Susanne Müller-Bomke, Michael Sperling, Heiko Hayen[a] and Uwe Karst*

Biolabeling with cobaltocinium tags

https://doi.org/10.1515/znb-2018-0093
Received May 20, 2018; accepted July 19, 2018

Abstract: A label for amino and thiol functionalities of peptides and proteins based on the activated cobaltocinium hexafluorophosphate succinimide ester (CoS) is presented. Despite the known selectivity of a succinimide ester towards amines, CoS also modifies cysteine residues under the same reaction conditions. The derivatized biomolecules were investigated using liquid chromatography with subsequent electrospray-mass spectrometric detection (LC/ESI-MS). In combination with their remarkable stability under physiological conditions, easy handling and good spectroscopic properties, cobaltocinium ions provide all requirements for a powerful labeling reagent. Furthermore, in direct comparison to the isoelectronic well-established ferrocene reagents, the higher redox potential and the chemical stability of the cobaltocinium moiety add to the benefits as a derivatizing agent for bioanalysis.

Keywords: amines; biolabeling; cobaltocinium; electrospray ionization-mass spectrometry (ESI-MS); thiols.

Dedicated to: Professor Bernt Krebs on the occasion of his 80th birthday.

1 Introduction

Peptides and proteins are frequent targets for chemical modification by covalent conjugation. In general, labels can be used to address various fundamental biological questions: particularly regarding diagnostics and therapeutics, selective and efficient protein labeling techniques are required [1–5]. In the past, this labeling approach was often based on the use of transition metal organometallic complexes [6–12]. They were employed to label functional groups in a biomolecule, typically the ε-amino function of the lysine residue and the thiol function of a cysteine residue.

In this work, a cobaltocinium tag is introduced as a new promising labeling tag. For this purpose, the activated cobaltocinium succinimide ester (CoS) with its counterion hexafluorophosphate (PF_6^-) was used. Therefore, the reaction of CoS with several biomolecules including peptides and proteins was in the focus of interest. Cobaltocinium ions are isoelectronic with ferrocenes, thus providing high stability even against strong oxidizing agents, such as fuming nitric acid, potassium permanganate or ozone [13, 14]. Hence, it is not surprising that cobaltocinium salts were proposed early on as haptens and tracers in biological systems [15]. Metzler-Nolte et al. described the successful synthesis of a cobaltocinium ion conjugate with an antigen nuclear localization signal by solid-phase peptide synthesis, which specifically delivers the organometallic species into the cell nucleus [16]. Thereby, the much higher redox potential and the chemical stability of the cobaltocinium ion over the ferrocene moiety were used for enhancing the cellular uptake of bioconjugates. More detailed information on the synthesis of cobaltocinium-peptide bioconjugates prepared by solid phase peptide synthesis has also been described by Metzler-Nolte et al. [17, 18].

The potential for biological assays was discussed when Moise et al. synthesized metallohaptens of the activated cobaltocinium ester with psychostimulating drugs like amphetamine or desipramine [19]. Moreover, the inherent positive charge of the cobaltocinium moiety is beneficial. Degrand et al. labeled two antiepileptic drugs, phenobarbital and phenytoin, with a cobaltocinium salt and proposed that the presence of the positive charge on the two drugs increases their hydrophilicity at the expense of their hydrophobic character [20]. This effect should prevent non-specific interactions with the proteins present in sera or antisera. Moreover, the influence of a positive charge on the structure and stability of peptides, stabilized by hydrogen bonds, has been investigated by van Staveren et al. [21], whereby a positive charge introduced by a cobaltocinium tag instead of a ferrocene one does not significantly influence hydrogen bonds that stabilize a peptide coil structure.

The synthesis of cobaltocinium-peptide nucleic acids (PNA) has been reported in literature [22]. Here,

[a]This author's part of the work was carried out at Leibniz-Institut für Analytische Wissenschaften – ISAS – e.V., Otto-Hahn-Straße 6b, 44227 Dortmund, Germany
***Corresponding author: Uwe Karst,** Westfälische Wilhelms-Universität Münster, Institute of Inorganic and Analytical Chemistry, Corrensstraße 30, 48149 Münster, Germany,
e-mail: uk@uni-muenster.de
Susanne Müller-Bomke, Michael Sperling and Heiko Hayen: Westfälische Wilhelms-Universität Münster, Institute of Inorganic and Analytical Chemistry, Corrensstraße 30, 48149 Münster, Germany

Brought to you by | York University Libraries
Authenticated
Download Date | 12/4/18 9:33 AM

ferrocene-PNA conjugates are compared to the isostructural but positively charged cobaltocinium derivatives to avoid solubility problems associated with the lipophilic ferrocene group. Indeed, the cobaltocinium-PNA conjugate showed excellent chemical stability upon storage in solution. Together with an amide, the positive charge of the cobaltocinium unit has been found to be the essential component for molecular and electrochemical recognition of anionic guest species [23, 24]. In combination with their extraordinary stability under physiological conditions, the easy handling and good spectroscopic properties, cobaltocinium ions provide excellent features for a powerful labeling reagent.

No attempt has been made yet to use CoS as a labeling reagent for peptides and proteins and subsequent analysis of the derivatized biomolecules by liquid chromatography with electrospray ionization-mass spectrometric detection (LC/ESI-MS). In this paper, the investigation of CoS including the optimization of labeling conditions as well as the investigation of the selectivity of CoS towards individual amino acids and functional groups are described. Additionally, CoS is used as a labeling reagent for peptides and proteins. The modification of the functional groups of the acidic peptide hormone insulin and the basic model protein ß-lactoglobulin A with CoS and their subsequent mass spectrometric analysis is presented.

2 Experimental

2.1 Chemicals

ß-Lactoglobulin A (ß-LGA) (from bovine milk), insulin (from bovine pancreas), trypsin (from bovine pancreas), octylamine, L-lysine, boric acid, dimethylcyclopentadiene, cobalt(II) bromide and *N*,*N*′-dicyclohexylcarbodiimide were purchased from Sigma Aldrich Chemie GmbH (Steinheim, Germany). Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) was obtained from ABCR GmbH & Co. KG (Karlsruhe, Germany). L-Cysteine, dicyclopentadiene, pyrrolidine and sodium hexafluorophosphate were ordered from Acros Organics (Geel, Belgium). L-Histidine, *N*-hydroxysuccinimide, piperidine and pyrrolidine were obtained from Merck KGaA (Darmstadt, Germany). Urea and formic acid were purchased from Fluka Chemie GmbH (Buchs, Switzerland). PD-10 Sephadex G-25 M desalting columns for protein purification were purchased from GE Healthcare (Buckinghamshire, UK). All chemicals were ordered in the highest purity available and were used without further purification. Water was purified using a Milli-Q Gradient A 10 system and filtered through a 0.22 µm Millipak 40 filter (Millipore, Billerica, MA, USA).

