

---

## AWARD ADDRESS

# 2005 Emil Thomas Kaiser Award

---

RONALD T. RAINES

Departments of Biochemistry and Chemistry, University of Wisconsin—Madison, Madison, Wisconsin 53706, USA

### Abstract

Collagen is the most abundant protein in animals. The conformational stability of the collagen triple helix is enhanced by the hydroxyl group of its prevalent (2*S*,4*R*)-4-hydroxyproline residues. For 25 years, the prevailing paradigm had been that this enhanced stability is due to hydrogen bonds mediated by bridging water molecules. We tested this hypothesis with synthetic collagen triple helices containing 4-fluoroproline residues. The results have unveiled a wealth of stereoelectronic effects that contribute markedly to the stability of collagen, as well as other proteins. This new understanding is leading to synthetic collagens for a variety of applications in biotechnology and biomedicine.

**Keywords:** stability and mutagenesis; hydrogen bonds; collagen; biotechnology; synthetic proteins

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 (for review, see Jenkins and Raines 2002). In natural collagen, three such strands are wound into a tight triple helix in which each strand assumes the conformation of a polyproline type II (PPII) helix. The strands are held together by a ladder of interstrand GlyN–H···O=CXaa hydrogen bonds, each contributing ~2.0 kcal/mol to the conformational stability of the collagen triple helix (Jenkins et al. 2005).

The hydroxyl group of Hyp is installed in collagen strands by the enzyme prolyl 4-hydroxylase. The prevalence of this post-translational modification is extraordinary, as the abundance of Hyp in humans exceeds that of seven “common” amino acids: Cys, Gln, His, Met, Phe, Trp, and Tyr. Its importance is likewise evident from prolyl 4-hydroxylase (and thus Hyp itself) being essential for animal life (Friedman et al. 2000).

In 1973, Darwin Prockop and his coworkers (Sakakibara et al. 1973) showed that the hydroxyl group of the prevalent

Hyp residues in collagen increases markedly the conformational stability of the triple helix. During the subsequent 25 years, this enhanced stability was believed to arise from water molecules that form bridges between the hydroxyl group and a main-chain oxygen. Doubt about this paradigm led us to explore the chemical forces that stabilize the collagen triple helix. The results of our exploration were reported on August 2, 2005, in the Emil Thomas Kaiser Award lecture at the 19<sup>th</sup> Annual Symposium of The Protein Society in Boston, Massachusetts, and are reiterated below.

### 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 (Bella et al. 1994). In addition, the electronegative oxygen in a hydroxyl group is effective at withdrawing electron density by through-bond and through-space interactions (Stock 1972). To distinguish between the contributions of hydrogen bonding and inductive effects to collagen stability, we used chemical synthesis to replace 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

---

Reprint requests to: Ronald T. Raines, Department of Biochemistry, University of Wisconsin—Madison, 433 Babcock Drive, Madison, WI 53706-1544, USA; e-mail: raines@biochem.wisc.edu; fax: (608) 262-3453.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.062139406>.

organic fluorine forms only weak hydrogen bonds (Welch and Eswarakrishnan 1991; Resnati 1993; Ojima et al. 1996; O'Hagan and Rzepa 1997; Marsh 2000; Yoder and Kumar 2002). This latter attribute of fluorine warrants elaboration.

Anionic fluoride forms strong hydrogen bonds. Indeed, the hydrogen bond in gas phase  $[F\cdots H-F]^-$  is the strongest known (Harrell and McDaniel 1964; Shan et al. 1996). In contrast to anionic fluoride, organic fluorine is a poor hydrogen-bond acceptor. X-ray diffraction analyses (Murray-Rust et al. 1983; Shimoni et al. 1994) as well as extensive surveys of structural databases (Shimoni and Glusker 1994; Howard et al. 1996; Dunitz and Taylor 1997) have revealed but few crystalline organofluorine compounds that display short C-F $\cdots$ H-X distances, where X = C, N, or O. In addition, a presumably intimate C-F $\cdots$ H-N interaction does not stabilize DNA double helices (Moran et al. 1997). The weakness of the C-F $\cdots$ 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 the structure and properties 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) (Eberhardt et al. 1996). The nitrogen of AcFlpOMe is more pyramidal than that of AcHypOMe or AcProOMe (Panasik et al. 1994), indicating 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 (Eberhardt et al. 1996). The value of  $\Delta H^\ddagger$  for amide bond isomerization is smaller for AcFlpOMe than for AcProOMe (Eberhardt et al. 1996). Each of these results is consistent with a traditional picture of amidic resonance (Pauling 1939), 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. This coupling suggested to us that an inductive effect could contribute to the conformational stability of collagen.

