

Structural Insights from  $^{15}\text{N}$  Relaxation Data for an Anisotropic Collagen Peptide

Jianxi Xiao and Jean Baum\*

Department of Chemistry and Chemical Biology, BioMaPs Institute for Quantitative Biology, Rutgers University, Piscataway, New Jersey 08854

Received July 9, 2009; E-mail: jean.baum@rutgers.edu

Knowledge of the structure and dynamics of the collagen triple helix is important for understanding its many interactions with receptors and other matrix molecules as well as the perturbations that are caused by collagen disease mutations.<sup>1</sup> Collagen possesses a unique triple-helical conformation that consists of three supercoiled polyproline II (PPII)-like helices with a repetitive Gly-X-Y sequence. The unique triple-helical conformation is stabilized by interchain hydrogen bonding and an extensive hydration network, as seen in the high-resolution X-ray structures of collagen model peptides.<sup>2a</sup> Gly residues from the three chains are closely packed in the center of the triple helix, and the backbone amide proton of Gly forms H bonds with the carbonyl oxygen of the X residue in the adjacent chain. The degree of the H bonding may play an important role in recognition of collagen by other molecules and in determining the severity of collagen disease arising from Gly mutations; aspects of the frequency and strength of this H bonding have been studied by NMR, X-ray, and computational approaches.<sup>1d,2</sup> Herein we report  $^{15}\text{N}$  relaxation NMR measurements on a collagen-like model peptide that reveal that the orientations of the Gly N-H bonds relative to the protein backbone have an unanticipated geometry and suggest that the H bonding may be responsible for this effect.

Standard NMR approaches that use nuclear Overhauser effects (NOEs) and  $J$  couplings to obtain high-resolution structures are limited for collagen model peptides.<sup>1d</sup> Complications include their rodlike nature and repetitive sequences. The close packing of the three chains and the linear nature results in only a small number of short-range distances from NOE experiments. For globular proteins with intermediate anisotropy, there have been a few examples of the use of  $^{15}\text{N}$  relaxation rates for global structure refinement and determination of interdomain orientations.<sup>3</sup> Here we demonstrate that  $^{15}\text{N}$  relaxation measurements and their dependence on rotational diffusion anisotropy can be used to obtain novel structural information about N-H bonds in the very anisotropic collagen model peptide.

The triple-helical peptide T3-785 [with sequence (POG)<sub>3</sub>ITGARGLAG(POG)<sub>4</sub>Y, where O stands for hydroxyproline] was designed to model an imino acid-poor region occurring one triplet C-terminal to the unique collagenase cleavage site in type-III collagen.<sup>4a</sup> The crystal structure of T3-785 shows that the peptide is a long straight rod,<sup>2a</sup> and calculation of the relative ratio of the principal components of the inertia tensor from the crystal structure coordinates indicates that the peptide can be modeled as an axially symmetric rotor [see the Supporting Information (SI)].<sup>4b</sup> These rotational properties can be described by a cylinder model and a prolate ellipsoid model.<sup>4c</sup> NMR hydrogen exchange data indicate that the Gly residues in the peptide have a rigid backbone and that Gly is H-bonded, as indicated by the high protection factors.<sup>4a</sup>

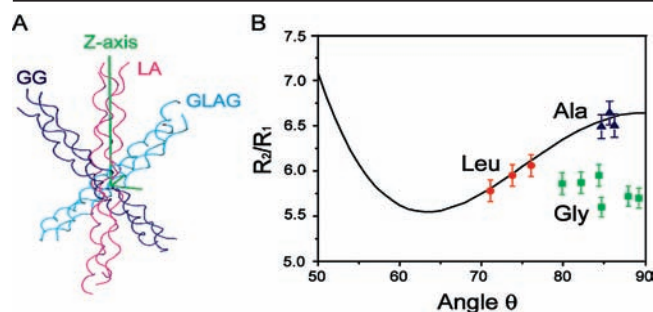
In a protein with axially symmetric diffusion, the  $^{15}\text{N}$  relaxation parameters  $R_1$  and  $R_2$  depend on the orientation of the N-H bond relative to the unique axis of the diffusion tensor (angle  $\theta$ ) (Figure S1 in the SI). From the experimental  $^{15}\text{N}$   $R_2/R_1$  ratios and the known

