University of Alberta

Affinity Capture and Mass Spectrometry Analysis of Proteins Binding to Arsenic

by

Huiming Yan ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of

Master of Science

Department of Chemistry

Edmonton, Alberta

Fall 2008

Your file Votre référence
ISBN: 978-0-494-47445-7
Our file Notre référence
ISBN: 978-0-494-47445-7

ABSTRACT

Exposure to high levels of arsenic causes a wide range of health effects. Arsenic binding to the sulfhydryl group of cysteine in proteins, affecting protein's function, is one of the possible mechanisms. This thesis focuses on understanding arsenic binding to proteins and developing affinity capture techniques for the analysis of arsenic binding proteins.

An arsenic affinity technique was developed to capture the specific arsenic binding proteins. The captured proteins were subsequently identified by mass spectrometry. Over seventy proteins in A549 cells were captured and identified. Using the available structural information of the proteins, the possible cysteines binding to arsenic were proposed. This technique was further applied to identify proteins in the nuclear extract of A549 cells treated with either benzo[α]pyrene diol expoxide alone or in combination with arsenite.

The binding of arsenicals with p53 and PARP-1 was studied using HPLC-ICP-MS, affinity capture and Western blot analysis. Weak binding of p53 and PARP-1 to phenylarsine oxide was observed.

ACKNOWLEDGMENTS

I would like to thank my research supervisor, Dr. X. Chris Le. I also appreciate Dr. Liang Li, Dr. Robert E. Campbell, and Dr. Michael Weinfeld for the helpful discussion and suggestions. This work also received support from Dr. Liang Li's group, Dr. David Bundle's group, Dr. Tod Lowary's group, and Dr. Michael Weinfeld's group. Financial support from the Department of Chemistry at the University of Alberta, the Natural Sciences and Engineering Research Council of Canada (NSERC), and the National Cancer Institute of Canada is also gratefully acknowledged.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

List of Schemes

Abbreviations

ACN	acetonitrile
As^{III}	arsenite
As^{V}	arsenate
BAL	British anti-Lewisite (2,3-dimercapto-1-propanol)
BPDE	benzo[a]pyrene diol epoxide
BSA	bovine serum albumin
DMA^{III}	dimethylarsinous acid
DMA^{V}	dimethylarsinic acid
DMSO	dimethylsulfate oxide
DTNB	5,5'-dithiobis(2-nitrobenzoic acid)
DTT	dithiothreitol
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
HSA	human serum albumin
IAA	iodoacetamide
K_d	dissociation constant
MMA^{III}	monomethylarsonous acid
MMA^{V}	monomethylarsonic acid
$NPAO^{III}$	4-aminophenylarsine oxide
$NPAO^{V}$	4-aminophenylarsonic acid
PAO^{III}	phenylarsine oxide
PAO^{V}	phenylarsonic acid
PARP-1	poly (ADP-ribose) polymerase
PMSF	phenylmethanesulfonyl fluoride
SDS	sodium dodecyl sulfate
TCA	trichloroacetic acid

TCEP	tris(2-Carboxyethyl) phosphine
TFA	trifluoroacetic acid
Tris	tris(hydroxymethyl)aminomethane

Chapter 1

Introduction

1. Arsenic species in nature

Arsenic ranks twentieth in abundance among the elements in the earth's crust.[1] It does not often exist in its elemental state and is far more common in sulfides and sulfur-containing salts in mineral ores, for example, arsenopyrite (FeAsS), realgar (AsS), enaegite ($CuAsS_4$), orpiment (As_2S_3), and tennantite ($Cu_{11}FeAs_4S_{13}$). Usually, arsenic-containing ores are associated with other mineral ores, such as gold, silver, dyscrasite, barite, cinnabar, and nickeline mineral ore. Coal has an average arsenic concentration of about 10 mg/kg, but some coals contain more than 1000 mg/kg. [1] Although arsenic ore is not intentionally mined for arsenic compound, arsenic is usually released as a byproduct in processing the ores of other elements. This "accidental" source results in the release of most of the arsenic in the environment.

Arsenic is widely distributed in food, water, air, and soil. It exists in inorganic and organic forms and in different oxidation states (-3, 0, +3, +5). Inorganic arsenic is the common form in the environment, and is found in groundwater, surface water, and many foods, such as rice and grains. Organic arsenicals are often biotransformation products (metabolites) of inorganic arsenic. Table 1.1 shows arsenic species that are common in the environment and are relevant to this thesis.

1.1 Arsenic in water

The concentration of arsenic in surface water, such as rivers, lakes, and oceans is not constant. Andreae[2] reported a mean of 1.4 μg/L, with a range of 0.1-75 μg/L in some European and American rivers. But in Chile, arsenic concentration in surface water is up to 1000 μg/L due to the pollution of mining[3]. Arsenic concentration in the groundwater is also variable. High concentration of arsenic was found in the groundwater of Bangladesh,

northeastern India, Vietnam, China, and the western U.S.[4]. In the blackfoot-disease (BFD) area of Taiwan, the average arsenic in wells was 671±149 μg/L; but outside the BFD area, it dropped to 0.7 μg/L[5]. Table 1.2[3] shows the arsenic concentration in the groundwater and surface water in some contamination areas. Arsenite and arsenate are the predominant forms of arsenic found in water although methylarsonic and dimethylarsinic acids have been found at low levels[6]. The distribution between two main arsenic species in water, arsenite and arsenate, highly depends on the redox potential and pH[7, 8]. As^{III} is the predominant form in the typical anoxic groundwater condition at pH less than 7, but As^{V} will dominate in the oxic groundwater.

Table 1.1 List of common arsenic species

Full name	Abbreviation	Formula
Arsenous acid (Arsenite)	As^{III}	$As(OH)_3$
Arsenic acid (Arsenate)	As^{V}	$AsO(OH)_3$
Monomethylarsonous acid	MMA^{III}	$CH_3As(OH)_2$
Dimethylarsinous acid	DMA^{III}	$(CH_3)_2AsOH$
Monomethylarsonic acid	MMA^{V}	$CH_3AsO(OH)_2$
Dimethylarsinic acid	DMA^{V}	$(CH_3)_2AsO(OH)$
Phenylarsine oxide	PAO^{III}	C_6H_5AsO
Phenylarsonic acid	PAO^{V}	$C_6H_5AsO(OH)_2$
4-aminophenylarsine oxide	$NPAO^{III}$	$4\text{-}NH_2C_6H_4AsO$
4-aminophenylarsonic acid	$NPAO^{V}$	$4\text{-}NH_2C_6H_4AsO(OH)_2$
Arsenobetaine	AsB	$(CH_3)_3As^{+}CH_2COO^{-}$

Table 1.2. Worldwide occurrences of arsenic contamination in water[3]

Location	Potential exposed population	Concentration (μg/L)	Environmental conditions	Source
Argentina	2,000,000	1-2,900	Natural; volcanic rocks and thermal springs	Groundwater
Bangladesh	>29,000,000	1-4,730	Natural, alluvia	Groundwater
Bolivia	50,000	-	Natural and anthropogenic	Surface water and groundwater
Chile	500,000	100-1,000	Natural and anthropogenic;basin lakes, thermal springs, mining	Surface water
Greece	150,000	-	Natural and anthropogenic; thermal springs and mining	Surface water
Hungary, Rumania	400,000	2-176	Natural; alluvial sediments; organics	Surface water
China	>400,000	1-2,400	Natural; alluvial and lake sediments; high alkalinity	Groundwater
Mexico	400,000	8-620	Natural and anthropogenic; volcanic sediments, mining	Surface water and groundwater
Nepal	-	-	Natural, alluvia	Groundwater
Spain	>50,000	1-100	Natural; alluvial sediments	Surface water
Taiwan	>100,000	1-1,820	Natural	Groundwater
Thailand	15,000	1-5,000	Anthropogenic, mining	Surface water
Vietnam	>1,000,000	1-3,050	Natural, alluvia	Groundwater
West Bengal, India	>1,000,000	10-3,880	Natural, alluvia	Groundwater

1.2 Arsenic in food

Generally, arsenic in food is less than 0.25 mg/kg with the exception of seafood. Plant foods may be contaminated by using arsenical pesticides or by the deposition of arsenic in air from industry. Poultry may contain high level of arsenic concentration because of using organic arsenical feed additives. The commonly used additives in feed of poultry and swine include 4-aminophenylarsonic acid, 3-nitro-4-hydroxyphenylarsonic acid and

4-nitrophenylarsonic acid. So the arsenic concentration in pig and poultry liver and kidney is often over 1 mg/kg and even up to 10 mg/kg. Average concentration of arsenic in fish and shellfish is often greater than 5 mg/kg. Dabeka[9] gave some details in studies of arsenic in food: fish contains the highest mean arsenic concentration 1662 μg/kg; meat and poultry is about 24.3 μg/kg; bakery goods and cereals are about 24.5 μg/kg.

In addition to the total arsenic in food, the nature of arsenic species in food is also an important consideration. Yost et al[9] determined the total arsenic in various food items and separated arsenic crudely to inorganic (extractable by dilute hydrochloric acid) and organic fraction in Table 1.3.

Table 1.3 Arsenic concentrations in food (mg/kg, Wet Weight)

Food	Total	Inorganic	Organic	% Inorganic
Chicken	0.022	0.0090	0.012	41
Sole	4	0.022	4.4	1
Tuna	1.1	0.025	1.2	2
Pickerel	0.14	0.022	0.086	15
Shrimp	0.65	0.10	0.52	>16
Rice	0.24	0.1	0.16	43
Apple juice	0.012	0.0088	<0.002	73
Tea	0.035	0.0091	0.025	26

The most significant source of dietary arsenic is marine animals, and the organic arsenic dominates in marine animals. Arsenobetaine (AsB), arsenosugars, and tetramethylarsonium are found as the major arsenic compounds in marine organisms. Generally, AsB is over 50% and up to 95% of total arsenic in marine animals.

2. Human exposure to Arsenic

Human exposure to arsenic is mainly through oral and inhalation routes, the oral route of exposure includes intake from water, food, soil (and dust), and medicinal products. Inhalation occurs primarily by smelters and in occupations involved in mining/smelting

operations, agriculture or microelectronics. Limited studies have shown that dermal absorption may also be important. Both inorganic and organic arsenic compounds can be absorbed through human and animal skin. [9]

For the general populations, the main sources of arsenic exposure are from water and food. The World Health Organization (WHO), the United States Environmental Protection Agency (EPA) and Health Canada have set 10 μg/L as the guideline level for arsenic in drinking water. For a daily consumption of 2 L water containing <10 μg/L arsenic, the maximum daily intake of arsenic from water is <20 μg [9]. Most of this arsenic is likely to be inorganic. But individuals consuming water from groundwater containing high concentrations of arsenic (0.2-0.5 mg/L) will have much high daily intakes (0.4-1 mg).

In the general population, food is also a main source of arsenic exposure. Table 1.4[9] shows the daily dietary ingestion of total arsenic in some countries.

Table 1.4 Daily ingestion of total arsenic

Country	Arsenic intake (μg/day)	Reference
United States	62	Gartrell, 1985
United Kingdom	89	Food Additives and Contaminants Committee, 1984
Austria	27	Pfannhauser and Pechanek , 1977
New Zealand	55	Dick. 1978
Canada	38.1	Dabeka, 1993

The total arsenic intakes and the species of arsenic intakes from food highly depend on individuals' dietary habits. Regular consumption of offal from pigs or poultry which have received arsenicals in their feed could result in high arsenic intake; the nature of this arsenic species is unknown. Fish is also the major source of the average dietary arsenic intake. In some cases, the arsenic intakes from fish are likely to be about 1 mg/day assuming a balanced fish diet containing about 5 mg/kg of arsenic. Normally, most of the arsenic intake from fish is in the form of organoarsenic compounds, such as

arsenobetaine.

3. Arsenic speciation analysis

There are over thirty arsenic compounds existing in the environmental and biological systems. These arsenic compounds have a variety of toxicity. To understand the biogeochemistry, toxicity and metabolism of these arsenic compounds, various analytical technologies have been used to determine the individual arsenic compound. To achieve high selectivity and sensitivity in analysis, a combination of analytical techniques is often used in the determination of arsenic speciation.

3.1. Separation techniques

High performance liquid chromatography (HPLC) is the most commonly used technique in the separation of arsenic species. The major HPLC techniques used for arsenic speciation are ion exchange and ion pairing. Other separation techniques, such as gas chromatography, supercritical fluid chromatography, and capillary electrophoresis, have also been used, but less common for arsenic speciation.

Both anion and cation exchange chromatography have been used for separating arsenic compounds. Anion exchange [10-12] is most commonly used to separate As^{III}, As^{V}, MMA^{V}, and DMA^{V}, based on the difference in their pKa values. Anion exchange can separate arsenic species in a wide pH range. Cation exchange[13, 14] is used to separate arsenobetaine, arsenocholine, trimethylarsine oxide, and tetramethylarsonium ion.

Ion pair chromatography [15-17] has become a routine method to separate neutral and ionic arsenic species. Tetrabutylammonium (TBA)[18, 19] is the common ion pairing reagent for separation of As^{III}, As^{V}, MMA^{V}, DMA^{V}, DMA^{III}, and MMA^{III} in 7 minutes. Tetraethylammonium hydroxide (TEAH)[20-23] is another ion-pair reagent, used often for the separation of arsenobetaine and arsenosugars.

3.2. Detection techniques

High sensitivity is necessary for detection of trace amounts of arsenic in environmental and biological samples. Atomic spectrometry and mass spectrometry provide an excellent sensitivity for arsenic detection.

Inductively coupled plasma mass spectrometry (ICPMS) provides high sensitivity, multi-element capability and isotope ratio measurement capability. The optimum flow rate (1 mL/min) for both ICPMS and HPLC is compatible, making the interface between HPLC and ICPMS easy to accomplish. HPLC-ICPMS has become the most effective approach in arsenic analysis. HPLC coupled with electrospray ionization mass spectrometry (HPLC-ES-MS) has been used to identify arsenic species.[24-26]

Hydride generation (HG) is a chemical derivatization process to generate volatile hydrides by chemical reaction, typically by sodium borohydride. When HG is coupled with atom spectrometry, it can enhance the sensitivity by 10 to 100 folds. HG is also used as sample introduction interface between HPLC and atomic absorption spectrometry (AAS) and atomic fluorescence spectrometry (AFS).[27-29] Detection limit can be improved by coupling HG with ether AAS or AFS.[30-33] But the limitation of HG is that it only can detect the arsenic species that form gaseous hydride, such as As^{III}, As^{V}, MMA^{III}, MMA^{V}, DMA^{III}, DMA^{V}, and can not detect all arsenic species, such as arsenobetaine.

4. Arsenic metabolism and biomarker of arsenic

4.1. Distribution and biomarker of arsenic

The total arsenic in human body is about 3 to 4 mg[34], widely distributed in the body including the liver, kidney, lung, spleen and skin, with the highest concentration in the hair and nails because of high sulfydryl content of keratin. Other sites, such as uterus, bone, muscle and neural tissue, have been shown to contain detectable arsenic. Total arsenic in tissues can be detected, but the accurate distribution between trivalent and pentavalent arsenic still can not be elucidated due to the change of the original valence

state of the arsenic during the digestions of the tissues [35].

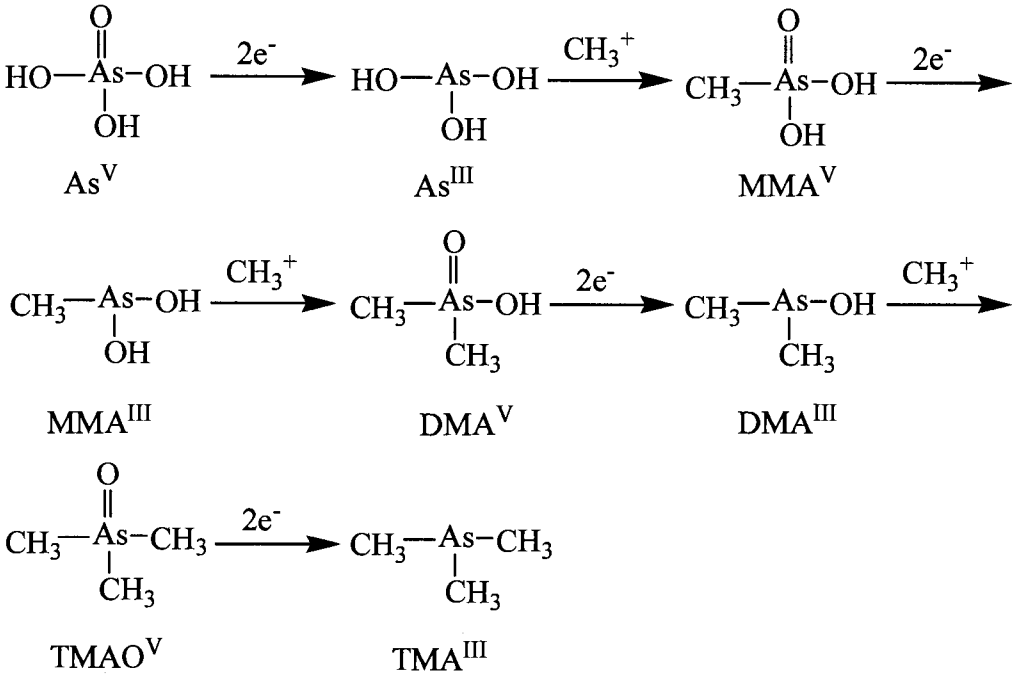
Arsenic is eliminated to urine, feces, nails and hair. Excretion by urine is the primary route for elimination of most arsenic species. Approximately 55–80% of the absorbed arsenic is excreted in urine and 5% in feces, DMA^V is the major form and accounts for 60 –80% of total arsenic in urine. The nontoxic AsB in marine animals can be excreted rapidly in urine without change, but arsenosugars are metabolized to DMA and other arsenic compounds[36]. Thus, high levels of arsenic in urine can indicate recent exposure to high concentration of inorganic arsenic species or certain food consumption. Generally, arsenic concentration in urine is used as a biomarker for the arsenic exposure because average concentrations of arsenic metabolites in the urine correlate with the levels of exposure, for example, the concentration of arsenic in drinking water.

Most of the absorbed arsenic has a short half-time in blood. Blood is a more difficult matrix than urine for speciation analysis[9], and thus, only the total arsenic concentration is generally reported. Partial separation[37, 38] has been used to detected total arsenic and some arsenic species in plasma, red blood cells, serum, and whole blood. But blood is not a reliable biomarker of arsenic exposure due to the wide variation of arsenic concentrations with respect to the sampling time. Although arsenic is also excreted to hair and fingernail, hair is not the accurate index for arsenic exposure due to the external contamination via air, water, soap, and shampoos.

4.2. Arsenic metabolism and toxicity

The metabolism of inorganic arsenic species involves redox reactions and methylation. A variety of species of yeast, fungi, algae, plants and animals can transform inorganic arsenic compounds to the methyl derivatives [7, 39-41]. Arsenosugars, arsenobetaine, arsenocholine and arsenolipids, commonly present in marine organisms, including lobsters, shrimps, mussels, oysters, and sea-weeds,[42, 43] may also be biotransformation products of inorganic arsenic. But how they are formed is not understood.

Methylation of arsenic is a process of a two-electron reduction of the pentavalent arsenic species to the trivalent arsenic species followed by oxidative addition of a methyl group to arsenic (Scheme 1) [7, 44]. S-adenosylmethionine (SAM)[45] is the methyl donor, and the reactions are catalyzed by methyltransferases in the presence of glutathione[39].



Scheme 1. Methylation process of arsenic (Reproduced from reference 18)

The toxicity of arsenic species is highly dependent on its oxidation state and chemical composition. For many years it was believed that the acute toxicity of inorganic arsenic was greater than organic arsenic. However, Cullen et al.[46] found that MMA^{III} is more toxic than arsenite to a microorganism. Human cells are also more sensitive to the cytotoxic effects of MMA^{III} than arsenite[47, 48]. DMA^{III} is at least as cytotoxic as arsenite in several human cell types[48].

Trivalent arsenic species are reactive to thiols, such as glutathione and cysteine. Methylated trivalent arsenicals such as MMA^{III} are potent inhibitors of GSH reductase and thioredoxin reductase. The pentavalent arsenate has a similar structure to phosphate and may interfere with oxidative phosphorylation by forming an unstable arsenate ester.

Thus, arsenate could influence phosphotransfer reactions, which are required for ATP generation. Arsenate accelerates the rate of ITP (inosine triphosphate) hydrolysis and inhibits both Ca^{2+} and Sr^{2+} uptake[49]. The perturbation of intracellular Ca^{2+} homeostasis activates protein kinase C (PKC) activity, which may play an important role in arsenic-induced genotoxicity.[50]

5. Arsenic carcinogenesis

Epidemiologic studies have shown associations between the exposure to high level arsenic and the prevalence of cancers of the bladder, lung, skin, and kidney. However, laboratory studies showed that arsenic itself was not mutagenic, although some deleterious effects of arsenic on DNA have been reported, including potentiation of DNA damage, inhibition of DNA repair, sister chromatid exchange, and gene ampliation. The exact mechanism of action responsible for arsenic carcinogenesis is not clear. Numerous hypotheses have been proposed, including oxidative stress, genotoxic damage, DNA repair inhibition, and activation of signal transduction pathways.

5.1. Effects of DNA damage and DNA repair inhibition by arsenic

Arsenic interferes with DNA methyltransferases[51], resulting in inactivation of tumor suppressor genes through DNA hypermethylation. Other studies suggest that arsenic-induced malignant transformation is linked to DNA hypomethylation subsequent to depletion of S-adenosyl-methionine. Both hypo- and hyper-methylation of DNA could cause aberrant expression of genes (such as oncogenes or tumour-suppressor genes) which in turn cause abnormality in cell proliferation leading to carcinogenesis. Metabolic methylation of inorganic arsenic is a genotoxic-enhancing process because it is involved in induction of DNA damage and DNA single-strand breaks resulting from the inhibition of repair polymerization.[52, 53] It is possible that arsenic-mediated DNA–protein interactions may play a major role in arsenic carcinogenesis, and the induced protein

associated DNA-strand breaks could provide an explanation for chromosome aberration[54]. Another recent research[55] showed that arsenite and MMA^{III} increased the formation of DNA adduct with benzo[α] pyrene diol epoxide (BPDE).

DNA repair is inhibited in arsenic-treated cells, and this inhibition may result in a co-mutagenic effect with X-rays, UV light, and several chemicals. Recent research[56] showed pre-treatment of human cells with arsenic impairs the nucleotide excision repair pathway and leads to enhanced UV mutagenesis. Other research [57] showed a decreased expression of critical genes involved in nucleotide excision repair of damaged DNA in human cells treated by arsenite and the trivalent methylated metabolites. It is possible that the inhibition of DNA repair by arsenite is via effects on DNA ligase. DNA ligase activity in nuclear extracts has been shown to be decreased by arsenite[58]; but the enzyme level was not directly affected[58, 59], suggesting that arsenite may indirectly inhibit DNA ligase activity by altering cellular redox levels or by affecting signal transduction pathways and phosphorylation of proteins related to DNA ligase activity[59]. Arsenite (≥5 μM), the trivalent (≥2.5 μM) and the pentavalent (≥250 μM) methyl arsenic metabolites diminished their repair at non-cytotoxic concentrations. All trivalent arsenicals were able to release zinc from the human XPA protein (XPAzf); furthermore, MMA^{III} and DMA^{III} inhibited the activity of isolated *E. coli* zinc finger protein Fpg, both involved in DNA repair (NER and BER).

Eukaryotic cells respond to DNA strand breaks by stimulating the enzyme poly(ADP-ribose) polymerase, and this enzyme may have a role in DNA repair. A significant dose-dependent decrease in activity of this enzyme in human T-cell lymphoma cells has been observed after treatment with arsenic [60]. At a dose of 10μM arsenite, there is approximately a 50% decrease in enzyme activity and 80% decrease in cell viability. The enzyme contains two vicinal sulfhydryl groups, and arsenite may bind one or both groups and inhibit the enzyme. Similarly, poly(ADP-ribosyl)ation[61], which is predominantly mediated by poly(ADP-ribose) polymerase-1 (PARP-1), is inhibited at

concentrations as low as 10 nM arsenite in cultured HeLa cells. This concentration of arsenic is similar to concentrations in blood and urine of the general population. Since poly(ADP-ribosyl)ation is an immediate cellular response to DNA damage, playing a major role in DNA base excision repair and the maintenance of genomic stability, its inhibition by arsenite may add to the risk of cancer formation under low-exposure conditions.

5.2. Signal transduction pathways and gene expression induced by arsenic

Signal transduction is a process by which an extracellular signal is transmitted into the cell through the plasma membrane and along an intracellular chain of signaling molecules to stimulate a cellular response. Cell transformation is a complex process involving a variety of transcription factors and signaling pathways.

Arsenic affects the gene transcription, expression and activation of numerous signaling proteins, including growth factor receptors , G-proteins such as ras, tyrosine kinases such as c-src, mitohen-activated protein kinase (MAPKs), and nuclear transcription factors such as NFκB, AP-1, and p53. These effects may involve either activation or inactivation. Effects may be directly through the interaction of arsenic with proteins, or indirectly through the formation of arsenic-induced reactive oxygen species (ROS).

The p53 gene is mutated in greater than 50% of all human cancers. The p53 protein functions as a transcription factor, and is involved in the activation of apoptosis-inducing genes. The binding of p53 to DNA requires zinc, which may make the p53 protein susceptible to inactivation by transition metals since they are known to substitute for zinc. In addition, p53 can be activated by changes in oxidative conditions within the cell, due to its content of labile cysteine residues.

The role of p53 in arsenic-related cellular effects is still not well understood. The effects of arsenic on p53 are dependent on the cell type and the chemical form of arsenic.

Conflicting results indicate that arsenic has no effect on p53, decreases p53 expression, or induces p53 phosphorylation and accumulation [62]. Boonchai et al[62]. compared the immunohistochemical expression of the p53 protein in patients with arsenic-related basal cell carcinomas (BCCs) to p53 expression in patients with other types of BCCs. Results indicated a lower level of p53 expression in arsenic-related BCCs compared to expression in sporadic BCCs. In contrast, another group found[62] that the p53 gene was overexpressed in patients from the blackfoot disease endemic area of Taiwan with arsenic related skin cancers. In addition, a report[62] indicated a high incidence of p53 gene mutations (28.6–55.5% of patients) in all 3 skin cancer conditions. Another group[62] examined p53 in premalignant or malignant skin lesions from patients treated with arsenic, and the accumulation of the p53 protein was detected in 78% of the lesions from cases with arsenic exposure.

6. Arsenic binding to peptides and proteins

The binding of trivalent arsenicals to thiol groups in proteins, resulting in possible inhibition and alteration of biological functions of the proteins, can be considered as a possible mechanism of arsenic toxicity. Many studies have investigated the binding between the trivalent arsenicals and various thiols, including small thiol ligands peptides, and proteins.

6.1 Arsenic binding to small molecules containing thiols

Trivalent arsenic species have a high affinity for thiol-containing ligands. The studies [63-65] on arsenic affinity to different thiol ligands have demonstrated the thermodynamics of the As^{III}-thiol interaction: arsenicals can form moderately stable complexes with glutathione (GSH), dithiols dimercaptosuccinic acid (DMSA) and dihydrolipoic acid (DHLA); MMA^{III} has a higher affinity than arsenite for thiol ligands; the formation of As-thiolate bond is exothermic. $As^{III}(GS)_3$ and $MMA^{III}(GS)_2$ are more stable than $DMA^{III}(GS)$. The behavior of As^{III}-glutathione system is very similar to that of

As^{III}-cysteine. At physiological pH, almost 80% of the ligand (GSH and cysteine) is present as $As(L)_3$.

6.2 Arsenic binding to peptide models containing thiols

An important issue concerning the biological effects of arsenic[66] is the protein conformational changes accompanying the binding. Binding of As^{III} might induce global changes in protein structure and consequent loss of biological activity. To understand how arsenic binding influences the conformations of individual secondary structural elements of proteins, different peptide models have been synthesized to study the arsenic-peptide binding behavior.

Daniel [66, 67] studied the effect of As^{III} binding on α-helical structure and β-hairpin structure in synthesized peptide models. They first studied the binding of As^{III} to α-helical peptides containing two cysteine (cys) residues in i, and i + 1, i + 2, or i + 3 arrangements and various spatial locations. Binding occurred with alteration of the helical backbone for nearly all of the cysteine-containing peptides with exception of cysteines in the i and i + 4 sequential positions. Arsenic binding to the i and i + 3 positions resulted in disrupting helical structure, and forming a relatively stable alternate fold. Circular dichroism (CD) spectroscopy showed that As^{III} binding caused helical destabilization when cysteine residues were located at central or C-terminal regions of the helix. To understand the influences of arsenic binding on β-structure, pairs of cysteines were introduced into a model monomeric β-hairpin to yield peptides that have arsenic binding occurring either across the strands or within the same strand of the β-hairpin. Results showed cysteines at non-hydrogen bond (NHB) positions across the β-hairpin with the same face of the sheet enhanced the β-structure. Cross-strand cysteine residues on opposite faces close to the termini of the hairpin could still bind arsenic tightly and showed modestly increased β-sheet content. However, As^{III} binding to nonopposed cysteines, or to cysteines at HB and NHB positions along one strand of the hairpin, led to loss of β-structure.

Kitchin[68-70] studied the binding of arsenic to a series of synthesized peptides in α-helix on the effect of increasing length between two cysteines in the peptide. Results showed that peptides with 0-14 intervening amino acids between two cysteines bound to arsenite with K_d values of 2.7-20.1 μM, there was a little difference in arsenic binding affinity between peptides with 0-14 intervening amino acids. The binding capacity of peptides was proportional to the number of free cysteines. Schmidt[71] studied the different complexes of peptides and proteins with arsenic, copper, and zinc species by mass spectrometry. They showed that covalent bond between arsenic and sulfur was stable at both acidic and neutral pH values. The arsenic binding strongly depended on the reduction environment of the system. In the presence of Cu^{2+} and Zn^{2+}, no arsenic binding was detected in model protein lysozyme, and the Cu- and Zn-lysozyme complexes with a stoichiometry of 1:1 and 2:1 for Zn^{2+} ions and Cu^{2+} ions were observed, which implied arsenic could not replace zinc in the zinc finger structure.

6.3. Arsenic binding to proteins

Many studies have focused on arsenic binding to proteins including binding of arsenicals to proteins during reductive methylation[72-74]. Studies [75-77] also show that ArsR, ArsC and ArsA ATPase can bind to trivalent arsenic by one to three cysteines. Rosen[77] studied the spatial distance of three cysteines in ArsA ATPase, and found that cys113, cys172 and cys422 are located within a distance of 3-6 Å from each other. They also proposed an arsenite-cysteine interaction model that As^{III} interacts with the three cysteines in a trigonal pyramidal geometry. The study[78] of the binding of phenylarsinoxide to Arg-tRNA protein transferase indicated that vicinal thiol groups were not necessary for arsenic binding.

Thioredoxin[79], tubulin and actin[80], thioredoxin peroxidase II and galectin I[81, 82] have been shown to bind to arsenic in vitro. The study of sodium arsenite binding to galectin[82] shows that the binding causes the structural change of galectin I, and also slightly affects

the biological activity of GALI. Hemoglobin[83] also binds to trivalent arsenicals in vitro and in vivo in both human and rat bloods. Metallothionein[84, 85] can bind to trivalent arsenicals in acidic condition, in which Zn^{2+} and Cd^{2+} coordinated in metellothionein are released.

Only very small numbers of proteins have been identified as arsenic binding proteins. Without identifying more arsenic-binding proteins, it is difficult to obtain any comprehensive understanding of the effects of arsenic on biological function of proteins in cells. Thus, selection and identification of arsenic specific binding proteins will play an important role in understanding arsenic effect on human health.

