C-terminal modification of fully unprotected peptide hydrazides via in situ generation of isocyanates

Alexander A. Vinogradov, Mark D. Simon, and Bradley L. Pentelute

Department of Chemistry, Massachusetts Institute of Technology, 18-596, 77 Massachusetts Avenue, Cambridge, Massachusetts, 02139, United States.

Supporting Information

Contents

1. Materials

2-(1H-Benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), 2-(7-Aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) and N^{α}-Fmoc protected D-amino acids were purchased from Chem-Impex International. N^{α}-Fmoc protected L-amino acids were obtained through Advanced ChemTech, and Novabiochem.

2-chlorotrityl chloride resin (200-400mesh, 1.2 mmol/g) was purchased from Chem-Impex International and was subsequently used to prepare 2-chlorotrityl hydrazine resin adhering to the protocol reported by Stavropoulos, George et al. *Letters in Peptide Science*. 1995, 2, 315-318. The loading of the resin was determined to be 0.75 mmol/g.

N,N-Dimethylformamide (DMF), dichloromethane (DCM), diethyl ether, and HPLC-grade acetonitrile were from EMD Millipore. Triisoproyl silane (TIPS), and 1,2-Ethanedithiol (EDT) were from Alfa Aesar. Solvents for HPLC-MS were purchased from and Fluka. All other reagents were purchased from Sigma-Aldrich. Ni-NTA Agarose beads were from Qiagen.

All reagents were used as received.

2. Peptide synthesis

2.1. Fmoc Solid Phase Peptide Synthesis

Peptides were synthesized on an enhanced version of our recently reported automated flow-based platform using a 2nd generation vessel (Simon, Heider et al. *ChemBioChem*, **2014**, 15, 713-720). During synthesis all solvents and reagents were preheated to 60°C immediately before reaching the synthesis vessel. The following standard 94 second cycle was used to assemble all peptides, unless noted:

- Amide bond formation (coupling) – 14 seconds at 23mL/min
- Removal of coupling reagent (wash) – 30 seconds at 20mL/min
- N^{α}-Fmoc removal (deprotection) – 20 seconds at 20mL/min
- Removal of deprotection reagent and products (wash) – 30 seconds at 20mL/min

*Coupling. C*oupling was performed by delivering the following coupling solution at 20 mL/min for 14 s. The coupling solution consisted of of N^{α}-Fmoc and side chain protected amino

acid dissolved in equimolar 0.4 M HBTU in DMF. To activate the amino acid, DIEA was injected at 4ml/min. The mixed fluid then passed through a static mixer and preheat loop to the synthesis vessel. In order to minimize racemization of Cys and His residues, only 2ml/min of DIEA were injected when activating these amino acids. In all cases, this coupling solution contained at least four equivalents of activated amino acid with respect to the resin.

To bond C-terminal Val and Ile residues to the resin, 1 mmol of protected amino acid dissolved in 2.5 mL of 0.4 M HBTU in DMF was manually delivered to the synthesis vessel, then synthesis was paused for 10 minutes. An increased reaction time was necessary to quantitatively couple these sterically hindered residues to the resin.

Side chain protection was as follows: Arg(Pbf), Asn(Trt), Asp(OtBu), Cys(Trt), Gln(Trt), Glu(OtBu), His(Trt), Lys(Boc), Ser(tBu), Thr(tBu), Trp(Boc), Tyr(tBu).

Wash. Excess reagent and reaction by-products were washed out from the synthesis vessel with 10 mL of DMF delivered at 20 mL/min over 30 s.

Deprotection. N^{α}-Fmoc protecting groups were removed with 6.6 mL of 20% (v/v) piperidine in DMF delivered at 20 mL/min over 20 seconds.

Three peptides: H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$, H_2N-Ala-Lys-Val-Ile-Asn-Thr-Phe-Asp-Gly-Ile-$CONHNH_2$, and H_2N-Ala-His-Val-Ile-Asn-Thr-Phe-Asp-Gly-Thr-$CONHNH_2$ were synthesized on the manual version of the flow system, described in the same manuscript. The synthesis procedure was similar, except 0.38 M HATU was used as an activator, and the following three minute synthetic cycle was used:

- Amide bond formation (coupling) – 30 seconds at 6mL/min
- Removal of coupling reagent (wash) – 60 seconds at 20mL/min
- N^{α}-Fmoc removal (deprotection) – 20 seconds at 20mL/min
- Removal of deprotection reagent and products (wash) – 60 seconds at 20mL/min

Peptide H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$ was additionally synthesized on a larger scale using standard batch Fmoc SPPS procedures: 2.5g of Chlorotrityl hydrazide resin (0.25mmol/g) was pre-swollen in DCM for 30 minutes, then drained. To bond C-terminal Val to the resin, 5 mmol of Fmoc protected valine was dissolved in 12.5 mL of 0.38 M HATU in DMF and 1.5mL of DIEA, then added to the swollen resin. After 1.5 hours, the resin

was drained and a second 5 mmol aliquot of Fmoc-Val-OH, prepared as above, was added. After a second 1.5 hour coupling, the resin was drained and washed three times with 30mL of DMF each. Fmoc was removed by treatment with 20mL of 20% piperidine in DMF for 20 minutes, and the resin was drained and washed with DMF three times (30mL each). Synthesis then proceeded according to the following procedure: 5mmol of each protected amino acid was dissolved in 12.5 mL of 0.38 M HATU in DMF and 1.5mL of DIEA, then added to the resin. After a 30 minute coupling step, the resin was drained and washed with DMF three times (30mL each). Fmoc was removed by treatment with 30mL of 20% piperidine in DMF for 20 minutes, and the resin was again drained and washed as above. After the final Fmoc removal, the resin was washed four times with 40mL of DCM each, and dried *in vacuo*.

All other peptides mentioned herein were synthesized and purified as described previously in Mong, Vinogradov et al, *ChemBioChem*, **2014**, 15, 721-733.

2.2. Cleavage of peptides from resin

All peptides were cleaved from the resin and side chain deprotected with a standard cleavage cocktail of 2.5% (v/v) EDT, 2.5% (v/v) H_2O, and 1% (v/v) TIPS in TFA for 7 minutes at 60 °C. In all cases, compressed nitrogen was used to evaporate the cleavage solution to dryness together with resin. The resulting mixtures were triturated and washed three times with cold diethyl ether, and, unless noted, dissolved in 95% water / 5% acetonitrile (0.1% TFA added). Resin was filtered and solutions were purified without intermediate lyophilization.

2.3. Purification of Peptides

All peptides were purified on a Waters 600 HPLC system with a Waters 484 or 486 UV detector using water with 0.1% TFA added (solvent A) and acetonitrile with 0.1% TFA added (solvent B) as solvents.

An Agilent Zorbax 300SB preparative C_3 column (300Å, 7μm, 21.2 x 250mm) was used at 20 mL/min linear flow rate for the purification of peptide hydrazides. Unless noted, the following gradient was used to purify peptide hydrazides: 5% B in A for 5 minutes, then 5%-35% B ramping linearly over 90 minutes.

An Agilent Zorbax 300SB semi-preparative C3 column (300 Å, 5 μm, 9.4 x 250 mm) was used at 5 mL/min linear flow rate for the purification of isocyanate conjugation reactions. Unless noted, the following gradient was used to purify isocyanate conjugation reactions: 5% B in A for 10 minutes; then 5%-15% B ramping linearly over 10 minutes; then 15%-45% B ramping linearly over 90 minutes.

Fractions were collected and screened for the desired material using HPLC-ESI-TOF.

3. HPLC-MS Analysis

Unless noted, all peptides, proteins and reaction mixtures were analyzed on an Agilent 6520 Accurate Mass Q-TOF LC-MS using an Agilent Zorbax 300SB C_3 column (300 Å, 5 μm, 2.1 x 150 mm) run with the following method. At 40 °C and a flow rate of 0.8 mL/min, the following gradient was used: 1% acetonitrile with 0.1% formic acid added (FA, solvent B') in water with 0.1% FA (solvent A') for 2 min, 1-61% B' in A' ramping linearly over 11 min, 61% B' in A' for 1 minute. Typically, 100 – 1000 μg/mL solutions of peptides and proteins were subject to analysis. More concentrated solutions were diluted to appropriate concentrations with 50% A/50% B. Unless noted, all chromatograms shown in this work are plots of total ion current (TIC) versus time. HPLC yields of reactions were calculated via manual integration of the area under peaks corresponding to reaction products on TIC vs. time chromatograms. Generally, TIC and UV_{214} integration gave converging results, consistent with previous reports (D. Cohen, C. Zhang, *et al. JACS*, **2015**, 137, 9784-9787); TIC integration was a preferred method of characterization for practical reasons. The data were analyzed using Agilent MassHunter Qualitative analysis software. All MS deconvolution spectra were obtained using the maximum entropy algorithm.

4. Peptide synthesis: analytical data

4.1. H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$

The peptide was synthesized and purified as described above (Section 2) on 180 mg of resin. The crude peptide was lyophilized to give 105 mg of crude peptide. 26.2 mg of pure peptide was obtained after purification.

The batch synthesis of this peptide additionally yielded 140 mg of pure compound after lyophilization.

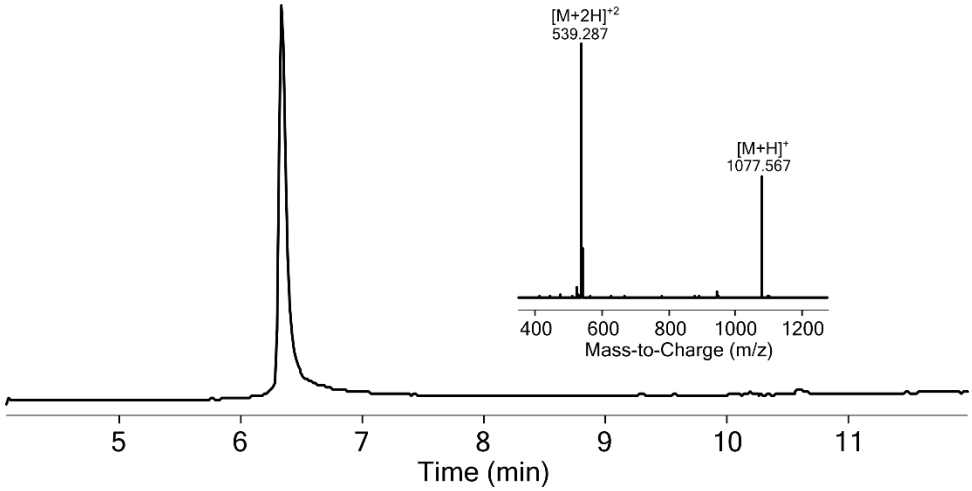


Figure S1. HPLC-MS (TIC) chromatogram for the purified peptide H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$ with MS inset on the right. Calc. monoisotopic mass = 1076.6 Da.

4.2. H_2N-Ala-Lys-Val-Ile-Asn-Thr-Phe-Asp-Gly-Ile-$CONHNH_2$

The peptide was synthesized and purified as described above (Section 2) on 190 mg of resin. Purification of crude peptide yielded 16.5 mg of pure peptide after lyophilization.

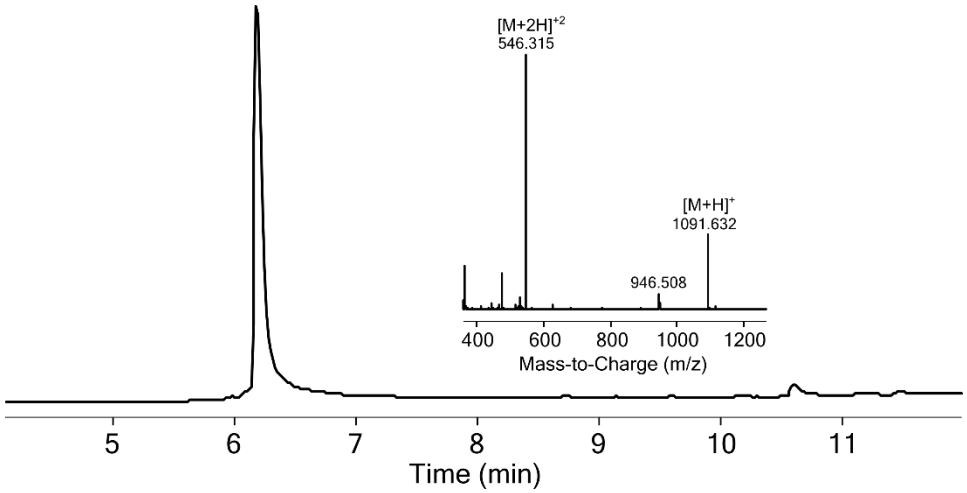


Figure S2. HPLC-MS (TIC) chromatogram for the purified peptide H_2N-Ala-Lys-Val-Ile-Asn-Thr-Phe-Asp-Gly-Ile-$CONHNH_2$ with MS inset on the right. Calc. monoisotopic mass = 1090.6 Da.

4.3. H_2N-Ala-His-Val-Ile-Asn-Thr-Phe-Asp-Gly-Thr-$CONHNH_2$

The peptide was synthesized and purified as described above (Section 2) on 180 mg of resin. Purification of crude peptide yielded 16.2 mg of pure peptide after lyophilization.

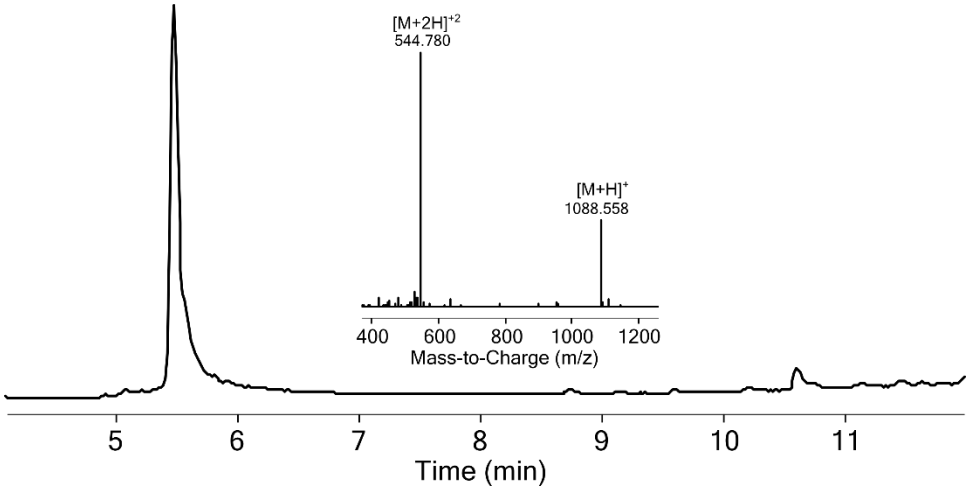


Figure S3. HPLC-MS (TIC) chromatogram for the purified peptide H_2N-Ala-His-Val-Ile-Asn-Thr-Phe-Asp-Gly-Thr-$CONHNH_2$ with MS inset on the right. Calc. monoisotopic mass = 1087.54 Da.

4.4. H_2N-Ala-Lys-Val-Ile-Asn-Thr-Phe-Asp-Gly-Trp-$CONHNH_2$

The peptide was synthesized and purified as described above (Section 2) on 239 mg of resin. Purification of crude peptide yielded 4.5 mg of pure peptide after lyophilization.

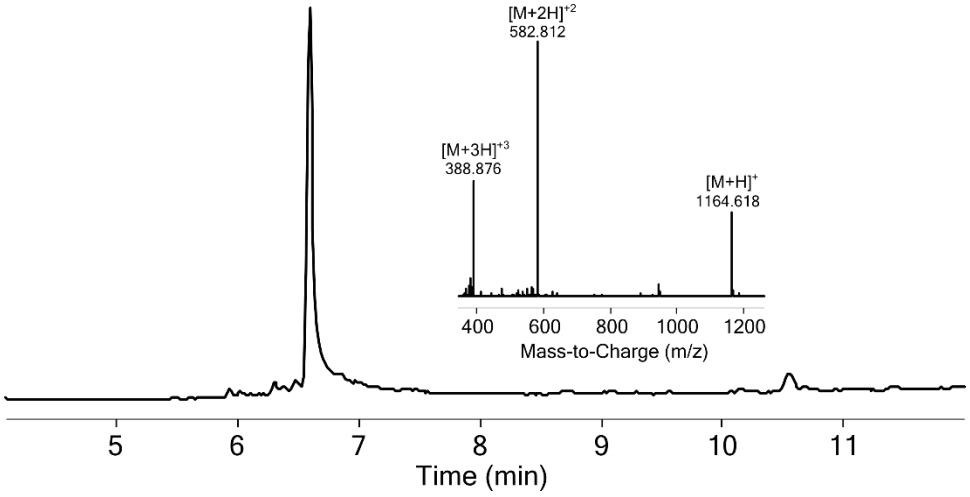


Figure S4. HPLC-MS (TIC) chromatogram for the purified peptide H_2N-Ala-Lys-Val-Ile-Asn-Thr-Phe-Asp-Gly-Trp-$CONHNH_2$ with MS inset on the right. Calc. monoisotopic mass = 1163.6 Da.

