Thermodynamics of Formation of the Triple Helix from Free Chains and from Template-Constrained Chains of Collagen-like Monodisperse Poly(Gly-Pro-Hyp) Structures

Elsa Locardi,[†] Juliann Kwak,[†] Harold A. Scheraga,*[,‡] and Murray Goodman*[,†]

Department of Chemistry and Biochemistry, University of California at San Diego, La Jolla, California 92093-0343, and Baker Laboratory of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853-1301

Received: July 22, 1999; In Final Form: September 16, 1999

Statistical thermodynamic methods, developed for treating the α-helix−coil transition, are applied herein to describe the formation of the triple helix from short free chains and short template-constrained chains of collagen-like monodisperse poly(tripeptides), using poly(Gly-Pro-Hyp) as the example. For such short chains, application of the one-helical-sequence approximation indicates that there is very little unwinding from the ends, so that an all-or-none model is adequate to treat this transition. From the dependence of the helix nucleation and propagation parameters on chain length, concentration, and temperature, the thermodynamic parameters for formation of the triple helix from both free chains and template-constrained monodisperse poly(Gly-Pro-Hyp) chains are similar, and also similar to those for free poly(Gly-Pro-Pro) chains.

Introduction

The statistical thermodynamics of the formation of a triple helix has been studied extensively by a number of authors.[1−8] In this paper, we describe an evaluation of the thermodynamics for the formation of the triple helix of collagen-like monodisperse poly(tripeptides) not only from N-terminally capped free chains but also from template-constrained chains. The template-assembled analogues exhibit higher melting transition temperatures as well as faster triple helical folding rates, thereby demonstrating more efficient triple helical folding than their single-chain counterparts. In addition, we include an analysis considering the effect of the length of the peptide chains as well as the concentration of peptide in solution on triple helical folding.

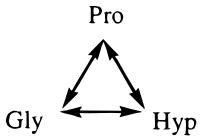
The compounds investigated are the acetyl analogues Ac-(Gly-Pro-Hyp)$_n$-NH$_2$ (where n = 3, 5, 6, 9) and the KTA conjugates KTA-[Gly-(Gly-Pro-Hyp)$_n$-NH$_2$]$_3$ (where n = 3, 4, 5, 6 and KTA denotes the Kemp triacid, *cis*,*cis*-1,3,5-trimethylcyclohexane-1,3,5-tricarboxylic acid). The chain-length dependence and concentration dependence of the equilibrium transition curves, published earlier,[9,10] were used for analysis of the thermodynamics of triple-helix formation.

It is well-known that the helix−coil transition in very short chains can be treated by an all-or-none (AON) theory, whereas short chains of moderate length can be treated by a one-helical-sequence model, where the unwinding occurs from the ends.[11] Since the chain lengths of the peptides are considered short, we treat a system of free chains initially by the one-helical-sequence model and then by the AON theory. These two treatments show that, for the short chains considered here, the AON model is more appropriate for free chains. Therefore, in treating the template-constrained chains, we use only the AON model.

We can define the state of the system by specifying the state of each structural unit as being either in the triple-helical region or in the non-triple-helical region. The structural unit of the growing helix is defined as the tripeptide unit (intrastrand)

Gly-Pro-Hyp

for the assembly of free-chain collagen-like sequences, and as the tripeptide register (interstrand)



for the template-assembled collagen-like sequences.

In both systems, free-chain and template-assembled collagen-like structures, triple-helix equilibria are pictured as a process of winding/unwinding from the ends of the peptide chains. Stepwise zippering of the triple helix requires the formation of a nucleus to initiate the triple helix, followed by folding of the chains to propagate the growth of the helix. The three peptide chains in the triple helix are assumed to be staggered by one residue from chain to chain.

Triple Helix−Coil Transition for Free Chains of Collagen-like Peptides

One-Helical-Sequence Model. Following Schwarz and Poland,[4] the partition function in this approximation may be written as

$$Z(n) = 1 + \sum_{m=1}^{n} (n - m + 1)^3 \sigma s^m \qquad (1)$$

The first term accounts for non-triple-helical structure; n is the number of tripeptide units for each strand; the factor $(n - m + 1)$ gives the number of ways to pick m tripeptide units in a helical conformation from a total of n; σ is the nucleation parameter indicating how much more difficult nucleation is

[†] University of California at San Diego.
[‡] Cornell University.

10.1021/jp9925435 CCC: $18.00
Published on Web 11/05/1999

compared with propagation; and s is the equilibrium constant for adding a tripeptide unit from each of the three chains to the triple helix.

