Characterization of the Nucleation Step and Folding of a Collagen Triple-Helix Peptide†

Yujia Xu, Manjiri Bhate, and Barbara Brodsky*

Department of Biochemistry, UMDNJ−Robert Wood Johnson Medical School, 675 Hoes Lane, Piscataway, New Jersey 08854

Received November 16, 2001; Revised Manuscript Received April 26, 2002

ABSTRACT: Peptide T1-892 is a triple-helical peptide designed to include two distinct domains: a C-terminal $(Gly\text{-}Pro\text{-}Hyp)_4$ sequence, together with an N-terminal 18-residue sequence from the α1(I) chain of type I collagen. Folding experiments of T1-892 using CD spectroscopy were carried out at varying concentrations and temperatures, and fitting of kinetic models to the data was used to obtain information about the folding mechanism and to derive rate constants. Proposed models include a heterogeneous population of monomers with respect to cis−trans isomerization and a third-order folding reaction from competent monomer to the triple helix. Fitting results support a nucleation domain composed of all or most of the $(Gly\text{-}Pro\text{-}Hyp)_4$ sequence, which must be in trans form before the monomer is competent to initiate triple-helix formation. The folding of competent monomer to a triple helix is best described by an all-or-none third-order reaction. The temperature dependence of the third-order rate constant indicates a negative activation energy and provides information about the thermodynamics of the trimerization step. These CD studies complement NMR studies carried out on the same peptide at high concentrations, illustrating how the rate-limiting folding step is affected by changes in concentration. This sequence preference of repeating Gly-Pro-Hyp units for the initiation of triple-helix formation in peptide T1-892 may be related to features in the triple-helix folding of collagens.

The triple helical collagen motif offers a unique system for studying protein folding because it has a linear structure, limited side chain interactions, and no typical hydrophobic core. The collagen triple helix has a uniform backbone conformation consisting of three supercoiled left-handed polyproline II-like chains (*1*, *2*). The three chains are staggered by one residue with respect to each other, and the close packing of the three chains requires that every third residue in each chain be a Gly, producing the characteristic $(Gly\text{-}X\text{-}Y)_n$ repeating sequence. All Gly residues are buried near the central axis, while the side chains of residues in the X and Y positions are largely exposed to solvent. Interchain hydrogen bonding of peptide groups helps stabilize the trimeric nature of the molecule. The extended polyproline II nature of individual chains is favored by the frequent occurrence of proline and hydroxyproline (Hyp)[1] in the X and Y positions (*3*), respectively, and the most common and stabilizing sequence is Gly-Pro-Hyp (*4*, *5*).

The high content of imino acids needed for the stabilization of collagen is a distinctive feature that must be addressed during the folding process. In contrast to amide bonds, imide bonds of Pro have a significant probability of adopting a cis conformation in an unfolded chain. The slow rate of cis−trans isomerization is known to contribute to the slow phase of folding in globular proteins with only a few Pro residues (*6*, *7*). In imino acid-rich collagens, this isomerization is a dominant feature in folding (*8*−*13*).

The folding of the triple helix has been best characterized for types I and III collagens, which are found in characteristic fibrils with an axial 670 Å repeat (*12*−*16*). In vivo folding is facilitated by the synthesis of procollagens containing N- and C-terminal globular propeptides flanking the $(Gly\text{-}X\text{-}Y)_n$ central domain. The self-association of three C-propeptides leads to trimerization and proper registration. It appears that Gly-Pro-Hyp-rich sequences are required at the C-terminus of the type III collagen triple-helix domain for the initiation of triple-helix formation (*15*). The triple-helix conformation is then propagated along the remaining tripeptide units in a zipper-like manner from the C- to the N-terminus, limited by cis−trans isomerization at imino acid residues (*8*, *9*, *12*, *13*).

Peptides with Gly as every third residue and a high content of Pro and Hyp adopt stable triple-helix structures, and their folding can be monitored by spectroscopic techniques, including CD and NMR (*11*, *17*, *18*). Although peptides lack the C-propeptide-mediated mechanism of trimerization, they can self-associate to form triple helices in correct register (*2*, *19*, *20*). The folding mechanisms of triple helical peptides such as $(Pro\text{-}Pro\text{-}Gly)_n$ and α1-CB2, a CNBr fragment from type I collagen, have been previously investigated using the helix−coil transition theory and zipper models successfully applied to α helices and DNA double helices (*21*−*25*). The helix−coil transition theory predicts that the formation of the triple helix of short peptides from three monomer chains includes a nucleation reaction, where three chains initiate

† This work was supported by grants from NIH (GM60048 to B.B.) and the Children's Brittle Bone Foundation (to B.B.).

* To whom correspondence should be addressed. E-mail: brodsky@umdnj.edu. Phone: (732) 235-4048. Fax: (732) 235-4783.

[1] Abbreviations: CD, circular dichroism; Hyp, hydroxyproline; RMS, root-mean-square deviation of the fit; three letter codes have been used to denote common amino acids.

10.1021/bi015952b CCC: $22.00
Published on Web 05/31/2002

triple-helix formation, followed by a propagation process described as stepwise zippering of a triple helix. In contrast, folding studies on cross-linked peptides or peptides at very high concentrations are dominated only by the propagation process (*8*, *9*, *11*, *26*), where the slow cis−trans isomerization step is rate-limiting.

Here, we extend previous folding studies to a collagen-like peptide T1-892 with two distinct domains: a C-terminal $(Gly\text{-}Pro\text{-}Hyp)_4$ which models the C-terminal sequence of type I collagen chains together with an N-terminal 18-residue sequence from the αl(I) chain of type I collagen (residues 892−909). This peptide adopts a rigid triple-helical conformation, with a melting temperature near 26 °C (*27*, *28*). CD folding of T1-892 is monitored over a range of concentrations and temperatures to elucidate features of the folding mechanism. Models for folding are introduced that include cis−trans isomerization before monomers are competent to initiate triple-helix formation, followed by the formation of triple-helical molecules. The dominant features emerging from this analysis relate to the early steps of folding. In particular, residues in most or all of the $(Gly\text{-}Pro\text{-}Hyp)_4$ sequence, but not the whole peptide, must be in the trans form before monomers can start the formation of the triple helix.

MATERIALS AND METHODS

Sample Preparation. Peptide T1-892, Ac-Gly-Pro-Ala-Gly-Pro-Ala-Gly-Pro-Val-Gly-Pro-Ala-Gly-Ala-Arg-Gly-Pro-Ala-$(Gly\text{-}Pro\text{-}Hyp)_4$-Gly-Val-$NH_2$, was synthesized by Synpep Corporation (Dublin, CA) and was purified and characterized as previously described (*28*).

Circular Dichroism. Circular dichroism spectra were recorded on an Aviv model 62DS spectrometer. Cells of path length 0.1 cm were used, and the temperature in the cell was controlled using a Hewlett-Packard Peltier thermoelectric temperature controller. Samples were prepared at varying concentrations, with peptides dried in vacuo over P_2O_5 for at least 48 h prior to weighing. The concentrations were calibrated using the ellipticity at 225 nm based on the study of a standard peptide with the same sequence plus a Tyr residue at the C-terminus (T1-892Y). The concentration of the standard T1-892Y was determined from the absorbance at 280 nm with an extinction coefficient $\epsilon_{280\ nm} = 1179\ M^{-1}\ cm^{-1}$. Other than the absorbance of the Tyr residue, the T1-892Y show the same CD spectral properties in the 210−240 nm range and the same thermal stability as those of the T1-892 samples used for folding.

