Enzymatic digestion and chromatographic analysis of arsenic species released from proteins

Meiling Lu a, Hailin Wang b, Xing-Fang Li a, Xiufen Lu a, X. Chris Le a,∗

a Department of Laboratory Medicine and Pathology, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Alberta, T6G 2C3, Canada
b State Key Laboratory for Environmental Chemistry and Eco-toxicology, Research Centre for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, 100085, China

ARTICLE INFO

Article history:
Received 22 December 2008
Received in revised form 4 March 2009
Accepted 6 March 2009
Available online 13 March 2009

Keywords:
Reactive arsenic metabolites
Protein–arsenic complex
Speciation
Protease digestion
Protein-bound arsenic species
Mass-spectrometry

ABSTRACT

A method combining gel filtration chromatography (GFC), protease digestion, and ion pair chromatography with inductively coupled plasma mass spectrometry detection was developed for the determination of arsenic species bound to proteins. The method was first established by examining the interactions of two model proteins, metallothionein (MT) and hemoglobin, with three reactive trivalent arsenic species. It was then successfully applied to the speciation of arsenic in red blood cells of rats. Inorganic arsenite (iAs III ), monomethylarsonous acid (MMA III ), and dimethylarsinous acid (DMA III ) were efficiently released from the proteins by protease digestion at pH 8.0, with the recovery ranging from 93% to 106%. There was no oxidation of iAs III or MMA III during the protease digestion process. Up to 61% DMA III (the least stable arsenic species) was unchanged, and the rest was oxidized to the pentavalent dimethylarsinic acid (DMA V). The arsenic species in the red blood cells of control rats was present as DMA III complex with hemoglobin. The method enabling the determination of the specific arsenic species that bind to cellular proteins is potentially useful for studying arsenic distribution, metabolism, and toxicity.

1. Introduction

Millions of people around the world are exposed to high levels of inorganic arsenic from drinking water [1]. Inorganic arsenic can be metabolized in the body to chemically reactive trivalent monomethylarsonous acid (MMA III ) and dimethylarsinous acid (DMA III ) in addition to the less reactive pentavalent monomethylarsonic acid (MMA V) and dimethylarsinic acid (DMA V) [2,3]. These various arsenic species have very different toxicological profiles. Generally, MMA V and DMA V are less toxic than inorganic arsenite (iAs III ) and arsenate (iAs V) [4]. However, their trivalent analogs, MMA III and DMA III, are more reactive and toxic by inducing or enhancing DNA damage [5–8], causing cell death [9–11], inhibiting cellular enzymes [8,12–15], and causing oxidative stress [16–18].

Understanding the health effects arising from exposure to arsenic requires speciation analysis of arsenic metabolites. Analytical techniques often involve chromatographic separation of the individual arsenic species followed by specific detection of arsenic [19–25]. High performance liquid chromatography (HPLC) linked with inductively coupled plasma mass spectrometry (ICPMS) has been widely used for speciation of trace levels of arsenic [19,21–24]. Arsenic metabolites in urine can be directly analyzed by using HPLC–ICPMS to provide useful information for the assessment of human exposure to arsenic [22,23]. There is no or minimum treatment of urinary samples prior to HPLC–ICPMS analysis. However, for the analysis of arsenic species in most other biological fluids and tissues, appropriate treatments of samples are necessary. Methods involving ultrasonic and/or microwave digestion with strong acids have been generally applied [22,25–27]. Strong acids could change the actual arsenic species present in the original sample. Mild conditions, on the other hand, may not be appropriate to extract arsenic species efficiently from the tissue samples. There has been much research in improving sample treatment methods for arsenic speciation analysis [28–30].

