Research Article

Journal of MASS SPECTROMETRY

Received: 7 February 2012 Revised: 16 April 2012 Accepted: 6 May 2012 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI 10.1002/jms.3025

Qualitative and quantitative characterization of the arsenic-binding behaviour of sulfur-containing peptides and proteins by the coupling of reversed phase liquid chromatography to electrospray ionization mass spectrometry

Anne-Christine Schmidt* and Kathleen Mickein

Phenylarsenic-substituted cysteine-containing peptides and proteins were completely differentiated from their unbound original forms by the coupling of reversed phase liquid chromatography with electrospray ionization mass spectrometry. The analysis of biomolecules possessing structure-stabilizing disulfide bridges after reduction provides new insights into requirements concerning the accessibility of cysteine residues for reducing agents as well as for arsenic compounds in a spatial protein structure. Complementary binding studies performed using direct ESI-MS without chromatographic coupling in different solvent systems demonstrated that more than one binding site were activated for aprotinin and lysozyme in denaturing solvents because of a stronger defolding. From the intensities of the different charge states occurring in the mass spectra as well as from the LC elution behaviour, it can be deduced that the folding state of the arsenic-bound protein species resembles the native, oxidized conformation. In contrast, although the milk protein α-lactalbumin has several disulfide bridges, only one phenylarsenic moiety was bound under strongly denaturing conditions. Because of the charge state distribution in the ESI mass spectra, a conformational change to a molten globule structure is assumed. For the second considered milk protein ß-lactoglobulin, a noncovalent interaction with phenylarsine oxide was detected.

In general, smaller apparent binding constants for the condensation reactions of the biomolecules with phenylarsine oxide leading to covalent arsenic–sulfur bindings were determined from direct injection ESI-MS measurements than from LC-ESI-MS coupling. The following order of binding affinities for one phenylarsenic group can be assumed from both ESI-MS and LC-ESI-MS: nonapeptide vasopressin > nonapeptide vasotocin > lysozyme > aprotinin > α-lactalbumin > thioredoxin. Kinetic investigations by LC-ESI-MS yielded a partial reaction order of 2 for vasopressin, Lys and α-lactalbumin and corresponding half-lives of 0.93, 2.56 and 123.5 min, respectively. Copyright © 2012 John Wiley & Sons, Ltd.

Supporting information may be found in the online version of this article.

Keywords: reversed phase liquid chromatography coupled to electrospray ionization mass spectrometry; arsenic-binding peptides and proteins; binding constants; rate constants; folding state

INTRODUCTION

Proteins act as enzymes, transporters, regulators and structure components in all living organisms. Peptides are involved in signal transfer processes and regulate important physiological functions. The elucidation of interactions between these biomolecules and environmental pollutants contributes to the understanding of the fatal effects of human activities on natural systems in the industrial era. Binding equilibrium constants and kinetic parameters are regarded as quantitative measures characterizing the effect of environmental pollutants and of pharmaceuticals on biomolecules.

Phenylated arsenicals have been produced as chemical warfare agents in the First and Second World War. Soils, and ground waters at former ammunition sites are contaminated with different phenylarsenic compounds such as diphenylarsinic acid, phenylarsonic acid and phenylarsine oxide (PAO) originating from degradation processes of diphenylarsine chloride (CLARK I), diphenylarsine cyanide (CLARK II) and phenylarsine dichloride (PFIFFIKUS).[1–3] Phenylated arsenic compounds were also used as drugs against syphilis and sleeping sickness.[4] Further, the use of arsenic compounds in the semiconductor technology, in glass production, in metal alloys and as pesticides leads to a dissemination of toxic arsenic in the environment.[5] Trivalent arsenic compounds react with thiol groups of biomolecules and inhibit their biological activity.[6] Reactions of the warfare degradation product PAO containing trivalent arsenic with several cysteine-containing

* *Correspondence to: Anne-Christine Schmidt, Institute of Analytical Chemistry, Technical University Bergakademie Freiberg, Leipziger Straße 29, D-09599 Freiberg, Germany. E-mail: Anne-Christine.Schmidt@chemie.tu-freiberg.de*

Institute of Analytical Chemistry, Technical University Bergakademie Freiberg, Leipziger Straße 29 D-09599 Freiberg, Germany

Copyright © 2012 John Wiley & Sons, Ltd.

peptides and proteins were elucidated in our previous works using electrospray ionization mass spectrometry (ESI-MS).[7,8] Bindings of arsenic compounds to biomolecules can be easily detected using ESI-MS due to the change of molecular mass.[9–12] However, the solvent compositions needed for spray formation and analyte ionization are not compatible with a native environment for biomolecules. Moreover, quantitative evaluations of binding equilibria were impeded by mutual ionization suppressions of reactants and reaction products.[13] To provide more native analysis conditions, a size exclusion chromatographic (SEC) method was developed for separation of arsenic-binding peptides and proteins, but the separation efficiency achieved by using an aqueous buffer as eluent was not sufficient for a complete differentiation of unbound and arsenic-substituted forms of biomolecules by the noninvasive UV detection.[14] The coupling of SEC to ESI-MS proved to be experimentally challenging but enabled a substance-specific detection of both arsenic-substituted biomolecules and unsubstituted forms.[14] Because of an incomplete chromatographic separation of reactants and products, these analyte pairs arrived simultaneously into the ion source so that competing effects during ionization occurring in case of ESI-MS without chromatographic coupling persist in SEC-ESI-MS. To ensure a temporally displaced entry into the ion source, a higher separation performance was desired and has been achieved by reversed phase liquid chromatography (RPLC).[15] For some nonapeptides, a complete separation of phenylarsenic-substituted forms from the original biomolecules was obtained by an RPLC gradient elution onto a monolithic column. In this case, the noninvasive UV detection sufficed to estimate binding constants from the peak areas of respective reactants and products, and the ESI-MS coupling could be omitted. However, larger proteins did not yield baseline-separated peaks for unbound and arsenic-bound forms during LC-UV method. For this reason as well as to further elucidate the arsenic-binding behaviour of peptides and proteins under varying experimental circumstances, we coupled the RPLC system with the ESI mass spectrometer in the current work. This approach completes our studies about the methodical influences of different analytical procedures on the quantitative evaluation of ligand bindings to biomolecules at the example of arsenic-binding peptides and proteins. Also in other works, LC-ESI-MS couplings were used for the analysis of arsenic species interactions with biomolecules.[9–12] However, the main part of these studies is restricted to qualitative aspects such as the structure and stability of compounds emerging from these interactions. An unelucidated stoichiometry and a lack of standard compounds contribute to the main problems concerning a quantitative analysis.[12] As important advantages of the mentioned analysis techniques focused on liquid chromatographic and mass spectrometric methods for bioanalytical purposes can be listed that very small sample amounts in the microliter range and low, biologically relevant analyte concentrations in the micromolar range suffice for the measurements. The gentle ionization process of the ESI-MS allows the analysis of intact biomolecules and their covalent and noncovalent ligand bindings. According to our previous works, the environmentally relevant trivalent phenylated arsenic compound PAO was chosen for the binding experiments because this arsenic species forms stable covalent arsenic–sulfur bonds in a condensation reaction with thiol groups of biomolecules.[7] An EC_{50} value of 0.03 mg l^{-1} determined for PAO in a toxicity assay with luminous bacteria[3] demonstrates the ecotoxicological potential of this substance. On the basis of our previous work,[15] a monolithic column was chosen for biomolecule separation because this type of chromatographic column delivers a high separation performance at short separation times for peptides and proteins.[16,17] The stoichiometry and apparent binding constants as important quantitative data of the reactions of the arsenic compound with different cysteine-containing peptides and proteins were deduced from the mass spectrometric signals of the original and the respective phenylarsenic-substituted forms. The influence of graduated denaturing properties of the solvent on the arsenic affinity of proteins is demonstrated in comparative measurements carried out using ESI-MS without chromatographic coupling. In addition, the reaction kinetics of the condensation reactions of selected biomolecules with PAO were elucidated by the help of the novel RPLC-ESI-MS method.

EXPERIMENTAL

Sample preparation

Deionized water prepared by a mixed-bed water demineralizer (TKA Wasseraufbereitungssysteme, Niederelbert, Germany) was used for all sample and solvent preparations. Biomolecules were purchased as solid substances from Sigma-Aldrich (Steinheim, Germany) and dissolved in 1 ml deionized water to make stock solutions of 1 mg ml^{-1} in case of vasopressin (Vpr) and vasotocin (Vtc) (acetate salts, ≥ 97% purity) as well as of thioredoxin (Trx; 2.48 U/mg protein, from *Spirulina* sp.), or of 10 mg ml^{-1} in case of aprotinin (Apr; 4 TIU/mg solid, from bovine lung), α-lactalbumin (α-Ltb; type I, ≥85%, from bovine milk), β-lactoglobulin A and B (β-Ltg; ~90%, from bovine milk), cytochrome *c* (Cyt c; ≥ 95%, from horse heart) and lysozyme (Lys; 96381 U/mg, from hen egg white). These peptide and protein stock solutions were aliquoted into 200 µl portions and stored frozen at −18 °C. A 2.98-mM stock solution of the arsenic compound PAO (solid substance in a purity of ≥97% from Sigma-Aldrich) in deionized water was stored at 4 °C. The reducing agent *tris*(carboxyethyl)phosphine (TCEP) was obtained as a 0.5-M hydrochloride solution from Sigma-Aldrich and stored at 4 °C. To perform binding studies, peptides and proteins were firstly reduced with a fivefold molar excess of TCEP for 20 min and thereafter incubated with PAO in varying molar ratios for 35 min in (i) 100% H_2O for injection into the LC system, (ii) 49 vol% H_2O/50 vol% acetonitrile (ACN; high-performance liquid chromatography gradient grade from VWR, Darmstadt, Germany)/1 vol% formic acid (HCOOH, p.a. from Merck, Darmstadt, Germany) for direct ESI-MS analysis and (iii) varying volume ratios of H_2O/ACN with 0.1 vol% HCOOH to simulate the solvent composition at the time point of elution in LC for direct ESI-MS analysis (Supplementary Table 1). In Supplementary Table 1, the initial concentrations of the biomolecules and the molar ratios to the arsenic compound are outlined, which were applied for direct ESI-MS measurements and for LC-MS coupling. Because of varying substance-specific ionization yields, slightly modified molarities of the different types of biomolecules were chosen. Using LC-MS, additional competing binding experiments were performed in which 25 µM Vpr, 10 µM Vtc, 10 µM Apr and 5 µM Lys were concomitantly incubated with molar PAO excesses to the summed biomolecule concentration of 1, 5, 10, 15 and 20 after reduction with a fivefold molar excess of TCEP to the summed biomolecule molarity.

