Revised manuscript

The advantage of 7-diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin fluorogenic tagging of sulfhydryl groups in oligopeptides for tandem mass spectrometry

Krisztina Knapp,[a] Zsuzsa Majer,[a] Gitta Schlosser[b*]

[a] *Department of Organic Chemistry, Laboratory for Chiroptical Structure Analysis, Eötvös Loránd University, Budapest, Hungary*

[b] *MTA-ELTE Research Group of Peptide Chemistry, Hungarian Academy of Sciences, Eötvös Loránd University, Budapest, Hungary*

[*] *Correspondence to:*

Gitta Schlosser, Ph.D., MTA-ELTE Research Group of Peptide Chemistry, Hungarian Academy of Sciences , Eötvös Loránd University, Budapest, 1117, Pázmány Péter sétány 1/A, Hungary. E-mail: sch@chem.elte.hu

Dear Sir,

There has been significant interest on the effects of UV-light excitation on the structure and function of proteins and peptides.[1] UV-light excitation has diverse and emerging roles in biological, medical and industrial processes. Photo-induced damage of proteins occurs via endogenous chromophores, such as tryptophan, tyrosine, phenylalanine, histidine, cysteine and/or cystine.[2] Data derived from investigation of various proteins (e.g. goat α-lactalbumin,[3] immunglobulins,[4] cutinase[5] and bovine growth hormone[6]) show that tryptophan residues may be the primary motifs in the initiation of protein photodegradation. Free cysteine thiyl radicals (CysS•) can be formed as a result of disulfide photolysis due to photoionization of the Trp residues of the polypeptide chain. However, photolytic split of disulfide bonds could strongly depend on the microenviroment of both moieties. Trp-mediated photolysis may occur at appropriate distance from the disulfide bridge. Several photolysis experiments were carried out using lower (λ=253.7 nm) and/or higher (λ=280-300 nm) wavelengths to provide evidence for disproportionation mechanism of disulfide photolysis.[7] Due to the reactivity of free thiyl radicals and of sulfhydryl groups, instant and sensitive detection of disulfide splitting is still a challenge.

Various methods and reagents have been reported in the literature for revealing free sulfhydryl moiety. Depending on the required detection limit, various thiol-reactive probes have been used for detection of sulfhydryl species, such as Ellman's reagent,[8] iodoacetamide- and maleimide-derivatives.[9] Among these probes, several fluorogenic agents are available for derivatization followed by detection. Fluorescent probes provide sensitive methodology for indirect detection of free sulfhydryl groups; however, these reagents are generally not suitable for the localization of cysteine residues in the polypeptide sequence. 7-diethylamino-3-(4'-maleimidophenyl)-4-methylcoumarin (CPM)[10] could be considered as a fast and very sensitive fluorogenic reagent for capturing free sulfhydryl groups. The CPM-peptide/protein derivative formed in the addition reaction has a measurable, strong fluorescence, while the CPM reagent does not show any fluorescence in this wavelength

3

region (λ_{ex}=384 nm, λ_{em}=475 nm). Thus, the increased fluorescence intensity can be utilized for quantitative determination of thiols.

The aim of our work was to investigate the effect of CPM tagging on peptide MS/MS fragmentation behaviour, because a combination of fluorescent labelling and peptide sequencing with tandem mass spectrometry could be an excellent tool for a sensitive sulfhydryl detection as well as localization of the free Cys sulfhydryl group. Therefore, in this study, a set of Cys containing model oligopeptides and their fluorescent derivatives with CPM tags were synthesized and analysed by tandem mass spectrometry (Figure 1).

Fmoc-Cys(Trt)-COOH was obtained from Reanal (Budapest, Hungary), all other amino acid derivatives and Fmoc-Rink-Amid MBHA resin were from Iris Biotech GmbH (Marktredwitz, Germany). 7-diethylamino-3-(4'-maleimidylphenyl)-4-ethylcoumarin (CPM) was purchased from Sigma-Aldrich (Saint Louis, USA). Acetonitrile (ACN) and acetic acid were analytical grade and were purchased from Molar Chemicals (Budapest, Hungary). Water was purified and deionized using Milli-Q system from Millipore (Billerica, MA, USA).

Synthetic linear pentapeptide derivatives, Ac-CXAKC-NH$_2$, Ac-CXAKC(Acm)-NH$_2$ and Ac-C(Acm)XAKC-NH$_2$ (where X: Trp or Phe; Acm: acetamidomethyl) were prepared on solid phase using Fmoc / ^{t}Bu chemistry.[11] After cleavage from the resin with trifluoroacetic acid in the presence of scavengers (triisopropylsilane, H$_2$O, phenol, 1,2-ethanedithiol) the crude products were lyophilized and peptides were purified by RP-HPLC on a C$_{18}$-silica preparative column (5 μm, 150 mm x 21.2 mm, 110 Å, Gemini-NX, Phenomenex).

Linear (Ac-CWAKC-NH$_2$, and Ac-CFAKC-NH$_2$) peptides were oxidized without further purification under mild alkaline conditions (0.1-0.5 mg/mL, pH 7.8-8.0 where the pH was adjusted by NH$_4$HCO$_3$). The cyclization reaction was monitored by analytical RP-HPLC (5 μm, 150 mm x 4.6 mm, 110 Å, Gemini-NX, Phenomenex). After lyophilisation, the crude cyclic peptides were purified by preparative RP-HPLC. All purified peptides were shown >95% homogeneity by analytical RP-HPLC and ESI-MS. Characteristics of the synthesized peptides are summarized in Supplementary Table 1.

CPM-labelled peptides were synthesized by Michael-addition. Linear peptide derivatives were incubated at 37°C for maximum 24 hrs with 2 mol equivalent of CPM reagent in dimethyl sulfoxide. Completion of reactions was monitored by analytical RP-HPLC (3.5 μm, 100 mm x 3.0 mm, Zorbax SB-C18). The CPM-peptide conjugates were

finally purified by semi-preparative RP-HPLC (5 μm, 250 mm x 10.0 mm, Alltima C18) and characterized by analytical RP-HPLC and ESI-MS. The characteristics of the CPM-peptide derivatives are summarized in Supplementary Table 2.

