Identification of Arsenic−Binding Proteins in Human Cells by Affinity Chromatography and Mass Spectrometry

ARTICLE *in* ANALYTICAL CHEMISTRY · MAY 2009

Impact Factor: 5.64 · DOI: 10.1021/ac900352k · Source: PubMed

5 AUTHORS, INCLUDING:

Nan Wang
University of Alberta
22 PUBLICATIONS **428** CITATIONS
SEE PROFILE

Michael Weinfeld
University of Alberta
146 PUBLICATIONS **6,250** CITATIONS
SEE PROFILE

X. Chris Le
University of Alberta
248 PUBLICATIONS **7,753** CITATIONS
SEE PROFILE

Technical Note

Identification of Arsenic-Binding Proteins in Human Cells by Affinity Chromatography and Mass Spectrometry

Huiming Yan, Nan Wang, Michael Weinfeld, William R. Cullen, and X. Chris Le

Anal. Chem., **2009**, 81 (10), 4144-4152• DOI: 10.1021/ac900352k • Publication Date (Web): 16 April 2009

Anal. Chem. **2009,** *81,* 4144–4152

Identification of Arsenic-Binding Proteins in Human Cells by Affinity Chromatography and Mass Spectrometry

Huiming Yan,[†] Nan Wang,[†] Michael Weinfeld,[‡] William R. Cullen,[§] and X. Chris Le*[,†,||]

Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2G3, Department of Oncology, University of Alberta, Cross Cancer Institute, 11560 University Avenue, Edmonton, Alberta, Canada T6G 1Z2, Department of Chemistry, University of British Columbia, 2036 Main Mall, Vancouver, British Columbia, Canada V6T 1Z1, and Department of Laboratory Medicine and Pathology, 10-102 Clinical Sciences Building, University of Alberta, Edmonton, Alberta, Canada, T6G 2G3

Exposure to high levels of arsenic can cause a wide range of health effects, including cancers of the bladder, lung, skin, and kidney. However, the mechanism(s) of action underlying these deleterious effects of arsenic remains unclear. Arsenic binding to cellular proteins is a possible mechanism of toxicity, and identifying such binding is analytically challenging because of the large concentration range and variety of proteins. We describe here an affinity selection technique, coupled with mass spectrometry, to select and identify specific arsenic-binding proteins from a large pool of cellular proteins. Controlled experiments using proteins either containing free cysteine(s) or having cysteine blocked showed that the arsenic affinity column specifically captured the proteins containing free cysteine(s) available to bind to arsenic. The technique was able to capture and identify trace amounts of bovine biliverdin reductase B present as a minor impurity in the commercial preparation of carbonic anhydrase II, demonstrating the ability to identify arsenic-binding proteins in the presence of a large excess of non-specific proteins. Application of the technique to the analysis of subcellular fractions of A549 human lung carcinoma cells identified 50 proteins in the nuclear fraction, and 24 proteins in the membrane/organelle fraction that could bind to arsenic, adding to the current list of only a few known arsenic-binding proteins.

Arsenic compounds occur naturally in the environment. Human exposure to high levels of arsenic from drinking water can cause a wide range of health effects,[1−3] most seriously, cancers of bladder, lung, urinary tract, and skin.[4−6] The biochemical mechanisms responsible for these effects caused by arsenic remain unclear,[7] but may be mediated by the binding of trivalent arsenicals to thiol groups in proteins, thereby changing the conformation of these proteins and inhibiting their functions. If some of the affected proteins are responsible for cellular repair of DNA damage, for example, the inhibition of these proteins could lead to carcinogenesis.

Recent studies have shown binding of trivalent arsenicals to cysteines in proteins. Hemoglobin,[8] metallothionein,[9,10] galectin-1 and thioredoxin peroxidase II,[11,12] ArsR protein,[13] GLUT4,[14−17] tubulin and Actin[18] have been demonstrated to bind to trivalent arsenic species, including inorganic arsenite and its methylated metabolites monomethyarsonous acid (MMA^{III}) and dimethylarsinous acid (DMA^{III}). In principle, DMA^{III} [$(CH_3)_2AsOH$] can bind to a single thiol group; MMA^{III} [$CH_3As(OH)_2$] can bind to two thiols; and inorganic arsenite [$As(OH)_3$] can bind to a

* To whom correspondence should be addressed. E-mail: xc.le@ualberta.ca.
† Department of Chemistry, University of Alberta.
‡ Department of Oncology, University of Alberta.
§ Department of Chemistry, University of British Columbia.
|| Department of Laboratory Medicine and Pathology, University of Alberta.

(1) National Research Council, *Arsenic in Drinking Water*; National Academy Press: Washington, DC, 1999; and 2001 Update.
(2) Tseng, C.-H.; Tseng, C.-P.; Chiou, H.-Y.; Hsueh, T.-M.; Chong, C.-K.; Chen, C.-J. *Toxicol. Lett.* **2002**, *133*, 69–76.
(3) Chen, C.-J.; Chuang, Y.-C.; Lin, T.-M.; Wu, H.-Y.; Tseng, C.-H. *Cancer Res.* **1985**, *45*, 5895–5899.
(4) Chiou, H.-Y.; Hsueh, Y.-M.; Liaw, K.-F.; Horng, S.-F.; Chiang, M.-H.; Pu, Y.-S.; Lin, J. S.-N.; Huang, C.-H.; Chen, C.-J. *Cancer Res.* **1995**, *55*, 1296–1300.
(5) Smith, A. H.; Goycolea, M.; Haque, R.; Biggs, M. L. *Am. J. Epidemiol.* **1998**, *147*, 660–669.
(6) Cantor, K. P.; Lubin, J. H. *Toxicol. Appl. Pharmacol.* **2007**, *222*, 252–257.
(7) Kitchin, K. T.; Wallace, K. *Toxicol. Appl. Pharmacol.* **2005**, *206*, 66–72.
(8) Lu, M.; Wang, H.; Li, X.-F.; Lu, X.; Cullen, W. R.; Arnold, L. L.; Cohen, S. M.; Le, X. C. *Chem. Res. Toxicol.* **2004**, *17*, 1733–1742.
(9) Jiang, G.; Gong, Z.; Li, X.-F.; Cullen, W. R.; Le, X. C. *Chem. Res. Toxicol.* **2003**, *16*, 873–880.
(10) Merrifield, M. E.; Ngu, T.; Stillman, M. J. *Biochem. Biophys. Res. Commun.* **2004**, *324*, 127–132.
(11) Lin, C.-H.; Huang, C.-F.; Chen, W.-Y.; Chang, Y.-Y.; Ding, W.-H.; Lin, M.-S.; Wu, S.-H.; Huang, R.-N. *Chem. Res. Toxicol.* **2006**, *19*, 469–474.
(12) Chang, K. N.; Lee, T. C.; Tam, M. F.; Chen, Y. C.; Lee, L. W.; Lee, S. Y.; Lin, P. J.; Huang, R. N. *Biochem. J.* **2003**, *371*, 495–503.
(13) Shi, W.; Dong, J.; Scott, R. A.; Ksenzenko, M. Y.; Rosen, B. P. *J. Biol. Chem.* **1996**, *271*, 9291–9297.
(14) Hoffman, R. D.; Lane, M. D. *J. Biol. Chem.* **1992**, *267*, 14005–14011.
(15) Rey, N. A.; Howarth, O. W.; Pereira-Maia, E. C. *J. Inorg. Biochem.* **2004**, *98*, 1151–1159.
(16) Spuches, A. M.; Kruszyna, H. G.; Rich, A. M.; Wilcox, D. E. *Inorg. Chem.* **2005**, *44*, 2964–2972.
(17) Raab, A.; Meharg, A. A.; Jaspars, M.; Genney, D. R.; Feldmann, J. *J. Anal. At. Spectrom.* **2004**, *19*, 183–190.
(18) Menzel, D. B.; Hamadeh, H. K.; Lee, E.; Meacher, D. M.; Said, V.; Rasmussen, R. E.; Greene, H.; Roth, R. N. *Toxicol. Lett.* **1999**, *105*, 89–101.

