Title: *N-propionyl thiosuccinimide linker upon hydrolysis becomes into a cleavable crosslinker by a gas phase metastable fragmentation and makes more reliable the conjugation sites identification.*

Short Title: A linker cleaved by a metastable fragmentation

Auhtors: Pablo Enmanuel Ramos[1,*], Satomy Pousa[2,*], Paulo Carvalho[4], Rodrigo Soares Caldeira Brant[4], Michel Batista[4], Hironobu Hojo[3], Alina Rodríguez Mallón[5], Vladimir Besada[2], Toshifumi Takao[6] and Luis Javier González[2,¥].

Affiliations:

[1]Bioinformatics Department and [2] Mass Spectrometry Laboratory. Proteomics Department, and [5]Animal Biotechnology Department. Center for Genetic Engineering and Biotechnology. Avenida 31 e/ 158 y 190. Cubanacán. Playa. Havana. Cuba.

[3]Laboratory for Protein Synthesis and [6]Functional Proteomics and Protein Profiling, Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan.

[4]Carlos Chagas Institute/FioCruz Rua Professor Algacyr Munhoz Mader, 3775 CIC 81350-010, Curitiba, Brasil.

(*) Authors with equal contributions

(¥) Corresponding author:

Dr. Luis Javier González PhD

Mass Spectrometry Laboratory, Department of Proteomics,

Center for Genetic Engineering and Biotechnology, Avenida 31 e/ 158 y 190,

Cubanacán, Playa, PO Box. 6162. La Habana, Cuba

Email: luis.javier@cigb.edu.cu

Abstract: (250 words)

Maleimide thiol chemistry is widely used for the synthesis of conjugate vaccines and antibody-drug-conjugates. Proteolysis of a conjugate vaccine mainly generates linear and type 2 peptides (cross-linked peptides) with a thiosuccinimide linker and its stabilized forms (thiazine, and/or hydrolyzed thiosuccinimide). Type 2 peptides contain valuable information on the conjugation sites. Cleavable crosslinkers could make more reliable the conjugation sites identifications, but based on a different chemistry, they are used to study protein-protein interactions and not in the synthesis of bioconjugates. Hydrolysis not only stabilizes the thiosuccinimide linker, but also turns it into a cleavable crosslinker via a gas phase metastable fragmentation observed in MALDI- and ESI-MS/MS analysis. The carboxyl group in the hydrolyzed thiosuccinimide linker provides a mobile proton that destabilizes the nearest pseudopeptide bond and, through the rearrangement of a five-membered ring intermediate, yields two intense linker fragment ions named here as P+71 and C+98. Specific ^{18}O-labeling of the hydrolyzed thiosuccinimide linker confirms that P+71 ion, has the epsilon amino group of Lys modified by a β-alanine residue while C+98 ion has the Cys residue *S*-alkylated by a succinic anhydride molecule. P+71 and C+98 provide information on the molecular masses of the crosslinked peptides pair, decrease the number of unassigned signals in the MS/MS spectra; and their backbone fragment ions increased the sequence coverage and make more reliable the assignment of conjugation sites. Stepped HCD is the fragmentation method of choice because in a single MS/MS spectrum several linker fragment ions and b_n/y_n ions were observed.

Keywords

MS-cleavable crosslinkers, thiosuccinimide, hydrolyzed linker, metastable fragmentation, conjugation sites, maleimide-thiol chemistry, conjugate vaccine.

Introduction:

The maleimide-thiol chemistry [1] is widely used to synthesize bioconjugates, including conjugate vaccines and antibody drug conjugates (ADC) [2]. The resultant thiosuccinimide linker in these bioconjugates, (**Figure 1**, structure III) is unstable *in vivo* due to the retro-Michael reaction [3]. Molecules with free thiols that are available in appreciable concentrations *in vivo*, react with the thiosuccinimide linker, releasing the antigen or the drug that is linked to the carrier protein. This process known as thiol exchange [4] could have a negative impact for the biological activity and cytotoxicity for the synthesized bioconjugate [5].

One strategy to overcome this limitation is the stabilization of the thiosuccinimide linker by its conversion into more stable forms that are resistant to the retro-Michael addition through two reactions known as transcyclization [6, 7] or hydrolysis [8] (**Figure 1**).

The transcyclization reaction transforms the thiosuccinimide linker into a thiazine linker [6, 7] (**Figure 1**, structure IV). This reaction also generates a new pseudopeptide bond in the thiazine linker structure (highlighted in red in **Figure 1**). For the transcyclization reaction the cysteine residue involved in the thiosuccinimide linker should be located at the *N*-terminal end of the peptide; and its amino terminal group should be free, and non-protonated [6].

The five-membered ring of the thiosuccinimide linker can be also hydrolyzed by two different pathways to yield two positional isomers of the thiosuccinamic acid thioether (**Figure 1**, structures I and II) [9]. Both positional isomers have in the linker, a newly formed pseudopeptide bond (highlighted in red in **Figure 1**) and a carboxylic group. This reaction takes place independently the position of the Cys residue within the

peptide sequence that is linked to the succinimide ring. The hydrolysis of the thiosuccinimide linker is favored at basic pH [8]. Both reactions, hydrolysis and transcyclization, proceed during sample processing prior to LC-MS/MS analysis [10] generally carried out at basic pH and 37ºC.

All the linker forms have a common pseudopeptide bond (highlighted in blue, **Figure 1**) and its fragmentation when detected in the MS/MS spectra provides information on the molecular mass of the crosslinked peptides.

The fragmentation in gas phase of the newly generated pseudopeptide bond present in the thiazine and the hydrolyzed thiosuccinimide linkers (**Figure 1**) generates an ion (named here as P+71) with identical molecular mass, and 71.0347 Da mass higher that the unmodified Lys containing peptide [7]. This linker fragment ion has been proposed for the differentiation by mass spectrometry of isobaric type 2 peptides cross-linked by a thiosuccinimide and thiazine linkers [7, 10].

The identification of the conjugation sites is a quality attribute of conjugate vaccines and ADC because it provides information on the conjugation efficiency; the linkage positions of the payload and it could modulates other properties such as the stability, toxicity, and pharmacokinetics [11, 12].

When a conjugate vaccine is digested with specific proteases a complex mixture of peptides is generated [13]. The type 2 peptides [14] are those that contain useful information on the conjugation sites [10, 15, 16]. They are composed by two cross-linked peptides, one derived from the carrier protein and the other from the antigen covalently linked by the thiosuccinimide linker or its stabilized forms (**Figure 1**). To identify the conjugation sites of conjugate vaccines by the LC-MS/MS analysis we have

used [10, 15, 16] the same software [17-21] that assign MS/MS spectra to type 2 peptides [22].

The identification of the conjugation sites and the studies of protein-protein interactions based on the chemical crosslinking and mass spectrometry share some challenges because both rely on the identification of type 2 peptides. According to a recent collaborative study [23] the automatic assignment of the MS/MS spectra to type 2 peptides although filtered by FDR criteria requires manual inspection even though this process has a subjective nature and it is a time-consuming task. This validation process is mainly based on the sequence coverage observed in the MS/MS spectra for both crosslinked peptides. When this validation process is supported by additional criteria, such as the extracted ion chromatograms [10], specific isotope labeling at MS level [24, 25]; and/or other pattern of signals observed in the MS/MS spectra, the output becomes more reliable.

The usage of cleavable cross-linkers for MS/MS analysis [26] also aims to this goal [27]. In general, these cross-linkers contain a very labile bond that is fragmented very efficiently in gas phase to yield intense signals of reporter ions detected in a silent region of the MS/MS spectra [28, 29]. Cleavable cross-linkers have also been designed with two labile bonds in their structures to yield a signal signatures in the MS/MS spectra that provides information on the molecular masses of the cross-linked peptides [26, 30]. With the same goal, other option has been the development of cross-linkers that introduces an spacer containing peptide bonds that can be easily cleavable in CID conditions such as the cases of Asp-Pro [31, 32] or Gly-Pro [33].

Although most of these cleavable cross-linkers are commercially available and have a tremendous analytical value in the context of studies of protein-protein interactions, they are not affordable for producing a veterinary conjugate vaccines, characterized by their low prices [34]. In addition, most of these cleavable cross-linkers have been designed to react with two primary amino groups separated by certain distance constrains, in the context of a 3D structure of multi-protein complexes, and they react randomly in a non-controlled manner and in single step. This uncontrolled reactions conditions could disrupt the linear and conformational epitopes in the antigen which could be deleterious for the biological activity of a conjugate vaccine. In addition, to our knowledge there are not previous toxicology records and clinical experiences using the above-mentioned MS cleavable cross-linkers in conjugate vaccines or ADC.

Rarely, a crosslinker reagent for human conjugate vaccines had been designed simultaneously to have a high efficiency in the chemical conjugation reaction and also to facilitate the analytical characterization of the conjugation sites.

In our group we have been working on the development of a broad spectrum veterinary vaccine against ticks [15, 16, 35] based on the chemical conjugation of a conserved antigenic P0 peptide (^{1}CAAGGGAAAAKPEESKKEEAK^{21}) [36] to three carrier proteins. For the synthesis of these vaccine candidates we have used different synthetic strategies [35, 37], including the maleimide-thiol chemistry [15, 16]. In the latter, we used BMPS (*N*-(β-maleimidopropyloxy)succinimide ester), a crosslinker reagent that have been used previously in the synthesis of effective human conjugate vaccines against *H. influenzae* type B [38], shiguellosis [39], and more recently against SARS-CoV-2 [40, 41].

BMPS allows the synthesis of conjugate vaccines by two well-controlled steps [15, 16]. In a first step, BMPS introduces several *N*-propionyl maleimide groups in the carrier protein, via the primary amino groups. In the second step, the maleimide groups react with the added Cys^{1}pP0 peptide containing a free thiol group at the *N*-terminal cysteine. By this procedure the structural integrity of the pP0 peptide is preserved being linked to the carrier protein by a thiosuccinimide linker.

In the LC-MS/MS analysis of conjugates where two stabilization pathways coexist (**Figure 1**) we noticed that the number of the conjugation sites based on the identification of type 2 peptides with the hydrolyzed thiosuccinimide linker and thiazine linkers were very similar [16]. However, the number of MS/MS assigned to type 2 peptides with the hydrolyzed linker increased considerably the number of supporting MS/MS spectra thus allowing a more reliable identification of the conjugation sites [16]. For these reasons in this work we studied the fragmentation of type 2 peptides with the hydrolyzed thiosuccinimide linker in gas phase for a more reliable characterization of the conjugate vaccines by LC-MS/MS analysis.

Experimental

BMPS crosslinker reagent (*N*-(β-maleimidopropyloxy)succinimide ester) was purchased from MERCK. Bm86 carrier protein (P20736) from *R. microplus* [42] was cloned and expressed in *P. pastoris* at the Center for Genetic Engineering and Biotechnology (CIGB) in Havana, Cuba [43]. Other solvent and reagents were analytical grade and were purchased from recognized suppliers.

Peptide synthesis and specific partial ^{18}O-labelling of the type 2 peptides with the hydrolyzed thiosuccinimide linkers

The Ac-Cys^{1}pP0 peptide (CH_3CO-^{1}CAAGGGAAAAKPEESKKEEAK^{21}) derived from *Rhipicephalus sp*. ticks (AJQ18906) acetylated at the *N*-terminal end was synthesized in the Peptide Synthesis Laboratory at CIGB [44]. Peptide was purified by preparative rp-HPLC and analytical fractions with a purity higher than 95 % were pooled and the identity was verified by ESI-MS using an experimental setup reported previously [45]. Other peptides P1 (AAAGGGAAAAK), P2(CAAGGGAAAAK), P3 (NCAGHK) and P4 (CH_3CO-GANAPKEPQR) as well as their derived type 2 peptides with the thiosuccinimide linker (P1-P2, P3-P4, P2-P4 and P1-P3) were synthesized as described previously [46]. Type 2 peptides with thiosuccinimide linker were purified by rp-HPLC, dried in the speed-vac, reconstituted in 100 mM Tris/HCl, pH=8.3 prepared alternatively with normal water or with 30 % (v/v)^{18}O-labeled water (98 % APE, Taiyo Nippon Sanso Corporation, Japan) and later incubated for 24 hours at 37^{o}C. The resultant type 2 peptides with thiazine and/or hydrolyzed thiosuccinimide linkers, were purified by analytical rp-HPLC and their identities were confirmed by MALDI-MS and MS/MS analyses [46] as shown in **Table S1**.

Synthesis of the Bm86-Ac-Cys[1]pP0 conjugate vaccine

The broad-spectrum conjugate vaccine against ticks, Bm86-Ac-Cys[1]pP0, was synthesized by the maleimide thiol chemistry using a procedure similar to the one described previously [15]. Very briefly, in a first step, the Bm86 carrier protein [43] was modified by adding an excess of BMPS crosslinker reagent to introduce several *N*-propionyl maleimide groups in lysine residues and at *N*-terminal end. In a second step, the activated carrier protein reacted with the Cys[1]pP0 peptide acetylated at the *N*-terminal end (Ac-Cys[1]pP0). The conjugate was purified by size exclusion chromatography to discard excess of reagents and dimeric Ac-Cys[1]pP0 generated as a byproduct of the synthesis.

Processing of Bm86-Ac-Cys[1]pP0 conjugate vaccine before LC-MS/MS analysis

The lyophilized conjugate vaccine Bm86-Ac-Cys[1]pP0 was dissolved in 1% (w/v), pH=8.3 ammonium bicarbonate solution and was heated at 37°C during 12 hours to transform the thiosuccinimide linker in the hydrolyzed form [9]. The resultant conjugate was reduced and *S*-alkylated by using a standard protocol [16] and later digested in tandem with Lysyl-endopeptidase (LEP, Wako, Japan) overnight and trypsin for additional six hours at the same temperature. The mixture of proteolytic peptides was acidified with 2 % of formic acid, desalted by using C18-Ziptip from Millipore following manufacturer instructions. The peptides eluted in an 60 % acetonitrile/1 % of formic acid solution were freeze and lyophilized, and kept at -20°C until LC-MS/MS analysis.

LC-MS/MS analysis

LC-MS/MS analysis was carried out at the mass spectrometry facility RPT02H/Carlos Chagas Institute–Fiocruz Paraná. The runs were performed on an Ultimate 3000 nLC

Thermo Scientific coupled with an Orbitrap Fusion Lumos Mass Spectrometer (Thermo Scientific). The peptide mixture was resuspended in formic acid 0.1% and quantified by absorbance in nanodrop (Thermo Scientific). Then, 0.5 µg was loaded on a 15 cm length, 75 µm I.D. emitter, packed in-house with ReproSil-Pur C18-AQ 3 µm resin (Dr Maisch GmbH HPLC) for peptide separation. The column was equilibrated with 250 nL/min of 95% of mobile phase A (0.1% formic acid in water) and the mobile phase B (0.1% formic acid, 5% water in acetonitrile) was increased linearly up to 40% during 60-minutes gradient. The instrument was set to data-dependent acquisition mode with time between master scans set to 2 s. Only the potential type 2 peptides derived from Bm86-Ac-Cys1pP0 included in a list (**Table S2**) were targeted for MS/MS analysis by using a normalized collision energy of 30 in collision induced dissociation (CID) and higher-energy collisional dissociation (HCD) fragmentation modes. Survey scans (300–2000 m/z) were acquired in the Orbitrap system with a resolution of 60,000 at m/z 200. The fragment ions were analyzed with a resolution of 15,000 at 200 *m/z*. The general mass spectrometric conditions were as follows: 2.30 kV spray voltage, capillary temperature of 175 ºC, predictive automatic gain control (AGC) set to 250% on MS1 and 220% for MS2.

Identification of type 2 peptides with the hydrolyzed thiosuccinimide linker in the Bm86-Ac-Cys^{1}pP0 conjugate:

The identification of type 2 peptides was carried out by pLink2 software [17] in the LC-MS/MS analysis of the tryptic peptides derived from Bm86-Ac-Cys^{1}pP0 conjugate. The elemental composition of the hydrolyzed thiosuccinimide linkers was $C_7H_7O_4N$. The definition for the crosslinked amino acids were the primary amino groups and a modified amino acid (named Z) with the elemental composition ($C_5H_7NO_2S$) corresponding to the

N-terminal acetylated Cys residue. A 5% FDR and other search parameters such as mass error tolerance, limits for molecular masses, fixed and variable modifications, and maximum and minimum length of type 2 peptides were the same as we described previously [16]. The set of MS/MS spectra assigned to type 2 peptides were manually validated by considering the sequence coverage of both crosslinked peptides in the MS/MS spectra and also the presence of an intense y"$_n$ fragment ion at the *N*-terminal side of Pro^{12} residue [47] revealing the size of the cysteinyl peptide (Ac-Cys^{1}pP0) [16] as well as the XIC patterns of all type 2 peptides identified [10]. Additionally, in the validation process, the fragmentation of the pseudopeptide (P+71, and C+98 ions) and thioether bonds (P+169, P+203, C-SH and C-34 ions) in the hydrolyzed thiosuccinimide linker were considered. In addition to this, in the manual validation other backbone fragment ions derived from the linearized crosslinked peptides containing the modified lysine and cysteine residues were also considered.

MALDI-MS and MS/MS analysis

MALDI-TOF MS was carried out with a 4800 MALDI-TOF/TOF mass spectrometer from Applied Biosystems (Framingham, MA, USA). All mass spectra were obtained by averaging 2500 laser shots from each sample well in the positive-ion mode. The instrument was controlled using 4000 series Explorer software (version 3.6; Applied Biosystems). Data were processed by using Data Explorer software (version 4.8; Applied Biosystems). The calibration of the instrument in MS and MS/MS modes were carried out by using the ions derived from mixture of several synthetic peptides as described previously [46]. Two MALDI-MS/MS spectra were acquired for all type 2 peptides with and without the CID gas (air). Peptides dissolved in water/acetonitrile/TFA

solutions (0.5 µL) were mixed with the matrix solution (0.5 µL), (the supernatant of a 50 % acetonitrile solution saturated with α–cyano-4–hydroxycinnamic acid), and then air dried on the flat surface of a stainless-steel plate. The ions were generated by irradiating the sample area with the output of a Nd/YAG laser at a wavelength of 355 nm, and generated ions were accelerated at 20 kV potential in the ion source.

Results and Discussions.

Metastable origin of the pseudopeptide bond fragmentation in type 2 peptides crosslinked by a hydrolyzed thiosuccinimide linker

The MALDI-MS/MS spectra of the P1-P2 type 2 peptide crosslinked by a thiazine ($(P1\text{-}P2)_T$) and hydrolyzed thiosuccinimide ($(P1\text{-}P2)_H$) linkers were acquired with and without collision gas (**Figure 2**).

The MALDI-MS/MS spectra of $(P1\text{-}P2)_T$ fragmented by collision induced dissociation (**Figure 2a**) allowed a full-sequence coverage of both crosslinked peptide sequences (**Fig. S1**). The fragmentation of the pseudopeptide bond newly generated by the transcyclization reaction [6, 7] (**Figure 1**) generates two ions at *m/z* 886.37 and *m/z* 927.30 that were tentatively assigned as (P1+71) and (P2+80), respectively (highlighted in red, see **Figures 2a** and **2e**). The intensities of these fragment ions (**Figure 2a**) were comparable to other backbone fragment ions (b_n or y_n) of P1 and P2 peptides. The expanded region (*m/z* 804-1034) of these two MALDI-MS/MS spectra (**Figure 2a** and **2b**) are shown in **Fig. S2**.

The linker fragment ions generated by the dissociation of the other pseudopeptide bond between the amino terminal group of P1 and the thiazine linker were detected at *m/z* 815.31 and *m/z* 998.34 and they were assigned to P1 and P2+151 (see empty and filled rhombuses in **Figure 2a,** and **2e**), respectively. These linker fragment ions P1 and P2+151), have intensities lower than the ones (P1+71 and P2+80) generated by the fragmentation of the other pseudopeptide bond in the thiazine linker, and newly generated after the transcyclization reaction [6, 7]. P1+71 is approximately 2.3 times more intense than P1, while P2+80 and P2+151 have similar intensities (**Figure 2a**).

In the MALDI-MS/MS spectrum of $(P1\text{-}P2)_T$ acquired without collision gas (**Figure 2b**), the sequence coverage decreased as expected (**Fig. S1a**) showing almost no fragment ions in the low-mass region.

In this condition, where only metastable fragment ions are detected (**Figure 2b**), it is evident that all backbone fragment ions as well (P1+71) and (P2+80), have a similar contribution of a metastable origin. In particular, (P1+71) linker fragment ion has an intensity 0.61 times lower than the signal (*m/z* 1684.54) assigned to the *C*-terminal rearrangement of peptides ($b_{10\alpha}+H_2O$ and $b_{10\beta}+H_2O$) generated mainly by a metastable fragmentation as reported previously by Gaskell *et al.* [48].

The MALDI-MS/MS spectra of $(P1\text{-}P2)_H$ acquired with (**Figure 2c**) and without (**Figure 2d**) collision gas yield linker fragment ions P1+71 (*m/z* 886.46) and P2+98 (*m/z* 945.40) confirmed the sequences of P1 and P2 peptides crosslinked by a hydrolyzed thiosuccinimide linker (**Fig. S1b**). In particular (P1+71) and (P2+98) were the two-most intense fragment ions in both MALDI-MS/MS spectra (**Figures 2c** and **2d**). In fact, (P1+71) linker fragment ion was 9.6 times more intense than the *C*-terminal rearrangement ions ($b_{10\alpha}+H_2O$ and $b_{10\beta}+H_2O$) [48].

Linker fragments derived from the other pseudopeptide bond (*m/z* 815.31 and *m/z* 1016.35) were detected as a very low-intense fragment ions (**Figures 2c** and **2d**), suggesting that their formation is less favorable than the pseudopeptide bond newly generated after linker hydrolysis (highlighted in red, see **Figures 2c** and **2d**).

The P1+71 and P2+98 fragment ions in the MALDI-MS/MS acquired without collision gas (**Figure 2d**) were considerably more intense than other linker fragment ions detected at *m/z* 984.46 and *m/z* 1018.44 and tentatively assigned to P1+169 and

P1+203. These linker fragment ions are originated by the fragmentation of the thioether bond in the hydrolyzed thiosuccinimide linker [46] (**Figure 2f** and **2g**) and their tentative structures are shown in **Fig. S3a**.

The results obtained for three other type 2 peptides with the hydrolyzed thiosuccinimide linker ($(P2\text{-}P4)_H$, $(P1\text{-}P3)_H$, and $(P3\text{-}P4)_H$) and analyzed by MALDI-MS/MS with and without collision gas (**Fig. S4** – **S8**) confirmed the results obtained for $(P1\text{-}P2)_H$ (**Figures 2c** and **2d**). In all of them P_n+71 were the most intense signal in the MALDI-MS/MS spectra. The relative intensity of P_n+71 and P_n+98 in all MALDI-MS/MS spectra seems to be modulated by the relative basicity of the *C*-terminal amino acids present in the two crosslinked peptides.

Although the linker structures for $(P1\text{-}P2)_T$ and $(P1\text{-}P2)_H$ are different (**Figures 2e** and **2f**), the structure of the resultant P1+71 fragment ions should be identical, however, the result shown in **Figure 2** led us to conclude that a metastable fragmentation generated more favorably P+71 in $(P1\text{-}P2)_H$ than the case of $(P1\text{-}P2)_T$.

Considering that P1-P2 with thiosuccinimide linker does not have this pseudopeptide bond in its linker structure we could say that hydrolysis turns the thiosuccinimide linker into a cleavable crosslinker in MALDI-MS/MS analysis.

Mechanism that could explain the metastable linker fragmentation in type 2 peptides with the hydrolyzed thiosuccinimide linker

To explain origin of the metastable fragmentation that generates intense P1+71 and P2+98 fragment ions (**Figure 2c**, **2d**) in $(P1\text{-}P2)_H$ (**Figures 2c** and **2d**) we hypothesized that the proton located at the linker carboxyl group become mobile and migrates spontaneously to the nearest carbonyl group in the newly formed pseudopeptide bond

by hydrolysis (highlighted in red, see **Figure 3**). This process destabilizes favorably this pseudopeptide bonds in a similar way that protonation does the same for Asp-Pro linker peptide bonds in linear peptides to generate backbone fragment ions (**Figure 3**) [49, 50].

Next, an spontaneous internal rearrangement of a cyclic five-membered ring intermediate leads to the dissociation of this destabilized linker pseudopeptide bond (**Figure 3**) to yield the two intense signals assigned to (P1+71) and (P2+98) in $(P1\text{-}P2)_H$ (**Figures 2c** and **2d**). Both linearized peptides P1 and P2 contain modified amino acids that reveal information on the conjugation site. The P+71 fragment ion contains a β-alanine amino acid linked by a pseudopeptide bond to the epsilon amino group of the crosslinked lysine residue (**Figure 3**). In the case of P2+98 fragment ion, the formerly crosslinked cysteine residue, is alkylated by a succinic anhydride structure (**Figure 3**). This rearrangement proposed for the hydrolyzed thiosuccinimide linker shares similarities with the described by Soberblom *et al*. for an Asp-Pro cleavable crosslinker [31, 32].

Although the mobile proton provided by the linker carboxyl group could also be transferred to the other pseudo peptide bond in the linker (between the amino terminal group of P1) or any peptide bond $(P1\text{-}P2)_H$, its fragmentation seems to be not equally favored. Probably, the linker carboxyl group involved in the internal rearrangement and other amide bonds are distant in the structure and could not yield a five membered ring intermediate required for this metastable fragmentation (**Figure 3**).

