Conformational Stability of Collagen Relies on a Stereoelectronic Effect

Lynn E. Bretscher,[†,‡] Cara L. Jenkins,[§] Kimberly M. Taylor,[||] Michele L. DeRider,[§] and Ronald T. Raines*[,†,§]

Departments of Biochemistry and Chemistry and Graduate Program in Biophysics, University of Wisconsin−Madison, Madison, Wisconsin 53706

Received August 22, 2000

A polypeptide chain can adopt many conformations. Yet, the sequence of its amino acid residues directs folding to a particular native state.[1] The loss of conformational entropy associated with folding destabilizes the native state. This destabilization is overcome by the hydrophobic effect, hydrogen bonds, other noncovalent interactions, and disulfide bonds.[2] We have identified another force that can contribute to the conformational stability of a protein.

The structure and reactivity of an organic molecule can rely on the stereochemistry of its electron pairs, bonded or nonbonded.[3] Such stereoelectronic effects, which arise from the mixing of an electron pair with the antibonding σ^* of an adjacent polar bond (C−X, where X = N or O), endow nucleic acids and carbohydrates with conformational stability.[4] For example, the multiple gauche effects (X−C−C−X) arising from a 2′ oxygen distinguish RNA·RNA and RNA·DNA duplexes from DNA·DNA duplexes.[5] The anomeric effect (X−C−X) enhances the stability of the α anomer of glycosides.[6] Here, we demonstrate for the first time that a stereoelectronic effect is critical for the conformational stability of a protein.

Collagen is the most abundant protein in animals.[7] For example, collagen comprises one-third of the protein in humans and three-fourths of the weight of human skin. The polypeptide chains of collagen are composed of approximately 300 repeats of the sequence XaaYaaGly, where Xaa is often an L-proline (Pro) residue and Yaa is often a 4(*R*)-hydroxy-L-proline (Hyp) residue. These chains are wound in tight triple helices, which are organized into fibrils of great tensile strength.

Pro and Hyp comprise nearly one-fourth of the residues in common types of collagen.[8] This prevalence of tertiary amides has dichotomous consequences for conformational stability. Pro and Hyp residues are constrained by their pyrrolidine rings, and this rigidity stabilizes triple-helical collagen.[9] Yet, the trans and cis conformations of the peptide bonds to Pro and Hyp residues are of nearly equal free energy, which destabilizes collagen because all peptide bonds in triple-helical collagen are in the trans conformation (Figure 1).[10] The 4(*R*)-hydroxyl group of the prevalent Hyp residues increases dramatically the conformational stability of collagen.[11] We had shown previously that this increase arises from inductive effects.[12]

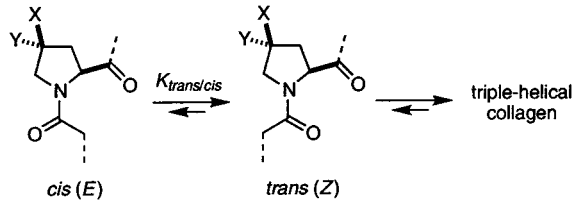

Figure 1. Relationship between cis−trans prolyl peptide bond isomerization and the formation of a collagen triple helix, which contains only trans peptide bonds. Pro: X = H, Y = H. Hyp: X = H, Y = OH. hyp: X = OH, Y = H. Flp: X = H, Y = F. flp: X = F, Y = H.

Can a stereoelectronic effect influence $K_{trans/cis}$? To answer this question, we synthesized residue mimics of the form AcYaaOMe, where Yaa is Pro, Hyp, 4(*S*)-hydroxy-L-proline (hyp), 4(*R*)-fluoro-L-proline (Flp), or 4(*S*)-fluoro-L-proline (flp).[13] We chose Flp and flp because fluorine is small and electronegative and forms only weak hydrogen bonds when bound to carbon.[14,15] We chose a methyl ester rather than an amide to prevent γ-turn formation, as had been observed in AcProNHMe.[16] We find that the electronegativity and stereochemistry of the 4-substituent in the Yaa mimics has a significant effect on $K_{trans/cis}$ (Table 1). Compared to a flp residue, a Pro residue is twice as likely and a Flp residue is 3 times as likely to have a trans peptide bond.

Does the value of $K_{trans/cis}$ have an impact on collagen stability? To answer this question, we synthesized $(ProYaaGly)_7$ strands containing Flp or flp residues in the Yaa position.[19] These strands are diastereomeric, differing only in the stereochemistry at C_γ of the Yaa residues. We found that a $(ProFlpGly)_7$ triple helix has a T_m of 45 °C.[20] In contrast, a $(ProflpGly)_7$ triple helix has a T_m of <2 °C (Table 1). A $(ProHypGly)_7$ triple helix has an intermediate T_m of 36 °C. Thus, both the electronegativity and

* To whom correspondence should be addressed. Telephone: (608) 262-8588. Fax: (608) 262-3453. E-mail: raines@biochem.wisc.edu.
† Department of Biochemistry.
‡ Current address: Department of Biochemistry, Medical College of Wisconsin, Milwaukee, WI 53226.
§ Department of Chemistry.
|| Graduate Program in Biophysics.

