

**CORRECTION: The “4(S)” stereochemistry herein is actually “4(R)”.**

## Inductive effects on the structure of proline residues

NICHOLAS PANASIK, JR.<sup>1</sup>, ERIC S. EBERHARDT<sup>1</sup>, ARTHUR S. EDISON<sup>2</sup>, DOUGLAS R. POWELL<sup>3</sup> and RONALD T. RAINES<sup>1,2</sup>

<sup>1</sup>Department of Biochemistry, <sup>2</sup>Graduate Program in Biophysics, <sup>3</sup>Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin, USA

Received 9 December 1993, accepted for publication 11 February 1994

4(S)-Hydroxyproline (Hyp) residues constitute about 10% of most forms of collagen, the most abundant protein in vertebrates. X-Ray diffraction analysis was used to ascertain how the structure of proline residues is affected by the inductive effect elicited by the hydroxyl group of Hyp residues. *N*-Acetylproline methylester (**1**), *N*-acetyl-4(S)-hydroxyproline methylester (**2**) and *N*-acetyl-4(S)-fluoroproline methylester (**3**) were synthesized, and their crystalline structures were determined at high resolution. The amide bond of crystalline **1** was in the *cis* conformation, which is the minor isomer in solution, and the pyrrolidine ring of **1** had C $\gamma$ -endo pucker. In crystalline **2** and **3** the amide bonds were in the *trans* conformation, and the pyrrolidine rings had C $\gamma$ -exo pucker. The lengths of the bonds between sp<sup>3</sup>-hybridized carbon atoms in the pyrrolidine ring were significantly shorter in **2** and **3** than in **1**, as was predicted by *ab initio* molecular orbital calculations at the RHF/3–21G level of theory. No significant change in bond length was observed in the other bonds of **1**, **2** or **3**. The pyramidalization of the nitrogen atom increased dramatically in the order: **1** < **2** < **3**. Together, these results indicate that electron-withdrawing substituents in the 4-position of proline residues can have a significant influence on the structure of these residues. In particular, the change in pyramidalization suggests that such substituents increase the sp<sup>3</sup>-character of the prolyl nitrogen atom and could thereby alter the rate of prolyl peptide bond isomerization. © Munksgaard 1994.

**Key words:** collagen; fluoroproline; hydroxyproline; molecular orbital calculation; peptide bond isomer; prolyl peptide bond; pyramidalization; ring pucker; X-ray diffraction

Collagen is the most abundant protein in vertebrates (1–3). Each polypeptide chain of collagen is composed of approximately 300 repeats of the sequence: X–Y–Gly, where X is often a proline (Pro) residue and Y is often a 4(S)-hydroxyproline (Hyp) residue. In connective tissue, collagen chains form triple-helical arrays in which all of the peptide bonds are in the *trans* (*Z*) conformation. These arrays are then organized into fibrils of great tensile strength (4).

The role of the hydroxyl group of Hyp residues in collagen structure and function is uncertain. The hydroxylation of proline residues occurs after collagen biosynthesis, but before the chains begin to form a triple helix. The effect of prolyl hydroxylation on the rate of triple helix formation is not known. Hydroxylation is known, however, to increase the thermal stability of triple-helical collagen (5, 6), but the molecular basis for this increased stability is not clear. Models based on the structure (7, 8) of triple-helical collagen and conformational energy calculations suggest that no hydrogen bonds can be formed between the hydroxyl group of Hyp residues and any backbone groups of the same

triple helix (9). Several other models have been proposed, however, in which a water molecule forms a bridge between the hydroxyl group and a mainchain carbonyl group (3).

The electronegative oxygen atom of a hydroxyl group is effective at withdrawing electron density by through-bond and through-space interactions. This inductive effect is apparent in the acid dissociation constants of appropriate derivatives. For example, the pK<sub>a</sub> of the secondary ammonium group of Pro is 10.64, while that of Hyp is 9.66; and the pK<sub>a</sub> of the carboxylic acid group of Pro is 1.95, while that of Hyp is 1.82 (10). These differences in acidity are consistent with the manifestation of an inductive effect that draws electron density toward the hydroxyl group of Hyp.

Any inductive effect from the hydroxyl group of Hyp residues should be apparent in the structure of these residues. To determine the effect of electron withdrawal on the structures of proline residues, *N*-acetylproline methylester (**1**), *N*-acetyl-4(S)-hydroxyproline methylester (**2**) and *N*-acetyl-4(S)-fluoroproline methylester (**3**) were synthesized. The methylester was used to mini-

mize intramolecular hydrogen bonding, as has been observed in *N*-acetylproline (11) and *N*-acetylproline *N'*-methylamide (12, 13). Compound 3 was studied because the fluoro group is a small substituent that extends the range of inductive effects. Inductive effects that arise either through bonds or through space have been described by many empirical parameters (14). For example, the acid dissociation constants of 4-substituted quinuclidines are related by the parameter  $\sigma_1$ , where  $\sigma_1 = (\text{p}K^{\text{H}_a} - \text{p}K^{\text{X}_a})/5.15$  (ref. 15). Hydrogen, hydroxyl and fluoro groups have  $\sigma_1$  values of 0.00, 0.27 and 0.51, respectively.

The structures of 1–3 were determined by X-ray diffraction analysis, and were also studied by *ab initio* molecular orbital calculations at the RHF/3–21G level of theory. The results of this work provide a high-resolution picture of the inductive effect on the structure of analogous proline residues.

## EXPERIMENTAL PROCEDURES

*N*-Acetylproline methylester (1), 4(*S*)-hydroxyproline methylester hydrochloride, and 4(*R*)-hydroxyproline methylester hydrochloride were from Bachem Biosciences (Philadelphia, PA), and were used without further purification. All other reagents and all solvents were from Aldrich Chemical (Milwaukee, WI), and were used without further purification. Flash chromatography (16) was carried out on silica gel 60 (230–400 mm) from Fisher (Pittsburgh, PA). Analytical chromatography was carried out on pre-coated plates of silica 60 F-254 (0.25 mm) from EM Science (Gibbstown, NJ).  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{19}\text{F}$  NMR spectra were recorded on a Bruker AM400, AM500 or AM600 spectrometer. Chemical shifts for  $^1\text{H}$  and  $^{13}\text{C}$  spectra are reported relative to the isotopic impurity peak  $\text{CDCl}_3$  (7.26 ppm for  $^1\text{H}$  and 77.0 ppm for  $^{13}\text{C}$ ). Chemical shifts for the  $^{19}\text{F}$  spectrum are reported relative to trifluoroacetic acid (0 ppm). Samples contained approximately 0.1 M solute in  $\text{CDCl}_3$ .

