DESIGNING ARSENIC BASED INHIBITORS
OF REDOX-ACTIVE ENZYMES

by

Aparna Sapra

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemistry and Biochemistry

Spring 2014

© 2014 Aparna Sapra
All Rights Reserved
DESIGNING ARSENIC BASED INHIBITORS
OF REDOX-ACTIVE ENZYMES

by

Aparna Sapra

Approved:

Murray V. Johnston, Ph.D.
Chair of the Department of Chemistry and Biochemistry

Approved:

George H. Watson, Ph.D.
Dean of the College of Arts and Sciences

Approved:

James G. Richards, Ph.D.
Vice Provost for Graduate and Professional Education
I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

Colin Thorpe, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

Harold B. White III, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

Joseph M. Fox, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

Catherine B. Kirn-Safran, Ph.D.
Member of dissertation committee
ACKNOWLEDGMENTS

This thesis has developed with the support and contributions of many individuals and it is my pleasure to extend my appreciation and gratitude to everyone involved. First and foremost, is my advisor Dr. Colin Thorpe. His continuous support and guidance has helped me in completing this thesis and the challenging research behind it. I have benefited greatly from his valuable insights and discussions. I feel fortunate to have worked under his mentorship during this journey, which has been a great learning experience and a memorable one. His dedication towards research will continue to inspire me.

I would like to thank my committee members Dr. Fox, Dr. White and Dr. Kirn-Safran for their time. Their advice and feedback have enriched this research work as well as my experience.

Being a part of the Department of Chemistry and Biochemistry has been a wonderful experience and I would like to extend my gratitude for giving me this opportunity.

I would like to thank the past and present members of the Thorpe Laboratory for their help and support throughout my PhD. Sharing this journey with them has been a truly enjoyable experience. I have made some wonderful friends here and the memories of the time spent with them will always remain with me.

I am grateful for the constant love and support of my parents, my sister Anupriya and my brother-in-law Nikhil. Finally, I would like to thank my best friend
and my husband Piyush. Without his continuous support this work would not have been possible.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>xi</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>xii</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>xv</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>xviii</td>
</tr>
</tbody>
</table>

Chapter

1  ARSENIC IN BIOLOGICAL SYSTEMS ................................................................. 1
   1.1  Arsenic Exposure ................................................................. 1
   1.2  Arsenic in Nature ............................................................... 3
   1.3  Arsenic Metabolism ............................................................. 5
   1.4  Cellular Targets of Arsenic ................................................ 6
   1.5  Arsenic as a Carcinogen ....................................................... 9
   1.6  Arsenic as a Therapeutic ..................................................... 11
      1.6.1  Salvarsan ................................................................. 11
      1.6.2  Melarsoprol ............................................................... 12
      1.6.3  Arsenic Trioxide and Acute Promyelocytic Leukemia .... 12
   1.7  New Arsenic Drugs .............................................................. 14
      1.7.1  Darinaparsin (S-dimethylarsino-glutathione) ................. 15
      1.7.2  GSAO (4-(N-(S-glutathionylacetyl) amino) phenylarsonous acid) .................................. 15
      1.7.3  Arsenic Nanoparticles ................................................. 16
   1.8  Arsenic Based Biotinylated Pull Down .................................... 17
   1.9  Arsenic Tomography ........................................................... 18
   1.10  Fluorescent Arsenic Molecules ............................................ 19

REFERENCES ............................................................................................................. 23

2  OXIDATIVE PROTEIN FOLDING ................................................................ 38
   2.1  Disulfide Bond Formation .................................................... 38
   2.2  Oxidative Protein Folding ..................................................... 41
   2.3  Oxidative Protein Folding in Prokaryotes .............................. 42
2.4 Oxidative Protein Folding in Eukaryotes ................................. 44
   2.4.1 PDI – First Model of Oxidative Protein Folding .................. 44
   2.4.1.1 Ero1 .............................................................................. 44
   2.4.1.2 Peroxiredoxin IV .......................................................... 45
   2.4.1.3 Glutathione Peroxidases ............................................... 46
   2.4.1.4 Vitamin K Epoxide Reductase ..................................... 46
   2.4.2 QSOX-PDI pathway – A Alternative Model for Oxidative
   Protein Folding ............................................................................ 49
2.5 Quiescin Sulfhydryl Oxidase (QSOX) ............................................. 50
   2.5.1 Discovery of QSOX .............................................................. 50
   2.5.2 Localization of QSOX .......................................................... 51
   2.5.3 Domain Structure of QSOX ................................................. 52
   2.5.4 Mechanism of QSOX ......................................................... 54
   2.5.5 Substrates of QSOX ........................................................... 57
   2.5.6 QSOX’ s Role in Cancer ...................................................... 59
2.6 Protein Disulfide Isomerase .......................................................... 60
   2.6.1 Discovery of PDI ................................................................. 60
   2.6.2 Localization of PDI ............................................................. 61
   2.6.3 Structure of PDI ................................................................. 62
   2.6.4 Functions of PDI ................................................................. 65
   2.6.5 PDI’ s Role in Cancer .......................................................... 67
REFERENCES .................................................................................. 68
3 INHIBITION OF OXIDATIVE PROTEIN FOLDING BY
MULTIVALENT ARSENICALS ............................................................... 82
   3.1 Introduction ............................................................................ 82
   3.2 Materials and Methods ......................................................... 86
   3.2.1 General Methods and Materials ................................. 86
   3.2.2 Synthesis of Arsenicals .................................................... 87
   3.2.2.1 Synthesis of PSAO ............................................. 87
   3.2.2.2 Scheme for Synthesis of 2,2’-(ethane-1,2-diylbis((2-
   ((4-arsinephenyl)amino)-2-oxoethyl)azanediyl))diacetic acid ....... 87
3.2.2.3 Scheme for Synthesis of 3',4-bis((4-arsinephenyl)carbamoyl)-[1,1'-biphenyl]-3,4'-dicarboxylic acid ......................................................... 88
3.2.2.4 Strategy for Synthesis of 1,3,5-((benzenetricarbonyltris(azanediyl)) tris(benzene-4,1-diyl))triarsineous acid ........................................... 89

3.2.3 Expression, Purification and Handling of Proteins ..................... 89
3.2.4 Preparation of Reduced Proteins ................................................. 90
3.2.5 Oxidative Protein Folding Assay using RfBP ............................ 91
3.2.6 Oxidative Protein Folding Assay using RNase ....................... 91
3.2.7 Turbidometric Insulin Reductase Assay .................................... 91
3.2.8 Binding of Arsenicals to Reduced PDI ...................................... 92
3.2.9 QSOX Activity Assay .............................................................. 92
3.2.10 COPASI Simulation ............................................................... 92
3.2.11 Thioflavin T Assay ................................................................. 93
3.2.12 Congo Red Spectroscopic Assay ............................................. 93
3.2.13 Transmission Electron Microscopy ......................................... 94

3.3 Results and Discussion.................................................................. 94

3.3.1 Effect of MVA's in Oxidative Protein Folding Assay using RfBP ................................................................. 96
3.3.2 Effect of Bis- and Tris-Arsenicals on the Oxidative Refolding of RNaseA ................................................................. 100
3.3.3 Bis- and Tris-Arsenicals as Inhibitors of QSOX ...................... 102
3.3.4 Interaction between Multivalent Arsenicals and PDI ............. 104
3.3.5 Interaction between Multivalent-Arsenicals and RfBP .......... 111
3.3.6 Fibril Formation with Reduced RNase and Multivalent-Arsenicals .................................................................................. 113

3.4 Discussion and Conclusion......................................................... 118

REFERENCES ......................................................................................... 121

4 AN ARSENICAL-MALEIMIDE FOR THE GENERATION OF NEW TARGETED BIOCHEMICAL REAGENTS ........................................ 125

4.1 Introduction ................................................................................. 125
4.2 Methods and Materials ............................................................. 127

4.2.1 General Methods and Reagents ............................................. 127
4.2.2 Synthesis of Arsenical Reagents ........................................... 127
4.2.2.1 Synthesis of Arsenical-Maleimide (As-Mal), 5 (4-(2, 5-dioxo-2, dihydro-1 H-pyrrol-1-yl) phenylarsonous acid) ........................................................................................................... 128

4.2.3 Stopped Flow Spectrophotometry .......................................................... 129
4.2.4 Mass Spectroscopic Analysis of Conjugated Proteins ......................... 129
4.2.5 Expression, Purification and Handling of Proteins .............................. 130
4.2.6 Preparation of Reduced Proteins and Labeling with Arsenical-Maleimide ................................................................. 130
4.2.7 Inhibition of Thioredoxin Reductase by the As-Mal Trx Conjugate ......................................................................................... 131
4.2.8 Inhibition of Protein Disulfide Isomerase by As-Mal RNase Peptides ................................................................. 131
4.2.9 $K_d$ determination of 1,8 bis-As-Mal-RNase with Reduced PDI ......................................................................................... 132

4.3 Results and Discussion ........................................................................... 132

4.3.1 Binding of As-Mal to Thiols ............................................................... 132
4.3.2 Inhibition of Thioredoxin Reductase ............................................... 135
4.3.3 Inhibition of Protein Disulfide Isomerase ......................................... 142

4.4 Conclusion .............................................................................................. 154

REFERENCES ........................................................................................................... 155

5 ARSENIC AFFINITY CHROMATOGRAPHY ............................................. 160

5.1 Introduction ............................................................................................. 160
5.2 A New Arsenic Sepharose Affinity Resin ............................................... 163
5.3 Materials and Methods ........................................................................... 163

5.3.1 General Methods and Materials .......................................................... 163
5.3.2 Preparation of As-Mal-Sepharose Resin ........................................... 164
5.3.3 Binding Capacity of As-Mal-Sepharose Resins ................................... 165
5.3.4 Mass Spectroscopic Analysis of Peptides .......................................... 165
5.3.5 Regeneration of Arsenical resins ........................................................ 166
5.3.6 Expression, Purification and Handling of Proteins ........................... 166
5.3.7 Purification of Disulfide-Containing Peptides from Yeast Extract ......................................................................................... 166
5.3.8 Analysis of As-Mal-Sepharose 4B with *Escherichia coli* Thioredoxin ......................................................................................... 168
5.3.9 Purification of Human Thioredoxin Reductase ................................ 168
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1.1</td>
<td>Representation for binding of As(III) species to proteins and small molecular weight thiols.</td>
<td>8</td>
</tr>
<tr>
<td>Table 2.1</td>
<td>Catalytic activity of QSOX</td>
<td>58</td>
</tr>
<tr>
<td>Table 5.1</td>
<td>Protein eluted from As-Mal-Sepharose 6B resin</td>
<td>172</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1.1 Arsenicals discussed in the text.................................................................2
Figure 1.2 Interaction of As(III) species with thiols....................................................4
Figure 1.3 Proposed mechanisms for arsenic biomethylation. ......................................5
Figure 1.4 Therapeutic arsenicals discussed in text....................................................10
Figure 1.5 Reaction scheme for labeling arsenic isotopes on antibody.........................18
Figure 1.6 Fluorescent arsenicals discussed in the text .............................................21
Figure 1.7 Fluorescent arsenicals discussed in the text (continued).............................22
Figure 2.1 Structure of some disulfide bonds-containing secreted proteins.................40
Figure 2.2 Oxidative protein folding.................................................................41
Figure 2.3 The Dsb system of oxidative protein folding in prokaryotes....................43
Figure 2.4 Eukaryotic pathways of PDI-first model of oxidative protein folding.........48
Figure 2.5 QSOX-PDI model for oxidative protein folding pathway .......................49
Figure 2.6 Domain structure of QSOX proteins......................................................52
Figure 2.7 Mechanism of QSOX.................................................................54
Figure 2.8 Crystal structure of *Trypanosoma brucei* QSOX.....................................56
Figure 2.9 Domain organization and crystal structure of PDI from
*Saccharomyces cerevisiae*. .............................................................................64
Figure 2.10 Reactions catalyzed by PDI.................................................................66
Figure 3.1 Structure of multivalent arsenicals..........................................................84
| Figure 3.2 | Design strategy for arsenic based inhibitors of QSOX and PDI. .......... 85 |
| Figure 3.3 | Synthesis of bis-arsenical-1 ..................................................................... 87 |
| Figure 3.4 | Synthesis of bis-arsenical-2 ..................................................................... 88 |
| Figure 3.5 | Synthesis of tri-arsenical-1 ...................................................................... 89 |
| Figure 3.6 | Quantification of MVAs using DTT titration ................................................. 95 |
| Figure 3.7 | Inhibition of oxidative folding of reduced RfBP in the presence of MVAs ..................................................................................................... 98 |
| Figure 3.8 | Comparison of monoarsenicals and MVAs as inhibitors of oxidative protein folding of RfBP ........................................................................... 99 |
| Figure 3.9 | Inhibition of oxidative refolding of RNase by arsenicals ...................... 101 |
| Figure 3.10 | Inhibition of QSOX reactivity by MVAs and monoarsenicals .............. 103 |
| Figure 3.11 | The effect of multivalent arsenicals on the reductase activity of PDI . . 105 |
| Figure 3.12 | MVA compounds do not significantly delay the onset of insulin B-chain aggregates ............................................................................................................................... 106 |
| Figure 3.13 | Binding stoichiometry of MVAs with reduced PDI .................................... 109 |
| Figure 3.14 | $K_d$ determination of reduced PDI with MVAs ......................................... 110 |
| Figure 3.15 | Titration of reduced RfBP with arsenicals .............................................. 112 |
| Figure 3.16 | Aggregation of reduced RNase with arsenicals followed with Thioflavin T ................................................................................................................................. 114 |
| Figure 3.17 | Congo red spectral shift assay with reduced RNase and MVAs .......... 115 |
| Figure 3.18 | TEM of multivalent arsenicals with reduced RNase ................................. 116 |
| Figure 3.19 | TEM of multivalent arsenicals with reduced RNase in presence of 5 mM GSH. ................................................................. 117 |
| Figure 3.20 | Interaction of MMA with dithiols in presence of GSH .............................. 119 |
| Figure 4.1 | Structures of 4-(N-(S-glutathionylacetyl)amino) phenylarsonous acid (GSAO) and arsenic-maleimide (As-Mal) ................................................................. 126 |
Figure 4.2  Synthesis of arsenical-maleimide (As-Mal) ............................................. 128
Figure 4.3  Titration of As-Mal with GSH .......................................................... 133
Figure 4.4  Reaction of N-phenylmaleimide and As-Mal with GSH ...................... 134
Figure 4.5  Kinetics of As-Mal conjugation to C35S thioredoxin ....................... 137
Figure 4.6  Mass spectroscopy of As-Mal Trx .................................................. 138
Figure 4.7  Quantification of As-Mal Trx ......................................................... 139
Figure 4.8  Mass spectrum of As-Mal-Trx after DTT treatment ......................... 140
Figure 4.9  Inhibition of thioredoxin reductase by arsenicals ............................. 141
Figure 4.10 Mass spectrum of (As-Mal)_8 RNase ............................................. 143
Figure 4.11 As-Mal conjugated RNase (As-Mal RNase) .................................. 144
Figure 4.12 Mass spectra of As-Mal RNase derivatives .................................... 145
Figure 4.13 Quantification of As-Mal RNase derivatives ................................... 146
Figure 4.14 Insulin reductase assay with selected arsenicals ............................... 149
Figure 4.15 Insulin reductase assay with cysteine RNase peptides ....................... 150
Figure 4.16 Insulin reductase assay in presence of 5 mM GSH ............................ 151
Figure 4.17 K_d determination of reduced PDI with 1, 8 bis-As-Mal RNase ....... 153
Figure 5.1  Preparation of As-Mal-Sepharose resins .......................................... 170
Figure 5.2  Thioredoxin binding and elution from As-Mal-Sepharose 4B resin ....... 174
Figure 5.3  Purification of human thioredoxin reductase by As-Mal-Sepharose 4B. ........................................................................................................ 175
Figure 5.4  Binding of RNase proteins with As-Mal-Sepharose 4B ................. 177
Figure 5.5  Regeneration of arsenic labeled resins ............................................. 178
Figure 5.6  Binding of reduced Trx with regenerated As-Mal-Sepharose 4B ....... 180
## LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALR</td>
<td>Augmenter of liver regeneration</td>
</tr>
<tr>
<td>APL</td>
<td>Acute promyelocytic leukemia</td>
</tr>
<tr>
<td>ATO</td>
<td>Arsenic trioxide</td>
</tr>
<tr>
<td>BA-1</td>
<td>Bis arsenical-1</td>
</tr>
<tr>
<td>BA-2</td>
<td>Bis arsenical-2</td>
</tr>
<tr>
<td>BAL</td>
<td>British anti-Lewisite</td>
</tr>
<tr>
<td>β-ME</td>
<td>β-Mercaptoethanol</td>
</tr>
<tr>
<td>CR</td>
<td>Clinical remission</td>
</tr>
<tr>
<td>CxxC</td>
<td>Cys-x-x-Cys motif, x is any other amino acid except cysteine</td>
</tr>
<tr>
<td>DMA (III)</td>
<td>Dimethylarsinous acid</td>
</tr>
<tr>
<td>DMA (V)</td>
<td>Dimethylarsinic acid</td>
</tr>
<tr>
<td>DTNB</td>
<td>5, 5'-dithiobis (2-nitrobenzoate)</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>ECM</td>
<td>Extracellular matrix</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>ER</td>
<td>Endoplasmic reticulum</td>
</tr>
<tr>
<td>Ero1</td>
<td>Endoplasmic reticulum oxidoreductin 1</td>
</tr>
<tr>
<td>ERV</td>
<td>Essential for growth and respiration</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>FAD</td>
<td>Flavin adenine dinucleotide</td>
</tr>
<tr>
<td>GuHCl</td>
<td>Guanidine hydrochloride</td>
</tr>
<tr>
<td>GSH</td>
<td>Reduced glutathione</td>
</tr>
<tr>
<td>GSSG</td>
<td>Oxidized glutathione</td>
</tr>
<tr>
<td>HRR</td>
<td>Helix rich region</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>$K_{cat}$</td>
<td>Catalytic rate constant</td>
</tr>
<tr>
<td>Km</td>
<td>Michaelis–Menten binding constant</td>
</tr>
<tr>
<td>LB</td>
<td>Lysogeny broth</td>
</tr>
<tr>
<td>MMA (III)</td>
<td>Monomethylarsonous acid</td>
</tr>
<tr>
<td>MMA (V)</td>
<td>Monomethylarsonic acid</td>
</tr>
<tr>
<td>MVAs</td>
<td>Multivalent arsenicals</td>
</tr>
<tr>
<td>NEM</td>
<td>N-Ethyl maleimide</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>NADPH</td>
<td>Nicotinamide adenine dinucleotide phosphate</td>
</tr>
<tr>
<td>PDI</td>
<td>Protein disulfide isomerase</td>
</tr>
<tr>
<td>PAO</td>
<td>Phenylarsine oxide</td>
</tr>
<tr>
<td>PSAO</td>
<td>p-Succinylamidephenyl arsenoxide</td>
</tr>
<tr>
<td>QSOX</td>
<td>Quiescin sulfhydryl oxidase</td>
</tr>
<tr>
<td>RfBP</td>
<td>Riboflavin binding protein</td>
</tr>
<tr>
<td>RNase</td>
<td>Ribonuclease A</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>ROS</td>
<td>Reactive oxygen species</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>TA-1</td>
<td>Tris arsenical-1</td>
</tr>
<tr>
<td>TCEP</td>
<td>Tris(2-carboxyethyl)phosphine</td>
</tr>
<tr>
<td>THP</td>
<td>Tris(hydroxypropyl)phosphine</td>
</tr>
<tr>
<td>TLC</td>
<td>Thin layer chromatography</td>
</tr>
<tr>
<td>Trx</td>
<td>Thioredoxin</td>
</tr>
<tr>
<td>TrxR</td>
<td>Thioredoxin reductase</td>
</tr>
</tbody>
</table>
ABSTRACT

The surprising efficacy of arsenic trioxide in the treatment of acute promyelocytic leukemia has renewed interest in the synthesis and testing of arsenicals as chemotherapeutic agents for other types of cancers. The biological effects of arsenicals largely reflect coordination of As(III) species to vicinal thiols. Several enzymes of oxidative protein folding are both up-regulated in certain cancer cells, and contain catalytically essential vicinal thiols. Here, I sought to develop arsenical-based inhibitors that would capture these redox-active motifs.

