Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

69451 Weinheim, Germany

Supporting Information

Cysteine-Free Peptide and Glycopeptide Ligations by Direct Aminolysis

*Richard J. Payne,[1] Simon Ficht,[1] William A. Greenberg[1] and Chi-Huey Wong[1,2]**

[1]Department of Chemistry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA. [2]The Genomics Research Center, Academia Sinica, 128 Section 2, Academia Road, Nankang, Taipei 115, Taiwan.

wong@scripps.edu

Table of contents

General Materials and Methods

Materials

Water was taken from a Milli-Q ultra pure water purification system (Millipore corp.). DMF was purchased as biotech grade. Commercial reagents were purchased from Sigma-Aldrich or Acros Organics and were used without further purification. Anhydrous grade solvents were purchased from Sigma-Aldrich or Acros Organics and were used directly. Resins, protected amino acids and PyBOP were purchased from Novabiochem. Deuterated solvents were purchased from Cambridge Isotope Laboratories Inc.

Mass spectrometry

MALDI-TOF mass spectra were measured on a Voyager-DE Pro biospectrometry workstation by PerSeptive Biosystems. A solution of 10 mg/ml α-cyano-4-hydroxy cinnamic acid containing 1:1 v/v acetonitrile + 0.1 % TFA: water + 0.1 % TFA was used for generating the probe-matrix mixture.

High resolution mass spectra were measured on an Agilent 6210 Time of Flight Mass Spectrometer.

HPLC

Analytical HPLC was run on a Hitachi (D-7000 HPLC system) instrument using an analytical column (Grace Vydac "Protein & Peptide C18", 150 x 4.6 mm, 10 μm particle size, flow rate 1.5 ml/min, 50 ^{o}C). Semi preparative HPLC was run on a Hitachi (D-7000 HPLC system) instrument using a semi preparative column (Grace Vydac "Protein & Peptide C18", 250 x 10 mm, 10-15 μm particle size, flow rate 4 mL/min). Preparative HPLC was run on a Hitachi (D-7000 HPLC system) instrument using a preparative Column (Grace Vydac "Protein & Peptide C18", 250 x 22 mm, 10-15 μm particle size, flow rate 8 mL/min). Detection of the signal was achieved with either photodiode array or UV at a wavelength, $\lambda = 280$ nm (detection of Tyr). For the purification and analytical traces of MUC1

peptides and glycopeptides, detection was performed at λ = 230 nm. Eluents A (0.1 % TFA in water) and B (0.1 % TFA in acetonitrile) were used in a linear gradient. Gradient A: 0 % B → 50 % B in 30 min; Gradient B: 0 % B → 35 % B in 30 min; Gradient C: 0 % B → 80 % B in 30 min.

Synthesis of Glycosyl Amino Acids.

Synthesis of Fmoc-Ser(Ac_3AcNH-β-Glc)-OH and Fmoc-Thr(Ac_3AcNH-α-Gal)-OH was carried out as previously described.[1, 2]

General Procedures for SPPS

General procedure for SPPS of peptides and glycopeptides following the Fmoc strategy. Solid-phase peptide synthesis was carried out in syringes, equipped with teflon filters, purchased from Torviq. Rink amide resin was initially washed (5x DCM, 5x DMF), followed by removal of the Fmoc group by treatment with 10% piperidine/DMF (2x 5 min) followed by a further washing step (5x DMF, 5x DCM, 5x DMF). For pre-activation of the first amino acid, 4 eq. of PyBOP and 8 eq. of NMM were added to a solution of 4 eq. protected amino acid (0.1 M) in DMF. After 5 min of pre-activation, the mixture was added to the resin. After 2 h the resin was washed (5x DMF, 5x DCM, 5x DMF), capped with acetic anhydride/pyridine (1:9) (2x 5 min) and washed (5x DMF, 5x DCM, 5x DMF).

Iterative peptide assembly: *Deprotection:* The resin was treated with 10% piperidine/DMF (2x 5 min) and subsequently washed (5x DMF, 5x DCM, 5x DMF). *Amino acid coupling:* A preactivated solution of 4 eq. protected amino acid (final concentration 0.1 M in DMF) using 4 eq. PyBOP and 8 eq. NMM was added to the resin. After 45 min, the resin was washed with DMF (5x), DCM (5x) and DMF (5x). *Capping*: Acetic anhydride/pyridine (1:9) was added to the resin. After 5 min the resin was washed with DMF (5x), DCM (5x) and DMF (5x). *Coupling of the glycosyl amino acid building*

blocks was carried out by adding a preactivated solution of 1 eq. of the glycosylated amino acid (final concentration 0.1 M in DMF) using 1 eq. PyBOP and 2 eq. NMM with respect to the loaded resin. After 12 h, the resin was washed with DMF (5x), DCM (5x) and DMF (5x). *Acetate deprotection:* The resin was washed with DCM (5x), MeOH (10x) before treating with 6:1 v/v methanol/hydrazine hydrate for 6 h to remove the acetate groups on the sugar. The resin was finally washed with MeOH (5x) and DCM (10x). *Cleavage*: A mixture of TFA: thioanisole: triisopropylsilane: water (17:1:1:1 v:v:v:v) was added. After 2 h, the resin was washed with TFA (4x 4 mL) *Work-up*: The combined solutions were concentrated in vacuo. The residue was dissolved in water containing 30% MeCN + 0.1% TFA and purified by preparative HPLC (gradient A) and analyzed by MALDI-TOF/MS (matrix: α-Cyano-4-hydroxycinnamic acid).

General procedure for the Boc synthesis of peptide thioesters

Preloading of 3-(tritylthio) propanoic acid onto the MBHA-linker: Resin loadings were aimed at approximately 300 μmol/g by adding the resin in excess. First, the resin was washed (5x DCM, 3 min 5% DIPEA/DCM, 5x DCM, 5x DMF). For preactivation of the 3-(tritylthio) propanoic acid, PyBOP (1 eq.) was added to a 0.1 M solution of the 3-(tritylthio) propanoic acid in DMF containing 2 eq. NMM. After 5 min of preactivation, the mixture was added to the resin. After 2 h the resin was washed (5x DMF, 5x DCM, 3 min 5% DIPEA/DCM, 5x DCM, 5x DMF). For capping the resin was treated with acetic anhydride/pyridine (1:9) (2x 10 min), washed (5x DMF, 10x DCM) and the resin dried in vacuo.

Solid-phase synthesis according to Boc-strategy: *Trt cleavage*: After treatment with TFA: thioanisole: triisopropylsilane: water (17:1:1:1 v:v:v:v; 2x 4 min), the resin was washed with DCM (8x) and DMF (5x). *Boc cleavage*: After treatment with 5% *m*-Cresol/TFA (2x 4 min) the resin was washed with DCM (8x) and DMF (5x). *Coupling*: After preactivation of 4 eq. protected amino acid (final concentration 0.1 M in DMF) for 5 min using 4 eq. PyBOP and 8 eq. NMM, the solution was added to

the resin. After 30 min, the resin was washed with DMF (5x), DCM (5x) and DMF (5x). *Capping*: Acetic anhydride/pyridine (1:9) was added to the resin. After 5 min the resin was washed with DMF (5x) and DCM (5x). *Terminal capping*: Acetic anhydride/pyridine (1:9) was added to the resin. After 10 min the resin was washed with DMF (5x) and DCM (8x). *Cleavage*: A mixture of TFMSA:TFA:thioanisole (2:8:1 v:v:v) was added to the resin. After 2 h, the resin was washed with TFA (4x) *Work-up*: The combined solutions were concentrated in vacuo. The residue was dissolved in water, purified by preparative HPLC and analyzed by MALDI-TOF/MS (matrix: α-Cyano-4-hydroxycinnamic acid).

Thioester Hydrolysis Studies

Selection of an Appropriate Buffer System

A number of buffers were assessed for their ability to prevent hydrolysis of the thioester component in the ligation reactions. Buffers were selected that possessed buffering capacities at a pH range of 7.0-8.5 (so that they would be able to facilitate the direct aminolysis reaction).

Buffers analyzed were:

Potassium phosphate (buffering pH range = 6.0-13.0)

Tris (buffering pH range = 7.0-9.0)

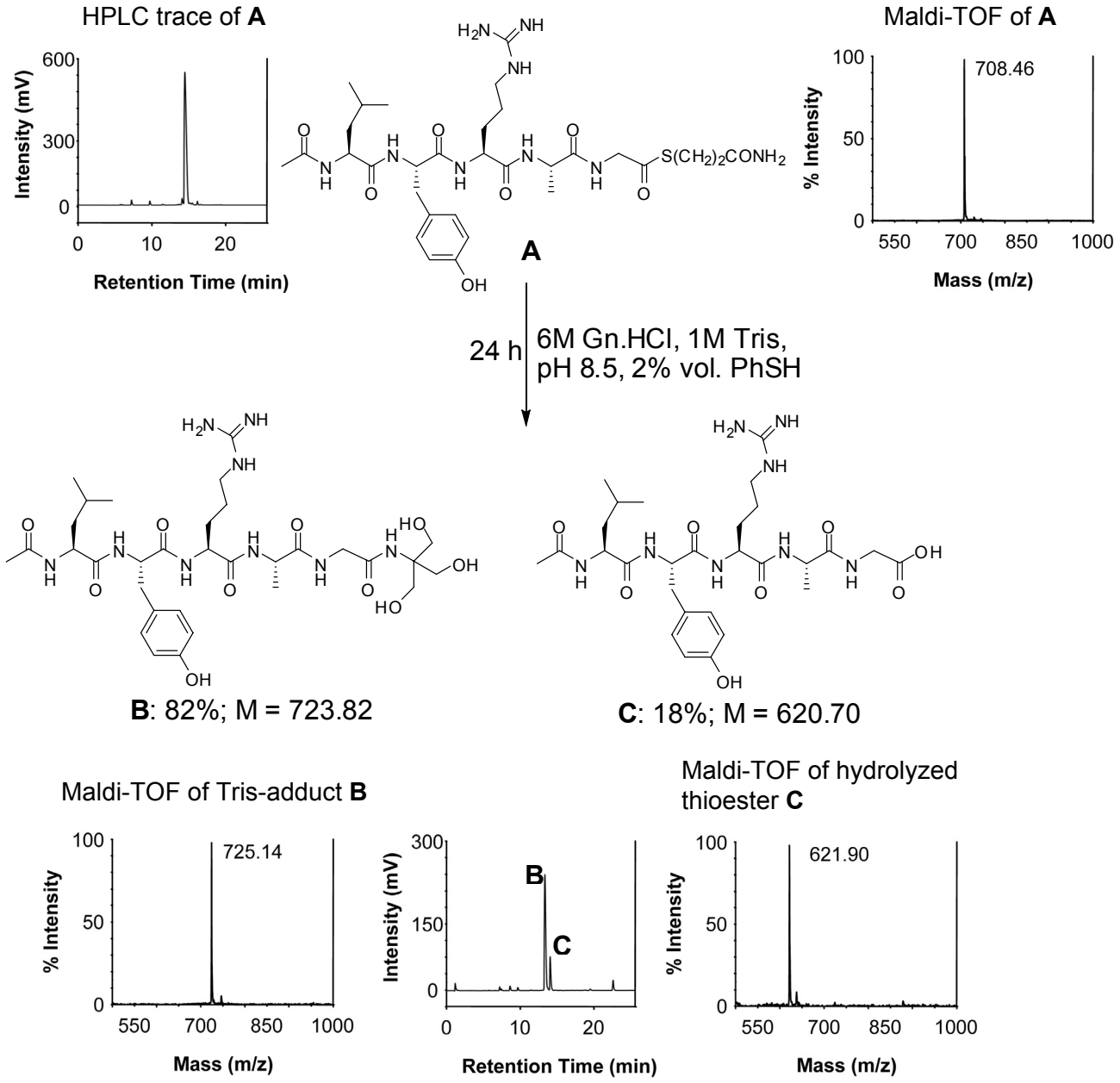
Imidazole (buffering pH range = 6.2-7.8)

HEPES (buffering pH range = 6.8-8.2)

Tris Aminolysis/Hydrolysis Studies

We were concerned that the primary amine of Tris buffer could carry out direct aminolysis of the peptide thioester, thereby resulting in unacceptable levels of by-product in the reaction mixture. To determine whether this was indeed the case, peptide thioester Ac-LYRAG-S$(CH_2)_2CONH_2$ (**A**) was

incubated in 1 M Tris buffer at pH 8.5. After 24 h, the peptide thioester had been converted to the product where Tris had carried out a direct aminolysis reaction on the peptide thioester (**B**, 82%) and hydrolyzed thioester (**C**, 18%) (Scheme S1). As a result of this study, Tris was eliminated as a candidate buffer.



Scheme S1. Aminolysis and hydrolysis of peptide thioester in the presence of Tris buffer (pH 8.5)

Thioester Hydrolysis Studies using potassium phosphate, HEPES and imidazole

Ac-LYRAA-S$(CH_2)_2CO_2NH_2$ (0.5 mg) was dissolved in a variety of different buffers (50 µL) and incubated at 37 ^oC. Aliquots of 10 µL were quenched with 0.1% TFA in water (90 µL) and analyzed by analytical HPLC (Gradient A) every 12 h for 36 h. An endpoint was taken after 72 h. Percentage hydrolysis was analyzed by integration of the peaks corresponding to the peptide thioester Ac-LYRAA-S$(CH_2)_2CO_2NH_2$ and the hydrolyzed product Ac-LYRAA-OH (Figure S1).

NB: Potassium phosphate buffer proved to be insoluble in the cosolvent (NMP) therefore abolishing its ability to buffer the solution on addition of the peptide thioester. Therefore hydrolysis of Ac-LYRAA-S$(CH_2)_2CO_2NH_2$ in 1:1 v/v NMP:6 M Gn.HCl, 400 mM potassium phosphate pH 8.5 and 4:1 v/v NMP:6 M Gn.HCl, 1 M potassium phosphate pH 8.5 were not monitored in these studies. Additionally, since imidazole was incapable of buffering at pH 8.5, hydrolysis studies of imidazole were conducted at pH 7.5 and compared to HEPES buffer at the same concentration, pH and percentage cosolvent (NMP).