Folding Experiments. Samples were heated to dissociate the trimers to monomers and cooled to desired temperatures to initiate refolding. About 0.3 mL of sample was denatured in a 70 °C water bath for 15 min and then rapidly quenched in an ice−water bath kept at 0−2 °C. The quenching time to reach the desired temperature was estimated for a given folding temperature by using a blank solution of the same volume with a microthermal probe. For example, for folding at 5 °C, the quenching takes 14 s. These time intervals were used for quenching in the actual experiment. The heating and quenching were carried out in the CD cuvette or in a small glass tube followed by a quick transfer into a CD cuvette that had been equilibrated at the folding temperature in the instrument. The ellipticity at 225 nm was monitored immediately after quenching with a time constant of 2 s and time interval of 10 s (or 30 s under slow folding conditions). The dead time was typically about 20−40 s depending on the refolding temperature. The concentration ranged from 0.12 to 1.57 mM, with the lower end limited by the long refolding time and poor signal. The observed ellipticity is expressed in millidegrees and is converted to mean residue ellipticity $[\theta]$ for fraction folded or molar ellipticity $\theta°$ for kinetic studies.

Fraction Folded and the Half-Time of Folding. The refolding half-time, $t_{1/2}$, was determined as the time for the fraction folded (FF) to reach 50%, with the fraction folded FF defined as

$$FF = ([\theta_{obs}] - [\theta_m])/([\theta_{TH}] - [\theta_m]) \qquad (1)$$

where $[\theta_{obs}]$ is the observed mean residue ellipticity. The $[\theta_{TH}]$ and $[\theta_m]$ are the mean residue ellipticities of the trimer and the monomer, respectively, at a given temperature. The $[\theta_{TH}]$ at a given folding temperature is measured directly before unfolding. There is a significant linear temperature dependence of monomer signal in the high-temperature range of equilibrium-melting curve and this is extrapolated back to the folding temperature to obtain $[\theta_m]$.

Analysis of the Folding Data. The CD traces of folding, corrected by the baseline, were directly fit to the folding schemes by expressing the change in observed ellipiticity over time $\theta(t)$ as

$$\left(\frac{d\theta(t)}{dt}\right) = \theta_{tr}°\left(\frac{d[TH]}{dt}\right) + \theta_M°\left(\sum_i \frac{d\{M_i\}}{dt}\right) \qquad (2)$$

Here, $\theta_{tr}°$ is the molar ellipticity of trimer, in millidegree/mol, estimated from the CD spectrum in conditions (0.37 mM, 5 °C) where the peptide appears to be fully folded and the mean residue ellipticity unchanged by increasing concentration. The $\theta_M°$ is the molar ellipticity of monomer, determined from the denatured baseline of the equilibrium-melting curve. Both were corrected for the temperature dependence according to the baselines obtained from the equilibrium-melting curve. The [TH] is the molar concentration of the trimer. In the folding model, the monomer is separated into populations with varying cis−trans configurations and each population is designated $\{M_i\}$ (see Results). The summation is over all major populations of monomer chains at the equilibrium unfolded state. It was assumed that all of the monomer chains have the same $\theta_M°$ value.

Global Fit of the Data. The microscopic rate constants of a folding scheme are estimated by numerically solving the corresponding differential equations (Appendix) using the software SCIENTIST by MicroMath. The data at each concentration and temperature are first fit individually. The convergence of such fit is relatively fast and a large selection of models can be tested. Because the model selection based on a single fit of data at a particular concentration and temperature can be ambiguous and even misleading, a global fit of groups of data at several different concentrations and a specific temperature is conducted in the second step. The global fit uses a set of equations with the microscopic rate constants as the common parameters, which lowers the degree of freedom for the overall fitting process. This global fitting approach also ensures that the derived model is consistent with the entire experimental data set. The goodness

of fit is judged on both the small root-mean-square (RMS) deviation value of fit and the random residual plots. To avoid the trap of local minimums during convergence, cautions were specifically taken to start the fit with several sets of randomly chosen initial values for the estimation parameters.

Fitting of the Kinetic Data to Scheme 2. To limit the number of simultaneously estimated parameters and to increase the precision of the estimates, the cis−trans isomerization rates of the $\{M_{GP}\}$ population, which have been well-characterized, were treated as known parameters. On the basis of the analysis of Creighton (*7*), the $k_{ct,GP}$ was set equal to the cis−trans rate estimated from oligopeptides (*29*), while the $k_{tc,GP}$ value taken from oligopeptide studies was multiplied by the number of Gly−Pro bonds in a given sequence. The CD folding traces were fit directly to Scheme 2, holding the rate constants $k_{ct,GP}$ and $k_{tc,GP}$ as fixed values. The other microscopic rate constants $k_{ct,slow}$, $k_{tc,slow}$, k_{rf}, and k_{uf} are estimated by numerically solving the corresponding differential equations (Appendix). The maximum number of simultaneously estimated parameters during general fitting is limited to $k_{ct,slow}$, $k_{tc,slow}$, and k_{rf}, because the value of k_{uf} is negligibly small and not included except at 15 °C. Thus, despite the additional monomer subpopulation, the number of fitting parameters are kept to be the same in Scheme 2 and in Scheme 1 (see Results).

Estimation of the Monomer Populations. The T1-892 peptide has two types of residues preceding an imino acid: Gly−Pro bonds and Pro−Hyp bonds. On the basis of the study of oligopeptides, the probability of a Gly−Pro bond to be in cis form is 14% at equilibrium and the probability of Pro−Pro is 6% (*29*). There are no data available regarding the cis/trans equilibrium of Pro−Hyp bonds from oligopeptide models. However, NMR studies of the monomer distribution in T1-892 (*11*) and in collagens (*30*) indicate that the equilibrium population of *cis*-Pro−Hyp bonds is close to the 6% Pro−Pro value, despite recent amide studies suggesting that Hyp will have a larger trans/cis ratio than Pro (*31*−*34*). Thus, the probability of 6% is used as an estimate of *cis*-Pro−Hyp bonds. Assuming isomerization of each imino bond is independent in the monomer state, the population of a specific cis−trans conformation in a given sequence can be calculated.

Temperature Dependence of the Rate Constant. The change of the third-order rate constant of folding at different temperatures was analyzed using the Arrhenius equation predicting a linear dependence of ln *k* versus *1/T*

$$k = A\mathrm{e}^{-E_a/RT} \tag{3}$$

where *k* is the rate constant, E_a the activation energy, *R* the gas constant, *T* the temperature, and *A* the pre-exponential factor.

