SUPPLEMENTARY INFORMATION

Unexpected reactivity and mechanism of carboxamide activation in bacterial N-linked protein glycosylation

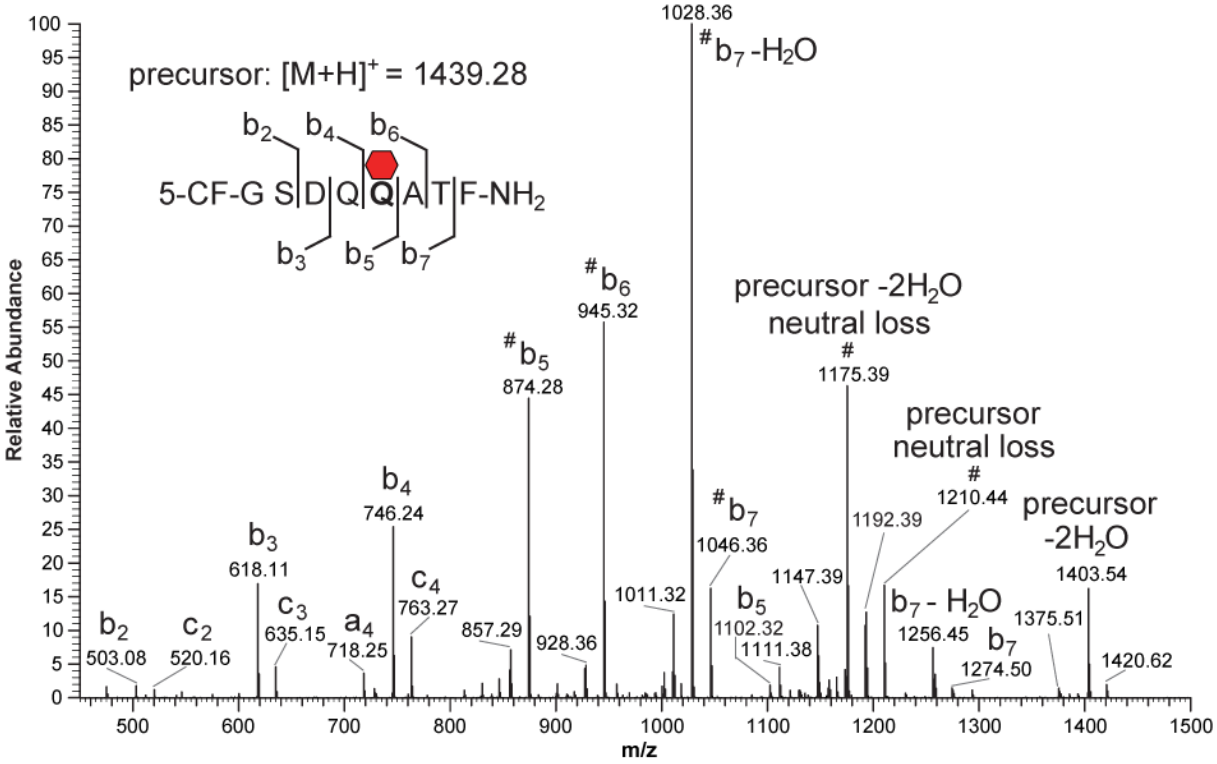
Christian Lizak[#], Sabina Gerber[#], Gaëlle Michaud, Mario Schubert, Yao-Yun Fan, Monika Bucher, Tamis Darbre, Markus Aebi, Jean-Louis Reymond, Kaspar P. Locher[*]

[#]These authors contributed equally

*Correspondence:
Phone: +41-44-633-3991
Fax: +41-44-633-1182
Email: locher@mol.biol.ethz.ch

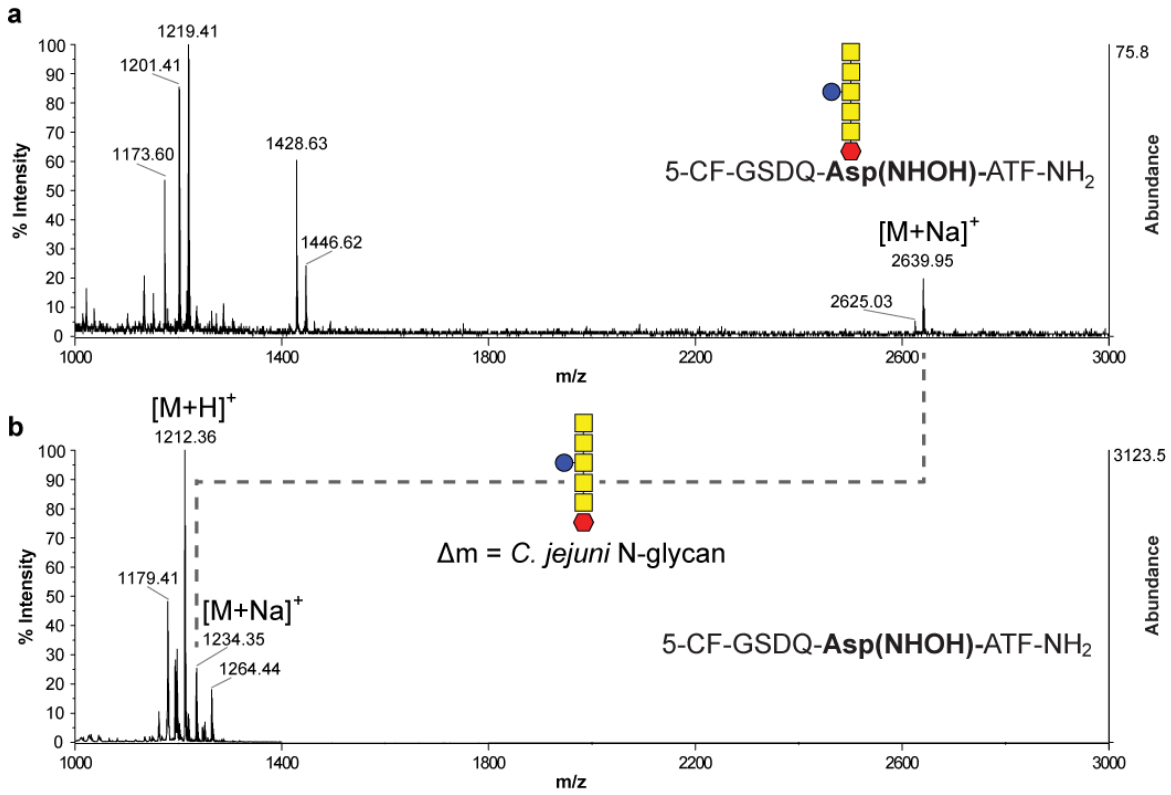
Supplementary Figures

Supplementary Figure S1



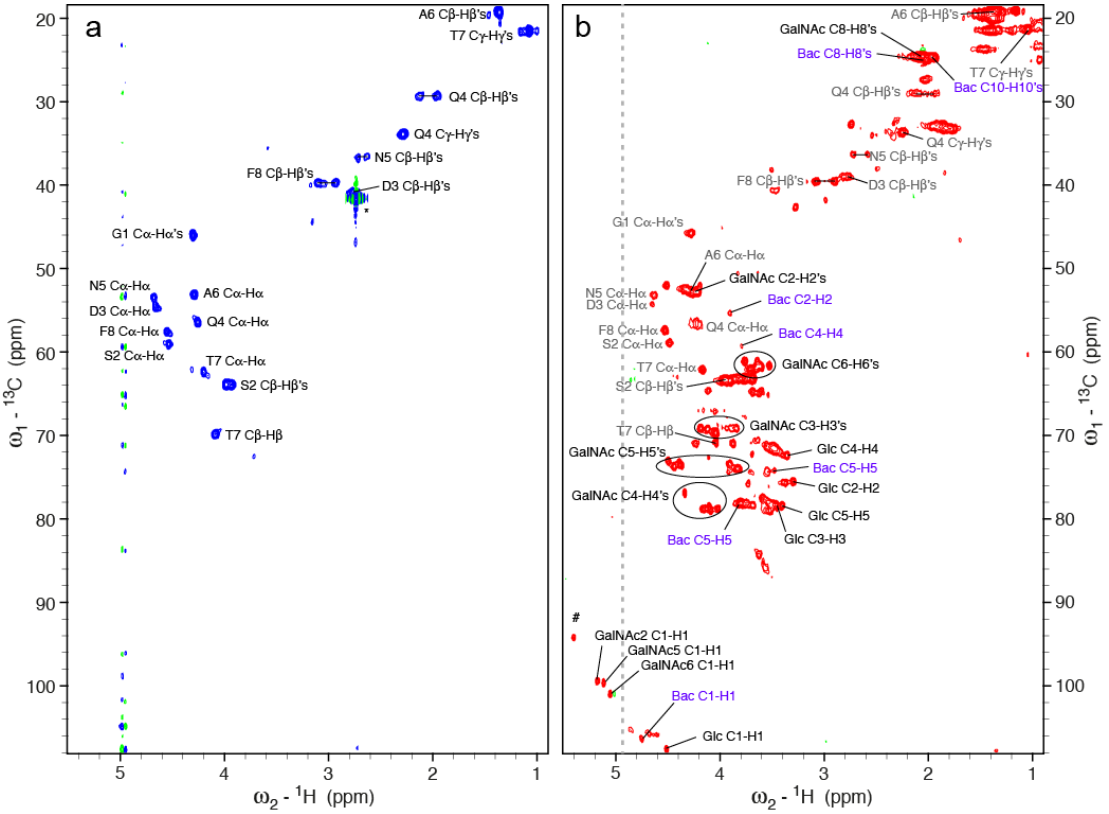
Supplementary Figure S1: LC-ESI-MS3 sequencing of glutamine-linked glycopeptide. *In vitro* glycosylation product of m/z = 873.01 (see Fig. 3d) from substrate peptide containing the DQQAT acceptor site was selected for MS2, and the product ion of 5-CF-GSDQQ(diNAcBac)ATF-NH_2 (m/z = 1439.28) was selected for MS3. A sequence assignment is schematically illustrated in the figure. Red hexagon: di-*N*-Acetyl-bacillosamine (diNAcBac); 5-CF: 5-carboxyfluorescein; # indicates the fragment ions arising from precursor glycopeptide neutral loss.

