Theory of the Helix–Coil Transition in Doubly Cross-Linked, Two-Chain, Coiled Coils. A Globular Protein Model

Jeffrey Skolnick†

Department of Chemistry, Washington University, St. Louis, Missouri 63130.
Received July 20, 1985

ABSTRACT: A statistical mechanical theory of the α-helix-to-random coil transition in doubly cross-linked, two-chain, coiled coils has been developed in the context of two models for which expressions are given for the partition function, overall helix content, and helix probability profiles. The more restrictive, "all-or-none" (AN) model a priori asserts that the net effect of loop entropy is so prohibitive that either the molecule contains a pair of interacting helices that propagates at a minimum between the two cross-linked pairs of α-helical turns (blocks) or there are no interacting helices present. If we identify the interacting helical conformation as the native state and the conformations lacking any interacting α-helices as the denatured state, the AN model bears a strong resemblance to the two-state model of globular protein folding. We also develop the more general, doubly cross-linked, interior eyelet model that allows, in addition to those conformational states permitted in the AN model, constrained random coil loops plus interacting α-helices between cross-linked blocks. Application is made to doubly cross-linked, homopolymeric, coiled coils, and a comparison with the helix–coil transition in singly cross-linked, coiled coils is presented. Furthermore, over a wide set of conditions the range of validity of the AN model is assessed. On the basis of the robustness of the AN model, we conclude that doubly cross-linked, coiled coils possess many of the qualitative features of the equilibrium globular protein folding process. In particular, they demonstrate the role of both "short- and long-range" interactions in determining the stability of the native state and graphically point out the crucial role played by topological constraints in determining the character of the native → denatured transition.

I. Introduction

Due to the structural complexity of globular proteins, it is very difficult to determine the relative importance of the various kinds of interactions that ultimately produce the equilibrium native conformation. A globular protein may possess regions of α-helices, β-sheets, loops, and bends juxtaposed in a complicated three-dimensional arrangement; as a consequence, the development of even the equilibrium theory of their native → denatured transition or vice versa is a difficult task, beyond existing capabilities. What are clearly required, then, are simpler model systems, amenable to both experimental and theoretical examination, that possess a sufficient number of representative features of globular proteins that their study is relevant and timely. For some time now, along with Holtzer and Hodges et al., we have argued that α-helical, two-chain, coiled coils are just such a class of molecules.[1-3] From a structural point of view they are quite simple,[4-11] consisting of two parallel, α-helical chains wound around each other with a slight supertwist. Furthermore, a salient member of the two-chain, coiled coils is the muscle regulatory protein tropomyosin, whose primary sequence possesses a quasi-repeating heptet, designated by the letters a–g.[11-14] Positions a and d tend to be occupied by hydrophobic residues, and position e (g) tends to be occupied by anionic (cationic) residues. When wound up into a coiled coil, the hydrophobic streaks of the two chains are in contact, and the possibility of salt bridge formation also exists, thereby allowing one to assess the relative importance of hydrophobic vs. electrostatic interactions. Moreover, as has been pointed out by Hodges and co-workers, tropomyosin is but one of a substantial number of proteins of biological importance that possess a coiled coil structure and/or a quasi-repeating heptet in portions of their primary sequence.[3] Thus, we have embarked on the development of a physical/statistical mechanical theory of the helix–coil transition in coiled coils; in the present paper, we extend the theory to doubly cross-linked molecules.

In a series of papers,[15-19] the theory of the helix–coil transition in non-cross-linked and singly cross-linked, two-chain, coiled coils has been developed in the spirit of single-chain, helix–coil theory.[20] That is, the helix–coil transition is fundamentally viewed to be of a continuous rather than an all-or-none nature. In such a system, "short-range" interactions characterizing the intrinsic stability of the helical as compared to the random coil state of a residue in an isolated single chain are accounted for by the Zimm–Bragg helix initiation parameter, σ, and propagation parameter s.[21] Consideration of the "long-range" interactions between the two helices gives rise to an interchain helix initiation parameter u_ϕ and a helix–helix propagation parameter w.[2] Physically u_ϕ is equal to the effective volume in configuration space accessible to an α-helical turn (block) in the interacting configuration when the α-helical turn in the neighboring chain is held fixed, and $-kT \ln w$ is the free energy of a pair of positionally fixed and interacting α-helical turns relative to the same pair of positionally fixed, but noninteracting α-helical turns. If the coiled coil is non-cross-linked, then in addition to the Zimm–Bragg parameters, u_ϕ and w enter directly into the calculation of the mean overall helix content. If, however, the coiled coil is singly cross-linked, then the relevant interchain parameters are w and r_ϕ,[18] with the latter equal to the ratio of the effective volume accessible to one of the two interacting, helical cross-linked blocks when the other is held fixed relative to that in all noninteracting states and includes the difference in configurational entropy of the cross-link itself between the two kinds of conformations.

The theory of non-cross-linked two-chain, coiled coils includes both the effects of loop entropy and the possibility of mismatched association of the two chains in the loops-excluded, imperfect matching model, LEIM.[17] Application of the LEIM model to rabbit α-tropomyosin has been quite successful.[22] A single algorithm for $w(T)$ extracted from circular dichroism measurements gives a reasonably good fit to the experimentally determined helix content vs. temperature curves at near-neutral and acidic pH. Thus the LEIM model passes the test of self-consistency. Moreover, with the same $w(T)$, the calculated weight fraction of single chains vs. temperature is in accord with light scattering data[23] and also gives a fraction of chains that are essentially in-register at low temperature

† Alfred P. Sloan Foundation Fellow.

0024-9297/86/2219-1153$01.50/0

and near-neutral pH that is consistent with the experimentally observed high fraction of chains that can be cross-linked under these conditions.[24] The latter two comparisons are significant in that they provide extrinsic tests of the validity of the theoretical approach.

More recently, buoyed by the success of the theory as applied to non-cross-linked coiled coils (dimers), we extended the theory to singly cross-linked coiled coils, initially a priori assumed to be in-register[18] and later extended to include mismatched states.[19] The important qualitative conclusions that emerged from these studies are the following. For singly cross-linked, coiled coils of short-to-moderate length, because of the effect of loop entropy,[25-27] constrained random coil loops originating at the cross-link and terminating on an interacting pair of helices have negligible statistical weight. Furthermore, since the presence of out-of-register, interacting helical states in a cross-linked molecule by necessity requires constrained random coil loops, mismatched configurations are unimportant throughout the entire course of the helix–coil transition. Thus, in a cross-linked dimer in an interacting configuration, there is just a single interacting helical stretch that must include the cross-linked pair of blocks; these states are in equilibrium with the noninteracting conformation of the two chains. This is to be contrasted with the case of a non-cross-linked dimer, where the single interacting helical stretch can occur anywhere and where mismatched configurations of the chains can be important.

If, as discussed above, one cross-link modifies the character of the helix–coil transition in two-chain, coiled coils, it is important to inquire about the effect of a second cross-link. In addition, since there exists a genetic variant of tropomyosin, rabbit β-tropomyosin, that possesses cross-linkable cysteines at residues 36 and 190,[28,29] consideration of the doubly cross-linked, coiled coil is of practical interest as well. As pointed out long ago by Poland and Scheraga,[30] loop entropy can change the character of the helix–coil transition in short, closed random coil loops that may contain noninteracting helices by destabilizing the random coil states. Perhaps, however, a far more telling example of the influence of loop entropy on conformational stability is the elegant study by Lin, Konishi, Denton, and Scheraga on the thermal stability of RNase A.[31] They introduced an extrinsic cross-link between Lys-7 and Lys-41 and found that the cross-linked protein exhibits a reversible thermal denaturation profile whose midpoint is 25 °C higher than that in the non-cross-linked protein. They interpret this result as implying that the cross-link probably destabilizes the denatured conformation by reducing the chain entropy. Clearly, then, an analysis of the model system of doubly cross-linked, two-chain, coiled coils should be of interest.

Because of the effect of loop entropy, the presence of a second cross-link in a two-chain, coiled coil should destabilize random coil states relative to the fully helical state between cross-links and thus should profoundly affect the character of the helix–coil transition. Among the questions we investigate in this paper are the following: What qualitative differences does the theory predict between the helix–coil transition in doubly and singly cross-linked and non-cross-linked coiled coils? In doubly cross-linked chains, what is the effect of variation in cross-link location, chain length, and helix initiation parameter on the character of the helix–coil transition? Finally, under what conditions, if any, is the all-or-none model, in which the molecule either is fully helical and interacting between cross-linked blocks or lacks any pairs of interacting helices, valid? When the all-or-none model holds, the helix–coil transition behaves in a certain sense, amplified below, like the two-state model of globular proteins.[25,32-35] In the following sections, in the context of the AN model and the more general, doubly cross-linked interior eyelet model (DCIEM), which includes loops containing constrained random coils and α-helices, we hope to address the above questions.

Due to the topological constrains introduced by a second cross-link, we cannot a priori neglect interior random coil states occurring between cross-links. Indeed, at very low helix content, where the probability of having an interacting helical stretch is small, such constrained random coil states will dominate the population. The DCIEM explictly accounts for such constrained random coil states as well as constrained random coil and interacting helical states described below in greater detail. Thus, the DCIEM is a generalization for doubly cross-linked dimers of the "interior eyelet" model developed in previous papers to treat singly cross-linked chains.[18,19]

To incorporate the quasi-repeating heptet structure of the primary sequence[11-14] into the theory, each chain is divided into blocks, the ith block of which contains m_i residues.[36] For a further discussion of this coarse graining procedure, we refer to previous work.[2] Finally, to account for the geometric constraints due to fixed bond lengths, bond angles, and rotational potentials present in a real protein, we assume that residues in the cross-linked block pair and participating in the constrained random loop are completely randomly coiled.

The outline of the rest of this paper is as follows. Section II develops the formalism required to treat the helix–coil transition in doubly cross-linked, coiled coils; here the emphasis is on a qualitative understanding of the features of the model. Technical details, as much as possible, are relegated to Supplementary Material, Appendices A–D. Expressions are presented for the partition function, overall helix content, and helix probability profiles. Then in section III, we present illustrative calculations on hypothetical, homopolymeric, doubly cross-linked, two-chain, coiled coils; comparison is made with singly cross-linked dimers. Finally, section IV summarizes the conclusions of this paper and suggests several avenues of future work.

II. Theory

A. General Considerations. In what follows, the theory of the helix–coil transition of a doubly cross-linked, heteropolymeric, two-chain, coiled coil, in the perfect matching limit, is developed. We assume for convenience that both chains are identical; extension to nonidentical chains is straightforward. Each chain is assumed to contain N_T residues divided into N_B blocks in which block i contains m_i residues. Block 1 contains the N-terminal residue, and block N_B contains the C-terminal residue. For descriptive purposes, the N terminus (C terminus) shall be referred to as the left (right) part of the molecule. The two chains contain cross-linked residues in block N_{C1} (N_{C2}), in which there are $m_{l1} - 1$ ($m_{l2} - 1$) residues preceding the cross-link and $m_{r1} - 1$ ($m_{r2} - 1$) residues after the cross-link in block N_{C1} (N_{C2}). Finally, let l_c be the ratio of the length of the cross-link to the virtual bond length (=3.80 Å)[37] of an amino acid residue; in calculations presented in section III l_c is set equal to 2.

In the following subsections, we divide the calculation into parts depending on the conformational state of the cross-linked blocks and whether there are interacting helical stretches present in the subpopulation of interest. In each subsection, expressions are given for the partition function, helix content, and helix probability profiles appropriate to the subpopulation of interest. We shall begin

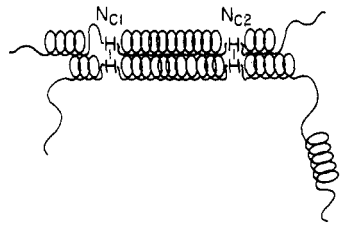

Figure 1. Schematic representation of the allowed conformations in H=H type states.

in section IIB with the simplest case, that is, where all four cross-linked blocks are helical and part of the fully helical interacting stretch that propagates between cross-links; such states are designated H=H (see Figure 1). These are the only interacting helical states that would be important if constrained loops containing both interacting helices and interior random coils between cross-linked blocks are prohibited. We then consider in section IIC states designated schematically as CHC, in which there is at least one interacting pair of helical blocks per molecule and in which one or more of the cross-linked blocks are randomly coiled (see Figure 2). Section IID examines the subpopulation of states labeled HCH, in which both pairs of cross-linked blocks are fully helical and interacting and contain constrained random coil loops between the cross-linked blocks (see Figure 3). Section IIE treats the subpopulation of doubly cross-linked molecules lacking any interacting helical states, designated X≈X states, with X representing any conformation (either coil or helix) of the cross-linked pair of blocks (see Figure 4). Sections IIB through IIE allow us to examine the relative contribution of the subpopulations described above to the total partition function, overall helix content, and helix probability profiles, expressions for which are presented in section IIF.

B. H=H Type States. In the following, we consider H=H type states, that is, states where the interacting helical stretch propagates between the two cross-linked blocks; a schematic representative is shown in Figure 1, where we denote the helical state of a cross-linked block by H. The interacting helical stretch may propagate beyond a cross-linked block, as occurs in Figure 1 for both chains to the right of block N_{C2}, or may be helical and noninteracting, as is present in the lower chain to the left of block pair N_{C1}. Finally, random coil states of the type indicated in the figure are also allowed.

