Acetyl-Terminated and Template-Assembled Collagen-Based Polypeptides Composed of Gly-Pro-Hyp Sequences. 3. Conformational Analysis by ^{1}H-NMR and Molecular Modeling Studies

Giuseppe Melacini, Yangbo Feng, and Murray Goodman*

Contribution from the Department of Chemistry & Biochemistry, University of California at San Diego, La Jolla, California 92093-0343

Received April 17, 1996[⊗]

Abstract: Using 1D and 2D ^{1}H-NMR and molecular modeling, we characterized the conformational features of template-assembled collagen-like polypeptides of the type KTA-[Gly-(Gly-Pro-Hyp)$_n$-NH$_2$]$_3$ (n = 1, 3, 5, 6; KTA denotes the Kemp triacid) and of the corresponding acetylated single-chain polypeptides Ac-(Gly-Pro-Hyp)$_n$-NH$_2$ (n = 1, 3, 5, 6, 9) in water. We established the presence of triple-helical conformations on the basis of consistent experimental observations including the appearance of a set of distinct assembled resonances and the measurement of low hydrogen-exchange rates for the assembled Gly NH of the longer chain analogs. In addition, following the pioneering work of Brodsky *et al*. we proved the consistency of the NOESY spectra with the interchain NOEs anticipated by the X-ray model for triple-helical (Gly-Pro-Hyp) sequences. For the KTA-terminated structures the triple helicity is further supported by the KTA signal splitting detected for KTA-[Gly-(Gly-Pro-Hyp)$_n$-NH$_2$]$_3$ (n = 3, 5, 6) and caused by the triple-helical screw symmetry which breaks the rotational symmetry of KTA. Thermal melting studies indicate that the KTA template leads to a significant gain in the free energy of triple-helix formation. This free energy gain results in a remarkable increase of the thermal stabilities of the KTA terminated compounds as compared to the acetyl analogs. Our NMR results are fully consistent with our previous investigations based on CD, UV, and optical rotation spectroscopic methods.

Introduction

This paper is part of a series of investigations of assembled collagen-like triple helices N-terminated by KTA (*cis*,*cis*-1,3,5-trimethylcyclohexane-1,3,5-tricarboxylic acid, also known as the Kemp triacid[1]). First we presented KTA-[Gly-(Gly-Pro-Hyp)$_3$-NH$_2$]$_3$ and its *incipient* triple-helical properties.[2] Subsequently we described the synthesis of the KTA-[Gly-(Gly-Pro-Hyp)$_n$-NH$_2$]$_3$ (n = 1, 3, 5, 6) compounds and of the corresponding acetylated single-chain polypeptides Ac-(Gly-Pro-Hyp)$_n$-NH$_2$ (n = 1, 3, 5, 6, 9) together with the biophysical characterization of these compounds based on circular dichroism (CD), ultraviolet measurements (UV), optical rotation spectroscopy, and preliminary nuclear magnetic resonance (NMR) studies.[3] In this paper we present an extensive ^{1}H-NMR and modeling investigation with the aim of understanding how the KTA can enhance the assembly of triple-helical structures. Herein, the acetyl analogs will be denoted as **Ac-*n***, where *n* indicates the number of Gly-Pro-Hyp repeats per chain; the KTA analogs will be abbreviated as **KTAg-*n*,*m***, where **g** denotes the glycine spacer, the first number ***n*** refers to the number of Gly-Pro-Hyp repeats on each chain, and the second number ***m*** denotes the number of substitutions on the Kemp triacid derivative.

The collagen triple helix consists of three polypeptide chains, each in an extended polyproline II-like helix which is left-handed and has approximately three residues per turn.[4] Slight distortions of the backbone torsions from the values of the regular polyproline II conformation allow the three individual chains to supercoil around a common axis, resulting in a right-handed coiled-coil structure with screw symmetry. In the supercoiled structure the three chains are staggered by one residue. For instance in the Gly-Pro-Hyp sequence shown in Figure 1 at any

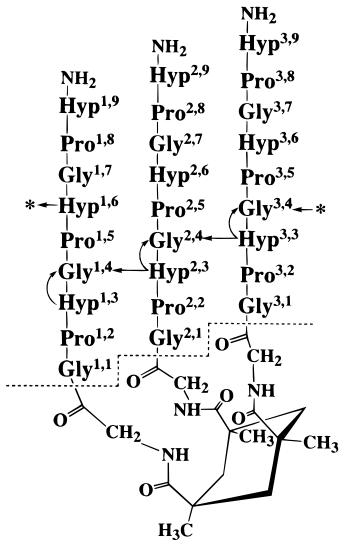

Figure 1. Schematic diagram of the triple-helical array of **KTAg-3,3** with its typical one-residue register shift. In the numbering convention of this figure, the first number refers to the chain position, while the second number indicates the residue position within a given chain. The dashed lines separate the template and the Gly spacers from the rest of the molecule. The arrows indicate different intrachain and interchain interactions between the Gly NH and Hyp C$_\alpha$H protons. An observed NOE between Gly NH and Hyp C$_\alpha$H can therefore arise from any one or all of these interactions.[4]

[⊗] Abstract published in *Advance ACS Abstracts,* October 15, 1996.

