Supporting Information for

Palladium Assisted Removal of a Solubilizing Tag from a Cys Side Chain to Facilitate Peptide and Protein Synthesis

Suman kumar Maity,[†] Guy Mann,[†] Muhammad Jbara, Shay Laps, Guy Kamnesky and Ashraf Brik*

Schulich Faculty of Chemistry, Technion-Israel Institute of Technology

Haifa, 3200008 (Israel)

E-mail: abrik@technion.ac.il

[†]S.K.M. and G.M. contributed equally

Table of Contents

General Methods	S2
List of the protected amino acids	S3
Synthesis of Phenylacetamidomethyl (Phacm) linker	S4-S7
Synthesis of the model peptide	S7-S8
Removal of the Phacm tag from the model peptide	S8
Synthesis of fragment 2	S9-S10
Synthesis of unmodified fragment 1	S11-S12
Synthesis of fragment 1 bearing the Arg tag	S12-S13
One-pot ligation, Phacm removal and desulfurization	S14-S15
One-pot ligation and desulfurization with unmodified fragment 1	S15-S16
Circular dichroism spectroscopy	S16
Inductively coupled plasma (ICP) experiment	S16
NMR and HRMS spectra	S17-S20

Experimental Section

Materials and methods

General methods:

SPPS was carried out manually in syringes, equipped with teflon filters, purchased from Torviq or by using an automated peptide synthesizer (CS336X, CSBIO). Analytical grade N,N-dimethylformamide (DMF) was purchased from Biolab. Resins were purchased from CreoSalus, protected amino acids were purchased from GL Biochem and activating [(2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate reagents (HBTU), hydroxybenzotriazole (HOBt), [(6-chlorobenzotriazol-1-yl)oxy-(dimethylamino)methylidene]-dimethylazanium hexafluorophosphate (HCTU), [Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate (HATU)] were purchased from Luxembourg Bio Technologies and GL Biochem. Chemicals were purchased from Aldrich, Strem Chemicals and Alfa Aesar. Analytical HPLC was performed on a Thermo instrument (Dionex Ultimate 3000) using analytical Xbridge (waters, BEH300 C4, 3.5 µm, 4.6 × 150 mm) and XSelect (waters, CSH C18, 3.5 μ m, 4.6 \times 150 mm) columns at flow rate of 1.2 mL/min. Preparative HPLC was performed on a Waters instrument using XSelect C18 (10μm, 19 × 250 mm) and semi preparative HPLC was performed on a Thermo Scientific instrument (Spectra System SCM1000) using XBridge BEH300 C4 (10 μ m, 150 \times 10 mm) and Jupiter C4 (10 μm, 300 Å, 250 × 10 mm) column, at flow rate of 15 and 4 mL/min respectively. All synthetic products were purified by HPLC and characterized by mass spectrometry using LCQ Fleet Ion Trap (Thermo Scientific). All calculated masses have been reported as an average isotope composition.

 1 H and 13 C NMR spectra were recorded using CDCl₃ and DMSO- d_6 as solvents. Chemical shifts were reported in δ units (ppm) with reference to TMS as an internal standard and J values are given in Hz. 1 H and 13 C-NMR spectra were recorded on a Bruker AMX-400 MHz. Flash column chromatography was carried out with silica gel (220–440 mesh). The reactions were carried out in oven-dried glassware under nitrogen. Analytical thin-layer chromatography (TLC) was performed on pre-coated plates (0.25 mm, silica gel 60 F254) and monitored by UV.

Buffer A: 0.1% TFA in water; buffer B: 0.1% TFA in acetonitrile.

List of the protected amino acids used in peptides synthesis: Fmoc-Gly-OH, Fmoc-Ala-OH, Fmoc-Val-OH, Fmoc-Leu-OH, Fmoc-Ile-OH, Fmoc-Phe-OH, Fmoc-His(Trt)-OH, Fmoc-Asn(Trt)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Lys(Boc)-OH, Fmoc-Tyr(tBu)-OH, Fmoc-Ser(tBu)-OH, Fmoc-Thr(tBu)-OH, Fmoc-Asp(OtBu)-OH, Fmoc-Glu(OtBu)-OH, Boc-Cys(Trt)-OH, Fmoc-Nle-OH, Boc-Ser(tBu)-OH, Boc-Ala-OH, Fmoc-Val-Thr(ψMe,MePro)-OH, Fmoc-Ile-Thr(ψMe,MePro)-OH, Fmoc-Glu(tBu)-Thr(ψMe,MePro)-OH, Fmoc-Ile-Ser(ψMe,MePro)-OH.

