Theory of the Cooperative Transition between Two Ordered Conformations of Poly(L-proline). II. Molecular Theory in the Absence of Solvent[1]

Seiji Tanaka[2a] and Harold A. Scheraga*[2b]

Department of Chemistry, Cornell University, Ithaca, New York 14853.
Received February 18, 1975

ABSTRACT: Phenomenological theories of the transition between helical form I (cis peptide bond) and helical form II (trans peptide bond) of poly(L-proline), which is a typical order ⇄ order transition, have been presented by Schwarz (using the parameters s, σ, β', and β'' in a 2 × 2 matrix formulation) and by the present authors (using the parameters s, σ, β_C, and β_N in a 4 × 4 matrix formulation). A molecular theory of the same transition has been formulated to account for the phenomenological parameters. The statistical weights of regular helical sequences with and without junctions between the two forms were computed from empirical potential energy functions. Two puckering conformations of the pyrrolidine ring, i.e., with the C^γ atom down and up, were allowed, and the free energy was computed for chains with four types of puckering, viz., regular down, regular up, random A, and random B, in the latter two of which the up and down puckerings were randomly distributed. The random A and random B chains have higher energy than those with regular down or up puckering, in both forms I and II. From both an energetical and a free energetical point of view, form I is more stable than form II under vacuum at room temperature. The dependence of the relative stabilities of form I and form II under vacuum on chain length was examined from both an energy and free energy point of view. The four parameters, s, σ, β', and β'', which describe the transitions in Schwarz's theory, were calculated from the statistical weights of various types of sequences. It was found that the thermally induced transition between form I and form II under vacuum occurs with the pyrrolidine rings remaining in the down conformation. The calculated values of s suggest that form I is more stable than form II in the regular down chain, while form II is more stable than form I in the regular up chain under vacuum at room temperature. The calculated values of σ for regular down and regular up pyrrolidine ring puckering are in good agreement with experimental observations, whereas those for random A and random B puckering are much smaller than the experimental values. A theory for the effect of solvent on the parameters s, σ, β', and β'' (at constant temperature) is developed, and the computations involving solvent effects are described in the next paper.

In the previous paper[3] of this series, we presented a general treatment of a one-dimensional phase transition in order to make clear and explicit the meaning of the parameters which appear in Schwarz's phenomenological theory[4] of the poly(L-proline) I ⇄ II interconversion, based on a nearest-neighbor Ising model. In the present paper, we present a molecular theory of this transition, in the absence of solvent. In paper III,[5] we will introduce the role of solvent in the molecular theory.

A molecular theory of the order ⇄ order transition in poly(L-proline) is of interest for the following reasons. (1) The form I ⇄ form II transition in poly(L-proline) differs from the helix ⇄ coil transition in the absence of both the intramolecular hydrogen bond of the helix and the entropy of the random coil. (2) In the theoretical formulation, and calculations therefrom, we do not have to consider the very large number of states in the random coil conformation. In the form I ⇄ form II transition, we can compute the long-range interactions which had to be neglected in the theory of the helix–coil transition because of the difficulty of treating the "excluded volume" effect and the large number of conformational states in the random coil in a polymer chain. Furthermore, whereas we had to assume that helical and coil sequences behave independently in the helix–coil transition,[6-9] we can take into account the long-range interactions between two neighboring sequences, each of which belongs to a different conformational phase, in the order ⇄ order transition. (3) In poly(L-proline), there is only one type of functional group (i.e., the carbonyl group) which can serve as a binding site in hydrogen bond forming solvents. This allows us to treat solvent effects in the poly(L-proline) transition more easily than in the helix–coil transition in polyamino acids that have two types of functional groups, i.e., the carbonyl CO and amide NH of the peptide group. Thus, while it was necessary to introduce an assumption that the binding constants K_{CO} and K_{NH}, for binding solvent molecules to these groups, are equal,[7-11] such an assumption is not required for poly(L-proline); this will make it easier to treat the effect of solvent on the transition in poly(L-proline).

The meaning of the parameters s, σ, β', and β'' of Schwarz's theory[4] was discussed in paper I.[3] In section I of this paper, we formulate the molecular theory of these phenomenological parameters. In section II, the numerical calculations are carried out. Finally, in section III, the numerical results are presented and discussed. Reference to equations in paper I will be made as eq I-1, I-2, etc.

I. Theoretical Formulation

A. Parameter s for the Growth Process. Consider the growth process illustrated by IV-1 in Table IV of paper I[3] for a segment consisting of $(2j + 1)$ residues, as depicted[12] in Figure 1, and re-define the statistical weights for the sequences[13] on the left- and right-hand sides of the equation representing the elementary process as $v_{j+1,j(\mathrm{II,I})}$ and $v_{j,j+1(\mathrm{II,I})}$, respectively. (See paper I for the relationship between the statistical weights used in the matrix of the one-dimensional Ising model and the statistical weights that are redefined here.) From eq I-48, the statistical weights of the segments can be written as

$$v_{j+1,j(\mathrm{II,I})} = w_{\mathrm{II,II}}{}^{j} w_{\mathrm{II,I}} w_{\mathrm{I,I}}{}^{j-1} \tag{1}$$

and

$$v_{j,j+1(\mathrm{II,I})} = w_{\mathrm{II,II}}{}^{j-1} w_{\mathrm{II,I}} w_{\mathrm{I,I}}{}^{j} \tag{2}$$

From eq I-49, I-50, I-62, 1, and 2, the parameter s may be calculated from the relation[14]

$$s = \lim_{j\to\infty} [v_{j,j+1(\mathrm{II,I})}/v_{j+1,j(\mathrm{II,I})}] \tag{3}$$

A factor u has been omitted from each of eq 1 and 2 (and from all subsequent similar equations) because it would cancel in eq 3 (and analogs thereof).

B. Nucleation Parameters. First, we consider the nu-

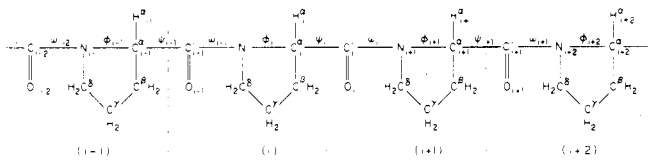

Figure 1. An interior segment of the poly(L-proline) chain. The IUPAC conventions and nomenclature[12] are used, except for the definition of a residue. A junction between form I (cis peptide bonds) and form II (trans peptide bonds) can occur, for example, at the dashed line between residues i and $i + 1$.

cleation process at the N terminus of a sequence. For the elementary process IV-2 in Table IV of paper I, the statistical weights of the sequences on the left- and right-hand sides (for sequences of $2j$ residues) are defined as $v_{2j(\mathrm{II,II})}$ and $v_{2j(\mathrm{I,II})}$, respectively. In a similar manner to eq 1 and 2, these statistical weights can be written as

$$v_{2j(\mathrm{II,II})} = w_{\mathrm{II,II}}{}^{2j-1} \tag{4}$$

and

$$v_{2j(\mathrm{I,II})} = w_{\mathrm{I,I}}{}^{j-1} w_{\mathrm{I,II}} w_{\mathrm{II,II}}{}^{j-1} \tag{5}$$

From eq I-49, I-50, I-53, I-62, 4, and 5, the nucleation parameter for form I at the N terminus of a sequence, σ'_{I}, may be calculated from the relation

$$\sigma'_{\mathrm{I}} = \lim_{j\to\infty} (1/s^j)[v_{2j(\mathrm{I,II})}/v_{2j(\mathrm{II,II})}] \tag{6}$$

In a similar manner, we obtain the nucleation parameter for form II at the N terminus of a sequence, σ'_{II}, by assigning the statistical weights $v_{2j(\mathrm{I,I})}$ and $v_{2j(\mathrm{II,I})}$ to the sequences on the left- and right-hand sides of elementary process IV-4 of Table IV of paper I. The resulting expression is

$$\sigma'_{\mathrm{II}} = \lim_{j\to\infty} s^j[v_{2j(\mathrm{II,I})}/v_{2j(\mathrm{I,I})}] \tag{7}$$

From eq I-64, 6, and 7, we immediately obtain the parameter β' as

$$\beta' = \lim_{j\to\infty} (1/s^{2j})\left[\left(\frac{v_{2j(\mathrm{I,II})}}{v_{2j(\mathrm{II,II})}}\right)\Big/\left(\frac{v_{2j(\mathrm{II,I})}}{v_{2j(\mathrm{I,I})}}\right)\right] \tag{8}$$

Second, we consider the nucleation process at the C terminus of a sequence. Defining the statistical weights, $v_{2j(\mathrm{II,II})}$ and $v_{2j(\mathrm{II,I})}$ for the sequences in elementary process IV-3 of Table IV of paper I, the nucleation parameter for form I at the C terminus of a sequence, σ''_{I}, and proceeding in a manner similar to the derivation of eq 6, we obtain

$$\sigma''_{\mathrm{I}} = \lim_{j\to\infty} (1/s^j)[v_{2j(\mathrm{II,I})}/v_{2j(\mathrm{II,II})}] \tag{9}$$

Similarly, the nucleation parameter for form II at the C terminus of a sequence, σ''_{II}, for elementary process IV-5 of Table IV of paper I is obtained as

$$\sigma''_{\mathrm{II}} = \lim_{j\to\infty} s^j[v_{2j(\mathrm{I,II})}/v_{2j(\mathrm{I,I})}] \tag{10}$$

where $v_{2j(\mathrm{I,II})}$ and $v_{2j(\mathrm{I,I})}$ are the statistical weights for the two sequences depicted in elementary process IV-5 of Table IV of paper I. From eq I-65, 9, and 10, the parameter β'' is

$$\beta'' = \lim_{j\to\infty} (1/s^{2j})\left[\left(\frac{v_{2j(\mathrm{II,I})}}{v_{2j(\mathrm{II,II})}}\right)\Big/\left(\frac{v_{2j(\mathrm{I,II})}}{v_{2j(\mathrm{I,I})}}\right)\right] \tag{11}$$

Third, the nucleation parameter for an isolated residue in the interior of a sequence, σ, may be computed from eq I-63, 6, and 7 or I-63, 9, and 10, viz.

$$\sigma = \lim_{j\to\infty} \{v_{2j(\mathrm{II,I})}/v_{2j(\mathrm{II,II})}\}\{v_{2j(\mathrm{I,II})}/v_{2j(\mathrm{I,I})}\} \tag{12}$$

Thus, all the parameters, s, σ, β', and β'', which determine the equilibrium behavior of the system of constant chain length N are expressed in terms of the statistical weights of the sequences.

C. Statistical Weights of the Helical Sequences with and without a Junction. In order to compute s, σ, β', and β'', it is first necessary to evaluate the statistical weights $v_{j,j+1(\mathrm{II,I})}$, $v_{j+1,j(\mathrm{II,I})}$, $v_{2j(\mathrm{I,I})}$, $v_{2j(\mathrm{II,II})}$, $v_{2j(\mathrm{I,II})}$, and $v_{2j(\mathrm{II,I})}$. In formulating the procedure to calculate the analogous statistical weights for the helix–coil transition,[6] using classical statistical mechanics to evaluate the partition function, integration of the Boltzmann factor over the coordinates and momenta led to two terms,[6] det **F** and det **G**. The evaluation of the term det **G** for the random coil presented difficulty because of the large number of conformations and the associated long-range interaction problem (excluded volume effect) in the random coil. In the order $\rightleftharpoons$ order transition in poly(L-proline), we do not encounter this difficulty because both conformations are regular helical ones (see also the discussion of section IB of paper I). While it is, thus, possible to evaluate the term det **G** for the two helical conformations, we nevertheless follow previous practice[6] and neglect it, since it is expected not to affect the computed parameters s, σ, β', and β'' significantly (because we expect it not to vary with the various sequences depicted in Table IV of paper I).

