# Nuclear Magnetic Resonance Shows Asymmetric Loss of Triple Helix in Peptides Modeling a Collagen Mutation in Brittle Bone Disease†

Xiaoyan Liu,‡ Seho Kim,§ Qing-Hong Dai,‡ Barbara Brodsky,|| and Jean Baum*,‡

*Department of Chemistry, Rutgers University, Piscataway, New Jersey 08855-0939, Department of Molecular Biology and Biochemistry, Rutgers University, Piscataway, New Jersey 08854, and Department of Biochemistry, University of Medicine and Dentistry of New Jersey-Robert Wood Johnson Medical School, Piscataway, New Jersey 08854*

*Received May 15, 1998; Revised Manuscript Received August 10, 1998*

ABSTRACT: To investigate a human folding disease, NMR studies were carried out on collagen-like peptides to define the structural consequences of a single amino acid change found in patients with osteogenesis imperfecta (OI), a disease characterized by fragile bones. One peptide included a normal collagen sequence, while a second peptide included a Gly → Ser substitution as found in a nonlethal case of OI. Residue specific internal dynamics and conformational studies indicate that the normal collagen-like sequence forms a triple helix which is rigid along its entire length. The introduction of a Gly → Ser substitution induces an asymmetric disruption of the uniform triple helix. While the C-terminal end of the peptide retains the triple helix, the Ser substitution site and residues N-terminal to it exhibit the mobility of a random chain. This equilibrium state indicates that a Gly substitution can terminate the C to N propagation of the triple helix and suggests that renucleation is required for folding to continue. Defective folding has been implicated in brittle bone disease, and these results begin to characterize the folding process in OI collagens. OI collagen studies may also provide insights about defective protein folding, assembly, and aggregation in other human diseases.

Recent elucidation of numerous mutations in collagen which result in connective tissue disorders has stimulated an interest in relating molecular features to pathology (*1*). The most well-characterized collagen disease is osteogenesis imperfecta (OI)[1] or brittle bone disease, a genetic disorder characterized by fragile bones. All forms of OI result from mutations in type I collagen, the major structural protein in bone (*2*). The type I collagen triple helix is a supercoiled assembly of three extended polyproline II-like chains, and is a heterotrimeric molecule composed of two α1(I) chains and one α2(I) chain (*3*, *4*). The close packing of the three chains can only accommodate Gly as every third residue, leading to the repetitive pattern $(Gly\text{-}X\text{-}Y)_n$. In most cases of OI, a single Gly in the Gly-X-Y repeating sequence is replaced by another amino acid, such as Ser, Cys, Asp, or Arg. The phenotype of OI shows extreme variation, ranging from mild to lethal. Gly substitution sites resulting in OI are found all along the triple-helix domain, with lethal mutations more prevalent near the C-terminal portion of the molecule (*2*, *5*).

Characterization of mutant OI collagens has shown that all have increased levels of post-translational modification of lysine (*2*, *6*). Post-translational modifications, such as hydroxylation of lysine, can take place on the unfolded polypeptide chain only, so the increased level of modification suggests a delay of triple-helix folding. Protease digestion experiments for monitoring the triple-helix content of a number of OI collagens support such a delay in folding (*7*). The mechanism by which a Gly substitution alters folding is not known. Most OI collagens show a small decrease in thermal stability which may result from conformational changes, and a number of models have been proposed (*4*, *6*). Molecular definition of the mutation site could provide a basis for defining the relationship between the site of the mutations, the decreased stability, the delayed folding, and the severity of the clinical phenotype.

NMR investigations of conformation, dynamics, and folding at specific sites within triple-helical peptides offer an opportunity to define detailed molecular changes that result from Gly → X mutations (*8*−*11*). A set of triple-helix peptides was designed to model a specific OI mutation in the αl(I) chain (*12*). One peptide contains the native αl-(I) sequence, while a homologous peptide contains a Gly → Ser substitution at site 901 that is known to result in a nonlethal case of OI. These peptides were synthesized with $^{15}N$-enriched residues at specific sites to allow NMR investigations of the dynamic and structural changes at the mutation site, and at defined positions N- and C-terminal to it. Major changes in dynamics and conformation resulted from the Gly → Ser substitution, and the degree of disruption increased from the C to N terminus of the peptide.

† This work was supported by NIH Grants GM45302 to J.B. and AR19626 to B.B. J.B. is a Camille and Henry Dreyfus Research Scholar.

* To whom correspondence should be addressed.

‡ Department of Chemistry, Rutgers University.

§ Department of Molecular Biology and Biochemistry, Rutgers University.

|| University of Medicine and Dentistry of New Jersey-Robert Wood Johnson Medical School.

[1] Abbreviations: OI, osteogenesis imperfecta; Hyp or O, hydroxyproline; $R_1$, longitudinal relaxation rate $(1/T_1)$; $R_2$, transverse relaxation rate $(1/T_2)$.

