Two-Dimensional NMR Assignments and Conformation of (Pro-Hyp-Gly)$_{10}$ and a Designed Collagen Triple-Helical Peptide†

Ming-Hua Li,‡ Pei Fan,‡ Barbara Brodsky,§ and Jean Baum*,‡

Department of Chemistry, Rutgers University, Piscataway, New Jersey 08855-0939, and Department of Biochemistry, UMDNJ–Robert Wood Johnson Medical School, Piscataway, New Jersey 08854

Received March 16, 1993; Revised Manuscript Received May 12, 1993

ABSTRACT: Homonuclear and heteronuclear 2D NMR methods are used to study two triple-helical peptides. One peptide, (POG)$_{10}$, is considered to be the most stable prototype of a triple helix. The second peptide, (POG)$_3$ITGARGLAGPOG(POG)$_3$ (denoted T3-785), was designed to model an imino acid poor region of collagen and contains 12 residues from near the unique collagenase cleavage site in type III collagen. Both peptides associated as trimers, with melting temperatures of 60 °C for (POG)$_{10}$ and 25 °C for the T3-785 peptide. Sequence-specific assignments were made for a tripeptide unit POG in (POG)$_{10}$, and 80% of the POG triplets are found to be in an equivalent environment. In T3-785, with nonrepeating X-Y-Gly units incorporated in the sequence, the three chains of the homotrimer can be distinguished from one another by NMR. The solution conformation of (POG)$_{10}$ is very similar to the model derived from X-ray fiber diffraction data, although the peptide contains less ordered regions at the peptide ends. In the trimer form of T3-785, the central residues of the three chains are closely packed, and the data are consistent with a triple-helical model with a one-residue stagger of three parallel chains. For T3-785, in contrast to (POG)$_{10}$, there are also resonances from a less ordered form, which are probably due to the presence of a small amount of monomer. The similarity of the backbone conformations of T3-785 and (POG)$_{10}$ suggests that an alternative conformation is not present in the imino acid poor region.

The triple helix is one of the regular structural motifs found in proteins. This motif forms a major domain in all collagens and is also found in a variety of other proteins, including C1q (Brodsky-Doyle et al., 1976), acetylcholinesterase (Schumacher et al., 1986), pulmonary surfactant apoprotein (Benson et al., 1985), mannose binding protein (Drickamer et al., 1986), and the macrophage scavenger receptor (Kodama et al., 1990). In collagen, the rodlike triple-helical domains associate to form fibrils and higher order structures found in connective tissues (van der Rest & Garrone, 1991; Linsenmayer, 1991). The most well-studied collagens are those which form D-periodic fibrils. Type I collagen, a heterotrimer, is the major collagen of tendon and bone, and type III collagen, a homotrimer, is found together with type I in the collagen fibrils of skin and blood vessels (van der Rest & Garrone, 1991; Linsenmayer, 1991). In addition to forming the structural basis of collagen, certain regions of the triple helix show specific binding to integrin receptors in different cells and to other extracellular matrix molecules, such as fibronectin (Guidry et al., 1990; Kirchhofer et al., 1990; Staatz et al., 1991). The triple-helical conformation is also found in peptides, such as (Pro-Hyp-Gly)$_{10}$ (referred to as (POG)$_{10}$) and (Pro-Pro-Gly)$_{10}$ (referred to as (PPG)$_{10}$), which can serve as models for more complex biological systems (Sakikabara et al., 1973; Engel et al., 1977; Heidemann & Roth, 1982; Long et al., 1992; Brodsky et al., 1992).

The triple helix has an unusual amino acid sequence where glycine is every third residue (X-Y-Gly)$_n$. Fibrillar collagens such as types I and III contain the (X-Y-Gly)$_n$ pattern along a 1000-residue central domain, and substitution of a Gly by another amino acid results in connective tissue diseases (Kuivaniemi et al., 1991). A high proportion of imino acids are present in the X and Y positions of triple helices (Fraser & McRae, 1973; Fietzek & Kuhn, 1975). In type I collagen, Pro-Hyp-Gly triplets are the most common triplets, accounting for 12% of the sequence. The rest of the sequence includes about 22% Pro-Y-Gly triplets, 22% X-Hyp-Gly triplets, and 44% X-Y-Gly triplets with no imino acids. Sequences with relatively high imino acid content alternate with regions containing few imino acids (Dolz & Heidemann, 1986). The most imino acid rich region is the C-terminus of the triple helices of the α1 chain of type I collagen, which consists of (Pro-Hyp-Gly)$_5$. Other regions of the collagen chain include stretches of up to 24 residues which contain no imino acids. Imino acid rich regions may provide stability to the triple helix, while the imino acid poor regions may be less stable and could constitute the sites for proteinase or receptor binding (Miller, 1984; Dolz & Heidemann, 1986; Fields, 1991).

The general conformation of the triple helix is deduced from studies of collagen tissues (Rich & Crick, 1961; Ramachandran, 1967). High-angle fiber X-ray diffraction studies of stretched tendon (Fraser et al., 1979), with a linked atom least-squares refinement to intensities, have resulted in the current model of the triple helix. As seen in Figure 1, the triple helix consists of three polypeptide chains, each in an extended left-handed polyproline II-like helix with close to three residues per turn. The three chains are supercoiled about each other in a right-handed manner. This extended chain conformation is stabilized by the steric restrictions of imino acids (Nemethy & Scheraga, 1984). The three chains are staggered by one residue. For example, for (POG)$_{10}$ at any given level (Figure 1b) there is a glycine from the first chain, a proline from the second chain, and a hydroxyproline from the third chain. The three chains are closely packed near a central axis such that only glycine can be accommodated. Hydrogen bonding between the three chains and close packing

† This work was supported by the Searle Scholar Fund/Chicago Community Trust (J.B.) and NIH Grants GM45302 (J.B.) and AR19626 (B.B.).

* Author to whom correspondence should be addressed.
‡ Rutgers University.
§ UMDNJ–Robert Wood Johnson Medical School.

0006-2960/93/0432-7377$04.00/0

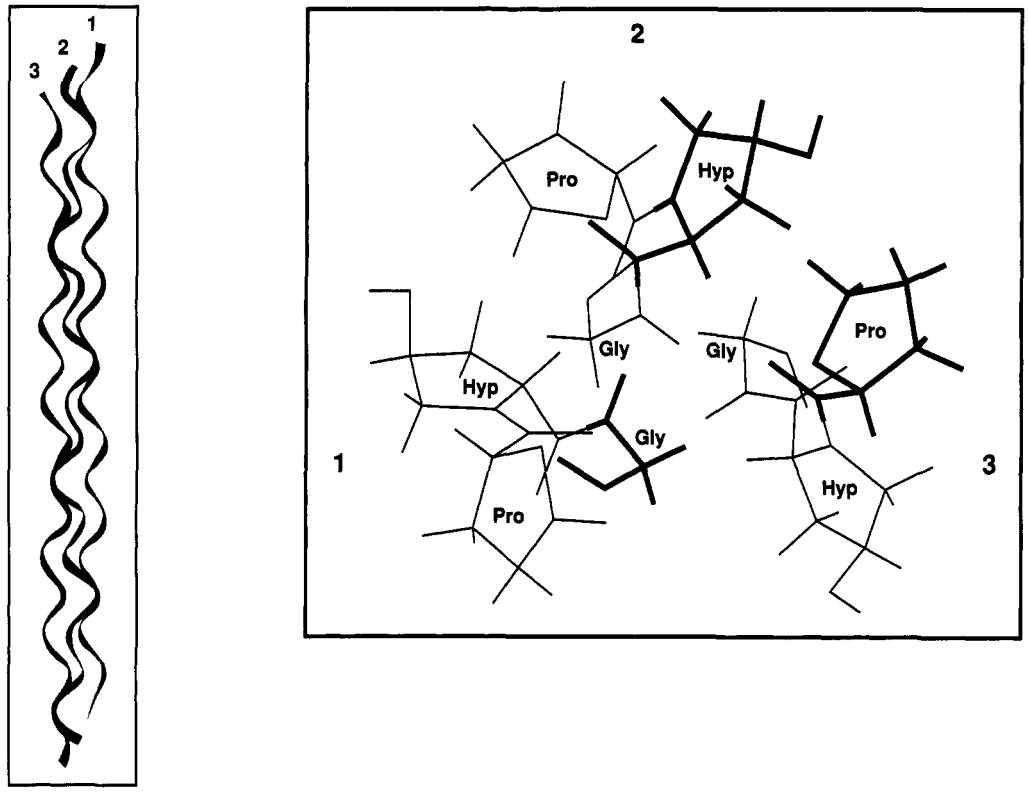

FIGURE 1: (a, left) Backbone ribbon diagram of the model triple-helical peptide $(POG)_{10}$ generated from X-ray fiber diffraction data (Fraser et al., 1979). Three chains are staggered by one residue and labeled with numbers 1, 2, and 3. The dimensions of the triple helix based on the fiber diffraction are 93 × 13 Å. (b, right) Cross-section of the triple-helical structure. One tripeptide unit Pro-Hyp-Gly from each chain is shown. The three residues at a given level come from three different chains and are Gly, Pro, and Hyp (in heavy lines).

contribute to the stability of the structure. Hydrodynamic and spectroscopic studies (von Hippel, 1967) suggest that the conformation of the triple helix in solution is generally similar to that seen by fiber diffraction in the solid state. The fiber diffraction and solution studies lend insight into the general features of the triple helix, but cannot provide details on specific interactions in the triple helix.

Solid-state and solution NMR can be used to study specific conformational aspects of the triple helix at the individual residue level. Dynamics studies in native tissues have shown significant mobility of the collagen backbone as well as side-chain residues (Jelinski & Torchia, 1979; Torchia et al., 1982; Sarkar et al., 1987). Using small collagen peptides and polypeptides such as the α1-CB2 fragment of rat skin collagen, 1D[1] ^{13}C assignments of aliphatic resonances in the random coil state have been obtained, and the mobility of residues in the coil and helix forms has been examined (Torchia & Lyerla, 1974; Torchia et al., 1975). NMR has also proved useful in detecting and quantitating the cis form of X-Pro (Hyp) peptide bonds in collagen (Roques et al., 1977; Di Blasi & Verdini, 1979; Sarkar et al., 1984; Torchia et al., 1985) and in studying single-chain X-Y-Gly repeating peptides (Mayo et al., 1991).

