(2) Armstrong, D. W.; Bui, K. H. *Anal. Chem.* **1982**, *54*, 706.
(3) Boehm, R. E.; Martire, D. E.; Armstrong, D. W.; Bui, K. H. *Macromolecules* **1983**, *16*, 466.
(4) Bui, K. H.; Armstrong, D. W.; Boehm, R. E. *J. Am. Chem. Soc.*, submitted for publication.
(5) Hill, T. L. "An Introduction to Statistical Thermodynamics"; Addison-Wesley: Reading, MA, 1960; Chapters 14, 19, and 21.
(6) Carnahan, N. F.; Starling, K. E. *J. Chem. Phys.* **1969**, *51*, 635.
(7) Barker, J. A.; Henderson, D. *Rev. Mod. Phys.* **1976**, *48*, 587 and references contained therein.
(8) Flory, P. J. "Principles of Polymer Chemistry"; Cornell University Press: Ithaca, NY, 1971.
(9) Tagami, Y.; Casassa, E. F. *J. Chem. Phys.* **1969**, *50*, 2206.
(10) Yamakawa, H. "Modern Theory of Polymer Solutions"; Harper and Row: New York, 1971; Chapter 4.
(11) Flory, P. J.; Krigbaum, W. R. *J. Chem. Phys.* **1950**, *18*, 1086.
(12) A slightly improved equation of state for hard disks has been developed. See: Henderson, D. *Mol. Phys.* **1975**, *30*, 971.
(13) Schoenmakers, P. J.; Billiet, H. A. H.; Tijssen, R.; De Galan, L. *J. Chromatogr.* **1978**, *149*, 519.

Changes in Unperturbed Dimensions upon Formation of Regular β Meanders

Wayne L. Mattice

Department of Chemistry, Louisiana State University, Baton Rouge, Louisiana 70803.
Received April 18, 1983

ABSTRACT: Configuration partition functions have been formulated for the formation of regular β meanders by homopolypeptides. A regular β meander is a segment of pleated sheet containing three antiparallel strands that are of the same length and in register. The strands are assumed to be connected by tight bends. Generator matrix formalism is used to assess the mean square unperturbed radii of gyration. Radii of gyration of short chains decrease upon formation of β meanders. The extent of the reduction is determined primarily by the number of the residues in the β meanders; it is only slightly dependent upon the cooperativity of the ordering process. For long chains, however, the dimensions are strongly dependent on the cooperativity. At any degree of polymerization, the lowest dimensions are obtained at high β meander content and low transition cooperativity. Unperturbed mean square radii of gyration are examined for chains containing the same number of residues as chicken and T4 lysozymes. Squared radii of gyration decrease by about 10% when each chain contains average β meanders similar to those found in the conformations adopted by these proteins in the cystalline state.

Helices[1] and pleated sheets[2] are two types of ordered structures commonly found in proteins and synthetic polypeptides. While both structures have been subjects of numerous investigations, the process of helix formation is much better understood than is the origin of the pleated sheet.[3] The helix–coil transition in homopolypeptides can be studied in dilute solution by classic techniques used to provide a physical characterization of macromolecules.[4,5] Comparable studies of homopolypeptides undergoing pleated sheet formation proceed with much greater difficulty because aggregation usually accompanies formation of this ordered structure.[6,7]

Backbone hydrogen bonding in the classic α helix[1] always occurs between amino acid residues i and $i + 4$. However, if amino acid residue i is in a pleated sheet, it might be hydrogen bonded to virtually any other amino acid residue in the polypeptide chain. Consequently a theoretical description of the helix–coil transition can focus on interactions of comparatively short range, but a comprehensive treatment of pleated sheet formation must address potentially important interactions of much longer range. For this reason the theory of pleated sheet formation remains in a rather rudimentary form while helix–coil transition theory has become quite sophisticated.[3]

The situation just described is brought well into focus by consideration of a specific example. Poly(L-glutamic acid) undergoes a pH-induced helix–coil transition in aqueous media. The intrinsic viscosity at the midpoint of the transition is significantly lower than that of either the helical or completely disordered form.[5] This observation implies that the mean square radius of gyration must pass through a minimum as the conformational transition occurs. Theoretical support has been provided for this conclusion. Using helix–coil transition theory,[8] the measured unperturbed dimensions of disordered homopolypeptides containing CH_2R side chains[9,10] and generator matrix formalism,[11] Miller and Flory[12] were able to show that a minimum in the unperturbed dimensions would occur if the cooperativity parameter, σ, was not too small. The change in unperturbed dimensions has been analyzed in terms of the principal moments of the gyration tensor.[13] Poly(L-glutamic acid) can also form a β structure in aqueous solution,[14] but detailed studies are not available. On the one hand, experimental characterization is rendered difficult by aggregation. On the other hand, theoretical computation of the unperturbed dimensions is made difficult by the important role played by long-range interactions.

The objective here is characterization of changes in the mean square radius of gyration upon formation of a specific type of pleated sheet in homopolypeptides. Restriction to a special class of pleated sheet is required so that consequences of important long-range interactions can be handled in a satisfactory manner. The pleated sheet dealt with here is among those described by Schultz[15] as a β meander. A β meander is a sheet comprised of three antiparallel strands. The β meander has been identified in crystal structures of several proteins, including both chicken[16] and T4[17] lysozyme. A further restriction that the three strands be in register and connected by tight turns will be imposed here.

Given these restrictions, one can evaluate the manner in which the mean square unperturbed radius of gyration depends on pleated sheet content for homopolypeptides of varying degree of polymerization. Several aspects of the calculations are of particular interest. The first objective

0024-9297/84/2217-0407$01.50/0

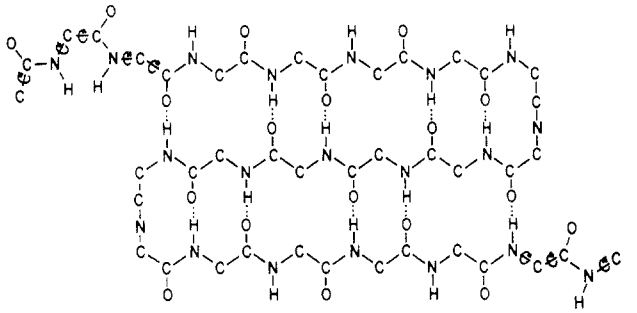

Figure 1. Diagrammatic representation of a portion of a polypeptide chain folded into a β meander. Substituents on the C^α atoms have been omitted.