10.1021/bi981147u CCC: $15.00 
Published on Web 10/09/1998

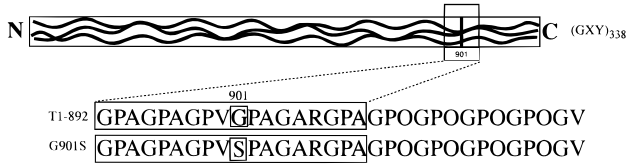


FIGURE 1: Schematic representation of the collagen triple helix showing the Gly → Ser substitution site at position 901. The expanded box region shows the 18-residue collagen sequence surrounding the position 901 substitution site. The peptide T1−892 has the sequence shown in the box with a $(GPO)_4$ sequence at the C-terminal end to ensure stability. Homologous peptide T1−892[G901S] has the same sequence as T1−892 except for the substitution of a Gly to Ser at position 901.

## MATERIALS AND METHODS

*Sample Preparation.* Peptides were purchased from Synpep Corp. (Dublin, CA). The concentrations of the peptide solutions used in all NMR experiments were 8 mM with 10% $D_2O$/90% $H_2O$ at pH 2.3.

*NMR Spectroscopy.* NMR experiments were performed on a Varian Unity Plus 500 MHz spectrometer equipped with a triple-resonance probe and pulsed field gradient. Two-dimensional data sets were processed on a Silicon Graphics workstation using the FELIX 2.05 software package (Biosym, Inc., San Diego, CA). Heteronuclear single-quantum coherence spectroscopy (HSQC) spectra were recorded at 5 °C using gradients for coherence pathway selection (*13*, *14*). HSQC−NOESY experiments were performed at 5 °C with two mixing times of 100 and 150 ms, and HSQC−TOCSY experiments were performed with a 50 ms mixing time (*9*, *15*, *16*).

*$^{15}N$ Relaxation.* The pulse sequences used to determine $^{15}N$ $T_1$, $T_2$, and heteronuclear NOE were described previously (*8*, *17*, *18*). The spectral widths were 2 and 5 kHz in the $^{15}N$ and $^{1}H$ dimensions, respectively. The time domain data sets consisted of 2048 complex points in the $t_2$ dimension and 96 increments in the $t_1$ dimension. Sixteen scans per $t_1$ increment were used in the $T_1$ and $T_2$ relaxation experiment, and 32 scans per $t_1$ increment were used in the NOE experiment. A recycle delay of 3 s was used in all relaxation experiments. Nine relaxation delays (40, 120, 200, 300, 400, 500, 600, 800, and 1000 ms) were used in the $T_1$ measurement. Nine relaxation delays (8.8, 17.7, 26.5, 35.3, 44.1, 52.9, 70.6, 88.2, and 105.8 ms) were used in the $T_2$ measurement. For the NOE measurement, spectra were recorded in the presence and absence of $^{1}H$ saturation. All the relaxation experiments were performed twice, and the $T_1$, $T_2$, and NOE values were obtained on the basis of the average of two data sets.

*Hydrogen Exchange.* The amide hydrogen exchange experiments were carried out at 5 °C and pH 2.3. The samples were equilibrated in $H_2O$ at 4 °C for 48 h to ensure triple-helix formation. The dried sample was quickly dissolved in $D_2O$ to initiate the H to D exchange, whereupon the HSQC spectra were acquired with 4 min intervals.

## RESULTS

*NMR Assignments of the $^{15}N$-Labeled Residues in the Native and Gly → Ser Substituted Peptides.* NMR studies were carried out on two peptides, one (T1−892) containing a Gly at residue 901 of the αl(I) chain, as found in normal sequences, and the second peptide (T1−892[G901S]) including Ser at position 901 as found in the OI mutated chain (Figure 1). Peptides include 18 residues from the sequence surrounding the αl(I) 901 site, together with four Gly-Pro-

