

Thermoelastic Properties of Helical Protein Molecules

H. Reiss¹

Received November 12, 1970

The effect of stress on the helix-coil transition in a protein or polypeptide is investigated using the methods of statistical mechanics. A case is treated in which the helical sections are regarded as flexible chains with *very long*, freely "orienting" segments and another in which they are considered to be rigid rods. Thermoelastic relations are derived; and it turns out that, depending upon conditions, stress can induce the helix-coil transition in one or another direction or do nothing at all. The most probable situations either involve stress applied to a molecule initially helical, in which case the helix is stabilized, or stress applied to the coil form, in which case transformation to the helical form is induced. The helical form exhibits a very low modulus of elasticity (which we also compute), and it is speculated that preservation of, or transition to, the helical form under stress aids in the protection of living tissue from disruption when subjected to large applied strain. Real tissues involve highly organized or quasirandom networks of protein chains. The results of this analysis suggest that, insofar as the mechanical properties of the networks are concerned, the chains can be treated as quasiharmonic strings whose configurations (weighted by potential energy) can be enumerated in order to include entropy effects in the calculation of the network modulus.

KEY WORDS: Protein; helix-coil transition; thermoelastic properties; statistical mechanics.

1. INTRODUCTION

The mechanical properties of living tissue are rooted in the biopolymeric molecules of which they are composed. These molecules are themselves assembled into superstructures, the forms of which may vary all the way from highly organized assemblies such as are found in striated muscle⁽¹⁾ to quasirandom networks typified by fibrin clots.⁽²⁾ In connective tissue such as collagen, cross-linking may be augmented by crystallization and other nonbond-specific processes.

In 1956, Flory⁽³⁾ considered elastic mechanisms in fibrous protein networks,

Work performed as part of a 1970 summer study project in connection with the Advanced Research Projects Agency Materials Research Council.

¹ Department of Chemistry, University of California, Los Angeles, California.

employing a quasithermodynamic approach. Although the features of the network itself exert a profound influence on the mechanical properties of tissue, characterization of these properties should begin with an examination of the mechanical behavior of the individual chains. For example, when these consist of proteins capable of adopting the α -helical form, we might ask whether the application of stress is able to induce transitions between helix and coil, or vice versa. If this is possible, the mechanical properties of the network should be very unusual. As we show later, there is a strong possibility that the application of stress induces the transition, coil-to-helix, and, for molecules initially in the helical state, merely further stabilizes that condition.

It is not unusual to discover that nature has adapted certain life processes to serve more than one purpose. The helix-coil transition may fall into this category. It may be fundamental, for example, to certain kinds of enzyme activity. At the same time, if stress tends to stabilize the helical form (or to induce it in molecules initially coiled), the transition may represent one of nature's methods for protecting tissues from the application of large, disruptive stress. This follows, in a qualitative way, from the following considerations. The helix is a fairly rigid structure which, in the final analysis, cannot be thought of (strictly) as having rubber or gaslike elasticity. On the other hand, if we regard the matter loosely and consider it to behave approximately in this manner, the effective length⁽⁴⁾ of the freely orienting segment will have to be quite large. For gaslike elasticity, this implies a low elastic modulus (since the relative numbers of configurations and hence the relative entropy are not so great in the contracted state) so that appreciable stress can only be supported, under equilibrium conditions, when the strain is large. In contrast, the coiled form of the protein will have a small, freely orienting segment and a large elastic modulus. Application of stress to the coil therefore causes transition to the helical form; the natural direction for the process since it leads to a reduction in elastic modulus and relief of stress. It is therefore difficult to apply large stress to the protein in *either* the coil or helical state.

The above argument may apply to globular as well as fibrous proteins. The helical form, and the possibility of coil-to-helix transformation, may represent a safety feature protecting globular molecules in fluid biological environments such as plasma. The extreme malleability of erythrocytes, enabling them to slip through tiny blood vessels and capillaries smaller than the cell itself, may be facilitated by this low-modulus mechanism.

These conclusions, elaborated below in a more quantitative manner, indicate that the elastic, or quasielastic, behavior of an individual globular protein molecule, or of a polypeptide chain in a network, might be modeled as a quasilinear elastic string (harmonic from the point of view of the dependence of potential energy on local curvature) in which entropic effects can still be evaluated through an enumeration of configurations properly weighted by potential energy. Such a model would of course have its value confined to the description of thermoelastic behavior. We shall defer the examination of this question to a later investigation. In the present paper, considerations will be limited to the investigation of stress on the helix-coil transition in individual protein or polypeptide chains.

2. CHOICE OF MODEL

In order to make contact with earlier work on the helix-coil transition in unstressed systems, we shall attempt to model the protein molecule, as much as possible, on the Zimm-Bragg scheme.⁽⁵⁾ The Zimm-Bragg model has in fact been developed for polypeptide rather than protein molecules, and we shall assume also that the protein can be approximated by a polypeptide. This amounts to replacing the actual distinct amino acids in the chain by an "average" amino acid; the exact effect of the improper order of averaging not being entirely clear. One effect demonstrated with certainty,⁽⁶⁾ however, is a narrowing of the range over which the helix-coil transition takes place. The "average" polypeptide is more like a crystal and has a "sharp" melting point, whereas the true protein may be likened to a disordered amorphous solid and possesses a softening "range."

In the α -helix, hydrogen bonding generally occurs between carbonyl and amino groups situated respectively on amino acid units which are third nearest neighbors. For the purpose of this article, we assume that only *first* nearest neighbors are involved, a simplification which is known⁽⁵⁾ not to influence the *form* of the conclusions—in view of the considerable averaging already involved in the model. At the moment, we are primarily involved with form.