We next compared the stability conferred upon 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 (Holmgren et al. 1998, 1999). We found that Flp residues allow for triple helix formation. Sedimentation equilibrium experiments with an analytical ultracentrifuge indicated that (Pro-Flp-Gly)<sub>10</sub> chains form a complex of molecular mass  $(8.0 \pm 0.1)$  kDa. The expected molecular mass of a (Pro-Flp-Gly)<sub>10</sub> trimer (C<sub>360</sub>H<sub>480</sub>N<sub>90</sub>O<sub>90</sub>F<sub>30</sub>) is 8078 Da. The fluorescence of 1-anilinonaphthalene-8-sulfonate (Brand and Gohlke 1972), which has affinity for molten globules (Semisotnov et al. 1991), was unchanged in the presence of an excess of






(Pro-Flp-Gly)<sub>10</sub> 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-Flp-Gly)<sub>10</sub> chains was indistinguishable from that of complexes composed of (Pro-Hyp-Gly)<sub>10</sub> or (Pro-Pro-Gly)<sub>10</sub> 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 (Piez and Sherman 1970). 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 carboxylate groups 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-Pro-Gly)<sub>10</sub> < (Pro-Hyp-Gly)<sub>10</sub> < (Pro-Flp-Gly)<sub>10</sub> (Holmgren et al. 1998, 1999). This order is inconsistent with collagen stability arising largely from bridging water molecules (Engel and Prockop 1998), but is consistent with the manifestation of an inductive effect from the electronegative substituent. The stability of the (Pro-Flp-Gly)<sub>10</sub> triple helix far exceeds that of any untemplated collagen mimic of similar size.

*Does Flp in the Yaa position increase triple helix stability because of a stereoelectronic effect?*

In other words, is the mere presence of an electron-withdrawing group on C<sup>γ</sup> 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-flp-Gly)<sub>7</sub>, unlike (Pro-Flp-Gly)<sub>7</sub>, does not form a stable triple helix (Fig. 1) (Bretscher et al. 2001). This result provided the first example of a stereoelectronic effect on the conformational stability of a protein. Moreover, the result led us to an explanation for the effect of Flp residues on collagen stability.

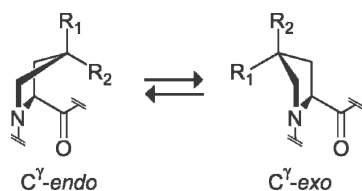
We have determined that the remarkable stability of triple helices with (Pro-Flp-Gly)<sub>*n*</sub> strands derives from the interplay of several factors, all of which arise from the inductive effect of the fluorine atom (Jenkins and Raines 2002): First, the *gauche* effect prescribes a favorable pyrrolidine ring pucker (Eberhardt et al. 1996; Bretscher et al. 2001). 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

	Residue Ring Pucker		Triple Helix $T_m$	
	$C^\gamma$ -endo	$C^\gamma$ -exo	(XaaProGly) <sub>7</sub>	(ProYaaGly) <sub>7</sub>
<b>Flp</b> 	14%	86% $\phi = -55^\circ$ $\psi = 140^\circ$	no helix	45°C
<b>Hyp (natural)</b> 			no helix	36°C
<b>Pro (natural)</b> 	66%	34%	6°C	6°C
<b>hyp</b> 			no helix	no helix
<b>flp</b> 		95% $\phi = -76^\circ$ $\psi = 172^\circ$	33°C	no helix

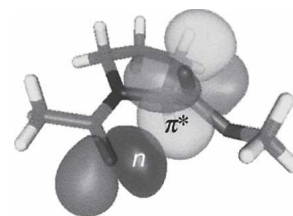
**Figure 1.** Correlation of ring pucker with collagen triple helix stability (Inouye et al. 1976; Bretscher et al. 2001; DeRider et al. 2002). In a crystalline triple helix, proline residues have  $\phi = -73^\circ$ ,  $\psi = 164^\circ$  in the Xaa position, and  $\phi = -60^\circ$ ,  $\psi = 150^\circ$  in the Yaa position (Bella et al. 1994).

$\sigma$  orbitals of more electropositive substituents, such as hydrogen, and the  $\sigma^*$  orbitals of the electronegative substituents. As expected from the manifestation of the *gauche* effect, the  $C^\gamma$ -exo ring pucker is predominant in Hyp residues in collagen-like peptides (Vitagliano et al. 2001; Schumacher et al. 2005), as well as in small-molecule structures of AcHypOME and AcFlpOME (Fig. 2) (Panasik et al. 1994). The *gauche* effect between fluoro and amide groups is especially strong (O'Hagan et al. 2000; Briggs et al. 2003).