Since both conformations are helical, we may use the assumption that the conformational fluctuations around the regular helical conformations of poly(L-proline) (forms I and II) will be small.[6] Thus, the statistical weight v_n for a regular helical sequence[12b] of n proline residues, or for two regular sequences separated by a junction, is[6]

$$v_n = \exp[-E^0{}_n(Q_0)/RT][(2\pi RT)^m/\det \mathbf{F}_n(Q_0)]^{1/2} \tag{13}$$

where m is the number of variable dihedral angles in a sequence, $E^0{}_n(Q_0)$ is the value of the conformational energy at the minimum-energy conformation described by the set of coordinates Q_0, and the elements of the matrix $\mathbf{F}_n(Q_0)$ are the second derivatives of the conformational energy at the corresponding minimum-energy conformation Q_0.

In calculating $E^0{}_n(Q_0)$, the dihedral angle for rotation about the peptide bond,[12] ω, and that for rotation about the C^α–C′ bond, ψ, are taken as variables, and the dihedral angle for rotation about the N–C^α bond, ϕ, is fixed in either of two conformations (see section IIA). However, when det $\mathbf{F}_n(Q_0)$ is calculated, ω is assumed to be fixed at the values obtained by the energy minimization in which ϕ and ψ were allowed to vary, ϕ is fixed, and only ψ is allowed to vary; i.e., only the small conformational fluctuations around the C^α–C′ bond are taken into account in computing the entropy with *and* without junctions for the sequences of the poly(L-proline) chain depicted in Figure 1. Thus, in the calculation of $E^0{}_n(Q_0)$, there are $4j - 1$ variable dihedral angles in a sequence of $2j$ proline residues and, in the calculation of det $\mathbf{F}_n(Q_0)$, $m = 2j - 1$ for a sequence of $2j$ residues; $n = 2j$ in calculating both $E^0{}_n(Q_0)$ and det $\mathbf{F}_n(Q_0)$. (n is the number of pyrrolidine rings and m is the number of ψ's in a chain of $n = 2j$ residues.) In addition, we allow for variation in U and D puckering in the calculation of $E^0{}_n(Q_0)$ and det $\mathbf{F}_n(Q_0)$.

Consider first the statistical weights of regular helical sequences $v_{2j(\mathrm{I,I})}$ and $v_{2j(\mathrm{II,II})}$ which appear in eq 6, 7, 9, and 10. These are evaluated by computing the quantities $E^0{}_{2j(\mathrm{I,I})}$, $E^0{}_{2j(\mathrm{II,II})}$, det $\mathbf{F}_{2j(\mathrm{I,I})}$, and det $\mathbf{F}_{2j(\mathrm{II,II})}$, as described in sections IIC and IID.

Next consider the statistical weights of helical sequences with a junction between two different helical states, $v_{j,j+1(\mathrm{II,I})}$, $v_{j+1,j(\mathrm{II,I})}$, $v_{2j(\mathrm{I,II})}$, and $v_{2j(\mathrm{II,I})}$. Such a junction can occur at any C^α–C′ bond, e.g., at the dashed line be-

tween residues i and $i + 1$ in Figure 1. According to our previous computations[15] of the conformational energy of L-proline oligomers, there is only one low-energy sharp minimum (at $\psi \simeq 160°$) in the whole range of ψ for every combination of peptide bond (trans and cis) and pyrrolidine ring puckering conformations. In addition, the dependence of the energy on ψ at a cis–trans or trans–cis junction is sharper than that for a trans–trans junction in the di-L-proline fragment. The general features of these results[15] for the di-L-proline fragment can be expected to hold for poly(L-proline). Indeed, in the computations carried out here, the dependence of the energy of a junction residue on ψ (with two regular helices attached), and the dependence of the energy of each regular helix on ψ, are all very similar to that found for di-L-proline; i.e., the energy is restricted to a narrow range around $\psi \sim 160°$. In poly(L-proline), there is no additional type of interaction (such as an intramolecular hydrogen bond) determining the helical conformation, beyond the nonbonded and electrostatic types of interactions in the di-L-proline fragment. These features allow us to treat the fluctuations in ψ at a junction in the same manner as the fluctuations in ψ in a regular helix, as far as the computation of the entropy is concerned; i.e., the entropy may be evaluated by considering only the small fluctuations in ψ around the minimum-energy conformation at a junction between two helical sequences, as well as those in the regular helix portions of the sequences on both sides of the junction. This assumption gives us the advantage of using the interdependent nearest-neighbor sequence model. If we were to take into account the large conformational fluctuation at a junction, we would have to employ the independent sequence model.[14] With the above assumption, we compute the statistical weights for sequences that include a junction, e.g., $v_{2j(\mathrm{I,II})}$, $v_{j,j+1(\mathrm{II,I})}$, etc. For example, we calculate $v_{2j(\mathrm{I,II})}$ from eq 13, using the minimum-energy value $E^0{}_{2j(\mathrm{I,II})}$ and the corresponding det $\mathbf{F}_{2j(\mathrm{I,II})}$, for the sequence consisting of j residues in form I, j residues in form II, and a form I–form II junction at the middle residue.

D. Computation of s, σ, β', and β''. In this section, we discuss the relation of the thermodynamic quantities s, σ, β', and β'' to molecular quantities. Consider first the quantity s which is defined in terms of the statistical weights appearing in eq 3. From eq 13, we may write $v_{j,j+1(\mathrm{II,I})}$ as

$$v_{j,j+1(\mathrm{II,I})} = \exp[-E^0{}_{j,j+1(\mathrm{II,I})}/RT] \times [(2\pi RT)^{2j}/\det \mathbf{F}_{j,j+1(\mathrm{II,I})}]^{1/2} \quad (14)$$

In a similar manner, $v_{j+1,j(\mathrm{II,I})}$ can be expressed in terms of $E^0{}_{j+1,j(\mathrm{II,I})}$ and det $\mathbf{F}_{j+1,j(\mathrm{II,I})}$. From these equations and eq 3, we obtain

$$-RT \ln s = f_0 + RT \ln g_0 \quad (15)$$

where

$$f_0 = \lim_{j\to\infty}[E^0{}_{j,j+1(\mathrm{II,I})} - E^0{}_{j+1,j(\mathrm{II,I})}] \quad (16)$$

and

$$g_0 = \lim_{j\to\infty}[\det \mathbf{F}_{j,j+1(\mathrm{II,I})}/\det \mathbf{F}_{j+1,j(\mathrm{II,I})}]^{1/2} \quad (17)$$

The quantity $-R \ln g_0$ may be regarded as the change in the conformational entropy when a residue in form II is converted to form I. Assuming that the potential energy (and hence f_0 and g_0) is independent of temperature,[16–18] and using the relations $\Delta H_s = RT^2(\partial \ln s/\partial T)$ and $\Delta S_s = \partial(RT \ln s)/\partial T$, we obtain the change in enthalpy and entropy associated with the quantity s as

$$\Delta H_s = f_0 \quad (18)$$

and

$$\Delta S_s = -R \ln g_0 \quad (19)$$

Consider next the quantity σ which is defined in terms of the statistical weights appearing in eq 12. Expressing these statistical weights in terms of the corresponding quantities E^0 and det $\mathbf{F}$ in eq 13, and then substituting in eq 12, we obtain

$$-RT \ln \sigma = f_1 + f_2 + RT \ln (g_1 g_2) \quad (20)$$

where

$$f_1 = \lim_{j\to\infty}[E^0{}_{2j(\mathrm{I,II})} - E^0{}_{2j(\mathrm{II,II})}] \quad (21)$$

$$f_2 = \lim_{j\to\infty}[E^0{}_{2j(\mathrm{II,I})} - E^0{}_{2j(\mathrm{I,I})}] \quad (22)$$

$$g_1 = \lim_{j\to\infty}[\det \mathbf{F}_{2j(\mathrm{I,II})}/\det \mathbf{F}_{2j(\mathrm{II,II})}]^{1/2} \quad (23)$$

and

$$g_2 = \lim_{j\to\infty}[\det \mathbf{F}_{2j(\mathrm{II,I})}/\det \mathbf{F}_{2j(\mathrm{I,I})}]^{1/2} \quad (24)$$

Therefore, from eq 20, the change in enthalpy and entropy associated with the quantity σ is

$$\Delta H_\sigma = f_1 + f_2 \quad (25)$$

and

$$\Delta S_\sigma = -R \ln (g_1 g_2) \quad (26)$$

Combining eq 15 and 21–24 with eq 8, we obtain the quantity β' as

$$-RT \ln \beta' = (f_1 - f_2 - 2jf_0) + RT \ln (g_1/g_2) - 2jRT \ln g_0 \quad (27)$$

The change in enthalpy and entropy associated with the quantity β' may be expressed as

$$\Delta H_{\beta'} = f_1 - f_2 - 2jf_0 \quad (28)$$

and

$$\Delta S_{\beta'} = -R \ln (g_1/g_2) + 2jR \ln g_0 \quad (29)$$

The quantity β'', defined in eq 11, is obtained in a similar manner as

$$-RT \ln \beta'' = (f_3 - f_4 - 2jf_0) + RT \ln (g_3/g_4) - 2jRT \ln g_0 \quad (30)$$

where

$$f_3 = \lim_{j\to\infty}[E^0{}_{2j(\mathrm{II,I})} - E^0{}_{2j(\mathrm{II,II})}] \quad (31)$$

$$f_4 = \lim_{j\to\infty}[E^0{}_{2j(\mathrm{I,II})} - E^0{}_{2j(\mathrm{I,I})}] \quad (32)$$

$$g_3 = \lim_{j\to\infty}[\det \mathbf{F}_{2j(\mathrm{II,I})} - \det \mathbf{F}_{2j(\mathrm{II,II})}]^{1/2} \quad (33)$$

and

$$g_4 = \lim_{j\to\infty}[\det \mathbf{F}_{2j(\mathrm{I,II})} - \det \mathbf{F}_{2j(\mathrm{I,I})}]^{1/2} \quad (34)$$

Hence, the change in enthalpy and entropy associated with the quantity β'' is

$$\Delta H_{\beta''} = f_3 - f_4 - 2jf_0 \quad (35)$$

and

$$\Delta S_{\beta''} = -R \ln (g_3/g_4) + 2jR \ln g_0 \quad (36)$$

Hence, we have expressed all of the quantities s, σ, β', and β'', which describe the equilibrium behavior of the system (for a chain length of N), in terms of molecular quantities. Thus far, the discussion has pertained to quantities

under vacuum. In section E, we discuss the effect of solvent on the transition of poly(L-proline).

E. Effect of Solvent. Even though the effect of solvent on the form I ⇄ form II transition curve will not be treated explicitly until paper III,[5] an initial discussion of solvent effects is introduced here to consider how the parameters are affected by solvent.

It has been amply demonstrated by experiment that the solvent plays an important role in determining the relative stabilities of forms I and II in solution. For example, the form I helix is stable in *n*-aliphatic alcohols and pyridine, whereas the form II helix is stable in water, benzyl alcohol, trifluoroethanol, acetic acid, and formic acid.[19-25] Furthermore, poly(L-proline) can undergo an interconversion between forms I and II by a change of solvent composition.[19,23] While many experimental studies have been carried out on the thermally induced helix–coil transition in other polyamino acids,[26] as far as we know no reports have appeared on a thermally induced interconversion between forms I and II in poly(L-proline).

