THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Polymorphism of Collagen Triple Helix Revealed by ^{19}F NMR of Model Peptide [Pro-4(*R*)-Hydroxyprolyl-Gly]$_3$-[Pro-4(*R*)-Fluoroprolyl-Gly]-[Pro-4(*R*)-Hydroxyprolyl-Gly]$_3$

Kazuki Kawahara,[†,×] Nobuaki Nemoto,[‡] Daisuke Motooka,[†] Yoshinori Nishi,[§] Masamitsu Doi,[∥] Susumu Uchiyama,[⊥] Takashi Nakazawa,[⊗] Yuji Nishiuchi,[○] Takuya Yoshida,[†] Tadayasu Ohkubo,[†] and Yuji Kobayashi*[,§,⊥]

[†]Graduate School of Pharmaceutical Sciences, Osaka University, Suita, Osaka 565-0871, Japan

[‡]JEOL RESONANCE Inc., Akishima, Tokyo 196-8558, Japan

[§]Osaka University of Pharmaceutical Sciences, Takatsuki, Osaka 569-1094, Japan

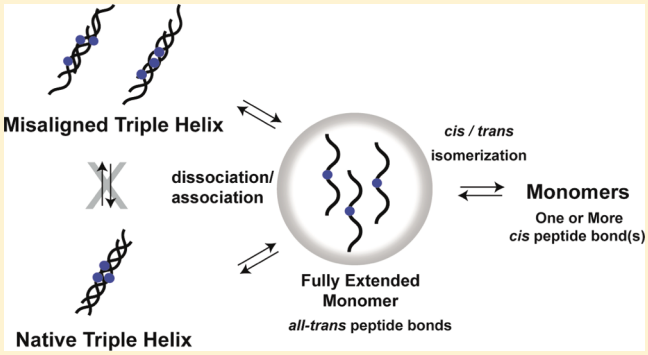
[∥]Department of Materials Science, Wakayama National College of Technology, Gobo, Wakayama 644-0023, Japan

[⊥]Graduate School of Engineering, Osaka University, Suita, Osaka 565-0871, Japan

[⊗]Department of Chemistry, Nara Women's University, Nara 630-8506, Japan

[○]Peptide Institute Inc., Minoh, Osaka 562-8688, Japan

ABSTRACT: We have characterized various structures of (Pro-Hyp^R-Gly)$_3$-Pro-fPro^R-Gly-(Pro-Hyp^R-Gly)$_3$ in the process of *cis*–*trans* isomerization and helix–coil transition by exploiting the sole ^{19}F NMR probe in 4(*R*)-fluoroproline (fPro^R). Around the transition temperature (T_m), we detected a species with a triple helical structure distinct from the ordinary one concerning the alignment of three strands. The ^{19}F–^{19}F exchange spectroscopy showed that this misaligned and that the ordinary triple helices were interchangeable only indirectly via an extended monomer strand with all-*trans* peptide bonds at Pro–fPro^R, Pro–Hyp^R, and Gly–Pro in the central segment. This finding demonstrates that the helix–coil transition of collagen peptides is not described with a simple two-state model. We thus elaborated a scheme for the transition mechanism of (Pro-Hyp^R-Gly)$_n$ that the most extended monomer strand can be the sole source both to the misaligned and correctly folded triple-helices. The staggered ends could help misaligned triple helices to self-assemble to higher-order structures. We have also discussed the possible relationship between the misaligned triple helix accumulating maximally at T_m and the kinetic hysteresis associated with the helix–coil transition of collagen.



1. INTRODUCTION

Collagen is a fibrous protein characterized by its unique triple-helix, which consists of three polyproline II-like left-handed helices intertwining to form a right-handed supercoiled triple-helical structure.[1,2] It provides the Anfinsen dogma with a challenging issue concerning the relationship between the amino acid sequence and higher-order structure.[3] The characteristic feature of primary sequences of collagen is approximately 300 units of X-Y-Gly triplet repeats and, about 20% of X and Y residues are occupied by imino acids Pro and Hyp.

Synthesizing a series of polytripeptides (Pro-Pro-Gly)$_n$ with a varied length of *n*, we showed a strong dependency of thermal transition between the triple-helical structure and single-coil on the chain length of collagen models.[4,5] Substitution of Pro to 4(*R*)-hydroxyproline (Hyp^R) at the Y position dramatically enhanced the thermal stability of the triple-helix,[6] demonstrating the usefulness of this model system for elucidating the stability of native collagen in the triple-helical state.

In addition to the replacement of Pro in (Pro-Pro-Gly)$_n$ with Hyp^R, a variety of model peptides having 4(*S*)-hydroxyproline (Hyp^S), 4(*R*)-fluoroproline (fPro^R), and 4(*S*)-fluoroproline (fPro^S) in place of Pro have been prepared to study the effects of the hydroxyl group and the fluorine atom at the 4-position (γ-position) of Pro on the thermodynamic stability of triple helical structures.[2,7−12] These studies led to illuminate an empirical rule concerning the preference of pyrrolidine-ring conformation of Pro for the triple helix formation,[2,13] as well as

Special Issue: B: Harold A. Scheraga Festschrift

Received: December 31, 2011
Revised: March 1, 2012
Published: March 1, 2012

different thermodynamic effects of hydration on collagen model peptides in single-coil and triple-helix states.[11] We also correlated thermodynamic features of hydration with the network of hydrogen bonds and the ring-puckering of Pro (Hyp^R) residues, based on X-ray studies of $(Hyp^R\text{-}Hyp^R\text{-}Gly)_{10}$ and $(Pro\text{-}Hyp^R\text{-}Gly)_4\text{-}(Hyp^S\text{-}Pro\text{-}Gly)_2\text{-}(Pro\text{-}Hyp^R\text{-}Gly)_4$.[14,15] Owing to these studies, there is a gradual progress in the understanding of the relationship between the primary structure of collagen molecule and its higher-order structure. Because of the lack of experimental techniques, however, it still remains to clarify the mechanism of triple-helix formation or self-assembly to form a higher-order structure at the atomic level, instead of analyzing the conformational transitions of the polypeptides with a conventional all-or-none two-state model.[16] In this context, $fPro^R$ residues in the model peptides allow an innovative approach exploiting ^{19}F atoms as probes for ^{19}F NMR studies in solution;[17] individual ^{19}F signals in a peptide could possibly inform us of a distinctive conformational state at each $fPro^R$ residue in the molecule that could help to understand the detailed picture of their conformational transition. The number of ^{19}F signals could certainly be much fewer than that of ^1H and ^{13}C signals, thus reducing the need of assigning a large number of signals to the relevant atoms reflecting a variety of conformational states such as those involved in the possible *cis*−*trans* isomerism of the peptide chain.[18−25]

