] results from studies of the thermal dependence on $[\theta]_{222}$ for short (5- or 6-residue) peptides (J. M. Scholtz and R. L. Baldwin, unpublished results). The expression for the complete helix [Eq. (3)] also contains a term for the dependence of $[\theta]_{222}$ on chain length. The value of −40,000 deg cm^2 dmol^{-1} is used for the infinite helix,[19] and a chain-length dependence is introduced for shorter chains. Several different values of x were used, ranging from 0 to 3,
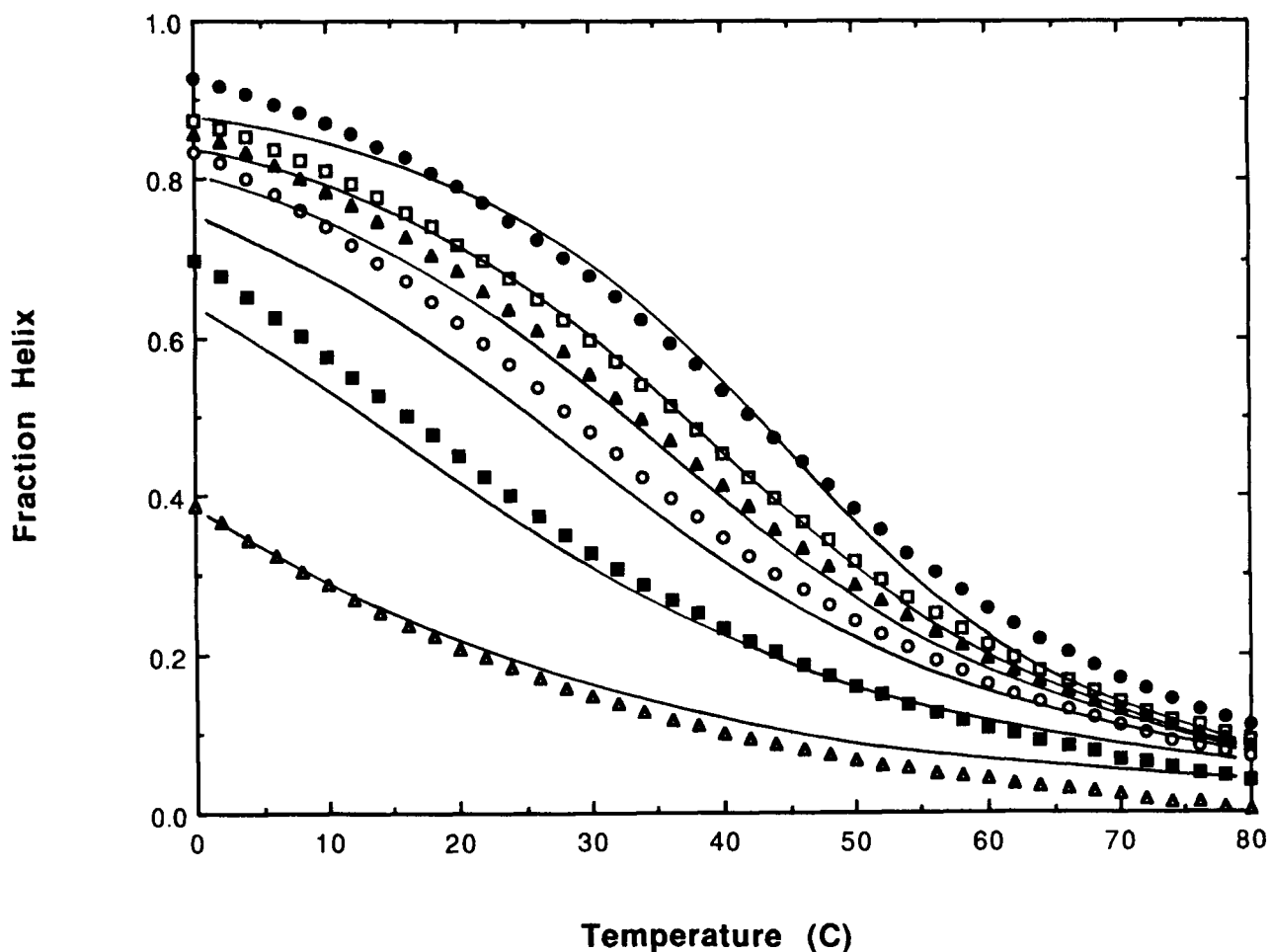