To investigate the reaction kinetics of the condensation reactions of PAO with biomolecules by LC-MS, 25 µM Vpr, 5 µM Lys or 7.5 µM α-Ltb were firstly reduced by the fivefold TCEP concentration and then incubated with 25, 5 or 22.5 µM of the arsenic compound in case of Vpr, Lys and α-Ltb, respectively, during a period of 165 min at room temperature (22 °C). The first

Copyright © 2012 John Wiley & Sons, Ltd.

Journal of MASS SPECTROMETRY

LC-MS measurement was carried out immediately after starting the arsenic–sulfur reaction. Further samples were measured every 15 min. All sample measurements were performed in minimum three parallels to get standard deviations.

ESI-MS measurements

The Mariner™ ESI time of flight mass spectrometer (Applied Biosystems, Houston, TX, USA) was tuned and mass calibrated daily in the positive ionization mode before the beginning of both direct ESI-MS sample measurements and LC-MS coupling as described by Schmidt and Steier.[13] Experimental parameters adjusted at the mass spectrometer for direct injection using a syringe pump and for LC coupling are listed in Supplementary Table 2.

After each sample measurement in the direct injection mode, the ESI needle was flushed with a solvent mixture consisting of 49 vol% H_2O/50 vol% ACN/1 vol% HCOOH at flow rates between 10 and 15 μl min^{-1}.

LC-ESI-MS measurements

For LC measurements, an Agilent 1100 series liquid chromatograph (Agilent Technologies, Santa Clara, USA) comprising a binary pump, a degasser, an autosampler and a variable wavelength detector (VWD) was used. Biomolecules were separated onto a 4.6 × 50-mm monolithic reversed phase column (ProSwift® RP-2 H, Dionex, Sunnyvale, CA, USA) consisting of a polystyrene-divinylbenzene polymeric support material and a phenylic stationary phase. Eluents were prepared by mixing deionized water and ACN, acidified by formic or acetic acid (p.a. from Merck, Darmstadt, Germany), vacuum filtrated over polyamide filters (grade 290; Sartorius AG, Goettingen, Germany) and finally degassed in an ultrasonic bath for 20 min. A sample volume of 25 μl was injected onto the LC column. A detection wavelength of 215 nm was chosen for the VWD.

RESULTS AND DISCUSSION

Method development for LC-ESI-MS coupling

The monolithic column used in the preceding LC-UV-based binding study[15] was adopted for the setup of the LC-ESI-MS coupling because this type of column offers beneficial conditions for separation of peptides and proteins with regard to separation performance and separation time.[15–17] The high flow rate of the mobile phase as well as the ion pairing reagent that were favourable for efficient separation of the considered biomolecules on this monolithic column in case of UV detection[15] could not be adopted for ESI-MS coupling because the eluent flow had to be reduced to achieve a stable and reproducible spray formation, whereas the ion pairing reagent caused a strong signal suppression of the analytes owing to a very high own ionization yield. Therefore, the gradient elution had to be changed to make it ESI compatible. The fundamental composition of both eluents (eluent A, 95 vol% H_2O/5 vol% ACN; eluent B, 5 vol% H_2O/95 vol% ACN) and the gradient variation from 1% up to 50% eluent B were retained, whereas the ion pairing reagent was substituted for the addition of formic acid or acetic acid in varying amounts (0.1, 0.25 and 0.5 vol% in eluents A and B). The gradient time was varied from 5 to 25 min, and flow rates of 0.5 and 1.0 ml min^{-1} were tested. Steps during the gradient holding a fixed composition of the mobile phase over a defined duration were also tested regarding an improvement of the separation efficiency. An insufficient separation of the biomolecules, particularly a coelution of the nonapeptides, was observed upon addition of acetic acid, whereas the replacement by formic acid resulted in an improved separation performance. The following optimized conditions are proposed for separation of Vpr, Vtc, Apr, Cyt c, Lys, α-Ltb, β-Ltg and Trx by RPLC-ESI-MS coupling:

Eluent A: 94.9 vol% H_2O/5 vol% ACN/0.1 vol% HCOOH
Eluent B: 4.9 vol% H_2O/95 vol% ACN/0.1 vol% HCOOH
Gradient: 0% to 50% eluent B in 7.5 min.
Flow rate: 1 ml min^{-1} with a split ratio of 1:10 for the LC effluent resulting in a flow rate of 100 μl min^{-1} introduced into the ion source.

The retention order results according to the p*I* values and molar masses of the biomolecules listed in the next section and has been discussed in detail by Schmidt and Mickein.[15]

RPLC-ESI-MS analysis of interactions of PAO with peptides and proteins

Thiol groups of peptides and proteins are functional groups of the amino acid constituent cysteine and act as reducing agents in redox processes in cells. Besides this redox activity mentioned also at the discussion of Trx (see following paragraphs), cysteine pairs often form structure-stabilizing disulfide bridges in the three-dimensional structure of proteins. On the market for biochemicals, both larger number and larger amounts of this type of proteins can be obtained than redox-active protein species. Therefore, five proteins (Apr, Lys, Cyt c, α-Ltb and β-Ltg) and two nonapeptides (Vpr and Vtc) containing a differing number of cysteine pairs, which are linked to structure-determining disulfide bridges, were investigated here concerning their arsenic-binding behaviour after reduction. Although the reaction of such reduced forms of non-redox-active peptides and proteins with arsenic seems to be somewhat artificial, valuable conclusions can be drawn regarding position and distance of binding-capable cysteine residues in a spatial protein structure. Although the reactivity of As(III) compounds for thiols is well known for a long time, the requirements for the accessibility of thiol groups for arsenic in the tertiary structure of peptides and proteins could not be adequately explained up to now.[18–20] The non-redox-active biomolecule types can serve as less expensive and easily manageable model systems to elucidate potential requirements for arsenic bindings in defolded protein structures.

For two of the mentioned proteins, Apr and α-Ltb, as well as for the nonapeptides, Vtc and Vpr, condensation reactions with PAO were observed in mass spectrometric and liquid chromatographic analyses in our preceding investigations.[14,15] Some further protein types (Lys, Cyt c and ß-Ltg) for which no arsenic interactions have been detected before were considered now both in ESI-MS and LC-ESI-MS measurements. To perform the binding studies presented in this work, structure-stabilizing disulfide bridges were reduced using TCEP. This type of reducing agent was found to be more compatible with ESI-MS measurements compared with dithiothreitol and tributylphosphine.[8,13] Moreover, in contrast to dithiothreitol and β-mercaptoethanol often used in biochemical experiments, lower concentrations of TCEP suffice to ensure an effective reducing activity in a wide range of pH conditions.[21] The molar excess of TCEP was chosen according to our previous work[14] to

 Copyright © 2012 John Wiley & Sons, Ltd.

ensure a high reduction degree of the cysteines. Larger TCEP amounts were omitted because of the very high ionization yield of this substance.

The novel RPLC-ESI-MS method introduced in the Method Development for LC-ESI-MS coupling section enables the differentiation of arsenic-substituted forms of biomolecules from the respective unsubstituted forms using specific mass detection in addition to the measurement of unspecific UV absorption. Because for LC analyses, the reaction system was incubated in pure water, the native protein conformation served as an initial state for reduction and exhibited a limited accessibility for reducing agents. During the LC separation process, the protein is exposed to increasingly denaturing conditions but the contemporaneous separation from the reducing agent as well as from the arsenic compound prevents a subsequent reaction.

The nonapeptide hormones Vpr ($M = 1.084$ kDa, $pI \approx 8.05$) and Vtc ($M = 1.05$ kDa, $pI \approx 8.06$) possess one structure-stabilizing disulfide bridge each. The mass-to-charge ratios of oxidized and the reduced forms of the nonapeptides differ by one or two units for $[M + 2H]^{2+}$ and for $[M + H]^{+}$, respectively, evidencing the reduction of the disulfide bridge by TCEP. A decrease of both the peak areas recorded by UV detection and the signal intensities detected by ESI-MS of the unsubstituted initial forms was observed accompanied by an increase of the UV peak areas and mass signal intensities of the arsenopeptides with rising PAO molarities (Fig. 1). The phenylarsenic-containing peptide forms elute as double peaks (Figs 1b and 1c). Different conformers that arise from changing chemical environment during gradient elution occur in the form of multiple peaks.[22,23] In agreement with the former RPLC-UV analyses in which the water–acetonitrile gradient elution was supported by an ion pairing reagent,[15] the phenylarsenic-substituted peptide forms elute later than the unsubstituted forms owing to a larger molecular size.

The small basic protein Apr (6.51 kDa, $pI = 10.5$) acts as a protease inhibitor in different organs such as lung, liver and pancreas. Although Apr possesses three disulfide bridges in its amino acid sequence, which connect the polypeptide chain to a ravelled structure, only the binding of one phenylarsenic residue could be detected using LC-ESI-MS (Supplementary Figure 1). This finding agrees with previous SEC-MS measurements.[14] A prerequisite for the reaction with the arsenic compound is also the accessibility of the disulfide bridges for the reducing agent. Moreover, the spatial distance of the cysteines plays a role for incorporation of a phenylarsenic residue in a polypeptide chain.[14] In the mass spectra, a slight shift of the mass signal intensities towards higher charge states was observed after the reduction of the protein, whereas the intensity distribution of the phenylarsenic-substituted protein again resembles the oxidized protein (Supplementary Figure 1). A marked discrepancy in the retention behaviour of the oxidized and the reduced protein species was registered (Fig. 2a). The phenylarsenic-containing product elutes in two distant peaks, which also differ in their shape (Fig. 2b). In agreement with this, in RPLC separations using another eluent system, both the reduced and the phenylarsenic-containing Apr eluted over a wide time interval.[15]

Lys serves as enzyme for bacterial resistance in mucous membranes, tear fluid, saliva and hen egg white, and hydrolyses 1,4-ß-glycosidic bonds in cell walls of bacteria. The Lys type used in the current study originating from hen egg white ($M = 14.3$ kDa, $pI = 10.7$) contains four disulfide bridges, which hold the native protein in a globular conformation. In contrast to former studies,[7,15] one phenylarsenic substitution was

