Loïc Dayon
Christophe Roussel
Michel Prudent
Niels Lion
Hubert H. Girault

Laboratoire d'Electrochimie
Physique et Analytique,
Institut des Sciences et
Ingénierie Chimiques,
Ecole Polytechnique Fédérale
de Lausanne,
Lausanne, Switzerland

On-line counting of cysteine residues in peptides during electrospray ionization by electrogenerated tags and their application to protein identification

The electrochemically induced mass spectrometric tagging of cysteines by substituted hydroquinones was studied for peptides in a classical electrospray solvent (*i.e.*, MeOH/H₂O/AcOH 50/49/1). The tagging efficiency was tested with different hydroquinone compounds on an undecapeptide containing one cysteine residue. 2-Carboxymethylhydroquinone was the most reactive probe and revealed to be suitable for cysteine quantification in peptides containing up to three cysteine residues, even in the case of two consecutive cysteines in the sequence. We demonstrate the compatibility of the on-line electrochemical tagging method for the cysteine content analysis of peptides coming from gel-free protein digestion procedures. The identification of bovine serum albumin and human α -lactalbumin digest samples in a peptide mapping strategy was greatly improved by the application of the electrotagging technique as post-column treatment. Indeed, the determination of cysteine content in the tryptic peptides provided powerful information in order to enhance the identification score as well as the discrimination against other protein candidates. The tagging method was then applied to the determination of four proteins in a model mixture.

Keywords: Electrochemistry / Mass fingerprinting / Mass spectrometry / Mass tag / Proteomics
DOI 10.1002/elps.200406144

1 Introduction

Protein identification is one of the main tasks in proteomics, which refers to the study of the protein content of a biological sample. With this aim in view, some identification procedures based on mass spectrometry (MS) [1, 2] have been developed, such as the study of the peptides coming from the proteolytic digestion of proteins [3, 4]. In a typical gel-free bottom-up approach [2], the proteolysis can lead from 30 to more than 100 peptides per digested proteins [5]. A protocol based on reversed-phase high-performance liquid chromatography-mass spectrometry (RP-HPLC-MS) is commonly used to separate and analyze the generated peptides [6, 7]. The parent protein is found due to the mass fingerprint of the tryptic peptides and comparison with DNA and protein databases [8]. Besides, it is advantageous to use supplementary information (in addition to the peptides masses) to constrain the database search. Sechi and Chait [9] have demon-

strated that protein identification can be greatly improved when knowing the content of one specific amino acid residue within tryptic peptides since it rules out many potential matches. It can be used for a large range of proteins when quantifying a well-chosen amino acid in proteomes [10]. Cysteine, contained in 89.3%, or histidine, contained in 83% of all proteins in humans, are good candidates [5]. The knowledge that a peptide contains at least one cysteine residue, together with a mass accuracy analysis of 0.5 Da, reduces by a factor of 2.5–10 the number of possible protein candidates from which such a peptide is derived [11].

The chemical tagging of free cysteines by iodoacetamide or by other alkylating reagents [12–14] was first introduced in protein 2-D gel electrophoresis procedures [15]. Cysteine alkylation avoids in particular the formation of acrylamide adducts during electrophoresis. The tagging of free cysteines can be also used for the counting of cysteines in peptides. MS methods to quantify cysteines are based on the mass shift between the modified and the untagged biomolecule or can rely upon the isotopic distribution pattern specific to the alkylated peptides [9, 16–18]. They often require a time-consuming sample preparation [19], and when used with a quantitative label, a twofold dilution of the sample, one-half unlabeled, the other labeled.

Correspondence: Professor Hubert H. Girault, Laboratoire d'Electrochimie Physique et Analytique, Institut des Sciences et Ingénierie Chimiques, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland
E-mail: hubert.girault@epfl.ch
Fax: +41-21-693-36-67

Abbreviation: β -hCG, human chorionic gonadotropin β (109–119)

Figure 1. Reaction mechanism of the electrochemical tagging of cysteines.

