of the two minima located at $\phi \simeq \pm 120°$ in PA; instead, the energy of OCPP increases continously as ϕ departs from ±75° reaching two maxima[23] at $\phi = 0$ and 180°.

According to the results of Figure 4, $\mu \simeq 1.80$–1.85 D for $\phi \simeq \pm 75°$, which is in excellent agreement with the experimental values of 1.75 and 1.79 D measured in cyclohexane and benzene solutions, respectively. The two minima of energy are identical both in energy and in value of μ; consequently, as in the case of MCPP, the dipole moment of this molecule should be almost independent of temperature despite the strong variation of μ with ϕ. The experimental results indicate that μ is roughly independent of temperature when measured in cyclohexane solutions, while in benzene it seems to have a very small positive temperature coefficient which could be produced by little variations (ca. 1–2°) of the valence angles.

We therefore conclude that the theoretical values of the dipole moments of these four molecules are in excellent agreement with the experimental results. In the case of MCPP and OCPP, the agreement could be taken as a experimental support for the conformational analysis of PA, reported by Hummel and Flory.[10]

Acknowledgment. The partial support of this work by the CICYT, through Grant 87051 is grateful acknowledged.

Registry No. PP, 937-41-7; OCPP, 60202-89-3; MCPP, 6603-24-3; PCPP, 61469-49-6.

References and Notes

(1) San Román, J.; Madruga, E. L. *Eur. Polym. J.* **1982**, *18*, 481.
(2) San Román, J.; Madruga, E. L.; del Puerto, M. A. *J. Polym. Sci. Polym. Chem. Ed.* **1983**, *21*, 691.
(3) San Román, J.; Madruga, E. L.; del Puerto, M. A. *J. Polym. Sci., Polym. Chem. Ed.* **1983**, *21*, 3303.
(4) San Román, J.; Madruga, E. L.; Guzmán, J. *Polym. Commun.* **1984**, *25*, 373.
(5) Saiz, E.; Hummel, J. P.; Flory, P. J.; Plavsic, M. *J. Phys. Chem.* **1981**, *85*, 3211.
(6) McClellan, A. L. *Table of Experimental Dipole Moments*; Rahara: El Cerrito, CA, 1974; Vol. II.
(7) Aroney, M.; LeFevre, R. J. W.; Chang, S. S. *J. Chem. Soc.* **1960**, 3173.
(8) Exner, O.; Fidlerová, Z.; Jehlicka, V. *Collect. Czech. Chem. Commun.* **1968**, *33*, 2019.
(9) Krishna, B.; Bhargava, S. K.; Prakash, P. *J. Mol. Struct.* **1971**, *8*, 195.
(10) Hummel, J. P.; Flory, P. J. *Macromolecules* **1980**, *13*, 479.
(11) Patai, S.; Bentor, M.; Reichmann, M. E. *J. Am. Chem. Soc.* **1952**, *74*, 845.
(12) Riande, E. *J. Polym. Sci., Polym. Phys. Ed.* **1976**, *14*, 2231.
(13) Timmermans, J. *Physico-Chemical Constants of Pure Organic Compounds*; Elsevier: Amsterdam, 1965; Vols. 1, 3.
(14) Guggenhein, E. H. *Trans. Faraday Soc.* **1949**, *45*, 714.
(15) Smith, J. W. *Trans. Faraday Soc.* **1950**, *46*, 394.
(16) LeFèvre, R. J. W.; Sundaran, A. *J. Chem. Soc.* **1962**, 3904.
(17) Vij, J. K.; Srivastava, K. K. *Bull. Chem. Soc. Jpn.* **1970**, *43*, 2313.
(18) Farmer, D. B.; Holt, A.; Walker, S. *J. Chem. Phys.* **1966**, *44*, 4116.
(19) LeFèvre, R. J. W.; Radford, D. V.; Ritchic, G. L. D.; Stiles, P. J. *J. Chem. Soc. B* **1968**, 148.
(20) Jones, R. A. Y.; Katrizky, A. R.; Ochkin, A. V. *J. Chem. Soc. B* **1971**, 1795.
(21) See Figure 4 of ref 10.
(22) Conformational energy calculations were made with the same geometry and energy functions given in ref 10. Since energy minimization by allowing bond angle relaxation was not performed, unrealistic high values of energy when $\phi = 0$ or 180° due to the interactions of O* with either H or Cl were obtained. However, the energies calculated in the range $\phi = 40$–140° should be reasonably good since there are no strong interactions that require relaxation of either bond lengths or bond angles in this range of ϕ.
(23) A third difference between PA and OCPP molecules is the nonequivalence between $\phi = 0°$ and $\phi = 180°$ in OCPP. However, this difference is irrelevant for the present analysis.

Conformational Entropy Associated with the Formation of Internal Loops in Collagen

Wayne L. Mattice,† George Némethy,‡ and Harold A. Scheraga*,‡

Department of Polymer Science, University of Akron, Akron, Ohio 44325, and Baker Laboratory of Chemistry, Cornell University, Ithaca, New York 14853-1301.
Received November 19, 1987; Revised Manuscript Received February 16, 1988

ABSTRACT: The conformational entropy of formation of internal loops in the collagen triple helix is estimated. Computer simulation is used to generate random conformations of short polypeptide chains that are connected at both ends to unbroken triple helices. The internal disruption of the triple helix has been modeled as a trifunctional star (i.e. a molecule in which three disordered chains emanate from a single trifunctional branch point, not counting the triple helix itself) that is constrained so that the three ends remote from the branch point meet at a common site. Parameters describing short-range interactions in the branches of the trifunctional star are adjusted so that the characteristic ratio of a linear random chain agrees with that determined for denatured collagen. Excluded volume is introduced by considering atoms of the chain that participate in long-range interactions as hard spheres. The diameter of these spheres is selected so that the chain expansion is at the upper end of the observed range for denatured collagen in aqueous media. Excluded volume interactions reduce by no more than an order of magnitude, relative to the formation of three unrestricted random chains, the probability of formation of internal disruptions for loops containing 30 residues per branch.

Introduction

The triple-helical structure of collagen may contain localized regions of reduced stability, even under conditions where the triple-stranded helix is thermodynamically stable. Transient local unfolding of the triple-stranded helix may occur in such regions, resulting in the formation of internal loops by the three constituent chains. The probability of formation of such a loop depends on the local amino acid sequence, and hence it varies along the molecule. The frequency and location of prolyl and hydroxy-

* To whom correspondence should be addressed.
† University of Akron.
‡ Cornell University.

