Supplementary Information (SI) for Analyst.

Supplementary Figures

Characterization by LC-MS/MS analysis of KLH vaccine conjugated with a tick antigen peptides.

Authors: Satomy Pousa[a,*], Pablo E. Ramos-Bermúdez[b], Vladimir Besada[a], Ania Cabrales-Rico[c], Osmany Guirola Cruz[b], Hilda Elisa Garay[d], Alina Rodríguez-Mallón[e], Katharina Zettl[f], Jacek R Wiśniewski[f] and Luis Javier González[a,*, §].

(*) Authors contributed equally to this work.

(§) Corresponding Author

Affiliations:

[a]Departments of Proteomics, Mass Spectrometry Laboratory, [b]Bioinformatics, [c]Purification and Analytic Group, [d]Laboratory of Peptide Synthesis and [e]Animal Biotechnology. Center for Genetic Engineering and Biotechnology. 31 Avenue, Cubanacan, Playa, Havana, Cuba.

[f]Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, 82152 Munich, Germany

Email:

Satomy Pousa (satomy.pousa@cigb.edu.cu) / 0000-0001-8365-9844
Pablo E. Ramos-Bermúdez (pablo.ramos@cigb.edu.cu) / 0000-0002-8439-984X
Vladimir Besada (vladimir.besada@cigb.edu.cu) / 0000-0003-3584-7714
Ania Cabrales-Rico (ania.cabrales@cigb.edu.cu) / 0000-0002-9868-1830
Osmany Guirola: (osmany.guirola@cigb.edu.cu) / 0000-0001-9987-3394
Hilda Elisa Garay (hilda.garay@cigb.edu.cu) / 0000-0003-3101-255X
Alina Rodríguez-Mallón (alina.rodriguez@cigb.edu.cu) / 0000-0002-6950-6793
Katharina Zettl (zettl@biochem.mpg.de) / 0000-0002-8103-2711
Jacek R Wiśniewski (jwisniew@biochem.mpg.de) / 0000-0002-8452-5095
Luis Javier González (luis.javier@cigb.edu.cu) / 0000-0002-8875-3642

Index

Fig. S1. Synthesis the KLH-Cys[1]pP0 conjugates using the maleimide-thiol chemistry [1]. KLH represents the carrier proteins from *M. crenulate:* KLH1 (Q10583) and KLH2 (Q10584). pP0 represents the sequence of the antigen peptide (H-AAGGGAAAAKPEESKKEEAK-NH_2) derived from *Rhipicephalus sp*. ticks, it was obtained by solid phase peptide synthesis and a Cys residue was added at the *N*-terminal end (Cys[1]pP0) [2]. BMPS is a hetero-bifunctional crosslinker (*N*-(β-maleimidopropyloxy) succinimide ester) that is able to crosslink primary amino groups in the carrier protein (Lys residues and the *N*-terminal end) with the free thiol group present in the sequence of Cys[1]pP0 peptide. In a first step, the carrier proteins KLH1 and KLH2 are activated with BMPS, introducing multiple maleimide groups at the lysine residues and the *N*-terminal end. In a second step, the activated carrier proteins react with pP0 analogue containing a Cys residue at the *N*-terminal end (Cys[1]pP0), linking several copies of the antigen in the resultant conjugate vaccine.

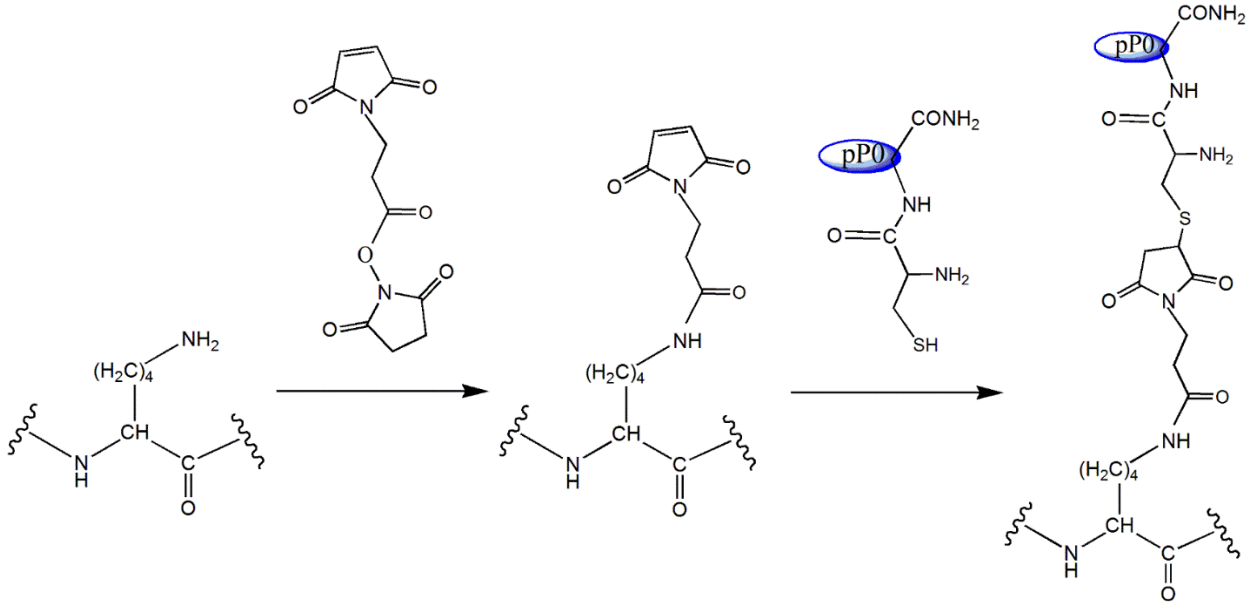


Fig. S2. The structure I represents the *N*-propionyl succinimide linker, named in this manuscript as thiosuccinimide linker. Through a pseudopeptide and a thioether bonds, it links the primary amino groups in the carrier protein with the free thiol group of the *N*-terminal cysteine residue in the antigenic Cys[1]pP0 peptide. In all structures the pseudopeptide bond between Lys side chain and the thiosuccinimide linker can be fragmented to yield the P and C+L fragment ions [3, 4] where "P" and "C" represent the Lys- and Cys-containing peptides derived from the carrier protein and the Cys-containing peptides, respectively. "L" represents the thiosuccinimide linker. The six-membered ring transcyclized linker (structure II) is generated by transcyclization [5] and contains an additional pseudopeptide bond (highlighted in red) that can be fragmented yielding two fragment ions P+71 and C+80 [6-8]. The thiosuccinimide linker (structure I) can be hydrolyzed through two pathways, "a" and "b" to yield two isomeric succinamic acid thioethers (structures III and IV). The thioether bond of cysteine residue in the hydrolyzed linker can be also fragmented in gas phase to yield the P+203 and C-34 fragment ion [9]. The structure of all the fragment ions is shown in **Fig. S3-4**.

