Theory of Helix–Coil Transitions of α-Helical, Two-Chain, Coiled Coils

Jeffrey Skolnick*

Department of Chemistry, Louisiana State University, Baton Rouge, Louisiana 70803

Alfred Holtzer

Department of Chemistry, Washington University, St. Louis, Missouri 63130.
Received June 23, 1981

ABSTRACT: A theory of the helix–coil transition for in-register, two-chain, α-helical, coiled coils such as tropomyosin and paramyosin is developed. The treatment differs from those formulated previously for DNA- or collagen-like double helices; in the present treatment, isolated single chains and each of the two strands in the dimer may be partially helical. We calculate the fraction of helix in the two-chain, coiled coil as a function of Zimm–Bragg cooperativity (σ) and helix stability (s) parameters appropriate to single chains and of a new parameter w that takes account of the enhanced helix stability in a paired chain vis-à-vis an isolated chain. The importance of the quasi-repeating heptet (observed in the amino acid sequence of rabbit tropomyosin) in stabilizing the helix conformation in a two-chain, coiled coil is accounted for via a coarse-graining approximation. Singly cross-linked homopolypeptide chains and both singly cross-linked and non-cross-linked chains with the rabbit α-tropomyosin sequence are treated in detail and estimates of the stability per helical residue are given. For rabbit tropomyosin, the helix probability profile is also calculated, and the possible location of a reported extrastable region is identified. In the case of non-cross-linked molecules, general expressions for two experimentally accessible quantities, the degree of chain association and the overall helix content, are derived. Application of the treatment to tropomyosin predicts that the principal thermal transition involves simultaneous dissociation and denaturation and that any thermal transition at higher temperature must be due to melting of extrastable region(s) in isolated, single, partially α-helical chains.

I. Introduction

Paramyosin and tropomyosin are salient examples of a class of protein molecules that possess a particularly simple molecular architecture. In benign media, these molecules comprise two, parallel, in-register, α-helical chains having a supertwist with a 140-Å repeat distance.[1-15] Hence, the secondary structure is an α helix, there is no additional tertiary structure, and the quaternary structure consists of parallel, in-register, supercoiled association of the two α helices. Because of their structural simplicity, two-chain, α-helical, coiled coils are natural model systems for developing insight into the factors giving rise to conformational stability in proteins. We therefore develop a theoretical treatment for estimating the stability of two-chain coiled coils relative to the isolated single-stranded molecules; it is principally to this point that the current work is addressed.

Experiments indicate that at sufficiently low temperature, tropomyosin, Tm, exists in solution as an essentially completely α-helical, two-chain, coiled coil.[5,6,16,17] Yet, Mattice, Srinivasan, and Santiago[18] estimate from theory that the helix content of an isolated, single chain of rabbit α-tropomyosin in benign aqueous solution at room temperature would be only about 14–19%. Perhaps, the theory is faulty; but if it is not, then clearly some preferential interhelix interactions, presumably of a hydrophobic and electrostatic nature and related to the quasi-repeating heptet that is characteristic of the primary structure,[13-15] must be responsible for the enormously increased helix content in the two-chain molecule. What is therefore required is a theory of the helix–coil transition that takes account of this enhanced helix stability brought about by dimerization and simultaneously allows for the possibility of helix to random coil transitions in single-stranded as well as in dimerized chains.

The stabilization of helices by interhelical interactions has been treated in a different context by Poland and Scheraga.[19] They considered helical stabilization brought about by interhelical contacts formed within the same single-stranded molecule (e.g., a hairpin configuration). At that time, the importance of the quasi-repeating heptet in effecting helix stabilization in molecules such as Tm was unknown, and thus, unlike in our case, it was not explicitly incorporated into any of the molecular models Poland and Scheraga formulated.

We present here several models of the helix–coil transition in α-helical, coiled coils; in each the overall helix content (which is observable experimentally) is dependent upon a parameter w related to the free energy difference of a helical residue in the isolated single chain and in the two-chain, coiled coil. In the following we shall refer for convenience to the two-chain associated species (which may have any helix content) as a "dimer" and the isolated single-chain species (which also may have any helix content) as a "monomer". Two kinds of systems are considered. The first and simpler case is the cross-linked dimer, in which the two chains are connected by a single covalent bond. Cross-linking may actually be achieved in tropomyosins by oxidation of sulfhydryl groups on the cysteine residues,[10,11] so this is a practical case. The second case, that of un-cross-linked molecules, allows for an equilibrium to exist between monomers and dimers. Since all disulfide cross-links can be broken by reduction, this is also a practical case. At a given temperature, both the monomer and dimer species may be partially helical. Thus, our treatment differs from those developed previously for the helix–coil transition for simple polypeptides[20-24] in that dimers are accounted for; and it differs from those developed previously for DNA- or collagen-like helices[25-29] in that both isolated strands and double strands may be helical in the present work.

In order to make the models physically simple and mathematically tractable, an approach is used which we believe preserves the essential physics. This approach has the following features.

(1) The theory is developed in terms of the Zimm–Bragg parameters[21] σ and s appropriate to each type of amino acid in the primary structure of an isolated, single chain. Thus, any values of σ and s deemed appropriate may be employed. However, to obtain numerical results, we have

0024-9297/82/2215-0303$01.25/0

used the values compiled by Mattice et al.[18] from the published values obtained by the host–guest technique of Scheraga et al.[30-33]

(2) It is assumed that cross-linking has no effect on σ and s of the cysteine residues, simply producing a "molecular staple". Strictly speaking, this approximation must be incorrect.[34,35] However, we point out that we have made this assumption merely to obtain numerical results; the formulation of the theory does not require it, and our results can be easily modified once better values of σ and s appropriate to cross-linked chains become available.

(3) The statistical weight of a randomly coiled residue is assigned a value of unity. This cannot be rigorously correct for a coiled residue in a partially helical dimer where the possibility of loop formation exists.[20] We do expect the reduction in entropy associated with loop formation in a randomly coiled segment somewhere in a polypeptide, two-chain, coiled coil to be less than that experienced by a loop in a DNA of the same size. In the DNA case, the chains are rather tightly twisted about one another, whereas in the polypeptide case the supertwist is gentle and effects of loop formation are probably minor. This point will have to be addressed at a later date.

(4) Since completely helical, two-chain, coiled coils are composed of parallel, in-register α helices, the possibility of mismatched association between residues on different chains is excluded.[1-15] That is, we ignore out-of-register associations in partially helical dimers.

With the foregoing in mind, we formulate in section II several models of the helix–coil transition in cross-linked dimers. We begin with the simplistic assumption that any helical residue on a given chain experiences an increased stabilization relative to an isolated chain when the corresponding residue on the neighboring chain is also helical. This model is of course physically unrealistic in that it ignores the fact that in a true coiled coil only one side of each helix can be in contact with the adjacent helix.[9,13] Nevertheless, some of the essential features emerge from this oversimplified model. To improve it, several coarse-graining approximations based on the repeating heptet are developed. The fraction of helix in the dimer is then calculated for typical homopolypeptides and for rabbit α-tropomyosin, Tm(Rα). In addition, the helix probability profiles are obtained for isolated Tm(Rα) chains with overall helix content of 19% and for cross-linked Tm(Rα) with overall helix content of 50%.

In section III we treat non-cross-linked chains, thus taking into account the possibility of equilibrium between monomers and dimers. Statistical mechanical expressions for the equilibrium constant for dimerization, K, and for the fraction of helix in monomers (f_{hm}) and in dimers (f_{hd}^{u}) are derived. Moreover the concentration and temperature dependence of the overall helix content and of the degree of chain dissociation for the equilibrium mixture is related to K, f_{hm}, and f_{hd}^{u}. Finally, in section IV we summarize the salient qualitative features of the study and suggest possible directions for future theoretical and experimental work.

II. Helix–Coil Transitions in Cross-Linked, Two-Chain, Coiled Coils

Interacting Individual-Residue Model. We begin with the consideration of a two-chain, coiled coil, each chain of which contains n residues. We denote the chains as 1 and 2, respectively, and denote a randomly coiled (or helical) residue on chain j as c(j) (or h(j)).

When the ith residue (for the amino terminus, $i = 1$) on both chains is in a helical state, h(1)h(2), there is assumed to be an enhanced probability, w^i, of this conformation relative to that in the two isolated chains. That is, if we designate $P_{hd}(i)$ as the statistical weight of simultaneous occurrence of a helical conformation at residue i in both chains and if we let the statistical weight be $P_{hm}(j,i)$ for a helical conformation at residue i in the chain j when it is cross-linked to, but not interacting with, the adjacent chain, then, with k_B as Boltzmann's constant and T the Kelvin temperature

$$w^i = P_{hd}(i)/[P_{hm}(1,i)P_{hm}(2,i)] \tag{1a}$$

$$w^i = \exp(-\Delta G_{hh}^{\,i}/k_B T) \tag{1b}$$

Thus, $\Delta G_{hh}^{\,i}$ is the standard free energy change and w^i is the corresponding equilibrium constant associated with transfer of a pair of helical segments h(1)h(2) at residue i from the noninteracting, cross-linked chains to the interacting, cross-linked chains. At the current level of development, w^i is assumed to be independent of whether helical residues i occur in the middle or at the beginning of a stretch of h(1)h(2) pairs. This is consistent with the assertion that the helix initiation parameter of residue i, σ_j^i, in chain j is the same for both the isolated chain and the dimer.

We may write the relevant statistical weight matrix[20] $\mathbf{U}_i$ for the ith pair of segments in terms of helix initiation (σ_j^i) and stability (s_j^i) parameters of chain j:

$$\mathbf{U}_i = \begin{array}{c|cccc} & & & i & \\ i-1 & \text{c(1)c(2)} & \text{c(1)h(2)} & \text{h(1)c(2)} & \text{h(1)h(2)} \\ \hline \text{c(1)c(2)} & 1 & \sigma_2 s_2 & \sigma_1 s_1 & \sigma_1 s_1 \sigma_2 s_2 w \\ \text{c(1)h(2)} & 1 & s_2 & \sigma_1 s_1 & \sigma_1 s_1 s_2 w \\ \text{h(1)c(2)} & 1 & \sigma_2 s_2 & s_1 & s_1 \sigma_2 s_2 w \\ \text{h(1)h(2)} & 1 & s_2 & s_1 & s_1 s_2 w \end{array} \tag{2}$$

Of course, in the more general case, the various σ_j^i associated with the residue in the dimer might be different from that in the monomer; i.e., it might be easier or more difficult to initiate the helix in the dimer relative to the monomer. If so, it would entail a trivial modification of the fourth column of eq 2. However, at this point we take the view that until it proves necessary we shall employ the σ_j^i appropriate to the isolated chains.

