

## Effect of 3-Hydroxyproline Residues on Collagen Stability

Cara L. Jenkins,<sup>†</sup> Lynn E. Bretscher,<sup>‡,§</sup> Ilia A. Guzei,<sup>||</sup> and Ronald T. Raines<sup>\*,†,‡</sup>

Contribution from the Department of Chemistry and Department of Biochemistry,  
University of Wisconsin-Madison, Madison, Wisconsin 53706

Received January 2, 2003; E-mail: raines@biochem.wisc.edu

**Abstract:** Collagen is an integral part of many types of connective tissue in animals, especially skin, bones, cartilage, and basement membranes. A fibrous protein, collagen has a triple-helical structure, which is comprised of strands with a repeating Xaa-Yaa-Gly sequence. L-Proline (Pro) and 4(R)-hydroxy-L-proline (4-Hyp) residues occur most often in the Xaa and Yaa positions. The 4-Hyp residue is known to increase markedly the conformational stability of a collagen triple helix. In natural collagen, a 3(S)-hydroxy-L-proline (3-Hyp) residue occurs in the sequence: 3-Hyp-4-Hyp-Gly. Its effect on collagen stability is unknown. Here, two host-guest peptides containing 3-Hyp are synthesized: (Pro-4-Hyp-Gly)<sub>3</sub>-3-Hyp-4-Hyp-Gly-(Pro-4-Hyp-Gly)<sub>3</sub> (peptide 1) and (Pro-4-Hyp-Gly)<sub>3</sub>-Pro-3-Hyp-Gly-(Pro-4-Hyp-Gly)<sub>3</sub> (peptide 2). The 3-Hyp residues in these two peptides diminish triple-helical stability in comparison to Pro. This destabilization is small when 3-Hyp is in the natural Xaa position (peptide 1). There, the inductive effect of its 3-hydroxyl group diminishes slightly the strength of the interstrand 3-HypC=O...H-NGly hydrogen bond. The destabilization is large when 3-Hyp is in the nonnatural Yaa position (peptide 2). There, its pyrrolidine ring pucker leads to inappropriate mainchain dihedral angles and interstrand steric clashes. Thus, the natural regioisomeric residues 3-Hyp and 4-Hyp have distinct effects on the conformational stability of the collagen triple helix.

### Introduction

Collagen is the most abundant protein in animals, comprising approximately one-third of the total protein by weight.<sup>1,2</sup> Collagen has a unique tertiary structure, which consists of three parallel left-handed polyproline-II-type strands wound tightly around a common axis.<sup>3–6</sup> Maintenance of this triple helix is essential for many biological functions.<sup>7</sup>

Each strand of collagen consists of the repeating sequence Xaa-Yaa-Gly, where Xaa and Yaa are often proline (Pro) residues. The post-translational hydroxylation of some of these proline residues generates 4(R)-hydroxy-L-proline (4-Hyp) residues and, to a lesser extent, 3(S)-hydroxy-L-proline (3-Hyp) residues. 3-Hyp, which is found in the natural triplet (3-Hyp-4-Hyp-Gly),<sup>8</sup> arises from the action of the enzyme prolyl-3-hydroxylase<sup>9</sup> rather than by regiochemical ambiguity by the enzyme prolyl-4-hydroxylase.<sup>10</sup>

In seminal work, Prockop and co-workers demonstrated that 4-Hyp residues enhance greatly the conformational stability of

the collagen triple helix.<sup>11</sup> This additional stability arises from stereoelectronic effects that preorganize a 4-Hyp residue in a conformation that befits a triple helix.<sup>12–15</sup> In contrast to 4-Hyp, no information is available on the effect of 3-Hyp residues on collagen stability.

Here, we synthesize collagen-like peptides that contain a 3-Hyp residue. We use these peptides to assess the contribution of 3-Hyp to the conformational stability of the collagen triple helix. We find that the presence of a 3-Hyp residue, in surprising contrast to a 4-Hyp residue, *destabilizes* a collagen triple helix. Like the stability endowed by 4-Hyp residues, the instability imposed by 3-Hyp residues appears to arise largely from inductive effects of its pendant hydroxyl group.

### Experimental Section

**General.** Reagents were obtained from Aldrich Chemical (Milwaukee, WI) or Fisher Scientific (Hanover Park, IL) and used without further purification. Amino acids and their derivatives were obtained from Fisher Scientific, Bachem Bioscience (King of Prussia, PA), or Novabiochem (San Diego, CA). Dichloromethane was distilled over

<sup>†</sup> Department of Chemistry, University of Wisconsin-Madison.

<sup>‡</sup> Department of Biochemistry, University of Wisconsin-Madison.

<sup>§</sup> Present address: Department of Biochemistry, Medical College of Wisconsin, Milwaukee, WI 53226.

<sup>||</sup> To whom correspondence should be addressed regarding X-ray diffraction analysis. Molecular Structure Laboratory, Department of Chemistry, University of Wisconsin-Madison, 1101 University Avenue, Madison, WI 53706-1396.

(1) Prockop, D. J.; Kivirikko, K. I. *Annu. Rev. Biochem.* **1995**, *64*, 403–434.

(2) Myllyharju, J.; Kivirikko, K. I. *Ann. Med. (Helsinki)* **2001**, *33*, 7–21.

(3) Ramachandran, G. N.; Kartha, G. *Nature* **1954**, *269*–270.

(4) Ramachandran, G. N.; Kartha, G. *Nature* **1955**, *176*, 593–595.

(5) Rich, A.; Crick, F. H. C. *Nature* **1955**, *176*, 915–916.

(6) Rich, A.; Crick, F. H. C. *J. Mol. Biol.* **1961**, *3*, 483–506.

(7) Fields, G. B. *Connect. Tissue Res.* **1995**, *31*, 235–243.

(8) Gryder, R. M.; Lamon, M.; Adams, E. *J. Biol. Chem.* **1975**, *250*, 2470–2474.