Transition-state theory relates the rate constant with the free energy $\Delta G^{\ddagger}$ of the quasithermodynamic equilibrium between the ground state and the activated complex (marked with ‡) at the transition state.

$$k = \left(\frac{k_B T\kappa}{h}\right)\mathrm{e}^{-\Delta G^{\ddagger}/RT} \tag{4a}$$

The pre-exponential factor of an elementary reaction is determined by the Boltzmann constant k_B, Planck constant *h*, and the transmission coefficient κ. The effects of temperature on $\Delta G^{\ddagger}$ can be expressed as

$$\Delta G^{\ddagger} = \Delta H^{\ddagger}(T^{\circ}) - T\Delta S^{\ddagger}(T^{\circ}) + \Delta C_p^{\ddagger}\left[(T - T^{\circ}) - T\ln\left(\frac{T}{T^{\circ}}\right)\right] \tag{4b}$$

where $\Delta H^{\ddagger}$, $\Delta S^{\ddagger}$ and $\Delta C_p^{\ddagger}$ are the enthalpy, the entropy, and the heat capacity of the activation, respectively, at the reference temperature T°. Temperature dependence of the folding rate constant are fit to the modified Eyring equation (*35*) by combining eqs 4a and 4b

$$\ln k = \ln\left(\frac{k_B T\kappa}{h}\right) - \left\{\Delta H^{\ddagger}(T^{\circ}) - T\Delta S^{\ddagger}(T^{\circ}) + \Delta C_p^{\ddagger}\left[(T - T^{\circ}) - T\ln\left(\frac{T}{T^{\circ}}\right)\right]\right\}/RT \tag{5}$$

For complex reactions such as protein folding, the value of κ is not known and a value of 1, as found for elementary reactions, is often assumed (*35*−*37*). As a consequence, only relative values of $\Delta S^{\ddagger}$, and thus also of $\Delta G^{\ddagger}$, can be calculated from direct treatment of the kinetic data with eq 5.

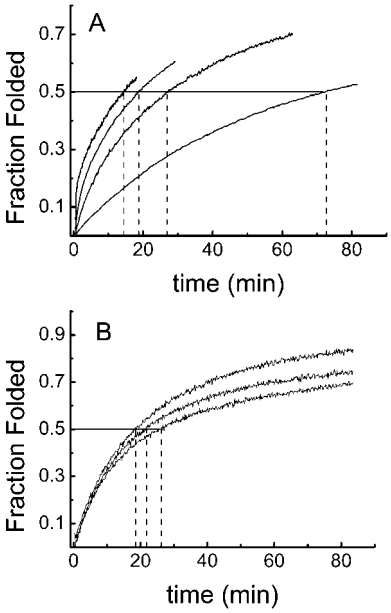

FIGURE 1: (A) CD refolding curves of peptide T1-892 at varying concentrations (left to right) 1.57, 0.80, 0.33, and 0.12 mM, at 0 °C. (B) Refolding curves at varying temperatures (left to right) 5, 10, and 15 °C (fixed concentration 0.37 mM). The vertical dashed lines mark the time at which the fraction folded reaches 50% ($t_{1/2}$).

RESULTS

CD Folding of Peptide T1-892 at Varying Concentration and Temperature. Refolding experiments of T1-892 were monitored using CD spectroscopy, at concentrations ranging from 0.12 to 1.57 mM (Figure 1A). The observed folding rate is strongly concentration-dependent, decreasing with the decrease in concentration. This is described by the empirical parameter $t_{1/2}$, the half-time of folding, which increases at lower concentrations, indicating a longer folding time. Folding experiments were carried out at different temperatures below the melting temperature (Figure 1B). The folding

Scheme 1: Kinetic Model for T1-892[a]

$$3\{M_{cis}\} \underset{k_{tc}}{\overset{k_{ct}}{\rightleftharpoons}} 3\{M_{tr}\} \underset{k_{uf}}{\overset{k_{rf}}{\rightleftharpoons}} TH$$

[a] In this scheme, $\{M_{cis}\}$ is the collection of monomers containing one or more *cis*-imino bonds in the initiation domain; $\{M_{tr}\}$ is the collection of monomers with the trans bonds capable of initiating triple-helix formation (i.e., the competent monomer); and TH is the fully folded triple helix. k_{ct} is the apparent rate constant monitoring the cis−trans isomerization of one or more *cis*-imino bonds of $\{M_{cis}\}$ needed to acquire a conformation competent for folding; k_{tc} is the trans−cis rate constant; k_{rf} is the third-order folding rate constant; and k_{uf} is the first-order rate constant of unfolding.

is faster at lower temperatures. For example, the $t_{1/2}$ at 5 °C is shorter than the value at 15 °C. Thus, non-Arrenius behavior of the folding rate was observed for T1-892.

The kinetics of peptide T1-892 at different concentrations and temperatures is not consistent with a simple first-, second-, or third-order folding reaction, indicating the presence of more than one kinetic process.

Kinetic Models Including cis−trans Isomerization of Monomer. To fit the kinetic data for peptide T1-892, a model was constructed that includes cis−trans isomerization of the monomer, as well as third-order folding to trimers in an all-or-none process (Scheme 1). In monomeric globular proteins with only a few Pro residues, the slow rate of cis−trans isomerization is known to contribute to the slow phase of folding (*6*, *7*). In imino acid-rich triple-helical peptides, this isomerization could be important at the monomer level, as well as for the propagation in trimers. Given the nine Gly−Pro and four Pro−Hyp bonds in peptide T1-892, ∼80% of monomers will contain at least one cis bond, which must be taken into account in folding models.

In Scheme 1, the folding starts from a heterogeneous monomer population with regards to the cis−trans configuration in peptide T1-892. Because of the presence of cis bonds, only a subpopulation of the monomer is expected to be in a form competent to initiate trimer formation and is designated as $\{M_{tr}\}$. The remaining subpopulation which requires one or more cis−trans isomerization events prior to becoming folding competent is designated $\{M_{cis}\}$. The ratio $\{M_{tr}\}/\{M_{cis}\}$ is used to describe the starting population of competent monomers. On the basis of the third-order two-state folding mechanism reported for other triple-helical peptides (*17*, *23*, *25*), an all-or-none third-order reaction is assumed from the competent monomer $\{M_{tr}\}$ to the complete triple helix, with no significant accumulation of kinetic intermediates.

The folding data of T1-892 at 5 °C for three concentrations are fit to Scheme 1. Given an initial ratio for $\{M_{tr}\}/\{M_{cis}\}$, values for k_{rf}, k_{ct}, and k_{tc} are obtained by numerically solving the differential equation using SCIENTST (Table 1), while k_{uf} was found to be negligible. For example, by specifying the ratio of $\{M_{tr}\}/\{M_{cis}\} = 40{:}60$ as the initial condition (see Material and Methods), the data are fit to a model assuming that 40% of the monomer chains have the correct cis−trans configuration which can directly nucleate and fold into the native triple helix, while 60% must undergo one or more cis−trans isomerizations prior to folding. The RMS of both individual (data not shown) and global fits is used to select the best ratio of $\{M_{tr}\}/\{M_{cis}\}$ (Table 1), and the corresponding microscopic rate constants were obtained by fitting. Residual plots are also used to judge the goodness of fit (data not shown).

Table 1: Global Fit of Data at 5 °C to Scheme 1 with Different $\{M_{tr}\}/\{M_{cis}\}$ Ratios

$\{M_{tr}\}/\{M_{cis}\}$	RMS[a]	$k_{rf} \times 10^{-5}$ ($s^{-1}\,M^{-2}$)	$k_{ct,slow} \times 10^{3}$ (s^{-1})	$k_{tc,slow} \times 10^{3}$ (s^{-1})
20:80	1.56	1.40 ± 0.14	1.66 ± 0.04	1.29 ± 0.07
30:70	1.68	1.11 ± 0.07	1.07 ± 0.02	0.52 ± 0.03
40:60	0.88	1.30 ± 0.04	1.21 ± 0.02	0.89 ± 0.02
50:50	0.86	0.58 ± 0.10	1.06 ± 0.02	0.34 ± 0.03
60:40	1.67	0.54 ± 0.02	0.84 ± 0.03	0.34 ± 0.02
80:20	4.38	0.93 ± 0.13	0.96 ± 0.24	0.98 ± 0.36

[a] The RMS of global fit of data with concentration ranging from 0.33 to 1.57 mM. For samples at 0.8 and 1.57 mM, the CD folding data were taken with a 30 s time average for improved signal-to-noise ratio.