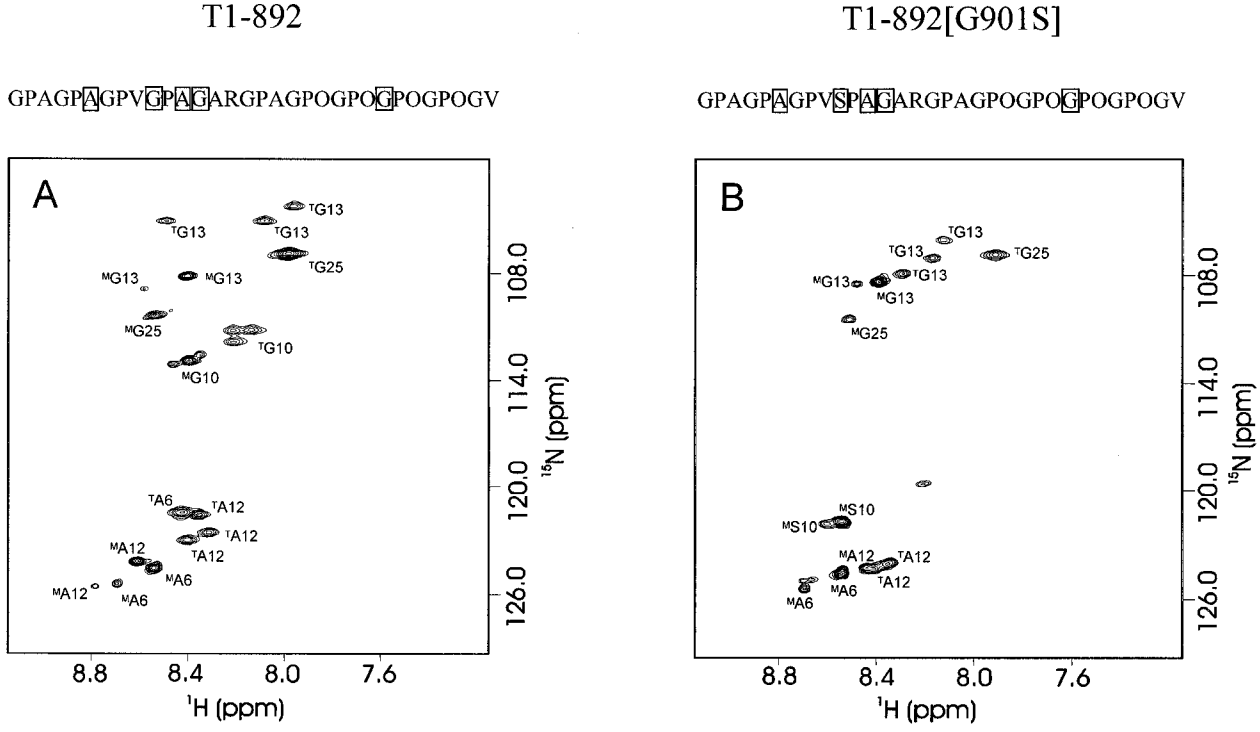


FIGURE 2: HSQC spectra in 90% $H_2O$/10% $D_2O$ at 5 °C and pH 2.3 for (A) T1−892 and (B) T1−892[G901S]. The peaks corresponding to the monomer or trimer state are denoted with a superscript M or T, respectively. The peptide sequence and the positions of the $^{15}N$-labeled amino acids are boxed. For T1−892, two peptides were synthesized with labeled positions at A6, and G10 for peptide 1 and A12, G13, and G25 for peptide 2. The HSQC spectra of peptides 1 and 2 are overlaid to produce the spectrum shown in the figure. For T1−892[G901S], two peptides were synthesized with labeled positions at A6, G13 for peptide 1 and G10, A12, G13, and G25 for peptide 2. The spectrum shown in the figure is the superposition of the spectra of peptides 1 and 2.

Table 1: Hydrogen Exchange Parameters for T1−892 and T1−892[G901S]

| peptide | residue | $k_{exch}$[a] ($min^{-1}$) | $P$[b] |
|---|---|---|---|
| T1−892 | G10 | $5.0 \times 10^{-5}$ | $2.9 \times 10^{2}$ |
| | G13 | $3.0 \times 10^{-5}$ | $1.6 \times 10^{3}$ |
| | G25 | $4.0 \times 10^{-5}$ | $>3.8 \times 10^{4}$ |
| T1−892[G901S] | S10 | $1.2 \times 10^{-2}$ | 1.7 |
| | G13 | $1.5 \times 10^{-2}$ | 2.8 |
| | G25 | $7.8 \times 10^{-4}$ | $>2.0 \times 10^{3}$ |

[a] $k_{exch}$ is the average exchange rate for the trimer resonances of each labeled residue. [b] $P$ is defined by $P = k_{exch}(\text{monomer})/k_{exch}(\text{trimer})$.

Hyp (Hyp is hydroxyproline or O) tripeptide units added to the C terminus to promote triple-helix formation. $^{15}$N-labeled amino acids were incorporated into T1−892 and T1−892-[G901S] so the conformation and dynamics of the native and Gly → Ser substituted peptides could be studied at specific positions along the peptide chain. Going from the C- to N-terminal direction, the peptides are labeled (Figure 2) at five positions: in the very stable Gly-Pro-Hyp region (Gly25), C-terminal to the substitution site (Gly13 and Ala12), at the substitution site itself (Gly10 in T1−892 and Ser10 in T1−892[G901S]), and N-terminal to the substitution site (Ala6).

The $^{1}$H and $^{15}$N resonances of the triple-helical peptide T1−892 are assigned using HSQC, HSQC−NOESY, and HSQC−TOCSY experiments (*9*, *10*). The NMR spectra of T1−892 (Figure 2A) indicate that this peptide forms a mixture of single-stranded and triple-helical molecules in solution at low temperatures. The assignments of the monomer signals are obtained by extrapolating from a high-temperature spectrum that contains only monomer peaks back to the low-temperature spectrum. Ala6, Ala12, Gly10, Gly13, and Gly25 have more than one resonance assigned to the monomer, and these are likely to be associated with multiple conformations arising from Pro cis−trans isomerization (*19*). The peaks that disappear at temperatures that are higher than the $T_m$ (thermal melting temperature) are derived from the trimer form. All labeled residues showed trimer as well as monomer peaks, and a constant ratio of trimer to monomer intensity, indicating that the entire peptide is triple-helical.