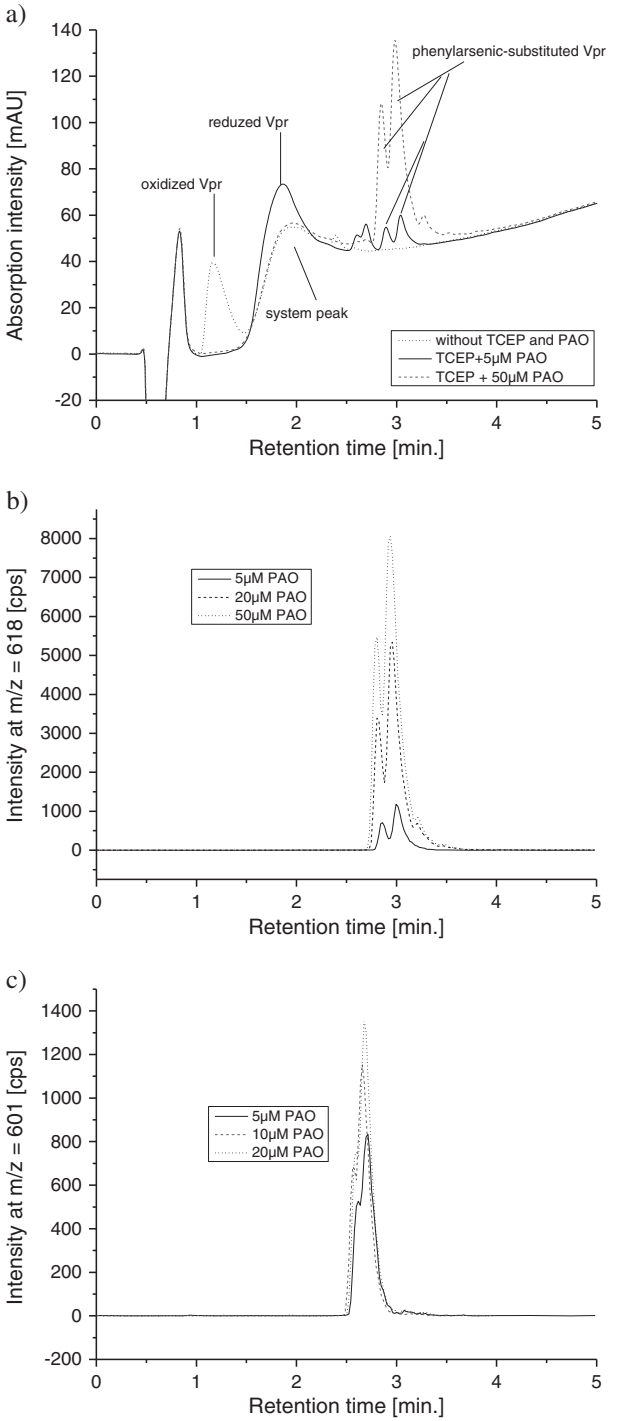

Figure 1. RPLC-UV-ESI-MS analysis of 25 µM (Vpr) or 10 µM (Vtc) peptide incubated with increasing PAO molarity at fivefold molar TCEP excess. (a) UV trace for Vpr at 215 nm. (b) XIC at *m/z*=618 obtained from ESI-MS for Vpr-AsPh. (c) XIC at *m/z*=601 obtained from ESI-MS for Vtc-AsPh.

discovered for Lys using RPLC-ESI-MS that contains four potential phenylarsenic-binding sites. The accessibility of the corresponding cysteine pairs for TCEP has been graduated in the following order: Cys30–Cys115 > Cys6–Cys127 > Cys64–Cys80 > Cys76–Cys94.[24] In the native folding state, the disulfide bridge of the Cys30–Cys115 pair is located at the surface of the molecule, whereas the other three disulfide bridges are covered in the inner sphere. The mass spectrum (Fig. 3) shows a changed

Copyright © 2012 John Wiley & Sons, Ltd.

Journal of MASS SPECTROMETRY

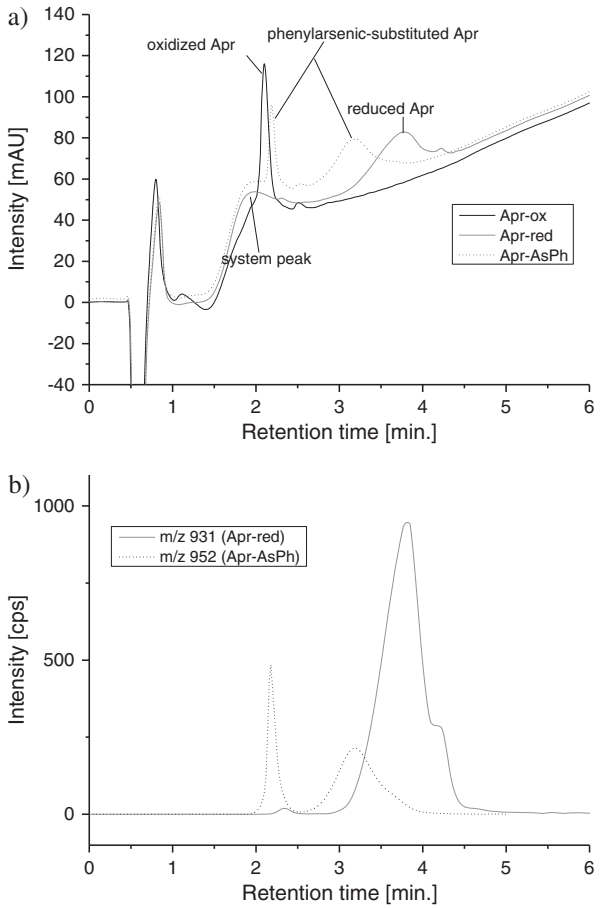

Figure 2. RPLC-UV-ESI-MS analysis of 10 μM Apr in the oxidized and reduced form and after incubation with 50 μM PAO. UV trace at 215 nm (a) and XIC traces at m/z = 931 for reduced Apr and at m/z = 952 for phenylarsenic-substituted Apr (b).

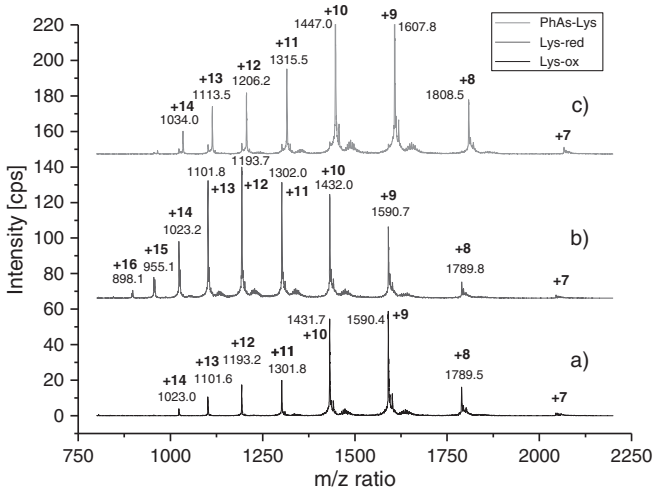

Figure 3. Mass spectra of oxidized (a), reduced (b), and phenylarsenic-substituted Lys (c) obtained from LC-ESI-MS coupling. Initial reactant concentrations: 5 μM Lys (a, b, c), 25 μM TCEP (b, c), and 5 μM PAO (c). 10 single spectra were averaged.

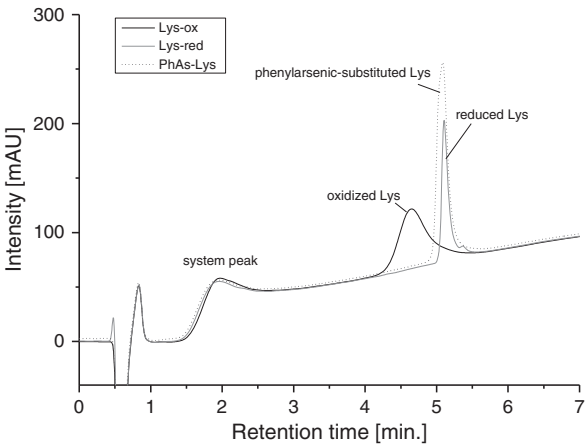

Figure 4. RPLC-UV-ESI-MS analysis of 5 μM Lys in the oxidized and reduced form and after incubation with 5 μM PAO. UV trace was recorded at 215 nm. Peaks were attributed to the reduced and phenylarsenic-substituted Lys by means of the corresponding XIC traces at m/z 1194 and 1447, respectively.

intensity distribution for arsenic-bound Lys: the centroid shifted to smaller charge states, indicating a stronger folding compared with the reduced, arsenic-free protein. On the other hand, the intensity distribution of the charge states of the oxidized form is similar to the phenylarsenic-substituted form (Figs 3a and 3c). From this finding, it can be assumed that the partial defolding entailed by the reduction step has been reversed by reconnection of two Cys residues via an S–As–S bridge. Despite their differing folding structures, the reduced and the arsenic-containing form coelute in RPLC around a retention time of 5.0 to 5.5 min (Fig. 4) because of the molecular size increase caused by the phenylarsenic moiety as ascertained using the corresponding mass spectra. The phenylarsenic-substituted form shows a slightly shorter retention time than the reduced form.

Both main components of cow milk, α-Ltb (14.178 kDa, $pI \approx 4.2\text{-}4.5$[25]) and ß-Ltg (isoform A: $M = 18.36$ kDa; isoform B: $M = 18.28$ kDa; $pI = 5.4$[26]), were also included in the arsenic-binding studies. In the polypeptide chain of α-Ltb, four structure-forming disulfide bridges occur that give a globular structure to the protein. α-Ltb contributes to the synthesis of lactose and belongs to the metalloproteins owing to its binding sites for Ca and Zn. Two disulfide bridges are located in the native conformation of ß-Ltg accompanied by a nonreactive thiol group in a hydrophobic region. Regarding the hitherto published arsenic-binding behaviour of α-Ltb, the phenylarsenic-substituted form could be partially separated from the initial form by using RPLC-UV and a calculation of consumption with PAO was possible.[15] In SEC-MS analysis using a solvent consisting of 75 vol% 10 mM ammonium formate, pH 5.0 and 25 vol% ACN, one phenylarsenic residue was bound to α-Ltb. In case of β-Ltg A + B, no changes of chromatograms appeared in former RPLC-UV analyses because of the incubation of the protein isoforms with PAO.[15] Such as in case of Lys, one phenylarsenic residue binds to α-Ltb despite four potential binding sites in the current LC-MS experiments. Differing accessibilities and reactivities of the four S-S bridges of α-Ltb for reducing agents have been shown.[27] A superreactivity was ascribed to the disulfide bridge between Cys6 and Cys120 because electrostatic effects resulting from the chemical environment as well as a bond tension caused by the geometric properties of the native protein conformation destabilize this disulfide bond. The intensity distribution of the charge states shifted from the oxidized to the reduced state in the same direction as recorded for Lys and

 Copyright © 2012 John Wiley & Sons, Ltd.

