Supporting Information

Oxo-Ester Mediated Native Chemical Ligation on Microarrays: An Efficient and Chemoselective Coupling Methodology

Martin J. Weissenborn,[a] Roberto Castangia,[a] Johannes W. Wehner,[b] Robert Šardzík,[a] Thisbe K. Lindhorst[b] and Sabine L. Flitsch*[a]*

[a] *School of Chemistry & Manchester Interdisciplinary Biocentre, The University of Manchester, 131 Princess Street, Manchester, M17DN (UK). Fax: 44- 161-275-1311; Tel: -44- 161-306-5172; E-mail: sabine.flitsch@manchester.ac.uk* [b] *Otto Diels Institute of Organic Chemistry, Christiana Albertina University of Kiel, Otto-Hahn-Platz 3/4, 24098 Kiel, Germany. Fax: 49-431-880-7410; E-mail: tklind@oc.uni-kiel.de*

1. General methods

Unless otherwise noted, all reagents including anhydrous solvents were obtained from commercial suppliers and used as delivered without further purification. Alkanethiol linkers [HS-$(CH_2)_{17}$-$(OC_2H_4)_3$-OH and HS-$(CH_2)_{17}$-$(OC_2H_4)_6$-OCH_2COOH] were purchased from Prochimia Surfaces (Poland). The peptide CQDSETRTFY **9** was purchased from Bachem (UK).

Reverse phase HPLC was performed on Agilent 1100 and Agilent 1200 systems using Phenomenex Luna 5 µm, 250 x 150 mm semi-preparative C18 column with UV detection at 210 nm.

Unless otherwise stated all MALDI-ToF MS experiments were carried out on an Ultraflex II instrument (Bruker Daltonics) in positive reflectron mode. A solution of matrix (2,4,6-trihydroxyacetophenone, 10 mg/mL in acetone) was applied on the gold and allowed to dry before analysis.

All spectra were analysed with FlexAnalysis software (Bruker, USA) using default integration settings. Calibration was either performed before the analysis at the Ultraflex II instrument or afterwards in FlexAnalysis. Unless otherwise noted, all m/z values refer to the $[M+Na]^+$ ion and the corresponding disulphide which is formed during ionisation.

2. Peptide synthesis

Cys-Ala-Lys-Gly-Val-Thr-Ser-Ala-Pro-Ala-Gly (11)
Gly-Ala-Lys-Gly-Val-Thr-Ser-Ala-Pro-Ala-Gly (10)
Gly-Ala-Pro-Gly-Pro-Thr-Pro-Gly-Pro-Ala-Gly-Lys (12)

Peptide synthesis was performed under standard Fmoc-based solid-phase peptide synthesis conditions using a CEM Liberty® peptide synthesizer starting from Fmoc-glycine loaded on polystyrene resin *via* the Wang linker (0.1 mmol, loading 0.79 mmol/g). Fmoc-protected amino acid (5 eq.), PyBOP (benzotriazol-1-yl-oxytripyrrolidino phosphonium hexafluorophosphate, 5 eq.) and DIPEA (N,N-diisopropylethylamine, 10 eq.) were used for the coupling steps. Piperidine (20 % solution in DMF) was used for the deprotection steps. After completion, the resin-bound peptide was thoroughly washed with DMF, DMF-MeOH (1:1), DCM and cleaved from the resin with trifluoroacetic acid / triisopropylsilane / water (95:2.5:2.5 v/v/v). Precipitation in cold ether, centrifugation and evaporation of the solvent afforded the desired crude peptide as a white solid. Semi-preparative HPLC purification (Eluents: A: 0.1 % formic acid in water; B: 0.1 % formic acid in acetonitrile, flow rate 4 ml/min, 0-5 min isocratic 10 % B; 5-20 min linear gradient from 10 to 40 % B) followed by freeze-drying afforded the peptides **10** and **11** as a white solid.

Cys-Ala-Lys-Gly-Val-Thr-Ser-Ala-Pro-Ala-Gly (11)
MALDI-ToF MS (THAP): calcd. for $C_{39}H_{67}N_{11}NaO_{14}S^+$: *m/z* 968.45 $[M+Na]^+$; found: *m/z* 967.90 $[M+Na]^+$;

Gly-Ala-Lys-Gly-Val-Thr-Ser-Ala-Pro-Ala-Gly (10)
MALDI-ToF MS (THAP): calcd. for $C_{38}H_{65}N_{11}NaO_{14}^+$: *m/z* 922.46 $[M+Na]^+$; found: *m/z* 922.02 $[M+Na]^+$;

Gly-Ala-Pro-Gly-Pro-Thr-Pro-Gly-Pro-Ala-Gly-Lys (12)
MALDI-ToF MS (THAP): calcd. for $C_{44}H_{71}N_{13}NaO_{14}^+$: *m/z* 1028.51 $[M+Na]^+$; found: *m/z* 1028.31 $[M+Na]^+$.

3. Saccharide synthesis

General methods used in saccharide synthesis.

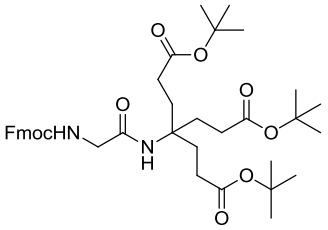
Commercially available starting materials and reagents were used without further purification. Reactions requiring dry conditions were performed under an atmosphere of nitrogen using oven-dried glassware. Dichloromethane (DCM) was dried over CaH_2 and anhydrous DMF was purchased.

L-Cystine-di[2-(α-D-mannopyranosyloxy)ethyl]amide (**13**)[1], the trivalent wedge **17**[2], 2-aminoethyl α-D-mannoside (**15**)[3] and 2-aminomethyl 2-acetamido-2-deoxy-β-D-glucopyranoside (**14**) [4] were prepared according to the literature. Amberlyst A-21 was treated as described by Srinivasan et. al.[5] before usage.

Reactions were monitored by thin-layer chromatography using either silica gel 60 GF254 on aluminium foil (Merck) or RP-18 F254s on aluminium foil (Merck) with detection by UV light and charring with sulfuric acid in EtOH (10 %). Silica gel 60 Å (Merck, 230-400 mesh) was used for flash chromatography. Preparative MPLC was performed on a Büchi apparatus using a LiChroprep RP-18 (40–60 mm, Merck) column for reverse phase and a LiChroprep Si 60 (40–60 mm, Merck) column for normal phase silica gel chromatography. Size exclusion chromatography (SEC) was carried out on Sephadex LH-20 from Pharmacia-Biotech.

^{1}H and ^{13}C NMR spectra were recorded using a Bruker DRX-500 or a Bruker AV-600 spectrometer. NMR spectra were calibrated with respect to the solvent peak (in case of $CDCl_3$ the reference was tetramethylsilane (TMS)). 2D NMR techniques (COSY, HSQC, HMBC) were used for full assignment of the spectra. ESI MS measurements were performed on a Mariner ESI-ToF 5280 instrument (Applied Biosystems). High-resolution mass spectra (HR-ESI MS) were obtained with the Waters Micromass LCT TOF mass spectrometer. MALDI-ToF mass spectra were recorded on a Bruker Biflex-III 19 kV instrument with Cl-CCA (4-chloro-α-cyanocinnamic acid) or DHB (2,5-dihydroxybenzoic acid) as matrix. Optical rotations were measured on a Perkin-Elmer polarimeter 341 (Na-D-line: 589 nm, length of cell 1 dm). IR spectra were recorded on a Perkin-Elmer Paragon 1000 FT-IR spectrometer. For sample preparation a Golden Gate-diamond-ATR unit with a saphire stamp was used. Melting points were measured on a MPD 350 (Fa. Gallenkamp) apparatus.

3-Cascade:*N*-fluorenylmethoxycarbonyl-glycyl-aminomethane[3]:propionic acid *tert*-butyl ester (18)



Fmoc-Glycine (**16**) (558 mg, 1.88 mmol), the trivalent wedge molecule **17** (860 mg, 2.07 mmol) and HBTU (1.07 g, 2.82 mmol) were dried under vacuum for 1 h. The reactants were dissolved in anhydrous DMF (20 mL), the mixture was cooled to 0°C and DIPEA (480 µL, 2.82 mmol) was added. The reaction mixture was stirred for 24 h at ambient temperature under a nitrogen atmosphere. The solvent was removed under reduced pressure and the crude product was dissolved in water and extracted twice with ethyl acetate (30 mL each). The combined organic phases were washed with brine, dried over $MgSO_4$ and after filtration solvents removed under reduced pressure. The crude product was purified by column chromatography (cyclohexane / ethyl acetate, 3 : 1) yielding **18** (1.20 g, 1.73 mmol, 92 %) as a colourless solid.

R*f*: 0.52 (cyclohexane / ethyl acetate, 1 : 1);
mp.: 113°C – 115°C;
^{1}H NMR (500 MHz, $CDCl_3$, 300 K): δ = 7.76 (d, ^3J = 7.5 Hz, 2H, $H_{aryl,Fmoc}$), 7.60 (d, ^3J = 7.5 Hz, 2H, $H_{aryl,Fmoc}$), 7.40 (t, ^3J = 7.5 Hz, 2H, $H_{aryl,Fmoc}$), 7.31 (t, ^3J = 7.5 Hz, 2H, $H_{aryl,Fmoc}$), 6.46 (s, 1H, NH), 5.44 (br s, 1H, NH), 4.41 (d, ^3J = 7.0 Hz, 2H, $CH_{2,Fmoc}$), 4.23 (t, ^3J = 7.0 Hz, CH_{Fmoc}), 3.80 (d, ^3J = 3.7 Hz, 2H, H_{Gly}), 2.23 (t, ^3J = 7.7 Hz, 6H, C*CH*$_2$CH$_2$C(O)), 1.99 (t, ^3J = 7.7 Hz, 6H, CCH$_2$*CH*$_2$C(O)), 1.43 (s, 27H, ^{t}*Butyl*) ppm;
^{13}C NMR (125 MHz, $CDCl_3$, 300 K): δ = 172.9 (*C*(O)O), 167.8 (*C*(O)NH), 156.4 (O*C*(O)NH), 143.8, 141.3 ($C_{aryl,Fmoc}$), 127.7, 127.1, 125.1, 120.0 ($CH_{aryl,Fmoc}$), 80.8 ($C(CH_3)_3$), 67.4 ($CH_{2,Fmoc}$), 57.8 (NH*C*(CH$_2$CH$_2$)$_3$), 47.1 (CH_{Fmoc}), 44.8 ($CH_{2,Gly}$), 30.1 (CCH$_2$*CH*$_2$C(O)), 29.8 (C*CH*$_2$CH$_2$C(O)), 28.1 (C(*CH*$_3$)$_3$)) ppm;
ESI MS: calcd. for $C_{39}H_{54}N_2NaO_9$: *m/z* 717.372 $[M+Na]^+$; found: *m/z* 717.379 $[M+Na]^+$;
IR (ATR-IR): $\tilde{\nu}$ = 2976, 1722, 1657, 1525, 1449, 1366, 1316, 1231, 1146, 1099, 1043, 953, 849, 801, 760, 739 cm^{-1}.

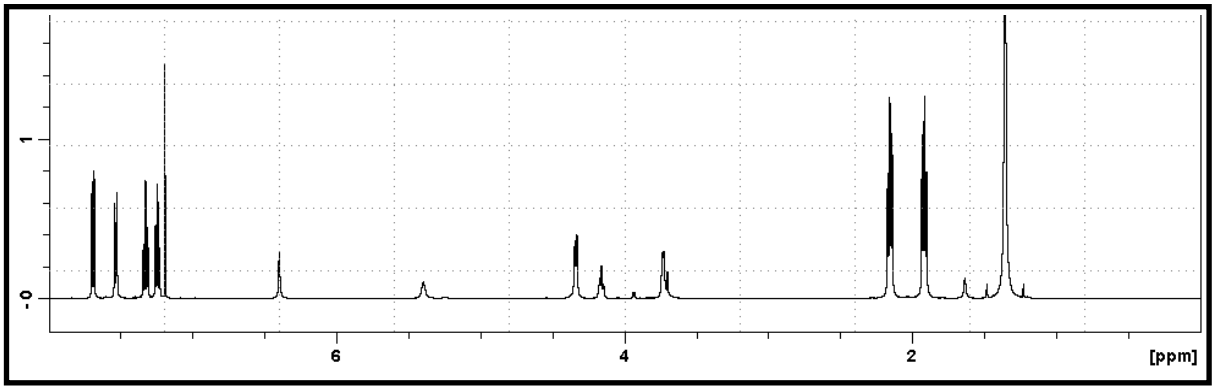


Figure S1: ^{1}H-NMR spectrum (500 MHz, $CDCl_3$) trivalent wedge **18**.

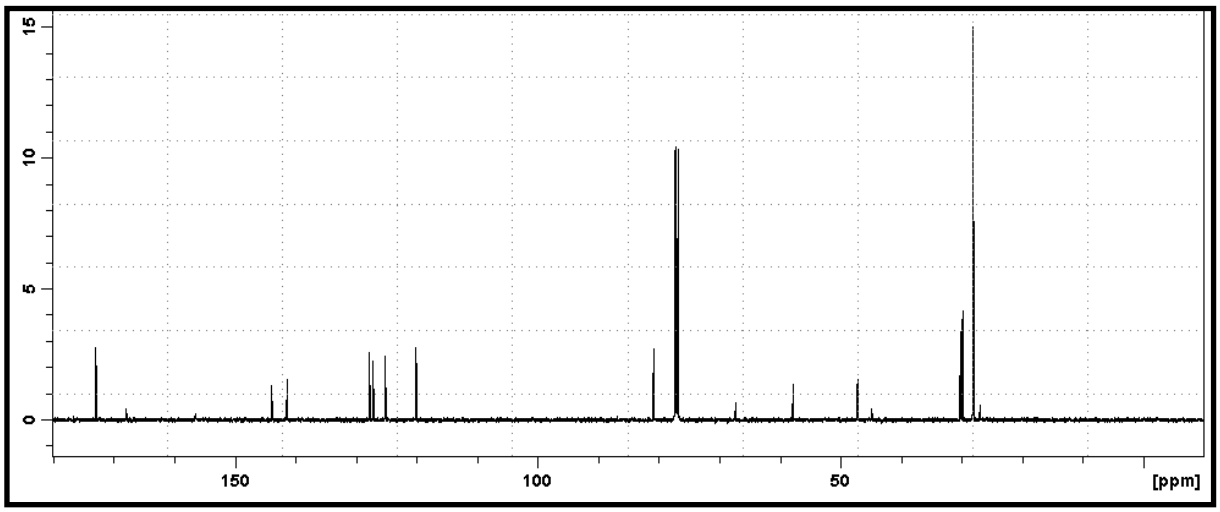
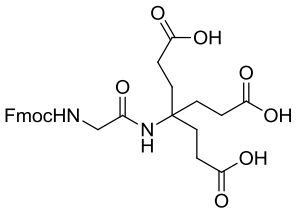


Figure S2: ^{13}C-NMR spectrum (125 MHz, $CDCl_3$) trivalent wedge **18.**

3-Cascade:*N*-fluorenylmethoxycarbonyl-glycyl-aminomethane[3]:propionic acid (19)



The *tert*-butyl ester **18** (140 mg, 201 µmol) was dissolved in formic acid (12 mL) and stirred overnight at room temperature. Then the solvent was removed under reduced pressure and the residue three times codestilled with toluene (10 mL each). The colourless syrup was dissolved in water and extracted three times with dichloromethane (10 mL each). The aqueous phase was lyophyllised yielding **19** (97 mg, 184 µmol, 92%) as a colourless solid.

R*f*: 0.40 (EtOAc);
mp.: 193°C;
^{1}H NMR (500 MHz, CD_3OD, 300 K): δ = 7.79 (d, ^{3}J = 7.5 Hz, 2H, $H_{aryl,Fmoc}$), 7.67 (d, ^{3}J = 7.5 Hz, 2H, $H_{aryl,Fmoc}$), 7.38 (t, ^{3}J = 7.5 Hz, 2H, $H_{aryl,Fmoc}$), 7.31 (t, ^{3}J = 7.5 Hz, 2H, $H_{aryl,Fmoc}$), 4.39 (d, ^{3}J = 6.8 Hz, 2H, $CH_{2,Fmoc}$), 4.23 (t, ^{3}J = 6.8 Hz, CH_{Fmoc}), 3.71 (s, 2H, H_{Gly}), 2.29 (m_c, 6H, $CCH_2CH_2C(O)$), 2,02 (m_c, 6H, $CCH_2CH_2C(O)$) ppm;
^{13}C NMR (125 MHz, CD_3OD, 300 K): δ = 177.1 (*C*(O)OH), 171.4 (*C*(O)NH), 159.2 (O*C*(O)NH), 145.4, 142.7 ($C_{aryl,Fmoc}$), 128.8, 128.2, 126.3, 120.9 ($CH_{aryl,Fmoc}$), 68.1 ($CH_{2,Fmoc}$), 58.8 ($NHC(CH_2CH_2)_3$), 45.2 ($CH_{2,Gly}$), 43.2 (CH_{Fmoc}), 30.7 ($CCH_2CH_2C(O)$), 29.2 ($CCH_2CH_2C(O)$) ppm;
MALDI-ToF MS (DHB): calcd. for $C_{27}H_{30}N_2NaO_9^+$: *m/*z 549.18 $[M+Na]^+$; found: *m/z* 549.02 $[M+Na]^+$;
IR (ATR-IR): $\tilde{\nu}$ = 3339, 2960, 1726, 1701, 1638, 1537, 1450, 1410, 1295, 1236, 1175, 1157, 1102, 1045, 931, 738 cm^{-1}.