Recently, high-resolution two-dimensional NMR has been successfully applied to the structural determination of many biomolecules (Wuthrich, 1986, 1989; Clore & Gronenborn, 1991). NMR is uniquely suited to studying the conformation of model triple-helical peptides in solution and to examining the effects of having different amino acid sequences in the X and Y positions of the $(X\text{-}Y\text{-}Gly)_n$ repeat. We report here the first application of homonuclear and heteronuclear 2D NMR to triple-helical molecules. NMR studies were performed on two triple-helical peptides which model different regions of a collagen chain. One is $(POG)_{10}$, which models the most imino acid rich region. The second peptide, $(POG)_3$-ITGARGLAGPOG$(POG)_3$, was designed with 12 X-Y-Gly residues from type III collagen. This region represents a biologically important imino acid poor region of collagen. The 12 residues are located near the unique collagenase cleavage site (Seyer et al., 1980; Fields, 1991). This sequence contains the site (residue 790) of a Gly to Ser amino acid substitution that results in Ehlers Danlos syndrome type IV (Tromp et al., 1989) and the single trypsin cleavage site (RG) in native type III collagen (Fields, 1991). Homonuclear and heteronuclear NMR experiments are performed on both peptides to obtain sequence-specific assignments as well as NOE information. The solution NOEs are compared to expected distances from the fiber diffraction data, and the conformations of the two peptides are compared to each other.

MATERIALS AND METHODS

Chemicals. Peptide $(POG)_{10}$ was purchased from Peptides International (Louisville, KY). The ^{15}N-enriched (99%) amino acids were purchased from Cambridge Isotope Laboratory. Deuterated water (D_2O) and deuterated acetic acid (AcOD) were obtained from Aldrich.

Peptide Synthesis. Two peptides with ^{15}N-enriched amino acids were synthesized at the Protein Microchemistry Laboratory at the Center for Advanced Biotechnology and Medicine (CABM) at Rutgers University. The peptides are

[1] Abbreviations: NMR, nuclear magnetic resonance; 1D, one-dimensional; 2D, two-dimensional; CD, circular dichroism; COSY, correlated spectroscopy; DQF-COSY, double quantum filtered correlation spectroscopy; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; TOCSY, total correlation spectroscopy; HMQC, ^1H-detected heteronuclear multiple quantum coherence spectroscopy; Hyp, hydroxyproline. Standard three-letter codes and one-letter codes are used for the amino acids; the one-letter code O is used for hydroxyproline.

$(POG)_3ITGARG^*L^*A^*GPOG(POG)_3$, which will be denoted T3-785, and $(POG)_4POG^*(POG)_5$ where the asterisk denotes an ^{15}N-labeled amino acid. The peptide synthesis was carried out on an Applied Biosystem 430A peptide synthesizer by stepwise solid-phase procedures on *t*-Boc-L-Gly PAM resin. All amino acids were double coupled, and the final coupling percents as assayed by a ninhydrin test were greater than 99%. The peptides were purified by high-pressure liquid chromatography (HPLC) using a Vydac C-18 column (330-Å pores, particle size 10 μm). The purity of the peptides synthesized at CABM was greater than 95% pure.

Peptide Characterization. Liquid secondary ion mass spectrometry was conducted using a VG Analytical ZAB-T instrument at the Center for Advanced Food Technology at Rutgers University. Circular dichroism (CD) spectra were recorded on Aviv Model 60DS and 62DS spectral polarimeters. Cell path lengths were 1 mm. For equilibrium melting transitions, the temperature in the cell was increased at a rate of 0.3 °C/3 min.

Sample Preparation. The purified peptide was dissolved in water for at least 48 h prior to any experiment in order to reach the trimer/monomer equilibrium. The CD samples were prepared in 0.1 M acetic acid with a concentration of 2 mg/mL (pH 2.6). The NMR samples were prepared in H_2O/D_2O (9:1) (99.996% D_2O) solution or 0.1 M acetic acid solution, and the peptide concentration was in the range of 14–28 mM (per single polypeptide chain). All NMR experiments were run at pH 2.5–3.5.

NMR Spectroscopy. All NMR experiments were performed on a Varian VXR-500 spectrometer. The data sets were processed on a Silicon Graphics workstation using FTNMR or FELIX. For homonuclear experiments, phase-sensitive double quantum filtered correlation spectroscopy (DQF-COSY) (Aue et al., 1976; Bax & Freeman, 1981; Rance et al., 1984; Shaka & Freeman, 1983), nuclear Overhauser enhancement spectroscopy (NOESY) (Jeener et al., 1979; Kumar et al., 1980; Wider et al., 1984; Otting et al., 1986), and total correlation spectroscopy (TOCSY) (Braunschweiler & Ernst, 1983; Bax & Davis, 1985) experiments were performed with the method of time proportional phase incrementation (TPPI) (Redfield & Kuntz, 1975; Marion & Wuthrich, 1983). NOESY spectra for $(POG)_{10}$ were obtained at mixing times of 5–300 ms. For T3-785, the NOESY mixing time was 150 ms. The 2D spectra were collected with 512 or 1024 t_1 increments and 2K data points in the t_2 dimension. The spectral width is 5 kHz. Data sets were multiplied by a phase-shifted sine-bell function and zero-filled to obtain a final digital resolution of 2.44 Hz/point in both dimensions.

The ^1H-detected heteronuclear multiple quantum coherence spectroscopy (HMQC) (Mueller, 1979; Bax et al., 1983; Bax & Subramanian, 1986), HMQC-TOCSY (Gronenborn et al., 1989), and HMQC-NOESY (Gronenborn et al., 1989) experiments were collected with 200 t_1 increments and 2K data points in the t_2 dimension using the TPPI method. For HMQC and HMQC-TOCSY experiments, the spectral widths are 1.25 and 5 kHz in the ^{15}N and ^1H dimensions, respectively. The final digital resolution, after zero-filling, is 1.22 and 2.44 Hz/point in the ^{15}N and ^1H dimensions, respectively. The isotropic mixing time of HMQC-TOCSY was 60 ms. HMQC-NOESY experiments were obtained at 10 °C, and NOESY mixing times ranged from 20 to 350 ms. The spectral widths are 2.5 and 5 kHz in the ^{15}N and ^1H dimensions, respectively. The final digital resolution, after zero-filling, is 2.44 Hz/point in both dimensions. For HMQC-COSY (Clore et al., 1988) experiments, 512 t_1 increments were collected, and the final digital resolution was 2.44 Hz/point in both dimensions.

Computer Modeling. The coordinates of $(POG)_{10}$ were obtained from X-ray fiber diffraction data (Fraser et al., 1979), and the conformation was generated on INSIGHT. A model of the design peptide T3-785 was obtained from $(POG)_{10}$ by replacing the nine central residues with the residues Ile-Thr-Gly-Ala-Arg-Gly-Leu-Ala-Gly. T3-785 was energy-minimized with IMPACT (Kitchen et al., 1990) using the method of steepest descent and conjugated gradient algorithms. After energy minimization, the backbone structure and dimensions of T3-785 are similar to those of $(POG)_{10}$. When the backbone atoms of $(POG)_{10}$ and T3-785 are superimposed, a root-mean-square deviation (RMSD) of 0.3 is obtained.

RESULTS

$(POG)_{10}$. $(POG)_{10}$ is considered to be the most stable prototype of a triple helix, and a model for this peptide is derived from the fiber diffraction data (Fraser et al., 1979). In this model, the three chains are staggered by one residue each and are related by a screw axis. Thus, the Pro-Hyp-Gly unit of chain 1 is rotated by 108° and is axially translated by 2.9 Å to give an equivalent Pro-Hyp-Gly unit of chain 2. The 30-mer $(POG)_{10}$ forms a trimer with a triple-helical conformation in aqueous solution and undergoes thermal transition to a monomer at 60 °C (see Figure 6b). Ultracentrifugation studies (2 mg/mL) indicate that the peptide is so highly associated as a trimer that it is not possible to detect monomer and determine an equilibrium constant at 10 °C (C. Long, E. Braswell, B. Brodsky, and J. Baum, unpublished results). 1D ^1H NMR (Kobayashi & Kyogoku, 1973) and preliminary 2D assignments (Brodsky et al., 1992) of this peptide have been reported. To facilitate the NMR assignments of the repeating model peptide $(POG)_{10}$, two related peptides with similar amino acid sequences were used. One is the peptide $(PPG)_{10}$, which forms a stable triple helix and is similar in structure to $(POG)_{10}$, but has a lower melting temperature of 30 °C (Kobayashi et al., 1970). The second is the peptide (Pro-Hyp-Gly)$_4$(Pro-Hyp)(Pro-Hyp-Gly)$_5$, which is very similar to $(POG)_{10}$ but contains a glycine deletion in the middle of the peptide. This is designated as the glycine deletion peptide. This deletion results in a destabilization of the triple helix (Long, 1992), such that the peptide is almost entirely monomer at 10 °C.

Assignments of $(POG)_{10}$. In $(POG)_{10}$, there are 10 Pro, Hyp, and Gly residues in each chain, but 10 sets of resonances for each residue type are not seen in the NMR spectra due to extensive overlapping of the resonances in the repeating tripeptide units. Spin systems of $(POG)_{10}$ were identified using DQF-COSY and TOCSY spectra. Glycines are distinguished from Pro and Hyp by a pair of nondegenerate $C_\alpha H$ resonances and by the amide proton resonances. Hyp has the same connectivity pattern as Pro, but is distinguished from Pro by the $C_\gamma H$ which is shifted considerably downfield by the OH group. In the TOCSY spectrum at 10 °C, two sets of major cross peaks for each residue type, i.e., Pro, Hyp, and Gly, are observed. One set of the major Pro cross peaks, designated as Pro1, has two distinct resonances for the $C_\beta H$ and $C_\delta H$ protons and only one resonance for the $C_\gamma H$ protons (Figure 2a). For the set of major Hyp cross peaks, designated as Hyp1, two distinct resonances can be observed for the C_β and C_δ protons. Two major Gly peaks are observed in the $C_\alpha H$–$C_\alpha H$ region of the spectrum (data not shown). The remaining sets of major cross peaks for Pro and Hyp are designated as Pro2 and Hyp2, as shown in Figure 2b. In addition, a number of less intense peaks corresponding to two Pro, one Hyp, and four Gly are also present.

