Electronic Supplementary Material (ESI) for Chemical Science.
This journal is © The Royal Society of Chemistry 2016

Supporting Information

Total chemical synthesis of proteins without HPLC purification

Simon F. Loibl, Ziv Harpaz, Robert Zitterbart, Oliver Seitz

Institut für Chemie, Humboldt-Universität zu Berlin, 12489 Berlin, Germany

Content

1. General information - 3 -
2. General procedures for synthesis of peptides - 3 -
   2.1 Peptide thioesters - 3 -
   2.2 Auxiliary-loaded peptides - 4 -
   2.3 Determination of peptide concentration - 5 -
3. Synthesis of C-and N-terminal fragments - 6 -
   3.1. His_6-MUC1 peptide thioesters - 6 -
   Investigation of racemization during thioester synthesis at C-terminal histidine - 9 -
   Non-chromatographic purification of peptide thioesters - 10 -
   3.2. Aux-MUC1 peptide hydrazides - 11 -
   Trifluoroacetylation of C-terminal peptide hydrazides during TFA-treatment - 14 -
4. Solid supported chemical ligation - 15 -
   4.1 Optimization of SPCL-conditions - 15 -
   4.1 Protocol for the synthesis of MUC1 proteins - 17 -
   4.2 Synthesis of MUC1 proteins by SPCL and non-chromatographic purification - 18 -
   4.3 Reactivity-based purification as an alternative to chromatographic purification - 23 -
   4.4 Stability of Ni-His_6-complex and Ni-NTA resin during ligation and auxiliary removal - 24 -
   4.5 Acidic elution of MUC1 proteins - 25 -
5. MALDI-MS analysis - 26 -
6. References - 26 -

1. General information

Materials and instruments

Commercially available compounds were used without further purification. Chemicals were acquired from Alfa Aesar (Karlsruhe, Germany), Sigma Aldrich (St. Louis, USA) and VWR Chemicals (Darmstadt, Germany). Coupling regents were purchased from Carl Roth (Karlsruhe, Germany) and Luxembourg Bio Technologies Ltd. **(**Ness Ziona, Israel). Amino acid building blocks were acquired from Carbolution Chemicals GmbH (Saarbrücken, Germany), Iris Biotech GmbH (Marktredwitz, Germany) and Novabiochem (Darmstadt, Germany). Resins for SPPS or catch&release purifications were purchased from Rapp Polymere GmbH (Tübingen, Germany), Cube Biotech (Monheim, Germany) and ABT (Madrid, Spain). Dimethylformamide (synthesis grade) for peptide synthesis was acquired from Biosolve (Valkenswaard, Netherland). Millipore water was obtained by using a Milli-Q Ultra Pure Water Purification system from membraPure (Henningsdorf, Germany). Acetonitrile (HPLC-grade) was purchased from VWR Chemicals (Darmstadt, Germany). Dry solvents were taken from a *MBraun* solvent purification system *SPS 800*.
UPLC-MS measurements were performed by using an *Acquity* system from *Waters* and a BEH130 C18 column (2.1 x 50 mm, 1.7 µm) with a binary mixture of A (0.1% TFA, 1% CH_3CN, 98.9% H_2O) and B (0.1% TFA, 1% H_2O, 98.9% CH_3CN) or CSH130 C18 column (2.1 x 50 mm, 1.7 µm) with a binary mixture of C (0.6% TFA, 1% CH_3CN, 98.9% H_2O) and D (0.6% TFA, 1% H_2O, 98.9% CH_3CN). For elution the solvents were used at a flow rate of 0.5 mL/min in a linear gradient as described and with the column oven temperature set to 50 °C.

MALDI-TOF mass spectra were recorded on a *Shimadzu Axima Confidence* in positive mode using a CHCA matrix (10 mg α-cyano-4-hydroxycinnamic acid suspended in 1 mL of H_2O/CH_3CN /TFA (1/1/0.001).

2. General procedures for synthesis of peptides

2.1 Peptide thioesters

Loading of chlorotrityl resin with Fmoc-His-OH.

The resin (~0.8 µmol/mg) was transferred into a syringe equipped with a fritted disc, allowed to swell in dry DCM (10 min) and was treated with a mixture of Fmoc-His(Trt)-OH (4 eq, c = 0.1 M in DCM) and DIPEA (20 eq). After 1 h the resin was capped with a solution of 10 % DIPEA in MeOH (50 eq base, 10 min) and subsequently washed (5 x DCM, 3x DMF). The Fmoc-group was removed by treatment of the resin with a solution of 20% piperidine in DMF (2 x 5 min). Afterwards the resin was washed (3 x DMF, 3 x DCM, 3x DMF) and the inital loading estimated by the UV-absorbance of the piperidine-fulvene adduct (λ = 301 nm, ε = 7800 $M^{-1}cm^{-1}$).

Automated solid-phase peptide synthesis was performed by using a *MultiPep* RS from *Intavis*.

- Coupling: Fmoc-protected amino acids (8 eq) were activated with HCTU/OxymaPure (8 eq each), NMM (20 eq) in DMF (final concentration = 0.3 M) and transferred to the resin. Single couplings from cycle 1 to 40 (coupling time = 15 min). Washing: 3 x 800 µL DMF.

  Double couplings from cycle 40 to 60 (coupling time = 1 x 15 min 1 x 10 min). Washing: 3 x 800 µL DMF.

  The His_6-unit was assembled by triple couplings (coupling time = 1 x 15 min, 1 x 10 min, 1 x 10 min). Washing: Washing: 3 x 800 µL DMF.

- Capping: The resin was treated with DMF/Ac_2O/lutidine (89/5/6) for 5 min. Washing: 5 x 800 µL DMF.

- Fmoc removal: The resin was treated with 20% piperidine in DMF for 1 x 8 min and 1x 6 min. Washing: 3 x 800 µL DMF.

Thioesterification and final work-up

After automated SPPS the free N-terminus was protected with the Boc-group by treatment of the resin-bound peptides with Boc_2O (~50 eq, c = 0.4 M) and DIPEA (~10 eq) in DMF for 30 min. The resin was washed (5 x DMF, 10 x DCM) and dried *in vacuo*. Cleavage of the fully protected peptide acid (~25 µmol scale) was achieved by adding a mixture of dichloromethane/2,2,2-trifluoroethanol/acetic acid (8/1/1, 10 mL) for 2 h to the resin. The resin was washed twice (2 x 4 mL) with a mixture of dichloromethane/2,2,2-trifluoroethanol/acetic acid (8/1/1). The combined solutions were poured into hexane (50 mL) and concentrated *in vacuo*. The crude product was dissolved in DMF (4 mL) and cooled to -35 °C. To the cooled solution PyBOP (5 eq), methyl 3-mercaptopropionate (30 eq) and DIPEA (5 eq) were added. After 4h the reaction was stopped by addition of TFA (20 eq). Volatiles were removed *in vacuo* and the crude peptide treated with a mixture of TFA/TIS/water (95/2.5/2.5; 8 mL) for 2 h. After 2 h the TFA-solution was concentrated to a volume of ~ 1-2 mL. Precipitation of peptides was achieved by addition of Et_2O (~8-10 fold volume) to the remaining residue. The suspension was cooled (in dry ice ~1 h), centrifuged (4000 rpm, 15 min, 4 °C) and the ether phase was decanted. The crude peptide was dissolved in a mixture of H_2O/CH_3CN /TFA (3/1/0.001) and lyophilized.

2.2 Auxiliary-loaded peptides

Pre-loading of chlorotrityl resin

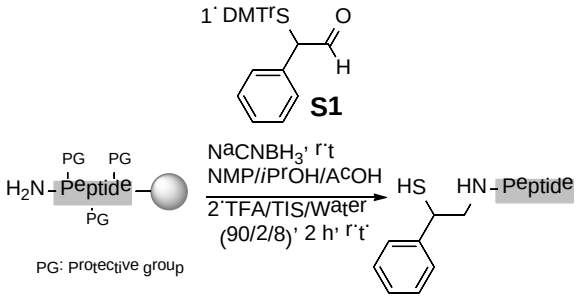
The resin (~0.8 µmol/mg) was transferred into a syringe equipped with a fritted disc, allowed to swell in DMF (10 min) and treated with a mixture of hydrazine hydrate (10 eq, c = 0.5 M in DMF) and NEt_3 (20 eq). After 1 h the resin was treated with a solution of 10 % DIPEA in MeOH (50eq base, 10 min) and subsequently washed (5 x DCM, 3x DMF). Fmoc-His(Trt)-OH (8 eq, 0.4 M in DMF) was coupled onto the hydrazine functionalized resin in presence of PyBOP (8 eq) and DIPEA (12 eq). After 45 min the resin was washed (3 x DMF, 3 x DCM, 3 x DMF) and capped with DMF/Ac_2O/lutidine (89/5/6, 5 min). The Fmoc-group was removed

by treatment of the resin with a solution of 20% piperidine in DMF (2 x 5 min). Afterwards the resin was washed (3 x DMF, 3 x DCM, 3x DMF) and the inital loading estimated by the UV-absorbance of the piperidine-fulven adduct (λ = 301 nm, ε = 7800 $M^{-1}cm^{-1}$).

Automated solid-phase peptide synthesis was performed by using a *MultiPep* RS from *Intavis*:

- Coupling: Fmoc-protected amino acids (8 eq) were activated with HCTU/OxymaPure (8 eq each), NMM (20 eq) in DMF (final concentration = 0.3 M) and transferred to the resin. Single couplings from cycle 1 to 40 (coupling time = 15 min). Washing: 3 x 800 µL DMF. Double couplings from cycle 40 to 60 (coupling time = 1 x 15 min 1 x 10 min). Washing: 3 x 800 µL DMF.
- Capping: The resin was treated with DMF/Ac_2O/lutidine (89/5/6) for 5 min. Washing: 5 x 800 µL DMF
- Fmoc-removal: The resin was treated with 20% piperidine in DMF for 1 x 8 min and 1x 6 min. Washing: 3 x 800 µL DMF.

Introduction of auxiliary onto peptide during SPPS and final work-up



The peptidyl-resin was allowed to swell in NMP for 5 min and then treated with a mixture of aldehyde **S1** (prepared as described previously[1]) and $NaCNBH_3$ (15 eq each, c = 0.4 M) in NMP/*i*PrOH (3/1) with 5% AcOH for 16 h. The resin was washed (3 x DMF, 10 x DCM), dried under vacuum and treated with a mixture (8 mL) of TFA/TIS/Water (90/2/8). After 2 h the resin was filteredand washed with TFA (3x 1 mL). The combined filtrates were concentrated to a volume of ~ 1-2 mL. Precipitation of peptides was achieved by addition of Et_2O (~8-10-fold volume) to the remaining residue. The suspension was cooled (in dry ice ~1 h), centrifuged (4000 rpm, 15 min, 4 °C) and the ether phase was decanted. The crude peptide was dissolved in a mixture of H_2O/CH_3CN /TFA (3/1/0.001) and lyophilized.

2.3 Determination of peptide concentration

Concentrations of peptides were estimated by weighing the crude material obtained after lyophilization and dissolving it afterwards in an appropriate volume. Arg-and His-residues as well as the free N-terminus were assumed to be present as their corresponding TFA-salts.

3. Synthesis of C-and N-terminal fragments

3.1. His_6-MUC1 peptide thioesters

Crude 26mer His_6-MUC1 peptide thioester **1a**

The peptide thioester **1a** was synthesized on a preloaded chlorotrityl-resin (loading 25.0 µmol) (see 2.1). After SPPS, the fully protected peptide acid was converted into the corresponding thioester and the progress of the reaction was monitored by UPLC-MS (see Fig S1b). TFA-treatment, ether precipitation and final lyophilization gave the crude peptide thioester **1a** as a white solid (96.7 mg, 25.2 µmol based on **1a** as 9*TFA-salt).

UPLC-MS: t_R = 2.3 min (3-40% B in 4 min, BEH-column); m/z = 938.2 ($C_{120}H_{178}N_{43}O_{35}S$ $(M+3H)^{3+}$, calcd.: 938.1 (100%)), 704.1 ($C_{120}H_{179}N_{43}O_{35}S$ $(M+4H)^{4+}$, calcd.: 703.8 (100%)), 563.5 ($C_{120}H_{180}N_{43}O_{35}S$ $(M+5H)^{5+}$, calcd.: 563.3 (100%)); $C_{120}H_{175}N_{43}O_{35}S$ (MW = 2812.0 $g \cdot mol^{-1}$).

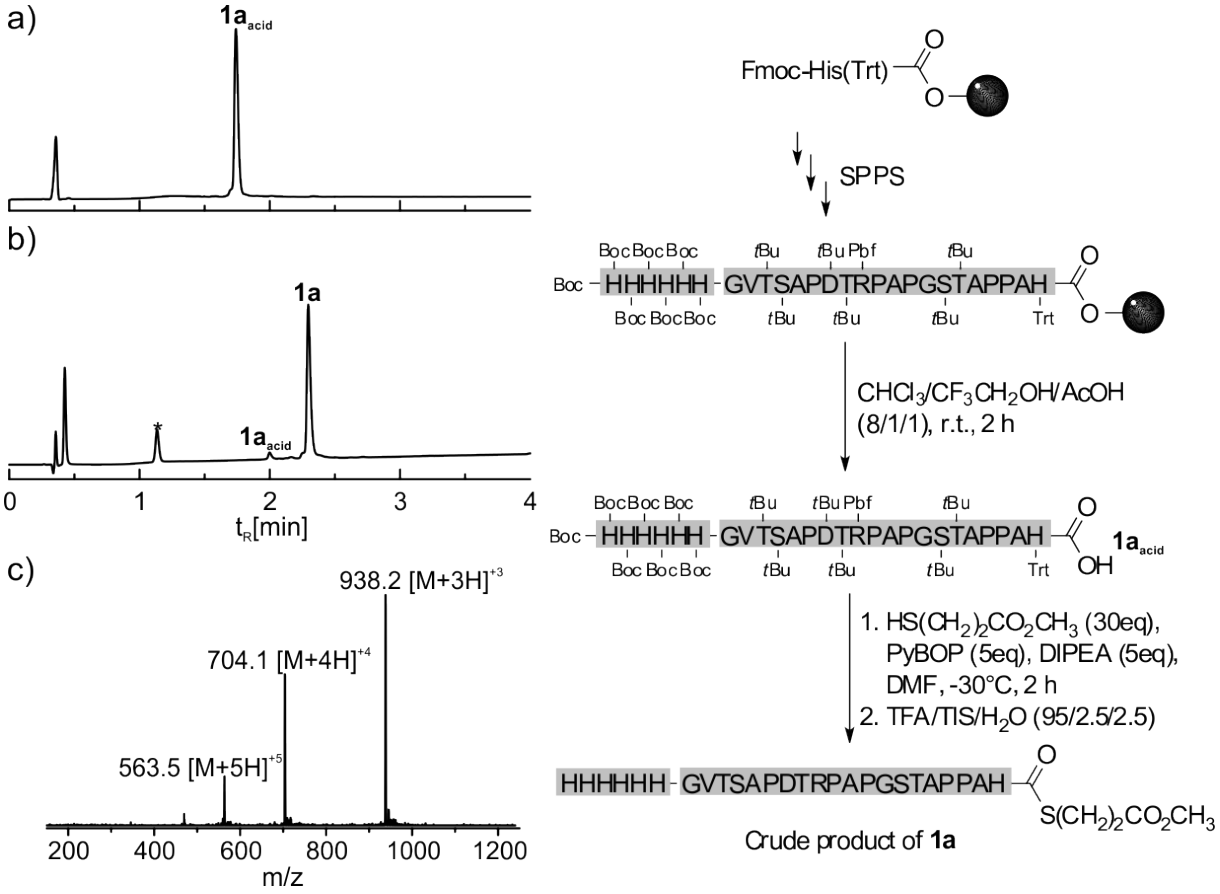


Figure S1: UPLC traces of cleaved peptides before (a) and after (b) in-solution thioesterification. ESI-MS analysis (c) of crude peptide thioester **1a**.