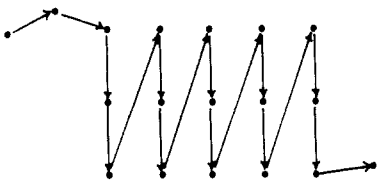

Figure 2. Virtual bond model for the chain configuration depicted in Figure 1.

is determination of whether β-meander formation causes the unperturbed dimensions to pass through a minimum, as is already known to be the case for the helix–coil transition. A second objective is assessment of the amount of β-meander formation permitted without significant change in $\langle r^2\rangle_0/\langle s^2\rangle_0$, the ratio of the mean square unperturbed end-to-end distance and radius of gyration. A third objective is characterization of values of the statistical weights required to produce β meanders comparable to those identified in crystalline globular proteins. The final objective is assessment of the extent of chain collapse when the sheet content resembles that found in typical globular proteins.

Formulation of the Configuration Partition Function and Unperturbed Dimensions

Virtual Bond Scheme. Figure 1 depicts a portion of a polypeptide chain folded into a β meander. The two bends are drawn with the hydrogen-bonding pattern described for a β bend.[18] Maintenance of the β meander places severe restrictions on ϕ and ψ for most of the amino acid residues shown. Circular arrows denote N–C^α and C^α–C′ bonds about which appreciable freedom of rotation is retained. Clearly, those residues that have severe restraints placed on ϕ and ψ should be classified as being part of the sheet. A somewhat more arbitrary decision is required in classification of the first flexible residue at the junction between sheet and random coil. In view of its flexibility, this residue might be classified as being a random coil. However, each of these residues participates in hydrogen bond formation and might therefore be considered to be part of the sheet. Unless specifically stated otherwise, the latter point of view will be adopted here. With this definition, Figure 1 depicts a β meander containing 18 amino acid residues, 6 in each of three strands.

A more diagrammatic representation of this structure is depicted in Figure 2. Here, positions of the C^α atoms are denoted by small filled circles. These circles could be connected by virtual bonds in the order in which amino acid residues occur in the polypeptide chain. With standard peptide geometry, all such virtual bonds would have a length of 380 pm.[1,2] A different virtual bond pattern has been employed in Figure 2. The C^α atoms at equivalent positions in the three strands are connected by two successive virtual bonds. These two virtual bonds are followed by a much longer one that returns the pattern to the initial strand, but moved one position further along the sheet.

Virtual bonds depicted in Figure 2 can be divided into three kinds on the basis of their lengths. Virtual bonds have a length of 380 pm if they connect two C^α atoms in amino acid residues that retain appreciable freedom of rotation about their N–C^α and C^α–C′ bonds.[9] Virtual bonds involving a C^α atom in the middle strand are assigned a length of 473 pm, which is the interchain distance in the pleated sheet formed by poly(L-alanine) in the solid state.[19] This virtual bond length must be in error by about 100 pm for those peptide units involved in the bends. The approximation used can be justified here because the slight error in positioning of C^α atoms at the bends had a trivial effect on computed values of the radius of gyration. The longest virtual bonds, which connect C^α atoms in the external strands, are assigned a length of 1007 pm. This is the length specified for a planar sheet with interchain distance 473 pm and unit height 344 pm. The values cited are those found for the pleated sheet formed by poly(L-alanine) in the solid state.[19] The poly(L-alanine) sheet is not really planar but is instead pleated. If the pleat were to be taken into account, the length of this virtual bond would be increased from 1007 to 1019 pm. Considering the sheet to be planar, and thereby ignoring the pleat, has little consequence for the computed values of the mean square radius of gyration.

Transformation Matrices. Averaging of the unperturbed dimensions over all configurations requires the use of matrices that transform a vector from its representation in the local coordinate system for virtual bond $i+1$ to its representation in the local coordinate system of virtual bond i. The local coordinate system for virtual bond i has its X axis running along that virtual bond, the Y axis in the plane of virtual bonds i and $i-1$, with a positive projection on virtual bond $i-1$, and the Z axis completing a right-handed coordinate system.[11] The structure described above for the β meander dictates that the identity matrix, $\mathbf{E}_3$, serve as the transformation matrix when virtual bonds i and $i+1$ are both of length 473 pm. The transformation matrix $\mathbf{T}_s$, which is required when either virtual bond i or $i+1$ is of length 1007 pm, is formulated with a "bond angle" of 20°.

$$\mathbf{T}_s = \begin{bmatrix} -\cos 20^\circ & \sin 20^\circ & 0 \\ \sin 20^\circ & \cos 20^\circ & 0 \\ 0 & 0 & -1 \end{bmatrix} \qquad (1)$$

Averaged transformation matrices are required when flexibility exists about bonds to the intervening C^α atom. The conformational energy surface described by Brant et al.[10] will be used for these amino acid residues. Several different averaged transformation matrices may result. If both virtual bonds involved in the transformation are of length 380 pm, the transformation matrix is that appropriate for a residue in an unperturbed poly(L-alanine) random coil. This matrix is denoted $\langle\mathbf{T}\rangle_{Ala}$. Averaging over the conformational energy surface at 10° intervals for ϕ and ψ, using $T = 300$ K, yields

$$\langle\mathbf{T}\rangle_{Ala} = \begin{bmatrix} 0.51 & 0.21 & 0.61 \\ -0.064 & -0.61 & 0.22 \\ 0.66 & -0.24 & -0.31 \end{bmatrix} \qquad (2)$$

Elements in this matrix differ slightly from those reported by Brant et al.[10] because we have sampled the conformational energy surface with smaller intervals for ϕ and ψ.

A different averaged transformation matrix is required if the transformation is from a virtual bond of length 473

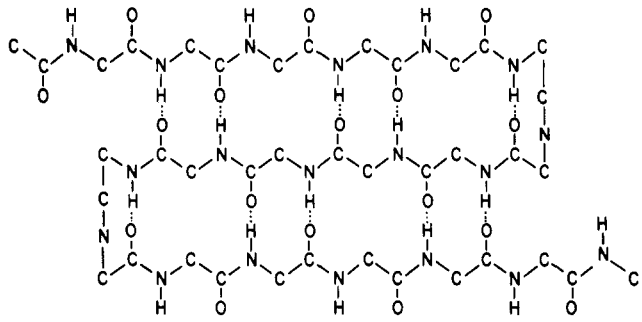

Figure 3. Diagrammatic representation of a β meander in which the first amino acid residue involved in hydrogen bond formation does so with N–H.

