Fluoroprolines as Tools for Protein Design and Engineering**

Christian Renner, Stefan Alefelder, Jae H. Bae, Nediljko Budisa, Robert Huber, and Luis Moroder*

The naturally occurring fluorine isotope, ^{19}F, has excellent nuclear magnetic resonance properties.[1] It is the second most sensitive nucleus after hydrogen and exhibits even higher sensitivity towards changes in its chemical environment. Despite some minor differences in the van der Waals radius and particularly in length and dipole moment of the C–F bond, fluorine is generally considered as isosteric to hydrogen that it replaces. Taking into account that fluorine occurrence in biological molecules is extremely rare,[2] natural background signals are absent, and thus ^{19}F NMR spectroscopy has become a useful tool for investigating protein structure and dynamics, as well as interactions between fluorinated ligands and receptors. In vivo incorporation of (4*S*)-fluoroproline ((4*S*)-FPro) into collagen and its proto- and procollagen precursors has been observed in cell cultures; its effect was to decrease the stability and/or folding rate of the collagen triple helix.[3] Conversely, in synthetic collagen model peptides that contain (4*R*)-4-fluoroproline at the positions of hydroxyproline, the stability of the triple helical fold was found to be markedly enhanced,[4] a fact which was attributed to the inductive effect of fluorine as derived from the kinetic and thermodynamic properties of (4*R*)-FPro in dioxane.[5]

For a more appropriate correlation between the thermodynamics of the *cis*/*trans* isomerization of fluorinated prolines and the stability of related protein mutants, we have characterized in a comparative study the kinetics and thermodynamics of the isomerization of Ac-L-Pro-OMe (**1**), Ac-(4*R*)-L-FPro-OMe (**2**), Ac-(4*S*)-L-FPro-OMe (**3**), and Ac-(4)-L-F_2Pro-OMe (**4**) in aqueous solution by FT-IR and $^1H/^{19}F$ NMR spectroscopy. The observed preference of Ac-(4*S*)-L-FPro-OMe for the *cis* conformation compared to the situation for **1** and particularly to that for **2** was confirmed by the differentiated thermostabilities of related barstar mutants.

In the model compounds the known pH effect on the isomerization of Xaa–Pro bonds[6] was prevented by the N- and C-derivatization. The inductive effect of the F substituents, reported previously for Ac-(4*R*)-L-FPro-OMe (**2**) in dioxane,[5] was observed in aqueous media for both monofluorinated proline derivatives **2** and **3**, and as expected, more intensively for the difluoro derivative **4** by a blue shift of the amide–I band (C=O stretching) compared to that of **1** in the FT-IR spectrum (**2**: $+8.5\ cm^{-1}$, **3**: $+12.5\ cm^{-1}$, **4**: $+18\ cm^{-1}$). This strengthening of the C=O bond accompanies a weakening of the double-bond character of the Ac–Pro amide bond. This inductive effect is also well evidenced by the pK_a value of the immonium group of the free amino acid L-4,4-F_2Pro (7.15 ± 0.1) compared to the known pK_a values of 9.2 for (4*R*)-L-FPro and of 10.8 for L-Pro.[5] The resulting pyramidization of the N atom of the Ac–Pro bond leads to an acceleration of the *trans*/*cis* isomerization in the order **1** < **2** < **3** < **4** (Figure 1). From Eyring plots of the isomerization rates

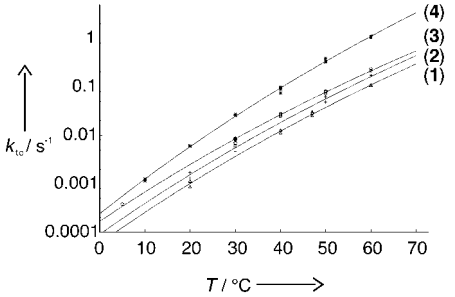

Figure 1. Logarithmic plot of the *trans*→*cis* isomerization rates k_{tc} of compounds **1**–**4** versus temperature. This type of plot is similar, but not equivalent to an Eyring plot; it was used to allow direct extraction of the isomerization rates k_{tc}. Multiple symbols at one temperature indicate multiple measurements with different NOESY mixing times (τ_m).

k_{tc} between 4 and 60 °C the activation enthalpy $\Delta H^{\neq}$ and entropy $\Delta S^{\neq}$ were determined (Table 1). The equilibrium constants K_{ZE} (*trans*/*cis* isomer ratios) show a completely different behavior, that is an order of **3** < **4** < **1** < **2** (Figure 2).