In the subsequent development, we will need the statistical weights of the conformational states not associated with constrained random coil loops. Let H(C) be a completely helical (random coil) block, and [H]C ([C]H) be an interfacial helix–coil (coil–helix) block. The statistical weights appropriate to block i in an isolated single chain are

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

The internal partition function of H=H type states follows from

$$Z_{hh} = \mathbf{J}^* \prod_{i=1}^{N_{C1}-1} \mathbf{U}_{hh,i}\mathbf{U}_{hh,N_{C1}} \prod_{i=N_{C2}+1}^{N_B} \mathbf{U}_{hhi}\mathbf{J} \qquad \text{(II-2)}$$

with $\mathbf{J}^*$ a row vector of dimension twelve composed of one followed by eleven zeroes, $\mathbf{J}$ a column vector of four zeros followed by eight ones, and $\mathbf{U}_{hhi}$ a partitioned 12 × 12 matrix of the form, if $i < N_{C1}$

$$\mathbf{U}_{hh,i} = \begin{bmatrix} \mathbf{U}_d & \mathbf{U}_{CH} & \mathbf{0}_4 \\ \mathbf{0}_4 & \mathbf{U}_{HH} & \mathbf{0}_4 \\ \mathbf{0}_4 & \mathbf{0}_4 & \mathbf{0}_4 \end{bmatrix}_i \qquad \text{(II-3)}$$

wherein $\mathbf{U}_d$, $\mathbf{U}_{CH}$, and $\mathbf{U}_{HH}$ are the standard 4 × 4 noninteracting, initiating, and propagating statistical weight matrices defined previously in eq II-3–II-5 of ref 16. In what follows, $\mathbf{O}_n$ is a null $n \times n$ matrix. Physically, eq II-3 ensures that the single interacting helical stretch does not terminate until at least block pair N_{C2}.

Setting $i = N_{C1}$, we have

$$\mathbf{U}_{hh,N_{C1}} = \begin{bmatrix} \mathbf{0}_4 & \mathbf{U}_{CHH} & \mathbf{0}_4 \\ \mathbf{0}_4 & \mathbf{U}_{HHH} & \mathbf{0}_4 \\ \mathbf{0}_4 & \mathbf{0}_4 & \mathbf{0}_4 \end{bmatrix}_{N_{C1}} \qquad \text{(II-4a)}$$

in which

$$\mathbf{U}_{CHH,N_{C1}} = \begin{bmatrix} & 0 & \tau^2_{N_{C1}} w \prod_{i=N_{C1}+1}^{N_{C2}} SM_i^2 w \\ \mathbf{0}_2 & \vdots & \\ & 0 & \tau_{N_{C1}} SM_{N_{C1}} w \prod_{i=N_{C1}+1}^{N_{C2}} SM_i^2 w \\ \text{---} & \vdots & \\ & 0 & \tau_{N_{C1}} SM_{N_{C1}} w \prod_{i=N_{C1}+1}^{N_{C2}} SM_i^2 w \\ \mathbf{0}_2 & \vdots & \\ & 0 & 0 \end{bmatrix}_{N_{C1}} \qquad \text{(II-4b)}$$

wherein, since for this subpopulation by hypothesis there are no constrained interior random coils to the right of the cross-link, we have to modify τ_i of eq II-1 by

$$\tau_{N_{C1}} = \sum_{j=1}^{ml_1} \sigma_j \prod_{k=j}^{m_{N_{C1}}} s_k \qquad \text{(II-4c)}$$

and

$$\mathbf{U}_{HHH,N_{C1}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & \prod_{i=N_{C1}}^{N_{C2}} SM_i^2 w \end{bmatrix}_{N_{C1}} \qquad \text{(II-4d)}$$

$\mathbf{U}_{hh,N_{C1}}$, defined above in eq II-4a, generates the completely interacting helical stretch propagating from at least block pair N_{C1} to block pair N_{C2}.

Finally, if $i > N_{C2}$

$$\mathbf{U}_{hh,i} = \begin{bmatrix} \mathbf{0}_4 & \mathbf{0}_4 & \mathbf{0}_4 \\ \mathbf{0}_4 & \mathbf{U}_{HH} & \mathbf{U}_{HC} \\ \mathbf{0}_4 & \mathbf{0}_4 & \mathbf{U}_d \end{bmatrix}_i \qquad \text{(II-5)}$$

where $\mathbf{U}_{HC}$ is the 4 × 4 terminating matrix of the interacting helix stretch and is given by eq II-6 of ref 16.

Overall Helix Content. The overall helix content for the H=H subpopulation is calculated by employing

$$f_{hh} = \mathbf{J}_s^* \prod_{i=1}^{N_{C1}-1} \mathbf{A}_{hh,i}\mathbf{A}_{hh,N_{C1}} \prod_{i=N_{C2}+1}^{N_B} \mathbf{A}_{hh,i}\mathbf{J}_s/(N_T Z_{hh}) \qquad \text{(II-6a)}$$

with $\mathbf{J}_s^*$ a row vector consisting of unity followed by twenty-three zeroes and $\mathbf{J}_s$ a column vector composed of sixteen zeroes followed by eight ones. $\mathbf{A}_{hh,i}$ is a partitioned

24 × 24 matrix of the form

$$\mathbf{A}_{hh,i} = \begin{bmatrix} \mathbf{U}_{hh,i} & \mathbf{U}'_{hh,i} \\ \mathbf{0}_{12} & \mathbf{U}_{hh,i} \end{bmatrix} \tag{II-6b}$$

with $\mathbf{U}_{hh,i}$ defined in eq II-3–II-5 for $i < N_{C1}$, $i = N_{C1}$, and $i > N_{C2}$, respectively.

If $i < N_{C1}$ or if $i > N_{C2}$

$$\mathbf{U}'_{hh,i} = \sum_{k=1}^{m_i} \frac{\partial \mathbf{U}_{hh,i}}{\partial \ln s_{ik}} \tag{II-7}$$

wherein s_{ik} is the helix propagation parameter of the kth residue in block i, and we use the notation of Flory[38] in evaluating the matrix derivative.

Since $\mathbf{U}_{hh,N_{C1}}$ defined in eq II-4a contains the statistical weight of the fully helical sequence between cross-links, we have to count all such states in $\mathbf{U}'_{hh,N_{C1}}$. Namely

$$\mathbf{U}'_{hh,N_{C1}} = \sum_{i=N_{C1}}^{N_{C2}} \sum_{k=1}^{m_i} \frac{\partial \mathbf{U}_{hh,N_{C1}}}{\partial \ln s_{ik}} \tag{II-8}$$

Helix Probability Profiles. The mean helix content of the jth block, $f_{hh}(j)$, where $j < N_{C1}$, is readily obtained from

$$f_{hh}(j) = \mathbf{J}^* \prod_{i=1}^{j-1} \mathbf{U}_{hh,i} \mathbf{U}'_{hh,j} \prod_{i=j+1}^{N_{C1}} \mathbf{U}_{hh,i} \prod_{i=N_{C2}+1}^{N_B} \mathbf{U}_{hh,i} \mathbf{J} / (m_j Z_{hh}) \tag{II-9}$$

and $\mathbf{J}^*$ and $\mathbf{J}$ are defined immediately following eq II-2.

Analogously, if $j > N_{C2}$

$$f_{hh}(j) = \mathbf{J}^* \prod_{i=1}^{N_{C1}} \mathbf{U}_{hh,i} \prod_{i=N_{C2}+1}^{j-1} \mathbf{U}_{hh,i} \mathbf{U}'_{hh,j} \prod_{i=j+1}^{N_B} \mathbf{U}_{hh,i} \mathbf{J} / (m_j Z_{hh}) \tag{II-10}$$

In both eq II-9 and II-10, $\mathbf{U}_{hh,i}$ is defined in eq II-3–II-5, and $\mathbf{U}'_{hh,i}$ is defined in eq II-7.

Finally, if $N_{C1} \leq j \leq N_{C2}$

$$f_{hh}(j) = \mathbf{J}^* \prod_{i=1}^{N_{C1}-1} \mathbf{U}_{hh,i} \mathbf{U}'_{hh,N_{C1}j} \prod_{i=N_{C2}+1}^{N_B} \mathbf{U}_{hh,i} \mathbf{J} / (m_j Z_{hh}) \tag{II-11a}$$

where we have

$$\mathbf{U}_{hh,N_{C1}j} = \sum_{k=1}^{m_j} \frac{\partial \mathbf{U}_{hh,N_{C1}}}{\partial \ln s_{jk}} \tag{II-11b}$$

This completes the development of the formal expressions for the partition function, overall helix content, and the helix probability profiles of H=H type states.

C. CHC Type States. The subpopulation of CHC type states contains at least one randomly coiled, cross-linked block and at least one (and perhaps up to three) pair of interacting helices. Typical kinds of conformational states generated by the serial matrix product expression developed below for blocks $i \leq N_{C1}$, $N_{C1} \leq i \leq N_{C2}$, and $i > N_{C2}$ are shown in Figure 2, parts A–C, respectively. Those states in Figure 2A lacking an interacting pair of helices are followed by states in Figure 2, parts B or C, that have such an interacting helical stretch. Similar considerations hold for Figure 2, parts B and C.

On the basis of results from singly cross-linked chains, we would expect those states in Figure 2 part A (**1a** and **2a**) and part C (**12a** and **13a**) that have a constrained random coil loop not between the pair of cross-links to have negligible weight relative to the configurations having unconstrained random coil loops. However, for completeness they are included.

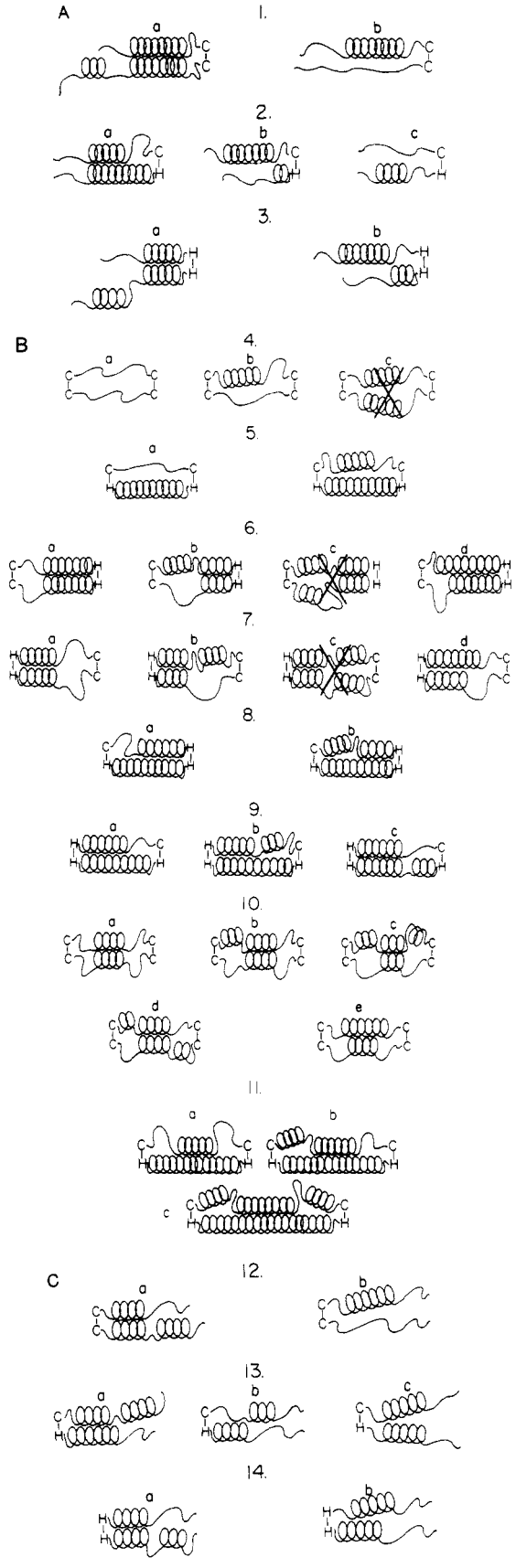

Figure 2. (A) Representative CHC conformations to the left of cross-linked block N_{C1}. (B) Representative CHC conformations between cross-linked blocks N_{C1} and N_{C2}. (C) Representative CHC conformations to the right of cross-linked block N_{C2}.

We have neglected CHC states having two noninteracting helices in constrained random coil loops; these are

the "crossed out" conformations **4c**, **6c**, and **7c**. Even for σ on the order of 10^{-2}, such constrained loops containing two noninteracting helices are of negligible statistical weight relative to the case where one helical stretch is transformed to the completely randomly coiled state, as in the second set of conformations depicted in **6b** and **7b**.

Partition Function. The internal partition function for CHC type states is given by

$$Z_{chc} = \mathbf{J}^* \prod_{i=1}^{N_B} \mathbf{U}_{CHC,i} \mathbf{J} - Z_{hh} \qquad \text{(II-12)}$$

where $\mathbf{J}^*$ and $\mathbf{J}$ are row and column vectors defined immediately below eq II-2 and $\mathbf{U}_{CHC}$ is a partitioned 12×12 matrix of the form given below. The Z_{hh} term removes states containing the fully helical, interacting helical stretch between cross-linked blocks. These states are members of the H=H subpopulation.

