TOWARDS SYNTHETIC COLLAGEN

Ronald T. Raines

Departments of Biochemistry and Chemistry
University of Wisconsin–Madison
Madison, WI 53706

Introduction

Collagen is the most abundant protein in animals. Each polypeptide chain of collagen is composed of repeats of the sequence: Xaa–Yaa–Gly, where Xaa is often a (2*S*)-proline (Pro) residue and Yaa is often a (2*S*,4*R*)-4-hydroxyproline (Hyp) residue. In natural collagen, three such strands are wound into a tight triple helix in which each strand assumes the conformation of a polyproline II-type helix. The hydroxyl group of its prevalent Hyp residues increases markedly the conformational stability of the collagen triple helix. For 25 years, the prevailing paradigm had been that the enhanced stability arises from water molecules that form bridges between the hydroxyl group and a main-chain oxygen. We have overturned this paradigm.

Results and Discussion

How does Hyp in the Yaa position increase triple helix stability? Hydroxyl groups can form hydrogen bonds with water, as observed in the structure of crystalline collagen. In addition, the electronegative oxygen in a hydroxyl group is effective at withdrawing electron density by through-bond and through-space interactions.[1] To distinguish between the contributions of hydrogen bonding and inductive effects to collagen stability, we replaced the hydroxyl groups in Hyp residues with fluorine atoms. We chose fluorine because it is the most electronegative atom and thus elicits a large inductive effect, and because organic fluorine does not form hydrogen bonds.[2-7] This latter attribute of fluorine warrants elaboration.

Anionic fluoride forms strong hydrogen bonds. Indeed, the hydrogen bond in gas-phase $[F\cdots H\text{–}F]^-$ is the strongest known.[8,9] In contrast to anionic fluoride, organic fluorine is a poor hydrogen bond acceptor. X-ray diffraction analyses[10,11] as well as extensive structure database surveys[12-14] have revealed but few crystalline organofluorine compounds that display short C–F···H–X distances, where X = C, N, or O. In addition, a presumably intimate C–F···H–N interaction does not stabilize DNA double helices.[15] The weakness of the C–F···H–X interaction is likely due to the high charge of the fluorine nucleus, which compacts the surrounding electrons.

The inductive effect of a fluoro group is apparent in our data on the structure[16] and properties[17] of proline derivatives. For example, the nitrogen pK_a of the conjugate acid of (2*S*,4*R*)-4-fluoroproline (FlpOH; 9.23) is lower than that of HypOH (9.68) and ProOH (10.8).[17] The nitrogen of AcFlpOMe is more pyramidal than that of AcHypOMe or AcProOMe.[16] This result indicates that the nitrogen of AcFlpOMe has greater sp^3 character and hence higher electron density. The amide I vibrational mode, which results primarily from the C=O stretching vibration, decreases in the order: AcFlpOMe > AcHypOMe > AcProOMe.[17] The value of $\Delta H^{\ddagger}$ for amide bond isomerization is smaller for AcFlpOMe than for AcProOMe.[17] Each of these results is consistent with the traditional picture of amide resonance[18] coupled with an inductive effect that increases the bond order in the amide C=O bond and decreases the bond order in the amide C–N bond. These data suggested to us that inductive effect could contribute to the conformational stability of collagen.

