360-MHz ^{1}H NMR Conformational Analysis of Gly-Pro-X Peptides (X = Ala, Cha, Phe)†

Marc J. O. Anteunis,[1a] Frans A. M. Borremans,[1a] John M. Stewart,[1b] and Robert E. London*[1c]

Contribution from the Laboratorium voor Organische Chemie, Rujksuniversiteit, Gent, Belgium, The Department of Biochemistry, University of Colorado Medical School, Denver, Colorado 80262, and the Los Alamos Scientific Laboratory, Los Alamos, New Mexico 87545. *Received November 8, 1979*

Abstract: 360-MHz ^{1}H NMR studies of Gly-Pro-X tripeptides (X = Ala, cyclohexylalanine, and Phe) have been undertaken in order to obtain a more specific understanding of the physical basis for the high cis/trans ratio observed in Gly-Pro-Phe and peptides such as [Gly^6]bradykinin containing this sequence. Limited resolution of Pro ^{1}H resonances even at high field required the addition of $EuCl_3$ as shift reagent; however, spectral simulations show that the Pro ^{1}H couplings are not perturbed by the presence of the $EuCl_3$. Results indicate that the cis ↔ trans interconversion is accompanied by an extensive conformational rearrangement of the entire tripeptide. A conformational rearrangement of Pro in which the γ-endo/γ-exo ratio increases in each case coincident with the trans → cis transition is indicated by the $^3J_{HH}$ values. A strong correlation between the preference of the Phe side chain, χ^1, for the E_N conformation (phenyl ring trans to carbonyl and extended toward nitrogen) and the cis configuration of the Gly–Pro peptide bond is also apparent. These studies provide a more complete picture of the factors which determine the cis/trans ratio in proline-containing peptides and indicate that simple steric arguments advanced previously must be refined. Thus, at neutral pH, the cis probability is 17%, 17.5%, and 27.5% for X = Ala, cyclohexylalanine, and Phe, respectively. The results indicate that an electrostatic interaction involving the π electrons of the Phe side chain and the Gly–Pro peptide bond destabilizes the trans conformer. This interaction is closely analogous to that which leads to a high cis/trans ratio for C-terminal proline residues.

There is currently great interest in identifying and evaluating the conformational determinants of linear peptides. Proline-containing peptides are of particular interest since (1) proline is widely distributed in biologically active peptides, (2) the conformational restrictions introduced by the pyrrolidine ring lead to a more tractable problem in which specific interactions are reflected in the cis/trans ratio of the X–Pro peptide bond, and (3) the work of Brandts and co-workers indicates that an understanding of the cis/trans equilibrium of peptide bonds is fundamental to an understanding of protein denaturation–renaturation kinetics[2] as well as to peptide degradation by prolidase.[3] For the latter reason, the cis content of proteins may be of physiological significance, governing the rate of inactivation affected by proteolases.

Although polypeptides related to Gly-Pro-Phe have been of primary interest as models for collagen,[4] our interest in this tripeptide was originally stimulated by the observation that [Gly^6]bradykinin, which contains this sequence at positions 6–8, exhibits a considerably larger cis/trans ratio (0.62) than either bradykinin (having a Ser at position 6) or most other linear peptides with internal proline residues.[5] In this case, ^{13}C NMR studies have demonstrated the importance of the Phe residue following the proline in determining both the high cis/trans ratio and the unusual upfield shift behavior of the gly methylene carbon in the cis peptide. These observations led to two alternative conformational proposals. Specifically, the upfield shift of the cis Gly methylene carbon can reflect either (1) a ring current contribution due to proximity of the benzyl ring of Phe[8] or (2) a conformation in which a hydrophobic Pro–Phe interaction places the Pro carbonyl oxygen in close proximity with the Gly methylene carbon, leading to a large electrostatic contribution to the shift. The ^{13}C observations indicated the possibility of a large Gly H_α nonequivalence[6] and prompted a high-field ^{1}H NMR study. The present 360-MHz ^{1}H NMR studies provide a much more detailed conformational picture of the peptide in solution, lending strong support to the first alternative noted above.

In addition to bradykinin, large cis/trans ratios have been observed in corticotropin,[7] [Pro^3]oxytocin, and [Pro^3,Gly^4]-oxytocin,[8] as well as in model peptides.[9,10] In the latter studies, the cis–trans isomerism about X-Pro in X-Pro-Y peptides has been correlated with the steric requirements of both the preceding and succeeding residues, the amount of cis increasing with increasing bulkiness of these residues. For this reason, a comparative conformational study of Gly-Pro-X (X = Ala, Cha, Phe; Cha = cyclohexylalanine) is of particular value, since effects associated with the steric requirements of the large Cha and Phe side chains can be separated from those specifically associated with the electronic structure of the aromatic residue.

