Multi-molecule reaction of serum albumin can occur through thiol-yne coupling

Mauro Lo Conte,[a] Samuele Staderini,[a] Alberto Marra,[a] Macarena Sánchez-Navarro,[b] Benjamin G. Davis*[b] and Alessandro Dondoni *[a]

[a] *Dipartimento di Chimica, Laboratorio di Chimica Organica, Università di Ferrara, Via L. Borsari 46, 44100 Ferrara, Italy.*
[b] *Department of Chemistry, University of Oxford, Chemistry Research Laboratory, Mansfield Road, Oxford OX1 3TA, UK.*

Supporting Information

Table of contents **Page**

General Experimental Section. Flash column chromatography[29] was performed on silica gel 60 (40-63 μm). Optical rotations were measured at 20 ± 2 °C in the stated solvent; $[\alpha]_D$ values are given in deg·mL·g^{-1}·dm^{-1}. ^{1}H NMR (300 and 400 MHz) and ^{13}C NMR spectra (75 MHz) were recorded from $CDCl_3$ solutions at room temperature unless otherwise specified. Peak assignments were aided by ^{1}H-^{1}H COSY and gradient-HMQC experiments. In the ^{1}H NMR spectra reported below, the *n* and *m* values quoted in geminal or vicinal proton-proton coupling constants $J_{n,m}$ refer to the number of the corresponding sugar protons.

For accurate mass measurements the compounds were analyzed in positive ion mode by electrospray hybrid quadrupole orthogonal acceleration time-of-flight mass spectrometer (Q-TOF) fitted with a Z-spray electrospray ion source. The capillary source voltage and the cone voltage were set at 3500 V and 35 V, respectively; the source temperature was kept at 80 °C; nitrogen was used as a drying gas at a flow rate of ca. 50 L/h. The time-of-flight analyzer was externally calibrated with NaI from *m/z* 300 to 2000 to yield an accuracy near to 5 ppm. When necessary an internal lock mass was used to further increase the mass accuracy. Accurate mass data were collected by directly infusing samples (10 pmol/μL in 1:1 CH_3CN-H_2O containing 10 mM ammonium formate) into the system at a flow rate of 5 μL/min. The monoisotopic masses were calculated according to the reported[30] atomic weights of the elements.

MALDI MS analysis was conducted on a Water MALDI Micro MX spectrometer with TOF detection, in positive linear mode. Unmodified protein was used as a lock mass calibrant. Data were processed using Mass Lynx 4.1. Mass spectra were smoothed with Savitzky-Golay smoothing prior to dispersity analysis. A 2 μL sample of the protein solution (1mg/mL) was diluted to 20 μL with 0.1% TFA in H_2O. This solution was mixed in a 1:10 ratio (v/v) with a solution of sinapinic acid (10 mg/mL in 2:3 H_2O/CH_3CN containing 0.1% TFA). 2.5 μL of this matrix/sample solution, was spotted onto a steel target and allowed to co-crystallize at r. t. over 3 h.

The commercially available photoinitiator DPAP (Aldrich 19611-8) and glutathione **6** were used without further purification. The known *N*-Boc cysteine ethyl ester **2**[31] and the fluorescein thiol **4**[32] were prepared as described.

For each sample of protein 1 μL of solution (~1mgmL^{-1}) was analyzed by LC ESI-QTOF MS/MS using a Waters SYNAPT HDMS system with a Waters nano-acquity UPLC system. Chromatography comprises mobile phase a 100% water with 0.1% formic acid and mobile phase B 100% acetonitrile with 0.1% formic acid.

A 90 min chromatographic gradient was used at 0.4μL/min to give a linear increase from 5% B to 40% B in 75 min, from 40% B to 95% B in 5 min, from 95% B to 1% B in 5 min and the column was conditioned again at 1% B for 8 min.

A Water BEHC18 75uM x 100mm trap column was used with an analytical Waters BEH130 C-18 (1.7 particle size) nanoAcquity UPLC column. Other mass spec parameters were as follows : Nanolockspray using Reserpine's ^{13}C peak as a lockmass; Capillary = 2.65 kV; Cone = 35 V.

