Application of the Augmented Theory of α-Helix-to-Random-Coil Transitions of Two-Chain, Coiled Coils to Extant Data on Synthetic, Tropomyosin-Analog Peptides

ALFRED HOLTZER* and JEFFREY SKOLNICK, *Department of Chemistry and the Institute of Macromolecular Chemistry, Washington University, St. Louis, Missouri 63130*

Synopsis

The statistical mechanical theory for the helix-to-random-coil transition in two-chain coiled coils is applied to extant data for two synthetic coiled-coil polypeptides. These peptides have the primary structure $K(LEALEGK)_n$, in which $n = 4, 5$. This repeating heptet sequence mimics the pattern of hydrophobic, acidic, and basic residues characteristic of the 284-residue tropomyosin molecule, the prototypical coiled-coil protein. Theoretical calculations for single chains show that such model peptides cannot be directly compared to proteins like tropomyosin because of differences in chain length (29 and 36 residues vs 284) and in intrachain interactions, the latter caused by the differences in amino acid composition and sequence between protein and model. Application of the theory to extant data on the two synthetic peptides provides a semiquantitative fit and results in an assessment of the interhelix interaction in the model peptides. The value obtained, ~ 2000 cal · (mol of turn pairs)$^{-1}$, is four to five times larger than has been obtained for tropomyosin. This probably is a result of greater regularity in the structure of the synthetics and of the exclusive presence of leucine in the hydrophobic interface. The theory employed here insists that this powerful interhelix interaction in the synthetic is the principal reason that such short chains can be so highly helical at moderate and low temperatures. Theory predicts, indeed, that a tropomyosin-length chain with a sequence homologous to these synthetics would be completely thermally stable in the entire temperature range accessible in aqueous solutions. Theory also predicts a much more pronounced effect of concentration on the 29- and 36-residue synthetic polymers than is predicted or observed in the case of tropomyosin, and it also predicts a pronounced stabilizing effect of pH-reduction on the thermal curves. On the last two points, sufficient data are not yet available with which to test the theory.

INTRODUCTION

Coiled-coil proteins have a strikingly simple, cablelike structure, comprising two essentially completely (right-handed) α-helical polypeptide chains, in parallel and in register, and with a gentle, overall, left-handed supertwist.[1-8] This conformation is intermediate in complexity between that of globular proteins and of α-helix-forming, synthetic homopolypeptides. Thus, studies of the thermal unfolding transitions of coiled coils not only supply data on some particular proteins that are of great biochemical interest, but can serve to develop a model on which to test our ideas on protein folding in general.

Because of the essential linearity of the coiled-coil structure, the relationship between the amino acid sequence and the structure is much more readily

*To whom reprint requests should be sent.

Biopolymers, Vol. 27, 87–96 (1988)

CCC 0006-3525/88/010087-10$04.00

discernible than in the case of globular proteins.[9-12] In coiled coils, the sequence is based on a pseudo-repeating heptet of amino acids (designated by the letters a–g) in which positions a and d are predominantly hydrophobic, position e is generally an acidic residue, and g is generally a basic residue. The α-helix resulting from such a sequence is markedly amphipathic, allowing two parallel registered helices in association to be strongly stabilized by packing of hydrophobic a and d residues on one helix among corresponding a′ and d′ residues on the other. Moreover, a negatively charged e residue of the nth heptet on one helix interacts electrostatically and favorably with a positively charged g residue of the $n - 1$th heptet on the other, providing not only additional stabilization but also the discrimination that dictates parallel over antiparallel assembly.[9-12]

Although these qualitative ideas are well established, no corresponding quantitative evaluation can yet be made definitively. There is, for example, no certain way as yet to measure this interhelix interaction directly by experiment. However, a statistical-mechanical theory of the helix–coil transition has been developed that allows that interaction to be dissected from data for helix content vs temperature.[13-20] This theory is closely related to the Zimm–Bragg helix–coil theory for single-chain polypeptides,[21] as extended to chains of arbitrary sequence by Mattice.[22] The "short-range" (intrachain) helix-stabilizing interactions are characterized by two parameters (σ and s), giving the standard free energy for initiation of a helical sequence in a chain as $-kT \ln(\sigma s)$ and for propagation as $-kT \ln s$. Since values of σ and $s(T)$ for each amino acid found in proteins are available from experiments on single-chain polypeptides,[23] the intrahelix interactions may be taken as a given. Use of the statistical-mechanical theory for two-chain coiled coils in conjunction with measurements of helix content then allow extraction of the interhelix interaction.[18,20] This interaction is generally given as the standard free energy for bringing together two positionally fixed, distant helical turns, to form a positionally fixed interacting pair, and is written as $-kT \ln w(T)$. Thus, $w(T)$ characterizes the interhelix interaction, much as σ and $s(T)$ characterize the intrahelix interaction.