Experimental Section

The tripeptides Gly-L-Pro-L-Ala and Gly-L-Pro-L-Cha were purchased from Research Plus and Bachem, Inc., respectively. Gly-L-Pro-L-Phe was synthesized by using the solid-phase technique and characterized as described previously.[5]

All spectra were acquired in FT mode on a Bruker WH360 instrument at a temperature of 35 ± 0.5 °C in D_2O solution with 5% TSP (sodium 2,2,3,3-tetradeuterio-3-(trimethylsilyl)propionate) as the the internal standard. Gly-Pro-Ala, Gly-Pro-Cha, and Gly-Pro-Phe were dissolved as such (Zwitterionic form) at 0.041, 0.016, and 0.019 M, respectively.

The information content of the Pro proton patterns *only* was increased by adding $EuCl_3$ as shift reagent. Concentrations of $EuCl_3$ were optimized for minimum overlap of Pro signals of cis and trans isomers, maximum shift difference of Pro patterns in each isomer, and negligible line

† This work performed in part under the auspices of the U.S. Department of Energy.

(1) (a) Laboratorium voor Organische Chemie, Rijksuniversiteit. (b) Department of Biochemistry, University of Colorado Medical School. (c) Los Alamos Scientific Laboratory.

(2) (a) Brandts, J. F.; Halvorson, H. R.; Brennan, M. *Biochemistry* **1975**, *14*, 4953. (b) Brandts, J. F.; Brennan, M.; Lin, L.-N. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 4178. (c) Schmid, F. X.; Baldwin, R. L. *Ibid.* **1978**, *75*, 4764. (d) Creighton, T. E. *J. Mol. Biol.* **1978**, *125*, 401. (e) Bachinger, H. P.; Bruckner, P.; Timpl, R.; Engel, J. *Eur. J. Biochem.* **1978**, *90*, 595.

(3) Lin, L.-N.; Brandts, J. F. *Biochemistry* **1979**, *18*, 43.

(4) (a) Tamburro, A. M.; Scatturin, A.; Del Pra, A. *Int. J. Pept. Protein Res.* **1977**, *9*, 310. (b) Brahmachari, S. K., Ananthanarayanan, V. S.; Rapaka, R. S.; Bhatnagar, R. S. *Biopolymers* **1978**, *17*, 2097.

(5) London, R. E.; Stewart, J. M.; Williams, R.; Cann, J. R.; Matwiyoff, N. A. *J. Am. Chem. Soc.* **1979**, *101*, 2455.

(6) (a) Anteunis, M. J. O. *Tetrahedron Lett.* **1977**, 1535. (b) Anteunis, M. J. O.; Becu, C.; Lala, A. K.; Verhegge, G.; Narayan-Lala, K. *Bull. Soc. Chim. Belg.* **1977**, *86*, 161.

(7) Toma, F.; Fermandjian, S.; Low, M.; Kisfaludy, L. *Biochim. Biophys. Acta* **1978**, *534*, 112.

(8) Deslauriers, R.; Smith, I. C. P.; Levy, G. C.; Orlowski, R.; Walter, R. J. *J. Am. Chem. Soc.* **1978**, *100*, 3912.

(9) Grathwohl, C.; Wuthrich, K. *Biopolymers* **1976**, *15*, 2025, 2043.

(10) Deslauriers, R.; Becker, J. M.; Steinfeld, A. S.; Naider, F. *Biopolymers* **1979**, *18*, 523.

0002-7863/81/1503-2187$01.25/0

Table I. Shift[a] of Tripeptides Gly¹-Pro²-X³ in D_2O at 35 °C (pH ~6.0)

	Gly¹		Pro²							X³		
	α^1_A	α^1_B	α^2	β^2_A	β^2_B	γ^2_A	γ^2_B	δ^2_A	δ^2_B	α^3	β^3_A	β^3_B
X = Ala, trans	3.99_6	4.00_3	4.47_3	2.31_1	2.02_5	2.04	2.04	3.57_5	3.59_5	4.14_9	1.35_8	
cis (17%)	3.93_7	3.74_2	4.48_5	2.22	2.41_4	~1.94	2.16	~3.56	~3.59	4.19_5	1.35_8	
X = Cha, trans	4.01_3	3.98_7	4.49_0	2.31_0	2.02	2.04	2.04	3.57_0	3.63_0	4.24_4	1.63_1[c]	1.57_7[c]
cis (17.5%)	3.95_4	3.74_6	4.50_0	2.40_6	~2.00	2.20	~1.90	~3.59	~3.67	4.30_5		
X = Phe, trans	3.98_6	3.95_4	4.41_9	2.16_7	1.84_0	1.95_6	1.80_8	3.48_9	3.55_1	4.47_8	3.18_3	3.00_6
$\Delta \times 10^{2}$[b]	−1.0	−5.0	−5.4	−14.4	−18.5	−8.4	−23.2	−8.6	−4.4			
												(mean: −9.9)
cis (27.5%)	3.73	3.22_4	4.31_3	2.24_7	1.80_6	2.13_9	1.58_3	~3.54	~3.50	4.58_3	3.32_5	2.90_7
$\Delta \times 10^{2}$[b]	−20.5	−51.8	−17.2	−16.7	−41.4	−2.1	−35.7	−2	−9			

[a] In ppm, with TSP as internal standard. [b] Shift difference with Gly-Pro-Ala, negative sign if at higher field for Gly-Pro-Phe. [c] Assignments assessed by double irradiation experiments, e.g., $\{\alpha\}\beta$.

broadening. Optimum molar concentrations in $EuCl_3$ were 0.04 and 0.120 M for Gly-Pro-Cha and 0.033 M for Gly-Pro-Phe. No suitable shift effects could be induced for Gly-Pro-Ala.