(9) The crystalline structure of a bacterial proline 3-hydroxylase has been determined by X-ray diffraction analysis (Clifton, I. J.; Hsueh, L.-C.; Baldwin, J. E.; Harlos, K.; Schofield, C. J. *Eur. J. Biochem.* **2001**, *268*, 6625–6636). The product of its catalysis is the amino acid 3(R)-hydroxy-L-proline, which has not been found in natural collagen.

(10) Tryggvason, K.; Risteli, J.; Kivirikko, K. *Biochem. Biophys. Res. Commun.* **1977**, *76*, 275–281.

(11) Berg, R. A.; Prockop, D. J. *Biochem. Biophys. Res. Commun.* **1973**, *52*, 115–120.

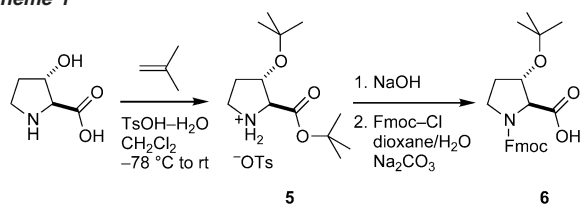
(12) Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E.; Raines, R. T. *Nature* **1998**, *392*, 666–667.

(13) Engel, J.; Prockop, D. J. *Matrix Biol.* **1998**, *17*, 679–680.

(14) Bretscher, L. E.; Jenkins, C. L.; Taylor, K. M.; DeRider, M. L.; Raines, R. T. *J. Am. Chem. Soc.* **2001**, *123*, 777–778.

(15) Jenkins, C. L.; Raines, R. T. *Nat. Prod. Rep.* **2002**, *19*, 49–59.

Scheme 1



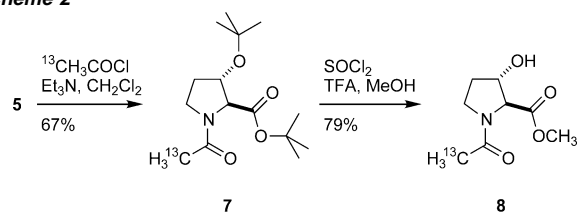
CaH<sub>2</sub>(s) or drawn from a Baker Cycletainer. Thin-layer chromatography was performed by using aluminum-backed plates coated with silica gel containing F<sub>254</sub> phosphor and visualized by UV illumination or staining with I<sub>2</sub>, *p*-anisaldehyde stain, or phosphomolybdic acid stain. NMR spectra were obtained with Bruker AC-300 and Varian UNITY-500 spectrometers. Mass spectra were obtained with a Micromass LCT ESI or Perkin–Elmer Voyager MALDI-TOF instrument.

**3(S)-tert-Butoxy-L-proline tert-Butyl Ester Tosylate (5).** *N*-(9H-Fluoren-9-ylmethoxycarbonyl)-3-tert-butoxy-L-proline (**6**) was synthesized by the route shown in Scheme 1. 2-Methyl-propene (approximately 10 mL) was condensed into a pear-shaped flask at  $-78\text{ }^{\circ}\text{C}$  and then added to a suspension of 3-hydroxyproline (0.50 g, 3.8 mmol) and *p*-toluenesulfonic acid hydrate (3.00 g, 15.8 mmol) in dichloromethane (25 mL) at  $-78\text{ }^{\circ}\text{C}$ . The reaction was stirred for 3 d, allowing the mixture to come to room temperature. The reaction mixture was then cooled to  $0\text{ }^{\circ}\text{C}$ , vented carefully, and then poured into a separatory funnel and washed twice with a saturated aqueous solution of NaHCO<sub>3</sub>. The combined aqueous layers were extracted once with dichloromethane, and the combined organic layers were then washed once with water, dried over MgSO<sub>4</sub>(s), filtered, and concentrated under reduced pressure to yield compound **5** as a yellow oil (1.142 g, 72%). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 7.78 (d, *J* = 8.1 Hz, 2H), 7.14 (d, *J* = 8.1 Hz, 2H), 4.29 (m, 1H), 3.99 (d, *J* = 2.9 Hz, 1H) 3.50–3.43 (m, 1H), 3.37–3.27 (m, 1H), 2.33 (s, 3H), 2.00–1.92 (m, 1H), 1.83–1.76 (m, 1H), 1.46 (s, 9H), 1.18 (s, 9H).

***N*-(9H-Fluoren-9-ylmethoxycarbonyl)-3(S)-tert-butoxy-L-proline (6).** Compound **5** (412 mg, 0.99 mmol) was suspended in 1.0 M NaOH (10 mL, 10 mmol) and stirred until the reaction mixture became homogeneous, that is, for about 4 h. The reaction mixture was then cooled to  $0\text{ }^{\circ}\text{C}$  and acidified with concentrated HCl to pH 0. Na<sub>2</sub>CO<sub>3</sub> (1.1 g) was added immediately, followed by dioxane (25 mL). Fmoc-Cl (292 mg, 1.13 mmol) was added in small portions, and the resulting mixture was stirred at room temperature for 4 h. The reaction mixture was then poured into a separatory funnel with 50 mL of H<sub>2</sub>O and extracted with diethyl ether (3  $\times$  25 mL). The aqueous layer was then combined with 50 mL of ethyl acetate in a beaker, cooled to  $0\text{ }^{\circ}\text{C}$ , and acidified with concentrated aqueous HCl to pH 0. The layers were separated, and the aqueous layer was extracted again with EtOAc (25 mL). The combined ethyl acetate layers were dried over MgSO<sub>4</sub>(s) and filtered, and the solvent was removed under reduced pressure to yield compound **6** as a white solid (0.148 g, 36.6% over two steps). <sup>1</sup>H NMR (two rotamers, 500 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 7.75 and 7.70 (d, *J* = 7 and 7.5 Hz, 2H), 7.60 and 7.55 (dd, *J* = 8, 8 Hz and 7.5, 10 Hz, 2H), 7.41–7.27 (m, 4H), 4.68–4.12 (m, 5H), 3.76–3.61 (m, 2H), 2.12–2.01 (m, 1H), 1.89–1.80 (m, 1H), 1.22 (s, 9H). <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>, two rotamers)  $\delta$  (ppm) 175.52, 174.12, 155.92, 154.62, 144.00, 143.84, 143.74, 141.26, 127.71, 127.57, 127.05, 126.96, 125.15, 125.06, 124.96, 119.95, 119.86, 75.26, 75.10, 73.67, 67.94, 67.62, 67.34, 66.05, 65.93, 61.15, 61.13, 47.12, 45.46, 45.30, 33.09, 32.23, 28.08. ESIMS (*m/z*) [*M* – H]<sup>+</sup> calculated, 408.1811; found, 408.1805.

