Supporting Information

for

Semi-Synthesis of Aminomethyl-Insulin: An Atom-Economic Strategy to Increase the Affinity and Selectivity of a Protein for Recognition by a Synthetic Receptor

Hayden R. Anderson,† Wei L. Reeves,† Andrew T. Bockus, Paolo Suating, Amy G. Grice, Madeleine Gallagher, and Adam R. Urbach*

Department of Chemistry, Trinity University, San Antonio, TX, 78212

† authors contributed equally to this work

* to whom correspondence should be addressed: aurbach@trinity.edu

Table of Contents

Materials and Methods

Materials. The following compounds were of ≥ 95% grade and used without purification: sodium sulfate, 2-(methylsulfonyl) ethanol, N-hydroxysuccinimide, biotech-grade dimethyl formamide (DMF), anhydrous dichloromethane (DCM), methanol, triisopropylsilane (TIS), trifluoroacetic acid (TFA), piperidine, formic acid, HPLC grade acetonitrile, sodium hydroxide, 4-nitropheynyl chloroformate, pyridine, O-(benzotriazol-1-yl)-N,N,N',N'- tetramethyluronium hexafluorophosphate (HBTU) (Sigma-Aldrich); ethyl acetate, 2-propanol, monobasic and dibasic sodium phosphate (VWR); anhydrous ethyl ether, 1,4-dioxane (Fisher Scientific); hydrochloric acid (J.T. Baker); human recombinant insulin (SAFC Biosciences); 4-ethylmorpholine (Fluka Analytical); phenyl isothiocyanate (PITC) (Acros Organics); (L)-4-aminomethyl phenylalanine (Chem-Impex International Inc.); rink amide MBHA resin, (L)-Fmoc-Val-OH, (L)-Fmoc-Phe-OH, (L)-Fmoc-Asn(Trt)-OH, (L)-Fmoc-Gln(Trt)-OH (Peptides International); deuterium oxide, "100%" (D, 99.96%) (Cambridge Isotopes). Water was obtained from a Barnstead Nanopure Infinity water system (18 MΩ cm). Cucurbit[7]uril (Q7) was synthesized according to a published procedure,[1] and standardized by calorimetric titration. Methylsulfonylethyl *p*-nitrophenyl-carbonate (Msc-ONp) was synthesized in one step from commercially available starting materials according to a published procedure.[2]

Circular Dichroism (CD) Spectroscopy. CD Spectra for native human insulin and AM-insulin were acquired on a Jasco J-815 circular dichroism spectropolarimeter at 25 °C with a 0.1 cm pathlength, 1.0 nm pitch, 2.0 nm bandwidth, 8 second response time, 50 nm/min scan rate. Both samples were at 10 μM protein in buffer containing 10 mM sodium phosphate, 1 mM lysine, and 4 mM EDTA, pH 7.0.

***RP-HPLC Analysis*.** Product mixtures in the insulin semi-synthesis were analyzed by analytical reversed-phase high performance liquid chromatography (RP-HPLC) using a C8 analytical column (Aeris WIDEPORE 3.6 μ XB-C8, 100×2.1 mm, Phenomenex, part No: 00D-4481-AN), in 0.1% aqueous formic acid with a gradient of 1%/min of a solution containing 0.1 % formic acid in 9:1 acetonitrile:H_2O, at a flow rate of 0.2 mL/min, and monitoring at 220 nm. Samples were prepared by dissolving dry protein in 0.1% aqueous formic acid to make a 1 mg/mL stock solution, and then diluting the stock with 0.1 % aqueous formic acid to a final protein concentration of 0.05 mg/mL.

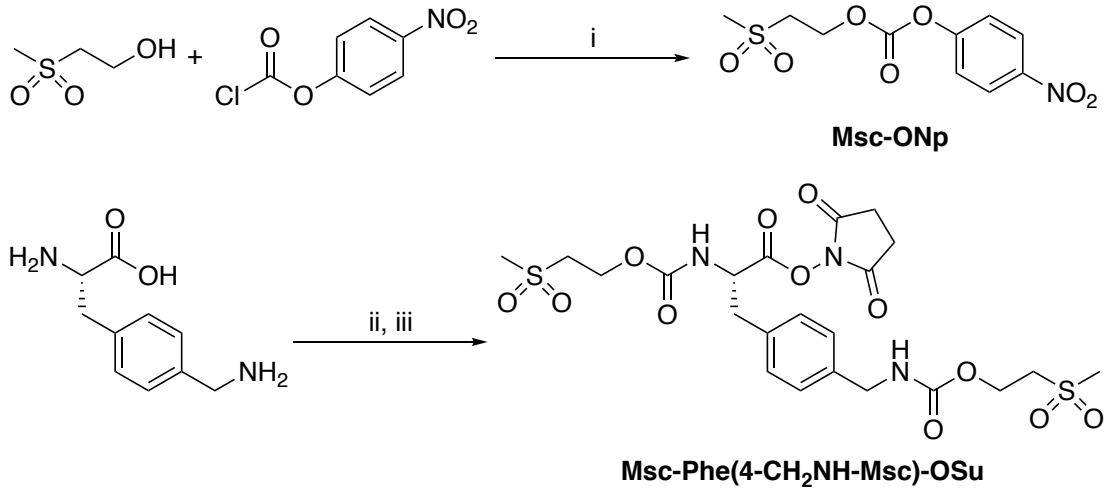
HPLC-ESI-MS Analysis. Product mixtures in the insulin semi-synthesis were analyzed via high-performance liquid chromatography with electrospray ionization time-of-flight mass spectrometry (HPLC-ESI-TOF-MS) using an Agilent 1260 Infinity II HPLC in tandem with an Agilent 6230 TOF LC/MS at the UT Health San Antonio Institutional Mass Spectrometry Core. The HPLC used a C8 analytical column (Aeris WIDEPORE 3.6 μ XB-C8, 100×2.1mm, Phenomenex, part No: 00D-4481-AN), in 0.1% aqueous formic acid with a gradient of 5.5%/min of a solution containing 0.1 % formic acid in 9:1 acetonitrile:H_2O, at a flow rate of 0.2 mL/min, and monitoring at 220 nm. The total ion chromatograms and mass spectra were used to confirm the purity and identity of the products and binding to Q7.

Nuclear Magnetic Resonance (NMR) Spectroscopy. ^{1}H NMR spectra were acquired at concentrations of 0.4 - 1 mM solute dissolved in either deuterium oxide (D_2O) or DMSO-d_6 at 25 °C on an Agilent 500 MHz spectrometer. Presaturation of signal from residual protiated solvent was used as necessary for spectra collected in D_2O solution.

Isothermal Titration Calorimetry (ITC). Titrations were carried out at 300 K in 10 mM sodium phosphate buffer, pH 7.0, using a VP-ITC calorimeter (Microcal, Inc). In each experiment, Q7 was in the injection syringe at a concentration in the range 0.5 – 1.2 mM, and the protein or peptide was in the cell at a concentration in the range 0.03 – 0.1 mM. All solutions were degassed prior to titration. The titration schedule consisted of 28 consecutive injections of 2-10 μL with at least a 200 s interval between injections. Heats of dilution were subtracted from each data set. The data were analyzed using Origin software and fit by nonlinear regression to a binary equilibrium model (non-interacting sites) supplied with the software in order to determine the molar enthalpy, equilibrium association constant, and stoichiometry of binding. These values were then used to calculate the free energies of binding and the entropic contributions to the free energies of binding.

Electrospray Ionization Mass Spectrometry (ESI-MS). Mass spectra of purified Msc-ONp, Msc-Phe(CH_2NH-Msc)-OH, and Msc-Phe(CH_2NH-Msc)-OSu products, as well as the tetrapeptides in the presence and absence of Q7, were acquired using a Thermo LCQ DECA XP mass spectrometer with an electrospray ion source in positive ion mode. All samples were dissolved in pure water.

Synthesis of Msc-Phe(4-CH_2NH-Msc)-OSu (6)



Scheme S1. Synthesis of Msc-Phe(4-CH_2NH-Msc)-OSu (**6**): i. pyridine, r.t., overnight; ii. MscONp, CH_3CN/H_2O, Et_3N, r.t., overnight; iii. DCC, NHS, DMF, r.t., overnight

i. Methylsulfonylethyl-*p*-nitrophenyl-carbonate (Msc-ONp, Scheme S1) was synthesized in one step from commercially available starting materials according to a published procedure.[2]

ii. 4-aminomethyl phenylalanine (429 mg, 2.21 mmol) and Msc-ONp (1.41 g, 4.87 mmol) were suspended in 40 mL acetonitrile/H_2O (4:1 v/v). Freshly distilled Et_3N (710 µL) was added to the suspension. The reaction mixture was stirred overnight at r.t. and then under high vacuum at r.t. The yellow residue was redissolved in 15 mL H_2O and adjusted to pH 5 with 6 M HCl. The resulting solution was washed with diethyl ether (15 mL, 3X). The aqueous layer was further acidified to pH 3 with 6 M HCl and extracted with ethyl acetate (15 mL, 3X). The organic layers were combined, dried over anhydrous Na_2SO_4, and concentrated under high vacuum at r.t. to obtain Msc-Phe(4-CH_2NH-Msc)-OH as a white solid, 928 mg (85% yield). ^{1}H NMR, 25 °C, DMSO-d^6, 500 MHz, δ (ppm): 12.77 (1H, s), 7.87 (1H, t, J = 4 Hz), 7.74 (1H, d, J = 5 Hz), 7.20 (4H, m), 4.32 (2H, t, J = 4 Hz), 4.24 (2H, t, J = 6 Hz), 4.17 (2H, d, J = 6 Hz), 4.11 (1H, m), 3.47 (2H, d, J = 3 Hz), 3.40 (2H, t, J = 3 Hz), 3.04 (3H, d, J = 4 Hz), 2.96 (3H, s), 2.80 (2H, m) (Figure S1). ESI-MS Calc. m/z: 517.10 ($M+Na^+$), Found: 517.20.