This theory has been applied to the best-studied coiled coil, tropomyosin.[18,20,24] Although many questions remain, the theory successfully fits a wide variety of data and explains many of their features. However, tropomyosin contains 18 different kinds of amino acid residues, resulting in a richness in interhelical contacts that is only very roughly approximated by the single site-averaged value of $w(T)$ that has thus far been obtained. A more complete view would account for the particular constellation of interfacial amino acids that appear at each helical turn.

In an endeavor to probe such details of the helix–helix interaction, specific-sequence, synthetic model peptides, such as those synthesized in Hodges' laboratory, can play a vital role.[25,26] These peptides feature a sequence designed to mimic the pattern of hydrophobic and charged groups characteristic of coiled coils. In the most recent study,[26] the peptides were of the sequence $K(LEALEGK)_n$, which we will abbreviate as KJ_n. The heptet of amino acids designated "J" precisely follows the form canonical in coiled coils, with a and d(hydrophobic) positions occupied by leucines, and e and g by glutamate and lysine, respectively.

In an earlier paper,[17] a crude early form of the statistical-mechanical theory was applied to data on a related model peptide synthesized and studied in earlier work from Hodges' laboratory.[25] This 43-residue peptide chain included a cysteine near the amino terminus and was joined to its neighboring chain by a disulfide link. Since application of the theory to cross-linked chains was particularly uncertain at that time, the results must be considered doubtful.

In the meantime, more detailed data on noncross-linked synthetic models of the KJ_n form have become available.[26] Moreover, the theory has since been augmented to include the effects of loop entropy and out-of-register conformations,[14-16] and has been extensively applied to noncross-linked tropomyosin.[20,24] In the present work, therefore, we apply this more complete theory to these noncross-linked model polymers. The uniformity of sequence in these models will allow us to characterize the helix–helix interaction for the specific case wherein hydrophobic interactions are entirely leucine–leucine and salt linkages are entirely glutamate–lysine. Moreover, the theory leads to several observations about how to account for differences in chain length and short-range interactions in drawing conclusions about proteins from such models. Finally, the theory makes predictions concerning dependence of thermal stability on pH and on polymer concentration, variables whose effects have not yet been thoroughly explored by experiment.

METHODS

Data Base

Experimental data employed were those reported by Lau et al.,[26] crucially supplemented by a communication to us from Prof. R. S. Hodges. Since tropomyosin is molecular dispersed at near-neutral pH only if the salt concentration is relatively high, detailed unfolding studies are normally carried out in the ionic strength range of 0.5–1.2M. These high ionic strengths are, in our case, necessary anyway, since they suppress charge–charge interactions, which are not included in the Zimm–Bragg theory employed here. Unfortunately, the small, synthetic, tropomyosinlike peptides are of limited supply, and the only data available to us in the proper ionic strength range are those for samples designated in Ref. 26 as TM-29 and TM-36 (in our notation, KJ_4 and KJ_5, respectively). These data refer to the solvent 1.1*M* KCl, 0.05*M* KP_i, pH 7.0. These conditions are comparable to those used in those studies of tropomyosin to which the theory has already been applied.[20,24]

The absence of appropriate data at the lower molecular weights is, however, probably not serious. The theory used here ignores all end effects and is therefore inherently polymeric in nature. This is true both with regard to the physical assumptions that enter its formalism and its realization, the latter employing the s and σ values obtained from guest inclusions in single-chain host polymers. Moreover, the heptet nature of the amino acid sequence of coiled coils requires that salt linkages exist between the fifth member of a given heptet and the seventh member of the previous heptet. Therefore, if a synthetic model chain is to be at all comparable in its helix–helix interaction per turn to the average over a long polymer like tropomyosin, it cannot

consist of only one or two heptets since end heptets are necessarily atypical. Indeed, it may be doubted that even the four- and five-heptet polymers for which we have appropriate data are sufficiently long to be quite comparable. Unfortunately, chemical synthesis of these specific sequence polymers to appreciably larger degree of polymerization is prohibitively difficult. These smaller peptides are the only models we have.