Synthesis of propargyl 1-thio-β-D-glucopyranoside (1). A solution of known[33] propargyl 2,3,4,6-tetra-*O*-acetyl-1-thio-β-D-glucopyranoside (200 mg, 0.50 mmol) in MeOH (5 mL), Et_3N (1 mL), and H_2O (1 mL) was kept at r. t. for 14 h and then concentrated. The residue was eluted from a C-18 silica gel cartridge with H_2O to give **1** (107 mg, 92%) as a colorless syrup; $[\alpha]_D$ = -86.2 (*c* 0.7, H_2O). ^1H NMR (300 MHz, D_2O): δ 4.59 (d, 1H, $J_{1,2}$ = 9.8 Hz, H-1), 3.76 (dd, 1H, $J_{5,6a}$ = 2.3, $J_{6a,6b}$ = 12.5 Hz, H-6a), 3.57 (dd, 1H, $J_{5,6b}$ = 5.8 Hz, H-6b), 3.47 (dd, 1H, *J* = 2.2, 17.0 Hz, 1 H of CH_2S), 3.40-3.24 (m, 5H), 2.52 (dd, 1H, *J* = 2.2, 2.7 Hz, C≡CH). ^{13}C NMR (75 MHz, D_2O): δ 84.1 (CH), 80.2 (C), 79.8 (CH), 77.1 (CH), 72.1 (CH), 71.9 (CH), 69.4 (CH), 60.7 (CH_2), 16.9 (CH_2). HRMS (ESI/Q-TOF) *m/z* calcd for $C_9H_{14}NaO_5S$ $(M+Na)^+$ 257.0460, found 257.0475.

Synthesis of 5. The reaction was carried out in a glass vial, located 2.5 cm away from the household UVA lamp apparatus equipped with four 15 W tubes (1.5 x 27 cm each). A solution of cysteine **2** (40 mg, 0.16 mmol), propargyl thioglucoside **1** (150 mg, 0.64 mmol), and 2,2-dimethoxy-2-phenylacetophenone (DPAP, 4.1 mg, 16 μmol) in MeOH (400 μL) was irradiated at r. t. for 10 min under magnetic stirring and then concentrated. The residue was eluted from a column of silica gel with AcOEt-CH_2Cl_2 (from 1.5:1 to 2:1) to give syrupy **3** (24 mg, 31%) as a ca. 1:1 *E/Z* mixture. Eluted second was unmodified **1** (112 mg, 75%). ^1H NMR (300 MHz, CD_3OD) selected data: δ 6.23 (bd, 0.4H, *J* = 15.0 Hz, CH=C*H*S), 6.18 (d, 0.6H, *J* = 9.4 Hz, CH=C*H*S), 5.78-5.64 (m, 1H, C*H*=CHS), 4.38 (d, 0.6H, *J* = 9.8 Hz, H-1), 4.36 (d, 0.4H, *J* = 9.8 Hz, H-1), 4.22 (q, 2H, *J* = 7.0 Hz, CH_2CH$_3$), 1.48 (s, 9H, *t*-Bu), 1.31 (t, 3H, *J* = 7.0 Hz, CH_2CH_3). HRMS (ESI/Q-TOF) *m/z* calcd for $C_{19}H_{33}NNaO_9S_2$ $(M+Na)^+$ 506.1494, found 506.1511.

A solution of **3** (24 mg, 49.6 μmol), **4** (93 mg, 0.20 mmol), and DPAP (5.1 mg, 20 μmol) in DMF (350 μL) was irradiated at r. t. for 30 min under magnetic stirring and then concentrated. The residue was eluted from a column of Sephadex LH-20 (1.5 x 40 cm) with MeOH to give **5** (16.5 mg, 35%) as a syrup. ^1H NMR (400 MHz, CD_3OD) selected data: δ 8.10 (bs, 1H), 7.79-7.74 (m, 1H), 7.18 (d, 1H, *J* = 8.2 Hz), 6.76 (bd, 2H, *J* = 8.4 Hz), 6.65 (d, 2H, *J* = 2.6 Hz), 6.57 (dd, 2H, *J* = 2.6, 8.4 Hz), 4.52 (d, 1H, *J* = 10.0 Hz, H-1), 4.18 (q, 2H, *J* = 7.0 Hz, CH_2CH$_3$), 1.42 (s, 9H, *t*-Bu). ^{13}C NMR (100 MHz, CD_3OD): δ 182.6, 172.8, 171.4, 154.8, 130.7, 126.5, 114.7, 112.1, 103.5, 88.3, 87.0, 82.0, 80.8, 79.6, 74.6, 74.5, 62.9, 62.6, 55.5, 45.3, 38.1, 35.6, 35.4, 31.2, 28.7, 14.6. HRMS (ESI/Q-TOF) *m/z* calcd for $C_{42}H_{52}N_3O_{14}S_4$ $(M+H)^+$ 950.2332, found 950.2298.