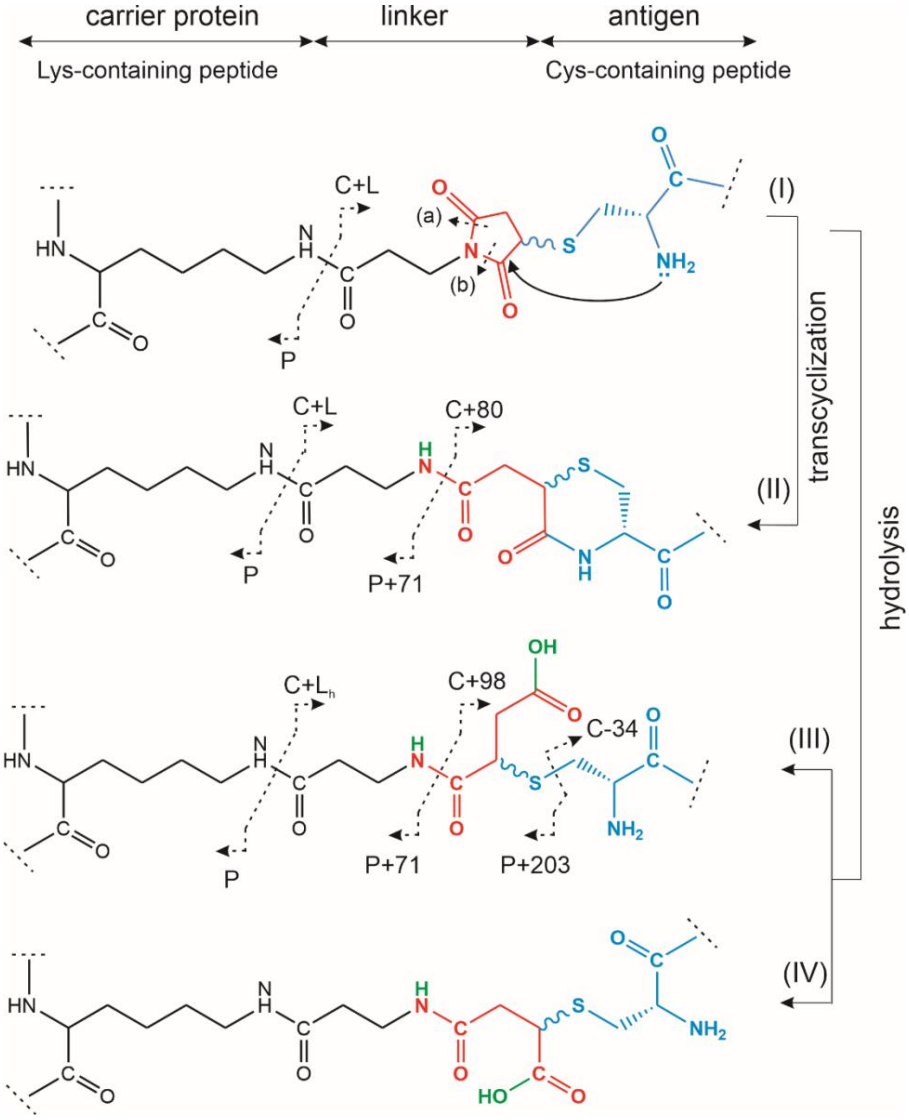


Fig. S3. The structure shown in **(a)** correspond to the five membered ring transcyclized linker. The ions named as P and C+L shown in **(b)** and **(c)**, respectively are common for the structures of thiosuccinimide and transcyclized linkers. The structure of the ion named P+71 shown in **(d)** is generated by the fragmentation of the pseudopeptide bond generated during the transcyclization reaction and the conversion of thiosuccinimide into transcyclized. The ion P+71 is also common for the hydrolyzed thiosuccinimide linker as shown in **Fig. S4**. The ion named here as C+80 and shown in **(e)** is exclusive of transcyclized linker. The P and P+71 fragment ions provide information on the molecular mass of the proteolytic peptides derived from the carrier protein present in the identified type 2 peptide. The C+L and C+80 fragment ions provide information on the molecular mass of the proteolytic peptides derived from $Cys^{1}pP0$ present in the identified type 2 peptide. The molecular masses of the expected C+L and C+80 are shown in **Table S2** and **S3**.

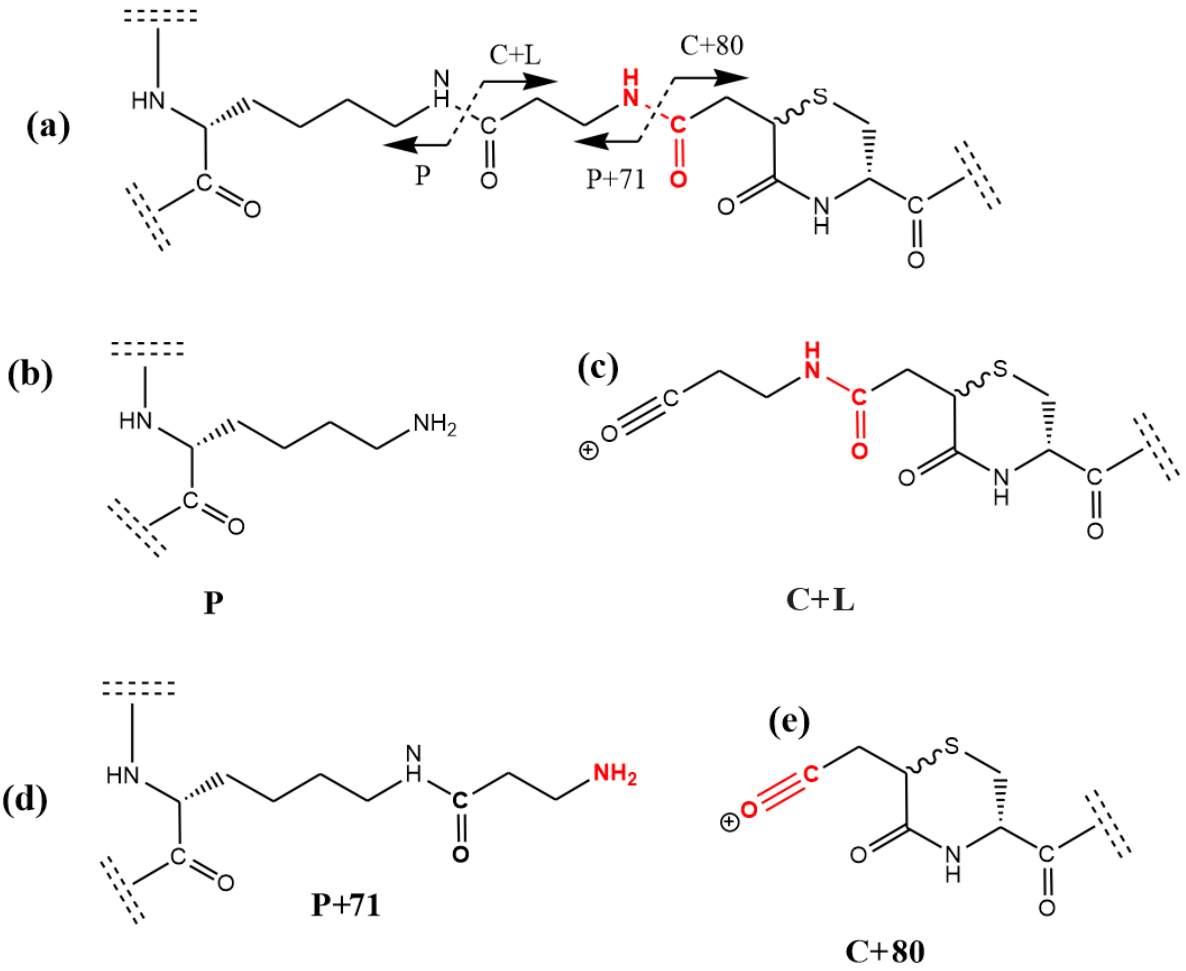
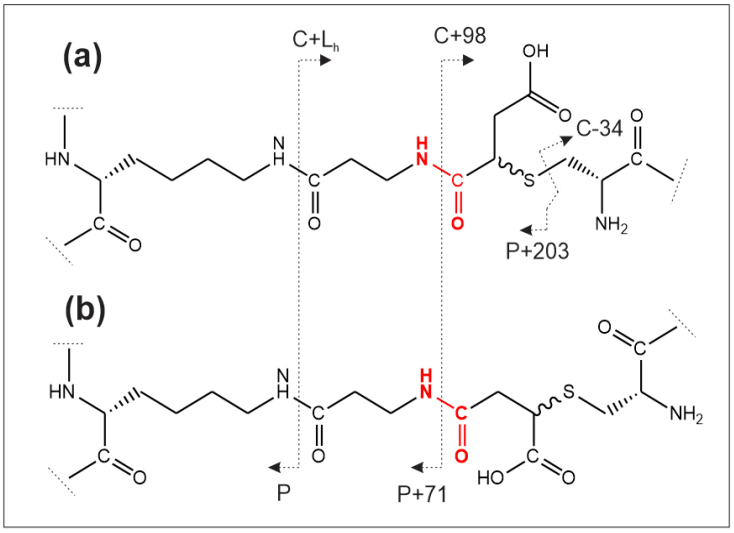
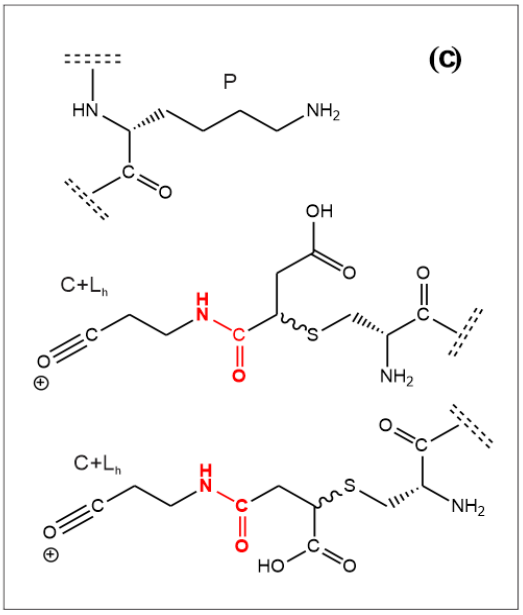
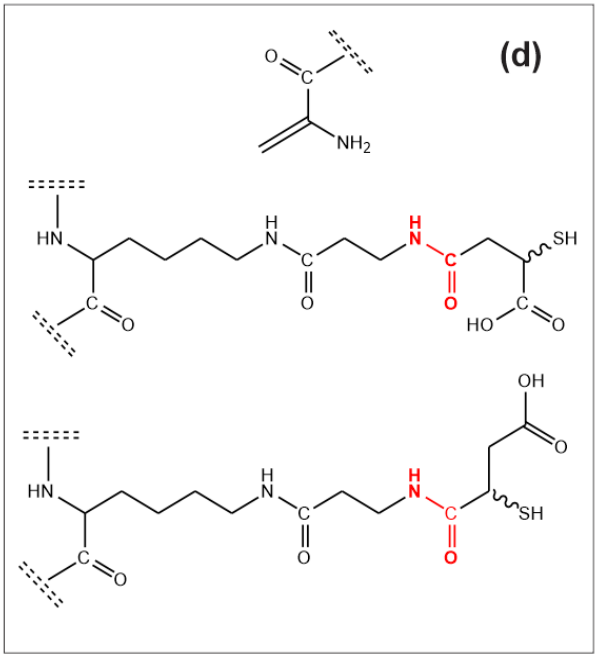