The number of trimer peaks for a specifically labeled residue depends on the amino acid sequence that surrounds that residue (*9*). In the low-temperature spectrum, Gly10, Ala12, and Gly13 have three peaks that are derived from the trimer form. Three peaks arise because the nonequivalent environment surrounding each residue gives rise to small differences in chemical shifts for the residues on the three different chains of the trimer. For Ala6 and Gly25, there is only a single peak associated with the trimer form because the repetitive GPO sequence that surrounds Gly25 and the repeating GPA sequences around Ala6 result in a more symmetric environment for these residues in the trimer. The amide hydrogen exchange rates for the labeled residues in T1−892 were determined from $^{1}$H−$^{15}$N HSQC experiments at 5 °C, and the exchange rates and protection factors (*20*) are listed in Table 1. In T1−892, all residues in the triple helix are more protected than the residues of the monomer chain, with the hydrogen exchange protection factors highest for the Gly-Pro-Hyp region.

Table 2: $^{15}$N Relaxation Parameters for the Labeled Residues in T1−892 and T1−892[G901S][a]

| peptide | residue | $R_1$ ($s^{-1}$) | $R_2$ ($s^{-1}$) | NOE |
|---|---|---|---|---|
| T1−892 | A6 | 1.55 (1.35) | 15.57 (2.02) | 0.53 (−1.02) |
| | G10 | 1.66 (1.30) | 15.40 (1.34) | 0.62 (−0.48) |
| | | 1.83 | 15.07 | 0.63 |
| | | 1.85 | 14.14 | 0.61 |
| | A12 | 1.53 (1.31) | 14.94 (2.85) | 0.70 (−0.51) |
| | | 1.39 | 15.49 | 0.62 |
| | | 1.49 | 15.42 | 0.68 |
| | G13 | 1.64 (1.48) | 15.48 (2.10) | 0.71 (−0.39) |
| | | 1.65 | 15.14 | 0.64 |
| | | 1.69 | 15.86 | 0.64 |
| | G25 | 1.65 (1.44) | 14.66 (3.26) | 0.65 (−0.37) |
| T1−892[G901S] | A6[b] | 1.37 | 1.39 | −1.00 |
| | S10[b] | 1.53 | 4.41 | −0.24 |
| | A12[c] | 1.54 (1.78) | 6.29 (2.67) | 0.07 (−0.26) |
| | | 1.45 | 8.96 | 0.21 |
| | G13 | 1.48 (1.67) | 7.91 (2.33) | 0.15 (−0.30) |
| | | 1.42 | 10.38 | 0.31 |
| | | 1.38 | 12.69 | 0.43 |
| | G25 | 1.39 (1.66) | 15.34 (2.71) | 0.60 (−0.17) |

[a] Relaxation parameters are reported for each trimer resonance of a specifically labeled residue. The values in parentheses are the relaxation parameters for the single-chain conformation. [b] The trimer and monomer peaks are overlapped. [c] Two trimer peaks are overlapped for A12.

For the substituted peptide T1−892[G901S], the HSQC spectrum (Figure 2B) shows both monomer and trimer peaks for some, but not all, labeled residues, indicating that only part of the peptide is in the triple-helical conformation at low temperatures. Gly25 at the C-terminal end of the substituted peptide has a chemical shift identical to that of the equivalent residue in T1−892. When we moved toward the N-terminal end, Gly13 and Ala12 in the substituted peptide give rise to trimer peaks that have chemical shifts that are closer to the monomer peak than the chemical shifts of the trimer peaks of Gly13 and Ala12 in T1−892. The monomer-to-trimer ratio increases at Ala12 and Gly13, indicating that conformational exchange is occurring between resonances in a trimer-like environment and monomer-like environment. Strikingly, the substitution site, Ser10, has a broad resonance with no well-defined trimer peak, and Ala6, N-terminal to the substitution site, has sharp resonances at the monomer position and no distinct trimer peak. From the C- to N-terminal end, the chemical shifts of the resonances arising from the trimer form of the mutant peptide are becoming less well dispersed relative to one another and are moving toward the monomer chemical shift. This trend continues to the point where the trimer peaks of the mutant peptide overlap with each other and with the monomer peaks, at the mutation site and at the N-terminal end of the peptide. In T1−892[G901S], the protection factors are all lower than those in T1−892, and the variation from the C- to N-terminal end is significantly larger than in T1−892. The amide protons at the N-terminal end are exchanging at rates that are similar to those expected for the monomer conformation.

