Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry

Electronic Supplementary Information

Synthesis of Homogeneous MUC1 Oligomers *via* a Bi-Directional Ligation Strategy

*Dima Al Sheikha, Brendan L. Wilkinson, Gajan Santhakumar, Morten Thaysen-Andersen and Richard J. Payne**
School of Chemistry, The University of Sydney, NSW 2006, AUSTRALIA
**richard.payne@sydney.edu.au*

Table of Contents

1.0 Preparative HPLC Gradients

Preparative reversed-phase HPLC was performed using a Waters 600 Multisolvent Delivery System and Waters 500 pump with 2489 photodiode array detector or Waters 486 Programmable wavelength detector operating at 230 and 214 nm. All purifications used a mobile phase of 0.1% TFA in water (Solvent A) and 0.1% TFA in acetonitrile (Solvent B) unless otherwise noted using a variety of gradients and columns as detailed below:

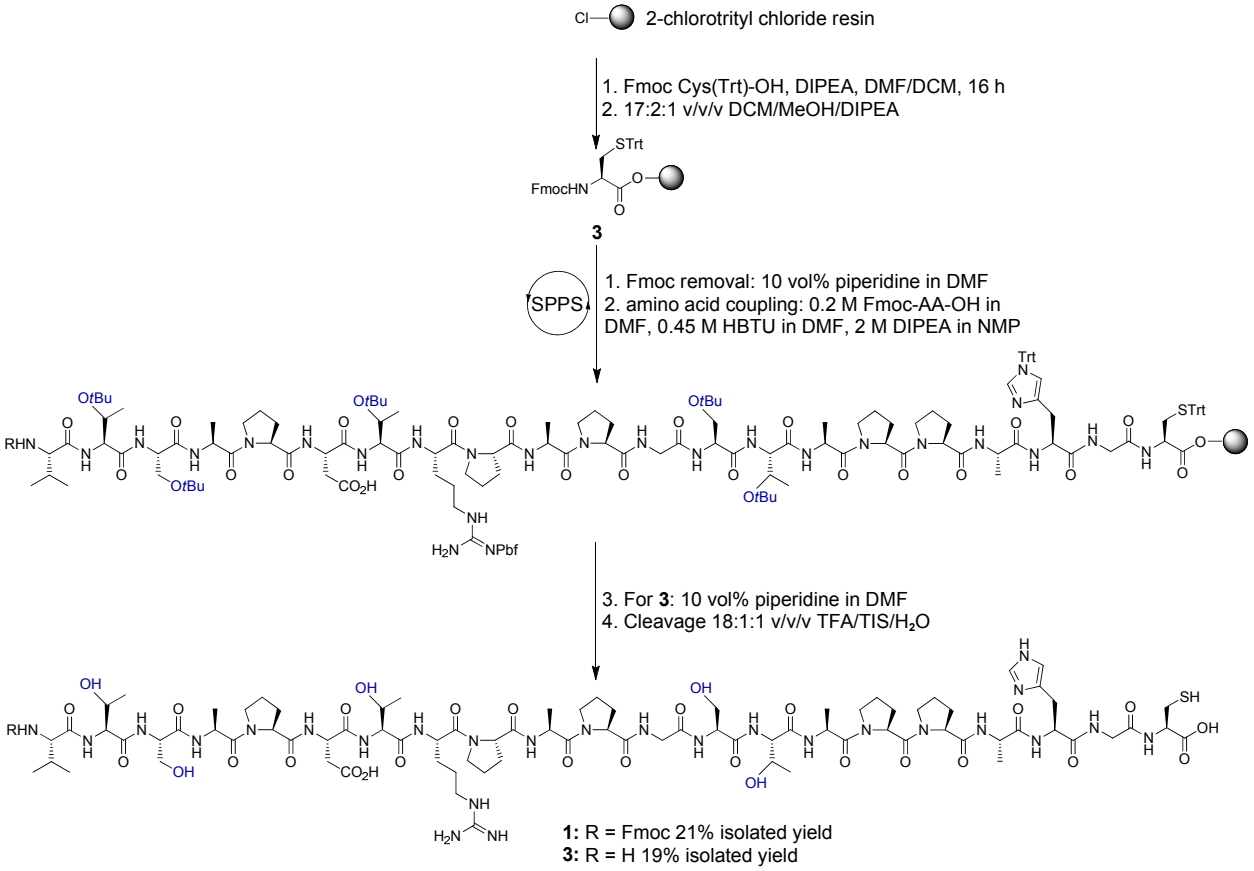
***Method A*:** Purification was performed on a Waters Sunfire™ 5 µm (C-18) preparative column operating at a flow rate of 7 mL min^{-1} using a linear gradient of 0% to 50% B over 40 min. ***Method B*:** Purification was performed on a Waters Sunfire™ 5 µm (C-18) preparative column operating at a flow rate of 7 mL min^{-1} using a linear gradient of 0% to 40% B over 60 min. ***Method C*:** Purification was performed on a Grace Vydac™ 10 µm (C-18) semi-preparative column operating at a flow rate of 4 mL min^{-1} using a linear gradient of 0% for 10 min to 45% B over 60 min. ***Method D*:** Purification was performed on a Waters Sunfire™ 5 µm (C-18) preparative column operating at a flow rate of 7 mL min^{-1} using a linear gradient of 0% to 30% B over 60 min a mobile phase of 0.1% formic acid in water (Solvent A) and 0.1% formic acid in acetonitrile (Solvent B). ***Method E*:** Purification was performed on a Grace Vydac™ 5 µm (C-4) semi-preparative column operating at a flow rate of 4 mL min^{-1} using a linear gradient of 0% for 10 min to 45% B over 60 min.

***Method F*:** Purification was performed on a Waters Sunfire™ 5 µm (C-18) semi-preparative column operating at a flow rate of 4 mL min^{-1} using a linear gradient of 0% for 10 min to 30% B over 60 min. ***Method G*:** Purification was performed on a Waters Sunfire™ 5 µm (C-18) preparative column operating at a flow rate of 7 mL min^{-1} using a linear gradient of 0% for 10 min to 50% B over 60 min. ***Method H*:** Purification was performed on a Grace Vydac™ 5 µm (C-4) semi-preparative column operating at a flow rate of 4 mL min^{-1} using a linear gradient of 0% to 15% B over 60 min. ***Method J*:** Purification was performed on a Waters Sunfire™ 5 µm (C-18) preparative column operating at a flow rate of 7 mL min^{-1} using a linear gradient of 0% for 10 min to 50% over 60 min.

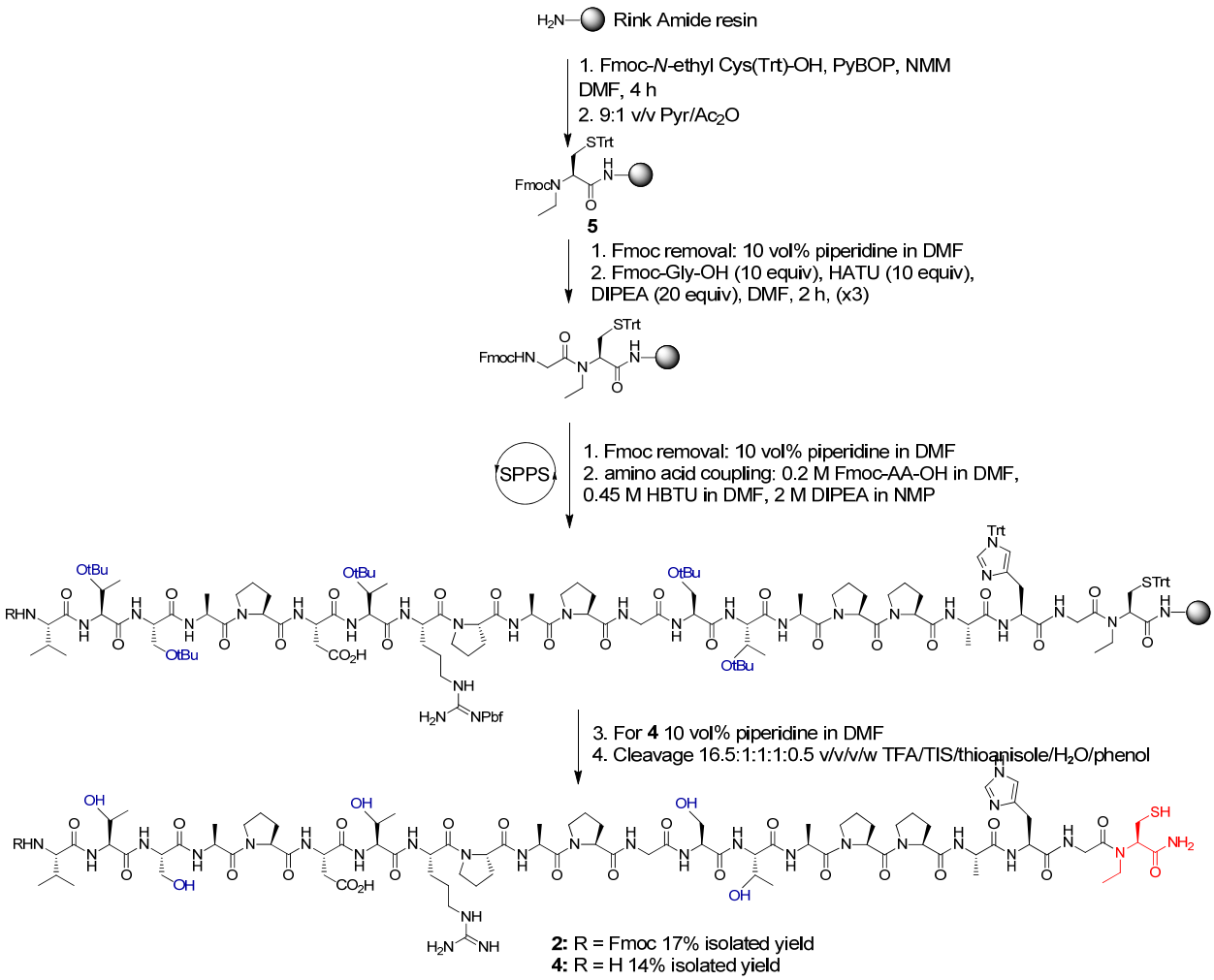
2.0 SPPS of peptides 1-4

Peptides **1-4** were synthesised using microwave assisted Fmoc-SPPS (Scheme 1 and 2).

2.1 Solid-phase peptide synthesis (SPPS) of MUC1 peptides 1 and 3



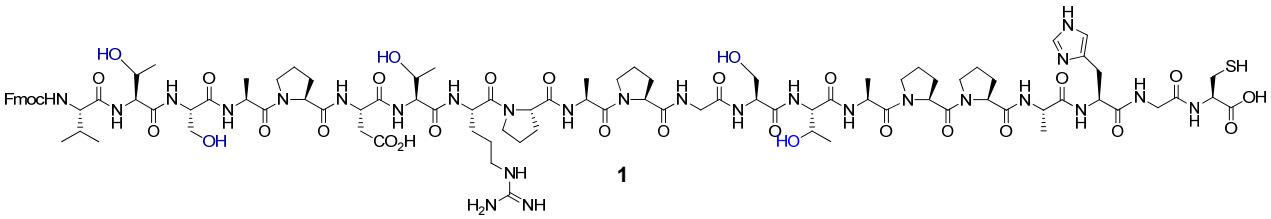
Scheme S1. Microwave-assisted Fmoc-SPPS of peptides **1** and **3.**



Scheme S2. Fmoc-SPPS of peptides **2** and **4** containing a C-terminal N-ethyl cysteine residue.

3.0 Analytical data

Peptide 1: Fmoc-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-Cys-OH (1)



1

Peptide **1** was prepared according to Fmoc-strategy SPPS on 100 μmol of 2-Cl Trt Cl resin as outlined in the general procedure and purified by preparative reversed-phase HPLC (*Method A*) to give a white solid following lyophilisation (45 mg, 21% isolated yield based on the original resin loading).

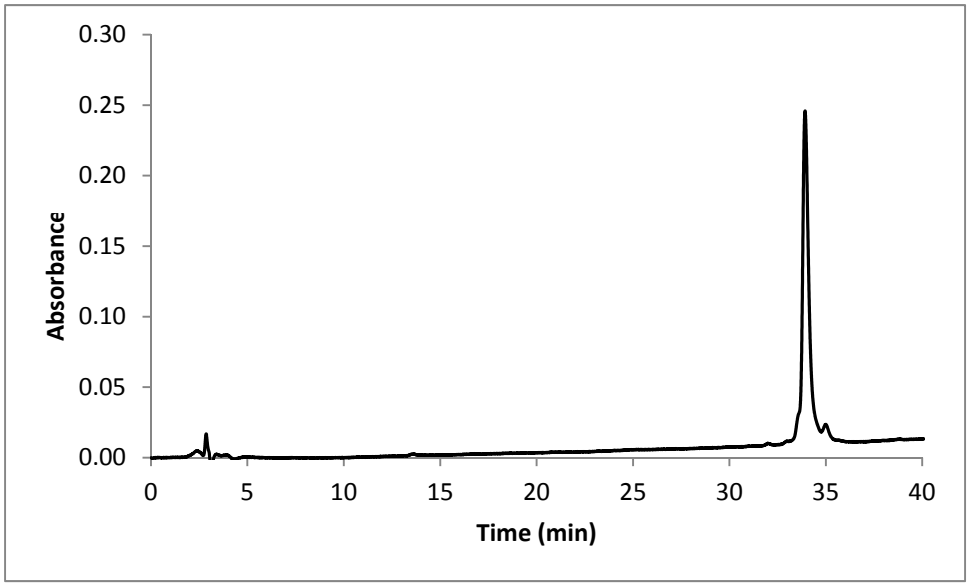


Figure S1. Analytical HPLC of peptide **1**: R_t = 34 min (0 to 50% B over 40 min, λ = 230 nm).