Table 1. Activation enthalpy ($\Delta H^{\neq}$) and entropy ($\Delta S^{\neq}$) as derived from Eyring plots. Error limits obtained from the residuals of the linear least-squares fitting were 1–2% for compounds **1**, **3**, and **4**, and 2–3% for compound **2**. Additionally the free energy of activation at 300 K ($\Delta G^{\neq}_{300\,K}$) is given.

Proline derivative	*trans*→*cis*			*cis*→*trans*		
	$\Delta H^{\neq}$ [kJ mol^{-1}]	$\Delta S^{\neq}$ [J mol^{-1} K^{-1}]	$\Delta G^{\neq}_{300\,K}$ [kJ mol^{-1}]	$\Delta H^{\neq}$ [kJ mol^{-1}]	$\Delta S^{\neq}$ [J mol^{-1} K^{-1}]	$\Delta G^{\neq}_{300\,K}$ [kJ mol^{-1}]
Ac-Pro-OMe **1**	92.3	12.8	88.4	87.2	9.00	84.5
Ac-(4*R*)-FPro-OMe **2**	91.8	14.8	87.4	84.2	5.37	82.6
Ac-(4*S*)-FPro-OMe **3**	87.5	4.17	86.3	84.7	2.42	84.0
Ac-4,4-F_2Pro-OMe **4**	103.7	66.38	83.8	98.5	59.11	80.8

Unexpectedly, the free energy difference between the *cis* and *trans* isomer is smallest for the (4*S*)-L-FPro derivative **3** (Table 2), while the K_{ZE} value for the difluoro derivative **4** is somewhat smaller than that of the proline compound **1**. van't Hoff plots served to derive the enthalpic and entropic contributions to the free energy difference between the *cis* and *trans* isomers (Table 2); these contributions are in good agreement with the corresponding differences between $\Delta H^{\neq}$ and $\Delta S^{\neq}$ for the *cis*→*trans* and the *trans*→*cis* isomerization.

[*] Prof. L. Moroder, Dr. C. Renner, Dr. S. Alefelder, J. H. Bae, Dr. N. Budisa, Prof. R. Huber
Max-Planck-Institut für Biochemie
Am Klopferspitz 18A, 82152 Martinsried (Germany)
Fax: (+49) 89-8578-2847
E-mail: moroder@biochem.mpg.de

[**] We thank Mrs. E. Weyher for skillful technical assistance in spectroscopic analyses and Mrs. W. Wenger for her excellent technical assistance in protein preparation. We are indebted to Dr. R. Golbik for providing us with barstar plasmid and protocols for its isolation and purification.

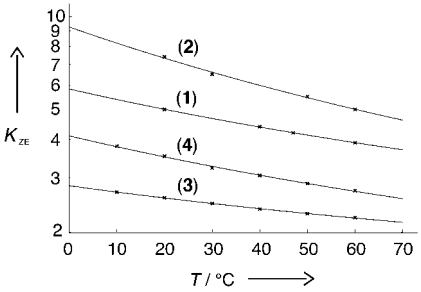

Figure 2. Logarithmic plot of the equilibrium constants K_{ZE} of the *trans*/*cis* isomerization of the amide bond of compounds **1**–**4** versus temperature. This type of plot is similar, but not equivalent to a van't Hoff plot; it was used to allow direct extraction of the K_{ZE} values.

Table 2. Enthalpy (ΔH^0) and entropy (ΔS^0) contributions to free energy difference between *trans* and *cis* isomers as derived from van't Hoff plots. Error limits were obtained from the residuals of the linear least-squares fitting. Additionally the free energy difference ΔG^0 at 300 K is given.

Proline derivative	ΔH^0 [kJ mol^{-1}]	ΔS^0 [J mol^{-1} K^{-1}]	$\Delta G^0_{300\,K}$ [kJ mol^{-1}]
Ac-Pro-OMe **1**	−5.04 (±1.0%)	−3.82 (±4.3%)	−3.90
Ac-(4*R*)-FPro-OMe **2**	−7.73 (±3.3%)	−9.81 (±8.3%)	−4.78
Ac-(4*S*)-FPro-OMe **3**	−3.04 (±1.1%)	−2.47 (±4.3%)	−2.30
Ac-4,4-F_2Pro-OMe **4**	−5.21 (±2.3%)	−7.32 (±5.3%)	−3.01