Figure 3. Comparison of the measured fractional helicity (symbols) with curves calculated using the ZB model. The symbols are the same as those in Figure 2. The data points were calculated from those shown in Figure 2 (only one thermal unfolding curve for each peptide is shown for clarity) using $x = 2.5$ in Eq. (3). The curves were generated using the ZB model with $s(0°C) = 1.35$, $\sigma = 0.0033$ and $\Delta H^0 = -955$ cal/mole residue.

and the results were compared (see below and Table I). The temperature dependence of θ_H was determined from studies of coiled-coil peptides that remain fully helical in the low temperature range (see Ref. 20 and P. S. Kim, personal communication).

The results fitting the data in Figure 2 to models for the helix–coil transition are shown in Table I for different values of the constant x in Eq. (3). Figure 3 shows, in graphical form, the experimental data (points) and the calculated curves (lines) based on the parameters in Table I for $x = 2.5$. From inspection of the results in Table I, it is clear that the parameters s, σ, and ΔH° are insensitive to the value of x, the constant expressing the chain-length dependence.

DISCUSSION

Helix–Coil Theory

In the late 1950s and early 1960s, several essentially identical models were developed to explain the α-helix to coil transition for polypeptides. Two of these formalisms, developed by Zimm and Bragg[9] and Lifson and Roig,[10] are especially well suited for our purposes, and have been used in this report. Each formalism is based on a statistical mechanical model for the α-helix to coil transition in which each residue can exist in only one of two conformations, either helix or coil. The overall transition for the entire molecule is not a two-state process from a fully helical molecule to one that is fully random coil, but rather a transition between populations of molecules that are mostly helical, with strongly frayed ends, and molecules that are almost fully random coils.

The parameters of the two helix–coil theories used here are described in Materials and Methods. The two models express the partition function for the helix–coil transition using correlation matrices; this procedure enables one to analyze heteropolymers containing residues with different helix propensities (s or w values). The complete theories allow for several stretches of helical residues in any single chain. If two simplifying assumptions are made, namely that we treat each peptide as a homopolymer and that each chain is allowed to have only one helical segment, then the fractional helicity f_H can be obtained easily from either partition function. We make these two assumptions here. Since we are concerned with short chains and helix nucleation is unfavorable, it is plausible that each chain will contain no more than one stretch of helical residues, and

comparison with the complete theory shows that this assumption is satisfactory (calculations not shown). Calculations that allow A, E, K, F and Y to have different helix propensities show that treating each peptide as a homopolymer does not affect either the applicability of standard helix–coil theory or the determination of σ and ΔH° (data not shown).

CD Measurements of Helix Content

CD has been used to measure the average helical content of a peptide at a given temperature. In order to relate $[\theta]_{222}$ to the fractional helicity, values of $[\theta]_{222}$ for the complete helix (θ_H) and the complete coil (θ_C) for each peptide must be known at every temperature. There are theoretical[21] as well as empirical[19] reasons for expecting a dependence of $[\theta]_{222}$ on the length of the complete helix, although the exact form of the dependence has not been demonstrated. Fortunately, our results (Table I) prove to be insensitive to the value of x, the length-dependence parameter in Eq. (3). Further work is required to determine the exact nature of the length dependence for the complete helix.

The temperature dependencies of $[\theta]_{222}$ for the complete helix and the coil forms of the peptide must also be known. The temperature dependence of θ_H [Eq. (3)] is obtained from the low temperature range of the thermal unfolding curves for some coiled-coil helices with high T_ms (see Ref. 20 and P. S. Kim, personal communication). A similar temperature dependence can be observed below the thermal transition zone for proteins that contain chiefly α-helical structure (J. A. Schellman, personal communication). The expression for the coil form of the peptide, θ_C [Eq. (4)] is obtained from the temperature dependence of $[\theta]_{222}$ for some short peptides (4–6 residues) that appear to be random coil at all temperatures in aqueous solution.