Crude 46mer His_6-MUC1 peptide thioester **1b**

The peptide thioester **1b** was synthesized on a preloaded chlorotrityl-resin (loading 25.0 µmol) via Fmoc-strategy (see 2.1). After SPPS, the fully protected peptide acid was converted into the corresponding thioester and the progress of the reaction was monitored by UPLC-MS (see Fig S2b). TFA-treatment, ether precipitation and final lyophilization gave the crude peptide thioester **1b** as a white solid (135 mg, 22.8 µmol based on **1b** as 11*TFA-salt).

UPLC-MS: t_R = 2.4 min (3-40% B in 4 min, BEH-column); m/z = 1171.1 ($C_{200}H_{304}N_{68}O_{62}S$ $(M+4H)^{4+}$, calcd.: 1171.3 (100%)), 937.1 ($C_{200}H_{305}N_{68}O_{62}S$ $(M+5H)^{5+}$, calcd.: 937.3 (100%)), 781.4 ($C_{200}H_{306}N_{68}O_{62}S$ $(M+6H)^{6+}$, calcd.: 781.2 (100%)), 669.8 ($C_{200}H_{307}N_{68}O_{62}S$ $(M+7H)^{7+}$, calcd.: 669.8 (100%)), 586.0 ($C_{200}H_{308}N_{68}O_{62}S$ $(M+8H)^{8+}$, calcd.: 586.2 (100%)); $C_{200}H_{300}N_{68}O_{62}S$ (MW = 4681.0 g·mol^{-1}).

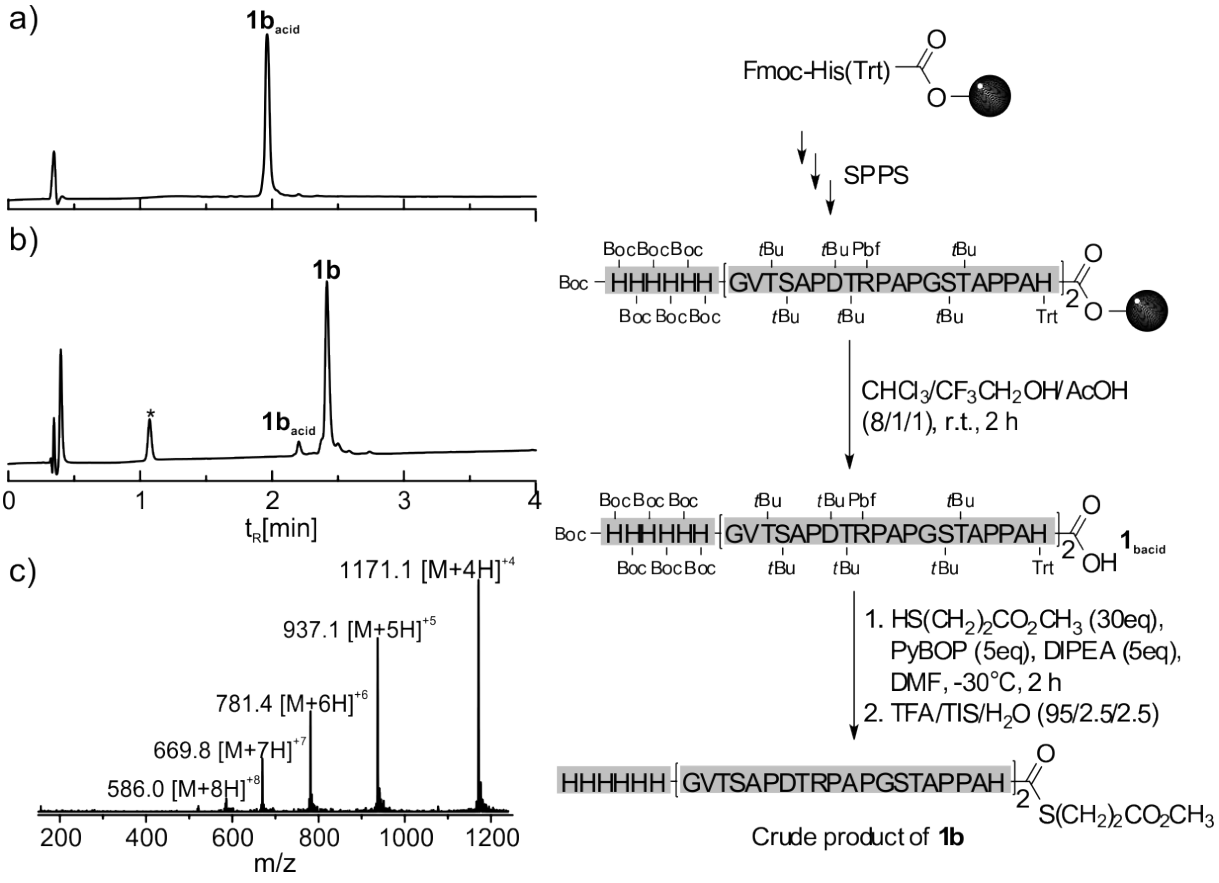


Figure S2: UPLC traces of cleaved peptides before (a) and after (b) in-solution thioesterification. ESI-MS spectrum (c) of crude peptide thioester **1b**.

Crude 66mer His_6-MUC1 peptide thioester **1c**

The peptide thioester **1c** was synthesized on a preloaded chlorotrityl-resin (loading 25.0 µmol) via Fmoc-strategy (see 2.1). After SPPS, the fully protected peptide acid was converted into the corresponding thioester and the progress of the reaction was monitored by UPLC-MS (see Fig S2b). TFA-treatment, ether precipitation and final lyophilization gave the crude peptide thioester **1c** as a white solid (136 mg, 22.3 µmol based on **1c** as 11*TFA-salt).

UPLC-MS: t_R = 2.5 min (3-40% B in 4 min, BEH-column); m/z = 1092.5 ($C_{280}H_{431}N_{93}O_{89}S$ $(M+6H)^{6+}$, calcd.: 1092.7 (100%)), 936.7 ($C_{280}H_{432}N_{93}O_{89}S$ $(M+7H)^{7+}$, calcd.: 936.7 (100%)), 819.9 ($C_{280}H_{433}N_{93}O_{89}S$ $(M+6H)^{6+}$, calcd.: 819.8 (100%)), 729.1 ($C_{280}H_{434}N_{93}O_{89}S$ $(M+9H)^{9+}$, calcd.: 728.8 (100%)); $C_{280}H_{425}N_{93}O_{89}S$ (MW = 6550.01g·mol^{-1}).

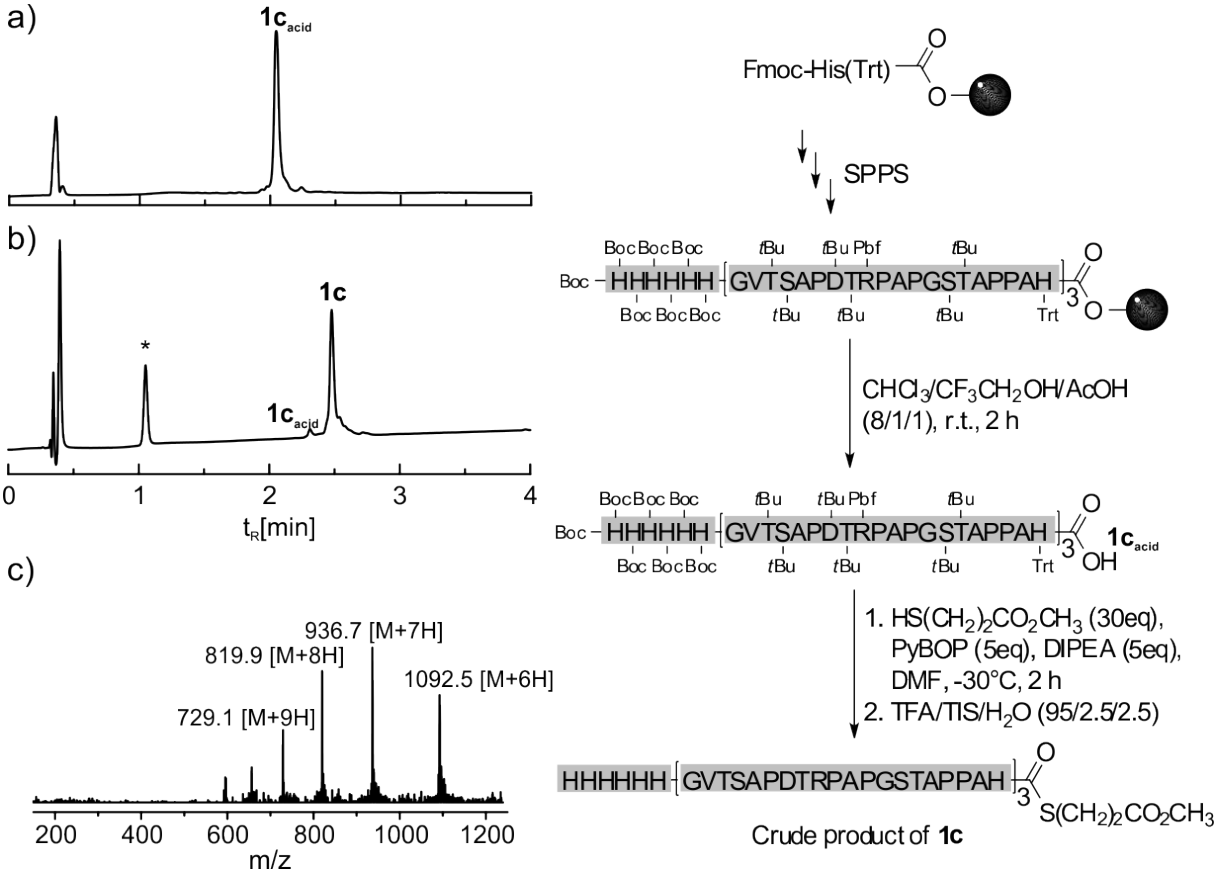


Figure S3: UPLC traces of cleaved peptides before (a) and after (b) in-solution thioesterification. ESI-MS spectrum (c) of crude peptide thioester **1c**.

Investigation of racemization during thioester synthesis at C-terminal histidine

The peptide thioesters **S2-L** and **S2-D** were synthesized on Fmoc-L-His(Trt) and Fmoc-D-His(Trt)-preloaded chlorotrityl-resins (loading 15.0 µmol) via Fmoc-strategy (see 2.2). After SPPS, the fully protected peptide acids were converted into the corresponding thioesters and the progress of the reaction was monitored by UPLC-MS (see Fig S2b). TFA-treatment, ether precipitation and final lyophilization gave the crude peptide thioester **S2-L** and **S2-D** as white solids. UPLC analysis of the crude products confirmed that in-solution thioesterfication proceeded without racemization of the C-terminal His-residue.

S2-L: UPLC-MS: t_R = 6.4 min (3-8% B in 10 min, CSH-column); m/z = 695.5 ($C_{30}H_{47}N_8O_9S$ $(M+H)^+$, calcd.: 695.3), 348.5 ($C_{30}H_{48}N_8O_9S$ $(M+2H)^{2+}$, calcd.: 348.2); $C_{30}H_{46}N_8O_9S$ (MW = 694.8 g·mol^{-1}).

S2-D: UPLC-MS: t_R = 6.1 min (3-8% B in 10 min, CSH-column); m/z = 695.5 ($C_{30}H_{47}N_8O_9S$ $(M+H)^+$, calcd.: 695.3), 348.5 ($C_{30}H_{48}N_8O_9S$ $(M+2H)^{2+}$, calcd.: 348.2); $C_{30}H_{46}N_8O_9S$ (MW = 694.8 g·mol^{-1}).

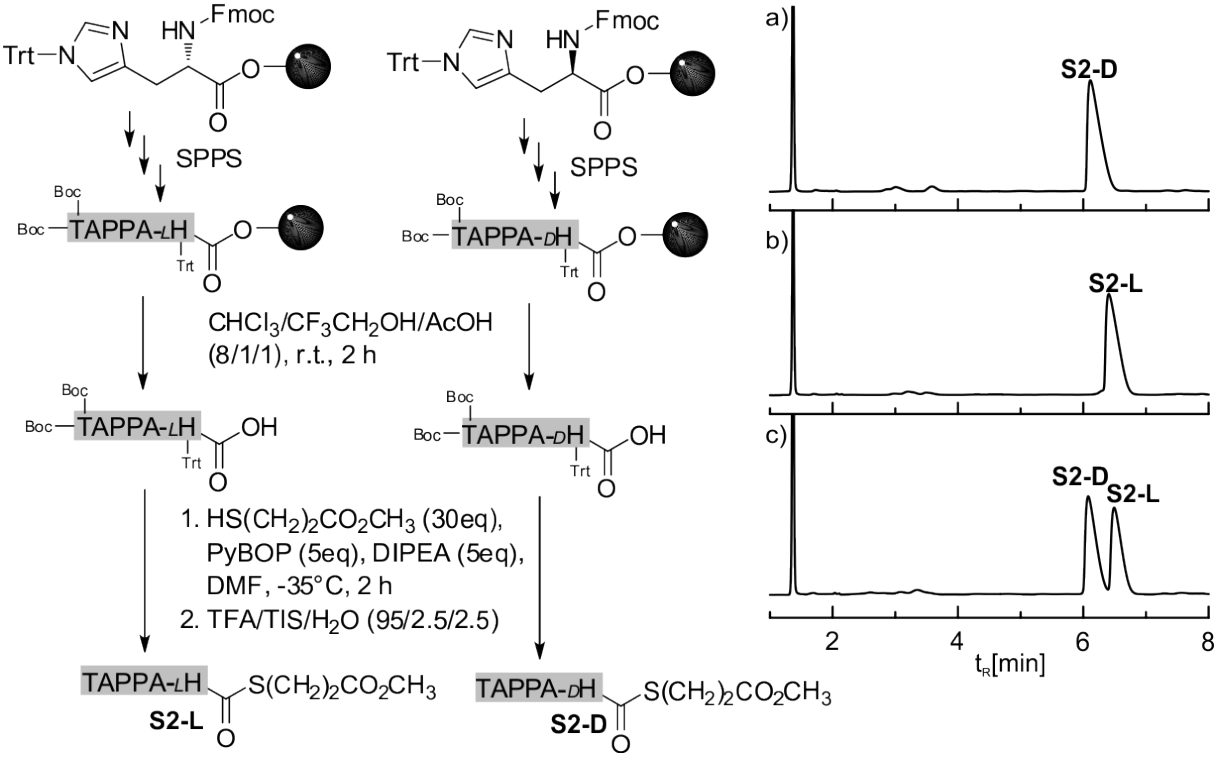


Figure S4: UPLC traces of crude materials obtained after SPPS, in-solution thioesterification and final TFA-treatment of peptide thioester **S2-D** (a), **S2-L** (b) and co-injection (c) of both samples.

