J|A|C|S
ARTICLES
Published on Web 06/06/2007

FTIR Studies of Collagen Model Peptides: Complementary Experimental and Simulation Approaches to Conformation and Unfolding

Michael A. Bryan,[†] Joseph W. Brauner,[§] Gloria Anderle,[‡] Carol R. Flach,[§] Barbara Brodsky,[†] and Richard Mendelsohn*,[§]

Contribution from the Department of Biochemistry, University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School, Piscataway, New Jersey, Department of Chemistry and Geology, Becton College, Fairleigh Dickinson University, Madison, New Jersey, and Department of Chemistry, Newark College of Arts and Science, Rutgers University, New Jersey

Received February 16, 2007; E-mail: mendelso@andromeda.rutgers.edu

Abstract: X-ray crystallography of collagen model peptides has provided high-resolution structures of the basic triple-helical conformation and its water-mediated hydration network. Vibrational spectroscopy provides a useful bridge for transferring the structural information from X-ray diffraction to collagen in its native environment. The vibrational mode most useful for this purpose is the amide I mode (mostly peptide bond C═O stretch) near 1650 cm^{-1}. The current study refines and extends the range of utility of a novel simulation method that accurately predicts the infrared (IR) amide I spectral contour from the three-dimensional structure of a protein or peptide. The approach is demonstrated through accurate simulation of the experimental amide I contour in solution for both a standard triple helix, $(Pro\text{-}Pro\text{-}Gly)_{10}$, and a second peptide with a Gly → Ala substitution in the middle of the chain that models the effect of a mutation in the native collagen sequence. Monitoring the major amide I peak as a function of temperature gives sharp thermal transitions for both peptides, similar to those obtained by circular dichroism spectroscopy, and the Fourier transform infrared (FTIR) spectra of the unfolded states were compared with polyproline II. The simulation studies were extended to model early stages of thermal denaturation of $(Pro\text{-}Pro\text{-}Gly)_{10}$. Dihedral angle changes suggested by molecular dynamics simulations were made in a stepwise fashion to generate peptide unwinding from each end, which emulates the effect of increasing temperature. Simulated bands from these new structures were then compared to the experimental bands obtained as temperature was increased. The similarity between the simulated and experimental IR spectra lends credence to the simulation method and paves the way for a variety of applications.

1. Introduction

The study of collagen by a wide range of physicochemical methods has provided insights into structure, self-association, and ligand binding, all of which are essential for its physiological functions.[1] Collagen forms the structural backbone of many tissues including bone, tendon, skin, blood vessels, and basement membranes. In addition to its structural roles, collagen binds a wide variety of ligands, including other extracellular matrix proteins, proteoglycans, and cell surface receptors. Collagen has been studied at many levels of structural organization and at many distance scales, from direct observation in tissues to detailed determination of its molecular structure and dynamics.

Infrared (IR) spectroscopy has been widely used to monitor elements of protein structure in films and solution. In addition, technological advances have permitted FTIR data to be acquired with a spatial resolution of ~10 micrometers from small tissue sections. Recently, FTIR microscopic data for collagen organization in bone and skin have been reported and have been used to characterize a pathological state of bone.[2] Thus, vibrational spectroscopy provides a bridge between the atomic level information acquired about collagen structure from X-ray diffraction and the examination of collagen in its native environment.

A fundamental understanding of the origin of the collagen vibrational modes and their sensitivity to altered structure and organization will aid in the interpretation of vibrational spectra of tissues. The collagen triple helix is a unique protein motif defined by the supercoiling of three polypeptide chains in a polyproline II conformation.[3] This conformation confers strict

[†] Robert Wood Johnson Medical School.
[‡] Fairleigh Dickinson University.
[§] Rutgers University.

(1) (a) Brodsky, B.; Persikov, A. V. *Adv. Protein Chem.* **2005**, *70*, 301−339. (b) Kielty, C. M.; Grant, M. E. In *Connective Tissue and Its Hereditable Disorders*; Royce, P. M., Steinmann, B., Eds.; Wiley-Liss: New York, 2002; pp 159−221.
(2) (a) Paschalis, E. P.; Verdelis, K.; Doty, S. B.; Boskey, A. L.; Mendelsohn, R.; Yamauchi, M. *J. Bone Miner. Res.* **2001**, *16*, 1821−1828. (b) Paschalis, E. P.; Shane, E.; Lyritis, G.; Skarantavos, G.; Mendelsohn, R.; Boskey, A. L. *J. Bone Miner. Res.* **2004**, *19*, 2000−2004.
(3) (a) Ramachandran, G. N.; Kartha, G. *Nature* **1955**, *176*, 593−595. (b) Rich, A.; Crick, F. H. *J. Mol. Biol.* **1961**, *3*, 483−506. (c) Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M. *Science* **1994**, *266*, 75−81.

10.1021/ja071154i CCC: $37.00

amino acid sequence constraints requiring Gly as every third residue, resulting in a $(Gly\text{-}XAA\text{-}YAA)_n$ sequence. The exclusive preservation of Gly at every third position in the amino acid sequence is required for close packing of the three chains since Gly residues are buried at the center of the triple helix. This position can accommodate no larger residue; thus, mutations which replace one Gly by another residue in collagens lead to disease, for example, osteogenesis imperfecta.[4] In contrast to the Gly position, the X and Y positions are exposed on the surface and can sterically accommodate any residue.[5] The X and Y positions are frequently occupied by Pro and its post-translationally modified form hydroxyproline (Hyp) which serve to stabilize the triple helix by promoting the polyproline II conformation. The triple helix is stabilized by intramolecular hydrogen bonds, with one direct hydrogen bond between the Gly N−H and the carbonyl of the X residue in an adjacent chain. Triplets of the form Pro-Hyp-Gly are the most common and the most stabilizing tripeptide sequence in collagens.