7 Affinity techniques and mass spectrometry

7.1 Affinity purification

Specific proteins (or group of proteins) can be purified from crude cell extracts by affinity chromatography. Affinity chromatography (affinity purification) applies specific binding interactions between molecules. A particular ligand is chemically immobilized or "coupled" to a solid support so that when a mixture is passed through the column, only those molecules having specific binding affinity to the ligand are retained. Affinity purification generally involves the following steps:

1. Incubate crude mixture with the immobilized ligand support to allow the target molecule in the sample to bind to the immobilized ligand.
2. Wash away non-bound components from solid support.
3. Elute the target molecules from the immobilized ligand by altering the buffer conditions that weaken the binding interaction.

A support or matrix in affinity purification is any material to which a biospecific ligand may be covalently attached. Typically, affinity matrix is a solid. Hundreds of substances have been described and employed as affinity matrices. Useful affinity supports are those with a high surface area to volume ratio, chemical groups that are

easily modified for covalent attachment of ligands, minimal nonspecific binding properties, good flow characteristics, and mechanical and chemical stability.

Ligands can be immobilized directly to solid support material by formation of covalent chemical bonds between particular functional groups on the ligand (e.g., primary amines, sulfhydryls, carboxylic acids, aldehydes) and reactive groups (isothiocyanates, isocyanates, NHS esters, carbodiimides, epoxides, anhydrides, sulfonyl chlorides) on the support. Ligands also can be coupled to support by simple adsorption, entrapment and encapsulation.

Different kinds of ligands can be immobilized on a solid according to the selection purposes. For example, to select metal-binding proteins, metal[86, 87] as a ligand was immobilized on chelating resin. Nucleic acids (RNA) and cyclic nucleotide derivatives[88, 89] are also used as ligands to select nucleotide interaction proteins. Phosphate donor[90] is also a ligand for affinity column to purify the thymidine kinase in rat liver. An arsenic affinity column[80, 81, 91-93] was prepared and applied to select arsenic binding proteins from protein mixture.

7.2 Mass spectrometry and proteomics

With the development of "soft" ionization methods, electrospray ionization (ESI)[94] and matrix-assisted laser desorption ionization (MALDI)[95] in the late 1980s, polypeptides were accessible to mass spectrometric analysis. This catalyzed the development of complex multistage instruments[96], for example, hybrid quadrupole time-of-flight (Q-Q-ToF) and tandem time-of-flight (ToF-ToF), to meet the challenges of protein and proteome analysis.

Mass spectrometers are used to measure the molecular mass of a molecule and to determine additional structural features including the amino acid sequence and type of posttranslational modifications. Single-stage mass spectrometers are used to measure molecular mass. After the initial mass determination, specific ions are selected and

subjected to fragmentation through collision in multistage instrument. In such experiments, known as tandem mass spectrometry (MS/MS), detailed structural features of the peptides are inferable from the analysis of the masses of the resulting fragments. The advantages of femtomole sensitivity, high resolution and mass accuracy, high-quality MS and MS/MS data, and rapid data collection enable mass spectrometry to be an effective tool for protein identification and detailed investigation of post-translational modifications. The availability of automatic sample analysis and database searching provide mass spectrometry high-throughput features. Mass spectrometry has quickly become a powerful tool of large-scale proteome research.

Two common approaches can achieve protein identification. The first way is gel digestion and Mass Spectrometry analysis. Proteins are separated by one-dimensional gel electrophoresis or two-dimensional gel electrophoresis. Then gel digestion is applied and peptides are extracted from the gel for mass spectrometry analysis. The second way is in-solution protein digestion. The digested peptide solution is fractionated and analyzed by two-dimensional HPLC separation and mass spectrometry. The MS/MS data from mass spectrometry analysis are submitted for database search.

8 Research rationale, objectives, and overview of the thesis

Arsenic has long been considered a human carcinogen. Exposure to high levels of arsenic can cause a wide range of health effects, including cancers of the bladder, lung, skin, and kidney. But the mechanisms of action regarding the effects of arsenic on health are still unclear. One possible mechanism is the binding of arsenic to the sulfhydryl group of cysteine in proteins, thus affecting the functions of the proteins. It has been shown that arsenic can suppress DNA repair in cells. It is possible that arsenic binds to the DNA repair proteins and affects their functions.

The primary objective of this thesis research is to study arsenic interaction with proteins. Two aspects are investigated, one is to study arsenic binding to purified proteins,

and the other is to select and identify proteins that bind to arsenic by developing an affinity technique.

We first chose two DNA repair proteins p53 and poly (ADP-ribose) polymerase (PARP-1) and studied their binding to arsenic (chapter 2). These proteins are involved in nucleotide-excision and base-excision DNA repair.

We then decided to select and identify all (or as many as possible) proteins that bind to arsenic. We developed an affinity technique to capture arsenic specific binding proteins from a large pool of many proteins in cells, and used mass spectrometry to identify the captured proteins. This technique allowed for the capture and mass spectrometry identification of over 70 arsenic-binding proteins (chapter 3). We further applied this technique to the capture and identification of arsenic-binding proteins in human cells treated with different carcinogens (chapter 4). We also explored the use of the affinity material for HPLC separation of arsenic-binding proteins (chapter 5).

9. References

1. Lenihan, J.; Fletcher, W. W., *Environment and man, Volume 6, The chemical environment.* Blackie, Glasgow and London: 1977.
2. Andreae, M. O.; Byrd, J. T.; Froehlich, P. N., Arsenic, antimony, germanium, and tin in the Tejo estuary, Portugal: modeling a polluted estuary. *Environ. Sci. & Technol.* **1983,** 17, (12), 731-737.
3. Rahman, M., Arsenic and contamination of drinking water in Bangladesh: a public health perspective. *J Health Population Nutrition* **2002,** 20, 193-197.
4. Wang, J. S.; Wai, C. M., Arsenic in Drinking Water—A Global Environmental Problem. *J. Chem. Educ.* **2004,** 81, (2), 207-213.
5. Chen, S. L.; Dzeng, S. R.; Yang, M. H.; Chiu, K. H.; Shieh, G. M.; Wai, C. M., Arsenic Species in Groundwaters of the Blackfoot Disease Area, Taiwan. *Environ. Sci. Technol.* **1994,** 28, (5), 877-881.
6. Smedley, P. L.; Nicolli, H. B.; Macdonald, D. M. J.; Barros, A. J.; Tullio, J. O., Hydrogeochemistry of arsenic and other inorganic constituents in groundwaters from La Pampa, Argentina. *Appl. Geochem.* **2002,** 17, (3), 259-284.
7. Cullen, W. R.; Reimer, K. J., Arsenic speciation in the environment. *Chem. Rev.* **1989,** 89, (4), 713-764.
8. Masscheleyn, P. H.; Delaune, R. D.; Patrick, W. H., Effect of redox potential and pH on arsenic speciation and solubility in a contaminated soil. *Environ. Sci. Technol.* 1991, (25), 8.
9. *Arsenic in Drinking Water.* National Research Council, National Academy Press, : Washington, DC, 1999, and 2001.
10. Mandal, B. K.; Ogra, Y.; Suzuki, K. T., Identification of dimethylarsinous and monomethylarsonous acids in human urine of the arsenic-affected areas in West Bengal, India. *Chem Res Toxicol* **2001,** 14, (4), 371-8.

11. Lintschinger, J.; Schramel, P.; Hatalak-Rauscher, A.; Wendler, I.; Michalke, B., A new method for the analysis of arsenic species in urine by using HPLC-ICP-MS *Fres.' J. Anal. Chem.* **1998.** , 362, 313.

12. Falk, K.; Emons, H., Speciation of arsenic compounds by ion-exchange HPLC-ICP-MS with different nebulizers *J. Anal. At. Spectrom.* **2000,** 15, 643.

13. Larsen, E. H.; Pritzl, G.; Hansen, S. H., Speciation of eight arsenic compounds in human urine by high-performance liquid chromatography with inductively coupled plasma mass spectrometric detection using antimonate for internal chromatographic standardization. *J. Anal. At. Spectrom.* **1993,** 8, 557.

14. Shibata, Y.; Morita, M., Characterization of organic arsenic compounds in bivalves. *Appl. Organomet. Chem.* **1992,** 6, 343.

15. Wangkarn, S.; Pergantis, S. A., High-speed separation of arsenic compounds using narrow-bore high-performance liquid chromatography on-line with inductively coupled plasma mass spectrometry *J. Anal. At. Spectrom.* **2000,** 15, 627.

16. Thomas, P.; Sniatecki, K., Determination of trace amounts of arsenic species in natural waters by high-performance liquid chromatography–inductively coupled plasma mass spectrometry. *J. Anal.l At. Spectrom.* **1995,** 10, 615.

17. Do, S. R.; Pradeau, D.; Guyon, F., Speciation of arsenic and selenium compounds by ion-pair reversed-phase chromatography with electrothermic atomic absorption spectrometry: Application of experimental design for chromatographic optimisation. *J. Chromatogr. A* **2001,** 918, 87.

18. Le, X. C.; Lu, X.; Ma, M.; Cullen, W. R.; Aposhian, H. V.; Zheng, B., Speciation of key arsenic metabolic intermediates in human urine. *Anal Chem.* **2000,** 72, 5172.

19. Le, X. C.; Ma, M.; Cullen, W. R.; Aposhian, H. V.; Lu, X.; Zheng, B., Determination of monomethylarsonous acid, a key arsenic methylation intermediate, in human urine. *Environ Health Perspect* **2000,** 108, 1015-8.

20. Shibata, Y.; Morita, M., Speciation of Arsenic by Reversed-Phase High Performance Liquid Chromatography-Inductively Coupled Plasma Mass Spectrometry. *Anal. Sci.* **1989,** 1, 107.

21. Le, X. C.; Cullen, W. R.; Reimer, K. J., Speciation of arsenic compounds in some marine organisms. *Environ. Sci. Technol.* **1994,** 28, 1598.

22. Ma, M. S.; Le, X. C., Effect of arsenosugar ingestion on urinary arsenic speciation *Clin. Chem.* **1998,** 44, 539.

23. Koch, I.; Wang, L.; Ollson, C. A.; Cullen, W. R.; Reimer, K. J., The Predominance of Inorganic Arsenic Species in Plants from Yellowknife, Northwest Territories, Canada. *Environ. Sci. Technol.* **2000,** 34, 22.

24. McSheehy, S. and Szpunar, J., Speciation of arsenic in edible algae by bi-dimensional size-exclusion anion exchange HPLC with dual ICP-MS and electrospray MS/MS detection. *J. Anal. At. Spectrom.* **2000,** 15, 79.

25. Madsen, A. D.; Goessler, W.; Pedersen, S. N.; Francesconi, K. A., Characterization of an algal extract by HPLC-ICP-MS and LC-electrospray MS for use in arsenosugar speciation studies *J. Anal. At. Spectrom.* **2000,** 15, 657.

26. McSheehy, S.; Pohl, P.; Obiski, R.; Szpunar, J., Investigation of arsenic speciation in oyster test reference material by multidimensional HPLC-ICP-MS and electrospray tandem mass spectrometry (ES-MS-MS) *Analyst,* **2001,** 126, 1055.

27. Howard, A. G.; Hunt, L. E., Coupled photooxidation-hydride AAS detector for the HPLC of arsenic compounds. *Anal Chem* **1993,** 65, 2995.

28. Tsalev, D. L.; Sperling, M.; Welz, B., Speciation determination of arsenic in urine by high-performance liquid chromatography–hydride generation atomic absorption spectrometry with on-line ultraviolet photooxidation. *Analyst* **1998,** 123, 1703.

29. Tsalev, D. L.; Sperling, M.; Welz, B., Flow-injection hydride generation atomic absorption spectrometric study of the automated on-line pre-reduction of arsenate, methylarsonate and dimethylarsinate and high-performance liquid chromatographic separation of their-cysteine complexes *Talanta* **2000,** 51, 1059.

30. Le, X. C. ; Ma, M., Short-Column Liquid Chromatography with Hydride Generation Atomic Fluorescence Detection for the Speciation of Arsenic. *Anal Chem* **1998,** 70, 1926.

31. Le, X. C.; Ma, M.; Wong, N. A., Speciation of Arsenic Compounds Using High-Performance Liquid Chromatography at Elevated Temperature and Selective Hydride Generation Atomic Fluorescence Detection. *Anal Chem* **1996,** 68, 4501.

32. Corns, W. T.; Stockwell, P. B.; Ebdon, L.; Hill, S. J., Development of an atomic fluorescence spectrometer for the hydride-forming elements. *J. Anal. At. Spectrom.,* **1993,** 8, 71.

33. Bohari, Y.; Astruc, A.; Astruc, M.; Cloud, J., Improvements of hydride generation for the speciation of arsenic in natural freshwater samples by HPLC-HG-AFS. *J. Anal. At. Spectrom.,* **2001,** 16, 774.

34. *Medical and biological effects of environmental pollutants - arsenic.* National Academy of Sciences, Washington, D.C.: 1977.

35. Lauwerys, R. R.; Buchet, J. P.; Roels, H., The determination of trace levels of arsenic in human biological material. *Arch. Toxicol.* **1979,** 41, 239-247.

36. Le, X. C.; Cullen, W. R.; Reimer, K. J. Human urinary arsenic excretion after one-time ingestion of seaweed, crab, and shrimp *Clin. Chem.* **1994,** 40, (4), 617-625.

37. Concha, G.; Vogler, G.; Lezcano, D.; Nermell, B.; Vahter, M., Exposure to Inorganic Arsenic Metabolites during Early Human Development *Toxicol. Sci.* **1998,** 44, (2), 185-190.

38. Zhang, X.; De Kimpe, R. C., J.; Mees, L.; Vanderbiesen, V.; De Cubber, A. and Vanholder, R., Accumulation of arsenic species in serum of patients with chronic renal disease *Clin. Chem.* **1996,** 42, (8), 1231-1237.

39. Styblo, M.; Thomas, D. J., *Arsenic: Exposure and Health Effects.* Chapman & Hall, UK: 1997.

40. Aposhian, H. V., Enzymatic methylation of arsenic species and other new approaches to arsenic toxicity. *Annu Rev Pharmacol Toxicol* **1997,** 37, 397-419.

41. Maeda, S., *Arsenic in Environment Part I: Cycling and Characterization* John Wiley & Sons Inc.: 1994.

42. Edmond, J. S.; Francesconi, K. A., Arseno-sugars from brown kelp (Ecklonia radiata) as intermediates in cycling of arsenic in a marine ecosystem *Nature* **1981,** 289, 602.

43. Edmonds, J. S.; Francesconi, K. A., Transformations of arsenic in the marine environment. *Experimentia* **1987,** 43, 553.

44. Thompson, D. J., A chemical hypothesis for arsenic methylation in mammals. *Chem Biol Interact* **1993,** 88, (2-3), 89-14.

45. Buchet, J. P.; Lauwerys, R., Study of inorganic arsenic methylation by rat liver in vitro: relevance for the interpretation of observations in man. *Arch Toxicol* **1985,** 57, (2), 125-9.

46. Cullen, W. R.; McBride, B. C., The metabolism of methylarsine oxide and sulfide. *Appl. Organomet. Chem.* **1989,** 3, 71.

47. Petrick, J. S.; Ayala-Fierro, F.; Cullen, W. R., Monomethylarsonous acid (MMA(III)) is more toxic than arsenite in Chang human hepatocytes. *Toxicol. Appl. Pharmacol.* **2000,** 163, 203.

48. Styblo, M.; Del Razo, L. M., Comparative toxicity of trivalent and pentavalent inorganic and methylated arsenicals in rat and human cells. *Arch. Toxicol.* **2000,** 74, 289.

49. Alves, E. W.; De Meis, L., Effects of arsenate on the Ca2+ ATPase of sarcoplasmic reticulum. *Eur J Biochem* **1987,** 166, (3), 647-51.

50. Liu, Y.-C.; Huang, H., Involvement of calcium-dependent protein kinase C in arsenite-induced genotoxicity in chinese hamster ovary cells. *J. Cell Biochem.* **1997,** 64, 423.

51. Goering, P. L.; Aposhian, H. V.; Mass, M. J.; Cebrian, M.; Beck, B. D.; Waalkes, M. P., The enigma of arsenic carcinogenesis: role of metabolism. *Toxicol Sci* **1999,** 49, (1), 5-14.

52. Yamanaka, K.; Hayashi, H.; Tachikawa, M.; Kato, K.; Hasegawa, A.; Oku, N.; Okada, S., Metabolic methylation is a possible genotoxicity-enhancing process of inorganic arsenics. *Mutat Res* **1997,** 394, (1-3), 95-101.

53. Hartwig, A.; Groblinghoff, U. D.; Beyersmann, D.; Natarajan, A. T.; Filon, R.; Mullenders, L. H., Interaction of arsenic(III) with nucleotide excision repair in UV-irradiated human fibroblasts. *Carcinogenesis* **1997,** 18, (2), 399-405.

54. Dong, J. T.; Luo, X. M., Arsenic-induced DNA-strand breaks associated with DNA-protein crosslinks in human fetal lung fibroblasts. *Mutat Res* **1993,** 302, (2), 97-102.

55. Schwerdtle, T.; Walter, I.; Hartwig, A., Arsenite and its biomethylated metabolites interfere with the formation and repair of sTable BPDE-induced DNA adducts in human cells and impair XPAzf and Fpg. *DNA Repair (Amst)* **2003,** 2, (12), 1449-63.

56. Danaee, H.; Nelson, H. H.; Liber, H.; Little, J. B.; Kelsey, K. T., Low dose exposure to sodium arsenite synergistically interacts with UV radiation to induce mutations and alter DNA repair in human cells. *Mutagenesis* **2004,** 19, (2), 143-8.

57. Andrew, A. S.; Karagas, M. R.; Hamilton, J. W., Decreased DNA repair gene expression among individuals exposed to arsenic in United States drinking water. *Int J Cancer* **2003,** 104, (3), 263-8.

58. Li, J. H.; Rossman, T. G., Inhibition of DNA ligase activity by arsenite: a possible mechanism of its comutagenesis. *Mol Toxicol* **1989,** 2, (1), 1-9.

59. Hu, Y.; Su, L.; Snow, E. T., Arsenic toxicity is enzyme specific and its affects on ligation are not caused by the direct inhibition of DNA repair enzymes. *Mutat Res* **1998,** 408, (3), 203-18.

60. Yager, J. W.; Wiencke, J. K., Inhibition of poly(ADP-ribose) polymerase by arsenite. *Mutat Res* **1997,** 386, (3), 345-51.

61. Hartwig, A.; Pelzer, A.; Asmuss, M.; Burkle, A., Very low concentrations of arsenite suppress poly(ADP-ribosyl)ation in mammalian cells. *Int J Cancer* **2003,** 104, (1), 1-6.

62. Bode, A. M.; Dong, Z., The paradox of arsenic: molecular mechanisms of cell transformation and chemotherapeutic effects. *Crit Rev Oncol Hematol* **2002,** 42, (1), 5-24.

63. Spuches, A. M.; Kruszyna, H. G.; Rich, A. M.; Wilcox, D. E., Thermodynamics of the As(III)-thiol interaction: arsenite and monomethylarsenite complexes with glutathione, dihydrolipoic acid, and other thiol ligands. *Inorg Chem* **2005,** 44, (8), 2964-72.

64. Rey, N. A.; Howarth, O. W.; Pereira-Maia, E. C., Equilibrium characterization of the As(III)-cysteine and the As(III)-glutathione systems in aqueous solution. *J Inorg Biochem* **2004,** 98, (6), 1151-9.

65. Raab, A.; Meharg, A. A.; Jaspars, M.; Genney, D. R.; Feldmann, J., Arsenic–glutathione complexes—their stability in solution and during separation by different HPLC modes. *Anal. At. Spectrom.* **2004,** 19, 183.

66. Ramadan, D.; Cline, D. J.; Bai, S.; Thorpe, C.; Schneider, J. P., Effects of As(III) binding on beta-hairpin structure. *J Am Chem Soc* **2007,** 129, (10), 2981-8.

67. Cline, D. J.; Thorpe, C.; Schneider, J. P., Effects of As(III) binding on alpha-helical structure. *J Am Chem Soc* **2003,** 125, (10), 2923-9.

68. Kitchin, K. T.; Wallace, K., Arsenite binding to synthetic peptides based on the Zn finger region and the estrogen binding region of the human estrogen receptor-alpha. *Toxicol Appl Pharmacol* **2005,** 206, (1), 66-72.

69. Kitchin, K. T.; Wallace, K., Dissociation of arsenite-peptide complexes: triphasic nature, rate constants, half-lives, and biological importance. *J Biochem Mol Toxicol* **2006,** 20, (1), 48-56.

70. Kitchin, K. T.; Wallace, K., Arsenite binding to synthetic peptides: the effect of increasing length between two cysteines. *J Biochem Mol Toxicol* **2006,** 20, (1), 35-8.

71. Schmidt, A. C.; Koppelt, J.; Neustadt, M.; Otto, M., 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, (2), 153-63.

72. Naranmandura, H.; Suzuki, N.; Suzuki, K. T., Trivalent arsenicals are bound to proteins during reductive methylation. *Chem Res Toxicol* **2006,** 19, (8), 1010-8.

73. Styblo, M.; Thomas, D. J., Binding of arsenicals to proteins in an in vitro methylation system. *Toxicol Appl Pharmacol* **1997,** 147, (1), 1-8.

74. Pizarro, I.; Gómez, M.; Cámara, C.; Palacios, M. A.; Roman-Silva, D. A., Evaluation of arsenic species–protein binding in cardiovascular tissues by bidimensional chromatography with ICP-MS detection. *J. Anal. At. Spectrom.* **2004,** 19, 292.

75. Shi, W.; Dong, J.; Scott, R. A.; Ksenzenko, M. Y.; Rosen, B. P., The role of arsenic-thiol interactions in metalloregulation of the ars operon. *J Biol Chem* **1996,** 271, (16), 9291-7.

76. DeMel, S.; Shi, J.; Martin, P.; Rosen, B. P.; Edwards, B. F., Arginine 60 in the ArsC arsenate reductase of E. coli plasmid R773 determines the chemical nature of the bound As(III) product. *Protein Sci* **2004,** 13, (9), 2330-40.

77. Bhattacharjee, H.; Rosen, B. P., Spatial proximity of Cys113, Cys172, and Cys422 in the metalloactivation domain of the ArsA ATPase. *J Biol Chem* **1996,** 271, (40), 24465-70.

78. Li, J.; Pickart, C. M., Binding of phenylarsenoxide to Arg-tRNA protein transferase is independent of vicinal thiols. *Biochemistry* **1995,** 34, (48), 15829-37.

79. Schmidt, A. C.; Neustadt, M.; Otto, M., 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, (6), 771-80.

80. Menzel, D. B.; Hamadeh, H. K.; Lee, E.; Meacher, D. M.; Said, V.; Rasmussen, R. E.; Greene, H.; Roth, R. N., Arsenic binding proteins from human lymphoblastoid cells. *Toxicol Lett* **1999,** 105, (2), 89-101.

81. Chang, K. N.; Lee, T. C.; Tam, M. F.; Chen, Y. C.; Lee, L. W.; Lee, S. Y.; Lin, P. J.; Huang, R. N., Identification of galectin I and thioredoxin peroxidase II as two arsenic-binding proteins in Chinese hamster ovary cells. *Biochem J* **2003,** 371, (Pt 2), 495-503.

82. Lin, C. H.; Huang, C. F.; Chen, W. Y.; Chang, Y. Y.; Ding, W. H.; Lin, M. S.; Wu, S. H.; Huang, R. N., Characterization of the interaction of galectin-1 with sodium arsenite. *Chem Res Toxicol* **2006,** 19, (3), 469-74.

83. Lu, M.; Wang, H.; Li, X. F.; Lu, X.; Cullen, W. R.; Arnold, L. L.; Cohen, S. M.; Le, X. C., Evidence of hemoglobin binding to arsenic as a basis for the accumulation of arsenic in rat blood. *Chem Res Toxicol* **2004,** 17, (12), 1733-42.

84. Jiang, G.; Gong, Z.; Li, X. F.; Cullen, W. R.; Le, X. C., Interaction of trivalent arsenicals with metallothionein. *Chem Res Toxicol* **2003,** 16, (7), 873-80.

85. Merrifield, M. E.; Ngu, T.; Stillman, M. J., Arsenic binding to Fucus vesiculosus metallothionein. *Biochem Biophys Res Commun* **2004,** 324, (1), 127-32.

86. She, Y. M.; Narindrasorasak, S.; Yang, S.; Spitale, N.; Roberts, E. A.; Sarkar, B., Identification of metal-binding proteins in human hepatoma lines by immobilized metal affinity chromatography and mass spectrometry. *Mol Cell Proteomics* **2003,** 2, (12), 1306-18.

87. Smith, S. D.; She, Y. M.; Roberts, E. A.; Sarkar, B., Using Immobilized Metal Affinity Chromatography, Two-Dimensional Electrophoresis and Mass Spectrometry to Identify Hepatocellular Proteins with Copper-Binding Ability. *J. Proteome Res.* **2004,** 3, 834.

88. Dills, W. L. J.; Beavo, J. A.; Bechtel, P. J.; Myers, K. R.; Sakai, L. J.; Krebs, E. G., Binding of adenosine 3',5'-monophosphate dependent protein kinase regulatory subunit to immobilized cyclic nucleotide derivatives. *Biochemistry* **1976,** 15, (17), 3724-31.

89. Rodgers, J. T.; Patel, P.; Hennes, J. L.; Bolognia, S. L.; Mascotti, D. P., Use of biotin-labeled nucleic acids for protein purification and agarose-based chemiluminescent electromobility shift assays. *Anal Biochem* **2000,** 277, (2), 254-9.

90. Kizer, D. E.; Holman, L.; Ebner, K., Phosphate Donor Specificity and Some Kinetic Properties of Thymidine Kinase Purified from Regenerating Rat Liver by a Procedure Which Included Affinity-Gel Chromatography. *Proceedings of the Oklahoma Academy of Science* **1977,** 57, 44.

91. Hoffman, R. D.; Lane, M. D., Iodophenylarsine oxide and arsenical affinity chromatography: new probes for dithiol proteins. Application to tubulins and to components of the insulin receptor-glucose transporter signal transduction pathway. *J Biol Chem* **1992,** 267, (20), 14005-11.

92. Kalef, E.; Walfish, P. G.; Gitler, C., Arsenical-based affinity chromatography of vicinal dithiol-containing proteins: purification of L1210 leukemia cytoplasmic proteins and the recombinant rat c-erb A beta 1 T3 receptor. *Anal Biochem* **1993,** 212, (2), 325-34.

93. Donoghue, N.; Yam, P. T.; Jiang, X. M.; Hogg, P. J., Presence of closely spaced protein thiols on the surface of mammalian cells. *Protein Sci* **2000,** 9, (12), 2436-45.

94. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M., Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989,** 246, (4926), 64-71.

95. Karas, M.; Hillenkamp, F., Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* **1988,** 60, (20), 2299-301.

96. Domon, B.; Aebersold, R., Mass spectrometry and protein analysis. *Science* **2006,** 312, (5771), 212-7.

Chapter 2

Arsenic binding to DNA repair proteins in vitro

1. Introduction

1.1 Effect of arsenic on DNA damage and repair

Although the mechanism of arsenic carcinogenicity is still not clear, several modes of action have been proposed, including the induction of oxidative stress, inhibition of DNA repair, alteration of DNA methylation, enhanced cell proliferation and suppression of protein expression (e.g. p53)[1]. Some studies have shown that inorganic arsenic increased mutation frequency in *E. coli.* after UV light exposure[2], and also increased the mutagenicity and clastogenicity after UV light, benzo(α)pyrene, X-ray, and DNA cross-linking agents exposure in mammalian cells[3, 4]. Several studies revealed that arsenite interfered with the repair of UV-induced DNA damage [5-7]. Relatively low, non-cytotoxic concentrations of arsenic compounds inhibited base excision repair (BER) and nucleotide excision repair (NER)[5, 8, 9]. DNA ligase activity was inhibited by arsenite in nuclear extracts from arsenic-treated V79 cells[10]. A delayed rejoining of repair-mediated DNA strand breaks after UV irradiation and arsenic treatment in CHO cells was observed[11]. Research by Schwerdtle[12] showed that arsenite and MMA^{III} increased BPDE-DNA adduct, while both trivalent and pentavalent arsenic metabolites diminished DNA repair at non-cytotoxic concentrations. Fpg activity was inhibited by MMA^{III} and DMA^{III}, and all trivalent arsenic could release zinc from zinc finger domain of XPA with MMA^{III} and DMA^{III} being more reactive than arsenite.

1.2 P53 and PARP

Poly(ADP-ribosy)ation of nuclear proteins is one of the immediate cellular responses to DNA damage induced by chemotherapy agents (alkylating agents and topoisomerase I inhibitors). The reaction is mainly catalyzed by Poly (ADP-ribose) polymerase (PARP-1),

an abundant nuclear enzyme that mediates repair of DNA single strand breaks via the activation and recruitment of DNA repair enzymes. PARP is required for efficient DNA repair of single strand breaks and BER. Stimulated by single or double strand breaks of DNA, PARP-1 catalyzes the transfer of ADP-ribose units from NAD to histone H1 and H2B, to form protein-conjugated poly(ADP-ribose)[13-15]. This post-translational modification leads to the dissociation of proteins from DNA, which become more accessible for repair enzymes[16]. Two zinc finger motifs in the 170-N-terminal amino acid domain of PARP-1 are required for recognition of DNA strand breaks [17-19]. These two zinc finger structures contain three cysteine residues and one histidine (cys 21, 24, 56 and his 53; cys 125, 128, 162 and his 159).

Research on poly(ADP-ribose)polymerase (PARP-1) showed that poly(ADP-ribosy)ation is suppressed at low nanomolar concentration of arsenite in HeLa S3 cells[20]. Research by Yager[21] showed that arsenite decreased PARP-1 activity in a dose-dependent manner. Recent study by Walter[22] showed that the trivalent methylated metabolites MMA^{III} and DMA^{III} inhibited PARP-1 activity in cultured human HeLa S3 cells at concentrations as low as 1 nM. One possible reason[23] for DNA repair inhibition of PARP-1 by arsenic may be direct reaction of arsenic and cysteine in its DNA binding domain due to the high affinity of arsenic to the SH group

P53 is a transcription factor that regulates the cell cycle and functions as a tumor suppressor. It can induce or inhibit the expression of over 150 genes. These genes regulate the arrest of mammalian cell in G1 near the border of S phase or in G2, right before mitosis. P53 is also involved in DNA repair processes, cell growth /proliferation, cell cycle arrest, and apoptosis. P53 is comprised of 393 amino acids and has five domains: N-terminal transcription-activation domain (residues 1-42), which activates transcription factors; proline rich domain (residues 80-94), which is important for the apoptotic activity of p53; central DNA-binding core domain (Residues 102-292), which contains one zinc atom and several arginine amino acids; homo-oligomerisation domain

(residues 307-355), which is essential for the tetramerization of p53, and C-terminal residues (356-393) which is involved in downregulation of DNA binding of the central domain. The DNA binding core domain directly interacts with DNA in a sequence specific manner. X-ray crystallography[24] demonstrated the presence of tetrahedrally coordinated zinc in the DNA-binding domain. The DNA binding domain is made up of two β sheets supporting large loop/helix structures directly involved in contacting DNA, these loops are bridged together by the tetrahedral coordination of a zinc atom by three cysteines (cys 176, 238, and 242) and one histidine (His 179)[24, 25]. The DNA binding domain contains another five cysteines (cys 124, 135, 141, 275, and 277). Cysteines 275 and 277 form a C-X-C motif within the loop binding in the major groove of DNA. Cys 277 is exposed at the protein surface and donates a hydrogen bond to base in major group of DNA[24]. Site-directed mutagenesis of the residues cys 124, 135, 141, 275 in murine p53 showed that these cysteines may cooperate to modulate the structure of the DNA binding domain[26], these cyseines may be more accessible for redox modifications than those in binding of zinc. Other research [27, 28] confirmed that the redox status of a critical set of cysteines in the DNA-binding domain regulates p53 DNA-binding activity.

