CROATICA CHEMICA ACTA CCACAA **72** (2–3) 259–277 (1999)

ISSN-0011-1643
CCA-2576 *Original Scientific Paper*

Conformational Studies in Solid State and Solution of Protected C-terminal Dipeptide Fragment (Boc-Phe-Pro-NH_2) of Morphiceptin*

Biserka Kojić-Prodić,[a,**] *Snježana Antolić,*[a] *Marina Kveder,*[a] *Ivanka Žigrović,*[a] *Jurka Kidrič,*[b] *and Štefica Horvat*[a]

[a]*Ruđer Bošković Institute, P.O.B. 1016, HR-10001 Zagreb, Croatia*

[b]*National Institute of Chemistry, P.O.B. 30, 61015 Ljubljana, Slovenia*

Received September 14, 1998; revised January 8, 1999; accepted January 20, 1999

The crystal structure of the protected C-terminal dipeptide fragment (Boc-Phe-Pro-NH_2) of the μ-opioid receptor highly selective agonist, morphiceptin (Tyr-Pro-Phe-Pro-NH_2), was determined; the crystals are monoclinic with space group $P2_1$ and unit cell dimensions: a = 11.5731(5), b = 6.4490(3), c = 15.4082(5) Å, β = 100.359(5)° and Z = 2. To examine the influence of proline on the conformation of peptide bond, the molecular conformation was studied in solid state and solution (using ^{1}H and ^{13}C NMR data). The X-ray analysis revealed the following conformations of peptide backbone: ϕ_1 = –63.2(5)°, ψ_1 = 156.1(4)°, ω_1 = –174.3(4)°, ϕ_2 = –66.0(5)° and ψ_2 = 152.0(4)°. The conformation of the Boc group is *trans-trans*. Experimental data revealed the *trans* conformation about the Phe-Pro amide bond, both in solid state and solution (DMSO). The possibility of *cis*/*trans* isomerization about the peptide bond (ω_1) was examined by theoretical calculations using BIOSYM software. Molecular modelling, including molecular dynamics simulations of the title dipeptide, is in favour of *trans* peptide bond.

Key words: C-terminal dipeptide, *cis*/*trans* isomerization, morphiceptin, conformational analysis, X-ray structure, NMR, molecular dynamics simulations.

---

* Dedicated to Professor Boris Kamenar on the occasion of his 70th birthday.

** Author to whom correspondence should be addressed. (E-mail: kojic@rudjer.irb.hr)

INTRODUCTION

Proline plays a unique role in peptide and protein structures having a capacity to affect local polypeptide environment, where it often participates in the formation of loops, turns and the polyproline helix.[1–3] While most peptide bonds exist in the *trans* geometry, those involving proline residue (Xaa-Pro) display a tendency to assuming both *cis* and *trans* conformations due to the small difference in free energy between these two conformers.[3–5] It has been suggested that *cis*/*trans* proline isomerization plays many important biochemical roles, *e.g.* controlling the rate of protein folding, triggering receptor-mediated transmembrane signalling, and providing a recognition element in peptide antigens.[6–8] Due to the unique function of proline in folding and refolding of proteins, *cis*/*trans* isomerization has been studied from various aspects.[3,5,9] The recent spectroscopic studies and molecular mechanics and dynamics calculations on model peptides[4,5,10–12] demonstrated that *cis*/*trans* equilibria depend on the nature of amino acid residues flanking the Pro-peptide unit and on the ionization state of terminal groups. Therefore, the equilibrium should be determined for each Xaa-Pro sequence individually. MacArthur and Thornton[3] performed extensive analysis using the Protein Data Bank[13], which showed strong sequence preferences for *cis versus trans* prolines.

Protected C-terminal dipeptide (Boc-Phe-Pro-NH_2) is a constitutional fragment of morphiceptin (Tyr-Pro-Phe-Pro-NH_2), one of the most selective agonists for μ-opioid receptor discovered in bovine milk proteins.[14,15] Spectroscopic studies of morphiceptin revealed the presence of four configurational isomers (*trans*/*trans, cis*/*trans, trans*/*cis,* and *cis*/*cis*) both in water and dimethyl sulphoxide.[16] According to ^{1}H and ^{13}C NMR data, the largest proportion (amounting to 55%) has been assigned as the all-*trans* isomer. The second configurational isomer, accounting for 25%, adopts a *cis* conformation around the Tyr-Pro amide bond. A topochemical model developed to explain the bioactivity of morphiceptin suggested the requirement of a *cis* amide bond linking N-terminal Tyr-Pro residues.[17] In terms of 'bioactive conformations', the C-terminal address sequence may stabilize a specific conformation among various conformations accessible to the N-terminal message sequence and, for full understanding of the opioid activity, it is necessary to identify the specific conformational states for rotatable bonds present in this part of the opioid peptide molecule.

This work has focused on the comparative conformational analysis of Boc-Phe-Pro-NH_2, based on X-ray crystallographic studies, NMR spectroscopic data and molecular dynamics simulations.

EXPERIMENTAL

Dipeptide Synthesis

Boc-Phe-Pro-NH_2 was synthesized by the conventional liquid-phase method. Amidation of the Boc-Pro-OH[18] was carried out using mixed anhydride activation followed by treatment with conc. ammonia solution. Boc protection was removed with 90% TFA and the obtained product was coupled with Boc-Phe-OH by the mixed anhydride method to yield Boc-Phe-Pro-NH_2 in 60% yield; m.p. 326–329 K, $[\alpha]_D$ = –42.5° (c = 1, MeOH), Lit.[19] m.p. 328–331 K.

NMR Experimental Details

NMR spectra were acquired on Varian spectrometers VXR-300, Unity $^+$300 and INOVA 600 at 298.15 K. 10 mmol dm^{-3} solution in DMSO-d_6 was used. Chemical shifts are reported in ppm relative to tetramethylsilane as internal standard. Data were processed using Varian VNMR software. COSY, NOESY and ROESY were acquired with 512 increments and 8 (at 600 MHz) or 32 (at 300 MHz) transients. For NOESY and ROESY experiments, mixing times of either 250 or 500 ms with different transmitter offsets were used. Heteronuclear multiple-quantum coherence (HMQC with 512 increments and 32 transients) experiments were used for the detection of ^{13}C–^1H correlations. 1D ^{13}C NMR spectrum was acquired at 75.4 MHz with 103000 transients.