Peptides have proven to be useful models for elucidating the principles of collagen structure and organization.[1a,6] Crystallographic studies of collagen-like peptides have confirmed the basic triple-helical conformation and have provided detailed information on hydration, side chain interactions, and packing geometry. The crystal structures of $(Pro\text{-}Pro\text{-}Gly)_{10}$, $(Pro\text{-}Hyp\text{-}Gly)_{10}$, and $(Pro\text{-}Hyp\text{-}Gly)_4$-Glu-Lys-Gly-$(Pro\text{-}Hyp\text{-}Gly)_5$ all show similar triple-helical conformations with 7-fold symmetry, a characteristic pattern of backbone hydrogen bonding, and a significant level of interaction with water.[7] A high-resolution structure was also determined for a peptide designed to model a collagen mutation, $(Pro\text{-}Hyp\text{-}Gly)_4$-Pro-Hyp-Ala-$(Pro\text{-}Hyp\text{-}Gly)_5$ [denoted as the Gly → Ala peptide], which has a decreased stability compared to $(Pro\text{-}Hyp\text{-}Gly)_{10}$.[3c,8] The Pro-Hyp-Gly ends form standard 7/2 triple helices while there is local unwinding, and direct peptide hydrogen bonds are replaced by water-mediated hydrogen bonding at the Ala substitution site.[3c]

The availability of high-resolution structures for collagen model peptides provides atomic coordinates needed (see below) for simulations of the amide I contour in the IR spectra. This vibrational mode occurs between 1600 and 1700 cm^{-1} and arises largely from peptide bond C═O stretch. The mode has been used extensively in analysis of protein secondary structures. In previous IR studies of collagen and collagen model peptides, it was proposed that the observed three major amide I bands are derived from the three nonequivalent amide C═O groups in each of the Gly-X-Y units.[9] The hydration of these groups has also been considered as a contributing factor to the spectra.[10]

Recent studies[11] reported a novel simulation method that permits accurate prediction of the amide I spectral contour from the three-dimensional structure of a protein. The method, which incorporates through space transition dipole coupling (TDC) along with through valence bond and through hydrogen bond coupling, simulates the important interactions between the peptide group (C═O) oscillators that define the amide I mode. In the current work, the TDC and H-bond interactions are both inter- and intrachain. For the current set of peptides, interchain hydrogen bonds and transition dipole coupling are explicitly included in the model. In previous work, the amide I contours of six globular proteins with a variety of secondary structure motifs were accurately simulated. Fair agreement was also seen between the simulated and the experimental amide I spectra of the collagen model peptide $(Pro\text{-}Pro\text{-}Gly)_{10}$.

The prior preliminary work[11b] on $(Pro\text{-}Pro\text{-}Gly)_{10}$ is refined here to improve the fit and to elaborate on the assignment of its amide I features. With an improved fit in hand, simulations were then done on the early stages of unfolding of $(Pro\text{-}Pro\text{-}Gly)_{10}$. To confirm the parameter variations made in improving the fit of the amide I contour, a similar model peptide with a Gly to Ala substitution in the middle of the chain was also simulated. This variant is of intrinsic interest as a model for a mutant form of collagen.

2. Experimental Section

2.1. Peptides. The synthetic collagen model peptide $(Pro\text{-}Pro\text{-}Gly)_{10}$ was purchased from Peptides International (Louisville, KY). This peptide is denoted as $(PPG)_{10}$. The synthetic collagen model peptide $(Pro\text{-}Hyp\text{-}Gly)_4$-Pro-Hyp-Ala-$(Pro\text{-}Hyp\text{-}Gly)_5$, containing one Gly to Ala replacement in a $(Pro\text{-}Hyp\text{-}Gly)_{10}$ context, was purchased from Tufts University Core Facility (Boston, MA). This peptide is denoted as the Gly → Ala peptide. Poly-L-proline was purchased from Sigma-Aldrich (St. Louis, MO). This peptide lot contained an average molecular weight of 6400 g/mol by viscosity. Deuterated water (D_2O) and deuterated hydrochloric acid (DCl), both at 99.9% isotopic purity, were purchased from Sigma-Aldrich Corporation. Peptides were purified using a reverse-phase high-performance liquid chromatography (HPLC) system (Shimadzu) on a C-18 column and were eluted in 0.1% trifluoroacetic acid with a binary gradient of 0−40% (v/v) water/acetonitrile. The purity was confirmed by laser desorption mass spectrometry. The concentration was determined by monitoring absorbance at 214 nm using $\epsilon^{214} = 2200\ M^{-1}\ cm^{-1}$ per peptide bond. Peptide samples were dissolved in DCl/D_2O, pD = 3, at 5 mg/mL (2 mM) for both the FTIR and circular dichroism (CD) measurements

2.2. FTIR Spectroscopy. Spectra were collected on a Mattson Instruments Research Series spectrometer equipped with a sample shuttle and a mercury-cadmium-telluride detector. Sample solutions (5 mg/mL; total sample volume of ~35 μL) were placed between CaF_2 windows separated with a 25 μm spacer. Spectra were acquired over the temperature range 12−60 °C by coadding four blocks of 256 scans over the mid-infrared region. To ensure that $(Pro\text{-}Pro\text{-}Gly)_{10}$ formed a completely native trimer at 12 °C, an IR spectrum was also obtained at 5 °C. The profiles of the two low-temperature scans were identical, validating 12 °C to be an adequate native starting point for thermal unfolding studies. The conditions for thermal perturbation used temperature steps of 1.5 °C with a 3 min instrument equilibration time and a 5 min incubation period at each temperature, along with a 7 min data collection time. This procedure gives an average heating rate of 0.1 °C/min. Spectra were acquired at 4 cm^{-1} resolution, were apodized

(4) Byers, P. H.; Cole, W. G. In *Connective Tissue and Its Hereditable Disorders*; Royce, P. M., Steinmann, B., Eds.; Wiley-Liss: New York, 2002; pp 385−430.
(5) Persikov, A. V.; Ramshaw, J. A.; Kirkpatrick, A.; Brodsky, B. *Biochemistry* **2000**, *39*, 14960−14967.
(6) (a) Baum, J.; Brodsky, B. *Curr. Opin. Struct. Biol.* **1999**, *9*, 122−128. (b) Jenkins, C. L.; Raines, R. T. *Nat. Prod. Rep.* **2002**, *19*, 49−59.
(7) (a) Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari, A. *Protein Sci.* **2002**, *11*, 262−270. (b) Kramer, R. Z.; Venugopal, M. G.; Bella, J.; Mayville, P.; Brodsky, B.; Berman, H. M. *J. Mol. Biol.* **2000**, *301*, 1191−1205. (c) Nagarajan, V.; Kamitori, S.; Okuyama, K. *J. Biochem. (Tokyo)* **1998**, *124*, 1117−1123. (d) Nagarajan, V.; Kamitori, S.; Okuyama, K. *J. Biochem. (Tokyo)* **1999**, *125*, 310−318.
(8) Long, C. G.; Braswell, E.; Zhu, D.; Apigo, J.; Baum, J.; Brodsky, B. *Biochemistry* **1993**, *32*, 11688−11695.
(9) (a) George, A.; Veis, A. *Biochemistry* **1991**, *30*, 2372−2377. (b) Jakobsen, R. J.; Brown, L. L.; Hutson, T. B.; Fink, D. J.; Veis, A. *Science* **1983**, *220*, 1288−1290. (c) Payne, K. J.; Veis, A. *Biopolymers* **1988**, *27*, 1749−1760. (d) Lazarev, Y. A.; Grishkovsky, B. A.; Khromova, T. B. *Biopolymers* **1985**, *24*, 1449−1478.
(10) Susi, H.; Ard, J. S.; Carroll, R. J. *Biopolymers* **1971**, *10*, 1597−1604.
(11) (a) Brauner, J. W.; Dugan, C.; Mendelsohn, R. *J. Am. Chem. Soc.* **2000**, *122*, 677−683. (b) Brauner, J. W.; Flach, C. R.; Mendelsohn, R. *J. Am. Chem. Soc.* **2005**, *127*, 100−109.