Apr, indicating a partial defolding after reduction (Figs 5a and 5b). In agreement with Apr and Lys, the charge distribution and the related folding structure of arsenic-containing α-Ltb equal to the reduced protein species (Fig. 5). According to preceding LC-UV studies in presence of an ion pairing reagent,[15] the arsenic-free initial forms and the phenylarsenic-substituted form of α-Ltb appear as an incompletely separated double peak in the chromatograms. With the help of the corresponding mass spectra, the first ($t_R = 5.61$ min) and the second part ($t_R = 5.72$ min) of the double peak could be associated to the unbound, reduced α-Ltb and to the reaction product, respectively (Supplementary Figure 2). In agreement with the other considered biomolecules, a decreasing peak area of the initial protein and a simultaneously increasing peak area of the arsenic-containing protein were observed both using UV and ESI-MS detection with increasing PAO concentrations in the initial samples. Although β-Ltg A and B have five cysteines each and one disulfide bridge is located in a position that is accessible for reducing agents,[28] no reaction with PAO could be detected using RPLC-UV-ESI-MS. Because of a somewhat higher molar mass ($\Delta m \approx 87\ g\ mol^{-1}$) of β-Ltg A, this isoform eluted somewhat later ($\Delta t_R \approx 0.08$ min in an incompletely separated double peak) than the isoform B (see Fig. 7).

Because arsenic bindings of the redox-active protein Trx were detected in our former studies using ESI-MS,[7,8] this slightly acidic (pI = 4.67) 11.69-kDa protein was also included in the current experiments. Trx transfers electrons in biological redox reactions and regulates the enzyme activation during photosynthesis and synthesis of desoxyribonucleotides. Trx elutes in form of multiple peaks that can be ascribed to different conformers (Fig. 6). On the basis of the extracted ion chromatograms (XICs) of the most intensive m/z ratios, it could be elucidated that all three Trx forms elute as a double peak (Fig. 6b). Because of a low ionization yield, larger Trx concentrations (25 μM) must be used compared with the other proteins (5 or 7.5 μM).

The novel RPLC-ESI-MS method is also suited for the analysis of complex reaction mixtures consisting of several biomolecules and the arsenic compound (Fig. 7). Despite of some overlapping peaks in the UV detection, the reactants

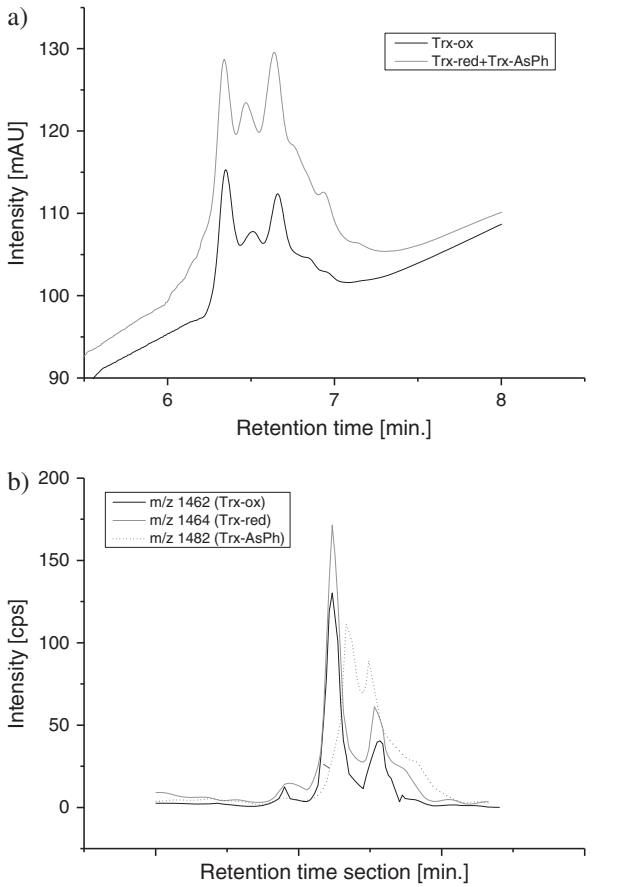

Figure 6. RPLC-UV-ESI-MS analysis of Trx in the oxidized, reduced, and phenylarsenic-substituted form. UV trace at 215 nm (a). XIC traces at m/z 1462 for oxidized Trx, at m/z 1464 for reduced Trx, and m/z 1482 for phenylarsenic-substituted Trx (b). Initial reactant concentrations: 25 μM Trx; 25 μM Trx and 125 μM TCEP; 25 μM Trx, 125 μM TCEP, and 125 μM PAO to monitor the oxidized, the reduced, and the phenylarsenic-substituted protein form, respectively.

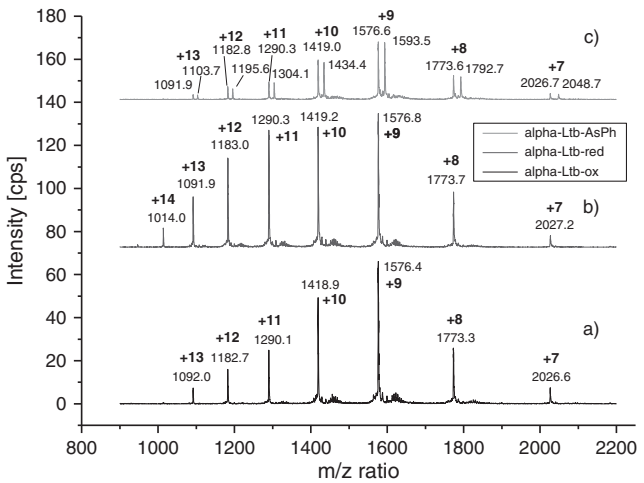

Figure 5. Mass spectra of oxidized (a), reduced (b), and phenylarsenic-substituted α-Ltb (c) obtained from RPLC-ESI-MS coupling. Initial reactant concentrations: 7.5 μM α-Ltb (a,b,c), 37.5 μM TCEP (b,c), 37.5 μM PAO (c). 10 single spectra were averaged. Due to an incomplete consumption, the mass signals of the reduced protein form occur besides the mass signals of the phenylarsenic-substituted form in spectrum c).

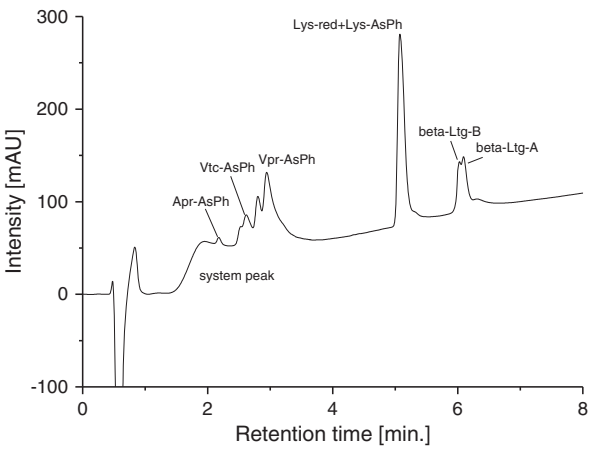

Figure 7. RPLC-UV chromatogram of a complex reaction mixture. The sample contains 20 μM Vpr, 10 μM Vtc, 10 μM Apr, 5 μM Lys, 5 μM β-Ltg, 250 μM TCEP, and 50 μM PAO in water.

Copyright © 2012 John Wiley & Sons, Ltd.

Journal of MASS SPECTROMETRY

and products can be reliably identified using the mass spectrometric detection.

ESI-MS with direct injection for analysis of interactions of PAO with peptides and proteins

In accordance with the LC-ESI-MS coupling (RPLC-ESI-MS Analysis of Interactions of PAO with Peptides and Proteins section), the increasing mass signal intensities of reaction products at simultaneously decreasing intensities of reactants were observed during arsenic-binding studies of biomolecules with increasing arsenic concentrations by using ESI-MS with direct sample injection.

ESI-MS of single reaction systems in different solvent compositions

For samples to be injected into the LC system, the reactions took place in pure water, and additions of an organic solvent and a strong acid were necessary for direct ESI-MS measurements to produce a stable spray and an efficient protonation of the analyte molecules. Because of this, the reactants in the direct ESI-MS-based binding studies were already exposed to a denaturing environment at the start of the arsenic–sulfur reaction. In contrast to this, the components of reaction systems prepared for LC-MS-based analysis are subjected to the denaturing eluents after the reaction took place in a nondenaturing aqueous solution. If the biomolecules were sprayed into the ion source, the solvent composition conforms to the eluent composition at the time point of elution from the column. To elucidate the effect of these solvent mixtures onto the mass signal ratios of reactants and products, the reaction systems were also measured by direct ESI-MS in such water–ACN mixtures, which reflect the respective elution conditions of the unreacted oxidized form of the biomolecule (see Supplementary Table 1). Second, all examined biomolecules were analyzed in a uniform solvent type, which was chosen according to a water–ACN–formic acid mixture optimized for ESI-MS measurements before[7] to ensure a uniform basis of comparison of the arsenic-binding behaviour of the different biomolecules. Here, the same type of acid was used as proton donor for ionization of biomolecules as it was found to be effective for biomolecule separation in the LC-MS coupling (compare Method Development for LC-ESI-MS Coupling section). A substitution of formic acid by the formerly used acetic acid[7] did not change the quality of the mass spectra regarding the signal to noise ratio for the analytes.