Fig. S4. The structure shown in **(a)** and **(b)** correspond to the isomers of the thiosuccinamic acids formed during the hydrolysis of the thiosuccinimide linker (**Fig. S1**). The ions shown in panel **(d)** and the ion named as C+98 in panel **(e)** are exclusive of the hydrolyzed thiosuccinimide linker. The ion named as P+71 in panel **(e)** is exclusive for the transcyclized and hydrolyzed thiosuccinimide linkers structures. The ion P shown in panel **(c)** is common for the linker studied in this manuscript: transcyclized and hydrolyzed thiosuccinimide (**Fig. S2**).







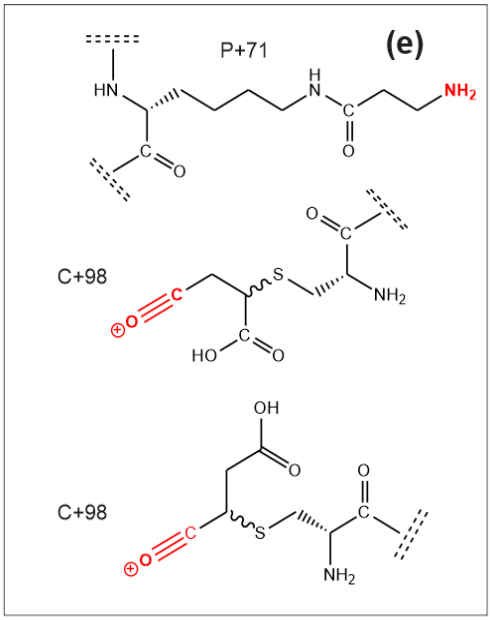


Fig. S5. Total ion current chromatogram obtained for LC-MS/MS analysis of KLH-Cys[1]pP0 conjugate digested with Lys-C **(A)**, trypsin **(B)** and v8 **(C)**. The digestion method used is the MED-FASP described by Wiśniewski and collaborators [10]. The dashed lines that appear in the Lys-C digestion (A) indicate an average of the intensity of the major peaks obtained in the trypsin (blue) and Glu-C (red) digestions.

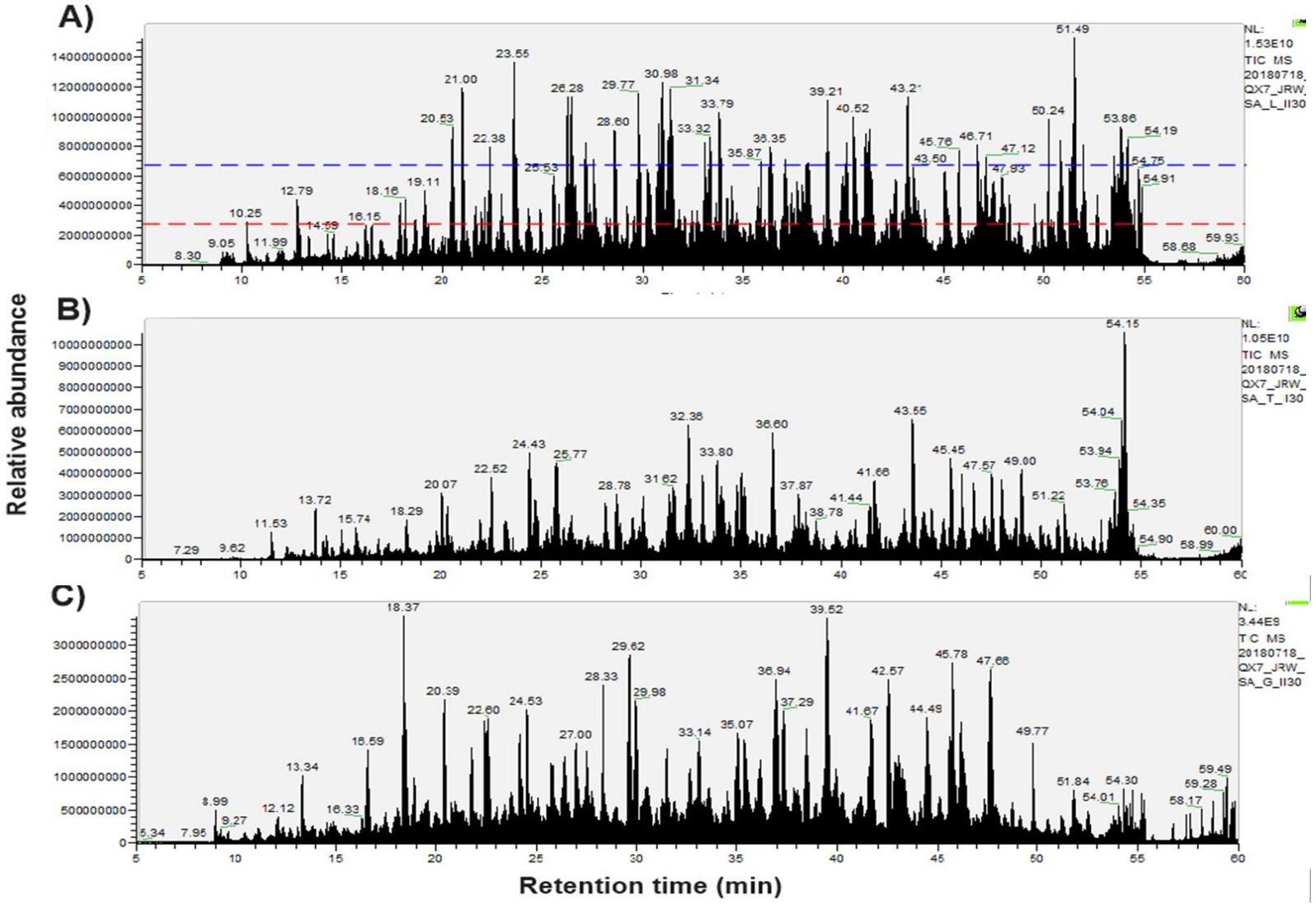


Fig. S6. Assignment of conjugation site located at K2633 in the sequence of KLH1 in the KLH1-Cys[1]pP0 conjugate. **(A)** MS/MS spectrum of type 2 peptide corresponding to KLH1 [D[2632]-R[2640]] linked to Cys[1]pP0 (C[1]-K[11]), detected at *m/z* = 710.997, 3+. Diagnostic ions that validate the assignment (P, P+71, C+L and C+80) are highlighted with arrows. The pink dot corresponds to the scan number # 222591, where the MS/MS mode was triggered to obtain the MS/MS spectrum shown in A. **(B)** Isotopic distribution of both peaks obtained in the extracted ion chromatogram of the type 2 peptide shown in **Fig.1A**.