The recently developed on-line electrochemical tagging of free cysteine residues using a microspray chip [20] as electrospray ionization (ESI) source is a practical alternative to the previous sequential methods. Such a process yields instantaneously both tagged and untagged biomolecules directly from a single sample. A microspray device with an integrated carbon electrode is used to electrogenerate species able to react specifically with cysteine moieties in acidic spray condition [21]. As the device behaves like an electrolytic cell [22, 23], benzoquinone compounds are generated at the electrode by oxidation of hydroquinone added to the analyte. The tagging of free cysteine by hydroquinones coupled to MS detection involves an electrochemical-chemical-electrochemical (ECE) mechanism as it was previously described (Fig. 1) [21, 24, 25]. First, benzoquinones are electrogenerated at the electrode integrated into the microspray MS chip. Then, the benzoquinone forms undergo a 1,4-Michael addition of free cysteine before being analyzed by MS. The method was originally applied to the target protein β -lactoglobulin A with only one free cysteine residue using hydroquinone [21]. The obtained spectrum presented two peak distributions: the unmodified one, corresponding to the starting biomolecule, and a shifted one, corresponding to the adduct resulting from the labeling of the free cysteine residue contained in the protein. Several studies were conducted to optimize the tagging reaction and several substituted hydroquinones, presenting different reactivities toward L-cysteine, were employed. The kinetics of the addition of L-cysteine was determined by electrochemical studies and the final on-line tagging efficiency was investigated by MS measurements [25–27].

In this paper, the study of substituted hydroquinones as tags for cysteine-containing peptides was carried out. The reactivity/tagging efficiency was investigated on a target peptide, the human chorionic gonadotropin β (109–119) (β -hCG) containing one cysteine residue, using several hydroquinones: 2-methoxyhydroquinone, 2-methylhydroquinone, hydroquinone, and 2-carboxymethylhydroquinone. Detailed investigations were then pursued on MS electrotagging using hydroquinone and 2-carboxymethylhydroquinone for the counting of cysteine units in peptides. Peptides containing up to three cysteines were tested and the tagging of a peptide containing two consecutive cysteine residues was studied. The electrotagging with 2-carboxymethylhydroquinone was thus applied to bovine serum albumin (BSA) and human α -lactalbumin digestions in a gel-free peptide mapping strategy in order to highlight the improvement of the protein identification by the knowledge of the cysteine content of some tryptic peptides in comparison with the classical procedure. The post-column MS electrotagging was then tested for the identification of four proteins in a model mixture.

2 Materials and methods

2.1 Chemicals

Hydroquinone (> 98%), ammonium bicarbonate (> 98%), 1,4-dithio-DL-threitol (DTT) (≥ 99.5), and myoglobin from horse heart (> 90%) were purchased from Fluka (Buchs, Switzerland), 2-methylhydroquinone (99%), 2-methoxyhydroquinone (99%) from Acros (Pittsburgh, PA, USA),

2-carboxymethylhydroquinone (99%) from Aldrich (Milwaukee, WI, USA), and iodoacetamide, bradykinin (99%), Met-enkephalin-Arg-Phe (98%), BSA, human α -lactalbumin (> 90%), albumin from chicken eggs (ovalbumin, > 98%), and β -lactoglobulin A from bovine milk from Sigma (St. Louis, MO, USA). Chorionic gonadotropin- β (109–119) (> 94%), Leu-enkephalin (> 99%), and KCTCCA (70%) were bought from Bachem (Bubendorf, Switzerland). Porcine trypsin was from Promega (Madison, WI, USA). Methanol (> 99.8%, Riedel-de Haën, Seelze, Germany), acetic acid (> 99.5%, Fluka), trifluoroacetic acid (TFA) (99%, Merck, Darmstadt, Germany), and acetonitrile (> 99.5%, Fluka) were used without any further purification. Deionized water (18.5 M Ω) was prepared using a Milli-Q system from Millipore (Bedford, MA, USA). Solvents used for HPLC were water (HPLC-grade) from Sds (Peypin, France) and acetonitrile ultra-gradient HPLC-grade from J.T. Baker (Deventer, Netherlands). The synthetic peptides ALRCTCS and ACKCTCM were prepared by Catherine Servis at the Institut de Biochimie (Faculté de Médecine, Epalinges, Switzerland).

2.2 Microchip device

The fabrication of the microspray interface has been previously described [20]. Briefly, a polyethylene terephthalate (PET) substrate (100 μ m thick Melinex[®] sheet from Dupont (Wilmington, DE, USA)) was photoablated with a UV excimer laser (ArF 193 nm from Lambda Physik (Göttingen, Germany)). The carbon electrode was 70 $\times$ 25 μ m² integrated in the channel of a 35 $\times$ 30 μ m² section. The distance from the electrode to the outlet of the microchip was 1.8–2 cm.