Furthermore, defining the statistical matrix of the ith residue on isolated chain j by

$$\mathbf{U}_j^{\,i} = \begin{bmatrix} 1 & \sigma_j^i s_j^i \\ 1 & s_j^i \end{bmatrix} \tag{3}$$

and the three-dimensional identity matrix by

$$\mathbf{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

allows us to write $\mathbf{U}_i$ in the compact form

$$\mathbf{U}_i = \mathbf{U}_1^{\,i} \otimes \mathbf{U}_2^{\,i} \begin{bmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{0} & w^i \end{bmatrix} \tag{5}$$

where $\mathbf{U}_1^{\,i} \otimes \mathbf{U}_2^{\,i}$ is the direct-product matrix[36] formed from $\mathbf{U}_1^{\,i}$ and $\mathbf{U}_2^{\,i}$; $\mathbf{0}$ represents a rectangular field of zeroes.

Following the method of analysis presented by Flory,[37] we write the partition function for the cross-linked, two-chain molecule (the "stapled" dimer), Z_{sd}, as

$$Z_{sd} = \text{Row}(1,0,0,0)\left(\prod_{i=1}^{n} \mathbf{U}_1^{\,i} \otimes \mathbf{U}_2^{\,i} \begin{bmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{0} & w^i \end{bmatrix}\right) \text{Col}(1,1,1,1) \tag{6}$$

with Row a row vector and Col, a column vector.

Using the well-known relationship for direct-product matrices,[36]

$$(\mathbf{A}_1 \otimes \mathbf{B}_1)(\mathbf{A}_2 \otimes \mathbf{B}_2) = \mathbf{A}_1\mathbf{A}_2 \otimes \mathbf{B}_1\mathbf{B}_2$$

it is straightforward to demonstrate that if the chains do not interact, i.e., if all $w^i = 1$, then Z_{sd} reduces, as it must, to the product of the partition function of the isolated chain; i.e.

$$Z_{sd} = Z_{m1}Z_{m2} \tag{7a}$$

where

$$Z_{mj} = \text{Row } (1,0)\prod_{i=1}^{n}\mathbf{U}_j^i \text{ Col } (1,1) \tag{7b}$$

Furthermore, the fraction of helix in the stapled dimer, f_{hd}, is readily derived by elementary statistical mechanics from Flory's relationship[37] for the same quantity in the monomeric case; we obtain

$$f_{hd} = \frac{1}{2n}\sum_{j=1}^{2}\sum_{i=1}^{n}\left(\frac{\partial \ln Z_{sd}}{\partial \ln s_j^i}\right) \tag{8}$$

Following Flory,[37] we can convert eq 8 to an equivalent expression that is more suitable for numerical analysis:

$$f_{hd} = \frac{1}{2nZ_{sd}}\boldsymbol{J}^*\left[\prod_{i=1}^{n}\mathbf{U}_{si}\right]\boldsymbol{J} \tag{9}$$

wherein

$$\boldsymbol{J}^* = \text{Row } (1,0,0,0,0,0,0,0) \tag{10a}$$

and

$$\boldsymbol{J} = \text{Col } (0,0,0,0,1,1,1,1) \tag{10b}$$

and $\mathbf{U}_{si}$ is an 8 × 8 block matrix defined by

$$\mathbf{U}_{si} = \begin{bmatrix} \mathbf{U}_i & \mathbf{U}_i' \\ \mathbf{0} & \mathbf{U}_i \end{bmatrix} \tag{11}$$

with

$$\mathbf{U}_i' = \begin{bmatrix} 0 & \sigma_2 s_2 & \sigma_1 s_1 & 2\sigma_1 s_1 \sigma_2 s_2 w \\ 0 & s_2 & \sigma_1 s_1 & 2\sigma_1 s_1 s_2 w \\ 0 & \sigma_2 s_2 & s_1 & 2 s_1 \sigma_2 s_2 w \\ 0 & s_2 & s_1 & 2 s_1 s_2 w \end{bmatrix} \tag{12a}$$

which may be rewritten as

$$\mathbf{U}_i' = \left(\frac{\partial \mathbf{U}_i}{\partial \ln s_1^i}\right) + \left(\frac{\partial \mathbf{U}_i}{\partial \ln s_2^i}\right) \tag{12b}$$

To develop a qualitative understanding of the dependence of Z_{sd} and f_{hd} on w, we examined a simple case: a cross-linked homopolymer with identical σ, s, and w for every residue. We begin with some limiting cases. In the limit that $n \to \infty$, Flory has shown that[37]

$$\lim_{n\to\infty} Z_{sd} \sim \lambda_{1d}{}^n \tag{13}$$

wherein λ_{1d} is the largest eigenvalue of the $\mathbf{U}_i$ defined in eq 5. Furthermore, applying eq 8 to this stapled, homopolymeric dimer of infinite molecular weight, we have

$$\lim_{n\to\infty} f_{hd} = \frac{1}{2n}\left(\frac{\partial \ln Z_{sd}}{\partial \ln s}\right) = \frac{1}{2}\left(\frac{\partial \ln \lambda_{1d}}{\partial \ln s}\right) \tag{14}$$

If, in addition, the chains are indifferent to one another, we may set w equal to unity in eq 5, whereupon

$$\lambda_{1d} = \lambda_{+s}{}^2 \tag{15}$$

with

$$\lambda_{\pm s} = \frac{1 + s \pm [(1-s)^2 + 4\sigma s]^{1/2}}{2} \tag{16}$$

wherein + and − are to be associated with the positive and negative radical, respectively, and λ_{+s} and λ_{-s} are the eigenvalues of the statistical weight matrix obtained from eq 3 for the isolated homopolymer. Using eq 14 gives

$$f_{hd} = f_{hs} = \frac{\lambda_{+s} - 1}{\lambda_{+s} - \lambda_{-s}} \tag{17}$$

As required, f_{hs}, the fraction of helix in an isolated single chain, is the same as in the noninteracting, stapled dimer.

In the opposite limit of strong chain–chain interaction, w becomes very large. We have found by numerical computation in that case that λ_{1d} converges to λ_{+w}, the largest eigenvalue of the statistical weight matrix $\mathbf{U}_2$:

$$\mathbf{U}_2 = \begin{array}{c|cc} & i & \\ i-1 & \text{c(1)c(2)} & \text{h(1)h(2)} \\ \hline \text{c(1)c(2)} & 1 & \sigma^2 s^2 w \\ \text{h(1)h(2)} & 1 & s^2 w \end{array} \tag{18}$$

$$\lambda_{\pm w} = \frac{1 + s^2 w \pm [(1 - s^2 w)^2 + 4 s^2 w \sigma^2]^{1/2}}{2} \tag{19}$$

The physical basis for this equality is that, in the limit $w \to \infty$, a h(1)h(2) conformation is far more stable than either a c(1)h(2) or h(1)c(2) conformation. Thus, once a single helical state occurs in one of the chains, the decrease in free energy accompanying helix formation on the adjacent chain overwhelms the increase in free energy associated with helix initiation. Of course, since λ_{+w} excludes certain conformations of weight greater than zero, it follows that $\lambda_{+w} \leq \lambda_{1d}$.

Comparison of numerically determined values of λ_{1d} with λ_{+w} defined in eq 19 provides the following observations. For values of s and σ in the ranges $1.14 \geq s \geq 0.6$ and $\sigma \leq 10^{-3}$, λ_{1d} and λ_{+w} agree within 2% or better when w is greater than 2.0. If $\sigma = 10^{-2}$ and $s = 0.6$, w must exceed 2.5 for λ_{+w} to be within 2% of λ_{1d}. The dependence on σ is to be expected. Larger values of σ imply that helix initiation is easier and that the average length of stretches of helix is smaller. Consequently, if helical states are disfavored ($s < 1$) but helix initiation is fairly easy, one would expect a significant contribution of conformations such as c(1)h(2) and h(1)c(2) to Z_{sd}. It is only in the limit of large w that hh conformations of weights s^2w or σ^2s^2w dominate. As σ decreases, once a single hh pair is initiated, h(1)h(2) conformations will tend to propagate over long stretches of the two chains and the relative contribution, at the same value of w, of h(1)h(2) conformations to Z_{sd} will be larger.

Approximating λ_{1d} by λ_{+w} thus corresponds to adoption of a model allowing only two possible states [c(1)c(2) or h(1)h(2)] for each residue pair. Thence, substitution of eq 19 into eq 14 gives the approximate helix fraction f_{h2} for this model in the limit of large n:

$$f_{h2} = (\lambda_{+w} - 1)/(\lambda_{+w} - \lambda_{-w}) \tag{20}$$

In Figure 1 we show the dependence on w of f_{hd} as determined numerically from eq 9 for a stapled dimer with $s = 0.9797$, $\sigma = 10^{-4}$, and $n = 284$. We also give the $n = \infty$ result, using eq 20 for the two-state-per-residue-pair model. As in the case of isolated chains, increasing σ broadens the transition. The statistical weight of an h(1)h(2) pair, s^2w, plays the same role in the two-chain, coiled coil as s does in an isolated chain. As is evident from

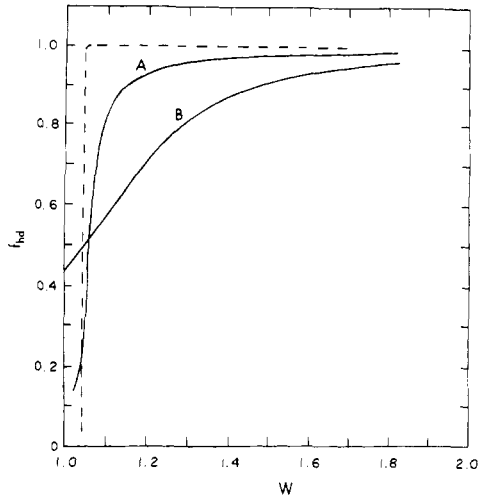

Figure 1. Fraction of helix in homopolymeric, cross-linked dimer as a function of w for the interacting individual residue model with $s = 0.9797$. Solid curve A: f_{hd} with $\sigma = 10^{-4}$, $n = 284$. Solid curve B: f_{hd} with $\sigma = 10^{-2}$, $n = 284$. Dashed curve: f_{h2} in the two-state-per-residue pair (cc–hh) approximation with $\sigma = 10^{-4}$, $n \rightarrow \infty$.

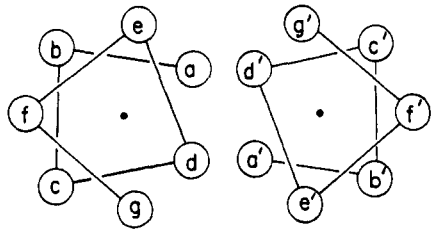

Figure 2. Schematic view of cross section of two-chain, coiled coil from carboxyl end. For details of side-chain packing see ref 13.