Mass spectrometric experiments were performed on a Bruker Esquire 3000+ ion trap mass spectrometer (Bruker, Bremen, Germany) equipped with electrospray ionisation (ESI) source. Samples were dissolved in a mixture of ACN/water = 1/1 (V/V), containing 0.1% acetic acid and introduced into the ESI source by a syringe pump with a flow rate of 10 μl/min.

In summary, for this study, a series of linear peptides and their CPM-labelled analogues, as well as the cyclic peptides bearing disulfide bridges were synthesized based on the model sequence: Ac-CXAKC-NH$_2$, where X stands for tryptophan or phenylalanine (Supplementary Tables 1. and 2.). These two aromatic residues could play a crucial role in the UV-light induced split of the disulfide bridge. To study the role of tryptophan (or other aromatic amino acids) in the photolysis of proteins, cyclic peptides are useful models. The CPM adducts of the evolved sulfhydryl groups upon illumination of model cyclic peptides can be detected for example with fluorescence or HPLC. Model peptides chosen in this work contain a single Trp (or Phe) residue in the sequences nearby aliphatic amino acid residues. A series of linear peptides (precursors of cyclic models) and their CPM-labelled analogues were synthesized, as well as the cyclic analogues. To avoid recombination and also side reactions related to the sulfhydryl group, in these experiments, one of the cysteine side-chains was protected by acetamidomethyl (Acm) group (Figure 1.; Supplementary Tables 1 and 2). These peptides were subjected to collision-induced dissociation in an ion trap mass spectrometer.

Sequence fragmentation of cyclic peptides is regulated by the presence of the disulfide bridge: the results reported here show that internal fragments and disulfide-linked fragments can be detected in the MS/MS spectra, together with intensive b_2, b_3, b_4 fragment ions and neutral losses (Supplementary Figure 1). MS/MS spectra of the corresponding linear Ac-CXAKC-NH$_2$ peptides are simpler, b_2, b_3, b_4 fragment ions and neutral losses are observed (Supplementary Figure 1). Linear peptides bearing Acm protecting group (Ac-C(Acm)XAKC-NH$_2$ and Ac-CXAKC(Acm)-NH$_2$) show analogous spectra compared to the unprotected linear compounds; the cleavage of the Acm group from the cysteine side chain is detected even at low collision energy, however this phenomenon does not affect the

fragmentation of the peptides in general (Supplementary Figure 2). The loss of the Acm group is detected as a 71 mass difference.

Peptides bearing CPM tags and Acm protecting groups (Ac-C(Acm)XAKC(CPM)-NH$_2$ and Ac-C(CPM)XAKC(Acm)-NH$_2$) exhibit a highly different fragmentation pattern (Figure 2.; Supplementary Figure 3.). Interestingly, the fragmentation in this case is driven by the presence/absence of the CPM group: fragment ions bearing the CPM moiety are the most prominent in the MS/MS spectra. In the case of [1]Cys labelling, *a* and *b* ion series can be detected (Figure 2B). In the case of [5]Cys labelling, *y* ion series is dominant in the MS/MS spectrum (Figure 2A). Cleavage of the CPM group is also observed with characteristic marker ion at *m/z* 435, corresponding to the CPM unit bearing the sulfur atom of the cysteine side chain. Linear pentapeptides in which both cysteine residues are modified with CPM tags (Ac-C(CPM)XAKC(CPM)-NH$_2$) show a characteristic combination of sequence fragments in the MS/MS spectrum of the singly protonated peptides: CPM-tagged *a*, *b* and *y* ion series are detected with high intensity, and the cleaved CPM marker ion could be observed as well (Figure 2C). Due to its positive effect on peptide fragmentation, fluorogenic tagging with 7-diethylamino-3-(4'-maleimidophenyl)-4-methylcoumarin combined with tandem mass spectrometry could be considered as a promising choice to catch free sulfhydryl groups in oligopeptides and perhaps in proteins. Characteristic profile of the MS/MS fragments in combination with the detection of the cleaved CPM marker ions could allow an optimal methodology for sulfhydryl tracking with high sensitivity both in singly or double labelled peptides.

Taken together, localization of cysteine residues bearing free sulfhydryl groups with high sensitivity is a challenging task due to the spontaneous recombination of disulfide bridges in peptides and in proteins. In this study, we demonstrated that CPM tagging combined with tandem mass spectrometry could have multiple advantages to monitor and identify sulfhydryl groups in peptides. A unique reporter ion (m/z 435) containing CPM moiety, as well as characteristic sequence ions bearing the CPM tag are generated by CID during MS analysis, offering an excellent tool for a sensitive localization of sulfhydryl groups. It is perhaps important to note, that the application of CPM reagent in proteomic studies might have some limitations as well, due to the large size of the coumarin ring and possible steric restrictions in the case of hidden functional groups.

Yours,

Krisztina Knapp,[a] Zsuzsa Majer,[a] Gitta Schlosser[b]

[a] *Department of Organic Chemistry, Laboratory for Chiroptical Structure Analysis, Eötvös Loránd University, Budapest, Hungary*

[b] *MTA-ELTE Research Group of Peptide Chemistry, Hungarian Academy of Sciences, Eötvös Loránd University, Budapest, Hungary*

Acknowledgements:

TABLES AND FIGURES LEGENDS:

Figure 1. Scheme of the model peptides, where X stands for tryptophan or phenylalanine. *A.)* unlabelled peptides, *B.)* CPM labelled peptides. Please note that cyclopeptides are referred to Ac-c(CXAKC)-NH$_2$ throughout the text.

Figure 2. MS/MS spectrum of the singly protonated, CPM-labeled peptides: *a.)* Ac-C(Acm)WAKC(CPM)-NH$_2$, *b.)* Ac-C(CPM)WAKC(Acm)-NH$_2$, *c.)* Ac-C(CPM)WAKC(CPM)-NH$_2$. -Acm is used for the loss of the acetamidomethyl group (71 mass difference), ^{o} label is used for 18 (H$_2$O) loss.

Supplementary material:

Supplementary table 1. Analytical characteristics of unlabelled peptides.

Supplementary table 2. Analytical characteristics of CPM-labelled peptides.