In general, solvent effects on transitions in polypeptides and proteins may be classified into three categories:[7] (1) direct binding of solvent molecules to the functional groups of the polymer, e.g., to the NH and CO groups; (2) effect of the solvent on the intramolecular interactions of the polymer chain, e.g., the effect of the dielectric constant on electrostatic interactions; and (3) mutual effect of the solvent and polymer on each other, e.g., hydrophobic bonding in aqueous solution.

The type (1) solvent effect has been known to play an important role in conformational transitions in polypeptides and proteins in solvents which form hydrogen bonds with the peptide group. Indeed, the helix–coil transition in polyamino acids in mixtures of active and inactive solvents has been explained in terms of this type of solvent effect.[7,8,10,11] Various experiments have shown that the type (1) solvent effect plays a dominant role in influencing the form I ⇄ form II transition in poly(L-proline).[22,24,25] According to the equilibrium study by Ganser et al.[23] the form I helix of poly(L-proline), which is stable in *n*-butyl alcohol, is transformed to the form II helix by the addition of benzyl alcohol or trifluoroethanol. Infrared studies[22,24,25] have proven that the alcohol OH groups are hydrogen bonded to the peptide CO groups, the latter being the only available functional groups in the poly(L-proline) chain. Moreover, these studies[22,24] lead to the conclusion that the transition is induced by the preferential binding of benzyl alcohol or trifluoroethanol molecules to residues which are in the form II conformation; i.e., whereas *both* *n*-butyl alcohol and benzyl alcohol bind to the peptide CO groups, the standard free energy of binding is sufficiently different for both alcohols in forms I and II so that variation in the relative amounts of the two alcohols can induce the I ⇄ II interconversion (this point will be discussed in paper III[5]).

It may be assumed that there is no cooperativity when solvent is bound to the peptide groups, and hence that interactions between bound solvent molecules may be neglected; i.e., to a good approximation, the equilibrium constant for the binding of a solvent molecule to a given peptide group may be assumed to be independent of the conformation of the neighboring peptide groups, but to depend only on the conformational state of the residue to which the solvent molecule binds. With this assumption, the statistical weight of a residue can be expressed by two terms, one resulting from the free energy of a peptide residue in the absence of solvent binding and the other from the free energy of binding a solvent molecule to the peptide CO group. The contribution of solvent binding to the statistical weight of a residue in forms I and II, respectively, is

$$z_{\mathrm{I}} = 1 + \sum_i K^{\mathrm{I}}{}_i a_i \tag{37}$$

and

$$z_{\mathrm{II}} = 1 + \sum_i K^{\mathrm{II}}{}_i a_i \tag{38}$$

where $K^{\mathrm{I}}{}_i$ and $K^{\mathrm{II}}{}_i$ are the equilibrium constants for formation of a hydrogen bond between solvent species i at activity a_i and the ith peptide CO group in forms I and II, respectively.[27]

Using eq 37 and 38, the statistical weight $v^{(s)}{}_{j,j+1(\mathrm{II,I})}$ (including the contribution of solvent binding) for the sequence consisting of j residues in form II and $j + 1$ residues in form I can be expressed as

$$v^{(s)}{}_{j,j+1(\mathrm{II,I})} = v_{j,j+1(\mathrm{II,I})} z_{\mathrm{II}}{}^{j} z_{\mathrm{I}}{}^{j+1} \tag{39}$$

The corresponding expression for the sequence consisting of $j + 1$ residues in form II and j residues in form I is

$$v^{(s)}{}_{j+1,j(\mathrm{II,I})} = v_{j+1,j(\mathrm{II,I})} z_{\mathrm{II}}{}^{j+1} z_{\mathrm{I}}{}^{j} \tag{40}$$

Therefore, using eq 3, 39, and 40, the parameter $s^{(s)}$ for growth in solution becomes

$$s^{(s)} = s(z_{\mathrm{I}}/z_{\mathrm{II}}) \tag{41}$$

In eq 39–41, the values without the superscript (s) are quantities computed in the absence of solvent binding.

Applying the same considerations to the statistical weights $v^{(s)}{}_{2j(\mathrm{II,I})}$, $v^{(s)}{}_{2j(\mathrm{II,II})}$, $v^{(s)}{}_{2j(\mathrm{I,II})}$, and $v^{(s)}{}_{2j(\mathrm{I,I})}$ in eq 8, 11, and 12 leads to the results

$$\beta'^{(s)} = \beta' \tag{42}$$

$$\beta''^{(s)} = \beta'' \tag{43}$$

and

$$\sigma^{(s)} = \sigma \tag{44}$$

Equations 42–44 show that, while the quantity s depends on solvent (and solvent composition), the quantities σ, β', and β'' do not depend on the solvent, as far as a type (1) solvent effect is concerned. The same conclusion was deduced for σ in the helix–coil transition.[7]

The type (2) solvent effect probably does not play an important role in the solvent-induced transition between form I and form II in poly(L-proline) because the transition takes place in an extremely narrow range of variation of solvent composition,[23] in which the effective dielectric constant would not be expected to vary very much. However, when the thermodynamic parameters are computed "in vacuum" [in order to be combined with the type (1) solvent effect, as indicated in eq 41, or to be used directly, as shown in eq 42–44], a dielectric constant must be selected to represent the type (2) solvent effect. The value chosen for the dielectric constant D will affect not only s but also σ, β', and β''. One could optimize the value selected for the dielectric constant (and the approximate "cut-off distance" used for computing electrostatic interactions) by trying to achieve a best fit of the computed and experimental values of σ, β', and β'' which are not influenced by a type (1) solvent effect [or by a type (3) solvent effect; see below]. However, because of the ambiguities in defining a local dielectric constant and a "cut-off distance", we have resorted to the usually used[17] value of $D = 2$, and will discuss the role of electrostatic interactions in determining the conformation of poly(L-proline) in section IIB.

The typical type (3) solvent effect is the hydrophobic bond,[28] which has been shown to affect the helix–coil transition of, say, poly(L-alanine) in water.[29-31] However, we would not expect a type (3) solvent effect to play a significant role in the poly(L-proline) transition in organic solvents because organic solvents do not have the structural

Table I
Chains with Specific Random Puckering

Notation	Random sequence of up (U) and down (D) for 30 residues[a]
Random A	DUUUUDDUDUUDDDUDUDUUDUUUDD-UUDU
Random B	DUDDUDUDUUDDUDDUUUDDUUUDUU-DDUD

[a] Obtained with a pseudo-random number generator.[33]

features of liquid water. If one were to evaluate the conformational stability of poly(L-proline) in aqueous solution, one would have to take into consideration the hydrophobic interactions between the large nonpolar groups of the pyrrolidine rings. In fact, Nemethy and Scheraga[28] estimated a large free energy change for the interaction of two pyrrolidine rings in water.

II. Numerical Calculations

In this section, we describe the numerical calculation of the thermodynamic quantities "in vacuum", i.e., with the inclusion of the type (2) solvent effect (by the choice of $D = 2$). The resulting numerical values of s, σ, β', and β'' are given in section III together with the effect of solvent on s.

A. Geometry of the Proline Residue and the Energy Functions. In the interior segment of the poly(L-proline) chain depicted in Figure 1, the bond lengths and bond angles are assumed to be fixed at the values used in ref 15; similarly, the energy functions are those used in ref 15. Further discussion of the geometry of the proline residue and the energy functions is given in ref 17. In the present paper, we allow ω to be a variable; i.e., we take account of possible slight deviations of the peptide group from planarity. The torsional energy function for rotation about the peptide bond is given in ref 17, and no intrinsic torsional energy function is used[17] for variation of ψ. As in our previous paper,[15] in the calculation of E^0 and det **F**, we employ two types of puckering conformations of the pyrrolidine ring, which are designated "down or D" and "up or U", respectively, depending upon whether the C^{γ}_i atom is in the position where ϕ is fixed at $-75.0°$ and χ^1 at $18.7°$, or in the position where ϕ is fixed at $-67.6°$ and χ^1 at $-6.1°$, with respect to the carbonyl C′ atom of the $(i-1)$th proline residue.[32]

In order to examine the effect of a distribution of ring puckering conformations on the conformation of the regular helix and on the parameters σ, β', and β'', the calculations were performed not only for the regular U and D conformations but also for two chains with randomly distributed puckering states; in the latter two chains, the distribution of puckering conformations was obtained on a computer using a uniform pseudo-random-number generator.[33] The two specific random chains ($n = 30$), designated as "random A" and "random B", have the puckering conformations shown in Table I. In the calculation of s, the statistical weight can be computed only for a regular sequence; therefore, only regular D and regular U ring puckering were considered when computing s.

B. Energy Minimization for Regular Helices. We consider first the rate of convergence, with increasing chain length, of the conformational energies and minimum-energy conformations, and then the relative stabilities of various regular helices, in order to obtain the behavior at infinite chain length.

To obtain the regular form I and form II helices (i.e., regular except, in some cases, with respect to ring puckering), the conformational energies were minimized with respect to ψ and ω (with the constraint that these be the same in every residue). The procedure, applied to chains beginning with a C′ atom and ending with a C^{α} atom, was the same as that used previously,[15] except that end groups were included in the calculations reported in ref 15. First, the computation was carried out for form I and form II sequences of poly(L-proline) with regular D puckering. The dependence of ψ, ω, and the conformational energies (expressed as energy per residue) on the number of residues in a helical sequence are shown in Figure 2, where E_{tot}, E_{elec}, E_{nb}, and E_{tor} denote the total, electrostatic, nonbonded, and torsional energies, respectively. The starting values of ω in this energy minimization were taken as 0° (cis) and 180° (trans) for form I and form II, respectively. The starting values of ψ were taken as those for minimum-energy cis–cis and trans–trans dipeptides, as computed previously.[15] As seen in Figure 2, the values of ψ, ω, and the conformational energies are almost constant for $n > 15$, and the convergence is satisfactory at $n = 30$. Similar calculations were carried out for other regular helical structures of poly(L-proline) with different ring-puckering conformations, and the results for $n = 30$ are shown in Table II.

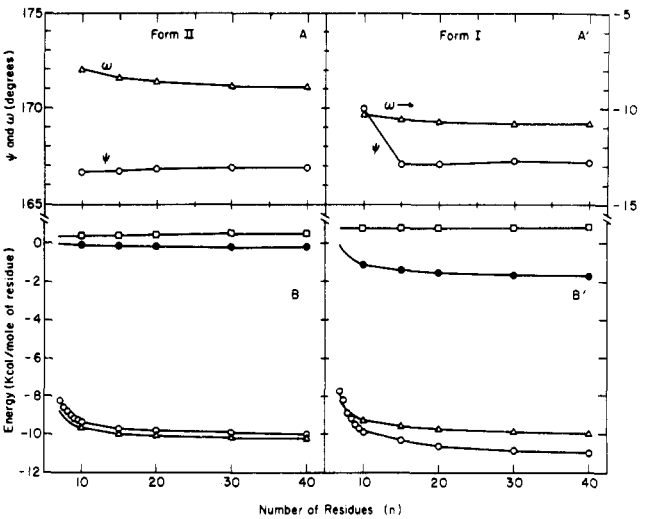

Figure 2. Dependence of the minimum-energy conformation and energy (per residue) on the number of residues of L-proline with regular D puckering, for form I and form II helices. In B and B′, the symbols represent: E_{tot} (○); E_{elec} (△); E_{nb} (●); and E_{tor} (□).