0024-9297/88/2221-2811$01.50/0

prolyl residues may play an especially important role in determining this probability.[1-4] Because of the limited conformational freedom of these two amino acid residues, stretches of the polypeptide chain that do not contain these residues are expected to be more flexible than other regions; hence, the formation of loops would be more favorable in such stretches than in other regions of the sequence. This relationship between the amino acid sequence and the ease of formation of the loop has been proposed repeatedly in a qualitative way, on the basis of experimental evidence.[1-4] The possible existence of highly localized internal loops is suggested by the observation that limited enzymatic proteolytic cleavage of triple-stranded type I collagen by mammalian collagenase,[1-6] as well as by trypsin and thermolysin,[1,7] is confined to one particular 12-residue sequence that contains no prolyl or hydroxyprolyl residues. Other regions that do not contain Pro or Hyp may also be relatively unstable, and it has been suggested that two such regions play a role during fibrillogenesis in vitro.[8]

A comparison of the ease with which internal loops are formed in various regions of the collagen sequence requires the evaluation of $\Delta G_x(y)$, the free energy of formation of a three-stranded loop that contains a segment of x residues from each of the three strands (i.e. a total of $3x$ residues) and starts with residue y

$$\Delta G_x(y) = G_{c,x}(y) - G_{h,x}(y) \tag{1}$$

where $G_{h,x}(y)$ is the free energy of a triple-helical segment, containing $3x$ residues (x residues in each of the three strands) and starting with residue y, that forms part of a longer triple helix. The free energy of the same segment in a loop, i.e., a disordered state, with both ends of each of the three strands connected to unbroken triple helices, is $G_{c,x}(y)$. A plot of $\Delta G_x(y)$ as a function of y for collagen would be expected to possess minima, indicating the position of the sequences with relatively high probabilities of formation of a loop. If the (positive) value of $\Delta G_x(y)$ is not very high, the loop has marginal, transient stability, and it may become accessible for biologically important reactions.

The total free energy of formation of a loop is composed of several different contributions. For the process of going from the triple helix to the loop structure, these contributions are (i) the energy of breaking noncovalent interactions between the three strands in the triple helix, (ii) the change of conformational energy of the individual strands, (iii) the change in the free energy of hydration of the strands, and (iv) the conformational entropy of formation of the loop. Theoretical values of the energy of the triple helix have been computed earlier for several collagen-like poly(tripeptide) triple helices.[9-13] The energy of various conformations of the individual polypeptide strands can also be obtained by means of conformational energy computations.[9,14] The contribution of hydration is now being investigated by means of recently derived procedures for the computation of free energies of hydration.[15] The change in conformational entropy (contribution iv) for the formation of a triple loop from a triple helix can be expressed as the sum of three contributions: (iva) the entropy gain of converting the triple-stranded helix to three random flight chains, (ivb) a reduction of the entropy because of the restriction of the ends to form a triple loop, and (ivc) a further reduction of the entropy because of short- and long-range excluded volume effects within and between the chains. In the present study, formation of a loop is modeled by means of a computer simulation involving the generation of random chains, and the entropy contributions ivb and ivc are estimated for several factors (described below) that contribute to the excluded volume effect.

The basis for the treatment of this problem is provided by computational approaches to the entropy of ring closure for the formation of a simple loop, that is applicable to the formation of internal loops in a double-stranded helix, e.g. DNA.[16] This problem is equivalent to that of computing the probability for the meeting of two ends of a linear polymer molecule. The solution is well-known for the case of a long random-flight chain,[17,18] and it has been applied to the phase transition in double-stranded nucleic acids,[19] including a correction for the excluded volume.[20] The formation of an internal disruption in a triple helix presents a similar problem. In this case, three disordered chains emanate from the origin of the disruption, and the remote ends of the three chains must meet again at the site where the disruption ends and the triple helix resumes. Random flight treatments for the cyclization of linear chains can be generalized to treat the case of multiple correlated loops.[21] These treatments are based, however, on Gaussian random-flight statistics,[17,18,21,22,23] and/or they use[19,20,22] the method of sequence-generating functions[24] that is valid in the limit of infinitely long chains. Therefore, they apply only to long loops. Other theoretical studies of the triple helix–coil transition of collagen treated primarily either the all-or-none transition or a zipperlike unfolding from the ends of the triple helix rather than internal loop formation.[25-28]

In the consideration of internal loops in collagen, only short loops with x of the order of about 20–30 residues are of primary interest because the longest observed sequences not containing Pro or Hyp do not exceed this size.[2] Therefore, the entropy of the loops must be evaluated by means of a formulation that is applicable to short chains. A computer simulation is used here, instead of an analytical formulation of the entropy, to generate random conformations of a short chain. This work deals primarily with the consideration of several corrections to the random flight statistics of an open chain in order to take into account the excluded volume effects of real polypeptide chains and the formation of triple-stranded loops.

As the first correction, random chain statistics must be modified to reflect the short-range interactions that occur in the collagen chain. In particular, the computation must lead to the correct characteristic ratio for an unconstrained single chain. The polypeptide chain is represented here in terms of virtual bonds of length $l = 3.80$ Å, connecting neighboring C^α atoms. For a polypeptide chain of $n + 1$ residues, consisting of n virtual bonds of length l, the characteristic ratio is defined[29] as

$$C = \langle r^2 \rangle_0 / nl^2 \tag{2}$$

where $\langle r^2 \rangle_0$ is the mean-square end-to-end distance in the absence of long-range excluded volume interactions. For an infinitely long chain, the characteristic ratio is denoted C_∞. Random-flight chains have a characteristic ratio of unity, but experimental measurements on collagen gave values of 2.5–3.3 for denatured ichthyocol[30,31] and calf skin[31,32] collagen, respectively. Actually, the observed mean-square end-to-end distance, $\langle r^2 \rangle$, in good solvents is greater than $\langle r^2 \rangle_0$ by an expansion factor, α, defined as

$$\alpha^2 = \langle r^2 \rangle / \langle r^2 \rangle_0 \tag{3}$$

The value of α can be estimated from the measured molecular weight, intrinsic viscosity, and second virial coefficient.[33] This expansion can be attributed to repulsive long-range interactions between parts of the chain. It is treated in the computation in terms of an excluded volume. This treatment follows the analogous use of excluded

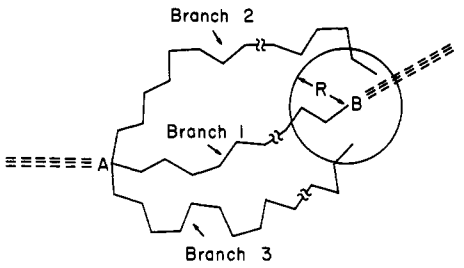

Figure 1. Diagrammatic representation of an internal loop (solid lines) in a triple helix (parallel dashed lines). The disruption commences at the tetrafunctional branch point A at the left. The remote ends of the three branches must meet at the point B where the triple helix resumes. Ends of chains 2 and 3 that fall inside the sphere with radius R are accepted as satisfying the condition for loop closure.

volume for polymethylene stars in an earlier study.[34]

A second correction takes into account the effect of excluded volume on the probability of cyclization of a linear chain. A third correction accounts for the excluded volume at the point of ring closure, i.e. at the point in space where the ends of the three separate chains meet (Figure 1).

Methods

Generation of Random Chains. The method used for generating chains has been adapted from the one recently employed for the study of the expansion factors of trifunctional polymethylene stars.[34] In that work, the unperturbed polymethylene chains were those described in the well-known rotational isomeric state model introduced by Abe et al.[35] In the present work, the only important modifications have been changes in the geometry of the chain and the weighting of the rotational isomers, in order to represent disordered collagen chains. Formally, the beginning of an internal loop in collagen corresponds to a tetrafunctional branching point. One of the four branches is, however, the triple helix having a fixed conformation. The triple helix is not considered here. Therefore, the branch point can be treated as trifunctional in the computation.