2.3 MS setup

An LCQ DUO ion-trap mass spectrometer (Finnigan, San José, CA, USA) was used. The heated capillary was kept at 200°C. In each experiment, the ion transmission parameters are optimized automatically in order to improve the detection of the analyte of interest (the unmodified peptide). The ESI interface removed, the microchip holder was mounted on the probe slide adapter of the mass spectrometer. The device was coupled to a syringe pump (KdScientific, Holliston, MA, USA) to introduce the solution. The flow rate was set to 250 nL $\cdot$ min⁻¹ and the voltage applied was 3.5–4 kV. The distance from the outlet of the microchip to the entrance of the spectrometer varied between 1 and 2 cm in order to optimize the signal and the trap injection time. In a typical experiment, a given peptide was mixed with a given hydroquinone in order to obtain 50 μ M and 20 mM respectively. The spray medium used was MeOH/H₂O/CH₃COOH 50/49/1. The solutions were not degassed.

2.4 Digestion and peptide analysis

Two mg of protein were dissolved in 2 mL ammonium bicarbonate solution (5 mM, pH 8), and 1.23 mg DTT (4 mM) and 20 μ g trypsin (protein ratio of 1:100 w/w) was added. For the mixture of four proteins, 1 mg of each protein was dissolved in 4 mL buffer solution (40 μ g trypsin was used). The digestion was run at 37°C for 4 h. The solution was lyophilized and the obtained solid was then diluted in 400 μ L H₂O; 100 μ L were injected for separation. Before separation, the BSA digest sample was divided into two equal parts, and one of them was subjected to iodoacetamide alkylation (+1.41 mg iodoacetamide (38 mM), reaction performed at 45°C for 30 min in the dark). The HPLC separation (Alliance 2690 Waters system using Millennium software) was performed using a C₁₈ reverse phase column. The separation was run for 90 min using a gradient of H₂O/TFA 99.88/0.12 (solvent A) and acetonitrile (solvent B). The gradient was run as follows: 0–5 min 100 % A, then to 62 % A and 38 % B at 65 min, 43 % A and 57 % B at 75 min, 20 % A and 80 % B at 90 min at a flow rate of 1 mL $\cdot$ min⁻¹. Fractions were collected every minute. To 90 μ L of a given fraction, 10 μ L 2-carboxymethylhydroquinone (200 mM) in CH₃CN/H₂O/TFA 50/49/1 were added to be analyzed by microspray MS. To 90 μ L of a fraction of the alkylated sample was added 10 μ L CH₃CN/H₂O/TFA 50/49/1.

2.5 Protein identification

Single protein samples were identified using the Mascot search software [28, 29]. The model mixture was recovered using ProFound [30, 31]. Mascot search was performed with the following parameters: charge states of +1, +2, and +3, 1 maximum miscleavage, possible oxidation of methionine as variable modification, no fixed modification (except for the iodoacetamide alkylated sample for which carbamidomethyl was chosen), peptide mass tolerance of $\pm$ 0.4 Da. The main chosen parameters for the mixture are 1 miscleavage, tolerance of 0.4 Da, taxonomy Chordata and the search was focused with one charge state peptides. The search was performed in the Swiss-Prot database [32]. In all tagging experiments, the cysteine content information was only entered for peptides detected as containing some cysteines. The information on the supposed absence of cysteine was not specified.

3 Results and discussion

3.1 Tagging of β -hCG

The cysteine selective electrochemical tagging method [27] was successfully assessed with a target peptide containing one cysteine residue, β -hCG (**3b**) (M =

1269.3 $\text{g} \cdot \text{mol}^{-1}$) using 2-methoxyhydroquinone (**1a**), 2-methylhydroquinone (**1b**), hydroquinone (**1c**), and 2-carboxymethylhydroquinone (**1d**) (Fig. 1). The electrochemical MS tagging reaction was monitored (Fig. 2) using the microspray interface described above and a classical MS medium (MeOH/H₂O/AcOH 50/49/1). The tagging occurs with all hydroquinones **1a–d**, and the presence of both the untagged peptide as well as the tagged product was observed (Fig. 2). The presence of sodium and potassium adducts on the peptides accounts for extra peaks in the spectra. The tagging efficiency for each probe was compared, by calculating the average conversion yield on 40 min (0.55, 0.24, 0.22, and 0.56 for **1a–d**, respectively; the conversion yield was calculated as the ratio of the abundance of the tagged peptide over the sum of the untagged and tagged peptide on MS spectra). The tagging conversion

yields follow the electronic effect of each substituent excepted for **1a**, which should normally exhibit the lowest yield.