Synthesis of 8. To a solution of glutathione **6** (23 mg, 75 μmol) and **1** (70 mg, 0.30 mmol) in H_2O (200 μL) was added a solution of DPAP (2 mg, 7.5 μmol) in MeOH (100 μL). The solution was irradiated at r. t. for 5 min under magnetic stirring and then eluted from a column of Sephadex LH-20 with 1:1 H_2O-MeOH to give syrupy **7** (26 mg, 64%) as a ca. 1:1 *E/Z* mixture. Eluted second was unmodified **1** (56 mg, 80%). ^{1}H NMR (300 MHz, D_2O) selected data: δ 6.12-6.04 (m, 1H, CH=C*H*S), 5.72-5.56 (m, 1H, C*H*=CHS), 4.31 (d, 0.5H, *J* = 10.0 Hz, H-1), 4.29 (d, 0.5H, *J* = 10.0 Hz, H-1), 2.44-2.35 (m, 2H), 2.05-1.96 (m, 2H). HRMS (ESI/Q-TOF) *m/z* calcd for $C_{19}H_{32}N_3O_{11}S_2$ $(M+H)^+$ 542.1478, found 542.1501.

A solution of **7** (26 mg, 48 μmol), **4** (86 mg, 0.19 mmol), and DPAP (4.9 mg, 19 μmol) in DMF (350 μL) was irradiated at r. t. for 30 min under magnetic stirring and then concentrated. The residue was eluted from a column of Sephadex LH-20 (1.5 x 40 cm) with MeOH to give **8** (7.2 mg, 15%) as a syrup. ^{1}H NMR (400 MHz, CD_3OD) selected data: δ 8.10 (bs, 1H), 7.84-7.80 (m, 1H), 7.18-7.13 (m, 1H), 6.69-6.62 (m, 3H), 6.57-6.53 (m, 2H), 4.52 (d, 0.5H, *J* = 10.0 Hz, H-1), 4.50 (d, 0.5H, *J* = 10.0 Hz, H-1). HRMS (ESI/Q-TOF) *m/z* calcd for $C_{42}H_{50}N_5O_{16}S_4$ $(M+H)^+$ 1008.2135, found 1008.2098.

Synthesis of 12. To a solution of commercial BSA (30 mg, 0.46 μmol) in 20 mM phosphate buffer at pH 7.4 (3.0 mL) was slowly added a solution of **1** (3.5 mg, 15.0 μmol) and DPAP (0.3 mg, 1.3 μmol) in DMSO (150 μL). The mixture was irradiated at r. t. for 5 min under magnetic stirring, then filtered twice through Ultrafree-MC microcentrifuge filter (nominal MW limit of 5,000 Da) for 30 min at 5,000 rpm. To a solution of the protein glycoconjugate **10** in 20 mM phosphate buffer at pH 7.4 (3.0 mL) was added a solution of **4** (34 mg, 73 μmol) and DPAP (1.9 mg, 7.3 μmol) in DMSO (150 μL). The mixture was irradiated at r. t. for 10 min under magnetic stirring, then filtered twice through Ultrafree-MC microcentrifuge filter (nominal MW limit of 5,000 Da) for 30 min at 5,000 rpm to give **12**.

Synthesis of 13. To a solution of commercially available BSA (0.5 mg, 7.5 nmol) in 20 mM phosphate buffer at pH 7.4 (125 μL) was slowly added a solution of cyclooctyne **9** (0.2 mg, 64.0 μmol) in DMSO (4μL) and DPAP (1 μL of a 0.025 M solution in DMSO, 25.0 nmol). The mixture was irradiated at r. t. for 50 min, then filtered twice through Ultrafree-MC microcentrifuge filter (nominal MW limit of 10,000 Da) for 10 min at 5,000 rpm. To a solution of the protein conjugate **11** in 20 mM phosphate buffer at pH 7.4 (125 μL) was added a solution glutathione **6** (0.47 mg, 1.5 μmol) in DMSO (4 μL) and DPAP (2 μL of a 0.025 M solution in DMSO, 50.0 nmol). The mixture

was irradiated at r. t. for 10 min, then filtered twice by Ultrafree-MC microcentrifuge filter (nominal MW limit of 10,000 Da) for 10 min at 5,000 rpm to give **13**.