Thiazine linker does not have any source of mobile protons in their structure that promote its own destabilization therefore the metastable fragmentation for $(P1\text{-}P2)_T$ is

not favored. The formation P1+71 is probably also mediated by another mobile proton available in $(P1\text{-}P2)_T$. In this sense, this pseudopeptide bond and other peptide bonds in $(P1\text{-}P2)_T$ seem to have a similar affinity for protonation taking into account their similar intensities observed in the MALDI-MS spectra (**Figures 2a** and **2b**).

Complementary linker fragment ions (P1+203 and P2-34) and (P1+169 and P2-SH) generated by the thioether bond fragmentation also have a metastable component for their origin (see filled and empty circles and stars, **Figures 2d, 2f, 2g)**. However, their formation seems to be not so favored like the cases P1+71 and P2+98 in $(P1\text{-}P2)_H$ (**Figures 2c** and **2d)** and probably other mechanisms involving beta elimination and proton transfer mediated their formation [46, 51] (**Fig. S9** and **Fig. S10**).

MALDI-MS/MS analysis of type 2 peptides partially labeled with ^{18}O in the hydrolyzed thiosuccinimide linker

Figures 4a and **4b** show the isotopic ion distribution of $(P1\text{-}P3)_H$ obtained in a buffer prepared with normal and 30 % of ^{18}O-labeled water, respectively. The MALDI-MS/MS spectrum (**Figure 4c**) of the ^{18}O-labeled $(P1\text{-}P3)_H$ shows that P1+71 (*m/z* 886.38) and P3+98 (*m/z* 727.29) are the two-most intense fragment ions. The isotopic ion distributions of the backbone fragment ions were pooled into two groups: (1) signals showing their natural isotopic ion distribution (with no linker in their structures), and (2) showing the incorporation of ^{18}O in the linker-containing fragment ions (**Fig. S11**).

All the ^{18}O-labelling is retained in P3+98 linker fragment ion, specifically in the succinic anhydride group (**Figure 3**) that is alkylating the Cys (**Figure 4d**) while P1+71 shows its natural isotopic ion distribution (**Figure 4e**). This difference agrees with the fragmentation mechanism depicted in **Figure 3** and the structures proposed for both

linker fragment ions. The reduced P3 peptide (P3-SH, *m/z* 629.20) does not contain the linker in their structure in consequence showed its natural isotopic distribution (**Figure 4d**). Other linker fragment ions generated by the thioether bond fragmentation detected at *m/z* 984.34 and *m/z* 1018.38 and assigned as P3+169 and P3+203, respectively, showed isotopic ion distributions that evidenced the incorporation of ^{18}O like the ion detected at *m/z* 940.32 assigned as $b_{2\alpha}$ (**Figure 4f**). The linker fragment ions P3+169 and P3+203 contain the hydrolyzed thiosuccinimide linker (**Fig. S9** and **S10)**.

These results evidenced that the hydrolysis of the thiosuccinimide linker turns it into a cleavable crosslinker in MALDI-MS/MS analysis. Like other cleavable crosslinkers with similar structural characteristics [31, 32], this attribute could be conveniently used to perform a more reliable identification of the conjugation sites.

The MALDI-MS of the three other type 2 peptides synthesized with the thiosuccinimide linker in a buffer prepared with 30 % of ^{18}O-labeled water (**Fig. S12-S14**) showed similar characteristics (**Figure 4**).

Detection of the pseudopeptide bond fragmentation in LC-MS/MS experiments: CID vs HCD

Most of the conjugate vaccines and ADC, owing to their complexity are analyzed by LC-MS/MS and not by MALDI-MS. Considering that the multiply-charged ions in ESI-MS have a greater availability of mobile protons, we wonder which of these regularities described above for the type 2 peptides with a hydrolyzed thiosuccinimide linker can be extended to LC-MS/MS analysis.

We synthesized Bm86-Ac-Cys^{1}pP0 conjugate as a model to study the fragmentation type 2 peptides with the hydrolyzed thiosuccinimide linker. In this conjugate the

hydrolysis is the only stabilization pathway allowed for the thiosuccinimide linker (**Figure 1**) because the *N*-terminal end of the cysteinyl peptide (Ac-Cys^{1}pP0) is blocked with an acetyl group. This conjugate was incubated at basic pH to promote the linker hydrolysis and after reduction and *S*-alkylation it was digested with trypsin and analyzed by LC-MS/MS. The precursor ions of thirty-two 32 different type 2 peptides (see **Table S2**) with the hydrolyzed thiosuccinimide linker were fragmented by CID and HCD in two independent scans using the same normalized collision energy of 30.

The MS/MS spectra of the of the precursor ion detected at *m/z*=1222.56, 3+ were assigned by pLink2 software to a type 2 peptide [^{70}A-K^{87}]-(^{1}Z-K^{16}) with the hydrolyzed thiosuccinimide linker when fragmented by CID (**Figure 5a**) and by HCD (**Figure 5b**) were used. ^{1}Z means an *N*-terminal acetylated Cysteine residue in Cys^{1}pP0 peptide.

In the MS/MS spectrum shown in **Figure 5a** the two-most intense fragment ions detected at *m/z*=1085.01, 2+ and *m/z*=1557.69, 1+ were not assigned automatically by pLink2. In addition, the poor sequence coverage of the peptide [^{70}A-K^{87}] make this assignment questionable. Knowing the lability of the hydrolyzed thiosuccinimide linker in gas phase (**Figure 3**) the signals detected at *m/z*=1085.01, 2+ and *m/z*=1557.69, 1+ were manually assigned to the linker fragment ions ([^{70}A-K^{87}]+71)$^{2+}$ and ((^{1}Z-K^{16})+98)$^{1+}$, respectively. Also the third-most intense signal detected at *m/z*=779.34, 2+ was assigned to ((^{1}Z-K^{16})+98)$^{2+}$. All these linker fragment ions were also more intense than the signal detected at *m/z*=589.28, 1+ and assigned to the βy_5, which is originated by the fragmentation of a very labile peptide bond at the *N*-terminal of proline [47]. This result confirms that hydrolyzed thiosuccinimide linker is also cleavable in ESI-MS analysis when CID is used.

The fragmentation of the other pseudopeptide bond present in the linker between the epsilon amino group of Lys and the β-Ala residue in the hydrolyzed linker (highlighted in blue, see **Figure 1**) that would yield the $([^{70}A\text{-}K^{87}]+nH)^{n+}$ and $((^{1}Z\text{-}K^{16})+169+nH)^{n+}$ fragment ions were not observed. It confirms that despite they are also a pseudopeptide bond itd fragmentation by CID is not favored like the cases of $([^{70}A\text{-}K^{87}]+71)^{2+}$ and $((^{1}Z\text{-}K^{16})+98)^{1+}$.

The MS/MS spectrum of this triply-charged peptide (*m/z*= 1222.56) fragmented by HCD (**Figure 5b**) was also assigned automatically by pLink2 to $[^{70}A\text{-}K^{87}]\text{-}(^{1}Z\text{-}K^{16})$. Only four backbone fragment ions (b_3, b_4, b_6 and y_2) derived from the eighteen amino acids peptide $[^{70}A\text{-}K^{87}]$ were automatically assigned and several signals remained unassigned.

This assignment could be also questionable despite of a more efficient fragmentation method such as HCD was used [52, 53]. In this MS/MS spectrum, the most intense fragment ion (*m/z*=1557.69, 1+) was also assigned manually to $(^{1}Z\text{-}K^{16})+98)^{1+}$. The intense signal detected at *m/z*=1085.01, 2+ and assigned to $([^{70}A\text{-}K^{87}]+71)^{2+}$ in **Figure 5a** decreased considerably the intensity in **Figure 5b**, suggesting that it was further fragmented by HCD. When this possibility was considered the sequence coverage of $([^{70}A\text{-}K^{87}]+71)^{2+}$ was improved by matching new eleven $*y_n$ ions (from $*y_{14}$ to $*y_4$) containing a β-alanine modified lysine (conjugation site at Lys^{85}) with the molecular mass increased by 71Da (**Figure 5b**). Also four $*b_n$ fragment ions (from $*b_6$ to $*b_9$) of the peptide $(^{1}Z\text{-}K^{16})+98)^{1+}$ were assigned manually after considering the backbone fragment ions derived from this linearized peptide containing a cysteine residue alkylated by a succinic anhydride molecule (**Figure 3**).

The linker fragment ion ([^{70}A-K^{87}]+71)$^{2+}$ was additionally fragmented more efficiently than (^{1}Z-K^{16})+98)$^{1+}$ probably because it is a doubly-charged ions with more availability of mobile protons. Considering the favorable cleavability of the hydrolyzed thiosuccinimide linker, when multiply-charged precursor ions are fragmented the number of protons attached to the resultant P+71 and C+98 linker fragment ions could not be evenly distributed and in consequence their relative fragmentation efficiency could be different. Although the additional fragmentation of P+71 and C+98 linker fragment ions could affect their detection in the MS/MS spectra, their derived fragment ions contain very valuable information on the conjugation site, decrease the number of non-assigned signals in the MS/MS spectra and increase the sequence coverage, an important aspect in the data validation. This phenomenon was more noticeable in HCD (**Figure 5b**) than CID (**Figure 5a**), because HCD is more efficient CID fragmentation method.

The assignment of the fragment ions derived from the linearized peptides should be included in the scoring methods and algorithms of software dedicated to the identification of type 2 peptides improve the confidence of the results and to rescue some dubious assignments such as the depicted in **Figure 5**.

Influence of the number of mobile protons on the intensities of the P+71 and C+98 linker fragment ions of peptides fragmented by sHCD

Figure 6 shows the MS/MS spectra same type 2 peptide ([^{192}T-K^{200}]-(Ac-^{1}C-K^{16})) with different precursor charge states (from 2+ to 4+) fragmented by sHCD at a normalized collision energies of 27, 30 and 33.

In the MS/MS spectrum of the 2+ precursor ion (**Figure 6a**) the two-most intense signals detected at *m/z* 1110.55, 1+ and *m/z* 1557.69, 1+ were assigned to linker

fragment ions ([^{192}T-K^{200}]+71)$^{+}$ and ((^{1}Z-K^{16})+98)$^{+}$, respectively. Both signals were more intense that the $y_{5}\alpha$ (*m/z* 589.28, 1+), a fragmentation of a very labile peptide bond at the *N*-terminal end of proline [47].

The scarce source of mobile protons in ([^{192}T-K^{200}]-(^{1}Z-K^{16}))$^{2+}$ precursor ion makes the newly generated carboxyl group as the most prominent source of a mobile proton, therefore it destabilizes the pseudopeptide bond (**Figure 3**) and in consequence linker fragment ions ([^{192}T-K^{200}]+71)$^{+}$ and ((^{1}Z-K^{16})+98)$^{+}$ were detected as the two-most intense in the MS/MS spectrum (**Figure 6a**). The sequence coverage of (^{1}Z-K^{16}) was very modest (40 %). When the MS/MS is dominated by intense signals of the crosslinked peptide pair due to the linker fragmentation, the sequence coverage deduced from the backbone fragment ions is compromised. In this MS/MS spectrum other linker fragment ions ([^{192}T-K^{200}]+169)$^{+}$, ([^{192}T-K^{200}]+203)$^{+}$, (^{1}Z-K^{16})$^{+}$ and ((^{1}Z-K^{16}) - 34)$^{+}$, generated by the thioether bond fragmentation [46], as well as the fragmentation of the other linker pseudopeptide bond (highlighted in blue, see **Figure 1**) were not observed.

In the MS/MS spectra of 3+ (**Figure 6b**) and 4+ (**Figure 6c**) precursor ions, the intensity of ([^{192}T-K^{200}]+71)$^{+}$ and ((^{1}Z-K^{16})+98)$^{+}$ decreased, although their corresponding doubly-charged ions were also detected and with intensities comparable to other backbone fragment ions (see also **Table S3** and **S4**).

As expected, this result confirms that a higher availability of mobile protons activates multiple fragmentation pathways [49] and it is not favorable for the detection of the linker fragment ions [^{192}T-K^{200}]+71)$^{+}$ and ((^{1}Z-K^{16})+98)$^{+}$ which require only the exclusive fragmentation of the pseudopeptide bond in the hydrolyzed thiosuccinimide linker.

In addition, in **Figure 6b** and **6c** the $([^{192}T\text{-}K^{200}]+203)^{+}$ linker fragment ion generated by the thioether bond fragmentation (m/z = 1242.55, 1+) was detected and with an intensity comparable to the linker fragment ions $[^{192}T\text{-}K^{200}]+71)^{2+}$ and $((^{1}Z\text{-}K^{16})+98)^{2+}$. The MS/MS spectra of the same precursor ions fragmented by HCD at a collision energy of 30 showed similar characteristics (**Fig. S15, Table S5-S7**).

All these linker fragment ions have an analytical value by confirming the presence of the hydrolyzed thiosuccinimide linker and providing information on the molecular masses of the crosslinked peptides.

Conclusions:

Hydrolysis turns the resultant thiosuccinimide linker into a cleavable crosslinker in MALDI-MS/MS and ESI-MS/MS analysis. Type 2 peptides with the hydrolyzed thiosuccinimide linker undergoes a linker fragmentation of a metastable origin in gas phase. This fragmentation is initiated by a proton transfer of the newly generated linker carboxyl group to the nearest pseudopeptide bond. The destabilized amide bond in the linker undergoes an internal rearrangement through a five-membered ring that generates two linker fragment ions named here as P+71 and C+98. P+71 ion contains a β-Ala residue linked to the epsilon amino group of the conjugated lysine residue. In the structure of C+98 ion, the cysteine residue is *S*-alkylated by a succinic anhydride molecule.

These linker fragment ions provide information on the molecular mass of the two crosslinked peptides and their structures have two modified amino acids (lysine and cysteine) residues involved in the crosslinking reaction.

The assignment of backbone fragment ions derived from P+71 and C+98 ions increases the sequence coverage; decreases the number of unassigned signals in the MS/MS spectra and also provides valuable information on the position in the sequence of the conjugation site.

sHCD could be the most appropriated fragmentation method to analyze the type 2 peptides with a hydrolyzed thiosuccinimide linker because in the same MS/MS spectra, we observed a high sequence coverage of the crosslinked peptide pair as well as linker fragment ions generated by a metastable fragmentation and others requiring collision-activated decomposition for their detection.

The fragmentations described here for the hydrolyzed thiosuccinimide linker as well as their derived backbone fragment ions must be incorporated in the scoring methods and algorithms of software developed to assign the MS/MS spectra of type 2 peptides to obtain a more reliable assignment of the conjugation sites of conjugate vaccines.

The regularities observed here could be also applicable to bioconjugates synthesized by using self-hydrolyzing maleimides that have developed for an efficient stabilization of the bioconjugates through linker hydrolysis.

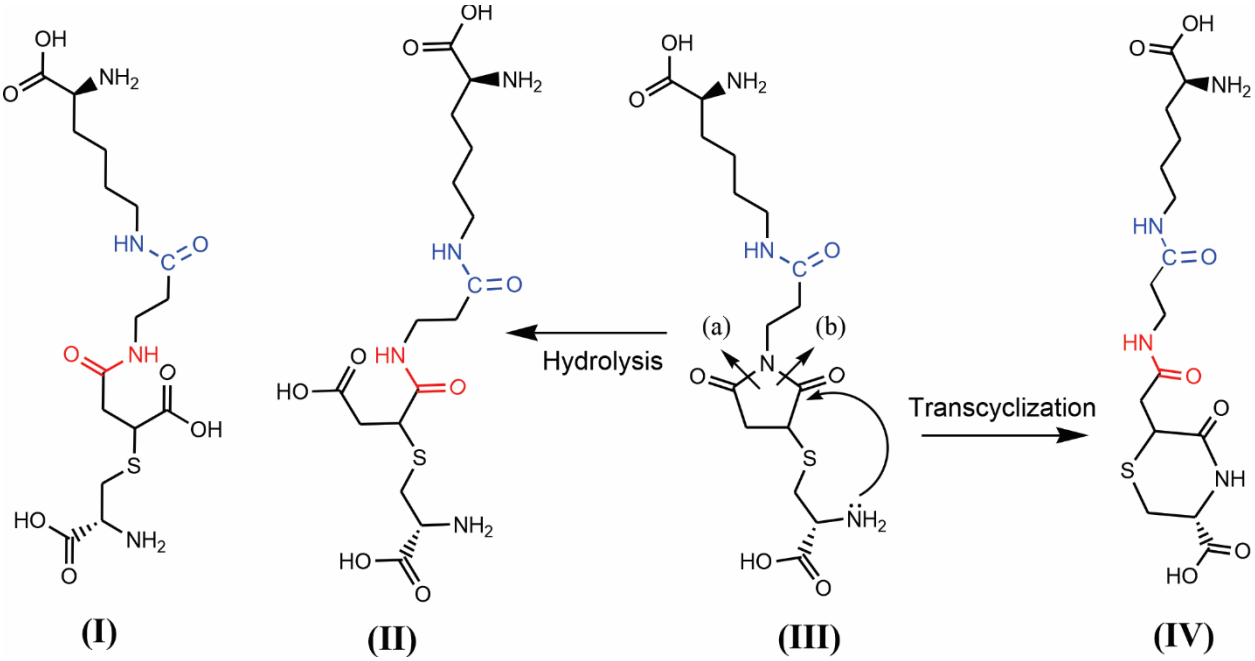


Figure 1. Structures of thiosuccinimide (III), thiazine (IV) and hydrolyzed thiosuccinimide linkers (I and II). Highlighted in red, the newly formed pseudopeptide bonds, generated by transcyclization and hydrolysis reactions, are indicated. The thiosuccinimide ring has two possible pathways for hydrolysis denoted here as (a) and (b) to yield two positional isomers of the hydrolyzed thiosuccinimide linker (structures I and II). Highlighted in blue the other pseudopeptide bond common for all the linker structures shown in this figure.

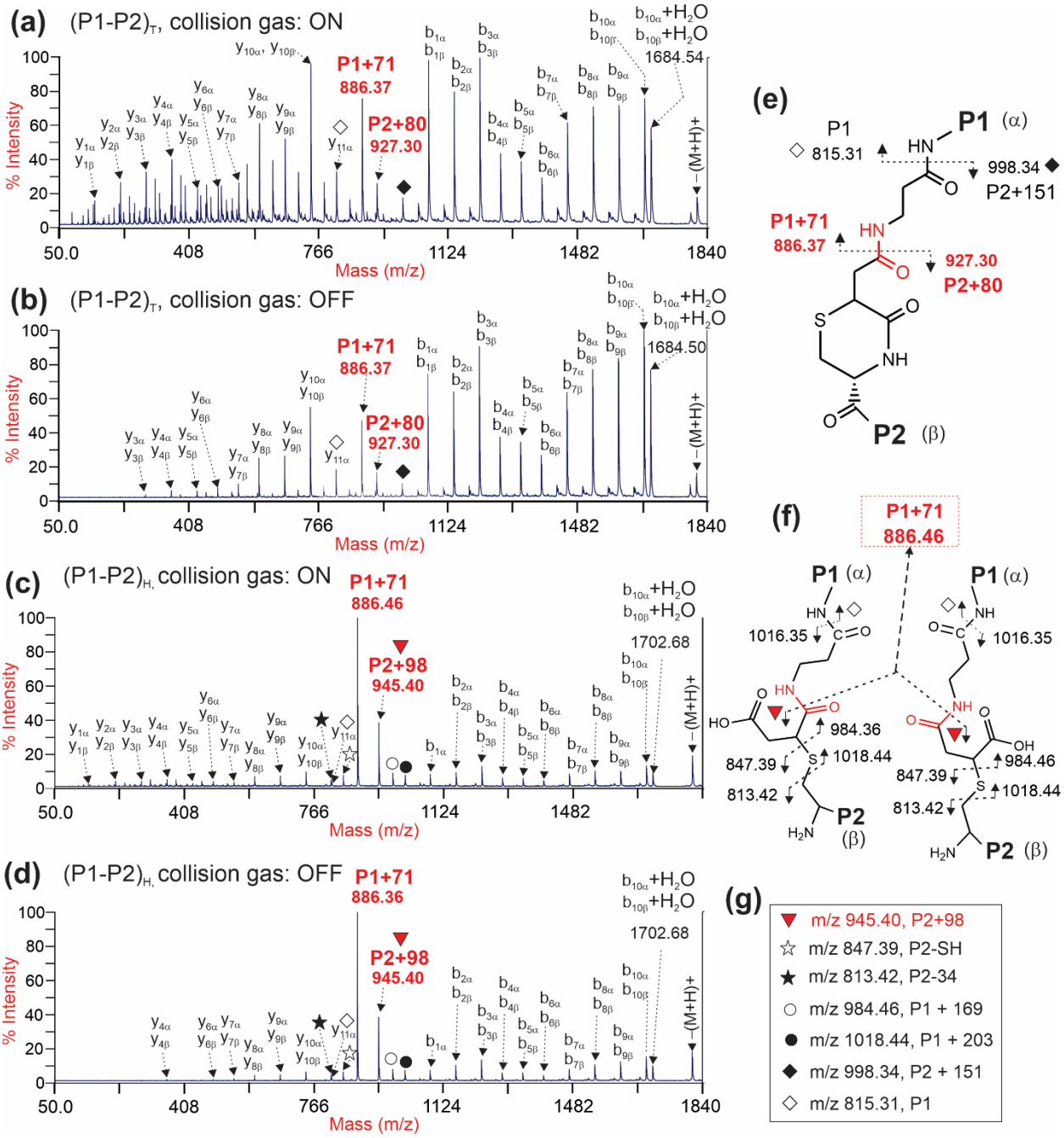


Figure 2. Metastable and collision-activated fragmentation origin of P1+71, P2+80 and P2+98 fragment ions of a type 2 peptide crosslinked by a thiazine ($(P1\text{-}P2)_T$) and hydrolyzed thiosuccinimide ($(P1\text{-}P2)_H$) linkers in MALDI-MS/MS experiments. MALDI-MS/MS spectra in **(a)** and **(b)** correspond to the $(P1\text{-}P2)_T$ acquired with and without collision gas, respectively. MALDI-MS/MS spectra in **(c)** and **(d)** correspond to the (P1-

P2)$_H$ acquired with and without collision gas, respectively. The structures shown in **(e)** and **(f)** correspond to (P1-P2)$_T$ and (P1-P2)$_H$, respectively. The newly pseudopeptide bond generated by the transcyclization and hydrolysis are highlighted in red in **(e)** and **(f)**. The signals labeled with a red triangle in **(c)**, **(d)** and **(f)** correspond to the P2+98 linker fragment ion. The nomenclature of the fragment ions agrees with the reported by Schilling *et al* [14]. The *m/z* of all fragment ions of the hydrolyzed thiosuccinimide linker is shown in the legend **(g)**.

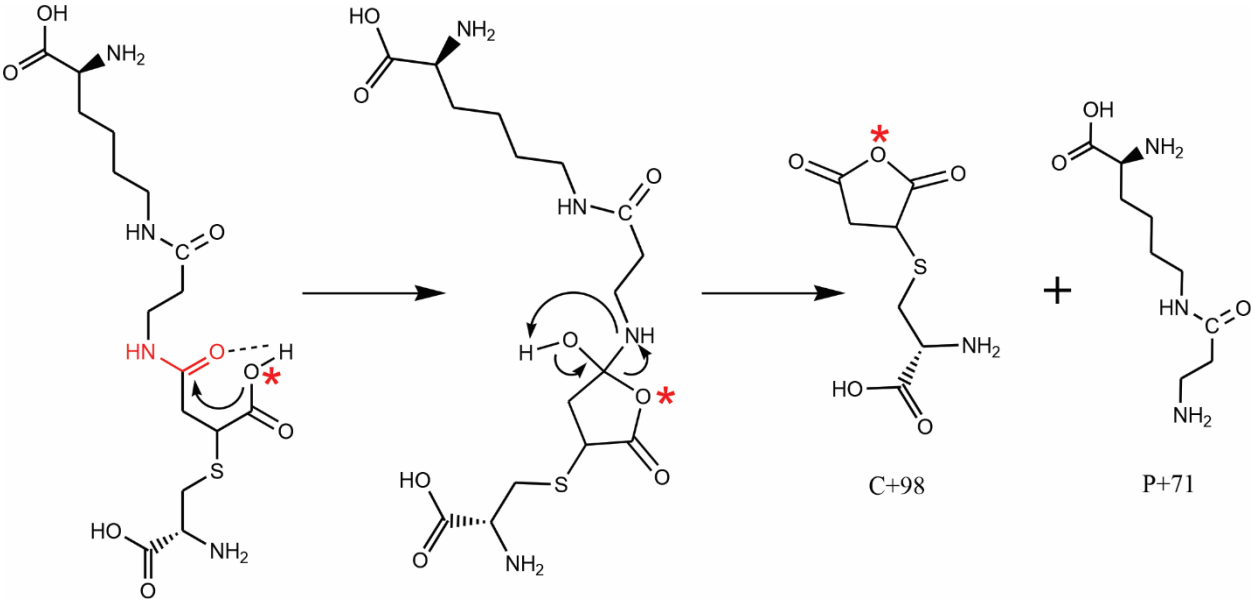


Figure 3. Proposed mechanism for the metastable fragmentation of the hydrolyzed thiosuccinimide linker that yields two fragment ions C+98 and P+71 corresponding to the cysteine- and lysine-containing peptides, respectively. The structure on the left represents the lysine and cysteine residues linked by a hydrolyzed thiosuccinimide linker. The red asterisk indicates the incorporation of ^{18}O upon hydrolysis of the original thiosuccinimide linker in a buffer prepared with 30 % (APE) of $H_2{}^{18}O$ water. The

mechanism shown in this figure is common for the two positional isomers generated by the hydrolysis of the thiosuccinimide linker (structures I and II, **Figure 1**).

