Journal Pre-proofs

Enrichment of histidine containing peptides by on-line immobilised metal affinity solid-phase extraction capillary electrophoresis-mass spectrometry

Roger Pero-Gascon, Estela Giménez, Victoria Sanz-Nebot, Fernando Benavente

PII: S0026-265X(20)31046-8
DOI: https://doi.org/10.1016/j.microc.2020.105013
Reference: MICROC 105013

To appear in: *Microchemical Journal*

Received Date: 29 March 2020
Revised Date: 8 May 2020
Accepted Date: 9 May 2020

Please cite this article as: R. Pero-Gascon, E. Giménez, V. Sanz-Nebot, F. Benavente, Enrichment of histidine containing peptides by on-line immobilised metal affinity solid-phase extraction capillary electrophoresis-mass spectrometry, *Microchemical Journal* (2020), doi: https://doi.org/10.1016/j.microc.2020.105013

Enrichment of histidine containing peptides by on-line immobilised metal affinity solid-phase extraction capillary electrophoresis-mass spectrometry

Roger Pero-Gascon, Estela Giménez*, Victoria Sanz-Nebot, Fernando Benavente

Department of Chemical Engineering and Analytical Chemistry, Institute for Research on Nutrition and Food Safety (INSA·UB), University of Barcelona, Martí i Franqués 1-11, 08028, Barcelona, Spain

*Corresponding author: estelagimenez@ub.edu (E. Giménez, PhD)

Tel: (+34) 934039378

Keywords: capillary electrophoresis / histidine / immobilised metal affinity / mass spectrometry / on-line solid-phase extraction / peptides

Abstract

The analysis and detection of targeted peptides has received due attention in many fields of proteomics research, such as discovery of biomarkers. In this regard, capillary electrophoresis-mass spectrometry (CE-MS) is widely used, however, direct analysis of low-abundance peptides such as histidine-containing peptides (His-peptides) in complex matrixes by CE-MS remains an analytical challenge. In the present study, an immobilised metal affinity solid-phase extraction containing Ni(II) coupled on-line to capillary electrophoresis-mass spectrometry (IMA-SPE-CE-MS) method has been developed to selectively enrich His-peptides from protein tryptic digests and enhance sensitivity. The method was optimised with α-casein and validated with other standard proteins (β-casein and κ-casein). Later, it was applied to an Escherichia Coli (*E. Coli*) whole cell lysate. IMA-SPE-CE-MS was very selective and allowed an enrichment factor up to 100-fold. The on-line enrichment and separation method coupled to MS detection is straightforward and advantageous over off-line pretreatment methods in terms of simplicity, cost-effectiveness and throughput.

1. Introduction

In proteomics, depletion of the most abundant proteins is recommended to reduce sample complexity and identify low-abundance proteins that usually provide more biologically and clinically relevant information. Among the proposed bottom-up strategies, the isolation of a subproteome based on the enrichment of peptides containing less abundant amino acids such as cysteine, histidine, methionine or tryptophan has always been of great interest [1,2]. In the case of histidine, it is present in the sequence of 85-95% of all the proteins in human and yeast proteomes, but only in 15-20% of the peptides that can be generated from these proteins after tryptic digestion [1,2]. Hence, the analysis of histidine containing peptides (His-peptides) from enzymatic digests enables to simplify the peptide mixture while keeping a significant coverage of the proteome accessible to quantification [3–6]. This could be an important complement to the comprehensive bottom-up proteomics studies supported by state-of-the-art high-end MS instruments, especially when specific parts of the protein structure need to be targeted.

Immobilised metal affinity (IMA) sorbents have been often used in off-line solid-phase extraction (SPE) for the selective enrichment of His-peptides before liquid chromatography-mass spectrometry (LC-MS) analysis [7,8]. In particular, pretreatment of enzymatic digests by IMA-SPE is regarded as an interesting approach to improve accuracy in quantitative proteomics because it is a simple and effective strategy to decrease sample complexity, hence reducing both interference and ion suppression effects in LC-MS analysis [5,6]. The choice of the metal ion immobilised on the support depends on the peptide to be purified. In the case of His-peptides, the most used are divalent cations such as Cu(II) and Ni(II) with iminodiacetic acid (IDA) as chelating ligand [4–11]. The on-line coupling of SPE to LC-MS allows minimising sample handling and

increasing analysis throughput but require complex instrumental set-ups with valves, which are especially delicate when working in capillary or nano LC [12]. On-line solid-phase extraction capillary electrophoresis-mass spectrometry (SPE-CE-MS) can be regarded as an excellent alternative for automated sample clean-up and sensitive analyses [13–16]. In the typical SPE-CE-MS configuration, which is valve-free, a microcartridge filled with a sorbent that selectively retains the target analytes, is integrated in-line near the inlet of the separation capillary [15]. After loading a large sample volume (~50-100 μL), retained compounds are eluted into the separation capillary with a small volume of a suitable solution (~50 nL) before separation and detection by CE-MS. Reversed-phase sorbents (e.g., C18) have been preferentially used to analyse peptides by SPE-CE-MS [13–15]. Immobilised metal affinity solid-phase extraction capillary electrophoresis (IMA-SPE-CE) has been mostly demonstrated with UV detection [17–20]. IMA-SPE-CE-MS has been explored to a lesser extent [21–23]. So far, only Cao et al have demonstrated the analysis of phosphopeptides from complex protein digests by IMA-SPE-CE-MS with a Fe(III) sorbent [21,22]. More recently, we successfully established a IMA-SPE-CE-MS method with a Ni(II) sorbent for the analysis of β-amyloid peptides as model compounds with histidine residues in standards, plasma and serum samples [23].

In the present work, the potential of IMA-SPE-CE-MS with a Ni(II) sorbent is investigated to selectively enrich His-peptides from complex protein digests. The method was first optimised with tryptic digests of α-casein (α-CSN), and then its applicability was further validated with β-casein (β-CSN) and κ-casein (κ-CSN) digests and a digest of an Escherichia Coli (*E. Coli*) whole cell lysate containing a 6x-His tagged protein.

2. Materials and Methods

2.1. Chemicals

All chemicals used in the preparation of buffers and solutions were of analytical reagent grade. Acetic acid (HAc, glacial), formic acid (HFor, 98-100%), sodium hydroxide, phosphoric acid (85%) and ammonium hydroxide (25%) were supplied by Merck (Darmstadt, Germany). DL-Dithiothreitol (DTT, ≥99%), iodoacetamide (IAA, ≥98%), imidazole (99.5%) and ammonium hydrogen carbonate (≥99.9%) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Isopropanol was provided by Scharlab (Barcelona, Spain) while acetonitrile and water by Sigma-Aldrich (all of them of LC-MS quality grade). Trypsin (sequencing grade modified, 16000 $U \cdot mg^{-1}$) and chymotrypsin were purchased from Promega (Madison, WI, USA). ESI low concentration (ESI-L) tuning mix was supplied by Agilent Technologies (Waldbronn, Germany) for tuning and calibration of the mass spectrometer.