2.2 LC/ESI-MS

For LC/MS measurements of derivatized biomolecules, a HPLC system from Shimadzu (Duisburg, Germany) and a triple-quadrupole ion trap (QTRAP) mass spectrometer (Applied Biosystems, Darmstadt, Germany), equipped with a Turbo IonSpray (pneumatically assisted ESI) source were used. The LC system consisted of two LC-10ADVP pumps, a DGC-14A degasser, a SIL-HTA autosampler, a CTO-10AVP column oven and a SPD-10AVVP UV detector. The software used for controlling was Analyst 1.4.1 (Applied Biosystems, Darmstadt, Germany). On the QTRAP instrument, (+)-ESI-MS was performed either in the full-scan mode ($m/z = 600–1700$, integration time 1 s), in the selected ion monitoring (SIM) mode or in MS/MS (precursor ion scan) mode. The analytes were ionized in the ESI interface with an ion spray voltage of 4200 V, using 50 psi nebulizer gas (N_2), 25 psi curtain gas (N_2), and 50 psi dry gas (N_2) with a temperature of 400°C. In SIM mode, collision gas (N_2) was set to low, declustering potential to 60 V, entrance potential to 12 V, collision cell entrance potential to 19 V, and collision energy to 30 eV. For MS/MS experiments, collision gas was set to high, declustering potential to 70 V, entrance potential to 12 V, collision cell entrance potential to 16 V, and collision energy to 50 eV. These conditions were optimized for CoS and its derivatives under HPLC conditions and were used for all further LC/MS measurements.

2.3 FTICR-MS

Labeled intact proteins were analyzed by direct-infusion nanospray and ultra-high resolution-mass spectrometric detection. The hybrid linear quadrupole ion trap-Fourier transform ion cyclotron resonance (FTICR) mass spectrometer (LTQ FT, Thermo Scientific, Bremen, Germany) was equipped with a 7.0 Tesla actively shielded superconducting magnet and a nanospray ionization source. The nanospray capillaries (tip inner diameter (i.d.) ~ 5 µm) were from MasCom (Bremen, Germany). The instrument was operated in the positive ion mode and the mass resolution was set to 400.000 (at $m/z = 400$, FWHM). The spectra were recorded under the following conditions: source voltage 1.2–1.5 kV, capillary voltage 40 V, capillary temperature 250°C, and tube-lens voltage 250 V.

Brought to you by | York University Libraries
Authenticated
Download Date | 12/4/18 9:33 AM

2.4 LC/FTICR-MS

The tryptic peptide mixtures were separated by means of HPLC on a Surveyor HPLC system consisting of a Surveyor MS pump and Surveyor autosampler (Thermo Scientific, San Jose, CA, USA). For MS and MS/MS experiments, the hybrid mass spectrometer was operated in the data-dependent mode, analogous to a previously described method [25]. Briefly, the instrument was set to automatically switch between MS and MS/MS acquisition. Survey mass spectra in the mass range $m/z = 300–1500$ were acquired in the FTICR. The three most intense ions were sequentially isolated for accurate mass measurements by an FTICR "SIM scan", in a narrow mass window. Subsequent fragmentation was carried out in the linear ion trap by collision-induced dissociation. The general mass spectrometric conditions were: spray voltage, 3.5 kV; sheath gas flow, 30 (arbitrary units); auxiliary gas flow, 5 (arbitrary units); and sweep gas flow, 2 (arbitrary units); ion-transfer tube temperature, 275°C; normalized collision energy at 35% for MS/MS.

2.5 ESI-TOF-MS

For the additional determination of accurate masses, a Bruker microTOF mass spectrometer (Bruker Daltonics, Bremen, Germany) was used. The instrument was controlled by microTOFControl v.1.1. The flow-injection analysis was performed using a syringe pump model 74900 (Cole Parmer, Vernon Hills, IL, USA). Full-scan spectra were recorded using (+)-ESI-MS under the following conditions: end plate offset: –500 V, capillary: –4000 V, nebulizer gas (N_2): 0.5 bar, drying gas (N_2): 4 L min^{-1}, drying temperature: 180°C, skimmer 1: 70 V, skimmer 2: 25 V, capillary exit: 210 V, hexapole 1: 23.0 V, hexapole 2: 21.6 V, hexapole RF: 300 V, transfer time: 70 µs, pre-pulse storage: 22 µs, detector: 1300 V. Prior to each measurement sequence, external mass calibration was carried out using sodium-formate clusters.

2.6 HPLC separation

Conditions A. For the determination of the limits of detection (LOD) and the limits of quantification (LOQ), a SeQuant® ZIC®-HILIC column (Merck, Darmstadt, Germany) with the following dimensions was employed: 150 mm × 2.1 mm i.d., particle size 5 µm. In this case, an isocratic elution with 65% aqueous acetonitrile at a flow rate of 0.3 mL min^{-1} and an analysis time of 15 min was applied.

Conditions B. Separation of analytes was carried out using a Discovery® C8 column (Supelco, Taufkirchen, Germany) with the following dimensions: 150 mm length × 2.1 mm i.d., 5 µm particle size. To prevent contamination of the analytical column, a guard column with the following dimensions was used: 20 mm × 2.1 mm i.d., particle size of 5 µm. The flow rate of the mobile phase was 0.3 mL min^{-1}. The column was operated at ambient temperature. The injection volume was 10 µL. For all separations, solvent A of the mobile phase was 0.1% formic acid in deionized water, solvent B was acetonitrile. The following gradient was applied: after one min. with an isocratic composition of 20% organic mobile phase, the acetonitrile content was increased to 90% within 8 min and held at 90% for 1 min. Finally, the system was equilibrated for 3 min.

Conditions C. The separation of tryptic peptides was carried out using a SeQuant® ZIC®-HILIC column (150 mm length × 2.1 mm i.d., particle size 3.5 µm) from Merck equipped with a guard column (5 mm × 1 mm i.d., particle size 5 µm). The injection volume was set to 2 µL. The following binary gradient was used at a flow rate of 75 µL min^{-1}: an isocratic step of 10% solvent B for 2 min was followed by an increase from 10 to 80% in 36 min. Then, the composition of 80% B was held constant for 7 min and a decrease from 80 to 10% in 2 min followed. Finally, the column was re-equilibrated at 10% with solvent B for 13 min. Mobile phase A contained 5% water and 95% acetonitrile (*v/v*) with 10 mM ammonium acetate-acetic acid (pH 4.8). Mobile phase B consisted of 5% acetonitrile and 95% H_2O (*v/v*) with 10 mM ammonium acetate-acetic acid (pH 4.8).

2.7 Synthesis of the reagents

2.7.1 Cobaltocinium hexafluorophosphate succinimide ester (CoS)

The synthesis of CoS, starting from 1-cobaltociniumcarboxylic acid hexafluorophosphate, was performed according to reference [19] with slight modifications. 1-Cobaltociniumcarboxylic acid ($CoCp_2COOH$) was synthesized as reported by Sheats et al. [13]: 300 mg of $CoCp_2COOH$ (0.79 mmol) was dissolved in 25 mL acetone p.a. and treated with 91.2 mg (0.8 mmol) *N*-hydroxysuccinimide and 160 mg (0.79 mmol) dicyclohexylcarbodiimide. The suspension was stirred for 24 h in the dark. Afterwards, the precipitated dicyclohexylurea was filtered off and the solvent was evaporated *in vacuo*. The yellow solid was washed with dry pentane and recrystallized

Brought to you by | York University Libraries
Authenticated
Download Date | 12/4/18 9:33 AM

in toluene/hexane (1:1 (v/v)). The obtained yellow product was dried *in vacuo*. Yield: 262.2 mg (70%). – MS (TOF): $m/z = 330.0168$ $[M]^+$, calcd. $m/z = 330.0171$ for $[CoC_{15}H_{13}O_4N]^+$. – ^{1}H NMR (ppm, 200 MHz, CD_3CN): $\delta = 4.19$ (s, 9H, Cp, Cp*), 2.92 (s, 4H, 8H, H9), 2.71 (s, 4H, 6H, 7H).

