Supplementary Material (ESI) for Chemical Communications

Electronic supplementary information

Synthesis and structural characterization as 12-helix of the hexamer of a β-amino acid tethered to a pyrrolidin-2-one ring

Ileana Menegazzo, Stefano Mammi and Alexander Fries*

Dipartimento di Scienze Chimiche - Università di Padova

Via Marzolo 1 - I-35131 Padova, Italy

Roberta Galeazzi, Gianluca Martelli, Mario Orena and Samuele Rinaldi*

Dipartimento di Scienze dei Materiali e della Terra - Università Politecnica delle Marche

Via Brecce Bianche - I-60131, Ancona, Italy

Table of Contents:

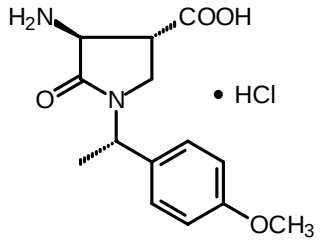
General: Commercial-grade reagents and solvents were used without further purification except as indicated. DCM was dried prior to use by percolation through anhydrous Al_2O_3. All reactions were stirred magnetically; moisture-sensitive reactions were performed under nitrogen in flame-dried glassware. Thin-layer chromatography (TLC), was used to monitor reactions. Visualization was accomplished by either ultraviolet light or by immersing the plate in a 5% aqueous solution of potassium permanganate and heating. Flash chromatography with silica gel was performed following the conditions described by Still and coworkers.[2] ^{1}H NMR (200 MHz) and ^{13}C NMR (50 MHz) spectra of compounds **2** – **10** were measured on a Varian Gemini 200 spectrometer, whereas ^{1}H NMR (600 MHz) and ^{13}C NMR (150 MHz) spectra of compound **1** were measured on a Bruker Avance DMX 600 spectrometer. Proton chemical shifts are reported as δ values relative to tetramethylsilane (0.00 ppm) or to the particular solvent used in the experiment. Carbon chemical shifts are reported as δ values relative to the particular solvent used in the experiment ($CDCl_3$: 77.0 ppm). Data are reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, dd = doublet of doublet, br = broad), coupling constant, and integration. The samples were analyzed with a liquid chromatography Agilent Technologies HP1100 equipped with a Zorbax Eclipse XDB-C8 Agilent and Technologies column (flow rate 0.5 mL/ min) and equipped with a diode-array UV detector (220 and 254 nm). Acetonitrile and methanol for HPLC were purchased from a commercial supplier. All the samples were prepared by diluting 1 mg in 5 mL of a 1:1 mixture of H_2O and acetonitrile in pure acetonitrile or in pure methanol. The MSD1100 mass detector was utilized under the following conditions: mass range 100-2500 uma, positive scanning, energy of fragmentor 50 V, drying gas flow (nitrogen) 10.0 mL/min, nebulizer pressure 45 psig, drying gas temperature 350 °C, capillary voltage 4500 V.

(3*S*,4*R*,1'*S*)-4-methoxycarbonyl-3-trichloroacetylamino-1-[1'-(4-methoxyphenylethyl)]pyrrolidin-2-one was prepared according to ref. 5 (R. Galeazzi, G. Martelli, M. Orena, S. Rinaldi and P. Sabatino, *Tetrahedron*, 2005, **61**, 5465-5473).

Abbreviations: DCM = dichloromethane, TEA = triethylamine, TFA = trifluoroacetic acid, EDAC = *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride.

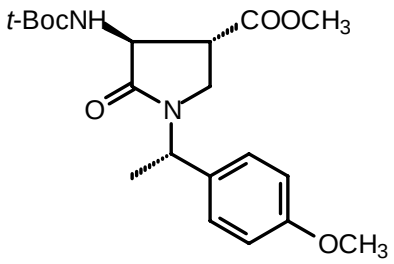
Experimental

(3*S*,4*R*,1'*S*)-3-Amino-2-oxo-1-[1'-(4-methoxyphenylethyl)]pyrrolidinecarboxylic acid [(3*S*,4*R*,1'*S*)-AOMPC], (2) hydrochloride.



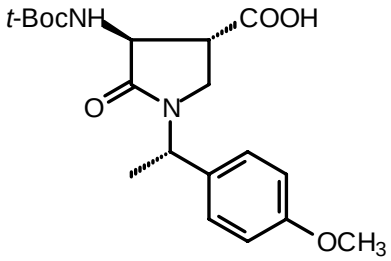
In a flask containing (3*S*,4*R*,1'*S*)-4-methoxycarbonyl-3-trichloroacetylamino-1-[1'-(4-methoxyphenylethyl)]pyrrolidin-2-one (R. Galeazzi, G. Martelli, M. Orena, S. Rinaldi and P. Sabatino, *Tetrahedron*, 2005, **61**, 5465-5473) (0.5 g, 1.14 mmol), 6 M NaOH (5 mL) was added at rt and the resulting mixture was stirred for 24 h. The pH of the homogeneous solution was adjusted to 1 by addition of 6 M HCl, water was removed under reduced pressure and by addition of MeOH (10 mL) the inorganic salts were precipitated and subsequently removed by suction filtration, followed by washing of the filter with MeOH (5 mL). The solvent was eliminated under reduced pressure and the residue was crystallized (MeOH-AcOEt 1:1) to give the hydrochloride of **2** (0.33 g, 93%) as white crystals: mp 272-273 °C (dec); ^{1}H NMR (200 MHz, D_2O) δ 1.53 (d, J = 7.2 Hz, 3H), 2.72 - 2.86 (m, 1H), 3.22 - 3.49 (m, 2H), 3.84 (s, 3H), 3.98 (d, J = 9.7 Hz, 1H), 5.26 (q, J = 7.2 Hz, 1H), 7.02 (d, J = 8.8 Hz, 2 ArH), 7.34 (d, J = 8.8 Hz, 2 ArH); ^{13}C NMR (50 MHz, D_2O) δ 18.8, 46.3, 51.3, 52.9, 58.5, 59.2, 117.2, 131.3, 135.0, 161.3, 177.1, 181.6; $[\alpha]_D$ -127.0 (c 1.0, H_2O); ESI-MS: *m/z* 279.2 $[MH]^+$, 301.4 $[M+Na]^+$. Anal. Calcd for $C_{14}H_{19}ClN_2O_4$ (314.10): C 53.40; H 6.09; N 8.90. Found: C, 53.34; H, 6.03; N, 8.95.

(3*S*,4*R*,1'*S*)-3-*t*-Butoxycarbonylamino-4-methoxycarbonyl-1-[1'-(4-methoxyphenyl)ethyl]pyrrolidin-2-one (3).



To a solution of compound **2** hydrochloride (0.31 g; 1.0 mmol) in dry MeOH (5 ml), Boc_2O (0.33 g, 1.5 mmol) and TEA (0.4 mL, 3.0 mmol) were added at rt. After 12 h water (5 mL) was added, MeOH was removed *in vacuo* and the solution was acidified with 1 M HCl (0.3 mL). After extraction with ethyl acetate (3 x 10 mL), the organic layer was dried (Na_2SO_4) and removal of the solvent under reduced pressure gave a residue that was taken off in methanol (5 mL). Subsequent esterification by excess diazomethane in ethyl ether followed by removal of the solvent and purification of the residue by silica gel chromatography (cyclohexane-ethyl acetate 7:3) gave compound **3** (0.36 g, 79%) as a viscous oil [***CAUTION****: Diazomethane is an explosive and a highly toxic gas. Explosions may occur if the substance is dry and undiluted. All operations involving diazomethane should be carried out in an efficient fumehood following appropriate precautions*]: IR ($CHCl_3$) ν 3341, 1731, 1725, 1668 cm^{-1}; ^1H NMR (200 MHz, $CDCl_3$) δ 1.42 (s, 9H), 1.50 (d, *J* = 7.2 Hz, 3H), 2.91 - 3.10 (m, 1H), 3.16 (dd, *J* = 8.8, *J* = 9.5 Hz, 1H), 3.38 (dd, *J* = 9.3, *J* = 9.5 Hz, 1H), 3.71 (s, 3H), 3.79 (s, 3H), 4.43 (dd, *J* = 6.3, *J* = 9.6 Hz, 1H), 5.20 (d, *J* = 6.3 Hz, 1H, NH), 5.44 (q, *J* = 7.2 Hz, 1H), 6.86 (d, *J* = 8.8 Hz, 2 ArH), 7.22 (d, *J* = 8.8 Hz, 2 ArH); ^{13}C NMR (50 MHz, $CDCl_3$) δ 16.1, 28.1, 41.5, 45.2, 49.2, 52.2, 55.1, 56.3, 79.9, 113.8, 128.1, 131.0, 155.2, 160.0, 169.2, 172.3; $[\alpha]_D$ -117.1 (c 1.9, $CHCl_3$); ESI-MS: *m/z* 393.2 $[MH]^+$, 415.1 $[MH+Na]^+$. Anal. Calcd for $C_{20}H_{28}N_2O_6$ (392.19): C, 61.21; H, 7.19; N 7.14. Found: C, 61.24; H, 7.16; N, 7.09.

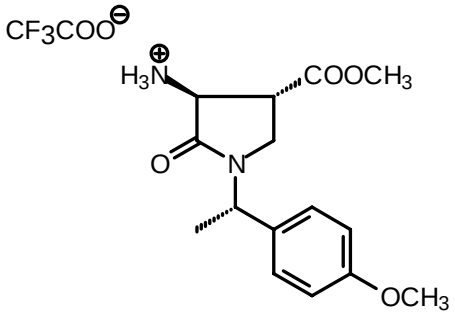
(3*S*,4*R*,1′*S*)-{3-*t*-Butoxycarbonylamino-1-[1′-(4-methoxyphenyl)ethyl]pyrrolidin-2-on-4-yl} carboxylic acid [(3*S*,4*R*,1′*S*)-*t*-Boc-AOMPC], (4).



A solution of the compound **3** (2.0 g, 5.1 mmol) in MeOH (1 mL) was added to a 0.5 M NaOH solution (40 mL, 20 mmol) at rt, followed by an additional amount of MeOH (10 mL) until complete solubilisation of the suspension was achieved, and the mixture was stirred for 12 h. Then, after evaporation of MeOH under reduced pressure, the aqueous solution was cooled to 0 °C and pH adjusted to 1 by addition of 6 M HCl. Eventually, the mixture was extracted with ethyl acetate (3 x 20 mL), the organic layer dried over Na_2SO_4 filtered and the filtrate was evaporated at reduced pressure, to give the compound **4** (1.75 g, 90%) as a white solid: mp. 72-74 °C; ^1H NMR (200

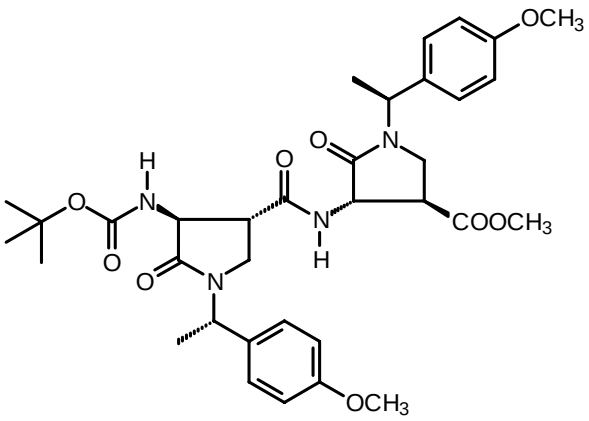
MHz, $CDCl_3$) δ 1.43 (s, 9H), 1.54 (d, *J* = 7.0 Hz, 3H), 3.08 - 3.35 (m, 2H), 3.38 - 3.42 (m, 1H), 3.80 (s, 3H), 4.52 (dd, *J* = 6.1 Hz, *J* = 8.4 Hz, 1H), 5.80 (d, *J* = 6.1 Hz, 1H, NH), 6.88 (d, *J* = 8.6 Hz, 2 ArH), 7.25 (d, *J* = 8.6 Hz, 2 ArH), 11.55 (br s, 1H, COOH); ^{13}C NMR (50 MHz, $CDCl_3$) δ 16.1, 28.1, 42.1, 43.9, 46.4, 50.2 (50%), 50.3 (50%), 55.3, 81.6, 114.1, 128.3, 130.1, 159.3, 163.9, 174.6; $[\alpha]_D$ -140.0 (c 0.5, $CHCl_3$); ESI-MS: *m/z* 379.2 $[MH]^+$, 401.2 $[MH+Na]^+$. Anal. Calcd for $C_{19}H_{26}N_2O_6$ (378.18): C, 60.30; H, 6.93; N 7.40. Found: C, 60.25; H, 6.86; N, 7.45.