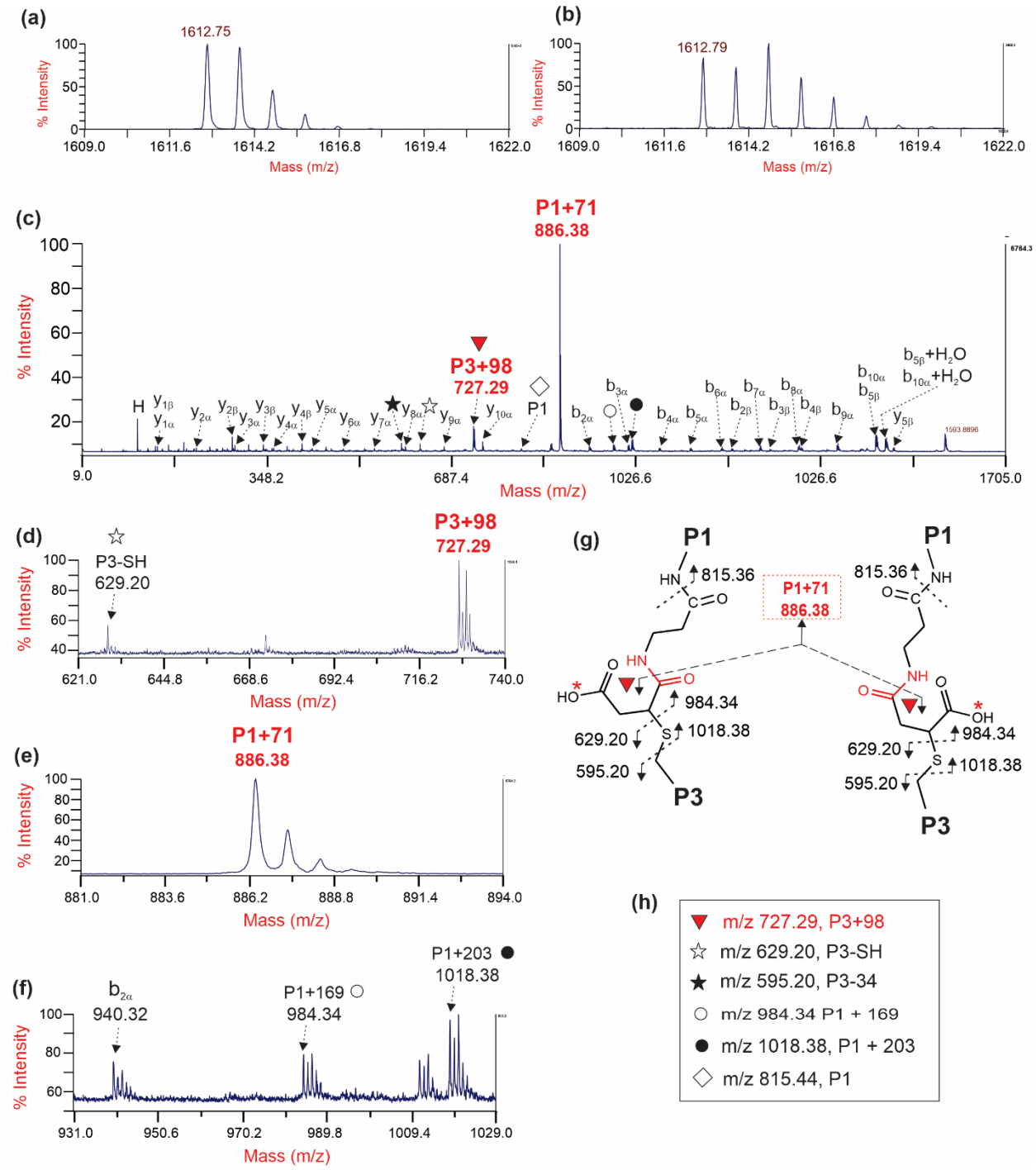


Figure 4. Isotopic ion distributions of $(P1\text{-}P3)_H$ generated by the hydrolysis of the original P1-P3 with the thiosuccinimide linker in a 100 mM Tris/HCl pH=8 buffer prepared with normal water **(a)** and ^{18}O-labeled water (30 % APE) **(b)**. **(c)** MALDI-MS/MS of the $(P1\text{-}P3)_H$ partially labeled with ^{18}O at the newly carboxylic group generated upon hydrolysis of the succinimide ring. The expanded regions (*m/z* 621-740), (*m/z* 881-894) and (m/z 931-1029) of the MALDI-MS/MS spectrum in **(c)** are shown in **(d)**, **(e)** and **(f)**, respectively. **(g)** An schematic representation of the structures of the two positional isomers of $(P1\text{-}P3)_H$ shows the linker fragment ions of the hydrolyzed thiosuccinimide linker. The *m/z* of all fragment ions of the hydrolyzed thiosuccinimide linker is shown in the legend **(h)**. The nomenclature of the fragment ions agrees with the reported by Schilling *et al* [14].

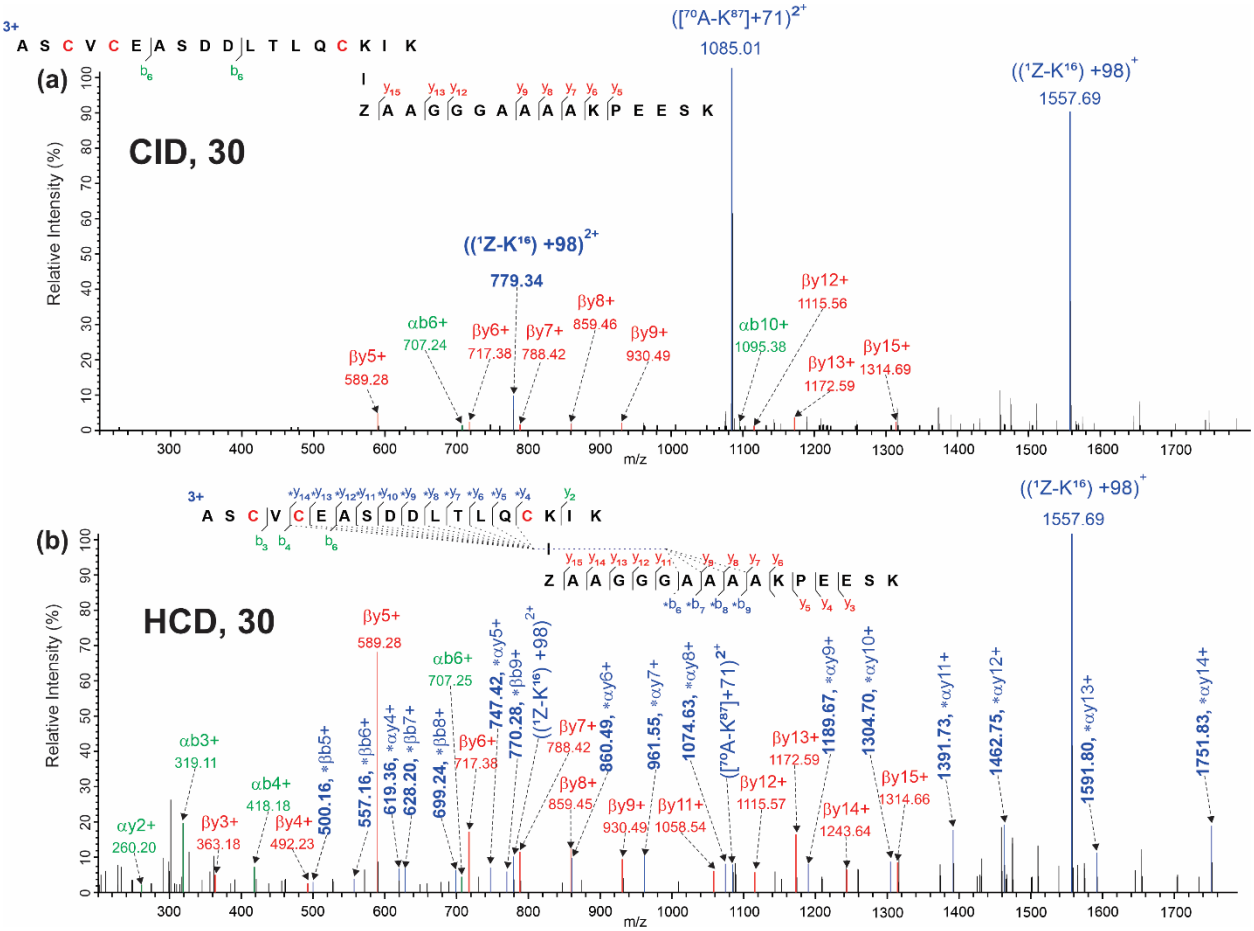


Figure 5. The MS/MS spectra assigned by the pLink2 software to a tryptic type 2 peptide [^{70}ASCVCEASDDLTLQC**K**IK^{87}]-(^{1}**Z**AAGGGAAAAKPEESK^{16}) (*m/z*=1242.5640, 3+) of the Bm86-Ac-Cys^{1}pP0 conjugate with the hydrolyzed thiosuccinimide linker. The MS/MS spectra shown in **(a)** and **(b)** were acquired when this peptide was fragmented by CID and HCD, respectively by using a normalized collision energy of 30. In the amino acid sequences of the type 2 peptides, Cys residues highlighted in red are carbamidomethylated. *Z* indicates an *N*-terminal acetylated of the cysteinyl peptide Ac-Cys^{1}pP0 peptide crosslinked to lysine 85 in the sequence of the Bm86 carrier protein. The signals detected at *m/z* 1085.01 and *m/z* 1557.69 were assigned to the linker fragment ions ([^{70}A-K^{81}]+71)$^{2+}$ and (Ac-^{1}C-K^{16}) + 98)$^{+}$, respectively. Backbone fragments

highlighted in green and red correspond to the alpha [^{70}ASCVCEASDDLTLQC**K**IK^{87}] and beta (^{1}**Z**AAGGGAAAAKPEESK^{16}) peptides, respectively. Fragments highlighted in blue as *(α/β)y_n and *(α/β)b_n) were manually assigned and they are derived from either the alpha of beta peptides containing a lysine and cysteine with their molecular masses increased by 71 and 98 Da, respectively. Broken lines in **(b)** indicate the fragment ions that were generated by the fragmentations of the pseudopeptide bond in the linker and a peptide backbone in one of the crosslinked peptides. The nomenclature for the other backbone fragment ions agrees with the reported by Schilling *et al* [14].

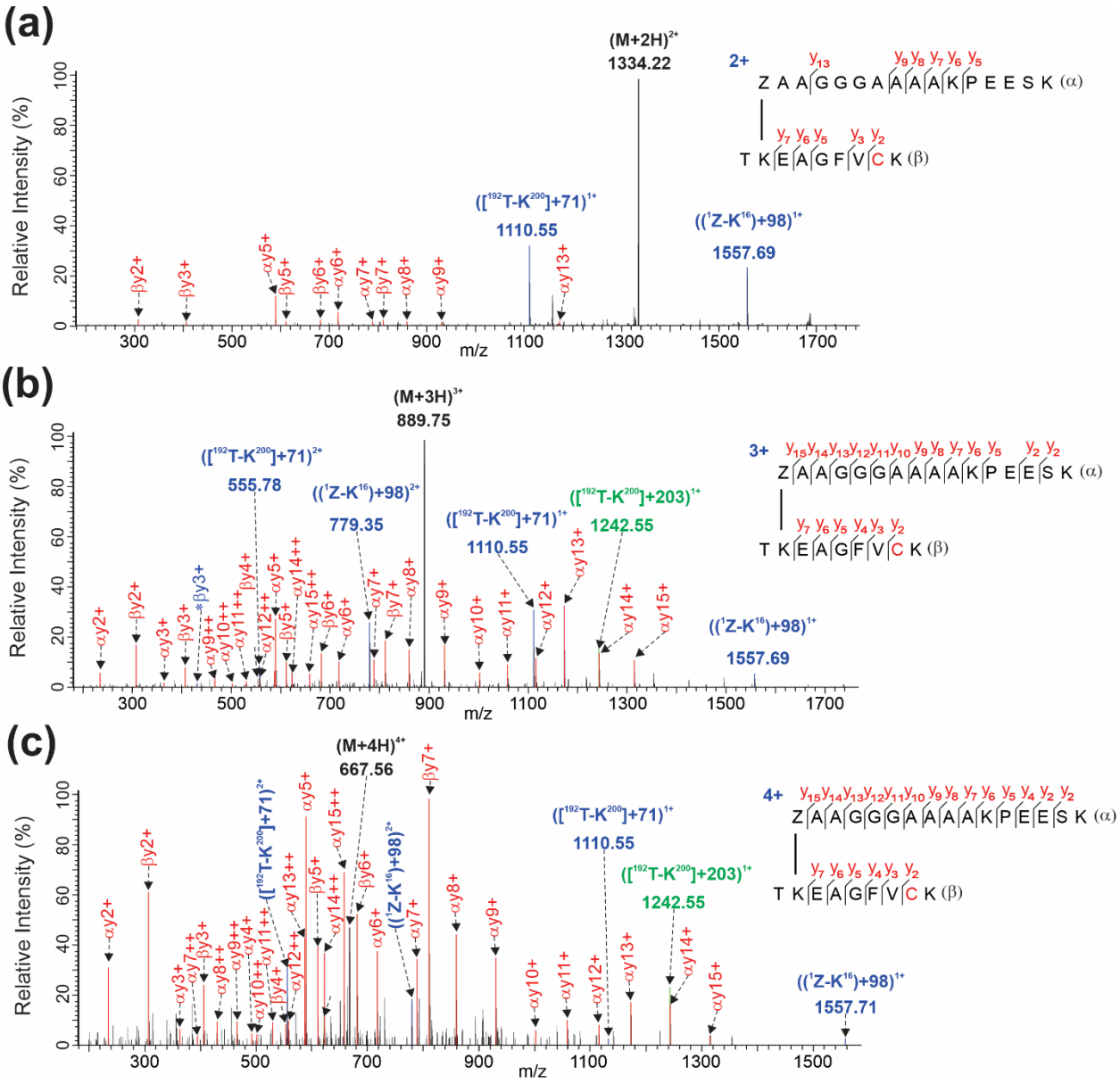


Figure 6. MS/MS spectra of the type 2 peptide [^{192}T-K^{200}]-(^{1}Z-K^{16}) with different charge states (2+ **(a)**, 3+ **(b)** and 4+ **(c)**) and fragmented by stepped HCD using normalized collision energies of 27, 30 and 33. The fragment ions assigned as ([^{192}T-K^{200}]+71)$^{n+}$, and ((^{1}Z-K^{16})+98)$^{n+}$, were originated by a metastable fragmentation of the linker pseudopeptide bond (see **Figure 3**). The amino acid indicated as Z corresponds to an *N*-terminal acetylated cysteine residue in Ac-Cys^{1}pP0 peptide. The sequence coverages obtained from each MS/MS spectrum. Other linker fragment ions ([^{192}T-

$K^{200}]+203)^+$, ($[^{192}T\text{-}K^{200}]+169)^+$, and $(^1Z\text{-}K^{16})^+$) generated by the thioether bond fragmentation were only observed in **(b)** and **(c)**.

CRediT

PERB: Data curation, Formal Analysis, Software, Writing - review & editing. **SP:** Data curation, Investigation, Methodology, Validation, Visualization. **PC:** Software, Resources, Funding Acquisition. **RSCB:** Investigation. **MB:** Investigation, Writing - original draft. **HH:** Investigation. **ARM:** Resources, Project Administration, Funding Acquisition. **VB:** Writing - review & editing. **TT:** Investigation, Supervision, Funding Acquisition, Writing - review & editing. **LJG:** Conceptualization, Supervision, Investigation, Methodology, Validation, Visualization, Writing - original draft; and Writing - review & editing.

Acknowledgements

The authors LJG, SP and ARM belong to the CYTED Network INCOGARR (118RT0541) and mobility was partially supported by CYTED. The authors thank FIOCRUZ for using their Network of Technological Platforms.

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, http://dx.doi.org/10.1039/C5PY00168D.
2. Hermanson, G., *Chapter 3-The Reactions of Bioconjugation. in (ed. Hermanson, GTBT-BT (Third E.) 229–258*. 2013, Academic Press, http://dx.doi.org/10.1016/B978-0-12-382239-0.00003-0.
3. Baldwin, A.D. and K.L. Kiick, *Tunable degradation of maleimide–thiol adducts in reducing environments.* Bioconjugate chemistry, 2011. **22**(10): p. 1946-1953, http://dx.doi.org/10.1021/bc200148v.
4. Ryan, C.P., et al., *Tunable reagents for multi-functional bioconjugation: reversible or permanent chemical modification of proteins and peptides by control of maleimide hydrolysis.* Chemical Communications, 2011. **47**(19): p. 5452-5454, http://dx.doi.org/10.1039/C1CC11114K.
5. Szijj, P.A., C. Bahou, and V. Chudasama, *Minireview: Addressing the retro-Michael instability of maleimide bioconjugates.* Drug Discovery Today: Technologies, 2018. **30**: p. 27-34, http://dx.doi.org/10.1016/j.ddtec.2018.07.002.
6. Lahnsteiner, M., et al., *Improving the stability of maleimide–thiol conjugation for drug targeting.* Chemistry–A European Journal, 2020. **26**(68): p. 15867-15870, http://dx.doi.org/10.1002/chem.202003951.
7. 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, http://dx.doi.org/10.1002/psc.3323.
8. Boyatzis, A.E., et al., *Limiting the hydrolysis and oxidation of maleimide–peptide adducts improves detection of protein thiol oxidation.* Journal of proteome research, 2017. **16**(5): p. 2004-2015, http://dx.doi.org/10.1021/acs.jproteome.6b01060.
9. Fontaine, S.D., et al., *Long-term stabilization of maleimide–thiol conjugates.* Bioconjugate chemistry, 2015. **26**(1): p. 145-152, http://dx.doi.org/10.1021/bc5005262.
10. González, L.J., et al., *On the utility of the extracted ion chromatograms for assigning the conjugation sites in bioconjugates synthesized by the maleimide-thiol chemistry.* 2023, http://dx.doi.org/10.26434/chemrxiv-2023-3g9g6-v2.
11. Han, L., Y. Zhao, and Q. Zhang, *Conjugation Site Analysis by MS/MS Protein Sequencing.* Antibody-Drug Conjugates: Methods and Protocols, 2020: p. 221-233, http://dx.doi.org/10.1007/978-1-4939-9929-3_15.
12. Strop, P., et al., *Location matters: site of conjugation modulates stability and pharmacokinetics of antibody drug conjugates.* Chemistry & biology, 2013. **20**(2): p. 161-167, https://doi.org/10.1016/j.chembiol.2013.01.010.
13. Yılmaz, Ş., et al., *Cross-linked peptide identification: A computational forest of algorithms.* Mass spectrometry reviews, 2018. **37**(6): p. 738-749, http://dx.doi.org/10.1002/mas.21559.
14. Schilling, B., et al., *MS2Assign, automated assignment and nomenclature of tandem mass spectra of chemically crosslinked peptides.* Journal of the American Society for Mass Spectrometry, 2003. **14**(8): p. 834-850, http://dx.doi.org/10.1016/S1044-0305(03)00327-1.
15. Rodriguez Mallon, 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), http://dx.doi.org/10.3390/pathogens9060513.
16. 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.* Analytical and Bioanalytical Chemistry, 2021. **413**(23): p. 5885-5900, http://dx.doi.org/10.1007/s00216-021-03569-0.

17. Chen, Z.L., et al., *A high-speed search engine pLink 2 with systematic evaluation for proteome-scale identification of cross-linked peptides.* Nat Commun, 2019. **10**(1): p. 3404, http://dx.doi.org/10.1038/s41467-019-11337-z.
18. Gotze, M., et al., *StavroX--a software for analyzing crosslinked products in protein interaction studies.* J Am Soc Mass Spectrom, 2012. **23**(1): p. 76-87, http://dx.doi.org/10.1007/s13361-011-0261-2.
19. Hoopmann, M.R., et al., *Kojak: efficient analysis of chemically cross-linked protein complexes.* J Proteome Res, 2015. **14**(5): p. 2190-8, http://dx.doi.org/10.1021/pr501321h.
20. Trnka, M.J., et al., *Matching cross-linked peptide spectra: only as good as the worse identification.* Mol Cell Proteomics, 2014. **13**(2): p. 420-34, http://dx.doi.org/10.1074/mcp.M113.034009.
21. Lima, D.B., et al., *SIM-XL: A powerful and user-friendly tool for peptide cross-linking analysis.* Journal of proteomics, 2015. **129**: p. 51-55, http://dx.doi.org/10.1016/j.jprot.2015.01.013.
22. Rappsilber, J., *The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes.* Journal of structural biology, 2011. **173**(3): p. 530-540, http://dx.doi.org/10.1016/j.jsb.2010.10.014.
23. Iacobucci, C., et al., *First community-wide, comparative cross-linking mass spectrometry study.* Analytical chemistry, 2019. **91**(11): p. 6953-6961, http://dx.doi.org/10.1021/acs.analchem.9b00658.
24. Müller, D., et al., *Isotope-tagged cross-linking reagents. A new tool in mass spectrometric protein interaction analysis.* Analytical chemistry, 2001. **73**(9): p. 1927-1934, http://dx.doi.org/10.1021/ac001379a.
25. Zelter, A., et al., *Isotope signatures allow identification of chemically cross-linked peptides by mass spectrometry: a novel method to determine interresidue distances in protein structures through cross-linking.* Journal of proteome research, 2010. **9**(7): p. 3583-3589, http://dx.doi.org/10.1021/pr1001115.
26. Sinz, A., *Divide and conquer: cleavable cross-linkers to study protein conformation and protein–protein interactions.* Analytical and bioanalytical chemistry, 2017. **409**(1): p. 33-44, http://dx.doi.org/10.1007/s00216-016-9941-x.
27. Arlt, C., et al., *Integrated workflow for structural proteomics studies based on cross-linking/mass spectrometry with an MS/MS cleavable cross-linker.* Analytical chemistry, 2016. **88**(16): p. 7930-7937, http://dx.doi.org/10.1021/acs.analchem.5b04853.
28. Back, J., et al., *A new crosslinker for mass spectrometric analysis of the quaternary structure of protein complexes.* Journal of the American Society for Mass Spectrometry, 2001. **12**(2): p. 222-227, http://dx.doi.org/10.1016/S1044-0305(00)00212-9.
29. Tan, D., et al., *Trifunctional cross-linker for mapping protein-protein interaction networks and comparing protein conformational states.* Elife, 2016. **5**: p. e12509, http://dx.doi.org/10.7554/eLife.12509.
30. Muller, M.Q., et al., *Cleavable cross-linker for protein structure analysis: reliable identification of cross-linking products by tandem MS.* Analytical chemistry, 2010. **82**(16): p. 6958-6968, http://dx.doi.org/10.1021/ac101241t.
31. Soderblom, E.J. and M.B. Goshe, *Collision-induced dissociative chemical cross-linking reagents and methodology: applications to protein structural characterization using tandem mass spectrometry analysis.* Analytical chemistry, 2006. **78**(23): p. 8059-8068, http://dx.doi.org/10.1021/ac0613840.
32. Soderblom, E.J., et al., *Tandem mass spectrometry acquisition approaches to enhance identification of protein-protein interactions using low-energy collision-induced dissociative*

chemical crosslinking reagents. Rapid Communications in Mass Spectrometry, 2007. **21**(21): p. 3395-3408, http://dx.doi.org/10.1002/rcm.3213.
33. Dreiocker, F., et al., *Collision-induced dissociative chemical cross-linking reagent for protein structure characterization: applied Edman chemistry in the gas phase.* Journal of mass spectrometry, 2010. **45**(2): p. 178-189, http://dx.doi.org/10.1002/jms.1702.
34. Lewis, C.E. and J.A. Roth, *Challenges in having vaccines available to control transboundary diseases of livestock.* Current Issues in Molecular Biology, 2021. **42**(1): p. 1-40, http://dx.doi.org/10.21775/cimb.042.001.
35. Rodríguez-Mallon, A., et al., *A chemical conjugate of the tick P0 peptide is efficacious against Amblyomma mixtum.* Transboundary and emerging diseases, 2020. **67**: p. 175-177, http://dx.doi.org/10.1111/tbed.13455.
36. Rodríguez-Mallon, A., et al., *A novel tick antigen shows high vaccine efficacy against the dog tick, Rhipicephalus sanguineus.* Vaccine, 2012. **30**(10): p. 1782-1789, http://dx.doi.org/10.1016/j.vaccine.2012.01.011.
37. Rodríguez-Mallon, A., et al., *High efficacy of a 20 amino acid peptide of the acidic ribosomal protein P0 against the cattle tick, Rhipicephalus microplus.* Ticks and Tick-borne Diseases, 2015. **6**(4): p. 530-537, http://dx.doi.org/10.1016/j.ttbdis.2015.04.007.
38. Verez-Bencomo, V., et al., *A synthetic conjugate polysaccharide vaccine against Haemophilus influenzae type b.* Science, 2004. **305**(5683): p. 522-525, http://dx.doi.org/10.1126/science.1095209.
39. van Der Put, R.M., et al., *A synthetic carbohydrate conjugate vaccine candidate against shigellosis: improved bioconjugation and impact of alum on immunogenicity.* Bioconjugate chemistry, 2016. **27**(4): p. 883-892, http://dx.doi.org/10.1021/acs.bioconjchem.5b00617.
40. Valdes-Balbin, Y., et al., *SARS-CoV-2 RBD-Tetanus toxoid conjugate vaccine induces a strong neutralizing immunity in preclinical studies.* ACS Chemical Biology, 2021. **16**(7): p. 1223-1233, http://dx.doi.org/10.1021/acschembio.1c00272.
41. Toledo-Romaní, M.E., et al., *Safety and efficacy of the two doses conjugated protein-based SOBERANA-02 COVID-19 vaccine and of a heterologous three-dose combination with SOBERANA-Plus: a double-blind, randomised, placebo-controlled phase 3 clinical trial.* The Lancet Regional Health–Americas, 2023. **18**, http://dx.doi.org/10.1016/j.lana.2022.100423.
42. Suarez, M., et al., *High impact and effectiveness of Gavac™ vaccine in the national program for control of bovine ticks Rhipicephalus microplus in Venezuela.* Livestock Science, 2016. **187**: p. 48-52, http://dx.doi.org/10.1016/j.livsci.2016.02.005.
43. Canales, M., et al., *Large-scale production in Pichia pastoris of the recombinant vaccine Gavac™ against cattle tick.* Vaccine, 1997. **15**(4): p. 414-422, http://dx.doi.org/10.1016/S0264-410X(96)00192-2.
44. 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, http://dx.doi.org/10.1111/j.1399-3011.1990.tb00939.x.
45. Garay, H., et al., *Characterization of low-abundance species in the active pharmaceutical ingredient of CIGB-300: A clinical-grade anticancer synthetic peptide.* Journal of Peptide Science, 2018. **24**(6): p. e3081, https://doi.org/10.1002/psc.3081.
46. 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, http://dx.doi.org/10.1002/rcm.9660.
47. Harrison, A.G. and A.B. Young, *Fragmentation reactions of deprotonated peptides containing proline. The proline effect.* J Mass Spectrom, 2005. **40**(9): p. 1173-86, http://dx.doi.org/10.1002/jms.891.

