Role of Carbohydrate in Stabilizing the Triple-Helix in a Model for a Deep-Sea Hydrothermal Vent Worm Collagen†

James G. Bann,‡,§,|| Hans Peter Bächinger,‡,§ and David H. Peyton*,⊥

Department of Biochemistry and Molecular Biology, Oregon Health and Science University, Portland, Oregon 97201, Research Department, Shriners Hospital for Children, Portland, Oregon 97201, and Department of Chemistry, Portland State University, Portland, Oregon 97207-0751

Received October 22, 2002; Revised Manuscript Received January 22, 2003

ABSTRACT: The glycopeptide Ac-(Gly-Pro-Thr(β-Gal))$_{10}$-NH_2 forms a collagen-like triple-helix. A ^{1}H NMR structural analysis is reported for the peptides Ac-(Gly-Pro-Thr)$_n$-NH_2 and Ac-(Gly-Pro-Thr(β-Gal))$_n$-NH_2, where $n = 1$, 5, and 10. NMR assignments for the individual peptides are made using one- and two-dimensional TOCSY, ROESY, and NOESY experiments. The NMR and corroborating CD data show that Ac-(Gly-Pro-Thr)$_n$-NH_2, $n = 1$, 5, or 10, as well as Ac-(Gly-Pro-Thr(β-Gal))$_n$-NH_2, $n = 1$ or 5 peptides are unable to form collagen-like triple-helical structures. Furthermore, the equilibrium ratio of cis to trans isomers of the Pro residues is unaffected by the presence of carbohydrate. For Ac-(Gly-Pro-Thr(β-Gal))$_{10}$-NH_2, the kinetics of amide ^{1}H exchange with solvent deuterium indicate a slow rate of exchange for both the Gly and the Thr amide. The data are thus consistent with a model in which the carbohydrate stabilizes the triple helix through an occlusion of water molecules and by hydrogen bonding but not through an influence on the cis to trans isomer ratio.

The cuticle collagen of the deep-sea hydrothermal vent worm *Riftia pachyptila* is unusual in that the Yaa position of the triple-helix signature repeat, Gly-Xaa-Yaa, has a mono-, di-, or tri-galactoside of Thr, rather than the hydroxyproline characteristic of vertebrate collagen (*1*). We have shown previously that substitution of Thr(β-Gal) for threonine in the peptide Ac-(Gly-Pro-Thr)$_{10}$-NH_2 induces formation of a triple helix with an observed T_m of 39 °C (*2*). Thus, circular dichroism (CD),[1] analytical ultracentrifugation, and nuclear magnetic resonance (NMR) data were used to show that Ac-(Gly-Pro-Thr(β-Gal))$_{10}$-NH_2 can form a triple-helical structure. Here, we expand our NMR analysis to understand the structure-inducing effect of glycosylation. The structures of Ac-Gly-Pro-Thr(β-Gal)-NH_2, Ac-(Gly-Pro-Thr(β-Gal))$_5$-NH_2, and Ac-(Gly-Pro-Thr(β-Gal))$_{10}$-NH_2 are investigated by CD spectroscopy to show the dependence of chain length on the stability of the triple helix. Following the earlier protocols of others (*3*−*9*), NMR experiments are used to assign the resonances of the peptides Ac-Gly-Pro-Thr-NH_2 and Ac-Gly-Pro-Thr(β-Gal)-NH_2; these assignments allow subsequent assignment of protons that reside in the longer chain analogues, Ac-(Gly-Pro-Thr)$_5$-NH_2, Ac-(Gly-Pro-Thr(β-Gal))$_5$-NH_2, and Ac-(Gly-Pro-Thr(β-Gal)$_{10}$-NH_2. The effect of glycosylation on the proline peptide bond cis to trans isomer ratio is also investigated in the single-chain peptides. The data provided here gives evidence that the relative change in conformational energy on going from a single chain to the triple helix is not due to a change in the cis to trans isomer ratio of proline. We also investigated the kinetics of amide exchange for the various peptides and find that the rate of exchange of both the Gly and the Thr NH is slow only in Ac-(Gly-Pro-Thr(β-Gal))$_{10}$-NH_2. Although the (Yaa) Thr NH is projected to be solvent exposed (*4*), these data suggest that the carbohydrate, in the triple-helical structure, inhibits water from strong, direct, and rapid interaction with the backbone. This would result in an increase in stability by providing a more favorable environment for the amide to form intrachain or interchain hydrogen bonds rather than to form amide−water hydrogen bonds.

MATERIALS AND METHODS

FMOC-Glycosyl Amino Acid Preparation. The method of Elofsson was used in the preparation of N^{α}-9-fluorenylmethoxycarbonyl (FMOC)-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-Thr-OH (FMOC-Thr((Ac)$_4$-β-D-Gal)-OH) (*10*−*12*). Briefly, the glycosyl amino acid was prepared by reacting FMOC-Thr-OH (PerSeptive Biosystems) with β-D-galactose pentaacetate (Sigma) using boron-trifluoride di-

† This work was supported by a grant from Shriners Hospitals (to H.P.B.).

* Corresponding author. E-mail: peytond@pdx.edu.
‡ Oregon Health and Science University.
§ Shriners Hospital for Children.
|| Current address: Department of Biochemistry and Molecular Biophysics, Washington University in St. Louis School of Medicine, St. Louis, MO 63110.
⊥ Portland State University.

