Sweet is stable: glycosylation stabilizes collagen

James G. Bann[a], David H. Peyton[c], Hans Peter Bächinger[a,b,*]

[a] *Department of Biochemistry and Molecular Biology, Oregon Health Sciences University, 3108 SW Sam Jackson Park Road, Portland, OR 97201, USA*
[b] *Research Department, Shriners Hospital for Children, 3101 SW Sam Jackson Park Road, Portland, OR 97201, USA*
[c] *Department of Chemistry, Portland State University, Portland, OR 97267, USA*

Received 20 March 2000

Edited by Felix Wieland

Abstract For most collagens, the melting temperature (T_m) of the triple-helical structure of collagen correlates with the total content of proline (Pro) and 4-*trans*-hydroxyproline (Hyp) in the Xaa and Yaa positions of the -Gly-Xaa-Yaa- triplet repeat. The cuticle collagen of the deep-sea hydrothermal vent worm *Riftia pachyptila*, despite a very low content of Pro and Hyp, has a relatively high thermal stability. Rather than Hyp occupying the Yaa position, as is normally found in mammalian collagens, this position is occupied by threonine (Thr) which is *O*-glycosylated. We compare the triple-helix forming propensities in water of two model peptides, Ac-(Gly-Pro-Thr)$_{10}$-NH$_2$ and Ac-(Gly-Pro-Thr(Galβ))$_{10}$-NH$_2$, and show that a collagen triple-helix structure is only achieved after glycosylation of Thr. Thus, we show for the first time that glycosylation is required for the formation of a stable tertiary structure and that this modification represents an alternative way of stabilizing the collagen triple-helix that is independent of the presence of Hyp.

Key words: Collagen; Glyco-peptide; Structure; Thermal stability; *Riftia pachyptila*

1. Introduction

The collagens of invertebrates have been shown to exhibit unusual and interesting properties [1,2]. An example of this is from the deep-sea hydrothermal vent worm *Riftia pachyptila*. This organism lives under extreme conditions (high pressure, low oxygen and steep temperature gradients) but is protected from its environment by a thick cuticle [3,4]. The *R. pachyptila* cuticle is mainly composed of a collagen that forms a plywood-like network of fibrils and exhibits a unique amino acid composition. While having a low content of proline (Pro) and 4-*trans*-hydroxyproline (Hyp) ($\sim$5%), this collagen has a very high content of threonine (Thr) residues ($>$18%) [5]. Biochemical analyses and partial sequencing of the cuticle collagen's triple-helical region revealed that Thr occupied the Yaa position of the -Gly-Xaa-Yaa- triplet repeat and the Thr residues were glycosylated, mainly with di- and tri-saccharides of galactose [6]. In most collagens, there is a positive correlation between the total content of Pro and Hyp imino acids and the melting temperature (T_m) of the triple-helix [7,8]. However, the *R. pachyptila* cuticle collagen, with an observed T_m of 37°C, deviates considerably from this trend (Fig. 1).

The finding of a di- and tri- pattern of galactose carbohydrate units is consistent with the pattern found in the cuticle collagens of both *Lumbricus terrestris* (earthworm) and *Nereis virens* (clamworm) [9,10]. Also consistent with the annelid cuticle collagens is the position of Hyp in the -Gly-Xaa-Yaa- triplet repeat, occurring only in the Xaa position [11,12]. In vertebrates, Hyp only occurs in the Yaa position, and Hyp in the Xaa position is thought to either not contribute to or to destabilize the thermal stability of the triple-helix [11,13]. For example, although the amount of Hyp in the *L. terrestris* collagen is very high ($\sim$16–18%), this collagen has a low thermal stability (T_m = 22°C) [14]. Therefore, in *R. pachyptila*, glycosylation of Thr must somehow compensate for the lack of Hyp in the Yaa in order to achieve the high-observed thermal stability.

In these studies our main goal was to develop model collagen-like peptides of the *R. pachyptila* cuticle collagen that would allow us to determine the influence of glycosylation on the stability of the triple-helix. In addition, *O*-glycosylation of proteins has been shown to impart an enhanced resistance to enzymatic degradation and both the *L. terrestris* and *N. virens* cuticle collagens in the native state are resistant to degradation by clostridial collagenase [11,15]. Thus, these collagen-like peptides may also exhibit an enhanced resistance to proteolysis and could lead to the development of novel biomaterials [16].