Second, the  $C^\gamma$ -exo ring pucker preorganizes the main-chain torsion angles of Flp residues. The  $\phi$  angle correlates with ring pucker, with a  $C^\gamma$ -exo pucker giving a high (that is, less negative) value of  $\phi$ , and a  $C^\gamma$ -endo pucker giving a low value of  $\phi$  (Vitagliano et al. 2001; DeRider et al. 2002). The  $\psi$  angle also correlates with ring pucker, as a  $C^\gamma$ -exo pucker gives a low value of  $\psi$ , and a  $C^\gamma$ -endo pucker gives a high value of  $\psi$  (DeRider et al. 2002). The  $\phi$  and  $\psi$  angles in crystalline AcFlpOME



**Figure 2.** Ring conformations of 4-substituted proline residues. The  $C^\gamma$ -endo conformation is favored strongly when  $R_1 = H$  and  $R_2 = F$  (as in flp). The  $C^\gamma$ -exo conformation is favored strongly when  $R_1 = OH$  (Hyp) or  $F$  (Flp) and  $R_2 = H$ . The  $C^\gamma$ -exo: $C^\gamma$ -endo ratio is  $\sim 1:2$  when  $R_1 = R_2 = H$  (Pro) (DeRider et al. 2002).

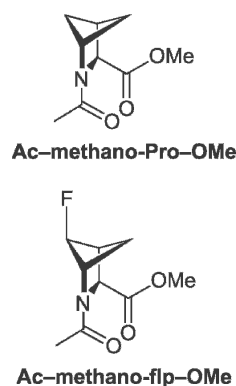


**Figure 3.** Natural bond orbitals depicting the  $n \rightarrow \pi^*$  interaction between  $O'_{i-1}$  and  $C'_i$  in the *trans* peptide bond isomer of AcFlpOME having  $C^\gamma$ -exo ring pucker (DeRider et al. 2002; Hinderaker and Raines 2003).

(Panasik et al. 1994) do not differ significantly from those of residues in the Yaa position of triple-helical collagen (Bella et al. 1994).

The  $\psi$  angle in AcFlpOME is not only preorganized for triple helix formation but also establishes a geometry conducive for a favorable interaction between a nonbonding electron pair ( $n$ ) of the amide oxygen ( $O'_{i-1}$ ) and the  $\pi$  anti-bonding orbital ( $\pi^*$ ) of the ester carbon ( $C'_i$ ). The  $O'_{i-1} \cdots C'_i = O'_i$  angle in AcFlpOME is  $98^\circ$ , which is close to the ideal angle for an  $n \rightarrow \pi^*$  interaction (Fig. 3) (Bürgi et al. 1973, 1974a, b), which is not to be confused with an  $n \rightarrow \pi^*$  electronic transition. 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  $6 \text{ cm}^{-1}$  in AcFlpOME than in AcflpOME, presumably because the  $n \rightarrow \pi^*$  interaction decreases the  $C=O$  bond order (Bretscher et al. 2001). The  $n \rightarrow \pi^*$  interaction stabilizes not only the ideal  $\psi$  angle for triple-helix formation, but also the requisite *trans* peptide bond isomer ( $\omega = 180^\circ$ ) of the Flp peptide bond. In the *cis* isomer ( $\omega = 0^\circ$ ),  $C^\alpha_{i-1}$  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 (Eberhardt et al. 1996). The inverse trend holds for electronegative 4*S* substituents, which impose a  $C^\gamma$ -endo pucker (Bretscher et al. 2001).

Another factor could contribute to the correlation of the configuration of the fluoro group on  $C^\gamma$  and the *trans*:*cis* ratio—the orientation of the C–F dipole. To dissect the relative contributions of ring pucker and the C–F dipole orientation to the *trans*:*cis* ratio, we synthesized methanoproline derivatives that, in essence, incorporate both pyrrolidine ring puckers into a single framework (Fig. 4). Our fluoro-substituted methanoproline is an analog of (2*S*,4*S*)-4-fluoroproline (flp), respectively, except that the fluoro group is fixed in an antithetical conformation—*anti* rather than *gauche* to the pyrrolidine nitrogen about the  $C^{\delta 2}$ – $C^{\gamma 1}$  bond. We found that the presence of the fluoro group has an insignificant effect on the *trans*:*cis* ratio (Jenkins et al. 2004). These data again support the



**Figure 4.** Structures of Ac-methano-Pro-OMe and Ac-methano-flp-OMe, which have amide bonds with indistinguishable *trans:cis* ratios (Jenkins et al. 2004).

idea that an electronegative substituent on the flexible pyrrolidine ring of proline affects the *trans:cis* ratio by altering the pucker of the ring. Finally, it is noteworthy that the association of the *trans:cis* ratio with pyrrolidine ring pucker explains the well-known observation that *cis* prolyl peptide bonds tend to have *endo* ring puckers in crystalline proteins (Milner-White et al. 1992).

