(12) Djerassi, C.; Riniker, R.; Riniker, B. *J. Am. Chem. Soc.* **1956**, *78*, 6377.
(13) Cohn, E. J.; Edsall, J. T. "Proteins, Amino Acids, and Peptides"; Reinhold Publishing Co.: New York, 1943; Chapter 16, p 371.
(14) von Dreele, P. H.; Poland, D.; Scheraga, H. A. *Macromolecules* **1971**, *4*, 396.
(15) Types I and III β-turn conformations in *dipeptides* exhibit similar CD spectra.[16] Since the backbone dihedral angles of such β-turns are similar to those of the α-helix,[17] there is no inconsistency in assigning the polymer CD spectra of Figure 2 to a mixture of α-helix and coil.
(16) Bandekar, J.; Evans, D. J.; Krimm, S.; Leach, S. J.; Lee, S.; McQuie, J. R.; Minasian, E.; Némethy, G.; Pottle, M. S.; Scheraga, H. A.; Stimson, E. R.; Woody, R. W. *Int. J. Pept. Protein Res.* **1982**, *19*, 187.
(17) Zimmerman, S. S.; Scheraga, H. A. *Biopolymers* **1977**, *16*, 811.
(18) Lehman, G. W.; McTague, J. P. *J. Chem. Phys.* **1968**, *49*, 3170.
(19) Platzer, K. E. B.; Ananthanarayanan, V. S.; Andreatta, R. H.; Scheraga, H. A. *Macromolecules* **1972**, *5*, 177.
(20) Ananthanarayanan, V. S.; Andreatta, R. H.; Poland, D.; Scheraga, H. A. *Macromolecules* **1971**, *4*, 417.
(21) It should be noted that, in eq 2, we use average errors, which are more significant than mean-square deviations. Figure 4 and Table IV show that the data scatter around the line for $\Delta\theta_h = 0$ rather than lying all on one side of this line. The mean-square deviations would all be positive and mask this trend in the data.

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

Jeffrey Skolnick†

Department of Chemistry, Washington University, St. Louis, Missouri 63130.
Received September 12, 1984

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

I. Introduction

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

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

† Alfred P. Sloan Foundation Fellow.

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

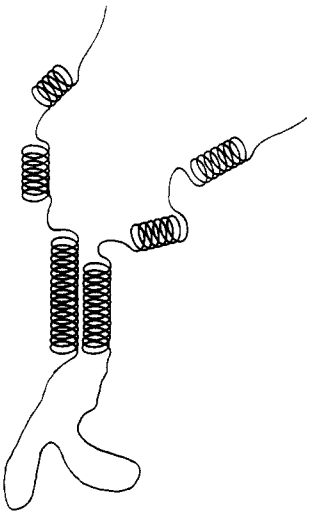

Figure 1. Schematic representation of the accessible conformations in the broken α-helical hairpin model.

Poland and Scheraga assume that the randomly coiled loop between helices occurs essentially in the middle of the chain, we allow for the possibility of hairpins composed of strands of unequal length such as are depicted in Figure 1. Again, because of the effect of loop entropy, essentially the only important mismatched states are these that are out-of-register.

In the following, we develop the theory of the helix–coil transition in broken α-helical hairpins subject to the assumptions listed below. The reader may readily recognize the analogy to the equilibrium theory of the helix–coil transition in two-chain, coiled coils.[8-10]

(1) Each individual amino acid is assumed to have a Zimm–Bragg helix initiation parameter, σ, and propagation parameter, $s(T)$, that are characteristic of the individual amino acids in isolated chains where the possibility of hairpinlike configurations is excluded.[11] Practical applications to proteins require the primary sequence and the table of σ and $s(T)$ values for the amino acids determined by Scheraga et al.;[12-25] a convenient algorithmic form for these quantities has recently been compiled.[26,27]

(2) To account for the enhanced stability of the side-by-side pair of interacting helices in the hairpin conformation, as in the case of two-chain, coiled coils, we introduce the interhelical interaction parameter w.[8-10,28] Physically, $-RT \ln w$ is the interaction free energy of a positionally fixed, side-by-side pair of interacting α-helical turns in the α-helical hairpin configuration relative to the free energy of the pair of positionally fixed helical turns that are noninteracting. In principle w should be both temperature and site dependent. However, in the calculations discussed below, for simplicity, we assume a site-independent w.

In a broken α-helical hairpin, the reduction in configurational entropy of the random coils between the interacting helices must be explicitly accounted for,[29-31] a situation that is not true in the case of non-cross-linked, two-chain, coiled coils. We would expect, and as demonstrated below indeed find, that the incorporation of an interior random coil loop between interacting helices makes the helix–coil transition more cooperative relative to the case where the interior random coil loop is cut and the ends are free. Nevertheless, for typical values of σ and s, the ranges of the helix–helix interaction parameters required to accomplish the coil-to-helix transition from a noninteracting single chain to an essentially completely α-helical hairpin and from a noninteracting pair of chains to an essentially completely α-helical, two-chain, coiled coil are similar. This would seem to indicate that a fairly modest (on the order of 500 cal per mole of pairs of α-helical turns) interhelical interaction could result in a rather augmented helix content.

The remainder of the paper develops the qualitative picture of the helix–coil transition in α-helical hairpin polypeptides. The organization of the paper is as follows. In section II we present the general formalism required to calculate the partition function, overall helix content, helix probability profiles, and fraction of random coils in end random coiled sequences in a polypeptide chain having arbitrary amino acid sequence. We first formulate the perfect matching limit and then examine the more general case that includes the contribution of out-of-register conformations to the aforementioned averages. The reader that is interested only in qualitative results may easily skip this section as it just presents the computational framework required for the calculations. Section III presents an application of the theory to homopolymeric, broken α-helical hairpins, again examining the perfect matching limit before considering the more general situation. We present representative calculations designed to give qualitative insight into the behavior of the helix–coil transition of polypeptides in the context of the broken α-helical hairpin model. Section IV summarizes the results of this paper and indicates several possible directions of future work.

II. General Theory

Perfect Matching Limit. In this section we begin by developing the theory of the α-helix to random coil transition of broken α-helical hairpins in the perfect matching limit. Consider a single heteropolypeptide chain composed of N_T residues divided into $2N_B + l$ blocks, the ith block of which contains m_i residues. As in the case of two-chain, coiled coils, we invoke coarse graining[32] to account for the fact that, in a pair of side-by-side interacting helices, all the m_i residues in the α-helical turn must be in the helical conformation in order to bring about the helix–helix interaction.[28] Furthermore, l is the minimum number of randomly coiled blocks in the turn between interacting helical sequences. (We shall set l equal to unity in the calculations presented in section III.) In the perfect matching limit, the maximum number of interacting α-helical blocks is $2N_B$, i.e., a maximum of N_B blocks per noninterrupted α-helical sequence. In a typical conformation, there will be a single interacting pair of α-helices containing $2L$ residues, with $L < N_B$, and a completely randomly coiled loop between the interacting pairs of α-helices of length p, where $l < p \leq 2N_B + l - 2$. In addition to the interior random coil loop between interacting helical stretches, each of the interacting helices may be preceded by alternating random coil and noninteracting helical sequences.

Statistical Weights of the Block Conformations. Before proceeding to the construction of expressions for the partition function, overall helix content, and fraction of end coils, we summarize the statistical weights necessary in the calculations that follow. Let H(C) symbolize a helical (randomly coiled) block, i.e., a block containing m_i residues in the helical (random coiled) configuration, and let [H]C and [C]H depict random coil and helical blocks located at the interface between a sequence of helical and coiled blocks and randomly coiled and helical blocks, respectively. The statistical weights of the various conformations accessible to the ith block that is neither in the interacting pair of helices nor in the randomly coiled loop between the interacting helices are

conformation	statistical weight
[C]C	1
[H]C	$\mathcal{S}_i = 1 + \sum_{j=1}^{m_i-1} \prod_{k=1}^{j} s_k$
[C]H	$\tau_i = \sum_{j=1}^{m_i} \sigma_j \prod_{k=j}^{m_i} s_k$ (II-1)
[H]H	$SM = \prod_{k=1}^{m_i} s_k$

σ and s are the standard Zimm–Bragg helix initiation and propagation parameters.[11] It proves convenient to place the interior random coil loop at the "left-hand side" of the molecule and number the interacting block pairs from left to right. That is, the pair of blocks that is at the free ends of the molecule is block pair N_B. Furthermore, the statistical weight of the p block randomly coiled stretch and the pair of interfacial [CC]HH helical states adjoining the randomly coiled loop is given by

$$\tau_l^2 w = w \sum_{j=1}^{m_i} \sum_{n=1}^{m_i} \sigma_{j1}\sigma_{n2} \prod_{k=j}^{m_i} \prod_{r=n}^{m_i} s_k s_r \delta(i,j,n) \quad \text{(II-2a)}$$

in which w is the helix–helix interaction parameter, 1 and 2 label the members of the ith helical block pair, and $\delta(i,j,n)$ is the statistical weight of the interior random coil loop of the p blocks constrained to form a closed ring. Following the work of Schellman,[29] Flory,[30] and Jacobson and Stockmayer[31] $\delta(i,j,n)$ is of the form

$$\delta(i,j,n) = C_1(lm + 2\sum_{k=1}^{i-1} m_k + j + n - 2)^{-\alpha} \quad \text{(II-2b)}$$

Hence, $\delta(i,j,n)$ accounts for the reduction in configurational entropy of the p randomly coiled blocks constrained to form a closed loop, that is, whose ends lie within a given volume relative to the case where the ends are free. The orientational and distance constraint on the ends is reflected in the factor C_1.

$$C_1 = u_\phi(3/2\pi)^{3/2} b_0^{-3} \quad \text{(II-2c)}$$

u_ϕ is the volume accessible to the center of mass of one of the interacting [C]H type blocks when the other block is held fixed times an orientational factor that specifies the allowed range of relative orientations of the two helical blocks which permits the interhelical interaction to occur.[33-35] u_ϕ is very much analogous to an interchain helix initiation parameter. In a Gaussian chain $\alpha = 1.5$,[29-31] and if excluded volume effects are considered $1.5 < \alpha \leq 2$.[36]

Construction of the Partition Function. In the calculation of the partition function Z_{hl} of a broken α-helical hairpin, it is convenient to divide up Z_{hl} into two pieces, namely

$$Z_{hl} = Z_m + Z_{int}(N_B) \quad \text{(II-3)}$$

Here Z_m, the partition function of an isolated single chain lacking any helical contacts whatsoever, is given by eq II-9 of ref 9 for a chain consisting of $2N_B + l$ blocks.