10.1021/ac900352k CCC: $40.75
Published on Web 04/16/2009

maximum of three thiols in the protein.[9,19−21] However, studies aimed at characterizing arsenic-binding proteins from treated cells have only been able to identify a few highly abundant proteins (galectin-1, thioredoxin peroxidase, GLUT4, tubulin and Actin).[12−14,18,22] It is possible that a large variety of potential arsenic-binding proteins present at lower concentrations are undetectable by current analytical techniques. Identification of low-abundance arsenic-binding proteins in the presence of a large excess of abundant proteins in cells would be facilitated by the development of new affinity media.

We describe here a method that combines an improved arsenic-affinity selection medium with tandem mass spectrometry. With the improvement of the arsenic immobilization efficiency, many of the arsenic binding proteins in cell lysates can be captured for mass spectrometry analysis. We further demonstrate the application of this affinity technique to the identification of arsenic-binding proteins in subcellular fractions of A549 human lung carcinoma cells.

EXPERIMENTAL SECTION

Materials. Bovine serum albumin (BSA), human serum albumin (HSA), carbonic anhydrase II, transferrin, dithiothreitol (DTT), iodoacetamide (IAA), Triton X-100, sodium dodecyl sulfate (SDS), citric acid, sodium chloride, silver nitrate, 2,3-dimercapto-1-propanol (BAL), 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB), formaldehyde solution (37%), trifluoroacetic acid (TFA), dimethylsulfate oxide (DMSO), benzylamine, ammonium bicarbonate, trichloroacetic acid (TCA), formic acid (HPLC grade, Fluka), and Eupergit C (Fluka) were purchased from Sigma-Aldrich Canada (Oakville, ON). HPLC-grade acetonitrile and acetone, HPLC-grade acetic acid, calcium chloride, modified trypsin (Promega), hydrochloric acid and HEPES buffer were from Fisher Scientific Canada (Ottawa, ON). ProteoExtract subcellular proteome extraction kit was purchased from CALBIOCHEM (San Diego, CA). RC DC protein assay was from Bio-Rad (Hercules, CA). 4-Aminophenylarsine oxide ($NPAO^{III}$) was synthesized by reducing arsanillic acid (Sigma-Aldrich) as previously described.[23]

Instrumentation. A PerkinElmer 200 series HPLC system (PE Instruments, Norwalk, CT, U.S.A.), equipped with a pump and an autosampler, was used with an Elan 6100 DRC plus ICPMS (PE/Sciex, Toronto, ON, Canada). A Biosep-SEC-S 2000 column (300 × 4.6 mm, Phenomenex, Torrance, CA) and a ZORBAX GF-250 column (250 × 4.6 mm, Agilent) were used for separation of protein-bound and unbound arsenic species. The following conditions of ICPMS were used: rf power (1150 W), plasma gas flow (13 L/min), auxiliary gas flow (1.1 L/min), and nebulizer gas flow (0.79 L/min). A QStar Pulsar-i mass spectrometer (Applied Biosystems) equipped with a nanospray ionization source was used for the identification of proteins.

Preparation of Affinity Columns. Two arsenic affinity media were prepared by reaction of Eupergit C beads with either 4-aminophenylarsine oxide ($NPAO^{III}$) or arsenite (As^{III}). 4-Aminophenylarsine oxide (0.1 g) was dissolved in 1 mL DMSO (pH adjusted to 2 by HCl), mixed with 3 mL deionized water, and the solution pH was adjusted to 4 by NaOH solution. The solution was poured into a gravity column (Bio-Rad) containing 0.5 g reactive Eupergit C beads. Oxygen in the mixture was removed by purging with nitrogen for 10 min. The column was then sealed, and shaken slowly at room temperature for 24 h. The affinity column was washed sequentially with 500 mL of 100 mM NaCl solution, 500 mL of deionized water, 200 mL of 50% acetonitrile solution, and 500 mL of deionized water. The effluent from the column was analyzed for arsenic using ICP-MS, and the repeated washing continued until no arsenic (<0.01 μg/L) was detectable in the washing solution. The affinity column containing arsenite was prepared similarly, except that 0.07 g of $NaAsO_2$ in 4 mL of water (pH 7.5) was used instead of $NPAO^{III}$.

The amount of $NPAO^{III}$ and As^{III} immobilized were determined by the amount of BAL reacted with the trivalent arsenicals (see Supporting Information, Figure S2). Free BAL can reduce DTNB immediately, and the reduced product thionitrobenzoate ion can be detected at 412 nm by spectrometry[22] (Bio-Rad Smartspec 3000).

Benzylamine instead of 4-aminophenylarsine oxide was reacted with Eupergit C to prepare a non-specific control column without arsenic.

Test of Specificity of Affinity Columns. The specificity of the affinity column was tested by proteins containing free cysteine or no free cysteine. Human serum albumin (HSA) and bovine serum albumin (BSA) with free cysteine were used as positive control and applied to the arsenic affinity columns. Negative controls were performed from HSA and BSA, which were carbamidomethylated with dithiothreitol (DTT) and iodoacetamide (IAA) to block the free cysteines. Briefly, approximately 1 mg HSA/BSA was dissolved in 1 mL of 50 mM NH_4HCO_3 containing 1% SDS, to which solution 10 mM DTT was added. After 30 min incubation at 56 °C, IAA solution was added until its final concentration was 50 mM. After incubation in the dark for 1 h, protein was precipitated with acetone. Protein transferrin, in which all cysteines exist in disulfide bonds, provided another negative control. The membrane protein fraction of A549 cells was also applied to the non-specific control column to test the hydrophobic interaction between the media and proteins. Carbonic anhydrase II, which lacks a cysteine residue, was used to optimize the affinity technique parameters and to test the capture of trace amounts of thiol-containing proteins in the presence of large excess of carbonic anhydrase.

Determination of Protein-Bound and Unbound Arsenic Species. NPAO bound to the proteins (HSA and BSA) and the unbound NPAO species were separated by size-exclusion chromatography (Phenomenex Biosep-SEC-S 2000 or Agilent ZORBAX GF-250) with 10 mM ammonium bicarbonate as the mobile phase at a flow rate of 0.7 mL/min. The effluent from HPLC was directly introduced to ICPMS for simultaneous detection of AsO^+ (m/z 91) and SO^+ (m/z 48). Detection of SO^+ (because of proteins) along with AsO^+ was used to support the identification of protein and arsenic in the same molecule.