Supplementary figure 1. MS/MS spectrum of the singly protonated *a.)* Ac-c(CWAKC)-NH$_2$ cyclopeptide, *b.)* the linear Ac-CWAKC-NH$_2$ peptide, *c.)* Ac-c(CFAKC)-NH$_2$ cyclopeptide and *d.)* the linear Ac-CFAKC-NH$_2$ peptide. Internal fragments are labelled with the corresponding 1 letter amino acid codes. * label is used for 17 (NH$_3$) loss, ^{o} label is used for 18 (H$_2$O) loss. ( ) brackets are used for indication of disulfide-linked fragment ions.

Supplementary figure 2. MS/MS spectrum of the singly protonated peptides: *a.)* Ac-C(Acm)WAKC-NH$_2$ *b.)* Ac-CWAKC(Acm)-NH$_2$. *c.)* Ac-C(Acm)FAKC-NH$_2$ *d.)* Ac-CFAKC(Acm)-NH$_2$.-Acm is used for the loss of the acetamidomethyl group (71 mass difference), * label is used for 17 (NH$_3$) loss, ^{o} label is used for 18 (H$_2$O) loss.

Supplementary figure 3. MS/MS spectrum of the singly protonated, CPM-labeled peptides: *a.)* Ac-C(Acm)FAKC(CPM)-NH$_2$, *b.)* Ac-C(CPM)FAKC(Acm)-NH$_2$, *c.)* Ac-C(CPM)FAKC(CPM)-NH$_2$. -Acm is used for the loss of the acetamidomethyl group (71 mass difference), * label is used for 17 (NH$_3$) loss.
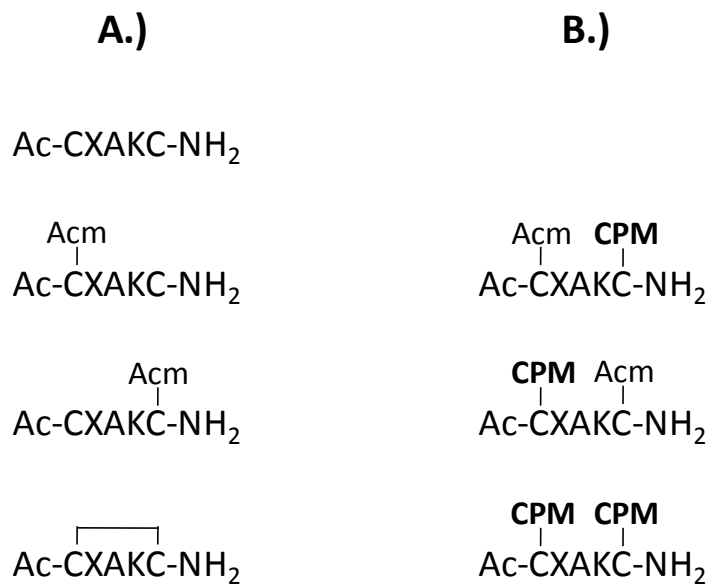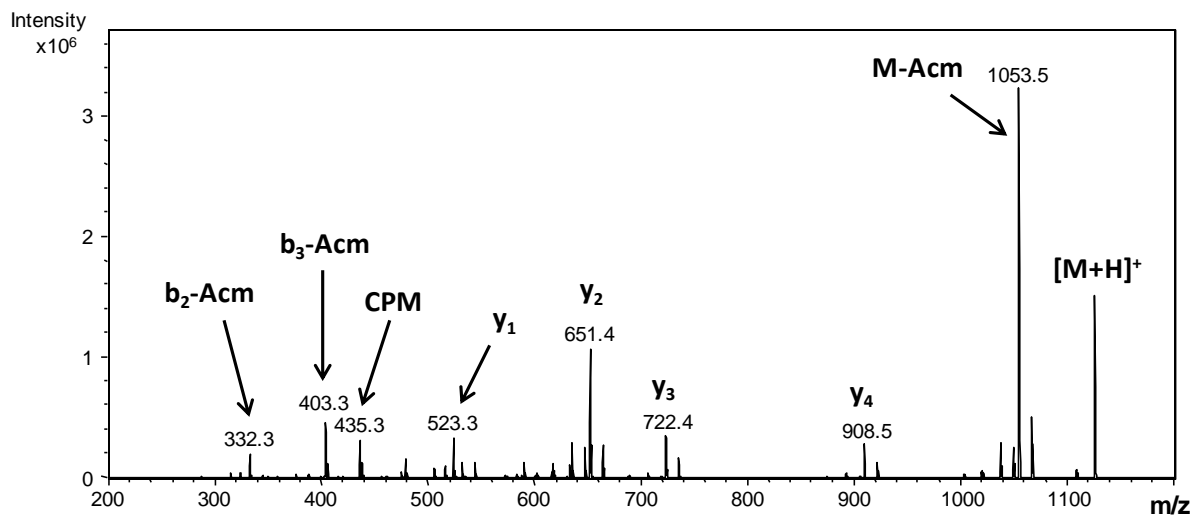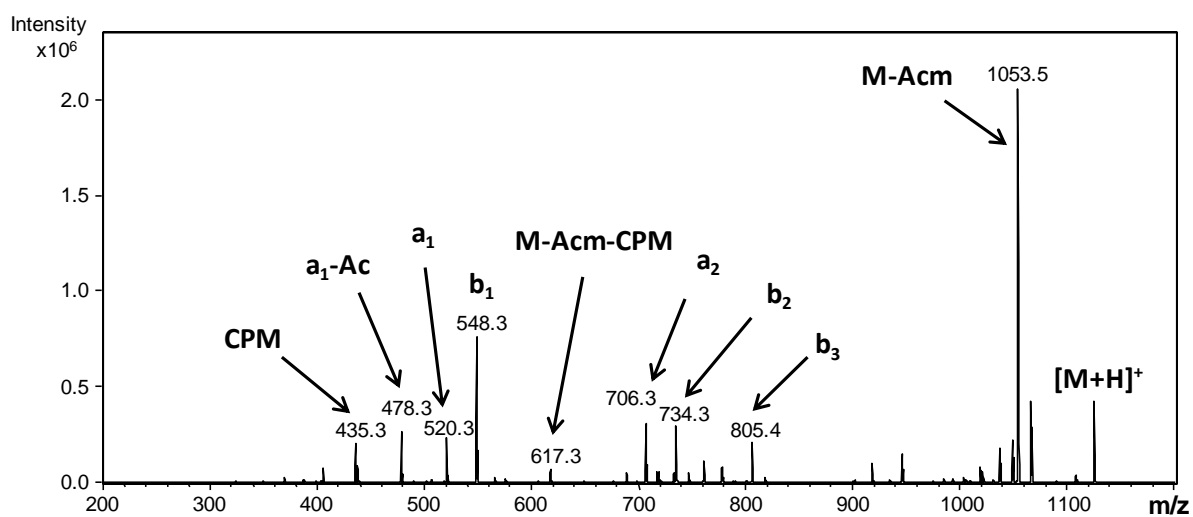
FIGURES:

Figure 1.