### *N*-Acetyl-4(*S*)-acetoxypoline methylester

4(*S*)-Hydroxyproline methylester hydrochloride (4.0 g, 22 mmol) was dissolved in acetic anhydride (40 mL), and the resulting solution was stirred for 12 h at room temperature. The mixture was then concentrated under reduced pressure, and the resulting residue was subjected to flash chromatography (ethyl acetate–hexane, 1:1 v/v) to yield product as a white solid (3.6 g, 75%;  $R_f$  0.40, ethyl acetate).  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , ppm) 5.16 (br s, 1H,  $\text{C}_1^{\text{H}}$ ; minor isomer at 5.10), 4.33 (t,  $J = 2$  Hz, 1H,  $\text{C}_2^{\text{H}}$ ), 3.74 (dd,  $J = 12$  Hz,  $J = 4.5$  Hz, 1H,  $\text{C}_1^{\beta 2}\text{H}$ ), 3.55 (s, 3H,  $\text{C}_2\text{H}_3$ ; minor isomer at 3.61), 3.43 (br d,  $J = 11$  Hz, 1H,  $\text{C}_1^{\beta 2}\text{H}'$ ; minor isomer at 3.42), 2.21 (m, 1H,  $\text{C}_1^{\beta}\text{H}$ ; minor isomer at 2.34), 2.03 (m, 1H,  $\text{C}_1^{\beta}\text{H}'$ ), 1.90 [s, 3H,  $\text{C}_0^{\text{H}}_3$ ; minor isomer at 1.80], 1.89 [s, 3H,  $\text{O}^{\delta 1}\text{C}(\text{O})\text{CH}_3$ ; minor isomer at 1.87];  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , ppm) 171.85 ( $\text{C}_0$ ), 169.91 [ $\text{O}^{\delta 1}\text{C}(\text{O})$ ], 168.90

( $\text{C}_1$ ), 72.40 ( $\text{C}_1^{\gamma}$ ; minor isomer at 70.80), 56.82 ( $\text{C}_1^{\alpha}$ ; minor isomer at 58.02), 52.75 ( $\text{C}_1^{\beta 2}$ ; minor isomer at 51.15), 51.92 ( $\text{C}_2$ ; minor isomer at 52.42), 34.54 ( $\text{C}_2^{\beta}$ ; minor isomer at 36.54), 21.85 ( $\text{C}_0^{\alpha}$ ), 20.45 [ $\text{O}^{\delta 1}\text{C}(\text{O})\text{CH}_3$ ].

### *N*-Acetyl-4(*S*)-hydroxyproline methylester (2)

Anhydrous potassium carbonate (17) (0.151 g, 1.09 mmol) was added to a solution of *N*-acetyl-4(*S*)-acetoxypoline methylester (2.5 g, 11 mmol) in methanol (12 mL), and the resulting slurry was stirred for 30 min at room temperature. The mixture was then filtered, the filtrate was concentrated under reduced pressure, and the resulting residue was subjected to flash chromatography (ethyl acetate–hexane, 1:2 v/v) to yield product as a white solid (1.50 g, 74%;  $R_f$  0.17, ethyl acetate).  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ) 4.62 (d,  $J = 2.5$  Hz, 1H,  $\text{C}_1^{\text{H}}$ ; minor isomer at 4.49), 4.33 (t,  $J = 8$  Hz, 1H,  $\text{C}_1^{\alpha}\text{H}$ ; minor isomer at 4.40), 3.59 (d,  $J = 9$  Hz, 1H,  $\text{C}_1^{\beta 2}\text{H}$ ), 3.55 (s, 3H,  $\text{C}_2\text{H}_3$ ; minor isomer at 3.57), 3.35 (d,  $J = 5.5$  Hz, 1H,  $\text{C}_1^{\beta 2}\text{H}'$ ), 2.91 (br s, 1H,  $\text{O}^{\delta 1}\text{H}$ ), 2.14–2.09 (m,  $\text{C}_1^{\beta}\text{H}$ , 1H; minor isomer at 2.28–2.23), 1.92 (s, 3H,  $\text{C}_0^{\text{H}}_3$ ; minor isomer at 1.78), 1.88–1.83 (m, 1H,  $\text{C}_1^{\beta}\text{H}'$ ; minor isomer at 2.04–2.00);  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , ppm) 172.55 ( $\text{C}_0$ ; minor isomer at 172.38), 170.46 ( $\text{C}_1$ ; minor isomer at 169.90), 69.28 ( $\text{C}_1^{\gamma}$ ; minor isomer at 67.70), 57.24 ( $\text{C}_1^{\alpha}$ ; minor isomer at 518.40), 55.54 ( $\text{C}_1^{\beta}$ ; minor isomer at 54.01), 51.90 ( $\text{C}_2$ ; minor isomer at 52.36), 37.45 ( $\text{C}_2^{\beta}$ ; minor isomer at 39.27), 21.80 ( $\text{C}_0^{\alpha}$ ; minor isomer at 21.20). MS (EI,  $m/e$ ): 187.2594 (calcd. for  $\text{C}_8\text{H}_{13}\text{NO}_4$  187.0845).

### *N*-Acetyl-4(*R*)-acetoxypoline methylester

4(*R*)-Hydroxyproline methylester hydrochloride (0.25 g, 1.4 mmol) was dissolved in acetic anhydride (4 mL), and the resulting solution was stirred for 24 h at room temperature. The mixture was concentrated under reduced pressure, and the resulting residue was subjected to flash chromatography (ethyl acetate) to yield product as a white solid (0.295 g, 93%;  $R_f$  0.75, chloroform–methanol–acetic acid, 18:2:1 v/v/v).  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ) 5.20 (br s, 1H,  $\text{C}_1^{\text{H}}$ ), 4.61 (d,  $J = 4.5$  Hz, 1H,  $\text{C}_1^{\alpha}\text{H}$ ; minor isomer at 4.41), 3.78 (dd,  $J = 6$  Hz, 1H,  $\text{C}_1^{\beta 2}\text{H}$ ), 3.65 (s, 3H,  $\text{C}_2\text{H}_3$ ; minor isomer at 3.70), 3.602 (br d,  $J = 5.5$  Hz, 1H,  $\text{C}_1^{\beta 2}\text{H}'$ ), 2.42–2.32 (m, 1H,  $\text{C}_1^{\beta}\text{H}$ ), 2.22 (br d,  $J = 7$  Hz, 1H,  $\text{C}_1^{\beta}\text{H}'$ ; minor isomer at 2.36), 2.01 (s, 3H,  $\text{C}_0^{\text{H}}_3$ ; minor isomer at 1.96), 1.94 [s, 3H,  $\text{O}^{\delta 1}\text{C}(\text{O})\text{CH}_3$ ; minor isomer at 1.89];  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , ppm) 171.60 ( $\text{C}_0$ ), 170.36 [ $\text{O}^{\delta 1}\text{C}(\text{O})$ ]; minor isomer at 169.98], 169.72 ( $\text{C}_1$ ; minor isomer at 169.88), 72.81 ( $\text{C}_1^{\gamma}$ ; minor isomer at 71.48), 52.35 ( $\text{C}_2$ ), 56.97 ( $\text{C}_1^{\alpha}$ ; minor isomer at 58.87), 53.16 ( $\text{C}_1^{\beta}$ ; minor isomer at 52.68), 34.84 ( $\text{C}_2^{\beta}$ ; minor isomer at 37.08), 22.23 ( $\text{C}_0^{\alpha}$ ; minor isomer at 22.09), 20.90 [ $\text{O}^{\delta 1}\text{C}(\text{O})\text{CH}_3$ ].