To determine whether epimerization of the  $\alpha$ -carbon had occurred upon saponification of the *tert*-butyl ester **5**, the hydroxyl group was deprotected by acidolysis of its *tert*-butyl protecting group and the resulting hydroxyl group was converted into its Mosher ester. Briefly, Fmoc-3-HypOH (17 mg, 0.048 mmol) was dissolved in 1 mL of dioxane. K<sub>2</sub>CO<sub>3</sub> (20 mg, 0.14 mmol) was added to this solution, followed by H<sub>2</sub>O (2 drops). A Mosher acid chloride, *R*-(–)- $\alpha$ -methoxy-

Scheme 2



$\alpha$ -(trifluoromethyl)phenylacetyl chloride (0.010 mL, 0.053 mmol), was added, and the resulting mixture was stirred at room temperature for 15 min. EtOAc (1 mL) was then added to the reaction mixture, followed by concentrated HCl (2 drops) and H<sub>2</sub>O (1 mL). The layers were mixed thoroughly and then separated. The aqueous layer was extracted twice more with EtOAc, and the combined organic extracts were dried over MgSO<sub>4</sub>(s), filtered, and concentrated to a colorless oil. The <sup>19</sup>F NMR spectrum of the oil contained only a single peak at  $-72\text{ ppm}$ , indicating that compound **6** had maintained its stereochemical integrity.

***N*-(<sup>13</sup>C<sub>2</sub>-Acetyl)-3(S)-tert-butoxy-L-proline tert-Butyl Ester (7).** [<sup>13</sup>CH<sub>3</sub>]Ac-3-Hyp-OMe (**8**) was synthesized by the route shown in Scheme 2. The <sup>13</sup>C in the acetyl group facilitated the measurement of amide bond *trans/cis* ratios because both the acetyl and ester methyl groups overlap with other resonances in the <sup>1</sup>H spectrum. Compound **5** (771 mg, 1.86 mmol) was dissolved in 25 mL of dry dichloromethane and cooled to  $0\text{ }^{\circ}\text{C}$ . <sup>13</sup>CH<sub>3</sub>COCl (200  $\mu$ L, 2.78 mmol) was added, followed by triethylamine (800  $\mu$ L, 5.74 mmol). The reaction was stirred at  $0\text{ }^{\circ}\text{C}$  for 1 h and then at room temperature overnight. The reaction mixture was transferred to a separatory funnel and washed twice with water. The organic layer was dried over MgSO<sub>4</sub>(s), filtered, and concentrated under reduced pressure to a brown oil. The product was purified by silica gel chromatography, eluting with CH<sub>2</sub>Cl<sub>2</sub> containing MeOH (5% v/v) to yield compound **7** as a clear, colorless oil (355 mg, 66.8%). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) major isomer 4.28 (bs, 1H), 4.17 (m, 1H), 3.65 (m, 2H), 2.20–1.80 (m, 2H) 2.09 (d, <sup>1</sup>*J*<sub>CH</sub> = 132 Hz, 3H), 1.46 (s, 9H), 1.22 (s, 9H); minor isomer  $\delta$  (ppm) 4.30 (m, 1H), 4.07 (s, 1H), 3.65 (m, 2H), 2.20–1.80 (m, 2H), 1.96 (d, <sup>1</sup>*J*<sub>CH</sub> = 132 Hz, 3H), 1.48 (s, 9H), 1.24 (s, 9H). <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>, 8 scans)  $\delta$  (ppm) 22.10, 22.25 (<sup>13</sup>C-labeled carbon).

***N*-(<sup>13</sup>C<sub>2</sub>-Acetyl)-3(S)-hydroxy-L-proline Methyl Ester (8).** Compound **7** (355 mg, 1.24 mmol) was dissolved in MeOH. Trifluoroacetic acid (0.10 mL) was added, followed by a dropwise addition of SOCl<sub>2</sub> (1.00 mL, 13.7 mmol). The reaction mixture was heated at reflux for 1 h and then cooled to room temperature and concentrated under reduced pressure. The resulting oil was purified by chromatography on silica gel, eluting with CH<sub>2</sub>Cl<sub>2</sub> containing MeOH (5% v/v) to yield amide **8** as a clear, colorless oil (183 mg, 79%). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>, two rotamers)  $\delta$  (ppm) major isomer 4.43 (s, 1H), 4.38 (d, 4 Hz, 1H), 3.71–3.61 (m, 2H), 3.68 (s, 3H), 2.05 (d, <sup>1</sup>*J*<sub>CH</sub> = 128.5 Hz, 3H), 2.14–2.06 (m, 1H), 2.01–1.94 (m, 1H); minor isomer  $\delta$  (ppm) 4.50 (d, 4 Hz, 1H), 4.33 (s, 1H), 3.73 (s, 3H), 3.71–3.61 (m, 2H), 1.94 (d, <sup>1</sup>*J*<sub>CH</sub> = 128.5 Hz, 3H). <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>, two rotamers)  $\delta$  (ppm) 174.44, 174.03, 172.98, 75.08, 73.66, 69.46, 67.68, 54.12, 53.83, 46.72, 45.15, 32.64, 30.93, 21.84, 21.60.