In textbooks, the well-known preference of the peptide bond for the *trans* conformation is attributed to steric effects. This picture is incomplete, as a proline residue with an N-formyl group ( $H_{i-1}-C'_{i-1}=O_{i-1}$ ), in which  $H_{i-1}$  presents less steric hindrance than does  $O_{i-1}$ , likewise prefers a *trans* conformation (Fig. 5) (Hinderaker and Raines 2003). Hence, the preference of the peptide bond for the *trans* conformation cannot be explained by steric effects alone. Rather, an  $n \rightarrow \pi^*$  interaction, which is only operative in the *trans* conformation, also contributes significantly to this preference.

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:  $\phi$ ,  $\psi$ , and  $\omega$ . Density functional theory (DFT) calculations with the (hybrid) B3LYP method are in gratifying agreement with this explanation and all experimental data (DeRider et al. 2002). These same arguments apply to the prevalent Hyp residues in natural collagen.

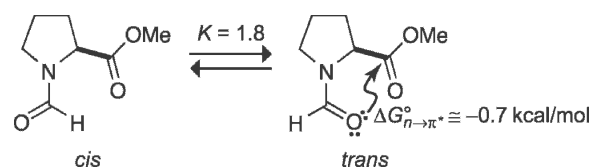
#### Can stereoelectronic effects in the Xaa position increase triple helix stability?

Having established a link between the  $C^\gamma$ -*exo* ring pucker in the Yaa position and triple helix stability, we next focused our attention on the Xaa position of the collagen triple helix, in which proline residues have  $C^\gamma$ -*endo* pucker (Bella et al. 1994). 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)<sub>10</sub> with (2*S*,4*S*)-4-hydroxyproline (hyp) was known to produce strands that fail to form triple helices (Inouye et al. 1976). 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 (Vitagliano et al. 2001). Replacing hydrogen with fluorine, on the other hand, typically results in little steric destabilization (Welch and Eswarakrishnan 1991; Resnati 1993; Ojima et al. 1996; O'Hagan and Rzepa 1997; Marsh 2000; Yoder and Kumar 2002).

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)<sub>7</sub> and (flp-Pro-Gly)<sub>7</sub>, where Flp and flp refer to the 4*R* and 4*S* diastereomers, respectively. We found that (flp-Pro-Gly)<sub>7</sub>, but not (Flp-Pro-Gly)<sub>7</sub>, forms a stable triple helix (Fig. 1) (Hodges and Raines 2003). Apparently, stereoelectronic effects can indeed operate adventitiously (or deleteriously) in the Xaa position of collagen. There, flp is able to preorganize the  $\phi$  and  $\psi$  dihedrals as in a triple helix without encountering the steric conflicts that appear to plague hyp in this position (Vitagliano et al. 2001). Altogether, the conformational stability gained by replacing hyp with flp in the Xaa position exceeds that gained by replacing Hyp with Flp in the Yaa position (Fig. 1).

The conformational stability of a (flp-Pro-Gly)<sub>7</sub> triple helix is, however, less than that of a (Pro-Flp-Gly)<sub>7</sub> triple helix (Fig. 1). This finding is again explicable by a consideration of preorganization in the two strands. First, Flp in the Yaa position causes favorable preorganization of all three main-chain dihedral angles ( $\phi$ ,  $\psi$ , and  $\omega$ ). In the Xaa position, flp increases the probability of the peptide bond adopting a *cis* ( $\omega = 0^\circ$ ) conformation (Bretscher et al. 2001), thus mitigating somewhat the benefit accrued from the preorganization of  $\phi$  and  $\psi$ . Second, a  $C^\gamma$ -*endo* pucker is already favored in Pro (DeRider et al. 2002), and flp only increases that preference. In contrast, Flp has the more dramatic effect of reversing the preferred ring pucker of Pro, thereby alleviating the entropic penalty of triple-helix formation to a greater degree.



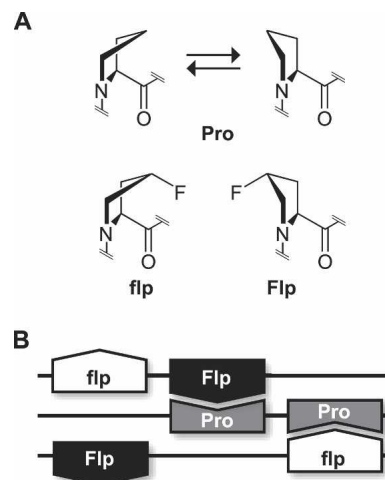
**Figure 5.** *Trans:cis* ratio of FmProOMe. The free energy of the  $n \rightarrow \pi^*$  interaction was estimated from experimental and theoretical data (DeRider et al. 2002; Hinderaker and Raines 2003).

