

Supporting Information for

**A novel PEG-based solid support enables the synthesis of > 50  
amino acids peptide thioesters and the total synthesis of a  
functional SUMO-1 peptide conjugate**

**Emmanuelle Boll,<sup>a</sup> Hervé Drobecq,<sup>a</sup> Nathalie Ollivier,<sup>a</sup> Laurent Raibaut,<sup>a</sup> Rémi Desmet,<sup>a</sup>  
Jérôme Vicogne,<sup>a</sup> Oleg Melnyk<sup>\*a</sup>**

*UMR CNRS 8161, Pasteur Institute of Lille 59021 Lille, France.*

*oleg.melnyk@ibl.fr.*

## Table of content

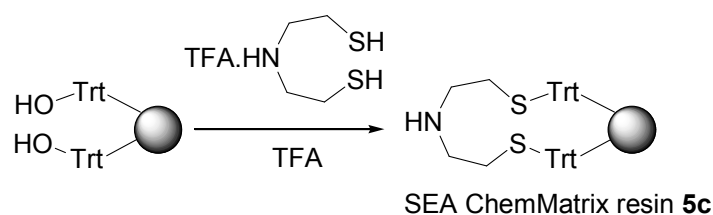
Synthesis and characterization of bis(2-sulfanylethyl)amino ChemMatrix resin (1 mmol scale).....	3
Synthesis of <i>bis</i> (2-sulfanylethyl)amino ChemMatrix resin 5c .....	3
Synthesis of <i>bis</i> (2-sulfanylethyl)amino Novasyn resin 5d .....	4
Synthesis of SUMO-1(2-97)-SEA <sup>off</sup> 11 and of Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH <sub>2</sub> 15 proteins .....	4
Synthesis of C-terminal SEA <sup>off</sup> peptides 7a and 7b.....	4
Coupling of the first amino acid residue to the SEA ChemMatrix resin.....	4
Automated peptide elongation step using Fmoc-SPPS .....	5
Deprotection and cleavage step .....	6
Purification of SEA <sup>off</sup> peptide segments 7a and 7b (Fig. S1-6).....	6
Conversion of SEA <sup>off</sup> peptide segment 7a into MPA thioester segment 10 (Fig. S7-9).....	13
Ligation of SUMO (2-50)-MPA thioester segment 10 with SUMO (51-97)-SEA <sup>off</sup> peptide segment 8b. Isolation of SUMO-1 (2-97)-SEA <sup>off</sup> protein 11 .....	16
Experimental procedure (Fig. S10).....	17
Characterization of SUMO-1 (2-97)-SEA <sup>off</sup> 11 (Fig. S11-14) .....	18
One-pot sequential ligation of three peptide segments for the synthesis of Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH <sub>2</sub> protein 15 (Fig. S15-20) .....	23
MALDI-TOF analysis of Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH <sub>2</sub> 15, in source fragmentation as a proof for the structure (Fig. S21-23).....	28
Biochemical characterization of Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH <sub>2</sub> protein 15 (Fig. S24-28) .....	31
Cleavage of the conjugate by Ulp1 enzyme .....	31
Characterization of Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH <sub>2</sub> protein 15 by circular dichroism .....	36
Synthesis of 20 kD peptide 9 .....	37
Synthesis of peptides 6d and 8 (Fig. S29-30).....	37
Ligation of peptides 6d and 8. Isolation of the 20 kD peptide 9 .....	39
Synthesis of peptide 7c using SEA NovaSyn resin .....	40

## Synthesis and characterization of bis(2-sulfanylethyl)amino ChemMatrix resin (1 mmol scale)

The bis(2-tritylsulfanylethyl)amine was prepared as described previously by reacting bis(2-chloroethyl)amine hydrochloride with triphenylmethyl mercaptan in the presence of DBU.<sup>1</sup> Both bis(2-chloroethyl)amine hydrochloride and triphenylmethyl mercaptan are commercially available.

Caution: The synthesis must be carried out in an efficient fume hood with the appropriate protection.

### Synthesis of bis(2-sulfanylethyl)amino ChemMatrix resin **5c**



Scheme S1

Trityl-OH ChemMatrix resin (2.22 g, 1 mmol, 0.45 mol/g) was conditioned in CHCl<sub>3</sub> by washing the resin three times during 2 min with the solvent (40mL) in a glass reactor equipped with a sintered glass and a tap in the bottom and a stopper in the top.

The bis(2-tritylsulfanylethyl)amine (207.30 mg, 0.333 mmol) was dissolved in a mixture of trifluoroacetic acid (TFA)/triisopropylsilane (TIS) : 97.5/2.5 by vol, 30 mL). The mixture was stirred 30 min at room temperature until which the solvent was removed in vacuo in a rotatory evaporator. Cyclohexane was added to the residue and evaporated to remove the residual TFA (twice, 2 × 50 mL). The residual white solid corresponding to dithiol **2** was immediately dissolved in CHCl<sub>3</sub> (20 mL) and added to the Trityl-OH ChemMatrix resin. Then, TFA (4 mL, 52 mmol) was added to the bead suspension which was shaken for 2 h (the color of the beads turns yellow due to the formation of trityl carbocations).

The reaction was quenched by addition of 1,5-lutidine (12.04 mL, 104 mmol, caution: this step is highly exothermic). The resin was shaken for 2 min and then filtered and washed

<sup>1</sup> N. Ollivier, J. Dheur, R. Mhidia, A. Blanpain and O. Melnyk, *Org. Lett.*, 2010, **12**, 5238-5241.

successively with  $\text{CHCl}_3$  ( $2 \times 2$  min) ; diisopropylethylamine (DIEA) 5% by vol in  $\text{CH}_2\text{Cl}_2$  (2 min) ;  $\text{CH}_2\text{Cl}_2$  ( $3 \times 2$  min), diethylether ( $3 \times 2$  min) and dried in vacuo.

The SEA ChemMatrix resin furnished a strongly positive Chloranyl assay and a negative Ellman assay (absence of free thiols).

SEA ChemMatrix resin and SEA PS resins were compared in an attempt to produce long SEA peptides derived from nerve growth factor (NGF). In particular SEA ChemMatrix resin enabled the successful synthesis of SEA peptide CTTTHTFVKALTMDGKQAAWRFIRIDTA-SEA (NGF 201-228, uniprot identifier P01138, 28 amino acids, 5.3% yield after RP-HPLC purification) which could not be detected in the crude produced using SEA PS resin. The synthesis was performed using an automated column peptide synthesizer as for SUMO-1 SEA peptides **7a,b**.

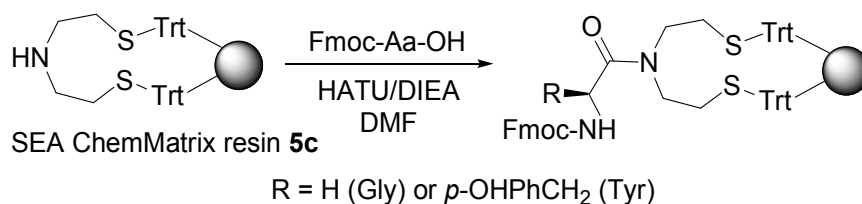
## **Synthesis of *bis*(2-sulfanylethyl)amino Novasyn resin 5d**

SEA Novasyn resin was prepared and characterized using the procedure described above for the preparation of SEA ChemMatrix resin. In this case, we used *bis*(2-tritylsulfanylethyl)amine (248.4 mg, 0.4 mmol) and Novasyn TGT alcohol resin (3.8 g, 1 mmol, 0.26 mmol/g) as starting materials. The loading of the resin was 0.11 mmol/g for Fmoc-Gly-SEA Novasyn TGT resin.

## **Synthesis of SUMO-1(2-97)-SEA<sup>off</sup> 11 and of Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub> 15 proteins**

### **Synthesis of C-terminal SEA<sup>off</sup> peptides 7a and 7b.**

#### **Coupling of the first amino acid residue to the SEA ChemMatrix resin**



Scheme S2

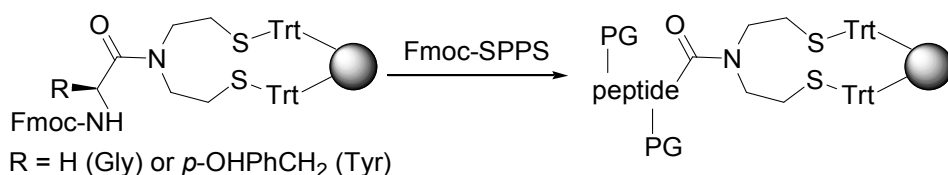
SEA ChemMatrix resin **5c** (200 mg) was conditioned in DMF by washing the resin three times during 2 min with the solvent (2 mL) in a glass reactor equipped with a sintered glass and a tap in the bottom and a stopper in the top.

Fmoc-Aa-OH (89.19 mg for Fmoc-Gly-OH, 137.88 mg for Fmoc-Tyr(*t*Bu)-OH, 0.3 mmol) was dissolved in DMF (1 mL). *N*-[(dimethylamino)-1*H*-1,2,3-triazolo-[4,5-*b*]pyridin-1-ylmethylene]-*N*-methylethaniminium hexafluorophosphate *N*-oxide (HATU, 112.93 mg, 0.297 mmol) was dissolved in the minimal volume of DMF and added to the amino acid solution. Finally, DIEA (104.6  $\mu$ L, 0.6 mmol) was added to the above solution to start the activation of the amino acid (1 min). Then, the activated amino acid solution was added to the resin which was shaken during 1 h 30 min at room temperature. The resin was subsequently washed with DMF (5  $\times$  2 min) and controlled using the chloranil assay which was negative.<sup>2</sup> The resin was subsequently acetylated with a mixture of acetic anhydride/DIEA/DMF : 10/5/85 by vol (2  $\times$  10 mL, 2 min and then 20 min) before being washed successively with DMF (3  $\times$  2 min), CH<sub>2</sub>Cl<sub>2</sub> (3  $\times$  2 min) and diethylether (3  $\times$  2 min). The resin was then dried in vacuo during 1 h 30 min.

The loading of the resin was determined by UV quantification at 290 nm of the dibenzofulvene-piperidine adduct formed by treating aliquots of the resin with piperidine. We found 0.20 mmol/g for Fmoc-Gly-SEA ChemMatrix resin and 0.17 mmol/g for Fmoc-Tyr(*t*Bu)-SEA ChemMatrix resin.

### Automated peptide elongation step using Fmoc-SPPS

<sup>2</sup> The coupling proceeds efficiently for unhindered amino acids such as Fmoc-Gly-OH or Fmoc-Tyr(*t*Bu)-OH used in this study. More crowded amino acids may require a second coupling step. In any case, we recommend to cap the resin afterwards to avoid the presence of free secondary amino groups on the resin and thus the potential formation of deletion side-products in the subsequent steps.



**sequence for SUMO-1 (2-50) :** SDQEAKPSTE DLGDKKEGEY IKLKVIGQDS SEIHFVKVMT  
THLKKLKESY

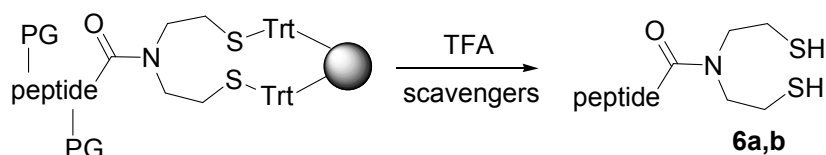
**sequence for SUMO-1 (51-97) :** C(*S*tBu)QRQGVPMNS LRFLFEGQRI ADNHTPKELG  
MEEEDVIEVY QEQTGG

PG : standard protecting groups for Fmoc-SPPS

Scheme S3

The peptide elongation step was performed using an automated column peptide synthesizer using standard Fmoc-SPPS protocols (0.25  $\mu\text{mol}$  scale). The amino acids were activated using HATU/DIEA in DMF.

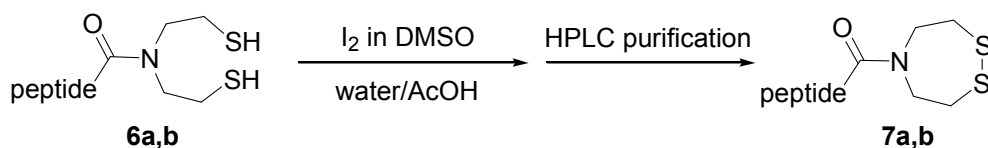
### Deprotection and cleavage step



Scheme S4

The peptidyl resins (25  $\mu\text{mol}$  scale) were deprotected and cleaved in a mixture of TFA/triisopropylsilane (TIS)/dimethylsulfide (DMS)/thioanisole/H<sub>2</sub>O : 90/2.5/2.5/2.5/2.5 by vol (10 mL) during 2 h 30 min. The crude peptides were precipitated in an ice-cold mixture of Et<sub>2</sub>O/*n*-heptane : 1/1 by vol to give 80 mg (44%) of SEA<sup>on</sup> peptide segment **6a** and 62 mg (40%) of SEA<sup>on</sup> peptide segment **6b**.

### Purification of SEA<sup>off</sup> peptide segments 7a and 7b (Fig. S1-6)



Scheme S5

The crude SEA<sup>on</sup> peptide segment (25  $\mu$ mol) was dissolved in AcOH/H<sub>2</sub>O: 1/4 by vol (final peptide concentration 0.5 mM). Iodine solution (200 mM in DMSO, 50  $\mu$ mol, 250  $\mu$ L) was added in one portion. After 30 s, dithiothreitol (DTT, 65 mM in water, 50  $\mu$ mol, 775  $\mu$ L) was added to quench the excess of iodine. Once DTT was added, the mixture was immediately purified by reversed-phase HPLC to give the corresponding SEA<sup>off</sup> peptide segment.

Preparative HPLC conditions for SEA<sup>off</sup> peptide segment **7a,b**: C18 Nucleosil 300 Å, eluent A water containing 0.05% of TFA, eluent B CH<sub>3</sub>CN/water : 4/1 by vol containing 0.05% of TFA, gradient: 0-18% B in 1 min, then 18-45% B in 45 min, flow rate 25 mL/min, UV detection at 215 nm.

Oxidation and HPLC purification of the crude SEA<sup>on</sup> peptide segment **6a** (30.4 mg) furnished SEA<sup>off</sup> peptide segment **7a** (9.7 mg, 6%).

Oxidation and HPLC purification of the crude SEA<sup>on</sup> peptide segment **6b** (62 mg) furnished SEA<sup>off</sup> peptide segment **7b** (4.0 mg, 3%).

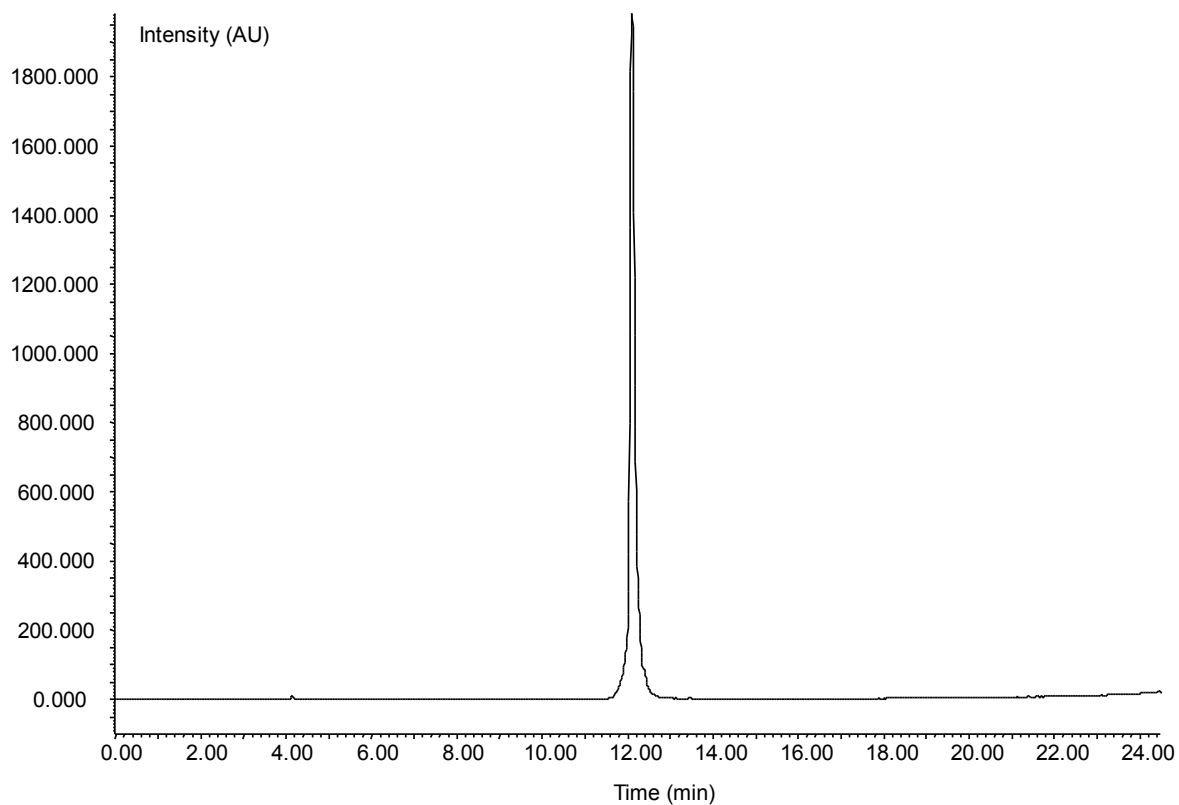
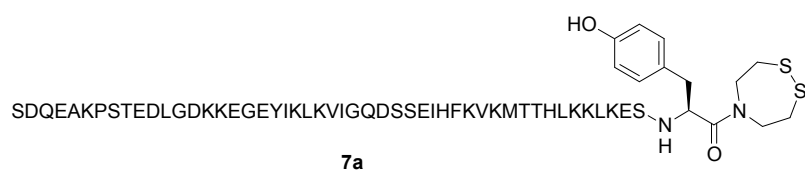


Fig. S1. LC-MS analysis of SEA<sup>off</sup> peptide segment **7a**. HPLC trace (light scattering detection).