Since the current objective is not the description of a chain with a uniquely specified sequence but a general formulation that provides an approximation for the various sequences found in collagen, atomic detail is not necessary. It is sufficient to generate chains with mean-square dimensions that are typical of disordered collagens. Within the framework of the model described by Abe et al.,[35] such chains are obtained by an appropriate choice of the virtual bond geometry, as described below. Representative trifunctional stars, in which the individual chains have unperturbed dimensions corresponding to denatured collagen, are generated by using a random number generator and the same formalism for the a priori and conditional probabilities as that employed for trifunctional polymethylene stars.[36]

Excluded volume is introduced by the assignment of a hard sphere to each C^α atom. The radius of the hard sphere is denoted by r^*. Only the C^α atoms connected by at least three intervening bonds contribute to long-range interactions. Because of the assumption of fixed virtual bond length and bond angle, C^α atoms connected by one or two bonds have fixed distances of separation and do not contribute to the energy. The weight assigned to a particular chain conformation is unity if all pairs of atoms that participate in long-range interactions are separated by distances larger than $2r^*$. A zero weight is assigned if any of these pairs of atoms is separated by a distance smaller than $2r^*$. An ensemble, unperturbed by long-range interactions, is obtained in the special case where $r^* = 0$.

For a given value of r^*, a simulation run yields a value of $\langle r^2 \rangle$ for the ensemble of chains generated ($\langle r^2 \rangle_0$ for $r^* = 0$) and hence, using eq 3, a value of α. Conversely, in order to determine the value of r^* that corresponds to a given α, it is necessary to use a trial and error process, i.e. to repeat the simulation runs for several trial values of r^*.

Definition of the Radius of Gyration of Chain Ends. The end-to-end distance of a particular conformation of a linear chain is r. A corresponding radius of gyration of chain ends, s_e, can be defined, based on the distances of *only the two atoms at the end of the chain* from the center of mass. Since $r = 2s_e$, these two parameters can be used interchangeably for a linear chain. For a trifunctional star, there exist three different end-to-end distances r, but the atoms at the ends of the three branches specify a unique value of s_e. This quantity should be distinguished from the customary radius of gyration s, specified by all atoms of the chain. For unperturbed chains only, $\langle r^2 \rangle_0 = 6\langle s^2 \rangle_0$, while $\langle r^2 \rangle_0 = 4\langle s_e^2 \rangle_0$.

Incorporation of the Consequences of Short-Range Interactions. The characteristic ratio of denatured collagen can be estimated from experimental data, as summarized by Flory.[31] The most reliable estimates are deduced from studies of ichthyocol collagen[30] in 2 M KCNS at pH 3.7 and 25 °C and calf skin collagen[32] in 0.15 M citrate at pH 3.7 and 38 °C. Both samples were studied in good solvents, in which the chain is expanded (i.e. $\alpha > 1$) because repulsive long-range polymer–polymer interactions are only partially compensated by polymer–solvent interactions. Values of $\alpha = 1.24$ and 1.36 were obtained for denatured ichthyocol and calf skin collagens, respectively.[31] Combination of these values of α with the measured values of $\langle r^2 \rangle$ and substitution in eq 2 and 3, give the observed characteristic ratios $C = 2.5$ for calf skin and $C = 3.3$ for ichthyocol collagen.[31] Smaller values, in the range $C = 1.5$–1.7, were estimated[31] from experimental data reported by Gouinlock et al. in better solvents.[37] Greater weight is given, however, to the characteristic ratios of 2.5–3.3 because this value of C is in reasonable agreement with the theoretical range of 2.8–3.6 estimated for sequential copolypeptides in which every third residue is glycyl.[38] It is also supported by measured values near $C \approx 2.8$ for two related sequential copolypeptides, poly(Pro-Gly) and poly(Gly-Gly-Pro-Gly).[38]

In order to generate unbranched chains having $C_\infty \approx 3$, using the rotational isomeric state model,[34] it is necessary to introduce short-range interactions that modify the behavior from that of a random-flight chain ($C_\infty = 1$). This can be achieved by considering chains with fixed bond angles and independent, symmetric rotation potentials for internal bonds. Writing θ for the bond angle and ϕ for the dihedral angle about the *virtual bonds* and employing the convention where $\phi = 180°$ for the trans conformation, the expression for C_∞ becomes[39,40]

$$C_\infty = \left(\frac{1 - \cos\theta}{1 + \cos\theta}\right)\left(\frac{1 - \langle\cos\phi\rangle}{1 + \langle\cos\phi\rangle}\right) \tag{4}$$

The term involving $\langle\cos\phi\rangle$ vanishes when there are three equally probable rotational states with $\phi = 180°$ and $\pm 60°$. The value of C_∞ will then be unity if $\theta = 90°$, and it increases to 3 as θ increases to 120°. These values of θ are physically reasonable because they occur frequently for the low-energy ranges of the dihedral angles (ϕ, ψ) of a polypeptide.[41] With use of the values and assumptions described here, chains with any desired number n of virtual bonds of length l can be generated by a random number generator.

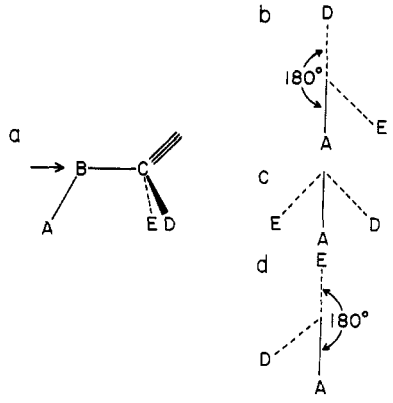

Figure 2. Depiction of the rotational states for the virtual bond B–C, preceding atom C located at the trifunctional branch point. Actually, atom C is tetrafunctional, because it is located at the beginning of a loop located at the end of a triple-helical sequence (indicated by the triple line). For the present purpose, however, only the three branches in the loop have to be considered. (a) Structure of the chain near the branch point. The three rotational states, viewed along the bond B–C, from the direction shown by the arrow, are shown in parts b–d, where the bond A–B is depicted by a full line and bonds C–D and D–E are depicted by dashed lines.

The rotational isomers have to be defined in a slightly different manner for the three bonds next to the trifunctional branch point. These bonds are depicted as CB, CD, and CE in Figure 2. For each of these three bonds, two of the rotational isomers correspond to arrangements with two substituents (e.g. A and D or A and E) in a trans position (parts b and d of Figure 2). The third rotational isomer has two gauche substituents (part c of Figure 2). The three rotational isomers are converted into each other by rotations of 120° around the B–C bond only if the bond angles BCD, BCE, and DCE are all tetrahedral. For other bond angles θ, the values of ϕ must be determined from the geometry around atom C.