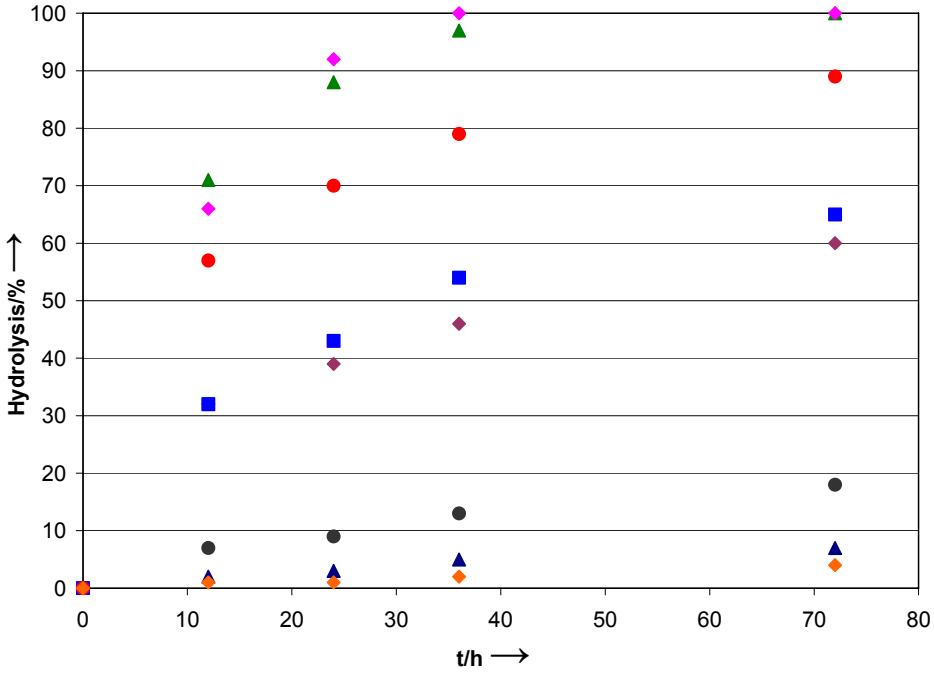


Figure S1. Hydrolysis of Ac-LYRAA-S$(CH_2)_2CO_2NH_2$ in a variety of different buffer systems. ♦ 6 M Gn.HCl, 200 mM imidazole pH 7.5; ▲ 6 M Gn.HCl, 200 mM HEPES pH 8.5; ● 6 M Gn.HCl, 200 mM potassium phosphate pH 8.5; ■ 1:1 v/v NMP:6 M Gn.HCl, 400 mM imidazole pH 7.5; ♦ 1:1 v/v NMP:6 M Gn.HCl, 400 mM HEPES pH 8.5; ● 4:1 v/v NMP:6 M Gn.HCl, 1 M HEPES pH 8.5; ▲4:1 v/v NMP:6 M Gn.HCl, 1 M imidazole pH 7.5; ♦ 4:1 v/v NMP:6 M Gn.HCl, 1 M HEPES pH 7.5.

These studies show that both HEPES and imidazole, when mixed with the organic cosolvent NMP, can drastically suppress hydrolysis of the thioester over a period of 72 h. Due to the slower rate of hydrolysis and more suitable buffering range accessible by HEPES buffer, the remainder of the experiments were carried using 4:1 v/v NMP:6 M Gn.HCl, 1 M HEPES as the solvent system.

General procedure for cysteine-free aminolysis reactions. Peptides or glycopeptides (1.5 equiv, approx. 3 μmol) were dissolved in 150μL of deoxygenated ligation buffer [4:1 v/v *N*-methyl-2-pyrrolidinone (NMP): 6 M guanidine hydrochloride, 1M HEPES, pH = 8.5][a]. This solution was transferred to an eppendorf tube containing the thioester (approx. 2 μmol). Thiophenol (2% by volume, 3 μL) was added and the reaction mixed gently. The ligation mixture was incubated at 37 °C with gentle mixing every 12 h until the reaction was confirmed to be complete by LC-MS. The ligation reactions were quenched by the addition of TCEP solution (0.6 mL of a 10 mg/mL solution) if the products contained a cysteine residue or 0.1% TFA in water (0.6 mL) if the products were free of cysteine residues. The products were purified by semi-preparative HPLC (Gradient B).

[a] An aqueous buffer containing 6 M Gn.HCl and 1 M HEPES was prepared and adjusted to pH 8.5 using 25% aqueous sodium hydroxide solution. The resulting solution (1 mL) was diluted with NMP (4 mL) to produce the final buffer for use in the direct aminolysis reactions.

Kinetics of Cysteine-Free Ligation in the Presence of Internal Cysteine Residues

Cysteine-free peptide **2** and cysteine-containing peptides **4-9** were reacted with peptide thioester **1** (Ac-LYRAG-S$(CH_2)_2CO_2NH_2$) under the reaction conditions described above. Aliquots of 10 µL were removed from the reaction mixture and quenched with TCEP solution (90 µL of a 10 mg/mL solution). These aliquots were taken after 1 h, and then every 2 h for 12 h. An endpoint aliquot was quenched after 24 h. These samples were analyzed by analytical HPLC (Gradient A) and the percentage conversion calculated by integrating the peaks corresponding to starting peptide and the ligated peptide product.

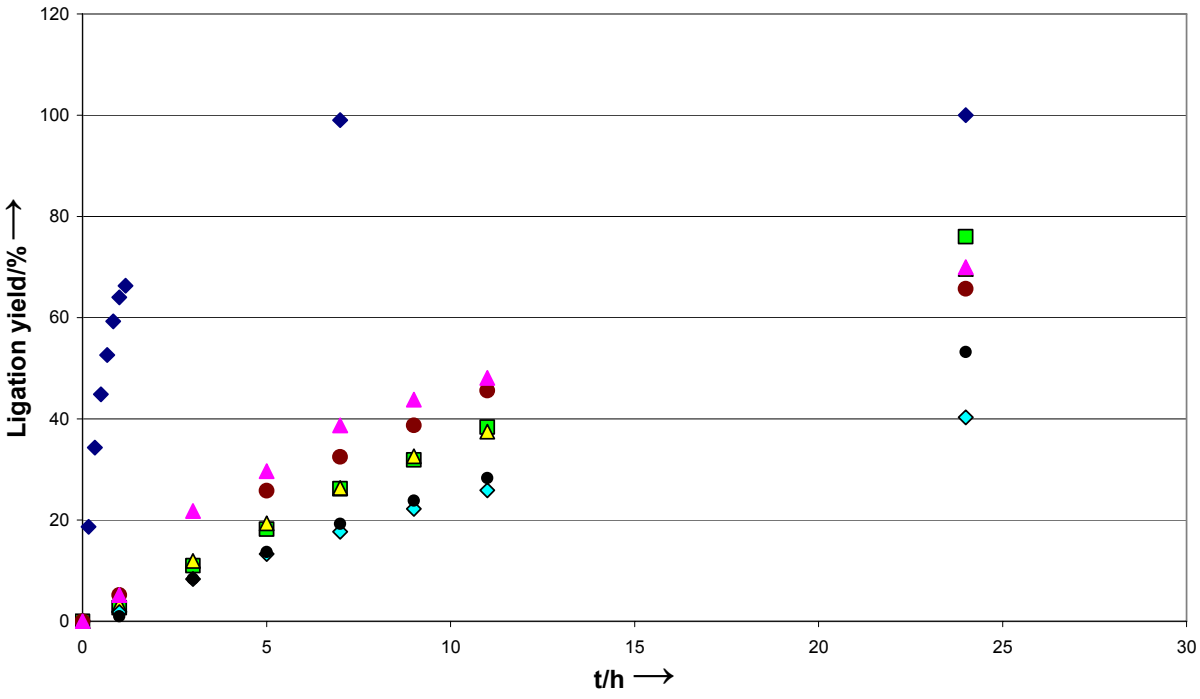
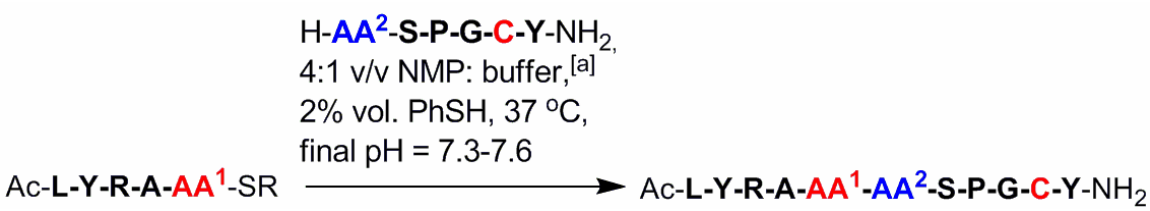


Figure S2. Kinetics of cysteine-free ligation reactions between peptide thioester **1** and peptides **2** and **4-9** containing cysteine residues. ▲ peptide **2**; ♦ peptide **4**; ● peptide **5**; ● peptide **6**; ♦ peptide **7**; ▲ peptide **8**; ■ peptide **9**.

Direct Aminolysis Reactions of Cysteine-Containing Peptides

Table S1



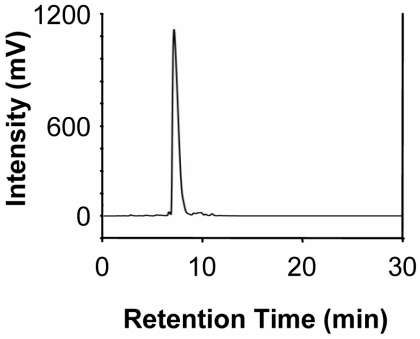
Entry	Peptide thioester AA^1	Peptide AA^2	Ligation junction AA^1-AA^2	Isolated Yield [%]
1	Gly	Gly	Gly-Gly	quant.[b]
2	Ala	Gly	Ala-Gly	84[b]
3	His	Gly	His-Gly	69[b]
4	Gly	His	Gly-His	52[c]
5	Ala	His	Ala-His	74[c]
6	His	His	His-His	41[c]
7	Gly	Ala	Gly-Ala	35[c]
8	Ala	Ala	Ala-Ala	35[c]
9	His	Ala	His-Ala	55[c]
10	Gly	Asp	Gly-Asp	77[b]
11	Ala	Asp	Ala-Asp	64[b]
12	His	Asp	His-Asp	83[b]
13	Gly	Glu	Gly-Glu	84[b]
14	Ala	Glu	Ala-Glu	68[b]
15	His	Glu	His-Glu	quant.[b]

[a] buffer = 6M Gn.HCl, 1M HEPES, pH 8.5; [b] t = 48 h; [c] t = 96 h.

Analytical Data:

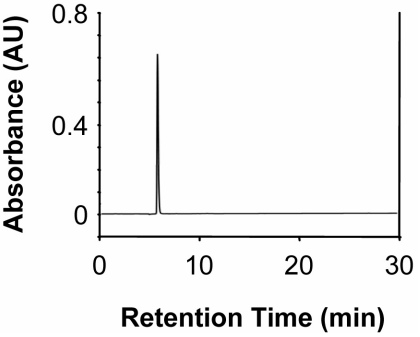
Peptides from Table 1

H-Gly Ser Pro Gly Tyr Ser-NH_2



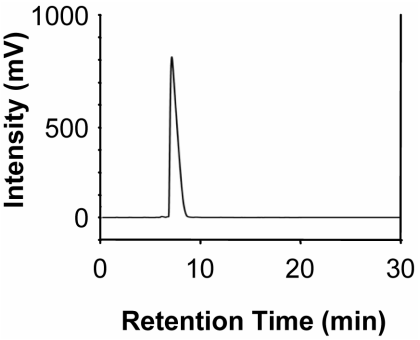
Starting from 100 µmol Fmoc protected Rink Amide resin, the linear assembly was performed following the Fmoc-strategy. Yield: 44.0 mg (78 µmol, 78%). ESI-TOF high-acc. (*m/z*): 566.2574 ($[M+H]^+$, theor. 566.2569). HPLC: t_R: 7.1 min (Gradient B); $C_{24}H_{35}N_7O_9$ (565.58).

H-Ala Ser Pro Gly Tyr Ser-NH_2



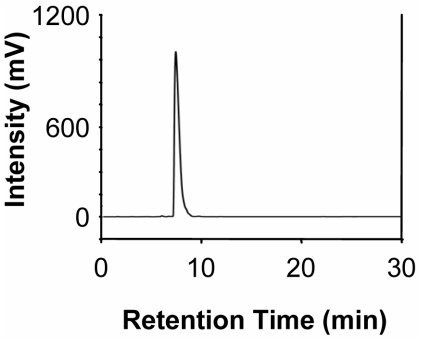
Starting from 80 µmol Fmoc protected Rink Amide resin, the linear assembly was performed following the Fmoc-strategy. Yield: 30.0 mg (52 µmol, 65%). ESI-TOF high-acc. (*m/z*): 580.2731 ($[M+H]^+$, theor. 580.2725). HPLC: t_R: 5.7 min (Gradient A); $C_{25}H_{37}N_7O_9$ (579.60).

H-His Ser Pro Gly Tyr Ser-NH_2



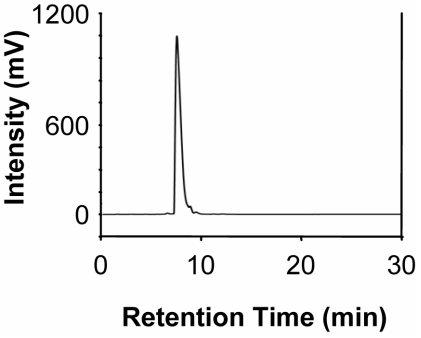
Starting from 80 µmol Fmoc protected Rink Amide resin, the linear assembly was performed following the Fmoc-strategy. Yield: 41.0 mg (64 µmol, 80%). ESI-TOF high-acc. (*m/z*): 646.2941 ($[M+H]^+$, theor. 646.2943). HPLC: t_R: 7.0 min (Gradient B); $C_{28}H_{39}N_9O_9$ (645.66).

H-Asp Ser Pro Gly Tyr Ser-NH_2



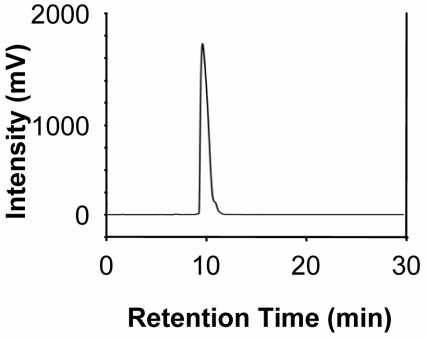
Starting from 80 µmol Fmoc protected Rink Amide resin, the linear assembly was performed following the Fmoc-strategy. Yield: 43.7 mg (70 µmol, 88%). ESI-TOF high-acc. (*m/z*): 624.2636 ($[M+H]^+$, theor. 624.2624). HPLC: t_R: 7.3 min (Gradient B); $C_{26}H_{37}N_7O_{11}$ (623.61).

H-Glu Ser Pro Gly Tyr Ser-NH_2



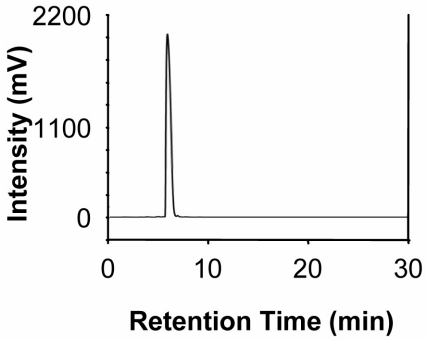
Starting from 80 µmol Fmoc protected Rink Amide resin, the linear assembly was performed following the Fmoc-strategy. Yield: 36.0 mg (56 µmol, 70%). ESI-TOF high-acc. (*m/z*): 638.2787 ($[M+H]^+$, theor. 638.278). HPLC: t_R: 7.4 min (Gradient B); $C_{27}H_{39}N_7O_{11}$ (637.64).