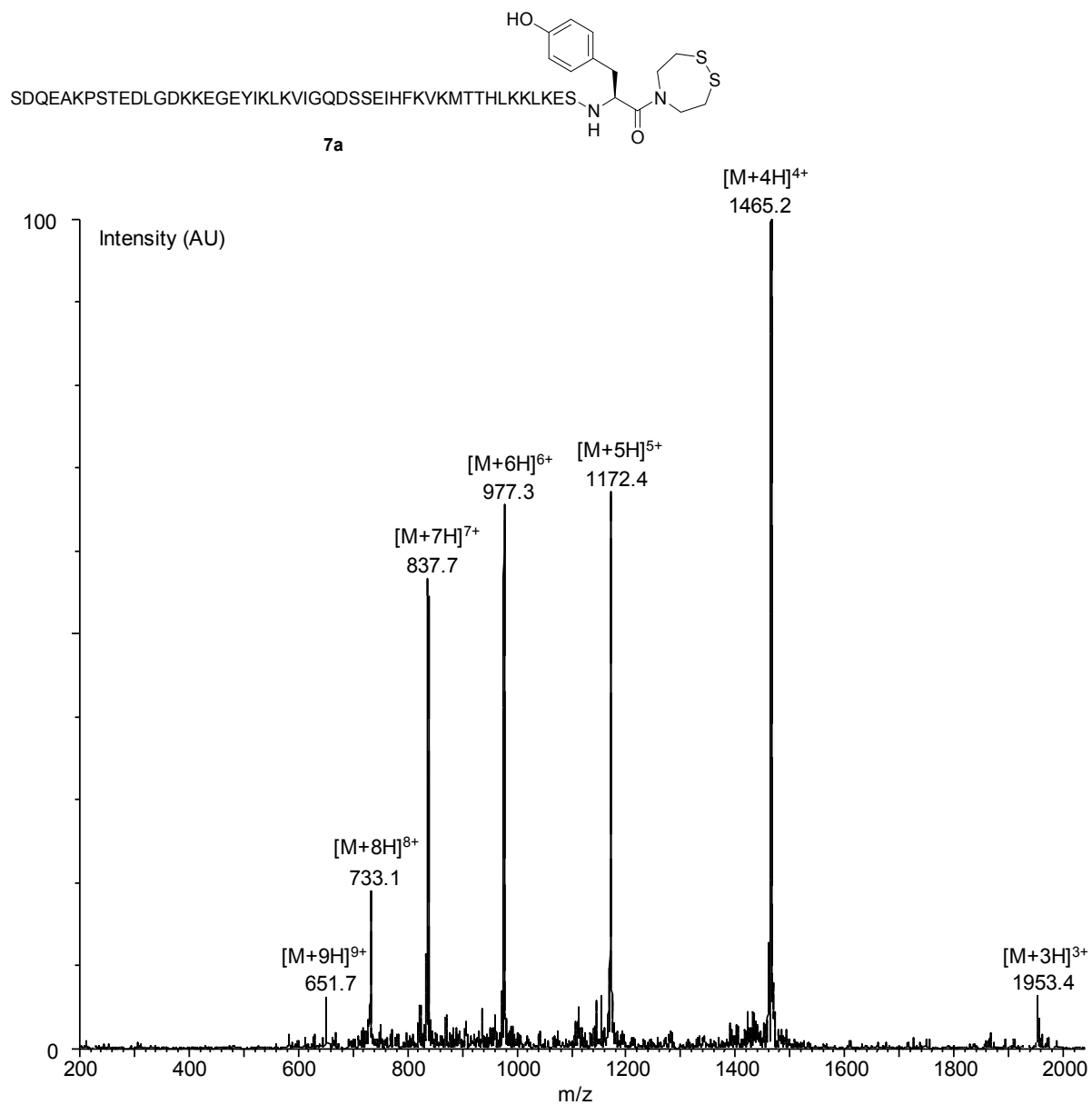


Figure S2. MS trace for SEA<sup>off</sup> peptide segment **7a**. Calculated for M (average mass) 5855.7, observed 5856.9 after deconvolution.

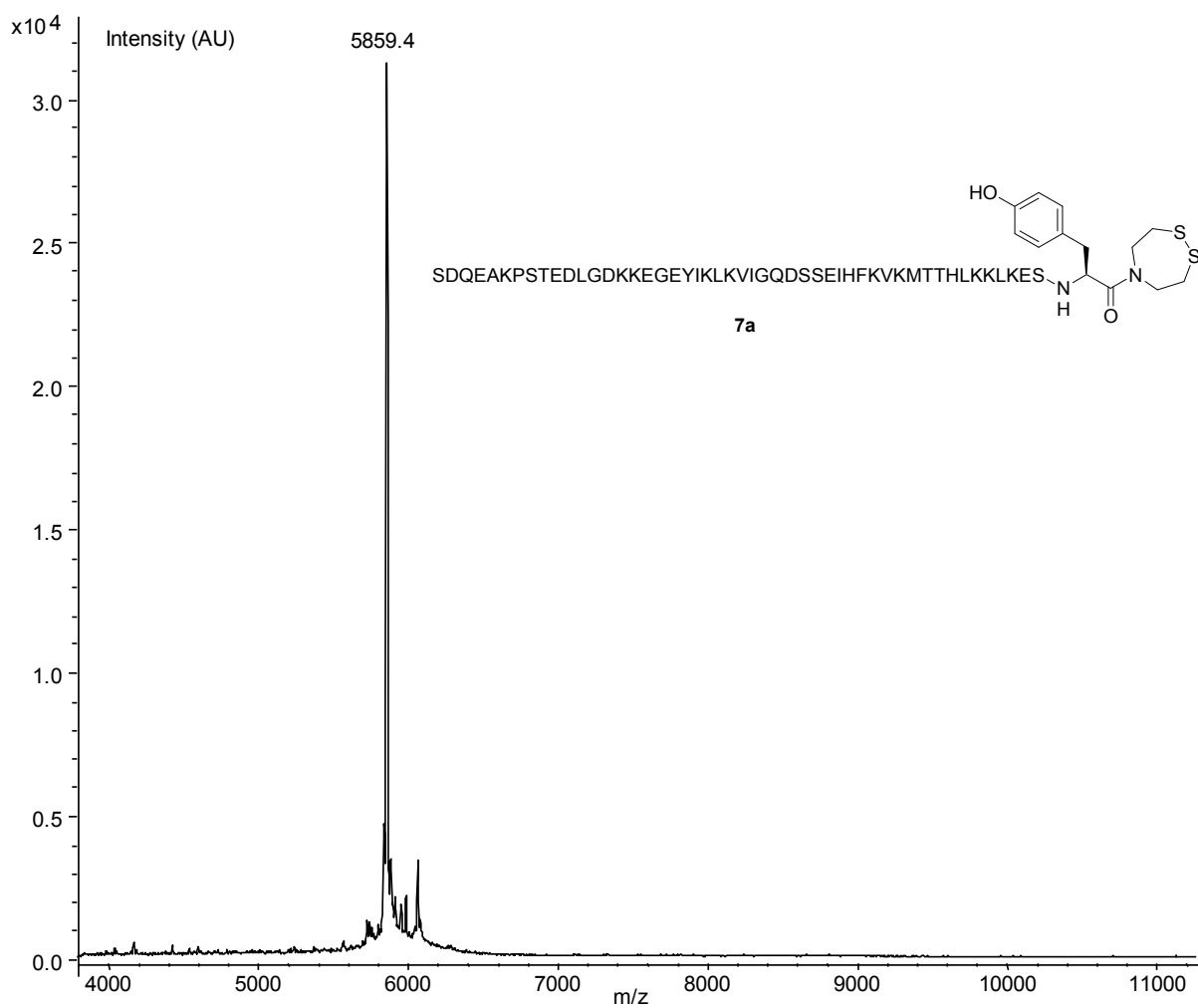


Fig. S3. MALDI-TOF analysis of SEA<sup>off</sup> peptide segment **7a**. Calculated for  $[M+H]^+$  (average mass) 5856.7.0,  $[M+H]^+$  observed 5859.4

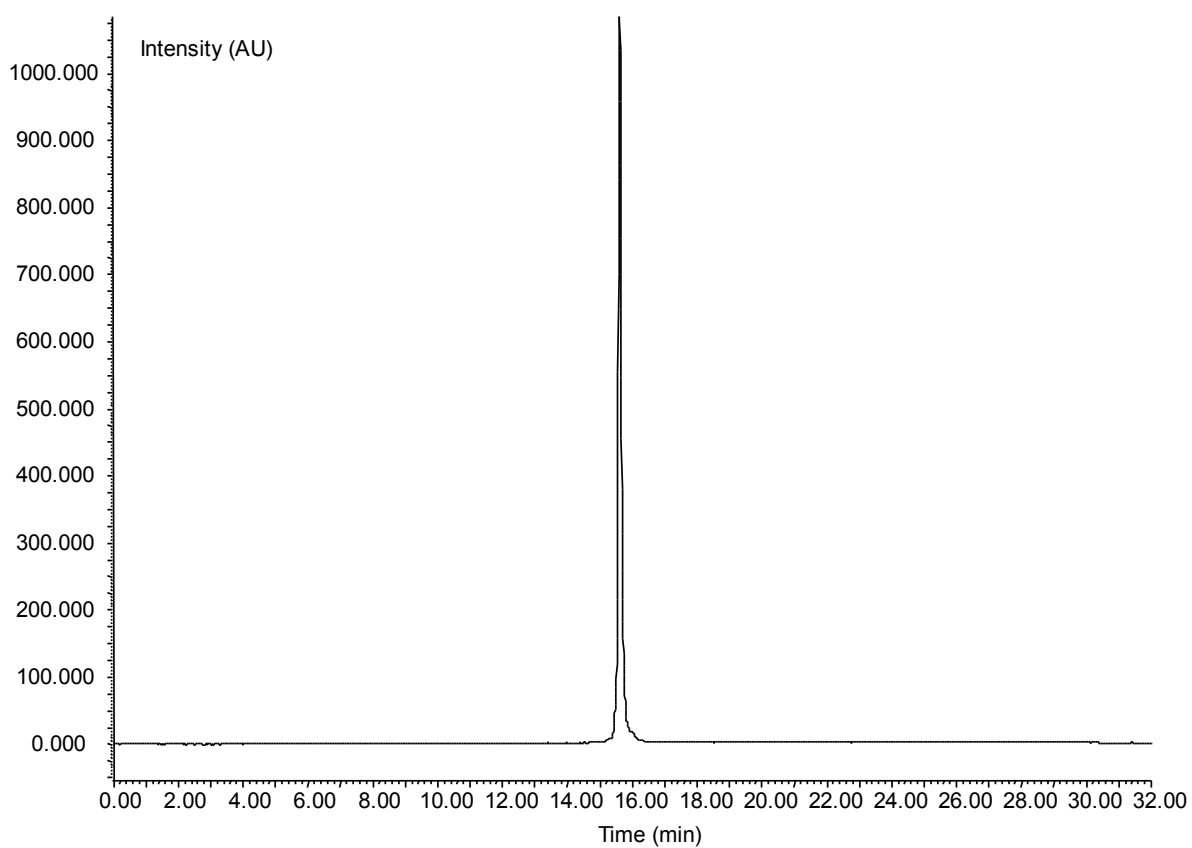


Fig. S4. LC-MS analysis of SEA<sup>off</sup> peptide segment **7b**. HPLC trace (light scattering detection).

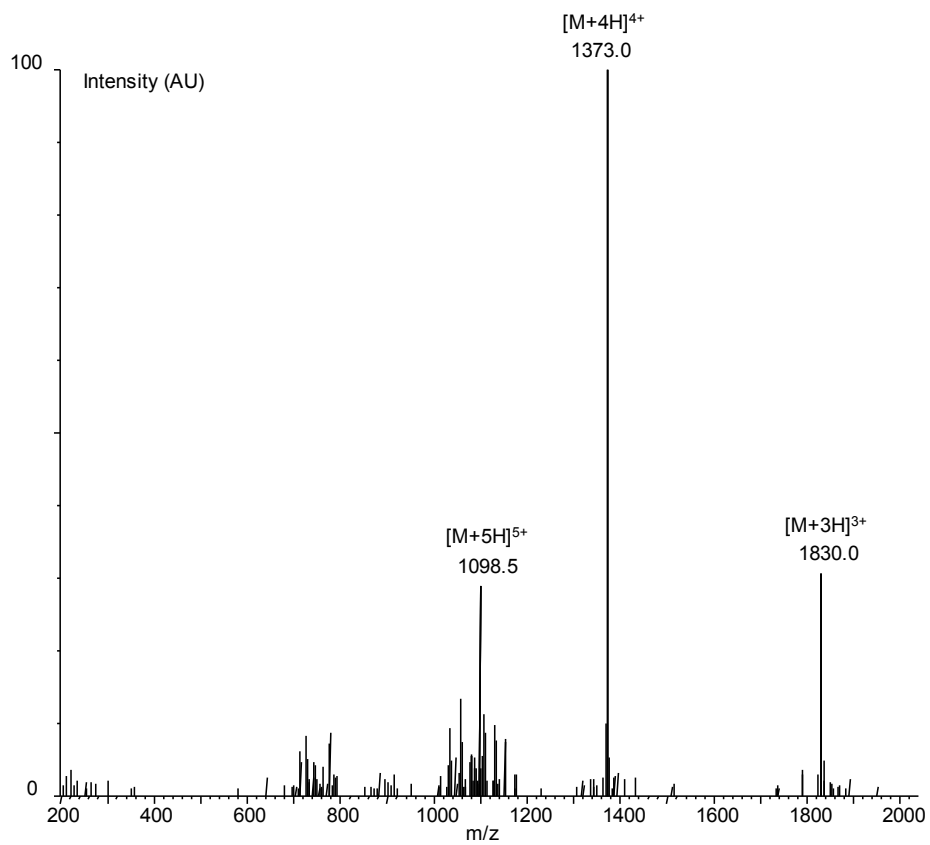


Fig. S5. MS trace for SEA<sup>off</sup> peptide segment **7b**. Calculated for M (average mass) 5486.2, observed after deconvolution 5487.5.

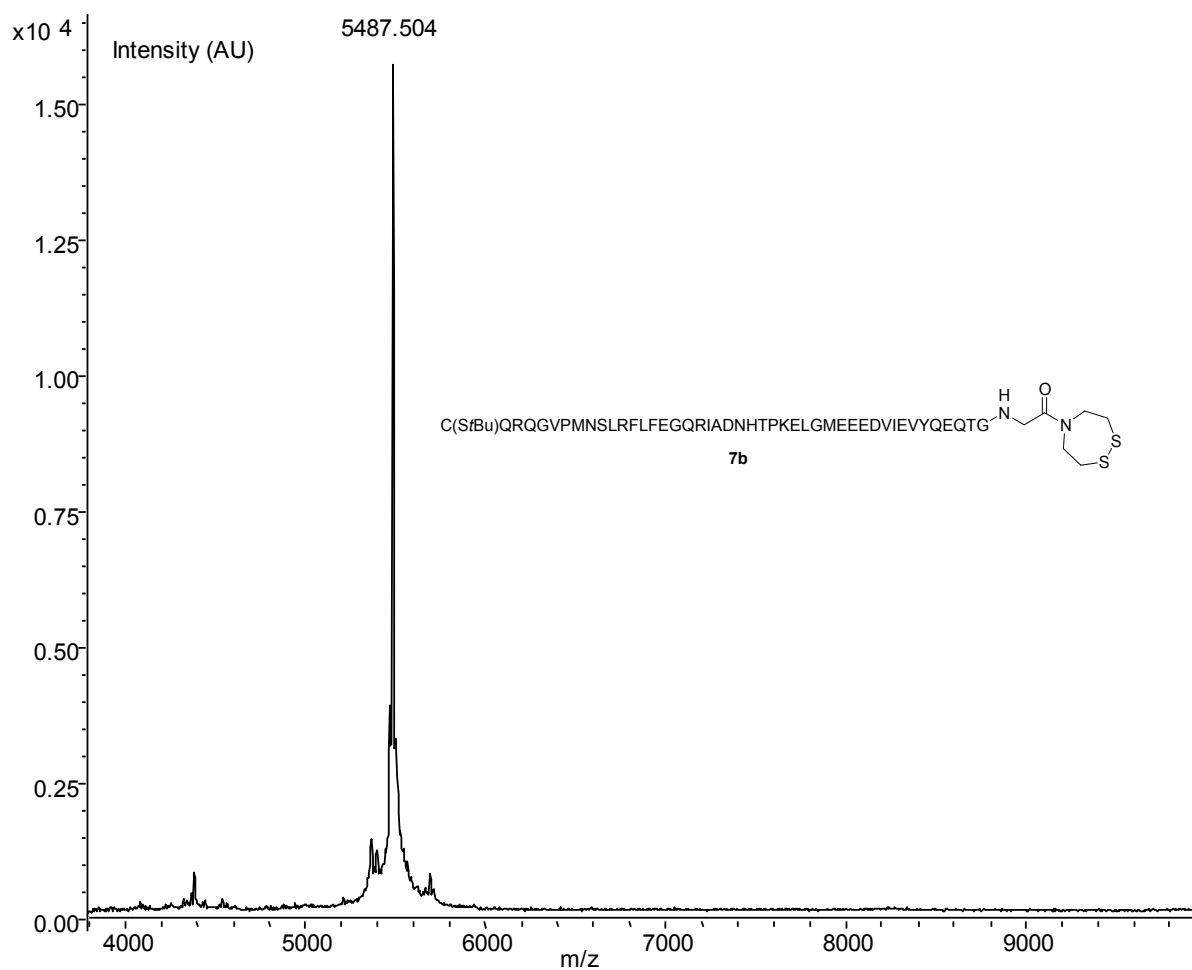
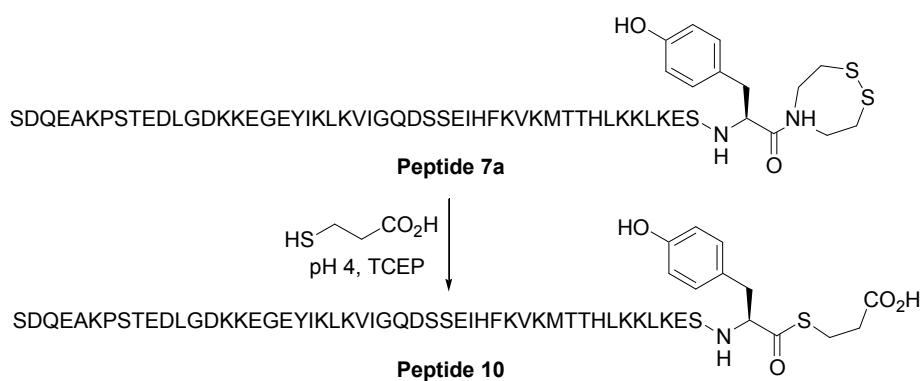


Fig. S6. MALDI-TOF analysis of SEA<sup>off</sup> peptide segment **7b**. Calculated for [M+H]<sup>+</sup> (average mass) 5487.2, observed 5487.5.