Non-chromatographic purification of peptide thioesters

50 µL of Ni-NTA resin (50% suspension, *Cube Biotech*, *Germany*) were transferred into a reaction reactor equipped with a filter frit. The solution was removed by filtration and the resin allowed to swell in 200 µL of aqueous buffer (6M GuHCl, 200 PBS, pH 7.5) for 2 min. Lyophilized His_6-MUC1-peptid thioesters **17₁**, **17₂** and **17₃** (approx. 200 nmol) were dissolved in 100 µL of aqueous buffer (6M GuHCl, 200 PBS, pH 7.5) and added to the resin. After immobilization (~15min) the buffer was removed by filtration and the resin was washed (5 x 200 µL of aqueous buffer (6M GuHCl, 200 PBS, pH 7.5, then 5 x 200 µL water). The resin was treated for 10 min with 200 µL of 0.1% aqueous TFA. The solution was filtered and the eluate subjected to UPLC-MS analysis. The results depicted in scheme S5 show that non-peptidic and peptidic impurities are easily removed.

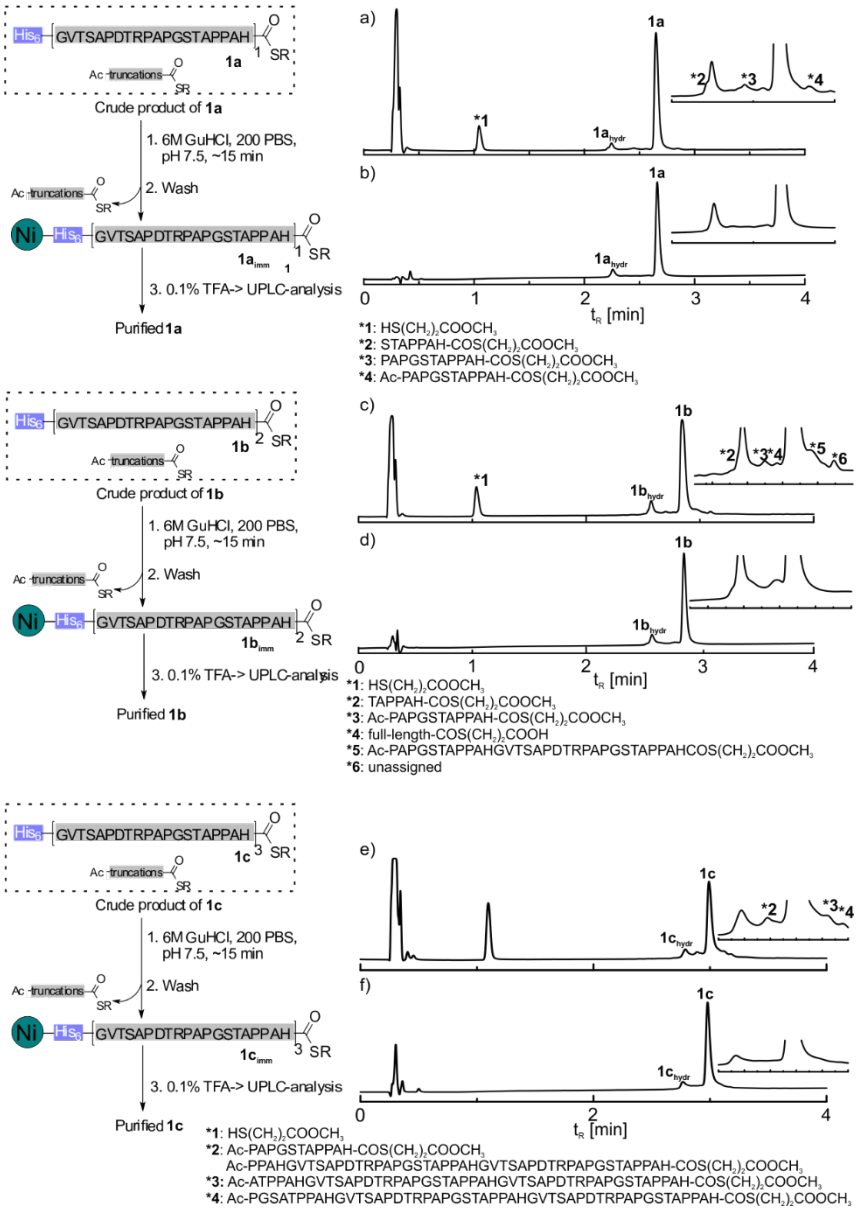


Figure S5: Non-chromatographic purification of His_6-MUC1-peptid thioesters **1a**, **1b** and **1c**. UPLC traces of crude materials (a,c,e) obtained after SPPS, in-solution thioesterification and final TFA-treatment and after non-chromatographic purification (b,d,f).

3.2. Aux-MUC1 peptide hydrazides

Crude 20mer Aux-MUC1 peptide hydrazide **2a**

The auxiliary-armed peptide **2a** was synthesized on a Fmoc-His(Trt)-hydrazine-loaded trityl-resin (loading: 25.0 µmol) as described in 2.2. After SPPS the auxiliary was introduced by reductive amination and the conversion was checked by UPLC-MS (see Fig S6b). Cleavage of the peptide from the resin, ether precipitation and final lyophilization gave the crude peptide **2a** as a white solid (44.7 mg, 18.8 µmol based on **2a** as 3*TFA-salt).

UPLC-MS: t_R = 2.5 min (3-40% B in 4 min, BEH-column); m/z = 1019.4 ($C_{88}H_{139}N_{27}O_{27}S$ $(M+2H)^{2+}$, calcd.: 1019.5 (100%)), 680.1 ($C_{88}H_{140}N_{27}O_{27}S$ $(M+3H)^{3+}$, calcd.: 680.0 (100%)), 510.5 ($C_{88}H_{141}N_{27}O_{27}S$ $(M+4H)^{4+}$, calcd.: 510.3 (100%)), 408.5 ($C_{88}H_{142}N_{27}O_{27}S$ $(M+5H)^{5+}$, calcd.: 408.4 (100%)); $C_{88}H_{137}N_{27}O_{27}S$ (MW = 2037.3 g·mol^{-1}).

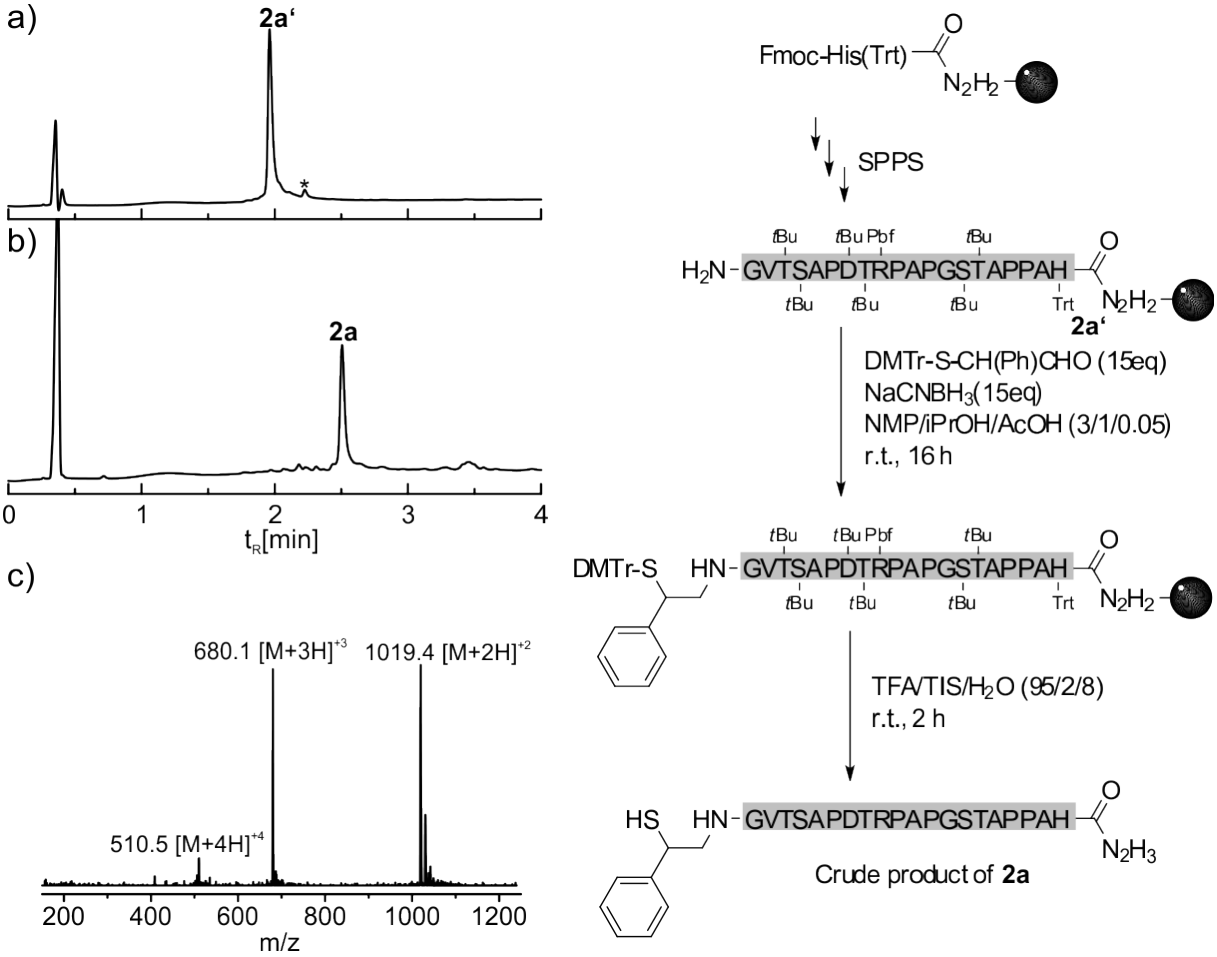


Figure S6: UPLC traces of liberated peptides before (a) and after (b) reductive amination. ESI-MS spectrum (c) of crude auxiliary peptide **2a**.

Crude 40mer Aux-MUC1 peptide hydrazide **2b**

The auxiliary-armed peptide **2b** was synthesized on a Fmoc-His(Trt)-hydrazine-loaded trityl-resin (loading: 25.0 µmol) as described in 2.2. After SPPS the auxiliary was introduced by reductive amination and the conversion was checked by UPLC-MS (see Fig S2b). Cleavage of the peptide from the resin, ether precipitation and final lyophilization gave the crude peptide as a white solid (68.5 mg, 15.3 µmol based on **2b** as 5*TFA-salt).

UPLC-MS: t_R = 2.6 min (3-40% B in 4 min, BEH-column); m/z = 977.6 ($C_{168}H_{266}N_{52}O_{54}S$ $(M+4H)^{4+}$, calcd.: 977.5 (100%)), 782.3 ($C_{168}H_{267}N_{52}O_{54}S$ $(M+5H)^{5+}$, calcd.: 782.2 (100%)), 652.4 ($C_{168}H_{268}N_{52}O_{54}S$ $(M+6H)^{6+}$, calcd.: 652.0 (100%)); $C_{168}H_{262}N_{52}O_{54}S$ (MW = 3906.3 $g \cdot mol^{-1}$).

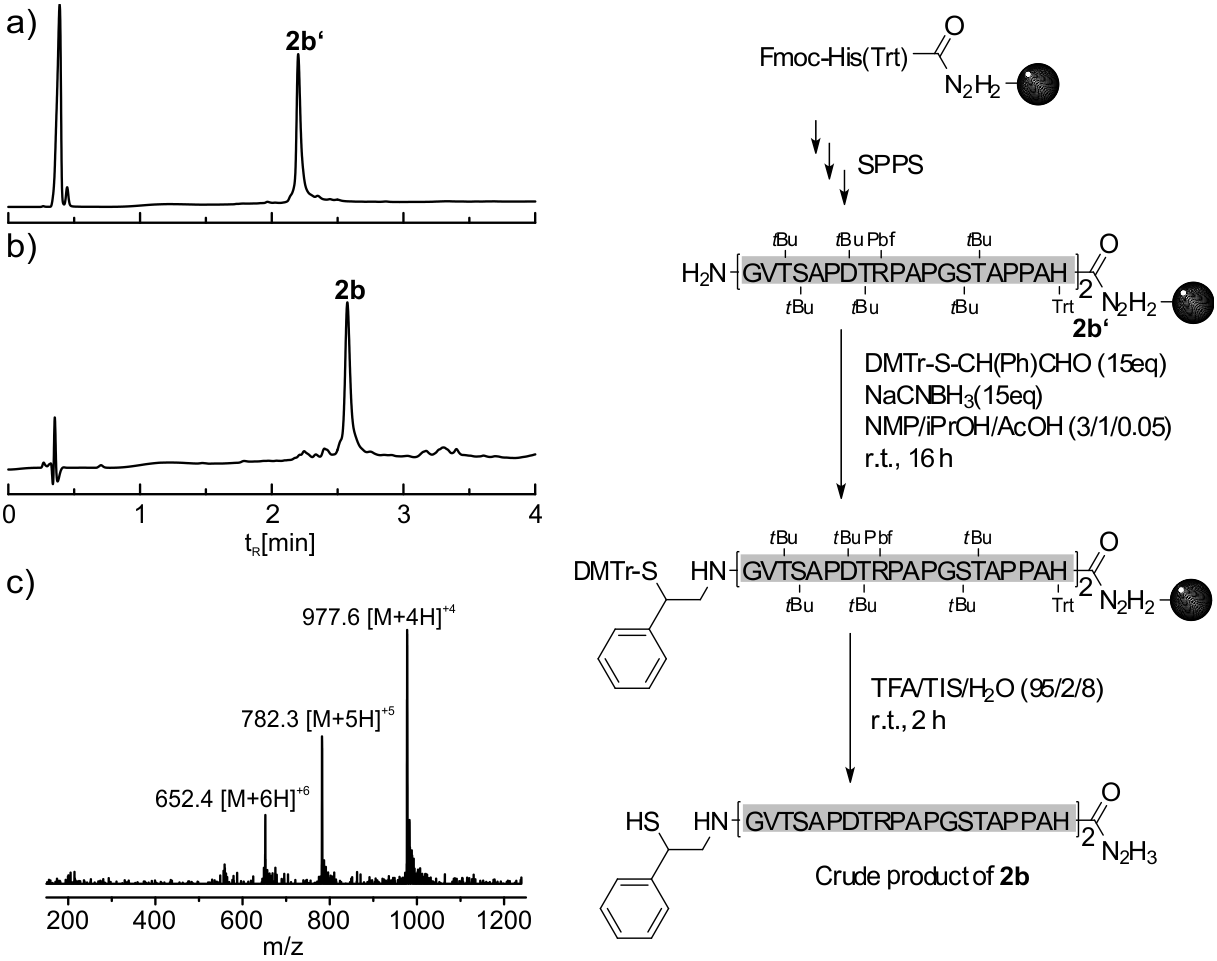


Figure S7: UPLC traces of cleaved peptides before (a) and after (b) reductive amination. ESI-MS spectrum (c) of crude auxiliary peptide **2b**.

