

https://dspace.org/

Qualitative and quantitative characterization of arsenic-binding peptides by LC-ESI-MS for the development of an arsenic point of care device for the South African mines.

Moepya, Refilwe

2022
https://hdl.handle.net/10539/37027

Downloaded from a DSpace repository. https://dspace.org/



WITS
UNIVERSITY

Qualitative and quantitative characterization of arsenic-binding peptides by LC-ESI-MS for the development of an arsenic point of care device for the South African mines.

By

Refilwe Moepya

A dissertation submitted to the faculty of Science, University of the Witwatersrand, Johannesburg in fulfilment of the requirements for the degree of Master of Science.

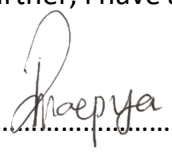
Supervised by Dr Maya Makatini

2022

Disclaimer

I hereby declare that this dissertation is my own original work and has not been submitted before to any institution for assessment purposes.

Further, I have acknowledged all sources used and I have cited these in the reference section.



……. 06-07-2022

Refilwe J. Moepya Date

Acknowledgements

I have been inspired by a number of people throughout my MSc journey. I would first and foremost like to thank God for granting me the strength and wisdom to conduct this research and write the dissertation. I also want to thank God for my supervisor, **Dr Maya Makatini**. Thank you for patiently mentoring me through this project and affording me the platform and space to grow as a scientist.

I would like to thank my lovely parents **Matthews Moepya** and **Lizzy Moepya** for supporting me and always encouraging me to do better and be better. **Tshegofatso Moepya**, **Lerato Moepya**, and **Kagiso Moepya**, words cannot express how grateful I am to have you as my support system.

I would like to send my heartfelt gratitude to everyone in the peptide synthesis group (**Khanani Machumela**, **Ntombizanele Ngqinayo**, **Priscilla Nyembe**, and **Ntlama Lesotho**) your assistance and advice is much appreciated. A special thanks goes to **Dr Thabo Peme** for proofreading my work.

My sincere gratitude goes out to the Wits Mass Spec team (**Dr Eric Morifi** and **Thapelo Mbhele**), who have always gone above and beyond to help me.

I would like to thank **Mahlohonolo Sekoto**, **Lebogang Letselebe**, **Bongani Malinga**, **Mudzuli Maphupha**, **Lerato Mokoloko**, **Lindelo Mguni** and **Moeketsi Kgotle** for their emotional and mental support during this journey.

Lastly, I would like to acknowledge the National Research Foundation (**NRF**) for its financial support.

LIST OF TABLES

Table 2.1: Common amino acids in nature. 29
Table 2.2: Summary of synthesized ArsR-derived peptides. 50
Table 3.1: Detailed summary of the complexes formed during arsenic binding to ArsR-derived peptides. 72
Table 4.1: Concentrations of pep1-RJM with individual As(III) species used in the titration reactions. 87
Table 4.2: Observed masses of Pep1-As complexes. 92
Table 4.3: Binding constant reaction of Pep1-RJM with iAs(III) obtained by means of peak areas using LC-ESI-MS. 94
Table 4.4: Binding constant reaction of Pep1-RJM with DMA(III) obtained by means of peak areas using LC-ESI-MS. 95
Table 4.5: Binding constant reaction of Pep1-RJM with PAO(III) obtained by means of peak areas using LC-ESI-MS. 96
Table 6.1: Absorbance values used to determine the loading capacity of the 2-CTC resin. 109
Table 6.2: LCMS results for ArsR derived peptides. 110
Table 6.3: Optimum As(III) binding conditions. 114
SD-1: Table 7.2:Absorbance values used for determining loading capacity of 2-CTC. 116
SD-19: Table 7.3: Effect of pH on arsenic(III) binding. 134
SD-20: Table 7.4: Effect of reaction temperature on binding. 134
SD-21: Table 7.5: Effect of reaction time on binding. 135
SD-22: Table 7.6: Effect of initial As(III) on binding. 135
SD-23: Table 7.7: Effect of peptide dose on As(III) binding. 136