(1) Kemp, D. S.; Petrakis, K. S. *J. Org. Chem.* **1981**, *46*, 5140−5143.
(2) Goodman, M.; Feng, Y.; Melacini, G.; Taulane, J. P. *J. Am. Chem. Soc*. **1996**, *118*, 5156−5157. This is paper 1 of our series on collagen-based structures.
(3) Feng, Y.; Melacini, G.; Taulane, J. P.; Goodman, M. *J. Am. Chem. Soc*. **1996**, *118*, 10351−10358.
(4) Li, M.-H.; Fan, P.; Brodsky, B.; Baum, J. *Biochemistry* **1993**, *32* (29), 7373−7387.

S0002-7863(96)01261-9 CCC: $12.00

given level the triple-helix packing involves a Gly from one chain, a Pro from a second chain, and a Hyp from the remaining chain. This structure is stabilized by interchain hydrogen bonding and by the close packing of the three chains.[5]

In our design of template-assembled synthetic collagens, we took advantage of the close interchain packing to link the N-termini of the three polypeptide chains through a trifunctional organic molecule, such as the Kemp triacid. In order to reduce steric hindrance effects, glycine was introduced as a spacer residue between the KTA and each polypeptide chain of the triple helix, resulting in the KTA conjugates **KTAg-*n*,3** (n = 1, 3, 5, 6). We also studied the *N*-acetyl analogs **Ac-*n*** (n = 1, 3, 5, 6, 9) as control molecules. In this paper we report the characterization of the compounds mentioned above using 1D and 2D ^{1}H-NMR to obtain sequence-specific assignments, NOE, and solvent shielding information. The triple helicity of the analogs under investigation is studied as a function of chain length, template, and temperature. The resulting studies yield new insight on the effect of the KTA template on collagen-like triple helices and on the effect of triple helicity on the KTA template.

Materials and Methods

Sample Preparation. All compounds were synthesized as described in previous papers.[2,3] The NMR samples were prepared in H_2O/D_2O (9:1) and D_2O (purchased from Isotec, Inc.) with a peptide concentration in the range 1.4−4.0 mg/mL. After dissolution in water, all samples were kept at 5 °C for at least 24 h prior to any experiment, in order to allow for a proper equilibration of the sample.[6] The pH was adjusted to 3.4 ± 0.1 (direct pH meter reading without correction for isotope effects).

NMR Spectroscopy. All NMR experiments were carried out on an AMX-500 Bruker spectrometer. The 1D and 2D standard experiments were acquired and processed according to routine schemes as described in the Supporting Information.[7−12] For the equilibrium melting transition measurements, the temperature of the probe was calibrated using a DORIC 450 thermocouple, and the temperature was increased in steps of 2 °C allowing 12 min of equilibration time for each step. For the hydrogen-exchange measurements, 1D spectra were collected at 5 °C over a period of 19−20 days. The integrals were normalized using the peaks in the spectral region from 1.5 to 3 ppm which contains a known number of protons from the Pro $C_\beta H$, the Pro $C_\gamma H$, and the Hyp $C_\beta H$. In the case of the KTA analogs, this spectral region also includes the equatorial methylene protons of the KTA ring.

Molecular Modeling. Energy minimizations were carried out using the Discover program[13] and the force field CFF91 with a distance dependent dielectric constant to approximate the solvent effects.[14] For the molecular modeling of KTA-assembled triple helices, we focused on **KTAg-3,3** which among the KTA analogs under investigation is the lowest molecular weight compound showing evidence of a collagen-like triple-helical structure. The modeling of this analog has the advantage of requiring shorter computational times than the longer chain analogs such as **KTAg-5,3** and **KTAg-6,3**.

The 1D NMR integrations show that for **KTAg-3,3** at 5 °C the average number of Pro residues per molecule in a triple-helix like environment is 6. Considering possible end effects, we took a conservative approach of assuming only five Pro residues per three chains packed in the triple-helix array. For **KTAg-3,3** (Figure 1) there are three possible ways to pack five Pro residues in such a triple-helical array: residues 1,3−1,7, residues 2,2−2,6 and residues 3,1−3,5 are in a triple-helical conformation (Figure 1); residues 1,4−1,8, residues 2,3−2,7, and residues 3,2−3,6 are in a triple-helical conformation (Figure 1); residues 1,5−1,9, residues 2,4−2,8, and residues 3,3−3,7 are in a triple-helical conformation (Figure 1). Since we can not discriminate among these three possibilities, we carried out three independent simulations, one for each possibility. In each simulation 100 structures were initially generated by a distance geometry (DG) algorithm,[15] constraining the residues assumed to be in a triple-helical conformation to have the same backbone torsions and the same hydrogen-bonding features as the structure proposed by Miller, Nemethy, and Scheraga for $(Gly\text{-}Pro\text{-}Hyp)_n$.[16] The DG structures were then minimized and clustered (see Supporting Information).[5,17,18]

Results and Discussion

Assignments and Conformational Characterizations. The spin systems of the compounds studied were identified using DQF-COSY and TOCSY spectra as previously described by Brodsky *et al.*[4] For **Ac-1** at 5 °C two sets of resonances can be observed. The major set (80−85%) contains all *trans* peptide bonds, while the minor set represents conformations in which the Gly−Pro peptide bond is *cis* and the Pro−Hyp peptide bond is *trans* as indicated by the sequential NOEs (see Supporting Information). For **Ac-3** the assignment is essentially the same as for **Ac-1**. In **Ac-3** the overlapping of the resonances in the repeating tripeptide units results in the appearance of only three spin systems for Gly, Pro, and Hyp, with some dispersion of the $C_\alpha H_{l,h}$ and NH resonances. Also for **KTAg-1,3** the assignment is essentially the same as for **Ac-1** with the NH and the degenerate $C_\alpha H$ resonances of the spacer Gly at 8.20 and 3.79 ppm, respectively.