Synthesis of Alloc-Phenylacetamidomethyl (Phacm) linker (1):

Synthesis of compound 3:

$$H_2N$$
 allyl chloroformate, iPr_2NEt , CH_2Cl_2 Alloc N H 3

To a solution of 3.25 g (24.6 mmol) of 4-aminophenylacetonitrile (2) in 30 mL of dry dichloromethane (DCM), 5.13 mL (29.5 mmol) of N, N-diisopropylethylamine (DIEA) and 3.14 mL (29.5 mmol) of allylchloroformate were added dropwise and the reaction mixture was stirred for 1 h at room temperature under nitrogen atmosphere. The completion of the reaction was monitored by TLC. After completion of the reaction, the solvent was evaporated and the residue was dissolved in ethylacetate followed by washing with 5 % HCL and water. The combined organic layers were dried over Na₂SO₄, filtered and solvent was removed under vacuum to afford the compound **3** as a brown solid (5.26 g, 99% yield). The compound was used for next step without purification. 1 H NMR (400 MHz, CDCl₃, δ ppm): 3.69 (s, 2H), 4.66 (d, J = 5.6 Hz, 2H), 5.24-5.27 (m, 1H), 5.33-5.37 (m, 1H), 5.90-6.00 (m, 1H), 6.95 (s, 1H), 7.23 (d, J = 8 Hz, 2H), 7.40 (d, J = 8 Hz, 2H); 13 C NMR (100 MHz, CDCl₃ δ ppm): 23.10, 66.07, 118.14, 118.46, 119.32, 124.75, 128.74, 132.41, 137.92, 153.35. HRMS exact mass calculated for $C_{12}H_{13}N_{2}O_{2}$ ([M+H] $^{+}$): 217.0977 Da, observed: 217.0975 Da.

Synthesis of compound 4:

Alloc N 3
$$\frac{35\% \text{ HCl}}{\text{H}_2\text{O}}$$
 Alloc N 4

In a round bottom flask equipped with a reflux condenser, 5.26 g (24.3 mmol) of compound **3** and 15 mL of 35% HCl were placed and the reaction mixture was stirred vigorously at 40 °C for 1 h in an oil bath. Subsequently the oil bath was replaced with cold water bath and 15 mL water was added dropwise. When the total amount of water was added, the mixture was cooled externally with ice bath for 30 minutes. The precipitate appeared was dissolved in ethyl acetate and washed with 10% Na₂CO₃ solution (3 × 50 mL). The combined organic layers were dried over Na₂SO₄, filtered and solvent was removed under reduced pressure followed by purification using flash column chromatography (20% MeOH : CHCl₃) to afford the compound **4** as a brown solid (4.8 g, 85% yield). ¹H NMR (400 MHz, DMSO- d_6 , δ ppm): 3.29 (s, 2H), 4.59-4.60 (m, 2H), 5.21-5.24 (m, 1H), 5.33-5.38 (m, 1H), 5.94-6.01 (m, 1H), 6.84 (s, 1H), 7.15 (d, J = 8 Hz, 2H), 7.39 (d, J = 8 Hz, 2H); ¹³C NMR (100 MHz, DMSO- d_6 δ ppm): 41.56, 64.55, 117.51, 118.12, 129.32, 130.51, 133.34, 137.34, 153.21, 172.38. HRMS exact mass calculated for C₁₂H₁₄N₂O₃Na ([M+Na]⁺): 257.0902 Da, observed: 257.0932 Da.