X-ray Structure Determination

The peptide was crystallized from the solvent mixture ethanol/diisopropyl ether (1 : 1 v/v) by slow evaporation at room temperature for 1 day. The crystal data and details of data collection and refinement are listed in Table I. The X-ray intensity data were collected with an Enraf-Nonius CAD4 diffractometer with graphite-monochromatized Cu-Kα radiation. There were no significant variations in intensity for standard reflections. The data were corrected for Lorentz and polarization effects.[20] The structure was determined by direct methods using the SHELX86 program.[21] Refinement was performed by full-matrix least-squares with the SHELX93[22] system of programs on F^2 values. Most hydrogen atoms were calculated on stereochemical grounds and refined riding on their respective C atoms. Atomic scattering factors were those included in SHELX93.[22] Details of the refinement procedure are given in Table I. During structure determination, the enantiomer with the absolute configuration S on C_1^α and C_2^α according to the synthesis procedure, was selected. The solvent molecule (ethanol) was located in a Fourier map. The molecular geometry was calculated by the EUCLID program.[23] Drawings were prepared by the PLUTON program incorporated in EUCLID[23] and ORTEP II.[24] The final atomic coordinates and equivalent isotropic thermal parameters are listed in Table II.

Molecular Mechanics and Molecular Dynamics Simulations

In order to examine the possibility of *cis*/*trans* imide isomerization of *N-tert*-butyloxycarbonyl-L-phenylalanyl-L-proline amide and to compare the conformations in solid state and solution, molecular mechanics and dynamics calculations were carried out using DISCOVER[25] (BIOSYM, ver. 2.9.7 & 95.0 / 3.00, 1995) with the AMBER[26] force field.

TABLE I

General and crystal data and details of the diffraction procedure and structure determination

Crystal data	
Empirical formula	$C_{19}N_3O_4H_{27} \cdot C_2H_6O$
Molecular weight, M_r	407.51
a / Å	11.5731(5)
b / Å	6.4490(3)
c / Å	15.4082(5)
β / °	100.359(5)
V / Å^3	1131.25(8)
Unit cell contents, Z	2
Dc / g cm^{-3}	1.196
F (000)	440
Crystal system	monoclinic
Space group	$P2_1$
Crystal size / mm	0.11 × 0.07 × 0.34
μ (Cu-Kα) / cm^{-1}	6.98
Data collection	
Diffractometer	Enraf-Nonius CAD4
Radiation / Å	λ(Cu-Kα) = 1.54184 graphite monochromator
Temperature / K	295(3)
θ_{min} , θ_{max} / ° for cell det.	12, 47
No. of reflections used for cell det.	25
θ_{min} , θ_{max} / °	2.92, 74.36
(ω/2θ scan) / °	$\Delta\omega$ = 0.62 + 0.26 tanθ
h, k, l limits	0, 14; 0, 8; –19, 19
Refinement	
Independent reflections observed with $I > 4\sigma(I)$	1736
No. of parameters	309
Refinement on F^2	
$w = 1 / [\sigma^2(F_o^2) + (0.1037 \cdot P)^2]$, $P = (F_o^2 + 2F_c^2) / 3$	
$R(F)$, $wR(F^2)$	0.053, 0.156
Goodness of fit, S	0.974
Max. shift / error $(\Delta / \sigma)_{max}$	< 0.05
Residual electron density, $\Delta\rho_{max}$, $\Delta\rho_{min}$ / e Å^{-3}	0.26, –0.22

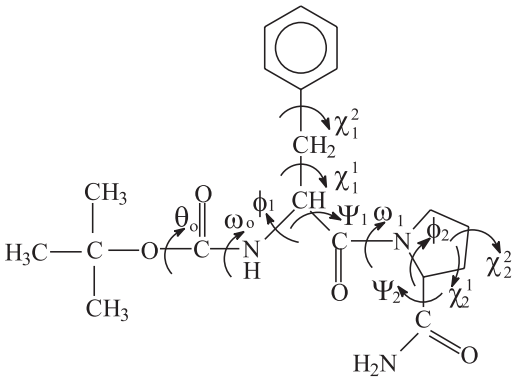
TABLE II

Final atomic coordinates and equivalent isotropic thermal parameters of the nonhydrogen atoms with estimated standard deviations given in parentheses

Atom	*x*	*y*	*z*	U_{eq} / Å^2
O	0.8987(2)	0.2419(7)	0.2082(1)	0.051(1)
O_1	0.5714(2)	0.1860(7)	0.1167(1)	0.0505(9)
O_2	0.2874(3)	0.1576(8)	0.0847(2)	0.066(1)
O_0	0.8151(3)	–0.0773	0.1935(2)	0.051(1)
N_1	0.7403(2)	0.1860(7)	0.2645(1)	0.0407(9)
N_2	0.4802(3)	–0.0906(7)	0.1593(1)	0.0406(9)
N_3	0.3106(3)	0.062(1)	–0.0522(2)	0.062(1)
C1	0.9650(5)	0.116(1)	0.0753(3)	0.082(2)
C_1'	0.5583(3)	0.0618(8)	0.1737(2)	0.035(1)
C2	1.0622(5)	0.400(1)	0.1674(4)	0.077(2)
C_2'	0.3309(3)	0.0465(9)	0.0359(2)	0.044(1)
C3	1.0785(5)	0.039(2)	0.2264(4)	0.097(3)
C4	1.0028(3)	0.188(1)	0.1679(2)	0.049(1)
C5	0.8187(3)	0.0995(8)	0.2194(2)	0.039(1)
C_1^{α}	0.6315(3)	0.0791(8)	0.2675(2)	0.034(1)
C_1^{β}	0.5611(3)	0.1950(9)	0.3289(2)	0.044(1)
C_1^{γ}	0.6173(3)	0.1820(9)	0.4249(2)	0.044(1)
$C_1^{\delta 2}$	0.6755(5)	0.349(1)	0.4675(3)	0.067(2)
$C_1^{\varepsilon 2}$	0.7283(6)	0.336(2)	0.5551(3)	0.091(3)
C^{ζ}	0.7220(5)	0.157(2)	0.6014(3)	0.089(3)
$C_1^{\varepsilon 1}$	0.6627(6)	–0.011(2)	0.5608(3)	0.090(3)
$C_1^{\delta 1}$	0.6097(5)	0.001(1)	0.4720(3)	0.068(2)
C_2^{α}	0.4145(3)	–0.1313(8)	0.0705(2)	0.043(1)
C_2^{β}	0.3469(6)	–0.329(1)	0.0803(3)	0.085(2)
C_2^{γ}	0.3765(6)	–0.394(1)	0.1688(3)	0.092(2)
C_2^{δ}	0.4585(4)	–0.2521(9)	0.2211(2)	0.060(2)
O6[a]	0.2112(3)	0.1166(7)	0.6865(2)	0.058(1)
C6[a]	0.0995(5)	0.105(1)	0.6294(3)	0.077(2)
C7[a]	0.0790(7)	0.280(2)	0.5673(4)	0.110(3)

[a] Ethanol.