The nonapeptides Vpr and Vtc have short retention times so that a low concentration of organic solvent (4–5 vol%) predominates during elution from the chromatographic column and subsequent entry into the ion source. Because the folding state of these peptides is less affected by the solvent composition compared with larger proteins, the arsenic-binding behaviour was similar in both tested solvent systems and corresponds to the LC-MS-based data presented in the RPLC-ESI-MS Analysis of Interactions of PAO with Peptides and Proteins section. The small protein Apr reacted with one phenylarsenic residue in the solvent system 90.4 vol% H_2O/9.5 vol% ACN/0.1 vol% HCOOH that reflects the eluent composition at the retention time of Apr in LC-MS (Table 1). Therewith, the direct ESI-MS measurements confirmed the LC-MS results discussed previously (RPLC-ESI-MS Analysis of Interactions of PAO with Peptides and Proteins section). If the portion of the organic solvent was elevated to 50 vol%, additional phenylarsenic-substituted products appeared (Fig. 8) in the mass spectra, including

Table 1. Phenylarsenic substitutions of proteins in dependence on the solvent composition demonstrated by ESI-MS with direct sample injection

Protein	Solvent H_2O/ACN/HCOOH (v/v/v)	Oxidized form M (g mol^{-1})	Reduced form M (g mol^{-1})	Δm, monosubstituted—reduced form (g mol^{-1})	Δm, disubstituted—reduced form (g mol^{-1})	Δm, trisubstituted—reduced form (g mol^{-1})
Apr	90.4/9.5/0.1	6511	6515 ± 0	149.8 ± 0.84 (150)	—	—
	49/50/1		6517.33 ± 0.52	151.11 ± 0.32 (150)	302.04 ± 0.38 (300)	453.19 ± 0.64 (450)
Lys	63.4/36.5/0.1	14306	14311.83 ± 0.98	151.18 ± 0.47(150)	302.12 ± 0.58 (300)	452.28 ± 0.68 (450)
	49/50/1		14314.43 ± 0.79	150.57 ± 0.53 (150)	302.0 ± 0.71 (300)	454.0 ± 1.73 (450)
α-Ltb	58.9/41.0/0.1	14178	14181.67 ± 0.50	151.38 ± 0.52 (150)	—	—
	49/50/1		14181.11 ± 0.78	151.5 ± 0.93 (150)	—	—
β-Ltg	54.4/45.5/0.1	18279 (B) 18365 (A)	18280 ± 0 (B)	167.0 ± 1.41 (168) (B)	—	—
			18365.5 ± 0.55 (A)	167.67 ± 2.08 (168) (A)		
	49/50/1		18280.5 ± 1.1 (B)	166.33 ± 0.58 (168) (B)	—	—
			18366.17 ± 0.98 (A)	168.0 ± 1.73 (168) (A)		

Molar masses and mass differences obtained from mass deconvolutions including standard deviations from $n = 5$–8 parallel measurements are shown. For the phenylarsenic-substituted forms, theoretical mass differences are given in parentheses.

Copyright © 2012 John Wiley & Sons, Ltd.

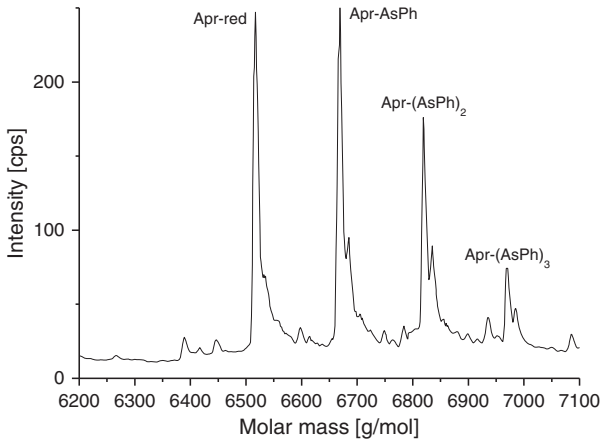

Figure 8. Deconvoluted mass spectra obtained from direct ESI-MS analysis of unbound, reduced Apr and of mono-, di-, and trisubstituted Apr in 49 vol% H_2O / 50 vol% ACN / 1 vol% HCOOH. Initial reactant concentrations: 10 μM Apr, 50 μM TCEP, 10 μM PAO. 10 single spectra were averaged.

monosubstituted, disubstituted and trisubstituted species. The measured mass difference deviations of the phenylarsenic substitutions from the theoretical value can be ascribed to the restricted mass resolution of the time-of-flight analyzer, especially at larger *m/z* ratios. In presence of the higher ACN concentration, an improved accessibility of the cysteine residues for the reducing agent exists because of the defolding of the native conformation. Moreover, the higher acid content of the solvent system optimized for direct ESI-MS potentiates the denaturing activity. Therefore, six protons were attached to the protein (Table 1), resulting in a reduction of all three disulfide bridges. After the incubation of the protein with the reducing agent in the solvent characterized by a high water content of 90.4%, only two disulfide bridges have been broken upon reduction. In case of Lys, the reduction of three disulfide bridges and the subsequent reaction with up to three PAO molecules was already detected in a solvent composition consistent with the eluent for Lys elution in RPLC (36.5 vol.% ACN) (Supplementary Figure 3). Similar mass spectra were obtained after elevation of the ACN content to 50 vol%. In this solvent type, the mass difference between the reduced and the oxidized form of Lys (Table 1) indicates a complete reduction of all four disulfide bridges, but because of steric or electronic hindrance in the reduced defolded protein, no fourth phenylarsenic moiety was bound. In contrast to this binding behaviour, only one phenylarsenic binding was found using the LC-MS coupling because the corresponding samples were prepared in pure water before injection into LC (compare RPLC-ESI-MS Analysis of Interactions of PAO with Peptides and Proteins section). α-Ltb reacted with one PAO in all tested solvent systems independent of the content of the organic solvent (Table 1). For α-Ltb, a specific partially folded protein structure termed molten globule is known.[29] This protein conformation involving three disulfide bridges shows a high stability in the acidic pH range[27] prevailing in the tested ESI-MS solvents. Because of this, only one of the four disulfide bridges of α-Ltb has been reduced by TCEP, and no further phenylarsenic residues were bound under increasing denaturing conditions in contrast to the other studied protein types. This assumption was substantiated by the detection of the same molar mass of the reduced protein form (14181 g mol^{-1}) both with LC-MS (reduction in pure water) and with direct ESI-MS (reduction in acidic ACN/water mixture).

In case of β-Ltg, no phenylarsenic substitutions were ascertained both in the eluent mixture prevailing at the retention time of the protein and in the special ESI-MS solvent. This finding is in agreement with the LC-MS analysis of β-Ltg-PAO reaction systems (see RPLC-ESI-MS Analysis of Interactions of PAO with Peptides and Proteins section). However, in both solvent systems used for the direct MS measurements, a noncovalent binding of PAO with both β-Ltg isoforms was observed (Table 1). Neither noncovalent nor covalent bindings with the arsenic compound were detected for Cyt c both in the corresponding eluent mixture (67.9 vol% H_2O/34.0 vol% ACN/0.1 vol% HCOOH) and in the ESI-MS solvent (49 vol% H_2O/50 vol% ACN/1 vol% HCOOH). The two cysteines of Cyt c (Cys-14 and Cys-17) are involved in two thioether bindings to a heme group and show no reactivity for reducing agents. The basic Cyt c (p*I* between 10.0 and 10.5; $M = 12.384$ kDa) is involved in the electron transfer in the respiration chain.

A correlation between the stoichiometry of arsenic-protein conjugates and the number of available cysteine residues was also found by Wang *et al.*[10] By using a coupling of affinity chromatography to ESI-MS, several properties of proteins were assumed to define their arsenic-binding behaviour: (i) the three-dimensional structure of proteins that determines if cysteine residues are located at the surface of the macromolecule or enclosed in its inner sphere; (ii) the position of other types of amino acids in proximity of free thiol groups, which affect the electron density and the hydrophobicity or can sterically hinder the thiol reactivity; and (iii) the orientation of the thiol functions that can be directed to the protein surface or to the inner sphere.

In general, ligand bindings can also be favoured by increasing reactant concentrations occurring during ESI spray formation. Therefore, in ESI-MS-based analysis, it is not known if the recorded bindings are formed in the original sample solution or if they represent analytical artefacts that are generated during the ESI process. Despite this, in case of the reaction systems considered in the current study, the measured reaction kinetics (see Investigation of Reaction Kinetics by LC-ESI-MS section) imply a reaction in the original sample solution because it resulted in half-life values in the range of <1 min up to 124 min.

Determination of binding constants from direct ESI-MS and LC-ESI-MS analysis for reactions of PAO with peptides and proteins

The quantitative evaluation of the reactions of PAO with the different peptides and proteins considered in the current study is based on a calculation model proposed in our previous works.[8,14] The main disadvantage of this approach consisting in the use of similar ionization efficiencies for respective reactants and products (Eqn (7)) as well as for different charge states of same analytes (Eqn (6)) was discussed in detail in a critical article.[13] However, these assumptions are necessary up to now because no standard substances are available for the calibration of ion yields of arsenic-containing biomolecules in ESI-MS. Our approach delivers no thermodynamic constants but apparent values that suffice for a comparison of binding affinities of different biomolecules under similar experimental conditions. XICs obtained from RPLC-ESI-MS coupling and peak intensities of deconvoluted mass spectra obtained from ESI-MS measurements without LC coupling for unbound and arsenic-substituted forms of peptides and proteins were used for calculation of binding constants. In case of LC-MS, the sum of all XIC peak areas

Copyright © 2012 John Wiley & Sons, Ltd.

originating from the various *m/z* ratios of reactants and products was used. A quantitation of mass spectrometric data on the basis of only one intensive mass signal is less elaborate indeed, but the intensity distribution of the charge states of a protein can fluctuate between different measurements.[30] For quantitative LC-MS analysis of the reaction systems, further uncertainties must be kept in mind because the ionization yields of reactants and products may vary because of the solvent gradient and the coelution.