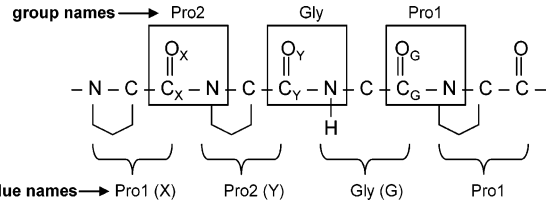

Figure 1. Collagen peptide schematic showing the nomenclature used in the current work. The residue sequence is Pro1 (X), Pro 2 (Y), and Gly (G). The group name comes from the residue providing the N in the peptide bond.

with a triangular function, and were Fourier transformed with one level of zero-filling to yield spectra encoded every ~2 cm^{-1}. Using Grams/32 software (Galactic industries), a spectrum of D_2O/DCl, acquired at the same temperature as the sample spectrum, was subtracted from each peptide spectrum. The fraction folded $F(T)$ at each temperature was calculated from FTIR melting curves as

$$F(T) = (\theta_{obs} - \theta_M)/(\theta_N - \theta_M)$$

where θ_{obs} is the observed intensity, which changes as a function of temperature. θ_N and θ_M are the uppermost and lowermost intensities representative of the native and the monomer states respectively. The melting temperature T_m was obtained as a midpoint of the transition, that is, $F(T_m) = 0.5$.

2.3. CD Spectroscopy. CD spectra were recorded on an AVIV Model 62DS Spectrophotometer (Aviv Biomedical, Inc.) regulated by a Hewlett-Packard Peltier temperature controller. The same sample concentrations and solvents were used as described in FTIR spectroscopy. Prior to all melting experiments, collagen model peptides were incubated at 0 °C for 48 h. The conditions for thermal melting includes a temperature step of 0.3 °C with 2 min equilibration time at each temperature and 10 s data collection time at 225 nm. This procedure gives an average heating rate of 0.1 °C/min, and the T_m was obtained as a midpoint of the transition, that is, $F(T_m) = 0.5$ as described in a previous paper.[12]

2.4. Simulation of the Low-Temperature Amide I Bands of Two Collagen Model Peptides, $(PPG)_{10}$ and Gly → Ala. A labeled schematic diagram of the peptide structure of $(PPG)_{10}$ is shown in Figure 1 to clarify the two nomenclatures used here. The novel IR simulations focus on peptide groups while earlier analyses of spectra focused on the carbonyl group of a particular residue.

The first step in improving the amide I simulation is to start with a realistic structure. Therefore, NCCO atomic coordinates of $(PPG)_{10}$ were obtained from a 1 ns molecular mechanics simulation in explicit water at 16 °C starting from the X-ray crystal structure 1K6F in the Protein Data Bank.

In our original work,[11b] identical parameters were assumed for the two different peptide groups whose nitrogen atom comes from the first and second proline residues in the sequence.

In the simulation focused nomenclature, the peptide group name is derived from the residue supplying the nitrogen atom. To improve the simulation, the two different proline-based groups had to be distinguished from each other. It was found necessary to set the inverse mass factor for Pro1 groups to 0.986 while that for Pro2 groups is set to 0.963. The transition dipole strength of the Pro1 groups is set to 400 $cm^{3/2}/s$ (0.40 $m^{3/2}/s$) as before but for the Pro2 groups 300 $cm^{3/2}/s$ (0.30 $m^{3/2}/s$) is used. This change may reflect the effect of the hydrogen bond between the carbonyl C_XO_X and the amide hydrogen of the Gly groups in another chain. The bond may decrease the electron density flux in the peptide group during the group vibration. The transition dipole strength of the carbonyl groups of the peptide bond preceding the Gly (C_YO_Y) was decreased to 220 $cm^{3/2}/s$ (0.22 $m^{3/2}/s$) from the value of 300 $cm^{3/2}/s$ (0.30 $m^{3/2}/s$) used previously. The hydrogen atom on the α carbon does not supply as much electron density to the group as does a proline ring. The amide I contour is the total of the summed contributions from the three kinds of peptide groups, each again designated by the residue supplying the nitrogen atom (Gly, Pro1, or Pro2). The width factor for the Gly group sub-bands was increased to 9 cm^{-1} from 7 cm^{-1}, the value used for the Pro1 and Pro2 group sub-bands.

To test the validity of the parameters chosen for $(PPG)_{10}$, a simulation was also carried out on the Gly → Ala peptide, for which experimental amide I contours were also acquired. The NCCO atomic coordinates of the Gly → Ala peptide were as before obtained from a 1 ns molecular mechanics run in explicit water at 16 °C starting from the X-ray crystal structure 1CGD from the PDB.

The parameters should be similar for the two molecules. Three small changes were found necessary. The inverse mass factor for the Pro1 groups was set to 0.976, the transition dipole strength of the Gly groups was set to 250 $cm^{3/2}/s$ (0.25 $m^{3/2}/s$), and the width factor for all three sub-bands is 7 cm^{-1}. The Ala group transition dipole strength was set to 400 $cm^{3/2}/s$ (0.40 $m^{3/2}/s$).

The only justification for the new parameters is the improvement in the agreement between the simulated and experimental spectra. On the basis of our experience, it is very unlikely that wrong parameters would give a good fit along the entire amide I contour. It is impossible to tweak one isolated part of the contour with a parameter value change. Any change affects the whole contour.

