β-Turn Structure in Glycinylphenylalanine Dipeptide Based N-Amidothioureas

Xiao-Sheng Yan, Kun Wu, Yuan Yuan, Ying Zhan, Jin-He Wang, Zhao Li and Yun-Bao Jiang*

Department of Chemistry, College of Chemistry and Chemical Engineering, and the MOE Key Laboratory of Analytical Sciences, Collaborative Innovation Centre of Chemistry for Energy Materials (iChEM), Xiamen University, Xiamen 361005, China.
Tel: +86 592 218 8372; Fax: +86 592 218 5662; E-mail: ybjiang@xmu.edu.cn

Electronic Supplementary Information (ESI)
Scheme S1. Syntheses of 2, 1-3, 1-4a and 1-4b

**NMe₂-GFOEt**: To a chilled solution of NMe₂-Gly-OH·HCl (0.56 g, 4.0 mmol) and Et₃N (0.62 mL, 4.4 mmol) in CHCl₃ (10 mL) was added isobutyl chloroformate (0.65 mL, 5.0 mmol) at 0°C. After 30 min, a solution of (L or D) H-Phe-OEt·HCl (0.69 g, 3.0 mmol) and Et₃N (0.46 mL, 3.3 mmol) in CHCl₃ (20 mL) was added. The mixture was left to stand at room temperature for 4 hours, evaporated in vacuo, and the solid residue was dissolved in AcOEt. The solution was washed successively with 1% NH₃·H₂O, saturated NH₄Cl and water, dried over anhydrous Na₂SO₄.
and concentrated under reduced pressure, to obtain 0.76 g oily product NMe₂-GFOEt, 91.0%.

NMe₂-GFNHH₂: Excess aqueous hydrazine (80%) was added to NMe₂-GFOEt in ethanol (15.0 mL) and then refluxed for 24 hours, evaporated in vacuo, and the viscous liquid was dissolved in water. The solution was extracted by CH₂Cl₂, dried over anhydrous Na₂SO₄ and concentrated under reduced pressure, to obtain 0.56 g NMe₂-GFNHH₂, 77.8%.

2: NMe₂-GFNHH₂ then reacted with excess phenyl isothiocyanate in CH₂Cl₂. The solution was stirred at room temperature for 24 hours, after which the solvent was removed under reduced pressure. The product was washed by petroleum ether, water and hot diethyl ether to obtain 0.59 g 2, 69.7%.

1: ¹H NMR (400 MHz, DMSO-d₆): δ (ppm) 10.30 (s, 1H), 9.78 (s, 1H), 9.33 (s, 1H), 7.98 (s, 1H), 7.53 (d, J = 7.7 Hz, 2H); 7.37 – 7.11 (m, 8H), 4.59 (s, 1H), 3.20 (dd, J = 13.9, 4.5 Hz, 1H), 3.05 – 2.88 (m, 2H), 2.78 (d, J = 15.5 Hz, 1H), 2.07 (s, 6H); ¹³C NMR (101 MHz, CD₂CN): δ (ppm) 182.00, 172.27, 170.68, 138.73, 136.77, 129.28, 128.65, 128.36, 126.94, 125.74, 124.98, 62.37, 54.01, 45.04, 36.31; HRMS (ESI): calcd for [C₂₀H₂₆N₂O₂S]⁺: 400.1802, found: 400.1807.

Crystal data for compound 1: C₂₀H₂₅N₅O₂S, Mᵣ = 399.51, T = 293(2) K, Orthorhombic, space group P₂₁2₁₂₁, a = 8.7524(4), b = 15.1376(7), c = 15.8697(8) Å, V = 2102.58(17) Å³, ρ_c = 1.262 Mg.m⁻³, µ(MoKα) = 0.179 mm⁻¹, Z = 4, reflections collected: 5437, independent reflections: 3517 (Rint = 0.0211), S = 0.814, final R indices [I > 2σ(I)]: R₁ = 0.0360, wR₂ = 0.1040, R indices (all data): R₁ = 0.0411, wR₂ = 0.1088, Flack parameter = -0.03(8), CCDC No. 914552.

2: ¹H NMR (400 MHz, DMSO-d₆): δ (ppm) 10.31 (s, 1H), 9.78 (s, 1H), 9.33 (s, 1H), 7.97 (s, 1H), 7.52 (d, J = 7.7 Hz, 2H), 7.41 – 7.11 (m, 8H), 4.59 (s, 1H), 3.20 (dd, J = 13.9, 4.6 Hz, 1H), 3.00 – 2.89 (m, 2H), 2.78 (d, J = 15.5 Hz, 1H), 2.07 (s, 6H); ¹³C NMR (101 MHz, CD₂CN): δ (ppm) 182.00, 172.26, 170.71, 138.73, 136.77, 129.29, 128.65, 128.36, 126.94, 125.75, 125.00, 62.36, 53.99, 45.04, 36.33; HRMS (ESI): calcd for [C₂₀H₂₆N₂O₂S]⁺: 400.1802, found: 400.1809.