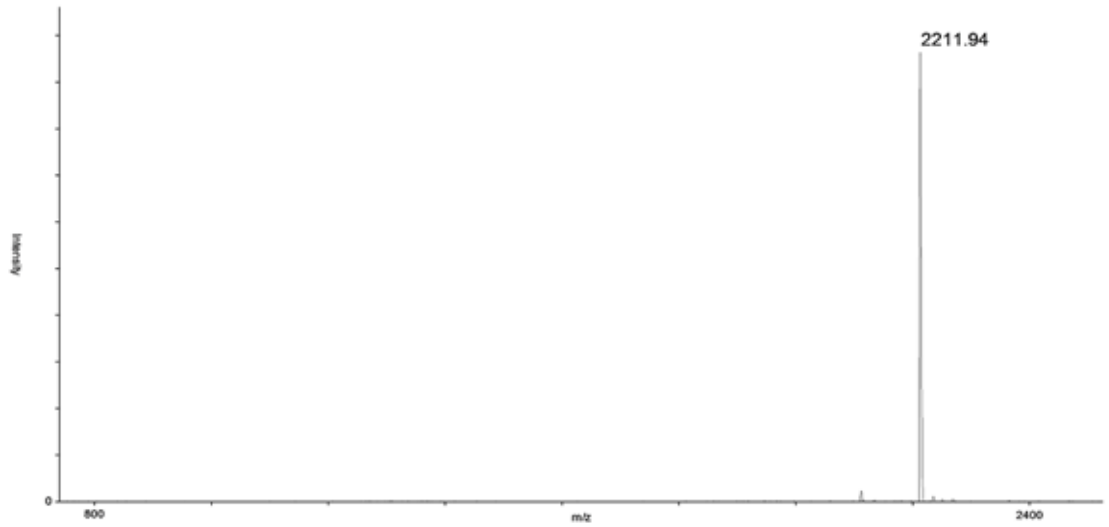


Figure S2. MALDI-Tof mass spectrum of peptide **1**.

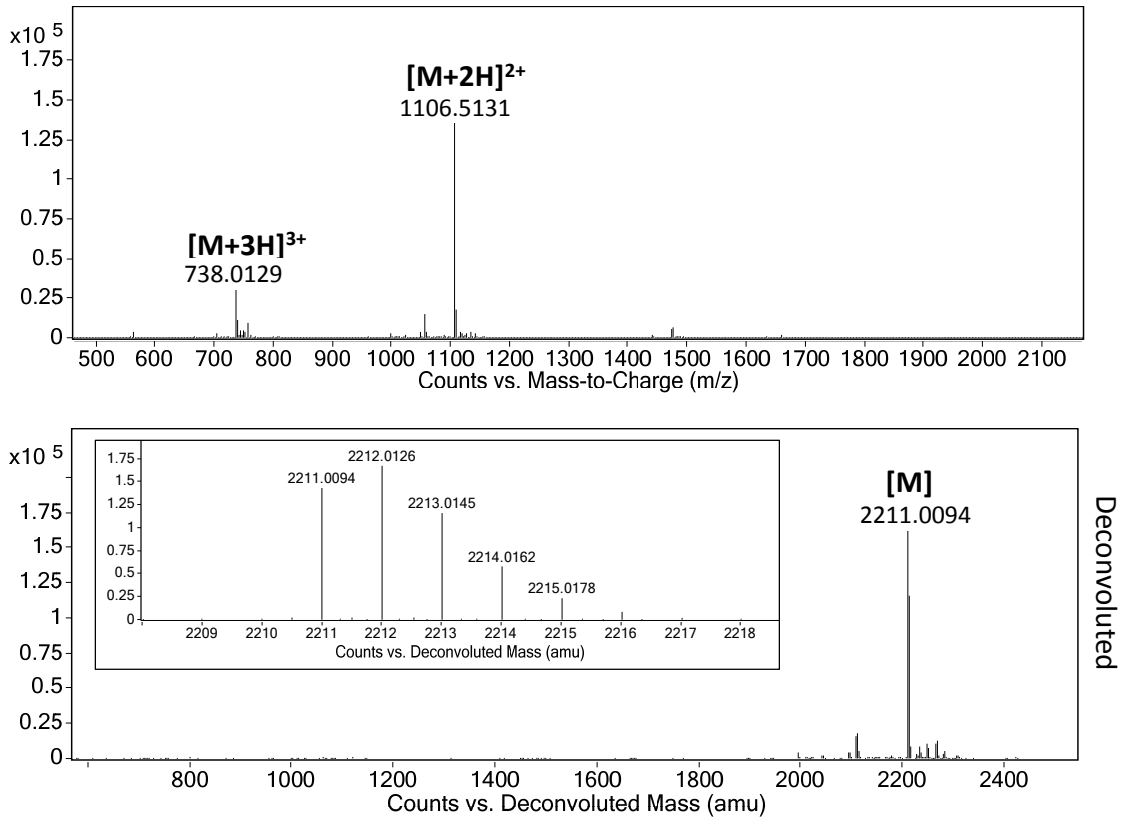
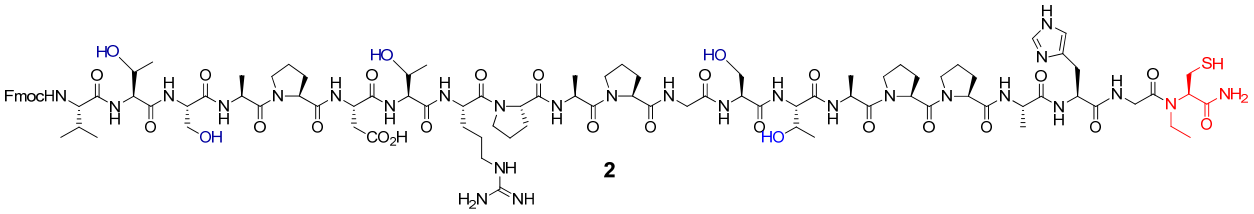


Figure S3. ESI-Q-Tof mass spectrum (top) and deconvoluted spectrum (bottom) with zoomed region (inset) of peptide 1.

Peptide 2: Fmoc-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-Cys(*N*-ethyl)-NH_2 (2)



Peptide **2** was prepared according to Fmoc-strategy SPPS on 100 μmol of 2-Cl Trt Cl resin as outlined in the general procedure and purified by preparative reversed-phase-HPLC (*Method A*) to give a white solid following lyophilisation (36 mg, 17% isolated yield based on the original resin loading).

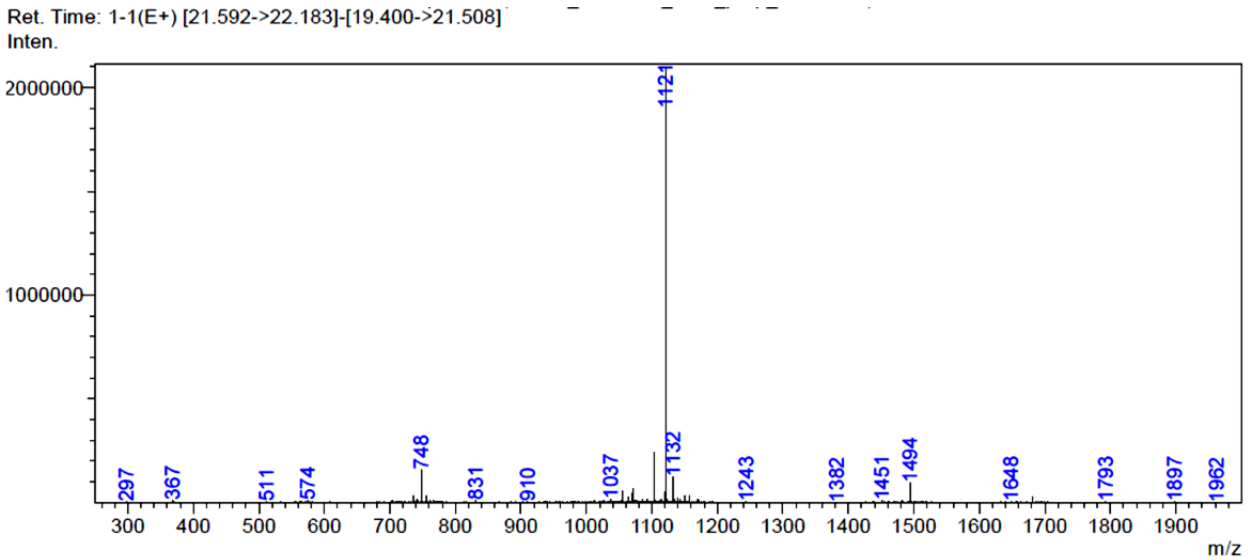
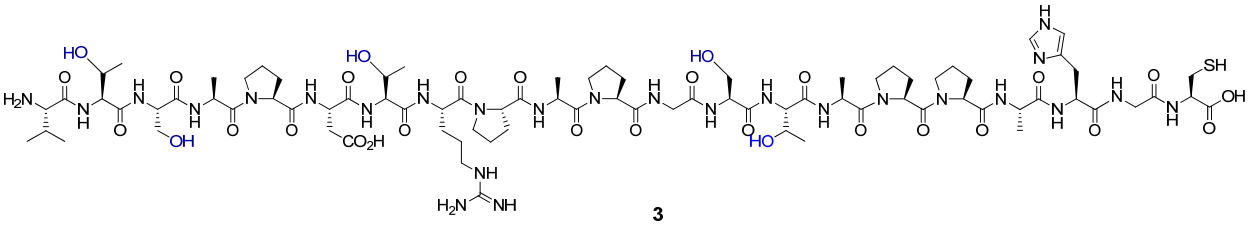


Figure S4. ESI-MS spectrum of peptide **2**

Peptide 3: H-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-Cys-OH (3)



3

Peptide **3** was prepared according to Fmoc-strategy SPPS outlined in the general procedure purified by preparative RP-HPLC (*Method B*) to give a white solid after lyophilisation (36 mg, 19% isolated yield based on the original resin loading).

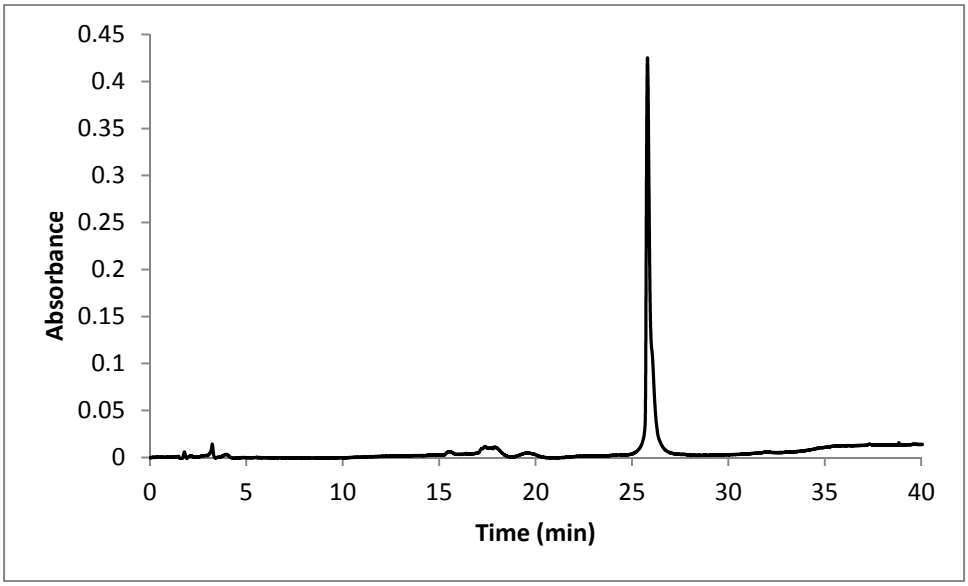


Figure S5. Analytical HPLC chromatogram of peptide **3**: R_t = 26 min gradient (0 to 50% B over 40 min, λ = 230 nm).

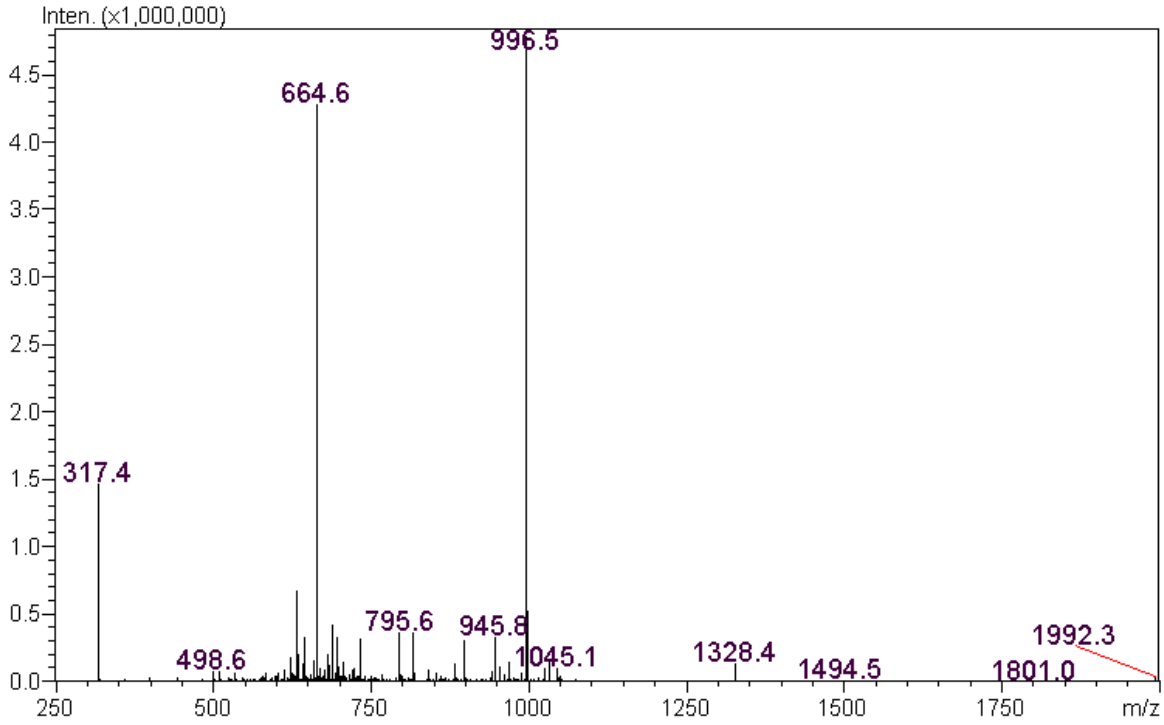
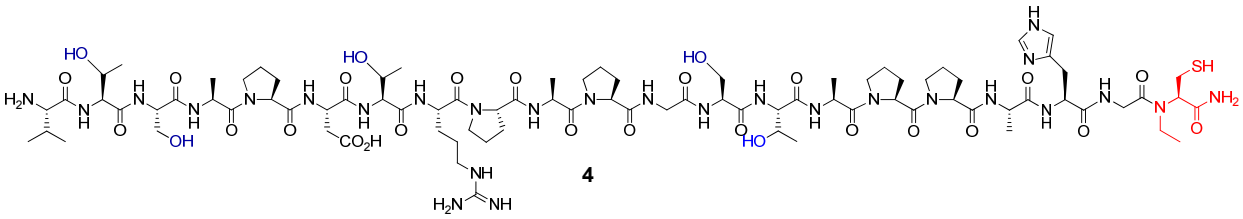


Figure S6. ESI-MS spectrum of peptide **2**.