Designating a single strand of (Gly-Xxx-Yyy)$_n$ as S_1 and the three-stranded complex as S_3, the intermolecular ensemble can be described by the following equilibrium

$$3S_1 \overset{K}{\rightleftharpoons} S_3 \tag{2}$$

with the constraint

$$[S_1] + 3[S_3] = C \tag{3}$$

where C is the total concentration of single strands. Designating the function Q as

$$Q = \sum_{m=1}^{n} (n - m + 1)^3 s^m \tag{4}$$

the equilibrium constant is then

$$K = \sigma Q = [S_3]/[S_1]^3 \tag{5}$$

The net fraction of helix θ_h is determined by

$$\theta_h = \frac{[\alpha]_{observed} - [\alpha]_{monomer}}{[\alpha]_{trimer} - [\alpha]_{monomer}} \tag{6}$$

where $\alpha_{observed}$ is the observed specific rotation, α_{trimer} is the specific rotation when the peptide is fully associated, i.e., in a triple helical structure (given by linear extrapolation of the premelting baseline from lower temperatures to the temperature under investigation), and $\alpha_{monomer}$ is the specific rotation of the monomer (determined by linear extrapolation of the specific rotation at high temperature to the temperature under investigation). This parameter describes the quantity of the fraction of molecules in the triple helix and the fraction of the helical array in the complex.

Following the derivation described by Schwarz and Poland,[4] eqs 7 and 8 are obtained

$$\left(\frac{\partial \ln \theta_h}{\partial \ln C}\right)_{T,n} = \frac{2\zeta}{3 - 2\zeta} \tag{7}$$

$$\frac{Q(s)}{Q'(s)} = \frac{1 - \zeta}{n\theta_h} \tag{8}$$

where ζ denotes the fraction of free single strands, and Q and Q' (where Q' denotes $\partial Q/\partial \ln s$) are expressed as finite polynomials in s, n, and m to obtain $s(T,C,n)$. Thus, the logarithmic slope of the association isotherm at each value of C at a given temperature gives $\zeta(T,C,n)$ (eq 7) which, with $\theta_h(T,C,n)$ in eq 8, gives $s(T,C,n)$ at a given temperature. In this calculation, it should be noted that, in contrast to the (single strand) α-helix−coil transition, where $s(T)$ is independent of n and C, $s(T)$ is a function of n and C for the triple helix−coil transition. Once $s(T,C,n)$, and hence the values of Q and Q', are known, the value for $\sigma(T,C,n)$ can be determined.

Measurement of $(\partial \ln \theta_h/\partial \ln C)_{T,n}$ at different temperatures yields the temperature dependence of $s(T,C,n)$ and hence the respective thermodynamic parameter ΔH_s.

$$\left(\frac{\partial \ln s}{\partial(1/T)}\right)_{C,n} = -\frac{\Delta H_s}{R} \tag{9}$$

While, in principle, σ should depend on temperature,[12] the experimental data are too sparse to detect the temperature dependence. Therefore, following the usual practice[13] we take

$$\Delta H_\sigma = 0 \tag{10}$$

Using eqs 9 and 10, s and σ can be written as

$$s(C,n) = \exp\left(\frac{-\Delta H_s}{RT} + \frac{\Delta S_s}{R}\right) = \exp\left(\frac{\Delta H_s(T - T_c)}{RTT_c}\right) \tag{11}$$

$$T_c = \frac{\Delta H_s}{\Delta S_s} \tag{12}$$

$$\sigma(C,n) = \exp\left(\frac{\Delta S_\sigma}{R}\right) \tag{13}$$

where it is assumed that ΔH_σ, ΔH_s, ΔS_s, and ΔS_σ are independent of temperature, and T_c is the melting temperature where $s = 1$ and $\theta_h = {}^1/_2$ for the infinite polymer ($n \rightarrow \infty$). The value of T_c is obtained by plotting the reciprocal of the melting temperature vs $1/n$ and extrapolating to $1/n \rightarrow 0$.

By repeating the above procedure at various temperatures, the value of ΔH_s can be obtained from the slope of the curves of $\ln s(T,C,n)$ vs $1/T$. The values of ΔS_s and ΔS_σ are calculated from eqs 12 and 13, respectively.

All-or-None Model. In the AON model, only the triple helix and free chains exist, without any intermediate (partially unwound) chains. Following the derivation described by Schwarz and Poland,[4] the equilibrium constant for the all-or-none model is written as

$$K = \sigma s^n = \theta_h/3C^2(1 - \theta_h)^3 \tag{14}$$

Implicit differentiation of eq 14 leads to

$$\left(\frac{\partial \ln \theta_h}{\partial \ln C}\right)_{T,n} = \frac{2(1 - \theta_h)}{1 + 2\theta_h} \tag{15}$$

$$\left(\frac{\partial \ln \theta_h}{\partial(1/T)}\right)_{C,n} = \left(-\frac{n\Delta H_s}{R}\right)\left(\frac{1 - \theta_h}{1 + 2\theta_h}\right) \tag{16}$$

At $\theta_h = {}^1/_2$, eqs 14−16 become

$$K = 4/3C^2 \tag{17}$$

$$\left(\frac{\partial \ln \theta_h}{\partial \ln C}\right)_{T,n} = 1/2 \tag{18}$$

$$\left(\frac{\partial \ln \theta_h}{\partial(1/T)}\right)_{C,n} = -\frac{n\Delta H_s}{4R} \tag{19}$$

Equation 18, evaluated at $\theta_h = {}^1/_2$, provides a useful test for the all-or-none approximation. The slope of the thermal transition curve evaluated at the midpoint as a function of chain length gives ΔH_s (eq 19). Knowing ΔH_s, the parameters ΔS_s and σ can be evaluated from the chain-length dependence of T_m as described in eq 20.