48. Thorne, G.C., K.D. Ballard, and S.J. Gaskell, *Metastable decomposition of peptide [M+ H]+ ions via rearrangement involving loss of the C-terminal amino acid residue.* Journal of the American Society for Mass Spectrometry, 1990. **1**(3): p. 249-257, http://dx.doi.org/10.1002/(SICI)1096-9888(199602)31:2<150::AID-JMS287>3.0.CO;2-5.
49. Wysocki, V.H., et al., *Mobile and localized protons: a framework for understanding peptide dissociation.* Journal of Mass Spectrometry, 2000. **35**(12): p. 1399-1406, https://doi.org/10.1002/1096-9888(200012)35:12<1399::aid-jms86>3.0.co;2-r
50. Dongré, A.R., et al., *Influence of peptide composition, gas-phase basicity, and chemical modification on fragmentation efficiency: evidence for the mobile proton model.* Journal of the American Chemical Society, 1996. **118**(35): p. 8365-8374, https://doi.org/10.1021/ja9542193
51. Mormann, M., et al., *Fragmentation of intra-peptide and inter-peptide disulfide bonds of proteolytic peptides by nanoESI collision-induced dissociation.* Analytical and bioanalytical chemistry, 2008. **392**: p. 831-838, http://dx.doi.org/10.1007/s00216-008-2258-7.
52. Jedrychowski, M.P., et al., *Evaluation of HCD-and CID-type fragmentation within their respective detection platforms for murine phosphoproteomics.* Molecular & Cellular Proteomics, 2011. **10**(12), http://dx.doi.org/10.1074/mcp.M111.009910.
53. McAlister, G.C., et al., *Higher-energy collision-activated dissociation without a dedicated collision cell.* Molecular & Cellular Proteomics, 2011. **10**(5), https://doi.org/10.1074/mcp.O111.009456 – 1.

Supplemental files

Title: *N*-propionyl thiosuccinimide linker upon hydrolysis becomes into a cleavable crosslinker by a gas phase metastable fragmentation and makes more reliable the conjugation sites identification

Auhtors: Pablo Enmanuel Ramos[1,*], Satomy Pousa[2,*], Paulo Carvalho[4], Rodrigo Soares Caldeira Brant[4], Michel Batista[4], Hironobu Hojo[3], Alina Rodríguez Mallón[5], Vladimir Besada[2], Toshifumi Takao[6] and Luis Javier González[2,¥].

Affiliations:

[1]Bioinformatics Department [2] Mass Spectrometry Laboratory. Proteomics Department, and [5]Animal Biotechnology Department. Center for Genetic Engineering and Biotechnology. Avenida 31 e/ 158 y 190. Cubanacán. Playa. PO. Box 6162. Havana. Cuba.
[3]Laboratory for Protein Synthesis and [6]Functional Proteomics and Protein Profiling, Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan.
[4]Carlos Chagas Institute/FioCruz Rua Professor Algacyr Munhoz Mader, 3775 CIC 81350-010, Curitiba, Brasil.

(*) Authors with equal contributions
(¥) Corresponding author

Luis Javier González PhD
Mass Spectrometry Laboratory, Department of Proteomics,
Center for Genetic Engineering and Biotechnology, Avenida 31 e/ 158 y 190,
Cubanacán, Playa, PO Box. 6162. La Habana, Cuba
Email: luis.javier@cigb.edu.cu

Index:

Fig. S1. Summary of the fragment ions observed in the MALDI-MS/MS spectra of the type 2 peptides **(a)** $(P1\text{-}P2)_T$ and **(b)** $(P1\text{-}P2)_H$ crosslinked by thiazine and the hydrolyzed thiosuccinimide linkers, respectively. The amino acid sequences of P1 and P2 are βAAAGGGAAAAK and CAAGGGAAAAK, respectively. The m/z values highlighted in bold correspond to fragment ions observed when the MALDI-MS/MS was acquired in presence and absence of a collision gas while non-bold characters correspond to fragment ions that were only detected when the collision gas was ON. In the linker structures the new pseudopeptide bond generated by the transcyclization **(a)** and hydrolysis **(b)** are highlighted in red. To simplify the figure, in **(b)** the represented linker structure corresponds to only one of the two possible positional isomers of $(P1\text{-}P2)_H$.

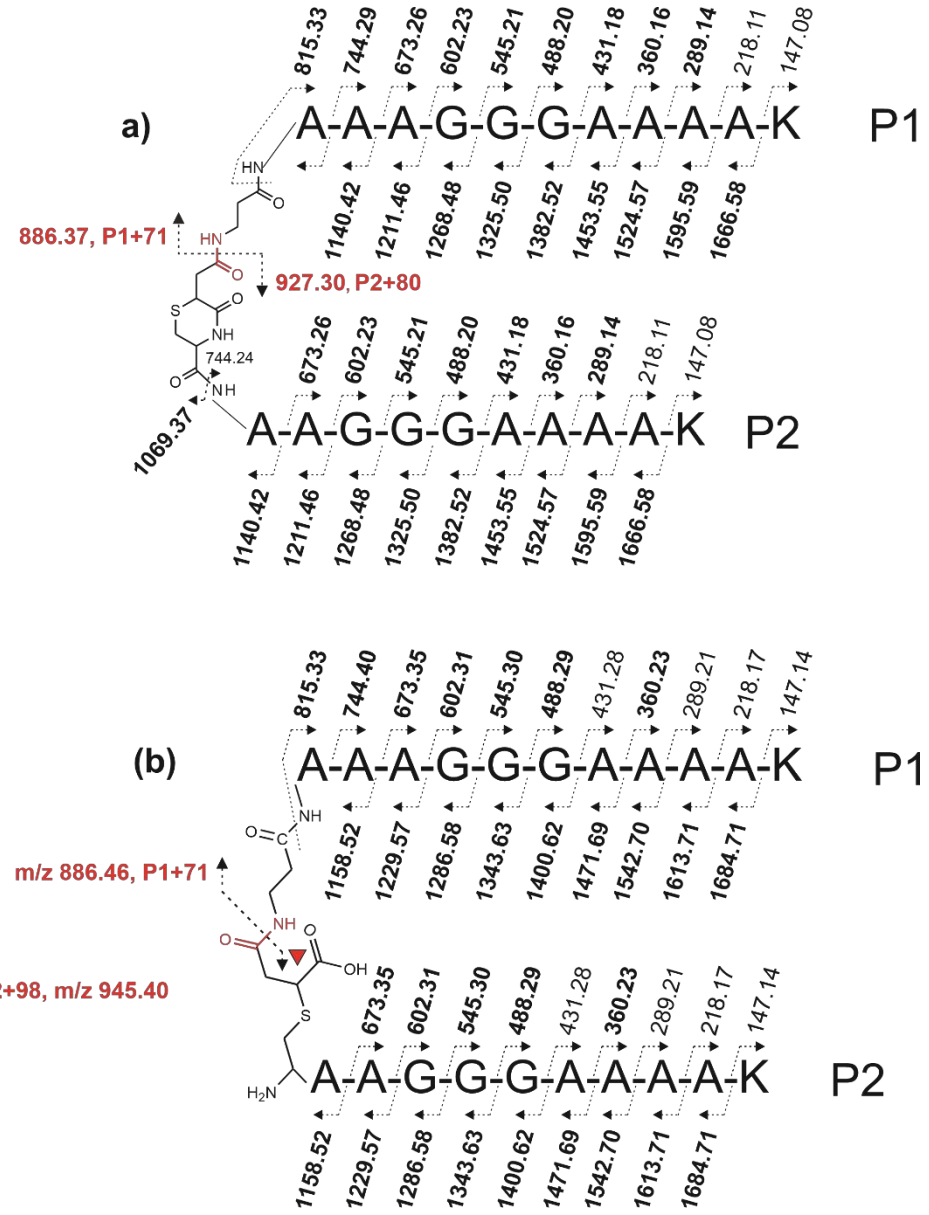
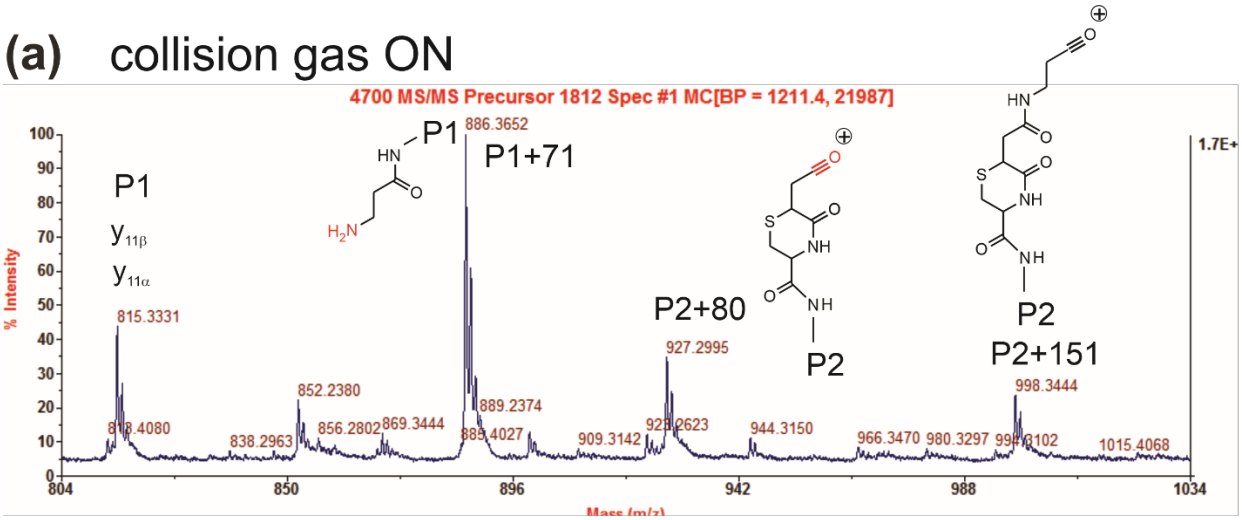


Fig. S2. Expanded regions (*m/z* 804-1034) in **(a)** and **(b)** correspond to the MALDI-MS/MS spectra of type 2 peptide $(P1\text{-}P2)_T$ acquired by using and not using collision gas, respectively. The nomenclature and the proposed structures for the observed fragment derived from the linker are in agreement with previous manuscript [1].



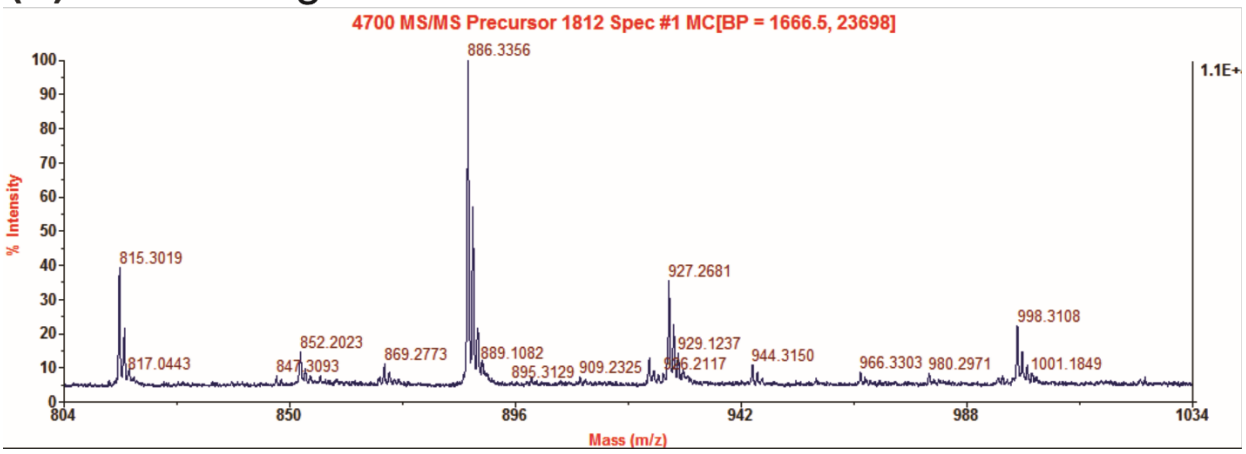
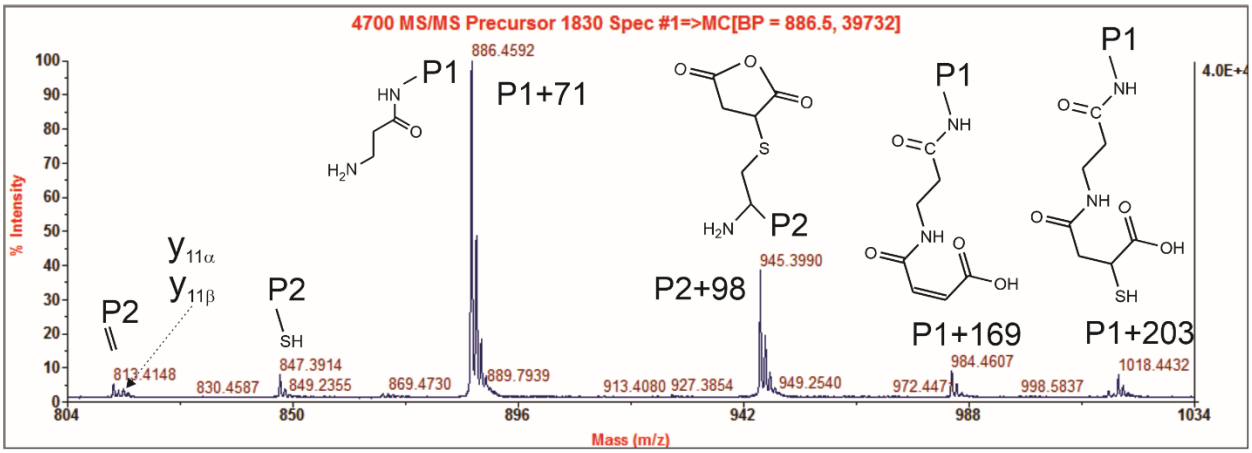


Fig. S3. Expanded regions (m/z 804-1034) in **(a)** and **(b)** correspond to the MALDI-MS/MS spectra of type 2 peptide $(P1\text{-}P2)_H$ with the hydrolyzed thiosuccinimide linker acquired by using and not using collision gas, respectively. The nomenclature and the proposed structures for the observed fragment derived from the linker are in agreement with previous manuscript [1]. The fragment ion represented as P2= corresponds to P2 peptide with the N-terminal Cys residue transformed as a dehydroalanine residue.

(a) collision gas ON



(b) collision gas OFF

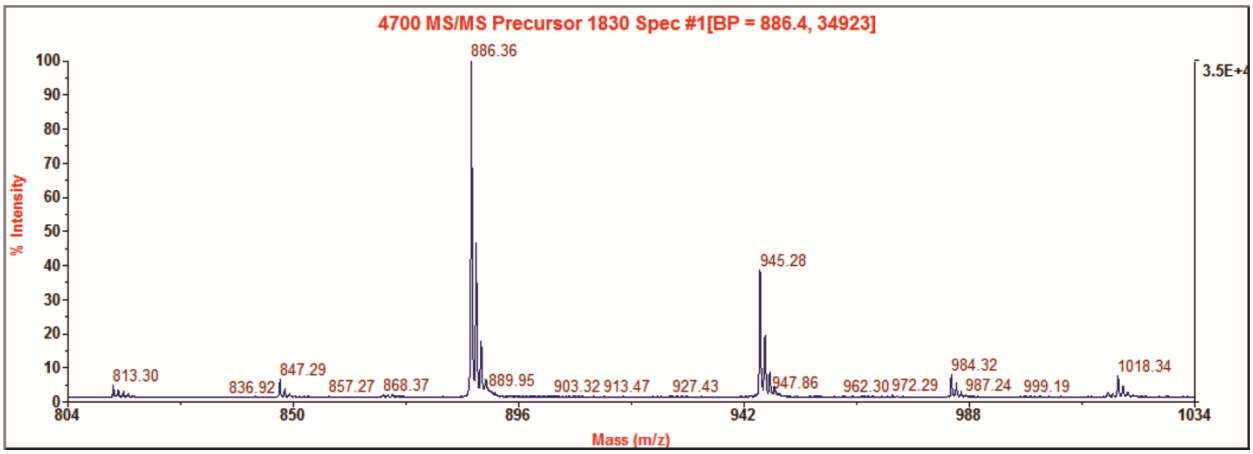


Fig. S4. MALDI-MS/MS analysis of the type 2 peptides $(P2\text{-}P4)_T$ and $(P2\text{-}P4)_H$ crosslinked by a thiazine and the hydrolyzed thiosuccinimide linkers, respectively. The amino acid sequences of P2 and P4 are Ac-GANAPKEPQR and CAAGGGAAAAK, respectively, where Ac- means an acetyl group at the N-terminal end of the P4 peptide. The figure also shows a comparison of the MS/MS spectra acquired with collision gas (see **(a)** and **(c)**) and without collision gas (see **(b)** and **(d)**). The fragment ions P4+71, P2+80 and P2+98 generated by the fragmentation of the new pseudopeptide bond present in linkers is highlighted in red. The fragment ion assigned as $b_{10\beta}+H_2O$ present in all MS/MS spectra corresponds to a C-terminal metastable fragmentation of the $(M+H)^+$ parent ions as described previously [2, 3]. The fragment ions $y_{3\alpha}$ and $y_{10\alpha}$ are originated by a favorable y_n fragmentation at the *N*-terminal of the two proline residues present in the sequence of P4 peptide [4]. The fragment ion indicated as P4 correspond to a fragmentation of the pseudopeptide bond between the lysine residue in P4 peptide and the linkers, either thiazine or hydrolyzed thiosuccinimide. The fragment ions labeled with an empty and filled circles correspond the P4+169 and P4+203, respectively [1]. The fragment ions labeled with an empty and filled stars correspond the reduced P2 peptide (P2-SH) and also to the P2 peptide with the Cys residue transformed as a dehydroalanine, respectively [1]. The nomenclature of the fragment ions is in agreement with the proposed by Schilling *et. al.* for the crosslinked peptides [5]. The legend in **(e)** indicates the assignment as well as the *m/z* of the corresponding ions.

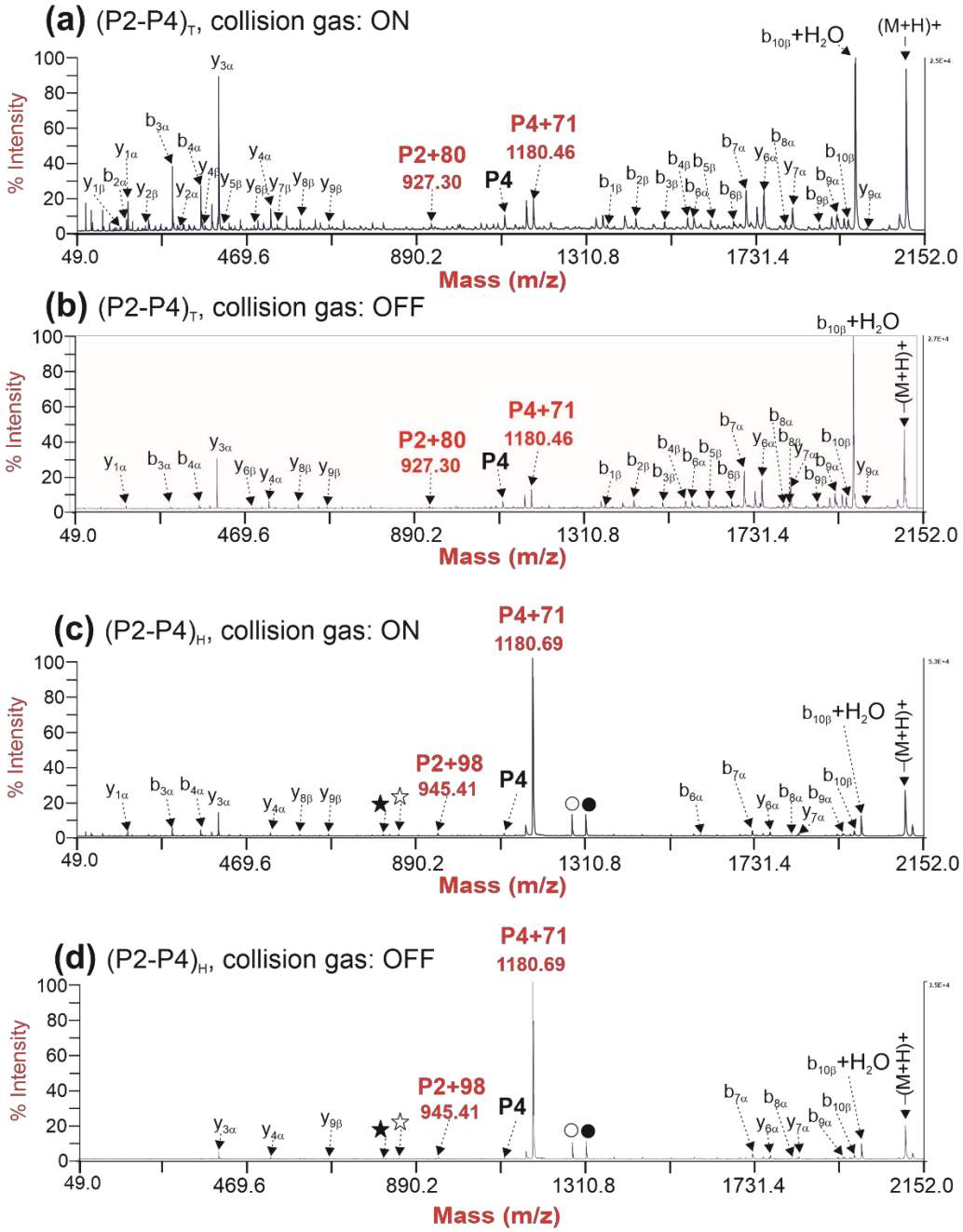
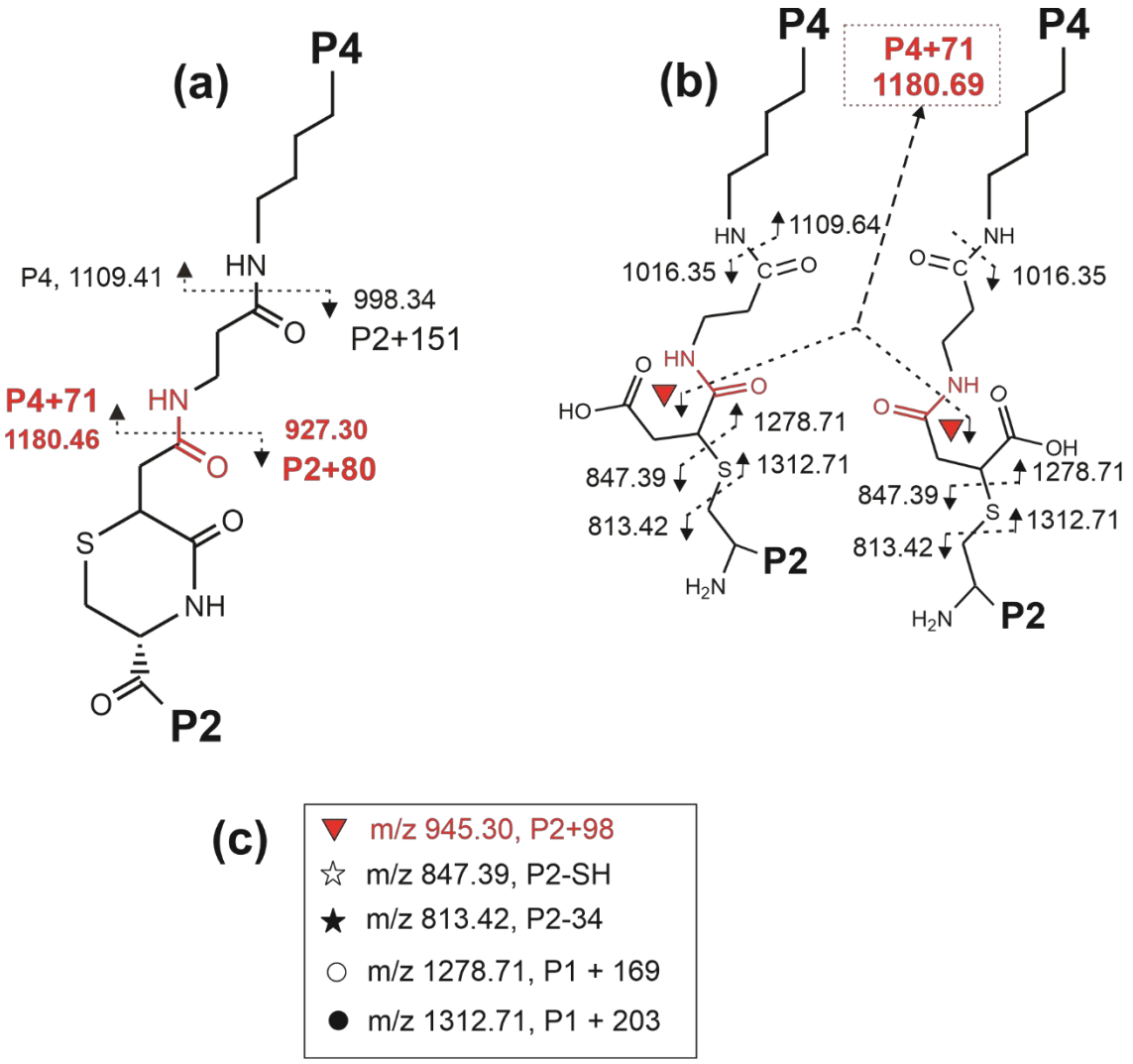


Fig S5. Structures of thiazine **(a)** and the positional hydrolyzed thiosuccinimide linkers **(b)** in the $(P2\text{-}P4)_T$ and $(P2\text{-}P4)_H$ type 2 peptides, respectively and their corresponding fragmentations observed in the MALDI-MS/MS spectra shown in **Fig. S4**. The legend in **(c)** shows the assignment of the fragment ions derived from both linkers.