Peptide 4: H-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-Cys(*N*-ethyl)-NH_2 (4)



4

Peptide **4** was prepared according to Fmoc-strategy SPPS outlined in the general procedure and purified by preparative RP-HPLC (*Method B*) to give a white solid after lyophilisation (26 mg, 14% isolated yield based on the original resin loading).

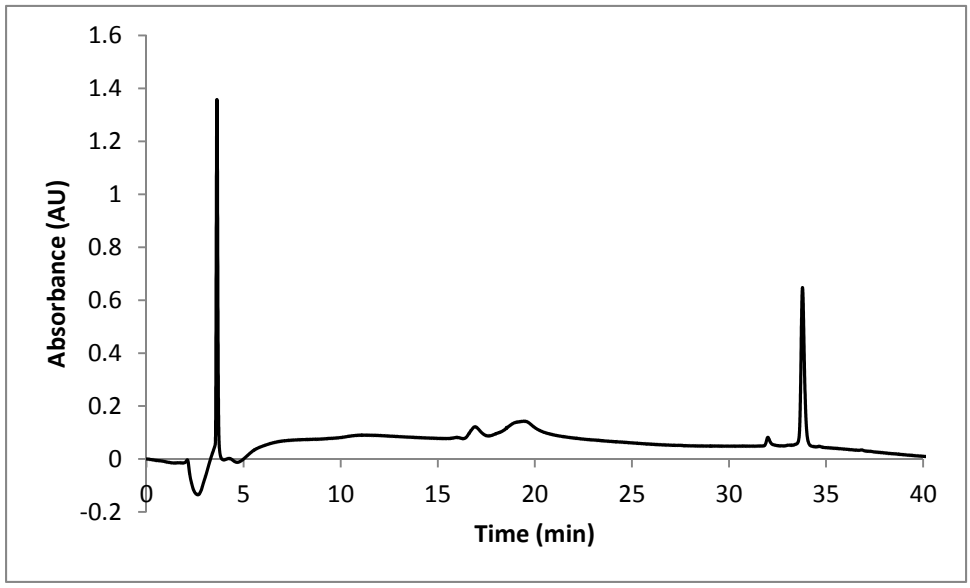


Figure S7. Analytical HPLC chromatogram of peptide **4**: R_t = 33.7 min gradient (0 to 50% B over 40 min, λ = 230 nm).

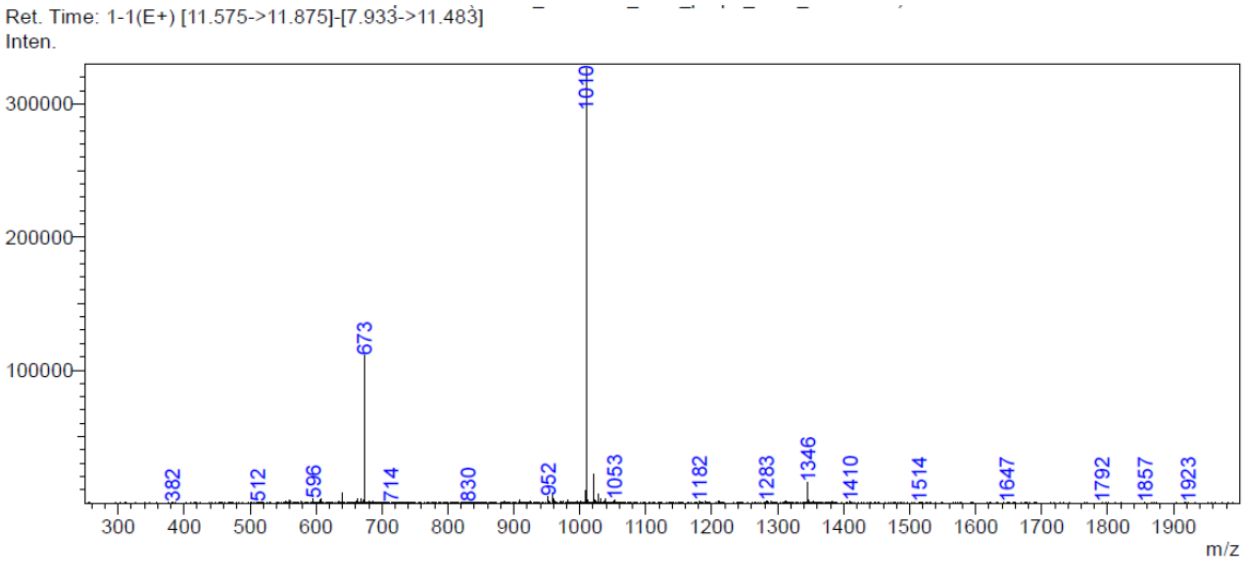
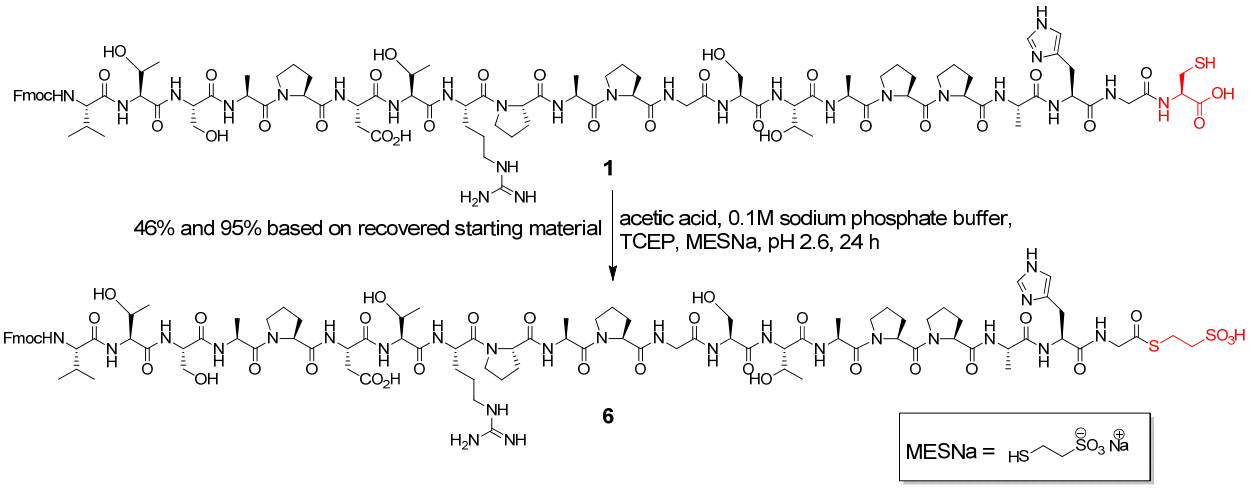


Figure S8. ESI-MS spectrum of peptide **4.**

Peptide thioester 6: Fmoc-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly- $S(CH_2)_2SO_3H$ (6)

Peptide **1** (25 mg, 11.3 µmol, 1.0 equiv.) was subjected to the N→S thioesterification conditions outlined in the general procedure to afford peptide thioester **6** as a white solid following purification by RP-HPLC (*Method B*) and lyophilisation (Scheme 3, 11.6 mg, 46% isolated yield).



Scheme S3. Synthesis of peptide thioester **6** by N→S acyl transfer.

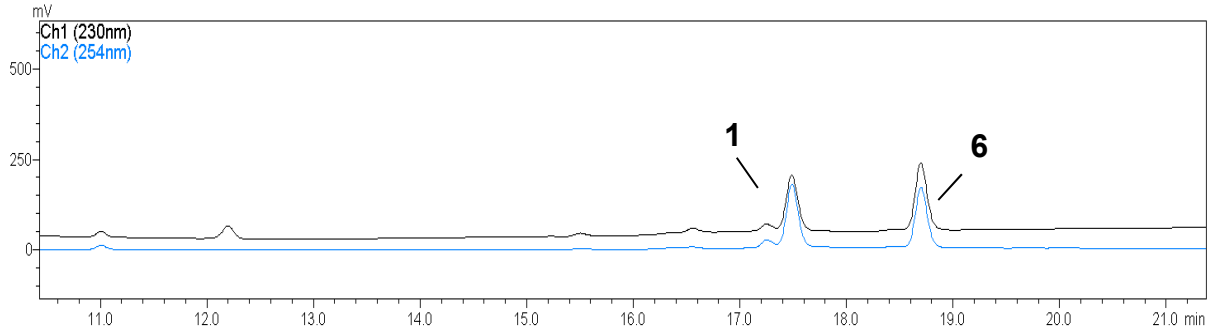


Figure S9. Chromatogram extracted from LC-MS of N→S thioesterification reaction after 24 h. (MESNa elutes at R_t =2 min).

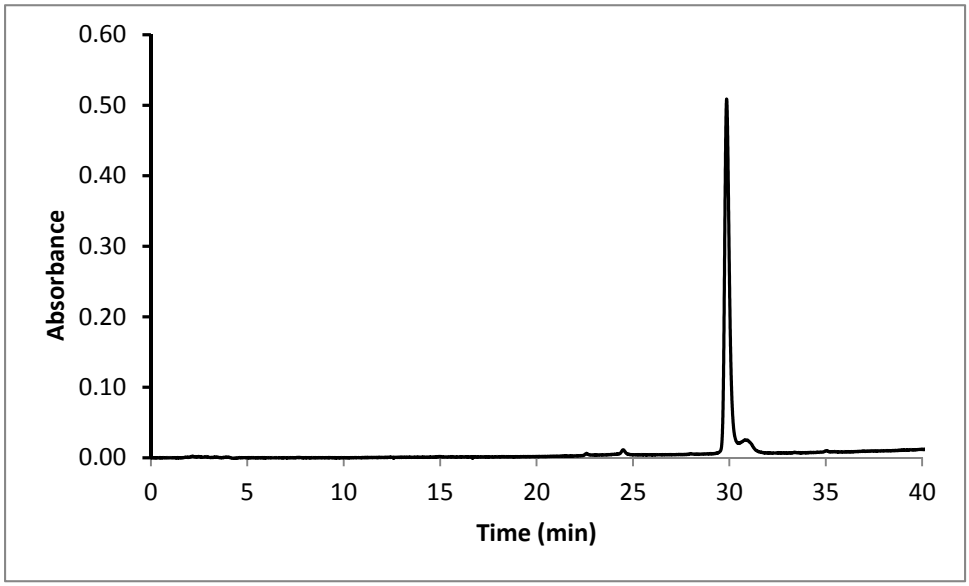


Figure S10. Analytical HPLC chromatogram of peptide thioester **6**: R_t = 34 min (0 to 50% B over 40 min, λ = 230 nm).

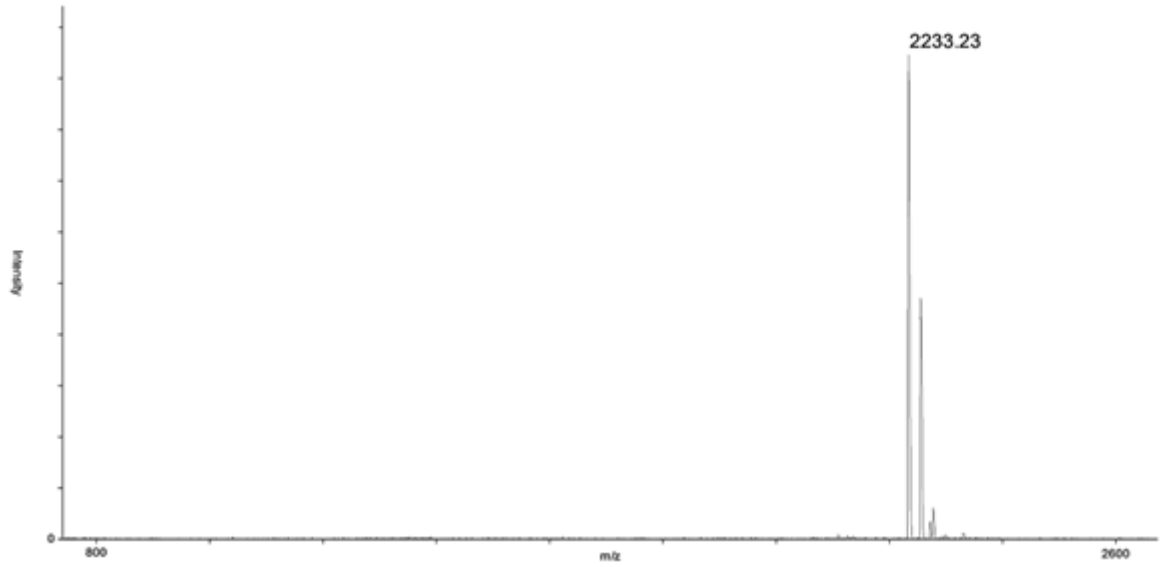


Figure S11. MALDI-Tof mass spectrum of peptide thioester **6**.