Methyl (3*S*,4*R*,1′*S*)-{3-ammonium-1-[1′-(4-methoxyphenyl)ethyl]pyrrolidin-2-on-4-yl}carboxylate trifluoroacetate [(3*S*,4*R*,1′*S*)-H-AOMPC-OMe trifluoroacetate], (5).



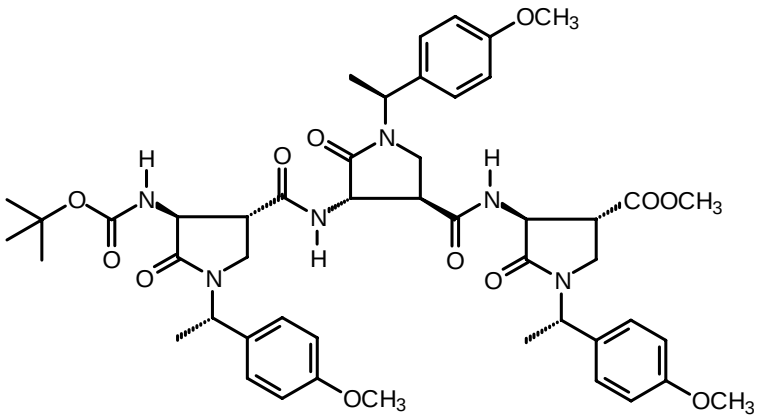
To a solution of compound **3** (784 mg, 2.0 mmol) in DCM (5 mL), TFA (2 mL, 3.07 g, 27.0 mmol) was slowly added and the clear solution was stirred for 1 h at rt. Then volatiles were eliminated under reduced pressure, and the residue was washed with diethyl ether, to give the product **5** (812 mg, quantitative yield) as a low-melting solid: ^{1}H NMR (200 MHz, $CDCl_3$) δ 1.52 (d, *J* = 7.0 Hz, 3H), 3.34 - 3.55 (m, 3H), 3.74 (s, 3H), 3.79 (s, 3H), 4.54 (d, *J* = 8.9 Hz, 1H), 5.36 (q, *J* = 7.0 Hz, 1H), 6.86 (d, *J* = 8.6 Hz, 2 ArH), 7.16 (d, *J* = 8.6 Hz, 2 ArH), 9.3 (br s, 3H); ^{13}C NMR (50 MHz, $CDCl_3$) δ 15.9, 41.5, 42.1, 50.9, 53.2, 54.1, 55.2, 114.4, 115.2 (q, *J* = 285.0 Hz), 128.1, 129.2, 159.5, 160.9 (q, *J* = 40.2 Hz), 169.9; $[\alpha]_D$ -109.8 (c 1.0, $CHCl_3$); ESI-MS: *m/z* 293.4 $[MH]^+$, 315.2 $[M+Na]^+$. Anal. Calcd for $C_{17}H_{21}F_3N_2O_6$ (406.14): C, 50.23; H, 5.21; N, 6.90. Found: C, 50.16; H, 5.17; N, 6.83.

***t*-Boc-[(3*S*,4*R*,1′*S*)-AOMPC]$_2$-OMe, (6).**



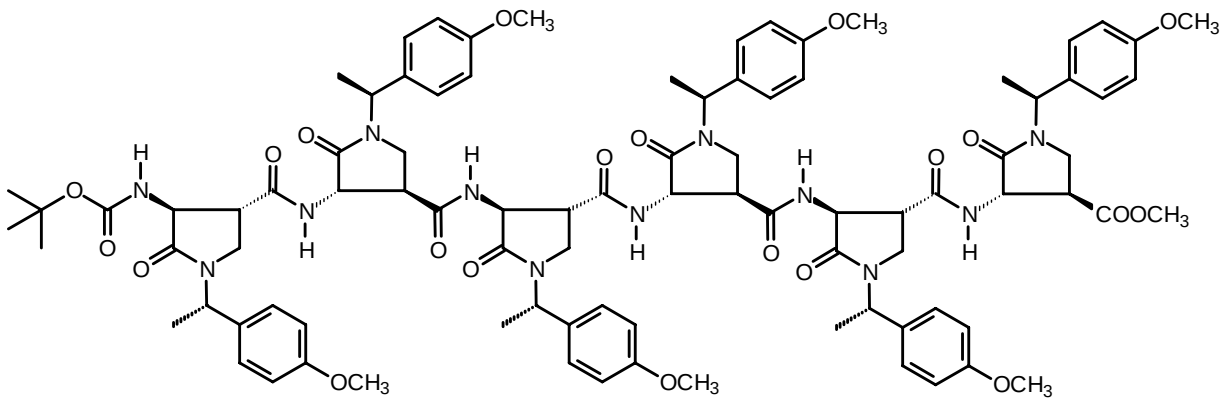
To a solution containing Et_3N (0.26 mL, 2.0 mmol) and trifluoroacetate **5** (812 mg, 2.0 mmol) in dry DCM (12 mL) at 0 °C, compound **4** (756 mg, 2.0 mmol) dissolved in dry DCM (3 mL) was added, followed by EDAC (400 mg, 2.1 mmol). The temperature raised to rt and the mixture was stirred for an additional 12 h. Water (15 mL) was then added, and the mixture was extracted with ethyl acetate (3 x 30 mL). The organic layer was dried (Na_2SO_4) and removed *in vacuo* and the residue was purified by silica gel chromatography (ethyl acetate 100%) to give the dimer **6** (0.93 g, 71%) as a white solid: mp 78-80 °C; ^{1}H NMR (200 MHz, $CDCl_3$) δ 1.42 (s, 9H), 1.46 (d, *J* = 7.0 Hz, 3H), 1.48 (d, *J* = 7.0 Hz, 3H), 2.85 - 3.02 (m, 1H), 3.08 - 3.35 (m, 4H), 3.55 (dd, *J* = 7.2 Hz, *J* = 8.1Hz, 1H), 3.69 (s, 3H), 3.78 (s, 6H), 4.43 (dd, *J* = 6.4 Hz, *J* = 8.3 Hz, 1H), 4.62 (dd, *J* = 5.9 Hz, *J* = 7.5 Hz, 1H), 5.33 (d, *J* = 5.9 Hz, 1H, NH), 5.42 (q, *J* = 7.0 Hz, 1H), 5.45 (q, *J* = 7.0 Hz, 1H), 6.84 (d, *J* = 8.3 Hz, 4 ArH), 7.21 (d, *J* = 8.3, 2 ArH), 7.23 (d, *J* = 8.3 Hz, 2 ArH), 8.45 (d, *J* = 6.4 Hz, 1H, NH); ^{13}C NMR (50 MHz, $CDCl_3$) δ 16.2, 16.3, 28.2, 41.5, 41.7, 41.8, 42.1, 43.8, 45.0, 49.1, 49.5, 52.3, 55.2, 55.3, 55.7, 56.2, 80.9, 144.0, 128.0,128.1, 128.2, 131.1, 131.2, 159.0 (50%), 159.1 (50%), 168.9 (50%), 169.3 (50%), 170.4, 170.9, 172.2 (50%), 172.6 (50%); $[\alpha]_D$ -235.0 (c 0.5, $CHCl_3$); ESI-MS: *m/z* 653.3 (M^+), 675.2 $[M+Na]^+$. Anal. Calcd for $C_{34}H_{44}N_4O_9$ (652.31): C, 62.56; H, 6.79; N, 8.58. Found: C, 62.48; H, 6.72; N, 6.86.

_t_-Boc-[(3*S*,4*R*,1'*S*)-AOMPC]$_3$-OMe, (8).



To a solution of compound **6** (784 mg, 2.0 mmol) in DCM (5 mL), TFA (2 mL, 3.06 g, 26.8 mmol) was slowly added and the clear solution was stirred for 1 h at rt. Volatiles were then eliminated under reduced pressure, and the residue was washed with diethyl ether, to give the trifluoroacetate **7** (812 mg, quantitative yield) as a low-melting solid. Then, to a solution containing Et_3N (0.128 mL, 1.0 mmol) and the trifluoroacetate **7** (666 mg, 1.0 mmol) in dry DCM (5 mL) at 0 °C, compound **4** (378 mg, 1.0 mmol) dissolved in dry DCM (2 mL) was added, followed by EDAC (210 mg, 1.1 mmol). The temperature was raised to rt and the mixture was stirred for an additional 12 h. Water (15 mL) was then added, and the mixture extracted with ethyl acetate (3 x 30 mL). The organic layer was dried over Na_2SO_4, filtered and concentrated under reduced pressure and the residue was purified by silica gel chromatography (ethyl acetate 100%) to give the trimer **8** (611 mg, 67%) as a white solid: mp 134-136 °C; ^{1}H NMR (200 MHz, $CDCl_3$) δ 1.38 (s, 9H), 1.49 (d, *J* = 7.2 Hz, 3+3 H), 1.51 (d, *J* = 7.2 Hz, 3H), 2.58 - 2.87 (m, 2H), 3.05 - 3.35 (m, 6H), 3.51 (dd, *J* = 9.8 Hz, *J* = 9.8 Hz, 1H), 3.73 (s, 3H), 3.79 (s, 3H), 3.80 (s, 3H), 3.81 (s, 3H), 4.23 (dd, *J* = 5.9 Hz, *J* = 9.6 Hz, 1H), 4.71 (dd, *J* = 6.9 Hz, *J* = 8.9 Hz, 1H), 4.81 (m, 1H), 5.40 (m, 3 H + 1 NH), 6.78 - 6.92 (m, 6 ArH), 7.18 - 7.24 (m, 6 ArH + NH), 8.92 (d, *J* = 6.9 Hz, 1H, NH); ^{13}C NMR (50 MHz, $CDCl_3$) δ 16.1, 16.2, 16.3, 28.1, 41.5, 42.0, 43.5, 46.9, 48.2, 48.3,48.9, 49.3, 49.6, 52.2, 55.0, 55.1,55.2, 55.4, 563, 80.6, 113.8, 113.9, 127.8, 127.9, 128.0, 128.2, 131.0, 131.1, 156.4, 158.9, 159.0, 168., 169.1, 170.9, 171.1, 172.3; $[\alpha]_D$ -453.3 (c 0.3, $CHCl_3$); ESI-MS: *m/z* 913.4 $[MH]^+$, 935.3 $[M+Na]^+$. Anal. Calcd for $C_{48}H_{60}N_6O_{12}$ (912.43): C, 63.14; H, 6.62; N, 9.20. Found: C, 63.07; H, 6.71; N, 9.14.

***t*-Boc-[(3*S*,4*R*,1'*S*)-AOMPC]$_6$-OMe (1).**



The compound **8** (456 mg, 0.5 mmol) dissolved in MeOH (1 mL) was added to a 0.5 M NaOH solution (5 mL, 2.5 mmol) and then MeOH (1 mL) was added until a clear solution, that was stirred for an additional 12 h. MeOH was subsequently removed under reduced pressure, the aqueous solution was cooled to 0 °C and the pH of the aqueous layer adjusted to 1 by adding 6 M HCl. Eventually, the mixture was extracted with ethyl acetate (3 x 10 mL), the organic layer dried over Na_2SO_4, filtered, and removed under reduced pressure, to give *t*-Boc-[(3*S*,4*R*,1′*S*)-AOMPC]$_3$, **9** (427 mg, 90%) as a white solid: mp 76-78 °C; ESI-MS: *m/z* 899.4 $[MH]^+$, 921.3 $[M+Na)]^+$. Anal. Calcd for $C_{47}H_{58}N_6O_{12}$ (898.41): C, 62.79; H, 6.50; N, 9.35. Found: C, 62.70; H, 6.41; N, 9.43.

