Electronic supplementary information

Promotion of the collagen triple helix in a hydrophobic environment

by Vladimir Kubyshkin[a] and Nediljko Budisa[a,b]

a Technical University of Berlin, Berlin, Germany

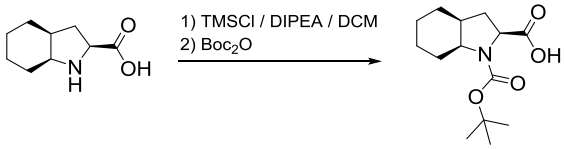
b University of Manitoba, Winnipeg, Canada

Index

For the graphical abstract, we used the ribbon structure generated from the crystal structure reported in M. D. Shoulders, K. A. Satyshur, K. T. Forest and R. T. Raines, *Proc. Natl. Acad. Sci. U.S.A.* 2010, **107**, 559-564 (main text reference 4). PDB entry: 3ipn.

Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

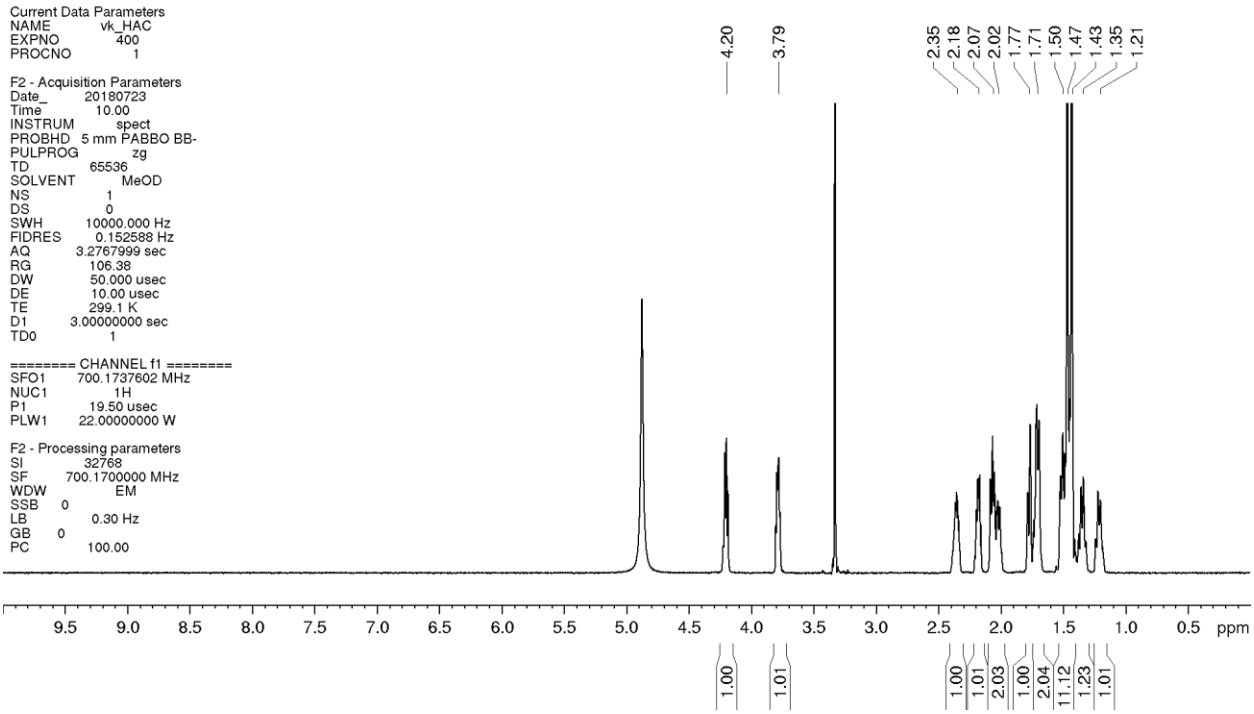
Solution phase peptide synthesis



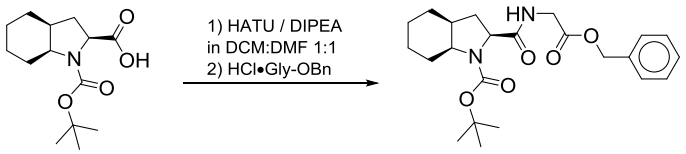
Oic Boc-Oic

Boc-Oic: (*2S,3aS,7aS*)-octahydroindole-2-carboxylic acid (2 g; 11.8 mmol; 1 equiv.) was mixed with dichloromethane (45 ml) and *N,N*-diisopropylethylamine (DIPEA, 8.2 ml; 47.1 mmol; 4 equiv) under a nitrogen atmosphere. Resulting mixture was cooled down in an ice bath. Trimethylsilyl chloride (TMSCl, 3.3 ml; 26 mmol; 2.2 equiv) was added slowly over 2 min, and the mixture was stirred for additional 25 min under cooling. Solution of di-*tert*-butyl dicarbonate (Boc_2O, 2.9 g; 13.3. mmol; 1.1 equiv.) in dichloromethane (5 ml) was added, and resulting mixture was stirred for additional 4 hours, while the bath was allowed to warm to 15 °C. Dichloromethane was removed gently under reduced pressure (temperature ≤ 35°C). The residue was dissolved in a mixture of 5% aqueous sodium hydrogen carbonate (50 ml), water and some acetone to the total volume of 300 ml. Resulting solution was washed with *tert*-butylmethylether (2x100 ml), acidified by adding potassium hydrogen sulphate (15 g) and extracted with ethyl acetate (4x50 ml). Extracts were dried over sodium sulphate, filtered and concentrated in vacuum. Acetonitrile and water were added to give a homogenous solution, which was then freeze-dried. Boc-Oic was obtained as a lightly yellowish powder (2.95 g; 11.0 mmol; 93% yield).

^{1}H NMR (MeOD, 700 MHz), δ: 4.20 (m, 1H, α-CH), 3.79 (m, 1H, δ-CH), 2.35 (m, 1H), 2.18 (m, 1H), 2.07 (m, 1H), 2.02 (m, 1H), 1.77 (m, 1H), 1.71 (m, 2H), 1.53-1.44 (m, 2H), 1.47 and 1.43 (two s, 9H, $(CH_3)_3C$), 1.35 (m, 1H), 1.21 (m, 1H).



Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

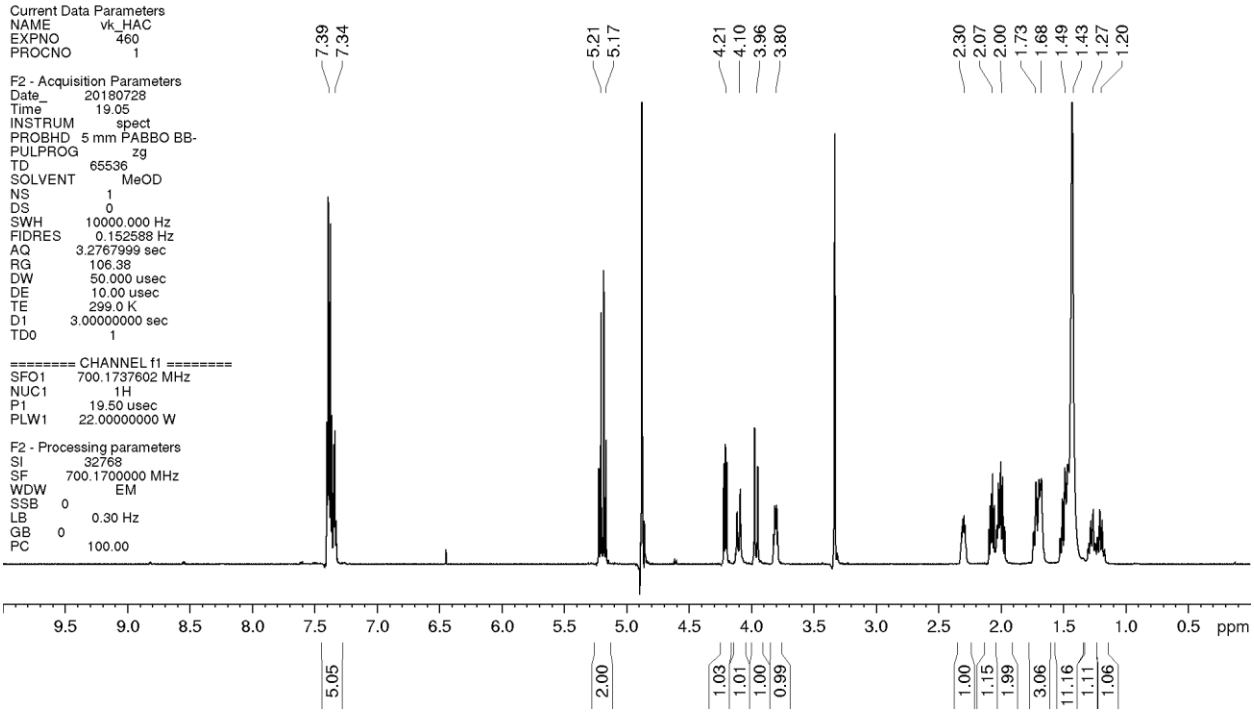


Boc-Oic Boc-Oic-Gly-OBn

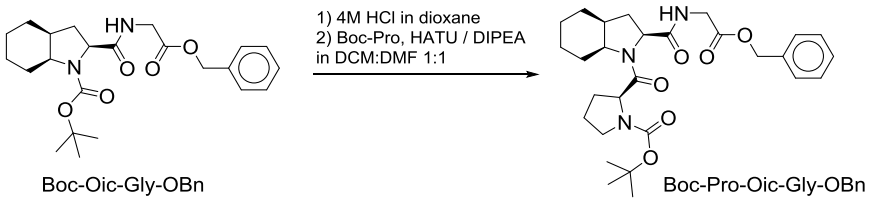
HATU = 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate

Boc-Oic-Gly-OBn: Boc-Oic (1.01 g; 3.75 mmol; 1 equiv) and HATU (1.41 g; 3.71 mmol; 1 equiv) were mixed with dichloromethane – *N,N*-dimethylformamide (DMF) 1:1 (v/v) mixture (20 ml), DIPEA (1.0 ml; 5.7 mmol; 1.5 equiv), and the mixture was cooled down in an ice bath under a nitrogen atmosphere. The mixture was stirred for 5 min under cooling, then 5 min under ambient temperature, and then again 5 min under cooling. DIPEA (1.3 ml; 7.5 mmol; 2 equiv) was added followed by solid HCl·Gly-OBn (1.13 g, 5.60 mmol; 1.5 equiv.). The mixture was stirred under cooling for the next 1 hour, and then the next 4 hours under ambient temperature (21-25 °C). Water (15 ml) and dichloromethane (100 ml) were added; organic phase was separated, washed with 10% aqueous potassium hydrogen suplhate (2x15 ml), 1M sodium hydrogen carbonate (2x15 ml) and brine (1x15 ml), dried over sodium sulphate, filtered and concentrated in vacuum. The residue was freeze-dried from acetonitrile – water mixture. Resulting crude material was purified on a silica gel column using ethyl acetate – methanol (25:1) mixture as an eluent ($R_f \geq 0.8$). Boc-Oic-Gly-OBn was obtained as a sticky yellowish solid. The yield was 1.33-1.47 g in different attempts, average 1.40 g; 3.36 mmol, 90%.

^{1}H NMR (MeOD, 700 MHz), δ: 7.41-7.32 (m, 5H, Ph), 5.21 and 5.17 (two d, AB, J = 12.3 Hz, 2H, CH_2Ph), 4.21 (dd, J = 10.0, 7.7 Hz, 1H, α-CH in Oic), 4.10 and 3.96 (two d, AB, J = 18.0 Hz, 2H, CH_2 in Gly), 3.80 (m, 1H, δ-CH in Oic), 2.30 (m, 1H), 2.07 (m, 1H), 2.00 (m, 2H), 1.72 (m, 1H), 1.68 (m, 2H), 1.53-1.43 (m, 2H), 1.43 (br s, 9H, $(CH_3)_3C$), 1.27 (m, 1H), 1.20 (m, 1H).

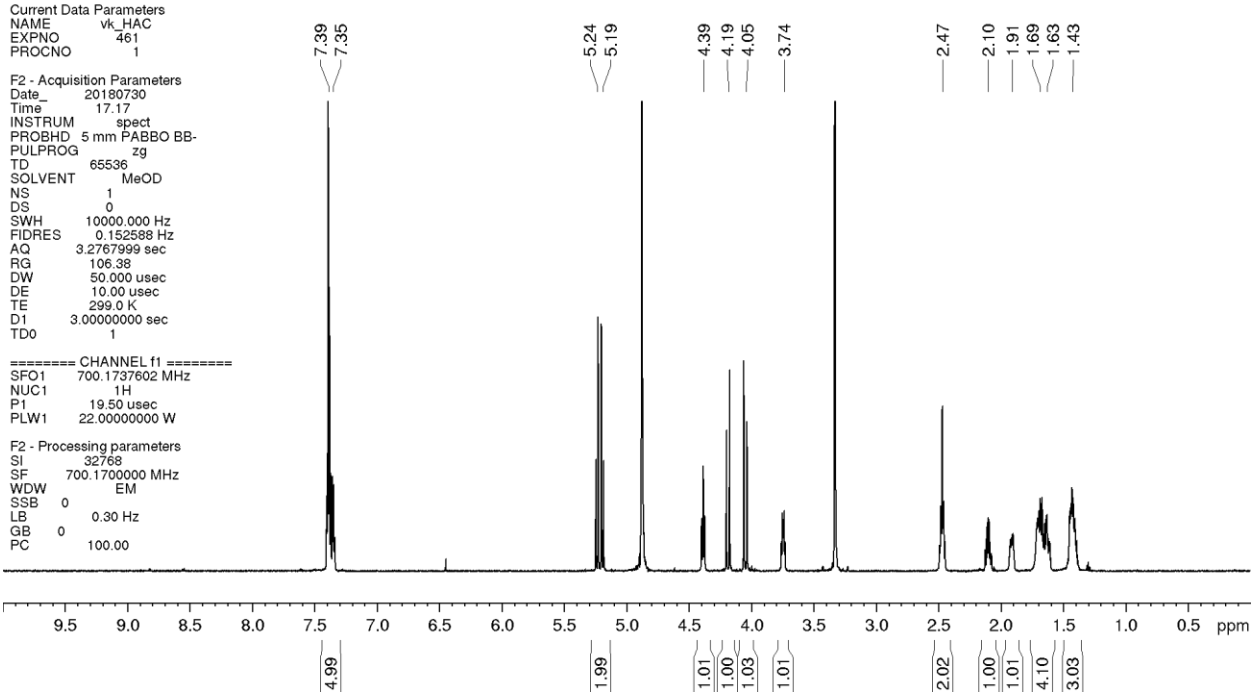


Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d



Boc-Pro-Oic-Gly-OBn: 1) Boc-Oic-Gly-OBn (1.47 g; 3.53 mmol) was mixed with 4 M hydrogen chloride solution in dioxane (10 ml; 40 mmol). The mixture was stirred at ambient temperature for 1.5 hours, bulk solvent was removed under reduced pressure; the mixture was dissolved in acetonitrile – water mixture and freeze-dried. HCl·Oic-Gly-OBn was obtained as a colorless solid and launched in the next step.

^{1}H NMR (MeOD, 700 MHz), δ: 7.41-7.33 (m, 5H, Ph), 5.24 and 5.19 (two d, AB, *J* = 12.3 Hz, 2H, CH_2Ph), 4.39 (t, *J* = 8.4 Hz, 1H, α-CH in Oic), 4.19 and 4.05 (two d, AB, *J* = 17.7 Hz, 2H, CH_2 in Gly), 3.74 (m, 1H, δ-CH in Oic), 2.47 (m, 2H), 2.10 (m, 1H), 1.91 (m, 1H), 1.72-1.60 (m, 4H), 1.45-1.38 (m, 3H).