The particular reactivity of 2-methoxyhydroquinone (**1a**) had been previously observed with L-cysteine (**3a**) and attributed to a better ionization efficiency of the formed adduct **3a+2a** compared with adducts coming from the reaction with **2b** and **2c** [27]. However, no ionization differences between the peptide and the peptide+benzoquinone adducts could be seen (data not shown). As previously observed, the higher reactivity of **1a** could be explained by its natural oxidation by oxygen in the spray solution, leading to a chemical tagging in the infusion line prior to the electrochemical MS tagging [27]. As a proof, the tagging yield involving **1a** was observed to evolve through time, from 0.38 to 0.64 after 40 min infusion (MS experiments were not run under inert atmosphere).

Figure 2. MS spectra of β -hCG infused at $250 \text{ nL} \cdot \text{min}^{-1}$ in the presence of (a) **1a**, (b) **1b**, (c) **1c**, and (d) **1d**.

3.2 Tagging of peptides containing several cysteines

Since lots of proteins contain many cysteine residues, the probability to find a peptide coming from a protein digestion having more than one cysteine is relevant. Thus, in order to check if the method could be applied to the quantification of cysteine units in tryptic peptides, the on-line electrochemical tagging of peptides containing more than one cysteine residue as well as a peptide presenting two consecutive cysteines was investigated. In the following, we focused only on the reaction of **1c** and **1d** with several peptides. This choice was justified by the fact that, on one hand, **1d** always provided the best tagging efficiency because of the kinetics of the addition reaction and, on the other hand, **1d** is the less naturally oxidizable hydroquinone compound [27]. Hydroquinone (**1c**) is also taken as reference since it is the simplest, and the most studied species for on-line tagging [21, 25–27]. Moreover, it is not quickly oxidized in the presence of oxygen [21].

First, the electrochemical tagging of two synthetic peptides was studied: one contains two cysteine residues (ALRCTCS, $M = 752.33 \text{ g} \cdot \text{mol}^{-1}$) and the other presents three cysteine moieties (ACKCTCM, $M = 758.26 \text{ g} \cdot \text{mol}^{-1}$).

Hydroquinone compounds **1c** and **1d** were investigated for the electrochemical tagging of ALRCTCS. As can be observed in the spectra of Fig. 3, **1d** provides again the best tagging efficiency, evidenced by the fact that the second tagging peak presents a good intensity compared to the one obtained with **1c**. Then, ACKCTCM was infused in the presence of hydroquinone compounds **1c** and **1d**. The tagging results were followed by MS. Using **1c**, it was not possible to obtain the tagging of all cysteine residues (Fig. 4a). In contrast, the use of **1d** showed that it is possible to tag the three residues at the same time, providing the quantification of free cysteines present in this peptide. Indeed, the signals coming from the singly, doubly and triply tagged peptide are well observable on the MS spectrum; the presence of the untagged peptide signal is also visible (Fig. 4b).

Figure 3. MS spectra of peptide ALRCTCS infused in the presence of (a) **1c** and (b) **1d**.

Figure 4. MS spectra of peptide ACKCTCM infused in the presence of (a) **1c** and (b) **1d**.

The on-line electrochemical tagging scheme was then assessed in the challenging case of the synthetic peptide KCTCCA ($M = 627.81 \text{ g} \cdot \text{mol}^{-1}$), containing three cysteines, two of which being consecutive. Despite possible steric hindrance, Figs. 5a and b show clearly that

Figure 5. MS spectra of peptide KCTCCA infused in the presence of (a) **1c** and (b) **1d**. MS/MS spectrum at m/z 960.2 ($[M+2\text{tags}+H]^+$) corresponding to the two times tagged peptide. 40% collision energy was employed. α -, β - and γ -peptides represent the different two tagged peptide possibilities (* means that the cysteine is labeled). The peak at m/z 928.1 corresponds to the loss of sulfur [35].

the tagging happens in an efficient manner with **1d** whereas with **1c** only the singly tagged peak is well observed. Yet, on a set of experiments, one could notice that the triply tagged peptide peak was sometimes poorly intense (*i.e.*, the triply tagged compound is not greatly produced). As a matter of facts, the tagging of two consecutive cysteines may be difficult. In this way, the tandem mass spectrometric (MS/MS) spectrum (Fig. 5c) of the peak at m/z 960.2 (corresponding to the doubly tagged peptide $[M+2\text{tags}+H]^+$) lets think that the two consecutive cysteine residues are not tagged at the same time (*i.e.*, it is easier to tag distant cysteine residues). For the doubly tagged peptide, three possible adducts can be generated. Indeed, α -peptide (KC^*TC^*CA), β -peptide (KC^*TCC^*A), and γ -peptide ($KCTC^*C^*A$) can be formed, in which the asterisk represents the cysteine residue labeled by 2-carboxymethylhydroquinone (**1d**). The predicted peptide fragments obtained in the three cases were calculated and gathered together in Table 1 [33]. The mass fragments obtained by MS/MS do not match with