The free energy of a hydrogen-bonded helical segment may be denoted by F_H , while if the bond is dissociated, it becomes F_0 . The contributions, as factors in the molecular partition function of bonded and nonbonded segments, are then

$$e^{-F_H/kT}, \quad e^{-F_0/kT} \quad (1)$$

respectively. Here, k is the Boltzmann constant and T is temperature. Zimm and Bragg select the nonbonded state as standard (choosing the zero of energy appropriately) and set $F_0 = 0$. Then,

$$F_H = \Delta F \quad (2)$$

where ΔF is the free energy of hydrogen bonding. Since it is a *free* energy, ΔF is itself temperature-dependent. We write

$$e^{-F_H/kT} = e^{-\Delta F/kT} = s \quad (3)$$

$$e^{-F_0/kT} = 1 \quad (4)$$

using therefore the same notation as Zimm and Bragg.

Every uninterrupted sequence of helical segments has two ends which interface with sequences of nonhydrogen-bonded segments (coil sequences). The free energy associated with these two interfaces may be denoted by Ψ , and the contribution to the partition function is then

$$\sigma = e^{-\Psi/kT} \quad (5)$$

where σ is a quantity which Zimm and Bragg have called the "nucleation parameter." Several different interpretations have been given to Ψ and σ . In the simplest approach,⁽⁵⁾ Ψ is assumed to include both the stacking energy (associated with the difference between having a bonded and nonbonded segment and two bonded

segments in sequence) and the entropy associated with the multiplicity of configurations which the coil sequence may possess. In some treatments,⁽⁷⁾ however, Ψ is limited to the stacking energy and the coil entropy is evaluated separately. In the present paper, we adopt the latter approach.

Since we are interested primarily in the form of the phenomena, it is convenient to reduce the mathematical complexity to the point where simple analytic expressions can indeed be obtained and the forms in question are directly visible. Therefore, we will limit both coil and helical configurations (the latter now being regarded as somewhat flexible) to those corresponding to *one-dimensional* random flights. This is by no means an essential restriction, and it can be removed without incurring insuperable difficulties. On the other hand, the formulas involved become considerably more unwieldy. In the helix, we shall denote the effective length of the freely orienting segment by γ , while in the coil it will be denoted by λ . Further, we shall assume that γ corresponds to m amino acid (hydrogen-bonded) units and that λ involves n such units. A certain necessary loss of precision accompanies this procedure in that (a) m and n are chosen to be integers, (b) m and n may not be commensurate, (c) fractional segments involving numbers of amino acid units less than m or n will have to be identified as full segments. If, for example, the number of dissociated hydrogen bonds in a coil sequence is less than n , the coil is assumed to have *one* random flight segment. If the number of dissociated bonds lies between n and $2n$, the coil has two segments, and so on.

In the one-dimensional random flight model, the end-to-end distance of a coil, denoted by l , is an integral number (positive or negative) of lengths λ . Suppose the coil contains ν random flight segments, or

$$x = n\nu \quad (6)$$

amino acid units. Then, the number of random flight configurations consistent with given ν and l is

$$q_{l,x} = q_{l,\nu} = \frac{\nu!}{[\frac{1}{2}(\nu + l)]! [\frac{1}{2}(\nu - l)]!} \quad (7)$$

where l is even if ν is even, odd if ν is odd, and can vary between $+\nu$ and $-\nu$. A simple notation results from defining

$$r = (\nu - l)/2 \quad \text{or} \quad l = \nu - 2r \quad (8)$$

Then,

$$q_{l,\nu} = q_{r,\nu} = \nu! / (\nu - r)! r! \quad (9)$$

where now

$$r = 0, 1, 2, \dots, \nu \quad (10)$$

The quantity $q_{l,\nu}$ is the factor that the coil sequence (l, ν) contributes to the partition function of the molecule. Because of the lack of precision associated with the need to deal with fractional random flight segments, $q_{l,\nu}$ will have to go with coils having end-to-end distances l , but with

$$(\nu - 1)n < x \leq \nu n \quad (11)$$

The contribution to the partition function of a helical sequence can be derived in a similar fashion. Suppose there are μ random flight segments in the helical sequence. Then, the helix will consist of

$$h = m\mu \quad (12)$$

amino acid (hydrogen-bonded) units. The configurational part of the contribution to the partition function will be similar to the corresponding quantity for the coil exhibited in Eq. (9). Thus, we will have

$$q'_{t,h} = \frac{\mu!}{[\frac{1}{2}(\mu+t)]! [\frac{1}{2}(\mu-t)]!} = \frac{\mu!}{(\mu-r)! r!} \quad (13)$$

where t is the end-to-end distance for the helix (in units of γ), the counterpart of l for the coil, and

$$r = 0, 1, 2, \dots, \mu \quad (14)$$

However, the helix partition function also has contributions from the stacking interfacial energy Ψ , and the free energy of bonding. As in the original Zimm-Bragg model, this contribution may be represented as

$$q'_h = \sigma s^h \quad (15)$$

The total contribution of a helical sequence is then the product of the quantities appearing in Eqs. (13) and (15), namely

$$q_{t,h} = [\mu! / (\mu-r)! r!] \sigma s^h \quad (16)$$

If in a given total configuration of the molecule there are $n_{t,h}$ helices of length t containing h amino acid units and $n_{l,x}$ coils of length l containing x units, the term contributed to the partition function by this configuration is

$$\prod_{t,h} (q_{t,h})^{n_{t,h}} \prod_{l,x} (q_{l,x})^{n_{l,x}} \quad (17)$$

The partition function in the canonical ensemble is obtained by summing Eq. (17) over all sets $n_{t,h}, n_{l,x}$ consistent with the requirement (conservation of the N of amino acid units in the molecule),

$$\sum_{t,h} h n_{t,h} + \sum_{l,x} x n_{l,x} = N \quad (18)$$

including permutation of helices among helices and coils among coils. Thus, for the partition function in the canonical ensemble, we obtain

$$Q = \sum_{\mathbf{n}} \prod_{t,h} (q_{t,h})^{n_{t,h}} \prod_{l,x} (q_{l,x})^{n_{l,x}} \quad (19)$$

where the formal vector \mathbf{n} is meant to indicate the allowable sets of n 's, including the above-mentioned permutations.