As seen in Table II, form I helices are *energetically* more stable than form II helices, no matter what the ring-puckering conformation is (as shown in section IID, form I helices are *entropically* less stable than form II helices, for each ring-puckering conformation). The stability of form I helices arises mainly from the nonbonded attractions because of their more compact structure (see Figures 3 and 4, where the minimum-energy structures of forms I and II with regular D puckering are shown). As can be seen in Figure 2, the electrostatic energies of both forms (for $D = 2$) are almost the same. Therefore, the relative stability of the two forms is insensitive to the choice of the value of D, and would become even more so if higher values of D were used to simulate the type (2) solvent effect discussed in section IE. In addition, it can be seen from Table II that the chains with regular D puckering are energetically more stable than the other three for both forms I and II.

It is of interest to point out that the relative stability of form I and form II varies with chain length. While the cis form is *energetically* more stable for long chains, as indicated above, the energy of the trans form is lower for short chains (in the absence of solvent), as shown in Table III and in Figure 2. As the chain length increases, the relative-

Table II
Regular Form I and Form II Conformations of Minimum Energy in Poly(L-proline)

Helical form	Puckering[a] conformation	Dihedral angles, deg		Energies, kcal/mol of sequence[b]			
		ω	ψ	E_{tot}	E_{elec}	E_{nb}	E_{tor}
Form II	Regular D	171.1	166.9	−297.2	−305.3	−6.0	14.2
	Regular U	171.5	176.0	−287.6	−306.8	5.9	13.4
	Random A	170.0	171.8	−283.5	−306.9	5.4	18.0
	Random B	170.1	171.3	−284.7	−306.8	4.3	17.8
Form I	Regular D	−10.8	167.2	−325.5	−296.8	−49.7	21.0
	Regular U	−17.2	170.6	−300.7	−304.4	−48.5	52.2
	Random A	−14.4	169.9	−302.4	−301.1	−38.4	37.2
	Random B	−13.9	169.8	−302.5	−301.0	−36.3	34.8

[a] D denotes the "down" conformation with respect to pyrrolidine ring puckering ($\phi = -75.0°$ and $\chi^1 = 18.7°$) and U denotes the "up" ($\phi = -67.6°$ and $\chi^1 = -6.1°$). (See text and ref 15 for more details.) Random A and B are the chains with random distribution of puckering conformation, given in Table I. [b] The energy values are given for sequences of $n = 30$.

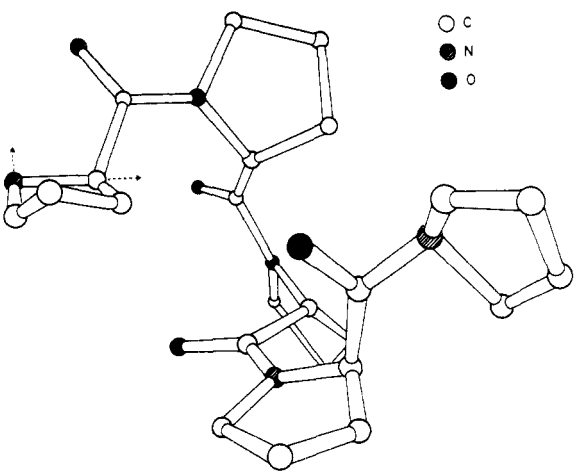

Figure 3. The form I (cis) helical conformation of minimum energy, with regular D ring puckering and $\phi = -75.0°$, $\chi^1 = 18.7°$, obtained in this paper. A sequence consisting of five pyrrolidine rings (~1⅔ turns of right-handed helix) is shown. The x and y axes lie in the plane of the paper at the first N–C$^\alpha$ bond, as shown. The size of the atomic symbols (with hydrogen atoms omitted) was selected to present a perspective along the z axis. The dihedral angles used for the drawing are those of Table II.

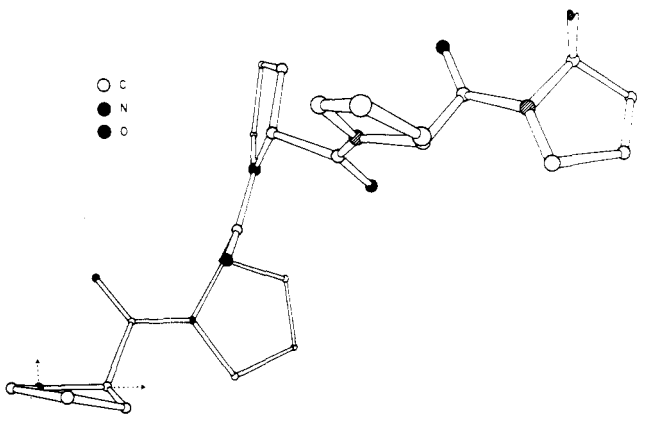

Figure 4. Same as Figure 3, but for the left-handed form II (trans) helix.

ly long-range nonbonded interactions stabilize the cis form over the trans because of the compact structure of the cis form.

C. Energy Minimization for Sequences with a Junction. In this section, we determine the conformation of the residue at the junction between two different forms. Let us suppose that the junction occurs at the dashed line between residues i and $i + 1$ in Figure 1. Up to (and including) residue i, the conformations (i.e., the values of the ψ's and ω's) are in form I; the conformation of the sequence consisting of residues $n \geq i + 1$ is in form II (likewise for II, I junctions). It is assumed that the dihedral angle ψ_i is not involved in either of the two helical sequences. The regular helical conformations on both sides of this junction were held fixed at the values given in Table II, and the energy was minimized by varying only ψ_i. The energy minimizations were carried out for $n = 30$, i.e., for 15 residues in each of forms I and II, and the results are summarized in Table IV (calculations for $n = 20$, i.e., for 10 residues on each side of the junction, yielded the same results).

Thus, all minimum-energy conformations, both for regular helical sequences of forms I and II and for the junction between them, in a sufficiently long chain, have been determined.

Table III
Differences in Energies and Free Energies of Short Oligomers[a]

No. of residues[b]	Difference in energy (trans − cis), kcal/mol of conformer	Difference in free energy[c] (trans − cis), kcal/mol of conformer
4	−3.0	−3.2
5	−3.1	−2.8
6	−1.0	−2.1
7	0.07	−1.3
8	1.3	−0.5
9	2.0	0.3
10	3.6	1.0

[a] For structures of the type shown in Figure 1, i.e., without end groups, and for D puckering. [b] For residues as defined in Figure 1, i.e., beginning with a C′ atom and ending with a C$^\alpha$ atom. [c] To be discussed in section IIIA.

In order to evaluate s, σ, β', and β'', we must examine the chain-length dependence and convergence of the energy terms f_0, f_1, f_2, f_3, and f_4 in eq 16, 21, 22, 31, and 32, respectively. For f_0, we need the conformational energies $E^0_{j,j+1(II,I)}$ and $E^0_{j+1,j(II,I)}$; these were computed for sequences of poly(L-proline) with regular D ring puckering as a function of the number of proline residues ($n = 2j + 1$), using the appropriate values of ψ, ω given in Tables II and IV, with j residues in form I and j in form II (Table II) and the junction residue from Table IV. The results, shown in

Table IV
Minimum-Energy Conformations of Sequences with a Junction

Junction at	Puckering[a] conformation	Dihedral angles ψ^b at the junction, deg	Energies, kcal/mol of sequence[c] E_{tot}[d]	E_{elec}	E_{nb}	E_{tor}
II–I	Regular D	165.9	−308.8	−299.1	−27.4	17.7
	Regular U	178.6	−289.8	−303.4	−18.7	32.3
	Random A	179.4	−287.7	−302.1	−12.6	27.0
	Random B	179.4	−283.9	−302.1	−7.5	25.7
I–II	Regular D	166.5	−309.1	−299.6	−27.3	17.7
	Regular U	179.1	−292.3	−303.5	−19.8	31.0
	Random A	179.6	−289.2	−302.0	−13.5	26.3
	Random B	179.6	−292.2	−301.6	−15.8	25.1

[a] See footnote *a* of Table II. [b] The other dihedral angles in a sequence are maintained at those in the regular helices (given in Table II) corresponding to form I and form II, and the puckering conformations of the pyrrolidine rings. [c] The energy values are given for the sequence of $n = 30$. [d] These values of E_{tot} are defined in the text as the junction energies $E_{2j(II,I)}$ and $E_{2j(I,II)}$.

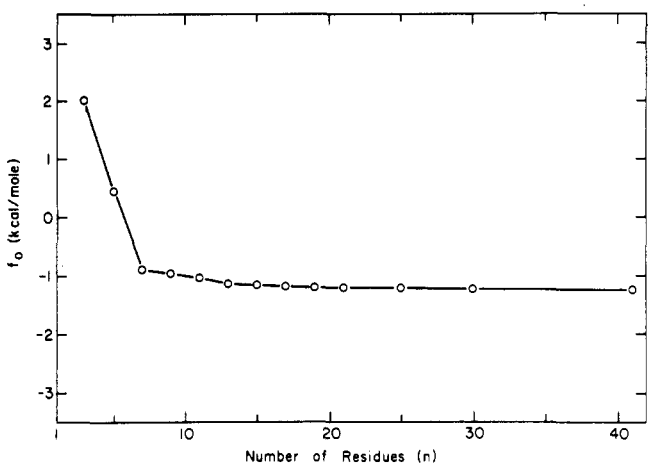

Figure 5. Dependence of $f_0 = E^0_{j,j+1(II,I)} - E^0_{j+1,j(II,I)}$ on the number of proline residues n ($=2j + 1$) for regular D ring puckering. The asymptotic limit of f_0 for large j gives ΔH_s (see eq 18).

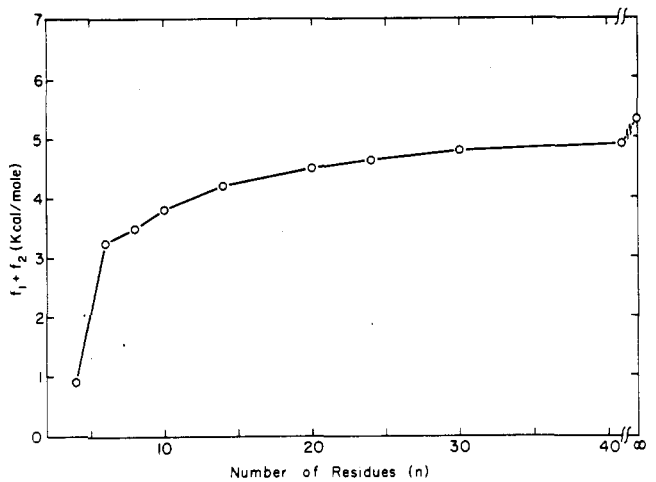

Figure 6. Dependence of $f_1 + f_2$ on the number of proline residues n ($=2j$) for regular D ring puckering. The values of $f_1 + f_2$, extrapolated to $n \rightarrow \infty$, give ΔH_σ for the infinite chain (see eq 25).

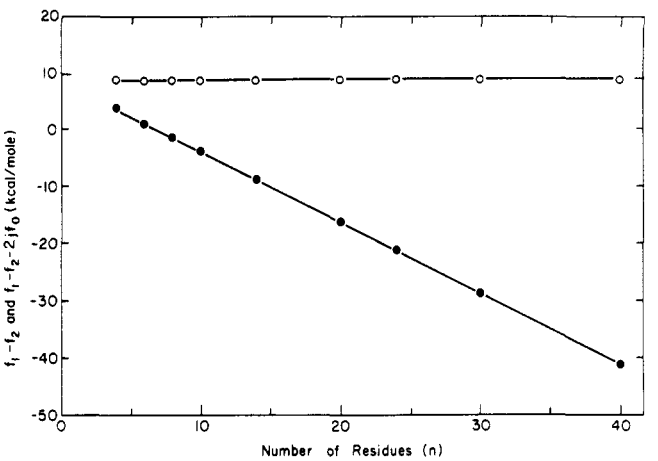

Figure 7. Dependence of $f_1 - f_2$ (●) and $f_1 - f_2 - 2jf_0$ (○) on the number of proline residues ($n = 2j$) for regular D ring puckering. The constant value of $f_1 - f_2 - 2jf_0$ gives $\Delta H_{\beta'}$ (see eq 28).