Synthesis of 14. To a solution of commercially available BSA (0.5 mg, 7.5 nmol) in 20 mM phosphate buffer at pH = 7.4 (125 µL) was slowly added a solution of cyclooctyne **9** (0.1 mg, 64.0 µmol) in DMSO (2 µL). The mixture was placed in the end-over-end rotator during 10h and then filtered twice by Ultrafree-MC microcentrifuge filter (nominal MW limit of 10,000 Da) for 10 min at 5,000 rpm to give **14**.

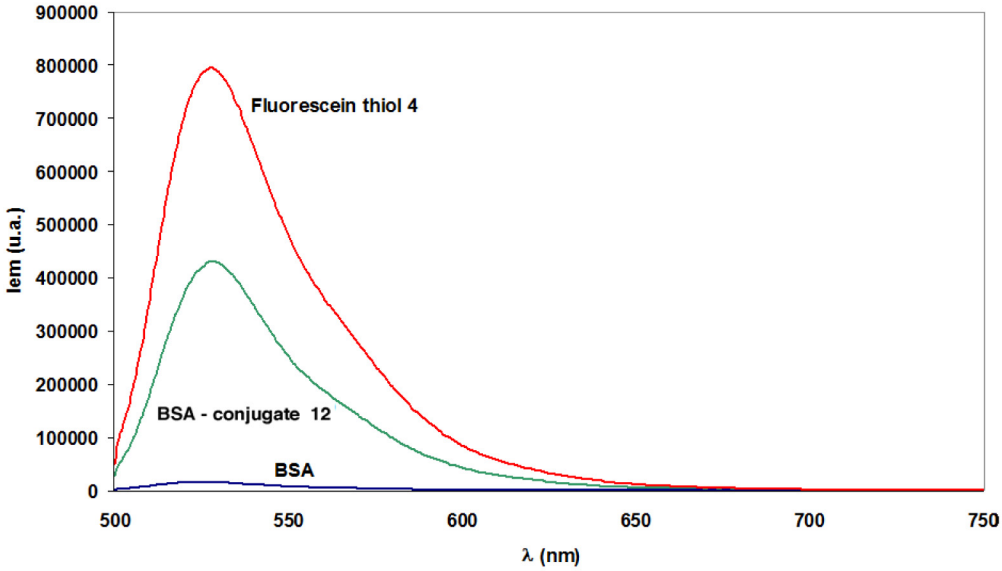

Fig. S1 Fluorescence emission spectra (λ_{ex} = 490 nm) of phosphate buffer solutions (pH 7.40) of BSA, BSA-conjugate **12**, and thiol **4**.

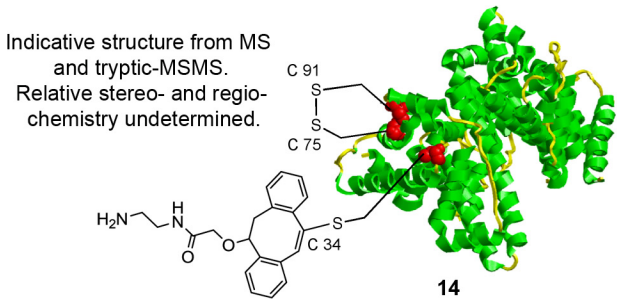

Figure S2. Structure of **14**.