pm to one of length 380 pm. This transformation occurs at the beginning of the β meander. The conformational energy surface determined by Brant et al.[10] will tend to direct the random coil away from the β meander. Consequently, it will be assumed that the conformational energy surface provides an adequate approximation to the behavior of the disordered residue immediately adjacent to the sheet. The resulting transformation matrix is

$$\langle \mathbf{T} \rangle_{b,CO} = \begin{bmatrix} -0.13 & 0.54 & 0.61 \\ 0.60 & -0.16 & 0.22 \\ 0.34 & 0.62 & -0.31 \end{bmatrix} \tag{3}$$

if the β meander has the structure depicted in Figure 1. A different averaged transformation matrix would be required if the β meander were to begin in the manner depicted in Figure 3. In this figure, the first amino acid residue involved in hydrogen bond formation does so with its N–H. In contrast, the equivalent amino acid residue in Figure 1 uses its C′–O to form the first hydrogen bond. Thus the nature of the transformation from the first virtual bond of length 473 pm to its predecessor depends upon the structure of the β meander. For the structure depicted in Figure 3, this transformation is

$$\langle \mathbf{T} \rangle_{b,NH} = \begin{bmatrix} 0.072 & 0.55 & -0.61 \\ -0.58 & -0.22 & -0.22 \\ -0.41 & 0.57 & 0.31 \end{bmatrix} \tag{4}$$

A similar situation arises at the other end of the β meander. For the structure depicted in Figure 1, this transformation is

$$\langle \mathbf{T} \rangle_{e,NH} = \begin{bmatrix} 0.20 & 0.65 & -0.055 \\ 0.48 & 0.044 & 0.65 \\ 0.66 & -0.24 & -0.31 \end{bmatrix} \tag{5}$$

but in Figure 3 it would be

$$\langle \mathbf{T} \rangle_{e,CO} = \begin{bmatrix} -0.14 & 0.64 & -0.13 \\ 0.47 & -0.32 & -0.50 \\ -0.66 & -0.24 & -0.31 \end{bmatrix} \tag{6}$$

Formulation Based on Single Residues. The completely disordered chain is assigned a statistical weight of 1. A β meander with i residues in each strand has a statistical weight that can be written in first approximation as $(\tau t^3)^i b^2$. Each set of three residues connected by virtual bonds of length 473 pm contributes a factor τt^3. These three amino acid residues form two hydrogen bonds. Thus τt^3 has a contribution of t from each of the two amino acid residues arbitrarily considered to complete hydrogen bond formation, and a factor of τt from the remaining amino acid residue. A factor b arises from each of the two bends.

Further consideration is required in order to properly account for flexible residues at junctions between a β meander and regions of random coil. The preceding scheme attributes a statistical weight of τt^3 to amino acid residues 3–5 in Figure 2, and it assigns the same statistical weight to residues 6–8. Reference to Figure 1 brings this assignment into question because flexibility is present at amino acid residue 3 but not at amino acid residue 6. A similar problem arises at the other end of the meander. With a factor denoted a to correct for the flexibility of each of the amino acid residues at the ends of the β meander, the statistical weight becomes $(\tau t^3)^i (ab)^2$. For present purposes this statistical weight will be condensed to $A^i B$, where $A = \tau t^3$ and $B = (ab)^2$.

This scheme attributes identical statistical weights to the two β meanders depicted in Figures 1 and 3. No difference in weighting is allowed due to the types of hydrogen bonds formed by the first and last amino acid residues in the β meander. While such refinements might eventually become of interest, they will not be employed here.

The value of t is expected to be dependent on solvent and temperature, just as is its analogue, s, in the treatment of the transition from random coil to α helix.[8] The loss in configurational entropy upon immobilization of an amino acid residue should cause τ and b to be less than 1. One further expects $b < \tau$ because the ϕ,ψ adopted by amino acid residues in the pleated sheet formed by poly(L-alanine) in the solid state[19] fall in a region of lower conformational energy than do the ϕ,ψ required for formation of a sharp bend.[18,20,21] The parameter a is expected to have a value near $1/\tau$. Thus B is expected to be less than 1, and solvent- or temperature-induced conformational transitions result from changes in A.

A formulation for the configuration partition function, Z, which incorporates the above weighting scheme and is also in a form amenable to use in computation of the mean square unperturbed dimensions, is

$$Z = \text{row } (1, \mathbf{0})\mathbf{U}^{n+1} \text{ col } (1, \mathbf{0}) \tag{7}$$

$$\mathbf{U} = \begin{bmatrix} 1 & AB & 0 & 0 & AB & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & A & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & A & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{8}$$

There are n virtual bonds ($n + 1$ amino acid residues) in the homopolypeptide. A rectangular null matrix is denoted by $\mathbf{0}$. Its dimensions are determined by the requirement that matrices be conformable for multipication.

Computation of the unperturbed mean square end-to-end distance, $\langle r^2 \rangle_0$, and mean square radius of gyration, $\langle s^2 \rangle_0$, would require supermatrices in which each element in $\mathbf{U}$ is expanded through multiplication by a matrix of dimensions 5×5 (for $\langle r^2 \rangle_0$) or 7×7 (for $\langle s^2 \rangle_0$).[11] These matrices are

$$\begin{bmatrix} 1 & 2\mathbf{l}^T\mathbf{T} & l^2 \\ \mathbf{0} & \mathbf{T} & \mathbf{l} \\ 0 & \mathbf{0} & 1 \end{bmatrix} \tag{9}$$

$$\begin{bmatrix} 1 & 1 & 2\mathbf{l}^T\mathbf{T} & l^2 & l^2 \\ 0 & 1 & 2\mathbf{l}^T\mathbf{T} & l^2 & l^2 \\ \mathbf{0} & \mathbf{0} & \mathbf{T} & \mathbf{l} & \mathbf{l} \\ 0 & 0 & \mathbf{0} & 1 & 1 \\ 0 & 0 & \mathbf{0} & 0 & 1 \end{bmatrix} \tag{10}$$

Table I
Triplets Used in Z_3

Q_{ij}	transformation matrices	virtual bond lengths, pm	statistical weight	may be followed by
Q_{11}	$\langle T\rangle_{Ala}\langle T\rangle_{Ala}\langle T\rangle_{Ala}$	380, 380, 380	1	Q_{11}, Q_{12}, Q_{13}
Q_{12}	$\langle T\rangle_{b,CO}E_3T_s$	380, 473, 473	AB	Q_{21}, Q_{23}
Q_{13}	$\langle T\rangle_{b,NH}E_3T_s$	380, 473, 473	AB	Q_{31}, Q_{32}
Q_{23}	$T_sE_3T_s$	1007, 473, 473	A	Q_{31}, Q_{32}
Q_{32}	$T_sE_3T_s$	1007, 473, 473	A	Q_{21}, Q_{23}
Q_{21}	$T_sE_3\langle T\rangle_{e,CO}$	1007, 473, 473	A	Q_{11}, Q_{12}, Q_{13}
Q_{31}	$T_sE_3\langle T\rangle_{e,NH}$	1007, 473, 473	A	Q_{11}, Q_{12}, Q_{13}