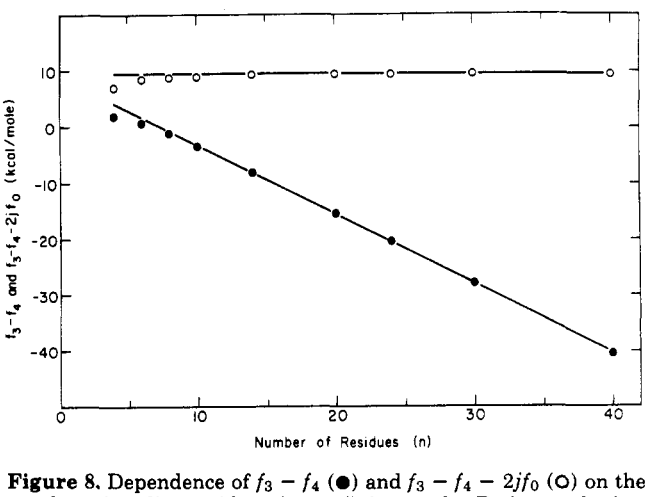

Figure 8. Dependence of $f_3 - f_4$ (●) and $f_3 - f_4 - 2jf_0$ (○) on the number of proline residues ($n = 2j$) for regular D ring puckering. The asymptotic value of $f_3 - f_4 - 2jf_0$ gives $\Delta H_{\beta''}$ (see eq 35).

Figure 5, indicate that f_0 converges rapidly as n increases, and that the convergence is satisfactory already at $n = 31$.

In a similar manner, f_1, f_2, f_3, and f_4, defined in eq 21, 22, 31, and 32, respectively, were computed for sequences with regular D ring puckering as a function of the number of proline residues. The resulting quantities, in the forms $f_1 + f_2$, $f_1 - f_2 - 2jf_0$, and $f_3 - f_4 - 2jf_0$ which appear in eq 20 or 25, 27 or 28, and 30 or 35, respectively, are plotted against the number of proline residues ($n = 2j$) in Figures 6, 7, and 8. The value of f_0, used in $f_1 - f_2 - 2jf_0$ and $f_3 - f_4 - 2jf_0$, is that for $n = \infty$ (of Figure 5). Figure 6 shows that the quantity $f_1 + f_2$, required for the computation of σ, converges slowly with increasing j, and has not yet converged at $n = 40$. Therefore, the value of $f_1 + f_2$ at infinite chain length was obtained by extrapolating these values (against $1/n$) to $n \rightarrow \infty$. The quantities $f_1 - f_2 - 2jf_0$ and $f_3 - f_4 - 2jf_0$, required for the computation of β' and β'', re-

spectively, become constant for $n > 14$ (see Figures 7 and 8).

Similarly, calculations of f_0, $f_1 + f_2$, $f_1 - f_2 - 2jf_0$, and $f_3 - f_4 - 2jf_0$ were carried out for sequences with regular U puckering to obtain s, σ, β', and β'', and calculations of $f_1 + f_2$ were carried out for sequences with random A and random B puckering to obtain σ. Since a chain length of $n = 30$ was sufficiently long to yield the asymptotic limits at $n = \infty$ for regular D puckering, a similar chain length was used here. The asymptotic values of f_0 and $f_1 + f_2$ give ΔH_s and ΔH_σ, respectively (see eq 18 and 25); thus, the asymptotic numerical values of f_0 and $f_1 + f_2$ for regular U puckering, and of $f_1 + f_2$ for random A and random B puckering, were taken from Table VIII (for ΔH_s) and Table IX (for ΔH_σ) below.

It can be seen in Table IV that, for each type of ring puckering, a I–II junction is *energetically* more stable than a II–I junction; i.e., it is preferable to have a helical sequence of cis followed by trans, rather than vice versa (for a chain considered from the N to the C terminus, as in Figure 1). This result, which was discussed in our previous paper,[15] is in good agreement with experimental NMR observations.[34] This point will be discussed from a free energy point of view in section IIIA.

D. Calculation of $\mathbf{F}_n$ for Sequences with and without a Junction. In order to evaluate the entropic contributions to s, σ, β', and β'' (by means of eq 19, 26, 29, and 36, respectively), it is necessary to calculate the contribution to the free energy of the polymer chain from the small fluctuations in the minimum-energy helical conformation. This entropic contribution to the free energy (or the statistical weight, in the form given in eq 13) may be evaluated from the curvature of the energy function at the minimum,[6] which requires the computation of the second derivatives of the energy function at the minimum; i.e., the matrix of second derivatives, $\mathbf{F}_n$, and its determinant must be calculated.

In calculating $\mathbf{F}_n$ for a sequence of n proline residues, we neglect the small conformational fluctuations in ϕ and ω, and consider only the fluctuations in ψ. The values of the ψ_i's (constrained to be equal in the regular helical segments), and of the ψ at the junction, were allowed to undergo small fluctuations around the minimum-energy conformation. Under these conditions, det $\mathbf{F}_n$ was computed; i.e., for the sequences of regular helices, det $\mathbf{F}_{2j(\mathrm{I,I})}$ and det $\mathbf{F}_{2j(\mathrm{II,II})}$, and for the sequences with a junction at the midpoint, det $\mathbf{F}_{2j(\mathrm{I,II})}$ and det $\mathbf{F}_{2j(\mathrm{II,I})}$ were computed as a function of the number of proline residues, $n = 2j$. The results are shown in Figure 9.

The following approximate method was used to compute det $\mathbf{F}_n$ for large j ($n = 2j \geq 20$) in order to save computer time. In the matrix $\mathbf{F}_n$ (of order $n \times n$), the (i,j) element is the second derivative of the energy, viz., $\partial^2 E/\partial\psi_i\partial\psi_j$ (with $1 \leq i, j \leq n$). Letting $k = |i - j|$, we know[6,35] that the terms for large k do not contribute to det $\mathbf{F}_n$ since the corresponding elements $\partial^2 E/\partial\psi_i\partial\psi_j$ are nearly equal to zero in a sufficiently long chain. At the same time, the product of the diagonal elements is large enough compared to the contribution from the nondiagonal elements that the latter may be neglected. This is not true if the polypeptide chain has side chains containing several bonds about which rotation can take place; also, the contribution from the nondiagonal elements of small k cannot be neglected for a *short* sequence. In the present case of poly(L-proline), with fixed geometry of the pyrrolidine ring and fixed dihedral angles ϕ and ω, two single bonds about which rotation takes place (even two successive ones, viz., ψ_i and ψ_{i+1}) are separated from each other by a large distance. As a result, the contributions to det $\mathbf{F}_n$ from the elements of large k decrease rapidly as k increases. In order to assess the contribution of the nondiagonal elements, we have computed the value of det $\mathbf{F}_n$ for $n = 10$ and 20 at various degrees of approximation, i.e., by neglecting all nondiagonal elements of k values larger than a particular one. The results are shown in Table V for $n = 10$. In the kth approximation (in which nondiagonal elements of all values of k larger than the value indicated are taken as zero), the values of det $\mathbf{F}_n$ are given for both form I and form II sequences with regular D puckering of the pyrrolidine ring.

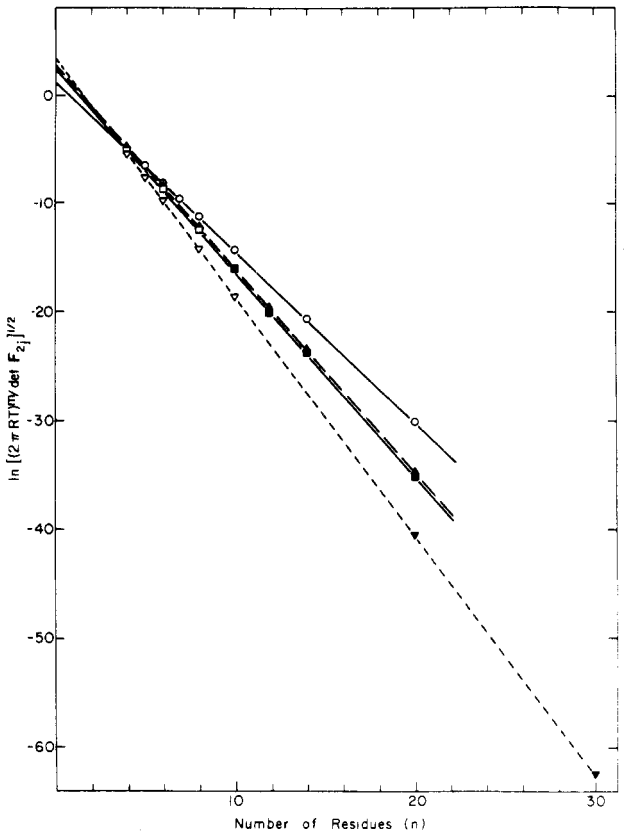

Figure 9. Dependence of ln $[(2\pi RT)^m/\det \mathbf{F}_{2j}]^{1/2}$ on the number of proline residues, n ($=2j$), for regular D ring puckering at a temperature of 300°K. The value of m is taken as $2j - 1$. The four straight lines correspond to the following sequences: I–I (▽); II–II (○); I–II (□); and II–I (△). The solid symbols were obtained by an approximate method described in the text.

Table V
Values of ln $[(2\pi RT)^m/\det \mathbf{F}_n]^{1/2}$ at Various Degrees of Approximation for the Off-Diagonal Elements

Approximation[a] k	$\{\ln [(2\pi RT)^m/\det \mathbf{F}_n]^{1/2}\}^b$	
	For form II	For form I
0	−40.02	−44.29
1	−40.02	−44.27
2	−40.02	−44.24
3	−40.02	−44.24
4	−40.02	−44.24
5	−40.02	−44.24
6	−40.02	−44.24
7	−40.02	−44.24
8[c]	−40.02	−44.24

[a] Terms up to the given value of k are included, and all terms of values of k greater than the indicated value are taken as zero. [b] $m = n - 1$ and $n = 10$ for the sequence of proline residues whose pyrrolidine rings have the regular D puckering conformation. [c] Values from the exact calculation.

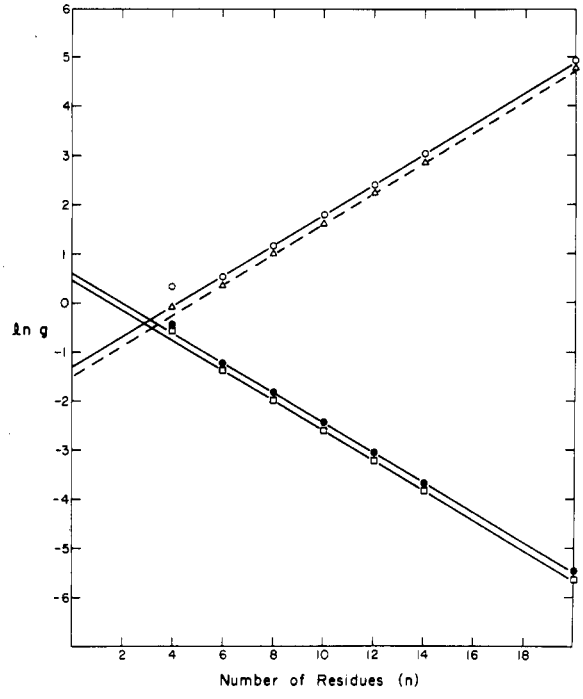

Figure 10. Dependence of ln g_1 (○), ln g_2 (□), ln g_3 (△), and ln g_4 (●) on the number of proline residues, n (=$2j$), for regular D puckering of the pyrrolidine rings.