The work in this dissertation introduces synthesis of simple multivalent arsenicals that could take advantage of the chelate effect leading to more effective inhibition of enzymes with multiple CxxC motifs like quiescin sulphhydryl oxidase (QSOX) and protein disulfide isomerase (PDI). The results from this study showed that the small molecule multivalent arsenicals were able to inhibit the oxidative folding pathways of the reduced unfolded proteins, riboflavin binding protein (RfBP) and ribonuclease (RNase). These compounds were effective even in the presence of millimolar concentration of glutathione, which is an effective competitor for the arsenic functionality. However, the main targets of these small molecule arsenicals turned out to be the reduced unfolded protein substrates and not the redox-active enzymes.
Next, to increase the specificity towards the more structured CxxC motifs, an As(III)-containing maleimide (As-Mal) was synthesized which could be readily conjugated to exposed cysteine residues in peptides and proteins. The conjugation provided a scaffold for directing the As(III) species to the target proteins for their inhibition. Experiments with thioredoxin reductase and protein disulfide isomerase show that their CxxC motifs can be efficiently captured by their cognate arsenical-carrying substrates. These peptide-based arsenicals as well as small molecule multivalent arsenicals could be used to inhibit cell-surface redox-active enzymes where the extracellular concentration of reduced GSH is in low micromolar range.

The As-Mal reagent was also used to generate two new arsenic-labeled affinity purification resins by reaction with thiol-activated Sepharose beads. These resins were subsequently used for purifying thiol-containing peptides from yeast extracts and proteins containing redox active CxxC motifs. A new technique for regenerating these resins was also developed that makes this affinity purification method highly cost effective.
Chapter 1
ARSENIC IN BIOLOGICAL SYSTEMS

1.1 Arsenic Exposure

Arsenic (As) is a metalloid found in the earth crust, mainly as sulfide ores such as orpiment (As₂S₃) or realgar (As₂S₂) [1]. It is regarded as a poison by the World Health Organization (WHO) and its contamination in our food system comes mainly from drinking water. The range of arsenic in fresh water could range from 1-50 μg/L [1]. The WHO has recommended the maximum contamination limit of arsenic in drinking water to be 10 μg/L but in some parts of Bangladesh, Taiwan and eastern India it can reach up to 50 μg/L. The most common symptoms for chronic exposure to arsenic are skin lesions, pigmentation and thickening of skin known as hyperkeratosis. Arsenic exposure can also lead to various types of cancers including skin, lung, bladder and kidney [2]. A severe form of peripheral vascular disease known as blackfoot disease is found in the population of the southwest coastal region of Taiwan, caused by exposure to high levels of arsenic in artesian well water [3].
Figure 1.1  Arsenicals discussed in the text.
1.2 Arsenic in Nature

Arsenic resides below Phosphorus in group 15 of the periodic table. It has two oxidation states, As(V) and As(III) whose oxyanions are known as arsenate and arsenite respectively; both species impart toxicity by different mechanism.

Arsenate is structurally similar to inorganic phosphate which leads to the replacement of P(V) by As(V) in several pathways especially phosphorylation reactions involving phosphate esters [4]. One aqueous form of arsenic(V) oxide is arsenic acid (H$_3$AsO$_4$) and under physiological conditions it may exists as H$_3$AsO$_4$ ≈ H$_2$AsO$_4$⁻ ≈ HAsO$_4$²⁻ (pK$_{a1}$ = 2.19, pK$_{a2}$ = 6.94 and pK$_{a3}$ = 11.5) [5].

Phosphate has similarly charged oxygen atoms and almost identical pKa values hence the analogous esters of As(V) are able to replace P(V) in cases of arsenic exposure. Their structural and chemical similarities also lead to the uptake of arsenate inside the cell via phosphate transporters and phosphate ion pumps [6–8]. As(V) esters are unstable at neutral pH and hydrolyze rapidly contributing towards toxicity [9].

However the toxic effects of arsenic are believed to be mainly due to the interaction of As(III) species with sulphhydryl groups. For example arsenous acid (Figure 1.1) can form stable covalent bond with sulfur atoms of the cysteine residues in proteins as shown in Figure 1.2. The resulting As-S bond is more resistant to hydrolysis than an As-O bond. Due to the chelate effect As(III) can bind tightly to proteins containing thiols in close proximity which allows it to form stable ring
structures. As(V) can be reduced to As(III) inside the cells or arsenic trioxide (As$_2$O$_3$) can be directly transported into the mammalian cells by aquaglyceroporins [10][11]. This transporter belongs to the aquaporin superfamily that transports neutral solutes like glycerol and urea inside the cell [12][13]. Arsenic trioxide when dissolved in water exists largely as the protonated, unionized neutral species, arsenous acid As(OH)$_3$ ($pK_{a1}$ = 9.4, $pK_{a2}$ = 13.5 and $pK_{a3}$ = 14.0) [5]. The aquaporin hAPQ9 was shown to uptake As(OH)$_3$ in xenopus oocytes [11] while aquaporin mAPQ7 and mAPQ9 are involved in As(OH)$_3$ uptake in mice and rats respectively [14].

![Diagram](image.png)

**Figure 1.2** Interaction of As(III) species with thiols. Panel A shows binding of an arsenoxide species with dithiols. Panel B shows arsenous acid coordinating to three thiol groups.
Arsenic can undergo methylation inside a cell. In cases of acute toxicity, appearance of methylated arsenic species [(monomethylarsonic acid MMA(V); monomethylarsonous acid MMA(III); dimethylarsinic acid DMA(V); and dimethylarsinous acid DMA(III), shown in Figure 1.2] in human urine suggested that this phenomena was a detoxification process [16][17]. Later studies have shown that these forms are more toxic than their non-methylated arsenic precursors [18][19]. In human cells only two enzymes are required for these biotransformations, arsenate/MMA(V) reductase and arsenite/MMA(III) methyltransferase [20][21].

Initially intracellular arsenate is reduced to arsenite by arsenate/MMA(V) reductase. MMA(V) reductase has been purified from human liver and has also been
termed as glutathione-S-transferase $\omega$ [22]. After reduction the resulting arsenous acid undergoes oxidative methylation by arsenite/ MMA(III) methyltransferase, where S-adenosylmethionone (SAM) acts as the methyl donor, again followed by reduction with MMA(V) reductase resulting in MMA(III). This cycle is repeated to generate DMA(III) as a final product [23]. The overall mechanism is shown in Figure 1.3. MMA(III) methyltransferase has been purified from rabbit liver [24] and human hepatocytes [25]. Another detoxification process by which arsenite and its complexes are excreted from the cells is through multidrug resistant proteins (MRPs); however MRPs do not export arsenate species [26].

1.4 Cellular Targets of Arsenic

The biological interaction between arsenic (III) and thiols is well documented. A compilation of more than 100 enzymes that are inhibited by As(III) species was published by Webb in 1966 [27]. As(III) binding with various cysteine residue containing proteins and peptides [28–32] as well as small molecules like glutathione (GSH), dithiothreitol (DTT) [32] and dihydrolipoic acid (DHLA) [33] has been studied. In a thermodynamic study done by Spuches et al. GSH was used as a model to study its binding with arsenite, the stepwise stability constants were determined to be $20 \text{ M}^{-1}$, $430 \text{ M}^{-1}$ and $1200 \text{ M}^{-1}$ with an overall stability constant of $10^7 \text{ M}^{-3}$ [34]. In proteins, due to the favorable interaction of As(III) with cysteine residues that are in close proximity, several arsenicals have been shown to inhibit redox active enzymes containing CxxC motifs, where x represents an amino acid other than cysteine. While
this subject will be discussed in detail in future chapters a few illustrative examples are presented here.

As(III) inhibits thioredoxin reductase (TrxR), which is a part of the thioredoxin system and is the only enzyme known to directly reduce the thioredoxin (Trx) protein inside the cells. The proper functioning of this system is essential for maintaining the redox balance of the cell and its survival [35]. The mammalian TrxR contains one CxxC motif and one CU motif (U= selenocysteine), both these sites are required for the enzymes activity and are inhibited by arsenic (III) compounds. MMA(III) was found to be a potent and irreversible inhibitor of this enzyme with an IC₅₀ value of 100 nm [36]. Arsenic trioxide has also been shown to inhibit mammalian TrxR with an IC₅₀ value of 0.25 µM [37]. Glutathione reductase, a redox-active enzyme that catalyzes NADPH-dependent reduction of GSSG, is also inhibited by methylated As(III) species, MMA(III) and DMA(III) more potently than arsenite [38]. MMA(III) also inhibits the pyruvate dehydrogenase multienzyme complex with an IC₅₀ value of 17.6 µM [39]. This key enzyme system links the glycolysis pathway to the citric acid cycle. MMA(III) interacts with the dithiol present in the lipoic acid moiety of the multienzyme complex that is essential for its proper function [39]. In addition to binding to well structured catalytic motifs, previous work from our laboratory showed the ability of As(III) species to bind tightly to cysteine residues in unfolded reduced proteins [31], this work will be discussed further in section Chapter 3, sections 3.3.5 & 3.3.6 of this dissertation.
Table 1.1  Representation for binding of As(III) species to proteins and small molecular weight thiols.

<table>
<thead>
<tr>
<th>Arsenical</th>
<th>Protein and small molecules</th>
<th>$K_d$ (µM)</th>
<th>Method</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>As(OH)$_3$</td>
<td>Human thioredoxin</td>
<td>24</td>
<td>ICPMS</td>
<td>[40]</td>
</tr>
<tr>
<td>As(OH)$_3$</td>
<td>PDI</td>
<td>18</td>
<td>Fluorescence</td>
<td>[31]</td>
</tr>
<tr>
<td>MMA</td>
<td><em>E.coli</em> thioredoxin</td>
<td>23</td>
<td>ICPMS</td>
<td>[40]</td>
</tr>
<tr>
<td>MMA</td>
<td>PDI</td>
<td>25</td>
<td>Fluorescence</td>
<td>[31]</td>
</tr>
<tr>
<td>MMA</td>
<td>Thioredoxin reductase</td>
<td>0.25</td>
<td>UV-Vis</td>
<td>[41]</td>
</tr>
<tr>
<td>PAO</td>
<td>Human thioredoxin</td>
<td>0.09</td>
<td>ICPMS</td>
<td>[40]</td>
</tr>
<tr>
<td>PAO</td>
<td><em>E.coli</em> thioredoxin</td>
<td>0.01</td>
<td>ICPMS</td>
<td>[40]</td>
</tr>
<tr>
<td>PSAO</td>
<td>PD1</td>
<td>1.1</td>
<td>Fluorescence</td>
<td>[31]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Small molecular weight thiols</th>
</tr>
</thead>
<tbody>
<tr>
<td>As(OH)$_3$</td>
</tr>
<tr>
<td>As(OH)$_3$</td>
</tr>
<tr>
<td>As(OH)$_3$</td>
</tr>
<tr>
<td>MMA</td>
</tr>
<tr>
<td>MMA</td>
</tr>
<tr>
<td>MMA</td>
</tr>
<tr>
<td>GSAO</td>
</tr>
<tr>
<td>GSAO</td>
</tr>
</tbody>
</table>

Structures of As(OH)$_3$, MMA, PAO and PSAO are shown in Figure 1.1. Structure of GSAO is shown in Figure 1.4.
1.5 Arsenic as a Carcinogen

Arsenic is classified as a human carcinogen by the International Agency for Research on Cancer and the US Environmental Protection Agency. The classification is based on various epidemiological studies that associate arsenic exposure to the development of cancer, but its precise mechanism of action is still not known. The presence of arsenic influences various signaling pathways that are involved in cell proliferation or apoptosis including, p53, mitogen-activated protein kinases (MAPK), down regulation of anti-apoptotic Bcl-2 and upregulation of apoptotic Bax proteins [43]. In addition, arsenic induces apoptosis through oxidative stress by increasing the concentration of reactive oxygen species (ROS). Generation of ROS leads to activation of transcription factor AP-1 and nuclear factor- \( \text{kB} \) [44–46]. These transcription factors control cellular pathways involving cell differentiation and proliferation and their activation have been shown to induce apoptosis [47–49]. An increase in ROS also disrupts membrane potential leading to caspase 3 activation [50]. Arsenic has been shown to interact with all the MAPK pathways i.e. ERKs, JNKs and p38 kinases [51]. Arsenic-induced activation of ERK leads to cell proliferation, while activation of JNK leads to induction of apoptosis and inhibition of cell transformation [52–54].
Figure 1.4  Therapeutic arsenicals discussed in text
1.6 Arsenic as a Therapeutic

Even though arsenic is considered as a poison, it also has had a long history as a therapeutic. In fact arsenic has been used in medicinal chemistry for over 2000 years and can be regarded as one of the oldest drugs in the world. It was a key ingredient in antiseptics, antispasmodics, antiperiodic and tonics in old medicines. In traditional Chinese medicine inorganic forms of arsenic were used to treat skin diseases and asthma [55]. One of the earliest examples of an arsenic based drug is “Fowlers solution” made by Thomas Fowler, it consisted of 1% solution of potassium arsenite and was used for the treatment of asthma, epilepsy, hysteria, syphilis, ulcers and cancer [56]. Arsenic-based drugs were also used for treatment of leukemia until they were replaced by radiotherapy in the early 20th century [57].

1.6.1 Salvarsan

Salvarsan was first discovered by Elrich and Hata in 1909 as being effective against the causative agent of syphilis, Treponema pallidum [58]. The original published structure of this drug had a chemically unstable As=As bond; However later molecular weight determination showed that Salvarsan was actually a polymeric form, containing As-As bonds [59], one such structure is shown in Figure 1.4. Salvarsan was used as a treatment for syphilis until penicillin became available in 1943 [60]. Paul Ehrlich also known as the “Father of Chemotherapy” introduced the idea of “magic bullet” compounds that only targets the microbes without harming the host. In this case the magic bullet consisted of arsenic in the form of Salvarsan. Salvarsan is
considered as the first chemotherapeutic drug ever made.

### 1.6.2 Melarsoprol

Melarsoprol shown in Figure 1.4 was synthesized by Freidheim in 1949 and used for the treatment of Human African sleeping sickness or trypanosomiasis [61]. This arsenical is still one of the most widely used drug for this disease. Human African sleeping sickness is caused by *Trypanosoma brucei* and is transmitted by the bite of an infected tsetse fly. There are two forms of this disease, the West and Central African form (*Trypanosoma brucei gambiense*) and the East and South African form (*Trypanosoma brucei rhodesiense*). Around 10,000 people are affected by this disease each year [62] [63]. The exact mechanism of action of melarsoprol is still unknown, but there have been studies that show the active metabolite of the drug to be melarsen oxide [64] and the glutathione analog trypanonthione as one of its primary targets [65]. Freidheim actually had first synthesized melarsen oxide in 1948, but added the dithiols BAL group to reduce the cytotoxicity of the drug [66].

### 1.6.3 Arsenic Trioxide and Acute Promyelocytic Leukemia

A surprisingly successful use of arsenic based therapeutic has been the treatment of acute promyelocytic leukemia (APL) by arsenic trioxide (ATO) (Figure 1.4). APL is caused by an arrest of maturation of myeloid cells at the promyelocytic stage. This leads to accumulation of primary granules in promyelocytes leading to blood coagulation and resulting in hemorrhagic death [67]. The most common cause of APL
is due to a specific chromosomal translocation t(15;17) (q22;q21) which results in the fusion protein PML/RARα. This chimeric protein arrests the expression of genes required for normal cell differentiation [68].

The promyelocytic leukemia protein (PML) is encoded by the PML gene on chromosome 15 and is a redox-sensitive tumor suppressor. It is found in the nucleus in distinct structures known as PML nuclear bodies (PML-NBs) [69]. The function of PML–NBs is still not clear but they are thought to be involved in diverse functions including apoptosis, angiogenesis, protein storage and post translation modification of proteins [70]. The second protein involved in the fusion causing APL is RARα. It encode for a nuclear hormone receptor that acts as a transcriptional regulator upon binding of its ligand retinoic acid (RA). The fusion protein disrupts the function of RARα resulting in inhibition of transcription factors of its target genes.

The first therapy used for APL was treatment with all trans retinoic acid (ATRA) in 1986 [71]. ATRA binds to the chimeric protein triggering a change in its conformation and activation of transcription factors leading to proteasome mediated degradation of the fusion protein and relocation of the PML protein to the nuclear bodies [72]. In some cases ATRA is not sufficient alone and has to be combined with DNA damaging chemotherapeutic like anthracyclines leading to a 30-40% relapse within first five years [73].

Arsenic trioxide had been used in ancient Chinese medicines for the treatment of various cancers. It was formally introduced as the treatment for APL in 1970s as
“Ai-Lin I” which was a mixture of crude ATO and herbal extracts [74]. From 1994 clinical trials were done with pure ATO with patients that have undergone relapse with ATRA plus additional chemotherapy [75]. The clinical remission rate on treatment with ATO alone was 65-90% [73] [76] whereas combined therapy by ATRA and ATO resulted in clinical remission rate of > 90% [72] [75] [76]. ATO exerts dose dependent effects on APL cells; at low concentration (0.1-0.5 µM) it induces partial differentiation, whereas at high concentrations (0.5-2 µM) it induces apoptosis [79]. ATO can also be used without any DNA-damaging chemotherapy, hence it is a more potent inhibitor of APL cells than ATRA [80] [81].

Arsenic reacts with the PML portion of the fusion protein [82]. The gene encoding the PML protein consists an N -terminus cysteine-rich RING finger motif followed by two cysteine-rich B box ZF domains [83]. ATO was shown to form intermolecular As-S bonds with the cysteine residues of the zinc finger domain resulting in oligomerization [84]. The polymeric structure then interact with the SUMO ligase ubiquitin BC9 leading to sumoylation and degradation of APL cells [85]. Due to its remarkable success ATO was approved by the FDA in September 2000 and is sold under the commercial name of Trisenox by Cell Therapeutics Inc.

1.7 New Arsenic Drugs

Despite the success of arsenic trioxide in the treatment of leukemia, it has not been successful with solid tumors due to rapid renal clearance and dose limitations
because of its toxicity [86]. To improve the efficiency and absorption of arsenical drugs new techniques have been tried. The following section discusses two new arsenicals that are in clinical trials and arsenic nanoparticles as an alternate means of drug delivery.

1.7.1 **Darinaparsin (S-dimethylarsino-glutathione)**

In 2008 a new arsenical drug known as Darinaparsin (Figure 1.4) was introduced. It is a conjugate of glutathione with dimethylarsinous acid [87]. Dissociation of glutathione leads to the activated form of the drug and to its cellular uptake [88]. It acts by arresting GM-2 cycle in cells and is a more potent pro-apoptotic agent than arsenic trioxide. Darinaparsin induces apoptosis mainly through disruption of mitochondrial function resulting in ROS, and interacting with various signal transduction pathways [89] [90]. This arsenical was recently shown to be effective against solid tumors under conditions of hypoxia [91]. In addition darinaparsin is not transported out of the cells by the MRP1 proteins and its accumulation was found to be 5-fold higher than ATO in multiple myeloma cell lines [88]. The drug is currently in phase II clinical trial for primary liver cancer, advanced myeloma [92] and lymphomas [89] [93]. Darinaparsin is commercially known as ZIO-101 and is licensed by Ziopharma Oncology Inc.

1.7.2 **GSAO (4-(N-(S-glutathionylacetyl) amino) phenylarsonous acid)**

GSAO (Figure 1.4) is a conjugate of glutathione and phenylarsonous acid and
was synthesized by the Hogg laboratory [94]. It targets adenine nucleotide translocase (ANT) [94] which is the most abundant protein in the mitochondrial intermembrane space and a key component of the mitochondrial permeability transition pore (MPTP) [95]. ANT contains one cysteine residue in each of its three loops (Cys 57, Cys 160, Cys 257) facing in the mitochondrial intermembrane space. GSAO forms intramolecular cross links with Cys 57 and Cys 257 that leads to opening of the MPTP resulting in an increase of ROS and enhanced apoptosis of proliferating endothelial cells [96]. The same cysteine residues are targets of PAO and arsenite [97–100]. GSAO is an angiogenesis inhibitor of quiescent endothelial cells and hence acts indirectly by blocking blood supply to tumors, since endothelial cells do not have MRP1/2 [101] they are not able to export the drug out of the cell. In contrast the tumor cells overexpress MRPs and tend to be resistant to a wide range of therapeutics. A prodrug approach has also been applied with a new arsenical PNEAO. PNEAO was shown to accumulate 85-times faster inside the cell with a 44-fold increase in antiproliferative activity and 20-fold increase in antitumor activity in mice compared to GSAO [102][103]. GSAO is currently under phase 1 clinical trial with patients with solid tumors undertaken by Cancer Research UK, Department of Medical oncology, Christie Hospital.