**pK<sub>a</sub> Determinations.** The pK<sub>a</sub> values of the carboxylic acid groups of 3-HypOH and 4-HypOH were determined by <sup>1</sup>H NMR spectroscopy performed in aqueous solutions of different pH. D<sub>2</sub>O stock solutions consisted of 210 mM 3-HypOH (27.6 mg in 1.0 mL), 209 mM 4-HypOH (27.4 mg in 1.0 mL), 104 mM sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS, 40.8 mg in 1.0 mL), 100 mM D<sub>3</sub>PO<sub>4</sub> (68.5  $\mu$ L of 85% w/v D<sub>3</sub>PO<sub>4</sub> in 10.0 mL), 100 mM DCl (156  $\mu$ L of 20% DCl in 10.0 mL), and 100 mM NaOD (67.6  $\mu$ L of fresh 40% w/v NaOD in 10.0 mL). Each sample contained 49  $\mu$ L of the 3-Hyp or 4-Hyp stock solution, 1  $\mu$ L of the DSS stock solution, and 950  $\mu$ L of buffer, according to Tables S1 and S2 in the Supporting Information. The pH of each solution was measured by using a standard pH electrode calibrated in nondeuterated buffers. <sup>1</sup>H NMR spectra were obtained

(32 scans, 24 °C), and the chemical shift of the  $\alpha$ -proton was determined for each sample. The resulting pH and chemical shift data were fitted to eq 1 to yield the  $pK_a$  values.

$$\delta = \frac{\delta_{\text{low}} + \delta_{\text{high}} 10^{(\text{pH}-pK_a)}}{1 + 10^{(\text{pH}-pK_a)}} \quad (1)$$

The acidity of an aqueous solution of a molecule having a titratable group with a fraction factor near unity can be measured in D<sub>2</sub>O by using a standard pH meter equilibrated against H<sub>2</sub>O buffers and adding a correction factor of 0.4 units to the reading.<sup>16</sup> The difference in  $pK_a$ 's of amino acids in H<sub>2</sub>O and D<sub>2</sub>O is approximately 0.5 units.<sup>17</sup> Thus, a correction of about 0.1 units can be applied to a  $pK_a$  measurement obtained by measuring apparent pH values with a standard glass electrode. To determine a more precise correction factor, we measured the  $pK_a$  of 3-HypOH and 4-HypOH in D<sub>2</sub>O in the same manner and found that the 4-HypOH value was within 0.02 units of the value reported previously.<sup>18</sup> Thus, we did not apply any correction factor herein to the 3-Hyp  $pK_a$  value.

**Crystallization of *N*-(<sup>13</sup>C<sub>2</sub>-Acetyl)-3(S)-hydroxy-L-proline Methyl Ester (8).** Equal portions of a solution of amide **8** (67 mg) in 3.0 mL of dichloromethane were placed in six half-dram vials. Approximately 15 drops (Pasteur pipet) of a different cosolvent (hexanes, ethyl acetate, diethyl ether, tetrahydrofuran, *tert*-butyl methyl ether) or no cosolvent was added to each vial. The vials were then capped loosely and allowed to sit at room temperature undisturbed for several days. The vial with ethyl acetate as a cosolvent contained the crystals most suitable for X-ray diffraction analysis.

**X-ray Diffraction Analysis.** A colorless crystal with approximate dimensions 0.50 × 0.50 × 0.41 mm<sup>3</sup> was selected under oil under ambient conditions and attached to the tip of a glass capillary. The crystal was mounted in a stream of N<sub>2</sub>(g) at (173 ± 2) K and centered in the X-ray beam by using a video camera. Crystal evaluation and data collection were performed on a Bruker CCD-1000 diffractometer with Mo K $\alpha$  ( $\lambda$  = 0.710 73 Å) radiation and a diffractometer-to-crystal distance of 4.9 cm.

Initial cell constants were obtained from three series of  $\omega$  scans at different starting angles. Each series consisted of 20 frames collected at intervals of 0.3° in a 6° range about  $\omega$  with an exposure time of 10 s per frame. A total of 102 reflections were obtained. The reflections were indexed successfully by an automated indexing routine built into the SMART program. The final cell constants were calculated from a set of 1508 strong reflections from the actual data collection.

Data were collected by using the hemisphere data collection routine. The reciprocal space was surveyed to the extent of a full sphere to a resolution of 0.80 Å. A total of 2911 data sets were harvested by collecting one set of frames with 0.3° scans in  $\omega$  with an exposure time of 10 s per frame. This redundant data set was corrected for Lorentz and polarization effects. The absorption correction was based on fitting a function to the empirical transmission surface as sampled by multiple equivalent measurements.<sup>19</sup> All software and sources of the scattering factors were contained in the SHELXTL (version 5.1) program library (G. Sheldrick, Bruker Analytical X-ray Systems, Madison, WI).

**Synthesis of Peptides 1 and 2.** Peptide synthesis was performed on an Applied Biosystems Synergy synthesizer using standard Fmoc/*t*Bu methodology and HBTU as the coupling reagent. 3-Hyp residues were incorporated by using compound **6**.

Peptide **1**, which was (Pro-4-Hyp-Gly)<sub>3</sub>-3-Hyp-4-Hyp-Gly-(Pro-4-Hyp-Gly)<sub>3</sub>, was purified by reversed-phase HPLC on a C-18 column, eluting with a linear gradient (5–60% v/v) of solvent B, where solvent

A was H<sub>2</sub>O containing TFA (0.1% v/v) and solvent B was CH<sub>3</sub>CN containing TFA (0.1% v/v). MALDI MS ( $m/z$ ) [M + H]<sup>+</sup> calculated, 1905; found, 1906. Peptide **2**, which was (Pro-4-Hyp-Gly)<sub>3</sub>-Pro-3-Hyp-Gly-(Pro-4-Hyp-Gly)<sub>3</sub>, was purified likewise. MALDI MS ( $m/z$ ): [M + H]<sup>+</sup> calculated, 1889; found, 1890.