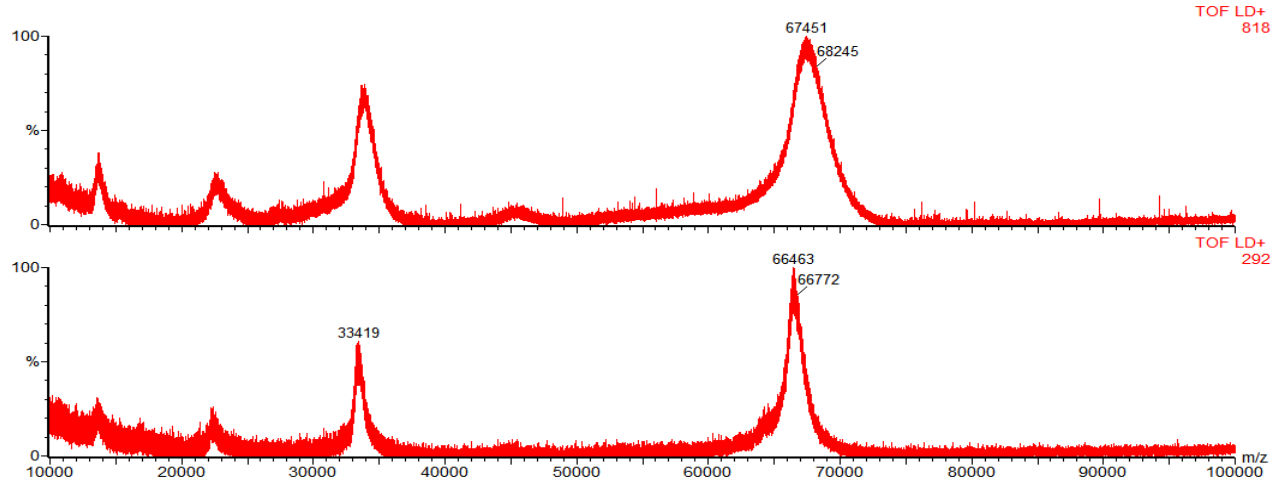

Fig. S2. MALDI-TOF MS spectra of BSA (bottom) and BSA-conjugate **11** (top).

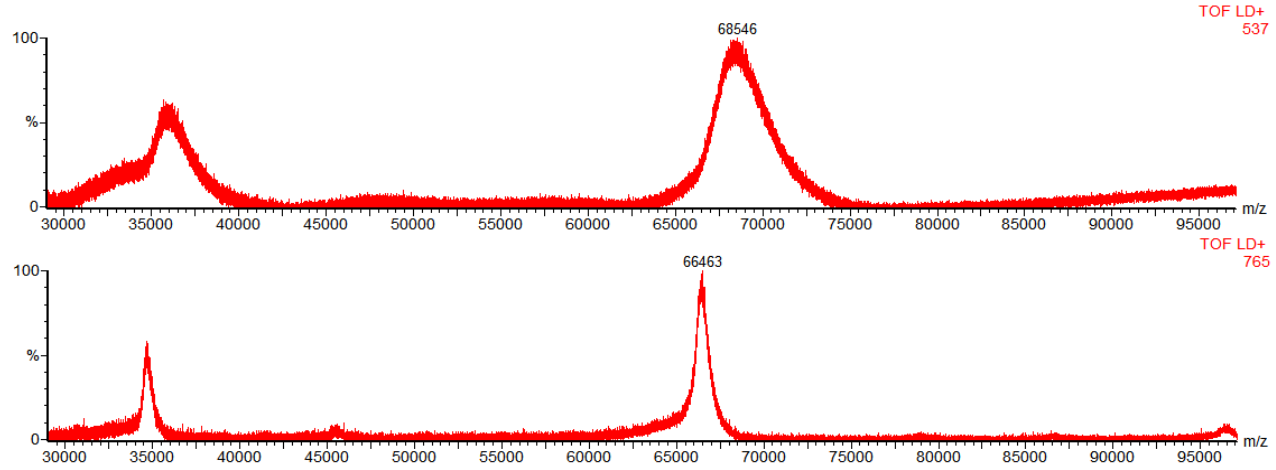

Fig. S3. MALDI-TOF MS spectra of BSA (bottom) and BSA-conjugate **12** (top).