1.7.3 Arsenic Nanoparticles

Encapsulation of ATO in lipid based nanoparticles known as “nanobins” has been used as a method to improve its absorption and delivery. The major drawback of
this technique is the substantial loss of ATO within a few hours from these nanobins. Since the predominate form of ATO in aqueous solution at physiological pH is the protonated, neutral arsenous acid [As(OH)$_3$] species, it leads to the rapid diffusion of encapsulated ATO across the lipid bilayers [104]. In work done by the O’Halloran group, ATO was encapsulated along divalent transition metal ions like Ni$^{2+}$ which led to the formation of stable insoluble complexes [105]. The addition of metal ions increases the stability of these nanobins with an increase in shelf life from a few hours to up to 6 months at 4°C [105]. The nanobins were tested against triple negative breast cancer cells and exhibited greater cytotoxicity and enhanced tumor uptake along with reduced plasma clearance in comparison to ATO alone [106]. Co-encapsulation of ATO and cisplatin has also been achieved in nanobins [107]. Novel arsenic compounds have recently been synthesized by O’Halloran and colleagues including arsenoplatin that contains an As(III)–Pt(II) bond. This compound was found to be more than two-fold more cytotoxic against cisplatin resistant ovarian cancer cell line A2780$^{CP}$ than cisplatin alone [108].

1.8 Arsenic Based Biotinylated Pull Down

To identify proteins that are targeted by arsenic, Zhang et al. synthesized an arsenic-biotin conjugate by coupling the pentafluorophenol ester of biotin with PAO and testing it against the human breast cancer cell line MCF-7 [109]. The bound proteins were eluted from a streptavidin resin, separated by SDS-PAGE and identified by MALDI mass spectroscopy. Western blotting identified β-tubulin and pyruvate
kinase M2 as two arsenic binding proteins. Donoghue et al. used biotin attached with
the GSAO arsenical to identify cell-surface proteins that contained closely spaced
thiols. After pull down with streptavidin labeled beads, 10 and 12 distinct proteins
were identified from endothelial and fibrosarcoma cell surfaces respectively, Protein
disulfide isomerase (PDI), a protein that will be introduced later in Chapter 2 of this
dissertation was identified from both cell types [42].

1.9 Arsenic Tomography

![Reaction scheme for labeling arsenic isotopes on antibody](image)

**Figure 1.5 Reaction scheme for labeling arsenic isotopes on antibody.** Adapted
from [110]

Arsenic isotopes have been used for molecular imaging of solid tumors in rats
[111]. A method to attach As$^{77}$ on Bavituximab, a chimeric antibody was developed by
Thorpe & Rösch (Figure 1.5) [110]. The antibody specifically targets
phosphatidylserines exposed on the external surface of vascular endothelial cells in
tumors [112]. More than 99.9% of the labeling could be achieved with stability up to
72 hrs making it highly suitable for in vivo molecular imaging. The concentration of arsenic used for the labeling was 10-fold below the toxicity level. It is surprising that the arsenical shows such stability in biological fluids given the general reversibility of As-thiol complexes in aqueous solution.

1.10 Fluorescent Arsenic Molecules

Fluorescent arsenicals have been used as probes to study the interaction of proteins with arsenic. One of the first of these arsenicals to be synthesized was PAO-DNS (Figure 1.6), in which phenylarsine oxide was complexed to dansyl chloride to generate a fluorescent conjugate [113]. It was used for studying melarsoprol like activity against trypanosomes. Hogg and coworkers had prepared two fluorescent versions of GSAO drug (GSAO-F and GSAO-Cy5.5, Figure 1.6). GSAO-F is a fluorescein-bound analog that was originally used for finding the localization of the drug in rapidly growing endothelial cells, whereas the cyanine labeled analog GSAO-Cy5.5 was used for live tumor imaging in mice [114].

Fluorescent bis-arsenicals have been popular probes for studying various protein-protein interactions in vitro and in vivo. Roger Tsien’s laboratory designed the first bis-arsenical molecule FlAsH in 1998 (Figure 1.7) [115]. FlAsH itself is non-fluorescent and contains two 1,2-ethanedithiol (EDT) molecules bound to arsenic. The compound becomes fluorescent on displacement of the EDT molecules by a peptide tag containing CCxxCC motif that is genetically introduced in the protein of interest.
The rigid spacing of the cys residues allows the arsenical to bind tightly ($K_d \sim 10^{-11}$ M) [116]. Originally the tetracysteine motif was sited within an $\alpha$-helix conformation but later studies showed that the brightest and highest affinity complexes were formed with the $\beta$ hairpin sequence of CCPGCC. Various red and blue analogs of FlAsH have been prepared that are useful for distance measurements by Förster resonance energy transfer (FRET) [117] [116]. These reagents have been used for study of protein localization [118] [119], protein folding [120] [121], live cell fluorescent imaging [122] and in the investigation of amyloid formation of proteins involved in Alzheimer and Huntington diseases [123–126]. Affinity purification of CCxxCC tagged proteins with immobilized FlAsH on sepharose resin has also been performed [127] [116].

Many cyanine labeled fluorescent arsenicals have been prepared by Thomas Squier's laboratory, for example AsCy3 (Figure 1.7), is a bis-arsenical that can be labeled with protein of interest with a genetically fused peptide tag CCKAEACC ($K_d \sim 80 \pm 10$ nm) and further used as FRET pair with FlAsH labeled proteins [128]. The cell permeability of this compound was increased by substituting the sulfonate groups with uncharged methoxy ester groups [129]. A two cyanine (Cy3 and Cy5) derivatives of this compound AsCy3Cy5 (Figure 1.7) has also been made which can act as photoswitchable fluorescent probe for studying protein trafficking [130]. Another fluorescent monoarsenical thiol reactive affinity probe (TRAP_Cy3, Figure 1.7) was also developed by the same group and used for determining the redox-active cysteines in microbes [131].
Figure 1.6  Fluorescent arsenicals discussed in the text
Figure 1.7  Fluorescent arsenicals discussed in the text (continued)
REFERENCES


[33] A. Von Döllen and H. Strasdeit, “Models for the Inhibition of Dithiol-Containing Enzymes by Organoaarsenic Compounds : Synthetic Routes and the Structure of \([\text{PhAs (HlipS2)}] (\text{HlipS2 Reduced Lipoic Acid})\),” *European journal of inorganic chemistry*, vol. 1, p. 61, 1998.


2.1 Disulfide Bond Formation

In eukaryotic cells about 25% of their total proteins encoded are secreted [1], with the majority of these proteins requiring disulfide bond formation to attain their native state. Disulfide bonds play a variety of roles. Firstly, they frequently serve structural roles increasing the thermal stability of proteins and suppressing proteolytic digestion of extracellular proteins. Some examples of proteins containing structural disulfides are shown in Figure 2.1. Secondly, some disulfide bonds have regulatory roles and control the function of the protein when they break and/or form, these are termed as “allosteric” for this reason [2]. Thirdly, a number of enzymes contain redox-active disulfides that play essential roles in catalysis.

Chemically the generation of a disulfide bond from a pair of sulfhydryl functional groups involves a 2-electron oxidation as shown in equation 2.1, where X is the electron acceptor.

\[2 \text{RSH} + X \rightarrow \text{RSSR} + XH_2\]  
(Equation 2.1)

When molecular oxygen is the oxidant this reaction is very slow because dioxygen is a ground state triplet whereas thiols are singlet [3]. While the reaction is
spin forbidden this restriction can be circumvented in the presence of transition metal ions such as Cu$^{2+}$ and Fe$^{3+}$. The catalysis of thiol autoxidation by transition metal ions is well known [3] and a number of iron- and copper-dependent sulphydryl oxidase enzymes were described in the earlier literature. However while the existence of these oxidases is chemically appealing a reexamination of the evidence for their existence suggests that metal content of these enzymes is artifactual. This chapter discusses oxidative protein folding and some of the catalyst known to introduce and rearrange disulfide bonds in nascent protein chains.
Figure 2.1  Structure of some disulfide bonds-containing secreted proteins. The disulfide bonds are shown in yellow.
2.2 Oxidative Protein Folding

Figure 2.2 Oxidative protein folding. Disulfide bonds are introduced in the reduced unfolded protein in the oxidation step by an oxidant (represented by X). Any mispaired disulfides are then isomerized to generate their correct pairing in the isomerization step. The insertion of correct disulfide pairings leads to the attainment of the native state of protein.

The process by which reduced unfolded proteins form disulfide bonds to attain their three dimensional native state is known as oxidative protein folding. Oxidative protein folding pathway consists of two steps as shown in Figure 2.2. The first step represents the oxidation where disulfide bonds are inserted in the reduced protein. However all know oxidizing systems known are error prone and a second step is required to rearrange the mis-paired disulfide bonds prior to the emergence of the native fold. As the number of disulfide bonds to be introduced increases, the number of potential mispairings rises dramatically. Hence for proteins containing, 2, 4, 8 and 16 cysteine residues, complete oxidation yields 1, 3, 105 and > 2 million possibilities respectively [4]. The following section reviews selected oxidative folding pathways in prokaryotes and eukaryotes.
2.3 Oxidative Protein Folding in Prokaryotes

In gram negative prokaryotes a family of thiol oxidoreductase known as disulfide bond forming (Dsb) proteins drives oxidative folding in the periplasmic space [5]. Figure 2.3 depicts a representation of the pathway in *Escherichia coli*. DsbA introduces disulfide bonds in reduced protein substrate through a redox-active CxxC motif [6]. The reducing equivalents are then passed to the membrane-bound protein DsbB by disulfide exchange [7] and then further passed on to ubiquinone under aerobic conditions or menaquinone under anaerobic conditions, hence linking the oxidative folding to the electron transport chain [8]. Any mispaired disulfides in the protein substrate originally introduced by DsbA are isomerized by DsbC [9], which is a homodimer with two redox-active CxxC motifs and V shaped peptide binding site [10]. DsbC is maintained in its reduced form by another transmembrane protein DsbD [11] which receives its reducing equivalents from cytosolic thioredoxin [8] as shown in Figure 2.3. In bacteria that lack DsbB, a eukaryotic homolog of vitamin K epoxide reductase (VKOR), a protein required for blood clotting in humans, accepts electrons from DsbA and passes them to vitamin K epoxide [12] [13].
Figure 2.3  The Dsb system of oxidative protein folding in prokaryotes. Reprinted from Biochimica et Biophysica Acta-Molecular Cellular Research, Volume 1783, Issue 4, pages 520-529, K.Inaba, K.Ito, “Structure and mechanisms of the DsbB-DsbA disulfide bond generation machine”.
2.4 Oxidative Protein Folding in Eukaryotes

2.4.1 PDI – First Model of Oxidative Protein Folding

In eukaryotes there are various modes of oxidative protein folding in the endoplasmic reticulum (ER), at least four of which use the enzyme protein disulfide isomerase (PDI) as the immediate oxidant of the reduced protein substrate. PDI also acts as the isomerase for the next step of oxidative folding. The enzymes that regenerate oxidized PDI and their respective oxidative folding pathways are discussed in the next section. PDI will be discussed in detail in section 2.6 of this chapter.

2.4.1.1 Ero1

The yeast endoplasmic reticulum oxidoreductin1 (Ero1) was discovered independently by two groups [14] [15] and subsequently shown to be a flavin adenine dinucleotide (FAD) linked enzyme that oxidizes PDI. The interaction between Ero1 and PDI was first established in yeast in a study done by Kaiser group [16] where the mixed disulfide between PDI-Ero1 was trapped. Two orthologs of yeast Ero1 have been described in mammals, Ero1α and Ero1β [17] [18]. In this pathway, oxidized PDI receives reducing equivalents from the reduced protein substrate, Ero1 then regenerates the reduced PDI with concomitant reduction of molecular oxygen to hydrogen peroxide [19] as depicted in Figure 2.4; panel A. Reduced PDI also isomerizes the mispaired disulfides in the second step to produce the native form of the protein substrate.
Even though this pathway was rapidly accepted as the major, or even sole, generator of disulfide bonds in mammals, the Ero1-PDI system presents some interesting issues. For example both oxidized and reduced PDI are needed for “PDI-first” pathways of oxidative folding and yet the glutathione redox poise predicted from modern measurements of ER redox status [20] [21] predict that glutathione will be completely reduced. Since PDI is rapidly reduced by glutathione, the use of Ero1 as an oxidase of PDI will lead to potential of a “futile cycle” in which reduced glutathione and reduced protein substrates compete for oxidized PDI. Another interesting issue is that reduced PDI seems a very poor substrate of Ero1[22]. Lastly a simultaneous knockout of Ero1α and β isoforms in mouse by Ron and colleagues [23] produced a very mild phenotype. These findings led to the consideration of other pathways for disulfide bond generation.

2.4.1.2 Peroxiredoxin IV

Peroxiredoxin IV (PrxIV) is an abundant ER resident enzyme and exists as a decamer with five dimers [24]. It protects the cell from oxidative stress by reducing H$_2$O$_2$ produced by Ero1-PDI oxidative folding pathway [25] discussed above. A study by Zito et al. showed that PrxIV could drive PDI catalyzed refolding of RNase in vitro by utilizing the H$_2$O$_2$. The overall reaction catalyzed by PrxIV is shown in Figure 2.4; panel B. An active cysteine of PrxIV reacts with H$_2$O$_2$ and become sulfenylated. The sulfenic acid then reacts with a cysteine in the adjoining chain thereby forming an
intermolecular disulfide which can be reduced by PDI generating oxidized PDI which can then drive the oxidative folding of RNase [26].

2.4.1.3 Glutathione Peroxidases

A study done by the Ruddock lab showed that two human glutathione peroxidases GPx7 and GPx8 were ER resident proteins that could utilize the peroxide generated by the Ero1-PDI pathway to make disulfide bonds. The overall reaction is shown in Figure 2.4; panel C. The addition of GPx7 or GPx8 along with PDI and H$_2$O$_2$ enabled efficient oxidative refolding of a reduced denatured bovine pancreatic trypsin inhibitor (BPTI) in vitro [27]. GPx7 and GPx8 were also shown to interact with Ero1α in vivo, and GPx7 significantly increased the rate of oxygen consumption by Ero1α in vitro [27].

2.4.1.4 Vitamin K Epoxide Reductase

Vitamin K epoxide reductase (VKOR) is a transmembrane ER resident protein [28]. The enzyme reduces vitamin K epoxide to a hydroquinone which is an essential cofactor for gamma-glutamyl carboxylase mediated catalysis of glutamate carboxylation. The carboxylase catalyzes an essential post-translation modification in blood clotting [29]. In vitro experiments show that VKOR drives the oxidative folding of reduced RNase catalyzed by PDI [30]. In this pathway, the electrons are donated by reduced protein substrate to PDI which transfer them to VKOR through disulfide exchange. The reducing equivalents are finally passed to vitamin K epoxide
regenerating oxidized PDI. The reaction catalyzed by VKOR is depicted in Figure 2.4; panel D. Mixed disulfide with VKOR and two members of PDI family, TMX4 and TMX, has been trapped in vivo [28].
Figure 2.4  Eukaryotic pathways of PDI-first model of oxidative protein folding. Four eukaryotic enzymes, Ero1 (A), PrxIV (B), GPx7/8 (C) and VKOR (D) regenerate oxidized PDI in the ER, which subsequently drive the oxidative folding of reduced protein substrates.
2.4.2 QSOX-PDI pathway – A Alternative Model for Oxidative Protein Folding

In this pathway Quiescin Sulphhydryl Oxidase (QSOX), a flavoprotein sulphhydryl oxidase acts as the immediate oxidant and directly inserts disulfide bonds in the reduced protein substrate. Here PDI acts as an isomerase to correct mispaired disulfides (Figure 2.5). PDI itself is not a substrate of QSOX and hence both these enzymes work independently [31]. QSOX can insert disulfides with turnover numbers of ~2000 thiols oxidized per minute with a $K_m$ of 150 $\mu$M (Table 2.1) making it the most efficient sulphhydryl oxidase known to date. In vitro low nanomolar concentrations of QSOX and micromolar concentrations of reduced PDI are all that is required for the efficient refolding of riboflavin binding protein RfBP [31]. Since this substrate has 18 free thiols and hence 34 million possibilities for disulfides pairing, the QSOX-PDI pathway deserves consideration as one of the multiple pathways for disulfide bond formation in the ER. The following section describes in more detail the structure and mechanism of QSOX.
2.5 Quiescin Sulphhydryl Oxidase (QSOX)

2.5.1 Discovery of QSOX

This flavoprotein sulphhydryl oxidase was first purified from avian egg white by following the yellow color of its cofactor FAD [32]. The FAD content of egg white had previously been described by Dr. Harold White and coworkers [33]. Enzymatic studies showed that the protein was a disulfide bond generator for reduced unfolded proteins and was shown to catalyze the following reaction

\[
2 \text{R-SH} + \text{O}_2 \rightarrow \text{R-S-S-R} + \text{H}_2\text{O}_2
\]  
(Equation 2.2)

Determination of the sequence of tryptic peptides from the avian sulphhydryl oxidase led to the finding that it was homologous to the human gene Quiescin Q6, a protein that was shown by Coppock and coworkers to be secreted from human fibroblasts as they enter quiescence [34][35]. This gave the enzyme its name, quiescin sulphhydryl oxidase and its abbreviation QSOX. Discovery of QSOX from egg white lead to the reinvestigation of previously reported Cu and Fe-dependent sulphhydryl oxidase from rabbit skin [36] and bovine milk [37] respectively. These studies suggested that these enzymes were also QSOX family members [38]. In the case of Fe-dependent sulphhydryl oxidase described in bovine milk, a reexamination showed that essentially all of the enzymes activity was due to QSOX and the iron content reflected contamination from the iron-binding protein lactoferrin [39]. These oxidases were then recognized as members of the newly formed flavoenzyme family QSOX. Another sulphhydryl oxidase from seminal vesicles [40] was also found to be a member of
QSOX family [41]. Detailed study on the QSOX enzymes from avian egg white [42], bovine milk [39] and recombinant human [43] and trypanosomal brucei [44] have now been conducted. The next section discusses the cellular location of mammalian QSOX before dealing with the structure of this oxidase.

### 2.5.2 Localization of QSOX

In humans two paralogs of QSOX (QSOX1 and QSOX2) have been identified, with a 37% sequence identity [45]. QSOX1 is found to be generally expressed at higher levels than QSOX 2 [45] and has two splice variants. QSOX1a retains the transmembrane region at the C-terminus of the protein and a shorter form QSOX1b that is secreted. In situ hybridization and immunohistochemistry showed that QSOX is present in rat immune, endocrine, reproductive and respiratory systems as well as digestive and urinary tracts [46]. High levels of expression were also observed in epithelial cells that specialize in protein secretion such as Islet of Langerhans and the bronchiolar epithelial layer of the lung [46]. QSOX 1 is also secreted into a variety of biological fluids including milk, human tears and seminal vesicles [47].

Immunohistochemistry shows that human QSOX1 is abundant in epidermis, pancreas, small intestine, placenta, lungs and hair follicle [47]. QSOX1 is prominent in tissues with a heavy secretory load, consistent with its role in oxidative protein folding [48]. It is also believed to play a role an important role in the formation of extracellular matrix and its activity was required for incorporation of laminin into the matrix [49]. Intracellularly QSOX is found in the ER, golgi and at the cell surface [45], [50–52].
2.5.3 Domain Structure of QSOX

![Domain Structure of QSOX](image)

**Figure 2.6 Domain structure of QSOX proteins.** Metazoans QSOXs have two thioredoxin domains (Trx1 and Trx2) followed by a helix-rich region and an ERV/ALR domain (Panel A). Plant and protists lack the second redox-inactive Trx2 domain (Panel B). Cysteines of the redox-active CxxC motifs are represented by solid yellow lines. The flavin cofactor is shown in orange. The transmembrane region (TM) is shown by solid black line. The approximate position of signal sequence is depicted in red.

QSOX protein family are multi-domain enzymes and represent a fusion of two ancient proteins thioredoxin and Erv1[48] [53] (Figure 2.6). The N-terminal starts with a signal sequence followed by four recognizable domains. In metazoan QSOXs the first domain is a PDI-like Trx domain which contains the first redox-active CxxC motif (labeled as C\textsubscript{I}xxC\textsubscript{II} in Figure 2.6; also known as the distal disulfide because of its distance from the flavin cofactor) [54]. The second domain is also a PDI-like Trx domain but it lacks a redox-active disulfide. The function of this domain is unknown and it is absent from plant and protist QSOXs [48]. It is followed by a helix rich region (HRR), which appears to have evolved from an Erv/ALR like domain but lacks redox-active centers [55]. The last domain is an Erv/ALR domain which contains the non-covalently bound flavin cofactor and the second CxxC motif (labeled as C\textsubscript{III}xxC\textsubscript{IV}}
in Figure 2.6; also known as the proximal disulfide). This last structural module also contains a third CxxC motif of currently unknown function, which was found to be catalytically non-essential at least in vitro [43]. The discovery of the QSOX family led to the uncovering of a number of small stand-alone sulfhydryl oxidases collectively termed as Erv/ALR. Erv1p and Erv2p are found in yeast and the mammalian homolog of Erv1p, augmenter of liver regeneration (ALR) is largely found intra-mitochondrially [56–58]. The three CxxC motifs are conserved in all the QSOX proteins [55]. Genes encoding for one or more QSOXs are present in all metazoans and plants sequenced to date [45]. However QSOX is absent from fungi but have Ero1p that is essential for growth of Saccharomyces cerevisiae [14].
2.5.4 Mechanism of QSOX

As mentioned earlier, QSOX directly oxidizes reduced proteins with the concomitant reduction of molecular oxygen to hydrogen peroxide. Studies done by this laboratory leads to the following overall catalytic mechanism (summarized in Figure 2.7).