2.7.2 Cobaltocinium hexafluorophosphate

A volume of 12.5 mL (0.15 mol) freshly cracked cyclopentadiene was dissolved in 50 mL dried pyrrolidine. The reaction mixture was cooled down to 0°C and 16.40 g (0.075 mol) $CoBr_2$ was added in small portions. The solution was allowed to warm up to room temperature and stirred for further 12 h. The solvent was evaporated *in vacuo* and the residue was dissolved in 250 mL hot water and clarified with 5 g Norit®. After the dropwise addition of a solution containing 3.25 g of sodium hexafluorophosphate in 25 mL of water, the solution was concentrated and stored for 12 h at 4°C. The resulting yellow precipitate was dried *in vacuo*. Yield: 8.1 g (0.042 mol, 57%). MS (TOF): $m/z = 189.0110$ $[M]^+$, calcd. $m/z = 189.0109$ for $[CoC_{10}H_{10}]^+$. – ^{1}H NMR (ppm, 200 MHz, CD_3CN): $\delta = 5.72$ (s, 10H, Cp*). – ^{13}C NMR (ppm, 200 MHz, CD_3CN): $\delta = 86.52$ (s, 10C, Cp).

2.8 Examination of hydrolysis

The hydrolysis of CoS to the corresponding cobaltocinium-carboxylic acid ($CoCp_2COOH$) was investigated for different ratios of organic to aqueous solvent. A stock solution of CoS with a concentration of 10^{-2} M was prepared in acetonitrile. Different volumes of borate buffer (20 mM, pH 9) were added resulting in final ratios of organic to aqueous solvent between 1:1, 2:1, 5:1 and 10:1. The final concentration of CoS was kept constant at 10^{-3} M. The course of hydrolysis was analyzed for 15 min every 90 s via flow-injection (FI)-LC/ESI-MS monitoring the analytes of interest in the SIM mode.

2.9 Examination of low-molecular weight amines and thiols

Stock solutions of the amines and thiols of interest with concentrations between 10^{-3} M and 10^{-6} M of each analyte were prepared by dilution of the respective amount of solution with borate buffer (20 mM, pH 9). CoS was dissolved in acetonitrile to obtain a solution with a concentration of 10^{-3} M. The respective amine or thiol (100 µL) and 100 µL of CoS were mixed in a vial. After a reaction time of 20 min, the reaction products were analyzed via LC/ESI-MS.

2.10 Derivatization of peptides and proteins

The peptide hormone insulin (51 amino acids) and the protein ß-LGA(162 amino acids) were dissolved in urea solution [7 M, in borate buffer (20 mM, pH 9)] to form stock solutions with concentrations between 0.125 and 0.200 mmol L^{-1}. The derivatization was carried out using a 10-fold molar excess of CoS (10 mM, in acetonitrile). After stirring the solution for 30 min, the excess of derivatizing agent was removed by passing the reaction mixture through a PD-10 desalting column according to the supplier's protocol, using 0.01% formic acid as the mobile phase. The experiments were repeated, but a reduction of the intramolecular disulfide bonds was performed prior to the derivatization with CoS. Therefore, the protein was treated for 60 min at ambient temperature with a ninefold molar excess of tris(2-carboxyethyl)phosphine hydrochloride (TCEP, 100 mM, in 0.3 M aqueous ammonia solution) with respect to the thiol groups available. All reaction products were diluted 1:1 with ammonium acetate buffer (100 mM, pH 4.8)-methanol (1:1) before they were analyzed by means of LC/ESI-MS.

2.11 Tryptic digest

Before the tryptic digest of the proteins could be performed, the intramolecular disulfide bonds had to be reduced. The reduction of disulfide bridges was performed for 30 min at ambient temperature using a ninefold molar excess of TCEP (100 mM, in 0.3 M aqueous ammonia solution) with respect to the thiol groups available. Nine hundred and sixty milligram of urea were added to 2 mL of the protein and the calculated amount of TCEP was adjoined. The reduction was terminated by passing the mixture through a PD-10 desalting column. Digestion with trypsin was performed by adding 5 µL of a stock solution of 2.3 mg trypsin in 1 mL water to 1 mL of the protein solution containing approximately 5.0×10^{-5} M of reduced protein and 50 mM ammonium bicarbonate. After incubating the mixture for 2 h at 37°C, the digest was terminated by adding 200 µL of ice-cold acetonitrile. The resulting tryptic peptides were subsequently analyzed by means of FI-TOF-MS.

Brought to you by | York University Libraries
Authenticated
Download Date | 12/4/18 9:33 AM

3 Results and discussion

3.1 Optimization of the labeling procedure

The activated cobaltocinium succinimide ester selectively reacts with amines under basic conditions. Furthermore, CoS may react with free thiol groups to form stable thioether bonds (Fig. 1). As cobaltocinium ions can be synthesized from the respective 19 electron compound, the cobaltocene, under aqueous acidic conditions, CoS was synthesized starting from 1-cobaltociniumcarboxylic acid and *N*-hydroxysuccinimide. Due to the moderate solubility in both water and organic solvents, hexafluorophosphate was selected as the counterion.

The limits of detection (LOD) of cobaltocinium derivatives for LC/ESI-MS are lower compared to the ferrocene analogues. Obviously, the good mass spectrometric response of this compound class is based on the inherent positive charge of the cobaltocinium unit. The LOD was exemplarily investigated for cobaltocinium hexafluorophosphate $[Co(Cp)_2]PF_6$, which could be obtained without any impurities from freshly cracked cyclopentadiene, cobalt(II) bromide and sodium hexafluorophosphate. For the determination of LOD and LOQ, hydrophilic interaction liquid chromatography (HILIC, see HPLC conditions A) was used prior to the mass spectrometric detection, as only this zwitterionic stationary phase established a good retention for the $[CoCp_2]^+$ cation. The LOQ, reasonably defined as three times the LOD, was 5.0×10^{-9} M, whereas the LOQ could be defined as 1.5×10^{-8} M under optimized MS conditions even on an instrument of an earlier generation.

Initially, the derivatization of CoS with low-molecular weight amines (lysine, octylamine) was performed to optimize the reaction conditions according to the excess of derivatizing agent, the suitable buffer system, pH value and reaction time. The rate of the hydrolysis reaction of CoS to the corresponding 1-cobaltociniumcarboxylic acid ($CoCp_2COOH$, $m/z = 233.0013$ $[M]^+$, calcd. $m/z = 233.0007$ for $[CoC_{11}H_{10}O_2]^+$) is dependent of the solvent composition, regarding the ratio of organic to aqueous part. With equal volumes of acetonitrile and aqueous borate buffer (1:1), hydrolysis is completed within 10 min. In acetonitrile, CoS is stable over several weeks and no hydrolysis product could be identified. Under aqueous conditions, the derivatization reaction with amino functions is significantly faster than the hydrolysis of the derivatizing agent.