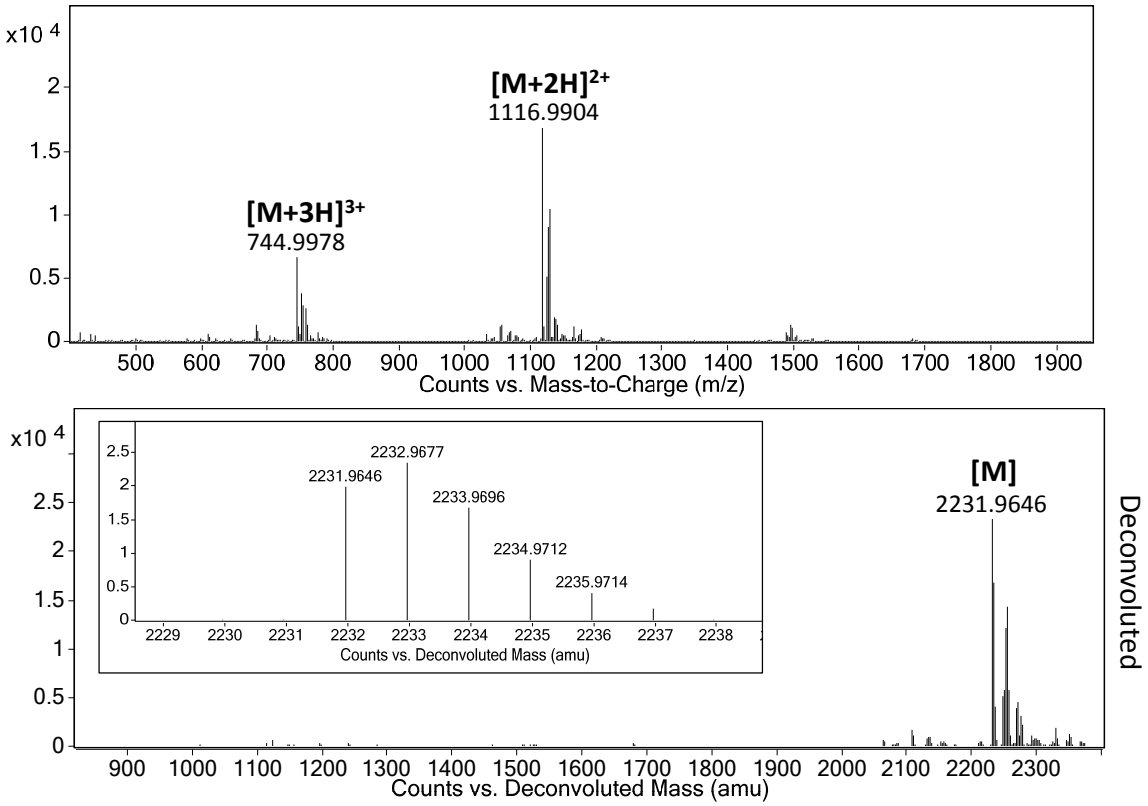
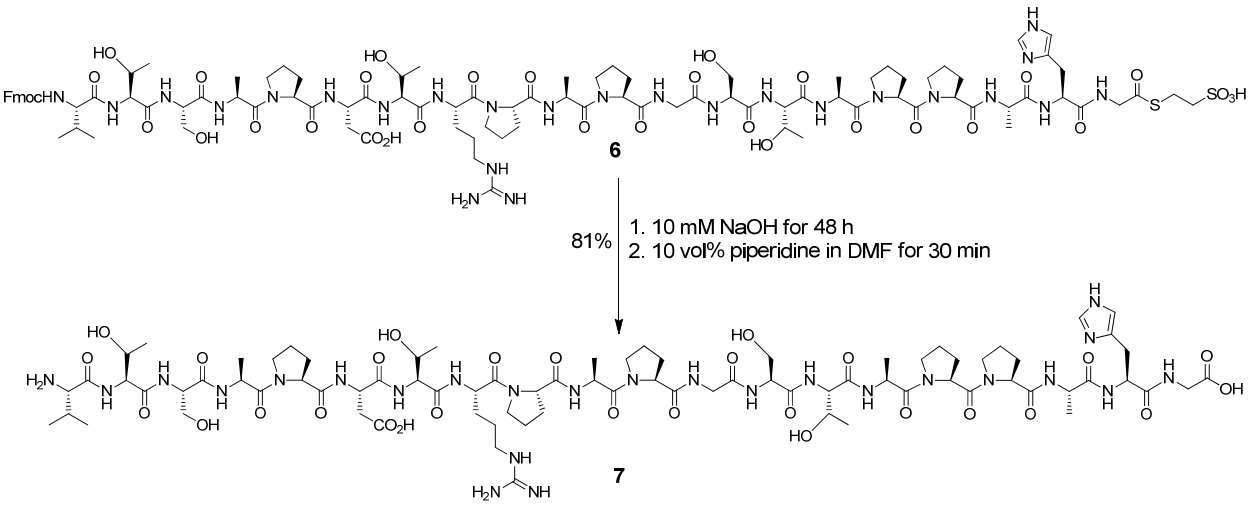


Figure S12. ESI-MS spectrum of peptide thioester **6.**

Peptide 7: H-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-OH (7)

Peptide thioester **6** (4 mg, 1.79 μmol) was treated with 10 mM NaOH (1 mL) and stirred at rt for 48 h. The reaction mixture was lyophilised and treated with 10 vol% piperidine in DMF for 30 min. Subsequent purification by RP-HPLC (*Method F*) and lyophilisation provided the desired peptide **7** as a white solid (2.8 mg, 81%). R_t = 30 min (0 to 30% B over 60 min).



Scheme S4. Synthesis of peptide **7.**

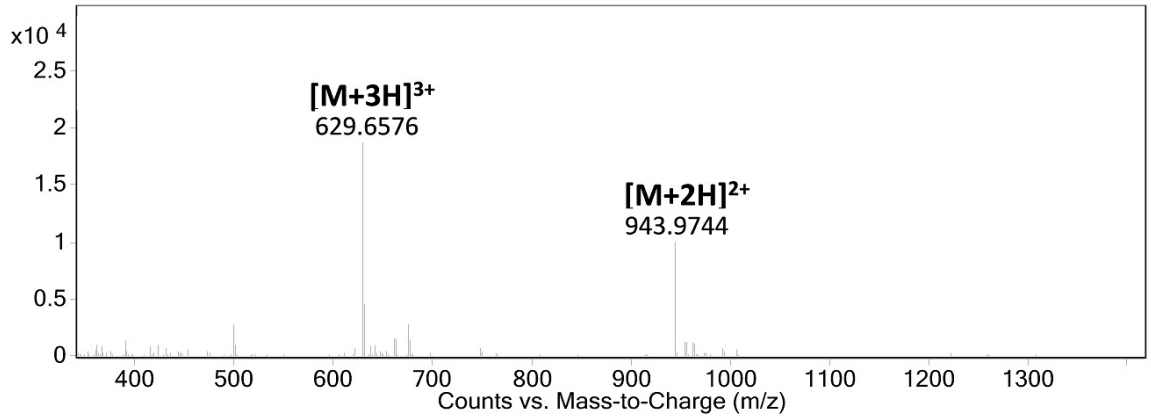
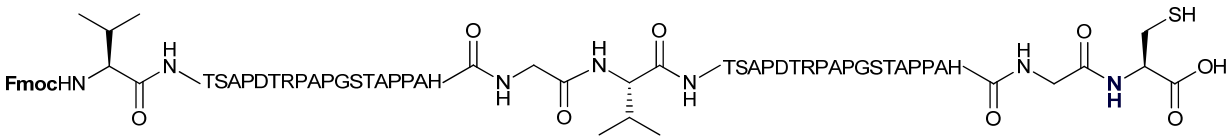


Figure S13. ESI-MS spectrum of peptide **7**.

Peptide 8: Fmoc-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-Cys-OH (8)



8

Peptide **8** was prepared by reacting unprotected peptide **3** (2.2 mg, 1.1 μmol, 1.0 equiv) and peptide thioester **6** (3 mg, 1.3 μmol, 1.2 equiv) *via* the general Ag(I)-assisted ligation procedure (20 h) outlined in the general procedure to give a white solid following purification by preparative RP-HPLC (*Method C*) and lyophilisation (3.5 mg, 77%).

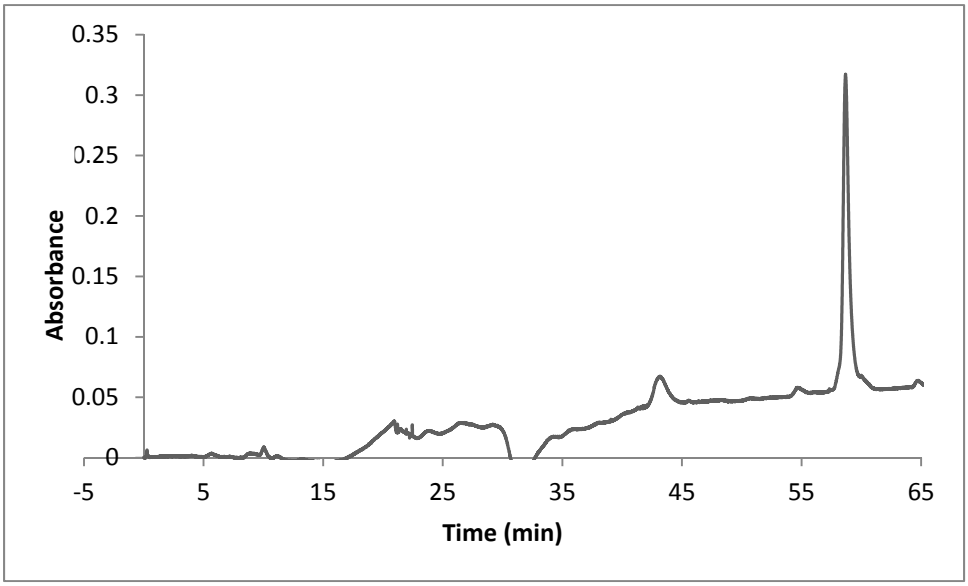


Figure S14. Analytical HPLC chromatogram: R_t = 58 min (0 to 45% B over 60 min).

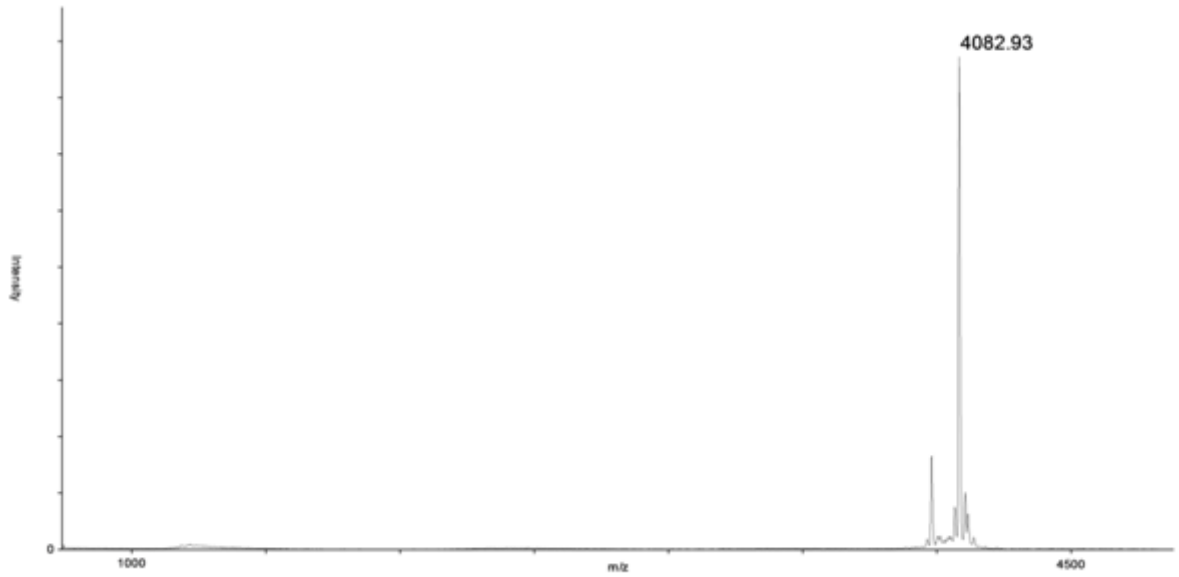


Figure S15. MALDI-Tof mass spectrum of peptide **8**.

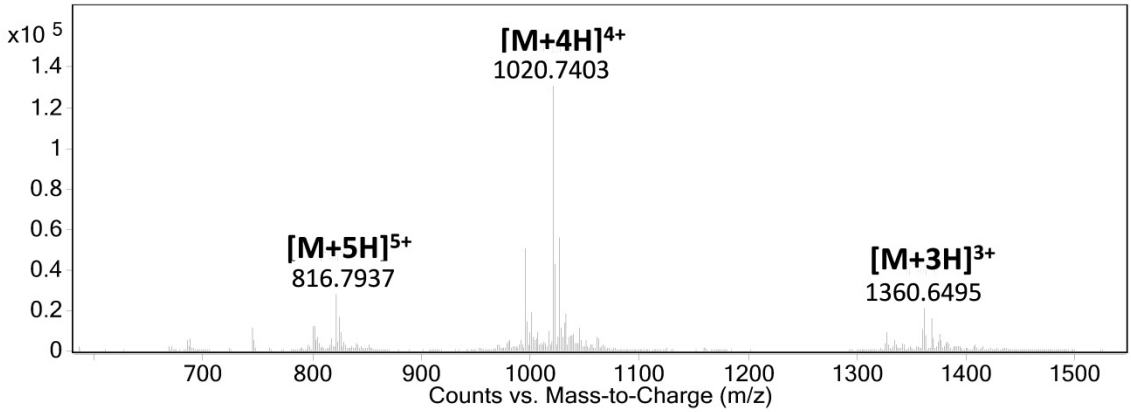
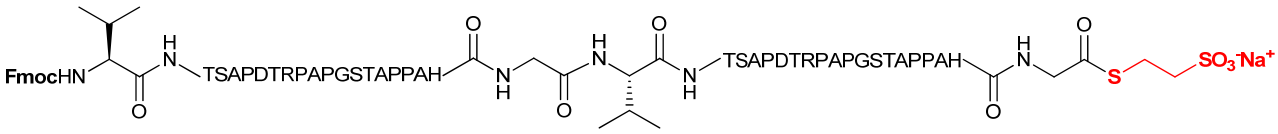


Figure S15. ESI-Q-TOF MS of peptide **8**.