A.)   **B.)**

Ac-CXAKC-NH$_2$

    Acm   Acm **CPM**
    | | |
Ac-CXAKC-NH$_2$ Ac-CXAKC-NH$_2$

        Acm   **CPM** Acm
        | | |
Ac-CXAKC-NH$_2$ Ac-CXAKC-NH$_2$

                   CPM CPM
    ┌──────┐ | |
Ac-CXAKC-NH$_2$ Ac-CXAKC-NH$_2$

Figure 2.

A.)

B.)

C.)
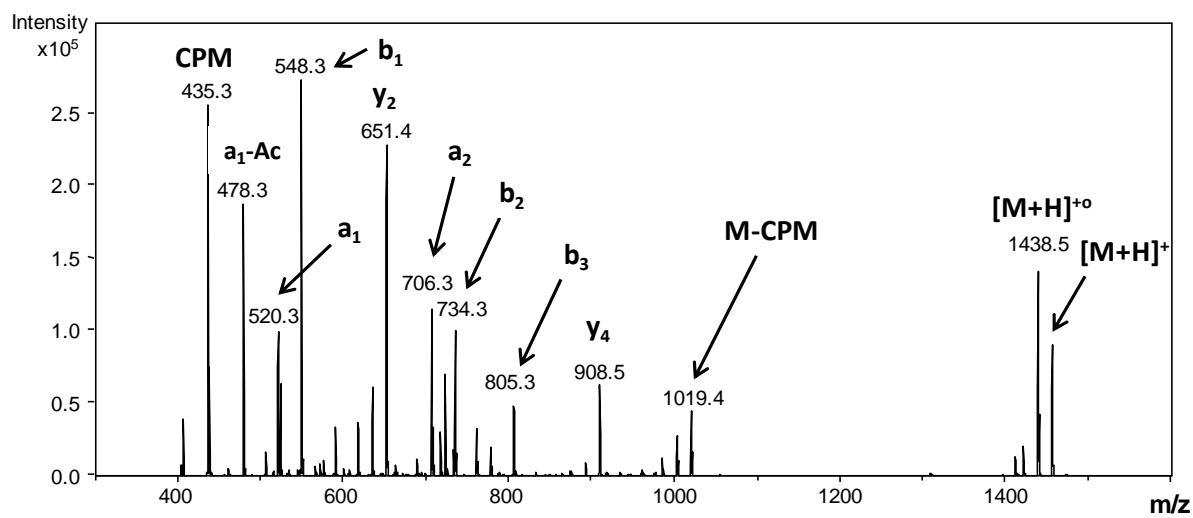
Supplementary table 1.:

Compounds	HPLC retention time (min) [a]	Molecular mass, theoretical [b]
Ac-CWAKC-NH$_2$ (**1**)	21.5	650.3
Ac-C(Acm)WAKC-NH$_2$ (**2**)	16.4	721.3
Ac-CWAKC(Acm)-NH$_2$ (**3**)	16.4	721.3
Ac-c(CWAKC)-NH$_2$ (**4**)	18.0	648.2
Ac-CFAKC-NH$_2$ (**5**)	18.2	611.3
Ac-C(Acm)FAKC-NH$_2$ (**6**)	17.1	682.3
Ac-CFAKC(Acm)-NH$_2$ (**7**)	17.1	682.3
Ac-c(CFAKC)-NH$_2$ (**8**)	16.1	609.2

a) Column type: 5 μm, 150 mm x 4.6 mm, Gemini NX
b) Monoisotopic molecular mass; measured masses were within 0.3 Da of the theoretical ones.

Supplementary table 2.:

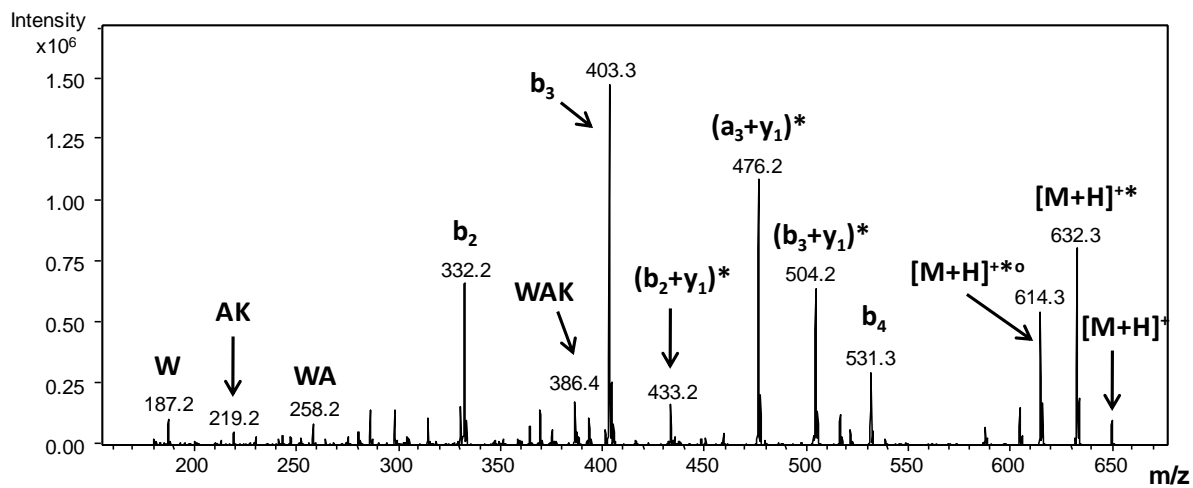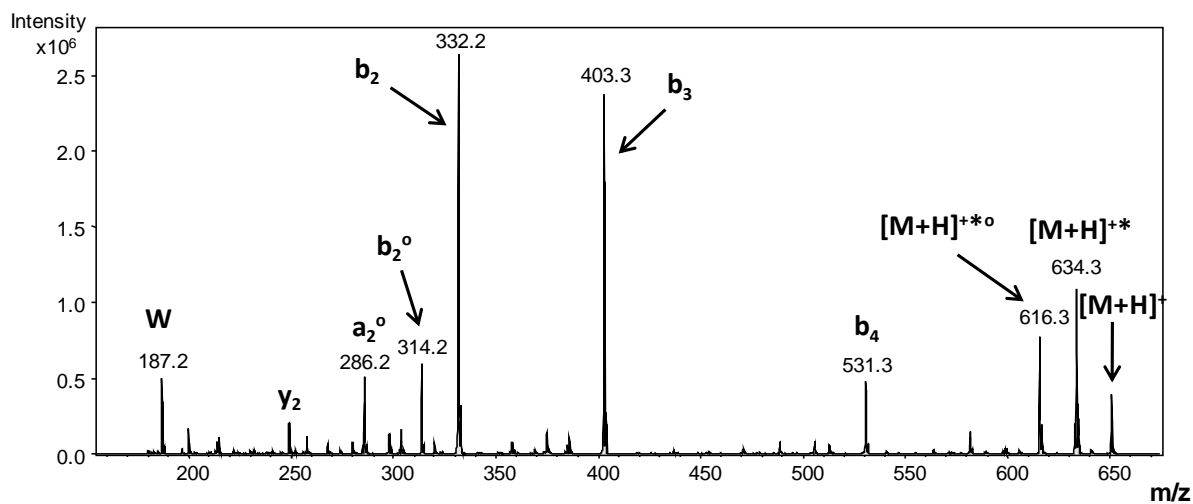| CPM-labelled compounds	HPLC retention time (min)	Molecular mass, theoretical [c]
Ac-C(CPM)WAKC(CPM)-NH$_2$ (**1A**)	16.0 [b]	1454.6
Ac-C(Acm)WAKC(CPM)-NH$_2$ (**2A**)	26.6 [b]	1123.5
Ac-C(CPM)WAKC(Acm)-NH$_2$ (**3A**)	26.6 [b]	1123.5
Ac-C(CPM)FAKC(CPM)-NH$_2$ (**5A**)	34.1 [a]	1415.6
Ac-C(Acm)FAKC(CPM)-NH$_2$ (**6A**)	26.1 [a]	1084.4
Ac-C(CPM)FAKC(Acm)-NH$_2$ (**7A**)	26.1 [a]	1084.4

a) Column type: 5 μm, 150 mm x 4.6 mm, Gemini NX
b) Column type: 3.5 μm, 100 mm x 3.0 mm, Zorbax SB-C18
c) Monoisotopic molecular mass; measured masses were within 0.3 Da of the theoretical ones.
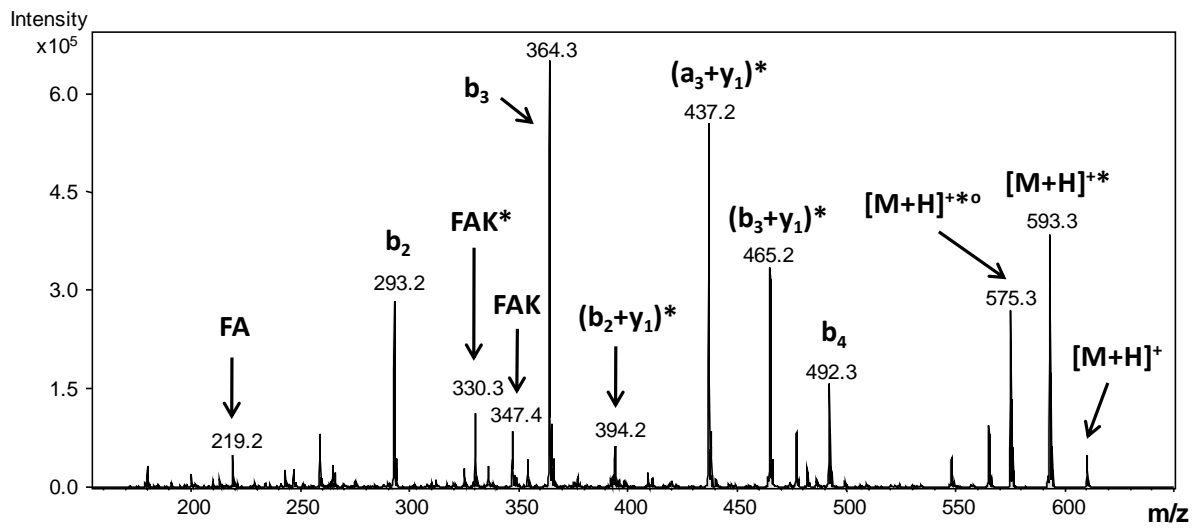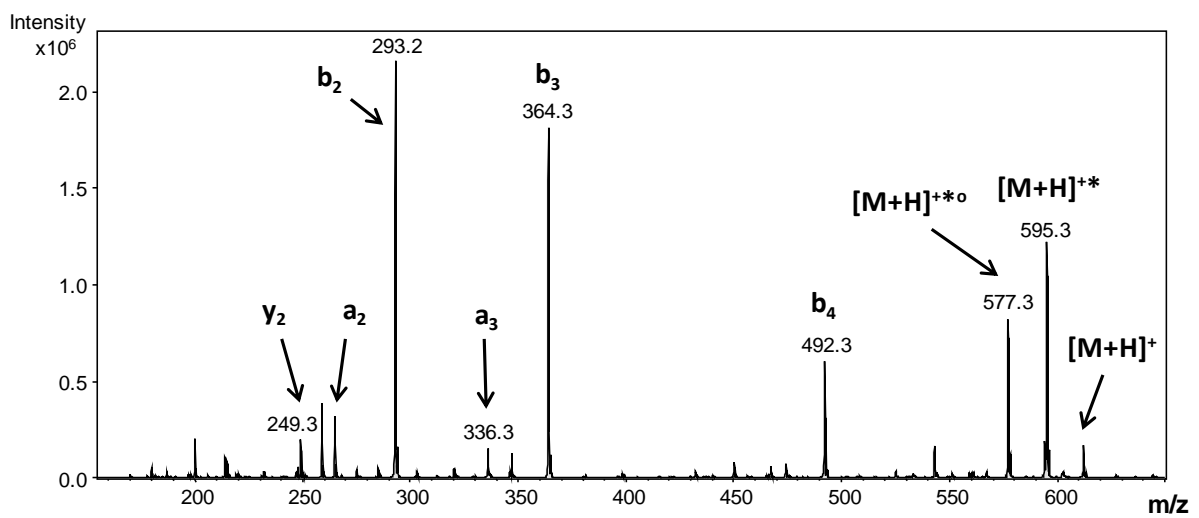
Supplementary figure 1.