Crude 60mer Aux-MUC1 peptide hydrazide **2c**

The auxiliary-armed peptide **2c** was synthesized on a Fmoc-His(Trt)-hydrazine-loaded trityl-resin (loading: 25.0 µmol) as described in 2.2. After SPPS the auxiliary was introduced by reductive amination and the conversion was checked by UPLC-MS (see Fig S2b). Cleavage of the peptide from the resin, ether precipitation and final lyophilization gave the crude peptide as a white solid (112 mg, 17.1 µmol based on **2c** as 7*TFA-salt).

UPLC-MS: t_R = 2.7 min (3-40% B in 4 min, BEH-column); m/z = 1156.1 ($C_{248}H_{392}N_{77}O_{81}S$ $(M+5H)^{5+}$, calcd.: 1155.8 (100%)), 963.9 ($C_{248}H_{393}N_{77}O_{81}S$ $(M+6H)^{6+}$, calcd.: 963.3 (100%)), 826.2 ($C_{248}H_{394}N_{77}O_{81}S$ $(M+7H)^{7+}$, calcd.: 825.8 (100%)), 723.2 ($C_{248}H_{395}N_{77}O_{81}S$ $(M+8H)^{8+}$, calcd.: 722.7 (100%)), 642.8 ($C_{248}H_{396}N_{77}O_{81}S$ $(M+9H)^{9+}$, calcd.: 642.5 (100%)); $C_{248}H_{387}N_{77}O_{81}S$ (MW = 5775.3 g·mol^{-1}).

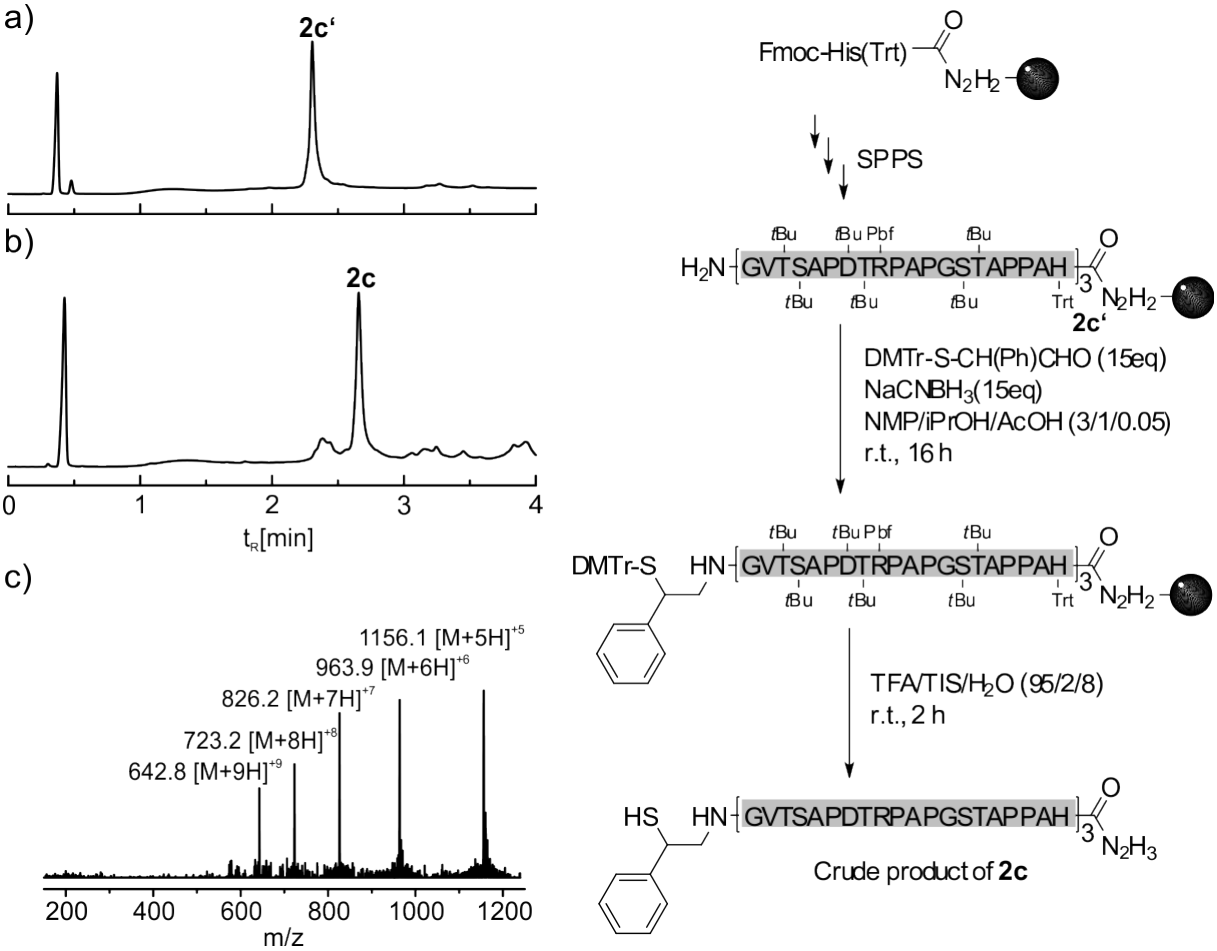


Figure S8: UPLC traces of cleaved peptides before (a) and after (b) reductive amination. ESI-MS spectrum (c) of crude auxiliary peptide **2c**.

Trifluoroacetylation of C-terminal peptide hydrazides during TFA-treatment

The peptide hydrazide **S3-N_2H_3** (without auxiliary) was synthesized as described in 2.2. For comparison, the peptide acid **S3-OH** was synthesized on a histidine-loaded trityl-resin as described in 2.1. After SPPS, cleavage of the peptides from the resin was performed with TFA/TIS (95/5) or TFA/TIS/Water (90/2/8) at room temperature for 1 h. The crude materials obtained after ether precipitation were analyzed by UPLC-MS. The peptide hydrazide **S3-N_2H_3** suffered from trifluoroacetylation when the cleavage was performed with TFA/TIS (95/5) (Fig. S9b). By comparison, trifluoroacatylation did not occur with the peptide acid **S3-OH** (Fid. S9c). The amount of trifluoroacetylated by-product **S3-TFA** was reduced when high amounts (8 vol%) of water were included during TFA-treatment (Fig. S9a).

S3-N_2H_3: UPLC-MS: t_R = 1.9 min (3-40% B in 6 min, BEH-column); m/z = 951.8 ($C_{80}H_{131}N_{27}O_{27}$ $(M+2H)^{2+}$, calcd.: 951.0 (100%)), 634.9 ($C_{80}H_{132}N_{27}O_{27}$ $(M+3H)^{3+}$, calcd.: 634.3 (100%)), 476.4 ($C_{80}H_{133}N_{27}O_{27}$ $(M+4H)^{4+}$, calcd.: 476.0 (100%)); $C_{80}H_{129}N_{27}O_{27}$ (MW = 1901.0 g·mol^{-1}).

S3-TFA: UPLC-MS: t_R = 2.2 min (3-40% B in 6 min, BEH-column); m/z = 999.6 ($C_{82}H_{130}F_3N_{27}O_{28}$ $(M+2H)^{2+}$, calcd.: 999.0 (100%)), 666.7 ($C_{82}H_{131}F_3N_{27}O_{28}$ $(M+3H)^{3+}$, calcd.: 666.3 (100%)), 500.3 ($C_{82}H_{134}F_3N_{27}O_{28}$ $(M+4H)^{4+}$, calcd.: 500.0 (100%)); $C_{82}H_{128}F_3N_{27}O_{28}$ (MW = 1997.1 g·mol^{-1}).

S3-OH: UPLC-MS: t_R = 1.9 min (3-40% B in 6 min, BEH-column); m/z = 944.5 ($C_{80}H_{129}N_{25}O_{28}$ $(M+2H)^{2+}$, calcd.: 944.0 (100%)), 630.0 ($C_{80}H_{130}N_{25}O_{28}$ $(M+3H)^{3+}$, calcd.: 629.7 (100%)), 472.9 ($C_{80}H_{131}N_{25}O_{28}$ $(M+4H)^{4+}$, calcd.: 472.5 (100%)); $C_{80}H_{127}N_{25}O_{28}$ (MW = 1887.0 g·mol^{-1}).

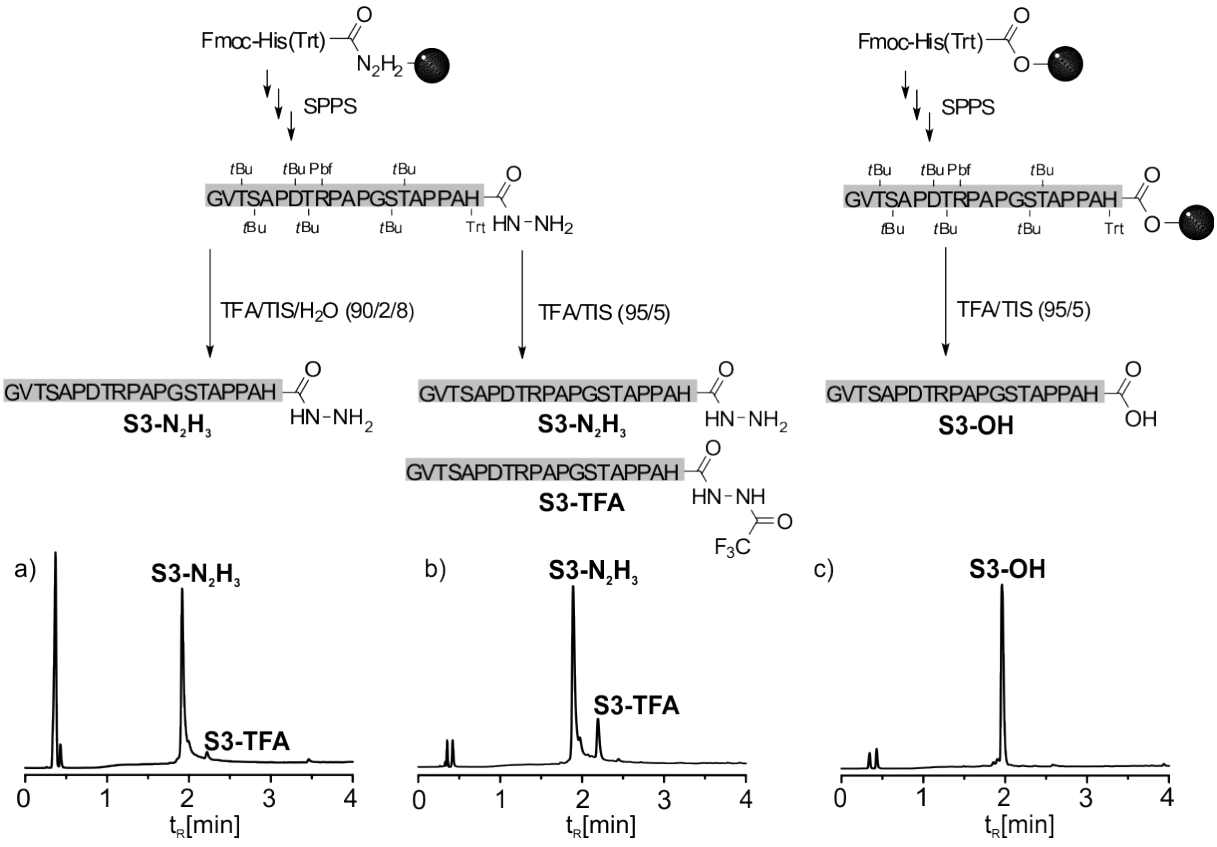


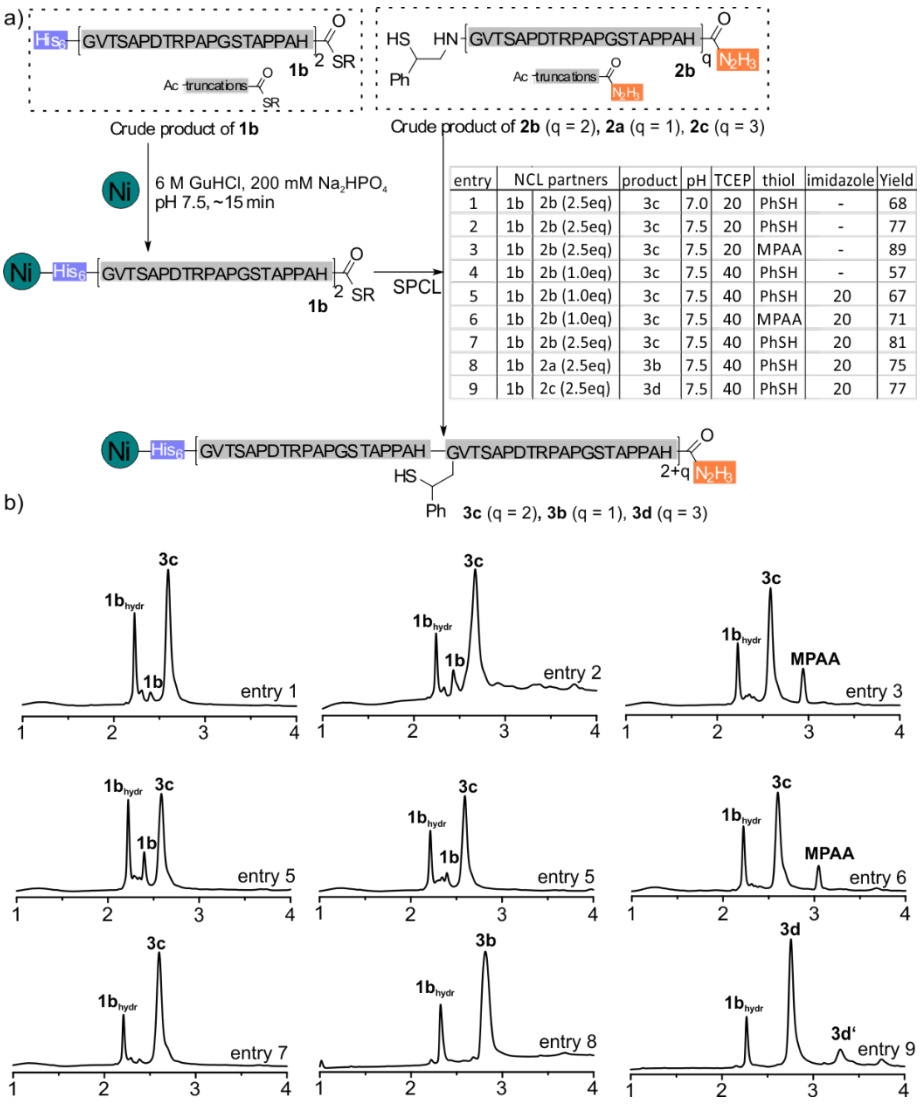
Figure S9: UPLC traces of cleaved peptide hydrazide **S3-N_2H_3** in presence (a) and absence (b) of water during TFA-treatment. UPLC trace of cleaved peptide acid **S3-OH** after TFA-treatment.