$$\frac{1}{T_m} = \frac{1}{n}\left(\frac{R \ln \sigma - R \ln(4/3C^2)}{\Delta H_s}\right) + \frac{\Delta S_s}{\Delta H_s} \tag{20}$$

Triple Helix−Coil Transition of Template-Assembled Collagen-like Peptides

One-Helical-Sequence Model. By analogy with the formation of the triple helix from free collagen-like chains, the

assembly of triple-helical structures in covalently bridged sequences is cooperative since the formation of the first stretch of structure (nucleation) is thermodynamically more difficult than the growth of structure via propagation. Hence, following Roth and Heidemann[7]

$$C \overset{\sigma s}{\rightleftharpoons} H \overset{s}{\rightleftharpoons} H_2 \overset{s}{\rightleftharpoons} - - - \overset{s}{\rightleftharpoons} H_{3n-2} \tag{21}$$

where C is the unassembled structure, H is the nucleus, H_{3n-2} is the triple helix with $3n - 2$ units from chains of n tripeptides each in the triple-helical register (the "2" accounts for both ends being staggered by one residue per chain), and σ is the nucleation parameter which, in this case, is independent of the concentration of the single-stranded chains since the entire process proceeds intramolecularly.

The partition function for this model is particularly simple. By analogy with the α-helix model,[11] it is useful to define the elementary structure-formation step as the change of a structural unit (tripeptide register) from the unassembled to the triple-helical assembled structure. This step is associated with the formation of a hydrogen bond between a Gly and a Hyp of a neighboring chain. That is, $Z(n)$ can be expressed as a function of the number (m) of trimeric units in a defined triple-helical structure out of $3n - 2$, the total number of trimeric units that can possibly take part in the triple helix. A factor of $(3n - m - 1)$ arises from the number of ways of placing a sequence of m trimeric units within three chains of n units each.

$$Z(n) = 1 + \sum_{m=1}^{3n-2} (3n - m - 1)\sigma s^m \tag{22}$$

If the template-assembled sequences constrain the helical section to start at the template junction, then the factor $(3n - m - 1)$ in eq 22 would not apply; hence, it is omitted for the short chains.

All-or-None Model. As will be shown in the Results section, the AON model is adequate to describe the helix−coil transition for the short free chains. Therefore, it is reasonable to expect that the AON model will also be applicable to treat template-constrained chains.

The range of validity for the two-state theory can be evaluated using[7]

$$K = \sigma s^{3n-2} \tag{23}$$

$$K = \theta_h/(1 - \theta_h) \tag{24}$$

The values of s and σ can be expressed in terms of their respective thermodynamic parameters, enthalpy and entropy, using eqs 11 and 13. It is assumed that σ does not depend on temperature ($\Delta H_\sigma = 0$). In analogy with eq 16, the following is derived for intramolecular transitions:

$$\left(\frac{\partial \ln \theta_h}{\partial(1/T)}\right)_n = \left(-\frac{(3n-2)\Delta H_s}{R}\right)\left(\frac{1-\theta_h}{2-\theta_h}\right) \tag{25}$$

When $\theta_h = {}^1/_2$, $K = 1$ and eq 25 becomes

$$\left(\frac{\partial \ln \theta_h}{\partial(1/T)}\right)_n = -\frac{(3n-2)\Delta H_s}{3R} \tag{26}$$

Knowing ΔH_s, ΔS_s and σ and can be evaluated from the chain-length dependence of T_m from the following equation

$$\frac{1}{T_m} = \frac{1}{3n-2}\frac{R \ln \sigma}{\Delta H_s} + \frac{\Delta S_s}{\Delta H_s} \tag{27}$$

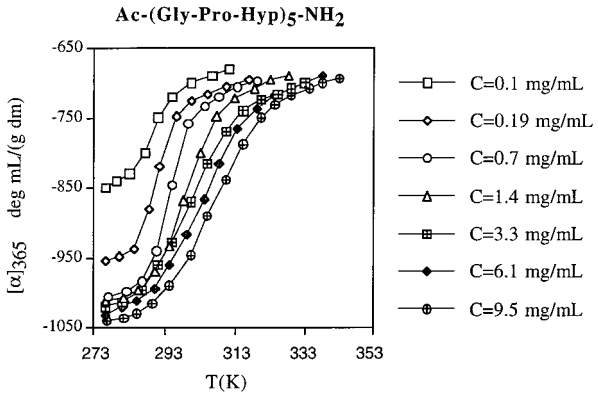


Figure 1. Equilibrium transition curves measured by specific rotation [α] at 365 nm for Ac-(Gly-Pro-Hyp)$_5$-NH$_2$ at different concentrations.