To a solution of compound **8** (456 mg, 0.5 mmol) in DCM (5 mL), TFA (0.5 mL, 0.77 g, 6.8 mmol) was slowly added and the clear solution was stirred for 1 h at rt. Volatiles were then eliminated under reduced pressure, and the residue was washed with diethyl ether, to give H-[(3*S*,4*R*,1′*S*)-AOMPC]$_3$-OCH_3 trifluoroacetate, **10** (440 mg, 95%) as a low-melting solid. Eventually, to a solution containing Et_3N (0.05 mL, 45 mg, 0.4 mmol) and the trifluoroacetate **10** (370 mg, 0.4 mmol) in dry DCM (5 mL) at 0 °C, the acid **9** (359 mg, 0.4 mmol) was added, dissolved in dry DCM (2 mL) followed by EDAC (210 mg, 1.1 mmol). The temperature raised to rt and the mixture was stirred for an additional 12 h. Water (15 mL) was then added, and the mixture was extracted with ethyl acetate (3 x 30 mL). The organic layer was dried over Na_2SO_4, filtered and removed *in vacuo* and the residue was purified by silica gel chromatography (ethyl acetate 100%) to give the hexamer **1** (440 mg, 65%) as a white solid; mp. 176-178 °C ; $[\alpha]_D$ -176.0 (c 0.6, $CHCl_3$); ESI-MS: *m/z* 1694.1 $[MH]^+$, 1716.1 $[M+Na]^+$. Anal. Calcd for $C_{90}H_{108}N_{12}O_{21}$ (1692.77): C, 63.82; H, 6.43; N, 9.92. Found: C, 63.75; H, 6.35; N, 10.03.

Nuclear Magnetic Resonance of Hexamer 1

Experimental

The NMR spectra of hexamer **1** were acquired in $CDCl_3$ (4 mM solution), at 298 K, on a Bruker Avance DMX 600 spectrometer (600.09 MHz proton frequency) equipped with a 5 mm TXI probe. Data processing was performed on an Indy workstation using the XWIN-NMR software. ^{1}H and ^{13}C resonances assignment was obtained using 2D homonuclear correlation spectroscopy [Double-quantum filtered correlation (DQF-COSY),[1] total correlation (TOCSY) [2] and rotating-frame Overhauser enhancement (ROESY) [3] experiments] and 2D heteronuclear correlation spectroscopy [heteronuclear multiple quantum correlation (HMQC) [4] and heteronuclear multiple bond correlation (HMBC) [5]]. The proton and carbon chemical shifts were referenced to the residual solvent signal at 7.26 ppm (^{1}H) and at 77.0 ppm (^{13}C).

The ^{1}H 1D spectrum was acquired with 8 transients, a spectral width of 6613 Hz, 32k data points. DQF-COSY, TOCSY and ROESY spectra were recorded using a total of 512 experiments, of 1, 8 and 16 scans, respectively; 4k data points were recorded in the acquisition dimension (F2) and the spectral width was 11 ppm. The data was zero-filled in both dimensions to final sizes of 4k×1k points. Multiplication with a gaussian function in F2 and a sin function (DQF-COSY) or $\cos^2$ function (ROESY and TOCSY) in F1 was performed prior to 2D Fourier transformation. The TOCSY mixing time was 70 ms. In the ROESY experiment, a spin lock sequence of small flip angle pulses was applied for a mixing period of 300 ms.[6]

The ^{1}H-^{13}C HMQC spectrum was acquired using a total of 200 experiments of 8 scans each. The TPPI method was used, [7] and ^{13}C-decoupling during acquisition was achieved with a GARP sequence. The spectral width was 11 ppm in F2, 150 ppm in F1. The data matrix was zero-filled in both dimensions to a final size of 1k×1k points. Multiplication with a gaussian function (F2) and a $\cos^2$ function (F1) was performed prior to 2D Fourier transformation.

The ^{1}H-^{13}C HMBC was acquired using a total of 300 experiments of 100 scans each and processed in the magnitude mode. The evolution delay for ^{1}H-^{13}C long range coupling was 67 ms; the spectral width was 11 ppm in F2, 250 ppm in F1. The data was zero-filled in both dimensions to final sizes of 2k×1k points. Multiplication with a gaussian function (F2) and a sin function (F1) was performed prior to 2D Fourier transformation.

Conformational analysis

Homonuclear and heteronuclear two-dimensional NMR experiments were used to obtain the ^{1}H and ^{13}C backbone resonance assignment of **1** (Table1).

Table 1. ^{1}H and ^{13}C backbone resonance assignment of hexamer **1** in $CDCl_3$ at 298 K.

Residue		1	2	3	4	5	6
C**H**$_3$ *Boc*	1.35						
CH$_3$ *Boc*	28.5						
Cquat*Boc*	81.5						
N**H**		5.39	6.53	7.83	8.48	8.33	8.42
C**H**(3)		4.35	4.58	4.69	4.78	4.94	4.98
CH(3)		57.3	56.1	55.8	55.5	55.2*	55.2*
C**H**(4)		2.43	2.55	2.73	2.92	2.98	3.32
CH(4)		49.9	49.3	48.3	48.1	46.4	44.3
CH(5)		2.47 3.27	2.80 3.43	2.94 3.68	3.25 3.71	3.18 3.82	3.38
CH(5)		41.1	42.3	41.5	42.3	41.7	42.4
CH(3)**C**O		168.2	168.2	168.2	169.5	169.9	169.9
CH(4)**C**O		172.5	171.9	171.7	170.9	171.4	173.1

* Superimposed peaks

The secondary structure was investigated using the information from the ROESY experiment.

The assignment of the backbone proton resonances of each monomer was achieved with the DQF-COSY experiment and confirmed by the TOCSY spectrum; these peaks were well resolved and the pattern of signals pertaining to each monomer was easily identified. The 1'-(4-methoxyphenylethyl) side chains give overlapping signals and were not sequentially assigned.

An analysis of the ROESY spectrum was attempted, but the high number of cross peaks did not allow us to distinguish between sequential and medium-range cross peaks. Sequence-specific assignment was obtained comparing $^{1}H/^{13}C$ HMBC experiments with HMQC and DQF-COSY spectra. Detection of long-range couplings between **C**ONH(*i*) and **H**4(*i*) and between **C**O(*i)* and N**H**(*i*+1) in the HMBC experiment was the key to sequence-specific assignment.

The NH protons of the various residues are quite different, with chemical shifts spread between 5.39 and 8.48 ppm. Complete backbone assignment revealed that the amide NH protons resonating at higher chemical shift values belong to residues 3-6, and are probably involved in hydrogen bonds. In fact, in a non-hydrogen bonding solvent such as $CDCl_3$ NH hydrogen bond formation causes a downfield shift of 2-3 ppm.[8] To exclude intermolecular hydrogen bond formation, the IR spectra of **1** in $CDCl_3$ at different concentrations in the range 10^{-2} - 10^{-4} M were recorded. The shape of the spectra in the N—H stretching (amide A) region is very similar in all cases, indicating that no aggregation occurs (data not shown).

ROESY data strongly supported the existence of a secondary structure. Some medium-range ROEs are indicative a 12-helical conformation: H3(i) — NH(i+3), H5(i) — NH(i+3), H3(i) — NH(i+2) and H3(i) — H4(i+2).

References

1 B. Ancian, I. Bourgeois, J. F. Dauphin and A. A. Shaw, *J. Magn. Reson.*, 1997, **A 125**, 348-354.

2 A. Bax and D. G. Davis, *J. Magn. Reson.*, 1985, **65**, 355-360.

3 A. Bax and D.G. Davis, *J. Magn. Reson.*, 1985, **63**, 207-213.

4 A. Bax, R. H. Griffey and B. L. Hawkins, *J. Magn. Reson.*, 1983, **55**, 301-315.

5 A. Bax and M. F. Summers, *J. Am. Chem. Soc.*, 1986, **108**, 2093-2094.

6 H. Kessler, C. Griesinger, R. Kerssebaum, K. Wagner and R. R. Ernst, *J. Am. Chem. Soc.*, 1987, **109**, 607-609.

7 D. Marion and K. Wüthrich, *Biochem. Biophys. Res. Commun.*, 1983, **113**, 967-974.

8 J. D. Fisk, D. R. Powell and S. H.Gellman, *J. Am. Chem. Soc.*, 2000, **122**, 5443-5447.

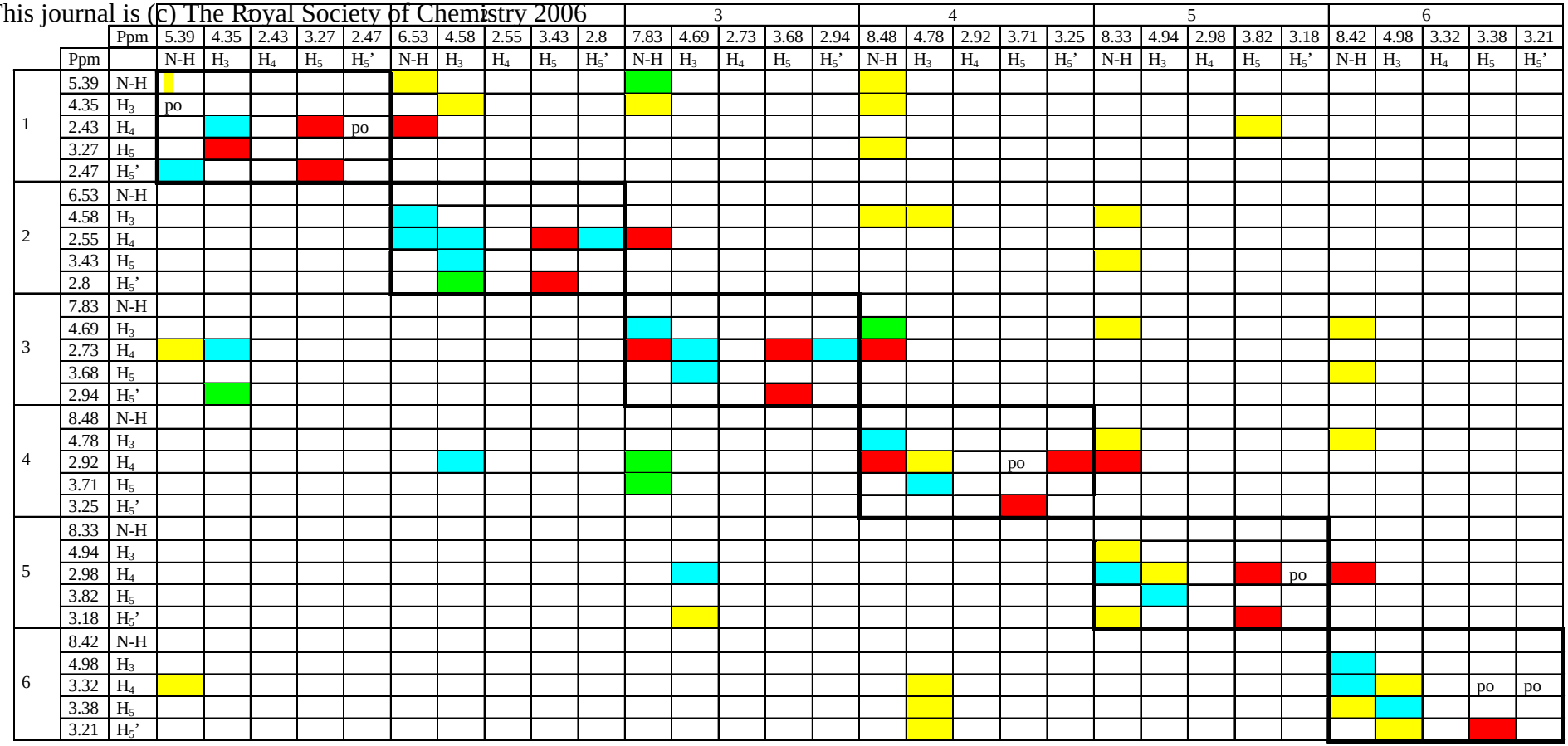


Figure. Intensities of the ROESY peaks (arbitrary units) detected for **1** in $CDCl_3$.