Selection of Arsenic Binding Proteins by Affinity Column. Protein solution (2 mL) was poured into an affinity column containing 2 mL of affinity resin. The mixture was incubated at

(19) Bhattacharjee, H.; Rosen, B. P. *J. Biol. Chem.* **1996**, *271*, 24465–24470.
(20) Guo, Y. Z.; Ling, Y.; Thomson, B. A.; Siu, K. W. M. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1787–1794.
(21) Ngu, T. T.; Stillman, M. J. *J. Am. Chem. Soc.* **2006**, *128*, 12473–12483.
(22) Kalef, E.; Walfish, P. G.; Gitler, C. *Anal. Biochem.* **1993**, *212*, 325–334.
(23) Ehrlich, P.; Bertheim, A. *Bericht* **1910**, *43*, 917–927.

room temperature for 2 h with slow shaking. The column was washed sequentially with 10 mL of buffer A (20 mM Hepes, 50 mM NaCl, pH 7.5), 5 × 5 mL of buffer B (1% SDS in buffer A). The column was finally eluted twice with 2 mL of 10 mM DTT in Buffer A, followed by 2 mL of 100 mM DTT in buffer B. Each fraction of washing and elution solutions was separately collected (for 10 min) for the subsequent gel electrophoresis analysis, or in-solution digestion and LC-ESI-MS/MS analysis. Selected elution fractions were also quantified for the amount of protein by the Bio-Rad RC DC protein assay.

SDS-Polyacrylamide Gel Electrophoresis (PAGE) and Protein in-Gel Digestion. All fractions collected by affinity chromatography were freeze-dried. The freeze-dried samples were dissolved in deionized water (0.2−1 mL), and loaded onto SDS-PAGE gels. SDS-PAGE separation was carried out using 4%/10% stacking/separation polyacryamide mini gels and a Bio-Rad mini protein III system. The protein bands were silver stained for visualization.

To identify the proteins after SDS-PAGE, protein bands were excised from the gel and digested with trypsin. Briefly, gel bands were cut into about 1 × 1 mm pieces and washed 3 times in deionized water for 15 min each. After reduction and alkylation, gel pieces were dehydrated and resuspended in 0.02 μg/μL trypsin, 1 mM $CaCl_2$, and 50 mM NH_4HCO_3 buffer for overnight digestion at 37 °C. Peptides were extracted three times with 60% acetonitrile and 0.1% TFA solution, and then with 80% acetonitrile and 0.1% TFA solution. Each extraction was performed for 20 min with sonication, and the extracts were pooled for the subsequent LC-ESI-MS/MS analyses.

Protein Purification and in-Solution Digestion. The arsenic binding proteins in the elution fractions collected from the arsenic column were also precipitated with 20% TCA at 4 °C overnight, followed by centrifugation at 15,000 × *g* for 30 min at 4 °C. The precipitate was washed with ice cold acetone and dissolved in 100 μL of 0.5% SDS and 100 mM NH_4HCO_3 solution. After standard reduction and alkylation procedure, the proteins were precipitated again by 8 volumes of acetone (−20 °C) overnight and centrifuged for 30 min at 4 °C at 15,000 × *g*. Protein precipitate was washed by ice cold acetone and redissolved in 50 μL of 100 mM NH_4HCO_3 (pH 8.5) solution with vortexing. Finally, 3 μL of 1 μg/μL trypsin solution and 2.5 μL of 20 mM $CaCl_2$ were added for digestion overnight at 37 °C. After the overnight digestion, the reaction was stopped by adjusting the pH to 3 using formic acid.

SDS-Assisted Digestion and Peptide Fractionation. Proteins that could not be redissolved in NH_4HCO_3 solution after alkylation were dissolved in 7 μL of 1% SDS and 100 mM NH_4HCO_3 solution. Then the solution was diluted to 100 μL with 100 mM NH_4HCO_3 (pH 8.5), to which 3 μL of 1 μg/μL trypsin solution and 5 μL of 20 mM $CaCl_2$ were added for tryptic digestion. After overnight digestion at 37 °C, the reaction was stopped by adjusting the pH to 3. SDS in the digestion solution was removed and peptides in the solution were fractionated by strong cation exchange chromatography. A BioBasic SCX column from Thermo Scientific was used with a gradient elution consisting of 100% A (mobile phase A: 20% acetonitrile, 0.1% TFA) for 15 min, 0−30% B (mobile phase B: 20% acetonitrile, 0.1% TFA, 1 M NaCl) for 3 min, 30−50% B for 4 min, and finally 50% B and 50% A for 38 min. The flow rate was 0.2 mL/min. Fractions containing the peptides, as monitored at 214 nm, were collected for the subsequent LC-ESI-MS/MS analysis.

Mass Spectrometry Analysis of Proteins and Peptides. In-gel digestion samples were analyzed by nanospray-ESI-MS and LC-ESI-MS/MS. Peptides from in-gel digestion solution were desalted by Zip-tip (Millipore) following the instruction of the supplier. 50% methanol and 0.2% formic acid were used in nanospray-ESI experiments. Proteins were identified from both peptide fingerprints and MS/MS data.

LC-ESI-MS/MS was used to analyze in-solution digestion samples. A 1 μL volume of peptide solution was injected and separated by reversed-phase chromatography on a 0.3 × 150 mm Agilent C18 column at a flow-rate of 2 μL/min, and detected by QSTAR mass spectrometer. Gradient elution was performed with solvent A (0.1%, v/v, acetic acid in deionized water) and B (0.1%, v/v, acetic acid in acetonitrile). The gradient program consisted of 5% solvent B for 10 min for desalting, then increasing B to 25% over 30 min, and 25%−60% B over 100 min, and finally 60%−95% B over 60 min.

Database Searches. The Mascot search program (http://www.matrixscience.com/cgi/search_form) was used for all database searches. For peptide fingerprint, peptide mass tolerance was 100 ppm. For collision induced dissociation (CID) spectra search, the precision tolerance was set at 0.3 Da for both parent peptide and fragment ions. Trypsin was set as proteolytic enzyme, carbamidomethylation of cysteine was set as fixed modification, and methionine oxidation was set as variable modification, one missed cleavage site per peptide was allowed. An automatic database search followed by manual inspection was applied if more than one potential match was reported for one spectrum. Only the peptide hit with the highest score was examined. The rules used to inspect the peptide spectral identification were reported by Chen et al.[24] A protein was considered identified only when at least two peptides were found to identify the same protein. Possible cysteine positions in sequence and protein conformation were obtained from Swissprot.