Synthesis of compound 1:

To a solution of 1.3 g (5.62 mmol) of compound 4 and 37 mg (0.66 mmol) of KOH in 20 mL H2O/dioxane mixture (1:3 v/v) 0.72 mL (8.43 mmol) formaldehyde (35%) was added and the mixture was heated at 70 °C for 5 min and then at 25 °C for overnight. The pH of the reaction mixture was adjusted to 7 with 1 (N) HCl and the solvent was evaporated under reduced pressure. The residue was dissolved in acetone and dried over Na₂SO₄, filtered and solvent was evaporated under vacuum to give the compound 5. The crude solid was used for the next reaction without purification.

The resulting compound **5** and 1.54 g (4.49 mmol) of Fmoc-Cys-OH were dissolved in 20 mL H2O/dioxane mixture (1:3 v/v) and cooled in ice bath. 7.5 mL of trifluoromethanesulfonic acid (TFMSA) / trifluoroacetic acid (TFA) mixture (1:19 v/v) was added. The reaction mixture was stirred under nitrogen atmosphere for 6 h. The solvent was then evaporated and the residue was dissolved in ethyl acetate followed by washing with water. The combined organic layer was dried over Na₂SO₄, filtered and solvent was removed under reduced pressure followed by purification using flash column chromatography (50% ethyl acetate : hexane) to afford the compound **1** as a white solid (1.4 g, 42% yield over the two steps). H NMR (400 MHz, DMSO- d_6 , δ ppm): 2.77-2.83 (m, 1H), 3.00-3.04 (m, 1H), 3.38 (s, 2H), 4.18-4.31 (m, 6H), 4.59 (d, J = 5.2 Hz, 2H), 5.21-5.24 (m, 1H), 5.33-5.37 (m, 1H), 5.94-6.01 (m, 1H), 7.16 (d, J = 8 Hz, 2H), 7.32 (t,

 $J = 7.6 \text{ Hz}, 2\text{H}), 7.37-7.43 \text{ (m, 4H)}, 7.72-7.75 \text{ (m, 2H)}, 7.77 \text{ (s, 1H)}, 7.89 \text{ (d, } <math>J = 7.2 \text{ Hz}, 2\text{H}), 8.66 \text{ (t, } J = 6 \text{ Hz}, 1\text{H}), 9.67 \text{ (s, 1H)}, 12.86 \text{ (s, 1H)}. $^{13}\text{C NMR} \text{ (100 MHz, DMSO-} d_6 \delta \text{ ppm)}: 31.72, 40.46, 41.55, 46.58, 54.06, 64.56, 65.74, 117.51, 118.18, 120.09, 125.28, 127.06, 127.63, 129.29, 129.95, 133.33, 137.49, 140.70, 143.77, 153.20, 156.60, 170.50, 172.29. HRMS exact mass calculated for <math>C_{31}H_{32}N_3O_7S$ ([M+H]⁺): 590.1961 Da, observed: 590.1962 Da.

Synthesis of the model peptide 1 (ACLYRAGLYRAG) bearing the Alloc-Phacm linker:

The synthesis was carried out using Fmoc-SPPS on Rink amide resin (0.43 mmol/g, 0.1 mmol scale). The pre-swollen resin was treated with 20% piperidine in DMF containing 0.1 mmol HOBt (3-5-3 min) to remove the Fmoc-protecting group. Peptide synthesis was performed on an automated peptide synthesizer in presence of 4 equiv of amino acid, HCTU and 8 equiv of N,N'-diisopropylethylamine (DIEA). After coupling of the first ten amino acids, the Phacm linker (compound 1) was coupled manually using 2.5 equiv of amino acid, HATU and 5 equiv DIEA for 1.5 h. The last amino acid Ala was coupled in the Boc protected form. Subsequently, Alloc deprotection was carried out by using 0.2 equiv tetrakis(triphenylphosphine)palladium(0) (Pd(PPh3)4) and 20 equiv phenylsilane

in dry DCM for 1 h. Fmoc-Gly-OH and Boc-Arg(Pbf)-OH were coupled on the free amine respectively using HATU as a coupling agent.

Total deprotection and cleavage from the resin: The resin was washed with DMF, MeOH, DCM and dried. The peptide was cleaved using TFA:triisopropylsilane (TIS):water (95:2.5:2.5) cocktail for 2 h. The cleavage mixture was filtered and the combined filtrate was added dropwise to a 10-fold volume of cold ether and centrifuged. The precipitated crude peptide was dissolved in acetonitrile-water (1:1) and was further diluted to ~30% with water and lyophilized. The HPLC analysis was carried out on a C18 analytical column using a gradient of 0-60% B over 30 min.