We directly compared the stability conferred to a collagen triple helix by a 4*R* fluoro group and 4*R* hydroxyl group. To do so, we synthesized a collagen-like peptide containing Pro–Flp–Gly units.[19,20] We found that Flp residues allow for triple helix formation. Sedimentation equilibrium experiments with an analytical ultracentrifuge indicated that $(Pro\text{–}Flp\text{–}Gly)_{10}$ chains form a complex of molecular mass (8.0 ± 0.1) kDa. The expected molecular mass of a $(Pro\text{–}Flp\text{–}Gly)_{10}$ trimer ($C_{360}H_{480}N_{90}O_{90}F_{30}$) is 8078 Da. The fluorescence of 1-anilinonaphthalene-8-sulfonate,[21] which has affinity for molten globules,[22] was unchanged in the presence of an excess of $(ProFlpGly)_{10}$ trimer. This result suggested that the tertiary structure of the trimer is packed tightly. At low temperature, the circular dichroism (CD) spectrum of the complex formed by $(Pro\text{–}Flp\text{–}Gly)_{10}$ chains was indistinguishable from that of complexes composed of $(Pro\text{–}Hyp\text{–}Gly)_{10}$ or $(Pro\text{–}Pro\text{–}Gly)_{10}$ chains. All three polymers had a CD spectrum with a positive peak at 225 nm and a stronger negative peak at 200–210 nm, which are defining characteristics of a collagen triple helix.[23] The ellipticity at 225 nm of each triple helix decreased in a sigmoidal manner with increasing temperature, which is characteristic of denaturation of the triple helix. This temperature-dependent change in conformational stability was observed in two solvents: 50 mM acetic acid, which stabilizes triple helices by protonating the C-terminal carboxylates and thereby eliminating unfavorable Coulombic interactions, and phosphate-buffered saline (PBS), which mimics a physiological environment.

Flp residues enhance triple helix stability. In both 50 mM acetic acid and PBS, the values of T_m and $\Delta\Delta G_m$ for the three triple helices differ dramatically, increasing in the order: $(Pro\text{–}Pro\text{–}Gly)_{10} < (Pro\text{–}Hyp\text{–}Gly)_{10} < (Pro\text{–}Flp\text{–}Gly)_{10}$.[19,20] This order is inconsistent with collagen stability arising largely from bridging water molecules, but is consistent with the manifestation of an inductive effect from the electronegative substituent. The stability of the $(Pro\text{–}Flp\text{–}Gly)_{10}$ triple helix far exceeds that of any untemplated collagen mimic of similar size.

Does Flp in the Yaa position endow collagen with hyperstability because of a stereoelectronic effect? In other words, is the mere presence of an electron-withdrawing group on C^{γ} enough, or does the group have to be in the *R* configuration? To answer this question, we synthesized collagen strands containing (2*S*,4*S*)-4-fluoroproline (flp), which is a diastereomer of Flp. We found that $(Pro\text{–}flp\text{–}Gly)_7$, unlike $(Pro\text{–}Flp\text{–}Gly)_7$, does not form a stable triple helix (Table 1).[24] This result provided the first example of a stereoelectronic effect on protein conformational stability. Moreover, the result led us to an explanation for the effect of Flp residues on collagen stability.

We have determined that the remarkable stability of triple helices with $(Pro\text{–}Flp\text{–}Gly)_n$ strands derives from the interplay of several factors, all of which arise from the inductive effect of the fluorine atom.[25] First, the gauche effect prescribes a favorable pyrrolidine ring pucker.[17,24] The gauche effect arises when two vicinal carbons bear electronegative substituents. These electronegative substituents prefer to reside *gauche* (60°) to each other so that there is maximum overlap between the σ orbitals of more electropositive substituents, such as hydrogen, and the σ^* orbitals of the electronegative substituents. As expected from the manifestation of the gauche effect, the C^{γ}-exo ring pucker is predominant in Hyp residues in the Yaa position of collagen-like peptides,[26] as well as in small-molecule structures of AcHypOMe and AcFlpOMe.[16] The gauche effect between fluoro and amide groups is especially strong.[27,28]

Second, the C^{γ}-exo ring pucker preorganizes the main-chain torsion angles of Flp residues. The ϕ angle correlates with ring pucker, with a C^{γ}-exo pucker giving a high (*i.e.*, less negative) value of ϕ, and a C^{γ}-endo pucker giving a low value of ϕ.[26,29] The ψ angle also correlates with ring pucker, as a C^{γ}-exo pucker gives a low value of ψ, and a C^{γ}-endo pucker gives a high value of ψ.[29] The ϕ and ψ angles in crystalline AcFlpOMe,[16] do not differ significantly from those of residues in the Yaa position of triple-helical collagen.[30]