RESULTS

Preparation of Arsenic Affinity Column and Characterization of Immobilization Efficiency. Our strategy of preparing an affinity column for capturing arsenic-binding proteins was to use trivalent arsenicals known to have high affinity for thiols as affinity ligands to bind with free thiols in target proteins. Previous work with arsenite, dimethylarsinous acid (DMA^{III}), monomethylarsonous acid (MMA^{III}), and phenylarsine oxide (PAO^{III}) has shown that of the four trivalent arsenicals tested; [8] PAO^{III} is relatively stable and has the highest affinity for hemoglobin. Conversely, DMA^{III} and MMA^{III} are not stable in solutions. Therefore, we chose an analogue of PAO^{III}, 4-aminophenylarsine oxide ($NPAO^{III}$), as the affinity ligand for capturing arsenic-binding proteins. The presence of a reactive primary amine group in $NPAO^{III}$ makes it readily immobilizable on a solid chromatographic support (resin) by the reaction of $NPAO^{III}$ with an epoxy group on the resin. Because the amine group

(24) Chen, Y.; Kwon, S. W.; Kim, S. C.; Zhao, Y. *J. Proteome Res.* **2005**, *4*, 998–1005.

and arsenic group are at the *para* positions, the arsenic group is kept exposed on the resin after the amine reacts with the epoxide. (Supporting Information, Figure S1).

We determined the amount of immobilized arsenic by treating the column with BAL and then treating the eluent with DTNB. The absorbance of the reaction product, thionitrobenzoate ion (λ = 412 nm, ε = 14,125), was measured to determine the amount of BAL that was not retained by the column (Supporting Information, Figure S2). We found that the average amount of $NPAO^{III}$ immobilized on the resin was 27.3 ± 2.3 μmol/g wet resin, measured from eight replicate preparations.

For comparison, we also immobilized inorganic arsenite (As^{III}) on the resin. We found that the amount of immobilized As^{III} was only about 4 nmol/g. The lower amount of immobilization of As^{III} is probably because the hydroxyl group in As^{III} is less reactive than the amine group in $NPAO^{III}$ with the epoxy group on the resin. In addition, the O−As bond formed between the epoxy group, and arsenic is easily hydrolyzed. Thus, $NPAO^{III}$ is more suitable than As^{III} for preparing an affinity column.

Specificity of the Arsenic Affinity Columns. To investigate the specificity of the arsenic affinity column, we tested the capture and elution behavior of several proteins that either contain free cysteine(s) or have no free cysteine. Initially we compared HSA, which has a free cysteine, and HSA after its cysteine was blocked by carbamidomethylation. The former is expected to be retained by the arsenic affinity column, while the latter is not expected to be captured.

Figure 1 shows results from gel electrophoresis analyses of HSA (lane 1), washing solutions (with buffer containing 1% SDS) from the column after HSA was passed through the column (lanes 2−7), and the elution fraction containing the captured HSA (lanes 8−10). As shown in Figure 1a, the non-specific binding can be washed off with a buffer containing 1% SDS (lanes 2−7). The fractions eluted with 10 mM DTT (lanes 8 and 9) and subsequently with 100 mM DTT (lane 10) show the detection of the captured HSA. These results indicate that the arsenic affinity column is able to capture the unmodified HSA (containing a free cysteine) and that the captured HSA can be subsequently eluted for analysis.

In contrast, once the thiol in the cysteine is blocked by carbamidomethylation, the modified HSA cannot be captured by the arsenic affinity column (Figure 1b). There is no HSA detected in the DTT elution fractions (lanes 8−10). These results suggest that a free cysteine in the protein is necessary for the capture of the protein by the arsenic affinity column.

Further support was obtained from LC-ICPMS analyses of two reaction mixtures of $NPAO^{III}$ with either unmodified HSA or cysteine-blocked HSA (Figure 2). We observed that the HSA containing a free cysteine was able to bind to $NPAO^{III}$ to form a NPAO-HSA complex, which was separated from the unbound $NPAO^{III}$ using size exclusion chromatography. There was no complex between $NPAO^{III}$, and the HSA that was pretreated to block its reactive cysteine. These results confirm that the specific capture of HSA on the $NPAO^{III}$ column is due to the binding of $NPAO^{III}$ to the free thiol in HSA.

Similarly, we conducted parallel experiments using BSA (containing one free cysteine) to examine the capture of BSA on the arsenic affinity column and the binding between BSA and $NPAO^{III}$. Both LC-ICPMS analysis and gel electrophoresis

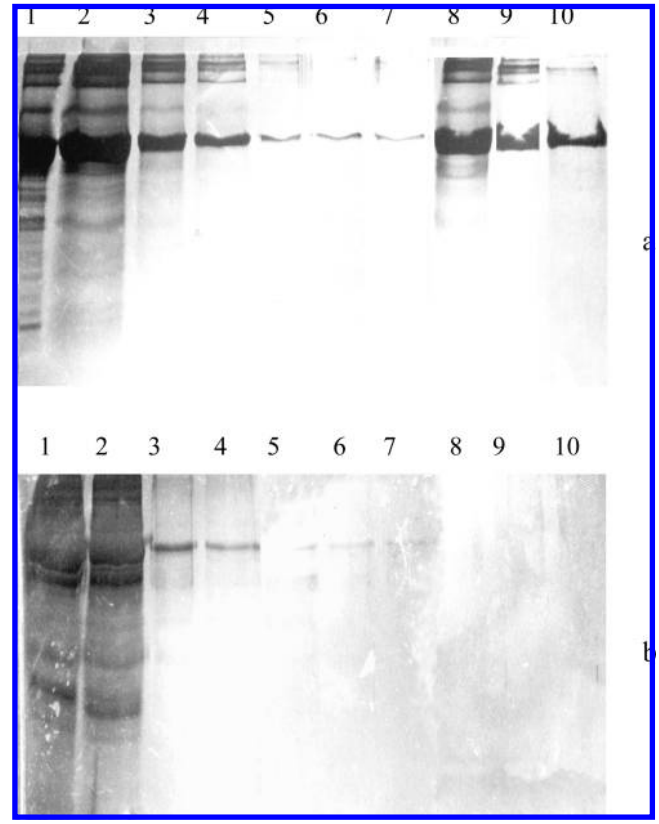

Figure 1. Gel electrophoresis image from the analysis of HSA (lane 1), the washing solution (lanes 2−7), and eluting solution (lanes 8−10) collected from the affinity column after applying an HSA sample to the column. (a) HSA containing a free cysteine; (b) HSA treated with carbamidomethylation to block the thiol group. Lane 1: HSA sample without passing through the affinity column; lanes 2−7: 6 rounds of washing with buffer A (20 mM Hepes, 50 mM NaCl, 1% SDS, pH 7.5); lanes 8 and 9: elution with 10 mM DTT in buffer A; lane 10: elution with 100 mM DTT in buffer A.

results corroborate that BSA can bind to $NPAO^{III}$ and can be captured by $NPAO^{III}$ affinity column.

To confirm the identity of the captured proteins in the above experiments, we carried out in-gel digestion (lanes 8−10 in Figure 1), and analyzed the products by nanoelectrospray ionization mass spectrometry (nano-ESI-MS). Peptide fingerprint maps from these analyses show 73% peptide coverage for BSA and 63% for HSA, indicating that the protein bands from DTT elution solutions are indeed BSA and HSA, respectively.