### *N*-Acetyl-4(*R*)-hydroxyproline methylester

Anhydrous potassium carbonate (17) (0.015 g, 0.11 mmol) was added to a solution of *N*-acetyl-4(*R*)-

TABLE 1  
Crystallographic parameters for 1–3

Parameter	1	2	3
Empirical formula	C <sub>8</sub> H <sub>13</sub> NO <sub>3</sub>	C <sub>8</sub> H <sub>13</sub> NO <sub>4</sub>	C <sub>8</sub> H <sub>12</sub> NO <sub>3</sub> F
Molecular weight	171.2	187.2	189.2
Crystal system	Orthorhombic	Monoclinic	Orthorhombic
Crystal size (mm)	0.3 × 0.4 × 0.6	0.2 × 0.4 × 0.4	0.1 × 0.3 × 0.5
Space group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Z, molecule/unit cell	4	4	8
<i>a</i> (Å)	7.328(3)	6.5077(9)	6.6060(9)
<i>b</i> (Å)	10.052(3)	12.1992(16)	6.6603(7)
<i>c</i> (Å)	11.643(3)	12.0478(14)	41.076(3)
$\beta$ (°)		98.881(10)	
<i>V</i> (Å <sup>3</sup> )	857.6(5)	945.0(2)	1807.2(2)
<i>d</i> (calc) (g/cm <sup>3</sup> )	1.326	1.316	1.391
Peaks to determine cell	25	25	26
Reflections collected	2444	1463	2132
Independent reflections	1146	1334	1901
<i>R</i> (int)	3.06%	2.69%	2.26%
Observed reflections, { <i>F</i> > 4.0σ( <i>F</i> )}	1135	1273	1800
<i>R</i> ( <i>F</i> ) (obs. data) (%)	3.08	3.53	3.68
w <i>R</i> ( <i>F</i> ) (obs. data) (%)	4.29	4.67	5.06
<i>s</i>	1.59	1.31	1.06
$\Delta\rho$ (max) (Å <sup>-3</sup> )	0.23	0.17	0.25
$\Delta\rho$ (min) (Å <sup>-3</sup> )	-0.29	-0.15	-0.36

acetoxypoline methylester (0.251 g, 1.10 mmol) in methanol (12 mL), and the resulting slurry was stirred for 10 min at room temperature. The mixture was then filtered, concentrated under reduced pressure, and subjected to flash chromatography (ethyl acetate) to yield product as a white solid (0.145 g, 75%; *R*<sub>f</sub> 0.39, chloroform–methanol–acetic acid, 18:2:1 v/v/v). <sup>1</sup>H NMR (CDCl<sub>3</sub>) 4.48 (d, *J* = 5 Hz, 1H, C<sup>γ</sup>H), 4.41 (m, 1H, C<sup>α</sup>H), 3.77–3.72 (m, 1H, C<sup>δ2</sup>H), 3.68–3.61 (m, 1H, C<sup>δ1</sup>H'), 3.66 (s, 3H, C<sub>2</sub>H<sub>3</sub>), 2.80 (s, 1H, O<sup>δ</sup>H), 2.34–2.25 (m, 2H, C<sup>β</sup>H<sub>2</sub>), 2.06 (s, 3H, C<sup>α</sup>H<sub>3</sub>; minor isomer at 2.00); <sup>13</sup>C NMR (CDCl<sub>3</sub>, ppm) 174.4 (C<sub>0</sub>), 169.6 (C<sub>1</sub>), 71.2 (C<sub>1</sub><sup>γ</sup>; minor isomer at 68.8), 57.2 (C<sub>1</sub><sup>α</sup>; minor isomer at 59.0), 56.2 (C<sub>1</sub><sup>δ</sup>; minor isomer at 55.2), 52.8 (C<sub>2</sub>), 37.0 (C<sub>1</sub><sup>β</sup>; minor isomer at 39.4), 22.0 (C<sub>0</sub><sup>α</sup>).

#### *N*-Acetyl-4(*S*)-fluoropoline methylester (3)

The hydroxyl group in *N*-acetyl-4(*R*)-hydroxypoline methylester was converted into a fluoro group by using the morpholino analogue of diethylaminosulfur trifluoride (DAST) (18). Morpholinisulfur trifluoride (0.564 g, 3.22 mmol) was added dropwise over 10 min and with stirring over 10 min to a solution (5 mL) at -80 °C under nitrogen of dichloromethane containing *N*-acetyl-4(*R*)-hydroxypoline methylester (0.120 g, 0.645 mmol). The reaction was allowed to warm to room temperature, and was then stirred for 48 h. The mixture was concentrated under reduced pressure, quenched with water (2 mL), concentrated again under reduced pressure, and subjected to flash chromatography (ethyl acetate–hexanes, 1:4 v/v) to yield product as a color-

TABLE 2

Atomic coordinates ( $\times 10^5$ ) and equivalent isotropic temperature factors ( $\text{\AA}^2 \times 10^4$ ) for nonhydrogen atoms of 1