$Z_{int}(N_B)$ is the partition function of the broken α-helical hairpin containing up to a maximum of N_B pairs of interacting, α-helical blocks.

Proceeding by analogy to the supermatrix formulation of the helix–coil transition in two-chain, coiled coils[10] we have

$$Z_{int}(N_B) = \mathbf{J}^* \prod_{i=1}^{N_B} \mathbf{U}_{hl,i} \mathbf{J} \quad \text{(II-4a)}$$

$\mathbf{J}^*$ is a row vector consisting of one followed by 11 zeros, and $\mathbf{J}$ is a column vector composed of 4 zeros followed by 8 ones. $\mathbf{U}_{hl}$ is a partitioned 12×12 matrix of the form

$$\mathbf{U}_{hl,i} = \begin{bmatrix} \mathbf{E}_4 & \mathbf{U}_{CH}{}^l & \mathbf{0}_4 \\ \mathbf{0}_4 & \mathbf{U}_{HH} & \mathbf{U}_{HC} \\ \mathbf{0}_4 & \mathbf{0}_4 & \mathbf{U}_d \end{bmatrix}_i \quad \text{(II-4b)}$$

$\mathbf{E}_4$ and $\mathbf{O}_4$ are the 4×4 identity and null matrices, respectively. In the following $\mathbf{O}_n$ denotes an $n \times n$ null matrix. $\mathbf{E}_4$ specifies that if the ith block pair is noninteracting, then it must be part of the interior random coil loop. Moreover, $\mathbf{U}_{CH}{}^l$ is a 4×4 matrix of the form

$$\mathbf{U}_{CH,i}{}^l = \begin{bmatrix} \mathbf{0}_2 & \begin{matrix} 0 & \tau_l^2 w \\ 0 & 0 \end{matrix} \\ \mathbf{0}_2 & \mathbf{0}_2 \end{bmatrix} \quad \text{(II-4c)}$$

Equation II-4c specifies that the ith block pair is at the beginning of the interacting helical sequence and is immediately preceded by the random coil loop containing $l + 2(i-1)$ randomly coiled blocks. In writing eq II-4c we assume that the interior loop is completely randomly coiled and contains no noninteracting helical states. While extension to a more general case is possible, since loop entropy will tend to keep the interior random coil loop quite small, this is a reasonable approximation.

To the right of the interior random coil loop-interacting helical block pair junction, the heteropolypeptide is isomorphic to the loops-excluded, perfect matching model of a two-chain, coiled coil developed previously.[10] Thus, $\mathbf{U}_d$ and $\mathbf{U}_{HH}$, $\mathbf{U}_{HC}$ may be found in eq II-3, II-5a, and II-5b of ref 10.

Overall Helix Content. The overall helix–helix content, f_{hl}, of a broken, α-helical hairpin polypeptide is given by

$$f_{hl} = Z_{hl}^{-1}\{Z_m f_{hm} + \mathbf{J}_s^* \prod_{i=1}^{N_B} \mathbf{A}_{hl,i} \mathbf{J}_s / N_T\} \quad \text{(II-5)}$$

in which Z_m and f_{hm} are the partition function and helix content, respectively, of a noninteracting single-chain polypeptide. The formula for the calculation of f_{hm} may be found in eq II-12 of ref 9. $\mathbf{J}_s^*$ of eq II-5 is a row vector consisting of one followed by 23 zeros. $\mathbf{J}_s$ is a column vector of 16 zeros followed by 8 ones.

$\mathbf{A}_{hl}$ in eq II-5 is a supermatrix of dimension 24×24 in which

$$\mathbf{A}_{hl,i} = \begin{bmatrix} \mathbf{U}_{hl} & \mathbf{U}'_{hl} \\ \mathbf{0}_{12} & \mathbf{U}_{hl} \end{bmatrix}_i \quad \text{(II-6a)}$$

wherein $\mathbf{U}_{hl}$ is defined in eq II-4b ff and

$$\mathbf{U}'_{hl,i} = \begin{bmatrix} \mathbf{0}_4 & \mathbf{U}_{CH}{}^{l'} & \mathbf{0}_4 \\ \mathbf{0}_4 & \mathbf{U}_{HH}{}' & \mathbf{U}_{HC}{}' \\ \mathbf{0}_4 & \mathbf{0}_4 & \mathbf{U}_d{}' \end{bmatrix}_i \quad \text{(II-6b)}$$

Here, we count the contribution to the overall helix content of both blocks "1" and "2" in block pair "i". Now

$$\mathbf{U}_{CH,i}{}^{l'} = \begin{bmatrix} \mathbf{0}_2 & \begin{matrix} 0 & \tau_l^{2\prime} w \\ 0 & 0 \end{matrix} \\ \mathbf{0}_2 & \mathbf{0}_2 \end{bmatrix} \quad \text{(II-7a)}$$

with

$$\gamma_l{}^{2\prime} = \sum_{j=1}^{m_i} \sum_{n=1}^{m_i} \sigma_{j1}\sigma_{n2}(2m_i - j - n + 2) \prod_{k=j}^{m_i} \prod_{r=n} s_k s_r \delta(i,j,n) \quad \text{(II-7b)}$$

Finally, $\mathbf{U}_d'$, $\mathbf{U}_{HH}'$, and $\mathbf{U}_{HC}'$ are defined in eq II-12c, II-12e, and II-12f of ref 10.

Calculation of Helix Probability Profiles. In the single chain under consideration there are $2N_B + l$ distinct blocks. Let the two ends of the chain be blocks 1 and $2N_B + l$, respectively, and let the l completely random coiled blocks in the hairpin be blocks $N_B + 1 \leq k \leq N_B + l$. It is straightforward to show that if $1 \leq k \leq N_B$

$$f_{hl}(k) = Z_{hl}^{-1}\{Z_m f_{hm}(k) + m_k^{-1}\mathbf{J}^* \prod_{i=1}^{j-1} \mathbf{U}_{hl,i}\mathbf{U}'_{hl,1} \prod_{i=j+1}^{N_B} \mathbf{U}_{hl,i}\mathbf{J}\} \quad \text{(II-8)}$$

with $j = N_B - k + 1$ and in which $f_{hm}(k)$ is the helix content of the kth residue in the conformation in which there are no interacting helices (see eq 48 of ref 28 with $w = 1$). m_k is the number of residues in the kth block. Furthermore

$$\mathbf{U}'_{hl,1} = \sum_{n=1}^{m_k} \frac{\partial \mathbf{U}_{hl,k}}{\partial \ln s_n^1} \quad \text{(II-9)}$$

$\mathbf{J}^*$ and $\mathbf{J}$ are defined following eq II-4a.

Now if $N_B + 1 \leq k \leq N_B + l$

$$f_{hl}(k) = Z_m f_{hm}(k) Z_{hl}^{-1} \quad \text{(II-10)}$$

That is, the only contribution to $f_{hl}(k)$ comes from the conformation lacking any interhelical contacts, and if $N_B + l < k \leq 2N_B + l$

$$f_{hl}(k) = Z_{hl}^{-1}\{Z_m f_{hm}(k) + m_k^{-1}\mathbf{J}^* \prod_{i=1}^{j-1} \mathbf{U}_{hl,i}\mathbf{U}'_{hl,2} \prod_{i=j+1}^{N_B} \mathbf{U}_{hl,i}\mathbf{J}\} \quad \text{(II-11a)}$$

with $k = N_B + l + j$; $f_{hm}(k)$ was previously defined.

$$\mathbf{U}'_{hl,2} = \sum_{n=1}^{m_k} \frac{\partial \mathbf{U}_{hl,k}}{\partial \ln s_n^2} \quad \text{(II-11b)}$$

It should be pointed out that if both halves of the molecule are identical, then

$$f_{hl}(N_B - k + 1) = f_{hl}(N_B + l + k) \quad \text{(II-12)}$$

and one need only employ eq II-8 and II-10 to construct the helix probability profile.