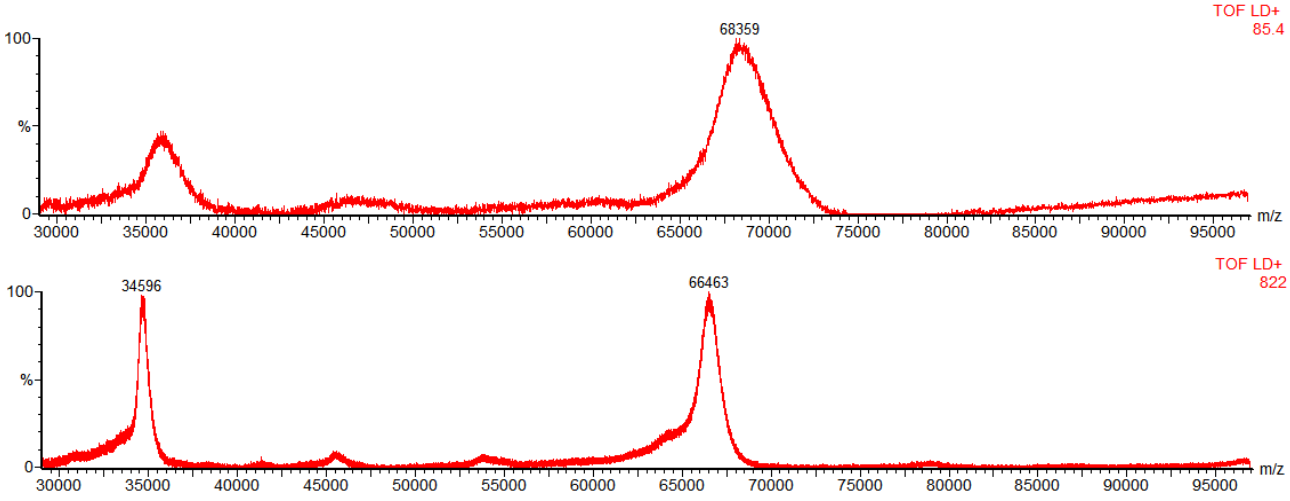

Fig. S4. MALDI-TOF MS spectra of BSA (bottom) and BSA-conjugate **13** (top).

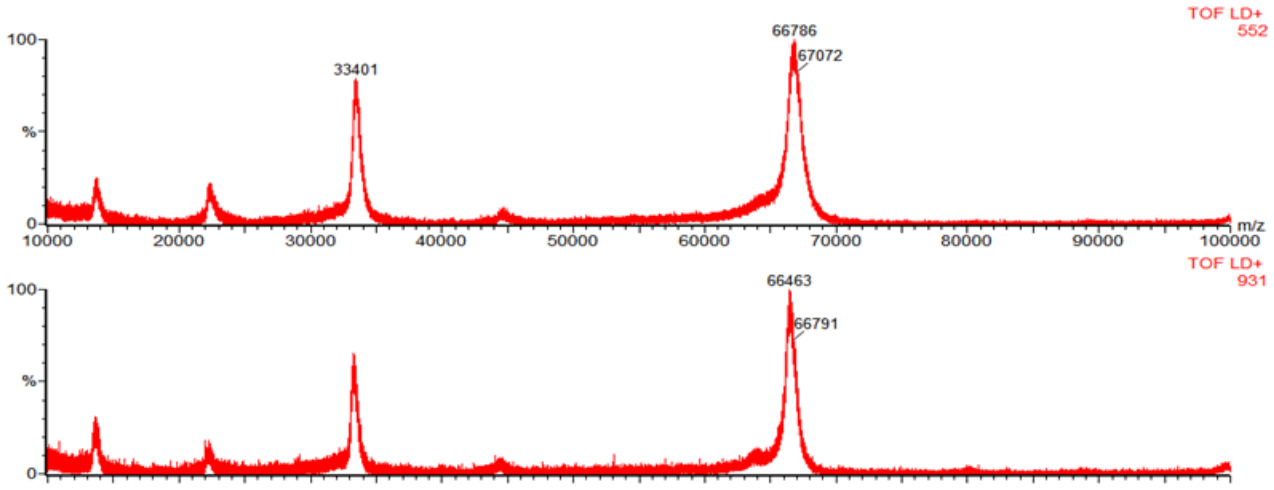

Fig. S5. MALDI-TOF MS spectra of BSA (bottom) and BSA-conjugate **14** (top).

Tryptic digestion and LC ESI-QTOF MS analysis of derivatives 13 and 14 .

The enzymatic digestion of BSA derivatives **13** and **14** was performed by tryptic digestion of the samples without reduction of Cys-Cys disulfide bridges. A 100 µL solution of each sample (**13** and **14**) at 50 ng/µL in 20 mM sodium phosphate buffer were incubated overnight (12 h) with shaking at 37 ˚C after addition of trypsin enzyme (Promega) for a ratio of 1/50 enzyme/protein (*w*/*w*).