Data were obtained from Prof. Hodges in the form of mean residue ellipticity at 222 nm vs temperature. Unfortunately, data at only a single peptide concentration is available for each polymer. These concentrations were reported to us to be 0.9 mg/mL for KJ_4 and 1.0 mg/mL for KJ_5. Mean residue ellipticities were converted to fraction helix using the same algorithm employed in our laboratory for similar studies of tropomyosin.[18,27] Again, this provides maximum comparability between models and protein. However, the results are such that no CD-to-helix-content algorithm in common use would yield results significantly different from those obtained here.

THEORY

The theoretical equations and data-fitting techniques employed were exactly as used previously for the tropomyosin data.[13,20,24] Algorithms giving $s(T)$ for each amino acid residue type were the same as have been used previously[17,18]; these were designed to fit the experimental values obtained in Scheraga's laboratory.[23] Values of σ were obtained from the same source.

As in previous work,[18,20,24] a trial value of $w(T)$ was employed in the theory until one was found such that the resulting theoretical value for helix content agreed with the experimental value at that temperature. As usual, only data points showing helix content in the range of 15–90% were used because of the extreme sensitivity of $w(T)$ to helix content outside that range. All the resulting values of $RT \ln w(T)$ for both polymers were then fit, as usual, to the equation

$$-\Delta G^{\infty} = RT \ln w(T) = BT \ln T + A_0 + A_1 T \tag{1}$$

wherein B, A_0, and A_1 are constants, and ΔG^{∞} is in cal · (mol of turn pairs)$^{-1}$. The resulting algorithm was reinserted into the theory to generate theoretical curves of helix content for each peptide as a function of concentration and temperature for comparison with experiment.

RESULTS AND DISCUSSION

Single Chains

Theory allows us to explore the effects of intrachain interactions in the absence of interchain interactions. Figure 1 shows the results of calculations on single chains. It has already been pointed out that direct comparison between coiled-coil proteins and synthetic models is not possible unless chain lengths are the same.[17] In the present instance, model chains of 29 and of 36 residues are to be compared to tropomyosin, which has 284 residues/chain.

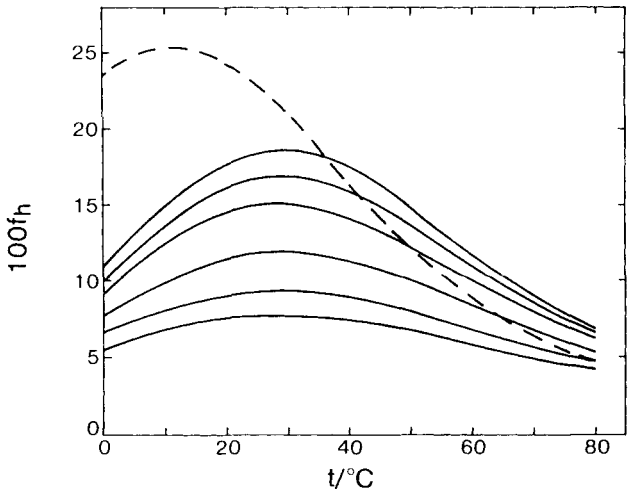

Fig. 1. Percent helix vs temperature for single chains at neutral pH. Solid curves are from theory for peptides of sequence $K(LEALEGK)_n$, with (from bottom up) n = 4, 5, 7, 12, 19, and 40. Dashed curve is from theory for tropomyosin.

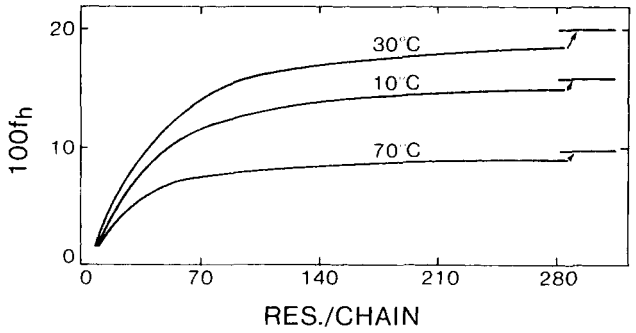

Fig. 2. Percent helix vs chain length for single chains of sequence $K(LEALEGK)_n$ at representative temperatures and neutral pH. All curves are from theory. Asymptotic ($n \rightarrow \infty$) value of helix content is shown for each temperature.