H-Tyr Ser Pro Gly Tyr Ser-NH_2



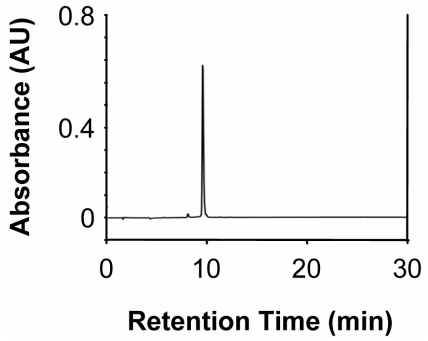
Starting from 80 µmol Fmoc protected Rink Amide resin, the linear assembly was performed following the Fmoc-strategy. Yield: 46.0 mg (68.5 µmol, 86%). ESI-TOF high-acc. (*m/z*): 672.2994 ($[M+H]^+$, theor. 672.2988). HPLC: t_R: 9.6 min (Gradient B); $C_{31}H_{41}N_7O_{10}$ (671.70).

H-Gly Ser Pro Gly Lys Tyr-NH_2



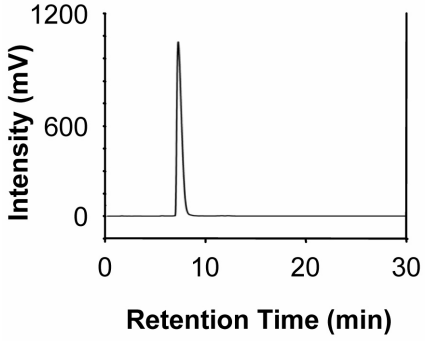
Starting from 100 µmol Fmoc protected Rink Amide resin, the linear assembly was performed following the Fmoc-strategy. Yield: 52 mg (85.7 µmol, 86%). ESI-TOF high-acc. (*m/z*): 607.3204 ($[M+H]^+$, theor. 607.3198). HPLC: t_R: 5.8 min (Gradient B); $C_{27}H_{42}N_8O_8$ (606.67).

H-Gly Ser Pro Gly Lys(ivDde) Tyr-NH_2



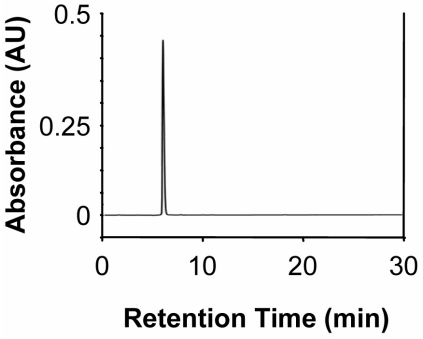
Starting from 100 µmol Fmoc protected Rink Amide resin, the linear assembly was performed following the Fmoc-strategy. Yield: 80 mg (94 µmol, 94%). ESI-TOF high-acc. (*m/z*): 813.4503 ($[M+H]^+$, theor. 813.4511). HPLC: t_R: 9.5 min (Gradient A); $C_{40}H_{60}N_8O_{10}$ (812.95).

H-Gly Ser(β-GlcNAc) Pro Gly Tyr Ser-NH_2



ESI-TOF high-acc. (*m/z*): 769.3377 ([M+H]$^+$, theor. 769.3368). HPLC: t_R: 7.1 min (Gradient B); $C_{32}H_{48}N_8O_{14}$ (768.77).

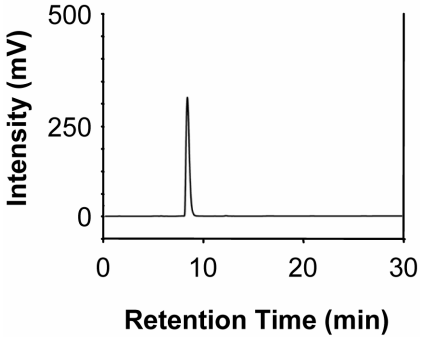
H-Gly Thr(α-GalNAc) Pro Gly Tyr Ser-NH$_2$



ESI-TOF high-acc. (*m/z*): 783.3521 ([M+H]$^+$, theor. 783.3525). HPLC: t_R: 5.9 min (Gradient A); $C_{33}H_{50}N_8O_{14}$ (782.80).

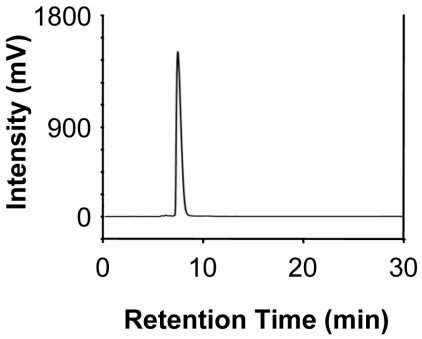
Peptides from Table 2

H-CysSerProGlyTyrSer-NH$_2$



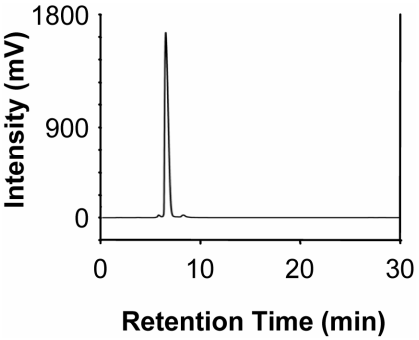
Starting from 100 µmol Fmoc protected Rink Amide resin, the linear assembly was performed following the Fmoc-strategy. Yield: 59.1 mg (97 µmol, 97%). ESI-TOF high-acc. (*m/z*): 612.2446 ([M+H]$^+$, theor. 612.2446). HPLC: t_R: 8.3 min (Gradient B); $C_{25}H_{37}N_7O_9S$ (611.67).

H-GlyCysProGlyTyrSer-NH$_2$



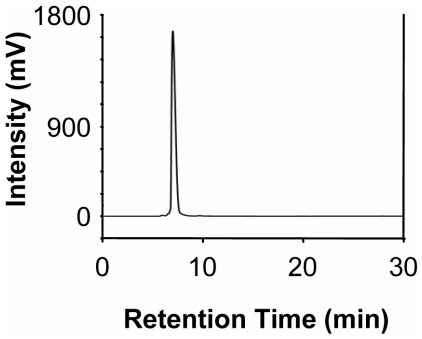
Starting from 100 µmol Fmoc protected Rink Amide resin, the linear assembly was performed following the Fmoc-strategy. Yield: 54.4 mg (94 µmol, 94%). ESI-TOF high-acc. (*m/z*): 582.2355 ($[M+H]^+$, theor. 582.234). HPLC: t_R: 7.3 min (Gradient B); $C_{24}H_{35}N_7O_8S$ (581.64).

H-GlySerCysGlyTyrSer-NH$_2$



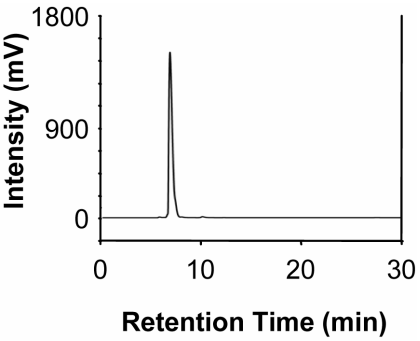
Starting from 100 µmol Fmoc protected Rink Amide resin, the linear assembly was performed following the Fmoc-strategy. Yield: 54.1 mg (95 µmol, 95%). ESI-TOF high-acc. (*m/z*): 572.2144 ($[M+H]^+$, theor. 572.2133). HPLC: t_R: 6.4 min (Gradient B); $C_{22}H_{33}N_7O_9S$ (571.60).

H-GlySerProCysTyrSer-NH$_2$



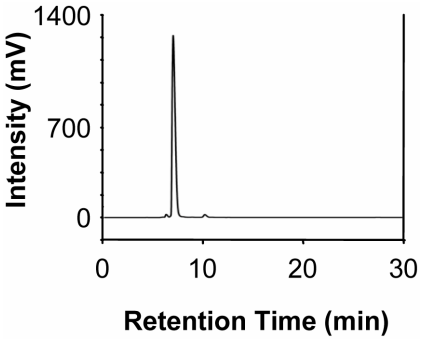
Starting from 100 µmol Fmoc protected Rink Amide resin, the linear assembly was performed following the Fmoc-strategy. Yield: 59.3 mg (97 µmol, 97%). ESI-TOF high-acc. (*m/z*): 612.2456 ($[M+H]^+$, theor. 612.2446). HPLC: t_R: 6.8 min (Gradient B); $C_{25}H_{37}N_7O_9S$ (611.67).

H-GlySerProGlyCysTyr-NH$_2$



Starting from 100 µmol Fmoc protected Rink Amide resin, the linear assembly was performed following the Fmoc-strategy. Yield: 54.2 mg (93 µmol, 93%). ESI-TOF high-acc. (*m/z*): 582.2346 ($[M+H]^+$, theor. 582.234). HPLC: t_R: 6.8 min (Gradient B); $C_{24}H_{35}N_7O_8S$ (581.64).

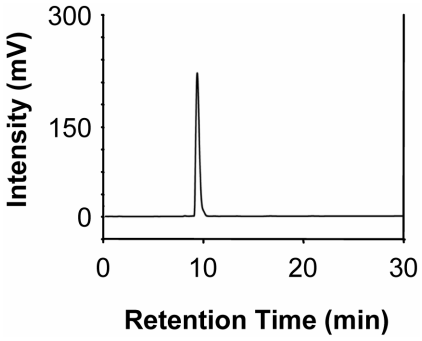
H-GlySerProGlyTyrCys-NH_2



Starting from 100 µmol Fmoc protected Rink Amide resin, the linear assembly was performed following the Fmoc-strategy. Yield: 55.2 mg (95 µmol, 95%). ESI-TOF high-acc. (*m/z*): 582.2348 ($[M+H]^+$, theor. 582.234). HPLC: t_R: 6.9 min (Gradient B); $C_{24}H_{35}N_7O_8S$ (581.64).

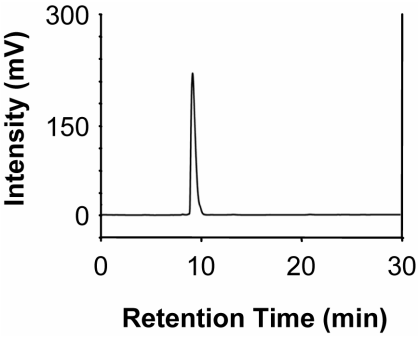
Peptides from Table 3

H-Ala Ser Pro Gly Cys Tyr-NH_2



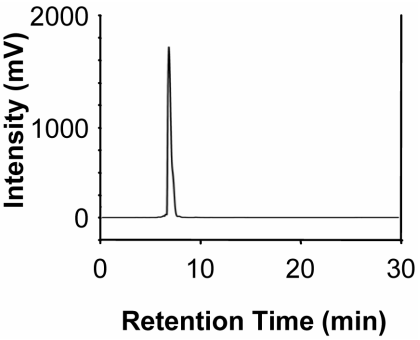
Starting from 100 µmol Fmoc protected Rink Amide resin, the linear assembly was performed following the Fmoc-strategy. Yield: 48.1 mg (81 µmol, 81%). ESI-TOF high-acc. (*m/z*): 596.2503 ($[M+H]^+$, theor. 596.2497). HPLC: t_R: 9.3 min (Gradient B); $C_{25}H_{37}N_7O_8S$ (595.67).

H-His Ser Pro Gly Cys Tyr-NH_2



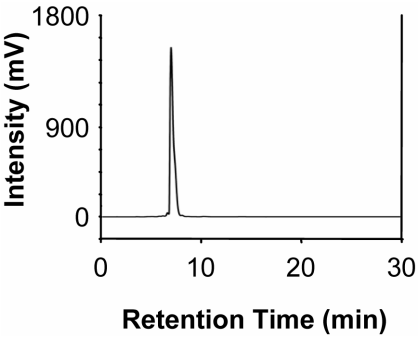
Starting from 100 µmol Fmoc protected Rink Amide resin, the linear assembly was performed following the Fmoc-strategy. Yield: 53.7 mg (81 µmol, 81%). ESI-TOF high-acc. (*m/z*): 662.2726 ($[M+H]^+$, theor. 662.2715). HPLC: t_R: 9.1 min (Gradient B); $C_{28}H_{39}N_9O_8S$ (661.73).

H-Asp Ser Pro Gly Cys Tyr-NH_2



Starting from 100 µmol Fmoc protected Rink Amide resin, the linear assembly was performed following the Fmoc-strategy. Yield: 38.3 mg (60 µmol, 60%). ESI-TOF high-acc. (*m/z*): 640.2407 ($[M+H]^+$, theor. 640.2395). HPLC: t_R: 6.7 min (Gradient B); $C_{26}H_{37}N_7O_{10}S$ (639.68).

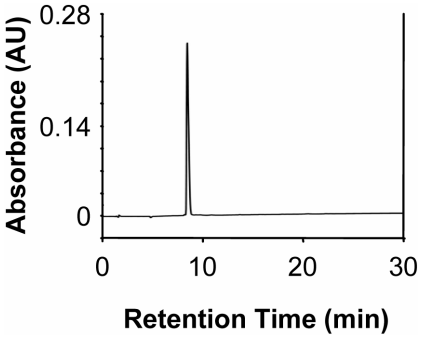
H-Glu Ser Pro Gly Cys Tyr-NH_2



Starting from 100 µmol Fmoc protected Rink Amide resin, the linear assembly was performed following the Fmoc-strategy. Yield: 52.8 mg (81 µmol, 81%). ESI-TOF high-acc. (*m/z*): 654.2561 ($[M+H]^+$, theor. 654.2552). HPLC: t_R: 6.9 min (Gradient B); $C_{27}H_{39}N_7O_{10}S$ (653.70).