Scheme 2: Branched Kinetic Model for the Folding of T1-892[a]

$$3\{M_{GP}\} \underset{4\,k_{tc,GP}}{\overset{k_{ct,GP}}{\rightleftharpoons}} 3\{M_{tr}\} \underset{k_{uf}}{\overset{k_{rf}}{\rightleftharpoons}} TH$$

$$3\{M_{tr}\} \underset{k_{tc,slow}}{\overset{k_{ct,slow}}{\rightleftharpoons}} 3\{M_{slow}\}$$

[a] $\{M_{GP}\}$ is the collection of monomers containing one Gly−Pro *cis*-imino bond in the nucleation domain; $\{M_{slow}\}$ is the collection of monomers containing one or more *cis*-Pro−Hyp or multiple *cis*-Gly−Pro in the nucleation domain; $\{M_{tr}\}$ is the competent monomer; and TH is the fully folded triple helix. $k_{ct,GP}$ and $k_{tc,GP}$ are the cis−trans isomerization rates of the subpopulation of monomer $\{M_{GP}\}$ and are held constant during fitting (see Material and Methods), while $k_{ct,slow}$ and $k_{tc,slow}$ are the cis−trans isomerization rates of the subpopulation $\{M_{slow}\}$. k_{rf} is the third-order folding rate constant, and k_{uf} is the first-order rate constant of unfolding.

As shown in Table 1, the $\{M_{tr}\}/\{M_{cis}\}$ ratios 40:60 and 50:50 give the smallest global RMS values. These results suggest that, at 5 °C, 40−50% of the monomer population in the equilibrium unfolded state is competent for nucleation and folding. These results hold for fits of folding at other temperatures as well (data not shown). Calculations indicate that 20% percent of the population of T1-892 molecules would have all-trans bonds, yet the ratio of 20:80 does not fit the data. The higher value of the competent monomer population obtained from fitting indicates that nucleation requires only a critical subdomain of T1-892 to contain all-trans bonds.

Taking advantage of what is known about the cis/trans populations and isomerization rate constants in the monomer chains, it is possible to further refine the proposed kinetic model (Scheme 2). The Gly−Pro (∼14%) and Pro−Pro (∼6%) populations obtained from small oligopeptides (see Material and Methods) are in good agreement with values seen for Gly−Pro and Pro−Hyp in unfolded collagen (*30*) and monomeric T1-892 (*11*). Thus, the percent cis residues in any defined sequence portion of unfolded T1-892 chains can be calculated. The rates of cis−trans isomerization have also been well-determined for model peptides (*29*). At 4 °C, k_{ct} is estimated to be 29×10^{-4} s^{-1} for Gly−Pro and much

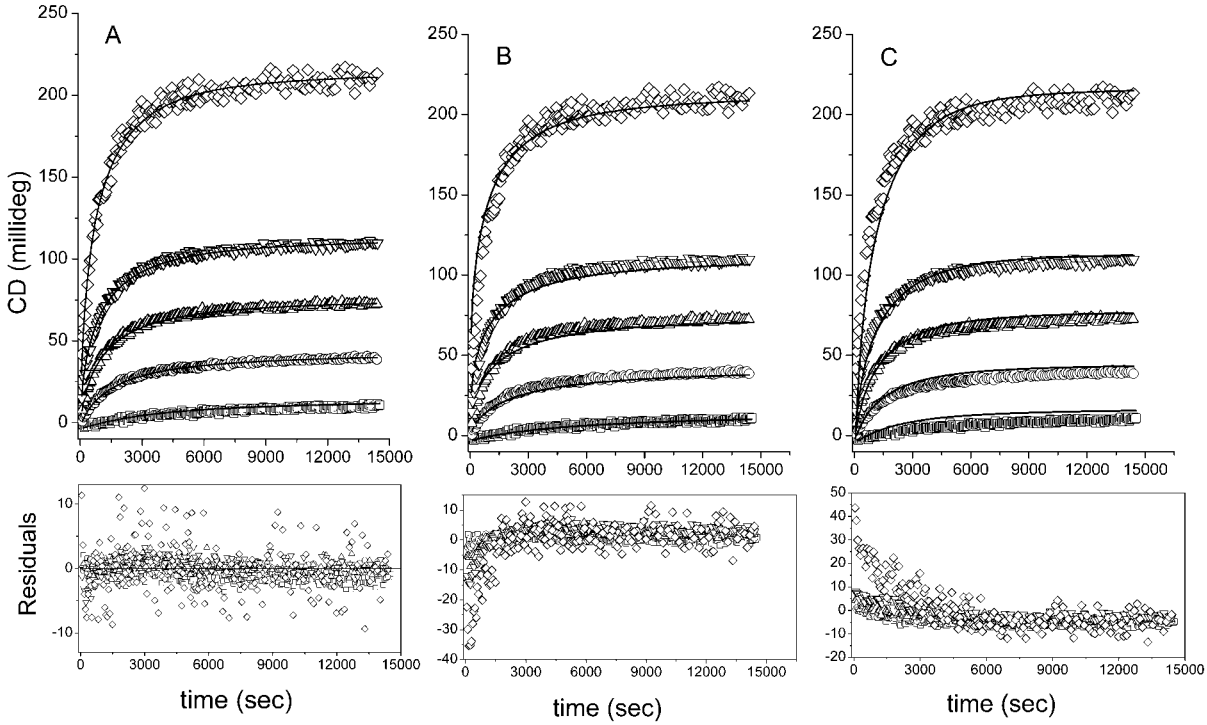

FIGURE 2: Global fitting of folding data at different concentrations to models for the size of the nucleation domain (solid curves, upper panels) and the corresponding residual plots (lower panels). The concentrations are (bottom to top) 0.13 mM (□), 0.33 mM (○), 0.54 mM (△), 0.8 mM (▽), and 1.57 mM(◇). (A) Model assuming a requirement for all-trans peptide bonds in (Gly-Pro-Hyp)$_4$ domain. (B) Model assuming a requirement for all-trans in a consecutive (Gly-Pro-Hyp)$_3$ region. (C) Model assuming a requirement for trans peptide bonds in the entire T1-892 peptide. The fitting results are given in Table 2. The noisier CD signal of 1.57 mM sample is due to the increased Dynode voltage at this concentration.