The stereospecific assignments of the β and δ methylene protons of Pro1 and Hyp1 are obtained from the NOESY

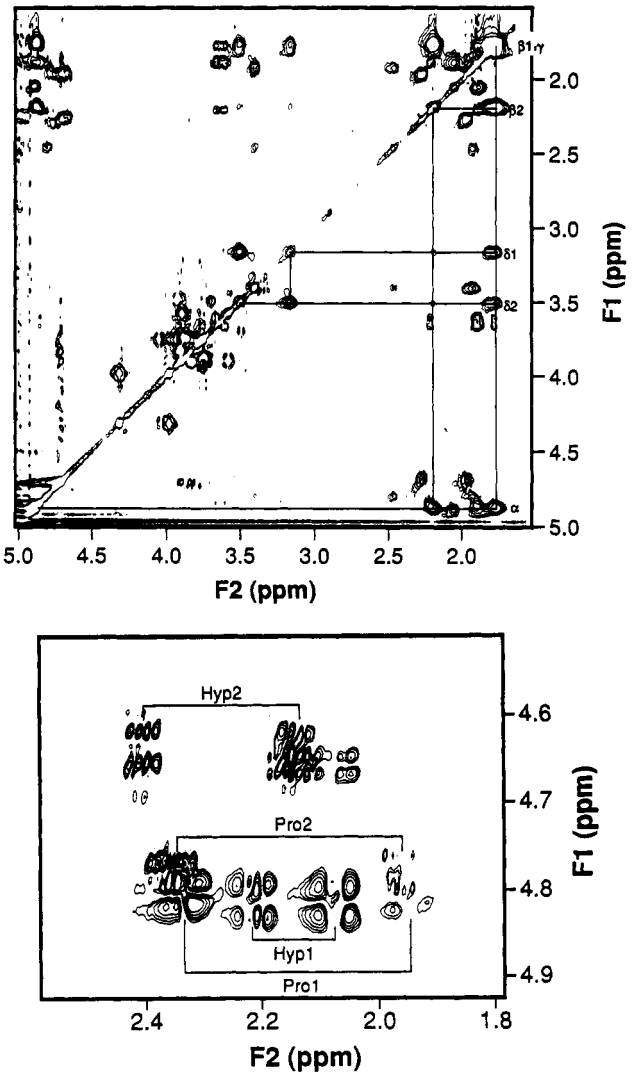

FIGURE 2: (a, top) Aliphatic region of the TOCSY spectrum of $(POG)_{10}$ dissolved in D_2O at pH 3.4 and 10 °C, with a mixing time of 40 ms. A representative Pro1 spin system is shown with a solid line. The proton assignments are labeled along the *F*1 axis. (b, bottom) Expanded region of a DQF-COSY spectrum of $(POG)_{10}$ dissolved in D_2O at pH 3.4, 10 °C, showing the $C_\alpha H$–$C_\beta H$ cross peaks of Pro2 and Hyp2 as well as Pro1 and Hyp1. The solid lines connect the pair of β protons from each imino acid. The intensity ratios of Pro2/Pro1 and Hyp2/Hyp1 in this spectrum appear to be greater than the 20%/80% which is observed in the 1D spectrum (Brodsky et al., 1992), and the discrepancy is likely due to the short T_2's of Pro1 and Hyp1 (see text).

spectrum (Wagner et al., 1987). Because the $C_\beta H_2$ is always closer to the $C_\alpha H$ due to ring restriction, the protons $C_\beta H_1$ and $C_\beta H_2$ can be distinguished on the basis of their NOE intensities with the $C_\alpha H$. The two C_δ protons can be differentiated by their different NOE cross-peak intensities with the C_β protons. The $C_\delta H_1$ is always closer in space to $C_\beta H_1$ and therefore shows a stronger cross peak with the $C_\beta H_1$. Stereospecific assignments of β and δ protons for Pro2 and Hyp2 could not be obtained because the NOE intensities of $C_\alpha H$ to $C_\beta H_1$ and $C_\beta H_2$ could not be differentiated. The chemical shifts of Pro1 and Pro2, Hyp1 and Hyp2, and Gly1 and Gly2 are listed in Table I.

Sequential connectivities obtained from the NOESY spectrum led to the definition of two distinct tripeptide units. Sequential connectivities are observed between Pro1, Hyp1, and Gly1 (Figure 3), defining this as tripeptide unit 1. Sequential connectivities were also observed between Pro2, Hyp2, and Gly2 (Figure 3), defining a second tripeptide unit, designated as tripeptide unit 2.

Table I: Chemical Shifts of Assigned Proton Resonances for $(POG)_{10}$ at 10 °C

	chemical shift (ppm)[a]				
amino acid	NH	$C_\alpha H$	$C_\beta H$	$C_\gamma H$	$C_\delta H$
Pro1		4.84	2.35, 1.96	2.01	3.57, 3.25
Hyp1		4.85	2.23, 2.08	4.69	3.94, 3.86
Gly1	7.99	3.94, 3.86			
Pro2		4.81	2.37, 1.97	2.08	3.71, 3.66
Hyp2		4.66	2.43, 2.14	4.69	3.93, 3.86
Gly2	8.60	4.32, 4.00			

[a] ^1H chemical shifts are reported relative to sodium 3-(trimethylsilyl)tetradeuteriopropionate (STP).

A number of experimental observations indicated that one of the tripeptide units (1) corresponds to residues in a triple helix, which constitutes the major portion of the molecule, while the second (2) corresponds to residues in a less ordered form. First, the relative intensity of the $C_\delta H$ from the Pro residue in tripeptide 1 versus the $C_\alpha H$ of residue Gly in tripeptide 2 can be quantified by 1D ^1H NMR, and it was found that the resonance from tripeptide 1 constitutes 80% of the intensity (Brodsky et al., 1992). This suggests that tripeptide 1 represents a number of overlapping residues and constitutes the major component of the trimer, which is triple-helical according to various spectroscopic and X-ray studies (Sakikabara et al., 1973). Second, at high temperature (85 °C), where only monomer is present in the solution, a unique tripeptide unit can be assigned from the NMR spectra, and this tripeptide unit is similar to the tripeptide 2 observed at low temperature. Third, we have studied the glycine deletion peptide, and this highly monomeric peptide has sequential connectivities which can be overlaid directly on tripeptide 2. Tripeptide 1 does not occur in the high-temperature spectrum or in the spectrum of the glycine deletion peptide. These data are consistent with tripeptide 1 being in a triple-helical conformation and tripeptide 2 in a different, less structured environment.

A $(POG)_{10}$ sample containing a single ^{15}N-labeled Gly residue at position 15, in the center of the chain, was studied. The HMQC spectrum of this peptide at 10 °C showed only a single cross peak with the amide proton chemical shift at 7.99 ppm, corresponding to the amide position of the glycines in tripeptide 1. There are no cross peaks in the HMQC spectrum corresponding to tripeptide 2. This result confirms that $(POG)_{10}$ at 10 °C is composed uniquely of trimers. It also indicates that the triple-helical region corresponds to tripeptide 1 resonances, which represent 80% of the chain. Since the labeled glycine in the center of the chain shows only one conformation, tripeptide 2, representing about 20% of the chain, must represent less ordered regions which are located near the peptide ends.

Conformation of $(POG)_{10}$. Even though the assignments of tripeptide unit 1 are obtained, it has not been possible at this stage to perform a structure determination using NOEs as distance constraints. Difficulties in direct structure determination or modeling arise from the presence of three closely packed chains in this molecule, all with identical repetitive sequences. A given NOE cross peak may result from interaction between atoms in one chain and/or between atoms in different chains. For example, as seen in Figure 4, an NOE between Gly NH and Hyp $C_\alpha H$ can arise from either an interchain interaction or an intrachain interaction. Given this complication, we have taken an initial structural approach of examining the observed NOEs within the tripeptide unit and comparing these with the model proposed for $(POG)_{10}$ from X-ray fiber diffraction data. We limited our comparison

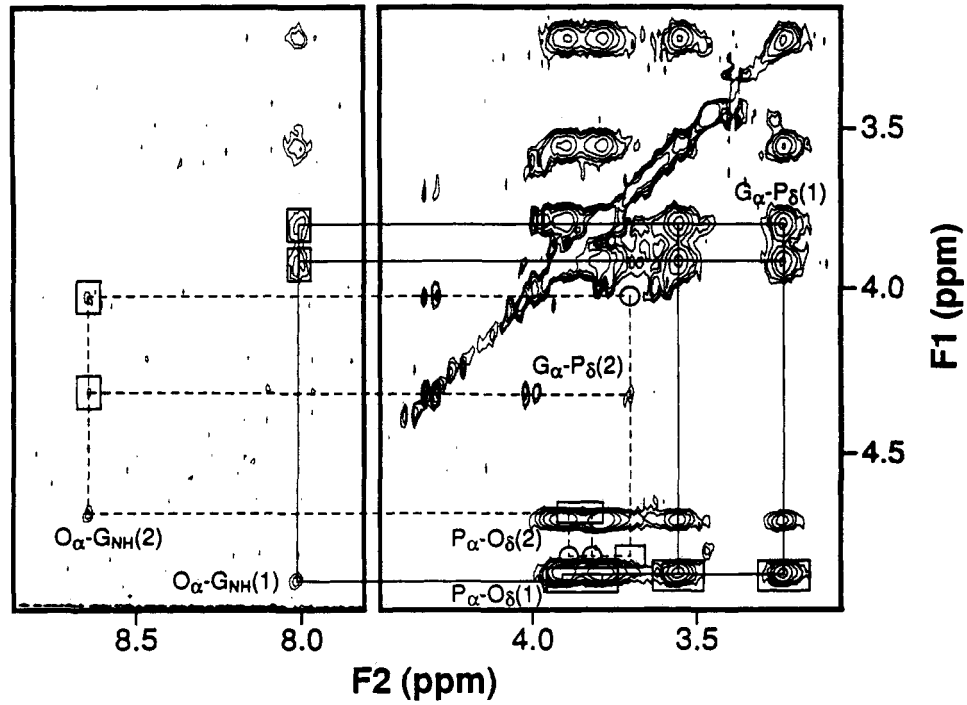

FIGURE 3: Aliphatic region of the NOESY spectrum (mixing time 300 ms) of $(POG)_{10}$ dissolved in D_2O (pH 3.4, 10 °C) together with the amide region of the NOESY spectrum (mixing time 300 ms) obtained in 90% H_2O/10% D_2O solution showing the sequential connectivities of two tripeptide units P-O-G. The solid lines connect tripeptide 1 and dashed lines connect tripeptide 2. Sequential NOEs are indicated and labeled in the spectra. The boxes represent the positions of the TOCSY cross peaks. The circles indicate the positions of the sequential NOEs that are missing in the $(POG)_{10}$ spectrum, but are observed in the spectrum of the Gly deletion peptide. For tripeptide 1, because the $C_\alpha H$'s of P and O have similar chemical shifts, the sequential NOEs between P_α and O_δ overlap with the intraresidue NOE from O_α to O_δ.