**Thermal Denaturation Experiments.** Values for  $T_m$  for each triple helix were determined in triplicate by thermal denaturation experiments monitored by CD spectroscopy on an Aviv 202 SF instrument equipped with an automated temperature controller. A 0.2 mM solution of peptide **1** or **2** in 50 mM HOAc was incubated at 4 °C for >24 h. Aliquots of 300  $\mu$ L were placed in 0.1-cm path length quartz cuvettes that had been equilibrated at 5 °C. Wavelength scans were performed from 200 to 260 nm at 5 and 50 °C, with a slit width of 1 nm and an averaging time of 3 s. Thermal denaturation experiments were performed by raising the temperature from 5 to 50 °C in 3 °C steps, equilibrating for 5 min at each temperature, and monitoring at 225 nm with a 20-s averaging time. Values of  $T_m$ , which is the temperature at the midpoint of the thermal transition, were determined by fitting the data to a two-state model using the software package NLREG v. 4.0 (Philip Sherrod).

**Molecular Modeling.** Molecular models of triple-helical **1** and **2** were created by modifying the structure of crystalline collagen (PDB entry 1CAG<sup>20</sup>). The atomic coordinates were imported into the program SYBYL (Tripos, St. Louis, MO). The side-chain methyl group of the three alanine residues were replaced with a hydrogen atom, and a section of the (Pro-Hyp-Gly) triple helix was excised from the C-terminus. This section was subjected to energy minimization with the Tripos force field. A hydroxyl group was added to C $\beta$  of the proline residue in the Xaa position of the central triplet in each strand (to mimic peptide **1**) or moved from C $\gamma'$  to C $\beta$  in the Yaa position of the central triplet in each strand (to mimic peptide **2**). The dihedral angles of the 3-Hyp residues were then altered to match those of the structure of crystalline amide **8**. The resulting structures were not subjected to further minimization.

## Results and Discussion

Synthetic collagen mimics have been effective in dissecting the basis for the conformational stability of the collagen triple helix.<sup>15,21</sup> In particular, host–guest studies have revealed important insights on the contribution of individual amino acid residues, both natural and nonnatural.<sup>22–27</sup> To reveal the effect of 3-Hyp residues on triple-helical stability, we synthesized two host–guest peptides in which the central triplet of (Pro-4-Hyp-Gly)<sub>7</sub> was replaced with one containing 3-Hyp. These peptides were (Pro-4-Hyp-Gly)<sub>3</sub>-3-Hyp-4-Hyp-Gly-(Pro-4-Hyp-Gly)<sub>3</sub> (**1**) and (Pro-4-Hyp-Gly)<sub>3</sub>-Pro-3-Hyp-Gly-(Pro-4-Hyp-Gly)<sub>3</sub> (**2**). Peptide **1** was designed to elucidate the role of the 3-Hyp-4-Hyp-Gly triplet as is found most often in natural collagen. Peptide **2** was designed to reveal the effect of the regiochemistry (3-Hyp versus 4-Hyp) of the pendant hydroxyl group. These peptides were synthesized by standard Fmoc/*t*Bu coupling strategies using compound **6** (Scheme 1) to introduce 3-Hyp residues.

**Conformational Stability.** The unfolding of triple helices of peptides **1** and **2** was monitored by CD spectroscopy. The

(16) Glasoe, P. K.; Long, F. A. *J. Phys. Chem.* **1960**, *64*, 188–190.

(17) Hyman, H. H.; Kaganove, A.; Katz, J. J. *J. Phys. Chem.* **1960**, *64*, 1653–1655.

(18) Fasman, G. D., Ed. *Practical Handbook of Biochemistry and Molecular Biology*; CRC Press: Boca Raton, FL, 1989.

(19) Blessing, R. H. *Acta Crystallogr.* **1995**, *A51*, 33–38.

(20) Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M. *Science* **1994**, *266*, 75–81.

(21) Fields, G. B.; Prockop, D. J. *Biopolymers* **1996**, *40*, 345–357.

(22) Shah, N. K.; Ramshaw, J. A. M.; Kirkpatrick, A.; Shah, C.; Brodsky, B. *Biochemistry* **1996**, *35*, 10262–10268.

(23) Ramshaw, J. A. M.; Shah, N. K.; Brodsky, B. *J. Struct. Biol.* **1998**, *122*, 86–91.

(24) Kwak, J.; Jefferson, E. A.; Bhumralkar, M.; Goodman, M. *Bioorg. Med. Chem.* **1999**, *7*, 153–160.

(25) Persikov, A. V.; Ramshaw, J. A. M.; Brodsky, B. *Biopolymers* **2000**, *55*, 436–450.

(26) Persikov, A. V.; Ramshaw, J. A. M.; Kirkpatrick, A.; Brodsky, B. *Biochemistry* **2000**, *39*, 14960–14967.

(27) Kersteen, E. A.; Raines, R. T. *Biopolymers* **2001**, *59*, 24–28.