All spectra were accumulated in a 32K (24 bit) data table (Aspect 2000 computer), with quadrature detection, pulse width 2.0 μs (10° flip angle), spectral width 4000 Hz, and acquisition time (AQ) 4.096 s. The digital resolution was 0.244 Hz/point. Spectral resolution was enhanced by Gaussian multiplication (GM), i.e., the FID was multiplied by a two-term exponential of the form $\exp(-at - bt^2)$ where $a = \pi$ LB and $b = -a/(2GB \cdot AQ)$. GM parameters (LB, GB) were Gly-Pro-Ala (−3.2, 0.15), Gly-Pro-Cha (−2.5, 0.17), and Gly-Pro-Phe (−2.8, 0.10).

Proline patterns were analyzed as seven-spin systems, with an iterative simulation program supplied by "Bruker Spectrospin NV" for the ASPECT 2000 computer. For Gly-Pro-Cha and Gly-Pro-Phe, simulations were started on the $EuCl_3$-doped samples, until the observed and calculated spectra were in good agreement. By keeping the thus obtained set of coupling constants and varying the chemical shifts, the nondoped spectra were simulated and found to be in excellent agreement with the experimental patterns. This procedure overcomes the problem of conformational perturbation induced by the shift reagent.[11] For Gly-Pro-Ala the couplings found in the Pro patterns for Gly-Pro-Cha gave a satisfactory fit with altered shifts. In a few cases of very strong coupling, e.g., the glycine protons in the trans peptides, data obtained in the presence of $EuCl_3$ and extrapolated to zero Eu^{3+} concentration were also used to analyze the coupling constants.

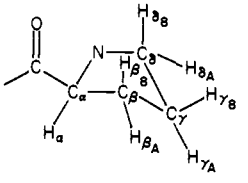
All assignments presented in Table I and II were confirmed by double-resonance decoupling experiments. Protons are labeled with Greek letters to indicate the carbon position to which they are bonded, with numerical superscripts to indicate the amino acid residue, and, in the case of degeneracy, with subscripts, A, B, ... thus, e.g., $\alpha_B{}^1$ corresponds to one of the Gly methylene protons, $\gamma_A{}^2$ to one of the protons bonded to Pro C_γ, etc.

Results

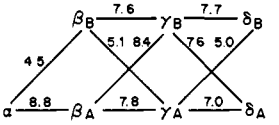
Trans Isomers. ^1H chemical shift and ^1H–^1H coupling constants deduced for the peptides Gly-Pro-X (X = Ala, Cha, Phe) are summarized in Tables I and II, respectively. The percent of cis isomer in each case is also included. In general, the shifts and coupling constants for the trans peptides indicate no special effects to be present.

^{13}C relaxation measurements[5] indicate considerable internal mobility for the N-terminal glycine residue. The average conformation of the Gly C_α–C′ bond (ψ^1) is reflected in the geminal coupling constant of the α protons, $^2J_{\alpha\alpha'}$.[12] The values of (−) 16.5 and 16.6 Hz (Table II) are typical for N-terminal glycine residues[6b] and correspond to an average value of $\psi^1 \sim 180°$.

The use of vicinal coupling constants to evaluate the conformation of the pyrrolidine ring of Pro has been discussed by Pogliani and co-workers.[13] Using the nomenclature indicated below,



the vicinal coupling constants may be represented by using a figure analogous to that of Pogliani et al., e.g., for *trans*-Gly-Pro-Cha.



Protons above (below) the plane of the ring correspond to the upper (lower) level of the diagram. Vicinal coupling constants connecting the upper and lower levels can be related to the ring conformation. The above pattern indicates approximate symmetry of the couplings relative to the $C_\gamma{}^2$–$\gamma_A{}^2$–$\gamma_B{}^2$ plane but significant asymmetry relative to the main plain of the ring. This pattern suggests that the observed couplings reflect a weighted average of the C_γ endo and C_γ exo conformers. Using the dihedral angles and Karplus relations deduced by Pogliani et al.,[13b] the fractional endo probability can be calculated from the observed vicinal coupling constants as summarized in Table III. On this basis, the conformation is concluded to reflect a C_s-C^γ_{endo}–C_s-C^γ_{exo} equilibrium near 60:40 endo:exo for X = Ala or Cha but a somewhat higher endo population, 85:15, if X = Phe. The deviation from symmetry about the $C_\gamma{}^2$–$\gamma_A{}^2$–$\gamma_B{}^2$ plane indicates small probabilities of other puckered conformations.