TABLE 1: Concentration Dependence of the Fraction of Helix (θ_h) for Ac-(Gly-Pro-Hyp)$_5$-NH$_2$ at the Corresponding Melting Temperatures

T (K)	ln C (C (mg/mL))	ln θ_h	T (K)	ln C (C (mg/mL))	ln θ_h
291	−2.30 ± 0.01	−1.07 ± 0.01	295	−2.30 ± 0.01	−2.35 ± 0.01
	−1.66 ± 0.01	−0.72 ± 0.01		−1.66 ± 0.01	−1.82 ± 0.01
	−0.36 ± 0.01	−0.23 ± 0.01		−0.36 ± 0.01	−0.77 ± 0.01
	0.34 ± 0.01	−0.18 ± 0.01		0.34 ± 0.01	−0.36 ± 0.01
	1.19 ± 0.01	−0.16 ± 0.01		1.19 ± 0.01	−0.34 ± 0.01
	1.81 ± 0.01	−0.05 ± 0.01		1.81 ± 0.01	−0.18 ± 0.01
	2.25 ± 0.01	−0.04 ± 0.01		2.25 ± 0.01	−0.15 ± 0.01
299	−2.30 ± 0.01	−5.81 ± 0.01	300	−2.30 ± 0.01	−6.10 ± 0.01
	−1.66 ± 0.01	−3.33 ± 0.01		−1.66 ± 0.01	−3.56 ± 0.01
	−0.36 ± 0.01	−1.91 ± 0.01		−0.36 ± 0.01	−2.28 ± 0.01
	0.34 ± 0.01	−0.76 ± 0.01		0.34 ± 0.01	−0.87 ± 0.01
	1.19 ± 0.01	−0.62 ± 0.01		1.19 ± 0.01	−0.71 ± 0.01
	1.81 ± 0.01	−0.35 ± 0.01		1.81 ± 0.01	−0.40 ± 0.01
	2.25 ± 0.01	−0.27 ± 0.01		2.25 ± 0.01	−0.31 ± 0.01
304	−2.30 ± 0.01	−6.71 ± 0.01	307	−2.30 ± 0.01	−7.24 ± 0.01
	−1.66 ± 0.01	−4.12 ± 0.01		−1.66 ± 0.01	−4.64 ± 0.01
	−0.36 ± 0.01	−2.94 ± 0.01		−0.36 ± 0.01	−3.60 ± 0.01
	0.34 ± 0.01	−1.40 ± 0.01		0.34 ± 0.01	−2.03 ± 0.01
	1.19 ± 0.01	−1.12 ± 0.01		1.19 ± 0.01	−1.53 ± 0.01
	1.81 ± 0.01	−0.66 ± 0.01		1.81 ± 0.01	−1.00 ± 0.01
	2.25 ± 0.01	−0.53 ± 0.01		2.25 ± 0.01	−0.72 ± 0.01

Determination of Parameters

Ac-(Gly-Pro-Hyp)$_n$-NH$_2$. *Concentration Dependence of* θ_h. The experimental transition curves reported in Figure 1 for $n = 5$, using specific rotations at 365 nm as a function of temperature for different concentrations, were converted to net fraction of helix (θ_h) as a function of temperature, using eq 6.

Values of the fraction of helix (θ_h) as a function of the concentration are reported as ln C and ln θ_h in Table 1 and Figure 2 at different temperatures. Thus, the logarithmic slope of the association isotherm can be computed at each concentration for each temperature from which the fraction of free single strands $\zeta(T,C)$ can be obtained by using eq 7. Substituting these values in eq 8, $s(T,C)$, the equilibrium constant for the propagation step, is obtained. Having also determined $\sigma(T,C)$ from an equation reported by Schwarz and Poland,[4] the temperature dependence of $s(T,C)$ (eqs 9, 11) leads to $\Delta H_s(C)$.

In order to check that the AON model is applicable, we evaluated $\partial \ln \theta_h/\partial \ln C$ at the temperature where $\theta_h = {}^1/_2$ ($T = T_m$) at each concentration. A comparison between the AON and the one-helical-sequence models from concentration dependence

TABLE 2: Comparison between the AON and the One-Helical-Sequence Models from Concentration Dependence Data Acquired for Ac-(Gly-Pro-Hyp)$_5$-NH$_2$

	C (mg/mL)[a]	T_m (K)	$\partial \ln \theta_h / \partial \ln C$	ζ	θ_h
AON	1.40 ± 0.01	299 ± 1	0.50	0.50	0.50
one-helical-sequence	1.40 ± 0.01	299 ± 1	0.49 ± 0.02	0.49 ± 0.01	0.47 ± 0.01

[a] This concentration was chosen because the best fit of the data ($\ln\theta_h$ vs $\ln C$) was obtained in Figure 2.

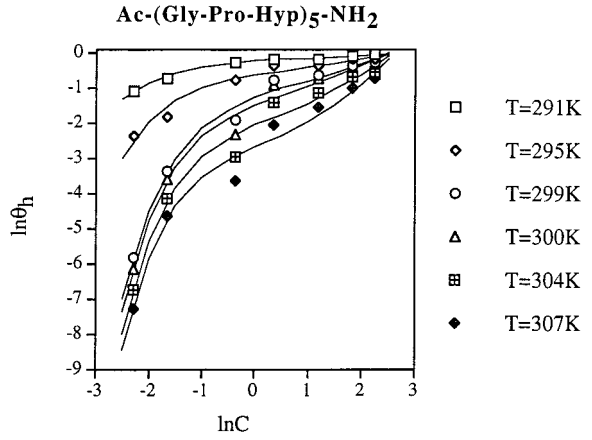


Figure 2. Fraction of triple helix θ_h plotted as a function of concentration for Ac-(Gly-Pro-Hyp)$_5$-NH$_2$ at different temperatures. The symbols represent the experimental values and the lines represent fitted curves obtained by using the nonlinear least-squares method.