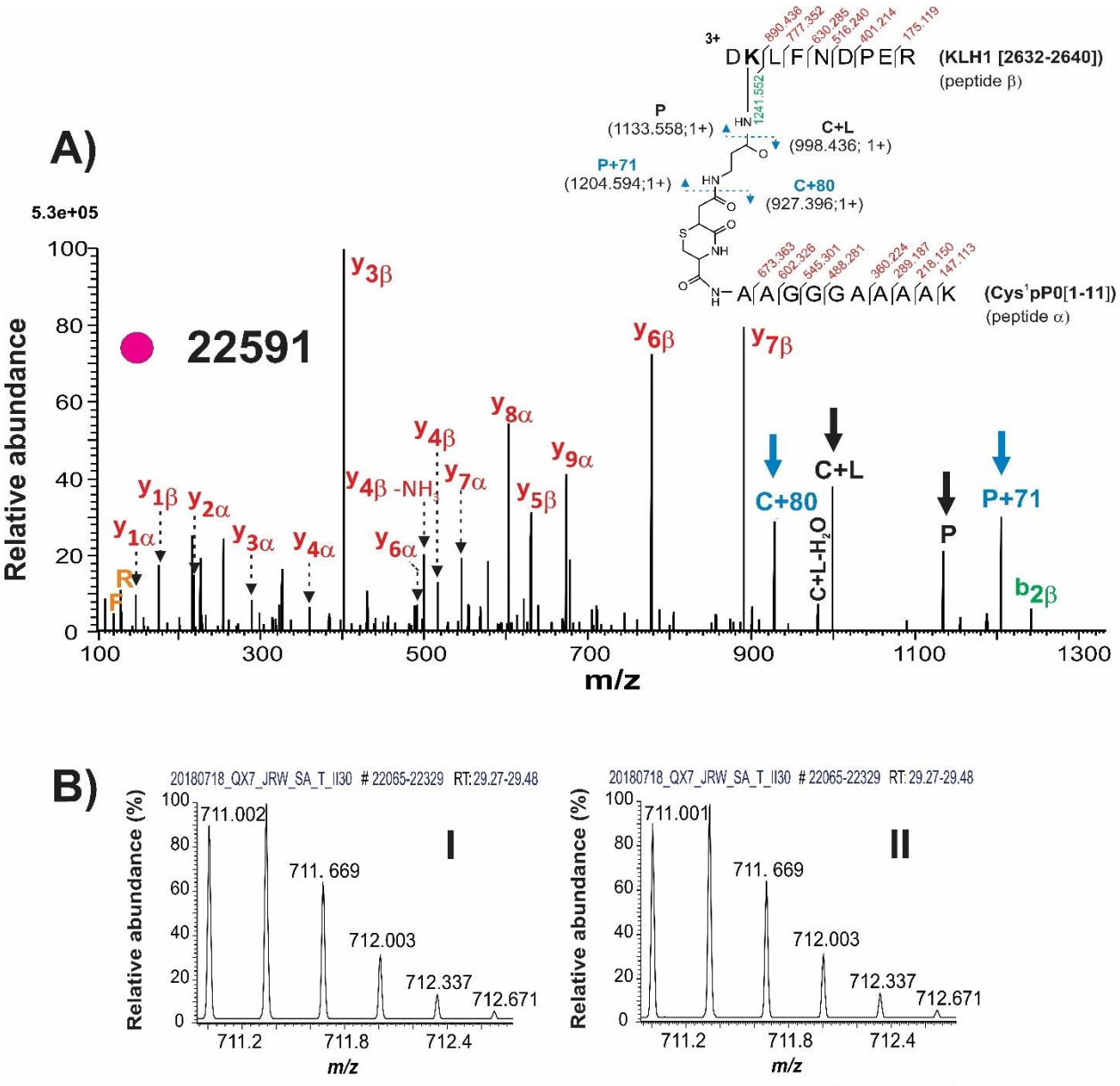


Fig. S7. Isotopic distribution of four-peak XIC pattern corresponding to KLH2 [D^{2097}-R^{2102}] linked to Cys^1pP0 (C^1-K^{16}), detected at m/z=893.101, 3+, with the hydrolyzed linker in KLH-Cys^1pP0 conjugates. XIC pattern obtained is shown in **Fig. 1C.**

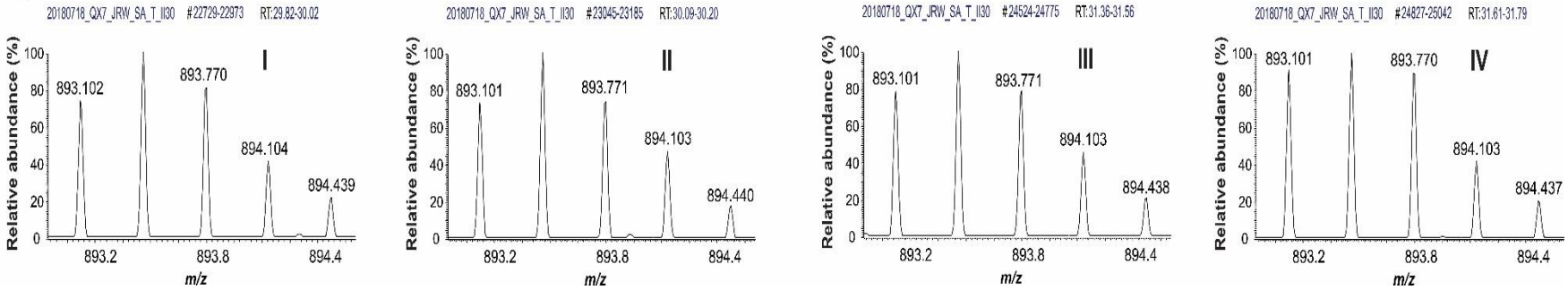


Fig. S8. *N*-terminus ragged peptides were identified in the linear peptides of KLH1 [2172-2183] **(A)** and KLH2 [1475-1489] **(B)** using the Peaks software [10]. MS/MS spectra showing the *N*-terminus ragged peptides KLH1 [2177-2183] **(C)** and KLH2 [1477-1489] **(D)** resulted from spontaneous ion source fragmentations at *N*-terminus of KLH1 [2172-2183] and KLH2 [1475-1489], respectively [11]. Letters a and c in (A) and (B) correspond to ammonia loss and carbamidomethylation, respectively.