2.5. Molecular Mechanics Simulations. All computations were performed on a Linux workstation. The PDB templates used for the initial structure were 1K6F for $(PPG)_{10}$ and 1CGD for the Gly → Ala peptide. Both 1K6F and 1CGD were modified in SwissPro Deep View[13] by removing solvent molecules. 1K6F was further modified by the removal of chains D, E, and F. The models were each solvated in LEAP in the Amber 8 package using a periodic box of TIP3P water with a box dimension of 9 Å from the protein surface, and the forcefield *ff03* was applied.[14] The models were minimized at 289 K and 1 bar for 600 cycles of Steepest Descent and 200 cycles of Conjugate Gradient with a light restraint of 4 kcal/mol on all residues. Long-range electrostatics were modeled using the particle-mesh Ewald method.[15] The bonds to hydrogen atoms in the model were constrained using the Shake constraint; the temperature-coupling algorithm used was Langevin; the pressure-coupling algorithm used was the Berendsen algorithm.[16] Additional minimization was done on each model for 25 ps with end residues constrained with a force of 10 kcals/mol to prevent unraveling of the triple helix using conditions cited above. Finally, each model was subjected to 1 ns of unrestrained simulation at 289 K and 1 bar. The protein backbone atom coordinates from the resulting trajectory were extracted at 1 ps intervals and were averaged.

2.6. Simulations of the Early Stages of Unfolding. A molecular mechanics simulation of $(PPG)_{10}$ in explicit water at 27 °C provided the dihedral angles of the peptide at this higher temperature. The dihedral angles at 16 °C were calculated from the starting structure. This gives a set of dihedral angle changes.

With the improved low temperature $(PPG)_{10}$ amide I simulation in hand, we modified the 16 °C molecular mechanics structure to simulate

(12) Persikov, A. V.; Xu, Y.; Brodsky, B. *Protein Sci.* **2004**, *13*, 893−902.

(13) Guex, N.; Peitsch, M. C. *Electrophoresis* **1997**, *18*, 2714−2723.

(14) Duan, Y.; Wu, C; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. *J. Comput. Chem.* **2003**, *24*, 1999−2012.

(15) (a) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092. (b) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. *J. Chem. Phys.* **1995**, *103*, 8577−8592.

(16) (a) Izaguirre, J.; Catarello, D.; Wozniak, J.; Skeel, R. *J. Chem. Phys.* **2001**, *114*, 2090−2098. (b) Loncharich, R.; Brooks, B. R.; Pastor, R. *Biopolymers* **1992**, *32*, 523−535. (c) Pastor, R.; Brooks, B. R.; Szabo, A. *Mol. Phys.* **1988**, *65*, 1409−1419. (d) Berendsen, H. J. C.; Postma, J. P. M.; vanGunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3584−3590.

(17) (a) Loncharich, R. J.; Brooks, B. R.; Pastor, R. W. *Biopolymers* **1992**, *32*, 523−535. (b) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph.* **1996**, *14*, 33−38.

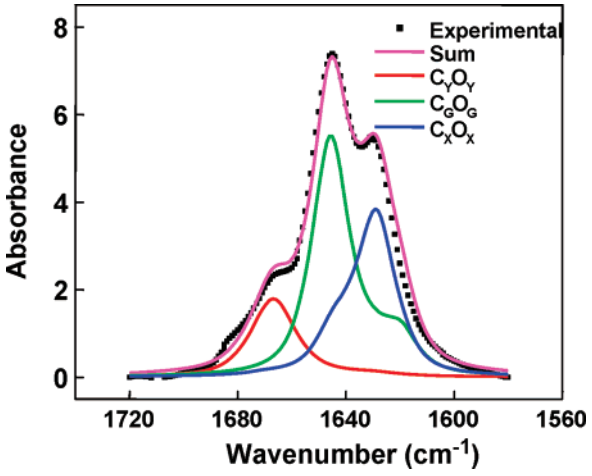

Figure 2. The experimental (black dots) and simulated (magenta line) amide I contour for the peptide $(PPG)_{10}$ in D_2O solution in its triple-helical state. The total simulation is broken into the contributions of the peptide bond preceding residue Gly (C_YO_Y-red line), Pro1 (C_GO_G-green line), and Pro2 (C_XO_X-blue) sub-bands. Notation is as described in Figure 1.

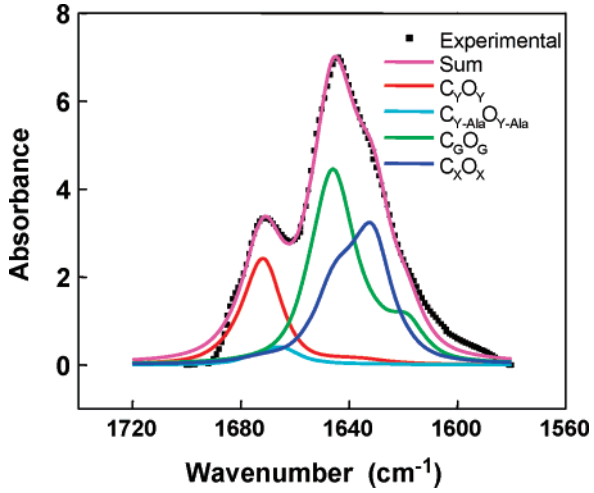

Figure 3. The experimental (black dots) and simulated (magenta line) amide I contour for the Gly → Ala peptide in the triple-helical state. Simulation is broken into contributions of the peptide bonds preceding Gly (C_YO_Y-red line), Ala ($C_{Y-Ala}O_{Y-Ala}$-light blue), Pro1 (C_GO_G-green line), and Pro2 (C_XO_X-blue) sub-bands, using notation analogous to that in Figure 1.

the early stages of thermal denaturation by changing the dihedral angles at the α carbon atoms in each chain. The changes were made in steps of one α carbon atom at a time starting with the outermost α carbon atoms at each end and working in toward the middle of each chain simultaneously.

For example, changing the phi angle at an α carbon atom near the N-terminus involves a rotation around an N−C bond of all the off-axis atoms going out toward the N-terminus from the rotation axis. A change in the phi angle at an α carbon atom near the C-terminus involves a rotation of all the off-axis atoms going out to the C-terminus from the axis of rotation.