structure, the diffusion tensor can be derived using the fitting program r2r1\_diffusion, and the angle  $\theta$  that defines the orientation of the N-H bond can be obtained.<sup>4b</sup> Relaxation rates  $R_1$  and  $R_2$  were obtained at 500 MHz for labeled residues in the central region of T3-785, including G15, L16, A17, and G18, as well as for G24 at the C-terminal end (Table S1 in the SI). The  $R_2$  values are almost identical for all of the labeled residues, while the  $R_1$  values show small variations, with G15 and G18 having the largest values ( $R_1 = 2.02$  and  $2.04$ , respectively) and A17 the smallest value ( $R_1 = 1.79$ ). The differences in  $R_1$  result in a range of  $R_2/R_1$  ratios from 5.6 to 6.64.

The derivation of the diffusion tensor from  $R_2/R_1$  requires the use of residues for which there is an absence of conformational exchange and large-amplitude internal motion. All 12 labeled residues (four labeled residues per chain) showed no evidence of conformational exchange on the millisecond time scale, as determined from  $R_2^{\text{Hahn-echo}}$  experiments, and no evidence of large-amplitude internal motions, as seen from the NOE values, which were uniformly greater than 0.6. These results suggest no significant dynamics of the peptide in solution. Therefore, all of the labeled residues could in principle be included in the derivation of the diffusion tensor, but it was not possible to use the relaxation data for all four residues and obtain meaningful results.

The selection of which labeled residues to use in the derivation of the diffusion tensor was based on the following three criteria: an  $F$ -value analysis for selection of an isotropic or anisotropic model to describe the triple helix, the proper alignment of the unique axis of the diffusion tensor ( $\mathbf{D}$ ) with the long symmetric axis of the peptide, and a comparison of the experimental  $D_{\parallel}/D_{\perp}$  and  $\tau_c$  values with the ones calculated from the cylinder and prolate ellipsoid theoretical models. When the relaxation data for the four labeled residues G15-L16-A17-G18 or for only the two glycine residues G15 and G18 were used, an anisotropic model for the triple helix showed no improvement in the fit relative to an isotropic model, as indicated by the high  $p$  values (0.23 and 0.30, respectively), and the unique axis of the diffusion tensor was not parallel to the long symmetric axis of the peptide (Figure 1A). However, when only the six data points for L16-A17 were used, a low  $p$  value of 0.03 indicated that the anisotropic model was significantly better than the isotropic model, and the unique axis of the diffusion tensor was aligned parallel to the long axis of the peptide (Figure 1A), both of which are expected for this very anisotropic system. From the experimental data, the ratio of the principal values of the diffusion tensor,  $D_{\parallel}/D_{\perp}$ , is 13.1, and the overall correlation time  $\tau_c$  is 6.92 ns. These values are comparable to those obtained from a cylinder model ( $D_{\parallel}/D_{\perp} = 12.3$  and  $\tau_c = 6.98$  ns) and a prolate ellipsoid model ( $D_{\parallel}/D_{\perp} = 11.9$ ), indicating that the experimental values are consistent with those obtained from theoretical hydrodynamic models (see the SI).<sup>4</sup> These three criteria all indicate that using the relaxation data for L16-A17 only is the best selection for derivation of the diffusion tensor.

$R_2/R_1$  values were back-calculated for all of the labeled residues G15-L16-A17-G18 given the diffusion tensor obtained above (Figure



**Figure 1.** (A) Relative positions of the triple-helical peptide in the diffusion-tensor frame obtained by fitting the relaxation data of G15 and G18 only (purple), L16 and A17 only (red), and all the labeled residues GLAG (blue). (B) Plots of  $R_2/R_1$  vs  $\theta$ . The line shows  $R_2/R_1$  back-calculated on the basis of the diffusion tensor, and the colored dots (Gly in green, Leu in red, Ala in blue) show the experimental values. The figure assumes that  $\theta' = \theta$ .