Supplementary Figure S2



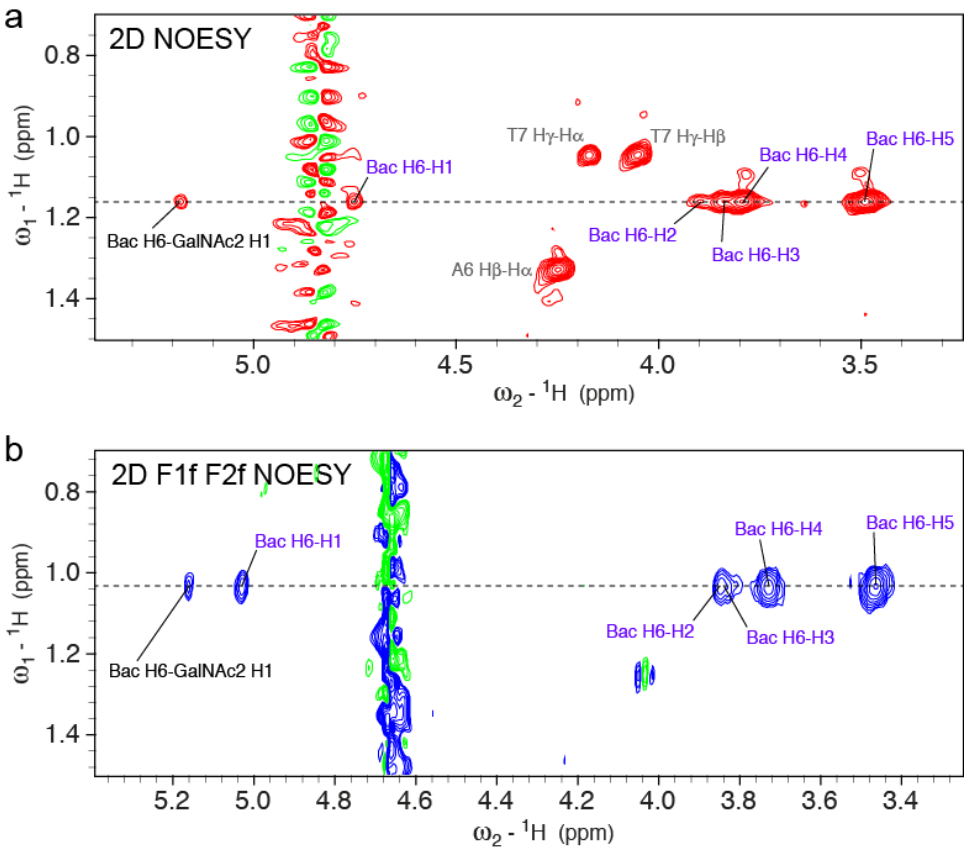
Supplementary Figure S2: MALDI-MS profile of Asp(NHOH) *in vitro* glycosylation. After the *in vitro* glycosylation reaction, fluorescently labelled peptide and glycopeptide were extracted and subjected to MALDI-MS analysis. (a) Glycosylation product of substrate 5-CF-GSDQ-Asp(NHOH)-ATF-NH_2 corresponds to m/z = 2639.95. (b) The product of reaction without PglB corresponds to m/z = 1234.35. The resulting mass difference between the peptide and glycopeptide is indicated and corresponds to the mass of the *C. jejuni* N-glycan. Red hexagon: diNAcBac; yellow squares: *N*-Acetyl-galactosamine; blue circles: glucose. 5-CF: 5-carboxyfluorescein.

Supplementary Figure S3



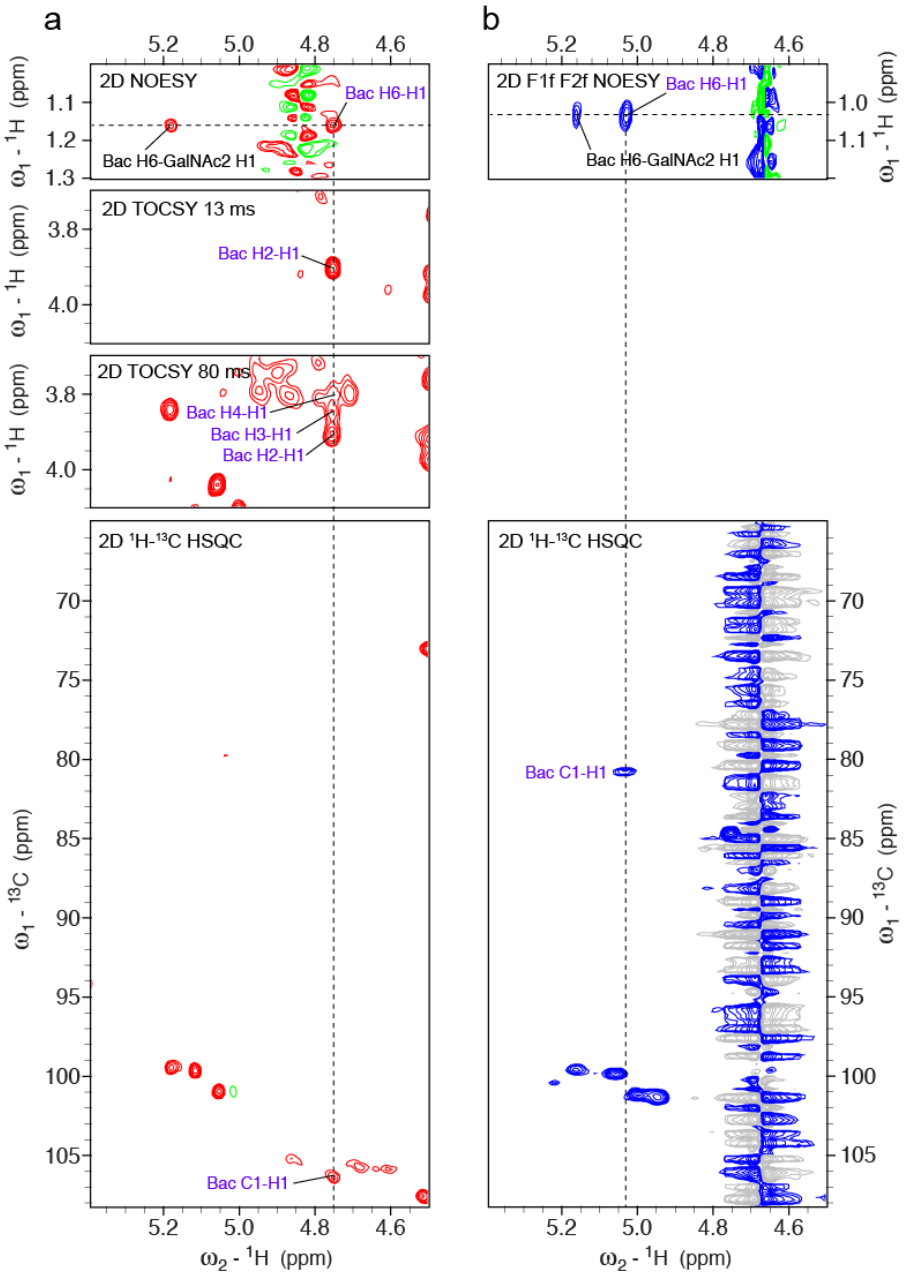
Supplementary Figure S3: Comparison of the natural abundance 2D ^{1}H-^{13}C HSQC spectra of non-glycosylated and *in vitro* glycosylated Asp(NHOH) peptide. (a) One-bond ^{1}H-^{13}C correlations of the non-glycosylated peptide measured at 283 K and 500 MHz with 240 transients. The signal marked with * is coming from a DMSO impurity that could not be removed by lyophilisation. (b) Comparable spectrum of the *in vitro* glycosylated peptide recorded at 283 K and 900 MHz with 260 transients. The signals of the bacterial glycan are clearly visible and correspond to the signals observed of the *in vitro* glycosylated protein AcrA61-210ΔΔ, except the C1-H1 correlation of diNAcBac[18]. Axes are calibrated to DSS according to Markley et al[56].

Supplementary Figure S4



Supplementary Figure S4: Comparison of NOESY cross-peaks of di-*N*-Acetyl-bacillosamine of *in vitro* glycosylated Asp(NHOH) peptide and protein AcrA61-210ΔΔ. (a) Region of 2D NOESY spectrum of the *in vitro* glycosylated Asp(NHOH) peptide showing NOE signals of the diNAcBac methyl group H6. The spectrum was measured at 293 K and 900 MHz with 96 transients and a mixing time of 150 ms. Peptide signals are labelled in grey. Only one set of signals is found at the typical chemical shift the resonance of diNAcBac H6 (between 1.0 and 1.2 ppm). (b) 2D ^{13}C-filtered-filtered NOESY spectrum of the *in vitro* glycosylated protein AcrA61-210ΔΔ showing NOE signals of the diNAcBac methyl group H6 at 303 K [18]. The resonance assignment of AcrA61-210ΔΔ is found at the BioMagRes databank (accession number 15737) and was reported previously[18].

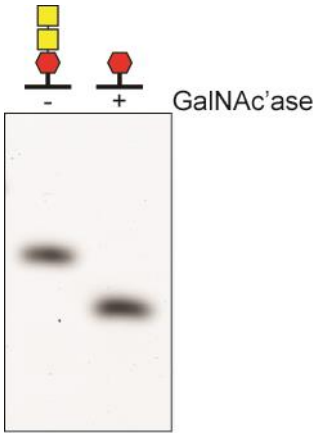
Supplementary Figure S5



Supplementary Figure S5: Assignment of the ^{13}C chemical shift of the anomeric C1 of di-*N*-Acetyl-bacillosamine. (a) Spectra of *in vitro* glycosylated Asp(NHOH) peptide. The H1 chemical shift of diNAcBac observed first in the 2D NOESY spectrum (top) was confirmed by 2D TOCSY spectra with different mixing times. A natural abundance 2D ^{1}H-^{13}C HSQC (bottom) shows only one correlation at the diNAcBac H1 resonance (dashed line) with a ^{13}C chemical shift of 106.4 ppm that is typical for an O-glycosidic linkage to hydroxamates[21,22]. (b) Spectra of the *in vitro* glycosylated protein AcrA61-210ΔΔ [18]. The 2D ^{13}C-filtered-filtered NOESY spectrum (top) displays the H1 chemical shift of diNAcBac more downfield than in the peptide. The natural abundance 2D ^{1}H-^{13}C HSQC (bottom) shows a correlation at the

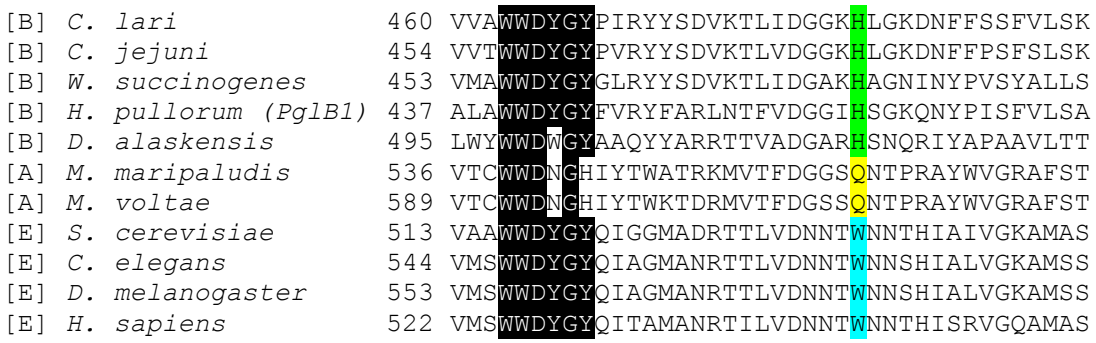
diNAcBac H1 resonance (dashed line) with a ^{13}C chemical shift of 80.8 ppm that is typical for an N-glycosidic linkage[18]. Axes are calibrated to DSS according to Markley et al[56].