[1] Abbreviations: CD, circular dichroism; DIPEA, *N*,*N*-diisopropylethylamine; DQF-COSY, two-dimensional double-quantum-filtered J-correlated spectroscopy; FMOC, N^{α}-9-fluorenylmethyloxycarbonyl; Gal, galactose; HATU, (*O*-(7-azabenzotriazol-1-yl)-1.1.3.3-tetramethyluronium hexafluorophosphate; HPLC, high-performance liquid chromatography; MALDI-TOF, matrix-assisted laser desorption/ionization-time-of-flight; NMR, nuclear magnetic resonance; NOESY, two-dimensional nuclear Overhauser effect spectroscopy; ROESY, two-dimensional rotating-frame nuclear Overhauser spectroscopy; TFA, trifluoroacetic acid; TPPI, time-proportional phase incrementation; TOCSY, two-dimensional total correlation J spectroscopy.

10.1021/bi027050w CCC: $25.00
Published on Web 03/19/2003

ethyl-etherate (Aldrich) as a catalyst in dry acetonitrile (Aldrich). The glycosyl amino acid was then purified by preparative HPLC (Phenomenex Luna C8, 5u, 100 Å, 250 × 22.5 mm i.d.; A/B 55:45, v/v (A: 0.1% TFA/H_2O, B: 0.1% TFA/CH_3CN) and dried three times from dichloromethane and once from toluene to remove traces of trifluoroacetic acid. The product was then characterized by electrospray mass spectrometry, in addition to ^{13}C and ^{1}H NMR.

Peptide Synthesis. Peptides were synthesized using a Milligen 9050 peptide synthesizer. Peptide couplings were done on a PAL-PEG-PS resin (Perseptive Biosystems, 0.16 mmol/g) to afford C-terminal amides. The glycopeptide and peptide were synthesized using FMOC dipeptides of Gly-Pro-OH (Novabiochem, 4.0 equiv), FMOC-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-Thr-OH (1.3−2.0 equiv), and FMOC-Thr-OH (4.0 equiv) and HATU (O-(7-azabenzotriazol-1-yl)-1.1.3.3-tetramethyluronium hexafluorophosphate, Perseptive Biosystems, 4.0 equiv)/DIPEA mediated peptide couplings. The peptides were cleaved from the resin and purified by semipreparative HPLC (Phenomenex Luna C18, 5 μ, 100 Å, 250 × 10 mm i.d.). The glycopeptide was subsequently treated with 6 mM sodium methoxide in methanol overnight to remove the acetyl groups and again purified by semipreparative HPLC. All peptides were verified by analytical HPLC, electrospray, or MALDI-TOF mass spectrometry and amino acid analysis.

CD Measurements. Circular dichroism spectra were recorded on an Aviv 202 spectropolarimeter using a thermostated cell of 1 mm path length (Hellma). All measurements were performed in water at 5 °C, and the concentrations were all 100 μM. Peptide concentrations were determined by amino acid analysis. The thermal titrations were recorded from 5−70 °C at 223 nm by raising the temperature at a rate of 12 °C/h.

NMR Measurements. NMR spectra were recorded on a Bruker AMX-400 spectrometer, operating at 400.14 MHz, between pH 4.3 and 5.2 and at 25 °C unless otherwise noted. The 90° pulse width was 8 μs, and a low-power 2 s presaturation pulse was applied to suppress the H_2O (HOD) resonance. The spectra were recorded as 16 384 points for the 1-D spectra and as 1024 × 512 data point sets for the 2-D spectra. The NOESY data were collected with TPPI (*13*) in the indirect dimension, at mixing times between 30 and 150 ms, and a total recording time of about 17 h. ROESY data (*14*) were collected similarly but using a 200 ms mixing time. TOCSY data (*15*) were collected with various mixing times, ranging from 30 to 90 ms. The data were processed with Swan-MR (*16*, *17*) or nmrPipe (*18*) to 1024 × 1024 real data sets after application of a 65°-phase-shifted $\sin^2$ function and Fourier transformation for the 2-D spectra; baselines were straightened with polynomials as needed. Spectra were referenced to 0 ppm via internal 2,2-dimethylsilapentane-5-sulfonate or via the water resonance (4.76 ppm at 298 K). Final visualization and analyses of the 2-D data sets were performed using NMRView (*19*) or Swan-MR. Peptides were lyophilized and then quickly dissolved in D_2O at pH 5.2, and 1-D spectra were recorded at different time points at 25 °C throughout. The rates were obtained by monitoring the change in peak height as a function of time, fitting the data to a single-exponential function (*20*).

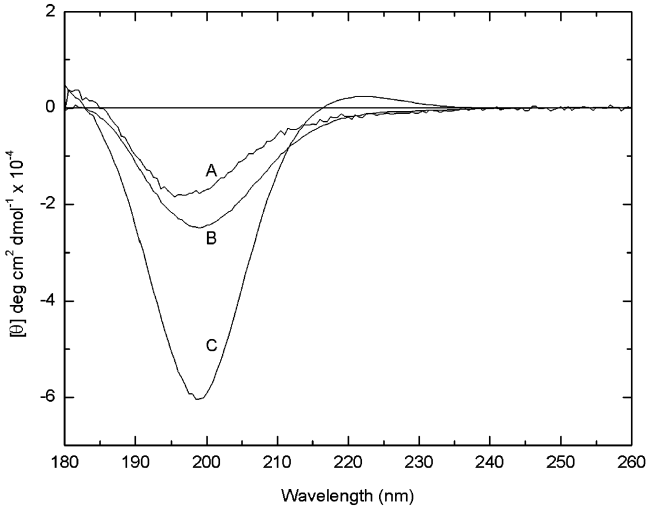

FIGURE 1: CD spectra of collagen-like peptides at 5 °C. (A) Ac-(Gly-Pro-Thr(β-D-Gal))$_1$-NH_2. (B) Ac-(Gly-Pro-Thr(β-D-Gal))$_5$-NH_2. (C) Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH_2. Note the similarity in shape between panels A and B, indicating that these peptides do not adopt triple-helical conformations under these conditions, but the peptide Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH_2 in panel C does.