(1) Anfinsen, C. B. *Science* **1973**, *181*, 223−230.
(2) Dill, K. A. *Biochemistry* **1990**, *29*, 7133−7155.
(3) Deslongchamps, P. *Stereoelectronic Effects in Organic Chemistry*; Pergamon Press: New York, 1983.
(4) Gorenstein, D. G. *Chem. Rev.* **1987**, *87*, 1047−1077.
(5) Plavec, J.; Thibaudeau, C.; Chattopadhyaya, J. *J. Am. Chem. Soc.* **1994**, *116*, 6558−6560.
(6) Thatcher, G. R. J., Ed. *The Anomeric Effect and Associated Stereoelectronic Effects*; American Chemical Society: Washington, DC, 1993.
(7) *Collagen 1−4*; Nimni, M. E., Ed.; CRC Press: Boca Raton, FL, 1988.
(8) Hyp alone comprises approximately 10% of the residues in collagen. Fietzek, P. P.; Kuhn, K. *Mol. Cell. Biochem.* **1975**, *8*, 141−157.
(9) Shah, N. K.; Ramshaw, J. A. M.; Kirkpatrick, A.; Shah, C.; Brodsky, B. *Biochemistry* **1996**, *35*, 10262−10268.
(10) (a) Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M. *Science* **1994**, *266*, 75−81. (b) Nagarajan, V.; Kamitori, S.; Okuyama, K. *J. Biochem.* **1999**, *125*, 310−318.
(11) Sakakibara, S.; Inouye, K.; Shudo, K.; Kishida, Y.; Kobayashi, Y.; Prockop, D. J. *Biochim. Biophys. Acta* **1973**, *303*, 198−202.
(12) (a) Eberhardt, E. S.; Panasik, N., Jr.; Raines, R. T. *J. Am. Chem. Soc.* **1996**, *118*, 12261−12266. (b) Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E.; Raines, R. T. *Nature* **1998**, *392*, 666−667. (c) Holmgren, S. K.; Bretscher, L. E.; Taylor, K. M.; Raines, R. T. *Chem. Biol.* **1999**, *6*, 63−70.
(13) AcProOMe, AcHypOMe, and AcFlpOMe were synthesized as described (ref 23). AchypOMe was synthesized by a route analogous to that used to prepare AcHypOMe. AcflpOMe was synthesized as follows: AcHypOMe (2.11 g, 11.3 mmol) was dissolved in CH_2Cl_2 (40 mL) and pyridine (5.2 mL), and the solution was cooled to −78 °C. Morpholino-sulfurtrifluoride (morph-DAST, 2.2 mL, 18 mmol) was added via syringe, and the reaction mixture was stirred for 24 h, warming to room temperature. The reaction was quenched at 0 °C with methanol (5 mL), and the resulting mixture was washed with 50-mL portions of 2 M H_2SO_4, water, and saturated $NaHCO_3$. The first two washes were extracted twice with CH_2Cl_2, and the combined organic layers were dried over $MgSO_4$, filtered, and concentrated to give a yellow oil. The oil was subjected to column chromatography on silica gel, eluting with 3:1 EtOAc:hexanes, then EtOAc, then 10% MeOH in EtOAc, and yielding 721 mg (34%) of 95% pure AcflpOMe.
(14) For comprehensive analyses of organic fluorine as a hydrogen bond acceptor, see: (a) Howard, J. A. K.; Hoy, V. J.; O'Hagan, D.; Smith, G. T. *Tetrahedron* **1996**, *52*, 12613−12622. (b) Dunitz, J. D.; Taylor, R. *Chem. Eur. J.* **1997**, *3*, 89−98.
(15) For reviews of bioorganic fluorine chemistry, see: (a) Welch, J. T.; Eswarakrishnan, S. *Fluorine in Bioorganic Chemistry*; Wiley: New York, 1991. (b) Resnati, G. *Tetrahedron* **2000**, *49*, 9385−9445. (c) Ojima, I., McCarthy, J. R., Welch, J. T., Eds.; *Biomedical Frontiers of Fluorine Chemistry*; American Chemical Society: Washington, DC, 1996. (d) O'Hagan, D.; Rzepa, H. S. *Chem. Commun.* **1997**, 645−652. (e) March, E. N. G. *Chem. Biol.* **2000**, *7*, R153−R157.
(16) (a) Matsuzaki, T.; Iitaka, Y. *Acta Crystallogr., Sect. B* **1971**, *27*, 507−516. (b) Liang, G.-B., Rito, C. J.; Gellman, S. H. *Biopolymers* **1992**, *32*, 293−301.