Supplementary Figure S6



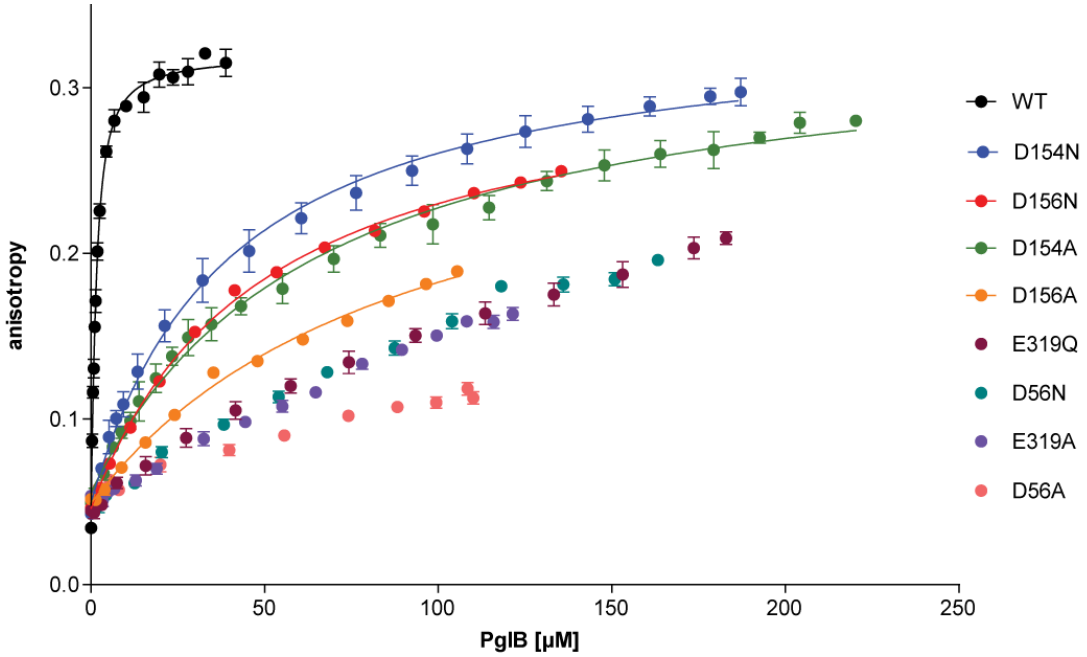
Supplementary Figure S6: Glycan trimming on the glycopeptide. Fluorescently labelled glycopeptide (DQNAT sequon) carrying the diNAcBac-$GalNAc_2$ N-glycan was incubated with or without exo-α-*N*-Acetyl-galactosaminidase to remove terminal GalNAc residues of the glycan. The resulting glycopeptides were analysed by Tricine SDS-PAGE and bands were visualized by a fluorescence gel scan (488 nm excitation and 526 nm emission). Removal of terminal GalNAc residues causes a mobility shift to reduced molecular weight. Red hexagon: diNAcBac; yellow squares: *N*-Acteyl-galactosamine.

Supplementary Figure S7



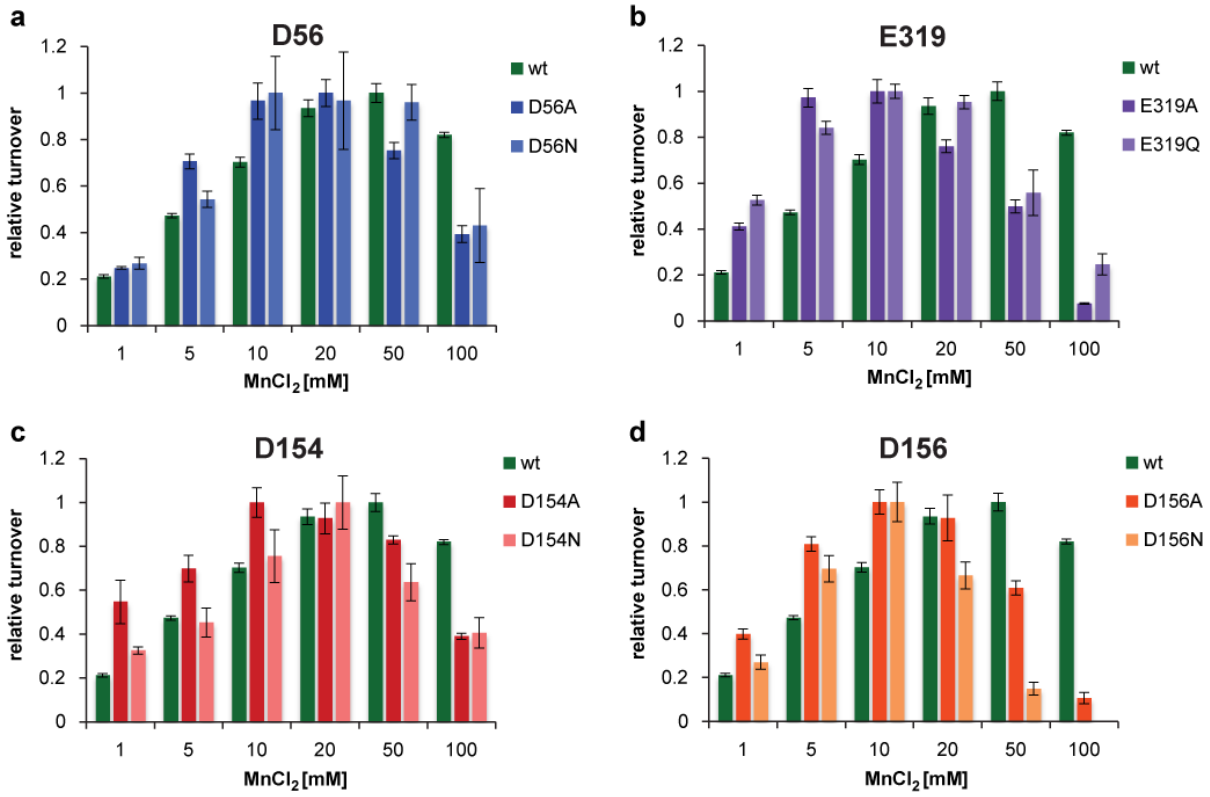
Supplementary Figure S7: Sequence alignment showing the conservation of H485. Alignment of different bacterial [B], archaeal [A] and eukaryotic [E] STT3 homologues. The conserved WWD-motif is highlighted in black.

Supplementary Figure S8



Supplementary Figure S8: Peptide binding of different enzyme variants affecting the catalytic site. To quantify peptide binding using fluorescence anisotropy, purified PglB variants (wt or mutants as indicated in the figure) were titrated into a solution containing 1 µM fluorescently labelled peptide (containing the DQNAT sequon) and 10 mM $MnCl_2$. Data points reflect the mean of 20 measurements of the same sample, error bars indicate standard deviations. Curve fitting was performed assuming a single binding site. Data for wt PglB (black curve) are as presented in[15].

Supplementary Figure S9



Supplementary Figure S9: Influence of $MnCl_2$ concentration on *in vitro* activity of different PglB active site mutants. The turnover rate of PglB variants (wt or mutants as indicated in the figure) on fluorescently labelled substrate peptide (DQNAT sequon) was investigated at various concentrations of $MnCl_2$. Each data point results from a single turnover rate determination and the error bars account for the standard error (SEM) of each fit. Note that the relative activities were calculated for different PglB mutants individually. The turnover rates of investigated mutants are different, but in order to compare the $MnCl_2$ activity profiles, individual relative activities of several mutants are presented in one diagram. Data for wt PglB (green bars) are as presented in[15]. (a) Comparison of wt PglB to mutant D56A and D56N, respectively. (b) Comparison of wt PglB to mutant E319A and E319Q, respectively. (c) Comparison of wt PglB to mutant D154A and D154N, respectively. (d) Comparison of wt PglB to mutant D156A and D156N, respectively.

Supplementary Table

Supplementary Table S1: *In vitro* glycosylation turnover rates of PglB active site mutants

mutant	number of purifications	variations in relative activity (reduction compared to wt)
D56A	5	11,000 – 150,000-fold [a]
D56N	5	8,000 – 25,000-fold
D154A	10	3,500 – 30,000-fold
D154N	5	17,000 – 60,000-fold
D156A	3	450 – 500-fold
D156N	3	1,500 – 2,400-fold
E319A	3	900 – 1,000-fold [a]
E319Q	4	10,000 – 34,000-fold

[a] Data as presented in[15].

Supplementary Methods

Synthesis of acceptor peptides labelled with 5-carboxyfluorescein

Amino acids were used as the following derivatives: Fmoc-Phe-OH, Fmoc-Thr(tBu)-OH, Fmoc-Ala-OH, Fmoc-Asn(Trt)-OH, Fmoc-Asp(OAll)-OH, Fmoc-L-β-Homoserine($O^{t}Bu$), Fmoc-Ala-Thr($\Psi^{Me,Me}$pro)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Asp(OtBu)-OH, Fmoc-Ser(tBu)-OH and Fmoc-Gly-OH.