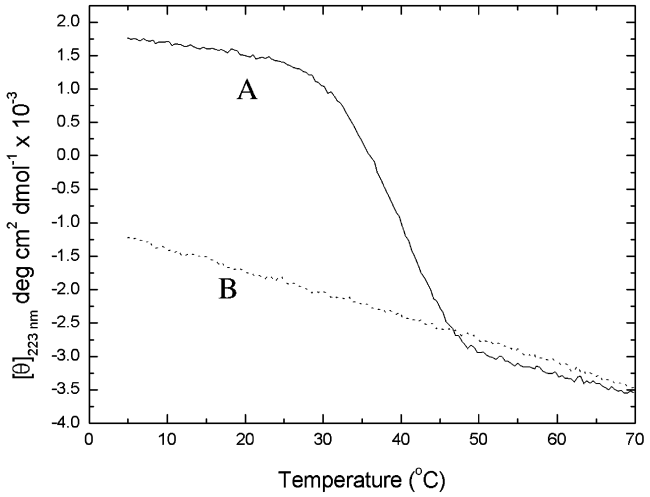

FIGURE 2: Ellipticity at 223 nm as a function of temperature. (A) Solid line: Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH_2. (B) Broken line: Ac-(Gly-Pro-Thr(β-D-Gal))$_5$-NH_2. Only the longer glycosylated peptide exhibits a cooperative transition within this temperature range in water solvent.

RESULTS

As shown in Figure 1, CD spectroscopy at 5 °C indicates that Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH_2 can readily form a triple helix but that the Ac-(Gly-Pro-Thr(β-D-Gal))$_5$-NH_2 is more like Ac-(Gly-Pro-Thr(β-D-Gal))$_1$-NH_2, which clearly does not form a triple helix. Figure 2 shows, in even more dramatic form, the different behaviors of Ac-(Gly-Pro-Thr-(β-D-Gal))$_5$-NH_2 and Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH_2. From these experiments, it can be seen that Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH_2 has a T_m ~38 °C but that Ac-(Gly-Pro-Thr-(β-D-Gal))$_5$-NH_2 has no observable T_m, even to near the freezing point of the water solvent.

Figure 3 shows a comparison of the three model molecular systems, Ac-(Gly-Pro-Thr(β-D-Gal))$_1$-NH_2, Ac-(Gly-Pro-Thr-(β-D-Gal))$_5$-NH_2, and Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH_2. The bottom spectrum (trace 3A) shows the form of the upfield region of Ac-(Gly-Pro-Thr(β-D-Gal))$_1$-NH_2, which is not

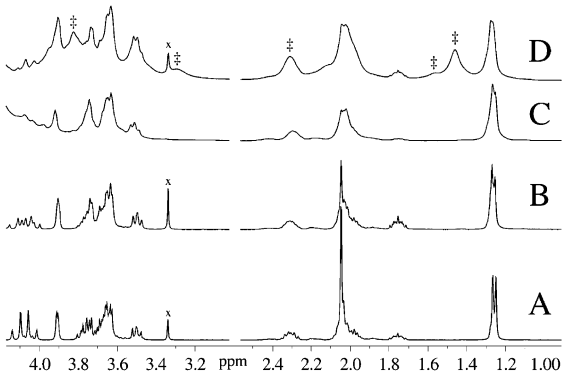

FIGURE 3: 1-D ^{1}H NMR spectra of collagen model peptides. (A) Ac-(Gly-Pro-Thr(β-D-Gal))$_1$-NH_2, which is not capable of forming a triple helix. (B) Ac-(Gly-Pro-Thr(β-D-Gal))$_5$-NH_2. The lines are somewhat broadened, relative to those of Ac-(Gly-Pro-Thr(β-D-Gal))$_1$-NH_2, as is expected for the larger molecular weight of Ac-(Gly-Pro-Thr(β-D-Gal))$_1$-NH_2. (C) Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH_2, after taking the sample to 60°, then dropping the temperature to 25° to record the spectrum immediately. (D) Ac-(Gly-Pro-Thr-(β-D-Gal))$_{10}$-NH_2, after allowing the sample to incubate at 4 °C for 3 days. The most obvious of the new peaks arising from presence of the triple helix are marked with the symbol ‡.

capable of forming a triple helix. The amino acid peaks are readily assignable to the usual position of each of the protons in the constituent compounds (Gly, Pro, and Thr), as summarized in Table S-I in Supporting Information. The NMR assignment strategy follows the protocols established by others (*3*−*9*). Thus, individual spin systems are found from DQF-COSY and TOCSY spectra and their known approximate chemical shifts. It is not possible to obtain the complete sequence-specific assignment of each of the 30 residues in Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH_2; the overlap is so strong that assignments are made for the major composite resonance positions. This set of assignments formed the basis for the assignments that follow. Trace 3B shows the spectrum of Ac-(Gly-Pro-Thr(β-D-Gal))$_5$-NH_2, which is almost identical to that of Ac-(Gly-Pro-Thr(β-D-Gal))$_1$-NH_2, as can be seen by comparing to Trace 3A. The lines are broadened relative to those of Ac-(Gly-Pro-Thr(β-D-Gal))$_1$-NH_2, as expected for the larger molecular weight of Ac-(Gly-Pro-Thr(β-D-Gal))$_5$-NH_2, but the lines are in nearly the same positions as in Ac-(Gly-Pro-Thr(β-D-Gal))$_1$-NH_2. This result is completely expected, given the lack of a T_m for Ac-(Gly-Pro-Thr(β-D-Gal))$_5$-NH_2, as recorded by CD spectroscopy.