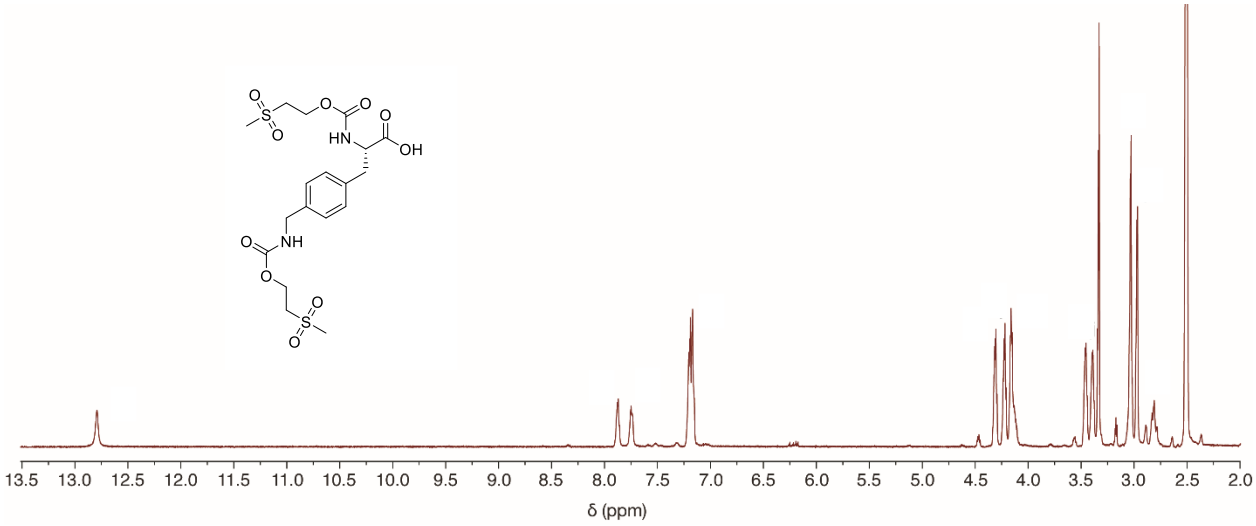


Figure S1. ^{1}H NMR spectrum of Msc-Phe(4-CH_2NH-Msc)-OH at 25 °C in DMSO-d_6.

iii. Msc-Phe(4-CH_2NH-Msc)-OH (55 mg, 0.111 mmol) and N-hydroxysuccinimide (12.8 mg, 0.111 mmol) were dissolved in biotech-grade DMF (300 µL). DCC (25.2 mg, 0.122 mmol) was added to the solution, and the reaction mixture was stirred overnight at r.t.. The white precipitate was removed by centrifugation, and the supernatant was concentrated overnight under high vacuum at r.t. The yellow residue was suspended in 10 mL methanol and stirred for 1 hour at r.t. The white precipitate was collected by centrifugation and dried under high vacuum at r.t. to obtain Msc-Phe(4-CH_2NH-Msc)-OSu as a white solid, 52 mg (80% yield). H^1 NMR 25 °C, DMSO-d^6, 500 MHz, δ (ppm): 8.25 (1H, d, J = 8 Hz), 7.88 (1H, t, J = 6 Hz), 7.30 (2H, d, J = 8 Hz), 7.19 (2H, d, J = 8 Hz), 4.65 (1H, m), 4.32 (2H, t, J = 6 Hz), 4.27 (2H, t, J = 4 Hz), 4.18 (2H, d, J = 6), 3.46 (2H, t, J = 6), 3.41 (2H, t, J = 6), 3.04 (3H, s), 3.01(2H, d, J = 3), 2.96 (3H, s) 2.84 (4H, s). (Figure S2). ESI-MS Calc. m/z: 614.12 ($M+Na^+$), Found: 614.20.

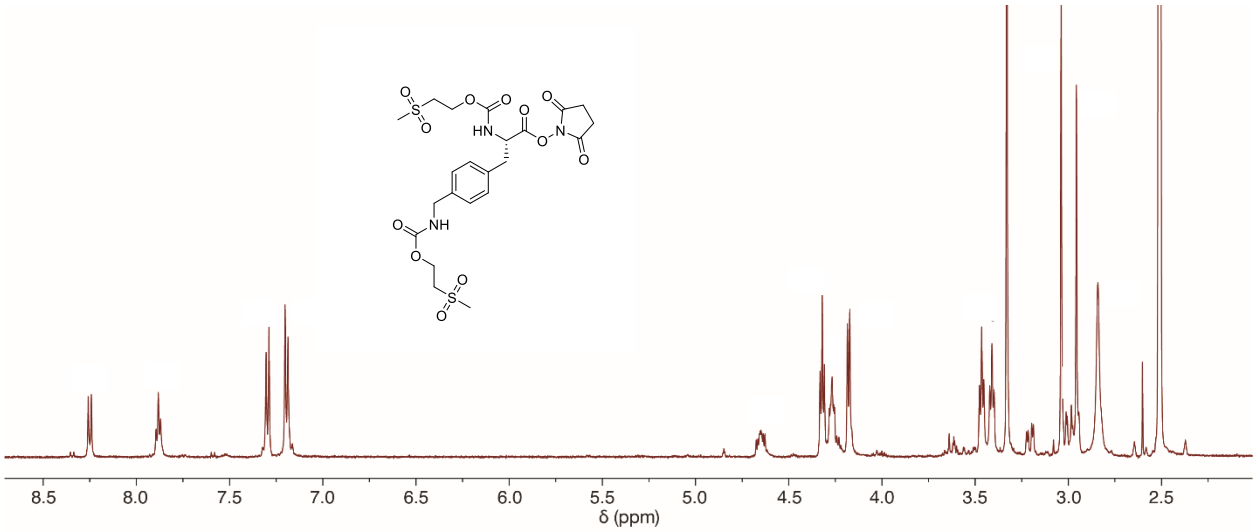
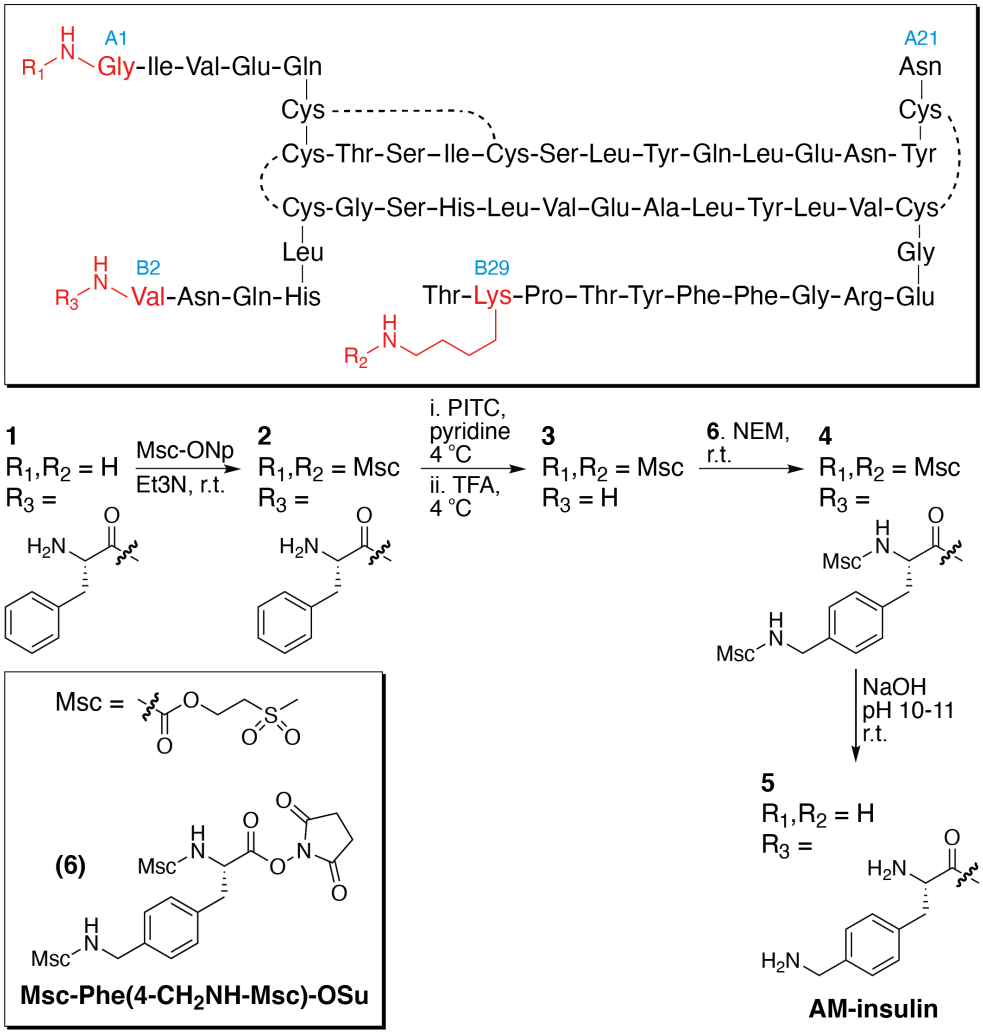


Figure S2. ^{1}H NMR spectra of Msc-Phe(4-CH_2NH-Msc)-OSu (**6**) at 25 °C in DMSO-d_6.

Semi-synthesis of AM-insulin



Scheme S2. Four-step semi-synthesis of AM-insulin from native insulin. Abbreviations: Msc-ONp, methyl sulfonylethyl p-nitrophenyl carbonate; Et3N, triethylamine; r.t., room temperature; PITC, phenyl isothiocyanate; TFA, trifluoroacetic acid; NEM, N-ethyl morpholine.

Procedures for the semi-synthesis of AM-insulin were developed based on prior work.[3-5]

i. A 5% (v/v) solution of Et_3N in DMF/H_2O (4:1 v/v) was prepared, and the pH was adjusted to 9-10 with 6 M HCl. Recombinant human insulin (490 mg, 84.4 μmol) was dissolved in 43 mL of the Et_3N solution. Msc-ONp (161 mg, 556 μmol) was added to the mixture, and the reaction was stirred for 20 h at r.t. The mixture was concentrated under high vacuum at r.t. until white crystals (the Et_3N•HCl salt) formed. The solid was isolated by centrifugation, and the supernatant (~10 mL) was poured into cold ethyl ether (200 mL). The precipitate was collected by centrifugation

and washed with cold ethyl ether. The resulting white solid was dried over high vacuum at r.t. to obtain 506 mg of the crude product as a white solid containing ~82% diMsc-insulin (**2**) based on HPLC traces detected at 220 nm (~81% yield **2**) (Figures S3 and S4). ESI-TOF-MS Calc. m/z: 6107.73, Found: 6108.30 (deconvoluted) (Figure S5).

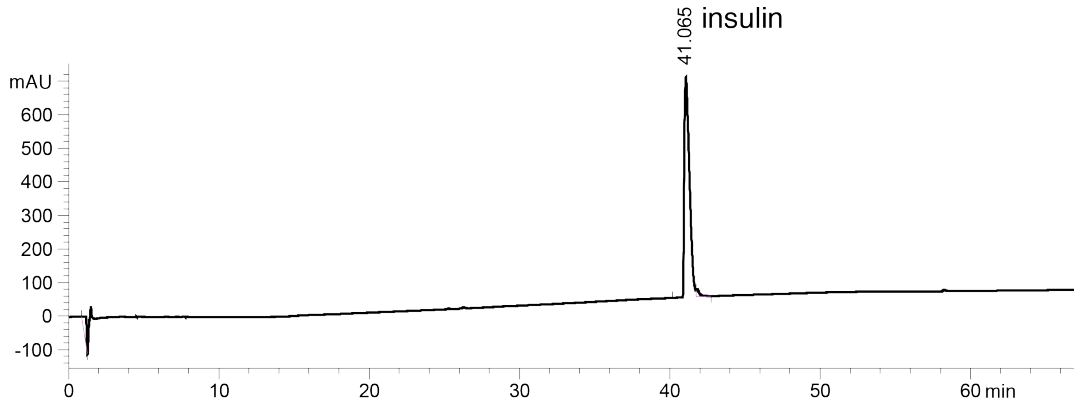


Figure S3. Analytical HPLC trace at 220 nm of pure, native human insulin **1**.

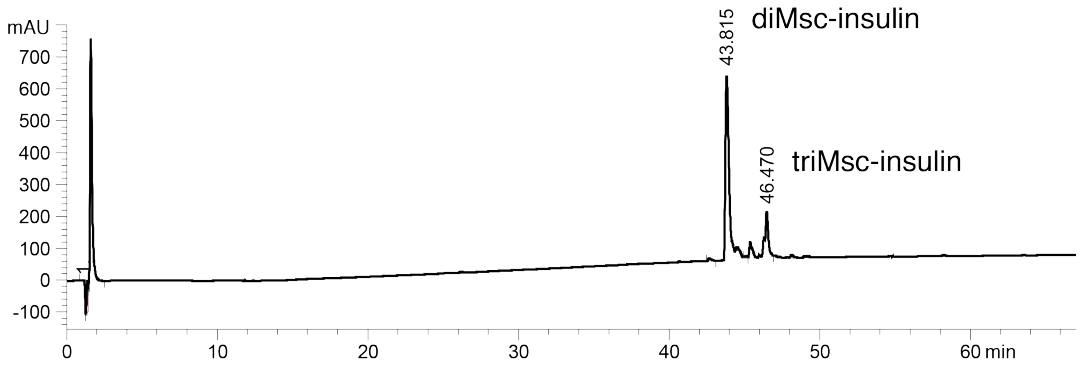


Figure S4. Analytical HPLC trace at 220 nm of the crude reaction mixture from the synthesis of diMsc-insulin (**2**).