### Conversion of SEA<sup>off</sup> peptide segment **7a** into MPA thioester segment **10** (Fig. S7-9)

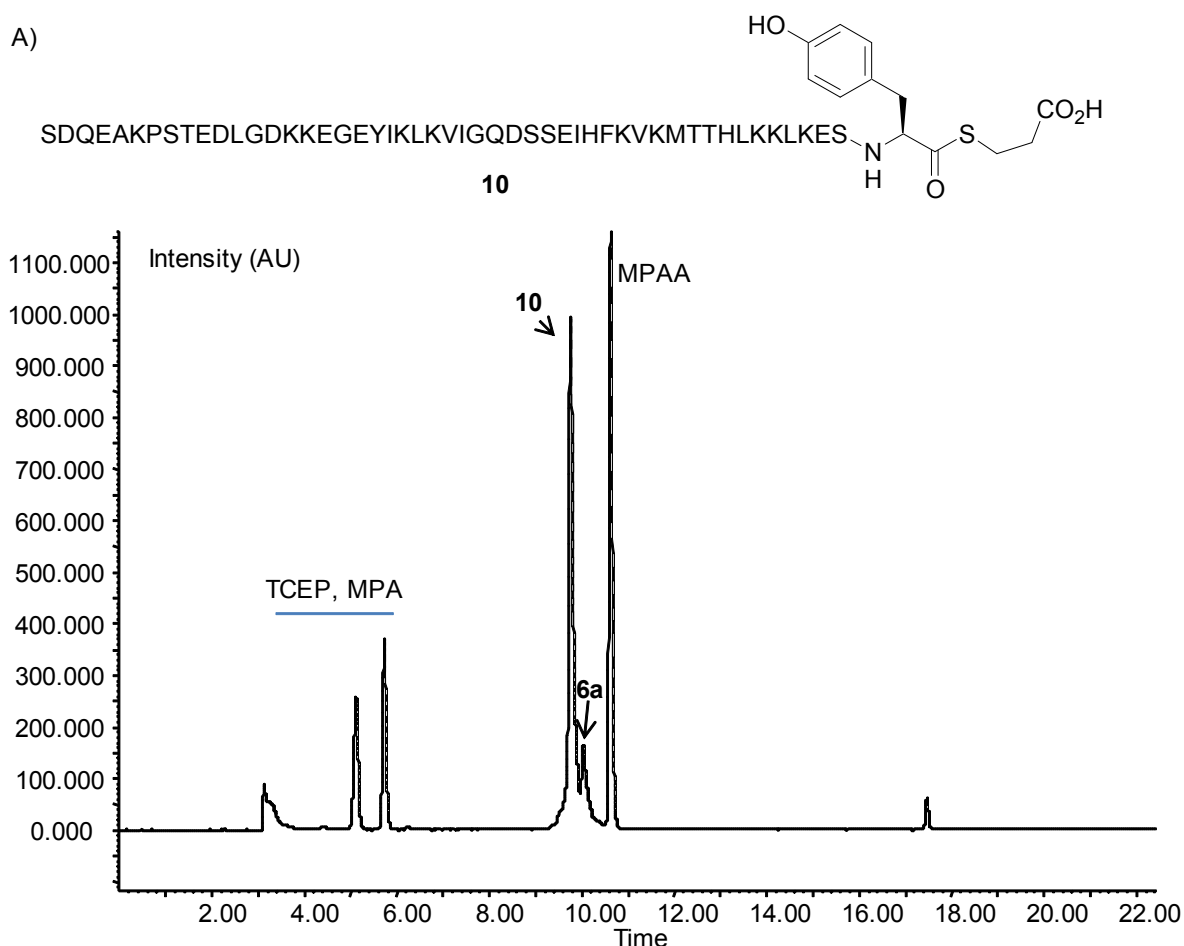


Scheme S6

The reaction was carried out under nitrogen atmosphere. Tris(2-carboxylethyl)phosphine hydrochloride (TCEP.HCl, 114.7 mg, 80.0 mmol) was dissolved in 0.2 M pH 7.3 sodium phosphate buffer (5 mL). SEA<sup>off</sup> peptide segment **7b** (10.90 mg) was dissolved in the above solution (4.46 mL). Then, 3-mercaptopropionic acid (MPA, 222  $\mu$ L) was added and the pH of the reaction mixture was adjusted to 3.8 by addition of 5 M NaOH. The reaction was agitated at 37  $^{\circ}$ C for 24 h and then acidified with 10% aqueous TFA (510  $\mu$ L). The mixture was extracted with diethylether (3 $\times$  10 mL) to remove the excess of MPA and immediately purified by RP-HPLC using a C18 Nucleosil column (120  $\text{\AA}$ , 5  $\mu$ m) to give 3.0 mg (28%) of MPA thioester segment **10**.

Gradient used for the HPLC purification: eluent A water containing 0.05% of TFA, eluent B CH<sub>3</sub>CN/water : 4/1 by vol containing 0.05% of TFA, gradient: 0-10% B in 5 min, then 10-100% B in 150 min, flow rate 6 mL/min, UV detection at 215 nm.

A)



B)

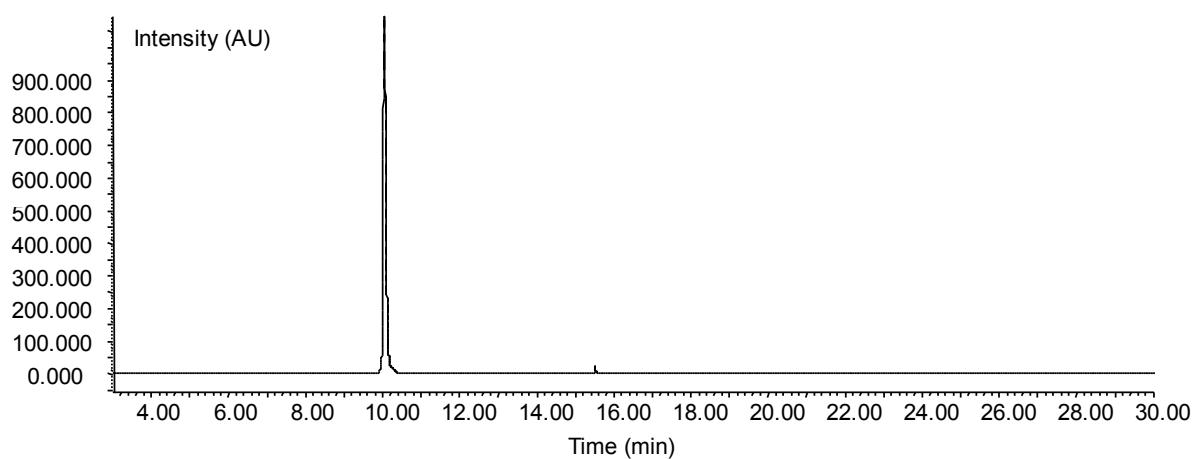


Fig. S7. LC-MS analysis of MPA peptide thioester **10**. A) HPLC trace (light scattering detection) of the crude exchange mixture, B) HPLC trace (light scattering detection) of the purified product.

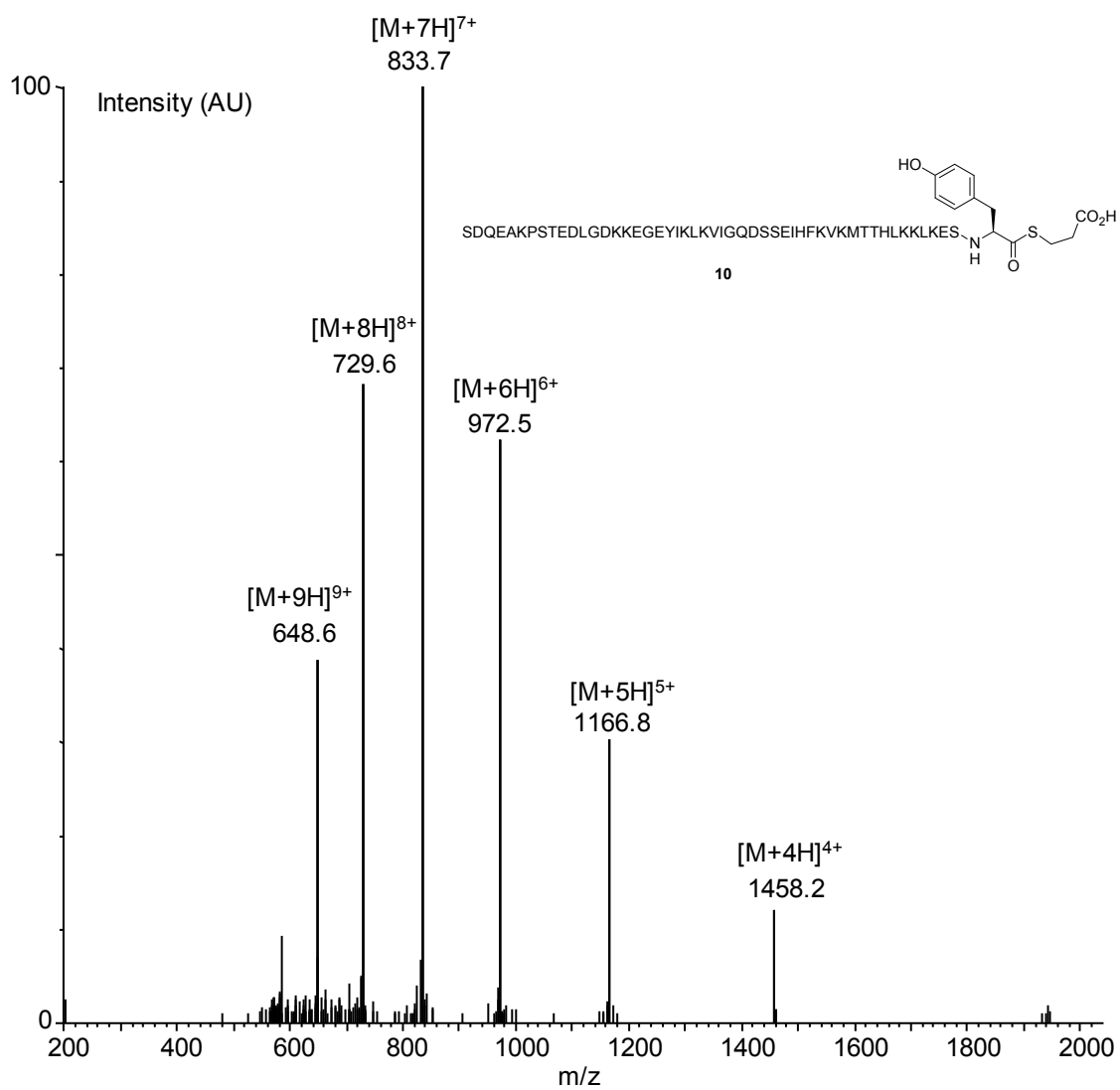


Fig. S8. LC-MS analysis of MPA peptide thioester **10**. MS trace for the purifier product (Fig S7B), calculated for M (average mass) 5826.6, observed 5828.9 after deconvolution.

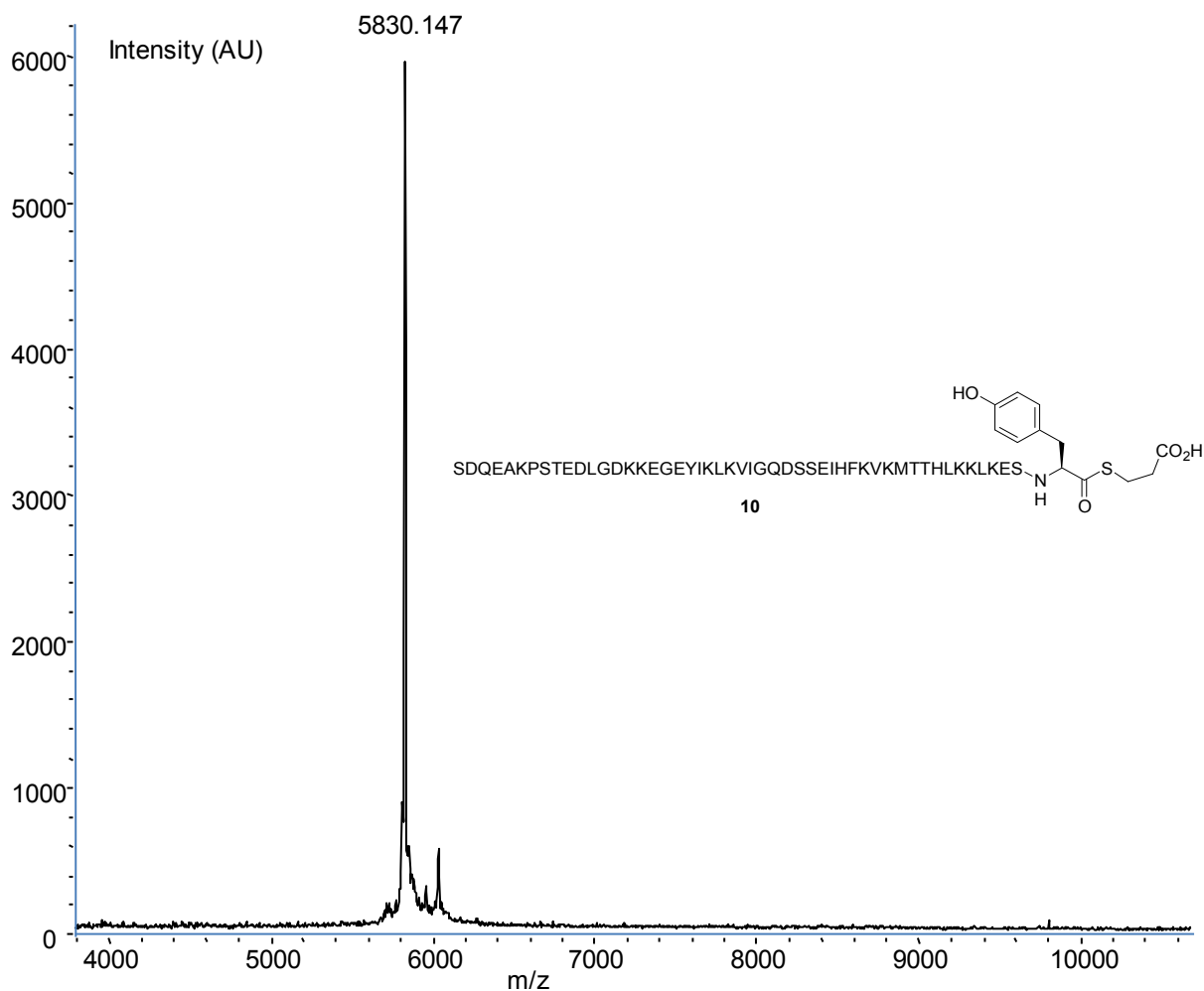
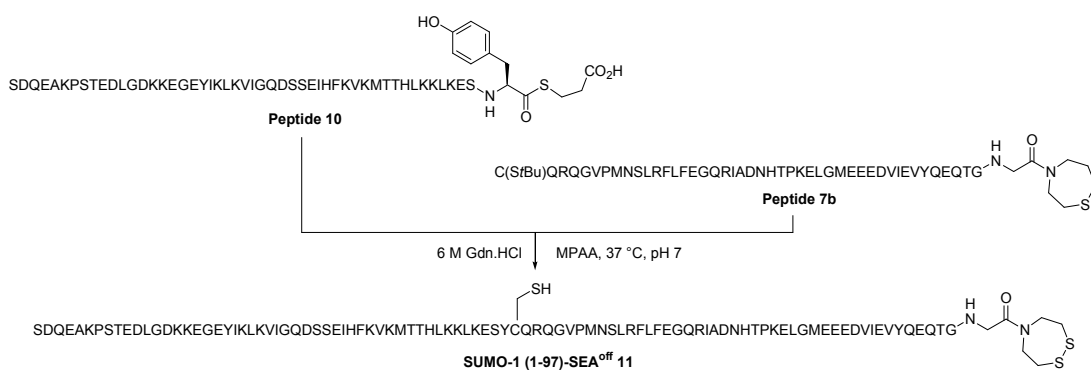


Fig. S9. MALDI-TOF analysis of MPA peptide thioester **10**. Calculated for  $[M+H]^+$  (average mass) 5827.6, observed 5830.1.

**Ligation of SUMO (2-50)-MPA thioester segment **10** with SUMO (51-97)-SEA<sup>off</sup> peptide segment **8b**. Isolation of SUMO-1 (2-97)-SEA<sup>off</sup> protein **11****





Scheme S7

### Experimental procedure (Fig. S10)

The reaction was performed under nitrogen atmosphere.

A solution of 4-mercaptophenylacetic acid (MPAA, 16.82 mg, 0.1 mmol) in 6 M guanidinium chloride/0.1 M pH 7.0 sodium phosphate buffer was prepared (0.5 mL, final concentration of MPAA 200 mM). Peptide thioester **10** (2.6 mg, 0.355  $\mu$ mol) and SEA<sup>off</sup> peptide segment **7b** (2.19 mg, 0.355  $\mu$ mol) were dissolved in the above solution (118  $\mu$ L) and the pH was adjusted to 7.01 by addition of aqueous NaOH (6 M). The final peptide segment concentration was 3 mM. The reaction was agitated for 18 h at 37 °C, and then diluted with water (2 mL), acidified to pH 1-2 by adding 10% aqueous TFA (30  $\mu$ L) and extracted with diethylether (4  $\times$  8 mL) to remove the excess of MPAA. The mixture was immediately purified by HPLC to yield 0.974 mg of SUMO-1 (2-97)-SEA<sup>off</sup> peptide **11** (22%).

HPLC conditions: C3 Zorbax 300 Å column, detection at 215 nm, flow rate 6 mL/min, eluent A water containing 0.05% of TFA, eluent B CH<sub>3</sub>CN/water : 4/1 by vol containing 0.05% of TFA, gradient 0-10% B in 5 min, 10-40% B in 60 min.