Incorporation of the Consequences of Long-Range Interactions. Long-range interactions are represented in terms of a hard-sphere excluded volume assigned to each C^α atom. The radius of the hard sphere, r^*, is an adjustable parameter. There is no a priori exact way of specifying r^*, but a lower and an upper bound can be specified, as follows.

When $r^* = 0$, the chains are unperturbed by long-range interactions and have the properties expected in a Θ solvent. On the other hand, solvents in which collagen is denatured are found to be good solvents rather than Θ solvents.[30-32,37] Apparently, the poor solvent power of a Θ solvent is insufficient to produce disruption of the triple helix. Thus, $r^* = 0$ is a lower bound because any solvent, which is good enough to cause internal disruptions in the triple helix, must be a better solvent.

An upper limit on the interesting range for r^* is established by the measured expansion factors for completely denatured collagen chains of high molecular weight.[30-32,37] The values of α cited earlier for denatured ichthyocol (M = 135000) and calf skin (M = 195000) collagen correspond to $\alpha = 1.10 \pm 0.02$ at $M = 5000$ (i.e. 60 residues, which is the size of the longest linear chain in the trifunctional star, i.e. the size of a linear chain composed of any *two* branches in Figure 1) when the molecular weight dependence is that described by Flory's relationship[42]

$$(\alpha^5 - \alpha^3)M^{-1/2} = \text{constant} \qquad (5)$$

A similar treatment of the data for the more highly perturbed gelatin samples[37] yields α in the range 1.17–1.25 for $M = 5000$. A value of $r^*/l = 0.45$, used in the present work, yields an expansion factor $\alpha = 1.2$ for a collagen chain of this size. Therefore, it represents an upper limit because it would correspond to very good solvent conditions, where the triple helix is completely disrupted and the denatured chains are highly expanded. Therefore, r^* for an exact calculation should be less than the high value used here as a test case that gives an extreme upper limit of the effect of excluded volume.

Closure of the Three-Stranded Loop. Exact ring closure in Figure 1 implies that the terminal atoms of all three chains coincide at point B. As discussed below, the computed probability of achieving this in the numerical simulation procedure was very small for chains on a lattice, while no chains with exact loop closure were generated for off-lattice chains. Therefore, a less stringent criterion must be used in practice; viz., a loop is accepted as closed when the chain ends fall within a specified cutoff distance. For a double-stranded loop, i.e. for the joining of two chain ends, such a criterion can be satisfied by centering a sphere of radius R on the atom at the remote end of branch 1, as depicted in Figure 1, and requiring that the end of branch 2 should fall within this sphere. The fraction of chains with acceptable ring closure in a simulation run is computed for various choices of R, and this fraction is extrapolated to R approaching zero, in order to obtain the probability of exact ring closure.

This procedure is not applicable for the joining of three chain ends because it results in the acceptance of configurations in which the ends of both chains 2 and 3 fall inside the sphere of radius R, i.e. have a distance less than R from the end of chain 1, but the distance between the terminal atoms of chains 2 and 3 is greater than R. Such a situation is shown in Figure 1. Furthermore, there is no reason to single out any of the chains as chain 1. Therefore, it is preferable to use a criterion that treats all three chains equivalently. Such a criterion is provided by the radius of gyration of chain ends, s_e. A configuration is accepted as a closed three-stranded loop if s_e is less than a cutoff value. The fraction of such chains is computed for various choices of the cutoff in a simulation run. By extrapolating the cutoff to zero in the manner described below in the subsection "Probabilities of Cyclization for Finite, Perturbed Linear Chains", the probability for exact ring closure is obtained.

For finite values of R and of the cutoff in s_e, the two criteria give different results because of the reason described above. After extrapolation to zero, the chains retained by using both criteria have exact ring closure. Therefore, the two methods are equivalent. In the present approximation, bond angle restrictions between the three chain ends as well as between the triple helix and the three chains at both ends of the loop are disregarded.

Results and Discussion

Distribution functions are shown and compared first for chains on a lattice with tetrahedral bond angles (θ = 109.47°) and then for off-lattice chains with larger bond angles, viz., θ = 115.38° and 120.0°, in order to demonstrate the effects of short-range interactions on the characteristic ratio C and on the probability of cyclization. These bond angles correspond to characteristic ratios of C = 2.0, 2.5, and 3.0, respectively. The latter two values cover the observed range for denatured collagens, cited above. The effects of long-range interactions are analyzed for all kinds of chains by considering atoms with $r^*/l = 0$ (no excluded volume) and $r^*/l = 0.45$, corresponding to $r^* = 1.71$ Å (with excluded volume). For each of the above choices of the parameters, the effect of short- and long-range interactions on the distribution functions for the radii of gyration of

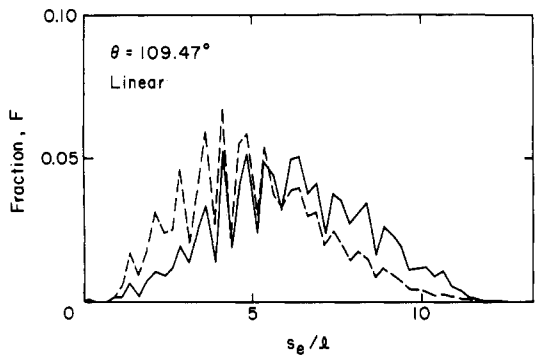

Figure 3. Distribution functions for the radius of gyration of the two terminal atoms of an unbranched polymer, i.e. chains 1 and 2, when bond angles are tetrahedral and each branch contains 30 bonds, in the absence of chain 3. The unperturbed distribution function (dashed line), computed for $r^* = 0$, is based on 100 000 chains. Only 15 772 chains survive when $r^*/l = 0.45$ (full line). In drawing the distribution function, chains were collected at intervals of 0.25 for s_e/l.

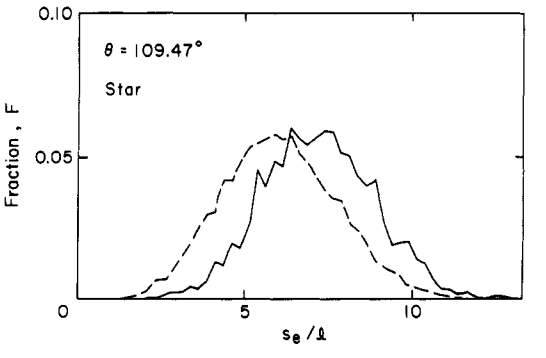

Figure 4. Distribution functions for the radius of gyration of the three terminal atoms in branches 1, 2, and 3 of a trifunctional star when bond angles are tetrahedral and each branch contains 30 bonds. Unperturbed ($r^* = 0$, dashed line) and perturbed ($r^*/l = 0.45$, full line) distribution functions are based on 100 000 and 7 120 chains, respectively.

open linear (unbranched) chains is discussed first, followed by an analysis of the probabilities of cyclization of the open chains. Finally, the effect of going to a triple loop (instead of a double loop) will be discussed in the last subsection.