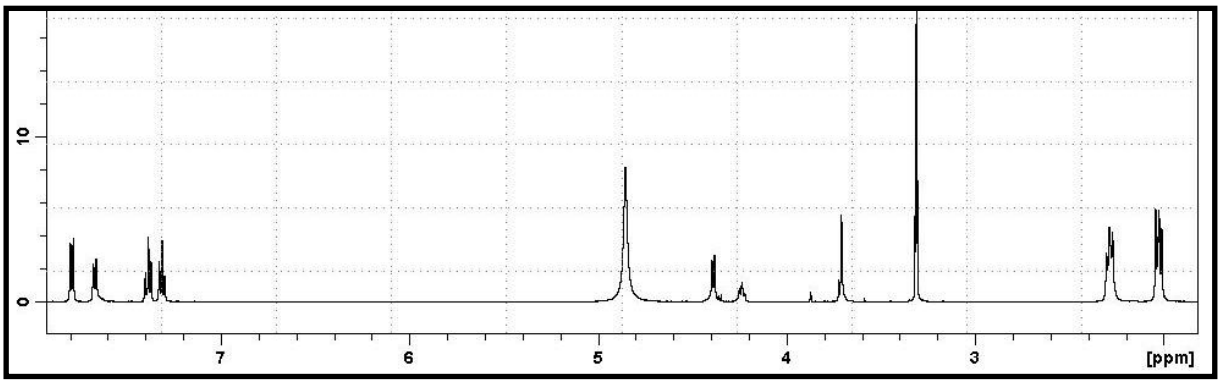


Figure S3: ^{1}H-NMR spectrum (500 MHz, CD_3OD) trivalent wedge **19.**

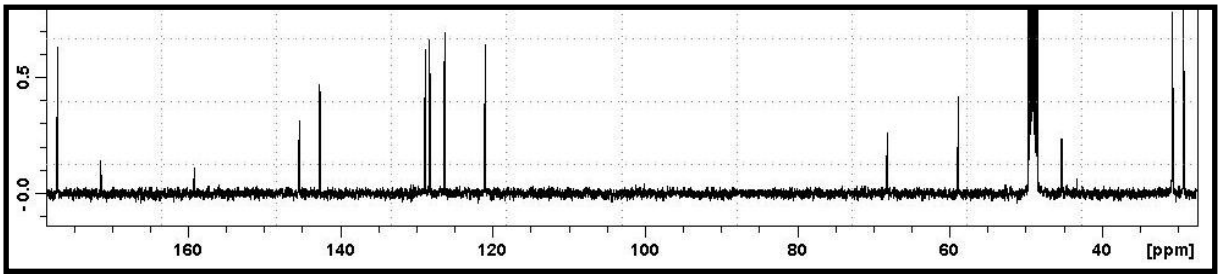
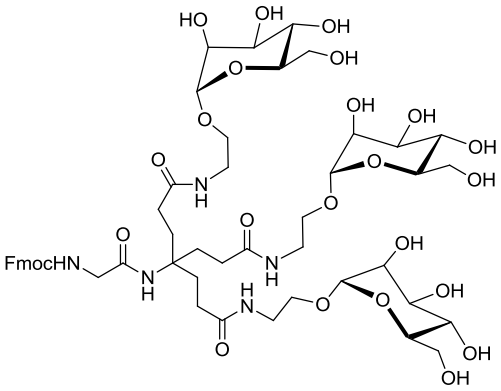


Figure S4: ^{13}C-NMR spectrum (125 MHz, CD_3OD) trivalent wedge **19**.

3-Cascade:*N*-fluorenylmethoxycarbonyl-glycyl-aminomethane[3]:4-oxo-3-aza-hexyl-α-D-mannopyranoside (20)



The triacid **19** (152 mg, 289 µmol) and HATU (497 mg, 1.31 mmol) were dried under vacuum for 1 h. The reactants were dissolved in anhydrous DMF (5 mL), DIPEA (200 µL, 1.17 mmol) was added and it was cooled to 0°C under a nitrogen atmosphere. Parallely in a separate reaction vessel 2-aminoethyl α-D-mannopyranoside (**15**) (200 mg, 896 µmol) was dried under vacuum for 1 h, dissolved in anhydrous DMF (5 mL) and DIPEA (100 µL, 584 µmol) was added. After 20 min the mixture of **15** in anhydrous DMF was added to the cooled mixture of **19** and HATU. The reaction mixture was stirred overnight at ambient temperature. The solvent was removed under reduced pressure and the crude product was purified by reverse phase MPLC (A: methanol, B: water, A: 40 % => 70 %, 120 min) yielding **20** (210 mg, 184 µmol, 64 %) as a colourless lyophyllisate.

R*f*: 0.60 (RP-18, water / methanol, 1 : 2);
$[\alpha]_D^{20}$ = + 39° (*c* = 0.5, methanol);

^{1}H NMR (500 MHz, CD_3OD, 300 K): δ = 7.80 (d, ^3J = 7.5 Hz, 2H, $H_{aryl,Fmoc}$), 7.68 (dd, ^3J = 7.5 Hz, ^4J = 4.8 Hz, 2H, $H_{aryl,Fmoc}$), 7.39 (t, ^3J = 7.5 Hz, 2H, $H_{aryl,Fmoc}$), 7.32 (t, ^3J = 7.5 Hz, 2H, $H_{aryl,Fmoc}$), 4.75 (d, ^3J = 1.7 Hz, 3H, $H1_{man}$), 4.40 (m_c, 2H, $CH_{2,Fmoc}$), 4.26 (t, ^3J = 6.9 Hz, CH_{Fmoc}), 3.84 (dd, ^2J = 11.7 Hz, ^3J = 2.2 Hz, 3H, $H6a_{man}$), 3.82 (dd, ^3J = 1.7 Hz, ^3J = 3.4 Hz, 3H, $H2_{man}$), 3.78-3.67 (m, 14H, $HNCH_2CHHO$, H_{Gly}, $H3_{man}$, $H6b_{man}$), 3.60 (dd~t, ^3J = 9.2 Hz, ^3J = 9.7 Hz, 3H, $H4_{man}$), 3.57-3.48 (m, 6H, $H5_{man}$, $HNCH_2CHHO$), 3.43-3.32 (m, 6H, $HNCH_2CH_2O$), 2.20 (m_c, 6H, $CCH_2CH_2C(O)$), 2.02-1.94 (m, 6H, $CCH_2CH_2C(O)$) ppm;
^{13}C NMR (125 MHz, CD_3OD, 300 K): δ = 176.0 (3 *C*(O)NH), 171.5 (*C*(O)NH), 159.4 (O*C*(O)NH), 145.2, 142.6 ($C_{aryl,Fmoc}$), 128.8, 128.2, 126.2, 121.0 ($CH_{aryl,Fmoc}$), 101.7 ($C1_{man}$), 74.8 ($C5_{man}$), 72.6 ($C3_{man}$), 72.1 ($C2_{man}$), 68.8 ($C4_{man}$), 68.3 ($CH_{2,Fmoc}$), 67.3 ($HNCH_2CH_2O$), 63.0 ($C6_{man}$), 59.4 ($NHC(CH_2CH_2)_3$), 48.4 (CH_{Fmoc}), 45.3 ($CH_{2,Gly}$), 40.4 ($HNCH_2CH_2O$), 31.8 ($CCH_2CH_2C(O)$), 31.3 ($CCH_2CH_2C(O)$) ppm;
HR-ESI MS: calcd. for $C_{51}H_{75}N_5NaO_{24}$: *m/z* 1164,4694 $[M+Na]^+$; found: *m/z* 1164,4686 $[M+Na]^+$;
IR (ATR-IR): $\tilde{\nu}$ = 3286, 2927, 2885, 1638, 1542, 1449, 1243, 1175, 1157, 1131, 1052, 804, 739 cm^{-1}.

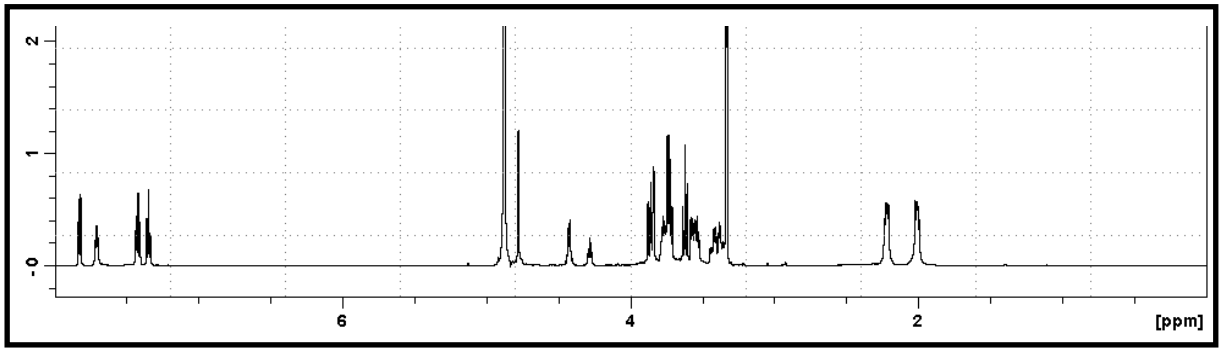


Figure S5: ^{1}H-NMR spectrum (500 MHz, CD_3OD) trivalent cluster mannoside **20**.

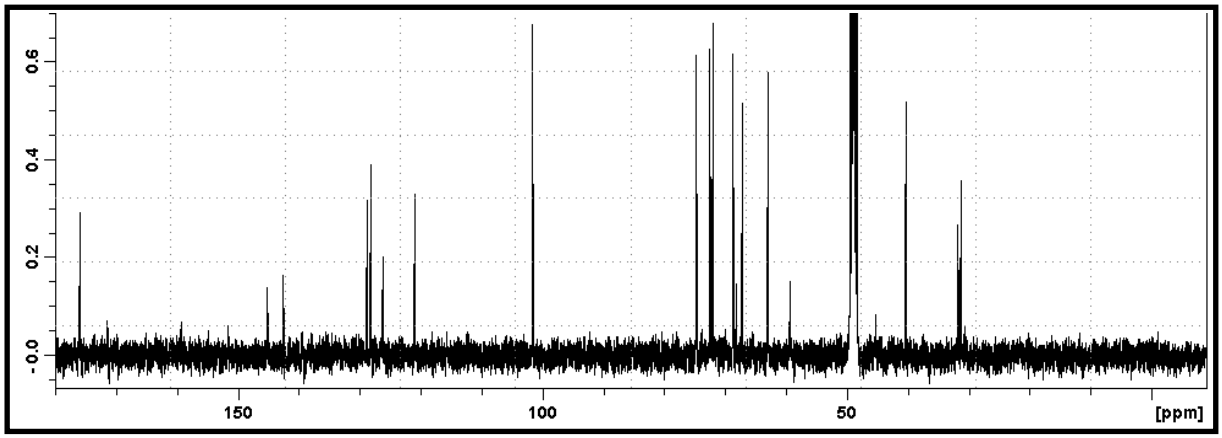


Figure S6: ^{13}C-NMR spectrum (125 MHz, CD_3OD) trivalent cluster mannoside **20**.

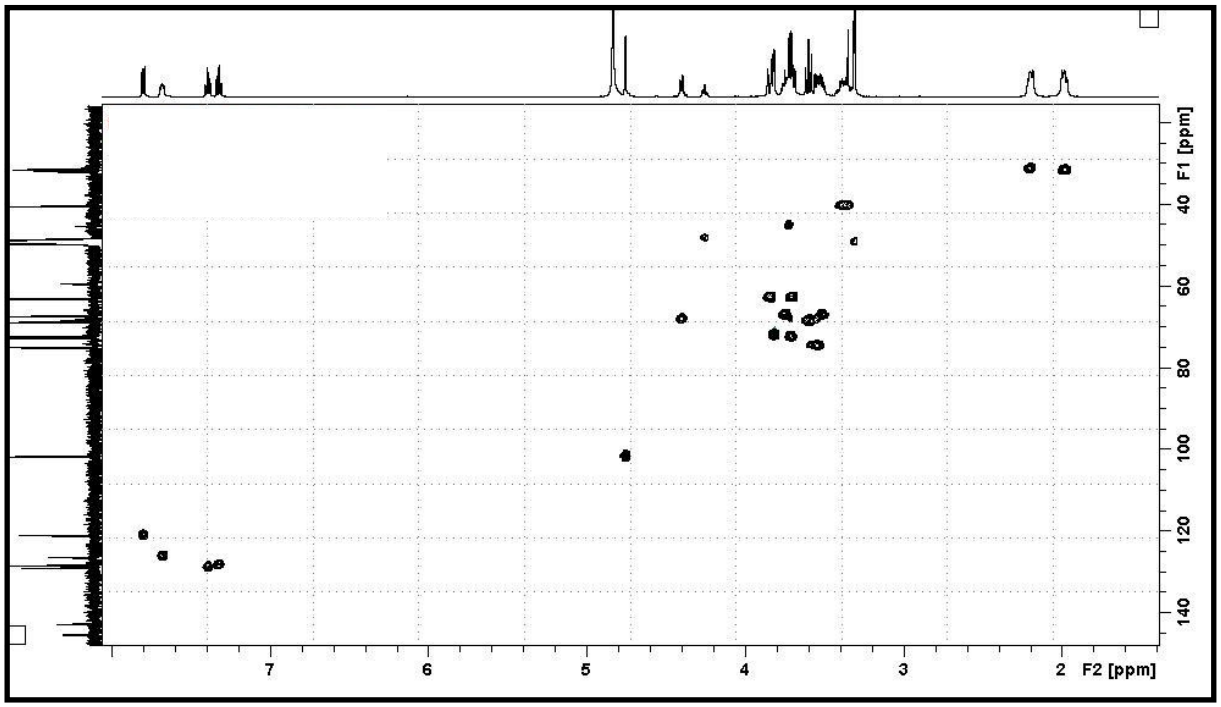
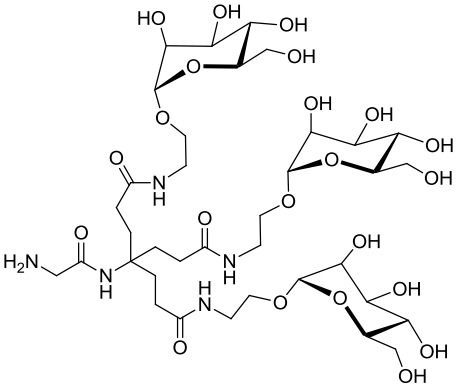


Figure S7: ^{13}C-^{1}H-HSQC NMR spectrum (500 MHz, CD_3OD) trivalent cluster mannoside **20**.

3-Cascade:glycyl-aminomethane[3]:4-oxo-3-aza-hexyl-α-D-mannopyranoside (21)



Fmoc-protected trivalent mannoside **20** (14 mg, 12.3 µmol) was dissolved in DMF (1 mL) and piperidine (100 µL) was added. The reaction mixture was stirred at ambient temperature. After 2 h additional piperidine (100 µL) was added and the reaction mixture was stirred for another 30 min. The solvent was removed under reduced pressure and the crude product was dissolved in water (4 mL) and extracted with dichloromethane (2 mL). The aqueous layer was lyophillised and the lyophyllisate was subjected to size exclusion chromatography (Sephadex LH-20, methanol) yielding **21** (10 mg, 10.9 µmol, 89 %) as a colourless lyophyllisate.

$[\alpha]_D^{20}$ = + 33° (*c* = 0.175, methanol);

^{1}H NMR (600 MHz, D_2O, 298 K): δ = 4.78 (d, ^3J = 1.7 Hz, 3H, $H1_{man}$), 3.85 (dd, ^3J = 1.7 Hz, ^3J = 3.4 Hz, 3H, $H2_{man}$), 3.79 (dd, ^2J = 12.2 Hz, ^3J = 2.1 Hz, 3H, $H6a_{man}$), 3.72-3.64 (m, 14H, $H3_{man}$, $HNCH_2CHHO$, $H6b_{man}$, H_{Gly}), 3.56 (dd~t, ^3J = 9.6 Hz, 3H, $H4_{man}$), 3.53-3.49 (m, 6H, $H5_{man}$, $HNCH_2CHHO$), 3.38 (ddd, ^2J = 14.3 Hz, ^3J = 7.3 Hz, ^3J = 3.8 Hz, 3H, $HNCHHCH_2O$), 3.27 (ddd, ^2J = 14.3 Hz, ^3J = 5.9 Hz, ^3J = 3.8 Hz, 3H, $HNCHHCH_2O$), 2.18-2.12 (m, 6H, $CCH_2CH_2C(O)$), 1.93-1.87 (m, 6H, $CCH_2CH_2C(O)$) ppm;

^{13}C NMR (150 MHz, D_2O, 298 K): δ = 175.9 (3 *C*(O)NH), 166.0 (*C*(O)NH), 99.7 ($C1_{man}$), 72.9 ($C5_{man}$), 70.6 ($C3_{man}$), 70.0 ($C2_{man}$), 66.7 ($C4_{man}$), 65.8 ($HNCH_2CH_2O$), 60.9 ($C6_{man}$), 58.6 ($(NHC(CH_2CH_2)_3)$), 40.7 ($CH_{2,Gly}$), 39.0 ($HNCH_2CH_2O$), 30.0 ($CCH_2CH_2C(O)$), 30.0 ($CCH_2CH_2C(O)$) ppm;

HR-ESI MS: calcd. for $C_{36}H_{66}N_5O_{22}$: *m/z* 920.4194 $[M+H]^+$; found: *m/z* 920.4160 $[M+H]^+$;

MALDI-ToF MS (DHB): calcd. for $C_{36}H_{65}N_5NaO_{22}$: *m/z* 942.40 $[M+Na]^+$; found: *m/z* 942.20 $[M+Na]^+$;

IR (ATR-IR): $\tilde{\nu}$ = 3282, 2935, 1634, 1634, 1575, 1558, 1456, 1350, 1204, 1130, 1055, 1030, 973, 806, 676, 580 cm^{-1}.