Theoretical calculations of the helix content (f_h) vs temperature for single isolated chains illustrate the substantial effect of chain length. Calculations for KJ_n polymers in which n = 4, 5, 7, 12, 19, and 40 are shown as solid curves in Fig. 1. The last has a chain length of 281 residues, whereas the largest chain actually synthesized was KJ_5 with a length of only 36 residues.

It is perhaps even more apparent from Fig. 2, directly showing helix content vs chain length, that the synthetic model peptides have chain length in a range where the helix content is most sharply dependent upon it. It is noteworthy that, according to the Zimm–Bragg theory employed here, none of these single-chain model polymers is expected to be highly helical (maximum ~ 20%), a result that serves to emphasize the importance of interhelix interactions in maintaining the full coiled-coil structure in aqueous media. We might add that a recent theory developed to supplement Zimm–Bragg theory

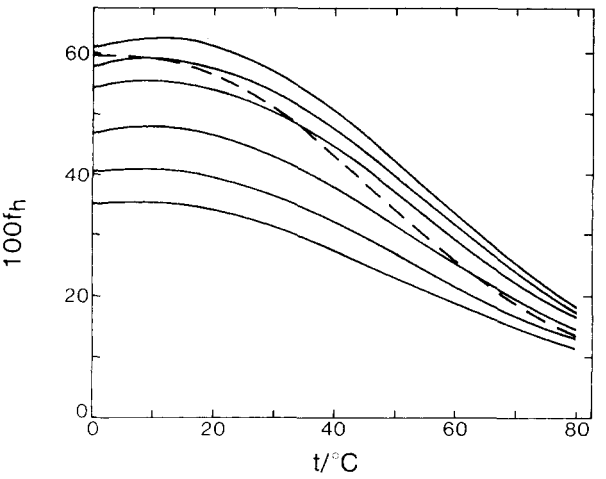

Fig. 3. Percent helix vs temperature for single chains at strongly acidic pH. Solid curves are as in Fig. 1, i.e., from theory for peptides of sequence $K(LEALEGK)_n$, with (from bottom up) $n = 4$, 5, 7, 12, 19, and 40. Dashed curve is from theory for tropomyosin.

by inclusion of physical effects not originally considered would predict an even lower helix content for these particular single-chain polymers.[28]

In addition to chain length, amino acid composition and sequence also complicate the comparison of model peptides with proteins. The results of a theoretical calculation of helix content vs temperature for tropomyosin single chains is also shown (dashed curve) on Fig. 1. Comparison of this tropomyosin curve with that for KJ_{40}, the putative model peptide of comparable chain length, dramatically illustrates the dangers of direct comparison of model and pattern peptides, even at matched chain length. The much broader range of amino acids in tropomyosin and consequent differences in the "short-range" interactions embodied in the parameters σ and $s(T)$ lead to very different expectations for the thermal unfolding curves of the 284-residue tropomyosin chain and the 281-residue KJ_{40}. The much lower helix content of KJ_{40} at lower temperatures demonstrates the effect of having a strong helix breaker (glycine) in every heptet. Tropomyosin has very few glycines. However, the decline in $s(T)$, with temperature, of leucine, alanine, glutamate, and lysine is not so great as that of the mean over the residues in tropomyosin. Hence, above room temperature, the predicted helix content of KJ_{40} single chains exceeds that for tropomyosin.

Thus, even KJ_{40}, which is a repetitive synthetic model of the appropriate chain length, is not expected to be directly comparable to tropomyosin in the thermal unfolding of its single chains. Since single chains are unavailable for experimentation (the dimer always forms spontaneously at moderate temperatures in the practical range of concentration), *some* theory is necessary in order to eliminate the effect of differences in intrachain interactions if conclusions about interchain interactions are to be drawn from studies of model peptides.

Figure 3 demonstrates how this picture changes at strongly acidic pH where all carboxyl groups are in protonated form. The usual augmentation of helix content is apparent.[19,20] This comes about, in both model and natural protein,

in spite of the increase in net charge, because of the well-known enhancement in σ and $s(T)$ for glutamic over glutamate and aspartic over aspartate.[23] The glutamic content of the synthetic KJ_n polymers is such that the enhancement overwhelms the pernicious influence of glycine; the result is that, at acidic pH, single chains of KJ_{40} are expected to have a higher helix content than tropomyosin chains even at low temperatures. Thus, material differences in behavior between the natural and synthetic polymers not only result from differences in short-range interactions, but these are themselves altered materially by pH. It is therefore absolutely necessary to have some formal way of sorting these effects out if information about tropomyosin is to be obtained from experiments on model peptides. This necessity is independent of the virtues or deficiencies of the particular theory employed here.