The calculation pathway is described in the following equations (Eqns (2)–(7)) at the example of LC-MS analysis of the reaction of Apr with PAO (Eqn (1)):

$$\mathrm{Apr(SH)_2} + \mathrm{(C_6H_5)AsO} \rightleftharpoons \mathrm{Apr(S(C_6H_5)As)S)} + \mathrm{H_2O} \quad (1)$$

$$K = [\mathrm{Apr-PhAs}]_{eq} \cdot [\mathrm{H_2O}] / [\mathrm{Apr}]_{eq} \cdot [\mathrm{PAO}]_{eq} \quad (2)$$

with3a

$$[\mathrm{Apr}]_{eq} = [\mathrm{Apr}]_0 \cdot \mathrm{PA(XIC)}_{\mathrm{Apr,\ unbound}} / \left(\mathrm{PA(XIC)}_{\mathrm{Apr,\ unbound}} + \mathrm{PA(XIC)}_{\mathrm{PhAs-Apr}}\right) \quad (3a)$$

$$[\mathrm{Apr-PhAs}]_{eq} = [\mathrm{Apr}]_0 \cdot \mathrm{PA(XIC)}_{\mathrm{PhAs-Apr}} / \left(\mathrm{PA(XIC)}_{\mathrm{Apr,\ unbound}} + \mathrm{PA(XIC)}_{\mathrm{PhAs-Apr}}\right) \quad (3b)$$

$$[\mathrm{PAO}]_{eq} = [\mathrm{PAO}]_0 - [\mathrm{Apr-PhAs}]_{eq} \quad (3c)$$

where

$$[\mathrm{Apr}]_0 = [\mathrm{Apr}]_{eq} + [\mathrm{Apr-PhAs}]_{eq} \quad (4)$$

$$[\mathrm{Apr}]_{eq} = b_{\mathrm{Apr,\ unbound}} \cdot \mathrm{PA(XIC)}_{\mathrm{Apr,\ unbound}} \quad (5a)$$

$$[\mathrm{Apr-PhAs}]_{eq} = b_{\mathrm{PhAs-Apr}} \cdot \mathrm{PA(XIC)}_{\mathrm{PhAs-Apr}} \quad (5b)$$

$$\begin{aligned}\mathrm{PA(XIC)}_{\mathrm{Apr,\ unbound}} &= \mathrm{PA(XIC)}\left([\mathrm{Apr}+5\mathrm{H}]^{5+}\right) \\ &+ \mathrm{PA(XIC)}\left([\mathrm{Apr}+6\mathrm{H}]^{6+}\right) \\ &+ \mathrm{PA(XIC)}\left([\mathrm{Apr}+7\mathrm{H}]^{7+}\right) \\ &+ \mathrm{PA(XIC)}\left([\mathrm{Apr}+8\mathrm{H}]^{8+}\right)\end{aligned} \quad (6a)$$

$$\begin{aligned}\mathrm{PA(XIC)}_{\mathrm{PhAs-Apr}} = \mathrm{PA(XIC)}\big([&\mathrm{PhAs-Apr}+5\mathrm{H}]^{5+} \\ +[&\mathrm{PhAs-Apr}+6\mathrm{H}]^{6+} \\ +[&\mathrm{PhAs-Apr}+7\mathrm{H}]^{7+} \\ +[&\mathrm{PhAs-Apr}+8\mathrm{H}]^{8+}\big)\end{aligned} \quad (6b)$$

$$b_{\mathrm{Apr,\ unbound}} = b_{\mathrm{PhAs-Apr}} = b \quad (7)$$

The symbols for Eqns (2)–(7) are defined as follows: $[\mathrm{Apr-PhAs}]_{eq}$ is the equilibrium concentration of phenylarsenic-bound Apr, $[\mathrm{H_2O}]$ is the water concentration of the sample, $[\mathrm{Apr}]_{eq}$ is the equilibrium concentration of unbound Apr, $[\mathrm{PAO}]_{eq}$ is the equilibrium concentration of PAO, $[\mathrm{Apr}]_0$ is the initial concentration of Apr, PA is the peak area, b is the sensitivity coefficient for ESI-MS detection and PhAs is the protein-bound phenylarsenic moiety.

The mass signal for monoprotonated PAO appeared in the XICs of the LC-MS only with a very low intensity because the coelution with TCEP in the front region of the chromatogram and the overlap with other low-molecular impurities suppressed the ionization of the arsenic compound. Because of this, the PAO signal was hardly evaluable for quantitative purposes. Besides, an external calibration used for determination of free PAO amounts in reaction equilibria with a tripeptide resulted in an underestimation.[13]

K values obtained from ESI-MS with direct injection and from the coupling with LC are summarized in Table 2. In case of LC-MS analysis, negative *K* values resulted for PAO– biomolecule ratios <1.0. These negative *K* values were not included in Table 2 because they cannot be regarded as reliable. The reason can be ascribed to the approximated estimation of the concentration of the unreacted arsenic compound (see Eqn 3c): because of the high ion yield of the arsenopeptides and arsenoproteins in the ESI, large XIC peak areas resulted for the reaction products. In Eqn 3b, high concentrations of the reaction products were associated to the XIC peak areas, which were then subtracted from the initial PAO concentration to get the free PAO concentration remaining in the reaction equilibrium. Negative, unrealistic concentrations for unreacted PAO occur at low initial PAO amounts and lead to negative *K* values in Eqn 2. In case of direct ESI-MS analysis, no negative PAO concentrations occur so that *K* values can also be given for PAO–biomolecule ratios <1. From ESI-MS without chromatographic coupling, somewhat lower *K* values were obtained than from LC-MS coupling. Similar to the LC-MS-based *K* values, binding constants calculated from ESI-MS signal intensities are characterized by a decreasing trend with rising PAO molarity. This phenomenon was also observed for other ESI-MS-based arsenic-binding studies.[8,13,14] This trend was attenuated at higher PAO excess. Also, the standard deviation of the parallel measurements diminished with rising PAO concentration. In contrast to the other considered biomolecules, a rising *K* trend with increasing PAO molarities was observed for the Apr-PAO reaction system, possibly due to strongly increasing signal intensities of the arsenic-containing product in direct ESI-MS analysis. Somewhat larger binding constants resulted from the competing experiment than from the analyses of single reaction systems at similar molar PAO–biomolecule ratios, presumably due to larger PAO excess in the competing experiment. Because of relative standard deviations higher than 100%, no reliable *K* values for the reaction of α-Ltb with PAO can be deduced from LC-MS analysis. Kinetic investigations revealed that this reaction system had not reached its equilibrium state within the time scale of the LC-MS experiments (Investigation of Reaction Kinetics by LC-ESI-MS section).

For several parallel reaction equilibria arising in solvent A in case of Apr and in both solvent systems in case of Lys (Table 1), the calculation model did not suffice. Consumption numbers were proposed by Schmidt and Steier[13] as alternative quantitative indication for reaction equilibria. On the basis of this approach, the gradual formation of the individual reaction products is demonstrated in Figs 9 and 10. At low PAO molarities, Apr reacted with one or two PAO molecules. The threefold phenylarsenic substitution prevails at higher PAO excess. Similar to Apr, a diminution of the fraction of the unbound protein and a concomitant growth of the fractions of the monosubstituted, disubstituted and trisubstituted reaction products was observed for Lys (Fig. 10). However, in contrast to Apr, the monosubstituted protein species remains dominant. For the reaction of α-Ltb with the arsenic compound, a *K* calculation was possible for both solvents because only one reaction product occurred in each case (see Table 1). A comparison of both solvent systems used for ESI-MS offers slightly larger *K* values in solvent B that contains lower ACN and acid concentrations than those in solvent A. This binding behaviour does not correspond with the stronger

 Copyright © 2012 John Wiley & Sons, Ltd.

Table 2. Binding constants *K* for the condensation reaction of PAO with different peptides and proteins determined from ESI-MS with direct injection and from RPLC-ESI-MS coupling

Biomolecule, initial concentration	Molar ratio PAO–biomolecule	*K* values from LC-ESI-MS (RSD)		*K* values from ESI-MS (RSD)	
		Single experiment[a]	Competing experiment[b]	Solvent A[c]	Solvent B[d]
Vpr 25 μM (LC-MS) 2.5 μM (direct MS)	0.5				$1.8 \times 10^9 \pm 1.9 \times 10^8$ (10.5%)
	0.75				$1.2 \times 10^8 \pm 8.0 \times 10^6$ (7.0%)
	1.0	$7.9 \times 10^9 \pm 1.1 \times 10^9$ (13.6%)			$5.8 \times 10^7 \pm 1.9 \times 10^6$ (3.2%)
	2.0	$7.5 \times 10^8 \pm 5.3 \times 10^7$ (7.1%)	$3.5 \times 10^9 \pm 3.2 \times 10^8$ (9.4%)		$2.4 \times 10^7 \pm 2.2 \times 10^6$ (9.4%)
	3.0	$8.4 \times 10^8 \pm 3.4 \times 10^7$ (4.1%)			$1.4 \times 10^7 \pm 1.5 \times 10^6$ (10.6%)
	10.0		$2.5 \times 10^8 \pm 1.7 \times 10^7$ (6.7%)		
	20		$1.7 \times 10^8 \pm 4.3 \times 10^7$ (25.3%)		
	30		$9.8 \times 10^7 \pm 2.0 \times 10^7$ (21%)		
Vtc 10 μM (LC-MS) 2.5 μM (direct MS)	0.5				$1.7 \times 10^7 \pm 1.7 \times 10^6$ (9.8%)
	0.75				$1.2 \times 10^7 \pm 6.7 \times 10^5$ (5.4%)
	1.0	$4.5 \times 10^9 \pm 1.5 \times 10^9$ (33.3%)			$1.1 \times 10^7 \pm 1.2 \times 10^6$ (10.2%)
	2.0	$1.5 \times 10^8 \pm 4.8 \times 10^7$ (32.4%)			$5.5 \times 10^6 \pm 5.5 \times 10^5$ (9.9%)
	3.0	$5.7 \times 10^7 \pm 3.3 \times 10^6$ (5.8%)			$3.4 \times 10^6 \pm 4.4 \times 10^5$ (13%)
	4.0	$3.9 \times 10^7 \pm 3.3 \times 10^5$ (0.9%)			
	5.0		$1.5 \times 10^9 \pm 3.1 \times 10^8$ (21.5%)		
	25.0		$2.9 \times 10^8 \pm 1.3 \times 10^7$ (4.3%)		
	50.0		$1.4 \times 10^8 \pm 1.7 \times 10^7$ (12.4%)		
	75.0		$1.2 \times 10^8 \pm 9.3 \times 10^6$ (7.8%)		
Apr 10 μM (LC-MS) 10 μM (direct MS)	0.25				$3.5 \times 10^6 \pm 2.8 \times 10^5$ (8.1%)
	0.5				$8.4 \times 10^6 \pm 1.9 \times 10^5$ (2.2%)
	0.75				$1.5 \times 10^7 \pm 1.9 \times 10^6$ (13.3%)
	1.0	$4.9 \times 10^8 \pm 2.4 \times 10^7$ (4.8%)			$1.2 \times 10^7 \pm 2.8 \times 10^5$ (2.4%)
	3.0	$1.3 \times 10^8 \pm 6.5 \times 10^6$ (4.9%)			$1.9 \times 10^7 \pm 1.0 \times 10^6$ (5.2%)
	5.0	$7.0 \times 10^7 \pm 6.7 \times 10^6$ (9.6%)	$4.5 \times 10^8 \pm 6.0 \times 10^7$ (13.5%)		$2.9 \times 10^7 \pm 9.8 \times 10^5$ (3.4%)
	7.5	$4.9 \times 10^7 \pm 1.1 \times 10^6$ (2.2%)			
	10.0	$3.4 \times 10^7 \pm 2.0 \times 10^6$ (6.0%)			

(Contitnues)

Copyright © 2012 John Wiley & Sons, Ltd.