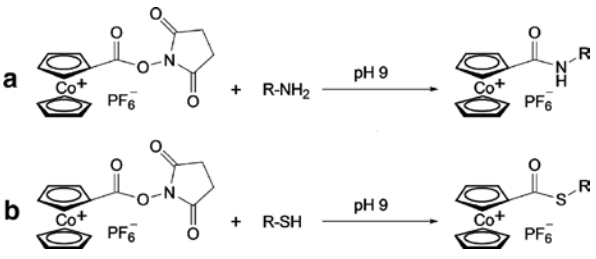

Fig. 1: Reaction of CoS with (a) an amine to the corresponding amide and with (b) a thiol to the corresponding thioether.

With a 10-fold molar excess of the derivatizing agent, only few lysine residues were labeled. CoS is assumed not to react with the α-amino group of the amino acid lysine. Labeling is more likely at the ε-amino group because it is the more basic moiety with a higher pK_a value of 10.3 compared to 8.9 for the α-amino group. The determination of the *m/z* value enables the calculation of molecular formulas with a deviation between the measured and calculated exact mass for the lysine derivative of −0.6 ppm (determined: 361.0955 $[M]^+$, calcd. $m/z = 361.0957$ for $[CoC_{17}H_{22}O_3N_2]^+$). An increase of the used molar excess of CoS up to the 50 fold with respect to the free amino groups available does not have any influence on the degree of labeling. A reaction of CoS with secondary amino groups, e.g. as part of the basic amino acid L-histidine (pK_a 9.0 of α-amino group), was not observed.

With respect to the optimum pH value for the derivatization reaction, the degree of labeling of amino functionalities with CoS was found to increase with the pH of the solution. For these investigations, the reaction of CoS with lysine and octylamine ($pK_a = 10.7$) was performed in a borate buffer-acetonitrile mixture (1:1, *v/v*) from pH 7 up to pH 11 in steps of 0.5 pH units. Fewer amino functions are protonated when the pH increases. This coincides with the experimental results, because the highest transformation rates of lysine and octylamine with CoS were achieved under basic conditions, which corresponds to the chemistry of succinimide esters known from literature [26]. However, when applying a pH above 9, hydrolysis is increasing. Therefore, the optimum pH for this reaction is 9, and for all further experiments, borate buffer (20 mM, pH 9) was used.

The model substance octylamine with a similar pK_a value compared to lysine (10.7 vs. 10.3) was used for further studies regarding the selection of the optimum excess of CoS over the number of free amino groups available. Octylamine is only slightly polar, thus the separation of derivatized octylamine could easily be achieved on a reversed-phase column (Fig. 2) (HPLC conditions B). The accurate mass of this CoS derivative ($m/z = 344.1415$; $[M]^+$), is in good agreement with the theoretical value (calcd. $m/z = 344.1419$ for $[CoC_{19}H_{27}ON]^+$). Whereas the octylamine derivative shows sufficient retention on a reversed-phase

Brought to you by | York University Libraries
Authenticated
Download Date | 12/4/18 9:33 AM

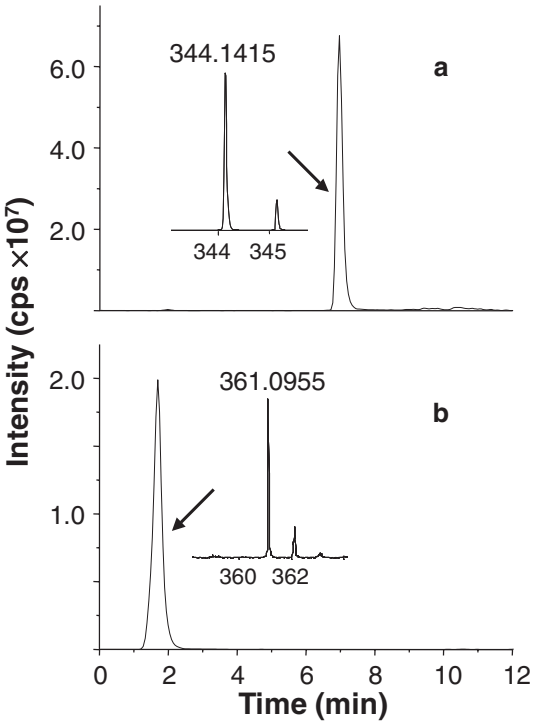

Fig. 2: LC chromatogram showing products of the reaction of octylamine with CoS (a) and of lysine with CoS (b). Additionally, the corresponding TOF-MS spectra are integrated and illustrate the exact masses of both derivatives. The separation was performed using reversed-phase conditions.

column, the lysine derivative ($m/z = 361.0955$ $[M]^+$, calcd. $m/z = 361.0957$ for $[CoC_{17}H_{22}O_3N_2]^+$) elutes within the void volume under reversed phase conditions.

All other cobaltocinium compounds including the precursors of the CoS synthesis ($Co(Cp)_2$, $CoCp_2CH_3$, $CoCp_2(CH_3)_2$, $CoCp_2(COOH)_2$) elute within the void volume as well. However, the positively charged cobaltocinium matrix can be separated by means of capillary electrophoresis (CE) or HILIC (see HPLC conditions A). The separation on a HILIC column was accomplished, and LODs and LOQs for the cobaltocinium salts in LC/ESI-MS were determined.

The derivatization of the model substance octylamine was tested with a 1-, 10-, 100- and 1000-fold molar excess of CoS relative to the free amino group. At first, the ratio of underivatized to derivatized amine was considered. In all cases, the ratio of underivatized to derivatized octylamine stayed constant after a reaction time of 10 min. Only if the ratio of underivatized to derivatized amine is constant, quantification of the analytes is possible. For the CoS-octylamine derivative, the largest intensities in the mass spectrum were observed using a 10-fold molar excess of derivatizing agent.

Typically, succinimide esters react selectively with amines under basic conditions (pH 9). However, reactions with thiols are not excluded under certain conditions (large excess of derivatizing agent, long reaction times) [6]. It was therefore surprising to observe a doubly derivatized cysteine derivative already under optimized reaction conditions for the amino function (10-fold molar excess of CoS, 20 min reaction time). Both the *N*-terminus and the thiol group were completely derivatized with CoS. The major signals in the mass spectra correspond to the singly derivatized cysteine ($m/z = 336.0103$ $[M]^+$, calcd. $m/z = 336.0099$ for $[CoC_{14}H_{15}O_3NS]^+$). In summary, a lower excess of CoS leads to higher transformation rates with amino and thiol groups.

3.2 MS/MS experiments

Further structural characterization of labeled amines was carried out by MS/MS applying collision-induced dissociation (CID). All derivatized amines show clear and similar fragmentation spectra. The fragmentation spectra of the CoS-lysine and CoS-cysteine derivatives are illustrated in Fig. 3. By means of the observed fragmentation pattern, cobaltocinium derivatives can easily be assigned since the fragment $m/z = 232.2$ ($CoCp_2CONH_2$) is always observed.

