The Crystal Structure of a Collagen-like Polypeptide with 3(S)-Hydroxyproline Residues in the Xaa Position Forms a Standard 7/2 Collagen Triple Helix

Collagen has a triple helical structure comprising strands with a repeating Xaa-Yaa-Gly sequence. 1-L-Proline (Pro) and 4(R)-hydroxyl-1-L-proline (4(R)Hyp) residues are found most frequently in the Xaa and Yaa positions. However, in natural collagen, 3(S)-hydroxy-L-proline (3(S)Hyp) occurs in the Xaa positions to varying extents and is most common in collagen types IV and V. Although 4(R)Hyp residues in the Yaa positions have been shown to be critical for the formation of a stable triple helix, the role of 3(S)Hyp residues in the Xaa position is not well understood. Indeed, recent studies have demonstrated that the presence of 3(S)Hyp in the Xaa positions of collagen-like peptides actually has a destabilizing effect relative to peptides with Pro in these locations. Whether this destabilization is reflected in a local unfolding or in other structural alterations of the collagen triple helix is unknown. Thus, to determine what effect the presence of 3(S)Hyp residues in the Xaa positions has on the overall conformation of the collagen triple helix, we determined the crystal structure of the polypeptide (8-Gly-Pro-4(R)Hyp)9-(Gly-Pro-4(R)Hyp)2-(Gly-Pro-4(R)Hyp)4-OH to 1.80 Å resolution. The structure shows that, despite the presence of the 3(S)Hyp residues, the peptide still adopts a typical 7/2 superhelical symmetry similar to that observed in other collagen structures. The puckering of the Xaa position 3(S)Hyp residues, which are all down (Cγ-endo), and the ϕ/ψ dihedral angles of the Xaa 3(S)Hyp residues are also similar to those of typical collagen Pro Xaa residues. Thus, the presence of 3(S)Hyp in the Xaa positions does not lead to large structural alterations in the collagen triple helix.

Collagens are the most abundant proteins in animals, comprising an estimated one-third of the total protein by weight (1, 2). At least 27 collagen types, which are formed from 42 distinct polypeptide chains, exist in vertebrates (1, 2). In addition, more than 20 additional proteins that adopt collagen-like structures, such as collectins, ficolins, and scavenger receptors, are also found (1). Collagen is an essential molecule in vertebrates, as it plays the dominant role in maintaining the structure of tissues. However, collagen and collagen-like proteins have many other important roles, such as cell adhesion, chemotaxis, cell migration, and the regulation of tissue remodeling during cell growth, differentiation, morphogenesis, and wound healing (1).

All collagen molecules consist of three polypeptide chains. These chains, called α chains, contain at least one domain composed of repeating Gly-Xaa-Yaa sequences (1, 2). Three α chains, which are each coiled into a polyproline II-like left-handed structure, twist around each other to form a loose right-handed superhelix, the so-called collagen triple helix. The presence of the small glycine residue at every third position in collagen is essential, because the compact triple helix places every third residue near the superhelical central axis. Collagen also has a high imino acid content (~20%) and typically contains Pro and (2S,4R)-4-hydroxyproline (4(R)Hyp)2 in the Xaa and Yaa positions, respectively. The presence of such imino acids places additional restraints on the collagen structure important for triple helix formation and stability. Extensive studies have been carried out examining collagen structure and stability using polypeptides as model systems (3–5). High resolution crystal structures of collagen peptides have been determined, including several Gly-Pro-Pro and Gly-Pro-4(R)Hyp peptide structures as well as a peptide with a Gly→Ala mutation, the biologically relevant (Pro-Hyp-Gly)3-Ile-Thr-Gly-Ala-Arg-Gly-Leu-Ala-Gly-Pro-Hyp-Gly-(Pro-Hyp-Gly)3 peptide, the charged peptide (Pro-Hyp-Gly)4-Glu-Lys-Gly-(Pro-Hyp-Gly)3, and more recently, the structure of H-(Gly-4(R)Hyp-Pro-4(R)Hyp)5-OH and H-(4(R)Hyp-4(R)Hyp-Gly)10-OH (6–16).

These structures, when combined with stability studies, have provided a detailed account of amino acid propensities for the collagen triple helix.