4.5. H_2N-Ala-Trp-Val-Ile-Asn-Thr-Phe-Asp-Gly-Arg-$CONHNH_2$

The peptide was synthesized and purified as described above (Section 2) on 232 mg of resin. Purification of crude peptide yielded 3.8 mg of pure peptide after lyophilization.

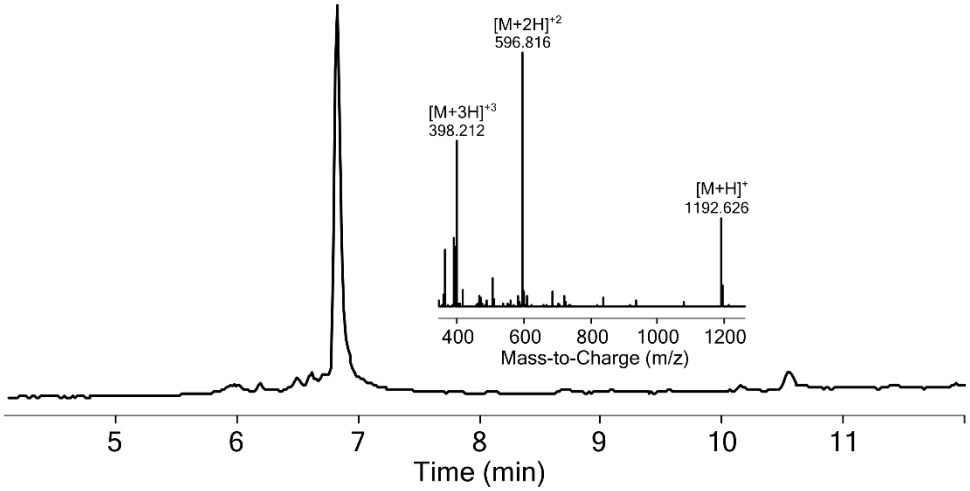


Figure S5. HPLC-MS (TIC) chromatogram for the purified peptide H_2N-Ala-Trp-Val-Ile-Asn-Thr-Phe-Asp-Gly-Arg-$CONHNH_2$ with MS inset on the right. Calc. monoisotopic mass = 1091.6 Da.

4.6. H_2N-Ala-Tyr-Val-Ile-Asn-Thr-Phe-Asp-Gly-Lys-$CONHNH_2$

The peptide was synthesized and purified as described above (Section 2) on 370 mg of resin. Purification of crude peptide yielded 18.5 mg of pure peptide after lyophilization.

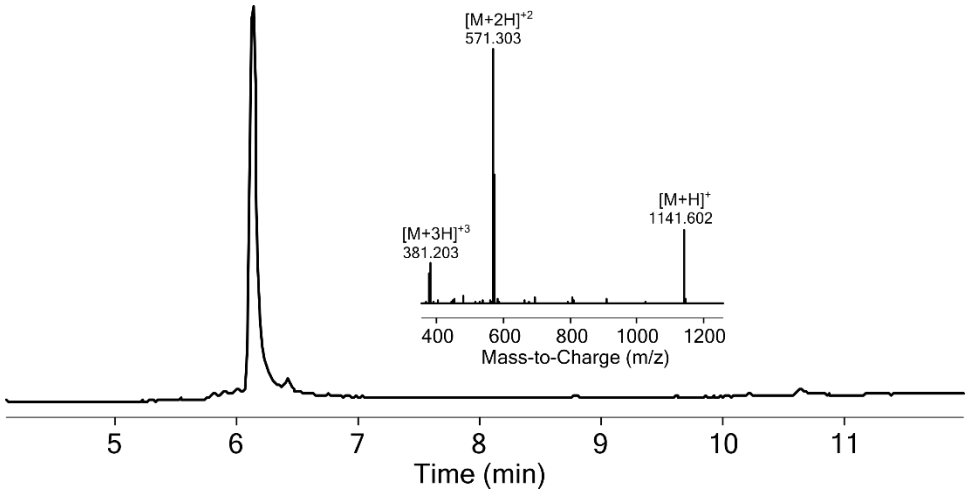


Figure S6. HPLC-MS (TIC) chromatogram for the purified peptide H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$ with MS inset on the right. Calc. monoisotopic mass = 1076.6 Da.

4.7. H_2N-Ala-Asp-Val-Ile-Asn-Thr-Phe-Asp-Gly-Leu-$CONHNH_2$

The peptide was synthesized and purified as described above (Section 2) on 225 mg of resin. Purification of crude peptide yielded 4.4 mg of pure peptide after lyophilization.

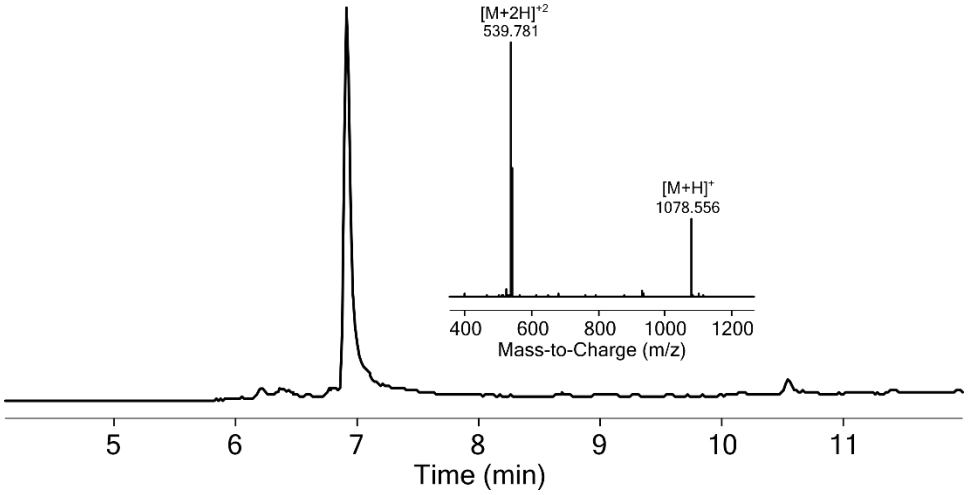


Figure S7. HPLC-MS (TIC) chromatogram for the purified peptide H_2N-Ala-Asp-Val-Ile-Asn-Thr-Phe-Asp-Gly-Leu-$CONHNH_2$ with MS inset on the right. Calc. monoisotopic mass = 1077.54 Da.

4.8. H_2N-Ala-Gly-Val-Ile-Asn-Thr-Phe-Asp-Gly-Met-$CONHNH_2$

The peptide was synthesized and purified as described above (Section 2) on 386 mg of resin. Purification of crude peptide yielded 9.1 mg of pure peptide after lyophilization.

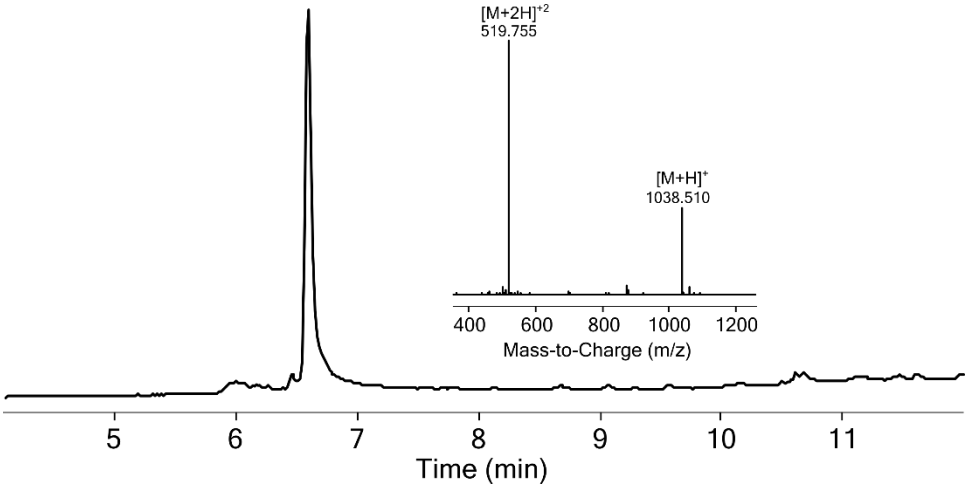


Figure S8. HPLC-MS (TIC) chromatogram for the purified peptide H_2N-Ala-Gly-Val-Ile-Asn-Thr-Phe-Asp-Gly-Met-$CONHNH_2$ with MS inset on the right. Calc. monoisotopic mass = 1037.5 Da.

4.9. H_2N-Ala-Glu-Val-Ile-Asn-Thr-Phe-Asp-Gly-Ala-$CONHNH_2$

The peptide was synthesized and purified as described above (Section 2) on 182 mg of resin. Purification of crude peptide yielded 20.0 mg of pure peptide after lyophilization.

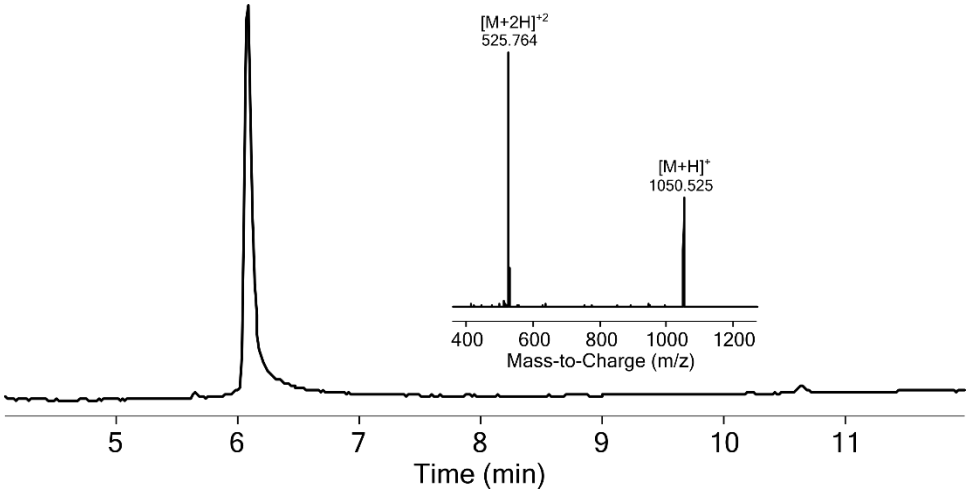


Figure S9. HPLC-MS (TIC) chromatogram for the purified peptide H_2N-Ala-Glu-Val-Ile-Asn-Thr-Phe-Asp-Gly-Ala-$CONHNH_2$ with MS inset on the right. Calc. monoisotopic mass = 1049.5 Da.

4.10. H_2N-Ala-Val-Val-Ile-Asn-Thr-Phe-Asp-Gly-His-$CONHNH_2$

The peptide was synthesized and purified as described above (Section 2) on 260 mg of resin. Purification of crude peptide yielded 5.5 mg of pure peptide after lyophilization.

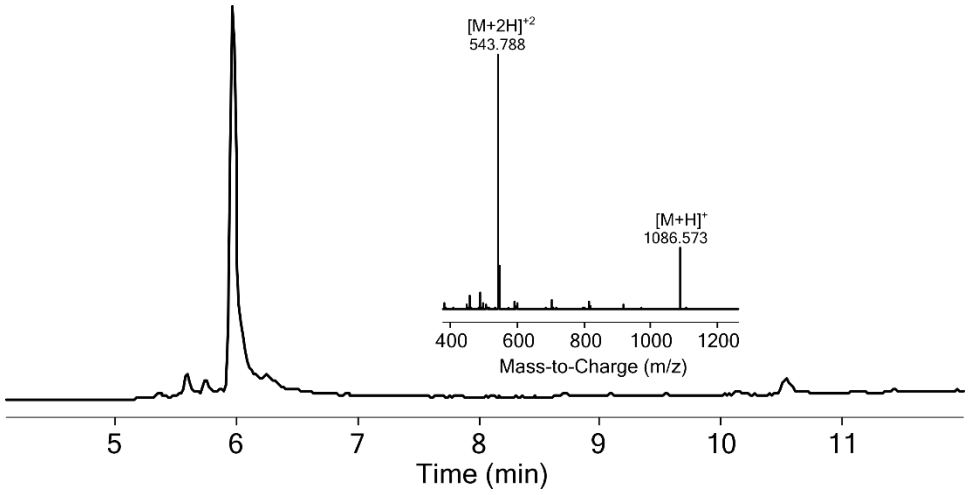


Figure S10. HPLC-MS (TIC) chromatogram for the purified peptide H_2N-Ala-Val-Val-Ile-Asn-Thr-Phe-Asp-Gly-His-$CONHNH_2$ with MS inset on the right. Calc. monoisotopic mass = 1085.6 Da.

4.11. H_2N-Ala-Met-Val-Ile-Asn-Thr-Phe-Asp-Gly-Phe-$CONHNH_2$

The peptide was synthesized and purified as described above (Section 2) on 385 mg of resin. Purification of crude peptide yielded 12.4 mg of pure peptide after lyophilization.

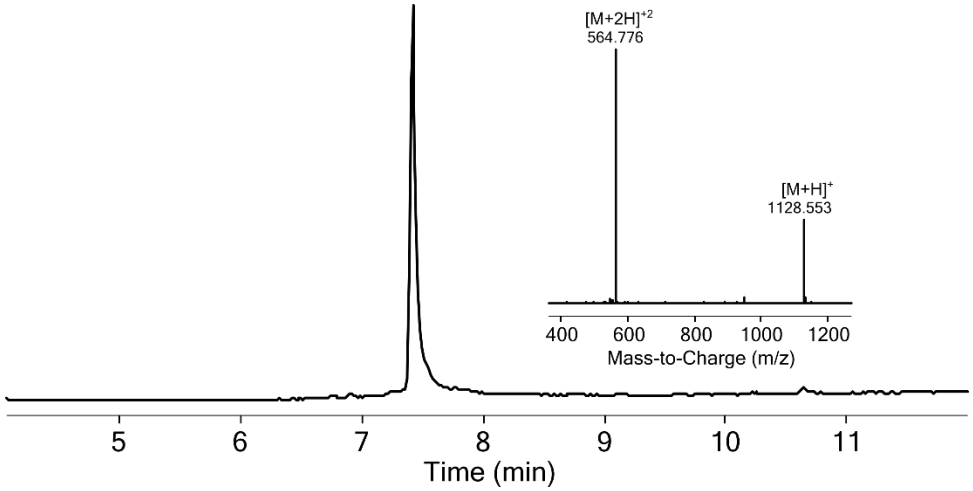


Figure S11. HPLC-MS (TIC) chromatogram for the purified peptide H_2N-Ala-Met-Val-Ile-Asn-Thr-Phe-Asp-Gly-Phe-$CONHNH_2$ with MS inset on the right. Calc. monoisotopic mass = 1127.54 Da.

4.12. H_2N-Ala-Lys-Val-Ile-Asn-Thr-Phe-Asp-Gly-Pro-$CONHNH_2$

The peptide was synthesized and purified as described above (Section 2) on 236 mg of resin. Purification of crude peptide yielded 3.7 mg of pure peptide after lyophilization.

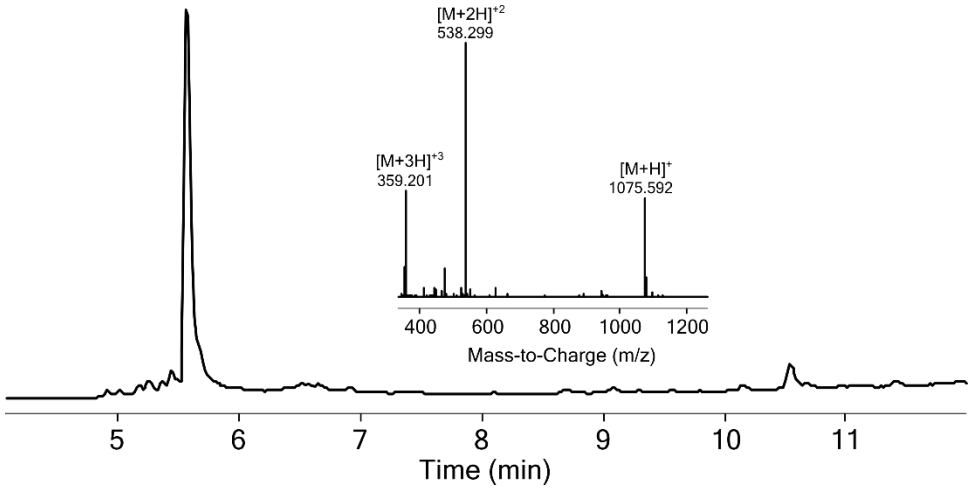


Figure S12. HPLC-MS (TIC) chromatogram for the purified peptide H_2N-Ala-Lys-Val-Ile-Asn-Thr-Phe-Asp-Gly-Pro-$CONHNH_2$ with MS inset on the right. Calc. monoisotopic mass = 1074.6 Da.

4.13. H_2N-Ala-Ser-Val-Ile-Asn-Thr-Phe-Asp-Gly-Gly-$CONHNH_2$

The peptide was synthesized and purified as described above (Section 2) on 194 mg of resin. Purification of crude peptide yielded 10.5 mg of pure peptide after lyophilization.