2. Materials and methods

2.1. Peptide synthesis

The synthesis of N^{α}-9-fluorenylmethoxycarbonyl (FMOC)-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-Thr-OH was done according to [17–19], purified by preparative high-performance liquid chromatography (HPLC) and characterized by electrospray mass spectrometry, ^{1}H NMR (nuclear magnetic resonance) and ^{13}C NMR. The glycopeptide and peptide were synthesized on a PAL-PEG-PS resin (Perseptive Biosystems, 0.16 mmol/g) using FMOC dipeptides of Gly-Pro-OH (Novabiochem, 4.0 eq), FMOC-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-Thr-OH (1.3–2.0 eq) and FMOC-Thr-OH (4.0 eq) and HATU (*O*-(7-azabenzotriazol-1-yl)-1.1.3.3-tetramethyluronium hexafluorophosphate, Perseptive Biosystems, 4.0 eq) mediated amino acid couplings on a Milligen 9050 peptide synthesizer. The peptides were cleaved from the resin and purified by semi-preparative HPLC. The glycopeptide was subsequently treated with 6 mM sodium methoxide in methanol overnight to remove the acetyl groups and again purified by semi-preparative HPLC. The peptides were verified by analytical HPLC, electrospray mass spectrometry, ^{1}H-NMR and amino acid analysis.

2.2. Circular dichroism (CD) measurements

CD spectra were recorded on an Aviv 202 spectropolarimeter using a thermostatted cell of 1 mm pathlength (Hellma). All measurements were performed in water. The concentrations of peptides were determined by amino acid analysis. The thermal transitions were recorded

*Corresponding author. Fax: (1)-503-221 3451.
E-mail: hpb@shcc.org

PII: S0014-5793(00)01493-9

from 5 to 70°C at 223 nm by raising the temperature at a rate of 12°C/h.

2.3. Analytical ultracentrifugation

Sedimentation equilibrium measurements were carried out on a Beckman Model E analytical ultracentrifuge, using double-sector cells. Peptides were run at two speeds (26 000 and 34 000 rpm), at three concentrations (0.5, 0.35 and 0.2 mg/ml) in water, monitoring at 230 nm at 25.0°C. Data were fitted using Scientist® (Micromath, Salt Lake City, UT, USA) with a non-linear least squares algorithm, assuming a partial specific volume of 0.667 cm^3/g based upon the amino acid sequence and the partial specific volume of galactose.

2.4. NMR measurements

NMR spectra were recorded on a Bruker AMX-400 spectrometer, operating at 400.14 MHz, using a dedicated 5 mm ^1H probe. The 90° pulse width was 8 ms; a low-power 2 s presaturation pulse was applied to suppress the H_2O resonance. The spectra were recorded as 16384 points for the 1D spectra and as 1024×512 data point sets for the 2D spectra. The NOESY data were collected with TPPI [20] in the indirect dimension with a mixing time of 150 ms and a total recording time of about 17 h. ROESY data [21,22] were collected similarly, but using a 200 ms mixing time. The data were processed with Swan-MR software [23] to 1024×1024 real data point sets after application of a 65°-phase-shifted sin^2 function and Fourier transformation for the 2D spectra; baselines were straightened with polynomials as needed. Spectra were referenced to 0 ppm via internal 2,2-dimethylsilapentane-5-sulfonate.

3. Results and discussion

As an initial step, we synthesized the peptides Ac-(Gly-Pro-Thr)$_{10}$-NH_2 and Ac-(Gly-Pro-Thr(Galβ)$_{10}$-NH_2. Although the peptides (Pro-Pro-Gly)$_{10}$ and (Pro-Hyp-Gly)$_{10}$ have been well characterized as forming stable collagen triple-helices, we chose to use acetyl/amide-terminated peptides since the removal of end charges increases the thermal stability of the triple-helix and also provides a more accurate model of an internal triple-helix motif [24]. The glycopeptide was prepared by first synthesizing *O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-N^{α}-9-fluorenylmethoxycarbonyl (FMOC)-Thr-OH and then using this as a building block in the solid phase peptide synthesis [17–19]. Monosaccharides of galactose are found in the *R. pachyptila* and annelid collagens, however the anomeric linkage (α versus β) to the Thr is not known [9,10]. Thus, the facile synthesis and purification of the β-anomer and the lack of any contaminating α-anomer provided a straightforward starting point for our investigations.