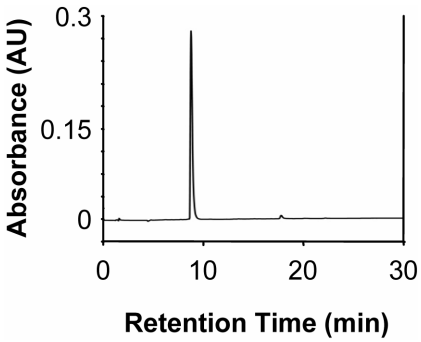
Peptide thioesters

Ac-Leu Tyr Arg Ala Gly-S$(CH_2)_2$CONH$_2$



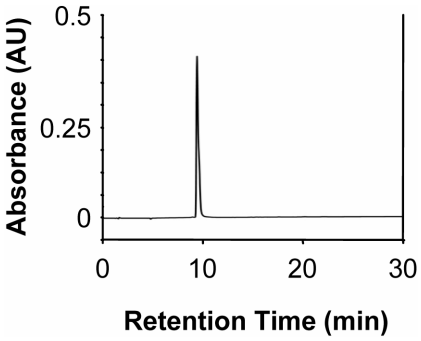
Starting from 100 µmol MBHA-resin preloaded with 3-(tritylthio) propanoic acid, the linear assembly was performed following the Boc-strategy. Yield: 50.6 mg (72 µmol, 72%). ESI-TOF high-acc. (*m/z*): 708.3499 ([M+H]$^+$, theor. 708.3497). HPLC: t_R: 8.4 min (Gradient C); $C_{31}H_{49}N_9O_8S$ (707.3425).

Ac-Leu Tyr Arg Ala Ala-S$(CH_2)_2$CONH$_2$



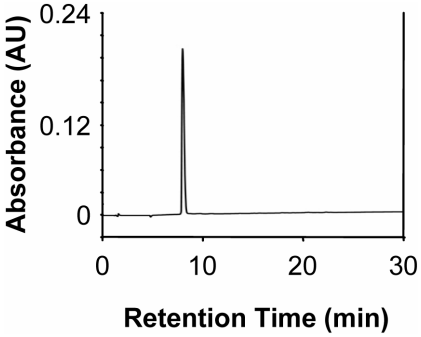
Starting from 100 µmol MBHA-resin preloaded with 3-(tritylthio) propanoic acid, the linear assembly was performed following the Boc-strategy. Yield: 39 mg (52 µmol, 52%). ESI-TOF high-acc. (*m/z*): 722.3644 ([M+H]$^+$, theor. 722.3654). HPLC: t_R: 8.8 min (Gradient C); $C_{32}H_{51}N_9O_8S$ (721.3581).

Ac-Leu Tyr Arg Ala Tyr-S$(CH_2)_2$CONH$_2$



Starting from 100 µmol MBHA-resin preloaded with 3-(tritylthio) propanoic acid, the linear assembly was performed following the Boc-strategy. Yield: 33.3 mg (41 µmol, 41%). ESI-TOF high-acc. (*m/z*): 814.3901 ([M+H]$^+$, theor. 814.3916). HPLC: t_R: 9.3 min (Gradient C); $C_{38}H_{55}N_9O_9S$ (813.3843).

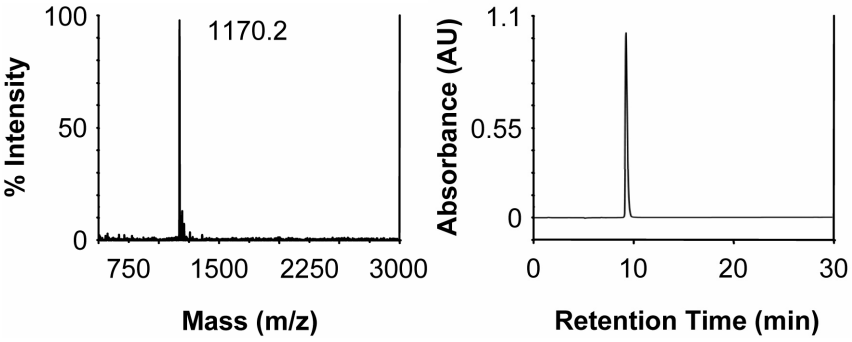
Ac-Leu Tyr Arg Ala His-S$(CH_2)_2$CONH$_2$



Starting from 100 µmol MBHA-resin preloaded with 3-(tritylthio) propanoic acid, the linear assembly was performed following the Boc-strategy. Yield: 31 mg (39 µmol, 39%). ESI-TOF high-acc. (*m/z*): 788.3863 ([M+H]$^+$, theor. 788.3872). HPLC: t_R: 8.0 min (Gradient C); $C_{35}H_{53}N_{11}O_8S$ (787.3799).

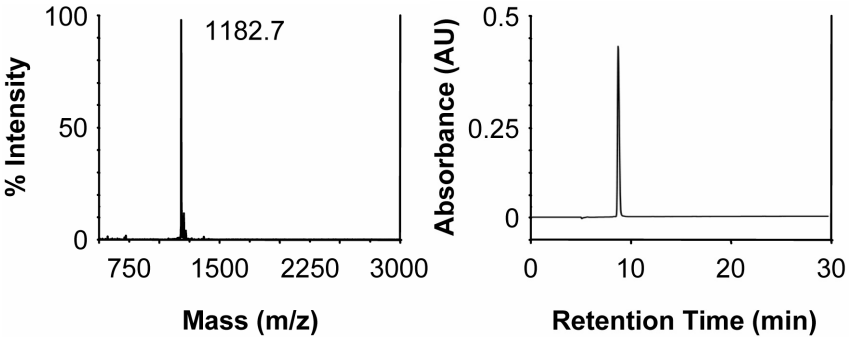
Ligation products from Table 1

Ac-LeuTyrArgAla**Gly-Gly**SerProGlyTyr Ser-NH$_2$



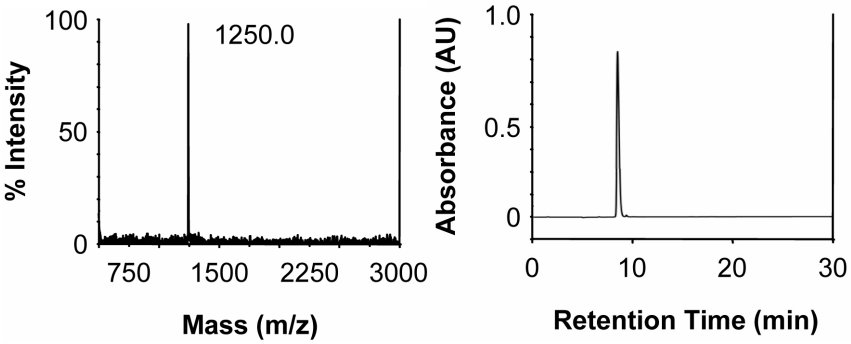
Yield: 89%. MALDI-TOF (*m/z*): 1170.2 ([M+H]$^+$, theor. 1169.3). HPLC: t_R: 9.2 min (Gradient A); $C_{52}H_{77}N_{15}O_{16}$ (1168.26).

Ac-LeuTyrArgAla**Ala-Gly**SerProGlyTyr Ser-NH_2



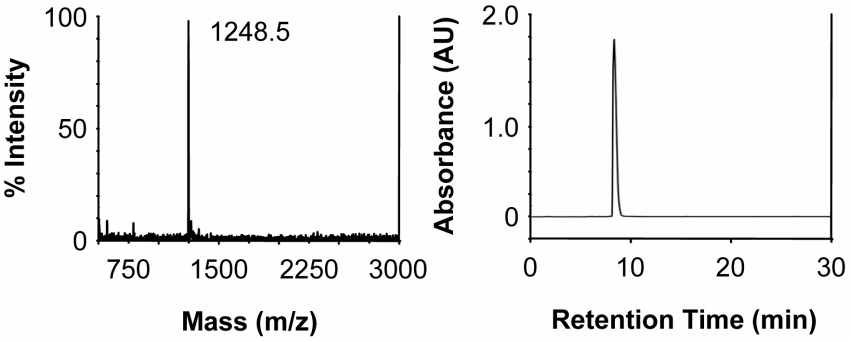
Yield: quant. MALDI-TOF (*m/z*): 1182.7 ($[M+H]^+$, theor. 1183.3). HPLC: t_R: 8.7 min (Gradient A); $C_{53}H_{79}N_{15}O_{16}$ (1182.3).

Ac-LeuTyrArgAla**His-Gly**SerProGlyTyr Ser-NH_2



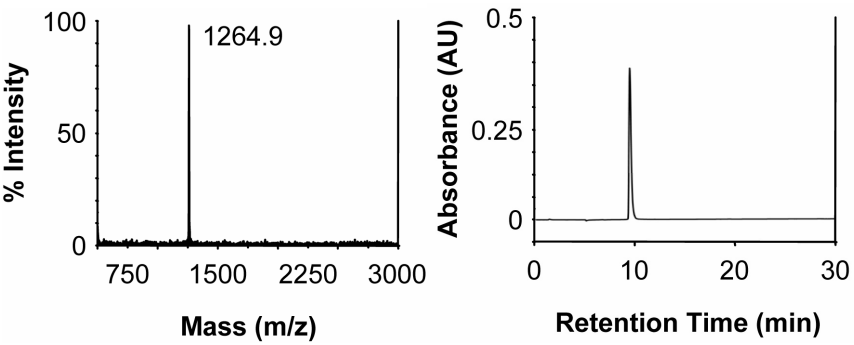
Yield: quant. MALDI-TOF (*m/z*): 1250.0 ($[M+H]^+$, theor. 1249.4). HPLC: t_R: 8.4 min (Gradient A); $C_{56}H_{81}N_{17}O_{16}$ (1248.3).

Ac-LeuTyrArgAla**Gly-His**SerProGlyTyr Ser-NH_2



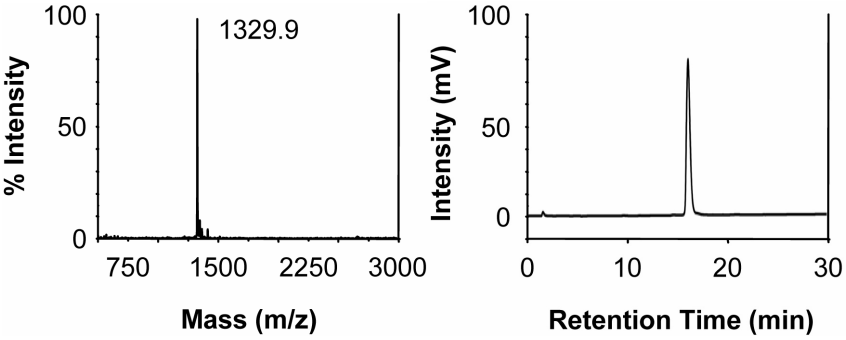
Yield: 57% MALDI-TOF (*m/z*): 1248.5 ($[M+H]^+$, theor. 1249.4). HPLC: t_R: 8.2 min (Gradient A); $C_{56}H_{81}N_{17}O_{16}$ (1248.3).

Ac-LeuTyrArgAla**Ala-His**SerProGlyTyr Ser-NH_2



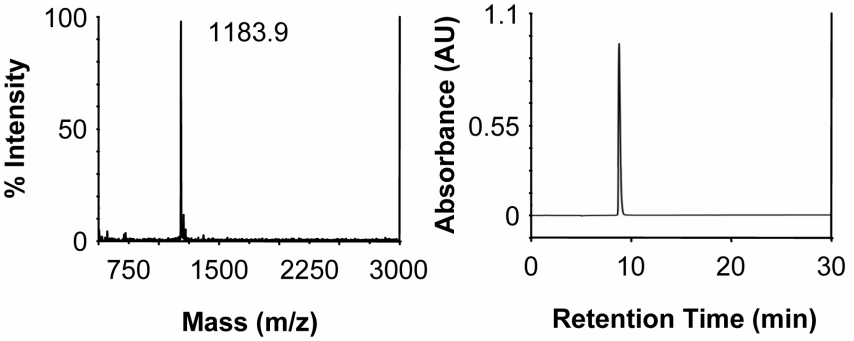
Yield: 40% MALDI-TOF (*m/z*): 1264.9 ($[M+H]^+$, theor. 1263.4). HPLC: t_R: 9.4 min (Gradient A); $C_{57}H_{83}N_{17}O_{16}$ (1262.4).

Ac-LeuTyrArgAla**His-His**SerProGlyTyr Ser-NH_2



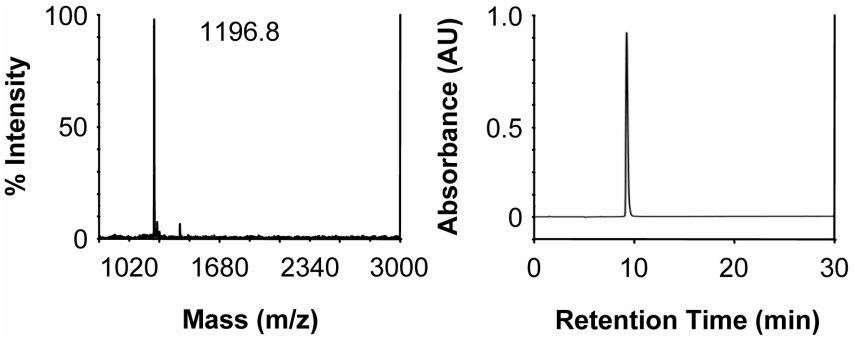
Yield: 41% MALDI-TOF (*m/z*): 1329.9 ($[M+H]^+$, theor. 1329.4). HPLC: t_R: 16.0 min (Gradient B); $C_{60}H_{85}N_{19}O_{16}$ (1328.4).

Ac-LeuTyrArgAla**Gly-Ala**SerProGlyTyr Ser-NH_2



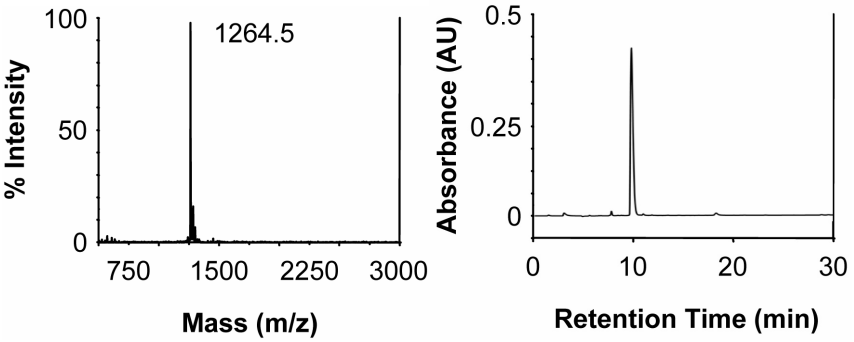
Yield: 66% MALDI-TOF (*m/z*): 1183.9 ($[M+H]^+$, theor. 1183.3). HPLC: t_R: 8.7 min (Gradient A); $C_{53}H_{79}N_{15}O_{16}$ (1182.3).