Removal of the Alloc-Phacm tag from the model peptide:

1 mg of the peptide 1 was dissolve in 300 μ L of 6M Gn·HCl, 200 mM phosphate buffer and treated with a solution of 1 mg PdCl₂ in Gn·HCl buffer (50 μ L). The reaction mixture was incubated at 37 °C for 20 min. Progress of the reaction was monitored by taking aliquot from the reaction mixture and treated with a small amount of DTT (dithiothreitol), followed by centrifugation and injection of the supernatant into analytical HPLC (C4 column) using a gradient of 0-60% B over 30 min. HPLC and mass analysis showed the complete removal of the tag within 20 mins.

Synthesis of H4(1-37)-Nbz, fragment 2:

The synthesis was carried out according to the following scheme:

Fragment 2 was prepared by N-acylurea methodology using Alloc protected 3,4diaminobenzoic acid (Dbz) linker. Fmoc protected Dbz was coupled manually on preswollen Rink amide resin (0.2 mmol) using 4 equiv amino acid, 4 equiv HBTU/HOBT and 8 equiv DIEA for 1 h (2 cycles). Subsequently, the resin was washed with DMF and DCM (3 × 5 mL) and was treated with a solution of allyl chloroformate (50 eq) and DIEA (2 eq) in DCM for 12 h. Following the Fmoc removal, the Leu³⁷ was manually coupled using 10 equiv amino acid, 10 equiv HATU and 20 equiv DIEA for 1 h (2) cycles). The remaining amino acids were coupled on an automated peptide synthesizer. Unless specified otherwise, for each coupling cycle, amino acid and the coupling agent (HCTU) were used in 4-fold excess while DIEA was used in 8-fold excess to the initial loading of the resin. Pseudoproline dipeptide Fmoc-Ile-Thr($\psi^{\text{Me Me}}_{,}$ pro)-OH was coupled at position Ile^{29} -Thr³⁰ by using 2.5 equiv of dipeptide, 2.5 equiv of HATU and 5 equiv of DIEA for 1.5 h. The last amino acid was coupled as Boc-Ser(tBu)-OH. Analytical cleavage and HPLC analysis were performed after these couplings to ensure complete reactions.

On-resin Alloc removal and Dbz cyclization: The Alloc-protected resin was washed and swollen in DCM. A solution of PhSiH₃ (20 eq with respect to initial loading of the resin) and Pd(PPh₃)₄ (0.35 eq) in DCM was added to the resin and shaken for 1 h at 25 °C. After Alloc removal the resin was washed with DCM and then with DMF followed by the treatment with a solution of *p*-nitrophenyl chloroformate (5eq, 120 mg, for 0.1 mmol scale) in 4 mL DMF and shaken for 30 min at room temperature and washed with DMF. Following this, the resin was treated with a solution of DIEA (25eq, 0.45mL) in DMF (4 mL) and was shaken for 10 min and washed with DMF. The peptides were cleaved from resin, lyophilized and purified on a preparative C18 column using a gradient 20-60% B over 60 min to afford the pure peptide in ~30% yield.

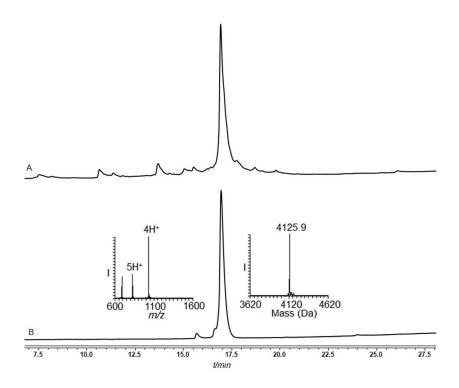


Figure S1: HPLC and mass analysis of (A) crude H4(1-37)-Nbz and (B) purified H4(1-37)-Nbz with the observed mass 4125.9 ± 0.5 Da, calcd 4124.6 Da (average isotopes).