*Backbone Dynamics of the Native and Substituted Peptides.* $^{15}$N $T_1$, $T_2$, and NOE relaxation experiments were performed to investigate the backbone dynamics of T1−892 (Table 2). Relaxation experiments are interpreted in terms of both the overall motion and the internal motions of the protein on the picosecond to millisecond time scale. In T1−

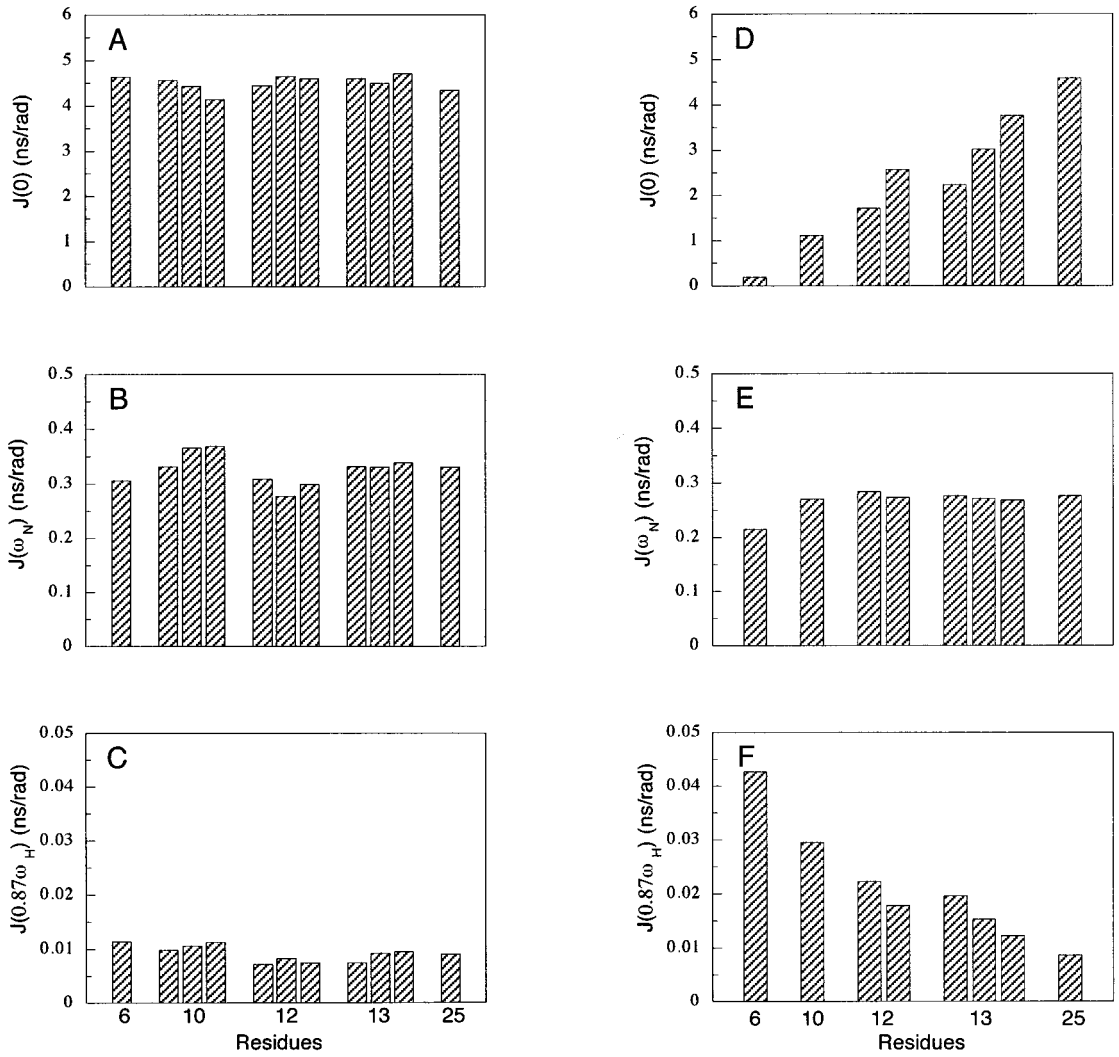


FIGURE 3: Bar graph of the spectral density values $J(0)$, $J(\omega_N)$, and $J(0.870\omega_H)$ vs $^{15}N$-labeled residues for T1−892 (A−C) and for T1−892[G901S] (D−F). Values of $R_1$, $R_2$, and NOE of the trimer (Table 2) are used to calculate the spectral density values.

892, the $R_1$ ($1/T_1$) and $R_2$ ($1/T_2$) values of all five labeled residues are similar in magnitude to one another with values of approximately 1.5 and 15, respectively. The NOE values are also similar to one another and close to the theoretical maximum, suggesting that the contribution of the internal motion to the relaxation process is negligible in T1−892.