TABLE 3: Concentration Dependence of the Fraction of Single Strands (ζ) on the Logarithmic Slope of the Association Isotherm ($\partial \ln \theta_h / \partial \ln C$) and of the Propagation Parameter (s) for Ac-(Gly-Pro-Hyp)$_5$-NH$_2$

T (K)	C (mg/mL)	θ_h	ζ	$\partial \ln \theta_h / \partial \ln C$	s
295 ± 1	0.70 ± 0.01	0.46 ± 0.01	0.35 ± 0.01	0.31 ± 0.02	4.8 ± 0.3
	1.40 ± 0.01	0.69 ± 0.01	0.26 ± 0.01	0.21 ± 0.02	26 ± 7
	3.30 ± 0.01	0.71 ± 0.01	0.25 ± 0.01	0.20 ± 0.02	38 ± 15
300 ± 1	0.70 ± 0.01	0.10 ± 0.01	0.65 ± 0.01	0.76 ± 0.02	0.7 ± 0.1
	1.40 ± 0.01	0.42 ± 0.01	0.52 ± 0.01	0.53 ± 0.02	12 ± 2
299 ± 1	0.70 ± 0.01	0.15 ± 0.01	0.64 ± 0.01	0.74 ± 0.02	1.6 ± 0.1
	1.40 ± 0.01	0.47 ± 0.01	0.49 ± 0.01	0.49 ± 0.02	18 ± 3
304 ± 1	0.70 ± 0.01	0.05 ± 0.01	0.65 ± 0.01	0.78 ± 0.02	~0
	1.40 ± 0.01	0.24 ± 0.01	0.57 ± 0.01	0.61 ± 0.02	2.9 ± 0.3

data acquired at $T = 299$ K is reported in Table 2. All the parameters computed from the one-helical-sequence approximation are in good agreement, within the estimated errors, with those derived from the AON theory.

The concentration dependence of the fraction of single strands $\zeta(T,C)$, of the logarithmic slope of the association isotherm $(\partial \ln \theta_h / \partial \ln C)_{T,n}$ and of the propagation parameter $s(T,C)$ at different temperatures is reported in Table 3.

A comparison between the thermodynamic parameters obtained for Ac-(Gly-Pro-Hyp)$_5$-NH$_2$ and the values previously reported by Schwarz and Poland[4] for (Gly-Pro-Pro)$_n$ sequences is shown in Table 4. The concentration of Ac-(Gly-Pro-Hyp)$_5$-NH$_2$ was such that the molarities of (Gly-Pro-Hyp) and (Gly-Pro-Pro) tripeptide units were similar.

Chain-Length Dependence of θ_h. The experimental transition curves reported in Figure 3, which show specific rotation at 365 nm as a function of temperature for peptides for $n = 3$, 5, 6, 9 at $C = 0.19$ mg/mL, were converted to net fraction of helix (θ_h) as a function of temperature using eq 6. The shorter sequence ($n = 3$) does not form a triple helix at any temperature. The experimental conditions were such that the concentration of (Gly-Pro-Hyp) tripeptide units was held constant, i.e., $nC = C_0$ where C is the total molar concentration of single strands.

The sharpness of the thermal transition, $(\partial \ln \theta / \partial(1/T))_{C,n}$,

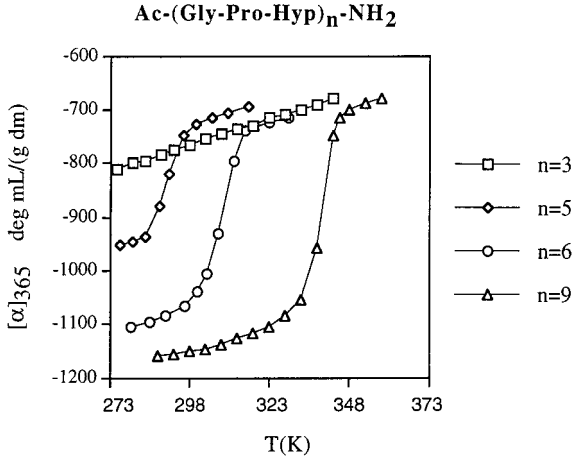


Figure 3. Equilibrium transition curves measured by specific rotation [α] at 365 nm for Ac-(Gly-Pro-Hyp)$_n$-NH$_2$ at 0.19 mg/mL with $n = 3$, 5, 6, 9.