In these matrices, $\mathbf{l} = \text{col}\,(l, \mathbf{0})$, $\mathbf{l}^T$ is the transpose, and l denotes the length of a virtual bond. The lengths, in pm, corresponding to each nonzero element in the statistical weight matrix are

$$\begin{bmatrix} 380 & 380 & 0 & 0 & 380 & 0 & 0 \\ 0 & 0 & 473 & 0 & 0 & 0 & 0 \\ 473 & 0 & 0 & 473 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1007 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 473 & 0 \\ 473 & 0 & 0 & 0 & 0 & 0 & 473 \\ 0 & 1007 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (11)$$

Similarly, the transformation matrices employed are

$$\begin{bmatrix} \langle T\rangle_{Ala} & \langle T\rangle_{b,CO} & 0 & 0 & \langle T\rangle_{b,NH} & 0 & 0 \\ 0 & 0 & E_3 & 0 & 0 & 0 & 0 \\ \langle T\rangle_{e,CO} & 0 & 0 & T_s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & T_s & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & E_3 & 0 \\ \langle T\rangle_{e,NH} & 0 & 0 & 0 & 0 & 0 & T_s \\ 0 & T_s & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (12)$$

Grouping in Triplets. The above formulation permits β meanders containing as few as three amino acid residues. Since each strand should contain at least two residues, the smallest β meander should have six residues. The generator matrices, of dimensions 35 × 35 (for $\langle r^2\rangle_0$) and 49 × 49 (for $\langle s^2\rangle_0$), are also quite large in the above formulation. These two problems can be alleviated in an alternative formulation that considers the amino acid residues in groups of three. The new formulation introduces a new approximation in that it retains in the configuration partition function only those configurations in which β meanders begin at residues 2, 5, ..., $n - 4$. The effect of this approximation will be brought out below.

Define a generator matrix $\mathbf{Q}$ as

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{11} & AB\mathbf{Q}_{12} & AB\mathbf{Q}_{13} \\ A\mathbf{Q}_{21} & \mathbf{0} & A\mathbf{Q}_{23} \\ A\mathbf{Q}_{31} & A\mathbf{Q}_{32} & \mathbf{0} \end{bmatrix} \quad (13)$$

Its dimensions are 15 × 15 (for computation of $\langle r^2\rangle_0$) or 21 × 21 (for computation of $\langle s^2\rangle_0$). Each $\mathbf{Q}_{ij}$ is formulated as the product of three matrices of the type specified in (9) or (10). Virtual bond lengths and transformation matrices, in sequential order, are specified in Table I. Thus $\mathbf{Q}_{11}$ represents the contribution of three successive residues in the random coil state. The remaining $\mathbf{Q}_{ij}$ represent the contribution by groups of three residues in the β meander. This group is at the beginning of the β meander in $\mathbf{Q}_{12}$ and $\mathbf{Q}_{13}$, at the end in $\mathbf{Q}_{21}$ and $\mathbf{Q}_{31}$, and at an internal position in $\mathbf{Q}_{23}$ and $\mathbf{Q}_{32}$. The latter two must be differentiated from one another so that the hydrogen-bonding pattern of the last residue in the β meander is consistent with the size of the sheet and the hydrogen-bonding pattern adopted by the first amino acid residue.

The configuration partition function suggested by the formulation in eq 13 is

$$Z_3 = \text{row}\,(1, \mathbf{0}) \begin{bmatrix} 1 & AB & AB \\ A & 0 & A \\ A & A & 0 \end{bmatrix}^{n/3} \text{col}\,(1, \mathbf{0}) \quad (14)$$

which in turn suggests

$$\langle r^2\rangle_0 = Z_3^{-1}\,\text{row}\,(1, \mathbf{0})\mathbf{Q}^{n/3}\,\text{col}\,(0, 0, 0, 0, 1, \mathbf{0}) \quad (15)$$

$$\langle s^2\rangle_0 = Z_3^{-1}(n+1)^{-2}\,\text{row}\,(1, \mathbf{0})\mathbf{Q}^{n/3}\,\text{col}\,(0, 0, 0, 0, 0, 0, 1, \mathbf{0}) \quad (16)$$

where the $\mathbf{Q}$'s in eq 15 and 16 are the 15 × 15 and 21 × 21 versions, respectively. Considerable numerical labor is saved for two reasons: First, the matrices are of much smaller dimensions than those required in absence of grouping of the residues in triplets. Second, the exponents for the matrices are $n/3$ rather than n.

An added bonus arising from the grouping of residues in triplets is that the smallest allowed β meander now contains six residues. It arises from $\mathbf{Q}_{12}\mathbf{Q}_{21}$ or $\mathbf{Q}_{13}\mathbf{Q}_{31}$. In contrast, the formulation that does not utilize grouping allowed a β meander containing as few as three residues. For this reason the formulation expressed in eq 14–16 may be more realistic than that based on the larger matrices. On the other hand, the more compact formulation has introduced new restrictive features. It is applicable only to chains in which n is divisible by three, and a β meander can only begin at residues 2, 5, ..., $n - 4$. The consequences of these new restrictions will be brought out in illustrative calculations described below.