The dihedral angle changes as calculated were used in going out to the C-terminus. The negatives of these changes were used in going back to the N-terminus so as to produce an unraveling of the triple helix at both ends. Since this procedure is arbitrary, some of the structures so generated had exaggerated interaction energies (presumably because of unrealistic van der Waals' contacts) and therefore gave rise to anomalous band shapes. Fortunately, most of the modified structures were well behaved and a sequence of contours representing 10 steps was generated that closely matches the experimental series of contours up to about 36 °C. Three of the steps were ignored for the reasons just mentioned. At these early stages of the unfolding, visualization of the structure indicated that the triple helix remains intact at the center but is unraveled at each end. The generally good agreement (see below) between the simulated and experimental bands suggests that this proposed mechanism is reasonable for the early stages of the thermal denaturation of these collagen model peptides.

3. Results

3.1. FTIR Experimental and Simulation Studies of Native Triple-Helical States of the (Pro-Pro-Gly)$_{10}$ and Gly → Ala Peptides. The $(Pro\text{-}Pro\text{-}Gly)_{10}$ peptide contains Gly as every third residue while the Gly → Ala peptide contains a Gly to Ala substitution in a $(Pro\text{-}Hyp\text{-}Gly)_{10}$ context, which perturbs the triple-helical structure.[3c,8] The amide I region (1600−1700 cm^{-1}) of the FTIR solution spectrum of $(Pro\text{-}Pro\text{-}Gly)_{10}$ at 16 °C, where the collagen peptide is known to exist in a triple-helical state, is shown in Figure 2. This spectrum shows three maxima at 1629, 1645, and 1667 cm^{-1} as described previously.[9d] The central peak (1645 cm^{-1}) has the highest absorbance, followed by the incompletely resolved shoulders at 1629 cm^{-1} and 1667 cm^{-1}.

The assignment of this experimental amide I region to the collagen peptide was further investigated using FTIR simulation methods described above. Figure 2 illustrates the excellent agreement between the experimental and the simulated FTIR spectra for $(PPG)_{10}$.

The simulation program provides clear identification of the oscillators giving rise to the three observed features. The 1629 cm^{-1} peak comes from the summed contributions of the C_XO_X groups with a small contribution from the C_GO_G groups (see Figure 1 for notation). The main peak at 1645 cm^{-1} comes from the C_GO_G groups with minor contributions from the C_XO_X and C_YO_Y groups. The high frequency incompletely resolved sub-band is mainly due to the C_YO_Y groups as shown in Figure 2, the carbonyl in the peptide bond preceding Gly.

The FTIR experimental amide I data for the Gly → Ala peptide in D_2O solution at 16 °C are shown in Figure 3 together with the simulated band. This peptide also adopts a triple-helical conformation at low temperatures. Three main sub-bands are again suggested although the low-frequency shoulder is much less resolved compared to the $(Pro\text{-}Pro\text{-}Gly)_{10}$ spectrum. The experimental spectrum includes a very broad feature centered at 1644 cm^{-1} with a weak shoulder near 1630 cm^{-1} and a peak near 1671 cm^{-1}. The low wavenumber shoulder observed in $(Pro\text{-}Pro\text{-}Gly)_{10}$ at 1629 cm^{-1} is broader in the Gly → Ala variant than in the unsubstituted derivative but is still evident as a shoulder.

Comparison of the FTIR experimental data with the amide I simulation again shows excellent agreement and demonstrates the accuracy of the simulation. The contributions of the three amide/imide groups in the repeating Gly-Pro-Hyp unit as well as the Ala contribution were summed and compared to the experimental spectrum in Figure 3.

The success of the simulations engenders confidence that the assignment of the spectral features to particular peptide groups or carbonyl oscillators as noted above is indeed correct. The quality of these results permits us to examine the thermal unfolding of the $(Pro\text{-}Pro\text{-}Gly)_{10}$ peptide by simulating bands from higher temperature structures suggested by molecular mechanics simulations.

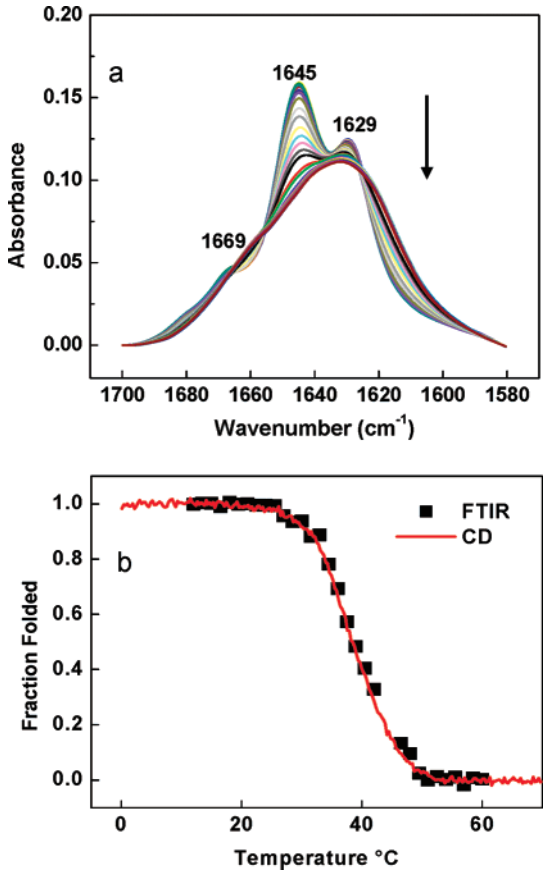

Figure 4. (a) Set of amide I spectra of $(PPG)_{10}$ in D_2O/DCl monitored by FTIR over the temperature range of 12−60 °C at 1.5 °C temperature intervals. (b) Melting curve derived from FTIR spectra monitored by intensity changes at 1645 cm^{-1} (filled squares) overlaid with CD derived melting curve monitored at 225 nm (red line).

3.2. FTIR Studies of the Thermal Transition of Triple-Helical Peptides. With the successful assignments of the amide I bands of these collagen model peptides available as a starting point, FTIR spectroscopy was used to monitor the thermal denaturation of these peptides. $(Pro\text{-}Pro\text{-}Gly)_{10}$ is known to undergo a triple helix to monomer transition at ~32 °C in dilute acetic acid.[12]

To investigate this transition, FTIR spectra were recorded in D_2O/DCl (pD = 3) from 12 to 60 °C in increasing temperature increments of 1.5 °C. The spectra are overlaid in Figure 4a. During denaturation, the band that undergoes the largest change is the central 1645 cm^{-1} feature, which significantly decreases in intensity as temperature is increased. The 1629 cm^{-1} band shows small changes with increasing temperature, and the high wavenumber band (1667 cm^{-1}) appears to broaden and disappear into the underlying broad feature as a function of temperature.