Fraction of Random Coils in Terminal Sequences. In this section we calculate f_{hlce}, the fraction of randomly coiled blocks in terminal random coil sequences via

$$f_{hlce} = f_{mce}Z_m Z_{hl}^{-1} + l(Z_{hl} - Z_m)/(2N_B + l)Z_{hl} + [\mathbf{J}_s^* \prod_{i=1}^{N_B} \mathbf{A}_{fl1,i}\mathbf{J}_s + \mathbf{J}_s^* \prod_{i=1}^{N_B} \mathbf{A}_{fl2,i}\mathbf{J}_s + \mathbf{J}_s^* \prod_{i=1}^{N_B} \mathbf{A}_{bl1,i}\mathbf{J}_s + \mathbf{J}_s^* \prod_{i=1}^{N_B} \mathbf{A}_{bl2,i}\mathbf{J}_s]/(2N_B + l)Z_{hl} \quad \text{(II-13)}$$

f_{mce} is the fraction of randomly coiled blocks in an isolated single chain lacking any helical hairpins whatsoever and may be found in eq III-13 of ref 8. The second term in eq II-13 counts the average contribution to f_{hlce} of the l randomly coiled blocks in the loop between interacting helices. $\mathbf{J}_s$ and $\mathbf{J}_s^*$ are defined following eq II-5. The bracketed terms count the contribution to f_{hlce} of blocks that may participate in interhelical interactions. More specifically

$$\mathbf{A}_{fl,q} = \begin{bmatrix} \mathbf{U}_{fl,q} & \mathbf{U}_{fl,q} \\ \mathbf{0}_{12} & \mathbf{U}_{hl} \end{bmatrix}_i \quad \text{(II-14a)}$$

$q = 1, 2$, $\mathbf{U}_{hl}$ is defined in eq II-4b, and

$$\mathbf{U}_{fl,q} = \begin{bmatrix} \mathbf{E}_4 & \mathbf{U}_{fCH,q}^{\,l} & \mathbf{0}_4 \\ \mathbf{0}_4 & \mathbf{0}_4 & \mathbf{0}_4 \\ \mathbf{0}_4 & \mathbf{0}_4 & \mathbf{0}_4 \end{bmatrix}_i \quad \text{(II-14b)}$$

Here

$$\mathbf{U}_{fCH,q}^{\,l} = \left[\begin{array}{cc:cc} \mathbf{0}_2 & & 0 & \tau_{qp}{}^2 w \\ & & 0 & 0 \\ \hdashline \mathbf{0}_2 & & & \mathbf{0}_2 \end{array}\right]_i \quad \text{(II-14c)}$$

with

$$\tau_{qp}{}^2 = \sum_{j=1}^{m_i}\sum_{n=1}^{m_i} \sigma_{j_1}\sigma_{n_2} \prod_{\substack{k=j \\ r=n}}^{m_i} s_k s_r \delta(i,j,n) \times \begin{cases} (j-1)/m_i; & \text{if } q = 1 \\ (n-1)/m_i; & \text{if } q = 2 \end{cases} \quad \text{(II-14d)}$$

Moreover

$$\mathbf{A}_{bl,q} = \begin{bmatrix} \mathbf{U}_{hl} & \mathbf{U}_{bl,q} \\ \mathbf{0}_{12} & \mathbf{U}_{bl,q} \end{bmatrix}_i \quad \text{(II-15a)}$$

for which

$$\mathbf{U}_{bl,q} = \begin{bmatrix} \mathbf{0}_4 & \mathbf{0}_4 & \mathbf{0}_4 \\ \mathbf{0}_4 & \mathbf{0}_4 & \mathbf{U}_{bHC,q} \\ \mathbf{0}_4 & \mathbf{0}_4 & \mathbf{U}_{bd,q} \end{bmatrix}_i \quad \text{(II-15b)}$$

$\mathbf{U}_{bHC,q}$ is a 4×4 matrix defined by

$$\mathbf{U}_{bHC,q} = \left[\begin{array}{cc:cc} \multicolumn{2}{c:}{\mathbf{0}_2} & \multicolumn{2}{c}{\mathbf{0}_2} \\ \hdashline 0 & 0 & 0 & 0 \\ (\mathcal{S}_{p1}\mathcal{S}_{2}\delta_{1q} + \mathcal{S}_{1}\mathcal{S}_{p2}\delta_{2q}) & SM_2\mathcal{S}_{p1}\delta_{1q} & SM_1\mathcal{S}_{p2}\delta_{2q} & 0 \end{array}\right] \quad \text{(II-15c)}$$

with δ_{jk} the Kronecker delta and $\mathcal{S}_p$ is defined in eq III-16b of ref 8.

Finally

$$\mathbf{U}_{bd,q} = \begin{bmatrix} 1 & 0 \\ \mathcal{S}_p & 0 \end{bmatrix} \otimes \mathbf{U}_{m_2}; \quad \text{if } q = 1$$
$$= \mathbf{U}_{m_1} \otimes \begin{bmatrix} 1 & 0 \\ \mathcal{S}_p & 0 \end{bmatrix}; \quad \text{if } q = 2 \quad \text{(II-15d)}$$

with

$$\mathbf{U}_m = \begin{bmatrix} 1 & \tau \\ \mathcal{S} & SM \end{bmatrix} \quad \text{(II-15e)}$$

In the case that the two halves of the molecule are the same, then the bracketed expression of eq II-13 may be replaced with

$$2[\mathbf{J}_s^* \prod_{i=1}^{N_B} \mathbf{A}_{fl1}\mathbf{J}_s + \mathbf{J}_s^* \prod_{i=1}^{N_B} \mathbf{A}_{bl1}\mathbf{J}_s] \quad \text{(II-16)}$$

to reduce the amount of computer time required in the calculation of f_{hlce}. This completes the development of the broken α-helical hairpin model in the perfect matching limit.

Extension to Imperfect Matching. In this section, we shall extend the theory of the helix-to-random-coil transition in broken α-helical hairpins to include the possibility of out-of-register states. Just as in the case of the two-chain, coiled coils, there are two kinds of out-of-register states, denoted by type "a" and type "b" and depicted schematically in Figure 2. Let N be maximum possible number of interacting block pairs in a given conformation, $1 \leq N \leq N_B$. On setting $N = N_B$, we recover

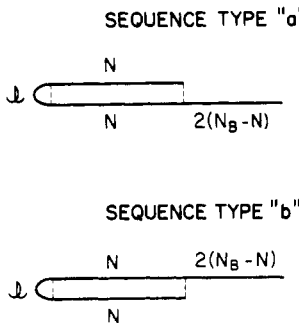

Figure 2. Schematic representation of the two kinds of out-of-register sequences containing a maximum of N interacting pairs of blocks. In sequence type a, for the Nth block pair, blocks 1 and $2N + l$ can interact, and for $i \leq N$, upper block $N - i + 1$ and lower block $N + i + l$ can interact. In sequence type b, upper block $2N_B - N - i + 1$ and lower block $2N_B - N + l + i$ can interact.

the perfect matching limit discussed above. If the single chain consists of identical blocks, only type a sequences need be considered. In the development presented below, we treat the more general case of a heteropolypeptide.

Construction of the Partition Function. Based on previous work in two-chain, coiled coils,[10] the partition function including all in- and out-of-register states is given by

$$Z_{imp} = Z_m + \sum_{N=1}^{N_B} Z_{int}(N) \qquad \text{(II-17)}$$

where $Z_{int}(N_B)$, the contribution of in-register hairpins, may be found in eq II-4a. For all $N < N_B$, we have

$$Z_{int}(N) = \mathbf{J}^* \prod_{i=1}^{N} \mathbf{U}^a{}_{hl,i}\mathbf{J} + \mathbf{J}^* \prod_{i=1}^{N} \mathbf{U}^b{}_{hl,i}\mathbf{J} \qquad \text{(II-18a)}$$

In all that follows, the superscript a (b) refers to a type a (b) sequence. We begin with the type a sequence statistical weight matrices. Let us label the rightmost, "upper" block that can interact as block 1. Thus, for the Nth block pair, block 1 and block $2N + l$ interact, or in general for the ith label in eq II-18a block $N - i + 1$ and block $N + i + l$ can interact. Hence $\mathbf{U}^a{}_{hl,i}$ with $i < N$ is given by eq II-4b ff in which the subscript 1 corresponds to the $N - i + 1$th block and the subscript 2 corresponds to the $N + i + l$th block. When $i = N$, we have to account for the $2(N_B - N)$ "dangling blocks" that cannot form interhelical contacts. Hence

$$\mathbf{U}^a{}_{hl,N} = \begin{bmatrix} \mathbf{U}^{ar} & \mathbf{U}_{CH}{}^{l}\mathbf{U}^{ar} & \mathbf{0}_4 \\ \mathbf{0}_4 & \mathbf{U}_{HH}\mathbf{U}^{ar} & \mathbf{U}_{HC}\mathbf{U}^{ar} \\ \mathbf{0}_4 & \mathbf{0}_4 & \mathbf{U}_d\mathbf{U}^{ar} \end{bmatrix}_N \qquad \text{(II-18b)}$$

where $\mathbf{U}_{CH}{}^{l}$ is given in eq II-4c ($\mathbf{U}_d$, $\mathbf{U}_{HH}$, and $\mathbf{U}_{HC}$ are given in eq II-3, II-5a, and II-5b of ref 10) in which the subscripted block 1 (i) and block 2 ($i - \Delta$) are identified with blocks 1 and $2N + l$, respectively. Moreover

$$\mathbf{U}^{ar} = \mathbf{E}_2 \otimes \prod_{k=2N+l+1}^{2N_B+l} \mathbf{U}_{mk} \qquad N < N_B \qquad \text{(II-18c)}$$

with $\mathbf{E}_2$ the 2×2 identity matrix, $\otimes$ denotes the direct product, and $\mathbf{U}_m$ has been defined in eq II-15e. As a practical matter, since eq II-18c can be cast in the form of a recursion relation, the sum over N in eq II-17 should start with $N = N_B$ and decrease to $N = 1$, thus cutting down the number of computations required to set up $\mathbf{U}^{ar}$ (as well as $\mathbf{U}^{br}$; see eq II-18e below).

We next examine type b sequences as are shown in Figure 2 and return to the previous labeling of blocks i. $\mathbf{U}_{hl,i}{}^{b}$ is defined for all $i < N$ in eq II-4b wherein the subscript 1 corresponds to block $2N_B - N - i + 1$ and subscript 2 corresponds to block $2N_B - N + l + i$. Moreover

$$\mathbf{U}^b{}_{hl,N} = \begin{bmatrix} \mathbf{U}^{br} & \mathbf{U}_{CH}{}^{l}\mathbf{U}^{br} & \mathbf{0}_4 \\ \mathbf{0}_4 & \mathbf{U}_{HH}\mathbf{U}^{br} & \mathbf{U}_{HC}\mathbf{U}^{br} \\ \mathbf{0}_4 & \mathbf{0}_4 & \mathbf{U}_d\mathbf{U}^{br} \end{bmatrix}_N \qquad \text{(II-18d)}$$

$$\mathbf{U}^{br} = \prod_{k=2(N_B-N)}^{1} \mathbf{U}_{mk} \otimes \mathbf{E}_2 \qquad \text{(II-18e)}$$

This completes the construction of the partition function Z_{imp}.