Interestingly, installing both flp in the Xaa position and Flp in the Yaa position does not allow for the formation of a stable triple helix, presumably because of an unfavorable steric interaction between fluoro groups on adjacent strands (Hodges and Raines 2005). Density functional theory calculations indicate that (2*S*,3*S*)-3-fluoroproline (3*S*-flp), like flp, should preorganize the main chain properly for triple-helix formation but without a steric conflict. Synthetic strands containing 3*S*-flp in the Xaa position and Flp in the Yaa position do form a triple helix. This helix is, however, less stable than one with Pro in the Xaa position, presumably because of an unfavorable inductive effect that diminishes the strength of the interstrand GlyN–H...O=C3*S*-flp hydrogen bond (as observed with (2*S*,3*S*)-3-hydroxyproline [Jenkins et al. 2003]). Thus, other forces can counter the benefits derived from the proper preorganization.

Although (Pro–Pro–Gly)<sub>7</sub> and (flp–Flp–Gly)<sub>7</sub> do not form stable homotrimeric helices, mixtures of these two peptides form stable heterotrimeric helices containing one (Pro–Pro–Gly)<sub>7</sub> strand and two (flp–Flp–Gly)<sub>7</sub> strands (Hodges and Raines 2005). This stoichiometry can be understood by considering the cross sections of the two possible heterotrimeric helices and their intrinsic stabilizing and destabilizing interactions. This unexpected finding portends the development of a “code” for the self-assembly of determinate triple helices from two or three strands. This code would not be based on hydrogen-bonding patterns, as in the Watson–Crick paradigm for the DNA double helix. Rather, the code would rely on a combination of stereoelectronic effects (Fig. 6A) and their steric consequences (Fig. 6B).

#### *Do stereoelectronic effects contribute to the stability of other proteins?*

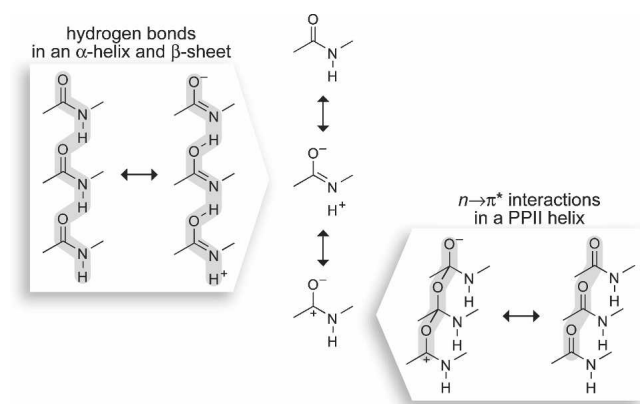
The PPII helix adopted by collagen strands is a prevalent conformation in both folded and unfolded proteins, and is known to play important roles in a wide variety of biological processes. Polyproline itself can also form a type I (PPI) helix, which has a disparate conformation. Our CD spectral analyses of (Pro)<sub>10</sub>, (Hyp)<sub>10</sub>, (Flp)<sub>10</sub>, and (flp)<sub>10</sub> show that 4*R* electron-withdrawing substituents stabilize a PPII helix relative to a PPI helix, even in a solvent that favors the PPI conformation, such as *n*-propanol (Hornig and Raines 2006). The stereochemistry at C<sup>γ</sup> dictates the relative stability of PPI and PPII helices, as (flp)<sub>10</sub> forms a mixture of PPI and PPII helices in water and a PPI helix in *n*-propanol. The conformational preferences of (Pro)<sub>10</sub> are intermediate between those of (Hyp)<sub>10</sub>/(Flp)<sub>10</sub> and (flp)<sub>10</sub>. (Interestingly, PPI helices of [flp]<sub>10</sub> exhibit cold denaturation in *n*-propanol with a value of *T*<sub>s</sub> near 70°C.) Together, these data show that stereoelectronic effects can have a substantial impact on



**Figure 6.** Basis of a code for triple helix formation. (A) Prevalent ring conformations of Pro, flp, and Flp (DeRider et al. 2002). (B) Depiction of the favorable (Flp...Pro) and (Pro...flp) and unfavorable (flp...Flp) steric interactions within cross-sections of a triple helix (Hodges and Raines 2005).

polyproline conformation and provide a rational means to stabilize a PPI or PPII helix.