In nature, collagen in its triple-helical structure further assembles into fibrils. The process of self-assembly was partly reproduced by using a model peptide $(Pro\text{-}Hyp^R\text{-}Gly)_n$ ($n > 7$), while the resulting branched filamentous structures were not quite the same as the axially periodic structure of collagen fibrils.[26−28] Expecting that ^{19}F NMR might serve as a means to corroborate the molecular mechanism of not only the simple two-state transition between triple helical and single coil structures but also such a complex phenomenon as self-association involving a variety of conformational isomers, we have synthesized model peptides including $(Pro\text{-}Hyp^R\text{-}Gly)_3\text{-}Pro\text{-}fPro^R\text{-}Gly\text{-}(Pro\text{-}Hyp^R\text{-}Gly)_3$, which we refer to as $fPro^{11}\text{-}[Pro\text{-}Hyp^R\text{-}Gly]_7$. This peptide is designed to contain $fPro^R$ at the center of the peptide chain, so that the sole ^{19}F nucleus in $fPro^R$ can be analyzed as a probe for a variety of conformational states. The chain length of this peptide is the same as $(Pro\text{-}Hyp^R\text{-}Gly)_7$, which has been reported to take a stable triple helix structure with $T_m = 36$ °C and not to proceed further to a higher-order structure.[29]

Owing to the high sensitivity and unambiguous assignment of ^{19}F signals, we could monitor the transition of a number of conformational states including *cis*−*trans* isomers of a monomeric peptide and misaligned triple helices alongside the ordinary collagen-like triple helix in solution. We report here the results of ^{19}F NMR experiments suggesting that the monomer segment with all-*trans* configuration of the peptide bond N-terminal of each imino acid is responsible to form a misaligned triple-helix. We also discuss the possible self-assembly of triple-helical structures to higher-order structures.

2. MATERIALS AND METHODS

2.1. Preparation of Model Peptides. Tripeptide units, $Pro\text{-}fPro^R\text{-}Gly$ and $Pro\text{-}Hyp^R\text{-}Gly$, were synthesized as reported previously.[9−11] $(Pro\text{-}fPro^R\text{-}Gly)_n$ ($n = 1 - 3$), $(Pro\text{-}Hyp^R\text{-}Gly)_7$, and $(Pro\text{-}Hyp^R\text{-}Gly)_3\text{-}Pro\text{-}fPro^R\text{-}Gly\text{-}(Pro\text{-}Hyp^R\text{-}Gly)_3$ [$fPro^{11}\text{-}(Pro\text{-}Hyp^R\text{-}Gly)_7$] were synthesized using an Applied Biosystems 433A peptide synthesizer. Couplings were carried out on an Alko-PEG resin (0.26 mmol g^{-1}, 0.1 mmol, Watanabe Chemical Industries). The synthesis scale was 0.14 mmol, and the Fmoc-tripeptide unit was activated with HATU (*O*-7-azabenzotriazol-1-yl-1,1,3,3-tetramethyluronium hexafluorophosphate) (4.0 equiv)/DIEA (*N*,*N*-diisopropylethylamine) (8.0 equiv) in DMF. Cleavage of the peptide resin proceeded for 1 h using a TFA/water/triisopropylsilane mixture (95:2.5:2.5). The peptide was purified by high performance liquid chromatography (HPLC) on a ZORBAX C-18 reversed-phase column (21.2 × 250 mm, Agilent). The purity of the peptides used was more than 97% as judged by HPLC. The molecular weight of each peptide was determined by MALDI-TOF mass spectrometry.

2.2. CD Spectroscopy. Measurements of CD spectra for $(Pro\text{-}Hyp^R\text{-}Gly)_7$ and $fPro^{11}\text{-}(Pro\text{-}Hyp^R\text{-}Gly)_7$ were carried out at 4 °C on an Aviv Model 202 Spectropolarimeter. Spectra were obtained by averaging 8 scans from 190 to 260 nm in a quartz cell of 1 mm path length. The lyophilized powders of peptide samples were dissolved in 100 mM AcOH to make the solution concentration approximately 0.045 mM. The precise concentrations were 0.039 and 0.036 mM for $(Pro\text{-}Hyp^R\text{-}Gly)_7$ and $fPro^{11}\text{-}(Pro\text{-}Hyp^R\text{-}Gly)_7$, respectively, as determined further by amino acid analysis of the acid hydrolyzate. The temperature-dependent CD spectra were recorded from 4 to 90 °C at a heating rate of 0.1 °C min^{-1}. The measurements were repeated three times for the same sample solution to check the reproducibility.

2.3. NMR Spectroscopy. The peptides, $(Pro\text{-}fPro^R\text{-}Gly)_n$ ($n = 1 - 3$) and $fPro^{11}\text{-}(Pro\text{-}Hyp^R\text{-}Gly)_7$, were dissolved in D_2O at concentrations of about 20 mM. Trace DSS (sodium 2, 2-dimethyl-2-silapentane-5-sulfonate) and TFA (trifluoroacetic acid) were added as internal references for ^1H and ^{19}F, respectively. Although NMR spectra were acquired on a JNM-ECA500 or JNM-ECA600 spectrometer (JEOL Ltd., Tokyo, Japan) operating at 11.4 T (500 MHz for ^1H) or 14.7 T (600 MHz for ^1H) equipped with 5 mm $^{13}C\text{-}\{^1H,^{19}F\}$ triple resonance probe, all the spectra shown here were measured at 14.7 T. Spectra were recorded in the temperature range of 15 to 75 °C. At each temperature, data acquisition was performed after equilibration of the sample for more than 30 min in the probe. All the ^{19}F NMR spectra were measured with ^1H WALTZ16 decoupling with 6.25 kHz field strength. Two-dimensional (2D) $^{19}F-^1H$ heteronuclear Overhauser effect spectroscopy (HOESY) was performed at 65 °C with the mixing time (τ_m) of 300 ms. 2D [$^1H-^1H$ total correlation spectroscopy]−[$^{19}F\{^1H\}$ insensitive nuclei enhancement by polarization transfer] (TOCSY−INEPT) spectra were measured at 65 °C with DIPSI-2 isotropic mixing with a 10.5 kHz field strength.[30] The experiments of $^{19}F-^{19}F$ exchange spectroscopy (EXSY) were performed at 65 °C with τ_m = 500 ms. Modifying the QUIET-NOESY sequence,[31−33] a nonselective 180° (^1H) pulse was applied at the midpoint of τ_m in the standard $^{19}F-^{19}F$ NOESY sequence: $90°-t_1-90°-\tau_m-90°-$(acquisition), to eliminate NOE cross-peaks arising from spin diffusion, so that correlations due to conformational exchange can be detected. All the data were processed using Delta-NMR software (JEOL Ltd.).