Coiled Coils

Experimental data for KJ_4 and KJ_5 polymers at neutral pH and high ionic strength[26] are shown in Fig. 4 as percent helix vs temperature. As expected for these homologous substances, the 36-residue chain (solid circles) has a higher helix content than the 29-residue chain (open circles), although both show very high helix content at benign temperatures. Neither thermal transition is as steep as in tropomyosin, a result also compatible with the relative shortness of the chains.

These experimental values are the only ones extant at conditions of pH and ionic strength comparable to studies of tropomyosin. They therefore form the data base for all the theoretical results reported in this section. The theoretical curves in Fig. 4 are discussed below. The discrete points on Fig. 5 show the values for the negative of the free energy of interhelix interaction required if

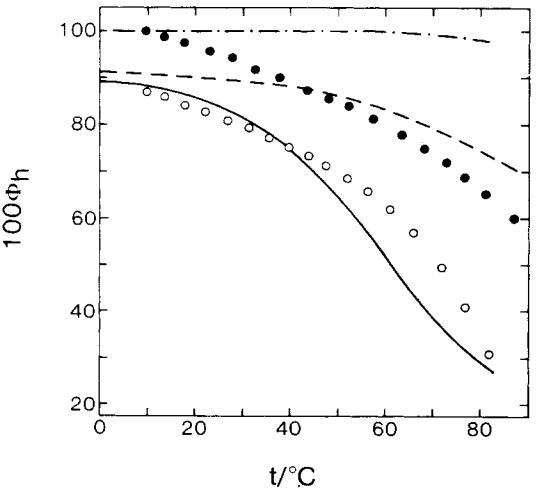

Fig. 4. Percent helix vs temperature for coiled coils at neutral pH. Data points are from Ref. 26 and supplementary information for 29-residue chain KJ_4 (open circles, 0.9 mg/mL) and 36-residue chain KJ_5 (solid circles, 1.0 mg/mL). All data are in aqueous solvent medium with 1.1*M* KCl, 0.05*M* KP_i, pH 7.0. Percent helix was calculated from ellipticity values as in Ref. 18. Full, dashed, and dot-dashed curves are from theory using the algorithm for interhelix interaction free energy given in Fig. 5. Full curve is for KJ_4 (29-residue chain) at 0.9 mg/cm^3, dashed curve is for KJ_5 (36-residue chain) at 1.0 mg/mL, and dot-dashed curve is for KJ_{40} (281-residue chains) at 1.0 mg/mL.

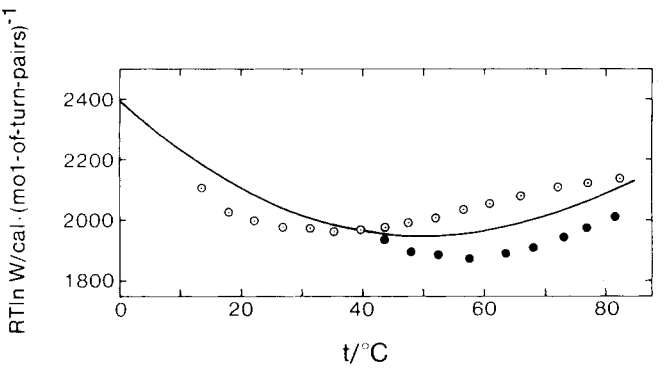

Fig. 5. Helix–helix interaction free energy in cal · (mol of turn pairs)$^{-1}$ vs temperature for synthetic coiled coils. Data points show values that must be inserted into theory to give the experimental values of helix content shown in Fig. 4. Open circles are for KJ_4, solid circles for KJ_5. Solid curve is best fit to all the points; it is given by Eq. (1) with $B = 110.656442$, $A_0 = 37{,}721.7211$, and $A_1 = -750.106658$.

the theory is to reproduce the helix content given by each corresponding data point on Fig. 4 in the feasible range of helix content. Since the same helix–helix interactions exist at each helical turn in these homologous samples, both sets of points should lie on the same curve. Instead, there seems to be a small systematic difference between the two sets. The cause for this is not known, but the difference amounts to only ~ 150 cal · (mol turn pairs)$^{-1}$, which is less than 10% of the total interaction free energy. Thus, the solid curve shown fits all the points rather well.