**Table 1.** Values of  $T_m$  for Synthetic (Pro-Hyp-Gly)<sub>3</sub>-(Xaa-Yaa-Gly)-(Pro-Hyp-Gly)<sub>3</sub> Triple Helices

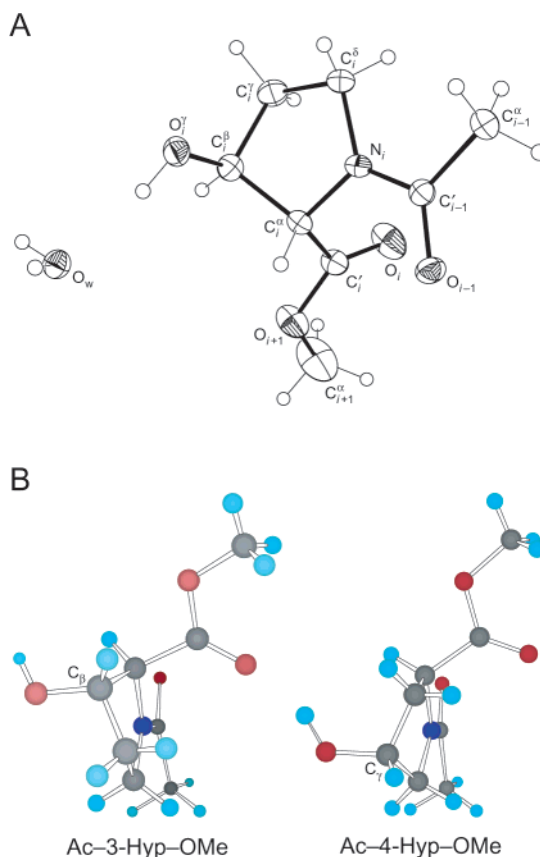
peptide	Xaa-Yaa-Gly	$T_m$ (°C) <sup>a</sup>
<b>3</b>	Pro-4-Hyp-Gly	36 ± 2 <sup>b</sup>
<b>1</b>	3-Hyp-4-Hyp-Gly	32.7 ± 0.9
<b>4</b>	Pro-Pro-Gly	30.5 ± 2.2 <sup>b</sup>
<b>2</b>	Pro-3-Hyp-Gly	21.0 ± 2.2

<sup>a</sup> Values of  $T_m$  were determined by CD spectroscopy for peptides (0.2 mM) in 50 mM acetic acid and are the average (±SE) of at least three determinations. <sup>b</sup> From ref 27.

resulting  $T_m$  values, along with those of triple helices of (Pro-4-Hyp-Gly)<sub>7</sub> (**3**) and (Pro-4-Hyp-Gly)<sub>3</sub>-Pro-Pro-Gly-(Pro-4-Hyp-Gly)<sub>3</sub> (**4**), are listed in Table 1. The values confirm that, in its natural (i.e., Yaa) position, 4-Hyp provides more conformational stability to a triple helix than does Pro. In contrast, 3-Hyp provides less conformational stability than does Pro in either its natural (i.e., Xaa) or a nonnatural (i.e., Yaa) position. The additional stability endowed by 4-Hyp is known to arise from stereoelectronic effects.<sup>14,15</sup> What is the origin of the detrimental effect of 3-Hyp on collagen stability?

**Peptide Bond Isomerization.** All of the peptide bonds in triple-helical collagen are in the trans (i.e., *Z*) conformation. The 4-hydroxyl group of a 4-Hyp residue increases the trans/cis ratio of an Xaa-4-Hyp peptide bond by approximately 50% compared to that of a Xaa-Pro bond.<sup>28</sup> We used NMR spectroscopy to determine the effect of the 3-hydroxyl group of a 3-Hyp residue on its trans/cis ratio. We found that the trans/cis ratios of amide **8** and Ac-Pro-OME in D<sub>2</sub>O were 4.9 and 4.6, respectively. These similar values indicate that peptide bond isomerization makes a negligible contribution to the instability of triple-helical **1** and **2**.

**Structure of a 3-Hyp Residue.** We used X-ray diffraction analysis to determine the structure of crystalline amide **8** (Figure 1A).<sup>29,30</sup> The pyrrolidine ring of crystalline amide **8** is puckered such that its N-C<sup>α</sup>-C<sup>β</sup>-O<sup>γ</sup> bond has a dihedral angle of -81.46 ± 0.16°, as would be expected from the manifestation of a gauche effect.<sup>14,31</sup> The ring pucker in crystalline Ac-3-Hyp-OME (**8**) differs significantly from the ring pucker in crystalline Ac-4-Hyp-OME (Figure 1B). The conformation of the pyrrolidine ring in amide **8** is intermediate between a <sup>1</sup>E envelope and <sup>1</sup>E<sub>2</sub> twisted conformation.<sup>32</sup> In the envelope conformation, the



**Figure 1.** (A) Ortep diagram of crystalline *N*-(<sup>13</sup>C<sub>2</sub>-acetyl)-3(*S*)-hydroxy-*L*-proline methyl ester (**8**) drawn with 30% probability ellipsoids. (B) Pyrrolidine ring pucker in crystalline amide **8** (left) and crystalline *N*-acetyl-4(*R*)-hydroxy-*L*-proline methyl ester (right).<sup>33</sup>

flap atom is C<sup>β</sup>. In the twisted conformation, atoms N<sub>*i*</sub>, C<sup>β</sup>, and C<sup>δ</sup> form the basal plane. Atom C<sup>β</sup> resides 0.456 ± 0.004 Å above that plane, and atom C<sup>δ</sup> resides 0.153 ± 0.005 Å below that plane. The phase angle  $\phi_2 = 9.6 \pm 0.3^\circ$ , and the puckering amplitude is  $q_2 = 0.378 \pm 0.002$  Å.

The structure of crystalline Ac-4-Hyp-OME contains two symmetry-independent molecules.<sup>33</sup> (Thus, there are two numbers herein for each parameter.) In both molecules, the pyrrolidine ring is in the <sup>5</sup>T<sub>4</sub> (that is, <sup>C</sup>T<sub>C<sup>γ</sup></sub>) twisted conformation with a characteristic phase angle of  $\phi_2 = 129.8 \pm 0.7^\circ$  (127.8 ± 0.7°) and puckering amplitude of  $q_2 = 0.354 \pm 0.004$  Å (0.389 ± 0.006 Å).<sup>32</sup> The basal plane is determined by atoms N<sub>*i*</sub>, C<sup>δ</sup>, and C<sup>γ</sup>. Atom C<sup>β</sup> resides 0.28 ± 0.01 Å (0.22 ± 0.01 Å) above that plane, and atom C<sup>γ</sup> resides 0.35 ± 0.01 Å (0.36 ± 0.01 Å) below the plane.