It is worthwhile noting that the expression for Z_3 in eq 14 utilizes a statistical weight matrix that is equivalent to that frequently used for a bond in a chain with a symmetric threefold rotation potential.[11]

$$A\begin{bmatrix} \tau & \sigma & \sigma \\ 1 & \sigma\psi & \sigma\omega \\ 1 & \sigma\omega & \sigma\psi \end{bmatrix} \quad (17)$$

The matrices are identical if $\tau = 1/A$, $\sigma = B$, $\psi = 0$, and $\omega = 1/B$. By taking advantage of the symmetry of the matrix,[11] we obtain

$$Z_3 = \text{row}\,(1, 0) \begin{bmatrix} 1 & 2AB \\ A & A \end{bmatrix}^{n/3} \text{col}\,(1, 0) \quad (18)$$

which is numerically identical with the result specified by eq 14. The eigenvalues for the 2 × 2 matrix in eq 18 are

$$\lambda_1, \lambda_2 = \frac{1}{2} + A/2 \pm [(1 - A)^2 + 8A^2B]^{1/2}/2 \quad (19)$$

and, for infinitely long chains, the probability that a residue will be part of a β meander is

$$p_\beta = \partial \ln \lambda_1/\partial \ln A = (\lambda_1 + 1 - \lambda_2)(\lambda_1 - 1)/[\lambda_1(\lambda_1 - \lambda_2)] \quad (20)$$

The average number of residues in these β meanders is

$$\nu = 3(\partial \ln \lambda_1/\partial \ln A)/(\partial \ln \lambda_1/\partial \ln B) = 3(\lambda_1 + 1 - \lambda_2)/(1 - \lambda_2) \quad (21)$$

Matrix expressions for p_β and ν, which are valid at any n, are[11]

$$p_\beta = 3(n + 1)^{-1}Z_3^{-1} \times \text{row}(1, 0) \begin{bmatrix} 1 & 2AB & 0 & 2AB \\ A & A & A & A \\ 0 & 0 & 1 & 2AB \\ 0 & 0 & A & A \end{bmatrix}^{n/3} \text{col}(0, 1, 0) \quad (22)$$

$$\nu = (1 + n)p_\beta/x \quad (23)$$

$$x = Z_3^{-1} \times \text{row}(1, 0) \begin{bmatrix} 1 & 2AB & 0 & 2AB \\ A & A & 0 & 0 \\ 0 & 0 & 1 & 2AB \\ 0 & 0 & A & A \end{bmatrix}^{n/3} \text{col}(0, 1, 0) \quad (24)$$

The condensation effected in eq 18 becomes inaccessible if a more detailed set of statistical weights were employed. The last two elements in the top row of the statistical weight matrix in eq 14 become nonidentical if the statistical weight of a sheet depends upon whether the first amino acid residue uses N–H or C–O in hydrogen bond formation. Similarly, the last two elements in the first column become nonidentical if that statistical weight depends upon whether the last amino acid residue uses N–H or C–O in hydrogen bond formation. Either refinement would be incompatible with the condensation effected in going from eq 14 to eq 18.

Strategy employed in formulation of Z is similar in some respects to the treatment of helical hairpins by Poland and Scheraga.[22] Statistical weights, which they denote by s and $\sigma^\dagger$ for a helical hairpin of two strands, have their counterparts in A and B, respectively, in regular β meanders. The present treatment differs in that the β meander can begin in two different ways (Figures 1 and 3) and geometric considerations are incorporated to enable computation of $\langle s^2 \rangle_0$.

Conformational Transitions Specified by Z and Z_3

Behavior of systems described by the configuration partition functions defined in eq 7 and 14 is illustrated by transitions produced through changes in A at constant B and n. Results obtained with chains comprised of 193 and 1537 amino acid residues are depicted in Figure 4.

Consideration is given first to the transition in the longer chain when B has the small value of 10^{-4}. Both configuration partition functions yield a highly cooperative transition centered near $A = 1$ (Figure 4B). The transition obtained with Z_3 has a higher cooperativity than that obtained with Z. The formulation of Z permits β meanders that originate at amino acid residues 1, 2, ..., $n - 1$, but Z_3 only allows the origin of a β meander at amino acid residues 2, 5, ..., $n - 4$. Since Z_3 offers fewer possibilities for the coexistence of random coil and β meander within a given configuration, it is expected that it will produce a more highly cooperative transition than does Z. Differences between transitions specified by Z and Z_3 are somewhat larger than those typically seen in applications of coarse graining[23] to the helix–coil transition.[24,25] Presumably, this difference arises because chh, cch, hcc, and hhc are allowed triplet blocks in applications to the helix–coil transition. Mixed triplets are not permitted in the present treatment of regular β meanders because the geometry is handled in a manner that assumes the number of residues in each ordered region is divisible by 3.

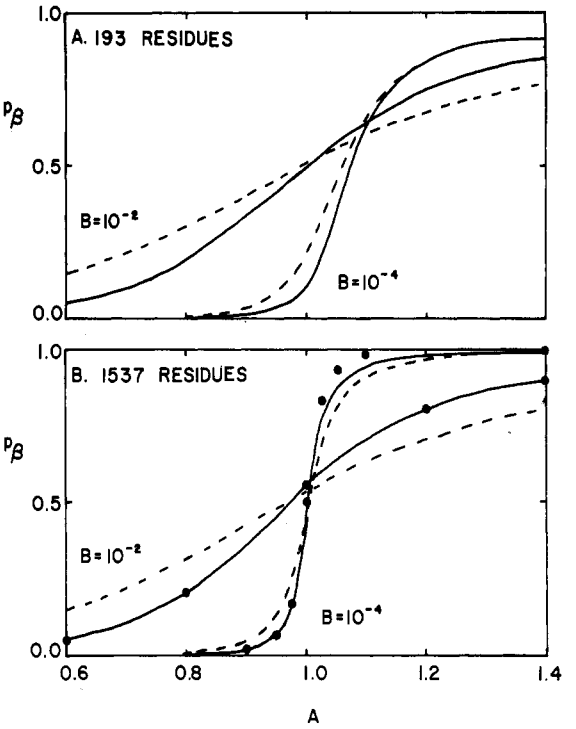

Figure 4. Fraction of the residues in the β meander for chains of 193 residues (A) or 1537 residues (B). Solid lines depict results obtained with Z_3, while dashed lines denote those obtained with Z. Filled circles in (B) denote the asymptotic limit at infinite n, as evaluated via eq 20.

The fact that Z permits the occurrence of a β meander with as few as three amino acid residues, while for Z_3 a β meander must contain at least six amino acid residues, is of little consequence when $B = 10^{-4}$. The average number of residues in a β meander is 49 from Z and 57 from Z_3 for a chain of 1537 amino acid residues, $A = 0.95$, $B = 10^{-4}$. The β meanders become larger as A increases. Consequently, both configuration partition functions show that the β meanders that play the important role in the transition contain many more than 3–6 amino acid residues.

If B is retained at 10^{-4} but the chain is shortened to 193 amino acid residues, transitions specified by both configuration partition functions become somewhat less cooperative and are shifted to higher A (Figure 4A). The more highly cooperative transition specified by Z_3 experiences the larger shift, as expected. Large β meanders still dominate the transition. For example, at $A = 1$, the average number of amino acid residues in a β meander is 57 (from Z) or 65 (from Z_3).