3. RESULTS AND DISCUSSION

3.1. Triple-Helix of $fPro^{11}\text{-}[Pro\text{-}Hyp^R\text{-}Gly]_7$. The CD spectrum of the model peptide $(fPro^{11}\text{-}[Pro\text{-}Hyp^R\text{-}Gly]_7)$ at 4 °C and the concentration of 0.045 mM showed an absorption band at around 225 nm, which was almost indistinguishable

dx.doi.org/10.1021/jp212631q | J. Phys. Chem. B 2012, 116, 6908−6915

with that of (Pro-Hyp^R-Gly)$_7$ in the characteristic collagen-like triple-helical structure (Figure 1). Thus, the influence of the

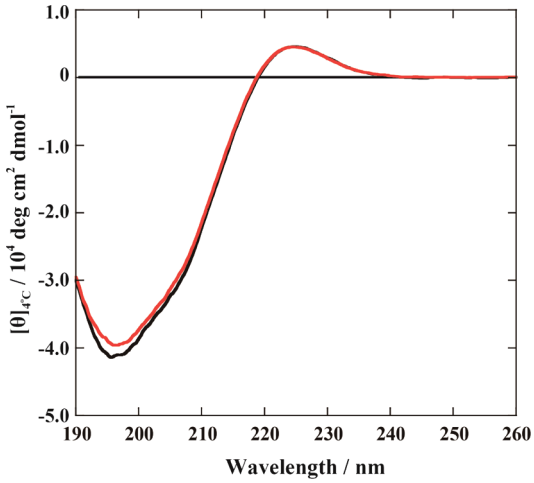


Figure 1. CD spectra of (Pro-Hyp^R-Gly)$_7$ (black) and fPro^{11}-[Pro-Hyp^R-Gly]$_7$ (red) at 4 °C. Concentration of each peptide is approximately 0.04 mM in 100 mM acetic acid.

substitution was nominal. The transition temperature (T_m) of fPro^{11}-[Pro-Hyp^R-Gly]$_7$) was about 36 °C and was comparable to those of (Pro-Hyp^R-Gly)$_7$ (data not shown).[29]

3.2. *cis*−*trans* Isomerism of Peptide Bonds in fPro^{11}-[Pro-Hyp^R-Gly]$_7$. The temperature dependence of the ^{1}H-decoupled 1D ^{19}F NMR spectra of fPro^{11}-[Pro-Hyp^R-Gly]$_7$ is shown in Figure 2. As the temperature increased, the relative intensities of the peaks around −179.0 ppm decreased, and those in the range from −180.0 to −182.0 ppm increased, apparently reflecting the transition of triple helix to single coil. Thus, we assigned the former peaks as the signals from species in triple helix and the latter from species in single coil. This is consistent with the findings that (Pro-fPro^R-Gly)$_n$ (n = 1 − 3), which does not have the ability to form a triple helix and exists in single-coil states even at 4 °C, gave the ^{19}F peaks only in the range −180.0 to −182.0 ppm and none around −179.0 ppm in the spectrum at 75 °C as shown in Figure 2. In the spectrum of (Pro-fPro^R-Gly)$_1$, there were two peaks at −180.80 and −181.14 ppm with the relative intensities of 95% and 5%, respectively.

Allowing for about 9:1 preference in favor of *trans* over *cis* isomer measured by ^{1}H NMR of Ac-fPro^R-OMe,[29,34,35] we assigned the ^{19}F signal observed with the higher intensity at −180.80 ppm to the fluorine in the *trans* isomer with respect to the Pro-fPro^R linkage. Accordingly, the minor peak at −181.14 ppm should represent the *cis* isomer. We also have the Gibbs free-energy difference $\Delta G^\circ_{298\,K}$ between *cis* and *trans* isomers at Pro-fPro^R in (Pro-fPro^R-Gly)$_1$ being 1.46 kcal mol^{-1} calculated using the equation $\Delta G^\circ = -RT\ln(K_{cis/trans})$ in which $K_{cis/trans}$ stands for the equilibrium constant of *cis*/*trans* isomerism. This value is within the range −2.0 to 2.0 kcal mol^{-1} reported for those of Pro−Pro and Pro−Hyp bonds in various oligopeptides.[36,37]

As shown in Figure 2, one major and several minor ^{19}F signals were observed in (Pro-fPro^R-Gly)$_2$ in addition to the two ^{19}F signals in the ^{19}F NMR spectrum of (Pro-fPro^R-Gly)$_1$. The two major peaks at −180.78 and −180.05 ppm were assigned to Fγ atoms of fPro^R at the N- and C-terminal side of the tripeptide unit, respectively, by their different responses to a

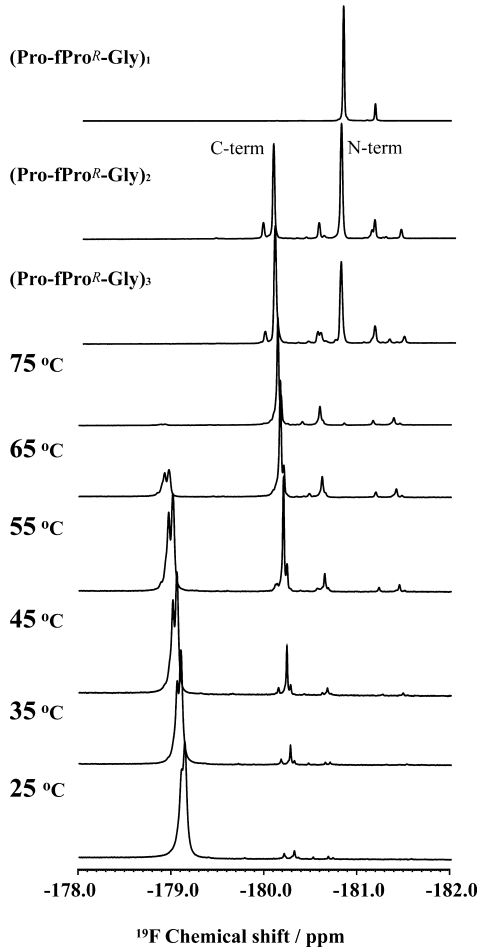


Figure 2. Temperature dependence of 1D ^{19}F NMR spectra of fPro^{11}-[Pro-Hyp^R-Gly]$_7$. Spectra recorded in the temperature range 25−75 °C are shown. 1D ^{19}F NMR spectra of (Pro-fPro^R-Gly)$_n$ (n = 1−3) recorded at 75 °C are also shown.

paramagnetic shift reagent, Pr(NO$_3$)$_3$, which is known to bind to the C-terminal carboxyl group of collagen polypeptide chain (data not shown).[38] In the H-Pro-fPro^R-Gly-Pro-fPro^R-Gly-OH sequence, if *cis*/*trans* isomer for Gly−Pro and Pro−fPro^R bonds were considered, a maximum of 2^3 isomers could be observed as 8 individual peaks; however, these signals were difficult to assign at this stage. In the ^{19}F NMR spectra of (Pro-fPro^R-Gly)$_2$, (Pro-fPro^R-Gly)$_3$, and even fPro^{11}-[Pro-Hyp^R-Gly]$_7$, ^{19}F signals of fPro^R in monomeric strands gave about the same set of peaks, almost irrespective of the number of triplet repeats.