[b] $U_{eq} = (1 / 3)\ \Sigma_i\Sigma_j U_{ij}\ a_i^* a_j^* \mathbf{a}_i.\mathbf{a}_j$.



Scheme 1. Torsion angle assignments used in the conformational analysis.

The following torsion angles of the peptide backbone were chosen for conformational search: ω_0, ϕ_1, ψ_1, ω_1, ϕ_2 and ψ_2 (Scheme 1). The structural flexibility of the Phe was examined by inspection of its side chain torsion angles χ_1^1 and χ_1^2.

Energy optimization of the crystallographically determined structure was performed in vacuo without constraint. The results are summarized in Table III. In order to examine the possibility of *cis*-proline orientation, simultaneous rotations about the torsion angles N_1–C_1^α–C_1^β–C_1^γ (χ_1^1) and C_1^α–C_1'–N_2–C_2^α (ω_1) in 24 steps (15° each) were calculated. To simulate the conditions of NMR measurements, molecular dynamics simulations were performed in DMSO with AMBER[26] force field using periodic boundary conditions. The explicit image model applying two values of the cut-off parameters (14 and 15 Å) was used with a cube (a = 32 Å) as a unit cell. The distance dependent dielectric constant was chosen during the molecular dynamics calculations. The sterically relaxed X-ray structure was soaked in the DMSO box[47] filled with about 210 molecules of DMSO. At the beginning, the entire system was subjected to minimization that included a combination of steepest descent and conjugate gradient method and in the final stage the BFGS (Broydon-Fletcher-Goldfarb-Shanno)[25] algorithm (gradient norm < 10^{-3} kcal / mol Å). After 5 ps of equilibration, the simulation over 320 ps was carried out at 300 K. To sample the larger conformational space, the simulation over 320 ps was also performed at elevated temperatures (300 K, 350 K, 400 K over 50 ps each, 450 K over 100 ps and again 300 K over 70 ps). In order to compare the influence of solvent on the conformational flexibility with respect to molecular dynamics calculations performed in DMSO, molecular dynamics simulations were also performed in vacuo.

RESULTS AND DISCUSSION

NMR Studies, Molecular Conformation in Solution

^{1}H and ^{13}C resonances were assigned using 1D and 2D experiments mentioned in the Experimental section. Experimental data were compared with those obtained for Boc-Phe-OH, H-Pro-NH_2 as well as with proton

TABLE III

Comparative conformational analysis based on values obtained from the X-ray structure analysis, molecular mechanics and molecular dynamics simulations

Torsion angles / °	ω_0	ϕ_1	ψ_1	χ_1^1	χ_1^2	ω_1	ϕ_2	ψ_2	χ_2^1	χ_2^2
X-ray data	166.1(3)	–63.2(5)	156.1(4)	–69.5(5)	–75.9(5)	–174.3(4)	–66.0(5)	152.0(4)	–1.1(6)	0.8(8)
Molecular mechanics calculations										
Molecule optimized in vacuo without constraint	169	–64	121	–62	79	–176	75	64	34	39
Molecular dynamics simulations										
in DMSO at 300 K predominant conformers (± 20°)	180	–120, –90	150	–60, 60	110	180	–60	–45, 150	–30, 30	–30, 30

[a] Calculations were performed by DISCOVER (AMBER).

chemical shifts of random-coil peptides.[27] The diastereotopic β protons of Phe were assigned *via* αH–βH couplings and NOEs with Pro-C_2^{δ} protons. Chemical shifts and coupling constants are given in Tables IV and V. Although *cis* and *trans* conformations across the X-Pro amide bond are often similar in energy, and their interconversion is slow on the NMR time scale,[5,10–12,16] only one set of resonances in ^{13}C NMR spectrum was observed for dipeptide Boc-Phe-Pro-NH_2, suggesting the presence of one isomer in DMSO-d_6 solution (Figure 1). On the basis of ^{13}C chemical shift resonances of the proline ring carbons, reflecting the configuration about X-Pro amide bond,[28] Phe-Pro bond in *trans* conformation was found in solution. This assumption was borne out by the small ^{13}C chemical shift difference, $\Delta\delta$ = 4.66 ppm, observed between Pro-C_2^{β} and Pro-C_2^{γ} carbons (Table IV) characteristic of *trans* amide linkages, while larger differences ($\Delta\delta$ = 8–10 ppm) are indicative of *cis*-proline linkages.[28] This is also supported by: a) large downfield shifts of proton resonances of Phe-C_1^{α} and Pro-C_2^{α} (4.34 ppm and 4.24 ppm, respectively), which are very sensitive to *cis* ↔ *trans* isomerization due to the influence of the anisotropic proline carbonyl group;[12,16,29–32] b) splitting of Phe-C_1^{β} protons (2.73 and 2.98 ppm);[29] c) strong NOE between the Phe-C_1^{α} and Pro-C_2^{δ} protons.

TABLE IV

^{13}C and ^{1}H chemical shifts (δ/ppm)

Residue	δ_C	δ_H
Phe-N_1		7.04
Phe-C_1^{α}	53.69	4.34
Phe-C_1^{β}, $C_1^{\beta'}$	36.16	2.73, 2.98
Phe-C_1^{γ}	138.10	
Phe-$C_1^{\delta 1}$, $C_1^{\delta 2}$	129.21	7.20
Phe-$C_1^{\varepsilon 1}$, $C_1^{\varepsilon 2}$	127.94	7.29
Phe-C^{ζ}	126.08	7.29
Phe-C_1'	170.26	
Pro-C_2^{α}	59.48	4.24
Pro-C_2^{β}, $C_1^{\beta'}$	29.04	1.81, 2.03
Pro-C_2^{γ}	24.38	1.90
Pro-C_2^{δ}	46.51	3.62
Pro-N_3		6.91, 7.13
Pro-C_2'	173.47	
Boc-CH_3	28.04	1.28
Boc-C5	155.19	
Boc-C4	77.84	

TABLE V

Experimental values of coupling constants ^3J for the conformation detected in solution (DMSO-d_6) and derived from the solid state (X-ray structure) and calculated values from molecular dynamics simulations

Residue	Coupling constants / Hz			Refs.
	Experimental		Calculated	
	NMR	X-ray	Molecular dynamics simulations	
Phe-N_1–Phe-C_1^α	8.4	3.14	[8 ± 3]	42
Phe-C_1^α–Phe-C_1^β	3.6	0.4	[2 ± 1]	43
Phe-C_1^α–Phe-$C_1^{\beta'}$	10.2	8.7	[8.7 ± 0.8]	43
Pro-C_2^α–Pro-C_2^β	3.6	2.7	[1.2 ± 0.6]	44
Pro-C_2^α–Pro-$C_2^{\beta'}$	8.4	8.2	[7 ± 2]	43

Calculations of coupling constants were performed for each time step of the trajectory and subsequently averaged[45] whereas the equations are used as given in Refs. 42–44.