3. THE CONSTANT-FORCE ENSEMBLE

Since we are interested in the thermoelastic properties of the molecule, the most convenient path does not involve the use of the canonical ensemble and the partition function Q appearing in Eq. (19). It is more convenient to work in the constant-force ensemble,⁽⁸⁾ where the relevant partition function is

$$\Delta = \sum_L e^{fL/kT} Q(L, T) \quad (20)$$

where f is the tensile force to which the molecule is subjected, L is the end-to-end distance of the molecule, and $Q(L, T)$ is the canonical ensemble partition function at fixed length L and temperature T . It is easy to show⁽⁸⁾ that

$$G = A - fL = -kT \ln \Delta \quad (21)$$

where G is the analog of a Gibbs free energy and A is the Helmholtz free energy of a molecule. It may also be shown⁽⁸⁾ that

$$\langle L \rangle = kT \{ \partial(\ln \Delta) / \partial f \}_T \quad (22)$$

where $\langle L \rangle$ represents the *average* end-to-end distance of the molecule at constant temperature T and force f . In the constant-force ensemble, L is permitted to fluctuate, but for a large enough molecule the fluctuation will be negligible, and L will have an essentially constant observed value (at given f and T) prescribed by $\langle L \rangle$. Equation (22) is the desired elastic equation of state.

It is necessary, however, to evaluate Δ , and for this purpose the form appearing in Eq. (20) is inconvenient since it requires the evaluation of $Q(L, T)$, the canonical ensemble partition function at L , and this is the very problem we wish to avoid in working with Δ rather than Q .

The various terms in the sum in Eq. (19) correspond to different values of L . The partition function Q in Eq. (19) corresponds to the zero-stress case. The value of L , corresponding to a given term in the sum in Eq. (19), is given by

$$L = \gamma \sum_{i,h} t n_{i,h} + \lambda \sum_{i,x} l n_{i,x} \quad (23)$$

so that Eq. (20) can be rewritten in the form

$$\Delta = \sum_{\mathbf{n}} \prod_{i,h} z_{i,h}^{n_{i,h}} z_{i,x}^{n_{i,x}} \quad (24)$$

where again the formal vector \mathbf{n} implies all sets of n 's consistent with Eq. (18), including permutations of helices among helices and coils among coils, i.e., over all allowable configurations. In Eq. (24),

$$\begin{aligned} z_{i,h} &= (e^{f\gamma/kT})^t q_{i,h} = v^t q_{i,h} \\ z_{i,x} &= (e^{f\lambda/kT})^l q_{i,x} = u^l q_{i,x} \end{aligned} \quad (25)$$

By counting the allowable permutations of helices among helices and coils among coils (the number of permutations going with a given set of n 's), Eq. (24) can be further rearranged to read

$$\Delta = \sum_{\mathbf{n}} \left\{ M! \prod_{t,h} \frac{(z_{t,h})^{n_{t,h}}}{n_{t,h}!} \right\} \left\{ M! \prod_{l,x} \frac{(z_{l,x})^{n_{l,x}}}{n_{l,x}!} \right\} \quad (26)$$

in which we have used the symbol

$$M = \sum_{t,h} n_{t,h} = \sum_{l,x} n_{l,x} \quad (27)$$

the number of coils equalling the number of helices except for negligible end effects. Since we are interested in the logarithm of Δ , we can represent it by the logarithm of the largest term in Eq. (26). Introducing Stirling's approximation for the factorials, and taking the variation with respect to the n 's subject to the constraints Eqs. (18) and (27), we obtain for the set of n 's going with the maximum term

$$n_{t,h} = \frac{M}{a} y^{-h} z_{t,h} = \frac{M\sigma}{a} \left(\frac{s}{y}\right)^h \frac{\mu!}{(\mu-r)! r!} u^{\mu-2r} \quad (28)$$

$$n_{l,x} = M\alpha y^{-x} z_{l,x} = M\alpha y^{-x} \frac{\nu!}{(\nu-r)! r!} v^{\nu-2r} \quad (29)$$

where α and y are undetermined multipliers whose values are fixed by Eqs. (18) and (27). Substitution of Eqs. (28) and (29) into the typical term of Eq. (26) yields

$$\ln \Delta = N \ln y \quad (30)$$

[since Eqs. (28) and (29) refer to the maximum term], so that determination of Δ is reduced to the determination of y . To accomplish this, we substitute Eqs. (28) and (29) into (27) and perform the indicated sums, which involve only geometric series. After summing, the quantity α can be eliminated between the two equations contained in Eq. (27), with the result

$$y^2 - \{V + sU\}y + s(1 - \sigma) VU = 0 \quad (31)$$

where

$$V = \left(v + \frac{1}{v}\right)^{1/n} \quad (32)$$

$$U = \left(u + \frac{1}{u}\right)^{1/m} \quad (33)$$

According to Eq. (21), combined with (30), the lowest free-energy solution corresponds to the largest root of Eq. (31). Thus, for the stable solution, we require the larger of the two roots. This is

$$y = \frac{V + sU}{2} + \left[\left(\frac{V - sU}{2} \right)^2 + \sigma s VU \right]^{1/2} \quad (34)$$