In addition, quantum mechanical ab initio calculations have elucidated the specific roles of stereochemical effects, ring puckering, cis/trans stabilization, and ϕ/ψ dihedral angles in Pro and 4(R)Hyp residues and the effect of these parameters in the stability of the collagen triple helix (17–20). These studies

2 The abbreviations used are: 4(R)Hyp, (2S,4R)-4-hydroxyproline; 3(S)Hyp, (2S,3S)-3-hydroxyproline; Fmoc, N-(9-fluorenylethoxycarbonyl); ASU, asymmetric unit.
reveal that electronegative (4S) and 4(R) substituents on Pro residues strongly favor down ring and up ring puckering, respectively. Thus, collagen stability is linked to the interplay between the pyrrolidine ring pucker, φ/ψ dihedral angles, and the cis/trans conformation. An important conclusion from these combined studies is that down pucker is preferred in the Xaa positions, and up pucker is preferred in the Yaa positions of the collagen triple helix (17–20).

These studies are consistent with the finding that 4(R)Hyp, which prefers the up pucker conformation, is preferred in the Yaa position and, in fact, is necessary for the stabilization of the collagen triple helix. Because 4(R)Hyp residues adopt preferentially the up pucker, they preorganize the main chain dihedral angles in a conformation that favors triple helix structure (21, 22). The importance of this stability is reflected in the finding that Gly-Xaa-4(R)Hyp triplets account for ~40% of the amino acid sequence of collagens (23). 4(R)Hyp residues are formed from post-translational modification of Pro residues in unfolded collagen, as mediated by prolyl 4-hydroxylases (1, 2, 25–28). The important role played by 4(R)Hyp in collagen structure/function is further underscored by the finding that knocking out the prolyl 4-hydroxylase activity in the nematode Caenorhabditis elegans is lethal (28).

In addition to prolyl 4-hydroxylases, two additional post-translational modifications mediated by hydroxylases are observed and important in collagen, the formation of hydroxylysine and (25,35)-3-hydroxyproline (3(S)Hyp). The hydroxylation of lysine residues, in the sequence Gly-Xaa-Lys, is carried out by one of three well characterized lysyl hydroxylases (29). This modification allows for the formation of intermolecular collagen cross-links and the attachment of carbohydrate units to lysyl hydroxyl groups. Hydroxylation of Pro residues in the Xaa positions to 3(S)Hyp is mediated by the recently characterized prolyl 3-hydroxylase enzymes (30). The occurrence of 3(S)Hyp is significantly less frequent than that of 4(R)Hyp in the total amino acid content of collagens. The fibrillar collagens, types I–III, contain a single 3(S)Hyp residue per chain, and higher modification levels are typically restricted to collagen types IV and V and are, thus, most abundant in basement membrane collagens (31, 32).

Acetyl-3(S)Hyp-O-methylester strongly favors a down puckered ring conformation (20), similar to the preferred pucker for the Xaa position of the collagen triple helix. Because of this, it was predicted that placing 3(S)Hyp in the Xaa position, similar to placing 4(R)Hyp in the Yaa position, should preorganize the main chain dihedral angles properly for triple helix formation. Indeed, it was thought that a triple helix with 3(S)Hyp residues in the Xaa position would be more stable than one with Pro residues at this position, as Pro does not favor one pucker type (up or down) over the other. Thus, it was a surprise when it was found that peptides with sequences containing long stretches of (3(S)Hyp)-(4(R)Hyp)-Gly do not form stable triple helices (33). However, when Raines and co-workers (20) examined this issue via host guest peptides, specifically, with (Pro-4(R)Hyp-Gly)3-3(S)Hyp-4(R)Hyp-Gly-(Pro-4(R)Hyp-Gly)3, they found that the presence of 3(S)Hyp residues embedded in the Xaa position of this peptide only decreased the transition midpoint temperature (Tm) value of the triple helix formation by 3 °C compared with a peptide with a Pro in the corresponding position. Nonetheless, the presence of the 3(S)Hyp in the Xaa position was clearly destabilizing rather than stabilizing. This finding leads to the suggestion that the presence of 3(S)Hyp in collagen functions to modulate the stability of basement membranes to allow for the meshwork-like structures they form. On the other hand, 3(S)Hyp may locally destabilize areas of the collagen network to allow for protein-protein interactions. Whatever its role, 3(S)Hyp is clearly an integral part of the collagen triple helix in fibrillar and basement membrane collagens and is also present at low levels in other collagen molecules such as type X collagen (34).