For **Ac-5**, **Ac-6**, **Ac-9**, **KTAg-3,3**, **KTAg-5,3**, and **KTAg-6,3** at 5 °C an additional set of resonances can be observed which is absent in **Ac-1**, **Ac-3**, and **KTAg-1,3**. In this new set of resonances the $C_\delta H$ protons of Pro are shifted upfield and become nondegenerate (see Supporting Information). The sequential connectivities between Gly and Pro, Pro and Hyp, and Hyp and Gly of the new set of resonances are consistent with the NOESY spectrum (Figure 2) leading to the definition of this new set of resonances as a distinct tripeptide unit. A number of experimental observations indicate that the new set of resonances detected for **Ac-5**, **Ac-6**, **Ac-9**, **KTAg-3,3**, **KTAg-5,3**, and **KTAg-6,3** corresponds to residues in a collagen-like triple helix. First, the same set of resonances constitutes 80% of the intensity for $(Pro\text{-}Hyp\text{-}Gly)_{10}$ in water solution at 10

(5) Miller, M. H.; Scheraga, H. A. *J. Polym. Sci.* **1976**, Symp. No. 54, 171−200.
(6) Long, C. G.; Braswell, E.; Zhu, D.; Apigo, J.; Baum, J.; Brodsky, B. *Biochemistry* **1993**, *32*, 11688−11695.
(7) (a) Aue, W. P.; Bartholdi, E.; Ernst, R. R. *J. Chem. Phys.* **1976**, *64*, 2229−2246. (b) Bax, A.; Freeman, R. *J. Magn. Reson.* **1981**, *44*, 542−561. (c) Rance, M.; Sorensen, O. W.; Bodenhausen, G.; Wagner, G.; Ernst, R. R.; Wuthrich, K. *Biochem. Biophys. Res. Commun.* **1984**, *117*, 479−485. (d) Shaka, B.; Freeman, R. *J. Magn. Reson.* **1983**, *51*, 169−173.
(8) (a) Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. *J. Chem. Phys.* **1979**, *71*, 4546−4553. (b) Kumar, A.; Ernst, R. R.; Wuthrich, K. *Biochem. Biophys. Res. Commun.* **1980**, *95*, 1−6. (c) Otting, G.; Widmer, H.; Wagner, G.; Wuthrich, K. *J. Magn. Reson.* **1986**, *66*, 187−193. (d) Wider, G.; Macura, S.; Kumar, A.; Ernst, R. R.; Wuthrich, K. *J. Magn. Reson.* **1984**, *56*, 207−234.
(9) (a) Braunsweiler, L.; Ernst, R. R. *J. Magn. Reson.* **1983**, *53*, 521−528. (b) Bax, A.; Davis, D. *J. Magn. Reson.* **1985**, *65*, 355−360.
(10) Bothner-By, A. A.; Stephens, R. L.; Lee, J.; Warren, C. D.; Jeanloz, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 811−813.
(11) (a) Redfield, A. G.; Kuntz, S. D. *J. Magn. Reson.* **1975**, *19*, 250−254. (b) Marion, D.; Wuthrich, K. *Biochem. Biophys. Res. Commun.* **1983**, *113*, 967−974.
(12) Dietrich, W.; Rudel, C. H.; Neumann, M. *J. Magn. Reson.* **1991**, *91*, 1−11.
(13) Hagler, A. T. In *The Peptides*; Udenfriends, S., Meienhofer, J., Hruby, V. J., Eds.; Academic Press: Orlando, FL, 1985; Vol. 7, pp 214−296.
(14) McCammon, J. A.; Wolynes, P. G.; Karplus, M. *Biochemistry* **1979**, *18*, 927−942.
(15) Quantum Chemistry Program Exchange No. 590, Indiana University Department of Chemistry.
(16) Miller, M. H.; Nemethy, G.; Scheraga, H. A. *Macromolecules* **1980**, *13*, 470−478.
(17) Baker, E. N.; Hubbard, R. E. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97−179.
(18) Polinsky, A.; Cooney, M. G.; Toy-Palmer, A.; Osapay, G.; Goodman, M. *J. Med. Chem.* **1992**, *35*, 4185−4194.

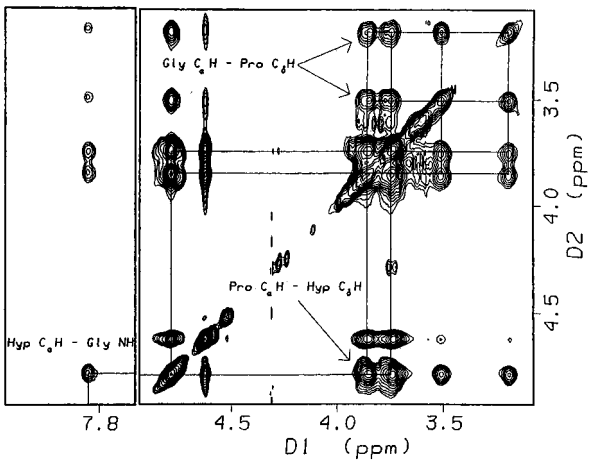

Figure 2. Amide and aliphatic region of the NOESY spectrum (mixing time 300 ms) of **KTAg-5,3** in D_2O at pH 3.4 ± 0.1 and 5 °C. The sequential NOEs for the assembled set are indicated.