A.)

B.)

C.)
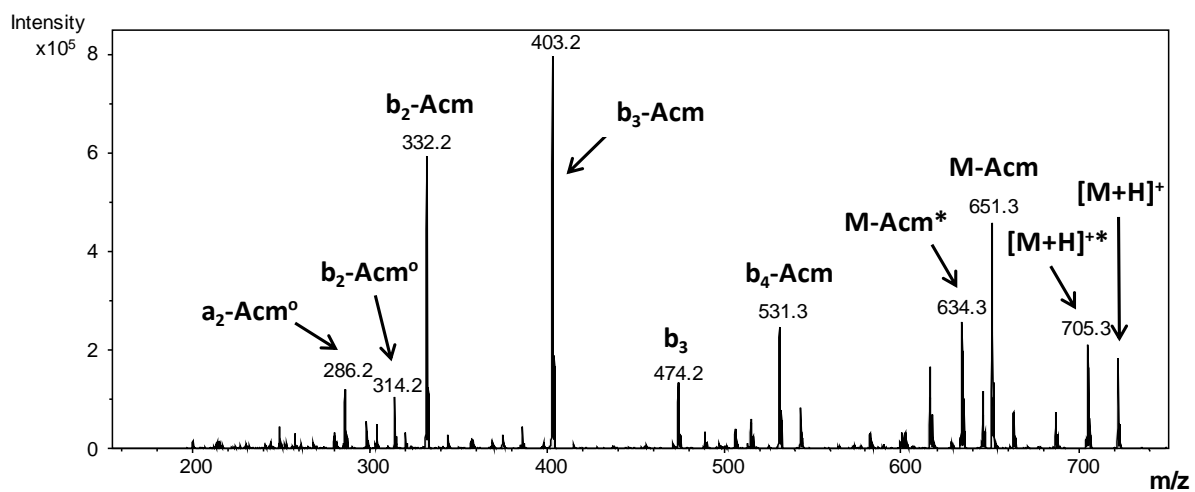

D.)

Supplementary figure 2.

A.)

B.)

C.)

Intensity
x10⁵

b₂-Acm
293.3

b₃-Acm

b₄-Acm
492.4

M-Acm*
M-Acm*ᵒ
577.4

595.3

M-Acm
612.4

[M+H]⁺*
666.3

[M+H]⁺

364.3

5
4
3
2
1
0

200   250   300   350   400   450   500   550   600   650   700   **m/z**

D.)