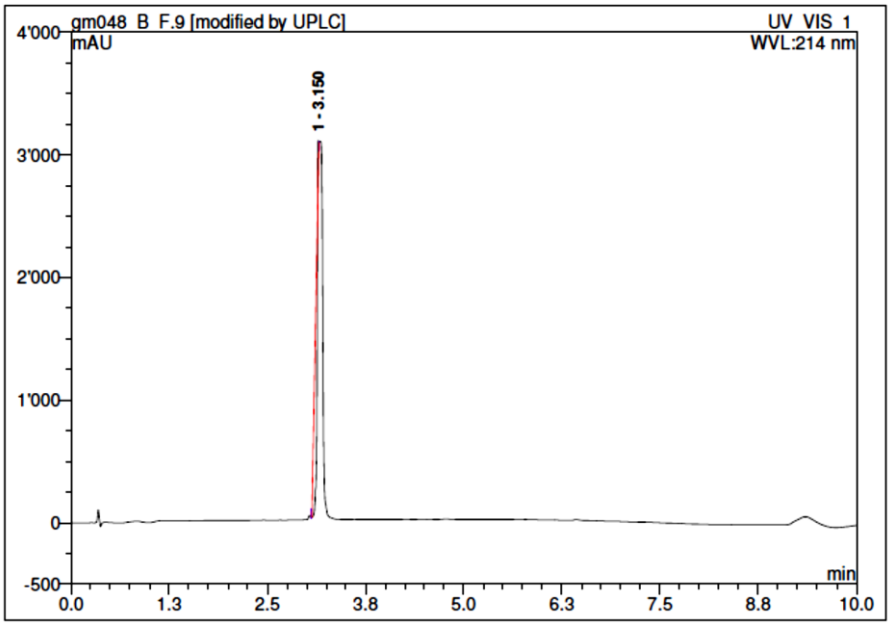
Analytical RP-UPLC was performed with an ULTIMATE Rapid Separation LC System (Dionex, ULTIMATE-3000RS diode array detector) using an Acclaim® RSLC 120 C18 column (Dionex, 2.2 µm, 120 Å, 3.0 × 50 mm, flow 1.2 $mL \cdot min^{-1}$). Compounds were detected by UV absorption at 214 nm. Data recording and processing was performed with Chromeleon Management System Version 6.80 (Dionex).

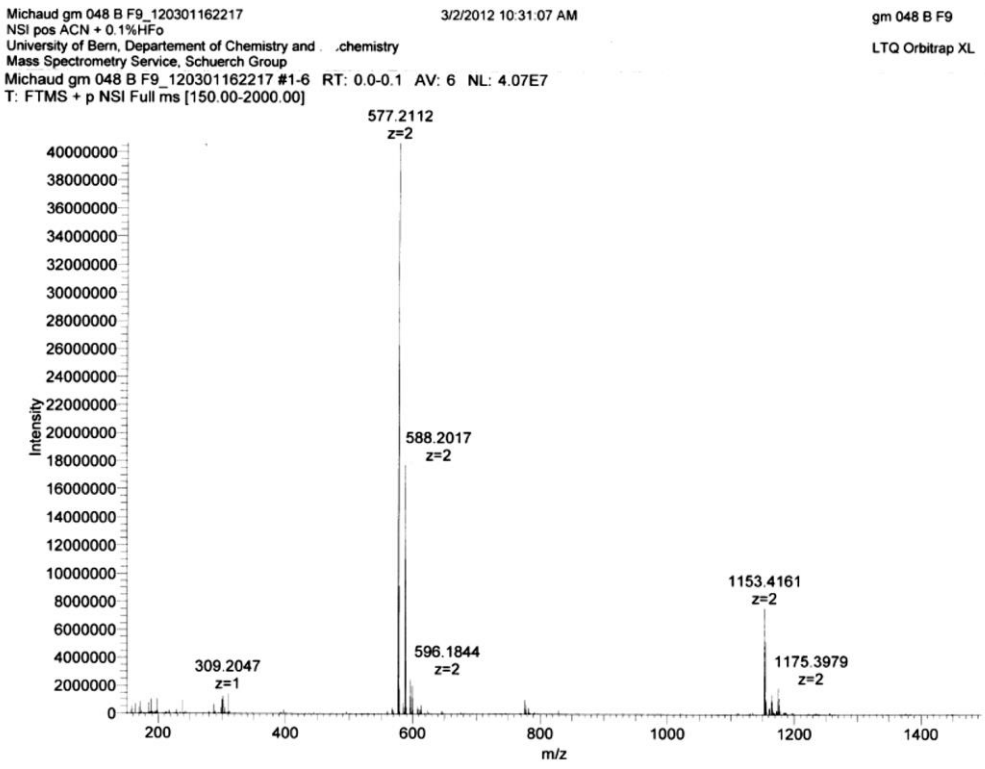
Preparative RP-HPLC was performed with a PrepLC4000 chromatography system (Waters) using a Reprospher Column (Dr. Maish GmbH, C18-DE, 5 µm, 100 × 30 mm, pore size 100 Å, flow rate of 60 $mL \cdot min^{-1}$). Compounds were detected by UV absorption at 214 nm using a 486 Tunable Absorbance detector (Waters). The elution solutions were: A H_2O with 0.1% TFA; B H_2O/MeCN (50:50); C H_2O/MeCN (10:90) with 0.1% TFA; D H_2O/MeCN (40:60) with 0.1% TFA.

MS spectra were recorded on a LTQ OrbitrapXL (ThermoFischer Scientific).

Synthesis of 5-CF-Gly-Ser-Asp-Gln-Ala-Ala-Thr-Phe-NH_2 (compound 2)

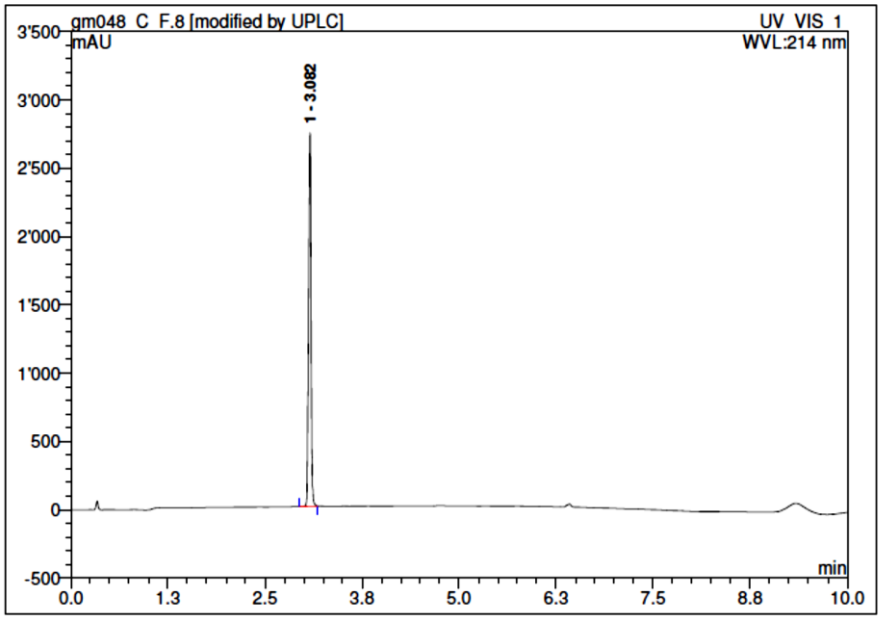
The purest fraction was isolated as a yellow foamy solid after preparative RP-HPLC purification (11 mg, 7.6 % yield). The gradient used for preparative RP-HPLC was A/D = 100/0 to 60/40 in 45 min, A/D 60/40 to 0/100 in 10 min with a flow rate of 40 $mL \cdot min^{-1}$. Analytical. UHPLC: t_R = 3.15 min (A/C = 100/0 to 0/100 in 7.5 min, flow rate 1.2 $mL \cdot min^{-1}$).
MS (ESI^+) calc. for $C_{54}H_{60}N_{10}O_{19}$ $[M+H]^+$: 1152.40, obsd: 1153.42 (z=1), 577.21 (z=2).

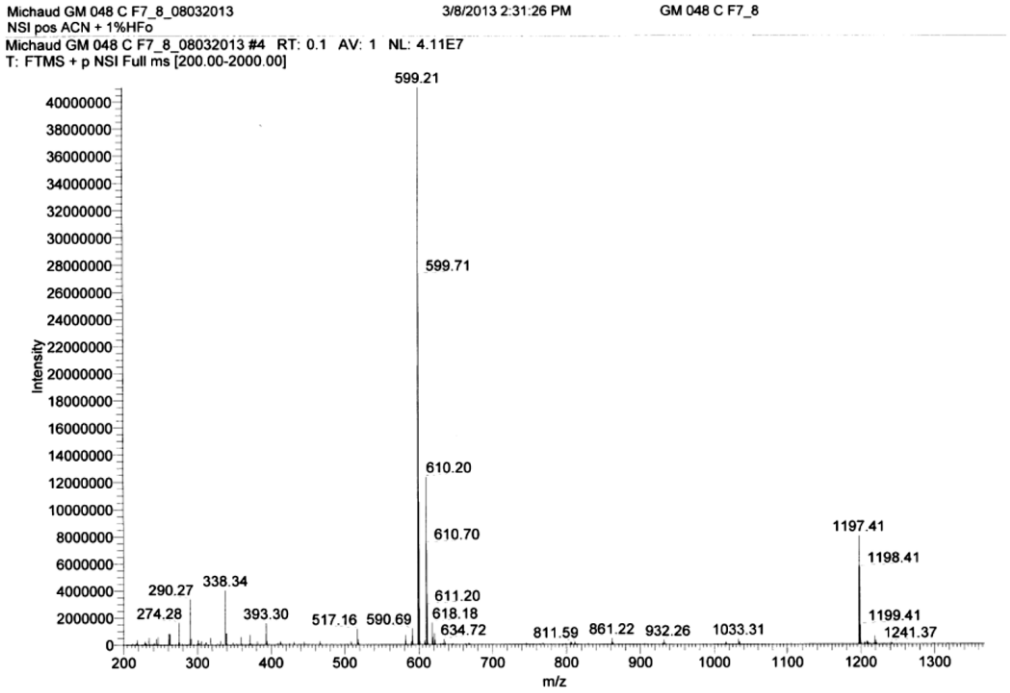




Synthesis of 5-CF-Gly-Ser-Asp-Gln-Asp-Ala-Thr-Phe-NH_2 (compound 3)