2.2. Protein samples

α-Casein (α-CSN, 70%), β-casein (β-CSN, 90%) and κ-casein (κ-CSN, ≥80%) were obtained from Sigma-Aldrich. Stock solutions of 1000 $\mu g \cdot mL^{-1}$ were prepared in water and aliquoted. Aliquots were evaporated to dryness using a Savant SPD-111V SpeedVac concentrator (Thermo-Fisher Scientific, Waltham, MA, USA) and stored at -20 ºC until enzymatic digestion.

Escherichia Coli positive control whole cell lysate - expressing 6X His tag protein was obtained from Abcam (Cambridge, UK). The whole cell lysate of *E. Coli* contained 250 μg of total protein. The content of the vial was dissolved in water to obtain a 1000

113 μg·mL^{-1} protein solution. Excipients of low-molecular mass were removed by
114 ultracentrifugation using Microcon YM-10 centrifugal filters from Millipore (Mr cut-off
115 10,000, Bedford, MA, USA) [24]. Centrifugations were carried out in a Mikro 20
116 centrifuge (Hettich, Tuttligen, Germany) at 25 ºC. Finally, aliquots of 1000 μg·mL^{-1} of
117 *E. Coli* whole cell lysate were evaporated to dryness by SpeedVac and stored at −20 °C
118 until enzymatic digestion.

119 κ-Casein and *E. Coli* whole cell lysate were firstly reduced and alkylated to facilitate
120 digestion. Briefly, an aliquot of 100 μg of dried protein was dissolved in 100 μL of 50
121 mM NH_4HCO_3 (pH 7.9) and 2.5 μL of 0.5M DTT in the same buffer was added. The
122 mixture was incubated in a TS-100 thermoshaker (Biosan, Riga, Latvian Republic) at 56
123 ºC for 30 min and then alkylated by adding 7 μL of 0.73 M IAA (in the same buffer) and
124 shaking for 30 min at room temperature in the dark. Excess of low-molecular mass
125 reagents was removed with Microcon YM-10 centrifugal filters as described in [24]. The
126 final protein residue was dissolved in 50 mM NH_4HCO_3 (pH 7.9) to obtain a final
127 concentration of 1000 μg·mL^{-1}.

128 A 100 μL aliquot of 1000 μg·mL^{-1} protein solution in 50 mM NH_4HCO_3 (pH 7.9)
129 was digested. Trypsin was added in an enzyme to protein ratio of 1:40 m/m. In the case
130 of trypsin:chymotrypsin digestion, chymotrypsin was added simultaneously at 1:25 m/m.
131 The mixture was vortexed and subsequently incubated at 37 ºC for 18 h. Digestions were
132 stopped by heating at 100 ºC for 10 min, and the digest was evaporated to dryness by
133 SpeedVac and stored at -20 ºC until analysis. All digestions were performed in triplicate.

134
135
136
137

2.3. CE-MS

CE-MS experiments were performed in a 7100 CE system coupled to a 6220 oa-TOF LC/MS mass spectrometer with an orthogonal sheathflow interface (Agilent Technologies). The sheath liquid was delivered at a flow rate of 3.3 μL·min^{-1} by a KD Scientific 100 series infusion pump (Holliston, MS ,USA) and degassed for 10 min by sonication before use. CE control and separation data acquisition (e.g., voltage, temperature and current) were performed using Chemstation software (Agilent Technologies) that was running in combination with the MassHunter workstation software (Agilent Technologies) for control, data acquisition and processing of the mass spectrometer. The mass spectrometer was tuned and calibrated following the manufacturer's instructions. The operational conditions in positive electrospray ionization (ESI) mode were: capillary voltage 4000 V, drying gas (N_2) temperature 200 ºC, drying gas flow rate 4 L min^{-1}, nebuliser gas (N_2) 7 psig, fragmentor voltage 190 V, skimmer voltage 60 V and OCT 1 RF Vpp voltage 300 V. Data were collected in profile (continuum) at 1 spectrum·s^{-1} (approx. 10,000 transients·spectrum^{-1}) between m/z 100 and 3200 working in the highest resolution mode (4 GHz).

The extracted ion electropherogram of each peptide was obtained considering the theoretical m/z of the molecular ions and a mass tolerance window of 20 ppm. Mass spectra were checked manually to verify the isotopic distribution and accurate identification (mass error <10 ppm).

A bare fused-silica capillary of 72 cm total length (L_T) x 75 μm internal diameter (I.D.) x 360 μm outer diameter (O.D.) (Polymicro Technologies, Phoenix, AZ, USA) was used in CE-MS. Activation and conditioning procedures were carried out off-line in order to avoid contamination with NaOH of the mass spectrometer. New capillaries were activated by flushing (930 mbar) sequentially for 30 min each with 1M NaOH, water and

background electrolyte (BGE). Samples were reconstituted in water and injected at 50 mbar for 15 s (approximately 80 nL, i.e., 2.5% of the capillary, estimated using the Hagen–Poiseuille equation [25]. This volume corresponds to 4 ng of protein for a 50 $\mu g \cdot mL^{-1}$ protein digest). Electrophoretic separations were performed at 25 ºC and 15 kV under normal polarity (cathode in the outlet). Between runs, capillaries were flushed with water (1 min), 1 M HAc (3 min), water (1 min) and BGE (5 min). Before CE-MS, all solutions were passed through a 0.22-µm nylon filter (MSI, Westboro, MS, USA).

Two BGEs were used for the analysis of the protein digests: an acidic BGE of 50 mM HFor and 50 mM HAc (pH 2.2) with a sheath liquid of 60:40 (v/v) iPrOH:H_2O with 0.05% (v/v) of HFor, and a neutral BGE of 25 mM H_3PO_4 (adjusted to pH 7.5 with ammonium hydroxide) with a sheath liquid of 60:40 (v/v) iPrOH:H_2O with 0.5% (v/v) of HFor.

pH measurements were carried out with a Crison 2002 potentiometer and a Crison electrode 52-03 (Crison instruments, Barcelona, Spain).