LIST OF SCHEMES

Scheme 1.1: Avidin-biotin interaction. † Denotes that Avidin is also often conjugated to an immobilized support.[65] 21
Scheme 2.1: General structure of ligand-induced conformational change of a protein.[5] 31
Scheme 2.2: Basic principle of SPPS.[11] 36
Scheme 2.3: Mechanism for anchoring 1st AA. 44
Scheme 2.4: Fmoc deprotection mechanism using 20% piperidine where R represents the amino acid sequence. 45
Scheme 2.5: Mechanism for the activation of the Fmoc AA by HBTU, producing an OBt leaving group. 46
Scheme 2.6: Mechanism for the activation of the Fmoc AA by HATU, producing an OAt leaving group. 47
Scheme 3.1: Reduction of As(V) to As(III) using thiol-containing agents. [6] 62
Scheme 3.2: Dithiothreitol induced disulfide bond reduction in Ars-derived peptides. 65
Scheme 3.3: Mechanistic representation of a disulfide exchange reaction using DTT.[9] 65
Scheme 3.4: Prereduction of As(V) to As(III) where reagent A=L-cystine, EDT, or DTT. 67
Scheme 4.1: Pep1-RJM biotinylation. 90

LIST OF FIGURES

Figure 1.1: Anthropogenic and natural sources of arsenic.[9] 2
Figure 1.2: Orbital diagram of 75As illustrating the number of electrons inside the orbitals of each energy level where a) represents (-3) Oxidation state, b) represents the elemental arsenic with oxidation state (0), c) represents oxidation state (+3) and lastly d) represents (+5) oxidation state.[14]. 4
Figure 1.3: Structural similarities between arsenate and phosphate anions.[18] 5
Figure 1.4: Pathway of arsenic biotransformation in humans.[23,24] 6
Figure 1.5: Structure of common chelating agents used as for the treatment of arsenic poisoning. 8
Figure 1.6: Schematic diagram with basic components of HPLC.[37] 10
Figure 1.7: Interaction mechanism of peptides with reverse phase HPLC stationary phase.[39] 11
Figure 1.8: The schematic diagram of the fundamental components of a mass spectrometer.[41] 12
Figure 1.9: Schematic diagram of Electrospray ionization process.[43] 13
Figure 1.10: Schematic diagram of a Q-TOF mass analyser.[48] 15
Figure 1.11: Structure of a biosensor: a) Biological signal, b) Physicochemical signal, c) Electrical signal.[54] 16
Figure 1.12: Schematic representation of conventional three-electrochemical cell where WE, RE, and CE stands for working electrode, reference electrode, and counter electrode.[56] 17
Figure 2.1: A conformational change as a result of iAsIII binding to the E. coli ArsR repressor (P15905).[9] 32
Figure 2.2: Protein alignment of two ArsR proteins ARSR2_ECOLX and ARSR_ECOLI obtained from the Uniprot protein database. ARSR2_ECOLX is the identifier name for ArsR pR773 and ARSR_ECOLI is the identifier name for a strain of pR773 ArsR (strain K12). 33
Figure 2.3:Peptide sequence of the ArsR-derived peptides with amino acids derived from the metal-binding domain. 33
Figure 2.4: The concept of Orthogonality using Fmoc/t-Bu chemistry.[14] 37
Figure 2.5: Common resins used for Fmoc-based SPPS. Resins (1-5) are used for the synthesis of peptide acids.[11] 39
Figure 2.6: Structures of commercially available coupling reagents for SPPS. Where A. Phosphonium salts, B. Uronium/ Aminium salts, C. Immonium salts, and D. Carbidiimides. 41
Figure 2.7:Basic equipment for SPPS. 42
Figure 2.8: Mass spectra of Fmoc protected Pep1-RJM coupling up to 9th AA. $[M+H]^+$=1277.5006 and $[M+2H]^{2+}$=639.2508 corresponds to GELCVCDLC-Fmoc sequence. 49
Figure 2.9: Mass spectra of Fmoc protected Pep2-RJM coupling up to 9th AA. $[M+H]^+$= 1063.4166 and $[M+2H]^{2+}$=532.2163 corresponds to CVCSGSSKA-Fmoc sequence. 49
Figure 2.10: Mass spectra of Fmoc protected Pep3-RJM coupling up to 9th AA. [M]=1133.4592 corresponds to the CVCSGDSKN-Fmoc sequence. 49
Figure 2.11: Mass spectra of Pep1-RJM with the complete sequence GELCVCDLCTAL-NH_2 51
Figure 2.12: Mass spectra of Pep2-RJM with the complete sequence CVCSGSSKAVCI-NH_2. 51
Figure 2.13: Mass spectra of Pep3-RJM with complete sequence CVCSGDSKNICS-NH_2. 51
Figure 2.14: Chromatogram showing the UV absorbance of Pep1-RJM at R_t=6.40 min. 52
Figure 2.15: HPLC chromatogram of purified Pep2-RJM. Peak 1 at 5.86 is an impurity from the synthesis whereas peak 2 at 6.31 minutes is Pep2-RJM. 53
Figure 2.16: HPLC chromatogram of purified Pep2-RJM. Peak 1-3 appearing at 5.23 min, 5.68 min and 6.43 min correspond to [M+2H]=578.7129. 54