The helix–helix interaction free energy is therefore well approximated by the algorithm for the solid curve, i.e., with Eq. (1) and the values $B = 110.656442$, $A_0 = 37{,}721.7211$, and $A_1 = -750.106658$. We present these values to nine figures because they were the actual values used in the calculations and because rounding off individual parameters in a three-term algorithm can lead to surprisingly large discrepancies, particularly when all terms are not of the same algebraic sign. We certainly do *not* harbor the view that the resulting free energy is of nine-figure significance. Indeed, we would be ecstatic if it came within 20%. The general shape of the curve is similar to those found previously.[20,24] The shape of the free energy curve has implications for the corresponding entropy and enthalpy of interaction, but a firm physical interpretation in molecular terms is still lacking.[17]

The most striking feature of this helix–helix interaction is its magnitude, which is some four or five times the mean value obtained for tropomyosin.[20,24] This is probably a result of two factors. For one, the extreme regularity of the synthetic polymers likely promotes better packing and assures that every heptet has the full canonical complement of hydrophobic and electrostatic interactions. In tropomyosin, many heptets possess one or more noncanonical residue types in a, d, e, or g positions, some of which can actually lead to repulsive interactions. For another, the choice of leucine for all a and d positions in the synthetic peptides assures very strong hydrophobic interactions. Moreover, it is apparent from the discussion of the last section that the high helix content of the synthetic peptides, in view of their short chains and modest intrahelix interactions, is in good part a consequence of the extremely

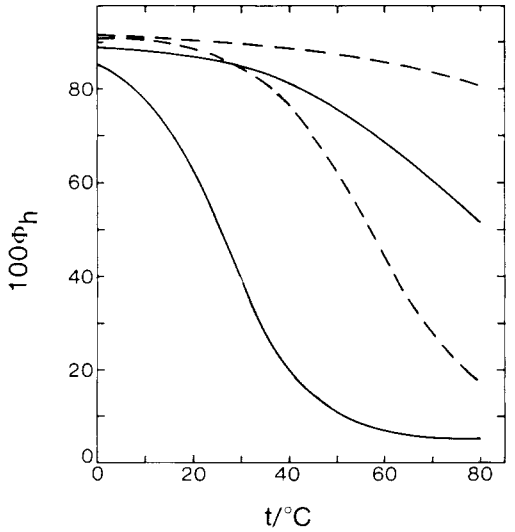

Fig. 6. Concentration dependence of percent helix vs temperature curves for coiled coils at neutral pH. All curves are from theory using the algorithm for interhelix interaction free energy given in Fig. 5. Full curves are for KJ_4 (29-residue chains), dashed curves for KJ_5 (36-residue chain). For each substance, upper curve is for 5 mg/mL, lower curve for 0.005 mg/mL.

strong interhelix interactions. This conclusion agrees with that drawn earlier from examination of a cross-linked peptide by a much cruder theory.[17]

The extent to which the single algorithm for helix–helix interactions fits all the data may be assessed from Fig. 4, wherein theoretical curves for each polymer are shown. Although agreement is not quantitative, the observed behavior is mimicked semiquantitatively except for the KJ_5 polymer above 90% helix (below 35°C). The reason for this discrepancy is unclear, and it is not even possible to be sure at present whether the fault is in the theory or in the experiments. The data themselves show an odd dislocation for that sample at ~ 35°C.

Calculations were also made for a putative 281-residue (KJ_{40}) polymer using the same helix–helix interaction algorithm; the results are also shown on Fig. 4. It is apparent that theory predicts that such a polymer would be essentially completely helical over the entire practical temperature range. Although such a long polypeptide is unlikely to be made by chemical synthesis, it may be possible to produce it by other means. It would be of great interest to see if the result is a coiled coil that is essentially indestructable thermally, as predicted by theory.