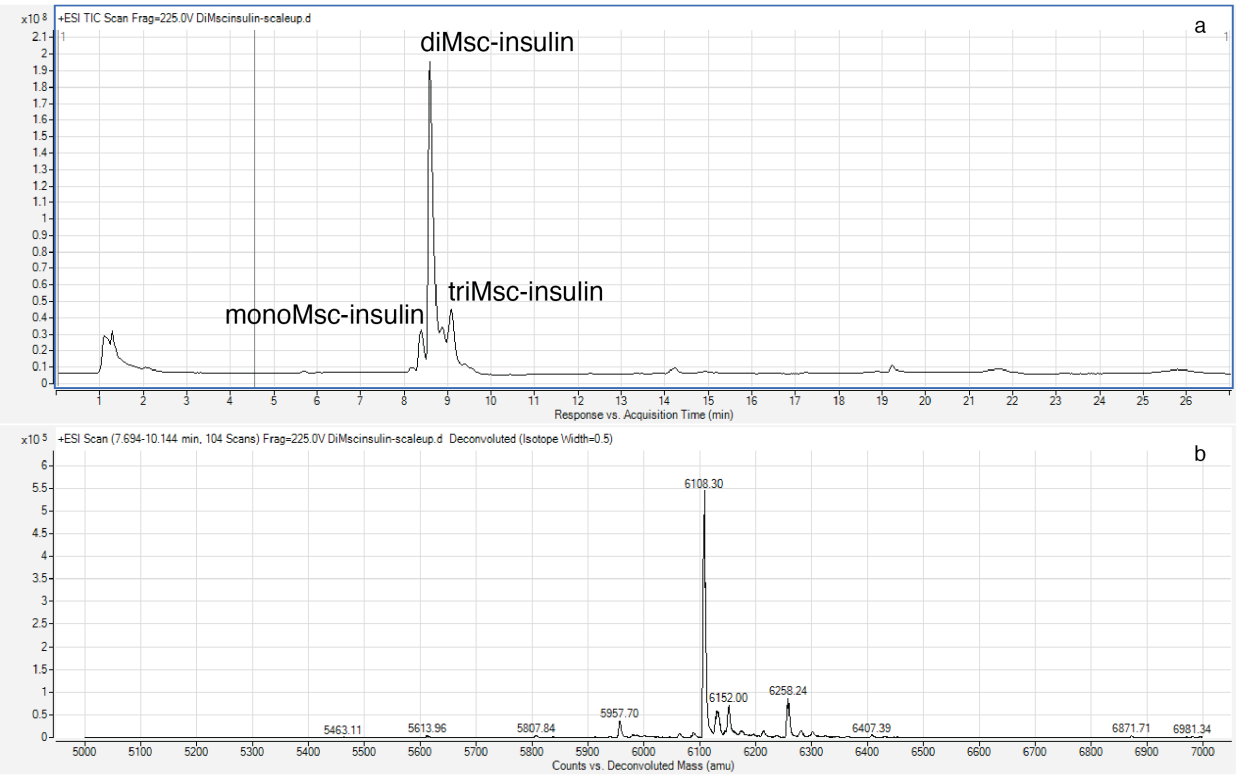


Figure S5. (top) Total ion chromatogram (TIC) of the crude reaction mixture from the synthesis of diMsc-insulin (**2**). (bottom) Deconvoluted ESI-TOF mass spectrum of the TIC at 7.694-10.144 minutes. Calc. m/z: 6107.73.

Synthesis of des(Phe^{B1})-diMsc-insulin (3)

ii. The crude product mixture from (i) (506 mg, ~82% diMsc-insulin, **2**, 68 μmol) was dissolved in 4 mL pyridine/ H_2O (3:1, v/v). PITC (133 mg, 118 μL) was added to the solution and the reaction mixture was kept at 4 °C without stirring for 24 h. The sample was concentrated under high vacuum at r.t. The addition of 4 mL 0.01 mM HCl (pH 5) to the residue caused the formation of white crystals. The precipitate was collected by centrifugation and dried under high vacuum at r.t.

iii. All of the crude product (450 mg) was dissolved in 30 mL pure TFA, and the solution was transferred to a 250 mL round bottom flask. The solution was allowed to stand for 45 min at 4 °C without stirring, and was then concentrated under high vacuum at r.t. Cold diethyl ether (40 mL) was added to precipitate the protein as a white solid. The solid was separated by centrifugation, washed with cold ethyl ether (40 mL), and dried under high vacuum at r.t to yield 420 mg of the crude product as a white solid containing ~87% des(Phe^{B1})-diMsc-insulin (**3**) based on HPLC traces detected at 220 nm (~90% conversion **2** → **3**) (Figure S6). ESI-TOF-MS Calc. m/z: 5960.66, Found: 5961.14 (deconvoluted) (Figure S7).

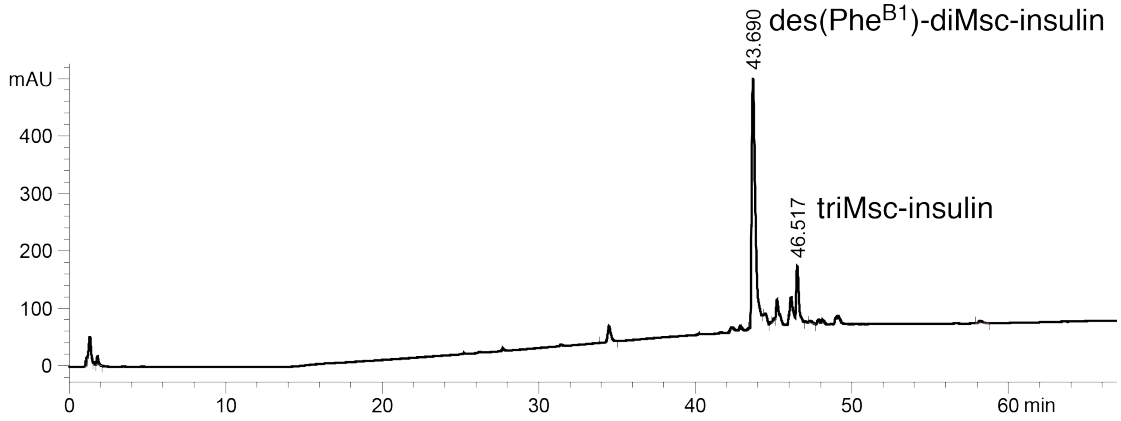


Figure S6. Analytical HPLC trace at 220 nm of the crude reaction mixture from the synthesis of des(Phe^{B1})-diMsc-insulin (**3**).

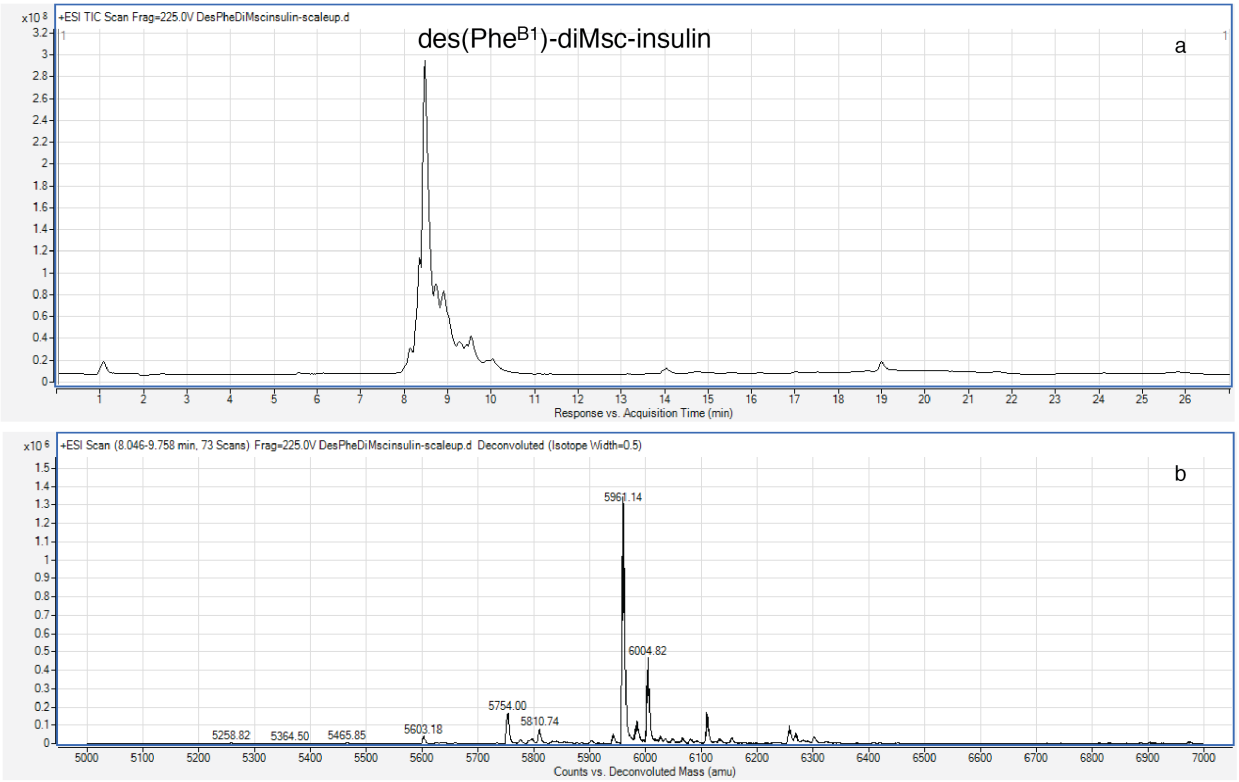


Figure S7. (top) Total ion chromatogram (TIC) of the crude reaction mixture from the synthesis of des(Phe^{B1})-diMsc-insulin (**3**). (bottom) Deconvoluted ESI-TOF mass spectrum of the TIC at 8.046-9.758 minutes. Calc. m/z: 5960.66.

Synthesis of tetraMsc-AM-insulin (4)

iv. The crude mixture from (iii) (385 mg, ~87% des(Phe^{B1})-diMsc-insulin, **3**, 56 µmol) and Msc-Phe(4-CH_2NH-Msc)-OSu (**6**, 254 mg, 429 µmol) were dissolved in 18 mL DMF. 4-Ethylmorpholine (307 µL, 2.42 mmol) was added to adjust to pH 8.5-9. The reaction mixture was stirred for 20 h at r.t., and then 4 mL H_2O were added. The mixture was concentrated under high vacuum at r.t. Addition of 80 mL cold diethyl ether resulted in the formation of white precipitate, which was collected by centrifugation and washed with cold diethyl ether (30 mL, 3X) followed by acetone (30 mL, 3X). The product was dried under high vacuum at r.t to yield 368 mg of the crude product as a white solid containing ~78% tetraMsc-AM-insulin (**4**) based on HPLC traces detected at 220 nm (~79% conversion **3** → **4**) (Figure S8). ESI-MS-TOF Calc. m/z 6437.77, Found: 6437.51 (deconvoluted) (Figure S9).