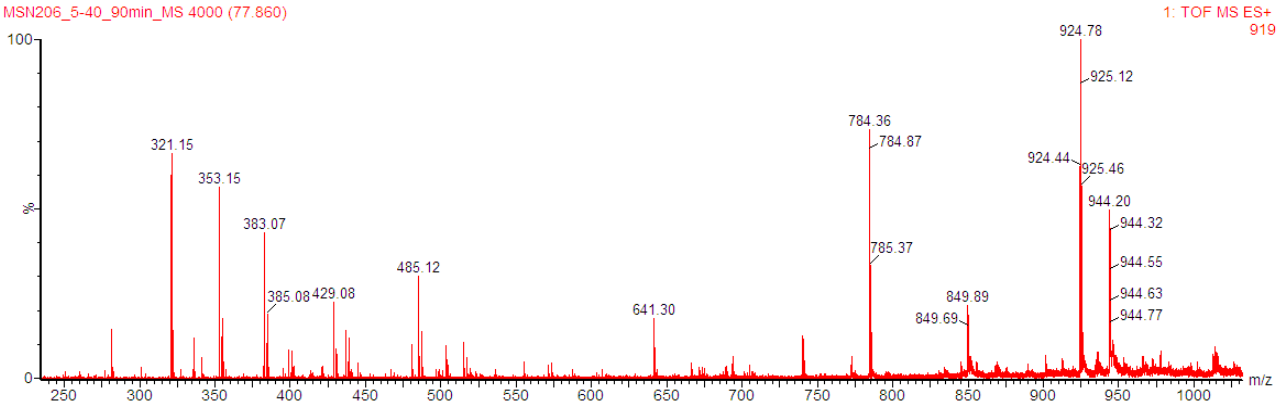

Fig. S6. ESI-QTOF MS spectrum of the peptide T5: 21GLVLIAFSQYLQQ**[Cys-CyOct-]**PFDEH VK41obtained by tryptic digestion of **14** and containing the modified Cys 34 (B= CysCyOCt). m/z. Value of the triply charged peptide = 924.47.

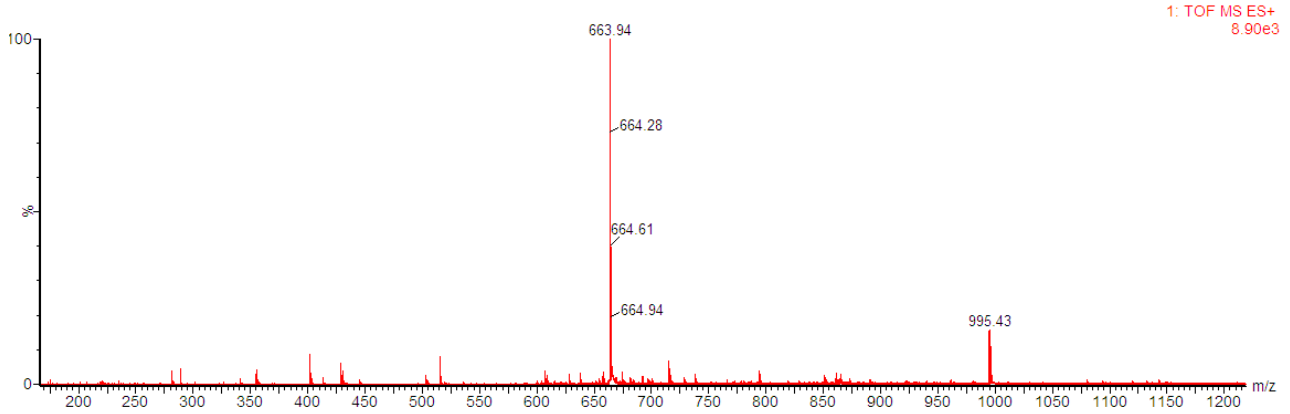

Fig. S7. ESI-QTOF MS spectrum of the peptide 65 SLHTLFGDEL**[Cys-CyOct-GSH]**K76 obtained by tryptic digestion of **13** and containing the modified Cys 75 (B= CysCyOCt-GSH). m/z. Value of the doubly charged peptide: 663.97 and the triply charged peptide = 995.46.

Table 1. MS - Analysis of tryptic digestion of compounds **13** and **14**.

Protein	Frag #	Res #	Sequence[a]	Calc. mass	Obs. mass	Charge[b]	Error (Δ)
13	**T8**	65-76	(K) SLHTLFGDEL**[Cys-CyOct-GSH]**K (V)	1988.90	995.97	2+	0.03
13	**T8**	65-76	(K) SLHTLFGDEL**[Cys-CyOct-GSH]**K (V)	1988.90	663.47	3+	0.05
14	**T5**	21-41	(K)GLVLIAFSQYLQQ**[Cys-CyOct]** PFDEHVK(L)	2770.38	924.44	3+	0.03