Currently, there is no structure available for a collagen peptide containing the triplet sequence Gly-3(S)Hyp-4(R)Hyp, and no structure of 3(S)Hyp has been determined in the context of a peptide or protein. We were interested in determining whether a peptide containing multiple 3(S)Hyp residues embedded in the Xaa positions could form a collagen triple helix or would, instead, form an altered conformation. To address this issue, we determined the crystal structure of the collagen-like peptide H-(Gly-Pro-4(R)Hyp)3-(Gly-3(S)Hyp-4(R)Hyp)2-(Gly-Pro-4(R)Hyp)2-OH. Surprisingly, our structure showed that this peptide forms a triple helix that conforms to 7/2 superhelix symmetry, similar to that of other collagen peptide-like structures. Moreover, the 3(S)Hyp residues in the Xaa position of our structure assumed a down pucker conformation essentially identical to that adopted by Pro residues in other structures with Pro-4(R)Hyp-Gly sequences (6–16). The φ/ψ dihedral angles of the Gly, Xaa, and Yaa residues were also similar to those observed in other peptides. This suggests that, although the presence of the 3(S)Hyp in the Xaa position is slightly unfavorable for the stability, a collagen sequence containing such a modification can fold into a collagen triple helix with standard 7/2 superhelix symmetry.

**MATERIALS AND METHODS**

**Peptide Synthesis**—Peptides were synthesized with an ABI433A synthesizer. Couplings were carried out on an ABI344 synthesizer using H-O-t-butyl-1-trans-4-hydroxyproline-2-chlorotrityl resin (AnaSpec, Inc. San Jose, CA) and using Fmoc-amino acids (Fmoc-Gly-OH, Fmoc-3(S)Hyp, and Fmoc-4(R)Hyp(t-butyl)-OH). HATU (O-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethylenuronium hexafluorophosphate (Perseptive Biosystems) (4.0 equivalents))/diisopropylethylamine mediated peptide couplings. The peptide was cleaved from the resin and purified by preparative high pressure liquid chromatography (Vydac® C18, 5 μm, 300 Å, 250 × 50 mm, W. R. Grace, Columbia, MD). All synthesized peptides were characterized by electrospray/quadrupole/time-of-flight mass spectrometry (Q-tof micro, Waters Associates) and amino acid analysis.

**Crystallization and Data Collection**—The purified and lyophilized peptide H-(Gly-Pro-4(R)Hyp)3-(Gly-3(S)Hyp-4(R)Hyp)2-(Gly-Pro-4(R)Hyp)2-OH was dissolved at a concentration of 10 mg/ml in water. The peptide was crystallized at 4 °C using the hanging drop vapor diffusion method. For crystallization, 2 μl of the peptide solution was mixed with 2 μl of the reservoir solution (25% of polyethylene glycol 4000). The crystals appeared as very thin plates in a period of 1–2 weeks.
Data collection at the synchrotron enabled us to index the full cell of the crystals as monoclinic, space group P2₁, with \( a = 18.12 \), \( b = 19.45 \) Å, and \( c = 81.50 \) Å and \( \beta = 91.26^\circ \).

For cryoprotection, glycerol was added to the drop to a final concentration of 29.7/32.8% to 1.80 Å resolution (38, 39). After the inclusion of several waters and additional rounds of refinement, the \( R_{work}/R_{free} \) converged to 26.8/29.4%. Multiple omit maps were calculated throughout the refinement process to confirm the correctness of the model. Notably, electron density for the hydroxyl moieties were clearly observed for all 3(S)Hyp and 4(R)Hyp residues. The current model has excellent stereochemistry (40) (Table 1).

**Differential Scanning Calorimetry**—Differential scanning calorimetry experiments were performed with a model CSC6100 Nano II differential scanning calorimeter (Calorimetry Sciences Corp., Lindon, UT) in 20 mM sodium phosphate buffer, pH 7.0, containing 135 mM NaCl. The scan rate was 0.5 K/min in all experiments. The calorimetric transition enthalpies were obtained by integration of the area under the excess heat capacity plots, employing a sigmoidal baseline to take the small heat capacity changes between the triple helical and unfolded state into account. To measure the heat capacity, a partial specific volume, 0.705 ml/g, was used for all peptides. Concentrations were determined by amino acid analysis.