2.4. SPE-CE-MS

A HisLink™ protein purification resin obtained from Promega (Madison, WI, USA) was used as immobilised metal affinity sorbent. It was made of 100 µm silica particles derivatised with nitrilotriacetic acid (NTA), which contained a high level of tetradentate-chelated Ni(II). Construction of the particle-packed microcartridge with frits for IMA-SPE-CE-MS was carried out as described elsewhere with little modifications [23]. Bare fused silica capillaries (Polymicro Technologies) were used for all the procedures. The microcartridge (0.7 cm L_T × 250 µm i.d. × 365 µm o.d. capillary) was inserted at 7.5 cm from the inlet of the separation capillary (72 cm L_T × 75 µm i.d. × 365 µm o.d., activated

as in CE-MS) using two plastic sleeves (Tygon® tube of 250 µm i.d., Thermo-Fisher Scientific). Previously the microcartridge was filled by vacuum with the IMA sorbent that was retained between two microfrits placed on each end of the microcartridge. These microfrits were obtained from the material of the original frits found in conventional C_{18} cartridges for SPE (Sep-Pack, Waters, Milford, MA, USA). As no glue was necessary to prevent microcartridge leaking, the microcartridge could be replaced to reuse the separation capillary.

The BGE for the IMA-SPE-CE-MS separation was the neutral BGE of 25 mM H_3PO_4 (adjusted to pH 7.5 with ammonium hydroxide) and the sheath liquid solution was the same as in CE-MS. Under the optimised conditions, the IMA-SPE-CE capillaries were first conditioned by applying pressure at 930 mbar (flush pressure) for 2 min with BGE. Afterwards, samples reconstituted in BGE and diluted to the desired concentration were loaded by flushing for 10 min (approximately 60 µL [25]). A final flush for 2 min with BGE was used for washing and removing non-specifically retained molecules. All these steps were performed with the nebuliser gas and the ESI capillary voltage switched off to prevent the introduction of contaminants into the MS. Then, both were switched on and a small volume of eluent (0.5% HAc) was injected at 50 mbar for 75 s (approximately 0.4 µL [25]). In order to prevent the eluent plug from traveling backwards due to the microcartridge backpressure, BGE was also injected at 25 mbar for 150 s. Separation was conducted at +15 kV for 40 min (cathode in the outlet). Postconditioning to avoid carryover was performed by flushing for 1 min with 50 mM imidazole and 1 min of water. This postconditioning step was also performed switching off the nebulizer gas and the ESI capillary voltage. IMA-SPE-CE-MS columns were filled with water when they were stored overnight to avoid salt precipitation.

3. Results and Discussion

A number of studies have taken advantage of the affinity selection of His-peptides with Cu(II)-IMA sorbents before LC-MS to reduce sample complexity and simplify identification of proteins in bottom-up proteomics, and virtually all these purification methods were off-line [3–6,9,10]. Very few studies have also explored the potential of off-line Ni(II)-IMA sorbents to preconcentrate His-peptides before LC-MS [11]. These Cu(II)- and Ni(II)-IMA sorbents typically contain IDA, which is a tridentate metal chelating agent. In general, it is well-known that chelating agents with a higher number of coordination sites prevent metal leaking because the sorbent become more stable at the expense of a weaker interaction with the target compounds [7,8]. In a previous study a Ni(II) sorbent with the tetradentate metal chelating agent NTA gave the best results in terms of reproducibility and sensitivity enhancement for the analysis by IMA-SPE-CE-MS of β-amyloid peptides, which contained histidine residues, in standards, plasma and serum samples [23]. In the present work, the potential of this Ni(II)-NTA sorbent was investigated to selectively enrich His-peptides from complex protein digests by IMA-SPE-CE-MS.

3.1 Analysis of His-peptides of α-CSN

With the aim of establishing a robust and reliable method to analyse His-peptides in complex protein digests, α-casein (α-CSN) was chosen as model protein because tryptic digestion yields a significant number of easily detected and well-characterised His-peptides. First, the coverage of peptides and, specifically His-peptides, in the tryptic digest of α-CSN was mapped by CE-MS using an acidic BGE (50 mM HFor and 50 mM HAc, pH 2.2) and a sheath-liquid of iPrOH:H_2O (60:40, with 0.05% HFor). Based on the

manufacturer information, α-CSN consists of two isoforms, i.e., α-CSN1 (23.7 kDa) and α-CSN2 (25.3 kDa), and α-CSN1 shows the highest abundance (80% *versus* 20% m/m, respectively in skim milk [26]). Table 1 shows the His-peptide sequences expected after trypsin digestion for α-CSN predicted by the PeptideMass tool from ExPASy (https://web.expasy.org/peptide_mass/). Additionally, Tables S1 and S2 show the peptide sequences expected after trypsin digestion of α-CSN, the theoretical m/z of the molecular ions and the peptides detected by CE-MS (and IMA-SPE-CE-MS). Fig. 1a shows the extracted ion electropherograms (EIEs) of the detected non-histidine containing peptides (non-His peptides; in black) and His-peptides (in colours) in the α-CSN digest using the acidic BGE. Under these conditions, at a protein concentration of 50 $\mu g \cdot mL^{-1}$, 17 non-His peptides from 35 (i.e., the total sum of α-CSN1 and α-CSN2 non-His peptides) were detected by CE-MS (48.6% of coverage). In the case of the His-peptides, 7 from 8 (i.e., the total sum of α-CSN1 and α-CSN2 His-peptides) were detected (87.5% of coverage): the 5 expected His-peptides for α-CSN1 and 2 out of 3 for α-CSN2. The α-CSN2 phosphorylated peptide [2-21] was the only His-peptide not detected, probably due to the lower ionization efficiency promoted by the presence of the phosphate groups and the relatively low-abundance of α-CSN2 in comparison to α-CSN1. Acidic BGEs with HAc or HFor are highly recommended for optimum detection sensitivity of peptides by CE-MS in positive ion mode [27]. However, acidic pH solutions promote metal ion leakage from the IMA sorbent in IMA-SPE-CE-MS. Hence, a neutral low ionic strength phosphate BGE (25 mM H_3PO_4, pH 7.5) was also evaluated by CE-MS, despite the lower volatility of phosphate solutions, as was the one which gave the best results for IMA-SPE-CE-MS in our previous study [23]. Fig. 1b shows the EIEs of the peptides and His-peptides of α-CSN detected by CE-MS using this neutral BGE. As can be seen, peaks were slightly wider due to the longer migration times promoted by the lower global

molecular charge that was not counterbalanced by the increase in the electroosmotic flow at pH 7.5. Furthermore, sensitivity was slightly lower and thus, the coverages of the peptides and His-peptides decreased (37.1% and 75%, respectively, see also Tables S1 and S2).