Figure 1, the midpoint of the transition for the large n limit of the two-state-per-residue-pair model is found where $s^2w = 1$, as required by eq 20. Furthermore, we point out the extreme sharpness of the transition predicted near $s^2w = 1$ at infinite n in the two-state-per-residue-pair model. Similar behavior is observed near $s = 1$ for isolated, infinitely long chains.[37] Note also that the midpoint for the more general model with $n = 284$ lies where s^2w is approximately unity, i.e., where $w = 1.04$. Finally, it is highly significant that a value of w of 1.15 or greater can, for the smaller σ ($=10^{-4}$), increase the helix content from its value of about 11% in the isolated chain to 90% in the two-chain, coiled coil. Otherwise stated, if dimerization raises the statistical weight of a helical residue in the h(1)h(2) pair by about 7%, essentially complete helix formation results. Thus, even this crude model, in which each residue on one chain can interact with its counterpart on the other, possesses some of the major properties that we seek to explain.

However, although the interacting individual residue model brings out some essential features of the helix–coil transition in two-chain, coiled coils, it cannot be taken literally. In fact (see Figure 2), it is physically impossible for every residue in a two-chain, coiled coil to be in contact (or directly interacting) with its neighbor on the opposite chain. In tropomyosin, essentially every first (a) and fourth (d) residue in each quasi-repeating heptet is hydrophobic. The geometry of the helix is such that the hydrophobic residues are in contact with corresponding ones on the opposite chain and experience a favorable interaction. Similarly, the possibility of salt bridge formation exists between every fifth (e) and every seventh (g) residue on one chain and every seventh (g′) and fifth (e′) residue, respectively, on the adjacent chain.[9,10,13] Consequently in order to have a net favorable interaction between chains, perhaps as many as seven consecutive helical residues on each strand must occur. Another plausible possibility is that the helical sequence of four residues (a–d) and the helical sequence of three residues (e–g) may be individually stabilized in a helical dimer. Thus, we must formulate a modified model of the helix–coil transition that explicitly incorporates these requirements.

The Coarse-Graining Approximation. The above requirements lead naturally to the coarse-graining approximation originally formulated by Crothers and Kallenbach[38] for heterogeneous block copolymers. The particular variant we employ has its origins in the treatment by Poland and Scheraga.[20,39] We next present a brief review of the method, describe an extension to isolated polypeptides of arbitrary sequence, and then extend the results to two-chain helices.

Consider an isolated polymer molecule conceptually divided into blocks containing m residues per block. Let H represent a helical block (i.e., m successive helical residues), C a random coil block, and [H]C and [C]H blocks preceded by an interface between residues of different conformations. For example, if $m = 4$

$$\begin{array}{ccccccccc} |\text{hhhh} & |\text{cccc} & |\text{cccc} & |\text{hhhh} & |\text{hhhh} \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ \text{H} & \text{[H]C} & \text{C} & \text{[C]H} & \text{H} \end{array} \tag{21}$$

It is central to the method that every residue in a helical (or coil) block *must* be h (or c), unless that block follows a block of different conformation; the allowed states for such [H]C and [C]H blocks are

[H]C	[C]H
cccc	ccch
hccc	cchh
hhcc	chhh
hhhc	hhhh

(22)

Note that the method neglects some, but not all, of those helical sequences within an [H]C or [C]H block that have length shorter than m. In the above, the conformation chhc, for example, is neglected. However, for the values of σ and s typical to two-chain, coiled coils, helical segments will tend to be appreciably longer than the values of m used, and we therefore expect that the neglect of these short helical sequences will result in a negligible error. It may be verified by inspection that the statistical weights of a specific-sequence, heteropolymeric block are given by

block	statistical weight
[C]C	1
[H]H	$\prod_{k=1}^{m} s_k \equiv \text{SM}$
[H]C	$1 + \sum_{j=1}^{m-1} \prod_{k=1}^{j} s_k \equiv \mathcal{S}$
[C]H	$\sum_{j=1}^{m} \sigma_j \prod_{k=j}^{m} s_k \equiv \tau$

(23)

s_k and σ_k refer to the values of s and σ for the kth residue in the block; and if $m = 1$, $\mathcal{S} = 1$. Observe that if all m units are identical, then

$$\begin{aligned} \text{SM} &= s^m \\ \mathcal{S} &= (s^m - 1)/(s - 1) \\ \tau &= \sigma s \mathcal{S} \end{aligned} \tag{24}$$

which are, as expected, the results obtained by Poland and

Scheraga for block copolymers.[20]

The statistical weight matrix $\mathbf{U}_c^i$ of the i block where $1 \le i \le n/m$ is in this coarse-graining approximation

$$\mathbf{U}_c^i = \begin{array}{c|cc} & C & H \\ \hline C & 1 & \tau^i \\ H & s^i & (SM)^i \end{array} \qquad (25)$$

and the partition function for an isolated, single chain (monomer) in the coarse-graining approximation is

$$Z_{mc} = \text{Row } (1,0) \prod_{i=1}^{n/m} \mathbf{U}_c^i \text{ Col } (1,1) \qquad (26)$$

Furthermore, the fraction of helix for a monomer in this approximation is

$$f_{hm,c} = \frac{1}{n} \sum_{l=1}^{n} \left(\frac{\partial \ln Z_{mc}}{\partial \ln s^l} \right) \qquad (27)$$

However, the form that is more conducive to numerical computation is

$$f_{hm,c} = \text{Row } (1,0,0,0) \prod_{i=1}^{n/m} \mathbf{Q}_i \text{ Col } (0,0,1,1) \qquad (28)$$

Here $\mathbf{Q}_i$ is a partitioned 4×4 matrix given by

$$\mathbf{Q}_i = \begin{bmatrix} \mathbf{U}_c^i & \mathbf{U}_{c,i}' \\ \mathbf{0} & \mathbf{U}_c^i \end{bmatrix} \qquad (29)$$

$\mathbf{0}$ represents a 2×2 null matrix, $\mathbf{U}_c^i$ is defined in eq 25, and

$$\mathbf{U}_{c,i}' = \begin{bmatrix} 0 & \tau^{i\prime} \\ s^{i\prime} & (SM)^{i\prime} \end{bmatrix} \qquad (30)$$

with

$$\tau^{i\prime} = \sum_{j=1}^{m} \left(\frac{\partial \tau^i}{\partial \ln s_j} \right) = \sum_{j=1}^{m} (m - j + 1)\sigma_j \prod_{k=j}^{m} s_k \qquad (31a)$$

$$(SM)^{i\prime} = \sum_{j=1}^{m} \left(\frac{\partial (SM)^i}{\partial \ln s_j} \right) = m(SM)^i \qquad (31b)$$

$$s^{i\prime} = \sum_{j=1}^{m} \left(\frac{\partial s^i}{\partial \ln s_j} \right) = \sum_{j=1}^{m-1} j \prod_{k=1}^{j} s_k \qquad (31c)$$

We note that these results can be trivially extended to the situation where $n = am + b$ and $b < m$ so polymers with a nonintegral number of blocks can easily be handled.

If all the residues are identical, the eigenvalues of eq 25 appropriate to a coarse-grained homopolymer are

$$\Lambda_\pm = \frac{1 + s^m \pm [(1 - s^m)^2 + 4\sigma s s^2]^{1/2}}{2} \qquad (32)$$

with + and − referring to the positive and negative radical, respectively. Now in the infinite, coarse-grained homopolymer limit

$$\lim_{n \to \infty} Z_{mc} = \Lambda_+^{n/m} \qquad (33)$$

By eq 27, it follows that

$$f_{hm,c} = \frac{\Lambda_+ - 1}{\Lambda_+ - \Lambda_-} + \frac{(\Lambda_+ - 1)(\Lambda_- - 1)}{(\Lambda_+)(\Lambda_+ - \Lambda_-)} \left\{ 1 - \frac{1}{m} - \frac{2}{m} H(m,s) \right\} \qquad (34)$$

and

$$H(m,s) = \frac{(m-1)s^{m+1} - ms^m + s}{s^{m+1} - s^m - s + 1} \qquad (35a)$$

and if $m = 1$

$$H(1,s) = 0 \qquad (35b)$$

and eq 34 reduces to the expression for f_{hs}, the fraction of helix in an isolated, non-coarse-grained chain as given by eq 17. We further observe that

$$H(m,1) = (m-1)/2 \qquad (35c)$$

thereby giving $f_{hm,c} = 0.5$ when $s = 1$.

At this juncture we are in a position to construct the statistical weight matrix $\mathbf{U}_{di}$, appropriate to the ith block in a cross-linked two-chain (dimer) coiled coil. Let w^i be the equilibrium constant for formation of a pair of helical blocks H(1)H(2) at block i in the interacting, cross-linked chains from the noninteracting, cross-linked chains. If $P_{Hd}(i)$ is the statistical weight of an H–H pair occurring in the dimer and $P_{Hm}(j,i)$ the statistical weight for the H conformation of block i in isolated chain j, then

$$w^i = P_{Hd}(i)/[P_{Hm}(1,i)P_{Hm}(2,i)] \qquad (36)$$

As a first approximation, w^i is assumed to be independent of whether the H(1)H(2) pair occurs at the beginning or middle of a stretch of "helical" blocks. In this formulation, all specific stretches of helical residues within paired helical blocks that contain at least one interfacial block are assigned the same w^i. That is, all individual conformational members of the pairings [C]H(1)H(2), H(1)[C]H(2), and [C]H(1)[C]H(2) are assigned the same w^i. For example, if $m = 3$, the ccc|hhh conformation of the [C]H block in one of the chains is enhanced by the same factor as is the ccc|cch (or any other) conformation of a [C]H block.

Strictly speaking, this cannot be correct. A sequence of helical residues of length less than m and located at the beginning of a stretch of helical blocks in both chains in the dimer experiences an enhanced helix probability relative to the isolated chains which will depend on the number and type of helical residues, but the corresponding equilibrium constant must be less than w^i. However, for values of σ typical of tropomyosin residues, helical conformations will propagate over distances appreciably greater than m. Hence, errors introduced by the somewhat incorrect treatment of helical end effects are likely to be small. Finally, we are motivated by the desire to keep the number of adjustable parameters to a minimum. If necessary, a more detailed treatment can be introduced at a later date.

A referee has pointed out quite correctly that there is a connection between coarse graining and neglect of the loop entropy for the dimer. By asserting that we can "lump" sequences of helical residues together, we are implicitly assuming, using the formulation given below, that the loop entropy contribution is unimportant (see assumption 3).