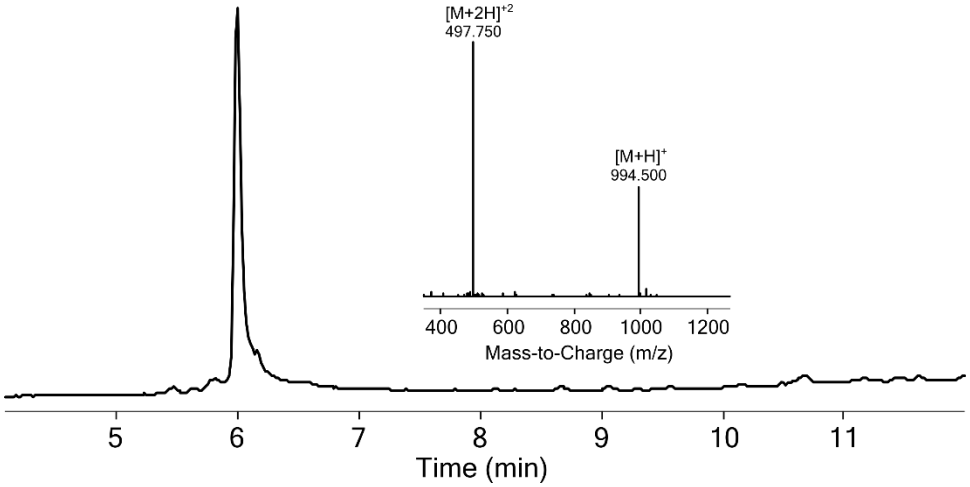


Figure S13. HPLC-MS (TIC) chromatogram for the purified peptide H_2N-Ala-Ser-Val-Ile-Asn-Thr-Phe-Asp-Gly-Gly-$CONHNH_2$ with MS inset on the right. Calc. monoisotopic mass = 993.5 Da.

4.14. H_2N-Ala-Asn-Val-Ile-Asn-Thr-Phe-Asp-Gly-Ser-$CONHNH_2$

The peptide was synthesized and purified as described above (Section 2) on 180 mg of resin. Purification of crude peptide yielded 5.6 mg of pure peptide after lyophilization.

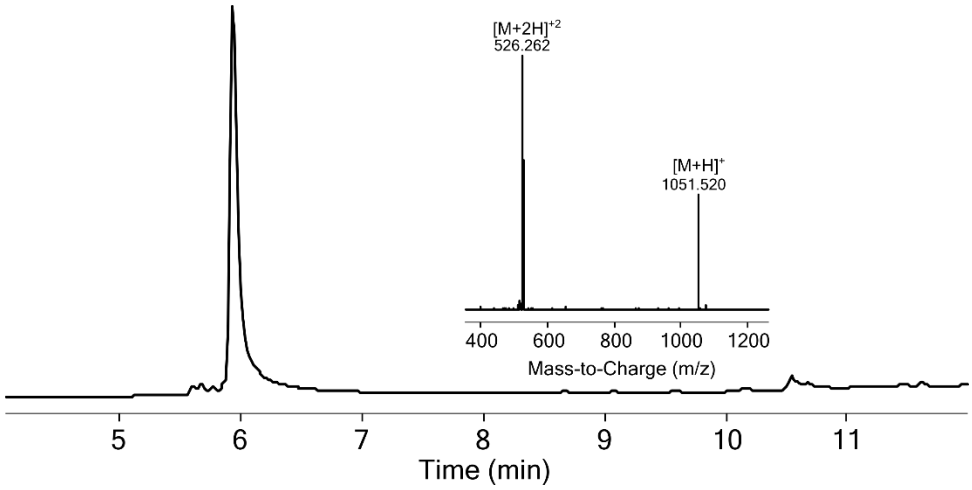


Figure S14. HPLC-MS (TIC) chromatogram for the purified peptide H_2N-Ala-Asn-Val-Ile-Asn-Thr-Phe-Asp-Gly-Ser-$CONHNH_2$ with MS inset on the right. Calc. monoisotopic mass = 1050.5 Da.

4.15. H_2N-Ala-Arg-Val-Ile-Asn-Thr-Phe-Asp-Gly-Tyr-$CONHNH_2$

The peptide was synthesized and purified as described above (Section 2) on 390 mg of resin. Purification of crude peptide yielded 28.0 mg of pure peptide after lyophilization.

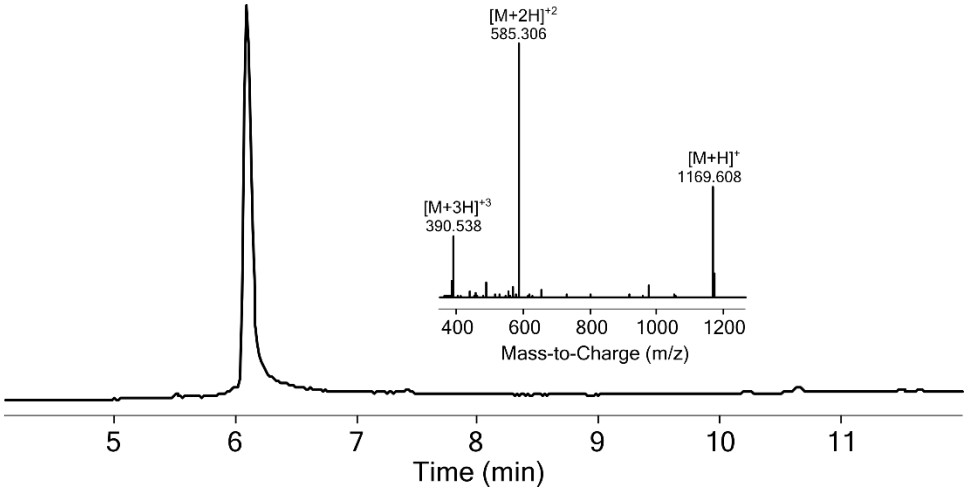


Figure S15. HPLC-MS (TIC) chromatogram for the purified peptide H_2N-Ala-Arg-Val-Ile-Asn-Thr-Phe-Asp-Gly-Tyr-$CONHNH_2$ with MS inset on the right. Calc. monoisotopic mass = 1168.6 Da.

4.16. H_2N-Ala-Pro-Val-Ile-Asn-Thr-Phe-Asp-Gly-Cys-$CONHNH_2$

The peptide was synthesized and purified as described above (Section 2) on 244 mg of resin. Purification of crude peptide yielded 2.8 mg of pure peptide after lyophilization.

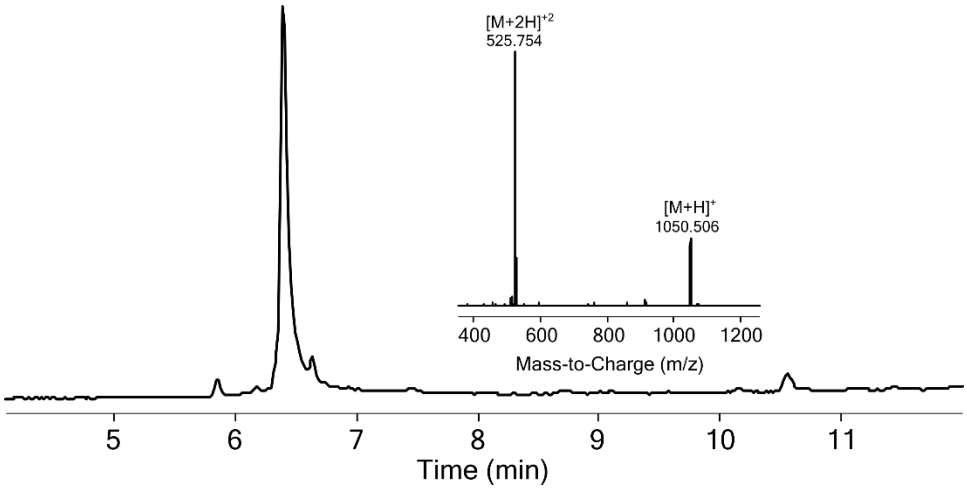


Figure S16. HPLC-MS (TIC) chromatogram for the purified peptide H_2N-Ala-Pro-Val-Ile-Asn-Thr-Phe-Asp-Gly-Cys-$CONHNH_2$ with MS inset on the right. Calc. monoisotopic mass = 1049.5 Da.

4.17. H_2N-Ala-Leu-Val-Ile-Asn-Thr-Phe-Asp-Gly-Glu-$CONHNH_2$

The peptide was synthesized and purified as described above (Section 2) on 396 mg of resin. Purification of crude peptide yielded 11.8 mg of pure peptide after lyophilization.

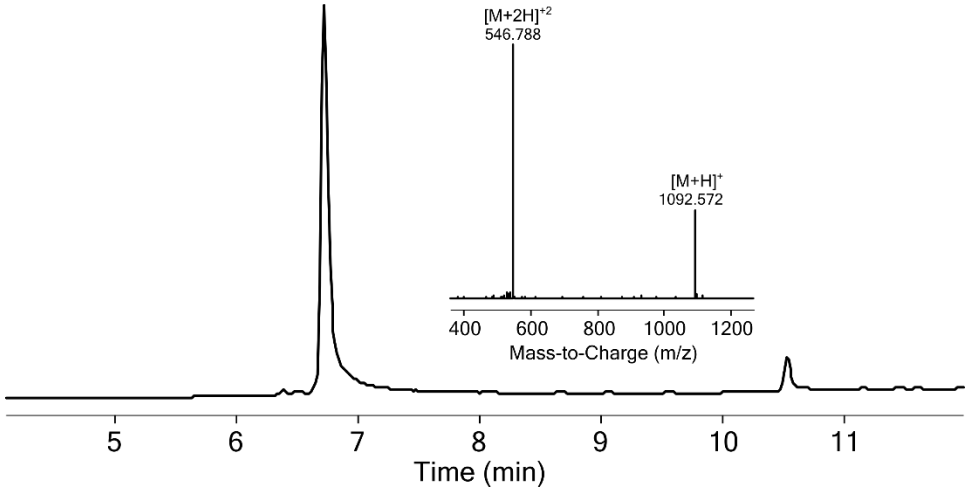


Figure S17. HPLC-MS (TIC) chromatogram for the purified peptide H_2N-Ala-Leu-Val-Ile-Asn-Thr-Phe-Asp-Gly-Glu-$CONHNH_2$ with MS inset on the right. Calc. monoisotopic mass = 1091.6 Da.

4.18. H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-^{D}Val-$CONHNH_2$

The peptide was synthesized as described above (Section 2) on 204 mg of resin, lyophilized and used without further purification. Lyophilization afforded 97 mg of peptide.

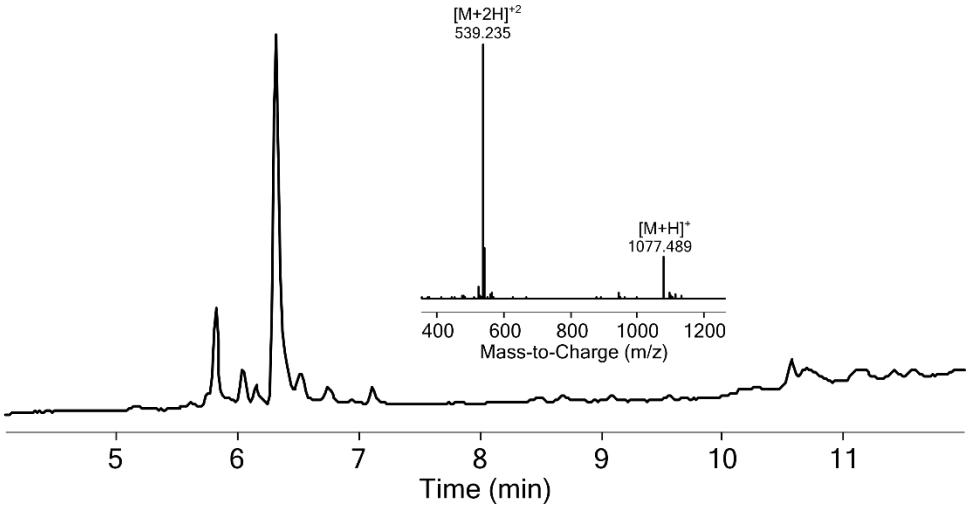


Figure S18. HPLC-MS (TIC) chromatogram for crude H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-^{D}Val-$CONHNH_2$ with MS inset on the right. Calc. monoisotopic mass = 1076.6 Da.

4.19. Peptides with Asp, Asn and Gln on the C-terminus

H_2N-Ala-Ala-Val-Ile-Asn-Thr-Phe-Asp-Gly-Asp-$CONHNH_2$, H_2N-Ala-Cys-Val-Ile-Asn-Thr-Phe-Asp-Gly-Asn-$CONHNH_2$, and H_2N-Ala-Ile-Val-Ile-Asn-Thr-Phe-Asp-Gly-Gln-$CONHNH_2$ were synthesized as described above. However, during cleavage and handling of the peptides, even at low pH, we observed an intramolecular cyclization that prevented us from isolating the desired hydrazides (Fig. S19). This side-reaction, accompanying synthesis of C-terminal hydrazides of Asp, Asn, and Gln, was described in detail by Fang, Li et al. *Angew. Chem. Int. Ed.* **2011**, 50, 7645 –7649.

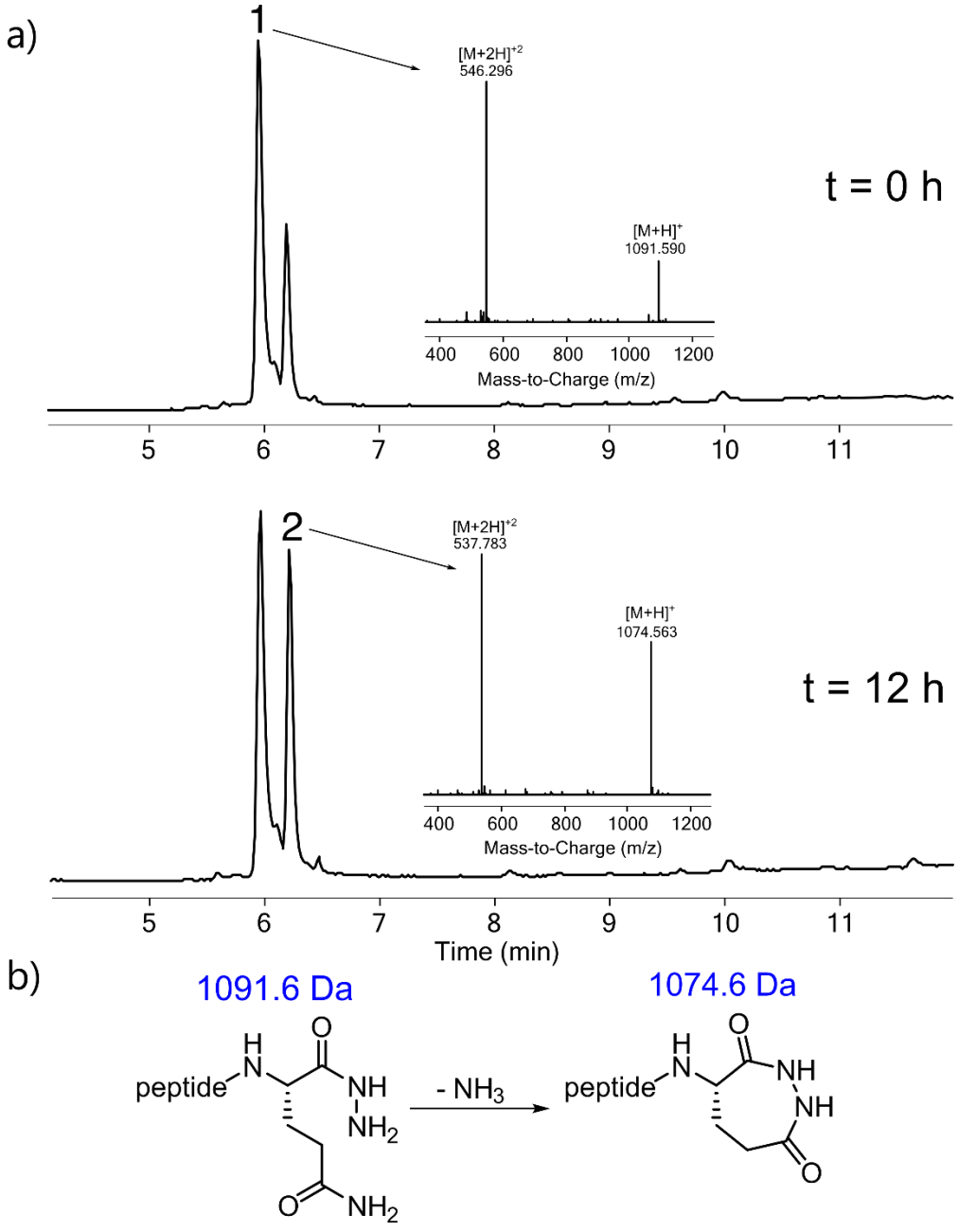


Figure S19. a) HPLC-MS (TIC) chromatograms obtained 12 hours apart from one HPLC fraction from the attempted purification of H_2N-Ala-Ile-Val-Ile-Asn-Thr-Phe-Asp-Gly-Gln-$CONHNH_2$ The amount of undesired cyclization product, corresponding to peak 2, increased significantly during this time. b) The reported explanation for the formation of the observed -17 Da product: cyclization with the Gln side chain.