Trace 3C is the upfield region of Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH_2, after taking this sample to 60°, then dropping the temperature to 25° to record the spectrum. This shows how the spectrum appears without the presence of significant amounts of triple helix: the peaks are about the same positions as they are in traces 3A,B for Ac-(Gly-Pro-Thr-(β-D-Gal))$_1$-NH_2 and Ac-(Gly-Pro-Thr(β-D-Gal))$_5$-NH_2 but with the additional line broadening expected for this larger molecule. After allowing another sample to incubate at 4 °C for 3 days, the result is trace 3D. Now this spectrum shows a new set of resonance positions and intensity shifts, as shown by the symbol ‡ for the more obvious of these peaks. We had used the new peak near 1.4 ppm and its ROE's, as evidence for triple-helix formation previously (*2*), but here we attempted to assign the resonances that appear for the triple-helical state. We note that the amount of folded material is enhanced by about 2-fold when the sample is

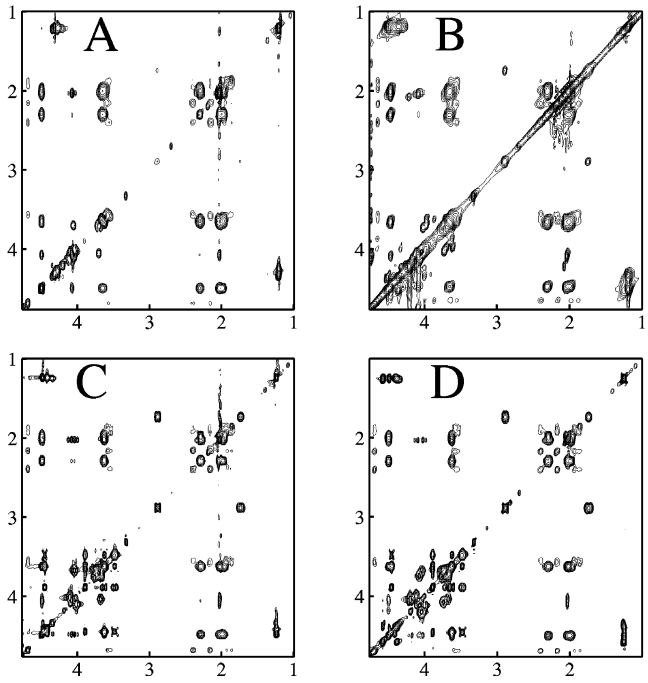

FIGURE 4: 2-D ^{1}H TOCSY regions for various model peptides that do not form a triple helix. (A) Ac-(Gly-Pro-Thr((Ac)$_4$-β-D-Gal))$_1$-NH_2. (B) Ac-(Gly-Pro-Thr(β-D-Gal))$_1$-NH_2. (C) Ac-(Gly-Pro-Thr)$_1$-NH_2. (D) Ac-(Gly-Pro-Thr((Ac)$_4$-β-D-Gal))$_5$-NH_2.

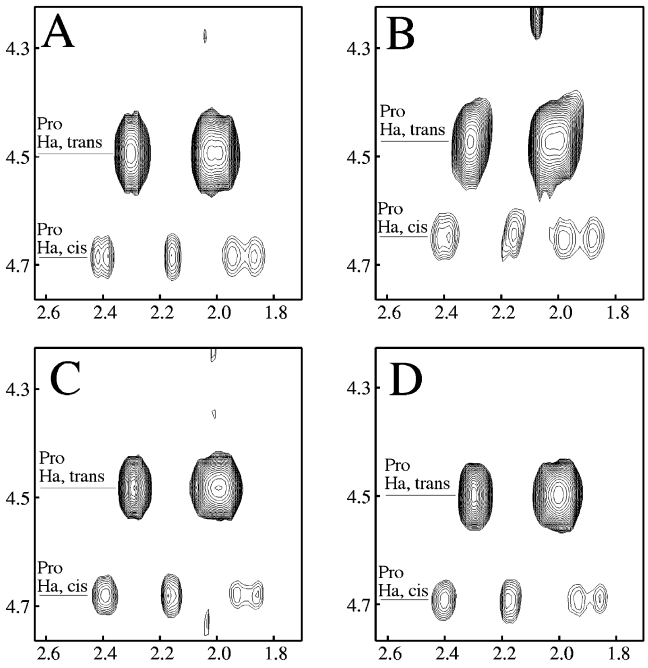

FIGURE 5: Expansions of the TOCSY spectra from Figure 4. (A) Ac-(Gly-Pro-Thr((Ac)$_4$-β-D-Gal))$_1$-NH_2. (B) Ac-(Gly-Pro-Thr(β-D-Gal))$_1$-NH_2. (C) Ac-(Gly-Pro-Thr)$_1$-NH_2. (D) Ac-(Gly-Pro-Thr-((Ac)$_4$-β-D-Gal))$_5$-NH_2. The cross-peaks are between the Pro Hα and the Pro Hβ,γ resonances.

dissolved in D_2O rather than in H_2O, analogous to what has been reported by others for (Pro-Pro-Gly)$_{10}$ (*21*, *22*).

Figure 4 shows the upfield region of 2D-TOCSY data sets recorded for Ac-(Gly-Pro-Thr)$_1$-NH_2, Ac-(Gly-Pro-Thr(β-D-Gal))$_1$-NH_2, protected Ac-(Gly-Pro-Thr(β-D-Gal))$_1$-NH_2, and Ac-(Gly-Pro-Thr(β-D-Gal))$_5$-NH_2, none of which forms a triple helix. From these spectra and corresponding NOESY spectra (not shown), assignments can be made for the nonlabile protons of the peptide portion of these compounds.