The ψ angle in AcFlpOMe is not only preorganized for triple helix formation, but also establishes a favorable interaction between a non-bonding electron pair (n) of the amide oxygen (O'_{i-1}) and the π antibonding orbital (π^*) of the ester carbon (C'_i). The $O'_{i-1}\cdots C'_i{=}O'_i$ angle in AcFlpOMe is 98°, which is close to the ideal angle for an $n\rightarrow\pi^*$ interaction.[31-33] Moreover, the $O'_{i-1}\cdots C''_i{=}O'_i$ distance in AcFlpOMe is only 2.76 Å, which predicates a meaningful interaction. Indeed, the ester carbonyl stretching vibration is lower by 6 cm^{-1} in AcFlpOMe than in AcflpOMe, presumably because the $n\rightarrow\pi^*$ interaction decreases the C=O bond order.[24] The $n\rightarrow\pi^*$ interaction stabilizes not only the ideal ψ angle for triple-helix formation, but also the requisite trans conformation ($\omega = 180°$) of the Flp peptide bond. In the cis conformation ($\omega = 0°$), C^{α}_{i-1} rather than O'_{i-1} would be proximal to C'_i, and no $n\rightarrow\pi^*$ interaction could occur. These stereoelectronic effects explain why the trans/cis ratio of the amide bond increases as the electronegativity of the substituent in the 4*R* position increases.[17] The reverse trend is true for electronegative 4*S* substituents, which impose a C^{γ}-endo pucker.[24] The association of ω angle with pyrrolidine ring pucker explains the well-known observation that cis prolyl peptide bonds tend to have endo ring puckers in crystalline proteins.[34]

In summary, Flp in the Yaa position stabilizes collagen by a stereoelectronic effect—the gauche effect—that fixes the pyrrolidine ring pucker and thus preorganizes all three main-chain torsion angles: ϕ, ψ, and ω. Density functional theory (DFT) calculations with the (hybrid) B3LYP method are in gratifying agreement with this explanation and all experimental data.[29] These same arguments apply to the prevalent Hyp residues in natural collagen.

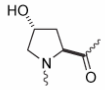
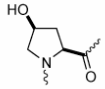
Can triple helix stability be increased by fixing the ring pucker of Pro in the Xaa position? Having established a link between the C^{γ}-exo ring pucker in the Yaa position and triple helix stability, we focused our attention on the Xaa position of the collagen triple helix, in which proline residues have C^{γ}-endo pucker.[30] The gauche effect can be used to preorganize this pucker by using proline residues with an electronegative 4*S* substituent. Yet, replacing Pro in the Xaa position of $(Pro–Pro–Gly)_{10}$ with (2*S*,4*S*)-4-hydroxyproline (hyp) is known to produce strands that fail to form triple helices.[35] We suspected that this result could be due to unfavorable steric interactions that develop upon replacing a hydrogen with a hydroxyl group. This suspicion is consistent with molecular modeling of hyp in the Xaa position.[26] Replacing hydrogen with fluorine, on the other hand, typically results in little steric destabilization.[2-7]

To search for a stereoelectronic effect in the Xaa position on collagen stability, we again used a fluoro group as a probe, synthesizing the peptides $(Flp–Pro–Gly)_7$ and $(flp–Pro–Gly)_7$, where Flp and flp refer to the 4*R* and 4*S* diastereomers, respectively. We found that $(flp–Pro–Gly)_7$ but not $(Flp–Pro–Gly)_7$ forms a stable triple helix (Table 1).[36] Moreover, only $(flp–Pro–Gly)_7$ shows the cooperative transition characteristic of triple helix unfolding upon thermal denaturation. The linear decrease in elipticity exhibited by $(Flp–Pro–Gly)_7$ is characteristic of the unfolding of a single polypeptide chain. Sedimentation equilibrium experiments confirm that $(Flp–Pro–Gly)_7$ but not $(flp–Pro–Gly)_7$ is a monomer at 4 °C, whereas both peptides are monomers at 37 °C.