4.20. H_2N-Gly-Gly-Gly-Gly-Gly-Leu-Glu-Ile-$CONHNH_2$

The peptide was synthesized as described above (Section 2) on 215 mg of resin. Following TFA cleavage the crude was redissolved in 99% A/ 1% B, and the resin was filtered out. The peptide was purified on the Waters system described in section 2.3 using the following gradient: 1% B in A for 5 minutes, then 1%-31% B ramping linearly over 90 minutes. Purification of crude peptide yielded 3.8 mg of pure peptide after lyophilization.

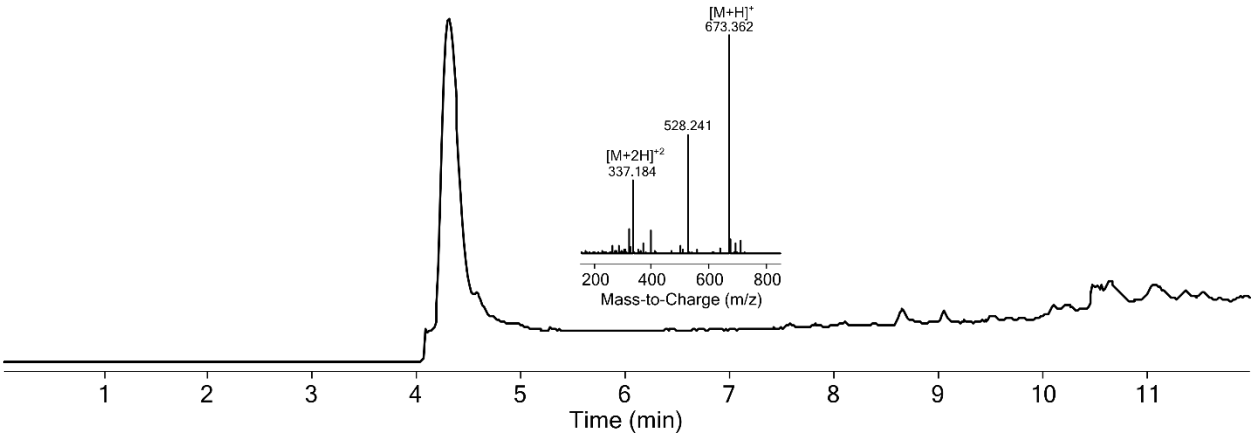


Figure S20. HPLC-MS (TIC) chromatogram for crude H_2N-Gly_5-Leu-Glu-Ile-$CONHNH_2$ with MS inset on the right. Calc. monoisotopic mass = 672.4 Da.

5. Additional experimental data

5.1. Discovery of isocyanate generation

Some data shown in this section are a more detailed reproduction of the data mentioned in Mong, Vinogradov et al, *ChemBioChem*, **2014**, 15, 721-733 describing the total synthesis of barnase. Originally, during oxidation/thioesterefication of barnase fragments, we consistently observed formation of MPAA thioester, contaminated with a side-product 15.0 Da more massive than thioester. It was observed for two different fragments.

In particular, during these studies we adhered to the standard oxidation/thioesterefication protocol described by Fang, Li et al. (*Angew. Chem. Int. Ed*. **2011**, 50, 7645 –7649). Thus, 2.00

mg (1.47 μmol) of H_2N-Gly_5-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$ were dissolved in 0.48 mL of oxidation buffer (200 mM Na_2HPO_4 and 6 M Gn·HCl in water, pH 3.2). The solution was incubated in an ice-salt bath at –18°C for 10 min and then 48 μL of 200 mM $NaNO_2$ solution in water was added to it dropwise while stirring. After 20 min an analytical sample for LC-MS was taken (*Fig. S21*) and then 0.48 mL of 0.2 M MPAA in ligation buffer (200 mM Na_2HPO_4 and 6 M Gn·HCl in water, pH 6.8) were added to the reaction. Upon complete addition of the MPAA solution, another analytical sample for the LC-MS was taken (*Fig. S22*).

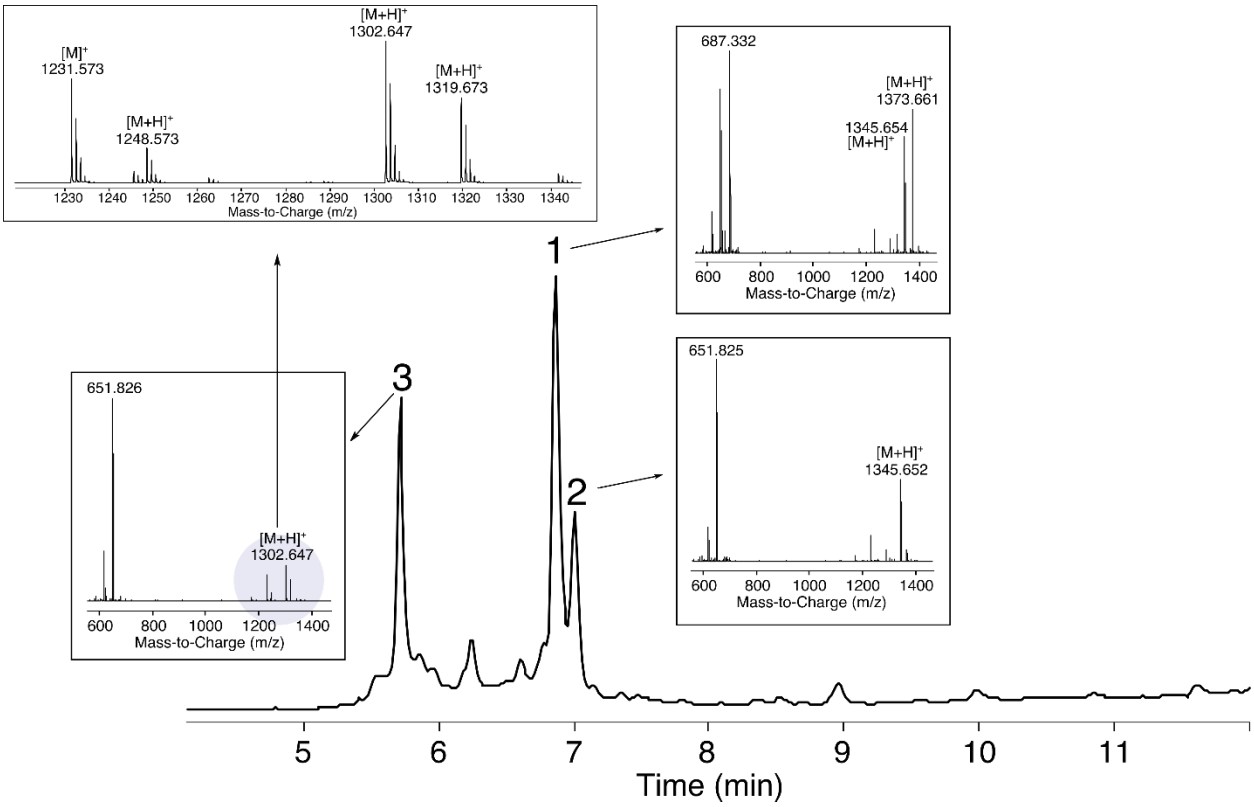
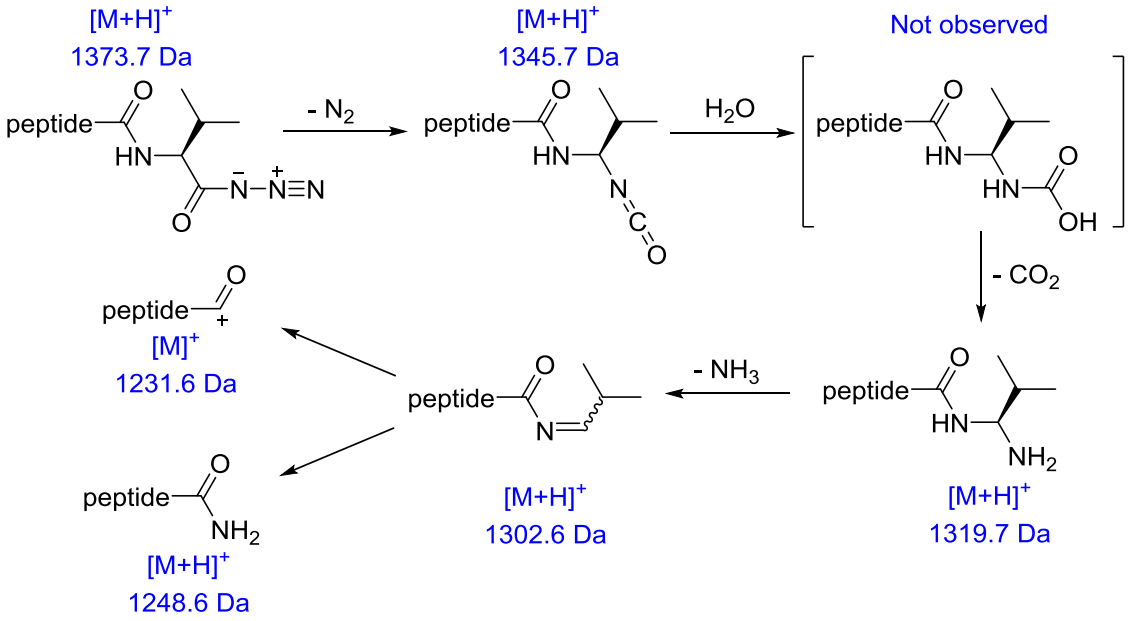


Figure S21 HPLC-MS (TIC) chromatogram for Gly_5-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$ oxidation reaction with MS insets for main products. Peak 1 corresponds to the expected acyl azide product (calc. monoisotopic mass = 1372.7 Da). The chromatogram shows the formation of isocyanate (peak 2 with MS inset of the peak apex on the bottom right) during hydrazide oxidation in the studied system. Peak 3 with MS insets on the left corresponds to isocyanates hydrolysis products.

The data presented in *Fig. S21* suggest that the hydrazide is successfully oxidized to form the acyl azide (peak 1). Two sets of ions (one is -28 Da relative to another) for peak **1** may be explained by the fragmentation of azide in the mass spectrum. Indeed, this behavior is quite common for azides (Li, Hoskins *et al*. *Macromolecules*, **2010**, 43, 6225–6228). In contrast, peak **2** is chromatographically resolved from the azide peak and there is no parent azide ion, which

allowed us to conclude that peak **2** is a real compound 28 Da less massive than the azide. The mass difference of 28 Da corresponds to a loss of a nitrogen molecule, which is consistent with the hypothesis that the acyl azide undergoes a Curtius rearrangement to release nitrogen and form an isocyanate. Analysis of peak **3** further supports this conclusion. The four major ions observed in the MS of the apex of peak **3** are consistent with isocyanate hydrolysis, which may be explained by the degradation of the C-terminal amino acid, which is accompanied by many intermediates, most of which can be observed in the MS (*Scheme S1*).

Scheme S1. Transformations of the C-terminal isocyanate in water.



We observed similar ions series for most peptides in this study. Because in most cases we observed several products stemming from isocyanate hydrolysis, we refer to them collectively as 'isocyanate hydrolysis products' in the manuscript.

Figure S22 shows the crude reaction mixture after addition of MPAA. Formation of peak 2 (+15 Da relative to the expected thioester) is consistent with the hypothesis that the acyl azide rearranges to isocyanate, because isocyanate may react with MPAA to give carbamothioate, which is 15 Da heavier than thioester (*Scheme S2*).

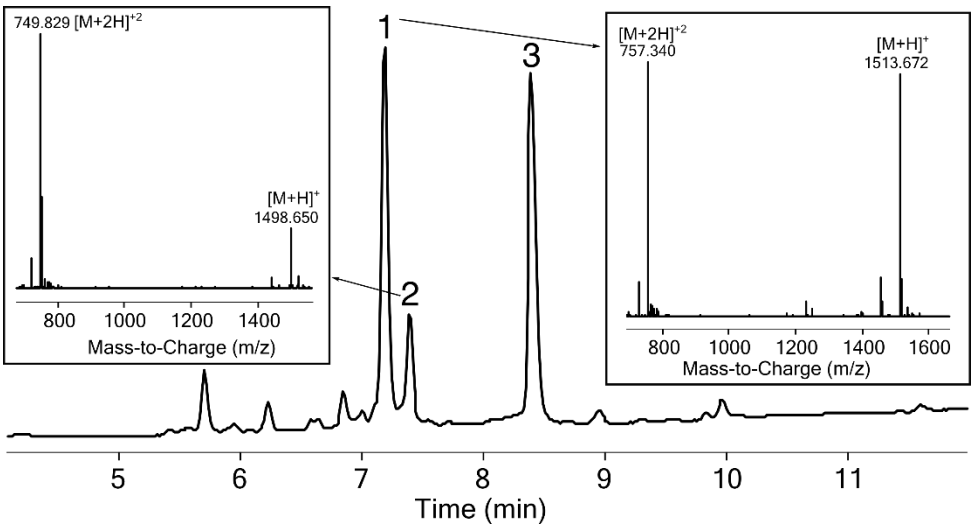
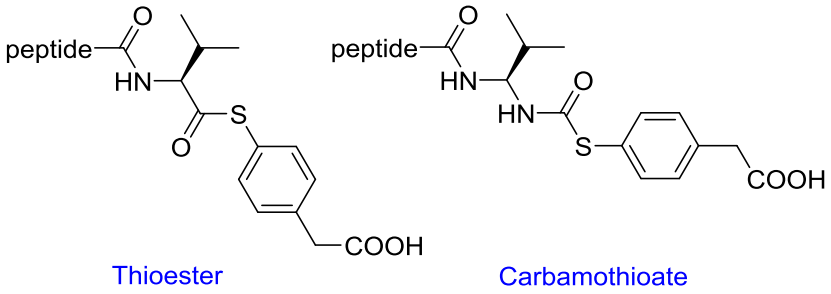


Figure S22: HPLC-MS (TIC) chromatogram for Gly_5-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$ oxidation/transesterification reaction with MS insets for main products. Peak 2 corresponds to the expected thioester product (calc. monoisotopic mass = 1498.6 Da). Peak 1 is the unexpected carbamothioate formed from the corresponding isocyanate. Peak 3 is MPAA disulfide.

Scheme S2. Chemical structures of C-terminal thioesters and carbamothioates: main products observed during thioesterefication of Gly_5-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$.



To exclude the possibility of side chains reacting with sodium nitrite or MPAA, we prepared the peptide H_2N-Gly_5-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Val-COOH and subjected it to the reaction conditions described above. The peptide was unreactive under these conditions, suggesting that the +15 Da product indeed stems from the C-terminal hydrazide oxidation.

Finally, we confirmed the structure of the conjugation product using MS/MS analysis. To this end, we performed oxidation/conjugation of a model peptide H_2N-Ala-Arg-Val-Ile-Asn-Thr-Phe-Asp-Gly-Tyr-$CONHNH_2$ under standard conditions (section 5.3) with 15 mM perfluorophenylhydrazine as a nucleophile at 57 °C. HPLC-MS chromatogram for the crude reaction mixture shown in *Figure S23* confirmed that the nucleophile was conjugated to the peptide. We then performed the secondary MS analysis of the main product $[M+H]^+$ and

$[M+2H]^{+2}$ ions. As *Figure S24* shows, the fragmentation patterns of both ions match the expected structure well. These data unambiguously indicate that perfluorophenylhydrazine conjugated to the C-terminus of the peptide. More importantly, we observed the extensive fragmentation of the C-terminal moiety, which confirmed the presence of the semicarbazide functional group. Thus, we concluded that C-terminal acyl azides rearrange to isocyanates, which react with external nucleophiles to give conjugation products.