Table 1. Predicted fragment masses from MS/MS experiments for KCTCCA peptide tagged on two cysteine residues

Fragment type	α -Peptide ^{a)} (KC^*TC^*CA) fragment mass (Da)	β -Peptide ^{a)} (KC^*TCC^*A) fragment mass (Da)	γ -Peptide ^{a)} ($KCTC^*C^*A$) fragment mass (Da)
a_n	a_5^* : 843 a_4 : 740 a_3 : 471 a_2 : 370	a_5 : 843 a_4 : 574 a_3 : 471 a_2 : 370	a_5 : 843 a_4 : 574 a_3 : 305 a_2 : 204
b_n	b_5 : 871 b_4 : 768 b_3 : 499 b_2 : 398	b_5 : 871 b_4 : 602 b_3 : 499 b_2 : 398	b_5 : 871 b_4 : 602 b_3 : 333 b_2 : 232
c_n	c_5 : 888 c_4 : 785 c_3 : 516 c_2 : 415	c_5 : 888 c_4 : 619 c_3 : 516 c_2 : 415	c_5 : 888 c_4 : 619 c_3 : 350 c_2 : 249
x_n	x_5 : 858 x_4 : 589 x_3 : 488 x_2 : 219	x_5 : 858 x_4 : 589 x_3 : 488 x_2 : 385	x_5 : 858 x_4 : 755 x_3 : 654 x_2 : 385
y_n	y_5 : 832 y_4 : 563 y_3 : 462 y_2 : 193	y_5 : 832 y_4 : 563 y_3 : 462 y_2 : 359	y_5 : 832 y_4 : 729 y_3 : 628 y_2 : 359
z_n	z_5 : 815 z_4 : 546 z_3 : 445 z_2 : 176	z_5 : 815 z_4 : 546 z_3 : 445 z_2 : 342	z_5 : 815 z_4 : 712 z_3 : 611 z_2 : 342

a) The peptide can be tagged twice in three different manners (α -, β -, and γ -peptide)
* represents a tagging site

those relative to the peptide tagged on the two consecutive cysteine moieties (Fig. 5c, Table 1). They all correspond to fragments of α - and β -peptides, which are tagged on distant residues. Thus, MS/MS experiments of on-line tagged products provide significant information for the localization of the cysteines [34]. These observations can be explained by steric hindrance reasons. When the first cysteine residue is tagged, the surrounding other sites become less accessible to the probe and especially the consecutive ones. Nevertheless, as the triply tagged product is all the same well observed (Fig. 5b), the good reactivity of 2-carboxymethylhydroquinone (**1d**) is able to counterbalance the previously mentioned steric hindrance factor.

This set of results confirms that the method is quite convenient due to its nonquantitative efficiency under such conditions, since it does not require the preparation of two samples to access the number of cysteines. Moreover, the electrolabeling is quasi-instantaneous compared to the chemical labeling which is often a time-consuming step [11]. The tool reveals itself as an efficient way to quantify cysteine in peptides.

To confirm that the on-line tagging can be used in other classical spray media, the synthetic peptide ACKCTCM was analyzed in $\text{CH}_3\text{CN}/\text{H}_2\text{O}/\text{TFA}$ 50/49.9/0.1 with **1d**. The tagging occurs as in the classical medium used for on-line tagging. Since the method can be quite useful for the study of protein digests, the tagging of ACKCTCM was also performed in $\text{CH}_3\text{CN}/\text{H}_2\text{O}/\text{TFA}$ 20/79.9/0.1 and $\text{CH}_3\text{CN}/\text{H}_2\text{O}/\text{TFA}$ 80/19.9/0.1 media. The tagging remains efficient: the three cysteines are labeled at the same time (data not shown). The selectivity for cysteine residues was tested under these different conditions on several peptides (bradykinin, Leu-enkephalin, and Met-enkephalin-Arg-Phe) and myoglobin from horse heart, which are cysteine-free. As previously obtained in $\text{MeOH}/\text{H}_2\text{O}/\text{CH}_3\text{COOH}$ 50/49/1 [27], the tagging reaction was found to be selective under the experimental conditions.