Previous studies⁽⁵⁾ show that, for most proteins, σ is a small number, of order 10^{-4} – 10^{-6} . Under these circumstances, the protein undergoes transition from coil to helix, or vice versa, abruptly over a range of sU or V of order $\sigma^{1/2}$. This is reflected in an abrupt change in the value of y [given by Eq. (34)] in this range. To determine the respective conditions under which coil and helix exist, we proceed as follows. We note that

$$s \left(\frac{\partial \ln \Delta}{\partial s} \right)_{u,v} = \frac{1}{\Delta} \left[\sum_n \left(\sum_{t,h} hn_{t,h} \right) \prod_{t,h} z_{t,h}^{n_{t,h}} z_{l,h}^{n_{l,h}} \right] = \langle \sum hn_{t,h} \rangle = m_H \quad (35)$$

where m_H is the average number of hydrogen-bonded amino acids in the molecule. When sU is large compared to V , Eq. (34) gives

$$y = sU \quad (36)$$

When V is large in comparison to sU , Eq. (34) requires

$$y = V \quad (37)$$

Substitution of Eqs. (36) and (37) alternatively into (30), and the results then substituted in Eq. (35), yields

$$m_H = N \quad (38)$$

for y having the value given by Eq. (36), and

$$m_H = 0 \quad (39)$$

when y is given by Eq. (37). The result in Eq. (38) indicates that the molecule is composed entirely of helix (all amino acids are hydrogen-bonded), while, in the case of Eq. (39), the molecule is all coil. Thus, when sU exceeds V , the molecule is helical; while, if the reverse is true, it is composed entirely of coil. For small values of σ , the transition takes place approximately when

$$sU = V \quad (40)$$

4. THERMOELASTIC RELATIONS

Although the parameter σ is small in most real systems, it is instructive to begin our examination of the thermoelastic relations which obtain in a protein molecule by considering the special case in which the helix-coil transition is noncooperative. This is the case in which $\sigma = 1$ or $\Delta F = 0$. (ΔF , it will be recalled, represents the free energy of bonding.) For this case, Eq. (34) gives

$$y = sU + V \quad (41)$$

Substituting this result into Eq. (35) and making use of Eqs. (32) and (33) yields for θ , the fraction of bonded or helical amino acids, the result

$$\theta = \frac{m_H}{N} = \frac{s[\cosh(\gamma f/kT)]^{1/m}}{s[\cosh(\gamma f/kT)]^{1/m} + [\cosh(\lambda f/kT)]^{1/n}} \quad (42)$$

If f is sufficiently large so that we may approximate the cosh terms in Eq. (42) by single exponentials, θ becomes

$$\theta = \left[1 + \exp \frac{[(\lambda/n) - (\gamma/m)]f + \Delta F}{kT} \right]^{-1} \quad (43)$$

where we have used the expression for s given in Eq. (3). We may consider Eq. (43) in four different cases. These correspond to:

$$\text{case a: } \Delta F < 0, \quad \gamma/m > \lambda/n$$

$$\text{case b: } \Delta F < 0, \quad \gamma/m < \lambda/n$$

$$\text{case c: } \Delta F > 0, \quad \gamma/m < \lambda/n$$

$$\text{case d: } \Delta F > 0, \quad \gamma/m > \lambda/n$$

In case a, at low enough temperatures or for sufficiently large negative values of ΔF , θ is approximately unity, so that the molecule is in the helical form. Since the first term in the exponential is negative, the application of additional stress (increasing f) merely brings θ closer to unity. Thus, in case a, stress stabilizes the helical form. In case b, the first term in the exponential is positive. As a result, for a sufficiently large value of f , θ will undergo a fairly abrupt reduction to zero, so that the molecule makes the transition, helix to coil, with the application of stress. The transition will be abrupt because we have assumed f large, and we are dealing with exponential behavior. Thus, even in the absence of cooperativity (σ), the application of stress, in case b, can lead to a sharp transition if ΔF is a sufficiently large negative number.

In case c, at low enough temperatures, θ will approximate zero and the molecule will be in the coil form. Since in this case the first term in the exponent in Eq. (43) is positive, the application of stress merely stabilizes the coil further. In case d, however, since the first term in the exponent is now negative, the application of sufficiently large stress will cause an increase of θ from zero to unity, so that a transition from coil to helix will have been induced.

Notice that, if $\lambda/n = \gamma/m$, stress has no effect on the system [the first term in the exponent in Eq. (43) vanishes], and the usual formula for the noncooperative, unstressed case is obtained.

In all cases, if we identify the transition by its midpoint, i.e., by $\theta = 1/2$, it occurs, according to Eq. (43), when

$$[(\lambda/n) - (\gamma/m)]f = -\Delta F \quad (44)$$

is satisfied.

We have elected to consider the situation for values of f sufficiently large so that Eq. (43) holds. Since this equation has a simple form, it is especially easy to discern the interplay of various physical effects. On the other hand, it should be clear that the same qualitative features of the phenomenon are to be expected when f is small, so that Eq. (42) must be used, or even when σ is small, so that Eq. (35) requires an even more complicated expression for θ than Eq. (42).

The important question to be answered at this juncture pertains to the probable sign of the inequality between γ/m and λ/n in the case of most proteins. These

quantities represent the length per amino acid in helical and coil, freely orienting segments, respectively. For a perfectly rigid helix (a situation for which we have not made allowance in our form of development), the symbol which would have to replace γ/m would be β , the length of an amino acid in the helix (see Appendix). Because of internal coiling in the freely orienting segment, it is probable that $\beta > \lambda/n$, and cases a and d would be involved. In these cases, as we have seen, the application of stress either stabilizes the helical form or induces it if the molecule is initially in the coil form.