**RESULTS AND DISCUSSION**

**Thermal Stability of the H-(Gly-Pro-4(R)Hyp)₃-(Gly-3(S)Hyp-4(R)Hyp)₃-(Gly-Pro-4(R)Hyp)₄-OH Peptide**—The thermal stability of the peptide H-(Gly-Pro-4(R)Hyp)₃-(Gly-3(S)Hyp-4(R)Hyp)₃-(Gly-Pro-4(R)Hyp)₄-OH was compared with the peptides H-(Gly-Pro-4(R)Hyp)₄-OH and H-(Gly-Pro-4(R)Hyp)₃-(Gly-3(S)Hyp-4(R)Hyp)-(Gly-Pro-4(R)Hyp)₄-OH using differential scanning calorimetry. At a scanning rate of 0.5 K/min, the transition midpoint temperature \( T_{m} \) of H-(Gly-

---

**Figure 1**. Differential scanning calorimetry of H-(Gly-Pro-4(R)Hyp)₃-(Gly-3(S)Hyp-4(R)Hyp)₃-(Gly-Pro-4(R)Hyp)₄-OH. The excess heat capacity of H-(Gly-Pro-4(R)Hyp)₃-(Gly-3(S)Hyp-4(R)Hyp)₃-(Gly-Pro-4(R)Hyp)₄-OH is shown as a function of temperature (black line). The scanning rate was 0.5 K/min, and the peptide concentration was 1.6 mM. For comparison, H-(Gly-Pro-4(R)Hyp)₄-OH at a concentration of 2.7 mM (red line) and H-(Gly-Pro-4(R)Hyp)₃-(Gly-3(S)Hyp-4(R)Hyp)-(Gly-Pro-4(R)Hyp)₄-OH at a concentration of 2.1 mM (blue line) are shown. The samples were measured in 20 mM sodium phosphate buffer, pH 7.0, containing 135 mM NaCl.
Pro-4(\(R\))Hyp\(9\)-OH at a concentration of 2.7 mM was 52 °C. The Tm of H-(Gly-Pro-4(\(R\))Hyp)\(4\)-(Gly-3(\(S\))Hyp-4(\(R\))Hyp)-(Gly-Pro-4(\(R\))Hyp)\(4\)-OH at a concentration of 2.1 mM was also 52 °C, and for H-(Gly-Pro-4(\(R\))Hyp)\(3\)-(Gly-3(\(S\))Hyp-4(\(R\))Hyp)\(2\)-(Gly-Pro-4(\(R\))Hyp)\(4\)-OH, a Tm of 51 °C was determined at a concentration of 1.6 mM. The Tm values of these three peptides were not significantly different. However, the transition enthalpies varied significantly. The \(\Delta H^\circ\) for H-(Gly-Pro-4(\(R\))Hyp)\(9\)-OH (350 kJ/mol trimer) was significantly larger than that of H-(Gly-Pro-4(\(R\))Hyp)\(4\)-(Gly-3(\(S\))Hyp-4(\(R\))Hyp)-(Gly-Pro-4(\(R\))Hyp)\(4\)-OH (210 kJ/mol trimer) and H-(Gly-Pro-4(\(R\))Hyp)\(3\)-(Gly-3(\(S\))Hyp-4(\(R\))Hyp)\(2\)-(Gly-Pro-4(\(R\))Hyp)\(4\)-OH (200 kJ/mol trimer) (Fig. 1). The decrease in \(\Delta H^\circ\) of the 3(\(S\))Hyp-containing peptides is compensated by a decrease in the transition entropy; therefore, the Tm values are very similar. The decrease of the transition enthalpy could be due to weaker hydrogen bonds or differences in the unfolded state. Further studies will be required to determine the cause of these thermodynamic differences.

Overall Structure of H-(Gly-Pro-4(\(R\))Hyp)\(3\)-(Gly-3(\(S\))Hyp-4(\(R\))Hyp)\(2\)-(Gly-Pro-4(\(R\))Hyp)\(4\)-OH Peptide—The structure of the H-(Gly-Pro-4(\(R\))Hyp)\(3\)-(Gly-3(\(S\))Hyp-4(\(R\))Hyp)\(2\)-(Gly-Pro-4(\(R\))Hyp)\(4\)-OH peptide was determined by molecular replacement (see “Materials and Methods”) and refined to an \(R\)_free of 29.4% to 1.80 Å resolution (Table 1). The crystallographic ASU consists of one complete triple helix (Fig. 2). Density is weak for the triplet repeats at the N- and C-terminal ends of each chain,
indicating higher flexibility of these residues. The current triple helical structure includes residues 3–27 of the A chain, 1–27 of the B chain, and 1–27 of the C chain. It also contains 152 water molecules. The packing of the triple helices in this structure appears to be pseudotetragonal (Fig. 3).