In the ^{19}F−^{19}F EXSY spectrum of (Pro-fPro^R-Gly)$_2$ at sufficiently long mixing time (500 ms) by which the multistep *cis*−*trans* isomerization event could be accomplished, we observed many exchange cross-peaks characteristic of *cis*−*trans* isomerization of Gly−Pro and Pro−fPro^R bonds in monomer states when temperature increased to 75 °C (Figure 3). We detected the conformational exchange between the largest peak (C2), which should reflect the major isomer with all *trans* peptide bonds and the other seven peaks (C1 and C3−C8), which should reflect isomers with one or more *cis* peptide bonds. From this exchange connectivity, we determined that peaks (C1−C8) are deduced as from 2^3 conformational states for ^{19}Fγ of fPro^R at the C-terminal triplet due to *cis*/*trans* duality for each one of three peptide linkages at Pro-fPro^R, Gly-Pro, and the N-terminal Pro-fPro^R. The other two ^{19}F signals

dx.doi.org/10.1021/jp212631q | J. Phys. Chem. B 2012, 116, 6908−6915

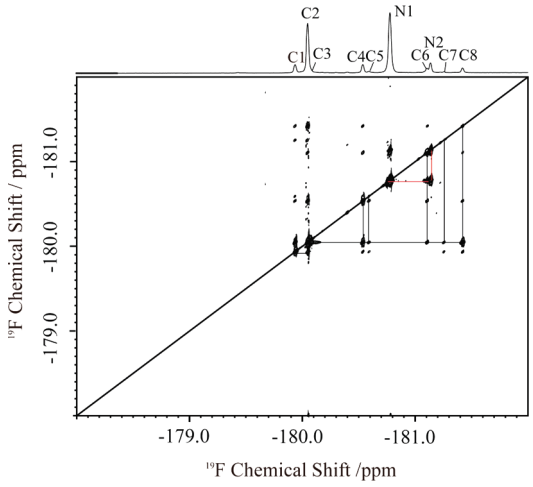


Figure 3. ^{19}F−^{19}F EXSY spectrum of (Pro-fPro^{R}-Gly)$_2$ recorded at 75 °C with a mixing time of 500 ms. The one-dimensional spectrum shown at the top of the spectra is the regular ^{19}F NMR spectrum of (Pro-fPro^{R}-Gly)$_2$ recorded at 75 °C. In the ^{19}F NMR spectrum, ^{19}F signals from fPro^{R} residue in triplets at N- and C-terminal sides of the sequence are labeled with N and C, respectively.

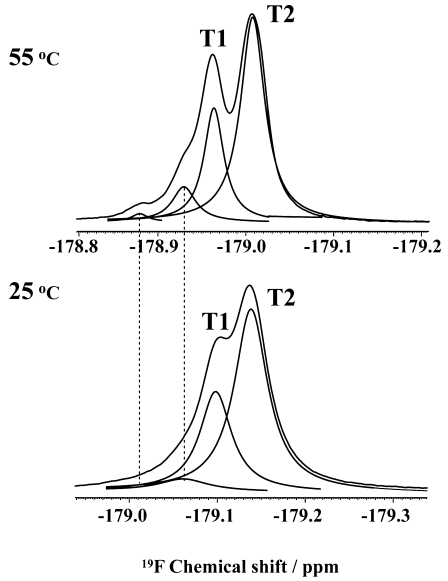


Figure 4. Deconvoluted peaks at around −179.0 ppm in the ^{19}F NMR spectra of fPro^{11}-[Pro-Hyp^{R}-Gly]$_7$ at 25 °C (bottom) and 55 °C (top). The two resolving peaks designated as T1 and T2 belong to the same triple-helix (see the text for detail), and the ratio of deconvoluted peak area of these peaks is close to 1:2.

(N1 and N2) of fPro^{R} had no such correlations but exchange between them. Thus, we determined N1 is *trans* and N2 is *cis* configuration of ^{19}Fγ of fPro^{R} at the N-terminal triplet attributed to *cis*−*trans* isomerization at the N-terminal Pro-fPro^{R} bond. As a similar effect has been observed in NMR spectra of backbone ^{1}H and ^{13}C signals of other collagen model peptides,[18−25,39] it is interesting that the ^{19}F signal of the N-terminal fPro^{R} was insensitive to the structural heterogeneity of the C-terminal side of the peptide chain.

For peptides (Pro-fPro^{R}-Gly)$_n$ with the larger number of n, the similar much higher preference of *trans* over *cis* configuration to that observed in (Pro-fPro^{R}-Gly)$_n$ (n = 1 − 3) could possibly be applied. On this ground, it is possible to assume that the most intense peak at −180.2 ppm in the spectrum of fPro^{11}-[Pro-Hyp^{R}-Gly]$_7$ represents the monomer in which all the peptide bonds throughout the full length of the chain take *trans* configuration because there is no particular evidence that contradicts this preference in the segments involving the Pro-Hyp^{R}-Gly unit. We therefore refer to such an all-*trans* configuration as a fully extended structure hereafter. The multiple ^{19}F peaks for the single site in the single coil state could be assigned.

3.3. Asymmetry of fPro^{11}-[Pro-Hyp^{R}-Gly]$_7$ in the Triple Helix. The ^{19}F NMR spectra of the model peptides provide information that has not been available from the analyses with conventional NMR experiments. That is, in ^{1}H, ^{13}C, and ^{15}N NMR spectra of model peptides in the triple-helical state, the signals of interest tend to converge into one broad peak. Because individualities from the position of each atom in the molecule are lost due to the symmetric and repetitive nature of the assembly, it has been difficult to obtain detailed information.[18,22,39] The ^{19}F spectra of fPro^{11}-[Pro-Hyp^{R}-Gly]$_7$ exhibited clearly two resolved peaks: T1 (−178.96 ppm) and T2 (−178.01 ppm), with an intensity ratio close to 1:2 at high temperatures as shown in Figure 4. The appearance of these two peaks could possibly be interpreted by the model depicted in Figure 5. That is, two of the three fPro^{R} residues in, say, chains A and B face the segment −Gly-Pro-fPro^{R}− of the adjacent chain, while the other fPro^{R} in chain C is in contact with the segment −Gly-Pro-Hyp^{R}− of another chain in the triple helix, in which these ^{19}F nuclei are *not* magnetically equivalent.

These interactions can be discerned in a ^{19}F−^{1}H HOESY spectrum displaying cross-correlations between the ^{19}Fγ signals of fPro^{R} in one chain and six ^{1}H peaks of Pro (Hα, Hβ1, Hβ2, Hγ, Hδ1, and Hδ2,) in another chain and one ^{1}H peak of Pro (Hα) in the same chain, if we assume the same distances between the Oγ of Hyp^{R} and hydrogen atoms of Pro in the crystals of (Pro-Hyp^{R}-Gly)$_n$ (Figure 6).[40] As shown in Figure 7, these seven correlations between the Fγ of fPro^{R} and hydrogen atoms of Pro in the neighboring chain within the distance of 0.5 nm were detected as cross-peaks in the present ^{19}F−^{1}H HOESY experiment. Thus, it might be possible to distinguish between the intra- and interchain interactions. Interestingly, the relative intensities of cross-peaks of ^{19}F peaks T1 and T2 were almost indistinguishable as judged from their 1D ^{1}H{^{19}F} traces in the ^{19}F−^{1}H HOESY spectrum (Figure 7B). Accordingly, three peptide chains in the triple helix are structurally identical, despite being magnetically non-equivalent around one of the three fPro^{R} residues due to the one residue staggering arrangement of each chain in the collagen triple-helix.