Table 2: Global Fit of the CD Folding Data at 5 °C to Scheme 2 with Different Assumptions about the Nucleation Domain

nucleation domain[a]	$\{M_{tr}\}/\{M_{GP}\}/\{M_{slow}\}$[b]	RMS[c]	$k_{rf} \times 10^{-5}$ (s^{-1} M^{-2})	$k_{ct,slow} \times 10^{3}$ (s^{-1})	$k_{tc,slow} \times 10^{3}$ (s^{-1})
$(GPO)_4$	43:28:29	2.50	1.37 ± 0.05	0.65 ± 0.01	0.10 ± 0.02
$(GPO)_3$	63:19:18	4.64	0.68 ± 0.04	0.67 ± 0.04	0.01 ± 0.03
whole peptide	20:29:51	6.08	6.17 ± 0.91	0.60 ± 0.02	0.10 ± 0.06

[a] The subdomain of T1-892 monomer chains that need to be in the all-trans conformation for the folding to initiate. Hydroxyproline is denoted by O. [b] Monomer populations calculated for each nucleation model as described in Materials and Methods. [c] The RMS of global fit of data ranging from 0.12 to 1.57 mM. The RMS values are greater than those shown in Table 1 because folding data at all concentrations shown here were collected with a time constant of 2 s.

slower for Pro−Hyp, 6×10^{-4} s^{-1}. On the basis of these findings, the $\{M_{cis}\}$ population is separated into two subspecies in Scheme 2. One species contains monomers with only one *cis*-Gly−Pro imide bond, $\{M_{GP}\}$, and is the dominant species (∼50% of $\{M_{cis}\}$). The rest of $\{M_{cis}\}$ includes peptides with one *cis*-Pro−Hyp bond or multiple cis bonds. Because of the slower isomerization rate of Pro−Hyp and the multiple isomerization events required for peptides containing more than one cis bond, the second species is designated $\{M_{slow}\}$. In Scheme 2, the isomerization rates for Gly−Pro in both directions are taken as constants, with values based on previous oligopeptide studies (*29*). Given an assumed nucleation domain and the monomer distribution of $\{M_{tr}\}/\{M_{GP}\}/\{M_{slow}\}$ calculated for that domain, the observed folding data is fit to three simultaneously estimated parameters, the isomerization rates for the slow species, $k_{ct,slow}$, $k_{tc,slow}$, and the folding rate k_{rf}. Again, the RMS and residual plots are used to evaluate the goodness of fit.

Fitting of the folding data to Scheme 2 at 5 °C is demonstrated in Figure 2 and Table 2 for three different assumptions about the portion of the T1-892 sequence which must be in all-trans form prior to folding. Models assuming a need for all-trans in the (Gly-Pro-Hyp)$_4$ sequence (Figure 2A), in the (Gly-Pro-Hyp)$_3$ sequence (Figure 2B), in the complete peptide (Figure 2C) or in the N-terminal sequence (Gly-Pro-Ala-Gly-Pro-Ala-Gly-Pro-Val-Gly-Pro-Ala) (data not shown) are examined. The model with (Gly-Pro-Hyp)$_4$ as the nucleation domain best fits the folding data for both individual concentrations and global fitting (see Materials and Methods), as monitored by RMS values (Table 2) and residual plots (Figure 2). The initial monomer distribution in the (Gly-Pro-Hyp)$_4$ domain is 43% trans and 57% cis (divided into 28% $\{M_{GP}\}$ and 29% $\{M_{slow}\}$), a value close to the 40−50% competent monomer that best fits Scheme 1. The folding at other temperatures also indicates that (Gly-Pro-Hyp)$_4$ is the best-fitting model (data not shown).

Microscopic cis−trans Isomerization Rate Constants. Because of the excellent fit of Scheme 2 and the more consistent values obtained for rate constants at varying concentrations and temperatures, the rate constants from this scheme were used for further analysis. Fitting the data at different temperatures indicates that the cis−trans isomer-

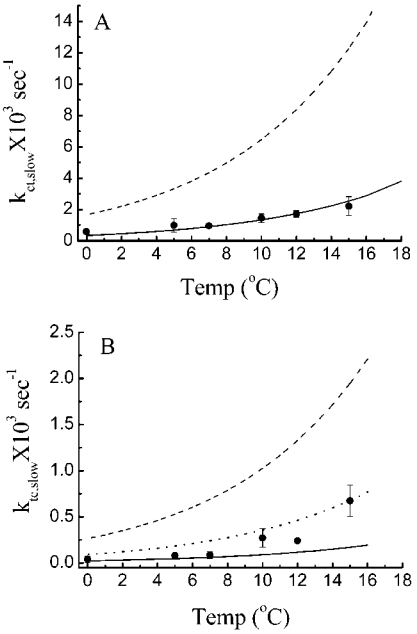

FIGURE 3: Temperature dependence of cis−trans isomerization rate constants for the monomer population. Values are shown from Scheme 2 ($k_{ct,slow}$ A and $k_{tc,slow}$ B) with the best-fitting (Gly-Pro-Hyp)$_4$ model (●). At each specified temperature, the average value of 2−7 samples at concentrations ranging from 0.12 to 1.58 mM are shown, together with error bars. The cis−trans isomerization rates determined for pentapeptides with one Gly−Pro or one Pro−Pro imide bond are presented by the dashed and solid lines, respectively (*29*). The trans−cis rate from all-trans to one cis conformation for a sequence with four Pro-Hyp bonds is expected to be 4 times larger than a single bond (*7*), as indicated by the dotted line in B.

ization rate constants ($k_{ct,slow}$ and $k_{tc,slow}$) are temperature-dependent (Figure 3). The estimated values increase at higher temperature (Figure 3, parts A and B), consistent with a cis−trans isomerization dominated reaction which shows a positive activation energy (∼84 kcal/mol) (*29*, *38*, *39*). In Scheme 2, where the monomer population not competent to fold is divided into faster and slower subpopulations, the cis−trans isomerization rate constants being estimated relate only to the slower subpopulation, which is dominated by species with one *cis*-Pro−Hyp bond. The $k_{ct,slow}$ value for the best-fitting (Gly-Pro-Hyp)$_4$ model is in good agreement with the rate of Pro−Pro bonds in oligopeptides (*29*) over the range of temperature from 0 to 15 °C (Figure 3A,B; Table 2). The trans−cis isomerization step is complicated by the need to take into account multiple ways of introducing one cis bond into the competent monomer. The $k_{tc,slow}$ values estimated for the best-fitting (Gly-Pro-Hyp)$_4$ nucleation model fall in the expected range of one to four Pro−Pro bonds over the temperature range studied. Thus, the rate constants obtained through fitting are consistent with those expected for a separate slow group and supports the validity of the branch included in Scheme 2.

Third-Order Folding Reaction. Following the cis−trans isomerization steps, the competent monomer $\{M_{tr}\}$ forms the triple helix, and the rate constants k_{rf} and k_{uf} which describe this two-state third-order reaction are obtained using Scheme 2. The value of k_{rf} at 5 °C is estimated to be $1.3-1.4 \times 10^5$ M^{-2} s^{-1} for Scheme 2 with the (Gly-Pro-Hyp)$_4$ model (Table 2). The value of k_{uf} is negligible under the current conditions, except at 15 °C where a small value of $k_{uf} \sim 3 \times 10^{-5}$ s^{-1}

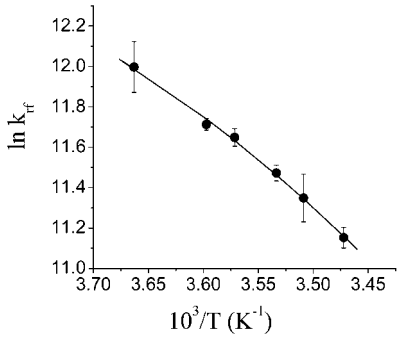

FIGURE 4: Arrhenius plot of k_{rf} values estimated by fitting to Scheme 2 with the (Gly-Pro-Hyp)$_4$ model (●). Each value at a specified temperature is the average of 2−7 samples at concentrations ranging from 0.12 to 1.58 mM. The solid line represents the fit of the data from Scheme 2 to a modified Eyring equation (eq 5).

is determined by fitting. The k_{rf} values, reflecting the third-order folding reaction, decrease with the increase of temperature for the best-fitting model (Figure 4), in contrast to the increasing rate of cis−trans isomerization at higher temperatures (Figure 3). This decrease in third-order folding rate overcomes the increase in the cis−trans reaction at high temperature, resulting in a slowing down of overall folding at elevated temperatures (Figure 1B).