If $i < N_{C1}$

$$\mathbf{U}_{CHC,i} = \begin{bmatrix} \mathbf{U}_d & \mathbf{U}_{CH} & \mathbf{0}_4 \\ \mathbf{0}_4 & \mathbf{U}_{HH} & \mathbf{U}^{\ell}_{HC} \\ \mathbf{0}_4 & \mathbf{0}_4 & \mathbf{E}_4 \end{bmatrix}_i \qquad \text{(II-13a)}$$

wherein $\mathbf{U}_d$, $\mathbf{U}_{CH}$, and $\mathbf{U}_{HH}$ are defined in eq II-3–II-5 of ref 16, $\mathbf{E}_4$ is the 4×4 identity matrix, and

$$\mathbf{U}^{\ell}_{HC} = \left[\begin{array}{cc|cc} & \mathbf{0}_2 & & \mathbf{0}_2 \\ \hline 0 & 0 & 0 & 0 \\ \mathcal{S}^2_{\ell} & (SM\mathcal{S})_{\ell} & (SM\mathcal{S})_{\ell} & 0 \end{array}\right]_i \qquad \text{(II-13b)}$$

where $\mathcal{S}_l^2$ and $(SM\mathcal{S})_l$ are defined in Appendix A, eq A-1 and A-2. For $i < N_{C1}$, $\mathbf{U}_{CHC,i}$ is identical with the statistical weight matrix, $\mathbf{U}_{se}$, to the left of a cross-link in a singly cross-linked, in-register chain in the "interior eyelet" model, defined in eq II-22a of ref 18.

Setting $i = N_{C1}$, we have

$$\mathbf{U}_{CHC,N_{C1}} = \begin{bmatrix} \mathbf{E}_4 & \mathbf{U}_{CH} & \mathbf{0}_4 \\ \mathbf{U}^{\ell}_{HC} & \mathbf{U}_{HH} & \mathbf{U}^{\ell}_{HC} \\ \mathbf{E}_4 & \mathbf{0}_4 & \mathbf{0}_4 \end{bmatrix}_{N_{C1}} \qquad \text{(II-14a)}$$

with τ_i in $\mathbf{U}_{CH,i}$ of eq II-4 of ref 16 replaced by

$$\tau_{N_C} = \sum_{j=1}^{m_{l1}} \sigma_j \prod_{k=j}^{m_{N_{C1}}} s_k \qquad \text{(II-14b)}$$

and

$$\mathbf{U}^{\ell k}_{HC,N_{C1}} = \left[\begin{array}{cc|cc} & \mathbf{0}_2 & & \mathbf{0}_2 \\ \hline 0 & 0 & 0 & 0 \\ \mathcal{S}^{2k}{}_{N_{C1}} & (SM\mathcal{S})^k{}_{N_{C1}} & (SM\mathcal{S})^k{}_{N_{C1}} & 0 \end{array}\right]_{N_{C1}} \qquad \text{(II-14c)}$$

Here $\mathcal{S}^{2k}{}_{N_{C1}}$ and $(SM\mathcal{S})^k{}_{N_{C1}}$ are given in eq A-3 and A-4 for $k = 1$ and 2, respectively, and sum over all noninteracting helical configurations as depicted in Figure 2B, structures **4a,b** and **5a,b**.

If $N_{C1} < i < N_{C2}$

$$\mathbf{U}_{CHC,i} = \begin{bmatrix} \mathbf{E}_4 & \mathbf{U}^{I}_{CH} & \mathbf{0}_4 \\ \mathbf{0}_4 & \mathbf{U}_{HH} & \mathbf{U}^{I}_{HC} \\ \mathbf{0}_4 & \mathbf{0}_4 & \mathbf{E}_4 \end{bmatrix}_i \qquad \text{(II-15a)}$$

in which

$$\mathbf{U}^{I}_{CH,i} = \left[\begin{array}{c|cc} \mathbf{0}_2 & 0 & \tau^2{}_{I,i} w \\ & 0 & SM_{N_{C1}} (\tau SM)_{I,i} w \\ \hline \mathbf{0}_2 & 0 & SM_{N_{C1}} (\tau SM)_{I,i} w \\ & 0 & 0 \end{array}\right]_i \qquad \text{(II-15b)}$$

$\tau^2{}_{I,i}$ and $(\tau SM)_{I,i}$ are defined explicitly in eq A-5 and A-6, and

$$\mathbf{U}^{I}_{HC,i} = \left[\begin{array}{cc|cc} & \mathbf{0}_2 & & \mathbf{0}_2 \\ \hline 0 & 0 & 0 & 0 \\ \mathcal{S}^2{}_I & (SM\mathcal{S})_I & (SM\mathcal{S})_I & 0 \end{array}\right]_i \qquad \text{(II-15c)}$$

$\mathcal{S}^2{}_I$ and $(SM\mathcal{S})_I$ are given in eq A-7 and A-8.

Representative conformations generated by the matrix product of eq II-15a are depicted schematically in Figure 2B in the non-crossed-out structures of **6**–**11**.

If we set $i = N_{C2}$, then

$$\mathbf{U}_{CHC,N_{C2}} = \begin{bmatrix} \mathbf{0}_4 & \mathbf{U}^{I}_{CH} & \mathbf{0}_4 \\ \mathbf{U}^{\ell_2}{}_{HC,N_{C2}} & \mathbf{U}_{HH} & \mathbf{U}^{\ell_2}{}_{HC,N_{C2}} \\ \mathbf{E}_4 & \mathbf{0}_4 & \mathbf{E}_4 \end{bmatrix}_{N_{C2}} \qquad \text{(II-16a)}$$

wherein the statistical weight matrices $\mathbf{U}_{HC,N_{C2}}$ and $\mathbf{E}_4$ in the lower left-hand corner of the first four columns of $\mathbf{U}_{CHC}$ allow for closed interior random coil loops preceding and following block N_{C2}. $\mathbf{U}_{HC,N_{C2}}$ is of the identical form as in eq II-15c, but $\mathcal{S}^2{}_{I,N_{C2}}$ and $(SM\mathcal{S})_{I,N_{C2}}$ are defined in eq A-9 and A-10, respectively, and account for the contribution of states in which at least one of the two members of block pair N_{C2} is randomly coiled and where both members of block pair $N_{C2} - 1$ are helical.

Furthermore, $\mathbf{U}^{I}_{CH,N_{C2}}$ is of the same form as eq II-15b but where we replace τ_i, in $\tau^2{}_{I,i}$ and $(\tau SM)_{I,i}$ of eq A-5 and A-6 by $\tau_{N_{C2}}$ wherein

$$\tau_{N_{C2}} = \sum_{j=1}^{ml_2} \sigma_j \prod_{k=j}^{N_{C2}} s_k \qquad \text{(II-16b)}$$

Finally, if $i > N_{C2}$

$$\mathbf{U}_{CHC,i} = \begin{bmatrix} \mathbf{E}_4 & \mathbf{U}^{r}_{CH} & \mathbf{0}_4 \\ \mathbf{0}_4 & \mathbf{U}_{HH} & \mathbf{U}_{HC} \\ \mathbf{0}_4 & \mathbf{0}_4 & \mathbf{U}_d \end{bmatrix}_i \qquad \text{(II-17a)}$$

with

$$\mathbf{U}^{r}_{CH,i} = \left[\begin{array}{c|cc} \mathbf{0}_2 & 0 & \tau^2{}_{\ell} w \\ & 0 & (\tau SM)_{\ell} w \\ \hline \mathbf{0}_2 & 0 & (\tau SM)_{\ell} w \\ & 0 & 0 \end{array}\right] \qquad \text{(II-17b)}$$

where $\tau^2{}_l$ and $(\tau SM)_l$ are defined in eq A-11 and A-12. $\mathbf{U}_{HC,i}$ of eq II-17a is the standard terminating matrix of an interacting helical stretch and is explicitly given in eq II-6 of ref 16.

Equation II-17a is identical with the statistical weight matrix $\mathbf{U}_{se}$ in the "interior eyelet" model of singly cross-linked chains for blocks following the cross-link; see eq II-25 of ref 18. This completes the construction of the partition function for CHC type states.

Overall Helix Content. The overall helix content for the CHC type states is calculated via

$$f_{chc} = \{\mathbf{J}_s^* \prod_{i=1}^{N_B} \mathbf{A}_{chc,i}\mathbf{J}_s - f_{hh}Z_{hh}N_T\}/(N_T Z_{chc}) \quad \text{(II-18)}$$

Wherein Z_{chc} is given in eq II-12 and $\mathbf{A}_{chc,i}$ is a partitioned 24 × 24 supermatrix of the form

$$\mathbf{A}_{chc,i} = \begin{bmatrix} \mathbf{U}_{CHC,i} & \mathbf{U}'_{CHC,i} \\ \mathbf{0}_{12} & \mathbf{U}_{CHC,i} \end{bmatrix} \quad \text{(II-19)}$$

$\mathbf{U}_{CHC,i}$ is defined in eq II-13a, II-14a, II-15a, II-16a, or II-17a for $i < N_{C1}$, $i = N_{C1}$, $N_{C1} < i < N_{C2}$, $i = N_{C2}$, or $i > N_{C2}$, respectively. The second class of terms removes the contribution of H=H states generated by the $\mathbf{A}_{chc,i}$ matrix product. When $i < N_{C1}$, we have

$$\mathbf{U}'_{CHC,i} = \sum_{j=i}^{N_{C1}-1} \sum_{k=1}^{m_j} \frac{\partial \mathbf{U}_{CHC,i}}{\partial \ln s_{jk}} \quad \text{(II-20)}$$

i.e., we have to count the contribution to the helix content of the **2a** sequences shown in Figure 2A that have a helical stretch propagating from block i to block N_{C1}.

If $N_{C1} \le i \le N_{C2}$, the statistical weight matrices count helical states to the left of, equal to, and to the right of block i, so that we require the total contribution of all such helical states. Thus

$$\mathbf{U}'_{CHC,i} = \sum_{j=N_{C1}}^{N_{C2}} \sum_{k=1}^{m_j} \frac{\partial \mathbf{U}_{CHC,i}}{\partial \ln s_{jk}} \quad \text{(II-21)}$$

Finally, if $i > N_{C2}$

$$\mathbf{U}'_{CHC,i} = \sum_{j=N_{C2}+1}^{i} \sum_{k=1}^{m_j} \frac{\partial \mathbf{U}_{CHC,i}}{\partial \ln s_{jk}} \quad \text{(II-22)}$$

Namely, we must count the contribution of helical sequences such as are shown in Figure 2c, type **13a** that start at block N_{C2} and propagate to block i. This completes the development of the formal expressions for the overall helix content of CHC type states.

Helix Probability Profiles. The helix content of the jth block can be obtained via a supermatrix expression of the form

$$f_{chc}(j) = \{\mathbf{J}_s^* \prod_{i=1}^{N_B} \mathbf{A}_{chc,ij}\mathbf{J}_s - f_{hh}(j)m_j Z_{hh}\}/m_j Z_{chc} \quad \text{(II-23)}$$

in which Z_{chc} is given in eq II-12 and $\mathbf{A}_{chc,ij}$ is a 24 × 24 supermatrix

$$\mathbf{A}_{chc,ij} = \begin{bmatrix} \mathbf{U}_{CHC,i} & \mathbf{U}'_{CHC,ij} \\ \mathbf{0}_{12} & \mathbf{U}_{CHC,i} \end{bmatrix}_i \quad \text{(II-24a)}$$

where $\mathbf{U}_{CHC,i}$ for $i < N_{C1}$, $i = N_{C1}$, $N_{C1} < i < N_{C2}$, $i = N_{C2}$, and $i > N_{C2}$ is defined in eq II-13a, II-14a, II-15a, II-16a, or II-17a, respectively. $f_{hh}(j)$ and Z_{hh} are given in eq II-9–II-11b and II-2, respectively, and remove the H=H states generated by the matrix product of the $\mathbf{A}_{chc,ij}$. Finally

$$\mathbf{U}'_{CHC,ij} = \sum_{k=1}^{m_j} \frac{\partial \mathbf{U}_{CHC,i}}{\partial \ln s_{jk}} \quad \text{(II-24b)}$$

D. HCH Type States. In HCH type states, the residues in block pairs N_{C1} and N_{C2} are fully helical and interacting, and there is at least one randomly coiled block contained between the two cross-links. To the left of the cross-linked block N_{C1} and to the right of cross-linked block N_{C2}, the range of allowed conformations accessible to HCH type states is identical with H=H states; thus, there is at least one (and possibly up to three) interacting helical stretches per molecule. However, unlike H=H states, there are a wide range of constrained interior random coil states between cross-links; the kinds of states we have included are schematically depicted in Figure 3. Since these states involve at least one random coil block between helical cross-links, at lower values of the helix content, we would expect them to be less important than CHC type states. Similarly, at higher values of the helix content, the H=H states should dominate. As shown below in section III, this is indeed the case.