Helix–Coil Theory and Short Peptides

Fitting the data to theory indicates that either the ZB or LR model for the helix–coil transition adequately describes the observed thermal unfolding transitions for short peptides in water. This study of length dependence is a basic test of helix–coil theory for short peptides in aqueous solution; in earlier work, the theory has been applied primarily to the unfolding transitions of long polypeptides in nonaqueous solutions (see Ref. 22, for example). Although the theory has been used successfully for long polypeptides, it has not been clear if the same model could be applied to short peptides since the

effects of the chain ends have to be taken into account. The end effects, which will be negligible for very long polypeptides, could contribute substantially in peptides that are only 14–50 residues long. Since we have not yet studied long polypeptides, these end effects require further investigation. A second problem in applying standard helix–coil theory to transitions in aqueous solution lies in the possible dependence of s, for a given amino acid, on neighboring residues, as a result of side-chain interactions that occur especially between charged residues.

The success of standard helix–coil theory in fitting the thermal unfolding transitions of short peptides in water is significant for several reasons. The ZB and LR models can now be used to determine helix propensities of amino acids from experiments using short peptides. It will be necessary, however, to find the cause of the large differences between results found by the host–guest method, using random copolymers, and by experiments with short peptides, using substitution of defined residues in unique-sequence peptides. It has been suggested[14] that the main reason for this difference lies in the special helix-forming properties of hydroxybutyl-L-glutamine, which has been used as the host residue in the host–guest copolymers.[7] According to this explanation, the values of s determined for guest residues in host–guest experiments are strongly affected by helix-stabilizing interactions that occur among the host residues. There is also a basic difference in methodology between the host–guest experiments and the ones reported here, as regards determining the nucleation constant σ. Our approach relies on the increase in helicity of a peptide as its length increases, an effect that arises directly from the difficulty of initiating the helix. This was the original approach used by Zimm, Doty, and Iso.[22] Determination of σ in the host–guest method is based on the shape of the transition curve for long random copolymers. An implicit assumption in the treatment of data from host–guest studies is that the copolymers employed are truly random in sequence. By using peptides of defined length and sequence, we circumvent the potential problems caused by sequence and length heterogeneity.

The results of fitting the data to helix–coil theory (Table I) have confirmed the value of ΔH° for helix formation[8] determined by calorimetry, and have provided a value of the helix nucleation constant σ for short peptides in water. The results also support the finding that s values in short peptides[1–5] are much larger than those found by the host–guest technique.[7] The average s value determined here,

1.35, is an average value for the repeat AEAAKA. The s value for alanine is larger than this, since adding increasing numbers of either glutamate or lysine residues to an alanine peptide decreases its helicity.[14] This average s is significantly larger than the value found in the host–guest studies[7] and it is consistent with average s values for other alanine-based peptides.[3]

The value of σ that best fits the data for these peptides agrees with that determined for the pH-induced helix–coil transition for long homopolymers of E[23–25] and K.[26] The σ value determined for these peptides in water is, however, an order of magnitude larger than the value determined for polymers of γ-benzyl-L-glutamic acid in dichloroacetic acid–dichloroethane mixtures.[22] This difference may be caused by the solvent. It has been suggested that σ should be independent of the residue type,[9] and that σ depends only on the polypeptide backbone. This suggestion stands in contrast to the σ values determined by the host–guest method,[7] which show large variations in σ among the amino acids. Further studies are required to determine if the same value for σ can be used for all residues.

Furthermore, the results confirm the value of ΔH° determined by calorimetry.[8] The enthalpy change, in helix–coil theory, results from the temperature dependence of s, the helix propagation parameter. Helix propagation includes peptide hydrogen-bond formation as well as van der Waals and hydrophobic interactions. The calorimetric value of ΔH° is independent of the model for the helix–coil transition, and the agreement of the present value with the calorimetric one supports the applicability of standard helix–coil theory to these experiments.