Figure 2.17: HPLC chromatogram of purified Pep3-RJM. Peak 1-2 at 5.66 min and 6.00 min correspond to Pep3-RJM, whereas peak 3 was an impurity from synthesis. 55
Figure 3.1: The covalent binding of cysteine-containing peptides with urinary arsenic metabolites where A represents the novel ArsR-derived peptides and B represents the covalently bonded Peptide-arsenic complex. 59
Figure 3.3.2: Structure of arsenic metabolites found in urine.[2] 60
Figure 3.3: The distribution of organic and inorganic arsenic as a function of pH.[3] 61
Figure 3.4: conformational folding as a result of disulfide bond formation. Diagram A represents a peptide with reduced thiol groups on the cysteine side chains. In structure B, the peptide folds to bring the two cysteine residues closer for the formation of a disulfide bond. 63
Figure 3.5: Structure of commonly used reducing reagents. BME and DTT are thiol-based reducing reagents and TCEP is a phosphine-cased reducing reagent.[8] 64
Figure 3.6: 6-vessel reactor where A. rubber seal, B. PTFE cap, C. round bottom flask, and D. Stirring hot plate. 68
Figure 3.7: A typical example of iAs binding using Pep1-RJM. iAs is coordinated to the three thiol groups of the cysteine residues on the peptide. 72
Figure 3.8: Mass spectrum of Pep1-iAs complex. 73
Figure 3.9: Interaction between iAs(III), DMA(III), and PAO(III) with ArsR-derived peptides containing three sulfur atoms from three different cysteine residues. 74
Figure 3.10: Effect of pH on the binding of arsenic(III) by Pep1-RJM. 76
Figure 3.11: Effect of temperature on the binding of arsenic(III) by Pep1-RJM. 77
Figure 3.12: Effect of reaction time on the binding of arsenic(III) by Pep1-RJM. 78
Figure 3.13: Effect of initial As(III) concentration on the binding of arsenic(III) by Pep1-RJM. 79
Figure 3.14: Effect of peptide concentration on the binding of arsenic(III) by Pep1-RJM. 80
Figure 4.1: structural example of Pep1-RJM coordination to iAs(III). Structure A shows a typical 1:1 coordination whereas structure B shows a cluster coordination. 91
Figure 4.2: Structure of Pep1-As complexes with covalent bonds contributing to the binding affinity of the complex. 93
Figure 4.3: Structure of biotinylated Pep1-RJM. 97
Figure 4.4: Mass spectrum of biotinylated Pep1-RJM. 98
Figure 6.1: Ultra-High-Performance Liquid Chromatography (Thermo Scientific Ultimate 3000, RS diode array detectors) with a Diode Array (190, 195, 215, 254, and 300 nm) coupled to a Bruker Compact Q-TOF high-resolution mass spectrometer. 106
Figure 6.6.2: Carousel 6 plus reaction station 107
Figure 6.6.3: Peak instruments pH meter used in As(III) binding studies 107
SD-2: Figure 7.1: Mass spectra of Pep1-RJM with sequence GELCVCDLCTAL-NH2. 117
SD-3:Figure 7.2: Mass spectra of Pep2-RJM with sequence CVCSGSSKAVCI-NH2. 118
SD-4: Figure 7.3: HPLC-MS chromatogram of Pep2-RJM with impurity. 119
SD-5: Figure 7.4: HPLC-MS chromatogram of purified Pep2-RJM. 120
SD-6: Figure 7.5: Mass spectra of Pep3-RJM with sequence CVCSGDSKNICS-NH_2. 121
SD-7: Figure 7.6: Chromatograms of Pep1_RJM after purification. A) Base peak chromatogram, b) UV spectra 215 nm, c) UV spectra 254 nm, d) UV spectra 300 nm and e) UV spectra 450 nm. 122
SD-8: Figure 7.7: Chromatograms of crude Pep2_RJM before purification. A) Base peak chromatogram, b) UV spectra 190 nm, c) UV spectra 195 nm, d) UV spectra 200 nm and e) UV spectra 200 nm 123
SD-9: Figure 7.8: Chromatograms of crude Pep3_RJM before purification. a) Base peak chromatogram, b) UV spectra 190 nm, c) UV spectra 195 nm, d) UV spectra 200 nm and e) UV spectra 200 nm 124