Actually, however, we may expect that the helix will be somewhat flexible. The value of γ/m will still, however, exceed λ/n , and so we may still expect cases a and d to prevail with most proteins. This is the basis for our conclusion that stress is likely to stabilize or induce the helical form. In passing, it should be noted that Eq. (44) provides the condition for locating the transition, even in the case that σ is small, since it is really identical with Eq. (40). If the stress is large enough or if γ/m exceeds λ/n by an appreciable amount, the transition can be shifted to significantly higher temperatures by the application of stress.

5. BEHAVIOR OF THE ELASTIC MODULUS

In order to examine the elastic modulus, it is convenient to substitute Eq. (30) into (22), obtaining

$$\langle L \rangle = NkT \left(\frac{\partial \ln y}{\partial U} \right)_T \left(\frac{\partial U}{\partial f} \right)_T + \left(\frac{\partial \ln y}{\partial V} \right)_T \left(\frac{\partial V}{\partial f} \right)_T \quad (45)$$

We may define the modulus as

$$E = (\partial \langle L \rangle / \partial f)^{-1} \quad (46)$$

Obviously, as may be seen from Eqs. (45) and (46), the modulus changes as y changes. If y changes gradually, as it would in the noncooperative case, the modulus changes gradually. On the other hand, if σ is small (cooperative case), the modulus changes abruptly (with the helix-coil transition) as y changes abruptly. In any event, substitution of Eqs. (32)–(34) into Eqs. (45) and (46) allow the modulus to be computed for the general situation. For the cooperative case, the formula for E in the transition region is quite complicated and not very illuminating since it only shows the specific course of the abrupt transition in E . Of more interest are the values *between* which E changes. If we examine these limits for the cooperative case, then we may use y prescribed alternatively by Eqs. (36) and (37) for the helix and coil, respectively. Substitution of these into Eqs. (45) and (46), making use of Eqs. (32) and (33), yields, for E in the helical case,

$$E = E_h = \frac{mkT}{N\gamma^2} \left(\cosh \frac{\gamma f}{kT} \right)^2 \quad (47)$$

and in the coil case,

$$E = E_c = \frac{nkT}{N\lambda^2} \left(\cosh \frac{\lambda f}{kT} \right)^2 \quad (48)$$

Since for most proteins we may expect

$$\gamma^2/m = \gamma(\gamma/m) \gg \lambda(\lambda/n) = \lambda^2/n \quad (49)$$

it is apparent that the coil modulus will be significantly larger than that of the helix. Thus, the application of stress to a molecule initially in coil form will induce an abrupt transition to helix, and with it an abrupt reduction in the value of E from E_c to E_h . The molecule will therefore yield (the flexible helix will unfold) and avoid buildup of disruptive stress. Alternatively, if the molecule is originally helical, that state will be retained while the flexible helix unfolds, again avoiding the buildup of undue stress. Of course, when the helix has unfolded completely, the individual chemical bonds begin to accept stress and eventually fail. We have not included this possibility in our formal analysis.

It may be of some interest to work out an example where the course of the change in modulus can be described by a simple analytic expression. This is possible in the noncooperative case ($\sigma = 1$) when the transition is able to take place at small enough values of f so that U and V can be approximated by the following formulas:

$$\begin{aligned} U &= [1 + (\gamma f/kT)^2]^{1/m} \\ V &= [1 + (\lambda f/kT)^2]^{1/n} \end{aligned} \quad (50)$$

in which it is assumed that, in view of the smallness of f , both U and V are close to unity. Substituting Eq. (41) together with Eq. (50) into Eqs. (45) and (46), and using the fact that U and V are both close to unity at the transition (this also requires that s be approximately unity), we find the result for the modulus

$$E = \left(\frac{kT}{2N} \right) \frac{mn}{s\gamma^2n + \lambda^2m} \quad (51)$$

Several kinds of interesting thermoelastic relations are possible. For example, it can be shown, by working out the complete analysis for the case where helices are completely inflexible (see Appendix), and where

$$\lambda/n > \beta, \quad \Delta F < (\ln 2)/n \quad (52)$$

(in which β is the length of amino acid in the helix), that it may be possible to apply stress to the molecule, initially in the coil form, cause a transition to helix, and, with continued increase of stress, induce coil to reappear.

In view of the foregoing treatment, it is highly probable that the main effect of stress will be to either induce helix from coil or to sustain the helix, thereby protecting the molecule against disruptive stress.

APPENDIX

In this appendix, we present a brief treatment of the case in which helical sequences are completely rigid. For the one-dimensional situation, this means that, as we follow the chain from one end to another, helices will lie in both positive and

negative directions. Instead of the n 's dealt with in the body of this paper, we must now deal with the set $n_{l,x}, n_h, n_h^*$. Here, $n_{l,x}$ has the same meaning as before, but n_h refers to the number of rigid helical sequences containing h amino acid units and pointing in the positive direction, while n_h^* refers to helical units with h amino acids but pointing in the negative direction. With these definitions, we can now define the appropriate constant-force ensemble partition function

$$\Delta = \sum_{\mathbf{n}} \prod_h z_h^{n_h} z_h^{*n_h^*} \prod_{l,x} z_{l,x}^{n_{l,x}} \quad (\text{A1})$$

where, as in the body of this paper, \mathbf{n} denotes sets of n 's consistent with

$$\sum_h (n_h + n_h^*) = \sum_{x,l} n_{l,x} = M \quad (\text{A2})$$

(where M is the common number of either kind of sequence) and

$$\sum h(n_h + n_h^*) + \sum x n_{l,x} = N \quad (\text{A3})$$

which ensures conservation of amino acids. \mathbf{n} also refers to permutations of coils among coils and helices among helices. In Eq. (A1),