To interpret the $T_1$, $T_2$, and NOE relaxation parameters in terms of the overall motion of the protein, and the intramolecular motions on the picosecond to millisecond time scale, one approach is to calculate the spectral density function $J(\omega)$ (*21*−*23*). The reduced spectral density mapping approach allows the derivation of the values of $J(\omega)$ at three different frequencies: $\omega = 0$, $\omega_N$, and $0.870\omega_H$ from the observed experimental relaxation parameters $T_1$, $T_2$, and NOE (*24*, *25*). At the five labeled sites in peptide T1−892, the spectral density $J(0)$ values are similar to one another, and such a similarity is observed at $J(\omega_N)$ and $J(0.870\omega_H)$ as well (Figure 3). The average value of $J(0)$ in peptide T1−892 is $4.51 \times 10^{-9}$ s/rad.

The theoretical simulation of the spectral density function in terms of specific motions requires a molecular shape, and in the case of rodlike triple-helical peptides, an anisotropic treatment is necessary (*8*). Hydrodynamic calculations indicate that a 32-mer triple-helical peptide such as peptide T1−892 can be accurately modeled by a cylinder with dimensions of 12.6 Å (diameter), 85 Å (length), and 83° (angle of a typical NH vector to the long principal axis). Calculation of rotational diffusion coefficients, using this cylinder model to account for the overall motion, resulted in a $J(0)$ of $4.8 \times 10^{-9}$ s/rad (*8*, *26*), assuming an order parameter of 1.0.

The similarity of the theoretical and experimental $J(0)$ values (Figure 3) indicates that the overall motion of the protein accounts entirely for the observed spectral density values and that there is no significant contribution to the relaxation from internal motions on the picosecond to millisecond time scale. Since the theoretical calculation of $J(0)$ assumes that the order parameter is 1.0, the order parameters of peptide T1−892 are likely to be in the range of 0.9−1.0. In summary, these results indicate that the dynamics of the five labeled residues are very similar to one another and that the triple helix is rigid along the entire T1−892 peptide.

Relaxation studies on the peptide with the Gly → Ser substitution show that the patterns of the $R_2$ and NOE values are very different from that of T1−892 (Table 2). In contrast to the relatively uniform values of $R_2$ and NOE across peptide T1−892, the $R_2$ and NOE values in T1−892[G901S]

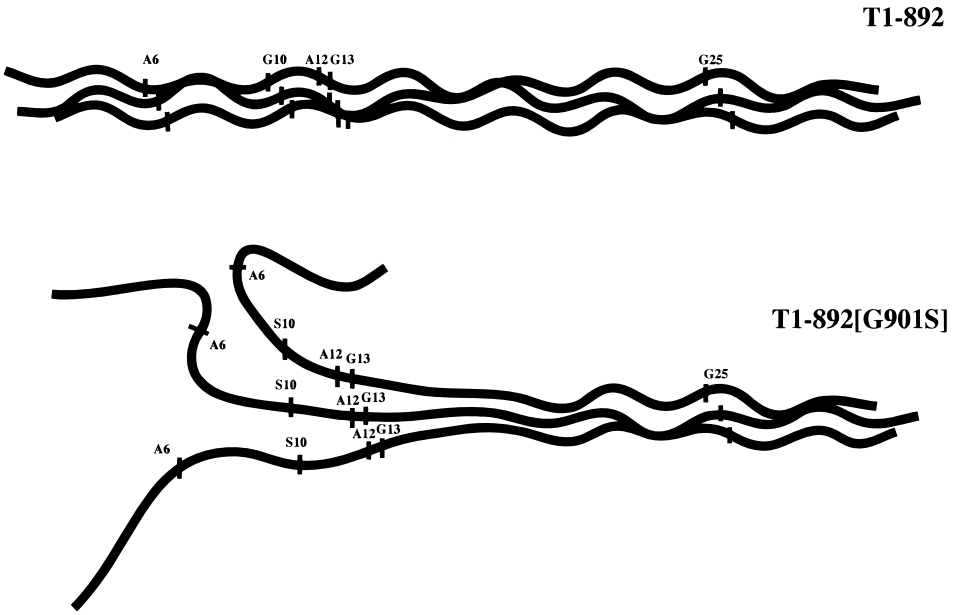


FIGURE 4: Schematic representation of the conformation of T1−892 and T1−892[G901S]. The $^{15}$N-labeled positions are denoted by the black vertical lines.

decrease substantially in a gradient manner in the C- to N-terminal direction. Specifically, the $R_2$ drops from 15.34 (Gly25) to 1.39 (Ala6), and the NOE drops from 0.6 (Gly25) to −1.0 (Ala6). The NOE and $R_2$ values obtained for Ala6 and Ser10 in the trimer form of the Gly → Ser substituted peptide are very similar to the values of the monomer form. In addition in T1−892[G901S], the three chains within the trimer have different $R_2$ and NOE values as opposed to T1−892 where the three chains within the trimer have $R_2$ and NOE values essentially identical to one another.