3.4. Partially Disordered Triple Helical State of fPro^{11}-[Pro-Hyp^{R}-Gly]$_7$. While deconvoluting the unsymmetrical broad peaks of T1 and T2, we recognized the existence of two additional minor peaks (Figure 4). These shoulder peaks have a temperature dependency such that their total intensity is maximal at 55 °C. This coincides precisely with T_m of the helix−coil transition estimated from the crossover of increasing and decreasing peak intensities of single coil and triple helical species as a function of temperature (Figure 8). Note that the ^{19}F−^{1}H HOESY correlations of shoulder peak at −178.88 ppm are almost indistinguishable from those of peaks T1 and T2, as their ^{1}H{^{19}F} traces show (Figure 7B). Therefore the structures of the species giving the shoulder peaks seem to be almost identical to the triple helical species giving T1 and T2 peaks. Furthermore, there is a clear correlation between shoulder

dx.doi.org/10.1021/jp212631q | J. Phys. Chem. B 2012, 116, 6908−6915

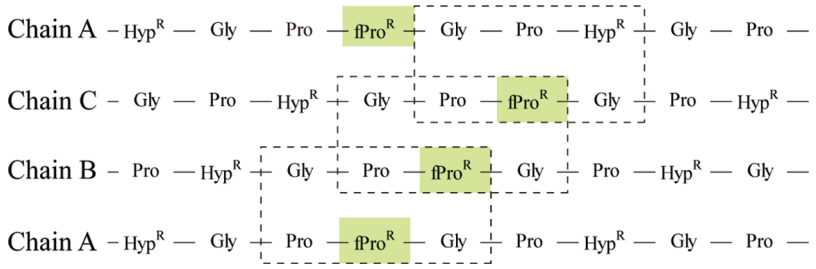


Figure 5. Schematic representation of alignment of chains in the triple helix of fPro^{11}-[Pro-Hyp^R-Gly]$_7$. The fPro^R residues are shaded by light green. Two fPro^R residues (in chains A and B) face the tripeptide unit −Gly-Pro-fPro^R−, whereas the other fPro^R residue (in chain C) faces the tripeptide unit −Gly-Pro-Hyp^R−.

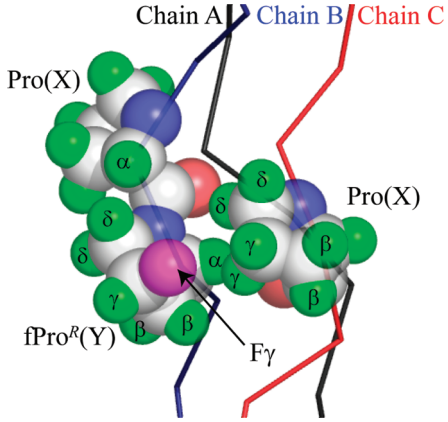


Figure 6. Model structure of triple helix of fPro^{11}-[Pro-Hyp^R-Gly]$_7$. The model is built by substituting the hydroxyl group of Hyp^R at the Y position of the crystal structure of (Pro-Hyp^R-Gly)$_n$ (PDB code: 1V6Q) with a fluorine atom. The protons within the fluorine−proton distance of 0.5 nm are labeled as α−γ. The three peptide strands may be distinguished by color (red, black, and blue).

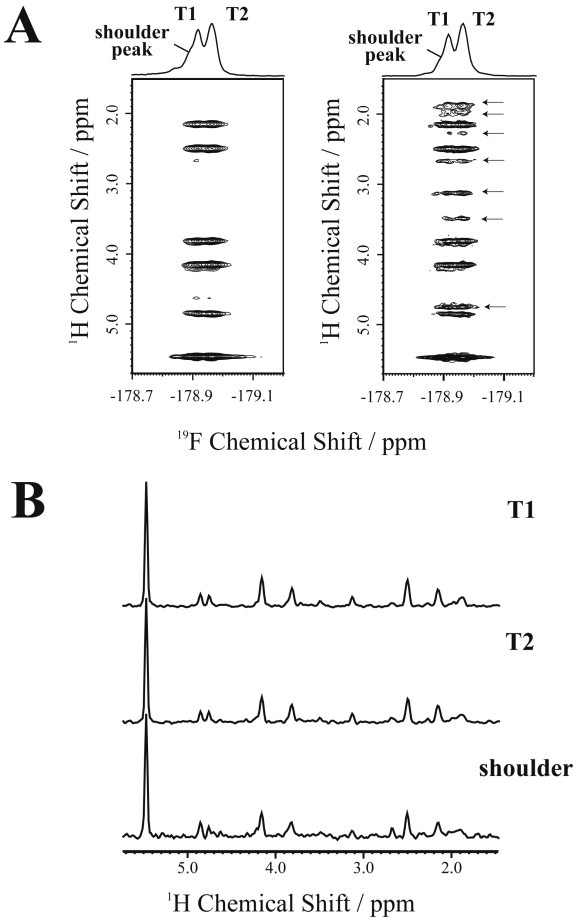


Figure 7. Expanded region of TOCSY-INEPT spectrum (left) and ^{19}F−^{1}H HOESY spectrum (right) of fPro^{11}-[Pro-Hyp^R-Gly]$_7$ at 65 °C (A). Inter-residual {^{1}H}−^{19}F HOE cross-peaks are indicated by arrows. The 1D traces relative to the ^{1}H dimension for peak T1 (−178.92 ppm), peak T2 (−178.97 ppm), and shoulder peak (−178.89 ppm) (B). The number and intensity of {^{1}H}−^{19}F HOE cross-peaks observed for the minor conformers are virtually identical with those for peak T1 and T2, indicating that fPro^R in the minor conformers have interchain contacts similar to those found in native triple-helix.

peaks and a peak representing monomer in addition to the correlation between major trimer peaks (T1 and T2) and monomer peaks in the ^{19}F−^{19}F EXSY spectrum (Figure 9). These results suggest that the shoulder peaks represent the misaligned triple helix with partially folded or unfolded triple-helical structure other than the ordinary triple helix, which explained the difference in their chemical shifts might be caused by the deviation of the relative position of three peptides from the correct alignment of strands in the triple helix depicted in Figure 5.