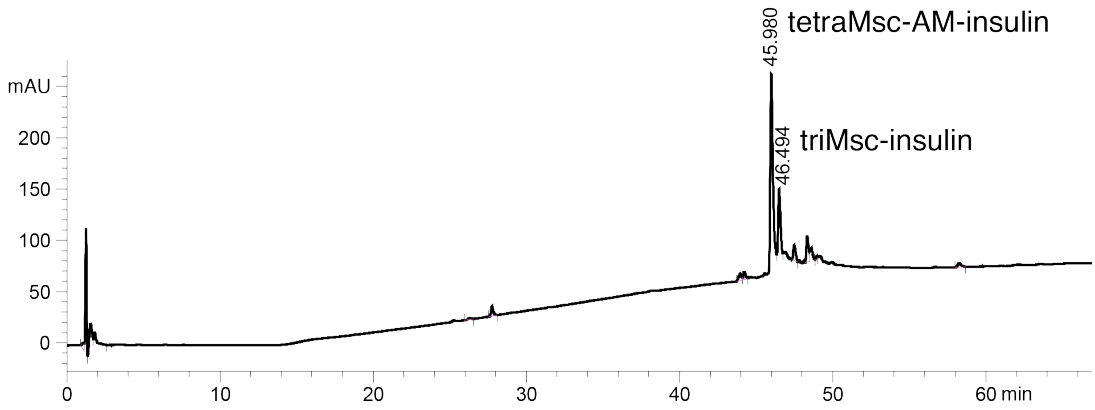


Figure S8. Analytical HPLC trace at 220 nm of the crude reaction mixture from the synthesis of tetraMsc-AM-insulin (**4**).

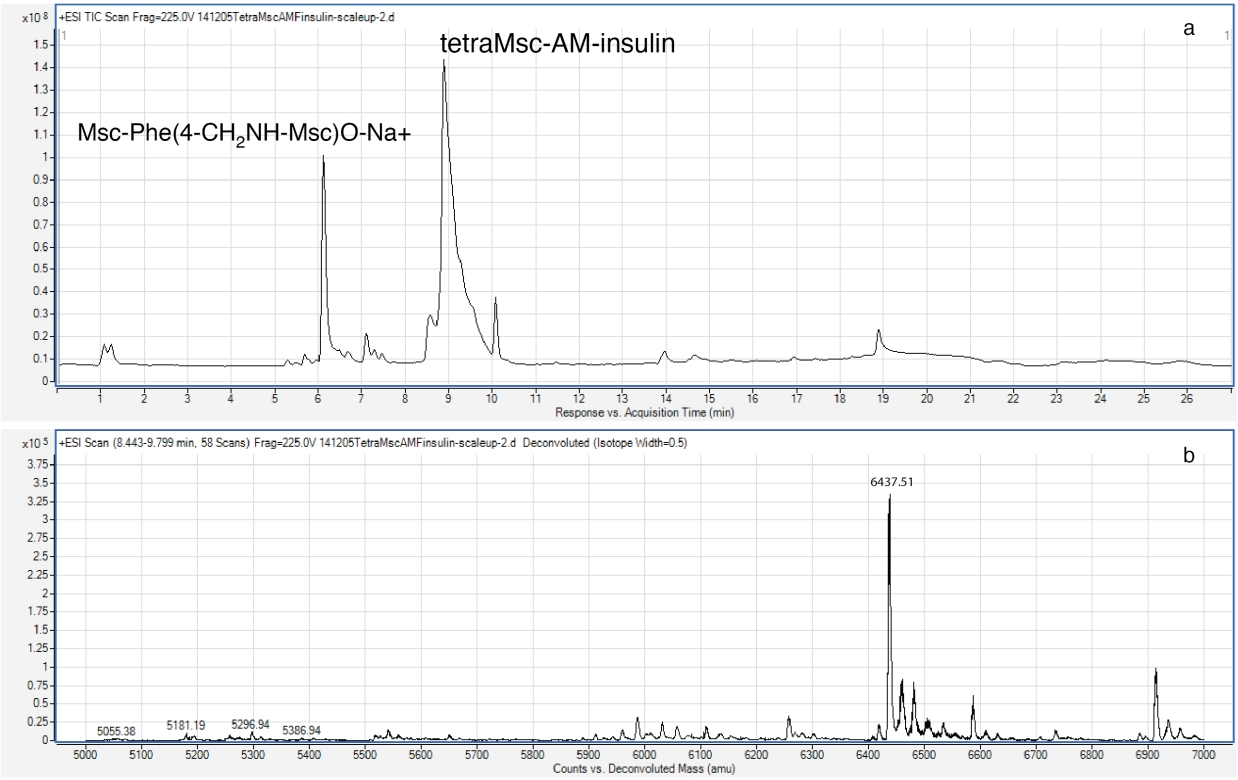


Figure S9. (top) Total ion chromatogram (TIC) of the crude reaction mixture from the synthesis of tetraMsc-AM-insulin (**4**). (bottom) Deconvoluted ESI-TOF mass spectrum of the TIC at 8.443-9.799 minutes. Calc. m/z 6437.77.

Synthesis of AM-insulin (5)

v. The crude mixture from (iv) (368 mg, ~78% tetraMsc-AM-insulin, **4**, 45 μmol) was dissolved in a solution of H_2O/methanol/1,4-dioxane (2:1:1, v/v/v, 18 mL). Aqueous NaOH (1.35 mL, 2 N) was added (resulting pH 10-11), and the mixture immediately turned light yellow. The reaction was carried out with gentle shaking for 3 h at r.t. The reaction mixture was neutralized by adding acetic acid, at which time the yellow color disappeared, and a white precipitate formed. The precipitate was separated by centrifugation and collected. The supernatant was concentrated under high vacuum at r.t. to 50% of its initial volume, at which point additional white precipitate formed. The precipitate was combined, washed with cold diethyl ether and acetone, and dried under high vacuum at r.t. to yield 300 mg of the crude product as a white solid containing 60% AM-insulin (**5**) based on HPLC traces detected at 220 nm (~70% conversion **4** → **5**) (Figure S10). ESI-MS-TOF Calc. m/z 5837.76, Found: 5837.62 (deconvoluted) (Figure S11).

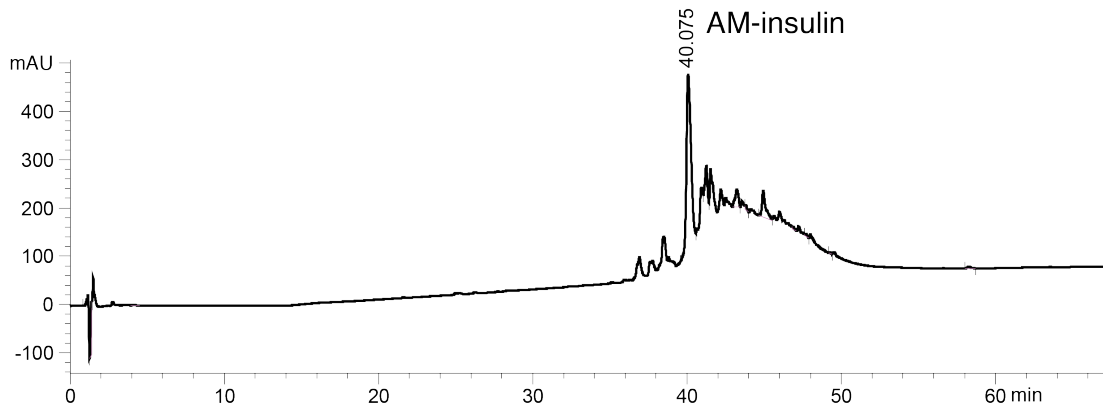


Figure S10. Analytical HPLC trace at 220 nm of the crude reaction mixture from the synthesis of AM-insulin (**5**).

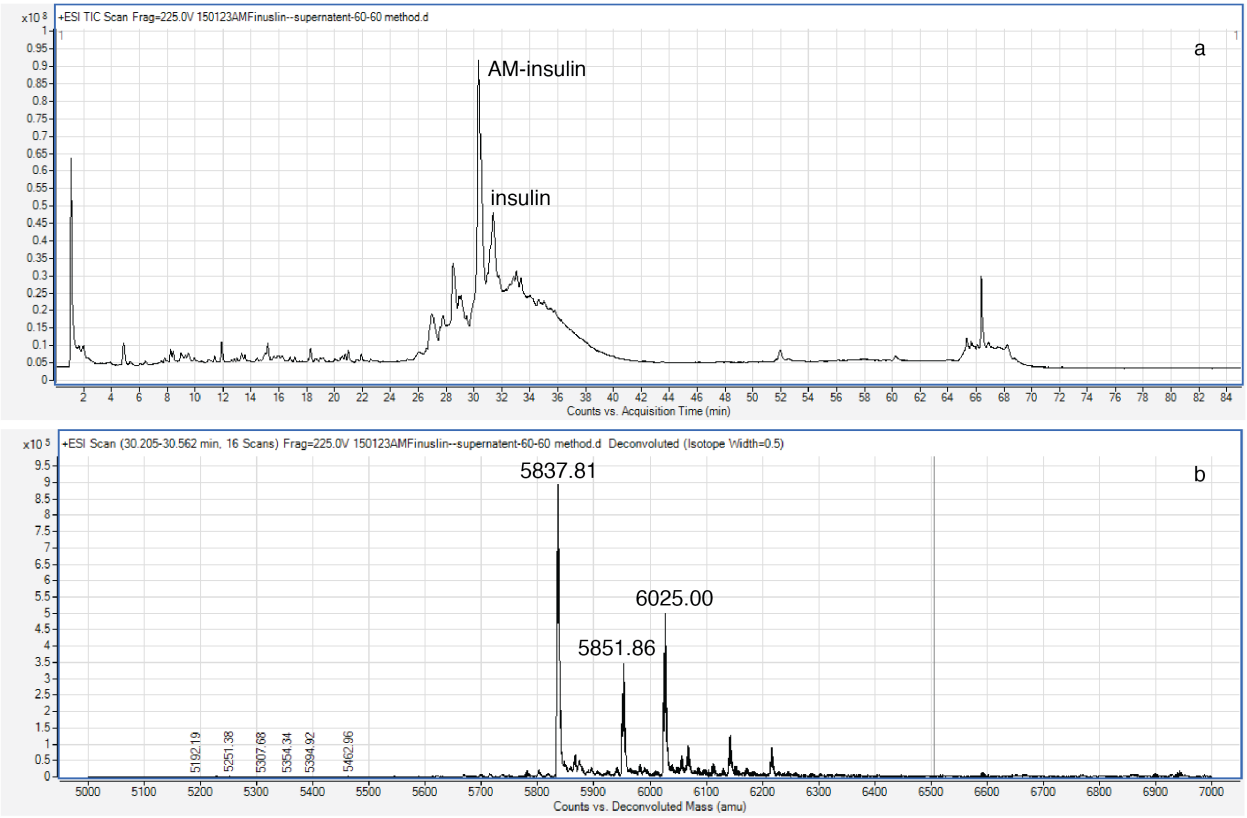


Figure S11. (top) Total ion chromatogram (TIC) of the crude reaction mixture from the synthesis of AM-insulin (**5**). (bottom) Deconvoluted ESI-TOF mass spectrum of the TIC at 30.205-30.562 minutes. Calc. m/z 5837.76.