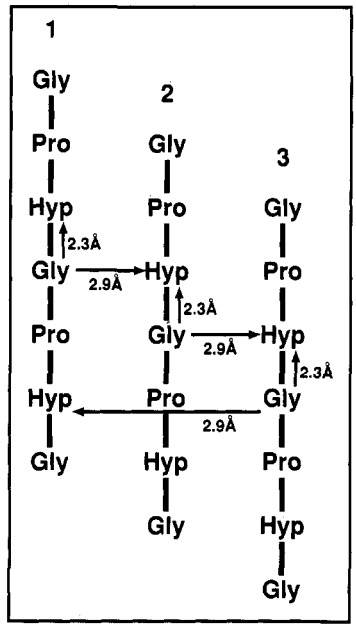

FIGURE 4: Parallel one-residue stagger schematic diagram of $(POG)_{10}$ showing the intrachain distances between GlyNH and Hyp$C_\alpha H$ (2.3 Å) and the interchain distances of GlyNH to Hyp$C_\alpha H$ (2.9 Å). It can be seen that an observed NOE between GlyNH and HypC_α can represent any one or all of these interactions.

to the backbone and/or side-chain NOEs that are predicted from interactions that can arise uniquely from interchain resonances, since these provide a critical test of the model (Table II). The spectrum of $(POG)_{10}$ has a significant amount of resonance overlap, and some of the unique interchain NOEs predicted from the model cannot be unambiguously assigned. For example, the backbone NOE between Hyp $C_\alpha H$ and Pro $C_\delta H$ cannot be distinguished from Pro $C_\alpha H$–Pro $C_\delta H$ because of the overlap of the Hyp $C_\alpha H$ and Pro $C_\alpha H$ resonances. To approach this problem, we used the assignments and NOE

Table II: Expected Unique Interchain NOEs (Distances Smaller than 4.5 Å) Based on the Model of Triple-Helical $(POG)_{10}$[a]

expected NOE	experimental NOE
Gly NH–Pro $C_\delta H_1$	+
Hyp $C_\delta H_2$–Pro $C_\delta H_1$	+* (overlap, Hyp $C_\delta H$ = Gly $C_\alpha H$)
Hyp $C_\alpha H$–Pro $C_\delta H_1$	+* (overlap, Hyp $C_\alpha H$ = Pro $C_\alpha H$)
Hyp $C_\alpha H$–Pro $C_\delta H_2$	+* (overlap, Hyp $C_\alpha H$ = Pro $C_\alpha H$)
Gly $C_\alpha H_2$–Pro $C_\alpha H$	+* (overlap, Gly $C_\alpha H$ = Hyp $C_\delta H$)
Gly $C_\alpha H_1$–Hyp $C_\beta H_1$	+* (overlap, Gly $C_\alpha H$ = Hyp $C_\delta H$)
Gly $C_\alpha H_1$–Hyp $C_\beta H_2$	+* (overlap, Gly $C_\alpha H$ = Hyp $C_\delta H$)
Pro $C_\beta H_1$–Hyp $C_\beta H_2$	+* (overlap, Pro $C_\beta H$ = Pro $C_\gamma H$)
Pro $C_\gamma H_1$–Hyp $C_\alpha H$	+* (overlap, Pro $C_\alpha H$ = Hyp $C_\alpha H$)
Pro $C_\gamma H_1$–Hyp $C_\beta H_1$	+
Pro $C_\gamma H_1$–Hyp $C_\beta H_2$	+
Pro $C_\gamma H_1$–Hyp $C_\gamma H$	+
Pro $C_\gamma H_2$–Hyp $C_\beta H_2$	+* (overlap, Pro γH_1 = Pro γH_2)
Pro $C_\delta \rho_1$–Hyp $C_\beta H_1$	+
Pro $C_\delta H_1$–Hyp $C_\beta H_2$	+
Pro $C_\delta H_1$–Hyp $C_\gamma H$	+
Pro $C_\delta H_2$–Hyp $C_\beta H_2$	+

[a] Observed NOEs are indicated with +. Peak overlap is often encountered in $(POG)_{10}$ (indicated in the parentheses). When the corresponding NOEs are observed between non-overlapping resonances in $(PPG)_{10}$, these are indicated with asterisks. Thus all predicted interchain NOEs are seen in either $(POG)_{10}$ or $(PPG)_{10}$ spectra.

data from $(PPG)_{10}$, a peptide closely related in conformation to $(POG)_{10}$ with NMR spectra very similar to those of $(POG)_{10}$. All backbone and side-chain NOEs that are overlapping in $(POG)_{10}$ are resolved in $(PPG)_{10}$ (M.-H. Li, & J. Baum, unpublished results). Over a range of mixing times, all of the expected NOEs arising from close interchain distances in the model are observed confirming the basic triple-helical conformation in solution.

The specific puckering of the Pro and Hyp rings could be determined from the observed distance constraints. Imino acid rings are never planar and can adopt two different conformations: a puckered up or puckered down conformation (Figure 5). If the C_γ atom is pushed out of the plane formed by the N, C_α, C_β, and C_δ atoms and is trans to the CO moiety,

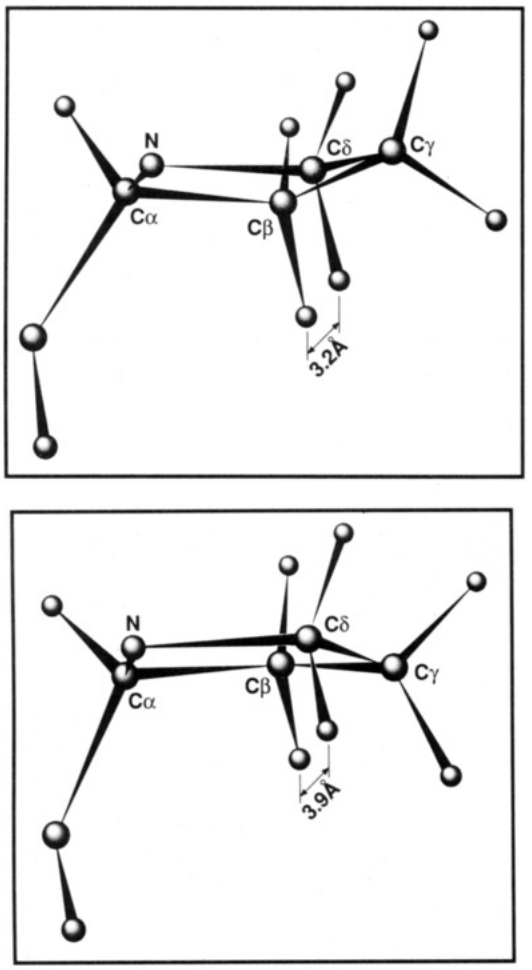

FIGURE 5: Ring diagram of Pro in (a, top) puckered up and (b, bottom) puckered down conformations. The distances between $C_\beta H_1$ and $C_\delta H_1$ in the two configurations are shown.

the conformation is puckered up; otherwise it is puckered down. The C_γ puckered up conformation has fixed dihedral angles of $\phi = -68°$ and $\chi_1 = -6°$ and exhibits a short distance between $C_\beta H_1$ and $C_\delta H_1$ (3.2 Å). The C_γ puckered down conformation has the fixed dihedral angles $\phi = -75°$ and $\chi_1 = +19°$ and the distance between $C_\beta H_1$ and $C_\delta H_1$ is 3.9 Å (Miller et al., 1980). Ring puckering has been determined by NMR for polyproline and polyhydroxyproline. It was found that polyproline adopts an averaged ring conformation, whereas polyhydroxyproline adopts a mostly puckered up conformation (Torchia, 1971, 1972). To determine the ring puckering of Pro and Hyp in solution, we searched for NOEs between the C_β and C_δ protons. The NOESY spectrum of $(POG)_{10}$ with a 10-ms mixing time shows one strong cross peak between $C_\beta H_1$ and $C_\delta H_1$ for Hyp, and no $C_\beta H$ and $C_\delta H$ cross peaks are observed for the Pro residue in this spectrum. In the NOESY spectrum with a 20-ms mixing time, one cross peak is observed between $C_\beta H_1$ and $C_\delta H_1$ of the Hyp and one between $C_\beta H_1$ and $C_\delta H_1$ of the Pro. The ratio of intensities between the Hyp and Pro cross peaks is approximately a factor of 3, which corresponds to the calculated intensity ratio for a puckered up versus a puckered down conformation (calculated to be 3.3). These results suggest that, in $(POG)_{10}$, the Hyp ring is in a puckered up conformation while the Pro ring is in a puckered down conformation.