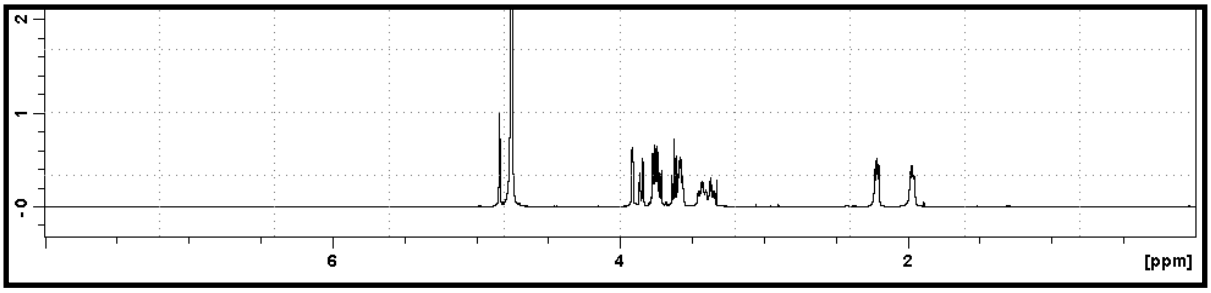


Figure S8: ^{1}H-NMR spectrum (600 MHz, D_2O) trivalent cluster mannoside **21**.

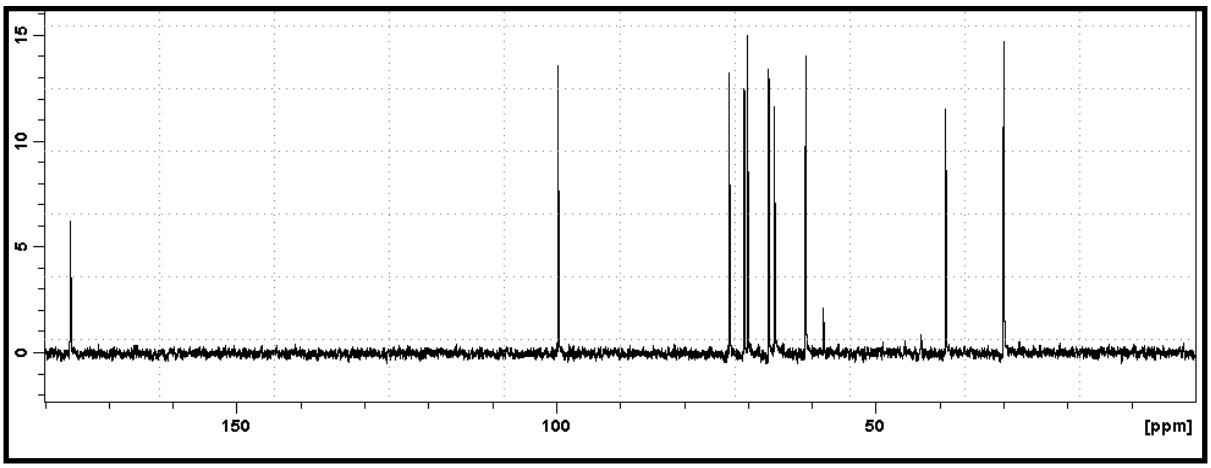


Figure S9: ^{13}C-NMR spectrum (150 MHz, D_2O) trivalent cluster mannoside **21**.

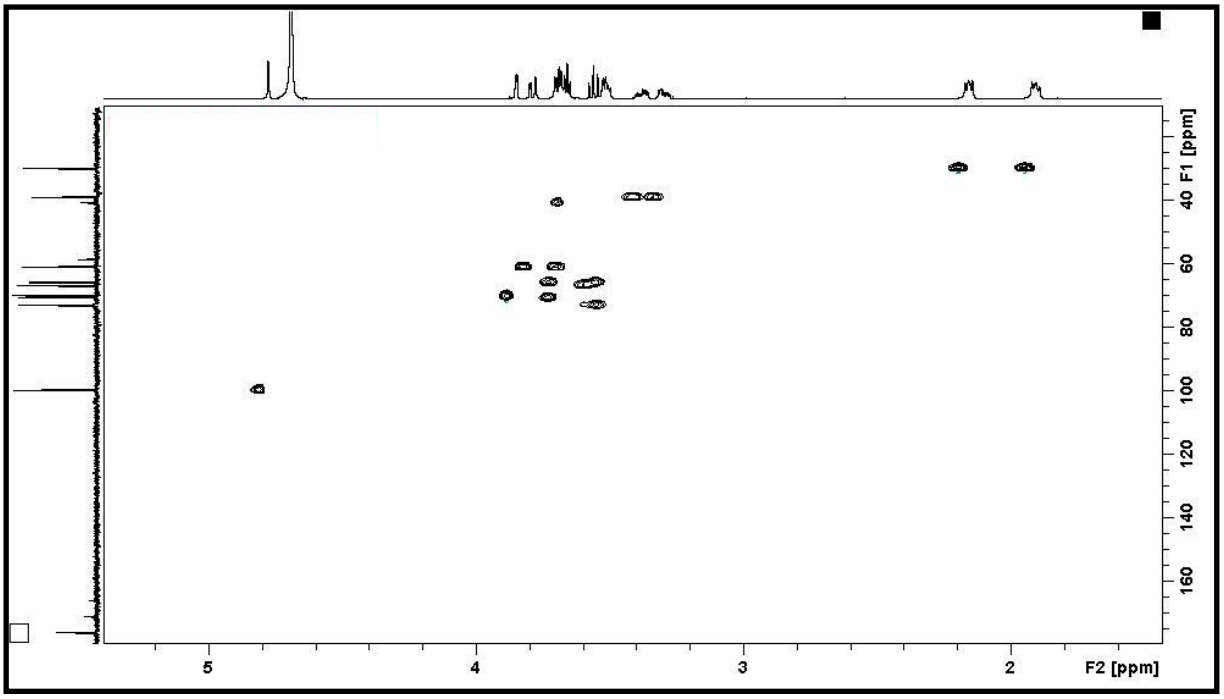
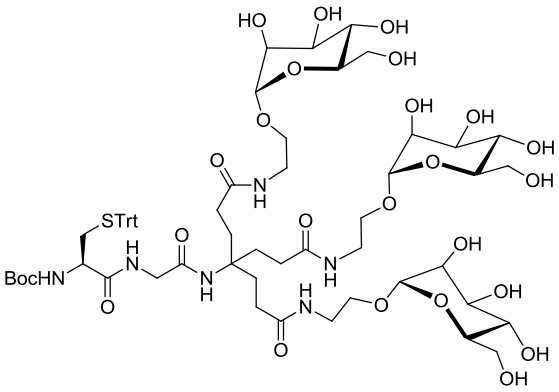


Figure S10: ^{13}C-^{1}H-HSQC NMR spectrum (600 MHz, D_2O) trivalent cluster mannoside **21.**

3-Cascade:*N*-(*tert*-butyloxycarbonyl)-*S*-(triphenylmethyl)-L-cysteinyl-glycyl-amino-methane[3]:4-oxo-3-aza-hexyl-α-D-mannopyranoside (22)



The trivalent mannoside **21** (120 mg, 130 µmol) and HATU (95 mg, 250 µmol) were dried under vacuum for 2 h. The reactants were dissolved in anhydrous DMF (3 mL), DIPEA (50 µL, 292 µmol) was added and it was cooled to 0°C under a nitrogen atmosphere. Parallely in another reaction vessel Boc-Cys(Trt)-OH (58 mg, 125 µmol) was dried under vacuum for 2 h, dissolved in anhydrous DMF (2 mL) and DIPEA (36 µL, 210 µmol) was added. After 20 min the mixture of Boc-Cys(Trt)-OH in anhydrous DMF was added to the cooled mixture of **21** and HATU. The reaction mixture was stirred overnight at ambient temperature under a nitrogen atmosphere. The solvent was removed under reduced pressure and the crude product was purified by reverse phase MPLC (A: methanol, B: water, A: 50 % => 95 %, 120 min) yielding **22** (75 mg, 54.2 µmol, 58 %) as a colourless lyophyllisate.

R*f*: 0.21 (RP-18, water / methanol, 1 : 2);
$[\alpha]_D^{20}$ = + 40° (*c* = 0.5, methanol);

^{1}H NMR (500 MHz, CD_3OD, 300 K): δ = 7.38 (d, ^3J = 7.8 Hz, 6H, $H_{aryl,Trt}$), 7.29 (m_c, 6H, $H_{aryl,Trt}$), 7.23 (m_c, 3H, $H_{aryl,Trt}$), 4.76 (d, ^3J = 1.7 Hz, 3H, $H1_{man}$), 4.08 (dd, ^3J = 5.2 Hz, ^3J = 8.4 Hz, 1H, $H_{Cys,\alpha}$), 3.84 (dd, ^2J = 11.8 Hz, ^3J = 2.0 Hz, 3H, $H6a_{man}$), 3.82 (dd, ^3J = 1.7 Hz, ^3J = 3.5 Hz, 3H, $H2_{man}$), 3.79-3.66 (m, 11H, $H_{Gly,\alpha a}$, $HNCH_2CHHO$, $H3_{man}$, $H6b_{man}$, $H_{Gly,\alpha b}$), 3.61 (dd, ^3J = 10.6 Hz, ^3J = 9.6 Hz, 3H, $H4_{man}$), 3.57-3.49 (m, 6H, $H5_{man}$, $HNCH_2CHHO$), 3.45-3.32 (m, 6H, $HNCH_2CH_2O$), 2.65 (dd, ^2J = 12.3 Hz, ^3J = 5.2 Hz, 1H, $H_{Cys,\beta a}$), 2.57 (dd, ^2J = 12.3 Hz, ^3J = 8.4 Hz, 1H, $H_{Cys,\beta b}$), 2.19-2.11 (m, 6H, $CCH_2CH_2(O)$), 2.04-1.96 (m, 3H, $CCH_2CHHC(O)$), 1.96-1.86 (m, 3H, $CCH_2CHHC(O)$), 1.46 (s, 9H, H_{Boc}) ppm;
^{13}C NMR (125 MHz, CD_3OD, 300 K): δ = 176.0 (3 *C*(O)NH), 174.6 ($C(O)NH_{Gly,Cys}$), 171.0 ($C(O)NH_{trival,Gly}$), 159.1 ($OC(O)NH$), 146.0 ($C_{aryl,Trt}$), 130.7, 129.0, 128.0 ($CH_{aryl,Trt}$), 101.7 ($C1_{man}$), 74.8 ($C5_{man}$), 72.6 ($C3_{man}$), 72.1 ($C2_{man}$), 68.8 ($C4_{man}$), 67.9 (($C(CH_3)_3$)), 67.2 ($HNCH_2CH_2O$), 63.0 ($C6_{man}$), 59.3 ($NHC(CH_2CH_2)_3$), 55.2 ($CH_{Cys,\alpha}$), 44.7 ($CH_{2,Gly}$), 40.5 ($HNCH_2CH_2O$), 35.6 ($CH_{2,Cys,\beta}$), 31.8 ($CCH_2CH_2C(O)$), 31.3 ($CCH_2CH_2C(O)$), 28.8 (($C(CH_3)_3$)) ppm;
HR-ESI MS: calcd. for $C_{63}H_{92}N_6NaO_{25}S$: *m/z* 1387.5725 $[M+Na]^+$; found: *m/z* 1387.5541 $[M+Na]^+$;
IR (ATR-IR): $\tilde{\nu}$ = 3297, 2933, 1645, 1540, 1445, 1367, 1249, 1134, 1055, 975, 806, 742, 700 cm^{-1}.

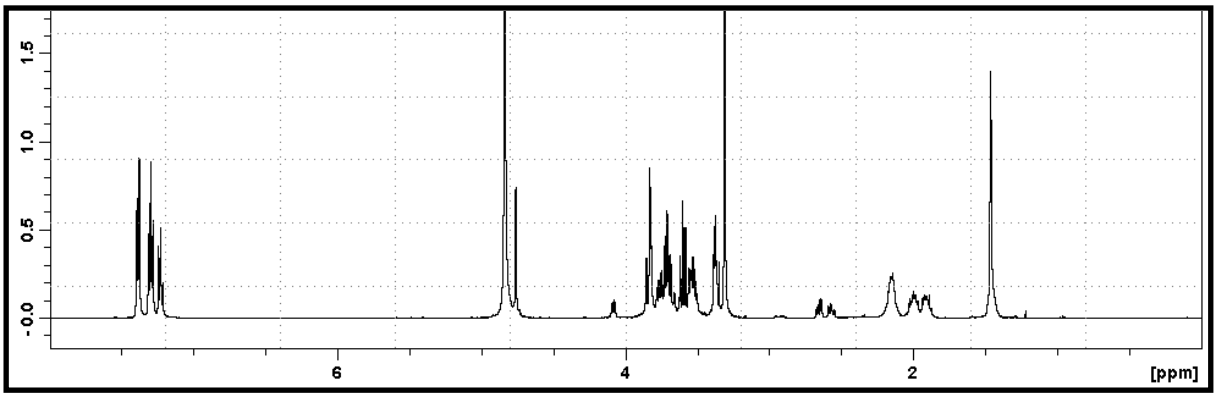


Figure S11: ^{1}H-NMR spectrum (500 MHz, CD_3OD) trivalent cluster mannoside **22**.

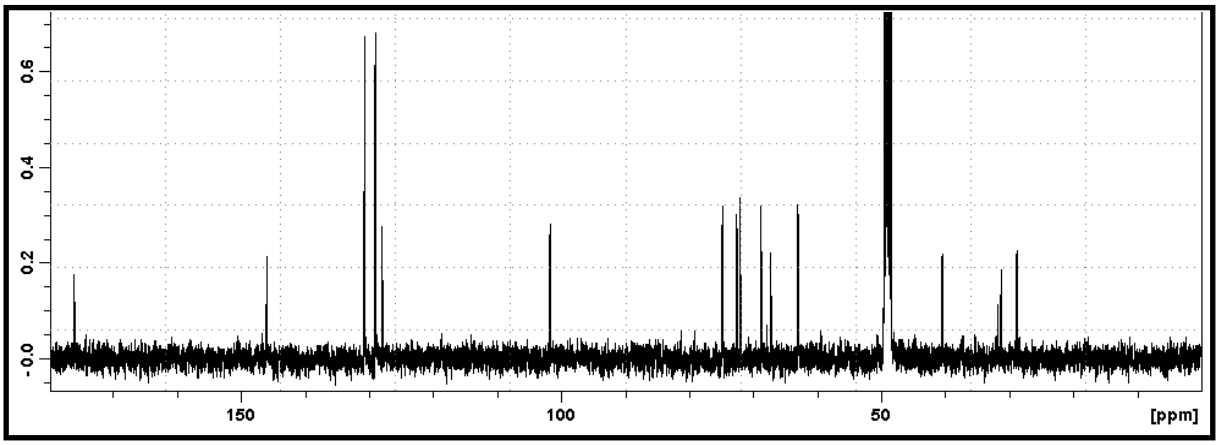


Figure S12: ^{13}C-NMR spectrum (125 MHz, CD_3OD) trivalent cluster mannoside **22**.

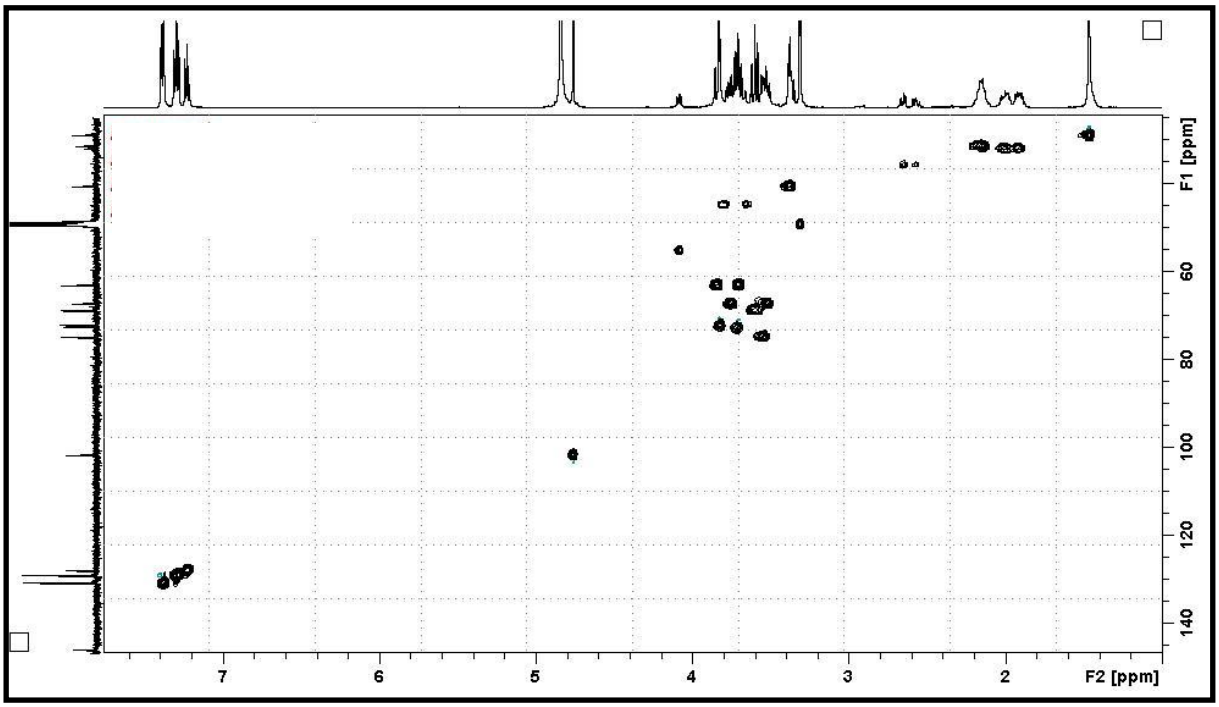
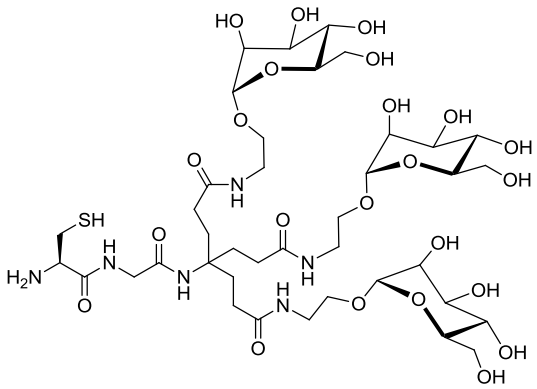


Figure S13: ^{13}C-^{1}H-HSQC NMR spectrum (500 MHz, CD_3OD) trivalent cluster mannoside **22.**

3-Cascade:L-cysteinyl-glycyl-aminomethane[3]:4-oxo-3-aza-hexyl-α-D-mannopyranoside (23)



The protected glycocluster **22** (66.0 mg, 48.3 µmol) was dried under vacuum for 1 h. It was dissolved in anhydrous dichloromethane (2 mL) under a nitrogen atmosphere. TFA (50 µL) and triethylsilane (46 µL, 290 µmol) were added and the reaction mixture was stirred at room temperature. After 3.5 h and 30 min later additional TFA (100 µL each) was added. After another 30 min RP-TLC (water / methanol, 1 : 3) indicated complete consumption of starting material and the reaction mixture was diluted with dichloromethane (5 mL). Amberlyst A-21 (3 g) was added and the mixture was stirred for another hour. The ion exchange resin was filtered off and it was washed with dichloromethane and methanol/dichloromethane (1:1) (10 mL each). The filtrate was evaporated and the residue was dissolved in water and lyophyllised. The lyophyllisate was subjected to size exclusion chromatography (Sephadex LH-20, methanol) yielding **23** (32 mg, 31.3 µmol, 65 %) as a colourless lyophyllisate.