4. Solid supported chemical ligation

4.1 Optimization of SPCL-conditions

200 µL of Ni-NTA resin (50% suspension, *Cube Biotech, Germany*) was transferred into a reaction reactor equipped with a filter frit. The solution was removed by filtration and the resin allowed to swell in 200 µL aqueous buffer (6 M GuHCl, 200 mM Na_2HPO_4, pH 7.5) for 2 min. After filtration crude 40mer His_6-MUC1peptide thioester **1b** (approx. 200 nmol) in 100 µL aqueous buffer (6M GuHCl, 200 PBS, pH 7.5) was added to the resin. After 15 min the buffer was separated from the resin by filtration and the resin was washed (5 x 200 µL of 6 M GuHCl, 200 mM Na_2HPO_4, pH 7.5). In the next step, auxiliary-mediated SPCL was performed by adding crude 40mer Aux-MUC1 peptide hydrazide **2b** (approx. 1 or 2.5 eq) in different ligation buffers (see table). The syringe reactor was flushed with argon, closed and vigorously shaken for 24 h. The solution was removed by filtration and the resin was washed (5 x 200 µL of aqueous buffer (6 M GuHCl, 200 mM Na_2HPO_4, pH 7.5), 3 x 200 µL water). To assess the conversion by UPLC analysis, a small amount of resin (~2 mg) was treated for 10 min with an aqueous solution of 100 mM imidazole (pH 8.5). The filtrate was acidified with 0.1% aqueous TFA, centrifuged and subjected to UPLC analysis. The yield of the reaction was calculated by peak-integration of the UPLC profile (λ =210 nm) using molar extinction coefficients according to Gruppen *et al.*. [2]

86mer ligation product $\mathbf{20_{22}}$: UPLC-MS: t_R = 2.6 min (3-40% B in 4 min, BEH-column); m/z = 1210.6 ($C_{364}H_{561}N_{120}O_{114}S$ $(M+7H)^{7+}$, calcd.: 1210.3 (100%)), 1059.4 ($C_{364}H_{562}N_{120}O_{114}S$ $(M+8H)^{8+}$, calcd.: 1059.1 (100%)), 942.0 ($C_{364}H_{564}N_{120}O_{114}S$ $(M+9H)^{9+}$, calcd.: 941.6 (100%)), 847.8 ($C_{364}H_{565}N_{120}O_{114}S$ $(M+10H)^{10+}$, calcd.: 847.5 (100%)), 771.2 ($C_{364}H_{566}N_{120}O_{114}S$ $(M+11H)^{11+}$, calcd.: 770.6 (100%)); $C_{364}H_{554}N_{120}O_{114}S$ (MW = 8467.1 $g \cdot mol^{-1}$).



entry	NCL partners		product	pH	TCEP	thiol	imidazole	Yield
1	1b	2b (2.5eq)	3c	7.0	20	PhSH	-	68
2	1b	2b (2.5eq)	3c	7.5	20	PhSH	-	77
3	1b	2b (2.5eq)	3c	7.5	20	MPAA	-	89
4	1b	2b (1.0eq)	3c	7.5	40	PhSH	-	57
5	1b	2b (1.0eq)	3c	7.5	40	PhSH	20	67
6	1b	2b (1.0eq)	3c	7.5	40	MPAA	20	71
7	1b	2b (2.5eq)	3c	7.5	40	PhSH	20	81
8	1b	2a (2.5eq)	3b	7.5	40	PhSH	20	75
9	1b	2c (2.5eq)	3d	7.5	40	PhSH	20	77

Figure S10: a) SPCL between immobilized 46mer His_6-MUC1 peptide thioester **1b** and auxiliary peptides **2a**, **2b** and **2c**. b) UPLC analysis of filtrate obtained upon imidazole treatment after SPCL. Conditions for SPCL: 6M GuHCl, 200mM Na_2HPO_4 , r.t. 24 h ; a: [mM]; b: 2vol% PhSH and 50 mM MPAA; c: Yield determined by peak-integration of UPLC profile (λ =210 nm). **3d'**: Thiophenol-disulfide of **3d**.

4.1 Protocol for the synthesis of MUC1 proteins

Immobilization onto Ni-NTA resin: 200 µL of Ni-NTA resin (50% suspension, *Cube Biotech, Germany*) was transferred into reaction reactor equipped with a filter frit. The solution was removed by filtration and the resin allowed to swell in 100 µL aqueous buffer (6 M GuHCl, 200 mM Na_2HPO_4, pH 7.5) for 2 min. After filtration a solution of His_6-peptide thioesters **1a**, **1b** or **1c** (approx. 500 nmol) in 100 µL aqueous buffer (6 M GuHCl, 200 mM Na_2HPO_4, pH 7.5) was added to the resin and the mixture was agitated for 15 min. The buffer was removed by filtration and the resin was washed (5 x 200 µL aqueous buffer (6 M GuHCl, 200 mM Na_2HPO_4, pH 7.5)).

SPCL: A solution (100 µL) of the auxiliary-armed peptide hydrazide **2a**, **2b** or **2c** (approx 2.5 eq) in ligation buffer (20 mM TCEP, 20 mM imidazole, 6 M GuHCl, 200 mM Na_2HPO_4, 2 vol% PhSH, pH 7.5) was added to the resin-bound peptide thioester. The syringe reactor was flushed with argon, closed and vigorously shaken for 24 h. The solution was removed by filtration and the resin was washed (5 x 200 µL aqueous buffer (6 M GuHCl, 200 mM Na_2HPO_4, pH 7.5), 3 x 200 µL water).

Auxiliary removal: The resin-bound ligation product was treated with 1000 µL of an aqueous solution (pH 8.5) of TCEP (200 mM) and morpholine (800 mM) at 40°C. After 24 h the cleavage solution was collected by filtration and acidified with 1M HCl to pH 7.0. To enable reimmobilisation of partially detached His_6-tagged proteins, the Ni-NTA resin remaining after filtration was added to the pH-adjusted solution (see also Fig. S17). After 15-30 min the solution was removed by filtration and the resin was washed (5 x 200 µl water). *Note: Alternatively auxiliary removal can be achieved at room temperature within 30 h by treatment of the resin-bound proteins with aqueous solutions of TCEP (400 mM) and morpholine (1.6 M).*

Acidic elution from Ni-NTA resin: The MUC1 protein obtained after auxiliary cleavage was released by treatment of the Ni-NTA resins with aqueous 0.25 M acetic acid (8 x 500 µL, 8 x 5 min). The acidic elution solutions were combined. *Note: Alternatively, release of protein can be achieved by treatment with 2.0 M aqueous buffer of AcOH/NaOAc (pH 3.6; 4 x 500µL, 4 x 10 min).*

Immobilization onto aldehyde-agarose resin: The combined acidic solutions were added to an aldehyde-agarose resin (purchased from *ABT, Spain*). After 30 min the acidic solution was separated by filtration and the resin washed (10 x 200 µL aqueous buffer (6 M GuHCl, 200 mM Na_2HPO_4, pH 7.5), 5 x 200 of water).

Final release of MUC1 proteins: The resin was treated with an aqueous solution of 0.5 vol% hydrazine hydrate (3 x 500 µL, 3 x 5 min). TFA (15 µL) was added to the combined eluates. Lyophilization afforded the proteins as white solids. *Note: Final release can also achieved by treatment of the resin with 1% aqueous TFA (4 x 1 mL, 4 x 30 min, see Fig. S18).*

4.2 Synthesis of MUC1 proteins by SPCL and non-chromatographic purification

1b + **2b** → **4c** (86mer MUC1 protein, 1.04 mg, 104 nmol (21 %, based on initial loading of SPPS–resin and **4c** as 15*TFA-salt)

UPLC-MS: t_R = 2.4 min (3-40% B in 4 min, BEH-column); m/z = 1191.1 ($C_{356}H_{553}N_{120}O_{114}$ $(M+7H)^{7+}$, calcd.: 1190.9 (100%)), 1042.7 ($C_{356}H_{554}N_{120}O_{114}$ $(M+8H)^{8+}$, calcd.: 1042.1 (100%)), 926.9 ($C_{356}H_{555}N_{120}O_{114}$ $(M+9H)^{9+}$, calcd.: 926.5 (100%)), 834.4 ($C_{356}H_{556}N_{120}O_{114}$ $(M+10H)^{10+}$, calcd.: 833.9 (100%)), 759.7 ($C_{356}H_{557}N_{120}O_{114}$ $(M+11H)^{11+}$, calcd.: 758.2 (100%)), 695.5 ($C_{356}H_{558}N_{120}O_{114}$ $(M+12H)^{12+}$, calcd.: 695.1 (100%)), 642.2 ($C_{356}H_{559}N_{120}O_{114}$ $(M+13H)^{13+}$, calcd.: 641.7 (100%)); MALDI-MS: 8331.6 ($C_{356}H_{547}N_{120}O_{114}$ $(M+1H)^{1+}$, calc.: 8332.0), 4167.0 ($C_{356}H_{548}N_{120}O_{114}$ $(M+2H)^{2+}$, calc.: 4166.5); $C_{356}H_{546}N_{120}O_{114}$ (MW = 8331.0 g·mol^{-1}).

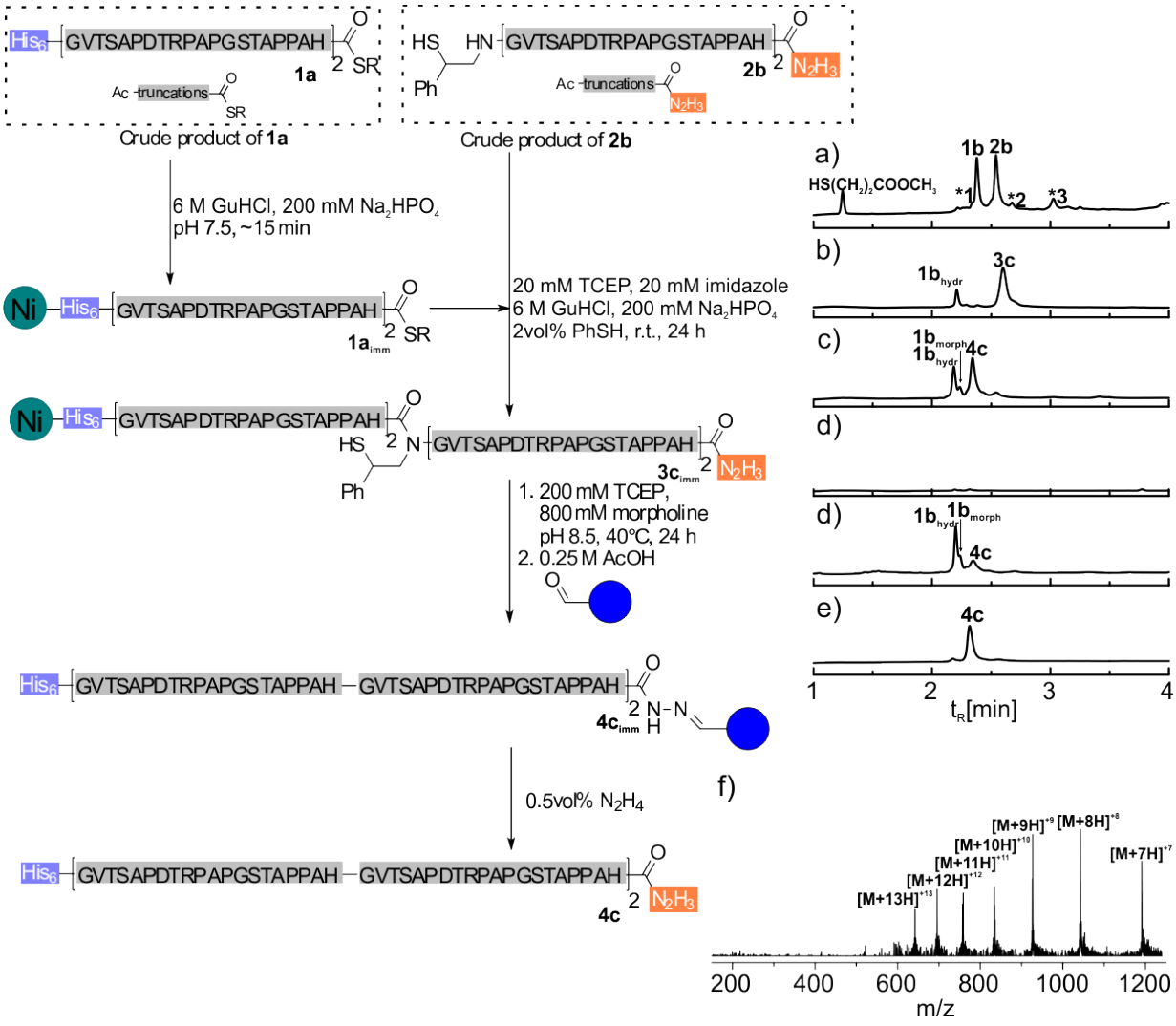


Figure S11: SPCL between immobilized 46mer His_6-MUC1 peptide thioester **1b** and 40mer auxiliary peptide **2b** and subsequent auxiliary-cleavage. The native 86mer MUC1 protein **4c** is isolated in high purity after performing a selective catch&release step, which is based on the reaction between the C-terminal peptide hydrazide moiety and the aldehyde groups of the agarose resin. UPLC analysis of the crude starting materials (co-injection) before (a) and after (b) SPCL. UPLC analysis of the crude product after auxiliary-cleavage (c) obtained by acidic elution from Ni-NTA resin with aqueous 0.25M acetic acid. UPLC analysis of the imin-based immobilization reaction indicates (d) unspecific binding of the hydrolyzed peptide thioester which is removed by washing of the resin (e). UPLC trace (f) and ESI-MS spectrum (g) of final MUC-1 protein. *1:H_2N-(GVTSAPDTRPAPGSTAPPAH)$_2$-CON_2H_3;*2:HS-Aux-HN-(GVTSAPDTRPAPGSTAPPAH)$_2$-CON_2HCOCF_3;*3:tBuS-Aux-HN-(GVTSAPDTRPAPGSTAPPAH)$_2$-CON_2H_3.