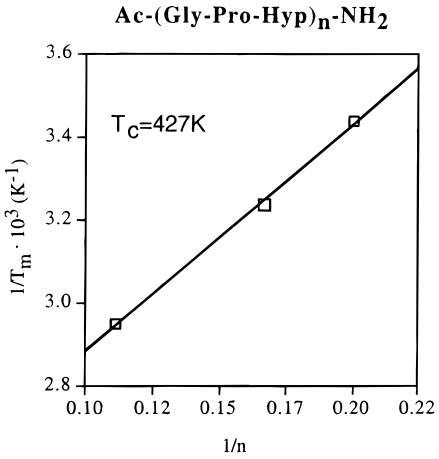


Figure 4. Chain-length dependence of T_m for Ac-(Gly-Pro-Hyp)$_n$-NH$_2$ at 0.19 mg/mL with $n = 5$, 6, 9.

was calculated at the midpoint as a function of n for $C = 0.19$ mg/mL. Using eq 16 evaluated at $\theta_h = 1/2$, i.e., eq 19, ΔH_s was computed as $\Delta H_s = -10 \pm 0.5$ kcal (mol trimeric unit)$^{-1}$ at this concentration from the slope of a plot of $\partial \ln \theta / \partial(1/T)$ vs n.

The transition curves readily provide the melting temperature T_c for the infinite polymer. By plotting the reciprocal of the melting temperature as a function of the reciprocal of the chain length (Figure 4), we obtain an extrapolated value at $n \rightarrow \infty$ ($1/n \rightarrow 0$) of $T_c = 427$ K. From a plot of $1/T_m$ vs $1/n$ at $C = 0.19$ mg/mL (eq 20), the equilibrium constant for the nucleation step $\ln \sigma = -12.5 \pm 1.4$, where σ has the units (mol/L)$^{-2}$, i.e., those of a trimolecular association constant, and $\Delta S_s = -23.4 \pm 1.2$ cal (mol trimeric unit)$^{-1}$ K^{-1} were obtained. The value $\Delta S_\sigma = -24.8 \pm 2.8$ cal (mol trimeric unit)$^{-1}$ K^{-1} was determined using eq 13.

From ΔH_s and ΔS_s, which were determined from the length dependence of θ_h, and with the assumption that these two parameters are independent of temperature, $s(T,C)$ can be computed at each temperature using eq 9. It should be pointed out the ΔH_s and ΔS_s were derived from two sets of measure-

TABLE 4: Comparison between the Thermodynamic Parameters Measured for Ac-(Gly-Pro-Hyp)$_5$-NH$_2$ from Concentration Dependence Data and the Values Measured for (Gly-Pro-Pro)$_n$ Sequences

	nC[a] (C(mM))	ln σ (σ(mol/L)$^{-2}$)	ΔH_σ (kcal/mol)	ΔS_σ (cal/(mol·K))	ΔH_s (kcal/mol)	ΔS_s (cal/(mol·K))
Ac-(Gly-Pro-Hyp)$_5$-NH$_2$	5.0 ± 0.1	−3.9 ± 1.0	0	−7.8 ± 1.4	−5.7 ± 0.3	−13.3 ± 0.8
H-(Gly-Pro-Pro)$_n$-OH	6.5 ± 0.1	−9.3 ± 2.0	0	−18.6 ± 2.0	−7.4[b] ± 0.8	−20.0 ± 2.0

[a] The nC is the total molar concentration of the tripeptide units. [b] Figure 3 of ref 4 shows essentially no dependence of ΔH_s on n.

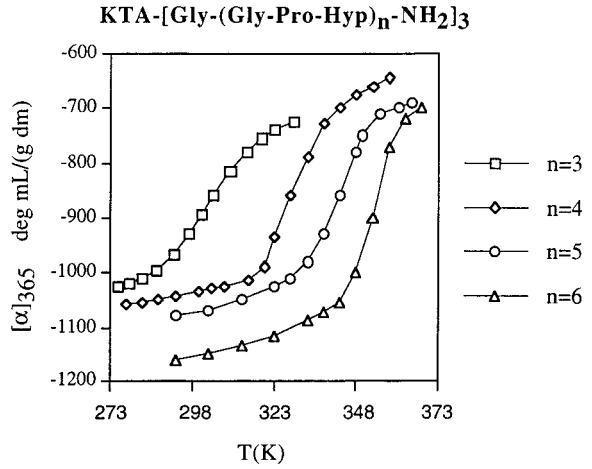


Figure 5. Equilibrium transition curves measured by specific rotation [α] at 365 nm for KTA-[Gly-(Gly-Pro-Hyp)$_n$-NH$_2$]$_3$ with n = 3, 4, 5, 6.

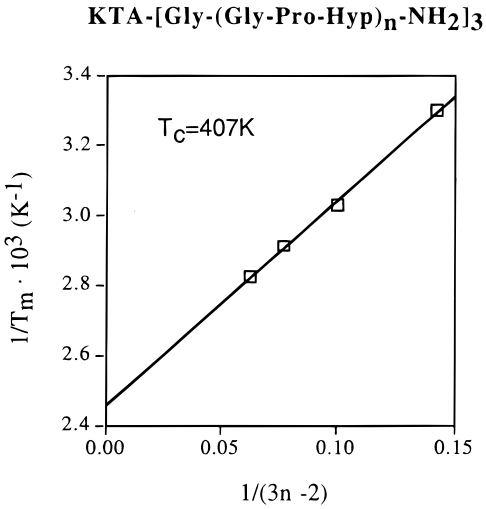


Figure 6. Chain-length dependence of T_m for KTA-[Gly-(Gly-Pro-Hyp)$_n$-NH$_2$]$_3$ with n = 3, 4, 5, 6.

ments: concentration and chain length dependence. Theoretically, the values obtained from both experimental approaches should be the same. However, the values differ slightly because the limited number of available data did not allow for the determination of ΔH_s and ΔS_s for the chain length-dependent analysis.