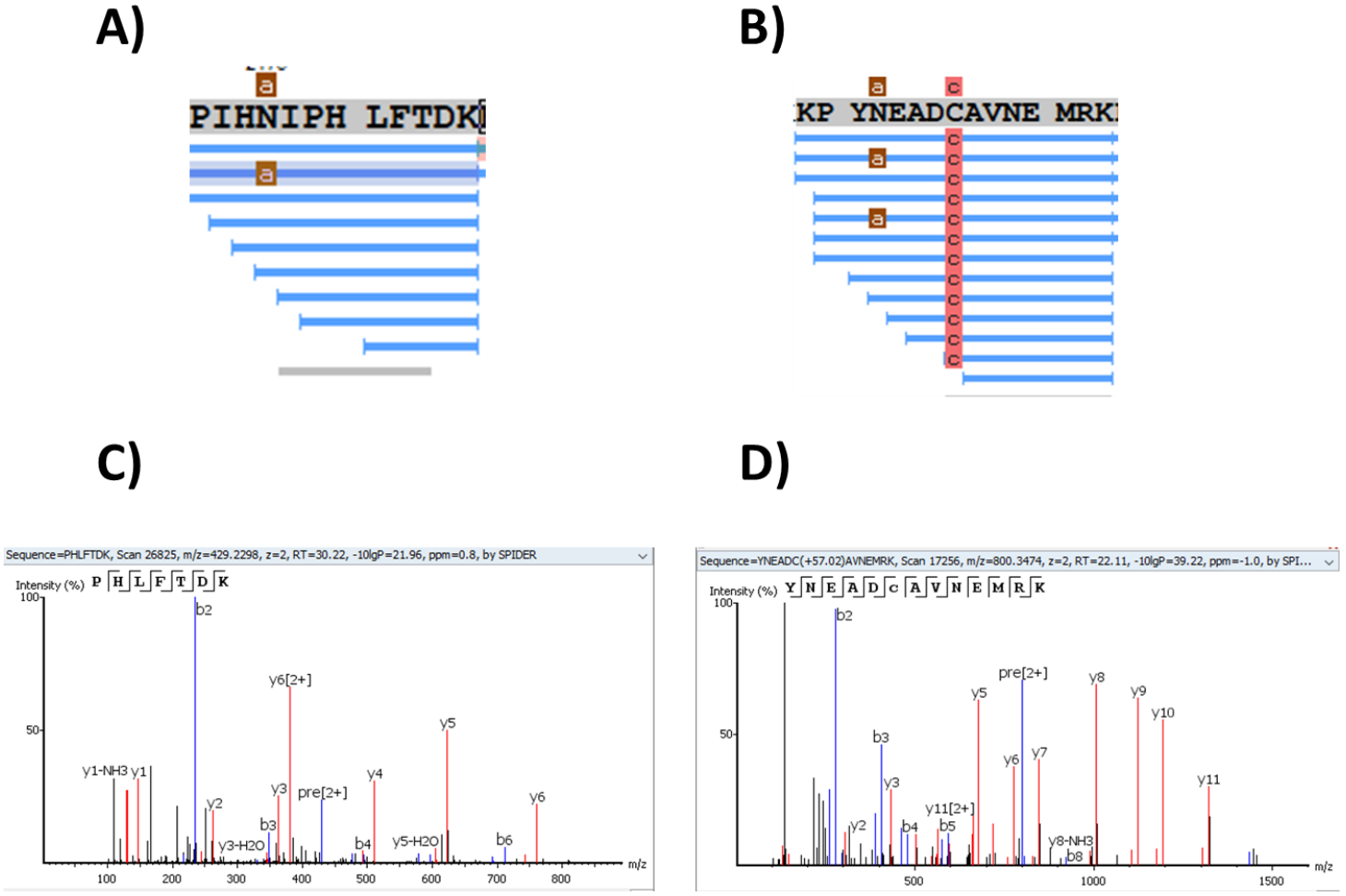


Fig. S9. Identification of amino acid changes in KLH1 [Y2331→H] **(A)** and KLH2 [G114→D] **(B)** by the Peaks software [10]. MS/MS spectra showing the amino acid change in the peptide KLH1 [2330-2340] **(C)** and KLH2 [110-119] **(D)**, respectively. The red letters (y"n series) and blue letters (b_n series) identify the amino acid change.

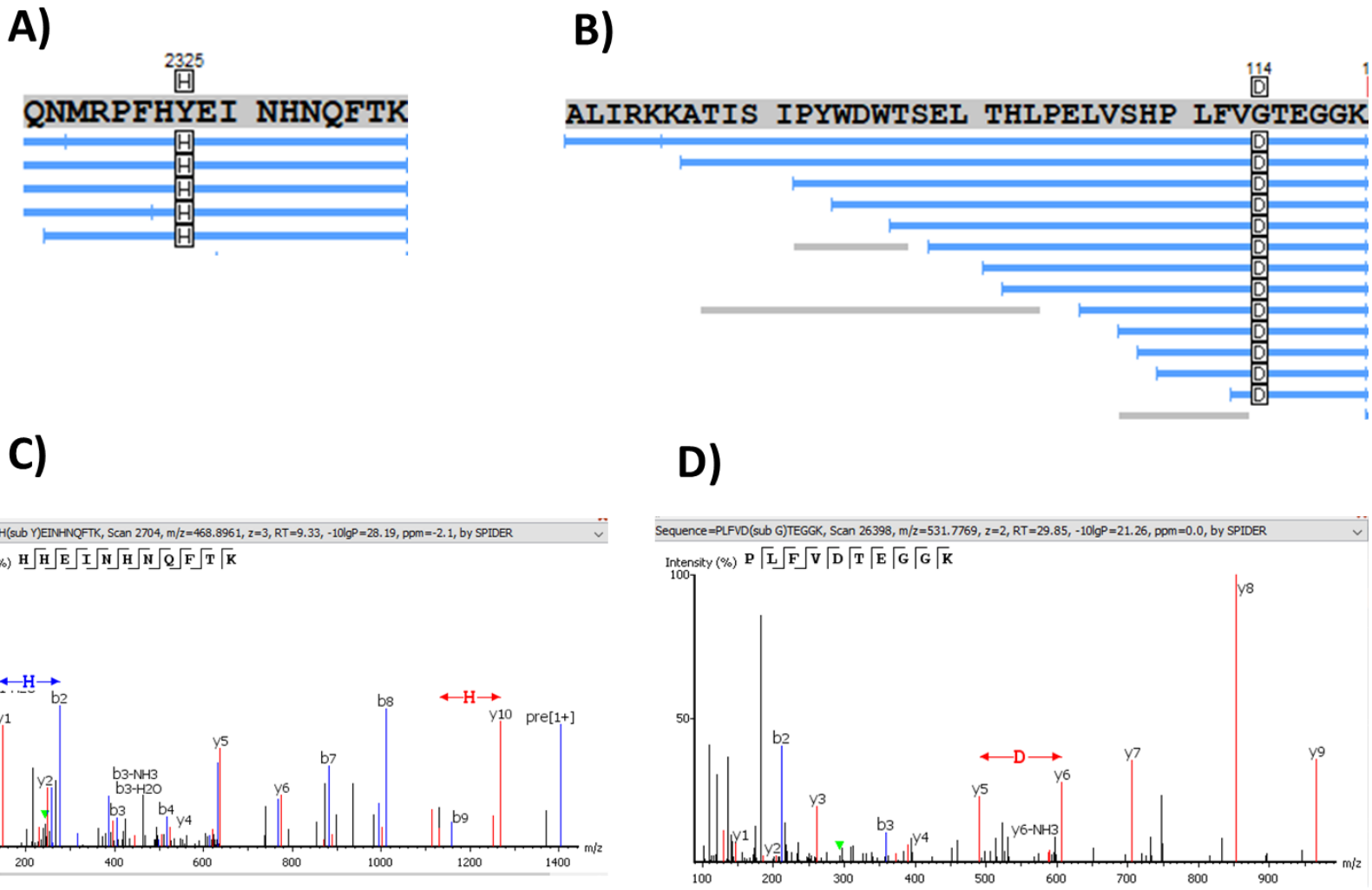


Fig. S10. Oxidized species of Trp. Isomers of the oxidized species of Trp: Oia (oxindolylalanine), Trp-OH (hydroxytryptophan) (+15.99 Da) and NFK (N-formylky”nurenine), DiOia (dioxindolylalanine) (+31.99 Da).