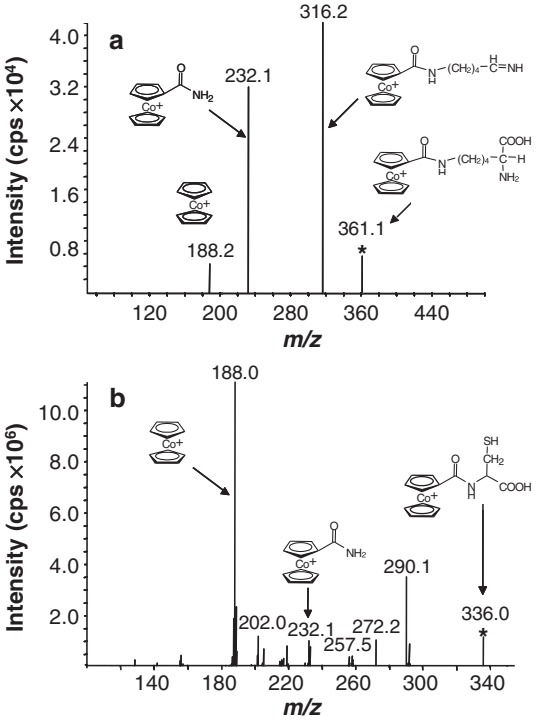

Fig. 3: ESI-MS/MS spectrum of singly derivatized lysine ($m/z = 361.1$) (a) and of singly derivatized cysteine ($m/z = 336.0$) (b).

Brought to you by | York University Libraries
Authenticated
Download Date | 12/4/18 9:33 AM

Additionally, the unsubstituted cobaltocinium ion ($CoCp_2$, $m/z = 188.2$) is visible in all mass spectra. These transitions offer good possibilities for the use of tandem mass spectrometry in quantitative analysis.

Fragmentation of the singly derivatized cysteine ($m/z = 336.0$, $[M+H]^+$) could not prove that the CoS label prefers the reaction with the thiol group over the *N*-terminus. The obtained fragmentation pattern was comparable to those observed for the fragmentation of the CoS-lysine derivative, because the fragment $m/z = 232.1$ ($CoCp_2CONH_2$) could be detected as well. However, this fragment will only be obtained if the label is directly located at the *N*-terminus of cysteine but will not occur if the thiol group is involved in the modification process. If the CoS-tag had been located at the thiol group, the fragment $m/z = 249.0$ ($CoCp_2COS$) would have been observed. In this case though, the main fragment in the spectrum is caused by the unsubstituted cobaltocinium ion ($m/z = 188.0$, $CoCp_2$). Further fragments could be assigned as follows: $m/z = 290.1$ $[M-CO_2]$, $m/z = 202.0$ $[CoCp_2CH_2]$, $m/z = 272.2$ $[M-CH_2CH_2SH]$, $m/z = 257.2$ $[M-CO_2-SH]$. Therefore, the preferential labeling position is the *N*-terminus of cysteine.

3.3 Derivatization of peptides and proteins

Furthermore, the chemical modification of peptides and proteins with CoS was investigated. The primary NH_2 group of lysine residues and the thiol group of cysteine residues should both be modified. Different model proteins like ß-LGA and insulin were treated with CoS. Initially, the reaction of CoS with the basic model protein ß-LGA (pI = 11.2) was investigated. ß-LGA contains 15 lysine residues, one *N*-terminus, two disulfide bridges and one free cysteine residue. Reactions with different molar excess of CoS with respect to the number of amino and thiol groups were tested. With equimolar amounts of CoS and ß-LGA, unlabeled protein could still be identified. When a 10-fold or even higher molar excess of CoS was used for modification, only the singly derivatized ß-LGA was identified and no unlabeled protein remained. Besides the singly derivatized ß-LGA, only signals of low intensity (<5%) could be observed which correspond to the doubly derivatized protein. In Fig. 4, the mass spectra of the unmodified (a) and the singly derivatized ß-LGA (b) detected by using nanospray FTICR-MS in the positive ion mode are compared.

Additionally, the calculated (Fig. 4, a2 and b2) and measured (Fig. 4, a1 and b1) isotopic patterns of the charge state +12 are presented. Due to the ultra-high resolution capabilities of the FTICR-MS, it is possible to resolve the 12+ charge state of the 18.4 kDa protein. For both unmodified and singly derivatized ß-LGA, the calculated and measured mass spectra are in good agreement. It can clearly be seen that the MS signals are shifted towards higher m/z values after reaction with CoS (a1 and b1), thus proving that a covalent CoS-protein adduct formation has occurred. No underivatized ß-LGA could be detected after derivatization with CoS. Deconvolution of the measured unmodified ß-LGA (a1) resulted in a molecular mass of the

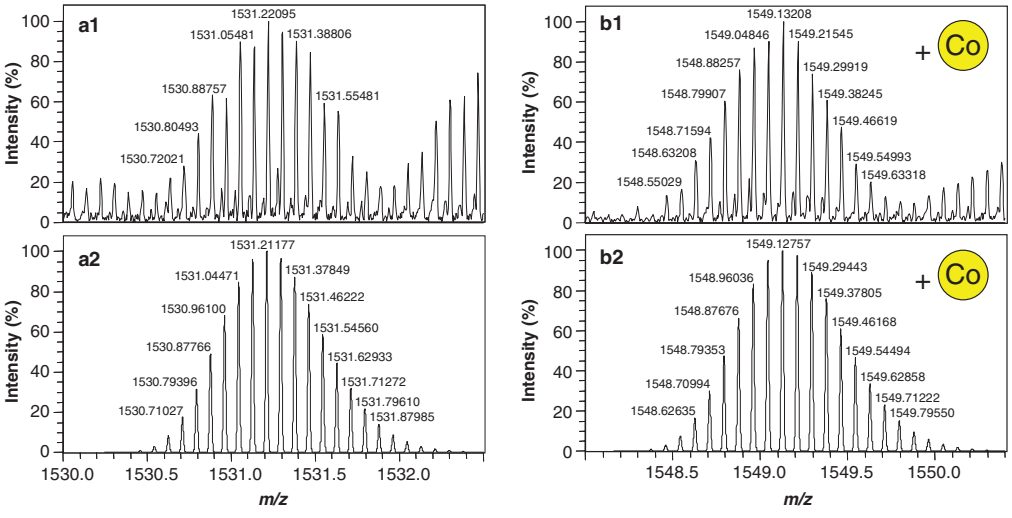

Fig. 4: FTICR-mass spectra of the +12 charge state of native (a) and derivatized (b) ß-LGA after reaction with CoS. Measured (a1, b1) and theoretical (a2, b2) isotope distributions are compared. The latter spectra were simulated at a resolution power of 110.000 (FWHM) for elemental compositions $C_{821}H_{1330}N_{206}O_{250}S_9$ and $C_{832}H_{1338}N_{206}O_{251}S_9Co$, respectively. The mass spectra of the derivatized protein were obtained after reaction with a 10-fold molar excess of derivatizing agent.