$[\alpha]_D^{20}$ = + 23° (*c* = 0.15, methanol);

^{1}H NMR (500 MHz, D_2O, 300 K): δ = 4.75 (d, ^3J = 1.7 Hz, 3H, $H1_{man}$), 4.12-4.06 (m, 1H, $H_{Cys,\alpha}$), 3.83 (dd, ^3J = 1.7 Hz, ^3J = 3.4 Hz, 3H, $H2_{man}$), 3.83 (s, 2H, H_{Gly}), 3.76 (dd, ^2J = 12.2 Hz, ^3J = 2.1 Hz, 3H, $H6a_{man}$), 3.68 (dd, ^3J = 3.4 Hz, ^3J = 9.3 Hz, 3H, $H3_{man}$), 3.68-3.64 (m, 3H, $HNCH_2CHHO$), 3.64 (dd, ^2J = 12.1 Hz, ^3J = 5.9 Hz, 3H, $H6b_{man}$), 3.53 (dd, ^3J = 9.3 Hz, ^3J = 9.9 Hz, 3H, $H4_{man}$), 3.52-3.46 (m, 6H, $H5_{man}$, $HNCH_2CHHO$), 3.35 (m_c, 3H, $HNCHHCH_2O$), 3.27 (m_c, 3H, $HNCHHCH_2O$) 3.18 (m_c, 1H, $H_{Cys,\beta a}$), 3.08-3.00 (m_c, 1H, $H_{Cys,\beta b}$), 2.13 (t, ^3J = 8.1 Hz, 6H, $CCH_2CH_2C(O)$), 1.90-1.84 (m, 6H, $CCH_2CH_2C(O)$) ppm;
^{13}C NMR (125 MHz, D_2O, 298 K): δ = 175.9 (3 *C*(O)NH), 169.4 (*C*(O)NH), 99.7 ($C1_{man}$), 72.9 ($C5_{man}$), 70.6 ($C3_{man}$), 70.0 ($C2_{man}$), 66.7 ($C4_{man}$), 65.8 ($HNCH_2CH_2O$), 60.9 ($C6_{man}$), 58.9 ($NHC(CH_2CH_2)_3$), 54.4 ($CH_{Cys,\alpha}$), 42.7 ($CH_{2,Gly}$), 39.1 ($CH_{2,Cys,\beta}$), 39.0 ($HNCH_2CH_2O$), 30.0 ($CCH_2CH_2C(O)$), 30.0 ($CCH_2CH_2C(O)$) ppm;
HR-ESI MS: calcd. for $C_{78}H_{140}N_{12}O_{46}S_2$ (disulfide): *m/z* 1022.4208 $[M+2H]^{2+}$; found: *m/z* 1022.4237 $[M+2H]^{2+}$;
MALDI-ToF MS (DHB): calcd. for $C_{39}H_{70}KN_6O_{23}S$: *m/z* 1061.38 $[M+K]^+$; found: *m/z* 1061.98 $[M+Na]^+$; calcd. for $C_{39}H_{70}N_6NaO_{23}S$: *m/z* 1045.41 $[M+Na]^+$; found: *m/z* 1046.02 $[M+Na]^+$;
IR (ATR-IR): $\tilde{\nu}$ = 3302, 2936, 1645, 1554, 1430, 1202, 1133, 1057, 976, 803, 679, 580, 510 cm^{-1}.

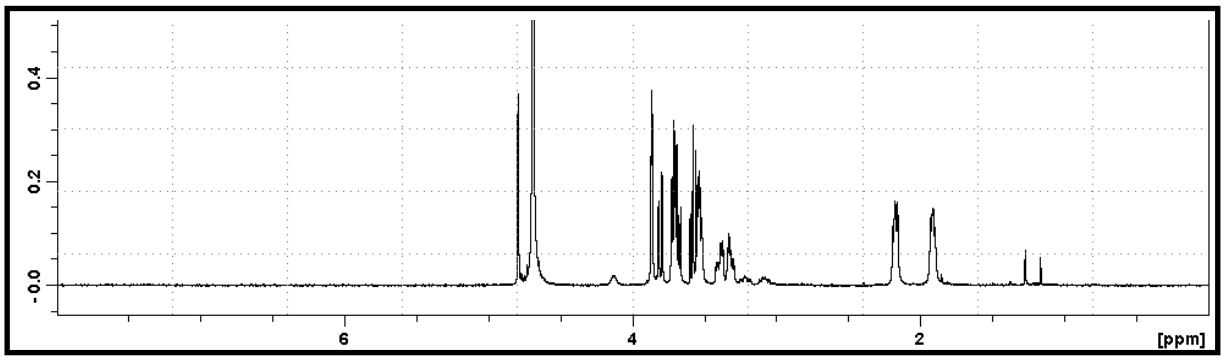


Figure S14: ^{1}H-NMR spectrum (500 MHz, D_2O) trivalent cluster mannoside **23**.

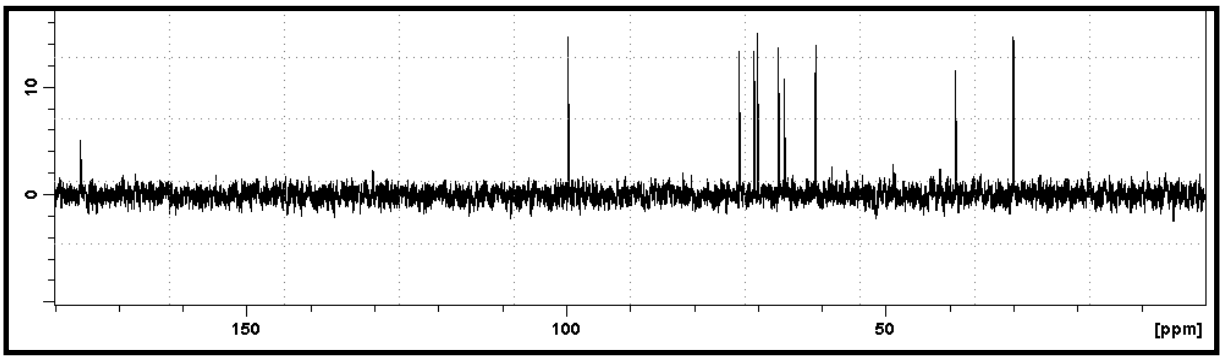


Figure S15: ^{13}C-NMR spectrum (125 MHz, D_2O) trivalent cluster mannoside **23**.

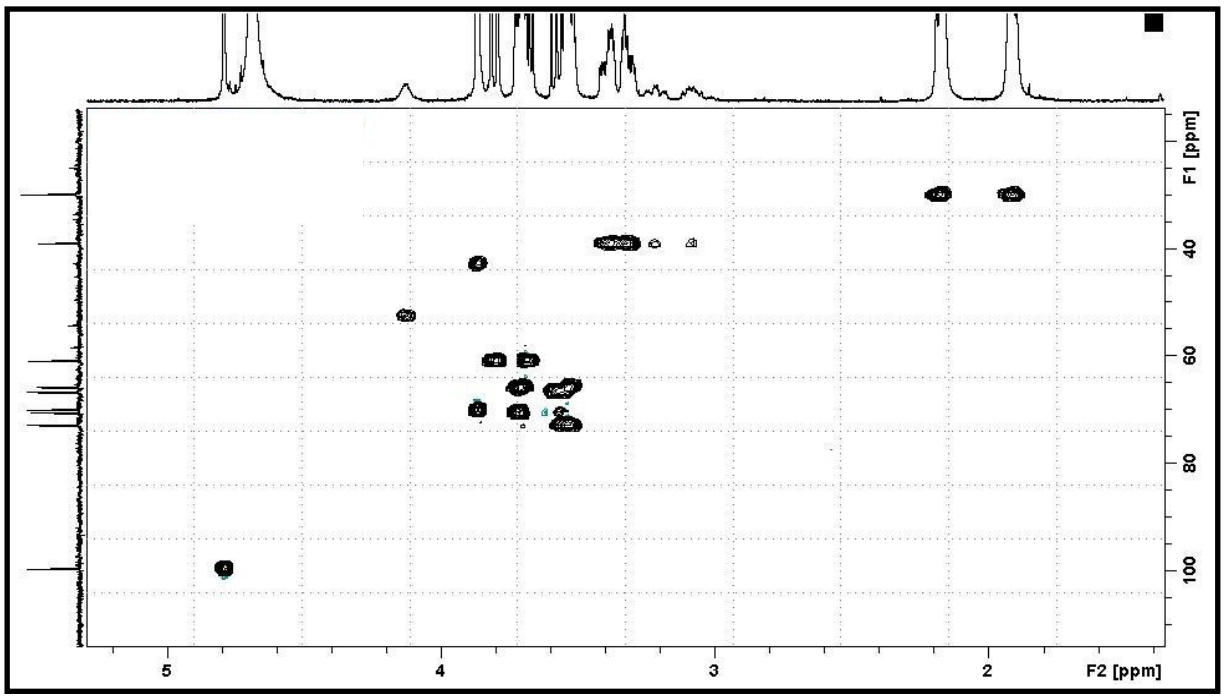


Figure S16: ^{13}C-^{1}H-HSQC NMR spectrum (500 MHz, D_2O) trivalent cluster mannoside **23**.

4. Coupling on Self-Assembled Monolayers (SAMs)

4.1 Preparation and activation of SAMs

A disposable 64-well gold plate (Applied Biosystems) was cleaned with Piranha solution (12 mL, 5:1 conc. H_2SO_4/ 30% H_2O_2) for 30 min, rinsed with distilled water, ethanol and dried under a stream of nitrogen. A DMSO solution of carboxylic acid-terminated [HS-$(CH_2)_{17}$-$(OC_2H_4)_6$-OCH_2-COOH] and tri(ethylene glycol) [HS-$(CH_2)_{17}$-$(OC_2H_4)_3$-OH] alkanethiols (final concentration 0.4 mg/mL, molar ratio 1:4) was applied on the plate (~1 µL per well) and left overnight at RT to form a mixed SAM. The plate was washed with ethanol and dried under nitrogen. The carboxylic group was activated with EDC and PFP (final concentrations 0.180 M and 0.174 M, respectively) in dry DMF for 0.5-1 h, followed by washing with water and ethanol and drying as above. The product formation was analysed by MALDI-ToF MS.

4.2 General coupling method of amino acids Cys, Gly, Phe and β-Ala

General Native Chemical Ligation procedure of Cys 7

An aqueous buffer containing 5 M guanidine hydrochloride (Gn·HCl), 75 mM Na_2HPO_4, 20% MeCN was added to imidazole (2.5 M) and stock solution of amino acid(s) and Tris(2-carboxyethyl)phosphine hydrochloride (TCEP, 7.5 eq. relative to Cys). The pH was adjusted to 7 using 2 M HCl. For Cys concentrations other than 4 mM, the solution was diluted with Gn·HCl (75 mM Na_2HPO_4) to obtain the different cysteine concentrations. The reactions were carried out overnight and were directly analysed.

Reactions in absence of imidazole
The Cys or Gly solutions without imidazole were prepared as above with the absence of imidazole and MeCN.

Coupling reactions of Gly
The coupling reactions of Gly were carried out in the same way as above but in the absence of TCEP.

Reactions in the presence of MPAA
According to the general procedure (see above) with the difference that the solution was used without imidazole and a saturated 4-mercaptophenylacetic acid (MPAA) solution was generated. Illustrated in Fig. 3 in the main document with the products **7** and **8**.

Cys and Gly experiments in different ratios
An aqueous buffer containing 5 M Guanidine (Gn·HCl), 75 mM Na_2HPO_4, 20% MeCN, and 2.5 M imidazole HCl (*for reactions in absence of imidazole, no imidazole HCl and MeCN was used*) was premixed and the pH was adjusted to 7 using 5 M NaOH. Cysteine (4 mM) and Gly (4-400 mM) with TCEP (7.5 eq. relative to Cys) were applied in various different ratios (1:1 → 1:100). The reactions were carried out overnight and were directly analysed.

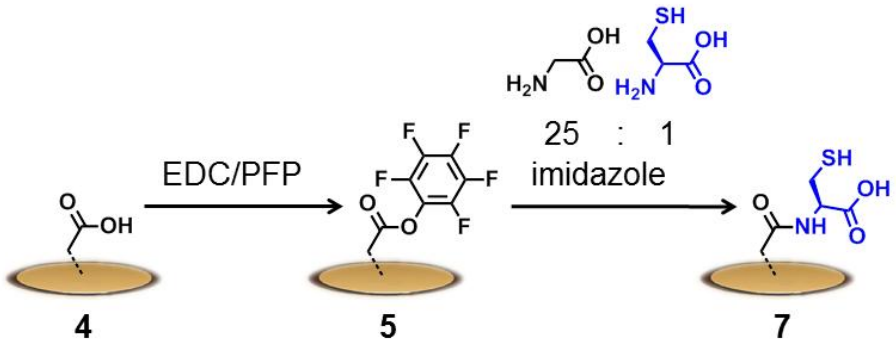


Figure S17: The chemoselective coupling on PFP activated gold plates at a 25 times excess of Gly over Cys.

4.3 Method for Native Chemical Ligation (NCL) of *N*-terminal Cys-peptides (9, 11) and –carbohydrates (13, 23)

Unless otherwise stated, an aqueous buffer containing 5 M Guanidine (Gn·HCl), 75 mM Na_2HPO_4, 20% MeCN, and 2.5 M imidazole HCl (*for reactions in absence of imidazole, no imidazole HCl and MeCN was used*) was premixed and the pH was adjusted to 7 using 5 M NaOH. Cysteine-derivatives (**9**, **11**, **13**, **23**) and/or amino-derivatives (**10**, **12**, **14**, **15**, **21**) with TCEP (7.5 eq. relative to Cys-derivative; *reactions of only the amino-derivative were carried out without TCEP*) were added in various concentrations. Reactions were carried out overnight and were directly analysed.

5. MALDI-ToF MS spectra

5.1 Amino acids

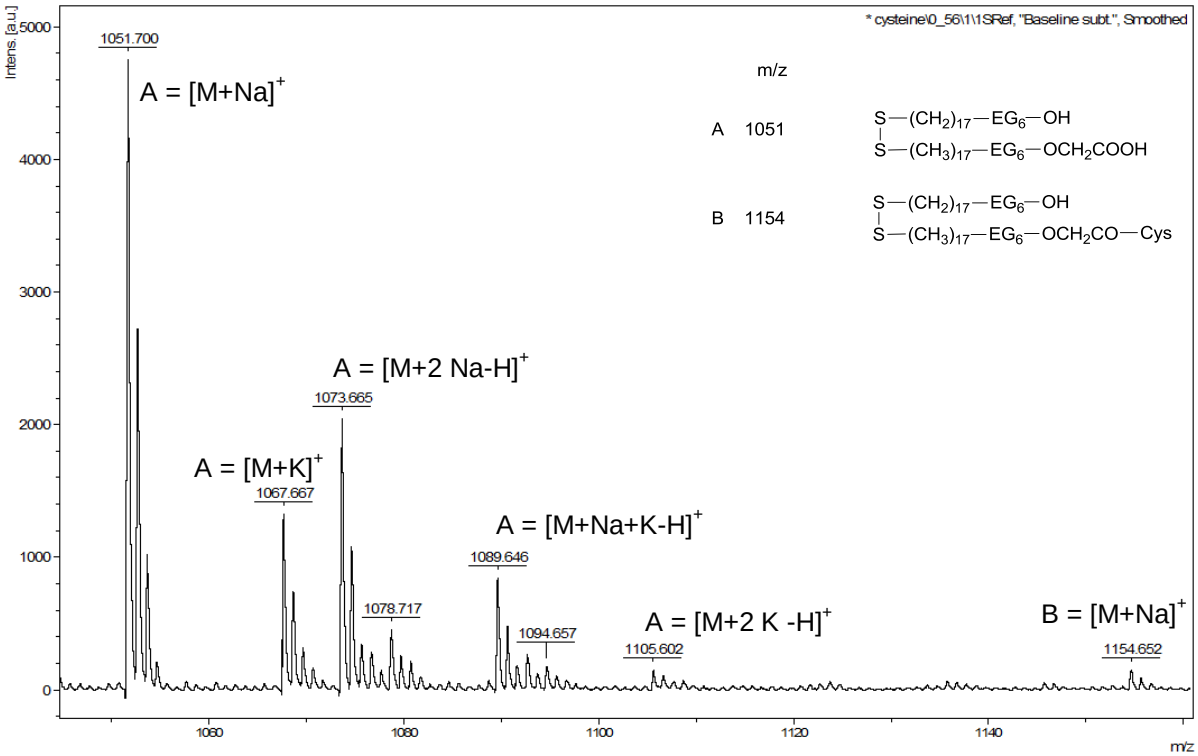


Figure S18 Coupling of Cys at a concentration of 0.25 mM. The different Na and K ions and salts are only here indicated for clarification. **A** is the disulphide of **4** and **B** is the disulphide of **7** as it is most commonly detected by MALDI-ToF MS on SAMs.