KTA-[Gly-(Gly-Pro-Hyp)$_n$-NH$_2$]$_3$. *Chain-Length Dependence of θ_h.* The experimental transition curves reported in Figure 5 show specific rotations at 365 nm as a function of temperature for peptides for which n = 3, 4, 5, 6. These values were converted to net fraction of helix (θ_h) as a function of temperature using eq 6.

The sharpness of the thermal transition, $(\partial \ln \theta/\partial(1/T))_n$, was calculated at the midpoint as a function of $3n - 2$. Using eq 26, when $\theta_h = {}^1/_2$, ΔH_s was computed as $\Delta H_s = -5.2 \pm 0.4$ kcal (mol trimeric unit)$^{-1}$.

The melting temperature for the infinite polymer (T_c = 407 K) was determined by plotting the reciprocal of the melting temperature vs $1/(3n - 2)$ (Figure 6). The equilibrium constant for the nucleation step ln $\sigma = -15.4 \pm 1.2$, $\Delta S_\sigma = -30.5 \pm 2.4$ cal (mol trimeric unit)$^{-1}$ K^{-1} and $\Delta S_s = -12.8 \pm 1.0$ cal (mol trimeric unit)$^{-1}$ K^{-1} were obtained by using eq 27 knowing ΔH_s. From ΔH_s and ΔS_s, and with the assumption that these two parameters are independent of temperature, s can be computed at each temperature by using eq 9.

Discussion

A thermodynamic treatment of the triple-helical folding for the N-terminal acetyl free-chain Ac-(Gly-Pro-Hyp)$_n$-NH$_2$ (where n = 3, 5, 6, 9) and for the template-assembled KTA-[Gly-(Gly-Pro-Hyp)$_n$-NH$_2$]$_3$ (where n = 3, 4, 5, 6) is discussed.

For such short chains, the application of the one-helical-sequence approximation indicates that there is very little unwinding from the ends of the triple helix, so that an all-or-none model is adequate to treat this transition. Table 2 shows a comparison between the values obtained by the one-helical-sequence and the all-or-none models which were applied to describe the triple-helix formation for the free-chain Ac-(Gly-Pro-Hyp)$_n$-NH$_2$. Furthermore, sequences where n = 5, 6, 9 were found to be completely triple-helical (θ_h = 1) at each lowest temperature of the experimental transition curves. This result justifies the all-or-none approximation which is even more reasonable in the case of the shorter template-assembled sequences.

For the free-chain sequences, the nucleation constant describes a trimolecular process (intermolecular); hence, σ has the units of (mol/L)$^{-2}$. In contrast, the template-assembled structures exhibit a monomolecular (intramolecular) transition from coil to triple helix which is independent of peptide concentration. Therefore, a comparison between the σ values and the corresponding thermodynamic parameters (ΔH and ΔS) for the triple helix transitions of free-chain and template-assembled structures is meaningless.

The σ values determined for free chains Ac-(Gly-Pro-Hyp)$_n$-NH$_2$ from the concentration dependence and chain-length dependence data are comparable. From the concentration dependence of θ_h for Ac-(Gly-Pro-Hyp)$_5$-NH$_2$, we obtained the most reliable result (best curve fit of the association isotherm) at C = 1.4 mg/mL, while the chain-length dependence was monitored at C = 0.19 mg/mL. Consequently, the $\Delta S_\sigma = -24.8 \pm 2.8$ cal (mol trimeric unit)$^{-1}$ K^{-1} determined from the chain-length dependence of θ_h is more negative (less favorable) than the $\Delta S_\sigma = -7.8 \pm 1.4$ cal (mol trimeric unit)$^{-1}$ K^{-1} determined from the concentration dependence of θ_h because of the lower molar concentration of the trimeric units. In fact, at the concentration of 0.19 mg/mL, based upon the data reported in Figure 2, θ_h is not at the plateau value but depends strongly on concentration.

In both cases, free-chain and template-assembled collagen peptides, the equilibrium constant for the propagation step, s, is unitless since it refers to an intramolecular transition. When s is greater, less than, or equal to unity, triple-helical formation

is favored, unfavored, and of equal probability with respect to unassembled structures, respectively. A comparison of the thermodynamic parameters associated with *s* ensures that the propagation step for the free-chain and template-assembled structures involves the same change in free energy, since the tripeptide unit is the same. The values of ΔH_s and ΔS_s determined from the concentration dependence of θ_h for Ac-(Gly-Pro-Hyp)$_5$-NH_2 ($C = 1.4$ mg/mL, $\Delta H_s = -5.7 \pm 0.3$ kcal (mol trimeric unit)$^{-1}$, $\Delta S_s = -13.3 \pm 0.8$ cal (mol trimeric unit)$^{-1}$ K^{-1}) and from the chain-length dependence of θ_h for KTA-[Gly-(Gly-Pro-Hyp)$_n$-NH_2]$_3$ ($n = 3, 4, 5, 6$, $\Delta H_s = -5.2 \pm 0.4$ kcal (mol trimeric unit)$^{-1}$, $\Delta S_s = -12.8 \pm 1.0$ cal (mol trimeric unit)$^{-1}$ K^{-1}) are similar.