1: NMe₂-GFNHH₂ (0.53 g, 2.0 mmol) then reacted with equivalent phenyl isocyanate (0.24 g, 2.0 mmol) in 50 mL CH₂CN. The solution was stirred at room temperature for 3 hours, after which the solvent was removed under reduced pressure. The product was washed by anhydrous ether and recrystallized in CH₂CN to obtain 0.61 g 1-3, 79.6%.

2: ¹H NMR (500 MHz, CD₂CN): δ (ppm) 8.41 (s, 1H), 7.86 (s, 1H), 7.60 (d, J = 4.8 Hz, 1H), 7.53 (d, J = 8.6, 0.9 Hz, 2H), 7.33 (d, J = 10.0, 4.6 Hz, 2H), 7.31 – 7.21 (m, 5H), 7.07 – 6.96 (m, 1H), 6.67 (s, 1H), 4.42 (dt, J = 9.0, 6.0 Hz, 1H), 3.21 (dd, J = 14.0, 5.7 Hz, 1H), 3.02 (dd, J = 13.9, 9.1 Hz, 1H), 2.94 (d, J = 16.3 Hz, 1H), 2.81 (d, J = 16.3 Hz, 1H), 2.13 (s, 5H); ¹³C NMR (126 MHz, MeOD): δ (ppm) 172.40, 171.88, 156.62, 138.60, 136.58, 129.97, 128.32, 128.24, 126.60, 122.88, 119.65, 61.91, 53.22, 44.50, 37.05; HRMS (ESI): calcd for [C₂₀H₂₆N₂O₁]⁺: 384.2036,
found: 384.2036. Crystal data for compound t-3: C20H25N5O3, \( M_r = 383.45 \), \( T = 273(2) K \),
Orthorhombic, space group P2\(_1\)2\(_1\)2\(_1\), \( a = 8.6444(15) \), \( b = 15.088(3) \), \( c = 15.404(3) \) \( \AA \), \( V=2009.1(6) \)
\( \AA^3 \), \( \rho_{\text{c}} = 1.268 \text{Mg.m}^{-3} \), \( \mu(\text{MoKα}) = 0.088 \text{mm}^{-1} \), \( Z = 4 \), reflections collected: 11532, independent
reflections: 4677 (\( R_{\text{int}} = 0.0451 \)), \( S = 1.0303 \), final \( R \) indices [1 > 2σ(1)]: \( R_1 = 0.0443 \), \( wR_2 = 0.1172 \), \( R \) indices (all data): \( R_1 = 0.0470 \), \( wR_2 = 0.1193 \), CCDC No. 926794.

**t-Boc-GFOEt**: To a chilled solution of Boc-Gly-OH (1.05 g, 6.0 mmol) and Et\(_3\)N (0.84 mL, 6.0 mmol) in CHCl\(_3\) (10 mL) was added isobutyl chloroformate (1.0 mL, 7.5 mmol) at 0 °C. After 30 min, a solution of L-Phe-OEt HCl (1.05 g, 4.5 mmol) and Et\(_3\)N (0.63 mL, 4.5 mmol) in CHCl\(_3\) (20 mL) was added. The mixture was left to stand at room temperature for 12 hours, evaporated in vacuo, and the residue was dissolved in AcOEt. The solution was washed successively with 1% NH\(_3\)H\(_2\)O, saturated NH\(_4\)Cl, 1% HCl and water, dried over anhydrous Na\(_2\)SO\(_4\) and concentrated under reduced pressure, to obtain 1.35 g t-Boc-GFOEt, 85.6%.

**t-Boc-GFNHH**: Excess aqueous hydrazine (80%) was added to t-Boc-GFOEt in ethanol (20.0 mL) and then refluxed for 24 hours, evaporated in vacuo, and the residue was washed by AcOEt and recrystallized in ethanol to obtain 0.98 g t-Boc-GFNHH, yield 76.0%.

**t-4a**: t-Boc-GFNHH then reacted with phenyl isothiocyanate in ethanol and then refluxed for 24 hours, evaporated in vacuo. The oily residue was washed by petroleum ether and then obtain solid product, which was washed by a little CH\(_2\)Cl\(_2\) to obtain 1.08 g t-4a, yield 78.8%.