Synthesis of unmodified fragment 1, Cys-H4(39-102):

$$H_2N$$
—

SPPS

Cleavage

 H_2N
 H_3
 $H_4(39-102)$

The synthesis of unmodified fragment 1 was carried out using Rink amide resin (0.43 mmol/g, 0.1 mmol scale). The pre-swelled resin was treated with 20% piperidine in DMF containing 0.1 mmol HOBt (3-5-3 min) to remove the Fmoc protecting group. Unless mentioned otherwise, all amino acids were coupled in an automated peptide synthesizer using 4 equiv of amino acid, 8 equiv of DIEA and 4 equiv of HCTU to the initial loading of the resin. Ala³⁸ was mutated to Cys to enable native chemical ligation. Pseudoproline dipeptides Fmoc-Ile-Ser(ψ^{Me}, Mepro)-OH, Fmoc-Glu(tBu)-Thr(ψMe,MePro)-OH and Fmoc-Val-Thr(ψ^{Me}, Mepro)-OH were coupled manually at positions Ile⁴⁶-Ser⁴⁷ and Glu⁵³-Thr⁵⁴, Val⁷⁰-Thr⁷¹ and Val⁸¹-Thr⁸² respectively, by using 2.5 equiv of dipeptides, 2.5 equiv of HATU and 5 equiv of DIEA for 1.5 h. Fmoc-Nle-OH was coupled at the position of Met⁸⁴ to avoid oxidation. Analytical cleavages and HPLC analyses were performed to ensure complete reactions.

Deprotection and cleavage from the resin: The peptide was deprotected and cleaved from resin as described above for model peptide and purified on a preparative C4 column using a gradient 20-60% B over 60 min to afford the unmodified fragment 1 in ~ 10% yield.

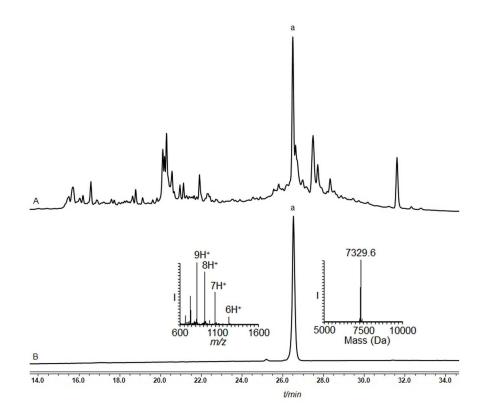


Figure S2: HPLC and mass analysis of (A) crude Cys-H4(39-102) and (B) purified Cys-H4(39-102) with the observed mass 7329.6 ± 0.8 Da, calcd 7329.5 Da (average isotopes).

Synthesis of fragment 1, Cys-H4(39-102) bearing the Arg tag:

The synthesis was carried out similarly as described for the unmodified fragment 1. The Phacm linker was coupled manually at the position of Ala⁶⁹ by using 2.5 equiv of amino acid, 2.5 equiv of HATU and 5 equiv of DIEA for 1.5 h. The N-terminal Cys was coupled as Boc-Cys(Trt)-OH. After completion of the sequence, Alloc deprotection was

carried out by using 0.2 equiv tetrakis(triphenylphosphine)palladium(0) (Pd(PPh3)₄) and 20 equiv phenylsilane in dry DCM for 1 h. Subsequently, three Arg residues were coupled on the free amine by using HATU as a coupling agent where the last Arg was coupled as Boc-Arg(Pbf)-OH.

Deprotection and cleavage from the resin: The peptide was deprotected and cleaved from resin as described above for model peptide and purified on a preparative C4 column using a gradient of 20-60% B over 60 min to afford the fragment 1 bearing the solubilizing tag in $\sim 20\%$ isolated yield.