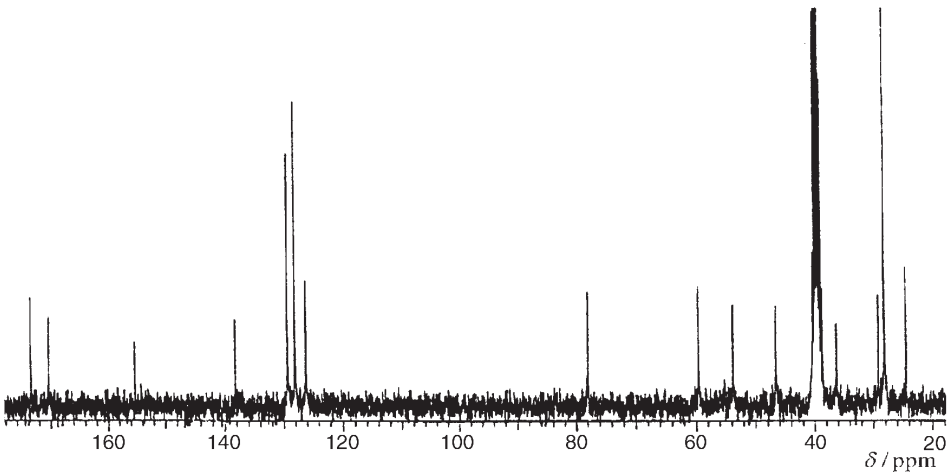


Figure 1. ^{13}C NMR spectrum of Boc-Phe-Pro-NH_2 in DMSO-d_6.

The relative populations of staggered conformations about Phe-C_1^α–Phe-C_1^β bond, characterized by torsional angle χ_1^1,[33,34] were calculated using Phe-C_1^α–Phe-C_1^β,$C_1^{\beta'}$ coupling constants. The calculated conformation defines the following rotamers: $\chi_1^1 = -60°$, predominant (populated 69%), $\chi_1^1 = 60°$, (po-

pulated 23%) and $\chi_1^1 = 180°$, (populated 8%). For the rotamer with $\chi_1^1 = -60°$, the orientation of C_1^α / $C_1^{\beta\,proR}$ protons is *trans* and this finding is in agreement with much stronger NOE between Pro-C_2^δ and Phe-$C_1^{\beta\,proS}$ protons than that between Pro-C_2^δ and Phe-$C_1^{\beta\,proR}$ protons which are further apart.

The analysis of NMR data shows that Boc-Phe-Pro-NH_2 prefers *trans* orientation (ω_1) in DMSO-d_6 solution, which together with the dominating rotamer ($\chi_1^1 = -60°$) about the C_1^α–C_1^β bond indicate the high conformational constraints in this part of the molecule.

X-ray Analysis

Molecular Structure

The ORTEP drawing[24] with thermal ellipsoids scaled at 30% probability level and atom numbering[35] is shown in Figure 2. It includes a molecule of solvent – ethanol. Selected bond lengths and angles are listed in Table VI. In the crystal structure studied, the peptide bond is planar: the mean value of deviation from planarity is 0.009(15) Å. Pyrrolidine ring is also planar

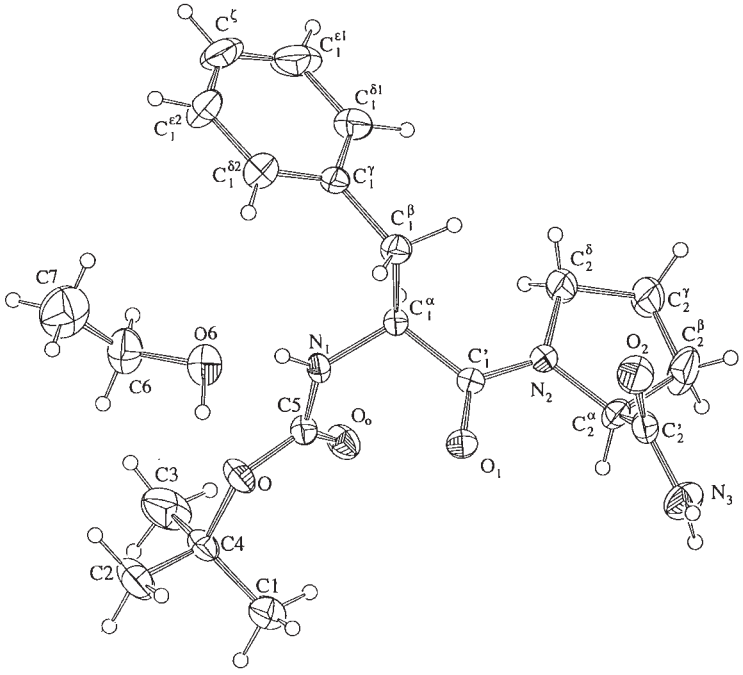


Figure 2. Molecular structure (ORTEP drawing) with atom numbering. Thermal ellipsoids are scaled at 30% probability level. Atom labeling is according to Ref. 35.