Atom	<i>x</i>	<i>y</i>	<i>z</i>	<i>U</i> <sub>eq</sub> <sup>a</sup>
C <sup>δ</sup>	6048(24)	87507(17)	33780(16)	277(6)
C <sub>0</sub>	-3483(22)	80827(16)	23945(15)	221(5)
O <sub>0</sub>	3793(17)	71793(12)	18445(10)	291(4)
N <sub>1</sub>	-20241(18)	85182(13)	21035(12)	202(4)
C <sup>γ</sup>	-30267(22)	79450(18)	11331(15)	246(5)
C <sup>γ</sup>	-48892(23)	86148(18)	11791(16)	282(6)
C <sup>γ</sup>	-45003(24)	99412(19)	17841(15)	272(6)
C <sup>γ</sup>	-30627(21)	95542(17)	26956(14)	221(5)
C <sub>1</sub>	-39464(23)	90069(17)	37783(15)	216(5)
O <sub>1</sub>	-39030(16)	78674(12)	40978(11)	308(4)
O <sub>2</sub>	-48147(15)	99782(11)	43494(10)	251(4)
C <sub>2</sub>	-57066(24)	95785(18)	54060(15)	261(5)

$$^a U_{eq} = (1/3) \sum_i \sum_j U_{ij} a_i a_j$$

less oil (0.060 g, 50%; *R*<sub>f</sub> 0.36, ethyl acetate). <sup>1</sup>H NMR (CDCl<sub>3</sub>) 5.23 (dt, *J* = 26.5 Hz, *J* = 3 Hz, 1H, C<sup>γ</sup>H; minor isomer at 5.17), 4.70 (d, *J* = 5 Hz, 1H, C<sup>α</sup>H; minor isomer at 4.45), 3.87–3.62 (m, 2H, C<sup>δ2</sup>H<sub>2</sub>), 3.67 (s, 3H, C<sub>2</sub>H<sub>3</sub>; minor isomer at 3.71), 2.43 (t, *J* = 15 Hz, broad, 1H, C<sup>β</sup>H; minor isomer at 3.64), 2.32 (m, 1H, C<sup>β</sup>H'; minor isomer at 3.45); 2.05 (s, 3H, C<sub>2</sub>H<sub>3</sub>; minor isomer at 1.98); <sup>13</sup>C NMR (CDCl<sub>3</sub>, ppm) 172.30 (C<sub>0</sub>), 169.20 (C<sub>1</sub>), 91.80 (d, *J*<sub>F</sub> = 180 Hz, C<sup>γ</sup>), 54.18 (C<sub>1</sub><sup>δ</sup>; minor isomer at 53.98), 52.31 (C<sub>2</sub>; minor isomer at 52.72), 52.11 (C<sub>1</sub><sup>α</sup>; minor isomer at 58.11), 35.89 (C<sub>1</sub><sup>β</sup>;

minor isomer at 36.06), 22.15 ( $C_0^2$ ; minor isomer at 22.10);  $^{19}\text{F}$  NMR ( $\text{CDCl}_3$ , ppm)  $\delta$  -114.302 (minor isomer at -114.943). MS (FAB,  $m/e$ ): 190.2535 [calcd. for  $\text{C}_8\text{H}_{13}\text{NO}_3\text{F}$  ( $\text{MH}^+$ ) 190.0879].

#### X-Ray diffraction analysis

Single crystals suitable for X-ray diffraction analysis were obtained by evaporation at ambient temperature of solutions of **1** or **3** in 2-propanol, or of **2** in ethyl acetate. Cell parameter data and intensity data were collected using a Siemens P4 diffractometer with  $\text{CuK}\alpha$  radiation ( $\lambda = 1.54178 \text{ \AA}$ ). The unit-cell parameters and their ESDs were derived from a least-squares treatment based on 25 (**1** and **2**) or 26 (**3**) high-angle reflections in the range  $48 > \theta > 50^\circ$  (**1**),  $45 > \theta > 50^\circ$  (**2**) or  $53 > \theta > 56^\circ$  (**3**). Integrated relative intensities for the independent reflections with  $2\theta < 114^\circ$  were measured in the Wyckoff (**1** and **2**) or the  $2\theta - \theta$  (**3**) scan mode. The intensity of three standard reflections, checked after each 100 (**1**), 50 (**2**) or 50 (**3**) reflections, varied by less than 5 (**1**), 4 (**2**) or 6% (**3**). Data were corrected for Lorentz and polarization effects but not for absorption. The structures were solved by direct methods using the program SHELXTL PLUS (19), and refined anisotropically with the full-matrix least-squares procedure.

TABLE 3

Atomic coordinates ( $\times 10^5$ ) and equivalent isotropic temperature factors ( $\text{\AA}^2 \times 10^4$ ) for nonhydrogen atoms of **2**

Atom	x	y	z	$U_{\text{eq}}^a$
$C_0^2(\text{A})$	4227(8)	7116(5)	11707(3)	68(2)
$C_0(\text{A})$	5135(6)	6330(5)	11076(3)	52(1)
$O_0(\text{A})$	5753(5)	5366(5)	11527(2)	72(1)
$N_1(\text{A})$	5283(5)	6412(4)	9984(3)	44(1)
$C_1^{22}(\text{A})$	4307(6)	7302(4)	9277(3)	50(1)
$C_1^1(\text{A})$	4137(6)	6841(4)	8104(3)	49(1)
$O_1^{21}(\text{A})$	2198(4)	6277(4)	7867(2)	55(1)
$C_1^2(\text{A})$	5978(6)	6072(4)	8173(3)	50(1)
$C_1^3(\text{A})$	6154(5)	5570(4)	9345(3)	41(1)
$C_1(\text{A})$	8416(6)	5325(4)	9809(3)	45(1)
$O_1(\text{A})$	9617(4)	5950(4)	10315(3)	65(1)
$O_2(\text{A})$	8869(4)	4316(4)	9498(2)	55(1)
$C_2(\text{A})$	11005(6)	3992(5)	9772(4)	67(2)
$C_2^3(\text{B})$	1017(7)	8943(5)	3665(3)	72(2)
$C_0(\text{B})$	-855(6)	9297(4)	4171(3)	51(1)
$O_0(\text{B})$	-2512(5)	9534(4)	3578(2)	67(1)
$N_1(\text{B})$	-695(4)	9340(4)	5284(2)	45(1)
$C_1^{22}(\text{B})$	1110(6)	9008(4)	6125(3)	48(1)
$C_1^1(\text{B})$	181(6)	8932(4)	7197(3)	45(1)
$O_1^{21}(\text{B})$	-834(4)	7914(4)	7299(2)	55(1)
$C_1^2(\text{B})$	-1467(6)	9812(5)	7052(3)	54(1)
$C_1^3(\text{B})$	-2440(5)	9715(4)	5811(3)	44(1)
$C_1(\text{B})$	-3323(6)	10788(4)	5358(3)	41(1)
$O_1(\text{B})$	-2329(4)	11538(4)	5071(3)	73(1)
$O_2(\text{B})$	-5354(4)	10814(4)	5338(2)	59(1)
$C_2(\text{B})$	-6402(7)	11836(5)	4997(4)	77(2)

$$^a U_{\text{eq}} = (1/3) \sum_i \sum_j U_{ij} a_i a_j (a_i a_j)$$

TABLE 4

Atomic coordinates ( $\times 10^5$ ) and equivalent isotropic temperature factors ( $\text{\AA}^2 \times 10^4$ ) for nonhydrogen atoms of **3**