The structure fits the data well as revealed by composite omit $2F_o - F_c$ maps (in which the model is omitted in stretches of 5% and then combined to build a final omit map covering the entire molecule), and, importantly, the high resolution of the data permits an unequivocal determination of the proline ring puckers in the structure (Fig. 4). As expected, the 4(R)Hyp residues in the Yaa positions all take the Cγ-exo or up pucker conformation, whereas the Xaa-Pro residues are all down puckerer. Notably, all of the Xaa 3(S)Hyp residues adopted a similar down pucker conformation observed for the Xaa-Pro residues, which is the same as that observed for Xaa-Pro residues in other collagen peptide structures (6–14). Indeed, the $H$-(Gly-Pro-4(R)Hyp)$_3$-(Gly-3(S)Hyp-4(R)Hyp)$_2$-(Gly-Pro-4(R)Hyp)$_4$-OH peptide adopted the standard super helix 7/2 symmetry observed in other collagen peptide structures (6–16). This is underscored by the fact that the structure can be superimposed onto that of [(Pro-Pro-Gly)$_{10}$]$_3$ (1K6F), with a root mean squared deviation for all C-α atoms of only 0.21 Å (Fig. 5). This indicates that these two structures are essentially identical.

**Water Contacts and Interhelical Residue Interactions**—The role that water plays in the stabilization and structure of the collagen triple helix has been highly controversial (21, 41, 42). Although it has now been established that water molecules do not play a predominant role in collagen stabilization, they do exist as an important and intrinsic feature of the triple helix (41, 43, 44). This is strongly supported by the fact that, in every collagen structure to date, there is a conserved pattern of water molecules in which the glycine carbonyl oxygens are singly hydrated and the Yaa carbonyl groups are doubly hydrated (6–16). In structures that contain a Yaa 4(R)Hyp, the hydroxyl moiety of the 4(R)Hyp is doubly hydrated. Our $H$-(Gly-Pro-4(R)Hyp)$_3$-(Gly-3(S)Hyp-4(R)Hyp)$_2$-(Gly-Pro-4(R)Hyp)$_4$-OH peptide structure contains these same hydration patterns. However, unlike most collagen structures, we did not observe any direct interhelical hydrogen bonding interactions between hydroxyl moieties of 4(R)Hyp symmetry-related molecules. This marks this structure as the third (along with the $H$-(Gly-4(R)Hyp-4(R)Hyp)$_3$-OH peptide structure and the Gly→Ala structure) in which such contacts are absent (14, 15). The lack of such contacts in yet another collagen peptide structure provides more support for the theory put forth by Raines and co-workers (21) that hydroxyl hydrogen bonds are not the crucial stabilizing feature of 4(R)Hyp Yaa residues but rather...
the inductive effect and preorganization potential of these hydroxylated residues. However, we cannot preclude the possibility that interhelical hydroxyl hydrogen bonds between 4(R)Hyp residues may play some important role in collagen function, as they have been observed in the majority of collagen peptide crystal structures containing 4(R)Hyp residues.

Interestingly, although there are no interhelical Yaa (4(R)Hyp)–(4(R)Hyp) hydroxyl hydrogen bonds in our structure, there are two interhelical Xaa (3(S)Hyp)–(3(S)Hyp) hydrogen bonds (OH–OH; 3.2 Å) (Fig. 6). Perhaps this is not surprising as the Xaa position in the collagen triple helix is the most exposed residue. However, such hydrogen bonds may play a role in collagen molecules that contain 3(S)Hyp in the Xaa position in the stabilization and creation of higher order structures. For example, it has been suggested that, even if Yaa4(R)Hyp–Yaa4(R)Hyp hydrogen bonds between hydroxyl groups are not essential for providing stability, they may play a role in the correct axial registration of helices (11–13). Perhaps interhelical hydrogen bonds between 3(S)Hyp hydroxyls might fulfill a similar function, or more likely, play a role in interactions with other matrix molecules.

The four 3(S)Hyp residues that do not form interhelical hydrogen bonds all share in common one very tight hydrogen bonding interaction to a water molecule (2.4 Å). In one case this water molecule is bridged to a hydroxyl moiety in a 4(R)Hyp residue on a adjacent, symmetry-related helix, whereas in the other cases the water bridges to other water molecules, which subsequently interact with the hydroxyl moieties of 4(R)Hyp or 3(S)Hyp residues or to carbonyl oxygens in adjacent chains. Therefore, unlike the conserved water networks observed for Gly and Yaa 4(R)Hyp residues in the collagen triple helix, there are no such solvent networks involving the 3(S)Hyp residues. However, because the Xaa position of the collagen helix is the most solvent accessible it might be expected that it could be involved in a larger variety of solvent interactions.