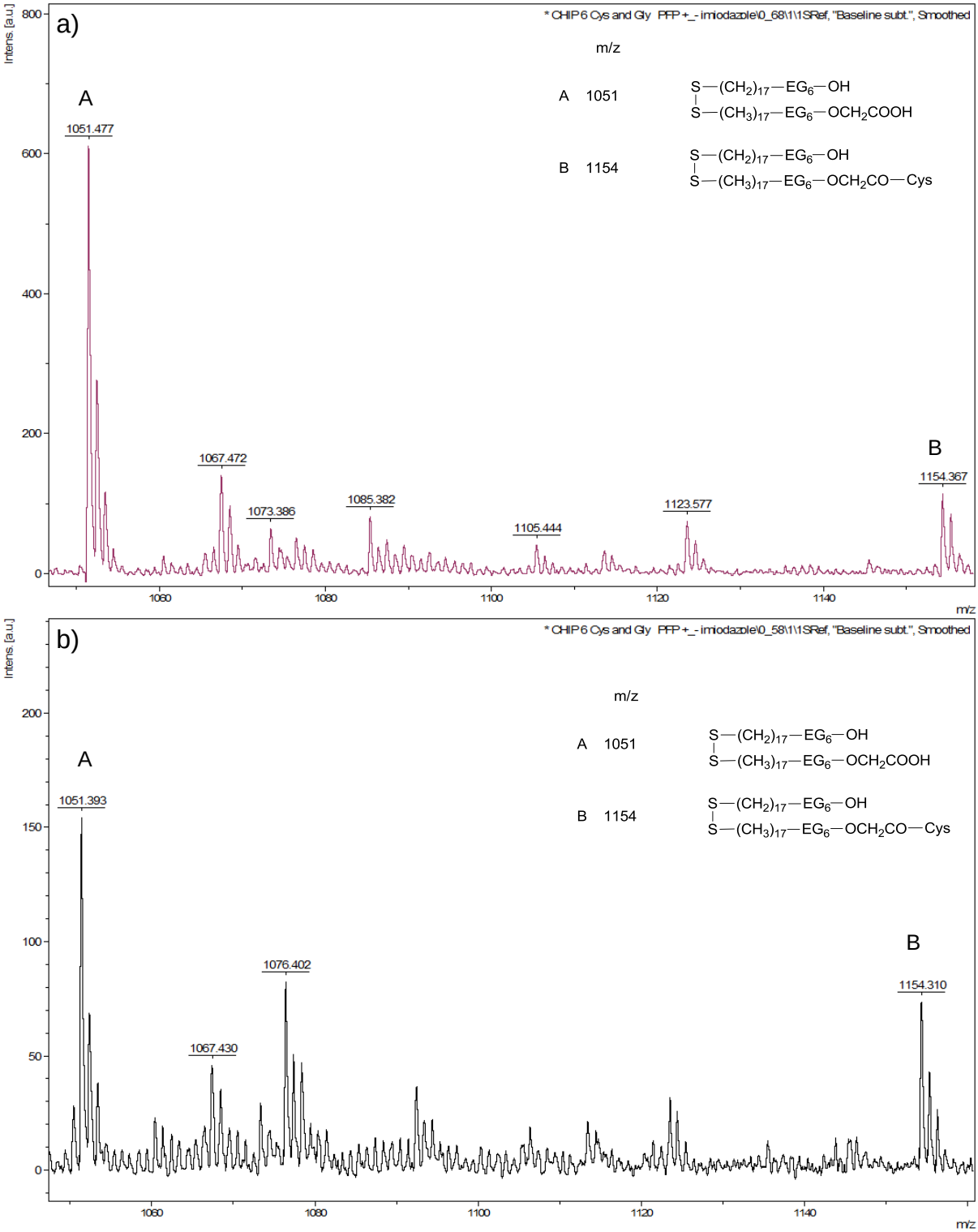


Figure S19 Coupling of Cys (2 mM) in presence (a) and absence (b) of imidazole. **A** is the disulphide of **4** and **B** is the disulphide of **7**.

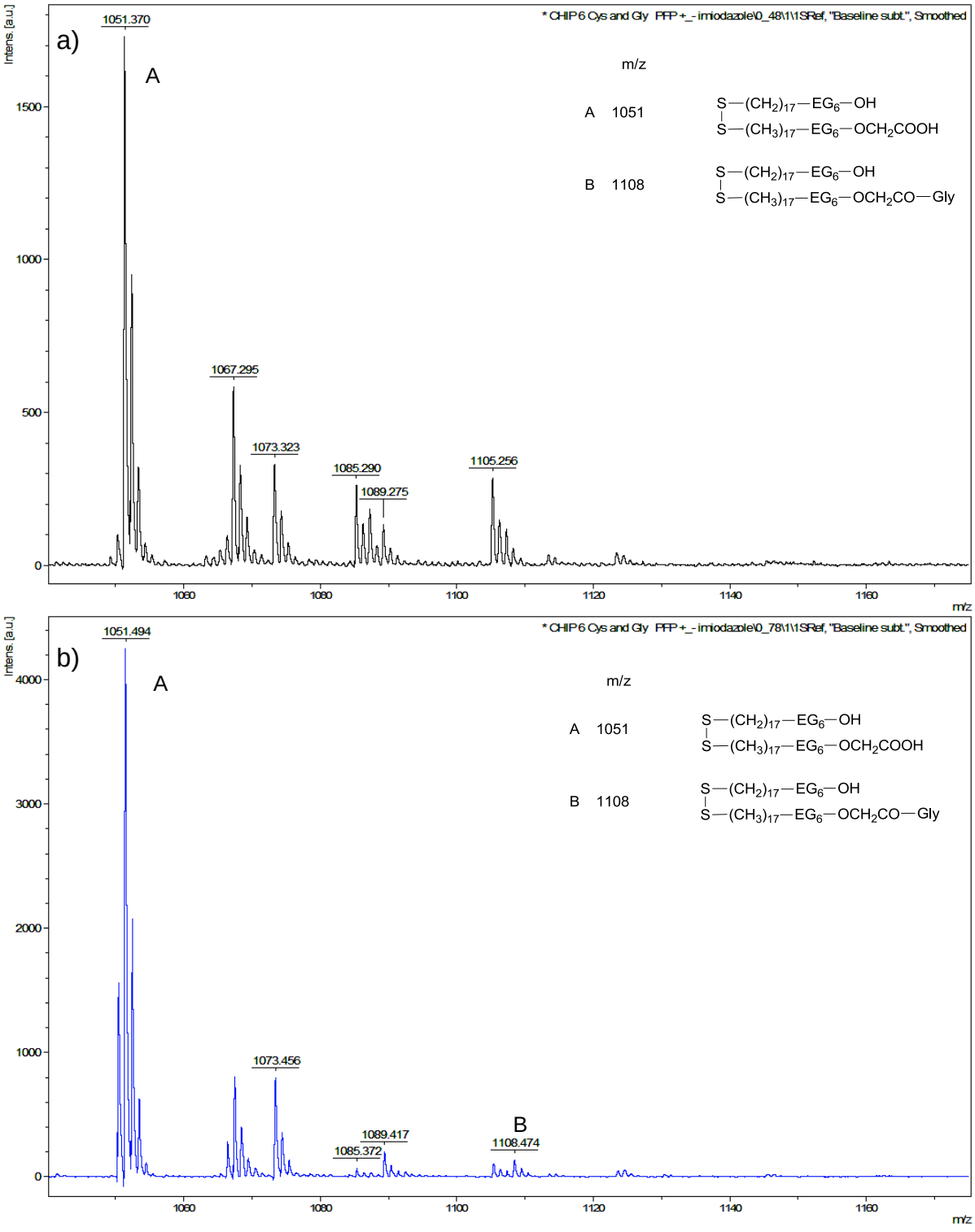


Figure S20 No coupling of Gly (2 mM) is observed in presence of imidazole (a) but in absence of imidazole (b).

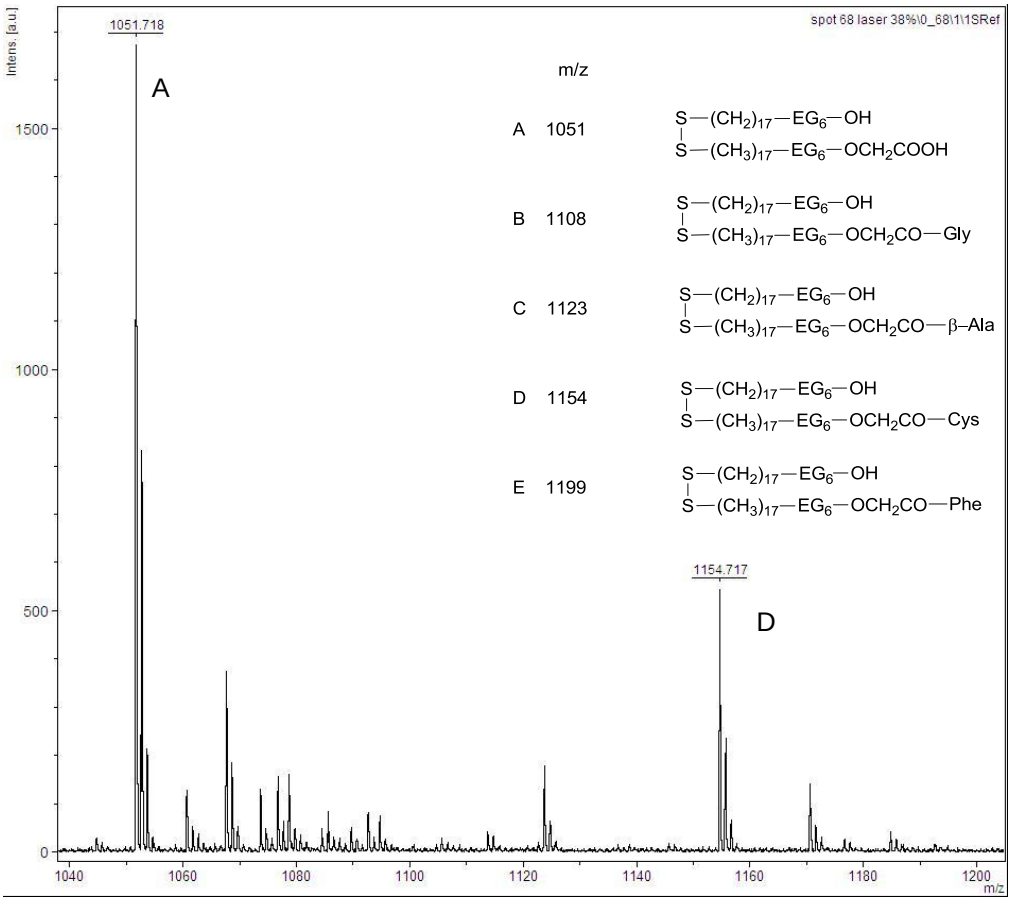


Figure S21 The competitive coupling reaction of β-Ala, Gly, Phe and Cys (all 1 mM) in presence of imidazole. Only the Cys product was observed. The peak at 1123 m/z was detected in the negative control as well and is therefore not the β-Ala product. **A** is the disulphide of **4** and **D** is the disulphide of **7**. This reaction is analogous to Fig. 3 in the main document with difference of imidazole instead of MPAA and additional β-Ala.

	m/z	
A	1051	S—$(CH_2)_{17}$—EG_6—OH S—$(CH_3)_{17}$—EG_6—OCH_2COOH
B	1108	S—$(CH_2)_{17}$—EG_6—OH S—$(CH_3)_{17}$—EG_6—OCH_2CO—Gly
C	1154	S—$(CH_2)_{17}$—EG_6—OH S—$(CH_3)_{17}$—EG_6—OCH_2CO—Cys
D	1199	S—$(CH_2)_{17}$—EG_6—OH S—$(CH_3)_{17}$—EG_6—OCH_2CO—Phe
E	1201	S—$(CH_2)_{17}$—EG_6—OH S—$(CH_3)_{17}$—EG_6—OCH_2CO—MPAA

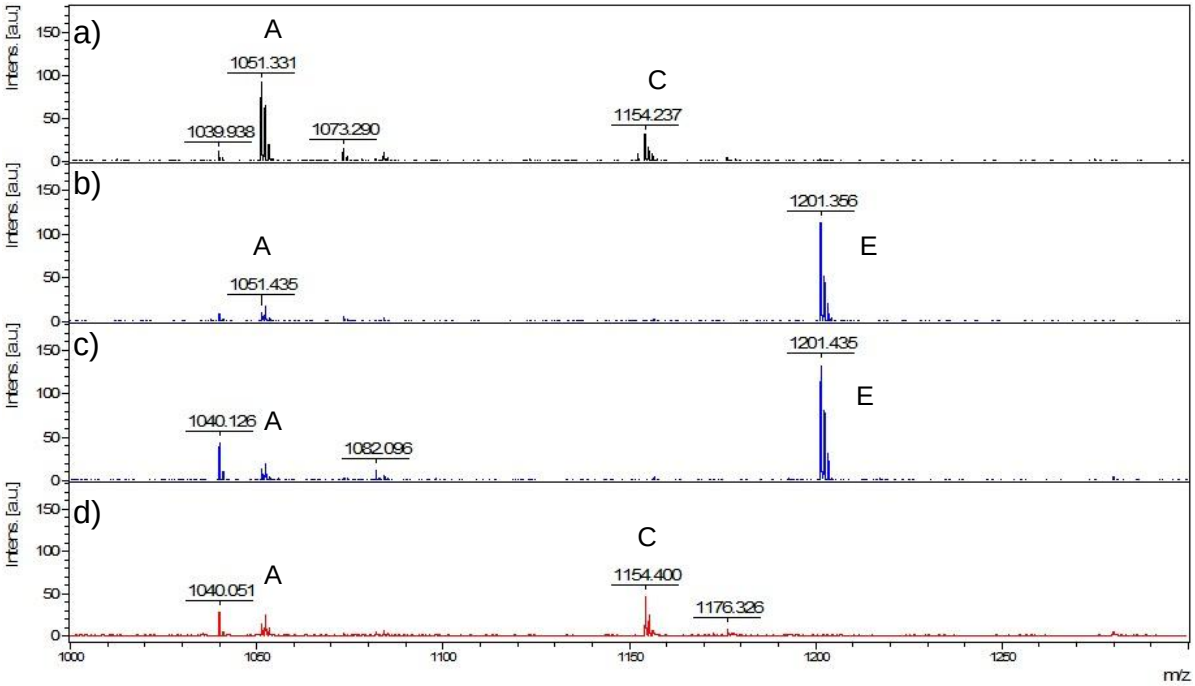


Figure S22 The reaction in presence of MPAA: a) Cys (4 mM) , b) Gly (4 mM), c) Phe (4 mM) and d) competitive with Gly, Phe and Cys (all 4 mM). **A** is the disulphide of **4**, **C** is the disulphide of **7** and **E** is the disulphide of **8**.

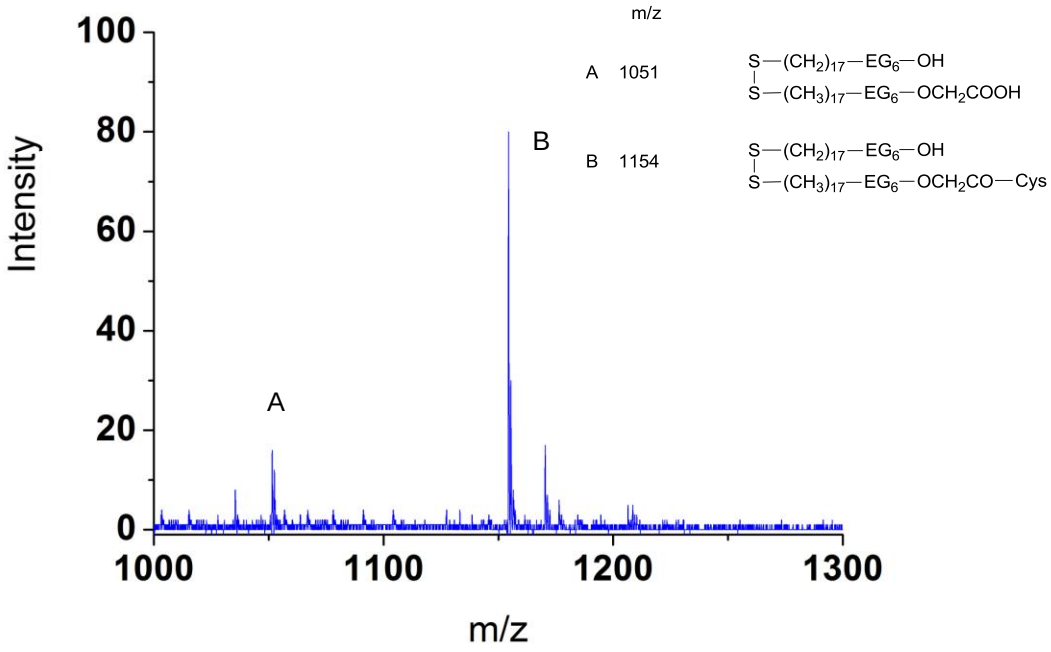


Figure S23 The chemoselective coupling on PFP activated gold plates at a 25 times excess of Gly over Cys. Only Cys product **7** was detected in MALDI-ToF MS**. A** is the disulphide of **4** and **B** is the disulphide of **7**.