Peptide thioester 10: Fmoc-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-S$(CH_2)_2SO_3H$ (10)



10

Peptide **8** (17.5 mg, 43.3 µmol, 1.0 equiv.) was subjected to the N→S thioesterification conditions outlined in the general procedure to afford peptide thioester **10** as a white solid following purification by RP-HPLC (*Method D*) and lyophilisation (11.6 mg, 34% isolated yield).

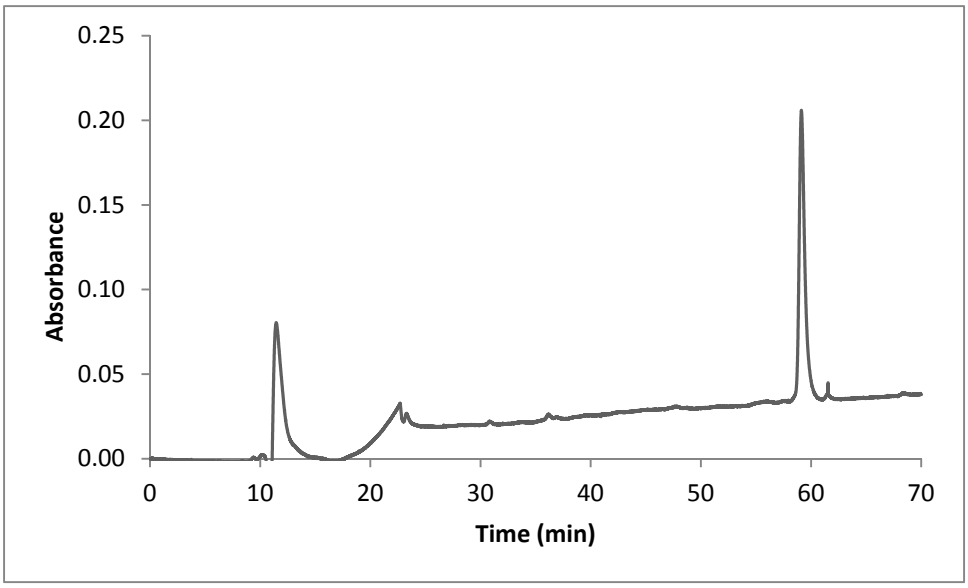


Figure S16. Analytical HPLC chromatogram of peptide thioester **10**: R_t = 60 min (0 to 45% B over 60 min).

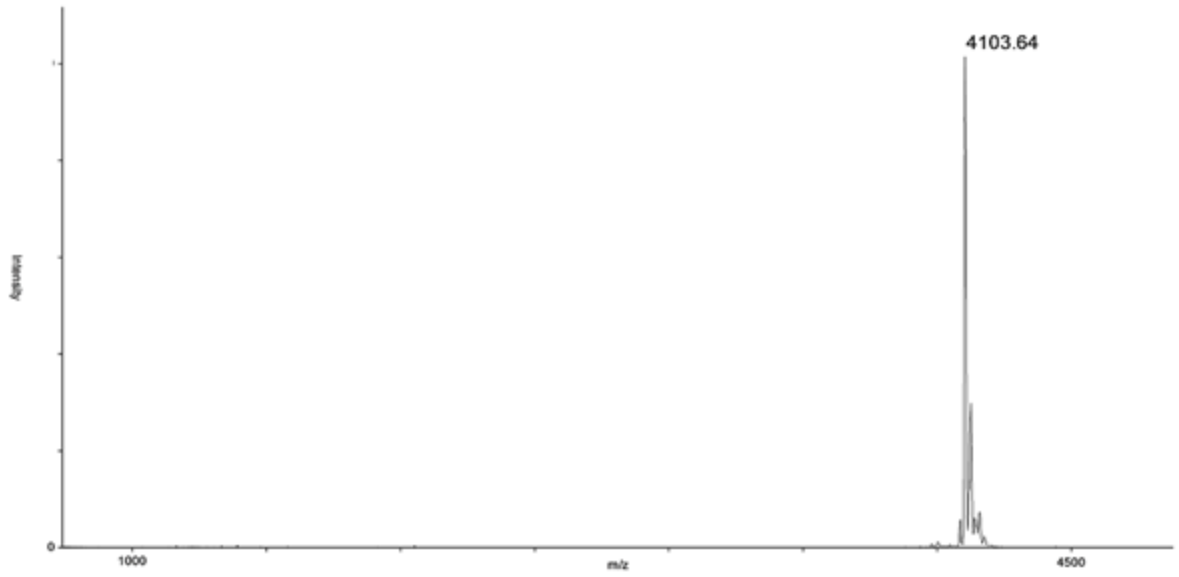


Figure S17. MALDI-Tof mass spectrum of peptide thioester **10**.

Peptide 12: H-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-OH (12)

Peptide **12** was prepared by reacting unprotected peptide **7** (1.9 mg, 1.0 μmol, 1.0 equiv) and peptide thioester **10** (5 mg, 1.2 μmol, 1.2 equiv) *via* the general Ag(I)-assisted ligation procedure outlined in the general procedure (24 h) to give a peptide **11** following purification by preparative RP-HPLC (*Method E*, 3 mg, 0.5 μmol, 50%). The lyophilised product was then treated with 10 vol% piperidine in DMF (1 mL) for 30 min and purified by RP-HPLC (*Method H*) to give **12** as a white solid after lyophilisation (2 mg, 36% over two steps).

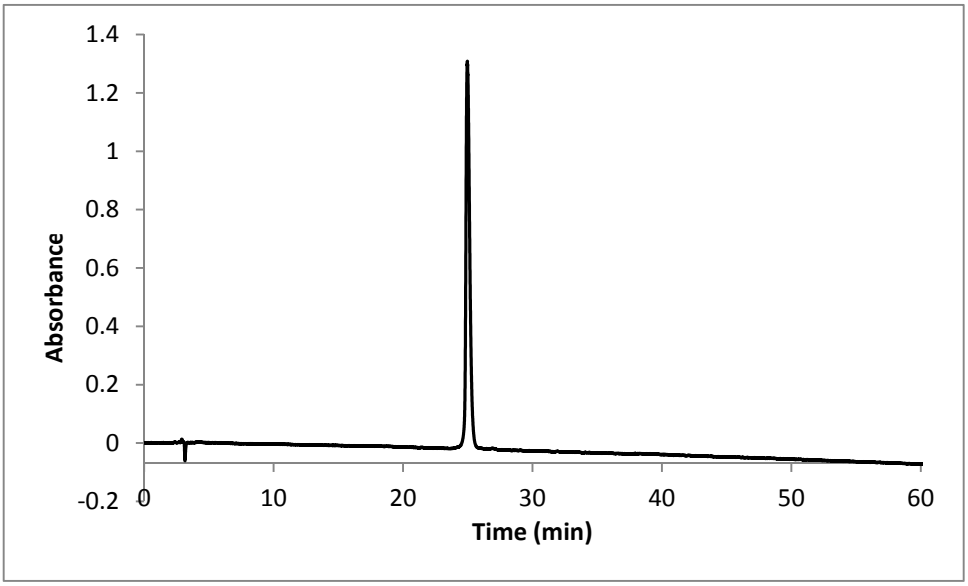


Figure S18. Analytical HPLC trace of peptide **12**: R_t = 25 min (0 to 45% B over 60 min).

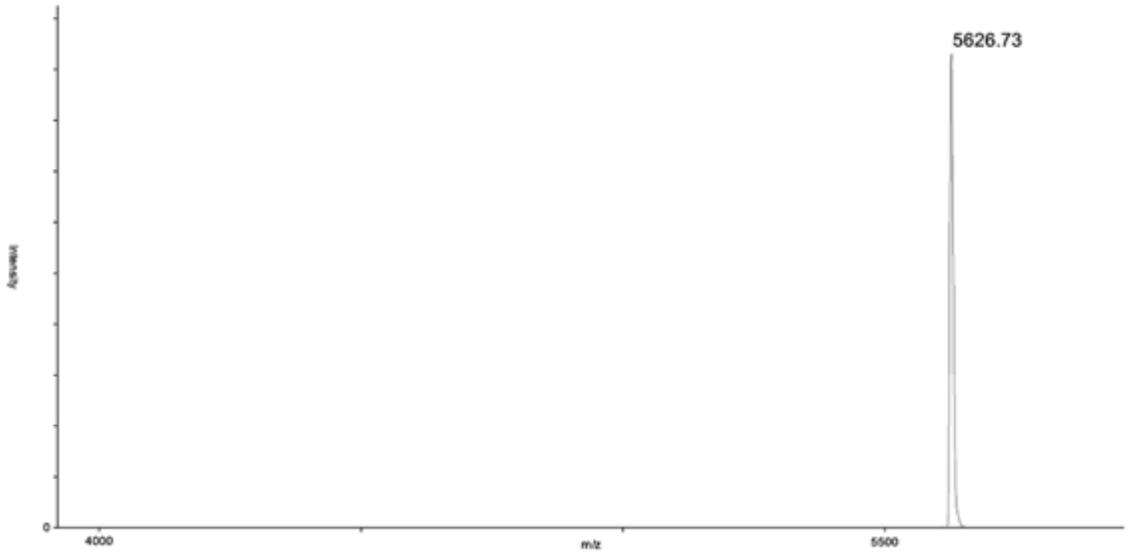


Figure S19. MALDI-Tof mass spectrum of peptide **12**.

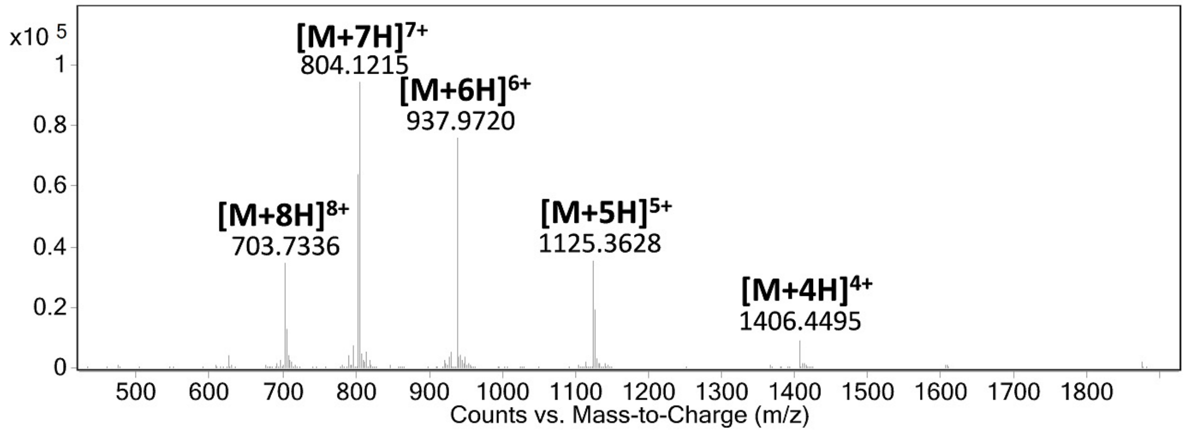
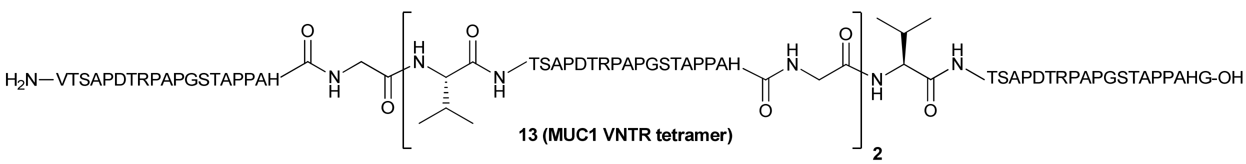


Figure S20. ESI-Q-Tof spectrum of peptide **12.**

Peptide 13: H-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-OH (13)



Peptide **13** was prepared by reacting unprotected peptide **12** (3 mg, 0.53 µmol, 1.0 equiv) and peptide thioester **6** (1.4 mg, 0.64 µmol, 1.2 equiv) *via* the general Ag(I)-assisted ligation procedure outlined in the general procedure (24 h) and purified by preparative RP-HPLC (*Method E*). The lyophilised product was then treated with 10 vol% piperidine in DMF (1 mL) for 30 min and purified by RP-HPLC (*Method E*) to give **13** white solid after lyophilisation (3 mg, 75% over two steps).

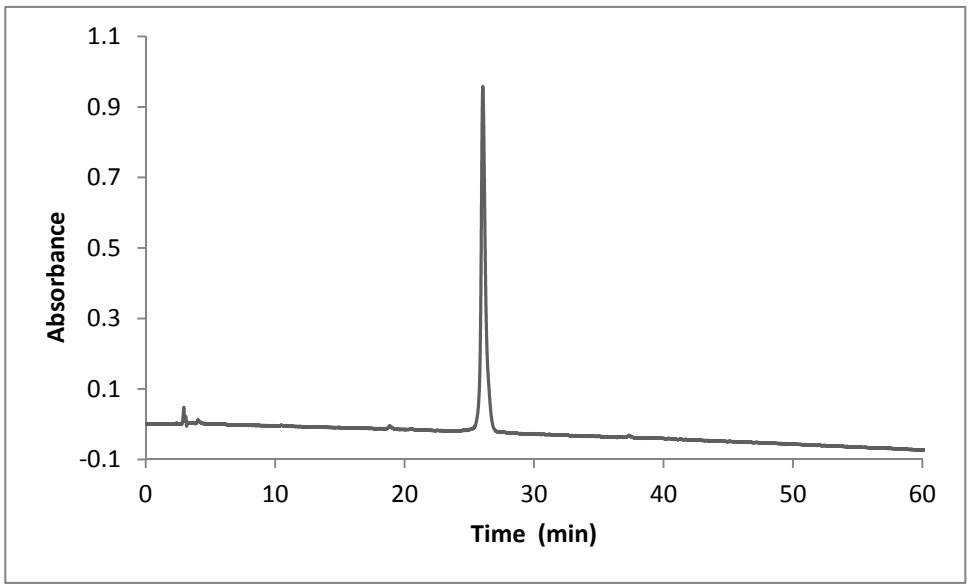


Figure S21. Analytical HPLC trace of peptide **13**: Rt= 26 min (0 to 45% B over 60 min).