Limits on the values of ψ^2 can be deduced from the ^{13}C NMR data.[5] The chemical shift difference for the proline C_β and C_γ carbons, $\Delta_{\beta\gamma}$, is correlated with the magnitude of the dihedral angle $\theta = \psi - 60°$.[14] Values for the trans isomers of 5.24, 5.36, and 5.36 ppm corresponding to X = Ala, Cha, and Phe, respectively, are consistent with $|\theta| \simeq 90°$ and $\psi \approx -30°$ or +150° corresponding to the cis′ or trans′ conformations and precluding a significant γ-turn probability.[14] The theoretical calculations of Madison[15a] suggest that if the Gly–Pro bond is trans, both the cis′ and trans′ conformations will have similar stabilities, so that both forms may be present in solution.

Information about ϕ^3 was obtained on the basis of observations of the amide proton couplings, $^3J_{HNC_\alpha H}$. Due to experimental difficulties, the values were measured in MeSO-d_6; however, in one case, Gly-Pro-Cha, similar values were also obtained in aqueous solution. Application of a Karplus-type relation[16] gives interproton torsional angles of ~0° or ~150°, corresponding to $\phi^3 = \theta + 60° = 60°$, 210°, or −90°. The +60° value in combination with the cis′ conformation of proline places the Phe side chain close to the glycine carbonyl, as discussed in greater detail below.

(11) Kessler, H.; Molter, M. *J. Am. Chem. Soc.* **1976**, *98*, 5969.
(12) Barfield, M.; Hruby, V. J.; Meraldi, J. P. *J. Am. Chem. Soc.* **1976**, *98*, 1308.
(13) (a) Pogliani, L.; Ellenberger, M.; Valat, J.; Bellocq, A. M. *Int. J. Pept. Protein Res.* **1975**, *7*, 345. (b) Pogliani, L.; Ellenberger, M.; Valat, J. *Org. Magn. Reson.* **1975**, *7*, 61.
(14) (a) Dorman, D. E.; Bovey, F. A. *J. Org. Chem.* **1973**, *38*, 2379. (b) Madison, V.; Atreyi, M.; Deber, C. M.; Blout, E. R. *J. Am. Chem. Soc.* **1974**, *96*, 6725. (c) Siemion, I. Z.; Wieland, T.; Pook, K. H. *Angew. Chem., Int. Ed. Engl.* **1975**, *14*, 702. (d) London, R. E. *Int. J. Peptide Protein Res.* **1979**, *14*, 377.
(15) (a) Madison, V. *Biopolymers* **1977**, *16*, 2671. (b) Hodes, Z. I.; Nemethy, G.; Scheraga, H. A. *Biopolymers* **1979**, *18*, 1565.
(16) (a) Bystrov, V. F. *Prog. NMR Spectrosc.* **1976**, *10*, 41. (b) Cung, M. T.; Marraud, M.; Neel, J. *Macromolecules* **1974**, *7*, 606. (c) DeMarco, A.; Llinas, M. *J. Magn. Reson.* **1980**, *39*, 253.

Table II. Coupling Constants (in Hz) of Tripeptides Gly^1-Pro^2-X^3 in D_2O at 35 °C (pH ~6.0)

	Gly^1		Pro^2										X^3					
	$\alpha_A\alpha_B$[a]	$\alpha\beta_A$	$\alpha\beta_B$	$\beta_A\beta_B$	$\beta_B\gamma_B$	$\beta_B\gamma_A$	$\beta_A\gamma_B$	$\beta_A\gamma_A$	$\gamma_A\gamma_B$	$\gamma_B\delta_B$	$\gamma_B\delta_A$	$\gamma_A\delta_B$	$\gamma_A\delta_A$	$\delta_A\delta_B$	$\alpha\beta_A$[a]	$\alpha\beta_B$	$\beta_A\beta_B$	$\beta_A\gamma$	$\beta_B\gamma$	$HNC_\alpha H$[e]
Ala, trans	16.5	*b*													7.1					~7.5
cis	16.1_5	8.8	2.8_5	−13.3	7.4	2.9_5	10.5	6.9	−12.5	8.5	9.2	3.0_5	8.9	−11.0	7.1					~6.8
Cha, trans	16.5	8.8	4.5	−13.1	7.6	5.1	8.4	7.8	−13.0	7.7	7.6	5.0	7.0	−10.1	5.2_5	10.2	−13.9	8.4	5.0	~8.2 (7.8)
cis	16.2	8.9	2.8_5	-13.3_5	8.7	2.9_5	10.8	7.0_5	−13.2	8.5	9.2	3.1_5	7.0	−11.0	5.8^c	9.2^c	?		?	~7.6 (~8)
Phe,[d] trans	16.6	8.7	3.7	−12.0	7.1	4.6	9.2	7.3_5	−12.2	7.8	7.8	4.4	7.3	−9.85	5.3_5	7.5_5	−13.9			~8.0
cis	16.2	8.8_5	2.4	−13.2	7.1	2.5	11.4	7.0	−13.4	8.8	9.4	2.5	7.1	−11.0	4.6	10.4_5	-14.0_5			~7.0

[a] The subscripts A and B in the Gly^1 and X^3 residues refer to low field and high field, respectively. In the Pro^2 residue the A and B refer to protons trans and cis respectively, with respect to the carboxylic group. [b] Same values as for *trans*-Gly-Pro-Cha within 0.2 Hz. [c] Apparent coupling constants. [d] Coupling constants in the Pro residue are obtained as best fits from iterative simulations. [e] Amide proton coupling constants are for Me_2SO-d_6 solutions. For the case of Gly-Pro-Cha, values obtained in H_2O–D_2O (5:95) are also included.