(c)

- ▼ m/z 945.30, P2+98
- ☆ m/z 847.39, P2-SH
- ★ m/z 813.42, P2-34
- ○ m/z 1278.71, P1 + 169
- ● m/z 1312.71, P1 + 203

Fig. S6. Summary of the backbone fragment ions observed in the MALDI-MS/MS spectra of the type 2 peptides **(a)** $(P2\text{-}P4)_T$ and **(b)** $(P2\text{-}P4)_H$ crosslinked by thiazine and the hydrolyzed thiosuccinimide linkers, respectively. The amino acid sequences of P2 and P4 are Ac-GANAPKEPQR and CAAGGGAAAAK, respectively, where Ac- means an acetyl group at the *N*-terminal end of P4. The alpha and beta peptides are indicated in **(a)** $(P2\text{-}P4)_T$ and **(b)** $(P2\text{-}P4)_H$. The m/z values highlighted in bold correspond to fragment ions observed when the MALDI-MS/MS was acquired in presence and absence of a collision gas. The legend in **(c)** corresponds to the linker fragment ions observed in **Fig. S5**. In the linker structures the new pseudopeptide bond generated by the transcyclization **(a)** and hydrolysis **(b)** are highlighted in red. To simplify the figure, in **(b)** the represented linker structure corresponds to only one of the two positional isomers of $(P2\text{-}P4)_H$.

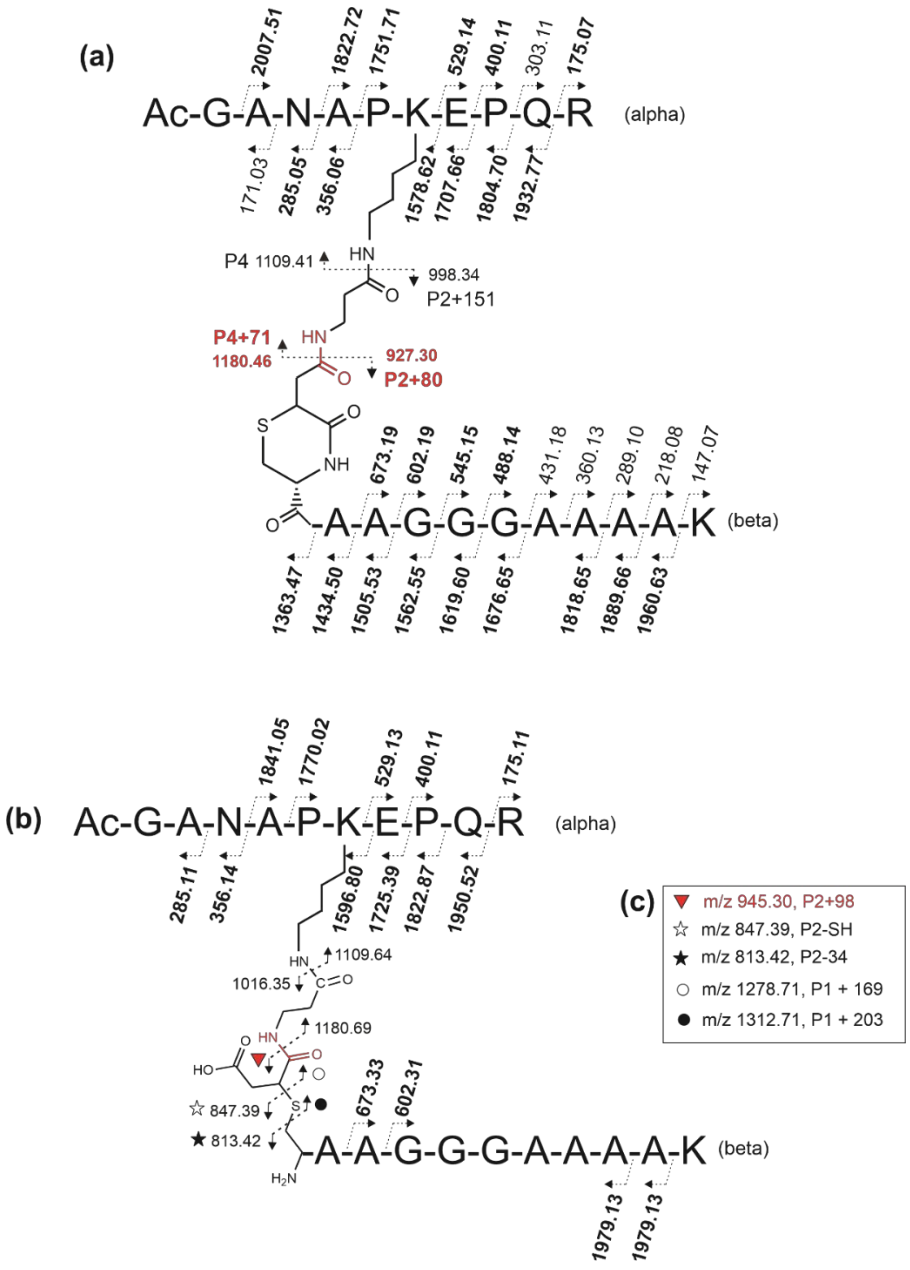
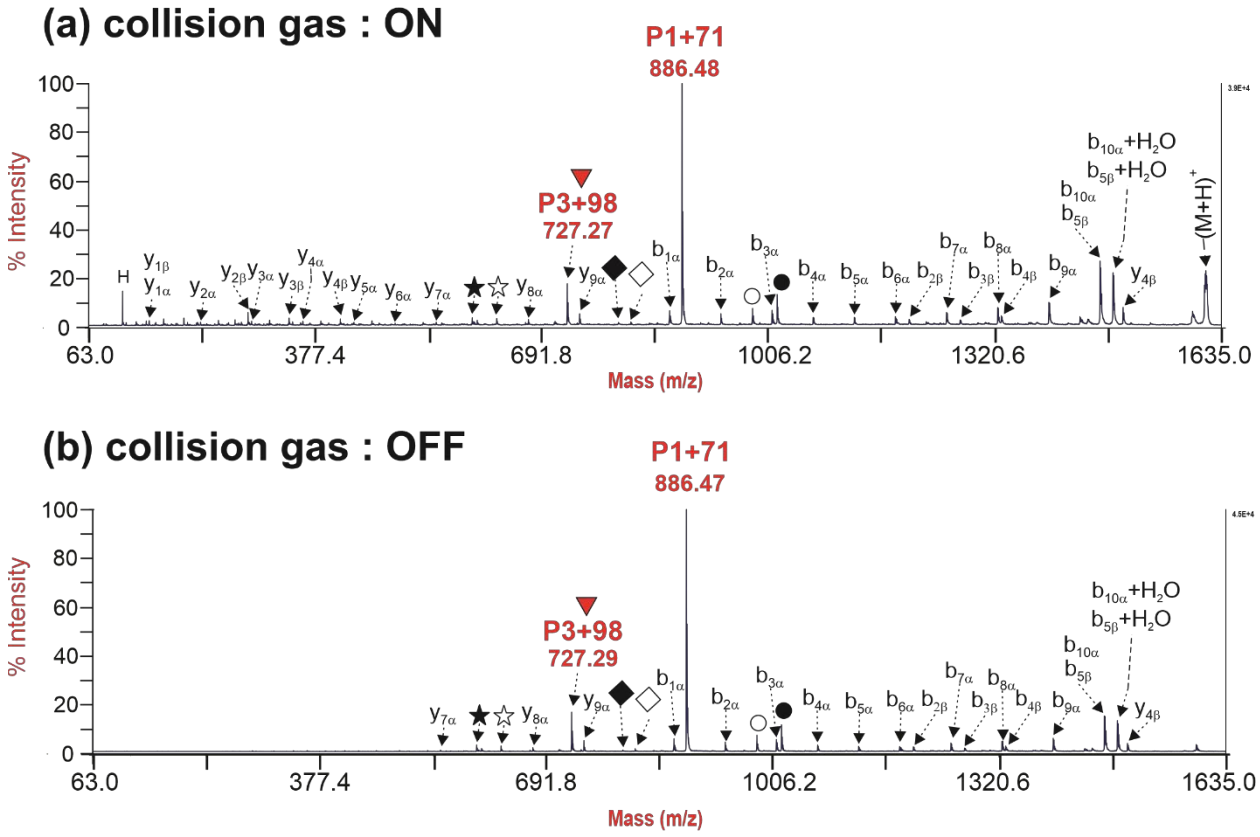


Fig. S7. MALDI-MS/MS analysis of the type 2 peptides $(P1\text{-}P3)_H$ with the hydrolyzed thiosuccinimide linker acquired with collision gas **(a)** and without collision gas **(b)**. The amino acid sequences of P1 and P3 are CAAGGGAAAAK and NCAGHK, respectively. The fragment ions P1+71 and P3+98 generated by the fragmentation of the newly generated pseudopeptide bond after the hydrolysis of the thiosuccinimide linker is highlighted in red. The isobaric fragment ion assigned as $b_{10\alpha}+H_2O$ and $b_{4\beta}+H_2O$ present in all MS/MS spectra corresponds to a C-terminal metastable fragmentation of the $(M+H)^+$ parent ions as described previously [2, 3]. The fragment ions labeled with an empty and filled circles correspond the P1+169 and P1+203, respectively [1]. The fragment ions labeled with an empty and filled stars correspond the reduced P3 peptide (P3-SH) and also to the P3 peptide with the Cys residue transformed as a dehydroalanine, respectively [1]. The structures of these fragment ions are the same shown in **Fig. S3**. The *m/z* values of all fragment ions are summarized in (e). The nomenclature of the fragment ions is in agreement with the proposed by Schilling *et. al.* for the crosslinked peptides [5]. The legend in **(d)** indicates the assignment as well as the *m/z* of the linker fragment ions. **(e)** Summary of the fragment ions observed in the MALDI-MS/MS spectra of the type 2 peptide $(P1\text{-}P3)_H$ crosslinked by a hydrolyzed thiosuccinimide linker. The alpha and beta peptides are indicated in **(e)**. The *m/z* values highlighted in bold correspond to fragment ions observed when the MALDI-MS/MS was acquired in presence and absence of a collision gas. To simplify the figure, in **(b)** the represented linker structure corresponds to only one of the two positional isomers of $(P1\text{-}P3)_H$.



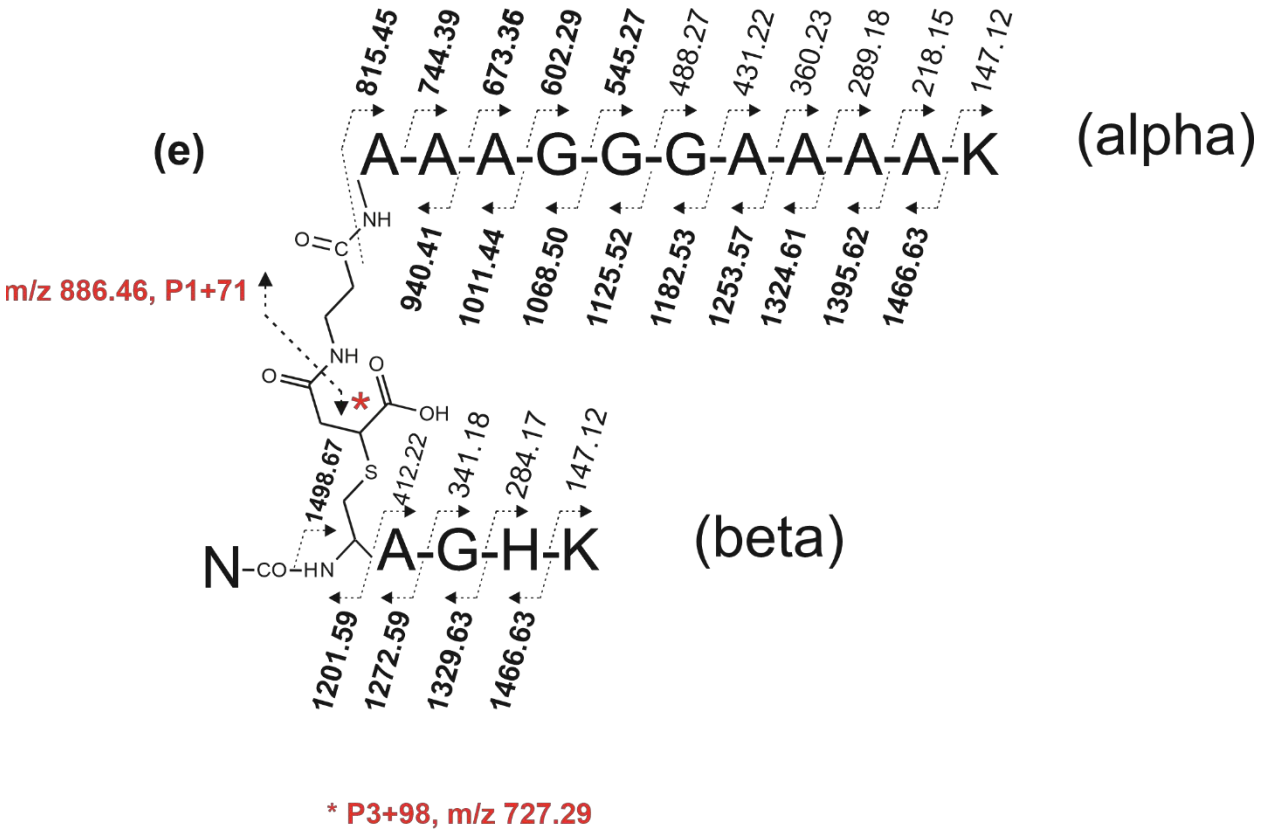
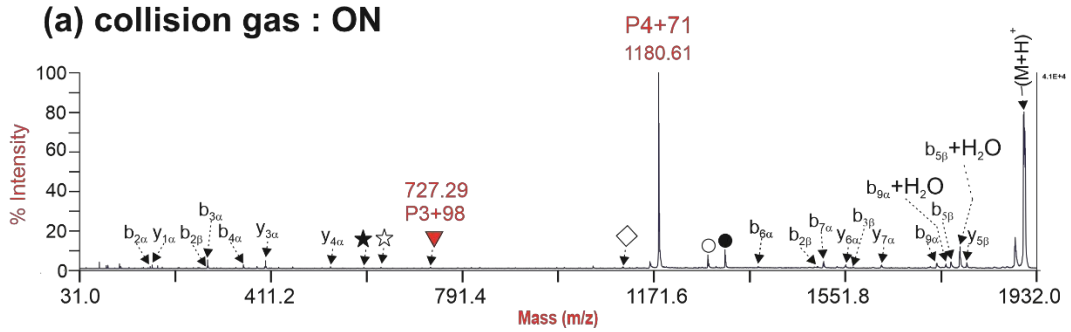
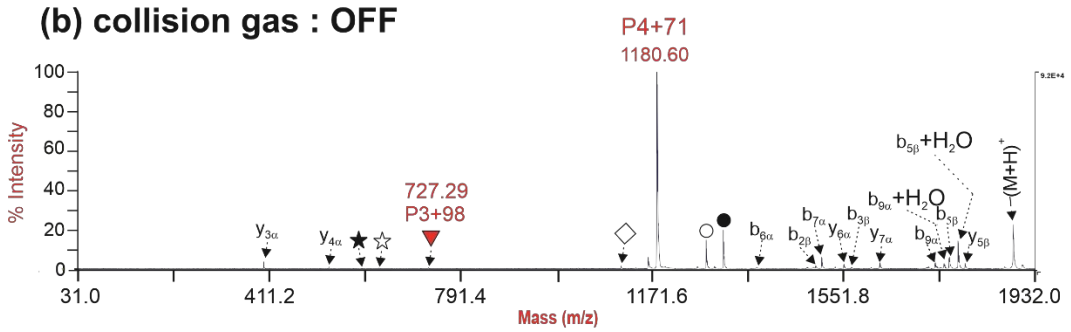
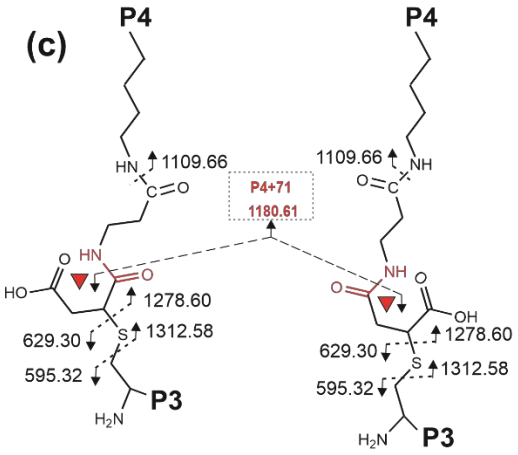
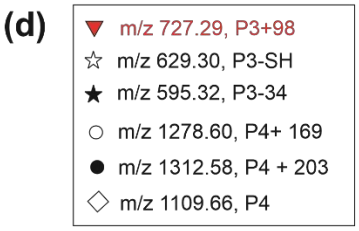


Fig. S8. MALDI-MS/MS analysis of the type 2 peptides $(P3\text{-}P4)_H$ with the hydrolyzed thiosuccinimide linker acquired with collision gas **(a)** and without collision gas **(b)**. The amino acid sequences of P3 and P4 are NCAGHK and CH_3CO-GANAPKEPQR, respectively, where CH_3CO- at the N-terminal end of P4. The fragment ions P4+71 and P3+98 generated by the fragmentation of the new pseudopeptide bond generated after the hydrolysis of the thiosuccinimide linker is highlighted in red. The isobaric fragment ion assigned as $b_{9\alpha}+H_2O$ and $b_{5\beta}+H_2O$ present in both MS/MS spectra (**(a)** and **(b)**) correspond to a *C*-terminal metastable fragmentation of the $(M+H)^+$ as described previously [2, 3]. The linker fragment ions labeled with an empty and filled circles in both MS/MS spectra correspond the P4+169 and P4+203 ions, respectively [1] as indicated in **(c)**. The fragment ions labeled with an empty and filled stars correspond the reduced P3 peptide (P3-SH) and also to the P3 peptide with the Cys residue transformed as a dehydroalanine, respectively [1]. The structures of these fragment ions are the same shown in **Fig. S3**. The *m/z* values of all fragment ions are summarized in **(e)**. The nomenclature of the fragment ions is in agreement with the proposed by Schilling *et. al.* for the crosslinked peptides [5]. The legend in **(d)** indicates the assignment as well as the *m/z* of the corresponding ions. **(e)** Summary of the fragment ions observed in the MALDI-MS/MS spectra of the type 2 peptide $(P3\text{-}P4)_H$ crosslinked by a hydrolyzed thiosuccinimide linker. The alpha and beta peptides are indicated. The m/z highlighted in bold correspond to fragment ions observed when the MALDI-MS/MS was acquired in presence and absence of a collision gas. To simplify the figure, in **(e)** the represented linker structure corresponds to only one of the two positional isomers of $(P3\text{-}P4)_H$.









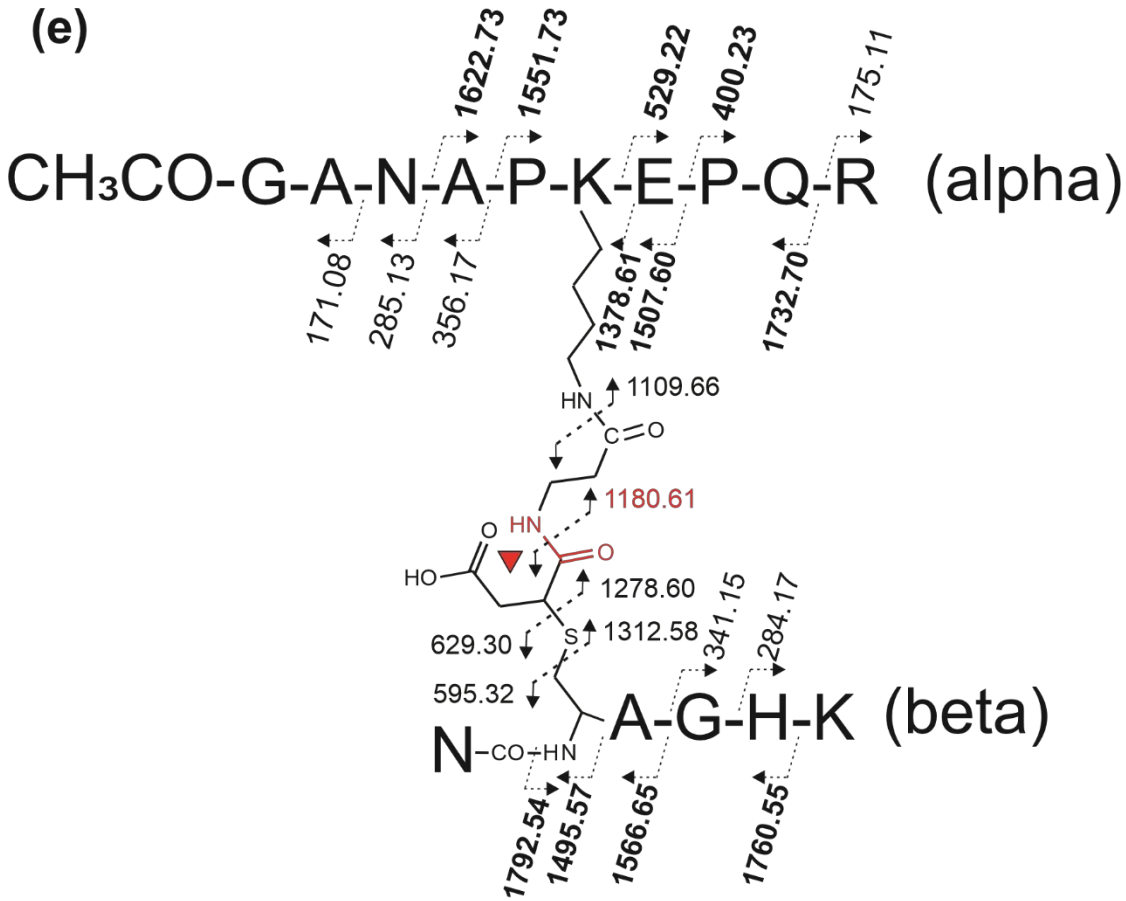


Fig. S9. Fragmentation of the thioether bond in a hydrolyzed thiosuccinimide linker between a lysine and cysteine residue. The Lysine and Cysteine residues in this structure represent the P1 and P2 peptides crosslinked by a hydrolyzed thiosuccinimide linker ($(P1\text{-}P2)_H$). The fragmentation of the thioether bond by a beta elimination and a proton transfer yields two linker fragment ions designed here as P1+169 Da and the reduced P2-SH.