1.3 Hypothesis

Trivalent arsenic compounds have high affinity to the sulfhydryl (SH) group. Therefore, it is possible that arsenic binds to cysteine in proteins, thereby affecting the activity and recruitment of DNA repair and signaling proteins. We hypothesize that trivalent arsenic species can bind to two DNA repair proteins (p53 and PARP-1).

2. Experimental

2.1 Materials

Protein p53 was from Dr. C. Arrowsmith, Department of Medical Biophysics, University of Toronto. PARP-1 was from Dr. G. Porier, Laval University.

Monomethylarsonous acid (MMA^{III}) and dimethylarsinous acid (DMA^{III}) were from Dr. W. R. Cullen, University of British Columbia. Monomethylarsonic acid (MMA^{V}) was purchased from ChemService (West Chester, PA). Dimethylarsinic acid (DMA^{V}), phenylarsine oxide (PAO), ammonia acetate, tris(2-Carboxyethyl) phosphine Hydrochloride (TCEP), 50% ammonia solution, Tris, and Micro dialyzer (Spectra/Pro® from Spectrum®) were obtained from Sigma-Aldrich Canada (Oakville, ON).

2.2 Instrument

A PerkinElmer 200 series HPLC system (PE Instruments, Norwalk, CT, USA), 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 ZOBAX GF-250 column (250×4.6 mm, Agilent) column 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).

2.3 Dialysis of P53 and PARP-1 Proteins

P53 protein was stored in buffer containing 50 mM Tris, 200 mM NaCl, 1 mM Benzamidine, 0.5 mM phenylmethanesulfonyl fluoride (PMSF), 0.5 mM TCEP, pH 6.5. PARP-1 was store in buffer containing 50 mM Tris, 200 mM NaCl, 0.5 mM PMSF, 1 mM β-mercaptoethanol, pH 7.4. Before incubating a protein with arsenic compounds, the protein was dialyzed over night in Buffer (20 mM Tris, 50 mM NaCl, 0.1 mM TECP) at 4°C protected by N_2. Microdialysis tube was prepared according to the instruction of the manufacture.

2.4 Incubation of arsenicals with proteins

Protein after dialysis was incubated with 0.5 mM TCEP (pH=2) for 1 hour at 4°C. The

concentration of p53 after dialysis was 12 μM, PARP-1 was 2 μM. Arsenic species (MMA^{III}, DMA^{III}, MMA^{V}, and PAO) of various concentrations (3-30 μM) were then added to either the p53 or PARP-1 solution. The mixtures were incubated at room temperature for varying incubation times (2 hours to 24 hours).

2.5 Determination of protein-bound and unbound arsenic species

Arsenic bound to the protein and the unbound arsenic species were separated by using a size exclusion column (Phenomenex Biosep-SEC-S 2000 or Agilent ZOBAX GF-250) with 10 mM ammonia 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 ArO^+ (m/z 91) and SO^+ (m/z 48). Detection of SO^+ (due to proteins) along with AsO^+ was to support the identification of protein and arsenic in the same molecule.

3. Result and discussion

3.1 Size exclusion chromatography separation and ICP-MS detection

Size exclusion chromatography separates analytes based on their size. In the protein-arsenic incubation sample, protein has a much larger size than arsenic molecule, thus protein is expected to be eluted earlier than the arsenic compounds. If the arsenic binds to the protein, the bound arsenic will be eluted with protein and separated from the unbound small arsenic compound.

Arsenic can be detected as As^+ (m/z 75) by ICP-MS. However, As^+ can suffer isobaric interference from $ArCl^+$ which is formed between the plasma gas (Ar) and Cl^- present in biological samples. To reduce this interference, we use dynamic reaction cell (DRC) function of the ICPMS instrument. We introduce O_2 to react with the As^+ ion produced by ICP. The conversion of As^+ (m/z 75) to AsO^+ (m/z 91) eliminated the interference from $ArCl^+$ (m/z 75). Another benefit of the introduction of O_2 was to achieve detection of protein by monitoring SO^+ signal. Because the proteins binding to arsenic all contain

cysteines, detection of SO^+ can be used to monitor proteins. Elements of H, C, O, N can not be used for specific detection of proteins because they are the elements existing in air and in the atmospheric pressure ICP. Detection of S by S^+ (m/z 32) suffers isobaric interference by O_2^+ (m/z 32). The dynamic reaction function using O_2 shifts S^+ to SO^+ (m/z 48) eliminating the interference from O_2^+. $^{32}S^{16}O^+$ has been used to detected sulfur signal by ICP-DRC-MS in both small molecules and biomolecules[29, 30], and the interference of O_2^+ is avoided. In this experiment, oxygen DRC mode is used to detect AsO^+ as arsenic signal, SO^+ as protein signal.

3.2 Arsenic binding to PARP-1

After dialysis of protein, 0.5 mM TCEP was added into protein solution and incubated for 1 hr at 4°C. This was to reduce protein from oxidation. Then PARP-1 (2 μM) was incubated with 3 μM PAO, MMA^{III}, DMA^{III}, and MMA^{V} at room temperature for 2 hour. Dialysis buffer was used as control with the same treatment as PARP-1 protein solution. Both protein incubation samples and control samples were analyzed by HPLC-ICP-MS.

Figure 2.1 shows the chromatograms from the analyses of PARP-1/PAO incubation sample and control sample. In the SO^+ detection channel, there was a small peak at 3.5-4 min, which corresponds to protein PARP-1. The large peak after the protein peak is due to phenylmethanesulfonyl fluoride (PMSF) presenting in the protein sample. From the detection of AsO^+ signal, four peaks were detected in the PARP-1/PAO incubation sample, and three peaks were detected in the PAO sample without PARP-1. A small peak at 3.5-4 min in the PARP-1/PAO incubation sample matched with the peak in the chromatogram with detection of SO^+. Therefore, this small peak is due to arsenic bound to the protein PARP-1. The other three peaks eluted after 4 minutes in the AsO^+ detection channel were present in both the PARP-1/PAO incubation sample and the PAO control sample without protein. These later-eluting peaks are due to PAO and its oxidation products.

In addition to PAO, 3μM MMA^{III}, DMA^{III} and MMA^{V} were also incubated separately

with dialyzed PARP-1 sample. Figure 2.2, 2.3, and 2.4 show the results of HPLC-ICP-MS analyses of the arsenic compounds with or without the incubation with PARP-1. In the AsO^+ detection channel, there is no peak at 3.5-4 min corresponding to PARP-1/MMA^{III}, PARP-1/DMA^{III} or PARP-1/MMA^{V} incubation samples. These results suggest that MMA^{III}, DMA^{III} and MMA^{V} can not bind to PARP-1 under the given conditions. Although it is not known why these arsenic species did not bind to PARP-1, previous work has shown that PAO had stronger binding than these other arsenic species to hemoglobin[31]. The stronger binding of PAO to PARP-1 is consistent with previous findings.

There are 17 cysteines in PARP-1. Cys 21, 24, 56 and His 53 form one zinc finger; cys 125, 128, 162 and His 159 form the second zinc finger. Comparing the electronegativity of zinc (1.650) and arsenic (2.180), zinc should bind tighter to the sulfphydryl groups than arsenic. Thus, the possibility of arsenic binding to the cysteine in zinc finger is low. The rest of cysteines in PARP-1 are cys 256, 295, 298, 311, 321, 429, 845, and 908. Cysteines 256, 295, 298, 311, 321 are in the DNA binding domain. The cysteines 295 and 298 form –CXXC- structure, which is optimal for arsenic binding to cysteine[32]. Therefore, it is possible that arsenic binds to these two cysteines. Because there is no three dimensional conformation data of PARP-1 in peptide between 225 and 384, we can not predict how arsenic may bind to cys 256, 295, 298, 311, 321. From the existing three dimensional data of PARP-1 peptide, we find that cys 429 and 908 are buried inside the protein, and are probably difficult for arsenic binding. cys 845 is exposed to exterior of protein, and is surrounded by water molecules. Thus, it is possible for arsenic binding to cys 845. Because there is not enough data for protein conformation, it is difficult to predict which cysteine(s) in PARP-1 bind to PAO.

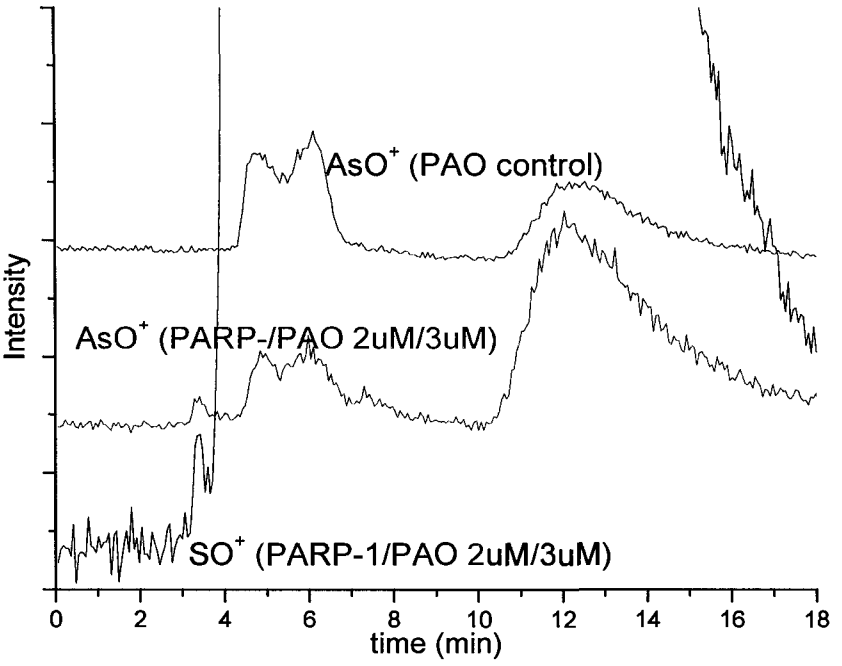


Figure 2.1 Chromatograms from the HPLC-ICPMS analyses of PARP-1/PAO incubation sample and PAO control sample. 3 μM PAO alone (blue trace) or 3 μM PAO and 2 μM PARP-1 were incubated in buffer (pH=2) for 2 hr (red and black trace). AsO^+ (m/z 91) was detected for As, and SO^+ (m/z 48) was detected for the present of protein

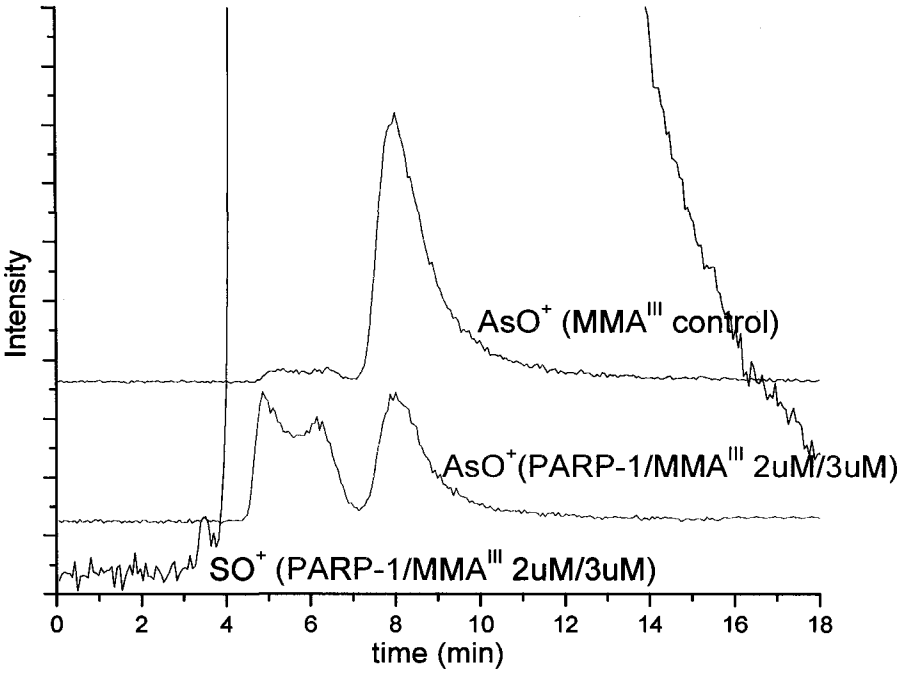


Figure 2.2 Chromatograms from the HPLC-ICPMS analyses of PARP-1/MMA^{III} incubation sample and MMA^{III} control sample. 3 μM MMA^{III} alone (blue trace) or 3 μM MMA^{III} and 2 μM PARP-1 were incubated in buffer (pH=2) for 2 hr (red and black trace). AsO^+ (m/z 91) was detected for As, and SO^+ (m/z 48) was detected for the present of protein

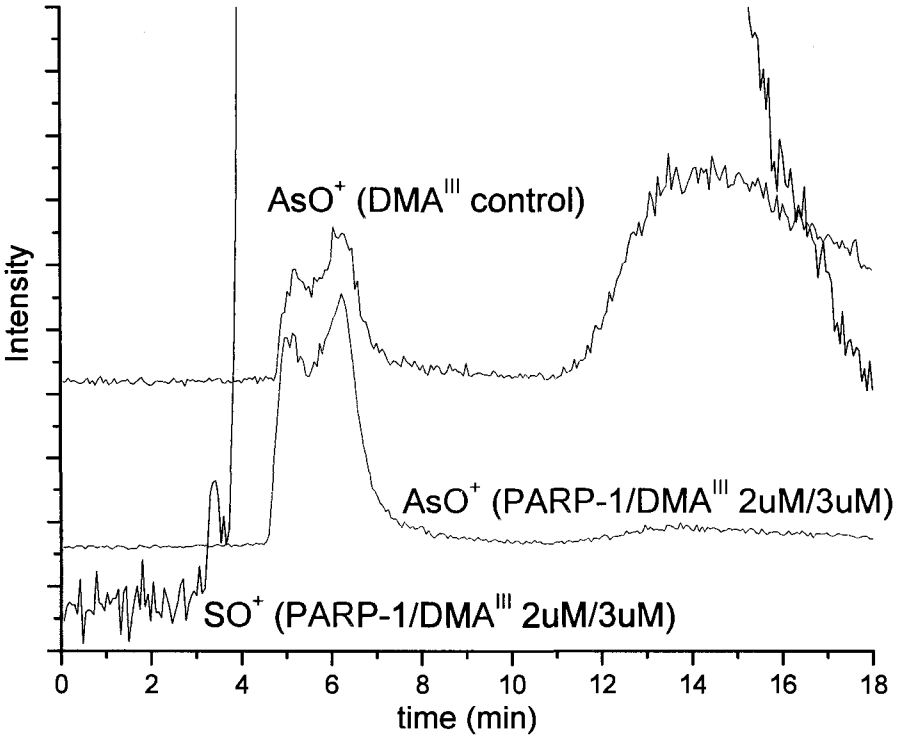


Figure 2.3 Chromatograms from the HPLC-ICPMS analyses of PARP-1/DMA^{III} incubation sample and DMA^{III} control sample. 3 μM DMA^{III} alone (blue trace) or 3 μM DMA^{III} and 2 μM PARP-1 were incubated in buffer (pH=2) for 2 hr (red and black trace). AsO^+ (m/z 91) was detected for As, and SO^+ (m/z 48) was detected for the present of protein

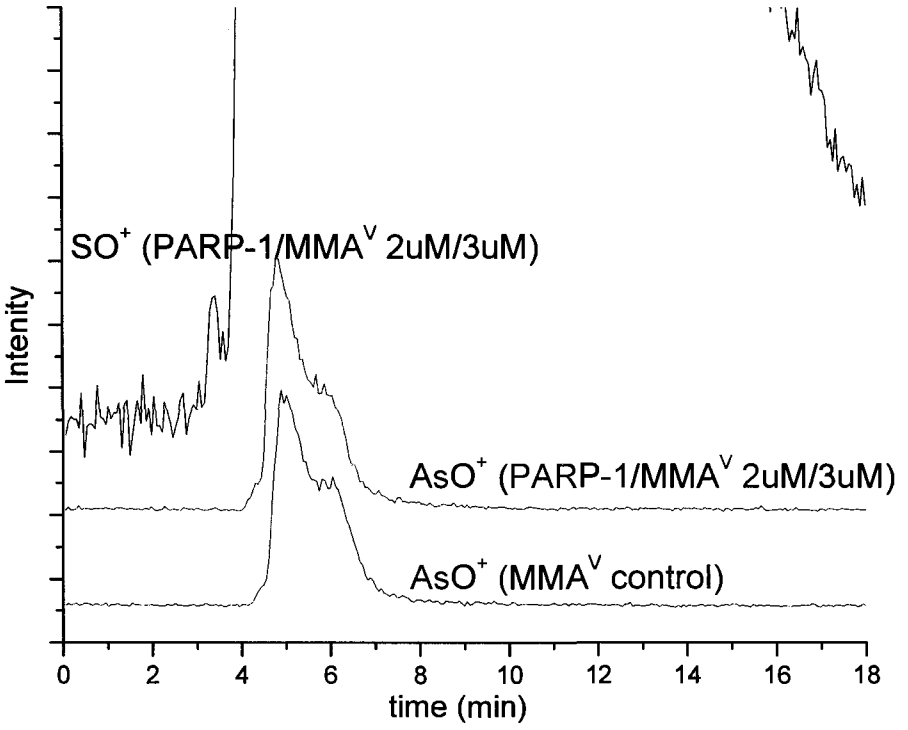


Figure 2.4 Chromatograms from the HPLC-ICPMS analyses of PARP-1/MMA^V incubation sample and MMA^V control sample. 3 μM MMA^V alone (blue trace) or 3 μM MMA^V and 2 μM PARP-1 were incubated in buffer (pH=2) for 2 hr (red and black trace). AsO^+ (m/z 91) was detected for As, and SO^+ (m/z 48) was detected for the present of protein

3.3 Arsenic binding to P53

P53 sample after dialysis was treated with 0.5mM TCEP, using the same procedures as for PARP-1. Then p53 (12 μM) was incubated separately with 3 μM PAO, MMA^{III}, DMA^{III}, and MMA^{V} at room temperature for 2 hour. Dialysis buffer was used as control with the same treatment as p53 protein solution. Both the protein incubation sample and the control sample were analyzed by HPLC-ICP-MS.

Figure 2.5 shows the chromatograms from the analyses of P53/PAO and P53/MMA^{III} incubation samples. In the SO^{+} detection channel, there is a small peak at 3-4 min, which corresponds to protein p53. The AsO^{+} detection shows a small peak at 3-4 min for both p53/PAO and p53/MMA^{III} incubation samples, which matches the peak in the chromatogram with detection of SO^{+}. Therefore, this small peak is considered as the arsenic from PAO or MMA^{III} bound to protein p53. The other peaks eluted after 4 minutes in the AsO^{+} detection channel are due to PAO or MMA^{III} and their oxidation products.

Besides PAO and MMA^{III}, 3 μM DMA^{III} and MMA^{V} were also incubated with dialyzed p53 sample. Figures 2.6 and 2.7 show the results of HPLC-ICPMS analyses of the arsenic compounds with or without the incubation with p53. In the AsO^{+} detection channel, there is no peak at 3-4 min in both p53/DMA^{III} and p53/MMA^{V} incubation samples. These results suggest that DMA^{III} and MMA^{V} can not bind to p53 under the given conditions.

There are 10 cysteines in P53 sequence (cys 124, 135, 141, 176, 182, 229, 238, 242, 275 and 277). All are in the DNA binding domain. Previous studies of P53 implicated cysteine residues in site-specific DNA binding via zinc coordination and redox regulation.[33, 34] As we discussed above, arsenic is difficult to substitute for zinc in zinc fingers due to the difference of electronegative property. Thus, cysteines in zinc finger may not be able to bind to arsenic. Other than the zinc finger cys 176, 238, 242, cys 124, 135, 141, 182, 229, 275, 277 are the possible sites for arsenic binding. It is possible that two

spatially proximal cysteines may bind to the same trivalent arsenic molecule simultaneously. From the three dimensional structure of p53 residues of domain 94-292, we find the spatial distance of cys 124, 135 and 141 is very close. The distance between cys135 and 141 is 3.49Å, the distance between cys 141 and 124 is 4.11Å, and the distance between cys135 and 124 is 5.44Å. Cys182 and 229 is on the surface of p53 protein individually. The distance between cys 275 and 277 is 7.13Å. Thus it is possible that two cysteines 135 and 141, or cysteines 141 and 124 may bind to one PAO molecule. Alternatively, cysteines that are far apart (cys 182, 229) may not bind to the same PAO molecule, but may each bind to a separate PAO molecule. Further data are needed to confirm this hypothesis.

Dose and time dependent experiments were carried out to study the P53/arsenic binding behavior. Dialyzed P53 (6 μM in 0.5 mM TCEP, pH=2) was incubated with PAO (3, 6, 18, 30 μM). These represent concentration ratio of p53/PAO ranging from 1:0.5 to 1:5. For time dependent experiment, P53/PAO (6μM/3μM) sample was incubated at different incubation times (2 hr, 6 hr, 12 hr and 24 hr). Figure 2.8 shows the dose dependent results, and Figure 2.9 shows the time dependent results. Both figures show that P53/PAO complex increases with the increase of PAO concentration and incubation time. Figure 2.10 summarizes the results of bound PAO in P53/PAO incubation sample of time and dose experiment.

Data from dose-dependent experiment are used to calculate the dissociation constant (K_d) based on the following equations.

$P + L \Leftrightarrow PL$ (where P is protein, L is ligand, and PL is complex)

$$K_d = \frac{[P][L]}{[PL]} \text{ and } [P] = [P]_0 - [PL],\ [L] = [L]_0 - [PL]$$

$$K_d = \frac{([P]_0 - [PL])[L]}{[PL]} = \frac{[P]_0[L]}{[PL]} - [L]\ ,$$

$$[PL] = \frac{[P]_0[L]}{K_d + [L]} \text{ , set } \upsilon = \frac{[PL]}{[P]_0} = \frac{[L]}{K_d + [L]}$$

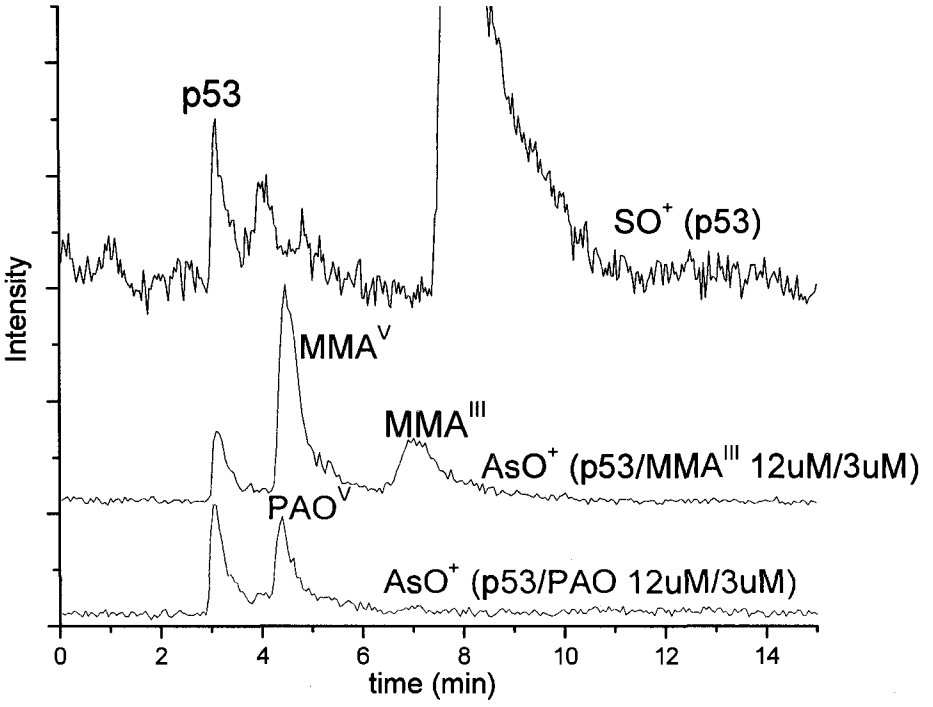


Figure 2.5 Chromatograms from the HPLC-ICPMS analyses of p53/PAO and p53/MMA^{III} incubation samples. 12 μM P53 (black trace) were incubated with 3 μM PAO (red trace) or 3 μM MMA^{III} (blue trace) in buffer (pH=2) for 2 hr. AsO^+ (m/z 91) was detected for As, and SO^+ (m/z 48) was detected for the present of protein

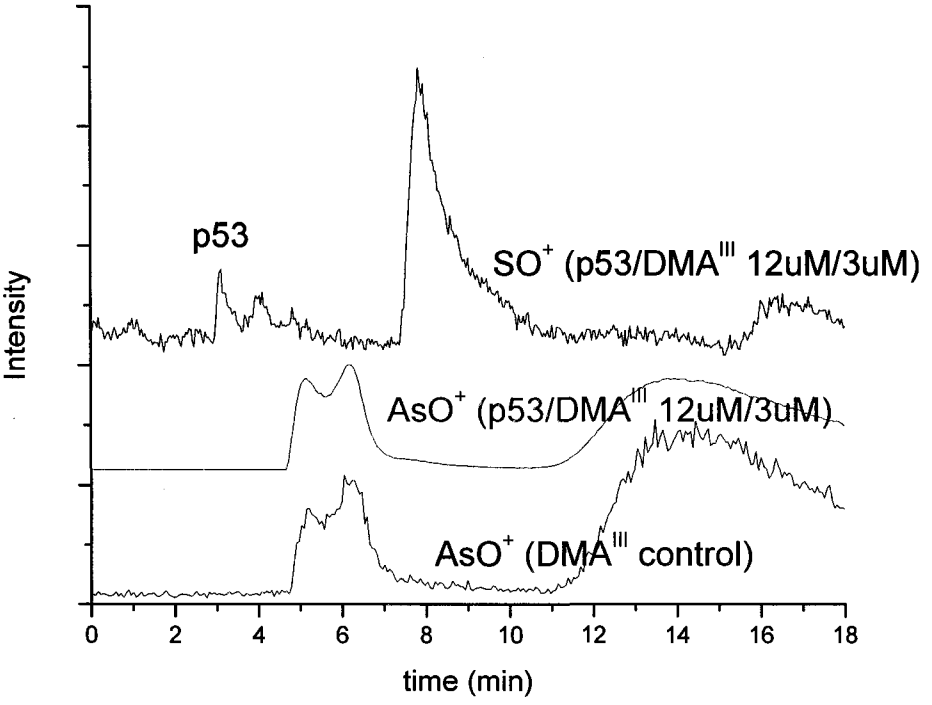


Figure 2.6 Chromatograms from the HPLC-ICPMS analyses of p53/DMA^{III} incubation sample and DMA^{III} control sample. 3 μM DMA^{III} alone (blue trace) or 3 μM DMA^{III} and 12 μM p53 were incubated in buffer (pH=2) for 2 hr (red and black trace). AsO^+ (m/z 91) was detected for As, and SO^+ (m/z 48) was detected for the present of protein

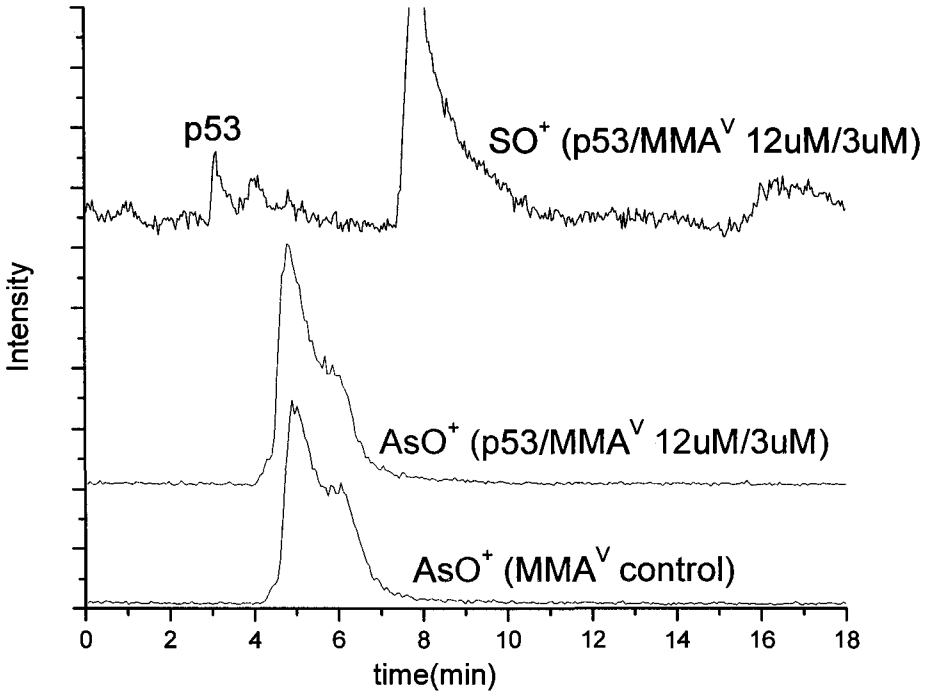


Figure 2.7 Chromatograms from the HPLC-ICPMS analyses of p53/MMA^V incubation sample and MMA^V control sample. 3 μM MMA^V alone (blue trace) or 3 μM MMA^V and 12 μM p53 were incubated in buffer (pH=2) for 2 hr (red and black trace). AsO^+ (m/z 91) was detected for As, and SO^+ (m/z 48) was detected for the present of protein

If the protein has n independent binding sites with the same affinity for ligand, then the average number of L bound per protein P becomes:

$$\upsilon = \frac{n[L]}{K_d + [L]}, \text{ and } \frac{1}{\upsilon} = K_d \frac{1}{[L]\, n} + \frac{1}{n}$$

From a linear plot of $\frac{1}{\upsilon}$ vs $\frac{1}{[L]}$, slope =K_d/n, Y-intercept=1/n. In above experiments, we detected the bound arsenic concentration as the concentration of complex (PL). Because the size of ligand (arsenic compound) is much smaller than the size of p53, and the amount of binding complex is much less than the amount of p53, we can assume that one arsenic molecule only can bind to one p53 molecule, and the concentration of complex is equal to the concentration of bound arsenic. Figure 2.11 shows the graph of $\frac{1}{\upsilon}$ vs $\frac{1}{[L]}$ using the data from the P53/PAO dose dependent experiment.

Figure 2.11 shows that with 12 hr and 24 hr incubations, the expected linear relations are achieved for binding mixtures containing 6 μM p53 and 3-30 μM PAO.
From the slope and intercept of these linear plots, we have obtained the K_d and n values (Table 2.1).

Data from Table 2.1 show that with the increase of incubation time from 12 to 24 hr, the number of cysteine in p53 bound to PAO increases from 1 to 2. It is possible for a PAO molecule to bind with two cysteines (cys 134 and 141 or cys 141 and 124). More data is needed to confirm the arsenic binding sites in P53. The affinity between p53 and PAO is moderate.

4. Conclusions

Trivalent arsenic compound PAO was confirmed to bind to PARP-1, but no evidence showed other trivalent arsenic compounds (MMA^{III}, DMA^{III}) bind to PARP-1. The arsenic binding sites in PARP-1 can not be estimated due to lack of the protein

conformation data. P53 can also bind to PAO and MMA^{III}, but not DMA^{III}. Time and dose dependent experiments showed that the amount of P53/PAO complex increases with the increase of PAO concentration and incubation time, but the binding affinity of p53/PAO complex is moderate. 3D conformation data of P53 suggest that there are four possible arsenic binding sites: (1) PAO binding to both cysteines 141 and 135; (2) PAO binding to both cysteines 141 and 124; (3) PAO binding to cysteine 182; and (4) PAO binding to cysteine 229. The number of PAO binding site in one p53 molecule calculated from dose dependent experiment is from 0.9 to 1.8, which means each PAO molecule may bind to one or two cysteines in p53.