The large difference in the K_{ZE} values of **2** and **3** relates to the different pucker of the proline ring. For 4-fluoro-L-proline the influence of the substitution site (*S* or *R*) on the proline ring pucker was determined as early as 1973 by NMR spectroscopy.[7] From $^1H-^1H$ and $^1H-^{19}F$ NMR coupling constants a *γ-exo*-puckered structure was derived for (4*R*)-L-FPro and a *γ-endo*-puckered structure for (4*S*)-L-FPro, both as free amino acids. By analyzing the acetyl methyl ester derivatives **1**–**4** in the present study, the $^1H-^1H$ and $^1H-^{19}F$ coupling patterns were indicative of a *γ-endo*-puckered structure for all compounds except for Ac-(4*R*)-L-Pro-OMe (**2**) which exhibits a *γ-exo*-pucker, and this independently of the *trans* or *cis* conformation of the Ac–Pro amide bond. The crystal structures of (4*R*)-L-Pro and L-Pro as free amino acids exhibit the same pucker as determined by NMR analysis, that is a *γ-exo*-pucker for (4*R*)-L-FPro and a *γ-endo*-pucker for Pro.[8] The *γ-exo*-pucker of the proline ring for (4*R*)-L-FPro and the *γ-endo*-pucker for (4*S*)-L-FPro is explained by the *gauche* effect.[4] For the difluoro derivative the *gauche* effects of both fluorine atoms counteract leading to behavior more similar to that of proline, namely a preference for the *endo*-pucker. Molecular modeling experiments show that due to steric hindrance, the energy difference between the *cis* and *trans* isomer is significantly larger for the *exo*-pucker than for the *endo*-pucker. This would suggest that a preference for the *endo*-pucker leads to a smaller energy difference between the *cis* and *trans* isomers and, therefore, to a larger population of the *cis* isomer state at equilibrium as indeed experimentally observed for (4*S*)-L-FPro in comparison to (4*R*)-L-FPro. These results fully agree with a recent Brookhaven data base analysis of peptide bonds in proteins[9] which showed that *cis* Xaa–Pro bonds are accompanied by an *endo*-pucker of the proline, whereas for *trans* Xaa–Pro bonds no preferred pucker was observed.

To correlate the thermodynamic data derived from the model compounds with the thermostability of 4-FPro-mutants of proteins, as a model protein we have chosen barstar, a single-domain extracellular RNAse inhibitor protein which contains two proline residues, that is Pro 27 in the *trans* and Pro 48 in the *cis* conformation. For the substitution experiments the barstar mutant b*C40A/C82A/P27A was used which lacks complications arising from oxidation of cysteine residues in the wild-type form and from isomerization affecting the peptidyl–Pro 27 bond. This triple mutant protein contains only one proline residue (Pro 48) being in *cis* conformation in the native state, located in a loop region almost completely exposed to the solvent.[10] The fluorinated prolines (4*R*)-L-FPro and (4*S*)-L-FPro were incorporated into the barstar (b*C40A/C82A/P27A) at the sequence position Pro 48 in a response to a CCG codon-containing DNA-template (plasmid) during fermentation with the proline-auxotrophic *E. coli* host strain DSM 1562 under efficient selective pressure;[11, 12] to our knowledge this is the first incorporation of fluorinated prolines into proteins through a bacterial expression system. The yields of mutants were comparable to those of the nonfluorinated barstar mutant (Figure 3). As expected, the (4*S*)-L-FPro mutant showed a higher thermostability ($T_m = 69.4\,^\circ C$) and the (4*R*)-L-FPro mutant a lower stability ($T_m = 61.5\,^\circ C$) than barstar b*C40A/C82A/P27A ($T_m = 64.3\,^\circ C$). In preliminary experiments difluoroproline was also incorporated into barstar, but without marked change in stability ($T_m = 64.2\,^\circ C$); the detailed biophysical characterization of all barstar mutants will be reported elsewhere.

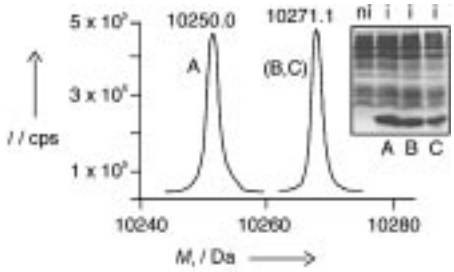

Figure 3. Left: Deconvoluted ESI-MS spectra of two separate measurements of barstar b*C40A/C82A/P27A (A; $M_r = 10250$ Da) and of the related P48(4*R*)-L-FPro (B) and P48(4*S*)-L-FPro (C) mutants, $M_r = 10271$ Da. Right: Expression profile of barstar b*C40A/C82A/P27A and its FPro-mutants: ni = noninduced cells (control); i = induced cells.