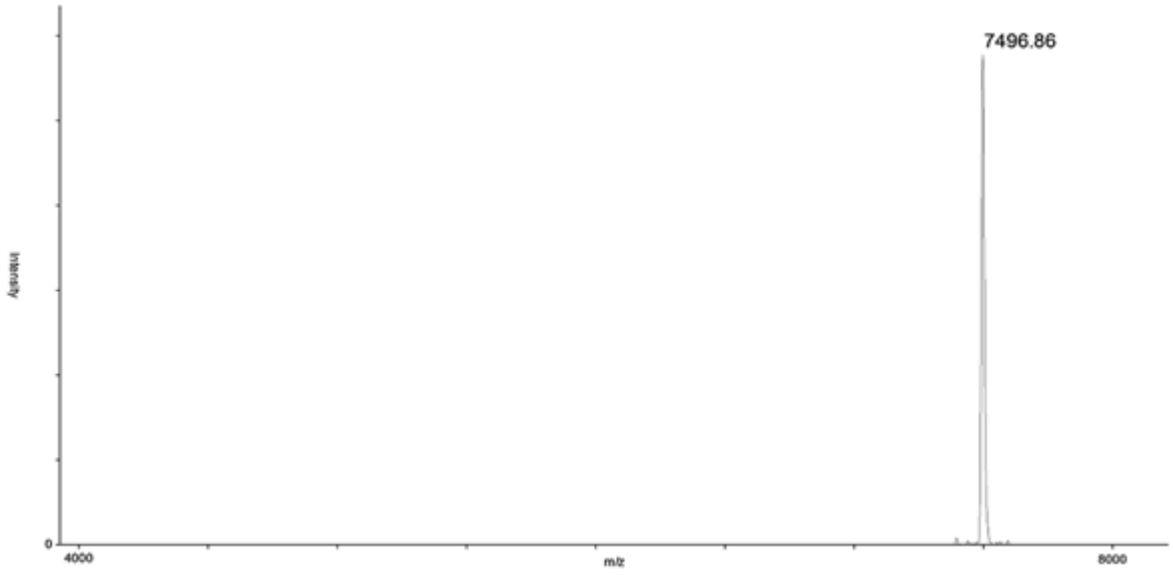


Figure S22. MALDI-Tof mass spectrum of peptide **13**.

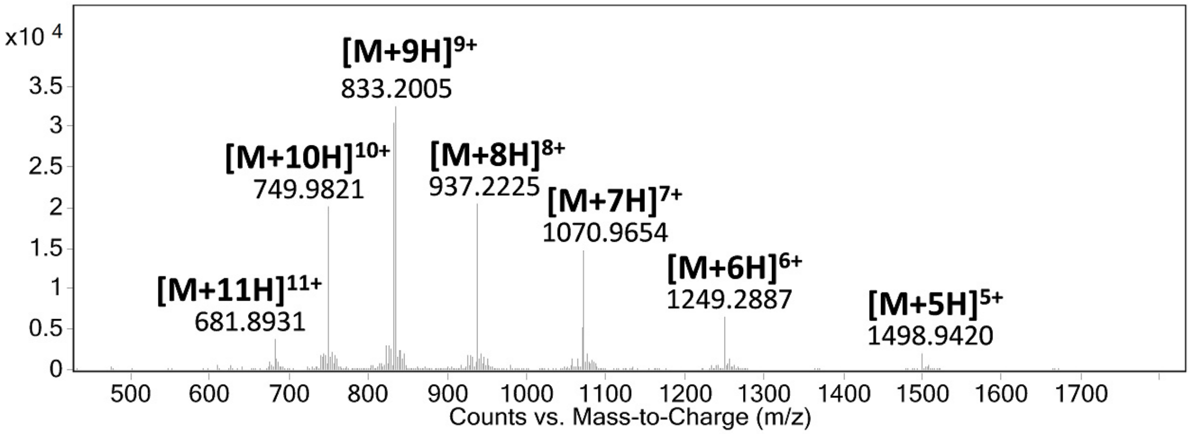
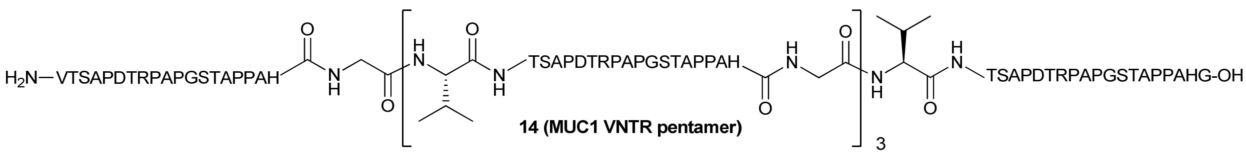


Figure S23. ESI-Q-Tof spectrum of peptide **13.**

Peptide 14: H-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly--Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-OH (14)



Peptide **14** was prepared by reacting unprotected peptide **13** (2.8 mg, 0.37 μmol, 1.0 equiv) and peptide thioester **6** (1 mg, 0.45 μmol, 1.2 equiv) *via* the general Ag(I)-assisted ligation procedure outlined in the general procedure (24 h) and purified by preparative RP-HPLC (*Method E*). The lyophilised product was then treated with 10 vol% piperidine in DMF (1 mL) for 30 min and purified by RP-HPLC (*Method E*) to give **14** white solid after lyophilisation (3.5 mg, 83% over two steps).

.

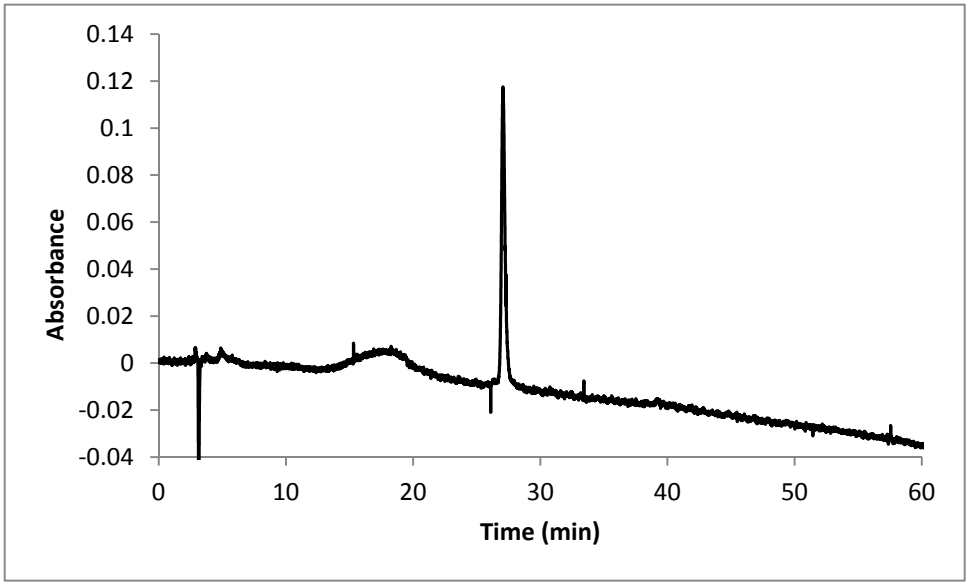


Figure S24. Analytical HPLC trace of peptide **14**: Rt= 28 min (0 to 45% B over 60 min).

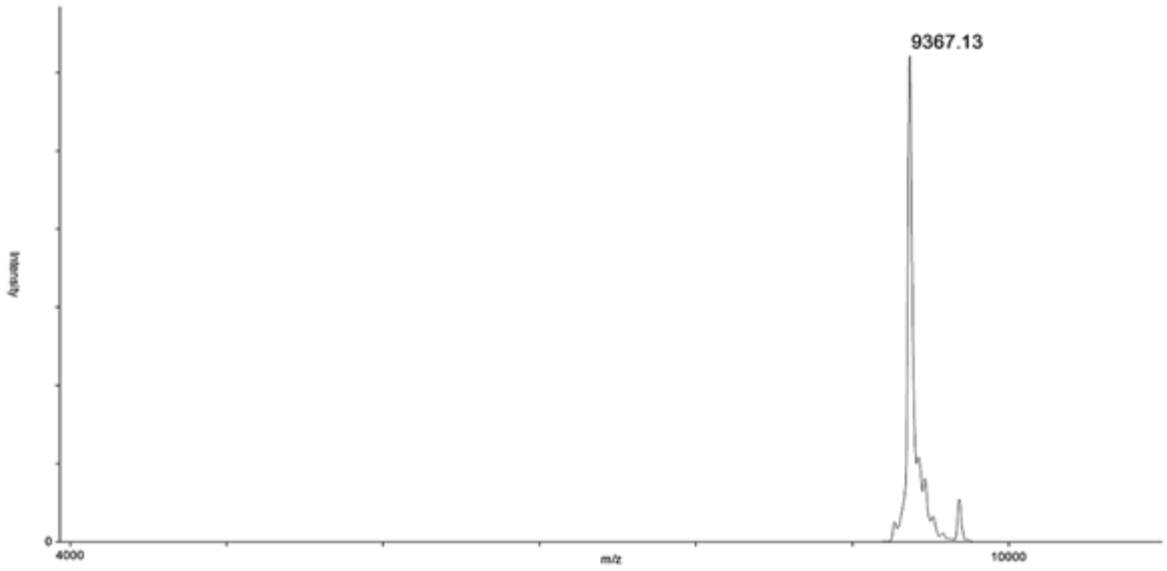


Figure S25. MALDI-Tof mass spectrum of peptide **14.**

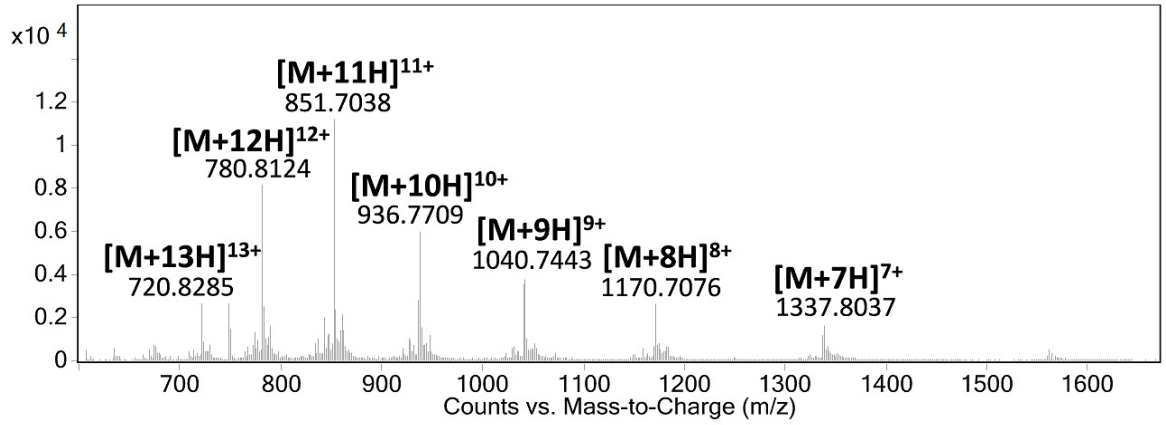
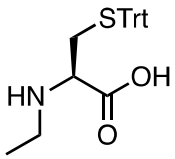


Figure S26. ESI-Q-Tof spectrum of peptide **14.**

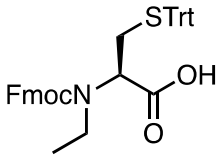
4.0 N→S thioesterification and ligation using *N*-ethyl cysteine- derivatised peptides

4.1 *N*-Ethyl-S-Trityl-L-Cysteine



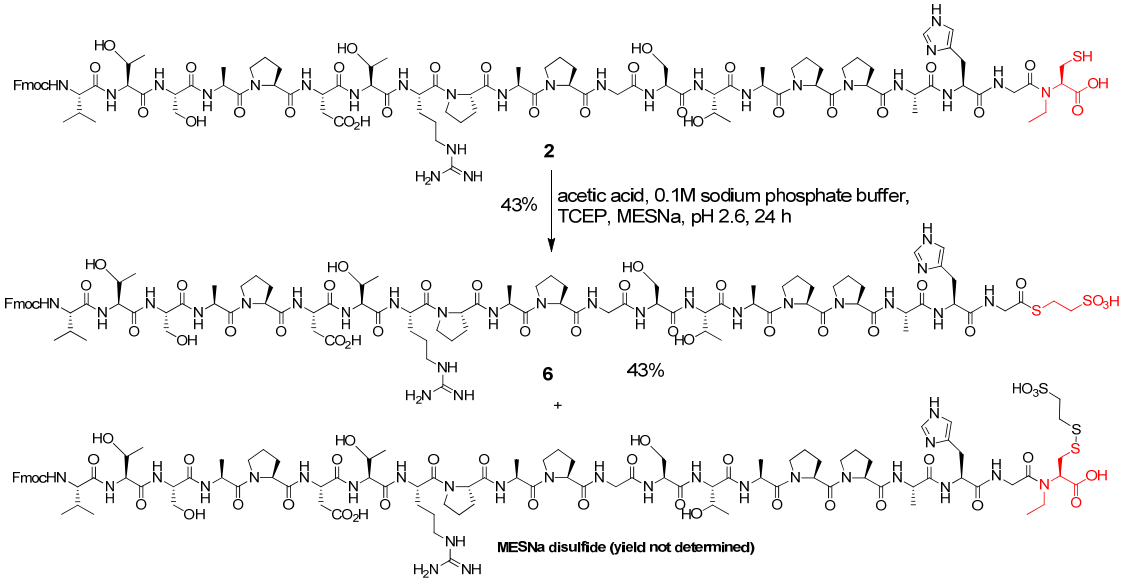
To a cooled (0 ºC) suspension of S-Trityl-L-cysteine (2 g, 3.41 mmol) and $NaBH_3CN$ (0.21 g, 3.41 mmol) in MeOH (6 ml) was added acetaldehyde (289 µl, 5.15 mmol). The resulting solution was allowed to stir at rt for 1 h, after which the RM was concentrated *in vacuo*, and the residue was taken up in 1-butanol. The organic layer was washed with distilled water (2 x 10 ml) and brine (2 x 10 ml). The organic layer was collected, dried and concentrated *in vacuo*. The crude material was purified via flash column chromatography (10% MeOH in CH_2Cl_2) to yield a white solid (1.2g, 80%) $[\alpha]_D$ = +25.3 º (*c* 1.1 in $CHCl_3$), ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.44 – 7.29 (5H, m, Trt), 7.24 – 7.16 (10H, m, Trt), 3.05 (1H, m, α-H), 2.77 (2H, m, β-H), 2.52 (2H, m, Et), 1.12 (3H, m, Et). These data are consistent with those previously reported by Hironobu *et al*[1]