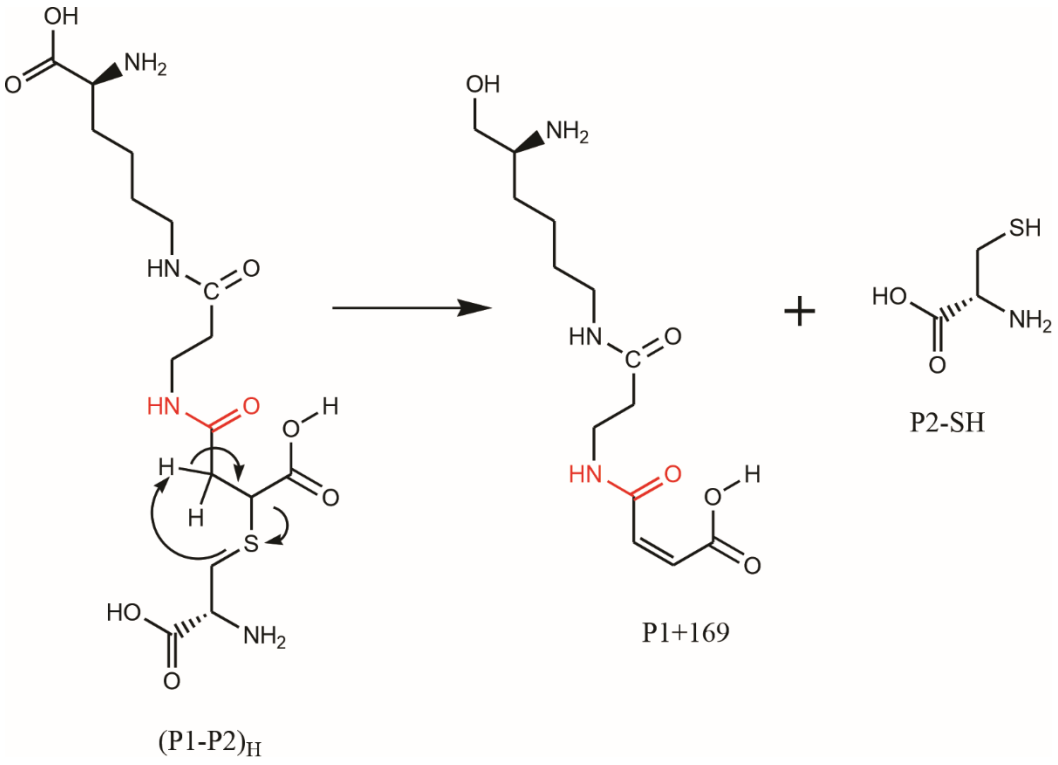


Fig. S10. Fragmentation of the thioether bond in a hydrolyzed thiosuccinimide linker between a lysine and cysteine residue. The Lysine and Cysteine residues in this structure represent the P1 and P2 peptides crosslinked by a hydrolyzed thiosuccinimide linker ($(P1\text{-}P2)_H$). The fragmentation of the thioether bond by a beta elimination and a proton transfer yields two linker fragment ions designed here as P1+203 Da and the reduced P2-34 Da.

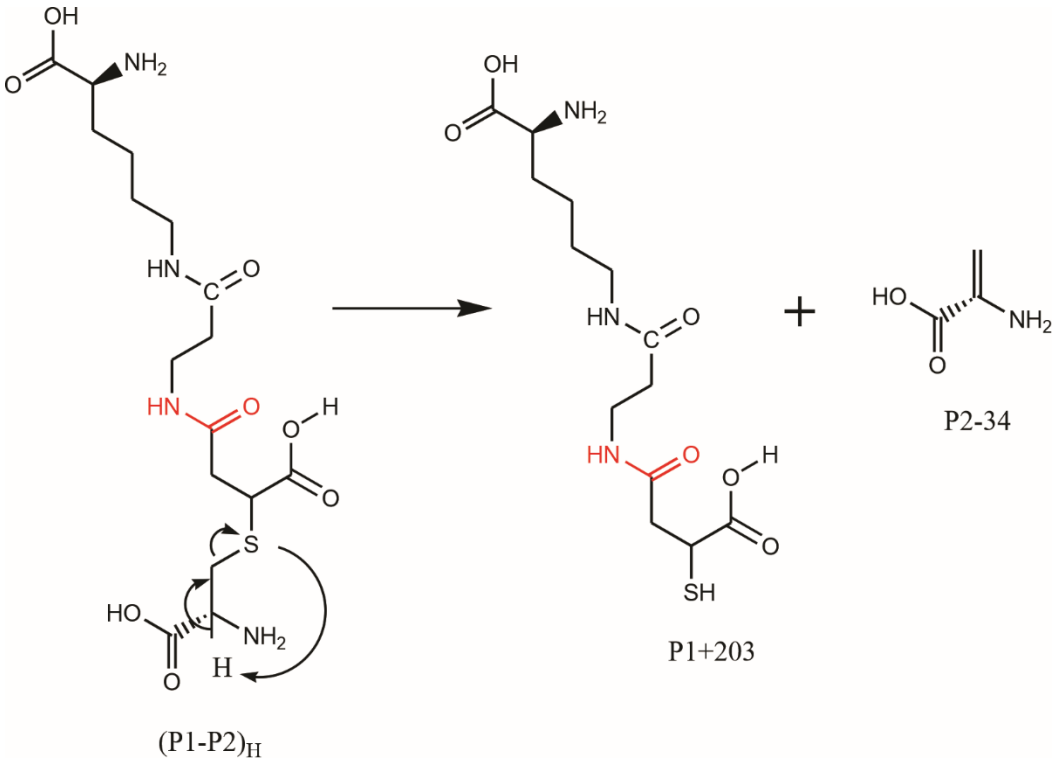


Fig. S11. Summary of the backbone fragment ions (b_n and y_n) observed in the MALDI-MS/MS spectrum of the partially ^{18}O-labeled $(P1\text{-}P3)_H$ at the new carboxyl group generated upon hydrolysis of five-membered ring in a buffer prepared with 30 % (v/v) of $H_2{}^{18}O$. Backbone fragment ions, either the b_n and y_n, written in bold are those that showed in the MALDI-MS/MS an isotopic ion distribution that evidenced the incorporation of ^{18}O in their structures. A red asterisk indicates in the structure of $(P1\text{-}P3)_H$ the position where the ^{18}O was incorporated. In $(P1\text{-}P3)_H$ the alpha and beta peptides are P1 and P3, respectively. To simplify the figure, the structure of only one of the two positional isomers of the hydrolyzed thiosuccinimide linker was represented.

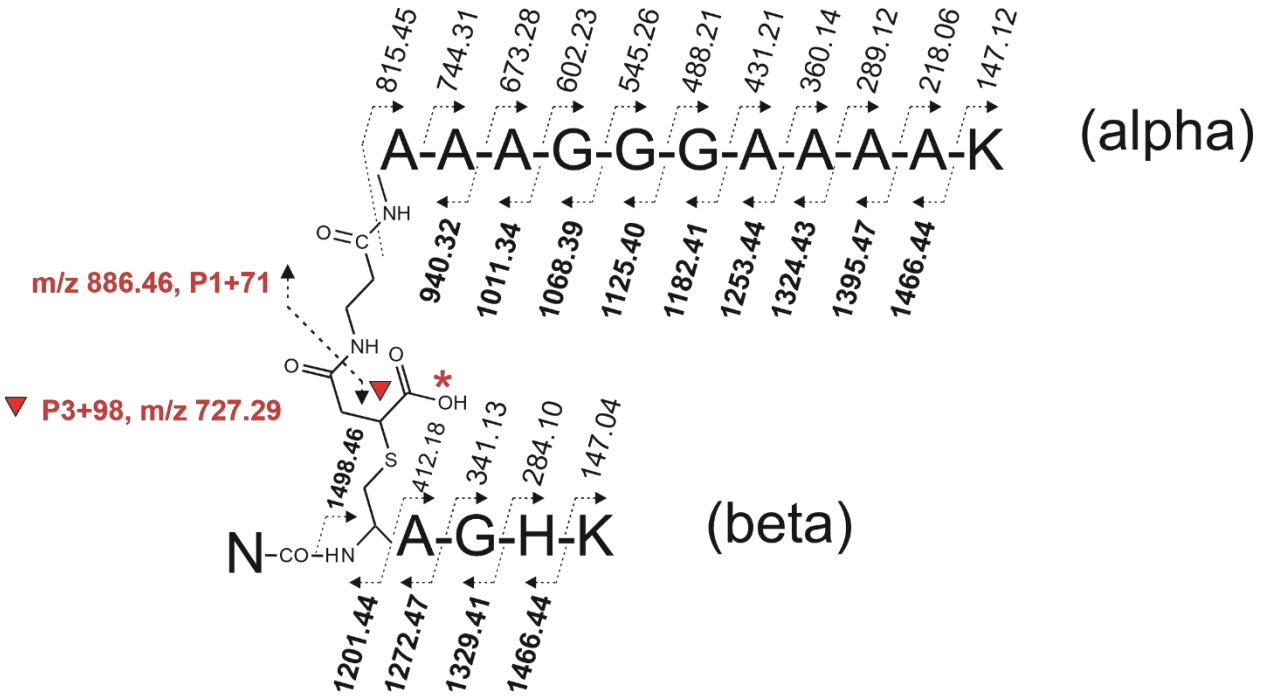
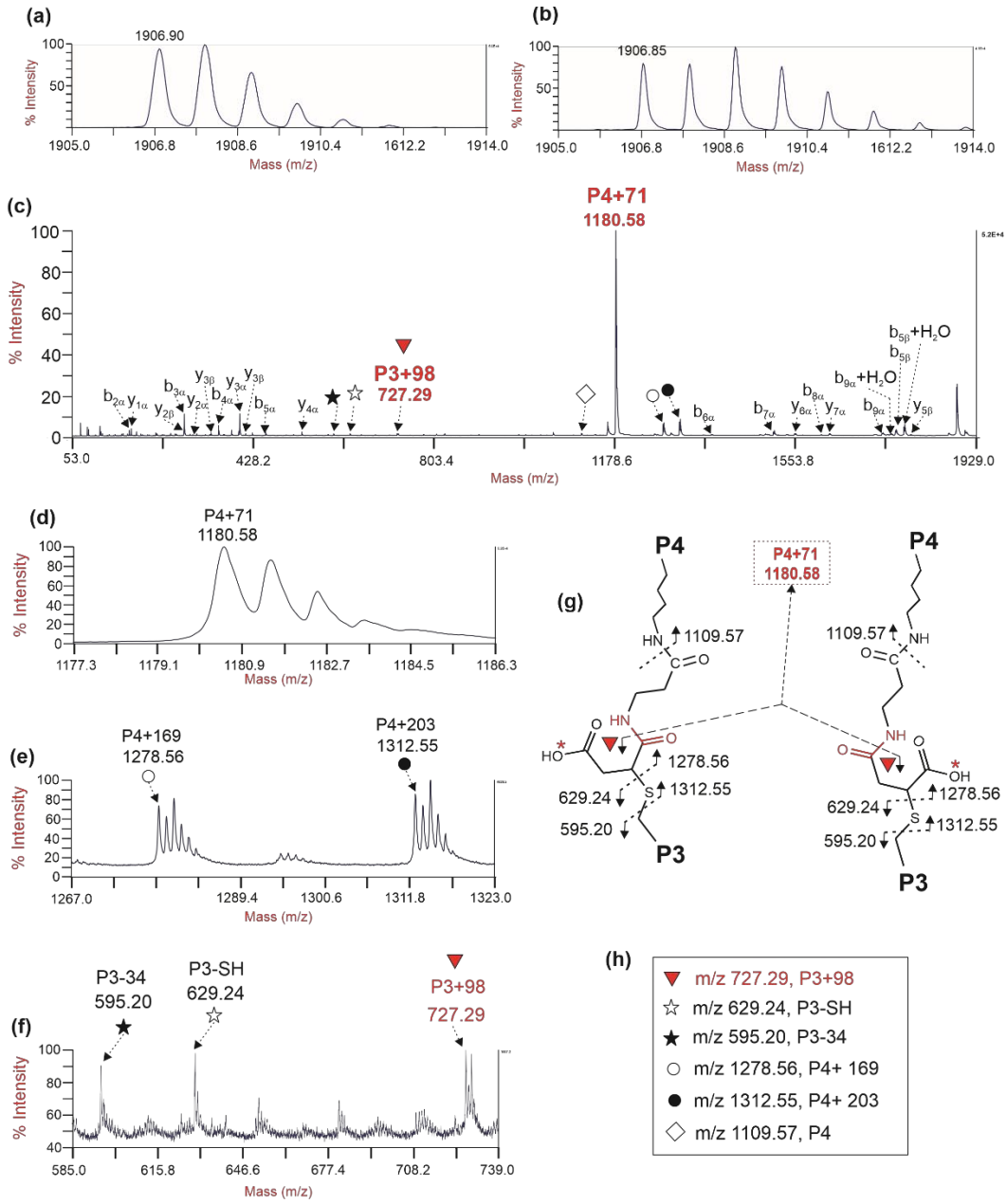


Fig. S12. Isotopic ion distributions of generated by the hydrolysis of the original P3-P4 with the thiosuccinimide linker in a 100 mM Tris/HCl pH=8 buffer prepared with normal water **(a)** and with ^{18}O-labeled water (30 % APE) **(b)**. **(c)** MALDI-MS/MS of the $(P3\text{-}P4)_H$ partially labeled with ^{18}O at the new carboxylic group generated upon hydrolysis of the succinimide ring. The expanded regions (*m/z* 1177.3-1186.3), (*m/z* 1267-1323) and (*m/z* 585-739) of the MALDI-MS/MS spectrum in **(c)** are shown in **(d)**, **(e)** and **(f)**, respectively. **(g)** An schematic representation of the structures of the two positional isomers of $(P3\text{-}P4)_H$ indicate the linker fragment ions (P4+71 and P3+98) as well as the thioether fragment ions (P4+169, P4+203, P3-SH and P3-34). **(h)** Legend of the linker fragment ions observed in **(c)**. **(i)** Summary of the backbone fragment ions (b_n and y_n) observed in the MALDI-MS/MS spectrum of the partially ^{18}O-labeled $(P3\text{-}P4)_H$ at the new carboxyl group generated upon hydrolysis of five-membered ring in a buffer prepared with 30 % (v/v) of $H_2{}^{18}O$. Backbone fragment ions, either the b_n and y_n, and linker fragment ions highlighted in bold characters showed in the MALDI-MS/MS an isotopic ion distribution that evidenced the incorporation of ^{18}O in their structures. A red asterisk indicates in the structure of $(P3\text{-}P4)_H$ the position where the ^{18}O was incorporated. In $(P3\text{-}P4)_H$ the alpha and beta peptides are P4 and P3, respectively. To simplify the figure the structure of only one of the two positional isomers of the hydrolyzed thiosuccinimide linker was represented.



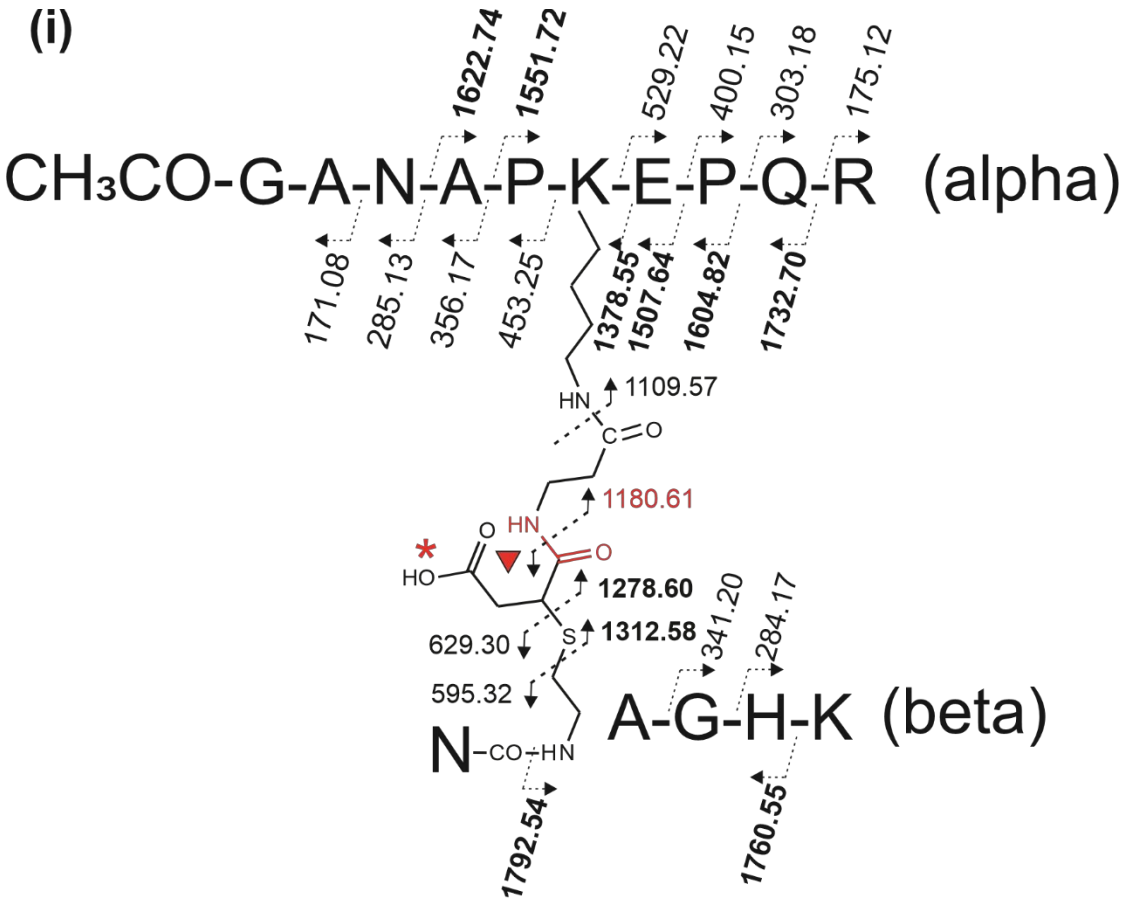
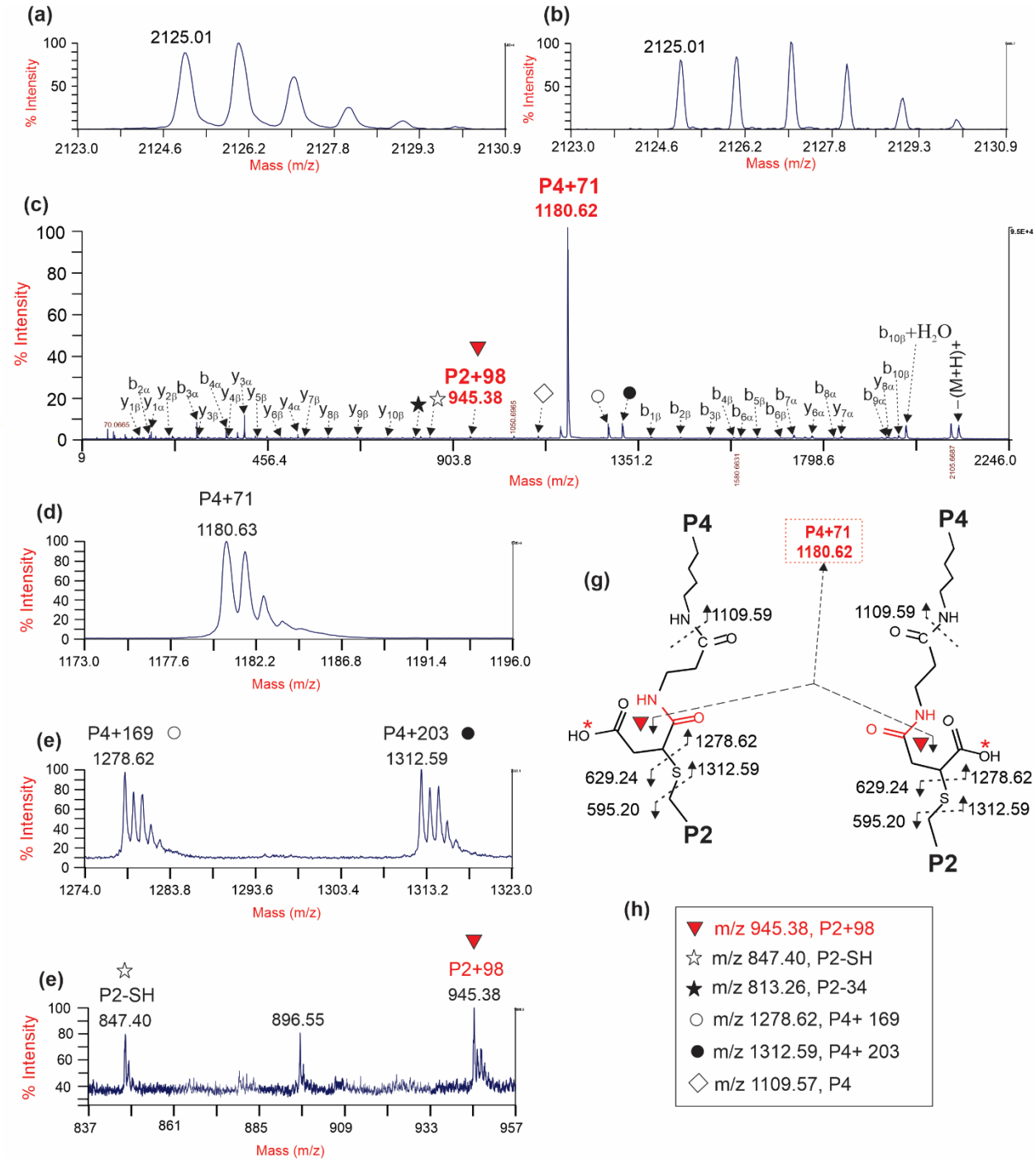


Fig. S13. Isotopic ion distributions of $(P2\text{-}P4)_H$ generated by the hydrolysis of the original P3-P4 with the thiosuccinimide linker in a 100 mM Tris/HCl pH=8 buffer prepared with normal water **(a)** and with ^{18}O-labeled water (30 % APE) **(b)**. **(c)** MALDI-MS/MS of the $(P2\text{-}P4)_H$ partially labeled with ^{18}O at the new carboxylic group generated upon hydrolysis of the succinimide ring. The expanded regions (*m/z* 1173-1196), (*m/z* 1274-1323) and (*m/z* 837-957) of the MALDI-MS/MS spectrum in **(c)** are shown in **(d)**, **(e)** and **(f)**, respectively. **(g)** An schematic representation of the structures of the two positional isomers of $(P3\text{-}P4)_H$ indicate the linker fragment ions (P4+71 and P2+98) as well as the thioether fragment ions (P4+169, P4+203, P2-SH and P2-34). **(h)** Legend of the linker fragment ions observed in **(c)**. **(i)** Summary of the backbone fragment ions (b_n and y_n) observed in the MALDI-MS/MS spectrum of the partially ^{18}O-labeled $(P2\text{-}P4)_H$ at the new carboxyl group generated upon hydrolysis of five-membered ring in a buffer prepared with 30 % (v/v) of $H_2{}^{18}O$. Backbone fragment ions, either the b_n and y_n, and linker fragment ions highlighted in bold characters showed in the MALDI-MS/MS an isotopic ion distribution that evidenced the incorporation of ^{18}O in their structures. A red asterisk indicates in the structure of $(P2\text{-}P4)_H$ the position where the ^{18}O was incorporated. The alpha and beta peptides are P4 and P2, respectively. To simplify the figure the structure of only one of the two positional isomers of the hydrolyzed thiosuccinimide linker was represented.