The successful expression of the barstar mutants confirms that fluorinated prolines are recognized by their cognate aminoacyl-tRNA synthetases and are incorporated into proteins as efficiently as fluorinated tyrosines,[13] tryptophans,[14] and phenylalanines.[15] Moreover, depending upon the (4*S*)- or (4*R*)-FPro isomer used, the *cis* or *trans* conformation of peptidyl–Pro bonds can be favored and disfavored, respectively, relative to proline itself, even in proteins. Conversely, bioincorporation of the difluoroproline into proteins is expected to yield mutants with unchanged thermostabilities, but bearing the valuable fluorine label for NMR structural analysis. The modulation of protein properties by synthetic amino acids with a priori defined properties,

  1433-7851/01/4005-0924 $ 17.50+.50/0

as demonstrated here, opens a rather new avenue to the rational design and engineering of proteins.

Experimental Section

The compounds Ac-(4*R*)-L-FPro-OMe (**2**), Ac-(4*S*)-L-FPro-OMe (**3**), and Ac-(4)-L-F_2Pro-OMe (**4**) were prepared according to the method developed by Demange et al.,[16] and Ac-Pro-OMe (**1**) by standard procedures. FT-IR spectra were recorded at 25 °C on a Perkin-Elmer 1760 X FT-IR spectrometer using an ATR cell (ATR = attenuated total reflection). For thermal unfolding experiments a JASCO J-715 spectropolarimeter equipped with a Peltier-type FDCD attachment, model PFD-350S/350L was employed and samples were measured in rectangular 110-QS Hellma quartz cells with an optical path length of 0.1 cm. The melting curves of wild-type as well as substituted proteins were acquired in the range 25–95 °C in 50 mM sodium dihydrogen phosphate with 20 mM Tris/HCl buffer (Tris = tris(hydroxymethyl)aminomethane) at pH 8.0. The unfolding profile was recorded at 222 nm at a protein concentration of 0.2 mg mL^{-1} using a temperature gradient of 50 °C h^{-1} as described elsewhere. Samples for NMR measurements were prepared by dissolving the compounds (3 mg) in D_2O (0.5 mL). All NMR experiments were performed on a Bruker DRX500 spectrometer using either a triple-resonance ($^{15}N/^{13}C/^{1}H$) inverse or a dual ($^{19}F/^{1}H$) inverse probe. ^{1}H, ^{13}C, and ^{19}F NMR signals were assigned based on homonuclear 2D $^{1}H-^{1}H$ NOESY and TOCSY experiments and heteronuclear 2D $^{19}F-^{1}H$ and $^{13}C-^{1}H$ COSY experiments.[17]

Thermodynamics and kinetics of amide bond isomerization: Equilibrium constants (K_{ZE}) for the *trans*/*cis* isomer ratios at various temperatures were determined by integration of signals in the 1D-^{1}H and 1D-^{13}C NMR spectrum for compounds **1**–**3**, and in the 1D-^{19}F NMR spectrum for **4**. For the 4-FPro derivatives ^{19}F NMR signals of the *cis* and *trans* isomer overlapped to such an extent that neither determination of the equilibrium constants nor measurement of the isomerization rates was possible by ^{19}F NMR spectroscopy. Comparison of K_{ZE} values obtained from 1D-^{1}H and 1D-^{13}C NMR spectra revealed that resonances from γ and δ protons were the most suitable for integration. The enthalpic and entropic contributions to the free energy difference between the *cis* and *trans* isomers were obtained from van't Hoff plots according to $\ln(K_{ZE}) = (-\Delta H^0/R)(1/T) + \Delta S^0/R$. The rate constants k_{EZ} for the *cis* $\rightarrow$ *trans* and k_{ZE} for *trans* $\rightarrow$ *cis* isomerization of the Ac–Pro amide bond were determined from 2D NOESY spectra. For the (4)-F_2Pro derivative **4** cross peaks between $^{19}F_{\gamma}^{trans}$ and $^{19}F_{\gamma}^{cis}$ in the $^{19}F-^{19}F$ NOESY spectrum were integrated and compared to integrals of the corresponding diagonal signals. To avoid the influence of differential T_1 times, only cross and diagonal peaks with the same F_1 frequency (indirect) and not with the same F_2 (direct) frequency were compared. For compounds **1**–**3** cross peaks between $^{1}H_{\alpha}^{trans}$ and $^{1}H_{\alpha}^{cis}$ in the $^{1}H-^{1}H$ NOESY spectrum were used since only these were separated well enough in the spectra. The cross peak to diagonal peak ratios thus obtained were scaled as follows to account for differential relaxation behavior of the resonances:[18] $k_{EZ} = (I_{ct}/I_{cc} + (I_{tc}/I_{tt})\,K_{ZE})/(2\tau_m)$, $k_{ZE} = ((I_{ct}/I_{cc})/K_{ZE} + I_{tc}/I_{tt})/(2\tau_m)$. I_{xx} denotes the peak integrals from the 2D NOESY spectra with the first subscript indicating the F_1 frequency (either *cis* or *trans*) and the second the F_2 frequency. At each temperature the NOESY mixing times τ_m were chosen such that $I_{ct} \ll I_{cc}$ and $I_{tc} \ll I_{tt}$. Activation enthalpy and entropy for the *cis* $\rightarrow$ *trans* and *trans* $\rightarrow$ *cis* isomerization were obtained from Eyring plots according to $\ln(k/T) = (-\Delta H^{\neq}/R)(1/T) + \Delta S^{\neq}/R + \ln(k_B/h)$.