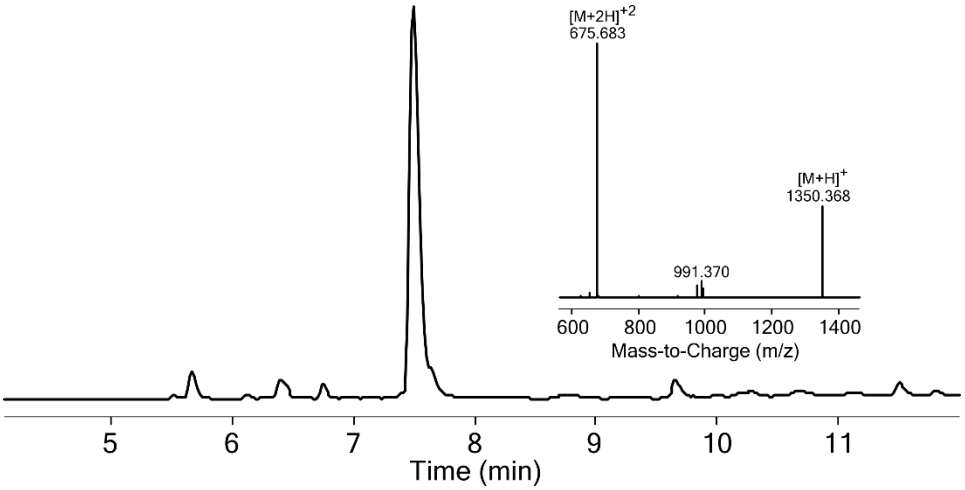


Figure S23. HPLC-MS (TIC) chromatogram for Ala-Arg-Val-Ile-Asn-Thr-Phe-Asp-Gly-Tyr-$CONHNH_2$ oxidation/conjugation reaction with perfluorophenylhydrazine (MS inset for the main peak apex is on the right; calc. monoisotopic mass of the expected semicarbazide ion is 1350.6 Da)

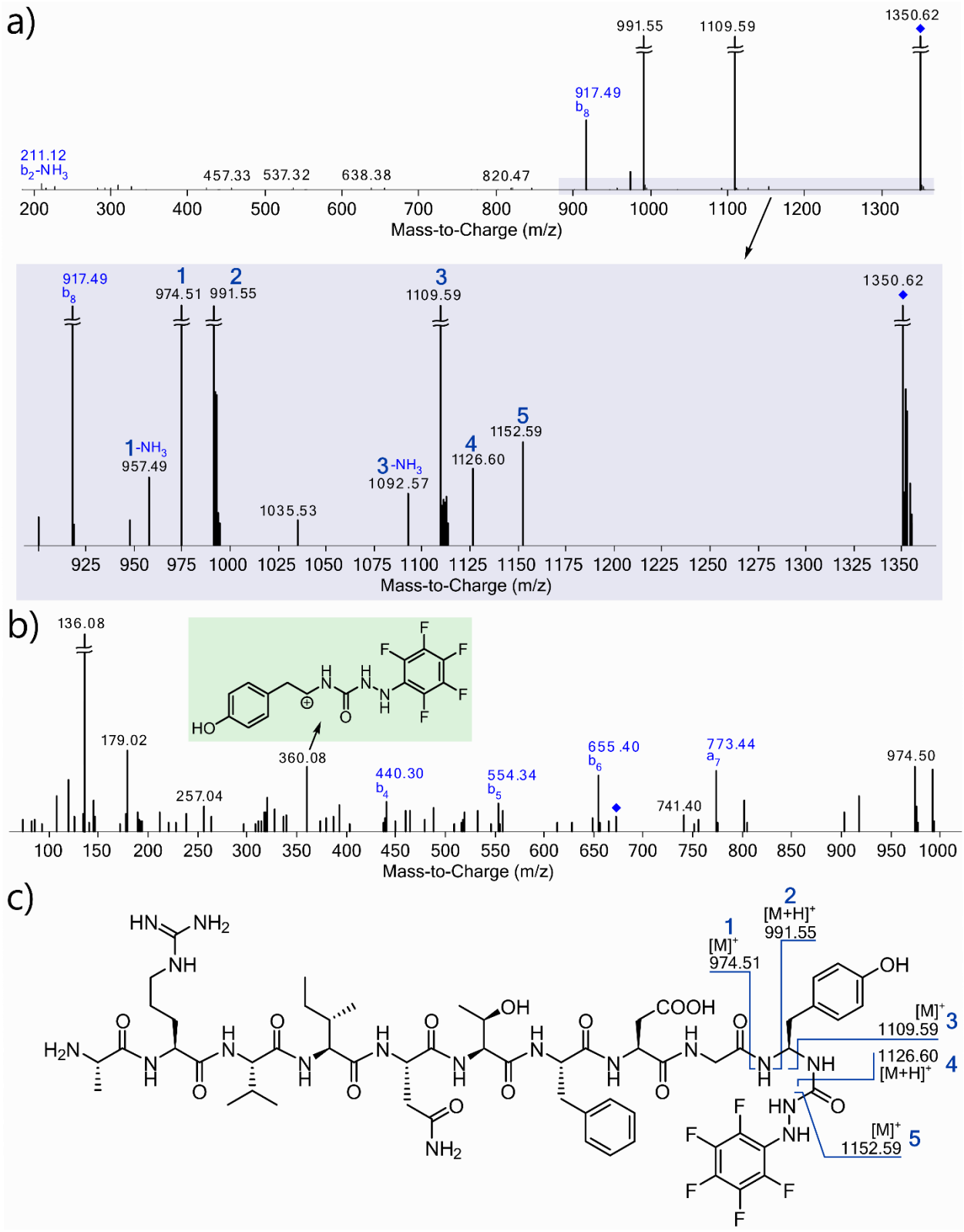


Figure S24. a) Top: MS/MS chromatogram for $[M+H]^+$ ion of the conjugation product. The parent ion is labelled as diamond. Bottom: zoom-in of the 900-1350 m/z region of the spectrum on the top. Peak labels correspond to fragments shown in panel c). b) MS/MS chromatogram for $[M+2H]^{+2}$ ion of the conjugation product. The parent ion is labelled as diamond. Some a_n and b_n

fragments are labelled. The C-terminal fragment containing semicarbazide was also identified and labelled. c) Structural formula of the conjugation product indicating the most important fragments.

5.2. Optimization of conjugation conditions

To optimize reaction conditions we chose H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$ as a model peptide and MPAA as a nucleophile. 4.14 mg (3.85 μmol) of peptide hydrazide were dissolved in 1.25 mL of oxidation buffer (200 mM Na_2HPO_4 and 6 M Gn·HCl in water, pH 3.2). The solution was incubated in an ice-salt bath at –10 °C for 5 min and then 125 μL of 200 mM $NaNO_2$ solution in water was added to it dropwise. After 10 min, 100 μl oxidized peptide solution was added to each of the following solutions:

a) 100 μl of 200 mM MPAA in 200 mM Na_2HPO_4, 6 M Gn·HCl buffer, pH 4.5 (thioester control experiment).

b) 100 μl of 200 mM MPAA in 200 mM Na_2HPO_4, 6 M Gn·HCl buffer, pH 9.4.

c) 100 μl of 200 mM MPAA in 200 mM Na_2HPO_4, 6 M Gn·HCl buffer, pH 4.5. pH of the resulting solution was immediately adjusted to 7.4.

d) 200 μl of acetonitrile. After 5 minutes at rt, 100 μl of 200 mM MPAA in 200 mM Na_2HPO_4, 6 M Gn·HCl buffer, pH 4.5 was added.

e) 300 μl of 67 mM MPAA in water/acn (1: 1, v/v) with 0.1% TFA added.

f) 100 μl of 200 mM MPAA in 200 mM Na_2HPO_4, 6 M Gn·HCl buffer, pH 4.5. The resulting solution was then immediately added to 1.8 mL water/acn (1: 1, v/v) with 0.1% TFA (final pH ~ 3.0).

In each case the reaction was allowed to proceed for 120 minutes before an HPLC-MS sample was taken.

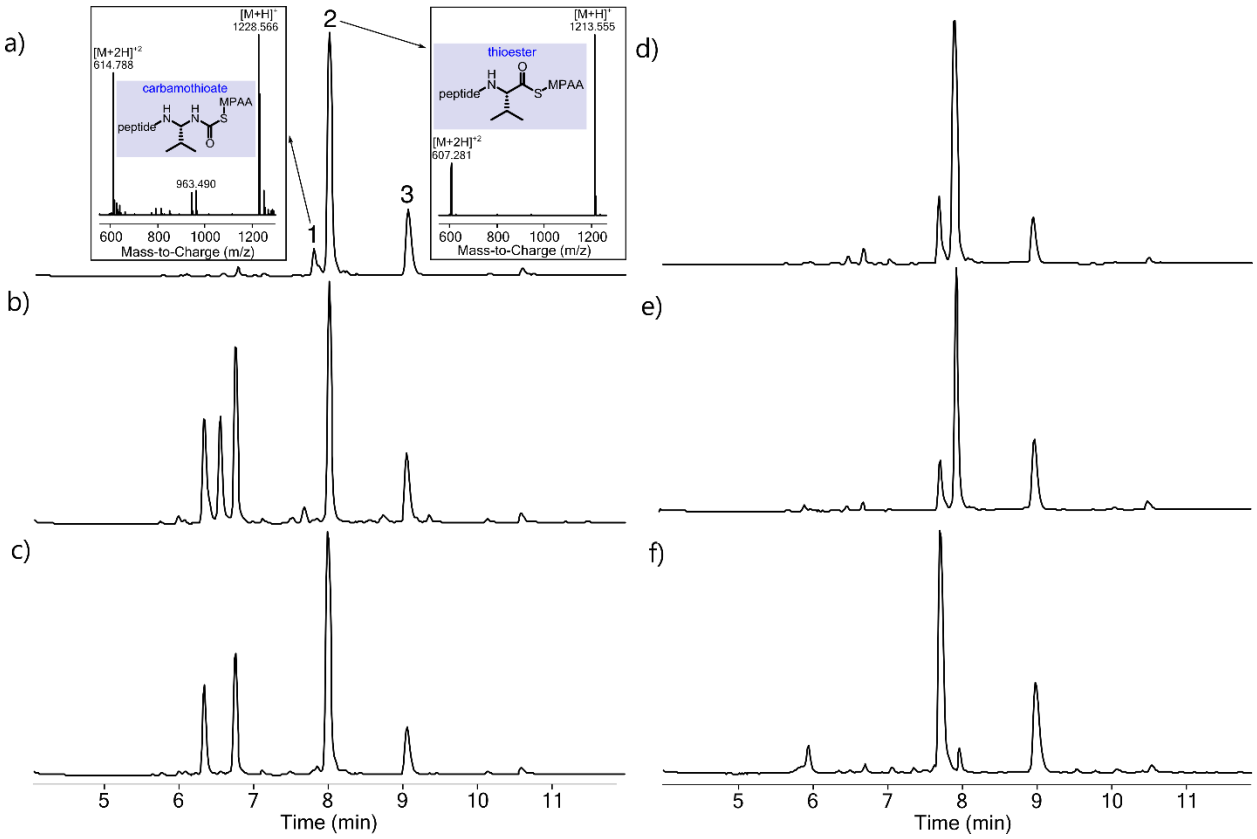


Figure S25. HPLC-MS (TIC) chromatograms for H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$ oxidation/conjugation reaction with MS insets for main products. Peak 1 is the carbamothioate, Peak 2 is the thioester, and Peak 3 is MPAA disulfide. Chromatogram indices correspond to the reaction conditions specified above.

Data shown in *Fig. S25* suggest that the exact order of reagent addition, pH of the reaction, or solvent composition had little influence on isocyanate generation. In fact, for reactions a) – e) we did not observe the formation of carbamothioate or other isocyanate derived products to any significant extent. In contrast, when we diluted both reagents (peptide to a final concentration of 133 μM, and MPAA to ~ 10 mM) we observed formation of carbamothioate as the primary reaction product.

To elucidate the solvent effects we then performed 3 more reactions. 0.65 mg peptide was oxidized as above, and then 35 μl of oxidized peptide solution was added to each of the following solutions:

a) 1.0 mL 8 mM MPAA solution in water/acn (1: 1, v/v) with 0.1% TFA added.

b) 1.0 mL 8 mM MPAA solution in water/acn (3: 1, v/v) with 0.1% TFA added.

c) 1.0 mL 8 mM MPAA solution in water with 0.1% TFA added.

In each case the reaction was allowed to proceed for 120 minutes before an HPLC-MS sample was taken (*Fig. S26*). We concluded that the concentration of acetonitrile in the solvent plays a minor role in the overall process and opted for water/acn (1: 1, v/v) as a universal solvent for further experiments.

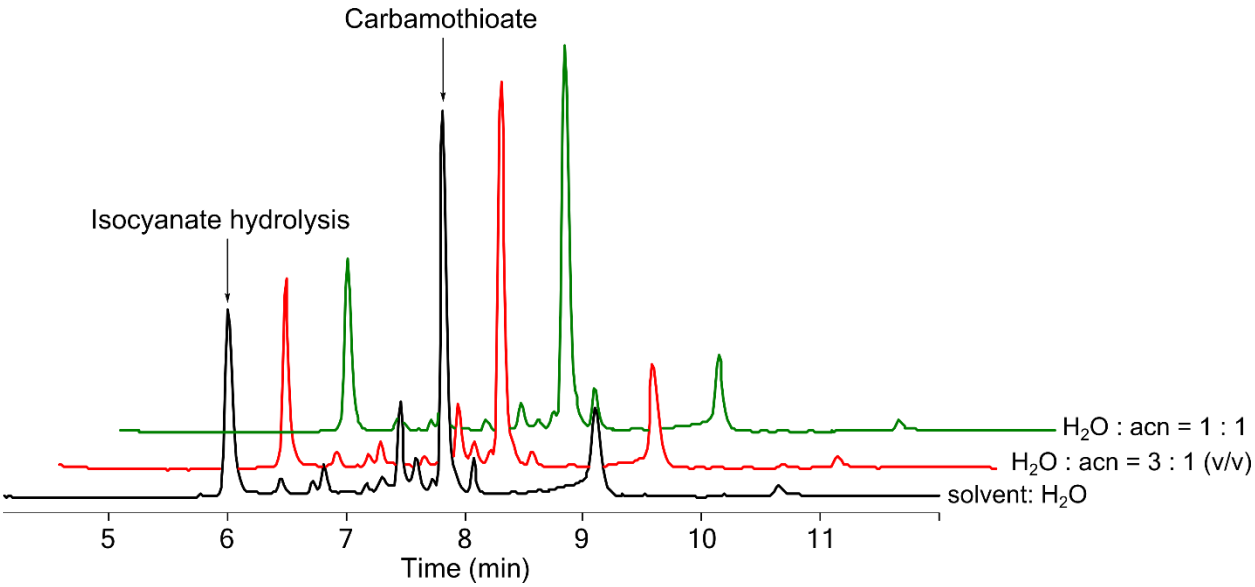


Figure S26. Overlaid HPLC-MS (TIC) chromatograms for H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$ oxidation/conjugation reaction in different solvents. Although the cleanest product was obtained in water/acn (1: 1, v/v), the influence of acetonitrile concentration is minor.

5.3. Standard reaction procedure for conjugation of nucleophiles to peptide isocyanates

Unless noted, all isocyanate conjugations below are performed under the following standard conditions: 3.07 mM peptide hydrazide solution in oxidation buffer (200 mM Na_2HPO_4 and 6 M Gn·HCl in water, pH 3.2) was incubated in an ice-salt bath at –10 °C for 5 min, and then 200 mM $NaNO_2$ solution in water was carefully added to the peptide to a final $NaNO_2$ concentration of 18 mM. The reaction was allowed to proceed for 10 minutes at –10 °C. Then, a solution of nucleophile in water/acetonitrile (1: 1, v/v, 0.1% TFA added, pH ~ 3.0-4.5) was added to the reaction mixture to a final peptide concentration of 133 μM and the mixture was warmed to the desired temperature (usually 17 or 57 °C). Depending on the temperature, the reaction mixture was analyzed using HPLC-MS after 20 minutes (57 °C reaction) or 120 minutes (17 °C reaction).

5.4. Nucleophile scope of the reaction

For this study the reactions were carried out following the standard conditions at 60 °C for 20 minutes except water/acetonitrile (1:3, v/v) was used as a solvent for the conjugation step. Following the reaction, the solutions were further acidified with citric acid monohydrate to its final concentration of 150 mM and cooled down to room temperature. The excess of acetonitrile was removed *in vacuo*; the remaining reaction mixtures were diluted into 95% A/5% B and purified by RP-HPLC as described in section 2.3.

5.4.1. H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$ + Perfluorophenyl-hydrazine

4.8 mg peptide was dissolved in 1.45 mL oxidation buffer, oxidized with 145 µL of 200 mM $NaNO_2$, and diluted into 33 mL of 25 mM perfluorophenylhydrazine in water/acetonitrile (1: 3, v/v). After 1.04 g of citric acid was added to the reaction mixture, the excess of acetonitrile was evaporated and the remaining solution (~10 mL) was diluted with 95% A/5% B to a total of 40 mL. RP-HPLC purification followed by lyophilization yielded 4.6 mg (82% yield) of the desired semicarbazide as a white powder (*Fig. S27*).

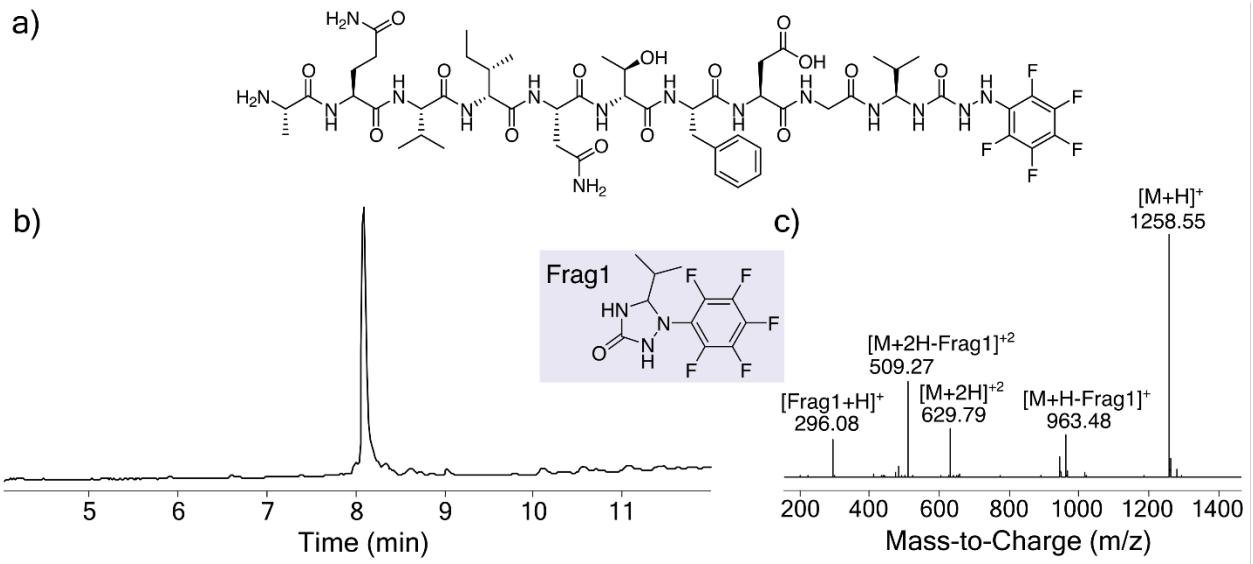


Figure S27. HPLC-MS (TIC) chromatogram (b) for the purified semicarbazide peptide H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-NH-CONHNH-C_6F_5 with the peptide structure shown in a). c) MS of the compound with major peaks assigned. ESI-induced peptide fragment Frag1 is postulated to exist in a cyclic form as drawn. Calc. monoisotopic mass = 1257.56 Da.