Purification of AM-insulin (5)

The crude AM-insulin reaction mixture was dissolved in 0.1% aqueous TFA and separated on a reversed-phase semi-preparative C12 column (Jupiter 4µ Proteo 90A, 250×10 mm, Phenomenex, Part No: 00G-4396-NO) using an Agilent 1260 Infinity instrument in 0.1% aqueous TFA under a gradient of 0.5%/min of 0.1 % TFA in acetonitrile at a flow rate of 5 mL/min (Figure S12).

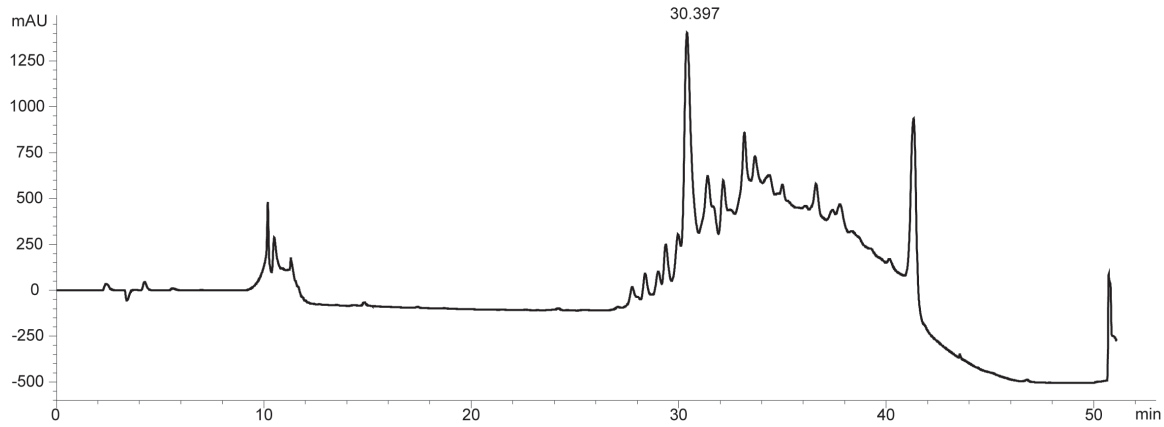


Figure S12. Representative Semi-Preparative HPLC trace at 220 nm of crude AM-insulin (**5**) during separation.

The purity of AM-insulin was checked by RP-HPLC on a C12 analytical column (Jupiter 4µ Proteo 90A, 50×2 mm, Phenomenex, Part No: 00G-4396-BO) using an Agilent 1260 Infinity instrument in 0.1% aqueous TFA under a gradient of 0.5%/min of 0.1% TFA in acetonitrile, at a flow rate of 1 mL/min) (Figure S13). Confirmation of the binding of AM-insulin with by LC-MS was performed using the same LC elution method (Figure S14).

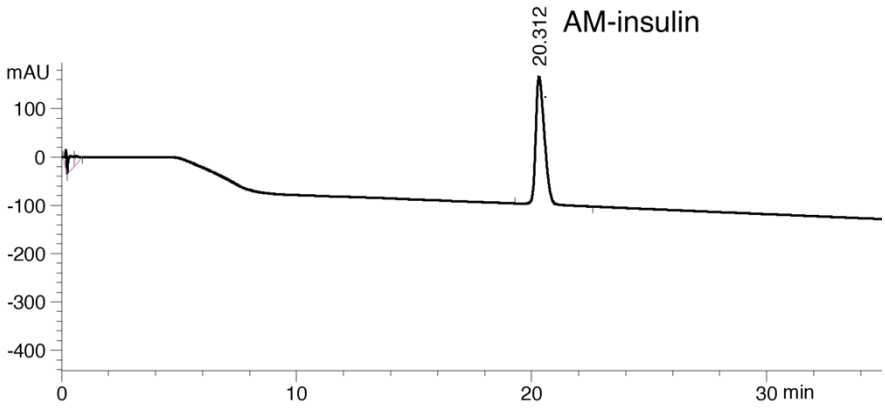


Figure S13. Example analytical HPLC trace at 220 nm of purified AM-insulin (**5**).

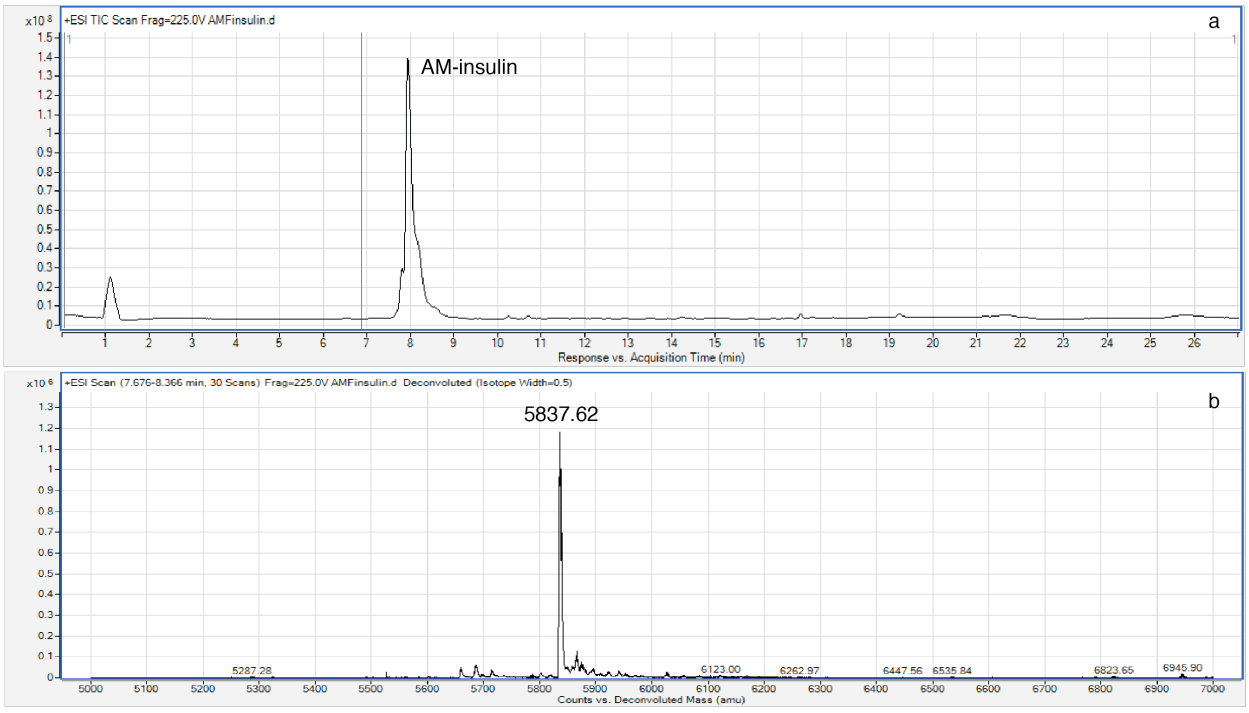


Figure S14. (top) Total ion chromatogram (TIC) of purified AM-insulin (**5**). (bottom) Deconvoluted ESI-TOF mass spectrum of the TIC peak at 7.990 minutes. Calc. m/z 5837.76.

Peptide Synthesis

Tetrapeptide synthesis was carried out using standard fmoc-scheme solid-phase methods using an automatic peptide synthesizer (CS Bio) on Rink amide MBHA resin. Each synthesis was carried out on the scale of 0.25 g resin (0.67 mmol/g). Fmoc deprotection was accomplished by shaking the resin in a solution of 20% piperidine in DMF at room temp for 20 min, followed by extensive rinsing with DMF. For each peptide coupling, a mixture of fmoc-protected amino acid, HBTU and DIEA (4 equiv. each) was dissolved in 10 mL DMF, mixed for 15 min, and added to the resin. Couplings proceeded at room temp for 60 min followed by extensive rinsing with DMF. After the final fmoc deprotection, the resin was transferred to a filter fitted with a medium porosity glass-frit and rinsed with dichloromethane. The peptide was cleaved from the resin by treatment with two portions of 10 mL of a mixture of 95% TFA, 2.5% TIS, and 2.5% water (v/v/v) for one hour each at room temperature with shaking. The filtrates from each deprotection were pooled and concentrated under high vacuum. The crude peptide was then dissolved in 30 mL 0.1% aqueous TFA and purified by reverse phase HPLC (Waters Delta 600 Semi-Prep system with a Waters Nova-Pak C18 column) using a 1%/min gradient of acetonitrile in 0.1% aqueous TFA. Fractions were checked by analytical HPLC (Agilent 1100 system with Waters Nova-Pak C18 column), and pure fractions were combined and lyophilized to dryness and confirmed by ESI-MS (Figures S15-18).

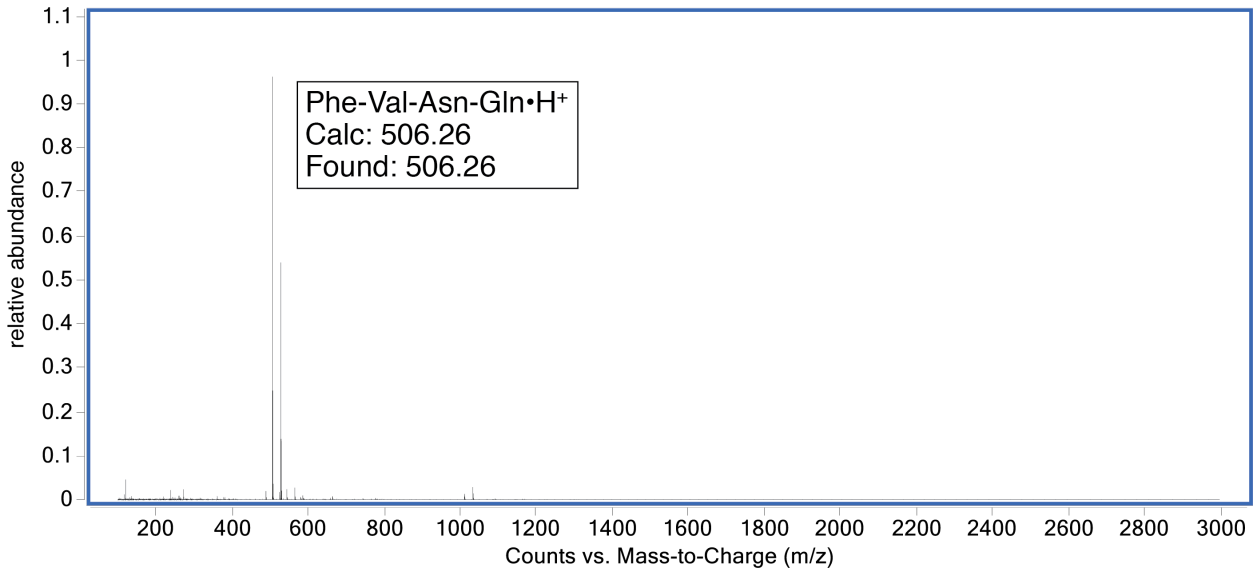


Figure S15. ESI-TOF Mass spectrum of purified H-Phe-Val-Asn-Gln-NH_2.