The ring pucker of 3-Hyp and 4-Hyp (Figure 1B) serve to preorganize the  $\phi$  (C<sup>γ</sup><sub>*i*-1</sub>-N<sub>*i*</sub>-C<sub>*i*</sub>-C<sup>γ</sup><sub>*i*</sub>) and  $\psi$  (N<sub>*i*</sub>-C<sub>*i*</sub>-C<sup>γ</sup><sub>*i*</sub>-N<sub>*i*+1</sub>) dihedral angles in a conformation that is appropriate for the Yaa position of a collagen triple helix.<sup>34-36</sup> Likewise, a 3-Hyp

(28) Eberhardt, E. S.; Panasiak, N., Jr.; Raines, R. T. *J. Am. Chem. Soc.* **1996**, *118*, 12261-12266.

(29) Systematic absences in the diffraction data were uniquely consistent with the space group *P*2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> for a crystal of amide **8**·H<sub>2</sub>O, which yielded chemically reasonable and computationally stable refinement results. A successful solution of the structure by the direct methods provided most of the non-hydrogen atoms from the *E*-map. The remaining non-hydrogen atoms were located in an alternating series of least-squares cycles and difference Fourier maps. All non-hydrogen atoms were refined with anisotropic displacement coefficients. All hydrogen atoms were included in the structure factor calculation at idealized positions and were allowed to ride on the neighboring atoms with relative isotropic displacement coefficients. The asymmetric unit contains one molecule of amide **8** and one molecule of water, which forms three hydrogen bonds. The final least-squares refinement of 138 parameters against 1963 data points resulted in residuals *R* (based on *F*<sup>2</sup> for *I* ≥ 2σ) and *wR* (based on *F*<sup>2</sup> for all data) of 0.0369 and 0.0965, respectively. The final difference Fourier map was featureless. The absolute configuration was assigned from the synthetic route.

(30) The conformation of crystalline amide **8** could be influenced by a hydrogen bond with water (Figure 1A). To search for such an influence, the geometry of amide **8** was optimized fully at the B3LYP/6-31+G\* level of theory. The structures obtained by experiment and theory are in good agreement (data not shown). Some bond distances vary beyond the estimated standard deviations but fall within the expected ranges. The pyrrolidine rings have an essentially identical conformation with the hydroxyl oxygens overlapping exactly.

(31) O'Hagan, D.; Bilton, C.; Howard, J. A. K.; Knight, L.; Tozer, D. J. *J. Chem. Soc., Perkin Trans. 2* **2000**, 605-607.

(32) Giacovazzo, C.; Monaco, H. L.; Artioli, G.; Viterbo, D.; Ferraris, G.; Gilli, G.; Zanotti, G.; Catti, M. *Fundamentals of Crystallography*, 2nd ed.; Oxford University Press: Oxford, UK, 2002.

(33) Panasiak, N., Jr.; Eberhardt, E. S.; Edison, A. S.; Powell, D. R.; Raines, R. T. *Int. J. Pept. Protein Res.* **1994**, *44*, 262-269.

(34) Vitagliano, L.; Berisio, R.; Mazzarella, L.; Zagari, A. *Biopolymers* **2001**, *58*, 459-464.

(35) Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari, A. *Protein Sci.* **2002**, *11*, 262-270.

(36) DeRider, M. L.; Wilkens, S. J.; Waddell, M. J.; Bretscher, L. E.; Weinhold, F.; Raines, R. T.; Markley, J. L. *J. Am. Chem. Soc.* **2002**, *124*, 2497-2505.

**Table 2.** Main-Chain Angles of Proline and Hydroxyproline Residues in Crystalline Amides and Triple Helices

angle	amide		triple helix			
	Ac-4-Hyp-OMe <sup>a</sup>	[ <sup>13</sup> CH <sub>3</sub> ]Ac-3-Hyp-OMe	Pro (Xaa) <sup>b</sup>	Pro (Yaa) <sup>b</sup>	Pro (Xaa) <sup>c</sup>	Hyp (Yaa) <sup>c</sup>
$\phi$	-50.9	-79.5	-75	-60	-69.8	-57.4
$\psi$	145.2	163.7	164	152	162	149.8

<sup>a</sup> From ref 33. <sup>b</sup> From ref 35. <sup>c</sup> From ref 41.

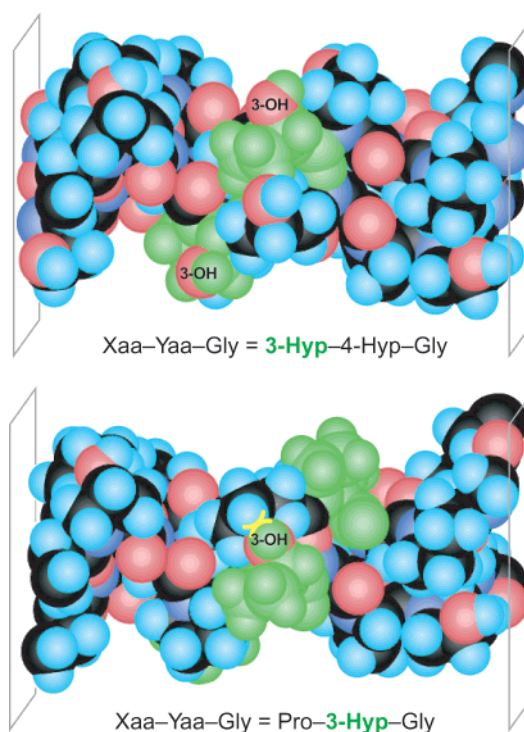
residue would be stabilizing in the Xaa position relative to Pro because its main-chain dihedral angles are favorable for that position, whereas a 3-Hyp residue would be destabilizing in the Yaa position because its dihedral angles differ significantly from the optimal (Table 2). Although 3-Hyp is destabilizing in both positions (Table 1), the preorganization that arises from the gauche effect does favor a triple helix with 3-Hyp in the Xaa rather than the Yaa position. Unlike 3-Hyp and 4-Hyp residues, a Pro residue is not constrained by a gauche effect to adopt a particular pucker and can, therefore, accommodate the different  $\phi$  and  $\psi$  angles of the Xaa and Yaa positions in a triple helix (Table 2).