To determine the transition temperature (T_m) of $(PPG)_{10}$ by FTIR, the intensity change of the 1645 cm^{-1} band was plotted as a function of temperature (see section 2.2). A cooperative transition was observed with a T_m value of ~37 °C (Figure 4b). This thermal transition using FTIR data was very similar to a CD melting curve acquired at the same average heating rate, solvent, and concentration. Previous studies suggest the system is close to equilibrium.[12] This transition temperature in D_2O is higher than that observed for the peptide in a non-deuterated solvent at the same pH, consistent with previous reports.[18]

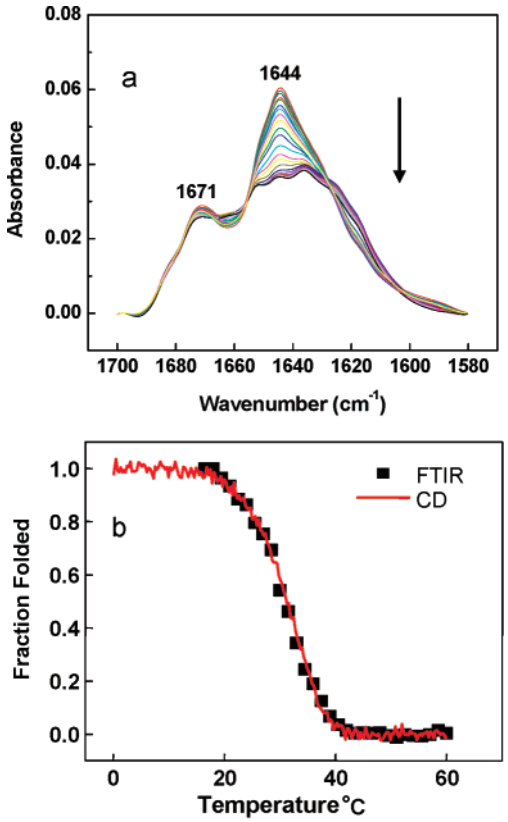

Figure 5. (a) Amide I profiles of the Gly → Ala peptide monitored by FTIR over the temperature range of 16−60 °C at 1.5 degree temperature intervals. (b) Melting curve of FTIR spectra monitored at 1644 cm^{-1} (filled squares) overlaid with CD melt at 225 nm (red line).

The same experimental approach was used to monitor the thermal transition of the G → A substituted peptide. FTIR spectra in the amide I region were collected over the temperature range of 16−60 °C. As seen in Figure 5a, at low temperatures the profile contains two main bands at 1644 and 1671 cm^{-1}. As the temperature increases, the central 1644 cm^{-1} undergoes a significant decrease in absorbance followed by a broadening accompanied by the appearance of additional shoulders. In contrast, the high wavenumber band (1671 cm^{-1}) shows only a small decrease in intensity and remains quite distinct with temperature. The thermal transition was investigated by monitoring the change in the central 1644 cm^{-1} band as a function of temperature. The plot in Figure 5b yields a T_m of 33 °C for both FTIR and CD.

At temperatures higher than the triple helix transition region, the FTIR spectra in the amide I region for $(Pro\text{-}Pro\text{-}Gly)_{10}$ and Gly → Ala reach a state which no longer changes significantly with increasing temperature. The $(Pro\text{-}Pro\text{-}Gly)_{10}$ FTIR spectra at temperatures of 50 °C and above reveal an asymmetric profile with a peak maximum at 1633 cm^{-1} and a weak shoulder at 1665 cm^{-1}. At temperatures greater than 42 °C, the FTIR spectra for Gly → Ala remain unchanged. The final Gly → Ala peptide spectrum shows a very broad band centered near 1630 containing a series of maxima at 1629, 1637, 1644, and 1652 cm^{-1} and a well-defined shoulder at 1671 cm^{-1} (Figure 6). Since 67% of both peptides are imino acids (Pro/Hyp), the unfolded spectra are compared to that of polyproline II (PPII) in D_2O/DCl. The

(18) Gough, C. A.; Bhatnagar, R. S. *J. Biomol. Struct. Dyn.* **1999**, *17*, 481−491.

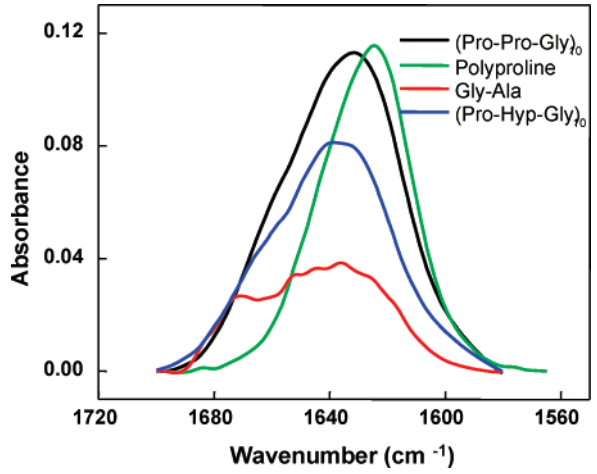

Figure 6. Amide I spectra of unfolded states of $(Pro\text{-}Pro\text{-}Gly)_{10}$ (black line), Gly → Ala (red line), and polyproline (green line) all at 54 °C, and $(Pro\text{-}Hyp\text{-}Gly)_{10}$ (blue line) at 80 °C.

PPII FTIR spectrum has a single maximum at 1621 cm^{-1} at low temperature (12 °C), which shifts to 1625 cm^{-1} at 54 °C. The thermal transition is linear, with precipitation after 60 °C, consistent with previous reports. The PPII spectrum at 54 °C shows some general resemblance to that of the unfolded $(Pro\text{-}Pro\text{-}Gly)_{10}$ peptide and also to that of the unfolded state of $(Pro\text{-}Hyp\text{-}Gly)_{10}$ (Figure 6). However, the $(Pro\text{-}Pro\text{-}Gly)_{10}$ and $(Pro\text{-}Hyp\text{-}Gly)_{10}$ unfolded spectra are significantly broader, a possible result of the presence of another peak at high wavenumbers which may originate from amide I vibrations of Gly. As seen in Figure 6, the Gly → Ala unfolded state differs from the unfolded $(Pro\text{-}Pro\text{-}Gly)_{10}$ and high-temperature PPII in having well-defined features.