The nucleation for the template-assembled structures can be described in terms of an intramolecular transition as noted above for the propagation. Therefore, the enthalpy and entropy involved in the two steps can be compared. Although covalent bridging of the peptide chains is assumed to favor nucleation, it is expected that σ is less than unity reflecting the fact that the ends of chain have a relatively low probability for triple helicity. The nucleation gives rise to an entropic loss, thereby rendering it more difficult than the growth of the helix by propagation steps; i.e., ΔS_σ is much more negative than ΔS_s.

Based upon the cooperativity of the process, *s* should depend strongly on the chain length for this association process. However, our results do not distinguish between sequences that differ only by the number of tripeptide units because the concentration dependence data of θ_h are available only for the acetyl analogue with five tripeptide repeats and the chain-length dependence data are available only for $n = 5, 6, 9$ for the acetyl free-chain analogues and $n = 3, 4, 5, 6$ for the KTA assembled structures. These data sets are not sufficient to allow a nonlinear fit of the curve obtained by plotting the slope of the thermal transition curve as a function of *n*. In addition, the range of *n* is limited.

The thermodynamic parameters associated with the propagation step for free chains are not very different from those published by Schwarz and Poland[4] for the (Gly-Pro-Pro)$_n$ sequences. The intrinsic lower tendency of the (Gly-Pro-Pro)$_n$ than the (Gly-Pro-Hyp)$_n$ sequences to form triple helix is probably balanced by the increased number of tripeptide units *n* (up to 20) as treated by Schwarz and Poland.

On the contrary, there is a difference in the thermodynamic parameters associated with the nucleation step. Specifically, comparison of the σ values in Table 4 shows that the nucleation step is much more favorable for our acetyl-terminated (Gly-Pro-Hyp)$_n$ sequences than for the free amine-terminated (Gly-Pro-Pro)$_n$ analogues. It has previously been seen in our laboratories[14] that acetylation at the N-termini of the peptide chains greatly enhances triple-helical propensity.

In this paper, we focused on the thermodynamics of triple-helical folding for single-stranded and template-assembled Gly-Pro-Hyp sequences for relatively short peptide lengths. The helical forming tendencies of peptide chains comprising different tripeptide units can be compared rigorously by these methods. Moreover, the influence of the capping mode by N-terminal acetylation and by template assembly of the peptide chains was evaluated. The effect of N-terminal acetylation versus the presence of the terminal amine structures has been shown to be more favorable for triple helical formation. Similarly, template assembly of the collagen peptide chains not only enhances but also stabilizes the triple helix, as demonstrated by their higher melting temperatures and faster triple helix folding rate.[9,10]

Material and Methods

The synthesis of the polypeptides and all the biophysical characterizations have been described elsewhere.[9,10]

The experimental equilibrium transition curves were fitted by using the least-squares method to obtain the thermodynamic parameters (σ, *s*, ΔH_s, $\Delta S_{\sigma,s}$).

All the numerical calculations, plots, curve fitting, and curve interpolation were carried out with the program *Mathematica* (Wolfram Research, Inc.) using an operating system based on UNIX.

Acknowledgment. This project is funded by a grant from the National Science Foundation (DMR-9802329).

References and Notes

(1) Flory, P. J.; Weaver, E. S. *J. Am. Chem. Soc.* **1960**, *82*, 4518.
(2) Harrington W. F.; Rao, N. V. *Biochemistry* **1970**, *9*, 3714.
(3) Gõ, N.; Suezaki, Y. *Biopolymers* **1973**, *12*, 1927.
(4) Schwarz, M. Jr.; Poland, D. *Biopolymers* **1974**, *13*, 687.
(5) Poland, D. *Biopolymers* **1974**, *13*, 1859.
(6) Schwarz, M. Jr.; Poland, D. *Biopolymers* **1974**, *13*, 1873.
(7) Roth, W.; Heidemann, E. *Biopolymers* **1980**, *19*, 1909.
(8) Nemethy, G.; Scheraga, H. A. *Biopolymers* **1989**, *28*, 1573.
(9) Feng, Y.; Melacini, G.; Taulane, J. P.; Goodman, M. *J. Am. Chem. Soc.* **1996**, *118*, 10351.
(10) Melacini, G.; Feng, Y.; Goodman, M. *J. Am. Chem. Soc.* **1996**, *118*, 10359.
(11) Poland, D.; Scheraga, H. A. *Theory of Helix-Coil Transitions in Biopolymers*, Academic Press: New York, 1970.
(12) Gõ, M.; Gõ, N.; Scheraga, H. A. *J. Chem. Phys.* **1971**, *54*, 4489.
(13) Von Dreele, P. H.; Lotan, N.; Ananthanarayanan, V. S.; Andreatta, R. H.; Poland, D.; Scheraga, H. A. *Macromolecules* **1971**, *4*, 408.
(14) Feng, Y. Ph.D. Thesis, University of California at San Diego, 1996.