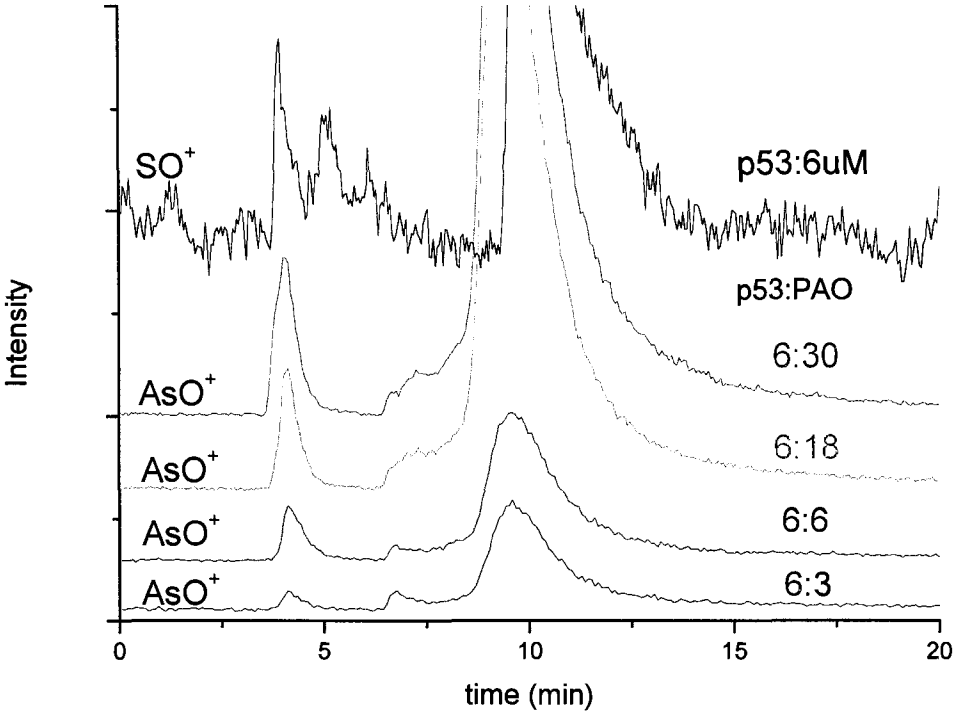


Figure 2.8 Chromatograms from the dose dependent HPLC-ICPMS analyses of incubation samples containing a constant concentration of p53 (6 μM) and increasing concentrations of PAO (3, 6, 18, 30 μM).

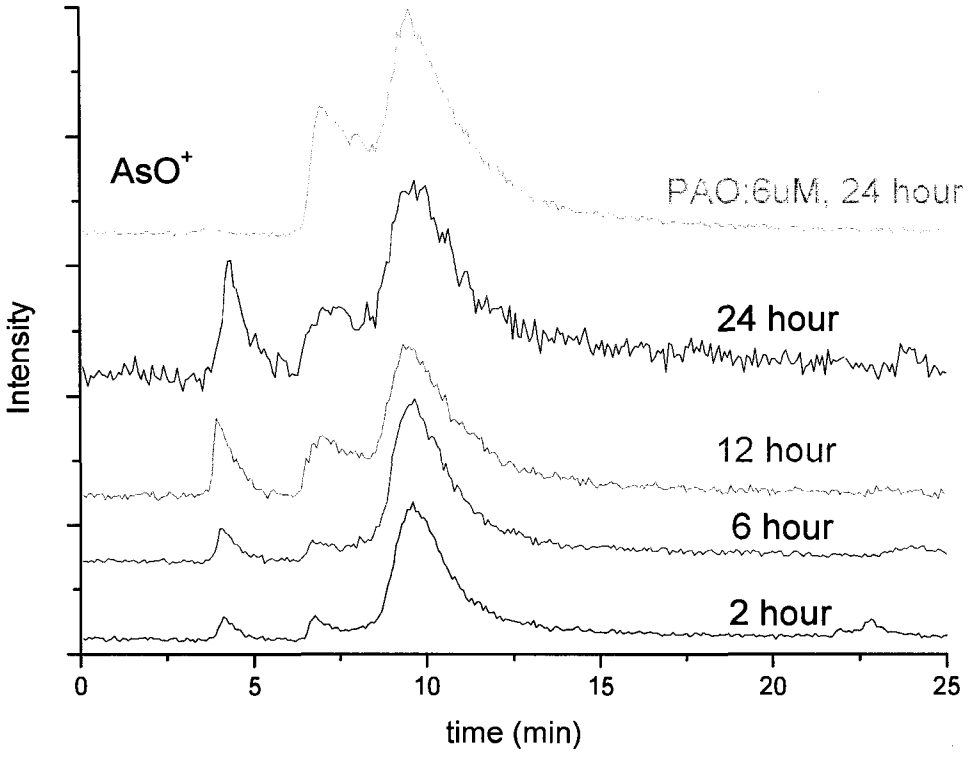


Figure 2.9 Chromatograms obtained from the time dependent HPLC-ICPMS analyses of a mixture containing 6 μM p53 and 3 μM PAO, incubated at room temperature for 2 hr, 6 hr, 12 hr, and 24 hr. AsO^+ (m/z 91) was detected for all the chromatograms

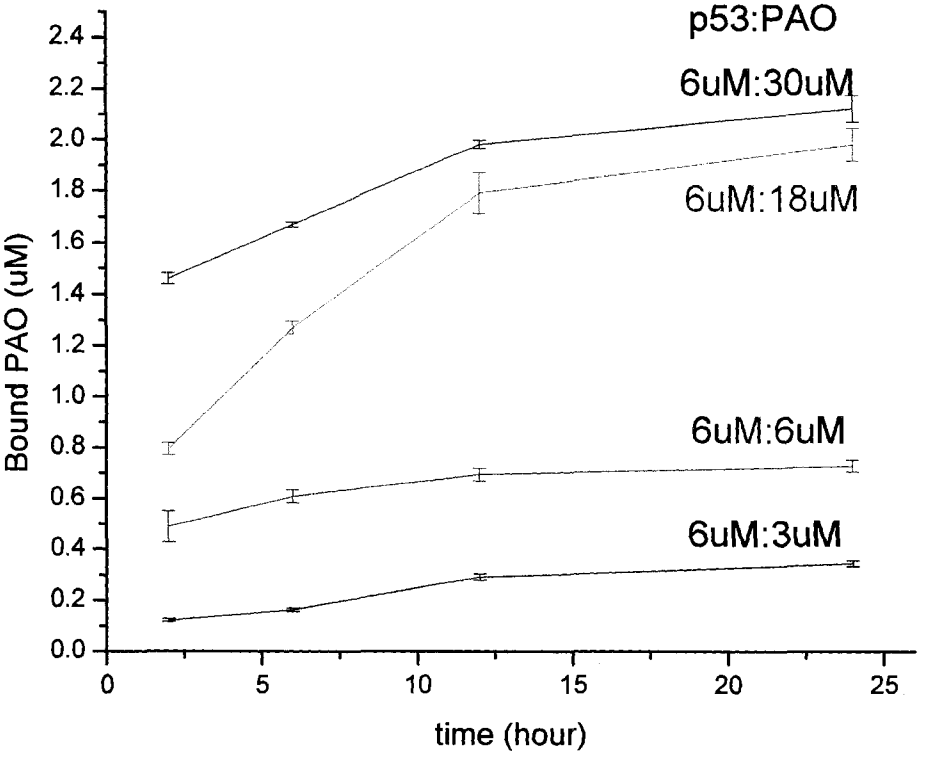


Figure 2.10 The amount of PAO bound to p53 in the various p53/PAO incubation samples incubated for 2-24 hours

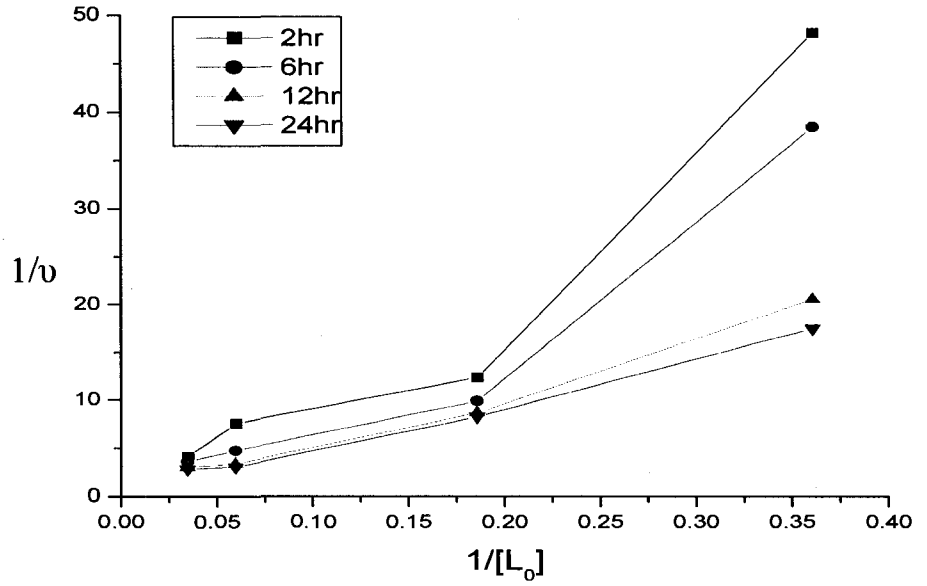


Figure 2.11 The graph of $\frac{1}{\upsilon}$ vs $\frac{1}{[L_0]}$ in P53/PAO dose dependent

Table 2.1 Dissociation constant (K_d) and number of cysteines in each p53 bound to PAO

Incubation time	K_d (uM)	n (No. of binding site)	R
12 hour	50.6	0.9	0.9897
24 hour	84.4	1.8	0.9952

5. References

1. Kitchin, K. T., Recent advances in arsenic carcinogenesis: modes of action, animal model systems, and methylated arsenic metabolites. *Toxicol Appl Pharmacol* **2001,** 172, (3), 249-61.
2. Rossman, T. G., Enhancement of UV-mutagenesis by low concentrations of arsenite in E. coli. *Mutat Res* **1981,** 91, (3), 207-11.
3. Hartmann, A.; Speit, G., Effect of arsenic and cadmium on the persistence of mutagen-induced DNA lesions in human cells. *Environ Mol Mutagen* **1996,** 27, (2), 98-104.
4. Hartwig, A., Current aspects in metal genotoxicity. *Biometals* **1995,** 3, 3-11.
5. Okui, T.; Fujiwara, Y., Inhibition of human excision DNA repair by inorganic arsenic and the co-mutagenic effect in V79 Chinese hamster cells. *Mutat Res* **1986,** 172, (1), 69-76.
6. Wang, T. C.; Huang, J. S.; Yang, V. C.; Lan, H. J.; Lin, C. J.; Jan, K. Y., Delay of the excision of UV light-induced DNA adducts is involved in the coclastogenicity of UV light plus arsenite. *Int J Radiat Biol* **1994,** 66, (4), 367-72.
7. Bau, D. T.; Gurr, J. R.; Jan, K. Y., Nitric oxide is involved in arsenite inhibition of pyrimidine dimer excision. *Carcinogenesis* **2001,** 22, (5), 709-16.
8. Gebel, T. W., Genotoxicity of arsenical compounds. *Int J Hyg Environ Health* **2001,** 203, (3), 249-62.
9. Hartwig, A.; Groblinghoff, U. D.; Beyersmann, D.; Natarajan, A. T.; Filon, R.; Mullenders, L. H., Interaction of arsenic(III) with nucleotide excision repair in UV-irradiated human fibroblasts. *Carcinogenesis* **1997,** 18, (2), 399-405.
10. Li, J. H.; Rossman, T. G., Inhibition of DNA ligase activity by arsenite: a possible mechanism of its comutagenesis. *Mol Toxicol* **1989,** 2, (1), 1-9.
11. Lee-Chen, S. F.; Yu, C. T.; Jan, K. Y., Effect of arsenite on the DNA repair of UV-irradiated Chinese hamster ovary cells. *Mutagenesis* **1992,** 7, (1), 51-5.
12. Schwerdtle, T.; Walter, I.; Hartwig, A., Arsenite and its biomethylated metabolites interfere with the formation and repair of sTable BPDE-induced DNA adducts in human cells and

impair XPAzf and Fpg. *DNA Repair (Amst)* **2003,** 2, (12), 1449-63.

13. Le Rhun, Y.; Kirkland, J. B.; Shah, G. M., Cellular responses to DNA damage in the absence of Poly(ADP-ribose) polymerase. *Biochem Biophys Res Commun* **1998,** 245, (1), 1-10.
14. Berg, N. A., Poly(ADP-ribose) in the cellular reponse to DNA damage. *Radiat.Res.* **1985,** 101, 4-15.
15. Boulikas, T., Poly(ADP-ribosyl)ation, DNA strand breaks, chromatin and cancer. *Toxicol Lett* **1993,** 67, (1-3), 129-50.
16. De Murcia, G.; Shall, S., *From DNA Damage and Stress Signalling to Cell Death.* Oxford University Press, Oxford, UK: 2000.
17. Ikejima, M.; Noguchi, S.; Yamashita, R.; Ogura, T.; Sugimura, T.; Gill, D. M.; Miwa, M., The zinc fingers of human poly(ADP-ribose) polymerase are differentially required for the recognition of DNA breaks and nicks and the consequent enzyme activation. Other structures recognize intact DNA. *J Biol Chem* **1990,** 265, (35), 21907-13.
18. Gradwohl, G.; Menissier de Murcia, J. M.; Molinete, M.; Simonin, F.; Koken, M.; Hoeijmakers, J. H.; De Murcia, G., The second zinc-finger domain of poly(ADP-ribose) polymerase determines specificity for single-stranded breaks in DNA. *Proc Natl Acad Sci U S A* **1990,** 87, (8), 2990-4.
19. Desjarlais, J. R.; Berg, J. M., Toward rules relating zinc finger protein sequences and DNA binding site preferences. *Proc Natl Acad Sci U S A* **1992,** 89, (16), 7345-9.
20. Hartwig, A.; Pelzer, A.; Asmuss, M.; Burkle, A., Very low concentrations of arsenite suppress poly(ADP-ribosyl)ation in mammalian cells. *Int J Cancer* **2003,** 104, (1), 1-6.
21. Yager, J. W.; Wiencke, J. K., Inhibition of poly(ADP-ribose) polymerase by arsenite. *Mutat Res* **1997,** 386, (3), 345-51.
22. Walter, I.; Schwerdtle, T.; Thuy, C.; Parsons, J. L.; Dianov, G. L.; Hartwig, A., Impact of arsenite and its methylated metabolites on PARP-1 activity, PARP-1 gene expression and poly(ADP-ribosyl)ation in cultured human cells. *DNA Repair (Amst)* **2007,** 6, (1), 61-70.
23. Knowles, F. C.; Benson, A, A.,The biochemistry of arsenic. *Trends Biochem. Sci.* **1983,** 8,

178.

24. Cho, Y.; Gorina, S.; Jeffrey, P. D.; Pavletich, N. P., Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science* **1994,** 265, (5170), 346-55.

25. Gorina, S.; Pavletich, N. P., Structure of the p53 tumor suppressor bound to the ankyrin and SH3 domains of 53BP2. *Science* **1996,** 274, (5289), 1001-5.

26. Rainwater, R.; Parks, D.; Anderson, M. E.; Tegtmeyer, P.; Mann, K., Role of cysteine residues in regulation of p53 function. *Mol Cell Biol* **1995,** 15, (7), 3892-903.

27. Delphin, C.; Cahen, P.; Lawrence, J. J.; Baudier, J., Characterization of baculovirus recombinant wild-type p53. Dimerization of p53 is required for high-affinity DNA binding and cysteine oxidation inhibits p53 DNA binding. *Eur J Biochem* **1994,** 223, (2), 683-92.

28. Meplan, C.; Richard, M. J.; Hainaut, P., Redox signalling and transition metals in the control of the p53 pathway. *Biochem Pharmacol* **2000,** 59, (1), 25-33.

29. Bandura, D. R.; Baranov, V. I.; Tanner, S. D., Detection of ultratrace phosphorus and sulfur by quadrupole ICPMS with dynamic reaction cell. *Anal Chem* **2002,** 74, (7), 1497-502.

30. Hann, S.; Koellensperger, G.; Obinger, C.; Furtmüller, P. G., SEC-ICP-DRCMS and SEC-ICP-SFMS for determination of metal–sulfur ratios in metalloproteins. *J. Anal. At. Spectrom.* **2004,** 19, (1), 74-79.

31. Lu, M.; Wang, H.; Li, X.; Le, X. C., Evidence of Hemoglobin binding to arsenic as a basis for the accumulation of arsenic in ret blood *Chem. Res. Toxicol.* **2004,** 17 1733-1742.

32. Cline, D. J.; Thorpe, C.; Schneider, J. P., Effects of As(III) binding on alpha-helical structure. *J Am Chem Soc* **2003,** 125, (10), 2923-9.

33. Hainaut, P.; Milner, J., Redox modulation of p53 conformation and sequence-specific DNA binding in vitro. *Cancer Res* **1993,** 53, (19), 4469-73.

34. Hupp, T. R.; Meek, D. W.; Midgley, C. A.; Lane, D. P., Activation of the cryptic DNA binding function of mutant forms of p53. *Nucleic Acids Res* **1993,** 21, (14), 3167-74.

Chapter 3

Development of an affinity technique to capture arsenic-binding proteins

1. Introduction

Arsenic is a ubiquitous contaminant in the environment because of its natural occurrence. Epidemiologic studies have shown that exposure to high levels of arsenic from drinking water could cause a wide range of health effects[1-3], most seriously, the cancers of the bladder, lung, skin, and kidney.[4-6] But the bio-chemical mechanisms of these effects caused by arsenic are still not clear[7]. A possible mechanism of arsenic action could be through the binding of trivalent arsenic to thiol groups in proteins, thereby changing the conformation of proteins and inhibiting protein functions. If the affected proteins are responsible for cellular repair of DNA damage, the inhibition of these proteins could lead to carcinogenesis.

Many reports have already shown evidence of trivalent arsenic binding to cysteines in proteins. It has been proposed[8] that inorganic arsenite can bind to three cysteines with three thiol groups arranged in 3-6Å by a tertiary structure of protein. The arsenic binding peptide model Cys-$(X)_n$-Cys, where X represents any amino acid and n is the number of amino acids, has been studied extensively for arsenic binding [7, 9-11].

Hemoglobin[12], metallothionein[13, 14], galectin-1 and thioredoxin peroxidase II[15, 16], ArsR protein[17], GLUT4[18], glutathione[19-21], tubulin and actin[22] have been demonstrated to bind to trivalent arsenic species, including inorganic arsenite and its methylation metabolites monomethyarsonous acid (MMA^{III}) and dimethylarsinous (DMA^{III}). It is possible that a large variety of potential arsenic binding proteins may exist with or without the Cys-$(X)_n$-Cys motif in the proteins. It is necessary to identify the potential arsenic binding proteins and to study how arsenic binding could affect activities of these

proteins.

Affinity techniques [16-18, 22, 23] have been used to select arsenic binding proteins from treated cells, and the affinity-selected proteins were subsequently identified by molecular weight from SDS PAGE and Western blot. Only a few proteins (galectin-1, thioredoxin peroxidase, GLUT4, tubulin and actin) have been identified by using this approach.

With the development of electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI), mass spectrometry has become a powerful method of large-scale proteome research. Its femtomole sensitivity, high-quality MS and MS/MS data, rapid data collection make it effective for protein identification. To identify arsenic binding proteins, we describe here a method that combines mass spectrometry and an improved arsenic-affinity selection technique. With the improvement of the arsenic immobilization efficiency from 14 umol/g to 27 umol/g wet resin, many of arsenic binding proteins in subcellular fraction of cell lysate can be identified by combination of affinity technique and mass spectrometry. The sequence and conformation of identified proteins were examined through Swissprot.

2. Experimental

2.1 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), formeldelye 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 were from Fisher Scientific Canada (Ottawa,

ON). ProteoExtract® subcellular proteome extraction kit was purchased from CALBIOCHEM (San Diego, CA, USA). RC DC protein assay was from Bio-Rad (Mississauga, ON). 4-aminophenylarsine oxide ($NPAO^{III}$) was provided by Dr. W. R. Cullen of University of British Columbia, Canada.

2.2 Affinity column preparation

Two kinds of arsenic affinity media were prepared by reaction of Eupergit® C beads with 4-aminophenylarsine oxide ($NPAO^{III}$) or arsenite (As^{III}). To prepare NPAO affinity media, arsenic compound 4-aminophenylarsine oxide (0.1 g) was dissolved in 1 mL DMSO, pH=2 by HCl, then mixed with 3 mL deionized water and adjusted to pH 4 by NaOH solution (pH measured by pH indicator). The solution was poured into gravity column containing 0.5 gram reactive Eupergit® C beads, oxygen in the mixture was removed by blowing nitrogen for 10 minutes , then the gravity column was sealed, and shaken slowly at room temperature for 24 hours. Arsenite affinity media was prepared by $NaAsO_2$ 0.07 g in 4 mL water, pH 7.5, following the NPAO affinity media preparation step. Both affinity media were washed by 500 mL 100 mM NaCl solution, 500 mL deionized water, 200 mL 50% acetonitrile solution and 500mL deionized water. The column was washed until no arsenic signal could be detected in the wash-off solution by ICP-MS (PE SCIEX, ELAN DRC Plus).

The amount of 4-aminophenylarsine oxide and arsenic oxide immobilized were determined by the amount of BAL reacted with arsine oxide group. Free BAL can reduce DTNB immediately and the reduced product yellow thionitrobenzoate ion can be detected at 412 nm by spectrometry [23] (Bio-Rad Smartspec™ 3000).

Benzylamine instead of 4-aminophenylarsine oxide was reacted with reactive Eupergit® C to prepare non-specific control media without arsenic compound.

2.3 Specific test

The specificity of the affinity column was tested by protein containing free cysteine or no free cysteine. Protein transferrin (all cysteine in dithiol bond) and cysteine blocked HSA were used as negative controls to apply to arsenic affinity media, membrane protein fraction of A549 cell was also applied to the non specific control column to test the hydrophobic interaction between the media and proteins. As negative controls, HSA was treated with carbamidomethylation to block the free cysteines in HSA. Human serum albumin and bovine serum albumin with free cysteine were used as positive control and applied to arsenic affinity column. Carbonic anhydrase II (lack of cysteine in molecule) was used to optimize the affinity technique parameter.

Arsenic binding to positive controls was confirmed by HPLC-ICP-MS. Protein in 50 mM NH_4HCO_3 buffer (pH 7.5) was incubated with $NPAO^{III}$ at room temperature for two hours. Protein/$NPAO^{III}$ incubation sample was injected into size exclusion column (Phenomenex Biosep-SEC-S 2000, 300×4.6 mm column or Agilent ZOBAX GF-250 250×4.6 mm column), eluted by 10 mM ammonia acetate to separate protein from unbound NPAO. In the Protein/$NPAO^{III}$ incubation sample, protein will be eluted earlier than unbound arsenic compounds. The $NPAO^{III}$ bound to protein will have the same retention time as protein, which is separated from the unbound small arsenic compounds. ICP-MS (Perkin Elmer Sciex, ELAN DRC, Plus) was used as detector in oxygen DRC mode. Sulfur oxide (SO^+, Mw 47.967) was used to detect protein because protein contains cysteine amino acid, and arsenic oxide (AsO^+, Mw 90.9165) was used to detect NPAO.

2.4 Subcellular fractionation of A549 cells

A549 cells were used as a model system for application. Subcellular extraction from A549 cells was carried out by following the procedure in the instruction of ProteoExtract® Subcellular Proteome Extraction kit. Total four fractions were extracted, including cytosol fraction, membrane/organelle fraction, nuclear fraction, and

cytoskeleton fraction.

2.5 Selection of arsenic-binding protein by affinity column

Protein solution (2 mL) was poured into affinity media (2 mL), and incubated at room temperature for two hours with slow shaking. The unbound proteins were washed out by 10 mL buffer A (20 mM Hepes, 50 mM NaCl, pH 7.5), then the non-specific binding proteins were washed out by 5 mL buffer B (1% SDS in buffer A) or 5 mL buffer C (1% Triton X-100 in 20 mM Hepes, 50 mM NaCl, pH 7.5) per round for 6 rounds. The specific binding proteins were eluted twice by 2 mL 10 mM DTT in Buffer A , followed by 2 mL 100 mM DTT in buffer B. Each washing and elution step was conducted for 10 minutes with slow shaking of the column. All the fractions were collected and quantitated by Bio-Rad RC DC protein assay.

2.6 SDS-polyacrylamide gel electrophoresis and protein in-gel digestion

All fractions collected from affinity column were dried by freeze-drying. The freeze-dried samples were dissolved in deionized water, and loaded in SDS-PAGE. Unbound protein fraction was dissolved in 1 mL deionized water; each 1% SDS buffer wash fraction was dissolved in 400 uL deionized water; each DTT elution fraction was dissolved in 200 uL deionized water. SDS-PAGE was carried out in a Bio-Rad mini protein III system in 4%/10% stacking/separation polyacryamide mini gel, followed by silver staining.

Protein bands were excised from the gel and digested with trypsin. Briefly, gel pieces were cut into about 1×1 mm pieces and washed 3 times in deionized water for 15 minutes. After reduction and alkylation, gel pieces was dehydrated and re-suspended in 0.02 μg/μL trypsin, 1 mM $CaCl_2$, and 50 mM NH_4HCO_3 buffer for overnight digestion at 37°C. Peptide was extracted three times by 60% acetonitrile and 0.1% TFA solution, one time in 80% acetonitrile and 0.1% TFA solution for 20 minutes of shaking each time.

2.7 Protein purification and in-solution digestion

Arsenic specific binding proteins in the concentrated elution fractions were precipitated by 20% TCA at 4℃ overnight, then centrifuged at 15,000×g for 30 minutes at 4℃, and washed with ice cold acetone. The precipitate was dissolved in 100μL 0.5% SDS 100mM NH_4HCO_3 solution, followed by the standard reduction and alkylation procedure. Then the proteins were precipitated again by 8 volumes of acetone (-20°C) overnight and centrifuged for 30 minutes in 4°C at 15,000×g, then washed by ice cold acetone. Protein precipitate was re-dissolved in 50μL 100mM NH_4HCO_3 (pH=8.5) solution with vortexing. Finally, 3 μL of 1 μg/μL trypsin solution and 2.5 μL 20 mM $CaCl_2$ was added for digestion overnight at 37°C. The digestion reaction was stopped by adjusting the pH to 3.

2.8 SDS-assisted digestion and peptide fractionation by SCX column

Proteins which could not be re-dissolved in NH_4HCO_3 solution after alkylation were dissolved in 7 μL 1% SDS 100 mM NH_4HCO_3 solution. Then the solution was diluted to 100 μL by 100 mM NH_4HCO_3(pH=8.5) solution, 3 μL 1 μg/μL trypsin solution and 5 μL 20 mM $CaCl_2$ were added for digestion overnight at 37°C. Digestion was stopped by adjusting the pH to 3. SDS in the digestion solution was removed, and peptides were fractionated by 60-minute gradient elution (mobile phase A: 20% acetonitrile, 0.1% TFA; mobile phase B: 20% acetonitrile, 0.1% TFA, 1 M NaCl) by SCX column (100x2.1 mm 5 μm BioBasic SCX Column from Thermo Scientific). The gradient was 100%A for 15 minute, 0-30% B for 3 minutes, 30-50% B for 4 minutes, 50% B for 38 minutes, flow rate was 0.2 mL/min. Peptide fractions were collected according to UV signal (214 nm).

2.9 Analysis of proteins and peptides

In-gel digestion samples were analyzed by nanospray-ESI-MS and nanospray-ESI-MS/MS. Peptides from in-gel digestion solution were desalted by Zip-tip

(Millipore) following the instruction procedure. 50% methanol and 0.2% formic acid was used in nanospray-ESI experiment. Proteins were identified from peptide fingerprint and MS/MS data generated by mass spectrometer (API QSTAR Pulsar-I, AB, PE SCIEX).

LC-ESI-MS/MS was used to analyze in-solution digestion samples. 1 μL peptide solution was injected and separated by reversed-phase chromatography on a 0.3×150 mm Agilent C18 column at a flow-rate 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). In 200 minutes gradient elution, the gradient was 5%B in 10 minute for desalting, 5%-25% B in 30 minutes, 25%-60% B in 100 minutes, 60%-95% B in 60 minutes.

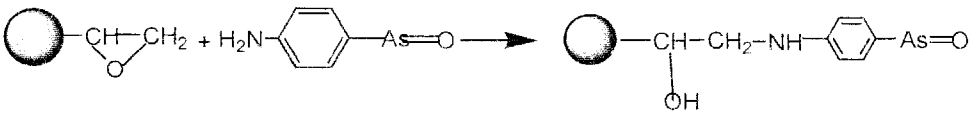
2.10 Database search

Mascot search program (http://www.matrixscience.com/cgi/search_form) was used for all database searches. For peptide fingerprint, peptide mass tolerance was 100ppm; for CID spectra, the precision tolerance is 0.3 Da for both parent peptide and fragment ions search. Trypsin was set as proteolytic enzyme, carbamidomethylation of cysteine was set as fixed modification, and methionine oxidation was set as variable modification, 1 missed cleavage site per peptide was allowed. An automatic database search followed by manual inspection was applied if more than one potential matches were reported for one spectrum, only the peptide hit with the highest score was examined. The rules used to inspect the peptide spectral identification was reported by Chen et al.[24] Protein was considered identified only when at least two peptides were found to identify the same protein. Possible cysteine position in sequence and protein conformation were also examined from Swissprot.

3 Results

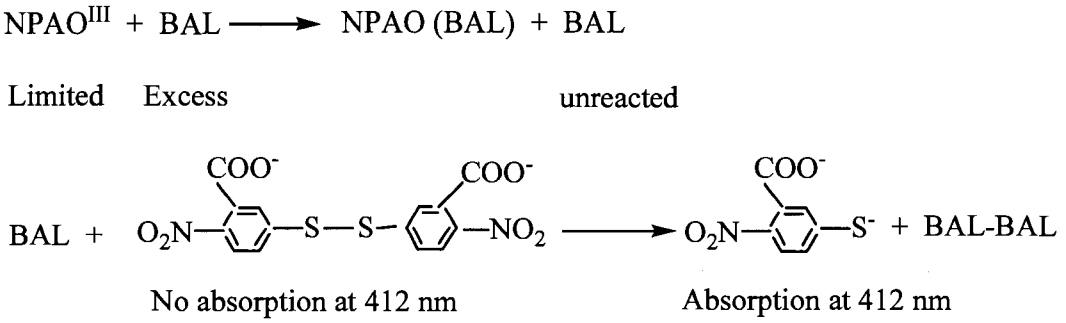
3.1 Preparation of arsenic affinity column and characterization of immobilization efficiency

Our strategy of preparing an affinity column for capturing arsenic-binding proteins was to use trivalent arsenicals as affinity ligands to bind with free thiols in target proteins, because it is known that trivalent arsenic species have high affinity for thiols. Several arsenic compounds, such as arsenite, dimethylarsinous acid (DMA^{III}), monomethylarsonous acid (MMA^{III}), and phenylarsine oxide (PAO^{III}), could serve as candidate affinity ligands. Previous work has shown that among the four trivalent arsenicals tested [12]; PAO^{III} is relatively stable and has the highest affinity to 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 of arsenic-binding proteins. The presence of a reactive primary amine group in $NPAO^{III}$ makes it readily immobilized onto a solid chromatographic support (resin) by the reaction of $NPAO^{III}$ with an epoxy group on the resin. Because the amine group and arsenic group are at the *para* positions, the arsenic group is kept exposed on the resin after amine reacts with the epoxide. (Scheme 3.1)



Scheme 3.1 The reaction involved for the immobilization of NPAO to a solid support

We determined the amount of immobilized arsenic by measuring the absorbance of thionitrobenzoate ion (λ=412 nm, ε=14,125), a reduction product of DTNB by BAL (scheme 3.2). The inhibition of BAL to reduce DTNB is a measure of arsenic because of the reaction of trivalent arsenicals with BAL. We found that the average amount of $NPAO^{III}$ immobilized on the resin was 27.3$\pm$2.3 μmole/g wet resin, measured from eight replicate preparations (Table 3.1).