The profiles of the spectral density functions in T1−892 and T1−892[G901S] are very different from one another. In T1−892[G901S], there is a dramatic monotonic decline in $J(0)$ along with a monotonic increase in $J(0.870\omega_H)$ from the C- to N-terminal end. $J(0)$ varies dramatically from 4.6 $\times 10^{-9}$ s/rad at the C-terminal end to 0.02 $\times 10^{-9}$ s/rad at the N-terminal end. In addition, the one-residue stagger of the triple helix can be seen by the similarity of the dynamic properties of Gly13 in strand 2 and Ala12 in strand 1, and the similarity of the $J(0)$ values of Gly13 in strand 3 and Ala12 in strand 2. The trends along the peptide chain indicate that internal mobility is making a significant contribution to the relaxation processes of T1−892[G901S] in a gradient manner from the C- to N-terminal end. At the Gly to Ser substitution site, the spectral density function is characteristic of extremely dynamic states. N-Terminal to the substitution site, Ala6 appears to be even more flexible than Ser10 with a spectral density function that is comparable to that of the monomer in solution. C-Terminal to the substitution site, at Ala12 and Gly13, the internal mobility decreases, and as one approaches Gly25, the spectral density values are very similar to those of T1−892, indicating that the C-terminal end is rigid. A comparison of the relaxation measurements of T1−892[G901S] and T1−892 suggests that the internal dynamics at the C-terminal ends are similar to one another whereas the dynamics of T1−892[G901S] at the substitution site and N-terminal to it are more mobile than those seen for rigid T1−892.

## DISCUSSION

Observations on OI collagens suggest that basic properties of the collagen triple helix, including conformation, stability, and folding, may be affected by a Gly substitution. For the first time, NMR studies on peptides modeling a collagen disease have been used to define at a molecular level the perturbation of the triple helix. The two peptides studied provide a realistic context for the Gly substitution and contain $^{15}$N-enriched residues at five defined positions that allow the changes to be monitored at individual sites.

CD spectroscopy showed that both peptide T1−892, with a Gly at site 901, and T1−892[G901S], with the Ser at position 901, formed triple helices at low temperatures (*12*). However, the Gly → Ser substitution resulted in an apparent 50% loss in triple-helix content and a decreased thermal stability. While CD spectroscopy monitors the global features of the peptide backbone, NMR spectroscopy offers the opportunity to focus on the properties of individual $^{15}$N-enriched residues at the substitution site and other specific sites in the peptide. As described below, the image obtained from NMR information explains the loss in triple-helix content seen by CD spectroscopy, and is consistent with the decreased stability.

The dynamic and conformational features of peptide T1−892, which represents a typical collagen Gly-X-Y repeating sequence, were probed by NMR. A picture of a rigid, uniform triple helix emerges from the similarity, across all five residues, of the trimer-to-monomer ratio and the $^{15}$N backbone dynamics (Figure 4). However, the observation of a difference in hydrogen exchange rates at the different sites suggests there must be some sequence dependence of backbone mobility arising from a local "breathing" motion. For example, repeating Gly-Pro-Hyp triplets appear to offer a much greater protection against hydrogen exchange than other triplets (*8*).

A disruption of the uniform triple helix is seen in T1−892[G901S] as a result of the introduction of a Ser in place of a Gly. The data indicate that a rigid triple helix is retained

at the C-terminal Gly-Pro-Hyp rich region, which has a rigidity similar to that of the triple helix seen for T1−892. In contrast, a monomer-like state, with no indication of any triple helix, is seen for the Ser10 substitution site and the Ala6 position near the N terminus. Just C-terminal to the substitution site, Ala12 and Gly13 are still in some triple-helical form, as indicated by their trimer chemical shifts. However, the decrease in the trimer-to-monomer ratio relative to Gly25 suggests the trimer consists of interconverting states in which Ala12 and Gly13 are sometimes in a triple-helical form and sometimes in a monomer-like environment (Figure 4). In addition, the triple-helical form of Ala12 and Gly13 is far more mobile than the triple helix of the corresponding residues in T1−892, as indicated by the $^{15}$N backbone dynamics. Thus, a dramatic change in conformation and dynamics from the C to N terminus is present along the peptide.

The simultaneous presence of triple helix at the C terminus with an absence of triple helix at the N terminus in the T1−892[G901S] trimer suggests a relationship between the NMR observed equilibrium state and the C to N folding of the collagen triple helix. The strong nucleating $(Gly\text{-}Pro\text{-}Hyp)_4$ domain in peptides T1−892 and T1−892[G901S] was designed to initiate triple-helix formation at the C terminus so that propagation could take place from the C to N terminus. This would serve as a good model for the physiological folding of collagen which begins with chain association of the globular C propeptides, and then folds from the C to N terminus with a proposed zipper-like mechanism (*27*). Direct NMR observations confirmed the C to N folding direction for both peptides (unpublished). From the NMR equilibrium data, it appears that the C- to N-terminal propagation is interrupted and terminated at the Gly to Ser substitution site. The studies on Ala12 and Gly13 suggest that the folding is perturbed before the Ser substitution site is reached. By preventing the completion of the folding process, the Gly → Ser mutation appears to have trapped a stage along the folding pathway.