3.3 Mass fingerprinting of BSA and human α -lactalbumin using on-line electrochemical tagging

As a proof of principle, the cysteine quantification tagging method was applied to a typical mass fingerprinting experiment of BSA, which contains 35 cysteines [3, 4]. BSA was reduced, digested, and the tryptic peptides separated by RP-HPLC. The fractions corresponding to retention times between 10 and 75 min on the chromatogram (Fig. 6a) were analyzed by chip-ESI-MS with the electrochemical tagging using 2-carboxymethylhydroquinone (**1d**). As a result, 22 peptides were tagged,

Figure 6. Chromatogram at 214 nm of (a) BSA digest and elution gradient used (acetonitrile evolution gradient is given), (b) α -Lactalbumin, and (c) a mixture of BSA, α -lactalbumin, ovalbumin, and β -lactoglobulin.

two of which corresponded to a doubly charge state of another tagged peptide. Seven of them were unidentified and 13 belonged to BSA. The tagged peptides peaks were well-defined. Three peptides belonging to BSA were tagged twice, showing the relevance of the study of the tagging of peptides with more than one cysteine.

The entire list of peptides was submitted to Mascot sequence query [29] for a search in the Swiss-Prot [32] database without giving any information on the taxonomy of the studied protein (see Section 2 for the search parameters). Without giving information on the cysteine content, BSA was identified with a score of 81 and protein coverage of 74% (63 peptides matched). When the information on the cysteine content of the tagged peptides was added (the number of cysteines is specified for cysteinyl-containing peptides by “comp(number[C])” entered after the peptide mass), the score raised to 206 (Fig. 7a). In comparison, the identification procedure was carried out on an iodoacetamide-alkylated BSA sample (see Section 2). In this case, BSA was identified with a score of 83 (coverage 61%, 56 peptides attributed to BSA). As this last experiment did not provide the same coverage from the precedent one, we decided to simulate the alkylation by iodoacetamide (+57 Da) of all cysteine-

Figure 7. Identification scores without and with the cysteine content information obtained by on-line electrochemical tagging for (a) BSA, (b) α -lactalbumin using Mascot, and (c) for a mixture of BSA, α -lactalbumin, ovalbumin, and β -lactoglobulin using ProFound.

containing peptides based on the peptide list of the untagged BSA search. The score gave 89 (coverage 74%). Thus, the alkylation with iodoacetamide does not bring the same information to constrain the search as the indication of the number of cysteines. The determination and indication of the peptide cysteine content [9, 18] with electrotagging was found to greatly improve the protein identification compared to (i) the untagged sample and (ii) the sample alkylated with iodoacetamide. Moreover, when only the 22 tagged peptides were entered with their cysteine content, the score was 160 with 22% coverage (13 peptides found) proving that the cysteine content

information is a powerful data and that the selection of cysteine-containing peptides coupled to the on-line tagging method could be considered.

Similarly, human α -lactalbumin, which contains only eight cysteines, was subjected to the mass fingerprinting electrochemical tagging experiment (Fig. 6b). Twenty-three peptides were found by MS analysis. Without specifying the cysteine content, the score was 94 with 44% coverage (12 peptides corresponded to the considered protein). When cysteine information was added, the score reached 118 (Fig. 7b). This time only

11 peptides were matched, due to the fact that the peptide at m/z 1693.5 ($[M+H]^+$) exhibited only one tag instead of two (m/z 2025.5 ($[M+2\text{tags}+H]^+$)) because of the limitation of the mass spectrometer to $m/z < 2000$. As this peptide contained in fact two cysteine residues, Mascot did not identify it as belonging to the protein of interest. Taking into account two cysteines in this peptide for the Mascot search, the score would have reached 147.

When the only six tagged peptides were entered, the score was 98 (16% coverage). Only four peptides were found. One peptide at m/z 1693.5 ($[M+H]^+$) was tagged once instead of twice (mass range limitation of the apparatus). It was not taken into account by the software. The other one was generated by two miscleavages and was excluded by Mascot (1 maximum miscleavage was specified for the search). As a conclusion, it illustrates the fact that cysteine content knowledge is a powerful additional information in database queries.