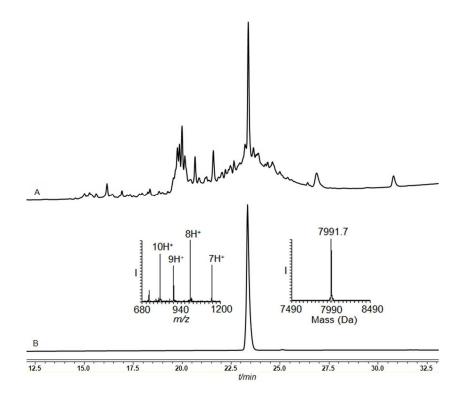


Figure S3: HPLC and mass analysis of (A) crude Cys-H4(39-102) containing the Arg tag and (B) purified Cys-H4(39-102) containing the Arg tag with the observed mass 7991.7 ± 0.8 Da, calcd 7992.1 Da (average isotopes).

One-pot ligation, Phacm removal and desulfurization for H4 synthesis:

In a typical ligation, fragment 1 (2.5 mg, 3.13×10^{-4} mmol) bearing Phacm linker and 2 $(1.29 \text{ mg}, 3.13 \times 10^{-4} \text{ mmol})$ were dissolved in argon purged 6 M Gn·HCl, 200 mM phosphate buffer (156 µL, 2 mM) containing 20 equiv of MPAA and 10 equiv of TCEP at pH 7.3. The reaction mixture was incubated at 37 °C for 4 h. Progress of the reaction was monitored by analytical HPLC using C4 column with a gradient of 0-60% buffer B over 30 min. After completion of the reaction, 50 equiv MgCl₂ (1.5 mg) dissolved in argon purged 6 M Gn·HCl, 200 mM phosphate buffer (25 uL) and added to the reaction mixture and the reaction mixture was incubated at 37 °C for 10 min. After, the reaction mixture was treated with 15 equiv PdCl₂ (0.83) dissolved in argon purged 6 M Gn·HCl, 200 mM phosphate buffer (25 μL). PdCl₂ was incubated at 37 °C for 30 min in Gn·HCl buffer for complete dissolution. The reaction mixture was incubated at 37 °C for 45 min for complete Phacm linker removal. Progress of the reaction was monitored by taking aliquot from the reaction mixture and treated with a small amount of DTT (dithiothreitol), followed by centrifugation and injection of the supernatant into analytical HPLC (C4 column) using a gradient of 0-60% B over 30 min. After completion of the reaction, 150 equiv of DTT was added and the resulting precipitate was separated by centrifugation. The precipitate was washed with Gn.HCl buffer twice and the combined supernatant solution was dialyzed in a Slide-A-Lyzer 3.5K dialysis cassette (Thermo scientific, 0.1-0.5 mL) in 6 M Gn·HCl, 200 mM phosphate buffer (500 mL) for overnight. Subsequently, the reaction mixture was desulfurized by treating with 21.5 mg TCEP, 10 mg (100 equiv) VA-044 and 30 µL tBuSH for 6 h. Progress of the reaction was monitored by analytical HPLC using C4 column with a gradient of 0-60% buffer B over 30 min. For semi-preparative HPLC, the same gradient was used to isolate the full H4 in \sim 20% isolated yield (0.7 mg).

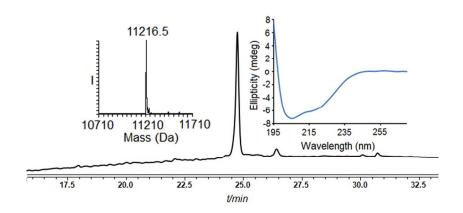


Figure S4: HPLC, mass-spectrometric and CD analysis of purified full length H4.

One-pot ligation and desulfurization using unmodified fragment 1:

In a typical ligation, unmodified fragment 1 (2.5 mg, 3.41×10^{-4} mmol) and fragment 2 (1.4 mg, 3.41×10^{-4} mmol) were dissolved in argon purged 6 M Gn·HCl, 200 mM phosphate buffer (170 µL, 2 mM) containing 20 eq of MPAA and 10 eq of TCEP at pH 7.3. The reaction mixture was incubated at 37 °C for 6 h. Progress of the reaction was monitored by analytical HPLC using C4 column with a gradient of 0-60% buffer B over 30 min. Following completion of the ligation, the reaction mixture was dialyzed in a Slide-A-Lyzer 3.5K dialysis cassette (Thermo scientific, 0.1-0.5 mL) in 6 M Gn·HCl, 200 mM phosphate buffer (500 mL) for overnight and the product was desulfurized by treating with 25 mg TCEP, 5.5 mg (50 equiv) VA-044 and 35 µL *t*BuSH (10% of volume) for 6 h. Progress of the reaction was monitored by analytical HPLC using C4 column with a gradient of 0-60% buffer B over 30 min.