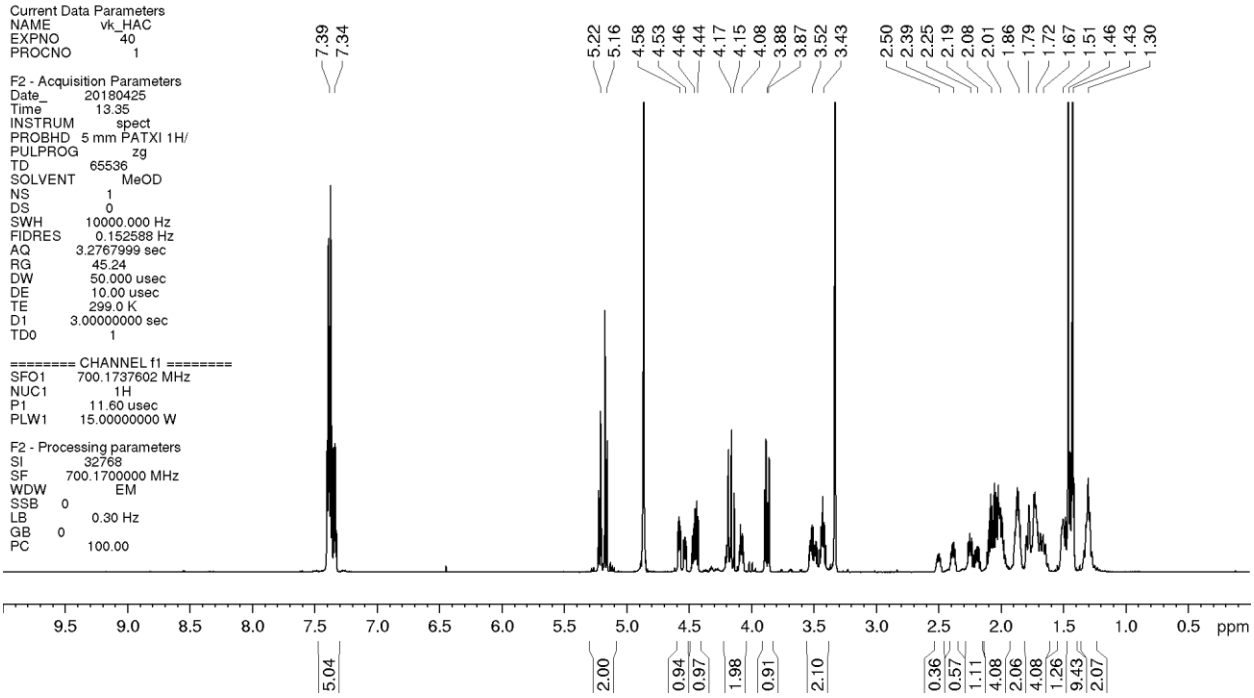


2) Boc-Pro (1.11 g; 5.16 mmol; 1.5 equiv.) was mixed with HATU (1.34 g; 3.53 mmol; 1 equiv.) and DIPEA (2.0 ml; 11.5 mmol; 3.3 equiv.) in dichloromethane - DMF 1:1 (v/v) mixture (15 ml) under a nitrogen atmosphere. Resulting mixture was stirred for 5 min under an ice batch cooling, 5 min at ambient temperature and then again 5 min at an ice batch cooling. HCl·Oic-Gly-OBn from the step 1) in the solvent mixture (5 ml) was added to the reaction mixture. Subsequent processing was done as in the previous step: the mixture was stirred under cooling for 1 h, then a few hours at the ambient temperature. Water was added, followed by dichloromethane, organic layer was separated and washed with potassium hydrogen sulphate, sodium hydrogen carbonate, brine, dried over sodium sulphate, filtered, concentrated in vacuum, the residue was freeze-dried from acetonitrile – water mixture. Resulting crude material was purified on a silica gel column using ethyl acetate – methanol (25:1) mixture as an eluent (R_f ≥ 0.8). After final freeze-drying from acetonitrile – water, the product Boc-Pro-Oic-Gly-OBn was obtained as a sticky yellowish solid (1.79 g; 3.49 mmol; 99% yield).

Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

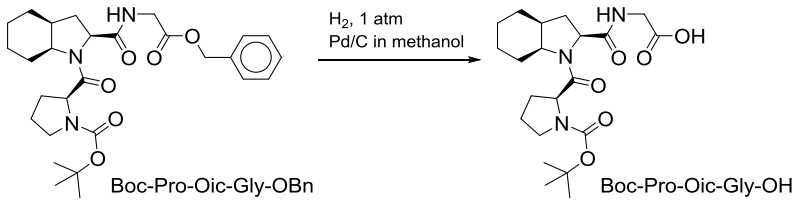
^{1}H NMR (MeOD, 700 MHz), δ: 7.40-7.32 (m, 5H, Ph), 5.22 and 5.16 (two d, AB, *J* = 12.4 Hz, 1H, CH_2Ph), 4.58 and 4.53 (two dd, J = 8.5, 4.4 Hz, 1H, α-CH in Pro, Boc-rotamers 3:2, exchange rates 0.9 and 0.6 Hz), 4.46 and 4.44 (two dd, *J* = 10.1, 8.0 Hz, α-CH in Oic, Boc-rotamers), 4.17 and 4.15 (two d, *J* = 18.0, 1H, CH in Gly, Boc-rotamers), 4.19 and 4.07 (two m, 1H, δ-CH in Oic, Boc-rotamers), 3.88 and 3.87 (two d, *J* = 18.0 Hz, 1H, CH in Gly, Boc-rotamers), 3.54-3.40 (m, 2H, δ-CH_2 in Pro), 2.50 and 2.39 (two m, 1H, γ-CH in Oic, Boc-rotamers 3:2, exchange rates 0.9 and 0.6 Hz), 2.25 and 2.19 (two m, 1H), 2.11-1.97 (m, 4H), 1.86 (m, 2H), 1.80-1.63 (m, 4H), 1.51 (m, 1H), 1.46 and 1.43 (two s, 9H, $(CH_3)_3C$), 1.30 (m, 2H).

Note that *N*-terminal Boc-group rotation creates a major source of the rotameric forms in the spectra (~ 3:2 ratio), however, there is a small amount of the minor rotameric forms coming from the *cis*-amide bond between Pro and Oic.



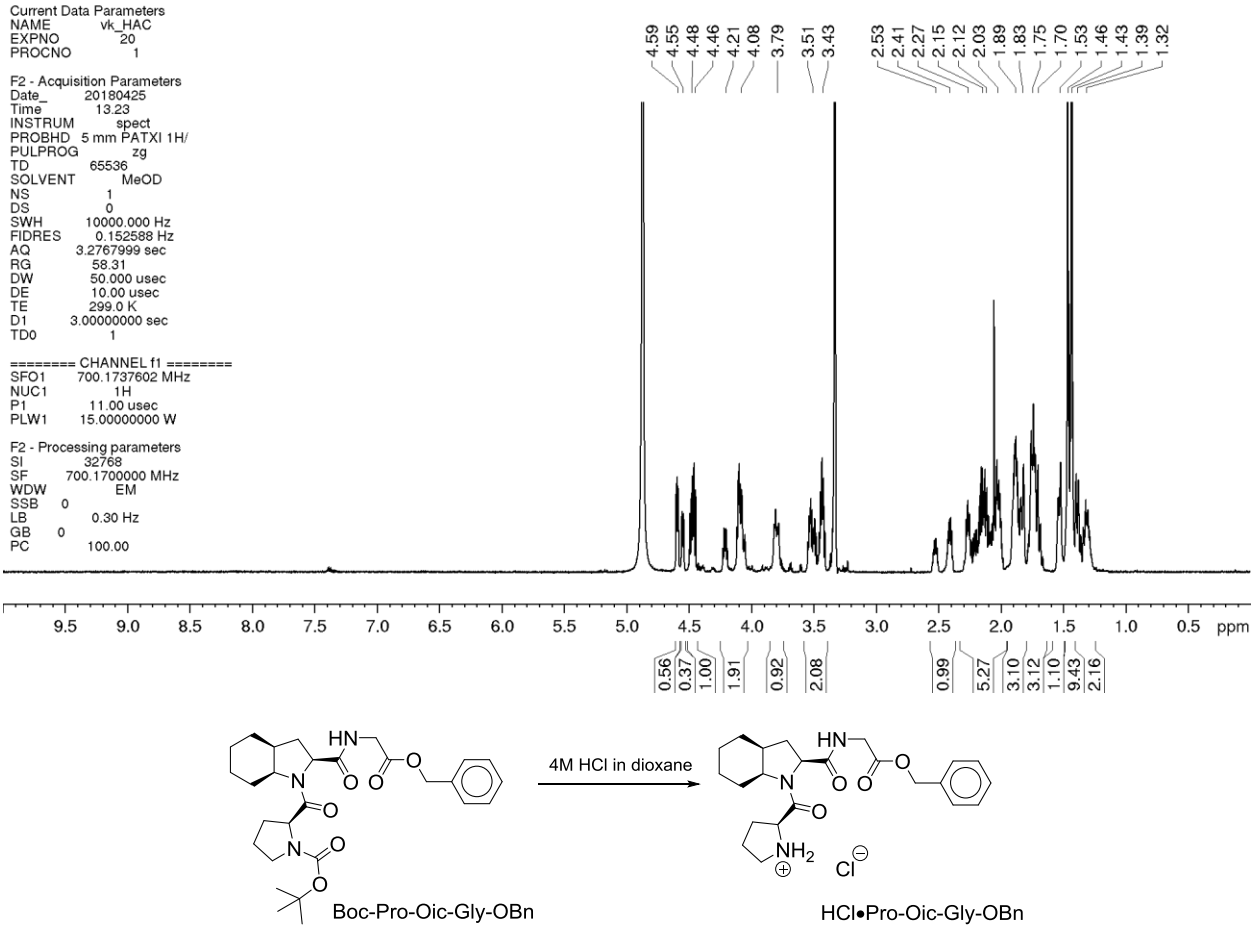
Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

Boc-(Pro-Oic-Gly)$_2$-OBn approach #1: triplet coupling

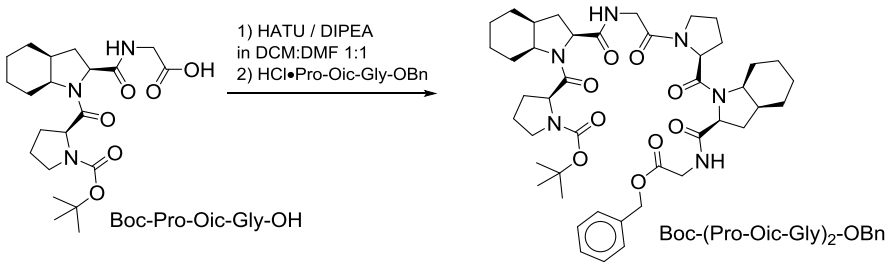


Boc-Pro-Oic-Gly-OBn (0.86 g; 1.67 mmol) and 10% palladium on charcoal (0.25 g) were mixed with methanol (6 ml), and resulting mixture was stirred under a hydrogen atmosphere for 2 hours. The mixture was filtered off, solvent was removed under reduced pressure, and the residue was freeze-dried from acetonitrile – water mixture. Boc-Pro-Oic-Gly-OH was obtained as a white powder (0.71 g; 1.67 mmol; quant.).

^{1}H NMR (MeOD, 700 MHz), δ: 4.59 and 4.55 (two dd, *J* = 8.6, 4.3 Hz, 1H, α-CH in Pro, Boc-rotamers 3:2), 4.48 and 4.46 (two dd, *J* = 10.2, 8.2 Hz, 1H, α-CH in Oic, Boc-rotamers), 4.21 and 4.12-4.05 (two m, 2H, δ-CH in Oic and CH in Gly), 3.79 (m, 1H, CH in Gly), 3.55-3.41 (m, 2H, δ-CH_2 in Pro), 2.53 and 2.41 (two m, 1H, Boc-rotamers 3:2), 2.27-1.99 (m, 5H), 1.91-1.81 (m, 3H), 1.78-1.67 (m, 3H), 1.53 (m, 1H), 1.46 and 1.43 (two s, 9H, $(CH3)_3$C, rotamers), 1.40-1.27 (m, 2H).



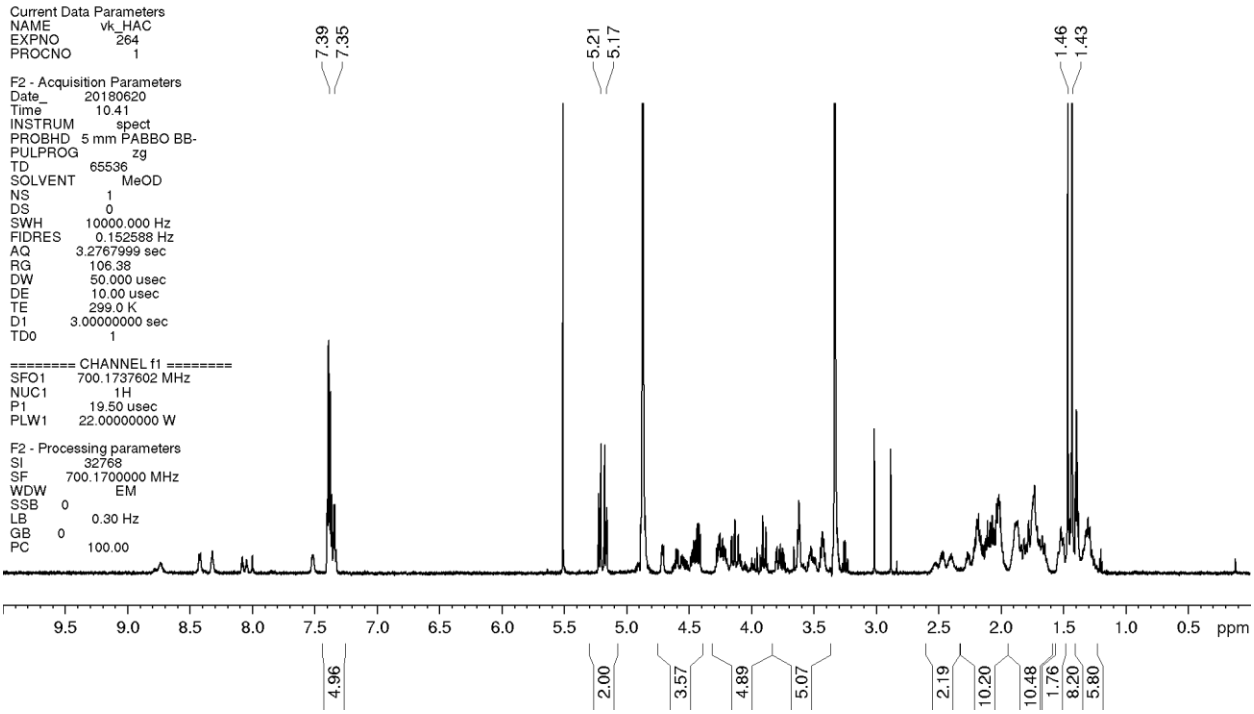
Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d



Boc-Pro-Oic-Gly-OBn (0.318 g; 0.62 mmol, 1 equiv.) was dissolved in 4M hydrogen chloride solution in dioxane (1.5 mol; 3 mmol). The mixture was stirred at the ambient temperature for 2 hours. Bulk solvent was blown off by nitrogen current, the residue was taken up to acetonitrile – water mixture and freeze-dried. HCl·Pro-Oic-Gly-OBn was obtained a sticky white solid, which was used in the next step.

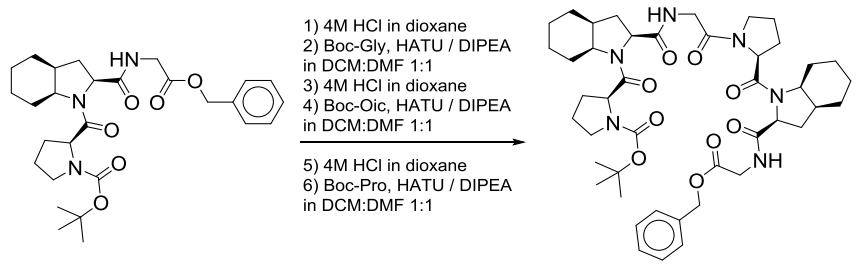
Boc-Pro-Oic-Gly-OH (0.262 g; 0.62 mmol; 1 equiv.) and HATU (0.235 g; 0.62 mmol; 1 equiv.) were mixed with DIPEA (0.32 ml; 1.84 mmol; 3 equiv.) in dichloromethane – DMF (1:1) mixture (1.5 ml) at ambient temperature. The mixture was shaken for 15 min, then HCl·Pro-Oic-Gly-OBn in dichloromethane – DMF mixture (0.5 ml) was added. Resulting mixture was shaken for 26 hours. Dichloromethane was blown off by nitrogen current, water was added and the mixture was freeze-dried. The residue was purified on a silica gel column using ethyl acetate – methanol (5:1) mixture as an eluent (R_f = 0.26). Boc-(Pro-Oic-Gly)$_2$-OBn was obtained after freeze-drying from acetonitrile – water as a white solid (0.27 g; 0.33 mmol; 53% yield).

^{1}H NMR (MeOD, 700 MHz), δ: 7.40-7.33 (m, 5H, Ph), 5.21 and 5.17 (two d, AB, *J* = 12.3 Hz, 2H, CH_2Ph), 4.70 and 4.63-4.41 (m, 4H), 4.28-3.87 (m, 5H), 3.81-3.73, 3.62 and 3.60-3.40 (series of multiplets, 5H), 2.55-2.36 (m, 2H), 2.29-1.97 (m, 10H), 1.92-1.63 (m, 10H), 1.51 (m, 2H), 1.46 and 1.43 (two s, 9H, $(CH_3)_3$C), 1.40-1.24 (m, 6H).



Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

Boc-(Pro-Oic-Gly)$_2$-OBn approach #2: stepwise peptide coupling



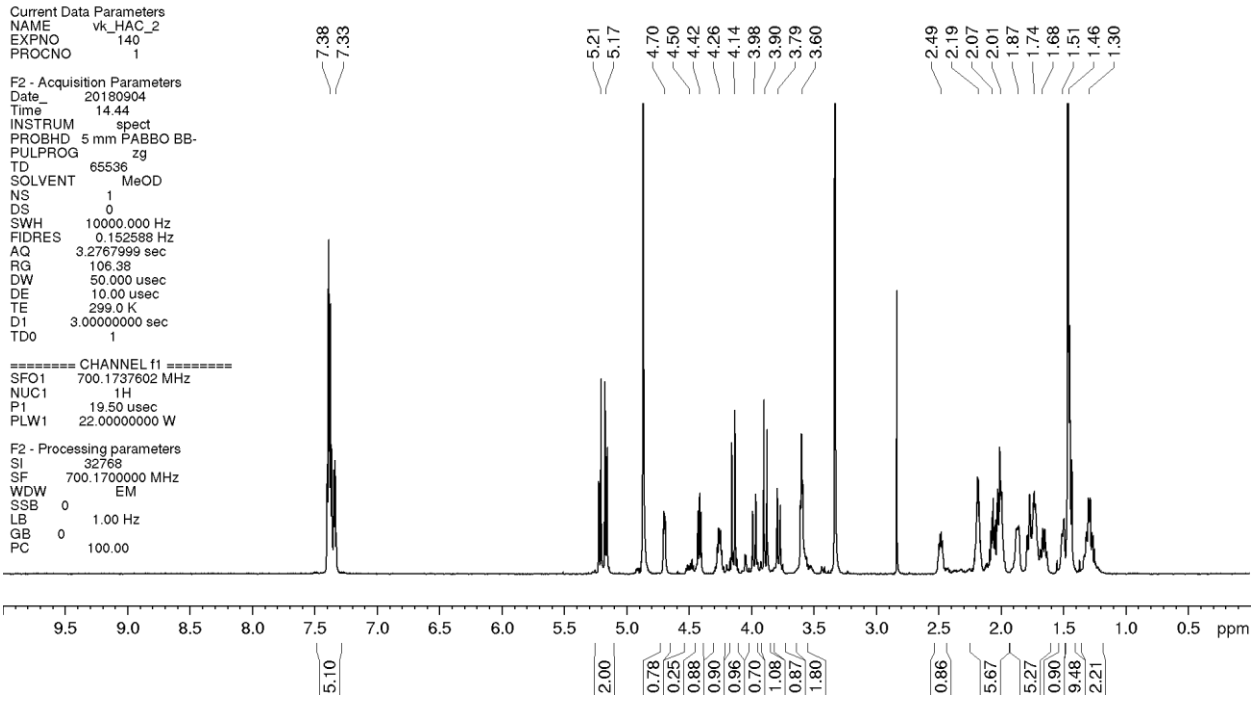
Boc-Pro-Oic-Gly-OBn

Boc-(Pro-Oic-Gly)$_2$-OBn

1) Boc-Pro-Oic-Gly-OBn (1.79 g; 3.49 mmol) was mixed with 4m hydrogen chloride in dioxane (6 ml) for 2 hours. The solvent was removed under reduced pressure; the residue was freeze-dried from acetonitrile – water mixture.

2) Boc-Gly (0.731 g; 4.18 mmol; 1.2 equiv.) and HATU (1.59 g; 4.18 mmol; 1.2 equiv.) were mixed with DIPEA (2.1 ml; 12.1 mmol; 3.5 equiv.) in dichloromethane – DMF (1:1) mixture (15 ml). The mixture was stirred 5 min unde an ice bath cooling, 5 min at the ambient temperature, then again 5 min under ice bath cooling. HCl·Pro-Oic-Gly-OBn in dichloromethane – DMF mixture (5 ml) was added, and the mixture was stirred 3 hours under ice batch cooling, then 3 hours at the ambient temperature. Water (25 ml) was added, followed by ethyl acetate (200 ml), organic layer was separated, washed with 10% potassium hydrogen sulphate (2x25 ml), 1 M sodium hydrogen carbonate (1x25 ml), brine (1x25 ml), dried over sodium sulphate, filtered and concentrated in vacuum. After freeze-drying from acetonitrile water the crude material was purified by silica gel chromatography using ethyl acetate – methanol (10:1) mixture as an eluent (R_f = 0.65). Boc-Gly-Pro-Oic-Gly-OBn was obtained as a yellowish sticky solid (1.65 g; 2.89 mmol; 83% yield).

For purity check, an unassigned ^{1}H NMR spectrum (MeOD, 700 MHz) is illustrated below:

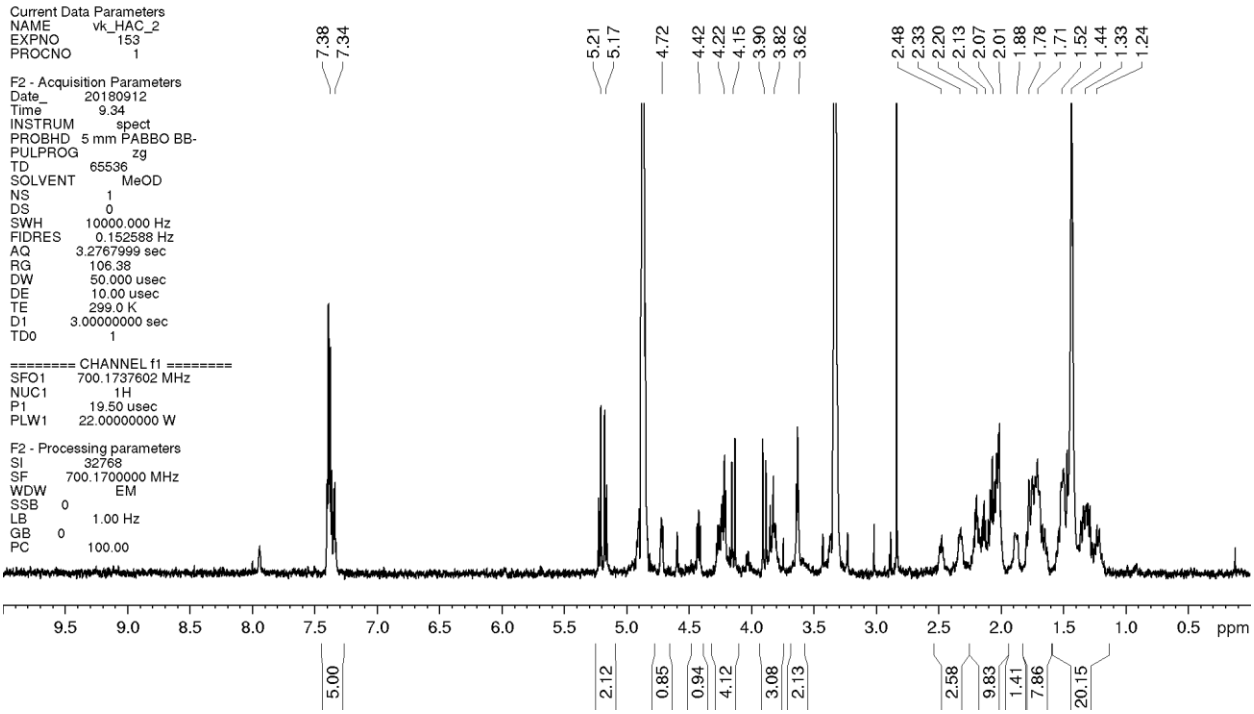


Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

3) Boc-Gly-Pro-Oic-Gly-OBn (1.65 g; 2.89 mmol) was mixed with 4M hydrogen chloride in dioxane (6 ml) for 1.5 hours at ambient temperature. The solvent was removed under reduced pressure, and the residue was freeze-dried from acetonitrile – water mixture.

4) Boc-Oic (0.86 g; 3.19 mmol; 1.1 equiv.) was mixed with HATU (1.21 g; 3.18 mmol; 1.1 equiv.) and DIPEA (3.0 ml; 17.2 mmol; 6 equiv.) in dichloromethane – DMF (1:1) mixture (15 ml). The mixture was stirred 5 min under ice bath cooling, then 5 min under ambient temperature and then again 5 min under cooling. HCl·Gly-Pro-Oic-Gly-OBn in the solvent mixture (5 ml) was added and the mixture was stirred for 3 hours under cooling and then 3 hour at the ambient temperature. Water (25 ml) and ethyl acetate (200 ml) were added, and further work-up was performed as in step 2). Final purification was performed by silica gel chromatography using ethyl acetate – methanol (10:1) mixture as an eluent (R_f = 0.58). Boc-Oic-Gly-Pro-Oic-Gly-OBn was obtained as a yellowish sticky solid (1.65 g; 2.29 mmol; 79% yield).

For purity check, an unassigned ^{1}H NMR spectrum (MeOD, 700 MHz) is illustrated below:

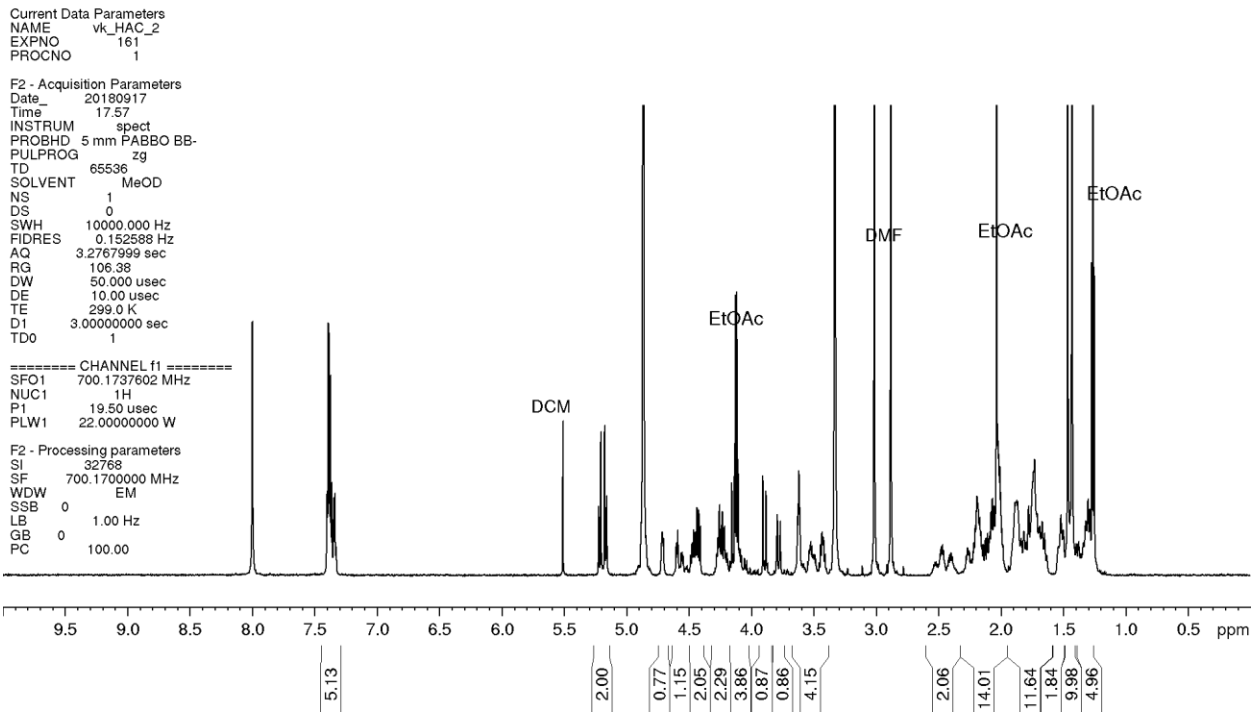


5) Boc-Oic-Gly-Pro-Oic-Gly-OBn (1.65 g; 2.29 mmol) was mixed with 4 M hydrogen chloride in dioxane (10 ml) for 1.5 hours. The solvent was removed under reduced pressure, and the residue was freeze-dried from acetonitrile – water.

6) Boc-Pro (0.59 g; 2.74 mmol; 1.2 equiv.) was mixed with HATU (1.04 g; 2.74 mmol; 1.2 equiv.) and DIPEA (2.0 ml; 11.5 mmol; 5 equiv.) in dichloromethane – DMF (1:1) mixture (15 ml). The coupling and the work-up was performed as in steps 2) and 4). The crude peptide material was purified on silica gel column using ethyl acetate – methanol (5:1) mixture as an eluent (R_f = 0.51). Boc-(Pro-Oic-Gly)$_2$-OBn was obtained after freeze-drying from acetonitrile – water as a white solid (1.57 g; 1.92 mmol; 84% yield).

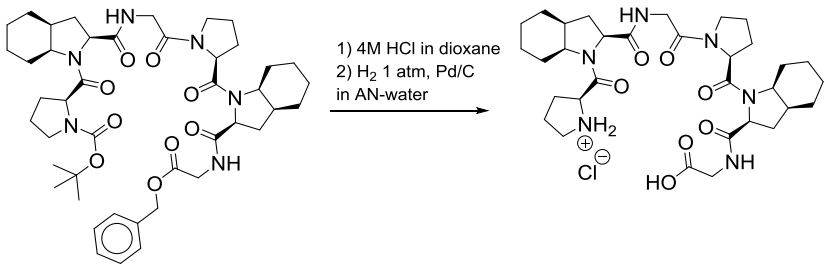
For purity check, an unassigned ^{1}H NMR spectrum (MeOD, 700 MHz) is illustrated below:

Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d



Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

Final deprotection:

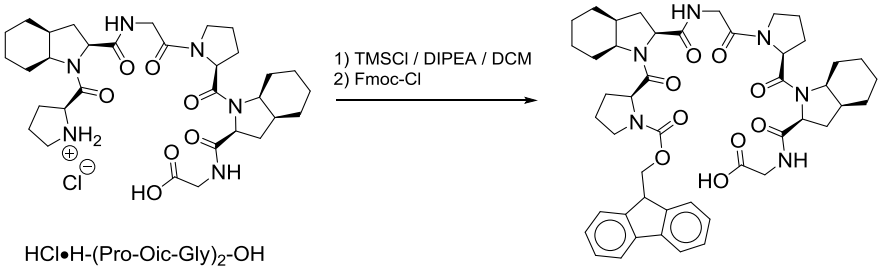


Boc-(Pro-Oic-Gly)$_2$-OBn

HCl•H-(Pro-Oic-Gly)$_2$-OH

The peptide Boc-(Pro-Oic-Gly)$_2$-OBn (1.57 g; 1.92 mmol) was mixed with 4 M hydrogen chloride in dioxane (11 ml) for 2.5 hours. The solvent was removed under reduced pressure, and the residues was freeze-dried from acetonitrile – water. The residue was mixed with 10% palladium on charcoal (0.51 g) in acetonitrile – water (1:1) mixture (20 ml) under hydrogen atmosphere under light heating (~ 30 °C). After 7 hours the mixture was filtered off, the liquid was freeze-dried to give HCl·H-(Pro-Oic-Gly)$_2$-OH as a brownish powder (1.17 g; 1.76 mmol; 92%).

Installation of the *N*-teminal Fmoc-group:



HCl•H-(Pro-Oic-Gly)$_2$-OH

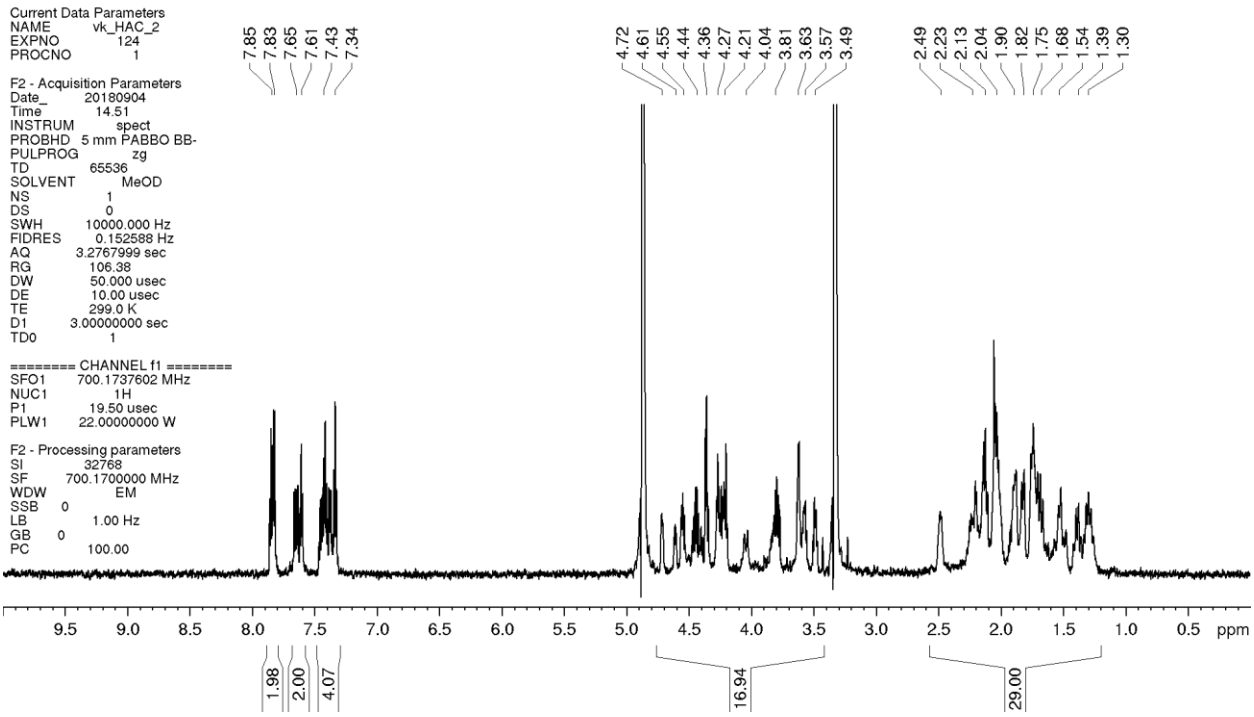
Fmoc-(Pro-Oic-Gly)$_2$-OH

HCl·H-(Pro-Oic-Gly)$_2$-OH (1.17 g; 1.76 mmol; 1 equiv.) from the previous step was mixed with DIPEA (2.0 ml; 11.5 mmol; 6.5 equiv.) in anhydrous dichloromethane (30 ml) under a nitrogen atmosphere. The mixture was cooled down in an ice bath, TMSCl (0.50 ml; 3.94 mmol; 2.2. equiv.) was added, and the mixture was stirred under cooling for additional 30 min. Solution of Fmoc chloride (0.50 g, 1.93 mmol; 1.1 equiv.) in anhydrous dichloromethane (5 ml) was added, the mixture was stirred for the next 7 hours, and it was allowed to warm up to 13 °C. 10% aqueous solution of potassium hydrogen sulphate (25 ml) was added, followed by addition of ethyl acetate (100 ml). Organic layer was separated, washed with brine (1x15 ml), dried over magnesium sulphate, filtered and concentrated in vacuum. The residue was dissolved in acetone (15 ml); 1 M sodium hydrogen carbonate (3.5 ml) and water (200 ml) were added. Resulting solution was stirred with hexane (150 ml) four times for over 5 hours, each time the hexane phase was separated and discarded. Potassium hydrogen suplhate (10 g) was added, and resulting mixture was extracted with ethyl acetate (4x150 ml). Organic phases were dried over sodium sulphate, filtered and concentrated in vacuum. The residue was freeze-dried from acetonitrile – water mixture to give Fmoc-(Pro-Oic-Gly)$_2$-OH as white powder (0.57 g, 0.67 mmol, 38%). In several attempts, the final yield varied between 35 and 47%. We analyzed various washes and fractions in course of the synthesis and found that there was a substantial compound left in the potassium hydrogen sulphate washes. Thus, the reaction with Fmoc-chloride was quite inefficient due to unknown reasons. In addition, tedious separation from residual Fmoc species was often incomplete even after extensive hexane washing, therefore reverse-phase HPLC separation was

Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

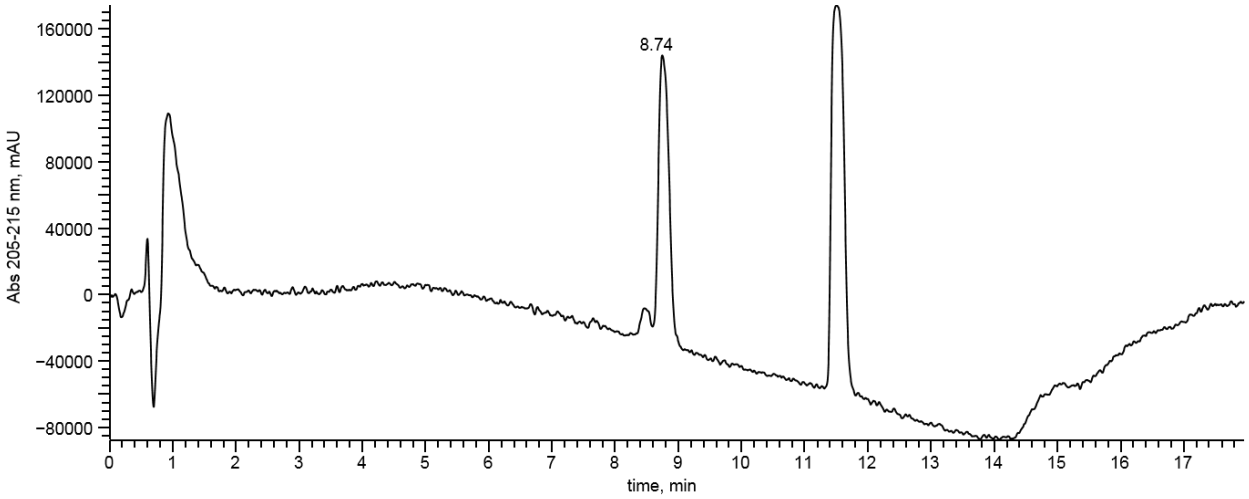
required. Curiously, all these problems were virtually absent in the synthesis of Fmoc-(Ash-Oic-Gly)$_2$-OH as described in the following text.

Mass-spectrum (ESI-Orbitrap): calcd. $[M+H]^+$ 851.4338, $[M+Na]^+$ 873.4157, found 851.4366, 873.4183. ^{1}H NMR (MeOD, 700 MHz), δ: 7.86-7.82 (m, 2H), 7.67-7.60 (m, 2H), 7.47-7.32 (m, 4H), 4.73-3.47 (series of multiplets, 17H), 2.50-1.25 (series of multiplets, 30H).

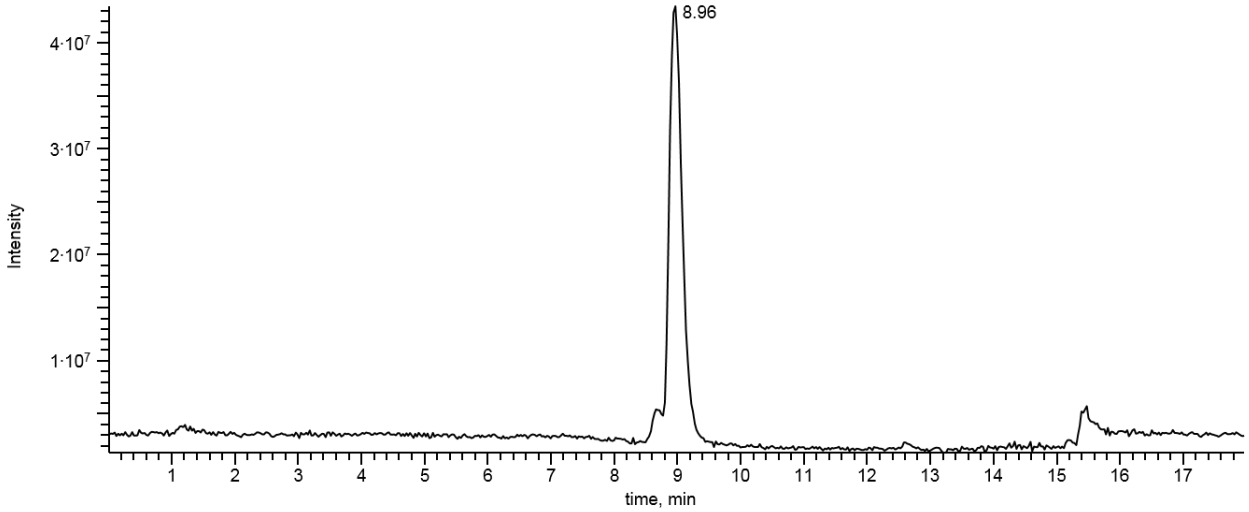


Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

RP-HPLC chromatogram, absorbance at 205-215 nm:

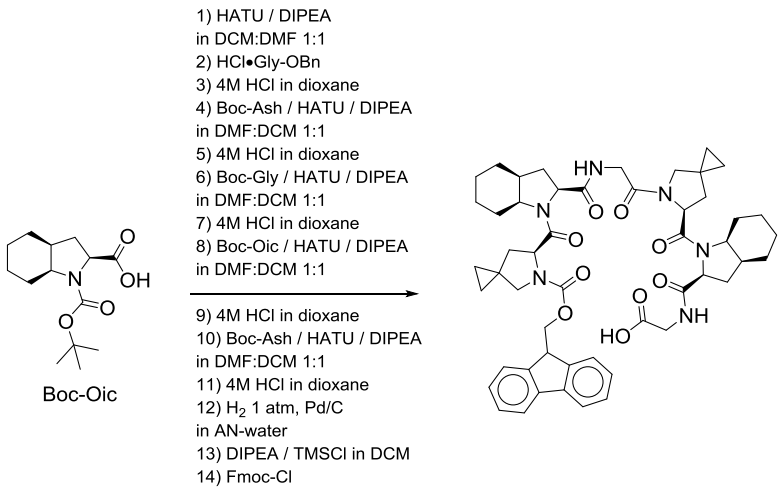


RP-HPLC chromatogram, total ion count:



Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

Synthesis of Fmoc-(Ash-Oic-Gly)-OH was performed using same protocols, the proportions were as follows:



target compound	starting materials	yield
Boc-Oic-Gly-OBn	Boc-Oic (1.01 g; 3.75 mmol; 1 equiv.) HATU (1.41 g; 3.71 mmol; 1 equiv.) DIPEA (2.3 ml; 13.2 mmol; 3.5 equiv.) HCl·Gly-OBn (1.13 g, 5.60 mmol; 1.5 equiv.)	1.40 g (3.36 mmol; 90% yield) Purification on silica gel. Eluent: ethyl acetate – methanol (25:1), $R_f \geq 0.8$
HCl·Oic-Gly-OBn	Boc-Oic-Gly-OBn (1.40 g; 3.36 mmol)	
Boc-Ash-Oic-Gly-OBn	Boc-Ash (0.983 g; 4.08 mmol; 1.2 equiv.) HATU (1.55 g; 4.08 mmol; 1.2 equiv.) DIPEA (3.0 ml; 17.2 mmol; 5.1 equiv.) HCl·Oic-Gly-OBn from the previous step	1.80 g (3.34 mmol; 99%) Purification on silica gel. Eluent: ethyl acetate – methanol (25:1), $R_f \geq 0.8$
HCl·Ash-Oic-Gly-OBn	Boc-Ash-Oic-Gly-OBn (1.77 g; 3.28 mmol)	
Boc-Gly-Ash-Oic-Gly-OBn	Boc-Gly (0.631 g; 3.61 mmol; 1.1 equiv.) HATU (1.37 g; 3.61 mmol; 1.1 equiv.) DIPEA (2.0 ml; 11.5 mmol; 3.2 equiv.) HCl·Ash-Oic-Gly-OBn from the previous step	1.72 g (2.88 mmol; 80%) Purification on silica gel. Eluent: ethyl acetate – methanol (10:1), $R_f \geq 0.8$ in 40:1 $R_f = 0.68$

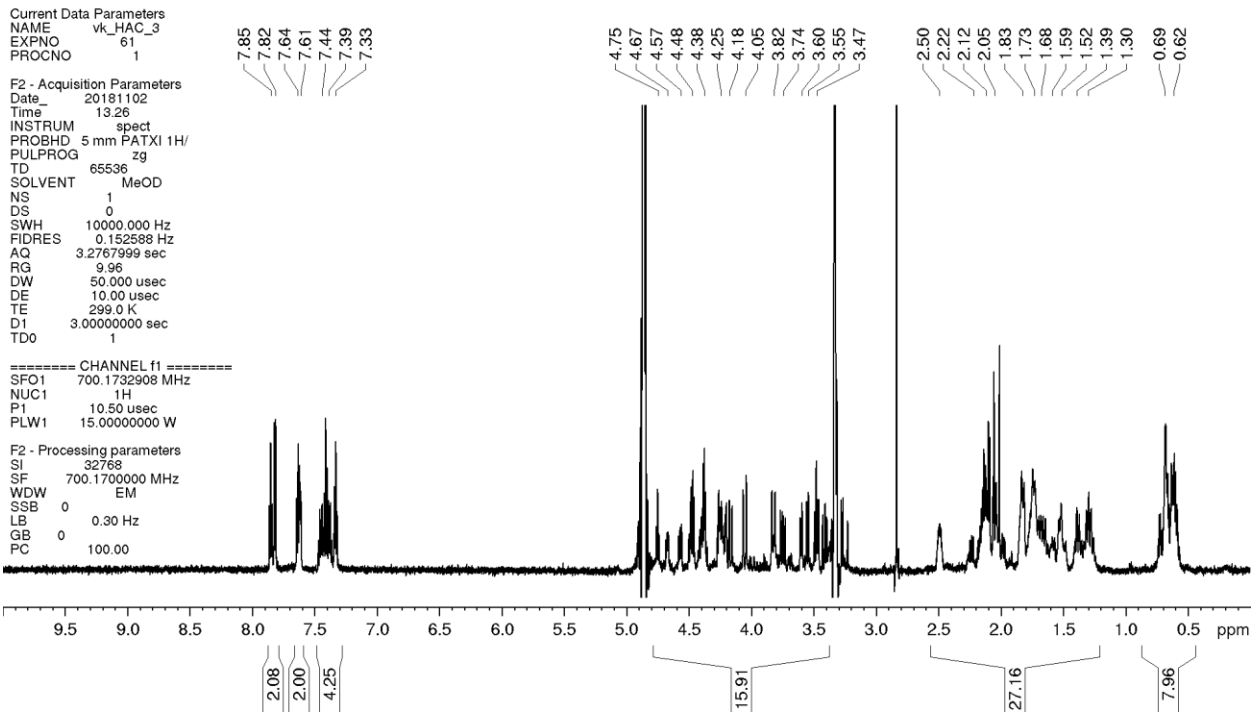
Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

HCl·Gly-Ash-Oic-Gly-OBn	Boc-Gly-Ash-Oic-Gly-OBn (1.72 g; 2.88 mmol)	
Boc-Oic-Gly-Ash-Oic-Gly-OBn	Boc-Oic (0.853 g; 3.17 mmol; 1.1 equiv.) HATU (1.20 g; 3.16 mmol; 1.1 equiv.) DIPEA (2.0 ml; 11.5 mmol; 4.0 equiv.) HCl·Gly-Ash-Oic-Gly-OBn from the previous step	2.00 g (2.67 mmol; 93%) Purification on silica gel. Eluent: ethyl acetate – methanol (10:1), R_f = 0.67
HCl·Oic-Gly-Ash-Oic-Gly-OBn	Boc-Oic-Gly-Ash-Oic-Gly-OBn (2.00 g; 2.67 mmol)	
Boc-Ash-Oic-Gly-Ash-Oic-Gly-OBn	Boc-Ash (0.691 g; 2.87 mmol; 1.07 equiv.) HATU (1.09; 2.87 mmol; 1.07 equiv.) DIPEA (2.0 ml; 11.5 mmol; 4.3 equiv.) HCl·Oic-Gly-Ash-Oic-Gly-OBn from the previous step	2.30 g (2.64 mmol; 99%) Purification on silica gel. Eluent: ethyl acetate – methanol (5:1), R_f = 0.6
HCl·Ash-Oic-Gly-Ash-Oic-Gly-OBn	Boc-Ash-Oic-Gly-Ash-Oic-Gly-OBn (2.30 g; 2.64 mmol)	
HCl·Ash-Oic-Gly-Ash-Oic-Gly-OH	HCl·Ash-Oic-Gly-Ash-Oic-Gly-OBn from previous step Pd/C (0.52 g)	1.75 g (2.44 mmol; 92%)
Fmoc-Ash-Oic-Gly-Ash-Oic-Gly-OH	HCl·Ash-Oic-Gly-Ash-Oic-Gly-OH (1.35 g; 1.88 mmol; 1.equiv.) DIPEA (1.5 ml; 8.6 mmol; 4.6 equiv.) TMSCl (0.53 ml; 4.18 mmol; 2.2 equiv.) Fmoc-Cl (0.536 g; 2.07 mmol; 1.1 equiv.)	1.41 g (1.56 mmol; 83%) Basic aqueous solution of the peptide was extensively washed with hexane from residual fluorenyl species.

Fmoc-Ash-Oic-Gly-Ash-Oic-Gly-OH was obtained as a white powder.

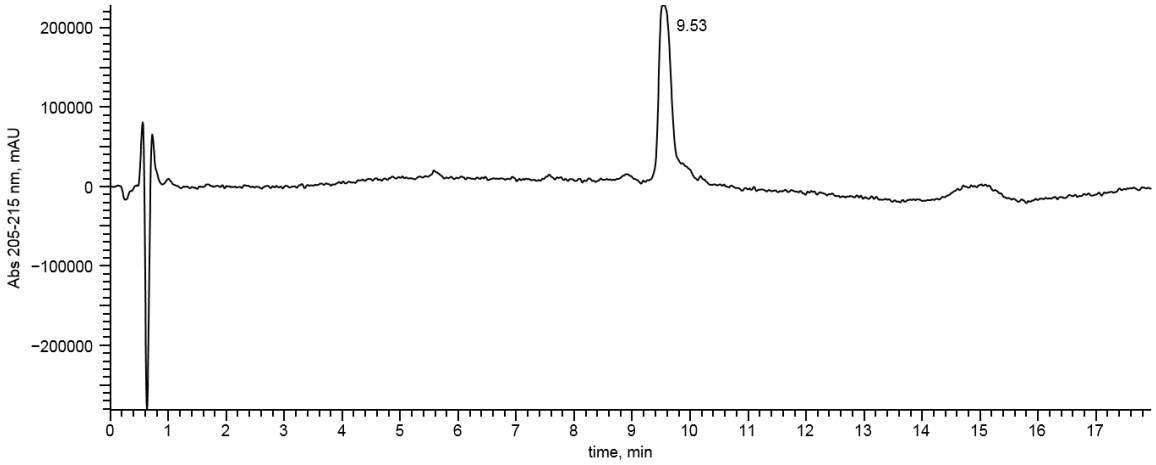
Mass-spectrum (ESI-Orbitrap): calcd. $[M+H]^+$ 903.4651, $[M+Na]^+$ 925.4470, found 903.4645, 925.4460. ^{1}H NMR (MeOD, 700 MHz), δ:7.86-7.81 (m, 2H), 7.65-7.60 (m, 2H), 7.47-7.32 (m, 4H), 4.77-3.36 (series of multiplets, 17 H), 2.51-1.25 (series of multiplets, 26H), 0.75-0.58 (m, 8H).

Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

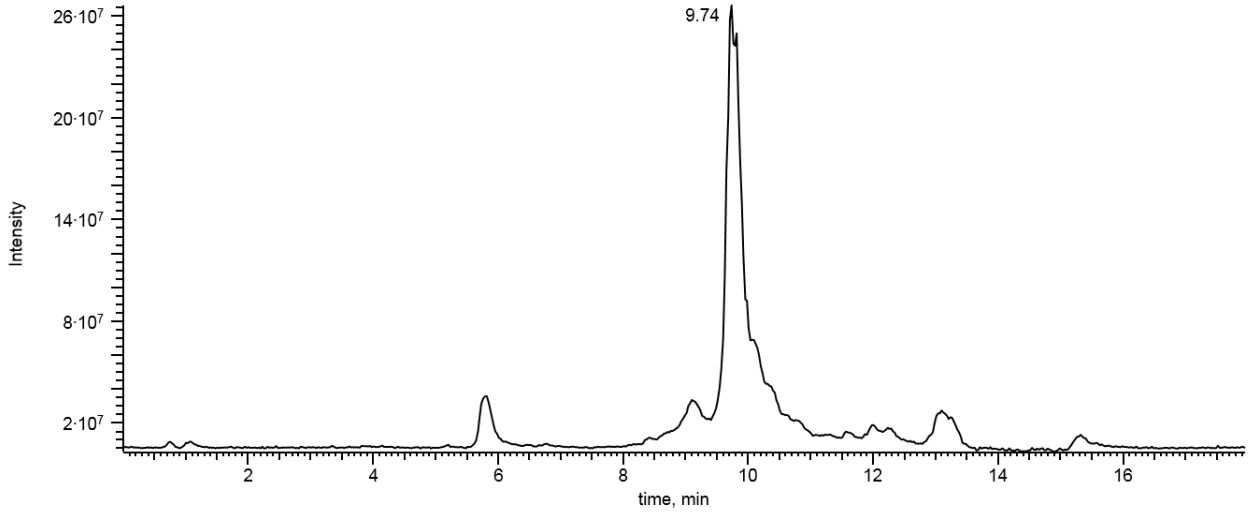


Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

RP-HPLC chromatogram, absorbance at 205-215 nm:



RP-HPLC chromatogram, total ion count:



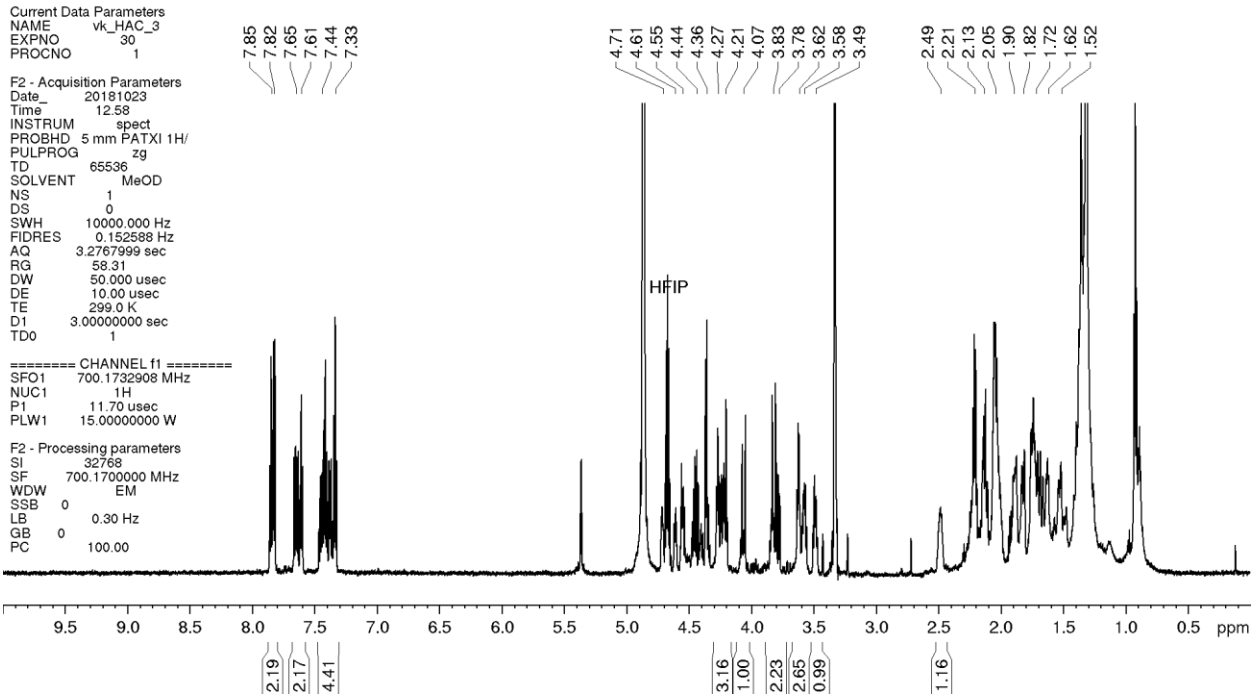
Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

Solid phase peptide synthesis

Chlorotrityl chloride resin (0.51 g, on polystyrene, PS) was mixed with DIPEA (0.25 ml) in anhydrous dichloromethane (5 ml). Fmoc-hexapepetide (54 mg; 0.06 mmol) was mixed with DIPEA (0.05 ml) in dichloromethane (2 ml), and resulting solution was added to the resin with extensive mixing. The mixture was shaken for the next 25 min, then methanol – dichloromethane (1:1) mixture (2 ml) was added, and shaking was continued for the next 20 min. The mixture was filtered off, washed with dichloromethane, methanol, dichloromethane and diethyl ether, and then it was dried in vacuum overnight.

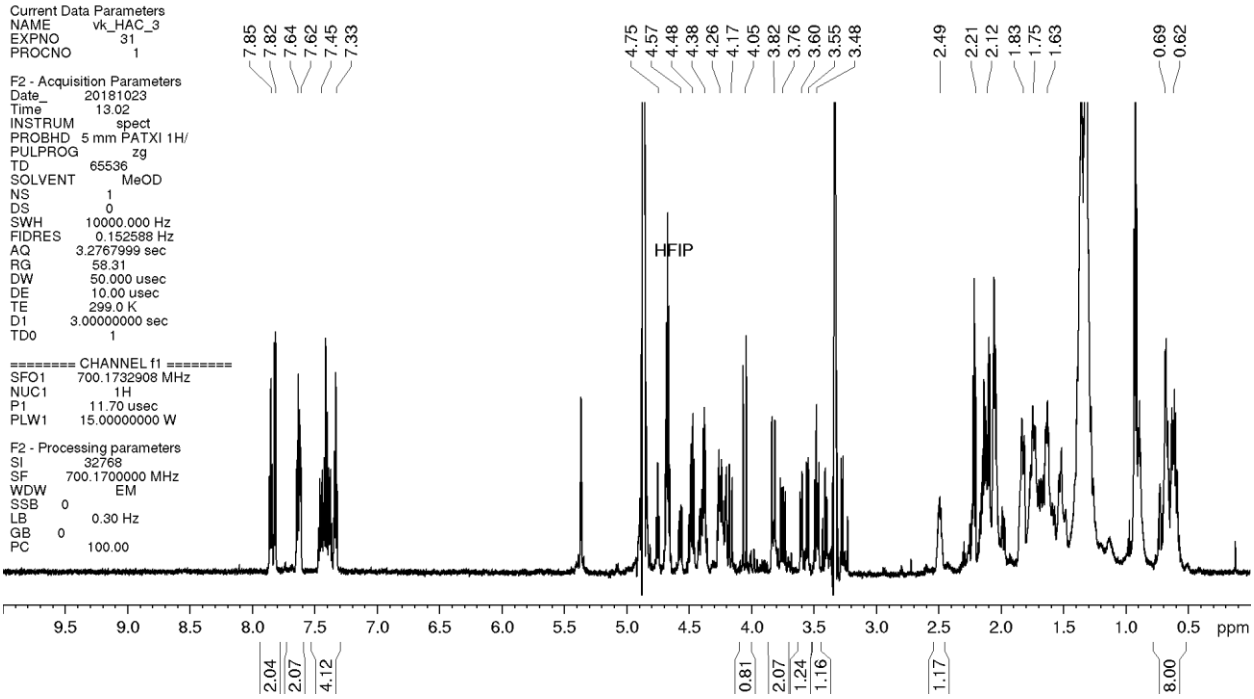
We performed test cleavage from resulting pre-loaded resins. The resin (0.02 g) was shaken with hexafluoroisopropanol (HFIP) – dichloromethane (1:2 v/v) mixture (3 ml) for 20 min, then it was filtered, washes were blown off by nitrogen current, and the residues were freeze-dried from acetonitrile – water. For both peptides approx. 2 mg of Fmoc-hexapeptides were obtained, and the spectra were consistent with the starting Fmoc-hexapeptides. ^{1}H NMR spectra are shown below (MeOD, 700 MHz):

Fmoc-(Pro-Oic-Gly)$_2$-OH:



Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

Fmoc-(Ash-Oic-Gly)$_2$-OH:



The peptide synthesis was performed with the following settings:

Fmoc-deptorection: by mixing with 22 vol % piperidine in DMF (2 ml) for 20-25 min

Hexapeptide coupling: Fmoc-hexaptide (100 mg Fmoc-(Pro-Oic-Gly)$_2$-OH or 106 mg Fmoc-(Ash-Oic-Gly)$_2$-OH) was mixed with HATU (43 mg) and DIPEA (65 μl) in dichloromethane – DMF (1:1) mixture (2 ml). Resulting mixture was shaken for about 5 min; then it was added to the resin for coupling. The coupling wascontinued for 5-7 h.

N-terminal acetylation: acetic anhydride (62 μl) was mixed with DIPEA (170 μl) in dichloromethane – DMF (1:1) mixture (2.5 ml) and this added to the resin. Shaking was continued for 4 hours.

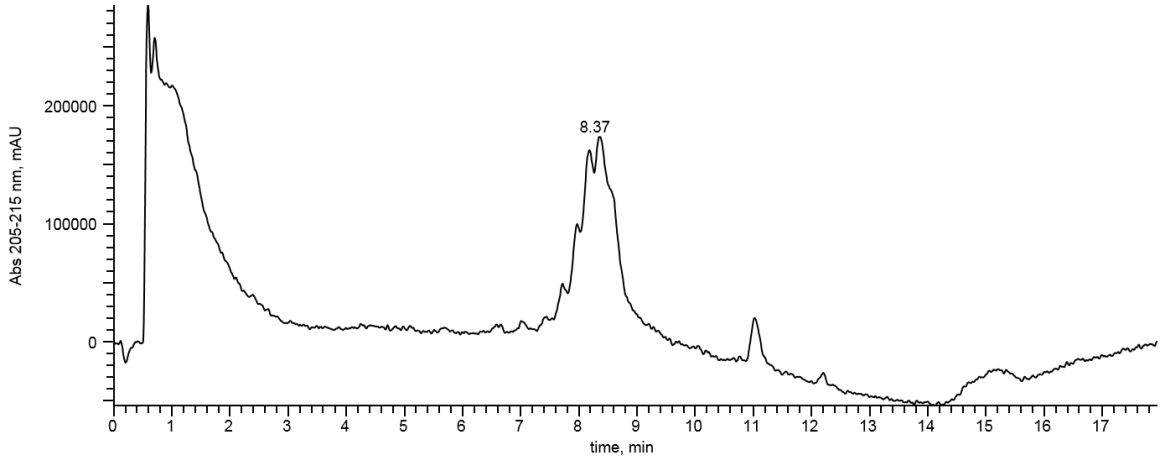
N-terminal pivaloylation: pivaloyl chloride (58 μl) was mixed with DIPEA (123 μl) in dichloromethane – DMF (1:1) mixture (2.5 ml) and this added to the resin. Shaking was continued for 3 hours. The step was repeated.

After the synthesis the resin was washed with DMF, dichloromethane, methanol, dichloromethane and diethyl ether; then it was dried in vacuum overnight.

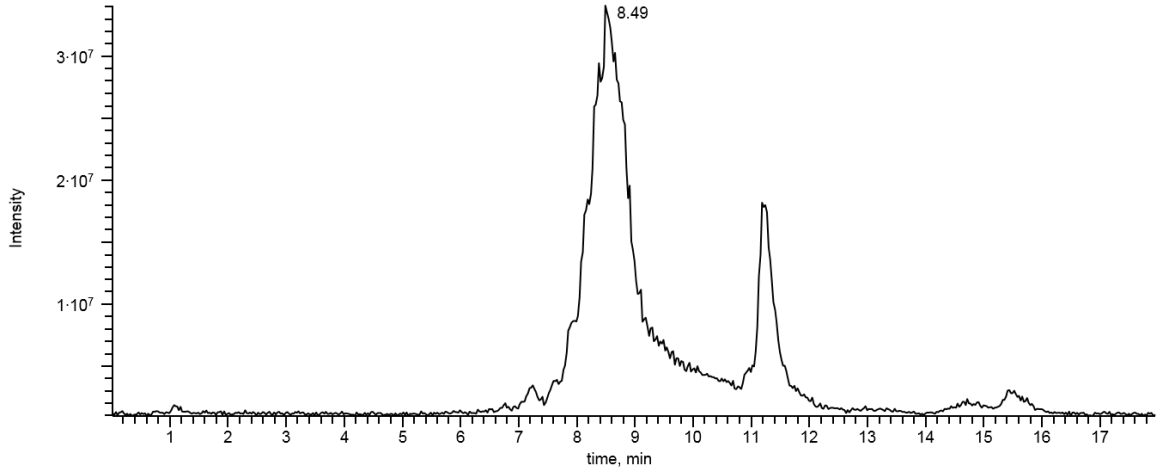
Peptide cleavage: The resin was mixed with hexafluoroisopropanol – dichloromethane (1:2 v/v) mixture (4 ml) for 20 min. It was filtered off, washed with dichloromethane twice, shaken with the HFIP – dichloromethane mixture again for 15 min, then filtered, washed with dichloromethane twice. Combined washes were blown off by nitrogen current under mild heating (~ 30 °C). The residue was suspended in acetonitrile – water mixture and freeze-dried. The peptide were obtained as white powders (approx. 170-180 mg each).

Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

RP-HPLC chromatogram for **AcPOG10**, absorbance at 205-215 nm:



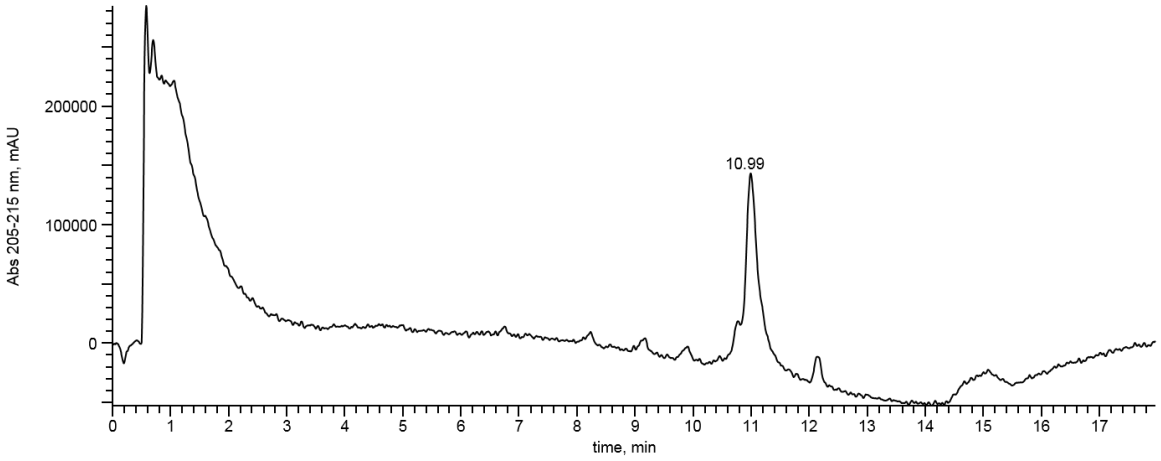
RP-HPLC chromatogram for **AcPOG10**, total ion count:



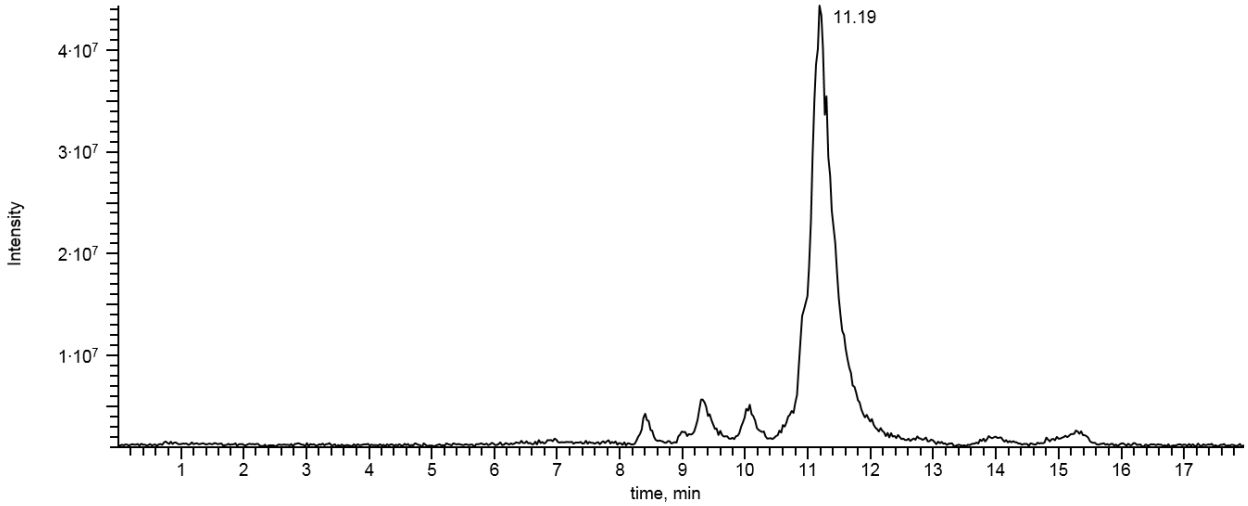
The reason for the complex shape of the peak is unknown.

Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

RP-HPLC chromatogram for **PivσOG10**, absorbance at 205-215 nm:



RP-HPLC chromatogram for **PivσOG10**, absorbance at 205-215 nm:



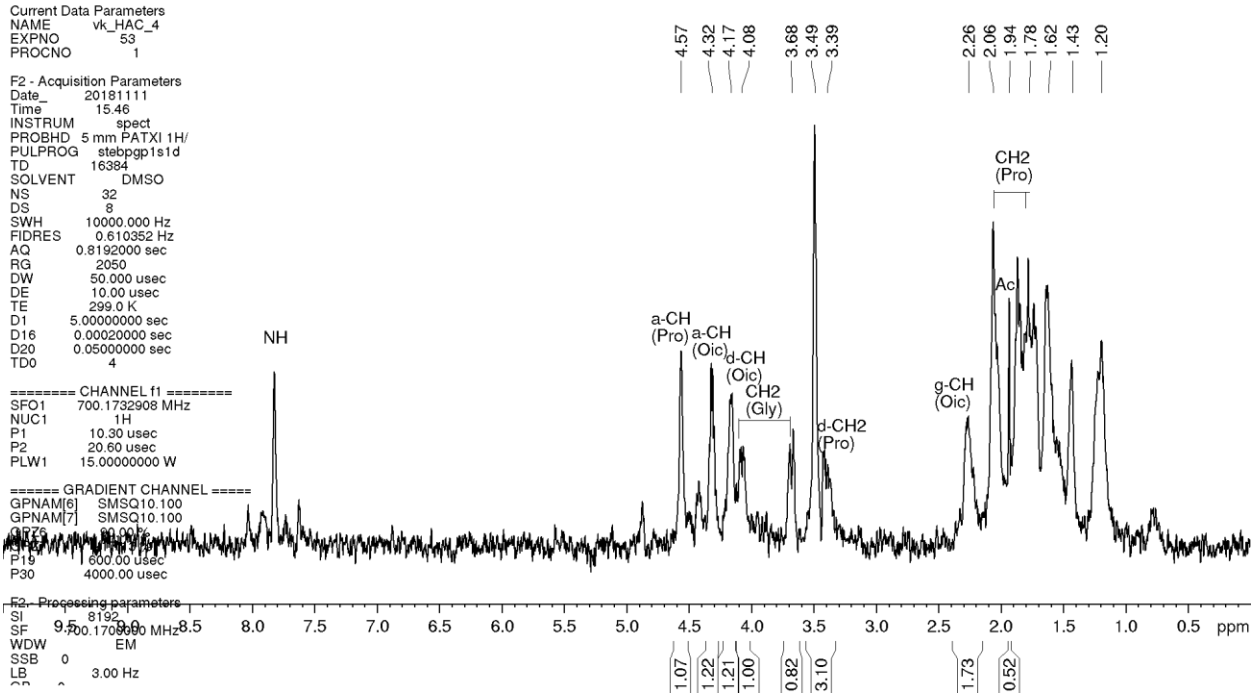
Detailed investigation for the behavior of hydrophobic collagen mimicking peptides under HPLC conditions is to be conducted. Here we considered the peptide samples after the synthesis pure enough for subsequent analysis of their structure and stability.

The peptides **AcσOG2**, **AcσOG4**, **AcσOG6**, **AcσOG8** and **AcσOG10** were prepared analogously, each from approx. 100 mg of pre-loaded resin.

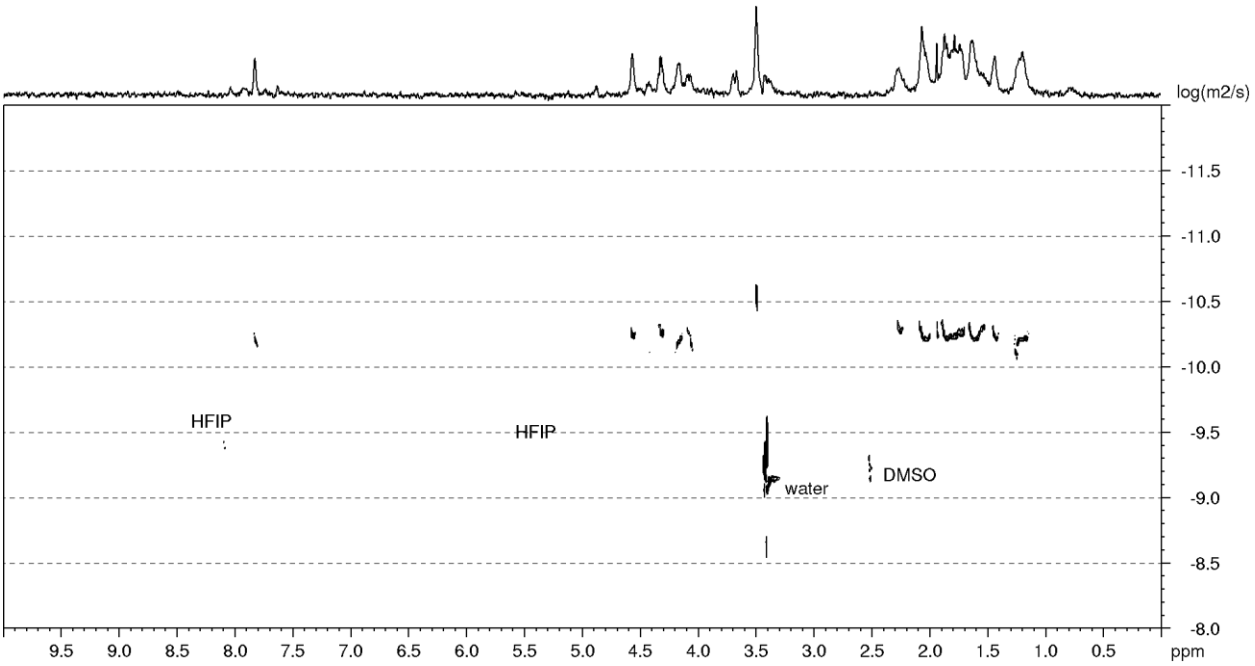
Characterization of peptides

^{1}H NMR analysis was performed in DMSO-d_6 solution at 2 mg ml^{-1} concentration at 25 °C.

^{1}H stimulated echo for **AcPOG10** (DMSO-d_6, 700 MHz):

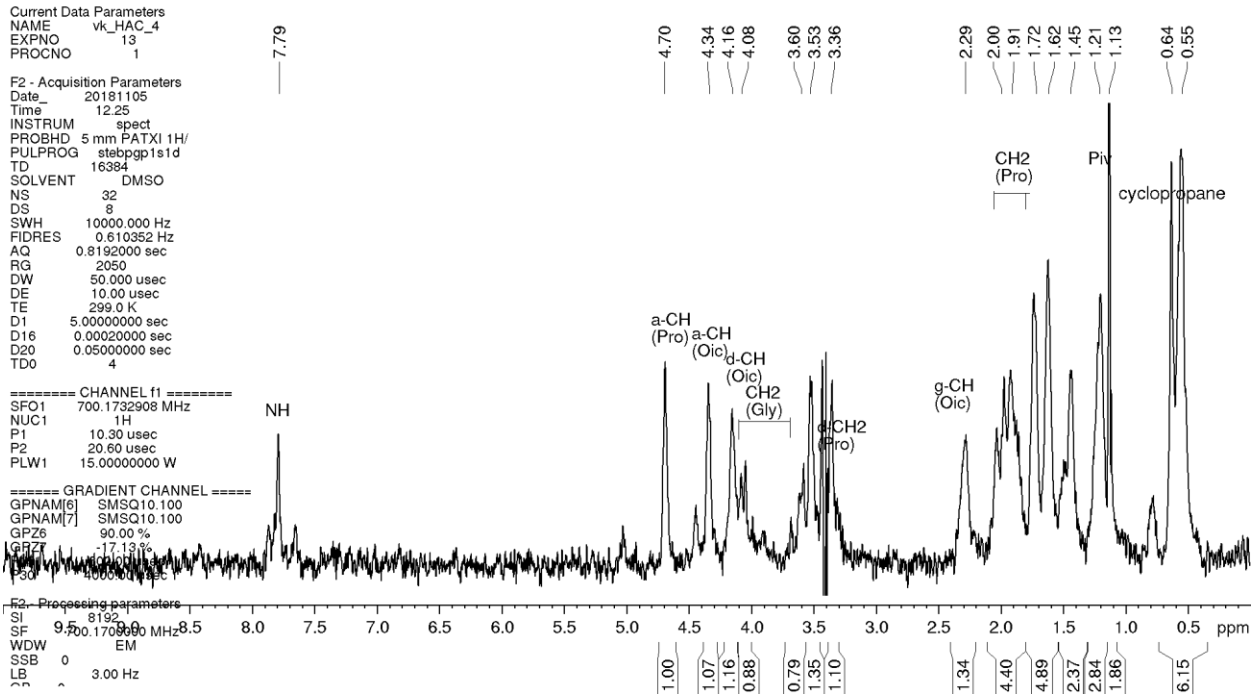


^{1}H DOSY for **AcPOG10** (DMSO-d_6, 700 MHz), δ = 50 ms, Δ/2 = 3 ms, 25 °C:

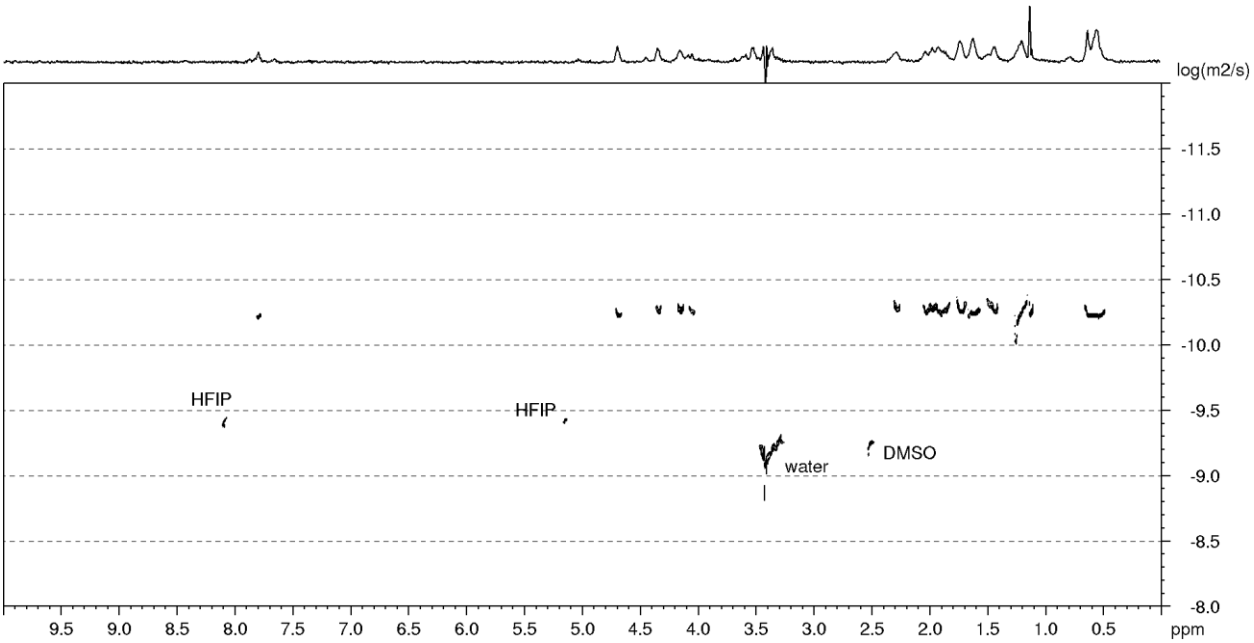


Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

^{1}H stimulated echo for **PivσOG10** (DMSO-d_6, 700 MHz):



^{1}H DOSY for **PivσOG10** (DMSO-d_6, 700 MHz), δ = 50 ms, $\Delta/2$ = 3 ms, 25 °C:

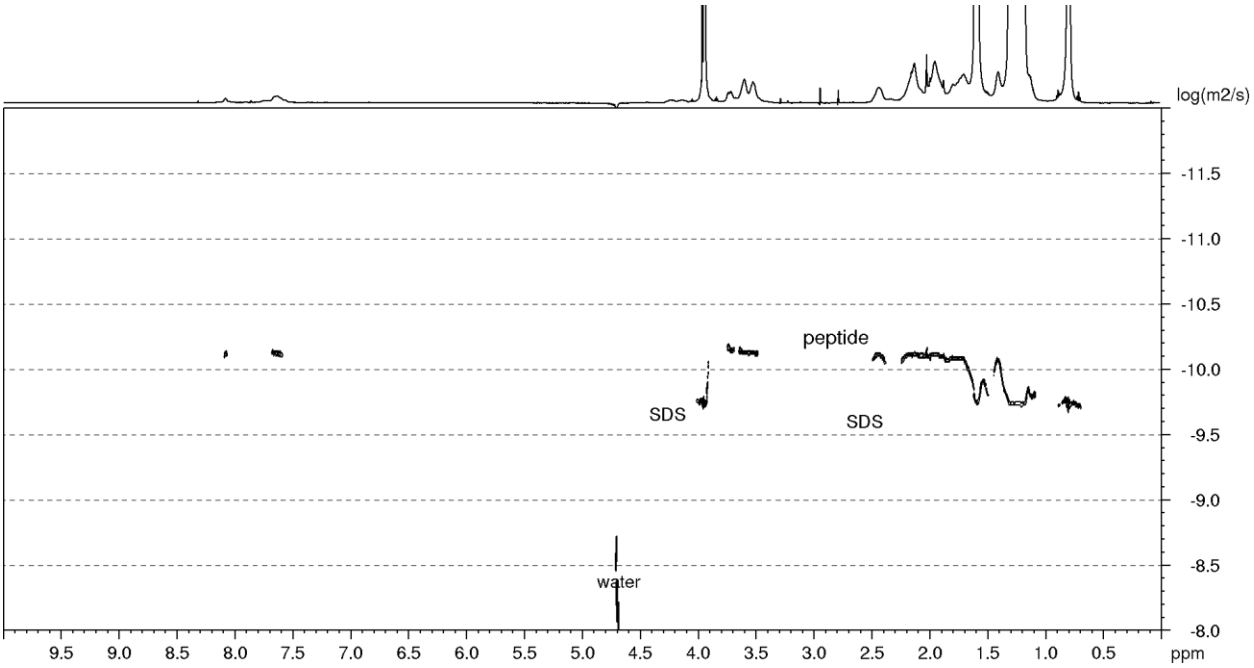


Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

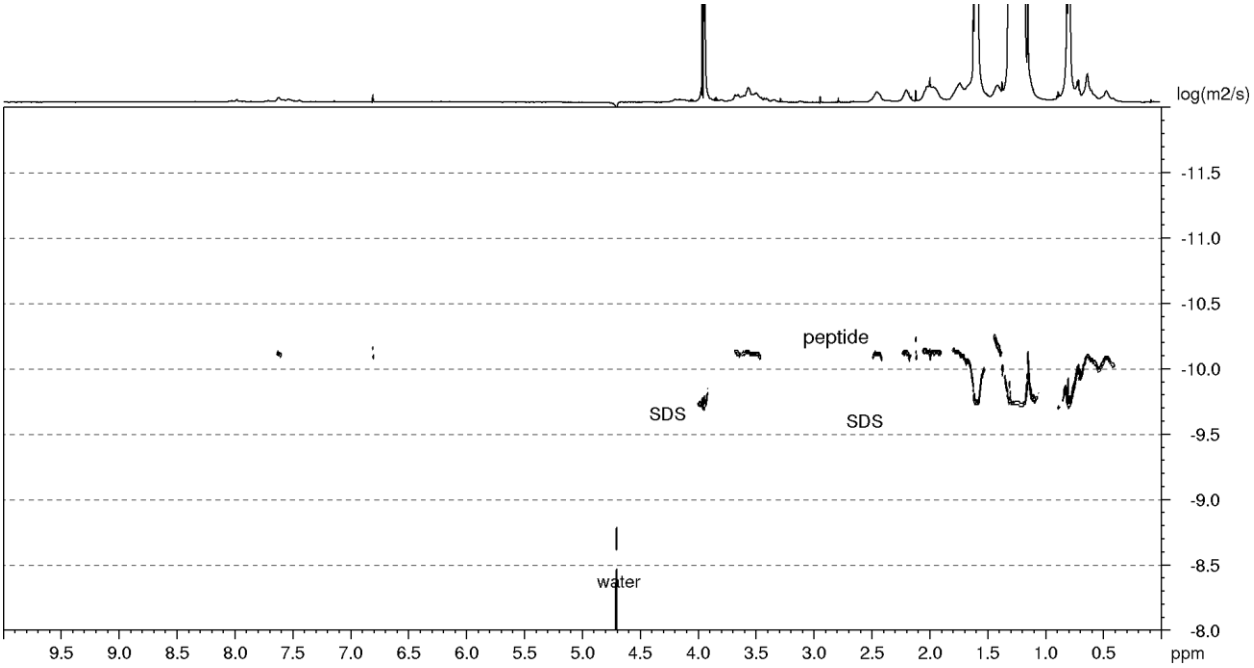
The diffusion analysis in sodium dodecyl sulphate (SDS) micelles was performed at 0.6 mM peptide and 22.5 mM SDS in H_2O / D_2O 10:1 mixture at 25 °C. Considering that a 'normal' SDS micelle consists of roughly 65 anions, this would make approx. 2 peptide molecules per micelle. Resulting particle size is approx. 24 kDa, and if this would be a spherical particle the diffusion coefficient (log*D*) would be about −9.89 log m^2 s^{-1} according to a previously derived equation.[S1] However, the linear size of the peptide seriously exceeds the micelle diameter (by more than twice). As the result, we expect some enlargement of the micelles. This can be concluded from observed slowed down detergent diffusion from about −9.60 log m^2 s^{-1} to about −9.72 log m^2 s^{-1}. The observed diffusion coefficient was −10.10±0.03 for **AcPOG10**, −10.11±0.03 for **PivσOG10**, −9.74±0.02 for SDS. This diffusion data makes us conclude that the peptides reside in micelles, since dissolved peptide log*D* would be about −9.60 log m^2 s^{-1}, which is much faster diffusion that the one observed.