Using this formulation, then, $\mathbf{U}_{di}$ is

$$\mathbf{U}_{di} = \begin{array}{c|cccc} & C(1)C(2) & C(1)H(2) & H(1)C(2) & H(1)H(2) \\ \hline C(1)C(2) & 1 & \tau_2 & \tau_2 & \tau_1\tau_2 w \\ C(1)H(2) & s_2 & (SM)_2 & \tau_1 s_2 & \tau_1 (SM)_2 w \\ H(1)C(2) & s_1 & s_1\tau_2 & (SM)_1 & (SM)_1\tau_2 w \\ H(1)H(2) & s_1 s_2 & s_1 (SM)_2 & (SM)_1 s_2 & (SM)_1 (SM)_2 w \end{array} \qquad (37)$$

Inspection of $\mathbf{U}_{di}$ reveals that it can be simply rewritten as

$$\mathbf{U}_{di} = (\mathbf{U}_{c,1}^i \otimes \mathbf{U}_{c,2}^i)\mathbf{E}_w \qquad (38)$$

$\mathbf{E}_w$ is a diagonal matrix with unity for each diagonal element except for $\mathbf{E}_w(4,4) = w$. Consequently, the partition

function of the cross-linked, two-chain, coiled coil in the coarse-graining approximation is

$$Z_{sd,c} = \text{Row } (1,0,0,0)[\prod_{i=1}^{n/m} \mathbf{U}_{di}] \text{ Col } (1,1,1,1) \quad (39)$$

Furthermore, the fraction of helix in this stapled, coarse-grained dimer is

$$f_{hd,c} = \frac{1}{2nZ_{sd,c}} \boldsymbol{J}^*[\prod_{i=1}^{n/m} \mathbf{A}_i]\boldsymbol{J} \quad (40a)$$

$$f_{hd,c} = \frac{1}{2n} \sum_{i=1}^{n} \sum_{j=1}^{2} \left(\frac{\partial \ln Z_{sd,c}}{\partial \ln s_j^i} \right) \quad (40b)$$

$\boldsymbol{J}^*$ and $\boldsymbol{J}$ having been defined in eq 10 and $\mathbf{A}_i$ being an 8×8 supermatrix of the form

$$\mathbf{A}_i = \begin{bmatrix} \mathbf{U}_{di} & \mathbf{U}_{di}' \\ \hline 0 & \mathbf{U}_{di} \end{bmatrix} \quad (41)$$

with

$$\mathbf{U}_{di}' = \sum_{l=mi+1}^{m(i+1)} \sum_{k=1}^{2} \left(\frac{\partial \mathbf{U}_{di}}{\partial \ln s_k^l} \right) \quad (42)$$

The explicit values of the matrix elements of eq 42 are listed in Appendix A.

If chains 1 and 2 are each homopolymeric, but not necessarily identical,

$$\lim_{n\to\infty} Z_{sd,c} \sim \lambda_{1c}^{n/m} \quad (43)$$

wherein λ_{1c} is the largest eigenvalue of $\mathbf{U}_{di}$.

In a fashion analogous to the $m = 1$ case, i.e., using the coarse-grained version of the two-step-per-residue-pair model, we proceed to construct an approximate solution for λ_{+c} based on the statistical weight matrix

$$\mathbf{U}_{2,c} = \begin{array}{c|cc} & \multicolumn{2}{c}{i} \\ i-1 & \text{C(1)C(2)} & \text{H(1)H(2)} \\ \hline \text{C(1)C(2)} & 1 & \sigma^2 s^2 \mathcal{S}^2 w \\ \text{H(1)H(2)} & \mathcal{S}^2 & s^{2m} w \end{array} \quad (44)$$

The treatment based on eq 44 will be called the CC–HH model in what follows. In the formulation of eq 44, it is assumed for convenience that both chains are the same, but this is not required. The eigenvalues of eq 44 are

$$\Lambda_{\pm}^{w} = \frac{1 + s^{2m}w \pm [(1 - s^{2m}w)^2 + 4\sigma^2 s^2 \mathcal{S}^4 w]^{1/2}}{2} \quad (45)$$

Comparison of numerically obtained values of λ_{1c} with Λ_+^w leads to these qualitative conclusions: Setting $m = 7$, we find, as expected, that $\lim_{w\to\infty} \lambda_{1c} = \Lambda_+^w$. For example, if $\sigma = 10^{-4}$ and $s = 0.9797$, Λ_+^w is within 1% of λ_{1c} when w is 2.0 or greater. Furthermore, the qualitative observation presented for the $m = 1$ case carries over to the $m = 7$ case; increasing σ results in a greater value of w for which Λ_+^w is a good approximation to λ_{1c}. Since the contributions of C(1)H(2) and H(1)C(2) configurations to the partition function are omitted from Λ_+^w, $\Lambda_+^w \leq \lambda_{1c}$ for any given value of w.

The fraction of helix in the CC–HH model, with $Z_{sd,2c} = (\Lambda_+^w)^{n/m}$ and $H(m,s)$ as defined in eq 35a, follows immediately from eq 40b:

$$f_{h2,c} = \frac{\Lambda_+^w - 1}{(\Lambda_+^w - \Lambda_-^w)} + \frac{(\Lambda_+^w - 1)(\Lambda_-^w - 1)}{\Lambda_+^w(\Lambda_+^w - \Lambda_-^w)}\left\{1 - 1/m - \frac{2}{m}H(m,s)\right\} \quad (46)$$

Notice that eq 46 is of precisely the same form as eq 34, the fraction of helix in an isolated chain, with Λ_+ and Λ_- replaced by Λ_+^w and Λ_-^w, respectively.

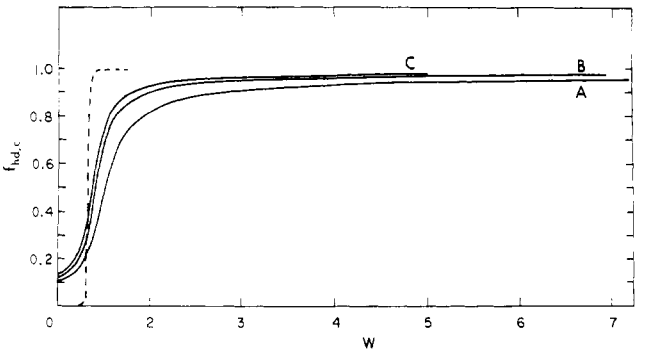

Figure 3. Fraction of helix in homopolymeric, cross-linked dimer as a function of w for the coarse-grained model with $s = 0.9797$ and $\sigma = 10^{-4}$. For all, $m = 7$. Solid curve A: $f_{hd,c}$ with $n = 287$. Solid curve B: $f_{hd,c}$ with $n = 574$. Solid curve C: $f_{hd,c}$ with $n = 861$. Dashed curve: $f_{h2,c}$ in the two-state-per-block-pair (CC–HH) approximation with $n \to \infty$.

Calculations on Coarse-Grained Homopolypeptides. Using the notion that the repeating heptet is the fundamental stabilized unit in the two-chain, coiled coil, we have performed calculations on homopolypeptides to gain qualitative insight into the dependence of $f_{hd,c}$ on w, σ, and n. As a bare minimum, the method must be able to reproduce the fraction of helix in isolated, noninteracting chains (f_{hs}). We shall examine the value of w at the midpoint of the helix-to-coil transition and the magnitude of w necessary for essentially complete helix formation.

In Figure 3 we plot vs. w $f_{hd,c}$ as determined numerically from eq 40a for $n = 287$, 574, and 861 and also $f_{h2,c}$ as determined for the approximate CC–HH model as given by eq 46 with $n = \infty$. In each case we use $m = 7$, $s = 0.9797$, and $\sigma = 10^{-4}$. Setting w equal to unity, we find $f_{hd,c} = 0.1075$, almost identical with the $m = 1$ value $f_{hs} = 0.1076$. In the limit of small w, $f_{h2,c}$ is grossly in error, as is expected. Nevertheless a study of $f_{h2,c}$ corresponding to $n = \infty$ is useful for estimating the midpoint of the helix–coil transition in the homopolymeric dimer. On the basis of eq 46, $f_{h2,c}$ should equal 0.5 when $s^{2m}w \simeq 1$ or $w = 1.33$, that is, when the statistical weight of the H(1)H(2) conformation equals unity. However, since there are other conformations that also contain helical states, e.g., [H(1)H(2)]C(1)C(2), the transition midpoint is shifted to slightly smaller values of w. In fact, $w = 1.31$ when $f_{h2,c} = 0.5$.

As one would predict, at a fixed large value of w, increasing n increases $f_{hd,c}$, and it may in the limit of infinite n indeed converge to $f_{h2,c}$ for $s^2w > 1$. At this point, it is worthwhile to investigate if the minimum value of w required to form completely helical dimers is a physically reasonable number. Recall that w is related to the standard free energy of transfer, ΔG_{HH}, of a H(1)H(2) pair from a cross-linked but noninteracting, two-chain, coiled coil to the H(1)H(2) conformation in the interacting dimer by

$$w = \exp\{-\Delta G_{HH}/kT\} \quad (47)$$

Then if $m = 7$ and $n = 287$, the value that corresponds nearly exactly to tropomyosin, Figure 3 shows $f_{hd,c} = 0.90$ when $w = 2.75$. Taking T to be 273 K, this corresponds to a $-\Delta G_{HH}$ of 548.80 cal/mol of heptet pairs or 39.20 cal/mol of residues. This is most certainly a reasonable number; it is, in fact, of the same order of magnitude as the difference in stability between the random coil and helical forms of discharged poly(glutamic acid).[40,41] Comparison with the $m = 1$ case is instructive. From Figure

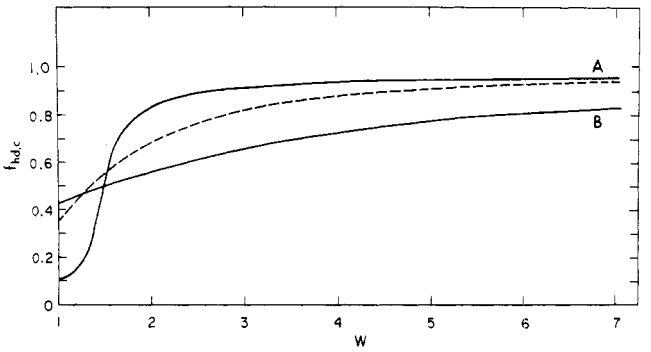

Figure 4. Fraction of helix in homopolymeric, cross-linked dimer as a function of w for the coarse-grained model with $s = 0.9797$. For all, $m = 7$. Solid curve A: $f_{hd,c}$ with $\sigma = 10^{-4}$, $n = 287$ (i.e., same as solid curve A on Figure 3). Solid curve B: $f_{hd,c}$ with $\sigma = 10^{-2}$, $n = 287$. Dashed curve: $f_{h2,c}$ in the two-state-per-block-pair (CC–HH) approximation with $\sigma = 10^{-2}$, $n \to \infty$.