Scheme 3.2 The reaction of measuring the amount of immobilized $NPAO^{III}$ on the affinity column

The reactant ratio, pH, and reaction time were optimized for the immobilization of NPAOIII. In this experiment, pH was fixed at 4 because NPAOIII could only be dissolved at pH less than 4. Reactant ratio was fixed at 1:1.5 (moles of epoxy group/moles of NPAOIII), which means epoxy group is the limiting reagent in the reaction. Figure 3.1 shows the amount of NPAOIII immobilized on polymer beads at different reaction times between 3 to 24 hours. The longer the reaction time, the more $NPAO^{III}$ immobilized on the polymer beads. 24 hour reaction time was chosen for immobilization of $NPAO^{III}$. After the affinity resin was stored at 4°C for 45 days, the amount of the bound $NPAO^{III}$ was decreased to 50%, probably due to the oxidation of $NPAO^{III}$ to $NPAO^{V}$. These results are shown in Figure 3.1.

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 nmole/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 $NPAO^{III}$ for the epoxy group on the resin. In addition, the ether bond between epoxy group and arsenic is easy to be hydrolyzed. Thus, $NPAO^{III}$ is more suitable than As^{III} for preparing an affinity column.

Table 3.1 Calibrated concentration of immobilized arsine oxide group

Times	1	2	3	4	5	6	7	8	AVE
Concentration (umole/g wet resin)	25.3	27.6	31.6	29.0	27.3	23.8	27.7	26.3	27.3±2.3

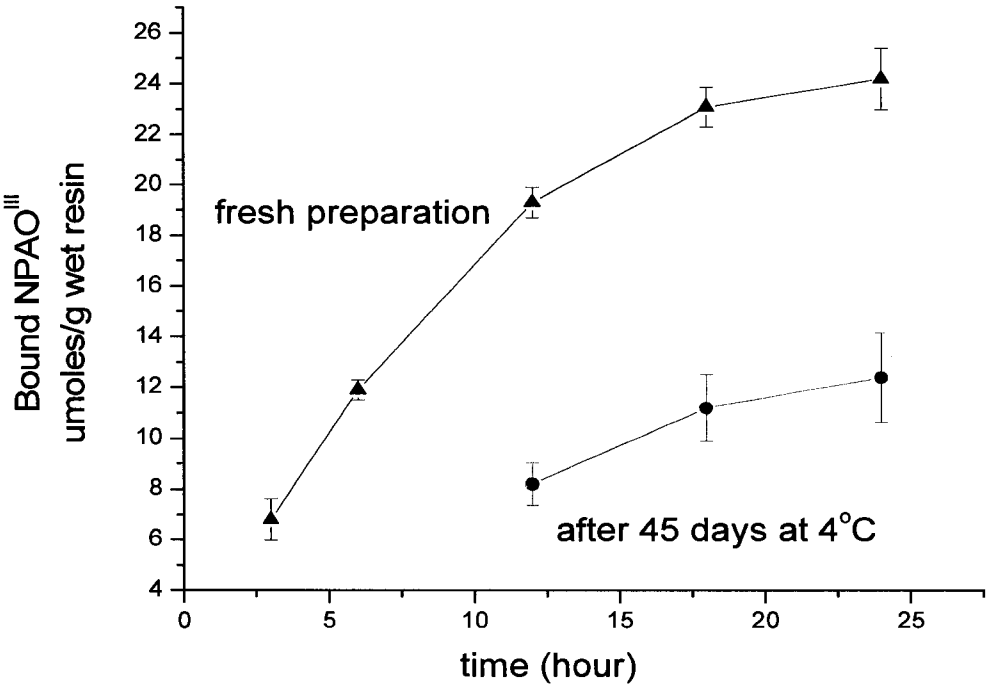


Figure 3.1 The amount of immobilized $NPAO^{III}$ at different reaction times (black trace), and the amount of immobilized $NPAO^{III}$ remaining after stored for 45 days at 4°C (red trace)

3.2 Optimization of washing conditions to remove non-specific binding from the affinity column

To optimize the affinity technique parameter, carbonic anhydrase II (no cysteine) was used as a control for removing non-specific binding on affinity column. Due to lack of cysteine in protein, carbonic anhydrase II should not be captured by affinity column. Figure 3.2 shows the electrophoresis image results of carbonic anhydrase II washed and eluted from affinity column. In figure 3.2(a), non-specific binding was washed off by buffer C (1% Triton X-100 in 20 mM Hepes, 50 mM NaCl, pH 7.5) for 5 rounds (lane 2 to 6), then eluted by 10 mM DTT (lane 7) in buffer A (20 mM Hepes, 50 mM NaCl, pH 7.5); a strong protein band of carbonic anhydrase II was detected in DTT elution (lane 7), which means buffer C could not wash off all non-specific binding of carbonic anhydrase II after 5 rounds of washes. But in figure 2.2(b), non-specific binding was washed off by buffer B (1% SDS in 20 mM Hepes, 50 mM NaCl, pH 7.5) for 5 rounds (lane 2 to 6), and eluted by 10 mM DTT (lane 7) in buffer A. No protein band at the same molecular weight as carbonic anhydrase II was detected in DTT elution (lane 7), which means that non-specific binding of carbonic anhydrase II had been washed off by buffer B after 5 rounds of washes;

To identify the DTT elution protein band (lane 7) shown in figure 3.2(b), this band was cut off for tryptic digestion and nano-ESI-MS/MS analysis. Tandem mass spectra of peptides (figure 3.3) from the digested protein, including m/z 606 for LQDVTDDHIR and m/z 586 for LPSEGPQPAHVVVGDVR, provided the identification of the protein as bovine biliverdin reductase B (MWt 22 kDa), an impurity present in the carbonic anhydrase II sample. Bovine biliverdin reductase B and carbonic anhydrase II are both located in cytoplasm. It is reasonable that carbonic anhydrase II sample contains 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. This cysteine in bovine biliverdin reductase B is the

possible site to bind with $NPAO^{III}$, resulting in the capture of bovine biliverdin reductase B on the affinity column. 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.

Non-denature condition (1% Triton X-100 in buffer) can not wash off all non specific binding proteins on the affinity column. In the denature wash condition (1% SDS in buffer), non specifically bound proteins on the affinity column can be removed efficiently. To optimize the washing condition, 0.1% and 0.5% SDS in buffer as washing solution was applied to the affinity column. Figure 3.4 shows the gel electrophoresis image results of carbonic anhydrase II washed by 0.5% SDS buffer from affinity column. A carbonic anhydrase II protein band was detected in DTT elution steps. Thus, 1%SDS buffer is the best washing condition to remove non specific binding proteins from the affinity column.

3.3 Specificity of the arsenic affinity column

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 (a thiol group available), and HSA after its cysteine has been blocked by carbamidomethylation. The former is expected to be captured on (retained by) the arsenic affinity column, while the latter is not expected to be captured.

Figure 3.5 shows results from gel electrophoresis analyses of HSA (lane 1), the unretained fraction (lane 2), washing solutions (with buffer containing 1% SDS) from the column after HSA was passed through the column (lanes 3-7), and the elution fraction containing the captured HSA (lanes 8-10). As shown in Figure 3.5(a), the non-specific binding can be washed off with a buffer containing 1% SDS (lanes 3-7). The elution fractions 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 show that the arsenic affinity

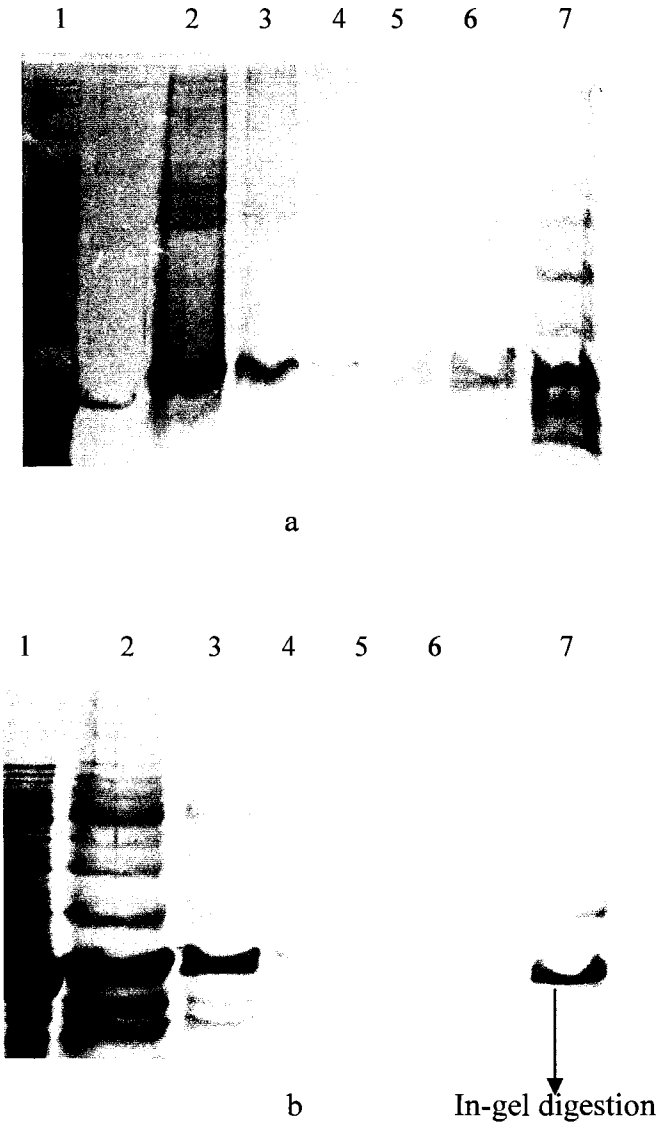


Figure 3.2 The electrophoresis image of carbonic anhydrase II washed and eluted from affinity resin. (a). lane 1: carbonic anhydrase II sample without passing through the affinity column; lane 2 to 6: five rounds of washing non-specific binding by buffer C (1% Triton X-100 in 20 mM Hepes, 50 mM NaCl, pH 7.5); lane 7: elution by 10 mM DTT in buffer A (20 mM Hepes, 50 mM NaCl, pH 7.5). (b). lane 1: carbonic anhydrase II sample without passing through the affinity column; lane 2 to 6: five rounds of washing non-specific binding by buffer B (1% SDS in 20 mM Hepes, 50 mM NaCl, pH 7.5); lane 7: elution by 10 mM DTT in buffer A.

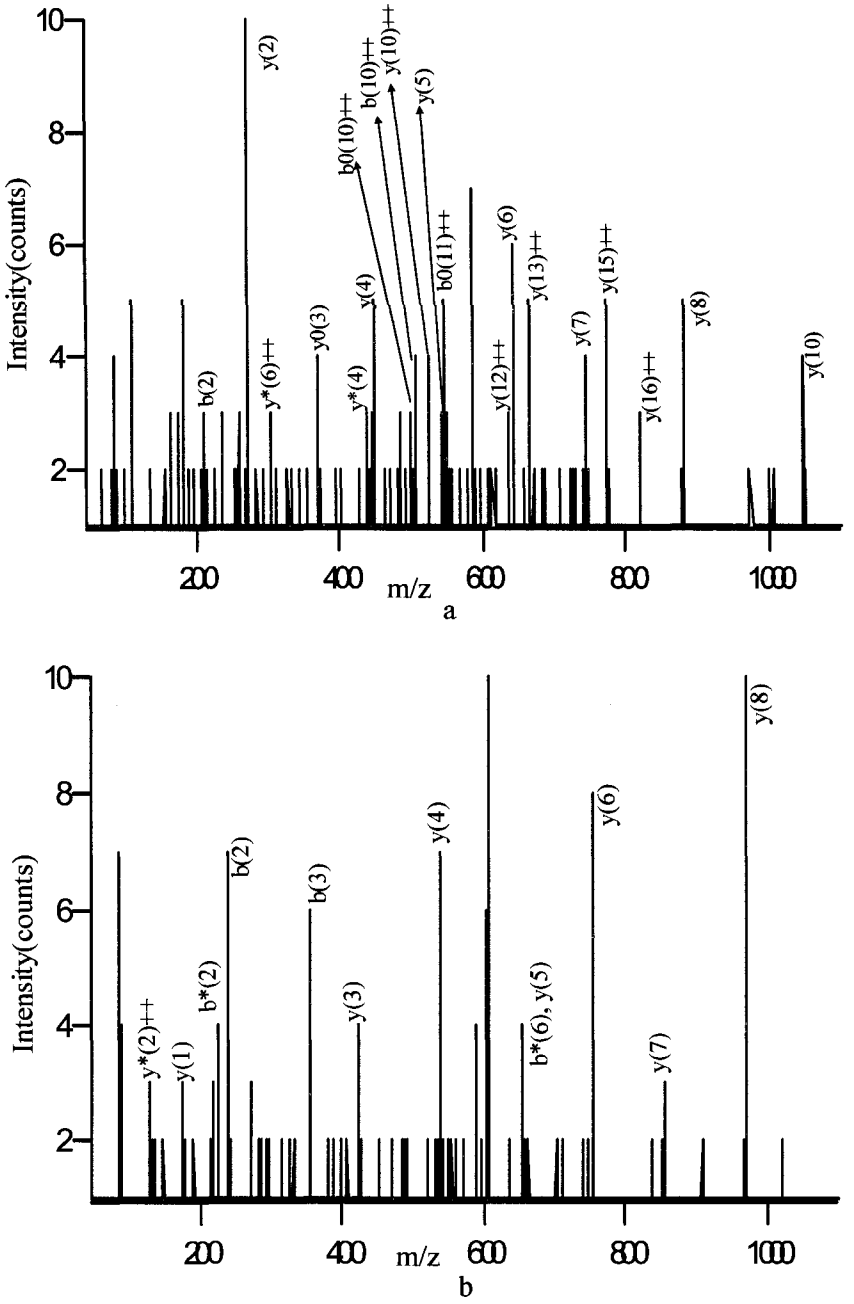


Figure 3.3 the MS/MS spectrum of peptide (a) m/z 586, identify as LPSEGPQPAHVVVGDVR, (b) m/z 606, identified as LQDVTDDHIR; both peptides belong to bovine biliverdin reductase B (Bovine BLVRB), an impurity present in the carbonic anhydrase sample.

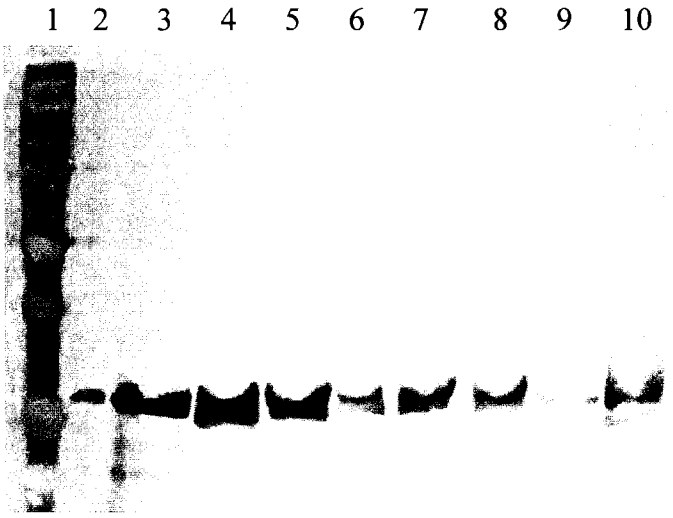


Figure 3.4 Electrophoresis image of carbonic anhydrase II washed and eluted from affinity column. Lane 1: ladder; lane 2: 3 ng carbonic anhydrase II sample without passing through the affinity column; lane 3 to 8: six rounds of washing non specific binding by 0.5%SDS in buffer A; lane 9: elution by 10 mM DDT in buffer A; lane 10: elution by 50 mM DTT in buffer B

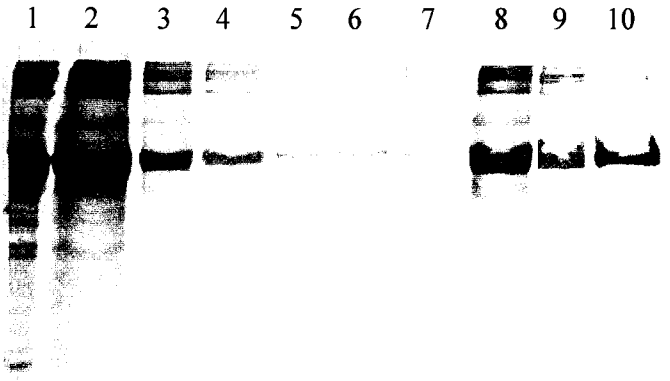
column is able to capture the unmodified HSA (containing a free cysteine) and that the captured HSA can be subsequently eluted for analysis.

In contrast, once the thiol in the cysteine is blocked by carbamidomethylation, the modified HSA cannot be captured by the arsenic affinity column (figure 3.5 (b)). 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.

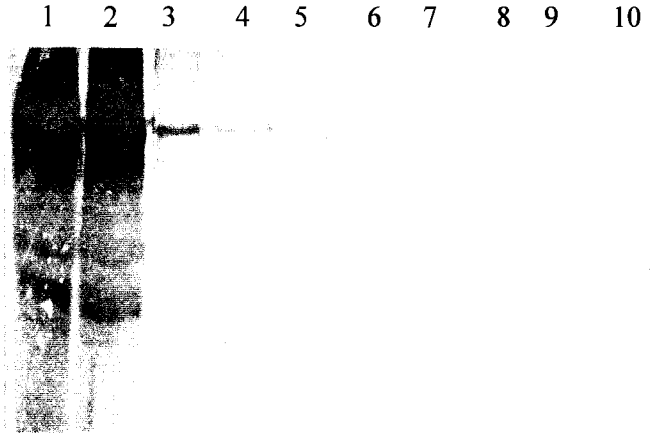
Further support was obtained from LC-ICPMS analyses of two reaction mixtures of $NPAO^{III}$ with either unmodified HSA or cysteine-blocked HSA (figure 3.6). We observed that the HSA containing a free cysteine was able to bind to $NPAO^{III}$ to form a NPAO-HSA complex, which was separated from the unbound $NPAO^{III}$ using size exclusion chromatography (figure 3.6(a)). There was no complex between $NPAO^{III}$ and the HSA that was pretreated to block its reactive cysteine (figure 3.6(b)). 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 results corroborate that BSA can bind to $NPAO^{III}$ and can be captured by $NPAO^{III}$ affinity column (figure 3.7).

To confirm the identity of the captured proteins in the above experiments, we carried out in-gel digestion, and analyzed the products by nanoelectrospray ionization mass spectrometry (nano-ESI-MS). Peptide fingerprint maps from these analyses

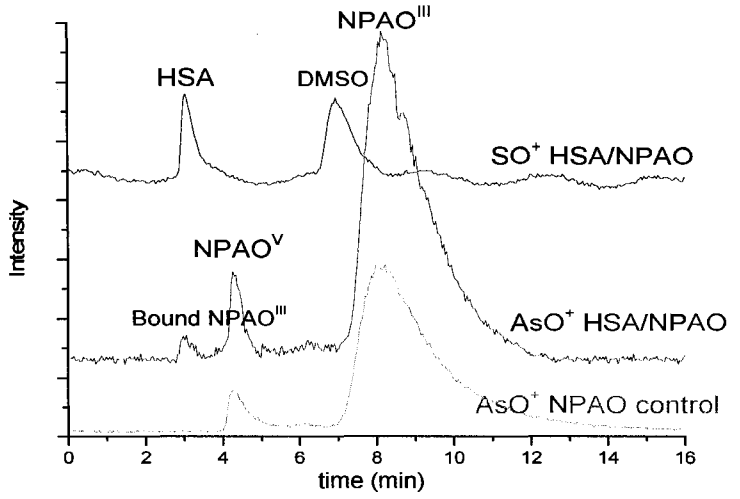


a



b

Figure 3.5. The gel electrophoresis image (a) positive control of HSA, (b)negative control of cysteine blocked HSA; lane1: HSA (a) and cysteine blocked HSA (b); lane 2 to lane 7: 6 rounds of washes, lane 8 & lane 9: 2 rounds of elution by 10 mM DTT in buffer A; lane 10: elution by 100 mM DTT in buffer B show 73% of peptide coverage for BSA and 63% for HSA, indicating that the protein bands from DTT elution solutions are indeed BSA and HSA, respectively.



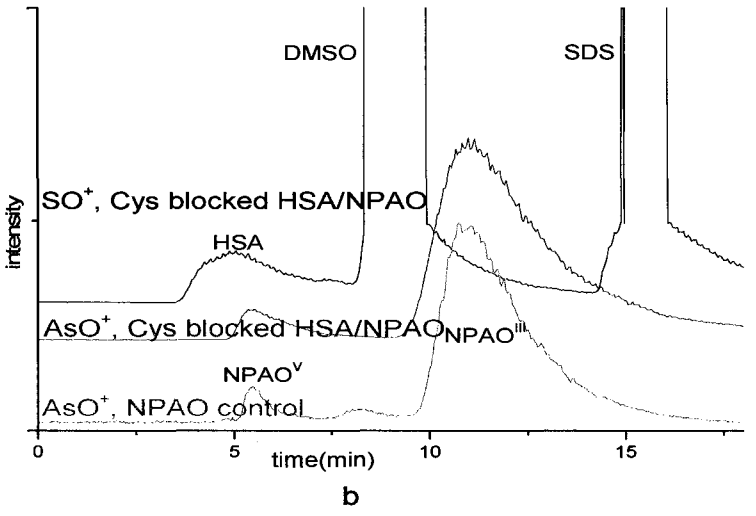


Figure 3.6 The chromatograms of HPLC-ICP-MS for (a)HSA/NPAO incubation sample, (b) cysteine blocked HSA/NPAO incubation sample, black trace: SO^+ signal for detection of HSA; red trace: AsO^+ signal for detection arsenic in incubation sample; green trace: AsO^+ signal for detection ofarsenic in NPAO control sample

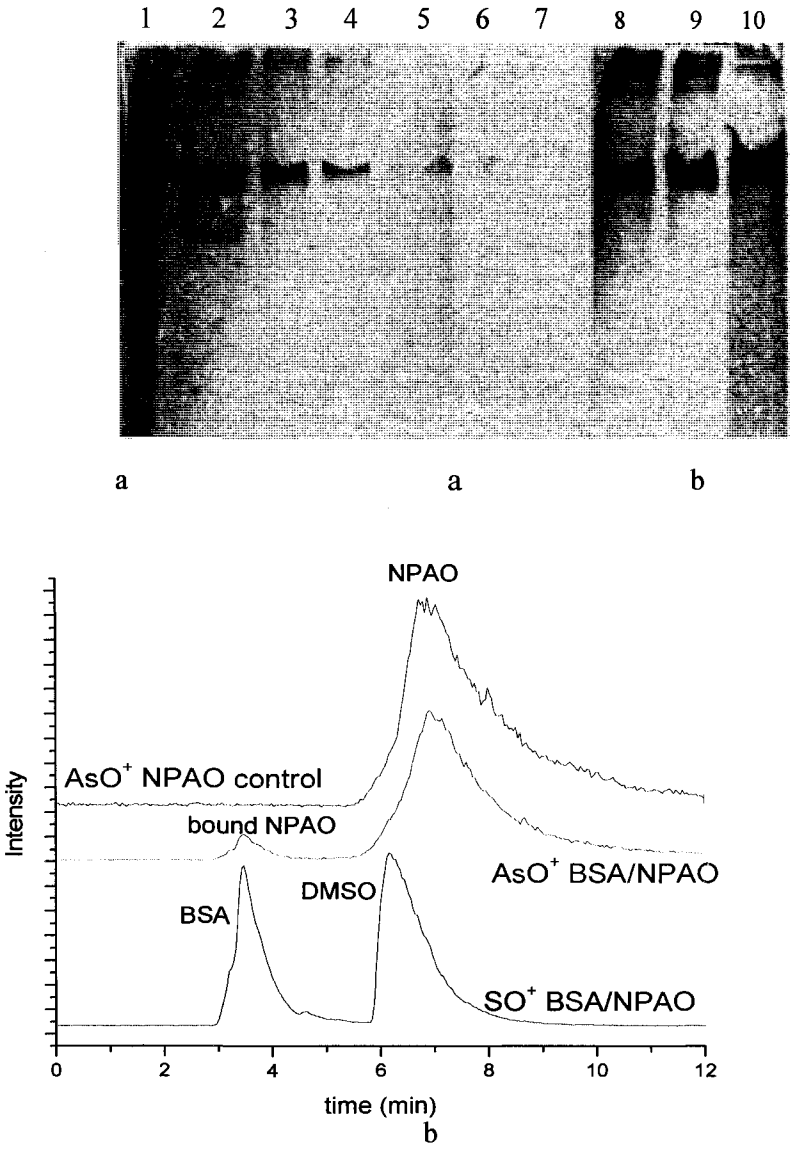
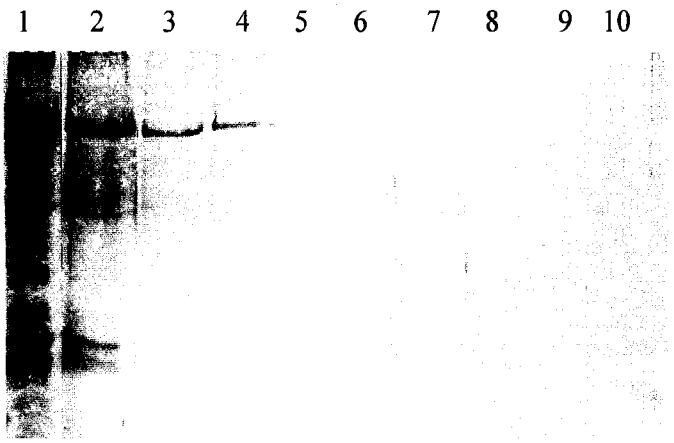


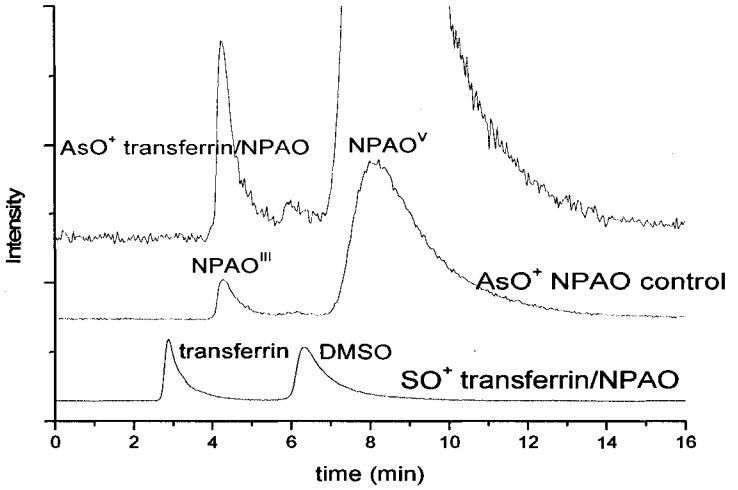
Figure 3.7 (a)The gel electrophoresis image of BSA washed and eluted from affinity column, lane1: BSA without passing through the column; lane 2 to lane 7: 6 rounds of washes, lane 8 & lane 9: 2 rounds of elution by 10 mM DTT in buffer A; lane 10: elution by 100 mM DTT in buffer B; (b) HPLC-ICP-MS chromatograms of BSA/NPAO incubation sample and NPAO control (Agilent ZOBAX GF-250 250×4.6 mm column); black trace: SO^+ signal for detection of BSA; red trace: AsO^+ signal for detection of arsenic in BSA/NPAO sample; blue trace: AsO^+ signal doe detection of arsenic in NPAO control.

In addition to cysteine blocked HSA negative control, transferrin (all cysteines in disulfide bonds) was applied to the affinity column. Transferrin should not be captured by affinity column due to no free cysteine. Figure 3.8 shows the electrophoresis result of transferrin applied to the affinity column and the HPLC-ICP-MS result of transferrin/NPAO incubation sample. No protein band was detected in DTT elution steps after 6 rounds of washes. The LC-ICP-MS result (b) of NPAO/transferrin incubation sample shows transferrin protein elution time was at 3 to 4 minute (SO^+ peak), but no AsO peak was detected at 3-4 min. Both affinity column and ICP-MS confirm that NPAOIII can not bind to transferrin.

To further confirm that the captured proteins are due to arsenic specific binding, not due to any non-specific binding caused by hydrophobic interaction between proteins and benzyl group on 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, and form a similar structure as the affinity resin but without 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. Figure 3.9 shows the over-developed electrophoresis gels from the analysis of the washing and elution solutions after applying the membrane/organelle sample to the non-specific control column. Although proteins are present in the initial washing solutions as expected, no protein band is detected in the DTT elution solution after 6 successive washes by 1% SDS buffer. 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.



a



b

Figure 3.8 (a) The gel electrophoresis image of transferrin washed and eluted from affinity column, lane1: transferrin without passing through the column; lane 2 to lane 7: 6 rounds of washes, lane 8 & lane 9: 2 rounds of elution by 10 mM DTT in buffer A; lane 10: elution by 100 mM DTT in buffer B; (b) HPLC-ICP-MS chromatograms of transferrin/NPAO incubation sample and NPAO control (Phenomenex Biosep-SEC-S 2000, 300×4.6 mm column); black trace: SO^+ signal for detection of transferrin; red trace: AsO^+ signal for detection of arsenic in transferrin/NPAO sample; blue trace: AsO^+ signal for detection of arsenic in NPAO control.

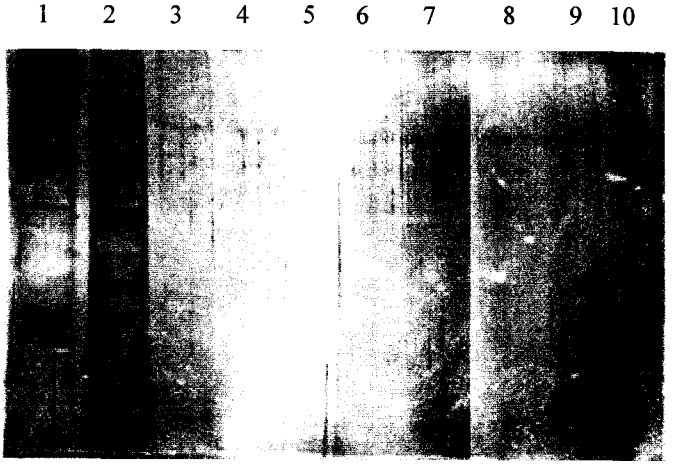


Figure 3.9 Electrophoresis image of Membrane/organelle extraction fraction of A549 cells applied to non affinity control column. lane1: unbound protein fraction washed off by buffer A; lane 2 to 7: 6 round washes by buffer B; lane 8: eluted by 50 mM DDT in buffer A; lane 9 & 10: eluted by 100 mM DTT in buffer B

3.4 Application of the arsenic affinity column to the capture and identification of cellular protein

The benefit of the arsenic affinity column is demonstrated in an application to the identification of arsenic binding proteins in sub-cellular fractions of human cells (A549 lung carcinoma cells). Gel electrophoresis results of the nuclear fraction and membrane/organelle fraction of A549 cells show that after applying the cell samples to the column and 6 successive washes with buffer containing 1% SDS, non-specific components were washed off. Subsequent DTT elution resulted in the release of arsenic-binding proteins from the column.