Once analysed α-CSN digest by CE-MS, the coverage for the non-His peptides and for the His-peptides was evaluated by IMA-SPE-CE-MS. The eluent composition and volume were optimised in our previous study [23]. In order to ensure the elution of the His-peptides from the Ni(II)-NTA sorbent and to obtain an adequate repeatability, eluent injection at 50 mbar for 75 s was required. When decreasing the amount of HAc or the injection time (e.g., 0.1% HAc or 50 s, respectively), elution of the His-peptides was incomplete and repeatability decreased, as was previously observed with β-amyloid peptides [23]. Fig. 1c shows the EIEs of the peptides and His-peptides detected by IMA-SPE-CE-MS in a 50 $\mu g \cdot mL^{-1}$ tryptic digest. As can be observed, only 2 non-His peptides (α-CSN1[1-3] and α-CSN2[189-197]) from a total of 35 were detected (5.7% of coverage). In contrast, 5 His-peptides from a total of 8 (62.5% of coverage) were detected (3 out of 5 and 2 out of 3 for α-CSN1 and α-CSN2, respectively), which demonstrated the selectivity of the methodology for the analysis of His-peptides (see Fig. 1c or Tables S1 and S2). With respect to the mass spectra of the His-peptides, the neutral BGE did not promote formation of peptide-phosphate adducts and protonated adducts were the most abundant as with the acidic BGE. To improve sensitivity and increase the His-peptide coverage, sample loading time was investigated loading the 50 $\mu g \cdot mL^{-1}$ α-CSN digest for 5, 10, 20 and 30 min at 930 mbar (Fig. 2a). For loading times up to 10 min, His-peptides peak area improved when the loading time increased, as expected due to the greater volume of sample loaded. However, longer loading times than 10 minutes did not favour detecting a larger number of His-peptides. In addition, the sample breakthrough volume

was exceeded and peak areas decreased because the amounts of His-peptides washed away were higher than the amounts retained. Therefore, to obtain the best enrichment factors, a loading time of 10 min was selected for the rest of the experiments. The coverage and linearity of IMA-SPE-CE-MS was also studied analysing α-CSN tryptic digests at concentrations from 20 to 100 $\mu g \cdot mL^{-1}$. At 100 $\mu g \cdot mL^{-1}$ the coverage was the same as for 50 $\mu g \cdot mL^{-1}$. In contrast, at 20 $\mu g \cdot mL^{-1}$, only one His-peptide (α-CSN1[4-7]) was detected, with very low intensity. Taking into account this His-peptide, the method was linear between 20 and 100 $\mu g \cdot mL^{-1}$ (R^2>0.98).

Consecutive analyses of the α-CSN digest were repeatable in terms of migration time and peak area (%*RSD* (n=3) were <3.0 and <11% at 50 $\mu g \cdot mL^{-1}$, respectively). The average lifetime of a microcartridge was 9 consecutive analyses (see supplementary Fig. S1). It was established by repeatedly analysing the 50 $\mu g \cdot mL^{-1}$ α-CSN tryptic digest until the peak areas for the His-peptides in the EIEs decreased more than 40% compared to the mean value of the first three analyses with the microcartridge under consideration.

The enrichment performance of IMA-SPE-CE-MS was examined with a 50 $\mu g \cdot mL^{-1}$ α-CSN digest. The graphic bar of Fig. 2b shows a comparison of the His-peptides peak area values obtained by CE-MS with the neutral BGE (Fig. 1b) and by IMA-SPE-CE-MS (Fig. 1c), and the corresponding preconcentration factors achieved. As can be observed, most His-peptides showed an increase in peak area by IMA-SPE-CE-MS, allowing detection of a short His-peptide (i.e., α-CSN2[77-80]), which was not detected by CE-MS. In contrast, His-peptides α-CSN1[8-22] and α-CSN1[125-132] were not detected by IMA-SPE-CE-MS probably due to its larger sequence and the position of the His residue (see the expected His-peptide sequences in Table 1). In addition to the sequence length, the number, accessibility and microenvironment of histidine residues may also be important in determining retention by the chelated metal [23,28]. Globally a lower

311 coverage of His-peptides was achieved by IMA-SPE-CE-MS compared to CE-MS (5 His-
312 peptides detected *versus* 6), but preconcentration factors from 25 to 98 were obtained for
313 those detected, demonstrating the potential of the IMA sorbent to selectively target His-
314 peptides, which are less abundant in the proteomes [1,2].

315

316 **3.2 Analysis of His-peptides of other model proteins**

317 To further validate the performance of the IMA-SPE-CE-MS method to selectively enrich
318 His-peptides from protein digests, the method was also applied to the analysis of other
319 standard proteins: β-casein (β-CSN) and κ-casein (κ-CSN) and it was compared to the
320 analysis by CE-MS. Fig. 2c shows a graphic bar with the expected His-peptides and the
321 peak area of those detected by CE-MS with the neutral BGE and IMA-SPE-CE-MS, when
322 analysing tryptic digests of β-CSN and κ-CSN at 50 $\mu g \cdot mL^{-1}$. In contrast to our previous
323 results with α-CSN, after His-peptide enrichment, no non-His peptides were detected and
324 the overall number of detected His-peptides was significantly reduced with regard to CE-
325 MS. While 3 His-peptides from a total of 4 were detected by CE-MS, only the His-peptide
326 [106-107] of β-CSN was detected by IMA-SPE-CE-MS. This was probably due to the
327 length of the His-peptides that were, in general, larger in these tryptic digests than in the
328 α-CSN digest (see Table 1). To study the influence of the peptide length in the retention
329 of the IMA sorbent, β-CSN and κ-CSN were also digested with a mixture of trypsin and
330 chymotrypsin. Table 1 shows the expected His-peptide sequences after trypsin and
331 chymotrypsin digestion for β-CSN and κ-CSN. As can be observed in Fig. 2d, after
332 digestion with trypsin:chymotrypsin, the number of His-peptides detected by IMA-SPE-
333 CE-MS increased compared to CE-MS (60% *versus* 40% of the His-peptide coverage).

Therefore, the reduction of the length of most of the His-peptides (see the sequences in Table 1) confirmed the selectivity of the IMA sorbent towards short His-peptides.