Our data conclusively show that a Gly → Ser mutation in a homotrimer peptide model of OI collagen stops triple-helix propagation at the mutation site and N-terminal to it. This is consistent with the observation of an increased post-translational modification level N-terminal to mutation sites in OI collagens (*2*). However, in collagens with a Gly to Ser mutation site, full-length triple-helical molecules are formed, indicating that folding does continue past the substitution site (*6*, *7*). If our homotrimer peptide model can be applied to collagen, then renucleation must occur N-terminal to the mutation site in OI collagen (*28*). The ease of renucleation may be critical in determining the overall rate of folding of full-length OI collagens, and the sequences N-terminal to the substitution site may be a principal factor in determining the severity of the disease. This work on homotrimer model systems provides an impetus for improved peptide design for better modeling of the heterotrimeric nature of mutant OI collagen trimers and indicates the need to investigate the length and sequence requirements for renucleation. The observed asymmetric disruption of the triple helix provides the first molecular description at a collagen disease site and begins the process of defining folding in OI collagens.

## SUPPORTING INFORMATION AVAILABLE

Chemical shifts of the labeled residues in Figure 2 (Table S1) and spectral density functions derived from $^{15}$N relaxation data in Table 2 (Table S2) (2 pages). Ordering information is given on any current masthead page.

## REFERENCES


1. Kuivaniemi, H., Tromp, G., Prockop, D. J., and Liams, C. (1997) *Hum. Mutat. 9*, 300−315.
2. Byers, P. H. (1993) in *Connective tissue and its heritable disorders: Molecular, genetic, and medical aspects* (Royce, P. M., and Steinmann, B., Eds.) pp 351−407, Wiley Liss, New York.
3. Rich, A., and Crick, F. H. C. (1961) *J. Mol. Biol. 3*, 483−506.
4. Bella, J., Eaton, M., Brodsky, B., and Berman, H. M. (1994) *Science 266*, 75−81.
5. Bonadio, J., and Byers, P. H. (1985) *Nature 316*, 363−366.
6. Bächinger, H. P., Morris, N. P., and Davis, J. M. (1993) *Am. J. Med. Gen. 45*, 152−162.
7. Raghunath, M., Bruckner, P., and Steinmann, B. (1994) *J. Mol. Biol. 236*, 940−949.
8. Fan, P., Li, M.-H., Brodsky, B., and Baum, J. (1993) *Biochemistry 32*, 13299−13309.
9. Li, M.-H., Fan, P., Brodsky, B., and Baum, J. (1993) *Biochemistry 32*, 7377−7387.
10. Liu, X., Siegel, D. L., Fan, P., Brodsky, B., and Baum, J. (1996) *Biochemistry 35*, 4306−4313.
11. Baum, J., and Brodsky, B. (1997) *Folding Des. 2*, R53−R60.
12. Yang, W., Battinenni, M., and Brodsky, B. (1997) *Biochemistry 36*, 6930−6935.
13. Bax, A., Ikura, M., Kay, L. E., Torchia, D. A., and Tschudin, R. (1990) *J. Magn. Reson. 86*, 304−318.
14. Davis, A. L., Keeler, J., Laue, E. D., and Moskau, D. (1992) *J. Magn. Reson. 98*, 207−216.
15. Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., and Clore, G. M. (1989) *Biochemistry 28*, 6150−6156.
16. Norwood, T. J., Boyd, J., Heritage, J. E., Soffe, N., and Campbell, I. D. (1990) *J. Magn. Reson. 87*, 488−501.
17. Palmer, A. G., III (1993) *Curr. Opin. Biotechnol. 4*, 385−391.
18. Farrow, N. A., Muhandiram, R., Singer, A., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) *Biochemistry 33*, 5984−6003.
19. Mayo, K. H., Parra-Diaz, D., McCarthy, J. B., and Chelberg, M. (1991) *Biochemistry 30*, 8251−8267.
20. Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) *Proteins: Struct., Funct., Genet. 17*, 75−86.
21. Wagner, G. (1993) *Curr. Opin. Struct. Biol. 3*, 748−754.
22. Palmer, A. G., III, Williams, J., and McDermott, A. (1996) *J. Phys. Chem. 100*, 13293−13310.
23. Dayie, K. T., Wagner, G., and Lefevre, J. F. (1996) *Annu. Rev. Phys. Chem. 47*, 243−282.
24. Farrow, N. A., Zhang, O., Szabo, A., Torchia, D. A., and Kay, L. E. (1995) *J. Biomol. NMR 6*, 153−162.
25. Peng, J. W., and Wagner, G. (1995) *Biochemistry 34*, 16733−16752.
26. Woessner, D. E. (1962) *J. Chem. Phys. 36*, 1−4.
27. Bächinger, H. P., Bruckner, P., Timpl, R., Prockop, D., and Engel, J. (1980) *Eur. J. Biochem. 106*, 619−632.
28. Wenstrup, R. J., Shrago-Howe, A. W., Lever, L. W., Phillips, C. L., Byers, P. H., and Cohn, D. H. (1991) *J. Biol. Chem. 266*, 2590−2594.


BI981147U