3.4 Mass fingerprinting of a mixture of four proteins using on-line electrochemical tagging

BSA, human α -lactalbumin, ovalbumin, and β -lactoglobulin from bovine milk were submitted to reduction and proteolytic digestion followed by RP-HPLC separation (Fig. 6c). The identification search was performed using ProFound [31], which allows searching for a protein mixture. The taxonomy Chordata was chosen for the search. Without tagging information, the mixture of four proteins was found with a probability of 53%. When cysteine information was specified in extra setting function, the probability reached 74% and the discrimination compared to other possibilities was enhanced (Fig. 7c). Moreover, the time search was improved (2.52 s and 1.56 s without and with tagging information, respectively). For ovalbumin, six peptides were matched without being tagged. For β -lactoglobulin, one cysteine residue was detected whereas in BSA, eight cysteines were found and seven in α -lactalbumin. The peptides, which were expected to be tagged because of their cysteine content, were labelled in all cases.

Another interesting characteristic is that the method potentially allows differentiation of isobaric peptides. For instance, the peptide at m/z 933.3 ($[M+H]^+$) could be attributed to both β -lactoglobulin and α -lactalbumin based on mass measurement only at the resolution of the ion trap. With the additional information on cysteine content, it could be unambiguously attributed to α -lactalbumin. As another example, the MS signal situated at m/z 701.3 ($[M+H]^+$) could be attributed to two peptides by mass

alone, one coming from β -lactoglobulin and the other one from BSA which contains one cysteine. In that case, the information on cysteine content was found as well to allow a correct attribution.

4 Concluding remarks

The successful on-line electrotagging of several cysteine residues in peptides was possible with 2-carboxymethylhydroquinone whereas hydroquinone and substituted hydroquinones, bearing an electron-donating group, were not efficient enough. Peptides containing up to three cysteine residues were successfully tagged with 2-carboxymethylhydroquinone, even if two cysteine moieties were consecutive. The counting of cysteine residues was directly obtained from one experiment through the direct observation of untagged and tagged products in a single mass spectrum.

The promising results of the tagging of synthetic peptides let apply the electrochemical tagging method to the identification of proteins through their proteolytic digest by peptide mass fingerprinting. The procedure is simple and efficient and does not cause loss of sample. The electrotagging is a quasi-instantaneous reaction contrary to any off-line alkylation step. Several proteins and a model mixture of four proteins were successfully identified by the determination of the peptide cysteine content by MS electrotagging. Even if the original peptide signal peak is divided according to the number of tags (by three at the maximum for a peptide containing two cysteine units for instance), the convenient information on cysteine content is greatly valuable. We showed that the identification score is improved compared to the case of untagged or iodoacetamide-alkylated digest samples. The discrimination with other candidates is also enhanced and the search time is reduced. The procedure is capable of enhancing mass spectrometric peptide mapping strategies.

The model experiments, despite the large amount of starting material, are quite realistic since the analyzed HPLC fractions, off-line collected, present peptide concentrations finally delivered to the MS comparable to those which would have been obtained from typical proteomics nanoLC procedures. However, the experiments described above were restricted to protein identification through peptide mass fingerprinting, which limits the approach to very simple protein mixtures, because of instrumental limitations: no on-line capillary or nanoLC system nor query tools (such as SEQUEST) were available to identify proteins from more complex mixtures in a true shotgun approach.

To push the technology a step further, the tag influence on the MS/MS sequencing will be investigated for gel-free bottom-up application. In any case, as the unmodified peptide is still present, MS/MS experiments could be performed on it as well as on the noncysteinyll-containing peptides. The method has to be validated in a real shotgun approach with adequate analysis system in order to study complex protein mixtures. In particular, the power of the approach has to be explored in conjunction with automated data-dependent MS/MS analysis, and should prove particularly useful for peptides that are not sequenced by MS/MS (two or three most abundant peptides in each spectrum).

The automation of the technique will be envisaged. It should be realized by the constant infusion of 2-carboxymethylhydroquinone (which was shown to be stable in solution [27]) by a T junction between the LC separation step and the microESI emitter. The tagging efficiency control can be adapted through the chip parameters or by the spray current regulation [21, 25–27]. Bioinformatics tools well adapted to the on-line electrotagging method could be also developed.

Catherine Servis is thanked for peptide synthesis. Tatiana C. Rohner, Eugen Damoc, and Henrik Jensen are thanked for fruitful discussions. Alain Razaname is thanked for technical support and the machine shop at EPFL for technical assistance. The EU is acknowledged for financial support through the project "MicroProteoMicS" (grant No. QL2-CT-2001-01903/OFES 01.0182-2).