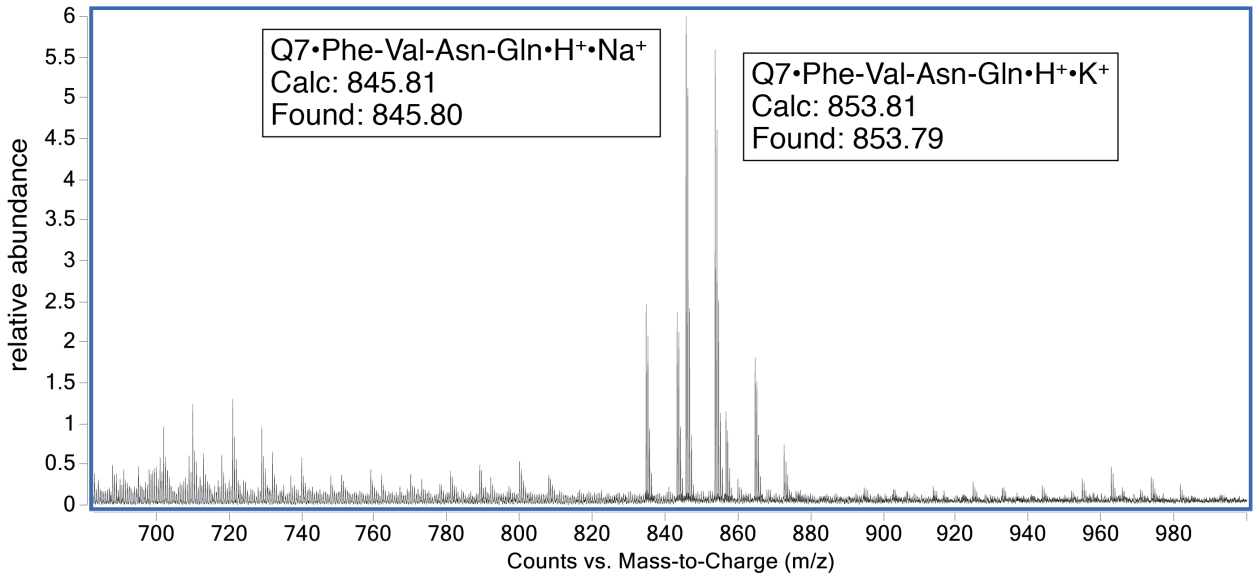


Figure S16. ESI-TOF mass spectrum of purified H-Phe-Val-Asn-Gln-NH_2 mixed with equimolar Q7.

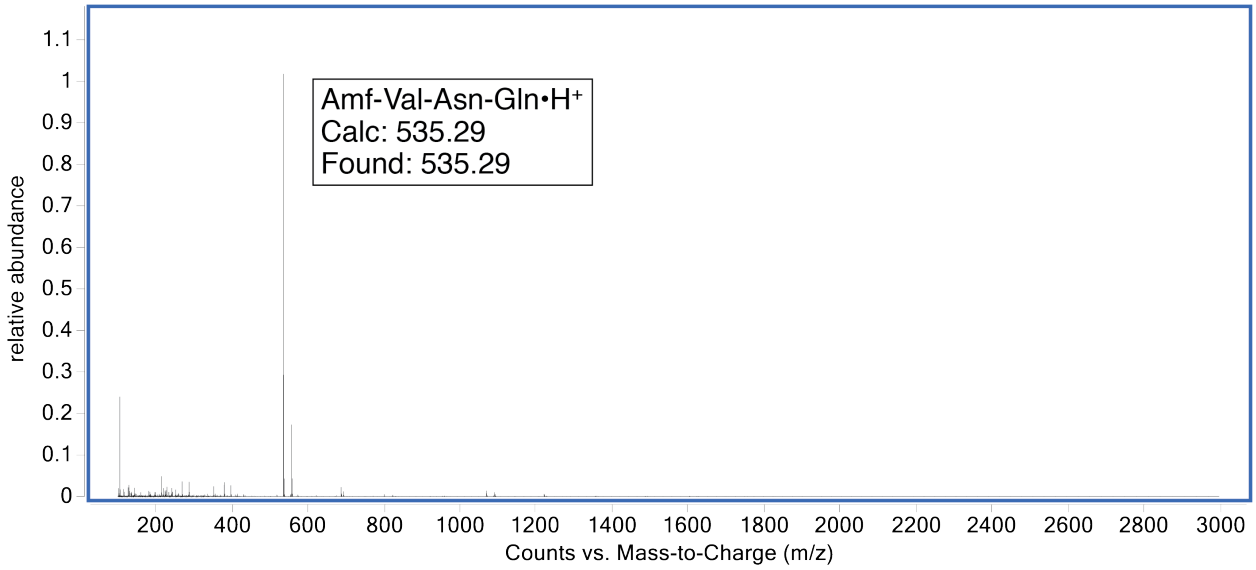


Figure S17. ESI-TOF Mass spectrum of purified H-Amf-Val-Asn-Gln-NH_2.

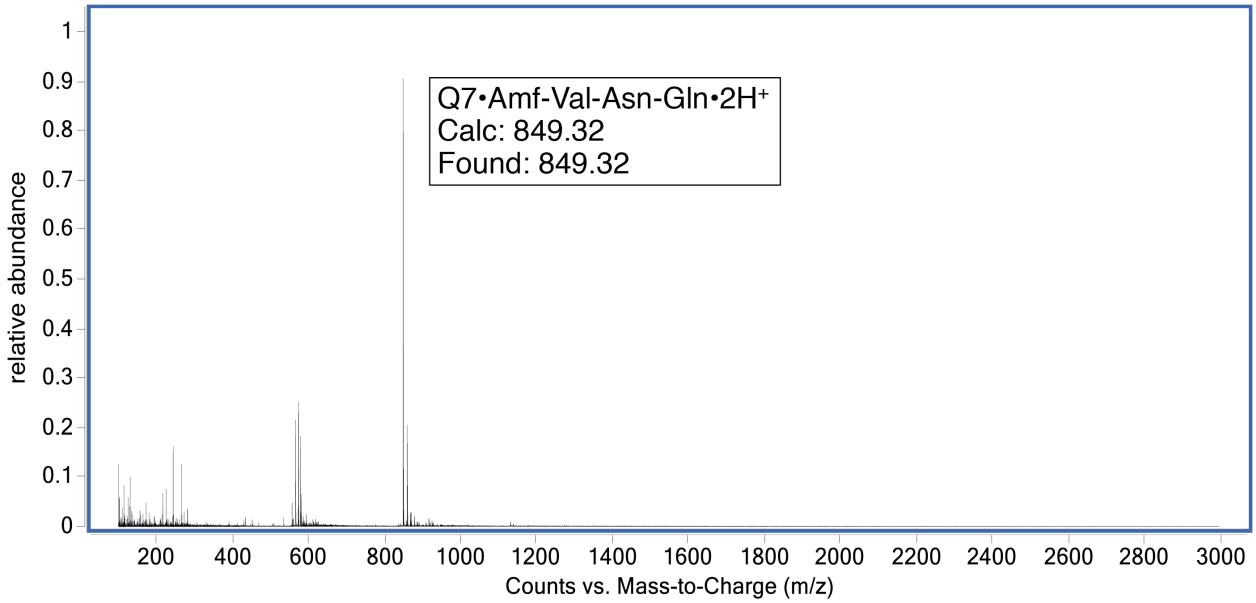


Figure S18. ESI-TOF mass spectrum of purified H-Amf-Val-Asn-Gln-NH_2 mixed with equimolar Q7.

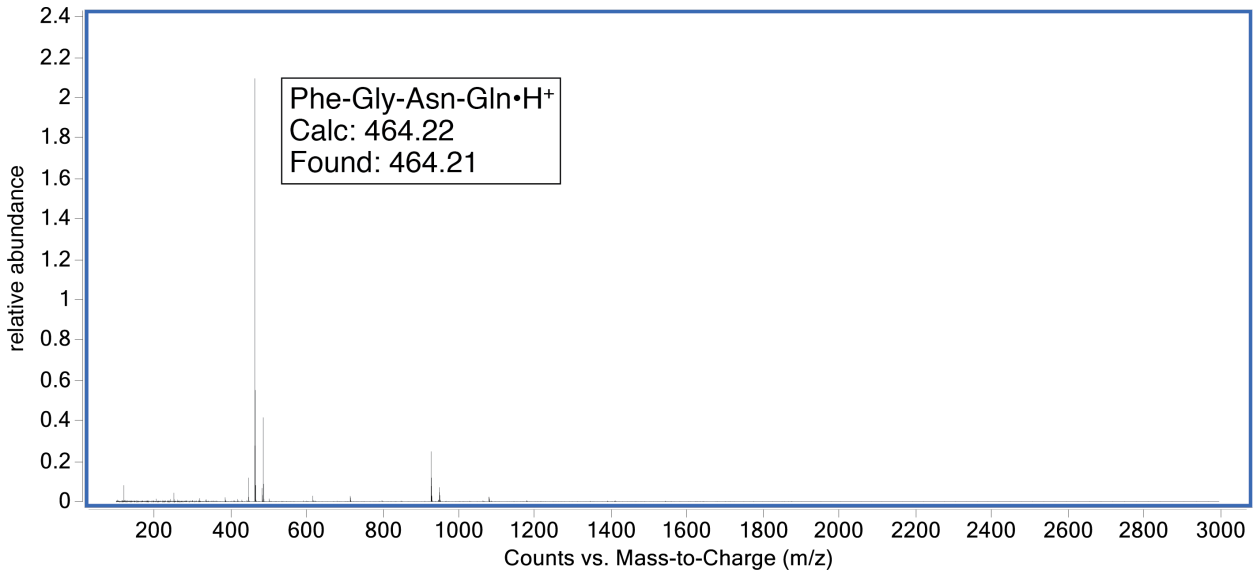


Figure S19. ESI-TOF Mass spectrum of purified H-Phe-Gly-Asn-Gln-NH_2.

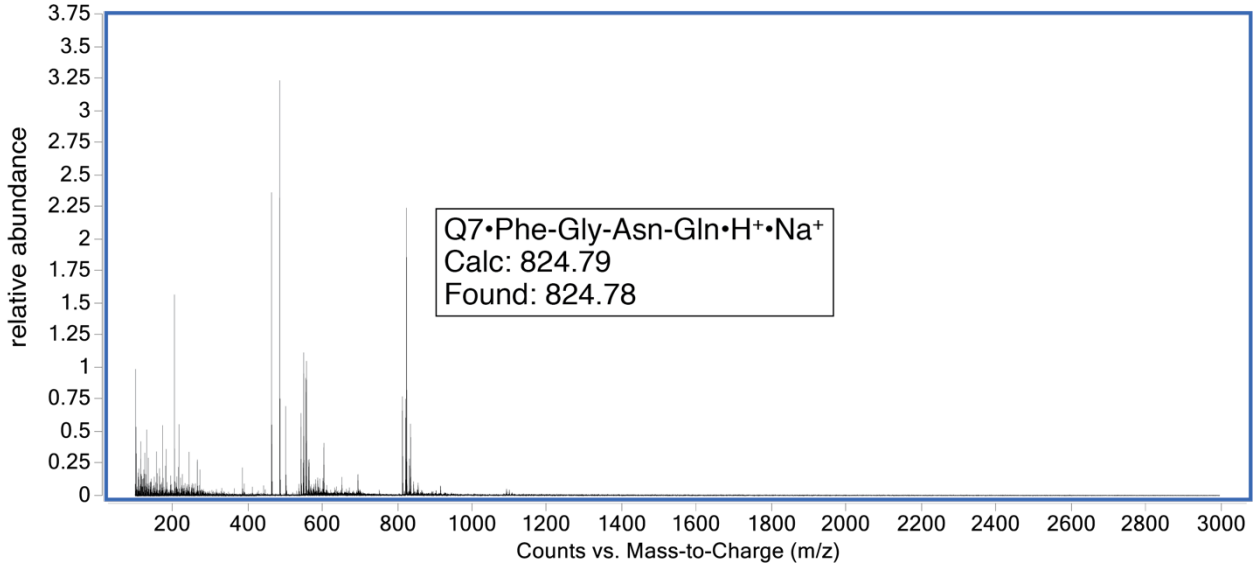


Figure S20. ESI-TOF mass spectrum of purified H-Phe-Gly-Asn-Gln-NH_2 mixed with equimolar Q7.