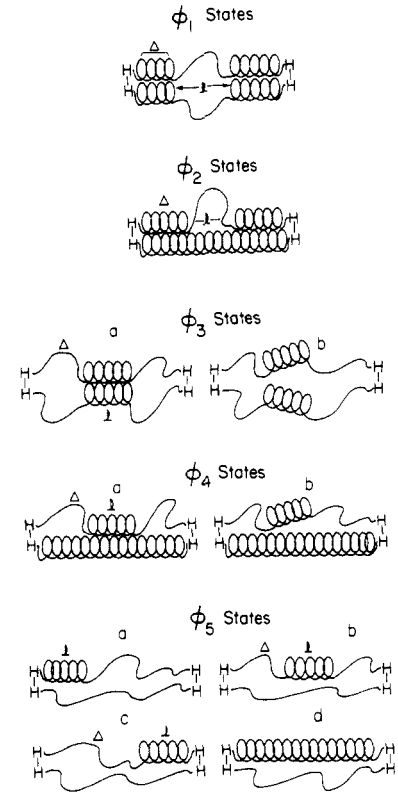

Figure 3. Schematic representation of the allowed conformations in HCH type states.

Partition Function. The internal partition function for HCH type states can be obtained from

$$Z_{hch} = \mathbf{J}^* \prod_{i=1}^{N_{C1}-1} \mathbf{U}_{hh,i}\mathbf{U}_{hch,N_{C1}} \prod_{i=N_{C2}+1}^{N_B} \mathbf{U}_{hh,i}\mathbf{J} \quad \text{(II-25)}$$

where $\mathbf{U}_{hh,i}$ is defined in eq II-3 for $i < N_{C1}$ and eq II-5 for $i > N_{C1}$. Furthermore, if $i = N_{C1}$

$$\mathbf{U}_{hch,N_{C1}} = \begin{bmatrix} \mathbf{0}_4 & \mathbf{U}_C & \mathbf{0}_4 \\ \mathbf{0}_4 & \mathbf{U}_H & \mathbf{0}_4 \\ \mathbf{0}_4 & \mathbf{0}_4 & \mathbf{0}_4 \end{bmatrix}_{N_{C1}} \quad \text{(II-26a)}$$

and sums over the constrained states between cross-links depicted in Figure 3. In eq II-26a

$$\mathbf{U}_{C,N_{C1}} = \left[\begin{array}{c|cc} \mathbf{0}_2 & 0 & \tau^2_{N_{C1}}\omega\phi \\ & 0 & \tau_{N_{C1}}SM_{N_{C1}}\omega\phi \\ \hline \mathbf{0}_2 & 0 & \tau_{N_{C1}}SM_{N_{C1}}\omega\phi \\ & 0 & 0 \end{array}\right]_{N_{C1}} \quad \text{(II-26b)}$$

with $\tau_{N_{C1}}$ given in eq II-4c and

$$\mathbf{U}_H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 0 & 0 \\ 0 & SM^2_{N_{C1}}\omega\phi \end{bmatrix}_{N_{C1}} \quad \text{(II-26c)}$$

ϕ is the total statistical weight associated with the constrained states in blocks $N_{C1} < i \leq N_{C2}$. In terms of the states depicted in Figure 3

$$\phi = \Phi_{12} + \phi_3 + \phi_4 + \phi_5 \qquad \text{(II-26d)}$$

Φ_{12} is defined in eq B-1 of Supplementary Materials and sums over ϕ_1 and ϕ_2 type states. Explicit expressions for ϕ_3, ϕ_4, and ϕ_5 may be found in eq B-5, B-6, and B-7. This completes the presentation of the partition function for HCH type states.

Overall Helix Content. The overall helix content for HCH type states follows from

$$f_{hch} = \{\mathbf{J}_s^* \prod_{i=1}^{N_{C1}-1} \mathbf{A}_{hh,i}\mathbf{A}_{hch,N_{C1}} \prod_{i=N_{C2}+1}^{N_B} \mathbf{A}_{hh,i}\mathbf{J}_s\}/N_T Z_{hch} \qquad \text{(II-27)}$$

where Z_{hch} is found in eq II-25, $\mathbf{A}_{hh,i}$ is a 24 × 24 supermatrix given in eq II-6b if $i < N_{C1}$ or $i > N_{C2}$. If $i = N_{C1}$, we have

$$\mathbf{A}_{hch,N_{C1}} = \begin{bmatrix} \mathbf{U}_{hch,N_{C1}} & \mathbf{U}'_{hch,N_{C1}} \\ \mathbf{0}_{12} & \mathbf{U}_{hch,N_{C1}} \end{bmatrix}_{N_{C1}} \qquad \text{(II-28a)}$$

wherein $\mathbf{U}_{hch,N_{C1}}$ is defined in eq II-26a and

$$\mathbf{U}'_{hch,N_{C1}} = \sum_{j=N_{C1}}^{N_{C2}} \sum_{k=1}^{m_j} \frac{\partial \mathbf{U}_{hch,N_{C1}}}{\partial \ln s_{jk}} \qquad \text{(II-28b)}$$

Helix Probability Profiles. The helix content of the jth block in the subpopulation of HCH type states, $f_{hch}(j)$ may be calculated as follows.

If $j < N_{C1}$

$$f_{hch}(j) = \mathbf{J}^* \prod_{i=1}^{j-1} \mathbf{U}_{hh,i}\mathbf{U}'_{hh,j} \prod_{i=j+1}^{N_{C1}-1} \mathbf{U}_{hh,i}\mathbf{U}_{hch,N_{C1}} \prod_{i=N_{C2}+1}^{N_B} \mathbf{U}_{hh,i}\mathbf{J}/m_j Z_{hch} \qquad \text{(II-29)}$$

wherein $\mathbf{U}_{hh,i}$ is given in eq II-3 and II-5 for $i < N_{C1}$ and $i > N_{C1}$, respectively. $\mathbf{U}'_{hh,j}$ is found in eq II-7, and $\mathbf{U}_{hch,N_{C1}}$ is defined in eq II-26a.

If $N_{C1} \leq j \leq N_{C2}$

$$f_{hch}(j) = \mathbf{J}^* \prod_{i=1}^{N_{C1}-1} \mathbf{U}_{hh,i}\mathbf{U}'_{hch,N_{C1}j} \prod_{i=N_{C2}+1}^{N_B} \mathbf{U}_{hh,i}\mathbf{J}/m_j Z_{hch} \qquad \text{(II-30a)}$$

where in addition to the $\mathbf{U}_{hh,i}$ given in eq II-3 and II-5 for $i \leq N_{C1}$ and $i > N_{C2}$, we have

$$\mathbf{U}'_{hch,N_{C1}j} = \sum_{k=1}^{m_j} \frac{\partial \mathbf{U}_{hch,N_{C1}}}{\partial \ln s_{jk}} \qquad \text{(II-30b)}$$

with $\mathbf{U}_{hch,N_{C1}}$ found in eq II-26a.

Finally, if $j > N_{C2}$

$$f_{hch}(j) = \mathbf{J}^* \prod_{i=1}^{N_{C1}-1} \mathbf{U}_{hh,i}\mathbf{U}_{hch,N_{C1}} \prod_{i=N_{C2}+1}^{j-1} \mathbf{U}_{hh,i}\mathbf{U}'_{hh,j} \prod_{i=j+1}^{N_B} \mathbf{U}_{hh,i}\mathbf{J}/m_j Z_{hch} \qquad \text{(II-31)}$$

for which $\mathbf{U}'_{hh,j}$ for $j > N_{C1}$ is given in eq II-7 and the relevant equations for the other statistical weight matrices are given above. This completes the formalism for HCH type states.

E. X≈X Type States. In X≈X type states, we consider the subpopulation of molecules without any interacting pairs of helices. Thus, while there may be helices occurring anywhere in the molecule, they are always in a noninteracting configuration; i.e., the side-by-side, essentially parallel arrangement of any two helices in adjacent chains

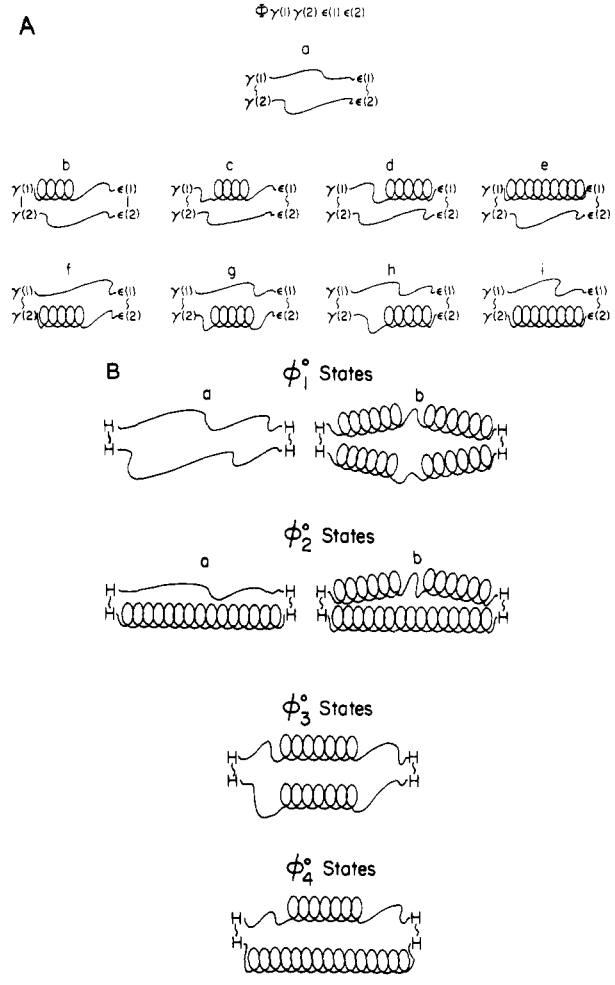

Figure 4. (A) Schematic representation of the allowed conformations in X≈X type states. (B) Schematic representation of additional allowed states that contribute to H(1)H(2)H(1)H(2) states in the X≈X subpopulation.

is excluded. Hence, prior to block N_{C1} and subsequent to block N_{C2}, the statistical weight matrix of the ith pair of blocks is simply that of the noninteracting states, namely $\mathbf{U}_{d,i}$ defined in eq II-3 of ref 16. The statistical weight of block N_{C1} is modified to reflect the fact that residues to the right of and including the cross-linked block in a [C]H state are entirely helical. Similarly, in an [H]C state, all the residues to the right of and including the cross-link are randomly coiled. There are 16 distinct conformations of the 4 cross-linked blocks in which $\gamma(1)$ ($\gamma(2)$) represents the conformation, either C or H, of the cross-linked block N_{C1} in chain one (two) and $\epsilon(1)$ ($\epsilon(2)$) represents the conformation of the cross-linked block N_{C2} in chain one (two). Between the two cross-links, there are a myriad of random coil and helical states; the ones we have included (and which, apart from the exceptions discussed below, make the dominant contribution to the partition function) are depicted schematically in Figure 4A for all configurations $\gamma(1)\gamma(2)$ $\epsilon(1)\epsilon(2)$. In addition to those states shown in Figure 4A, the fully helical conformation of all four cross-linked blocks also includes the noninteracting states shown in Figure 4B (such states must be included for consistency with the HCH conformations) and CCCC, CHCC, and CCCH type conformations also include the possibility of helical stretches in both chains; it is only in these states that multiple helical stretches between cross-linked blocks are important.

Partition Function. The internal partition function associated with X≈X type states can be expressed as

$$Z_{xx} = \text{row } (1, 0, 0, 0) \prod_{i=1}^{N_{C1}} \mathbf{U}_{d,i} \mathbf{U}_{X\approx X} \prod_{i=N_{C2}+1}^{N_B} \mathbf{U}_{d,i} \text{ col } (1, 1, 1, 1) \quad \text{(II-32)}$$

In the above equation, $\mathbf{U}_{d,i}$ is given by eq II-3 of ref 16. Furthermore, if $i = N_{C1}$, τ in $\mathbf{U}_{d,N_{C1}}$ is replaced by $\tau_{N_{C1}}$ defined previously in eq II-4c, and $\mathcal{S}_i$ in $\mathbf{U}_{d,N_{C1}}$ is replaced by

$$\mathcal{S}_{N_{C1}} = \sum_{d_1=1}^{ml_1-1} \prod_{k=1}^{d_1} s_k \quad \text{(II-33)}$$

$\tau_{N_{C1}}$ ($\mathcal{S}_{N_{C1}}$) state that if block N_{C1} is of the [C]H ([H]C) type, all the residues to the right of and including the cross-linked residue are helical (randomly coiled). This reflects the role of local topological constraints in enforcing helical and random coil conformations.

$\mathbf{U}_{X\approx X}$ is the 4 × 4 statistical weight matrix associated with the conformations of block pairs $N_{C1} + 1$ to N_{C2} inclusive and is of the form

$$\mathbf{U}_{X\approx X} = \begin{bmatrix} \phi_{cccc} & \phi_{ccch} & \phi_{cchc} & \phi_{cchh} \\ \phi_{chcc} & \phi_{chch} & \phi_{chhc} & \phi_{chhh} \\ \phi_{hccc} & \phi_{hcch} & \phi_{hchc} & \phi_{hchh} \\ \phi_{hhcc} & \phi_{hhch} & \phi_{hhhc} & \phi_{hhhh} \end{bmatrix} \quad \text{(II-34)}$$

wherein $\phi_{\gamma(1)\gamma(2)\epsilon(1)\epsilon(2)}$ is the statistical weight of block pairs $N_{C1} + 1$ to N_{C2} associated with conformation $\gamma(1)\gamma(2)$-$\epsilon(1)\epsilon(2)$ of the two cross-linked pairs of blocks. Explicit formulas for these terms are provided in Appendix C (Supplementary Materials). In the formulation of $\mathbf{U}_{X\approx X}$ care must be taken not to include interacting orientations associated with CHC type conformations in the matrix elements ϕ_{cchh}, ϕ_{hhcc}, ϕ_{chhh}, ϕ_{hchh}, ϕ_{hhch}, and ϕ_{hhhc} and with HCH type conformations in ϕ_{hhhh} states. All such details are attended to in Appendix C and need not concern us here.