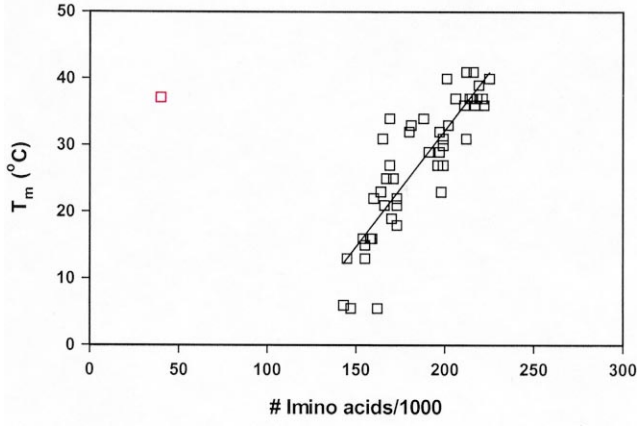

Fig. 1. Correlation between the total imino acid content and melting temperature (T_m) of the triple-helix of various collagens (□). Data taken from [7,8]. *R. pachyptila* cuticle collagen (red □). Data is from [5].

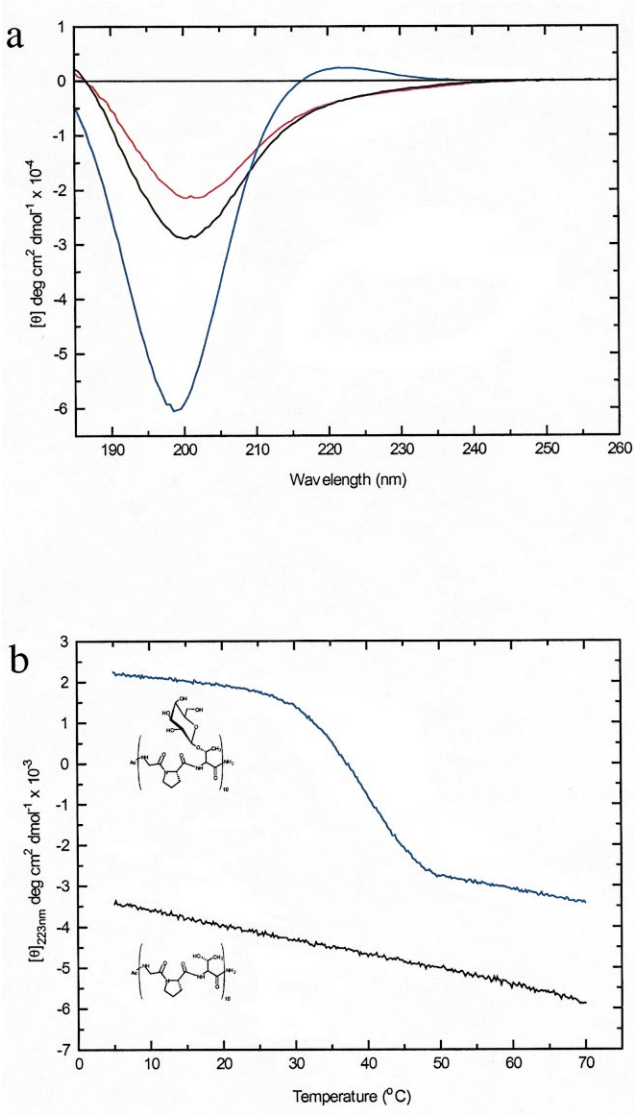

Fig. 2. a: CD spectra of collagen-like peptides. Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH_2 (blue curve) and Ac-(Gly-Pro-Thr)$_{10}$-NH_2 (black curve) at 5°C. The red curve represents Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH_2 at 70°C. Note the similarity in shape between the glycopeptide at 70°C and the peptide. b: Dependence of ellipticity [θ] at 223 nm as a function of temperature (°C). Ac-(Gly-Pro-Thr(β-D-Gal))$_{10}$-NH_2 (blue curve) and Ac-(Gly-Pro-Thr)$_{10}$-NH_2 (black curve).