Overall Helix Content. The overall helix content in the broken α-helical hairpin model including the possibility of out-of-register states can be obtained from

$$f_{imp} = f_{hl}Z_{hl}/Z_{imp} + \sum_{N=1}^{N_B-1} f_{int}(N)P(N) \qquad \text{(II-19a)}$$

where formulas f_{hl} and Z_{hl} have been previously given in eq II-5 and II-3, respectively. $f_{int}(N)$ is the average helix content of a molecule having a maximum of N interacting helical block pairs and may be constructed in an analogous fashion to the helix content of an out-of-register, two-chain, coiled coil; see eq II-11 ff of ref 10.[37] Furthermore, $P(N)$ is the probability of a molecule having N interacting pairs of helical blocks, that is

$$P(N) = Z_{int}(N)/Z_{imp}; \qquad \text{if } N > 0$$
$$= Z_m/Z_{imp}; \qquad \text{if } N = 0 \qquad \text{(II-19b)}$$

Calculation of Helix Probability Profiles. Let us consider the mean helix content of the kth block

$$f_{imp}(k) = f_{hl}(k)\frac{Z_{hl}}{Z_{imp}} + \sum_{N=1}^{N_B-1} Z_{int}(N)\{F^a(N,k) + F^b(N,k)\}/Z_{imp} \qquad \text{(II-20)}$$

wherein $f_{hl}(k)$ defined in eq II-8 ff is the contribution to the helix content of the kth block in the perfect matching limit. $F_a(N,k)$ is ($F^b(N,k)$) the helix content of the kth block in a helical hairpin containing a maximum of N interacting helical block pairs in a type a (b) sequence and may be constructed following the procedure presented in eq II-16b of ref 10, and applied to broken α-helical hairpins.[37] $Z_{int}(N)$ is given in eq II-18a. If the chain is homopolymeric, the contribution to the helix content of the kth block in a type a out-of-register sequence, $f_a(k)$, is related to the helix content in a type b sequence $f_b(k)$ by

$$f_a(k) = f_b(2N_B + l - k + 1) \qquad \text{(II-21)}$$

Thus, if homopolymeric chains are employed, one need only consider type a sequences. Moreover, $f_{imp}(k)$ with $k \leq N_B$ is related to $f_{imp}(k)$ for $k > N_B + l$ by the same relation as in eq II-12.

Fraction of Random Coils in Terminal Sequences. The fraction of random coils in terminal sequences f_{ice} may be calculated from

$$f_{ice} = f_{hlce}\frac{Z_{hl}}{Z_{imp}} + \sum_{N=1}^{N_B-1} P(N)f_{ice}(N) \qquad \text{(II-22)}$$

where $f_{ice}(N)$ is the fraction of random coils in terminal sequence or in a random coil interior loop between interacting helices containing a maximum of N interacting blocks. The procedure for the construction of $f_{ice}(N)$ is identical with that in an out-of-register, two-chain, coiled coil which may be found in eq II-27 ff of ref 10.[37]

Summarizing the results of this section, expressions have been presented for the partition function, the overall helix content, the helix probability profiles, and the fraction of

Table I[a]
Summary of Symbols for Various Calculated Quantities

quantity	single-chain, α-helical hairpins prohibited	broken α-helical hairpin, perfect matching limit	broken α-helical hairpin including out-of-register states
partition function	Z_m (eq II-9 of ref 9)	Z_{hl} (eq II-3)	Z_{imp} (eq II-17)
overall helix content	f_{hm} (eq II-12 of ref 9)	f_{hl} (eq II-5)	f_{imp} (eq II-19a)
helix content of the kth block	$f_{hm}(k)$ (eq 48 of ref 28)	$f_{hl}(k)$ (eq II-8)	$f_{imp}(k)$ (eq II-20)
fraction of blocks that are random coils and that propagate from an end or are in an interior random loop between interacting helices	f_{mce} (eq III-13 of ref 8)	f_{hlce} (eq II-13)	f_{ice} (eq II-22)
ratio of number of interior random coil blocks between the interacting pair of helices to the number of random coil blocks in free-end random coil sequences		$R_{hl} = N^0_i/N^0_e$	$R_i = N_i/N_e$

[a] The relevant equation is indicated in parentheses.

random coils in terminal random coil sequences for broken α-helical hairpins in both the perfect matching and imperfect matching cases. A summary of the symbols for these quantities for a single chain without interhelical contacts, including α-helical hairpins but no out-of-register states, and including out-of-register states may be found in Table I.

III. Application of the Broken α-Helical Hairpin Model to Homopolypeptides

In this section we present the results of calculations on a hypothetical homopolypeptide containing 284 residues, divided into 71 blocks, containing 4 residues per block. We shall take $N_B = 35$, and set l, the minimum number of blocks between interacting helical regions of the molecule, equal to 1. Furthermore, $s = 0.94$, and w is assumed to be uniform. We shall set C_1, defined in eq II-2c, equal to 0.6481, an intermediate value within the physically reasonable range of C_1. C_1 is obtained by setting $b_0 = 8.0$ Å,[38] and u_ϕ, the accessible volume equal to $\beta h\pi(d^2_{max} - d^2_{min})$, with h, the length of an α-helical turn, equal to 5.4 Å, $d_{min} = 7.0$ Å,[39] $d_{max} = 14$ Å, a value based on recent X-ray crystallographic data of Phillips on tropomyosin[40] and β, the allowed angular range between helices, equal to 0.403, a plausible estimate. It should be pointed out that the qualitative behavior discussed below is independent of the particular value of C_1; decreasing C_1 is analogous to decreasing the effective interhelical cooperativity parameter.

Perfect Matching Limit. In Figure 3 we plot f_{hl} vs. w calculated via eq II-5 with $\alpha = 1.5$ (see eq II-2b), $\sigma = 10^{-6}$, 10^{-5}, 10^{-4}, 10^{-3}, and 10^{-2} in curves A–E, respectively. The qualitative behavior of this curve is similar to that seen in two-chain, coiled coils.[9] We merely summarize these qualitative features here. At very small values of σ, the only appreciable helix content arises from those chains having interhelical contacts. When σ is increased, the minimum value of w necessary to promote helix formation decreases. In the limit of large w, the curves having larger values of σ approach those with smaller values of σ from below. However, as σ is increased further, such as in curve E, even a single chain lacking any interhelical contacts has an appreciable helix content; thus, the transition becomes broader and the value of w necessary for interhelical contacts to make an appreciable contribution to the partition function increases.

To further investigate the sensitivity of f_{hl} to the specific choice of α, the exponent in the probability for ring closure defined in eq II-2b, we have plotted f_{hl} vs. w for a broken, α-helical homopolypeptide having $s = 0.94$, $\sigma = 5 \times 10^{-4}$, with $\alpha = 1.5$ and $\alpha = 1.8$ in the solid and dashed lines of Figure 4a, respectively. As would be expected on increasing α, the coil–helix transition becomes more cooperative. Basically, increasing α makes it more difficult to form the interior random coil loop necessary to achieve the interhelical contact. For values of low to moderate helix content, this implies that the state lacking any interacting

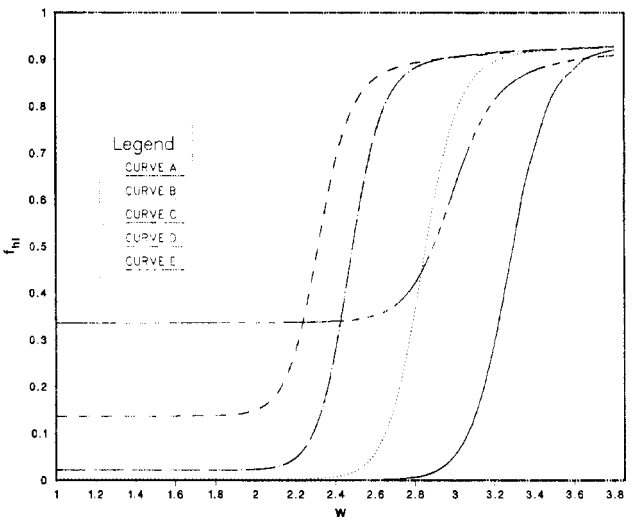

Figure 3. Plot of the helix content for a completely in-register, homopolymeric, broken α-helical hairpin, f_{hl}, vs. the helix–helix interaction parameter w calculated via eq II-5 with $\sigma = 10^{-6}$, 10^{-5}, 10^{-4}, 10^{-3}, and 10^{-2} in curves A–E, respectively. In all cases $N_B = 35$, $l = 1$, $s = 0.94$, and $\alpha = 1.5$.

helices becomes more favored, and the helix content is lower. As w increases to the point where noninteracting helical states are not favored, states having a smaller interior random coil loop (and concomitantly higher helix content) are more favored when $\alpha = 1.8$ than when $\alpha = 1.5$ and the f_{hl} vs. w curves cross. Nevertheless, as evidenced by Figure 4a, the qualitative shape of the f_{hl} vs. w curves remains unchanged.

To further study the effect of α on the nature of the helix–coil transition, it proves most useful to examine the ratio R_{hl}, in molecules containing at least one interacting block pair, of the number of end random coil blocks that unwind from the interior loop, N^0_i, to the number of random coil blocks that unwind from the two free ends of the molecule, N^0_e. R_{hl} is obtained by taking the ratio of the first pair of terms ($\mathbf{A}_{fl}$ type terms) to the second pair of terms ($\mathbf{A}_{bl}$ type terms) in the brackets of eq II-13. In Figure 4B we have plotted R_{hl} vs. w in the perfect matching limit for $\alpha = 1.5$ and $\alpha = 1.8$ in the solid and dashed lines, respectively. Clearly $R_{hl}(\alpha = 1.8) < R_{hl}(\alpha = 1.5)$, even in the limit where the helix content of the two cases is essentially indistinguishable.