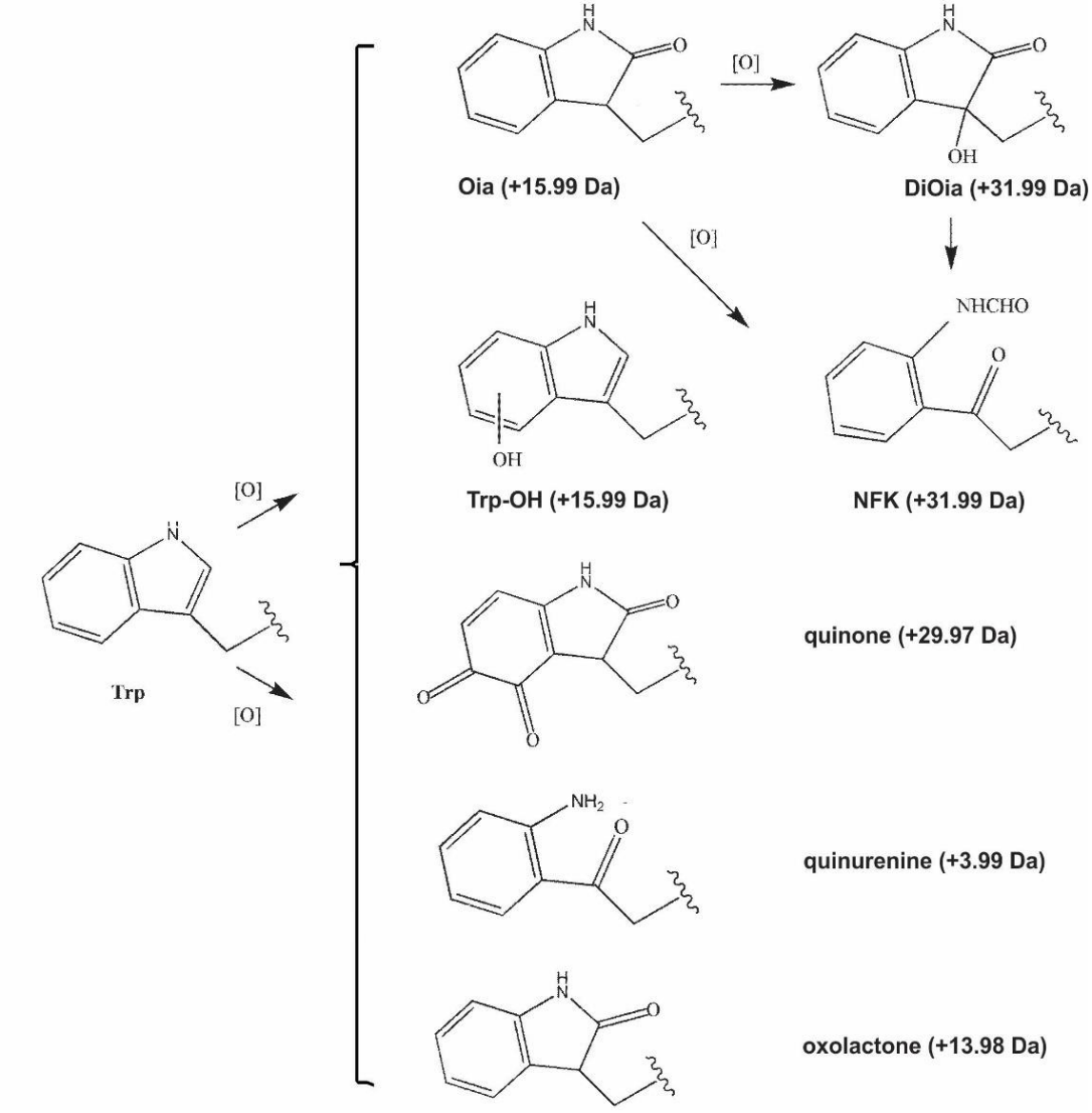


Fig. S11. Identification of the nitration of Tyr in KLH by the Peaks software [10] from intense y_n and b_n series belonging to the nitration of Y^{1731} in KLH1 [L^{1730}-R^{1740}] **(A)** and to the nitration of Y^{1324} in KLH2 [L^{1323}-R^{1335}] (B). The modified Tyr is represented by Y* in the sequence of the analyzed peptide. The letters in red (y_n series), green (b_n series) and orange (immonium ion) identify the signals in the MS/MS spectrum.

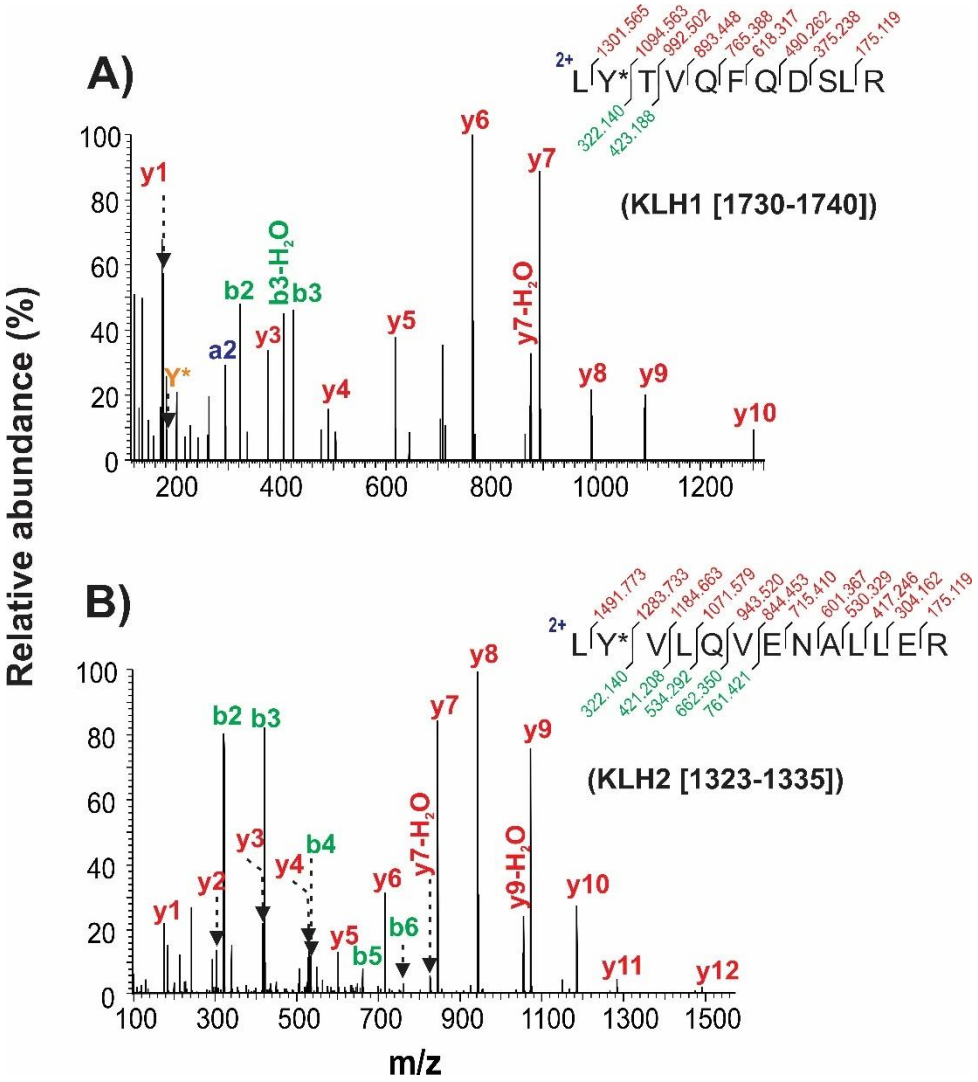


Fig. S12. MS/MS spectra that allowed the identification of phosphorylation in KLH2 by the Peaks software [10] of: **(A)** intense y"n series defining the increase of 79.97 Da of T^{1588} [A^{1584}-K^{1605}] and **(B)** the neutral loss of phosphoric acid (-97.98 Da) belonging to the phosphorylation of T^{2031} [Y^{2026}-K^{2033}]. The modified Thr is represented by the lower-case letter (t) in the sequence of the analyzed peptide.

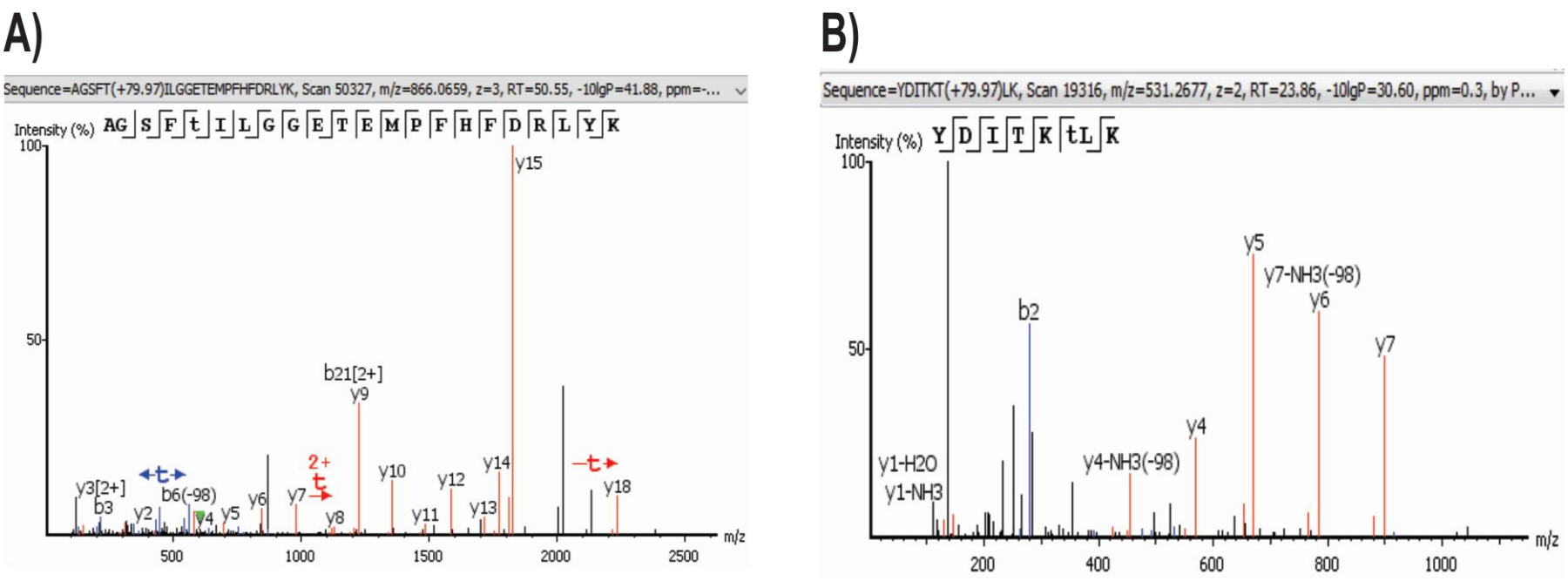


Fig. S13. MS/MS spectrum identifying the N-glycosylation of N^{389} in KLH2 [S^{383}-R^{399}] by the Peaks software [10] based on the presence of oxonium ions. The modified Asn is represented by the lower case letter (n) in the sequence of the analyzed peptide. Signals in green (m/z) correspond to oxonium ions.