°C,[4,19] and (Pro-Hyp-Gly)$_{10}$ is triple helical according to several spectroscopic and X-ray studies.[20] Second, for **Ac-6**, **Ac-9**, **KTAg-5,3**, and **KTAg-6,3** in D_2O (5 °C, pH 3.4 ± 0.1) the Gly NH of this resonance set has a low hydrogen-exchange rate ($k \leq 10^{-4}$ min^{-1} for **Ac-6** and **Ac-9** and $k \leq 10^{-5}$ min^{-1} for **KTAg-5,3** and **KTAg-6,3**; for **Ac-5** and **KTAg-3,3** the hydrogen-exchange rate is too fast to allow a reliable measurement of the corresponding kinetic constant), indicating shielding from the solvent as expected on the basis of the triple-helix hydrogen bond network.[16] Third, the NOESY crosspeaks for this set of resonances are consistent with a triple-helical structure as explained below.

For a proper analysis of NOESY spectra it is essential to consider that in triple-helical molecules three polypeptide chains are symmetrically and closely packed, and these polypeptide chains have identical repetitive sequences. A given NOESY crosspeak may therefore result from interactions between atoms of the same polypeptide chain and/or from interactions between atoms of two different polypeptide chains. For instance, as shown in Figure 1, the NOE between Hyp $C_\alpha H$ and Gly NH can result from either interchain or intrachain interactions.[4] Because of these complications it is difficult to proceed to a direct structure determination of triple-helical compounds using NOE-based distance constraints. However, the inverse process can be accomplished as follows: given a triple-helix model built on the basis of other structural techniques, such as X-rays, it is possible to predict which NOEs arise uniquely from interchain interactions (distances smaller than 4.5 Å) and use these NOEs to test the proposed model critically.

In an elegant and pioneering paper Brodsky *et al.*[4] identified the NOEs which are predicted to be interchain on the basis of the X-ray fiber diffraction triple-helical model of sequences containing Gly-Pro-Hyp repeats[21] and which can be used to test the triple-helical structure of (Pro-Hyp-Gly)$_{10}$ in water.[4] These NOEs are Gly NH−Pro $C_\delta H_1$, Pro $C_\gamma H_1$−Hyp $C_\beta H_1$, Pro $C_\gamma H_1$−Hyp $C_\beta H_2$, Pro $C_\gamma H_1$−Hyp $C_\gamma H$, Pro $C_\delta H_1$−Hyp $C_\beta H_1$, Pro $C_\delta H_1$−Hyp $C_\beta H_2$, Pro $C_\delta H_2$−Hyp $C_\beta H_2$, and Pro $C_\delta H_1$−Hyp $C_\gamma H$ (the hydrogens are named according to ref 4). These NOEs are all unambiguously observed for the major set of resonances in the NOESY spectra of **Ac-5**, **Ac-6**, **Ac-9**, **KTAg-5,3**, and **KTAg-6,3** at 5 °C and pH 3.4 ± 0.1 (for **KTAg-3,3** see ref 2). These observations confirm that this resonance set corresponds to the triple-helical conformation in solution and suggest that the triple-helical structure in solution is similar to that proposed on the basis of X-ray studies.[21] We therefore denote this set of resonances as the *assembled* set, and accordingly we define the *unassembled* sets as those observed for **Ac-1** and **KTAg-1,3**, which lack the structural requirements to form a triple helix and therefore assume less ordered conformations.

The equatorial and axial protons of the KTA methylene could be assigned on the basis of the known characteristics of cyclohexane derivatives in which the equatorial methylene protons have higher resonance frequencies than the axial protons.[22] Table 1 shows the assignments for the assembled and unassembled sets of resonances. It is notable that among the resonances of the unassembled *trans* set of the acetyl analogs, that of Gly $C_\alpha H_l$ at 4.25−4.16 ppm is not overlapped by any resonance of the other sets and can therefore be used to detect non-triple-helical conformations. However, in the case of the KTA analogs this resonance can be overlapped by a $C_\alpha H$ resonance of a spacer Gly. It is also notable that among the resonances of the assembled set, that of Pro $C_\delta H_h$ at 3.2 ppm is well resolved and not overlapped by any resonance belonging to the unassembled sets. The resonance at 3.2 ppm can therefore be used to identify and quantify triple-helical conformations.[19] We employed this resonance at 3.2 ppm to study the triple helicity as a function of chain length, template, and temperature as discussed below.

Triple Helicity as a Function of Chain Length and Template. Figure 3a shows the 1D spectral region containing the assembled Pro $C_\delta H_h$ and the unassembled Gly $C_\alpha H_l$ resonances of the acetyl analogs **Ac-1**, **Ac-3**, **Ac-5**, and **Ac-6** at 5 °C. It can be seen that only when there are at least five triplets per polypeptide chain the triple-helical Pro $C_\delta H_h$ resonance at 3.2 ppm appears and the intensity of the unassembled Gly $C_\alpha H_l$ at 4.3−4.2 ppm decreases. Also the spectrum of **Ac-9** confirms this trend. Figure 3b shows the same 1D spetral region as Figure 3a but for the KTA-terminated analogs. The assembled Pro $C_\delta H_h$ resonance at 3.2 ppm shows that in the KTA series three triplets (**KTAg-3,3**) are sufficient for the formation of an assembled structure.