A thiolate of the protein substrate attacks the solvent-exposed cysteine residue C₁ in the first Trx domain of QSOX generating a mixed disulfide [55] [59]. The mixed disulfide is resolved by the attack of another thiolate anion from the protein substrate resulting in the formation of disulfide bond in the substrate and reduction of the Trx domain [42] [43]. The reducing equivalents are then passed to the proximal CxxC motif via a mixed disulfide between C₁ and C₃III, this is the rate limiting step of the enzymes catalytic cycle [44]. This mixed disulfide is resolved by C₃II, regenerating

Figure 2.7 Mechanism of QSOX. The figure represents the flow of reducing equivalents in metazoan QSOX.
oxidized Trx domain and reduced proximal disulfide [43]. The reducing equivalents are then further passed to the flavin cofactor and, finally, to molecular oxygen [43] [44]. A recent crystal structure of the open and closed forms of TbQSOX shows the extensive rotation of the Trx domain necessitated by its role as an intermediary between the protein substrate and the ERV/ALR domain (Figure 2.8) [60].
Figure 2.8  Crystal structure of *Trypanosoma brucei* QSOX. Open (A) and closed (B) conformation of TbQSOX. The closed conformation represents the trapped mixed-disulfide between the Trx and ERV domain. The flexible loop (dashed line) between Trx and HRR domain is not visible in the crystal structure. Trx1 domain is shown in green, HRR is shown in grey and ERV/ALR domain is shown in blue. The cysteine residues involved in mixed disulfide are represented by spheres and their sulfur atoms are colored in yellow. PDB file: 3QDE.
2.5.5 Substrates of QSOX

*In vitro* kinetic studies done on avian, human and *Tb*QSOX show that the enzyme is able to oxidize reduced unfolded proteins efficiently, in contrast folded proteins like pyruvate kinase and aldolase, while they contain multiple thiols, are not good substrates [43][44][61]. In an extension of this work Codding et al. [62] showed that QSOX is able to form polymeric structures by forming intermolecular disulfides between unstructured polypeptide chains. In contrast QSOX was unable to link the thiols of folded protein substrates. This study also established that the QSOX apparently lacks a significant binding site for protein substrates. Table 2.1 summarizes the kinetic parameters of QSOX with a range of its substrates. The *K*$_{cat}$ value largely reflects the rate of inter-domain electron transfer [54] and hence does not vary greatly between substrates of the enzyme. A dominant aspect of the catalytic efficiency (*k*$_{cat}$/*K*$_{m}$; Table 2.1) is therefore the *K*$_{m}$ term. Here the monothiols have *K*$_{m}$ values from about 10-50 mM. In contrast dithiol and multi-thiol substrate show *K*$_{m}$ values that are generally much lower than 1 mM.
### Table 2.1  Catalytic activity of QSOX

<table>
<thead>
<tr>
<th>Substrate</th>
<th>$k_{\text{cat}}$ (min$^{-1}$)</th>
<th>$K_m$ (µM)</th>
<th>$k_{\text{cat}}/K_m$ (M$^{-1}$s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Avian QSOX</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cysteine$^{a,d,e}$</td>
<td>1275</td>
<td>10,900</td>
<td>2.0 x 10$^2$</td>
</tr>
<tr>
<td>b-Mercaptoethanol$^{a,d,e}$</td>
<td>1215</td>
<td>54,000</td>
<td>3.8 x 10$^2$</td>
</tr>
<tr>
<td>Dithiothreitol$^{a,e}$</td>
<td>1033</td>
<td>150</td>
<td>1.2 x 10$^3$</td>
</tr>
<tr>
<td>Glutathione$^{a,d,e}$</td>
<td>1385</td>
<td>20,000</td>
<td>1.2 x 10$^3$</td>
</tr>
<tr>
<td>Aldolase$^{b,f}$</td>
<td>200</td>
<td>160</td>
<td>2.1 x 10$^4$</td>
</tr>
<tr>
<td>Insulin chain A$^{b,f}$</td>
<td>1000</td>
<td>215</td>
<td>8.0 x 10$^4$</td>
</tr>
<tr>
<td>Insulin Chain B$^{b,f}$</td>
<td>700</td>
<td>300</td>
<td>3.9 x 10$^4$</td>
</tr>
<tr>
<td>Lysozyme$^{b,f}$</td>
<td>860</td>
<td>110</td>
<td>1.3 x 10$^5$</td>
</tr>
<tr>
<td>Pyruvate kinase$^{b}$</td>
<td>475</td>
<td>1250</td>
<td>6.5 x 10$^3$</td>
</tr>
<tr>
<td>RfBP$^{b,f}$</td>
<td>1100</td>
<td>230</td>
<td>8 x 10$^4$</td>
</tr>
<tr>
<td>RNase$^{b,f}$</td>
<td>610</td>
<td>115</td>
<td>9 x 10$^4$</td>
</tr>
<tr>
<td>PDI$^b$</td>
<td>---</td>
<td>---</td>
<td>&lt; 10$^2$</td>
</tr>
<tr>
<td><strong>Trypanosoma brucei QSOX</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trypanothione$^c$</td>
<td>480</td>
<td>3230</td>
<td>2.5 x 10$^3$</td>
</tr>
<tr>
<td>RNase$^c$</td>
<td>660</td>
<td>360</td>
<td>3.0 x 10$^4$</td>
</tr>
<tr>
<td>RfBP$^c$</td>
<td>420</td>
<td>300</td>
<td>2.3 x 10$^4$</td>
</tr>
<tr>
<td><strong>Human QSOX</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DTT$^{d,g}$</td>
<td>620</td>
<td>12,400</td>
<td>1.0 X 10$^5$</td>
</tr>
<tr>
<td>RNase$^{d,f}$</td>
<td>2160</td>
<td>320</td>
<td>1.1 x 10$^5$</td>
</tr>
</tbody>
</table>

a; Values compiled from reference[32],  b; values compiled from reference[61],  c; values compiled from reference[44], d; values compiled from reference[43], e; monothiols, f; reduced denatured protein, g; dithiols
2.5.6 QSOX’s Role in Cancer

Upregulation of QSOX has been found in pancreatic [63] prostate [64] and breast cancer cells [65]. In 2009 Lake and coworkers identified a peptide from the C-terminus of QSOX1 as a plasma biomarker for pancreatic ductal carcinoma patients [63]. Further studies done by the same group showed that QSOX1 activates posttranslation modifications in the metalloproteinases MMP-2 and MMP-9 that are contributors to the cell invasion potentials of pancreatic tumor cells [66]. QSOX1 was also identified as a biomarker in blood serum for patients with acute decompensated heart failure (ADHF) especially when combined with the levels of B-type natriuretic peptides that are released from stressed cardiomyocytes [67]. In prostate cancer cells QSOX1 overexpression is correlated with the loss of the NX31.1 gene product which is a tumor suppressor expressed in human prostate epithelial cells [64]. Recently the Lake group has also identified QSOX1 overexpression in luminal B breast cancer [68]. The knockout tumor cells were shown to have suppressed invasion and proliferation activity, while exogenous addition of QSOX1 restored their invasive potential. It is surprising that despite several correlations of enhanced QSOX1 levels with tumorigenesis, in one study QSOX1 upregulation was reported to protect breast cancer cells from invasion and also decreased tumor development in vivo [69]. Clearly more comprehensive studies need to be done to clearly determine the diverse role of QSOX1 in breast cancer.
2.6  Protein Disulfide Isomerase

2.6.1  Discovery of PDI

Pioneering work by Anfinsen and his colleagues led to the purification of an enzyme from rat liver extracts that catalyzed the refolding of RNase \textit{in vitro} [70]. The addition of this enzyme decreased the rate of half-time for the oxidation of RNase from several hours to 20 min [71]. An independent study from Venetianer and Straub also showed the reactivation of reduced RNase by a protein from chicken, pigeon and pig pancreatic tissues [72]. These proteins were collectively termed as thiol: protein disulfide oxidoreductases (TPOR). Later it was reported that this enzyme also facilitates the rearrangement of disulfide pairs in extensively mispaired scrambled RNase leading to the recovery of the activity of native protein. These observations led to renaming of the enzyme as “protein disulphide isomerase” (PDI) [73]. Prior to these discoveries, a previous enzyme with a similar molecular weight was purified from beef liver extract that promoted the degradation of insulin in presence of glutathione had been reported [74]. This latter catalyst was named as glutathione: insulin transhydrogenase (GIT) [75]. Many purifications of this enzyme were performed using liver and pancreas with little agreement on molecular properties. However each preparation showed some protein disulfide isomerase activity. It was later pointed out that many of these preparations were contaminated with partially degraded TPOR [76].
It took considerable time to achieve an improved purification of PDI from bovine liver [77] that led to sufficient amount of enzyme for a thorough characterization. A series of kinetic studies using the insulin reductase and RNase assays showed that the pure protein had GIT and TPOR-like activity. Detailed studies by Vardani’s group provided information of the enzymes tissue and its subcellular distribution [78]. Importantly antibodies that were raised against the three isolated proteins showed that they were immunologically identical [79]. Eventually it was concluded that all the activities were a reflection of pure protein disulfide isomerase [77].

2.6.2 Localization of PDI

PDI is an abundant protein in many mammalian cell types. It has been reported to be present in millimolar quantities in ER [80]. More than 20 PDI family members that have been identified in humans alone [81] and *Saccharomyces cerevisiae* has 5 PDI homologs [82]. Somewhat surprisingly PDI has several cellular locations. Intracellularly, a main locus of PDI is in the ER [83] [84] with smaller amounts found in the nucleus [85] and cytosol [86]. PDI is also found at the cell surface [87] and is also secreted from activated platelets [88], endothelial cells [89], hepatocytes [87] and pancreatic exocrine cells [90].
2.6.3 Structure of PDI

PDI is a member of the thioredoxin superfamily; all PDI family proteins contain one or more thioredoxin domain. Thioredoxin, itself is a single domain 12 kDa redox-active protein found seemingly universally in all prokaryotes and eukaryotes [91] [92]. The Trx domain contains a characteristic $\beta_4\alpha_4\beta_4\alpha\beta_4$ fold [93].

Mammalian PDI consists of four thioredoxin domains, $a$, $a'$, $b$, and $b'$, it also contains a short helix tail at the C-terminus. The domain structure of yeast PDI is depicted at the top of Figure 2.9. The $a$ and $a'$ domain have 37% sequence similarity and contain the redox-active CxxC motifs (here CGHC). The redox-inactive $b$ and $b'$ domain have 19% similarity [94]. The short acidic helix at the C-terminus contains the KDEL retention sequence for ER localization. While the full length crystal structure of a mammalian PDI is yet to appear, the structure of the highly similar yeast PDI is depicted in Figure 2.9. The crystal structure of yeast PDI shows the four Trx domains arranged in a “U” shape with the redox-active CxxC motifs facing each other in the $a$ and $a'$ domain [95]. The $b$ and $b'$ domain form the base of the U shape and constitutes a binding pocket that is large enough to accommodate an unfolded protein of about 100 amino acid residues [95]. The surface of the $b'$ domain primarily consists of hydrophobic residues and is essential for binding of the substrates. Larger substrates also interacts with the $a$ and $a'$ domain [96]. The $b$ domain does not have a clear role in binding but it is considered structurally important. All the domains, including the C-terminus tail are required for full catalytic activity of the enzyme [97].
As mentioned earlier, the human PDI is expected to be highly similar to yeast enzyme with which it share ~30% sequence similarity [98]. The NMR structure for the human $a$ and $a'$ domains shows similar Trx folds with the active site CxxC motifs present in the same position as the yeast PDI [99]. However there are few difference between the two enzymes, the regulatory disulfide present in $a$ domain of yeast enzyme is absent in the human form, but it contains two non redox-active cysteines in the $b'$ domain that are absent in the yeast structure [100].
Figure 2.9  Domain organization and crystal structure of PDI from *Saccharomyces cerevisiae*. The redox-active cysteines are represented by spheres and their sulfurs are shown in yellow. PDB code: 2B5E
2.6.4 Functions of PDI

PDI is involved in various oxidoreductase and isomerization reactions [84]. These are depicted in Figure 2.10 and briefly discussed in the following section.

Panel A in Figure 2.10 reflect the reversible oxidoreductase activity of PDI. As mentioned earlier, the “PDI-first” model of oxidative protein folding requires PDI to serve as a net oxidant. Nucleophilic attack by the reduced protein substrate generates a mixed-disulfide intermediate which can be resolved with a second thiol-disulfide exchange reaction leading to the insertion of a disulfide bond into the protein substrate and to reduction of PDI. In the case of the net generation of disulfide bonds, reduced PDI must then be reoxidized by Ero1, or the other oxidoreductases described in section 2.4.1 of this Chapter. However in terms of the isomerization activity of PDI the reactions in Panel A are freely reversible and so a mispaired disulfide introduced by PDI could then be removed by such reversal. A second round of reversible oxidation might generate a correct pairing which is less prone to reductive removal. In addition the disulfide introduced into the substrate protein might be made more thermodynamically stable with the introduction of additional correct pairing. Thus both the net oxidation of protein substrates and the isomerization of mispaired disulfide bonds require the involvement of both oxidized and reduced PDI. Figure 2.10; panel B illustrates a mode of isomerization of mispaired disulfide by PDI through disulfide bond “shuffling”. Here a series of mixed disulfide bonds can be resolved by different cysteine residues in the protein substrate. Here PDI does not cycle through discrete oxidized and reduced states.
**Figure 2.10** Reactions catalyzed by PDI. Panel A represents an oxidation of substrate protein by PDI. Panel B represents PDI catalyzed isomerization reactions through "disulfide shuffling".
2.6.5 PDI’s Role in Cancer

Protein disulfide isomerase upregulation has been reported in numerous cancer cells including, ovarian [101], brain [102], lung [103], prostate [104] [105], kidney [106] [107], melanoma [108] [109] and breast [110]. In breast cancer cells, PDI is one of the most upregulated proteins in interstitial fluids and hence has been proposed as a potential biomarker [111]. In addition it serves as a biomarker for the diagnosis of colorectal cancer [112]. PDI is also found on cell surfaces where it modulates extracellular redox poise [113–115] and is involved in platelet activation [113] [116] [117], enhancing metastasis [118] [102], and viral fusion [115] [119] [120]. PDI is reported to reduce disulfide bond in gp120 that helps in the entry of Human immunodeficiency virus (HIV) inside the cell [121]. Due to broad involvement of this enzyme in various diseases, it has been a target for antiviral [122], antithrombotic [123] and anticancer therapies [124][125].
REFERENCES


Chapter 3

INHIBITION OF OXIDATIVE PROTEIN FOLDING BY MULTIVALENT ARSENICALS

3.1 Introduction

The finding that administration of arsenic trioxide is remarkably effective in the treatment of acute promyelocytic leukemia [1] has stimulated renewed interest in the potential of arsenicals for the treatment of a variety of other types of cancer. As discussed in Chapter 1, arsenic trioxide yields arsonous acid in aqueous solution and this reagent can bind up to three thiols [2]. It is this coordination of thiols to As(III) species that is believed to underlie both the toxicity and the clinical potential of arsenicals. While arsenite (arsonous acid has a first pK of 9.4 [2]) binds many monothiols comparatively weakly, bis or tris-mercaptans, in which the sulfhydryl group can form stable ring structure with As(III) species, leads to tighter binding via the chelate effect [3]. Indeed a large number of enzymes with dithiol motifs have been shown to be inhibited by arsenite with dissociation constants of typically greater than 100 nM [4] [5]. In addition to arsenite, a number of monoalkyl and monoaryl As(III) species (including those previously studied in this laboratory) have been shown to strongly coordinate proteins and enzymes [5]. A number of these enzymes contain cysteine pairs contributed by redox-active disulfide motifs, including thioredoxins [6],
and their cognate reductases [7], dihydrolipoyl dehydrogenase [8] and glutathione reductase [9].

Previous work from this laboratory explored the interaction between the monoarsenicals (MMA, PSAO and arsenite) and the redox-active enzymes QSOX and PDI [10]. These enzymes, as discussed in Chapter 2, employ catalytically essential redox-active CxxC motifs that participate in oxidative protein folding. That study established that these monoarsenicals targeted unstructured cysteine rich reduced unfolded proteins in preference to the enzymes CxxC motifs directly. Further, the inclusion of 5 mM GSH was able to largely reverse the effects of these arsenicals on oxidative protein folding because the relatively high concentration of this monothiol provides for effective complexation of the monoarsenicals used [3].

This chapter explores the synthesis and characterization of simple bis- and tris-arsenical prototypes of reagents that might lead to more effective inhibition of enzymes with multiple CxxC motifs. Figure 3.1 shows the structure of three multivalent arsenicals synthesized here in comparison to the PSAO derivative widely used as a monoarsenical in our laboratory. During QSOX catalysis a flexible inter-domain linker allows the two catalytically essential CxxC motifs to share a mixed disulfide bond [11] and hence they are candidates for capture by small bis-arsenicals. Similarly the anticipated flexibility of PDI proteins [12] suggests that their CxxC motifs might be similarly captured by multivalent arsenicals (MVAs). Figure 3.2 provides cartoon depiction of the two enzymes used in this study.
Figure 3.1  **Structure of multivalent arsenicals.** The multivalent arsenicals (MVAs) synthesized in this work (BA-1; Bis arsenical-1, BA-2; Bis arsenical-2, TA-1; Tris arsenical-1). For comparison the structure of PSAO is also shown.
Design strategy for arsenic based inhibitors of QSOX and PDI.
Both QSOX and PDI contain four domains of similar size ~ 110 amino acids. The redox active CxxC motifs are present in first and fourth domain for both enzymes (panel A). Small multivalent arsenical molecules may be able to inhibit these enzymes by capturing the two active site CxxC motifs simultaneously (panel B) taking advantage of an inter-domain chelate effect.
3.2 Materials and Methods

3.2.1 General Methods and Materials

Ethylenediaminetetraacetic dianhydride, 1,3,5-benzenetricarbonyl trichloride, 3,3’,4,4’-biphenyltetracarboxylic dianhydride, guanidine hydrochloride (GuHCl), bovine pancreatic ribonuclease A (RNase A), bovine pancreatic insulin, thioflavin T, and EDTA were purchased from Sigma-Aldrich. 5,5’-dithiobis-2-nitrobenzoic acid (DTNB), dithiothreitol (DTT) and tris (2-carboxyethyl) phosphine hydrochloride (TCEP) were obtained from Gold Biotechnology. Size exclusion PD-10 columns were purchased from GE Healthcare and Congo Red was purchased from MP Biomedicals. All absorption spectra were collected on an Agilent 8453 UV/Vis spectrophotometer and the data were analyzed using ChemStation software. Experiments were performed at least in duplicate in 50 mM potassium phosphate buffer, pH 7.5 containing 1 mM EDTA unless otherwise noted. NMR was performed on a Bruker DRX 400 spectrometer. Mass spectroscopy was done on a QTOF Ultima instrument. Stock solutions of DTT, GSH and TCEP were standardized before use using DTNB [13]. Where appropriate the effectiveness of the arsenicals was compared at a constant concentration of 10 µM total As(III) (corresponding to 5 µM of BA-1 and -2 and 3.33 uM of TA-1 and 10 µM for PSAO, MMA and arsenite).
3.2.2 Synthesis of Arsenicals

Arsenic is classified as a human carcinogen and all arsenicals used here were handled with due care.

3.2.2.1 Synthesis of PSAO

Phenylarsine oxide (PAO) was synthesized as in Stevenson et al. [14] and converted to the more water soluble, and experimentally tractable analog PSAO using succinic anhydride [15].

3.2.2.2 Scheme for Synthesis of 2,2’-(ethane-1,2-diylbis((2-((4-arsinephenyl)amino)-2-oxoethyl)azanediyl))diacetic acid.

![Scheme for Synthesis of 2,2’-(ethane-1,2-diylbis((2-((4-arsinephenyl)amino)-2-oxoethyl)azanediyl))diacetic acid.]

Figure 3.3 Synthesis of Bis-Arsenical-1

PAO (236 mg, 1.01 mmol) was dissolved in 5 ml of dry DMF in a round bottom flask. Ethylenediaminetetraacetic dianhydride (100 mg, 0.39 mmol) was added to the reaction mixture and stirred overnight. The product was purified by reverse phase HPLC using a linear gradient of water with 0.1% formic acid to 90% acetonitrile containing 1% formic acid over 30 min at 1ml/min on a Phenomenex C18
column (5 µm, 250 x 4.60 mm, 300 Å). The fractions containing the product were collected and lyophilized. NMR (400 MHz D2O): \(^1\)H \(\delta\) (ppm); 3.39(4H), 3.47(4H), 4.00 (4H), 7.74(4H), 7.61(4H); \(^13\)C \(\delta\) (ppm); 50.6, 51.6, 54.8, 55.0, 56.7, 70.0, 107.9, 120.2, 126.8, 143.9, 147.6, 171.9; Expected Mass: 658.02, Mass found: 659.03.