1a + **2a** → **4a** (46mer MUC1 protein, 0.98 mg, 167 nmol (33 %, based on initial loading of SPPS–resin and **4a** as 11*TFA-salt)

UPLC-MS: t_R = 1,9 min (3-40% D in 4 min, CSH-column); m/z = 1149.2 ($C_{196}H_{230}N_{70}O_{60}$ $(M+4H)^{4+}$, calcd.: 1149.2 (100%)), 919.6 ($C_{196}H_{232}N_{70}O_{60}$ $(M+5H)^{5+}$, calcd.: 919.5 (100%)), 766.6 ($C_{196}H_{233}N_{70}O_{60}$ $(M+6H)^{6+}$, calcd.: 766.4 (100%)), 657.3 ($C_{196}H_{233}N_{70}O_{60}$ $(M+7H)^{7+}$, calcd.: 657.0 (100%)), 575.2 ($C_{196}H_{234}N_{70}O_{60}$ $(M+8H)^{8+}$, calcd.: 575.0 (100%)); MALDI-MS: 4592.5 ($C_{196}H_{297}N_{70}O_{60}$ $(M+1H)^{1+}$, calc.: 4594.0); $C_{196}H_{296}N_{70}O_{60}$ (MW = 4593.0 g·mol^{-1}).

1a + **2b** → **4b** (66mer MUC1 protein, 0.55 mg, 69.2 nmol (14 %, based on initial loading of SPPS–resin and **4b** as 13*TFA-salt)

UPLC-MS: t_R = 2.1 min (3-40% D in 4 min, CSH-column); m/z = 1077.9 ($C_{276}H_{427}N_{95}O_{87}$ $(M+6H)^{6+}$, calcd.: 1077.9 (100%)), 924.2 ($C_{276}H_{428}N_{95}O_{87}$ $(M+7H)^{7+}$, calcd.: 924.0 (100%)), 808.9 ($C_{276}H_{429}N_{95}O_{87}$ $(M+8H)^{8+}$, calcd.: 808.7 (100%)), 718.9 ($C_{276}H_{430}N_{95}O_{87}$ $(M+9H)^{9+}$, calcd.: 718.9 (100%)), 647.3 ($C_{276}H_{431}N_{95}O_{87}$ $(M+10H)^{10+}$, calcd.: 647.1 (100%)), 588.7 ($C_{276}H_{432}N_{95}O_{87}$ $(M+11H)^{11+}$, calcd.: 588.3 (100%)); MALDI-MS: 6462.1 ($C_{276}H_{422}N_{95}O_{87}$ $(M+1H)^{1+}$, calc.: 6463.0); $C_{276}H_{421}N_{95}O_{87}$ (MW = 6462.0 g·mol^{-1}).

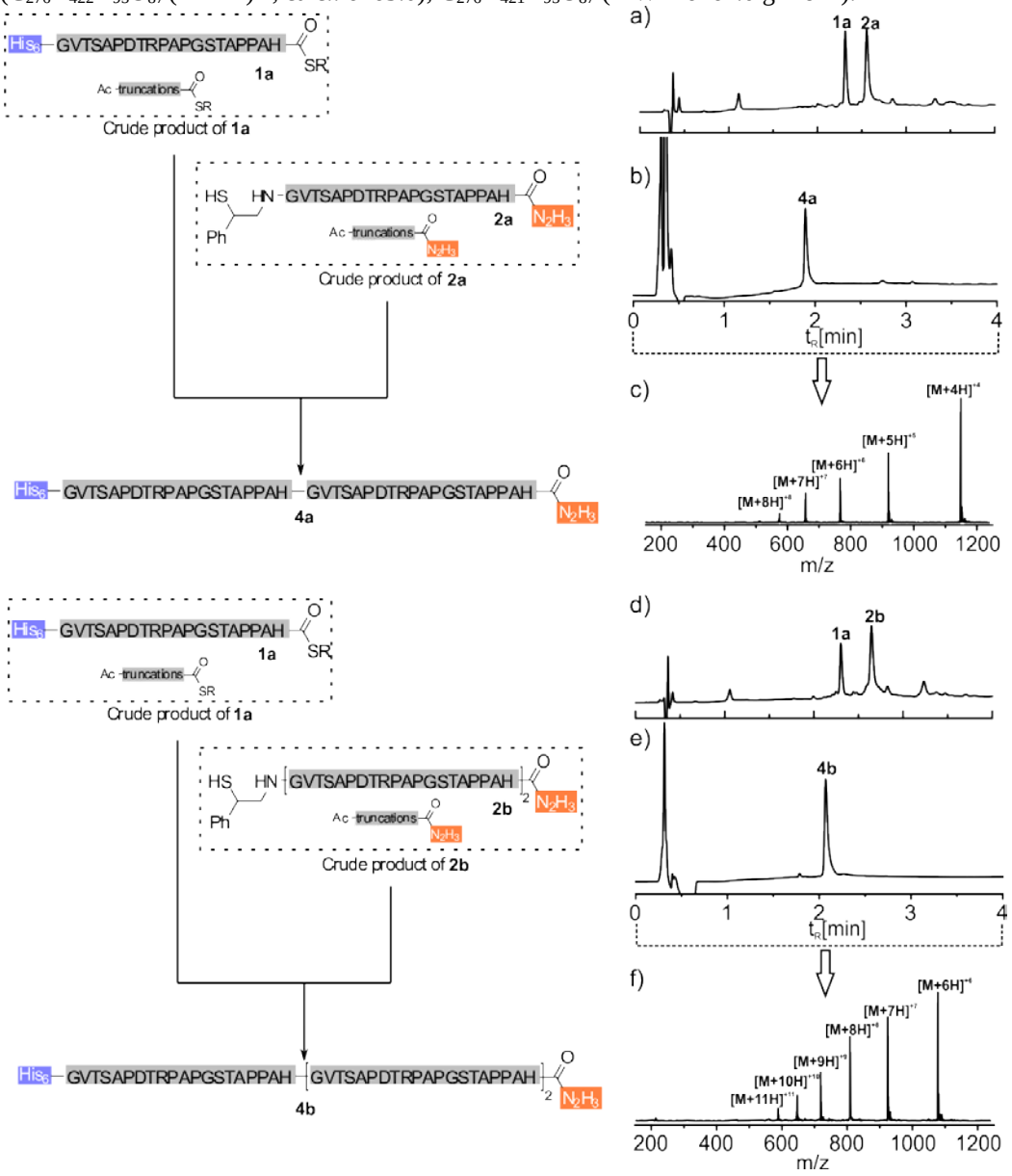


Figure S12: Synthesis of 46mer MUC1 protein **4a** and 66mer MUC1 protein **4b** by solid-supported auxiliary-mediated SPCL and subsequent non-chromatographic purification. UPLC analysis of a), d) co-injected crude starting materials and b), e) MUC1 proteins after release from aldehyde-agarose. ESI-MS analysis of c), f) product was performed over the entire run.

1a + **2c** → **4c** (86mer MUC1 protein, 0.87 mg, 86.6 nmol (17 %, based on initial loading of SPPS–resin and **4c** as 15*TFA-salt)

UPLC-MS: t_R = 2.1 min (3-40% B in 4 min, CSH-column); m/z = 1191.2 ($C_{356}H_{553}N_{120}O_{114}$ $(M+7H)^{7+}$, calcd.: 1190.9 (100%)), 1042.4 ($C_{356}H_{554}N_{120}O_{114}$ $(M+8H)^{8+}$, calcd.: 1042.1 (100%)), 926.8 ($C_{356}H_{555}N_{120}O_{114}$ $(M+9H)^{9+}$, calcd.: 926.5 (100%)), 834.4 ($C_{356}H_{556}N_{120}O_{114}$ $(M+10H)^{10+}$, calcd.: 833.9 (100%)), 758.6 ($C_{356}H_{557}N_{120}O_{114}$ $(M+11H)^{11+}$, calcd.: 758.2 (100%)), 695.4 ($C_{356}H_{558}N_{120}O_{114}$ $(M+12H)^{12+}$, calcd.: 695.1 (100%)), 642.1 ($C_{356}H_{559}N_{120}O_{114}$ $(M+13H)^{13+}$, calcd.: 641.7 (100%)); MALDI-MS: 8331.3 ($C_{356}H_{547}N_{120}O_{114}$ $(M+1H)^{1+}$, calc.: 8332.0); $C_{356}H_{546}N_{120}O_{114}$ (MW = 8331.0 g·mol^{-1}).

1b + **2a** → **4b** (66mer MUC1 protein, 0.73 mg, 91.9 nmol (18 %, based on initial loading of SPPS–resin and **4b** as 13*TFA-salt)

UPLC-MS: t_R = 2.0 min (3-40% D in 4 min, CSH-column); m/z = 1078.0 ($C_{276}H_{427}N_{95}O_{87}$ $(M+6H)^{6+}$, calcd.: 1077.9 (100%)), 924.3 ($C_{276}H_{428}N_{95}O_{87}$ $(M+7H)^{7+}$, calcd.: 924.0 (100%)), 808.9 ($C_{276}H_{429}N_{95}O_{87}$ $(M+8H)^{8+}$, calcd.: 808.7 (100%)), 719.1 ($C_{276}H_{430}N_{95}O_{87}$ $(M+9H)^{9+}$, calcd.: 718.9 (100%)), 647.4 ($C_{276}H_{431}N_{95}O_{87}$ $(M+10H)^{10+}$, calcd.: 647.1 (100%)), 588.6 ($C_{276}H_{432}N_{95}O_{87}$ $(M+11H)^{11+}$, calcd.: 588.3 (100%)); MALDI-MS: 6462.8 ($C_{276}H_{422}N_{95}O_{87}$ $(M+1H)^{1+}$, calc.: 6463.0), 3231.8 ($C_{276}H_{423}N_{95}O_{87}$ $(M+2H)^{2+}$, calc.: 3232.0); $C_{276}H_{421}N_{95}O_{87}$ (MW = 6462.0 g·mol^{-1}).

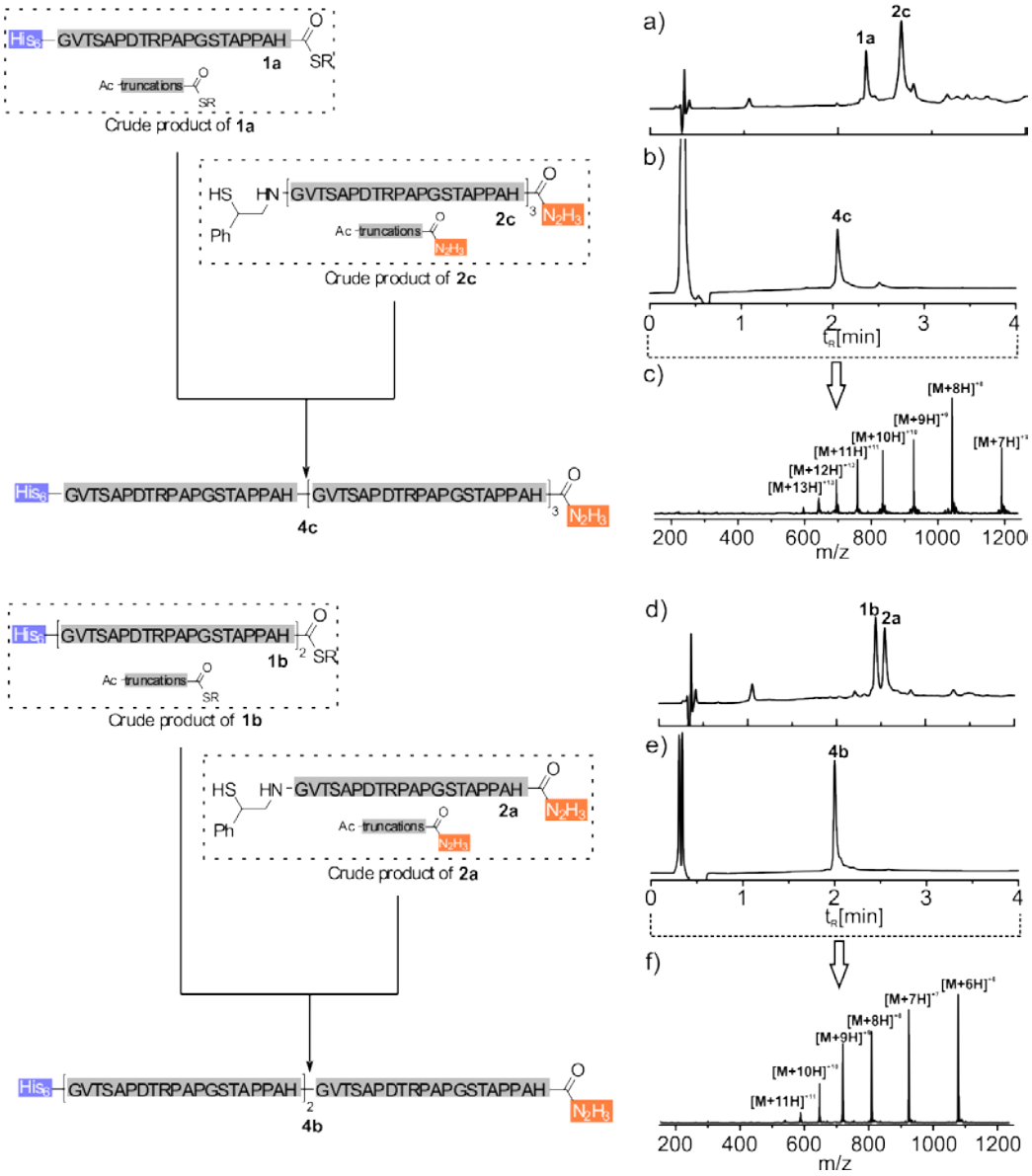


Figure S13: Synthesis of 86mer MUC1 protein **4c** and 66mer MUC1 protein **4b** by solid-supported auxiliary-mediated SPCL and subsequent non-chromatographic purification. UPLC analysis of a), d) co-injected crude starting materials and b), e) MUC1 proteins after release from aldehyde-agarose. ESI-MS analysis of c), f) product was performed over the entire run time.