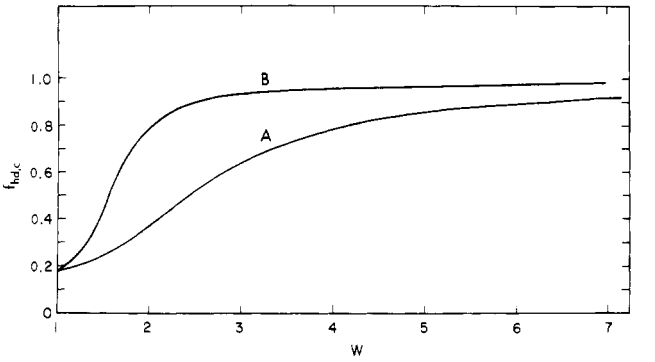

Figure 5. Fraction of helix in cross-linked rabbit α-tropomyosin as a function of w for the coarse-grained model. All values of s and of σ are appropriate to known primary structure. Curve A: $m = 7$. Curve B: $m = 4,3$.

1 we find $f_{hd} = 0.90$ at $w = 1.155$. Thus, $-\Delta G_{hh}$ equals 78.17 cal/mol of pairs of helical residues or 39.08 cal/mol of helical residues. Consequently, the per residue increment in helix stability is found to be essentially identical for the $m = 1$ and $m = 7$ cases. Of course, the $m = 1$ situation is nonphysical (see Figure 2), and the necessary stabilization per heptet pair is a factor of 7 larger than that required for a single helical pair.

The effect of a difference in transitional cooperativity may be investigated by changing the numerical value of σ. Figure 4 presents $f_{hd,c}$ as a function of w with $m = 7$ for a homopolymeric, cross-linked, two-chain, coiled coil with $n = 287$, $s = 0.9797$, but with $\sigma = 10^{-2}$; the curve with the previous value $\sigma = 10^{-4}$ is also shown for comparison. We find that $f_{hd,c}(w = 1)$ is 0.4302 compared with $f_{hs} = 0.4356$; certainly this is acceptable agreement. Relative to $\sigma = 10^{-4}$, the breadth of the transition is greater. Below 50% helix, the $f_{hd,c}$ is larger at a given w when $\sigma = 10^{-2}$ than when $\sigma = 10^{-4}$. However, above the transition midpoint the roles are reversed. The midpoint of the helix–coil transition as given by $f_{hd,c}$ and $f_{h2,c}$ is independent of σ, occurring at the same respective values of w for $\sigma = 10^{-2}$ and $\sigma = 10^{-4}$. In addition to the curves presented in Figure 4 for $n = 287$ and $\sigma = 10^{-2}$, we have calculated $f_{hd,c}$ for the same σ with $n = 574$. The curves are essentially indistinguishable. This is a different result from that seen in Figure 3, where at $\sigma = 10^{-4}$ appreciable dependence on n is found. This is reasonable, since the helix tends to propagate over a smaller number of segments when σ increases and we would expect the molecular weight dependence of $f_{hd,c}$ when $\sigma = 10^{-2}$ to be less than that when $\sigma = 10^{-4}$. If $m = 7$, the value of w required to form 80% helix is about 5.8 (90% helix is never reached for the range of w plotted). This corresponds to $-\Delta G_{HH} = 953.6$ cal/mol of heptet pairs or 68.11 cal/residue. If $m = 1$, $f_{hd} = 0.80$ when $w = 1.29$, and $-\Delta G_{HH} = 138.14$ cal/mol of helical pairs or 69.07 cal/mol of residues. Therefore, the per residue enhancement in helix stability with $\sigma = 10^{-2}$ is the same in the $m = 1$ and $m = 7$ models. However, comparison with the previous ($\sigma = 10^{-4}$) case shows that increasing σ increases the value of w required for complete helix formation.

Calculations on Rabbit α-Tropomyosin [Tm(Rα)]. We present here preliminary calculations for the fraction of helix in cross-linked $\alpha\alpha$ rabbit tropomyosin two-chain, coiled coils, each chain of which comprises 284 residues.[9,14,15] As already noted, the values of s and σ used are those tabulated by Mattice et al. for each residue type in the isolated chain.[18] The aims of the calculation are two: (1) we desire to assess the order of magnitude of w necessary for essentially complete helix formation; and (2) by examining the helix probability profile along the chain, we hope to locate helical regions of Tm(Rα) that are expected to be especially stable.

Two sets of calculations were performed. In the first set, m was taken to be 7 except for the first block at the N terminus, where it was taken to be 4. Thus, except for the first heptet, we did not distinguish between the first four members of the heptets, which tend to have hydrophobic residues at positions a and d (see Figure 2), and the latter three members, where salt bridge formation involving residues e and g with g′ and e′, respectively, is possible.

In the second set of calculations, we recognize this distinction by coarse graining the residues into alternating 4-residue and 3-residue blocks. While in time it may be possible to assign w a different value for the 4-residue and 3-residue blocks, i.e., to make a distinction between hydrophobic and electrostatic stabilizations or, even more exactly, to ascribe a priori site-dependent values of w, we make no attempts here to do so. Rather w is taken to be position independent.

In Figure 5 $f_{hd,c}$ as determined from eq 40a is plotted vs. w, with $m = 7$. For the noninteracting chains ($w = 1$), the $m = 7$ coarse-graining approximation is seen to give $f_{hm,c} = 0.183$, while in the exact 2×2 matrix result $f_{hs} = 0.189$, as is reported by Mattice et al.[18] This shows that this type of coarse graining does not invalidate the theory and generates confidence that it can be used to obtain the dependence of $f_{hd,c}$ on w. The coil-to-helix transition is seen to be broad and sigmoidal in character. The transition midpoint is at $w = 2.42$. Essentially complete helix formation is achieved, i.e., $f_{hd,c} > 0.90$, above $w = 6.25$, which at 0 °C corresponds to a $-\Delta G_{HH} = 994.2$ cal/mol of heptet pairs or 71.0 cal/mol of residues.

Figure 5 also shows $f_{hd,c}$ plotted against w in the 4,3 coarse-graining approximation. The curve is sigmoidal in shape and the transition is significantly sharper than the $m = 7$ coarse graining gives. At the midpoint of the coil-to-helix transition, $w = 1.55$. Furthermore, $f_{hd,c}$ exceeds 0.90 above the relatively modest w of 2.50. Translated into ΔG_{HH} at 0 °C this is $-\Delta G_{HH} = 497.1$ cal/mol of pairs of helical blocks. Taking the number of residues per block pair to be 7, this gives 71.0 cal/mol of residues. Therefore, on a stabilization per residue basis and assuming position-independent w, we cannot differentiate between the $m = 7$ and $m = 4,3$ coarse-graining models. Differences between the two models are likely to emerge when site-specific w^i are employed. This will form the basis of further work.

The helix probability profiles in the $m = 7$ and in the $m = 4,3$ coarse-graining approximations are obtained as

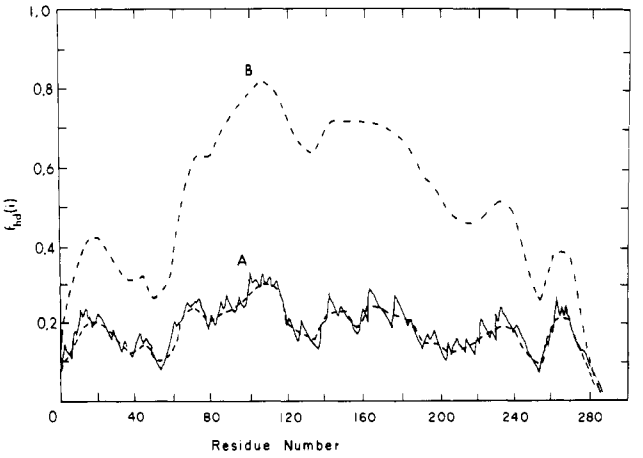

Figure 6. Helix probability profiles for rabbit α-tropomyosin. Solid curve A: "exact" (m = 1) results for noninteracting (w = 1) chains; f_{hd} = 0.189. Dashed curve A: m = 7 coarse-grained results for noninteracting (w = 1) chains; $f_{hd,c}$ = 0.183. Dashed curve B: m = 7 coarse-grained results for w = 2.5 interacting chains; $f_{hd,c}$ = 0.518.

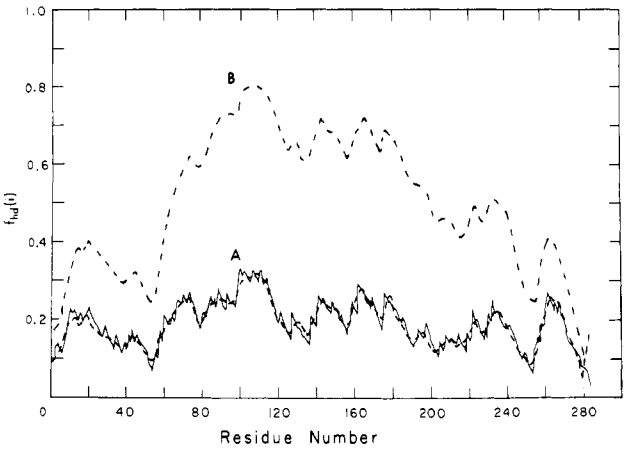

Figure 7. Helix probability profiles for rabbit α-tropomyosin. Solid curve A: "exact" (m = 1) results from noninteracting (w = 1) chains; f_{hd} = 0.189. Dashed curve A: m = 4,3 coarse-grained results for noninteracting (w = 1) chains; $f_{hd,c}$ = 0.188. Dashed curve B: m = 4,3 coarse-grained results for w = 1.55 interacting chains; $f_{hd,c}$ = 0.499.

follows. If the ith block contains m_i residues and there are N blocks, the probability that the ith block is helical is

$$f_{hd,c}(i) = \frac{1}{2m_iZ_{sd,c}} \text{Row } (1,0,0,0)[\prod_{k=1}^{i-1} \mathbf{U}_{dk}]\mathbf{U}_{di}'[\prod_{k=i+1}^{N} \mathbf{U}_{dk}] \text{ Col } (1,1,1,1) \quad (48)$$

where $Z_{sd,c}$ is defined in eq 39, $\mathbf{U}_{dk}$ is defined in eq 37, and $\mathbf{U}_{di}'$ may be found in eq A-1. Employing eq 48, we have determined $f_{hd,c}(i)$ at a mean helix content of near 50% as a function of residue location for the m = 7 coarse-graining approximation, shown in Figure 6 (as will be demonstrated in section III, $f_{hd,c}$ for both cross-linked and non-cross-linked chains is the same at this level of approximation), and for the 4,3 coarse-graining approximation, shown in Figure 7. In both figures we have also plotted the helix probability profile for noninteracting chains using the m = 1, 2 × 2 matrices and the coarse-graining approximation. The isolated-chain probability profiles are in agreement with the calculation of Mattice et al.[18] Furthermore, comparison of Figures 6 and 7 reveals that both the m = 7 and m = 4,3 coarse graining reproduce the general trends in the isolated-chain profile rather well. Moreover, because of smaller "mesh size" offered by the 4,3 treatment, it even displays the finer details. The helix probability profile for dimers at 50% mean helix content is seen to follow the general pattern evinced in the isolated chains; i.e., those regions having relatively large helix content in the isolated single strand also have a greater fraction of helix in the two-chain, coiled coil. This qualitative effect has been seen elsewhere in a different context by Mattice and Robinson;[42] if one acts to increase the stability of helical conformations by some enhanced interaction, the resulting helix probability profile displays regions of hills and valleys similar to those in the unperturbed system. Of course, in the perturbed system the overall helix formation is increased.