Overall Helix Content. The overall helix content of the X≈X subpopulation, f_{xx}, may be obtained from

$$f_{xx} = \{\mathbf{J}^*_8 \prod_{i=1}^{N_{C1}} \mathbf{A}_{d,i} \mathbf{A}_{X\approx X} \prod_{i=N_{C2}+1}^{N_B} \mathbf{A}_{d,i} \mathbf{J}_8\}/N_T Z_{xx} \quad \text{(II-35)}$$

Here $\mathbf{J}^*_8$ is a row vector consisting of unity followed by 7 zeroes, $\mathbf{J}_8$ is a column vector composed of 4 zeroes followed by 4 ones, and $\mathbf{A}_d$ is an 8 × 8 supermatrix of the form

$$\mathbf{A}_{d,i} = \begin{bmatrix} \mathbf{U}_{d,i} & \mathbf{U}'_{d,i} \\ 0_4 & \mathbf{U}_{d,i} \end{bmatrix} \quad \text{(II-36)}$$

wherein $\mathbf{U}'_{d,i}$ is defined in eq II-11c of ref 16 for all $i \neq N_{C1}$. If $i = N_{C1}$, we replace τ_i and $\mathcal{S}_i$ by $\tau_{N_{C1}}$ and $\mathcal{S}_{N_{C1}}$ defined in eq II-4c and II-33, respectively, and $\tau'_{N_{C1}}$ and $\mathcal{S}'_i$ by

$$\tau'_{N_{C1}} = \sum_{j=1}^{ml_1} \sigma_j \prod_{k=j}^{m_{N_{C1}}} s_k(m - j + 1) \quad \text{(II-37a)}$$

and

$$\mathcal{S}'_{N_{C1}} = \sum_{d_1=0}^{ml_1-1} d_1 \prod_{k=1}^{d_1} s_k \quad \text{(II-37b)}$$

Moreover

$$\mathbf{A}_{X\approx X} = \begin{bmatrix} \mathbf{U}_{X\approx X} & \mathbf{U}'_{X\approx X} \\ 0_4 & \mathbf{U}_{X\approx X} \end{bmatrix} \quad \text{(II-38a)}$$

in which $\mathbf{U}_{X\approx X}$ has been defined in eq II-34 and

$$\mathbf{U}'_{X\approx X} = \sum_{j=N_{C1}+1}^{N_{C2}} \sum_{k=1}^{m_j} \frac{\partial \mathbf{U}_{X\approx X}}{\partial \ln s_{jk}} \quad \text{(II-38b)}$$

Helix Probability Profiles. The mean helix content of the jth block, $f_{xx}(j)$ is obtained from

$$f_{xx}(j) = \text{row } (1, 0, 0, 0) \times \prod_{i=1}^{j-1} \mathbf{U}_{d,i} \mathbf{U}'_{d,j} \prod_{i=j+1}^{N_{C1}} \mathbf{U}_{d,i} \mathbf{U}_{X\approx X} \prod_{i=N_{C2}+1}^{N_B} \mathbf{U}_{d,i} \times \text{col } (1, 1, 1, 1)/m_j Z_{xx} \quad \text{(II-39)}$$

if $i \leq N_{C1}$, with $\mathbf{U}'_{d,j}$ defined in eq II-11c of ref 16. Similarly if $N_{C2} \geq j \geq N_{C1} + 1$, we have

$$f_{xx}(j) = \text{row } (1, 0, 0, 0) \prod_{i=1}^{N_{C1}} \mathbf{U}_{d,i} \mathbf{U}'_{X\approx X,j} \prod_{i=N_{C2}+1}^{N_B} \mathbf{U}_{d,i} \times \text{col } (1, 1, 1, 1)/m_j Z_{xx} \quad \text{(II-40a)}$$

with

$$\mathbf{U}'_{X\approx X,j} = \sum_{k=1}^{m_j} \frac{\partial \mathbf{U}_{X\approx X}}{\partial \ln s_{jk}} \quad \text{(II-40b)}$$

with $\mathbf{U}_{X\approx X}$ given in eq II-34.

Finally if $j \geq N_{C2} + 1$

$$f_{xx}(j) = \text{row } (1, 0, 0, 0) \prod_{i=1}^{N_{C1}} \mathbf{U}_{d,i} \mathbf{U}_{X\approx X} \prod_{i=N_{C2}+1}^{j-1} \mathbf{U}_{d,i} \mathbf{U}'_{d,j} \times \prod_{i=j+1}^{N_B} \mathbf{U}_{d,i} \text{ col } (1, 1, 1, 1)/m_j Z_{xx} \quad \text{(II-41)}$$

Again, $\mathbf{U}'_{d,j}$ may be found in eq II-11c of ref 16. This completes the discussion of X≈X type states.

F. Overall Average Quantities. In sections II-B–II-E, we have derived expressions for the partition function, overall helix content, and helix probability profiles appropriate to the H=H, CHC, HCH, and X≈X subpopulations. This decomposition of the total conformational space accessible to a doubly cross-linked molecule allows us to assess the relative importance of a given subpopulation to the overall partition function of the molecule. In this section, we assembly these expressions to provide the total internal partition function, helix content, and helix probability profiles, Z_{2XL}, f_{2XL}, and $f_{2XL}(j)$, respectively. (The subscript 2XL stands for doubly cross-linked.) Furthermore, we calculate the expressions relevant to the "all-or-none" model of the helix–coil transition in doubly cross-linked two-chain, coiled coils; that is, either the molecule is fully helical and interacting between cross-links (i.e., in an H=H state) or lacks any interacting pairs of helices whatsoever (i.e., in an X≈X state). The H=H type states are analogous to the "native" state of a globular protein and the X≈X type states are analogous to the "denatured" form of a globular protein. The internal partition function, helix content, and helix probability profiles in the all-or-none model are Z_{AN}, f_{AN}, and $f_{AN}(j)$, respectively, where the subscript AN stands for all-or-none in the context discussed above. A schematic representation of the allowed states in the all-or-none model is found in Figure 5.

Unlike the situation when we considered the internal partition function and helix content appropriate to the various subpopulations, when we assembly the subpopulations we have to account for the reduction in configurational entropy attendant to the formation of interacting helical and cross-linked blocks. We adopt the following conventions.

1. If both cross-linked blocks are in the subpopulation of molecules lacking any interacting helical state whatsoever, (X≈X type states, see Figure 4), assign a statistical

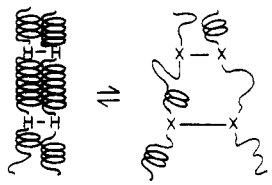

Figure 5. Schematic representation of the allowed states in the all-or-none model of the helix–coil transition in doubly cross-linked, coiled coils.

weight of unity for each cross-link. Thus the statistical weight associated with the cross-link itself, in the states described below, is normalized relative to the case having the greatest conformational degrees of freedom. While the statistical weight of the cross-link itself presumably may depend on the conformational states of the polypeptide backbone to which it is attached, we neglect this in the interest of keeping the number of adjustable (and unknown) parameters to an absolute minimum. However, the criterion of ring closure employed in Appendices C and D (Supplementary Material) assumes the random coils are next to an interacting pair of helices and possess the correct relative orientation that allows the interhelical interaction to occur. Thus, we must multiply the statistical weights of X≈X states calculated in section II-E by δ_Ω^{-1}, where δ_Ω is the ratio of the solid angle sustained by the interacting configurations of the two α-helices to 4π.

2. The statistical weight of a molecule in CHC type states (see Figure 2B, structures **6–9**) is assigned a statistical weight δ_{CHC}. An estimate of δ_{CHC} may be provided by calculating the relative probability, with respect to the noninteracting case, of having configurations of the cross-link whose ends lie within a given volume consistent with both a distance and orientation constraint on the cross-linked α-helical blocks that allows the interhelical interaction to occur. A reasonable estimate of δ_{CHC} may be made by employing the rotational isomeric calculation of Maroun and Mattice[39] on the allowed side-chain configurations of a cystinyl cross-link in a coiled coil and which leads to a value around 0.1. This is about the order of magnitude one might expect. We have also included in the CHC type states those states in which both cross-linked block pairs are in noninteracting states but in which there is at least one interacting helical stretch in the molecule (see Figure 2B, structures **4**, **5**, and **10**). While the configurational entropy of the cross-link is clearly not as reduced as it would be if one of the cross-linked blocks were in an interacting helical state, we would expect some reduction in the phase space accessible to the real molecule relative to X≈X type states due to the presence of the interacting helical stretch(es). This factor is unknown. However, these double eyelet states contribute at most a few percent to the Z_{chc} partition function. This, we feel it is a reasonable approximation to uniformly penalize all CHC states relative to X≈X states by the factor δ_{CHC}.

3. If both cross-linked block pairs are in interacting helical states, we assign a statistical weight δ_{HH}; thus H=H and HCH types states experience the same entropic penalty. Representatives of these states are shown in Figure 1 and Figure 3, respectively. If the available configurations of a given cross-link were independent of that of the other cross-link, then $\delta_{HH} = \delta_\Omega\delta_{CHC}^2$. Physically we would expect the number of configurations available to a real cross-link to be in fact less if both cross-linked block pairs are in interacting helical states than if just one were in an interacting helical state; i.e. $\delta_{HH} < \delta_\Omega\delta_{CHC}^2$. Unless otherwise indicated, we will ignore this subtlety in the calculations that follow; however, to retain this option, we employ the symbol δ_{HH} in the assessment of the entropic penalty experienced by H=H and HCH states.

With the above in mind, the total internal partition function of the doubly cross-linked molecule

$$Z_{tot} = Z_{2XL}/\delta_\Omega \qquad \text{(II-42a)}$$

where

$$Z_{2XL} = Z_{xx} + \delta_\Omega\delta_{CHC}Z_{chc} + \delta_\Omega\delta_{HH}(Z_{hh} + Z_{hch}) \qquad \text{(II-42b)}$$

wherein Z_{xx}, Z_{chc}, Z_{hch}, and Z_{hh} are defined in eq II-32, II-12, II-25, and II-2, respectively.

Similarly, in the approximate "all-or-none" model, the internal partition function is

$$Z_{AN} = Z_{xx} + \delta_\Omega\delta_{HH}Z_{hh} \qquad \text{(II-43)}$$

Now, the overall helix content of a double cross-linked molecule

$$f_{2XL} = \{f_{xx}Z_{xx} + f_{chc}\delta_\Omega\delta_{CHC}Z_{chc} + \delta_\Omega\delta_{HH}(f_{hh}Z_{hh} + f_{hch}Z_{hch})\}/Z_{2XL} \qquad \text{(II-44)}$$

wherein f_{xx}, f_{chc}, f_{hch}, and f_{hh} are given in eq II-35, II-18, II-27, and II-6a, respectively.

The helix content in the "all-or-none" model is

$$f_{AN} = \{f_{xx}Z_{xx} + \delta_\Omega\delta_{HH}f_{hh}Z_{hh}\}/Z_{AN} \qquad \text{(II-45)}$$

Finally, the helix content of the jth block of the doubly cross-linked molecule is

$$f_{2XL}(j) = \{f_{xx}(j)Z_{xx} + f_{chc}(j)\delta_\Omega\delta_{CHC}Z_{chc} + \delta_\Omega\delta_{HH}(f_{hh}(j)Z_{hh} + f_{hch}(j)Z_{hch})\}/Z_{2XL} \qquad \text{(II-46)}$$

where $f_{xx}(j)$, $f_{chc}(j)$, $f_{hch}(j)$, and $f_{hh}(j)$ may be found in subsections II-E, II-C, II-D, and II-B, respectively, entitled "Helix Probability Profiles".

Finally, the helix content of the jth block in the "all-or-none" model is given by

$$f_{AN}(j) = \{f_{xx}(j)Z_{xx} + \delta_\Omega\delta_{HH}f_{hh}(j)Z_{hh}\}/Z_{AN} \qquad \text{(II-47)}$$

With the completion of section II-F, the formal development of the theory of the helix–coil transition of doubly cross-linked, two-chain, coiled coils is finished. In section III, we shall present the application of the theory to doubly cross-linked, homopolymeric, two-chain, coiled coils.