TABLE VI

Interatomic bond distances and valence bond angles with estimated standard deviations given in parentheses

Bond lengths / Å		Bond angles / °	
O–C4	1.492(4)	C4–O–C5	121.7(4)
O–C5	1.337(6)	C5–N_1–C_1^α	119.1(4)
O_1–C_1'	1.218(5)	C_1'–N_2–C_2^α	120.9(3)
O_2–C_2'	1.212(6)	C_1'–N_2–C_2^δ	127.6(3)
O_0–C5	1.206(5)	C_2^α–N_2–C_2^δ	111.1(4)
N_1–C5	1.358(5)	O_1–C_1'–N_2	122.4(3)
N_1–C_1^α	1.444(5)	O_1–C_1'–C_1^α	120.6(4)
N_2–C_1'	1.327(6)	N_2–C_1'–C_1^α	117.0(3)
N_2–C_2^α	1.464(4)	O_2–C_2'–N_3	124.6(5)
N_2–C_2^δ	1.463(6)	O_2–C_2'–C_2^α	122.2(3)
N_3–C_2'	1.339(4)	N_3–C_2'–C_2^α	113.2(4)
C1–C4	1.490(6)	O–C4–C1	110.5(3)
C_1'–C_1^α	1.542(4)	O–C4–C2	101.1(5)
C2–C4	1.53(1)	O–C4–C3	109.2(3)
C_2'–C_2^α	1.533(7)	C1–C4–C2	109.3(5)
C3–C4	1.491(9)	C1–C4–C3	114.6(6)
C_1^α–C_1^β	1.547(6)	C2–C4–C3	111.3(4)
C_1^β–C_1^γ	1.507(4)	O–C5–O_0	126.2(4)
C_2^α–C_2^β	1.520(8)	O–C5–N_1	109.0(4)
C_2^β–C_2^γ	1.407(7)	O_o–C5–N_1	124.8(4)
C_2^γ–C_2^δ	1.451(8)	N_1–C_1^α–C_1'	109.9(3)
O6[a]–C6[a]	1.429(6)	N_1–C_1^α–C_1^β	110.3(4)
C6[a]–C7[a]	1.47(1)	C_1'–C_1^α–C_1^β	110.3(3)
C–C in phenyl ring	<1.380(9)>	C_1^α–C_1^β–C_1^γ	113.0(3)
		C_1^β–C_1^γ–$C_1^{\delta 2}$	120.9(5)
		C_1^β–C_1^γ–$C_1^{\delta 1}$	120.3(5)
		N2–C_2^α–C_2'	112.3(4)
		N_2–C_2^α–C_2^β	104.4(3)
		C_2'–C_2^α–C_2^β	111.1(4)
		C_2^α–C_2^β–C_2^γ	107.7(5)
		C_2^β–C_2^γ–C_2^δ	112.0(6)
		N_2–C_2^δ–C_2^γ	104.9(3)
		O6[a]–C6[a]–C7[a]	112.6(6)
		C–C–C in phenyl ring	<120.0(7)>

[a] Ethanol.

within the limits of experimental errors: the mean torsion angle value is 0.7(6)° (Table VII). The mean value of atom displacement from the best least-squares plane defined by pyrrolidine ring is 0.004(16) Å. The planarity of the pyrrolidine ring is also illustrated by the torsion angles χ_2^1 [–1.1(6)°] and χ_2^2 [0.8(8)°]. According to the literature,[36,37] puckered conformations of pyrrolidine rings of proline were detected. For *trans* conformation of L-proline in a peptide backbone, the pyrrolidine puckering of C$^\gamma$-*endo* and C$^\gamma$-*exo* were observed with a preference of C$^\gamma$-*endo*.[37,38] However, proline in the *cis* conformation exhibits a pyrrolidine ring in the C$^\gamma$-*exo* conformation only. In the title dipeptide, proline is a C-terminal amino acid with almost planar pyrrolidine ring.

TABLE VII

Selected torsion angles / °
with estimated standard deviations given in parentheses

θ_0	C4–O–C5–N_1	174.4(2)
ω_0	C_1^α–N_1–C5–O	166.1(3)
ϕ_1	C5–N_1–C_1^α–C_1'	–63.2(5)
	C5–N_1–C_1^α–C_1^β	175.0(3)
	C_1'–N_2–C_2^δ–C_2^γ	–172.5(5)
	C_2^α–N_2–C_2^δ–C_2^γ	–0.6(6)
	C_2^δ–N_2–C_2^α–C_2^β	1.1(5)
ω_1	C_2^α–N_2–C_1'–C_1^α	–174.3(4)
	C_2^δ–N_2–C_1'–C_1^α	–3.1(6)
	C_1'–N_2–C_2^α–C_2^β	173.5(4)
	C_2^δ–N_2–C_2^α–C_2'	121.5(4)
ϕ_2	C_1'–N_2–C_2^α–C_2'	–66.0(5)
ψ_1	N_2–C_1'–C_1^α–N_1	156.1(4)
	N_2–C_1'–C_1^α–C_1^β	–82.1(5)
χ_1^1	N_1–C_1^α–C_1^β–C_1^γ	–69.5(5)
χ_1^2	C_1^α–C_1^β–C_1^γ–$C_1^{\delta1}$	–75.9(5)
	C_1^α–C_1^β–C_1^γ–$C_2^{\delta2}$	104.9(6)
	C_1'–C_1^α–C_1^β–C_1^γ	169.0(4)
χ_2^1	N_2–C_2^α–C_2^β–C_2^γ	–1.1(6)
χ_2^2	C_2^α–C_2^β–C_2^γ–C_2^δ	0.8(8)
	C_2'–C_2^α–C_2^β–C_2^γ	–122.4(5)
	C_2^β–C_2^γ–C_2^δ–N_2	–0.1(8)
ψ_2	N_2–C_2^α–C_2'–N_3	152.0(4)

Dipeptide Conformation in Solid State

The overall molecular conformation is described by selected torsion angles listed in Table VII. The conformation of the Boc-group defined by the torsion angles ω_0 [166.1(3)°] and θ_0 [174.4(2)°] is *trans-trans,* as observed in other Boc-protected peptides.[39] The hydrogen bonds between ethanol and *N-tert*-butyloxycarbonyl group contribute to the *trans-trans* conformation of the Boc-group. The characteristic torsion angles for peptide backbone conformation are: $\phi_1 = -63.2(5)°$, $\psi_1 = 156.1(4)°$, $\omega_1 = -174.3(4)°$, $\phi_2 = -66.0(5)°$, and $\psi_2 = 152.0(4)°$ (Scheme 1 and Table VII). According to the values of ϕ and ψ, *trans* proline residue is of β type which is generally characterized by $\phi \approx -65°$ and $\psi \approx 150°$.[3] The extensive analysis of *cis*/*trans* isomerization and the proline conformations in proteins[5,9,40] revealed that the conformation adopted by the proline is influenced by the nature of the preceding residue. The recent analysis of MacArthur and Thornton,[3] based on non-homologous protein structures determined at ≤ 2.5 Å resolution, including 963 non-identical *trans* proline residues, showed that both α and β types are present in Phe-Pro sequence (51 : 49).

The conformation of the Phe side chain is described by χ_1^1 and χ_1^2 (Table VII). The angle of 63.4(3)° between bond C_1^α–C_1^β and an aromatic system defines its orientation towards a peptide backbone (Figure 2). The angle between the best least-squares planes of phenyl and pyrrolidine ring is 13.5°.