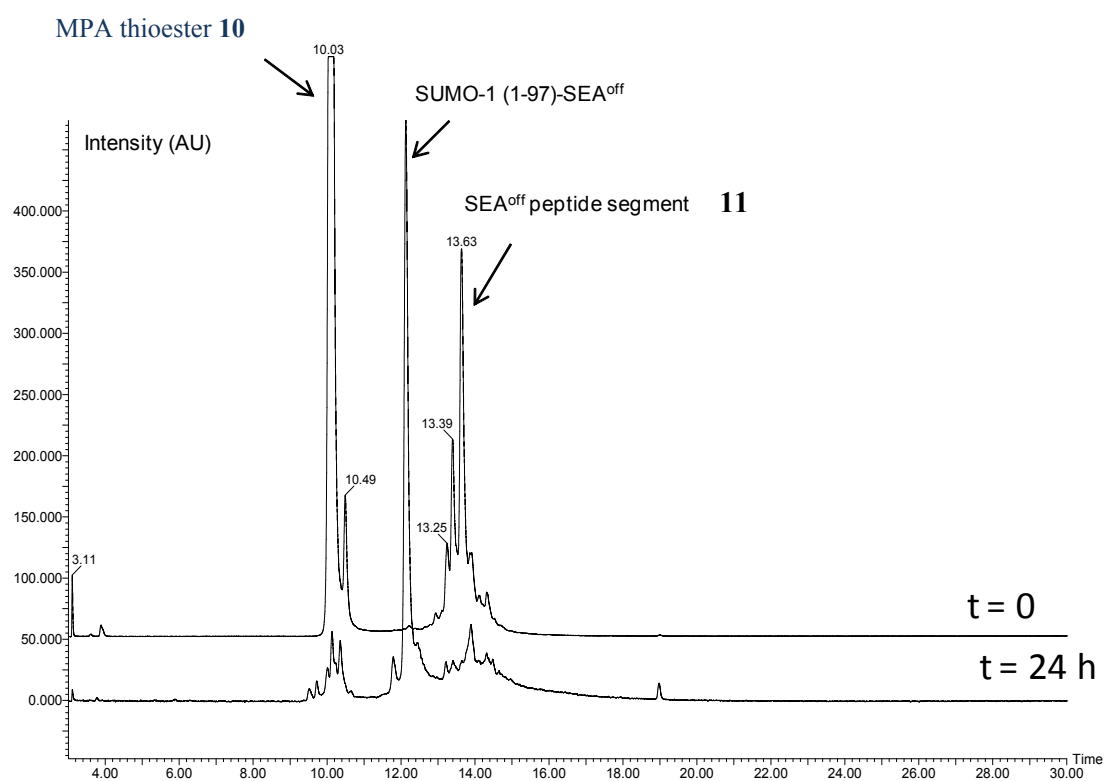


Fig. S10. LC-MS analysis of the ligation reaction of MPA peptide thioester **10** with SEA<sup>off</sup> peptide segment **7b**. The MS traces are in accordance with the proposed structures (data not shown for clarity, available on request).

### Characterization of SUMO-1 (2-97)-SEA<sup>off</sup> **11** (Fig. S11-14)

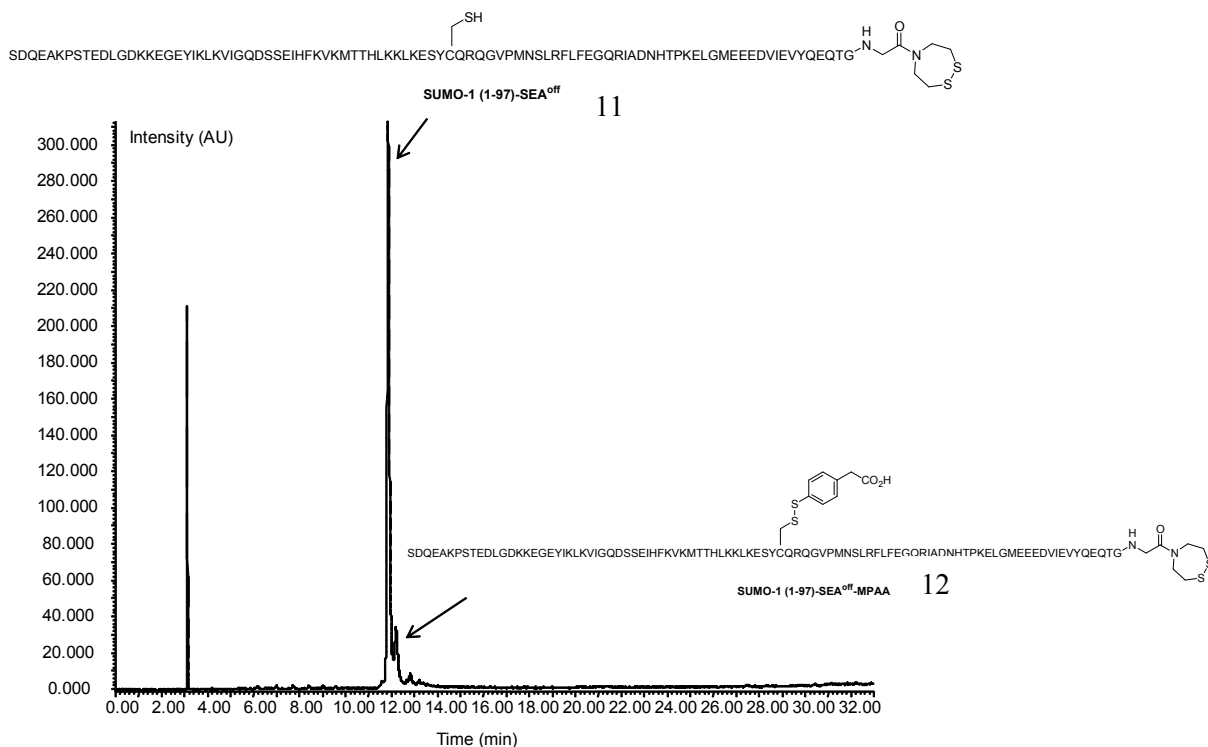


Fig. S11. LC-MS of purified SUMO-1 (2-97)-SEA<sup>off</sup> **11**. HPLC trace using light scattering detection. The minor peak corresponds to the disulfide with MPAA, i.e. SUMO-1 (2-97)-SEA<sup>off</sup>-MPAA **12**. *The presence of this product in the purified protein is not a problem because exploitation of SEA group functionality require the reduction of disulfide bonds in a subsequent step. In the one-pot process described later on, such mixed disulfides are always observed after the first ligation step. They are reduced in the second step by TCEP.*

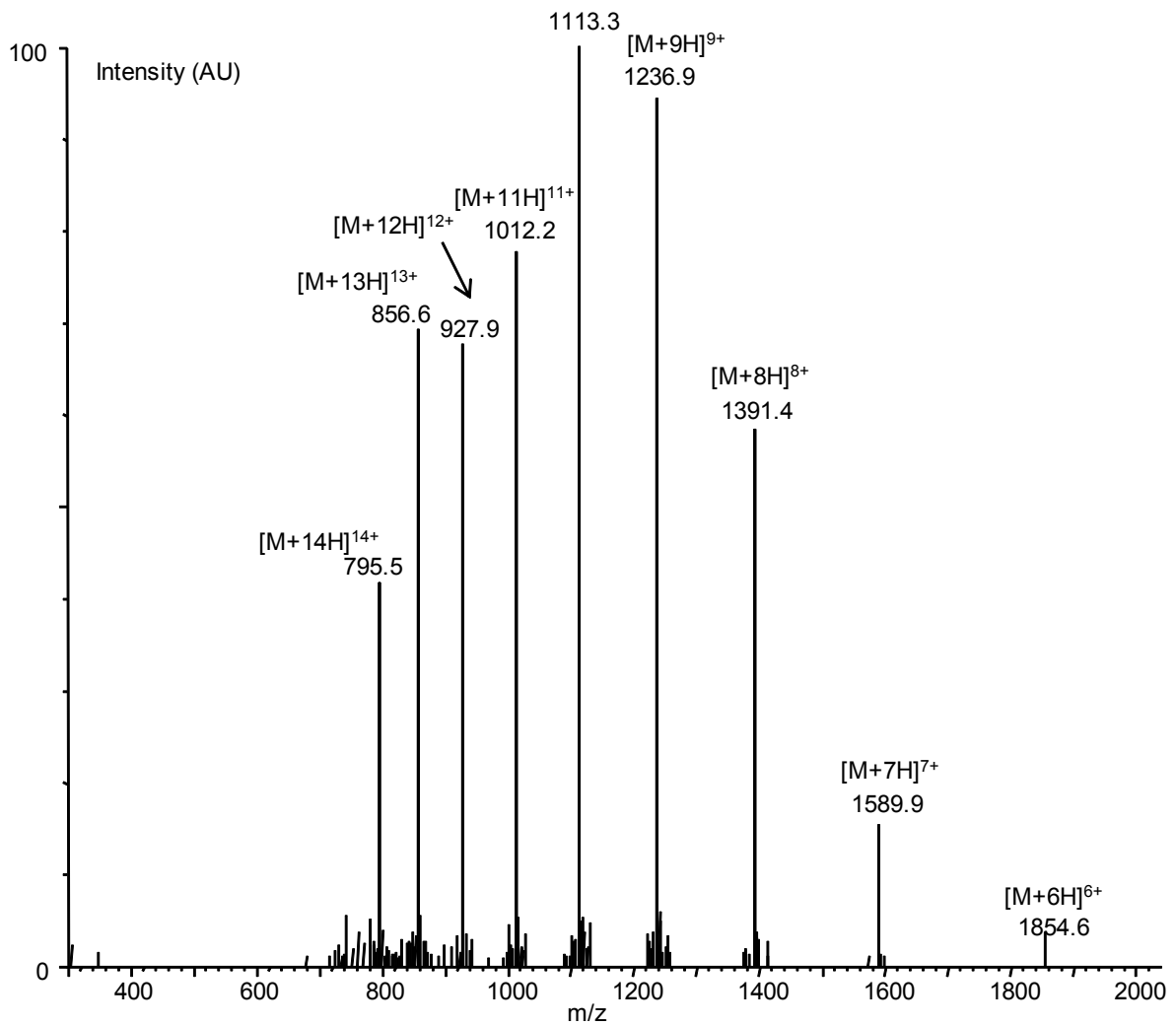
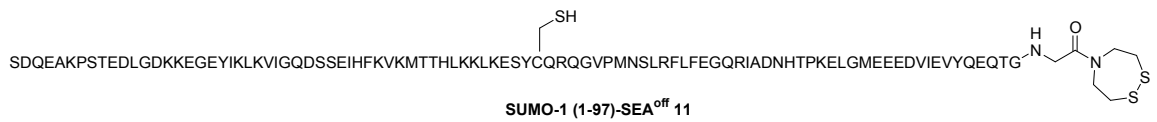


Fig. S12. LC-MS of purified SUMO-1 (2-97)-SEA<sup>off</sup> 11. MS trace of the major peak which corresponds to SUMO-1 (2-97)-SEA<sup>off</sup>. Calculated for M (average mass) 11118.5, observed after deconvolution 11122.7.

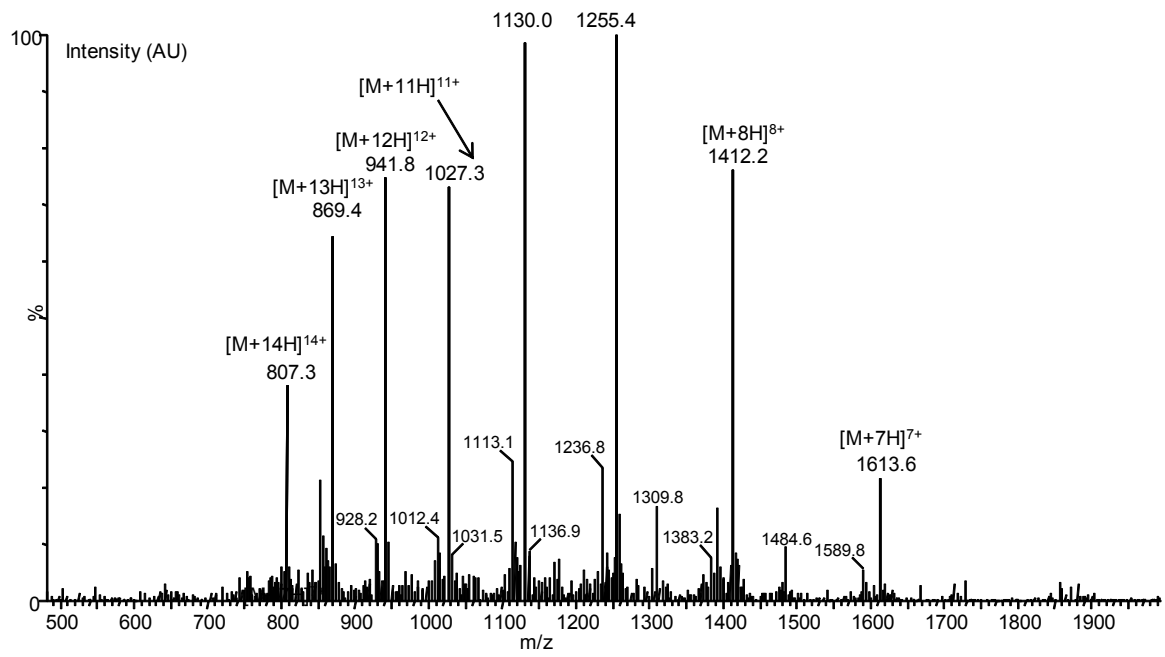
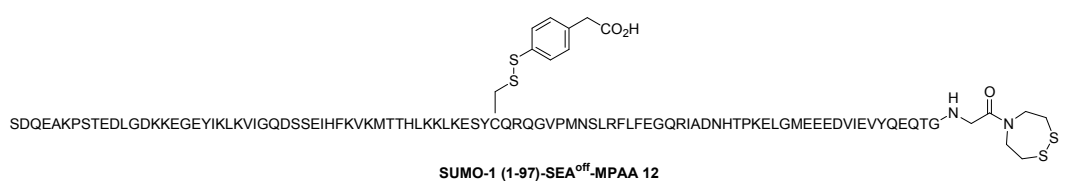


Fig. S13. LC-MS of purified SUMO-1 (2-97)-SEA<sup>off</sup>. MS trace of the minor peak which corresponds to SUMO-1 (2-97)-SEA<sup>off</sup>-MPAA 12. Calculated for M (average mass) 11284.7, observed after deconvolution 11288.9.

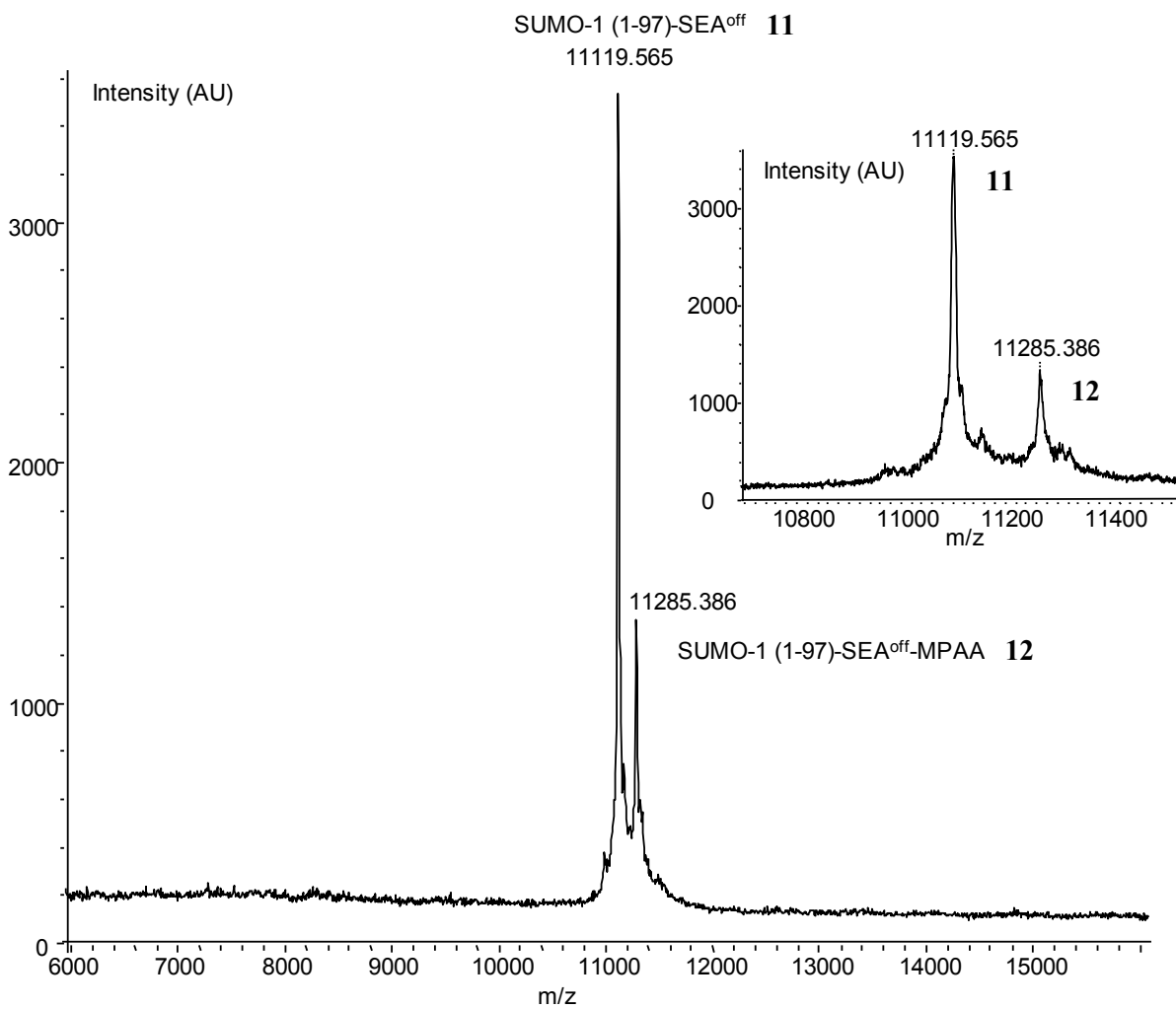
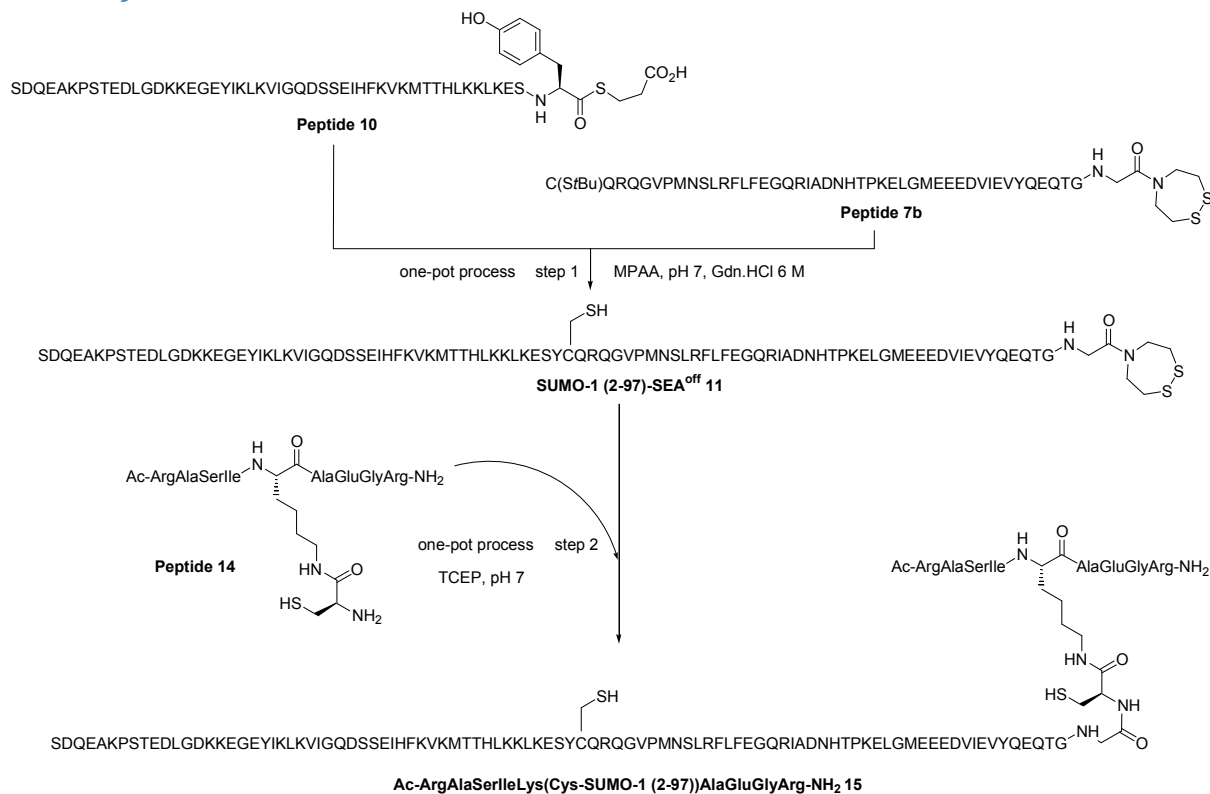


Fig. S14. MALDI-TOF analysis of purified SUMO-1 (2-97)-SEA<sup>off</sup>. SUMO-1 (2-97)-SEA<sup>off</sup> **11**, calculated for [M+H]<sup>+</sup> (average mass) 11119.5, observed 11119.56. SUMO-1 (2-97)-SEA<sup>off</sup>-MPAA **12**, calculated for [M+H]<sup>+</sup> (average mass) 11285.7, observed 11285.3.