Table III. Calculation of γ-Endo Conformational Probabilities for Proline[a]

coupled protons	calculation of P_{endo}
α,β_A	$^3J = 9.64P_{endo} + 7.78(1 - P_{endo})$
α,β_B	$^3J = 2.63P_{endo} + 9.28\ (1 - P_{endo})$
γ_B,β_A; γ_B,δ_A	$^3J = 10.02P_{endo} + 1.46(1 - P_{endo})$
γ_A,β_B; γ_A,δ_B	$^3J = 1.46P_{endo} + 10.02(1 - P_{endo})$

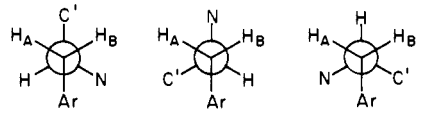
[a] Based on the treatment of Pogliani et al. and the nomenclature indicated in the text.

Table IV. Rotameric Distribution (χ^1) of Aromatic Residues in Peptides

peptide	pH/solvent	E_N	E_O	F	ref
cis-Gly-Pro-Phe	6	71	18	11	this work[a]
trans-Gly-Pro-Phe	6	45	25	30	this work[a]
Acetyl-Phe	ionized	75	22	2	21
Acetyl-Phe-*N*-methylamide	neutral	53	38	9	25
	Me_2SO	63	22	15	24
[Met^5]enkephalin[b]	neutral	60	26	14	25
	Me_2SO-d_6	69	13	18	25
TFA-Gly-Ala-Tyr-Gly OMe		61	21	18	23
TFA-Gly-Gly-Tyr-Ala OMe		65	19	16	23
cis-Gly-L-(4-*trans*-*F*)-Pro-Trp	2.7	major			26
trans-Gly-L-(4-*trans*-*F*)-Pro-Trp	2.7	46	33	21	26

[a] Data taken in D_2O at 35 °C. [b] Data for the zwitterionic form of the peptide at 35 °C.

Conformation of the aromatic side chain is readily deduced from the $^3J_{HH}$ values. The value for χ^1(Phe) is particularly relevant in the present context. Rotational isomers are illustrated below:



E_N(I) E_O(II) F(III)

Nomenclature for the rotameric states is based on that introduced by Young et al.,[17] in preference to the (nonuniform) use of Roman numeral designations. Thus *F* (folding) is equivalent to III above, E_O and E_N correspond to extended toward oxygen and nitrogen, respectively. Rotamer probabilities, summarized in Table IV, are calculated by using Pachler's formulation of the Karplus equation (J_g = 2.6 Hz, J_a = 13.56 Hz).[18] The two observed coupling constants, $\alpha^3\beta_A{}^3$ and $\alpha^3\beta_B{}^3$, coupled with the assignment of the high-field ($\beta_B{}^3$) proton as pro-R,[19] allow evaluation of all three rotamer probabilities. The rotameric distribution was found to be strongly dependent on whether the Gly–Pro bond was cis or trans. Results are summarized in Table IV. Although the numerical values are dependent on the particular parametrization of the Karplus equation,[19,20] the strong dependence on the isomeric state of the Gly–Pro bond is not altered.

The rotameric distribution of the Cha side chain was obtained on the basis of the assignment of the upfield resonance to $\beta_B{}^3$. This assignment is analogous to that for the Phe β protons and is consistent with the assignments of AcChaNHMe and AcChaOEt by Kobayashi et al.[21] on the basis of specifically deuterated analogues. The computed rotamer probabilities are then E_N (69%), E_O (24%), and *F* (6.5%). This E_N probability is therefore significantly greater than that observed in the trans Gly-Pro-Phe case (Table IV).

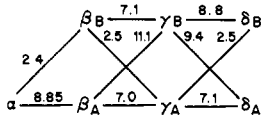
Proton shifts of the Gly and Pro residues in the trans isomers were all found to be similar, except for small upfield shifts (mean

(17) Young, P. E.; Madison, V.; Blout, E. R. *J. Am. Chem. Soc.* **1976**, *98*, 5365.
(18) Pachler, K. G. R. *Spectrochim. Acta* **1964**, *20*, 581.
(19) Feeney, J. *J. Magn. Reson.* **1976**, *21*, 473.
(20) Callens, R. E.; Anteunis, M. J. O. *Biochim. Biophys. Acta* **1979**, *577*, 324.
(21) Kobayashi, J.; Nagai, U. *Biopolymers* **1978**, *17*, 2265.