5.2Peptides

	m/z	
A	1051	$S—(CH_2)_{17}—EG_6—OH$ $S—(CH_3)_{17}—EG_6—OCH_2COOH$
B	1842	$HS—(CH_3)_{17}—EG_6—OCH_2CO—$Peptide **9**

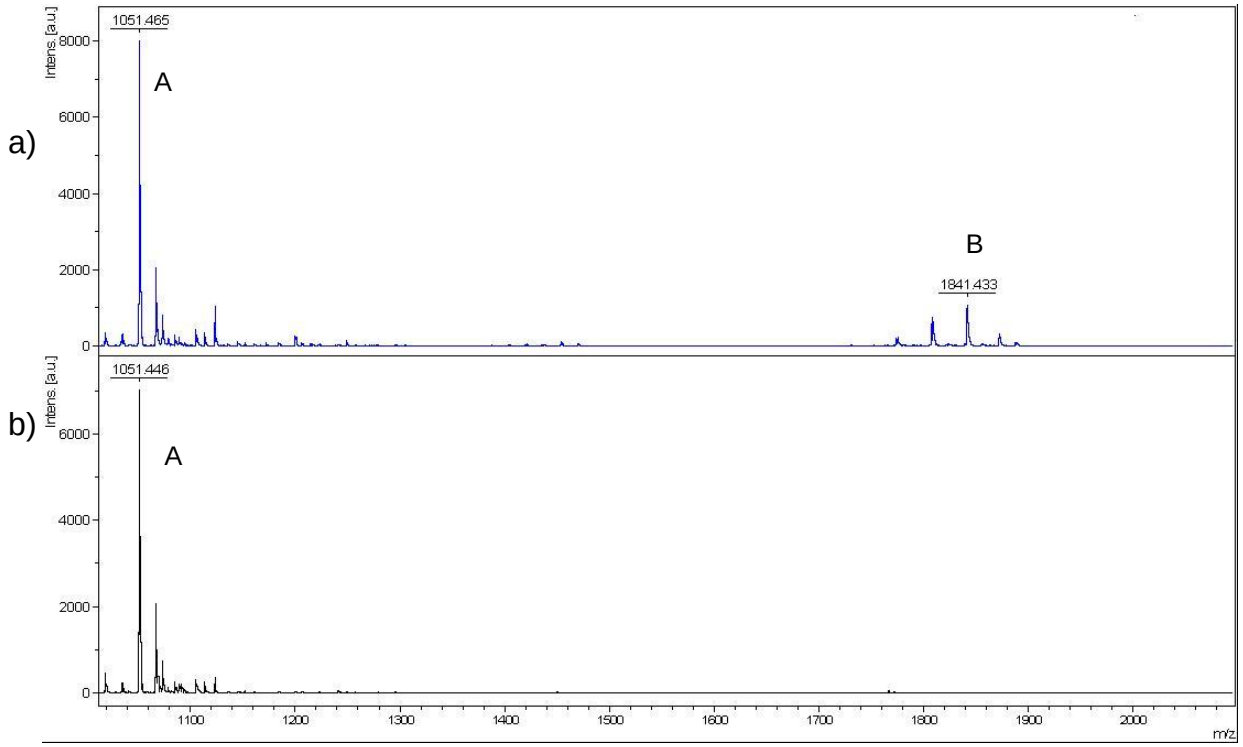


Figure S24 The coupling of the peptide **9** (0.25 mM) in presence (a) and absence (b) of imidazole. All detected ions are the $[M+H]^+$ ions. **A** is the disulphide of **4** and **B** is the disulphide of **24**.

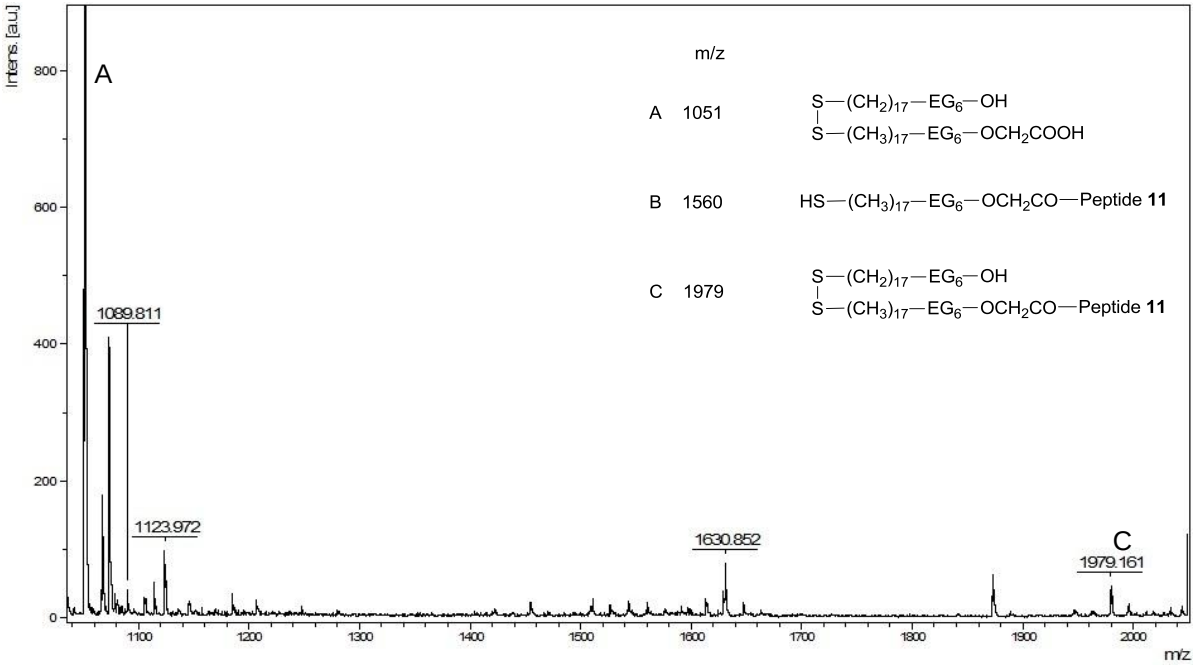


Figure S25 Coupling of the peptide **11** with imidazole at 0.25 mM. **A** is the disulphide of **4**, **B** the thiol of **25** and **C** is the disulphide of **25**.

	m/z	
A	1051	S—$(CH_2)_{17}$—EG_6—OH S—$(CH_3)_{17}$—EG_6—OCH_2COOH
B	1560	HS—$(CH_3)_{17}$—EG_6—OCH_2CO—Peptide **11**
C	1979	S—$(CH_2)_{17}$—EG_6—OH S—$(CH_3)_{17}$—EG_6—OCH_2CO—Peptide **11**

a)

b)

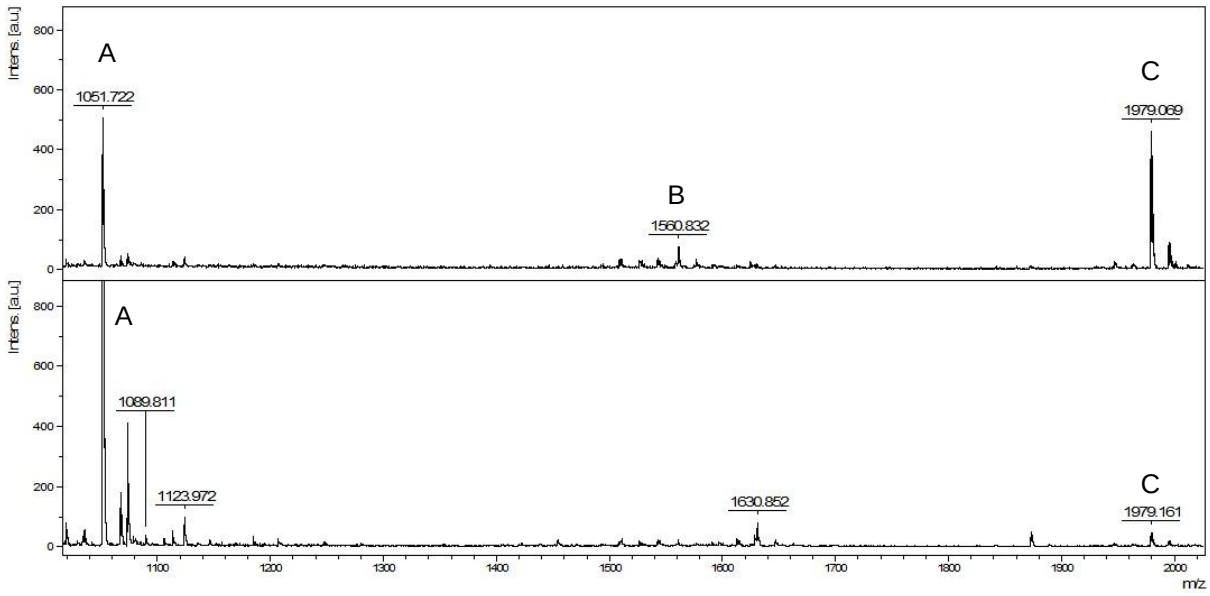


Figure S26 Coupling of the peptide **11** in presence of imidazole in a) 4 mM and b) 0.25 mM concentration. **A** is the disulphide of **4**, **B** the thiol of **25** and **C** is the disulphide of **25**.

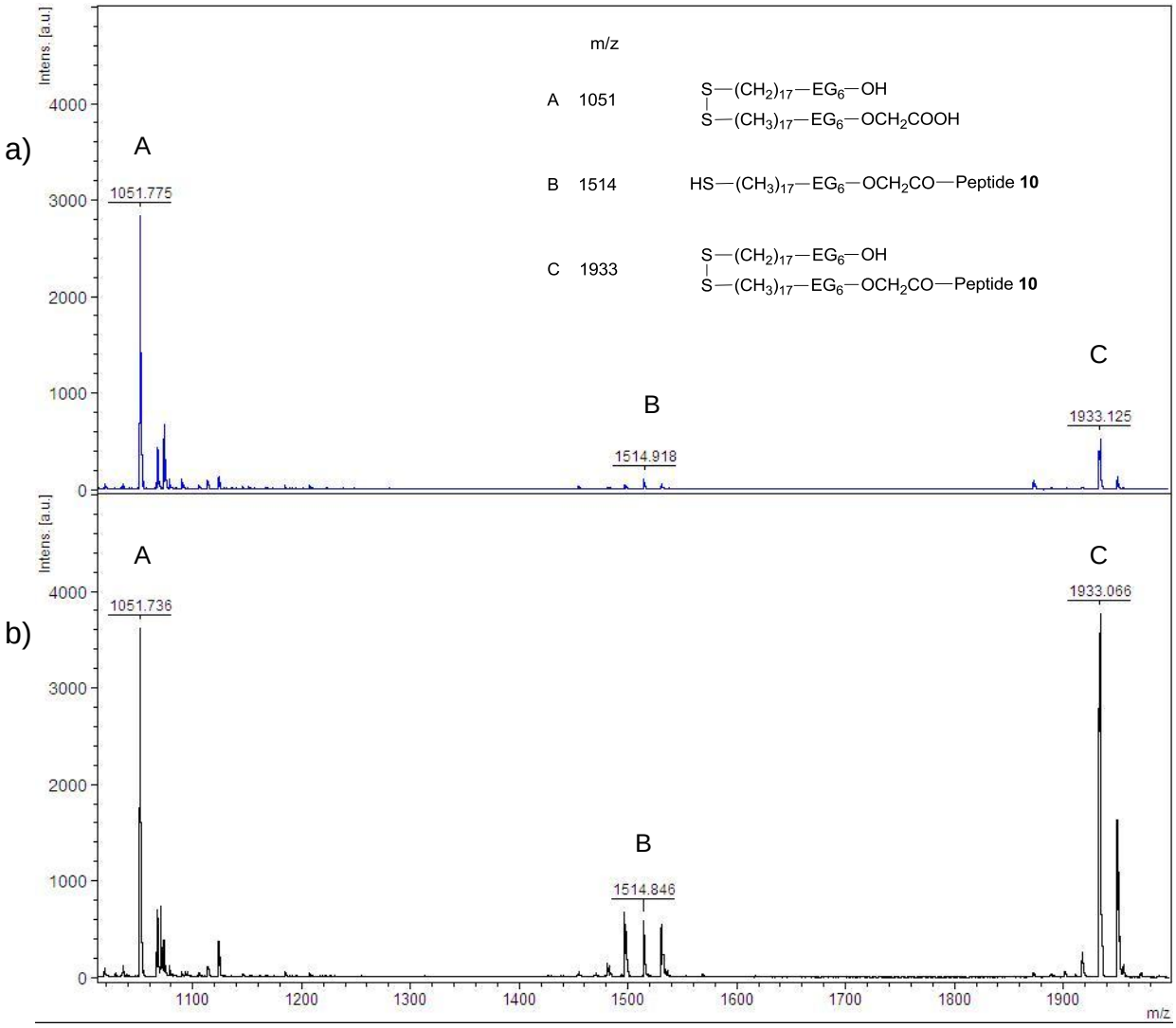


Figure S27 Coupling of the Gly-peptide **10** at 4 mM in a) presence and b) absence of imidazole. **A** is the disulphide of **4**, **B** the thiol of **10** and **C** is the disulphide of **10** on the surface.

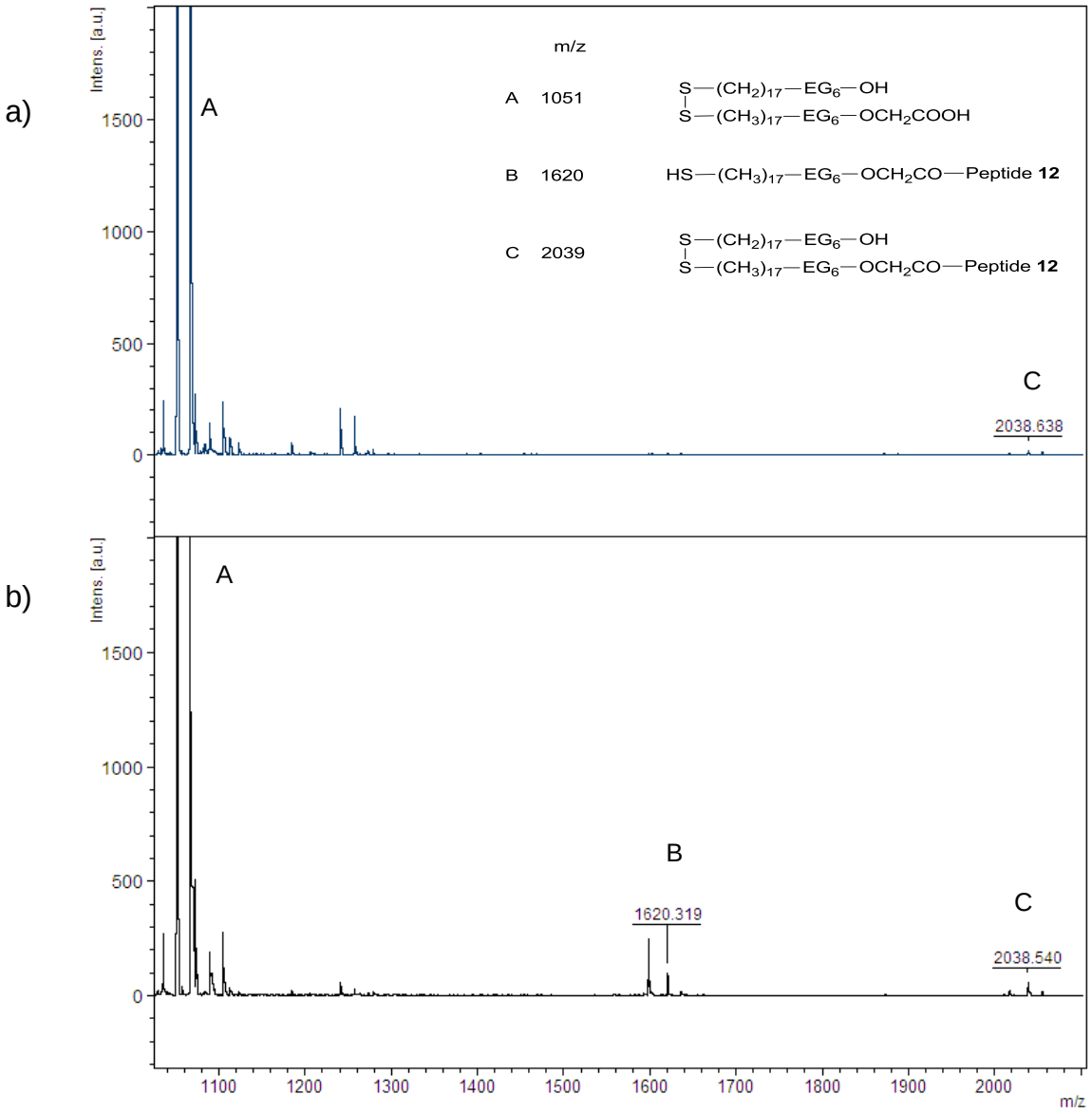


Figure S28 Coupling of the Lysine containing peptide **12** at 2 mM in a) presence and b) absence of imidazole. **A** is the disulphide of **4**, **B** the thiol of **25** and **C** is the disulphide of **25** on the surface.

	m/z	
A	1051	S—$(CH_2)_{17}$—EG_6—OH | S—$(CH_3)_{17}$—EG_6—OCH_2COOH
B	1560	HS—$(CH_3)_{17}$—EG_6—OCH_2CO—Peptide **11**
C	1842	HS—$(CH_3)_{17}$—EG_6—OCH_2CO—Peptide **9**
D	1979	S—$(CH_2)_{17}$—EG_6—OH | S—$(CH_3)_{17}$—EG_6—OCH_2CO—Peptide **11**