SD-10: Figure 7.9: Mass spectrum of iAs-Pep1 complex 125
SD-11:Figure 7.10: Mass spectrum of DMA-Pep1 complex 126
SD-12: Figure 7.11: Mass spectrum of PAO-Pep1 complex. 127
SD-13: Figure 7.12: Mass spectrum of iAs-Pep2 complex. 128
DS-14: Figure 7.13: Mass spectrum of DMA-Pep2 complex. 129
SD-15: Figure 7.14: Mass spectrum of PAO-Pep2 complex. 130
SD-16: Figure 7.15: Mass spectrum of iAs-Pep3 complex. 131
SD-17: Figure 7.16: Mass spectrum for DMA-Pep2 complex. 132
SD-18: Figure 7.17: Mass spectrum of PAO-Pep3 complex. 133
SD-24: Figure 7.18: LC-MS of biotinylated Pep1-RJM. 137

LIST OF ABBREVIATIONS

AA	amino acids
ACN	acetonitrile
Au	gold
Boc	tert-Butyl-oxycarbonyl
Cys (C)	cysteine
2-CTC	2-chlorotrityl chloride
DCM	dichloromethane
DIPEA	N,N-diisopropylethylamine
DMF	N,N-dimethylformamide
DMSO	dimethyl sulfoxide
EDT	1,2-Ethanedithiol
ESI	electrospray Ionisation
Fmoc	9-Fluorenylmethoxycarbonyl
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexaflourophosphate
HBTU	(2-(1H-benzotrial-1-yl-1,1,3,3-tetramethyluronium hexafluorophosphate
HPLC	high performance liquid chromatography
hrs	hours
LC-MS	liquid chromatography-mass spectrometry
iAs	inorganic arsenic
DMA	dimethylarsinous acid
MeOH	methanol