Molecular Dynamics Simulations

Computational details

The molecular dynamics simulations (MD) were performed on a Silicon Graphics Octane2 workstation using InsightII/Discover [1,2] and Macromodel [3] software packages and the conformational space was explored within both packages. Either simulated annealing protocol [4] and conformational sampling were used in order to generate stable conformations using NOE-derived interproton distances as restraints (force constant 100 kcal/mol). A cluster analysis of all the given structures was performed superimposing backbone atoms. Both programs and protocols were used, leading to the same results, and here we report the data obtained within Macromodel using AMBER* force field [5] since this package allowed us a more detailed cluster analysis of the final structures. In addition, unconstrained MD simulations were undertaken and the results obtained were compared with those arising from costrained simulations.

Thus, 40 constrained and 10 unconstrained MD simulations were performed and for each MD run 100 structure conformations were saved and optimized. A further cluster analysis of these structures was performed based on the RMS value for superimposing backbone atoms, and structures displaying a maximum backbone RMS < 0.6 were grouped into the same family. Before and after running MD simulations, the system was energy minimized using the same distance restraints. Typically, dynamics were initialized for 1 ps raising the temperature from 300 K to 700 K, then equilibrated at 700 K for 2 ns with time steps of 1 fs, then resumed at the final temperature of 300 K for 20 ps.

References

1. InsightII/Discover 2000, Accelrys, San Diego, USA.

2. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.*, 1983, **4**, 187-217.

3. F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson and W. C. Still, *J. Comput. Chem*., 1990, **11**, 440-467.

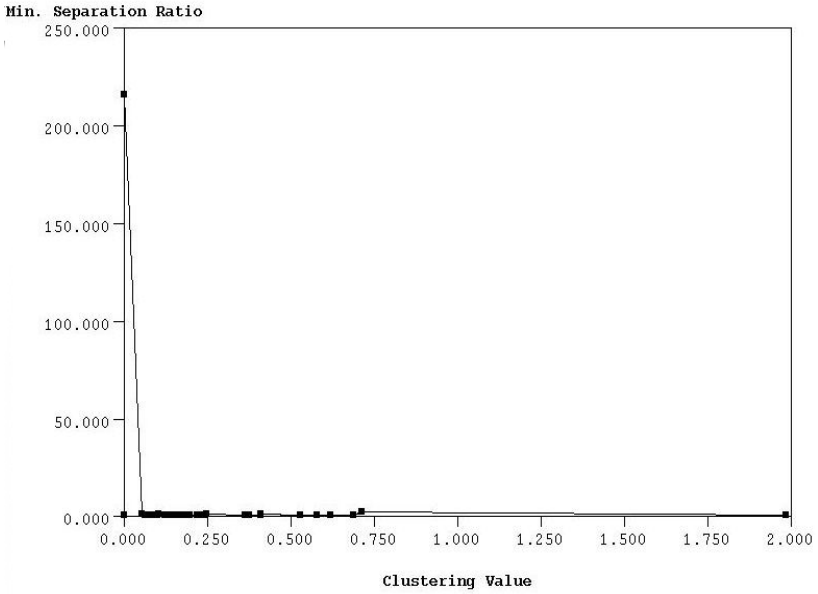
4. J. K. Young, R. P. Hicks and C. Anklin, *Biopolymers*, 1994, **34**, 1449-1462.

5. S. J. Weiner, P. A. Kollman, D. T. Nguyen and D. A. Case, *J. Comput. Chem*., 1986, **7**, 230-252.

<u>Restrained MD simulations:</u>

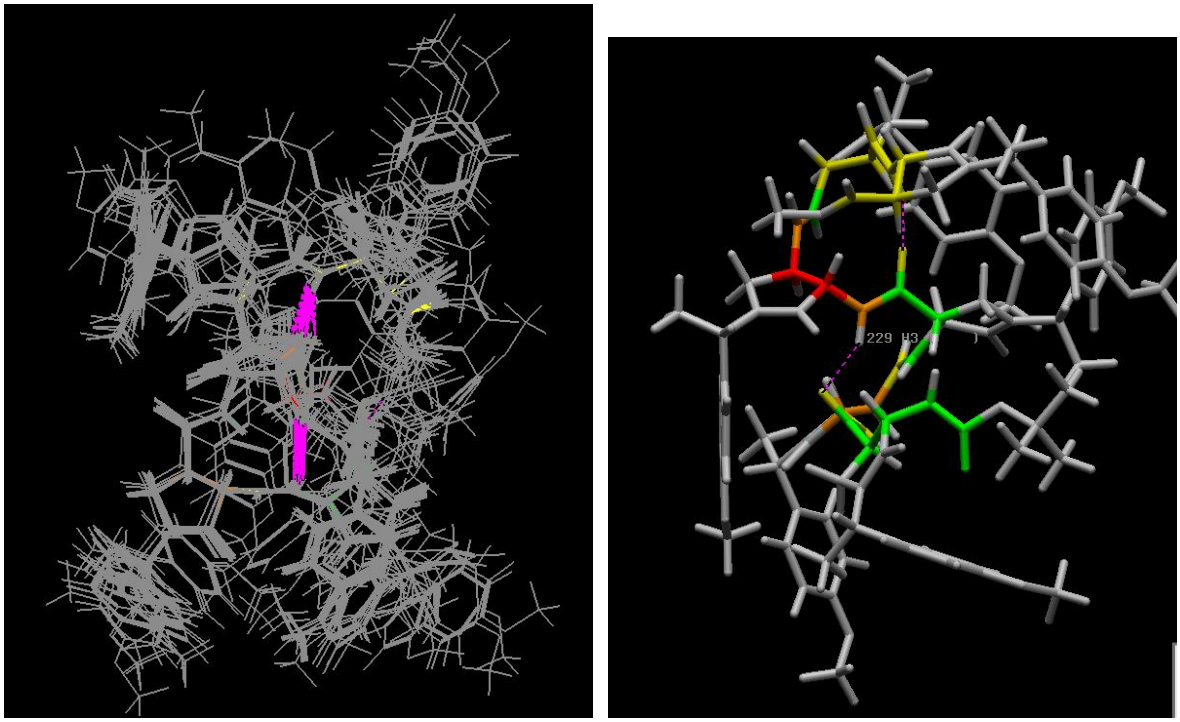
The cluster analysis (clustering level 42, RMS 0.53) showed the overall 48 structures grouped into 7 families (clusters). For most clusters a backbone arrangement perfectly compatible with a 12-helix structure was observed and this conformational behaviour seems to be strongly favoured from an

energetic point of view [Residues are numbered starting from the N-terminus. Reference backbone torsion angles: $\phi = 95$; $\theta = -94$; $\psi = 103$. See: (a) Roy, R. S.; Balaram, P. *J. Peptide Res.* **2004**, *63*, 279-289; (b) Cheng, R. P.; Gellman, S. H.; DeGrado, W. F. *Chem. Rev.* **2001**, *101*, 3219-3232]. The main differences in energy between the conformers belonging to clusters having a 12-helix structure are due to the lateral chains orientation.



Cluster No. 1

Members: 13. (lowest energy minimum cluster)

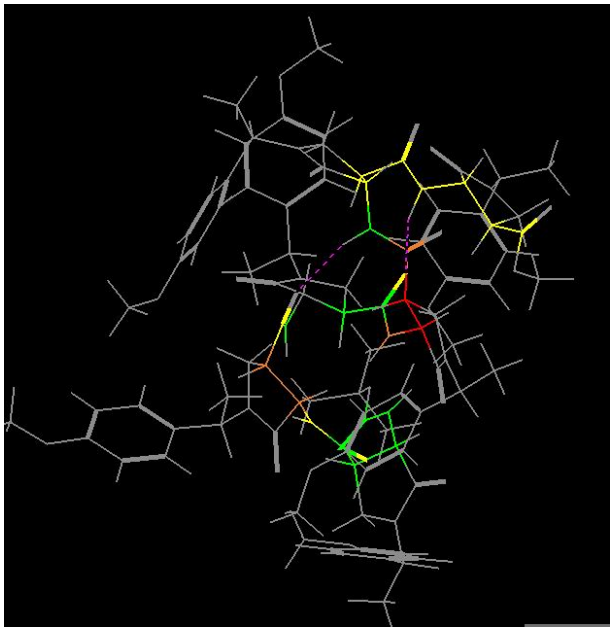


2 hydrogen-bonding interactions are present [(res6)NH···O=C(res3) and (res4)NH···O=C(res1)], leading to 12-membered rings.

Residue	ϕ	θ	ψ
1	-65.0	-88.7	132.3
2	152.7	-83.6	40.9
3	165.1	-86.8	100.7
4	91.5	-105.4	113.7
5	77.1	-87.3	102.4
6	145.3	-97.5	82.3

Cluster No. 2

Members: 1.

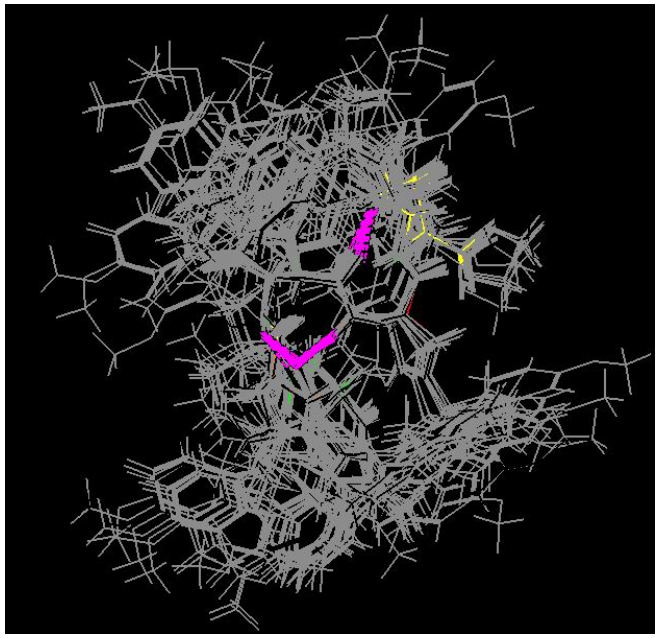


2 hydrogen-bonding interactions are present [(res6)NH···O=C(res3) and (res4)NH···O=C(res1)], leading to 12-membered rings.

Residue	ϕ	θ	ψ
1	-58.5	-104.1	108.8
2	167.8	-70.8	79.3
3	114.0	-84.9	84.2
4	84.1	-83.9	122.8
5	79.3	-102.7	108.8
6	140.9	-92.5	124.0

Cluster No. 3

Members: 15.

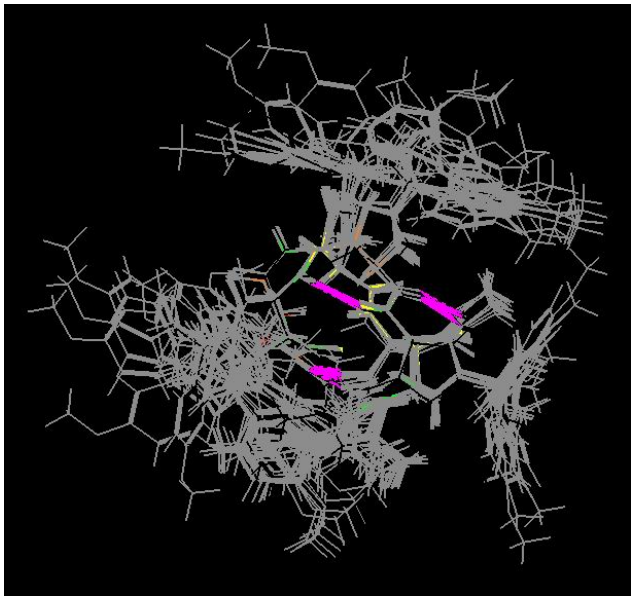


2 hydrogen-bonding interactions are present [(res6)NH···O=C(res3) and (res4)NH···O=C(res1)], leading to 12-membered rings; an additional hydrogen-bonding interaction takes place between (res3)NH···O=C(*t*-Boc).