Journal of MASS SPECTROMETRY

Table 2. *(Continued)*

Biomolecule, initial concentration	Molar ratio PAO–biomolecule	*K* values from LC-ESI-MS (RSD)		*K* values from ESI-MS (RSD)	
		Single experiment[a]	Competing experiment[b]	Solvent A[c]	Solvent B[d]
	25.0		$1.6 \times 10^7 \pm 8.3 \times 10^4$ (0.5%)		
	50.0		$9.4 \times 10^6 \pm 1.1 \times 10^6$ (11.6%)		
	75.0		$6.7 \times 10^6 \pm 2.9 \times 10^5$ (4.3%)		
Lys 5 μM (LC-MS)	1.0	$3.3 \times 10^9 \pm 9.2 \times 10^8$ (27.9%)			
	1.5	$6.1 \times 10^8 \pm 2.1 \times 10^7$ (3.4%)			
	2.0	$2.8 \times 10^8 \pm 9.8 \times 10^7$ (35%)			
	10.0		$5.2 \times 10^8 \pm 4.4 \times 10^7$ (8.5%)		
	50.0		$1.4 \times 10^7 \pm 1.5 \times 10^6$ (11.0%)		
	100		$6.3 \times 10^6 \pm 3.3 \times 10^5$ (5.2%)		
α-Ltb 7.5 μM (LC-MS)	0.5	$5.0 \times 10^7 \pm 5.6 \times 10^7$ (113%)		$4.5 \times 10^5 \pm 2.1 \times 10^4$ (4.6%)	$4.8 \times 10^5 \pm 7.3 \times 10^4$ (15.1%)
7.5 μM (direct MS)	1.0	$3.6 \times 10^6 \pm 7.0 \times 10^6$ (194%)		$2.8 \times 10^5 \pm 9.4 \times 10^3$ (3.4%)	$4.6 \times 10^5 \pm 6.6 \times 10^4$ (14.6%)
	3.0	$2.3 \times 10^6 \pm 1.3 \times 10^6$ (57.1%)		$1.2 \times 10^5 \pm 6.7 \times 10^3$ (5.4%)	$1.5 \times 10^5 \pm 6.7 \times 10^3$ (4.6%)
	5.0	$1.9 \times 10^6 \pm 1.2 \times 10^6$ (66%)		$9.4 \times 10^4 \pm 3.4 \times 10^3$ (3.6%)	$1.2 \times 10^5 \pm 1.1 \times 10^4$ (9.0%)
	7.5	$6.0 \times 10^5 \pm 6.5 \times 10^6$ (1076%)		$8.0 \times 10^4 \pm 2.5 \times 10^3$ (3.1%)	$8.4 \times 10^4 \pm 6.2 \times 10^3$ (7.4%)
	10.0	$6.4 \times 10^5 \pm 7.3 \times 10^6$ (1141%)		$2.6 \times 10^4 \pm 2.8 \times 10^3$ (10.8%)	$7.5 \times 10^4 \pm 2.2 \times 10^3$ (2.9%)
Trx 25 μM (LC-MS)	2.5	$5.1 \times 10^5 \pm 3.1 \times 10^4$ (6%)			
	5.0	$3.4 \times 10^5 \pm 3.8 \times 10^4$ (11.4%)			
	10.0	$2.4 \times 10^5 \pm 3.4 \times 10^4$ (14.5%)			

Average values and standard deviations from three parallel experiments are given for each concentration ratio of the reactants.
[a]Individual biomolecules were incubated with PAO.
[b]Five biomolecules were incubated with PAO in the same sample.
[c]Solvent A: 49 vol.% H_2O/50 vol.% ACN/1 vol.% HCOOH.
[d]Solvent B corresponds to the eluent composition at time point of elution of unreacted biomolecules in LC (see Table 1).

denaturing properties of solvent A, which should lead to an enhanced reactivity against arsenic. Presumably, the altered ionization behaviour of reactants and products plays a role.

To compare the binding affinities of the different biomolecules, the *K* values based on a PAO–biomolecule ratio of 1.0 or the averaged overall tested concentration ratios can be used. Hence, the following order of binding affinities can be deduced, Vpr > Vtc > Lys > Apr > α-Ltb > Trx, and substantiated by the smaller size of the peptides Vtc and Vpr and the related lower steric hindrance of two thiol groups to react with the arsenic compound compared with the larger proteins. The higher arsenic reactivity of Vpr compared with Vtc was also confirmed in former LC-UV measurements.[15] Within the four considered proteins, differing orientations of reactive thiol groups as well as the physicochemical environment caused by other amino acid side chains play a role.

For some of the considered biomolecules, a comparison with *K* values obtained from alternative methods can be made. For this purpose, ranges of *K* values averaged from different PAO–biomolecule concentration ratios were used to allow a comparison of different reactant molarities that have been measured by different analytical techniques. Despite the tenfold higher protein concentrations analysed with SEC-MS, the resulting binding constants for the reaction of PAO with Apr (≈1.6×10^8[14]) agreed with the corresponding data obtained from the novel RPLC-MS method (≈1.5×10^8). In contrast to this, higher binding constants

 Copyright © 2012 John Wiley & Sons, Ltd.

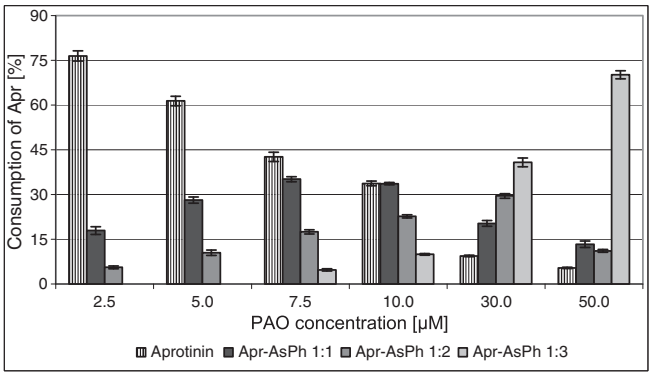

Figure 9. Consumption of Apr in the condensation reactions with rising PAO concentrations related to the initial Apr concentration of 10 µM. Direct ESI-MS analysis of $n = 3$ parallel prepared samples in a solvent consisting of 49 vol% H_2O/50 vol% ACN/1 vol% HCOOH.

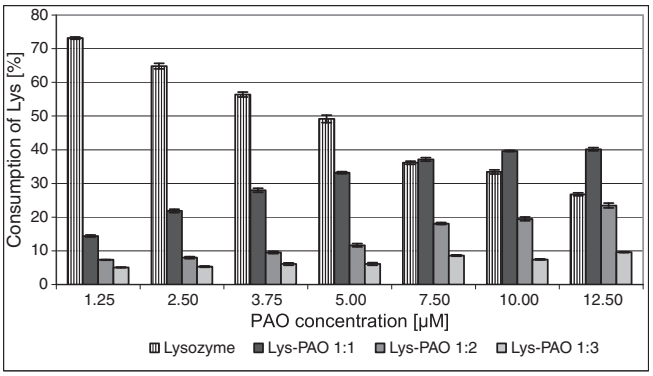

Figure 10. Consumption of Lys in the condensation reactions with rising PAO concentrations related to the initial Lys concentration of 5 µM. Direct ESI-MS analysis of $n = 3$ parallel prepared samples in a solvent consisting of 49 vol% H_2O/50 vol% ACN/1 vol% HCOOH.

resulted from SEC-MS (4.4×10^{10}–1.4×10^{11}[14]) in comparison to RPLC-MS ($\approx 3.6 \times 10^5$) for the reaction of PAO with Trx. Because the SEC eluent system caused a lower sensitivity of the MS detection, 13-fold higher molarities of this protein reactant were required compared with RPLC-MS. For the Trx-PAO reaction system, an average K value similarly to RPLC-MS was calculated from ESI-MS with direct injection ($\approx 6.4 \times 10^5$[7]). In case of the reactions of the nonapeptides Vtc and Vpr, the K value ranges determined by our new RPLC-MS method (3.9×10^7–4.5×10^9 for Vtc; $\approx 2.5 \times 10^9$ for Vpr) were somewhat higher than the corresponding values from RPLC-UV (9.6×10^5–1.2×10^8 for Vtc; 2.2×10^6–1.4×10^9 for Vpr[15]), whereby four- to ninefold higher peptide molarities had been used for the less sensitive UV detection. For the Lys-PAO system, no reference values are available up to now in the literature.

Investigation of reaction kinetics by LC-ESI-MS

The kinetic behaviour of the reactions of selected biomolecules with PAO was pursued using LC-MS coupling because a more native folding state of the initial biomolecules is recorded compared with ESI-MS without LC coupling (compare ESI-MS with Direct Injection for Analysis of Interactions of PAO with Peptides and Proteins section). To ascertain the partial reaction orders with respect to the individual biomolecules as well as the corresponding rate constants, the molar concentrations of the initial biomolecules and of their phenylarsenic-bound products had to be first determined from the XIC peak areas as described in Eqns (3)–(7). Then, the method of the integrated rate laws was applied. For all tested biomolecules, the plotting of the reciprocal concentration values of the unsubstituted original species against time passed after starting the reaction resulted in a straight line with a very good linearity: $R^2 = 0.978$, 1.000 and 0.983 for Vpr, Lys and α-Ltb, respectively. The reactions showed a partial reaction order of 2 with respect to the biomolecule reactant (Table 3). Beside these reaction orders, rate constants and half-lives are summarized in Table 3. α-Ltb reacted with PAO substantially more slowly than Apr and Lys. From the kinetic data obtained for α-Ltb, binding constants determined by LC-MS (Determination of binding constants from direct ESI-MS and LC-ESI-MS analysis for reactions of PAO with peptides and proteins section) does not reflect the state of equilibrium but represent the momentary state of the reaction system at a certain time point. The incubation time of the reaction system before injection into the LC was 35 min for these binding experiments (Sample Preparation section). Because no reactions of zeroth order occurred, the reaction rates are dependent on the reactant concentrations.