Examination of $f_{hd,c}(i)$ when $f_{hd,c}$ = 0.50 reveals a particularly stable region composed of a group of six heptets ranging from residues 82–123. A similar observation for residues 100–115 in the isolated molecule has been made previously.[18] In this particular stretch of dimer, the helix probability ranges from 64% to a maximum of 80%. Thus, we tentatively postulate that residues 82–123 form the extrastable region responsible for the high-temperature transition experimentally observed.[16b] These experiments were performed on non-cross-linked dimers, but, as will be seen in the next section, the theory predicts very little difference.

The lowest values in the probability profile, of course, are at the termini; local shallow minima are seen at residues 202 and 215; deeper minima occur near residues 54 and 254. The number of residues from the N terminus to the first deep minimum is 54 or 19% of the total number of amino acid residues. Similarly, the number of residues from the second deep minimum to the C terminus is 31 or 11% of the residues. The latter region is somewhat less stable than the former. At this point, pending further theoretical work and experimental studies on unhybridized Tm dimers, we hesitate to associate any particular portions of the chain with the low- and middle-temperature thermal transition regions observed in rabbit skeletal tropomyosin.[16b] Speculation about the general character of the middle-temperature thermal transition is deferred to section III.

III. Helix–Coil Transition and Dissociation Equilibrium in Un-Cross-Linked, Two-Chain, Coiled Coils

This section focuses on the relationship between the overall helix content of a solution, θ_h (which is what would ordinarily be experimentally observed), and the helix contents of the constituent, isolated-chain species and of the two-chain, coiled-coil species. Consider the dimerization of two identical chain molecules capable of forming a two-chain, α-helical, coiled coil:

$$A + A \rightleftharpoons A_2 \quad (49)$$

Let C_0 be the total formal concentration of protein in formula weights of chains per liter of solution, (A) the molar concentration of single-chain ("monomers") species, and (A_2) the molar concentration of two-chain, coiled-coil species ("dimers"). Then

$$C_0 = (A) + 2(A_2) \quad (50)$$

and the equilibrium constant for the reaction of eq 49 is

$$K = (A_2)/(A)^2 \quad (51)$$

Defining g_m as the fraction of chains that exist in the dissociated state, i.e., as single strands, $g_m = (A)/C_0$, we find therefore

$$K = (1 - g_m)/2g_m^2C_0 \quad (52)$$

Solving eq 52 for g_m and recognizing that g_m is positive, we find

$$g_m = \frac{-1 + [1 + 8KC_0]^{1/2}}{4KC_0} \tag{53}$$

Observe that $g_m \to 1$ as $C_0 \to 0$ and $g_m \to 0$ as $KC_0 \to \infty$, as required by the law of mass action.

Circular dichroism provides a means of measuring the total helix content of the system, Θ_h. Now, obviously

$$\Theta_h = g_m f_{hm} + (1 - g_m) f_{hd}^u \tag{54}$$

or

$$\Theta_h = f_{hd}^u + (f_{hm} - f_{hd}^u) g_m \tag{54a}$$

where f_{hd}^u is the helix fraction in the non-cross-linked helical dimers.

Ideally, one would like to obtain g_m, f_{hm}, and f_{hd}^u, as well as Θ_h, from experiment and then compare them with values obtained from theory. Several kinds of relevant experiments immediately come to mind. The measurement of g_m may be done by using, say, light scattering to determine the weight-average molecular weight. In this case $M_w = M_m(2 - g_m)$ and therefore g_m is immediately obtainable, knowing the single-chain molecular weight, M_m. However, eq 54 involves f_{hm} and f_{hd}^u as well, so more experiments are required. Studies of the concentration and temperature dependence of Θ_h must be undertaken. However, before proceeding to the analysis of the information provided by such studies, it will prove worthwhile to formulate the microscopic expressions for the important quantities K, f_{hm}, and f_{hd}^u.

Molecular Model of Un-Cross-Linked, Two-Chain, Coiled Coils. Among the various quantities involved in eq 54, we have already calculated f_{hm} in section II (see eq 27). If Z_d represents the partition function of an un-cross-linked, two-chain, coiled coil, then

$$f_{hd}^u = \sum_{i=1}^{n} \sum_{j=1}^{2} \frac{1}{2n} \left(\frac{\partial \ln Z_d}{\partial \ln s_j^i} \right) \tag{55}$$

Moreover, it is a standard result of statistical mechanics[43] that

$$K = VZ_d / Z_m^2 \tag{56}$$

wherein V is the volume of the system and Z_m is the partition function of a single chain which may be obtained from eq 26. Consequently, the problem reduces to the calculation of Z_d. As will be seen, our approach is very much in the spirit of Zimm's analogous treatment of helix-coil transitions in DNA.[29]

Z_d, the total configurational partition function of the unstapled dimer, differs from the internal partition function of the stapled dimer, $Z_{sd,c}$ (see eq 39), in that explicit account must be taken of the reduction in the configurational degrees of freedom brought about by dimerization. Furthermore, on physical grounds we expect the completely randomly coiled conformational state to be excluded from Z_d. Thus, we expect

$$Z_d = \frac{u}{V}(Z_{sd,c} - 1) \tag{57}$$

in which u represents the volume of configurational space accessible to the center of mass of one m-residue block when the center of mass of the other m-residue block is held fixed multiplied by the ratio of the volume of configurational space spanned by the internal degrees of freedom in the block in the dimer to that in the monomers.

Now u/V is independent of s_j^i. Thus, it follows from eq 55 and 57 that

$$f_{hd}^u = \frac{1}{2n} \sum_{i=1}^{n} \sum_{j=1}^{2} \left(\frac{\partial \ln (Z_{sd,c} - 1)}{\partial \ln s_j^i} \right) \tag{58}$$

However, $Z_{sd,c}$ is typically much greater than unity; thus

$$f_{hd} = \frac{1}{2n} \sum_{i=1}^{n} \sum_{j=1}^{2} \left(\frac{\partial \ln Z_{sd,c}}{\partial \ln s_j^i} \right) \tag{59a}$$

or

$$f_{hd}^u = f_{hd,c} \tag{59b}$$

where $f_{hd,c}$ is given in eq 40b. That is, the fraction of helix in the stapled and nonstapled dimers is the same. Finally, substituting eq 57 into eq 56 gives

$$K = \frac{u(Z_{sd,c} - 1)}{Z_{mc}^2} \simeq \frac{uZ_{sd,c}}{Z_{mc}^2} \tag{60}$$

Thus, we have derived expressions for K, f_{hm}, and f_{hd}^u.

Concentration Dependence of Θ_h. Let us rewrite eq 53 as

$$2KC_0 g_m^2 + g_m - 1 = 0 \tag{61}$$

Implicitly differentiating eq 61 with respect to C_0 yields

$$\frac{dg_m}{d \ln C_0} = \frac{g_m(g_m - 1)}{2 - g_m} \tag{62}$$

Now f_{hm} and f_{hd}^u have been assumed to be independent of concentration; therefore it follows from eq 54 that

$$\frac{d \ln \Theta_h}{d \ln C_0} = \left[1 - \frac{f_{hd}^u}{\Theta_h} \right] \frac{g_m - 1}{2 - g_m} \tag{63}$$

Therefore in order to obtain f_{hm} and f_{hd}^u from a dilution experiment, we further require an independent measurement of g_m by light scattering. Of course, we have implicitly assumed that C_0 is sufficiently high so that light scattering measurements can be done and yet sufficiently low so that appreciable numbers of dissociated chains exist. Whether these conditions are experimentally attainable remains to be seen.

In order to obtain an estimate for the range of the helix-coil transition where dilution studies would yield any information, we have performed the following calculation based on eq 60. As a first approximation, u may be obtained by calculating the excluded volume of two disks in contact. Each disk approximately represents a single turn of the α helix of radius $d/2$ and height h. Hence, $u = \pi d^2 h$. For a single turn of an α helix,[44] $h = 5.4$ Å and $d \simeq 4.6$ Å so that $u = 359$ Å^3. Substituting this value of u into eq 60 and assuming that a lower limit to light scattering measurement is at a concentration of about $C_0 = 10^{-5}$ M, we can calculate various values of g_m for a given $f_{hd,c}$. These are shown in Table I. It is clear that the treatment predicts that below $f_{hd,c}$ of 65%, only single chains are present and above $f_{hd,c}$ of 80% only dimers are found. One should therefore do dilution studies in the temperature regime where the helix content of the dimer is from 65% to 80%. This corresponds to the region where Θ_h is from 20% to 75%. We shall return to further ramifications of Table I in our discussion of the temperature dependence of Θ_h.

Measurement of Θ_h and g_m as functions of C_0 at a given temperature and use of eq 52 would provide us with K. Thence use of eq 54 with eq 63 or a plot of Θ_h vs. g_m and

Table I
Predicted Relationship of Helical Fraction in Dimer and Degree of Dissociation

f_{hd}^u	KC_0[a]	g_m	Θ_h
0.50	3.12×10^{-4}	0.999	0.19
0.66	0.029	0.948	0.21
0.70	0.706	0.846	0.27
0.75	1.68	0.420	0.51
0.80	2.96×10^{2}	0.120	0.73
0.85	1.01×10^{4}	7.00×10^{-3}	0.85
0.90	2.16×10^{5}	4.8×10^{-5}	0.90

[a] Calculated assuming $C_0 = 10^{-5}$ M and in the 4,3 coarse-graining approximation.

use of eq 54a alone would yield f_{hm} and f_{hd}^u. Then, the value of w necessary to produce the observed f_{hd}^u can be extracted from Figure 5. Using the w obtained in this way, we can calculate Z_{sd} and consequently obtain u from eq 60.