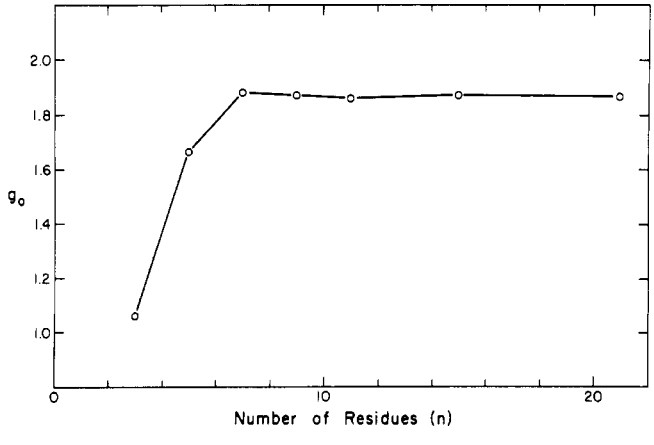

Figure 11. Dependence of g_0 on the number of proline residues, n (=$2j$ + 1), for regular D ring puckering. The asymptotic value of g_0 gives the entropic contribution, ΔS_s, to s (see eq 19).

The results in Table V (for n = 10) indicate that the k = 0 approximation, in which only the elements on the diagonal are taken into account, is sufficient to give the exact value of det $\mathbf{F}_n$ for the form II sequence, while the k = 2 approximation is required to reproduce the exact value for the form I sequence. A similar result was obtained for n = 20. Thus, for $n \geq 10$, we will use the k = 0 approximation for form II sequences, and the k = 2 approximation for form I sequences. For sequences with a junction, with j residues on each side of a II,I or I,II junction for g_1 and g_2 of eq 23 and 24, and with j and j + 1 residues on each side of a II,I (or I,II) junction for g_0 of eq 17, det $\mathbf{F}_n$ (where n = $2j$ or $2j$ + 1) was computed exactly for $n \leq 10$; for both II,I and I,II junctions, the k = 2 approximation was used for n = 11, and the k = 1 approximation was used for $n \geq 13$. These choices were based not only on the data of Table V, but also on the basis that values calculated exactly for ln $[(2\pi RT)^m/\det \mathbf{F}_{2j}]^{1/2}$ at small values of j extrapolate properly in Figure 9 (see below) to values calculated approxi-

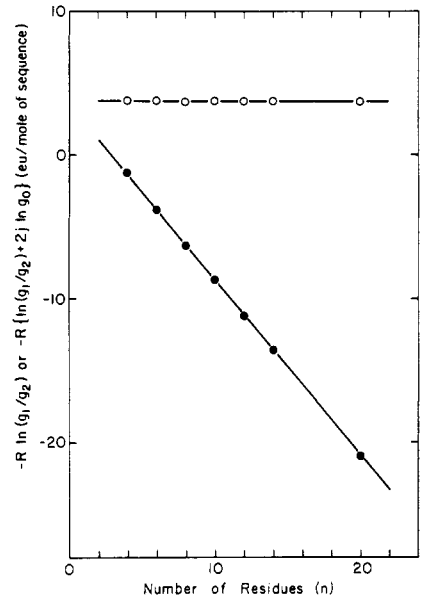

Figure 12. Dependence of $-R \ln (g_1/g_2)$ (●) and $-R\{\ln (g_1/g_2) + 2j \ln g_0\}$ (○) on the number of proline residues, n (=$2j$), for regular D ring puckering. The asymptotic value of $-R\{\ln (g_1/g_2) + 2j \ln g_0\}$ gives $\Delta S_{\beta'}$ (see eq 29).

mately at large values of j. Finally, it should be noted that the use of this approximate method does not mean that longer range interactions are neglected; all of the longer range interactions over the whole sequence have been taken into account in calculating the elements $\partial^2 E/\partial\psi_i\partial\psi_j$. We simply have eliminated the computation of terms of high k values. By this approximate method, we have saved considerable computer time. For example, in the computation of $\mathbf{F}_n$ for a form II sequence of n = 30, we have computed only the 30 diagonal elements, and not the 870 nondiagonal elements (and 144 of the 900 elements for form I).

As seen in Figure 9, all values of ln $[(2\pi RT)^m/\det \mathbf{F}_{2j}]^{1/2}$, for different combinations of sequences, rapidly approach straight lines at small values of n, which means that the interactions responsible for the determination of the entropy are mainly the short-range ones; i.e., the increment in ln $[(2\pi RT)^m/\det \mathbf{F}_{2j}]^{1/2}$ per residue becomes independent of n, so that additional contributions from long-range interactions are small or negligible. In addition, the approximate values obtained at large n lie on the extrapolation of the line based on exact values calculated at small n. It can also be seen that the entropic contribution is most positive for the form II structure for chains longer than about n = 5, while the form I helix is *energetically* more stable, as indicated in section IIB.

We have also examined the linearity of ln g_1, ln g_2, ln g_3, and ln g_4, defined in eq 23, 24, 33, and 34, respectively, with the number of proline residues. This linearity is shown in Figure 10.

The quantities g_0, $-R\{\ln (g_1/g_2) + 2j \ln g_0\}$, and $-R\{\ln (g_3/g_4) + 2j \ln g_0\}$, required in eq 15 or 19, 27 or 29, and 30 or 36, respectively, for sufficiently large n, are obtained from plots of these quantities against n (=$2j$) in Figures 11, 12, and 13. In computing the quantities plotted in Figures 12 and 13, the value of g_0 at sufficiently large n (from Figure 11) was used. Figure 11 illustrates that g_0 rapidly approaches its asymptote for $n > 8$. Figures 12 and 13 show that $-R\{\ln (g_1/g_2) + 2j \ln g_0\}$ and $-R\{\ln (g_3/g_4) + 2j \ln g_0\}$ are constant for all chain lengths plotted. The asymptotic values of g_0, $-R\{\ln (g_1/g_2) + 2j \ln g_0\}$, and $-R\{\ln (g_3/g_4) + 2j \ln g_0\}$ at sufficiently large n in Figures 11, 12, and 13, re-

Table VI
Stability of Helical Sequences of Form I and Form II[a]

Puckering[b] conformation	$\ln [(2\pi R)^m/\det \mathbf{F}_{2j(\mathrm{I,I})}]^{1/2}$[c]	$\ln [(2\pi R)^m/\det \mathbf{F}_{2j(\mathrm{II,II})}]^{1/2}$[c]	ΔH,[d] kcal/mol of sequence	$-T\Delta S$,[e] kcal/mol of sequence
Regular D	−145.1[f]	−129.0[g]	−28.3	9.7
Regular U	−145.2	−129.1	−13.1	9.7
Random A	−146.9	−129.7	−18.9	10.4
Random B	−146.9	−129.5	−17.8	10.5

[a] For a sequence with a number of residues, $n = 30$. [b] See footnote *a* of Table II. [c] For $m = 2j - 1$. The temperature T is not included in the factor $2\pi R$, but it is taken into account when ΔS is computed. [d] The values of ΔH are obtained from the energy differences of form I and form II, i.e., $\Delta H = E_{2j(\mathrm{I,I})} - E_{2j(\mathrm{II,II})}$, where $E_{2j(\mathrm{I,I})}$ and $E_{2j(\mathrm{II,II})}$ are given in Table II. [e] ΔS is the entropy difference (form I − form II), and the temperature T is 303°K. [f] Obtained by extrapolating the I-I line (▽) of Figure 9 to $n = 30$. [g] Obtained by extrapolating the II-II line (○) of Figure 9 to $n = 30$.

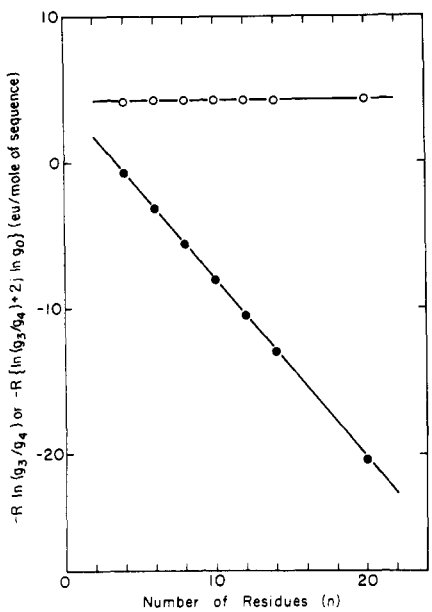

Figure 13. Dependence of $-R \ln (g_3/g_4)$ (●) and $-R\{\ln (g_3/g_4) + 2j \ln g_0\}$ (○) on the number of proline residues, n (=2j), for regular D ring puckering. The asymptotic value of $-R\{\ln (g_3/g_4) + 2j \ln g_0\}$ gives $\Delta S_{\beta''}$ (see eq 36).

spectively, give the entropic contributions (i.e., ΔS_s, $\Delta S_{\beta'}$, and $\Delta S_{\beta''}$) to s, β', and β'' (see eq 19, 29, and 36). Therefore, the asymptotic values of Figures 11, 12, and 13 are given by ΔS_s, $\Delta S_{\beta'}$, and $\Delta S_{\beta''}$ in Tables VIII, X, and XI. By using the values of $\ln g_1$ and $\ln g_2$ in Figure 10, the quantity $-R \ln g_1g_2$ can be computed. As can be seen in Figure 10, both $\ln g_1$ and $\ln g_2$ are linear in n, and vary in opposite directions; hence, $-R \ln g_1g_2$ is almost independent of n. The numerical values of $-R \ln g_1g_2$ (which gives ΔS_σ, according to eq 26) are given in Table I of paper III[5] for $n = 12$, 14, 30, and the asymptotic values are given in Table IX. Thus, the convergence of $-R \ln g_1g_2$ is satisfactory at $n = 12$.

Similar computations of g_0, $-R \ln g_1g_2$, $-R\{\ln (g_1/g_2) + 2j \ln g_0\}$, and $-R\{\ln (g_3/g_4) + 2j \ln g_0\}$ for regular U puckering, and only of $-R \ln g_1g_2$ for random A and random B puckering, were carried out for a chain length of $n = 30$, which is believed to be long enough to yield the asymptotic values (since $n = 12$ was satisfactory for convergence of the value of $-R \ln g_1g_2$ for chains with regular puckering). The asymptotic values of g_0, $-R \ln g_1g_2$, $-R\{\ln (g_1/g_2) + 2j \ln g_0\}$, and $-R\{\ln (g_3/g_4) + 2j \ln g_0\}$ for regular U puckering are given in the columns labeled ΔS_s of Table VIII, ΔS_σ of Table IX, $\Delta S_{\beta'}$ of Table X, and $\Delta S_{\beta''}$ of Table XI, respectively; the asymptotic values of $-R \ln g_1g_2$ for random A and random B puckering are given in the column labeled ΔS_σ of Table IX.

III. Results and Discussion

A. Stability of Form I and Form II Helices and of Sequences with a Junction. In section IIB, we found that the form I helix was energetically more stable than the form II helix and, in section IID, we saw that the form II helix was entropically favored over the form I helix. We now compare the stabilities of these two types of helices in terms of their free energies, by combining the results of sections IIB and IID; i.e., using eq 13, we calculate the free energy of the two helical conformations. In order to illustrate the effect of ring puckering, the enthalpic and entropic contributions to the free energy were calculated for four types of chains (regular D, regular U, random A, and random B) of $n = 30$ at 30°. In these calculations, we used the minimum-energy values of E_{tot} in Table II and $\ln [(2\pi R)^m/\det \mathbf{F}_{2j}]^{1/2}$ given in the second and third columns of Table VI.