CONCLUSIONS

An LC-ESI-MS coupling method for the separation and mass-specific detection of eight peptides and proteins as well as of their phenylarsenic-substituted reaction products was developed, which was suitable for the calculation of apparent binding constants and kinetic parameters for the reactions of the biomolecules with PAO. Complementary binding constants were determined from ESI-MS measurements with direct injections in two different solvent systems. Because of the addition of ACN and formic acid to the reactive sample, some proteins were stronger denatured, leading to the binding of up to three phenylarsenic residues instead of one binding observed in LC-MS. From charge state distributions in the ESI mass spectra, conclusions concerning the folding state of the phenylarsenic-substituted, reduced and oxidized protein forms can be drawn.

Two main advantages of the LC-ESI-MS coupling for the analysis of arsenic-binding peptides and proteins can be noticed. First, the reaction of the biomolecules with the arsenic compound can take place under nondenaturing conditions, for example, in pure water. Second, mixtures of biomolecules can be analyzed regarding their arsenic-binding behaviour. However, during the reduction step as well as the separation process, the biomolecules are exposed to denaturing agents.

To ensure more native separation conditions, arsenic-binding peptides and proteins should also be analyzed by electrophoretic

Table 3. Kinetic parameters for the reactions of biomolecules with PAO determined from time-graduated LC-ESI-MS analysis

Biomolecule	r	k (l mol^{-1} s^{-1})	$t_{1/2}$ (min)
Vpr	2	7.2×10^{-4}	0.93
Lys	2	1.3×10^{-3}	2.56
α-Ltb	2	1.8×10^{-5}	123.5

r, partial reaction order of the biomolecule; k, rate constant for the partial reaction of the biomolecule; $t_{1/2}$, half-life for the biomolecule.

Copyright © 2012 John Wiley & Sons, Ltd.

Journal of MASS SPECTROMETRY

methods such as capillary zone electrophoresis using aqueous buffers as separation media or blue native polyacrylamide gel electrophoresis. Gel electrophoresis in combination with element-specific mass spectrometry for arsenic detection as well as with ESI-MS for protein identification can also help to analyze real biological samples such as body fluids or serum containing thousands of proteins and several toxic arsenic compounds. In this context, the current LC-MS-based binding study demonstrates the general binding capability of arsenic compounds to sulfur-containing biomolecules. An attribution of arsenic-binding sites to the different cysteine residues within the amino acid sequence should succeed with the help of fragmentation experiments by using tandem mass spectrometry in the future. Because the phenyl ring of PAO exhibits a high UV absorptivity, titration experiments with UV detection could serve as an alternative technique to get binding constants for the considered reaction systems.

Acknowledgements

The authors thank the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) for financial support.

Supporting Information

Supporting information may be found in the online version of this article.

REFERENCES

[1] B. Daus, J. Mattusch, R. Wennrich, H. Weiss. Analytical investigations of phenyl arsenicals in groundwater. *Talanta* **2008**, *75*, 376.
[2] B. Daus, M. Hempel, R. Wennrich, H. Weiss. Concentrations and speciation of arsenic in groundwater polluted by warfare agents. *Environ. Pollut.* **2010**, *158*, 3439.
[3] R. Haas, M. Müller, L. Kaminski. Aquatische Ökotoxizität von Phenylarsinverbindungen—1. Neben- und Umwandlungsprodukte von Blaukreuzkampfstoffen. *UWSF–Z. Umweltchem. Ökotox.* **1996**, *8*, 62.
[4] R. L. Krauth-Siegel, H. Bauer, R. H. Schirmer. Dithiolproteine als Hüter des intrazellulären Redoxmilieus bei Parasiten: alte und neue Wirkstoff-Targets bei Trypanosomiasis und Malaria. *Angew. Chem.* **2005**, *117*, 698.
[5] HBM-Kommission. Stoffmonographie Arsen—Referenzwert für Urin. Stellungnahme der Kommission "Human-Biomonitoring" des Umweltbundesamtes. *Bundesgesundheitsbl. Gesundheitsforsch. Gesundheitsschutz* **2003**, *46*, 1098.
[6] M. F. Hughes. Arsenic toxicity and potential mechanisms of action. *Toxicol. Lett.* **2002**, *133*, 1.
[7] A. C. Schmidt, J. Koppelt, M. Neustadt, M. Otto. Mass spectrometric evidence for different complexes of peptides and proteins with arsenic(III), arsenic(V), copper(II) and zinc(II) species. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 153.
[8] A. C. Schmidt, M. Neustadt, M. Otto. Quantitative evaluation of the binding of phenylarsenic species to glutathione, isotocin, and thioredoxin by means of electrospray ionization time-of-flight mass spectrometry. *J. Mass Spectrom.* **2007**, *42*, 771.
[9] S. G. Park, D. J. Butcher. Investigation of the interaction between arsenic species and thiols via electrospray ionization tandem mass spectrometry. *Microchem. J.* **2010**, *95*, 57.
[10] Z. Wang, H. Zhang, X. F. Li, X. C. Le. Study of interactions between arsenicals and thioredoxins (human and E. coli) using mass spectrometry. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 3658.
[11] H. Yan, N. Wang, M. Weinfeld, W. R. Cullen, X. C. Le. Identification of arsenic-binding proteins in human cells by affinity chromatography and mass spectrometry. *Anal. Chem.* **2009**, *81*, 4144.
[12] K. Bluemlein, A. Raab, J. Feldmann. Stability of arsenic peptides in plant extracts: off-line versus on-line parallel elemental and molecular mass spectrometric detection for liquid chromatographic separation. *Anal. Bioanal. Chem.* **2009**, *393*, 357.
[13] A. C. Schmidt, S. Steier. Some critical aspects in the determination of binding constants by electrospray ionization mass spectrometry at the example of arsenic bindings to sulfur containing biomolecules. *J. Mass Spectrom.* **2010**, *45*, 870.
[14] A. C. Schmidt, B. Fahlbusch, M. Otto. Size exclusion chromatography coupled to electrospray ionization mass spectrometry for analysis and quantitative characterization of arsenic interactions with peptides and proteins. *J. Mass Spectrom.* **2009**, *44*, 898.
[15] A. C. Schmidt, K. Mickein. Optimization of peptide and protein separation with a monolithic reversed-phase column and application to arsenic-binding studies. *J. Chromatogr. A* **2011**, *1218*, 280.
[16] N. W. Smith, Z. Jiang. Developments in the use and fabrication of organic monolithic phases for use with high-performance liquid chromatography and capillary electrochromatography. *J. Chromatogr. A* **2008**, *1184*, 416.
[17] S. Xie, R. W. Allington, F. Svec, J. M. J. Fréchet. Rapid reversed-phase separation of proteins and peptides using optimized 'moulded' monolithic poly(styrene-co-divinylbenzene) columns. *J. Chromatogr. A* **1999**, *856*, 169.
[18] K. T. Kitchin, K. Wallace. Arsenite binding to synthetic peptides based on the Zn finger region of the human estrogen receptor-alpha. *Toxicol. Appl. Pharmacol.* **2005**, *206*, 66.
[19] K. T. Kitchin, K. Wallace. Arsenite binding to synthetic peptides: the effect of increasing length between two cysteines. *J. Biochem. Mol. Toxicol.* **2006**, *20*, 35.
[20] D. E. Carter, H. V. Aposhian, A. J. Gandolfi. The metabolism of inorganic arsenic oxides, gallium arsenide, and arsine: a toxicochemical review. *Toxicol. Appl. Pharmacol.* **2003**, *193*, 309.
[21] E. Burmeister Getz, M. Xiao, T. Chakraborty, R. Cooke, P. R. Selvin. A comparison between the sulfhydryl reductants tris(2-carboxyethyl) phosphine and dithiothreitol for use in protein biochemistry. *Anal. Biochem.* **1999**, *273*, 73.
[22] M. Kastner. *Protein Liquid Chromatography. J. Chromatogr., Library vol. 61*, Elsevier: Amsterdam, **2000**.
[23] K. D. Nugent, W. G. Burton, T. K. Slattery, B. F. Johnson, L. R. Snyder. Separation of proteins by reversed-phase high-performance liquid chromatography—II. Optimizing sample pre-treatment and mobile phase conditions. *J. Chromatogr.* **1988**, *443*, 381.
[24] C. David, S. Foley, M. Enescu. Protein S-S bridge reduction: a raman and computational study of lysozyme interaction with TCEP. *Phys. Chem. Chem. Phys.* **2009**, *11*, 2532.
[25] T. Sienkiewicz. Nomenklatur und einige Eigenschaften der Molkenproteine. 2. Mitteilung: α-lactalbumin, Immunoglobuline, Protease-Peptone, Minorproteine und Enzyme. *Die Nahrung* **1981**, *25*, 335.
[26] T. Sienkiewicz. Nomenklatur und einige Eigenschaften der Molkenproteine. 1. Mitteilung: ß-Lactoglobulin. *Die Nahrung* **1981**, *25*, 329.
[27] K. Kuwajima, M. Ikeguchi, T. Sugawara, Y. Hiraoka, S. Sugai. Kinetics of disulfide bond reduction in α-lactalbumin by dithiothreitol and molecular basis of superreactivity of the Cys6-Cys120 disulfide bond. *Biochemistry* **1990**, *29*, 8240.
[28] A. Tolkach, U. Kulozik. Kinetische Modellierung der thermischen Denaturierung von α-Lactalbumin im sauren pH-Bereich und in Anwesenheit eines Calcium-Komplexbildners. *Chem. Ing. Tech.* **2008**, *80*, 1165.
[29] E. A. Permyakov, L. J. Berliner. α-Lactalbumin: structure and function. *FEBS Lett.* **2000**, *473*, 269.
[30] I. Losito, T. Carbonara, L. Manaci, F. Palmisano. Evaluation of the thermal history of bovine milk from the lactosylation of whey proteins: an investigation by liquid chromatography-electrospray ionization mass spectrometry. *Anal. Bioanal. Chem.* **2007**, *389*, 2065.

 Copyright © 2012 John Wiley & Sons, Ltd.