Studies of a Designed Peptide. In the $(POG)_{10}$ NMR spectra, individual residues cannot be assigned due to a significant amount of resonance overlap. In order to break the symmetry, a peptide that incorporates different types of

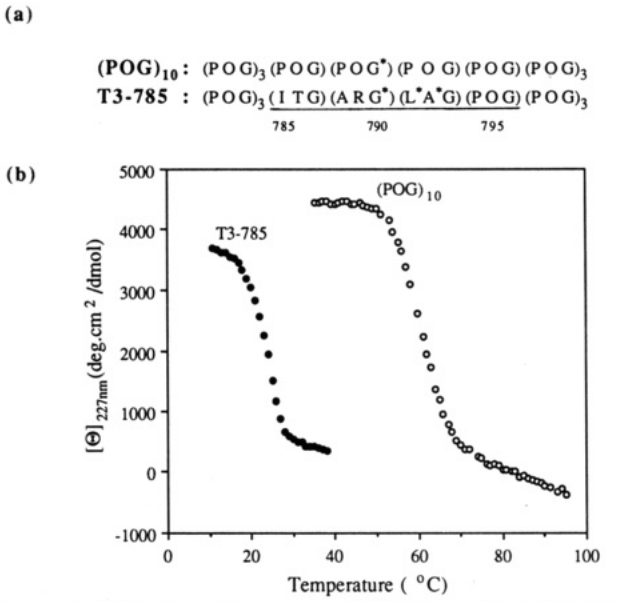

FIGURE 6: (a) Amino acid sequences of $(POG)_{10}$ and T3-785, with the sequence from type III collagen underlined and numbered according to the corresponding position in collagen (Ala-Kokko et al., 1989). The ^{15}N-labeled residues are indicated by asterisks. (b) CD equilibrium melting curves of peptides $(POG)_{10}$ (○) and T3-785 (●). The samples are dissolved in 0.1 M acetic acid solution at a concentration of 2 mg/mL. The melting rate is 0.3 °C/3 min. The melting temperatures of $(POG)_{10}$ and T3-785 are 60 and 25 °C, respectively.

amino acid residues into the triple-helical region is required. Our strategy is to design peptides that contain real collagen sequences flanked by POG caps at the C- and N-terminal ends to provide stability. The total number of residues per chain was kept at 30, so that comparison could be made with $(POG)_{10}$. The central collagen sequence was chosen from type III collagen because of its homotrimeric nature, since the trimeric peptides will have three identical chains. The peptide sequence is shown in Figure 6a. Only eight regions in type III collagen contain three or more triplets with no imino acids, and the sequence chosen comes from one of these.

The design peptide $(POG)_3$ITGARG*L*A*GPOG$(POG)_3$ (denoted T3-785) was synthesized with the ^{15}N-enriched amino acids for G-L-A at the positions indicated with asterisks. The molecular weight of the peptide is confirmed by mass spectrometry. Equilibrium ultracentrifugation indicates that the peptide is highly trimeric (>95%) (E. Braswell, unpublished results). To confirm that no further aggregation is occurring, the CD spectra were measured at a range of concentrations (0.8–6.4 mM) and found to be concentration independent. CD and NMR experiments confirm a sharp thermal transition for T3-785 near 25 °C, and its CD melting curve is shown in Figure 6b along with the melting curve of the $(POG)_{10}$ peptide. The design peptide T3-785 forms a stable trimer at temperatures suitable for study by NMR.

Assignments of T3-785. The ^{15}N chemical shifts of the three labeled residues are well resolved in the heteronuclear NMR spectra. The HMQC spectrum (Figure 7a) of the T3-785 peptide in H_2O at 10 °C, where the peptide is largely trimeric, shows a total of 12 resonances. For three labeled residues in a single trimer conformation, no more than nine peaks would be expected if all three chains of the trimer are distinguishable. The 12 resonances observed from the HMQC spectrum indicate that the peptide adopts more than one conformation in solution under the conditions of the experiment. When placed in D_2O, three of the twelve resonances exchange rapidly while nine remain (still present after 12 h).

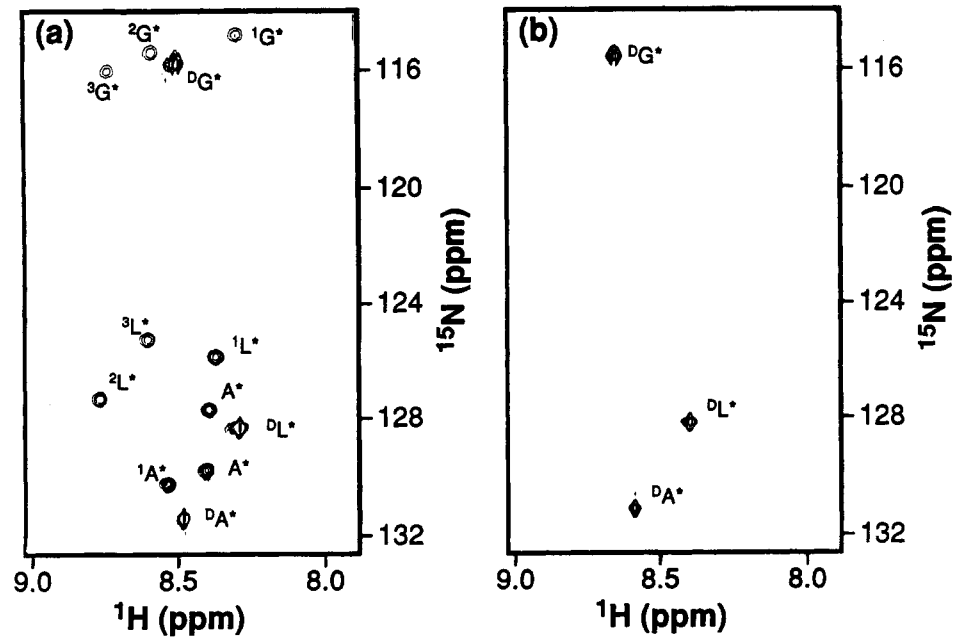

FIGURE 7: HMQC spectra of T3-785 dissolved in 90% H_2O/10% D_2O solution (pH 2.6) at (a) 10 °C and (b) 40 °C. The assignments of the ^{1}H–^{15}N correlation peaks of the three labeled residues G*, L*, and A* are shown.

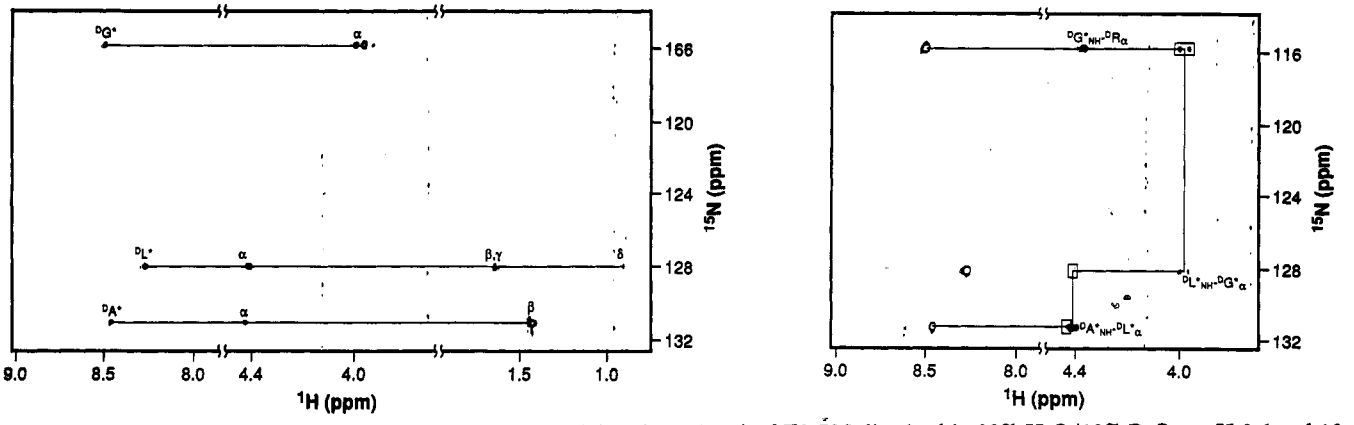

FIGURE 8: (a, left) HMQC-TOCSY spectrum (isotropic mixing time 60 ms) of T3-785 dissolved in 90% H_2O/10% D_2O at pH 2.6 and 10 °C showing the spin systems of ^{D}G*, ^{D}L*, and ^{D}A*. The solid line connects each spin system and the proton assignments are labeled along the ^{1}H axis. (b, right) HMQC-NOESY spectrum (mixing time 350 ms) of T3-785 dissolved in 90% H_2O/10% D_2O solution at pH 2.6, 10 °C, showing the sequential connectivities of tripeptide ^{D}G*-^{D}L*-^{D}A*. The boxes represent the positions of the TOCSY peaks. The absence of cross peaks from the other nine residues is due to the long mixing times used in these experiments and their short T_2's.

In H_2O, when the temperature of the solution is raised from 10 (trimer) to 40 °C (monomer), the intensities of the three rapidly exchanging amides increase while those of the nine slowly exchanging ones decrease and disappear (Figure 7b).

Using HMQC-TOCSY, shown in Figure 8a, and HMQC-COSY experiments in H_2O at 10 °C, only three spin systems corresponding to the three rapidly exchanging residues in the HMQC spectrum can be assigned. The spin systems of the nine remaining resonances could not be identified, probably because of the short transverse relaxation time of the amide protons (see the end of the Results section). The HMQC-NOESY experiment at a long mixing time (350 ms), shown in Figure 8b, indicates that these three residues are sequentially connected and correspond to a tripeptide unit of Gly*-Leu*-Ala*. This tripeptide unit will be referred to as tripeptide D and is likely to correspond to a non-triple-helical region.