The monomer peak correlated with the triple helical species corresponds to C2 of (Pro-fPro^R-Gly)$_2$, in which the peptide bond at the N-terminal side of each imino acid takes *trans* configuration (Figures 3 and 9). As discussed in the previous section, the chemical shift of the ^{19}F signal is not sensitive enough to probe the *cis*−*trans* isomerism in the full length of fPro^{11}-[Pro-Hyp^R-Gly]$_7$. Given the approximately 9:1 preference of the *trans* over *cis* configuration at the Pro−fPro^R bond in (Pro-fPro^R-Gly)$_1$, it is feasible that the most intense peak corresponds to C2 of (Pro-fPro^R-Gly)$_2$. Further, it is likely that this peak represents the monomer species of fPro^{11}-[Pro-Hyp^R-Gly]$_7$ with the all peptide bonds in *trans* configuration, existing in the fully extended state.

We observed no cross-correlations between the main peaks (T1 and T2) and the shoulder peaks in the ^{19}F−^{19}F EXSY spectrum (Figure 9). This fact leads to a new perception that the intermediate is an immature triple-helix due to the misalignment, which might be corrected not through sliding a peptide chain along the axis of the helix but via dissociation into the most extended monomer. If the difference between the genuine and immature collagen-like triple helix is due to the correct and incorrect alignment, respectively, we can explain the formation of higher-order structures observed for the model peptides (Pro-Hyp^R-Gly)$_n$ with longer chain lengths (n > 7)[26−28] than ours; a misaligned triple helix must have staggered

dx.doi.org/10.1021/jp212631q | J. Phys. Chem. B 2012, 116, 6908−6915

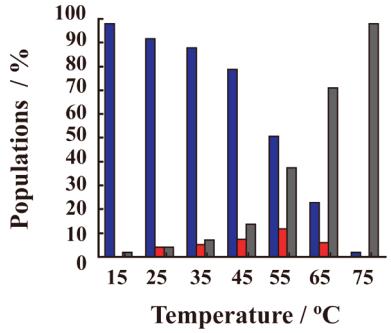


Figure 8. Histogram of individual conformational states (triple-helix, blue; minor (shoulder), red; and single-coil, gray) at different temperatures. The relative populations of triple helices (major and minor) and single coil were calculated from the peak areas of the relevant ^{19}F signals (Figures 2 and 4). The total population is normalized to 100%.

ends available for its self-assembling to extend the helix in a similar manner as the ligation of double-stranded DNA fragments with sticky ends.[41]

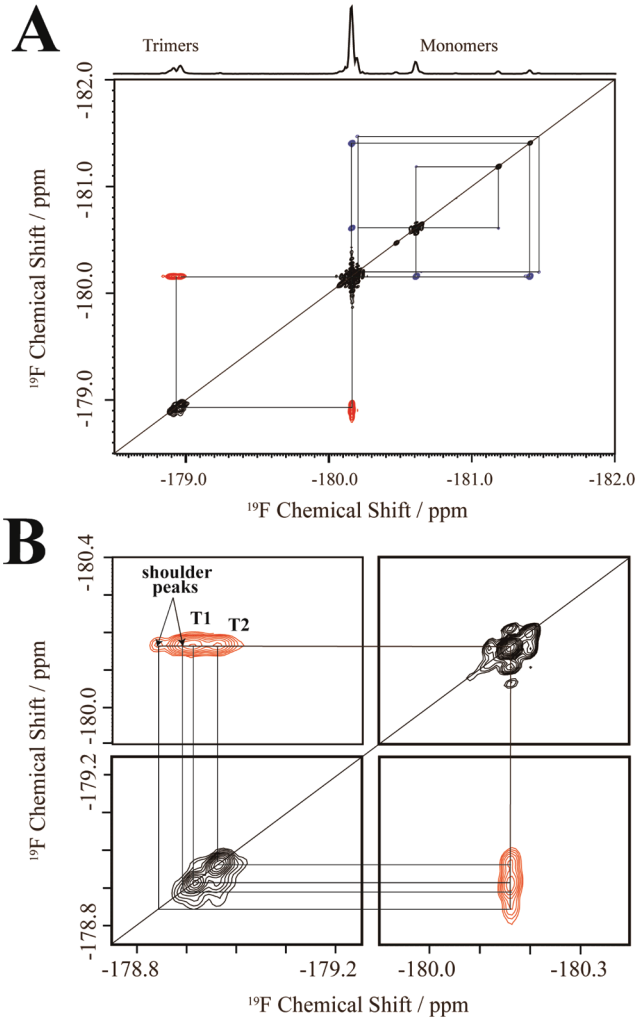


Figure 9. ^{19}F−^{19}F EXSY spectrum of fPro^{11}-[Pro-Hyp^{R}-Gly]$_7$ recorded at 65 °C with a mixing time of 500 ms (A). Expanded region of the ^{19}F−^{19}F EXSY spectra showing exchange cross-peaks between monomer segment with all-*trans* configuration and trimer species (B).

In fact, Raines and co-workers have demonstrated that fibrils resembling native collagen resulted from short synthetic peptides, which were designed to have staggered ends with different lengths of Pro−Pro−Gly or Pro−Hyp^{R}−Gly repeats by disulfide bonds.[42] Moreover, filamentous structures occurred appreciably by the self-association of trimeric (Pro-Hyp^{R}-Gly)$_{10}$ above the critical concentration of 1 mM, probably owing to a sufficient build-up of misaligned strands.[27] In this respect, the triple helical strands of fPro^{11}-[Pro-Hyp^{R}-Gly]$_7$ are less likely to self-assemble to higher-order structures because of the much lower stability of the requisite misaligned strands relative to the correctly assembled triple helix, in which the interchain contacts are maximal. That is, short misaligned triple helices would dissolve into the monomer much faster than assemble themselves to higher-order structures. In addition, the 20 mM solution of fPro^{11}-[Pro-Hyp^{R}-Gly]$_7$ can contain such misaligned strands at about 2 mM (10% of total concentration) at T_m = 55 °C (Figure 8). The maximal accumulation of the possible misaligned strands at the transition temperature is consistent with the optimal conditioning of (Pro-Hyp^{R}-Gly)$_n$ (n > 7) to form higher-order structures.[27]

3.5. Alignment of fPro^{11}-[Pro-Hyp^{R}-Gly]$_7$ Strand in Triple Helices Preceding the Formation of Higher-Order Structure. The majority of recent kinetic studies assume the involvement of a nucleation as the rate-determining step in the folding of the collagen-like triple helix.[23,43−49] The intermediate acting as a nucleus includes an all-*trans* chain, which builds up from those involving *cis*-peptide bonds to a considerable concentration to determine the rate.[23,43,44] However, it is insufficient to construct a model for explaining the kinetics by assuming a few intermediates without referring to the relevant experimental evidence that could specify the pair of species in equilibrium with each other. We thus suggest including the misaligned triple helix as a species, which can influence the rate of ordinary helix−coil transition. Kinetic hysteresis in collagen folding is one of such problems that could possibly be addressed by assuming the misaligned triple helix as an intermediate. That is, an appreciable accumulation of misaligned triple helix leads the concentration of fully extended monomer to lower so that the rate of assembling the monomer to the correctly folded triple helix should be reduced. If the concentration of monomer $[C]$ with the extended structure is lowered by the proportion spared to form misaligned structures maximally at T_m, the apparent $(T_m)_{cool}$ measured on cooling the solution of monomer should be below T_m. Even the change in $[C]$ is very small, its effect should be allowed because the rate of assembling three chains is a function of $[C]^3$, which might not be neglected to account for a minute difference between T_m and $(T_m)_{cool}$. Upon heating the solution so rapidly, the conversion of extended monomer to misaligned triple helices occurs only *after* the dissociation of correctly aligned triple helix, making the transition temperature on heating $(T_m)_{heat}$ higher than T_m. These arguments are based on our view that the ordinary triple helix occurs only through assembling the fully extended single strand but not from the misaligned structures by reorganizing the strands, at least in the experimental time scale of less than 1−1.5 s. The implication of our experimental result is in accord with a staggering zipper model in which rearranging the mismatched strands by sliding along the helices is theoretically prohibited.[50] This would suggest an alternative explanation to the kinetic hysteresis found on cooling and heating the solutions of (Gly-Pro-Pro)$_{10}$