Models which separate the cis−trans isomerization step from the third-order folding step, in principle, allow for the independent analysis of each reaction. The best-fitting model, Scheme 2 with a (Gly-Pro-Hyp)$_4$ initiation domain, gives a temperature dependence of the third-order rate constant (Figure 4) which provides a basis for an analysis of the thermodynamics of the folding reaction of the triple helix. By linear treatment of the Arrhenius plot of k_{rf} using eq 3, the estimated value of the activation energy ΔE is −8.70 kcal/mol. This negative activation energy is the basis of the decreased folding rate at higher temperature.

The best fit of the Arrhenius plot in Figure 4 indicates a slight negative curvature at low temperature, a feature observed for many globular proteins and α-helix coiled coils. In the literature, such curvature is attributed to small changes in the heat capacity $\Delta C_p^{\ddagger}$ from the unfolded state to the transition state and is treated using transition-state theory to derive the activation parameters (*35*−*37*, *40*−*43*). Application of the Eyring equation (eq 5) to the temperature dependence of the k_{rf} rate constants obtained from fits to the best model gives the following set of activation parameters: $\Delta H^{\ddagger} = -8.43$ kcal/mol, $\Delta S^{\ddagger} = -65.6$ cal/mol, and $\Delta C_p^{\ddagger} = -0.63$ kcal/mol. This raises the possibility that the energetically uphill step encountered in triple-helix formation is entropic in nature, because the parameters suggest negative unfavorable entropy, together with negative favorable enthalpy at all temperatures.

DISCUSSION

Folding of peptides has provided insight on the formation of secondary structure and provides a fruitful approach for understanding the complex processes leading to native protein structure. Studies on peptides with the collagen-like triple-helical conformation complement previous work, introducing new features related to oligomerization and an unusually high imino acid content. Folding rates of these trimeric peptides can be manipulated by changing concentra-

tion, to highlight different rate-limiting steps in the folding mechanism. Various folding models are tested through analysis of CD folding data at different concentrations and temperatures on a triple-helical peptide, T1-892.

The starting monomer population of T1-892, a 30-residue peptide with 13 X-imino acid bonds, is expected to be heterogeneous with respect to cis−trans configuration, and this has been directly confirmed by NMR studies (*11*). Incorporation of this initial heterogeneity in a folding model leads to the conclusion that only a subdomain of these imino acid bonds need to be in nativelike trans form prior to initiation of folding. This is demonstrated for the fitting of the CD data to a simple scheme (Scheme 1), which indicates that 40−50% of the starting monomer population has the required trans bonds to be competent. Extension to a scheme (Scheme 2) taking into account the hetergeneous kinetic and equilibrium behavior of Gly−Pro and Pro−Hyp bonds results in a best-fitting model where the C-terminus $(Gly\text{-}Pro\text{-}Hyp)_4$ subdomain has all-trans residues. About 42% of the initial monomer population will have all-trans conformation in the $(Gly\text{-}Pro\text{-}Hyp)_4$ region, giving close agreement with the 40−50% monomer derived from Scheme 1. The observation that the cis−trans isomerization rates at all temperatures for the best models in Scheme 1 are close to those expected for Pro−Pro and slower than expected for Gly−Pro (Table 1) is consistent with the slow Pro−Hyp isomerization being the rate-limiting step prior to initiation of folding. Taken together, these results indicate that all or nearly all of $(Gly\text{-}Pro\text{-}Hyp)_4$ domain in peptide T1-892 acts as the nucleation site, a subdomain where all residues must be trans and where folding starts. This conclusion is supported by host−guest peptides studies which show that the sequence Gly-Pro-Hyp has a propensity for the triple-helix conformation that is higher than the other Gly-X-Y sequences in T1-892 (Gly-Pro-Ala, Gly-Pro-Val, and Gly-Ala-Arg) (*5*). The initiation of folding at this $(Gly\text{-}Pro\text{-}Hyp)_4$ site is also directly demonstrated by NMR data (*11*). It is likely the $(Gly\text{-}Pro\text{-}Hyp)_4$ domain in peptide T1-892 functions as a highly efficient mechanism for nucleation because of the favorable entropy of adjacent Pro−Hyp residues, which are rigid, promote extended chains, and have φ,ψ angles close to those in native collagen (*3*, *44*).

A third-order rate constant describing the folding from the competent monomer to the triple-helix reaction can be derived from Scheme 2. While third-order rate constants have been previously reported for the folding of triple-helical peptides, third-order kinetics was limited to very slow folding peptides (*23*), at low concentrations (*17*) or to the initial stages (*24*, *25*). Here, by including an initial isomerization step prior to trimerization, previous work can be extended to provide a global fit to folding curves of T1-892 over a range of concentrations and temperatures and over an extended time course. The observation that a two-state all-or-none third-order reaction is sufficient to describe the folding from competent monomer to triple-helix formation suggests little accumulation of intermediates, such as partially folded dimers, even though they may be important in the kinetic process. Inclusion of additional steps between the competent monomer and the triple helix did not produce an improved fit. The lack of populated intermediates is also supported by an independent attempt to fit the data by multiexponential functions, showing that functions with two exponential terms fit significantly better than those involving more kinetic steps (data not shown).

Assessment of a range of models at different concentrations and temperatures, as judged by RMS values and random residual plots, establishes the robustness of the kinetic modeling approach and forms a basis for further analysis of the temperature dependence of the rate constants. The rates and activation energy of cis−trans isomerization of monomers leading to the competent monomer state were found to be in good agreement with values found for oligopeptides (*29*), as well as propagation-limited folding of trimeric cross-linked fragments of a collagen triple helix (*8*, *9*). The third-order folding reaction shows a strong temperature-dependence with a negative activation energy, a feature commonly seen in the folding of globular proteins and α-helix coiled coils (*35*−*37*, *40*−*42*). The negative activation energy could be attributed to rapid pre-equilibration between two or three chains before the first rate-limiting step (*45*) in the nucleation process of this peptide. Application of transition-state theory to this temperature dependence suggests an entropic barrier, in contrast with the enthalpic barrier seen in globular proteins (*35*, *36*, *46*, *47*). Recent reports on α-helix coiled-coil peptides suggest that folding mechanisms for this supercoiled motif with hydrophobic interfaces can involve an enthalpic barrier, as seen for a 62-residue C62GCN4 (*40*), or an entropic barrier, as seen for a 32-residue GCN4 variant (*37*).