Residue	ϕ	θ	ψ
1	-77.7	-81.8	-130.7
2	43.7	-98.9	108.3
3	104.0	-82.9	87.8
4	66.7	-94.6	130.0
5	75.3	-89.8	125.8
6	96.9	-92.0	125.8

Cluster No. 4

Members: 14.



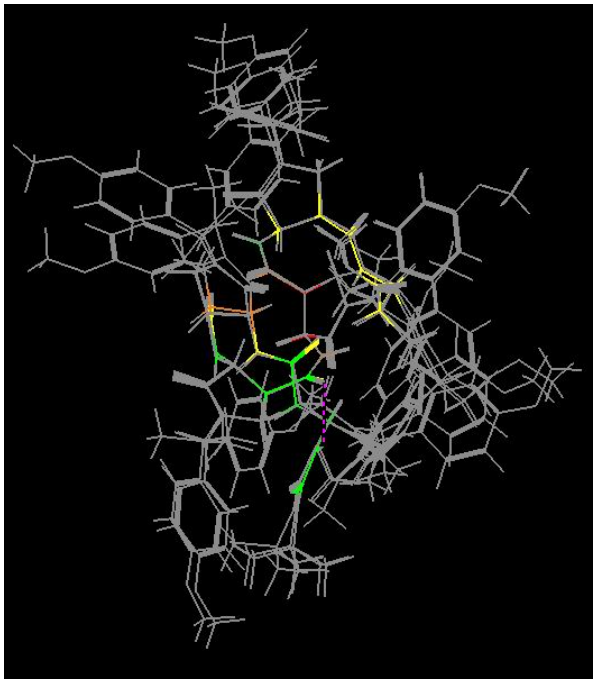
3 hydrogen-bondings interactions are present [(res3)NH···O=C(res1), (res2)NH···O=C (*t*-Boc) and (res4)NH···O=C(*t*-Boc)] but 12-termed rings are missing.

Residue	ϕ	θ	ψ
1	-35.7	-87.9	121.4
2	59.2	-139.6	59.4
3	133.2	-129.4	100.2
4	-75.2	-85.7	110.7
5	-88.3	-117.0	50.2
6	-94.8	-86.4	-69.0

Cluster No. 5

Members: 3.

Hydrogen-bonding interactions are missing.

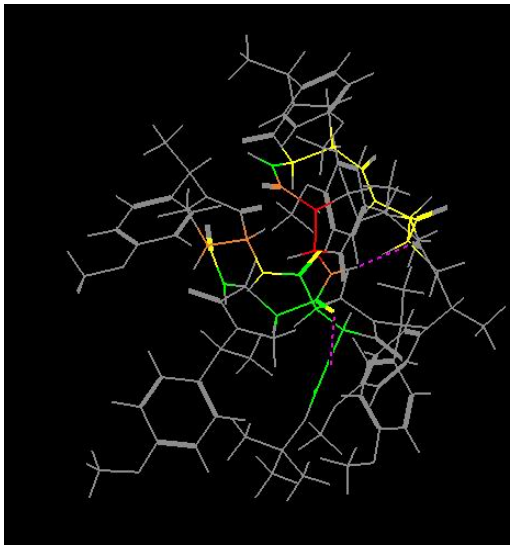


Residue	ϕ	θ	ψ
1	-64.7	-107.6	115.9
2	172.9	-91.7	72.4
3	132.4	-93.4	-61.7
4	105.9	-132.8	143.7
5	58.6	-132.2	71.9
6	93.6	-89.1	76.0

Cluster No. 6

Member: 1.

2 hydrogen-bonding interactions are present [(res3)NH···O=C(res1) and (res2)NH···O=C(*t*-Boc)] but 12-membered rings are missing.

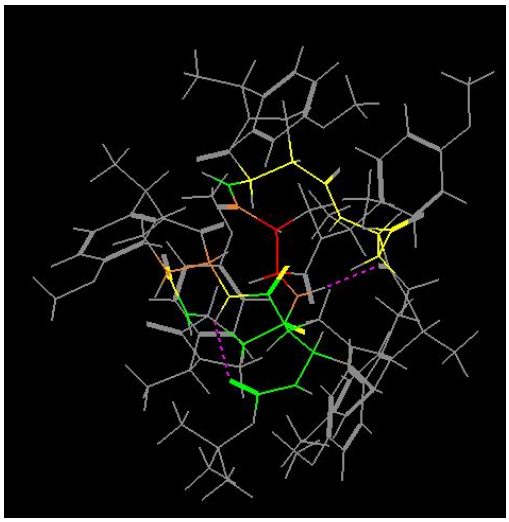


Residue	ϕ	θ	ψ
1	-60.6	-102.4	-92.9
2	159.8	-99.4	74.8
3	144.8	-101.6	-52.0
4	117.3	-129.6	133.9
5	61.9	-134.5	91.8
6	82.9	-92.9	110.2

Cluster No. 7

Members: 1.

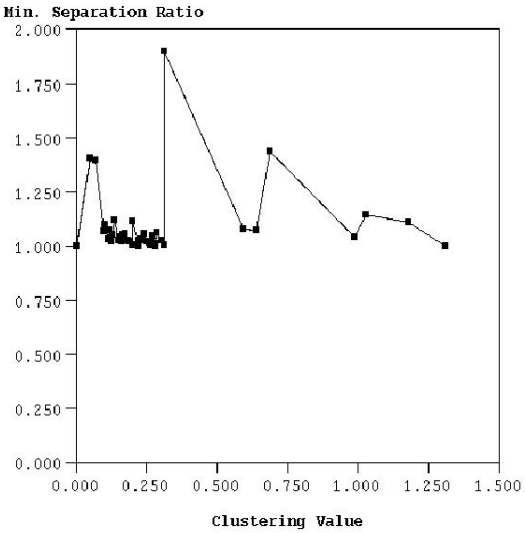
2 hydrogen-bonding interactions are present [(res3)NH···O=C(res1) and (res2)NH···O=C(*t*-Boc)], but 12-membered rings are missing.



Residue	ϕ	θ	ψ
1	16.1	-94.6	106.7
2	166.0	-99.1	73.6
3	46.4	-99.8	-57.9
4	113.2	-130.9	134.5
5	62.3	-134.3	78.1
6	78.2	-95.0	119.5

Unrestrained MD simulations

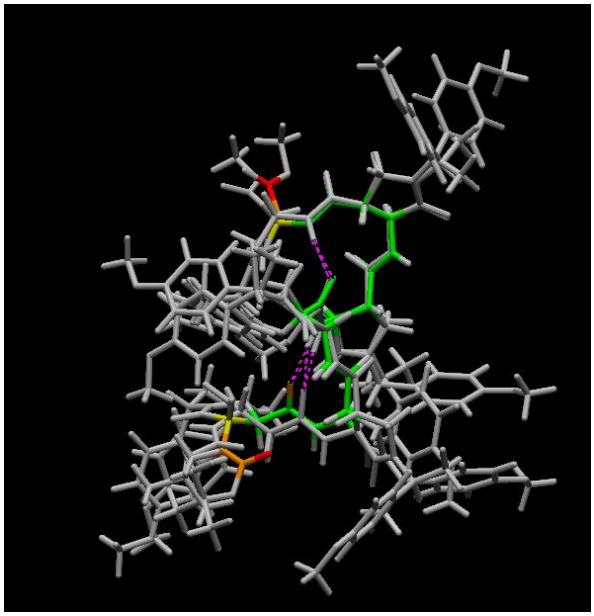
Within this group of unrestrained MD simulations, a conformer distribution was obtained close to that of the restrained ones, the 12-helix structure of the backbone being strongly favoured in the absence of restrictions, too. The data collected refer to cluster level 38, RMS 0.52, containing 44 structures grouped into 7 families.

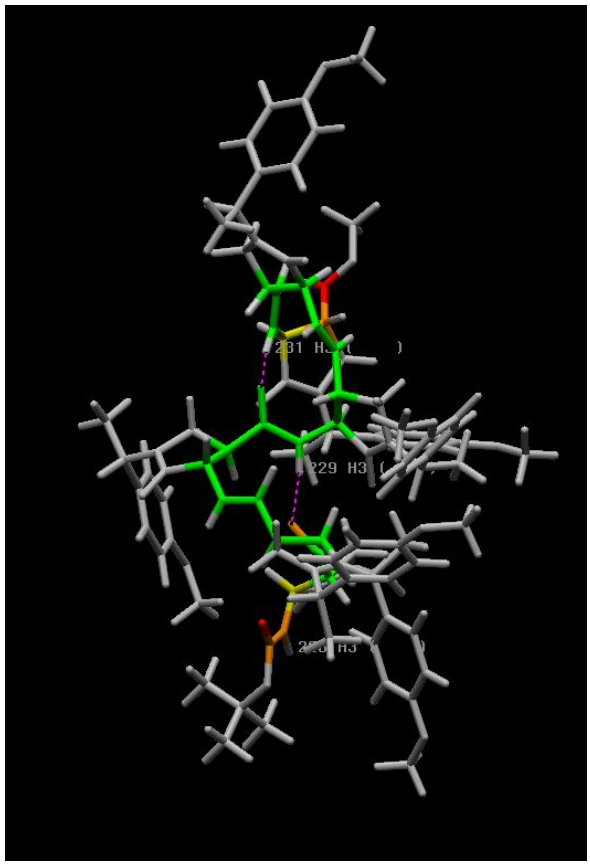


Cluster No. 1

2 hydrogen-bonding interactions [(res6)NH···O=C(res3) and (res4)NH···O=C(res1)] leading to 12-membered rings. These structures are compatible with a 12-helix secondary structure.

Members: 2.



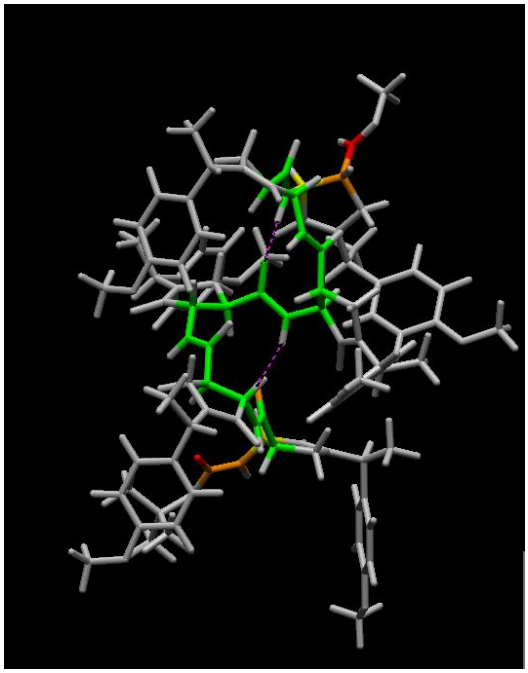


Residue	ϕ	θ	ψ
1	-111.1	-99.3	130.5
2	56.7	-103.9	135.0
3	75.3	-87.9	98.2
4	46.7	-102.8	144.7
5	75.9	-91.5	102.4
6	50.8	-107.7	139.3

Cluster No. 2

2 hydrogen bonding interactions [(res6)NH¨O=C(res3) and (res4)NH¨O=C(res1)], leading to formation of 12-membered rings. These structures are compatible with a 12-helix secondary structure.

Members: 5.