It is unfortunate that data are so far available only for single concentrations, because the theory also allows a strong prediction to be made about concentration dependence of the thermal unfolding curves, which results from dissociation into single chains. The feasible range for CD experiments is ~ 0.005–5 mg · cm^{-3}. Using the interhelix interaction algorithm given above, theoretical calculations were made of denaturation curves for KJ_4 and KJ_5 spanning the entire concentration range. The results can be seen in Fig. 6. It is immediately evident that the effect of concentration in these short-chain polymers is very large, far greater than is seen in tropomyosin.[20,24]

In future experimental studies of such models, it will be important to test this prediction as well. This could be done either by examining the concentration dependence of the thermal curves or by molecular weight measurements, since the theory also allows calculation of the extent of dissociation, once the interhelix interaction is available.[20,29] It is, of course, also of great interest to see if similar synthetic polymers in which hydrophobes other than leucine appear in the interfacial region show different interhelix interactions.

This work was supported in part by Grant GM-20064 from the Division of General Medical Sciences, United States Public Health Service; in part by a Grant from Muscular Dystrophy Association; and in part by Grant GM-37408 from the Division of General Medical Sciences, United States Public Health Service. We wish to thank Prof. Robert Hodges for supplying us with further information concerning the CD measurements reported in Ref. 26 and for a stimulating discussion.

References

1. Fraser, R. D. B. & MacRae, T. P. (1973) *Conformation in Fibrous Proteins*, Academic, New York, pp. 419–468.
2. Cohen, C. & Szent-Györgyi, A. G. (1957) *J. Am. Chem. Soc.* **79**, 248.
3. Holtzer, A., Clark, R. & Lowey, S. (1965) *Biochemistry* **4**, 2401–2411.
4. Woods, E. (1969) *Biochemistry* **8**, 4336–4344.
5. Caspar, D. L. D., Cohen, C. & Longley, W. (1969) *J. Mol. Biol.* **41**, 87–107.
6. Johnson, P. & Smillie, L. (1975) *Biochem. Biophys. Res. Commun.* **64**, 1316–1322.
7. Lehrer, S. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3377–3381.
8. Stewart, M. (1975) *FEBS Let.* **53**, 5–7.
9. Hodges, R., Sodek, J., Smillie, L. & Jurasek, L. (1972) *Cold Spring Harbor Symp. Quant. Biol.* **37**, 299–310.
10. Mak, A., Smillie, L. & Stewart, G. (1980) *J. Biol. Chem.* **255**, 3647–3655.
11. McLachlan, A. & Stewart, M. (1975) *J. Mol. Biol.* **98**, 293–304.
12. Parry, D. (1975) *J. Mol. Biol.* **98**, 519–535.
13. Skolnick, J. & Holtzer, A. (1982) *Macromolecules* **15**, 303–314.
14. Skolnick, J. (1983) *Macromolecules* **16**, 1069–1083.
15. Skolnick, J. (1983) *Macromolecules* **16**, 1763–1770.
16. Skolnick, J. (1984) *Macromolecules* **17**, 645–658.
17. Skolnick, J. & Holtzer, A. (1982) *Macromolecules* **15**, 812–821.
18. Holtzer, M. E., Holtzer, A. & Skolnick, J. (1983) *Macromolecules* **16**, 173–180.
19. Holtzer, M. E., Holtzer, A. & Skolnick, J. (1983) *Macromolecules* **16**, 462–465.
20. Skolnick, J. & Holtzer, A. (1985) *Macromolecules* **18**, 1549–1559.
21. Zimm, B. & Bragg, J. (1959) *J. Chem. Phys.* **31**, 526–535.
22. Mattice, W. (1980) *Macromolecules* **13**, 506–511.
23. Scheraga, H. (1978) *Pure Appl. Chem.* **50**, 315–324.
24. Holtzer, A. & Skolnick, J. (1986) *Macromolecules* **19**, 1769–1770.
25. Hodges, R. S., Saund, A. K., Chong, P. C. S., St.-Pierre, S. A. & Reid, R. E. (1981) *J. Biol. Chem.* **256**, 1214–1224.
26. Lau, S. Y. M., Taneja, A. K. & Hodges, R. S. (1984) *J. Biol. Chem.* **259**, 13253–13261.
27. Isom, L., Holtzer, M. E. & Holtzer, A. (1984) *Macromolecules* **17**, 2445–2447.
28. Vasquez, M., Pincus, M. & Scheraga, H. (1987) *Biopolymers* **26**, 351–371.
29. Yukioka, S., Noda, I., Nagasawa, M., Holtzer, M. E. & Holtzer, A. (1985) *Macromolecules* **18**, 1083–1086.

Received June 5, 1987
Accepted July 29, 1987