## One-pot sequential ligation of three peptide segments for the synthesis of Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub> protein 15 (Fig. S15-20)



Scheme S8

Peptide Ac-ArgAlaSerIleLys(Cys)AlaGluGlyArg-NH<sub>2</sub> **14** was synthesized using standard Fmoc SPPS on an automated peptide synthesizer and purified by preparative RP-HPLC.

A solution of 4-mercaptophenylacetic acid (MPAA, 16.82 mg, 0.1 mmol) in 6 M guanidinium chloride/0.1 M pH 7.0 sodium phosphate buffer was prepared (0.5 mL). Peptide thioester **10** (4.61 mg, 0.63 μmol) and SEA<sup>off</sup> peptide segment **7b** (3.87 mg, 0.63 μmol) were dissolved in the above solution (209 μL) and the pH was adjusted to 7.2 by addition of aqueous NaOH (6 M). The reaction was agitated for 8 h at 37 °C.

Then, a solution of TCEP (12.04 mg, 42 μmol) and MPAA (7.06 mg, 42 μmol) in guanidine hydrochloride (6 M)/0.1 M pH 7.0 sodium phosphate buffer (209 μL) was prepared and used to dissolve peptide **14** (1.71 mg, 1.26 μmol). The pH of this solution was 7.3. The peptide solution was then added to the above reaction mixture to start the second ligation step. The

final peptide concentration was 1.5 mM. The reaction was agitated for 72 h and then diluted with water (0.5 mL), acidified to pH 2 by adding 10% aqueous TFA and extracted with diethylether (3 × 2 mL) to remove the excess of MPAA. The mixture was immediately purified by HPLC to yield 1.16 mg of Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub> peptide **15** (13%).

HPLC conditions: C3 Zorbax 300 Å column, detection at 215 nm, flow rate 6 mL/min, eluent A water containing 0.05% of TFA, eluent B CH<sub>3</sub>CN/water : 4/1 by vol containing 0.05% of TFA, gradient 0-20% B in 5 min, 20-50% B in 60 min.

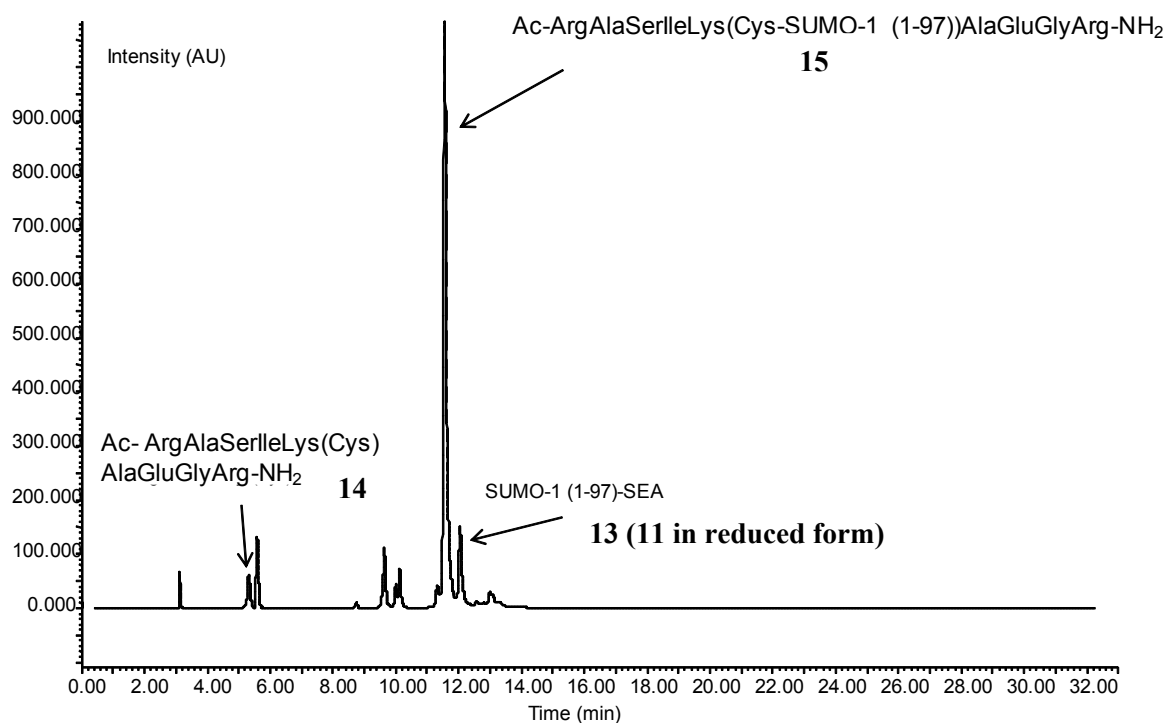


Fig. S15. LC-MS analysis of the crude reaction mixture after the two ligation steps (light scattering detection). MS data are in accordance with the proposed structures (data not shown, available on request). The major peak corresponds to the target sumoylated peptide Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub> **15** which was isolated by preparative HPLC and characterized (see below).



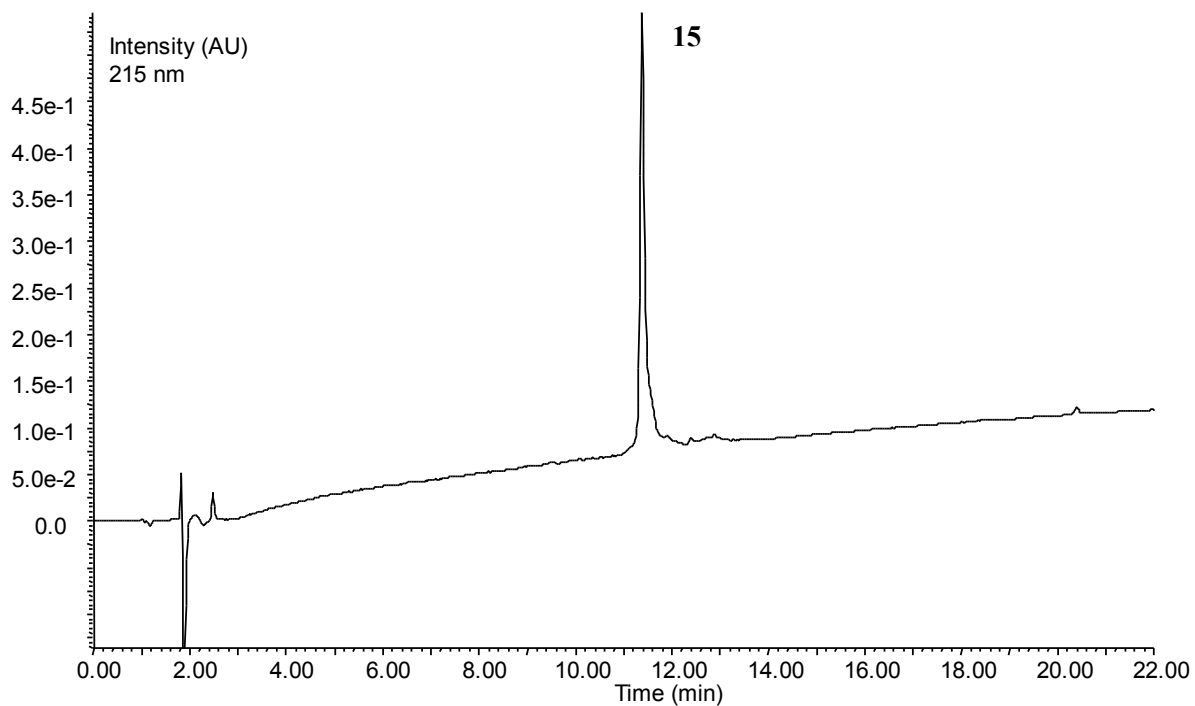


Fig. S16. LC-MS of the purified sumoylated peptide Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub> **15**. HPLC trace (UV detection at 215 nm).

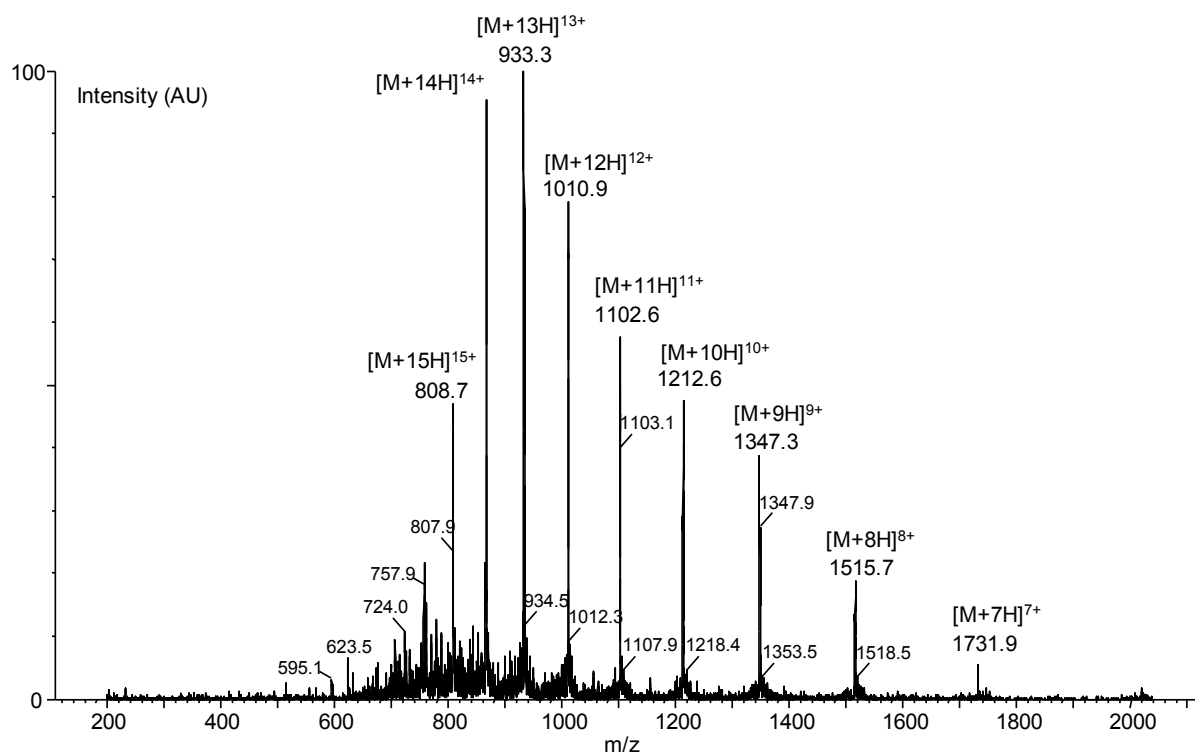


Fig. S17. LC-MS of the purified\_sumoylated peptide Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub> **15**. MS trace, calculated for M (average mass) 12115.5, observed 12118.69 after deconvolution.

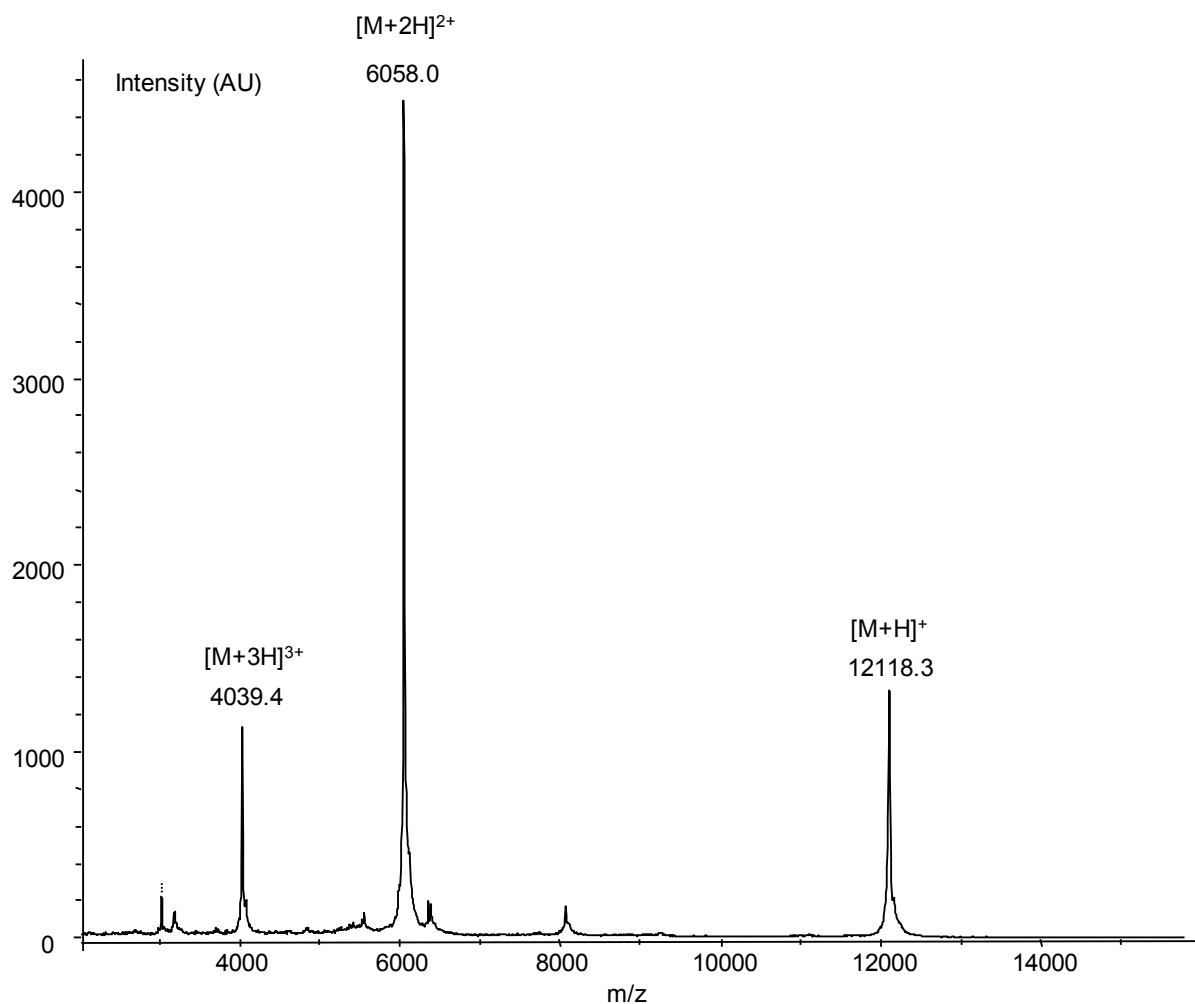


Fig. S18. MALDI-TOF analysis of the purified sumoylated peptide Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub> **15**. Matrix :  $\alpha$ -cyano-4-hydroxycinnamic acid. Calculated for  $[M+H]^+$  (average mass) 12116.5,  $[M+H]^+$  observed 12118.3.