[S1] V. Kubyshkin and N. Budisa, *J. Peptide Sci.*, 2018, **24**, e3076, doi: 10.1002/psc.3076

^{1}H DOSY for **AcPOG10** in 22.5 mM SDS in H_2O/D_2O 10:1 at 25 °C, δ = 50 ms, $\Delta/2$ = 4 ms:

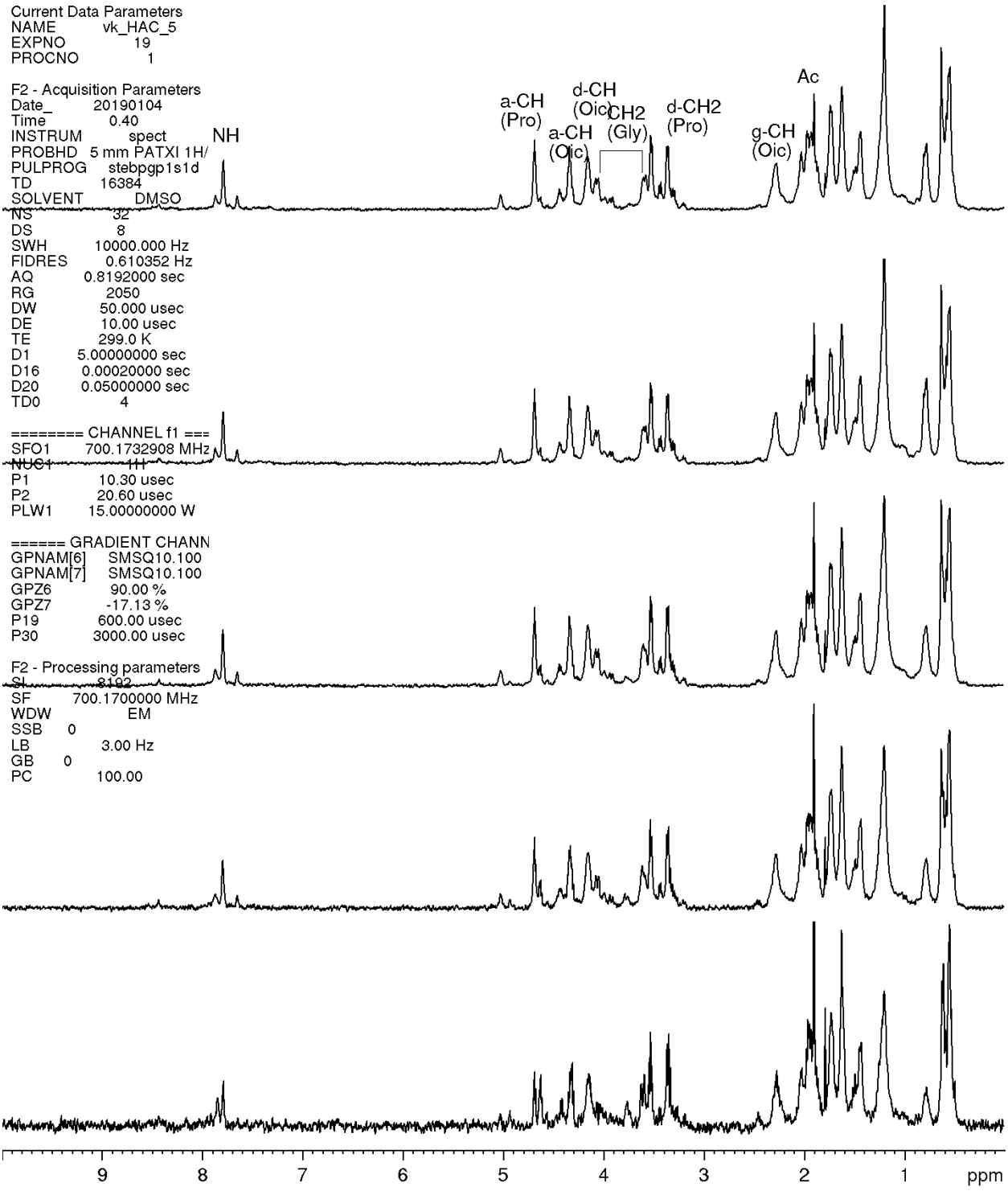


^{1}H DOSY for **PivσOG10** in 22.5 mM SDS in H_2O/D_2O 10:1 at 25 °C, δ = 50 ms, $\Delta/2$ = 4 ms:



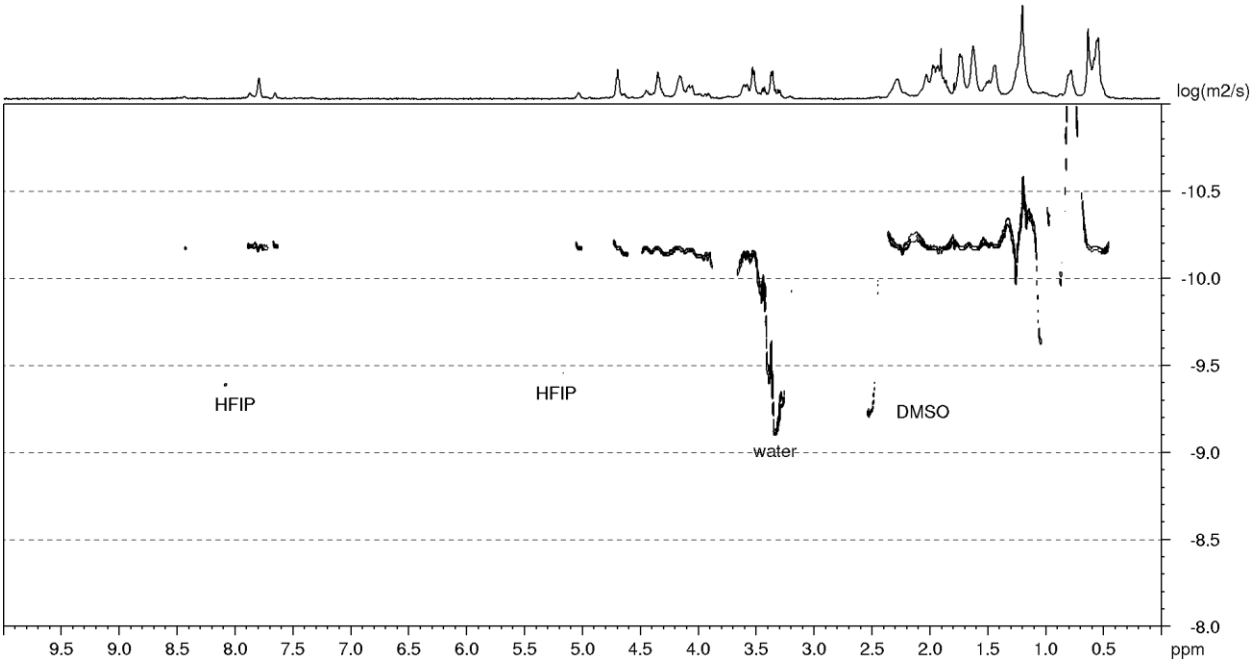
Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

^{1}H NMR spectra of the peptide series **AcσOG2**, **AcσOG4**, **AcσOG6**, **AcσOG8**, **AcσOG10** (from bottom to the top). The spectra were acquired in DMSO-d_6 solution at 4 mg ml^{-1} concentration.

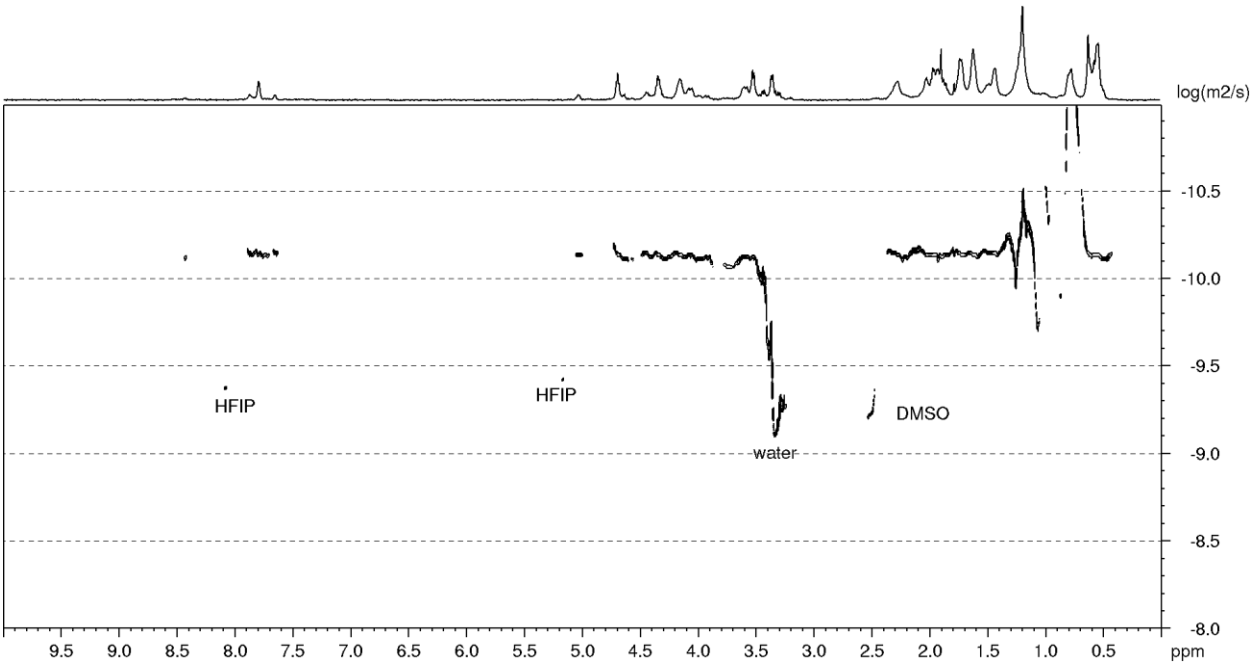


Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

^{1}H DOSY for **AcσOG10** (DMSO-d_6, 700 MHz), δ = 50 ms, Δ/2 = 4 ms, 25 °C:

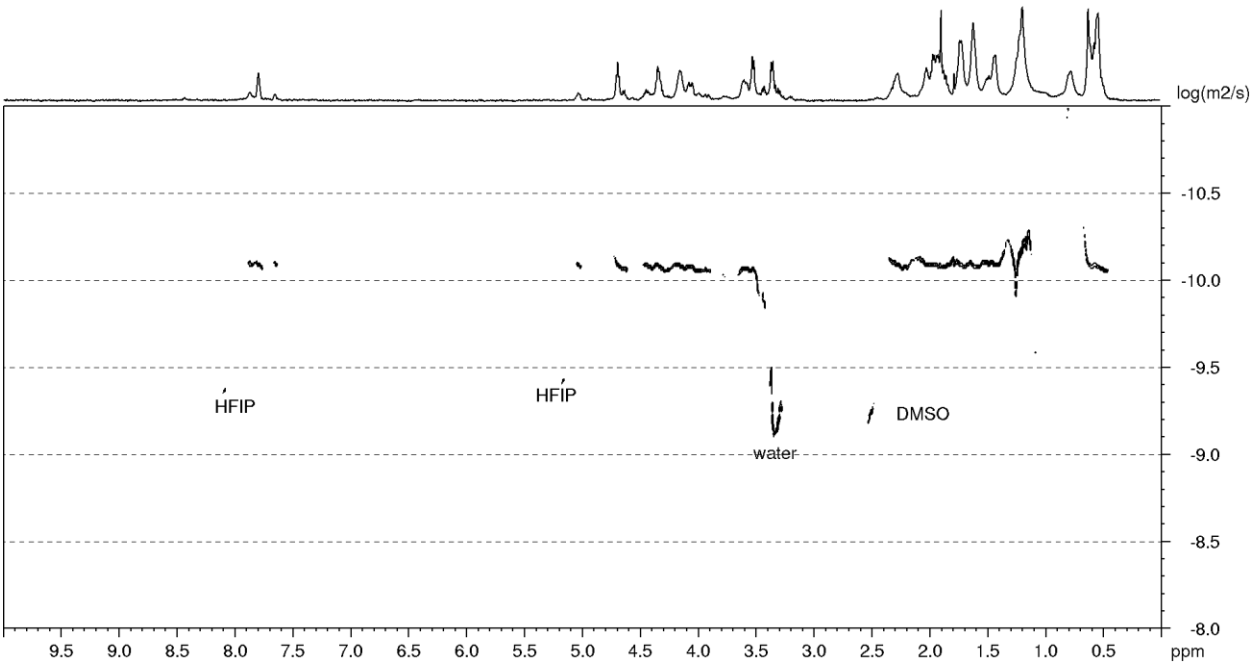


^{1}H DOSY for **AcσOG8** (DMSO-d_6, 700 MHz), δ = 50 ms, Δ/2 = 4 ms, 25 °C:

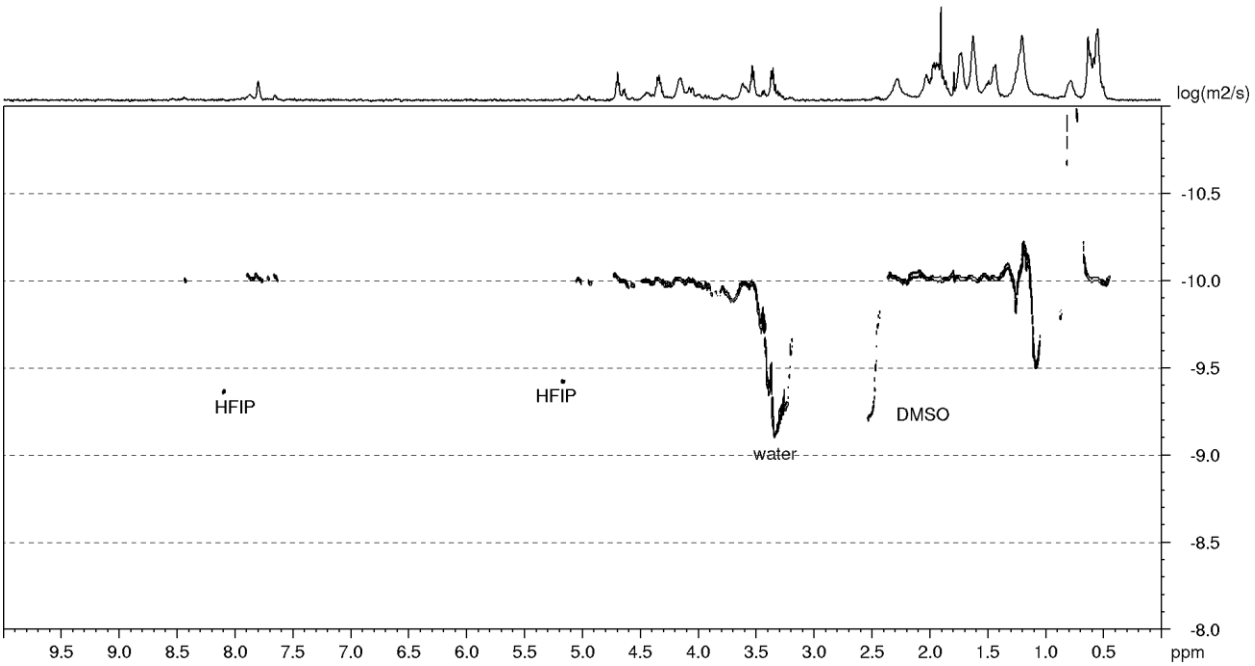


Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

^{1}H DOSY for **AcσOG6** (DMSO-d_6, 700 MHz), δ = 50 ms, Δ/2 = 4 ms, 25 °C:

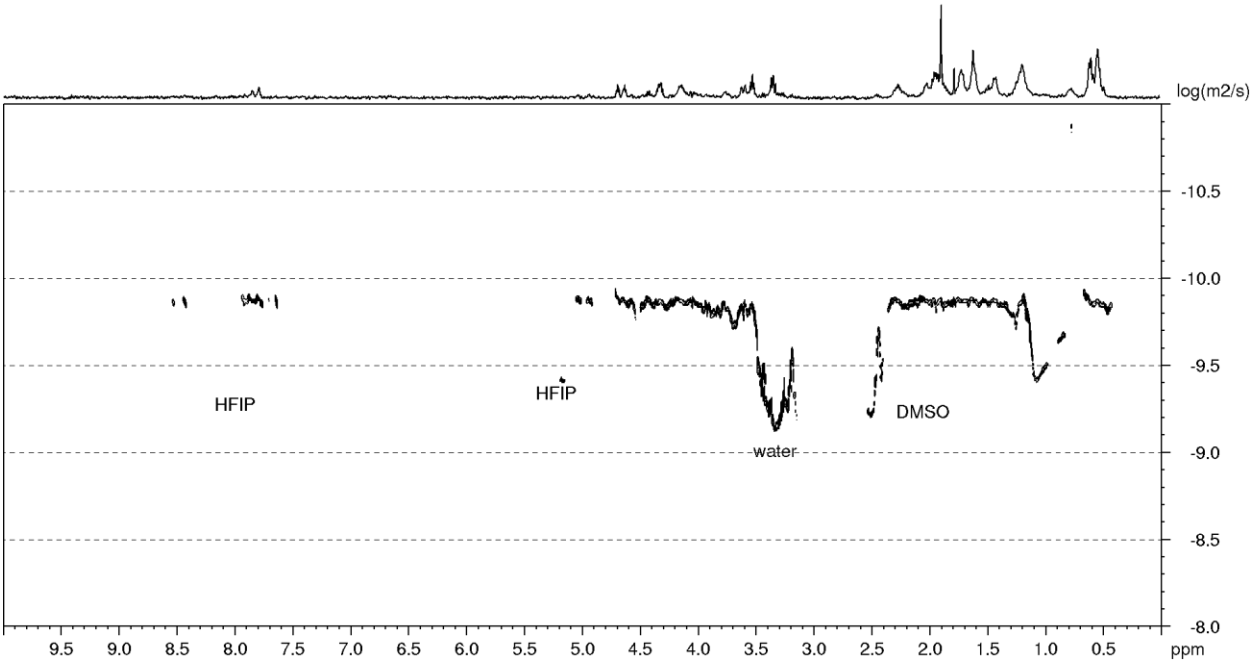


^{1}H DOSY for **AcσOG4** (DMSO-d_6, 700 MHz), δ = 50 ms, Δ/2 = 4 ms, 25 °C:



Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

^{1}H DOSY for **AcσOG2** (DMSO-d_6, 700 MHz), δ = 50 ms, Δ/2 = 4 ms, 25 °C:



Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

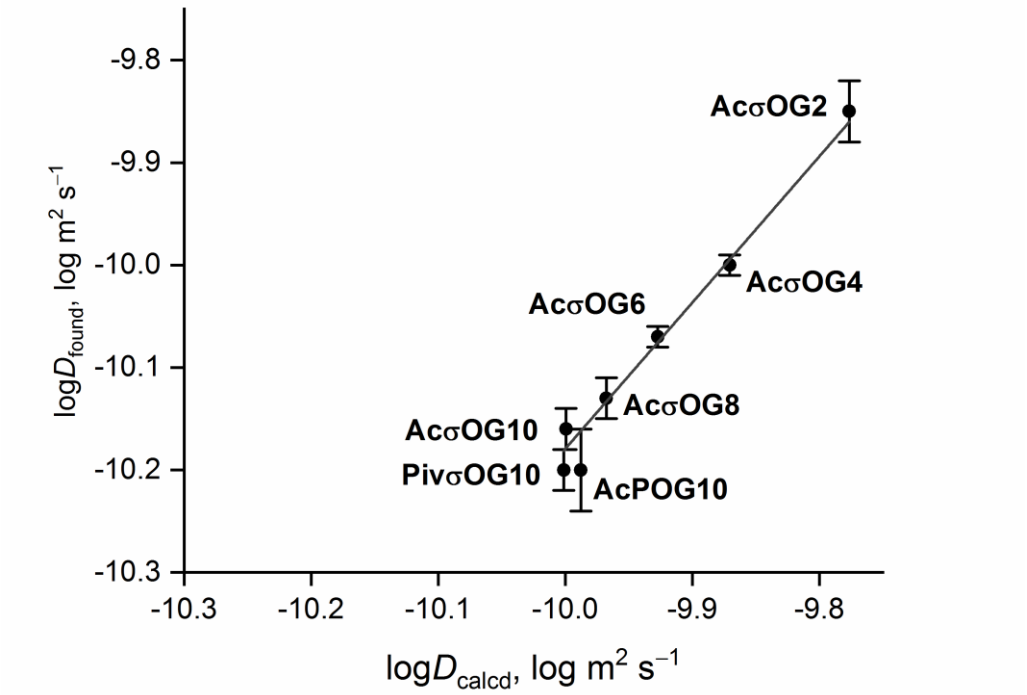
Analysis of the diffusion data

The experimental diffusion coefficients were determined in ^{1}H NMR spectra at 700 MHz in DMSO-d_6 solutions at 25 °C. Stimulated echo with bipolar gradients was performed in a pseudo-2D mode with diffusion time δ = 50 ms, gradient pulse $\Delta/2$ = 4 ms, linear gradient increment 10→98% over 256 data points. The spectra were processed conventionally in direct dimension, and conversion to log*D* projection was done using the software tool provided with the spectrometer machine. The peptide log*D* values were read out, and the half-high width of the log*D* projection was taken as the spread rather than error value.

Theoretical value was calculated using the rigid sphere Stokes-Einstein approximation, and the dynamic viscosity value for DMSO-d_6 η = 0.00219 kg m^{-1} s^{-1}.[S2] The molecular volumes were calculated from molecular weights using the equation previously derived for oligo-Oic peptides.[S3] The final equation is as follows:

$\log D_{calcd}$ [log m^2 s^{-1}]= −((log(1.223·MW [Da]))/3)−8.794

The experimental values are overall higher than the theoretical values, due to the deviations from the spherical behavior. The correlation with the spread values is as shown below:



Fitting to the correlation generally illustrates the homogeneity of the ^{1}H NMR signal in the spectra of the peptides.

[S2] R. Evans, Z. Deng, A. K. Rogerson, A. S. McLachlan, J. J. Richards, M. Nilsson and G. A. Morris, *Angew. Chem. Int. Ed.*, 2013, **52**, 3199-3202, doi: 10.1002/anie.201207403
[S3] V. Kubyshkin and N. Budisa, *Org. Biomol. Chem.*, 2017, **15**, 619-627, doi: 10.1039/C6OB02306A

Table S1. Summarized analytical data for the collagen mimicking peptides

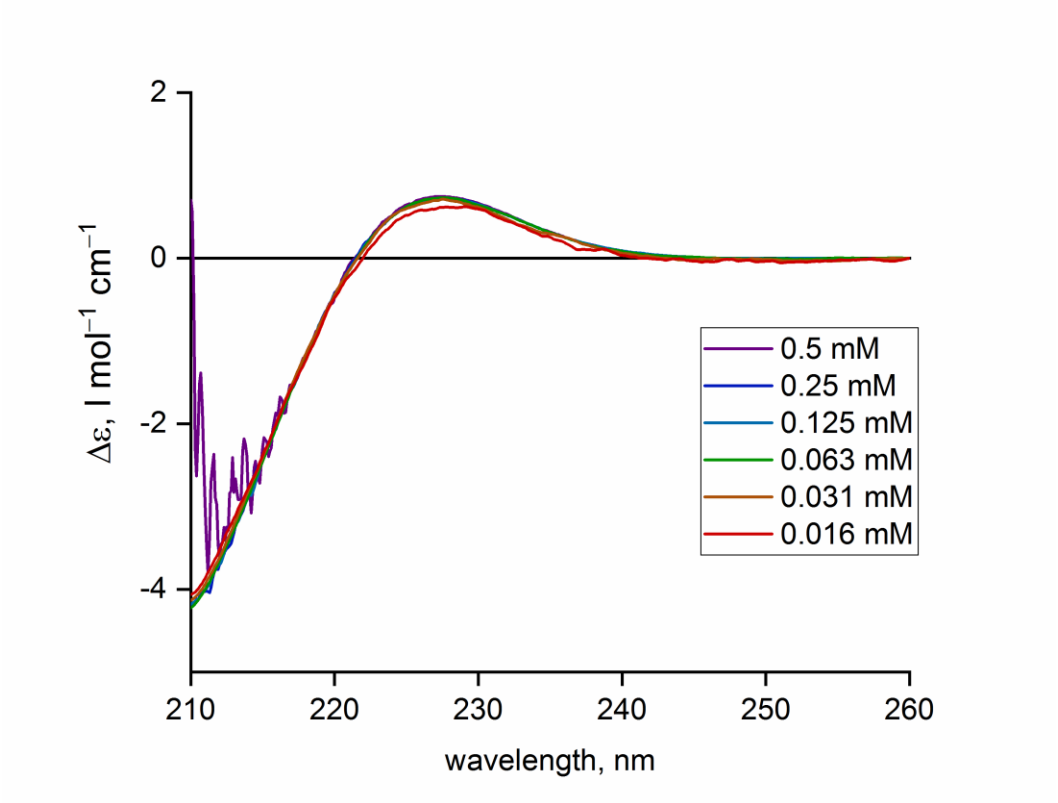
ID	Expected mass, Th	Observed mass, Th ESI-Orbitrap	log*D* in DMSO at 25 °C, log m^2 s^{-1}
AcPOG10	$[M+3H]^{3+}$ 1038.3952	1038.5972	−10.20
	$[M+2H]^{2+}$ 1557.3892	1557.3905	
PivσOG10	$[M+3H]^{3+}$ 1139.3297	1139.3302	−10.21
	$[M+2H]^{2+}$ 1708.4910	1708.4901	
	$[M+1H]^{+}$ 3415.9746	3415.9717	
AcσOG10	$[M+3H]^{3+}$ 1125.3141	1125.3162	−10.16
	$[M+2H]^{2+}$ 1687.4675	1687.4686	
	$[M+1H]^{+}$ 3373.9277	3373.9275	
AcσOG8	$[M+3H]^{3+}$ 904.5210	904.5218	−10.13
	$[M+2H]^{2+}$ 1356.2779	1356.2784	
AcσOG6	$[M+2H]^{2+}$ 1025.0883	1025.0885	−10.07
	$[M+1H]^{+}$ 2049.1693	2049.1663	
AcσOG4	$[M+2H]^{2+}$ 693.3970	693.3983	−10.00
	$[M+1H]^{+}$ 1385.7868	1385.7878	
AcσOG2	$[M+1H]^{+}$ 723.4076	723.4086	−9.85

Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

Circular dichroism spectra acquisition

The circular dichroism (CD) spectra were recorded in 1 mm quartz cell at 25 °C. The peptides amounts were determined gravimetrically. The peptide weights were approx. 0.20-0.60 mg scale. These were then mixed with calculated amounts of the solvents to give final desired concentration. The solvents were: 2,2,2-trifluoroethanol, 25 mM SDS in water, octan-1-ol. Resulting solutions were kept at the ambient temperature for about 24 hours prior to the measurements. For the initial solvent screening we set peptide concentration 0.2 mM, which is typically used for CMPs, because this concentration is higher than common critical concentration of the triple helical assembly. The spectra are as shown in the main text. In this way, we found that a stable triple helical assembly occurs in octan-1-ol (see Fig. 3A in the main text). For calculation of the CD we considered concentration of the amide bonds rather than the peptide concentration. Thus, our reported $\Delta\varepsilon$ values are equivalent to conventional mean-residue-elipticities (with a factor 3298).

For the peptide **PivσOG10** we tested dilution series. We dissolved the peptide at concentration 0.5 mM, and this was diluted twice to 0.25, 0.125, 0.063, 0.031, 0.016 mM concentrations, and this demonstrated no loss of the structure as shown below:



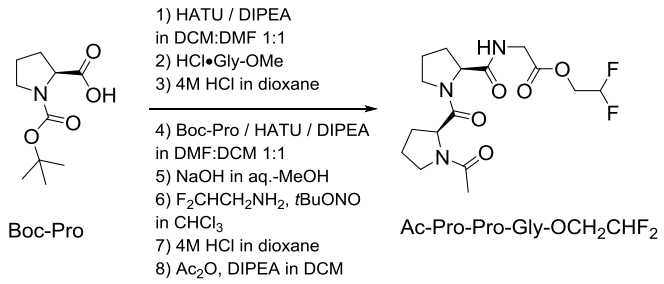
As we observed that the concentration can be set low, for the oligomer series, we tested 0.2 mM for each peptide as well as concentration of 1 mM amide, which is 0.033 mM for **AcσOG10**, 0.042 for **AcσOG8**, 0.056 mM for **AcσOG6**, 0.083 for **AcσOG4**, 0.167 mM for **AcσOG2**. Conclusions for both series were same.

The temperature measurements were at 227 nm, with 5 to 95 °C heating series. Sampling was at 0.5 °C, heating rate was 0.3 nm min^{-1} for **AcPOG10** and **PivσOG10** (Fig. 3B), and 1 nm min^{-1} for the oligomeric series **AcσOG2**, **AcσOG4**, **AcσOG6**, **AcσOG8** and **AcσOG10** (Fig. 4B).

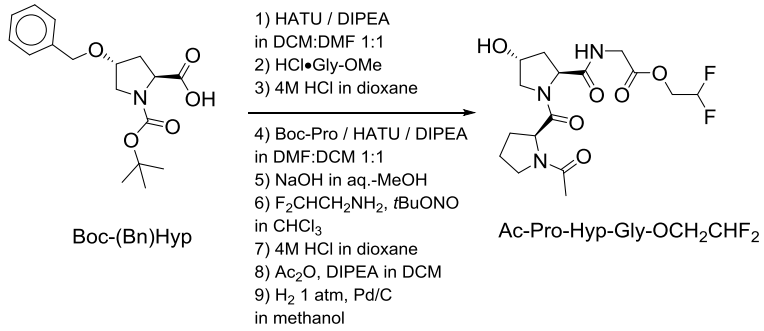
Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

Peptide partitioning

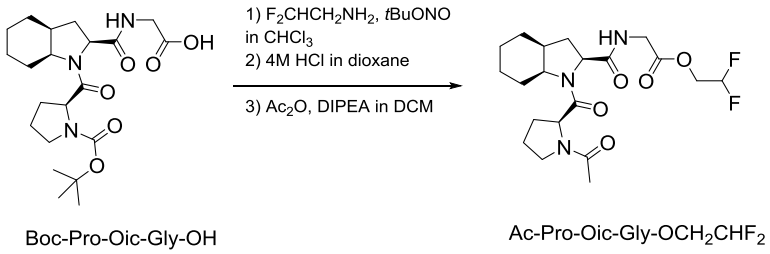
The peptides were synthesized in solution on a small scale (30-200 mg) as schematically shown below:



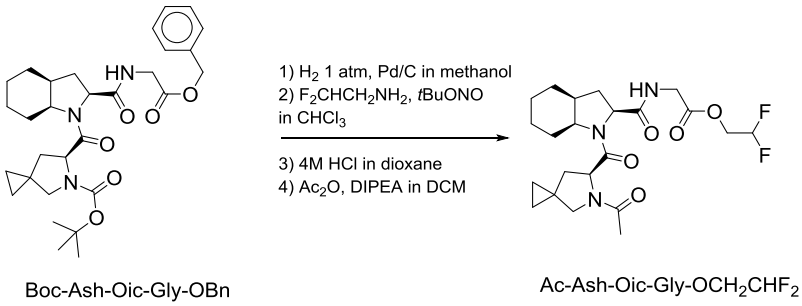
Mass-spectrum (ESI-Orbitrap): calcd. $[M+H]^+$ 376.1679, $[M+Na]^+$ 398.1498, found 376.1673, 398.1490.



Mass-spectrum (ESI-Orbitrap): calcd. $[M+H]^+$ 392.1628, $[M+Na]^+$ 414.1447, found 392.1644, 414.1463.



Mass-spectrum (ESI-Orbitrap): calcd. $[M+H]^+$ 430.2148, $[M+Na]^+$ 452.1967, found 430.2139, 452.1955.



Mass-spectrum (ESI-Orbitrap): calcd. $[M+H]^+$ 456.2305, found 456.2294.

Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d

The peptides were partitioned as follows. Octan-1-ol (1.00 ml) and water (1.00 ml) were added to a peptide (5 mg). The mixture was shaken for 20-24 h at the ambient temperature (22 °C). Three samples were partitioned for each analyzed peptide. Aliquotes of each phase (0.30 ml) were taken carefully, acetonitrile-d_3 (0.30 ml) was added to each sample. The samples were mixed in same type 5 mm NMR tubes. NMR spectra were recorded in a well-tuned BBFO probe at 25 °C. The $^{19}F\{^{1}H\}$ NMR spectra were recorded at 471 MHz in one-pulse sequence with inverse-gated decoupling (during acquisition), the transmitter was centered around −127 ppm. ^{19}F resonances were integrated and the absolute integral values were compared. The ratio was considered a partitioning constant, P. The values from the sample repeats, as well as ^{19}F NMR repeats with different settings were averaged to give the final value, and the root-mean square deviations are considered the error values.

Org. Biomol. Chem., **2019**, doi: 10.1039/c9ob00070d