A schematic model of the triple-helical array of **KTAg-3,3** is shown in Figure 1. Figure 1 shows that the register shift of each polypeptide chain along the triple-helix axis leads to the conclusion that at least one Gly-Pro-Hyp triplet per chain is only partially included in the triple-helix packing. This end effect is present also for **KTAg-5,3** and **KTAg-6,3**. However in these analogs it affects only one triplet out of five and six, respectively, whereas in **KTAg-3,3** one out of three Gly-Pro-Hyp triplets is affected. We therefore expect that the resonance dispersion induced by the triple-helix end effect is more significant in **KTAg-3,3** than in **KTAg-5,3** or **KTAg-6,3**. This interpretation is fully consistent with the broadening of the Pro $C_\delta H_h$ signal at 3.2 ppm observed for **KTAg-3,3** as compared to **KTAg-5,3** and **KTAg-6,3** (Figure 3b). In this regard, it is also interesting to note that for **KTAg-3,3** the interchain NOE Gly NH−Pro $C_\delta H_1$ was not observed indicating that structural distortions occur in **KTAg-3,3** as compared to the longer chain analogs for which this NOE was observed.[2]

Triple Helicity as a Function of Temperature. We monitored the melting of the triple-helical conformation of **Ac-5**, **Ac-6**, **KTAg-5,3**, and **KTAg-6,3** using the normalized integral of the assembled Pro $C_\delta H_h$ resonance at 3.2 ppm, which indicates the average number of collagen-like triple-helical Pro

(19) Brodsky, B.; Li, M.-H.; Long, G. G.; Apigo, J.; Baum, J. *Biopolymers* **1992**, *32*, 447−451.

(20) Sakikabara, S.; Inouye, K.; Shudo, K.; Kishida, Y.; Kobayashi, Y.; Prockop, D. J. *Biochim. Biophys. Acta* **1973**, *303*, 198−202.

(21) Fraser, R. D. B.; MacRae, T. P.; Suzuki, E. *J. Mol. Biol.* **1979**, *129*, 463−481.

(22) Morelle, N.; Gharbi-Benarons, J.; Acher, F.; Valle, G.; Crisma, M.; Toniolo, C.; Azerad, R.; Girault, J.-P. *J. Chem. Soc., Perkin Trans.* **1993**, *2*, 525−533.

Table 1. Chemical Shifts of Assigned Proton Resonances of Representative Compounds[a] in Water at 5 °C,[b] pH 3.4 ± 0.1

			chemical shift (ppm)[c]			
resonance set	residue	atom	**Ac-1**	**Ac-9**[d]	**KTAg-1,3**	**KTAg-6,3**[d]
assembled	Gly	NH		7.94		7.95
		$C_\alpha H_{l,h}$		3.85, 3.75		3.86, 3.76
	Pro	$C_\alpha H$		4.79		4.80
		$C_\beta H_{l,h}$		2.31, 1.91		2.31, 1.92
		$C_\gamma H_2$		1.96		1.96
		$C_\delta H_{l,h}$		3.52, 3.20		3.52, 3.20
	Hyp	$C_\alpha H$		4.80		4.80
		$C_\beta H_{l,h}$		2.18, 2.03		2.18, 2.02
		$C_\gamma H$		4.63		4.63
		$C_\delta H_{l,h}$		3.86, 3.75		3.86, 3.76
	Ac	Me		∼2.05		
	KTA	H_{ax}				1,27, 1.31, 1.30
		H_{eq}				2.88, 2.62, 2.47
		Me				1.25, 1.30, 1.19
unassembled (*trans*)	Gly	NH	8.35	8.60[e]	8.33	8.58[e]
		$C_\alpha H_{l,h}$	4.16, 4.01	4.27, 3.96[f]	4.12, 4.03	4.27, 3.96[f]
	Pro	$C_\alpha H$	4.77	4.79	4.76	4.80
		$C_\beta H_{l,h}$	2.32, 1.93	2.31, 1.92	2.31, 1.94	2.32, 1.93
		$C_\gamma H_2$	2.05	2.04	2.04	2.05
		$C_\delta H_{l,h}$	3.63, 3.63	3.67, 3.61	3.61, 3.61	3.67, 3.62
	Hyp	$C_\alpha H$	4.53	4.60[g]	4.50	4.61[g]
		$C_\beta H_{l,h}$	2.37, 2.08	2.37, 2.06	2.36, 2.07	2.37, 2.07
		$C_\gamma H$	4.63	4.61	4.62	4.61
		$C_\delta H_{l,h}$	3.94, 3.82	3.89, 3.82	3.90, 3.81	3.88, 3.82
	Ac	Me	2.05			
	KTA	H_{ax}			1.33	
		H_{eq}			2.66	
		Me			1.24	
unassembled (*cis*)	Gly	NH	8.34		8.34	
		$C_\alpha H_{l,h}$	4.04, 3.62		4.02, 3.66	
	Pro	$C_\alpha H$	4.95		4.97	
		$C_\beta H_{l,h}$	2.48, 2.20		2.45, 2.18	
		$C_\gamma H_{l,h}$	2.01, 1.87		1.99, 1.83	
		$C_\delta H_2$	3.56		3.56	
	Hyp	$C_\alpha H$	4.58		4.58	
		$C_\beta H_{l,h}$	2.42, 2.13		2.38, 2.06	
		$C_\gamma H$	4.66		4.64	
		$C_\delta H_2$	3.79		3.79	
	Ac	Me				
	KTA	H_{ax}				
		H_{eq}				
		Me				