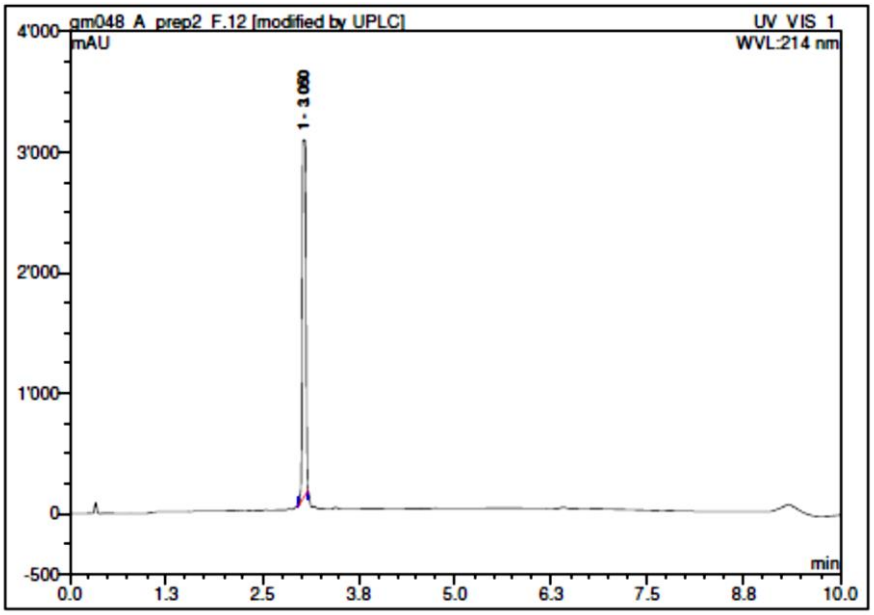
The purest fraction was isolated as a yellow foamy solid after preparative RP-HPLC purification (5.3 mg, 3.5 % yield). The gradient used for preparative RP-HPLC was A/D = 100/0 to 60/40 in 45 min, A/D 60/40 to 0/100 in 10 min with a flow rate of 40 $mL \cdot min^{-1}$. Analytical. UHPLC: t_R = 3.08 min (A/C = 100/0 to 0/100 in 7.5 min, flow rate 1.2 $mL \cdot min^{-1}$).
MS (ESI^+) calc. for $C_{55}H_{60}N_{10}O_{21}$ $[M+H]^+$: 1196.39, obsd: 1197.41 (z=1), 599.21 (z=2).

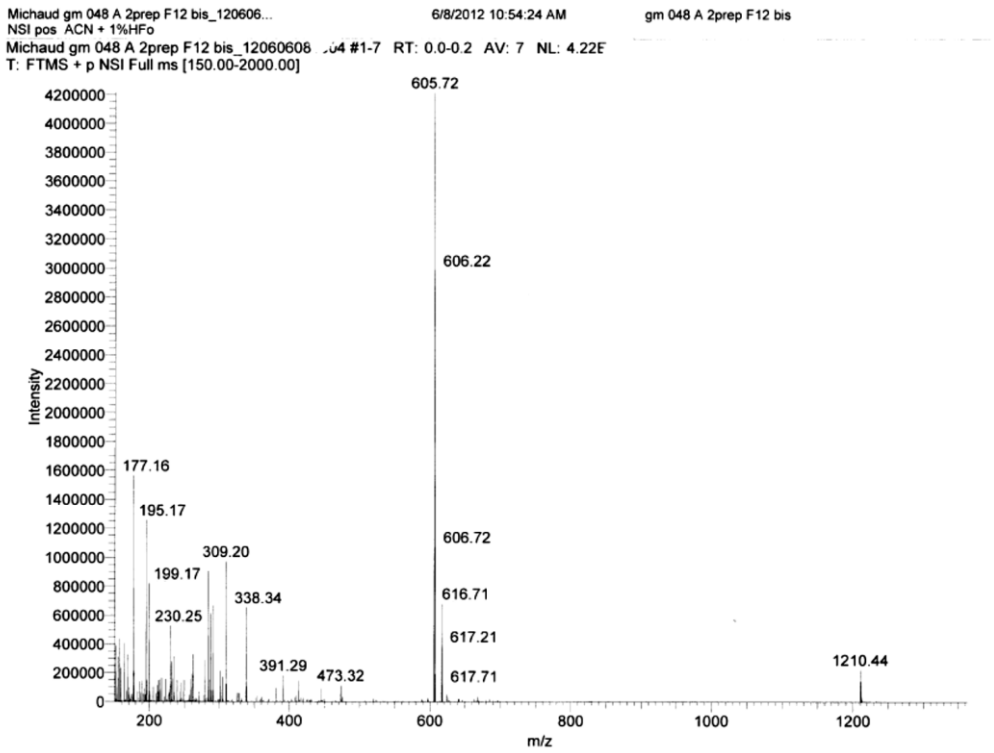




Synthesis of 5-CF-Gly-Ser-Asp-Gln-Gln-Ala-Thr-Phe-NH_2 (compound 4)

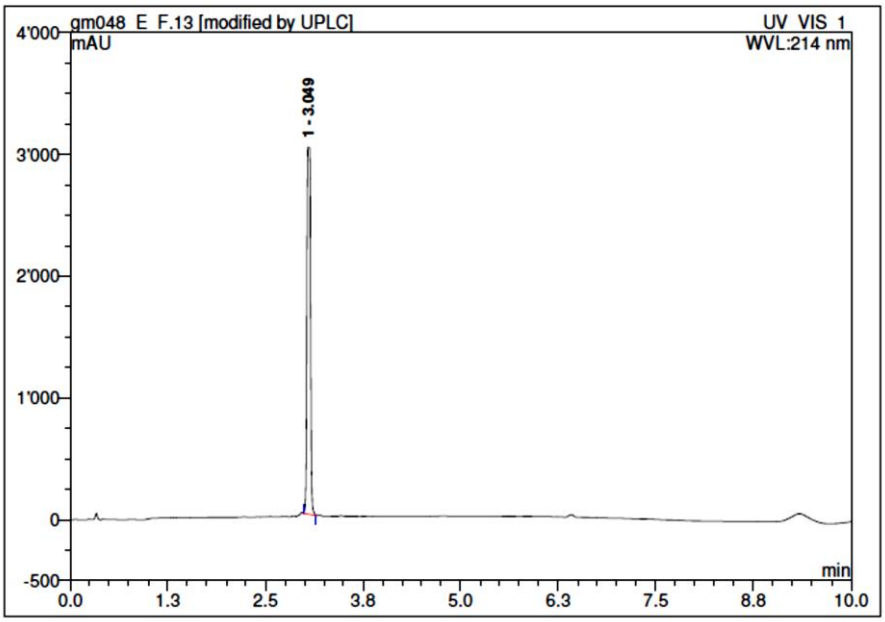
The purest fraction was isolated as a yellow foamy solid after preparative RP-HPLC purification (9 mg, 6 % yield). The gradient used for preparative RP-HPLC was A/D = 100/0 to 60/40 in 45 min, A/D 60/40 to 0/100 in 10 min with a flow rate of 40 $mL \cdot min^{-1}$. Analytical. UHPLC: t_R = 3.06 min (A/C = 100/0 to 0/100 in 7.5 min, flow rate 1.2 $mL \cdot min^{-1}$).
MS (ESI^+) calc. for $C_{56}H_{63}N_{11}O_{20}$ $[M+H]^+$: 1209.43, obsd: 1210.44 (z=1), 605.72 (z=2).

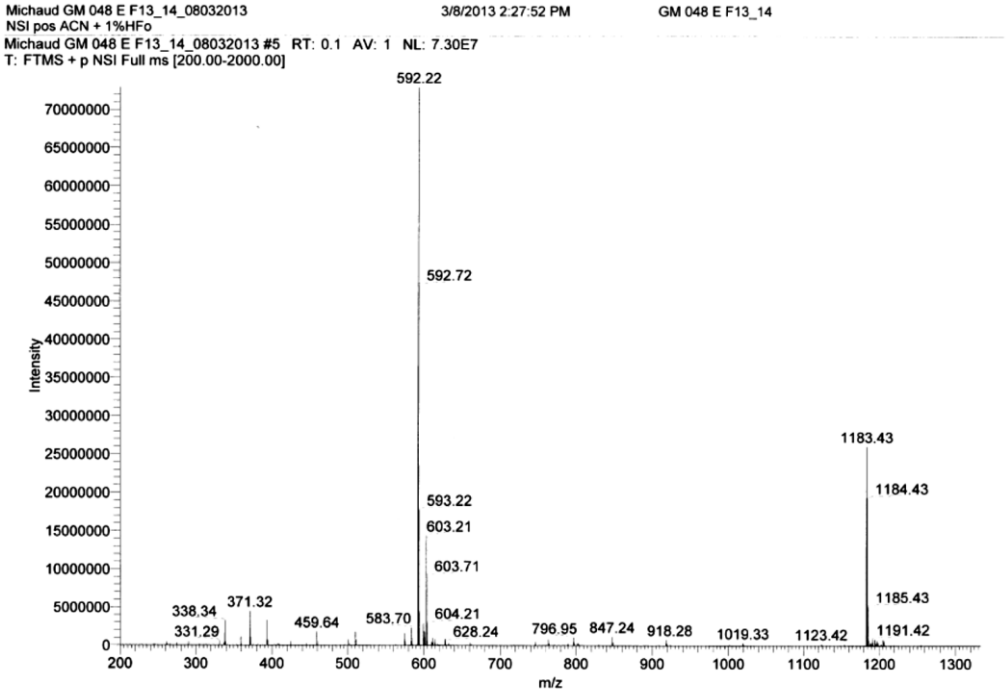




Synthesis of 5-CF-Gly-Ser-Asp-Gln-Homoserine-Ala-Thr-Phe-NH_2 (compound 5)

The purest fraction was isolated as a yellow foamy solid after preparative RP-HPLC purification (12.6 mg, 8.5 % yield). The gradient used for preparative RP-HPLC was A/D = 100/0 to 60/40 in 45 min, A/D 60/40 to 0/100 in 10 min with a flow rate of 40 $mL \cdot min^{-1}$. Analytical. UHPLC: t_R = 3.05 min (A/C = 100/0 to 0/100 in 7.5 min, flow rate 1.2 $mL \cdot min^{-1}$).
MS (ESI^+) calc. for $C_{55}H_{62}N_{10}O_{20}$ $[M+H]^+$: 1182.41, obsd: 1183.43 (z=1), 592.22 (z=2).