Ac-LeuTyrArgAla**Ala-Ala**SerProGlyTyr Ser-NH_2



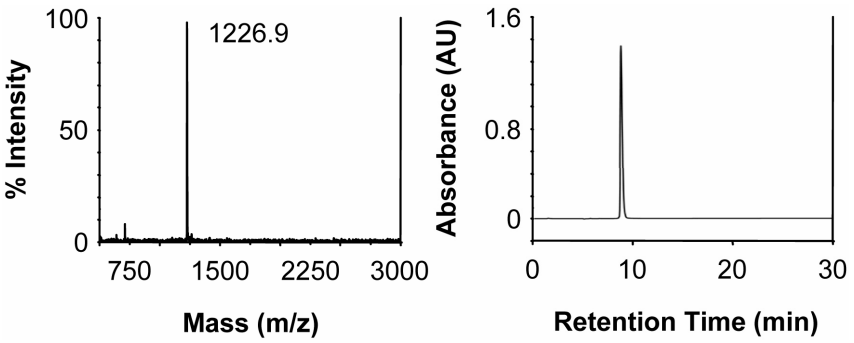
Yield: 54% MALDI-TOF (*m/z*): 1196.8 ($[M+H]^+$, theor. 1197.3). HPLC: t_R: 9.2 min (Gradient A); $C_{54}H_{81}N_{15}O_{16}$ (1196.3).

Ac-LeuTyrArgAla**His-Ala**SerProGlyTyr Ser-NH_2



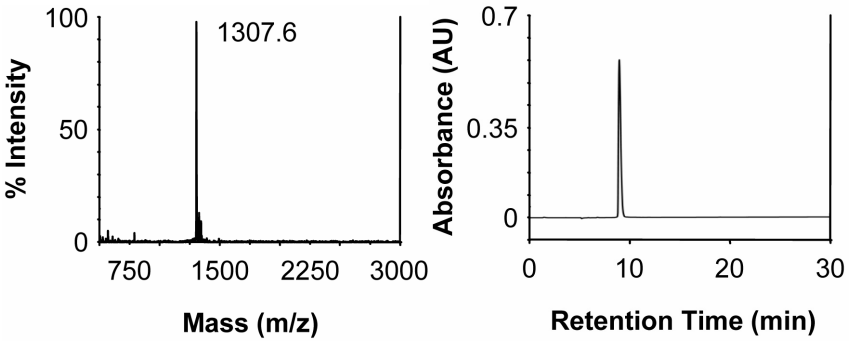
Yield: 37% MALDI-TOF (*m/z*): 1264.5 ($[M+H]^+$, theor. 1263.4). HPLC: t_R: 9.7 min (Gradient A); $C_{57}H_{83}N_{17}O_{16}$ (1262.4).

Ac-LeuTyrArgAla**Gly-Asp**SerProGlyTyr Ser-NH_2



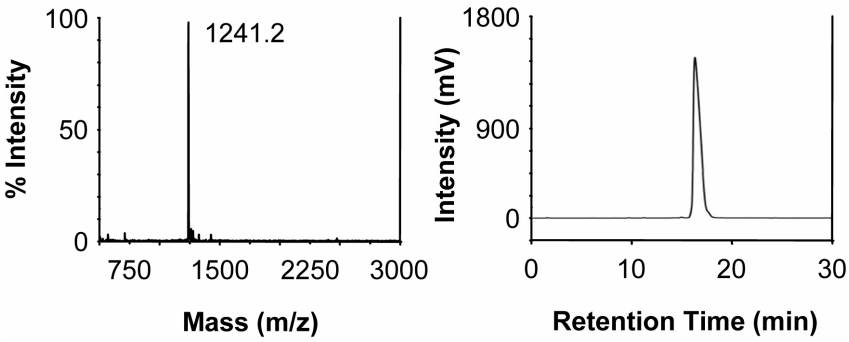
Yield: quant. MALDI-TOF (*m/z*): 1226.9 ($[M+H]^+$, theor. 1227.3). HPLC: t_R: 8.7 min (Gradient A); $C_{54}H_{79}N_{15}O_{18}$ (1226.3).

Ac-LeuTyrArgAla**His-Asp**SerProGlyTyr Ser-NH_2



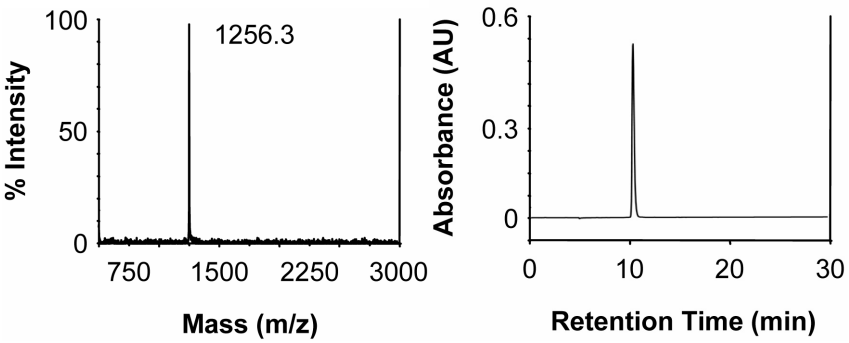
Yield: 89% MALDI-TOF (*m/z*): 1307.6 ($[M+H]^+$, theor. 1307.4). HPLC: t_R: 8.9 min (Gradient A); $C_{58}H_{83}N_{17}O_{18}$ (1306.4).

Ac-LeuTyrArgAla**Gly-Glu**SerProGlyTyr Ser-NH_2



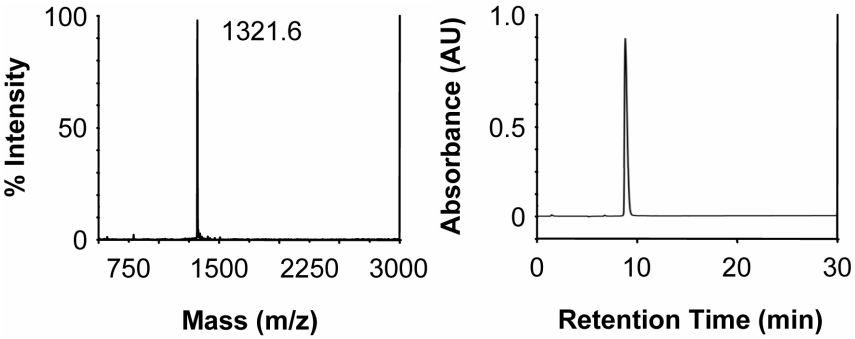
Yield: quant. MALDI-TOF (*m/z*): 1241.2 ($[M+H]^+$, theor. 1241.3). HPLC: t_R: 16.2 min (Gradient B); $C_{55}H_{81}N_{15}O_{18}$ (1240.3).

Ac-LeuTyrArgAla**Ala-Glu**SerProGlyTyr Ser-NH_2



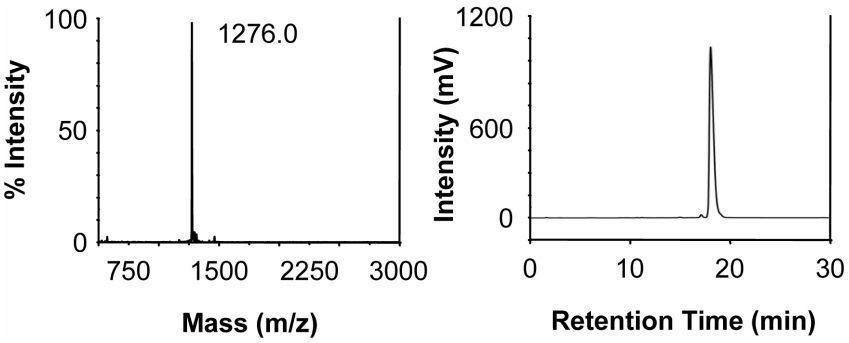
Yield: quant. MALDI-TOF (*m/z*): 1256.3 ($[M+H]^+$, theor. 1255.4). HPLC: t_R: 10.4 min (Gradient A); $C_{56}H_{83}N_{15}O_{18}$ (1254.3).

Ac-LeuTyrArgAla**His-Glu**SerProGlyTyr Ser-NH_2



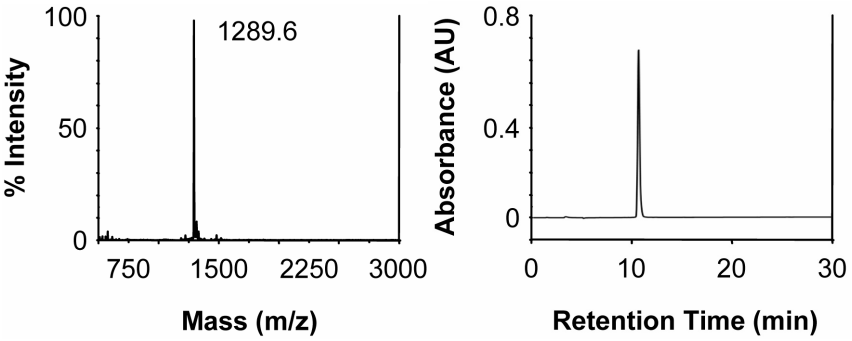
Yield: 90% MALDI-TOF (*m/z*): 1321.6 ($[M+H]^+$, theor. 1321.4). HPLC: t_R: 8.8 min (Gradient A); $C_{59}H_{85}N_{17}O_{18}$ (1320.4).

Ac-LeuTyrArgAla**Gly-Tyr**SerProGlyTyr Ser-NH_2



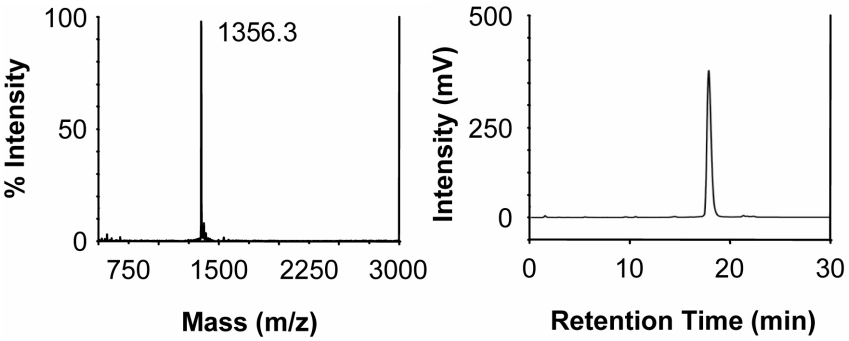
Yield: 39% MALDI-TOF (*m/z*): 1276.0 ($[M+H]^+$, theor. 1275.4). HPLC: t_R: 18.1 min (Gradient B); $C_{59}H_{83}N_{15}O_{17}$ (1274.4).

Ac-LeuTyrArgAla**Ala-Tyr**SerProGlyTyr Ser-NH_2



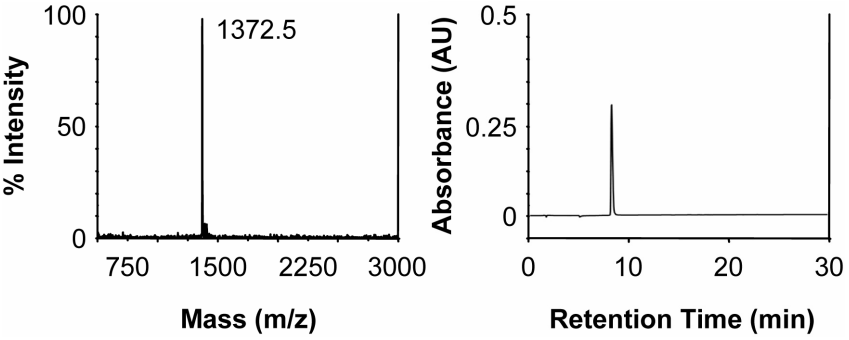
Yield: 34% MALDI-TOF (*m/z*): 1289.6 ($[M+H]^+$, theor. 1289.4). HPLC: t_R: 10.6 min (Gradient A); $C_{60}H_{85}N_{15}O_{17}$ (1288.4).

Ac-LeuTyrArgAla**His-Tyr**SerProGlyTyr Ser-NH_2



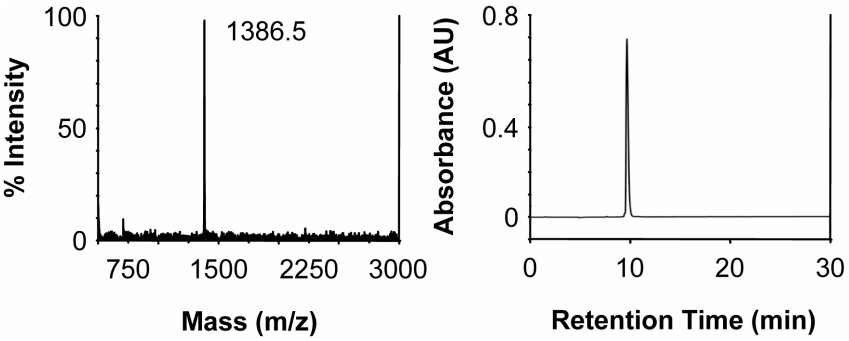
Yield: 12% MALDI-TOF (*m/z*): 1356.3 ($[M+H]^+$, theor. 1355.5). HPLC: t_R: 17.9 min (Gradient B); $C_{63}H_{87}N_{17}O_{17}$ (1354.5).

Ac-LeuTyrArgAla**Gly-Gly**Ser(β-GlcNAc)ProGlyTyr Ser-NH_2



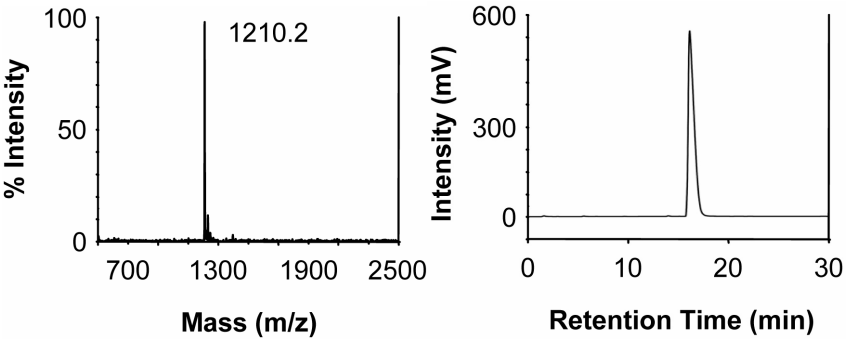
Yield: 98%. MALDI-TOF (*m/z*): 1372.5 ($[M+H]^+$, theor. 1372.5). HPLC: t_R: 8.2 min (Gradient A); $C_{60}H_{90}N_{16}O_{21}$ (1371.5).

Ac-LeuTyrArgAla**Gly-Gly**Thr(α-GalNAc)ProGlyTyr Ser-NH_2



Yield: 65%. MALDI-TOF (*m/z*): 1386.5 ($[M+H]^+$, theor. 1386.5). HPLC: t_R: 9.6 min (Gradient A); $C_{61}H_{92}N_{16}O_{21}$ (1385.5).