The CD folding studies reported here complement earlier NMR studies of peptide T1-892 containing ^{15}N-enriched residues at specific sites along the chain (*11*). Real-time NMR was used to monitor the folding of specific residues in T1-892 at high concentrations (10 mM), and each residue showed a fast phase lost in the dead time and a second slower phase dominated by cis−trans isomerization events. A Gly residue in the C-terminal $(Gly\text{-}Pro\text{-}Hyp)_4$ domain folds before residues in the central and N-terminal regions, as judged by its earlier monomer decay in the fast phase. The slow phase seen for all labeled residues is best described by a first-order reaction with rates similar to cis−trans isomerization, which is limiting the C- to N-terminal propagation of the triple helix. Because of the lower concentrations used, the current CD kinetic study can monitor early folding events not available to NMR. In particular, CD data can be analyzed to indicate the nature of the competent monomer, with respect to cis−trans isomerization, and to allow for the determination of the rate-limiting third-order rate constant. Even though propagation and additional cis−trans isomerization events associated with propagation must occur before the triple-helix formation is complete, the third-order nucleation reaction is the dominant energy barrier in CD folding studies as a result of the low concentrations used. The shifting of the folding landscape with changing peptide concentration illustrates how the rate-limiting folding step varies with the different experimental conditions of CD and NMR studies (Figure 5). Cross-linking of triple-helical peptides effectively increases the local concentration and promotes nucleation, again resulting in a propagation-limited reaction (*8*, *9*, *26*). The effect of cis−trans isomerization of a given imino acid bond on the kinetics of folding depends on its context and the rate-limiting conditions. For example, the *cis*-imide bonds in regions N-terminal to $(Gly\text{-}Pro\text{-}Hyp)_4$ in T1-892 are not involved in the rate-limiting nucleation step and therefore would have little effect on the third-order folding rate.

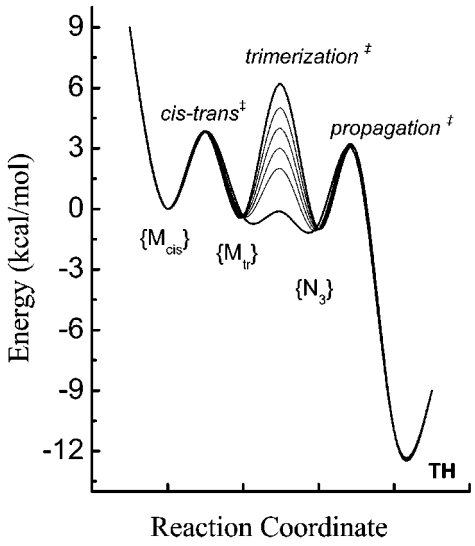

FIGURE 5: Activation energy diagram of the folding of T1-892. In this schematic diagram, monomers with cis bonds $\{M_{cis}\}$ are separated from competent monomer $\{M_{tr}\}$ by a barrier representing cis−trans isomerization within monomers. The association of three chains of competent monomer is characterized by the trimerization‡ barrier. The energy levels of trimerization‡ are shown for folding at concentrations of (top to bottom) 0.14, 0.37 (the common concentration used for CD), 0.95, 1.8, 2.6, and 10 mM (the NMR concentration), using the third-order rate constants k_{rf} estimated from Scheme 2. Because of the presence of an additional 1.2 *cis*-Gly−Pro bonds on average in the sequence N-terminal to the (Gly-Pro-Hyp)$_4$ domain, further cis−trans isomerization reactions with propagation‡ barrier are necessary during the C- to N-terminal propagation process before reaching the final triple-helix state (TH). Here, $\{N_3\}$ represents a postulated obligatory intermediate state that may populate transiently to a significant level when the relative barrier of propagation‡ is significantly higher than that of trimerization‡, as is the case at the high concentrations of NMR studies or for studies of cross-linked peptides. The barriers cis−trans‡ and propagation‡ were drawn corresponding to the activation energy of cis−trans isomerization of a Gly−Pro bond. Similar diagrams with a slightly higher energy cis−trans‡ barrier could be made with Pro−Hyp bonds or multiple cis bonds (not shown). The position of TH relative to $\{M_{cis}\}$ is based on equilibrium temperature melt experiment of T1-892 (*28*).

There is evidence that initiation of collagen triple-helix formation in fibril-forming collagens and in FACIT collagens are cooperative processes involving specific terminal Gly-X-Y sequences as well as adjacent noncollagenous domains (*14*, *15*, *48*−*52*). Studies on peptides, such as those presented here, clearly indicate that collagen model peptides with Gly-X-Y repeating sequences can initiate triple-helix formation in the absence of noncollagenous sequences and that peptides provide a system to characterize the features of nucleation within the triple-helix domain. In peptide T1-892, all or most of the (Gly-Pro-Hyp)$_4$ region acts as a highly efficient nucleation site, but this domain is not obligatory, because triple-helix formation is seen in peptides and collagens lacking this sequence (*14*, *15*, *23*, *25*, *52*). In mutant collagens, where triple-helix propagation is disrupted by a Gly substitution, it has been suggested that renucleation within the (Gly-X-Y)$_n$ domain is required to complete the triple-helical collagen molecule and that sequences N-terminal to the mutation may be critical in the facility of the renucleation step (*27*, *53*). The molecular interactions defined and quantitated in peptide folding studies may clarify features in normal collagen folding as well as the renucleation process in mutant collagens.

Appendix: Model Equations Used for SCIENTIST Fitting of Schemes 1 and 2

```
// fitting equation for Scheme 1 at one particular concentration
IndVars: t
DepVars: CD
Parameters: krf, kct, ktc
// krf for refolding; kuf for unfolding;
   ktc for trans → cis and kct the reverse
// implicit variables: (concentrations) TH, trimer;
   Mt, all-trans monomer; and Mc, monomer with cis bonds
// operator′ = d/dt

C0 = Mt + Mc + 3*TH
TH′ = (1/3)*krf*Mt∧3
Mt′ = kct*Mc − krf*Mt∧3 − ktc*Mt
Mc′ = ktc*Mt − kct*Mc
CD′ = e3*TH′ + e1*Mt′ + e1*Mc′ [a]

// initial conditions, with {Mtr}/{Mcis} ratio of 40:60 at 0.37 mM
t = 0.0
C0 = 0.000 37 //in Molar
Mt = 0.40*C0
Mc = 0.60*C0
TH = 0
CD = e1*C0

// from MicroMath Scientist Model File,
   fitting equation for Scheme 2
// branch model for the fitting of folding data, June 2001
// U0-all-trans, UG/UP-1 cis-GP/PP
IndVars: t
DepVars: CD
Parameters: kgct, kpct, kgtc, kptc, kN, kU

C0 = U0 + UG + UP + 3*TH

TH′ = (1/3)*KN*U0∧3 − Ku*TH
U0′ = kgct*UG + kpct*UP + 3*kU*TH − kgtc*U0 −
   kptc*U0 − kN*U0∧3
UG′ = kgtc*U0 − kgct*UG
UP′ = kptc*U0 − kpct*UP
CD′ = e3*1000*TH′ + e1*(U0′ + UG′ + UP′)

// initial conditions, GPO4 model at 0.13 mM
t = 0.0
C0 = 0.000 13
TH = 0
CD = e1*C0
U0 = 0.43*C0
UG = 0.28*C0
UP = 0.29*C0
```

[a] e3 and e1 are the molar ellipticity of the triple helix and the monomer, respectively. A particular set of values determined form equilibrium melt experiment (see Material and Methods) was assigned to these two parameters during each fit at different temperatures.

ACKNOWLEDGMENT

We thank Drs. Jean Baum, Jannette Carey, Angela Mohs, and Anton Persikov for helpful comments and critical reading of the manuscript. We are grateful to Dr. Smita Patel and her laboratory for their help with Scientist and for suggestions on presentation of the fitting results.