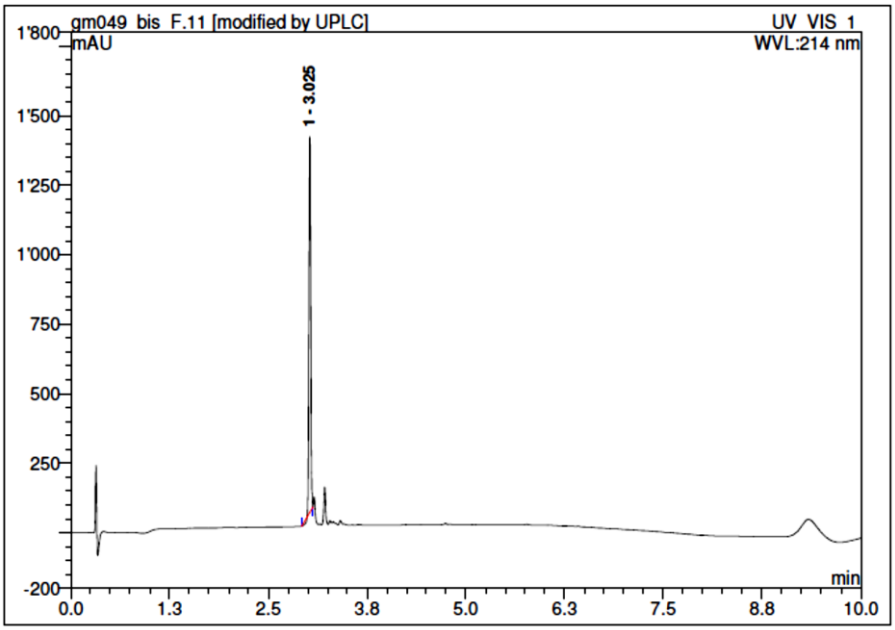


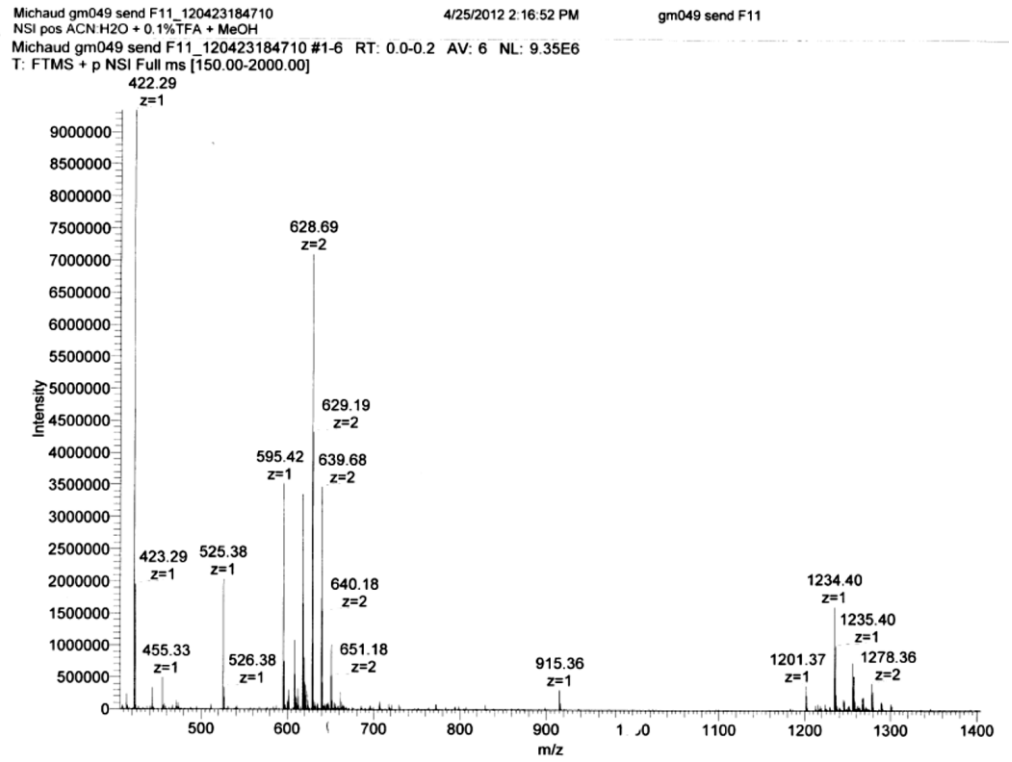
Synthesis of 5-CF-Gly-Ser-Asp-Gln-Asp(NHOH)-Ala-Thr-Phe-NH_2 (compound 6)

Peptide synthesis was performed by linear SPPS of the following peptide: FmocNH-Gly-Ser-Asp-Gln-Asp(OAllyl)-Ala-Thr-Phe-NH_2. The allyl protecting group was removed using the Allyl/Alloc deprotection procedure before *O*-(tert-butyldimethylsilyl)hydroxylamine (9 eq) was activated with PyBOP (5 eq) in DCM for 5 min and added to the resin. DIPEA (6 eq) was added and the resin was stirred for 3 h. The resin was washed with NMP (2 x 6 mL), MeOH (2 x 6 mL) and DCM (2 x 6 mL). Finally, the Fmoc group was removed and 5-CF was coupled to the peptide.

The purest fraction was isolated as a yellow foamy solid after preparative RP-HPLC purification (1 mg, 0.5 % yield). The gradient used for preparative RP-HPLC was A/D = 100/0 to 60/40 in 45 min, A/D 60/40 to 0/100 in 10 min with a flow rate of 40 $mL \cdot min^{-1}$. Analytical. UHPLC: t_R = 3.03 min (A/C = 100/0 to 0/100 in 7.5 min, flow rate 1.2 $mL \cdot min^{-1}$).

MS (ESI^+) calc. for $C_{55}H_{61}N_{11}O_{21}$ $[M+H]^+$: 1211.40; $[M+Na]^+$: 1234.39, obsd: 1234.40 (z=1), 628.69 (z=2).



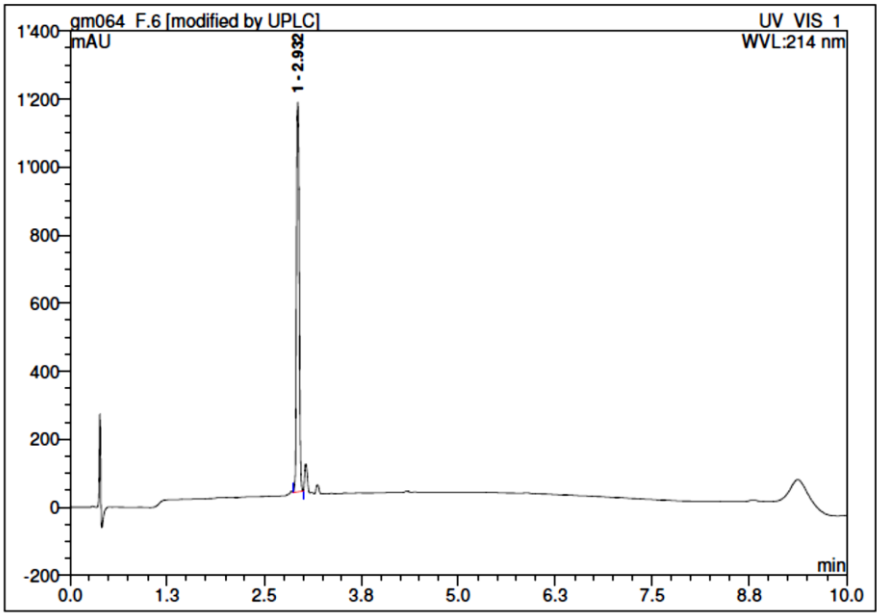


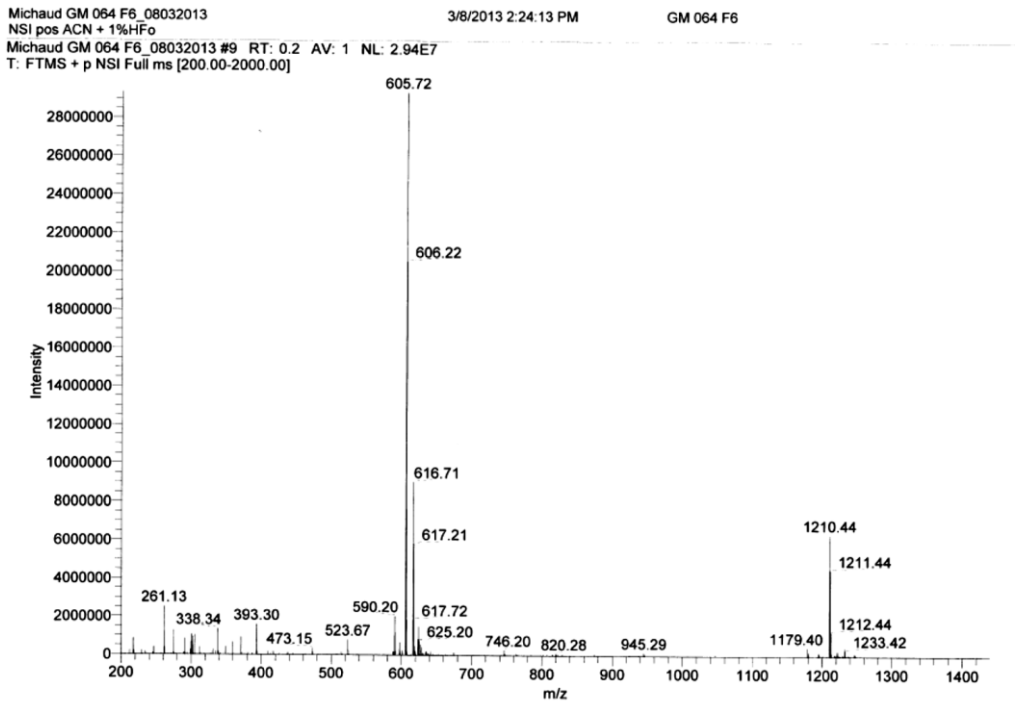
Synthesis of 5-CF-Gly-Ser-Asp-Gln-Asn(Me)-Ala-Thr-Phe-NH_2 (compound 7)

Peptide synthesis was performed by using the same peptide as described in 7. The allyl protecting group was removed using the Allyl/Alloc deprotection procedure. Then methylamine hydrochloride (3 eq) was dissolved in NMP and activated with PyBOP (5 eq) for 5 min. The further procedure was as described in 7.