$$\begin{aligned} z_h &= \sigma s^h e^{f\beta h/kT} = \sigma s^h \xi^h \\ z_h^* &= \sigma s^h e^{-f\beta h/kT} = \sigma s^h \xi^{-h} \\ z_{l,x} &= e^{f\lambda l/kT} q_{l,x} = v^l q_{l,x} \end{aligned} \quad (\text{A4})$$

where $q_{l,x}$ has the same meaning as in Eq. (7), and β is the length of an amino acid unit in a rigid helix. Note that β is not the same as γ/m , since now the helix is regarded as completely rigid. By way of contrast, γ/m was the "effective" length of freely orienting segment per amino acid unit. Once again, we may show the permutations of coils among coils and helices among helices explicitly, and write Eq. (A1) in the form

$$\Delta = \sum_{\mathbf{n}} \left\{ M! \prod_h \frac{z_h^{n_h} z_h^{*n_h^*}}{n_h! n_h^*!} \right\} \left\{ M! \prod_{l,x} \frac{z_{l,x}^{n_{l,x}}}{n_{l,x}!} \right\} \quad (\text{A5})$$

Selecting the largest term in this sum, subject to the constraints Eqs. (A2) and (A3) yields

$$\begin{aligned} n_h &= \frac{M\sigma}{a} \left(\frac{\xi s}{y} \right)^h \\ n_h^* &= \frac{M\sigma}{a} \left(\frac{s}{\xi y} \right)^h \\ n_{l,x} &= M a y^{-x} v^l q_{l,x} \end{aligned} \quad (\text{A6})$$

Note that

$$\xi = e^{f\beta/kT}, \quad v = e^{f\lambda/kT} \quad (\text{A7})$$

Substitution of Eq. (A6) into the typical term of Eq. (A5) yields, once again,

$$\ln \Delta = N \ln y \quad (\text{A8})$$

so that the evaluation now of Δ depends upon the determination of y . Substitution of Eqs. (A6) into (A2) and elimination of the undetermined multiplier α between the two equations which result yields

$$\sigma \sum_n \left\{ \left(\frac{\xi s}{y} \right)^n + \left(\frac{s}{\xi y} \right)^n \right\} = \sum_{x,l} y^{-xv} q_{l,x}^{-1} \quad (\text{A9})$$

Substituting for $q_{l,x}$ the explicit formula given in Eq. (7), and performing the sums (which involve only geometric series), we obtain

$$\frac{y - V}{\sigma V} = \frac{\xi s}{y - \xi s} + \frac{s/\xi}{y - s/\xi} \quad (\text{A10})$$

This is a cubic equation whose roots are the allowable values of y . As before, we require the largest root (largest allowable value of y) for the stable solution having the lowest free energy.

When σ is small (of order 10^{-4} – 10^{-6}), excellent approximations to the three roots of Eq. (A10) are immediately available. These are arrived at in the following way: If y is not within the distance of order σ of either ξs or s/ξ , neither term on the right side of Eq. (A10) can be of order $1/\sigma$, a very large number. Therefore, the left side of Eq. (A10) cannot be of order $1/\sigma$, and this can only happen if the numerator is of order σ , i.e., if

$$y = V + O(\sigma) \quad (\text{A11})$$

Thus, neglecting terms of order σ , we have for one root

$$y_1 = V \quad (\text{A12})$$

On the other hand, if y does not have this value, then the left side of Eq. (A10) is a very large number, of order σ , a condition which can only obtain when the denominator of one of the two terms on the right is of order σ , i.e.,

$$y = \xi s + O(\sigma), \quad y = s/\xi + O(\sigma) \quad (\text{A13})$$

which yield the two additional approximate roots

$$y_2 = \xi s \quad (\text{A14})$$

$$y_3 = s/\xi \quad (\text{A15})$$

These approximations fail only within small ranges of f and T (which determines f , v , and ξ) when any two roots fall within distances σ of one another. We may therefore adopt Eqs. (A12), (A14), and (A15), choosing the largest of y_1 , y_2 , and y_3 in any range for the evaluation of $\ln \Delta$. As we proceed from range to range, the roots “cross,”

but, in fact, if we had evaluated them exactly, there would be no real crossing. Instead, one root would always be the largest and would smoothly transfer its behavior, within the apparent “crossing” range of width σ , from that of y_1 to y_2 , etc; if, for example, y_2 “crossed” to exceed y_1 at the crossing. As in the body of this paper, such crossings correspond to places where the helix–coil transition occurs. Since these are not singular and infinitely sharp (except when $\sigma \rightarrow 0$), it is to be expected that there will be a smooth but narrow *range* of transition. In our approximation, we ignore the detailed behavior within these ranges, but it can be obtained from the exact solution of the cubic equation (A10). Since, independently of which root is involved, y will always be a function of ξ and v , we can, after substituting Eq. (A8) into Eq. (22), write

$$\frac{\langle L \rangle}{N} = \lambda \xi \left(\frac{\partial \ln y}{\partial \xi} \right)_T + \beta v \left(\frac{\partial \ln y}{\partial v} \right)_T \quad (\text{A16})$$

Equation (35) still applies in the present case, and when y_1 is substituted into it, the result is

$$m_H = 0 \quad (\text{A17})$$

which indicates that, when y_1 is the largest root, the molecule is in the coil form. Alternatively, if y_2 is substituted, we find

$$m_H = 1 \quad (\text{A18})$$

so that, when this root is larger, the molecule is helical.

Since f is positive, $v > 1$ and y_3 is always less than y_2 . Thus, we need only consider the competition between y_1 and y_2 ; y_3 corresponds to situations in which the molecule points in the negative direction.