Atom	x	y	z	$U_{\text{eq}}^a$
$C_0^2(\text{A})$	21664(47)	53590(46)	25263(6)	245(9)
$C_0(\text{A})$	34356(44)	50186(42)	22283(6)	195(8)
$O_0(\text{A})$	41468(34)	33685(30)	21578(4)	265(7)
$N_1(\text{A})$	37908(36)	66171(33)	20372(5)	163(6)
$C_1^{22}(\text{A})$	29143(44)	86328(38)	20698(6)	186(8)
$C_1^1(\text{A})$	32553(46)	95483(43)	17370(6)	223(8)
$F_1^{21}(\text{A})$	16273(27)	89927(26)	15329(4)	330(6)
$C_1^2(\text{A})$	51542(48)	85550(48)	16137(6)	239(9)
$C_1^3(\text{A})$	49625(45)	63882(45)	17358(6)	175(8)
$C_1(\text{A})$	70106(45)	54599(42)	17968(6)	194(8)
$O_1(\text{A})$	81310(31)	58832(32)	20199(5)	288(6)
$O_2(\text{A})$	74827(30)	41629(31)	15607(4)	244(6)
$C_2(\text{A})$	94687(48)	32699(56)	15884(7)	319(10)
$C_2^3(\text{B})$	67097(47)	53240(42)	-200(6)	225(8)
$C_0(\text{B})$	71000(42)	40544(43)	2759(6)	205(8)
$O_0(\text{B})$	87641(29)	33281(33)	3407(4)	265(6)
$N_1(\text{B})$	55071(34)	37171(36)	4756(5)	172(7)
$C_1^{22}(\text{B})$	34682(39)	45675(43)	4479(6)	180(8)
$C_1^1(\text{B})$	25886(44)	42092(43)	7801(6)	213(8)
$F_1^{21}(\text{B})$	31966(28)	57760(27)	9889(4)	314(5)
$C_1^2(\text{B})$	36080(48)	23063(44)	8950(6)	236(9)
$C_1^3(\text{B})$	57822(43)	25006(43)	7672(6)	174(8)
$C_1(\text{B})$	66760(40)	4643(41)	6876(6)	166(8)
$O_1(\text{B})$	61622(32)	-5881(30)	4677(5)	279(6)
$O_2(\text{B})$	80498(30)	-733(28)	9154(4)	213(6)
$C_2(\text{B})$	88789(52)	-20582(45)	8734(7)	274(9)

$$^a U_{\text{eq}} = (1/3) \sum_i \sum_j U_{ij} a_i a_j (a_i a_j)$$

Hydrogen atom positions were calculated and refined by using a riding model.

#### Ab initio calculations

Ab initio calculations were performed on an FPS-522 computer by using the electronic structure packages GAMESS (20) and SPARTAN (21). Geometries were first optimized at the STO-3G level of theory, and then fully optimized at the 3-21G level of theory. For each molecule (**1-3**), bond distances were calculated for the form of the particular peptide bond isomer (*cis* or *trans*) and ring pucker ( $C^\gamma$ -endo or  $C^\gamma$ -exo) observed in the crystalline state.

## RESULTS AND DISCUSSION

Approximately 10% of the residues in most forms of collagen are Hyp residues (**1-3**). The presence of the hydroxyl group on the  $C^\gamma$  atom of proline residues is known to increase the stability of collagen triple helices (5, 6). Here, we show that the inductive effect of the hydroxyl group affects the structure of the proline residue itself.

The crystalline structures of **1-3** were determined by

X-ray diffraction analysis. The crystallographic parameters for 1–3 are given in Table 1. Atomic coordinates and equivalent isotropic displacement coefficients for 1–3 are given in Tables 2–4. The molecular packing in crystals of 1–3 is shown in Fig. 1. Crystals of 1 contained one distinct molecule in each unit cell. In contrast, crystals of either 2 or 3 contained two distinct molecules (labeled A and B) in each unit cell. In crystals of 2, the hydroxyl group of molecule 2A donated a hydrogen bond to the amide carbonyl group of molecule 2B. The closest non-bonded distance from atom  $H_1^{\delta 2}$  (A) to  $O_2$  (B) was 1.680 Å. In molecule 2B, the  $C_0-N_1$  bond was  $0.023 \pm 0.007$  Å shorter and the  $C_0-O_0$  bond was  $0.007 \pm 0.007$  Å longer than were these bonds in molecule 2A. In crystals of 3, the corresponding bond lengths in molecules 3A and 3B did not differ significantly.

Perspective drawings of the crystalline structures of 1–3 are given in Fig. 2. References to atoms follow the recommendations of the IUPAC–IUB Commission on Biochemical Nomenclature (22), as indicated in Fig. 2. The X-ray diffraction analyses indicated that 2 and 3 crystallized with a *trans* (Z) amide bond. In contrast, 1 crystallized with a *cis* (E) amide bond. This conformation was observed in crystals from two independent preparations (second data set not shown). The *trans*

TABLE 5  
Length (Å) of bonds between nonhydrogen atoms of 1–3

Bond	1	2A	2B	3A	3B
$C_0^z-C_0$	1.500(3)	1.495(7)	1.507(7)	1.501(4)	1.503(4)
$C_0-O_0$	1.232(2)	1.225(5)	1.232(5)	1.230(3)	1.230(3)
$C_0-N_1$	1.347(2)	1.352(5)	1.329(5)	1.343(3)	1.353(3)
$N_1-C_1^{\delta 2}$	1.466(2)	1.463(6)	1.484(4)	1.468(3)	1.466(3)
$N_1-C_1^{\gamma}$	1.463(2)	1.450(6)	1.458(5)	1.468(3)	1.457(3)
$C_1^{\delta 2}-C_1^{\gamma}$	1.523(2)	1.509(5)	1.511(5)	1.513(3)	1.502(4)
$C_1^{\gamma}-X_1^{\delta 1}$	—	1.427(5)	1.421(6)	1.413(3)	1.409(3)
$C_1^{\gamma}-C_1^{\beta}$	1.535(3)	1.514(6)	1.509(7)	1.506(4)	1.511(4)
$C_1^{\beta}-C_1^{\alpha}$	1.545(2)	1.527(5)	1.535(5)	1.533(4)	1.535(4)
$C_1^{\alpha}-C_1$	1.520(2)	1.521(5)	1.499(7)	1.509(4)	1.515(4)
$C_1-O_1$	1.205(2)	1.190(6)	1.201(6)	1.211(3)	1.193(3)
$C_1-O_2$	1.342(2)	1.334(7)	1.319(5)	1.336(3)	1.352(3)
$O_2-C_2$	1.450(2)	1.433(5)	1.450(7)	1.445(4)	1.441(4)

isomer of 1 is about four-fold more abundant than is the *cis* isomer in a variety of solvents (23). Thus the *cis* isomer of 1 is disfavored, but the difference in energy between the *cis* and *trans* isomers is small. The crystallization of 1 in the *cis* conformation is not unusual, as 10–30% of prolyl amide bonds are in the *cis* conformation in small peptides of known crystalline structure (24).