Crystal Packing

The crystal packing is illustrated in Figure 3. Two-dimensional network is realized *via* hydrogen bonds of N–H···O and O–H···O types (Table VIII). The solvent molecule (ethanol) participates as a donor and an acceptor in hydrogen bonds to the peptide molecule *via* O6–H···O_0 and N_1–H···O_6 (Figure 3, Table VIII). Molecules related by the symmetry operation of 2_1 are connected by hydrogen bonds N_3–H331···O_0 and N_3–H332···O_1 into helices along ***b***, which include the eleven-membered ring structure $R_4^3(11)$ as described in the graph set notation.[41]

Comparative Conformational Analysis

The analysis includes comparison of the molecular conformations in solid state (X-ray data) and in DMSO-d_6 solution (NMR data) with the results of molecular mechanics and molecular dynamics simulations performed in DMSO (Table V).

The most interesting part of the conformational analysis is related to the peptide backbone, particularly the orientation of the proline residue (ω_1). In the solid state, the *trans* proline conformation [$\omega_1 = -174.3(4)°$] is observed. Sterical reasons related to the formation of hydrogen bonds between the

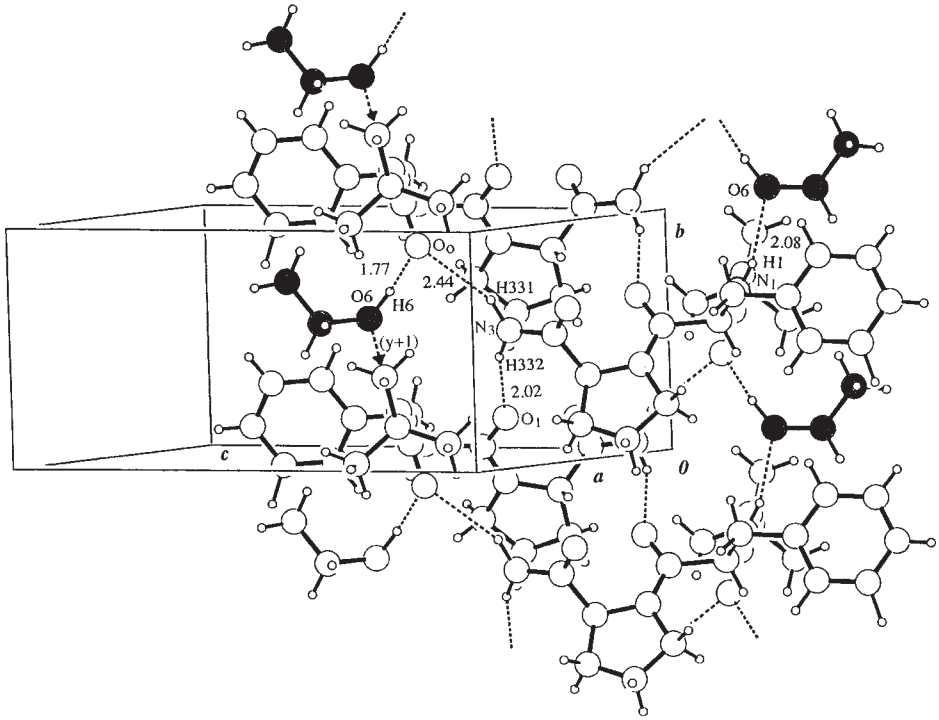


Figure 3. Crystal packing dominated by hydrogen bonds. Molecules of solvent (ethanol) are shaded. Hydrogen bonds are shown by dashed lines.

amide protons and neighbouring molecules [N_3–H331 $\cdots$ O_0, 3.335(6) Å and N_3–H332 $\cdots$ O_1, 3.038(7) Å (Figure 3, Table VIII)] are in favour of the *trans* proline orientation. An analogous example in the literature[46] is the crystal

TABLE VIII

Hydrogen bond geometry with estimated standard deviations given in parentheses

Type of H-bonds	D $\cdots$ A / Å	D–H / Å	H $\cdots$ A / Å	D–H $\cdots$ A / °	Symmetry operations on A
N_1–H1 $\cdots$ O6[a]	2.906(6)	0.860(6)	2.077(6)	162(1)	*x, y, z*
O6[a]–H6[a] $\cdots$ O_0	2.759(4)	1.01(6)	1.77(6)	168(4)	*x, y*+1, *z*
N_3–H331 $\cdots$ O_0	3.335(6)	0.99(9)	2.44(8)	150(5)	1–*x*, 1/2+*y*, –*z*
N_3–H332 $\cdots$ O_1	3.038(7)	1.07(7)	2.02(7)	157(5)	1–*x*, 1/2+*y*–1, –*z*

[a] Ethanol.

D: donor, A: acceptor

structure of Boc-Pro-Phe-Pro-OH with intermolecular hydrogen bonds between carbonyl oxygen of the C-terminal Pro and the amide proton of the Phe for both conformers (in the unit cell) with the *trans* proline orientation. The bulky, protected Boc-group restricts free rotation of the phenyl ring (Phe residue) and thus limits the conformational space for *trans*/*cis* conversion.

Experimental evidence from NMR data for the title dipeptide strongly confirms the *trans* Phe-Pro amide bond orientation ($\omega_1 \approx 180\,^\circ$) and this finding is supported by the results extracted from molecular mechanics and molecular dynamics simulations. The simultaneous rotations about two bonds C_1'–N_2 (ω_1) and C_1^α–C_1^β (χ_1^1) revealed energy minimum for $\omega_1 \approx 180\,^\circ$ and three energetically distinct rotamers with $\chi_1^1 = -60\,^\circ$, $+60\,^\circ$ and $180\,^\circ$ (Figure 4a). Molecular dynamics simulations confirmed that $\chi_1^1 = -60\,^\circ$ and $\chi_1^1 = +60\,^\circ$ are dominant rotamers, as it was experimentally deduced from NMR coupling constants (Table V). For the highly populated rotamer ($\chi_1^1 = -60\,^\circ$), the close contacts between Phe-$C_1^{\beta\,\mathrm{pro}S}$ and Pro-C_2^δ protons (Figure 4b) are only possible for the *trans* positioned Pro accompanied by the corresponding NOE. In solid state, the rotamer $\chi_1^1 = -69.5\,^\circ$ was observed as