Substitution of y_1 and y_2 alternatively into Eq. (A16) yields, respectively,

$$\langle L \rangle / N = (\lambda/n) \tanh(\lambda f/kT), \quad y_1 > y_2 \quad (\text{A19})$$

$$\langle L \rangle / N = \beta, \quad y_2 > y_1 \quad (\text{A20})$$

The first of these conditions is the familiar relation between strain and stress, well known in the theory of gaslike, one-dimensional elastic chains. This is to be expected when y_1 is the dominant root, since the molecule is then a rubberlike coil. Equation (A20) corresponds to the helical condition, and, as might be expected, the average length proves to be independent of stress and equal to $N\beta$, the number of helical segments times the length of one of them.

The relative behaviors of y_1 and y_2 are examined most conveniently by using $kT \ln y_1$ and $kT \ln y_2$,

$$kT \ln y_1 = (kT/n) \ln(e^{\lambda f/kT} + e^{-\lambda f/kT}) \quad (\text{A21})$$

$$kT \ln y_2 = \beta f - \Delta F \quad (\text{A22})$$

where we have used the definition of s given in Eq. (3). Four cases immediately distinguish themselves. The first is illustrated in Fig. 1. Here, $\lambda/n > \beta$ and

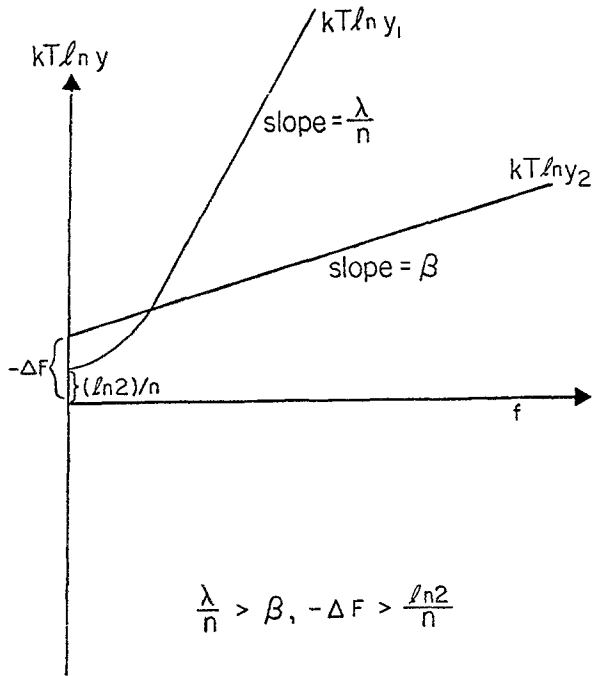


Fig. 1

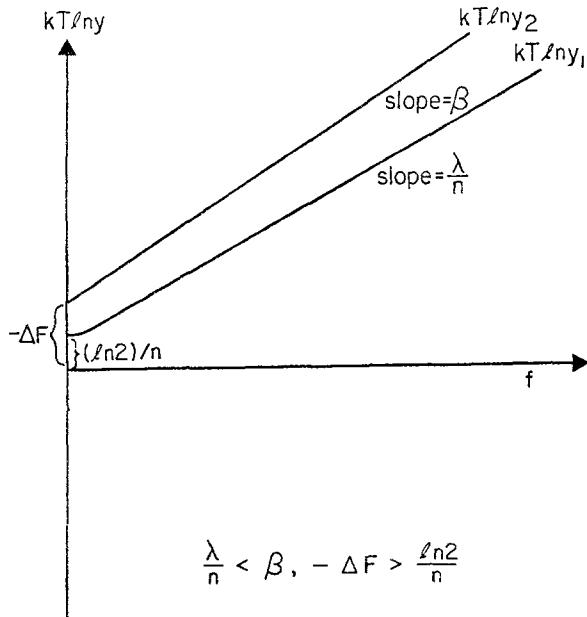


Fig. 2

$-\Delta F > (\ln 2)/n$. Under this circumstance, y_2 starts off (at $f = 0$) as the larger root. Thus, we begin with a molecule of helical form, but with increase of the applied stress, as may be seen from the plots of $kT \ln ny$ versus f in the figure, y_1 overtakes y_2 and the transition to the coil form takes place.

Figure 2 shows a case identical to the first, except that $\lambda/n < \beta$. Here, y_1 never overtakes y_2 and the molecule remains in the helical form regardless of how it is stressed. The freely orienting coil segment of length λ is already considerably "internally" coiled. Thus, λ/n , which represents the length per amino acid in this segment, is probably less than β , the corresponding length in the helix. Thus, for almost the same reasons involved in the treatment of the "flexible" helix, stress is seen to stabilize the helical form relative to coil.

Figures 3 and 4 illustrate cases in which $-\Delta F < (\ln 2)/n$. In Figure 3, $\lambda/n > \beta$, and there is no transition. The molecule starts out as a coil and stays that way. By the argument advanced above, however, we would expect Fig. 4 to represent the more typical situation. In this case, the molecule begins as a coil but undergoes a transition to helix upon the application of stress. Thus, we arrive at the same general conclusions for the case of the inflexible as the flexible helix.