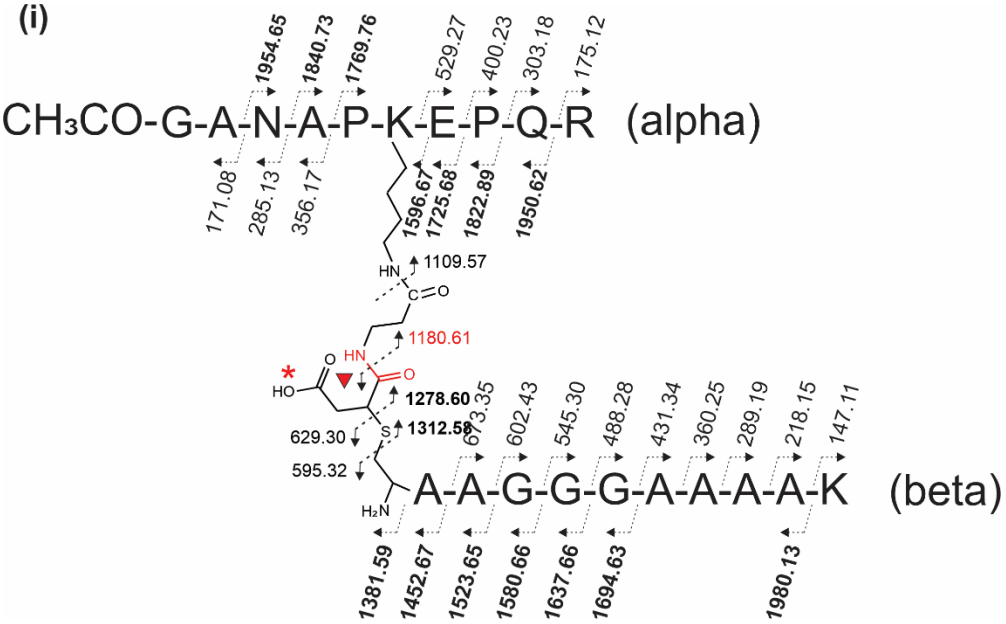
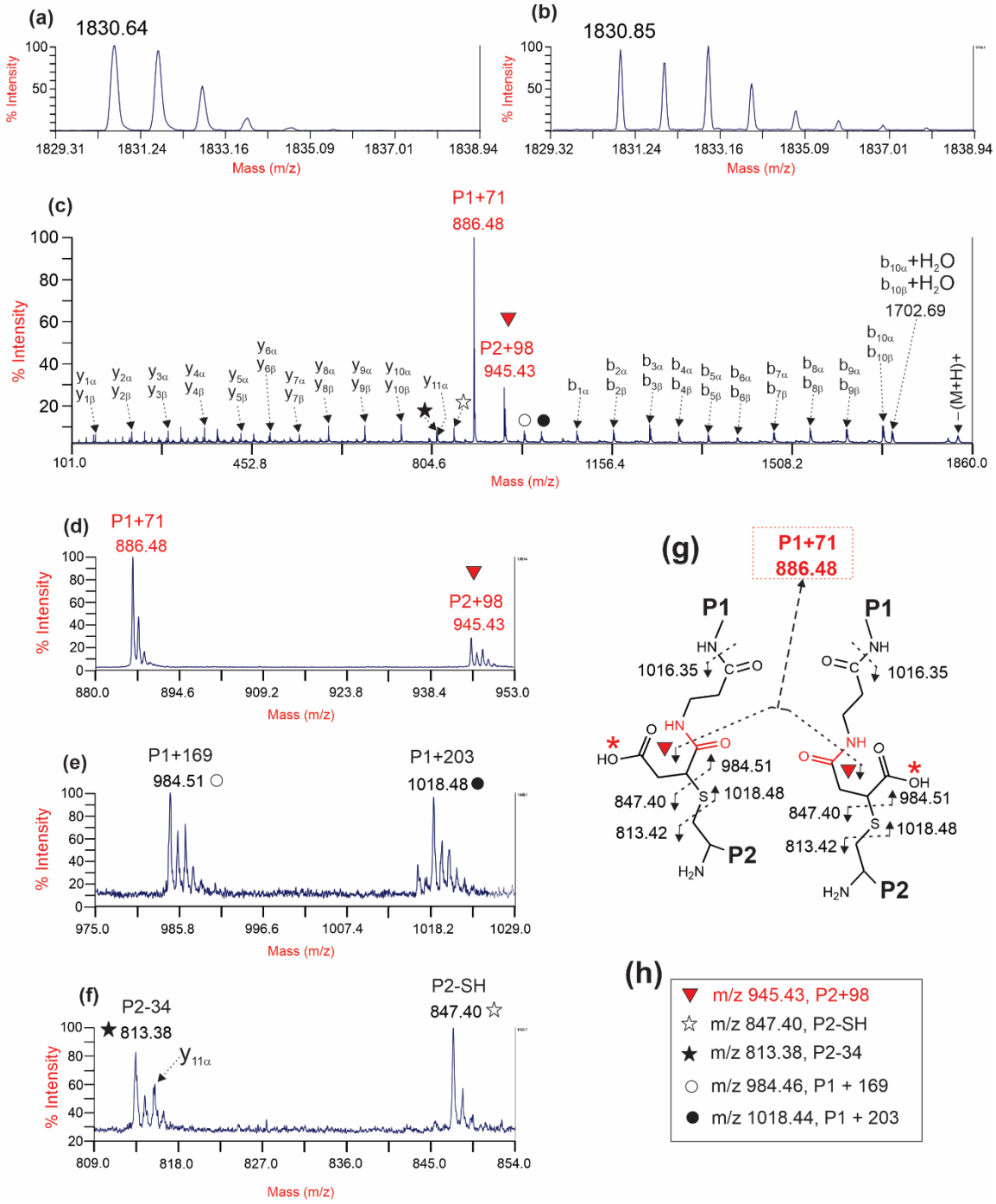


Fig. S14. Isotopic ion distributions of $(P1\text{-}P2)_H$ generated by the hydrolysis of the original P1-P2 with the thiosuccinimide linker in a 100 mM Tris/HCl pH=8 buffer prepared with normal water **(a)** and with ^{18}O-labeled water (30 % APE) **(b)**. **(c)** MALDI-MS/MS of the $(P1\text{-}P2)_H$ partially labeled with ^{18}O at the new carboxylic group generated upon hydrolysis of the succinimide ring. The expanded regions (*m/z* 1173-1196), (*m/z* 1274-1323) and (*m/z* 837-957) of the MALDI-MS/MS spectrum in **(c)** are shown in **(d)**, **(e)** and **(f)**, respectively. **(g)** An schematic representation of the structures of the two positional isomers of $(P1\text{-}P2)_H$ indicate the linker fragment ions (P1+71 and P2+98) as well as the thioether fragment ions (P1+169, P1+203, P2-SH and P2-34). **(h)** Legend of the linker fragment ions observed in **(c)**. **(i)** Summary of the backbone fragment ions (b_n and y_n) observed in the MALDI-MS/MS spectrum of the partially ^{18}O-labeled $(P1\text{-}P2)_H$ at the new carboxyl group generated upon hydrolysis of five-membered ring in a buffer prepared with 30 % (v/v) of $H_2{}^{18}O$. Backbone fragment ions, either the b_n and y_n, and linker fragment ions highlighted in bold characters showed in the MALDI-MS/MS an isotopic ion distribution that evidenced the incorporation of ^{18}O in their structures. A red asterisk indicates in the structure of $(P1\text{-}P2)_H$ the position where the ^{18}O was incorporated. The alpha and beta peptides are P1 and P2, respectively. To simplify the figure the structure of only one of the two positional isomers of the $(P1\text{-}P2)_H$ was represented.



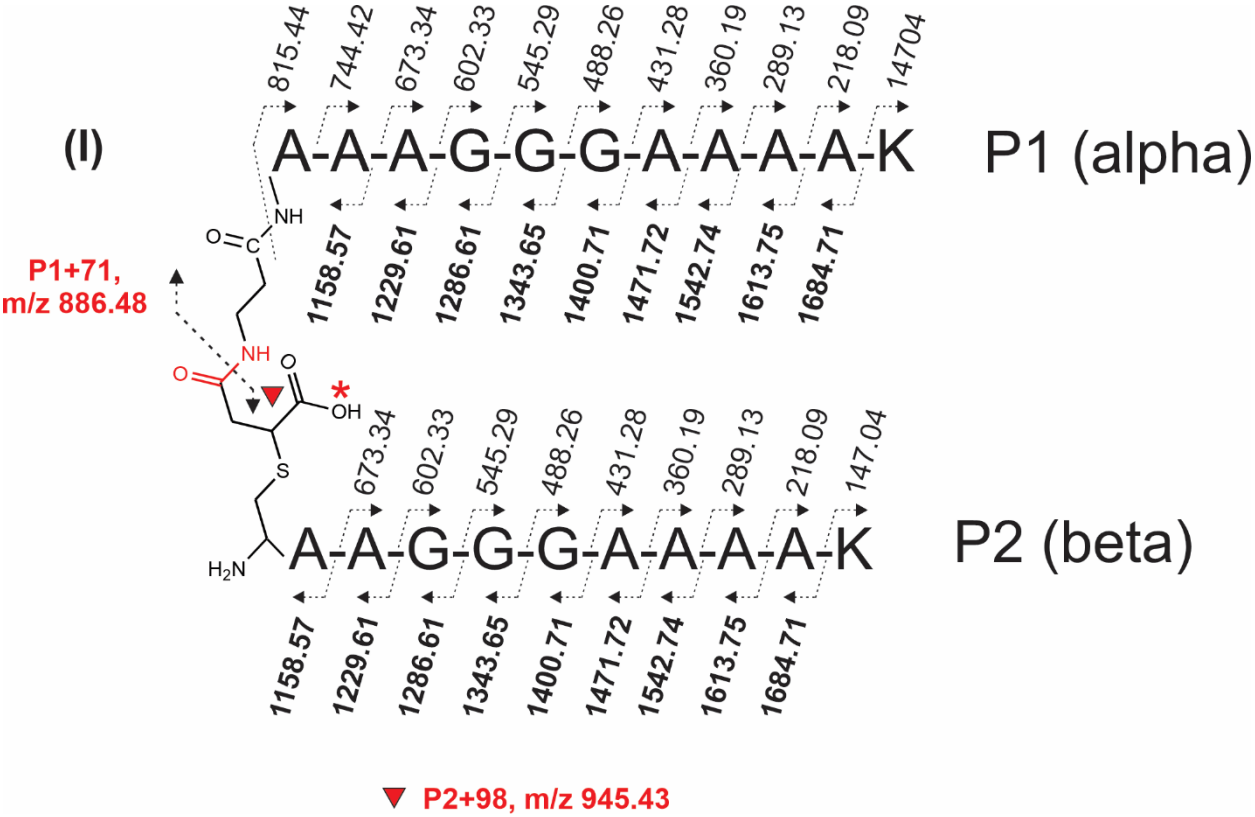
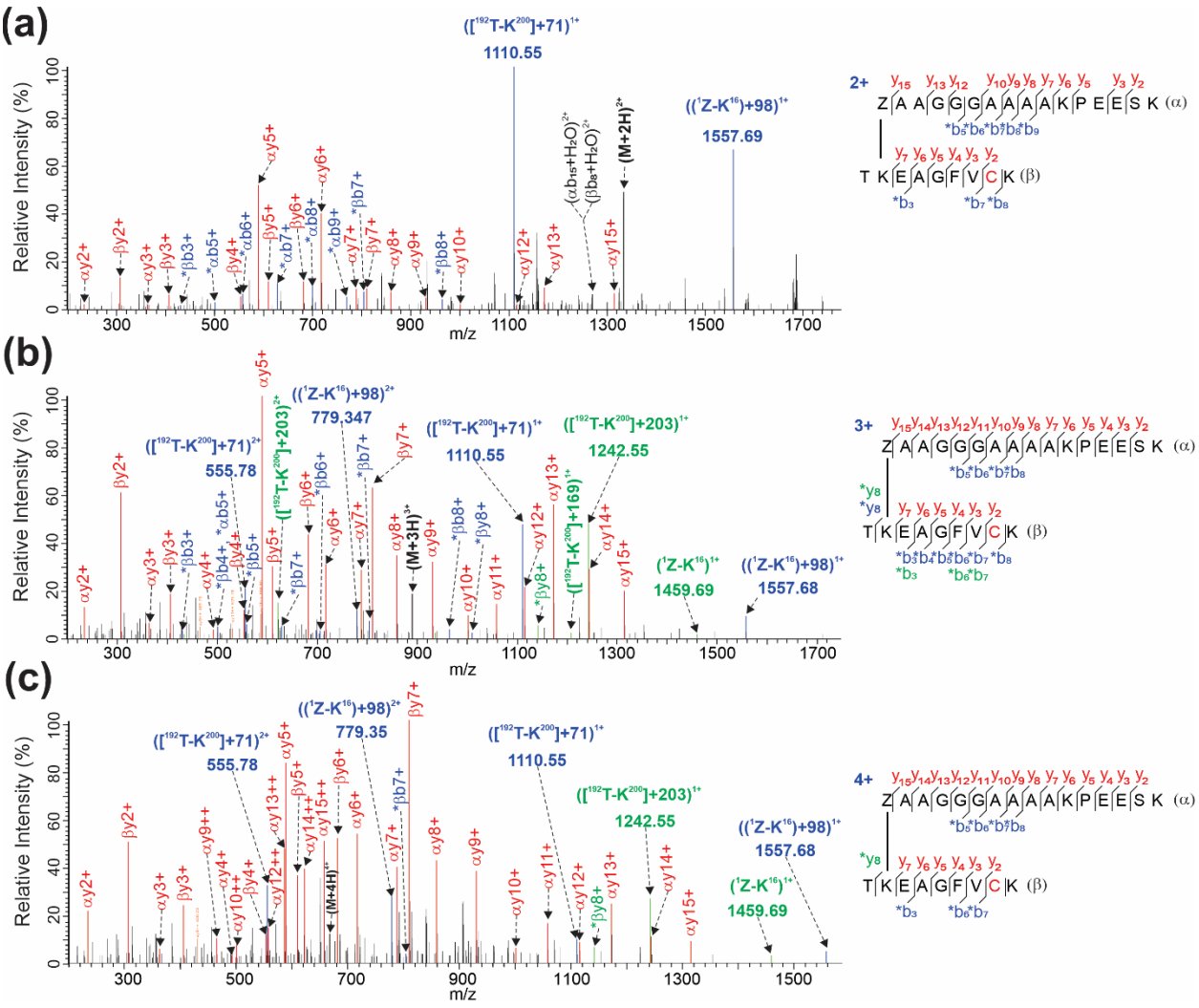


Fig. S15. MS/MS spectra of the type 2 peptide [^{192}T-K^{200}]-(^{1}Z-K^{16}) with different charge states (2+ **(a)**, 3+ **(b)** and 4+ **(c)**) and fragmented with the same normalized collision energy of 30. The amino acid indicated as Z corresponds to the acetylated *N*-terminal cysteine residue. The fragment ions assigned as ([^{192}T-K^{200}]+71)$^{n+}$, and ((Ac-^{1}C-K^{16})+98)$^{n+}$, were originated by a metastable fragmentation of the linker pseudopeptide bond, respectively (see **Figure 3**). Fragments ions *(α/β)y_n and *(α/β)b_n) highlighted in blue were manually assigned and they are derived from either the alpha of beta peptides containing a lysine and cysteine residues with their molecular masses increased by 71 (([^{192}T-K^{200}]+71)$^{+}$) and 98 Da ((^{1}Z-K^{16})+98)$^{+}$, respectively. The sequence coverages obtained from each MS/MS spectrum is shown on the left panels. Fragment ions highlighted in red correspond to (α/β)y_n of the alpha and beta peptides. Isobaric fragment ions αb_{15}+H_2O and βb_8+H_2O correspond to the C-terminal rearrangement of α and β peptides described previously by Gaskell et al. [2]. Other linker fragment ions ([^{192}T-K^{200}]+203)$^{+}$, ([^{192}T-K^{200}]+169)$^{+}$, and (^{1}Z-K^{16})$^{+}$) generated by the thioether bond fragmentation were only observed in **(b)** and **(c)**.



References:

1. 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, http://dx.doi.org/10.1002/rcm.9660.
2. Thorne, G.C., K.D. Ballard, and S.J. Gaskell, *Metastable decomposition of peptide [M+ H]+ ions via rearrangement involving loss of the C-terminal amino acid residue.* Journal of the American Society for Mass Spectrometry, 1990. **1**(3): p. 249-257, http://dx.doi.org/10.1002/(SICI)1096-9888(199602)31:2<150::AID-JMS287>3.0.CO;2-5.
3. Gonzalez, J., et al., *Effect of the position of a basic amino acid on C-terminal rearrangement of protonated peptides upon collision-induced dissociation.* Journal of mass spectrometry, 1996. **31**(2): p. 150-158,
4. Harrison, A.G. and A.B. Young, *Fragmentation reactions of deprotonated peptides containing proline. The proline effect.* J Mass Spectrom, 2005. **40**(9): p. 1173-86, http://dx.doi.org/10.1002/jms.891.
5. Schilling, B., et al., *MS2Assign, automated assignment and nomenclature of tandem mass spectra of chemically crosslinked peptides.* Journal of the American Society for Mass Spectrometry, 2003. **14**(8): p. 834-850, http://dx.doi.org/10.1016/S1044-0305(03)00327-1.

Supplemental Tables

Title: *N*-propionyl thiosuccinimide linker upon hydrolysis becomes into a cleavable crosslinker by a gas phase metastable fragmentation and makes more reliable the conjugation sites identification

Auhtors: Pablo Enmanuel Ramos[1,*], Satomy Pousa[2,*], Paulo Carvalho[4], Rodrigo Soares Caldeira Brant[4], Michel Batista[4], Hironobu Hojo[3], Alina Rodríguez Mallón[5], Vladimir Besada[2], Toshifumi Takao[6] and Luis Javier González [2,¥].

Affiliations:

[1]Bioinformatics Department, and [2] Mass Spectrometry Laboratory. Proteomics Department, and [5]Animal Biotechnology Department. Center for Genetic Engineering and Biotechnology. Avenida 31 e/ 158 y 190. Cubanacán. Playa. PO. Box 6162. Havana. Cuba.
[3]Laboratory for Protein Synthesis and [6]Functional Proteomics and Protein Profiling, Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan.
[4]Carlos Chagas Institute/FioCruz Rua Professor Algacyr Munhoz Mader, 3775 CIC 81350-010, Curitiba, Brasil.

(*) Authors with equal contributions
(¥) Corresponding author

Luis Javier González PhD
Mass Spectrometry Laboratory, Department of Proteomics,
Center for Genetic Engineering and Biotechnology, Avenida 31 e/ 158 y 190,
Cubanacán, Playa, PO Box. 6162. La Habana, Cuba
Email: luis.javier@cigb.edu.cu

Index:

Table S1. Summary of the results obtained in the MALDI-MS analysis of the synthetic peptides analyzed in this study

#	Code [1)]	Amino acid sequences [2)]	$(M+H)^+$ experimental	$(M+H)^+$ expected
1	P1	Mal-**A**AAGGGAAAAK	966.36	966.46
2	P2	**C**AAGGGAAAAK	847.29	847.41
3	P3	N**C**AGHK	629.16	629.28
4	P4	Ac-GANAP**K***EPQR	1260.57	1260.59
5	(P1-P2)	(**A**AAGGGAAAAK)-(**C**AAGGGAAAAK)	1812.88	1812.87
6	$(P1-P2)_T$		1812.61	1812.87
7	$(P1-P2)_H$		1830.64	1830.88
8	(P1-P3)	(**A**AAGGGAAAAK)-(N**C**AGHK)	1594.64	1594.74
9	$(P1-P3)_H$		1612.30	1612.75
10	(P2-P4)	(**C**AAGGGAAAAK)-(Ac-GANAP**K**EPQR)	2106.79	2106.99
11	$(P2-P4)_T$		2106.52	2106.99
12	$(P2-P4)_H$		2124.80	2125.01
13	(P3-P4)	(N**C**AGHK)-(Ac-GANAP**K**EPQR)	1888.42	1888.87
14	$(P3-P4)_H$		1906.46	1906.88

Table S2. Summary of the type 2 peptides derived from the trypsin digestion of the conjugate Bm86-Ac-Cys[1]pP0 conjugate contained in an inclusion list to be analyzed by MS/MS analysis.

#	*Peptide Sequence* [a)]	*m/z exp*	*z*	*m/z theor*
1	[[70]ASCVCEASDDLTLQC**K**IK[87]] - ([1]**Z**AAGGGAAAAKPEESK[16])	1242.5640	3	1242.5641
2	[[70]ASCVCEASDDLTLQC**K**IK[87]] - ([1]**Z**AAGGGAAAAKPEESK[16])	932.1750	4	932.175
3	([1]**Z**AAGGGAAAAKPEESK[16]) - [[579]AVC**K**EK[584]]	788.0380	3	788.0392
4	([1]**Z**AAGGGAAAAKPEESK[16]) - [[579]AVC**K**EK[584]]	591.2804	4	591.2813
5	([1]**Z**AAGGGAAAAKPEESK[16]) - [[599]D**K**DPDPGK[606]]	833.7142	3	833.7155
6	([1]**Z**AAGGGAAAAKPEESK[16]) - [[599]D**K**DPDPGK[606]]	625.5376	4	625.5386
7	([1]**Z**AAGGGAAAAKPEESK[16]) - [[142]DLCE**K**NLLQR[151]]	972.7989	3	972.7996
8	([1]**Z**AAGGGAAAAKPEESK[16]) - [[142]DLCE**K**NLLQR[151]]	729.8511	4	729.8517
9	([1]**Z**AAGGGAAAAKPEESK[16]) - [[425]DQEAAY**K**GQNK[435]]	960.4416	3	960.4424
10	([1]**Z**AAGGGAAAAKPEESK[16]) - [[425]DQEAAY**K**GQNK[435]]	720.5832	4	720.5838
11	([1]**Z**AAGGGAAAAKPEESK[16]) - [[194]EAGFVC**K**HGCR[204]]	983.4398	3	983.4415
12	([1]**Z**AAGGGAAAAKPEESK[16]) - [[194]EAGFVC**K**HGCR[204]]	737.8319	4	737.8331
13	([1]**Z**AAGGGAAAAKPEESK[16]) - [[194]EAGFVC**K**HGCR[204]]	590.4670	5	590.468
14	([1]**Z**AAGGGAAAAKPEESK[16]) - [[528]**K**LECVYK[534]]	856.4078	3	856.4093
15	([1]**Z**AAGGGAAAAKPEESK[16]) - [[528]**K**LECVYK[534]]	642.5579	4	642.5589
16	([1]**Z**AAGGGAAAAKPEESK[16]) - [[403]LLI**K**K[407]]	748.0630	3	748.0637
17	([1]**Z**AAGGGAAAAKPEESK[16]) - [[403]LLI**K**K[407]]	561.2992	4	561.2997
18	([1]**Z**AAGGGAAAAKPEESK[16]) - [[585]SEATTAATTTT**K**AK[598]]	1003.8138	3	1003.8K[16]
19	([1]**Z**AAGGGAAAAKPEESK[16]) - [[585]SEATTAATTTT**K**AK[598]]	753.1123	4	753.114
20	([1]**Z**AAGGGAAAAKPEESK[16]) - [[192]T**K**EAGFVCK[200]]	1334.1221	2	1334.1237
21	([1]**Z**AAGGGAAAAKPEESK[16]) - [[192]T**K**EAGFVCK[200]]	889.7507	3	889.7517
22	([1]**Z**AAGGGAAAAKPEESK[16]) - [[192]T**K**EAGFVCK[200]]	667.5650	4	667.5657
23	([1]**Z**AAGGGAAAAKPEESK[16]) - [[583]E**K**SEATTAATTTTK[596]]	1023.1485	3	1023.1511
24	[1]**Z**AAGGGAAAAKPEESK[16]) - [[583]E**K**SEATTAATTTTK[596]]	767.6133	4	767.6153
25	([1]**Z**AAGGGAAAAKPEESK[16]) - [[520]EIQECQD**K**K[528]]	935.7592	3	935.7609
26	([1]**Z**AAGGGAAAAKPEESK[16]) - [[520]EIQECQD**K**K[528]]	702.0714	4	702.0726
27	[[407]**K**NSATEIEEENLCDSLLK[423]] - ([1]**Z**AAGGGAAAAKPEESK[16])	1240.9118	3	1240.9126
28	[[407]**K**NSATEIEEENLCDSLLK[423]] - ([1]**Z**AAGGGAAAAKPEESK[16])	930.9358	4	930.9364
29	([1]**Z**AAGGGAAAAKPEESK[16]) - [[597]A**K**DKDPDPGK[606]]	900.0915	3	900.0929
30	([1]**Z**AAGGGAAAAKPEESK[16]) - [[597]A**K**DKDPDPGK[606]]	675.3206	4	675.32K[16]
31	([1]**Z**AAGGGAAAAKPEESK[16]) - [[597]A**K**DKDPDPGK[606]]	540.4580	5	540.4589
32	([1]**Z**AAGGGAAAAKPEESK[16]) - [[572]CVIENG**K**AVCK[582]]	969.1201	3	969.1217
33	([1]**Z**AAGGGAAAAKPEESK[16]) - [[572]CVIENG**K**AVCK[582]]	727.0920	4	727.0932
34	([1]**Z**AAGGGAAAAKPEESK[16]) - [[45]QCEY**K**DTCK[53]]	953.7453	3	953.7464
35	[1]**Z**AAGGGAAAAKPEESK[16]) - [[45]QCEY**K**DTCK[53]]	715.5609	4	715.5618
36	([1]**Z**AAGGGAAAAKPEESK[16]) - [[514]TTCNP**K**EIQECQDK[527]]	1126.8335	3	1126.8348
37	([1]**Z**AAGGGAAAAKPEESK[16]) - [[514]TTCNP**K**EIQECQDK[527]]	845.3771	4	845.3781
38	([1]**Z**AAGGGAAAAKPEESK[16]) - [[370]ARLIAE**K**PLSK[380]]	951.8294	3	951.8312
39	([1]**Z**AAGGGAAAAKPEESK[16]) - [[370]ARLIAE**K**PLSK[380]]	714.1240	4	714.1254
40	([1]**Z**AAGGGAAAAKPEESK[16]) - [[370]ARLIAE**K**PLSK[380]]	571.5008	5	571.5019

41	(^{1}**Z**AAGGGAAAAKPEESK^{16}) - [^{194}EAGFVC**K**HGCRSTGK^{208}]	1107.8390	3	1107.8402
42	(^{1}**Z**AAGGGAAAAKPEESK^{16}) - [^{194}EAGFVC**K**HGCRSTGK^{20}]^{8}	831.1312	4	831.1321
43	(^{1}**Z**AAGGGAAAAKPEESK^{16}) - [^{194}EAGFVC**K**HGCRSTGK^{208}]	665.1065	5	665.1073
44	(^{1}**Z**AAGGGAAAAKPEESK^{16}) - [^{194}EAGFVC**K**HGCRSTGK^{208}]	554.4234	6	554.424
45	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{45}QCEY**K**DTCK^{53}]	996.4416	3	996.4447
46	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{45}QCEY**K**DTCK^{53}]	747.5831	4	747.5855
47	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{45}QCEY**K**DTCK^{53}]	598.2681	5	598.27
48	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{192}T**K**EAGFVCK^{200}]	932.4484	3	932.45
49	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{192}T**K**EAGFVCK^{200}]	699.5882	4	699.5895
50	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{192}T**K**EAGFVCK^{200}]	559.8721	5	559.8732
51	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{585}SEATTAATTTT**K**AK^{598}]	785.1375	4	785.1377
52	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{585}SEATTAATTTT**K**AK^{598}]	628.3115	5	628.3117
53	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{403}LLI**K**K^{407}]	790.7608	3	790.762
54	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{403}LLI**K**K^{407}]	593.3225	4	593.3234
55	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{528}**K**LECVYK^{534}]	899.1056	3	899.1076
56	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{528}**K**LECVYK^{534}]	674.5812	4	674.5826
57	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{528}**K**LECVYK^{534}]	539.8665	5	539.8677
58	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{583}E**K**SEATTAATTTTK^{596}]	799.6374	4	799.639
59	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{583}E**K**SEATTAATTTTK^{596}]	639.9115	5	639.9128
60	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{520}EIQECQD**K**K^{528}]	978.4589	3	978.4594
61	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{520}EIQECQD**K**K^{528}]	734.0962	4	734.0964
62	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{520}EIQECQD**K**K^{528}]	587.4785	5	587.4787
63	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{142}DLCE**K**NLLQR^{151}]	761.8746	4	761.8754
64	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{142}DLCE**K**NLLQR^{151}]	609.7012	5	609.7019
65	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{598}D**K**DPDPGK^{606}]	657.5613	4	657.5623
66	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{598}D**K**DPDPGK^{606}]	526.2506	5	526.2515
67	[^{70}ASCVCEASDDLTLQC**K**IK^{87}] - (^{1}**Z**AAGGGAAAAKPEESKK^{17})	964.1951	4	964.1988
68	[^{70}ASCVCEASDDLTLQC**K**IK^{87}] - (^{1}**Z**AAGGGAAAAKPEESKK^{17})	771.5576	5	771.5606
69	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{528}KLECVY**K**NHK^{537}]	615.7060	5	615.707
70	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{514}TTCNP**K**EIQECQDK^{527}]	702.1213	5	702.123
71	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{194}EAGFVC**K**HGCRSTGK^{208}]	690.7254	5	690.7263
72	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{194}EAGFVC**K**HGCRSTGK^{208}]	575.7725	6	575.7732
73	(^{1}**Z**AAGGGAAAAKPEESKK^{17}) - [^{194}EAGFVC**K**HGCRSTGK^{208}]	493.6632	7	493.6638

a) type 2 peptides crosslinked by hydrolyzed thiosuccinimide linker. Z means an acetyl group blocking the N-terminal end of the cysteinyl peptide Cys^{1}pP0. Amino acid sequences inside brackets and square brackets correspond to tryptic peptides derived from the Ac-Cys^{1}pP0 peptide and the Bm86 carrier protein, respectively. Amino acids highlighted in bold and underlined are crosslinked by a hydrolyzed thiosuccinimide linker.