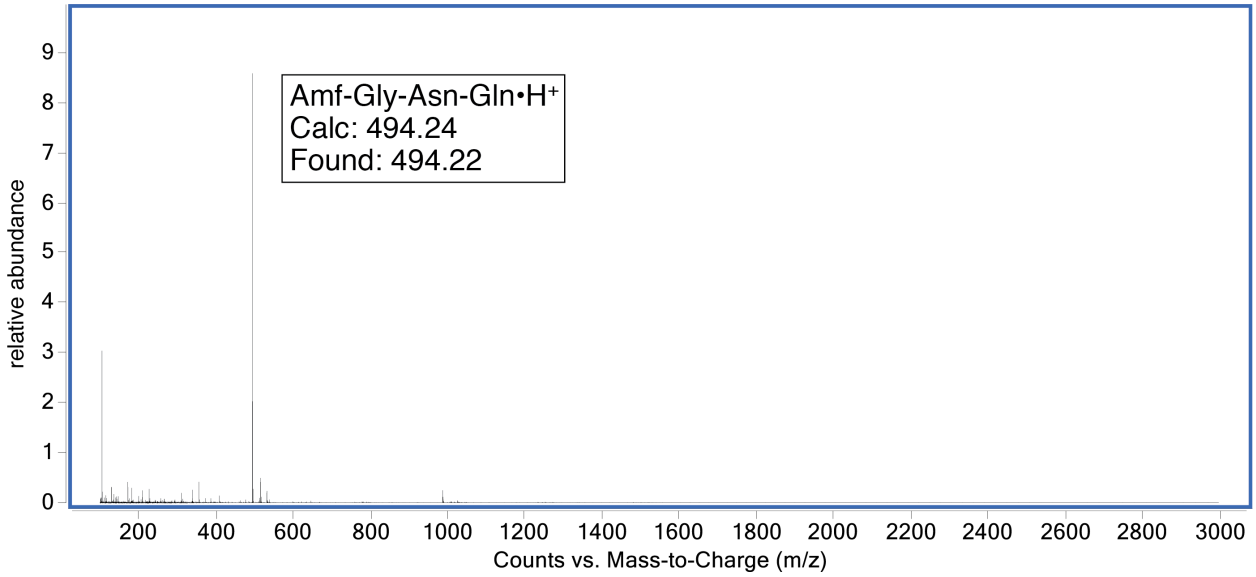


Figure S21. ESI-TOF Mass spectrum of purified H-Amf-Gly-Asn-Gln-NH_2.

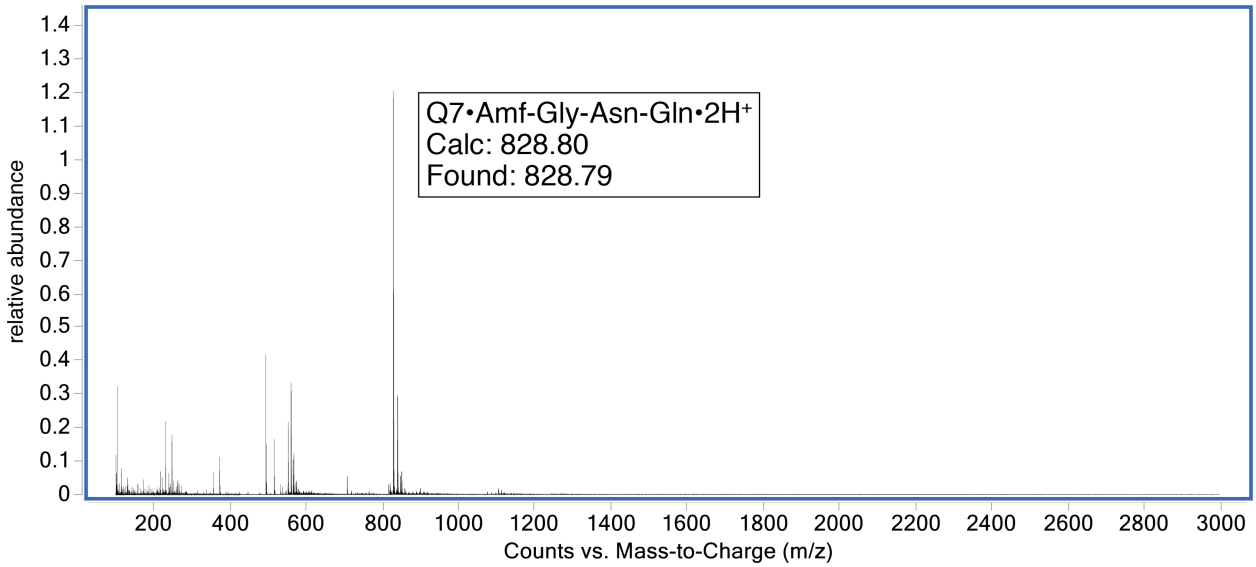


Figure S22. ESI-TOF mass spectrum of purified H-Amf-Gly-Asn-Gln-NH_2 mixed with equimolar Q7.

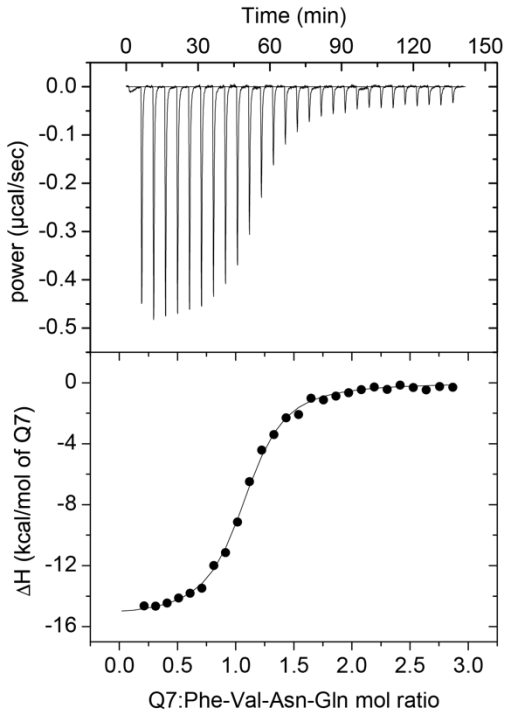


Figure S23. Representative ITC data of the titration of Phe-Val-Asn-Gln with Q7 at 300 K in 10 mM sodium phosphate, pH 7.0.

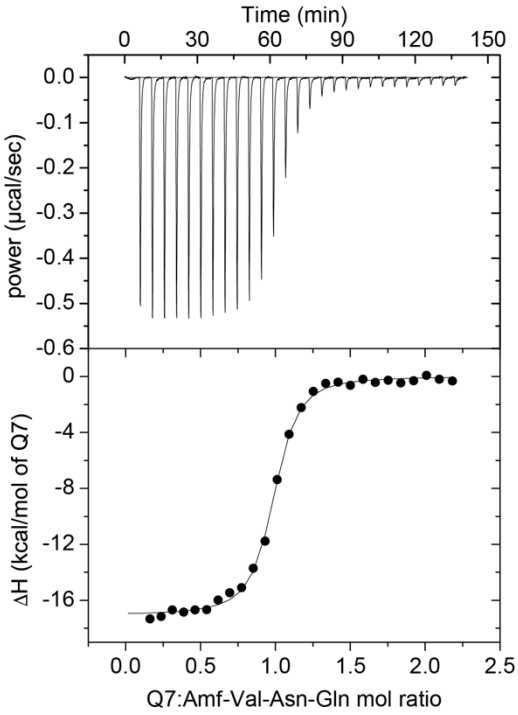


Figure S24. Representative ITC data of the titration of Amf-Val-Asn-Gln with Q7 at 300 K in 10 mM sodium phosphate, pH 7.0.

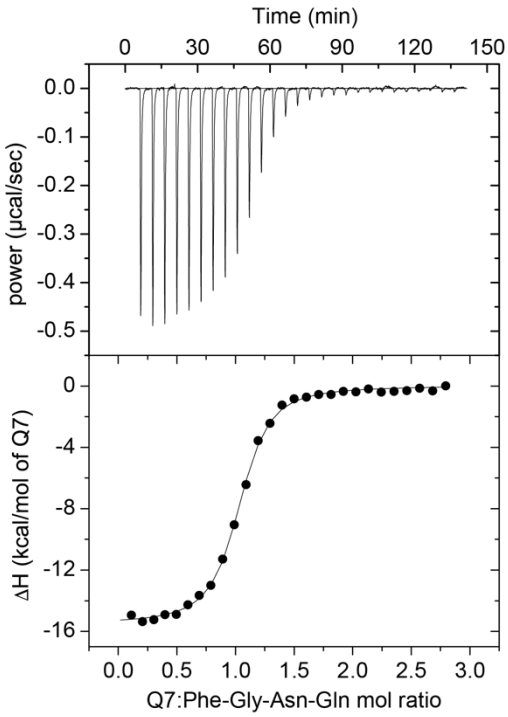


Figure S25. Representative ITC data of the titration of Phe-Gly-Asn-Gln with Q7 at 300 K in 10 mM sodium phosphate, pH 7.0.

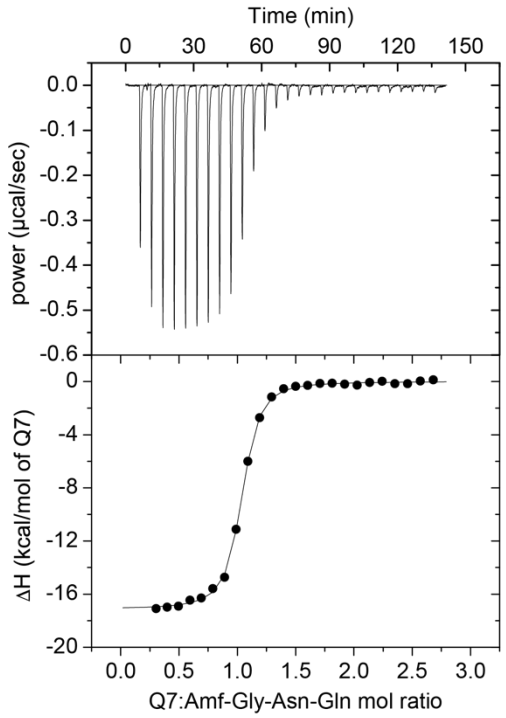


Figure S26. Representative ITC data of the titration of Amf-Gly-Asn-Gln with Q7 at 300 K in 10 mM sodium phosphate, pH 7.0.