Ring Puckering and \( \psi/\phi \) Values— The general trend revealed in the crystal structures of collagen peptides that have been determined thus far, with a few exceptions, are that the imino acid residues in Xaa and Yaa positions adopt down and up puckers respectively. Quantum mechanical \( \text{ab initio} \) studies have found that up puckered Pro residues favor smaller values of backbone dihedral angles with \( \phi \) values typically \( \approx -60^\circ \) and \( \psi \) values ranging from 149.8° to 153.3°. These \( \phi/\psi \) dihedral angles appear important in preorganizing the formation of the triple helical structure and also allow for the optimal interchain hydrogen bonds that are the signature of the collagen triple helix structure. As a result, the interchain hydrogen bonds, between the

---

**FIGURE 5**. C-α superimposition of the H-(Gly-Pro-4(R)Hyp)₃-(Gly-3(S)Hyp-4(R)Hyp)₃-(Gly-Pro-4(R)Hyp)₄-OH structure onto that of [(Pro-Pro-Gly)₁₀]₁₀. Superimposition of the H-(Gly-Pro-4(R)Hyp)₃-(Gly-3(S)Hyp-4(R)Hyp)₃-(Gly-Pro-4(R)Hyp)₄-OH structure onto that of [(Pro-Pro-Gly)₁₀]₁₀. Atoms of the H-(Gly-Pro-4(R)Hyp)₃-(Gly-3(S)Hyp-4(R)Hyp)₃-(Gly-Pro-4(R)Hyp)₄-OH structure are shown in green, except the 3(S)Hyp residues, which are shown in yellow, as in Fig. 1. Atoms of the of [(Pro-Pro-Gly)₁₀]₁₀ structure are shown in red. It is clear from this overlay that the structures adopt essentially the same triple helical conformation. This figure was made with PyMOL software (24).
glycine amide NH and the Xaa carbonyl, located one residue C-terminal to it on the adjacent chain, appear to be tightly restricted with obvious distance and angle requirements that are necessary for the precise formation of the triple helix structure.

For our \( H-(Gly-Pro-4(R)Hyp)_3-(Gly-3(S)Hyp-4(R)Hyp)_2-(Gly-Pro-4(R)Hyp)_2-OH \) peptide structure the average \( \phi/\psi \) values are respectively \(-73.4^\circ/174.7^\circ\) for Gly residues, \(-72.4^\circ/161.7^\circ\) for Pro residues in the Xaa positions and \(-61.8^\circ/153.8^\circ\) for 4(R)Hyp residues in the Yaa positions. For the 3(S)Hyp residues in the Xaa positions, the average \( \phi/\psi \) values are \(-71.8^\circ/161.9^\circ\). Comparison of these values to those of other collagen-like peptides shows that the \( \phi/\psi \) angles for the Gly residues and the 4(R)Hyp residues in the Xaa positions are essentially the same as those observed in these other collagen triple helix peptide structures (Table 2) and are consistent with values expected for idealized 7/2 collagen helices (7). Notably, the \( \phi/\psi \) angles for the 3(S)Hyp Xaa residues are essentially identical to those observed for the Pro Xaa residues and are also the same as those observed in the idealized 7/2 case and other high resolution crystal structures. Thus, the \( \phi/\psi \) angles observed in our structure are consistent with the finding that the structure adopts a standard 7/2 helix and adopts, essentially a standard collagen triple helical conformation.