Conformational analysis: The pucker of the proline ring was identified in analogy to the work by Gerig and McLeod[7] by the distinct pattern of the $^{1}H-^{1}H$ and $^{1}H-^{19}F$ coupling constants observed in 1D-^{1}H and 1D-^{19}F NMR spectra; for example, the γ-*exo*-pucker results in large and similar coupling constants $J(H_{\alpha},H_{\beta 1})$ and $J(H_{\alpha},H_{\beta 2})$, whereas for the *endo*-pucker one of the coupling constants is large and the other is small. Similarly, also other couplings show characteristic patterns for the *endo*- and *exo*-pucker, respectively.

Received: September 5, 2000 [Z 15760]

[1] J. T. Gerig, *Prog. NMR Spectrosc.* **1994**, *26*, 293–370.
[2] D. B. Harper, D. O'hagen, *Nat. Prod. Rep.* **1994**, *11*, 123–133.
[3] J. Uitto, D. J. Prockop, *Biochim. Biophys. Acta* **1974**, *336*, 234–251.
[4] S. K. Holmgren, L. E. Bretscher, K. M. Taylor, R. T. Raines, *Chem. Biol.* **1999**, *6*, 63–79.
[5] E. S. Eberhardt, N. Panasik, Jr., R. T. Raines, *J. Am. Chem. Soc.* **1996**, *118*, 12261–12266.
[6] C. Grathwohl, K. Wüthrich, *Biopolymers* **1976**, *15*, 2025–2041.
[7] J. T. Gerig, R. S. McLeod, *J. Am. Chem. Soc.* **1973**, *95*, 5725–5729.
[8] N. Panasik, Jr., E. S. Eberhardt, A. S. Edison, D. R. Powell, R. T. Raines, *Int. J. Pept. Protein Res.* **1994**, *44*, 262–269.
[9] D. Pal, P. Chakrabarti, *J. Mol. Biol.* **1999**, *294*, 271–288.
[10] R. Golbik, G. Fisher, A. R. Fersht, *Protein Sci.* **1999**, *8*, 1505–1514.
[11] N. Budisa, C. Minks, F. J. Medrano, J. Lutz, R. Huber, L. Moroder, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 455–459.
[12] N. Budisa, C. Minks, S. Alefelder, W. Wenger, F. Dong, L. Moroder, R. Huber, *FASEB J.* **1999**, *13*, 41–51.
[13] M. Ring, I. M. Armitage, R. E. Huber, *Biochem. Biophys. Res. Commun.* **1985**, *131*, 675–680.
[14] C. Minks, R. Huber, L. Moroder, N. Budisa, *Biochemistry* **1999**, *38*, 10649–10659.
[15] C. Minks, R. Huber, L. Moroder, N. Budisa, *Anal. Biochem.* **2000**, *284*, 29–34.
[16] L. Demange, A. Ménez, C. Dugave, *Tetrahedron Lett.* **1988**, *39*, 1169–1172.
[17] K. Wüthrich, *NMR of Proteins and Nucleic Acids*, Wiley, New York, **1986**.
[18] S. Macura, B. T. Farmer II, L. R. Brown, *J. Magn. Reson.* **1986**, *70*, 493–499.

  1433-7851/01/4005-0925 $ 17.50+.50/0