3.3 Analysis of *E. Coli* whole cell lysate

A whole cell lysate of *E. Coli* was analysed to explore the potential of the established methodology to selectively enrich His-peptides from complex and diluted biological samples. A whole cell lysate of *E. Coli* that expresses human α-lactalbumin with six additional histidine amino acids (6x-His) at its C-terminus was selected [29,30], which is typically used for quality control of purification of His-tagged proteins or as positive control for Western blots of 6x-His tagged protein samples. As in previous studies by LC-MS/MS [29] and CE-MS [30], 6x-His human α-lactalbumin, myoglobin C, cathepsin D, creatine kinase-MM and antithrombin III were identified by CE-MS and only the 6x-His tagged protein and those proteins producing short His-peptides were identified by IMA-SPE-CE-MS (i.e., 6x-His human α-lactalbumin, myoglobin and creatine kinase-MM). Table 2 shows the sequence of the peptides and His-peptides detected in the lysate tryptic digest (1000 μg mL^{-1}) by CE-MS using the acidic BGE and by IMA-SPE-CE-MS using the neutral BGE (50 μg·mL^{-1}). The findings by IMA-SPE-CE-MS demonstrated the selectivity of the methodology towards short His-peptides, which could be useful to specifically map certain regions of the proteomes. Moreover, the detection of the 6x-His tagged peptide of α-lactalbumin ([123-129]) only by IMA-SPE-CE-MS (see Fig. 3) indicate a possible additional application in the on-line preconcentration of intact His-tagged proteins or other His-tagged compounds such lipids [31] or carbohydrates [32]. J. Partyka et al recently reported improved sensitivity in CE-MS for the analysis of oligosaccharides and N-linked glycans using multicharged labeling by 6x-His tags [32].

Better limits of detection could be further achieved analysing these labelled carbohydrates by IMA-SPE-CE-MS.

Conclusions

We have developed an on-line IMA-SPE-CE-MS method to selectively enrich His-peptides from protein tryptic digests using a Ni(II)-NTA IMA sorbent. The method was evaluated and validated with tryptic digests of α-CSN, β-CSN and κ-CSN. The selectivity of the IMA sorbent towards short His-peptides was confirmed when casein proteins were digested with a mixture of trypsin and chymotrypsin. Preconcentration factors for the detected His-peptides ranged between 25 and 100 times, and the IMA-SPE microcartridges could be used up to 9 analyses with good repeatability of migration times and peak areas (%*RSD*<3.0 and <11%, respectively). The good performance of the IMA-SPE-CE-MS method was also demonstrated with an *E. coli* cell lysate, which indicated the great potential of this approach to reduce sample complexity in proteomics compared to off-line sample pretreatment methods, ensuring simplicity, cost-effectiveness and throughput.

Acknowledgments

This study was supported by a grant from the Spanish Ministry of Economy and Competitiveness (RTI2018-097411-B-I00) and the Cathedra UB Rector Francisco Buscarons Ubeda (Forensic Chemistry and Chemical Engineering). Roger Pero-Gascon acknowledges the Spanish Ministry of Education, Culture and Sport for a FPU (Formación del Profesorado Universitario) fellowship.

References

[1] H. Zhang, W. Yan, R. Aebersold, Chemical probes and tandem mass spectrometry: A strategy for the quantitative analysis of proteomes and subproteomes, Curr. Opin. Chem. Biol. 8 (2004) 66–75. https://doi.org/10.1016/j.cbpa.2003.12.001.

[2] F.E. Regnier, L. Riggs, R. Zhang, L. Xiong, P. Liu, A. Chakraborty, E. Seeley, C. Sioma, R.A. Thompson, Comparative proteomics based on stable isotope labeling and affinity selection, J. Mass Spectrom. 37 (2002) 133–145. https://doi.org/10.1002/jms.290.

[3] S. Wang, X. Zhang, F.E. Regnier, Quantitative proteomics strategy involving the selection of peptides containing both cysteine and histidine from tryptic digests of cell lysates, J. Chromatogr. A. 949 (2002) 153–162. https://doi.org/10.1016/S0021-9673(01)01509-6.

[4] R.R. Prasanna, S. Sidhik, A.S. Kamalanathan, K. Bhagavatula, M.A. Vijayalakshmi, Affinity selection of histidine-containing peptides using metal chelate methacrylate monolithic disk for targeted LC-MS/MS approach in high-

398 throughput proteomics, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 955–
399 956 (2014) 42–49. https://doi.org/10.1016/j.jchromb.2014.02.020.

400 [5] C. Mesmin, B. Domon, Improvement of the performance of targeted LC-MS
401 assays through enrichment of histidine-containing peptides, J. Proteome Res. 13
402 (2014) 6160–6168. https://doi.org/10.1021/pr5008152.

403 [6] J. Yang, F. Tian, M. Zhang, Y. Zhao, X. Qian, Y. Cai, W. Ying, Targeted
404 histidine-peptide enrichment improved the accuracy of isobaric-based
405 quantitative proteomics, Anal. Methods. 8 (2016) 5255–5261.
406 https://doi.org/10.1039/c6ay00332j.

407 [7] D. Todorova, M.A. Vijayalakshmi, Immobilized metal-ion affinity
408 chromatography, in: D.S. Hage (Ed.), Handb. Affin. Chromatogr., l, Taylor &
409 Francis, Boca Raton, FL, 2006: pp. 257–280.

410 [8] R.C.F. Cheung, J.H. Wong, T.B. Ng, Immobilized metal ion affinity
411 chromatography: a review on its applications, Appl. Microbiol. Biotechnol. 96
412 (2012) 1411–1420. https://doi.org/10.1007/s00253-012-4507-0.

413 [9] D. Ren, N.A. Penner, B.E. Slentz, H.D. Inerowicz, M. Rybalko, F.E. Regnier,
414 Contributions of commercial sorbents to the selectivity in immobilized metal
415 affinity chromatography with Cu(II), J. Chromatogr. A. 1031 (2004) 87–92.
416 https://doi.org/10.1016/j.chroma.2003.10.041.

417 [10] D. Ren, N.A. Penner, B.E. Slentz, F.E. Regnier, Histidine-rich peptide selection
418 and quantification in targeted proteomics, J. Proteome Res. 3 (2004) 37–45.
419 https://doi.org/10.1021/pr034049q.

420 [11] Y. Wu, G. Chang, Y. Zhao, Y. Zhang, Preparation of hollow nickel silicate

nanospheres for separation of His-tagged proteins, Dalt. Trans. 43 (2014) 779–783. https://doi.org/10.1039/c3dt52084f.

[12] J. Pan, C. Zhang, Z. Zhang, G. Li, Review of online coupling of sample preparation techniques with liquid chromatography, Anal. Chim. Acta. 815 (2014) 1–15. https://doi.org/10.1016/j.aca.2014.01.017.