Shown in Table 3 are the average lengths and angles for the Gly–NH–OC–Xaa interchain hydrogen bonds for the collagen structures solved to date. As can be seen, the hydrogen bond distances observed in these structures are all clustered in a very restricted range (2.88–3.01 Å). Likewise, the hydrogen bond angles exist in the narrow range between 164 and 169°. It was interesting to see whether the hydrogen bond distances and angles would be altered in our \( H-(Gly-Pro-4(R)Hyp)_3-(Gly-3(S)Hyp-4(R)Hyp)_2-(Gly-Pro-4(R)Hyp)_2-OH \) peptide structure, because Raines and co-workers (20) suggest that one explanation for the decreased stability of collagen peptides with 3(S)Hyp in the Xaa positions may be the reduction in the strength of the interchain hydrogen bond caused by the inductive effect of the 3(S)Hyp hydroxyl group. Specifically, they have determined that the \( pK_a \) of 3(S)Hyp is 1.62, which would make the 3(S)Hyp residue a weaker hydrogen bond acceptor than Pro (with a \( pK_a \) of 1.95). As shown in Table 3, the average Gly–NH–OC–3(S)Hyp Xaa interchain hydrogen bonds in our structure were 2.95 Å and 165°, which were not only essentially identical to the average interchain hydrogen bond distances and angles observed for the Gly–NH–OC–3(S)Hyp Xaa residues in our structure (2.96 Å, with an angle of 166°) but also basically the same as observed in other high resolution collagen peptide structures (Table 3). Thus, although the strength of the interchain hydrogen bonds involving the 3(S)Hyp residues may be diminished, the overall geometries of these hydrogen bonds are the same as observed in other collagen structures. Given the conservation of these parameters in the structures observed to date, it appears that the geometries of these interchain hydrogen bonds are tightly restricted and have precise distance and angle requirements necessary for the formation of the triple helix structure. There-
In conclusion, we have found that peptides with 3(S)Hyp residues would be able to form hydrogen bond interactions with Pro residues in the triple helical collagen structure. Moreover, the parameters, that define this helix are also essentially the same as those observed in other high resolution collagen peptide structures (6–16). Thus, what might explain the slightly reduced stability of a peptide with 3(S)Hyp residues compared with one with Pro residues? One likely contribution to the reduced stability of the Xaa 3(S)Hyp-4(R)Hyp Xaa (compared with Pro residues) may be the geometries of the structure. On the other hand, the fact that such interactions for the unfolded state. Perhaps leading to a preference for the unfolded state, the more hydrophobic Pro Xaa residues could not form such interactions in the unfolded state, perhaps leading to a contrast, the more hydrophobic Pro Xaa residues could not form such interactions in the unfolded state.
hydrogen bonds formed with 3(S)Hyp Xaa residues are not as favorable in the folded state as interchain hydrogen bonds formed by Pro Xaa residues. Clearly, further biophysical investigations will be needed to address this issue and clarify the basis for the reduced stability imparted to the collagen triple helix by 3(S)Hyp Xaa residues.

Fig. 7 shows that the hydroxyl groups of the 3(S)Hyp residues point away from the triple helix and are likely to be involved in collagen triple helix-protein interactions. The occurrence of a single residue of 3(S)Hyp per chain in the fibrillar collagens type I to III will most likely have a negligible effect on the stability of these collagens. However, the loss of this single 3(S)Hyp residue in a mouse model has significant biological consequences (45). The exact mechanism is not known, but a most likely explanation is that the 3(S)Hyp residue in fibrillar collagens is used for critical protein-protein interactions.

Acknowledgments—We thank the Advanced Light Source and their support staff. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences Division, of the United States Department of Energy under Contract No. DE-AC03-78SF00098 at Lawrence Berkeley National Laboratory. We thank Kerry Maddox for amino acid analysis.