$\simeq$ 0.10 ppm) in the Gly-Pro-Phe with β_B^2 and γ_B^2 most affected.

Cis Isomers. In contrast to the above observations, there are a number of pronounced anomalies in the cis isomers. The $^2J_\alpha$ Gly value is systematically 0.3–0.4 Hz less negative than in the trans isomer. The low −16.1 to −16.2 values in the cis isomers indicate a deviation of $\sim$15° for the mean value of ψ[6b,12] which may be ascribed to steric strain between one of the Gly methylene protons (α_B^1) and the Pro carbonyl carbon (vide infra).

The conformation of the pyrrolidine ring of Pro, deduced as noted above from the $^3J_{HH}$ values, is shifted toward higher γ-endo population in comparison to the corresponding trans isomer. For Gly-Pro-Phe, γ-exo is almost totally absent, and there is a substantial amount of C_2-C^γ_{endo}–C^β_{exo} or C_s-C^β_{exo} present, as indicated by the dramatic asymmetry of the couplings with respect to the main plane of the ring. For the γ-envelope conformation, $\phi^2 > 60°$, consistent with the steric interaction between Pro C′ and Gly methylene proton α_B^1 deduced independently from the low value of $^2J_\alpha$ as discussed above.



cis-Gly-Pro-Phe

Values for ψ^2 may be deduced from the ^{13}C $\Delta_{\beta\gamma}$ values of 9.69, 9.6, and 9.76 ppm, for X = Ala, Cha, and Phe, respectively. On the basis of the curve given by Siemion et al.,[14c] these correspond to $|\theta| = 120°$ and $\psi = +180°$ or −60°, again limiting the conformational possibilities to trans′ and cis′. In this case, however, the theoretical analysis given by Madison[15a] suggests that the cis–cis′ proline conformation will be significantly ($\sim$2 kcal/mol) more stable than the trans′ case.

Values for $^3J_{HNC_\alpha H}$ were found to be consistently smaller in the cis isomer. This difference could reflect a greater relative probability of the $\theta = 0°$ rotamer which, according to the more recent formulations of the Karplus curve, has a value of 6.4 Hz corresponding to this value of θ. The proline cis–cis′ conformation combined with a ϕ^3 value of +60° places the Phe side chain in close proximity with the Gly residue, as discussed in greater detail below.

The rotamer populations corresponding to χ^1(Phe) are markedly different in the cis relative to the trans conformation (Table IV). In particular, the E_N rotamer is strongly predominant in the former.

Chemical shift data for the three cis peptides also indicates a significant difference between the X = Phe and X = Ala, Cha cases. Values for $\Delta \times 10^2$ included in Table I, where Δ is the chemical shift difference relative to Gly-Pro-Ala, indicate that α_B^1 is strongly shifted upfield, with smaller upfield shifts also experienced by Pro β_B^2 and γ_B^2. These results suggest contributions from ring current effects arising from a nearby benzyl side chain of the Phe residue, as realized in the models discussed below.

Discussion

Conformation of Gly-Pro-Phe. It is apparent from the data summarized above that the trans → cis transition of the Gly–Pro bond is associated with many conformational readjustments involving all of the residues of the tripeptides. The difference in the cis/trans ratios obtained for X = Cha and X = Phe, as well as the conformational change in χ^1(Phe) accompanying the trans to the cis peptide, indicate that a specific interaction involving the aromatic side chain is important in determining this ratio. A model for the predominant conformation of Gly-Pro-Phe consistent with the data presented here is illustrated in Figure 1.

The relatively large cis/trans ratio in Gly-Pro-Phe reflects in part the preference of the Phe residue for the E_N conformation. As indicated in Figure 1a, this conformation places the face of the ring in close proximity with the Gly carbonyl oxygen. The resulting electrostatic repulsion between the π electron cloud of the former and the oxygen, illustrated in Figure 1a, destabilizes the trans conformer. Conversely, it is also possible that an attractive electrostatic interaction between the π cloud of Phe and the Gly carbonyl carbon stabilizes the cis conformation of the Gly–Pro bond. Similar electrostatic effects have been invoked to explain ASIS shifts accompanying aromatic solvation.[22] The large upfield shift of the Gly α_B^1 resonance in the *cis*-Gly–Pro–Phe peptide is a consequence of the ring current contribution arising from the proximity of the proton to the Phe side chain.