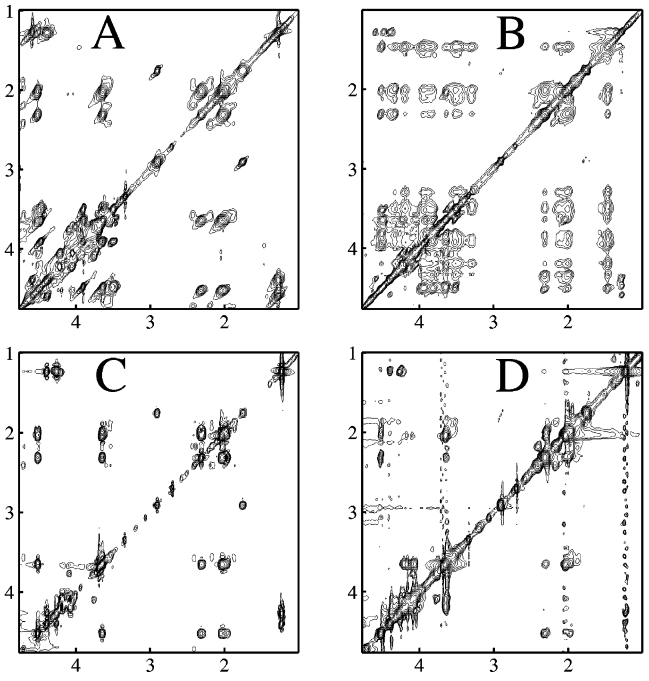

FIGURE 6: 2-D ^{1}H TOCSY (A and C) and NOESY (B and D) spectra of Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH_2. Spectra A and B are for the folded, triple-helical form, and spectra C and D are for the denatured form.

These assignments are summarized in Table S-I in the Supporting Information. Figure 5 shows an expanded region of these spectra; this set of spectra demonstrates clearly that the cross-peak intensities arising from the cis prolyl isomer conformation is present to about the same extent in each of these peptides. Strong overlap of resonances precludes integrating each of the resonances from the 1-D spectra, which would have provided the most accurate values of the cis to trans ratios. However, there is sufficient resolution between cis and trans isomer pair resonances for the Pro H_β at about 2.3−2.4 ppm to find about 12 and 15% trans for the unglycosylated and glycosylated forms, respectively. These numbers are unchanged whether the Gal groups are protected with acetates or not. The importance of these findings is that the cis proline conformation is present to nearly the same extent, whether the Thr residue is glycosylated or not.

Figure 6 is a comparison of 2-D NOESY and TOCSY data sets for Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH_2, recorded after allowing this molecule to attain a triple-helix conformation. This pair of spectra allows for assigning many of the predominating resonances, based upon the assignments made for the previously assigned shorter peptides and their NMR spectra depicted in Figure 4. It should be noted that the TOCSY spectrum is dominated by a different set of resonances than is the NOESY spectrum in Figure 6. This could be due to the different mechanisms giving rise to the cross-peaks, and so the cross-peak intensities reflect molecular mobility differently. Thus, the NOESY (TOCSY) cross-peaks would be expected to be strongest from the least (most) mobile parts of the triple helix; the NOESY spectra will be dominated by the middle of the structure, while the TOCSY will be dominated by the ends. There is, nevertheless, sufficient TOCSY observed for all of the resonances to help assign many of the strongest NOESY cross-peaks (see Table S-I in Supporting Information for the resulting resonance assignments). There are strong differences between the NOESY cross-peaks, depending upon whether the molecules are in the triple-helix form or not (Figure 6A,B as compared to 6C,D). The NOEs in the triple-helical system are large and extend from most nuclei to the other nuclei, as result from proximity of protons across chains in the triple helix, as well as many short intrachain distances. These NOEs persist down to the shortest mixing time of 30 ms.

To investigate the kinetics of amide proton exchange, peptides were lyophilized from H_2O, dissolved in 90% D_2O/10% H_2O, and spectra were recorded as a function of time. The appearance of the amide proton region of Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH_2 is shown in Figure 7. Both the Thr and the Gly peptide NH resonances take several hours to exchange H for D at room temperature. Integration of these

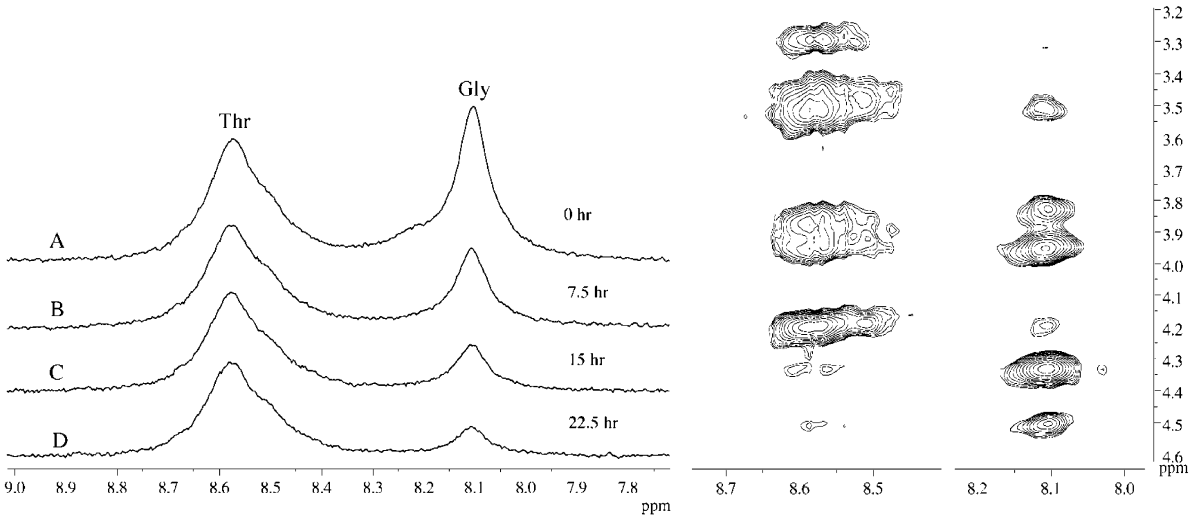

FIGURE 7: ^{1}H NMR spectra recorded after dispersing the triple helical form of Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH_2 into D_2O solvent. Left: (A−D) time-dependent exchange of Thr (8.57 ppm) and Gly (8.09 ppm) NH protons. Right: NOESY region, showing NOEs from these amide protons. Note that the intensities of these peaks relative to the aliphatic resonances in trace A (0 h) indicate that about 50% of both the Thr and the Gly amide H peaks have already exchanged for D; the remaining amides are presumably those toward the middle of the triple-helical chains.