To further confirm that the captured proteins are due to arsenic specific binding, and not due to any non-specific binding (e.g., caused by hydrophobic interaction between proteins and benzyl group on the affinity resin), we tested the retention of proteins on a control column that does not have arsenic affinity. Benzylamine was used to react with the epoxy group of the reactive resin Eupergit C, forming a similar structure as the affinity resin but without the reactive arsenic group. The membrane/organelle fraction of A549 cells was chosen to be tested as a worst case scenario because the membrane/organelle fraction of proteins is more hydrophobic and more likely to interact with the resin containing a benzyl group. Supporting Information, Figure S3 shows the overdeveloped electrophoresis gels from the analysis of the washing and elution solutions after applying the membrane/organelle sample to the non-specific control column. Although

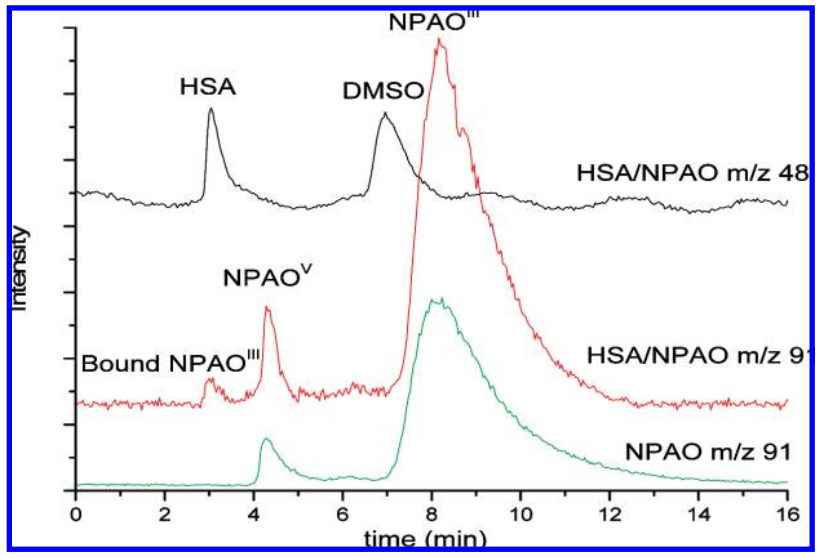

Figure 2. Chromatograms from HPLC-ICP-MS analyses of unbound NPAO^{III} and HSA-bound NPAO^{III}. Green trace shows AsO^+ signal from the analysis of NPAO; red trace and black trace are from the analysis of a mixture containing NPAO^{III} (3 μM) and HSA (8 μM). The black trace is from the monitoring of SO^+ (*m*/*z* 48), indicating the elution of HSA at 3−4 min. The red trace is from the monitoring of AsO^+ (*m*/*z* 91), indicating the elution of NPAO bound to HSA (3−4 min).

proteins are present in the initial washing solutions as expected, no protein band is detected in the DTT elution solution after 6 successive washes with buffer containing 1% SDS. These results indicate that the non-specific binding caused by hydrophobic interaction can be removed by extensive washing with buffers containing 1% SDS. Therefore, it can be concluded that the capture of particular proteins on the arsenic affinity column is due to specific binding of the proteins to arsenic, not due to the non-specific retention.

Separation of Arsenic-Binding Proteins from Large Excess of Non-Specific Proteins. Having established the specific retention of arsenic-binding proteins on the NPAO column, we further tested the possibility of capturing a small amount of arsenic-binding proteins in the presence of a large excess of proteins that do not bind to arsenic. To carry out this experiment, we chose a commercial preparation of bovine carbonic anhydrase II, which does not contain cysteine and is not expected to be retained on the arsenic affinity column. However, the minor proteins (present as the impurity) could be captured by the arsenic affinity column. Figure 3 shows the electrophoresis results of carbonic anhydrase II applied to the affinity column. Non-specifically binding proteins were removed by buffer containing 1% SDS after 5 successive washes (lanes 2−6). However, the elution with 10 mM DTT (lane 7) shows a strong band at a slightly lower molecular weight than that of the carbonic anhydrase II. This band was cut out for tryptic digestion and LC-ESI-MS/MS analysis. Tandem mass spectra of peptides from the digested protein (Supporting Information, Figure S4), including *m*/*z* 606 for LQDVTDDHIR and *m*/*z* 586 for LPSEGPQPAHVVVGDVR, provided the identification of the protein as bovine biliverdin reductase B (22 kDa). Bovine biliverdin reductase B and carbonic anhydrase II are both located in the cytoplasm. It is reasonable that the carbonic anhydrase II preparation contains a small amount of bovine biliverdin reductase B. Examination of the three-dimensional structure of bovine biliverdin reductase B from Swissprot Expasy show the presence of Cys188 at the C-terminal, which is exposed. Thus, it can be concluded that the arsenic affinity column can capture a small amount of arsenic-binding protein present in a sample containing excess amounts of other proteins.

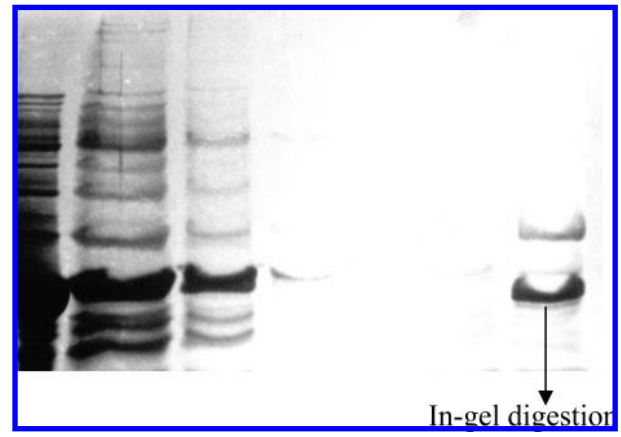

Figure 3. Gel electrophoresis images from the analyses of washing and eluting solutions from the affinity column after applying a carbonic anhydrase II sample. Lane 1: carbonic anhydrase II sample without passing through the affinity column; lanes 2 to 6: five rounds of washing with buffer B (1% SDS in 20 mM Hepes, 50 mM NaCl, pH 7.5); lane 7: elution by 10 mM DTT in buffer B.

Application of the Arsenic Affinity Column to the Capture and Identification of Cellular Proteins. The benefit of the arsenic affinity column is demonstrated in an application to the identification of arsenic binding proteins in subcellular fractions of human cells (A549 lung carcinoma cells). Gel electrophoresis of the nuclear fraction and membrane/organelle fraction of A549 cells showed that after applying the cell samples to the column and 6 successive washes with HEPES buffer containing 1% SDS, non-specific components were washed off. Subsequent DTT elution resulted in the release of arsenic-binding proteins from the column. In-solution digestion and LC-ESI-MS/MS were used to identify the arsenic-binding proteins eluted from the column. Using the combination of arsenic-affinity capturing and LC-ESI-MS/MS, we were able to identify 50 proteins in the nuclear fraction of A549 cells (Table 1), and 24 proteins in the membrane/