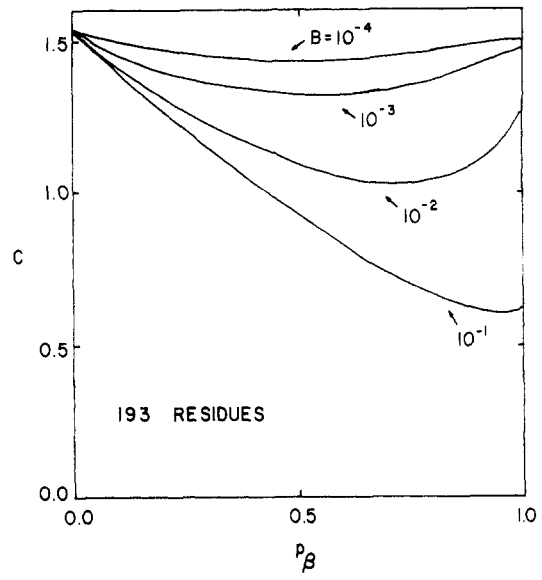

Figure 5. Unperturbed dimensions as a function of p_β for a chain containing 193 amino acid residues.

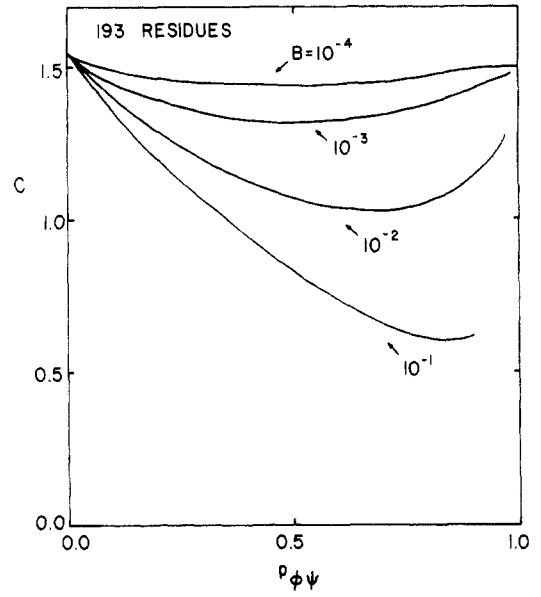

Figure 6. Unperturbed dimensions as a function of $p_{\phi\psi}$ for a chain containing 193 amino acid residues.

Figure 4 also depicts transitions obtained when B is assigned the larger value of 10^{-2}. As a result of the reduced cooperativity, there is little change in the transition upon lengthening the chain from 193 to 1537 amino acid residues. At both chain lengths the more highly cooperative transition is obtained with Z_3. Small β meanders now play a significant role in the transition. For example, when A is 0.6, the average number of amino acid residues in a β meander is 6.6–6.7 (from Z) or 10.1–10.3 (from Z_3). Consequently, the fact that Z permits occurrence of a β meander containing only three residues, while Z_3 demands that the smallest β meander contain six residues, now contributes to the differences in the transitions.

Filled circles in Figure 4B denote p_β calculated for the infinite chain via eq 20. These circles are essentially coincident with the curve calculated from eq 22 when $B = 10^{-2}$. If B is as small as 10^{-4}, limiting behavior is not obtained until the chain contains more than 1537 amino acid residues.

Changes in Unperturbed Dimensions

Figure 5 depicts dependence of the characteristic ratio on p_β for a chain of 193 amino acid residues. The characteristic ratio, C, is defined as $\langle s^2\rangle_0/nl_p^2$, where l_p = 380 pm and $\langle s^2\rangle_0$ is obtained via eq 16. At this degree of polymerization, nearly identical $\langle s^2\rangle_0$ are obtained for the random coil and for the chain folded into a single β meander. At vanishingly small B, these configurations would completely dominate the configuration partition function at any A, and the unperturbed dimensions would therefore be independent of p_β. The results depicted in Figure 5 show that a close approximation to this situation is obtained when B is as large as 10^{-4}.

Smaller β meanders are more compact than a random coil segment containing the same number of amino acid residues. Small β meanders play an increasingly important role in the conformational transition as B increases. Consequently, $\partial C/\partial p_\beta$, evaluated at $p_\beta = 0$, becomes more strongly negative as B increases from 10^{-4} to 10^{-1}. The characteristic ratio goes through a minimum, which becomes deeper and displaced to higher p_β as B increases.

Unperturbed dimensions depicted in Figure 5 depend strongly on B at large p_β. The origin of this dependence is revealed by examination of the definition of p_β. It is the fraction of the amino acid residues that participate in the hydrogen-bonding pattern of the β meander. Inspection of Figures 1 and 3 shows that, according to this definition, flexibility exists about ϕ and ψ at the first and last amino acid residues in each β meander. For vanishingly small B, the dominant chain configuration at the upper extreme for p_β is that single configuration in which amino acid residues 2 through $n + 1$ are part of the same β meander. With the exception of the residues 1, 2, and $n + 1$, this configuration is rigid. As B becomes larger, the dominant configurations at large p_β are those in which the chain contains several different β meanders without any intervening segments of random coil. Two residues with flexibility about ϕ,ψ occur at each junction between different meanders. The number of such junctions increases as B increases from 10^{-4} to 10^{-1}, causing the behavior seen in Figure 5 at large p_β.

Characteristic ratios depicted in Figure 5 are presented in an alternative fashion in Figure 6. Here they are shown as a function of $p_{\phi\psi}$, defined as the fraction of the amino acid residues that have lost rotational freedom at ϕ and ψ as a consequence of β-meander formation. Thus each β meander contributes two fewer amino acid residues to $p_{\phi\psi}$ than it does to p_β.

$$p_{\phi\psi} = (1 - 2/\nu)p_\beta \tag{25}$$

The upper limit for $p_{\phi\psi}$ is lower than the upper limit for p_β, with the displacement increasing as B increases from 10^{-4} to 10^{-1}. At $B = 10^{-1}$ and large A, the average β meander in a chain of 193 amino acid residues has about 21 residues involved in the hydrogen-bonding pattern. Only 19 of these 21 residues have severe restriction imposed on ϕ and ψ. Thus $p_{\phi\psi}$ terminates near 19/21.