peaks relative to the aliphatic resonances indicates that about 50% of both of these amide peaks have already exchanged before recording the first spectrum in Figure 7. (The fast exchanging amides are presumably those nearest the ends of the chains.) The first-order exchange rate of the amide proton of Thr(β-D-Gal) (k_{ex} $\sim 2.0 \times 10^{-4}$ min^{-1}) is similar to but even slower than for the amide proton of glycine (k_{ex} $\sim 1.3 \times 10^{-3}$ min^{-1}). For comparison, we note that the unfolded peptides exchange with solvent within the time it takes to dissolve, insert into the spectrometer, and record a 1-D spectrum. We also note that the resonances for Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH$_2$ are actually quite complicated, with some of the line width apparent in Figure 7 resulting from different chemical shifts, presumably from the various positions, ranging from the most interior toward the end amide positions. (Also, recall that the resonances nearest the ends have most likely already exchanged and so are not even represented in these spectra.) This heterogeneity is further illustrated by the shapes of the NOESY cross-peaks from these amide protons (right panel in Figure 7). This is not surprising, either from first principles, or from analogous results from ^{1}H and ^{1}H/^{15}N correlation spectroscopy of other collagen-like peptides (*23*−*25*).

DISCUSSION

The data presented here indicate that the proline cis to trans isomer ratio is unaffected by the presence of carbohydrate. Thus, in the case of Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH$_2$, the stability of the triple helix is likely due to other mechanisms known to influence collagen stability, such as proline ring pucker (*26*), inductive effects (*27*), or hydrogen bonding (*21*, *22*).

The triple helix has intermolecular hydrogen bonds and the possibility of stabilization by water-bridged hydrogen bond networks (*4*). Bridging water has been asserted to explain the relative increase in stability of (Pro-Pro-Gly)$_{10}$ in D$_2$O over H$_2$O solvent (*22*). The reason(s) for this are not yet clear but may arise from changes in dielectric constant, increased hydrophobic interactions, changes in amide−carbonyl, amide−water, or carbonyl−water bonding strengths (*28*). While general D$_2$O stabilization of proteins is known (*29*−*31*), the effect on Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH$_2$ is substantial and happens before the slowly exchanging Gly and Thr amide hydrogens are exchanged for deuterium (Figure 7). Furthermore, the species is a trimer even when in H$_2$O, as evidenced from ultracentrifuge data (*32*). So the D$_2$O solvent effect here is to enhance the triple helix throughout the trimeric assembly.

Both the Gly and the Thr amide hydrogens of Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH$_2$ from the middle of the sequence are strongly protected from solvent exchange. This is initially surprising because the NH in the Yaa position is projected to be exposed to solvent, but the Gly NH participates in an interchain hydrogen bond (*4*, *26*, *33*, *34*). (The intrinsic rates for amide proton exchange of unstructured Thr-NH-Gly and Gly-NH-Pro are predicted to be similar (*20*).) These slow rates of amide hydrogen exchange with solvent are only found for the triple helical structure, with much faster exchange rates found for Ac-(Gly-Pro-Thr(β-D-Gal))$_5$-NH$_2$; this eliminates the slight possibility that the slow exchange is intrinsic to Thr(β-D-Gal) in a single chain (Pro-Thr(β-D-Gal)-Gly)$_n$. Slow amide exchange has been observed in the designed collagen-like peptide T3−785 (*4*). The possibility