Intensity
x10⁶

b₂
293.3

b₃
364.3

b₄
492.4

M-Acm
612.4

[M+H]⁺

1.2
1.0
0.8
0.6
0.4
0.2
0.0

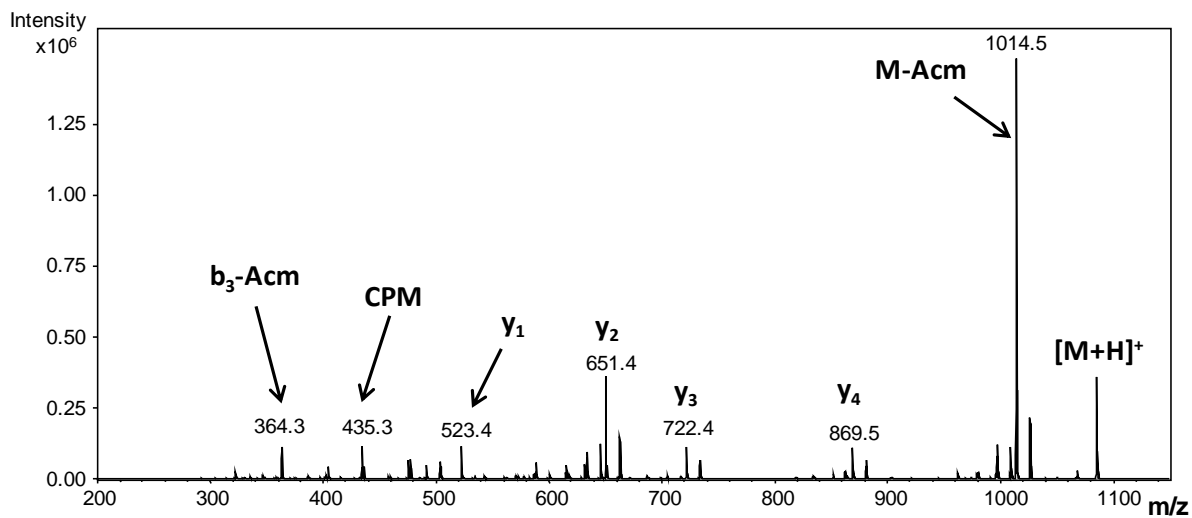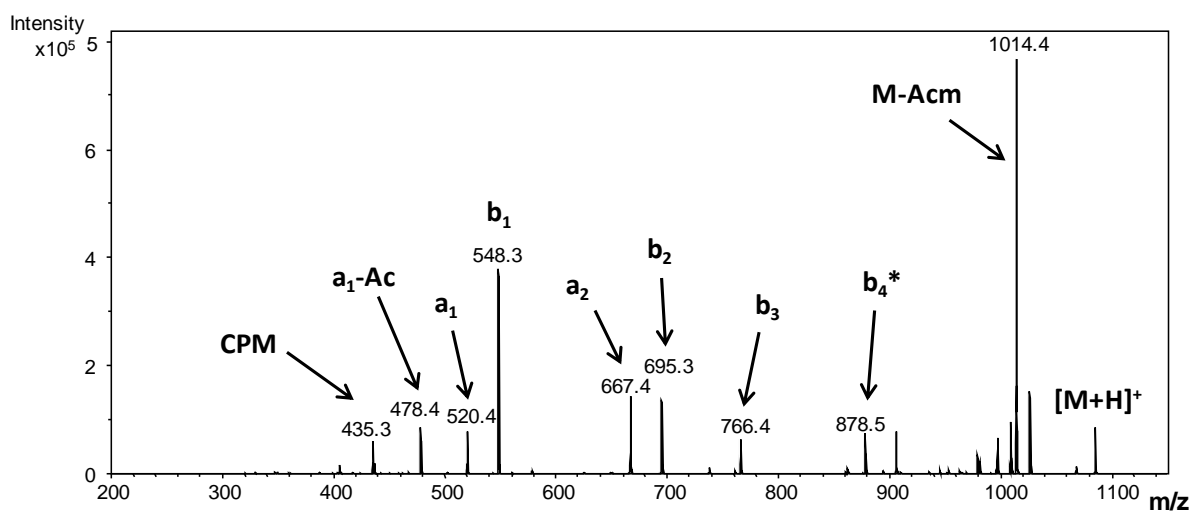200   250   300   350   400   450   500   550   600   650   700   **m/z**
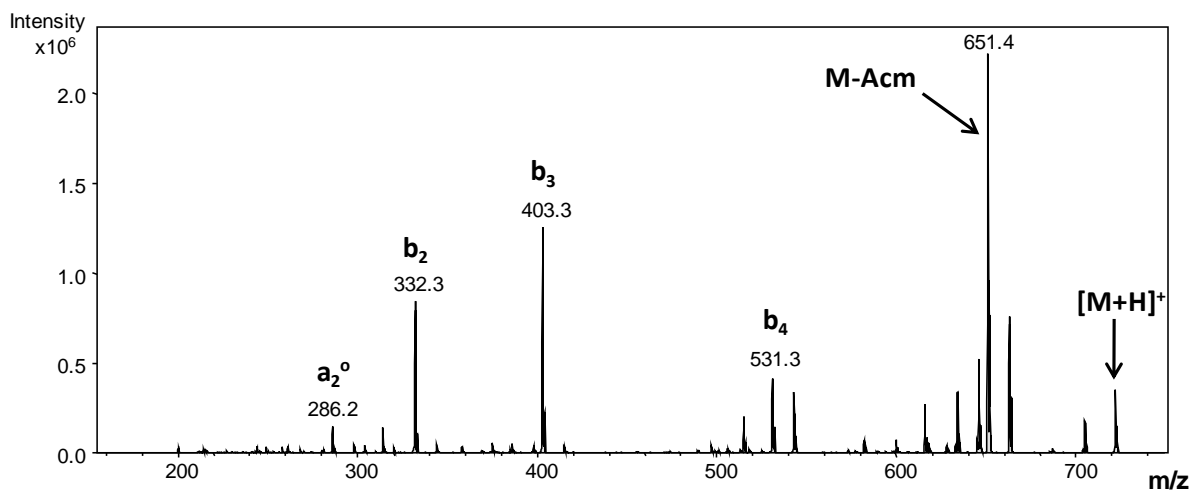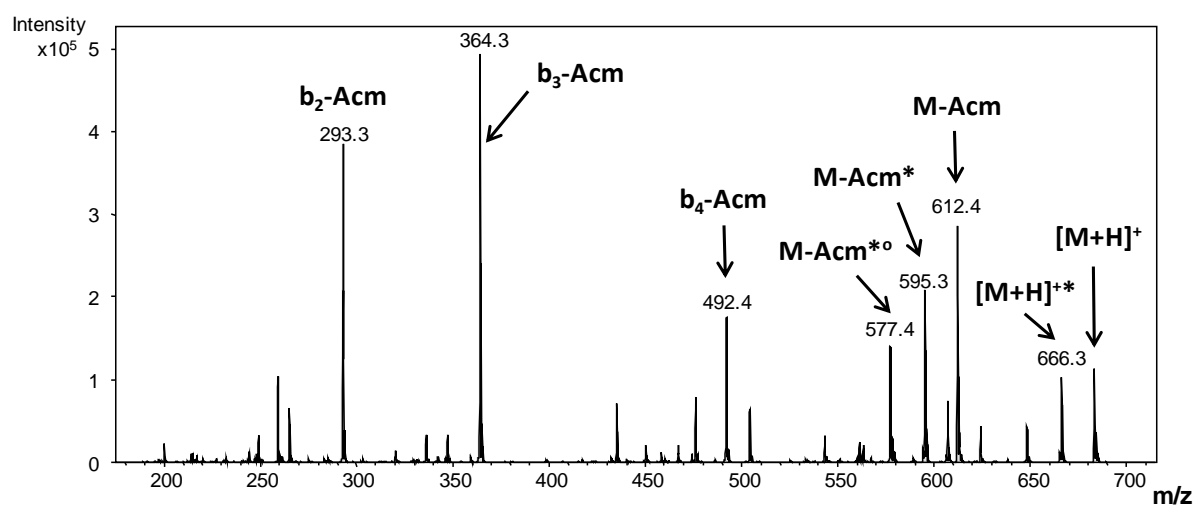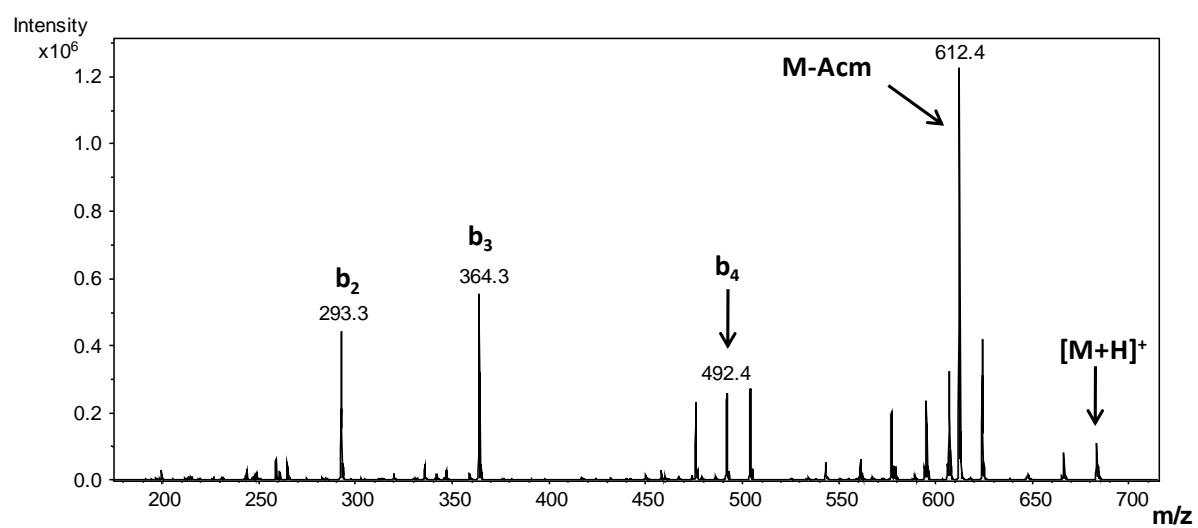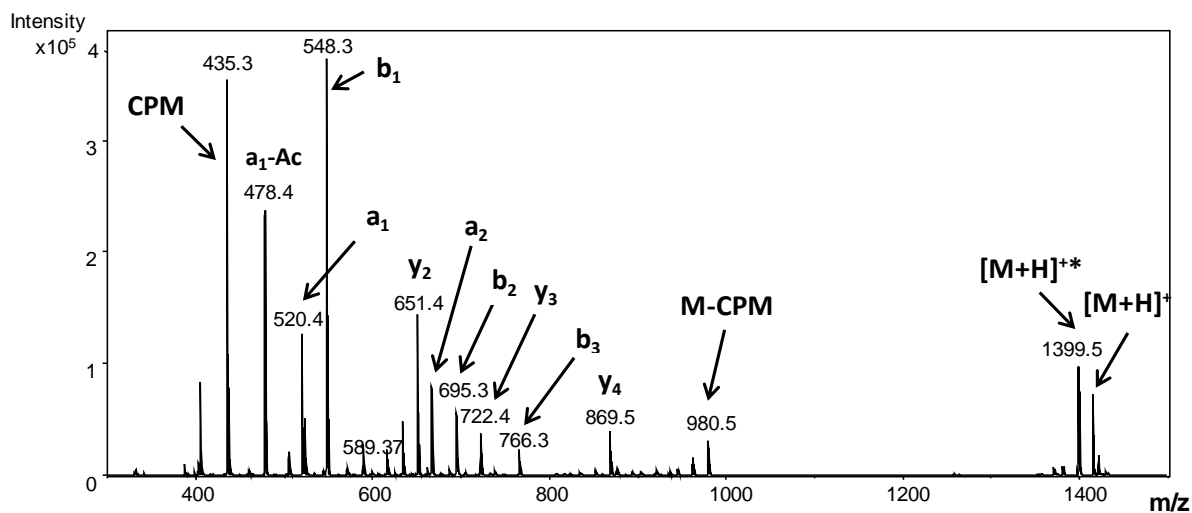
Supplementary figure 3.