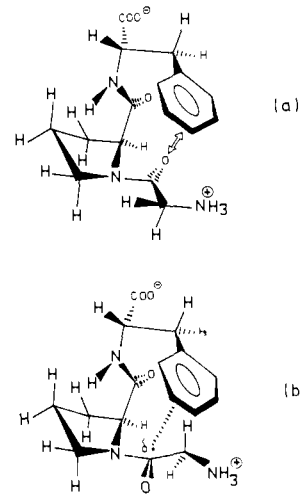


Figure 1. Predominant trans (a) and cis (b) conformation of Gly-Pro-Phe. Repulsive electrostatic interactions of the Gly carbonyl oxygen and the benzyl ring are indicated in (a) and a possible attractive interaction in (b). Upfield shift of the Gly C_α and α_B^1 resonances in (b) result from a ring current effect. As noted in Table IV, the E_N conformation of Phe illustrated in both (a) and (b) is less predominant in (a) than in (b).

The correlation of the cis/trans ratio of Gly-Pro with the E_N population of χ^1(Phe[3]) is central to the present analysis. The strong preference of the Phe side chain for the E_N conformation is a general characteristic of aromatic side chains.[21-25] Some typical rotamer populations are summarized in Table IV. There is a particularly close analogy with the tripeptide Gly-(4-*trans*-*F*)Pro-Trp studied by Gerig and McLeod.[26] In addition to the parallel dependence of the E_N conformational probability on the cis/trans ratio of the Gly–Pro peptide bond, a pronounced nonequivalence of the Gly H_α resonances in the cis isomer is observed.

The model illustrated in Figure 1 assumes a cis′ conformation for the Pro[2] residue, i.e., ψ^2(Pro) $\sim$ −50°. This is consistent with calculations for proline-containing peptides with the pyrrolidine ring in the γ-envelope form, and this indicates that the cis–cis′ conformation corresponds to the global minimum.[15] Upfield shifts of Pro β and γ resonances may reflect the presence of cis–cis′ isomer in which the benzyl side chain is located near the protons. The correlation of the pyrrolidine ring conformation of Pro[2] with the cis/trans nature of the peptide bond is similar to that observed in acetyl proline[13a] and predicted theoretically by Madison.[15a] Qualitatively similar conclusions can be reached from the ^{13}C relaxation data, as discussed below.

Correlation with Previously Reported ^{13}C NMR Data. ^{13}C NMR data including chemical shifts of the Gly residues, cis/trans ratios for each of the three tripeptides considered here, and spin–lattice relaxation measurements of Gly-Pro-Phe at pH 9.4 have previously been reported.[5] In general, there is a close analogy between many of the ^1H parameters reported here and the ^{13}C parameters. In particular, the large upfield shift of the Gly α_B^1 resonance parallels a similar upfield shift of the Gly C_α observed

(22) (a) Moriarty, R. M.; Kliegman, J. M. *J. Org. Chem.* **1966**, *31*, 3007. (b) Kliegman, J. M.; Young, L. B. *J. Phys. Chem.* **1965**, *69*, 1777. (c) Matton, H. V.; Richards, R. E. *Mol. Phys.* **1960**, *3*, 253; **1962**, *5*, 139. (d) La Planche, L. A.; Rogers, M. T. *J. Am. Chem. Soc.* **1965**, *85*, 3728; **1964**, *86*, 337. (e) Madison, V.; Schellman, J. R. *Biopolymers* **1970**, *9*, 511.
(23) Wuthrich, K.; De Marco, A. *Helv. Chim. Acta* **1976**, *59*, 2228.
(24) Kobayashi, J.; Nagai, U. *Tetrahedron Lett.* **1977**, 1803.
(25) Kobayashi, J.; Nagai, U.; Higashijima, T.; Miyazawa, T. Biochim. *Biophys. Acta* **1979**, *577*, 195.
(26) Gerig, J. T.; McLeod, R. S. *J. Am. Chem. Soc.* **1976**, *98*, 3970.

Table V. Chemical Shift Differences δ (Cis-Trans) for Glycine Nuclei of Gly-Pro-X

Gly nucleus[a]	Gly-Pro-Ala	Gly-Pro-Cha	Gly-Pro-Phe	net shielding[b] for Gly-Pro-Phe
$^1H\ \alpha_B$	−0.26	−0.24	−0.73	−0.5
$^1H\ \alpha_A$	−0.06	−0.06	−0.25	−0.2
$^{13}C\ \alpha$	−0.18	−0.32	−0.60	−0.35
$^{13}C_{C=O}$	+0.58	+0.44	0.0	−0.5

[a] Values for carbon shifts from ref 5. [b] Net shielding reflects shift difference between Gly-Pro-Phe and average for the other two peptides.

in the cis isomer. This was proposed to reflect either a ring current effect arising from proximity of the Gly C_α and the Phe side chain or an electrostatic effect arising from proximity of the Gly C_α and the Pro carbonyl oxygen. Using the chemical shifts of the trans conformers as an internal reference, it is now possible to evaluate the two proposed conformational models in greater detail. Values for δ (cis–trans) for the glycine nuclei are summarized in Table V, from which it is readily seen that the data for Gly-Pro-Phe differ significantly from that for the other two peptides. Entries reflecting this average difference are also included in Table V. These data are generally not consistent with a linear electric field shift contribution (LEFS) of the type discussed by Batchelor[27] arising from the Pro[2] carbonyl oxygen. In particular the LEFS of ^{13}C are generally many times greater than those observed for ^1H, in contrast with the last column of Table V. Further the ^{13}C and ^1H shift will frequently have opposite signs. For example, in carboxylic acids deprotonation of the carboxyl group typically shifts the carbon resonance downfield and the protons upfield.[28] The observation of a net shield for both glycine protons in *cis*-Gly-Pro-Phe would require that both C–H bonds be directed away from the Pro[2] carbonyl oxygen on the average, but such an orientation would lead to a downfield shift for Gly C_α. In contrast, the shifts are all more readily explained on the basis of a ring current contribution arising from proximity to the Phe side chain.