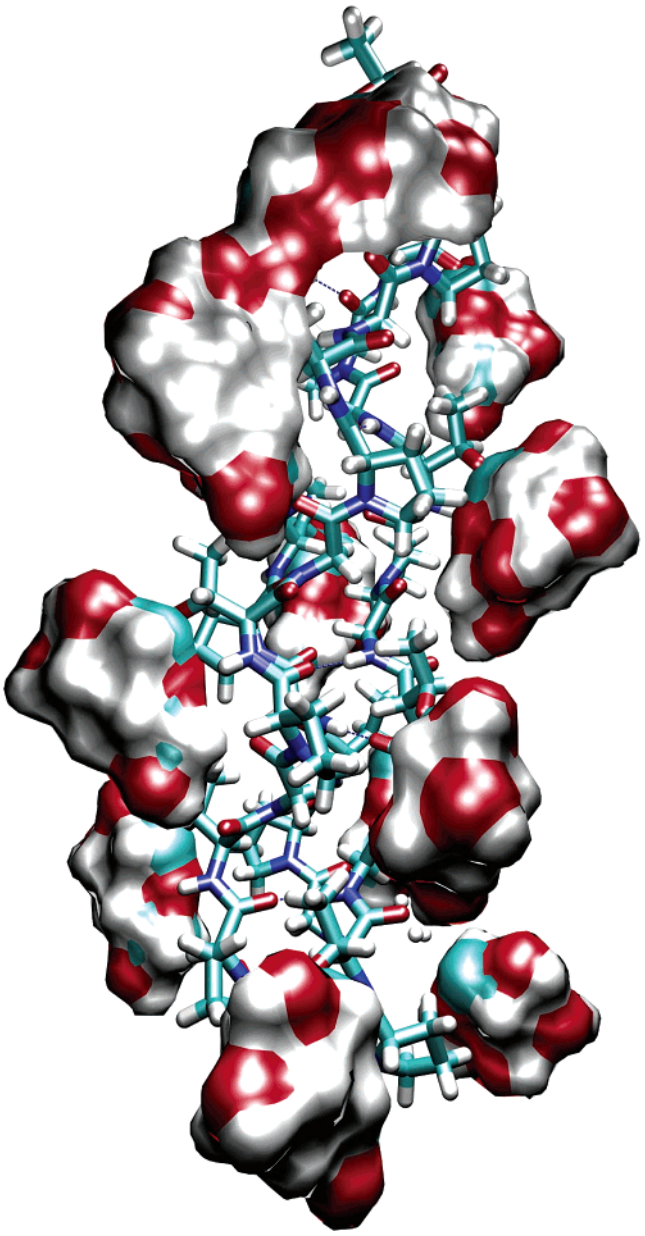

FIGURE 8: Model of Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH$_2$. Molecular modeling was performed using Biosym software programs from MSI (San Diego, CA). The structure was modeled using structure (Ac-(Gly-Pro-Pro)$_4$-NHMe)$_3$ as a template for the backbone conformation (*47*). The replacement of the Yaa position Pro for Thr was done with the Biopolymer program. Side-chain conformations of the Thr(β-D-Gal) were optimized by manually selecting the lowest energy conformer within the Biopolymer program. Subsequent energy minimization was performed with the Discover program, using the consistent-valence force field (CVFF). The backbone was tethered in place throughout with a force constant of 100 kcal/Å^2. A distant-dependent dielectric constant of 1.0 was used, and the side chains of all amino acids were uncharged, to approximate the effects of solvent shielding (*48*). The protocol for minimization was as follows: the method of steepest descents was used until a maximum derivative of <10 kcal/Å was reached, with the charge term included. Next, the method of conjugate gradients was used for 500 iterations until a maximum derivative of <1.0 kcal/Å was reached, with charges and cross-term energies included. Finally, the va09a quasi Newton−Raphson method was used for 500 iterations until a maximum derivative of <0.01 kcal/Å was achieved, with charges, cross-terms, and a Morse bond potential included.

of a strongly bound water inhibiting exchange was raised, and indeed a water molecule is observed in its crystal structure, forming an interchain hydrogen bond between the NH of the Xaa position and the CO of a neighboring Gly (*35*). A general role for water stabilizing the triple helix is still uncertain because 4(*R*)-fluoroproline also leads to substantial collagen triple-helix stability (*27*, *36*). Also, there is disagreement about the number of water molecules observed in collagen X-ray crystal structures, even from the same peptide (*33*, *34*, *37*). Finally, the triple-helix is more stable in anhydrous polyhydric solvents (*38*).

We believe that the most likely explanations for the even slower exchange of the triple-helical Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH_2 Thr amide hydrogen than the interchain hydrogen-bonded Gly amide hydrogen is that the amide proton of threonine forms a hydrogen bond with the galactose and/or that the carbohydrate almost completely excludes solvent from the threonine amide. Thus, the stabilizing influence of glycosylation is likely due to a change in the activity of water surrounding the polypeptide backbone, disfavoring amide−water hydrogen bonds (*32*). The interchain Gly amide hydrogen bond to the carboxyl group of a proline has a favorable geometry (*39*), so the carbohydrate would very effectively protect these from forming peptide−water hydrogen bonds, favoring the peptide−peptide hydrogen bonds that run like a ladder through the helix (*40*). A model for Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH_2 is shown in Figure 8. It is well-known that carbohydrate can stabilize or change secondary and tertiary structures (*41*−*44*), as well as folding processes (*45*), and that the carbohydrate can sweep a very large effective volume (*41*, *43*, *46*). Thus, Figure 8 probably under represents the degree to which the β-D-Gal can inhibit rapid solvent access to the amino acid residues. There is clear opportunity for the β-D-Gal to form hydrogen bonds, including an intraresidue hydrogen bond with the Thr amide, without a large distortion of the triple helix. Glycosylation could favor the conformations that maximize the interchain interactions, while minimizing water−peptide hydrogen bonds. Glycosylation has been shown to organize other polypeptides, perhaps most dramatically a mucin motif with α-*O*-GalNAc (but not β-*O*-GalNAc) linkages (*44*), for which the GalNAc acetyl NH groups were implicated in hydrogen bonding to the polypeptide backbone. This is similar to what we present for the effect of glycosylation of Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH_2, even though we have β-*O*-Gal linkages. It may be that the effect of organizing the three polypeptide chains amplifies the difference in stability between glycosylated and unglycosylated peptides.

ACKNOWLEDGMENT

The authors thank Prof. Juergen Engel for reading a draft of this manuscript.