III. Application to Homopolypeptides

The first and most important point we address is the validity of the all-or-none model of the helix–coil transition in doubly cross-linked, coiled coils. In all cases presented below, we have set $m = 4$, $m_{l1} = m_{l2} = 3$, $m_{r1} = m_{r2} = 2$, $l_c = 2$, $s = 0.94$, $\delta_\Omega\delta_{CHC} = 0.1$, $\delta_{HH} = \delta_\Omega\delta_{CHC}^2$ and for the singly cross-linked, coiled coils $r_\phi = \delta_\Omega\delta_{CHC}$. We begin in Figure 6 by examining the homopolymeric analogue of rabbit β-tropomyosin, and set $N_B = 71$, $N_{C1} = 9$, $N_{C2} = 48$, and $\sigma = 5 \times 10^{-4}$. In curve A, we have plotted the helix content obtained via the all-or-none model, f_{AN}, employing eq II-45, vs. w. Throughout the entire range or w, the helix content calculated in the all-or-none and the more general, doubly cross-linked interior eyelet model are completly indistinguishable on the scale of the graph and agree to 0.4% or better. For example, when $w = 1.75$ and where the disagreement between the two models is greatest, $f_{hh} = 0.7981$, $f_{chc} = 0.7091$, $f_{hch} = 0.7834$, and $f_{AN}/f_{2XL} = 1.0036$. Even in those states that possess an interacting helical stretch plus constrained random coils between cross-links, loop entropy acts to keep the contained random coil loop quite small. Thus, from the comparison of the overall helix content, we conclude that the all-or-none model, in which *either* the molecule is *fully helical* and *interacting* between the pair of cross-linked blocks or *lacks* any *interacting*

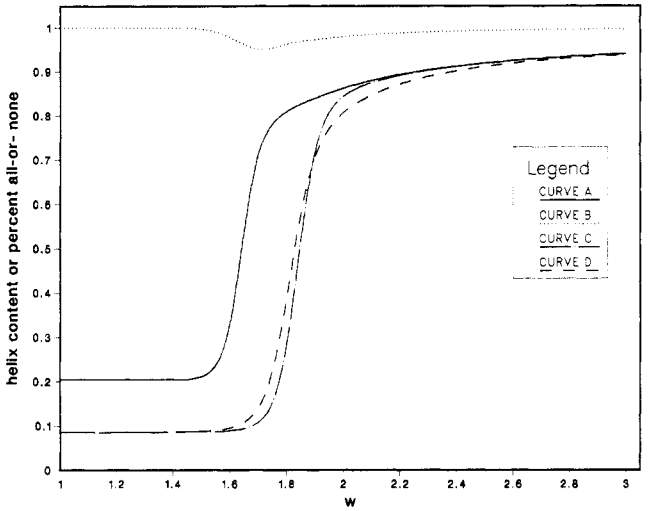

Figure 6. Plot of the overall helix content of a doubly (singly) cross-linked, homopolymeric, two-chain, coiled coil calculated in the all-or-none model, f_{AN}, (loops-excluded model, f^0_{he}) employing eq II-45 (eq II-8a ff of ref 18) vs. w in curve A (curves C and D) with $N_{C1} = 9$ and $N_{C2} = 48$. Curve B is the fraction of doubly cross-linked molecules that satisfy the all-or-none criterion, Z_{AN}/Z_{2XL}, vs. w. Z_{2XL} and Z_{AN} are defined in eq II-42b and II-43, respectively. In all cases, $N_B = 71$ and $\sigma = 5 \times 10^{-4}$. See text for additional parameters.

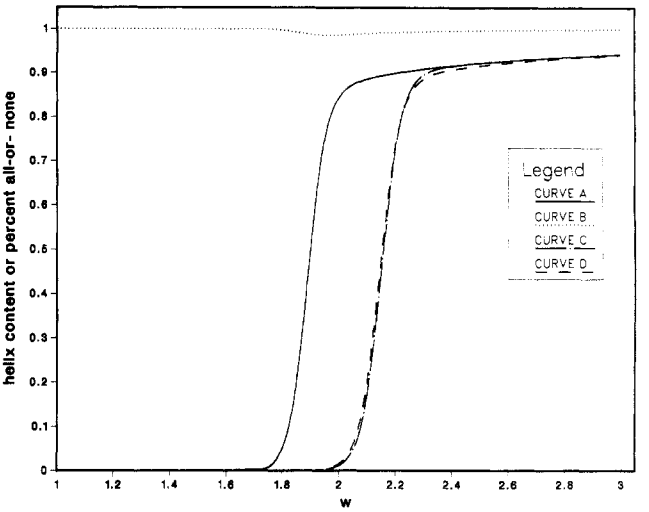

Figure 7. Plot of the overall helix content of a doubly (singly) cross-linked, homopolymeric, two-chain, coiled coil calculated in the all-or-none model, f_{AN}, (loops-excluded model, f^0_{he}) employing eq II-45 (eq II-8a ff of ref 18) vs. w in curve A (curves C and D) with $N_{C1} = 9$ and $N_{C2} = 48$. Curve B is the fraction of doubly cross-linked molecules that satisfy the all-or-none criterion, Z_{AN}/Z_{2XL}, vs. w. Z_{2XL} and Z_{AN} are defined in eq II-42b and II-43, respectively. In all cases, $N_B = 71$ and $\sigma = 10^{-6}$. See text for additional parameters.

pairs of *helices*, is an excellent description of the physics for this set of conditions.

A more stringent test is to calculate the relative population of molecules that satisfy the all-or-none requirement, Z_{AN}/Z_{2XL}. If this ratio lies close to unity throughout the entire course of the helix–coil transition, then we can conclude that the all-or-none approximation is really very good. In curve B of Figure 6, we plot Z_{AN}/Z_{2XL} vs. w with Z_{AN} and Z_{2XL} given in eq II-43 and II-42b, respectively. The ratio has a minimum very close to $w = 1.75$, where $Z_{AN}/Z_{2XL} = 0.9557$.

We next turn to an examination of the relative importance of the subpopulations that contribute to Z_{2XL}. As expected for the choice of σ and s employed here, in the limit that $w = 1$, the X≈X subpopulation dominates, and $Z_{xx} \gg Z_{chc} > Z_{hch} > Z_{hh}$. This ordering is easily rationalized. The statistical weight of the fully helical stretch between cross-linked blocks at $N_{C1} = 9$ and $N_{C2} = 48$ is very small, and thus HH states contribute the least. HCH states can have breaks in the side-by-pair of interacting helices and thus contain a large number of noninteracting helical and constrained random coil loops between cross-links; hence for this value of σ, $Z_{hch} > Z_{hh}$. However, we point out that in limit of very small σ, the likelihood of states containing two or more helical stretches becomes nil and then $Z_{hh} > Z_{hch}$. Similarly, CHC states can have constrained interior random coil loops and interacting helices but lack the constraint that both pairs of cross-linked blocks must be helical; therefore $Z_{chc} > Z_{hch}$.

In the limit of high overall helix content, the X≈X subpopulation becomes unimportant. Since the fully helical molecule is part of the H=H subpopulation, as the helix content increases to unity the all-or-none model must become correct. As w increases, the HCH subpopulation eventually becomes more important than the CHC subpopulation. The latter states satisfy the requirement that there be at least one randomly coil block located at a cross-link, whereas the former states satisfy the weaker condition that there be at least one randomly coiled block in the molecule somewhere between helical cross-links. If the cross-links are spaced sufficiently far apart, the combinatorial factor inherent in the HCH subpopulation will eventually dominate and $Z_{hch} > Z_{chc}$. Of course, the subpopulation lacking any breaks in the interacting pairs of helices whatsoever has an even larger statistical weight, and thus, we arrive at the ordering $Z_{hh} > Z_{hch} > Z_{chc}$ in the limit of high overall helix content.

On the basis of the preceding two paragraphs, it is clear that the all-or-none model must always hold at the two extremes of low and high helix content. Whether it holds throughout the entire course of the helix–coil transition depends on the relative weights of CHC and HCH states vis à vis H=H and X≈X states. Intuitively we would expect that by keeping everything the same, decreasing (increasing) σ (thereby making interior random coil loop formation in the CHC and HCH states more (less) difficult) should enhance (decrease) the validity of the all-or-none model. These expectations are realized in curves A and B of Figure 7 and 8, where we plot the overall helix content and the ratio Z_{AN}/Z_{2XL} vs. w respectively for $\sigma = 10^{-6}$ and $\sigma = 10^{-2}$. In all cases, $N_{C1} = 9$ and $N_{C2} = 48$.

Throughout the entire range of w, when $\sigma = 10^{-6}$, the all-or-none model is a very good representation of the physics. Z_{AN}/Z_{2XL} always exceeds 0.986, and, throughout the entire course of the helix–coil transition displayed in Figure 7, f_{AN} and f_{2XL} differ by at most 0.7%. Thus we display f_{AN} vs. w in Figure 7, with f_{AN} calculated via eq II-45. Turning to the case where $\sigma = 10^{-2}$ in Figure 8, we plot f_{2XL} vs. w in curve A calculated via eq II-44. The plot of Z_{AN}/Z_{2XL} vs. w, shown in curve B of Figure 8, has a minimum near $w = 2.0$, where it assumes the value 0.8398. We find as expected that the all-or-none approximation is not as robust as compared to the case when $\sigma = 5 \times 10^{-4}$ or 10^{-6}; but it is still a fairly good representation of the physical situation. Thus, by this most rigorous criterion, at most about 16% of the molecules have constrained interior random coil loops plus interacting α-helices between cross-links. The ratio of f_{AN} to f_{2XL} has a minimum in the vicinity of $w = 1.80$, where $f_{AN}/f_{2XL} = 0.9595$, $f_{2XL} = 0.3859$, and the helix content of the H=H, CHC, and HCH subpopulations are $f_{hh} = 0.7453$, $f_{chc} = 0.6799$ and $f_{hch} = 0.7283$. Furthermore, setting $w = 2.0$, $f_{AN}/f_{2XL} =$

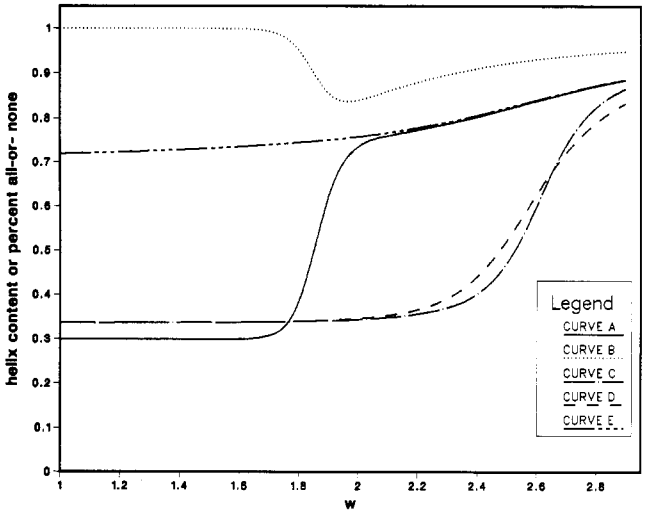

Figure 8. Plot of the overall helix content of a doubly (singly) cross-linked, homopolymeric, two-chain, coiled coil calculated in the doubly cross-linked interior eyelet model, f_{2XL}, (loops-excluded model, f^0_{he}) employing eq II-44 (eq II-8a ff of ref 18) vs. w in curve A (curves C and D) with $N_{C1} = 9$ and $N_{C2} = 48$. Curve B is the fraction of doubly cross-linked molecules that satisfy the all-or-none criterion, Z_{AN}/Z_{2XL}, vs. w. Z_{2XL} and Z_{AN} are defined in eq II-42b and II-43, respectively. Curve E presents the plot of f_{hh}, the helix content in the H=H subpopulation calculated via eq II-6a ff, vs. w. In all cases $N_B = 71$ and $\sigma = 10^{-2}$. See text for additional parameters.

1.0038, $f_{2XL} = 0.7534$, $f_{hh} = 0.7586$, $f_{chc} = 0.7041$, and $f_{hch} = 0.7431$. In other words, since loop entropy tends to keep the constrained random coil loops rather small, the measurement of the overall helix content itself is a very insensitive measure of the adequacy of the all-or-none approximation. The broad plateau region in the vicinity of 75% helix content reflects the difficulty in propagating helices substantially beyond the cross-linked blocks. This is verified in curve E where we plot f_{hh} vs. w. The f_{hh} curve coalesces with the f_{2XL} vs. w plot, curve A, near 77% helix. Finally, we observe that the qualitative behavior of the f_{2XL} vs. w curves as a function of σ and shown in Figures 6–8 is identical with that seen in singly cross-linked and non-cross-linked coiled coils and, having been addressed elsewhere, will not be discussed further.[16-18]

We have also plotted in curves C and D of Figures 6–8 the overall helix content, f^0_{he} obtained via eq II-8a ff of ref 18, of singly cross-linked homopolymeric, two-chain, coiled coils vs. w, cross-linked at $N_C = 9$ and 48, respectively, at the specified value of σ. Comparison of curve A (the doubly cross-linked molecule) with curves C and D (singly cross-linked molecules) reveals one particularly striking difference that holds independent of σ. The homopolymeric, doubly cross-linked, coiled coil is more stable than the analogous singly cross-linked, coiled coil with the two curves coalescing in the limit of high helix content. This is a general consequence of the destabilization of the noninteracting X≈X states relative to the interacting states due to the presence of constrained random coil loops. As mentioned in the introduction, similar behavior has been seen elsewhere in a real globular protein by Lin, Konishi, Denton, and Scheraga.[31] The slopes of the helix content vs. w curves in singly and doubly cross-linked molecules are quite similar. Whether or not the slope in the doubly cross-linked molecule is greater or less than that in a singly cross-linked molecule depends on the particular value of σ and the range of parameters employed (see below, in particular for the effect of variation of cross-link location).