### 3.2.2.3 Scheme for Synthesis of 3',4-bis((4-arsinephenyl)carbamoyl)-[1,1'-biphenyl]-3,4'-dicarboxylic acid

![Scheme](image)

**Figure 3.4 Synthesis of Bis-Arsenical-2**

PAO (201 mg, 1.01 mmol) was dissolved in dry DMF in a round bottom flask followed by addition of 100 mg (0.34 mmol) of 3,3',4,4'-biphenyltetra carboxylic dianhydride to the reaction mixture and stirred overnight. The reaction was monitored by TLC (run using pure methanol) and the final product was separated from the reaction mixture by trituration with ethyl acetate. The product was obtained as a yellow solid. NMR (400 MHz ,CD3OD); 1H \(\delta\) (ppm) 7.52 (4H), 7.66 (1H), 7.74 (4H), 7.78(1H), 7.86(1H), 7.88(1H), 8.07(2 H); 13C\(\delta\) (ppm) 121.17, 129.83, 129.22, 129.13, 127.34, 127.54, 131.74, 132.34, 141.39, 142.52, 144.04, 158.21, 168.09, 170.68 ; Expected Mass: 695.97, Mass obtained:696.93. BA-2 was a mixture of
regioisomers (3’, 3,4’ and 4,4’) which could not be readily separated chromatographically or by crystallization and was therefore used as is.

3.2.2.4 Strategy for Synthesis of 1,3,5-((benzenetricarbonyltris(azanediyl)) tris(benzene-4,1-diyl))triarsineous acid

Figure 3.5 Synthesis of Tri-Arsenical-1

PAO (298 mg, 1.48 mmol) was dissolved in 5 ml of dry acetone in a dry round bottom flask followed by the addition of 100 mg (0.37 mmol) of 1,3,5-benzenetricarbonyl trichloride. The reaction mixture was allowed to stir overnight. The insoluble product was removed by filtration and washed with acetone to remove excess PAO. NMR (400 MHz, CD3OD): \textsuperscript{1}H δ (ppm); 8.53(3H), 7.79 (6H), 7.59 (6H); \textsuperscript{13}C δ (ppm); 121.82, 124.26, 131.96, 133.18, 137.29, 167.12; Expected Mass: 758.92, Mass obtained: 759.89

3.2.3 Expression, Purification and Handling of Proteins

PDI was purified as Rancy et al. [16]. Avian QSOX and chicken riboflavin binding protein (RfBP) were prepared as described previously [17] [18] and were
generous gifts from Dr Karen Hoober and Dr Harold B. White, III respectively.

3.2.4 Preparation of Reduced Proteins

PDI was incubated with a 40-fold molar excess DTT at 25°C for 1 h in 50 mM potassium phosphate buffer, pH 7.5 containing 1 mM EDTA. The reduced enzyme was freed from excess reductant by size exclusion chromatography using a PD-10 column equilibrated with the same buffer without DTT and stored at -20°C until further use. Bovine pancreatic RNase A was incubated for 2 h at 37°C with an 80-fold molar excess of DTT in 100 mM Tris buffer, pH 8.0, containing 1 mM EDTA and 6 M GuHCl. The reduced denatured protein was separated from excess reductant and GuHCl by using a PD-10 column equilibrated with 10 mM sodium acetate buffer, pH 4.0 containing 1 mM EDTA and stored anaerobically under nitrogen at 4°C. Lyophilized RfBP was incubated for 2 h at 37°C with an 180-fold molar excess of DTT in 100 mM Tris buffer, pH 8.0, containing 1 mM EDTA and 6 M GuHCl. The reduced and denatured protein was separated by a PD-10 column equilibrated with the same buffer without the reductant and stored anaerobically under nitrogen at 4°C. Reduced protein concentrations were determined using the following extinction coefficients: RNase $\varepsilon_{278} = 9300 \text{ M}^{-1} \text{ cm}^{-1}$; RfBP $\varepsilon_{280} = 49,000 \text{ M}^{-1} \text{ cm}^{-1}$; reduced PDI $\varepsilon_{280} = 56,400 \text{ M}^{-1} \text{ cm}^{-1}$. Thiol titers were determined using 5-thio-2-nitrobenzoic acid anion, $\varepsilon_{412} = 14,150 \text{ M}^{-1} \text{ cm}^{-1}$. 
3.2.5 Oxidative Protein Folding Assay using RfBP

The refolding assay was performed by incubating 1 µM reduced and denatured chicken RfBP (18 µM free thiols) with 30 nM QSOX and 30 µM reduced PDI in the presence of 0.8 µM free riboflavin. The loss of riboflavin fluorescence over time was monitored for 60 min (excitation at 450 nm; 2 nm slit-width, and emission at 530 nm; 16 nm slit-width). Potential arsenicals inhibitors were added to the assay mixture where appropriate.

3.2.6 Oxidative Protein Folding Assay using RNase

Reduced unfolded RNase A (10 µM protein, 80 µM free thiols) was added to a mixture of 5 µM reduced PDI in 0.4 mL of a redox buffer comprised of 1 mM/0.2 mM or 5 mM/1 mM of reduced and oxidized glutathione respectively in 50 mM Tris buffer, pH 7.5, containing 1 mM EDTA, with or without 10 µM As(III) species. Aliquots of 60 µL were removed every 4 min and mixed with 60 µL of 2 mM cCMP solution and the RNase activity followed the hydrolysis of cCMP to CMP at 296 nm for 2 min.

3.2.7 Turbidometric Insulin Reductase Assay

Insulin solutions (50 µM) were mixed with either 100 µM TCEP or 5 mM GSH as reducing agents in the presence of arsenicals. The reaction was started by the addition of 1 µM PDI. The increase in light scattering was followed at 600 nm.
Control reactions were conducted under the same conditions in the absence of arsenicals. Lag times for the onset of turbidity were calculated as described previously [19].

### 3.2.8 Binding of Arsenicals to Reduced PDI

The binding of multivalent arsenicals to PDI was assessed by an increase in absorbance at 300 nm. The spectra were recorded 20 min after each addition, and the data were fit to a binding equation as described previously [10].

### 3.2.9 QSOX Activity Assay

Avian QSOX was assayed at 30 nM with either 5 mM TCEP or 5 mM GSH as its substrates in a Clark-type oxygen electrode in the presence of arsenicals as inhibitors. Enzyme activities are reported as a percentage of control experiments under the same conditions without arsenicals.

### 3.2.10 COPASI Simulation

The partition between dithiols and GSH complexes of monoarsenical MMA was simulated using COPASI [20] using the following equation.

\[
\begin{align*}
\text{MMA-dithiol} + 2 \text{H}_2\text{O} &\rightleftharpoons \text{MMA} + \text{dithiols} & K_d &= 10^{-5} \text{ to } 10^{-12} \text{ M} \\
\text{MMA-(SG)} + \text{H}_2\text{O} &\rightleftharpoons \text{MMA} + \text{GSH} & K_d &= 6.25 \times 10^{-5} \text{ M} \\
\text{MMA-(SG)}_2 + \text{H}_2\text{O} &\rightleftharpoons \text{MMA-(SG)} + \text{GSH} & K_d &= 5.26 \times 10^{-4} \text{ M}
\end{align*}
\]
Forward and reverse rate constants were chosen to give these $K_d$ values and the reaction was simulated to equilibrium starting with 10 μM of MMA and dithiols and with 5 mM GSH. The dissociation constant between MMA and GSH is reported from [3].

3.2.11 Thioflavin T Assay

Reduced RNase A (30 μM in 50 mM potassium phosphate buffer, pH 7.5, containing 1 mM EDTA) was mixed with 5 μM thioflavin T dye and multivalent arsenicals in the presence and absence of 5 mM GSH. TCEP (1 mM) was added to reaction mixtures without GSH to suppress the formation of any adventitious protein disulfides. The increase in fluorescence was monitored over time (excitation at 450 nm, emission at 485 nm). Control reactions were run under the same conditions in the absence of arsenicals.

3.2.12 Congo Red Spectroscopic Assay

Reduced RNase A (20 μM in 50 mM potassium phosphate buffer, pH 7.5, containing 1 mM EDTA) was incubated with 10 μM Congo Red dye and arsenical compounds for 30 minutes at room temperature in the absence and presence of 5 mM GSH. The spectrum of each reaction mixture was subtracted from spectrum of the same concentration of Congo Red in the absence of protein to observe the spectral shift at 540 nm.
3.2.13 Transmission Electron Microscopy

Imaging was done on a Zeiss LIBRA 120 transmission electron microscope equipped with a Gatan ultrascan 1000 2k x 2k CCD camera at the Delaware Biotechnology Institute. Reduced RNase A (50 µM) and arsenicals were incubated overnight in the absence and presence of 5mM GSH, the samples without GSH were maintained in the presence of 1 mM TCEP to stop the formation of any disulfides. Aliquots of 4 µL from samples were placed on Cu grids with a mesh size of 400. The excess liquid was wicked of with filter paper and the grid was washed 3 times with water before staining with 1% uranyl acetate. The grid was again wicked with filter paper to remove excess dye and air-dried for 1 h before taking the image.

3.3 Results and Discussion

Multivalent-arsenicals (MVAs) were obtained as stable solids. The stock solution were prepared in phosphate buffer or methanol and stored at 4°C prior to use. Concentrations of these reagents were determined gravimetrically and were confirmed via titration with a standardized solution of DTT by following the increase in absorbance in the near UV region of the spectrum that accompanies coordination of the arsenoxide species with thiols (Figure 3.6). These compounds were then tested as inhibitors of redox-active enzymes by monitoring their effect on the oxidative folding of RfBP and RNase.
Figure 3.6  **Quantification of MVAs using DTT titration.** The concentration BA-1 (panel A), BA-2 (panel B) and TA-1 (panel C) was established via DTT titration in 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA. Spectra were recorded immediately after each addition. The increase in absorbance was followed at 300 nm for BA-1, 320 nm for BA-2 and 330 nm for TA-1 and was plotted as a function of increasing concentration of DTT.
3.3.1 Effect of MVA's in Oxidative Protein Folding Assay using RfBP

The oxidative folding of RfBP can be achieved \textit{in vitro} using nanomolar concentrations of QSOX and micromolar levels of reduced PDI in aerobic solution [16]. The reduced denatured RfBP is converted to an apoprotein that is capable of rapid and stoichiometric binding of riboflavin with complete quenching of the characteristic fluorescence of the free vitamin [16]. Here the initial disulfide generation is catalyzed by QSOX and mispairings are corrected by PDI in its reduced form. A previous study done in this laboratory showed that monoarsenicals like arsenite, MMA and PSAO profoundly attenuated the ability of RfBP to bind riboflavin. However further studies showed that this effect did not reflect the direct inhibition of QSOX or PDI but was largely due to the unexpected ability of these monoarsenicals to capture the reduced conformationally mobile RfBP [10].

The inclusion of multivalent-arsenicals at a concentration of 10 µM As(III) species (either as 5 µM of the bis arsenicals BA-1 and BA-2, or 3.33 µM of TA-1) leads to the almost complete inhibition of riboflavin binding as shown by the fluorescence monitored with time (Figure 3.7; panel A). Under these conditions, the multivalent-arsenicals were comparably effective as monoarsenicals MMA, PSAO and arsenite at 10 µM concentration (Figure 3.8; panel A). However when oxidative folding mixtures were supplemented with 5 mM GSH, a concentration that is likely to be encountered intracellularly [21], the monothiols were notably less effective at inhibiting oxidative folding (Figure 3.8; panel B) whereas BA-1, BA-2 and TA-1 still
inhibited the rebinding of riboflavin almost completely (Figure 3.7; panel B). The monothiol glutathione has been shown to bind surprisingly tightly to monoalkylated As(III) species [5] [10] and hence lowers the concentration of arsenical available to bind its protein targets. Thus the data here show that the multivalent arsenicals are much more effective at inhibiting oxidative folding than their monoarsenical counterparts in the presence of competing glutathione (see discussion, section 3.4).
Figure 3.7  Inhibition of oxidative folding of reduced RfBP in the presence of MVAs. Oxidative protein folding of RfBP was monitored by the loss of riboflavin fluorescence as active holoprotein is generated by inclusion of 30 nM QSOX and 30 µM reduced PDI. The multivalent arsenicals strongly suppress oxidative folding under these conditions (panel A) and are comparably effective even in the presence of 5 mM (panel B). Control (no arsenical), black; BA-1 (5 µM), pink; BA-2 (5 µM), blue; TA-1 (3.33 µM), green.
Figure 3.8  **Comparison of monoarsenicals and MVAs as inhibitors of oxidative protein folding of RfBP.** Monoarsenicals PSAO, MMA and arsenite and the multivalent arsenicals (all at an aggregate As(III) concentration of 10 µM) strongly suppress oxidative folding in the absence of 5 mM GSH (panel A). In the presence of 5 mM GSH while the MVAs remain effective inhibitors, the monoalkyl and monoaryl arsenicals at 10 µM concentration become considerably less potent (panel B). Data for monoarsenicals taken from [10]. The continued strong inhibition of oxidative protein folding by arsenite when challenged with 5 mM GSH (far right, panel B) reflects the comparatively weak binding of glutathione to arsenite [3] [10].
3.3.2 Effect of Bis- and Tris-Arsenicals on the Oxidative Refolding of RNaseA

In a widely-used model for oxidative protein folding, reduced unfolded RNase is added to a glutathione redox buffer containing PDI to accelerate the oxidation of cysteine residues and also corrections of any mispaired disulfides [22]. Samples are then removed and assayed discontinuously for RNase enzymatic activity. In this system GSSG, not the oxidase QSOX, drives disulfide bond generation in the client protein. The multivalent-arsenicals prove rather comparable in inhibiting regain of RNase activity using either 1 mM GSH paired with 0.2 mM GSSG (where the redox poise is ~ -172 mV) or using 5 mM GSH and 1 mM GSSG (poised at ~ -193 mV). In the presence of multivalent arsenicals refolding is completely inhibited after about 8 minutes with either 1 mM GSH and 0.2 mM GSSH or 5 mM GSH and 1 mM GSSG and is limited to 30% or 40% recovery of RNase activity respectively (Figure 3.9). In contrast PSAO was unable to provide complete inhibition and was only able to slow the refolding process in both cases (Figure 3.9).
Inhibition of oxidative refolding of RNase by arsenicals. Reduced and denatured RNase (10 μM, 80 μM thiols) was incubated with 5 μM reduced PDI and a redox buffer containing 1 mM GSH and 0.2 mM GSSG (panel A) or 5 mM GSH and 1 mM GSSG (panel B). Multivalent arsenicals were able to inhibit the refolding of RNase in both cases while PSAO slows down the refolding but was unable to achieve complete inhibition. Control (no arsenical), black; BA-1 (5 μM), pink; BA-2 (5 μM), blue; TA-1 (3.33 μM), green; PSAO (10 μM), grey.
3.3.3  **Bis- and Tris-Arsenicals as Inhibitors of QSOX**

The above data showed strong inhibition of the refolding of a protein with complex disulfide connectivity (RfBP; 9 disulfides with 34 million possible combinations) in the presence of PDI and QSOX and marked inhibition of a much simpler refolding system involving RNase with only 4 disulfides (105 possible combinations) using PDI in the presence of a glutathione redox buffer. These data suggested that QSOX was probably not the major target of these multivalent-arsenicals. To address this issue, the effects on BA-1, BA-2 and TA-1 were tested on QSOX1 activity in isolation. Two alternate substrates of the oxidase were utilized; the first was 5 mM TCEP which is a phosphine-based model substrate and does not compete with free As(III). In this case a modest inhibition of QSOX was achieved. Next, when 5 mM GSH was used as the substrate that serve to compete for free As(III), QSOX activity was not significantly inhibited as depicted in Figure 3.10, panel A. These data further demonstrate that the inhibition of oxidative folding seen with RfBp is not principally due to QSOX.
Figure 3.10  Inhibition of QSOX reactivity by MVAs and monoarsenicals. Avian QSOX (30 nM) was assayed in a Clarke-type oxygen electrode with either 5 mM TCEP or 5 mM GSH in the presence of multivalent arsenicals (panel A) or monoarsenicals (panel B). Data for monoarsenicals (PSAO, MMA and arsenite) taken from [10].
3.3.4 Interaction between Multivalent Arsenicals and PDI

The standard reductase assay [23] was used to assess the potency of the multivalent arsenical as inhibitors of PDI. Reduced PDI drives isomerase-catalyzed reduction of the interchain disulfides of insulin with consequent accumulation of the weakly soluble B-chain of insulin. After a reproducible lag phase, a threshold concentration is reached leading to the appearance of B-chain aggregates and the onset of light scattering [24]. A marked delay in aggregation time is evident with all three multivalent arsenicals when the non-As(III)-coordinating reductant TCEP was used in this turbidometric assay (Figure 3.11; panel A1). In contrast the monoarsenical PSAO was ineffective at extending the lag time.

It is important to note that the delayed aggregation observed in panels A1 of Figure 3.11 could be due, in principle to either inhibition of PDI or a direct interference in the aggregation of the insulin B-chain by the multivalent arsenicals. Results shown in Figure 3.12 allows for the latter possibility to be excluded as a reason for delayed aggregation times.

Although the multivalent arsenicals provide significant inhibition of PDI using TCEP, replacement of this non-coordinating reductant by 5 mM GSH results in only marginal inhibition of the isomerase by BA-1, BA-2 and TA-1 (Figure 3.11; panel B1). These data show that while the multivalent arsenicals are more effective inhibitors of PDI than the monoarsenical PSAO when measured in the TCEP-driven reductase assay (Figure 3.11; panel A2), none of these arsenicals provide significant inhibition when 5 mM GSH is used 1 (Figure 3.11; panel B2).
Figure 3.11  **The effect of multivalent arsenicals on the reductase activity of PDI.** Porcine insulin (50 µM in 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA) was mixed with either 100 µM TCEP (Panel A1) or 5 mM GSH (Panel B1) in the absence or presence of arsenicals. The onset of turbidity was observed at 600 nm. The lag times obtained with different arsenicals are represented in the bar graph (panels A2 and B2). Control (no arsenical), black; BA-1 (5 µM), pink; BA-2 (5 µM), blue; TA-1 (3.33 µM), green; PSAO (10 µM), grey.
Figure 3.12  **MVA compounds do not significantly delay the onset of insulin B-chain aggregates.** Insulin (50 µM in 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA) was mixed with 5 mM TCEP in presence of MVAs. The onset of turbidity was observed at 600 nm (panel A). The MVAs do not significantly delay the development of light scattering in the reductase assay in the absence of reduced PDI. The lag times are also presented as a bar graph in panel B. Control (no arsenical), black; BA-1 (5 µM), pink; BA-2 (5 µM), blue; TA-1 (3.33 µM), green.
Since the dissociation constant between reduced PDI and the monoarsenical PSAO was determined previously as 1.1 µM [10], we then evaluated the binding of the phenylarsenoxide moiety in the context of bis- and tris arsenical derivatives (Figures 3.13 and 3.14). It should be noted that the CxxC motifs of the a and a’ thioredoxin domains of mammalian PDI have very similar active site sequences and essentially identical redox potentials [25] [26]. As might be expected, there was no evidence for heterogeneity in binding when PSAO was titrated with PDI containing these two reduced CxxC motifs. To facilitate an accurate quantitation of binding stoichiometry between the MVAs with reduced PDI, a relatively high concentration of reduced PDI was used. The titration with BA-1 yielded a stoichiometry of 1, suggesting that the two reduced CxxC motifs are captured by a single molecule of BA-1 (Figure 3.13, panel A). It should be noted that a recent paper shows that PDI can dimerize [27] and we cannot currently exclude the possibility that the stoichiometry observed here reflects arsenical bridging CxxC motifs on the two PDI chains. Whatever the origin of this stoichiometry, BA-1 inhibits PDI in the presence of TCEP (Figure 3.11; panel A1). The spectrophotometric titration was then repeated at a lower PDI concentration to permit a more accurate determination of dissociation constant. The increase in absorbance can be fit to a $K_d$ of 61.5 ± 5.3 nM for BA-1 with a stoichiometry of 1.01 ± 0.09 (Figure 3.14, panel A).