To purify and identify the arsenic binding proteins in the elution solution, TCA and acetone protein precipitation, SDS-assisted dissolution, in-solution digestion, 2D HPLC separation and LC-ECI-MS/MS were applied. TCA precipitation was used to purify the proteins in solution containing over 10% SDS. After the proteins underwent SDS-assisted dissolution and in-solution digestion, 2D HPLC separation was applied to remove SDS in the peptide solution and achieve peptide fractionation. Figure 3.10 shows the SCX chromatography for peptide purification and fractionation after in-solution digestion. SDS and other chemicals present in peptide solution after in-solution digestion were eluted before 10 minutes in SCX column, and the peptide solution was collected from 22 to 42 min by different time intervals based on the UV signal. The different peptide fractions were analyzed by RP HPLC-MS/MS. A 5min of desalting was applied to remove NaCl in peptide sample, and then the HPLC was switched to mass spectrometry for the IDA analysis. Figure 3.11 shows the baseline peak chromatography of peptide separation by C18 column. HPLC condition was optimized by different gradient procedure, and CID condition was optimized by different peptides from BSA.

CID spectra were submitted to Mascot for database search, followed by manual inspection. 50 proteins in the nuclear fraction of A549 cells (Table 3.2), and 24 proteins in the membrane/organelle fraction (Table 3.3) were identified as arsenic binding proteins.

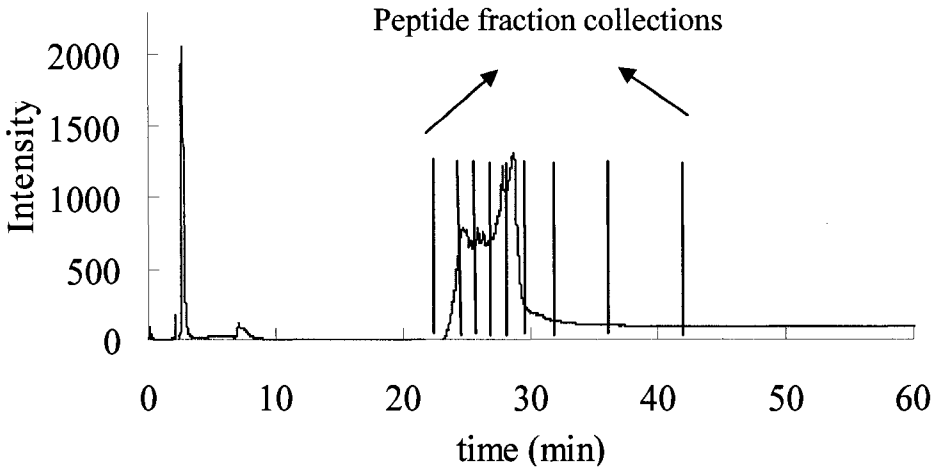


Figure 3.10 SCX chromatography for peptide purification and fractionation after in-solution digestion

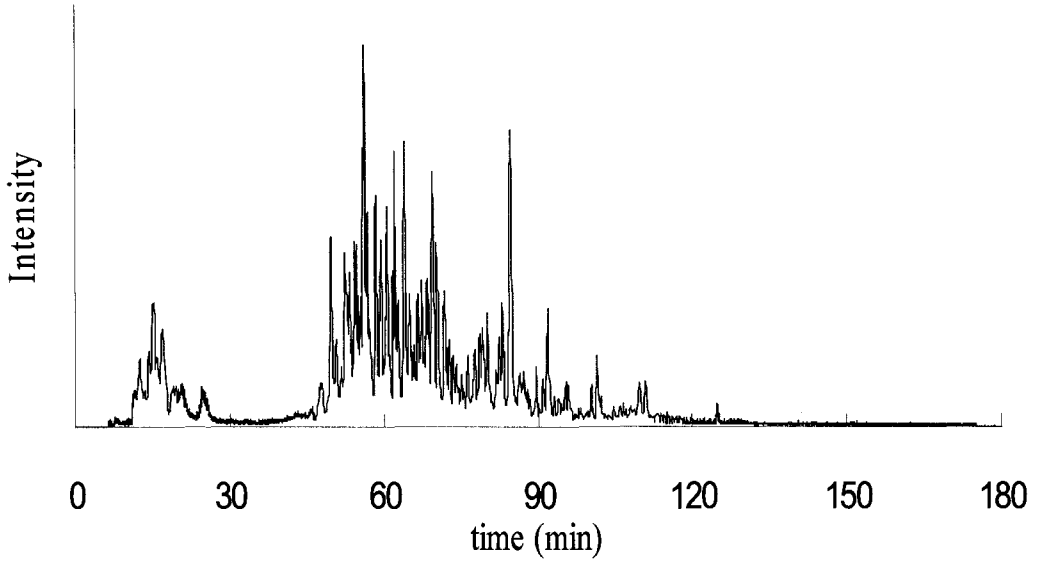


Figure 3.11 Base-peak chromatogram obtained from peptide separation by RP HPLC-MS/MS

Table 3.2 Arsenic binding proteins in A549 nuclear protein fraction

protein	access No.	Mw	Peptide	Possible cysteine position
①Actin, cytoplasmic 1	P60709	42KDa	GYSFTTTAER QEYDESGPSIVHR	17, 217, 257, 272*, 285*, 374
①Actin, aortic smooth muscle	P62736	42KDa	EITALAPSTMK + Oxidation (M) DSYVGDEAQSK	2, 12, 19, 219, 259, 287*, 376
①ADP/ATP translocase 1	P12235	33KDa	TAVAPIER YFPTQALNFAFK	57, 128*, 256*
①ADP/ATP translocase 3	P12236	33KDa	EQGVLSFWR DFLAGGIAAAISK	56, 128*, 159*, 256*
①Annexin A2	P07355	38KDa	AYTNFDAER TPAQYDASELK	9, 335*
③Anterior gradient protein 2 homolog precursor	O95994	20kDa	ADITGR IMFVDPSLTVR	81
③ATP-dependent RNA helicase A	Q08211	141kDa	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
①ATP-dependent DNA helicase 2 subunit 1	P12956	70kDa	ILELDQFK LGSLVDEFK	66*, 150, 389, 398, 585*
②DNA topoisomerase 2-alpha	P11388	174kDa	YGVFPLR YDTVLDILR	104*,170,216*300, 392 , 405* 427, 455, 733, 862,997, 1008*, 1145*
③DNA-dependent	P78527	469kDa	AQEPESGLSEETQVK ILELSGSSSEDSEK	78 cysteines

protein kinase catalytic subunit				
③Dynein heavy chain, cytoplasmic	Q14204	532kDa	GTFDNAETK ETVDQVEELR	36 cysteines
②4F2 cell-surface antigen heavy chain	P08195	58KDa	GQSEDPGSLLSLFR VAEDEAEAAAAAK	330*
②Filamin-A	P21333	280kDa	ENGVYLIDVK AGGPGLER	41 cysteines
①Fructose-bisphosphate aldolase A	P04075	39kDa	GILAADESTGSIAK GILAADESTGSIAK	73*,135*,150, 178*,202,240*290*, 339
②Heat shock cognate 71 kDa protein	P11142	71KDa	DAGTIAGLNVLR VEIIANDQGNR	17, 267*, 603
③Heterogeneous nuclear ribonucleoprotein A/B	Q99729	37KDa	GFGFILFK GFVFITFK	97, 223
②Heterogeneous nuclear ribonucleoprotein A3	P51991	40KDa	EDTEEYNLR DYFEKYGK	64*, 85*, 94*, 196
②Heterogeneous nuclear ribonucleoproteins A2/B1	P22626	37KDa	QEMQEVQSSR GGNFGFGDSR	50*
②Heterogeneous nuclear ribonucleoproteins C1/C2	P07910	34KDa	VPPPPPIAR SDVEAIFSK	46*
②Heterogeneous nuclear ribonucleoprotein C-like 1	O60812	32KDa	VDSLLENLEK VFIGNLNTLVVK	
②Heterogeneous	Q14103	39KDa	GFGFVLFK IFVGGLSPDTPEEK	126, 226*, 252*

nuclear ribonucleoprotein D0				
③Heterogeneous nuclear ribonucleoprotein K	P61978	51KDa	GSDFDCELR IDEPLEGSEDR	132, 145, 184, 185, 205
③Heterogeneous nuclear ribonucleoprotein L	P14866	60kDa	R.YYGGGSEGGR.A SDALETLGFLNHYQMK	120, 184, 187, 229, 230, 373, 421, 441, 490, 550
③Heterogeneous nuclear ribonucleoprotein R	O43390	71KDa	ENILEEFSK DLYEDELVPLFEK	99, 214, 226, 240, 292
③Heterogeneous nuclear ribonucleoprotein Q	O60506	69kDa	AGPIWDLR GFCFLEYEDHK	96,211,237, 289
③Interleukin enhancer -binding factor 2	Q12905	43KDa	ILPTLEAVAALGNK VKPAPDETSFSEALLK	37, 271, 291, 311
③Keratin, type I cytoskeletal 10	P13645	60KDa	LASYLDK SQYEQLAEQNR	25, 66, 401, 427
③Keratin, type I cytoskeletal 9	P35527	62KDa	SGGGGGGGLGSGGSIR FSSSSGYGGGSSR	3, 61, 406, 432
③Keratin, type II cytoskeletal 2 epidermal	P35908	66KDa	NLDLDSIIAEVK STSSFSCLSR	3, 7, 42, 423, 495
③Keratin, type II cytoskeletal 1	P04264	66KDa	YEDEINK AEAESLYQSK	48, 146, 496
②Lamin-A/C	P02545	74kDa	EGDLIAAQAR EDLQELNDR	552*, 570, 588, 591, 661
③Lamin-B1	P20700	66KDa	EYEAALNSK ESDLNGAQIK	110, 198, 317, 443, 583
Lamina-associated polypeptide 2	P42167	51KDa	PEFLEDPSVLTK HASPILPITEFSDIPR	362

③Leucine-rich repeat-containing protein 59	Q96AG4	35KDa	VAGDCLDEK ATILDLSCNK	48, 59, 131, 140, 266, 277
②Matrin-3	P43243	95KDa	SFQQSSLSR TEEGPTLSYGR	230, 293, 296, 552*, 563*,
□Non-POU domain-containing octamer-binding protein	Q15233	54kDa	NLPQYVSNELLEEAFSVF GQVER AVVIVDDR	145*,208
③Nucleolar RNA helicase 2	Q9NR30	89KDa	GVTFLFPIQAK EQLGEEIDSK	161, 291, 378, 418, 445, 537, 643, 682
②Nucleophosmin	P06748	32KDa	VDNDENEHQLSLR GPSSVEDIK	21, 104*, 275
②Peroxiredoxin-1	Q06830	22kDa	TIAQDYGVLK QITVNDLPVGR	71, 83*
③Phosphate carrier protein	Q00325	41KDa	IQTQPGYANTLR GIFNGFSVTLK	57, 68, 76, 91, 136, 237, 257, 277
②Polypyrimidine tract-binding protein 1	P26599	57kDa	KLPIDVTEGEVISLGLPFG K LPIDVTEGEVISLGLPFGK	23, 250*, 251
②Poly(rC)-binding protein 1	Q15365	37kDa	INISEGNCPER IANPVEGSSGR	54*, 109, 118, 158, 194, 201, 293*, 355
③Probable ATP-dependent RNA helicase DDX5	P17844	69kDa	TTYLVLDEADR GDGPICLVLAPTR	89, 170, 191, 194, 200, 221, 234, 320, 354
③Probable ATP-dependent RNA helicase DDX17	Q92841	72kDa	GDGPICLVLAPTR LIQLMEEIMAEK + 2 Oxidation (M)	87, 120, 168, 191, 198, 219, 240, 318, 352, 368, 505
①40S ribosomal protein S14	P62263	16KDa	IEDVTPIPSDSTR TPGPGAQSALR	30, 53*, 84
①Small nuclear ribonucleoprotein Sm D2	P62316	14KDa	SEMTPEELQK NNTQVLINCR	46*, 63*

③Spectrin alpha chain	Q13813	284KDa	VNSLGETAER DLTGVQNLR	158, 315, 466, 956, 1091, 1314, 1444, 1454, 1622, 1930, 2120, 2233, 2351, 2441
②Spectrin beta chain, brain 1	Q01082	275KDa	VLDNAIETEK LVSDGNINSDR	73, 112, 183, 604, 619, 624, 861, 964, 1284, 1389, 1900, 1970, 2227, 2262*
②Splicing factor, arginine/serine -rich 3	P84103	20KDa	AFGYYGPLR NPPGFAFVEFEDPR	6*, 10, 72*, 74
②U2 small nuclear ribonucleoprote in A'	P09661	29KDa	TFNPGAGLPTDK SLTYLSILR	77*, 89
②Vimentin	P08670	53KDa	QVDQLTNDK QDVDNASLAR	327*

Cys in Filamin-A are 205, 210, 444, 478, 543, 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

Cys 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

Dynein heavy chain, cytoplasmic : 220, 416, 867, 1059, 1484, 1888, 1932, 1956, 1977, 1999, 2076, 2186, 2659, 2454, 2466, 2594, 2639, 2663, 2712, 3147, 3325, 3389, 3507, 3693, 3712, 3808, 3940, 4044, 4121, 4170, 4216, 4510, 4540, 4556, 4570, 4644

Table 3.3 Arsenic binding proteins in A549 membrane protein fraction

protein	Access No.	Mw	peptide	Possible cysteine
①Actin	P60709 P62736	42KDa	EITALAPSTMK+Oxidation (M) DLTDYLMK+Oxidation (M)	17, 217, 257, 272*, 285*, 374 2, 12, 19, 219, 259, 287*, 376
①Annexin A2	P07355	38KDa	SYSPYDMLESIR TPAQYDASELK	9, 335*
① Elongation factor1-alpha1 ① Elongation factor 1-alpha 2	P68104 Q05639	50KDa 50KDa	IGGIGTVPVGR QLIVGVNK	31*, 111, 234*, 363, 370*, 411 31*, 111, 326, 363, 370*, 411
②4F2 cell-surface antigen heavy chain	P08195	60KDa	EEGSPLELER VAEDEAEAAAAAK	330*
② 78 kDa glucose-regulated protein precursor	P11021	72KDa	ELEEIVQPIISK VEIIANDQGNR	41, 420*
① Glucose-6-phosphate 1-dehydrogenase	P11413	59KDa	GYLDDPTVPR DNIACVILTFK	13*, 158, 232, 269, 294*,358, 385, 446*,
③ 60 kDa heat shock protein	P10809	61KDa	VGLQVVAVK NAGVEGSLIVEK	237, 442, 447
① 3-hydroxyacyl-CoA dehydrogenase type-2	Q99714	27KDa	DLAPIGIR DVQTALALAK	5, 58, 91*
③Keratin, type I	P35527	62KDa	TLLDIDNTR QEYEQLIAK	3, 61, 406, 432

cytoskeletal 9				
③Keratin, type I cytoskeletal 10	P13645	60KDa	LAADDFR LASYLDK	25, 66, 401, 427
③Keratin, type II cytoskeletal 1	P04264	66KDa	IEISELNR TLLEGEESR	48, 146, 496
③Keratin, type II cytoskeletal 2 epidermal	P35908	66KDa	FASFIDK TAAENDFVTLK	3, 7, 42, 423, 495
② Peroxiredoxin-1	Q06830	22KDa	ADEGISFR TIAQDYGVLK	71, 83*
③ Protein disulfide-isomerase A6 precursor	Q15084	48KDa	GSTAPVGGGAFPTIVER GESPVDYDGGR	11, 291, 292
①Pyruvate kinase isozymes	P14618	58KDa	GDYPLEAVR GDLGIEIPAEK	31*, 49, 152, 165, 317, 326, 357, 423*, 424*, 474*
③ Ras-related protein Rab-10 ③ Ras-related protein Rab-15 ③ Ras-related protein Rab-35 ② Ras-related protein Rab-1A ② Ras-related protein Rab-1B	P61026 P59190 Q15286 P62820 Q9H0U4	23KDa 24KDa 23KDa 23KDa 22KDa	LLLIGDSGVGK AFLTLAEDILR LQIWDTAGQER	24, 124, 199, 200 23, 26, 153, 210, 212 110, 114, 163, 200, 201 26, 126*, 204, 205 23*, 200, 201
② Ras-related protein Rab-7	P51149	24KDa	DEFLIQASPR EAINVEQAFQTIAR	83, 84, 143*, 205, 207
②Reticulon-4	Q9NQC3	130KDa	GPLPAAPPVAPER GPLPAAPPVAPER	424, 464, 559, 597, 699, 1101*

③ Stomatin-like protein 2	Q9UJZ1	39KDa	AEQINQAAGEASAVLAK DVQGTDASLDEELDR	167, 172
③ Stress-70 protein	p38646	74KDa	VQQTVQDLFGR AQFEGIVTDLIR	66, 317, 366, 487, 608
③ Thioredoxin domain-containing protein 1 precursor	Q9H3N1	32KDa	VDVTEQPGLSGR DFINFISDK	106, 165, 198, 205, 207
① Tubulin alpha-ubiquitous chain	P68363	51KDa	QLFHPEQLITGK IHFPLATYAPVISAEK	4, 20, 25*, 129*, 200, 213*, 295, 305, 315, 316, 347*, 376
③ Voltage-dependent anion-selective channel protein 2	P45880	32KDa	LTLSALVDGK YQLDPTASISAK	8, 13, 47, 76, 103, 133, 138, 210, 227,
③ Voltage-dependent anion-selective channel protein 3	Q9Y277	31KDa	LTLDTIFVPNTGK LTLSALIDGK	2, 8, 36, 65, 122, 229

Table 3.4 Affinity column recovery

	Incubation	Buffer A	Buffer B	DTT +buffer A	recovery
Carbonic anhydrase	2.27mg	1.68mg.	0.520mg		95.9±1.6%
Cys blocked HSA	1.132mg	0.776mg	0.244		91.8% (AVE)
Transferrin	4.03mg	3.85mg	0.300mg		107±4%
BSA	5.925mg	4.41mg	0.113mg	0.083mg	84.3±5.7%
HSA	1.03mg	0.752mg	0.029mg	0.070mg	84.7±1.9%
Nucleus fraction of A549 cell	1.14mg	0.545mg	0.085mg	0.113mg	66.2% (AVE)
Membrane/organelle fraction of A549 cell	1.21mg	0.245mg	0.449mg	0.078mg	65.1% (AVE)

4. Recovery of proteins from the affinity column

All the fractions from positive and negative control were collected to determine the protein concentration for the affinity column recovery by following the instruction of Bio-rad RC DC protein assay. Table 3.4 shows the column recovery from all samples. The proteins that were not captured by the affinity column were collected in the washing solutions, and the recovery was >90%. Two proteins (HSA and BSA) that were captured by the affinity column had about 80% recovery. Sample from the two cell extracts showed about 65% recovery. This lower recovery is probably because some proteins were retained on the affinity column and were not eluted by DTT solution. The eluant with higher affinity to phenylarsine oxide, such as BAL, could be tested.

5. Discussion

Several reports [16, 18, 22, 23] 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 Blot, 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 nuclear fraction and 24 proteins in organelle/membrane fraction of A549 cells as arsenic binding proteins.

Much research has focused on the arsenic binding to peptides or proteins containing vicinal dithiol[26]. The length of As-S bond is 2.23Å and the S-As-S angle is 92.7°. The two cysteines[8] in ArsA ATPase, having a distance between 3 Å and 6 Å can bind to trivalent arsenic. In the cys-$(x)_n$-cys peptide model ($n<14$), the location and the sequence

arrangement of the cysteines in an α helix (n<14) make very little difference for their binding to arsenite. But the peptides having a monothiol have lower binding affinity than peptides having dithiols; k_d values for arsenic binding to peptides containing a monothiol are at least 6 times less than those binding to peptides of dithiols.

The present study and previous research[12] from our group 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. X-ray crystallization structure of HSA[27] shows that cys34 does not participate in any disulfide bridge, its Sγ atom is toward interior and surrounded by Pro53, His39, Val77 and Tyr84 or the side chain. In solution, the side chain of Tyr84 may flip over and enable external counterparts, such as arsenic, to bind to the thiol and form arsenic complex. Another possibility is that the HSA backbone conformation may change to bring the sulfhydryl group toward the exterior of the protein. Hemoglobin is a tetramer consisting of two α chains and two β chains. The three-dimensional structure of hemoglobin in Swissprot database shows that the position of cys104 in the α chain is embedded inside the protein in a hydrophobic environment; the position of cys93 in the β chain is on the surface of protein with the sulfhydryl group towards the exterior of the protein surrounded by H_2O molecule, and the position of cys112 in the β chain is close to the surface of the protein with the sulfhydryl group pointing towards the exterior of the protein in a hydrophilic environment. It is possible for cys93 and cys112 in the β chain to bind to trivalent arsenicals. These examples suggest that cys-$(X)_n$-cys in α helix is not necessary for arsenic binding to proteins. The three dimensional structure of proteins, including the position of cysteine in the protein (embeded inside or close to surface), other amino acids in the side chain surrounding the cysteine (steric accessability and hindrance , hydrophobic or hydrophilic environment, the electron density and charge state), and the orientation of thiol group (towards exterior or interior), are more important to

arsenic binding.

In our study, the affinity resin Eupergit® C has a three-atom 6Å hydrophilic spacer connected to NPAO, so 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 in the protein, can not be selected by the affinity column. We examined the three-dimensional conformation of the identified proteins from Swiss Prot and literatures. The cysteine position column in Table 2 and 3 shows the free cysteines in the identified proteins, and the cysteine residues in the disulfide 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 dimension 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 3D conformation examination shows that protein binding to arsenic can happen by single cysteine in α helices, β sheet or loop. Cysteine position in protein tertiary structure plays an important role in arsenic binding affinity.

6. Conclusions

The average amount of NPAO immobilized is about 27.3±2.3 μmoles/g wet resin. The specificity studies of positive controls and negative controls confirmed that arsenic affinity column can be used to capture the arsenic specific binding proteins. The combination of the arsenic affinity technique and mass spectrometry is a practical method to identify multiple arsenic-binding proteins simultaneously. Total 50 proteins in nuclear

fraction and 24 proteins in membrane fraction of A549 cells were identified as arsenic-binding proteins. The results from examining protein 3D confirmation of the identified proteins show that the α helices peptide model cys-(X)n-cys is not necessary for arsenic binding. Arsenic can also bind to a single cysteine in α helices, β sheet or loop. The binding mainly depends on the cysteine position in the tertiary structure of protein.

7. References

1. *Arsenic in Drinking Water*. National Research Council, National Academy Press: Washington, DC: 1999 and 2001.
2. Tseng, C.H.; Tseng, C.P.; Chiou, H.Y.; Hsueh, T.M.; Chong, C.K., Epidemiologic evidence of diabetogenic effect of arsenic. *Toxicol. Lett.* **2002,** 133, 69-76.
3. Chen, C.J.; Chuang, Y.C.; Lin, T.M.; Wu, H.Y., Malignant neoplasms among residents of a blackfoot disease endemic area in Taiwan-high arsenic artesian well water and cancers. *Cancer Res.* **1985,** 45, 5895-5899.
4. Chiou, H.Y.; Hsueh, Y.M., Incidence of internal cancers and ingested inorganic arsenic: A 7 year follow-up study in Taiwan *Cancer Res.* **1995,** 55, 1296–1300.
5. Hopenhayn-Rich C. Bladder cancer mortality associated with arsenic in drinking water in Argentina. *Epidemiology* **1996,** 7, (117-124).
6. Smith, A.H.; Goycolea, M.; Haque, R.; M.L.; Marked, B.,increase in bladder and lung cancer mortality in a region of Northern Chile due to arsenic in drinking water *Am. J. Epidemiol* **1998,** 147, 660–669.
7. Kitchin, K. T.; Wallace,K., Arsenite binding to synthetic peptides based on the Zn finger region and the estrogen binding region of the human estrogen receptor-α. *Toxicol. Appl. Pharm.* **2005,** 206, 66-72.
8. Bhattacharjee, H.; Rosen, B. P., Spatial proximity of cys113, Cys172, and Cys422 in theMetalloactivation Domain of the ArsA ATPase. *J.Bio.Chem.* **1996,** 271, (40), 24465-24470.
9. Cline, D. J.; Thorpe, C.; Schneider, J. P., Effects of As(III) binding on α-helical structure *J. Am. Chem. Soc.* **2003,** 125, 2923-2929.
10. Kitchin, K. T.; Wallace,K., Dissociation of arsenite-peptide complexes: Triphasic nature, rate condtants, half-lives, and biological importance *J. Biochem. Mol. Toxicol.* **2006,** 20, (1), 48-55.
11. Kitchin, K. T.; Wallace,K., Arsenite binding to synthetic peptides: The effect of increasing

length between two cysteines. *J.Biochem.Mol.Toxicol* **2006,** 20, (1), 35-38.

12. Lu, M.; Wang, H.; Li, X.; Le, X. C., Evidence of Hemoglobin binding to arsenic as a basis for the accumulation of arsenic in ret blood *Chem,. Res. Toxicol.* **2004,** 17, 1733-1742
13. Jiang, G; Gong, Z.; Li, X.-F., Interaction of trivalent arsenicals with metallothionein *Chem. Res. Toxicol.* **2003,** 16, 873-880.
14. Merrifield, M. E.; Ngu, T.; Stillman, M. J., Arsenic binding to Fucus vesiculosus metallothionein. *Biochem.& Biophys. Res. Commun.* **2004,** 324, (127-132).
15. Lin, C.; Huang, C.; Chen, W.-Y., Characterization of the interaction of Galectin-1 with sodium arsenic *Chem.Res.Toxicol.* **2006,** 19, 469-474.
16. Chang, K. N.; Lee, T. C.; Chen, M. F., Identification of galectin I and thioredoxin peroxidase II as two arsenic binding proteins in Chinese hamster ovary cells *Biochem.J.* **2003,** 371, 495-503.
17. Shi, W; Dong, J.; Scott, R. A., The role of arsenic-thiol interactions in metalloregulation of the ars operon *J. Biol. Chem.* **1996,** 271, (16), 9291-9297.
18. Hoffman, R. D.; Lane, M. D., Iodophenylarsine oxide and arsenical affinity chromatography:New probes for dithiol proteins *J. Biol. Chem.* **1992,** 267, (20), 14005-14011.
19. Rey, N. A.; Howarth, O. W.; Pereira-Maia, E. C., Equilibrium characterization of the As(III)-cysteine and the As(III)-glutathione systems in aqueous solution *J. Inorg. Biochem.* **2004,** 98, 1151-1159.
20. Spuches, A. M.; Hruszyna, H. G; Rich, A. M.; Wilcox, D. E., Thermodynamics of the As(III)-thiol interaction: arsenite and monomethylarsenite complexes with glutathione, dihydrolipoic acid, and other thiol ligands. *Inorg. Chem.* **2005,** 44, 2964-2972.
21. Raab, A.; Meharg, A. A.; Jaspars, M., Arsenic-glutathione complexes-their stability in solution and during separation by different HPLC modes *J. Anal. At. Spectrom.* **2004,** 19, 183-190.
22. Menzel, D. B.; Hamadeh, H. K.; Lee, E., Arsenic binding proteins from human

lymphoblastoid cells *Toxicol. Lett.* **1999,** 105, 89-101.

23. Kalef, E.; Ealfish, P. G.; Gitler, C., Arsenical-Base Affinity chromatography of vicinal dithiol-containing proteins: purification of L1210 leukemia cytoplasmic proteins and the recombinant rat c-erb Aß1 T3 receptor. *Anal. Biochem.* **1993,** 212, 325-334.
24. Chen, Y.; Kwon, S. W.; Kim, S. C., Integrated approach for manual evaluation of peptides identified by searching protein sequence database with tandem mass spectra *J. Proteome Res.* **2005,** 4, 998-1005.
25. Forbus, J.; Spratt, H.; Wiktorowicz, J., Functional analysis of the nuclear proteome of human A549 alveolar epithelial cells by HPLC-high resolution 2-D gel electrophoresis *Proteomics* **2006,** 6, 2656-2672.
26. Sowerby, D. B., *The chemistry of organic arsenic antimony and bismuth compounds (patai, S.,ed)*. John Wiley and Sons, Inc., New York: 1994.
27. Sugio S., K. A., Mochizuki S., Noda M., Kobayashi K., Crystal structure of human serum albumin at 2.5-A resolution. *Protein Eng.* **1999,** 12, (439-446).

Chapter 4

Affinity capture and mass spectrometry identification of arsenic binding proteins from nucleus of A549 cells

1. Introduction

A few studies investigated the effect of arsenic on the formation of DNA adduct when cells were treated with benzo[a]pyrene (B[α]P) or its metabolite B[α]P diol epoxide (BPDE) in combination with arsenic. Arsenite was shown to increase BPDE-DNA adduct in both in vivo and in vitro studies.[1-4] Shen et al[5] found that arsenite enhanced the formation of BPDE-DNA adducts by increasing the cellular uptake of BPDE. Schwerdtle at al[4] found that all trivalent arsenicals were able to release zinc from XPAzf and that the activity of isolated Fpg was inhibited by MMA^{III} and DMA^{III}. Their results also suggest that trivalent arsenicals may contribute to diminished nucleotide excision repair by binding to DNA repair proteins.

To identify the DNA repair proteins binding to arsenic, we attempted to capture and characterize the arsenic binding proteins in the nucleus of control cells and cells treated with BPDE alone or in combination with arsenite. We then compared the difference of captured proteins in cells of the three different treatments. The objective is to detect arsenic-binding proteins that may be involved in the formation and/or repair of BPDE-DNA adduct.

2. Experimental

2.1 Materials

All the chemical materials used in this chapter are the same as in chapter 3. Nitrocellulose membrane (162-0115) was purchased from Bio-Rad (Mississauga, ON). P53 antibody (DO-1, Sc-126) and MDM2 antibody (C-18, Sc-812) were from Santa Cruz

Biotechnology (Santa Cruz, CA, USA). PARP-1 antibody (#9542) and β-Actin antibody (#4967) were from Cell Signaling Technology (Danvers, MA, USA). Lumi-Light Western Blotting substrate (12015200001) was from Roche (Switzerland).

2.2 cell culture and protein extraction

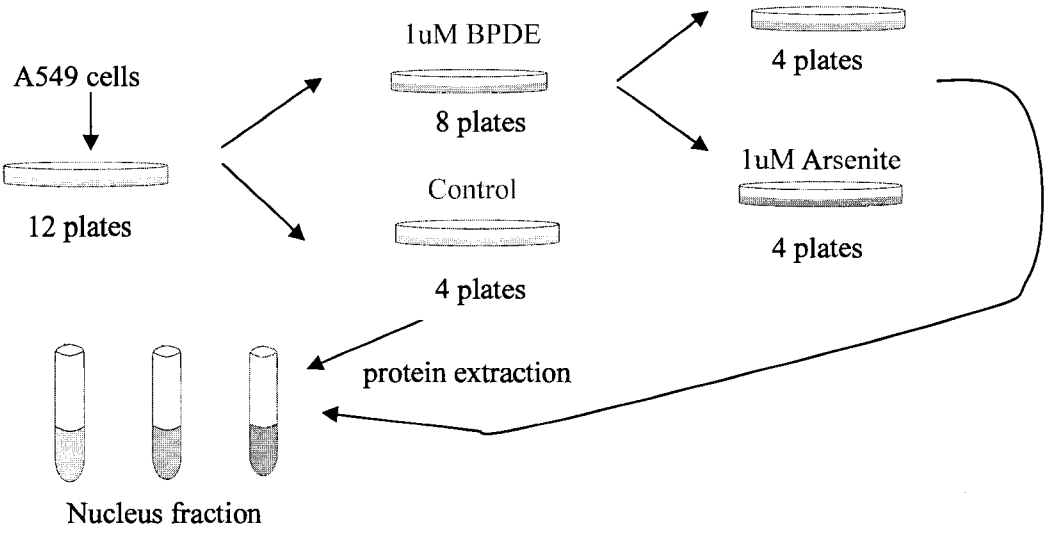


Figure 4.1 A549 cells incubated in the control medium or supplemented with BPDE, followed by incubation in the control medium or in medium supplemented with arsenite

A flow chart of cell treatment is shown in Figure 4.1. A549 cells (12 plates) were cultured in RPMI-1640 medium supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin. Eight plates of A549 cells were incubated in the medium with the addition of 1 μM BPDE, the other 4 plates of A549 cells stayed in the normal medium as control. After 0.5 hour incubation with 1 μM BPDE, the BPDE containing medium was removed from 8 plates of A549 cells, and the cells were rinsed with PBS for three times. Then the cells in four of these plates were incubated in the normal medium (RPMI-1640 containing 10% fetal bovine serum and 1% penicillin/streptomycin), and the cells in the other four plates were incubated in the medium (RPMI-1640 containing 10% fetal bovine

serum, 1% penicillin/streptomycin) with addition of 1 μM arsenite. These cells were incubated for another 12 hours. The cells in three types of treatments were separately harvested, and proteins were extracted by ProteoExtact® subcellular proteome extraction kit following the instruction of manufacture.