Another important question that must be addressed is the helix content of the X≈X conformations relative to the helix content of a non-cross-linked (or equivalently at the present level of approximation, singly cross-linked) molecule lacking any interacting helices. First of all, it must be recognized that we have focused on X≈X conformations that essentially contain at most a single helical stretch per chain between cross-links (see the introduction to section II-E). Thus, especially in the limit of very low helix content, the relevant reference state is a non-cross-linked molecule treated in the one helical sequence per chain approximation with helix content f_s.[20,40] (We point out that in the doubly cross-linked molecule, multiple helical sequences are permitted beyond the cross-links.) On the basis of a detailed analysis of the statistical weights of constrained loops containing helices and random coils, we find that if the mean length of the helical stretch in the absence of cross-links is somewhat less than the root mean square end to end vector of the unconstrained Gaussian chain formed by snipping the closed random coil loop in the doubly cross-linked molecule, then f_{xx} will exceed f_s. If this is not true, then $f_{xx} < f_s$. Moreover, for homopolymers, the ordering of f_{xx} vis à vis f_s is independent of σ (due to the one helical sequence approximation between cross-links) but is dependent on the size of the random coil loop between cross-links, the helix propagation parameter, and the presence or absence of unconstrained ends. Whether or not f_{xx} exceeds f_{hm} depends on the importance of multiple helical stretches relative to the enhanced stability of short, constrained helices in closed loops. We might expect f_{hm} to exceed f_{xx} in the limit of large σ.

The qualitative picture developed above is confirmed by the relative ordering of f_{xx}, f_{hm}, and f_s for the cases considered in Figures 6–8. When $\sigma = 10^{-6}$, $f_{xx} = 2.146 \times 10^{-4}$, whereas $f_s = 5.816 \times 10^{-5}$. Note however that $f_{hm} = 2.299 \times 10^{-4}$. On increasing $\sigma = 5 \times 10^{-4}$, $f_{xx} = 0.2041$ and $f_s = 1.925 \times 10^{-2}$. Here the agumented stability of helices in the loop is now sufficiently great that f_{xx} even exceeds $f_{hm} = 8.546 \times 10^{-2}$. Finally, if $\sigma = 10^{-2}$, $f_{xx} = 0.2989$, whereas $f_s = 5.174 \times 10^{-2}$. However, now the presence of multiple helical stretches in the noninteracting, singly cross-linked conformation now becomes important and f_{hm} exceeds f_{xx} with $f_{hm} = 0.3364$. As the foregoing discussion indicates, where f_{hm} exceeds or lies below f_{xx} is the result of a complicated interplay of various factors; there is no single ordering applicable to all conditions.

To examine the effect of variation of cross-link location on the helix–coil transition, we plot in Figure 9 f_{2XL} calculated via eq II-44 vs. w for $\sigma = 5 \times 10^{-4}$, $N_B = 71$, $N_{C1} = 9$, 45, and 47, and $N_{C2} = 48$ in curves A–C, respectively, as well as the singly cross-linked, coiled coil having $N_C = 48$ in curve D. As might be expected, decreasing the spacing between cross-links has made the all-or-none model better. When $N_{C1} = 45$, the ratio Z_{AN}/Z_{2XL} has a minimum near $w = 1.5$ with a value $Z_{AN}/Z_{2XL} = 0.9800$. Moreover, when $N_{C1} = 47$, the minimum Z_{AN}/Z_{2XL} occurs near $w = 1.6$ with a value $Z_{AN}/Z_{2XL} = 0.9946$. Observe that in the limit of high helix content, all the singly and doubly cross-linked helix content curves coalesce. In the transition region, the singly cross-linked molecule is less stable than the doubly cross-linked molecule; this results from the greater destabilization of the noninteracting states vis à vis the interacting helical states in doubly as compared to singly cross-linked molecules. This leads to the situation where the slope of the helix content vs. w curves is less in the doubly than in the singly cross-linked case, even though the doubly cross-linked molecules have in a sense a more cooperative transition. That is, if the doubly cross-linked molecule contains an interacting helical stretch, it is essentially fully helical between cross-links, whereas singly

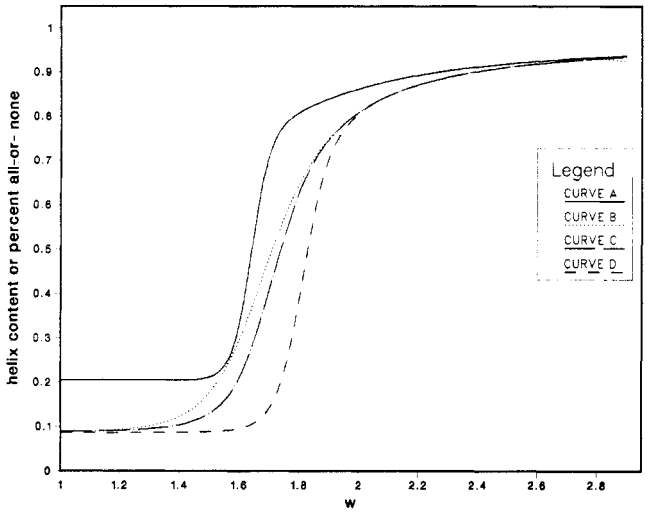

Figure 9. Plot of the overall helix content of a doubly cross-linked, coiled coil f_{2XL}, calculated via eq II-44, vs. w for $\sigma = 5 \times 10^{-4}$, $N_B = 71$, $N_{C2} = 48$ and $N_{C1} = 9$, 45, and 47 in curves A–C, respectively. In curve D, we plot the helix content of the singly cross-linked molecule obtained from eq II-8a ff of ref 18 vs. w having $\sigma = 5 \times 10^{-4}$, $N_B = 71$, and $N_C = 48$. See text for additional parameters.

cross-linked molecules obey the far less stringent requirement that the interacting helical stretch merely includes the cross-linked pair of blocks. Identical behavior is seen on setting $\sigma = 10^{-2}$; therefore we do not display this case here.

Part of the underlying origin of the validity of the all-or-none model in the examples discussed so far is the fact that the cross-links are not located at the chain ends. Hence, there are a large number of helical conformations that propagate beyond the cross-linked blocks (where the helix–coil transition is still continuous). The multiplicity of helical states acts in concert with the destabilization of constrained interior random coil states between the cross-links to produce the observed ratios of Z_{AN}/Z_{2XL}. Thus, it is of interest to examine a hypothetical coiled coil whose cross-links are located near the chain ends.

In Figure 10A (10B) we have plotted in curve A f_{2XL} (Z_{AN}/Z_{2XL}) vs. w for a homopolymeric, coiled coil having $N_B = 38$, $N_{C1} = 2$, $N_{C2} = 37$, $\sigma = 5 \times 10^{-4}$, and $s = 0.94$. All the other parameters remain unchanged. Observe that f_{xx} is 0.1618, whereas in the case where $N_B = 71$, $N_{C1} = 9$, and $N_{C2} = 48$, $f_{xx} = 0.2041$. This is a relative change in f_{xx} of 20%, whereas f_{hm} has decreased from 0.08546 to 0.07816, a relative change of about 8%. The greater relative decrease in f_{xx} is a reflection of the decrease in the number of noninteracting helical stretches that propagate beyond the cross-links. Furthermore we have also plotted in curve B of Figure 10A f_{AN} vs. w. Curves A and B differ throughout the course of the transition. The minimum value in f_{AN}/f_{2XL} occurs near $w = 1.80$, where $f_{AN}/f_{2XL} = 0.8502$ and $f_{2XL} = 0.3828$. Note that the all-or-none model in the vicinity of the minimum of the f_{AN}/f_{2XL} has a lower helix content than the full DCIEM, since HCH and CHC states having a higher helix content than X≈X states have been excluded. In the limit of high helix content, in this case near $w = 2.0$, the f_{AN} and f_{2XL} curves cross. Now X≈X states contribute negligibly; we are only interested in interacting helical states. Since H=H states have the highest helix content, we find that $f_{AN} \geq f_{2XL}$. In addition, in Figure 10B, Z_{AN}/Z_{2XL} has a minimum near $w = 1.95$, where this ratio has the value 0.7381. Here, $f_{2XL} = 0.8475$, $f_{chc} = 0.9074$, $f_{hch} = 0.9414$, and $f_{hh} = 0.9673$. The minimum in f_{AN}/f_{2XL} and Z_{AN}/Z_{2XL} occurs at different points in the

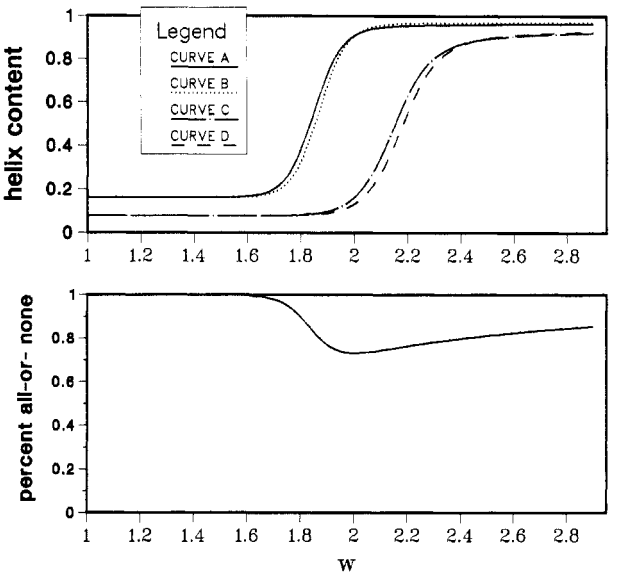

Figure 10. (A) Plot of the overall helix content of a doubly cross-linked, homopolymeric, two-chain, coiled coil calculated in the doubly cross-linked interior eyelet (all-or-none) model, f_{2XL} (f_{AN}) employing eq II-44 (II-45) vs. w in curve A (B) where $N_{C1} = 2$ and $N_{C2} = 37$. In curves C and D, we plot the helix content of a singly cross-linked chain, f^0_{he}, calculated via eq II-8a ff of ref 18, vs. w, having $N_C = 2$ and 37, respectively. In all cases $N_B = 38$ and $\sigma = 5 \times 10^{-4}$. (B) Plot of the fraction of doubly cross-linked, two-chain, coiled coils that satisfy the all-or-none criterion, Z_{AN}/Z_{2XL}, vs. w, with Z_{2XL} and Z_{AN} defined in eq II-42b and II-43, respectively. See text for additional parameters.

transition (near $w = 1.80$ and 1.95, respectively) as the two quantities reflect different averages. What is remarkable though is the robustness of the all-or-none model even here. The transition occurs in approximately the same place for both the DCIEM and all-or-none models. Basically, CHC and HCH states have quite a high content when they contribute appreciably to the overall population; due to the effect of loop entropy, constrained interior random coil loops are quite tiny.

Finally, in curves C and D of Figure 10A we have plotted the helix content of the singly cross-linked chains, f^0_{he}, calculated via eq II-8a of ref 18 vs. w with $N_B = 38$ and $N_C = 2$ and 37, respectively. Comparison with the analogous transition in the tropomyosin analogue (shown in curves A, C, and D of Figure 6) reveals that in the present case, the difference between the helix–coil transition in the singly and doubly cross-linked molecules is more appreciable. Basically, the X≈X states are less stable relative to the interacting helical states when free chain ends are clipped off. This is confirmed by the following calculation. The ratio of the Z_{xx} partition function with $N_B = 38$, $N_{C1} = 2$, and $N_{C2} = 37$ to that when $N_B = 71$, $N_{C1} = 9$, and $N_{C2} = 48$ is 0.0469, whereas for the noninteracting states in a singly cross-linked dimer, the ratio of partition function when $N_B = 38$ and $N_C = 37$ to that when $N_B = 71$ and $N_C = 48$ is 0.188. Thus, these calculations point out the interplay between the continuous and two-state character of the transition in determining the relative stability of singly vis-à-vis doubly cross-linked, coiled coils.

Helix probability profiles of doubly cross-linked, coiled coils display a broad plateau regime between cross-links, followed by a drop off at the ends. Since for homopolymers having a site independent w they are uninformative, in the interest of brevity, none are presented. Furthermore, since mismatched states between cross-links require the formation of two closed, random coil loops and since these states make a very small contribution to the total partition function, just as in singly cross-linked chains,[19] we con-

jecture that in doubly cross-linked chains, mismatched associations of interacting helices should be unimportant. This concludes the application of the theory to homopolymeric, doubly cross-linked, coiled coils.

IV. Discussion

In this paper, we have developed a statistical mechanical theory of the helix–coil transition in doubly cross-linked, coiled coils and have presented expressions for the calculation of the partition function, overall helix content, and helix probability profiles in the context of two models. The simpler "all-or-none" model asserts that the net effect of loop entropy is so overwhelming that either the coiled coil contains a pair of interacting helices between the two cross-linked pairs of α-helical turns, the "native" state, or it lacks any interacting helices whatsoever, the "denatured" state. Observe that beyond the cross-links, the helix–coil transition is still continuous and in the denatured state noninteracting helices are allowed. We have also developed the more general, doubly cross-linked, interior eyelet model that permits constrained random coil loops containing interacting helices between cross-linked blocks. On the basis of the application of the theory to the homopolymeric analogue, we believe that the helix–coil transition in rabbit β-tropomyosin should be well approximated by the "all-or-none" model. This is the major conclusion of this work and is to the best of our knowledge the first time a two-state model, albeit in the sense defined above, has emerged from a continuous statistical mechanical theory of a native → denatured transition.