3.3. FTIR Experimental and Simulation Studies of the Early Unfolding Stage of $(PPG)_{10}$. The uniformity and simplicity of the triple-helical structure of $(PPG)_{10}$ makes it possible to test unfolding schemes. The peptides studied are relatively short in length and both X-ray data and NMR data indicate that the terminal residues are not well ordered.[3c,19] In fact, the C-terminal glycines are missing from the 1K6F X-ray structure of the $(Pro\text{-}Pro\text{-}Gly)_{10}$ peptide. This suggests unfolding is likely to occur from one or both ends of the peptide.

The dramatic changes in the shape of these contours as temperature is increased are due to the structural changes occurring during thermal denaturation. We used a model based on an unraveling of the $(PPG)_{10}$ triple helix to obtain evolving structures from which a series of contours could be calculated and compared with the experimental contours. Agreement would indicate the correctness of the model. To test this model, we compared the calculated FTIR amide I spectra of partially unfolded species with the experimental data as generated during the early stages of the thermal transition discussed in the prior section. It is reiterated that our simulations of the FTIR data require a structure to permit calculation of the amide I contour.

Visualization of the early unfolded state depicts the triple helix in a partially unwound state (Figure 7a). Amide I contours

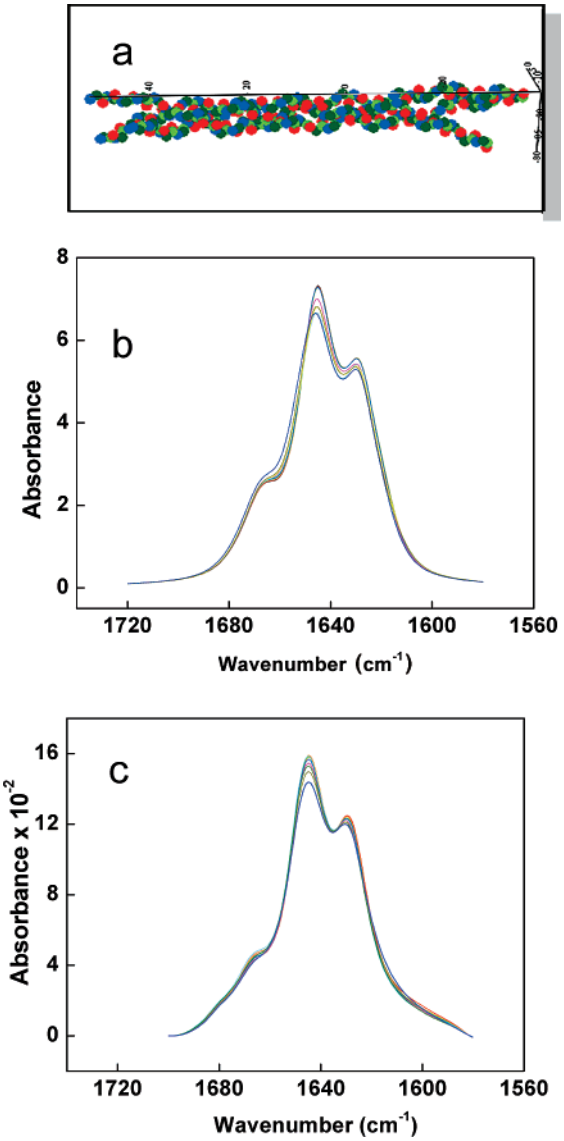

Figure 7. (a) Partially unwound structure of the $(PPG)_{10}$ molecule after 13 steps. (b) Changes in the simulated amide I contour for $(PPG)_{10}$ as a function of a gradual unraveling of the triple helix as described in the text. (c) Corresponding experimental thermally induced FTIR spectra of $(PPG)_{10}$ over the temperature range 12−34 °C.

were simulated for each progressively changed structure and were compared with experimental IR data. FTIR simulations (Figure 7b) for the first 10 steps in the unfolding (via the mechanism outlined above) reproduce the general pattern of intensity change seen experimentally in Figure 7c. In particular, the central 1645 cm^{-1} component of the amide I contour undergoes the most significant spectral changes in both the simulated and experimental spectra.

4. Discussion

A long-standing problem in protein IR spectroscopy has been the interpretation of the conformation-sensitive amide I contour. Hundreds of studies have attempted to determine the relative amount of secondary structures including helices, sheets, loops, and turns in proteins from this spectral feature (for reviews, see ref 20). Certain frequency ranges are assumed to correspond to particular secondary structures, for example, α-helices are assumed to absorb from 1650−1660 cm^{-1}. The relative band

(19) Brodsky, B.; Li, M. H.; Long, C. G.; Apigo, J.; Baum, J. *Biopolymers* **1992**, *32*, 447−451.
(20) (a) Miyazawa, T.; Shimanouchi, Y.; Mizushima, S. *J. Chem. Phys.* **1958**, *29*, 611−616. (b) Miyazawa, T.; Blout, E. R. *J. Am. Chem. Soc.* **1961**, *83*, 712−719. (c) Arrondo, J. L. R.; Muga, A.; Castresana, J.; Goni, F. M. *Prog. Biophys. Mol. Biol.* **1993**, *59*, 23−56.

areas for each of the sub-bands are often taken to represent the percentages of the particular secondary structures. There are problems with this approach.[21] First, extinction coefficients for different secondary structures may differ among proteins and within a single protein. Second, only regular, extensive regions (which rarely exist in proteins) would produce sharp bands at characteristic predictable frequencies. The correctness of the typically used set of spectra−structure correlations has been called into question many times. As an example, the effects of solvent exposure on the amide I modes of a synthetic helical coil have been demonstrated; shifts greater than 20 cm^{-1} from the "accepted" value have been observed.[22] Many such situations in which the standard spectra−structure correlations have been found wanting are well-documented and confound the traditional method of band analysis.

Our approach,[11] which requires a priori knowledge of the protein structure, regards the amide I contour as resulting from a collection of interacting amide I oscillators (one per peptide bond). The interactions between them are calculated from geometry-sensitive formulas suggested by physical considerations. The advantage of the approach is depicted in Figures 2 and 3 where the structural origin of observed spectral features can be elucidated.