The ratio $\langle r^2\rangle_0/\langle s^2\rangle_0$ is depicted as a function of p_β for chains composed of 193 amino acid residues in Figure 7. The completely disordered chain has a ratio slightly larger than 6 because the characteristic ratio determined by $\langle s^2\rangle_0$ converges more slowly to its asymptotic limit than does the characteristic ratio determined by $\langle r^2\rangle_0$.[9,10] The limiting value for $\langle r^2\rangle_0/\langle s^2\rangle_0$ at large p_β depends strongly on the cooperativity of the transition. This limit is close to 12 if the transition is highly cooperative. Under such

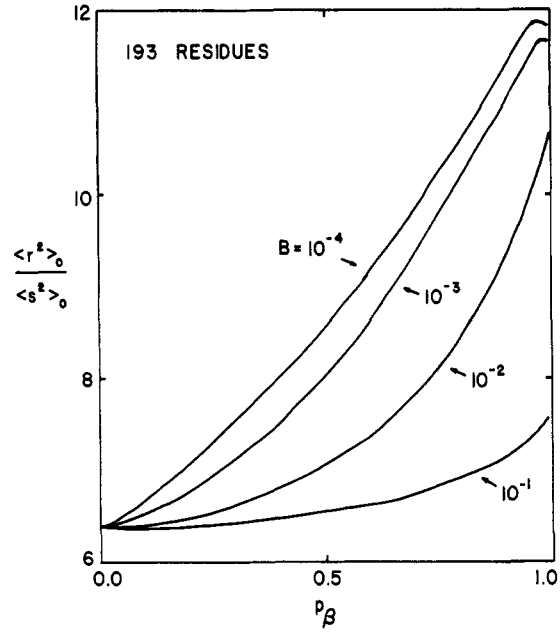

Figure 7. Ratio $\langle r^2\rangle_0/\langle s^2\rangle_0$ as a function of p_β for a chain containing 193 amino acid residues.

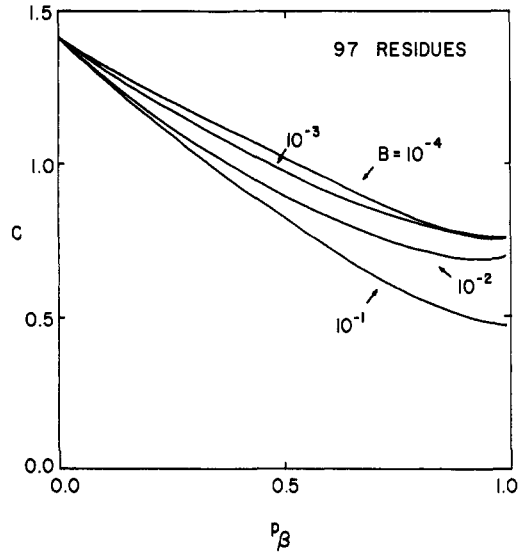

Figure 8. Unperturbed dimensions as a function of p_β for a chain containing 97 amino acid residues.

circumstances residues 2 through $n + 1$ are part of the same β meander. However, if B is 10^{-1}, $\langle r^2\rangle_0/\langle s^2\rangle_0$ remains smaller than 7.6 even at large p_β. The average chain contains nine β meanders at the largest p_β when B is 10^{-1}. Flexibility of the residues at the junctions causes the overall configuration to depart substantially from that of a rigid rod.

Figure 8 depicts the behavior of C for chains significantly shorter than those used in Figure 5. Unperturbed dimensions experience a nearly monotonic decrease as p_β increases. This decrease becomes more severe as B increases. However, the effect of B is less dramatic than was the case with the longer chain dealt with in Figure 5.

The behavior of much longer chains is depicted in Figure 9. The limiting characteristic ratios at high p_β, as well as the behavior at intermediate p_β, are now strongly dependent on B. If B is 10^{-1}, the characteristic ratio passes through a minimum near $p_\beta = 0.9$. This minimum is quite deep, with the change in C being greater than a factor of 2. When B is 10^{-4}, on the other hand, the minimum is

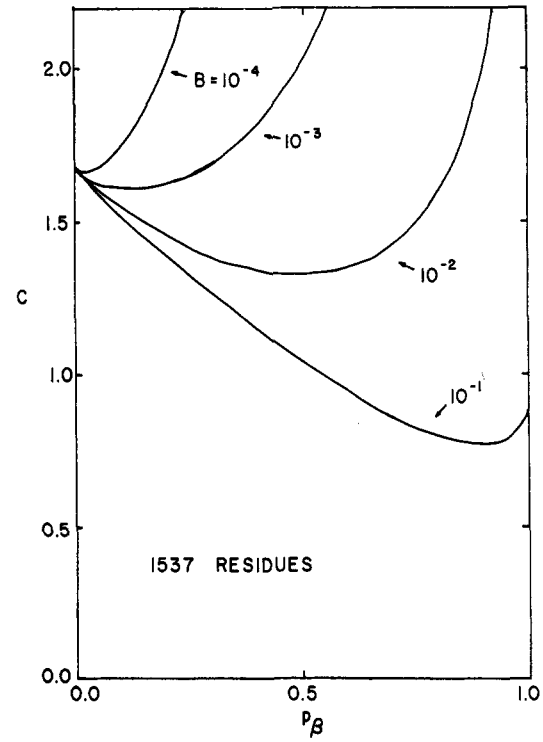

Figure 9. Unperturbed dimensions as a function of p_β for a chain containing 1537 amino acid residues.

extremely shallow and is displaced to small p_β. The dominant feature in the transition is now a sharp increase in C as p_β rises above 0.02. The sharp increase arises because extremely long β meanders dominate the configuration partition function when A and n are large and B is 10^{-4}.

Taken together, the behavior depicted in Figures 5, 8, and 9 leads to the following conclusions: For short chains, the unperturbed dimensions must decrease upon β-meander formation. Contraction of the chain is determined primarily by p_β; it is rather insensitive to plausible variation in B. For long chains, on the other hand, B exerts an increasingly important influence on the unperturbed dimensions as p_β increases. In any system likely to be of experimental interest, $\partial C/\partial p_\beta$ should be negative at small p_β. For any degree of polymerization likely to be of interest, the lowest unperturbed dimensions are obtained at large p_β and B.

Application to Lysozymes

Chicken[16] and T4[17] lysozyme contain a single β meander in the crystalline state. This section concerns homopolypeptides containing the same number of residues as each of the lysozymes. Statistical weights are adjusted so that the homopolypeptide has the same fraction of its residues in a β meander, with the average size of the β meander closely matching that found in the crystalline lysozyme. Since the remainder of the chain is unordered, the objective is not a modeling of globular lysozyme. Instead we wish to provide a qualitative estimate for B when the pleated sheet is to be of a size found in a globular protein.