5.4.2. H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$ + 2,4-dinitrophenyl-hydrazine

4.8 mg peptide was dissolved in 1.45 mL oxidation buffer, oxidized with 145 µL of 200 mM $NaNO_2$, and diluted into 33 mL of saturated solution of 2,4-dinitrophenylhydrazine in water/acetonitrile (1: 3, v/v). After 1.04 g of citric acid was added to the reaction mixture, the excess of acetonitrile was evaporated and the remaining solution (~10 mL) was diluted with 95% A/5% B to a total of 40 mL. RP-HPLC purification followed by lyophilization yielded 2.7 mg (48% yield) of the desired semicarbazide as an ochre powder (*Fig. S28*).

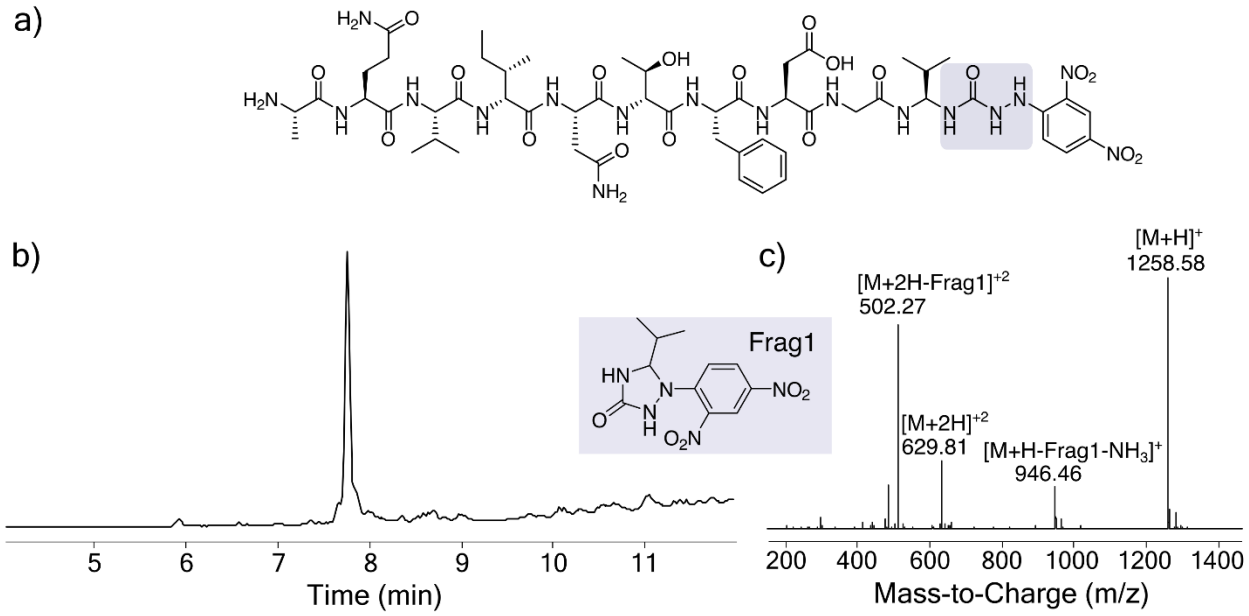


Figure S28. HPLC-MS (TIC) chromatogram (b) for the purified semicarbazide peptide H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-NH-CONHNH-C_6H_4-$(NO_2)_2$ with the peptide structure shown in a). c) MS of the compound with major peaks assigned. ESI-induced peptide fragment Frag1 is postulated to exist in a cyclic form as drawn. Calc. monoisotopic mass = 1257.56 Da.

5.4.3. H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$ + 4-hydroxybenzhydrazide

9.1 mg peptide was dissolved in 2.75 mL oxidation buffer, oxidized with 275 µL of 200 mM $NaNO_2$, and diluted into 64 mL of 30 mM 4-hydroxybenzhydrazide in water/acetonitrile (1: 3, v/v). After 1.89 g of citric acid was added to the reaction mixture, the excess of acetonitrile was evaporated and the remaining solution was diluted with 95% A/5% B to a total of 40 mL. RP-HPLC purification followed by lyophilization yielded 7.5 mg (73% yield) of the desired N'-acyl-semicarbazide as a white powder (*Fig. S29*).

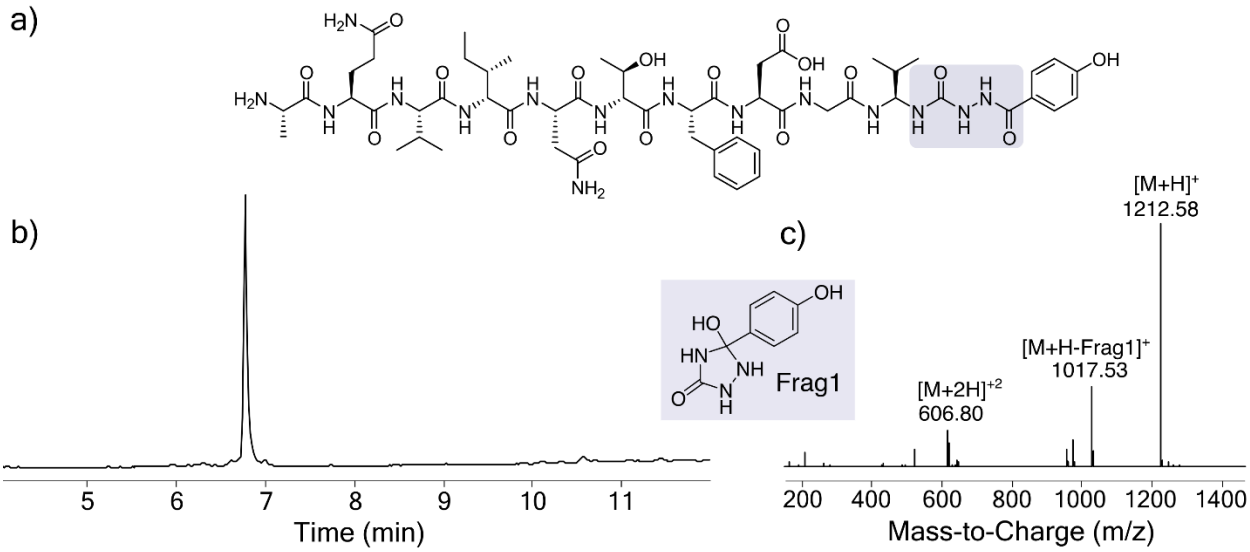


Figure S29. HPLC-MS (TIC) chromatogram (b) for the purified N'-acyl-semicarbazide peptide H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-NH-CONHNHCO-C_6H_5-OH with the peptide structure shown in a). c) MS of the compound with major peaks assigned. ESI-induced peptide fragment Frag1 is postulated to exist in a cyclic form as drawn. Calc. monoisotopic mass = 1211.59 Da.

5.4.4. H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$ + D-biotin hydrazide

3.1 mg peptide was dissolved in 0.94 mL oxidation buffer, oxidized with 94 µL of 200 mM $NaNO_2$, and diluted into 11.5 mL of 30 mM D-biotin hydrazide in water/acetonitrile (1: 3, v/v). After 360 mg of citric acid was added to the reaction mixture, the excess of acetonitrile was evaporated and the remaining solution was diluted with 95% A/5% B to a total of 25 mL. RP-HPLC purification followed by lyophilization yielded 2.6 mg (68% yield) of the desired N'-acyl-semicarbazide as a white powder (*Fig. S30*).

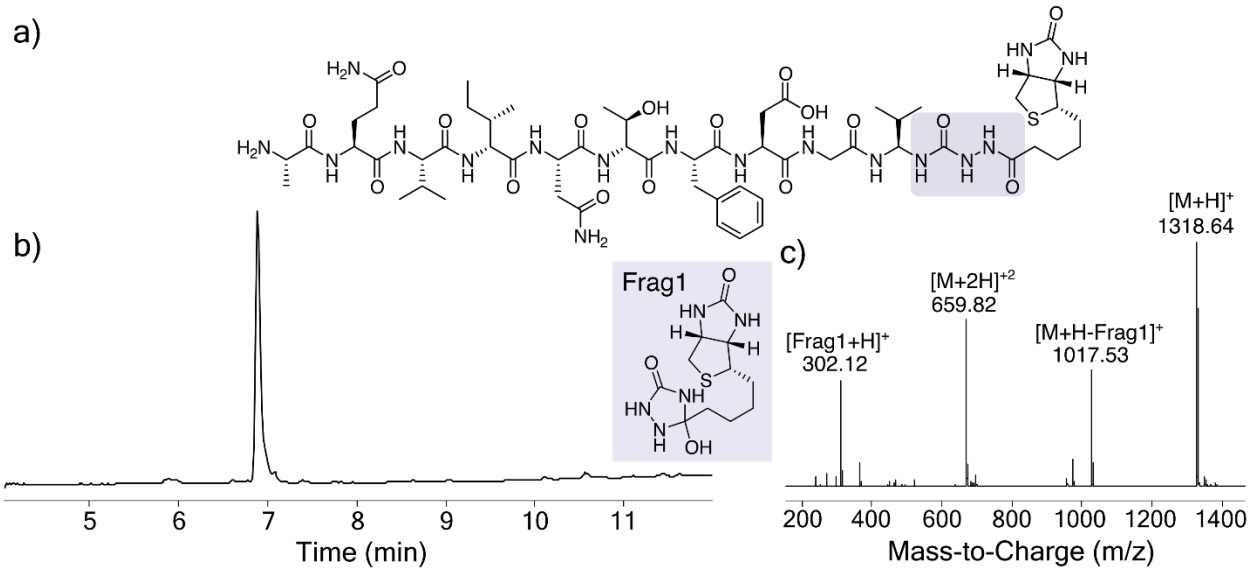


Figure S30. HPLC-MS (TIC) chromatogram (b) for the purified N'-acyl-semicarbazide peptide H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-NH-CONHNHCO-biotin with the peptide structure shown in a). c) MS of the compound with major peaks assigned. ESI-induced peptide fragment Frag1 is postulated to exist in a cyclic form as drawn. Calc. monoisotopic mass = 1317.65 Da.

5.4.5. H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$ + O-(2,3,4,5,6-pentafluorobenzyl)-hydroxylamine

9.0 mg peptide was dissolved in 2.72 mL oxidation buffer, oxidized with 272 µL of 200 mM $NaNO_2$, and diluted into 63 mL of 100 mM O-(2,3,4,5,6-pentafluorobenzyl)-hydroxylamine hydrochloride in water/acetonitrile (1: 3, v/v) pHed to 3.1 with 5 M NaOH in water. After 1.89 g of citric acid was added to the reaction mixture, the excess of acetonitrile was evaporated and the remaining solution (~15 mL) was diluted with 95% A/5% B to a total of 40 mL. RP-HPLC purification followed by lyophilization yielded 5.6 mg (53% yield) of the desired hydroxyurea derivative as a white powder (*Fig. S31*).

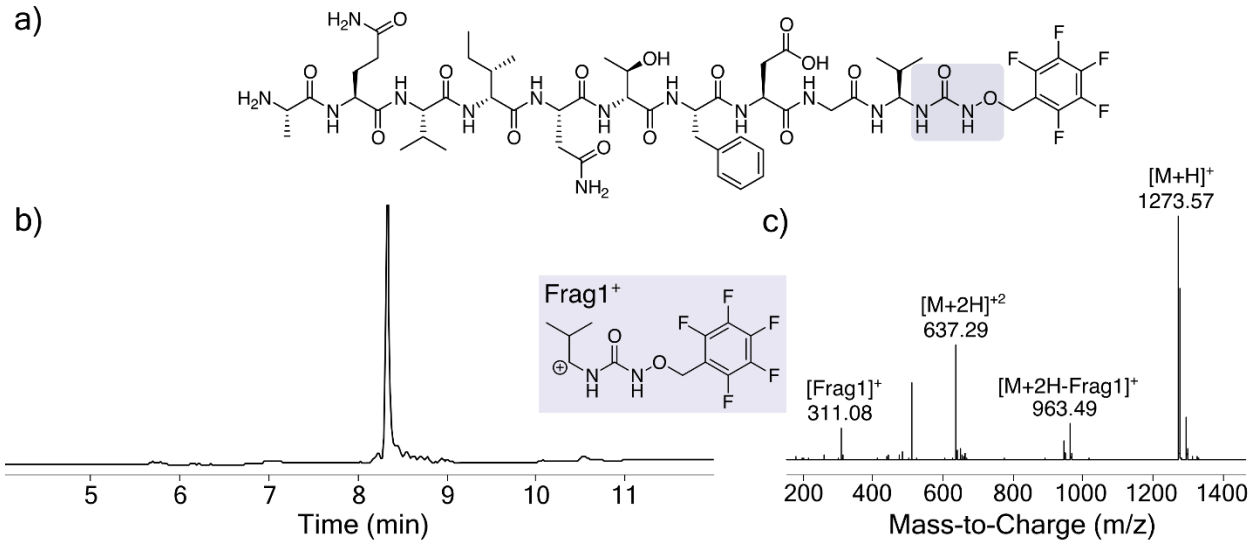


Figure S31. HPLC-MS (TIC) chromatogram (b) for the purified hydroxyurea derivative peptide H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-NH-CONHO-CH_2-C_6F_5 with the peptide structure shown in a). c) MS of the compound with major peaks assigned. ESI-induced peptide fragment Frag1 is postulated to exist in a linear form as drawn. Calc. monoisotopic mass = 1272.56 Da.

5.4.6. H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$ + methoxyamine

9.5 mg peptide was dissolved in 2.87 mL oxidation buffer, oxidized with 287 µL of 200 mM $NaNO_2$, and diluted into 67 mL of 100 mM methoxyamine hydrochloride in water/acetonitrile (1: 3, v/v) pHed to 3.1 with 5 M NaOH in water. After 2.0 g of citric acid was added to the reaction mixture, the excess of acetonitrile was evaporated and the remaining solution (~15 mL) was diluted with 95% A/5% B to a total of 40 mL. RP-HPLC purification followed by lyophilization yielded 4.1 mg (42% yield) of the desired hydroxyurea derivative as a white powder (*Fig. S32*).

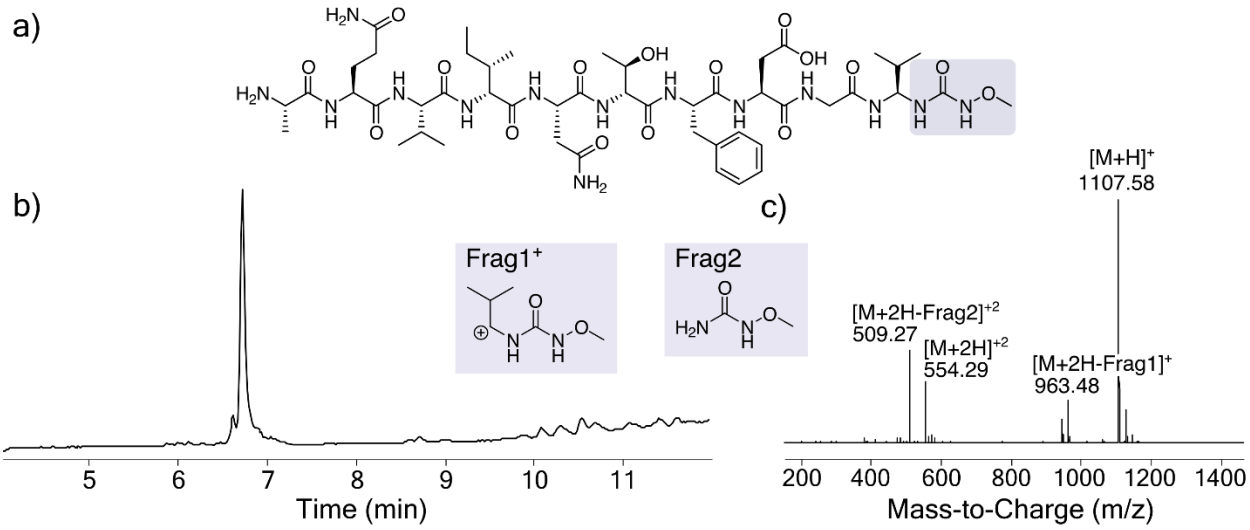


Figure S32. HPLC-MS (TIC) chromatogram (b) for the purified hydroxyurea derivative peptide H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-NH-$CONHOCH_3$ with the peptide structure shown in a). c) MS of the compound with major peaks assigned. ESI-induced peptide fragment Frag1 is postulated to exist in a linear form as drawn. Calc. monoisotopic mass = 1106.57 Da.