A.)

B.)

C.)

REFERENCES:

[1] J. Chen, P.R. Callis, J. King. Mechanism of the very efficient quenching of tryptophan fluorescence in human γD- and γS-Crystallins: The γ-Crystallin fold may have evolved to protect tryptophan residues from ultraviolet photodamage. *Biochemistry* **2009**, *48*, 3708.

[2] D.I. Pattison, A.S. Rahmanto and M.J. Davies. Photo-oxidation of proteins. *Photochem. Photobiol. Sci.* **2012**, *11*, 38.

[3] A. Vanhooren, B. Devreese, K. Vanhee, J. Van Beeumen, and I. Hanssens. Photoexcitation of tryptophan groups induces reduction of two disulfide bonds in goat α-lactalbumin. *Biochemistry* **2002**, *41*, 11035.

[4] E. Illyés, S. Staelens, A. Vanhooren, H. Deckmyn, I. Hanssens and Z. Majer. Role of the Trp-disulfide Triads in the UV Light Induced Degradation of a Monoclonal Antibody scFv. *Int. J. of Biochem. Res. Rev.* **2014**, *4*, 367.

[5] M.T. Neves-Petersen, Z. Gryczynski, J. Lakowicz, P. Fojan, S. Pedersen, E. Petersen, B.S. Petersen. High probability of disrupting a disulphide bridge mediated by an endogenous excited tryptophan residue. *Protein Sci.* **2002**, *11*, 588.

[6] B.L. Miller, M.J. Hageman, T.J. Thamann, L.B. Barron, C.J. Schöneich. Solid-state photodegradation of bovine somatotropin (bovine growth hormone): evidence for tryptophan-mediated photooxidation of disulfide bonds. *Pharm. Sci.* **2003**, *92*, 1698.

[7] O. Mozziconacci, B.A. Kerwin and C. Schöneich. Photolysis of an intrachain peptide disulfide bond: primary and secondary processes, formation of H_2S, and hydrogen transfer reactions. *J. Phys. Chem. B* **2010**, *114*, 3668.

[8] G.L. Ellman. Tissue sulphhydryl groups. *Arch. Biochem. Biophys.* **1959**, *82*, 70.

[9] E.M. Sletten, C.R. Bertozzi. Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. *Angew. Chem. Int. Ed.* **2009**, *48*, 6974.

[10] F.C. Ayers, G.L. Warner, K.L. Smith, D.A. Lawrence. Fluorometric quantitation of cellular and nonprotein thiols. *Anal. Biochem.* **1986**, *156,* 524.

[11] F. Albericio. *Solid-Phase Synthesis: A Practical Guide* (1st edition). CRC Press, Boca Raton, **2000**, 848.