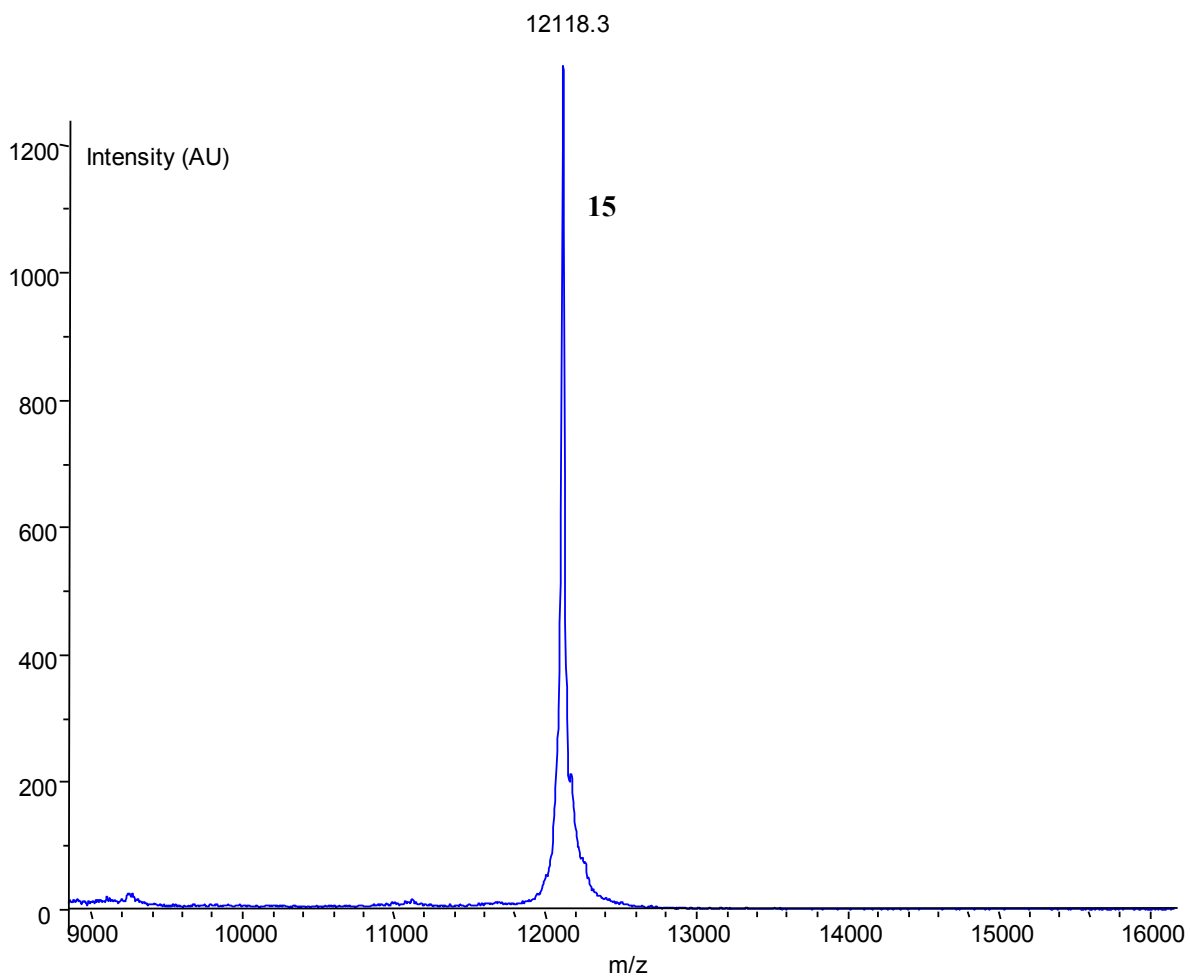


Fig. S19. MALDI-TOF analysis of the purified sumoylated peptide Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub> **15**. Matrix :  $\alpha$ -cyano-4-hydroxycinnamic acid. Calculated for [M+H]<sup>+</sup> (average mass) 12116.5, [M+H]<sup>+</sup> observed 12118.3.

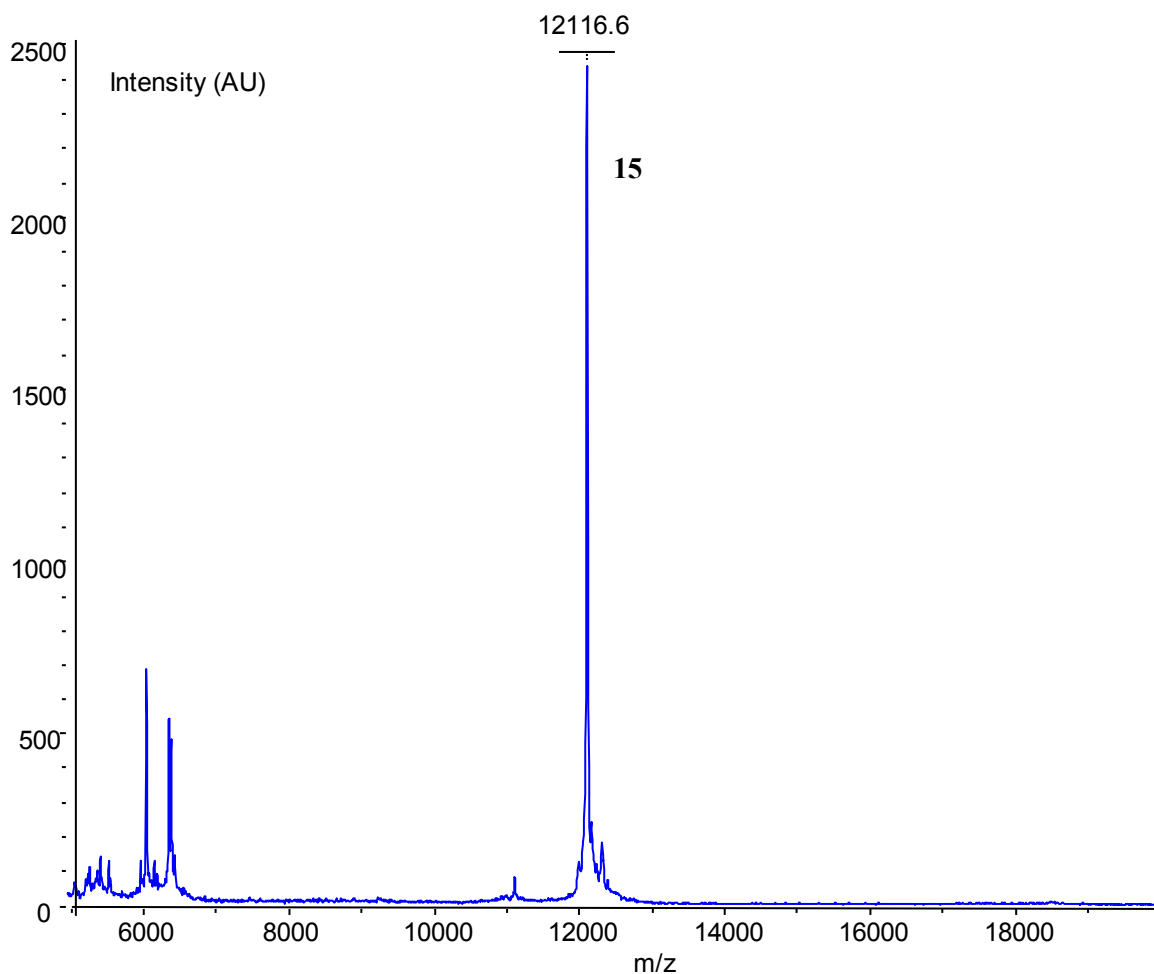


Fig. S20. MALDI-TOF analysis of the purified\_sumoylated peptide Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub> **15**. Matrix : sinapinic acid. Calculated for [M+H]<sup>+</sup> 12116.5, [M+H]<sup>+</sup> observed 12116.6.

### **MALDI-TOF analysis of Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub> **15**, in source fragmentation as a proof for the structure (Fig. S21-23)**

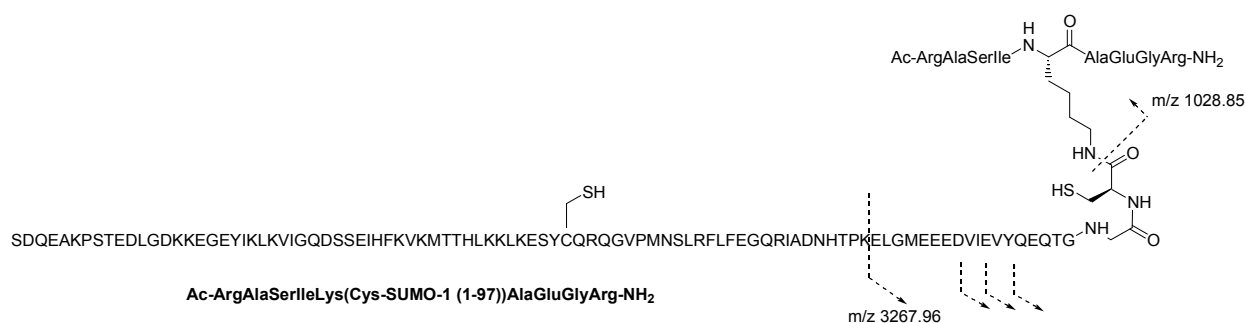
In this experiment, peptide Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub> was analyzed by MALDI-TOF mass spectrometry using in source fragmentation.

The analysis furnished the spectrum shown below in Fig. S21, in which we identified a fragment at m/z 1028 corresponding to the cleavage of the isopeptidic bond, as well as a series of fragments corresponding to the cleavage of SUMO-1 near its C-terminus, between Gly97 and Glu79 (these residues are underlined in Scheme S9).

Moreover, other ions corresponded to the fragmentation of the conjugate near the N-terminus of SUMO-1 peptide chain were identified as shown in Scheme S10 and Fig. S22 .

Finally, the ion at  $m/z$  1028 was selected and further fragmented. This experiment furnished the expected ion pattern for the short peptide part of the molecule (Fig. S23).

Overall, these data support the proposed structure for conjugate Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub>.



Scheme S9

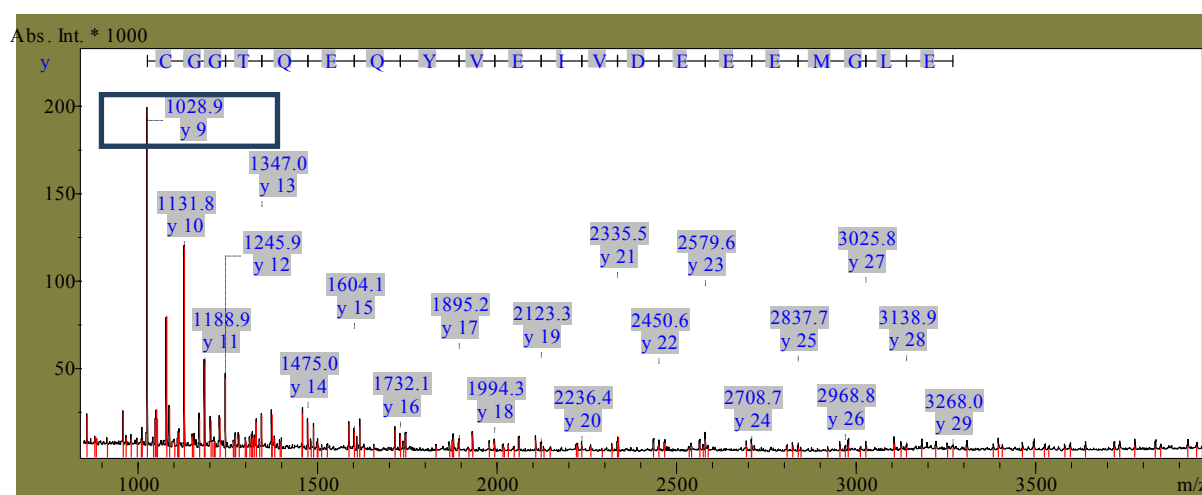


Fig. S21. MALDI-TOF analysis of the purified sumoylated peptide Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub> **15**. Matrix : 2,5-dihydroxybenzoic acid. The ion at  $m/z$  1028.85 corresponds to the cleavage of the isopeptide bond. Other ions correspond to the cleavage of SUMO-1 protein near the C-terminus as shown in the top of the figure.

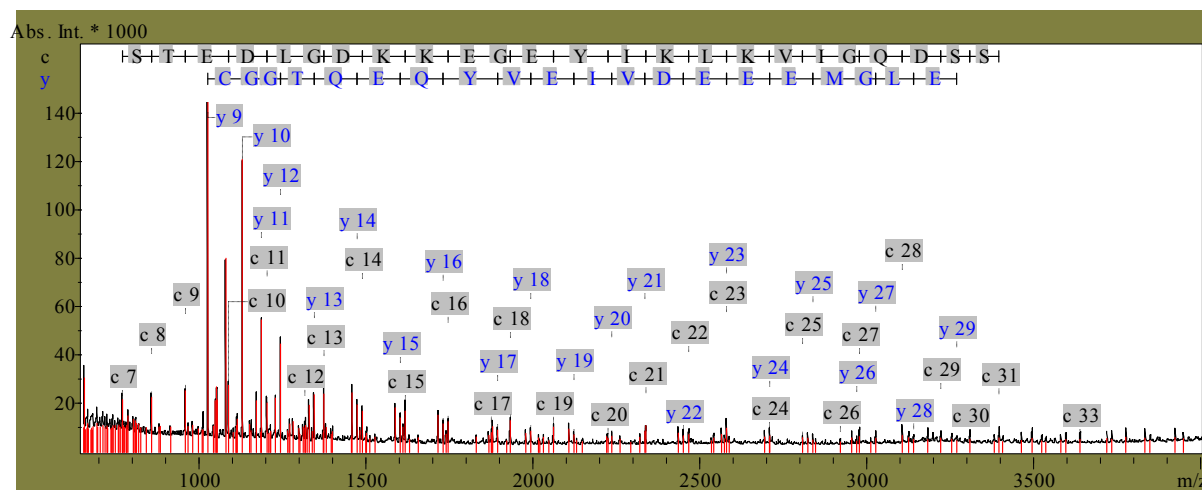
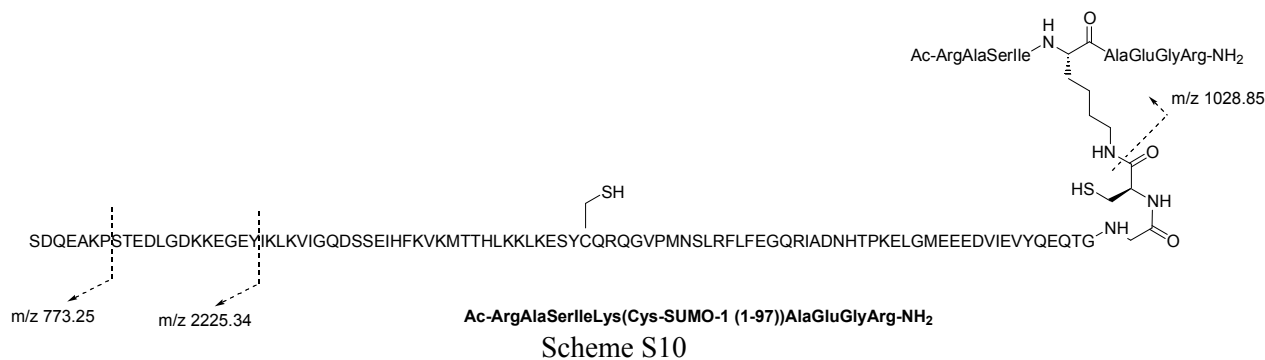


Fig. S22. MALDI-TOF analysis of the purified sumoylated peptide Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub> **15**. Matrix : 2,5-dihydroxybenzoic acid. Ions corresponding to the fragmentation of the N-terminal part of SUMO-1 were also detected. The residues concerned by this process are underlined in Scheme S10.

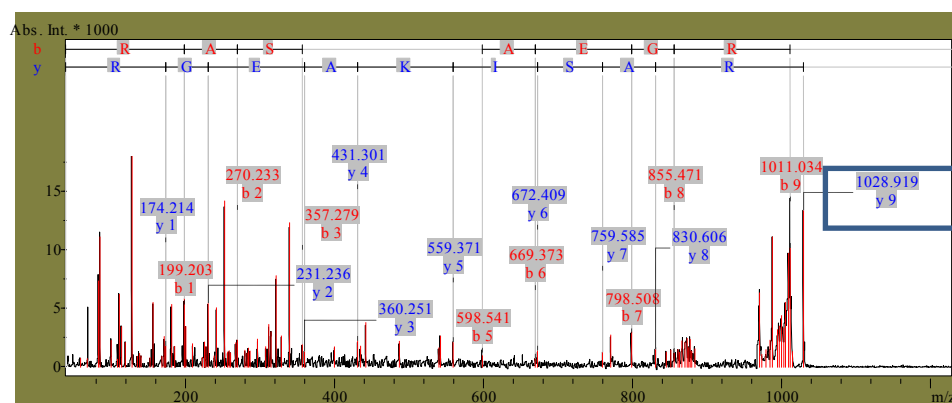


Fig. S23. MALDI-TOF analysis of the purified sumoylated peptide Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub> **15**. Matrix : 2,5-dihydroxybenzoic acid. The ion at m/z 1028.85 was selected to generate fragment ions corresponding to the short peptide.

## Biochemical characterization of Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub> protein 15 (Fig. S24-28)

### Cleavage of the conjugate by Ulp1 enzyme

We used in this experiment commercially available Ulp cysteine protease (Life technologies ref 12588-018), a recombinant fragment of Ulp1 (Ubl-specific protease 1) from *Saccharomyces cerevisiae*,<sup>3</sup> to probe the tertiary structure of SUMO-1 domain. Indeed, Ulp protease cleaves SUMO-1 conjugates in a highly specific manner by recognizing the tertiary structure of SUMO-1 modification.<sup>4</sup> Therefore, the cleavage of SUMO-1 conjugates by Ulp protease is a proof for the correct folding of the SUMO-1 domain within the conjugate.

The salt-free 10X SUMO protease buffer is composed of 500 mM Tris.HCl buffer, pH 8.0, 2% Igepal (NP-40) and 10 mM DTT.

The Ulp protease is provided in 25 mM Tris.HCl buffer, pH 8.0, 0.1% Igepal (NP-40), 250 mM NaCl, 50% (v/v) glycerol.

Experimental conditions for the microLC analysis: flow rate 50  $\mu$ L/min, gradient 0-60% of acetonitrile in 30min, Waters BEH300 C18 column, 5  $\mu$ m, 1  $\times$  150mm, UV detection at 215 nm.

The Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub> protein **15** (9  $\mu$ L, 1 mL/mL in water) was mixed with the 10X salt-free protease buffer (1  $\mu$ L). An aliquot was analyzed by microLC (trace c) in Fig. S24. The peak eluting at 24 min which corresponds to the conjugate was collected and analyzed by MALDI-TOF mass spectrometry as shown in Fig. S25.

The Ulp protease was then dissolved in deionized water (2  $\mu$ L, 1mg/mL) and added to the conjugate. The mixture was analyzed immediately by microLC (trace b in Fig. S24) and the peaks eluted at 14 min and 24 min were collected and analyzed by MALDI-TOF mass spectrometry. No change was observed 40 min, showing that the observed cleavage occurs in a few seconds.

The peak at 24 min corresponds to SUMO-1 (2-97) protein obtained by enzymatic cleavage (calcd. [M+H]<sup>+</sup> 11002.4, observed 11004.8) as shown in Fig. S25. This protein co-elutes with the starting Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub> conjugate.