Brought to you by | York University Libraries
Authenticated
Download Date | 12/4/18 9:33 AM

protein of 18362.6 Da, which is in good accordance with the theoretical value of 18361.1 Da calculated from the amino acid sequence. After reaction with CoS, a deconvoluted mass of 18577.5 Da could be observed for the modified protein. The mass shift of 214.9 Da corresponds well with the addition of one CoS molecule ($CoC_{11}H_8O$, calculated monoisotopic mass 215.90).

In another experiment, the two intramolecular disulfide bonds of ß-LGA were reduced with TCEP. Now, five free cysteine residues are potential nucleophilic reaction partners for the electrophilic CoS reagent. With a 10-fold molar excess of CoS with respect to the free amino and thiol groups available, a distribution of the labeling degree was achieved (Fig. 5a). The highest signals in the mass spectrum correspond to the fivefold derivatized protein. An enlargement of the charge state $z=+18$ ($[M+18H]^{18+}$) shows that the less intense signals can be assigned to the three-, four-, and sixfold derivatized ß-LGA (Fig. 5b). Further peaks appearing in the mass spectrum are related to sodium adduct formation. The replacement of a H^+ by Na^+ results in a mass shift of $m/z=21.9819$ for a singly charged species, and hence 1.2179 for the 18+ charge state. The respective measured m/z shift for the fivefold labeled ß-LGA of 1.2112 (1070.2066–1068.9954) is in good agreement with the theoretical value.

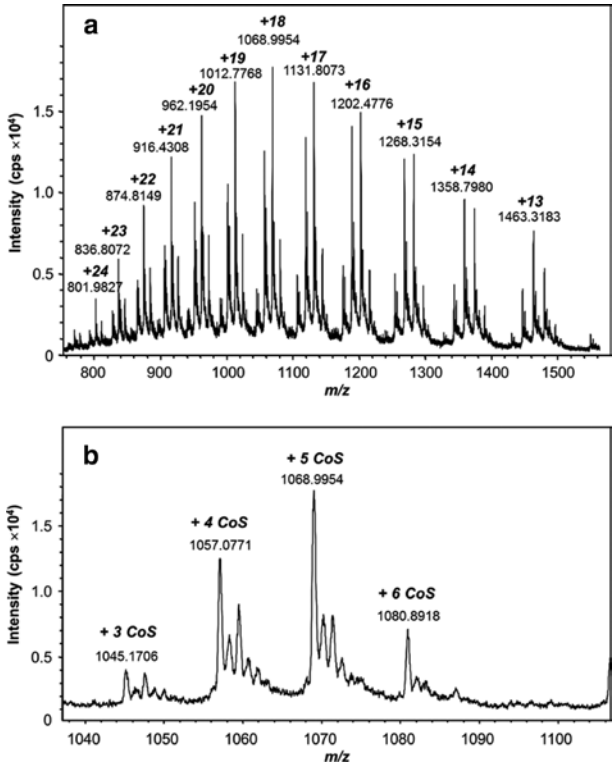

Fig. 5: (a) Time-of-flight mass spectrum of derivatized ß-lactoglobulin A after reduction of the two intramolecular disulfide bonds. (b) Close-up of the charge state $z=+18$.

Neither underivatized ß-LGA nor one- or twofold derivatized ß-LGA could be detected. The occurrence of a sixfold derivatized ß-LGA hints at partial modification of the *N*-terminus or one of the 15 free lysine residues in addition to almost quantitative modification of the free thiol groups.

To obtain more information about the modification site in ß-LGA and to answer the question whether the chemical modification is located at the *N*-terminus, at a lysine residue or at a free cysteine residue, a tryptic digest was carried out. The tryptic digest was performed for the following differently labeled analytes: (A) Singly derivatized ß-LGA (derivatization of ß-LGA with CoS was realized prior to the reduction of the two disulfide bridges), and (B): three- to sixfold derivatized ß-LGA (ß-LGA was obtained after reduction of the disulfide bridges). As ß-LGA contains in total 16 free amino groups (1 *N*-terminus and 15 free amino groups as being part of lysine residues) it is very likely that the CoS label for the singly derivatized ß-LGA is located either at the *N*-terminus or at the only free thiol group. ß-LGA contains one free cysteine residue, which is not bound in disulfide bridges. As could be shown for cysteine before under optimized reaction conditions, CoS quickly reacts with free thiol groups as well.

For the tryptic digest, urea and the excess of derivatizing agent were removed on a PD-10 desalting column. If the disulfide bonds were not reduced prior to the derivatization, they would be reduced afterwards using the reducing agent TCEP. Trypsin is a proteolytic enzyme, which cleaves peptides and proteins specifically at the *C*-terminal site of the amino acids lysine and arginine. The tryptic peptides can be analyzed with the help of mass spectrometric methods in more detail. In case of covalently modified proteins, some of the tryptic peptides carry the modification. Special attention must be paid to these digestion products. In this case, peptides bearing the free thiol group of cysteine, lysine residues and the *N*-terminus of ß-LGA were investigated. The T18 fragment ($m/z=829.9$) contains the amino acid residues 149–162 (LSFNPTQLEEQCHI), whereas a cysteine is located at position 160. The tryptic peptides were separated on a ZIC®-HILIC column under the conditions described above. The doubly charged T18 peptide modified with one CoS label ($[T18+CoS]^{2+}$) could be detected at $m/z=936.8848$ in both cases (ß-LGA + CoS and ß-LGA, reduced + CoS) in the positive ion mode using FTICR-MS.

To investigate the location of the CoS label in more detail, the peptides were fragmented by collision-induced dissociation in the linear-quadrupole ion trap of the hybrid

Brought to you by | York University Libraries
Authenticated
Download Date | 12/4/18 9:33 AM

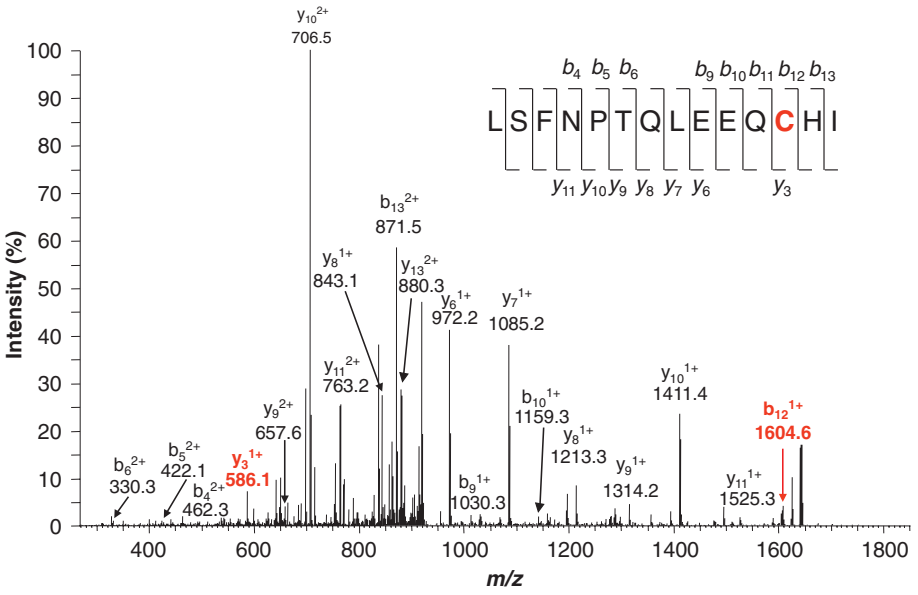

Fig. 6: Fragmentation of the T18 peptide at $m/z = 936.8848$ and assignment of *b*- and *y*-ion series. The fragment ions y_3^{1+} and b_{12}^{1+} containing the cobaltocinium tag are colored in red.