5.4.7. H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$ + MPAA

9.5 mg peptide was dissolved in 2.87 mL oxidation buffer, oxidized with 287 µL of 200 mM $NaNO_2$, and diluted into 67 mL of 10 mM MPAA in water/acetonitrile (1: 3, v/v), which was prepared by diluting 3.35 mL of 200 mM MPAA solution in 200 mM phosphate buffer, 6M Gn·HCl, pH 6.9 with 64 mL of water/acetonitrile (1: 3, v/v). After 2.0 mg of citric acid was added to the reaction mixture, the excess of acetonitrile was evaporated and the remaining solution was diluted with 95% A/5% B to a total of 40 mL. RP-HPLC purification followed by lyophilization yielded 6.5 mg (60% yield) of the desired carbamothioate as a white powder (*Fig. S33*).

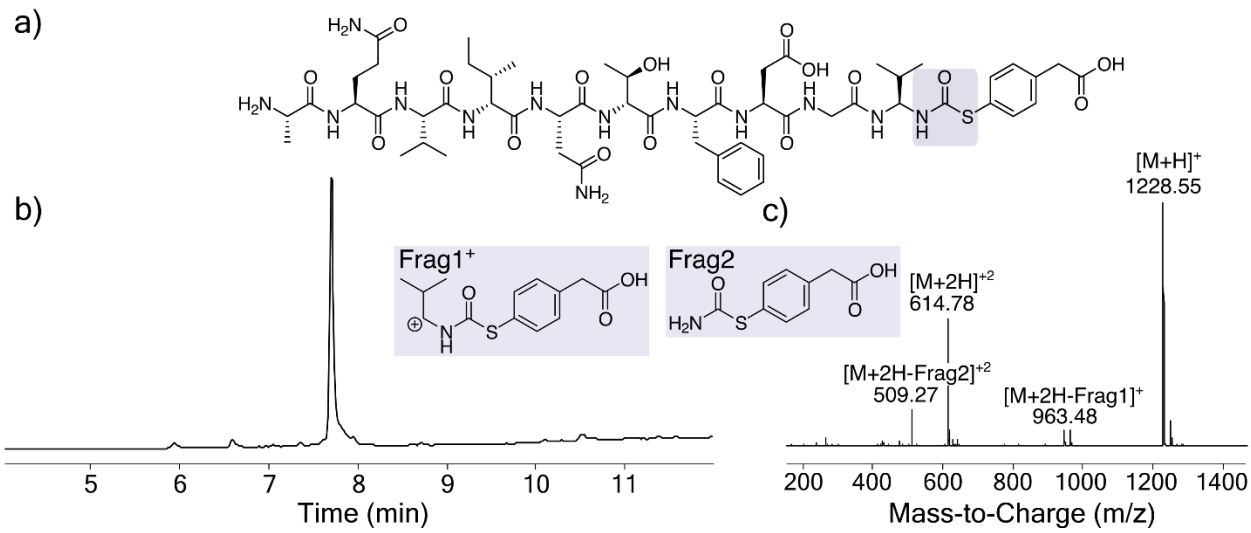


Figure S33. HPLC-MS (TIC) chromatogram (b) for the purified carbamothioate peptide H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-NH-COS-C_6H_4-CH_2-COOH with the peptide structure shown in a). c) MS of the compound with major peaks assigned. ESI-induced peptide fragment Frag1 is postulated to exist in a linear form as drawn. Calc. monoisotopic mass = 1227.56 Da.

5.4.8. H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$ + methanol

10.3 mg peptide was dissolved in 3.11 mL oxidation buffer, oxidized with 311 µL of 200 mM $NaNO_2$, and diluted into 72 mL of water/methanol (1: 19, v/v, 0.1% TFA added). After 2.1 mg of citric acid was added to the reaction mixture, the excess of acetonitrile was evaporated and the remaining solution was diluted with 95% A/5% B to a total of 40 mL. RP-HPLC purification followed by lyophilization yielded 6.9 mg (66% yield) of the desired semicarbazide as a white powder (*Fig. S34*).

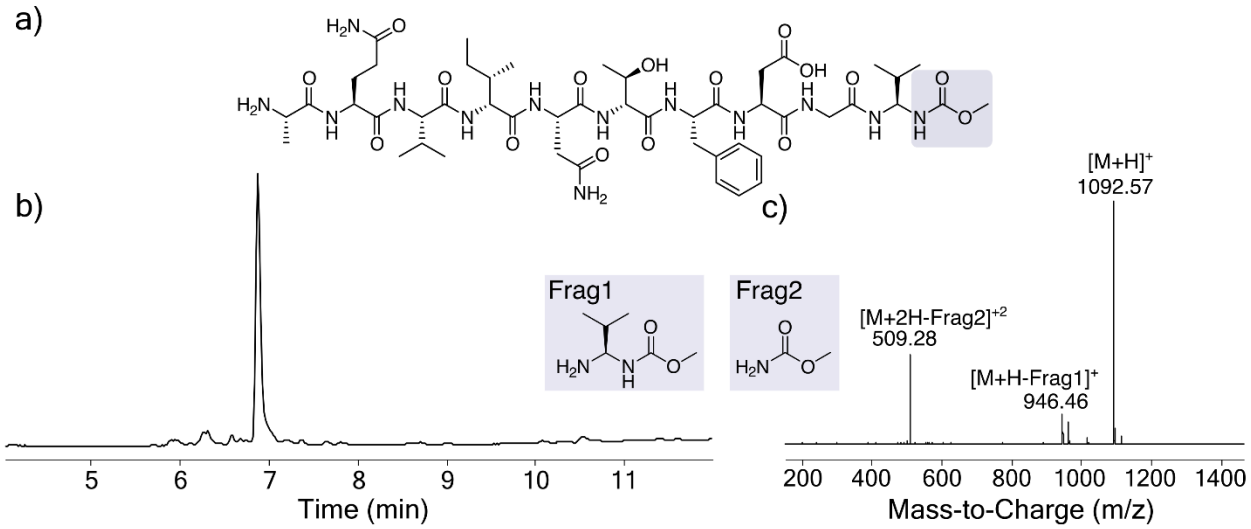


Figure S34. HPLC-MS (TIC) chromatogram (b) for the purified semicarbazide peptide H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-NH-COOMe with the peptide structure shown in a). c) MS of the compound with major peaks assigned. ESI-induced peptide fragment Frag1 is postulated to exist in a linear form as drawn. Calc. monoisotopic mass = 1091.56 Da.

5.4.9. H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$ + ethanol

1.0 mg peptide was dissolved in 0.30 mL oxidation buffer, oxidized with 30 µL of 200 mM $NaNO_2$, and diluted into 7.0 mL of water/ethanol (1: 19, v/v, 0.1% TFA added). The purification of the carbamate peptide was not attempted; RP-HPLC yield as estimated from the TIC chromatogram integration was 70%.

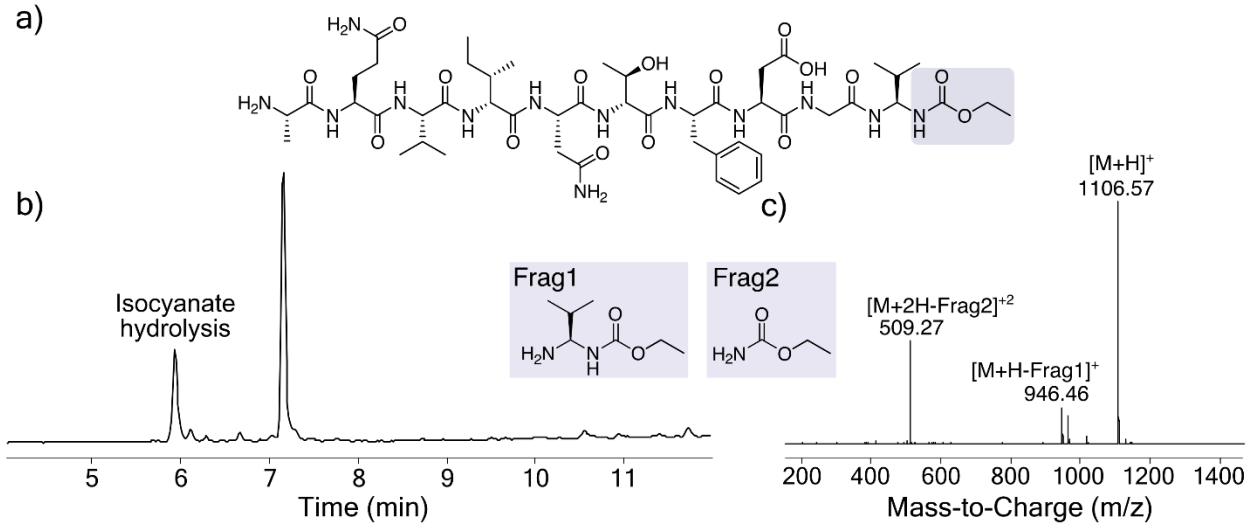


Figure S35. HPLC-MS (TIC) chromatogram (b) for the purified carbamate peptide H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-NH-COOEt with the peptide structure shown in a). c) MS of the compound with major peaks assigned. ESI-induced peptide fragment Frag1 is postulated to exist in a linear form as drawn. Calc. monoisotopic mass = 1105.57 Da.

5.5. Establishing the optimum nucleophile concentration for different peptide isocyanates

Correct nucleophile concentrations are crucial for the success of the conjugation. Here, we sought to compare optimum nucleophile concentrations for different peptides. To accomplish this we compared the MPAA conjugation results for two peptide hydrazides bearing different C-terminal residues: H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$ and H_2N-Ala-Arg-Val-Ile-Asn-Thr-Phe-Asp-Gly-Tyr-$CONHNH_2$. MPAA was chosen as a nucleophile sensitive to changes in concentration, i.e. there is a very narrow window of MPAA concentrations where conjugation to the isocyanate is high yielding (~5-20 mM). For

comparison, conjugation to perfluorophenylhydrazine is high yielding in the range 5-200 mM. Reactions were performed under standard conditions described in section 5.3. Reaction with H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$ were performed at 17 °C; with H_2N-Ala-Arg-Val-Ile-Asn-Thr-Phe-Asp-Gly-Tyr-$CONHNH_2$ — at 57 °C. The results of HPLC-MS analysis are summarized in *Fig. S36.*

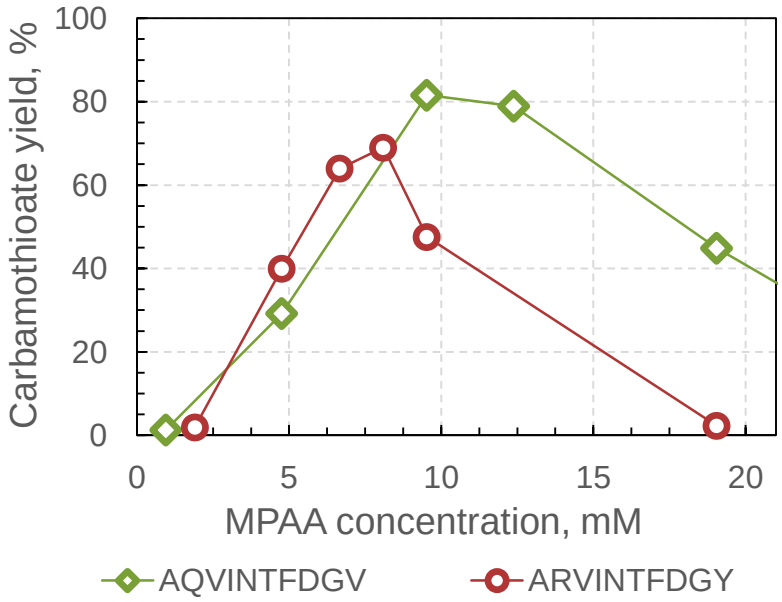


Figure S36. The plot of carbamothioate yield as a function of MPAA concentration for 2 peptides. Yields were obtained from HPLC-MS analysis; reactions were performed under standard (S. I. 5.3) conditions.

The MPAA concentration leading to the maximum carbamothioate yield was similar for the two peptides: 8.5 mM MPAA for H_2N-Ala-Arg-Val-Ile-Asn-Thr-Phe-Asp-Gly-Tyr-$CONHNH_2$ versus 10 mM for H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val-$CONHNH_2$. Thus, we concluded that the optimum nucleophile concentration is similar for different peptide isocyanates.

5.6. Experiment to detect racemization during the conjugation

To confirm that the Curtius rearrangement and all other steps in the process do not lead to racemization of the C-terminal residue we compared conjugation products formed from two peptide diastereomers: H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val^L-$CONHNH_2$ and H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-Val^D-$CONHNH_2$. Both reactions were performed under standard conditions (section 5.3) using perfluorophenylhydrazine as a nucleophile at 17 °C. To

resolve diastereomeric semicarbazides (conjugation products) we performed HPLC-MS analysis with the following gradient: 5% B' in A' for 2 min, 5-65% B' in A' ramping linearly over 70 min, 65% B' in A' for 1 minute. Diastereomers were resolved on HPLC, and each reaction led to the formation of a single diastereomer (*Fig. S37*). We concluded that the whole process is indeed racemization-free.

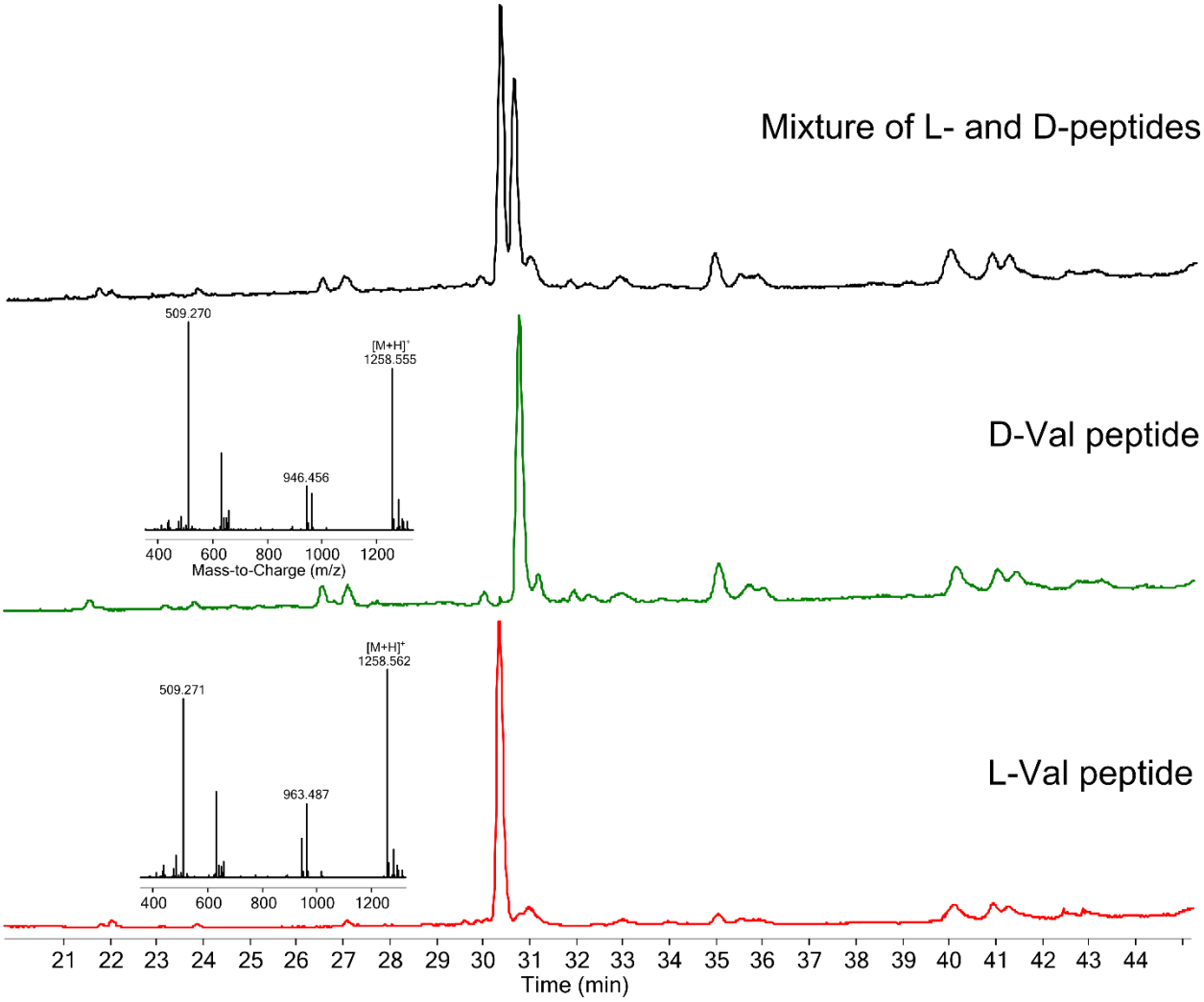


Figure S37. HPLC-MS (TIC) chromatograms for diastereomeric H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-^{L}Val-$CONHNH_2$ and H_2N-Ala-Gln-Val-Ile-Asn-Thr-Phe-Asp-Gly-^{D}Val-$CONHNH_2$ oxidation/conjugation reaction with MS insets for main products (semicarbazides).