Comparable experiments with BA-2 gave a stoichiometry of 1.02 ± 0.05 and a $K_d$ of 62.5 ± 3.9 nM (Figure 3.14, panel B). These dissociation constants are significantly lower than PSAO reflecting an approximately 17- to 18-fold increase in
affinity for the bis arsenicals BA-1 and BA-2. These data are consistent with the expression of a significant chelate effect with the bis-arsenicals although the strength of the interaction is insufficient to render them effective inhibitors of PDI at the glutathione concentrations of several millimolar that prevail intracellularly (see discussion, section 3.4). Experiments with TA-1 gave a $K_d$ of $64.2 \pm 2.2$ nM and a stoichiometry of $1.0 \pm 0.05$ (Figure 3.14, panel C) with reduced PDI. Evidently one of the three arsenical sites goes unused.
Figure 3.13 Binding stoichiometry of MVAs with reduced PDI. Reduced PDI (10 µM) was titrated with 2 µM aliquots of BA-1 (panel A), BA-2 (panel B) and TA-1 (panel C) in 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA. The increase in absorbance was followed at 300 nm for BA-1, 320 nm for BA-2 and 330 nm for TA-1. A clear end-point was observed at 10 µM concentration of PD1 with all the MVAs establishing the binding stoichiometry to be 1 MVA per reduced PDI.
**Figure 3.14**  *Kₐ determination of reduced PDI with MVAs.* Titrations of 1 µM reduced PDI were performed with increasing concentration of BA-1 (panel A), BA-2 (panel B) and TA-1 (panel C) in 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA and was followed by increase in absorbance at 300 nm. Data were fit according to [10]. All titrations were duplicated yielding the *Kₐ* of 61.5 ± 5.3 nM, 62.5 nM ± 3.9 and 64.2 ± 2.2 nM for BA-1, BA-2 and TA-1 respectively.
3.3.5 Interaction between Multivalent-Arsenicals and RfBP

The refolding experiments with RfBP and RNase show that oxidative folding is very strongly attenuated by the multivalent inhibitors studied here although neither QSOX nor PDI seem to be strongly inhibited by these arsenicals in the presence of 5 mM glutathione. Titration with reduced, denatured RfBP confirms that these multivalent-arsenicals bind avidly to unstructured reduced unfolded proteins as shown in Figure 3.15. Reduced RfBP has 18 free thiols and hence a stoichiometry of ~ 4.5 would be expected for both BA-1 and BA-2 per reduced RfBP (18 thiols bind 9 monoarsenicals or 4.5 bis arsenicals). Evidently both arsenical moieties in BA-1 and BA-2 can each capture a pair of thiols in reduced RfBP. In contrast the tris-arsenical shows a stoichiometry comparable to that of the bis-arsenicals suggesting that one arsenic site on average remains unused in each TA-1 molecule as observed previously with reduced PDI. Presumably the steric requirements of coordinating 6 cysteine peptide thiols preclude a significant contribution of this binding mode. In all cases the sharp endpoint of the titrations suggests that the effectively multivalent arsenical reagents bind very tightly to reduced RfBP. This binding appears to be tighter, on average than that of PSAO, because glutathione can effectively reverse the inhibition of the refolding assay with mono alkyl and aryl arsenicals but is unable to significantly reverse the inhibitions with the multivalent compounds (see discussion, section 3.4). Similar titrations with RNase could not be performed because addition of these MVAs led to a rapid onset of aggregation. These data are discussed in the following section.
Figure 3.15  **Titration of reduced RfBP with arsenicals.** Reduced denatured RfBP (1 μM, corresponding to 18 μM thiols) was titrated with increasing concentration of MVAs in 50 mM phosphate buffer pH 7.5, containing 1 mM EDTA. The absorbance at 300 nm is plotted as a function of added MVAs per RfBP. The dashed lines correspond to the end points of the titrations and range between 4.3 and 4.9 MVA molecules per RfBP. To ensure the reaction was complete, the spectrum was recorded 20 min after each addition. BA-1 (5 μM), pink; BA-2 (5 μM), blue; TA-1 (3.33 μM), green.
3.3.6 Fibril Formation with Reduced RNase and Multivalent-Arsenicals

Although the detailed titrations of MVAs with RfBP could be performed, comparable experiments with reduced RNase proved unworkable because of the onset of light scattering as the titration progressed. It should be noted that the induction of turbidity was noted earlier when the titrant was MMA [10]. Somewhat surprisingly, the formation of turbidity was not observed with the two other mono arsenicals (arsenite and PSAO) and hence aggregation is not universal consequence of thiol coordination by arsenicals. In contrast all three MVAs promoted formation of RNase aggregates that were insoluble and readily detected using thioflavin T fluorescence measurements (Figure 3.16; panel A) [28] [29]. Importantly these aggregates are still prominent with the inclusion of 5 mM GSH (Figure 3.16; panel B), whereas they do not form when solutions of reduced RNase and GSH are challenged with MMA [10]. Spectrophotometric experiments with congo red [30] are also consistent with an association of unfolded reduced RNase into structures with significant $\beta$-sheet content. Figure 3.17 shows the expected red-shift that occurs in the presence of multivalent arsenicals that is not observed with reduced protein alone. Again these effects are observed in both the absence and presence of GSH (Figure 3.17; panel A and B respectively). Finally the formation of fibrillar like structure of these aggregates was confirmed by TEM imaging [31] both in the absence (Figure 3.18) and presence of 5 mM GSH (Figure 3.19). Overall it appears that these MVAs are effective at promoting fibrillar and beta rich structure formation.
Figure 3.16  Aggregation of reduced RNase with arsenicals followed with Thioflavin T. Reduced RNase (30 µM) was mixed with 5 µM thioflavin T dye and MVAs in the absence (panel A) and presence (panel B) of 5 mM GSH. The increase in fluorescence was monitored over time at 485 nm. No significant increase in fluorescence was observed without the inclusion of the multivalent arsenicals in both cases. Control (no arsenical), black; BA-1 (5 µM), pink; BA-2 (5 µM), blue; TA-1 (3.33 µM), green.
Figure 3.17  Congo red spectral shift assay with reduced RNase and MVAs. Reduced RNase (20 µM) was incubated with 10 µM Congo Red dye and MVAs for 30 minutes at room temperature in 50 mM phosphate buffer pH 7.5, containing 1 mM EDTA in the absence (panel A) and presence of 5 mM GSH (panel B). The spectral shift at 545 nm indicates fibril like formation of RNase in presence of MVAs. Control (no arsenical), black; BA-1 (5 µM), pink; BA-2 (5 µM), blue; TA-1 (3.33 µM), green.
Figure 3.18  TEM of multivalent arsenicals with reduced RNase. Reduced RNase (50 μM in 50 mM potassium phosphate buffer, pH 7.5, containing 1 mM EDTA and 1 mM TCEP) was incubated with 5 μM BA-1 (panel A), BA-2 (panel B) and 3.33 μM of TA-1 (panel C) overnight at 4°C. TEM imaging shows fibril formation with all multivalent arsenicals. Reduced RNase in absence of any arsenical does not show fibril like structure (panel D).
Figure 3.19  TEM of multivalent arsenicals with reduced RNase in presence of 5 mM GSH. Reduced RNase (50 μM in 50 mM potassium phosphate buffer, pH 7.5, containing 1 mM EDTA and 5 mM GSH) was incubated with 5 μM BA-1 (panel A), BA-2 (panel B) and 3.33 μM of TA-1 (panel C) overnight at 4°C. TEM imaging shows fibril formation with all multivalent arsenicals. Reduced RNase in absence of any arsenical does not show fibril like structure (panel D).
3.4 Discussion and Conclusion

This work explored the design and implementation of small molecule multivalent-arsenicals on redox active enzymes containing multiple CxxC motifs. The three reagents BA-1, BA -2 and TA- 1 were readily synthesized by coupling 4-aminophenyl arsenoide with commercially available anhydrides or acyl chlorides in a one-pot reaction. These prototypical reagents showed promise as inhibitors of enzymes with multiple redox active disulfide motifs. These compounds were successfully able to inhibit the oxidative folding pathways of RfBP and RNase even in the presence of 5 mM GSH. The principal targets of these compounds were reduced unfolded protein substrates and not the CxxC motifs of QSOX or PDI.

A striking difference between the prior work with monoarsenicals and these current studies is that the multivalent arsenicals appear to bind avidly to reduced unfolded proteins so that they cannot be displaced by the millimolar levels of GSH that would prevail intracellularly. The studies of Wilcox and colleagues [3] allow general conclusions to be drawn concerning the competition between a dithiols and GSH for an arsenical center. Here we use MMA since the data of Spuches et al. [3] allow calculation of first and second dissociation constants for GSH (as depicted in the blue boxed area in Figure 3.20 , panel A).
Figure 3.20 Interaction of MMA with dithiols in presence of GSH. Panel A represents the complexation between 10 µM MMA and 10 µM dithiols and 5 mM GSH. Panel B represents the percentage of complexation between MMA and a hypothetical dithiols (green curve; exhibiting a binding with a range of $K_d$ values) in the presence of 5 mM GSH.

We now consider the competition for 10 µM MMA between 5 mM GSH and a hypothetical dithiols ligand such as that contributed by a CxxC motif showing $K_d$ values between $10^{-5}$ to $10^{-12}$ M. At a $K_d$ of $10^{-6}$ M only 1% of MMA will be complexed by 10 µM dithiols in competition with 5 mM GSH. A $K_d$ value of 6 nM would be required to secure 50% of MMA in a dithiol complex. A corresponding dissociation constant of 10 pM would result in 97% MMA complexed by the dithiol. These data
provide a rationalization for a number of observations. Thus while the multivalent
arsenicals bind considerably more tightly to reduced PDI than the monoarsenicals (e.g.
PSAO, $K_d = 1.1 \mu M$ [10]; BA-1, $K_d = 62 nM$) neither class of compounds are effective
inhibitors of the isomerase in presence of 5 mM GSH (with a prediction of 0% and
15% of PDI complexed with the arsenical respectively).

Given these thermodynamic constraints, the observations that multivalent
arsenicals remain effective inhibitors of oxidative protein folding of RfBP even in the
presence of 5 mM GSH suggests that they bind with low nanomolar to sub-nanomolar
$K_d$ values with the reduced RfBP. Further, the failure of 5 mM GSH to suppress or
reverse the aggregation of reduced RNase induced by low concentrations of
multivalent arsenicals is in contrast to the reversal of the effect when monoarsenical
MMA was used [10].

Thus the prevailing intracellular concentration of GSH requires that a target
dithiol motif bind to an arsenical with a $K_d$ of less that about 10 nM to be significantly
complexed. Extracellularly this requirement is greatly relaxed because the prevailing
concentration of GSH, and other small molecular weight monothiols, is likely to be
considerably less than 100 $\mu M$ [21]. Thus multivalent arsenicals may find utility as
inhibitors of extracellular oxidoreductases with two or more dithiols motifs.
REFERENCES


Chapter 4

AN ARSENICAL-MALEIMIDE FOR THE GENERATION OF NEW TARGETED BIOCHEMICAL REAGENTS

4.1 Introduction

The inability of small arsenical molecules to inhibit redox-active CxxC motifs as discussed in previous chapter raises the question of their specificity. To overcome this limitation we decided to use natural substrates of the redox-active enzymes as a vehicle for delivering arsenic. A continuing challenge in the efficient targeting of arsenicals is minimization of extraneous labeling and the toxicity associated with the many off-target thiol-containing proteins. For example, GSAO (Figure 4.1) uses the tripeptide glutathione as a vehicle for arsenical delivery via conjugation with 4-(2-bromoacetylamino)phenylarsonous acid [1] [2]. However, the bromoacetyl function reacts rather slowly under conditions typically used in thiol-bioconjugation reactions. Hence we sought a simple arsenical reagent that could be introduced rapidly and stoichiometrically at cysteine residues on peptides and proteins. Maleimides were chosen for this purpose because they exhibit high specificity towards cysteine residues and can be introduced rapidly and stoichiometrically, at pH range of 6.5-7.5 [3] [4]. Thiol - maleimide conjugation reaction has been an indispensable tool for biochemical
studies and is used extensively for conjugating cytotoxic agents to antibodies [3] [4] nanoparticles [7] [8], and dendrimers [9] [10]. This Chapter introduces an As(III) containing maleimide, As-Mal (Figure 4.1) that can be readily conjugated to exposed cysteine residues in proteins and peptides. This Conjugation provides a scaffold for directing the As(III) species to the target proteins for their inhibition.

Figure 4.1 Structures of 4-(N-(S-glutathionylacetyl)amino) phenylarsonous acid (GSAO) and arsenic-maleimide (As-Mal).
4.2 Methods and Materials

4.2.1 General Methods and Reagents

UV-Visible spectra were recorded on an Agilent 8453 diode-array instrument. Unless otherwise stated, enzyme assays were performed in 50 mM potassium phosphate buffer, pH 7.5, containing 1 mM EDTA. Enzyme assays and binding measurements were repeated at least in duplicate, or as noted. NMR spectra were performed on a Bruker AV 400. Mass spectroscopy utilized a QTOF Ultima instrument. Guanidine hydrochloride (GuHCl), reduced glutathione (GSH), bovine pancreas insulin, bovine pancreas ribonuclease A and NADPH were purchased from Sigma-Aldrich. 5, 5′-dithiobis (2-nitrobenzoate) (DTNB), dithiothreitol (DTT) and tris(2-carboxyethyl)phosphine (TCEP) were purchased from Gold Biotechnology Inc. Stock solutions of DTT, TCEP and GSH were standardized before use using DTNB [11]. Desalting PD-10 columns were purchased from GE Healthcare.

4.2.2 Synthesis of Arsenical Reagents

Arsenic is classified as human carcinogen and all arsenicals used here were handled with due care. p-Aminophenylarsenoxide (PAO) was synthesized as in Stevenson et al. [12]. Monomethylarsenous acid (MMA) and \( p \)-succinylamidophenylarsenoxide (PSAO) were made according to Cline et al. [13].
4.2.2.1 Synthesis of Arsenical-Maleimide (As-Mal), 5 (4-(2, 5-dioxo-2, dihydro-1 \(H\)-pyrrol-1-yl) phenylarsenous acid)

Figure 4.2 Synthesis of arsenical-maleimide (As-Mal)

PAO-maleamic acid was synthesized from 100 mg (0.49 mmol) of PAO dissolved in 5 mL of dry THF in a round bottom flask. Maleic anhydride (243 mg, 2.47 mmol) was added and the reaction and stirred at room temperature for 1 h. The insoluble product (PAO-maleamic acid) was removed by filtration, freed of excess maleic anhydride by washing with dry THF, and used without further purification (yield 144 mg, 95% from PAO). NMR (400 MHz \(\text{CD}_3\text{SOCD}_3\)): \(^1\text{H} \delta \) (ppm); 7.71 (d, 2H, \(J = 8\)Hz), 7.60 (d, 2H, \(J = 8\)Hz), 6.47(d, 1H, \(J = 12\) Hz), 6.32(d, 1H, \(J = 11.6\) Hz), 10.50(s, 1H), 12.96 (br), 3.43 (br); \(^{13}\text{C} \delta \) (ppm): 167.04, 163.44, 145.01, 140.41, 131.45, 130.48, 129.27, 119.14. Expected mass (as diol): 298.9774, Mass found: 298.9773.

Synthesis of As-Mal: 130 mg (0.43 mmol) of PAO-maleamic acid was added to a dry round bottom flask and dissolved in 5 ml acetic anhydride, followed by the addition of 107 mg of sodium acetate (3 equivalents). The reaction mixture was heated for 1 h at 120°C. The reaction was cooled, sodium acetate was filtered off and the
remaining acetic anhydride was removed by distillation followed by azeotroping with benzene to yield As-Mal as a yellow solid (98 mg, 80 % from 4). NMR (400MHz CD$_3$CN): $^1$H $\delta$ (ppm): 7.90 (d, 2H, j= 8.4Hz), 7.49 (d, 2H, j= 8.4 Hz) and 6.93 (s, 2H), 2.13 (br). $^{13}$C $\delta$ (ppm): 169.35, 148.92, 134.10, 133.41, 129.92, and 126.15. Expected mass (as arsenoxide): 262.9563, Mass found: 262.9563. Since $N$-phenylmaleimides are prone to ring-opening in solution, As-Mal was freshly dissolved in DMSO or 1 mM HCl before starting each experiment.

4.2.3 Stopped Flow Spectrophotometry

Absorbance changes at 320 nm were followed upon mixing equal volumes of 200 µM of GSH in 100 mM phosphate buffer, pH 7.5, containing 1 mM EDTA, with 200 µM As-Mal or $N$-phenylmaleimide in 1 mM HCl in a SF-61 DX2 double mixing stopped-flow (TgK Scientific) at 25°C. Data were collected in monochromator mode and analyzed using second-order kinetics with the KinetAsyst 3 software.

4.2.4 Mass Spectroscopic Analysis of Conjugated Proteins

As-Mal labeled proteins were chromatographed at 0.2 mL/min on a Phenomenex C18 column (5 µm, 250 x 4.60 mm, 300 Å) using a linear gradient from 0.1% formic acid in water to 90% acetonitrile with 0.1% formic acid over 60 min at 25°C and analyzed using a Waters LC-MS Q-Tof Ultima instrument in positive ion mode. Data were processed using MassLynx MaxEnt1 software.
4.2.5 Expression, Purification and Handling of Proteins

*Escherichia coli* thioredoxin (Trx) was obtained as in Tamura et al. [14]. *E. coli* thioredoxin reductase was a generous gift from Dr. Sharon Rozovsky of the University of Delaware. Human protein disulfide isomerase (PDI) was prepared according to Rancy et al. [15]. The septuple RNase mutant, retaining a single cysteine (the 4th cysteine, C4), was purified as in Codding et al. [16]. Additional RNase mutant proteins (the C1, 4; C1, 8 and C4, 8 di-cysteine derivatives, and the 1, 4, 8 tri-cysteine proteins) were obtained using QuikChange mutagenesis (Stratagene). Solutions of insulin were prepared as described previously [17].

4.2.6 Preparation of Reduced Proteins and Labeling with Arsenical-Maleimide

The C35S mutant of *E. coli* Trx was treated with a 10-fold molar excess DTT at 4°C for 2 h to ensure the absence of disulfide-bridged dimers of this single cysteine-containing protein. Excess DTT was removed using a PD-10 column equilibrated with 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA. The mutant protein was then incubated for 2 h with a 2-fold excess As-Mal at 4°C. Excess arsenical was removed using a PD-10 column as before. The labeled protein was stored at -20°C.

To remove any adventitious disulfides, all RNase mutant proteins were pre-incubated for 1 h at 37°C with a 10-fold molar excess of DTT in 100 mM Tris buffer, pH 8.0, containing 1 mM EDTA and 6 M GuHCl. The reduced proteins were freed
from excess reductant by size exclusion using PD-10 columns equilibrated with 50 mM potassium phosphate buffer, pH 7.0, containing 1 mM EDTA and 2 M GuHCl. The reduced RNase proteins were then incubated with 2-fold excess of As-Mal over total thiols for 2 h at room temperature. The As-Mal labeled RNase protein was freed from excess As-Mal using a PD-10 column run as before and stored at 4°C.

As-Mal RNase peptides and As-Mal Trx protein conjugates were characterized by mass spectroscopy and their ability to coordinate dithiols was verified using DTT titrations monitored at 300 nm. The stoichiometry of the conjugation reaction was verified by thiol titer using DTNB [11] before and after As-Mal labeling.

4.2.7 Inhibition of Thioredoxin Reductase by the As-Mal Trx Conjugate

Thioredoxin reductase was assayed following the Trx-mediated reduction of DTNB. In an assay volume of 120 µL, 1 µM wild type Trx was mixed with 200 µM DTNB, followed by 200 µM NADPH and where appropriate, 1 µM of the arsenicals were added as inhibitors. The assay was started by the addition of 25 nM of the reductase, and the accumulation of 5-thio-2-nitrobenzoate (TNB) was monitored at 412 nm.

4.2.8 Inhibition of Protein Disulfide Isomerase by As-Mal RNase Peptides

The inhibition of PDI by arsenicals was assessed using the standard insulin reductase turbidometric assay [18]. In a total reaction volume of 120 µL, 50 µM Insulin was mixed with a reducing agent (100 µM TCEP or 5 mM GSH). The reaction
was started by the addition of 1 μM PDI. Inhibition of PDI delays the onset of turbidity [18] which was followed at 600 nm at 25°C. The lag times were plotted to reflect the times taken to reach 25% of the maximal rates observed. Potential inhibitors were added to the assay mixture where appropriate immediately before the addition of the isomerase. All arsenicals were added to give a total arsenic concentration in the cuvette of 10 μM.

4.2.9 \(K_d\) determination of 1,8 bis-As-Mal-RNase with Reduced PDI

The binding of 1, 8 bis-As-Mal RNase was assessed by following the increase in absorbance at 300 nm. Reduced PDI (1 μM in 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA and 100 μM TCEP) was titrated with 0.2 μM aliquots of the 1,8 bis-arsenical and spectra were recorded 20 min after each addition. Data were analyzed as previously to yield stoichiometries and dissociation constants [19]. The CxxC motifs of the a and a' thioredoxin domains of PDI have identical sequences and the data fit a single class of binding sites.