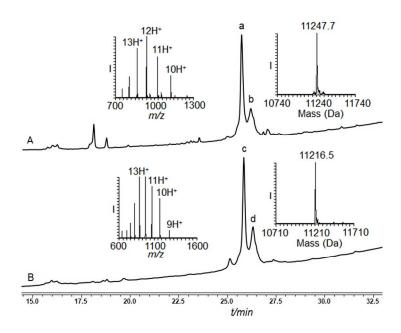


Figure S5: HPLC and mass analysis of (A) ligation of H4(1-37)-Nbz and unmodified Cys-H4(39-102) after 6 h; peak a corresponds to the ligation product with the observed mass 11247.7 ± 0.6 Da, calcd 11249.2 Da (average isotopes) and peak b corresponds to unreacted fragment 1; (B) desulfurization product with the observed mass 11216.5 ± 0.8 Da, calcd 11217.1 Da (average isotopes).

Circular dichroism spectroscopy: H4 was dissolved in 6M urea solution (5% of the total volume) and then further diluted with 20 mM Tris buffer containing 150 mM NaCl solution at pH 7.3. The exact concentration of the protein solution was determined by Pierce® BCA Protein Assay Kit, Thermo scientific. With this solution circular dichroism spectrum was recorded in a Chirascan (Applied Photophysics) instrument.

Inductively coupled plasma (ICP) experiment for detection of Pd content: H4 was dissolved in 6M urea solution (5% of the total volume) and then further diluted with 20 mM Tris buffer containing 150 mM NaCl solution at pH 7.3 and exact concentration of the protein solution was determined. The protein solution was subjected to ICP experiment in an iCAP 6000 SERIES ICP Spectrometer, Thermo Scientific. The amount of palladium content detected was 0.07% of the protein.



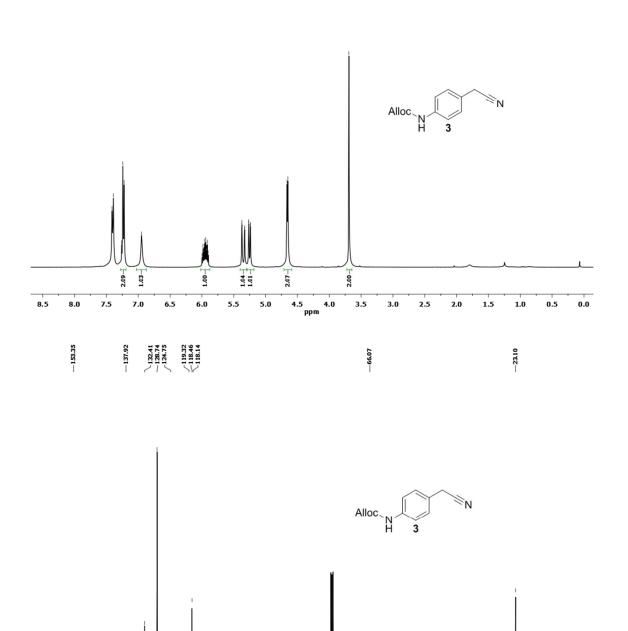


Figure S6: ¹H and ¹³C NMR spectra of compound 3 in CDCl3 (400 MHz and 100 MHz respectively).