The far-ultraviolet CD spectra of collagen molecules typically show a negative ellipticity (θ) around 198–200 nm and a maximum ellipticity near 220–225 nm, indicating the presence of a polyproline II helix [25,26]. In water at 5°C, the glycopeptide showed these CD spectral features, as seen in Fig. 2a. In contrast, a weaker negative ellipticity (−2900 deg cm^2 $dmol^{-1}$) around 200 nm and no positive ellipticity in the 220–225 nm range is observed for Ac-(Gly-Pro-Thr)$_{10}$-NH_2, indicating a mostly random-coil conformation. The ellipticity at 223 nm of both peptides was monitored as a function of temperature and this is shown in Fig. 2b. The melting curve of the glycopeptide in water showed the presence of a cooperative transition, with a T_m of 41°C. In contrast, no transition

was observed for the peptide Ac-(Gly-Pro-Thr)$_{10}$-NH_2, again consistent with the inability of this peptide and a similar peptide (H-(Gly-Pro-Thr)$_{10}$-Gly-Pro-Cys-Cys-OH) to form a triple-helix in aqueous buffer [6]. Sedimentation equilibrium measurements gave a weight-average molecular weight of 12 280 ± 200 Da, which is close to that predicted for three associated chains (12 699 Da). Finally, an independent test of a triple-helical conformation is the presence of a unique set of inter-chain nuclear Overhauser effects (NOEs) arising from the close packing of the triple-helix [27,28]. NOESY experiments of the glycopeptide revealed a unique set of NOEs (Fig. 3a), that disappear upon melting of the peptide as observed in the ROESY spectrum (Fig. 3b).

These results show that glycosylation of Thr is required for the formation of a triple-helix with respect to the non-glycosylated peptide and thus must be required to achieve the high thermal stability observed for the *R. pachyptila* cuticle collagen. There are several possible mechanisms, which may account for the stabilization. One possible mechanism is that the addition of galactose may restrict the conformational space available to the polypeptide backbone in a way similar to that of Pro, in which the steric restrictions imposed by the pyrrolidine ring stabilizes the conformation of the polyproline II helix [29]. A restriction of conformational space has been observed in other proteins and peptides that exhibit *O*-linked glycosylation [30–34], and in mucins, the high-degree of *O*-glycosylation results in a rigid, expanded structure [35]. Therefore, glycosylation would limit the number of conformations sampled in the unfolded state and thus *destabilize* this state.

Glycosylation may also influence the stability of the native state, through hydrogen bonds from the sugar to the polypeptide backbone. Crystallographic evidence has suggested that the higher thermal stability of (Pro-Hyp-Gly)$_{10}$ versus (Pro-Pro-Gly)$_{10}$ is due to the ability of the γ-OH group of Hyp to bridge water molecules to the amide carbonyls of the peptide backbone [36,37]. Another explanation for how Hyp stabilizes the triple-helix has recently been put forward [38,39]. These authors have substituted Hyp for 4-*trans*-fluoroproline, which does not form hydrogen bonds and are able to achieve a very stable collagen-like peptide. The conclusion from these experiments is that Hyp, rather than stabilizing hydrogen bonds, stabilizes the *trans* conformation of the imide peptide bond through a stereoelectronic inductive effect; because all peptide bonds in the triple-helix are *trans*, there is a cumulative effect on thermal stability.

In our case, it is also possible that the galactose hydroxyls will mediate interactions with water or that structural waters may be replaced by the sugar hydroxyls. However, recent NH exchange measurements by ^{1}H NMR indicate that both the Gly NH *and* the Thr NH are well protected from exchange with deuterium in the triple-helix. Since the Thr NH is projected to be solvent exposed [40], this suggests that the sugar prevents access of water to the peptide backbone. Further studies on whether glycosylation influences the *cis–trans* isomer ratio of Pro or the conformation of the single chains will help to clarify the mechanism of stabilization by glycosylation.