The nine remaining resonances of the HMQC spectrum in H_2O were assigned with HMQC-NOESY experiments (60-ms mixing time) in conjunction with homonuclear DQF-COSY and TOCSY. The glycine residues were tentatively assigned according to the ^{15}N chemical shifts of tripeptide D. Additional support for the assignments of the glycine residues is obtained from the HMQC spectrum of $(POG)_{10}$ with a single glycine labeled at position 15. Beginning with these glycine assignments, three distinct chains, referred to as chains 1, 2, and 3, can be traced out as shown in Figure 9. A general notation ^{n}X is used to differentiate the residues in different chains with the chain number (n = 1, 2, and 3) indicated as the superscript of residue X. Sequential connections between NH of the Gly* and $C_\alpha H$ of the preceding Arg (position 14) residue, and between $C_\alpha H$ of the Gly* and NH of the following Leu* (position 16), are observed for all three chains. Chain 1 can be traced out further to see the sequential connection between the ^{1}Leu* $C_\alpha H$ (position 16) and the ^{1}Ala* NH (position 17). The $C_\alpha H$ chemical shifts of ^{2}Leu* and ^{3}Leu* overlap exactly, making it impossible to distinguish the NH and $C_\alpha H$ chemical shifts of ^{2}Ala* and ^{3}Ala*. Changing the sample conditions did not resolve the ambiguity. The spin system identifications for Arg, Leu*, Ala*, and Gly* are confirmed in the homonuclear TOCSY experiment, and the chemical shifts are listed in Table III. It should be noted that the Pro-Hyp-Gly triplets which cap both ends of T3-785 and provide its stability give NMR peaks that are very similar to those seen for the $(POG)_{10}$ peptide.

Conformation of the Design Peptide. In studying the structure of the design peptide, we focus our attention on the

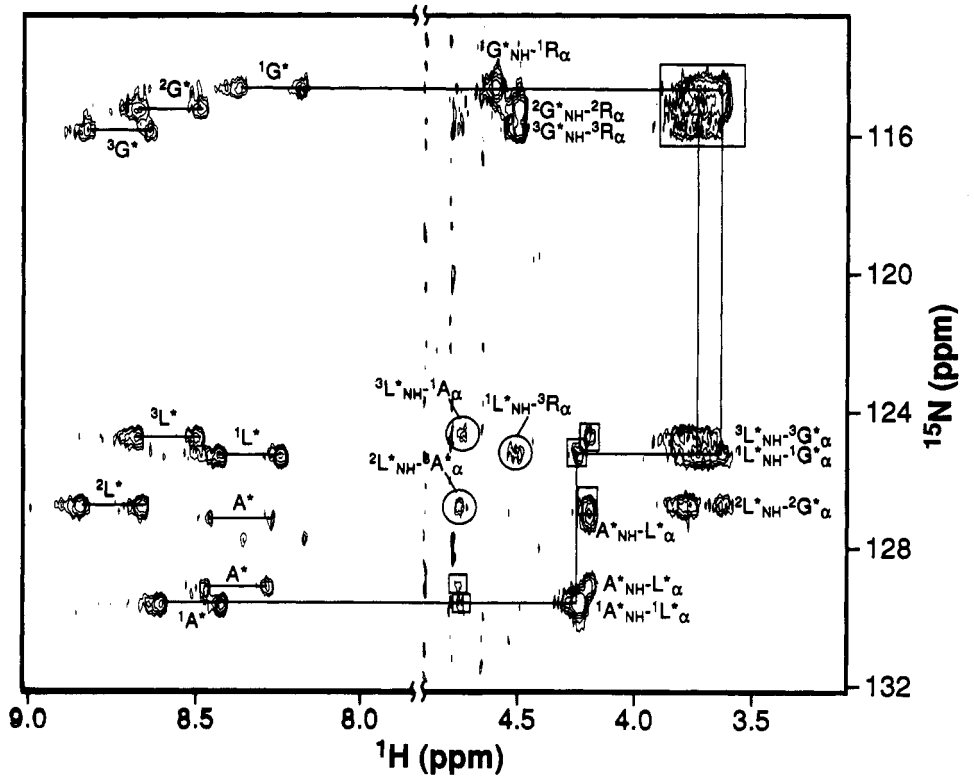

FIGURE 9: HMQC-NOESY spectrum (mixing time 60 ms) of T3-785 dissolved in D_2O at 10 °C, showing the NH and $C_\alpha H$ region of the spectrum. ^{15}N was not decoupled during the acquisition time. The sequential connectivities of chain 1 (see Figure 10) are fully traced out, and the sequence ^{1}R-$^{1}G^{*}$-$^{1}L^{*}$-$^{1}A^{*}$ is connected by solid lines. The boxed regions are the positions of HMQC-TOCSY peaks. $^{2}G^{*}$, $^{3}G^{*}$, $^{2}L^{*}$, and $^{3}L^{*}$ are indicated. Two of the alanines are not designated with chain numbers as they cannot be unambiguously assigned due to the overlap of C_α protons of $^{2}L^{*}$ and $^{3}L^{*}$. The circled peaks are the interchain NOEs (see text). The intraresidue NH–$C_\alpha H$ cross peak for the third Ala* is missing in this spectrum, but it can be observed at a mixing time of 100 ms (chemical shift of the $C_\alpha H$ is 4.45 ppm).

Table III: Chemical Shifts of Assigned Proton Resonances of Peptide T3-785 at 10 °C

	chemical shift (ppm)[a]					
amino acid	NH	$C_\alpha H$	$C_\beta H$	$C_\gamma H$	$C_\delta H$	others
^{1}Arg	nd	4.56	1.71	nd	nd	nd
^{1}Gly*	8.26 (114.7)	3.69, 3.63				
^{1}Leu*	8.32 (125.7)	4.22	1.50	1.61	0.95	
^{1}Ala*	8.50 (129.9)	4.71	1.37			
^{2}Arg	nd	4.46	1.69	1.83	3.57	nd
^{2}Gly*	8.56 (116.0)	3.76, 3.60				
^{2}Leu*	8.74 (127.0)	4.17	1.57	1.71	0.96	
^{3}Arg	nd	4.46	1.69	nd	3.57	nd
^{3}Gly*	8.72 (115.3)	3.76, 3.66				
^{3}Leu*	8.57 (125.0)	4.17	1.57	1.69	0.95	
^{b}Ala*	8.36 (127.3)	4.45	1.36			
^{b}Ala*	8.37 (129.4)	4.70	1.35			
^{D}Arg	nd	4.34	1.81	1.69	3.23	nd
^{D}Gly*	8.49 (115.7)	3.90				
^{D}Leu*	8.26 (127.9)	4.37	1.66	1.66		
^{D}Ala*	8.45 (131.1)	4.38	1.43			

[a] In the NH column, the ^{15}N chemical shifts are listed in parentheses. ^{1}H chemical shifts are reported relative to external sodium 3-(trimethylsilyl)tetradeuteriopropionate (STP) and ^{15}N chemical shifts relative to external liquid NH_3. The nondetermined resonances are indicated with nd. The superscript numbers (1, 2, and 3) represent the three chains, and superscript D represents the denatured form of the peptide. The asterisks indicate the ^{15}N-labeled amino acids. [b] Chain number cannot be identified because of overlap.

backbone conformation. In addition to the sequential NOEs, we observe three nonsequential or interchain backbone NOEs in the HMQC-NOESY spectra (Figure 9). These NOEs could be (1) ^{1,2,3}Ala*$C_\alpha H$ to ^{3}Leu*NH, (2) ^{1,2,3}Ala*$C_\alpha H$ to ^{2}Leu*NH, or (3) ^{2,3}Arg$C_\alpha H$ to ^{1}Leu*NH or ^{2,3}Ala*$C_\alpha H$ to ^{1}Leu*NH. The numbers in the superscripts indicate that, because of chemical shift overlap, it is impossible to distinguish between the different chains on the basis of the spectra. To further examine the structure adopted by the three chains, the computer model of the design peptide in the triple-helical conformation was generated. From this model, interchain backbone NOEs are expected between ^{1}Leu*NH and ^{3}Arg$C_\alpha H$, ^{2}Leu*NH and ^{1}Ala*$C_\alpha H$, and ^{3}Leu*NH and ^{2}Ala*$C_\alpha H$. These NOEs are confirmed in a homonuclear NOESY experiment based on the chemical shifts obtained from the HMQC spectra. The expected NOEs from the model are consistent with the observed NOEs in the HMQC-NOESY and homonuclear NOESY experiments and indicate that chains 1, 2, and 3 are packed against each other and represent the three different chains in a trimer.

Examination of the NOE intensities in homonuclear and heteronuclear spectra reveals a decreased intensity of interchain versus intrachain NOEs for peaks that are predicted to have similar distances on the basis of the computer model. The NOE intensity between two atoms is a function of the internuclear distance and the correlation time of the internuclear vector. The rod shape of the triple-helical peptides (Figure 1) results in anisotropic rotational motion. For axially symmetric anisotropic tumbling, the effective correlation time experienced by a proton pair depends on the orientation (θ) of the internuclear vector relative to the major axis of the molecule (Woessner, 1962). Sequential and interchain interactions, which have different internuclear orientations relative to each other, will have different effective correlation times and NOE intensities (Withka et al., 1990). On the basis of the computer model of T3-785, typical interchain and intrachain internuclear vectors have θ values of 30° and 50°, respectively, and the calculated difference in NOE intensity is 1.5, assuming equal distances. If, in addition, the small differences in distance between interchain and intrachain proton pairs are taken into account, the intensity ratio of sequential to interchain proton NOEs is calculated to be 5:1,

assuming equal internal motion (M.-H. Li, P. Fan, & J. Baum, unpublished data).

The homonuclear spectra exhibit the expected ratio of intensities for interchain and intrachain NOEs. However, for the same peaks, the HMQC-NOESY spectra show an interchain NOE intensity that is 10 rather than 5 times weaker than the intrachain NOE, for reasons that are not clearly understood. This may account for the fact that many expected NOEs with proton distances larger than 3.0 Å are not observed in the HMQC-NOESY experiment, as the observed NOEs that are less than 3.0 Å are already quite weak. It is important to be cautious in the interpretation of the data as the lack of NOEs may indicate the presence of differences between the solution and solid-state conformations or may be the result of overlap and artifacts in the NMR spectra. However, comparison of the existing NOE intensities in the homonuclear spectrum with the proposed model suggests that chains 1, 2, and 3 in the design peptide are likely to be in a triple-helical conformation. No interchain NOEs are observed between tripeptide D and chains 1, 2, and 3. This result, in conjunction with the earlier observation that the chemical shifts of tripeptide D correspond to the monomer chemical shift, suggests that tripeptide D is probably the monomer form of the peptide. It cannot be ruled out that tripeptide D may correspond to a trimer form which is conformationally averaged in the central region; however, there is no evidence for any conformation other than the monomer or triple helix.