It should also be pointed out that R_{hl} is a monotonically increasing function of w. This can be rationalized as follows. Consider a chain in the limit that all $2N_B$ blocks are helical and whose maximum helix content is 0.9859 (the four residues in the interior random coil loop of minimum length cannot be helical). In this case $R_{hl} = \infty$. As the chain starts to unwind because of the effect of loop entropy, the interior random coil loop is kept as small as possible. Most, but not all, of the unwinding takes place from the free ends. Thus, R_{hl} decreases with decreasing

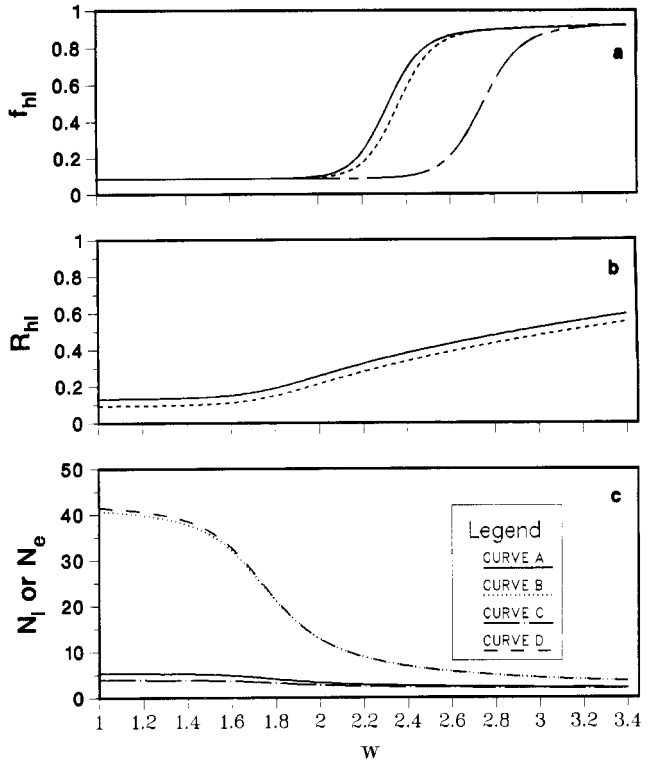

Figure 4. a: Plot of the helix content for a completely in-register, homopolymeric, broken α-helical hairpin, f_{hl}, vs. the helix–helix interaction parameter w calculated via eq II-5 with $C_1 = 0.6481$ and $\alpha = 1.5$ and 1.8 in the solid and dashed lines, respectively; $C_1 = 3.37 \times 10^{-3}$ and $\alpha = 1.5$ in the broken solid line. In all cases $N_B = 35$, $l = 1$, $\sigma = 5 \times 10^{-4}$, and $s = 0.94$. b: Plot of the ratio of the number of interior random coil blocks between the interacting pair of helices to the number of random coil blocks in free-end random coil sequences in a completely in-register, homopolymeric broken α-helical hairpin, R_{hl}, vs. the helix–helix interaction parameter w with $\alpha = 1.5$ and 1.8 in the solid and dashed lines, respectively. $N_B = 35$, $l = 1$, $\sigma = 5 \times 10^{-4}$, and $s = 0.94$. c: Plot of the number of interior random coil blocks between the interacting pair of helices, N^0_i, and the number of random coil blocks in free-end random coil sequences, N^0_e, in a completely in-register homopolymeric broken α-helical hairpin vs. the helix–helix interaction parameter w, with $\alpha = 1.5$ ($\alpha = 1.8$) in curves A and B (curves C and D), respectively. In all cases, $N_B = 35$, $l = 1$, $\sigma = 5 \times 10^{-4}$, and $s = 0.94$.

w. We have plotted in Figure 4c N^0_i and N^0_e vs. w with $\alpha = 1.5$ ($\alpha = 1.8$) in curves A and B (curves C and D), respectively, for the same homopolymeric chains as in Figure 4, a and b. As would be expected, $N^0_i(\alpha = 1.8) < N^0_i(\alpha = 1.5)$, and $N^0_e(\alpha = 1.5) \leq N^0_e(\alpha = 1.8)$ with the most drastic effect occurring in N^0_i. Increasing α makes the formation of interior loops more difficult and essentially leaves N^0_e unchanged. (There is a small increase in N^0_e on increasing α from 1.5 to 1.8, but the dominant effect is to diminish N^0_i.) Note that in both cases $N^0_i > l$, and N^0_i is a monotonically decreasing function of w decreasing from 5.30 (3.87) blocks when $\alpha = 1.5$ ($\alpha = 1.8$) to 2.17 (2.00) as w increases from 1.0 to 3.2. Clearly, setting $N^0_i = 1.0$ as Poland and Scheraga have done is incorrect.[1] There is a distinct tightening of the loop as the helix content of the broken hairpin increases.

To quantitively examine the effect of decreasing C_1 on the helix content we have also plotted in Figure 4a f_{hl} vs. w assuming $C_1 = 3.37 \times 10^{-3}$, with $\alpha = 1.5$ in the broken solid line. This corresponds to an orientational factor β, of 2.1×10^{-3} (obtained by assuming a maximum angle of 10° between the principal axes of the two helical turns and an allowed angular range of contact of 100°). At lower values of the helix content, the f_{hl} vs. w curve is smaller

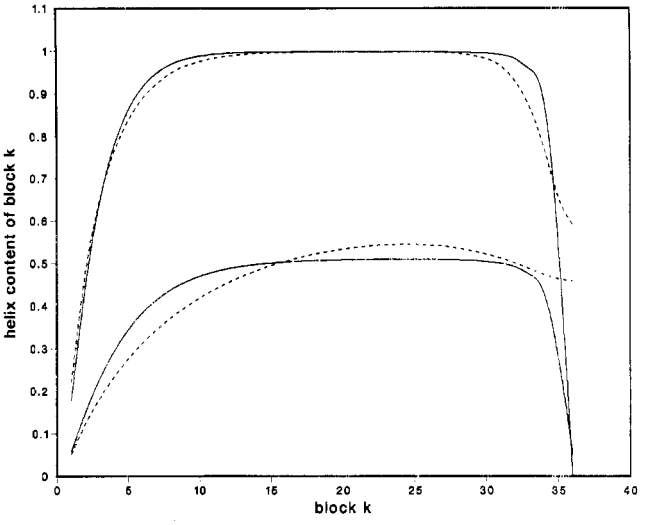

Figure 5. Plot of the helix content of the kth block vs. k, for a broken α-helical hairpin in the perfect matching (imperfect matching) limit $f_{hl}(k)$ ($f_{imp}(k)$) vs. k with $\alpha = 1.5$, $s = 0.94$, and $\sigma = 5 \times 10^{-4}$ and with $f_{hl} = 0.4371$ ($f_{imp} = 0.4426$) and 0.9019 (0.9028) in the lower and upper solid (dashed) curves, respectively.

when $C_1 = 3.37 \times 10^{-3}$ than when $C_1 = 0.6481$. However, in the limit of high helix content the two curves coalesce; the only important contribution to Z_{hl} arises from interacting helical states and there is only one interacting pair of helices per conformation. Hence, in the limit that w goes to infinity, $Z_{hl} \sim C_1 g(w,\alpha)$, wherein $g(w,\alpha)$ is a function of w and α, but is independent of C_1.

We now turn to the calculation of helix probability profiles. Since the chain is by hypothesis homopolymeric, if $1 \leq k \leq N_B$, $f_{hl}(k) = f_{hl}(2N_B - k + 1 + l)$. Hence, we need only consider the first $N_B + 1$ blocks. In Figure 5 we have plotted $f_{hl}(k)$ vs. k obtained from eq II-8 for a homopolymeric, broken α-helical hairpin in the perfect matching limit with $\alpha = 1.5$, $s = 0.94$, and $\sigma = 5 \times 10^{-4}$ and having f_{hl} equal to 0.4371 ($w = 2.3$) and 0.9019 ($w = 2.9$) in the lower and upper solid curves, respectively. Observe that the helix probability profiles at both moderate and high helix are not symmetric—they are skewed to larger values of k. Since the dominant unwinding occurs from the free ends (small k), this is to be expected on the basis of the above discussion. The helix content of block 36 is due entirely to the single chain conformations that lack any interhelical contacts whatsoever and goes to zero as the helix content increases. (The contribution of the state lacking any interhelical contacts goes to zero as the overall helix content approaches $2N_B/(2N_B + l)$.) We remind the reader that block 36 is located in the middle of the homopolymeric chain. As expected on increasing f_{hl}, the region where $f_{hl}(k)$ vs. k is essentially flat increases.