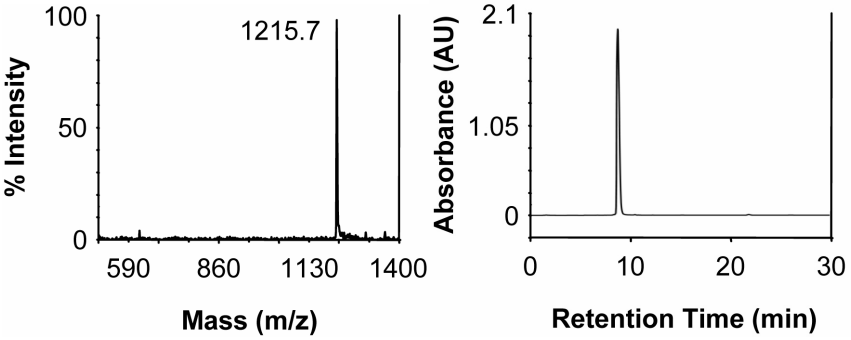
Ac-LeuTyrArgAla**Gly-Gly**SerProGlyLysTyr-NH_2



Yield: XX%. MALDI-TOF (*m/z*): 1210.2 ($[M+H]^+$, theor. 1210.4). HPLC: t_R: 16.1 min (Gradient B); $C_{55}H_{84}N_{16}O_{15}$ (1209.4).

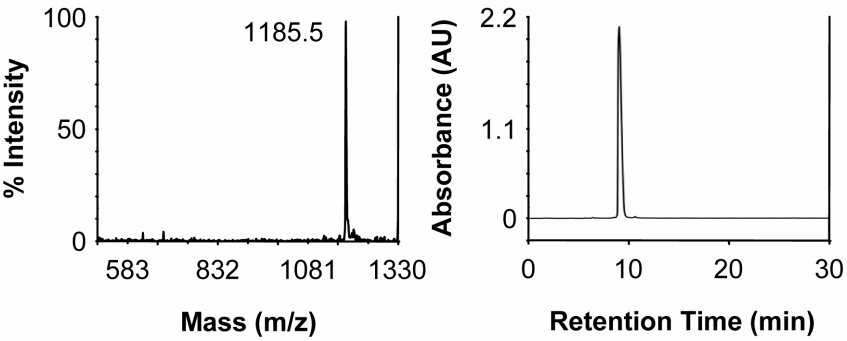
Ligation products from Table 2

Ac-LeuTyrArgAla**Gly-Cys**SerProGlyTyrSer-NH_2



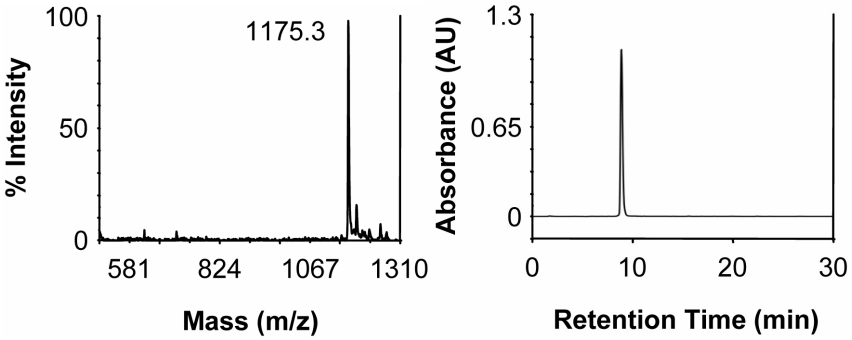
Yield: 94% MALDI-TOF (*m/z*): 1215.7 ($[M+H]^+$, theor. 1215.4). HPLC: t_R: 8.7 min (Gradient A); $C_{53}H_{79}N_{15}O_{16}S$ (1214.4).

Ac-LeuTyrArgAla**Gly-Gly**CysProGlyTyrSer-NH_2



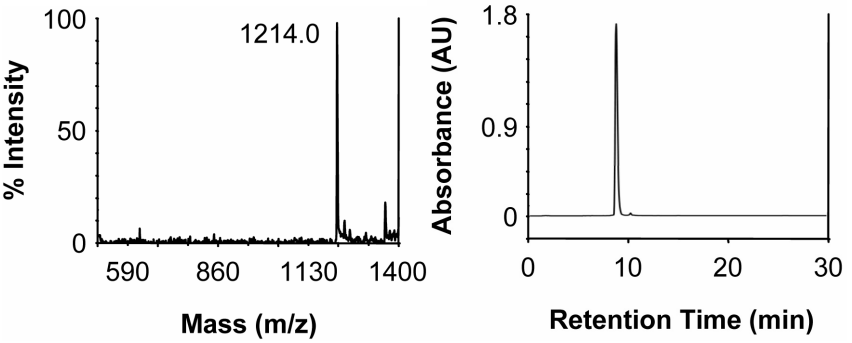
Yield: 69% MALDI-TOF (*m/z*): 1185.5 ($[M+H]^+$, theor. 1185.3). HPLC: t_R: 8.9 min (Gradient A); $C_{52}H_{77}N_{15}O_{15}S$ (1184.3).

Ac-LeuTyrArgAla**Gly-Gly**SerCysGlyTyrSer-NH_2



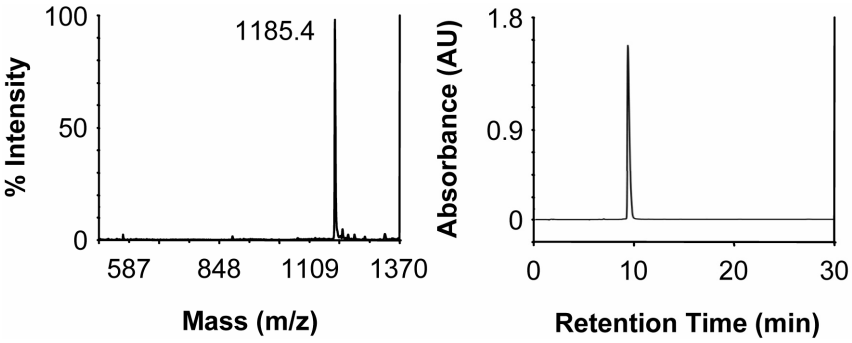
Yield: 69% MALDI-TOF (*m/z*): 1175.3 ($[M+H]^+$, theor. 1175.3). HPLC: t_R: 8.8 min (Gradient A); $C_{50}H_{75}N_{15}O_{16}S$ (1174.3).

Ac-LeuTyrArgAla**Gly-Gly**SerProCysTyrSer-NH_2



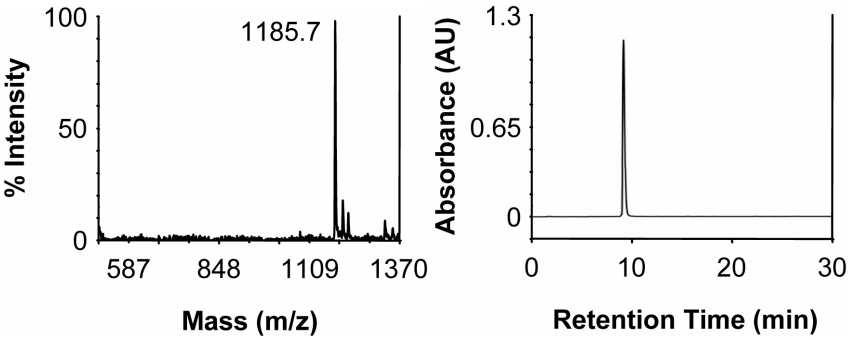
Yield: 44% MALDI-TOF (*m/z*): 1214.0 ($[M+H]^+$, theor. 1215.4). HPLC: t_R: 8.7 min (Gradient A); $C_{53}H_{79}N_{15}O_{16}S$ (1214.4).

Ac-LeuTyrArgAla**Gly-Gly**SerProGlyCysTyr-NH_2



Yield: 84% MALDI-TOF (*m/z*): 1185.4 ($[M+H]^+$, theor. 1185.3). HPLC: t_R: 9.2 min (Gradient A); $C_{52}H_{77}N_{15}O_{15}S$ (1184.3).

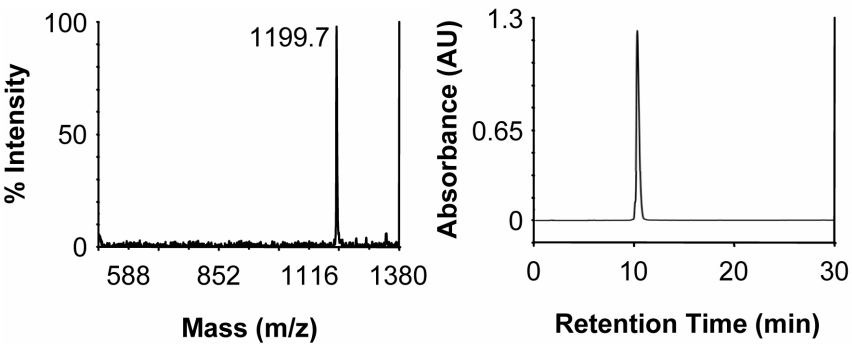
Ac-LeuTyrArgAla**Gly-Gly**SerProGlyTyrCys-NH_2



Yield: 87% MALDI-TOF (*m/z*): 1185.7 ($[M+H]^+$, theor. 1185.3). HPLC: t_R: 9.1 min (Gradient A); $C_{52}H_{77}N_{15}O_{15}S$ (1184.3).

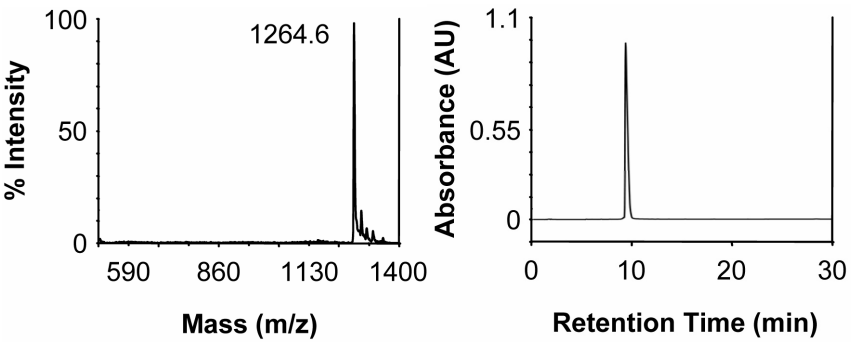
Ligation products from Table 3

Ac-LeuTyrArgAla**Ala-Gly**SerProGlyCysTyr-NH_2



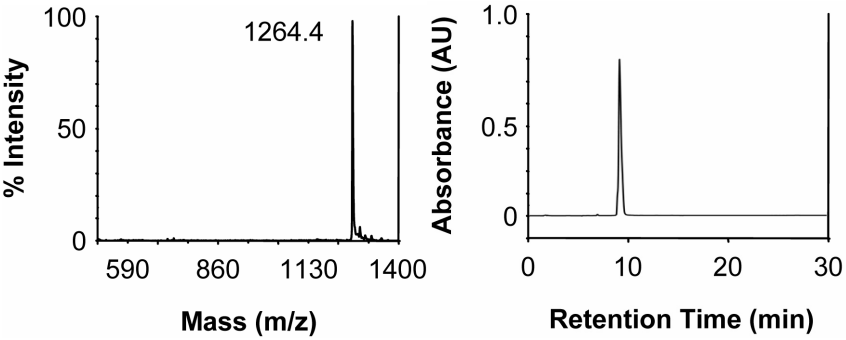
Yield: 84% MALDI-TOF (*m/z*): 1199.7 ($[M+H]^+$, theor. 1199.4). HPLC: t_R: 10.2 min (Gradient A); $C_{53}H_{79}N_{15}O_{15}S$ (1198.4).

Ac-LeuTyrArgAla**His-Gly**SerProGlyCysTyr-NH_2



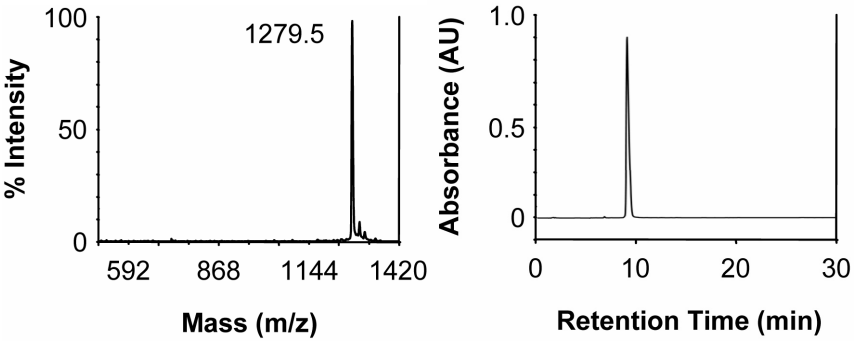
Yield: 69% MALDI-TOF (*m/z*): 1264.6 ($[M+H]^+$, theor. 1265.4). HPLC: t_R: 9.2 min (Gradient A); $C_{56}H_{81}N_{17}O_{15}S$ (1264.4).

Ac-LeuTyrArgAla**Gly-His**SerProGlyCysTyr-NH_2



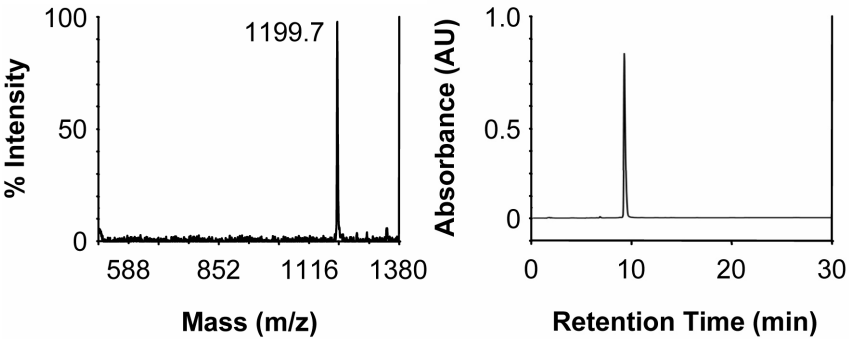
Yield: 52% MALDI-TOF (*m/z*): 1264.4 ($[M+H]^+$, theor. 1265.4). HPLC: t_R: 9.1 min (Gradient A); $C_{56}H_{81}N_{17}O_{15}S$ (1264.4).