The fact that an "all-or-none" model of the helix–coil transition in doubly cross-linked, coiled coils emerges as a reasonable approximation to the full physics for the various homopolymeric systems studied once again points out the crucial role played by loop entropy in these systems. In fact, this is the dominant theme that emerges from the study of coiled coils. In non-cross-linked chains of moderate length, loop entropy acts to produce a single interacting helical stretch per molecule (multiple, interacting stretches in coiled coils required constrained random coil loops, such loops pay a large entropic price). However, the interacting helical stretch may occur anywhere and mismatched, out-of-register conformations of the two chains can occur. Introduction of one cross-link acts to localize the helical stretch. Now, either the cross-linked pair of blocks is helical and interacting or there are essentially no interacting pairs of helices in the molecule. Moreover, mismatched, out-of-register states are excluded. The presence of a second cross-link further restricts the available phase space. The dominant species are either states that are fully helical between the two cross-linked blocks or states that lack any interacting helices. Moreover, mismatched states can be neglected. The destabilization of the "denatured" form due to constrained random coil loops makes the "native" form more stable vis â vis singly cross-linked chains. Overall then, there seems to be a strong isomorphism between doubly cross-linked, coiled coils and globular proteins. If so, it is a reasonable conjecture that an underlying cause of the all-or-none transition in globular proteins is also loop entropy; the introduction of favorable site-specific interactions only strengthens the case.

In the above, we have assumed that the cross-link leaves the coiled coil structure intact. In the real coiled coils, this may not be the case.[41,42] If so, multiple cross-links may reduce the enhanced stability of doubly, relative to singly, cross-linked chains. In future work, once experimental data on singly and doubly cross-linked rabbit β-tropomyosin becomes available, we shall apply the theory to these two molecules and attempt to assess the importance of regiospecific effects.[41-44]

In order to experimentally check the validity of the all-or-none model, it is clear from the above that measurement of the helix content alone is in itself inadequate. What are required at a minimum are calorimetric studies of doubly cross-linked and singly cross-linked coiled coils to determine the ratios of the Van't Hoff to calorimetric enthalpies for the entire thermal transition.[45] If the theory presented here is correct, singly and doubly cross-linked, coiled coils should behave quite differently. The detailed analysis of the predictions of the present theory for the differential heat capacity curves is beyond the scope of the present work. Finally, ideally, what one needs is a probe that correlates helix content between two spatially far apart regions. Unfortunately, such probes are unavailable at present and are unlikely to emerge in the near future.

In conclusion, we believe that doubly cross-linked, coiled coils have been shown to possess many of the qualitative features of globular proteins; they allow us to examine the relative importance of "short-range" vs. "long-range" interactions in determining the stability of the coiled coil structure and graphically demonstrate the importance of topological constraints in determining protein stability. The present study not only brings multiply cross-linked coiled coils under the auspices of a statistical mechanical theory but also points the way to the development of a statistical mechanical theory of the native → denatured transition in globular proteins.

Acknowledgment. This research was supported in part by a grant from the Biophysical Program of the National Science Foundation (No. PCM 82-12404). Thanks are due to Professors Holtzer and Yaris for useful discussions.

Supplementary Material Available: Appendix A: CHC Type States Statistical Weights, Appendix B: HCH Type States Statistical Weights, Appendix C: Statistical Weights of Completely Noninteracting Conformations Between Cross-Linked Blocks, and Appendix D: Configurational Entropy of Closed Loops Composed of Gaussian Random Coils and Phantom Rods (34 pages). Ordering information is given on any current masthead page.

References and Notes

(1) Crimmins, D.; Holtzer, A. *Biopolymers* **1981**, *20*, 925.
(2) Skolnick, J.; Holtzer, A. *Macromolecules* **1982**, *15*, 303.
(3) (a) Talbot, J. A.; Hodges, R. S. *Acc. Chem. Res.* **1982**, *15*, 224. (b) Lau, S. Y. M.; Taneja, A. K.; Hodges, R. S. *J. Biol. Chem.* **1984**, *259*, 13253.
(4) Szent-Györgyi, A. G.; Cohen, C.; Kendrick-Jones, J. *J. Mol. Biol.* **1971**, *56*, 239.
(5) Cohen, C.; Szent-Györgyi, A. G. *J. Am. Chem. Soc.* **1957**, *79*, 248.
(6) Cohen, C.; Holmes, K. *J. Mol. Biol.* **1963**, *6*, 423.
(7) Lowey, S.; Kucera, J.; Holtzer, A. *J. Mol. Biol.* **1963**, *7*, 234.
(8) Holtzer, A.; Clark, R.; Lowey, S. *Biochemistry* **1965**, *4*, 2401.
(9) Woods, E. *Biochemistry* **1969**, *8*, 4336.
(10) Caspar, D.; Cohen, C.; Langley, W. *J. Mol. Biol.* **1969**, *41*, 87.
(11) Hodges, R.; Sodek, J.; Smillie, L.; Jurasek, L. *Cold Spring Harbor Symp. Quant. Biol.* **1972**, *37*, 299.
(12) McLachlan, A.; Stewart, M. *J. Biol. Chem.* **1975**, *98*, 293.
(13) Stone, D.; Smillie, L. *J. Biol. Chem.* **1978**, *253*, 1137.
(14) Mak, A.; Lewis, W.; Smillie, L. *FEBS Lett.* **1979**, *105*, 232.
(15) Skolnick, J. *Macromolecules* **1983**, *16*, 1069.
(16) Skolnick, J. *Macromolecules* **1983**, *16*, 1763.
(17) Skolnick, J. *Macromolecules* **1984**, *17*, 645.
(18) Skolnick, J. *Macromolecules* **1985**, *18*, 1535.
(19) Chen, C. L.; Skolnick, J. *Macromolecules* **1986**, *19*, 242.
(20) Poland, D.; Scheraga, H. A. "Theory of Helix–Coil Transitions in Biopolymers"; Academic Press: New York, 1970.
(21) Zimm, B.; Bragg, J. *J. Chem. Phys.* **1959**, *31*, 526.
(22) Skolnick, J.; Holtzer, A. *Macromolecules* **1985**, *18*, 1549.
(23) Yukioka, S.; Noda, I.; Nagasawa, M.; Holtzer, M. E.; Holtzer, A. *Macromolecules* **1985**, *18*, 1083.
(24) Lehrer, S. *J. Mol. Biol.* **1978**, *118*, 209.
(25) Schellman, J. A. *C.R. Trav. Lab. Carlsberg., Ser. Chim.* **1955**, *29*, 230.
(26) Flory, P. *J. Am. Chem. Soc.* **1956**, *78*, 5222.

(27) Jacobson, H.; Stockmayer, W. H. *J. Chem. Phys.* **1950**, *18*, 1600.
(28) Cummins, P.; Perry, S. V. *Biochem. J.* **1973**, *133*, 765.
(29) Sodek, J.; Hodges, R. S.; Smillie, L. B.; Jurasek, L. *Proc. Natl. Acad. Sci. U.S.A.* **1972**, *69*, 3800.
(30) Poland, D. C.; Scheraga, H. A. *Biopolymers* **1965**, *3*, 379.
(31) Lin, S. H.; Konishi, Y.; Denton, M. E.; Scheraga, H. A. *Biochemistry* **1984**, *23*, 5504.
(32) Scheraga, H. A. *J. Phys. Chem.* **1960**, *64*, 1917.
(33) Tanford, C. *J. Am. Chem. Soc.* **1962**, *84*, 4240.
(34) Brandts, J.; Lumry, R. *J. Phys. Chem.* **1963**, *67*, 1484.
(35) Wetlaufer, D. B.; Malik, S. K.; Stoller, L.; Coffin, R. L. *J. Am. Chem. Soc.* **1964**, *86*, 508.
(36) Crothers, D.; Kallenbach, N. *J. Chem. Phys.* **1966**, *45*, 917.
(37) Flory, P. "Statistical Mechanics of Chain Molecules"; Wiley: New York, 1969; p 251.
(38) Reference 37, p 74.
(39) Maroun, R.; Mattice, W. *Biochim. Biophys. Acta* **1984**, *784*, 133.
(40) Calculated by assuming coarse graining.
(41) Betteridge, D.; Lehrer, S. *J. Mol. Biol.* **1983**, *167*, 481.
(42) Graceffa, P.; Lehrer, S. *Biochemistry* **1984**, *23*, 2606.
(43) Pato, M.; Mak, A.; Smillie, L. *J. Biol. Chem.* **1981**, *256*, 593.
(44) Williams, D. L., Jr.; Swenson, C. *Biochemistry* **1981**, *20*, 3856.
(45) Privalov, P. L. *Adv. Protein Chem.* **1979**, *33*, 167.
(46) Chandrasekhar, S. *Rev. Mod. Phys.* **1943**, *15*, 1.
(47) In the calculations presented in the text we set b_0 = 8.0 Å, see: Holtzer, M.; Holtzer, A. *Macromolecules* **1972**, *5*, 294.
(48) In the calculations presented in the text we employ C_1 = 0.2313, which is equivalent to setting u_ϕ = 359 Å^3; see ref 2.
(49) Abramowitz, M.; Stegun, I. A. "Handbook of Mathematical Functions"; Dover Publications: New York, 1965; Chapter 7.

Phase Separation in Ternary Systems Solvent–Polymer 1–Polymer 2. 1. Cloud Point and Critical Concentration

Karel Šolc

Michigan Molecular Institute, Midland, Michigan 48640. *Received October 25, 1985*

ABSTRACT: A novel approach is proposed for analysis of phase equilibria in Flory–Huggins (FH) ternary systems solvent–polymer 1–polymer 2. It is based on the separation factors σ_1 and σ_2, familiar from the theory of quasi-binary systems, and on two related variables $\eta^2 = \sigma_1\sigma_2$ and $\xi^2 = \sigma_2/\sigma_1$. While η is a measure of the distance of a cloud point from the critical point (CP), ξ characterizes the compatibility of the two polymers and their relative tendency to accumulate in one phase over the other due to molecular interactions. A single formally temperature-independent cloud-point equation, $F(\sigma_1,\sigma_2,\phi) = 0$, can be formulated; unlike in quasi-binary systems, however, here it is not self-contained and has to be solved in conjunction with two other equilibrium equations. Expansion of F around the CP leads to a simple closed expression for the critical concentration ϕ_c in terms of polymer mixture composition, chain lengths r_1 and r_2, and ξ_c^2. The formal separation of ϕ_c from the triplet of interaction parameters thus achieved simplifies greatly the analysis of ϕ_c. It is shown that for a given set of r_i's, the critical binodal direction depends only on the location of the CP in the composition triangle; this fact invites a simple test of the FH theory. Scott's segregating systems appear in our general theory as a special singular case.

1. Introduction

Most of the detailed liquid–liquid phase separation studies have been done in quasi-binary systems consisting of a solvent and a polydisperse polymer whose molecules differ only by their chain length. The chemical homogeneity of polymer segments in such solutions translates into a single parameter characterizing interactions with the solvent and a relatively simple form of phase equilibrium equations to be solved. Consequently, the effects of the molecular weight distribution on phase diagrams of quasi-binary solutions are fairly well understood.

The situation is markedly different for systems with a polymeric solute that is chemically inhomogeneous. Our understanding of even the simplest version of such systems, namely, of ternary systems solvent–polymer 1–polymer 2, is scant, and it has not seen much progress over the past 35 years. The only analytical results of which we are aware concern the approximate relations for the critical point published by Scott in 1949, valid for a certain class of ternary systems with segregating polymers.[1] Numerical calculation of binodals, the only type of data that can be legitimately used for discussion of solubility diagrams, is "notoriously difficult",[2] requiring "long and tedious approximation methods",[1] and for this reason usually avoided by authors at the expense of accuracy of their conclusions. The solubilities are routinely estimated just from the position of more easily accessible critical points or spinodals[3-7] and only exceptionally evaluated rigorously from phase equilibrium equations.[8] Even in this case, the numerical methods employed leave something to be desired: for instance, the "brute force" optimization technique for the sum of squared chemical potential differences modified by a penalty function[8] seems to fail around the critical point, presumably because of the computer round-off errors that become serious in this region.

The most fruitful approach to analyzing phase equilibria in Flory–Huggins quasi-binary systems seems to have been the one based on the separation factor σ. For instance, it led to the formulation of a single self-contained equation for the cloud point,[9] whose analysis revealed the existence of and criteria for multiple critical points,[10] and clarified their role in the mechanism of multiphase separations.[11,12] It also simplified numerical solution of cloud points by avoiding multiple iterations commonly employed by other authors. We wish to demonstrate that similar benefits are provided by applying an analogous σ-based analysis to the more complex problem of "true" ternary systems whose polymeric components differ also by the chemical nature of their monomer units.

In this paper the fundamentals of the new approach are presented and the conditions for the critical state are derived. Particular attention is paid to the interpretation of the expression for the critical concentration (that is formally separable from the triplet of interaction parameters g_1, g_2, and g_x) and to the so-called Scott systems and their place in the spectrum of all other cases. Other aspects of ternary equilibria such as the discussion of critical temperatures, criteria for double and triple critical points,

0024-9297/86/2219-1166$01.50/0