4.2 Fmoc-N-Ethyl-S-Trityl-L-Cysteine

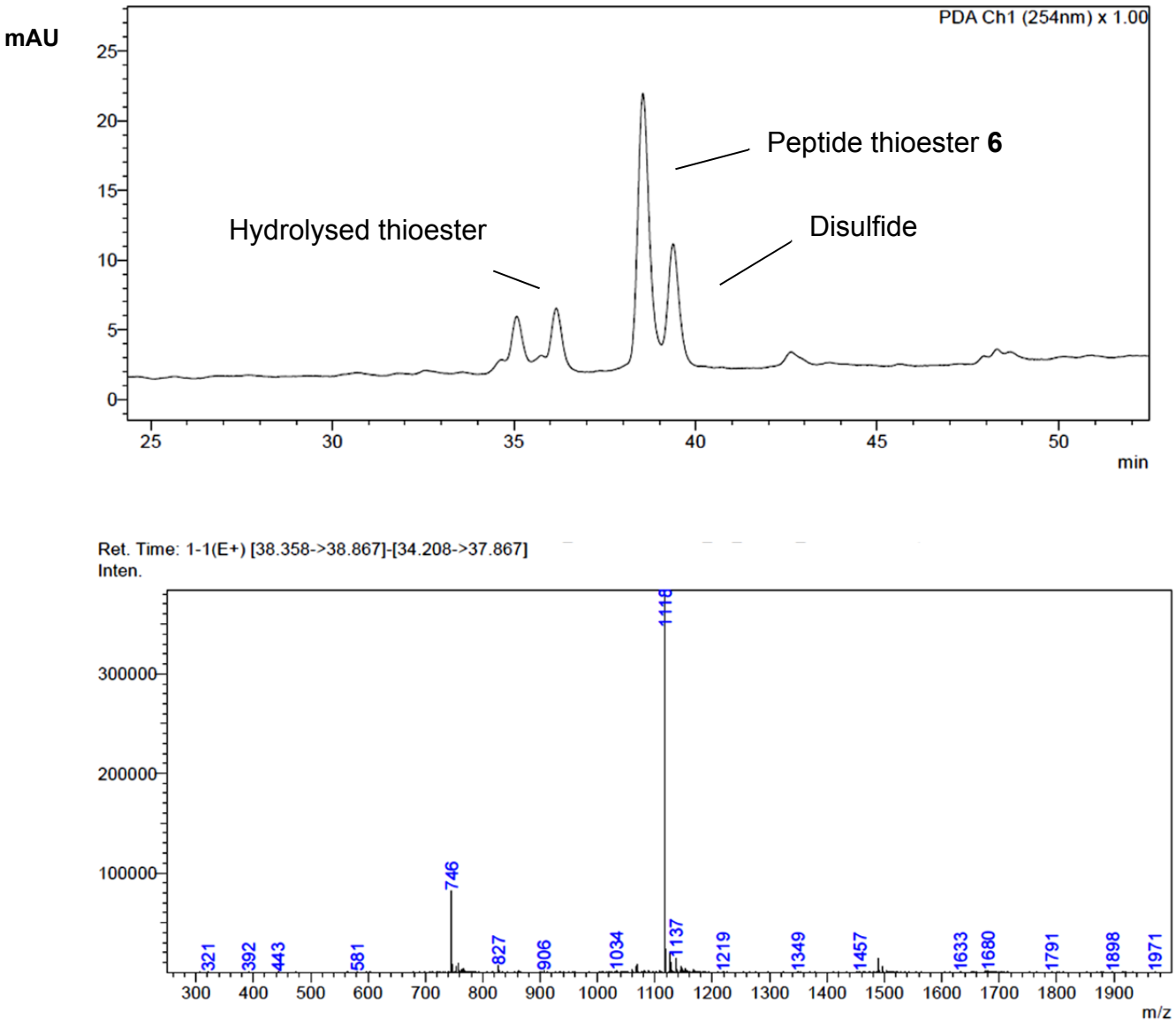


N-Ethyl-S-Trityl-L-Cysteine (436.3 mg, 1.12mmol) was dissolved in 10% $Na(CO_3)_{2(aq)}$ solution (6 ml) and 1,2-DME (3 ml). Fmoc-OSuc (639 mg, 1.79 mmol) dissolved in 1,2-DME (3 ml) was added to the reaction mixture. This was then allowed to stir for 16 h, after which it was filtered and the filtrate was neutralized with 1 M HCl. This was then extracted with EtOAc (3 x 5 ml), the organic layers were collected, dried and concentrated *in vacuo*. The crude material was purified via flash column chromatography (30vol.% EtOAc in Hexane with 1vol.% AcOH) to afford the title compound as a white solid (480 mg, 70%). $[\alpha]_D$ = -35.7 º (*c* 1.1 in $CHCl_3$), ^1H NMR (300 MHz, $CDCl_3$) δ 7.71 (2H, d, J_{HH} = 9 Hz, Fmoc), 7.52 (2H, d, J_{HH} = 9 Hz, Fmoc), 7.44 – 7.37 (7H, m, Fmoc + Trt), 7.33 – 7.12 (13H, m, Fmoc + Trt), 4.50 – 4.32 (2H, m, Fmoc-CH_2), 4.20 (0.6H, m, Fmoc-CH), 4.10 (0.4H, m, Fmoc-CH), 3.46 – 3.27 (0.4H, m, Et), 3.25 – 3.22 (0.6H, m, Et), 3.14 (0.6H, m, a-H), 3.08 (0.3H, m, a-H), 2.93 – 2.84 (1.2H, m, b-H), 2.75 – 2.63 (1.4H, m, Et + b-H), 2.45 (0.4H, 1H, b-H), 0.93 – 0.85 (3H, m, Et). These data are consistent with those previously reported by Hironobu *et al.*[1]

4.3 Thioesterification by N→S acyl transfer



Scheme S5. Synthesis of peptide thioester **6** by N→S acyl transfer.



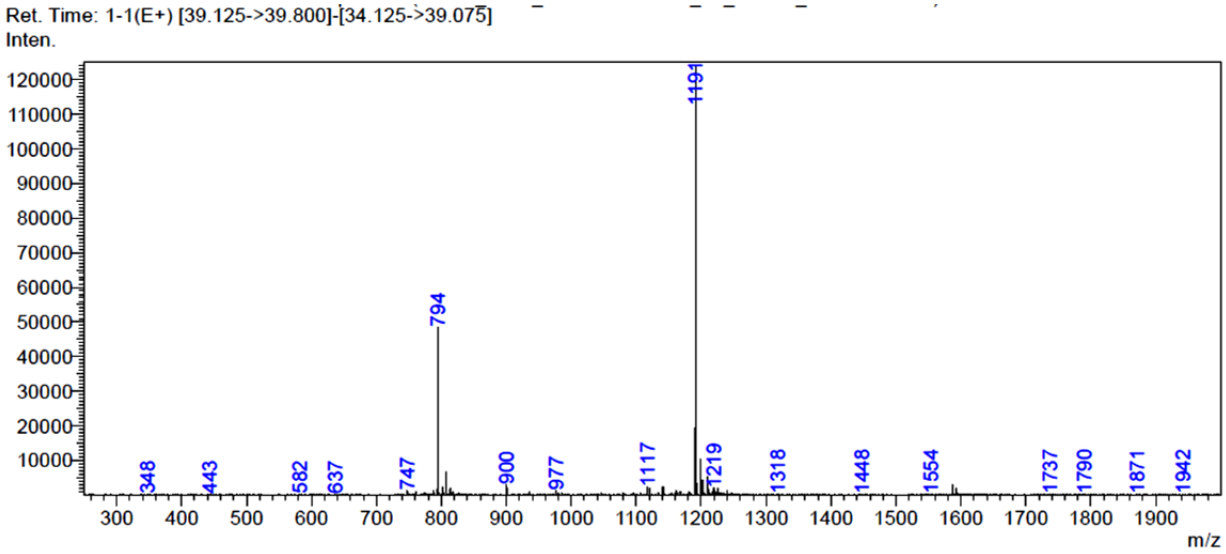


Figure 27. Extracted chromatogram from LC-MS of N→S thioesterification reaction after 4 h (0 to 50B over 60 min) (top), ESI-MS spectra of peptide thioester **6** (middle) and disulfide (bottom).

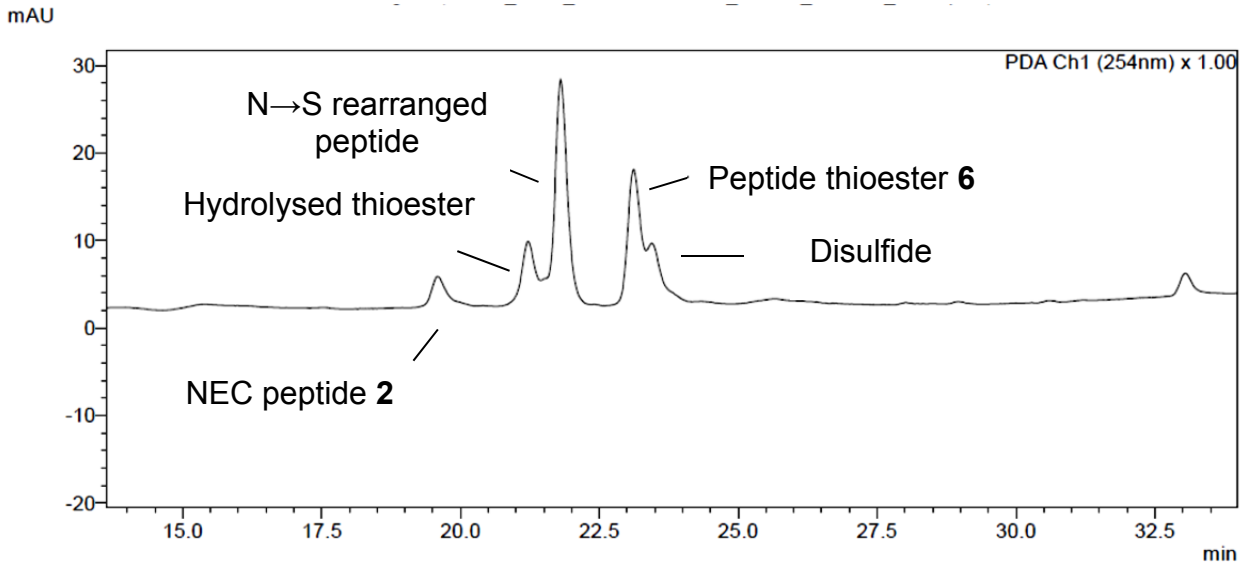
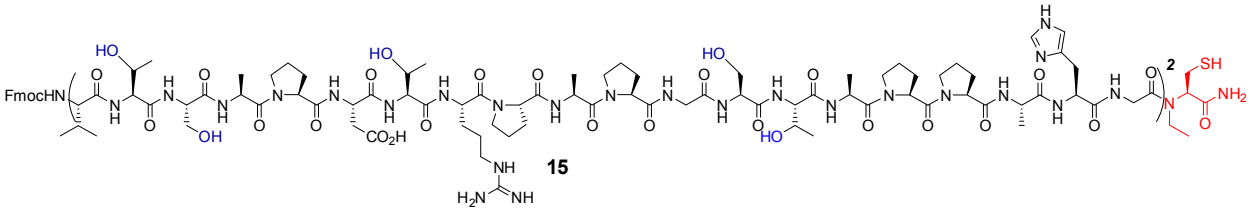


Figure 28. LC-MS of N→S thioesterification reaction after 4 h using 500 equiv MESNa.

Peptide 15: Fmoc-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Cys(*N*-ethyl)-NH_2 (15, NB: this experiment is not described in the main text**)**



Peptide **15** was prepared by reacting unprotected peptide **4** (2.8 mg, 1.4 μmol, 1.0 equiv) and peptide thioester **6** (3.1 mg, 1.3 μmol, 1.2 equiv) *via* the general Ag(I)-assisted ligation procedure outlined in the general procedure to give a white solid following purification by preparative RP-HPLC (*Method C*) and lyophilisation (3.9 mg, 95 μmol, 86%).