Ac-LeuTyrArgAla**Ala-His**SerProGlyCysTyr-NH_2



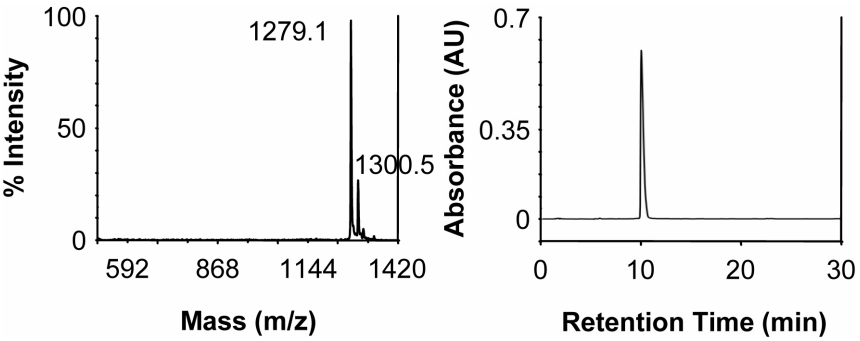
Yield: 74% MALDI-TOF (*m/z*): 1279.5 ($[M+H]^+$, theor. 1279.4). HPLC: t_R: 9.0 min (Gradient A); $C_{57}H_{83}N_{17}O_{15}S$ (1278.4).

Ac-LeuTyrArgAla**Gly-Ala**SerProGlyCysTyr-NH_2



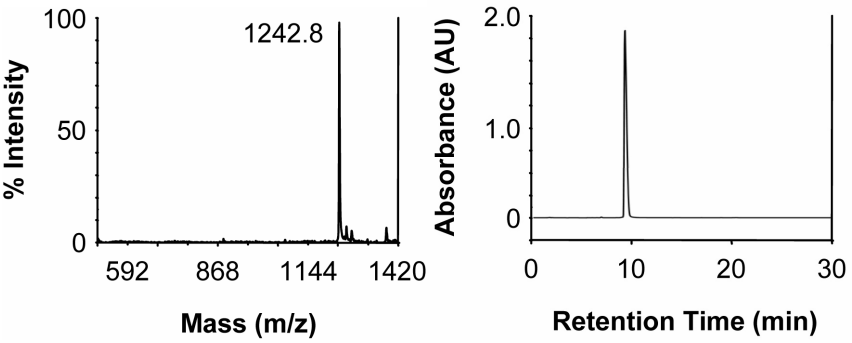
Yield: 35% MALDI-TOF (*m/z*): 1199.7 ($[M+H]^+$, theor. 1199.4). HPLC: t_R: 9.2 min (Gradient A); $C_{53}H_{79}N_{15}O_{15}S$ (1198.4).

Ac-LeuTyrArgAla**His-Ala**SerProGlyCysTyr-NH_2



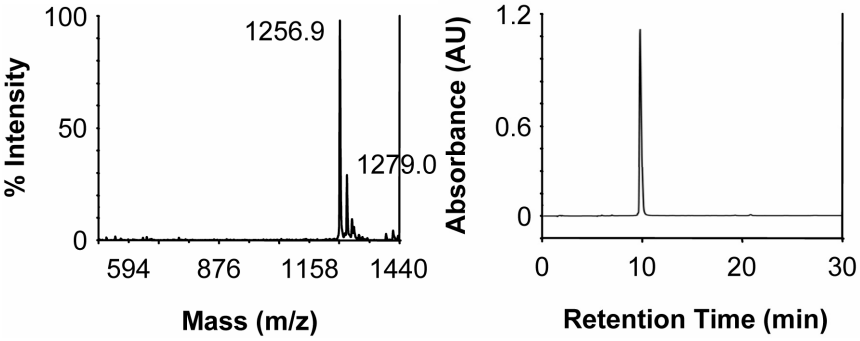
Yield: 55% MALDI-TOF (*m/z*): 1279.1 ($[M+H]^+$, theor. 1279.4; $[M+Na]^+$, theor. 1301.4). HPLC: t_R: 10.0 min (Gradient A); $C_{57}H_{83}N_{17}O_{15}S$ (1278.4).

Ac-LeuTyrArgAla**Gly-Asp**SerProGlyCysTyr-NH_2



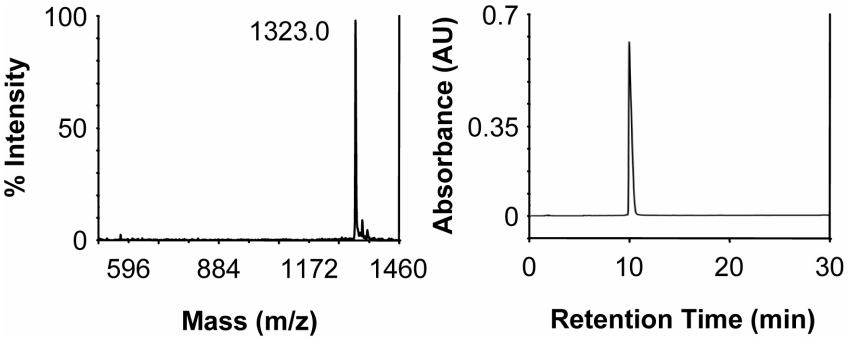
Yield: 77% MALDI-TOF (*m/z*): 1242.8 ($[M+H]^+$, theor. 1243.4). HPLC: t_R: 9.2 min (Gradient A); $C_{54}H_{79}N_{15}O_{17}S$ (1242.4).

Ac-LeuTyrArgAla**Ala-Asp**SerProGlyCysTyr-NH_2



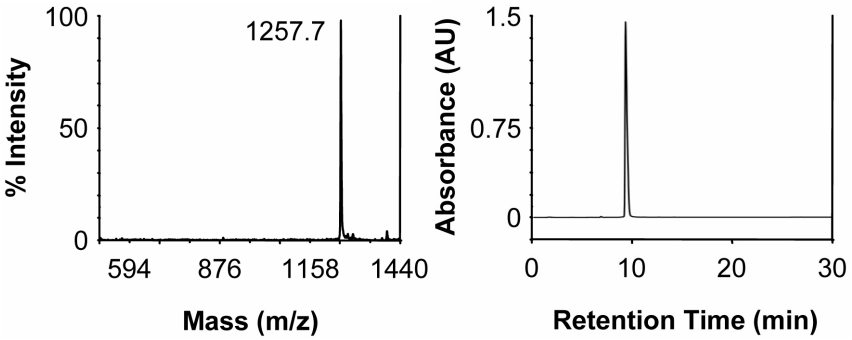
Yield: 64% MALDI-TOF (*m/z*): 1256.9 ($[M+H]^+$, theor. 1257.4; $[M+Na]^+$, theor. 1279.4). HPLC: t_R: 9.7 min (Gradient A); $C_{55}H_{81}N_{15}O_{17}S$ (1256.4).

Ac-LeuTyrArgAla**His-Asp**SerProGlyCysTyr-NH_2



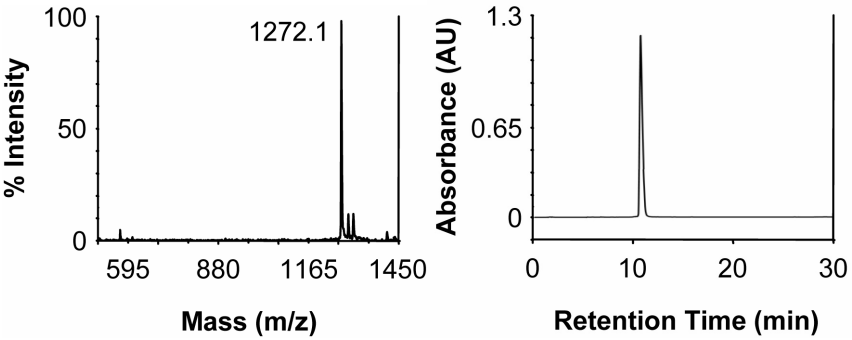
Yield: 83% MALDI-TOF (*m/z*): 1323.0 ($[M+H]^+$, theor. 1323.5). HPLC: t_R: 9.1 min (Gradient A); $C_{58}H_{83}N_{17}O_{17}S$ (1322.4).

Ac-LeuTyrArgAla**Gly-Glu**SerProGlyCysTyr-NH_2



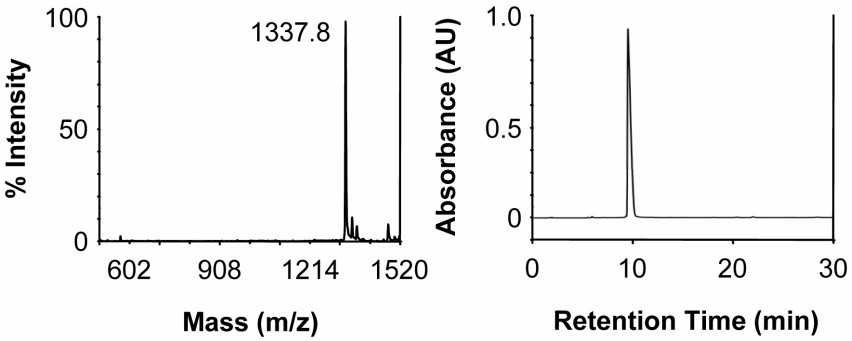
Yield: 84% MALDI-TOF (*m/z*): 1257.7 ($[M+H]^+$, theor. 1257.4). HPLC: t_R: 9.3 min (Gradient A); $C_{55}H_{81}N_{15}O_{17}S$ (1256.4).

Ac-LeuTyrArgAla**Ala-Glu**SerProGlyCysTyr-NH_2



Yield: 68% MALDI-TOF (*m/z*): 1272.1 ($[M+H]^+$, theor. 1271.4). HPLC: t_R: 10.7 min (Gradient A); $C_{56}H_{83}N_{15}O_{17}S$ (1270.4).

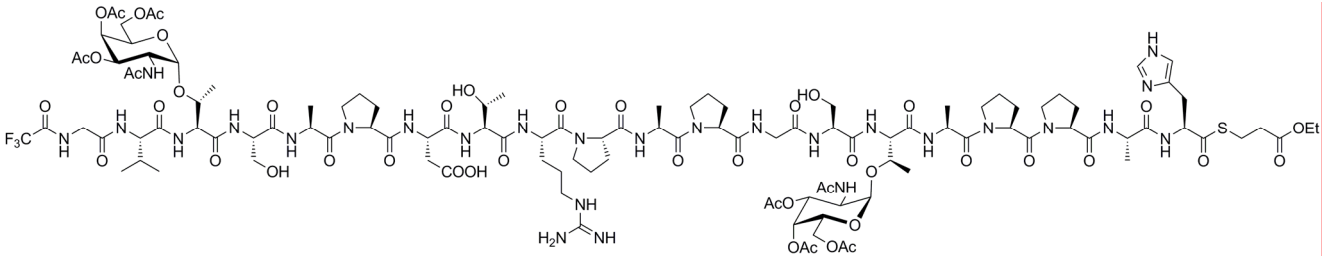
Ac-LeuTyrArgAla**His-Glu**SerProGlyCysTyr-NH_2



Yield: quant. MALDI-TOF (*m/z*): 1337.8 ($[M+H]^+$, theor. 1337.5). HPLC: t_R: 9.5 min (Gradient A); $C_{59}H_{85}N_{17}O_{17}S$ (1336.5).

Synthesis of MUC1 Glycopeptides

Synthesis of MUC1 glycopeptide thioester ***14***



Preloading of sulfamylbutyryl AM resin[3]

Fmoc-His(Boc)-OH (0.76 g, 1600 μmol) was dissolved in anhydrous chloroform (4.0 mL) and N, N-diisopropylethylamine (0.53 mL, 3200 μmol). This solution was cooled to -20 °C before the addition of PyBOP (0.83 g, 1600 μmol). The solution was transferred to a syringe containing 400 μmol of sulfamylbutyryl AM resin which had been pre-swelled in chloroform for 1 h. The syringe was stored at -20 °C for 8 h with gentle shaking every hour. The resin was washed with DCM (x5) and DMF (x10) before capping with 5:6:89 v/v/v acetic anhydride:2,6-lutidene:DMF for 4 min. The loading of the resin was 0.61 mmol/g as determined by deprotecting with 10% piperidine/DMF (2 x 5 min) and measuring the absorbance of the piperidine-fulvene adduct at 302 nm.

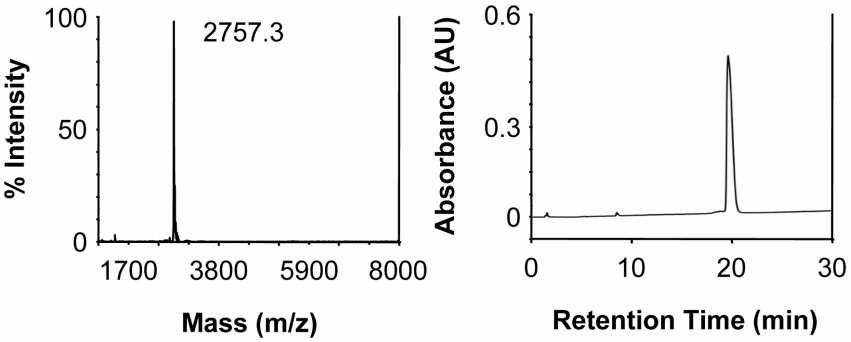
Amino acid residues were coupled by adding a preactivated solution containing 4 eq. of either Fmoc or trifluoroacetate protected amino acid (final concentration 0.1 M in DMF) using 4 eq. PyBOP and 8 eq. NMM to the resin. After 1 h, the resin was washed with DMF (5x), DCM (5x) and DMF (5x). After the sixth amino acid residue [Fmoc-Thr-α-GalNAc$(OAc)_3$-OH] all amino acids were double coupled (2 x 2 h) to maximize the overall yield of the peptide synthesis. *Capping*: 5:6:89 v/v/v acetic anhydride:2,6-lutidene:DMF. After 2 min the resin was washed with DMF (5x), DCM (5x) and DMF (5x). *Coupling of the glycosyl amino acid building blocks* was carried out by adding a preactivated solution of 1 eq. of the glycosylated amino acid (final concentration 0.1 M in DMF) using 1 eq. PyBOP and 2 eq. NMM with respect to the resin loading. After 12 h, the resin was washed with DMF (5x), DCM (5x) and DMF (5x).

Activation and thiolysis of sulfamylbutyryl AM resin[4]

16 μmol of the fully loaded resin (calculated based on the loading after coupling Fmoc-Val-OH) was washed with anhydrous DCM (x5) before treating with a 1 M solution of TMS-diazomethane in 1:1 v/v DCM:*n*-hexanes. After shaking for 1.5 h, the resin was washed with DCM (x5) and DMF (x10). Thiolysis was carried out by treating the resin with a mixture of mercaptopropionic acid ethyl ester (320 μL) and benzenethiol sodium salt (10 mg) in anhydrous DMF (2 mL). After shaking for 12 h, the solvent was removed in vacuo and dried under high vacuum for 12 h to afford the fully protected glycopeptide thioester as a yellow oil.