dx.doi.org/10.1021/jp212631q | J. Phys. Chem. B 2012, 116, 6908−6915

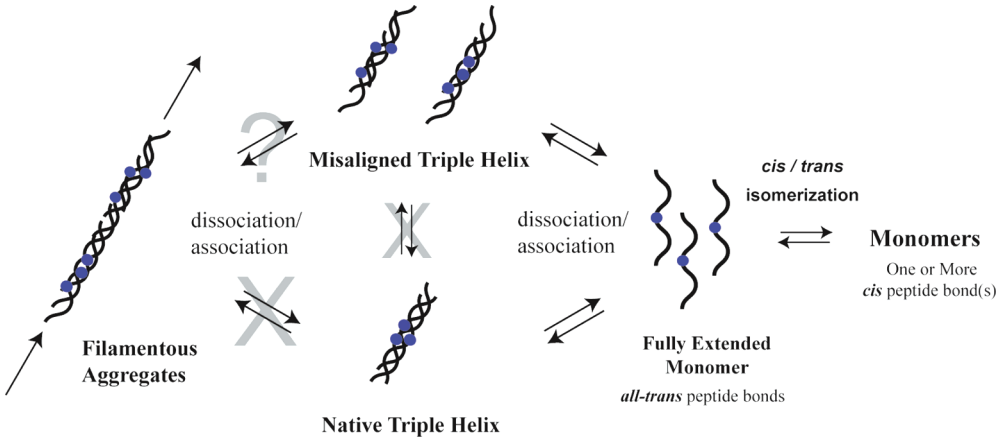


Figure 10. Schematic representation of the possible self-assembly mechanism of collagen model peptides (Pro-Hyp^R-Gly)$_n$. Considering that the nucleation event to form trimer species requires the monomer segment with all-*trans* configuration, the accumulation of misaligned trimer that could form filamentous structures are not negligible when the triplet repeats are longer than 7.

and (Gly-Pro-Hyp^R)$_{10}$ around T_m at a rate much faster than that required to establish equilibrium in the solution.[49]

In summary, we suggest a scheme depicting the relationship of monomer strands, triple helices, and higher-order structures (Figure 10). This scheme considers the fully extended single strand as a key intermediate leading to both correctly and incorrectly aligned triple helices. Figure 10 also includes the misaligned triple helix as the sole intermediate to be assembled into higher-order structures as discussed in the previous section. By extending the chain length of the peptide with [Pro-Hyp^R-Gly]$_n$-(Pro-fPro^R-Gly)-[Pro-Hyp^R-Gly]$_n$, we might be able to find higher-order structures such as filamentous (Figure 10) and randomly structured aggregates due to the possible increase in the proportion of misaligned triple helix.

4. CONCLUSIONS

Taking advantage of the high sensitivity and simplicity of a ^{19}F probe, we have characterized various conformational states of fPro^{11}-[Pro-Hyp^R-Gly]$_7$ in solution without ambiguity due to the difficulty in assigning a peak to each structure in NMR spectra with a large number of overlapped peaks. Individual species of the collagen model peptides undergoing isomerization, association, and dissociation were identified. These include a fully extended single strand with the all-*trans* configuration capable of forming not only the ordinary triple helix but also a misaligned one with staggered strands on both sides like the sticky ends of DNA, allowing the helices to self-assemble to form higher-order structures. The ^{19}F−^{19}F EXSY experiment showed that the misaligned and ordinary triple helices are interchangeable only indirectly via this fully extended single strand. We have clearly demonstrated that the transition of the collagen peptide between triple helical and single coil states is not described with a simple two-state model. We have also discussed a possibility that the misaligned triple helix accumulating maximally at T_m has some bearing on the kinetic hysteresis often found in helix−coil transition of collagen. An alternative mechanism of the transition of the collagen peptide, (Pro-Hyp^R-Gly)$_n$, was proposed.

AUTHOR INFORMATION

Corresponding Author

*Tel/Fax: +81-6-6879-4579. E-mail: yujik@protein.osaka-u.ac.jp.

Present Address

[×]Center for Research of Ancient Culture, Nara Women's University, Nara 630-8506, Japan.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. Evelyn R. Stimson for careful examination of the manuscript. We also thank Dr Yutaka Maeda and Professor Hisanori Shinohara for ^{19}F NMR measurements on a JNM-ECA600 spectrometer at the Research Center for Material Science of Nagoya University. One of the authors, Y.K., would like to dedicate this article to Professor Scheraga as a token of his gratitude to an outstanding and generous mentor in the field of Chemistry.

REFERENCES

(1) Brodsky, B.; Thiagarajan, G.; Madhan, B.; Kar, K. *Biopolymers* **2008**, *89*, 345−353.
(2) Shoulders, M. D.; Raines, R. T. *Annu. Rev. Biochem.* **2009**, *78*, 929−958.
(3) Anfinsen, C. B.; Haber, E.; Sela, M.; White, F. H. Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1961**, *47*, 1309−1314.
(4) Sakakibara, S.; Kishida, Y.; Kikuchi, Y.; Sakai, R.; Kakiuchi, K. *Bull. Chem. Soc. Jpn.* **1968**, *41*, 1273.
(5) Kobayashi, Y.; Sakai, R.; Kakiuchi, K.; Isemura, T. *Biopolymers* **1970**, *9*, 415−425.
(6) Sakakibara, S.; Inoue, K.; Shudo, K.; Kishida, Y.; Kobayashi, Y.; Prockop, D. J. *Biochim. Biophys. Acta* **1973**, *303*, 198−202.
(7) Holmgren, S. K.; Taylor, K. M.; Bretcher, L. E.; Raines, R. T. *Nature* **1998**, *392*, 667.
(8) Shoulders, M. D.; Kamer, K. J.; Raines, R. T. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 3859−3862.
(9) Doi, M.; Nishi, Y.; Uchiyama, S.; Nishiuchi, Y.; Nakazawa, T.; Ohkubo, T.; Kobayashi, Y. *J. Am. Chem. Soc.* **2003**, *125*, 9922−9923.
(10) Doi, M.; Nishi, Y.; Kiritoshi, N.; Iwata, T.; Nago, M.; Nakano, H.; Uchiyama, S.; Nakazawa, T.; Wakamiya, T.; Kobayashi, Y. *Tetrahedron* **2002**, *58*, 8453−8459.
(11) Nishi, Y.; Uchiyama, S.; Doi, M.; Nishiuchi, Y.; Nakazawa, T.; Ohkubo, T.; Kobayashi, Y. *Biochemistry* **2005**, *44*, 6034−6042.
(12) Persikov, A. V.; Ramshaw, J. A. M.; Kirkpatrick, A.; Brodsky, B. *J. Am. Chem. Soc.* **2003**, *125*, 11500−11501.
(13) Vitagliano, L.; Berisio, R.; Mazzarella, L.; Zagari, A. *Biochemistry* **2001**, *58*, 459−464.