Temperature Dependence of K and Θ_h. Unstapled rabbit skeletal tropomyosin is believed to have three thermal transition regions.[16b] The calculations presented in Table I suggest the following tentative interpretation of the experiment: The low-temperature transition region arises from partial melting within the nondissociated, two-chain, coiled coil. In the context of a site-independent w, the current treatment cannot account for this transition given the s and σ values appropriate to the primary structure of tropomyosin. Furthermore, the intermediate, and major, thermal transition most likely is accompanied by chain dissociation into single-stranded molecules that are less helical than the dimers with which they coexist. Finally, the high-temperature thermal transition is due to the denaturation of the residual helix content (~19%) of the isolated single chains. Consequently, to verify the theory, measurements of g_m as well as of Θ_h as a function of temperature are absolutely necessary.

We now examine what further information is provided by measuring Θ_h as a function of temperature and at fixed concentration.

Taking the derivative of Θ_h with respect to temperature reveals that

$$\frac{d\Theta_h}{dT} = \frac{g_m(g_m - 1)}{2 - g_m}(f_{hm} - f_{hd}^u)\frac{d \ln K}{dT} + (1 - g_m)\frac{df_{hd}^u}{dT} + g_m\frac{df_{hm}}{dT} \quad (64)$$

Now it is a standard thermodynamic result that

$$\frac{d \ln K}{dT} = \frac{\Delta H^\infty}{RT^2} \quad (65)$$

wherein ΔH^∞ is the standard enthalpy change of the association reaction (eq 49) using the infinitely dilute reference state and molarity concentration. However, from eq 60

$$\frac{d \ln K}{dT} = \frac{d \ln u}{dT} + \frac{d \ln (Z_{sd,c} - 1)}{dT} - \frac{d \ln Z_{mc}^2}{dT} \quad (66)$$

Presumably the difference of the latter two terms on the right-hand side of eq 66 reflects the temperature dependence of w. Flory[37] has shown that for a partition function Z

$$dZ/dT = \boldsymbol{J}^* \prod_{i=1}^{n} \hat{\mathbf{U}}_{Ti} \boldsymbol{J} \quad (67)$$

where $\boldsymbol{J}^*$ and $\boldsymbol{J}$ are defined in eq 10 and

$$\hat{\mathbf{U}}_{Ti} = \begin{bmatrix} \mathbf{U}_{di} & \mathbf{U}_{T,di}' \\ \mathbf{0} & \mathbf{U}_{di} \end{bmatrix} \quad (68)$$

is a partitioned 8 × 8 matrix with

$$\mathbf{U}_{T,di}' = d\mathbf{U}_{di}/dT \quad (69)$$

and $\mathbf{U}_{di}$ is given in eq 37. Presumably light scattering could provide $d \ln K/dT$. Equation 66 could perhaps be employed to elucidate the temperature dependence of u and w but, it might be difficult to sort them out. On the other hand, if a model of $w(T)$ and $u(T)$ is formulated, eq 66 provides the link between theory and experiment.

Flory[37] has also derived expressions for the temperature dependence of the fraction of helix:

$$\frac{df_{hd}(w)}{dT} = \frac{1}{2nZ_{sd}(w)}\boldsymbol{J}_{16}^*\left[\prod_{k=1}^{n} \bar{\bar{\mathbf{A}}}_{k,T}\right]\boldsymbol{J}_{16} - \frac{f_{hd}(w)}{2n}\frac{d \ln Z_{sd}}{dT} \quad (70)$$

wherein $\boldsymbol{J}_{16}^*$ is a row vector containing unity at the first position followed by 15 zeros and $\boldsymbol{J}_{16}$ is a column vector having 8 zeros, followed by 8 ones. Setting $w = 1$ gives f_{hm}, while setting $w \neq 1$ provides $f_{hd,c}$. Moreover $\bar{\bar{\mathbf{A}}}_{k,T}$ is a 16 × 16 matrix given by

$$\bar{\bar{\mathbf{A}}}_{k,T} = \begin{bmatrix} \mathbf{A}_k & \mathbf{A}_{k,T}' \\ \mathbf{0} & \mathbf{A}_k \end{bmatrix} \quad (71)$$

with $\mathbf{A}_k$ defined in eq 41 and

$$\mathbf{A}_{k,T}' = d\mathbf{A}_k/dT \quad (72)$$

Note that eq 70 is also applicable to cross-linked, two-chain, coiled coils. In the case of such stapled dimers, eq 70 provides the relevant equation for extracting the temperature dependence of w by measuring the temperature dependence of the circular dichroism alone. If non-cross-linked dimers are considered, it may be possible to estimate $df_{hd,c}/dT$ and df_{hm}/dT by making measurements in the vicinity of $g_m = 0$ and $g_m = 1$.

IV. Discussion

We have presented above a treatment of the helix–coil transition in two-chain, coiled coils that takes account of the enhanced helix stability arising from two-chain association. We began with a heuristic model of a cross-linked, two-chain, coiled coil that assumes that every nonbonded pair of helical segments in the dimer is stabilized, relative to the helical segments in the noninteracting chains, by the same amount ($-RT \ln w$). Although the model is physically unrealistic, the study of its qualitative features is nonetheless instructive. Calculation of the helix content in cross-linked (stapled) homopolymers, f_{hd}, reveals that f_{hd} depends on s^2w (the statistical weight of a helix–helix pair) and on σ in a manner qualitatively similar to the well-known dependence of f_{hm}, the helix content of an isolated chain, on σ and s.

We then employed a coarse-graining approximation that recognizes that one or two consecutive helical turns are physically required to effect the enhanced helix stability in two-chain, coiled coils. In other words, we incorporated into the model the importance of the quasi-repeating heptet in accomplishing the increased stabilization of helical conformations in the dimer. The helix content of

a stapled homopolymer, $f_{hd,c}$, was calculated in the 7 helical segments per block ($m = 7$) and in the alternating 4 helical segments and 3 helical segments per block ($m = 4,3$) approximation. If there are m residues per block, $f_{hd,c}$ displays the same dependence qualitatively on $s^{2m}w$ (the statistical weight of a pair of side-by-side helical blocks) and on σ as f_{hm} does on s and σ.

Using a site-independent w value and site-dependent s and σ values, we then calculated $f_{hd,c}$ of a singly cross-linked, rabbit α-tropomyosin dimer. The analogous theory for single chains of Tm(Rα) is known to predict a rather low (19%) helix content.[18] Our results on the dimer show that the enhanced stability required to induce essentially complete helix formation in the two-chain coiled coil is about 71 cal/mol of residues. Use of a site-independent w has several consequences. For one, it forces degeneracy of the stabilization per residue for the $m = 7$ and the $m = 4,3$ coarse-graining models. For another, it requires that the helix probability profile at 50% helix content mirror that of the isolated chains; while dimerization acts to increase the absolute stability of the helical regions, the relative stabilities of the various regions of the molecule remain the same as in the isolated chains.

A study of the helix probability profile indicates a region of high stability extending from residues 82–123. This may be the location of the extrastable region believed responsible for the high-temperature thermal transition that has been experimentally observed in unstapled chains.[16b] This residue 82–123 region comprises 14% of the residues, about the number in the extrastable region (13%)[18] and about the number predicted for the helix content at room temperature of isolated chains (14–19%).[18]

Next, we formulated the theory for un-cross-linked, two-chain, coiled coils, which are capable of undergoing dissociation into single, partially helical strands. Expressions were derived for the degree of chain dissociation (g_m) as well as for helix content for the entire equilibrium system (Θ_h), for constituent isolated chain species (f_{hm}), and for constituent dimer species (f_{hd}^u). At the level of approximation presented, the helix content is identical in both stapled and unstapled molecules. Then, ways of extracting the various model parameters and helix content from dilution and thermal denaturation experiments were developed. By examining the dependence of the overall helix content on w, the molecular processes associated with the middle- (major) and high-temperature thermal transitions likely to be present in α-tropomyosin were tentatively identified: we suggest that the middle-temperature thermal denaturation involves simultaneous partial denaturation in and the complete dissociation of the two-chain, coiled coil to isolated single chains. The high-temperature transition region involves the melting of the extrastable region in single-chain species.

Although this is not the place for a thorough discussion, a comment is in order about our use of helix probability profiles. As noted above, use of a site-independent w fixes the relative stability of local regions of the polypeptide chain essentially at the value for isolated chains. This approach thus ascribes the relative stability of local region of chain to the particular s and σ values that characterize its constituent amino acid residues. Local stability thereby becomes a result of that assembly of "short-range" interactions generally thought to contribute to s and σ. In most previous discussions of local helix stabilities for tropomyosin, a completely different way of obtaining helix probability profiles has been used in which helix probabilities for the particular amino acids are obtained by some variant of the Chou–Fasman approach.[45–48] In that approach, a statistical procedure is employed to obtain the likelihood that an individual residue is helical from tabulated crystal structures of many globular proteins. Thus, the resulting values for each amino acid represent some sort of global average over all physical environments in which that residue may be found in those globular proteins whose crystal structures have been included in the data base. Furthermore, at least in the case of tropomyosin, the resulting profiles have been subjected to certain local averaging procedures to smooth fluctuations.[47]

Clearly, then, profiles obtained through use of the s and σ values and through the use of Chou–Fasman parameters need not agree, since they employ different means of assessing the magnitude of the significant interactions. If, for example, the site dependence of w turns out to be crucial, then the profiles reported here will not faithfully mimic the situation in the real molecule. On the other hand, if the average physical environment in which an amino acid residue finds itself in globular proteins differs drastically from its environment in the α-helical, coiled coil, then Chou–Fasman parameters will not provide a meaningful profile. Indeed, part of the impetus for the present investigation is our feeling that the stabilizing interactions in the coiled coil may be sufficiently different from those in globular proteins to vitiate the Chou–Fasman approach. Furthermore, the Chou–Fasman approach, being "nonphysical", allows estimates of helix content only for native proteins; it cannot lead to calculation of, say, the temperature dependence of helix content. It is important therefore to see how far the more physical approach can go in explaining such properties in coiled coils.

At this juncture, much further theoretical and experimental work remains. With regard to the theory, we are currently extending the treatment to take explicit account of the presence of both charged and uncharged residues; this is required to take account of the pH dependence of the thermal denaturation profile. Additional avenues of development include a modification of the treatment of Mattice et al.[18] to better account for end effects and an extension of Zimm's treatment of DNA[29] to incorporate the contribution of loop entropy into the theory.