Trivalent arsenic species, iAs III, MMA III and DMA III can form protein–arsenic complexes because of their high affinity for sulfhydryl groups in proteins [31–34]. Unlike DNA adducts that are normally repaired in vivo, protein complexes are not repaired and have a longer lifetime in the body. Therefore, the arsenic species bound to proteins may be an important pool of reactive arsenic species in biological systems. Identifying what arsenic species are bound to the proteins is necessary to an understanding of the distribution and metabolism of arsenic in the body. However, releasing arsenic species from proteins by acid digestion cannot preserve the valence state of arsenic species. In preliminary analysis of digested red blood cell and plasma samples from patients under arsenic trioxide treatment, we were not able to observe any trivalent arsenic species when the samples were digested with nitric acid. The absence of the trivalent arsenic species might be due to
two factors: the low recovery or the conversion of these arsenic species. The trivalent arsenic species, especially DMA$^{III}$, are readily oxidized to their pentavalent forms [35]. Likewise, after treatment of human plasma and red blood cell samples with trichloroacetic acid (TCA), there were no trivalent arsenicals detected by the subsequent HPLC–ICPMS analysis, although trivalent arsenic species were present in urine samples that were not treated [22].

To identify the arsenic species bound to abundant proteins, we previously used electrospray ionization tandem mass spectrometry (ESI–MS/MS) [32,33]. We also used collision-induced dissociation and the unique mass defect of arsenic for the identification of arsenic binding to proteins [36]. Although this method is useful, it is very challenging to identify the arsenic species bound to proteins present at trace levels. An alternate approach is to release the arsenic species from proteins followed by highly sensitive HPLC–ICPMS analysis. A crucial requirement is to maintain the valence state and methylation status of arsenic species. The objective of this study was to develop a method combining protease digestion, HPLC separation, and ICPMS detection for the speciation analysis of arsenic bound to proteins. We achieve this objective first by testing model systems involving metallothionein (MT), hemoglobin, and three trivalent arsenicals, and then by applying the method to the analysis of arsenic species in red blood cells of rats. We demonstrate here that the method can be used to determine arsenic species bound to proteins.

2. Experimental

Caution. The arsenic species included in this study are toxic and are potential human carcinogens. Caution and care should be exercised when handling these materials.

2.1. Standards and analytical reagents

Four arsenic standards including sodium arsenite (iAs$^{III}$), sodium arsenate (iAs$^{V}$), MMA$^{V}$, and DMA$^{V}$, and standard proteins including rabbit metallothionein II, rat hemoglobin (rat Hb), and human serum albumin (HSA) were all purchased from Sigma–Aldrich (St. Louis, MO). Two reactive trivalent arsenic species, MMA$^{III}$ and DMA$^{III}$ were kindly provided by Dr. W.R. Cullen (University of British Columbia, Vancouver, Canada). F344 rats were fed with nor-
We chose hemoglobin, HSA, and metallothionein to represent proteins containing different types of cysteine residues as schematized in Fig. 1A: free cysteines, cysteines forming disulfide bonds, and cysteines bound to metal ions. Hemoglobin (Hb) contains several free cysteine residues; HSA contains a number of disulfide bonds; MT has about 20 cysteines that are usually chelated with zinc and cadmium ions. With the release of zinc and cadmium from MT by strong chelation reagent such as EDTA, the resultant apoMT contains many free cysteine residues close to one another. The adjacent free cysteines in apoMT could form complexes with iAs^{III} and MMA^{III}. Hemoglobin can form non-chelation type complex with DMA^{III} [31,32]. Therefore, the trivalent arsenic species having one, two, or three hydroxyl groups (Fig. 1B) and the selected three proteins represent a wide spectrum of possible interactions.

3.2. Interaction of model proteins with trivalent arsenic species

Having selected the model proteins, we studied their binding with each trivalent arsenic species. Fig. 2 shows chromatograms from GFC separation and ICP-MS analyses of reaction mixtures containing 1 μM arsenicals and 20 μM proteins. The protein-bound arsenic was excluded from the GFC column and eluted within 2 min, and the unbound arsenic species had retention times of

---

Fig. 1. Schematic diagram showing the representative cysteines in the three model proteins (A) and the molecular structures of the three reactive trivalent arsenic species (B).