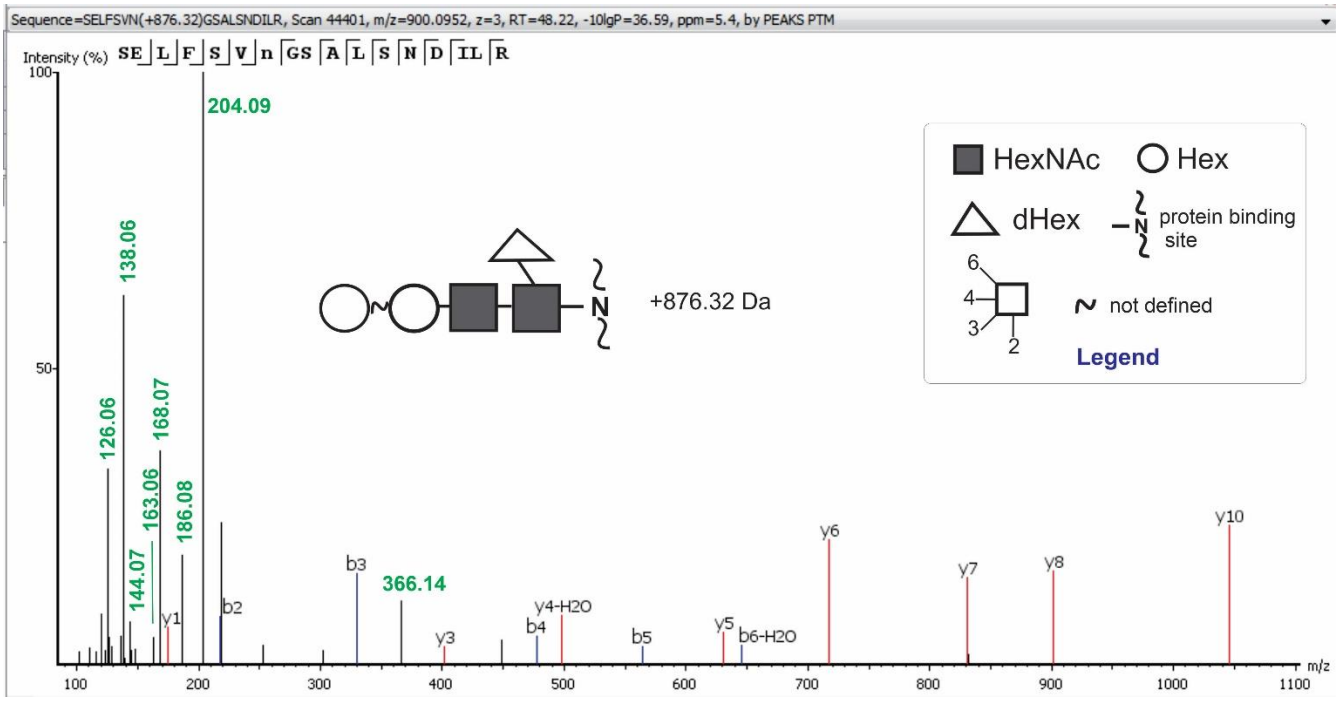


Fig. S14. Proposed structures for the *N*-glycans present in the glycopeptides identified in KLH1 and KLH2 by the Peaks software [10].

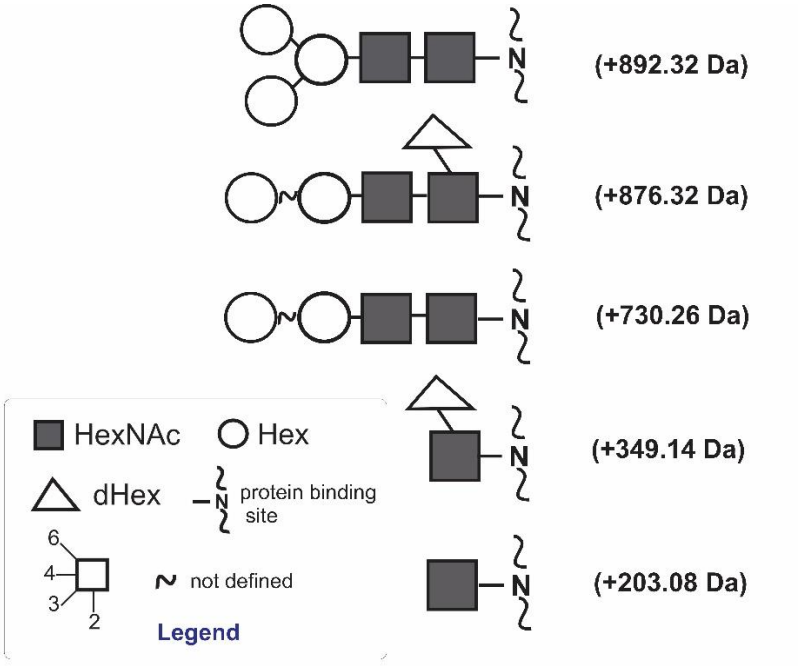


Fig. S15. MS/MS spectrum showing the identification by the Kojak software [10] of the thioether bond between residues C^{482}-H^{484} in peptide [Y^{477}-R^{494}] of KLH2. In the sequence of the analyzed peptide, the thioether bond between residues C^{482}-H^{484} connected by a line is identified; and C^{481} appears in its C[160] carbamidomethylated form. The letters in red (y_n series) and green (b_n series) identify the signals in the MS/MS spectrum.

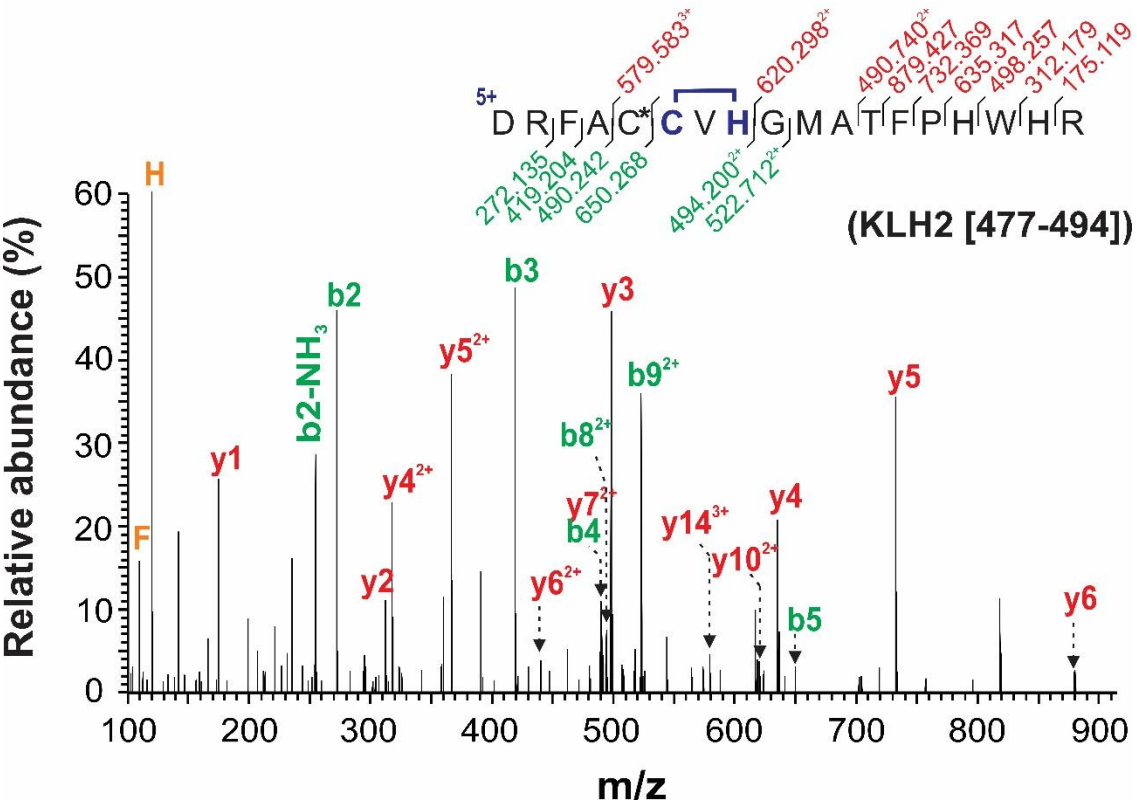


Fig. S16. Alignment of the sequences of the peptides containing the Cys-His thioether bond in the KLH1 and KLH2 sequences. In red letters, the conserved amino acid residues are highlighted and underlined, the Cys and His that form the thioether bond. This figure summarizes the Cys-His thioether bonds identified in the MS/MS spectra by the Kojak software [12]. FU (functional unit).