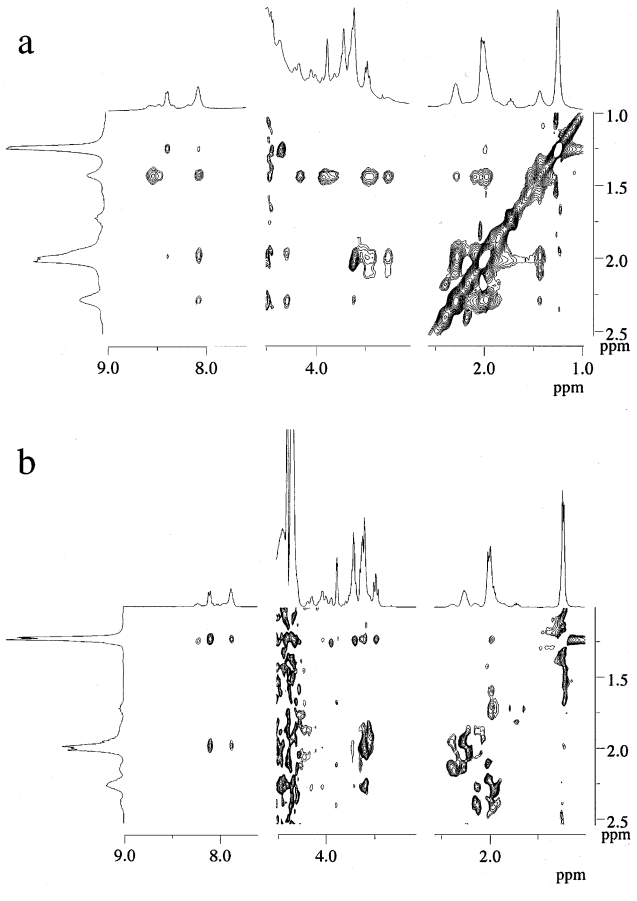

Fig. 3. a: 1D reference and 2D NOESY ^{1}H NMR spectroscopy (18) of Ac-(Gly-Pro-Thr(β-Gal))$_{10}$-NH_2 (1.8 mM) at 25°C and pH 4.2 (±0.1). NOE cross-peaks from the Thr-γCH_3 at 1.47 ppm peak to protons of the Pro aliphatic region, the Gal ring and the amide protons of the Thr and Gly indicate a strong packing of the triple-helix. b: 1D reference and 2D ROESY ^{1}H NMR spectroscopy (18) of Ac-(Gly-Pro-Thr(β-Gal))$_{10}$-NH_2 from (a) at 60°C. Note the loss of cross-peaks from the 1.47 ppm peak.

Acknowledgements: We are grateful to Herbert Waite (University of California, Santa Barbara, CA, USA) and Jan Kihlberg (Umea University, Sweden) for their comments on the manuscript. This work was supported by funds from the Shriners Hospital for Children.

References

[1] Coyne, K.J., Qin, X.X. and Waite, J.H. (1997) Science 277, 1830–1832.
[2] Engel, J. (1997) Science 277, 1785–1786.
[3] Gaill, F., Herbage, D. and Lepescheux, L. (1991) Matrix 11, 197–205.
[4] Gaill, F. (1993) FASEB J. 7, 558–565.
[5] Gaill, F., Mann, K., Wiedemann, H., Engel, J. and Timpl, R. (1995) J. Mol. Biol. 246, 284–294.
[6] Mann, K., Mechling, D.E., Bachinger, H.P., Eckerskorn, C., Gaill, F. and Timpl, R. (1996) J. Mol. Biol. 261, 255–266.
[7] Burjanadze, T.V. (1979) Biopolymers 18, 931–938.
[8] Privalov, P.L. (1982) Adv. Protein Chem. 35, 1–104.
[9] Muir, L. and Lee, Y.C. (1970) J. Biol. Chem. 245, 502–509.
[10] Spiro, R.G. and Bhoyroo, V.D. (1980) J. Biol. Chem. 255, 5347–5354.
[11] Goldstein, A. and Adams, E. (1970) J. Biol. Chem. 245, 5478–5483.
[12] Sharma, Y.D. and Tanzer, M.L. (1984) Anal. Biochem. 141, 205–212.
[13] Inouye, K., Kobayashi, Y., Kyogoku, Y., Kishida, Y., Sakakibara, S. and Prockop, D.J. (1982) Arch. Biochem. Biophys. 219, 198–203.
[14] Josse, J. and Harrington, W.F. (1964) J. Mol. Biol. 9, 269.
[15] Waite, J.H., Tanzer, M.L. and Merkel, J.R. (1980) J. Biol. Chem. 255, 3596–3599.