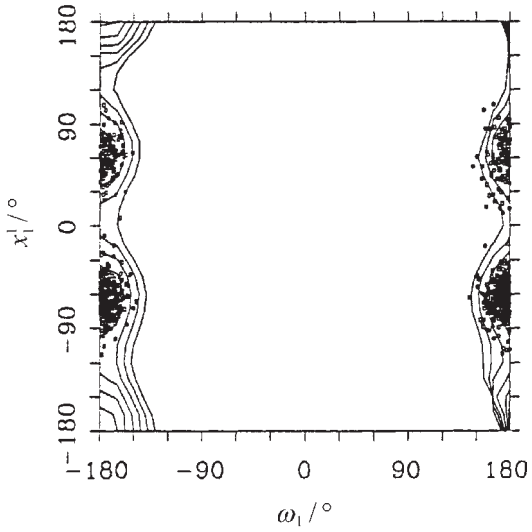


Figure 4. a) Molecular dynamic simulations performed over 320 ps at elevated temperatures (300 K, 350 K, 400 K over 50 ps each, 450 K over 100 ps and again 300 K over 70 ps) in DMSO. The values of torsion angles ω_1 and χ_1^1 obtained during the simulations are superimposed on the contour graph (obtained by rotations about the same torsion angles). Contour lines are drawn at an interval of 8.37 kJ / mol. Locations of energy minima in contour graph coincide with the mostly populated conformations in molecular dynamics simulation (high density dots).

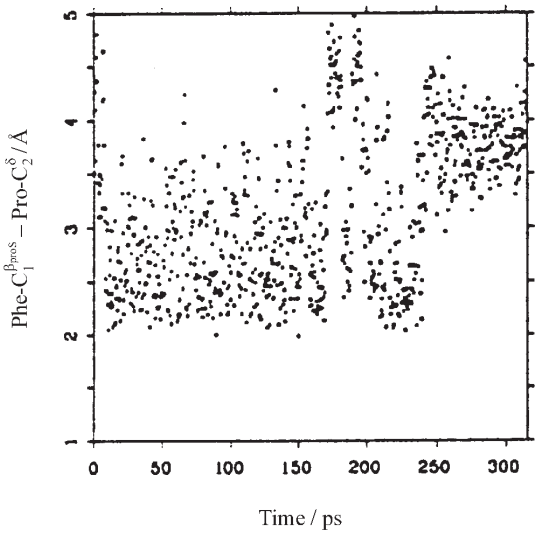


Figure 4. b) Molecular dynamics simulations performed over 320 ps at room temperature (300 K) in DMSO. Variations of the distance between Phe-$C_1^{\beta proS}$ and Pro-C_2^{δ} protons during the simulation time are shown. The corresponding NOE was observed in NMR experiments.

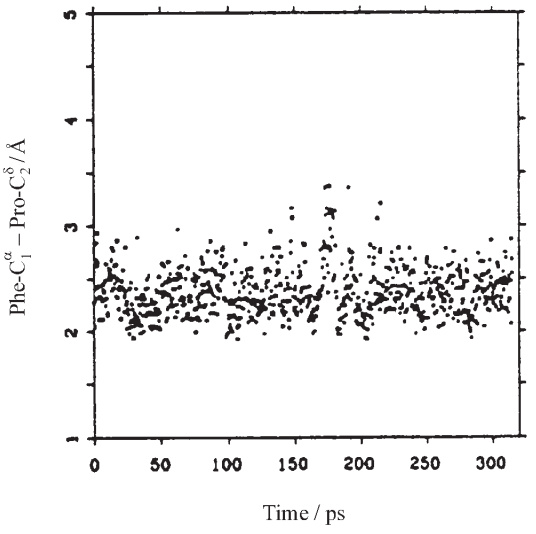


Figure 4. c) Molecular dynamics simulations performed over 320 ps at room temperatures (300 K) in DMSO.Variations of the distance between Phe-C_1^{α} and Pro-C_2^{δ} protons during the simulation time are shown. The corresponding NOE was observed in NMR experiments.

well. Molecular dynamics simulations performed even at higher temperatures revealed the *trans* proline orientation, only. The strong NOE between Phe-C_1^{α} and Pro-C_2^{δ} protons (Figure 4c) corresponds to the *trans* Pro orientation.

In order to examine the influence of a Boc-group on the *trans*/*cis* transition molecular dynamics simulations at 800 K of H-Phe-Pro-NH_2 was performed in vacuo. After 200 ps the *trans*/*cis* transition occurred with a two times lower barrier than in Boc-Phe-Pro-NH_2. It is important to mention that NMR data of unprotected H-Phe-Pro-NH_2 also detected both isomers in 70 : 30 ratio (in favour of *trans*).[19]

Supplementary Materials. – Crystallographic data for the structure reported in this paper have been deposited at Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 IEZ, UK (fax: + 44–1223–336033; e–mail: deposit@ccdc.cam.ac.uk) and can be obtained on request, free of charge, by quoting the publication citation and the deposition number 102928.

REFERENCES

1. T. E. Creighton, *Proteins: Structures and Molecular Properties,* 2nd ed., Freeman & Co., New York, 1993, pp. 312 and 324.
2. G. Vanhoof, F. Goossens, I. De Meester, D. Hendriks, and S. Scharpe, *FASEB J.* **9** (1995) 736.
3. M. W. MacArthur and J. M. Thornton, *J. Mol. Biol*. **218** (1991) 397; and the references cited therein.
4. K. Wüthrich and C. Grathwohl, *FEBS Lett.* **43** (1974) 337.
5. S. Fischer, R. L. Dunbrack Jr., and M. Karplus, *J. Am. Chem. Soc*. **116** (1994) 11931.
6. G. Schreiber and A. R. Fersht, *Biochemistry* **32** (1993) 11195.
7. G. Siligardi and A. F. Drake, *Biopolym. (Pept. Sci.)* **37** (1995) 281.
8. S. Feng, J. K. Chen, H. Yu, J. A. Simon, and S. L. Schreiber, *Science* **266** (1994) 1241.
9. G. D. Fasman, *Prediction of Protein Structure and the Principles of Protein Conformation*, Plenum Press, New York, 1990.
10. K. Tonan and S. -I. Ikawa, *J. Am. Chem. Soc.* **118** (1996) 6960.
11. D. Q. McDonald and W. C. Still, *J. Org. Chem.* **61** (1996) 1385.
12. C. Francart, J. -M. Wieruszeski, A. Tartar, and G. Lippens, *J. Am. Chem. Soc*.**118** (1996) 7019.
13. F. C. Bernstein, T. F. Koetzle, G. J. B. Williams, E. F. Jr. Meyer, M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi, The Protein Data Bank: a computer based archival file for macromolecular structures, *J. Mol. Biol.* **122** (1997) 535.
14. A. Henschen, F. Lottspeich, V. Brantl, and H. Teschemacher, *Hoppe-Seyler's Z. Physiol. Chem*. **360** (1979) 1217.
15. K. J. Su, Y. F. Chang, D. A. Brent, and J. K. Chang, *J. Biol. Chem*. **260** (1985) 9706.
16. M. Goodman and D. F. Mierke, *J. Am. Chem. Soc.* **111** (1989) 3489.