4.3 Results and Discussion

4.3.1 Binding of As-Mal to Thiols

GSH was used as a model peptide to test the binding of As-Mal with thiols. As-Mal shows two modes of reaction with thiol group: one via Michael addition at the
maleimide, and the other through coordination at the As (III) center. A titration of As-Mal with GSH showed a decline in absorbance at 320 nm (with a sharp endpoint at 1.0 GSH/As-Mal) as depicted in Figure 4.3. This reflects a rapid conjugation of thiol with maleimide without undue interference from reversible binding to the arsenical moiety. The reaction between 100 µM As-Mal and equimolar GSH at pH 7.5, 25°C was complete in 40 s, half-complete in 2.1 s, and was comparable to the behavior of N-phenylmaleimide which serves as a positive control (Figure 4.4). These experiments suggest that As-Mal should react rapidly with protein thiols providing that they are solvent-accessible.

![Figure 4.3](image)

**Figure 4.3**  **Titration of As-Mal with GSH.** As-Mal (100 µM) was titrated with 20 µM aliquots of GSH and the spectrum was recorded immediately after each addition. A clear end point was observed at 100 µM of GSH. The subsequent increase in absorbance at GSH concentration > 100 µM is due to reaction between the thiol group of GSH and the arsenical moiety of As-Mal. Reprinted from Journal of the American Chemical Society, 2013; 135(7):2415-8, with permission from the publisher.
Figure 4.4  Reaction of \( N \)-phenylmaleimide and As-Mal with GSH. Final concentrations of 100 \( \mu M \) GSH and either 100 \( \mu M \) \( N \)-phenylmaleimide (A) or 100 \( \mu M \) As-Mal (B) was mixed in stopped-flow spectrophotometer. The absorbance decline at 320 nm (black open circles) was fit to second-order kinetics (red solid line) yielding 8680 ± 80 M\(^{-1}\)s\(^{-1}\) and 4450 ± 415 M\(^{-1}\)s\(^{-1}\) respectively. Reprinted from Journal of
4.3.2 Inhibition of Thioredoxin Reductase

Thioredoxin reductase (TrxR) belongs to the flavoprotein family of pyridine nucleotide oxidoreductases [20] [21]. It is the only enzyme known to directly reduce thioredoxin intracellularly and is essential for maintaining the redox balance of the cell. TrxR is a homodimer, with each monomeric unit consisting of an FAD binding domain and an NADPH binding domain that also contains the redox-active disulfide. The NADPH binding domain also contains a binding pocket that helps in the recognition of Trx [20]. During catalysis, reducing equivalents are passed from NADPH to FAD moiety, followed by reduction of CxxC motif. A conformational change then leads to formation of a mixed disulfide with a thioredoxin docked against the reductase, and to the eventual release of reduced thioredoxin [22]. Reduced Trx in turn passes the electrons to its various protein substrates. Among these is ribonucleotide reductase, an enzyme that converts ribonucleotide into deoxyribonucleotides [23] and hence is essential for DNA synthesis and cell proliferation. Trx also reduces peroxiredoxins and methionone sulfoxide reductases that help in protecting the cells from ROS like H₂O₂ [24] [25]. Trx also regulates transcriptional factors like NF-Kβ [26], AP-1, AP-2, Ref-1[27] and estrogen receptors [28]. Due to their involvement in important pathways that are essential for cell viability and proliferation, TrxR and Trx are overexpressed in various cancers
including pancreatic, lung, breast and prostate [29–31]. TrxR has been viewed as an important target for cancer therapy, and many metal based compounds have been used for inhibiting the mammalian form [32] [33].

To deliver arsenic to TrxR, a single As-Mal group was introduced at the surface of *E. coli* Trx. Here Trx (an oxidoreductase containing a redox-active pair of cysteine residues, C32 and C35) was mutated to leave the surface-accessible C32 available for conjugation. Treatment of this C35S mutant with one equivalent of As-Mal rapidly generated a monolabeled derivative (at 100 µM concentrations the reaction was half-complete in <5 s, Figure 4.5). Monoalkylation was confirmed by mass spectrum (Figure 4.6), by the loss of the single DTNB reactive thiol group, and by titration of the conjugated arsenical reagent with dithiothreitol (Figure 4.7). Maleimides are prone to reverse Michael reactions and can undergo exchange with proteins containing reactive cysteine. To evaluate the stability As-Mal-Trx conjugate, it was incubated with 2-fold excess DTT for 2 h at room temperature. No change in mass spectrum (Figure 4.8) was observed indicating that exchange of maleimide between thioredoxin and DTT was not significant under these conditions.

Simple arsenicals along with As-Mal-Trx were tried as inhibitors of TrxR in an assay that follows the reduction of DTNB to the TNB anion, mediated by wild type Trx. The initial phase involving the reduction of Trx is shown in Figure 4.9; panel A. PSAO, MMA and ATO produced only modest inhibition of the TrxR assay whereas As-Mal-Trx was considerably more potent with a rapid and progressive loss of activity.
(Figure 4.9; panel B). This showed that in this case that the substrate protein labeled with an arsenic moiety was better inhibitors than just simple arsenical alone.

**Figure 4.5** Kinetics of As-Mal conjugation to C35S thioredoxin. As-Mal (100 µM) was added to 100 µM *E. coli* C35S Trx in 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA. The decline in absorbance at 320 nm reflects conjugation to the maleimide. Reprinted from Journal of the American Chemical Society, 2013; 135(7):2415-8, with permission from the publisher.
Figure 4.6  Mass spectroscopy of As-Mal Trx
Mass of the unconjugated C35S mutant Trx: 14783
Expected mass for As-Mal-Trx: 15046 (as arsenoxide)
Mass found: 15069 (15046 + Na)
Reprinted from Journal of the American Chemical Society, 2013; 135(7):2415-8, with permission from the publisher.
Figure 4.7  **Quantification of As-Mal Trx.** The As-Mal content of conjugated Trx was established via DTT titration in phosphate buffer pH 7.5, containing 1 mM EDTA (endpoint = 25 µM). Spectra were recorded immediately after each increment of DTT and the absorbance at 300 nm is plotted as a function of increasing concentration of DTT. Reprinted from Journal of the American Chemical Society, 2013; 135(7):2415-8, with permission from the publisher.
Figure 4.8  Mass spectrum of As-Mal-Trx after DTT treatment.
Expected mass of DTT-conjugated As-Mal-Trx: 15182,
Mass found 15183, plus mono, di- and trihydrates (inset).
No change in mass spectrum was found on incubation of 75 µM As-
Mal-Trx with 150 µM DTT over 2 h at room temperature. Reprinted
from Journal of the American Chemical Society, 2013; 135(7):2415-8,
with permission from the publisher.
**Figure 4.9** Inhibition of thioredoxin reductase by arsenicals. Panel A shows the reduction of Trx by NADPH catalyzed by thioredoxin reductase, together with the arsenicals tested at 1 µM. Reoxidation of reduced Trx by DTNB is followed in the assay shown in the panel B. The appearance of TNB anion was followed at 412 nm. Reprinted from Journal of the American Chemical Society, 2013; 135(7):2415-8, with permission from the publisher.
4.3.3 Inhibition of Protein Disulfide Isomerase

Next As-Mal labeled unfolded RNase constructs were tried as potential inhibitors of reduced PDI. As discussed in Chapter 2, human PDI is comprised of four thioredoxin domains with a and a' domains carrying CxxC motifs that are responsible for the varied oxidoreductase and isomerase activities of the enzyme. Typically, the substrates of PDI family members are proteins that retain disordered or conformationally mobile regions. A widely employed substrate of PDI is reduced and denatured, pancreatic ribonuclease A (RNase), and hence this protein was used as a vehicle for arsenic conjugation. Reduction of the four native disulfides of RNase leads to an unfolded protein that can be labeled with eight As-Mal moieties. This protein was prepared as described in methods section and its mass spectroscopy confirmed the presence of 8 arsenoxide functionalities on the polypeptide chain (Figure 4.10). Additionally a limited series of site-directed mutants were prepared which included RNase constructs with cysteines only at position 1,4, at 4,8, at 1,8 or at 1,4 and 8 (Figure 4.11). These proteins were subsequently labeled with As-Mal (As-Mal RNase) and characterized by mass spectroscopy (Figure 4.12) and quantified by titrations with DTT to confirm the presence of arsenic content (Figure 4.13).
Figure 4.10  Mass spectrum of (As-Mal)$_8$ RNase.
Mass of unconjugated reduced RNase: 13690
Expected mass for (As-Mal)$_8$ RNase: 15794 (as arsenoxides)
Mass found: 15796
Reprinted from Journal of the American Chemical Society, 2013; 135(7):2415-8, with permission from the publisher.
Figure 4.11  As-Mal conjugated RNase (As-Mal RNase). Cartoon depiction of As-Mal labeled cysteine residues along the 124 residue native chain of reduced RNase (at positions 26, 40, 58, 65, 72, 84, 95 and 110; labeled, for clarity, as cysteines 1-8).
**Figure 4.12** Mass spectra of As-Mal RNase derivatives. Mass spectra of mono (A), bis (B) and tris (C) conjugated As-Mal RNase. Reprinted from Journal of the American Chemical Society, 2013; 135(7):2415-8, with permission from the publisher.
Figure 4.13  Quantification of As-Mal RNase derivatives. The As-Mal content of conjugated RNase was established via DTT titration in phosphate buffer, pH 7.5, containing 1 mM EDTA. Spectra were recorded immediately after each addition. The increase in absorbance at 300 nm was plotted as a function of increasing concentration of DTT. Reprinted from Journal of the American Chemical Society, 2013; 135(7):2415-8, with permission from the publisher.
The widely-used insulin reductase assay was chosen to assess the effectiveness of As-Mal labeled RNase inhibitors [17]. PDI catalyzes the reduction of insulin disulfides driven by the water-soluble phosphine, TCEP, and the precipitation of the isolated B chain is followed by light scattering. In the absence of the As-Mal RNase derivatives the onset of turbidity occurs at ~ 230 s whereas it is strongly suppressed with the 1, 8 bis-As-Mal RNase and to a lesser extent with the 1, 4 and 4, 8 bis-As-Mal-RNase. The monoarsenical C4-As-Mal RNase was considerably less effective than 1, 8 bis-As Mal RNase. The 1, 4, 8-tris-As-Mal RNase was roughly comparable to the 1, 8 derivative (Figure 4.14). To confirm that inhibition reflected the conjugation of arsenicals to RNase the effect of the parent unconjugated cysteine RNase mutant proteins was evaluated and found no significant inhibition of insulin reduction (Figure 4.15). Further, the simple monoarsenical PSAO failed to significantly delay the onset of turbidity at a concentration of 10 µM. Thus the placement and spacing of arsenicals along a protein chain can be exploited to modulate inhibitory potency.

In addition, the presence of 10 µM GSH in the assay did not significantly attenuate the inhibition observed with the 1, 8 bis-As-Mal RNase. Since the levels of GSH in the extracellular matrix reach about 10 µM [34] [35], this inhibitor might be effective at targeting PDI present on the cell surface. In contrast intracellular concentrations of GSH are much higher and in the 0.5 – 10 mM range [36]. The inclusion of 5 mM GSH in the reaction assay led to insignificant inhibition of PDI by
all these arsenicals because at this concentration it is an effective competitor for the arsenic functionality (Figure 4.16).
Figure 4.14  **Insulin reductase assay with selected arsenicals.** Panel A shows the comparison between mono-, bis- and tris-As-Mal-RNase derivatives and the monoarsenical PSAO in the insulin reductase assay driven by TCEP. The same colors used are also employed in the bar diagram shown in panel B. The concentration of inhibitors was chosen to reflect a total of 10 μM arsenic in the assay. The lag times are calculated as described in methods. Reprinted from Journal of the American Chemical Society, 2013; 135(7):2415-8, with permission from the publisher.
**Figure 4.15** Insulin reductase assay with cysteine RNase peptides. Insulin (50 µM in 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA) was mixed with 100 µM TCEP. The reaction was started by the addition of 1 µM PDI. Unconjugated cysteine-RNase peptides were tested at a concentration 10 µM each. The arsenic free RNase peptides were unable to achieve any inhibition of the isomerase. The lag times are calculated as described in methods. Reprinted from Journal of the American Chemical Society, 2013; 135(7):2415-8, with permission from the publisher.
Figure 4.16 Insulin reductase assay in presence of 5 mM GSH. Insulin (50 µM in 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA) was mixed with 5 mM GSH and 5 µM 1, 8 bis-As-Mal RNase. The reaction was started by the addition of 1 µM PDI. The presence of 5 mM GSH as the ultimate reductant in the reaction mixture (blue) significantly attenuates the inhibition in comparison to when 100 µM TCEP is used as the reducing agent (red). Now the delay is comparable to the control assay (black) that does not contain any arsenical.

Since the 1, 8 bis-As-Mal RNase appeared to be the most effective inhibitor, a spectrophotometric titration of this bis-arsenical with reduced PDI was performed. The titration followed the increase in absorbance at 300 nm accompanying thiol coordination of the CxxC motifs of PDI to the hydrated As (III) species (Figure 4.17; panel A) [13][19][37]. The data in Figure 4.17; panel B were fit to a $K_d$ of 22 ± 7 nM with a stoichiometry of 1.21 ± 0.16 bis-arsenical RNase per PDI. These data are
consistent with the complexation of both the \( \text{a} \) and \( \text{a}' \) CxxC motifs of reduced PDI with the 1, 8 bis-As-Mal RNase. For comparison, a direct assessment of the binding of the monoarsenical PSAO to reduced PDI showed much weaker binding with a \( K_d \) of 1.1 \( \mu \text{M} \) [19], consistent with the failure to impact the insulin reductase assay significantly.
**Figure 4.17**  
*Kₐ determination of reduced PDI with 1, 8 bis-As-Mal RNase.* (A) A schematic representation of the interaction of 1, 8 bis-As-Mal RNase with reduced PDI. (B) Titration of 1 µM reduced PDI was performed with increasing concentration of 1, 8 bis-As-Mal RNase. The data were fit according to [19] yielding a *Kₐ* of 22 ± 7 nM. The redox-active cysteines are represented in spheres and their sulfur atoms are colored in yellow. Panel B reprinted from Journal of the American Chemical Society, 2013; 135(7):2415-8, with permission from the publisher.
4.4 Conclusion

In summary, this work introduces a simple way to conjugate arsenicals to cysteine-bearing peptides and proteins and demonstrates that the resulting adducts are more potent inhibitors than simple monoarsenical derivatives. Arsenicals might be delivered intracellularly via cell penetrating peptides [38] or thiol-bearing polymers [39] for targeting species that bind arsenical avidly in competition with GSH. While maleimide-thiol conjugates can be subject to reverse Michael reactions over extended times in physiological media, this is not necessarily a disadvantage because it provides the potential for targeted, controlled, release of cytotoxic species [40]. Finally, the facility of the conjugation chemistry described here could allow multiple peptidic and protein scaffolds to be screened prior to manipulation of linker stability [6] [41].
REFERENCES


Chapter 5

ARSENIC AFFINITY CHROMATOGRAPHY

5.1 Introduction

The ability of arsenic based molecules to bind proteins containing thiols in close proximity led to its use as a biochemical tool for affinity purification for enrichment of such proteins. However the earliest use arsenic affinity chromatography was in 1975 for the purification of alkaline phosphates from calf intestine by Sepharose 4B resin labeled with phenylarsonic acid [1]. This resin utilized As(V) as the affinity moiety since inorganic arsenate had been reported as a competitive inhibitor of the calf intestine enzyme [2]. Hannestad et al. prepared the first As(III) labeled resin in 1982 by reacting PAO with CNBr-activated Sepharose 6B. The resin had a binding capacity of about 3-4 µmol/ml and it was used to study the behavior of small molecular weight monothiols and dithiols towards arsenic. The column demonstrated a higher affinity for 1, 2 and 1, 3 dithiols like 2, 3-dimercaptopropanol and dihydrolipoic acid as compared to monothiols like cysteine and mercaptoacetate and 1, 4 dithiols like DTT [3]. In 1991 Zhou et al. prepared another As(III) affinity resin by covalently linking PAO to carboxymethyl (CM)-Bio-Gel A through an activated N-hydroxysuccinimide (NHS) ester [4]. This resin was used for the purification of lecithin cholesterol acyltransferase (LCAT), a plasma enzyme containing two catalytic active cysteine
residues (C 31, and C184). Plasma was first passed through various ion-exchange and hydrophobic columns and finally through arsenic affinity column to isolate the pure protein, resulting in a 13-fold increase in activity and an overall yield of 11%. Elutions were done with 2, 3-dimercaptopropanesulfonic acid (DMPS). In 1992 Hoffman et al. prepared an arsenic column by labeling PAO to the commercially available NHS activated agarose [5]. This column was used for separation of proteins containing vicinal thiols from 3T3-L1 adipocytes, which led to the identification of GLUT4 (insulin dependent glucose transporter) as an arsenic interacting protein. Elutions were performed by β-ME in 1% Triton or Sodium dodecyl sulfate (SDS) solution and the eluted proteins were detected on polyacrylamide gels by staining with 4-[¹²⁵I]iodophenylarsine oxide. This resin was stored in the presence of β-ME to ensure no oligomerization took place between the arsenic species themselves. In 1993 Gitler and co-workers introduced 6-aminohexanoic acid as spacer group between the Sepharose 4B resin and PAO via carbodiimide coupling [6]. The binding capacity of this resin was 14 µmol/ml and it was used for the purification of proteins containing vicinal dithiols from L1210 murine leukemia lymphoblast. The bound proteins were eluted with β-ME or DTT and identified by the thiol-specific labeling reagent N-iodoacetyl-3-[¹²⁵I]-iodotyrosine. In 1995 Winski and Carter used arsenic-affinity chromatography to identify hemoglobin as the main As(III) binding cytosolic protein in rat red blood cells [7]. In 1999 Menzel et al. also employed this technique to identify tubulin and actin as As(III) binding proteins from human lymphoblastoid cells.
PAO-Sepharose 4B resin has also been used to purify ubiquitin-protein ligase (E3) [9] and microsomal retinol dehydrogenase [10]. Both these proteins contain reactive cysteine residues and have been shown to be inhibited by phenylarsine oxide. In 1997 Halestrap et al. used PAO conjugated to Affigel-10 resin to identify adenine nucleotide translocase (ANT) as the target of phenylarsine oxide [11]. ANT as discussed in section 1.7.2 of Chapter 1 is the most abundant protein of mitochondrial intermembrane space and a key component of the mitochondrial permeability transition pore. It contains one cysteine residue in each of its three loops (Cys 57, Cys 160, Cys 257) that face the mitochondrial intermembrane space [12]. These cysteine residues are targets of PAO and arsenite [13–16]. In 1997 Gitler et al. introduced a general method to identify and enrich vicinal thiol containing proteins (VTPs) from cells by the use of arsenic chromatography [17]. In this study murine leukemia cells L1210 were incubated with cell permeable N-ethylmaleimide to alkylate any free thiols, followed by the reduction of disulfide bonds by DTT. Finally the enrichment of VTPs was carried out using a PAO-resin. However the use of the dithiol DTT leads to competition between the reductant and proteins for the arsenic resin. Foley et al. [18] improved this strategy by using TCEP, which is a non-competitive phosphine reducing agent in identifying VTPs from rat brain extract. In another study Eupergit C beads containing the oxirane functionality were reacted with PAO to make another arsenic-affinity resin. This resin was used to identify arsenic-binding proteins from A549 human lung carcinoma cells [19]. In 2003 Chang et al. [20] used a PAO-agarose affinity column to identified galectin1, glutathione S-transferase P-form and
thioredoxin peroxidase as As(III) binding protein from Chinese hamster ovary cells. Sepharose resins labeled with bis-arsenicals like the FlAsH compound and its derivatives (discussed in Chapter 1, section 1.10) have also been prepared and are used for purification of proteins genetically fused with a CCxxCC motif [21] [22].

5.2 A New Arsenic Sepharose Affinity Resin

As discussed in Chapter 4, As-Mal can be used to introduce arsenic rapidly and stoichiometrically on proteins and peptides as well as surfaces that contain free thiol groups [23]. This chapter discusses two new arsenical affinity resins that were made by reacting As-Mal with thiol-activated Sepharose beads. These resins were subsequently used in proof of principle experiments in a yeast proteomic study and for the purification of folded and unfolded proteins containing cysteine residues.