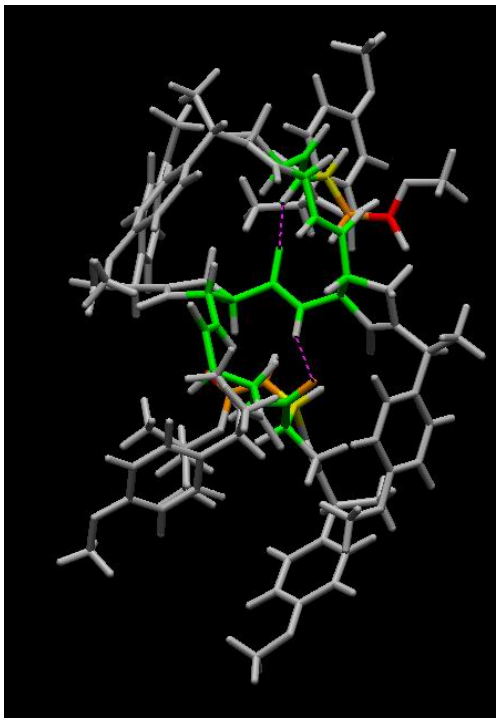


Residue	ϕ	θ	ψ
1	-129.1	-115.2	156.7
2	61.0	-101.0	136.9
3	72.1	-89.6	114.3
4	59.2	-96.2	144.1
5	68.9	-97.6	106.7
6	56.6	-140.6	142.8

Cluster No. 3

2 hydrogen bonding interactions [(res6)NH¨O=C(res3) and (res4)NH¨O=C(res1)], leading to formation of 12-termed rings. These structures are compatible with a 12-helix secondary structure.

Members: 14.

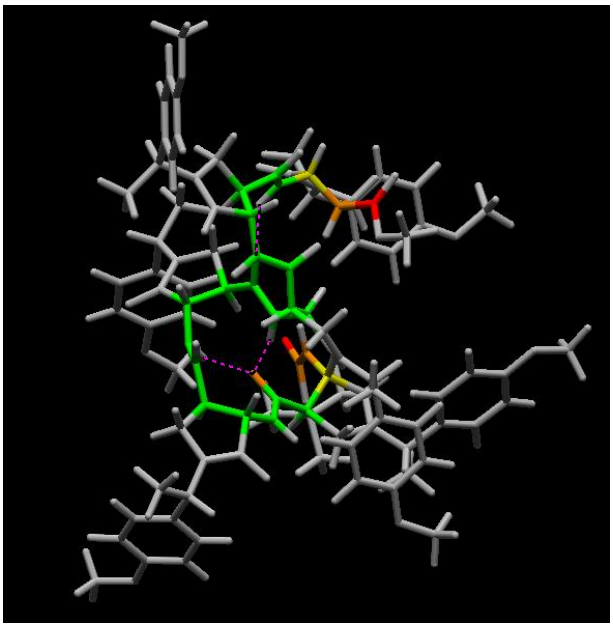


Residue	ϕ	θ	ψ
1	139.1	-76.1	124.5
2	152.0	-82.9	41.6
3	173.0	-86.1	94.2
4	85.2	-93.7	120.9
5	72.2	-90.8	121.0
6	115.3	-92.6	-30.5

Cluster No. 4

2 hydrogen bonding interactions [(res6)NH¨O=C(res3) and (res4)NH¨O=C(res1)], leading to 12-membered rings. These structures are compatible with a 12-helix secondary structure.

Members: 24.

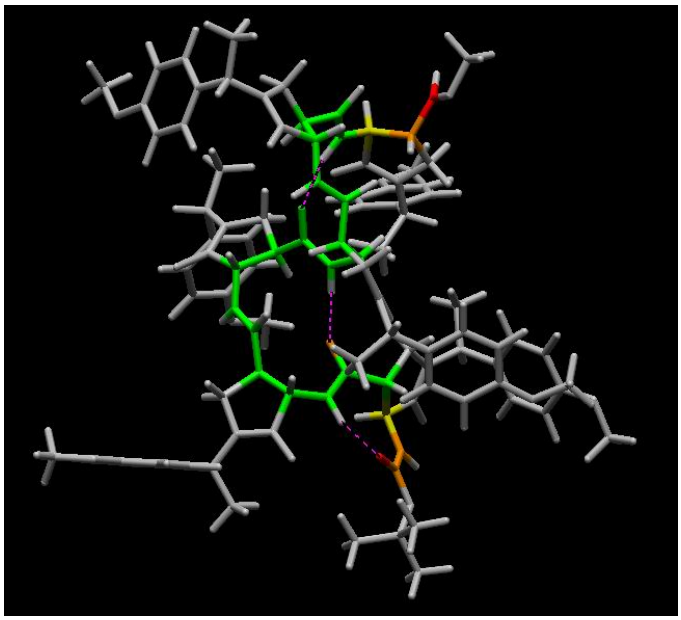


Residue	ϕ	θ	ψ
1	111.8	-76.2	-136.5
2	39.0	-96.7	112.8
3	98.1	-81.8	91.0
4	68.7	-92.4	127.6
5	76.3	-88.1	124.4
6	94.9	-90.8	129.1

Cluster No. 5

2 hydrogen bonding interactions [(res6)NH···O=C(res3) and (res4)NH···O=C(res1)], leading to 12-membered rings; an additional hydrogen bonding interaction takes place [(res2)NH···O=C(res1, *t*-Boc). These structures are compatible with a 12-helix secondary structure.

Members: 1.

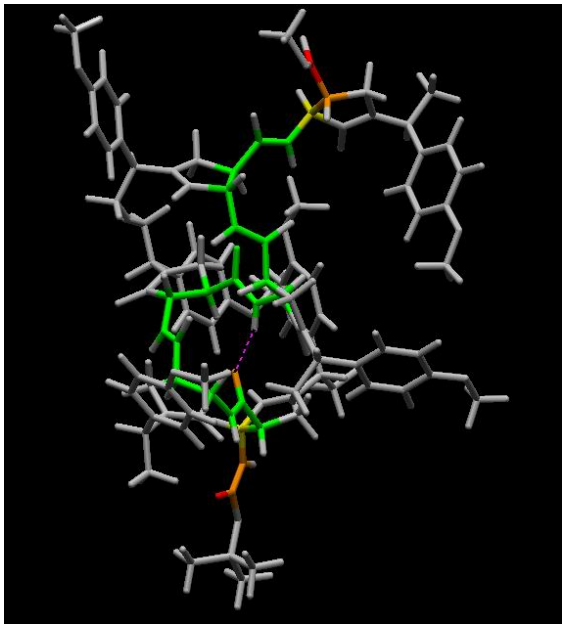


Residue	ϕ	θ	ψ
1	-142.5	-111.6	165.5
2	40.4	-98.9	112.8
3	73.5	-100.3	98.5
4	50.6	-102.0	144.5
5	72.0	-88.8	112.2
6	60.2	-107.3	131.3

Cluster No. 6

1 hydrogen bonding interaction occurs [(res4)NH···O=C(res1)], leading to a sole 12-termed ring. A helix-12 secondary structure takes place only in the C-terminus.

Members: 1.

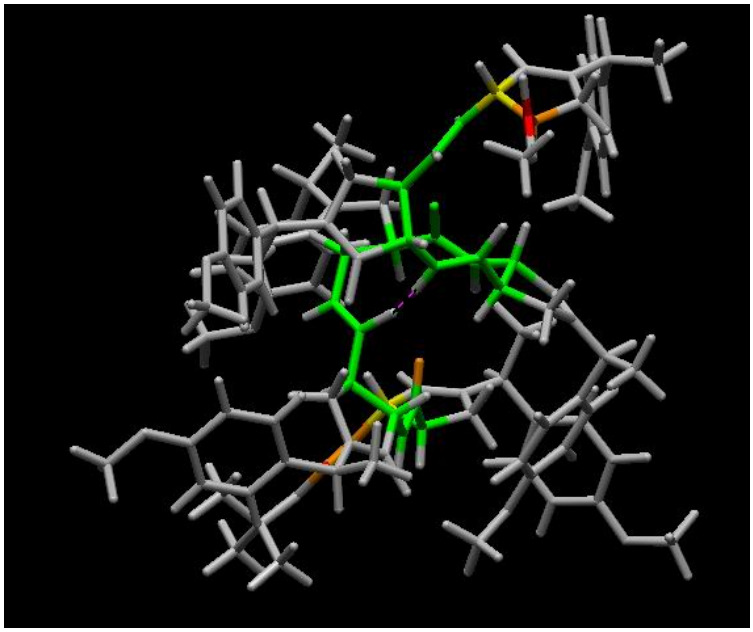


Residue	ϕ	θ	ψ
1	-124.1	-146.8	123.8
2	59.5	-103.0	132.4
3	76.4	-88.9	100.6
4	61.3	-95.0	132.0
5	55.0	-125.7	123.5
6	56.1	-102.2	138.4