For all four types of puckering (in long chains), the form I structure is enthalpically more stable but entropically less stable than that of form II. Since the entropic contributions do not compensate the enthalpic ones, as can be seen in Table VI, form I is always more stable than form II, i.e., has the lower free energy.

For *short* chains ($n < 7$), the trans form was shown to be *energetically* more stable (see Table III). We can now discuss the relative stabilities of cis and trans forms in short chains from a free energy point of view. Using the values of the conformational energies in Figure 2 and the entropic terms given in Figure 9, we obtain the differences in free energy listed in the last column of Table III. We see that, from a free energy point of view, the trans form is more stable for $n < 9$ but the cis form becomes the stable one for $n \geq 9$ (in the absence of solvent effects).

We next consider the effect of puckering on the stability of the form I and form II helices in long chains. For this purpose, we computed the relative statistical weights of each form by using eq 13 at 30°. The computed relative statistical weights of form II are 1, 10^{-7}, 10^{-10}, and 10^{-9} for the regular D, regular U, random A, and random B puckering, respectively, which implies that the regular D ring puckering conformation is the most stable one among these four types of form II helices under vacuum. In a similar manner, the relative statistical weights of form I are 1, 10^{-18}, 10^{-17}, and 10^{-17} for the regular D, regular U, random A, and random B puckering, respectively, again implying that the regular D ring puckering conformation is the most stable one among these four types of form I helices under vacuum. Finally, the computed statistical weights, relative to the *lowest energy* helix (form I with regular D puckering) are 1, 10^{-18}, 10^{-17}, 10^{-17}, 10^{-13}, 10^{-20}, 10^{-23}, 10^{-22} in the order: regular D, regular U, random A, random B of form I and in the same order of form II helices. This implies that the most stable helical conformation of poly(L-

Table VII
Stability of Sequences with a Junction between Form I and Form II[a]

Puckering[b] conformation	$\ln [(2\pi R)^m/\det \mathbf{F}_{2j(\mathrm{I,II})}]^{1/2c}$	$\ln [(2\pi R)^m/\det \mathbf{F}_{2j(\mathrm{II,I})}]^{1/2c}$	ΔH,[d] kcal/mol of sequence	$-T\Delta S$,[e] kcal/mol of sequence
Regular D	−136.5[f]	−136.6[g]	−0.37	0.10
Regular U	−137.3	−137.3	−2.45	0.01
Random A	−138.8	−138.7	−1.45	0.02
Random B	−137.4	−140.1	−8.28	−1.65

[a] For a sequence of $n = 30$. [b] See footnote *a* of Table II. [c] For $m = 2j - 1$. The temperature T is not included in the factor $2\pi R$, but is taken into account when ΔS is computed. [d] ΔH, i.e., enthalpy difference between I-II and II-I sequences, is calculated as $E_{2j(\mathrm{I,II})} - E_{2j(\mathrm{II,I})}$. The values of $E_{2j(\mathrm{II,I})}$ and $E_{2j(\mathrm{I,II})}$ for regular D, regular U, random A, and random B ($n = 30$) are given in Table IV. [e] Entropy difference between I-II and II-I sequences at 30°. [f] Obtained by extrapolating the I-II line (□) of Figure 9 to $n = 30$. [g] Obtained by extrapolating the II-I line (△) of Figure 9 to $n = 30$.

proline) under vacuum is the form I helix of regular D puckering. Unfortunately, this result cannot be compared with experimental data because poly(L-proline) is not soluble in nonpolar solvents, which would be the appropriate ones for this comparison. However, a relevant observation is the fact that the polymerization of the *N*-carboxyanhydride of L-proline in pyridine leads to poly(L-proline) as form I.[19] In addition, it is well known that the form I helix of this polymer can exist as a stable structure in the absence off solvent,[19] and that the transition of form I to form II takes place slowly when it is dissolved in such form II forming solvents as water, acetic acid, formic acid, and nonaliphatic alcohols;[19] no experimental results on the effect of ring puckering on the stability of the helical form are available.[36]

Next, we compare the stabilities of sequences with a junction. The enthalpic contributions (i.e., the values of E_{tot}) to the free energy of sequences with a junction are given in Table IV, and the entropic contributions, $\ln [(2\pi R)^m/\det \mathbf{F}_{2j}]^{1/2}$, are given in the second and third columns of Table VII; the resulting values of ΔH and $-T\Delta S$ are given in the last two columns of Table VII. As already pointed out in the discussion of Table IV, the data of Table VII show that a I–II junction is *energetically* more stable than a II–I junction. However, the data in the last column of Table VII demonstrate that the II–I junction (except for random B) is *entropically* more stable. Since the ΔH term outweighs the $-T\Delta S$ term, for all four types of puckering, the sequences with a I–II junction are very much more stable than those with a II–I junction, from a free energy point of view. This implies that the preferential direction of the transition of form II to form I is that from the *N* terminus to the *C* terminus and vice versa for the conversion from form I to form II. From NMR experiments,[34] it has been concluded that the form I helix of poly(L-proline) of low molecular weight in aqueous solution undergoes a transition from form I to form II starting at the carboxyl end of the chain. The theoretical results reported here are in good agreement with this observation; i.e., the transition from form I to form II takes place from the *C* terminus to the *N* terminus because I–II junctions are more stable than II–I junctions.

From a theoretical point of view, in the conversion from form II to form I, nucleation of I's can take place in *any* part of the chain, with a greater probability near the *N* rather than the *C* terminus. This statement is based on the form of the theory in which the probability of finding a I state can be computed from the free energy of the conformation of the chain; i.e., conformations in which I states are concentrated toward the *N* terminus of the chain are more stable than those in which II states are concentrated there, because I–II sequences are more stable than II–I sequences.

B. Results for *s* under Vacuum. In section II, we obtained all the numerical values which are required for the evaluation of the growth parameter, *s*, and the nucleation parameters, σ, β', and β''.

Table VIII
s and Related Quantities under Vacuum

Puckering[a] conformation	ΔH_s, kcal/mol of residue	ΔS_s, eu/mol of residue	T_m, °K	*s* at 25°	*s* at 70°
Regular D	−1.25	−1.24	980	4.28	3.26
Regular U	−0.07	−1.05	66	0.66	0.65

[a] See footnote *a* of Table II.

From the values of f_0 obtained in section IIC and g_0 obtained in section IID, we can calculate the enthalpic and entropic contributions to *s* (see eq 18 and 19). The quantities ΔH_s and ΔS_s for regular D and regular U chains are given in Table VIII; random A and random B puckering was not considered for the reason stated at the end of section IIA. The temperature dependence of *s* under vacuum is given by eq 15; the transition temperature, T_m, from form I to form II, defined as the temperature at which $s = 1$, is given in the fourth column of Table VIII. The resulting values imply that form I is more stable than form II in the regular D chain, while form II is more stable than form I in the regular U chain at room temperature under vacuum.[36]

In section IIIA, we concluded that forms I and II of poly(L-proline) with regular D puckering were more stable than those with regular U puckering. It is also possible to discuss the effect of puckering conformation on the stability of these two helical forms under vacuum from another point of view, viz., by comparing *s* under vacuum, obtained by conformational energy calculations, with the value obtained from experiments in which the effect of solvent is eliminated. This point will be discussed in paper III.[5] As indicated in section IE, *s* is markedly affected by a type (1) solvent effect. The dependence of *s* on solvent composition in binary solvent mixtures will be discussed in paper III.[5]

C. Results for σ, β', and β''. All the quantities required for the evaluation of σ, β', and β'' (by eq 20, 27, and 30, respectively) were computed in sections IIB, IIC, and IID. The enthalpic and entropic contributions to σ, computed by eq 25 and 26, are summarized in Table IX. As indicated in section IE, the values of σ are not expected to be affected by the dominant solvent effect [type (1)]. Therefore, we may compare the calculated values of σ under vacuum with those observed in solution. In making this comparison, it should be kept in mind that σ depends on temperature according to eq 20. From experimental studies[37] on the pressure dependence of the form I $\rightleftarrows$ form II transition, σ was determined as 3×10^{-3} at 27–28°. Ganser et al.[23] reported $\sigma = 1.0 \times 10^{-5}$ from equilibrium transition studies of

Table IX
σ and Related Quantities

Puckering[a] conformation	ΔH_σ, kcal/mol of sequence	ΔS_σ, eu/mol of sequence	σ at 25°	σ at 70°
Regular D	5.32	1.60	2.81×10^{-4}	9.12×10^{-4}
Regular U	6.23	−0.764	1.84×10^{-5}	7.30×10^{-5}
Random A	9.00	−1.96	9.36×10^{-8}	6.87×10^{-7}
Random B	11.1	−2.09	2.61×10^{-9}	3.04×10^{-8}

[a] See footnote *a* of Table II.

Table X
β' and Related Quantities

Puckering[a] conformation	$\Delta H_{\beta'}$, kcal/mol of sequence	$\Delta S_{\beta'}$, eu/mol of sequence	β' at 25°	β' at 70°
Regular D	8.64	3.73	3.01×10^{-6}	2.04×10^{-5}
Regular U	−13.5	−0.52	5.92×10^{9}	2.99×10^{8}

[a] See footnote *a* of Table II.

poly(L-proline) in *n*-butyl alcohol:benzyl alcohol at 70° (see paper III[5] for the reason why we do not consider here the values for σ for *n*-butyl alcohol:trifluoroethanol at 25 and 70°, where they[23] obtained 5×10^{-6} and $10^{-6} < \sigma < 10^{-5}$, respectively). The orders of magnitude of the theoretical values of σ for regular D and regular U puckering are in reasonable agreement with those observed experimentally.

The computed values of σ for random A and random B puckering are considerably smaller than the experimental values. This suggests that poly(L-proline) helices with this type of ring puckering do not exist. This conclusion is consistent with that obtained in section IIIA for the relative statistical weights. While we can thus rule out random A and random B ring puckering by these two criteria, we cannot conclude, from the above discussion of σ values, whether the preferred ring puckering is regular D or regular U in the poly(L-proline) helical forms. However, the results for s in section IIIB indicate that the puckering remains D in the I $\rightleftarrows$ II interconversion (see also ref 36).

The values of β' and β'', and their enthalpic and entropic contributions, can be computed from eq 28, 29, 35, and 36 by using the values obtained in sections IIB, IIC, and IID. The numerical results are given in Tables X and XI. In contrast to the values of σ, the values of β' and β'' computed for poly(L-proline) differ considerably between regular D and regular U ring puckering. The theoretical and experimental values of β' and β'' will be discussed in paper III[5] since, despite their independence of solvent (eq 42 and 43), some further discussion of the experimental data is required before the theoretical and experimental values can be compared.

IV. Conclusion

In this paper, the phenomenological theory for the order–order conformational transition proposed by Schwarz[4] and described by the present authors,[3] in which at least four parameters (s, σ, β', and β'') are required to describe the asymmetric nucleation, is interpreted in terms of a molecular theory. The dependence of the stabilities of the trans and cis forms of poly(L-proline) on chain length was discussed from both an energy and a free energy point of view. In the shorter chains of poly(L-proline), the trans form is more stable, but the cis one becomes more stable in long chains, under vacuum at room temperature. From a consideration of the effect of solvent on the parameters s, σ, β', and β'', it was concluded that solvent can affect the value of s markedly, whereas σ, β', and β'' cannot be influenced by the *dominant* solvent effect in which the solvent forms a hydrogen bond with the carbonyl group of a proline residue. Using empirical potential energy functions, the parameters s, σ, β', and β'' were computed under vacuum for various puckering conformations of the pyrrolidine ring. The computed values of σ for regular D and regular U chains are in good agreement with the experimental ones, whereas those for random A and random B chains are much smaller than the experimental ones. Finally, in principle, poly(L-proline) can undergo a thermally induced conformational transition from form I to form II with the puckering of the pyrrolidine ring remaining in the down conformation.