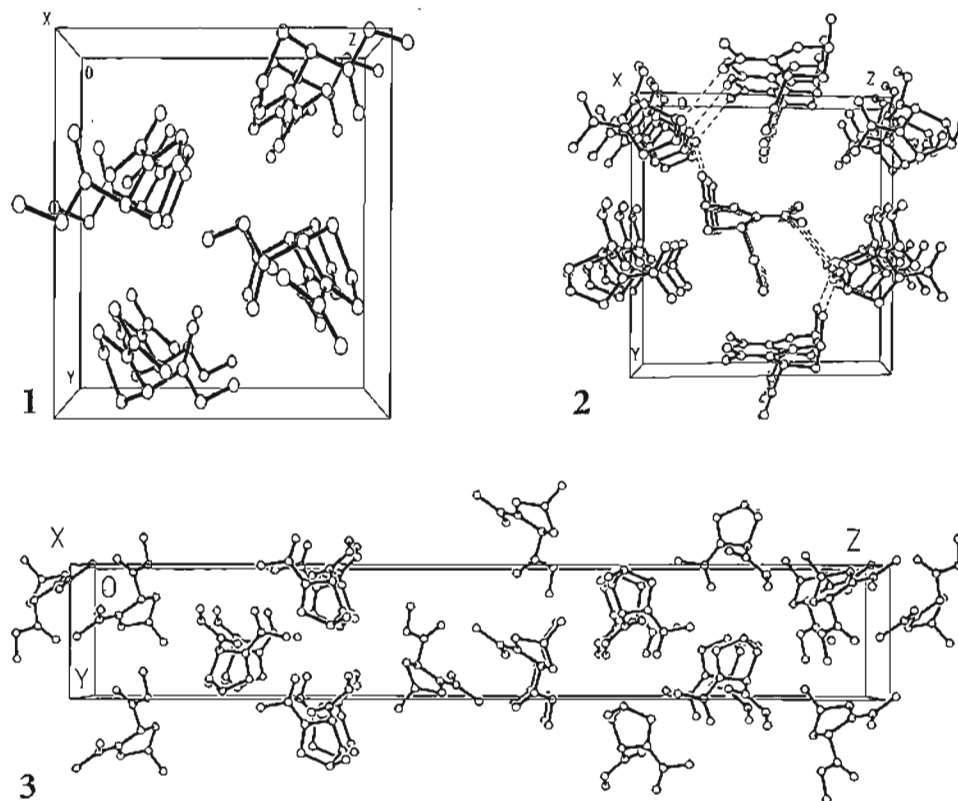


FIGURE 1  
Packing of crystalline 1, 2 and 3.

The pyrrolidine rings in crystalline 1–3 were puckered in one of two distinct conformations, as shown in Fig. 3. The pyrrolidine ring of crystalline 1 was puckered in the  $C^{\gamma}$ -endo conformation. (In the  $C^{\gamma}$ -endo conformation, the  $C_{\gamma}^{\delta}$  atom lies above the plane of the peptide bond, in the orientation drawn in Fig. 2.) This conformation is the one found most often in proline residues with *cis* peptide bonds (25). The pyrrolidine rings of both molecules A and B of crystalline 2 and 3 were puckered in the  $C^{\gamma}$ -exo conformation. This conformation is the one found most often in proline residues with *trans* peptide bonds (24). In 2 and 3, the  $C^{\gamma}$ -exo conformation places the hydroxyl or fluoro group in an axial position, which is also the preferred orientation of these groups in the 2' position of the ribose rings of nucleosides (26).

An electron-withdrawing substituent on the  $C_{\gamma}^{\delta}$  atom decreased the  $C_{\gamma}^{\delta 2}-C_{\gamma}^{\delta}$ ,  $C_{\gamma}^{\delta}-C_{\gamma}^{\beta}$  and  $C_{\gamma}^{\beta}-C_{\gamma}^{\alpha}$  bond lengths. The largest effect on bond length was a shortening of the  $C_{\gamma}^{\delta}-C_{\gamma}^{\beta}$  bond. Any substituent effect on bond length beyond  $C_{\gamma}^{\delta 2}$  or  $C_{\gamma}^{\delta}$  was too small to be resolved by the X-ray diffraction analysis. A plot of the  $C_{\gamma}^{\delta 2}-C_{\gamma}^{\delta}$ ,  $C_{\gamma}^{\delta}-C_{\gamma}^{\beta}$  and  $C_{\gamma}^{\beta}-C_{\gamma}^{\alpha}$  bond lengths determined from the crystalline structures versus the same bond lengths obtained from *ab initio* calculations is shown in Fig. 4. The proximity of the data to the dashed line in Fig. 4 indicates that the observed and calculated bond lengths largely agree. The  $C_{\gamma}^{\delta 2}-C_{\gamma}^{\delta}$ ,  $C_{\gamma}^{\delta}-C_{\gamma}^{\beta}$  and  $C_{\gamma}^{\beta}-C_{\gamma}^{\alpha}$  bond lengths in 1 and 2 did not differ significantly from those of other proline residues and

4-hydroxyproline residues, respectively, in the Cambridge Crystallographic Data Bank. (The Data Bank does not contain the structure of any 4-fluoroproline residues.)