[a] The spectra of the compounds not listed in this table include the same sets of resonances (see text). [b] The chemical shifts for the assembled (*cis*) set of resonances were measured at 27 °C in order to avoid the overlap of the Pro $C_\alpha H$ by the residual water signal. [c] ^{1}H-NMR chemical shifts are reported relative to sodium 2-(trimethylsilyl)tetradeuteriopropionate (STP). [d] For this compound the population of the unassembled (*cis*) set of resonances is too low to be detectable under our experimental conditions. [e] A minor resonance can be observed also at 8.34 ppm as the result of differences between residues in the core and residues at the ends of the polypeptide chain. [f] Minor resonances can be observed also at 4.17 and 4.01 ppm as the result of differences between residues in the core and residues at the ends of the polypeptide chain. [g] A minor resonance can be observed also at 4.53 ppm as the result of differences between residues in the core and residues at the ends of the polypeptide chain.

residues per chain (Figure 4). Figure 4 shows that the melting temperature of **Ac-5** is 26 °C, while that of **KTAg-5,3** is 70 °C. Similarly Figure 4 shows that the melting temperature of **Ac-6** is 42 °C and that of **KTAg-6,3** is at least 33 °C higher. The 1D integrals of Figure 4 also indicate that at 5 °C the average number of triple-helical Pro residues per chain is 4 for **Ac-5** and **KTAg-5,3**, and 5 for **Ac-6** and **KTAg-6,3**. For these four molecules the average number of assembled Pro residues per chain is therefore equal to the number of triplets per chain minus 1, as expected for the longest triple-helical structure allowed by the chain length. In this elongated triple helix one triplet per chain is only partially packed in the triple-helical array as a consequence of the register shift. For the **KTAg-3,3** analog the melting could not be quantitatively monitored using NMR because the $C_\delta H_h$ resonance of the assembled Pro not only changes in intensity as the temperature is increased but also shifts downfield (see Supporting Information).

The studies of the melting transitions monitored by NMR spectra are fully consistent with the results obtained using CD, UV, and optical rotation spectroscopic methods.[3] The CD spectra and optical rotation melting curves indicate that **Ac-5** and **Ac-6** are triple helical in water and that at a concentration of 0.2 mg/mL their melting temperature are 18 and 36 °C, respectively. The melting temperatures are 8 and 6 °C lower than the corresponding values measured by NMR, and this difference is accounted for by the different concentration ranges used for optical rotation and NMR experiments.[23] The CD spectra also show that **KTAg-3,3**, **KTAg-5,3**, and **KTAg-6,3** are triple helical in water at 20 °C as we observed by NMR. The melting transitions monitored by optical rotation and UV for the KTA analogs are also consistent with our NMR results.

Effect of the Triple Helix on the KTA Template. 1. Triple-Helix Register and KTA Symmetry. The ensemble of conformations generated for **KTAg-3,3** as described in the Materials and Methods section was clustered according to the torsional angles of the KTA cyclohexane ring. The clustering resulted in three conformational families, and Figure 5 shows the lowest energy structure of each cluster. It is interesting to note that in all the clustered structures the KTA ring is consistently bent out of the triple-helix axis in order to maintain

(23) Engel, J.; Chen, H.-T.; Prockop, D. J.; Klump, H. *Biopolymers* **1977**, *16*, 601−622.

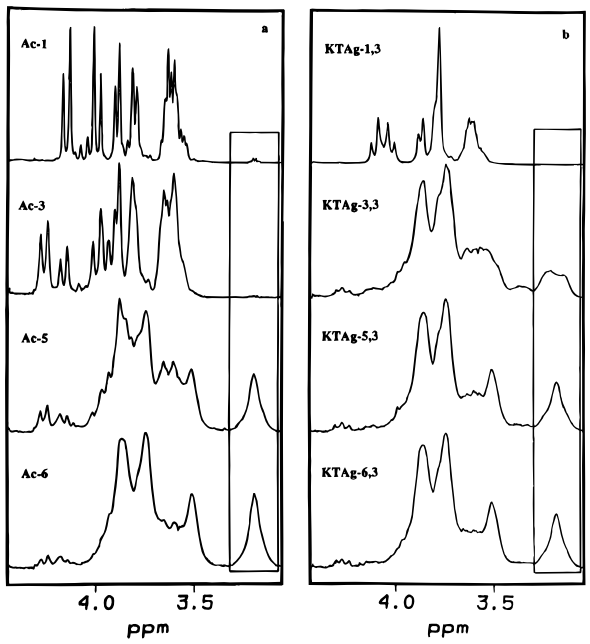

Figure 3. 1D ^{1}H-NMR spectral regions including the assembled Pro $C_\delta H_h$ signal at 3.2 ppm (boxed peaks) and the unassembled Gly $C_\alpha H_l$ signal at 4.3−4.2 ppm. All the spectra were acquired at pH 3.4 ± 0.1 and 5 °C.

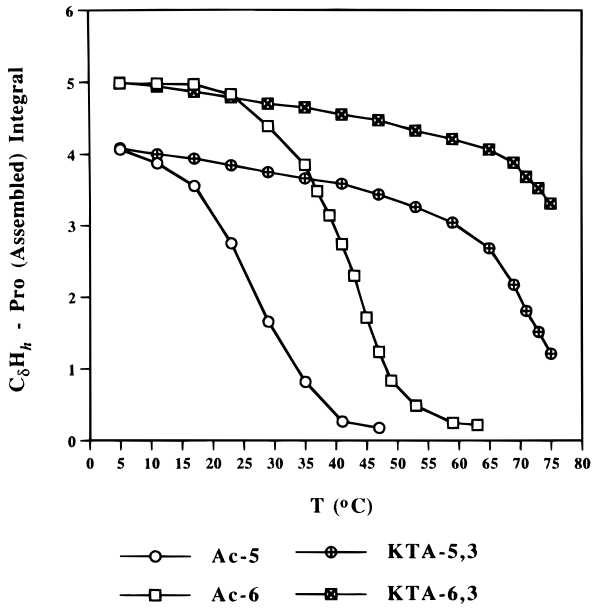

Figure 4. Thermal melting curves monitored by ^{1}H-NMR for the analogs **Ac-5**, **Ac-6**, **KTAg-5,3**, and **KTAg-6,3**. The vertical axis reports the normalized integral of the Pro $C_\delta H_h$ signal in the assembled set (see Materials and Methods for the details of the normalization). This integral represents the average number of proline residues per chain in a triple-helical environment at a given temperature.

the register shift imposed by the triple-helical array. This conformational feature is qualitatively outlined in Figure 1.