Table 1. Arsenic-Binding Proteins Identified from the Nuclear Fraction of A549 Cells[a,b,c]

protein	accession no.	MWt	peptide	possible cysteine position
① Actin, cytoplasmic 1	P60709	42 kDa	GYSFTTTAER QEYDESGPSIVHR	17, 217, 257, 272*, 285*, 374
① Actin, aortic smooth muscle	P62736	42 kDa	EITALAPSTMK + Oxidation (M) DSYVGDEAQSK	2, 12, 19, 219, 259, 287*, 376
① ADP/ATP translocase 1	P12235	33 kDa	TAVAPIER YFPTQALNFAFK	57, 129*, 257*
① ADP/ATP translocase 3	P12236	33 kDa	EQGVLSFWR DFLAGGIAAAISK	57, 129*, 160*, 257*
① Annexin A2	P07355	38 kDa	AYTNFDAER TPAQYDASELK	9, 335*
③ Anterior gradient protein 2 homologue precursor	O95994	20 kDa	ADITGR IMFVDPSLTVR	81
③ ATP-dependent RNA helicase A	Q08211	141 kDa	DINTDFLLVVLR ETPFELIEALLK	12, 36, 242, 415, 438, 469, 489, 552, 558, 578, 608, 612, 732, 773, 777, 872, 881, 888, 940, 973, 1004, 1029, 1099
① ATP-dependent DNA helicase 2 subunit 1	P12956	70 kDa	ILELDQFK LGSLVDEFK	66*, 150, 389, 398, 585*
② DNA topoisomerase 2-alpha	P11388	174 kDa	YGVFPLR YDTVLDILR	104*, 170, 216*, 300, 392, 405*, 427, 455, 733, 862, 997, 1008*, 1145*
③ DNA-dependent protein kinase catalytic subunit	P78527	469 kDa	AQEPESGLSEETQVK ILELSGSSSEDSEK	78 cysteines
③ Dynein heavy chain, cytoplasmic	Q14204	532 kDa	GTFDNAETK ETVDQVEELR	36 cysteines
② 4F2 cell-surface antigen heavy chain	P08195	58 kDa	GQSEDPGSLLSLFR VAEDEAEAAAAAK	330*
② Filamin-A	P21333	280 kDa	ENGVYLIDVK AGGPGLER	41 cysteines
① Fructose-bisphosphate aldolase A	P04075	39 kDa	GILAADESTGSIAK GILAADESTGSIAK	73*, 135*, 150, 178*, 202, 240*, 290*, 339
② Heat shock cognate 71 kDa protein	P11142	71 kDa	DAGTIAGLNVLR VEIIANDQGNR	17, 267*, 603
③ Heterogeneous nuclear ribonucleoprotein A/B	Q99729	37 kDa	GFGFILFK GFVFITFK	98, 224
② Heterogeneous nuclear ribonucleoprotein A3	P51991	40 kDa	EDTEEYNLR DYFEKYGK	64*, 85*, 94*, 196
② Heterogeneous nuclear ribonucleoproteins A2/B1	P22626	37 kDa	QEMQEVQSSR GGNFGFGDSR	50*
② Heterogeneous nuclear ribonucleoproteins C1/C2	P07910	34 kDa	VPPPPPIAR SDVEAIFSK	46*
② Heterogeneous nuclear ribonucleoprotein C-like 1	O60812	32 kDa	VDSLLENLEK VFIGNLNTLVVK	46
② Heterogeneous nuclear ribonucleoprotein D0	Q14103	39 kDa	GFGFVLFK IFVGGLSPDTPEEK	126, 226*, 252*
③ Heterogeneous nuclear ribonucleoprotein K	P61978	51 kDa	GSDFDCELR IDEPLEGSEDR	132, 145, 184, 185, 205
③ Heterogeneous nuclear ribonucleoprotein L	P14866	60 kDa	R.YYGGGSEGGR.A SDALETLGFLNHYQMK	151, 215, 218, 260, 261, 396, 404, 452, 472, 521, 581
③ Heterogeneous nuclear ribonucleoprotein R	O43390	71 kDa	ENILEEFSK DLYEDELVPLFEK	99, 214, 226, 240, 292
③ Heterogeneous nuclear ribonucleoprotein Q	O60506	69 kDa	AGPIWDLR GFCFLEYEDHK	96, 211, 237, 289
③ Interleukin enhancer -binding factor 2	Q12905	43 kDa	ILPTLEAVAALGNK	37, 271, 291, 311

Table 1. Continued

protein	accession no.	MWt	peptide	possible cysteine position
			VKPAPDETSFSEALLK	
③ Keratin, type I cytoskeletal 10	P13645	60 kDa	LASYLDK	25, 66, 401, 427
			SQYEQLAEQNR	
③ Keratin, type I cytoskeletal 9	P35527	62 kDa	SGGGGGGGLGSGGSIR	3, 61, 406, 432
			FSSSSGYGGGSSR	
③ Keratin, type II cytoskeletal 2 epidermal	P35908	66 kDa	NLDLDSIIAEVK	3, 7, 42, 423, 495
			STSSFSCLSR	
③ Keratin, type II cytoskeletal 1	P04264	66 kDa	YEDEINK	49, 147, 497
			AEAESLYQSK	
② Lamin-A/C	P02545	74 kDa	EGDLIAAQAR	552*, 570, 588, 591, 661
			EDLQELNDR	
③ Lamin-B1	P20700	66 kDa	EYEAALNSK	110, 198, 317, 443, 583
			ESDLNGAQIK	
Lamina-associated polypeptide 2	P42167	51 kDa	PEFLEDPSVLTK	362
			HASPILPITEFSDIPR	
③ Leucine-rich repeat-containing protein 59	Q96AG4	35 kDa	VAGDCLDEK	48, 59, 131, 140, 266, 277
			ATILDLSCNK	
② Matrin-3	P43243	95 kDa	SFQQSSLSR	230, 293, 296, 552*, 563*
			TEEGPTLSYGR	
② Non-POU domain-containing octamer-binding protein	Q15233	54 kDa	NLPQYVSNELLEEAFSVFGQVER	145*, 208
			AVVIVDDR	
③ Nucleolar RNA helicase 2	Q9NR30	89 kDa	GVTFLFPIQAK	161, 291, 378, 418, 445, 537, 539, 643, 682
			EQLGEEIDSK	
② Nucleophosmin	P06748	32 kDa	VDNDENEHQLSLR	21, 104*, 275
			GPSSVEDIK	
② Peroxiredoxin-1	Q06830	22 kDa	TIAQDYGVLK	71, 83*
			QITVNDLPVGR	
③ Phosphate carrier protein	Q00325	41 kDa	IQTQPGYANTLR	57, 68, 76, 91, 136, 237, 257, 277
			GIFNGFSVTLK	
② Polypyrimidine tract-binding protein 1	P26599	57 kDa	KLPIDVTEGEVISLGLPFGK	23, 250*, 251
			LPIDVTEGEVISLGLPFGK	
②Poly(rC)-binding protein 1	Q15365	37 kDa	INISEGNCPER	54*, 109, 118, 158, 194, 201, 293*, 355
			IANPVEGSSGR	
③ Probable ATP-dependent RNA helicase DDX5	P17844	69 kDa	TTYLVLDEADR	89, 170, 191, 194, 200, 221, 234, 320, 354
			GDGPICLVLAPTR	
③ Probable ATP-dependent RNA helicase DDX17	Q92841	72 kDa	GDGPICLVLAPTR	87, 120, 168, 191, 198, 219, 240, 318, 352, 368, 505
			LIQLMEEIMAEK +2 Oxidation (M)	
① 40S ribosomal protein S14	P62263	16 kDa	IEDVTPIPSDSTR	31, 54*, 85
			TPGPGAQSALR	
① Small nuclear ribonucleoprotein Sm D2	P62316	14 kDa	SEMTPEELQK	46*, 63*
			NNTQVLINCR	
③ Spectrin alpha chain	Q13813	284 kDa	VNSLGETAER	158, 315, 466, 956, 1091, 1314, 1444, 1454, 1622, 1930, 2120, 2233, 2351, 2441
			DLTGVQNLR	
② Spectrin beta chain, brain 1	Q01082	275 kDa	VLDNAIETEK	73, 112, 183, 604, 619, 624, 861, 964, 1284, 1389, 1900, 1970, 2227, 2262*
			LVSDGNINSDR	
② Splicing factor, arginine/ serine-rich 3	P84103	20 kDa	AFGYYGPLR	6*, 10, 72*, 74
			NPPGFAFVEFEDPR	
② U2 small nuclear ribonucleoprotein A′	P09661	29 kDa	TFNPGAGLPTDK	77*, 89
			SLTYLSILR	
② Vimentin	P08670	53 kDa	QVDQLTNDK	328*
			QDVDNASLAR	