The presence of an electron-withdrawing substituent on the pyrrolidine ring increased the pyramidalization of the nitrogen atom in crystalline 2 and 3. The pyramidalization of the nitrogen atom of an amide bond can be expressed by the parameter  $\delta_1$ , which here refers to the angle that the  $N_1-C_0$  bond makes with the plane defined by the  $N_1$ ,  $C_{\gamma}^{\delta}$  and  $C_{\gamma}^{\beta}$  atoms of the pyrrolidine ring (27). If  $\delta_1 = 0^\circ$ , then the atoms bound to  $N_1$  are planar. If  $|\delta_1| = 54.74^\circ$  [i.e.  $|\delta_1| = 1/2 \cos^{-1}(-1/3)$ ], then the atoms bound to  $N_1$  are tetrahedral. The value of  $\delta_1$  tends to be somewhat greater than  $0^\circ$  ( $-5 < \delta_1 < 10^\circ$ ) for prolyl peptide bonds in the *cis* conformation, and somewhat less than  $0^\circ$  ( $-10 < \delta_1 < 5^\circ$ ) for bonds in the *trans* conformation (27). In crystalline 1,  $\delta_1 = 1.05^\circ$ , indicating that  $C_0$  atom lies only slightly above the plane of the  $N_1$ ,  $C_{\gamma}^{\delta}$  and  $C_{\gamma}^{\beta}$  atoms, in the orientation is drawn in Fig. 2. In contrast, the values of  $\delta_1$  in crystalline 2 and 3 are  $-9.24$  (2A),  $-1.52$  (2B),  $-16.31$  (3A) and  $-3.54$  (3B). The deviation from planarity of the amide bonds of 2A and 3A is significantly larger than that of most prolyl peptide bonds (27). The trend in the pyramidalization at nitrogen ( $3 > 2 > 1$ ) is consistent with the inductive effects (14) of the fluoro, hydroxyl and hydrogen groups ( $3 > 2 > 1$ ). These results can be interpreted as indicating an increase in the  $sp^3$ -character of the prolyl nitrogen atom with increasing

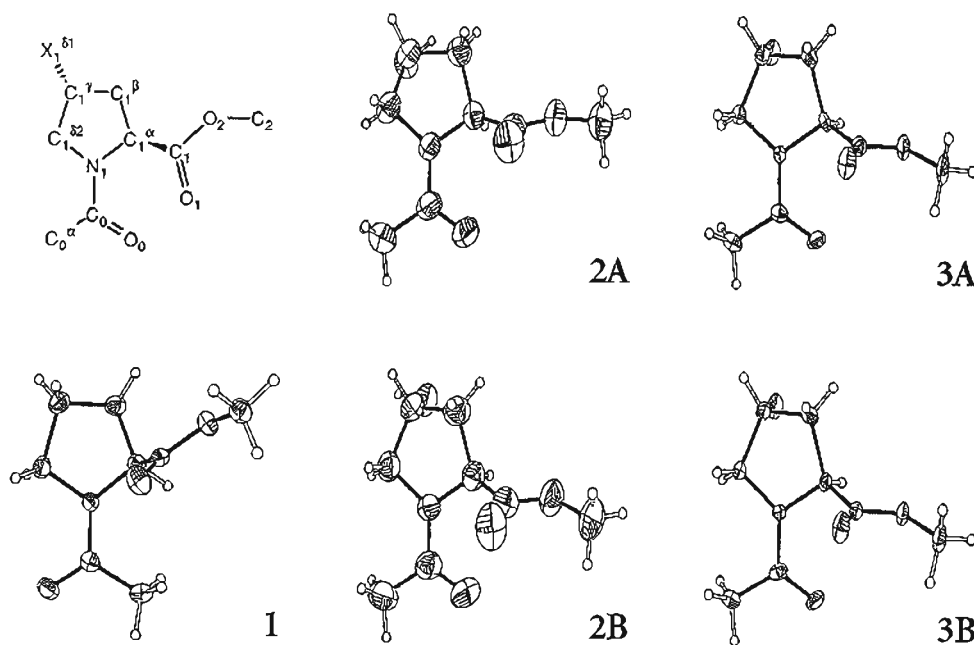


FIGURE 2

Labeling of nonhydrogen atoms in 1 ( $X = H$ ), 2 ( $X = OH$ ) and 3 ( $X = F$ ); and view showing amide bond conformation in crystalline 1, 2 (molecules A and B) and 3 (molecules A and B).

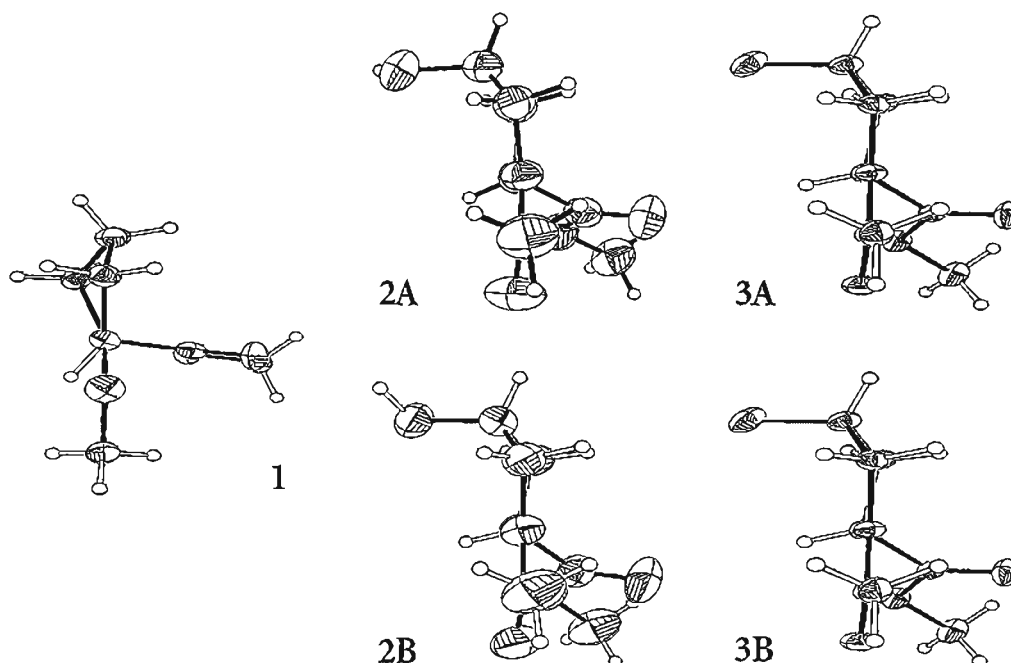


FIGURE 3

View showing pyrrolidine ring pucker in crystalline 1, 2 (molecules A and B) and 3 (molecules A and B).

electron withdrawal by a substituent in the 4-position of a proline residue.

The observed changes in nitrogen pyramidalization suggest that the inductive effects of substituents in the

4-position of proline residues can alter the distribution of electrons in a prolyl peptide bond. This alteration is likely to affect the kinetics of the *cis-trans* isomerization of peptide bonds that include the nitrogen atom of 4-substituted proline residues. The inductive effect on the energetics of prolyl peptide bond isomerization is now being determined with the analytical methods that we have described previously (28–30).