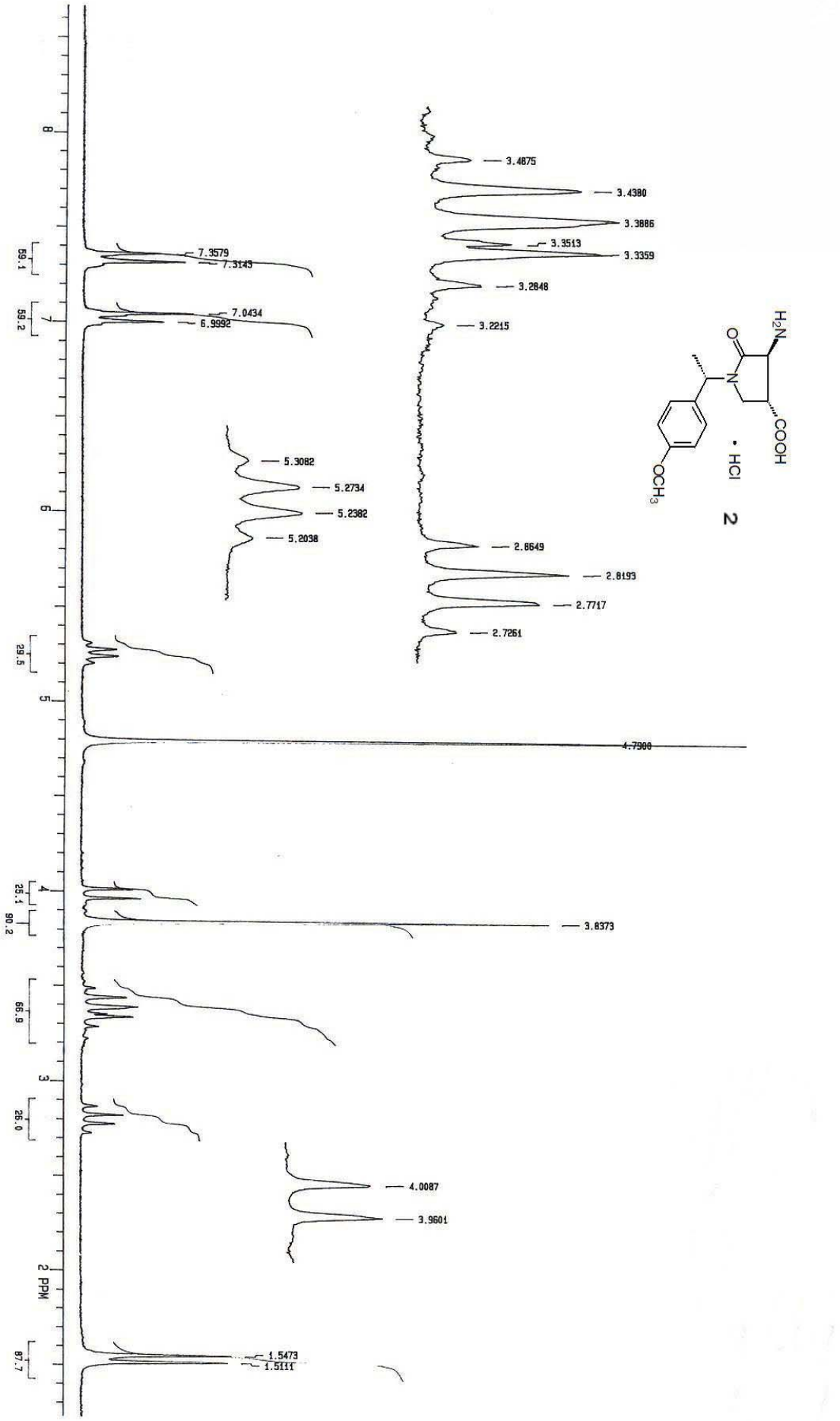
Cluster No. 7

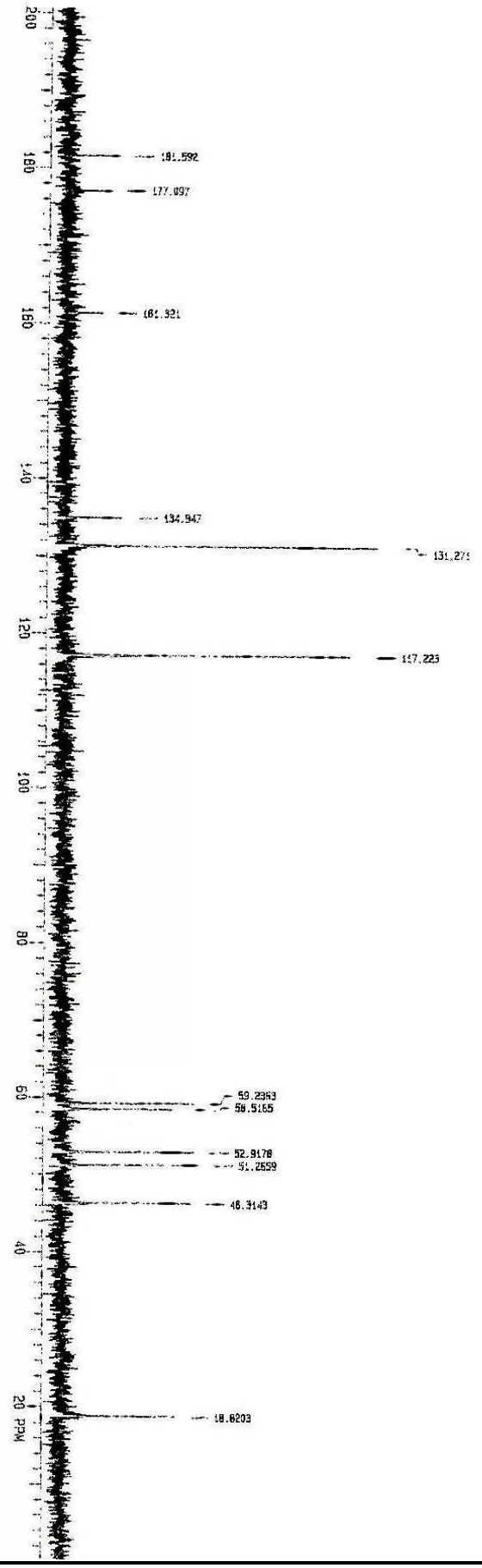
1 hydrogen bonding interaction occurs [(res5)NH···O=C(res2)] leading to a sole 12-termed ring. This conformer displays helix-12 secondary structure only in the middle of the sequence.

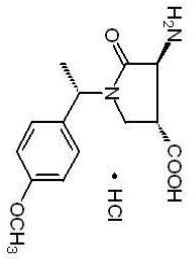
Members: 1.



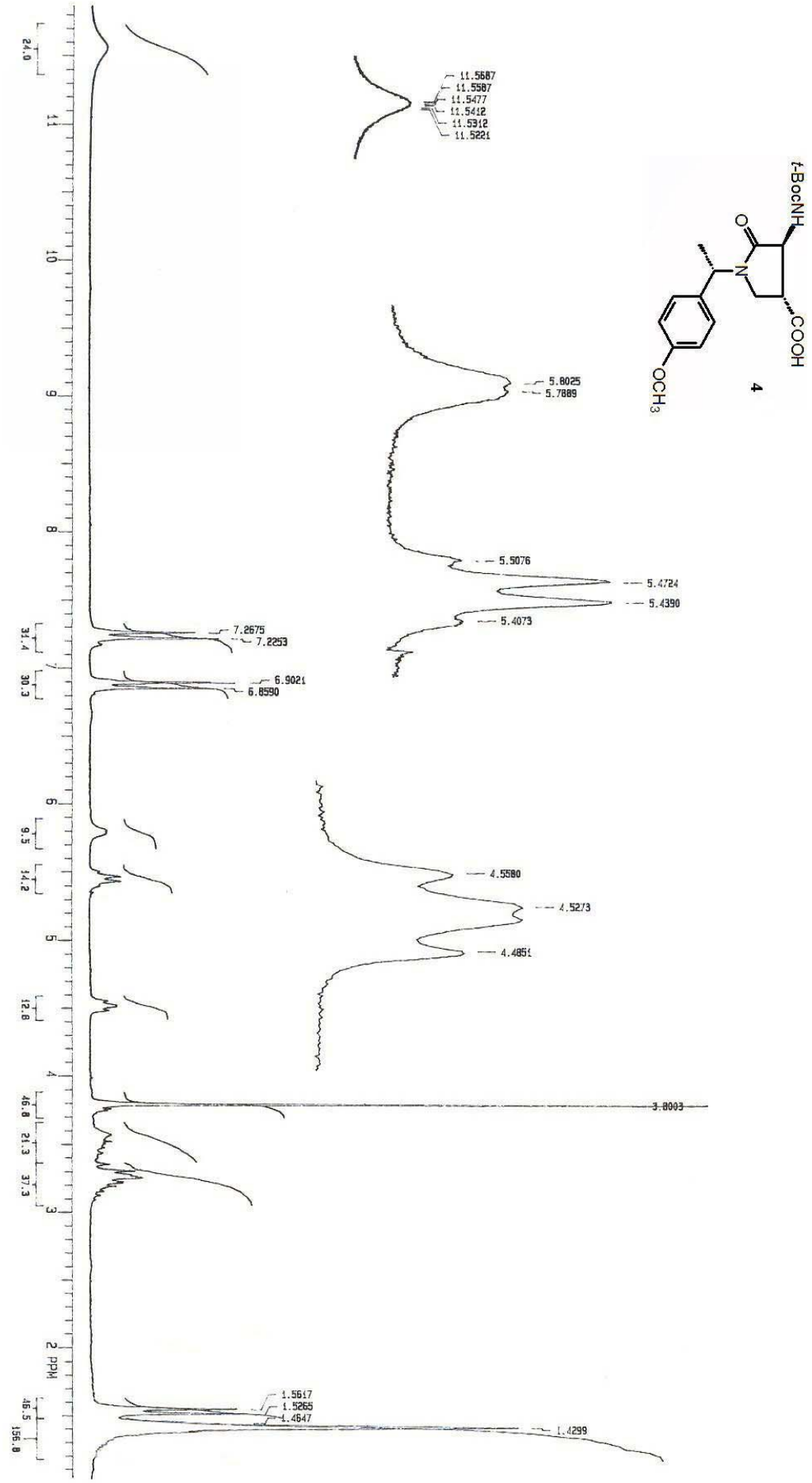
Residue	ϕ	θ	ψ
1	-129.3	-114.9	118.0
2	68.5	-97.9	132.4
3	64.3	-95.3	129.3
4	72.9	-91.6	114.4
5	60.4	-104.4	130.6
6	54.2	-109.3	136.4

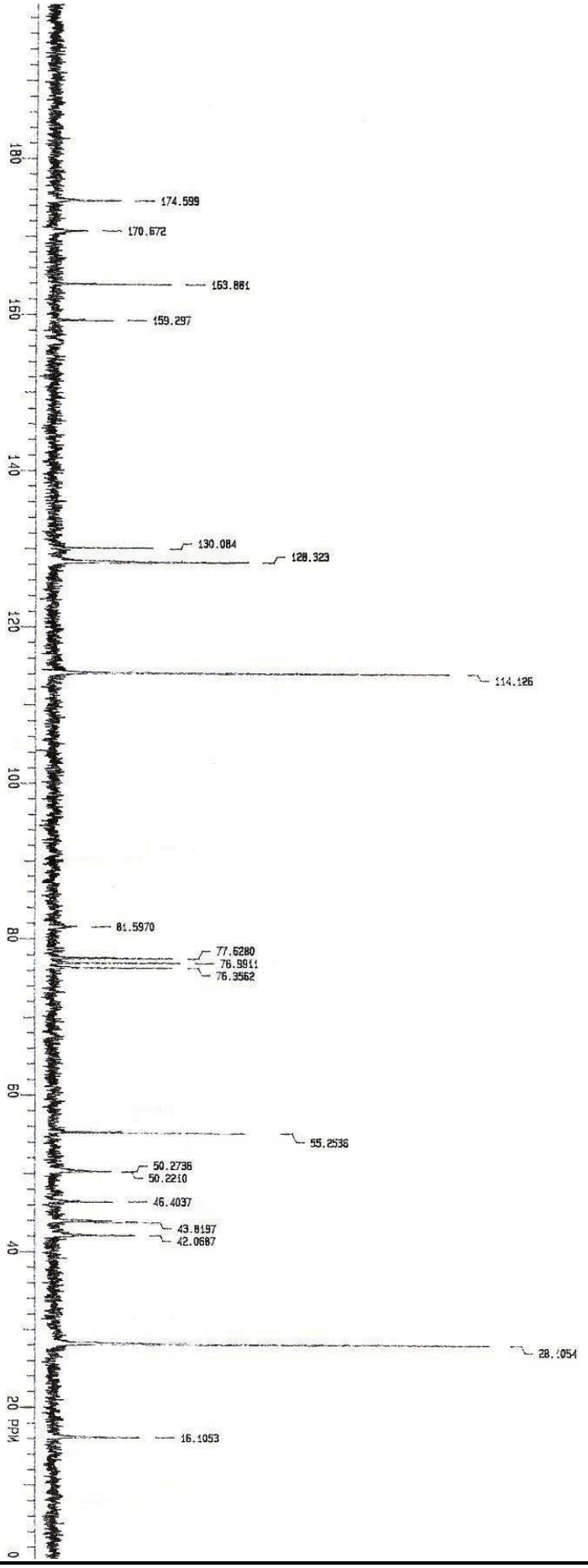

7.3579
7.3143
7.0434
6.9992
5.3082
5.2734
5.2382
5.2038
3.4875
3.4380
3.3886
3.3513
3.3359
3.2848
3.2215
2.8649
2.8193
2.7717
2.7261
4.7900
3.8373
4.0087
3.9601
1.5473
1.5111
59.1
59.2
29.5
25.1
90.2
66.9
26.0
87.7
8
7
6
5
4
3
2 PPM
H2N
O
N
COOH
OCH3
• HCl
2