In the study of triple-helical peptides, we have encountered many difficulties in obtaining COSY and TOCSY types of spectra, a problem that often occurs with large molecular weight proteins that have large correlation times. On the basis of the molecular weights of the triple-helical peptides, we would not expect to encounter these problems. However, the rod shape of the triple-helical peptides results in anisotropic rotational motion with effective correlation times between 9 and 42 ns. Therefore, many of the problems encountered in studying large molecules with NMR occur with these triple-helical peptides. In particular, the T_2 values of the amide protons for the triple-helical resonances in the peptide have been estimated from the line widths and are approximately 8 ms. These short T_2's reduce the sensitivity of the resonances in TOCSY, COSY, and HMQC relayed experiments. The T_2's for the non-triple-helical residues (e.g., tripeptide 2 in $(POG)_{10}$ and tripeptide D in T3-785) are much longer (>100 ms) than those of the triple-helical residues, and this is probably the reason that the intensities, in 2D correlation spectra, of the non-triple-helical residues are always stronger than expected on the basis of the equilibrium constants. These large differences in intensity between triple-helical and non-triple-helical resonances have been observed in α1-CB2 as well (Torchia et al., 1975).

DISCUSSION

Homonuclear and heteronuclear NMR experiments are used to study the conformation, in solution, of two different triple-helical peptides. $(POG)_{10}$ is a peptide that contains the highest possible imino acid content (67%), and the peptide T3-785 is a 30-mer that contains an imino acid poor central region surrounded by POG caps (47% imino acids). The central region of the peptide is composed of 12 residues taken from near the collagenase cleavage site of type III collagen. A comparison of the spectral features of the two peptides leads to specific information about the triple helix in solution.

To look at the specific conformation adopted by the trimer $(POG)_{10}$ and T3-785 peptides, NOEs were predicted from the model derived from fiber diffraction data. For $(POG)_{10}$, it was found that the expected interchain backbone and side-chain NOEs were, in fact, seen. The experimental observation of these peaks indicates that the repeating structure of the $(POG)_{10}$ trimer has a triple-helical conformation in solution that is similar to that of the proposed model (Fraser et al., 1979). The ring puckering of the imino acid residues was determined in solution from the distance constraints and intensities and showed that proline (X position) has a puckered down conformation, while hydroxyproline (Y position) has puckered up conformation. These results are consistent with the fiber diffraction model (Fraser et al., 1979) and theoretical predictions (Miller & Scheraga, 1976; Miller et al., 1980). Interestingly, NMR studies on $(PPG)_{10}$ show that the proline in the X position is puckered down while the proline in the Y position is puckered up (M.-H. Li, & J. Baum, unpublished results), supporting the concept that it is the position of the imino acid in the tripeptide sequence which is the determinant of ring puckering.

In T3-785, the presence of NOESY peaks between Ala*, Arg, and Leu* residues of different chains indicates that the chains are packed against one another in the central region of the peptide. In the homonuclear NMR spectrum, the intensities of the observed interchain NOEs are consistent with the intensities expected from the computer model, indicating that the three central residues are probably in a triple-helical conformation. The NOESY peaks observed between chains 1, 2, and 3 are definitive proof of a one-residue stagger between the three parallel chains. As shown in Figure 10, a stagger by four residues would not be consistent with the observed results. The distance constraints also rule out an antiparallel organization of the three chains. Only three chains in parallel with a one-residue stagger would give distances between neighboring chains consistent with the observed peaks. The NMR data indicate that, for the great majority of the peptide, the central region of T3-785 is in a triple-helical conformation not very different from that of $(POG)_{10}$ and that the introduction of variations in the X-Y-Gly sequence allows specifics such as the precise one-residue stagger and the parallel nature of the chains to be unambiguously determined.

In addition to the resonances that can be assigned to the triple-helical conformation, other resonances are seen for both $(POG)_{10}$ and T3-785 which can be assigned to tripeptide units which, according to a number of criteria, are in a less ordered or non-triple-helical environment. For $(POG)_{10}$, these less ordered tripeptide units account for about 20% of the molecule. Since the central ^{15}N-enriched glycine shows only a triple-helix resonance and $(POG)_{10}$ is known to be fully trimeric, one can conclude that the less ordered regions are not due to small amounts of monomer of $(POG)_{10}$ and are likely to be localized to the ends of triple-helical molecules. It is not surprising to find such end effects in relatively short peptides, and end effects similar to those observed for $(POG)_{10}$ have also been seen for the small collagen peptide α1-CB2 (Torchia et al., 1975). By contrast, the T3-785 peptide shows conformational heterogeneity in the central region of the peptide. The Arg, Gly*, Leu*, and Ala* residues show nonhelical resonances, in addition to the three helical resonances, and these have been assigned to tripeptide D. The observation of such nonhelical resonances for the T3-785 peptide is consistent with the presence of a small amount of monomer. This could arise directly from the shift in equilibrium constant and decreased thermal stability relative to $(POG)_{10}$, which results from including the imino acid poor sequence.

The equivalence of all three chains in the central region of $(POG)_{10}$ contrasts with their nonequivalence in peptide T3-

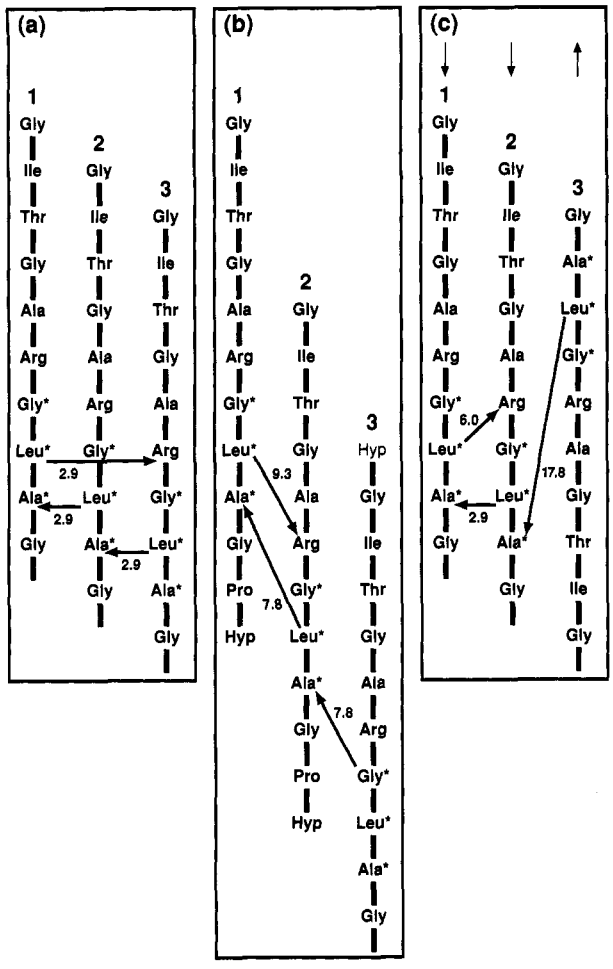

FIGURE 10: Possible chain stagger arrangements of T3-785 in a trimer: (a) parallel one-residue stagger; (b) parallel four-residue stagger; and (c) antiparallel one-residue stagger. The distances calculated between ^{1}Leu* NH and ^{3}Arg C_αH, ^{2}Leu* NH and ^{1}Ala* C_αH, and ^{3}Leu* NH and ^{2}Ala*C_αH are indicated for the three different models. Most of the distances in models b and c are too long to produce observable NOEs. Model a is most consistent with the observation of the three interchain NOEs.

785. At low temperature, when $(POG)_{10}$ is fully associated as a trimer, the 1D and 2D NMR spectra indicate that about 80% of the molecule consists of Pro-Hyp-Gly tripeptides that are in an equivalent environment. The $(POG)_{10}$ peptide with an ^{15}N-enriched glycine incorporated in the middle of the chain shows only one glycine peak at the chemical shift associated with the triple-helical position. Given the trimeric structure of the molecule, this reflects the equivalence of tripeptides in the three chains of one molecule, as well as the equivalence of tripeptide units in the same chain. In T3-785, the spectra of the central triple-helical Arg-Gly*-Leu*-Ala* residues show distinct differences from the central Gly in $(POG)_{10}$. Three cross peaks are seen for each residue in the T3-785 triple helix as compared to one peak for $(POG)_{10}$. Thus, while the three glycine residues in the center of $(POG)_{10}$ are in an equivalent environment, the glycine residues in the Gly*-Leu*-Ala* sequences are not equivalent for the three chains in a trimer, and the chains can be assigned and followed through the labeled unit. Examination of the triple-helix model of T3-785 (Figure 10a) shows that at any given level, two of the chains have the same type of neighboring residue and one is different. However, if the adjacent levels are considered, the same residue in all three chains has different neighbors, which could give rise to the three distinct triple-helical resonances for Gly*, Leu*, and Ala*. The precise equivalence of the three chains is only true when there are repetitive tripeptides, such as in $(POG)_{10}$, and is only an approximation in other cases, such as triple helices in collagen or other molecules. The nonequivalence of the same residue in the three chains of a homotrimer may have implications for the association of collagen molecules into larger aggregates in their functional state or the specific binding sites on collagens.