1B) in order to compare the experimental and theoretical values and understand the basis for why G15 and G18 needed to be eliminated from the diffusion tensor derivation. The back-calculated  $R_2/R_1$  values for G15 and G18 deviate from the experimental values and are too large by  $\sim 12\%$ . These deviations could result from either the noncollinearity of the principal axis of the shielding tensor with the N–H bond<sup>5a</sup> or the uncertainty in the positions of the H atoms relative to the backbone. Calculations suggest that the noncollinearity of the principal axis of the shielding tensor with the N–H bond does not account for the inability to fit the Gly residues (see the SI).

The divergence of the experimental data from the theoretical data can be decreased by altering the positions of the backbone amide protons of the Gly residues by adjusting the angle  $\theta$  to  $\sim 71^\circ$  (Figure 1B). The X-ray structure does not contain information about protons, as their electron density is too weak to permit accurate determination. The amide protons were added to the X-ray structure with the program REDUCE under the assumption that the amide proton is positioned in the  $C'-N-C\alpha$  plane while the N–H vector is inclined a small amount ( $\sim 4^\circ$ ) away from the line that bisects the  $C'-N-C\alpha$  angle toward the N– $C\alpha$  bond.<sup>5b</sup> In order to make the calculated  $R_2/R_1$  values consistent with the experimental values, the  $\theta$  values for G15 and G18 were adjusted by placing the amide proton out of the  $C'-N-C\alpha$  plane by  $17.0 \pm 5.3$  degrees. There is precedent for this, as a 1 Å structure of BPTI obtained by neutron diffraction and X-ray data have shown that some amide protons are out-of-plane by  $0.4 \pm 4.8^\circ$ .<sup>5c</sup> More recently, residual dipolar coupling (RDC) studies of the third IgG-binding domain of protein G (GB3) indicated that most of the N–H vectors were located in the  $C'-N-C\alpha$  plane, while 38 of them were out of the plane with a maximum out-of-plane angle of  $11.5^\circ$  for K13.<sup>6a,7</sup>

It has been suggested that deviations of the amide proton from its standard position may arise from secondary structure differences or from H bonding.<sup>6a</sup> After modification of the positions of the glycine amide protons according to the new  $\theta$  angle, the H-bond angles were recalculated and found to be improved relative to the ones without modification: on average, the N–H–O angle is closer to  $180^\circ$  and the H–O distances are shorter (Table 1). No clashes were caused by the repositioning of the amide protons, as checked by the program MolProbity.<sup>6b</sup> With the modified  $\theta$  angle for the Gly N–H bonds, the diffusion tensor could be obtained by using 500, 600, or 800 MHz NMR relaxation data simultaneously for all of the labeled residues without the need to consider the noncollinearity of the chemical shift anisotropy.

**Table 1.** Hydrogen-Bond Information for the Labeled Glycines before and after Modification of the Amide Protons

H bond	before modification		after modification	
	N–H–O angle	H–O distance	N–H–O angle	H–O distance
1G15 N–H...2A13 C=O	158	1.83	168	1.8
1G18 N–H...2L16 C=O	155	1.89	175	1.84
2G15 N–H...3A13 C=O	145	2.05	175	1.93
2G18 N–H...3L16 C=O	154	2.01	161	1.99
3G15 N–H...1L16 C=O	163	1.91	169	1.9
3G18 N–H...1P19 C=O	146	2.15	170	2.05

We propose that in peptide T3-785, the H bonding rather than uniform PPII secondary structure may be responsible for the deviation of the Gly amide protons from their standard positions. This is supported by the fact that only the H-bonded Gly residues required a modification of their  $\theta$  angles while the X and Y residues remained in their standard positions. Although the repositioning of the Gly N–H vectors resulted in improved H-bond angles and lengths, the distortion of the Gly amide protons may impact the H-bond strengths in this collagen recognition region.