In spite of the apparent success of helix–coil theory in fitting the thermal unfolding transitions of short peptides in water, there appear to be some deviations from the calculated and observed curves (Figure 3). It should be noted, that, although the theory appears to fit the observed transitions reasonably well, the values for σ and ΔH°, obtained from the fits, are subject to the assumptions employed in the analysis—namely, a temperature-independent ΔH° as well as an estimated θ_H. In order to determine more accurately θ_H, as well as ΔC_p, for helix formation, we need to investigate the helix–coil transitions of longer polypeptides. Nonetheless, our results indicate that classical helix–coil theory is able to describe the thermal unfolding transitions of short peptides in water.

REFERENCES

1. Padmanabhan, S., Marqusee, S., Ridgeway, T., Laue, T. M. & Baldwin, R. L. (1990) *Nature* (London) **344**, 268–270.
2. Padmanabhan, S. & Baldwin, R. L. (1991) *J. Mol. Biol.* **219**, 135–137.
3. Chakrabartty, A., Schellman, J. A. & Baldwin, R. L. (1991) *Nature* **351**, 586–588.
4. Merutka, G., Lipton, W., Shalongo, W., Park, S.-H. & Stellwagen, E. (1990) *Biochemistry* **29**, 7511–7515.
5. Lyu, P. C., Liff, M. I., Marky, L. A. & Kallenbach, N. R. (1990) *Science* **250**, 669–673.
6. O'Neil, K. T. & DeGrado, W. F. (1990) *Science* **250**, 646–651.
7. Wójcik, J., Altman, K.-H. & Scheraga, H. A. (1990) *Biopolymers* **30**, 121–134.
8. Scholtz, J. M., Marqusee, S., Baldwin, R. L., York, E. J., Stewart, J. M., Santoro, M. & Bolen, D. W. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 2854–2858.
9. Zimm, B. H. & Bragg, J. K. (1959) *J. Chem. Phys.* **31**, 526–535.
10. Lifson, S. & Roig, A. (1961) *J. Chem. Phys.* **34**, 1963–1974.
11. Stewart, J. M. & Young, J. D. (1984) *Solid Phase Peptide Synthesis,* Pierce Chemical Co., Rockland, IL.
12. Hudson, D. (1988) *J. Org. Chem.* **53**, 617–624.
13. Chen, G. C. & Yang, J. T. (1977) *Anal. Lett.* **10**, 1195–1207.
14. Marqusee, S., Robbins, V. H. & Baldwin, R. L. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 5286–5290.
15. Brandts, J. F. & Kaplan, L. J. (1973) *Biochemistry* **12**, 2011–2024.
16. Johnson, M. L., Correia, J. J., Yphantis, D. A. & Halvorson, H. R. (1981) *Biophys. J.* **36**, 575–588.
17. Marqusee, S. M. & Baldwin, R. L. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 8898–8902.
18. Woody, R. W. (1985) in *The Peptides,* Vol. 7, Udenfriend, S., Meienhofer, J. & Hruby, J. R., Eds., Academic Press, New York, pp. 15–114.
19. Chen, Y.-H., Yang, J. T. & Chau, K. H. (1974) *Biochemistry* **13**, 3350–3359.
20. O'Shea, E. K., Rutkowski, R. & Kim, P. S. (1989) *Science* **243**, 538–542.
21. Madison, V. & Schellman, J. (1972) *Biopolymers* **11**, 1041–1076.
22. Zimm, B. H., Doty, P. & Iso, K. (1959) *Proc. Natl. Acad. Sci. USA* **45**, 1601–1607.
23. Zimm, B. & Rice, S. (1960) *J. Mol. Phys.* **3**, 391–407.
24. Snipp, R. L., Miller, W. G. & Nylund, R. E. (1965) *J. Am. Chem. Soc.* **87**, 3547–3553.
25. Bychkova, V. E., Ptitsyn, O. B. & Barskaya, T. V. (1971) *Biopolymers* **10**, 2161–2179.
26. Bychkova, V. E. & Ptitsyn, O. B. (1971) *Biopolymers* **10**, 2181–2197.