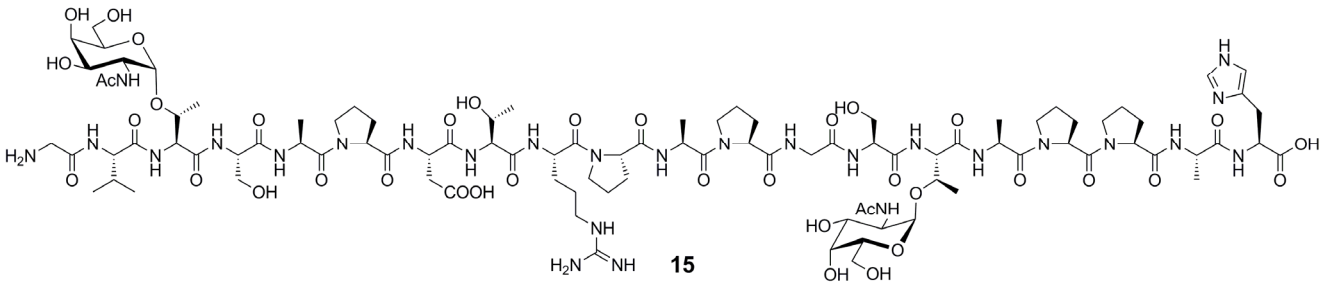
Protecting Group Cleavage: A 17:1:1:1 v/v/v/v mixture of TFA: thioanisole: triisopropylsilane: water was added to the residue above. After 1 h the solution was concentrated in vacuo. The resulting oil was dissolved in 20% MeCN in water + 0.1% TFA, purified by preparative HPLC (Gradient A, monitoring at 230 nm) and fractions analyzed by MALDI-TOF/MS (matrix: α-Cyano-4-

hydroxycinnamic acid). The desired glycopeptide thioester **14** was obtained, after lyophilization, as a fluffy white solid (12.5 mg, 27% yield).

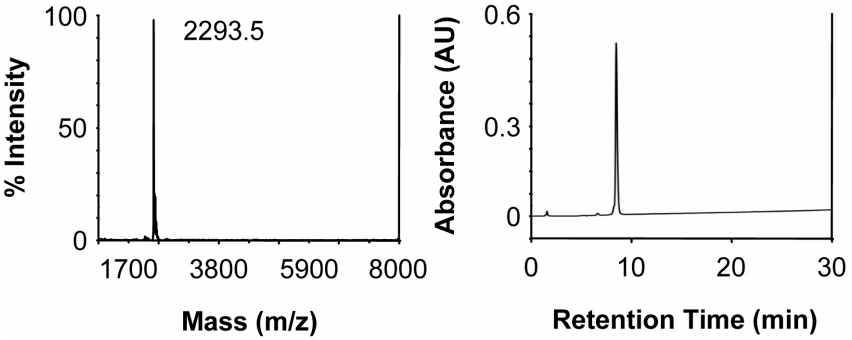


Yield: 27% MALDI-TOF (*m/z*): 2757.3 ([M+H]$^+$, theor. 2758.8). HPLC: t_R: 19.7 min (Gradient B); $C_{115}H_{172}F_3N_{27}O_{46}S$ (2757.8).

*Synthesis of Glycopeptide Starting Unit **15***

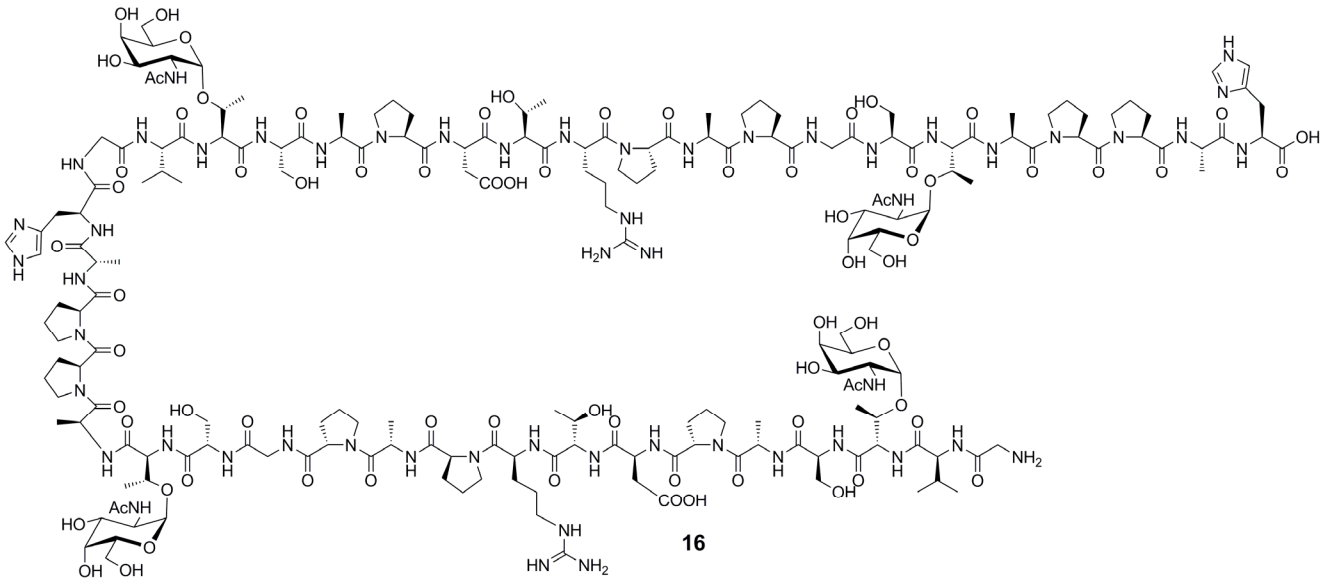


Glycopeptide thioester **14** (5.2 mg, 1.89 μmol) was dissolved in a 10 mM sodium hydroxide solution (1.7 mL) and incubated at rt for 24 h. After this time a second aliquot of 10 mM sodium hydroxide solution (1.5 mL) was added and the solution incubated at rt for a further 24 h, at which point the thioester functionality was completely hydrolyzed as determined by LC-MS analysis. A 5% hydrazine solution (160 μL) was added and the reaction incubated for a further 45 min at rt before the product was purified by preparative HPLC (Gradient A, monitoring at 230 nm) to afford the desired glycopeptide starting unit **15** as a fluffy white solid (4.0 mg, 92%).

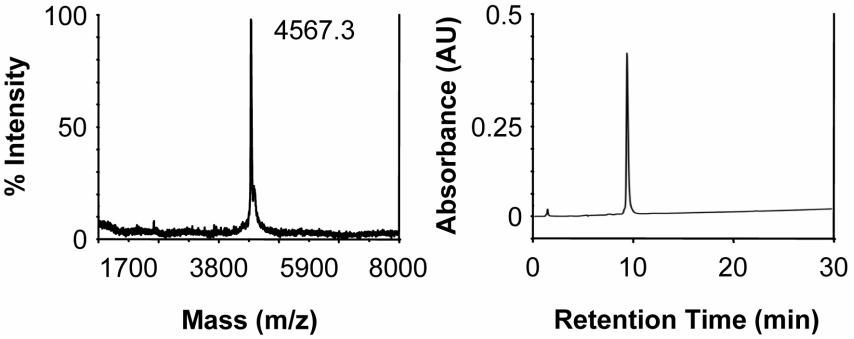


Yield: 92% MALDI-TOF (*m/z*): 2293.5 ([M+H]$^+$, theor. 2294.4). HPLC: t_R: 8.4 min (Gradient B); $C_{96}H_{153}N_{27}O_{38}$ (2293.4).

*Direct Aminolysis of Glycopeptide **15** and Glycopeptide Thioester **14***

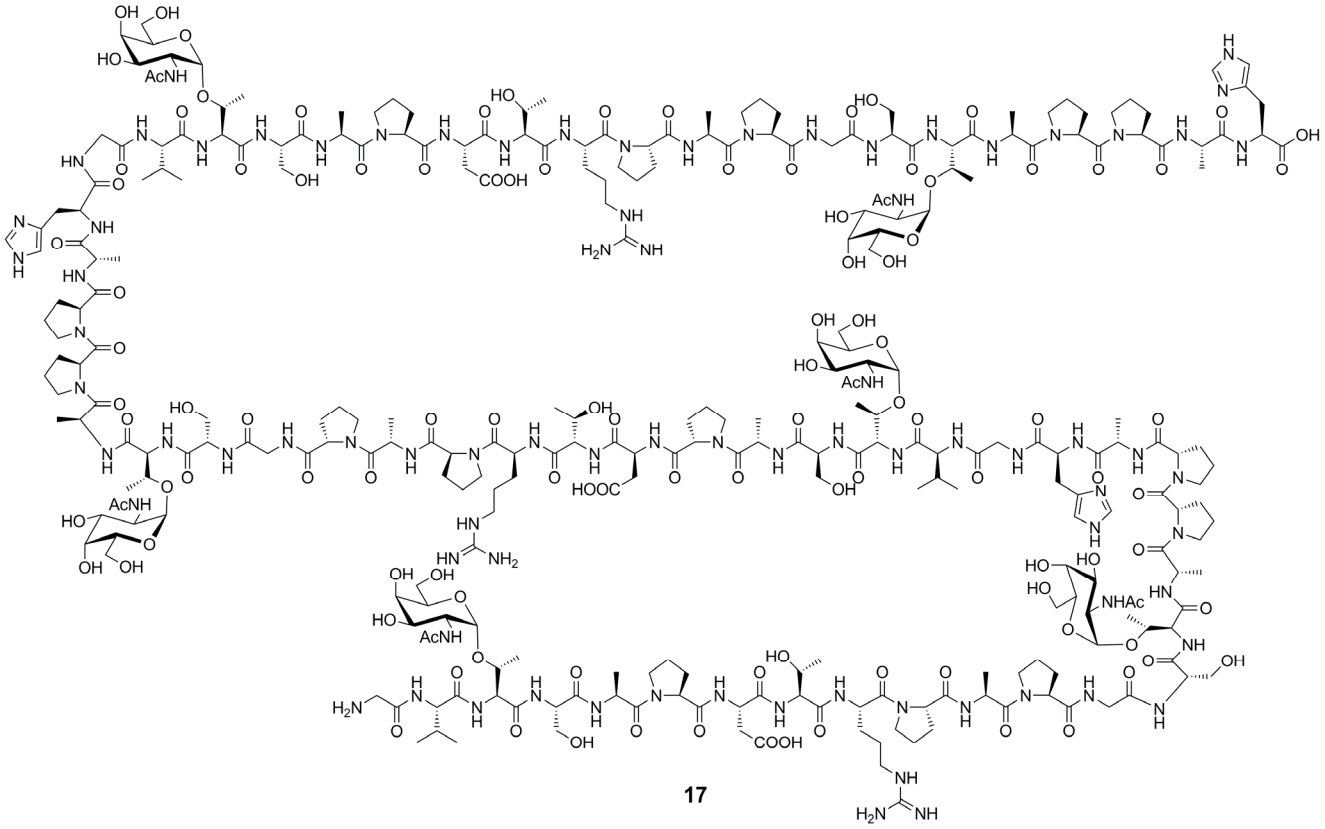


Glycopeptide **15** (1.5 mg, 0.65 μmol) and glycopeptide thioester **14** (3.6 mg, 1.31 μmol) were dissolved in 36 μL of 4:1 v/v NMP: 6 M Gn.HCl, 1 M HEPES, pH 8.5. The acidity of the peptide components caused the pH to drop to 5.5, therefore addition of 2.3 μL of a 2 M sodium hydroxide solution was required to bring the pH up to 7.6 where the ligation could take place. Thiophenol (0.67 μL) was added to the solution, which was mixed thoroughly and incubated at 37 ^oC for 60 h. A 5% hydrazine solution (400 μL) was added and the reaction incubated at rt for 1 h, before quenching with 0.1% TFA in water (400 μL). The reaction was purified by semi-preparative HPLC (Gradient A, monitoring at 230 nm) to afford the desired 40-mer ligation product **16** as a white solid (2.3 mg, 77%).



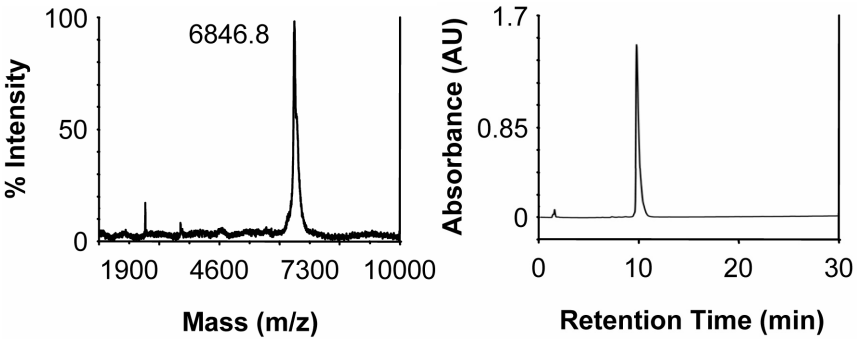
Yield: 77% MALDI-TOF (*m/z*): 4567.3 ($[M+H]^+$, theor. 4569.8). HPLC: t_R: 9.4 min (Gradient B); $C_{192}H_{304}N_{54}O_{75}$ (4568.8).

*Direct Aminolysis of Dimeric Glycopeptide **16** and Glycopeptide Thioester **14***



The 40-mer glycopeptide **16** (1.0 mg, 0.22 μmol) and glycopeptide thioester **14** (1.21 mg, 0.44 μmol) were dissolved in 12 μL of 4:1 v/v NMP: 6 M Gn.HCl, 1 M HEPES, pH 9.5. The acidity of the peptide components caused the pH to drop to 5.6, therefore addition of 1.0 μL of a 2 M sodium hydroxide solution was required to bring the pH up to 7.6 where the ligation could take place. Thiophenol (0.24 μL) was added to the solution, which was mixed thoroughly and incubated at 37 ^oC for 60 h. A 5%

hydrazine solution (400 μL) was added and the reaction incubated at rt for 1 h, before quenching with 0.1% TFA in water (400 μL). The reaction was purified by semi-preparative HPLC (gradient A, monitoring at 230 nm) to afford the desired 60-mer ligation product **17** as a white solid (1.2 mg, 80%).



Yield: 80% MALDI-TOF (*m/z*): 6846.8 ($[M+H]^+$, theor. 6845.2). HPLC: t_R: 9.7 min (Gradient B); $C_{288}H_{455}N_{81}O_{112}$ (6844.2).

References:

[1] U. K. Saha, R. R. Schmidt, *J. Chem. Soc., Perkin Trans. 1* **1997**, 1855-1860.

[2] R. J. Payne, S. Ficht, S. Tang, A. Brik, Y. Y. Yang, D. A. Case, C. H. Wong, *J. Am. Chem. Soc.* **2007**, *129*, 13527-13536.

[3] Novabiochem catalog 2006/2007.

[4] S. Mezzato, M. Schaffrath, C. Unverzagt, *Angew. Chem.* **2005***, 117,* 1677-1681; *Angew. Chem., Int. Ed.* **2005**, *44*, 1650-1654.