The purest fraction was isolated as a yellow foamy solid after preparative RP-HPLC purification (2 mg, 1.3 % yield). The gradient used for preparative RP-HPLC was A/D = 100/0 to 60/40 in 45 min, A/D 60/40 to 0/100 in 10 min with a flow rate of 40 $mL \cdot min^{-1}$. Analytical. UHPLC: t_R = 2.93 min (A/C = 100/0 to 0/100 in 7.5 min, flow rate 1.2 $mL \cdot min^{-1}$).

MS (ESI^+) calc. for $C_{56}H_{63}N_{11}O_{20}$ $[M+H]^+$: 1209.43; obsd: 1210.44 (z=1), 605.72 (z=2).

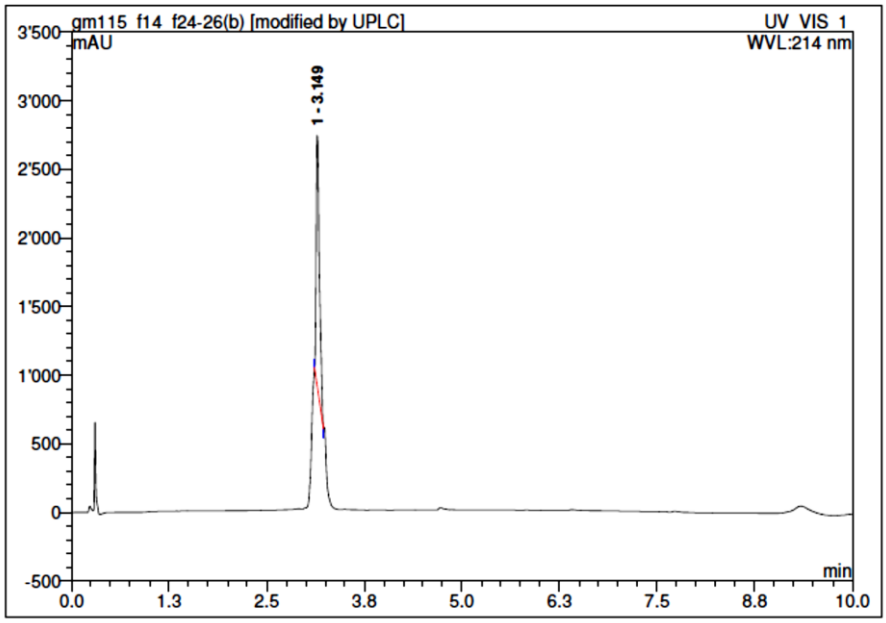


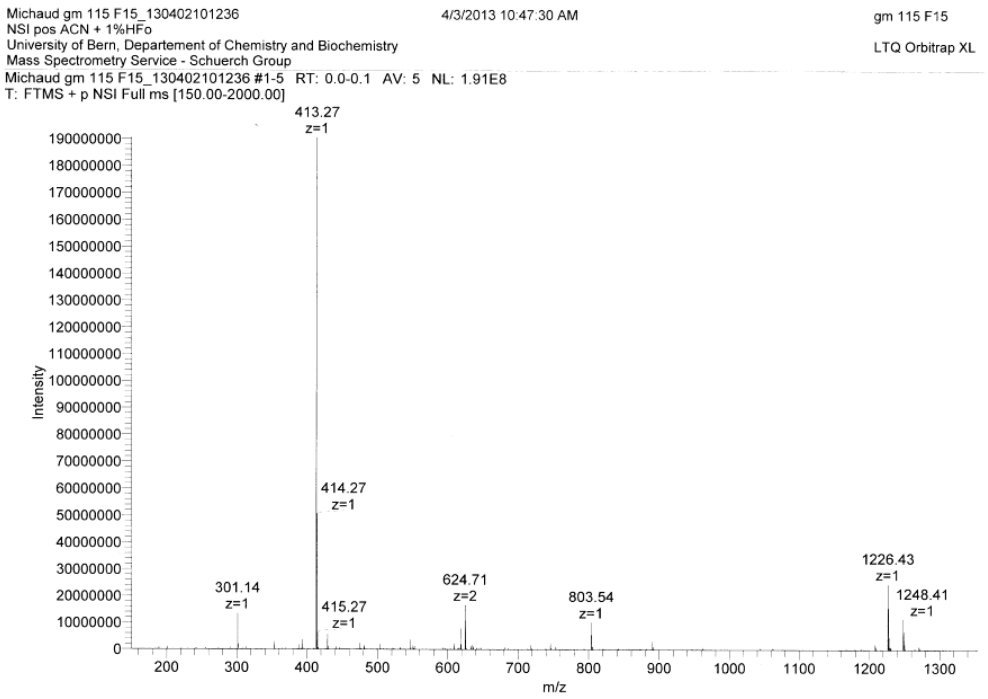


Synthesis of 5-CF-Gly-Ser-Asp-Gln-Asp(NHOMe)-Ala-Thr-Phe-NH_2 (compound 8)

Peptide synthesis was performed by linear SPPS of the following peptide: FmocNH-Gly-Ser-Asp-Gln-Asp(OAllyl)-Ala-Thr-Phe-NH_2, the dipeptide Ala-Thr was introduced using the following pseudoproline building block: Fmoc-Ala-Thr($\Psi^{Me,Me}$pro)-OH. The allyl protecting group was removed using the Allyl/Alloc deprotection procedure. Then *O*-methylhydroxylamine (9 eq) was activated with PyBOP (5 eq) in DCM for 5 min and added to the resin. The further procedure was as described in 7.

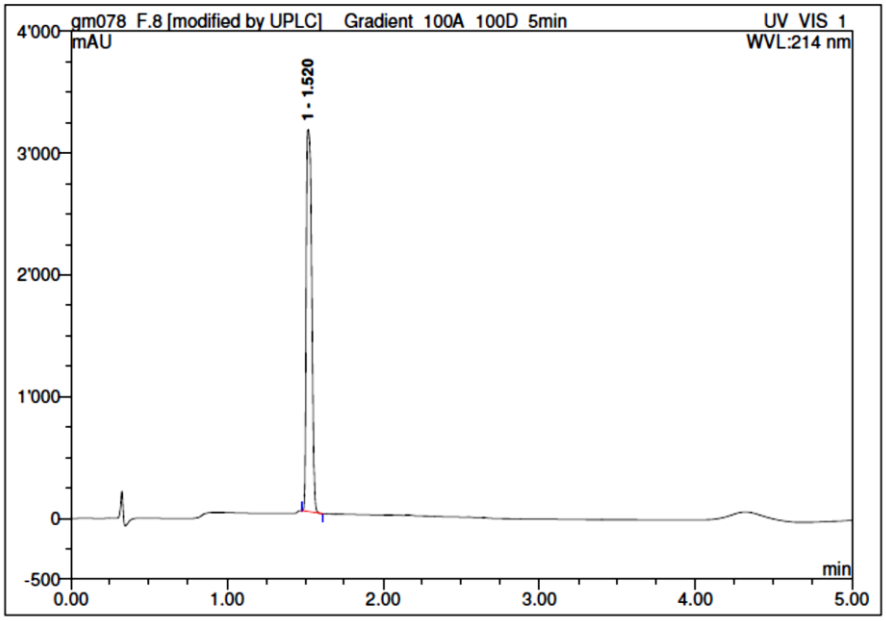
The purest fraction was isolated as a yellow foamy solid after preparative RP-HPLC purification (4.3 mg, 3.0 % yield). The gradient used for preparative RP-HPLC was A/D = 100/0 to 60/40 in 45 min, A/D 60/40 to 0/100 in 10 min with a flow rate of 40 $mL \cdot min^{-1}$. Analytical. UHPLC: t_R = 3.15 min (A/C = 100/0 to 0/100 in 7.5 min, flow rate 1.2 $mL \cdot min^{-1}$).
MS (ESI^+) calc. for $C_{56}H_{63}N_{11}O_{21}$ $[M+H]^+$: 1225.42; obsd: 1126.43 (z=1).

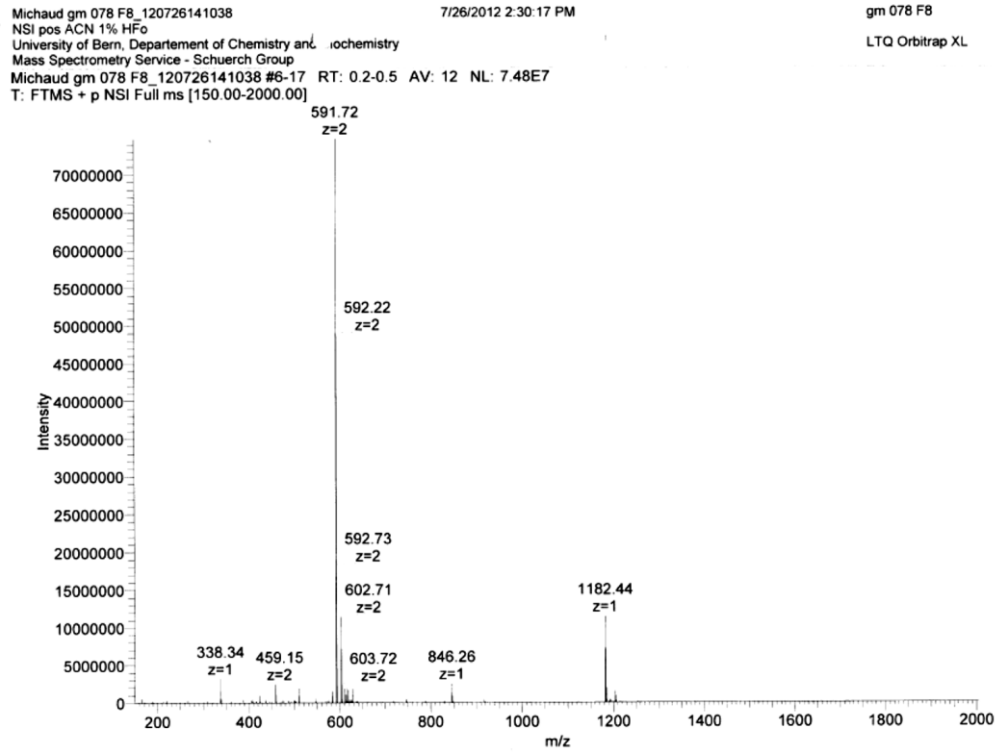




Synthesis of 5-CF-Gly-Ser-Asp-Gln-Dab-Ala-Thr-Phe-NH_2 (compound 9)