Distribution Functions for Chains on a Tetrahedral Lattice. By fixing $\theta = 109.47°$, all generated chain conformations can be represented on a tetrahedral lattice. First, unbranched chains of 60 bonds in a linear arrangement are generated. They correspond to branches 1 and 2 in Figure 1. Therefore, the distribution functions are given here as a function of s_e, rather than r, in order to facilitate comparison with distribution functions for three branches, to be presented later. The distribution function for s_e in 100 000 representative unperturbed chains ($r^* = 0$) is depicted in Figure 3 as a dashed line.

A total of 74 of the 100 000 chains that were generated have an end-to-end distance of zero, i.e. correspond to exactly closed rings, causing the point for the range $0 < s_e/l < 0.25$ to be above zero in Figure 3. No tetrahedral lattice chain with an even number of bonds can have s_e/l lying between 0.25 and 0.75 because the two smallest end-to-end distances are $r = 0$ and $r = (^8/_3)^{1/2}$, i.e. $s_e/l = 0$ and $s_e/l = (^2/_3)^{1/2}$, respectively. Consequently, the distribution function falls to zero in this range, but it rises above zero for s_e/l in the range 0.75–1.00. This interval contains 922 chains. The remainder of the distribution function is not smooth because the end-to-end distances accessible to a chain of 60 bonds on a tetrahedral lattice do not describe a continuum. The value of $C = 1.965$, obtained from the computed mean-square end-to-end distance, is in excellent agreement with the expected value for an infinitely long chain ($C_\infty = 2.000$) when allowance is made for the fact that the chain contains a finite, rather than an infinite, number of bonds.[43,44]

Only 15 772 of the 100 000 chains survive if the atoms in branches 1 and 2 participate in long-range interactions with $r^*/l = 0.45$. The normalized distribution function for the *surviving* chains is shown in Figure 3 as a full line. Among the chains that fail to survive are the 74 that had an end-to-end distance of zero in the unperturbed ensemble because of overlap of the first and last atoms. The perturbed distribution function retains the nonsmooth character of the unperturbed distribution function, but it is biased in favor of conformations with larger end-to-end distances. The effect of the choice of $r^*/l = 0.45$ on the mean dimensions can be assessed by focusing attention either on the expansion of the mean-square end-to-end distance or that of the mean-square radius of gyration, $\langle s^2 \rangle$, which do not differ much from each other. The two corresponding expansion factors will be distinguished by adopting the notation $\alpha_r^2 = \langle r^2 \rangle / \langle r^2 \rangle_0$ and $\alpha_s^2 = \langle s^2 \rangle / \langle s^2 \rangle_0$. All 61 atoms in the chain are considered in the evaluation of α_s^2, but only the two atoms at the ends contribute to α_r^2. Use of $r^*/l = 0.45$ produces $\alpha_r^2 = 1.49$ and $\alpha_s^2 = 1.40$. The finding that $\alpha_r^2 > \alpha_s^2 > 1$ is expected for short perturbed chains.[45,46] A value of $r^*/l = 0.45$ produces an expansion so large that it corresponds to collagen in a medium in which the triple helix is unstable because α_r and α_s are both larger than 1.1, as described above in the subsection "Incorporation of the Consequences of Long-Range Interactions". Thus, the calculations presented here give an upper limit for the magnitude of the excluded volume effect.

A trifunctional star is considered next. Figure 4 depicts distribution functions for the radius of gyration s_e of the three terminal atoms in branches 1, 2, and 3 when all bond angles are tetrahedral and each branch contains 30 bonds. As stated earlier, the quantity s_e has an unambiguous meaning for a trifunctional star because it can be defined for any number of chain ends, in contrast to r.

In the absence of excluded volume, for every conformation of branches 1 and 2 with 30 bonds each, there are up to 3^{29} distinguishable conformations of branch 3 with 30 bonds. Consequently, distribution functions for the radius of gyration of the *ends* of the trifunctional stars in Figure 4 are much smoother than are their counterparts for the two ends of the linear chains in Figure 3. None of the trifunctional stars generated has $s_e = 0$. The smallest values of s_e/l lie in the interval 1.00–1.25. This range was obtained for six of the 100 000 trifunctional stars. The fraction of the molecules in this interval, 0.00006, is so small that it is not distinguishable from zero on the scale used in Figure 4. The computed ratio of the mean-square radius of gyration $\langle s^2 \rangle_0$ of the 100 000 unperturbed ($r^* = 0$) *trifunctional stars* to $\langle s_e^2 \rangle_0$ for the corresponding unperturbed linear chain of 90 bonds is 0.52. This result agrees exactly with the theoretical value of 0.52, derived from the combination of the expressions $\langle r^2 \rangle_0 = 6\langle s^2 \rangle_0$ for a linear chain, $\langle r^2 \rangle_0 = 4\langle s_e^2 \rangle_0$ for the terminal atoms of a linear chain, and $g = {}^7/_9$. Here g is the ratio of the mean-square unpertured radii of gyration for the branched and linear molecules with the same number of bonds, as evaluated with random flight statistics.[47]

Only 7120 trifunctional stars survive the imposition of excluded volume with $r^*/l = 0.45$. The normalized distribution function is depicted in Figure 4. The smallest values of s_e/l, lying in the interval 1.50–1.75, are found in six of the surviving molecules.

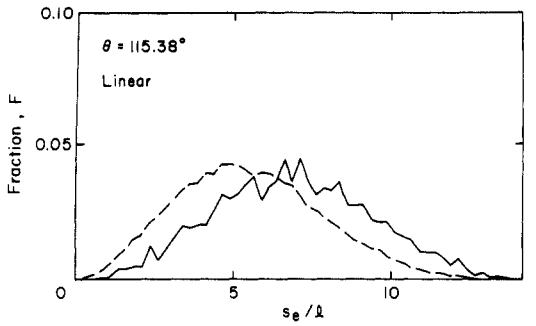

Figure 5. Distribution functions for the radius of gyration of the two terminal atoms of an unbranched polymer, i.e. chains 1 and 2 when $\theta = 115.38°$ and each branch contains 30 bonds. Unperturbed ($r^*/l = 0$) and perturbed ($r^*/l = 0.45$) distribution functions are based on 100 000 and 16 466 chains, respectively. See the legend of Figure 3 for further explanation.

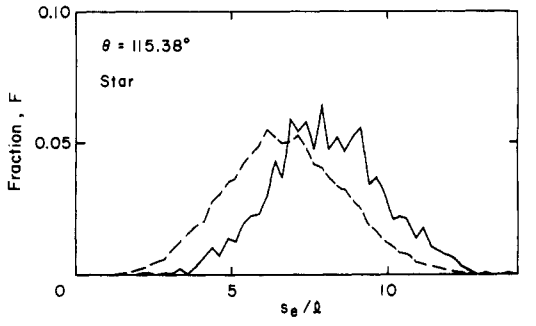

Figure 6. Distribution functions for the radius of gyration of the terminal atoms in branches 1, 2, and 3 of a trifunctional star when $\theta = 115.38°$ and each branch contains 30 bonds. Unperturbed ($r^*/l = 0$) and perturbed ($r^*/l = 0.45$) distribution functions are based on 100 000 and 7 491 chains, respectively.