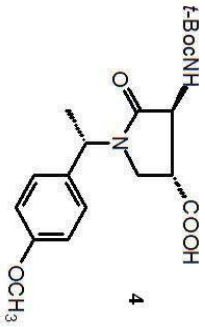

181.592
177.097
161.921
134.947
131.271
117.223
59.2363
58.5165
52.9178
51.2559
46.3143
18.8203
PPM


H2N
O
N
COOH
• HCl
OCH3

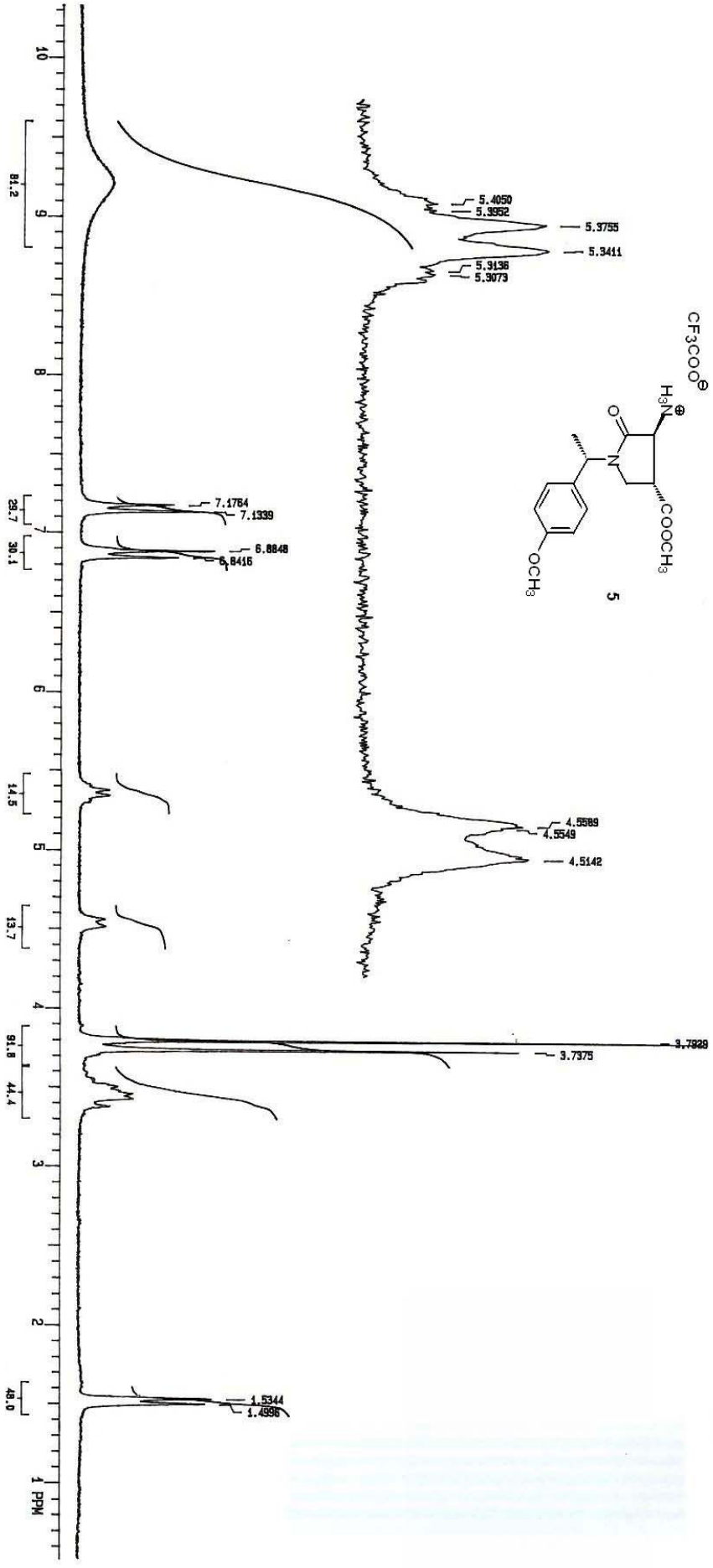
2

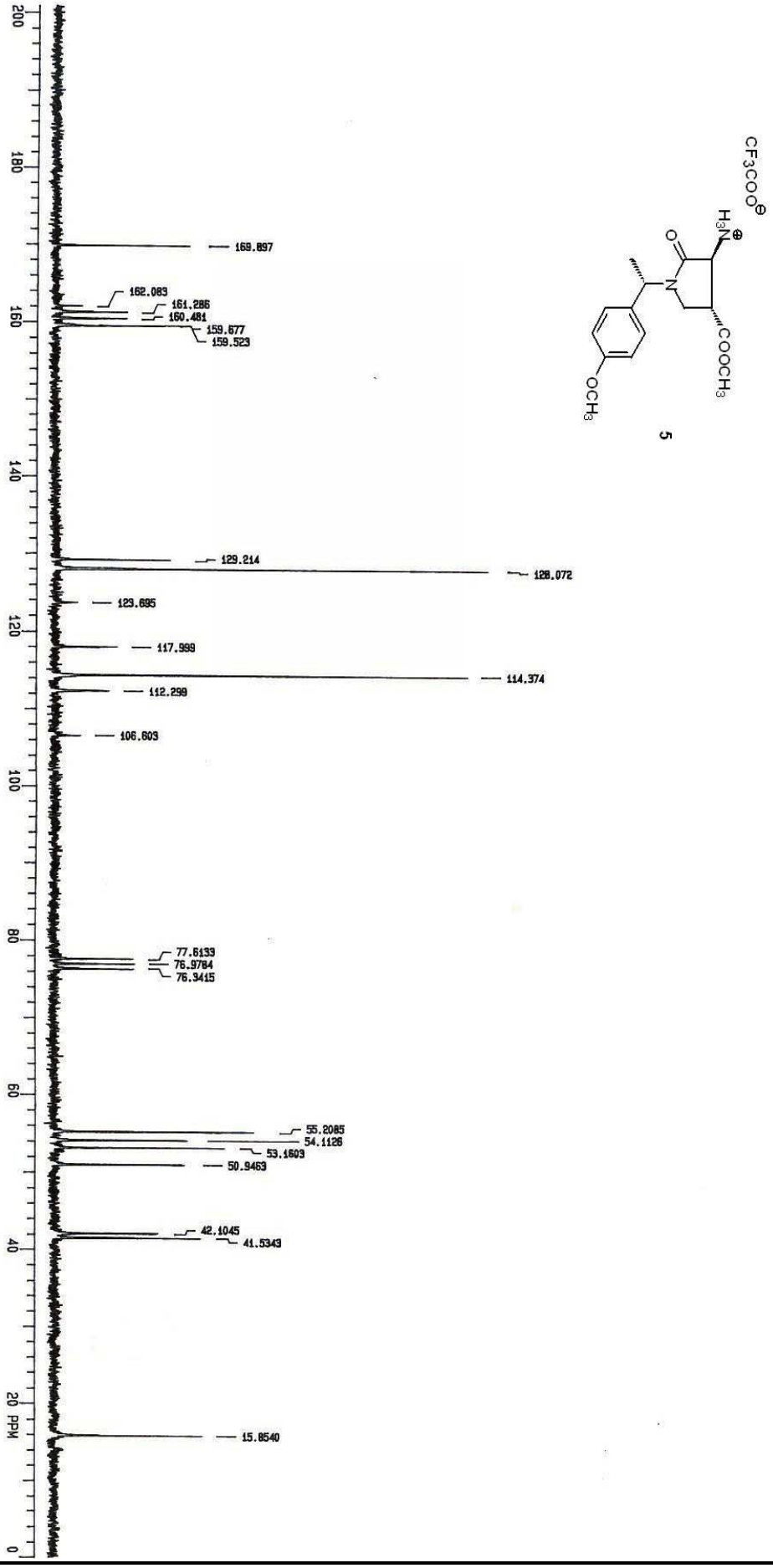

11.5687
11.5587
11.5477
11.5412
11.5312
11.5221
5.8025
5.7889
5.5076
5.4724
5.4390
5.4073
7.2675
7.2253
6.9021
6.8590
4.5580
4.5273
4.4851
3.8003
1.5617
1.5265
1.4647
1.4299
t-BocNH
COOH
OCH3
4
PPM

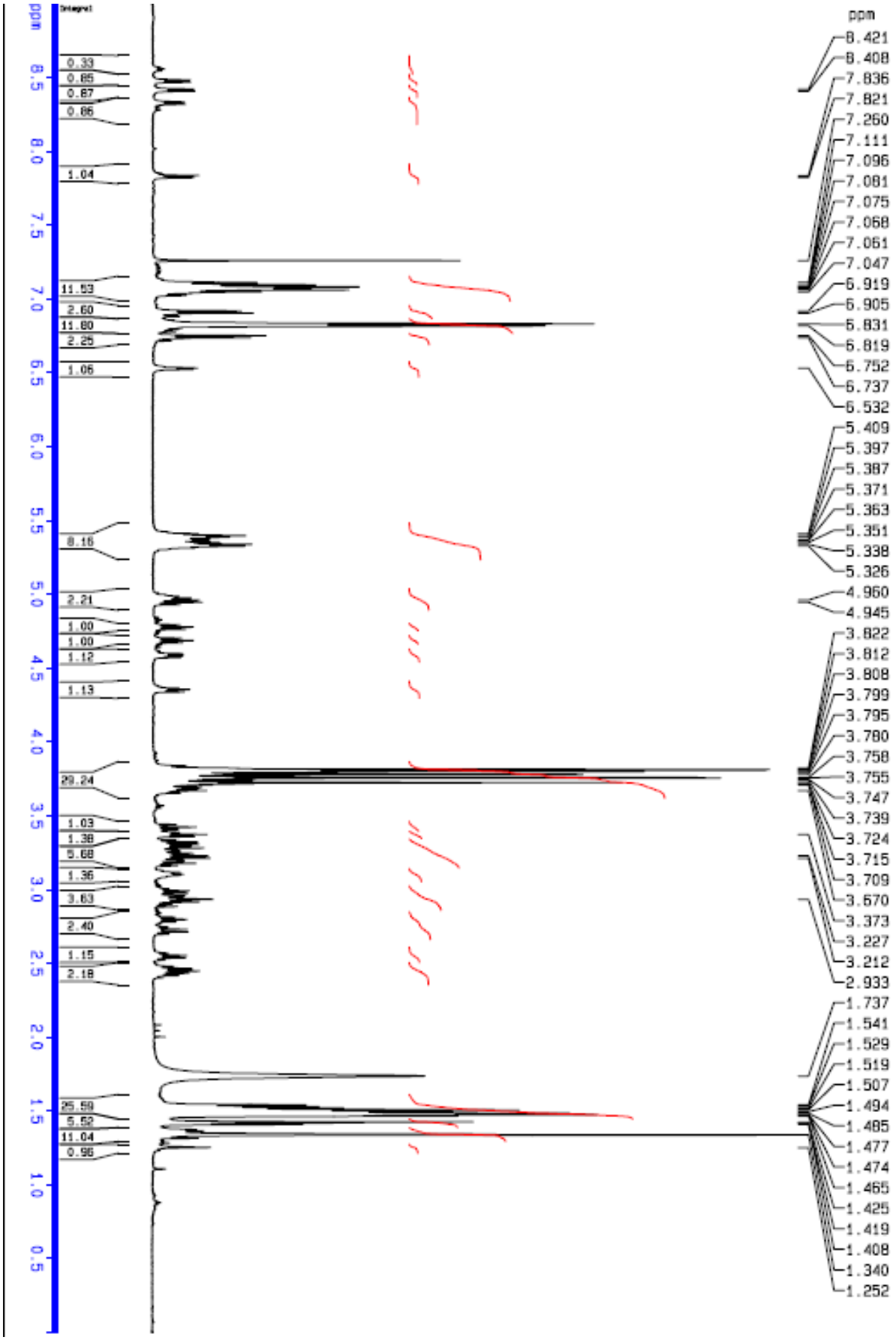

174.599
170.672
163.881
159.297
130.084
128.323
114.126
81.5970
77.6280
76.9911
76.3562
55.2536
50.2736
50.2210
46.4037
43.8197
42.0687
28.1054
16.1053
PPM


t-BocNH
O
N
COOH
OCH3

4


5.4050
5.3952
5.3755
5.3411
5.3136
5.3073
7.1764
7.1339
6.8848
6.8416
4.5589
4.5549
4.5142
3.7929
3.7375
1.5344
1.4996
CF3COO⊖
H3N⊕
COOCH3
OCH3
5
PPM


169.897
162.083
161.286
160.481
159.677
159.523
129.214
128.072
123.695
117.999
114.374
112.298
106.603
77.6133
76.9784
76.3415
55.2085
54.1126
53.1603
50.9463
42.1045
41.5343
15.8540
PPM
CF3COO⊖
H3N⊕
OCH3
COOCH3
5



^{1}H NMR of compound **1**, 600 MHz, 298 K