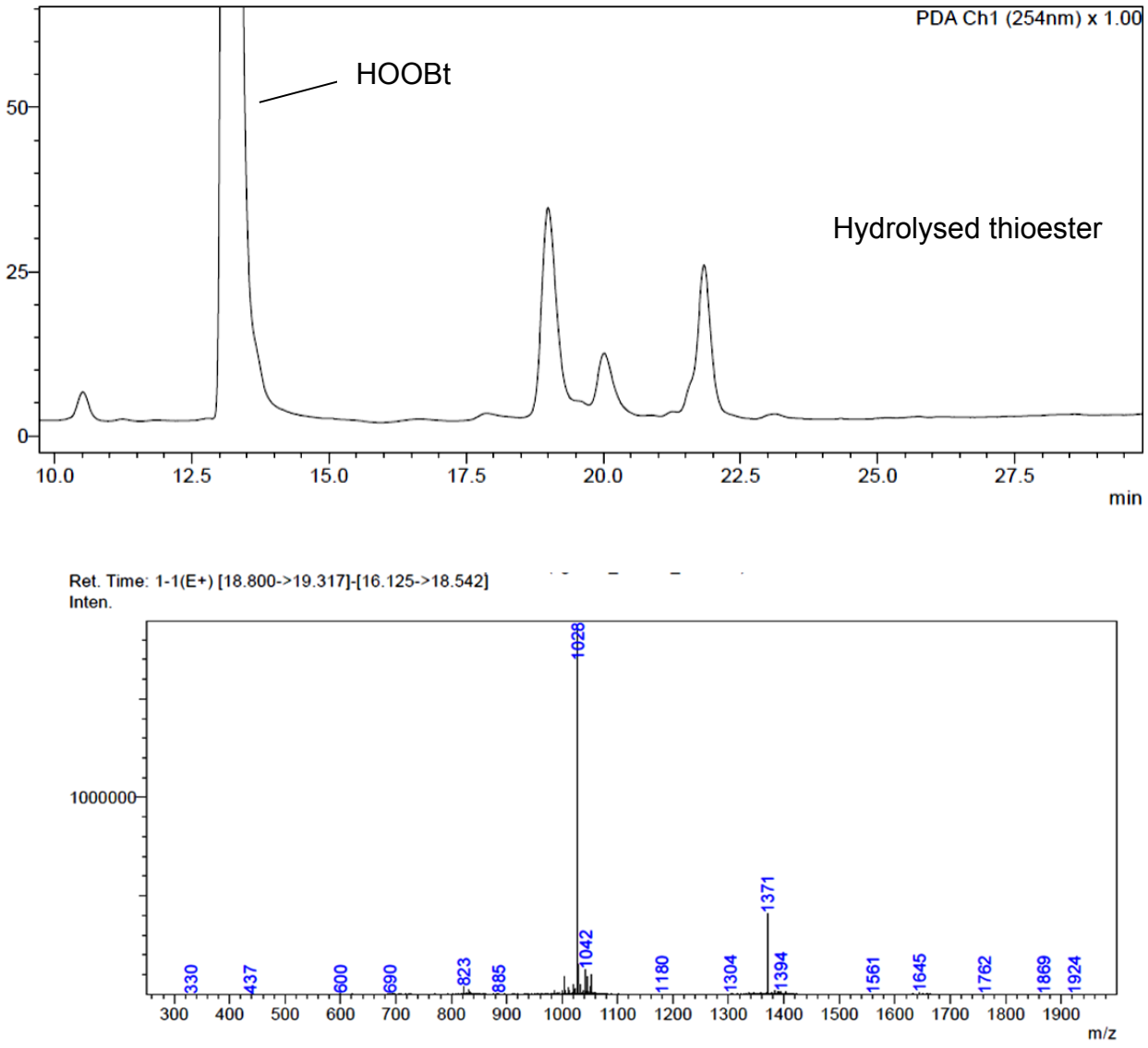
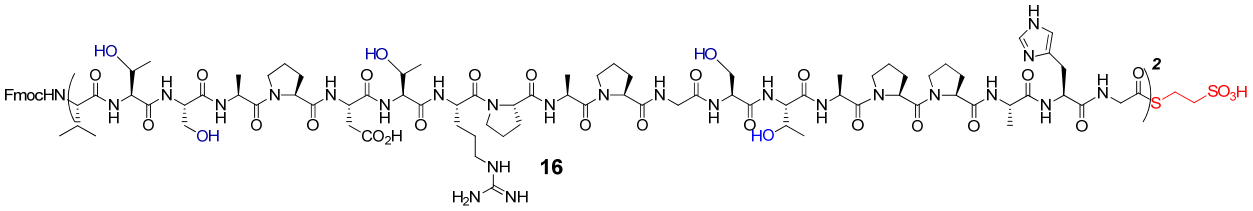
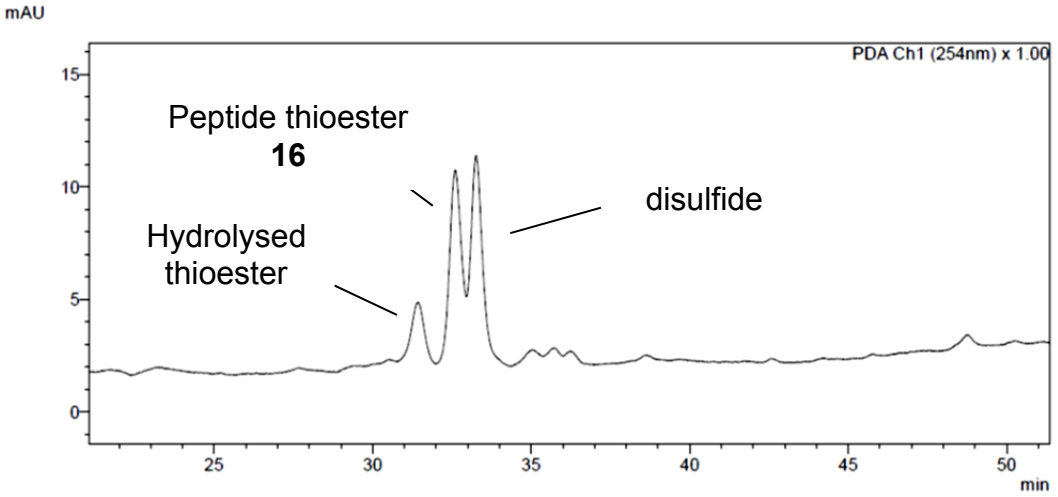


Figure S29. LC-MS of Ag(I)-assisted ligation after 20 h (0 to 50% B over 60 min).

Peptide thioester 16: Fmoc-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-S$(CH_2)_2SO_3H$ (16, NB: this experiment is not described in the main text**)**





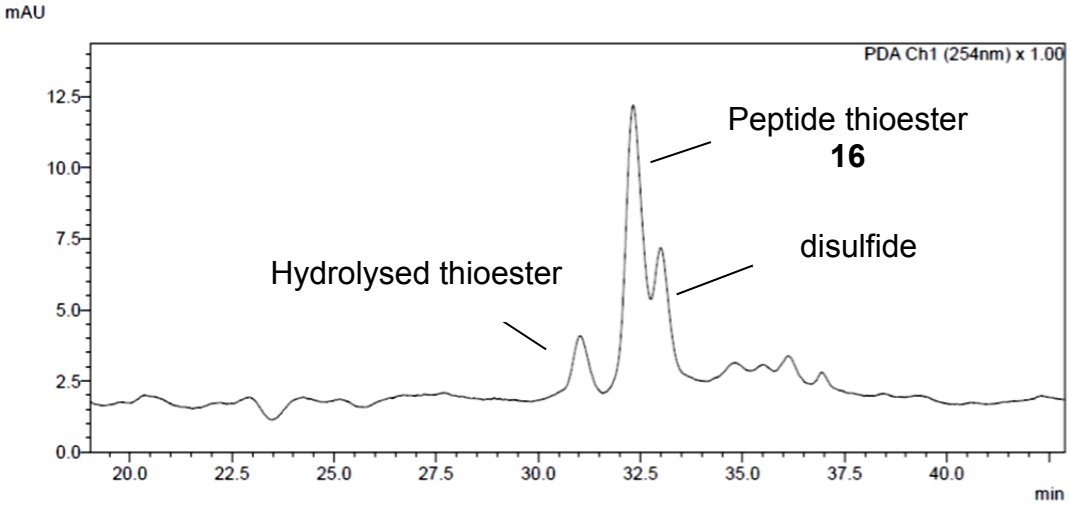
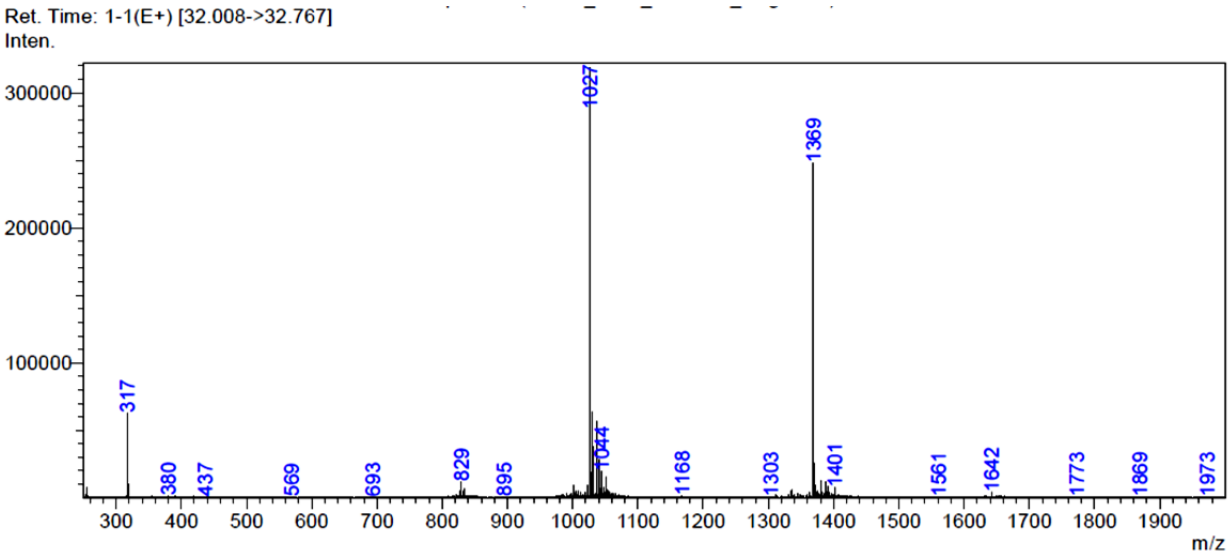


Figure 30. LC-MS of N→S thioesterification reaction after 4 h (0to50B over 60 min) (top) and after 16 h (bottom).



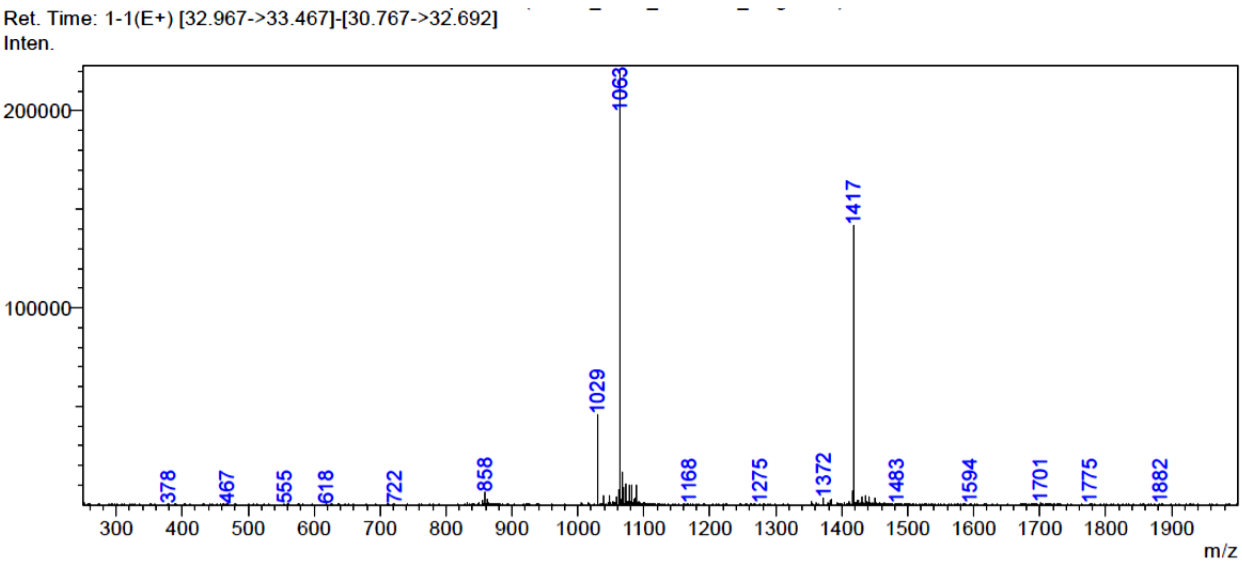


Figure 31. ESI-MS spectra of peptide thioester **16** (top) and disulfide (bottom).

References:

1. H. Hironobu, N. Yoshiaka, WO2008044628 (A1), 2008-4-17

		MALDI-MS			*ESI-MS*		
MUC1 code	**Formula**	**Exact mass from formula [M+H]**	**MALDI-TOF MS $[M+H]^+$**	**Deviation (ppm)**	**Exact mass from formula [M]**	**ESI-Q-TOF MS [M]**	**Deviation (ppm)**
MUC1 VNTR (**7**)	$C_{80}H_{127}N_{25}O_{28}$	1886.94[1]	1886.81[1,3]	69	1885.9282[1]	1885.9378[1,5]	5.0
(**1**)	$C_{98}H_{142}N_{26}O_{31}S$	2212.01[1]	2211.94[1,3]	32	2211.0055[1]	2211.0094[1,5]	1.7
MUC1 VNTR thioester (**6**)	$C_{97}H_{141}N_{25}O_{32}S_2$	2232.97[1]	2233.23[1,3]	116	2231.9616[1]	2231.9646[1,5]	1.3
(**8**)	$C_{178}H_{267}N_{51}O_{58}S$	4082.46[2]	4082.93[2,3]	115	4078.9232[1]	4078.9310[1,5]	1.9
(**10**)	$C_{177}H_{266}N_{50}O_{59}S_2$	4103.50[2]	4103.64[2,3]	34	4099.8793[1]	4099.8783[1,5]	0.2
MUC1 VNTR trimer (**12**)	$C_{240}H_{377}N_{75}O_{82}$	5626.09[2]	5626.73[2,4]	114	5621.7636[1]	5621.7942[1,5]	5.4
MUC1 VNTR tetramer (**13**)	$C_{320}H_{502}N_{100}O_{109}$	7495.12[2]	7496.86[2,4]	232	7489.6812[1]	7489.7274[1,5]	6.1
MUC1 VNTR pentamer (**14**)	$C_{400}H_{627}N_{125}O_{136}$	9364.14[2]	9367.13[2,4]	320	9357.5989[1]	9357.6455[1,5]	4.9

1) Monoisotopic mass, 2) Average mass, 3) Reflectron mode, 4) Linear mode, 5) Direct infusion