Distribution Functions for Off-Lattice Chains. Figures 5 and 6 depict distribution functions analogous to those presented in Figures 3 and 4 but for chains in which the bond angle is $\theta = 115.38°$, corresponding to $C_\infty = 2.5$. This value is in excellent agreement with the observed characteristic ratio of 2.5 for denatured calf skin collagen.[31] Chains with $\theta = 115.38°$ are no longer confined to a lattice. Freeing the chains from the confining influence of the lattice has two important effects on the distribution functions for the linear chain of 60 bonds (compare Figures 3 and 5). First, off-lattice chains have smoother distribution functions because the available end-to-end distances more nearly describe a continuum. While 74 of the 100 000 tetrahedral lattice chains had $s_e/l < 0.25$, none of the off-lattice chains have s_e/l falling into this range. Second, the guiding influence of the lattice abets the attainment of precisely cyclic conformations when the number of bonds is even and larger than 4, while no such constraint occurs for the off-lattice chains. The most nearly cyclic chains that were computed with $\theta = 115.38°$ have s_e/l in the interval 0.25–0.50. There were 48 such chains, out of 100 000.

While slightly more molecules survive the imposition of excluded volume interactions when $\theta = 115.38°$ than for $\theta = 109.47°$, there is little effect on the expansion factors. Linear chains of 60 bonds in Figure 5 have $\alpha_r^2 = 1.48$ and $\alpha_s^2 = 1.40$, and the trifunctional stars with 30 bonds in each branch in Figure 6 have $\alpha_s^2 = 1.36$.

Similar results have been obtained with $\theta = 120°$. For the latter bond angle, linear chains of 60 bonds have $\alpha_r^2 = 1.35$ and $\alpha_s^2 = 1.29$, and the trifunctional stars with 90 bonds have $\alpha_s^2 = 1.30$. The corresponding C_∞ for the linear chain is 3.0, which places it still within the range defined for denatured collagen.[30-32] The distribution functions do not differ significantly from those shown for $\theta = 115.38°$ in Figures 5 and 6 and are not shown here. The extent of the expansion produced with $r^*/l = 0.45$ is so large for both values of θ that it simulates collagen in a solvent medium favoring the denatured over the triple-helical form.

Probabilities of Cyclization for Finite, Perturbed Linear Chains. Successful exact cyclization of a linear chain is achieved when the two ends coincide. This objective was achieved with 74 of the 100 000 tetrahedral lattice chains, but not with any of the off-lattice chains. Therefore, a less stringent criterion for cyclization is needed in the latter case, as described in the Methods section. The distribution functions of Figures 3–6 are depicted in terms of the radius of gyration of chain ends s_e, expressed in the dimensionless form s_e/l. The fraction of molecules with s_e/l that is less than any given cutoff value can be obtained by summing the distribution function from zero to the cutoff value, in the manner described next.

The distribution functions are normalized

$$\sum_{i=1}^{\infty} F_i = 1 \tag{6}$$

where i indexes the intervals of width $0.25l$ and F_i is the fraction of the molecules having s_e/l between $(i-1)/4$ and $i/4$. The index of summation is defined as $i = 1$ for the interval having s_e/l in the range 0–0.25 and increases for successive intervals, i.e. for larger s_e. The compact notation W_j will be used to denote the cumulative distribution, i.e. the part of the sum generated from the first up through the jth interval, corresponding to $0 < s_e/l < j/4$, i.e.

$$W_j = \sum_{i=1}^{j} F_i \tag{7}$$

Thus, W_j is the fraction of molecules having a radius of gyration of chain ends less than $0.25j(s_e/l)$, i.e. accepted provisionally to have approximate ring closure. By extrapolating to the limit as $j \to 1$, in the manner described below, the probability for exact ring closure can be estimated. A zero subscript is appended to W_j if it is evaluated in the absence of excluded volume ($r^* = 0$); absence of this subscript signifies $r^* > 0$.

The influence of excluded volume on the ease of cyclization of a linear chain is approximated by $\lim_{j\to 1} W_j/W_{j,0}$ because this limit represents the ratio of distributions for the smallest possible s_e. The qualitative inspection of Figures 3 and 5 shows that this limit is <1 because compact conformations are strongly avoided in the perturbed distribution function. Cyclization probabilities based on random-flight chain statistics without excluded volume overestimate the ease of ring closure of perturbed chains.

A practical difficulty arises in making a quantitative assessment of this limit because very few chains have extremely small s_e. The distribution functions are not nearly as reliable at small s_e as they are at intermediate s_e. It is desirable to make use of the most accurately known portion of the distribution function in estimating the effect of excluded volume on cyclization probabilities. For this purpose, it is useful if $W_j/W_{j,0}$ is examined as a function of $W_{j,0}$ because this functional form can be extrapolated more reliably to $W_{j,0} = 0$ (Figure 7) than $W_{j,0}$ by itself. Three such plots, using $r^*/l = 0.45$, are depicted in Figure 7. The interior parts of the plots are nearly linear, even though the distribution functions used to prepare the plots are quite rough (e.g. compare part c with Figure 3 from which it was prepared). Dashed straight lines denote the extrapolation from the linear range.

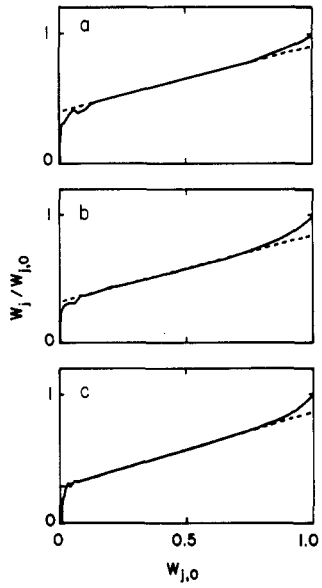

Figure 7. $W_j/W_{j,0}$ vs $W_{j,0}$ for linear chains of 60 bonds having $r^*/l = 0.45$ and the bond angles indicated. A bond angle of $\theta = 120°$ was used in part a. The corresponding distribution function is not shown here. Its qualitative appearance is similar to Figure 5. Parts b with $\theta = 115.38°$ and c with $\theta = 109.47°$ are based on the distribution functions in Figures 5 and 3, respectively. Dashed lines indicate linear extrapolations from the central range.

Random departures from linearity may be attributed to inaccuracies in the characterization of the distribution function at low values of j, i.e. at very small s_e. A systematic bias, lowering the curves, arises at very small $W_{j,0}$ because of the participation of the two atoms at the ends of the chain in long-range interactions. Thus, when $r^*/l = 0.45$, the interaction of the two terminal atoms causes W_1, W_2, and W_3 to be zero because of the excluded volume between the two terminal atoms.

The linearly extrapolated intercepts in Figure 7 range from 0.28 in part c to 0.40 in part a. This result implies that the probability of ring closure (i.e. the formation of a loop by *two* chains) is reduced by a factor of 0.28–0.40 when excluded volume is taken into account. The intercepts would be lowered even further if W_j were evaluated by using $r^*/l > 0.45$. However, $r^*/l = 0.45$ is already large enough to correspond to the expansion of collagen in a solvent that is so good that the denatured form is favored over the triple helix. If the triple helix is the dominant conformation, the solvent must be somewhat poorer, a condition that corresponds to a value of r^*/l that lies somewhere between 0 and 0.45.