The peak at 14 min corresponds to the peptide Ac-ArgAlaSerIleLys(Cys)AlaGluGlyArg-NH<sub>2</sub> **14** obtained by enzymatic cleavage as shown in Fig. S26. Trace a in Fig. S24 corresponds the reference peptide **14**.

---

<sup>3</sup> S. J. Li and M. Hochstrasser, *Nature*, 1999, **398**, 246-251.

<sup>4</sup> In this paper, the crystal structure of Ulp1-Smt3 complex (*Saccharomyces cerevisiae* Smt3 shares 47% of homology with mammalian SUMO-1) shows an extensive interface between the two proteins, with a total buried accessible surface area of 2400  $\text{Å}^2$ . E. Mossessova and C. D. Lima, *Mol Cell*, 2000, **5**, 865-876.

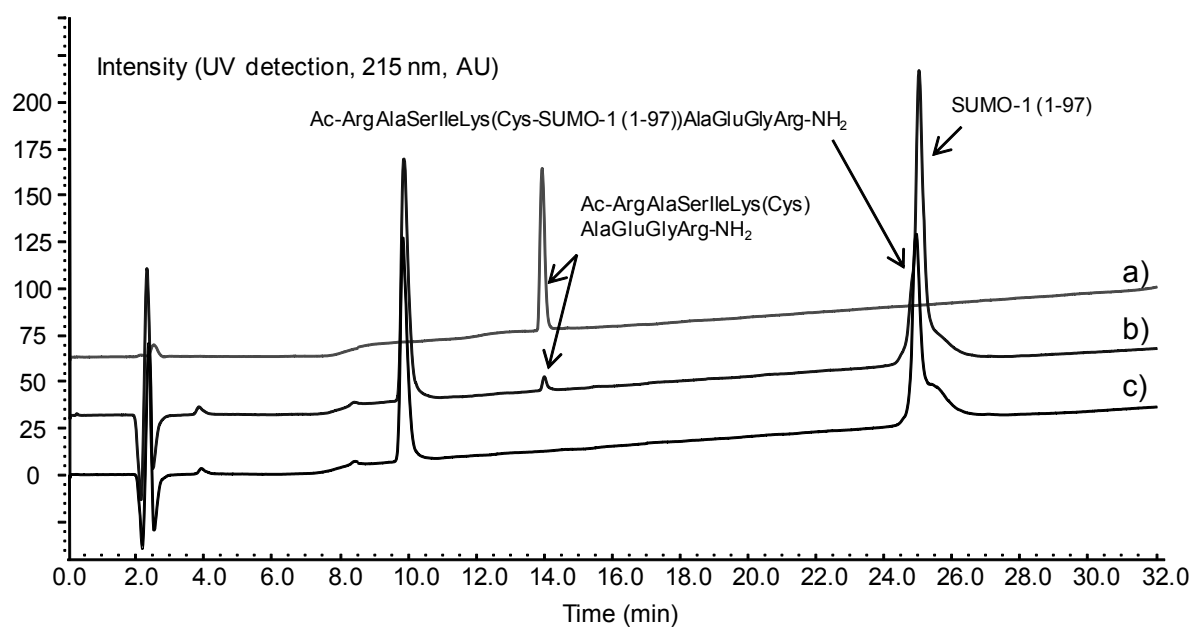


Figure S24. MicroLC analysis of the cleavage of Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub> protein **15** by Ulp protease. Trace a) Peptide **14**; trace b) Cleavage mixture few seconds after mixing; trace c) Conjugate **15** in the cleavage buffer. Experimental conditions for the microLC analysis: flow rate 50  $\mu$ L/min, gradient 0-60% of acetonitrile in 30 min, Waters BEH300 C18 column, 5  $\mu$ m, 1  $\times$  150 mm, UV detection at 215 nm.



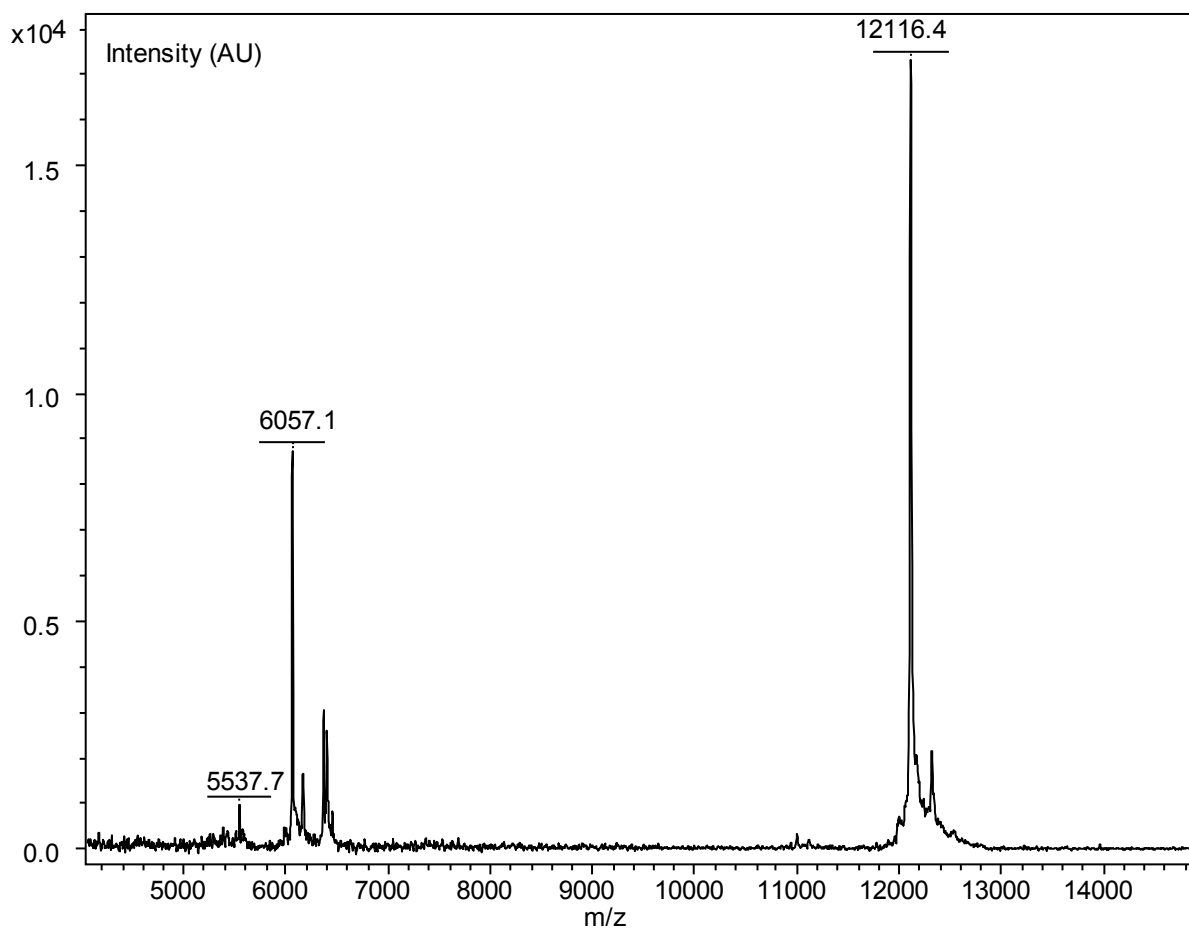


Fig. S25. Trace c in Fig. S24, MALDI-TOF analysis of the peak at 24 min. Sumoylated peptide Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub> **15**. Matrix : sinapinic acid. Calculated for [M+H]<sup>+</sup> 12116.5, [M+H]<sup>+</sup> observed 12116.4.

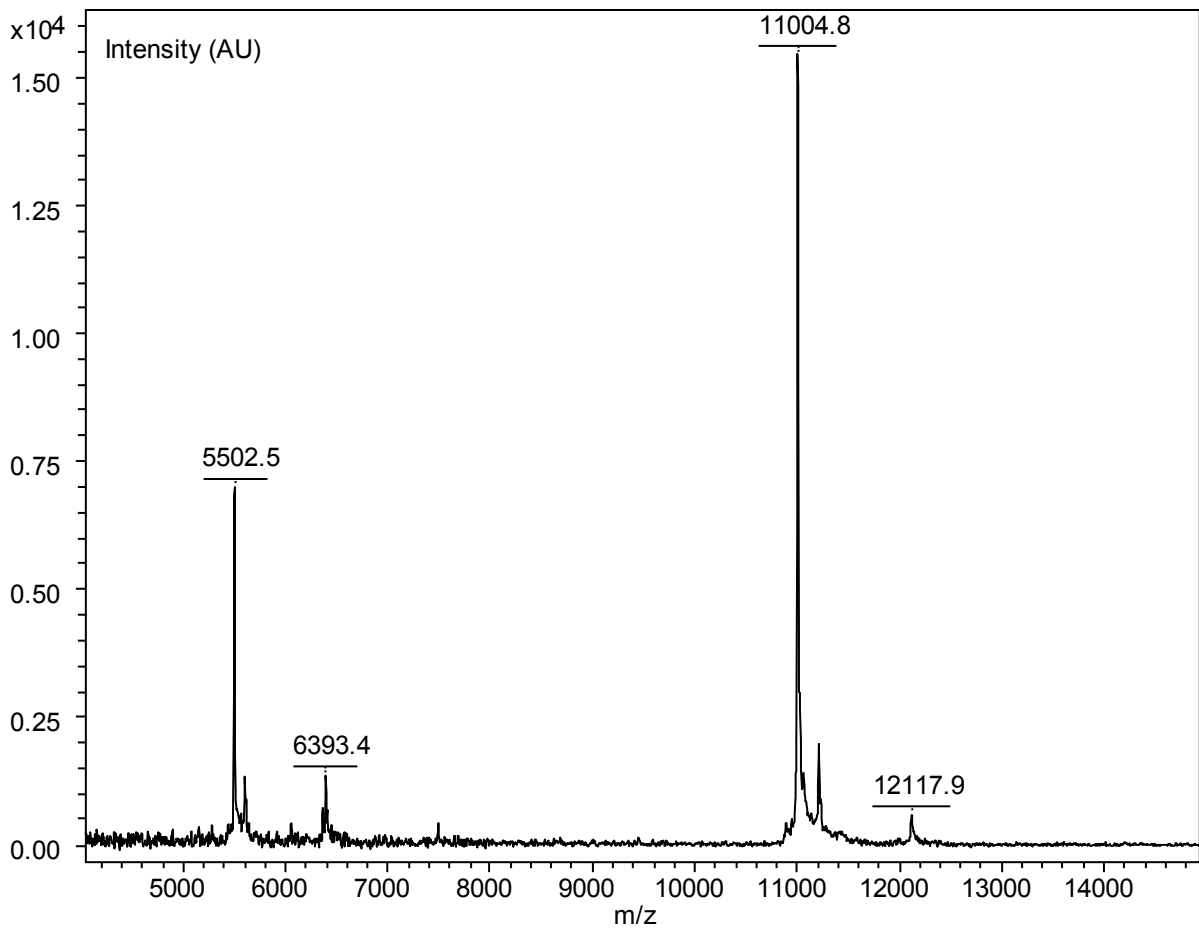


Fig. S26. Trace b in Fig. S24, MALDI-TOF analysis of the peak at 24 min. SUMO-1 (2-97) protein. Matrix : sinapinic acid. Calculated for  $[M+H]^+$  11002.4,  $[M+H]^+$  observed 11004.8.

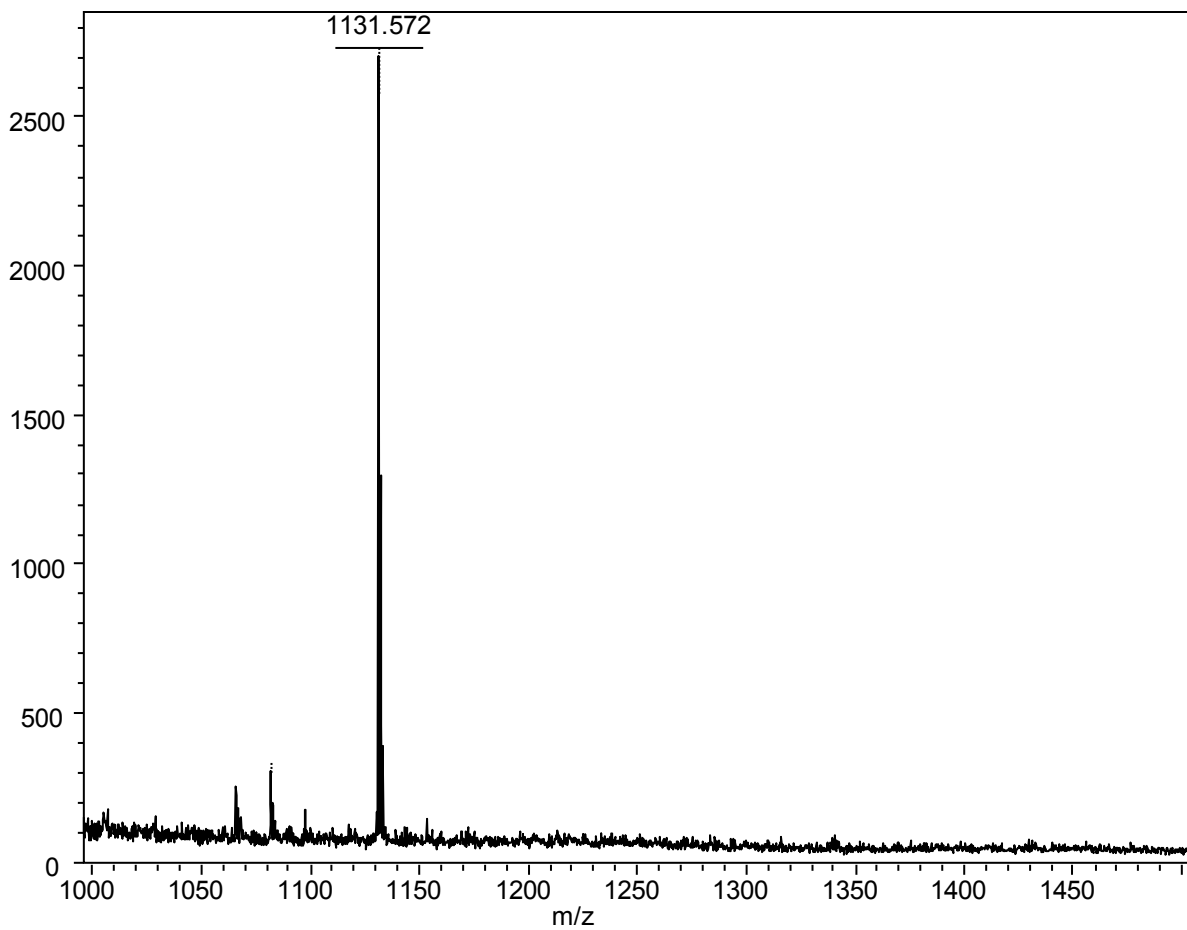


Fig. S27. Trace b in Fig. S24, MALDI-TOF analysis of the peak at 14 min. Peptide Ac-ArgAlaSerIleLys(Cys)AlaGluGlyArg-NH<sub>2</sub> **15**. Matrix : sinapinic acid. Calculated for [M+H]<sup>+</sup> 1131.605 [M+H]<sup>+</sup> observed 1131.57.

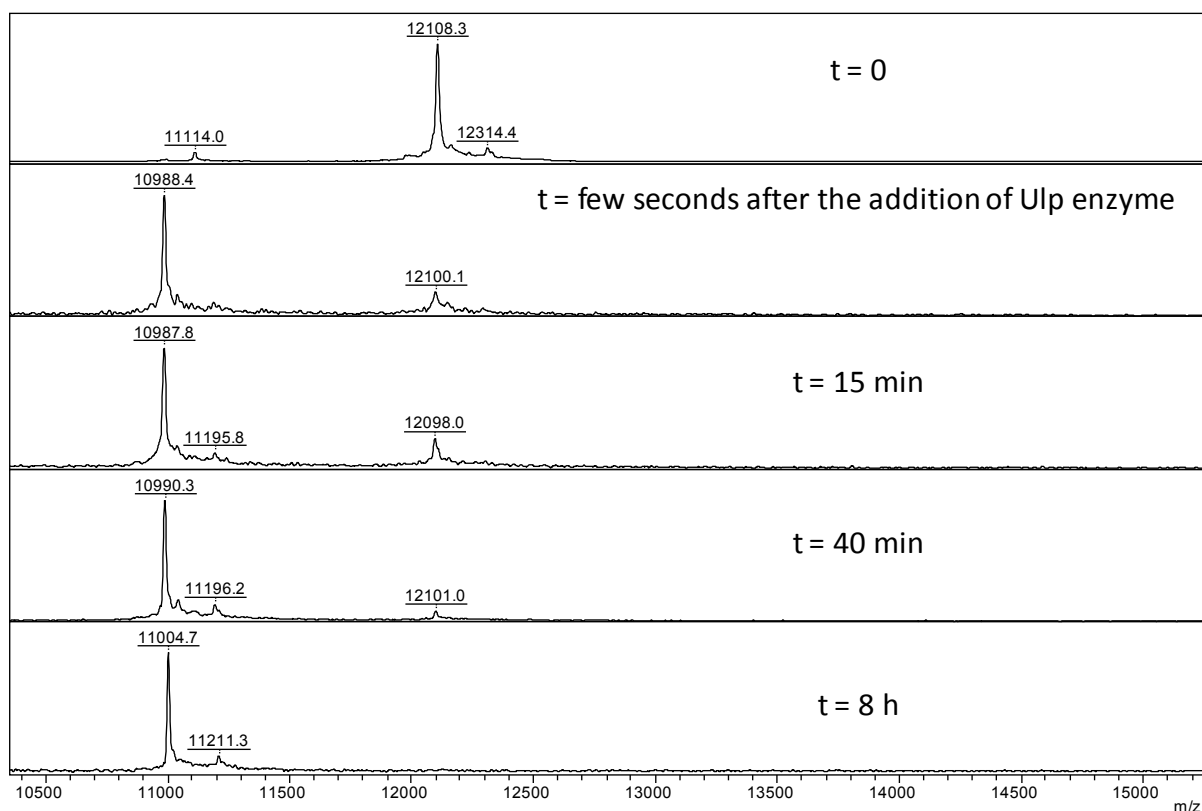


Fig. S28. Monitoring of the cleavage by MALDI-TOF mass spectrometry. Matrix : sinapinic acid.