The stability conferred on a PPII helix by  $n \rightarrow \pi^*$  interactions and that conferred on an  $\alpha$ -helix or a  $\beta$ -sheet by hydrogen bonds are explicable by the invocation of a single concept—resonance. The principle of resonance was first articulated by Linus Pauling, who wrote that a “molecule could be described as fluctuating rapidly between the two electronic formulas, and achieving stability greater than that of either formula through the ‘resonance energy’ of this fluctuation” (Pauling 1932). In modern quantum chemistry, the “rapid fluctuation” model has been displaced by a view of resonance as an electron donor–electron acceptor correction to the dominant Lewis structure (Weinhold and Landis 2005). In a consideration of protein structure, the relevant resonance is that of the peptide bond. The oxygen of a peptide bond is anionic; the hydrogen and carbon are cationic (Milner-White 1997). This charge distribution is apparent in the two dipolar resonance forms that contribute most to the structure of a peptide bond (Fig. 7). Favorable interactions between the anionic oxygen of one peptide bond and the cationic hydrogen or carbon of another peptide bond ordains the formation of an  $\alpha$ -helix/ $\beta$ -sheet or a PPII helix, respectively. These favorable interactions stabilize one of the dipolar resonance forms, leading to the known cooperative propagation of structure. The propensity of a particular residue to assume an  $\alpha$ -helix, a  $\beta$ -sheet, or a PPII helix arises from the interplay of factors that enhance or diminish the impact of the hydrogen bonds or  $n \rightarrow \pi^*$  interactions. For example, side chains can either shield the main chain from

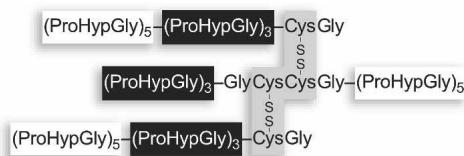


**Figure 7.** Interactions that stabilize the common elements of secondary structure as embedded within the major resonance forms of the peptide bond.

distracting hydrogen bonds with water or, conversely, themselves form distracting hydrogen bonds with the main chain.

#### *Can chemical synthesis produce a triple helix of biosynthetic length?*

By employing stereoelectronic effects, we produced collagen triple helices that are stronger than any found in nature. We then sought to use molecular self-assembly to produce collagen triple helices that are longer than any found in nature. There is a practical imperative for this quest. Collagen isolated from natural sources has long served as the basis for many biomaterials. Still, natural collagen is difficult to modify and can engender pathogenic and immunological side effects. Triple helices derived from synthetic peptides are much shorter (<10 nm) than natural collagen (~300 nm), limiting their utility. To overcome these limitations, we synthesized short collagen fragments in which the three strands are held in a staggered array by disulfide bonds (Fig. 8). Data from CD spectroscopy, dynamic light scattering, analytical ultracentrifugation, atomic force microscopy, and transmission electron microscopy indicate that these “sticky-ended” fragments self-assemble via intermolecular triple helix formation (Kotch and Raines 2006). The resulting fibrils resemble natural collagen, and some are longer (>400 nm) than any known collagen.



**Figure 8.** Structure of a synthetic collagen fragment that self-assembles into fibrils of 1-nm width and nearly 1- $\mu$ m length (Kotch and Raines 2006).

## Conclusions

Our work on collagen can be summarized by recapitulating its major conclusions. First, we have disproved an entrenched paradigm about the contribution of bridging water molecules to collagen stability. Second, we have demonstrated that stereoelectronic effects play a critical role in the conformational stability of collagen as well as PPII helices and the *trans* isomer of peptide bonds. Finally, we have used chemical synthesis to produce both hyperstable (strong) and hyperextended (long) collagens. These synthetic fibrils are now poised for use in biotechnology and biomedicine.

## Acknowledgments

The 2005 Emil Thomas Kaiser Award is dedicated to the outstanding students who have worked with me on collagen. Their names rightly dominate the list of references. I am also grateful to my mentors, Chris Walsh, Jeremy Knowles, and Bill Rutter, for their guidance and support throughout my career. I had the privilege of meeting Tom Kaiser in his laboratory at The Rockefeller University and elsewhere, and I thank the SynPep Corporation for sponsoring this award in his name. Work on collagen in my laboratory has been supported by grants from the NIAMS (NIH AR44276) and the Arthritis Foundation.