The schematic model shown in Figure 1 has important implications for the KTA template. Specifically, Figure 1 shows that the screw symmetry of the triple-helix coiled coils and the consequent one-residue register shift along the triple-helix axis break the ternary rotational symmetry of the KTA template. A positive proof of the loss of KTA symmetry as a result of triple-helix formation comes from the TOCSY spectrum of **KTAg-3,3** at 5 °C. At this temperature the triple helix is not denatured, and three distinct crosspeaks can be observed between the equatorial and axial protons of the KTA methylenes (Figure 6a). However when the triple helix is absent either because it has been denatured (Figure 6b) or because the chain is too short as for **KTAg-1,3** (Figure 6c), then the KTA signal splitting disappears indicating that the ternary rotational symmetry of KTA is maintained. Furthermore the same KTA splitting shown in Figure 6a could also be observed for **KTAg-5,3** and **KTAg-6,3**. On the basis of these observations we conclude that the KTA signal splitting can be considered as additional support of triple helicity for the KTA-terminated analogs.

2. KTA Conformation. Further insight into the conformation of the KTA ring can be obtained from the differences in methylene chemical shifts, which are known to be very sensitive to the conformational preferences of KTA analogs.[1] Using rigid, known structures it has been shown that when the three methyls are equatorial and the three carbonyls are axial the difference in methylene chemical shift is 1.3−1.6 ppm and the equatorial methylene proton is anomalously deshielded at 2.7 ppm by the anisotropy of the carbonyl groups. Conversely, when the three carbonyls are equatorial, the difference in the methylene chemical shift is just 0.6−0.8 ppm, as was also seen for simple cyclohexane derivatives,[24] and no anomalous deshielding is expected for the equatorial methylene proton which is observed at 2.0−2.2 ppm.[1]

For **KTAg-1,3** the difference in the methylene chemical shift is 1.34 ppm and the equatorial methylene proton is found at 2.66 ppm indicating that the KTA template adopts a chairlike conformation in which the three methyls are equatorial and the three carbonyls axial. For the other KTA conjugates, **KTAg-3,3**, **KTAg-5,3**, and **KTAg-6,3**, which are triple helical and show the KTA signal splitting, the difference in methylene chemical shift is 1.15−1.56 ppm and the equatorial methylene proton is found at 2.47−2.87 ppm indicating that the KTA template adopts a conformation similar to that proposed for **KTAg-1,3**. These conclusions are fully consistent with previous investigations on the Kemp triacid which was assigned a chair conformation for its protonated, monoionic, and dianionic forms with the three methyls equatorial.[1]

Effect of the KTA on the Triple Helix. The compensation for the different spacings of KTA and of collagen-like triple helices[25] allowed by the Gly spacers and the maintenance of the triple-helix register shift allowed by the KTA bending discussed above indicate that the triple-helix structure of our KTA conjugates is not significantly distorted as compared to canonical collagen-like triple helices. This conclusion is also supported by the similarity in chemical shift, NOE patterns, and amide hydrogen-exchange rates observed between the triple-helical acetyl analogs and the KTA conjugates.

The absence of significant structural distortions in the KTA-assembled triple helices suggests that the KTA template does not significantly alter the enthalpy of triple-helix formation. On the other hand the KTA template decreases the entropy loss associated with triple-helix formation which is an intramolecular process in the KTA analogs while it is an intermolecular process in the acetyl compounds. The KTA template leads therefore to a significant entropy gain without any significant enthalpy loss. Thus the KTA template allows a net free energy gain as confirmed by the induction of triple helicity in **KTAg-3,3** (Figure 3) and by the effect of the KTA on the melting temperatures (Figure 4).

(24) Jackman, L. M.; Sternhell, S. *Applications of Nuclear Magnetic Resonance Spectroscopy in Organic Chemistry*, 2nd ed.; Pergamon: Oxford, 1969; pp 238−240.

(25) Chen, J. M.; Kung, C. E.; Feairheller, S. H.; Brown, E. M. *J. Protein Chem.* **1991**, *10*, 535.