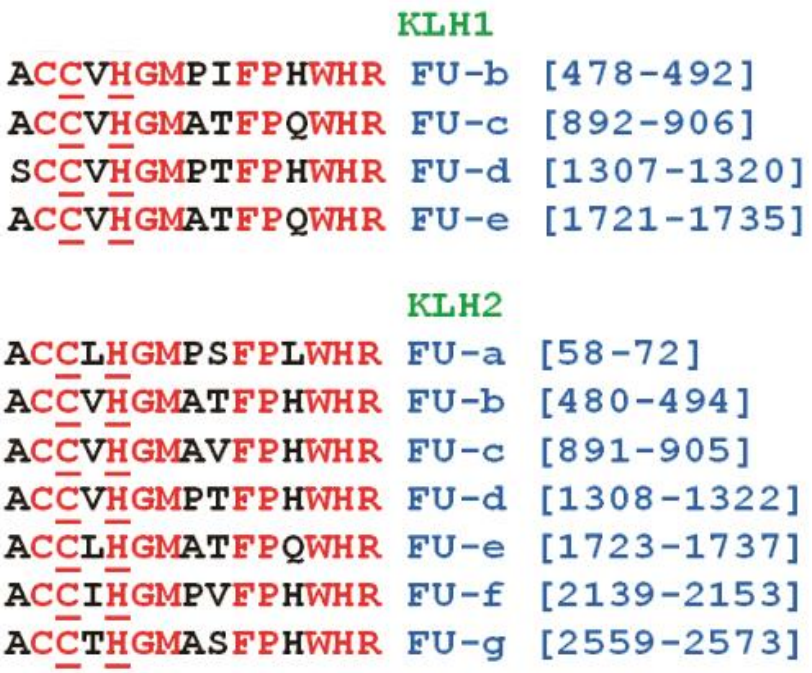
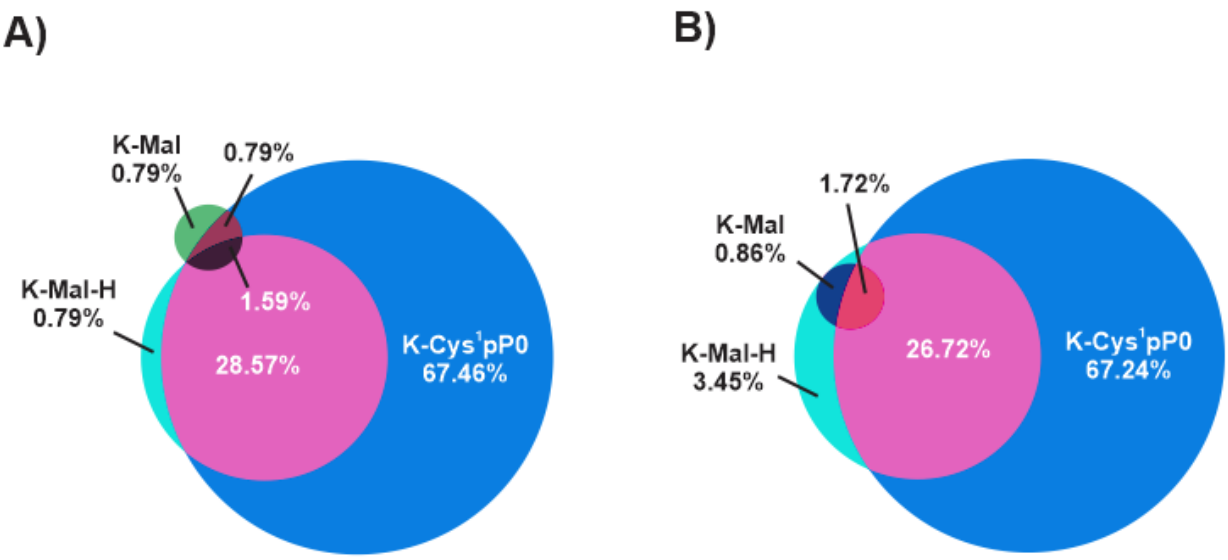


Fig. S17. Proportional Venn Diagram [13] overlaps the identification of Lys residues modified with Cys[1]pP0 (K-Cys[1]pP0), the maleimide group (K-Mal, ∆m= +151.027 Da) and the hydrolyzed maleimide group (K-Mal-H, ∆m= +169.038 Da), in peptides obtained by digestion of **(A)** KLH1-Cys[1]pP0 and **(B)** KLH2-Cys[1]pP0 conjugates with Lys-C, trypsin and v8 using the MED-FASP protocol [14] and analyzed by LC-MS/MS.



References

1. Northrop, B.H., S.H. Frayne, and U. Choudhary, *Thiol–maleimide "click" chemistry: evaluating the influence of solvent, initiator, and thiol on the reaction mechanism, kinetics, and selectivity.* Polymer Chemistry, 2015. **6**(18): p. 3415-3430, https://doi.org/10.1039/c5py00168d
2. Fields, G.B. and R.L. Noble, *Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids.* Int J Pept Protein Res, 1990. **35**(3): p. 161-214, https://doi.org/10.1111/j.1399-3011.1990.tb00939.x.
3. Rodríguez Mallón, A., et al., *Functional and mass spectrometric evaluation of an anti-tick antigen based on the P0 peptide conjugated to Bm86 protein.* Pathogens, 2020. **9**(6): p. 513, https://doi.org/10.3390/pathogens9060513
4. González, L.J., et al., *Synthesis, LC-MS/MS analysis, and biological evaluation of two vaccine candidates against ticks based on the antigenic P0 peptide from R. sanguineus linked to the p64K carrier protein from Neisseria meningitidis.* Anal Bioanal Chem, 2021: p. 1-16, https://doi.org/10.1007/s00216-021-03569-0
5. Lahnsteiner, M., et al., *Improving the stability of maleimide–thiol conjugation for drug targeting.* Chemistry–A European Journal, 2020. **26**(68): p. 15867-15870, https://doi.org/10.1002/chem.202003951
6. Gober, I.N., A.J. Riemen, and M. Villain, *Sequence sensitivity and pH dependence of maleimide conjugated N-terminal cysteine peptides to thiazine rearrangement.* Journal of Peptide Science, 2021. **27**(7): p. e3323, https://doi.org/10.1002/psc.3323
7. Pousa, S., et al., *On the utility of the extracted ion chromatograms for assigning the conjugation sites and side reactions in bioconjugates synthesized by the maleimide-thiol chemistry.* Microchemical Journal, 2024: p. 111025, https://doi.org/10.1016/j.microc.2024.111025
8. González, L.J., et al., *Differentiation of isobaric cross-linked peptides prepared via maleimide chemistry using MALDI-MS and MS/MS.* Rapid Communications in Mass Spectrometry, 2024. **38**(2): p. e9660, https://doi.org/10.1002/rcm.9660
9. Ramos-Bermúdez, P.E., et al., *A hydrolyzed N-propionylthiosuccinimide linker is cleaved by metastable fragmentation, increasing reliability of conjugation site identification in conjugate vaccines.* Rapid Communications in Mass Spectrometry, 2024. **38**(18): p. e9859, https://doi.org/10.1002/rcm.9859
10. Ma, B., et al., *PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry.* Rapid communications in mass spectrometry, 2003. **17**(20): p. 2337-2342, https://doi.org/10.1002/rcm.1196
11. Kim, J.S., et al., *In-source fragmentation and the sources of partially tryptic peptides in shotgun proteomics.* J Proteome Res, 2013. **12**(2): p. 910-6, https://doi.org/10.1021/pr300955f
12. Hoopmann, M.R., et al., *Kojak: efficient analysis of chemically cross-linked protein complexes.* Journal of proteome research, 2015. **14**(5): p. 2190-2198, https://doi.org/10.1021/pr501321h
13. Hulsen, T., J. de Vlieg, and W. Alkema., *BioVenn–a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams*. BMC genomics, 2008. 9: p. 1-6, https://doi.org/10.1186/1471-2164-9-488.
14. Wiśniewski, J.R., *Quantitative evaluation of filter aided sample preparation (FASP) and multienzyme digestion FASP protocols.* Analytical chemistry, 2016. **88**(10): p. 5438-5443,