Received May 28, 2004

5 References

- [1] Aebersold, R., Mann, M., *Nature* 2003, 422, 198–207.
- [2] Patterson, S. D., Aebersold, R. H., *Nature Genet.* 2003, 33, 311–323.
- [3] Ashcroft, A. E., *Nat. Prod. Rep.* 2003, 20, 202–215.
- [4] Gevaert, K., Vandekerckhove, J., *Electrophoresis* 2000, 21, 1145–1154.
- [5] Regnier, F. E., Riggs, L., Zhang, R. J., Xiong, L., Liu, P. R., Chakraborty, A., Seeley, E., Sioma, C., Thompson, R. A., *J. Mass Spectrom.* 2002, 37, 133–145.
- [6] Wolters, D. A., Washburn, M. P., Yates III, J. R., *Anal. Chem.* 2001, 73, 5683–5690.
- [7] Washburn, M. P., Wolters, D., Yates III, J. R., *Nat. Biotechnol.* 2001, 19, 242–247.
- [8] Fenyo, D., *Curr. Opin. Biotechnol.* 2000, 11, 391–395.
- [9] Sechi, S., Chait, B. T., *Anal. Chem.* 1998, 70, 5150–5158.
- [10] Zhang, H., Yan, W., Aebersold, R., *Curr. Opin. Chem. Biol.* 2004, 8, 66–75.
- [11] Hamdan, M., Righetti, P. G., *Mass Spectrom. Rev.* 2002, 21, 287–302.
- [12] Brune, D. C., *Anal. Biochem.* 1992, 207, 285–290.
- [13] Jue, R. A., Hale, J. E., *Anal. Biochem.* 1993, 210, 39–44.
- [14] Lundell, N., Schreitmuller, T., *Anal. Biochem.* 1999, 266, 31–47.
- [15] Herbert, B., Galvani, M., Hamdan, M., Olivieri, E., MacCarthy, J., Pedersen, S., Righetti, P. G., *Electrophoresis* 2001, 22, 2046–2057.
- [16] Adamczyk, M., Gebler, J. C., Wu, J., *Rapid Commun. Mass Spectrom.* 1999, 13, 1813–1817.
- [17] Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., Aebersold, R., *Nat. Biotechnol.* 1999, 17, 994–999.
- [18] Goodlett, D. R., Bruce, J. E., Anderson, G. A., Rist, B., Pasa-Tolic, L., Fiehn, O., Smith, R. D., Aebersold, R., *Anal. Chem.* 2000, 72, 1112–1118.
- [19] Galvani, M., Hamdan, M., Herbert, B., Righetti, P. G., *Electrophoresis* 2001, 22, 2058–2065.
- [20] Rohner, T. C., Rossier, J. S., Girault, H. H., *Anal. Chem.* 2001, 73, 5353–5357.
- [21] Rohner, T. C., Rossier, J. S., Girault, H. H., *Electrochem. Commun.* 2002, 4, 695–700.
- [22] Vanberkel, G. J., Zhou, F. M., *Anal. Chem.* 1995, 67, 2916–2923.
- [23] Zhou, F. M., Vanberkel, G. J., *Anal. Chem.* 1995, 67, 3643–3649.
- [24] Hignett, G., Threlfell, S., Wain, A. J., Lawrence, N. S., Wilkins, S. J., Davis, J., Compton, R. G., Cardosi, M. F., *Analyst* 2001, 126, 353–357.
- [25] Roussel, C., Rohner, T. C., Jensen, H., Girault, H. H., *ChemPhysChem* 2003, 4, 200–206.
- [26] Rohner, T. C., Josserand, J., Jensen, H., Girault, H. H., *Anal. Chem.* 2003, 75, 2065–2074.
- [27] Roussel, C., Dayon, L., Jensen, H., Girault, H. H., *J. Electroanal. Chem.* 2004, 570, 187–199.
- [28] Perkins, D. N., Pappin, D. J. C., Creasy, D. M., Cottrell, J. S., *Electrophoresis* 1999, 20, 3551–3567.
- [29] http://www.matrixscience.com/search_form_select.html
- [30] Zhang, W. Z., Chait, B. T., *Anal. Chem.* 2000, 72, 2482–2489.
- [31] http://prowl.rockefeller.edu/profound_bin/WebProFound.exe
- [32] <http://www.expasy.org/sprot/>
- [33] Biemann, K., *Methods Enzymol.* 1990, 193, 886–887.
- [34] Mason, D. E., Liebler, D. C., *Chem. Res. Toxicol.* 2000, 13, 976–982.
- [35] Silverstein, R. M., Basler, G. C., Morill, T. C., *Spectrometric Identification of Organic Compounds*, John Wiley & Sons, New York 1991.