10.1021/ja005542v CCC: $20.00
Published on Web 01/03/2001

Table 1. Effect of 4(*R*)- and 4(*S*)-Substituents of Proline Residues on the Conformational Stability of Triple-Helical Collagen and on Related Parameters

	AcYaaOMe		triple-helical $(ProYaaGly)_7$
Yaa	$K_{trans/cis}$[a]	ν_{ester} (cm^{-1})	T_m (°C)[c]
Flp	6.7	1748	45
Hyp	6.1	1746	36
Pro	4.6	1743	6−7[d]
hyp	2.4	1725	<5[e]
flp	2.5	1754	<2

[a] Values of $K_{trans/cis}$ (± 5%) were measured in D_2O at 25 °C by integration of ^1H NMR spectra. [b] Values of ν_{ester} (±1 cm^{-1}) were obtained in chloroform at 25 °C by FTIR spectroscopy with a Mattson Infinity instrument. [c] Values of T_m (±1 °C) were measured in 50 mM acetic acid by CD spectroscopy with an Aviv 62A DS instrument. [d] From ref 17. [e] Reported for $(ProhypGly)_{10}$ in ref 18.

the stereochemistry of the C_γ substituent have a significant effect on T_m. Moreover, larger T_m values of $(ProYaaGly)_7$ triple helices correlate with larger $K_{trans/cis}$ values of AcYaaOMe mimics. We conclude that a C_γ substituent can enhance conformational stability by favoring the trans isomer, thereby preorganizing individual strands to resemble more closely the strands in a triple helix (Figure 1).

How can a stereoelectronic effect influence $K_{trans/cis}$? The fluorine−amide gauche effect is especially strong.[21] In proline, the gauche effect from a 4(*R*)- or 4(*S*)-fluoro substituent imposes a C_γ-exo or C_γ-endo pyrrolidine ring pucker, respectively. Neither pucker creates any apparent unfavorable electronic or steric interactions in a collagen triple helix.[10] Instead, $K_{trans/cis}$ appears to be related to the main-chain dihedral angle, ψ.

The value of ψ is 141° in crystalline AcFlpOMe (Figure 2) and nearly 150° in the Hyp residues of triple-helical collagen.[10] Values of ψ near 150° provide a nearly optimal geometry for a favorable interaction of O_0 of a trans prolyl peptide bond with $C_1{=}O_1$. The $O_0{\cdots}C_1$ distance in crystalline AcFlpOMe is 2.76 Å, which puts the two atoms in van der Waals contact. The $O_0{\cdots}C_1{=}O_1$ angle in crystalline AcFlpOMe is 98°, which is remarkably similar to the Bürgi−Dunitz trajectory—the preferred angle of 109° for the approach of a nucleophile.[22] Thus, we propose that a favorable $O_0{\cdots}C_1$ interaction stabilizes both the trans prolyl peptide bond isomer and a ψ that is appropriate for triple helix formation.

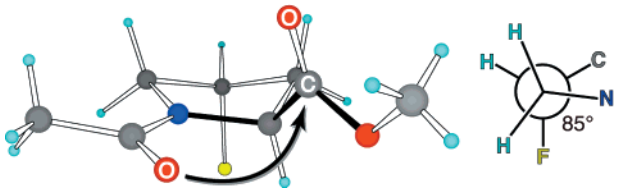

Figure 2. (Left) Structure of crystalline AcFlpOMe.[23] $r_{O_0\cdots C_1} = 2.76$ Å, $\angle O_0{\cdots}C_1{=}O_1 = 98°$. The ψ dihedral angle (solid black bonds) is 141°. (Right) Newman projection depicting the gauche effect. The $N_1{-}C_1^\delta{-}C_1^\gamma{-}F_1^{\delta 1}$ dihedral angle is 85°.

In a flp residue, fluorine and $C_1{=}O_1$ reside on the same side of the pyrrolidine ring. In the C_γ-endo conformation that arises from the gauche effect, unfavorable $F_1^{\delta 1}{\cdots}O_1$ interactions could disfavor ψ being near 150° and the $O_0{\cdots}C_1{=}O_1$ angle being near 109°. A weaker $O_0{\cdots}C_1$ interaction would decrease $K_{trans/cis}$, as is observed (Table 1). An improper ψ and low $K_{trans/cis}$ would destablize a collagen triple helix, as is also observed (Table 1).