**t-4b**: 10 mL CH\(_2\)Cl\(_2\) and 10 mL CF\(_3\)COOH was added in 0.5 g t-4a and then the solution was stirred at room temperature for 5 hours, evaporated in vacuo. The oily residue was dissolved in water. Then adjust the pH value to 8-9 by saturated NaHCO\(_3\) solution. The solution was extracted by CH\(_2\)Cl\(_2\), dried over anhydrous Na\(_2\)SO\(_4\) and concentrated under reduced pressure, to obtain 0.30 g t-4b, yield 76.1%.

**t-4a**: \( ^1\)H NMR (400 MHz, CD\(_3\)CN): \( \delta \) ppm: 8.84 (s, 1H), 7.61 (d, \( J = 7.8 \) Hz, 2H), 7.44 – 7.18 (m, 8H), 7.13 (s, 1H), 5.72 (s, 1H), 4.37 (dt, \( J = 8.7, 35.8 \) Hz, 1H), 3.70 – 3.53 (m, 2H), 3.21 (dd, \( J = 14.0, 5.4 \) Hz, 1H), 3.04 (dd, \( J = 14.0, 8.9 \) Hz, 1H), 1.37 (s, 9H); \( ^{13}\)C NMR (101 MHz, CD\(_3\)CN): \( \delta \) ppm: 181.85, 171.84, 170.45, 156.65, 138.73, 136.84, 129.29, 128.59, 128.35, 126.89, 125.64, 124.73, 79.63, 54.83, 43.92, 36.01, 27.53; HRMS (ESI): calcd for \([\text{C}_{23}\text{H}_{29}\text{N}_{3}\text{O}_{4}\text{SNa}]^{+}\): 494.1832, found: 494.1834.

**t-4b**: \( ^1\)H NMR (500 MHz, CD\(_3\)CN): \( \delta \) ppm: 9.08 (s, 1H), 7.91 (s, 1H), 7.57 – 7.55 (m, 2H), 7.35 – 7.17 (m, 8H), 4.41 (s, 1H), 3.21 – 3.13 (m, 3H), 3.01 (dd, \( J = 13.9, 8.7 \) Hz, 1H); \( ^{13}\)C NMR (101 MHz, CD\(_3\)CN): \( \delta \) ppm: 181.37, 174.47, 170.91, 138.61, 136.56, 129.18, 128.54, 128.31, 126.89, 125.65, 124.92, 54.06, 43.79, 36.65; HRMS (ESI): calcd for \([\text{C}_{18}\text{H}_{22}\text{N}_{3}\text{O}_{2}\text{S}]^{+}\): 372.1489, found: 372.1489.
Figure S1. CD spectra of 1 (a) and 2 (b) in CH₂CN. [1] = [2] = 40 μM.

Figure S2. ¹H NMR spectra of L-2 in CD₃CN, DMSO-d₆ and CDCl₃. [L-2] = 10 mM.

Table S1. Torsions in the X-ray crystal of L-2

<table>
<thead>
<tr>
<th></th>
<th>( \Phi_i )</th>
<th>( \Psi_{i+1} )</th>
<th>( \Phi_{i+1} )</th>
<th>( \Psi_{i+2} )</th>
<th>( \Phi_{i+2} )</th>
<th>( \Psi_{i+2} )</th>
<th>( \Phi_{i+2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \omega )</td>
<td>176.63°</td>
<td>-73.47°</td>
<td>117.51°</td>
<td>-173.39°</td>
<td>77.92°</td>
<td>14.84°</td>
<td>-178.10°</td>
</tr>
</tbody>
</table>
Figure S3. Calculated structures of 1-2 at B3LYP/6-31++G** level.

Figure S4. Influence on NH proton resonances of 1-2 in CD$_3$CN/DMSO-$d_6$ and CD$_3$CN/H$_2$O mixture by volume ratio of (a) DMSO-$d_6$ and (b) H$_2$O. Signals of NH$_i$ and NH$_k$ are invisible in CD$_3$CN/H$_2$O mixture. [1-2] = 10 mM. For NH proton numbering see Figure 2 in the text.
Figure S5. $^1$H-$^1$H NOESY spectrum of L-2 in CD$_3$CN. The nuclear Overhauser effect (NOE) signals between $H_a$ or $H_b$ and $H_f$ are in consistent with the existence of $\beta$-turn in L-2. [L-2] = 10 mM.

Figure S6. Absorption (a) and CD (b) spectra of L-2 in CH$_3$CN-H$_2$O binary solvents. [L-2] = 80 $\mu$M. Using CD signal at 270 nm as an indication of the $\beta$-turn structure, it is assumed that it exists in solution containing up to ca. 15% by volume water.
Figure S7. Temperature dependent CD spectra of \( L \cdot 2 \) in CH\(_3\)CN. \([L \cdot 2]\) = 80 \( \mu \)M. The CD spectrum does not change very much, suggesting the \( \beta \)-structure in 2 is stable over 25 - 45 °C.