## References

- Bella, J., Eaton, M., Brodsky, B., and Berman, H.M. 1994. Crystal and molecular structure of a collagen-like peptide at 1.9 Å resolution. *Science* **266**: 75–81.
- Brand, L. and Gohlke, J.R. 1972. Fluorescence probes for structure. *Annu. Rev. Biochem.* **41**: 843–868.
- Bretscher, L.E., Jenkins, C.L., Taylor, K.M., DeRider, M.L., and Raines, R.T. 2001. Conformational stability of collagen relies on a stereoelectronic effect. *J. Am. Chem. Soc.* **123**: 777–778.
- Briggs, C.R.S., O'Hagan, D., Howard, J.A.K., and Yufit, D.S. 2003. The C–F bond as a tool in the conformational control of amides. *J. Fluor. Chem.* **119**: 9–13.
- Bürgi, H.B., Dunitz, J.D., and Shefter, E. 1973. Geometrical reaction coordinates. II. Nucleophilic addition to a carbonyl group. *J. Am. Chem. Soc.* **95**: 5065–5067.
- Bürgi, H.B., Dunitz, J.D., Lehn, J.M., and Wipff, G. 1974a. Stereochemistry of reaction paths at carbonyl centres. *Tetrahedron* **30**: 1563–1572.
- Bürgi, H.B., Lehn, J.M., and Wipff, G. 1974b. An *ab initio* study of nucleophilic addition to a carbonyl group. *J. Am. Chem. Soc.* **96**: 1965–1966.
- DeRider, M.L., Wilkens, S.J., Waddell, M.J., Bretscher, L.E., Weinhold, F., Raines, R.T., and Markley, J.L. 2002. Collagen stability: Insights from NMR spectroscopic and hybrid density functional computational investigations of the effect of electronegative substituents on prolyl ring conformations. *J. Am. Chem. Soc.* **124**: 2497–2505.
- Dunitz, J.D. and Taylor, R. 1997. Organic fluorine hardly ever accepts hydrogen bonds. *Chem. Eur. J.* **3**: 89–98.
- Eberhardt, E.S., Panasiak Jr., N., and Raines, R.T. 1996. Inductive effects on the energetics of prolyl peptide bond isomerization: Implications for collagen folding and stability. *J. Am. Chem. Soc.* **118**: 12261–12266.
- Engel, J. and Prockop, D.J. 1998. Does bound water contribute to the stability of collagen? *Matrix Biol.* **17**: 679–680.
- Friedman, L., Higgin, J.J., Moulder, G., Barstead, R., Raines, R.T., and Kimble, J. 2000. Prolyl 4-hydroxylase is required for viability and morphogenesis in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci.* **97**: 4736–4741.
- Harrell, S.A. and McDaniel, D.H. 1964. Strong hydrogen bonds. II. The hydrogen difluoride ion. *J. Am. Chem. Soc.* **86**: 4497.