To discern whether hydroxyproline residues have similar conformations in small molecule and collagen-related peptides contexts, we compared the dihedral angles of Ac-4-Hyp-OMe with those of a 4-Hyp residue in crystalline collagen (PDB entry 1CAG<sup>20</sup>) by pseudorotational analysis.<sup>37</sup> We found that the two rings were essentially superimposable (data not shown), indicating that Hyp ring puckers are not dependent on context and providing validity to our assumption that 3-Hyp would have the same pucker in a peptide context as in amide **8**.

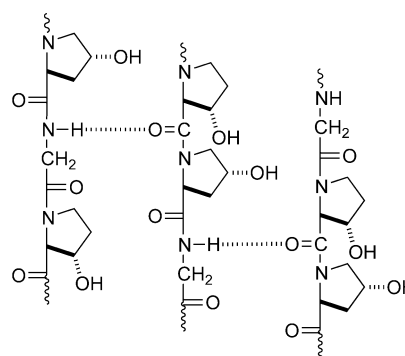
**Steric Effect on Collagen Stability.** Could steric effects contribute to the instability imparted by 3-Hyp? In a collagen triple helix, residues in the Xaa position are more solvent-exposed than are those in the Yaa position<sup>38</sup> and are, therefore, less likely to introduce unfavorable steric interactions. Indeed, residues with bulky side chains are found only rarely in the Yaa position of natural collagen.<sup>23</sup> A steric clash between 3-Hyp in the Yaa position with residues in neighboring strands (Figure 2, bottom) is likely to diminish the stability of triple-helical **2**. This steric clash is absent in triple-helical **1** (Figure 2, top).

**Effect of Hydrogen Bonds on Collagen Stability.** In a collagen triple helix, the C=O of the residue in the Xaa position accepts a hydrogen bond from the glycine N-H of another strand (Figure 3). The carboxyl pK<sub>a</sub> of XaaOH is a measure of the ability of the Xaa residue to accept such a hydrogen bond.<sup>39</sup> The pK<sub>a</sub> values of the carboxyl groups of 3-HypOH and 4-HypOH were determined by monitoring the effect of pH on the <sup>1</sup>H NMR chemical shift of their  $\alpha$ -protons. The carboxyl pK<sub>a</sub> of 4-HypOH was found to be 1.80, which is close to the value of 1.82 reported previously,<sup>18</sup> and that of ProOH is 1.95.<sup>18</sup> In contrast, the pK<sub>a</sub> of 3-HypOH is 1.62, making the 3-Hyp residue a weaker hydrogen bond acceptor than Pro. This attribute is likely to contribute to the instability conferred by a 3-Hyp residue in the Xaa position.

**Biological Implications.** 3-Hyp only decreases the T<sub>m</sub> value of triple-helical **2** by 3 °C compared to  $\geq 5$  °C for other residues in the Xaa position of (Pro-Hyp-Gly)-based host-guest triple helices.<sup>26</sup> Thus, the insertion of 3-Hyp could serve to modulate



**Figure 2.** Model of a segment of the triple helix formed by peptide **1** (top) and peptide **2** (bottom). Each 3-Hyp residue (green), the oxygen of its hydroxyl group (red, labeled), and the steric clash in triple-helical peptide **2** (yellow) are indicated.



**Figure 3.** Putative interstrand hydrogen bonds in triple-helical (3-Hyp-4-Hyp-Gly)<sub>n</sub>.

the local stability of triple helices. This idea is supported by the 10-fold increase in 3-Hyp residues in basement-membrane collagens as compared to fibrillar collagens.<sup>8,40</sup> Collagen triple helices in basement membranes interact with each other, as well as with other biomolecules, in a more varied and complex manner than do those in fibrils. This network of interactions could require regions of finely tuned conformational stability, and 3-Hyp residues could provide that tuning.

(37) Chacko, K. K.; Swaminathan, S.; Veena, K. R. *Curr. Sci.* **1983**, *52*, 660–663.

(38) Jones, E. Y.; Miller, A. J. *Mol. Biol.* **1991**, *218*, 209–219.

(39) Holmgren, S. K.; Bretscher, L. E.; Taylor, K. M.; Raines, R. T. *Chem. Biol.* **1999**, *6*, 63–70.

(40) Kefalides, N. A. *Int. Rev. Connect. Tissue Res.* **1973**, *6*, 63–104.

(41) Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari, A. *Biopolymers* **2001**, *56*, 8–13.

**Acknowledgment.** We are grateful to J. A. Hodges, R. Berisio, and L. Vitagliano for contributive discussions. This work was supported by Grant AR44276 (NIH). C.L.J. was supported by Chemistry–Biology Interface Training Grant GM08506 (NIH). CD spectra were obtained at the University of Wisconsin-Madison Biophysical Instrumentation Facility, which was supported by Grants BIR-9512577 (NSF) and RR13790 (NIH). NMR spectra were obtained at the Magnetic Resonance Facility in the Department of Chemistry, which was supported by Grants CHE-8813550, CHE-9629688, and CHE-

9208463 (NSF) and S10 RR08389-01 and S10 RR04981-01 (NIH).

**Supporting Information Available:** Data for pH titration of 3-HypOH and 4-HypOH, thermal denaturation of triple-helical peptides **1** and **2**, and X-ray diffraction analysis of amide **8**. This material is available free of charge via the Internet at <http://pubs.acs.org>.

JA034015J