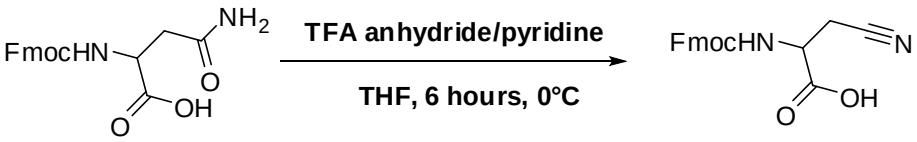
The purest fraction was isolated as a yellow foamy solid after preparative RP-HPLC purification (27.4 mg, 19 % yield). The gradient used for preparative RP-HPLC was A/D = 100/0 to 60/40 in 45 min, A/D 60/40 to 0/100 in 10 min with a flow rate of 40 $mL \cdot min^{-1}$. Analytical. UHPLC: t_R = 1.52 min (A/C = 100/0 to 0/100 in 2.2 min, flow rate 1.2 $mL \cdot min^{-1}$).
MS (ESI^+) calc. for $C_{55}H_{63}N_{11}O_{19}$ $[M+H]^+$: 1181.43; obsd: 1182.44 (z=1), 591.72.(z=2).





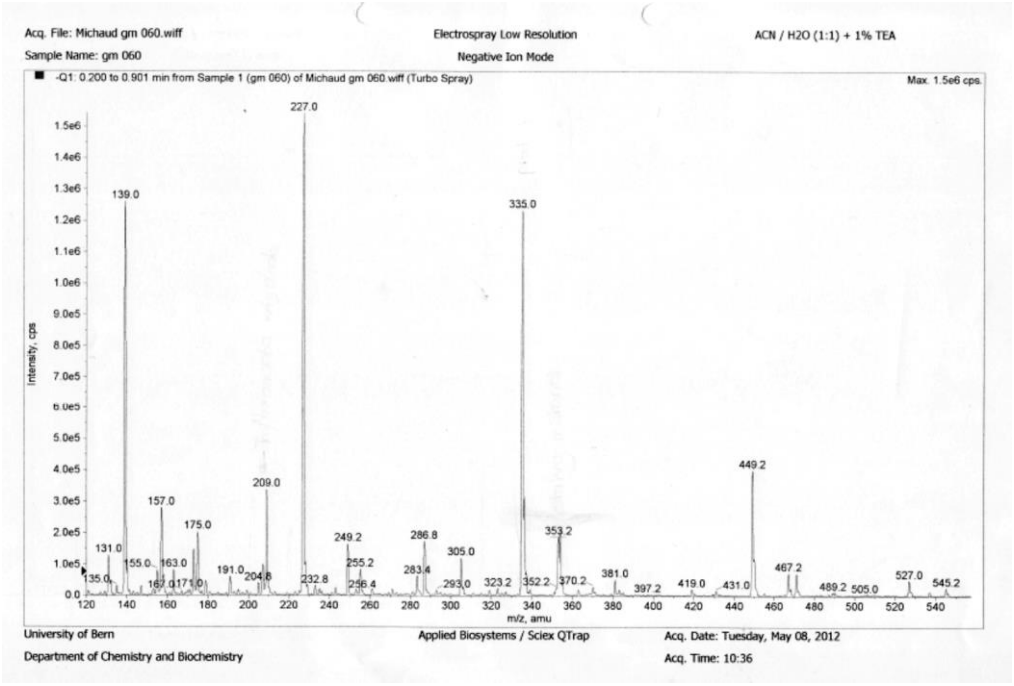
Synthesis of the Asn(Nitrile) amino acid building block:

Amino acid synthesis was performed as previously described[57].



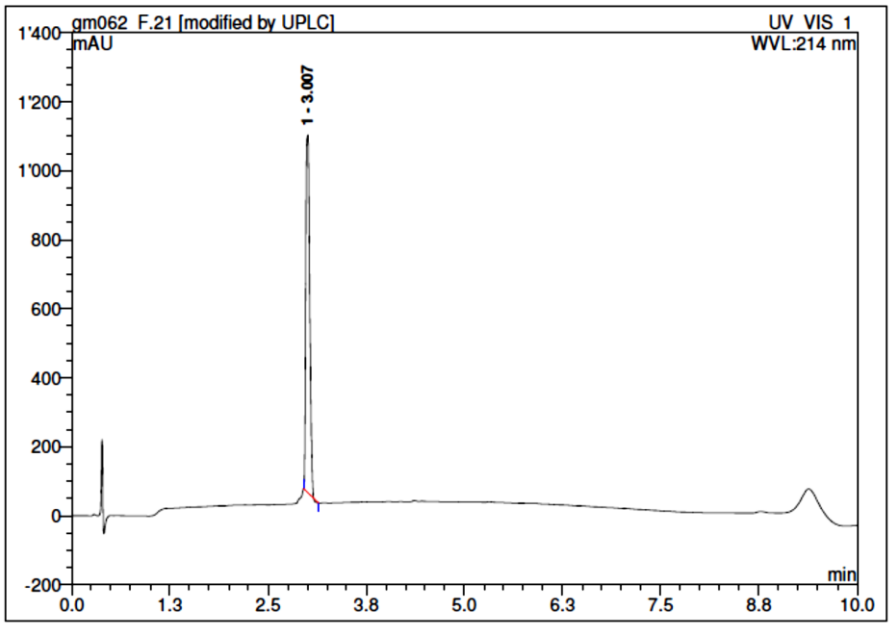
Following a slight modification of the reported procedure[57], Fmoc-Asn-OH (500 mg, 1.41 mmol, 1 eq) was loaded in a 50 mL flask under argon. The solid was dissolved in dry THF (11.4 mL) and cooled down in an ice bath for 5 min. TFA anhydride (0.22 mL, 1.55 mmol, 1.1 eq) was added drop wise. The reaction mixture was stirred for 5 min at 0°C. Dry pyridine (0.23 mL, 2.82 mmol, 2 eq) was added drop wise and the reaction mixture was stirred for 6 h in an ice bath. The solvent was removed under reduce pressure and the product dried under high vacuum for several hours. The product was used as crude.

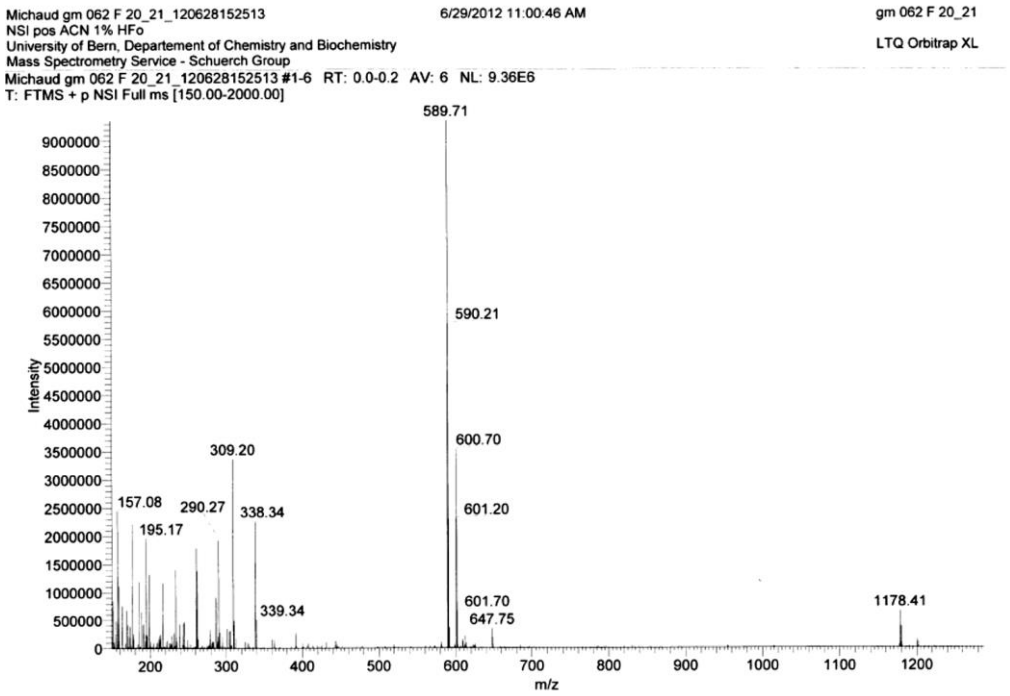
MS (ESI$^-$) calc. for $C_{19}H_{16}N_2O_4$ [M-H]$^-$: 336.11 obsd: 335.00.



Synthesis of 5-CF-Gly-Ser-Asp-Gln-Asn(C≡N)-Ala-Thr-Phe-NH_2 (compound 10)

The purest fraction was isolated as a yellow foamy solid after preparative RP-HPLC purification (1.8 mg, 1.2 % yield). The gradient used for preparative RP-HPLC was A/D = 100/0 to 60/40 in 45 min, A/D 60/40 to 0/100 in 10 min with a flow rate of 40 $mL \cdot min^{-1}$. Analytical. UHPLC: t_R = 3.01. min (A/C = 100/0 to 0/100 in 7.5 min, flow rate 1.2 $mL \cdot min^{-1}$).
MS (ESI^+) calc. for $C_{55}H_{59}N_{11}O_{19}$ $[M+H]^+$: 1177.40; obsd: 1178.41 (z=1), 589.71 (z=2).





Supplementary Reference

57. Sureshbabu, V.V., Venkataramanarao, R., Naik, S.A. & Chennakrishnareddy, G. Synthesis of tetrazole analogues of amino acids using Fmoc chemistry: isolation of amino free tetrazoles and their incorporation into peptides. *Tetrahedron Letters* **48**, 7038-7041 (2007).