[13] F. Benavente, S. Medina-Casanellas, E. Giménez, V. Sanz-Nebot, On-line solid-phase extraction capillary electrophoresis mass spectrometry for preconcentration and clean-up of peptides and proteins, in: N.T. Tran, M. Taverna (Eds.), Capill. Electrophor. Proteins Pept. Methods Protoc., Springer, New York, 2016: pp. 67–84. https://doi.org/10.1007/978-1-4939-4014-1_6.

[14] R. Ramautar, G.W. Somsen, G.J. de Jong, Developments in coupled solid-phase extraction-capillary electrophoresis 2013-2015, Electrophoresis. 37 (2016) 35–44. https://doi.org/10.1002/elps.201500401.

[15] L. Pont, R. Pero-Gascon, E. Gimenez, V. Sanz-Nebot, F. Benavente, A critical retrospective and prospective review of designs and materials in in-line solid-phase extraction capillary electrophoresis, Anal. Chim. Acta. 1079 (2019) 1–19. https://doi.org/10.1016/j.aca.2019.05.022.

[16] M.C. Breadmore, W. Grochocki, U. Kalsoom, M.N. Alves, S.C. Phung, M.T. Rokh, J.M. Cabot, A. Ghiasvand, F. Li, A.I. Shallan, A.S.A. Keyon, A.A. Alhusban, H.H. See, A. Wuethrich, M. Dawod, J.P. Quirino, Recent advances in enhancing the sensitivity of electrophoresis and electrochromatography in capillaries and microchips (2016–2018), Electrophoresis. 40 (2019) 17–39. https://doi.org/10.1002/elps.201800384.

[17] J. Cai, Z. El Rassi, Selective on-line preconcentration of proteins by tandem metal chelate capillaries-capillary zone electrophoresis, J. Liq. Chromatogr. 16 (1993) 2007–2024. https://doi.org/10.1080/10826079308019910.

[18] N.M. Vizioli, M.L. Rusell, M.L. Carbajal, C.N. Carducci, M. Grasselli, On-line affinity selection of histidine-containing peptides using a polymeric monolithic support for capillary electrophoresis., Electrophoresis. 26 (2005) 2942–2948. https://doi.org/10.1002/elps.200410416.

[19] K. Tsukagoshi, Y. Shimadzu, T. Yamane, R. Nakajima, Preparation of an iminodiacetic acid-modified capillary and its performance in capillary liquid chromatography and immobilized metal chelate affinity capillary electrophoresis, J. Chromatogr. A. 1040 (2004) 151–154. https://doi.org/10.1016/j.chroma.2004.03.063.

[20] L. Zhang, L. Zhang, W. Zhang, Y. Zhang, On-line concentration of peptides and proteins with the hyphenation of polymer monolithic immobilized metal affinity chromatography and capillary electrophoresis, Electrophoresis. 26 (2005) 2172–2178. https://doi.org/10.1002/elps.200410377.

[21] P. Cao, J.T. Stults, Phosphopeptide analysis by on-line immobilized metal-ion affinity chromatography-capillary electrophoresis-electrospray ionization mass spectrometry, J. Chromatogr. A. 853 (1999) 225–235. https://doi.org/10.1016/s0021-9673(99)00481-1.

[22] P. Cao, J.T. Stults, Mapping the phosphorylation sites of proteins using on-line immobilized metal affinity chromatography/capillary electrophoresis/electrospray ionization multiple stage tandem mass spectrometry, Rapid Commun. Mass Spectrom. 14 (2000) 1600–1606. https://doi.org/10.1002/1097-

0231(20000915)14:17<1600::AID-RCM68>3.0.CO;2-V.

[23] L. Ortiz-Martin, F. Benavente, S. Medina-Casanellas, E. Giménez, V. Sanz-Nebot, Study of immobilized metal affinity chromatography sorbents for the analysis of peptides by on-line solid-phase extraction capillary electrophoresis-mass spectrometry, Electrophoresis. 36 (2015) 962–970. https://doi.org/10.1002/elps.201400374.

[24] E. Giménez, R. Ramos-Hernan, F. Benavente, J. Barbosa, V. Sanz-Nebot, Analysis of recombinant human erythropoietin glycopeptides by capillary electrophoresis electrospray-time of flight-mass spectrometry, Anal. Chim. Acta. 709 (2012) 81–90. https://doi.org/10.1016/j.aca.2011.10.028.

[25] H.H. Lauer, G.P. Rozing, High performance capillary electrophoresis, second ed., Waldbronn, Germany, 2014, pp. 60-61, n.d.

[26] H.W. Modler, Functional properties of nonfat dairy ingredients - A review. Modification of products containing casein, J. Dairy Sci. 68 (1985) 2195–2205. https://doi.org/10.3168/jds.S0022-0302(85)81091-2.

[27] C. Wenz, C. Barbas, A. Lopez-Gonzalvez, A. Garcia, F. Benavente, V. Sanz-Nebot, T. Blanc, G. Freckleton, P. Britz-McKibbin, M. Shanmuganathan, F. De L'Escaille, J. Far, R. Haselberg, S. Huang, C. Huhn, M. Pattky, D. Michels, S. Mou, F. Yang, C. Neususess, N. Tromsdorf, E.E.K. Baidoo, J.D. Keasling, S.S. Park, Interlaboratory study to evaluate the robustness of capillary electrophoresis-mass spectrometry for peptide mapping, J. Sep. Sci. 38 (2015) 3262–3270. https://doi.org/10.1002/jssc.201500551.

[28] H. Block, B. Maertens, A. Spriestersbach, N. Brinker, J. Kubicek, R. Fabis, J.

Labahn, F. Schäfer, Immobilized-metal affinity chromatography (IMAC): A review, in: R.R. Burgess, M.P. Deutscher (Eds.), Guid. to Protein Purif., 2nd ed., Academic Press, Miami, FL, USA, 2009: pp. 439–473. https://doi.org/10.1016/S0076-6879(09)63027-5.

[29] J.C. Wright, M.O. Collins, L. Yu, L. Käll, M. Brosch, J.S. Choudhary, Enhanced peptide identification by electron transfer dissociation using an improved mascot percolator, Mol. Cell. Proteomics. 11 (2012) 478–491. https://doi.org/10.1074/mcp.O111.014522.

[30] L. Villegas, R. Pero-Gascon, F. Benavente, J. Barbosa, V. Sanz-Nebot, On-line protein digestion by immobilized enzyme microreactor capillary electrophoresis-mass spectrometry, Talanta. 199 (2019) 116–123. https://doi.org/10.1016/j.talanta.2019.02.039.