There is an interesting modification of Fig. 3 which we present in Fig. 5. This is the case in which $-\Delta F < (\ln 2)/n$, $\lambda/n > \beta$, but where $-\Delta F$ is only slightly less than $(\ln 2)/n$. Here, it is possible for the curve belonging to y_2 to cross that of y_1 in two places in the manner shown in the figure. In the approximation that f is small

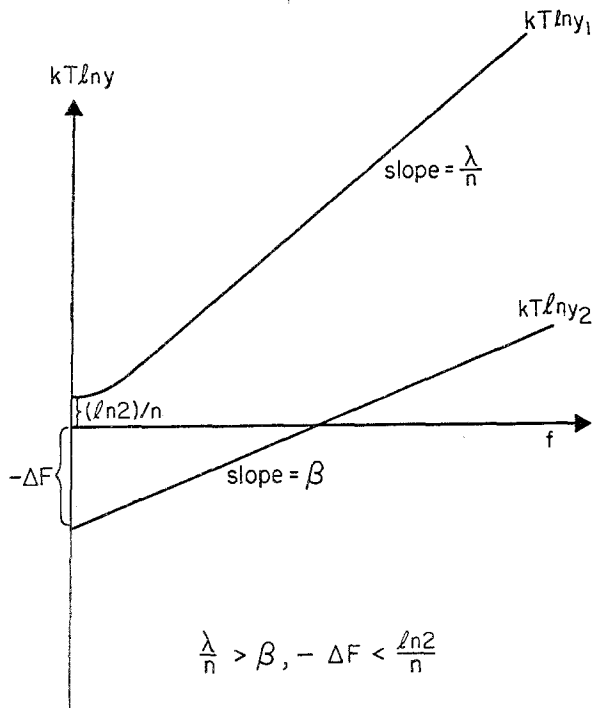


Fig. 3

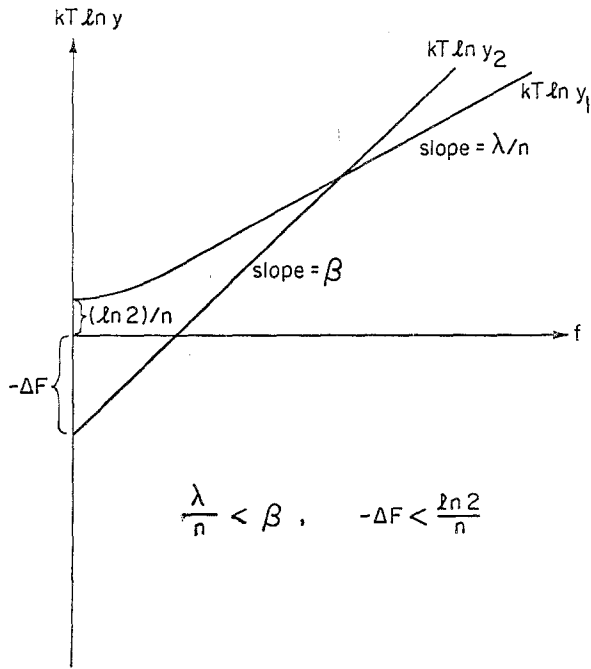


Fig. 4

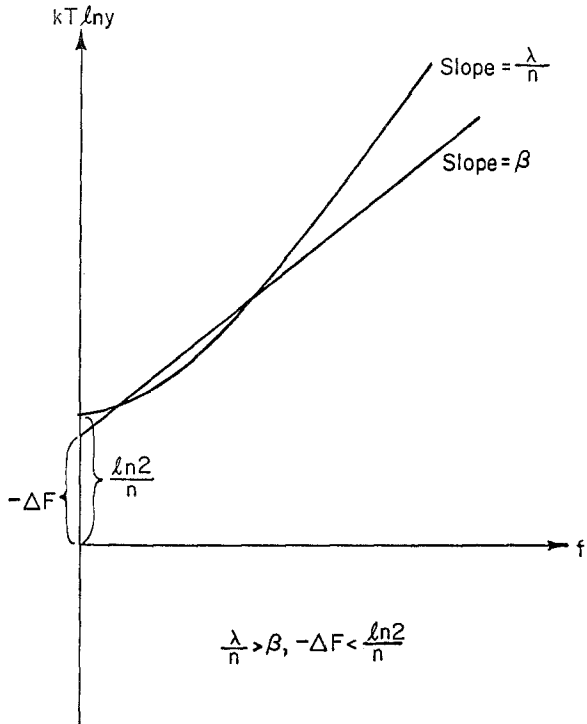


Fig. 5

at the locations of both crossings, the two values of f at these locations are approximately, respectively,

$$f = \frac{n\beta kT}{\lambda^2} \pm \left[\left(\frac{n\beta kT}{\lambda^2} \right)^2 - \frac{2kTn}{\lambda^2} \left(\frac{kT \ln 2}{n} + \Delta F \right) \right]^{1/2} \quad (\text{A23})$$

It is obvious that, for the value of f given by this equation to be real,

$$\frac{m\beta^2 kT}{\lambda^2} > 2 \left[\frac{kT \ln 2}{n} + \Delta F \right] \quad (\text{A24})$$

which therefore represents a condition which must be satisfied in order that the double crossing be possible.

When such a double crossing occurs, the molecule is initially in the coil form, makes a transition to helix at the first crossing, and makes a second transition back to coil at the second crossing.

REFERENCES

1. H. Curtis, *Biology* (Worth Publishers, New York, 1968), p. 80.
2. J. D. Ferry and P. R. Morrison, *J. Am. Chem. Soc.* **69**:400 (1947).
3. P. J. Flory, *J. Am. Chem. Soc.* **78**:5222 (1956).
4. P. J. Flory, *Principles of Polymer Chemistry* (Cornell University Press, Ithaca, New York, 1953), p. 411.
5. B. H. Zimm and J. K. Bragg, *J. Chem. Phys.* **31**:526 (1959).
6. G. W. Lehman and J. P. McTague, *J. Chem. Phys.* **49**:3170 (1968).
7. B. H. Zimm, *J. Chem. Phys.* **33**:1349 (1960).
8. T. L. Hill, *Introduction to Statistical Thermodynamics* (Addison Wesley, Reading, Massachusetts, 1960), Chapter 13.