Chicken lysozyme contains 129 amino acid residues, with residues 42–46, 50–54, and 58–60 involved in the sheet.[16] Residues 47–49 and 55–57 occur in the bends connecting strands of the β meander. If the residues in the bends are ignored, the sheet contains 13 residues and p_β is 0.101. Alternatively, if the residues in the bends are considered to be part of the β meander, it then contains 19 residues and p_β is 0.147. These values are shown in parentheses in the first two lines of Table II. The lysozyme from T4

Table II
Homopolypeptide Analogues of Lysozymes

residues	A	B	p_β [a]	ν [a]	contraction [b]
129	0.72	0.01	0.110 (0.101)	12.3 (13)	9.4
129	0.88	0.003	0.150 (0.147)	20.2 (19)	10.7
164	0.65	0.01	0.072 (0.073)	10.9 (12)	6.3
164	0.83	0.003	0.102 (0.104)	17.7 (17)	7.5

[a] See text for the origin of numbers in parentheses. [b] Percent contraction of $\langle s^2 \rangle_0$ from the value for the completely disordered chain.

contains 164 residues, and residues 18–20, 23–27, and 31–34 are in the sheet.[17] Bends are formed by residues 21–22 and 28–30. If the bends are neglected, the sheet contains 12 residues, yielding $p_\beta = 0.073$. Inclusion of the bends in the β meander increases ν to 17 and p_β to 0.104. These results for T4 are presented in parentheses in the last two lines of Table II.

Calculations using eq 22 and 24 yield average β meanders similar to those found in the lysozymes if B is in the range 0.003–0.01. The lower value for B is required to produce the larger β meanders that result from a definition which considers those residues in the bends to be part of the sheet. Recalling that B represents a compact notation for $(ab)^2$, this result implies that the effective ab is on the order of 0.1.

Mean square unperturbed radii of gyration were calculated via eq 16 for each combination of n, A, and B. In each case they were smaller than the dimensions calculated for the completely disordered chain containing the same number of amino acid residues. The percent reduction in C produced by formation of the β meander is shown in the last column of Table II. It is found to be in the range 6–11%. While formation of the β meander produces a contraction, it is only a small portion of the change accomplished in going from the random coil to the globular state. For example, the squared radius of gyration defined by the shape of chicken lysozyme in the crystalline state[26] is 93% smaller than the mean square radius of gyration of an unperturbed poly(L-alanine) random coil containing the same number of residues. Most of the reduction in dimensions in globular chicken lysozyme results from the conformations adopted by residues that are not involved in the β meander.

Qualitatively different conclusions might be forthcoming for certain other proteins. If the effective B is taken to be on the order of 0.01, extensive β-meander formation can lead to appreciable reduction in the unperturbed dimension of homopolypeptides containing the same numbers of residues as typical proteins (Figures 5, 8, and 9). For a B of this magnitude, a greater reduction would be possible with a chain of 10^2 residues than for a chain of 10^3 residues.

Refinements required to produce a more general treatment of sheet formation might also produce a greater contraction. Possible refinements include allowance of sheets in which the number of strands is different from three, letting some of the bends become loops, relaxation of the requirements that all strands be antiparallel and of the same length, providing for a twist to the sheet, and allowance of irregularities such as the β bulge.[27] Additional contraction would also result from formation of short helices.[12,13]

Acknowledgment. This work was supported by National Science Foundation Research Grant PCM 81-18197.

Registry No. Lysozyme, 9001-63-2.

References and Notes

(1) Pauling, L.; Corey, R. B.; Branson, H. R. *Proc. Natl. Acad. Sci. U.S.A.* **1951**, *37*, 205–211.
(2) Pauling, L.; Corey, R. B. *Proc. Natl. Acad. Sci. U.S.A.* **1951**, *37*, 729–735.
(3) Poland, D.; Scheraga, H. A. "Theory of Helix-Coil Transitions in Biopolymers"; Academic Press: New York, 1970.
(4) Doty, P.; Bradbury, J. H.; Holtzer, A. M. *J. Am. Chem. Soc.* **1956**, *78*, 947–954.
(5) Doty, P.; Wada, A.; Yang, J. T.; Blout, E. R. *J. Polym. Sci.* **1957**, *23*, 851–861.
(6) Maeda, H.; Kadono, K.; Ikeda, S. *Macromolecules* **1982**, *15*, 822–827.
(7) Auer, H. E.; Miller-Auer, H. *Biopolymers* **1982**, *21*, 1245–1259.
(8) Zimm, B. H.; Bragg, J. K. *J. Chem. Phys.* **1959**, *31*, 526–535.
(9) Brant, D. A.; Flory, P. J. *J. Am. Chem. Soc.* **1965**, *87*, 2788–2791.
(10) Brant, D. A.; Miller, W. G.; Flory, P. J. *J. Mol. Biol.* **1967**, *23*, 47–65.
(11) Flory, P. J. *Macromolecules* **1974**, *7*, 381–392.
(12) Miller, W. G.; Flory, P. J. *J. Mol. Biol.* **1966**, *15*, 298–314.
(13) Mattice, W. L. *Macromolecules* **1980**, *13*, 904–909.
(14) Rinaudo, M.; Domard, A. *J. Am. Chem. Soc.* **1976**, *98*, 6360–6364.
(15) Schultz, G. E. *J. Mol. Biol.* **1980**, *138*, 335–347.
(16) Blake, C. C. F.; Koenig, D. F.; Mair, G.; North, A. C. T.; Phillips, D. C.; Sarma, V. R. *Nature (London)* **1965**, *206*, 757–761.
(17) Matthews, B. W.; Remington, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 4178–4182.
(18) Venkatachalam, C. M. *Biopolymers* **1968**, *6*, 1425–1436.
(19) Arnott, S.; Dover, S. D.; Elliott, A. *J. Mol. Biol.* **1967**, *30*, 201–208.
(20) Chandrasekaran, R.; Lakshminarayanan, A. V.; Pandya, U. V.; Ramachandran, G. N. *Biochim. Biophys. Acta* **1973**, *303*, 14–27.
(21) Lewis, P. N.; Momany, F. A.; Scheraga, H. A. *Biochim. Biophys. Acta* **1973**, *303*, 211–229.
(22) Reference 3, pp 202–203.
(23) Crothers, D. M.; Kallenbach, N. R. *J. Chem. Phys.* **1966**, *45*, 917–927.
(24) Reference 3, pp 171–175.
(25) Skolnick, J.; Holtzer, A. *Macromolecules* **1982**, *15*, 303–314.
(26) Mattice, W. L.; Carpenter, D. K. *Biopolymers* **1977**, *16*, 81–94.
(27) Richardson, J. S.; Getzoff, E. D.; Richardson, D. C. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 2574–2578.