Inclusion of Imperfect Matching. In the following we again assume that there are 71 blocks per chain, with a maximum number of interacting pairs of helical turns, N_B, equal to 35. All out-of-register states are assumed to be accessible and no cognizance is taken of the possibility of the exclusion of salt-bridge-hydrophobic contacts such as is required in the treatment of the helix–coil transition in the two-chain, coiled coil tropomyosin.[41] This assumption is not an intrinsic limitation of the formalism, but rather is assumed here merely for definiteness. Hence, N, the maximum number of interacting blocks in a given out-of-register state, is allowed to range over all $1 \leq N \leq N_B$. We point out that fixing the total number of residues, N_T, at 284 is not really restrictive; calculations of the overall helix content as a function of the helix–helix interaction parameter are presented over a sufficiently wide

Table II
Population Distribution and Helix Content for Broken α-Helical Hairpins Having a Maximum of N Interacting Helical Block Pairs[a]

N	$P(N)^b$	$f(N)^b$	$P(N)^c$	$F(N)^c$
35	1.396-2	0.7584	2.032-1	0.9139
34	2.676-2	0.7515	3.238-1	0.9054
33	2.431-2	0.7364	2.053-1	0.8886
32	2.513-2	0.7173	1.196-1	0.8685
31	1.881-2	0.6962	6.713-2	0.8464
30	1.632-2	0.6739	3.696-2	0.8232
29	1.410-2	0.6510	2.016-2	0.7991
28	1.215-2	0.6279	1.093-2	0.7745
27	1.405-2	0.6047	5.916-3	0.7495
26	8.945-3	0.5816	3.196-3	0.7243
25	7.713-3	0.5586	1.725-3	0.6990
24	6.614-3	0.5358	9.308-4	0.6736
23	5.666-3	0.5132	5.021-4	0.6482
22	4.847-3	0.4909	2.708-4	0.6227
21	4.141-3	0.4690	1.461-4	0.5972
20	3.531-3	0.4473	7.876-5	0.5717
19	3.005-3	0.4260	4.248-5	0.5462
18	2.551-3	0.4052	2.290-5	0.5207
17	2.160-3	0.3847	1.235-5	0.4951
16	1.823-3	0.3647	6.661-6	0.4697
15	1.532-3	0.3451	3.591-6	0.4442
14	1.281-3	0.3261	1.936-6	0.4187
13	1.065-3	0.3077	1.044-6	0.3933
12	8.793-4	0.2896	5.624-7	0.3680
11	7.194-4	0.2723	3.028-7	0.3428
10	5.820-4	0.2556	1.629-7	0.3178
9	4.642-4	0.2396	8.740-8	0.2932
8	3.635-4	0.2242	4.671-8	0.2691
7	2.778-4	0.2097	2.478-8	0.2457
6	2.054-4	0.1960	1.297-8	0.2234
5	1.448-4	0.1832	6.626-9	0.2026
4	9.501-5	0.1714	3.231-9	0.1837
3	5.515-5	0.1610	1.436-9	0.1674
2	2.492-5	0.1520	5.159-10	0.1542
1	4.803-6	0.1450	8.770-11	0.1450
0	7.829-1	0.08546	8.933-6	0.08546

[a] We have used the notation $1.5 \times 10^{-3} \equiv 1.5\text{-}3$. [b] $w = 2.0$, $f_{imp} = 0.2061$. [c] $w = 3.2$, $f_{imp} = 0.8863$.

range of values of σ that the entire qualitative behavior of f_{imp} is displayed.

In Figure 6, we present curves of f_{imp} vs. w calculated via eq II-19a for a homopolymeric, broken α-helical hairpin with $\alpha = 1.5$, $C_1 = 0.6418$, and values of $\sigma = 10^{-6}$, 10^{-5}, 10^{-4}, 10^{-3}, and 10^{-2} in curves A–E, respectively. Comparison of the curves in Figure 6 (that include mismatched states) with those in Figure 3 (the perfect matching limit) at identical σ reveals that inclusion of mismatch has made the helix–coil transition broader and less cooperative. This is easily rationalized. At low values of the helix content the dominant contribution to the partition function comes from the noninteracting chain, a state whose helix content is the lowest. By including states that are out-of-register but interacting, the relative contribution of states whose helix content is higher increases, hence $f_{imp} > f_{hl}$. Consider now the case of sufficiently high helix content such that the noninteracting chain contribution to f_{imp} is negligible. Since the completely in-register state has the highest helix content, inclusion of states that are out-of-register acts to decrease the overall mean helix content. The above qualitative conclusions are verified in Table II, where we present $P(N)$, the probability of finding a molecule having a maximum of N interacting block pairs, and $f(N)$, the average helix content of a chain having a maximum of N interacting block pairs for $w = 2.0$ ($f_{imp} = 0.2061$), and $w = 3.2$ ($f_{imp} = 0.8863$) in columns two and three, and four and five, respectively.

To examine the effect of α (defined via eq II-2b) on the overall helix content, in Figure 7a we have plotted f_{imp}

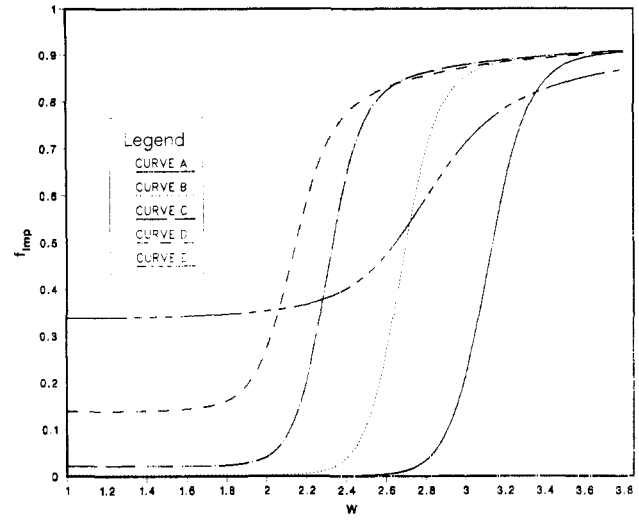

Figure 6. Plot of the helix content for a homopolymeric broken α-helical hairpin including out-of-register states, f_{imp}, vs. the helix–helix interaction parameter w calculated via eq II-19a with $\sigma = 10^{-6}$, 10^{-5}, 10^{-4}, 10^{-3}, and 10^{-2} in curves A–E, respectively. In all cases, $N_B = 35$, $l = 1$, $s = 0.94$, and $\alpha = 1.5$.

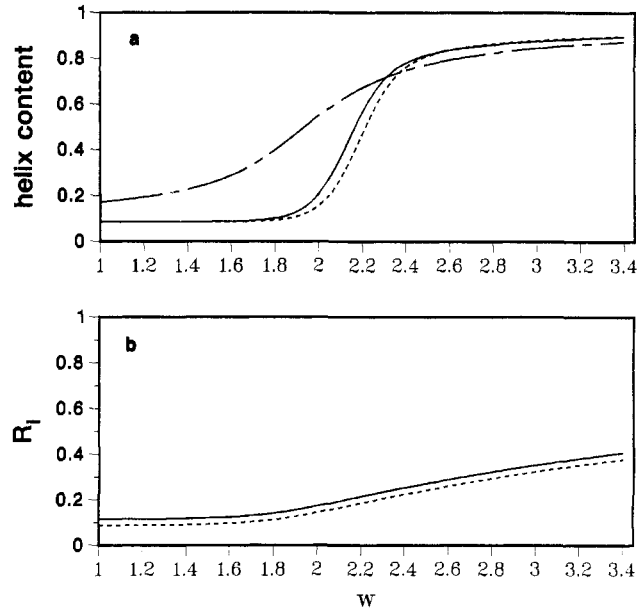

Figure 7. a: Plot of the helix content for a homopolymeric broken α-helical hairpin including out-of-register states, f_{imp}, vs. the helix–helix interaction parameter w calculated via eq II-19a with $C_1 = 0.6481$, $\alpha = 1.5$ and 1.8 in the solid and dashed lines, respectively, with $N_B = 35$, $l = 1$, $\sigma = 5 \times 10^{-4}$, and $s = 0.94$. The broken solid line is the plot of the helix content in a two-chain, coiled coil, f_{hd}, obtained from eq III-1, vs. the helix–helix interaction parameter w. Each of the two chains is assumed to contain 35 blocks, each having 4 residues per block, $\sigma = 5 \times 10^{-4}$ and $s = 0.94$. b: Plot of the ratio of the number of interior random coil blocks between the interacting pair of helices to the number of random coil blocks in free-end random coil sequences in a homopolymeric broken α-helical hairpin including out-of-register states, R_i, vs. the helix–helix interaction parameter w with $C_1 = 0.6481$, $\alpha = 1.5$ and $\alpha = 1.8$ in the solid and dashed lines, respectively; $N_B = 35$, $l = 1$, $\sigma = 5 \times 10^{-4}$, and $s = 0.94$.

obtained from eq II-19a as a function of w for $\alpha = 1.5$ and $\alpha = 1.8$ with C_1 equal to 0.6481 and $\sigma = 5 \times 10^{-4}$ in the solid and dashed curves, respectively. Furthermore, to compare the character of the helix–coil transition in broken, α-helical hairpins with that in two-chain, coiled coils, we have also plotted in the broken solid line of Figure 7a the helix content, f_{hd}, vs. w of a hypothetical two-chain, coiled coil containing 35 blocks per chain, with $s = 0.94$ and $\sigma = 5$

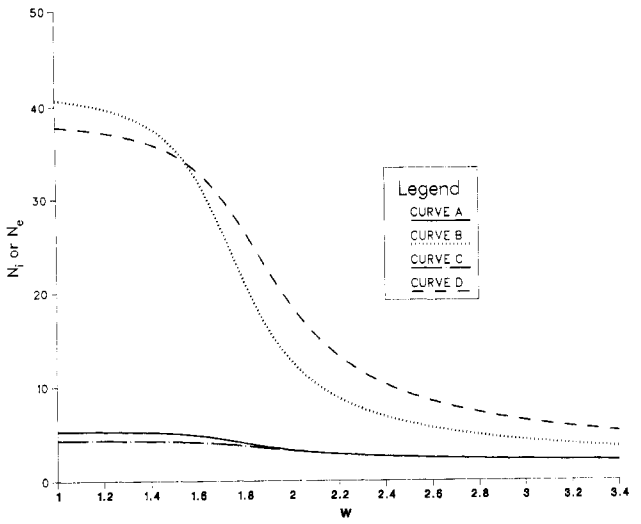

Figure 8. Plot of the number of interior random coil blocks between the interacting pair of helices, N^0_i (N_i), and the number of blocks in free-end random coil sequences N^0_e (N_e) excluding (including) out-of-register states vs. the helix–helix interaction parameter w in the curves A and B (C and D), respectively. In all cases $\alpha = 1.5$, $C_1 = 0.6481$, $\sigma = 5 \times 10^{-4}$, and $s = 0.94$.