Additional experimental investigations are needed. The temperature dependence of the helix content in cross-linked and un-cross-linked $\alpha\alpha$ and $\beta\beta$ tropomyosin should be measured over the entire accessible temperature range. Light scattering experiments must be employed to measure g_m. Comparison of the relative stabilities of the various tropomyosins may provide an estimate of the site specificity of w. In general, what is required is construction of a table of values of w appropriate to each constellation of particular, adjacent, interacting amino acid residues just as one now exists for σ and s. That w perhaps is site specific is supported by chain hybridization experiments on paramyosin;[49] similar studies on hybridization of $\alpha\alpha$ and $\beta\beta$ tropomyosin will perhaps be informative, as will further studies on the stability of excised segments of tropomyosin[50] and of synthetic polypeptides whose sequences are related to the tropomyosin heptet structure.[51] The latter seem to indicate that leucines in the core (a or d) positions provide a greater stabilization than other hydrophobes.

Acknowledgment. This study was supported in part by Grant No. GM-20064 from the Division of General Medical Sciences, United States Public Health Service. Partial support was also provided by the donors of the Petroleum Research Fund, administered by the American Chemical Society. We thank Professor Wayne Mattice for informative discussions.

Appendix A

Explicit Evaluation of U_{di}'

$U_{di}' =$

$$\begin{bmatrix} 0 & \tau_2' & \tau_1' & w(\tau_1'\tau_2+\tau_2'\tau_1) \\ \mathcal{S}_2' & m(SM)_2 & \tau_1'\mathcal{S}_2+\tau_1\mathcal{S}_2' & w(SM)_2(m\tau_1+\tau_2') \\ \mathcal{S}_1' & \mathcal{S}_1'\tau_2+\mathcal{S}_1\tau_2' & m(SM)_1 & w(SM)_1(m\tau_2+\tau_2') \\ \mathcal{S}_1'\mathcal{S}_2+\mathcal{S}_1\mathcal{S}_2' & (SM)_2(\mathcal{S}_1'+m\mathcal{S}_1) & (SM)_1(\mathcal{S}_2'+m\mathcal{S}_2) & 2m(SM)_1(SM)_2 \end{bmatrix} \quad \text{(A-1)}$$

The index i has been dropped for convenience. The subscripts 1 and 2 label chains 1 and 2, respectively. $\mathcal{S}_j'$ and τ_j' are defined in eq 31c and 31a, respectively.

If both chains 1 and 2 contain blocks composed at m identical residues (the residues need not be the same in both chains), then

$$\mathcal{S}' = s\frac{d\mathcal{S}}{ds} = G(m,s) \quad \text{(A-2)}$$

$$\tau' = \sigma s\mathcal{S} + \sigma sG(m,s) \quad \text{(A-3)}$$

$$SM = s^m \quad \text{(A-4)}$$

where

$$G(m,s) = \frac{(m-1)s^{m+1} - ms^m + s}{(s-1)^2} \quad \text{(A-5)}$$

Note that

$$G(1,s) = 0 \quad \text{(A-6)}$$

$$G(m,1) = m(m-1)/2 \quad \text{(A-7)}$$

References and Notes

(1) Szent-Györgyi, A. G.; Cohen, C.; Kendrick-Jones, J. *J. Mol. Biol.* **1971**, *56*, 239–58.
(2) Cohen, C.; Holmes, K. *J. Mol. Biol.* **1963**, *6*, 423–32.
(3) Lowey, S.; Kucera, J.; Holtzer, A. *J. Mol. Biol.* **1963**, *7*, 234–44.
(4) Cowgill, R. *Biochemistry* **1974**, *13*, 2467–74.
(5) Cohen, C.; Szent-Györgyi, A. G. *J. Am. Chem. Soc.* **1957**, *79*, 248.
(6) Holtzer, A.; Clark, R.; Lowey, S. *Biochemistry* **1965**, *4*, 2401–11.
(7) Woods, E. *Biochemistry* **1969**, *8*, 4336–44.
(8) Caspar, D.; Cohen, C.; Longley, W. *J. Mol. Biol.* **1969**, *41*, 87–107.
(9) Hodges, R.; Sodek, J.; Smillie, L.; Jurasek, L. *Cold Spring Harbor Symp. Quant. Biol.* **1972**, *37*, 299–310.
(10) Johnson, P.; Smillie, L. *Biochem. Biophys. Res. Commun.* **1975**, *64*, 1316–22.
(11) Lehrer, S. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 3377–81.
(12) Stewart, M. *FEBS Lett.* **1975**, *53*, 5–7.
(13) McLachlan, A.; Stewart, M. *J. Mol. Biol.* **1975**, *98*, 293–304.
(14) Stone, D.; Smillie, L. *J. Biol. Chem.* **1978**, *253*, 1137–48.
(15) Mak, A.; Lewis, W.; Smillie, L. *FEBS Lett.* **1979**, *105*, 232–4.
(16) (a) Woods, E. *Aust. J. Biol. Sci.* **1976**, *29*, 405–18. (b) Crimmins, D.; Isom, L.; Holtzer, A. *Comp. Biochem. Physiol.* **1981**, *69B*, 35–46.
(17) Wu, C.-S.; Ikeda, K.; Yang, J.-T. *Biochemistry* **1981**, *20*, 566–70.
(18) Mattice, W.; Srinivasan, G.; Santiago, G. *Macromolecules* **1980**, *13*, 1254–60.
(19) (a) Poland, D. C.; Scheraga, H. A. *Biopolymers* **1965**, *3*, 305. (b) Poland, D. C.; Scheraga, H. A. *Ibid.* **1965**, *3*, 335.
(20) Poland, D.; Scheraga, H. "Theory of Helix-Coil Transitions in Biopolymers"; Academic Press: New York, 1970.
(21) Zimm, B.; Bragg, J. *J. Chem. Phys.* **1959**, *31*, 526–35.
(22) Nagai, K. *J. Phys. Soc. Jpn.* **1960**, *15*, 407.
(23) Lifson, S.; Roig, A. *J. Chem. Phys.* **1961**, *34*, 1963–74.
(24) Gibbs, J.; DiMarzio, E. *J. Chem. Phys.* **1958**, *28*, 1247–8. *Ibid.* **1959**, *30*, 271–82.
(25) Lehman, G.; McTague, J. *J. Chem. Phys.* **1968**, *49*, 3170–9.
(26) Lifson, S.; Zimm, B. *Biopolymers* **1963**, *1*, 15–23.
(27) Lifson, S. *Biopolymers* **1963**, *1*, 25–32.
(28) Lifson, S.; Allegra, G. *Biopolymers* **1964**, *2*, 65–8.
(29) Zimm, B. *J. Chem. Phys.* **1960**, *33*, 1349–56.
(30) Ananthanarayanan, V.; Andreatta, R.; Poland, D.; Scheraga, H. *Macromolecules* **1971**, *4*, 417–24.
(31) Alter, J.; Taylor, G.; Scheraga, H. *Macromolecules* **1972**, *5*, 739–46.
(32) Matheson, R.; Nemenoff, R.; Cadinaux, F.; Scheraga, H. *Biopolymers* **1977**, *16*, 1567–85.
(33) Hecht, M.; Zweiful, B.; Scheraga, H. *Macromolecules* **1978**, *11*, 545–51.
(34) Lehrer, S. *J. Mol. Biol.* **1978**, *118*, 209–26.
(35) Edwards, B.; Sykes, B. *Biochemistry* **1980**, *19*, 2577–83.
(36) Byron, F.; Fuller, R. "Mathematics of Classical and Quantum Physics"; Addison-Wesley: Reading, MA, 1969; Vol. I.
(37) Flory, P. "Statistical Mechanics of Chain Molecules"; Wiley: New York, 1969.
(38) Crothers, D.; Kallenbach, N. *J. Chem. Phys.* **1966**, *45*, 917–27.
(39) Poland, D.; Scheraga, H. *Physiol. Chem. Phys.* **1969**, *1*, 389.
(40) Nagasawa, M.; Holtzer, A. *J. Am. Chem. Soc.* **1964**, *86*, 538–43.
(41) Olander, D.; Holtzer, A. *J. Am. Chem. Soc.* **1968**, *90*, 4549–60.
(42) Mattice, W.; Robinson, R. *Biopolymers* **1981**, *20*, 1421–34.
(43) Mayer, J.; Mayer, M. "Statistical Mechanics"; Wiley: New York, 1940; Chapter 9.
(44) Schultz, G.; Schirmer, R. "Principles of Protein Structure"; Springer-Verlag: New York, 1979.
(45) Chou, P.; Fasman, G. *Biochemistry* **1974**, *13*, 222–45.
(46) Levitt, M. *Biochemistry* **1978**, *17*, 4277–85.
(47) Parry, D. *J. Mol. Biol.* **1975**, *98*, 519–35.
(48) Smillie, L.; Pato, M.; Pearlstone, J.; Mak, A. *J. Mol. Biol.* **1980**, *136*, 199–202.
(49) Crimmins, D.; Holtzer, A. *Biopolymers* **1981**, *20*, 925–50.
(50) Pato, M.; Mak, A.; Smillie, L. *J. Biol. Chem.* **1981**, *256*, 593–601.
(51) Hodges, R.; Saund, A.; Chong, P.; St.-Pierre, S.; Reid, R. *J. Biol. Chem.* **1981**, *256*, 1214–24.

Shape and Surface Features of Globular Proteins

M. Prabhakaran and P. K. Ponnuswamy*

Department of Physics, Autonomous Postgraduate Centre, University of Madras, Tiruchirapalli 620 020, Tamil Nadu, India.

Received May 20, 1981

ABSTRACT: By constructing best-fitting ellipsoidal shapes and computing surface areas from volume packing considerations, we investigate shape anisotropy in globular proteins. The spatial positions from the centroids of the fitted ellipsoids and the water contact areas of the amino acid residues in a set of protein molecules are used to obtain information about the demarcation between buried and exposed spatial levels of the residues, the local surface roughness of the protein, and the intrinsic residue contribution to the molecular shape. Crystal data form the basis for this study.

Introduction

Size and shape are two exceedingly important characteristics of globular proteins in assessing the specificity of their interactions with other molecular entities and consequently, their biological functions. Studies have been made to investigate these subtle features from various points of view. The size and shape characters were studied by analyzing amino acid composition,[1] surface roughness,[2,3] solvent exclusion,[4] surface symmetry,[5] and topological considerations.[6] Spherical representation was found to be

0024-9297/82/2215-0314$01.25/0