2.3 Preparation of affinity column

Affinity column was prepared by following the steps described in chapter 3. Three affinity columns were prepared simultaneously to get the same NPAO affinity media for each column. In the end, each NPAO affinity column was about 2.5 ml (2 gram)

2.4 Selection of arsenic-binding protein by affinity column

Each nuclear protein extract (2 mL pH 7.5) was poured into an affinity column, and incubated at room temperature for two hour with slow shaking. The column was washed by 10 mL buffer A (20 mM Hepes, 50 mM NaCl, pH=7.5) to remove the unbound proteins. Then, the column was further washed by 5 mL buffer B (1% SDS in buffer A) for six times to remove the non-specific binding proteins. The remaining proteins were eluted for three times by 2 mL 10 mM DTT in Buffer A, 2 mL 50 mM DTT in buffer B, and 2 mL 100 mM DTT in buffer B. Each washing and elution step took 10 minutes with slow shaking. All the fractions were separately collected and proteins were determined by Bio-Rad RC DC protein assay.

To ensure that all the specific arsenic-binding proteins were captured by affinity column, a second affinity column was used to capture the proteins that were missed by the first column. Briefly, the wash fraction from the first affinity column by buffer A was freeze-dried, then re-dissolved in 1 mL deionized water, and incubated with 1 mM TCEP at room temperature for 1 hour to reduce any oxidized proteins. This solution was then applied to a newly prepared affinity column to capture the arsenic-binding proteins. The same wash and elution steps as above were applied. The results from the second affinity

selection were compared with the results from the first affinity selection.

2.5 SDS-polyacrylamide gel electrophoresis

All fractions eluted from affinity column were freeze-dried. SDS-PAGE and silver staining were performed to detect the proteins as described in chapter 3.

2.6 Protein purification and in-solution digestion

Protein purification and in-solution digestion were carried out as described in chapter 3.

2.7 Analysis of proteins and peptides

LC-ESI-MS/MS was performed using an Agilent 1100 capillary HPLC and Applied Biosystem mass spectrometry (API QSTAR Pulsar-I, AB, PE SCIEX). 1 μL peptide solution after in-solution digestion was injected and separated by reversed-phase chromatography on a 0.5×150 mm Agilent C18 column. Gradient elution was performed with solvent A (0.1%, V/V, acetic acid) and B (0.1%, V/V, acetic acid in ACN). The gradient program was 5%B for 5 minute (desalting), 5%-30% B between 5 to 40 minutes, 30%-55% B between 40 to 150 minutes, and 55%-95% B between 150 to 190 minutes. The flow-rate was 2 μL/min.

2.8 Database search

Mascot search program (http://www.matrixscience.com/cgi/search_form) was used for all database searches as described in Chapter 3.

2.9 Western blot

The nuclear proteins, after affinity capture, were also analyzed by Western blot to detect the arsenic binding proteins whose amount were lower than the detection limit of

mass spectrometry. The nuclear fractions of proteins from cells of all three treatments were applied to three affinity columns. Three DTT elution fractions were obtained. As a non-affinity control, BPDE treated cell sample was applied to non-affinity column (chapter 3), and the DTT elution fraction was also obtained. The three nuclear fractions of cell extraction, three DTT elution fractions from affinity column, and the DTT elution fraction from the non-affinity column were applied to Western blot analysis.

The DTT elution solutions from the same cell treatment were pooled and freeze-dried. A total 400 μL deionized water was added to re-dissolve the proteins. 10 μg proteins in this sample were loaded onto 4%/10% stacking/separation SDS-PAGE. Western blot was prepared by transfering proteins from the gel to a nitrocellulose membrane at 115 V for 75 minutes in transfer buffer (20% methanol in TG×1 buffer) at 4°C. The remaining protein binding sites on the membrane were blocked by rinsing the membrane with 5% nonfat dry milk in PBSMT buffer (PBS×1, pH 7.4, 0.1% Tween 20) for three times (15 min per time) at room temperature. Membrane was incubated with primary antibodies in 5% nonfat milk PBSMT buffer at 4°C overnight. The unbound primary antibodies were removed by rinsing the membrane with 5% nonfat milk PBSMT solution for three times (15 min per time) at room temperature. Secondary antibodies were diluted at 1:10000 and incubated with the membrane for 45 minutes at room temperature in 5% nonfat milk PBSMT buffer, and then the unbound secondary antibodies were removed by rinsing the membrane twice with PBSMT solution for 5 minutes and then with PBS for 5 minutes. For the detection step, the enhanced chemiluminescence (ECL) substrate was used according to the instruction of manufacture.

The primary antibody were a monoclonal mouse antibody for p53 (1:500 dilution), a polyclonal rabbit antibody for MDM2 (1:100 dilution), a polyclonal rabbit antibody for PARP-1 (1:1000 dilution), and a polyclonal rabbit antibody for the loading control protein actin-β (1:1000 dilution). The secondary antibody was used in 1:10000 dilution.

3 Results and discussion

3.1 The captured proteins from three treatment samples

The nuclear proteins from A549 cells of three treatments (control, BPDE, and BPDE followed by arsenite) were individually applied to three affinity columns prepared simultaneously. After non-specific binding proteins were removed, the arsenic specific binding proteins were eluted with DTT three times, and all the fractions were collected for freeze-dry and gel electrophoresis.

Figure 4.2 shows the gel image of three samples. In the 5th and 6th rounds of wash with 1% SDS (lanes 2 to 7), there were still some proteins remaining. With the subsequent elution by DTT, additional proteins were eluted from the column for the samples of control cells (lane 8), BPDE treated cells (lane 9), and BPDE + As^{III} treated cells (lane 10). Most bands in lanes 8-10 were similar, indicating that similar proteins were captured on the arsenic affinity column.

The HPLC-ESI-MS/MS results from the in-solution digestion samples of the three treatment samples confirmed that the arsenic specific binding proteins in the three different treatment samples were similar. Table 4.1 shows a list of 56 proteins that were captured by the arsenic affinity column and subsequently detected by HPLC-ESI-MS/MS. These proteins are common to all the nucleus extract samples from the three cell treatment procedures (control cells, BPDE treated cells and BPDE + As^{III} treated cells). A few captured and detected proteins were different between the cells of the three treatments. These are summarized in Tables 4.2, 4.3, 4.4.

Table 4.1 to 4.4 also show the free cysteines in the identified proteins. The cysteine residues in the disulfide and those involved in Zn finger binding are not included in the Tables. The cysteine positions labeled with * are the cysteines exposed at the exterior of the protein, and they are probably the arsenic binding sites in the protein. 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 dimension

conformation was examined by their peptide fragments. The proteins labeled with ③ have no conformational data available.

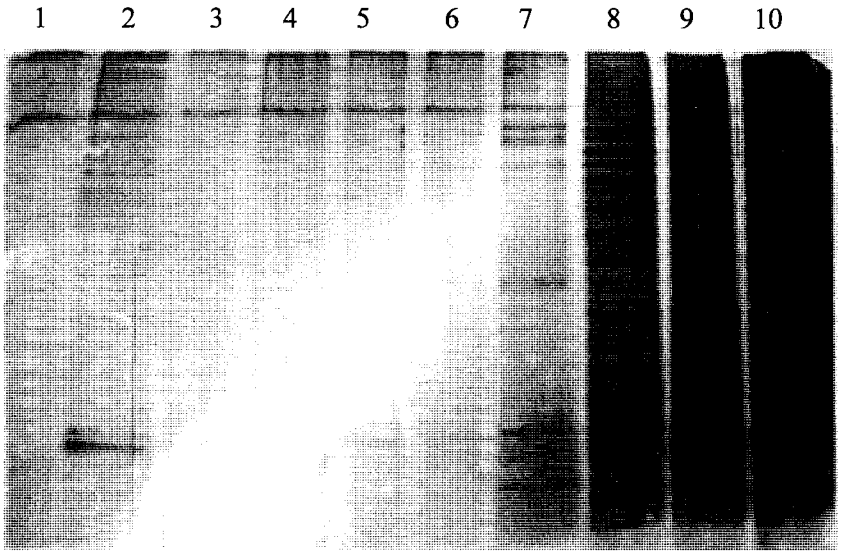


Figure 4.2 Gel image from SDS-PAGE analyses of washing solution (lanes 2-7) and elution fraction (lanes 8-10) from three cell samples. Nuclear extract of control cells BPDE treated cells, and BPDE+As treated cells were applied to three separate arsenic affinity column. The washing fraction and elution fractions were analyzed.

Lane 1: 2 ng BSA

Lane 2 and 3: the 5^{th} and 6^{th} washes from the control cell sample

Lane 4 and 5: the 5^{th} and 6^{th} washes from the BPDE treated cell sample

Lane 6 and 7: the 5^{th} and 6^{th} washes from the BPDE+ As^{III} treated cell sample

Lane 8: 50 mM DTT elution solution from the control cell sample

Lane 9: 50 mM DTT elution solution from the BPDE treated cell sample

Lane 10: 50 mM DTT elution solution from the BPDE+ As^{III} treated cell sample

Table 4.1 The captured and detected proteins that are common to three cell treatments (control cells, BPDE treated cells, and BPDE+ As^{III} treated cells)

protein	access No.	Mw	Peptide	cysteine position
①Actin, cytoplasmic 1	P60709	42kDa	EITALAPSTMK SYELPDGQVITIGNER	17, 217, 257, 272*,285*,374
①Actin, aortic smooth muscle	P62736	42kDa	QEYDESGPSIVHR GYSFTTTAER	2, 12, 19, 219, 259, 287*, 376
①Annexin A2	P07355	38kDa	SALSGHLETVILGLLK TNQELQEINR	9, 335*
③Anterior gradient protein 2 homolog precursor	O95994	20kDa	ADITGR IMFVDPSLTVR	81
③ATP-dependent RNA helicase A	Q08211	141kDa	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
①ATP-dependent DNA helicase 2 subunit 1	P12956	70kDa	ILELDQFK LGSLVDEFK	66*, 150, 389, 398, 585*
②DNA topoisomerase 2-alpha	P11388	174kDa	YGVFPLR YDTVLDILR	104*,170,216* 300, 392 , 405* 427, 455, 733, 862,997, 1008*, 1145*
③DNA-dependent protein kinase catalytic subunit	P78527	469kDa	AQEPESGLSEETQVK ILELSGSSSEDSEK	78 cysteines
②4F2 cell-surface antigen heavy chain	P08195	58kDa	EDFDSLLQSAK LLTSFLPAQLLR	330*

②Filamin-A	P21333	280kDa	ENGVYLIDVK AGGPGLER	41 cysteines
①Fructose-bisphosphate aldolase A	P04075	39kDa	GILAADESTGSIAK GILAADESTGSIAK	73*,135*,150, 178*,202,240* 290*,339
③Heterogeneous nuclear ribonucleoprotein A/B	Q99729	37Kda	FGEVVDCTIK GFVFITFK	97, 223
②Heterogeneous nuclear ribonucleoproteins A2/B1	P22626	37kDa	IDTIEIITDR GFGFVTFDDHDPVDK	50*
②Heterogeneous nuclear ribonucleoprotein A3	P51991	40kDa	GFAFVTFDDHDTVDK LFIGGLSFETTDDSLR	64*, 85*, 94*, 196
②Heterogeneous nuclear ribonucleoproteins C1/C2 ②Heterogeneous nuclear ribonucleoprotein C-like 1	P07910 O60812	34kDa 32kDa	GDDQLELIK SDVEAIFSK VDSLLENLEK SDVEAIFSK	46*
②Heterogeneous nuclear ribonucleoprotein D0	Q14103	39kDa	R.GFCFITFK.E R.EYFGGFGEVESIELPMDN K.T + Oxidation (M)	126, 226*, 252*
③Heterogeneous nuclear ribonucleoprotein K	P61978	51kDa	ILSISADIETIGEILK K.IILDLISESPIK.G	132, 145, 184, 185, 205
③Heterogeneous nuclear ribonucleoprotein L	P14866	60kDa	R.YYGGGSEGGR.A SDALETLGFLNHYQMK	120, 184, 187, 229, 230, 373, 421, 441, 490, 550
③Heterogeneous nuclear ribonucleoprotein R	O43390	71kDa	GFCFLEYEDHK ENILEEFSK	99, 214, 226, 240, 292
③Heterogeneous nuclear ribonucleoprotein Q	O60506	69kDa	AGPIWDLR GFCFLEYEDHK	96,211,237, 289
③Heterogeneous nuclear	Q00839	91kDa	GYFEYIEENK GNFTLPEVAECFDEITYVEL	287, 293, 333, 387, 389, 406,

ribonucleoprotein U			QK	448, 451, 495, 560, 592, 605, 646
①Histone H31T	Q16695	15kDa	VTIMPK+ Oxidation (M) EIAQDFK	97*, 111*
③Interleukin enhancer -binding factor 2	Q12905	43kDa	ILITTVPPNLR INNVIDNLIVAPGTFEVQIEE VR	37, 271, 291, 311
③Keratin, type II cytoskeletal 1	P04264	66kDa	TNAENEFVTIK LNDLEDALQQAK	48, 146, 496
③Keratin, type I cytoskeletal 9	P35527	62kDa	TLLDIDNTR DQIVDLTVGNNK	3, 61, 406, 432
③Keratin, type II cytoskeletal 2 epidermal	P35908	66kDa	NVQDAIADAEQR LALDVEIATYR	3, 7, 42, 423, 495
③Lamin-B1	P20700	66kDa	ALYETELADAR AGGPTTPLSPTR	110, 198, 317, 443, 583
②Lamin-A/C	P02545	74kDa	LADALQELR ITESEEVVSR	522*, 570, 588 591, 661
③Leucine-rich repeat-containing protein 59	Q96AG 4	35Kda	LVTLPVSFAQLK DKLDGNELDLSLSDLNEVP VK	48, 59, 131, 140, 266, 277
□Non-POU domain-containing octamer-binding protein	Q15233	54kDa	NLPQYVSNELLEEAFSVFGQ VER AVVIVDDR	145*,208
③Nucleolar RNA helicase 2	Q9NR30	89kDa	STYEQVDLIGK IGVPSATEIIK	161, 291, 378, 418, 445, 537, 643, 682
②Nucleophosmin	P06748	32kDa	DELHIVEAEAMNYEGSPIK + Oxidation (M) MTDQEAIQDLWQWR	21, 104*, 275
②Peroxiredoxin-1	Q06830	22kDa	TIAQDYGVLK QITVNDLPVGR	71, 83*
②Polypyrimidine tract-binding protein 1	P26599	57kDa	KLPIDVTEGEVISLGLPFGK LPIDVTEGEVISLGLPFGK	23, 250*, 251
③Probable	Q92841	72kDa	GDGPICLVLAPTR LIQLMEEIMAEK + 2	87, 120, 168, 191, 198, 219,

ATP-dependent RNA helicase DDX17			Oxidation (M)	240, 318, 352, 368, 505
①40S ribosomal protein S14	P62263	16kDa	IEDVTPIPSDSTR GEATVSFDDPPSAK	30, 53*, 84
②Splicing factor, arginine/serine-rich 3	P84103	20kDa	NPPGFAFVEFEDPR DAADAVR	6*, 10, 72*, 74
②U2 small nuclear ribonucleoprotein A'	P09661	29kDa	ELDLRGYK ERQEAEK	77*, 89
②Vimentin	P08670	53kDa	FLEQQNK LLEGEESR	327*

Cys in Filamin-A are 205, 210, 444, 478, 543, 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

Cys 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

Table 4.2 The proteins captured and detected from the control cells but not from the BPDE or BPDE+ As^{III} treated cells

protein	access No.	Mw	Peptide	cysteine position
①ADP/ATP translocase 1	P12235	33kDa	IFKSDGLR NVHIFVSWMIAQSVTAVAG LVSYP FDTVR	57, 128*, 256*
①ADP/ATP translocase 3	P12236	33kDa	GIVDCIVR SGT EREFR	56, 128*, 159* 256*
③Dynein heavy chain, cytoplasmic	Q14204	532kDa	GTFDNAETK ETVDQVEELR	36 cysteines
②Heat shock cognate 71 kDa protein	P11142	71Kda	HWPFMVVNDAGR STAGDTHLGGEDFDNR	17, 267*, 603
②Matrin-3	P43243	95kDa	TEEGPTLSYGR YQLLQLVEPFGVISNHLILN K	230, 293, 296, 552*, 563*,
③Phosphate carrier protein	Q00325	41Kda	VLYSNMLGEENTYLW R DAAPK	57, 68, 76, 91, 136, 237, 257, 277
②Poly(rC)-binding protein 1	Q15365	37kDa	INISEGNCPER IANPVEGSSGR	54*, 109, 118, 158, 194, 201, 293*, 355
③Probable ATP-dependent RNA helicase DDX5	P17844	69kDa	TTYLVLDEADR GDGPICLVLAPTR	89, 170, 191, 194, 200, 221, 234, 320, 354
①Small nuclear ribonucleoprotein Sm D2	P62316	13kDa	SEMTPEELQK NNTQVLINCR	46*, 63*
③Spectrin alpha chain	Q13813	284Kda	DVTGAEALLER AALLELWELR	158, 315, 466, 956,1091,1314 1444,1454, 1622,1930, 2120,2233,

				2351, 2441
②Spectrin beta chain, brain 1	Q01082	274kDa	LTTLELLEVR DGLNEAWADLLELIDTR	73, 112, 183, 604, 619, 624, 861, 964, 1284, 1389, 1900, 1970, 2227, 2262*

Dynein heavy chain, cytoplasmic : 220 , 416 , 867 , 1059 , 1484 , 1888 , 1932 , 1956 , 1977 , 1999 , 2076 , 2186 , 2659 , 2454 , 2466 , 2594 , 2639 , 2663 , 2712 , 3147 , 3325 , 3389 , 3507 , 3693 , 3712 , 3808 , 3940 , 4044 , 4121 , 4170 , 4216 , 4510 , 4540 , 4556 , 4570 , 4644

Table 4.3 The proteins captured and detected from the BPDE treated cells but not from the control cells or the BPDE+ As^{III} treated cells

protein	access No.	Mw	Peptide	cysteine position
③DNA repair protein RAD50	Q92878	154kDa	TVQQVNQEK VFQTEAELQEVISDLQSK	48, 53, 102, 133, 157, 221, 325, 633, 680, 681, 684, 788, 990,1201, 1222,1225, 1296, 1302
③Dynein heavy chain, cytoplasmic	Q14204	532kDa	GTFDNAETK ETVDQVEELR	78 cysteines
②rRNA 2'-O-methyltransferase fibrillarin	P22087	33kDa	VSISEGDDK TNIIPVIEDAR	99*, 268*

Table 4.4 The proteins captured and detected from the BPDE+ As^{III} treated cells but not from the control cells or the BPDE treated cells

protein	access No.	Mw	Peptide	cysteine position
③ELAV-like protein 1	Q15717	36kDa	SLFSSIGEVESAK VLVDQTTGLSR	13, 245, 284
③Probable ATP-dependent RNA helicase DDX5	P17844	69kDa	TTYLVLDEADR GDGPICLVLAPTR	89, 170, 191, 194, 200, 221, 234, 320, 354
①Small nuclear ribonucleoprotein Sm D2	P62316	13kDa	SEMTPEELQK GDSVIVVLR	46*, 63*

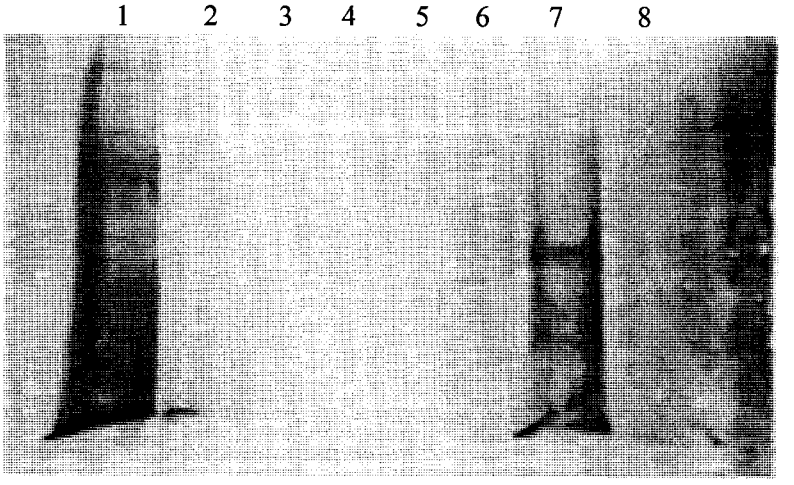


Figure 4.3 Gel image from the SDS-PAGE analyses of second affinity capture sample. The sample was from the control cells that have already gone through the first arsenic affinity column. The unbound (washing) fraction from the first affinity column was applied to the second arsenic affinity column. From this second affinity column, the 1% SDS washing solution (lanes 1-6), 50mM DTT elution (lane 7), and 100mM DTT elution (lane 8) were analyzed here.

3.2 Second affinity capture

In order to test whether any arsenic binding proteins may not be completely captured by a single arsenic affinity column, the washing solutions from the affinity column were collected and applied to a second arsenic affinity column. Figure 4.3 shows the gel image from the SDS-PAGE analyses of the washing and elution solutions from the second affinity column. By comparing figure 4.2 with figure 4.3, it is apparent that the captured proteins on the second affinity column (figure 4.3) is much less than the proteins captured on the first affinity column (figure 4.4). Table 4.5 shows the protein identification results from HPLC-ESI-MS/MS. All these proteins captured on the second affinity column are already captured and identified by using the first affinity column (Table 4.1 and 4.2). No new proteins were captured by using the second affinity column. Thus, the second affinity column is not necessary to captured different proteins; it only served to capture more amount of the same proteins.

3.3 Western blot

Mass Spectrometry can identify many arsenic-binding proteins simultaneously, but its detection limit is about nano-gram. It is difficult to detect the low abundance proteins. The detection limit of Western Blot is about picogram. Therefore, by combining arsenic affinity capture with Western blot analysis, it is possible to detect lower levels of arsenic-binding proteins.

To demonstrate the proof of principle, we choose to test the detection of three proteins, p53, MDM2, and PARP-1. Antibodies to these proteins are available to us. Nuclear proteins from the control cells, BPDE treated cells and the BPDE+As treated cells were analyzed. Nuclear protein extract from BPDE treated cells was also applied to a non-affinity column (chapter 3), which served as non-specific binding control.

Figure 4.4 shows the Western blot results of detecting p53. Actin loading control shows that the amount of proteins loaded was similar for the three samples. The level of

p53 in the BPDE treated cells was higher than in the control cells. This is understandable because DNA damage caused by BPDE could induce the expression of p53. When the cells were treated with BPDE followed by arsenite treatment, the level of p53 returned to normal. This observation is consistent with those of Shen et al[5], who found that arsenite inhibited the induction of p53 levels caused by BPDE. After the samples were applied to the arsenic affinity column to capture the arsenic binding proteins, p53 could be detected in the BPDE treated cell sample. This indicates that p53 in the cell sample can bind to arsenic. However, p53 was not captured on the affinity column from the control or BPDE+As treated cells. The reason may be because the binding affinity of p53 to PAO is low (Chapter 2), and p53 at low concentrations was not easy to be captured by affinity column. Therefore, only the p53 at high concentration (BPDE treated cell sample) could be captured by affinity column. The BPDE treated cell extraction sample was also applied to a non-affinity column (Chapter 3). No p53 was detected in the non-affinity control, which confirms that the p53 detected in BPDE treated cell sample is due to specific binding to arsenic.

Figure 4.5 shows the Western Blot results of MDM2. The level of MDM2 was higher in the control cells than in the BPDE or BPDE+As^{III} treated cells, whereas the p53 level was lower in the control cells than the treated cells (figure 4.4). MDM2 is the major regulator of p53, and it can trigger the degradation of p53 by the ubiquitin system. The control sample had a higher MDM2 expression (figure 4.5), which matches the lower p53 level (figure 4.4). BPDE may play a role in MDM2 expression because MDM2 expression was lower in both BPDE and BPDE+As^{III} treated cells. BPDE sample and BPDE+As^{III} sample had similar MDM2 expression, but the p53 expression level was lower in BPDE+As^{III} treated cells than that in BPDE treated cells, which suggests that the regulation of MDM2 is not the only factor to affect the difference in p53 expression between the BPDE treated cells and the BPDE+As^{III} treated cells. Arsenite may suppress the p53 expression, but not the MDM2 expression. Further research is needed to

understand the effect of As on p53 and MDM2. After the cell samples were applied to the arsenic affinity column to capture arsenic-binding proteins, no MDM2 was detected from any of the three cell samples, suggesting that MDM2 did not bind to arsenic (PAO).

Figure 4.6 shows the Western blot results of PARP-1. Cells under three treatments expressed similar, high levels of PARP-1. After the samples were applied to the arsenic affinity columns, a small amount of PARP-1 was captured by the arsenic affinity column. This result is consistent with the finding of PARP-1 interaction with arsenic (chapter 2), where weak binding affinity was observed between PARP-1 and trivalent arsenicals (PAO).

4 . Conclusion

More than 50 proteins were captured and detected from nuclear extract of A549 control cells and the cells treated with BPDE or BPDE+As^{III}. Small difference was observed in the number of captured proteins between the control cells and treated cells. The selection from the second affinity column was not necessary for capturing new proteins that might be missed by the first affinity column. P53 and PARP-1 were captured by arsenic affinity column and identified by Western Blot analysis. Arsenite and BPDE had no effect on the expression level of PARP-1, but arsenite suppressed the p53 expression in BPDE treated cells. MDM2 was not captured by affinity column, judging from the Western Blot analysis.

Table 4.5. The proteins captured by the second affinity column from control cell sample

protein	Access No.	Peptide
Actin, cytoplasmic 1	P60709	GYSFTTTAER SYELPDGQVITIGNER
Annexin A2	P07355	AYTNFDAER TPAQYDASELK
Heterogeneous nuclear ribonucleoproteins A2/B1	P22626	QEMQEVQSSR GGNFGFGDSR
Heterogeneous nuclear ribonucleoproteins C1/C2	P07910	GDDLQAIK VDSLLENLEK
Keratin, type II cytoskeletal 1	P04264	YEDEINK AEAESLYQSK
Keratin, type I cytoskeletal 10	P13645	LASYLDK SQYEQLAEQNR
Keratin, type I cytoskeletal 9	P35527	SGGGGGGGLGSGGSIR FSSSSGYGGGSSR
Lamin-A/C (70 kDa lamin)	P02545	EGDLIAAQAR EDLQELNDR
Lamin-B1	P20700	EYEAALNSK ESDLNGAQIK
Keratin, type II cytoskeletal 7	P08729	AEAEAWYQTK LEAAIAEAEER
Leucine-rich repeat-containing protein 59	Q96AG4	LVTLPVSFAQLK DKLDGNELDLSLSDLNEVPVK
Nucleophosmin	P06748	TVSLGAGAK GPSSVEDIK
Spectrin alpha chain	Q13813	DLTGVQNLR DLIGVQNLLK
Vimentin	P08670	QVDQLTNDK QDVDNASLAR

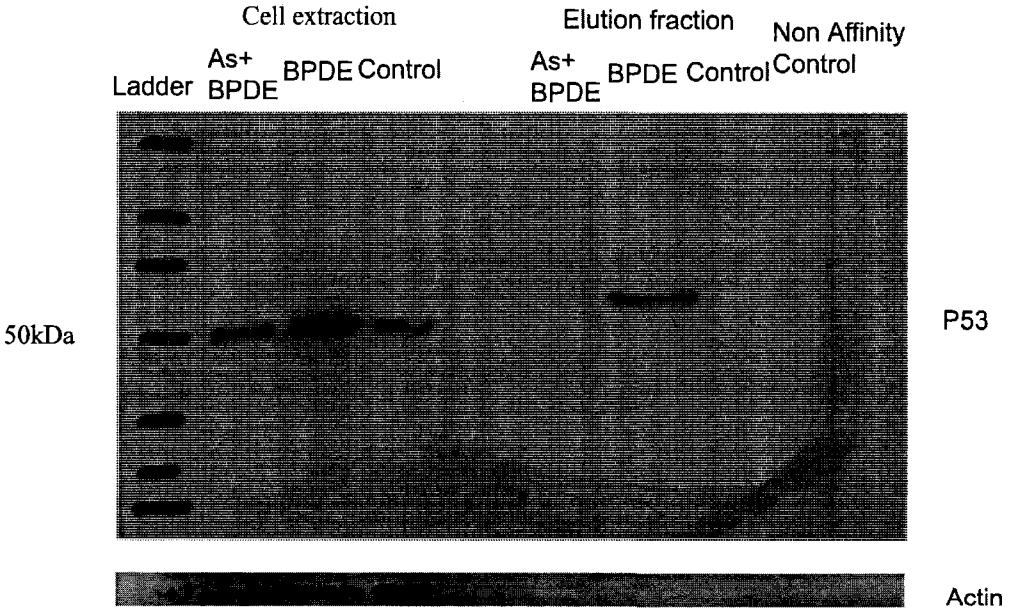


Figure4.4 Western Blot results of p53 and actin. Control A549 cells, A549 cells incubated with 1 μM BPDE for 0.5 hr and A549 cells incubated with 1 μM BPDE for 0.5 hr followed by 1 μM arsenite for 12 hr were analyzed in parallel. Cell nucleus extracts were either directly analyzed by Western blotting (left lanes labeled 2, 3 and 4), or applied to arsenic affinity column and the DTT elution fraction analyzed by Western blotting (right lanes 5, 6 and 7). A control column without arsenic affinity served to test any non-specific binding (lane 8)

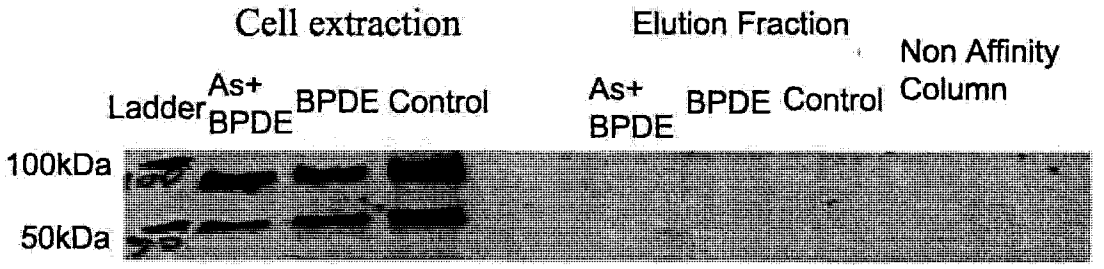


Figure 4.5 Western Blot results of protein MDM2. Control A549 cells, A549 cells incubated with 1 μM BPDE for 0.5 hr and A549 cells incubated with 1 μM BPDE for 0.5 hr followed by 1 μM arsenite for 12 hr were analyzed in parallel. Cell nucleus extracts were either directly analyzed by Western blotting (left lanes labeled 2, 3 and 4), or applied to arsenic affinity column and the DTT elution fraction analyzed by Western blotting (right lanes 5, 6 and 7). A control column without arsenic affinity served to test any non-specific binding (lane 8)

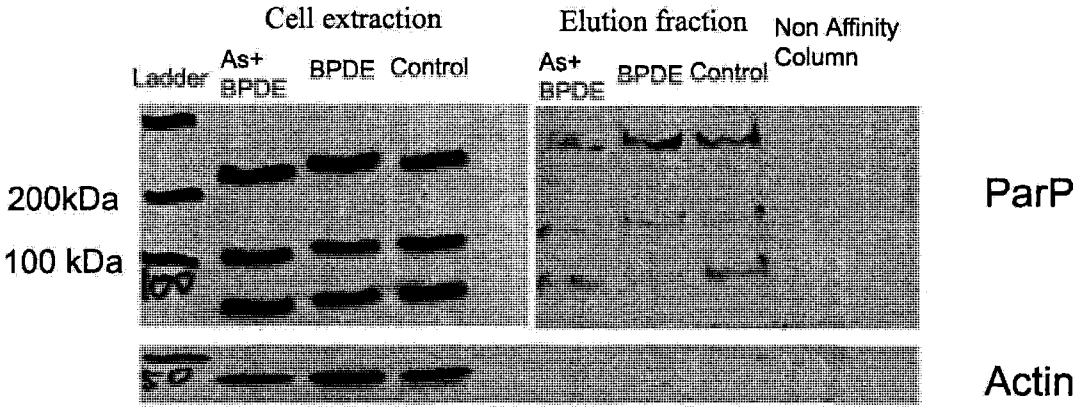


Figure 4.6 Western Blot results of protein PARP-1. Control A549 cells, A549 cells incubated with 1 μM BPDE for 0.5hr and A549 cells incubated with 1 μM BPDE for 0.5 hr followed by 1 μM arsenite for 12 hr were analyzed in parallel. Cell nucleus extracts were either directly analyzed by Western blotting (left lanes labeled 2, 3 and 4), or applied to arsenic affinity column and the DTT elution fraction analyzed by Western blotting (right lanes 5, 6 and 7). A control column without arsenic affinity served to test any non-specific binding (lane 8)

5 . References

1. Maier, A.; Schumann, B. L.; Chang, X.; Talaska, G; Puga, A., Arsenic co-exposure potentiates benzo[a]pyrene genotoxicity. Mutat Res 2002, 517, (1-2), 101-11.
2. Fischer, J. M.; Robbins, S. B.; Al-Zoughool, M.; Kannamkumarath, S. S.; Stringer, S. L.; Larson, J. S.; Caruso, J. A.; Talaska, G; Stambrook, P. J.; Stringer, J. R., Co-mutagenic activity of arsenic and benzo[a]pyrene in mouse skin. Mutat Res 2005, 588, (1), 35-46.
3. Evans, C. D.; LaDow, K.; Schumann, B. L.; Savage, R. E., Jr.; Caruso, J.; Vonderheide, A.; Succop, P.; Talaska, G., Effect of arsenic on benzo[a]pyrene DNA adduct levels in mouse skin and lung. Carcinogenesis 2004, 25, (4), 493-7.
4. Schwerdtle, T.; Walter, I.; Hartwig, A., Arsenite and its biomethylated metabolites interfere with the formation and repair of sTable BPDE-induced DNA adducts in human cells and impair XPAzf and Fpg. DNA Repair (Amst) 2003, 2, (12), 1449-63.
5. Shen, S.; Lee, J.; Sun, X.; Wang, H.; Weinfeld, M.; Le, X. C., Elevation of cellular BPDE uptake by human cells: a possible factor contributing to co-carcinogenicity by arsenite. Environ Health Perspect 2006, 114, (12), 1832-7.