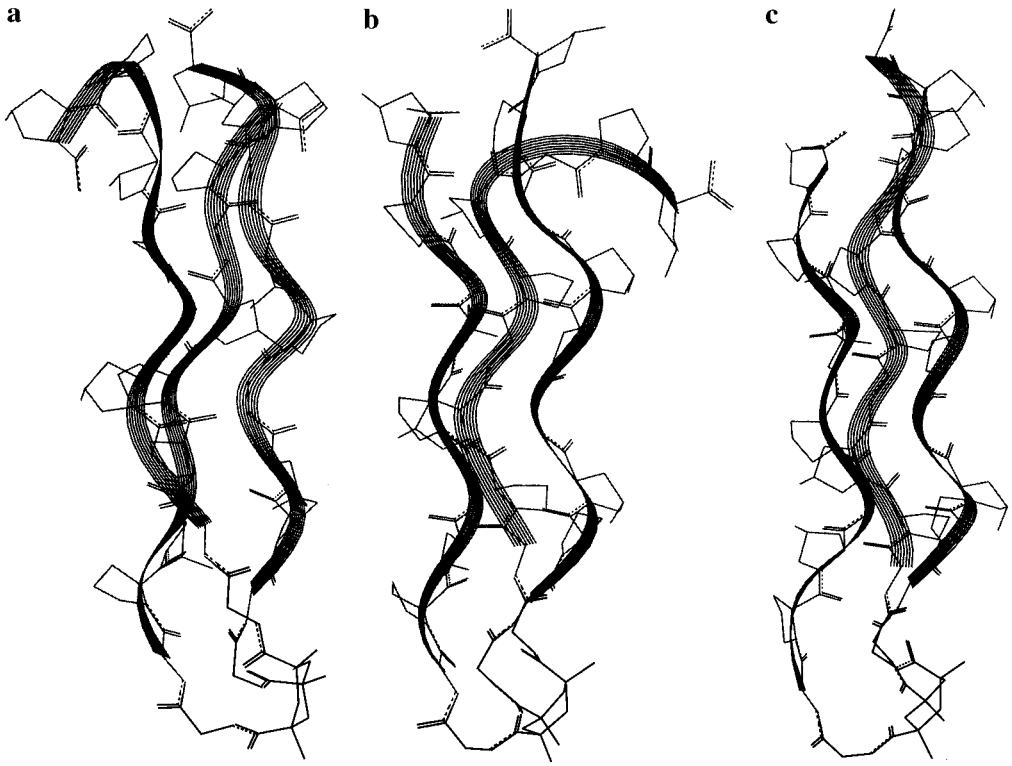

Figure 5. Lowest energy structures of the conformational families found for **KTAg-3,3**. The three families differ mainly for the conformation of the KTA ring. For each structure the KTA template is in the lower part of the figure. It can be seen that the KTA ring is consistently bent out of the triple-helix axis.

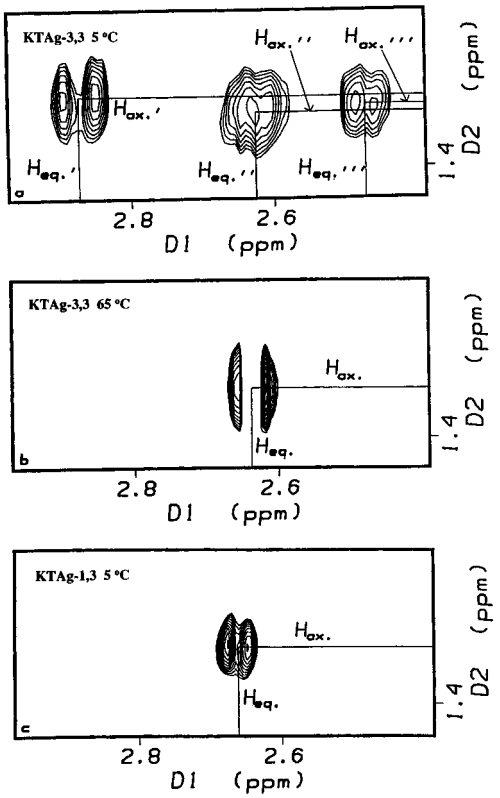

Figure 6. (a) Axial−equatorial region of the TOCSY spectrum of **KTAg-3,3** at pH 3.4 ± 0.1 and 5 °C. (b) Same as in panel a but at 65 °C. At this temperature the assembled structure of **KTAg-3,3** is fully denatured. (c) Axial−equatorial region of the TOCSY spectrum of **KTAg-1,3** at pH 3.4 ± 0.1 and 5 °C.

Conclusions

We have employed ^{1}H-NMR and molecular modeling to study the effect of collagen-like triple helicity on the KTA template and the effect of the KTA template on triple helicity in four KTA-terminated analogs of different chain lengths. With regard to the effect of triple helicity on the KTA template, we noticed that the screw symmetry of the coiled-coil triple helix breaks the ternary rotational symmetry of the KTA template. The loss of ternary rotational symmetry of KTA results in the splitting of the KTA TOCSY crosspeak into three distinct subgroups. This KTA signal splitting can be used as evidence for the triple helicity of the KTA compounds.

As to the effect of the KTA template on triple helicity, our studies indicate that the KTA template leads to a net gain in the free energy of triple-helix formation. This free energy gain results in the induction of an assembled collagen-like structure in the compound containing only three triplets per chain, **KTAg-3,3**, and in increased thermal stability: the melting temperature of **KTAg-5,3** is 44 °C higher than that of **Ac-5**, and similarly the melting temperature of **KTAg-6,3** is more than 33 °C higher than that of **Ac-6**. We conclude that the KTA template facilitates the structural requirements for triple-helix formation and opens new design opportunities for collagen mimetics, such as the introduction of novel sequences different from Gly-Pro-Hyp. We are now exploring these possibilities.

Acknowledgment. This project was funded by grants from the National Science Foundation (NSFDNR-9201133) and CIBA-Vision Inc. Dedicated to the memory of Professor Mario Farina.

Supporting Information Available: NMR and molecular modeling methods, assignment, TOCSY spectrum of the assembled set, and 1D spectra of **KTAg-3,3** at different temperatures (6 pages). See any current masthead page for ordering and Internet access instructions.

JA9612615