(14) Kawahara, K.; Nishi, Y.; Nakamura, S.; Uchiyama, S.; Nishiuchi, Y.; Nakazawa, T.; Ohkubo, T.; Kobayashi, Y. *Biochemistry* **2005**, *44*, 15812−15822.
(15) Motooka, D.; Kawahara, K.; Nakamura, S.; Doi, M.; Nishi, Y.; Nishiuchi, Y.; Kang, Y. K.; Nakazawa, T.; Uchiyama, S.; Yoshida, T.; Ohkubo, T.; Kobayashi, Y. *Biopolymers* **2011**, in press.
(16) Locardi, E.; Kwak, J.; Scheraga, H. A.; Goodman, M. *J. Phys. Chem. A* **1999**, *103*, 10561−10566.
(17) Gerig, J. T. *Prog. Nucl. Magn. Reson. Spectrosc.* **1994**, *26*, 293−370.
(18) Mayo, K. H. *Biopolymers* **1996**, *40*, 359−370.
(19) Melacini, G.; Bonvin, A. M. J. J.; Goodman, M.; Boelens, R.; Kaptein, R. *J. Mol. Biol.* **2000**, *300*, 1041−1048.
(20) Doi, M.; Nakazawa, T.; Watanabe, K.; Uchiyama, S.; Ohkubo, T.; Kobayashi, Y. *Peptide Science* **2001**, *2000*, 341−344.
(21) Fiori, S.; Saccà, B.; Moroder, L. *J. Mol. Biol.* **2002**, *319*, 1235−1242.
(22) Fan, P.; Li, M.; Brodsky, B.; Baum, J. *Biochemistry* **1993**, *32*, 13299−13309.
(23) Buevich, A. V.; Dai, Q.; Liu, X.; Brodsky, B.; Baum, J. *Biochemistry* **2000**, *39*, 4299−4308.
(24) Xu, Y.; Hyde, T.; Wang, X.; BHate, M.; Brodsky, B.; Baum, J. *Biochemistry* **2002**, *42*, 8969−8703.
(25) Buevich, A. V.; Baum, J. *J. Am. Chem. Soc.* **2002**, *124*, 7156−7162.
(26) Kishimoto, T.; Morihara, Y.; Osanai, M.; Ogata, S.; Kamitakahara, M.; Ohtsuki, C.; Tanihara, M. *Biopolymers* **2005**, *79*, 163−172.
(27) Kar, K.; Amin, P.; Bryan, M. A.; Persikov, A. V.; Mohs, A.; Wang, Y.; Brodsky, B. *J. Biol. Chem.* **2006**, *281*, 33283−33290.
(28) Kar, K.; Wang, Y.; Brodsky, B. *Protein Sci.* **2008**, *17*, 1086−1095.
(29) Bretscher, L. E.; Jenkins, C. L.; Taylor, K. M.; DeRider, M. L.; Raines, R. T. *J. Am. Chem. Soc.* **2001**, *123*, 777−778.
(30) Battiste, J.; Newmark, R. A. *Prog. Nucl. Magn. Reson. Spectrosc.* **2006**, *48*, 1−23.
(31) Zwahlen, C.; Vincent, S. J. F.; Di Bari, L.; Levitt, M. H.; Bodenhausen, G. *J. Am. Chem. Soc.* **1994**, *116*, 362−368.
(32) Vincent, S. J. F.; Zwahlen, C.; Bodenhausen, G. *J. Biomol. NMR* **1996**, *7*, 169−172.
(33) Vincent, S. J. F.; Zwahlen, C.; Post, C. B.; Burgner, J. W.; Bodenhausen, G. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 4383−4388.
(34) Eberhardt, E. S.; Panasik, N. Jr.; Raines, R. T. *J. Am. Chem. Soc.* **1996**, *118*, 12261−12266.
(35) Renner, C.; Alefelder, S.; Bae, J. H.; Budisa, N.; Huber, R.; Moroder, L. *Angew. Chem., Int. Ed.* **2001**, *40*, 923−925.
(36) Grathwohl, C.; Wüthrich, K. *Biopolymers* **1976**, *15*, 2025−2041.
(37) Grathwohl, C.; Wüthrich, K. *Biopolymers* **1981**, *20*, 2623−2633.
(38) Kobayashi, Y.; Kyogoku, Y. *J. Mol. Biol.* **1973**, *81*, 337.
(39) Sarkar, S. K.; Young, P. E.; Sullivian, C. E.; Torchia, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 4800−4803.
(40) Okuyama, K.; Hongo, C.; Fukushima, R.; Wu, G.; Narita, H.; Noguchi, K.; Tanaka, Y.; Nishino, N. *Biopolymers* **2004**, *76*, 367−377.
(41) Pandya, M. J.; Spooner, G. M.; Sunde, M.; Thorpe, J. R.; Rodger, A.; Woolfson, D. N. *Biochemistry* **2000**, *39*, 8728−8734.
(42) Kotch, F. W.; Raines, R. T. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 3028−3033.
(43) Baum, J.; Brodsky, B. *Fold. Des.* **1997**, *2*, R53−R60.
(44) Xu, Y.; Bhate, M.; Brodsky, B. *Biochemistry* **2002**, *41*, 8143−8151.
(45) Frank, S.; Boudko, S.; Mizuno, K.; Schulthess, T.; Engel, J.; Bächinger, H. P. *J. Biol. Chem.* **2003**, *278*, 7747−7750.
(46) Bachmann, A.; Kiefhaber, T.; Boudko, S.; Engel, J.; Bächinger, H. P. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13897−13902.
(47) Engel, J.; Bächinger, H. P. *Matrix Biol.* **2000**, *19*, 235−244.
(48) Mizuno, K.; Boudko, S.; Engel, J.; Bächinger, H. P. *Biophys. J.* **2010**, *98*, 3004−3014.
(49) Persikov, A. V.; Xu, Y.; Brodsky, B. *Protein Sci.* **2004**, *13*, 893−902.
(50) Weidner, H. *Biopolymers* **1975**, *14*, 763−780.