$\times 10^{-4}$, and in which the noninteracting helical state is allowed. More explicitly

$$f_{hd} = \frac{f_{hm}Z_m^2}{(Z_m^2 + Z_{dimer})} + \frac{f_{dimer}Z_{dimer}}{(Z_m^2 + Z_{dimer})} \qquad \text{(III-1)}$$

Z_{dimer} and f_{dimer} may be obtained from consideration of the internal partition function and helix content of a two-chain, coiled coil (eq II-1, II-2a, and II-11 of ref 10, respectively). As would be expected, since the broken α-helical hairpin contains an interior closed random coil loop between interacting helices, the helix–coil transition is more cooperative than in the corresponding hypothetical two-chain, coiled coil. Nevertheless, the magnitude of the helix–helix interaction parameter required to bring about the helix–coil transition in both systems is quite similar.

We next turn to the calculation of R_i, defined analogously to R_{hl} as the ratio of the number of interior random coil blocks between the interacting pair of helices, N_i, to the number of randomly coiled blocks in the free-end random coil sequences, N_e.[37]

In Figure 7b, we plot R_i vs. w for a homopolymeric broken α-helical hairpin having $s = 0.94$ and $\sigma = 5 \times 10^{-4}$, $C_1 = 0.6481$, with $\alpha = 1.5$ and 1.8 in the solid and dashed lines, respectively. Just as in the perfect matching limit $R_i(\alpha = 1.8) < R_i(\alpha = 1.5)$, with R_i a monotonically increasing function of w.

A more detailed understanding of the relationship between the R_{hl} and R_i emerges from Figure 8, where we plot, for $\alpha = 1.5$, $C_1 = 0.6481$, $\sigma = 5 \times 10^{-4}$, and $s = 0.94$, N^0_i and N^0_e, for chains in the perfect matching limit matching in curves A and B, and N_i and N_e, the corresponding quantities including out-of-register states, in curves C and D, respectively. Observe that N_i is always less than or equal to N^0_i, with the former quantity converging to the latter from below. The origin of this effect is as follows. In the limit that $w = 1$, all out-of-register states are essentially equally populated; the more out-of-register a state is, the smaller must be the corresponding N_i. For example $N_i = 1$ when $N = 1$; there is merely a loop composed of a single block. Thus, inclusion of out-of-register states must decrease N_i. However, as the helix content increases, the population is dominated by states of high helix content that are relatively close to being in-register. In Table II, column four, the states with $N \geq 30$ have essentially the same size interior random coil loop to minimize the effect of loop entropy. However, since the chain is out-of-register, it contains free ends that are more frayed than the perfect matching limit. This is also the reason that $N_e > N^0_e$ in the high helix content regime. In the other extreme where $w \to 1$, as mentioned above, all of the out-of-register states contribute essentially equally to the partition function. In particular, the same out-of-register states that make $N_i < N^0_i$ also make N_e less than N^0_e. For example, the chain with $N = 1$ has only one end that can be randomly coiled (there is a single pair of blocks that must be helical on one of the ends) and has a helix content of 0.1450; the in-register hairpin with $w = 1$ has a helix content of 0.1753. Hence, N_e has decreased relative to the in-register case in the limit of small w. The latter effect diminishes with increasing helix content. Thus, over the majority of the range of w, the decrease in R_i as compared to R_{hl} is due to the greater fraction of random coils that unwind from the free ends in the imperfect matching relative to the perfect matching limit. Finally, we again point out that N_i exceeds unity over the range of helix contents studied and decreases very slowly with increasing w. It does not converge to unity until the maximum possible helix content, $2N_B/2N_B + l$, is reached. Here again, in the imperfect matching limit, neglect of the variable loop size between interacting helices is incorrect.

In the imperfect matching case any one of blocks 2 to $2N_B + l - 1$ may be the interior random coil block between interacting pairs of helices. The fact that block $N_B + 1$ need not be in the all random coil state has important effects on the shape of the helix probability profile. This is clearly seen in Figure 5, where we have plotted $f_{imp}(k)$ vs. k calculated via eq II-20 for a homopolymeric broken α-helical hairpin with $\alpha = 1.5$, $C_1 = 0.6481$, and $f_{imp} = 0.4426$ and 0.9028 in the lower and upper dashed curves, respectively. Since two halves of the molecule with $k > N_B + l$ and $k \leq N_B$ have the same helix probability profile (see eq II-21), we need only examine $k \leq N_B + l$. Observe that the minimum in $f_{imp}(k)$, at $k = N_B + 1$, becomes more pronounced as w increases. Basically in the limit of high helix content $f_{imp}(k) \to f_{hl}(k)$; otherwise stated the imperfect matching case converges to the perfect matching limit as $w \to \infty$. For lower values of the helix content, the more out-of-register states make a greater contribution to the helix content. Furthermore, consistent with the behavior of N_e shown in Figure 8, at essentially the same helix content there is more melting from the ends in the imperfect matching case than in the perfect matching limit.

Summarizing the results of this section, we have applied the theory of the helix-to-random-coil transition of broken α-helical hairpins to homopolypeptides with and without the possibility of out-of-register states. In both the perfect matching limit and the more general case that allows for imperfect matching, if a uniform helix–helix interaction parameter is assumed, the coil-to-helix transition proceeds mostly by rewinding of the two free ends, but also to a lesser extent by the tightening of the interior randomly coil loop between the interacting pair of helices. The inclusion of "out-of-register" states is seen to make the helix–coil transition less cooperative relative to the "perfect matching limit", where the maximum length of the interacting helices is essentially half the length of the molecule. As in the case of two-chain, coiled coils, the imperfect matching helical hairpin converges to the perfect matching limit as the helix–helix interaction parameter w is increased; i.e., the relative population of the in-register states, $P(N_B)$, converges to unity in the limit that $w \to \infty$.

Comparison of the helix–coil transition in broken α-helical hairpins with an analogous two-chain, coiled coil reveals that as expected the helix–coil transition of the former is more cooperative; nevertheless, the magnitude of the helix–helix interaction parameters necessary to induce helix formation is quite similar. Thus, the broken α-helical hairpin model points out the plausibility of interhelical interactions as a means of stabilizing the tertiary structure in possible nascent forms of globular proteins in the early stages of protein folding.

IV. Discussion

This paper has developed a theory for the helix–coil transition of single-chain polypeptides in which interhelical contacts between noncontiguous pieces of the chain act to augment the helix content and stabilize the α-helical structure. We explicitly focused on the broken α-helical hairpin, a generalization of the Poland and Scheraga hairpin.[1-3] We allow for the possibility of a single interacting pair of α-helices, joined together by an interior random coil stretch of arbitrary length greater or equal to some minimum value, and where each of the interacting helices may perhaps have attached tails containing stretches of noninteracting helices punctuated by random coil stretches. Expressions for the partition function, the overall helix content, the helix probability profiles, and the fraction of random coils in free ends and in the interior loop between interacting helices have been derived both for the situation where the maximum length of the interacting helices is essentially half the length of the chain, the "perfect matching" case, and in the less restrictive case where any non-nearest-neighbor pair of helical turns in the chain are allowed to interact, the imperfect matching, out-of-register case. The formalism is applicable to single-chain heteropolypeptides and is similar in spirit to the theory of the helix–coil transition in α-helical, two-chain, coiled coils developed previously.[8-10,28]

We have applied the theory to a homopolypeptide, chosen such that its intrinsic helix content in the absence of interhelical stabilization is small. If Zimm–Bragg theory holds, this is likely to be the case for most proteins in which tertiary and quaternary interactions are absent.[42] The helix–coil transition in broken α-helical hairpins is seen to be substantially narrower than the analogous transition in a two-chain, coiled coil of identical length. This reflects the enhanced cooperativity introduced by the presence of an interior random coil loop; when interior random coil loops are unavoidably necessary to achieve a substantial helix content, due to the effect of loop entropy, the transition is greatly sharpened. It would be interesting to see how sharp the transition becomes on introduction of multiple bends due to preferential interactions between interacting helices and at what point the ratio of the van't Hoff to the calorimetric enthalpy, a traditional measure in proteins of whether a transition proceeds as "all or none", approaches unity within experimental error.[43] This of course brings up the question of the relative stability of two interacting helices as compared to three interacting helices in a single polyamino acid chain. The third helix either may interact with the two helices, as was treated by Poland and Scheraga, or may form a helical sheet in which helix i only interacts with helix $i - 1$. We consider the latter case for purposes of illustration and assume a uniform w. The formation of a third helix by necessity involves the creation of a closed interior random coil loop of length $n \geq l$, blocks. In the limit of high helix and assuming a uniform site-independent w, the most probable state consists of three interacting helices of equal length N; the ratio of the statistical weight of this state to the conformational state containing two fully helical interacting helices is approximately $\sigma C_1 n^{-3/2}$. Thus, since σ is on the order of 10^{-2}–10^{-4} in proteins and C_1 is estimated to be between $0.7 \geq C_1 \geq 3.4 \times 10^{-3}$, clearly, the pair of two interacting helices will dominate the population for typical proteins. Consider now the limit of moderate to low helix content where, assuming a uniform helix–helix interaction, out-of-register states contribute significantly to the population. In the present treatment, for an interacting helical stretch of maximum length N, we allow for the full complement of noninteracting states in the $2(N_B - N)$ blocks in the noninteracting tail. The statistical weights of these states must be compared to the relative statistical weight of the third interacting helical stretch of maximum length $2(N_B - N) - l$. An estimate of their relative importance may be made by comparing the ratio Q of the statistical weight of a helical state of length R preceded by l random coil blocks in the interfacial region between the second pair of interacting helices to that of the noninteracting states, including the fully helical but noninteracting state of length $R + l$. This gives $Q = w^R \tau S M^{R-1} \delta(l)/(SM^{R+l} + \tau SM^{R-1})$. Now plausible estimates of $\delta(l)$ range from about 10^{-1} to 10^{-4}. Hence, whether or not Q exceeds unity will depend on the competition between the enhanced stability of the interacting helical stretch relative to the loss of entropy on the formation of the loop; in general this is likely to be very sensitive to site-specific factors. The inclusion of multiple interacting helices in the formalism will be the object of future work.