17. T. Yamazaki, S. Ro, M. Goodman, N. N. Chung, and P. W. Schiller, *J. Med. Chem.* **36** (1993) 708.
18. L. Moroder, A. Hallet, E. Wünsch, O. Keller, and G. Wersin, *Hoppe-Seyler's Z. Physiol. Chem.* **357** (1976) 1651.
19. D. F. Mierke, G. Nössner, P. W. Schiller, and M. Goodman, *Int. J. Pept. Protein Res*. **35** (1990) 35.
20. A. L. Spek, HELENA. Program for Data Reduction, Utrecht University, The Netherlands, 1990.
21. G. M. Sheldrick, *Acta Crystallogr., Sect A* **46** (1990) 467.
22. G. M. Sheldrick, SHELXL93. Program for crystal structure refinement, University of Göttingen, Germany, 1993.
23. A. L. Spek, in: D. Sayre (Ed.), *The EUCLID Package. Computation Crystallography,* Claredon Press, Oxford, 1982, pp. 528.
24. C. K. Johnson, ORTEPII. Report ORNL-5138. Oak Ridge National Laboratory, Tennessee, USA, 1976.
25. BIOSYM, DISCOVER, Version 2.9.7., Biosym Technologies, 1065 Barnes Canyon Rd., San Diego, CA 92121, USA, 1995.
26. S. W. Homans, *Biochemistry* **29** (1990) 9110.
27. K. Wüthrich, *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York, 1986.
28. O. Jardetzky and G. K. Roberts, *NMR in Molecular Biology*, Academic Press, New York, 1981.
29. K. D. O' Neal, M. V. Chari, C. H. McDonald, R. G. Cook, L. -Y. Yu-Lee, J. D.Morrisett, and W. T. Shearer, *Biochem. J.* **315** (1996) 833.
30. L. Pogliani, M. Ellenberger, J. Valat, and A. M. Bellocq, *Int. J. Pept. Protein Res.* **7** (1975) 345.
31. M. Petersheim, R. L. Moldow, H. N. Halladay, A. J. Kastin, and A. J. Fischman, *Int. J. Pept. Protein Res.* **40** (1992) 41.
32. J. Poznanski, A. Ejchart, K. L. Wierzchowski, and M. Ciurak, *Biopolymers* **33** (1993) 781.
33. K. G. R. Pachler, *Spectrochim. Acta* **20** (1964) 581.
34. H. Kessler, C. Griesinger, and K. Wagner, *J. Am. Chem. Soc.* **109** (1987) 6927.
35. IUPAC-IUB Commission on Biochemical Nomenclature (CBN) *J. Mol. Biol.* **55** (1971) 299.
36. J. Trikha, H. C. Patel, and T. P. Singh, *Acta Crystallogr.*, *Sect. C* **46** (1990) 74.
37. T. Ashida and M. Kakudo, *Bull. Chem. Soc. Jpn*. **47** (1974) 1129.
38. E. J. Milner-White, L. H. Bull, and P. H. MacCallum, *J. Mol. Biol.* **228** (1992) 725.
39. E. Subramanian and J. J. Sahayamary, *Int. J. Pept. Protein Res.* **41** (1993)319.
40. J. S. Richardson and D. C. Richardson, *Science* **240** (1988) 1468.
41. J. Bernstein, R. E. Davis, L. Shimoni, and N. -L. Chang, *Angew. Chem*., *Int. Ed. Engl.* **34** (1995) 1555.
42. V. F. Bystrov, V. T. Ivanov, S. L. Portnova, T. A. Balashova, and Yu. A. Ovchinnikov, *Tetrahedron* **29** (1973) 873.
43. M. Karplus, *J. Chem. Phys*. **30** (1959) 11.
44. C. A. G. Haasnoot, F. A. A. M. de Leeuw, and C. Altona, *Tetrahedron* **36** (1989) 2783.
45. H. Kessler, C. Griesinger, J. Lautz, A. Müller, W. F. van Gunsteren, and H. J. C. Berendsen, *J. Am. Chem. Soc*. **110** (1988) 3393.
46. P. J. Milne, D. W. Oliver, H. M. Ross, *J. Cryst. Spectrosc. Res.* **23** (1993) 449.
47. S. Tomić and V. Milinković, personal communication.

SAŽETAK

Konformacijske studije u čvrstom stanju i otopini C-zaštićenog dipeptidnog dijela (Boc-Phe-Pro-NH_2) morficeptina

Biserka Kojić-Prodić, Snježana Antolić, Marina Kveder, Ivanka Žigrović, Jurka Kidrič i Štefica Horvat

Određena je kristalna struktura C-zaštićenog dipeptidnog dijela (Boc-Phe-Pro-NH_2) visoko selektivnog antagonista μ-opioidnog receptora morficeptina (Tyr-Pro-Phe-Pro-NH_2); kristali su monoklinski, prostorne skupine $P2_1$ i dimenzija jedinične ćelije: a = 11,5731(5), b = 6,4490(3), c = 15,4082(5) Å, β = 100,359(5) ° i Z = 2. Proučavana je molekulska konformacija u čvrstom stanju i otopini (uporabom ^{1}H and ^{13}C NMR podataka), da bi se ispitao utjecaj prolina na konformaciju peptidne veze. Rentgenska analiza pokazala je sljedeću konformaciju peptidne veze: ϕ_1 = –63,2(5) °, ψ_1 = 156,1(4) °, ω_1 = –174,3(4) °, ϕ_2 = –66,0(5) ° i ψ_2 = 152,0(4) °. Konformacija Boc-skupine je *trans-trans*. Eksperimentalni podatci pokazuju *trans* konformaciju oko amidne veze Phe-Pro u čvrstom stanju kao i u otopini (DMSO). Mogućnost *cis*/*trans* izomerizacije oko peptidne veze (ω_1) istražena je teorijskim metodama uporabom programskog paketa BIOSYM. Molekulsko modeliranje, uključujući simulaciju molekulske dinamike naslovnog peptida, također potvrđuje *trans*-peptidnu vezu.