REFERENCES
1. Myllyharju, J., and Kivirikko, K. I. (2004) Trends Genet. 20, 33–43
2. Myllyharju, J. (2003) Matrix Biol. 22, 15–24
3. Fields, G. B., and Prockop, D. J. (1996) Biopolymers 40, 345–357
4. Heidemann, E., and Roth, W. (1982) Adv. Polym. Sci. 43, 143–203
5. Mayo, K. H. (1996) Biopolymers 40, 359–370
6. Berisio, R., Vitagliano, L., Mazzarella, L., and Zagari, A. (2002) Protein Sci. 11, 262–270
7. Okuyama, K., Hongo, C., Fukushima, R., Wu, G., Narita, H., Noguchi, K., Tanaka, Y., and Nishino, N. (2004) Biopolymers 76, 367–377
8. Hong, C., Nagarajan, V., Noguchi, K., Kamitori, S., Okuyama, K., Tanaka, Y., and Nishino, N. (2001) Polymer J. 33, 812–818
9. Nagarajan, V., Kamitori, S., and Okuyama, K. (1998) J. Biochem. (Tokyo) 124, 1117–1123
10. Nagarajan, V., Kamitori, S., and Okuyama, K. (1999) J. Biochem. (Tokyo) 125, 310–318
11. Kramer, R. Z., Bella, J., Brodsky, B., and Berman, H. M. (2001) J. Mol. Biol. 311, 131–147
12. Kramer, R. Z., Venugopal, M. G., Bella, J., Mayville, P., Brodsky, B., and Berman, H. M. (2000) J. Mol. Biol. 301, 1191–1205
13. Kramer, R. Z., Vitagliano, L., Bella, J., Berisio, R., Mazzarella, L., Brodsky, B., Zagari, A., and Berman, H. M. (1998) J. Mol. Biol. 280, 623–638
14. Bella, J., Eaton, M., Brodsky, B., and Berman, H. M. (1994) Science 266, 75–81
15. Schumacher, M. A., Mizuno, K., and Bächinger, H. P. (2005) J. Biol. Chem. 280, 397–403
16. Kawahara, K., Nishi, Y., Nakamura, S., Uchiyama, S., Nishuchi, Y., Nakanawa, T., Ohkudo, T., and Kobayashi, Y. (2005) Biochemistry 44, 15812–15822
17. Improta, R., Mele, F., Crescenzi, O., Benzi, C., and Barone, V. (2002) J. Am. Chem. Soc. 124, 7857–7865
18. DeRider, M. L., Wilkens, S. J., Waddell, M. J., Bretscher, L. E., Weinhold, F., Raines, R. T., and Markley, J. L. (2002) J. Am. Chem. Soc. 124, 2497–2505
19. Jenkins, C. L., and Raines, R. T. (2002) Nat. Prod. Rep. 19, 49–59
20. Jenkins, C. L., Bretscher, L. E., Guzei, I. A., and Raines, R. T. (2003) J. Am. Chem. Soc. 125, 6422–6427
21. Holmgren, S. K., Taylor, K. M., Bretscher, L. E., and Raines, R. T. (1996) Nature 392, 501–503
22. Holmgren, S. K., Bretscher, L. E., Taylor, K. M., and Raines, R. T. (1999) Chem. Biol. 6, 63–70
23. Ramshaw, J. A. M., Shah, N. K., and Brodsky, B. (1998) J. Struct. Biol. 122, 86–91
24. DeLano, W. L. (2002) The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos, CA
25. Lamande, S. R., and Bateman, J. F. (1999) Semin. Cell Dev. Biol. 10, 455–464
26. Kivirikko, K. I., and Pihlajaniemi, T. (1998) Adv. Enzymol. Relat. Areas Mol. Biol. 72, 325–398
27. Kivirikko, K. I., and Myllyharju, J. (1998) Matrix Biol. 16, 357–368
28. Winter, A. D., and Page, A. P. (2000) Mol. Cell. Biol. 20, 4084–4094
29. Yeowell, H. N., and Walker, L. C. (2000) Mol. Metab. 71, 212–224
30. Vranka, J. A., Sakai, L. Y., and Bächinger, H. P. (2004) J. Biol. Chem. 279, 23615–23621
31. Rhodes, R. K., and Miller, E. J. (1978) Biochemistry 17, 3442–3448
32. Kefalides, N. A. (1973) Int. Rev. Connect. Tissue Res. 6, 103–114
33. Mizuno, K., Hayashi, T., Peyton, D. H., and Bächinger, H. P. (2004) J. Biol. Chem. 279, 282–287
34. Bos, K. J., Rucklidge, G. J., Dunbar, B., and Robins, S. P. (1999) Matrix Biol. 18, 149–153
35. Collaborative Computational Project X (1994) Acta Crystallogr. Sect. D Biol. Crystallogr. 50, 760–763
36. Leslie, A. G. W. (1992) J. Appl. Crystallogr. 25, 760–763
37. Kissinger, C. R., Gehlhaar, D. K., and Fogel, D. B. (1999) Acta Crystallogr. Sect. D Biol. Crystallogr. 55, 484–491
38. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gross, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D Biol. Crystallogr. 54, 905–921
39. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta Crystallogr. Sect. A 47, 110–119
40. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Crystallogr. 26, 283–291
41. Mogilner, I. G., Ruderman, G., and Grigera, J. R. (2002) J. Mol. Graph. Model. 21, 209–213
42. Engel, J., and Prockop, D. J. (1998) Matrix Biol. 17, 679–680
43. Bella, J., Brodsky, B., and Berman, H. M. (1995) Structure (Lond.) 3, 893–906
44. Kramer, R. Z., Bella, J., Mayville, P., Brodsky, B., and Berman, H. M. (1999) Nat. Struct. Biol. 6, 454–457
45. Morello, R., Bertin, T., Chen, Y., Hicks, J., Tonachini, L., Massimiliano, M., Castagnola, P., Rauch, F., Glorieux, F. H., Vranka, J., Bächinger, H. P., Pace, I. M., Schwarze, U., Byers, P. H., Weis, M., Fernandes, R. J., Eyre, D. R., Yao, Z., Boyce, B. F., and Lee, B. (2006) Cell, in press