[a] Cysteines in Filamin-A are 205, 210, 444, 478, 483, 574, 623, 649, 717, 733, 796, 810, 1018, 1108, 1122, 1157, 1165, 1185, 1198, 1225, 1260, 1353, 1402, 1410, 1453, 1645, 1686, 1689, 1723, 1865*, 1912*, 1920, 1997, 2102, 2107, 2160, 2199, 2293, 2476, 2479, 2601. [b] Cysteines in PRKDC are 10, 25, 42, 81, 90, 111, 123, 223, 232, 285, 301, 373, 392, 457, 458, 478, 491, 630, 729, 795, 931, 974, 1029, 1032, 1127, 1128, 1135, 1164, 1176, 1183, 1229, 1255, 1266, 1312, 1335, 1364, 1377, 1399, 1432, 1455, 1499, 1507, 1525, 1629, 1742, 1767, 1791, 1831, 1904, 1919, 1947, 1953, 1954, 2244, 2248, 2292, 2363, 2397, 2403, 2435, 2469, 2857, 2880, 3001, 3014, 3187, 3234, 3281, 3286, 3293, 3403, 3420, 3683, 3781, 3912, 4045, 4061, 4106. [c] Cysteines in Dynein heavy chain, cytoplasmic are 220, 416, 867, 1059, 1484, 1888, 1932, 1956, 1977, 1999, 2076, 2186, 2359, 2454, 2466, 2594, 2639, 2663, 2712, 3147, 3325, 3389, 3507, 3693, 3712, 3808, 3940, 4044, 4121, 4170, 4216, 4510, 4540, 4556, 4570, 4644.

Table 2. Arsenic-Binding Proteins Identified from the Membrane Fraction of A549 Cells

protein	accession no.	MWt	peptide	possible cysteine position
① Actin	P60709	42 kDa	EITALAPSTMK+Oxidation (M)	17, 217, 257, 272*, 285*, 374
	P62736		DLTDYLMK+Oxidation (M)	2, 12, 19, 219, 259, 287*, 376
① Annexin A2	P07355	38 kDa	SYSPYDMLESIR	9, 335*
			TPAQYDASELK	
① Elongation factor1-alpha1	P68104	50 kDa	IGGIGTVPVGR	31*, 111, 234*, 363, 370*, 411
			QLIVGVNK	
① Elongation factor 1-alpha 2	Q05639	50 kDa		31*, 111, 326, 363, 370*, 411
② 4F2 cell-surface antigen heavy chain	P08195	60 kDa	EEGSPLELER	330*
			VAEDEAEAAAAAK	
② 78 kDa glucose-regulated protein precursor	P11021	72 kDa	ELEEIVQPIISK	41, 420*
			VEIIANDQGNR	
① Glucose-6-phosphate 1-dehydrogenase	P11413	59 kDa	GYLDDPTVPR	13*, 158, 232, 269, 294*, 358, 385, 446*
			DNIACVILTFK	
③ 60 kDa heat shock protein	P10809	61 kDa	VGLQVVAVK	237, 442, 447
			NAGVEGSLIVEK	
① 3-hydroxyacyl-CoA dehydrogenase type-2	Q99714	27 kDa	DLAPIGIR	5, 58, 91*
			DVQTALALAK	
③ Keratin, type I cytoskeletal 9	P35527	62 kDa	TLLDIDNTR	3, 61, 406, 432
			QEYEQLIAK	
③ Keratin, type I cytoskeletal 10	P13645	60 kDa	LAADDFR	25, 66, 401, 427
			LASYLDK	
③ Keratin, type II cytoskeletal 1	P04264	66 kDa	IEISELNR	48, 146, 496
			TLLEGEESR	
③ Keratin, type II cytoskeletal 2 epidermal	P35908	66 kDa	FASFIDK	3, 7, 42, 423, 495
			TAAENDFVTLK	
② Peroxiredoxin-1	Q06830	22 kDa	ADEGISFR	71, 83*
			TIAQDYGVLK	
③ Protein disulfide-isomerase A6 precursor	Q15084	48 kDa	GSTAPVGGGAFPTIVER	11, 291, 297
			GESPVDYDGGR	
① Pyruvate kinase isozymes	P14618	58 kDa	GDYPLEAVR	31*, 49, 152, 165, 317, 326, 358, 423*, 424*, 474*
			GDLGIEIPAEK	
③ Ras-related protein Rab-10	P61026	23 kDa	LLLIGDSGVGK	24, 124, 199, 200
③ Ras-related protein Rab-15	P59190	24 kDa	AFLTLAEDILR	23, 26, 153, 210, 212
③ Ras-related protein Rab-35	Q15286	23 kDa	LQIWDTAGQER	110, 114, 163, 200, 201
② Ras-related protein Rab-1A	P62820	23 kDa		26, 126*, 204, 205
② Ras-related protein Rab-1B	Q9H0U4	22 kDa		23*, 200, 201
② Ras-related protein Rab-7	P51149	24 kDa	DEFLIQASPR	83, 84, 143*, 205, 207
			EAINVEQAFQTIAR	
② Reticulon-4	Q9NQC3	130 kDa	GPLPAAPPVAPER	424, 464, 559, 597, 699, 1101*
			GPLPAAPPVAPER	
③ Stomatin-like protein 2	Q9UJZ1	39 kDa	AEQINQAAGEASAVLAK	167, 172
			DVQGTDASLDEELDR	
③ Stress-70 protein	p38646	74 kDa	VQQTVQDLFGR	66, 317, 366, 487, 608
			AQFEGIVTDLIR	
③ Thioredoxin domain-containing protein 1 precursor	Q9H3N1	32 kDa	VDVTEQPGLSGR	106, 165, 198, 205, 207
			DFINFISDK	
① Tubulin alpha-ubiquitous chain	P68363	51 kDa	QLFHPEQLITGK	4, 20, 25*, 129*, 200, 213*, 295, 305, 315, 316, 347*, 376
			IHFPLATYAPVISAEK	
③ Voltage-dependent anion-selective channel protein 2	P45880	32 kDa	LTLSALVDGK	8, 13, 47, 76, 103, 133, 138, 210, 227
			YQLDPTASISAK	
③ Voltage-dependent anion-selective channel protein 3	Q9Y277	31 kDa	LTLDTIFVPNTGK	2, 8, 36, 65, 122, 229
			LTLSALIDGK	

organelle fraction (Table 2). Importantly, these are arsenic-binding proteins.