In this work, the above simulation approach is extended to model the structural differences between the collagen model peptide $(Pro\text{-}Pro\text{-}Gly)_{10}$ and a Gly → Ala variant which mimics the single Gly substitution mutations seen in collagen diseases. The high-resolution crystal structures available for these peptides permit simulations of the amide I region and an analysis of the contributions of the three nonequivalent amide/imide groups within each repeating X-Y-Gly unit. The excellent agreement between the simulated amide I bands and that observed experimentally in the native spectra confirms the validity of the computational method and provides confidence in the origin of the bands assigned in the simulations. Extensive studies have been reported on the amide I region of the FTIR spectrum for $(Pro\text{-}Pro\text{-}Gly)_{10}$.[9d] The spectrum reported here closely resembles that reported by Lazarev, showing three distinct bands at similar positions. On the basis of the presence of the three nonequivalent C═O groups within each repeating Pro-Pro-Gly unit, Lazarev assigned the three major bands, assigning 1628 cm^{-1} to C_GO_G (or C_1O_1 in Lazarev's notation), 1644 cm^{-1} to C_XO_X (C_2O_2 in Lazarev's notation), and 1664 cm^{-1} to C_YO_Y (C_3O_3 in Lazarev's notation). The large change in the 1644 cm^{-1} peak upon denaturation was considered to support the assignment of this peak to the carbonyl between the two Pro residues, since this carbonyl is involved in direct hydrogen bonding to the amine NH of the Gly. However, the analysis of contributions from the simulation indicates a reversal of the 1645 cm^{-1} and 1629 cm^{-1} assignments. The dominant contribution to the 1644 cm^{-1} sub-band comes from C_GO_G oscillators, not from the hydrogen-bonded C_XO_X carbonyls, while the dominant contribution to 1629 cm^{-1} comes from C_XO_X. However, both bands have contributions from C_XO_X and C_GO_G because of through space and through bond coupling.

To simulate spectral changes during the early stages of unfolding, molecular dynamics simulations at 16 °C and 27 °C were used to suggest dihedral angle changes at successive α carbon atoms in the chains to represent various degrees of peptide unwinding. The similarity of the calculated IR spectra between most of the partially unwound forms and the experimental spectra lends credence to the method.

Since a dramatic decrease in intensity was observed for the 1644 cm^{-1} band, and not for the other amide I bands, the intensity at 1644 cm^{-1} was monitored during the thermal transition and showed excellent agreement with the CD melting curve obtained under the same conditions. Even though there is a complex origin of the 1644 cm^{-1} band, it is largely from C_GO_G and the change in its vibrational intensity must reflect the unfolding which is also seen in the thermal dependence of the CD spectra.

This is the first reported FTIR spectrum of a collagen-like peptide with a Gly substitution, which models the mutations found in collagen diseases. The excellent agreement between the simulation of the Gly → Ala peptide and its low-temperature experimental spectrum suggests that the solution structure from molecular mechanics is reasonably accurate. The crystal structure shows a local perturbation at the Ala site, with a small unwinding and replacement of the direct hydrogen bond by a water-mediated hydrogen bond ($AlaNH\cdots H_2O\cdots C_XO_X$).[3c] The similarity of its amide I profile to that of $(Pro\text{-}Pro\text{-}Gly)_{10}$, with the same three components, supports the local nature of the changes.

Similar to the observations for $(Pro\text{-}Pro\text{-}Gly)_{10}$, the central 1644 cm^{-1} peak of the Gly → Ala peptide undergoes the most dramatic intensity change with increasing temperature, and when it was used to monitor the thermal transition, it also demonstrated excellent agreement with the CD melting curve obtained under the same conditions. This band may act as a marker which is highly sensitive to the change in the secondary structure as the collagen model undergoes a thermal transition from the native to the unfolded state.

The FTIR spectra of the unfolded states of these two peptides were characterized. It has been suggested that unfolded proteins and peptides may contain polyproline II-like structure,[23] and these peptides which have high imino acid content are expected to retain some polyproline II-like character in the unfolded state. At high temperatures, polyproline II, $(Pro\text{-}Pro\text{-}Gly)_{10}$, and $(Pro\text{-}Hyp\text{-}Gly)_{10}$ have generally similar broad bands near 1625−1633 cm^{-1}, although the latter two peptides have a greater breadth and asymmetry on the high wavenumber side which may result from the presence of a residual shoulder in the spectrum. The residual structure in the amide I spectra of the unfolded state of Gly → Ala is rather unexpected. These data suggest that the break in the central region restricts some accessible conformations in the unfolded state relative to $(Pro\text{-}Pro\text{-}Gly)_{10}$. The less random nature of the Gly → Ala monomer could be a factor in the decrease in thermal transition temperature of Gly → Ala

(21) (a) Braiman, M. S.; Rothschild, K. *J. Annu. Rev. Biophys. Chem.* **1998**, *17*, 541−570. (b) Surewicz, W. K.; Mantsch, H. H. *Biochim. Biophys. Acta* **1988**, *952*, 115−130. (c) Jackson, M.; Mantsch, H. H. *Crit. Rev. Biochem. Mol. Biol.* **1995**, *302*, 95−120.

(22) (a) Gilmanshin, R.; Gulotta, M.; Dyer, R. B.; Callender, R. H. *Biochemistry* **1988**, *40*, 5127−5136. (b) Manas, E. S.; Getahun, Z.; Wright, W. W.; DeGrado, W. F.; Vanderkooi, J. M. *J. Am. Chem. Soc.* **2000**, *122* (41), 9883−9890.

(23) (a) Sreerama, N.; Woody, R. W. *Biochemistry* **1994**, *33*, 10022−10025. (b) Shi, Z.; Woody, R. W.; Kallenbach, N. R. *Adv. Protein Chem.* **2002**, *62*, 163−240. (c) Ma, K.; Kan, L-s.; Wang, K. *Biochemistry* **2001**, *40*, 3427−3438. (d) Measey, T.; Schweitzer-Stenner, R. *Chem. Phys. Lett.* **2005**, *408*, 123−7.

relative to $(Pro\text{-}Hyp\text{-}Gly)_{10}$ (33 °C and 67 °C in D_2O, respectively).[8] Restriction of conformational freedom in the monomer is entropically unfavorable and could lead to a smaller free-energy difference between the native trimer and the unfolded monomer states. Alternatively, the lower transition temperature may simply reflect the disruptive effect of the Gly to Ala substitution.

A variety of extensions of the current experimental approach may be envisioned. Isotope labeling of particular peptide bond carbons will significantly change the observed amide I contour and provide insight into changes at defined positions as the conformation is being altered. Such an approach may also be profitably applied in ligand-binding applications.

Acknowledgment. This work was supported by PHS grant GM 29864 to R.M., GM60048 to B.B., and an NIH Predoctoral Fellowship Award (F31) to M.A.B.

JA071154I