Interaction of Branch 3 with Branches 1 and 2. The terminal atom in a third branch (branch 3 of Figure 1) must meet the terminal atoms in the other two branches in a triple loop, i.e. a loop formed by an *internal* disruption of the triple helix. The notation adopted in the preceding section will be extended here by appending a subscript b if the W was evaluated by using a distribution function for a trifunctional star, such as those depicted in Figure 4 and 6. Absence of b as a subscript signifies use of a distribution function for a linear chain (i.e. for a double loop), such as those depicted in Figures 3 and 5.

Consider first the possibility that branches 2 and 3 behave independently. This assumption should be legitimate in the absence of long-range interactions, but it becomes dubious when excluded volume effects are present. In the absence of excluded volume, the probability that the remote end of branch 2 will find the end of branch 1 should equal the probability that the end of branch 3 will find the end of branch 1. Squaring either of these probabilities

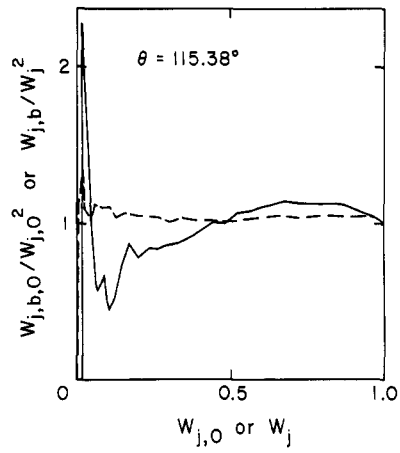

Figure 8. Plots of $W_{j,b,0}/(W_{j,0})^2$ vs $W_{j,0}$ for $r^*/l = 0$ (dashed line) and of $W_{j,b}/(W_j)^2$ vs W_j for $r^*/l = 0.45$ (full line), when the bond angle is $\theta = 115.38°$ and each branch contains 30 bonds. The corresponding distribution functions are depicted in Figures 5 and 6.

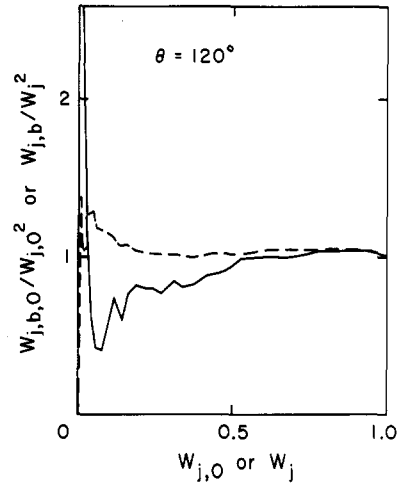

Figure 9. Plots of $W_{j,b,0}/(W_{j,0})^2$ vs $W_{j,0}$ for $r^*/l = 0$ (dashed line) and of $W_{j,b}/(W_j)^2$ vs W_j for $r^*/l = 0.45$ (full line), when the bond angle is 120° and each branch contains 30 bonds.

would give the probability that all three chain ends come together at one point. If the same concept is expressed in terms of W's, rather than probabilities, it becomes

$$W_{j,b,0} = (W_{j,0})^2 \tag{8}$$

Figure 8 depicts $W_{j,b,0}/(W_{j,0})^2$ as a function of $W_{j,0}$ by using unperturbed distribution functions for $\theta = 115.38°$ from Figures 5 and 6. The value of $W_{j,b,0}/(W_{j,0})^2$ falls between 1.0 and 1.1 for all except very small $W_{j,0}$. Overall, the behavior of $W_{j,b,0}/(W_{j,0})^2$ in Figure 8 supports the notion that branches 2 and 3 behave independently when r^* is zero. Figure 9 presents equivalent plots for the case where $\theta = 120°$. Here also, $W_{j,b,0}/(W_{j,0})^2$ remains between 1.0 and 1.1 for all but very small $W_{j,0}$. Thus, essentially the same behavior is seen for the two bond angles that cover the observed range of C in denatured collagen. No plot of $W_{j,b,0}/(W_{j,0})^2$ is presented here for the lattice chain ($\theta = 109.47°$) because its characteristic ratio C falls below the observed values.

The analogous ratio in the presence of exluded volume, $W_{j,b}/(W_j)^2$, is also depicted in Figures 8 and 9. Fluctuations become violent at very small W_j because of inadequacies in the characterization of the distribution functions themselves at small s_e. Therefore, the trends can be followed only for $W_j > 0.1$. It is apparent that excluded volume has abolished the independence of branches 2 and 3 because the perturbed curves reach values smaller than

unity. Interaction of branch 3 with branches 1 and 2 makes it more difficult to produce an internal disruption in the triple helix when $r^* > 0$ than when $r^* = 0$ because of the difficulty of bringing together the three chain ends. The size of the effect should be evaluated as the ratio of the intercept for perturbed to unperturbed curves in Figures 8 and 9. The unperturbed curves in Figures 8 and 9 extrapolate to an ordinate intercept of about 1.0. The intercepts on the ordinates of the perturbed curves are not well-defined, but reasonable visual extrapolation in both figures (using the portions of the curves for $W_j > 0.1$) results in intercepts in an approximate range from 0.25 to 0.5. Therefore, it seems that the interaction of branch 3 with branches 1 and 2 leads to a reduction of W by a factor that is not likely to be less than 0.25.

Conclusions

It has been shown here that the probability of cyclization of two 30-residue random-flight chains is reduced by a factor of up to 0.28–0.40 by long-range excluded volume effects. This factor has been calculated for an expansion factor (related to short-range interactions between neighboring residues) that corresponds to the observed characteristic ratios of denatured collagen in a good solvent. The probability of formation of a triple loop is smaller by another factor of about 0.25–0.50, because of the interaction of branch 3 with branches 1 and 2. Combination of these two numbers results in a maximal overall reduction of the probability of loop formation by a factor of about 0.07–0.20 for the combined effect of short- and long-range excluded volumes and of the presence of three branches in the loop, i.e. of all effects listed as contributions ivb and ivc in the Introduction. In other words, the probability of loop formation that is computed by using random-flight statistics (contribution iva) may overstate the probability of internal loop formation by about an order of magnitude.

The reduction would be stronger if one were to assume larger excluded volumes, i.e. higher values of r^*. However, as noted earlier, the value of r^* used here is already so large that it corresponds to collagen in a good solvent in which the denatured form is more stable than the triple helix. Under solution conditions where the triple helix is thermodynamically more stable, smaller values of r^* are more applicable. Therefore, the present calculation represents an extreme case. For solvent media in which the triple helix is the stable conformation, the correction of the probability of loop formation for excluded volume will be in the range 0.1–1.0.