Fig. 2. Chromatograms showing that different proteins have various binding affinity to the three trivalent arsenic species. (A) HSA; (B) MT; (C) rat Hb; and (D) apoMT. Trace (i) was from the analysis of the protein without added arsenic. The other three chromatographic traces in each graph were from the analysis of incubation mixtures containing the protein and iAs^{III} trace (ii), MMA^{III} trace (iii), and DMA^{III} trace (iv). The chromatograms, except for the bottom traces, were manually offset on the vertical axis for visual clarity.
4–6 min. Fig. 2A shows that HSA did not bind to iAsIII, MMAIII, or DMAIII. This is consistent with the fact that HSA does not contain a free cysteine. Similarly, the untreated MT did not bind to iAsIII or DMAIII (Fig. 2B traces ii and iv). Approximately 8.1% of total MMAIII (Fig. 2B, trace iii) was bound to MT, probably due to competition between MMAIII and the metal ions in the MT. However, after the metal ions were released from MT by EDTA chelation, the apoMT showed substantial increases in its binding to all the three trivalent arsenicals (Fig. 2D). Approximately 91% MMAIII and 59% iAsIII were bound to apoMT (traces ii and iii in Fig. 2D). Rat hemoglobin, having several free cysteines, showed predominant binding to DMAIII as well as some binding with MMAIII (traces iii and iv in Fig. 2C). The percentage of protein-bound arsenic accounted for 61% DMAIII and 29% MMAIII that were in the reaction mixtures. The percentage of protein-bound arsenic was obtained from the ratio of peak area corresponding to the protein-bound arsenic species over the total peak areas of both the protein-bound and free arsenic species.

We further measured the apparent binding constants of apoMT to trivalent arsenic species using the same method as described previously [32,33], and compared these values to those for rHb (Table 1). Among all the six protein–arsenic complexes, apoMT-MMAIII has the highest binding constant. MMAIII has two hydroxyl groups available to bind with two cysteines. The strong binding between the apoMT and MMAIII is probably due in part to the chelation of two cysteine residues in apoMT to the same MMAIII molecule. In the presence of EDTA, the metal ions such as zinc and cadmium in MT were chelated with EDTA, resulting in the free cysteine residues. Most of these cysteine residues are close to each other in the primary sequence (about four amino acids apart). Therefore, the adjacent two cysteines in apoMT are able to form a stable complex with a MMAIII molecule (with the loss of two H2O molecules). ApoMT can also form complex with iAsIII with higher affinity, showing that there were 2–3 cysteine residues close to each other. DMAIII has only one hydroxyl group available to bind with one cysteine. The complex is less stable without the chelating effect.

### 3.3. Purification of protein–arsenic complexes

Using GFC–ICPMS, we monitored the binding between apoMT and iAsIII (Fig. 3A) or MMAIII (Fig. 4A). In subsequent GFC separation (without ICPMS detection) of the reaction mixtures, we collected the fractions corresponding to the MT-iAsIII complex (Fig. 3A) and the MT-MMAIII complex (Fig. 4A) in the time window of 1.5–2.5 min. Protein-bound arsenic in each fraction was quantified by fluorescence ICPMS, and the fractions with most protein-bound arsenic (0.2 μM iAsIII or 0.4 μM MMAIII in Table 2) were used for the subsequent speciation experiments described below.

### 3.4. Speciation of arsenic released from the apoMT fractions

The apoMT-iAsIII protein fractions obtained above were digested with protease for 30 min, and the digest was analyzed for arsenic species using ion pair chromatography ICPMS [20]. Three trivalent arsenic species and their pentavalent analogs (20 μg/L or 0.27 μM each) in a standard mixture were well resolved from one another (top trace of Fig. 4B). Analysis of the digested apoMT-iAsIII solution showed a major peak appearing at 1.6 min (Fig. 3B, middle trace), corresponding to iAsIII. This peak was absent from the digested solution of apoMT itself (Fig. 3B, bottom trace). The detection of iAsIII released from the apoMT fraction by protease digestion suggests that iAsIII was the arsenic species originally bound to the apoMT. These results demonstrate that protease digestion did not signif-