1b + **2c** → **4d** (106mer MUC1 protein, 0.54 mg, 44.4 nmol (9 %, based on initial loading of SPPS–resin and **4d** as 17*TFA-salt)

UPLC-MS: t_R = 2.2 min (3-40% D in 4 min); m/z = 1134.4 ($C_{436}H_{680}N_{145}O_{141}$ $(M+9H)^{9+}$, calcd.: 1134.2 (100%)), 1021.1 ($C_{436}H_{681}N_{145}O_{141}$ $(M+10H)10^{+}$, calcd.: 1020.9 (100%)), 928.5 ($C_{436}H_{682}N_{145}O_{141}$ $(M+11H)^{11+}$, calcd.: 928.2 (100%)), 851.2 ($C_{436}H_{683}N_{145}O_{141}$ $(M+12H)^{12+}$, calcd.: 850.9 (100%)), 785.9 ($C_{436}H_{684}N_{145}O_{141}$ $(M+13H)^{13+}$, calcd.: 785.5 (100%)), 729.7 ($C_{436}H_{685}N_{145}O_{141}$ $(M+14H)^{14+}$, calcd.: 729.5 (100%)), 681.1 ($C_{436}H_{686}N_{145}O_{141}$ $(M+15H)^{15+}$, calcd.: 680.9 (100%)), 638.2 ($C_{436}H_{686}N_{145}O_{141}$ $(M+16H)^{16+}$, calcd.: 638.4 (100%)); MALDI-MS: 10199.6 ($C_{436}H_{672}N_{145}O_{141}$ $(M+1H)^{1+}$, calc.: 10201.0), 5102.8 ($C_{436}H_{673}N_{145}O_{141}$ $(M+2H)^{2+}$, calc.: 5101.0); $C_{436}H_{671}N_{145}O_{141}$ (MW = 10200.0 g·mol^{-1}).

1c + **2a** → **4c** (86mer MUC1 protein, 0.74 mg, 73.2 nmol (15 %, based on initial loading of SPPS–resin and **4c** as 15*TFA-salt)

UPLC-MS: t_R = 2.0 min (3-40% D in 4 min); m/z = 1191.2 ($C_{356}H_{553}N_{120}O_{114}$ $(M+7H)^{7+}$, calcd.: 1190.9 (100%)), 1042.4 ($C_{356}H_{554}N_{120}O_{114}$ $(M+8H)^{8+}$, calcd.: 1042.1 (100%)), 926.8 ($C_{356}H_{555}N_{120}O_{114}$ $(M+9H)^{9+}$, calcd.: 926.5 (100%)), 834.3 ($C_{356}H_{556}N_{120}O_{114}$ $(M+10H)^{10+}$, calcd.: 833.9 (100%)), 758.6 ($C_{356}H_{557}N_{120}O_{114}$ $(M+11H)^{11+}$, calcd.: 758.2 (100%)), 695.5 ($C_{356}H_{558}N_{120}O_{114}$ $(M+12H)^{12+}$, calcd.: 695.1 (100%)), 642.0 ($C_{356}H_{559}N_{120}O_{114}$ $(M+13H)^{13+}$, calcd.: 641.7 (100%)), 596.0 ($C_{356}H_{560}N_{120}O_{114}$ $(M+14H)^{14+}$, calcd.: 595.9 (100%)); MALDI-MS: 8332.3 ($C_{356}H_{547}N_{120}O_{114}$ $(M+1H)^{1+}$, calc.: 8332.0); $C_{356}H_{546}N_{120}O_{114}$ (MW = 8331.0 g·mol^{-1}).

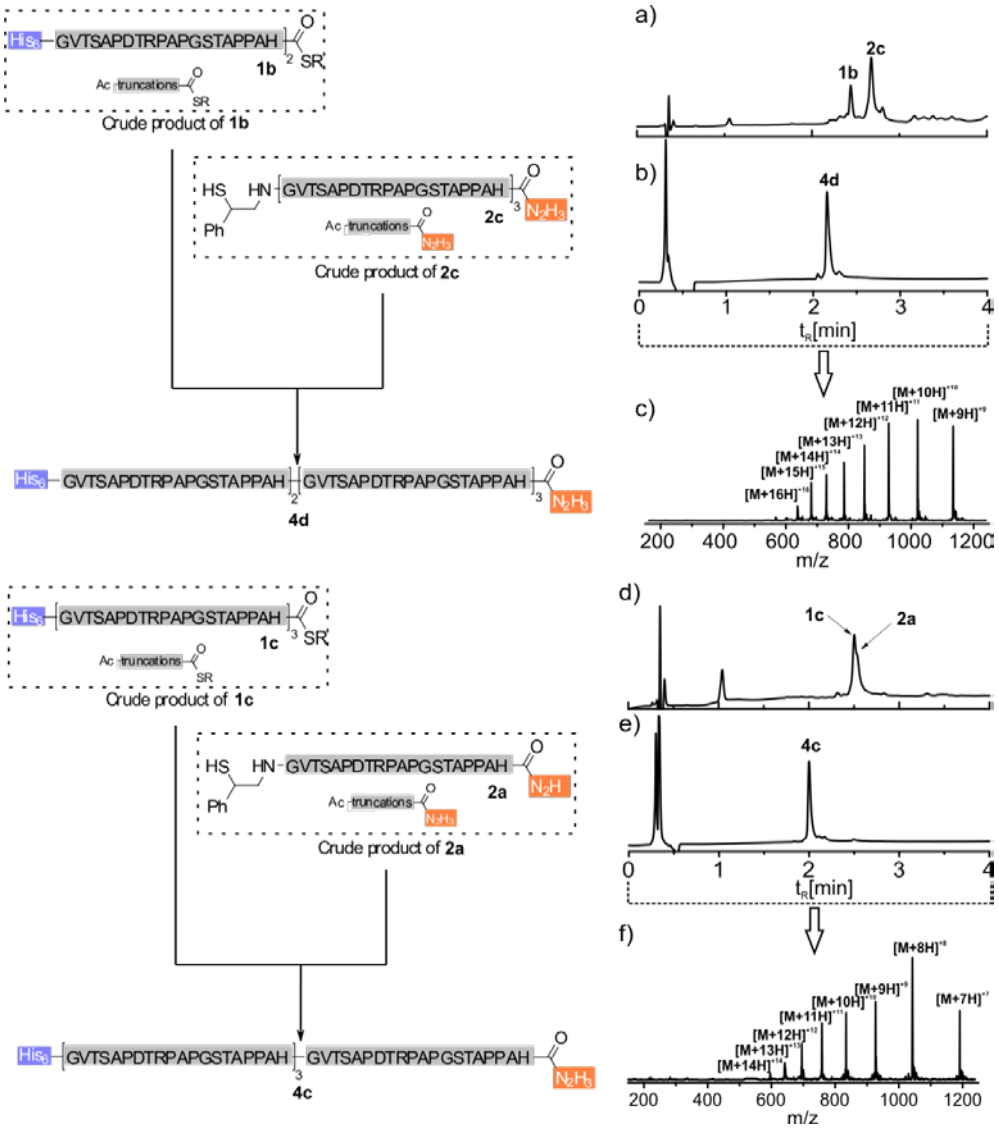


Figure S14: Synthesis of 106mer MUC1 protein **4d** and 86mer MUC1 protein **4c** by solid-supported auxiliary-mediated SPCL and subsequent non-chromatographic purification. UPLC analysis of a), d) co-injected crude starting materials and b), e) MUC1 proteins after release from aldehyde-agarose. ESI-MS analysis of c), f) product was performed over the entire run time.

1c + **2b** → **4d** (106mer MUC1 protein, 0.51 mg, 42.0 nmol (8 %, based on initial loading of SPPS–resin and **4e** as 17*TFA-salt)

UPLC-MS: t_R = 2.1 min (3-40% D in 4 min); m/z = 1134.5 ($C_{436}H_{680}N_{145}O_{141}$ $(M+9H)^{9+}$, calcd.: 1134.2 (100%)), 1021.1 ($C_{436}H_{681}N_{145}O_{141}$ $(M+10H)10^{+}$, calcd.: 1020.9 (100%)), 928.4 ($C_{436}H_{682}N_{145}O_{141}$ $(M+11H)^{11+}$, calcd.: 928.2 (100%)), 851.2 ($C_{436}H_{683}N_{145}O_{141}$ $(M+12H)^{12+}$, calcd.: 850.9 (100%)), 785.9 ($C_{436}H_{684}N_{145}O_{141}$ $(M+13H)^{13+}$, calcd.: 785.5 (100%)), 729.9, ($C_{436}H_{685}N_{145}O_{141}$ $(M+14H)^{14+}$, calcd.: 729.5 (100%)), 681.0 ($C_{436}H_{686}N_{145}O_{141}$ $(M+15H)^{15+}$, calcd.: 680.9 (100%)), 638.9 ($C_{436}H_{686}N_{145}O_{141}$ $(M+16H)^{16+}$, calcd.: 638.4 (100%)); MALDI-MS: 10200.2 ($C_{436}H_{672}N_{145}O_{141}$ $(M+1H)^{1+}$, calc.: 10201.0), 5102.0 ($C_{436}H_{673}N_{145}O_{141}$ $(M+2H)^{2+}$, calc.: 5101.0); $C_{436}H_{671}N_{145}O_{141}$ (MW = 10200.0 g·mol^{-1}).

1c + **2c** → **4e** (126mer MUC1 protein, 1.08 mg, 85.3 nmol (17 %, based on initial loading of SPPS–resin and **4e** as 19*TFA-salt)

UPLC-MS: t_R = 2.2 min (3-40% D in 4 min); m/z = 1208.1 ($C_{516}H_{806}N_{170}O_{168}$ $(M+10H)^{10+}$, calcd.: 1207.7 (100%)), 1098.4 ($C_{516}H_{807}N_{170}O_{168}$ $(M+11H)^{11+}$, calcd.: 1098.0 (100%)), 1006.8 ($C_{516}H_{808}N_{170}O_{168}$ $(M+12H)^{12+}$, calcd.: 1006.6 (100%)), 929.5 ($C_{516}H_{809}N_{170}O_{168}$ $(M+13H)^{13+}$, calcd.: 929.9 (100%)), 863.3 ($C_{516}H_{810}N_{170}O_{168}$ $(M+14H)^{14+}$, calcd.: 862.9 (100%)), 805.9 ($C_{516}H_{811}N_{170}O_{168}$ $(M+15H)^{15+}$, calcd.: 805.5 (100%)), 755.6 ($C_{516}H_{812}N_{170}O_{168}$ $(M+16H)^{16+}$, calcd.: 755.2 (100%)), 711.3 ($C_{516}H_{813}N_{170}O_{168}$ $(M+17H)^{17+}$, calcd.: 710.8 (100%)), 671.2 ($C_{516}H_{814}N_{170}O_{168}$ $(M+18H)^{18+}$, calcd.: 671.4 (100%)); MALDI-MS: 12069.5 ($C_{516}H_{797}N_{170}O_{168}$ $(M+1H)^{1+}$, calc.: 12069.9), 6035.6 ($C_{516}H_{798}N_{170}O_{168}$ $(M+2H)^{2+}$, calc.: 6035.5), 4025.5 ($C_{516}H_{799}N_{170}O_{168}$ $(M+3H)^{3+}$, calc.: 4024.0); $C_{516}H_{796}N_{170}O_{168}$ (MW = 12068.9 g·mol^{-1}).

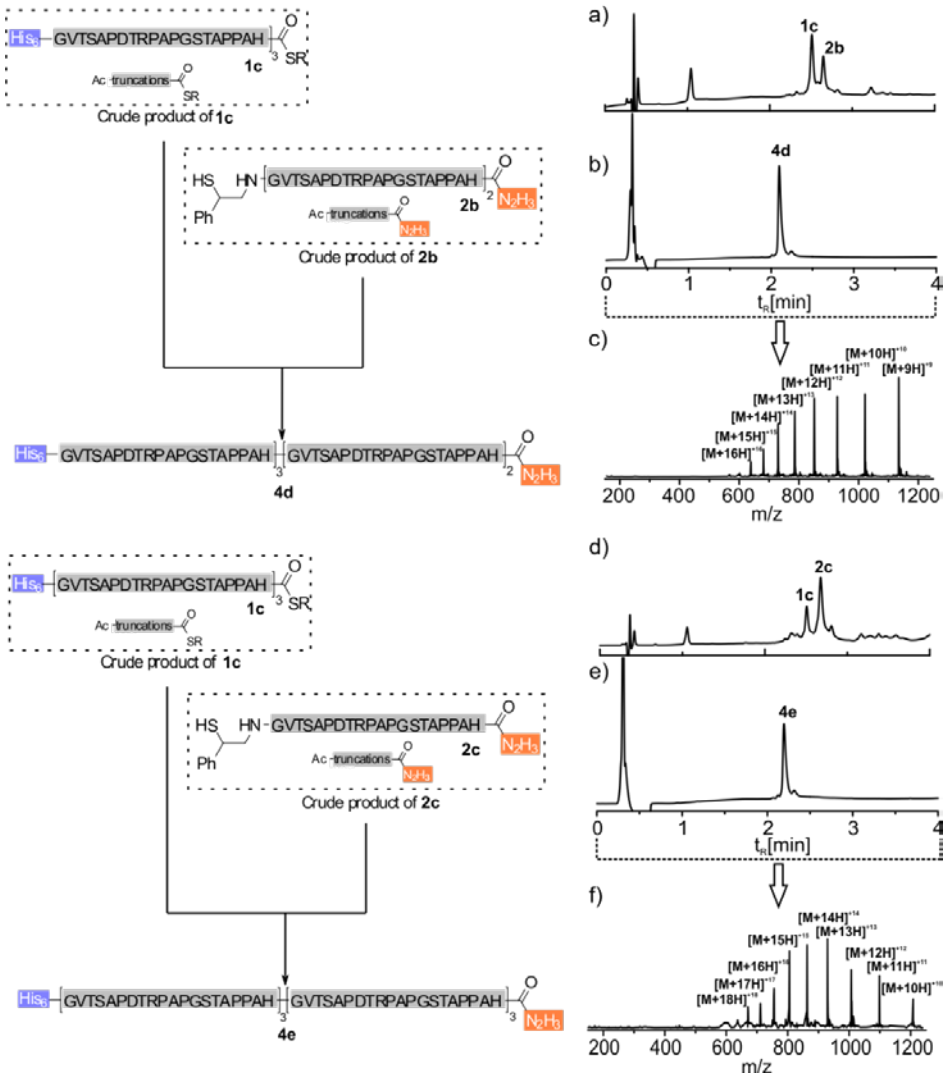


Figure S15: Synthesis of 106mer MUC1 protein **4d** and 126mer MUC1 protein **4e** by solid-supported auxiliary-mediated SPCL and subsequent non-chromatographic purification. UPLC analysis of a), d) co-injected crude starting materials and b), e) MUC1 proteins after release from aldehyde-agarose. ESI-MS analysis of c), f) product was performed over the entire run time.