[16] Ramshaw, J.A., Werkmeister, J.A. and Glattauer, V. (1996) Biotechnol. Genet. Eng. Rev. 13, 335–382.
[17] Kihlberg, J., Elofsson, M. and Salvador, L.A. (1997) Methods Enzymol. 289, 221–245.
[18] Salvador, L.A., Elofson, M. and Kihlberg, J. (1995) Tetrahedron 51, 5643–5656.
[19] Elofson, M., Walse, B. and Kihlberg, J. (1991) Tetrahedron Lett. 32, 7613–7617.
[20] Marion, D. and Wüthrich, K. (1983) Biochem. Biophys. Res. Commun. 113, 967–974.
[21] Bax, A. and Davis, D.G. (1985) J. Magn. Reson. 65, 355–360.
[22] Bothner-By, A.A., Stephens, R.L., Lee, J., Warren, C.D. and Jeanloz, R.W. (1984) J. Am. Chem. Soc. 106, 811–813.
[23] Balacco, G. (1994) J. Chem. Inf. Comput. Sci. 34, 1235–1241.
[24] Feng, Y., Melacini, G., Taulane, J.P. and Goodman, M. (1996) Biopolymers 39, 859–872.
[25] Engel, J. (1968) Das Leder 19, 21–26.
[26] Brown, F.R.d., Di Corato, A., Lorenzi, G.P. and Blout, E.R. (1972) J. Mol. Biol. 63, 85–99.
[27] Li, M.H., Fan, P., Brodsky, B. and Baum, J. (1993) Biochemistry 32, 7377–7387.
[28] Feng, Y., Melacini, G., Taulane, J.P. and Goodman, M. (1996) J. Am. Chem. Soc. 118, 10351–10358.
[29] Harrington, W.F. and Rao, N.V. (1967) in: Conformation of Biopolymers (Ramachandran, G.N., Ed.), pp. 513–531. Academic, New York.
[30] O'Conner, S.E. and Imperiali, B. (1998) Chem. Biol. 5, 427–437.
[31] Live, D.H. et al. (1999) Proc. Natl. Acad. Sci. USA 96, 3489–3493.
[32] Wu, W., Pasternack, L., Huang, D.-H., Koeller, K.M., Lin, C.-C., Oliver Seitz, O. and Wong, C.-H. (1999) J. Am. Chem. Soc. 121, 121–127.
[33] Biondi, L., Filira, F., Gobbo, M., Pavin, E. and Rocchi, R. (1998) J. Pept. Sci. 4, 58–71.
[34] Andreotti, A.H. and Kahne, D. (1993) J. Am. Chem. Soc. 115, 3352–3353.
[35] Jentoft, N. (1990) Trends Biochem. Sci. 15, 291–294.
[36] Bella, J., Eaton, M., Brodsky, B. and Berman, H.M. (1994) Science 266, 75–81.
[37] Bella, J., Brodsky, B. and Berman, H.M. (1995) Structure 3, 893–906.
[38] Holmgren, S.K., Taylor, K.M., Bretscher, L.E. and Raines, R.T. (1998) Nature 392, 666–667.
[39] Holmgren, S.K., Bretscher, L.Y., Taylor, K.M. and Raines, R.T. (1999) Chem. Biol. 6, 63–70.
[40] Kramer, R.Z., Bella, J., Mayville, P., Brodsky, B. and Berman, H.M. (1999) Nat. Struct. Biol. 6, 454–457.