### Table 1

<table>
<thead>
<tr>
<th>Species</th>
<th>$K_b$ (M$^{-1}$)</th>
<th>Rat Hb [32]</th>
<th>ApoMT</th>
</tr>
</thead>
<tbody>
<tr>
<td>iAsIII</td>
<td>$2.3 \times 10^4$</td>
<td>$7.2 \times 10^4$</td>
<td></td>
</tr>
<tr>
<td>MMAIII</td>
<td>$4.7 \times 10^4$</td>
<td>$5.0 \times 10^3$</td>
<td></td>
</tr>
<tr>
<td>DMAIII</td>
<td>$2.2 \times 10^3$</td>
<td>$2.9 \times 10^1$</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>Model system</th>
<th>Total complex (μM)</th>
<th>iAsIII (μM)</th>
<th>MMAIII (μM)</th>
<th>MMAV (μM)</th>
<th>DMAIII (μM)</th>
<th>DMAV (μM)</th>
<th>Total recovery (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ApoMT-iAsIII</td>
<td>0.21±0.01</td>
<td>0.22±0.01</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>106±6</td>
</tr>
<tr>
<td>ApoMT-MMAIII</td>
<td>0.43±0.04</td>
<td>0.36±0.02</td>
<td>0.05±0.01</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>95±1</td>
</tr>
<tr>
<td>rHb-DMAIII</td>
<td>0.88±0.02</td>
<td>0.90±0.06</td>
<td>0.54±0.01</td>
<td>0.32±0.01</td>
<td>0.37±0.12</td>
<td>93±3</td>
<td></td>
</tr>
<tr>
<td>rHb-DMAIII/MMMAIII</td>
<td>1.99±0.13</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>
Fig. 4. (A) Chromatograms from GFC–ICPMS analysis showing separation of the apoMT-MMAIII complex from the unbound MMAIII; (B) chromatograms from ion pair chromatographic separation and ICPMS detection of arsenic species released from the apoMT-MMAIII complex, showing that MMAIII was the species bound to apoMT. The chromatograms, except for the bottom traces, were manually shifted up on the vertical axis for visual clarity.

Fig. 5. (A) Chromatograms from GFC–ICPMS analyses of lysate of rat red blood cells (bottom trace) and the lysate incubated with MMAIII in vitro for 2 h (top trace); (B) chromatograms from ion pair chromatographic separation and ICPMS analyses of arsenic species released from rHb. The chromatograms, except for the bottom traces, were manually offset on the vertical axis for visual clarity.

3.5. Speciation of arsenic bound to hemoglobin in rat red blood cells

Having demonstrated the speciation of arsenic released from the purified protein–arsenic complexes, we further applied the technique to speciation of arsenic in red blood cells of rats. The F344 rats were exposed to background levels of arsenic from the normal diet and drinking water. Red blood cells of five such rats were collected and lysed in 20 mM ammonium acetate (pH 7.0). The cell debris was removed by centrifugation, and the supernatant was then subjected to GFC fractionation. Hemoglobin accounted for ~95% of total protein in red blood cells. As shown in Fig. 5A (bottom trace), there is clearly a protein-bound arsenic peak present in the GFC–ICPMS chromatogram from the analysis of rat Hb. After collecting the protein fractions and subjecting them to protease digestion for 30 min, we analyzed the released arsenic species by HPLC–ICPMS. Two major arsenic peaks appeared in the chromatogram (Fig. 5B, bottom trace), which corresponded to DMAV (39%) and DMAIII (61%). Unlike DMAIII, DMAV is not known to bind to any proteins [31]. Given the oxidative instability of DMAIII [35], the presence of DMAV is most likely due to the oxidation of DMAIII during sample processing. There were no monomethyl or inorganic arsenicals observed in the protein digest. Therefore, it was concluded that DMAIII was the arsenic species bound to hemoglobin in the rats.

To test whether hemoglobin extracted from the rats could bind to other arsenic species, we further incubated the lysate of red blood cells with MMAIII in vitro and analyzed the mixture. We found that the protein–arsenic complexes increased by 1.3-fold, and the
protein-bound arsenic accounted for about 86% of total arsenic in the reaction solution. After protease digestion of this protein fraction, three arsenic species were observed, which corresponded to MMAIII, DMAIII and DMAV (Fig. 5B, middle trace). These results indicate that MMAIII can react with rat Hb and form protein-MMAIII complex in vitro although in vivo DMAIII is the predominant arsenic species bound to rat Hb.

We also exposed fresh rat red blood cells to MMAIII overnight, extracted the hemoglobin and analyzed it by nanoESI-MS as described previously [32]. We found that MMAIII could indeed be taken up by red blood cells and form a complex with the α unit of rat hemoglobin (Supplementary Information Figure S1).