5.3 Materials and Methods

5.3.1 General Methods and Materials

As-Mal was synthesized as described in section 4.2.2.1 of Chapter 4. UV-Visible spectra were recorded on an Agilent 8453 diode-array instrument. Guanidine hydrochloride (GuHCl), oxidized glutathione (GSSG), N-ethylmaleimide (NEM), bovine pancreas ribonuclease A (RNase A) and thiopropyl Sepharose 6B were purchased from Sigma-Aldrich. Dithiothreitol (DTT), 5, 5′-dithiobis (2-nitrobenzoate) (DTNB), and tris (2-carboxyethyl) phosphine (TCEP) were purchased from Gold
Biotechnology Inc. Thiopropyl Sepharose 4B and desalting PD-10 columns were purchased from GE Healthcare and tris(3-hydroxypropyl)phosphine (THP) was purchased from Calbiochem. Stock solutions of DTT, TCEP, and THP were standardized before use using DTNB [24].

5.3.2 Preparation of As-Mal-Sepharose Resin

Dry thiopropyl Sepharose 4B and 6B beads (1 gm) were swollen and washed with distilled water to remove any additives that were present. The resins were converted to their free thiol form by reducing them with 1% DTT solution in 300 mM sodium bicarbonate buffer, pH 8.5, containing 1 mM EDTA for 2 h at room temperature (Figure 5.1). The resins were washed extensively with 100 mM sodium acetate buffer, pH 4.0, containing 500 mM sodium chloride and 1 mM EDTA to remove the free protecting group (2-pyridone) and excess DTT. The thiol activated beads were then equilibrated with 100 mM Tris buffer, pH 7.5, containing 500 mM sodium chloride and 1 mM EDTA followed by incubation with a freshly-prepared solutions of As-Mal (10 mM and 40 mM for Sepharose 4B and 6B respectively) in same buffer for 2 h at room temperature. The As-Mal labeled Sepharose resins were washed with 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA to remove excess compound and stored in the same buffer at 4°C (Figure 5.1).
5.3.3 Binding Capacity of As-Mal-Sepharose Resins

The binding capacity of the resins was determined by the concentration of DTT that bound per ml of resin. The DTT concentration was verified by thiol titer using DTNB before and after the addition to arsenic resin. Typical values obtained were: 4-5 µmol/ml of resin for As-Mal-Sepharose 4B and 30-35 µmol/ml of resin for As-Mal-Sepharose 6B.

5.3.4 Mass Spectroscopic Analysis of Peptides

The mass spectroscopy was performed by Dr Shawn Gannon. The peptides were separated using a Waters Acquity UPLC system equipped with a BEH C18 100 x 1 mm, 1.7 micron particle size column utilizing a linear gradient composed of 0.1% formic acid in water (solvent A) and 0.1% formic acid in methanol (solvent B). The separation method began with a 3 minute isocratic hold at the initial conditions of 5% solvent B. Following the three minute hold the gradient varied from 5% solvent B to 65% solvent B over 67 minutes (total time 70 minutes). The peptides were detected using a Thermo Scientific Orbitrap Velos mass spectrometer. The data were acquired using the data dependent analysis mode, which allowed for the collection of both full scan and ms-ms data of the peptides to obtain amino acid sequence information. The extracted ion chromatogram was generated based on the mass of the alkylated peptides and the resulting peaks were integrated. The reduced fraction was calculated by dividing the peak area of the reduced mass by the sum of the reduced and oxidized
peak areas. The ms-ms data were further analyzed using the SEQUEST peptide search algorithm of Thermo Scientific Proteome Discoverer (Version 1.3.0.339).

5.3.5 **Regeneration of Arsenical resins**

As-Mal-Sepharose resin saturated with DTT was incubated with 100 mM NEM solution in water for 2 h at room temperature. The regenerated resin was washed extensively with water and re-equilibrated with 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA.

5.3.6 **Expression, Purification and Handling of Proteins**

Human thioredoxin reductase was a generous gift from Dr Sharon Rozovsky of the University of Delaware. *Escherichia coli* thioredoxin (Trx) and the septuple RNase mutant, retaining a single cysteine (the 4\textsuperscript{th} cysteine, C4) was purified as in Codding et al. [25]. Additional RNase mutant proteins (cysteines at position 1,4, at 1,8 ,and at 3,4) were obtained using QuikChange mutagenesis (Stratagene) as described earlier [23].

5.3.7 **Purification of Disulfide-Containing Peptides from Yeast Extract**

*Saccharomyces cerevisiae* (S288C) was grown for 30 h at 25°C in YPP media [26] and the harvested cells were stored at -80°C prior to use. A portion of the pelleted cells (200 mg) was added to a 1.5 ml polypropylene Eppendorf tube containing 6 mg of dry oxidized glutathione followed by 0.5 ml of 50 mM potassium phosphate buffer, pH 7.5, containing 0.5 mM EDTA and 65 mM NaCl. The mixture was resuspended
and then 100 µl of a 4:1 v/v mixture of ethanol: toluene was added with vortexing for 2 min to further permeabilize the cells [27]. The cell suspension was incubated for 5 min at room temperature and then 2.5 mg of N-ethylmaleimide was added with mixing and left for a further 5 min. Under these native like conditions many redox-active dithiols would be oxidized to their disulfide forms and surface accessible thiols would then be rapidly alkylated with NEM. Proteins were then denatured by the addition of 600 mg of solid guanidine hydrochloride followed by vortexing intermittently for 10 min. The mixture was then dialyzed in 1 cm wide 14-16 kDa cutoff tubing against two 600 ml changes of distilled water followed by 600 ml of 25 mM potassium phosphate buffer, pH 7.5, containing 0.25 mM EDTA and 33 mM NaCl. The copious flocculent precipitate in the dialysis tube was resuspended by inversion and one-half was used for the subsequent digestion using 1 mg of TPCK-treated trypsin. Following a 12 h incubation at 37°C, an additional 0.5 mg of trypsin was added for further 6 h incubation. The suspension was then centrifuged for 2 min in an Eppendorf bench centrifuge and the clear supernatant (~1 ml) was mixed end-over-end for 1.5 h with 0.2 ml of As-Mal-Sepharose 6B resin in the presence of 7 mM THP. The resin was then washed in the centrifuge tube with three 1 ml aliquots of 50 mM potassium phosphate buffer, pH 7.5, containing 1 mM EDTA. Bound peptides were then eluted by exposing the packed resin to 1 ml of 3 mM DTT with rocking for 1 h followed by recovery of the peptides by centrifugation and alkylation of cysteine residues and excess DTT using a final concentration of 10 mM NEM. Preparation of peptides was
performed by Dr Colin Thorpe. The alkylated peptide mixture was then analyzed by Dr Shawn Gannon as described in section 5.3.4.

5.3.8 Analysis of As-Mal-Sepharose 4B with *Escherichia coli* Thioredoxin

*Escherichia coli* Trx (200 µM in 0.5 ml of 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA) was mixed end-over-end with 0.5 ml As-Mal-Sepharose 4B in the presence of 5 mM TCEP for 2 h at 4°C in a small 1.5 ml column. The packed resin was washed 4 times with 1 ml of 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA. The elutions were performed with 0.5 ml DTT solutions prepared in same buffer.

5.3.9 Purification of Human Thioredoxin Reductase

Impure human thioredoxin reductase protein fraction (100 µM in 0.5 ml of 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA) was incubated with 0.5 ml As-Mal-Sepharose 4B in the presence of 5 mM TCEP for 2 h at 4°C. The resin was washed 3 times with 1 ml of 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA followed by 0.5 ml elutions with DTT solutions prepared in the same buffer.

5.3.10 Reduction of RNase and its Interaction with As-Mal-Sepharose 4B

To remove any adventitious disulfides, all RNase mutant proteins and wild type RNase were pre-incubated with 80-fold molar excess DTT in 100 mM Tris buffer
containing 1 mM EDTA and 6 M GuHCl, pH 8.0 for 2 h at 37°C. The reduced and
denatured proteins were separated from the excess reductant and GuHCl by using a
PD-10 column equilibrated with 10 mM sodium acetate buffer, pH 4.0, containing 1
mM EDTA. Reduced proteins were stored anaerobically under nitrogen at 4°C until
further use.

Reduced wild type RNase (70 µM; 560 µM thiols) and 100 µM each of
dicysteine RNase mutants (corresponding to 200 µM thiols) with cysteines at position
1,4, at 3,4 and 1,8 were mixed end-over-end with 0.5 ml of As-Mal-Sepharose 4B in
the presence of 5 mM TCEP for 2 h at room temperature in a small 1.5 ml column.
The resin was washed 5 times with 1 ml of 50 mM phosphate buffer, pH 7.5,
containing 1 mM EDTA followed by 0.5 ml elutions with DTT solutions prepared in
same buffer.
Figure 5.1 Preparation of As-Mal-Sepharose resins. The preparation of As-Mal labeled resins consisted of two steps; the resins were first reduced with DTT to generate their thiol activated forms. The reduced resins were then incubated with As-Mal to generate As-Mal-Sepharose 6B/4B resins in the second step (see materials and methods).
5.4 Results and Discussion

As mentioned earlier the maleimide derivative As-Mal is readily conjugated to thiopropyl-Sepharose resins. In this work, we conjugated two supports with different exclusion limits. Sepharose 6B has a relatively high capacity of 30-35 µmol/ml resin and was used for the purification of small peptides. The lower capacity Sepharose 4B (4-5 µmol/ml resin) was used for experiments involving proteins.

5.4.1 Yeast proteomic Study with As-Mal-Sepharose 6B

The As-Mal-Sepharose 6B resin was used for a proof of principle for proteomic study to examine if it could be used to identify peptides from proteins with redox-active disulfide bonds. As described in methods section, a permeabilized suspension of Saccharomyces cerevisiae was treated with GSSG in an attempt to oxidize a portion of reduced redox-active dithiols motifs. Any structural proteins containing disulfide would already be oxidized and in principle might contribute to the peptides identified in this study. Proteins were then denatured and exposed to NEM so that as many free thiols, or thiols readily released from metal sequestration, would be alkylated and thus excluded from binding to the resin. The sample was then digested with trypsin and the resulting digest reduced in the presence of arsenical resin. Following displacement of bound peptides by DTT, the solutions were then treated with excess NEM and analyzed by MS-MS (see methods). Table 5.1 represents the list of peptides that were eluted from the resin and the corresponding proteins that were identified from the yeast genome.
Table 5.1  **Protein eluted from As-Mal-Sepharose 6B resin.** The eluted peptides were analyzed by mass spectroscopy and the corresponding proteins were identified from the yeast genome.

<table>
<thead>
<tr>
<th>Peptides eluted from As-Mal-Sepharose 6B resin</th>
<th>Corresponding protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>QNEGHEcQcQcGScK</td>
<td>Cu-Metallothionein</td>
</tr>
<tr>
<td>DTcILFNGQDSEK</td>
<td>Cox17p, Copper-binding metallothionein</td>
</tr>
<tr>
<td>cVESVNDLSLSSQEEQcLSNcVNR</td>
<td>Tim8p, Mitochondrial intermembrane space protein</td>
</tr>
<tr>
<td>DScHCGSTcLPScSGGEK</td>
<td>Crs5p, Copper-binding metallothionein</td>
</tr>
<tr>
<td>VVGANPAAIK</td>
<td>Trx2p, Cytoplasmic thioredoxin isoenzyme</td>
</tr>
<tr>
<td>DNAEGQcGESLADQAcR</td>
<td>Hsp12p, Plasma membrane localized protein</td>
</tr>
<tr>
<td>AETAAQDVQQK</td>
<td>YOL109W, Peripheral membrane protein of the plasma membrane</td>
</tr>
<tr>
<td>TYNALcPLDWieK</td>
<td>Cox12p, Subunit V1b of cytochrome c oxidase</td>
</tr>
<tr>
<td>VIDQLYcTccSK</td>
<td>Ymr194c-bp, Putative protein of unknown function</td>
</tr>
<tr>
<td>cFTDcVNDFFTSK</td>
<td>Tim9p, Mitochondrial intermembrane space protein</td>
</tr>
<tr>
<td>TGNAGPRPAcGVIGLTN</td>
<td>CuZn superoxide dismutase</td>
</tr>
<tr>
<td>NDAcIDQcLAK</td>
<td>Tim13p, Mitochondrial intermembrane space protein</td>
</tr>
<tr>
<td>SSSTcDSLNQVcTcYcEHENSAVK</td>
<td>Ccw14p, Covalently linked cell wall glycoprotein</td>
</tr>
<tr>
<td>IVSNAScTTcLAPLAK</td>
<td>glyceraldehyde-3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>DESDIDEIYPPRRVSSIIDK</td>
<td>orf2</td>
</tr>
<tr>
<td>NPVILADAccSR</td>
<td>Pdc1p, pyruvate decarboxylase isozymes</td>
</tr>
<tr>
<td>AcYVcGK</td>
<td>Gis2p, Putative zinc finger protein</td>
</tr>
<tr>
<td>ILETAREYILGSLIELER</td>
<td>non-clathrin coat protein</td>
</tr>
<tr>
<td>VVGLSSLPcYEK</td>
<td>alcohol dehydrogenase</td>
</tr>
<tr>
<td>SGDAALVK</td>
<td>Tef1p, Translational elongation factor EF-1 alpha</td>
</tr>
<tr>
<td>ALGGTcVNVGcVPK</td>
<td>glutathione reductase</td>
</tr>
<tr>
<td>TEDNEYQEISASALKK</td>
<td>Ded81p, Cytosolic asparaginyl-tRNA synthetase</td>
</tr>
<tr>
<td>IYcSGLQcGR</td>
<td>Map1p, Methionine aminopeptidase</td>
</tr>
<tr>
<td>ARSSTVTAGTPSSSSSIQKYKPSNVPR</td>
<td>Kic1p, Protein kinase of the PAK/Ste20 kinase family</td>
</tr>
<tr>
<td>GELcESTFSDEPTLVNVPScK</td>
<td>NHA1</td>
</tr>
<tr>
<td>IEEELGDK</td>
<td>Eno2p, Enolase II</td>
</tr>
<tr>
<td>EENFcNLVNRcLK</td>
<td>NIP80, subunit of the dynactin complex</td>
</tr>
<tr>
<td>YGlcGPNGcGK</td>
<td>elongation factor 3</td>
</tr>
</tbody>
</table>
5.4.2 Binding of Folded and Unfolded Proteins with As-Mal-Sepharose 4B

As mentioned earlier Sepharose 4B resin has a lower binding capacity and is more suitable for binding proteins. To test the As-Mal-Sepharose 4B resin, *Escherichia coli* thioredoxin with a single redox active centre at C32-C35 was incubated with the resin in the presence of TCEP. The reduced protein completely bound to the resin and no leakage was observed either in the flow through or in the washes. The bound protein could be eluted at 0.5 mM DTT concentration (Figure 5.2).

The performance of the resin was then tested with human thioredoxin reductase, an oxidoreductase containing a CxxC motif and a surface-exposed CU (U=selenocysteine) motif at the C-terminus [28]. Both these redox centers are required for this enzyme's activity [29] and have been shown to bind to arsenic [30]. The impure protein was incubated with As-Mal-Sepharose 4B in the presence of TCEP for 2 h at 4°C (section 5.3.9). A truncated form of the protein which lacks the selenocysteine residue at the C-terminus was not able to bind to the resin and was removed along with other impurities in the flow through and washes while the full length protein was able to bind to the resin and was eluted with DTT as shown in Figure 5.3.
Figure 5.2  Thioredoxin binding and elution from As-Mal-Sepharose 4B resin. Purified *E. coli* Trx bound with As-Mal-Sepharose 4B resin in presence of TCEP as described in section 5.3.8. The protein bound completely to the resin with no material evident in flow through or in 4 additional washes. The protein elutes at 0.5 mM DTT and higher concentration.
Figure 5.3  **Purification of human thioredoxin reductase by As-Mal-Sepharose 4B.** The impure protein was incubated with the resin in presence of TCEP as described in section 5.3.9. A truncated form of the protein that lacked selenocysteine along with other protein impurities was removed in flow-through and washes. The full length bound protein was eluted with DTT.
Next to evaluate the difference in the binding efficiency of proteins containing cysteine residues at varied position and spacing, reduced unfolded pancreatic ribonuclease A (RNase) with 8 cysteine residues was used a model protein. A series of site-directed mutants were prepared where only cysteines at position 1, 4, at 3, 4 and at 1, 8 positions were retained in the sequence. These dicysteine RNase peptides along with wild type RNase were tested with As-Mal-Sepharose 4B. Figure 5.4 shows a qualitative comparison of the binding affinities of these proteins. The binding affinity of 3, 4-RNase which had 15 amino acid residues in between the two cysteine residues was evidently higher than 1, 4-RNase which has a spacing of 44 amino acid residues. In contrast the 1, 8-RNase mutant with 84 amino acid residues separating the first and the eight cysteine residue was unable to bind to the resin; about 90% of the protein added to the column was recovered in flow through and washes (data not shown). As expected from these trends, the wild type RNase with 8 cys residues bound very tightly to the resin and required higher DTT concentration for elution than the mutant RNase constructs.
**Figure 5.4** Binding of RNase proteins with As-Mal-Sepharose 4B. The spacing between cysteine residues led to difference in binding affinities of RNase mutants. The 1, 4-RNase mutant was eluted at an initial 1 mM DTT concentration, while the 3, 4-RNase mutant required 5 mM DTT. The wild type RNase containing 8 cysteine residues required at least 20 mM DTT for elution. All the RNase proteins were reduced and denatured prior to loading to the arsenic affinity column.
5.5 Regeneration of Arsenical Resins

Despite the specificity of arsenic towards proteins containing vicinal thiols, arsenic affinity chromatography is not widely used. One likely disincentive is the costs associated with these resins. While monothiols such as β-ME can be used to elute proteins that bind arsenicals relatively weakly, many more tightly bound species require DTT as an eluant. This results in the formation of a stable cyclic complex between DTT and As(III). The suggestion from the manufacturers of arsenical resins states that elution with DTT would permanently inactivate the resin hence making these relatively costly resins a onetime use and an expensive chromatography technique. However we were able to successfully regenerate arsenic-labeled resins by

Figure 5.5 Regeneration of arsenic labeled resins. Schematic representation for regeneration of DTT bound arsenical resin.
use of the thiol-maleimide chemistry as depicted in Figure 5.9. To decouple the DTT-As(III) complex, the beads were incubated with excess NEM (100 mM) for 2 h at room temperature. The arsenic-thiol bond is reversible in nature therefore NEM can capture any DTT that transitionally releases one thiol ligand thereby leading to the eventual dissociation of monoalkylated DTT with subsequent generation of the bismaleimide-DTT adduct. Following incubation, excess NEM and DTT-NEM adducts were readily removed by washing the resin with water. This method provides an easy, effective and cost efficient way to regenerate arsenical resins. The regenerated resin showed 100% of the binding capacity as compared with the freshly prepared As-Mal-Sepharose resins as shown in Figure 5.6.
Figure 5.6  **Binding of reduced Trx with regenerated As-Mal-Sepharose 4B.**
The regenerated resin (0.5 ml) was incubated with 200 µM purified *E. coli* Trx in presence of 5 mM TCEP (see section 5.3.8). The protein bound completely and no leakage of protein was observed in flow-through or washes. The elution profile was comparable to that observed with fresh resin.
5.6 Conclusion

In summary, this work provides a simple way to prepare arsenic labeled resin for purification of proteins and peptides containing free thiols. As-Mal-Sepharose 6B resin was shown to bind cysteine-containing peptides in a trial yeast proteomic experiment. The corresponding Sepharose 4B resin showed promise in the purification of enzymes with dithiols redox motifs. Finally we show that the use of DTT does not inactivate the resin for subsequent uses suggesting that arsenic affinity chromatography might deserve more general attention.
REFERENCES


Appendix A

NMR DATA FOR MULTIVALENT ARSENICALS

1H NMR for Bis-arsenical-1

The peak at 4.04, 1.24, and 2.17 are due to ethyl acetate present as solvent impurity.
$^{13}$C NMR for Bis-arsenical-1
The peaks at 3.2, 2.05, and 1.07 are due to ethyl acetate present as solvent impurity.
13C NMR for Bis-arsenical-2
The peak at 2.06 is due to acetone present as a solvent impurity.
$^{13}$C NMR for Tris-arsesenical-1
Appendix B

NMR DATA FOR ARSENICAL-MALEIMIDE

1H NMR spectra for As-Mal
$^{13}$C NMR spectra for As-Mal
Appendix C

COPYRIGHTS

Chapter 1

Figure 1.3 is adapted from:


Figure 1.5 is adapted from:


Chapter 2

Figure 2.3 is reprinted from:


Chapter 4

Figures and text are reprinted with permission from:

Additionally Figure 4.15 is adapted from:

An Arsenic–Maleimide for the Generation of New Targeted Biochemical Reagents

Author: Aparna Sapra and Colin Thorpe
Publication: Journal of the American Chemical Society
Publisher: American Chemical Society
Date: Feb 1, 2013
Copyright © 2013, American Chemical Society

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from [COMPLETE REFERENCE CITATION]. Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

If credit is given to another source for the material you requested, permission must be obtained from that source.