[31] T. Stora, Z. Dienes, H. Vogel, C. Duschl, Histidine-tagged amphiphiles for the reversible formation of lipid bilayer aggregates on chelator-functionalized gold surfaces, Langmuir. 16 (2000) 5471–5478. https://doi.org/10.1021/la991711h.

[32] J. Partyka, J. Krenkova, R. Cmelik, F. Foret, Multi-charged labeling of oligosaccharides and N-linked glycans by hexahistidine-based tags for capillary electrophoresis-mass spectrometry analysis, J. Chromatogr. A. 1560 (2018) 91–96. https://doi.org/10.1016/j.chroma.2018.05.030.

Figure Captions

Fig. 1 Extracted ion electropherograms for the analysis of a 50 µg·mL^{-1} α-CSN tryptic digest by CE-MS using (A) an acidic BGE (50 mM HFor and 50 mM HAc (pH 2.2)) and (B) a neutral BGE (25 mM H_3PO_4 (adjusted to pH 7.5)), and (C) by IMA-SPE-CE-MS (BGE: 25 mM H_3PO_4 (adjusted to pH 7.5)

Fig. 2 Analysis of a 50 µg·mL^{-1} α-CSN tryptic digest. (A) Plot of His-peptides peak area vs loading time at 930 mbar for IMA-SPE-CE-MS. (B) Comparison of His-peptides peak areas for CE-MS and IMA-SPE-CE-MS (BGE: 25 mM H_3PO_4 (adjusted to pH 7.5) in both cases). Analysis of 50 µg·mL^{-1} β-CSN and κ-CSN tryptic digests. Comparison of His-peptides peak areas for CE-MS and IMA-SPE-CE-MS (BGE: 25 mM H_3PO_4 (adjusted to pH 7.5) in both cases). Digestion with (C) trypsin and (D) trypsin:chymotrypsin. Preconcentration factor is given in numbers. The asterisk indicates that the preconcentration factor was not calculated because the His-peptide was not detected (n.d.) by CE-MS. Standard deviation is given as error bars (n=3)

Fig. 3 Extracted ion electropherograms of the 6x-His tagged peptide of human α-lactalbumin ([123-129]) in the *E. Coli* whole cell lysate tryptic digest by CE-MS (1000 µg·mL^{-1}, BGE: 50 mM HFor and 50 mM HAc, pH 2.2) (green) and by IMA-SPE-CE-MS (50 µg·mL^{-1}, BGE: 25 mM H_3PO_4 (adjusted to pH 7.5)) (blue)

The authors declare no conflicts of interest.

CRediT authorship contribution statement

535 **R. Pero-Gascon:** Methodology, Investigation, Visualization, Writing - original draft. **E.**
536 **Giménez:** Conceptualization, Supervision, Visualization, Writing - original draft, Writing
537 - review & editing. **V. Sanz-Nebot:** Writing - review & editing, Project administration,
538 Funding acquisition **F. Benavente:** Conceptualization, Supervision, Writing - review &
539 editing, Project administration, Funding acquisition.
540

541 - An on-line IMA-SPE-CE-MS method to selectively enrich His-peptides has
542 been developed.
543 - The method is simple and cost-effective, with potential for high-throughput.
544 - It allows preconcentration factors for His-peptides up to 100-fold.
545 - It could be also applied in the analysis of 6x-His tagged proteins and other
546 compounds.

547

548 **Table 1** Sequence of the His-peptides expected after trypsin and trypsin:chymotrypsin digestion
549 for α-CSN, β-CSN and κ-CSN predicted by the PeptideMass tool from ExPASy
550 (https://web.expasy.org/peptide_mass/). His residues in His-peptides are marked in bold and
551 phosphorylations in orange colour

552

Protein	Trypsin		Tryp	
α-CSN1	[4-7]	**H**PIK	[4-7]	**H**PIK
	[8-22]	**H**QGLPQEVLNENLLR	[8-22]	**H**QGLP
	[80-83]	**H**IQK	[80-83]	**H**IQK
	[120-124]	L**H**SMK	[120-124]	L**H**SMK
	[125-132]	EGI**H**AQQK	[125-132]	EGI**H**AC
α-CSN2	[2-21]	NTME**H**VSSSEESIISQETYK	[2-20]	NTME**H**
	[77-80]	**H**YQK	[77-78]	**H**Y
	[182-188]	TVYQ**H**QK	[185-188]	Q**H**QK
β-CSN	[49-97]	I**H**PFAQTQSLVYPFPGPIPNSLPQNIPPLTQTPVVVPPFLQPEVMGVSK	[49-52]	I**H**PF
	[106-107]	**H**K	[106-107]	**H**K
	[114-169]	YPVEPFTESQSLTLTDVENL**H**LPLPLLQSWM**H**QP**H**QPLPPTVMFPPQSVLSLSQSK	[120-143]	TESQSL
			[144-169]	M**H**QP**H**
κ-CSN	[98-111]	**H**P**H**P**H**LSFMAIPPK	[98-105]	**H**P**H**P**H**

553

554

555 **Table 2** Detected peptides and identified proteins in the *E. Coli* whole cell lysate tryptic digest
556 by CE-MS (1000 µg mL^{-1}, BGE: 50 mM HFor:50 mM HAc, pH 2.2) and by IMA-SPE-CE-MS (50
557 $\mu g \cdot mL^{-1}$, BGE: 25 mM H_3PO_4 (adjusted to pH 7.5)). His residues in His-peptides are marked in bold

558

ntified otein[a]	**CE-MS (1000 µg mL^{-1}, acidic BGE)**		**IMA-SPE-CE-MS (50 µg mL^{-1}, neutral BGE)**	
s tagged albumin 14 kDa	[2-5] QFTK	[109-114] ALCTEK	[123-129] L**HHHHHH**	
globin C 17 kDa	[43-45] FDK	[48-50] **H**LK	[48-50] **H**LK	[97-98]
eatine se-MM 43 kDa	[1-9] MPFGNT**H**NK [26-32] **H**NN**H**MAK [42-43] LR [44-45] DK [131-132] VR [133-135] TGR [136-138] SIK	[149-151] GER [173-177] YYPLK [267-292] AG**H**PFMWNQ**H**LGYVLTCPSNLGTGLR [315-316] LR [317-319] LQK [359-365] LMVEMEK [367-369] LEK	[26-32] **H**NN**H**MAK [106-107] **H**K [305-307] **H**PK	
epsin D 38 kDa	[110-112] VER [113-120] QVFGEATK [131-141] FDGILGMAYPR [190-192] YYK [246-249] ELQK	[278-281] LGGK [282-284] GYK [285-293] LSPEDYTLK [294-299] VSQAGK [340-347] VGFAEAAR		
hrombin III 49 kDa	[14-24] DIPMNPMCIYR [48-53] VWELSK [54-57] ANSR [140-145] LVSANR [146-150] LFGDK	[223-226] GLWK [237-241] ELFYK [263-275] VAEGTQVLELPFK [351-359] LPGIVAGR [394-399] SLNPNR		

[177-183] ENAEQSR	
[184-188] AAINK	

559 ª Peptide and protein identification was based on the intact molecular ions of the detected
560 peptides and on previous studies by LC-MS/MS [28] and CE-MS [29].