#### ACKNOWLEDGMENTS

Spectra were recorded at NMRFAM, which is supported by grant RR02301 (NIH). E.S.E. is a Wharton Predoctoral Fellow. A.S.E. is supported by Molecular Biophysics Training Grant 5 T32 GM08293 (NIH). R.T.R. is a Presidential Young Investigator (NSF), Searle Scholar (Chicago Community Trust), and Shaw Scientist (Milwaukee Foundation).

#### REFERENCES

1. Ramachandran, G.N. & Reddi, A.H., Ed. (1976) *Biochemistry of Collagen*, Plenum Press, New York
2. Bornstein, P. & Traub, W. (1979) in *The Proteins*, Neurath, H. & Hill, R.L., Eds., 3rd edn., Vol. 4, Academic Press, New York, pp. 411–632
3. Nemethy, G. (1988) in *Collagen*, Nimni, M.E., Ed., Vol. I, CRC Press, Boca Raton, FL, and references therein
4. Jones, E.Y. & Miller, A. (1991) *J. Mol. Biol.* **218**, 209–219
5. Sakakibara, S., Inouye, K., Shudo, K., Kishida, Y., Kobayashi, Y. & Prockop, D. (1973) *Biochim. Biophys. Acta* **303**, 198–202
6. Berg, R.A. & Prockop, D.J. (1973) *Biochem. Biophys. Res. Commun.* **52**, 115–120
7. Fraser, R.D.B., MacRae, T.P. & Suzuki, E. (1979) *J. Mol. Biol.* **129**, 463–481

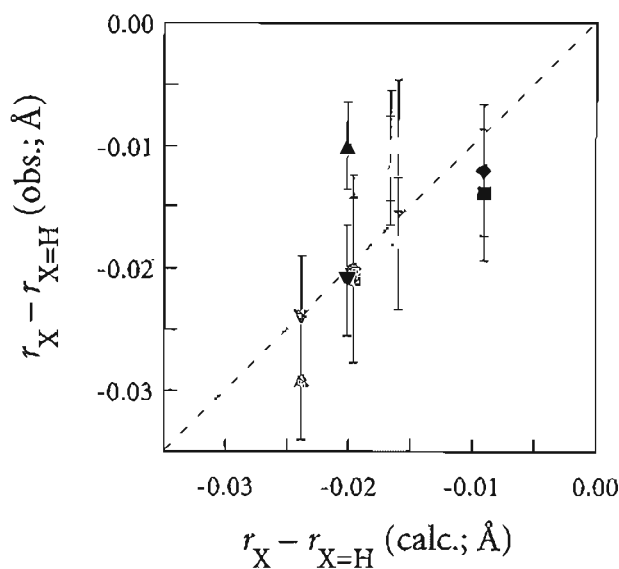


FIGURE 4

Difference in calculated and observed  $C_{\alpha}^{\beta 2}-C_{\gamma}^{\beta}$  (black),  $C_{\gamma}^{\beta}-C_{\delta}^{\beta}$  (dark grey) and  $C_{\delta}^{\beta}-C_{\epsilon}^{\beta}$  (light grey) bond lengths between 1 and 2A ( $\square$ ), 2B ( $\blacklozenge$ ), 3A ( $\blacktriangle$ ) or 3B ( $\blacktriangledown$ ). Values on the ordinate were observed in crystalline structures; values on the abscissa were calculated *ab-initio*. The dashed line (slope = 1) is for reference.

8. Li, M.-H., Fan, P., Brodsky, B. & Baum, J. (1993) *Biochemistry* **32**, 7377-7387
9. Okuyama, K., Okuyama, K., Arnott, S., Takayanagi, M. & Kakudo, M. (1981) *J. Mol. Biol.* **152**, 427-443
10. Fasman, G.D., Ed. (1989) *Practical Handbook of Biochemistry and Molecular Biology*, CRC Press, Boca Raton, FL, pp. 35, 39
11. Detar, D.F. & Luthra, N.P. (1977) *J. Am. Chem. Soc.* **99**, 1232-1244
12. Matsuzaki, T. & Iitaka, Y. (1971) *Acta Crystallogr., Sect. B* **27**, 507-516
13. Liang, G.-B., Rito, C.J. & Gellman, S.H. (1992) *Biopolymers* **32**, 293-301, and references therein
14. Exner, O. (1978) in *Correlation Analysis in Chemistry*, Chapman, N.B. & Shorter, J., Eds., Plenum, London, pp. 493-540
15. Ceppi, E., Eckhardt, W. & Grob, C.A. (1973) *Tetrahedron Lett.* **37**, 3627-3630
16. Still, W.C., Kahy, M. & Mitra, A. (1978) *J. Org. Chem.* **43**, 2923-2925
17. Ager, D.J. (1977) *J. Chem. Res. (S)* **6**, 6-7
18. Resnati, G. (1993) *Tetrahedron* **49**, 9385-9445
19. Sheldrick, G.M. unpublished results
20. Schmidt, M.W., Baldrige, K.K., Boatz, J.A., Jensen, J.H., Koseki, S., Gordon, M.S., Nguyen, K.A., Windus, T.L. & Elbert, S.T. (1990) *OCPE Bulletin* **10**, 52-54
21. Carpenter, J.E., McGrath, M.P. & Hehre, W.J. (1989) *J. Am. Chem. Soc.* **111**, 6154-6156
22. IUPAC-IUB Commission on Biochemical Nomenclature (1970) *Biochemistry* **9**, 3471-3479
23. Madison, V. & Schellman, J. (1970) *Biopolymers* **9**, 511-567
24. Stewart, D.E., Sarkar, A. & Warnpler, J.E. (1990) *J. Mol. Biol.* **214**, 253-260, and references therein
25. Milner-White, E.J., Bell, L.H. & Maccallum, P.H. (1992) *J. Mol. Biol.* **228**, 725-734
26. Cushley, R., Codrington, J.F. & Fox, J.J. (1968) *Can. J. Chem.* **46**, 1131-1140
27. Balaji, V.N., Rao, M.J., Dietrich, S.W. & Sasisekharan, V. (1986) *Biochem. Biophys. Res. Commun.* **140**, 895-900
28. Eberhardt, E.S., Loh, S.N., Hinck, A.P. & Raines, R.T. (1992) *J. Am. Chem. Soc.* **114**, 5437-5439
29. Eberhardt, E.S., Loh, S.N. & Raines, R.T. (1993) *Tetrahedron Lett.* **34**, 3055-3056
30. Eberhardt, E.S. & Raines, R.T. (1994) *J. Am. Chem. Soc.* **116**, 2149-2150

## Address:

Prof. Ronald T. Raines  
Department of Biochemistry  
University of Wisconsin - Madison  
420 Henry Hall  
Madison, WI 53706-1569  
USA