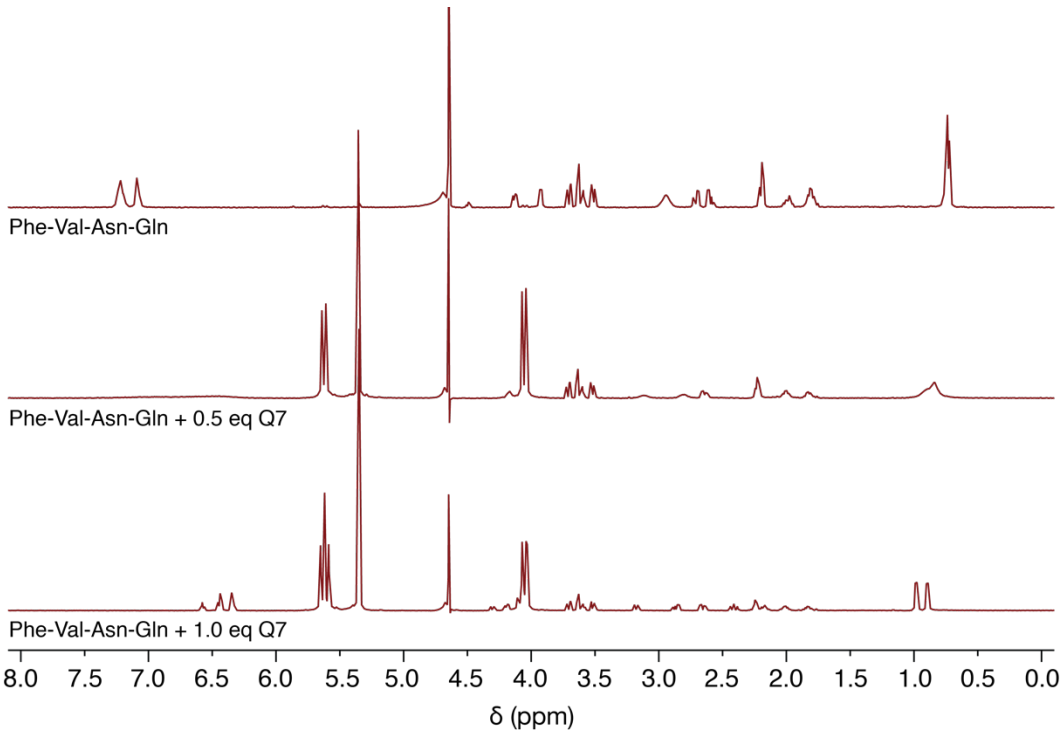


Figure S27. Overlay of ^{1}H NMR spectra of Phe-Val-Asn-Gln with 0 (top), 0.5 (middle), and 1.0 (bottom) equivalents of Q7 at 25°C in D_2O.

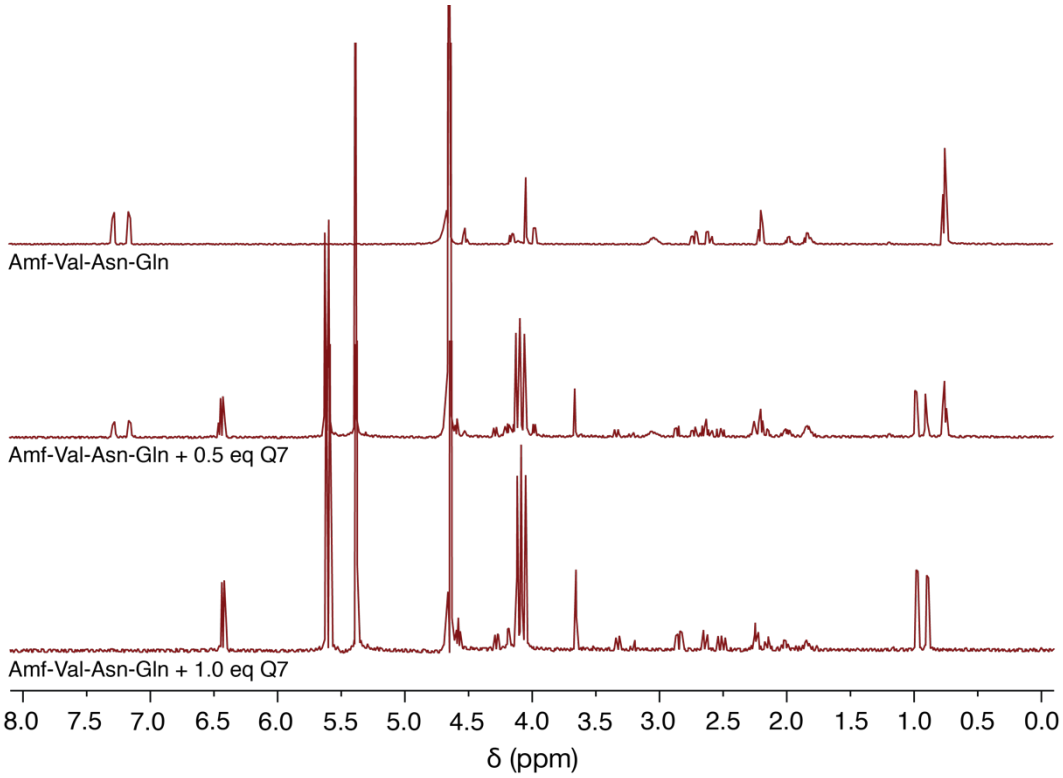


Figure S28. Overlay of ^{1}H NMR spectra of Amf-Val-Asn-Gln with 0 (top), 0.5 (middle), and 1.0 (bottom) equivalents of Q7 at 25°C in D_2O.

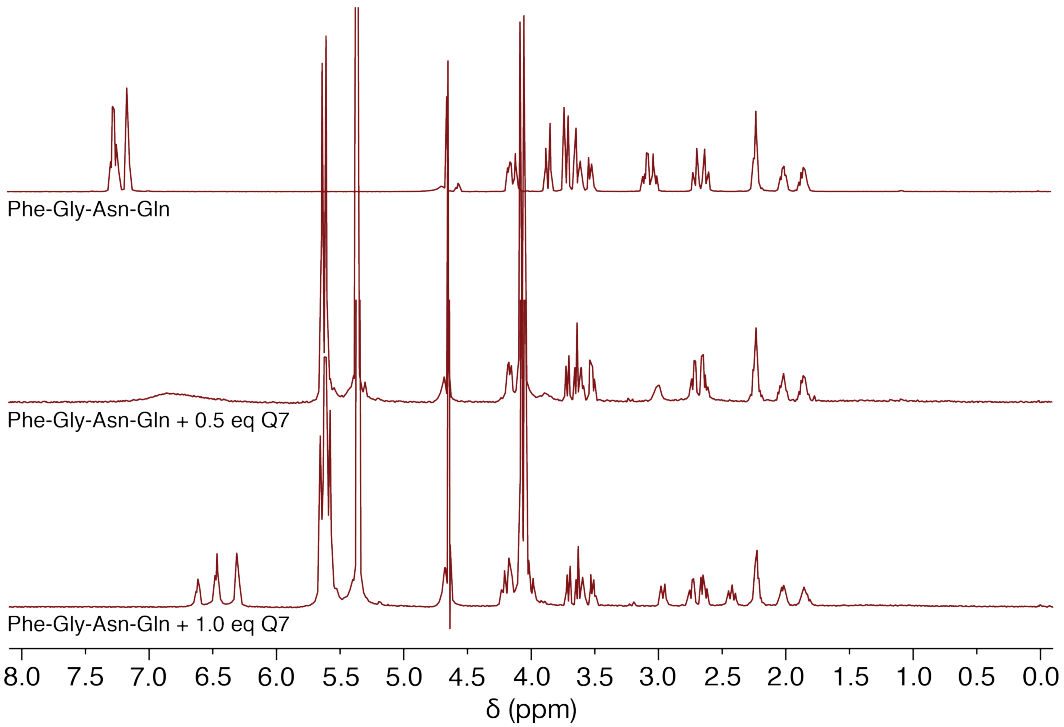


Figure S29. Overlay of ^{1}H NMR spectra of Phe-Gly-Asn-Gln with 0 (top), 0.5 (middle), and 1.0 (bottom) equivalents of Q7 at 25 °C in D_2O.

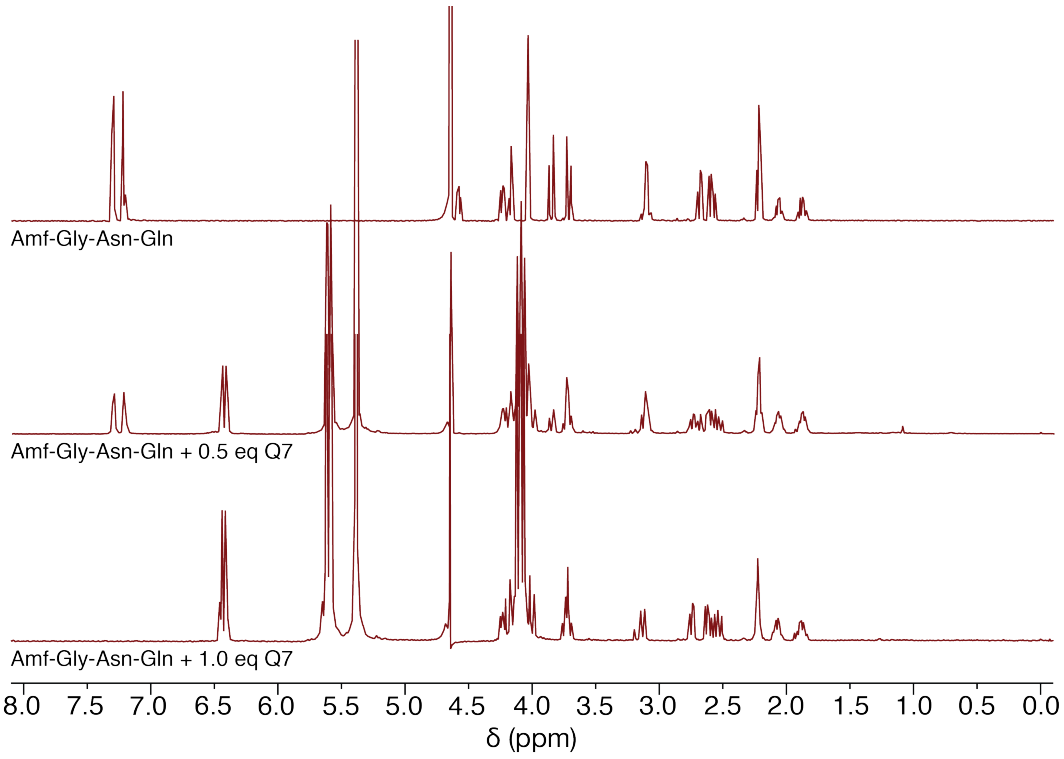


Figure S30. Overlay of ^{1}H NMR spectra of Amf-Gly-Asn-Gln with 0 (top), 0.5 (middle), and 1.0 (bottom) equivalents of Q7 at 25°C in D_2O.

References.

1. Day, A. I., Arnold, A. P., Blanch, R. J., Snushall, B. (2001) Controlling factors in the synthesis of cucurbituril and its homologues. J. Org. Chem. 66, 8094-8100.
2. Tesser, G. I., and Balvert-Geers, I. C. (1975) The methylsulfonylethyloxycarbonyl group, a new and versatile amino protective group. Int. J. Peptide Protein Res. 7, 295-305.
3. Offord, R. E. (1980) Semisynthetic Proteins. John Wiley & Sons Ltd., New York.
4. Davies, J. G. and Wallace, C. J. A. (2000) Protein Engineering by Semisynthesis (Wallace, C. J. A., Ed.) CRC Press LLC., Boca Raton, Florida.
5. Geiger, R., Obermeier, R., and Tesser, G. I. (1975) Der Methylfulfonyläthyloxycarbonyl-Rest als reversible Aminoschutzgruppe für insulin. Chem. Ber. 108, 2758-2763.