Chapter 5

Preparation and testing of arsenic affinity column for liquid chromatography separation of Arsenic Binding Proteins

1. Introduction

Chapter 3 and Chapter 4 have demonstrated the preparation of the low pressure affinity column and its application to capture arsenic binding proteins from cell extract. Here we attempt to extend its application to separate the arsenic binding proteins according to their arsenic affinity by continuously pumping DTT solution into LC affinity column.

2. Experimental

2.1 Preparation of LC affinity column

NPAO was immobilized on beads by following the steps in chapter 3. The affinity beads were packed into a 4.6×150 mm stainless steel column until the column pressure was constant at 75 bar at a flow rate 1 mL/min in mobile phase 50%ACN/50%H_2O.

2.2 Selection of arsenic binding proteins by LC affinity column

The LC affinity column was pre-equilibrated for 30min by buffer (20 mM Hepes, 50 mM NaCl, pH 7.4) at a flow rate of 1mL/min. Then 1mL protein extraction fraction was injected to the LC affinity column by syringe pump at a flow rate of 0.1 mL/min, and incubated 2 hours at room temperature. The affinity LC column was connected to Agilent 1100 HPLC, rinsed by buffer and then the capture proteins were eluted by DTT buffer.

2.3 SDS-PAGE

Each washing fraction (4 mL) and elution fraction (2 mL) was collected and

concentrated by freeze-drying. Then the dried washing fraction sample was dissolved in 400 μL deionized water, and the elution sample was dissolved in 200 μL deionized water. The samples were used for gel electrophoresis as described in chapter 3.

2.4 Purification and identification of captured proteins

The elution sample was precipitated by TCA, followed by the standard reduction and alkylation procedures, and in-solution digestion. The digested solution was injected to RP-HPLC connected to Mass Spectrometry for protein identification described in chapter 3.

3. Results

3.1 Separation of arsenic-binding proteins by affinity column

Initially, our attempt was to separate various proteins by using arsenic affinity column. The rationale was to make use of the differences in affinity between proteins and arsenic on the column to achieve separation. Unfortunately, the DTT in the elution buffer resulted in high background of UV absorption detection (λ=214 nm and 280 nm). Consequently, this approach was not pursued further.

3.2 Removal of non-specific binding proteins

First, we tried to use the gravity affinity column washing parameters (chapter 3) in LC affinity column to remove the non-specific binding proteins. The washing and elution conditions are shown in Table 5.1. Buffer A is 20 mM Hepes, 50 mM NaCl, pH 7.4, and buffer B is 1%SDS in Buffer A.

Carbonic anhydrase II (1 mL, 1.8 mg/mL) was injected to LC affinity column, incubated at room temperature for 2 hours. The washing and elution steps were followed as shown in Table 5.1. Six washing fractions and three DTT elution fractions were collected and concentrated for SDS-PAGE. Figure 5.1 shows the electrophoresis results.

After 40 min wash, no protein band was detected in lane 6 and lane 7. The subsequent elution with DTT resulted in the release of carbonic anhydrase II protein (lane 8 and 9).

We then tried to use acidic isopropanol to remove possible hydrophobic interaction. Because the arsenic-sulfide bond is more stable under the acidic condition than the basic condition, the captured proteins would not be lost during washing steps. Table 5.2 shows the modified washing and elution conditions for testing carbonic anhydrase II using the LC affinity column.

Figure 5.2 shows the gel electrophoresis of carbonic anhydrase II sample. The gel was over developed. Lanes 2 to 6 show that after 1% SDS buffer and H_2O/isopropanol washes, no carbonic anhydrase II band was detected. Although some protein bands were detected in the DTT elution steps (lanes 7 and 8), their size were different from that of carbonic anhydrase II. The amount of detected proteins in lanes 7 and 8 was too low to be identified. Nonetheless, these results show that carbonic anhydrase II can be completely washed off from the column by using the conditions shown in Table 5.2. Carbonic anhydrase II has no free cysteine, and is not expected to be specifically bound to the arsenic affinity column.

The second protein used to test the removal of non specific binding protein from the LC affinity column was transferrin. It did not bind to NPAO, as shown in chapter 3. Transferrin solution (1 mL, 2 mg/mL) was injected into the LC affinity column by syringe pump. After 2 hours incubation at room temperature, the LC affinity column was washed and eluted by following the procedure in Table 5.2. Figure 5.3 shows the gel electrophoresis result of washing and elution solution. No transferrin protein band was detected in the DTT elution after the column was washed with 1% SDS and isopropanol/H_2O. Thus, the washing and elution procedure in Table 5.2 can wash off the non-specific binding components.

Table 5.1 Washing and elution conditions for the analysis of carbonic anhydrase II by the LC affinity column

Time (min)	Bufffer	Sampling	Flow rate (mL/min)
0-30	Buffer A		1
30-70	Buffer B	32-36 min, 36-40 min, 42-46 min, 46-50 min, 62-66 min, 66-70 min,	1
70-74	10mM DTT in Buffer B	70-74 min	0.5
74-78	50mM DTT in Buffer B	74-78 min	0.5
78-82	100mM DTT in Buffer B	78-82 min	0.5

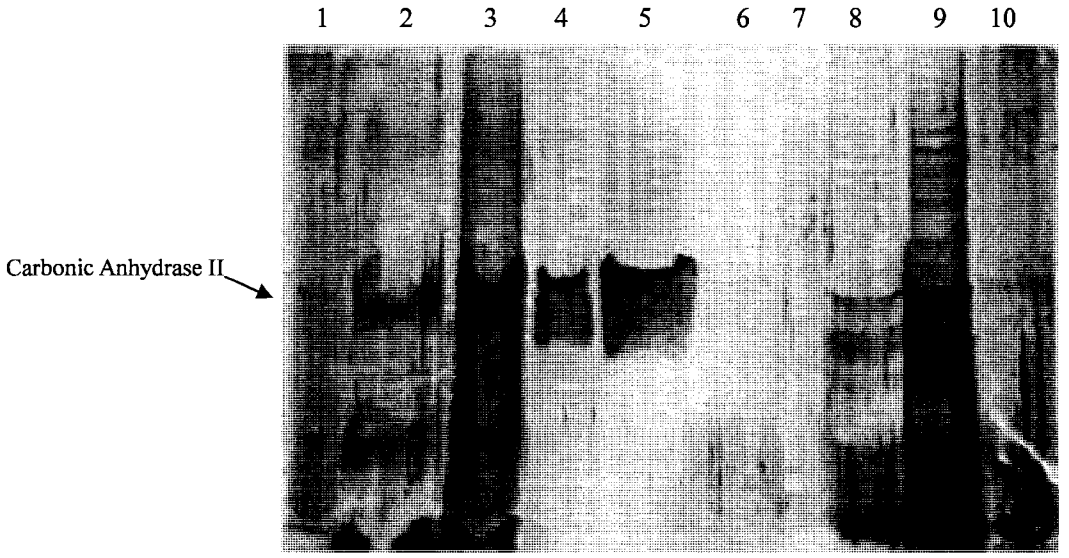


Figure 5.1. The image of gel electrophoresis of carbonic anhydrase II using the LC affinity column for separation (sampling time is shown in Table 5.1); lane 1: 2 ng carbonic anhydrase II; lanes 2 to lane 7: washing fractions from 30 to 70 minute; lane 8: elution with 10 mM DTT in 1%SDS; lane 9: elution with 50 mM DTT in 1% SDS; lane 10: elution with 100 mM DTT in 1% SDS

Table 5.2 Modified washing and elution conditions for carbonic anhydrase II

Time (min)	Bufffer	Sampling	Flow rate (mL/min)
0-30	Buffer A		1
30-70	Buffer B	42-46 min, 66-70 min,	1
70-80	(50% H_2O/50% isopropanol) with 0.1% acetate acid	76-80 min	1
80-90	(30% H_2O/70% isopropanol) with 0.1% acetate acid	86-90 min	1
90-110	(15% H_2O/85% isopropanol) with 0.1% acetate acid	96-100 min, 106-110 min	1
110-120	Buffer A		0.5
120-124	50 mM DTT in Buffer B	120-124 min	0.5
124-128	100 mM DTT in Buffer B	124-128 min	0.5

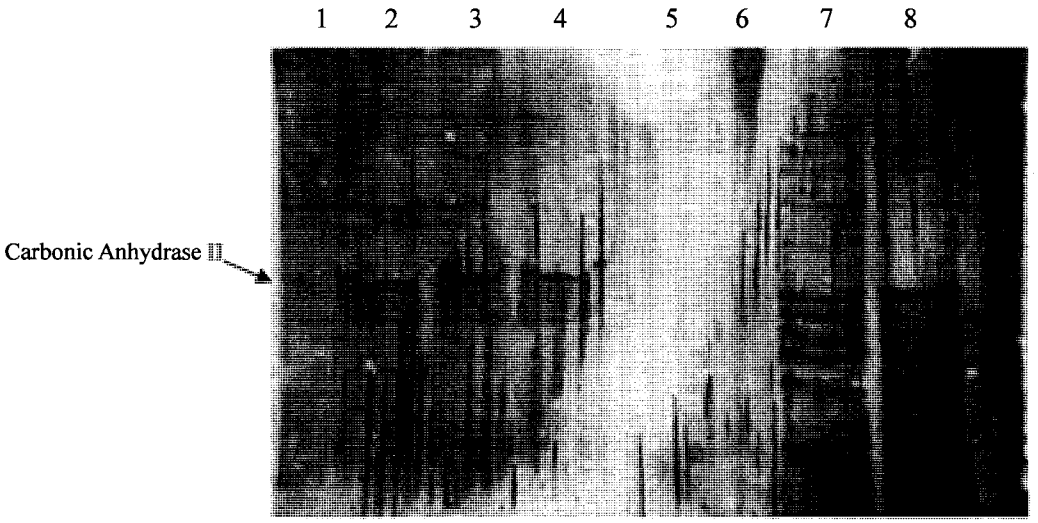


Figure 5.2. The image of gel electrophoresis of carbonic anhydrase II in LC affinity column (sampling time shown in Table 5.2); lane 1: 1 ng of carbonic anhydrase II; lanes 2 & 3: 1% SDS buffer washes from 30 to 70 minute; lanes 4 to 6: isopropanol/H_2O washes from 70 to 110 minute (no sampling for 96-100 min); lane 7: elution with 50 mM DTT in 1% SDS; lane 8: elution with 100 mM DTT in 1% SDS

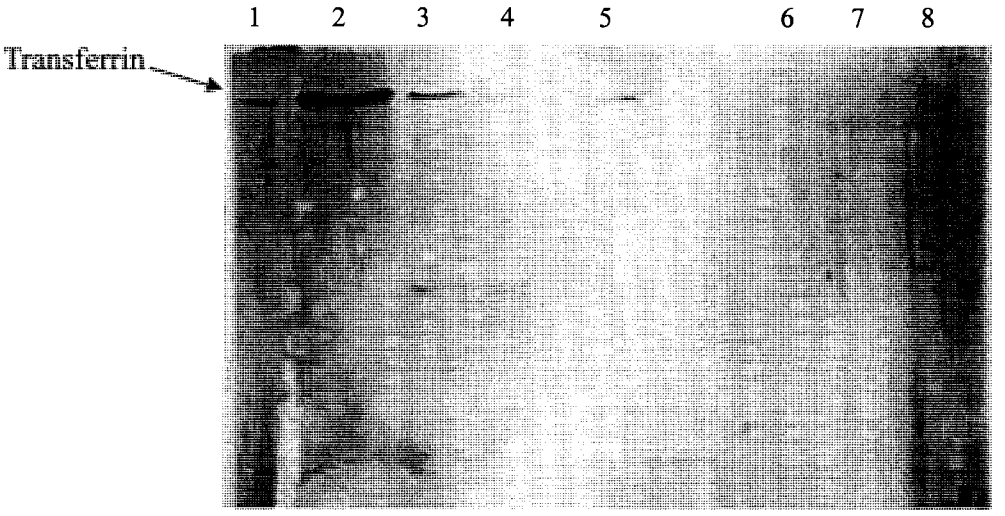


Figure 5.3. The image of gel electrophoresis of transferrin after separation using the LC affinity column (sampling time shown in Table 5.2); lane 1: 2 ng of transferrin; lanes 2 & 3: 1% SDS buffer washes from 30 to 70 minute; lanes 4 to 6: isopropanol/H_2O washes from 70 to 110 minute; lane 7: elution with 50 mM DTT in 1% SDS; lane 8: elution with 100 mM DTT in 1% SDS

3.3 Selection and identification of arsenic binding proteins by LC affinity column

Membrane extraction fraction (1 mL) from A549 cells was applied to the LC affinity column to separate the arsenic binding proteins by following the washing and elution steps listed in Table 5.2. Figure 5.4 shows the gel electrophoresis results of membrane extraction fraction of A549 cells after the LC affinity separation. 10 mM DTT solution could not elute the captured proteins from the LC affinity column. 50 mM DTT could elute some of the captured proteins, and 100 mM DTT led to the elution of more captured proteins. By comparing the 50 mM DTT elution (lane 9) and 100 mM DTT elution (lane 10) in figure 5.4, It seems that the 100 mM DTT can elute the captured proteins with larger molecular weight.

The three DTT elution fractions were purified for protein identification as described in chapter 3. The amount of proteins in 10 mM DTT elution fraction was too low to be detected. Table 5.3 and Table 5.4 show the identified proteins in 50 mM DTT elution and 100 mM DTT elution. The identified proteins in 50 mM DTT elution fraction are very similar to those in 100 mM DTT elution fraction. Identified proteins in the 100 mM DTT elution fraction included most of the identified proteins in 50 mM DTT elution fraction, except that a few proteins with larger molecular weight were also present in the 100 mM DTT elution..

We also applied the nuclear extraction fraction of A549 cells to the LC affinity column, and follow the procedure in Table 5.2 to wash the non specific binding proteins and to elute the captured proteins by DTT. To identify the captured proteins, we combined all the DTT elution fractions to obtain more amount of proteins for purification and identification. The identified captured proteins in nuclear extraction fraction of A549 cells are listed in Table 5.5.

The number of proteins identified after the separation by the LC affinity column was smaller than those identified after captured by the gravity affinity column shown in chapter 3. Probably the extra washing time and washing steps used in the LC affinity

approach have resulted in the removal of proteins that have low binding affinity for arsenic. It appears that the gravity affinity column is more efficient to capture arsenic binding proteins for subsequently analysis.

4. Conclusion

LC affinity column can be used to capture the specific arsenic binding proteins. But the LC affinity column needs extra washing steps to remove non specific binding proteins compared with the gravity affinity column, probably due to the uneven packing of the LC affinity column. The number of the identified proteins using the LC affinity column is fewer than that using the gravity affinity column.

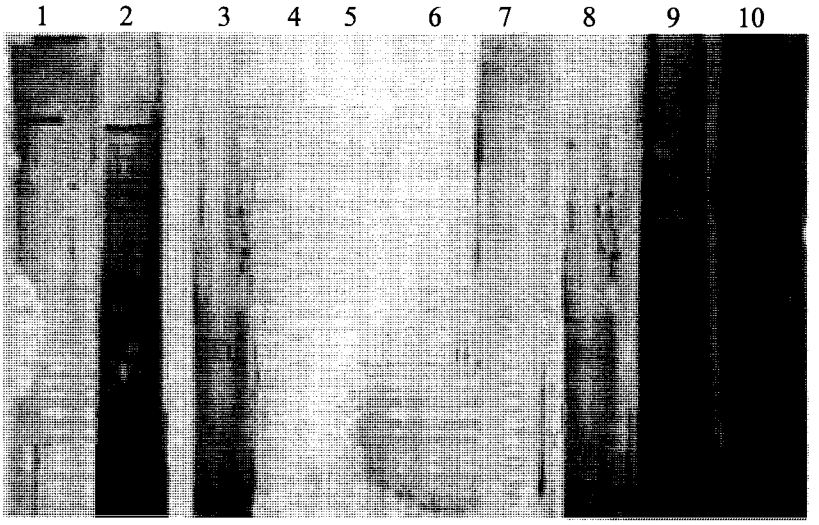


Figure 5.4. The image of gel electrophoresis of membrane extraction fraction after LC affinity separation (sampling time shown in Table 5.2); lane 1: 2 ng BSA; lanes 2 & 3: 1% SDS buffer washes from 30 to 70 minute; lanes 4 to 7: isopropanol/H_2O washes from 70 to 110 minute; lane 8: elution with 10 mM DTT in 1% SDS; lane 9: elution with 50 mM DTT in 1% SDS; lane 10: elution with 100 mM DTT in 1%SDS

Table 5.3 Proteins in membrane extraction fraction captured by the LC column and eluted by 50 mM DTT

Protein	access No.	Mw	Peptide	cysteine position
①Annexin A2	P07355	38KDa	TPAQYDASELK TNQELQEINR	9, 335*
①Elongation factor 2	Q05639	50kDa	TFCQLILDPIFK GEGQLGPAER	31*, 111, 326, 363, 370*, 411
□Elongation factor Tu	P49411	50kDa	LLDAVDTYIPVPAR DLEKPFLLPVEAVYSVP GR	40, 127, 147, 222*, 290*, 387*,
①Glucose-6-phosphate 1-dehydrogenase	p11413	59KDa	GGYFDEFGIIR DNIACVILTFK	13*, 158, 232, 269, 294*,358, 385, 446*,
③60 kDa heat shock protein	p10809	61KDa	GIIDPTK DDAMLLK	237, 442, 447
①3-hydroxyacyl-CoA dehydrogenase type-2	Q99714	27KDa	VMTIAPGLFGTPLLTSL PEK LGNNCVFAPADVTSEK	5, 58, 91*
③Keratin, type I cytoskeletal 9	p35527	62KDa	SDLEMQYETLQEELMA LK YCGQLQMIQEQISNLE AQITDVR + Oxidation (M)	3, 61, 406, 432
③Keratin, type I cytoskeletal 10	p13645	60KDa	VTMQNLNDR SQYEQLAEQNR	25, 66, 401, 427
③Keratin, type II cytoskeletal 1	p04264	66KDa	QISNLQQSISDAEQR THNLEPYFESFINNLR	48, 146, 496
②Nucleophosmin	P06748	32KDa	MTDQEAIQDLWQWR MSVQPTVSLGGFEITPP VVLR	21, 104*, 275
□Protein disulfide-isomerase A6 precursor	Q15084	48kDa	TGEAIVDAALSALR GSFSEQGINEFLR	11, 291, 292
□Ras-related protein Rab-2A	P61019	23kDa	TASNVEEAFINTAK GAAGALLVYDITR	21*, 211, 212
□Retinal dehydrogenase	P00352	55kDa	ANNTFYGLSAGVFTK	50*, 126*, 133*,

1			YCAGWADK	163, 186, 276, 302, 303, 456*, 464*,
□40S ribosomal protein SA	P08865	33kDa	AIVAIENPADVSVISSR FAAATGATPIAGR	148, 163,
□Stomatin-like protein 2	Q9UJZ1	38kDa	ATVLESEGTR ILEPGLNILIPVLDR	167, 172,
□Stress-70 protein	p38646	74KDa	EQQIVIQSSGGLSK SDIGEVILVGGMTR	66, 317, 366, 487, 608
□Thioredoxin-dependent peroxide reductase	P30048		GTAVVNGEFK DYGVLLEGSGLALR.	35, 45, 60, 127

Table 5.4. Proteins in membrane extraction fraction captured by the LC affinity column and eluted by 100mM DTT

Protein	access No.	Mw	Peptide	cysteine position
①Annexin A2	P07355	38KDa	QDIAFAYQR DIISDTSGDFR	9, 335*
①Elongation factor 2	Q05639	50kDa	IGGIGTVPVGR VETGVLKPGMVVTFAP VNVTTEVK	31*, 111, 326, 363, 370*, 411
□Elongation factor Tu	P49411	50kDa	LLDAVDTYIPVPAR DLEKPFLLPVEAVYSVP GR	40, 127, 147, 222*, 290*, 387*,
①Glucose-6-phosphate 1-dehydrogenase	p11413	59KDa	GGYFDEFGIIR DNIACVILTFK	13*, 158, 232, 269, 294*,358, 385, 446*,
□78 kDa glucose-regulated protein precursor	P11021	72KDa	VMEHFIKLYK LIPRN TVVPTK	41, 420*
③60 kDa heat shock protein	p10809	61KDa	ISSIQSIVPALEIANAHR ALMLQGVDLLADAVAV TMGPK	237, 442, 447
□3-hydroxyacyl-CoA dehydrogenase type-2	Q99714	27KDa	DVQTALALAK TYNLK	5, 58, 91*
□Keratin, type I cytoskeletal 9	P35527	62KDa	QEYEQLIAK HGVQELEIELQSQLSK	3, 61, 406, 432
□Keratin, type I cytoskeletal 10	P13645	60KDa	QSLEASLAETEGR YENEVALR	25, 66, 401, 427
□Keratin, type II cytoskeletal 1	P04264	66KDa	LNDLEDALQQAK SLNNQFASFIDK	48, 146, 496
□Keratin, type II cytoskeletal 2 epidermal	P35908	66KDa	AQYEEIAQR NLDLDSIIAEVK	3, 7, 42, 423, 495
□Protein disulfide-isomerase A6 precursor	Q15084	48kDa	AATALKDVVK GESPVDYDGGR	11, 291, 292
□Ras-related protein Rab-2A	P61019	23kDa	TASNVEEAFINTAK GAAGALLVYDITR	21*, 211, 212
□Reticulon-4	Q9NQC3	130KDa	VAVEAPMR GPLPAAPPVAPER	424, 464, 559, 597, 699, 1101*
□40S ribosomal protein SA	P08865	33kDa	GEATVSFDDPPSAK GDNITLLQSVSN	148, 163,

□Sarcoplasmic/ endoplasmic reticulum calcium ATPase 2	P16615	115kDa	EWGSGSDTLR AMGVVVATGVNTEIGK	70*, 268*, 344, 349, 365, 377, 404, 417, 420, 447, 471, 498, 524, 560, 595, 613, 635, 669*, 674, 773, 841, 875*, 887*, 909, 937*, 997, 1005, 1010
□Stomatin-like protein 2	Q9UJZ1	38kDa	GSLLASGR ATVLESEGTR	167, 172,
□Stress-70 protein	p38646	74KDa	VQQTVQDLFGR DAGQISGLNVLR	66, 317, 366, 487, 608
□Thioredoxin-dependent peroxide reductase	P30048	28kDa	GTAVVNGEFK DYGVLLEGSGLALR.	35, 45, 60, 127

Table 5.5. Proteins in nucleus extraction fraction captured by the LC affinity column and eluted by DTT

Protein	access No.	Mw	Peptide	cysteine position
□Annexin A2	P07355	38KDa	GVDEVTIVNILTNR GLGTDEDSLIEIICSR	9, 335*
□Anterior gradient protein 2 homolog precursor	O95994	20KDa	HLSPDGQYVPR IMFVDPSLTVR.	81
□ATP synthase subunit O, mitochondrial [Precursor]	P48047	23KDa	VAASVLNPYVK LVRPPVQVYGIEGR	15, 141*
□S100 calcium-binding protein A10	P60930	11KDa	EFPGFLENQKDPLAVDK EFPGFLENQK	62*, 83
□Heterogeneous nuclear ribonucleoprotein A/B	Q99729	37KDa	LFIGGLSFETTDESLR IEVIEIMTDR	97, 223
□Heterogeneous nuclear ribonucleoproteins A2/B1	P22626	37KDa	LTDCVVMR GGNFGFGDSR	50*
□Histone H31T	Q16695	15KDa	STELLIR EIAQDFK	97*, 111*
□Keratin, type I cytoskeletal 9	P35527	62KDa	TLLDIDNTR QVLDNLTMEK+Oxidation (M)	3, 61, 406, 432
□Keratin, type I cytoskeletal 10	P13645	59KDa	ALEESNYELEGK VLDELTLTK	25, 66, 401, 427
□Keratin, type II cytoskeletal 1	P04264	66KDa	SLDLDSIIAEVK THNLEPYFESFINNLR	48, 146, 496
□Keratin, type II cytoskeletal 2 epidermal	P35908	66KDa	NVQDAIADAEQR FLEQQNQVLQTK	3, 7, 42, 423, 495
□Mitochondrial transcription factor A	A9QXC6	29KDa	SWEEQMIEVGR SAYNVYVAER	22, 25, 246
□Small nuclear ribonucleoprotein E	P62304	11KDa	GDNITLLQSVSN VMVQPINLIFR	46
□Splicing factor, arginine/serine-rich 3	P84103	19KDa	NPPGFAFVEFEDPR AFGYYGPLR.S	6*, 10, 72*, 74

Chapter 6

Summary and Concluding Remarks

Chronic exposure to high levels of arsenic results in a wide range of health effects, including cancers of the bladder, lung, and skin. It has been estimated that as high as several hundred million people around the world are exposed to arsenic in drinking water at concentrations higher than the World Health Organization guideline level of 10 μg/L. How arsenic causes the wide range of adverse health effects remains poorly understood, and an acceptable level is still debatable. One possible mechanism of action is through the binding of arsenic to proteins, resulting in changes in the conformation, function, and /or activity of the proteins. It has been shown that arsenic can suppress cellular repair of DNA damage. It is possible that arsenic binds to the DNA repair proteins and affects their functions. Thus, this thesis focuses on studies of arsenic interaction with proteins. Two aspects have been investigated, one was to study arsenic binding to purified proteins, and the other was to capture and identify proteins that bind to arsenic by developing an affinity technique.

Two DNA proteins, p53 and poly (ADP-ribose) polymerase (PARP-1), were chosen to study their binding to trivalent arsenicals (Chapter 2). These proteins are involved in nucleotide-excision and base-excision DNA repair. Size exclusion chromatography separation combined with ICPMS detection enabled the differentiation of the protein-bound arsenic from the unbound arsenic species. p53 was found to bind to PAO and MMA^{III}, but not to DMA^{III}. The amount of p53-PAO complex increased with the increase of PAO concentration and incubation time. PAO was also shown to bind to PARP-1. There was no evidence of MMA^{III} or DMA^{III} binding to PARP-1. These differences in binding affinity may be related to the accessibility of the various arsenic compounds to the cysteine groups in the proteins. Information on the protein conformation would be useful for further study.

To identify potential cellular proteins that can bind to arsenic compounds, an arsenic affinity capture technique was developed and the captured proteins were identified by mass spectrometry. Because the affinity ligand was phenylarsine oxide, the captured proteins represented those that bound to this arsenic compound. The main benefit of the technique was to capture all (or as many as possible) specific arsenic-binding proteins from a large pool of many proteins in cells. This technique allowed for the capture and mass spectrometry identification of more than 70 arsenic-binding proteins (Chapter 3). All these proteins contained at least one cysteine, and most contained multiple cysteines. Using the available structural information of the proteins, the possible cysteines binding to arsenic were proposed. Because not all the identified proteins have structural information available for examination, this research was not able to pinpoint particular site binding to arsenic.

The technique of arsenic affinity capture and mass spectrometry analysis was further applied to identify proteins in the nucleus fraction of control A549 cells, and A549 cells treated with either benzo[α]pyrene diol epoxide alone or in combination with arsenite. The captured and identified proteins from the cells of these three treatments were similar (in number and identity) (Chapter 4). There was no quantitative information on the expression levels of these proteins.

Making use of the differences in binding affinity of the various proteins for phenylarsine oxide, the arsenic affinity technique was also tested for chromatographic separation of proteins. However, due to the high background of DTT when using absorbance detection (214 nm), the eluted proteins could not be reliably detected. Other elution solutions would be necessary in future research for this affinity HPLC application.

This thesis resulted in the development of analytical techniques, including HPLC-ICPMS for studying protein interaction with arsenic, affinity capture of arsenic-binding proteins, and mass spectrometry identification of arsenic-binding proteins. These analytical techniques will enable further biochemical and toxicological studies to

improve understanding of arsenic health effects.

Future research may include identification of arsenic-binding proteins in other cells, study of the arsenic effects on the biological activities of these arsenic-binding proteins, and the structural understanding of the consequence of arsenic binding to proteins. Site directed mutagenesis may be used to study the arsenic binding site(s) on proteins based on the predicted possible arsenic binding site(s).