1. Contents

Disclaimer II

Acknowledgements III

LIST OF TABLES IV

LIST OF SCHEMES V

LIST OF FIGURES VI

LIST OF ABBREVIATIONS IX

Abstract XIII

Chapter 1 1

1.1 Introduction 1

1.2 Arsenic exposure 2

1.3 The chemistry of arsenic 3

1.4 Kinetics and metabolism 5

1.5 Health effects 7

1.6 Clinical treatment 7

1.7 Biomarkers of exposure 8

1.8 Analytical methods and characterization 9

1.8.1 Detection of inorganic and organic arsenic 9

1.8.2 Detection of peptides and peptide-arsenic complexes 9

1.9 Biosensor 16

1.10 Electrochemical cell 17

1.11 Peptide-based bioreceptors 19

1.12 Modes of peptide immobilization 20

1.12.1 Biotin-(Strept) avidin interaction 21

1.13 Background and aims of dissertation 22

1.13.1 Problem statement 22

1.13.2 Aim 22

1.13.3 Objectives 22

1.13.4 Novelty of study 23

Chapter 2 : Design and synthesis of novel arsenic binding peptides 29

2.1 Introduction 29

2.2 Relationship between structure and function of a bioreceptor 30

2.3 Relationship between structure and function of arsenical resistance operon repressor (ArsR) 31

2.4 Sequence of ArsR-derived peptides 33

2.5 Basic principles of solid phase peptide synthesis (SPPS) 34

2.5.1 Peptide synthesis 34
2.5.2 Basic principle of SPPS 35
2.5.3 Boc/Bzl and Fmoc/tBu strategies 36
2.5.4 Solid support and linkers in SPPS 38
2.5.5 Amino acid activation and coupling reagents 39
2.5.6 Peptide cleavage 41
2.6 Methods 42
2.6.1 Synthesis of Pep1-RJM, Pep2-RJM, and Pep3-RJM 42
2.6.2 Swelling and activation of 2-CTC resin 43
2.6.3 Loading the first amino acid on a 2-CTC resin 43
2.6.4 Fmoc deprotection 44
2.6.5 Peptide ligation 45
2.6.6 Peptide cleavage 47
2.6.7 Peptide purification and characterization 48
2.7 Results and discussion 48
2.7.1 Peptide synthesis 48
2.7.2 UV absorbance of ArsR-derived peptides 52
2.7.3 Purification and characterization of Pep1-RJM 52
2.7.4 Purification and characterization of Pep2-RJM 53
2.7.5 Purification and characterization of Pep3-RJM 54
2.8 Concluding remarks 55
Chapter 3 : Qualitative determination of arsenic metabolites bound to novel Ars-R derived peptides in urine matrix by HPLC-ESI-MS 58
3.1 Introduction 58
3.2 Arsenic speciation 59
3.3 Reduction of As(V) to As(III) 62
3.4 Peptide disulfide bond reduction 62
3.5 Materials and methods 64
3.5.1 Reagents 64
3.5.2 Sample preparation 64
3.5.3 Reduction of peptides 65
3.5.4 Reduction As(V) to As(III) 65
3.5.5 Binding reaction setup 68
3.5.6 HPLC-ESI-MS 69
3.6 Results and discussion 70
3.6.1 Arsenic(V) reduction to As(III) 70

3.6.2 Characterization 71
3.6.3 Effect of pH 74
3.6.4 Effect of temperature 76
3.6.5 Effect of reaction time 77
3.6.6 Effect of initial As(III) concentration 78
3.6.7 Effect of peptide dosage 79
3.6.8 Competitive binding studies 80
3.6 Concluding remarks 81
Chapter 4 : Quantitative characterization of arsenic(III)-binding to Pep1-RJM using Liquid-chromatography hyphenated to electrospray ionization quadrupole time-of-flight mass spectrometry. 84
4.1 Introduction 84
4.2 Binding constants from LC-ESI-MS 85
4.3 Materials and methods 86
4.3.1 Sample preparation 86
4.3.2 LC-ESI-MS 87
4.3.3 Binding calculations[5] 88
4.3.3 Pep1-RJM biotinylation 90
4.4 Results and discussion 90
4.4.1 Binding constant calculations 91
4.4.2 Pep1-RJM biotinylation 97
4.5 Concluding remarks 98
Chapter 5 : Conclusion and recommendations 101
5.1 Summary 101
5.2 Conclusion 102
5.3 Future scope of research 103
5.4 Research contribution 104
Chapter 6 : Experimental 105
6.1 Reagents 105
6.6 Coupling of the 2nd-9th amino acid 109
APPENDICES 115

Abstract

Long-term occupational exposure to inorganic arsenic (iAs), through inhalation, skin contact, or ingestion, can induce a wide range of negative health effects, including cancer, neurological disorders, and cardiovascular diseases. Colorimetry, inductively coupled plasma mass spectrometry (ICP-MS), and atomic absorption spectrometry (AAS) are some of the currently used techniques for the detection of arsenic. However, these techniques are expensive, technically challenging and it is difficult to make real-time decisions due to the time that lapses between sample collection, transportation, analysis, and data processing. As a result, affordable and fast arsenic-monitoring systems are highly required. The research presented in this dissertation explores the development of peptides that are derived from arsenic resistance repressor protein (ArsR) as the molecular recognition element of an arsenic monitoring point of care device. This study aims to qualitatively and quantitatively investigate the binding affinity of arsenic to the ArsR derived peptides using the LC-ESI-MS.