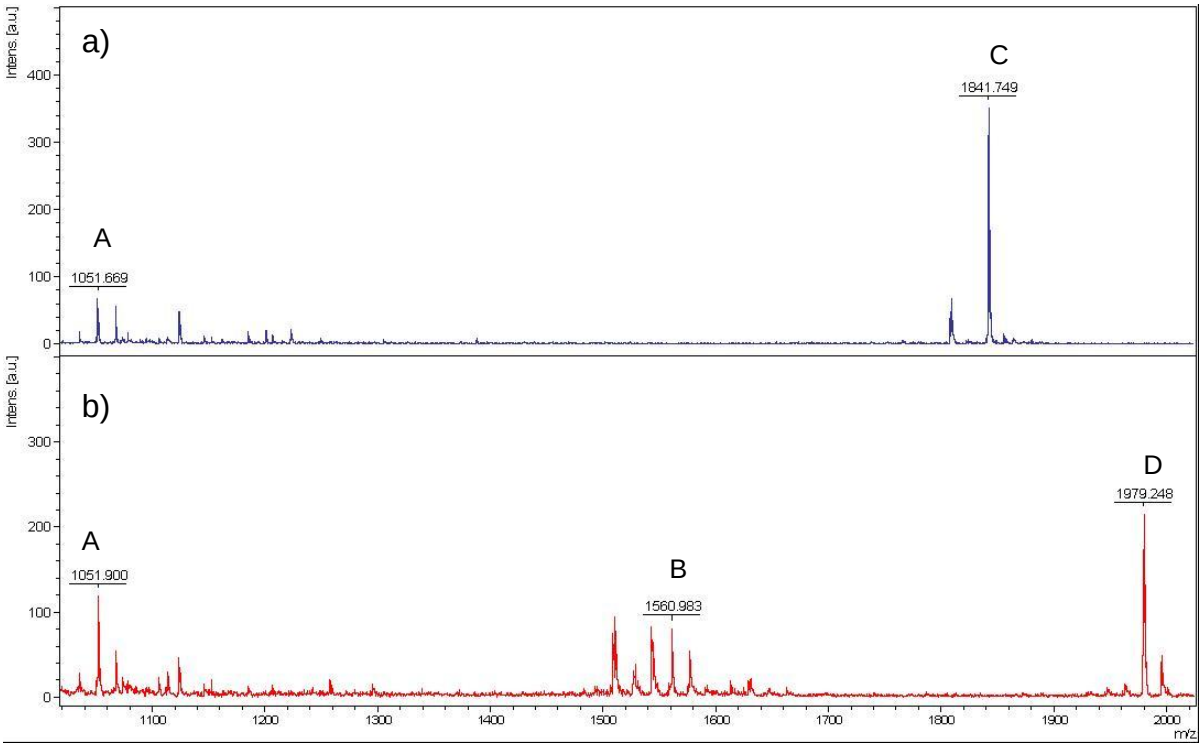


Figure S29 Competitive coupling reactions at 2 mM and in presence of imidazole: a) Cys-peptide **9** with Gly-peptide **10**. The detected ion E is corresponding to the $[M+H]^+$ ion. b) Cys-peptide **11** with Gly-peptide **10**. **A** is the disulphide of **4**, **B** the thiol of **25**, **C** is the disulphide of **24** and **D** is the disulphide of **25**.

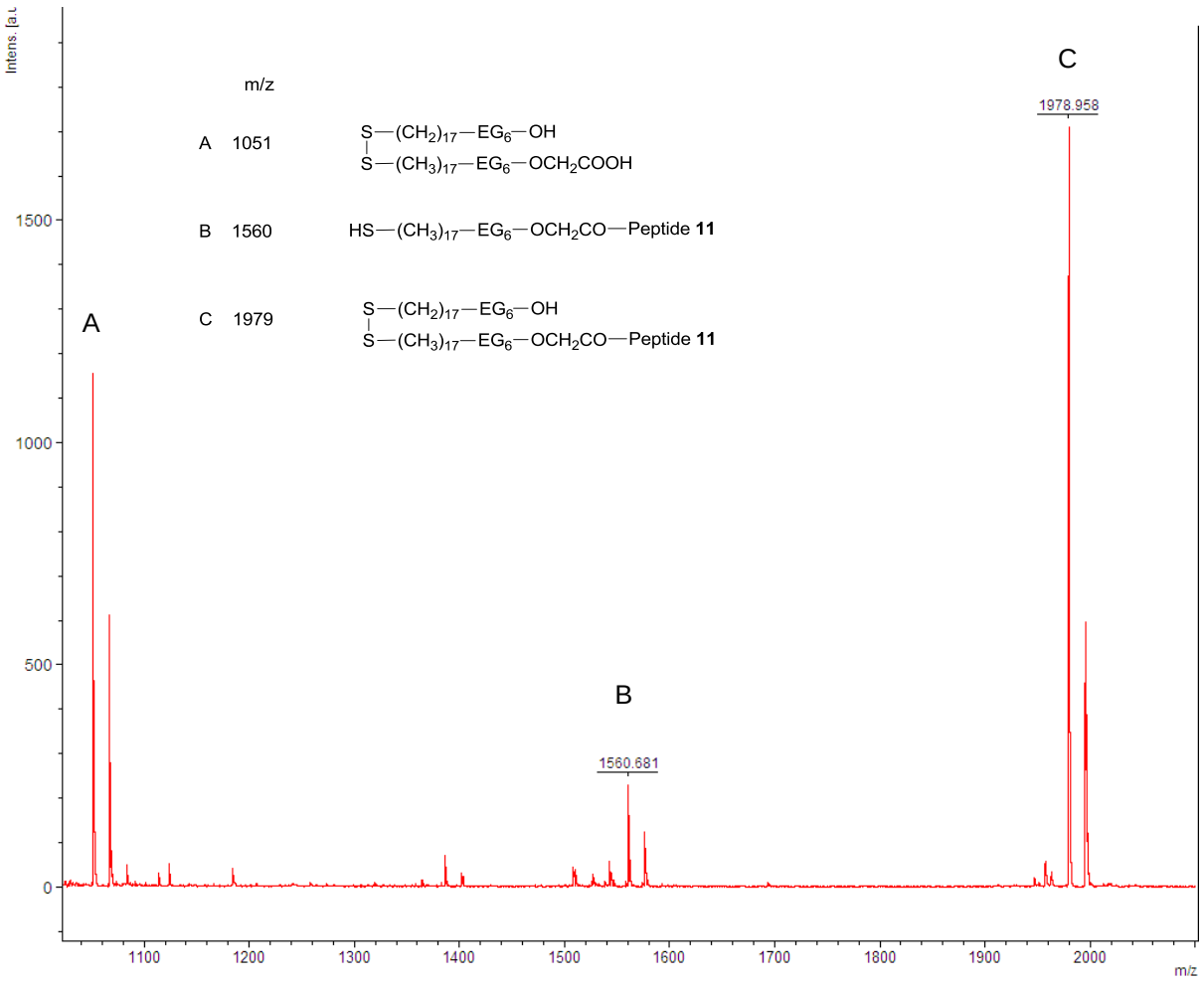


Figure S30 Competitive coupling reactions at 2 mM and in presence of imidazole: Cys-peptide **11** with Lysine containing peptide **12**. **A** is the disulphide of **4**, **B** the thiol of **25**, **C** is the disulphide of **25**.

5.3 Coupling of carbohydrates

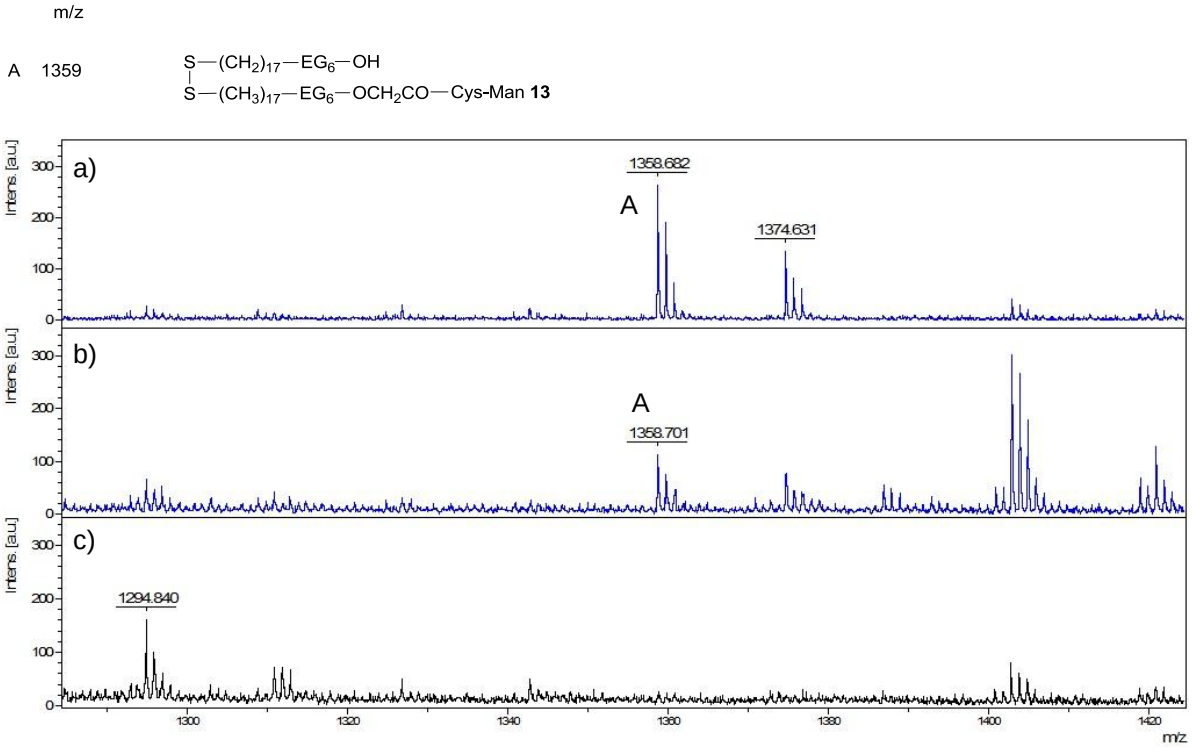


Figure S31 The coupling of the Cys-Man **13** in presence of imidazole at 0.1 (a) and 0.025 mM (b). In absence of imidazole at 0.1 mM concentration was no product detected (c). **A** is the disulphide of **26**.

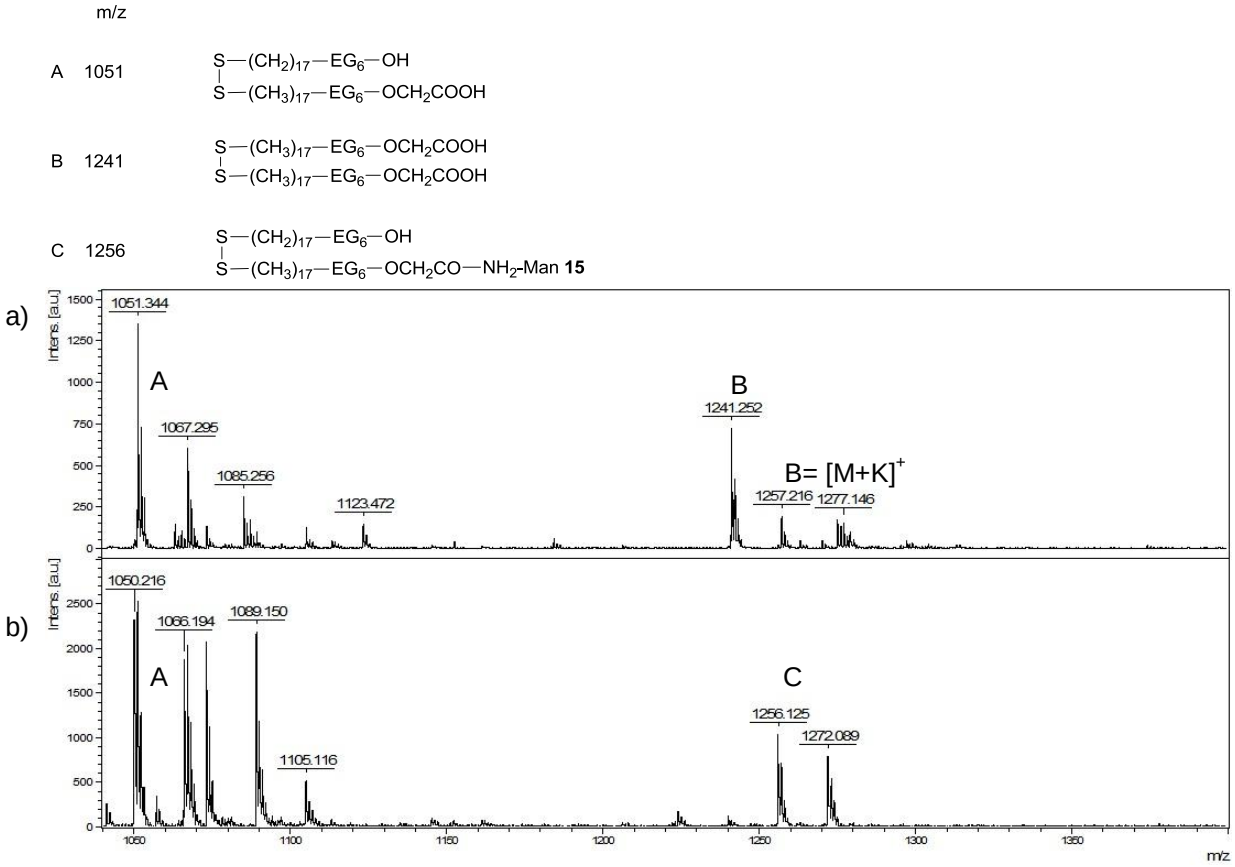


Figure S32 The coupling of aminoethyl-Man **15** showed no product formation at 4 mM in the presence of imidazole (a). In absence of imidazole at 4 mM was coupling of **15** observed. **A** is the disulphide of **4** and **C** is the disulphide of **15** on the surface.

	m/z	
A	1638	HS—$(CH_3)_{17}$—EG_6—OCH_2CO—Cys-Man_3 **23**
B	2056	S—$(CH_2)_{17}$—EG_6—OH S—$(CH_3)_{17}$—EG_6—OCH_2CO—Cys-Man_3 **23**

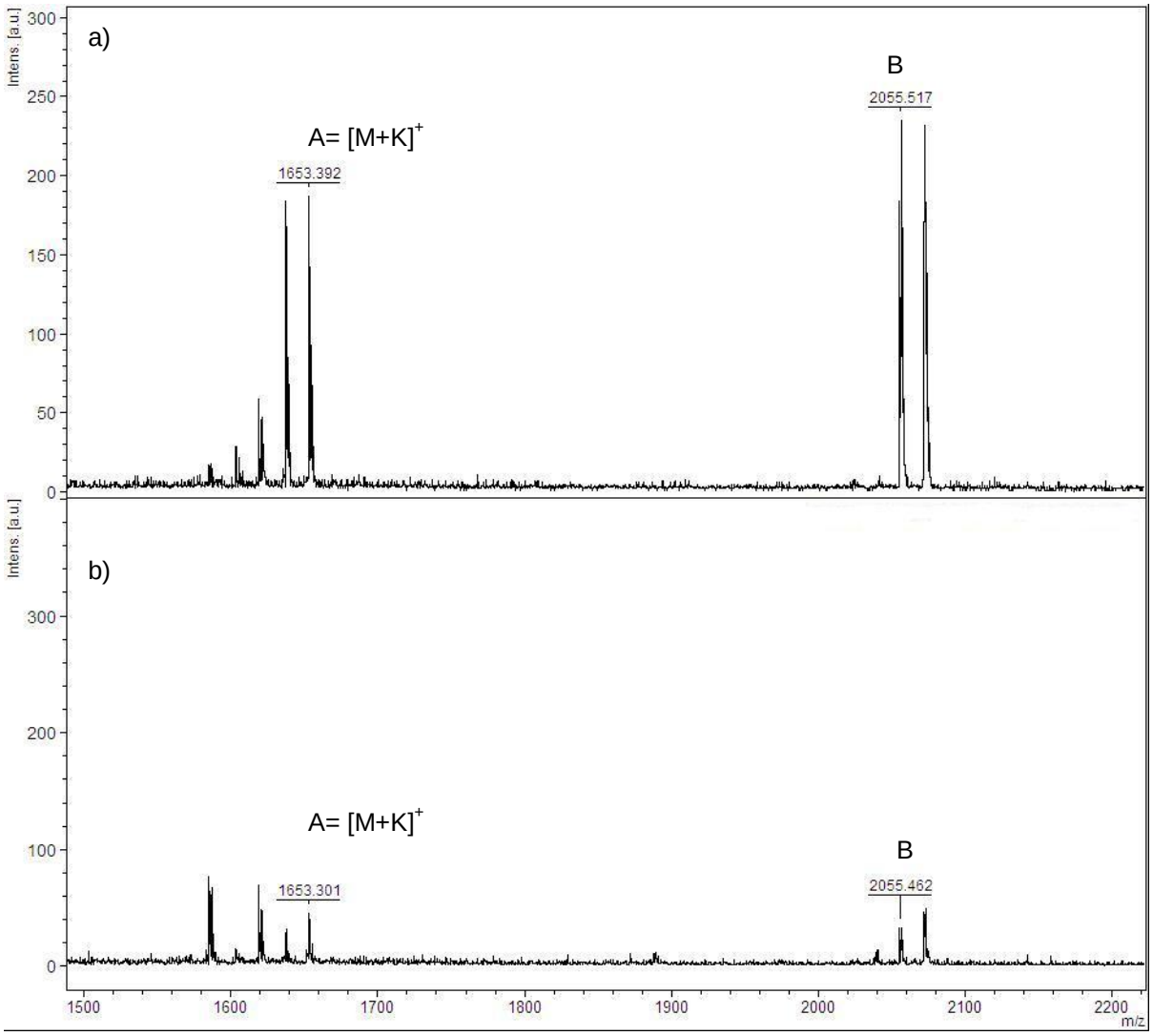


Figure S33 The coupling of trivalent Cys-Man **23** showed product formation in presence of imidazole at 1 mM (a) and 0.1 mM (b). **A** is the thiol of **27** and **B** the disulphide of **27**.