Apparently, stereoelectronic effects can operate adventitiously (or deleteriously) in the Xaa position of collagen. There, flp is able to preorganize the ϕ and ψ dihedrals as in a triple helix without encountering the steric conflicts that appear to plague hyp in this position.[26] In addition, the 4*S*-substituent in the Xaa position has limited access to solvent, thus making a fluoro group better suited than a hydroxyl group to occupy this position. Altogether, the gain in stability upon replacing hyp with flp in the Xaa position exceeds that of replacing Hyp with Flp in the Yaa position (Table 1).

The conformational stability of a $(flp–Pro–Gly)_7$ triple helix is less than that of a $(Pro–Flp–Gly)_7$ triple helix (Table 1). Two factors could contribute to this lower stability. First, Flp in the Yaa position causes favorable preorganization of all three main-chain dihedral angles (ϕ, ψ, and ω). In the Xaa position, flp increases the probability of ω adopting a cis (ω = 0°) conformation,[24] thus mitigating somewhat the benefit accrued from the preorganization of ϕ and ψ. Second, a C^{γ}-endo pucker is already favored in Pro,[7] and flp only increases that preference. In contrast, Flp has the the more dramatic effect of reversing the preferred ring pucker, thereby alleviating the entropic penalty of triple-helix formation to a greater degree.

Table 1. Correlation of ring pucker with collagen triple helix stability.[24,29,36] In a collagen triple helix, proline residues have ϕ = –73°, ψ = 161° in the Xaa position and ϕ = –58°, ψ = 152° in the Yaa position.

		Residue Ring Pucker		Triple Helix T_m	
		C^{γ}-endo	C^{γ}-exo	$(XaaProGly)_7$	$(ProYaaGly)_7$
Flp		14%	86% ϕ = –55° ψ = 140°	no helix	45 °C
Hyp (natural)				no helix	36 °C
Pro (natural)		66%	34%	6 °C	6 °C
hyp				no helix	no helix
flp		95% ϕ = –76° ψ = 172°	5%	33 °C	no helix

Conclusions. Our findings on collagen are notable on several fronts. First, they overturn a 25-year old paradigm. Second, they are the first to demonstrate that stereoelectronic effects are critical for the conformational stability of a protein. Finally, they could give rise to hyperstable synthetic collagens for a variety of applications in biotechnology and biomedicine.

Acknowledgement. Work on collagen in our laboratory is supported by grant AR44276 (NIH).