It should be pointed out that the formalism developed here for broken α-helical hairpins can be extended to treat the helix–coil transition of singly and doubly cross-linked, two-chain, coiled coils in which loop entropy is explicitly included. The presence of a cross-link affords the possibility of a single as well as double random coil loops between interacting helical stretches and is intimately related to the problem treated here. This is the subject of a future work.[44] Moreover, a cross-linked, synthetic, two-chain, coiled coil that is closely related to the broken α-helical hairpin has been synthesized by Hodges et al.[45] Since singly and perhaps, in the future, doubly cross-linked tropomyosin can be prepared, such a theory is required before the qualitative understanding of the helix-coil transition in two-chain, coiled coils where the stability is dominated by interhelical interactions is complete.

α-Helices have been invoked by Finkelstein and Ptitsyn[4,5] and Lim[6] as a nascent structure along the folding pathway of globular proteins and as structural elements or microdomains in the diffusion–collison–adhesion model of Karplus and Weaver.[46-48] Another possibility is that the diffusion between unstructured microdomains of low helix content occurs rapidly, but the formation of structured, interacting microdomains of higher helix content might be relatively slow. In this case, the microdomains of appreciable helix content are formed after rather than before the segments coalesce. If any of the approaches discussed above is correct, broken α-helical hairpins are likely to be found in the early stages of protein folding and thus their equilibrium properties and kinetic behavior are of interest. This paper has developed the necessary formalism to treat the former time-independent properties, and in forthcoming work we shall examine the kinetics of folding in broken α-helical hairpins. We believe the study of this simple model system is illustrative for pointing out some of the qualitative features of the far more complicated general protein folding problem.

Acknowledgment. This research was supported in part by a grant from the Biophysical Program of the National

Science Foundation (No. PCM 82-12404). Stimulating discussions with Professor Alfred Holtzer are gratefully acknowledged.

References and Notes

(1) Poland, D. C.; Scheraga, H. A. *Biopolymers* **1965**, *3*, 305.
(2) Poland, D. C.; Scheraga, H. A. *Biopolymers* **1965**, *3*, 335.
(3) Poland, D.; Scheraga, H. A. "Theory of Helix–Coil Transitions in Biopolymers"; Academic Press: New York, 1970; sections 9H and 9I.
(4) Finkelstein, A. V.; Ptitsyn, O. B. *J. Mol. Biol.* **1976**, *103*, 15.
(5) Ptitsyn, O. B.; Finkelstein, V. A. *Q. Rev. Biophys.* **1980**, *13*, 339.
(6) Lim, V. I. *FEBS Lett.* **1979**, *89*, 10.
(7) Matheson, R. R.; Scheraga, H. A. *Macromolecules* **1978**, *11*, 819.
(8) Skolnick, J. *Macromolecules* **1983**, *16*, 1069.
(9) Skolnick, J. *Macromolecules* **1983**, *16*, 1763.
(10) Skolnick, J. *Macromolecules* **1984**, *17*, 645.
(11) Zimm, B.; Bragg, J. *J. Chem. Phys.* **1959**, *31*, 526.
(12) Ananthanarayanan, V. S.; Andreatta, R. H.; Poland, D.; Scheraga, H. A. *Macromolecules* **1971**, *4*, 417.
(13) Platzer, K. E. B.; Ananthanarayanan, V. S.; Andreatta, R. H.; Scheraga, H. A. *Macromolecules* **1972**, *5*, 177.
(14) Alter, J. E.; Taylor, G. T.; Scheraga, H. A. *Macromolecules* **1972**, *5*, 739.
(15) Van Wart, H. E.; Taylor, G. T.; Scheraga, H. A. *Macromolecules* **1973**, *6*, 266.
(16) Alter, J. E.; Andreatta, R. H.; Taylor, G. T.; Scheraga, H. A. *Macromolecules* **1973**, *6*, 564.
(17) Maxfield, F. R.; Alter, J. E.; Taylor, G. T.; Scheraga, H. A. *Macromolecules* **1975**, *8*, 479.
(18) Scheule, R. K.; Cardinaux, F.; Taylor, G. T.; Scheraga, H. A. *Macromolecules* **1976**, *9*, 23.
(19) Dygert, M. K.; Taylor, G. T.; Cardinaux, F.; Scheraga, H. A. *Macromolecules* **1976**, *9*, 794.
(20) Matheson, R. R.; Nemenoff, R. A.; Cardinaux, F.; Scheraga, H. A. *Biopolymers* **1977**, *16*, 1567.
(21) van Nispen, J. W.; Hill, D. J.; Scheraga, H. A. *Biopolymers* **1977**, *16*, 1587.
(22) Hill, D. J.; Cardinaux, F.; Scheraga, H. A. *Biopolymers* **1977**, *16*, 2447.
(23) Konishi, Y.; van Nispen, J. W.; Davenport, G.; Scheraga, H. A. *Macromolecules* **1977**, *10*, 1264.
(24) Kobayashi, Y.; Cardinaux, F.; Zweifel, B. O.; Scheraga, H. A. *Macromolecules* **1977**, *10*, 1271.
(25) Hecht, M. H.; Zweifel, B. O.; Scheraga, H. A. *Macromolecules* **1978**, *11*, 545.
(26) Skolnick, J.; Holtzer, A. *Macromolecules* **1982**, *15*, 812.
(27) Holtzer, M. E.; Holtzer, A.; Skolnick, J. *Macromolecules* **1983**, *16*, 173.
(28) Skolnick, J.; Holtzer, A. *Macromolecules* **1982**, *15*, 303.
(29) Schellman, J. A. *C. R. Trav. Lab. Carlsberg.*, Ser. Chim. **1955**, *29*, 230.
(30) Flory, P. J. *J. Am. Chem. Soc.* **1956**, *78*, 5222.
(31) Jacobson, H.; Stockmayer, W. H. *J. Chem. Phys.* **1950**, *18*, 1600.
(32) Crothers, D.; Kallenbach, N. *J. Chem. Phys.* **1966**, *45*, 917.
(33) Mayer, J.; Mayer, M. "Statistical Mechanics"; Wiley: New York, 1940; Chapter 9.
(34) Zimm, B. *J. Chem. Phys.* **1960**, *33*, 1349.
(35) u_ϕ also contains the ratio, q_ϕ, of the volume of configuration space spanned by the internal degrees of freedom in the helical block in the hairpin to that in the noninteracting state. Since q_ϕ is unknown, we arbitrarily set it equal to unity.
(36) de Gennes, P. G. "Scaling Concepts in Polymer Physics"; Cornell University Press: Ithaca, NY, 1979.
(37) The reader interested in the specific details should write the author.
(38) Holtzer, M.; Holtzer, A. *Macromolecules* **1972**, *5*, 294.
(39) Chou, K. C.; Nemethy, G.; Scheraga, H. A. *J. Chem. Phys.* **1983**, *87*, 2869.
(40) Phillips, G. N. *Biophys. J.* **1984**, *45*, 392a.
(41) Skolnick, J.; Holtzer, A. *Macromolecules*, in press.
(42) Mattice, W. L.; Srinivasan, G.; Santiago, G. *Macromolecules* **1980**, *13*, 1254.
(43) Privalov, P. L. *Adv. Protein Chem.* **1979**, *33*, 167.
(44) J. Skolnick, *Macromolecules*, in press.
(45) Hodges, R.; Saund, A.; Chong, P.; St. Pierre, S.; Reid, R. *J. Biol. Chem.* **1981**, *256*, 1214.
(46) Karplus, M.; Weaver, D. L. *Nature (London)* **1976**, *260*, 404.
(47) Karplus, M.; Weaver, D. L. *Biopolymers* **1979**, *18*, 1421.
(48) Weaver, D. L. *Biopolymers* **1982**, *21*, 1275.

α-Helix-to-Random-Coil Transition of Two-Chain, Coiled Coils. Light Scattering Experiments on the Thermal Denaturation of α-Tropomyosin†

S. Yukioka, Ichiro Noda, and Mitsuru Nagasawa

Department of Synthetic Chemistry, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Japan 464

Marilyn Emerson Holtzer and Alfred Holtzer*

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

Received November 2, 1984

ABSTRACT: Light scattering experiments on solutions of α-tropomyosin in benign buffer near neutral pH are reported as a function of temperature. The Rayleigh ratio for all cases is almost independent of scattering angle, as expected for optically clear solutions whose constituent particles have a greatest dimension barely greater than one-tenth the wavelength. Terms in the virial expansion beyond the first are shown to be negligible, and the absolute value of the relevant function of the zero-angle excess Rayleigh ratio, $(Kc/R_0)^{-1}$, at 20 °C agrees satisfactorily with the known molecular weight of the two-chain, coiled-coil, native molecule. At the highest temperatures (≥60 °C), the molecular weight is half that value, indicating dissociation into two, separate polypeptide chains. Comparison of the full thermal course of weight-average molecular weight with studies of helix content (by circular dichroism) indicates that chain dissociation and cooperative loss of helix occur in the same temperature domain. Thus, it is likely that the two processes are closely coupled.

The native tropomyosin molecule comprises two right-handed α-helical polypeptide chains set side-by-side in parallel and in register and given a slight, left-handed supertwist, a structure that can be called a two-chain, coiled coil.[1] In the absence of interchain cross-links, the molecule is supposedly converted at elevated temperature to two separated chains of very low helix content, i.e., to

† We dedicate this paper to Prof. Paul Doty in this year of his 65th birthday.

0024-9297/85/2218-1083$01.50/0