- Hinderaker, M.P. and Raines, R.T. 2003. An electronic effect on protein structure. *Protein Sci.* **12**: 1188–1194.
- Hodges, J.A. and Raines, R.T. 2003. Stereoelectronic effects on collagen stability: The dichotomy of 4-fluoroproline diastereomers. *J. Am. Chem. Soc.* **125**: 9262–9263.
- . 2005. Stereoelectronic and steric effects in the collagen triple helix: Toward a code for strand association. *J. Am. Chem. Soc.* **127**: 15923–15932.
- Holmgren, S.K., Taylor, K.M., Bretscher, L.E., and Raines, R.T. 1998. Code for collagen's stability deciphered. *Nature* **392**: 666–667.
- Holmgren, S.K., Bretscher, L.E., Taylor, K.M., and Raines, R.T. 1999. A hyperstable collagen mimic. *Chem. Biol.* **6**: 63–70.
- Hornig, J.-C. and Raines, R.T. 2006. Stereoelectronic effects on polyproline conformation. *Protein Sci.* **15**: 74–83.
- Howard, J.A.K., Hoy, V.J., O'Hagan, D., and Smith, G.T. 1996. How good is fluorine as a hydrogen bond acceptor? *Tetrahedron* **52**: 12613–12622.
- Inouye, K., Sakakibara, S., and Prockop, D.J. 1976. Effects of the stereo-configuration of the hydroxyl group in 4-hydroxyproline on the triple-helical structures formed by homogenous peptides resembling collagen. *Biochim. Biophys. Acta* **420**: 133–141.
- Jenkins, C.L. and Raines, R.T. 2002. Insights on the conformational stability of collagen. *Nat. Prod. Rep.* **19**: 49–59.
- Jenkins, C.L., Bretscher, L.E., Guzei, I.A., and Raines, R.T. 2003. Effect of 3-hydroxyproline residues on collagen stability. *J. Am. Chem. Soc.* **125**: 6422–6427.
- Jenkins, C.L., Lin, G., Duo, J., Rapolu, D., Guzei, I.A., Raines, R.T., and Krow, G.R. 2004. Substituted 2-azabicyclo[2.1.1]hexanes as constrained proline analogues: Implications for collagen stability. *J. Org. Chem.* **69**: 8565–8573.
- Jenkins, C.L., Vasbinder, M.M., Miller, S.J., and Raines, R.T. 2005. Peptide bond isosteres: Ester or (E)-alkene in the backbone of the collagen triple helix. *Org. Lett.* **7**: 2619–2622.
- Kotch, F.W. and Raines, R.T. 2006. Self-assembly of synthetic collagen triple helices. *Proc. Natl. Acad. Sci.* **103**: 3028–3033.
- Marsh, E.N.G. 2000. Towards the nonstick egg: Designing fluorous proteins. *Chem. Biol.* **7**: R153–R157.
- Milner-White, E.J. 1997. The partial charge on the nitrogen atom in peptide bonds. *Protein Sci.* **6**: 2477–2482.
- Milner-White, J.E., Bell, L.H., and Maccallum, P.H. 1992. Pyrrolidine ring puckering in *cis* and *trans*-proline residues in proteins and polypeptides. *J. Mol. Biol.* **228**: 725–734.
- Moran, S., Ren, R.X., and Kool, E.T. 1997. A thymidine triphosphate shape analog lacking Watson–Crick pairing ability is replicated with high sequence selectivity. *Proc. Natl. Acad. Sci.* **94**: 10506–10511.
- Murray-Rust, P., Stallings, W.C., Monti, C.T., Preston, R.K., and Glusker, J.P. 1983. Intermolecular interactions of the C–F bond: The crystallographic environment of fluorinated carboxylic acids and related structures. *J. Am. Chem. Soc.* **105**: 3206–3214.
- O'Hagan, D. and Rzepa, H.S. 1997. Some influences of fluorine in bioorganic chemistry. *Chem. Commun.* **7**: 645–652.
- O'Hagan, D., Bilton, C., Howard, J.A.K., Knight, L., and Tozer, D.J. 2000. The preferred conformation of N- $\beta$ -fluoroethylamides. Observation of the fluorine amide gauche effect. *J. Chem. Soc. Perkin Trans. 1* **2**: 605–607.
- Ojima, I., McCarthy, J.R., and Welch, J.T. 1996. *Biomedical frontiers of fluorine chemistry*. American Chemical Society, Washington, DC.
- Panasik Jr., N., Eberhardt, E.S., Edison, A.S., Powell, D.R., and Raines, R.T. 1994. Inductive effects on the structure of proline residues. *Int. J. Pept. Protein Res.* **44**: 262–269.
- Pauling, L. 1932. The nature of the chemical bond III. The transition from one extreme bond type to another. *J. Am. Chem. Soc.* **54**: 988–1003.
- . 1939. *The nature of the chemical bond*, pp. 186–193. Cornell University Press, Ithaca, NY.
- Piez, K.A. and Sherman, M.R. 1970. Characterization of the product formed by renaturation of  $\alpha$ 1-CB2, a small peptide from collagen. *Biochemistry* **9**: 4129–4133.
- Resnati, G. 1993. Synthesis of chiral and bioactive fluoroorganic compounds. *Tetrahedron* **49**: 9385–9445.
- Sakakibara, S., Inouye, K., Shudo, K., Kishida, Y., Kobayashi, Y., and Prockop, D.J. 1973. Synthesis of (Pro–Hyp–Gly)<sub>n</sub> of defined molecular weights. Evidence for the stabilization of collagen triple helix by hydroxyproline. *Biochim. Biophys. Acta* **303**: 198–202.
- Schumacher, M., Mizuno, K., and Bächinger, H.P. 2005. The crystal structure of the collagen-like polypeptide (glycyl-4(R)-hydroxyprolyl-4(R)-hydroxyprolyl)<sub>9</sub> at 1.55 Å resolution shows up-puckering of the proline ring in the Xaa position. *J. Biol. Chem.* **280**: 20397–20403.
- Semisotnov, G.V., Rodionova, N.A., Razgulyaev, O.I., Uversky, V.N., Gripas', A.F., and Gilmanshin, R.I. 1991. Study of the "molten globule" intermediate in protein folding by a hydrophobic fluorescent probe. *Biopolymers* **31**: 119–128.
- Shan, S., Loh, S., and Herschlag, D. 1996. The energetics of hydrogen bonds in model systems: Implications for enzymatic catalysis. *Science* **272**: 97–101.
- Shimoni, L. and Glusker, J.P. 1994. The geometry of intermolecular interactions in some crystalline fluorine-containing organic-compounds. *Struct. Chem.* **5**: 383–397.
- Shimoni, L., Carrell, H.L., Glusker, J.P., and Coombs, M.M. 1994. Intermolecular effects in crystals of 11-(trifluoromethyl)-15,16-dihydrocyclopenta(a)phenanthren-17-one. *J. Am. Chem. Soc.* **116**: 8162–8168.
- Stock, L.M. 1972. The origin of the inductive effect. *J. Chem. Educ.* **49**: 400–404.
- Vitagliano, L., Berisio, R., Mazzarella, L., and Zagari, A. 2001. Structural bases of collagen stabilization induced by proline hydroxylation. *Biopolymers* **58**: 459–464.
- Weinhold, F. and Landis, C.R. 2005. *Valency and bonding: A natural bond orbital donor–acceptor perspective*. Cambridge University Press, Cambridge, UK.
- Welch, J.T. and Eswarakrishnan, S. 1991. *Fluorine in bioorganic chemistry*. Wiley, New York.
- Yoder, N.C. and Kumar, K. 2002. Fluorinated amino acids in protein design and engineering. *Chem. Soc. Rev.* **31**: 335–341.