Infrared spectroscopy provides support for a stronger $O_0{\cdots}C_1$ interaction in AcFlpOMe than in AcflpOMe. An ester carbonyl stretching vibration (ν_{ester}) decreases with decreasing C=O bond order. The value of ν_{ester} is 6 cm^{-1} lower in AcFlpOMe than in its diastereomer, AcflpOMe (Table 1). This decrease could arise from a stronger $O_0{\cdots}C_1$ interaction and consequent donation of electron density from O_0 to O_1 via C_1.[24]

The value of ν_{ester} is 21 cm^{-1} lower in AchypOMe than in its diastereomer, AcHypOMe (Table 1). This decrease is consistent with the formation of an $O_1^{\delta 1}{-}H{\cdots}O_1$ hydrogen bond within the hyp residue. Such a hydrogen bond would perturb the ψ angle of hyp residues in a $(ProhypGly)_{10}$ triple helix. Thus, the inability of $(ProhypGly)_{10}$ to form a stable triple helix[18] cannot be ascribed solely to a stereoelectronic effect.

Like nucleic acids and carbohydrates, proteins contain many polar bonds. The relatively high abundance of carbon in proteins does, however, insulate most polar bonds from one another. Still, natural side chains other than that of Hyp can manifest stereoelectronic effects. For example, the common amino acids L-serine and L-threonine, as well as 3(*S*)-hydroxy-L-proline and 5(*R*)-hydroxy-L-lysine (which are found in collagen), all contain X−C−C−X systems that are subject to a gauche effect. The subtle manipulation of these stereoelectronic effects with natural or non-natural amino acids could enable the creation of more stable proteins, as we have demonstrated here (Table 1) and elsewhere[12] for collagen.

Acknowledgment. This work was supported by Grant AR44276 (NIH) and the Arthritis Foundation. C.L.J. was supported by NIH Chemistry−Biology Interface Training Grant GM08506. K.M.T. was supported by an HHMI predoctoral fellowship. ^1H NMR spectra were obtained at the National Magnetic Resonance Facility at Madison, WI, which was supported by Grant RR02301 (NIH). CD and IR spectra were obtained at the University of Wisconsin−Madison Biophysical Instrumentation Facility, which was established by Grants BIR-9512577 (NSF) and RR13790 (NIH).

JA005542V

(17) Shaw, B. R.; Schurr, J. M. *Biopolymers* **1975**, *14*, 1951−1985.
(18) Inouye, K.; Sakakibara, S.; Prockop, D. J. *Biochim. Biophys. Acta* **1976**, *420*, 133−141.
(19) $(ProFlpGly)_7$, $(ProflpGly)_7$, and $(ProHypGly)_7$ were synthesized by segment condensation on a solid phase. FmocProYaaGlyOH units were assembled by standard solution-phase procedures from Flp, flp, and commercial reagents. 2-Chlorotrityl chloride resin (Advanced ChemTech) was modified with the trimer units using standard conditions. Six additional units were coupled to the modified resin using an Applied Biosystems 432A peptide synthesizer. Peptides were cleaved from the resin and purified by HPLC on a Vydac C-18 reversed-phase column to give the desired peptide. All three 21-mers were judged by HPLC and MALDI-TOF mass spectrometry to be >90% pure.
(20) Triple helices were formed by incubating 21-mers (0.23 mM) for 24 h at 4 °C in 50 mM acetic acid. The concentration of solutions of 21-mers was determined by measuring *A* at 214 nm ($\epsilon = 6.4 \times 10^4\ M^{-1}\ cm^{-1}$). Values of T_m, which is the temperature at the midpoint of the thermal transition, were determined as described (ref 12c).
(21) O'Hagan, D.; Bilton, C.; Howard, J. A. K.; Knight, L.; Tozer, D. J. *J. Chem. Soc., Perkin Trans. 2* **2000**, 605−607.
(22) (a) Bürgi, H. B.; Dunitz, J. D.; Shefter, E. *J. Am. Chem. Soc.* **1973**, *95*, 5065−5067. (b) Bürgi, H. B.; Dunitz, J. D.; Wipff, G. *J. Am. Chem. Soc.* **1974**, *96*, 1956−1957. (c) Bürgi, H. B.; Dunitz, J. D.; Lehn, J. M.; Wipff, G. *Tetrahedron* **1974**, *30*, 1563−1572.
(23) Panasik, N., Jr.; Eberhardt, E. S.; Edison, A. S.; Powell, D. R.; Raines, R. T. *Int. J. Pept. Protein Res.* **1994**, *44*, 262−269.
(24) The $\Delta\nu_{ester}$ of 6 cm^{-1} is replicated exactly by frequency calculations performed on the fully geometry-optimized structures with B3LYP/6-31+G*. These calculations do, indeed, reveal the donation of electron density from O_0 to O_1 via C_1. DeRider, M. L.; Wilkens, S. J.; Bretscher, L. E.; Weinhold, F. A.; Raines, R. T.; Markley, J. L. Unpublished results.