561

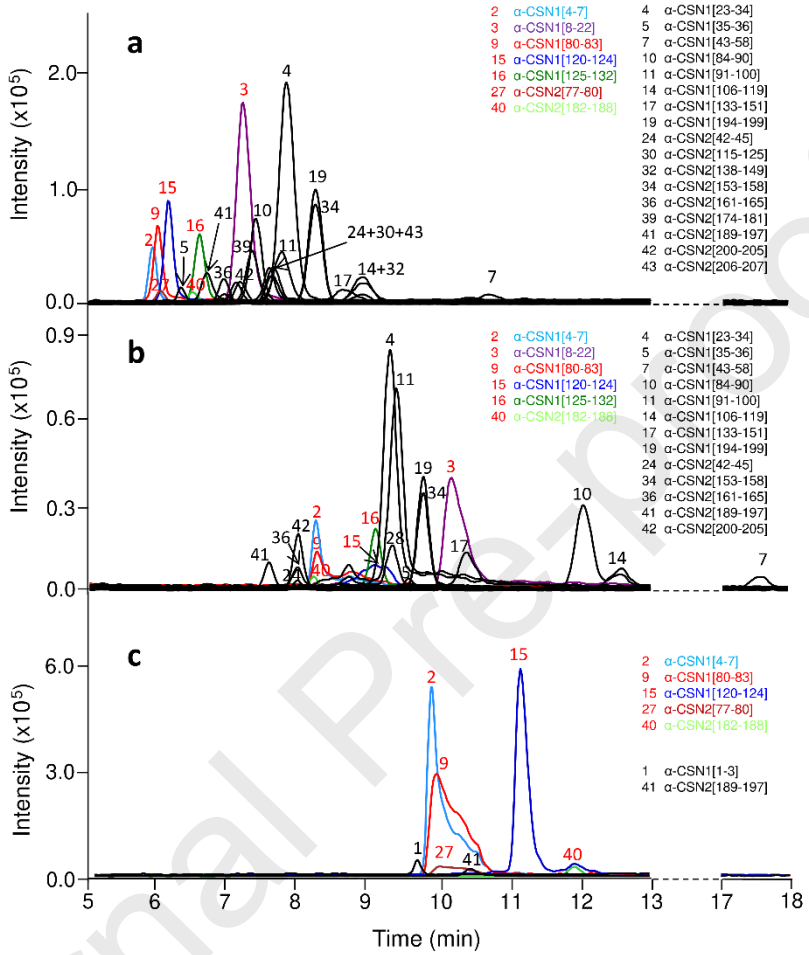
a
b
c
Intensity (x10^5)
Time (min)
2 α-CSN1[4-7]
3 α-CSN1[8-22]
9 α-CSN1[80-83]
15 α-CSN1[120-124]
16 α-CSN1[125-132]
27 α-CSN2[77-80]
40 α-CSN2[182-188]
4 α-CSN1[23-34]
5 α-CSN1[35-36]
7 α-CSN1[43-58]
10 α-CSN1[84-90]
11 α-CSN1[91-100]
14 α-CSN1[106-119]
17 α-CSN1[133-151]
19 α-CSN1[194-199]
24 α-CSN2[42-45]
30 α-CSN2[115-125]
32 α-CSN2[138-149]
34 α-CSN2[153-158]
36 α-CSN2[161-165]
39 α-CSN2[174-181]
41 α-CSN2[189-197]
42 α-CSN2[200-205]
43 α-CSN2[206-207]
1 α-CSN1[1-3]
24+30+43
14+32

Figure 1

562

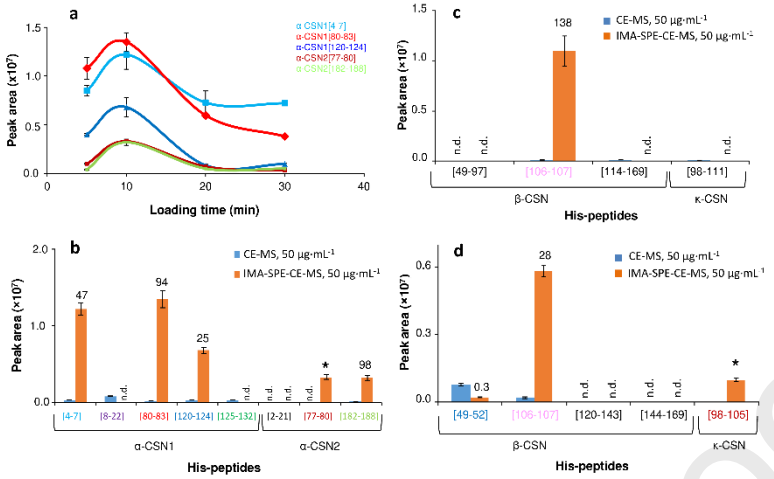

Figure 2

563

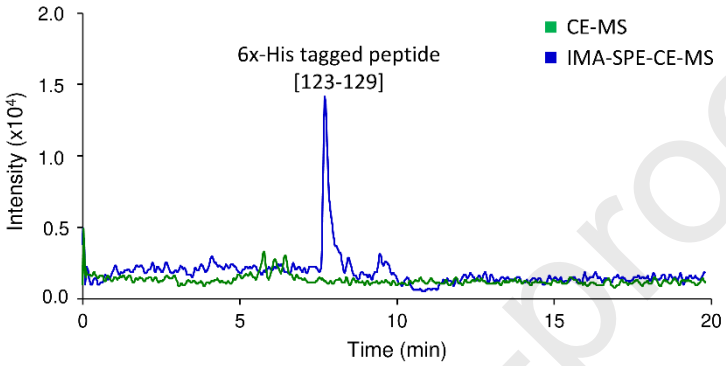

Figure 3

564