MS. For example, the peptide T18 at $m/z = 936.8848$ was fragmented in an MS/MS experiment. The *b*- and particularly the *y*-ion series are covered well. In the low-resolution MS/MS spectrum of this peptide (Fig. 6), the fragmentation product at $m/z = 586.1$ can be assigned to the y_3^{1+} ion modified with a CoS-label. Additionally, the b_{12}^{1+} ion at $m/z = 1604.6$ containing the CoS label was detected. Thus, the modification can be assigned to the cysteine residue C160. A modification of the secondary amino group of histidine (H) at the position 161 can be excluded.

Further peptides containing a cobaltocinium tag could be identified. For example, the T1 fragment bearing the CoS tag at the *N*-terminus and several tryptic peptides consisting of derivatized lysine residues were detected. Interestingly, the peptide T13, which contains the amino acid residues 102–124 (YLLF**C**MENSAEPEQSLV**C**Q-**C**LVR), could not be detected. However, the respective peptide containing one disulfide bridge was identified. In addition to the unmodified peptide ($m/z = 1337.1112$; calcd. $m/z = 1337.1126$ for $[C_{115}H_{183}O_{36}N_{29}S_4]^{2+}$) the singly labeled peptide was detected ($m/z = 1444.1062$; calcd. $m/z = 1444.1041$ for $[CoC_{126}H_{190}O_{37}N_{29}S_4]^{2+}$).

In summary, the tryptic digest demonstrates that the CoS is located at a cysteine residue, at the *N*-terminus and/or at lysine residues. These observations correspond to the results obtained from the analysis of derivatized free amino acids.

The reaction of the acidic peptide insulin (pI = 5.3) with CoS was investigated in the following. Insulin consists of an A and a B chain, which are both connected by two intramolecular disulfide bridges. Another disulfide bond is located within chain A. Moreover, the lysine residues are possible reaction partners for the derivatizing agent. Insulin contains two *N*-termini and one additional lysine residue within chain B. In the first step, insulin was treated with CoS without a preceding reduction of the disulfide bonds. As before, a 10-fold molar excess of CoS with respect to the free amino groups available was used. After purification, FI-ESI-TOF-MS measurements were performed. The mass spectra of the native insulin and its derivatives after the reaction with CoS are presented in Fig. 7. The determined accurate masses are in good agreement with the theoretical values. After the modification with CoS, singly, doubly and, at low intensity, triply derivatized insulin could be observed. The unmodified peptide could still be found in the mass spectra with an even 1.5-fold higher intensity than the singly labeled peptide.

In a second experiment, the derivatization was performed after reduction of the disulfide bonds. After the reduction, chain A and B had to be considered separately. Chain B contains two amino groups (one *N*-terminus and one free lysine residue) and two free cysteine residues. The A chain contains, apart from the *N*-terminus, no further amino functionalities from lysine residues, but four free thiol groups are available after the reduction with TCEP. In fact, after the derivatization with CoS, up to four labels were attached to the A chain. Taking the results of the derivatization with ß-LGA into account, a labeling of the four free cysteine residues can be assumed. However, the

Brought to you by | York University Libraries
Authenticated
Download Date | 12/4/18 9:33 AM

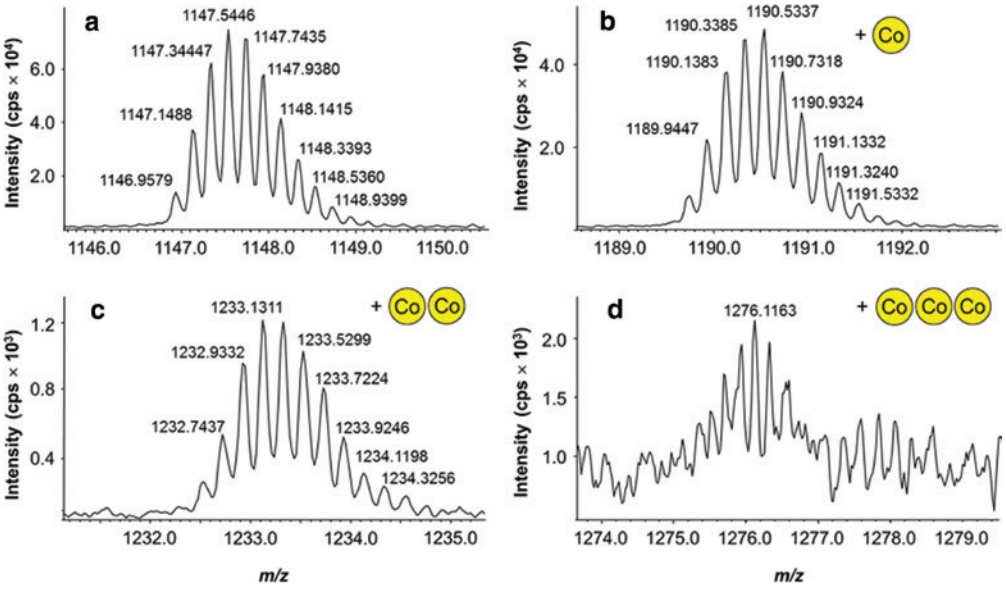

Fig. 7: Mass spectra recorded with a time-of-flight mass spectrometer before and after derivatization of insulin with CoS: underivatized (a), singly derivatized (b), doubly derivatized (c), triply derivatized (d). In each mass spectrum, the charge state +4 of the peptide is presented.

unmodified A chain was still detectable. In case of the B chain, the singly and doubly derivatized species could be identified.

4 Conclusions

The interaction of CoS with different low molecular weight amines, thiols, the peptide hormone insulin and the model protein *ß*-LGA was investigated. The labeling conditions were optimized initially. The labeling procedure using cobaltocinium salts features high efficiency and simplicity. All derivatized biomolecules were analyzed using mass spectrometry. Further structural characterization, i.e. localization of the labeling position, was carried out by a tryptic digestion of the labeled protein *ß*-LGA and fragmentation experiments of the obtained tryptic peptides by collision-induced dissociation. Besides the *N*-terminus and some lysine residues, the Co label could be located at the cysteine residue in the tryptic peptide T18. Despite the selectivity of the substituted succinimide ester group for amines, CoS labels cysteine residues as well. Therefore, it may potentially be used to label both amino and thiol functionalities in a large number of biomolecules offering a platform for proteome wide analysis in the future. As an additional benefit, the use of cobaltocinium tags provides the possibility for complementary detection by ESI-MS and ICP-MS. Whereas ESI-MS allows identification and structural elucidation, ICP-MS measurements provide the possibility for quantification by using the high sensitivity, the large dynamic range and the species-independent calibration of this hyphenated technique [27].