(a) Predicted peptide fragments containing [Cys-CyOCt] and [Cys-CyOct-GSH]were determined by MassLynx software (v. 4.0 from Waters) according to the manufacturer's instructions. (b) Charge corresponding to $[M+2H]^{2+}$or $[M+3H]^{3+}$

a	30.03	143.12	242.19	355.27	468.35	539.39	686.46	773.49	901.55	1064.61	1177.70	1305.76	1433.82	1872.97	1970.03	2117.09	2232.12	2361.16	2498.22
b*	60.04	173.13	272.20	385.28	498.37	569.40	716.47	803.50	931.56	1094.63	1207.71	1335.77	1463.83	1902.98	2000.04	2147.10	2262.13	2391.17	2528.23
c'	74.05	187.13	286.20	399.28	512.37	583.41	730.47	817.51	945.56	1108.63	1221.71	1349.77	1477.83	1916.99	2014.04	2161.11	2276.13	2405.18	2542.24
i	30.03	86.10	72.08	86.10	86.10	44.05	120.08	60.04	101.07	136.08	86.10	101.07	101.07	412.17	70.07	120.08	88.04	102.06	110.07
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
	Gly	Leu	Val	Leu	Ile	Ala	Phe	Ser	Gln	Tyr	Leu	Gln	Gln	@	Pro	Phe	Asp	Glu	His
	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3
x	-	2740.35	2627.26	2528.20	2415.11	2302.03	2230.99	2083.92	1996.89	1868.83	1705.77	1592.68	1464.63	1336.57	897.41	800.36	653.29	538.26	409.22
y*	-	2714.37	2601.29	2502.22	2389.13	2276.05	2205.01	2057.94	1970.91	1842.85	1679.79	1566.71	1438.65	1310.59	871.43	774.38	627.31	512.28	383.24
z'	-	2698.35	2585.27	2486.20	2373.11	2260.03	2188.99	2041.92	1954.89	1826.83	1663.77	1550.69	1422.63	1294.57	855.41	758.36	611.29	496.26	367.22

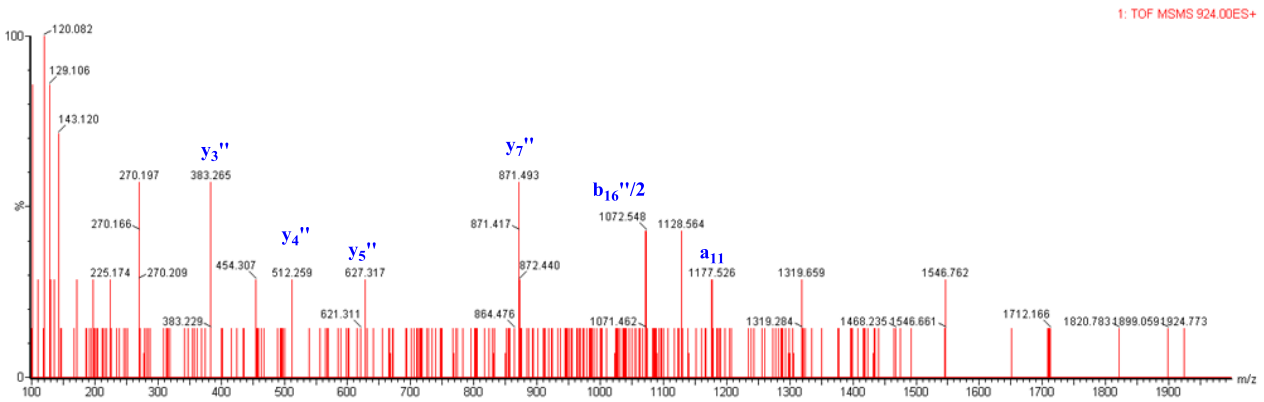

Fig. S8. ESI-QTOF MS-MS spectrum spectrum of the peptide T5: 21GLVLIAFSQYLQQ**[Cys-CyOct]**PFDEHVK41obtained by tryptic digestion of **14** and containing the modified Cys 34 (B= CysCyOCt. m/z. Value of the triply charged peptide = 924.47. The a'', b'', c', x, y'' and z' ion series of the fragments are reported.

References

29 W. C. Still, M. Kahn and A. Mitra *J. Org. Chem.*, 1978, **43**, 2923-2925.

30 *Pure Appl. Chem*. 1991, **63**, 975-990.

31 J. C. Namyslo and C. Stanitzek, *Synthesis* 2006, 3367-3369.

32 E. Wingender and A. Arellano, *Anal. Biochem.*, 1982, **127**, 351-360.

33 R. J. Kaufman and R. S. Sidhu, *J. Org. Chem.*, 1982, **47**, 4941-4947.