The use of  $^{15}\text{N}$  relaxation experiments to obtain long-range orientational restraints extends the structural tools available to the triple-helix system and may also be applied to nucleic acid systems that have anisotropic rodlike structural characteristics similar to those of the triple-helix protein. The experiments provide information about local N–H vector orientations and distortions that can be related to H-bonding properties. This powerful new tool may be also used to complement the short-range NOEs and  $J$ -coupling values found in anisotropic systems for more complete structure determination.

**Acknowledgment.** This work was supported by NIH Grant GM45302 and NSF Grants DBI-0403062 and DBI-0320746. We thank David Case, Seho Kim, and Barbara Brodsky for helpful discussions.

**Supporting Information Available:** Discussion of the inertia tensor, the diffusion tensor, hydrodynamic models, and the CSA contribution; Figure S1 and Table S1. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## References

- (a) Myllyharju, J.; Kivirikko, K. I. *Trends Genet.* **2004**, *20*, 33. (b) Lauer-Fields, J. L.; Juska, D.; Fields, G. B. *Biopolymers* **2002**, *66*, 19. (c) Di Lullo, G. A.; Sweeney, S. M.; Korkko, J.; Ala-Kokko, L.; San Antonio, J. D. *J. Biol. Chem.* **2002**, *277*, 4223. (d) Li, Y.; Brodsky, B.; Baum, J. *J. Biol. Chem.* **2009**, *284*, 20660.
- (a) Kramer, R. Z.; Bella, J.; Mayville, P.; Brodsky, B.; Berman, H. M. *Nat. Struct. Biol.* **1999**, *6*, 454. (b) Nerenberg, P. S.; Stultz, C. M. *J. Mol. Biol.* **2008**, *382*, 246. (c) Radmer, R. J.; Klein, T. E. *Biochemistry* **2004**, *43*, 5314.
- (a) Tjandra, N.; Garrett, D. S.; Gronenborn, A. M.; Bax, A.; Clore, G. M. *Nat. Struct. Biol.* **1997**, *4*, 443. (b) Fushman, D.; Xu, R.; Cowburn, D. *Biochemistry* **1999**, *38*, 10225. (c) Hashimoto, Y.; Smith, S. P.; Pickford, A. R.; Bocquier, A. A.; Campbell, I. D.; Werner, J. M. *J. Biomol. NMR* **2000**, *17*, 203. (d) Wu, H.; Blackledge, M.; Maciejewski, M. W.; Mullen, G. P.; King, S. M. *Biochemistry* **2003**, *42*, 57.
- (a) Fan, P.; Li, M. H.; Brodsky, B.; Baum, J. *Biochemistry* **1993**, *32*, 13299. (b) Tjandra, N.; Fella, S. E.; Pastor, R. W.; Bax, A. *J. Am. Chem. Soc.* **1995**, *117*, 12562. (c) Hall, J. B.; Fushman, D. *J. Biomol. NMR* **2003**, *27*, 261.
- (a) Boyd, J.; Redfield, C. *J. Am. Chem. Soc.* **1998**, *120*, 9692. (b) Word, J. M.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. *J. Mol. Biol.* **1999**, *285*, 1735. (c) Wlodawer, A.; Walter, J.; Huber, R.; Sjölin, L. *J. Mol. Biol.* **1984**, *180*, 301.
- (a) Ulmer, T. S.; Ramirez, B. E.; Delaglio, F.; Bax, A. *J. Am. Chem. Soc.* **2003**, *125*, 9179. (b) Davis, I. W.; Leaver-Fay, A.; Chen, V. B.; Block, J. N.; Kapral, G. J.; Wang, X.; Murray, L. W.; Arendall, W. B., 3rd; Snoeyink, J.; Richardson, J. S.; Richardson, D. C. *Nucleic Acids Res.* **2007**, *35*, W375.
- RDCs could potentially provide a complementary approach for obtaining orientational information about the N–H bond vectors in triple-helical peptides. Attempts to find a suitable alignment medium (bicelles, gels, and pf1 phage) have to date been unsuccessful.

JA9056823