4.3 Reactivity-based purification as an alternative to chromatographic purification

The limits of HPLC-based purification became apparent in the reaction between 66mer peptide thioester **1c** and 20mer auxiliary peptide **2a**. The hydrolysis product **1c$_{hydr}$** is difficult to separate from 86mer MUC1 protein **4c** (Fig. S16a). Such separations are readily achieved by using catch-and-release from aldehyde-agarose (Fig. S16b).

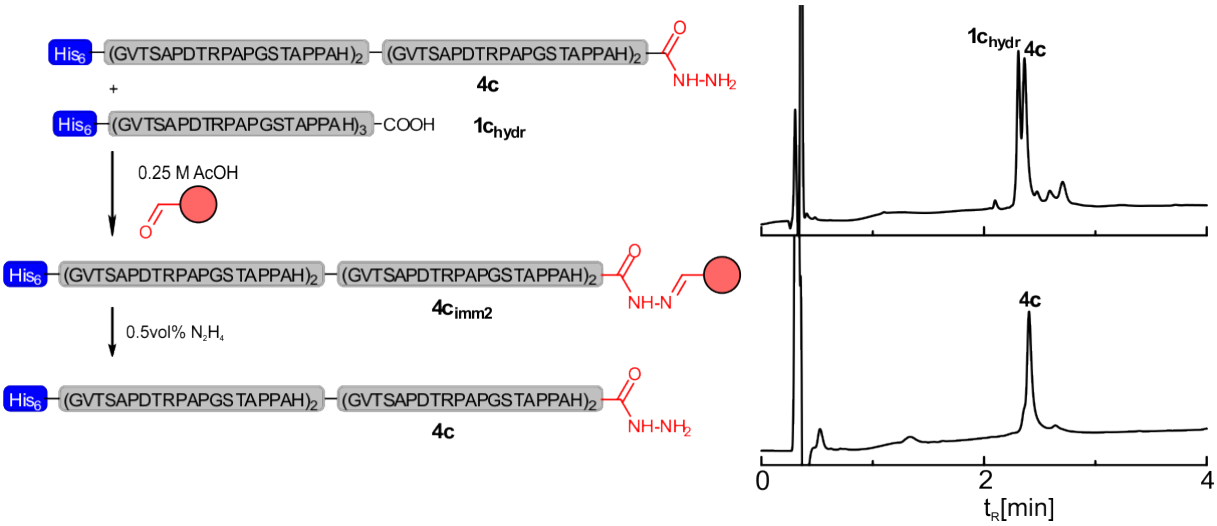


Figure S16: Reactivity-based purification as an alternative to chromatographic purification: Separation of MUC1-protein **4c** from hydrolyzed peptide thioester **1c$_{hydr}$** by using catch-and-release from aldehyde-agarose. UPLC analysis of the crude product (a) and the final MUC-1 protein **4c** (b).

4.4 Stability of Ni-His$_6$-complex and Ni-NTA resin during ligation and auxiliary removal

An aliquot of the supernatant collected during ligation of **2a** with resin-bound **1c** in a typical reaction buffer (20 mM TCEP, 20 mM imidazole, 6 M GuHCl, 200 mM Na_2HPO_4, 2 vol% PhSH, pH 7.5, r.t., 24 h) was analysed by UPLC-MS (Fig. S17). The absence of peaks for peptide thioester or ligation products indicates that the Ni·NTA·His$_6$ linkange remained staible under ligation conditions (Fig, S17b). product 4c no detachment of His$_6$-tagged proteins was observed, as indicated by the UPLC analysis of the ligation solution (see Fig S17b).

During auxiliary removal (200 mM TCEP, 800 mM morpholine, pH 8.5, 40°C, 24 h) partial detachment of the His$_6$-tagged proteins was observed (see Fig 17c). However, after adjusting the pH to 7.0 by addition of 1 M HCl the His$_6$-tagged proteins were re-immobilized (max. 30 minutes) onto the Ni-NTA resin (see Fig. S17d).

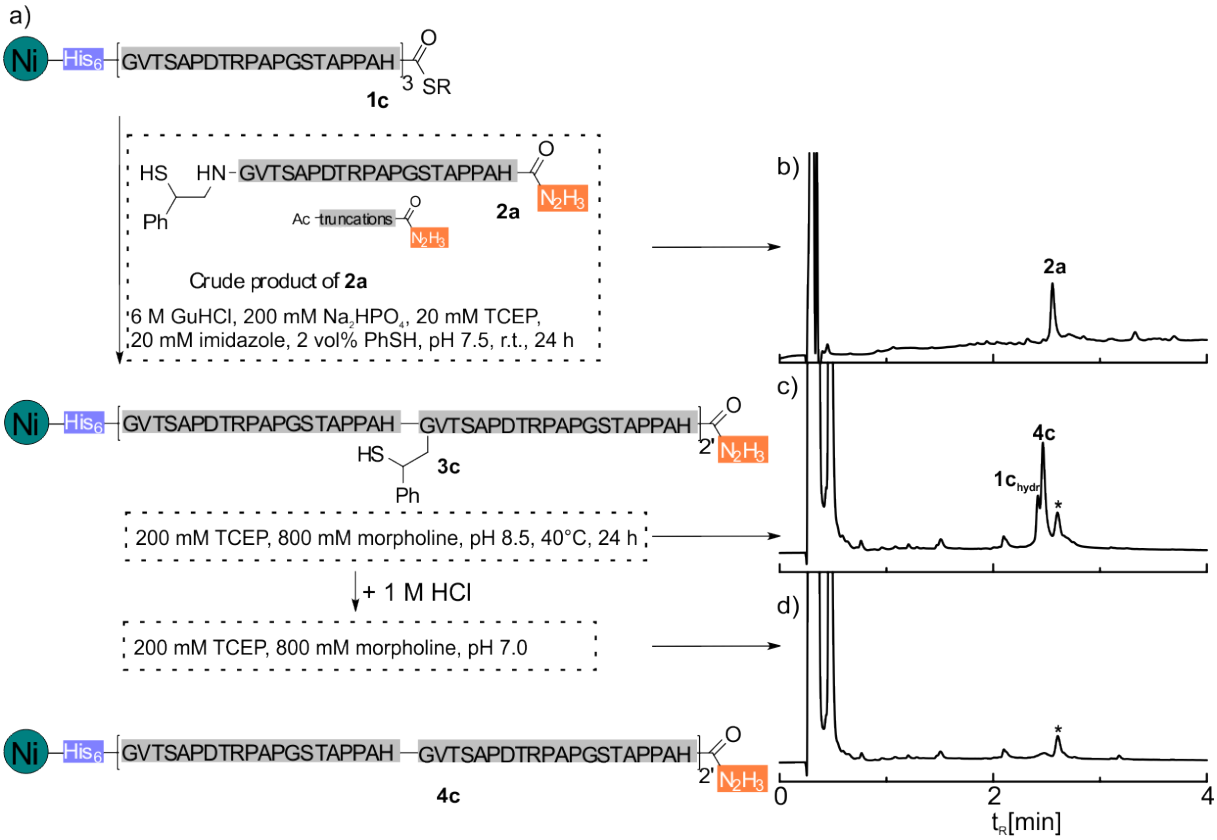


Figure S17: Stability of Ni-His$_6$-complex during ligation and auxiliary removal: UPLC analysis of (b) the ligation solution indicates that no His$_6$-tagged proteins are detached during the SPCL reaction. After removing the auxiliary partial detachment of the His$_6$-tagged proteins was detected (c). Re-immobilization was achieved by neutralization of the cleavage solution by addition of 1 M HCl (d).*:non-peptidic material

4.5 Acidic elution of MUC1 proteins

As an alternative to detachment by means of 0.5 vol% hydrazine treatment (Fig. S18b), final release was accomplished by acidolysis using 1 vol% TFA (Fig. S18c).

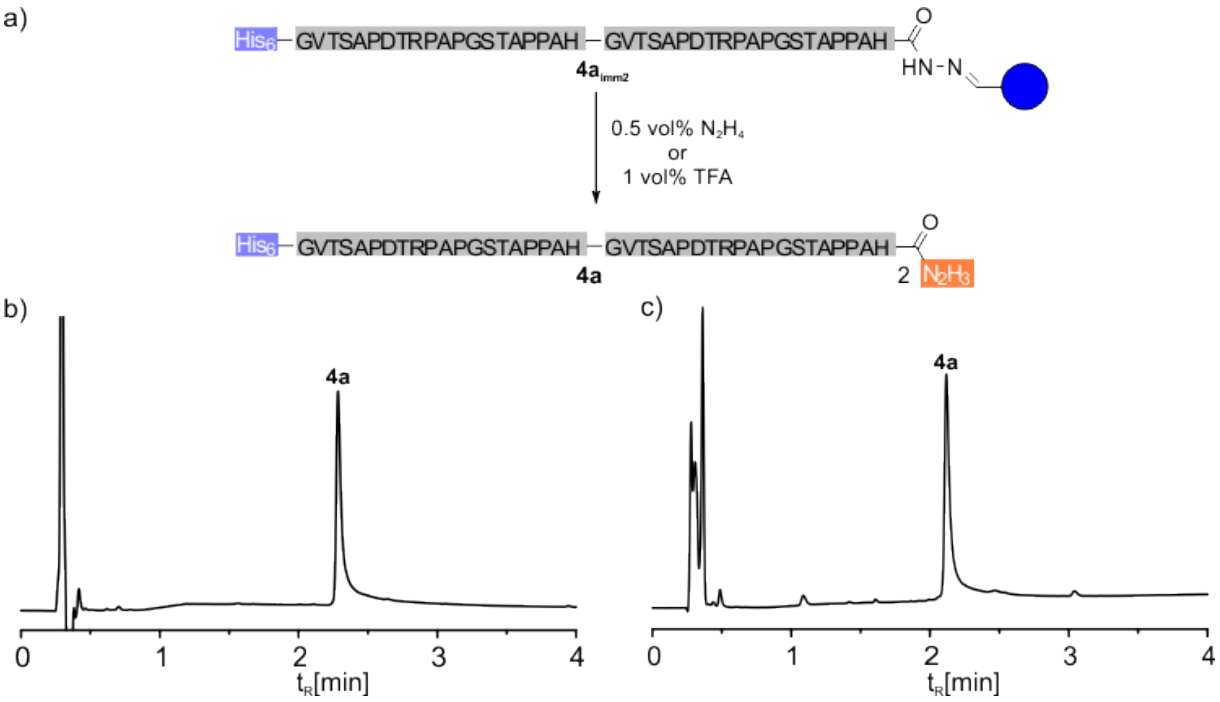


Figure S18: Final release of 46mer MUC1 protein **4a** and UPLC analysis of cleavage solution obtained by treatment with (b) 0.5 vol% hydrazine or (c) 1 vol% TFA.

5. MALDI-MS analysis

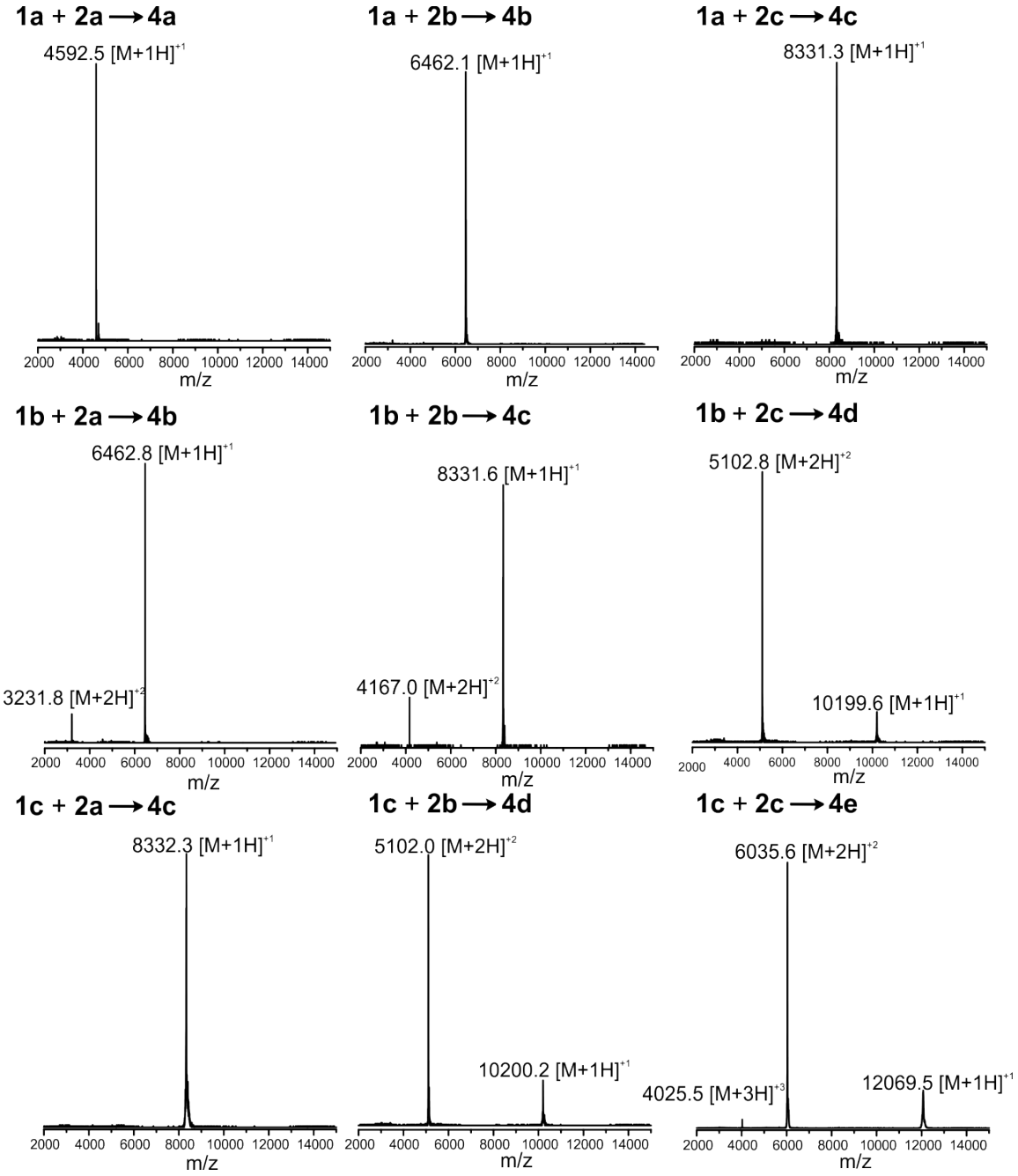


Figure S19: MALDI-MS analysis of final products. Conditions: linear positive mode, mass range (m/z) = 2000-15000, CHCA-matrix.

6. References

1. S. F. Loibl, Z. Harpaz and O. Seitz, *Angew. Chem. Int. Ed.*, 2015, **54**, 15055-15059.
2. B. J. H. Kuipers and H. Gruppen, *J. Agric. Food. Chem.*, 2007, **55**, 5445-5451.