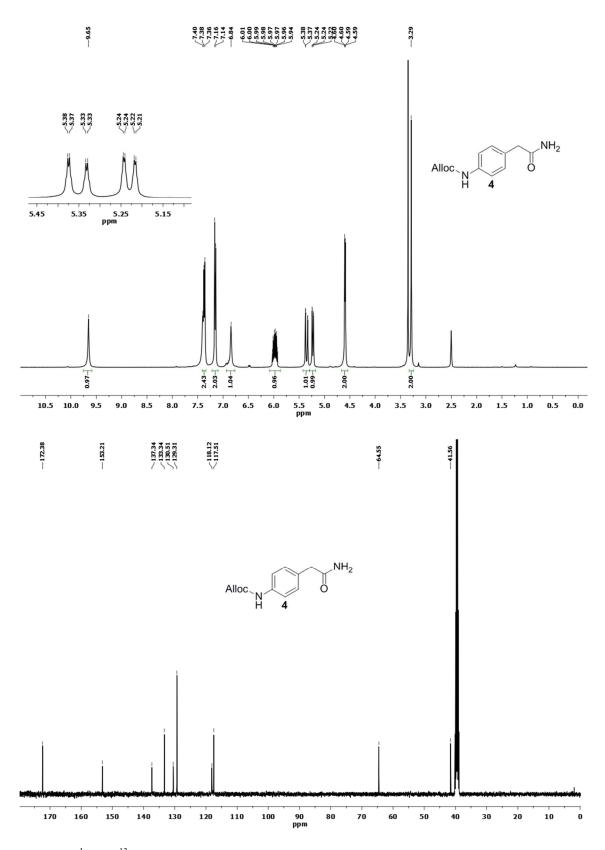


Figure S7: 1 H and 13 C NMR spectra of compound 4 in DMSO- d_{6} (400 MHz and 100 MHz respectively).

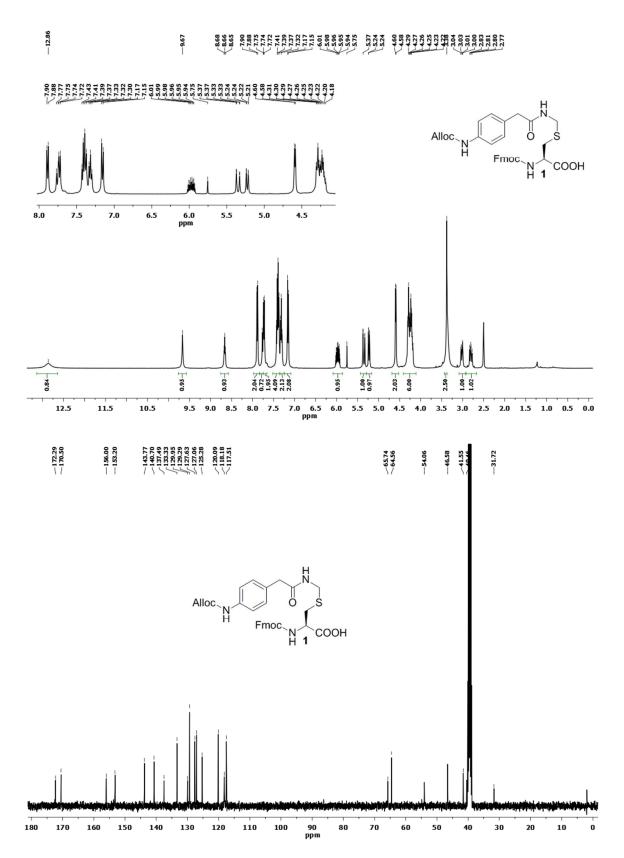


Figure S8: 1 H and 13 C NMR spectra of compound 1 in DMSO- d_{6} (400 MHz and 100 MHz respectively).

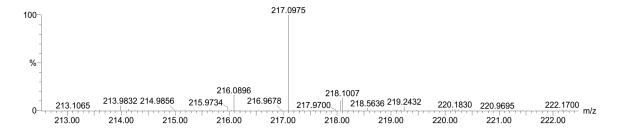


Figure S9: HRMS spectrum of compound 3.

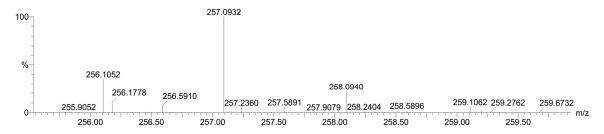


Figure S10: HRMS spectrum of compound 4.

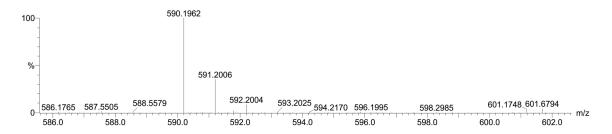


Figure S11: HRMS spectrum of compound 1.