REFERENCES

1. Rich, A., and Crick, F. H. C. (1961) *J. Mol. Biol. 3*, 483−506.
2. Bella, J., Eaton, M., Brodsky, B., and Berman, H. M. (1994) *Science 266*, 75−81.
3. Josse, J., and Harrington, W. F. (1964) *J. Mol Biol. 9*, 269−287.
4. Privalov, P. L. (1982) *Adv. Protein Chem. 35*, 1−104.
5. Persikov, A. V., Ramshaw, J. A. M., Kirkpatrick, A., and Brodsky, B. *Biochem. 39*, 14960−14967.

6. Brandts, J. F., Halvorson, H. R., and Brennan, M. (1975) *Biochemistry 14*, 4953−4963.
7. Creighton, T. E. (1978) *J. Mol. Biol. 125*, 401−406.
8. Bruckner, P., Bachinger, H. P., Timpl, R., and Engel, J. (1978) *Eur. J. Biochem. 90*, 595−603.
9. Bachinger, H. P., Bruckner, P., Timpl, R., and Engel, J. (1978) *Eur. J. Biochem. 90*, 605−613.
10. Baum, J., and Brodsky, B. (2000) in *Mechanisms of Protein Folding* (Pain, R. H., Ed.) 2nd ed., pp 330−351, Oxford University Press, Oxford, U.K.
11. Buevich, A. V., Dai, Q.-H., Liu, X., Brodsky, B., and Baum, J. (2000) *Biochemistry 39*, 4299−4308.
12. Bachinger, H. P., Bruckner, P., Timpl, R., Prockop, D. J., and Engel, J. (1980) *Eur. J. Biochem. 106*, 619−632.
13. Engel, J., and Prockop, D. J. (1991) *Annu. Rev. Biophys. Biophys. Chem. 20*, 137−152.
14. Bulleid, N. J., Dalley, J. A., and Lees, J. F. (1997) *EMBO J. 16*, 6694−6701.
15. Bulleid, N. J., Wilson, R., and Lees, J. F. (1996) *Biochem. J. 317*, 195−202.
16. Bruckner, P., and Eikenberry, E. (1984) *Eur. J. Biochem. 140*, 397−399.
17. Ackerman, M. S., Bhate, M., Shenoy, N., Beck, K., Ramshaw, A. M., and Brodsky, B. (1999) *J. Biol. Chem. 274*, 7668−7673.
18. Kobayashi, Y., Sakai, R., and Kakiuchi, K. (1970) *Biopolymers 9*, 415−425.
19. Kramer, R. Z., Bella, J., Brodsky, B., and Berman, H. M. (2001) *J. Mol. Biol. 311*, 131−147.
20. Li, M.-H., Fan, P., Brodsky, B., and Baum, J. (1993) *Biochemistry 32*, 7377−7387.
21. Go, N., and Suezaki, Y. (1973) *Biopolymers 12*, 1927−1930.
22. Schwarz, M., Jr., and Poland, D. (1974) *Biopolymers 13*, 687−701.
23. Sutoh, K., and Noda, H. (1974) *Biopolymers 13*, 2477−2488.
24. Weidner, H., Engel, J., and Fietzek, P. (1974) in *Peptides, Polypeptides and proteins* (Blout, E. R., Bovey, F. A., Goodman, M., and Lotan, N., Eds.) pp 419−435, Wiley-International, New York.
25. Piez, K. A., and Sherman, M. R. (1970) *Biochemistry 9*, 4134−4140.
26. Roth, W., and Heidemann, E. (1980) *Biopolymers 19*, 1909−1917.
27. Liu, X., Kim, S., Dai, Q.-H., Brodsky, B., and Baum, J. (1998) *Biochemistry 37*, 15528−15533.
28. Yang, W., Battineni, M. L., and Brodsky, B. (1997) *Biochemistry 36*, 6930−6935.
29. Reimer, U., Scherer, G., Drewello, M., Kruber, S., Schutkowski, M., and Fischer, G. (1998) *J. Mol. Biol. 279*, 449−460.
30. Sarkar, S. K., Young, P. E., Sullivan, C. E., and Torchia, D. A. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81* (15), 4800−4803.
31. Bretscher, L. E., Jenkins, C. L., Taylor, K. M., DeRider, M. L., and Raines, R. T. (2001) *J. Am. Chem. Soc. 123*, 777−778.
32. Eberhardt, E. S., Nicholas Panasik, J., and Raines, R. T. (1996) *J. Am. Chem. Soc. 118*, 12261−12266.
33. Nicholas Panasik, J., Eberhardt, E. S., Edison, A. S., Powell, D. R., and Raines, R. T. (1994) *Int. J. Pept. Protein Res. 44*, 262−269.
34. Renner, C., Alefelder, S., Bae, J. H., Budisa, N., Huber, R., and Moroder, L. (2001) *Angew. Chem., Int. Ed. 40*, 923−925.
35. Oliveberg, M., Tan, Y.-J., and Fersht, A. R. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 8926−8929.
36. Schindler, T., and Schmid, F. X. (1996) *Biochemistry 35*, 16833−16842.
37. Ibarra-Molero, B., Makhatadze, G. I., and Matthews, C. R. (2001) *Bichemistry 40*, 719−731.
38. Grathwohl, C., and Wuthrich, K. (1976) *Biopolymers 15*, 2043−2057.
39. Grathwohl, C., and Wuthrich, k. (1981) *Biopolymers 20*, 2623−2633
40. Bosshard, H. R., Durr, E., Hitz, T., and Jelesarov, I. (2001) *Biochemistry 40*, 3544−3552.
41. Goldberg, J. M., and Baldwin, R. L. (1998) *Biochemistry 37*, 2546−2555.
42. Bilsel, O., and Matthews, C. R. (2000) *Adv. Protein Chem. 53*, 153−207.
43. Locardi, E., Kwak, J., Scheraga, H. A., and Goodman, M (1999) *J. Phys. Chem. A 103*, 10561−10566.
44. Bansal, M., and Ananthanarayanan, V. S. (1988) *Biopolymers 27*, 299−312.
45. Cantor, C. R. and Schimmel, P. R. (1980) in *Physical Chemistry*, Part III, pp 1215−1219, W. H. Freeman and Company, New York.
46. Milla, M. E., and Sauer, R. T. (1994) *Biochemistry 33*, 1125−1133.
47. Chen, B.-L., Baase, W. A., and Schellman, J. A. (1989) *Biochemistry 28*, 691−699.
48. Doi, T., Higashino, K., Kerihara, Y., Wada, Y., Miyazaki, H. N., Uesugi, S., Imanishi, T., Kawabe, Y., Itakura, H., Yazaki, Y., Matsumoto, A., and Kodama, T. (1993) *J. Biol. Chem. 268*, 2126−2133.
49. Mazzorana, M., Gruffat, H., Sergeant, A., and van der Rest, M. (1993) *J. Biol. Chem. 268*, 3029−3032.
50. Hoppe, H. J., Barlow, P. N., and Reid, K. B. (1994) *FEBS Lett. 344*, 191−195.
51. Mechling, D. E., Gambee, J. E., Morris, N. P., Sakai, L. Y., Keene, D. R., Mayne, R., and Bachinger, H. P. (1996) *J. Biol. Chem. 271*, 13781−13785.
52. Lesage, A., Penin, F., Geourjon, C., Marion, D., and van der Rest, M. (1996) *Biochemistry 35*, 9647−9660.
53. Wenstrup, R. J., Shrago-Howe, A. W., Lever, L. W., Philips, C. L., Byers, P. H., and Cohn, D. H. (1991) *J. Biol. Chem. 266*, 2590−2594.

BI015952B