3.6. Effect of digestion time on the release of arsenic from proteins

To further refine the technique of releasing arsenic from proteins, we optimized the digestion time and tested its effect on the stability and recovery of arsenic species. As shown in Fig. 6, with an increase of the enzymatic digestion time from 30 min to 60 min and 90 min, the peak area of released iAsIII did not change significantly. The relative standard deviation of peak areas between the three time points was ~6%, which was within the range of experimental error. Similar results were obtained from two other binding systems: apoMT-MMAIII and rat red blood cells lysate with MMAIII. In the case of releasing DMAIII from rat red blood cells, the arsenic speciation profile showed a slight increase in the percentage of DMAV and decrease in the percentage of DMAIII with the increase of digestion time. Therefore, in order to better preserve the unstable DMAIII species, a digestion time of 30 min was selected as a compromise between the recovery and stability of arsenic species.

3.7. Recovery of arsenic released from protein complex by protease digestion

In addition to determining the concentration of arsenic species released from the proteins, we also measured total arsenic in the digested solution, and compared with the total protein-bound arsenic in the purified protein–arsenic complexes. The recovery of arsenic species (quantity of arsenic measured after releasing from protein over the quantity of protein–bound arsenic applied) ranged from 93% to 106% (Table 2). Even for the protein complex (apoMT-MMAIII) with the strongest binding, the recovery reached 95%, demonstrating the possibility to apply this method to the more complicated sample analysis. The complete release of arsenic species from the proteins after a 30 min protease digestion is consistent with the observation of no protein-bound arsenic in Figs. 3B, 4B and 5B that would otherwise appear at the void time (<1.5 min).

4. Discussion and conclusions

In summary, a method for determining arsenic species bound to proteins was developed and was successfully applied to detect the trace amount of protein-bound arsenic species in the red blood cells of rats. The recoveries for iAsIII, MMAIII, and DMAIII from both apoMT and hemoglobin systems were quantitative, and there was no conversion of iAsIII or MMAIII and minimum oxidation of DMAIII to DMAV. The proposed method has the potential to be used for analysis of arsenic metabolites bound to proteins in other tissue samples. The quantitative information on the protein-bound arsenic species will contribute to critical evaluations of arsenic distribution, metabolism, and toxicity.

Arsenic binding to proteins may play an important role in arsenic toxicity as arsenic compounds can inhibit the activity of several enzymes [8,12–15]. To understand dosimetry and biochemical effects arising from arsenic binding, it is crucial to determine the specific arsenic species that bind to cellular proteins. The method presented here allows the release of arsenic species from proteins followed by chromatographic separation and ICPMS detection of arsenic species.

To reliably determine the concentrations of arsenic species in biological system, some of the key considerations include achieving adequate recovery and maintaining the arsenic species unchanged before and during sample processing. The instability of some species such as MMAIII and DMAIII makes the task of preserving these species very challenging. In the previous arsenic speciation analysis in blood, these reactive arsenic species have not been detected, probably because oxidative changes of these species could occur during the extensive sample extraction and treatment processes [22]. Glutathione–arsenic complex was found not stable at alkaline pH [38]. We also found that sulfur-arsenic bond in the peptides and proteins was less stable under alkaline conditions, but the free DMAIII was more stable under such conditions (unpublished data). Therefore, we incorporated protease digestion under weak alkaline conditions to release arsenic species from proteins for the subsequent HPLC–ICPMS analysis of arsenic species. The combination of high efficiency of the protease, the mild alkaline conditions, and the short digestion time facilitated the preservation of iAsIII.
and MMAIII. We have achieved recoveries ranging from 93% to 106%. These recoveries were obtained from testing not only the standard protein–arsenic complexes but also the protein–arsenic complexes in red blood cells.

Acknowledgements

The authors thank Dr. William Cullen of the University of British Columbia for providing the MMAIII and DMAIII standards, Dr. Samuel M. Cohen and Ms. Lora Arnold of the University of Nebraska Medical Center for providing the rat red blood cells. This work was supported in part by Canadian Institutes of Health Research, Canadian Water Network, Metals in the Human Environment Strategic Network, Canada Research Chairs Program, Alberta Water Research Institute, and Alberta Health and Wellness.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.chroma.2009.03.015.

References