5.7. Additional data for some C-terminal amino scan experiments

All reactions described in this section were performed under standard conditions (section 5.3) using 20 mM perfluorophenylhydrazine as a nucleophile at 57 °C, except in case of H_2N-Ala-His-Val-Ile-Asn-Thr-Phe-Asp-Gly-Thr-$CONHNH_2$ the reaction was at 17 °C.

Peptides with C-terminal Thr and Ser amino acids conjugated poorly to nucleophiles. The reactions resulted in formation of products 17 Da less massive than starting hydrazides. We postulated that these products stem from intramolecular cyclization of isocyanates to C-terminal side chains, resulting in formation of oxazolidin-2-ones. The HPLC-MS trace of one of these reactions is demonstrated in *Figure S38*.

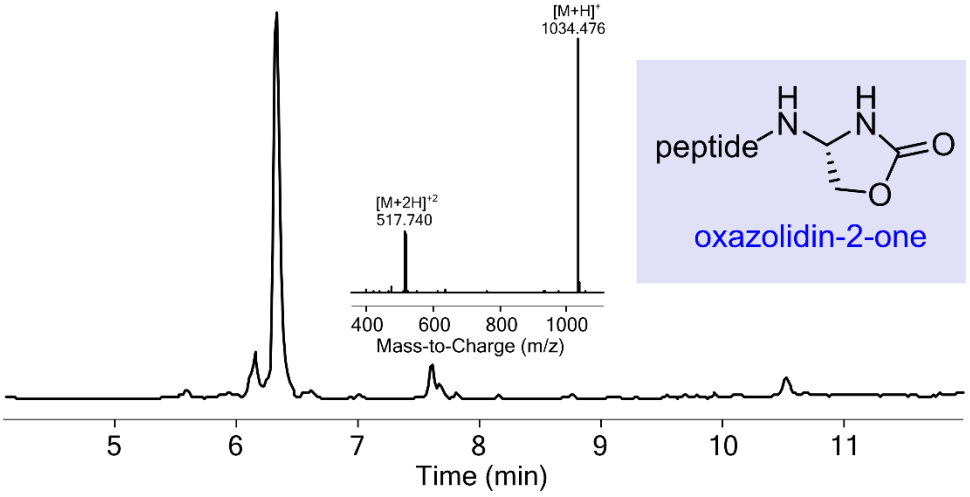


Figure S38. HPLC-MS (TIC) chromatogram for H_2N-Ala-Asn-Val-Ile-Asn-Thr-Phe-Asp-Gly-Ser-$CONHNH_2$ oxidation/conjugation reaction with MS inset for the main product.

Peptides with C-terminal His and Glu also conjugated poorly to perfluorophenylhydrazine. H_2N-Ala-Val-Val-Ile-Asn-Thr-Phe-Asp-Gly-His-$CONHNH_2$ and H_2N-Ala-Leu-Val-Ile-Asn-Thr-Phe-Asp-Gly-Glu-$CONHNH_2$ both showed significant products 14 Da less massive than the starting hydrazide, which were presumably the C-terminal carboxylate formed from hydrolysis of the azide (*Fig. S39, a*). We hypothesized that in both cases the azide cyclized to the C-terminal residue's side-chain, and then this cyclic intermediate was hydrolyzed to give the carboxylate (*Fig. S39, b*). To test this possibility, we performed oxidation/thioesterefication of H_2N-Ala-Leu-Val-Ile-Asn-Thr-Phe-Asp-Gly-Glu-$CONHNH_2$ under the standard conditions described in Fang, Li et al. *Angew. Chem. Int. Ed.* **2011**, 50, 7645 –7649. As shown in *Figure*

S39c, the reaction yielded the carboxylate and two thioester isomers. When H_2N-Ala-Val-Val-Ile-Asn-Thr-Phe-Asp-Gly-His-$CONHNH_2$ was subject to the same oxidation/thioesterefication conditions, a significant amount of the carboxylate was observed, which is not typical for peptides bearing other C-terminal amino acids. These data are consistent with our initial hypothesis that azides of His and Glu cyclize with their side-chains. Protection of glutamic acid as a cyclohexyl ester eliminated the formation of the carboxylate, as described in the next section.

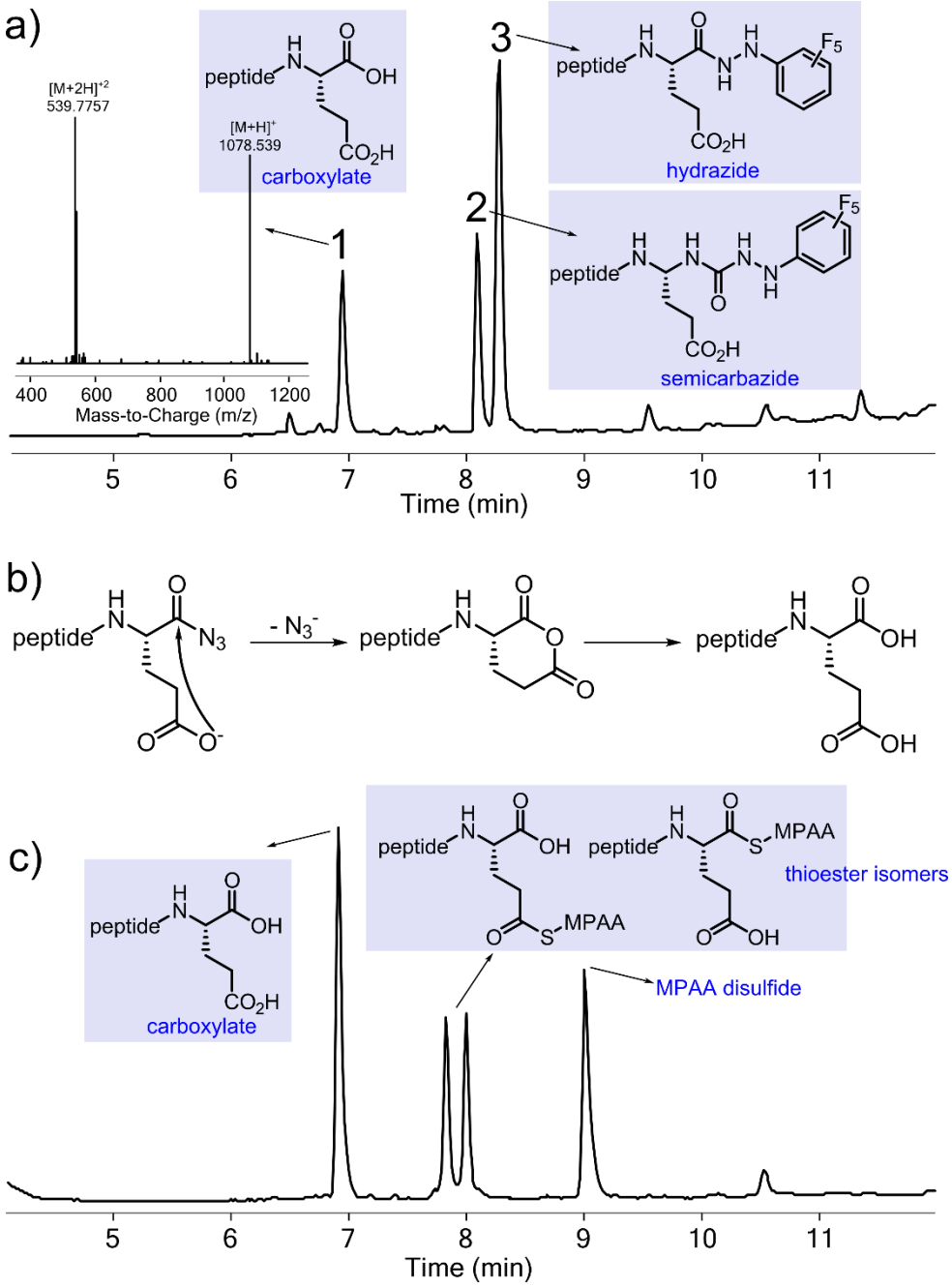


Figure S39. a) HPLC-MS (TIC) chromatogram for H_2N-Ala-Leu-Val-Ile-Asn-Thr-Phe-Asp-Gly-Glu-$CONHNH_2$ oxidation/conjugation reaction with MS inset for carboxylate. b) A proposed scheme for carboxylate formation. c) HPLC-MS (TIC) chromatogram for H_2N-Ala-Leu-Val-Ile-Asn-Thr-Phe-Asp-Gly-Glu-$CONHNH_2$ oxidation/thioesterefication reaction (200 mM MPAA).

5.8. Conjugation of perfluorophenylhydrazine to longer peptides

To study the applicability of our method to longer peptides, we performed oxidation/isocyanate conjugation for two barnase fragments: 26-mer H_2N-^{14}Cys-^{39}Val-$CONHNH_2$ and 37-mer H_2N-^{40}Cys(Acm)-^{76}Glu(Cy)-$CONHNH_2$. Full sequences in one letter codes were:

H_2N-CDYLQTYHKLPDNYITKSEAQALGWV-$CONHNH_2$ and

H_2N-(Acm)CSKGNLADVAPGKSIGGDIFSNREGKLPGKSGRTWRE(Cy)-$CONHNH_2$.

Perfluorophenylhydrazine was used as a nucleophile and reactions were performed under standard (section 5.3) conditions at 57 °C. HPLC-MS traces of the crude reaction mixtures are displayed in *Figure S40*.

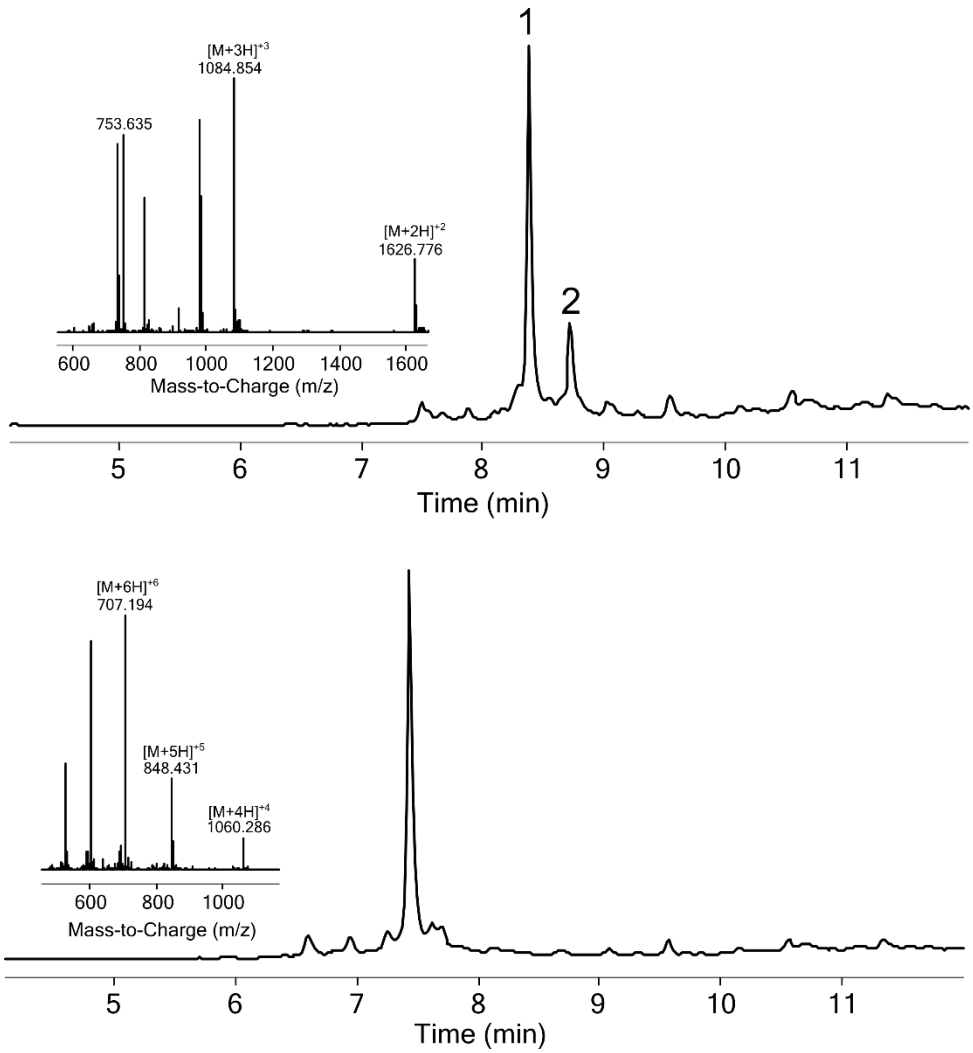


Figure S40. HPLC-MS (TIC) chromatograms for oxidation/conjugation reactions with the barnase fragments H_2N-^{14}Cys-^{39}Val-$CONHNH_2$ (26 residues, top) and H_2N-^{40}Cys(Acm)-^{76}Glu(Cy)-$CONHNH_2$ (37 residues, bottom). Top: peak 1 corresponds to the expected semicarbazide conjugation product with MS inset of the peak apex displayed on the left (calc. monoisotopic mass = 3250.5 Da). Peak 2 is an intermolecular disulfide of the conjugation product. Bottom: the main peak is the expected conjugation product with the MS inset of the peak apex on the left. (calc. monoisotopic mass = 4235.0 Da)

5.9. Sortase A mediated ligation of peptide hydrazide to LF_N-DTA

LF_N-DTA protein was expressed and purified as previously described (Ling, Policarpo *et al*. *J. Am. Chem. Soc.* **2012**, 134, 10749–10752). To produce a C-terminal protein hydrazide we employed a Sortase A* mediated ligation of H_2N-Gly_5-Leu-Glu-Ile-$CONHNH_2$ to LF_N-DTA-Leu-Pro-Ser-Thr-Gly_2-His_5. Ligation was carried out under the following conditions: a mixture of 300 µl Tris buffer (125 mM Tris, 100 mM NaCl, 1 mM $CaCl_2$, pH 7.5) containing 38 µM LF_N-DTA, 350 µM peptide hydrazide, 2.5 µM Sortase A*, and 80 µl Ni-NTA beads was

incubated at 17 °C for 30 minutes while nutating gently. Afterwards, the reaction mixture was centrifuged at 13000 rpm for 5 minutes. The beads were washed two times with Tris buffer (150 µl), and then with 40 mM imidazole in Tris buffer. The washes were combined with the reaction supernatant and concentrated over a Amicon Ultra-4 Ultracel-10K centrifugation filter. To remove excess peptide hydrazide and imidazole, the protein solution was subject to buffer exchange into Tris buffer. The reaction was analyzed with HPLC-MS to confirm complete formation of the desired C-terminal protein hydrazide. The MS spectrum of the resulting protein was deconvoluted using the maximum entropy algorithm to obtain deconvoluted the protein spectrum in *Figure S41*.

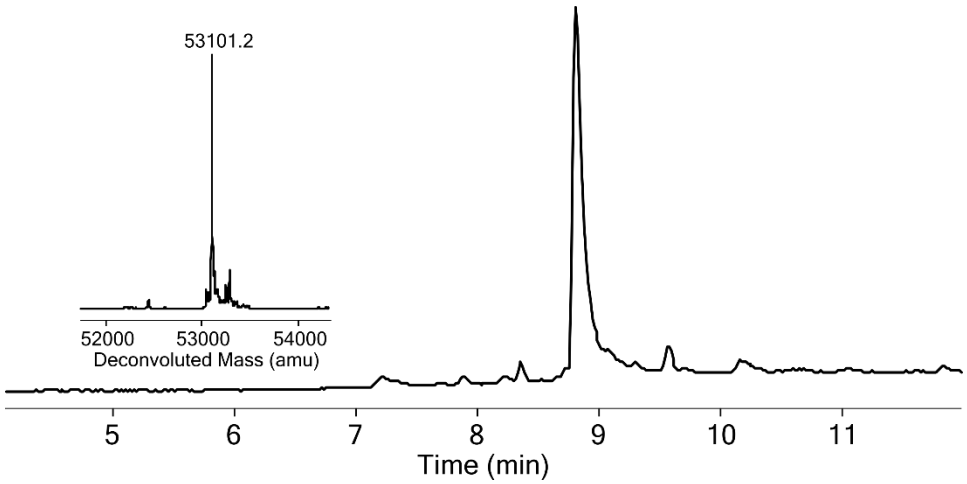


Figure S41. HPLC-MS (TIC) chromatogram for LF_N-DTA-Leu-Pro-Ser-Thr-Gly_5-Leu-Glu-Ile-$CONHNH_2$ obtained via Sortase A* mediated ligation of H_2N-Gly_5-Leu-Glu-Ile-$CONHNH_2$ to LF_N-DTA-Leu-Pro-Ser-Thr-Gly_2-His_5. The deconvoluted MS on the left was obtained using the maximum entropy algorithm.