References

(1) Stock, L. M. *J. Chem. Educ.* **1972,** *49*, 400–404.
(2) O'Hagan, D.; Rzepa, H. S. *Chem. Commun.* **1997**, 645–652.
(3) Ojima, I.; McCarthy, J. R.; Welch, J. T., Ed., *Biomedical Frontiers of Fluorine Chemistry*; American Chemical Society: Washington, DC, 1996.
(4) Welch, J. T.; Eswarakrishnan, S., *Fluorine in Bioorganic Chemistry*; Wiley: New York, 1991.
(5) Resnati, G. *Tetrahedron* **1993,** *49*, 9385–9445.
(6) Marsh, E. N. G. *Chem. Biol.* **2000,** *7*, R153–R157.
(7) Yoder, N. C.; Kumar, K. *Chem. Soc. Rev.* **2002,** *31*, 335–341.
(8) Harrell, S. A.; McDaniel, D. H. *J. Am. Chem. Soc.* **1964,** *86*, 4497.
(9) Shan, S.; Loh, S.; Herschlag, D. *Science* **1996,** *272*, 97–101.
(10) Murray-Rust, P.; Stallings, W. C.; Monti, C. T.; Preston, R. K.; Glusker, J. P. *J. Am. Chem. Soc.* **1983,** *1983*, 3206–3214.
(11) Shimoni, L.; Carrell, H. L.; Glusker, J. P.; Coombs, M. M. *J. Am. Chem. Soc.* **1994,** *116*, 8162–8168.
(12) Howard, J. A. K.; Hoy, V. J.; O'Hagan, D.; Smith, G. T. *Tetrahedron* **1996,** *52*, 12613–12622.
(13) Shimoni, L.; Glusker, J. P. *Struct. Chem.* **1994,** *5*, 383–397.
(14) Dunitz, J. D.; Taylor, R. *Chem. Eur. J.* **1997,** *3*, 89–98.
(15) Moran, S.; Ren, R. X.; Kool, E. T. *Proc. Natl. Acad. Sci. U.S.A.* **1997,** *94*, 10506–10511.
(16) Panasik, N., Jr.; Eberhardt, E. S.; Edison, A. S.; Powell, D. R.; Raines, R. T. *Int. J. Pept. Protein Res.* **1994,** *44*, 262–269.
(17) Eberhardt, E. S.; Panasik, N., Jr.; Raines, R. T. *J. Am. Chem. Soc.* **1996,** *118*, 12261–12266.
(18) Pauling, L., *The Nature of the Chemical Bond, 3rd ed.*; Cornell University Press: Ithaca, NY, 1960.
(19) Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E.; Raines, R. T. *Nature* **1998,** *392*, 666–667.
(20) Holmgren, S. K.; Bretscher, L. E.; Taylor, K. M.; Raines, R. T. *Chem. Biol.* **1999,** *6*, 63–70.
(21) Brand, L.; Gohlke, J. R. *Annu. Rev. Biochem.* **1972,** *41*, 843–868.
(22) Semisotnov, G. V.; Rodionova, N. A.; Razgulyaev, O. I.; Uversky, V. N.; Gripas', A. F.; Gilmanshin, R. I. *Biopolymers* **1991,** *31*, 119–128.
(23) Piez, K. A.; Sherman, M. R. *Biochemistry* **1970,** *9*, 4129–4133.
(24) Bretscher, L. E.; Jenkins, C. L.; Taylor, K. M.; DeRider, M. L.; Raines, R. T. *J. Am. Chem. Soc.* **2001,** *123*, 777–778.
(25) Jenkins, C. L.; Raines, R. T. *Nat. Prod. Rep.* **2002,** *19*, 49–59.
(26) Vitagliano, L.; Berisio, R.; Mazzarella, L.; Zagari, A. *Biopolymers* **2001,** *58*, 459–464.
(27) O'Hagan, D.; Bilton, C.; Howard, J. A. K.; Knight, L.; Tozer, D. J. *J. Chem. Soc., Perkin Trans. 2* **2000**, 605–607.
(28) Briggs, C. R. S.; O'Hagan, D.; Howard, J. A. K.; Yufit, D. S. *J. Fluorine Chem.* **2003,** *119*, 9–13.
(29) 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.
(30) Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M. *Science* **1994,** *266*, 75–81.
(31) Bürgi, H. B.; Dunitz, J. D.; Shefter, E. *J. Am. Chem. Soc.* **1973,** *95*, 5065–5067.
(32) Bürgi, H. B.; Dunitz, J. D.; Lehn, J. M.; Wipff, G. *Tetrahedron* **1974,** *30*, 1563–1572.
(33) Bürgi, H. B.; Lehn, J. M.; Wipff, G. *J. Am. Chem. Soc.* **1974,** *96*, 1965–1966.
(34) Milner-White, J. E.; Bell, L. H.; Maccallum, P. H. *J. Mol. Biol.* **1992,** *228*, 725–734.
(35) Inouye, K.; Sakakibara, S.; Prockop, D. J. *Biochim. Biophys. Acta* **1976,** *420*, 133–141.
(36) Hodges, J. A.; Raines, R. T. *J. Am. Chem. Soc.* **2003,** *125*, 9262–9263.