In collagen, some regions of low imino acid content appear to be involved in specific recognition, e.g., the collagenase cleavage site and the sites of hydroxylysine formation and cross-link formation (Fietzek & Kuhn, 1975). Studies on polypeptides and small peptides indicate that the imino acid content of a tripeptide influences the stability it confers on the molecule, with a general order of POG > XOG > PYG > XYG (Fietzek & Kuhn, 1975). This general scheme of stability is consistent with our studies, as the replacement of $(POG)_3$ triplets by ITGARGLAG markedly decreases the thermal stability of the peptide. It has been suggested that regions of low imino acid content could be more flexible and that they might even loop out and form a conformation alternative to the standard triple helix (Fields, 1991). The NMR studies presented here show that the backbone of $(XYG)_3$ in the center of the trimer formed by designed peptide T3-785 adopts just one well-defined conformation, where the three chains are in close contact and have a one-residue stagger with respect to each other. The distances seen are compatible with a standard triple-helix model. These results make it unlikely that imino acid poor regions such as that included in this peptide must always adopt an alternative conformation. However, the small number of NMR distance constraints observed only test the model at low resolution and do not preclude the possibility that there are small structural backbone variations in the imino acid poor region compared with $(POG)_{10}$. Variations in the degree of supercoiling or in the specifics of the hydrogen-bonding pattern (e.g., the one-bond versus the two-bond model) could be present in solution. On the basis of these observations, the recognition of binding sites along a collagen molecule could be entirely due to the specificity of side-chain interactions, with a constant backbone conformation, or due to subtle backbone changes which facilitate such interactions, but the presence of a defined alternative conformation as a basis for recognition (e.g., β-blend or loop) is unlikely in this case.

ACKNOWLEDGMENT

We are grateful to Dr. Chaun Wang and Dr. Cynthia G. Long for helpful discussions. We thank Dr. Ken Breslauer for the use of the CD instrument in his spectroscopy facility and Dr. Catherine Stolle for participating in the design of peptide T3-785.

REFERENCES

Ala-Kokko, L., Kontusaari, L. S., Baldwin, C., Kuivaniemi, H., & Prockop, D. J. P. (1929) *Biochem. J. 260*, 509–516.
Aue, W. P., Bartholdi, E., & Ernst, R. R. (1976) *J. Chem. Phys. 64*, 2229–2246.
Bax, A., & Freeman, R. (1981) *J. Magn. Reson. 44*, 542–561.
Bax, A., & Davis, D. (1985) *J. Magn. Reson. 65*, 355–360.
Bax, A., & Subramanian, S. (1986) *J. Magn. Reson. 67*, 565–569.
Bax, A., Griffey, R. H., & Hawkins, B. L. (1983) *J. Magn. Reson. 55*, 301–315.
Benson, B., Hawgood, S., Schilling, J., Clements, J., Damm, D., Cordell, B., & White, T. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 6379–6383.
Brahmachari, S. K., & Ananthanarayanan, V. S. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 5119–5123.

Braunsweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.
Brodsky, B., Li, M.-H., Long, C. G., Apigo, J., & Baum, J. (1992) *Biopolymers 32*, 447–451.
Brodsky-Doyle, B., Leonard, K. R., & Reid, K. B. (1976) *Biochem. J. 159*, 279–286.
Clore, G. M., & Gronenborn, A. M. (1987) *Protein Eng. 1*, 275–288.
Clore, G. M., & Gronenborn, A. M. (1991) *Science 252*, 1390–1399.
Clore, G. M., Bax, A., Wingfield, P., & Gronenborn, A. M. (1988) *FEBS Lett. 238*, 17–21.
Dedhar, S., Ruoslahti, E., & Pierschbacher, M. D. (1987) *J. Cell Biol. 104*, 585–593.
Di Blasi, R., & Verdini, A. S. (1979) *Biopolymers 18*, 735–738.
Dolz, R., & Heidemann, E. (1986) *Biopolymers 25*, 1069–1080.
Drickamer, K., Dordal, M. S., & Reynolds, L. (1986) *J. Biol. Chem. 261*, 6878–6887.
Engel, J., Chen, H., Prockop, D. J., & Klump, H. (1977) *Biopolymers 16*, 601–622.
Fields, G. B. (1991) *J. Theor. Biol. 153*, 585–602.
Fietzek, P. P., & Kuhn, K. (1975) *Mol. Cell. Biochem. 8*, 141–157.
Fraser, R. D. B., & MacRae, T. P. (1973) *Conformation in Fibrous Proteins*, Academic Press, New York.
Fraser, R. D. B., MacRae, T. P., & Suzuki, E. (1979) *J. Mol. Biol. 129*, 463–481.
Garbay-Jaureguiberry, C., Arnoux, B., Prange, T., Wehri-Altenburger, S., Pascard, C., & Roques, B. P. (1980) *J. Am. Chem. Soc. 102*, 1827–1837.
Gronenborn, A. M., Bax, A., Wingfield, P., & Clore, G. M. (1989) *FEBS Lett. 243*, 93–98.
Guidry, C., Miller, E. J., & Hook, M. (1990) *J. Biol. Chem. 265*, 19230–19236.
Heidemann, E., & Roth, W. (1982) *Adv. Polym. Sci. 43*, 143–203.
Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.
Jelinski, L. W., & Torchia, D. A. (1979) *J. Mol. Biol. 133*, 45–65.
Kirchhofer, D., Languino, L. R., Rouslahti, E., & Pierschbacher, M. D. (1990) *J. Biol. Chem. 265*, 615–618.
Kitchen, D. B., Hirata, F., Westbrook, J. D., Levy, R. M., Kofke, D., & Yarmush, M. (1990) *J. Comput. Chem. 11*, 1169–1180.
Kobayashi, Y., & Kyogoku, Y. (1973) *J. Mol. Biol. 81*, 337–347.
Kobayashi, Y., Sakai, R., Kabiuchi, K., & Isemura, P. (1970) *Biopolymers 9*, 415–425.
Kodama, T., Freeman, M., Rohrer, L., Zabrecky, P. M., & Krieger, M. (1990) *Nature 343*, 531–535.
Kuivaniemi, H., Tromp, G., & Prockop, D. J. (1991) *FASEB J. 5*, 2052–2060.
Kumar, A., Ernst, R. R., & Wuthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1–6.
Linsenmayer, T. (1991) in *Cell Biology of Extracellular Matrix* (Hay, E. D., ed.), pp 7–44, Plenum Press, New York.
Long, C. G. (1992) Ph.D. thesis, Rutgers University, Piscataway, NJ.
Long, C. G., Li, M. H., Baum, J., & Brodsky, B. (1992) *J. Mol. Biol. 225*, 1–4.
Marion, D., & Wuthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Mayo, K. H., Parra-Diaz, D., Chelberg, M., & MaCarthy, J. (1991) *Biochemistry 30*, 8251–8267.
Miller, E. J. (1984) in *Extracellular Matrix Biochemistry* (Piez, K. A., & Reddi, A. H., ed.) pp 41–82, Elsevier, New York.
Miller, M. H., & Scheraga, H. A. (1976) *J. Polym. Sci. 54*, 171–200.
Miller, M. H., Nemethy, G., & Scheraga, H. A. (1980) *Macromolecules 13*, 470–478.
Mueller, L. (1979) *J. Am. Chem. Soc. 101*, 4481–4484.
Nemethy, G., & Scheraga, H. A. (1984) *Biopolymers 23*, 2781–2799.
Otting, G., Widmer, H., Wanger, G., & Wuthrich, K. (1986) *J. Magn. Reson. 66*, 187–193.
Ramachandran, G. N. (1967) in *Treatise on Collagen* (Ramachandran, G. N., ed.) Vol. 1, pp 103–183, Academic Press, New York.
Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wuthrich, K. (1984) *Biochem. Biophys. Res. Commun. 117*, 479–485.
Redfield, A. G., & Kuntz, S. D. (1975) *J. Magn. Reson. 19*, 250–254.
Rich, A., & Crick, F. H. C. (1961) *J. Mol. Biol. 3*, 483–506.
Roques, B. P., Garbay-Jaureguiberry, C., Combrisson, S., & Oberlin, R. (1977) *Biopolymers 16*, 937–944.
Sakikabara, S., Inouye, K., Shudo, K., Kishida, Y., Kobayashi, Y., & Prockop, D. J. (1973) *Biochim. Biophys. Acta 303*, 198–202.
Sarkar, S. K., Young, P. E., Sullivan, C. E., & Torchia, D. A. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 4800–4803.
Sarkar, S. K., Hiyama, Y., Niu, C. H., Young, P. E., Gerig, J. T., & Torchia, D. A. (1987) *Biochemistry 26*, 6793–6800.
Schumacher, M., Camp, S., Maulet, Y., Newton, M., MacPhee-Quigley, K., Taylor, S. S., Friedmann, T., & Taylor, P. (1986) *Nature 319*, 407–409.
Seyer, J. M., Mainardi, C., & Kang, A. H. (1980) *Biochemistry 19*, 1583–1589.
Shaka, B., & Freeman, R. (1983) *J. Magn. Reson. 51*, 169–173.
Staatz, W. D., Fok, K. F., Zutter, M. M., Adams, S. P., Rodriguez, B. A., & Santoro, S. A. (1991) *J. Biol. Chem. 266*, 7363–7367.
Torchia, D. A. (1971) *Macromolecules 4*,, 440–442.
Torchia, D. A. (1972) *Macromolecules 5*, 566–569.
Torchia, D. A. (1982) *Mehods Enzymol.* (Struct. Contract. Proteins, Pt. A), *82*, 174–186.
Torchia, D. A., & Lyerla, J. R., Jr. (1974) *Biopolymers 13*, 97–114.
Torchia, D. A., Lyerla, J. R., Jr., & Quattrone, A. J. (1975) *Biochemistry 14*, 887–900.
Torchia, D. A., Hiyama, Y., Sarkar, S. K., Sullivan, C. E., & Young, P. E. (1985) *Biopolymers 24*, 65–75.
Tromp, G., Kuivaniemi, H., Shikata, H., & Prockop, D. J. (1989) *J. Biol. Chem. 264*, 1349–1352.
van der Rest, M., & Garrone, R. (1991) *FASEB J. 5*, 2814–2823.
von Hippel, P. H. (1967) in *Treatise on Collagen* (Ramachandran, G. N., Ed.) pp 253–338, Academic Press, New York.
Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N., & Wuthrich, K. (1987) *J. Mol. Biol. 196*, 611–639.
Wider, G., Macura, S., Kumar, A., Ernst, R. R., & Wuthrich, K. (1984) *J. Magn. Reson. 56*, 207–234.
Withka, J. M., Swaminathan, S., & Bolton, P. H. (1990) *J. Magn. Reson. 89*, 386–390.
Woessner, D. E. (1962) *J. Chem. Phys. 37*, 647–654.
Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wuthrich, K. (1989) *Acc. Chem. Res. 22*, 36–44.