Table S3. Assignment of the fragment ions derived from a triply-charged type 2 peptide [^{192}T-K^{200}]-(^{1}Z-K^{16})$^{3+}$ detected at *m/z* 889.7507 fragmented by stepped HCD (sHCD) at normalized collision energies of 27, 30 and of 33.

#	*m/z* experimental	*m/z* expected	Z	Assignment [a)]
1	234.144	234.1453	1	αy_2
2	307.143	307.1439	1	βy_2
3	363.185	363.1879	1	αy_3
4	406.211	406.2123	1	βy_3
5	430.230	430.2302	1	$*\beta b_3$
6	465.747	465.7487	2	αy_9^{2+}
7	501.267	501.2673	2	αy_{10}^{2+}
8	529.778	529.778	2	αy_{11}^{2+}
9	553.277	553.2807	1	βy_4
10	555.784	555.7848	2	$([^{192}T\text{-}K^{200}]+71)^{2+}$
11	557.165	557.1666	1	$*\alpha b_6$
12	558.288	558.2888	2	αy_{12}^{2+}
13	586.797	586.7995	2	αy_{13}^{2+}
14	589.282	589.2833	1	αy_5
15	610.302	610.3022	1	βy_5
16	622.306	622.3181	2	αy_{14}^{2+}
17	681.337	681.3393	1	βy_6
18	628.204	628.2037	1	$*\alpha b_7$
19	657.834	657.8366	2	αy_{15}^{2+}
20	699.237	699.2408	1	$*\alpha b_8$
21	705.346	705.3572	1	$*\beta b_6$
22	717.376	717.3782	1	αy_6
23	779.345	779.3467	2	$(([^{1}Z\text{-}K^{16})+98)^{2+}$
24	788.414	788.4153	1	αy_7
25	804.425	804.4256	1	$*\beta b_7$
26	810.380	810.3819	1	βy_7
27	859.449	859.4525	1	αy_8
28	930.486	930.4896	1	αy_9
29	964.448	964.4562	1	$*\beta b_8$
30	1001.520	1001.5267	1	αy_{10}
31	1110.557	1110.562	1	$([^{192}T\text{-}K^{200}]+71)^{1+}$
32	1058.545	1058.5267	1	αy_{11}
33	1115.566	1115.5696	1	αy_{12}
34	1172.586	1172.5911	1	αy_{13}
35	1242.546	1242.5497	1	$([^{192}T\text{-}K^{200}]+203)^{1+}$
36	1243.590	1243.6282	1	αy_{14}
37	1314.660	1314.6653	1	αy_{15}
38	1557.680	1557.6855	1	$(([^{1}Z\text{-}K^{16})+98)^{1+}$

a) Fragment ions highlighted in blue were assigned manually by considering the cleavability of the linker fragmentation and their derived fragment ions. Z corresponds to a *N*-terminal acetylated cysteine residue. In the structure of the type 2 peptide, the α and β peptides are (^{1}Z-K^{16}) and ([^{192}T-K^{200}], respectively. The nomenclature of the backbone fragment ions is in agreement with the reported by Schilling *et al* [1]. The linker fragment ions assigned as (^{1}Z-K^{16})+98)$^{1+}$ and ([^{192}T-K^{200}]+71)$^{1+}$ correspond to the α peptide and β peptides, containing the formerly crosslinked Lys and Cys modified by structures that increase their molecular mass by 98 Da and 71 Da, respectively. The linker fragment ion assigned as (^{1}Z-K^{16})+98)$^{1+}$contain the Cys^{1} S-alkylated by a succinic anhydride residue. The linker fragment ion assigned as ([^{192}T-K^{200}]+71)$^{1+}$ contain the Lys^{193} modified at its epsilon amino group by a residue of β-alanine.The linker fragment ion assigned as ([^{192}T-K^{200}]+71)$^{1+}$ contain the Lys^{193} modified at its epsilon amino group by a residue of β-alanine. The linker fragment ion assigned as ([^{192}T-K^{200}]+71)$^{1+}$ contain the Lys^{193} modified at its epsilon amino group by a residue of β-alanine. Fragment ions assigned as *(α/β)b_n or *(α/β)b_n correspond to the backbone fragment ions b_n or y_n derived from the linearized α and β peptides containing a modified Lys (+71 Da) or a Cys(+98 Da) residues as described above.

Table S4. Assignment of the fragment ions derived from the type 2 peptide [^{192}T-K^{200}]-(^{1}Z-K^{16})$^{4+}$ detected at *m/z* 667.5650 fragmented by stepped HCD (sHCD) at normalized collision energies of 27, 30 and of 33.

#	*m/z* experimental	*m/z* expected	Z	Assignment [a)]
	234.144	234.1453	1	αy_2
	307.143	307.1439	1	βy_2
	363.186	363.1879	1	αy_3
	394.711	394.7116	2	αy_7^{2+}
	406.211	406.2123	1	βy_3
	430.229	430.2302	2	αy_8^{2+}
	465.748	465.7487	2	αy_9^{2+}
	492.229	492.2305	1	αy_4
	501.265	501.2673	2	αy_{10}^{2+}
	529.785	529.778	2	αy_{11}^{2+}
	553.276	553.2807	1	βy_4
	555.784	555.7848	2	([^{192}T-K^{200}]+71)$^{2+}$
	558.288	558.2888	2	αy_{12}^{2+}
	586.797	586.7995	2	αy_{13}^{2+}
	589.282	589.2833	1	αy_5
	610.300	610.3022	1	βy_5
	621.779	621.7789	2	([^{192}T-K^{200}]+203)$^{2+}$
	622.314	622.3181	2	αy_{14}^{2+}
	657.835	657.8366	2	αy_{15}^{2+}
	681.337	681.3393	1	βy_6
	705.359	705.3572	1	$*\beta b_6$
	717.376	717.3782	1	αy_6
	779.345	779.3467	2	((^{1}Z-K^{16})+98)$^{2+}$
	788.413	788.4153	1	αy_7
	804.427	804.4256	1	$*\beta b_7$
	810.379	810.3819	1	βy_7
	859.448	859.4525	1	αy_8
	930.484	930.4896	1	αy_9
	964.440	964.4562	1	$*\beta b_8$
	1001.524	1001.5267	1	αy_{10}
	1058.545	1058.5267	1	αy_{11}
	1110.553	1110.562	1	([^{192}T-K^{200}]+71)$^{1+}$
	1115.561	1115.5696	1	αy_{12}
	1172.584	1172.5911	1	αy_{13}
	1242.547	1242.5497	1	([^{192}T-K^{200}]+203)$^{1+}$
	1243.562	1243.6282	1	αy_{14}
	1314.644	1314.6653	1	αy_{15}
	1557.710	1557.6855	1	((^{1}Z-K^{16})+98)$^{1+}$

a) Fragment ions highlighted in blue were assigned manually by considering the cleavability of the linker fragmentation and their derived fragment ions. Z corresponds to a *N*-terminal acetylated cysteine residue. In the structure of the type 2 peptide, the α and β peptides are (^{1}Z-K^{16}) and ([^{192}T-K^{200}], respectively. The nomenclature of the backbone fragment ions is in agreement with the reported by Schilling *et al* [1]. The linker fragment ions assigned as (^{1}Z-K^{16})+98)$^{1+}$ and ([^{192}T-K^{200}]+71)$^{1+}$ correspond to the α peptide and β peptides, containing the formerly crosslinked Lys and Cys modified by structures that increase their molecular mass by 98 Da and 71 Da, respectively. The linker fragment ion assigned as (^{1}Z-K^{16})+98)$^{1+}$contain the Cys^{1} S-alkylated by a succinic anhydride residue. The linker fragment ion assigned as ([^{192}T-K^{200}]+71)$^{1+}$ contain the Lys^{193} modified at its epsilon amino group by a residue of β-alanine. Fragment ions assigned as *(α/β)b$_n$ or *(α/β)b$_n$ correspond to the backbone fragment ions b$_n$ or y$_n$ derived from the linearized α and β peptides containing a Lys (+71 Da) or a Cys(+98 Da) modified residues as described above. The structure of the linker fragment ion ([^{192}T-K^{200}]+203)$^{n+}$ is generated by the thioether bond fragmentation as depicted in **Fig. S10.**

Table S5. Assignment of the fragment ions derived from a doubly-charged type 2 peptide [^{192}T-K^{200}]-(^{1}Z-K^{16}) detected at *m/z* 1334.1237 fragmented by HCD at a normalized collision energy of 30.

#	*m/z* experimental	*m/z* expected	Z	Assignment [a)]
1	234.14	234.1453	1	αy_2
2	307.142	307.1439	1	βy_2
3	363.19	363.1879	1	αy_3
4	406.211	406.2123	1	βy_3
5	430.23	430.2302	1	$*\beta b_3$
6	500.146	500.146	1	$*\alpha b_5$
7	553.272	553.2807	1	βy_4
8	557.165	557.1666	1	$*\alpha b_6$
9	589.282	589.2833	1	αy_5
10	610.302	610.3022	1	βy_5
11	681.340	681.3393	1	βy_6
12	628.203	628.2037	1	$*\alpha b_7$
13	699.241	699.2408	1	$*\alpha b_8$
14	705.359	705.3572	1	$*\beta b_6$
15	717.337	717.3782	1	αy_6
16	770.278	770.2780	1	$*\alpha b_9$
17	788.416	788.4153	1	αy_7
18	804.425	804.4256	1	$*\beta b_7$
19	810.382	810.3819	1	βy_7
20	859.449	859.4525	1	αy_8
21	930.484	930.4896	1	αy_9
22	964.448	964.4562	1	$*\beta b_8$
23	1001.508	1001.5267	1	αy_{10}
24	1110.561	1110.562	1	$([^{192}T\text{-}K^{200}]+71)^{1+}$
25	1115.576	1115.5696	1	αy_{12}
26	1172.592	1172.5911	1	αy_{13}
27	1314.657	1314.6653	1	αy_{15}
28	1557.688	1557.6855	1	$(([^{1}Z\text{-}K^{16})+98)^{1+}$

a) Fragment ions highlighted in blue were assigned manually by considering the cleavability of the linker fragmentation and their derived fragment ions. Z corresponds to a *N*-terminal acetylated cysteine residue. In the structure of the type 2 peptide, the α and β peptides are (^{1}Z-K^{16}) and ([^{192}T-K^{200}], respectively. The nomenclature of the backbone fragment ions is in agreement with the reported by Schilling *et al* [1]. The linker fragment ions assigned as (^{1}Z-K^{16})+98)$^{1+}$ and ([^{192}T-K^{200}]+71)$^{1+}$ correspond to the α peptide and β peptides, containing the formerly crosslinked Lys and Cys modified by structures that increase their molecular mass by 98 Da and 71 Da, respectively. The linker fragment ion assigned as (^{1}Z-K^{16})+98)$^{1+}$contain the Cys^{1} S-alkylated by a succinic anhydride residue. The linker fragment ion assigned as ([^{192}T-K^{200}]+71)$^{1+}$ contain the

Lys^{193} modified at its epsilon amino group by a residue of β-alanine.The linker fragment ion assigned as ([^{192}T-K^{200}]+71)$^{1+}$ contain the Lys^{193} modified at its epsilon amino group by a residue of β-alanine. The linker fragment ion assigned as ([^{192}T-K^{200}]+71)$^{1+}$ contain the Lys^{193} modified at its epsilon amino group by a residue of β-alanine. Fragment ions assigned as *(α/β)b_n or *(α/β)b_n correspond to the backbone fragment ions b_n or y_n derived from the linearized α and β peptides containing a modified Lys (+71 Da) or a Cys(+98 Da) residues as described above.

Table S6. Assignment of the fragment ions derived from a triply-charged type 2 peptide [^{192}T-K^{200}]-(^{1}Z-K^{16}) detected at *m/z* 889.7517 fragmented by HCD at a normalized collision energy of 30.

#	*m/z* experimental	*m/z* expected	Z	Assignment [a)]
	234.144	234.1453	1	αy_2
	307.143	307.1439	1	βy_2
	363.187	363.1879	1	αy_3
	406.212	406.2123	1	βy_3
	430.23	430.2302	1	$*\beta b_3$
	433.175	433.1757	1	$*\beta b_2$
	500.143	500.146	1	$*\alpha b_5$
	501.267	501.2673	1	$*\beta b_4$
	553.276	553.2807	1	βy_4
	555.785	555.7848	2	$([^{192}T\text{-}K^{200}]+71)^{2+}$
	557.165	557.1666	1	$*\alpha b_6$
	562.221	562.2183	1	$*\beta b_3$
	558.289	558.288	1	$*\beta b_5$
	589.283	589.2833	1	αy_5
	610.302	610.3022	1	βy_5
	681.340	681.3393	1	βy_6
	621.780	621.7789	1	$([^{192}T\text{-}K^{200}]+203)^{2+}$
	628.203	628.2037	1	$*\alpha b_7$
	705.359	705.3572	1	$*\beta b_6$
	717.379	717.3782	1	αy_6
	779.347	779.3467	2	$(([^{1}Z\text{-}K^{16})+98)^{2+}$
	788.417	788.4153	1	αy_7
	804.427	804.4256	1	$*\beta b_7$
	810.383	810.3819	1	βy_7
	936.347	837.3453	1	$*\beta b_6$
	859.453	859.4525	1	αy_8
	930.490	930.4896	1	αy_9
	936.422	936.4137	1	$*\beta b_7$
	964.454	964.4562	1	$*\beta b_8$
	1001.522	1001.5267	1	αy_{10}
	1009.500	1009.5140	1	$*\beta y_8$
	1058.546	1058.5481	1	αy_{11}
	1110.562	1110.562	1	$([^{192}T\text{-}K^{200}]+71)^{1+}$
	1115.569	1115.5696	1	αy_{12}
	1141.505	1141.5021	1	$*\beta y_8$
	1172.592	1172.5911	1	αy_{13}
	1242.552	1242.5497	1	$([^{192}T\text{-}K^{200}]+203)^{1+}$
	1243.604	1243.6282	1	αy_{14}
	1314.665	1314.6653	1	αy_{15}
	1459.691	1459.6851	1	$(^{1}Z\text{-}K^{16})^{1+}$
	1557.683	1557.6855	1	$(([^{1}Z\text{-}K^{16})+98)^{1+}$

a) Fragment ions highlighted in blue were assigned manually by considering the cleavability of the linker fragmentation and their derived fragment ions. Z corresponds to a *N*-terminal acetylated cysteine residue. In the structure of the type 2 peptide, the α and β peptides are (^{1}Z-K^{16}) and ([^{192}T-K^{200}], respectively. The nomenclature of the backbone fragment ions is in agreement with the reported by Schilling *et al* [1]. The linker fragment ions assigned as (^{1}Z-K^{16})+98)$^{1+}$ and ([^{192}T-K^{200}]+71)$^{1+}$ correspond to the α peptide and β peptides, containing the formerly crosslinked Lys and Cys modified by structures that increase their molecular mass by 98 Da and 71 Da, respectively. The linker fragment ion assigned as (^{1}Z-K^{16})+98)$^{1+}$contain the Cys^{1} S-alkylated by a succinic anhydride residue. The linker fragment ion assigned as ([^{192}T-K^{200}]+71)$^{1+}$ contain the Lys^{193} modified at its epsilon amino group by a residue of β-alanine.The linker fragment ion assigned as ([^{192}T-K^{200}]+71)$^{1+}$ contain the Lys^{193} modified at its epsilon amino group by a residue of β-alanine. The linker fragment ion assigned as ([^{192}T-K^{200}]+71)$^{1+}$ contain the Lys^{193} modified at its epsilon amino group by a residue of β-alanine. Fragment ions assigned as *(α/β)b$_n$ or *(α/β)b$_n$ correspond to the backbone fragment ions b$_n$ or y$_n$ derived from the linearized α and β peptides containing a modified Lys (+71 Da) or a Cys(+98 Da) residues as described above.

Table S7. Assignment of the fragment ions derived from the type 2 peptide $([^{192}T\text{-}K^{200}]\text{-}(^{1}Z\text{-}K^{16}))^{4+}$ detected at *m/z* 667.5657 fragmented by HCD at a normalized collision energy of 30.

#	*m/z* experimental	*m/z* expected	Z	Assignment [a)]
	234.144	234.1453	1	αy_2
	307.143	307.1439	1	βy_2
	363.186	363.1879	1	αy_3
	406.212	406.2123	1	βy_3
	430.23	430.2302	1	$*\beta b_3$
	465.747	465.7487	2	αy_9^{2+}
	492.225	492.2305	1	αy_4
	501.267	501.2673	2	αy_{10}^{2+}
	553.275	553.2807	1	βy_4
	555.784	555.7848	2	$([^{192}T\text{-}K^{200}]+71)^{2+}$
	557.165	557.1666	1	$*\alpha b_6$
	558.288	558.2888	2	αy_{12}^{2+}
	586.799	586.7995	2	αy_{13}^{2+}
	589.283	589.2833	1	αy_5
	610.302	610.3022	1	βy_5
	621.779	**621.7789**	**1**	$([^{192}T\text{-}K^{200}]+203)^{2+}$
	622.319	622.3181	2	αy_{14}^{2+}
	628.197	628.2037	1	$*\alpha b_7$
	657.837	657.8366	2	αy_{15}^{2+}
	681.339	681.3393	1	βy_6
	717.377	717.3782	1	αy_6
	779.346	779.3467	2	$((^{1}Z\text{-}K^{16})+98)^{2+}$
	788.415	788.4153	1	αy_7
	804.435	804.4256	1	$*\beta b_7$
	810.381	810.3819	1	βy_7
	859.451	859.4525	1	αy_8
	930.487	930.4896	1	αy_9
	1001.511	1001.5267	1	αy_{10}
	1058.558	1058.5481	1	αy_{11}
	1110.557	1110.562	1	$([^{192}T\text{-}K^{200}]+71)^{1+}$
	1115.562	1115.5696	1	αy_{12}
	1141.505	**1141.5021**	**1**	$*\beta y_8$
	1172.594	1172.5911	1	αy_{13}
	1242.552	**1242.5497**	**1**	$([^{192}T\text{-}K^{200}]+203)^{1+}$
	1243.589	1243.6282	1	αy_{14}
	1314.653	1314.6653	1	αy_{15}
	1459.668	**1459.6851**	**1**	$(^{1}Z\text{-}K^{16})^{1+}$
	1557.698	1557.6855	1	$((^{1}Z\text{-}K^{16})+98)^{1+}$

a) Fragment ions highlighted in blue were assigned manually by considering the cleavability of the linker fragmentation and their derived fragment ions. Z corresponds to a *N*-terminal acetylated cysteine residue. In the structure of the type 2 peptide, the α and β peptides are $(^{1}Z\text{-}K^{16})$ and

([^{192}T-K^{200}], respectively. The nomenclature of the backbone fragment ions is in agreement with the reported by Schilling *et al* [1]. The linker fragment ions assigned as (^{1}Z-K^{16})+98)$^{1+}$ and ([^{192}T-K^{200}]+71)$^{1+}$ correspond to the α peptide and β peptides, containing the formerly crosslinked Lys and Cys modified by structures that increase their molecular mass by 98 Da and 71 Da, respectively. The linker fragment ion assigned as (^{1}Z-K^{16})+98)$^{1+}$contain the Cys^{1} S-alkylated by a succinic anhydride residue. The linker fragment ion assigned as ([^{192}T-K^{200}]+71)$^{1+}$ contain the Lys^{193} modified at its epsilon amino group by a residue of β-alanine.The linker fragment ion assigned as ([^{192}T-K^{200}]+71)$^{1+}$ contain the Lys^{193} modified at its epsilon amino group by a residue of β-alanine. The linker fragment ion assigned as ([^{192}T-K^{200}]+71)$^{1+}$ contain the Lys^{193} modified at its epsilon amino group by a residue of β-alanine. Fragment ions assigned as *(α/β)b_n or *(α/β)b_n correspond to the backbone fragment ions b_n or y_n derived from the linearized α and β peptides containing a modified Lys (+71 Da) or a Cys(+98 Da) residues as described above.

Table S#. Assignment of the fragment ions derived from a doubly-charged type 2 peptide [^{192}T-K^{200}]-(^{1}Z-K^{16})$^{2+}$ detected at *m/z* 1334.218 fragmented by stepped HCD (sHCD) at normalized collision energies of 27, 30 and of 33.

#	*m/z* experimental	*m/z* expected	Z	Assignment [a)]
	307.142	307.1439	1	βy_2
	406.211	406.2123	1	βy_3
	430.23	430.2302	1	$*\beta b_3$
	500.146	500.146	1	$*\alpha b_5$
	557.165	557.1666	1	$*\alpha b_6$
	589.281	589.2833	1	αy_5
	610.297	610.3022	1	βy_5
	681.338	681.3393	1	βy_6
	628.203	628.2037	1	$*\alpha b_7$
	699.241	699.2408	1	$*\alpha b_8$
	705.359	705.3572	1	$*\beta b_6$
	717.376	717.3782	1	αy_6
	770.278	770.2780	1	$*\alpha b_9$
	788.412	788.4153	1	αy_7
	804.425	804.4256	1	$*\beta b_7$
	810.380	810.3819	1	βy_7
	859.451	859.4525	1	αy_8
	930.482	930.4896	1	αy_9
	964.448	964.4562	1	$*\beta b_8$
	1001.508	1001.5267	1	αy_{10}
	1110.558	1110.562	1	([^{192}T-K^{200}]+71)$^{1+}$
	1115.578	1115.5696	1	αy_{12}
	1172.592	1172.5911	1	αy_{13}
	1314.657	1314.6653	1	αy_{15}
	1557.682	1557.6855	1	((^{1}Z-K^{16})+98)$^{1+}$

a) Fragment ions highlighted in blue were assigned manually by considering the cleavability of the linker fragmentation and their derived fragment ions. Z corresponds to a *N*-terminal acetylated cysteine residue. In the structure of the type 2 peptide, the α and β peptides are (^{1}Z-K^{16}) and ([^{192}T-K^{200}], respectively. The nomenclature of the backbone fragment ions is in agreement with the reported by Schilling *et al* [1]. The linker fragment ions assigned as (^{1}Z-K^{16})+98)$^{1+}$ and ([^{192}T-K^{200}]+71)$^{1+}$ correspond to the α peptide and β peptides, containing the formerly crosslinked Lys and Cys modified by structures that increase their molecular mass by 98 Da and 71 Da, respectively. The linker fragment ion assigned as (^{1}Z-K^{16})+98)$^{1+}$contain the Cys^{1} S-alkylated by a succinic anhydride residue. The linker fragment ion assigned as ([^{192}T-K^{200}]+71)$^{1+}$ contain the Lys^{193} modified at its epsilon amino group by a residue of β-alanine.The linker fragment ion assigned as ([^{192}T-K^{200}]+71)$^{1+}$ contain the Lys^{193} modified at its epsilon amino group by a residue of β-alanine. The linker fragment ion assigned as ([^{192}T-K^{200}]+71)$^{1+}$ contain the Lys^{193} modified at its epsilon amino group by a residue of β-alanine. Fragment ions assigned as $*(\alpha/\beta)b_n$ or

*(α/β)b_n correspond to the backbone fragment ions b_n or y_n derived from the linearized α and β peptides containing a modified Lys (+71 Da) or a Cys(+98 Da) residues as described above.

References

1. Schilling, B., et al., *MS2Assign, automated assignment and nomenclature of tandem mass spectra of chemically crosslinked peptides.* Journal of the American Society for Mass Spectrometry, 2003. **14**(8): p. 834-850, http://dx.doi.org/10.1016/S1044-0305(03)00327-1.