Table XI
β'' and Related Quantities

Puckering[a] conformation	$\Delta H_{\beta''}$, kcal/mol of sequence	$\Delta S_{\beta''}$, eu/mol of sequence	β'' at 25°	β'' at 70°
Regular D	9.39	4.40	1.19×10^{-6}	9.52×10^{-6}
Regular U	−8.58	−0.45	1.56×10^{6}	2.34×10^{5}

[a] See footnote *a* of Table II.

References and Notes

(1) This work was supported by research grants from the National Institute of General Medical Sciences of the National Institutes of Health, U.S. Public Health Service (GM-14312), the National Science Foundation (BMS71-00872 A04), and Walter and George Todd.
(2) (a) From Kyoto University, 1972–1975; (b) address correspondence to this author.
(3) S. Tanaka and H. A. Scheraga, *Macromolecules*, part I, preceding paper in this issue.
(4) G. Schwarz, *Biopolymers*, **6**, 873 (1968).
(5) S. Tanaka and H. A. Scheraga, *Macromolecules*, part III, following paper in this issue.
(6) N. Gō, M. Gō, and H. A. Scheraga, *Proc. Natl. Acad. Sci. U.S.*, **59**, 1030 (1968).
(7) M. Gō, N. Gō, and H. A. Scheraga, *J. Chem. Phys.*, **52**, 2060 (1970).
(8) M. Gō, N. Gō, and H. A. Scheraga, *J. Chem. Phys.*, **54**, 4489 (1971).
(9) M. Gō, F. T. Hesselink, N. Gō, and H. A. Scheraga, *Macromolecules*, **7**, 459 (1974).
(10) D. Poland and H. A. Scheraga, *Biopolymers*, **3**, 275 (1965).
(11) M. Bixon and S. Lifson, *Biopolymers*, **4**, 815 (1966).
(12) (a) The conentions and nomenclature proposed by an IUPAC-IUB Commission [*Biochemistry*, **9**, 3471 (1970)] are adopted here, except for the definition of a residue and the grouping of the atoms in the molecule. (b) In the present paper, we use the term "regular" in two senses: a "regular helical sequence" indicates that the dihedral angles ϕ_i, ψ_i, ω_i (for $i = 1$ to n) are equivalent for every i in a sequence; "regular D" and "regular U", as defined later, indicate that the puckering conformations of the pyrrolidine rings are equivalent for every residue in a sequence.
(13) See ref 21 of paper I.[3]
(14) In the helix–coil transition, Gō et al.[6–9] employed the *independent-sequence* model in which they neglected the interactions between neighboring sequences (except for the inclusion of the interaction of a single coil residue with the following helical sequence). However, this is not so in the present case; here, we take into account the interactions between the nearest-neighboring sequences of different conformational states. If we were to employ the independent-sequence model, eq 3 would have to be altered to

$$s = \lim_{j \to \infty} \left[\left(\frac{l'_{j+1(\mathrm{I})}}{l'_{j+1(\mathrm{II})}} \right) \Big/ \left(\frac{l'_{j(\mathrm{I})}}{l'_{j(\mathrm{II})}} \right) \right]$$

(15) S. Tanaka and H. A. Scheraga, *Macromolecules*, **7**, 698 (1974).
(16) In this paper, we use the empirical potential energy functions[17] obtained by analysis of the crystal packing of small molecules.[18] This potential energy function may not be independent of temperature over a wide range of temperature (see ref 18 for a more detailed discussion of this problem). However, the temperature dependence of these poten-

tial energy functions may be small in the range in which they are usually used. Thus, the assumption that the potential energy is independent of temperature may not lead to significant errors since we are not interested in the temperature-induced transition, but mainly in the solvent-induced transition. (However, see ref 36.)
(17) F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, *J. Phys. Chem.*, submitted for publication.
(18) F. A. Momany, L. M. Carruthers, R. F. McGuire, and H. A. Scheraga, *J. Phys. Chem.*, 78, 1595, 1621 (1974).
(19) See ref 4 of paper I.[3]
(20) F. Gornick, L. Mandelkern, A. F. Diorio, and D. E. Roberts, *J. Am. Chem. Soc.*, 86, 2549 (1964).
(21) F. A. Bovey and F. P. Hood, *J. Am. Chem. Soc.*, 88, 2326 (1966).
(22) H. Strassmair, J. Engel, and G. Zundel, *Biopolymers*, 8, 237 (1969).
(23) V. Ganser, J. Engel, D. Winklmair, and G. Krause, *Biopolymers*, 9, 329 (1970).
(24) H. Strassmair, J. Engel, and S. Knuf, *Biopolymers*, 10, 1759 (1971).
(25) S. Knof, H. Strassmair, J. Engel, M. Rothe, and K. D. Steffen, *Biopolymers*, 11, 731 (1972).
(26) For example, see D. Poland and H. A. Scheraga, "Theory of Helix-Coil Transitions in Biopolymers", Academic Press, New York, N.Y., 1970.
(27) While there is no NH group in poly(L-proline), there is one in other polyamino acids. In such cases, the statistical weight may be written as the product of terms corresponding to the CO and NH groups. In this connection, we point out a typographical error in eq 29 of ref 7, which should have been written as

$$b = -RT \ln (1 + \sum_k K_{\mathrm{NH},k} a_k) - RT \ln (1 + \sum_l K_{\mathrm{CO},l} a_l)$$

This typographical error appeared only in the final paper, and does not affect any of the computations reported in ref 7.
(28) G. Nemethy and H. A. Scheraga, *J. Phys. Chem.*, 66, 1773 (1962).
(29) M. Bixon, H. A. Scheraga, and S. Lifson, *Biopolymers*, 1, 419 (1963).
(30) D. C. Poland and H. A. Scheraga, *Biopolymers*, 3, 283 (1965).
(31) R. T. Ingwall, H. A. Scheraga, N. Lotan, A. Berger, and E. Katchalski, *Biopolymers*, 6, 331 (1968).
(32) In the up and down geometries used here, the bond lengths and bond angles have been maintained close to those observed by X-ray diffraction, as described in ref 17; thus the ϕ and χ^1 values must differ for the up and down gometries to maintain the X-ray geometry for the two types of puckering. Although we mentioned only C^γ puckering in describing the proline geometry in our previous paper,[15] the position of the C^β atom also changes slightly as ϕ is altered by $-7.4°$ from -67.6 to $-75.0°$. However, the dominant puckering of the pyrrolidine ring may be described as "C^γ puckering" because χ^1 is altered by 24.8° from the up ($-6.1°$) to the down (18.7°) ring conformation, which is larger than the variation ($-7.4°$) in ϕ. This discussion will not change any other part of our previous paper.[15]
(33) Uniform pseudo-random numbers were generated by using RANDU of the IBM Scientific Subroutine Package.
(34) D. A. Torchia and F. A. Bovey, *Macromolecules*, 4, 246 (1971).
(35) N. Gō, M. Gō, and H. A. Scheraga, *Macromolecules*, 7, 137 (1974).
(36) The relative statistical weights for the four types of ring puckering in form I and form II chains were computed as a function of temperature to investigate the thermally induced transition under vacuum. In ref 16, we pointed out that we were not concerned with the *thermal* transition in the text of this paper, and thus were not concerned about the temperature dependence of the potential energy. However, we consider the thermal transition only in this reference, and retain the assumption that the potential energy is independent of temperature even though we are considering a large temperature range here. From an energy (i.e., enthalpy) point of view, the transition takes place in the following order as the temperature is raised: form I (regular D) → form I (random A or random B) → form I (regular U) → form II (regular D) → form II (regular U) → form II (random A or random B) (see Table II). However, from a free energy point of view, form I (D puckering) is overwhelmingly the most stable conformation between 0 and 700°K [e.g., see the statistical weights (at 303°K) of section IIIA]. At 700°K, the relative statistical weights become 1, 10^{-7}, 10^{-7}, 10^{-7}, 10^{-1}, 10^{-4}, 10^{-6}, and 10^{-5} for regular D, regular U, random A, and random B of form I and regular D, regular U, random A, and random B of form II, respectively. At 885°K, a transition from regular D (form I) to regular D (form II) takes place; this may be regarded as an all-or-none transition for a chain length of $n = 30$. Above 885°K, and up to ~5000°K, form II (regular D puckering) is the most stable form. Above 5000°K form II is the most stable, but with a mixture of ring puckerings (regular D, regular U, random A, and random B); however, between 885 and 5000°K, the regular U, random A, and random B puckerings do not exist. Thus, the only thermally induced transition in the range of 0 to 5000°K is that of form I to form II, with regular D puckering in both forms.
(37) J. M. Rifkind and J. Applequist, *J. Am. Chem. Soc.*, 90, 3650 (1968).

Theory of the Cooperative Transition between Two Ordered Conformations of Poly(L-Proline). III. Molecular Theory in the Presence of Solvent[1]

Seiji Tanaka[2a] and Harold A. Scheraga*[2b]

Department of Chemistry, Cornell University, Ithaca, New York 14853.
Received February 18, 1975

ABSTRACT: Phenomenological theories of the form I ⇄ form II interconversion in poly(L-proline) have been presented by Schwarz (using the parameters s, σ, β', and β'' in a 2 × 2 matrix formulation) and by the present authors (using the parameters s, σ, β_C, and β_N in a 4 × 4 matrix formulation). In addition, a molecular theory was developed to compute s, σ, β', and β'' under vacuum. In this paper, we take into account the effect of solvent on the parameters s, σ, β', and β'' of the isothermal poly(L-proline) form I ⇄ form II interconversion. The growth parameter s is sensitive to the binding of solvent molecules to the peptide CO groups, but the nucleation parameters σ, β', and β'' are not affected by this type of solvent effect. The calculated values of s and σ under vacuum are in good agreement with the corresponding values derived from experimental data. By combining the theoretical values of s, σ, β', and β'' under vacuum with experimentally determined equilibrium constants for the binding of alcohols to the peptide CO groups (which differ in magnitude for form I and form II), it was possible to reproduce the experimental transition curves satisfactorily. Alternatively, the binding constants for alcohols, obtained by combining our theoretically computed parameters under vacuum with experimental equilibrium transition curves, are in satisfactory agreement with those evaluated independently by infrared spectral measurements of the binding of alcohols to the peptide CO groups. It is pointed out that significant errors may arise in analyzing experimental data if short chains are included with long chains in the determination of s, σ, β', and β'' from the equilibrium transition curves. The transition of poly(L-proline) from form II to form I when n-butyl alcohol is added to a solution of the polymer in benzyl alcohol is brought about by the slight difference in the binding free energies of both alcohols to the carbonyl groups of form II. The different binding affinities of the two alcohols, ROH, to form II may arise from (a) the different hydrogen-bond strength between the alcohol and the proline carbonyl group, and (b) possible differences in nonbonded and electrostatic interactions between the R group and the binding-site environment of the proline carbonyl group. The greater binding affinity of form II (compared to form I) for a *given* alcohol is attributed to the more open and extended conformation of form II.

In accompanying papers, we considered the phenomenological theory (paper I[3]) and a molecular theory in the absence of solvent (paper II[4]) for the form I ⇄ form II interconversion of poly(L-proline). In this paper, we consider