Acknowledgement: Financial support by the Fonds der Chemischen Industrie (Frankfurt/Main, Germany) is gratefully acknowledged.

References

[1] A. Leitner, W. Lindner, Current chemical tagging strategies for proteome analysis by mass spectrometry, *J. Chromatogr. B* **2004**, *813*, 1.
[2] A. Tholey, D. Schaumlöffel, Metal labelling for quantitative protein and proteome analysis using inductively coupled plasma mass spectrometry, *Trends Anal. Chem.* **2010**, *29*, 399.
[3] D. Kretschy, G. Koellensperger, S. Hann, Elemental labelling combined with liquid chromatography inductively coupled plasma mass spectrometry for quantification of biomolecules: a review, *Anal. Chim. Acta* **2012**, *750*, 98.
[4] B. Campanella, E. Bramanti, Detection of proteins by hyphenated techniques with endogenous metal tags and metal chemical labeling, *Analyst* **2014**, *139*, 4124.
[5] Z. R. Liu, X. T. Li, G. Y. Xiao, B. B. Chen, M. He, B. Hu, Application of inductively coupled plasma mass spectrometry in the quantitative analysis of biomolecules with exogenous tags, a review, *Trends Anal. Chem.* **2017**, *93*, 78.
[6] M. Salmain, G. Jaouen, Side-chain selective and covalent labelling of proteins with transition organometallic complexes. Perspectives in biology, *C. R. Chimie* **2003**, *6*, 249.
[7] H. E. Amouri, Y. Besace, J. Vaissermann, G. Jaouen, Synthesis, reactivity and X-ray molecule structure of the activated ester

Brought to you by | York University Libraries
Authenticated
Download Date | 12/4/18 9:33 AM

complex [(η^5-C_5H_4COONS)Co(CO)$_2$] (NS = N-succinimidyl), *J. Organomet. Chem.* **1996**, *515*, 103.
[8] D. Pröfrock, A. Prange, Chemical labels and natural element tags for the quantitative analysis of bio-molecules, *J. Anal. At. Spectrom.* **2008**, *23*, 432.
[9] S. Bomke, M. Sperling, U. Karst, Organometallic derivatizing agents in bioanalysis, *Anal. Bioanal. Chem.* **2010**, *397*, 3483.
[10] G. Schwarz, L. Mueller, S. Beck, M. W. Linscheid, DOTA based metal labels for protein quantification: a review, *J. Anal. At. Spectrom.* **2014**, *29*, 221.
[11] T. C. de Bang, S. Husted, Lanthanide elements as labels for multiplexed and targeted analysis of proteins, DNA and RNA using inductively-coupled plasma mass spectrometry, *Trends Anal. Chem.* **2015**, *72*, 45.
[12] Y. Liang, L. M. Yang, Q. Wang, An ongoing path of element-labelling/tagging strategies towards quantitative bioanalysis using ICP-MS, *Appl. Spectrosc. Rev.* **2015**, *51*, 117.
[13] J. E. Sheats, M. Rausch, Synthesis and properties of cobalticinium salts. I. Synthesis of monosubstituted cobalticinium salt, *J. Org. Chem.* **1970**, 35, 3245.
[14] E. O. Fischer, G. E. Herberich, Über die Reaktivität des Di-cyclopentadienyl-kobalt(III)-Kations, *Chem. Ber.* **1961**, *94*, 1517.
[15] T. J. Gill, L. T. Mann, Studies of synthetic polypeptide antigens, *J. Immun.* **1966**, *5*, 906.
[16] F. Noor, A. Wüstholz, R. Kinscherf, N. Metzler-Nolte, A cobaltocenium-peptide bioconjugate shows enhanced cellular uptake and directed nuclear delivery, *Angew. Chem. Int. Ed.* **2005**, *44*, 2429.
[17] J. T. Chantson, M. V. V. Falzacappa, S. Crovella, N. Metzler-Nolte, Solid-phase synthesis, characterization, and antibacterial activities of metallocene-peptide bioconjugates, *Chem. Med. Chem.* **2006**, *1*, 1268.
[18] J. T. Chantson, M. V. V. Falzacappa, S. Crovella, N. Metzler-Nolte, Antibacterial activities of ferrocenoyl- and cobaltocenium-peptide bioconjugates, *J. Org. Chem.* **2005**, *690*, 4564.
[19] I. Lavastre, J. Bescancon, P. Brossiert, C. Moise, The use of metallocenic esters of N-hydroxysuccinimide for metallohapten synthesis, *Appl. Organom. Chem.* **1991**, *5*, 143.
[20] A. L. Bordes, B. Limoges, P. Brossier, C. Degrand, Simultaneous homogenous immunoassay of phenytoin and phenobarbital using a Nafion-loaded carbon paste electrode and two redox cationic labels, *Anal. Chim. Acta* **1997**, *356*, 195.
[21] D. R. van Staveren, T. Weyhermüller, N. Metzler-Nolte, Organometallic β-turn mimetics. A structural and spectroscopic study of inter-stand hydrogen bonding in ferrocene and cobaltocenium conjugates of amino acids and dipeptides, *Dalton Trans.* **2003**, 210.
[22] A. Maurer, H.-B. Kraatz, N. Metzler-Nolte, Synthesis and electrochemical characterization of metallocene-PNA oligomers, *Eur. J. Inorg. Chem.* **2005**, *16*, 3207.
[23] P. D. Beer, C. Hazlewood, D. Hesek, J. Hodacova, S. E. Stokes, Anion recognition by acyclic redox-responsive amide-linked cobaltocenium receptors, *J. Chem. Soc. Dalton Trans.* **1993**, *8*, 1327.
[24] P. D. Beer, S. E. Stokes, Anion recognition by amide-linked pyridyl and pyridinium substituted cobaltocinium receptors, *Polyhedron* **1995**, *7*, 873.
[25] S. Wiese, T. Gronemeyer, R. Ofman, M. Kunze, C. P. Grou, J. A. Almeida, M. Eisenacher, C. Stephan, H. Hayen, M. Pawlas, C. Bunse, L. Schollenberger, T. Korosec, H. R. Waterham, W. Schliebs, R. Erdmann, J. Berger, H. E. Meyer, W. Just, J. E. Azevedo, R. J. A. Wanders, B. Warscheid, Proteomic characterization of mouse kidney peroxisomes by tandem mass spectrometry and protein correlation profiling, *Mol. Cell. Proteomics* **2007**, *6*, 2045.
[26] M. Brinkley, A brief survey of methods for preparing protein conjugates with dyes, haptens and crosslinking reagents, *Bioconjugate Chem.* **1992**, *3*, 2.
[27] T. Pfeifer, R. Janzen, T. Steingrobe, M. Sperling, C. Engelhard, W. Buscher, Development of a novel low-flow ion source/sampling cone geometry for inductively coupled plasma mass spectrometry and application in hyphenated techniques, *Spectrochim. Acta, Part B* **2012**, *76*, 48.

Brought to you by | York University Libraries
Authenticated
Download Date | 12/4/18 9:33 AM