The present work demonstrates that excluded volume decreases the entropic contribution to internal loop formation by up to 5.3 cal/(deg·mol) and thus increases $G_{c,x}(y)$ by up to 1.5 kcal/mol, relative to the formation of random chains. Thus, the correction introduced is a minor one, decreasing the probability of loop formation by at most an order of magnitude. Preliminary calculations in the absence of solvent from our laboratory (not presented here) suggested that contributions i, ii, and iva result in an unreasonably high probability of loop formation. This is not reduced sufficiently by contributions ivb and ivc, as computed here. Therefore, it is likely that hydration, not taken into account heretofore, also makes a significant contribution to the free energy of loop formation. This contribution is now under study together with a reevaluation of the energetic factors. Initial results will be reported.[48]

Acknowledgment. This research was performed while W.L.M. was on sabbatical leave at Cornell University during 1983–1984. W.L.M. gratefully acknowledges sabbatical support from the National Institutes of Health (Senior National Research Service Award 1 F33-AM-07296-01). This research was also supported by research grants to H.A.S. from the National Institutes of Health (DK-08465) and from the National Science Foundation (DMB 84-01811), to G.N. from the National Institute on Aging (AG-00322), and to W.L.M. from the National Science Foundation (DMB 85-00338).

References and Notes

(1) Miller, E. J.; Finch, J. E., Jr.; Chung, E.; Butler, W. T. *Arch. Biochem. Biophys.* **1976**, *173*, 631.
(2) Bornstein, P.; Traub, W. In *The Proteins*, 3rd ed.; Neurath, H., Hill, R. L., Eds.; Academic: New York, 1979; Vol. IV, p 411.
(3) Ryhänen, L.; Zaragoza, E. J.; Uitto, J. *Arch. Biochem. Biophys.* **1983**, *223*, 562.
(4) French, M. F.; Mookhtiar, K. A.; Van Wart, H. E. *Biochemistry* **1987**, *26*, 681.
(5) Harris, E. D., Jr.; Cartwright, E. C. Mammalian collagenases, In *Research Monographs in Cell and Tissue Physiology*; Dingle, J. T., Ed.; Elsevier/North Holland Biomed: Amsterdam, 1977; Vol. 2, Chapter 6.
(6) Harper, E. *Annu. Rev. Biochem.* **1980**, *49*, 1063.
(7) Wang, H.-M.; Chan, J.; Pettigrew, D. W.; Sodek, J. *Biochim. Biophys. Acta* **1978**, *533*, 270.
(8) Helseth, D. L., Jr.; Veis, A. *J. Biol. Chem.* **1981**, *256*, 7118.
(9) Miller, M. H.; Scheraga, H. A. *J. Polym. Sci. (Polym. Symp.)*, **1976**, *54*, 171.
(10) Miller, M. H.; Némethy, G.; Scheraga, H. A. *Macromolecules* **1980**, *13*, 470.
(11) Miller, M. H.; Némethy, G.; Scheraga, H. A. *Macromolecules* **1980**, *13*, 910.
(12) Némethy, G.; Miller, M. H.; Scheraga, H. A. *Macromolecules* **1980**, *13*, 914.
(13) Némethy, G. *Biochimie* **1981**, *63*, 125.
(14) Momany, F. A.; McGuire, R. F.; Burgess, A. W.; Scheraga, H. A. *J. Phys. Chem.* **1975**, *79*, 2361.
(15) Kang, Y. K.; Némethy, G.; Scheraga, H. A. *J. Phys. Chem.* **1987**, *91*, 4105, 4109, 4118.
(16) Poland, D.; Scheraga, H. A. *Theory of Helix-Coil Transitions in Biopolymers*; Academic: New York, 1970.
(17) Jacobson, H.; Stockmayer, W. H. *J. Chem. Phys.* **1950**, *18*, 1600.
(18) Flory, P. J. *J. Am. Chem. Soc.* **1956**, *78*, 5222.
(19) Poland, D.; Scheraga, H. A. *J. Chem. Phys.* **1966**, *45*, 1464.
(20) Fisher, M. E. *J. Chem. Phys.* **1966**, *45*, 1469.
(21) Poland, D. C.; Scheraga, H. A. *Biopolymers* **1965**, *3*, 379.
(22) Chen, F. Y.; Huang, H. W. *Macromolecules* **1981**, *14*, 332.
(23) Chandrasekhar, S. *Rev. Mod. Phys.* **1943**, *15*, 1.
(24) Lifson, S. *J. Chem. Phys.* **1964**, *40*, 3705.
(25) Gō, N.; Suezaki, Y. *Biopolymers* **1973**, *12*, 1927.
(26) Schwarz, M., Jr.; Poland, D. *Biopolymers* **1974**, *13*, 687.
(27) Weidner, H.; Engel, J.; Fietzek, P. In *Peptides, Polypeptides and Proteins*; Blout, E. R., Bovey, F. A., Goodman, M., Lotan, N., Eds.; Wiley: New York, 1974; p 419.
(28) Shaw, B. R.; Schurr, J. M. *Biopolymers* **1975**, *14*, 1951.
(29) Brant, D. A.; Flory, P. J. *J. Am. Chem. Soc.* **1965**, *87*, 2788.
(30) Boedtker, H.; Doty, P. *J. Am. Chem. Soc.* **1956**, *78*, 4267.
(31) Flory, P. J. *Brookhaven Symp. Biol.* **1960**, *13*, 89.
(32) Doty, P.; Nishihara, T. In *Recent Advances in Gelatin and Glue Research*; Stainsby, G., Ed.; Pergamon: New York, 1958; p 92.
(33) Orofino, T. A.; Flory, P. J. *J. Chem. Phys.* **1957**, *26*, 1067.
(34) Mattice, W. L. *Macromolecules* **1984**, *17*, 415.
(35) Abe, A.; Jernigan, R. L.; Flory, P. J. *J. Am. Chem. Soc.* **1966**, *88*, 631.
(36) Mattice, W. L. *Macromolecules* **1975**, *8*, 644.
(37) Gouinlock, E. V., Jr.; Flory, P. J.; Scheraga, H. A. *J. Polym. Sci.* **1955**, *16*, 383.
(38) Mattice, W. L.; Mandelkern, L. *Biochemistry* **1971**, *10*, 1934.
(39) Oka, S. *Proc. Phys. Math. Soc. Jpn.* **1942**, *24*, 657.
(40) Flory, P. J. *Statistical Mechanics of Chain Molecules*, Wiley: New York, 1969; p 26.
(41) Nishikawa, K.; Momany, F. A.; Scheraga, H. A. *Macromolecules* **1974**, *7*, 797.
(42) Flory, P. J.; Fox, T. G., Jr. *J. Am. Chem. Soc.* **1951**, *73*, 1904.
(43) Benoit, H. *J. Chim. Phys.* **1947**, *44*, 18.
(44) Benoit, H.; Doty, P. *J. Phys. Chem.* **1953**, *57*, 958.
(45) Mattice, W. L.; Santiago, G. *Macromolecules* **1980**, *13*, 1560.
(46) Mattice, W. L. *Macromolecules* **1981**, *14*, 1485.
(47) Zimm, B. H.; Stockmayer, W. H. *J. Chem. Phys.* **1949**, *17*, 1301.
(48) Némethy, G.; Scheraga, H. A. submitted for publication in *Biopolymers*.