From the available protein/peptide sequence and structure information, we have confirmed the presence of free cysteines in the proteins that we have captured and identified; and these cysteines are indicated in Tables 1 and 2. The capture of these proteins by the arsenic affinity column further supports our conclusion of specific binding between arsenic and the proteins that resulted in their capture.

DISCUSSION

Our results show that the arsenic affinity technique can be used to select arsenic-binding proteins, and that the combination of the arsenic affinity chromatography and mass spectrometry is

a practical method to identify arsenic-binding proteins in the proteome. Several reports [12,14,18,22] have dealt with the selection and identification of arsenic-binding protein by using arsenic affinity media. However, only a few proteins (galectin-1, thioredoxin peroxidase, GLUT4, tubulin, and Actin) have been identified as arsenic binding proteins by using Western blotting, molecular weight determination, partial amino acid sequence analysis, and co-immunoprecipitation. Most arsenic binding proteins remain unknown. To identify many potential arsenic binding proteins, this study demonstrates the usefulness and capability of the improved arsenic affinity column (immobilizing 14−27 μmol of NPAO per gram of resin) and mass spectrometry. Using this powerful combination, we were able to capture and identify 50 proteins in the nuclear fraction and 24 proteins in the organelle/membrane fraction of A549 cells as arsenic binding proteins. The same strategy could be extended to study arsenic binding proteins in other cells and other metal-binding proteins.

We tested the column lifetime by comparing the capture of arsenic using freshly prepared columns with those stored at 4 °C for 30 and 45 days and by measuring the amount of $NPAO^{III}$ on the columns after the storage. We found that the amount of $NPAO^{III}$ bound to the resin was decreased to 50% after an affinity column was stored at 4 °C for 45 days, probably because of the oxidation of $NPAO^{III}$ to $NPAO^{V}$. However, this decrease did not affect the capture of arsenic. This result is understandable considering the large amount of $NPAO^{III}$ immobilized on the resin (27.3 ± 2.3 μmol/g wet weight resin). Thus, the arsenic affinity column can be stored at 4 °C and used a month later.

Much research has focused on arsenic binding to peptides or proteins containing vicinal dithiol.[25] The K_d values for arsenic binding to peptides containing a monothiol are at least 6 times lower than those binding to peptides of dithiols. Nonetheless, the present study and previous research[8] show that hemoglobin, HSA, and BSA can bind to trivalent arsenic, although these proteins do not have vicinal sulfhydryl structure. Thus, we examined the three-dimensional conformation of hemoglobin and HSA, and explored the possible binding cysteines in the protein sequence. The X-ray crystal structure of HSA[26] shows that Cys34 does not participate in any disulfide bridge, its Sγ atom lies toward the interior and surrounded by Pro53, His39, Val77, and Tyr84 of the side chain. In solution, the side chain of Tyr84 may flip over and enable other molecules, such as arsenite, to bind to the thiol and form an arsenic complex. Another possibility is that the HSA backbone conformation may change to bring the sulfhydryl group toward the exterior of protein. The three-dimensional structure of proteins, including the position of cysteines in the protein (embedded inside or close to the surface), other amino acids in the side chain surrounding the cysteine (steric hindrance, hydrophobic or hydrophilic environment, and the electron density), and the orientation of the thiol group (toward the exterior or interior), modulate arsenic binding.

In our study, the affinity resin Eupergit C has a three-atom 6 Å hydrophilic spacer connected to NPAO, and presumably only the cysteines on the protein surface or close to the protein surface can bind to the affinity column. In this case, the affinity column will select the proteins which have cysteine on the protein surface or close to the protein surface. Those proteins with cysteines embedded inside are unlikely to be selected by the affinity column. We examined the three-dimensional conformation of the identified proteins from Swissprot and literature. The column indicating cysteine positions in Tables 1 and 2 shows the free cysteines in the identified proteins. Other cysteine residues forming disulfide bonds and those involved in Zn finger binding are not included in the tables. The cysteine positions labeled with * are the cysteines on the protein surface. The proteins labeled with ① have three-dimensional structure information from the whole protein sequence in Swissprot. The proteins labeled with ② are those whose three-dimensional conformation was examined by their peptide fragments. The proteins labeled with ③ have no conformational data available. Our results show that all the proteins captured by the arsenic affinity column have at least one cysteine available for binding with trivalent arsenicals. The protein three-dimensional confirmation examination shows that protein binding to arsenic can be mediated by single cysteines in α helices, β sheets or loops.

A number of arsenic-binding proteins identified using the affinity chromatography tandem mass spectrometry approach (Tables 1 and 2) are of particular interest because of their important biological functions. For example, DNA-dependent protein kinase, ATP-dependent helicase II (Ku70), and topoisomerase 2 alpha, are involved in DNA repair and maintaining genome stability.[27,28] Several other proteins modulate the redox status of cells, for example, peroxiredoxin-1 and thioredoxin,[29] and apoptosis, for example, lamin A and heat shock cognate protein.[30,31] This work shows that arsenic can bind to these proteins in cell extracts. How arsenic affects the function of these proteins in biological systems will have to be confirmed by studying arsenic interaction with proteins in living cells.

ACKNOWLEDGMENT

The authors thank Ms. Jane Lee and Dr. Shengwen Shen for assistance on cell culture and cell extracts. This work was supported by Canadian Institutes of Health Research, the National Cancer Institute of Canada with funds from Terry Fox Foundation (to XCL and M.W.), Metals in the Human Environment Strategic Network, Canadian Water Network, Alberta Water Research Institute (Alberta Ingenuity), and Alberta Health and Wellness (to X.C.L.).

SUPPORTING INFORMATION AVAILABLE

Further details are given in Figures S1−S4. This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review February 15, 2009. Accepted March 27, 2009.

AC900352K

(25) Sowerby, D. B. *The chemistry of organic arsenic, antimony and bismuth compounds*; Patai, S., Ed.; John Wiley and Sons Inc.: New York, 1994.
(26) Sugio, S.; Kashima, A.; Mochizuki, S.; Noda, M.; Kobayashi, K. *Protein Eng.* **1999**, *12*, 439–446.
(27) Dip, R.; Naegeli, H. *FASEB J.* **2005**, *19*, 704–715.
(28) Degrassi, F.; Fiore, M.; Palitti, F. *Curr. Med. Chem. Anticancer Agents* **2004**, *4*, 317–325.
(29) Bindoli, A.; Fukuto, J. M.; Forman, H. J. *Antioxid. Redox. Signaling* **2008**, *10*, 1549–1564.
(30) Han, X.; Feng, X.; Rattner, J. B.; Smith, H.; Bose, P.; Suzuki, K.; Soliman, M. A.; Scott, M. S.; Burke, B. E.; Riabowol, K. *Nat. Cell Biol.* **2008**, *10*, 1333–1340.
(31) Rubio, E.; Valenciano, A. I.; Segundo, C.; Sanchez, N.; de Pablo, F.; de la Rosa, E. J. *Eur. J. Neurosci.* **2002**, *15*, 1646–1654.