## Characterization of Ac-ArgAlaSerIleLys(Cys-SUMO-1 (2-97))AlaGluGlyArg-NH<sub>2</sub> protein **15** by circular dichroism

The concentration of SUMO-1 conjugate **15** (7.4  $\mu$ M, 0.09 mg/mL in water) was determined by measuring the absorbance at 280 nm and validated with BCA assay (Thermo) using BSA as a reference. The circular dichroism spectrum of the protein was recorded on a Jasco Model J-810 spectropolarimeter. CD spectra were measured at 20 °C over the range 260–190 nm using 0.3 cm path-length cell and averaging of 10 scans. A 1 nm bandwidth, 0.5 nm data pitch and 50 nm/min scanning speed were used for spectral acquisition.

The  $\alpha$ -helical content of the synthetic protein SUMO-1 conjugate **15** (8.6 %) was estimated using the empirical equation of Greenfield & Fasman (Greenfield, N.; Fasman, G. D. Biochemistry 1969, 8, 4108.):

$$\alpha\text{-helix} = ([\theta]_{\text{MWR}222 \text{ nm}-3000}) / (-36,000 - 3000) \times 100.$$

## Synthesis of 20 kD peptide 9

### Synthesis of peptides 6d and 8 (Fig. S29-30)

#### Peptide sequences:

Peptide **6d**:

IRNC(S*t*Bu)IIGKGRSYKGTVSITKSGIKC(S*t*Bu)QPWSSMIPHEHSFLPSSYRGKDLQENY-SEA<sup>on</sup>  
(Fig S29)

Peptide **8**:

(CRNPRGEEGGPWCFTSNPEVRYEVCVIPQCSEVGG)<sub>2</sub>K-K(Biot)-NH<sub>2</sub> (Fig S30)

Peptide elongation was performed on ChemMatrix resin or SEA ChemMatrix resin (0.20 mmol/g) using standard Fmoc/tert-butyl chemistry on an automated peptide synthesizer.

Peptide **6d** was desalted on a short C18 column and lyophilized to give 120 mg (32 % yield, 50  $\mu$ mol scale) of crude peptide **6d**. The peptide (35 mg) was further purified by C18 HPLC (17.1 mg, 5%).

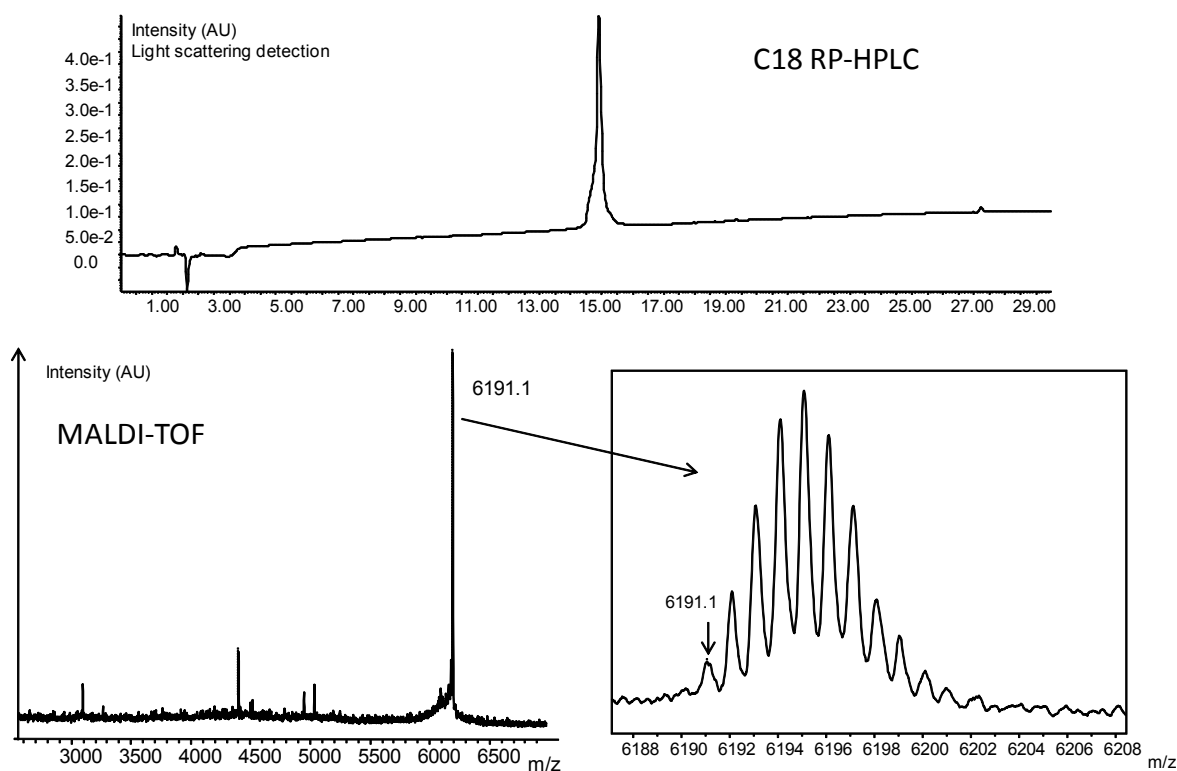


Fig. S29. Analytical HPLC profile and Maldi-TOF analysis of peptide-SEA<sup>on</sup> segment **6d**. Calcd. for [M+H]<sup>+</sup>: 6192.07, observed mass: 6191.1 ( monoisotopic composition).

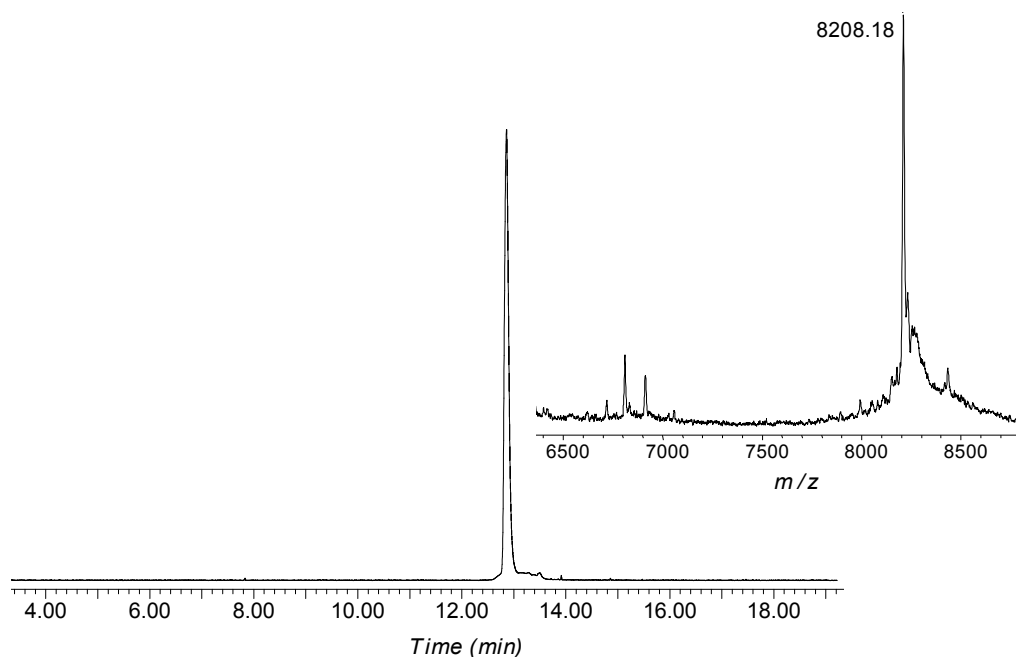
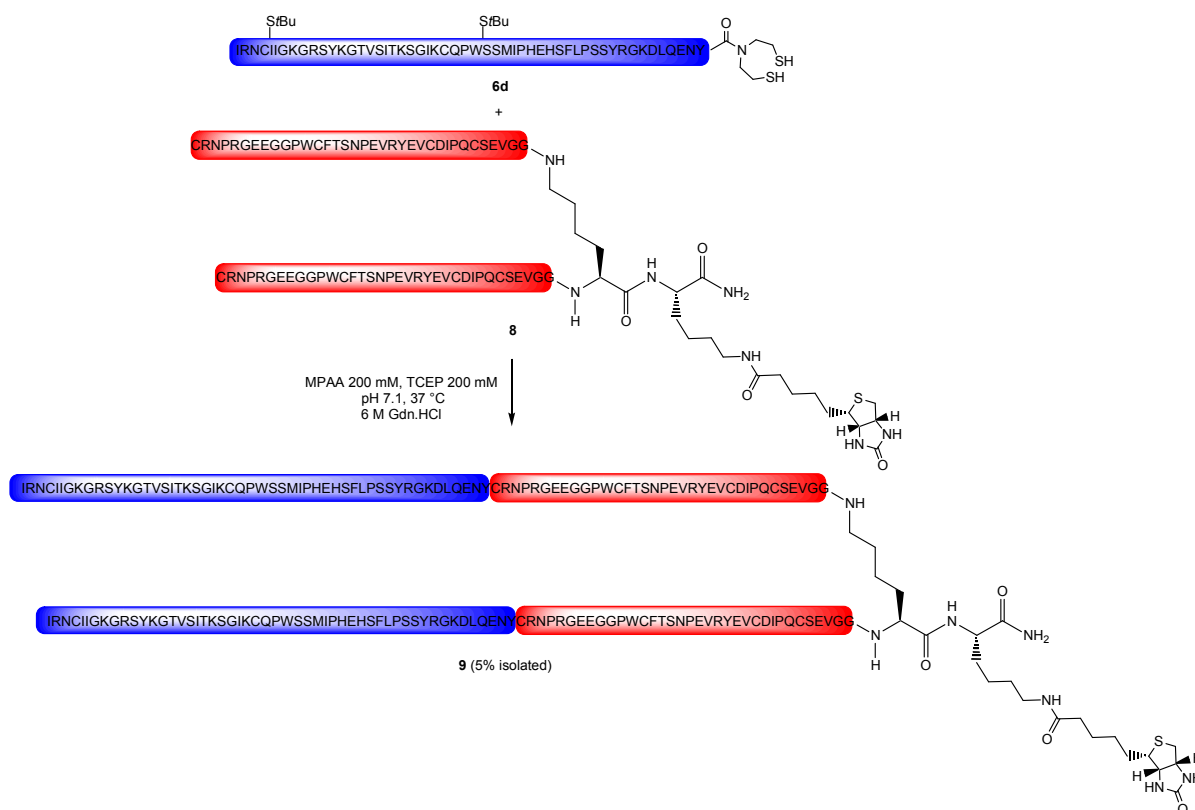


Fig. S30. Analytical HPLC profile ( $\lambda=215$  nm) and Maldi-TOF data of the purified biotinylated peptide segment **8**. 13 mg (6 % yield, 25  $\mu$ mol scale), calcd. for [M+H]<sup>+</sup>: 8208.1 observed : 8208.2 (average isotopic composition).

## Ligation of peptides **6d** and **8**. Isolation of the 20 kD peptide **9**



Scheme S11

TCEP.HCl (58 mg, 0.2 mmol) and MPAA (33 mg, 0.2 mmol) were dissolved in 6 M guanidinium chloride, 0.1 M pH 7.2 sodium phosphate buffer (1 mL). NaOH (5M) was then added to adjust the pH to 7.1.

Segment **6d** (5 mg, 0.7  $\mu$ mol) and Cys-segment **8** (12 mg, 0.35  $\mu$ mol, 0.5 eq) were dissolved in the above solution (224  $\mu$ L) and *t*BuOH (56  $\mu$ L) (final peptide concentration 2.5 mM). The reaction mixture was shaken at 37 °C under nitrogen atmosphere and monitored by LC-MS (Figure S26). After 48 h, the mixture was diluted with water-TFA 10% (2 mL) and purified by reversed-phase HPLC using a linear water-acetonitrile gradient containing 0.05% TFA to give the purified peptide **9** (0.82 mg, 5%) (Figure S27).

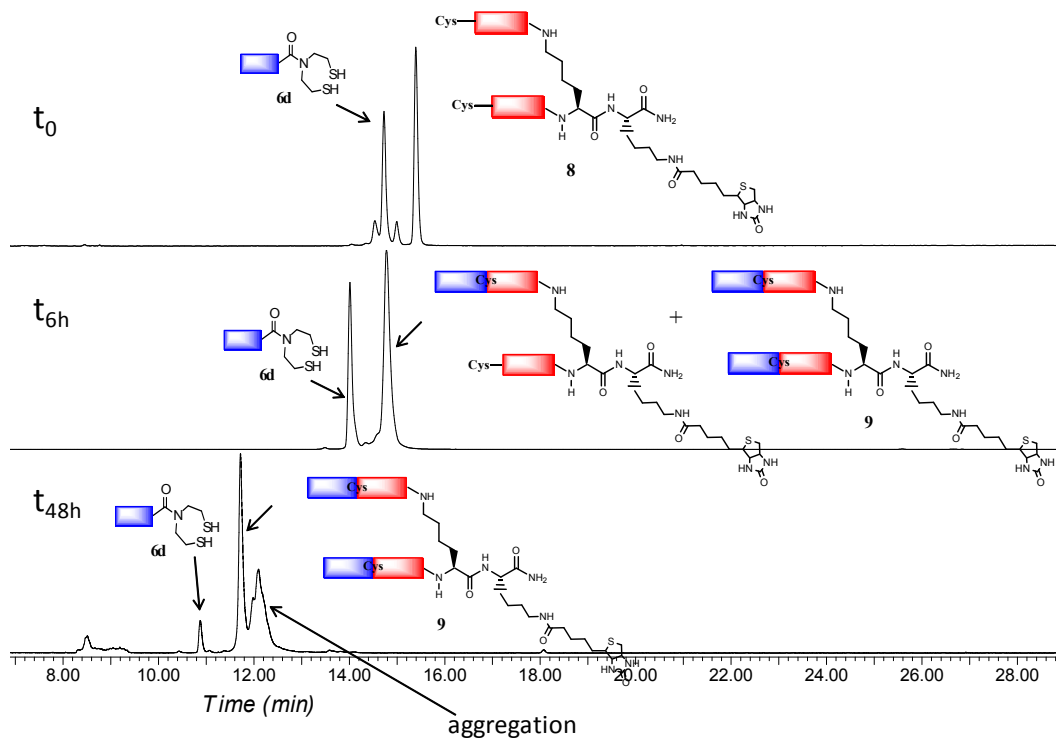


Fig. S29. Monitoring of the assembly of the 20 kD peptide **9** by LC-MS (light scattering detection).

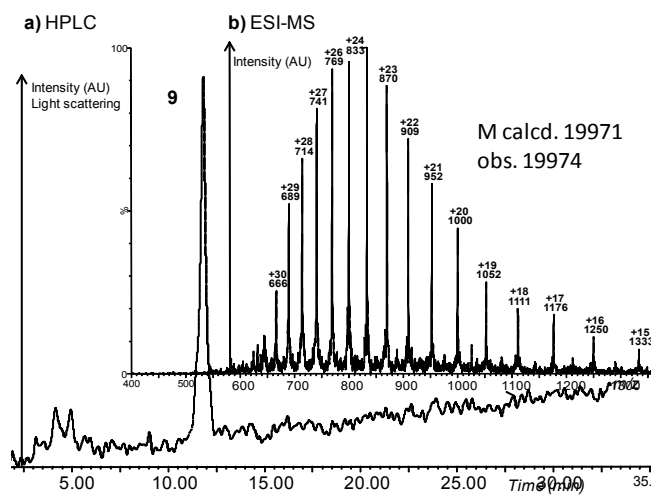


Fig. S30. LC-MS characterization of peptide **9**.

### Synthesis of peptide **7c** using SEA NovaSyn resin

Peptide sequence for **7c**: GQRKRRNTIHEFKKSAKTTLIKIDPALKIKTKKVNTADQ-SEA<sup>off</sup>

The peptide elongation was performed on the SEA Novasyn TGT resin (0.11 mmol/g) using standard Fmoc/*tert*-butyl chemistry on an automated batch peptide synthesizer.



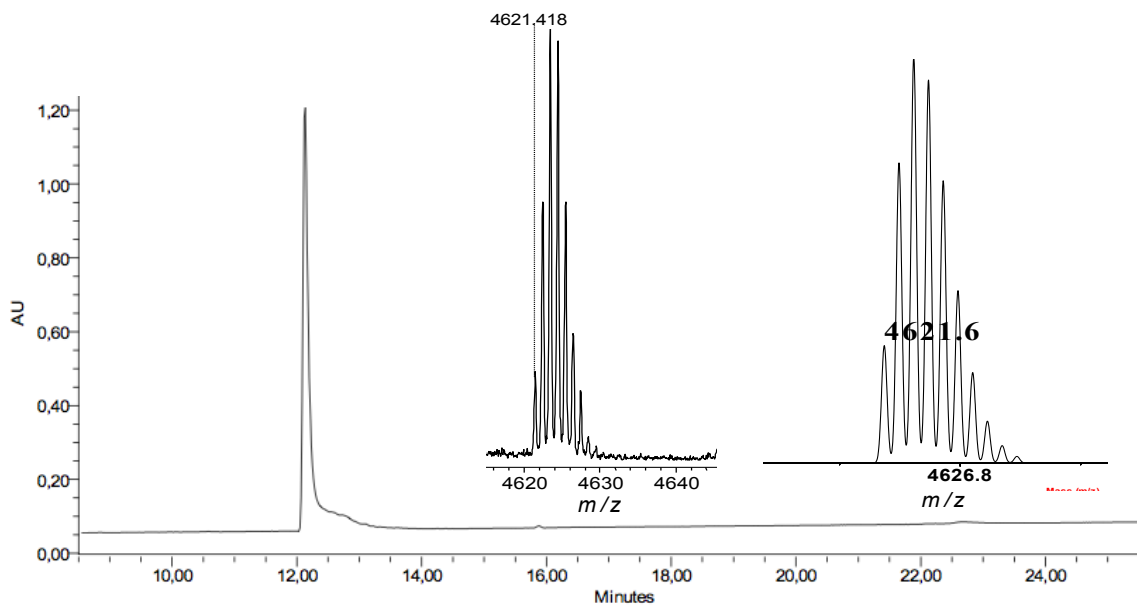


Fig. S31. Analytical HPLC profile ( $\lambda=215$  nm) and Maldi-TOF data for the purified peptide GQRKRRNTIHEFKKSAKTTLIKIDPALKIKTKKVNTADQ-SEA<sup>off</sup>. 45 mg (7 % yield, 100  $\mu$ mol scale), calcd. for  $[M+H]^+$ : 4621.42, observed: 4621.6 (isotopic composition).