Figure S8. \(^1\)H NMR spectra of \( L \cdot 2 \) in the presence of AcO\(^-\) in CD\(_3\)CN. \([L \cdot 2]\) = 10 mM.
**Figure S9.** Splitting of NMR signals of protons H<sub>b</sub>, H<sub>a</sub> and H<sub>d</sub>, H<sub>c</sub> in 1·2 in the presence of AcO<sup>-</sup> in CD<sub>3</sub>CN. [1·2] = 10 mM.

**Scheme S2.** AcO<sup>-</sup> binding with 2 and two possible models of hydrogen bonding networks in the anion binding complex.
Figure S10. X-ray crystal structure of l-3 and torsion angles of l-3. Dashed pink lines highlight the intramolecular hydrogen bonds.

Table S2. Bond lengths and angles of β-turns in the crystal structures of l-2 and l-3

<table>
<thead>
<tr>
<th></th>
<th>N5-H ( \cdot \cdot \cdot ) O1</th>
<th>( \angle ) N5H O1</th>
</tr>
</thead>
<tbody>
<tr>
<td>l-2</td>
<td>2.401 Å</td>
<td>152.72°</td>
</tr>
<tr>
<td>l-3</td>
<td>2.487 Å</td>
<td>146.53°</td>
</tr>
</tbody>
</table>
Figure S11. Influence of DMSO-$d_6$ volume ratio in CD$_3$CN/DMSO-$d_6$ mixture on the resonance of NH$_3$ proton in L-2 and L-3. [L-2] = [L-3] = 10 mM.

Figure S12. Absorption (a) and CD (b) spectra of L-3 in CH$_3$CN in the presence of AcO$^-$. [L-3] = 40 μM.
Table S3. Binding constants of L-2 and L-3 with AcO⁻ in CH₃CN

<table>
<thead>
<tr>
<th></th>
<th>( K_a ) ± sd, M⁻¹</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-2</td>
<td>(8.75 ± 5.41) ( \times 10^6 )</td>
<td>0.9973</td>
</tr>
<tr>
<td>L-3</td>
<td>(5.18 ± 0.39) ( \times 10^4 )</td>
<td>0.9983</td>
</tr>
</tbody>
</table>

\( a \). The constants were determined from the absorption titration at 296 nm (L-2) or 244 nm (L-3) as shown in Fig. 3 and Fig. S12. The following equation was used for fitting the data according to 1:1 binding model,

\[
A = A_c + \frac{A_{\text{lim}} - A_c}{2c_0} \left[ x + 1/K_a + c_0 - \sqrt{(x + 1/K_a + c_0)^2 - 4c_0x} \right]
\]

in which \( A_0 \), \( A \) and \( A_{\text{lim}} \) denote respectively the absorbance of host (L-2 or L-3), the bound complex and the limit value, \( c_0 \) is the molar concentration of host, \( x \) is the molar concentration of guest (AcO⁻), and \( K_a \) is the binding constant.

Figure S13. CD spectra of L-2 (a) and L-1 (b) in CH₃CN in the presence of AcO⁻. [L-2] = [L-1] = 40 \( \mu \)M.
Figure S14. Absorption spectra of L-2 (a), L-4a (b) and L-4b (c) in CH₃CN in the presence of AcO⁻.

\[ [\text{L-2}] = [\text{L-4a}] = [\text{L-4b}] = 40 \mu M. \]

Figure S15. CD spectra of L-2 (a), L-4a (b) and L-4b (c) in CH₃CN in the presence of AcO⁻.

\[ [\text{L-2}] = [\text{L-4a}] = [\text{L-4b}] = 40 \mu M. \]
$^1$H NMR and $^{13}$C NMR spectra of 2, L-3, L-4a and L-4b

$^1$H NMR of L-2 (400 MHz, DMSO-$d_6$)

$^{13}$C NMR of L-2 (101 MHz, CD$_3$CN)

Electronic Supplementary Material (ESI) for Chemical Communications
This journal is © The Royal Society of Chemistry 2013
$^1$H NMR of d-2 (400 MHz, DMSO-$d_6$)

$^{13}$C NMR of d-2 (101 MHz, CD$_3$CN)
$^{1}H$ NMR of L-3 (500 MHz, CD$_3$CN)

$^{13}C$ NMR of L-3 (126 MHz, MeOD)
$^1$H NMR of L-4a (400 MHz, CD$_3$CN)

$^{13}$C NMR of L-4a (101 MHz, CD$_3$CN)
$^1$H NMR of 1-4b (500 MHz, CD$_3$CN)

$^{13}$C NMR of 1-4b (101 MHz, CD$_3$CN)