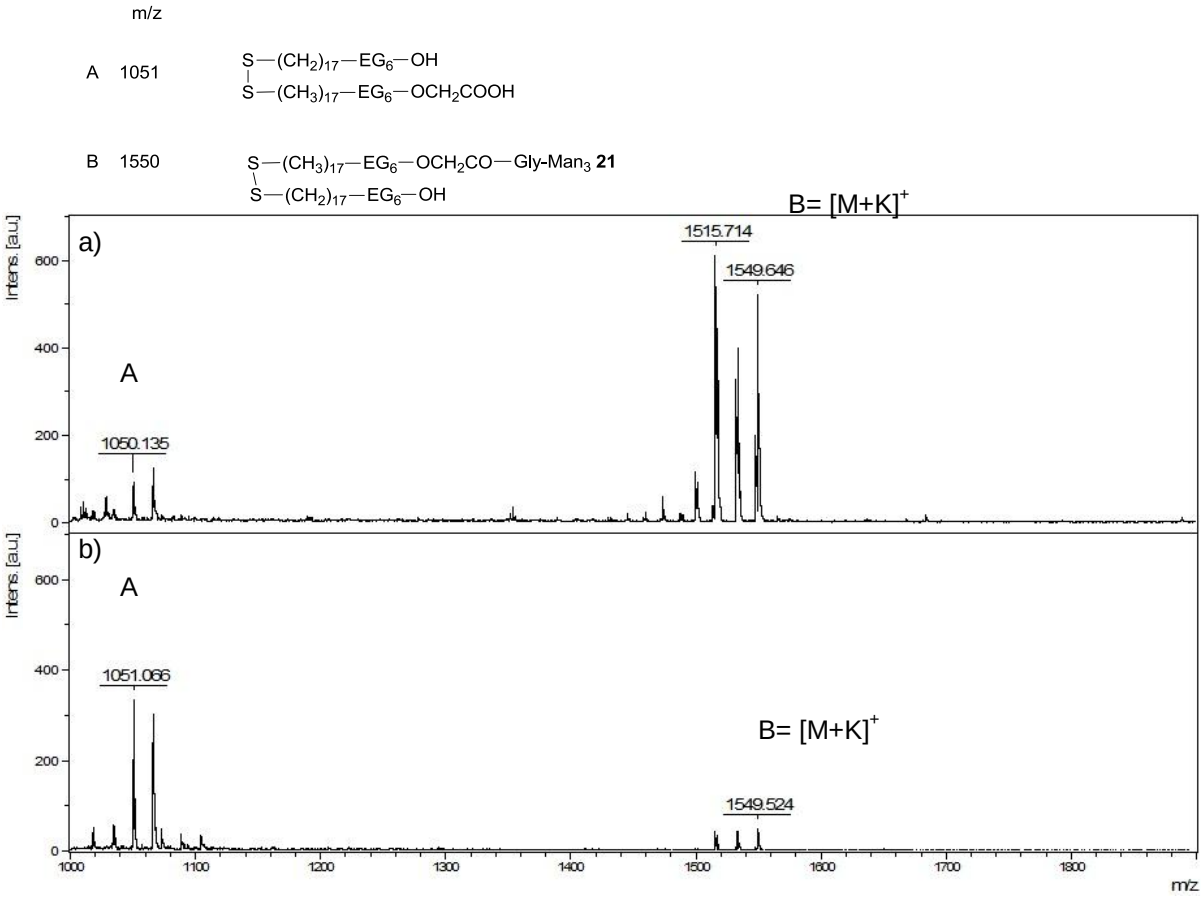


Figure S34 The coupling of trivalent Gly-Man **21** showed coupling in absence of imidazole at 4 mM (a) and 1 mM (b). **A** is the disulphide of **4**, **B** the disulphide of **21** on the surface

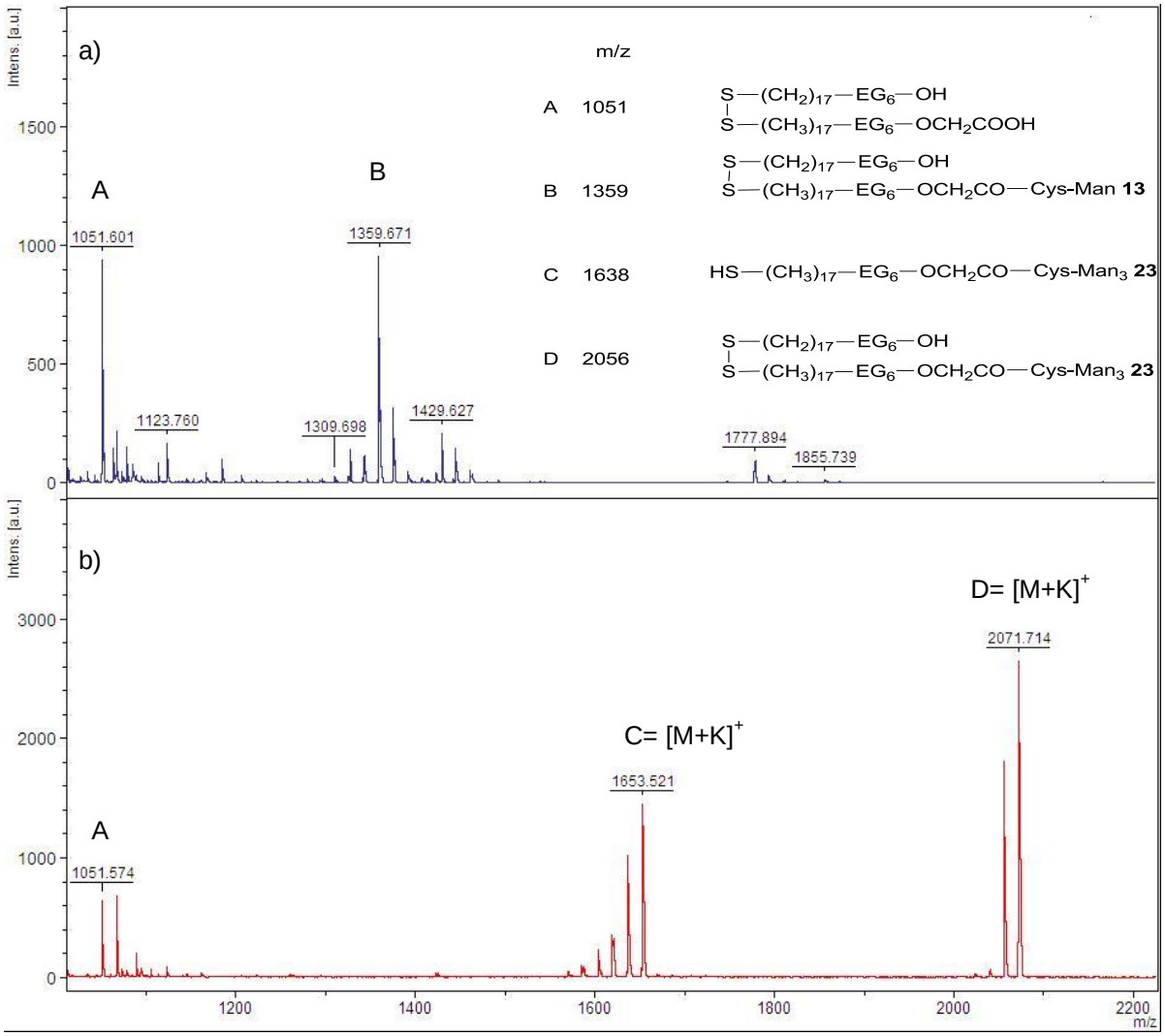


Figure S35 The competitive reactions in presence of imidazole: a) aminoethyl-Man **15** and Cys-Man **13** at 2 mM concentration. b) Trivalent Gly-Man **21** and trivalent Cys-Man **23** at a concentration of 4 mM. **A** is the disulphide of **4**, **B** the disulphide of **26**, **C** the thiol of **27** and **D** is the disulphide of **27**.

6. Surface Plasmon Resonance (SPR) Analysis

The experiments were performed on a Biacore 3000 system (GE Healthcare, Sweden) using Sensor Chip Au (GE Healthcare). The gold coated sensor was cleaned with piranha solution, rinsed with water, ethanol and dried in a stream of nitrogen. The formation of self-assembling monolayers was performed in the same way as described (see above). The chip was washed with ethanol, dried under nitrogen and mounted into a chip holder following instructions in the supplier's manual. After docking in the instrument the sensor was equilibrated with phosphate buffer saline (PBS, degassed and filtered) at a flow rate of 10 µL/min. For surface activation, 70 µl of a 1:1 mixture of freshly prepared NHS (0.4 M) and EDC (0.1 M) in water were injected (Channels 1-3). Reference spot (channel 1) was blocked by injecting 70 µl of amino ethanol HCl (1 M). Aminoethyl GlcNAc **14** (10 mM, Channel 2) and mannosyl-cysteine **13** (2 mM and 15 mM TCEP, Channel 3) were immobilised by injecting 100 µl of the substrate solution in PBS followed by blocking with amino ethanol (70 µl, 1 M). The SAMs in channel 4 were PFP-activation. In order to perform this activation the chip was undocked, and the gold surface coated with 2 µL of EDC and PFP (final concentrations 0.180 M and 0.174 M, respectively) in dry DMF at r.t. After 30 min the chip was washed with ethanol, and dried under nitrogen. After docking of the chip, the SPR chip was equilibrated with PBS buffer at a flow rate of 5 µL/min and 250 µL of the reaction mixture of 2 mM GlcNAc-NH_2 **14** and 2 mM Cys-Man **13** under NCL conditions (see section 4) were injected within an hour (channel 4).

Binding studies were carried out using HEPES buffer (10 mM HEPES, 0.15 M NaCl, 1 mM $CaCl_2$, 1 mM $MnCl_2$, pH 7.4) at a flow rate of 25 µl/min. 250 µl of lectin solution (50 µg/mL) were injected followed by 600 s dissociation sequence. The chip surface was regenerated by injecting 200 µl of 100 mM glycine-HCl solution (pH 2).

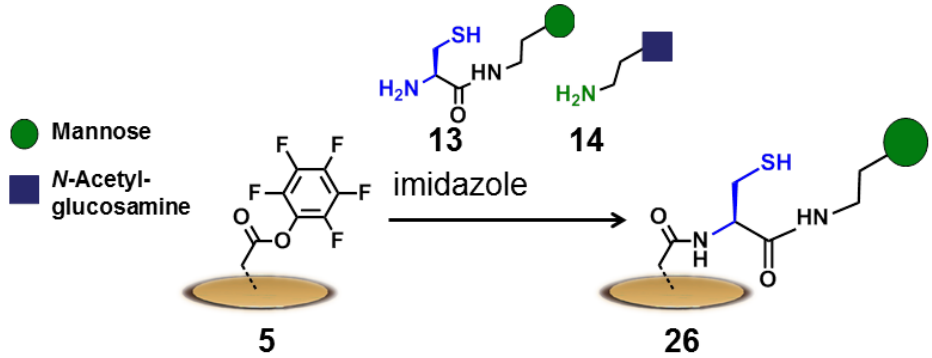
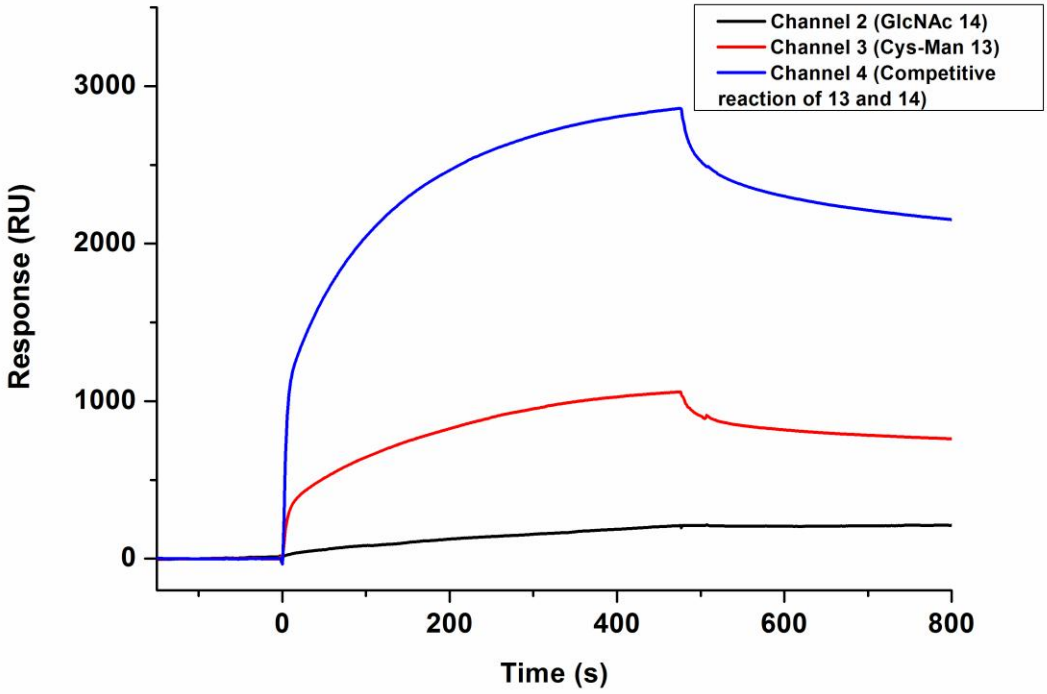


Figure S36 The cysteine compound Cys-Man **13** and the amino-compound GlcNAc **14** were applied in a competitive assay at 2 mM on PFP activated SAMs.

a)



b)

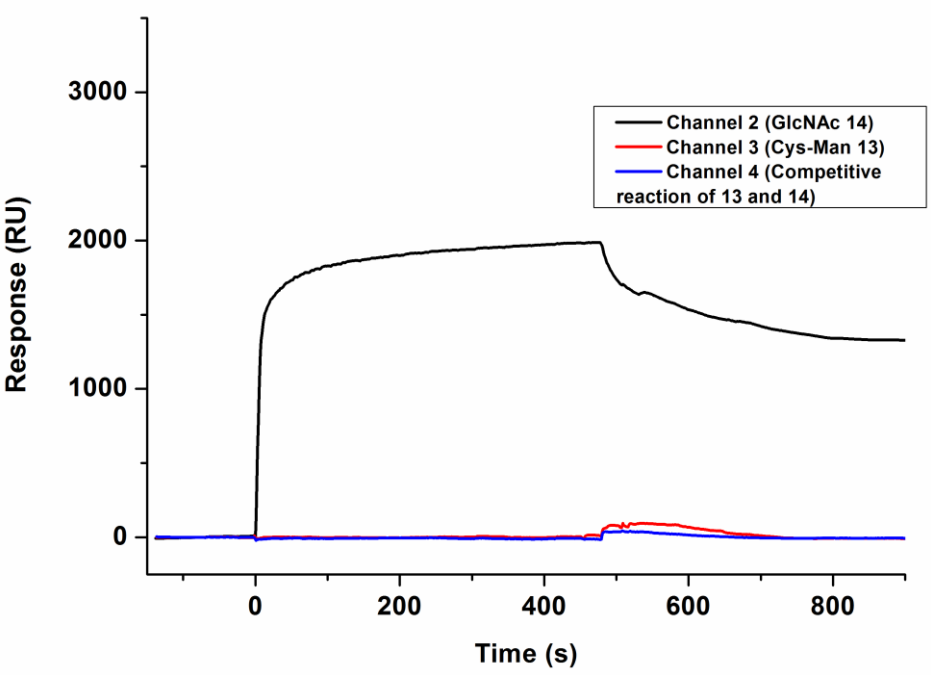


Figure S37 SPR analysis of the reaction in Figure S36 shown in Channel 4. Black: immobilised GlcNAc **14**, Red: immobilised Cys-Man **13**. a) Readout with the mannose binding lectin concanavalin A (ConA) (50 μg/mL). ConA showed strong binding to the spot (blue) were the competitive reaction was carried out. This shows the presence of the Cys-Man **13**. b) Readout with the GlcNAc binding lectin wheat germ agglutinin (WGA) (50 μg/mL) showed no binding to the spot (blue) were the competitive reaction was carried out. This shows that no GlcNAc **14** has bound to the surface.

References for supporting information

1. J. W. Wehner and T. K. Lindhorst, *Synthesis*, 2010, 3070-3082.
2. C. Akpo, E. Weber and U. Reiche, *New J. Chem.*, 2006, **30**, 1820-1833; G. R. Newkome, R. K. Behera, C. N. Moorefield and G. R. Baker, *J. Org. Chem.*, 1991, **56**, 7162-7167.
3. M. Kleinert, N. Röckendorf and T. K. Lindhorst, *Eur. J. Org. Chem.*, 2004, 3931-3940; T. K. Lindhorst, S. Kötter, U. Krallmann-Wenzel and S. Ehlers, *J. Chem. Soc. Perkin Trans. 1*, 2001, 823-831; J. Dahmen, T. Frejd, G. Gronberg, T. Lave, G. Magnusson and G. Noori, *Carbohydr. Res.*, 1983, **116**, 303-307.
4. R. Sardzik, G. T. Noble, M. J. Weissenborn, A. Martin, S. J. Webb and S. L. Flitsch, *Beilstein J. Org. Chem.*, 2010, **6**, 699-703.
5. N. Srinivasan, A. Yurek-George and A. Ganesan, *Mol. Divers.*, 2005, **9**, 291-293.