SUPPORTING INFORMATION AVAILABLE

Table of chemical shifts of protons for the collagen model peptides. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Mann, K., Mechling, D. E., Bächinger, H. P., Eckerskorn, C., Gaill, F., and Timpl, R. (1996) *J. Mol. Biol. 261*, 255−66.
2. Bann, J. G., Peyton, D. H., and Bächinger, H. P. (2000) *FEBS Lett. 473*, 237−40.
3. Anachi, R. B., Siegel, D. L., Baum, J., and Brodsky, B. (1995) *FEBS Lett. 368*, 551−5.
4. Fan, P., Li, M. H., Brodsky, B., and Baum, J. (1993) *Biochemistry 32*, 13299−309.
5. Feng, Y., Melacini, G., Taulane, J. P., and Goodman, M. (1996) *Biopolymers 39*, 859−72.
6. Feng, Y., Melacini, G., and Goodman, M. (1997) *Biochemistry 36*, 8716−24.
7. Li, M. H., Fan, P., Brodsky, B., and Baum, J. (1993) *Biochemistry 32*, 7377−87.
8. Long, C. G., Braswell, E., Zhu, D., Apigo, J., Baum, J., and Brodsky, B. (1993) *Biochemistry 32*, 11688−95.
9. Melacini, G., Feng, Y., and Goodman, M. (1997) *Biochemistry 36*, 8725−32.
10. Elofsson, M., Walse, B., and Kihlberg, J. (1991) *Tetrahedron Lett. 32*, 7613−16.
11. Kihlberg, J., Elofsson, M., and Salvador, L. A. (1997) *Methods Enzymol. 289*, 221−45.
12. Salvador, L. A., Elofsson, M., and Kihlberg, J. (1995) *Tetrahedron 51*, 5643−56.
13. Marion, D., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967−74.
14. Bothner-By, A. A., Stephens, R. L., Lee, J., Warren, C. D., and Jeanloz, R. W. (1984) *J. Am. Chem. Soc. 106*, 811−3.
15. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson. 65*, 355−60.
16. Balacco, G. (1994) *J. Chem. Inf. Comput. Sci. 34*, 1235−41.
17. Balacco, G. (2000) *Mol. Biol. Today 1*, 23−38.
18. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−93.
19. Johnson, B. A., and Blevins, R. A. (1994) *J. Biomol. NMR 4*, 603−14.
20. Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) *Proteins: Struct., Funct., Genet. 17*, 75−86.
21. Gough, C. A., Anderson, R. W., and Bhatnagar, R. S. (1998) *J. Biomol. Struct. Dyn. 15*, 1029−37.
22. Gough, C. A., and Bhatnagar, R. S. (1999) *J. Biomol. Struct. Dyn. 17*, 481−91.
23. Bhate, M., Wang, X., Baum, J., and Brodsky, B. (2002) *Biochemistry 41*, 6539−47.
24. Melacini, G., Bonvin, A. M., Goodman, M., Boelens, R., and Kaptein, R. (2000) *J. Mol. Biol. 300*, 1041−9.
25. Yu, Y. C., Roontga, V., Daragan, V. A., Mayo, K. H., Tirrell, M., and Fields, G. B. (1999) *Biochemistry 38*, 1659−68.
26. Vitagliano, L., Berisio, R., Mastrangelo, A., Mazzarella, L., and Zagari, A. (2001) *Protein Sci. 10*, 2627−32.
27. Holmgren, S. K., Bretscher, L. E., Taylor, K. M., and Raines, R. T. (1999) *Chem. Biol. 6*, 63−70.
28. Loh, S. N., and Markley, J. L. (1994) *Biochemistry 33*, 1029−36.
29. Parker, M. J., and Clarke, A. R. (1997) *Biochemistry 36*, 5786−94.
30. Chakrabarti, G., Kim, S., Gupta, M. L., Jr., Barton, J. S., and Himes, R. H. (1999) *Biochemistry 38*, 3067−72.
31. Itzhaki, L. S., and Evans, P. A. (1996) *Protein Sci. 5*, 140−6.
32. Bann, J. G., and Bächinger, H. P. (2000) *J. Biol. Chem. 275*, 24466−9.
33. Nagarajan, V., Kamitori, S., and Okuyama, K. (1998) *J. Biochem. (Tokyo) 124*, 1117−23.
34. Nagarajan, V., Kamitori, S., and Okuyama, K. (1999) *J. Biochem. (Tokyo) 125*, 310−8.
35. Kramer, R. Z., Bella, J., Mayville, P., Brodsky, B., and Berman, H. M. (1999) *Nat. Struct. Biol. 6*, 454−7.
36. Holmgren, S. K., Taylor, K. M., Bretscher, L. E., and Raines, R. T. (1998) *Nature 392*, 666−7.
37. Vitagliano, L., Berisio, R., Mazzarella, L., and Zagari, A. (2001) *Biopolymers 58*, 459−64.
38. Engel, J., Chen, H. T., Prockop, D. J., and Klump, H. (1977) *Biopolymers 16*, 601−22.
39. Persikov, A. V., Ramshaw, J. A., and Brodsky, B. (2000) *Biopolymers 55*, 436−50.
40. Brown, F. R., III, Hopfinger, A. J., and Blout, E. R. (1972) *J. Mol. Biol. 63*, 101−15.

41. Wormald, M. R., Petrescu, A. J., Pao, Y. L., Glithero, A., Elliott, T., and Dwek, R. A. (2002) *Chem. Rev. 102*, 371−86.
42. Van den Steen, P., Rudd, P. M., Dwek, R. A., and Opdenakker, G. (1998) *Crit. Rev. Biochem. Mol. Biol. 33*, 151−208.
43. Bush, C. A., Martin-Pastor, M., and Imberty, A. (1999) *Annu Rev. Biophys. Biomol. Struct. 28*, 269−93.
44. Coltart, D. M., Royyuru, A. K., Williams, L. J., Glunz, P. W., Sames, D., Kuduk, S. D., Schwarz, J. B., Chen, X. T., Danishefsky, S. J., and Live, D. H. (2002) *J. Am. Chem. Soc. 124*, 9833−44.
45. O'Connor, S. E., Pohlmann, J., Imperiali, B., Saskiawan, I., and Yamamoto, K. (2001) *J. Am. Chem. Soc. 123*, 6187−8.
46. Weller, C. T., Lustbader, J., Seshadri, K., Brown, J. M., Chadwick, C. A., Kolthoff, C. E., Ramnarain, S., Pollak, S., Canfield, R., and Homans, S. W. (1996) *Biochemistry 35*, 8815−23.
47. Nemethy, G., Gibson, K. D., Palmer, K. A., Yoon, C. N., Paterlini, G., Zagari, A., Rumsey, S., and Scheraga, H. A. (1992) *J. Phys. Chem. 96*, 6472−84.
48. Vitagliano, L., Nemethy, G., Zagari, A., and Scheraga, H. A. (1993) *Biochemistry 32*, 7354−9.

BI027050W