It is also interesting to consider the Gly-Pro-Phe relaxation data in light of the present results. Differences in the NT_1 values for the proline carbons can be explained by a model assuming at least two puckered conformations of the ring with roughly similar probabilities.[29] In general there are more parameters available than are necessary to fit only the ^{13}C NT_1 data. The ^{13}C relaxation data are most consistent with a jump intermediate between C_s-$C^\gamma_{exo} \leftrightarrow C_s$-$C^\gamma_{endo}$ and C_2-C^β_{exo}–$C^\gamma_{endo} \leftrightarrow C_2$–$C^\beta_{endo}$–$C^\gamma_{exo}$, a result consistent with the ^1H coupling data presented here. However, an interesting difference between the NT_1^γ/NT_1^α ratios for the cis and trans states, viz., 1.74 and 2.02, is measured. This small difference was originally analyzed by using a slightly greater value for the range of motion ($2\theta = 56°$) for the trans compared with the cis ($2\theta = 52°$) puckering jump. Such an effect can also reflect differences in the proportion of the two puckered forms in the cis vs. trans peptides. Thus, the increase in the γ-endo/γ-exo ratio deduced from ^1H NMR data to be coincident with the trans → cis transition of the Gly–Pro bond will produce a similar reduction in the NT_1^γ/NT_1^α value. For example, assuming a 56° range of motion for both the cis and trans isomers and rapid internal jumps with $\tau_A = 10^{-12}$ s, the data can be fit by using a 60:40 ratio for the probabilities of the two states, P_A:P_B, for the trans isomers and a 70:30 ratio for the cis isomers. Thus, the ^{13}C relaxation data show clearly that although the endo conformation becomes largely predominant in the cis peptide, a significant probability of the γ-exo is still present, since the NT_1^γ/NT_1^α ratio does not drop to 1.0. These results indicate that a more complete description of the proline ring puckering may be obtained by combining ^{13}C relaxation data and ^1H coupling constant data.

Factors Determining the Cis/Trans Ratio. An understanding of the factors which determine the cis/trans ratio of X–Pro peptide bonds is of value in connection with the suggested role of cis/trans isomerism in determining protein denaturation–renaturation kinetics[2] as well as proteolytic degradation of X–Pro bonds.[3] Previous studies have noted a correlation between the steric bulkiness of the residues preceding[9,10] or following[10] proline and the cis/trans ratio. However, in one of these studies[10] it was concluded that the effect of steric bulk on the cis content does not appear to be present in D_2O. The present study, as well as data reported previously,[5] underline the need for a more detailed understanding of this phenomenon, since the simple steric argument fails to explain the similar cis/trans ratios exhibited by Gly-Pro-Cha and Gly-Pro-Ala, as well as the difference between Gly-Pro-Cha and Gly-Pro-Phe.

The importance of electrostatic interactions in determining the cis/trans ratio in proline-containing peptides is supported by the studies of Grathwohl and Wuthrich.[9] For example, deprotonation of the carboxyl group of C-terminal proline residues typically results in an increased cis/trans ratio. The present data suggests a closely analogous effect in Gly-Pro-Phe such that an electrostatic interaction between the Gly carbonyl oxygen and the π electrons of the Phe side chain results in a destabilization of the trans conformation and an increase in the cis/trans ratio. Further, in the predominant E_N conformation, the Phe side chain maintains a spatial orientation relative to the X–Pro peptide bond which is closely analogous to that of the proline carboxyl group in peptides containing a C-terminal Pro residue (Figure 1). Thus, a pattern emerges in which the high cis/trans ratios observed for both C-terminal proline residues and internal proline residues followed by aromatic residues can be explained on the basis of an unfavorable electrostatic interaction in the trans isomer. This observation is of general interest, since predictions regarding the cis/trans nature of peptide bonds for internal proline residues in unfolded proteins are of value in the interpretation of the kinetics of protein folding and unfolding studies.[2]

(27) (a) Batchelor, J. G. *J. Am. Chem. Soc.* **1975**, *97*, 3410. (b) Batchelor, J. G.; Feeney, J.; Roberts, G. C. K. *J. Magn. Reson.* **1975**, *20*, 19.
(28) Hagen, R.; Roberts, J. D. *J. Am. Chem. Soc.* **1969**, *91*, 4504.
(29) London, R. E. *J. Am. Chem. Soc.* **1978**, *100*, 2678.