The ArsR-derived peptides (Pep1-RJM, Pep2-RJM, Pep3-RJM) were synthesized using the solid-phase peptide synthesis strategy based on the Fmoc approach. The synthesized peptides were purified using semi-preparative high-performance liquid chromatography (prep-HPLC) and characterized using liquid chromatography-tandem mass spectrometry (LCMS). The qualitative and quantitative binding activity of the ArsR-derived peptides to the arsenate (iAs(V), arsenite (iAs(III)), dimethylarsenic acid (DMA(V)), and Phenylarsine oxide (PAO(III)) standards was monitored by measuring the difference between the concentration of original unbound forms of the peptide and the concentration of peptide-arsenic complex using a combination of reversed-phase liquid chromatography and electrospray ionization mass spectrometry. The pentavalent arsenic metabolites were successfully reduced to the better binding trivalent state using L-cysteine while the peptides disulphides bond were reduced using a fivefold molar excess of dithiothreitol (DTT). The effect of various parameters such as pH, temperature, initial concentration of As(III) metabolites, reaction time, and peptide concentration on the binding efficiency was investigated.

Optimum As(III) binding was observed at pH8, at a temperature of 40 °C, an initial As(III) concentration of 60 µM, a binding reaction time of 15 minutes, and a peptide dosage of 25 µM. The condensation reaction of Pep1-RJM with arsenic showed a fixed 1:1 coordination without the formation of clusters even at higher molar concentrations of the As(III). The equilibrium constant for the interaction of Pep1-RJM with iAs(III), DMA(III), and PAO(III) that leads to the formation of covalent arsenic-sulphur bonds (As-S) was successfully measured using the peak areas of the extracted ion chromatograms of the LC-ESI-MS. The following order of binding affinities was obtained from the LC-ESI-MS: pep1-PAO (4.11×10^8 $M^{-1)}$ > pep1-iAs (3.47×10^8 M^{-1})> pep1-DMA (1.68×10^8 M^{-1}). Findings suggest that thiol-containing peptides derived from the ArsR protein are capable of effectively binding arsenic in urine and the LC-ESI-MS is an efficient tool in the quantification of arsenic binding.

Chapter 1

1.1 Introduction

South Africa is well-known for its abundant mineral resources.[1] Although mining contributes significantly to the country's economic prosperity, it also generates massive amounts of dust, which contains toxic and dangerous heavy metals such as arsenic that pollute the environment. Mineworkers are constantly exposed to hazardous dusts which results in the accumulation of metal toxins in their body's soft tissues, leading to a variety of illnesses in the long run. The type of minerals mined, as well as the pollutants present, their quantities and duration of exposure has an influence on the prevalence and severity of these illnesses.[2,3] As a result, occupational health and safety institutions and physicians must monitor mineral pollution levels in underground mines, as well as the uptake of these poisons by miners who work in them. The carcinogenic potential of arsenic makes it a major public health concern around the world. Arsenic is ubiquitous in the atmosphere, and it is impossible to avoid human exposure to it.[2,4]

A comprehensive overview of the various modes of arsenic exposure and the difference between levels of exposure caused by occupational activity and those caused by environmental exposure is presented in this chapter. This chapter also provides a general introduction to biomarkers and a description of the most suitable biological matrix for the detection of arsenic exposure. The conventional characterization techniques used for arsenic detection are also discussed, including their limitations, and how peptide-based electrochemical biosensors can be used to overcome these limitations. Furthermore, the basic principles of reverse-phase high-performance liquid chromatography and electrospray ionization are explored and their application to peptide-analyte binding studies.

1.2 Arsenic exposure

Arsenic is a naturally occurring element in the earth's crust.[5] As shown in Figure 1.1, it is released into the environment by a variety of natural sources and anthropogenic activities. Arsenicosis (arsenic poisoning) is a chronic illness caused by prolonged exposure to arsenic levels that are above the permissible exposure limits. It is a major public health issue that affects millions of people worldwide due to environmental and occupational exposure. Ingestion, inhalation, and dermal absorption are the most common routes of arsenic exposure. Although arsenic is found naturally in water, air, soil, sediments, and soil biota, the general public can be exposed to elevated levels of inorganic arsenic by drinking contaminated water, eating food crops irrigated with contaminated water and eating contaminated food. The majority of occupational exposure occurs through the inhalation of contaminated air.[6–8]

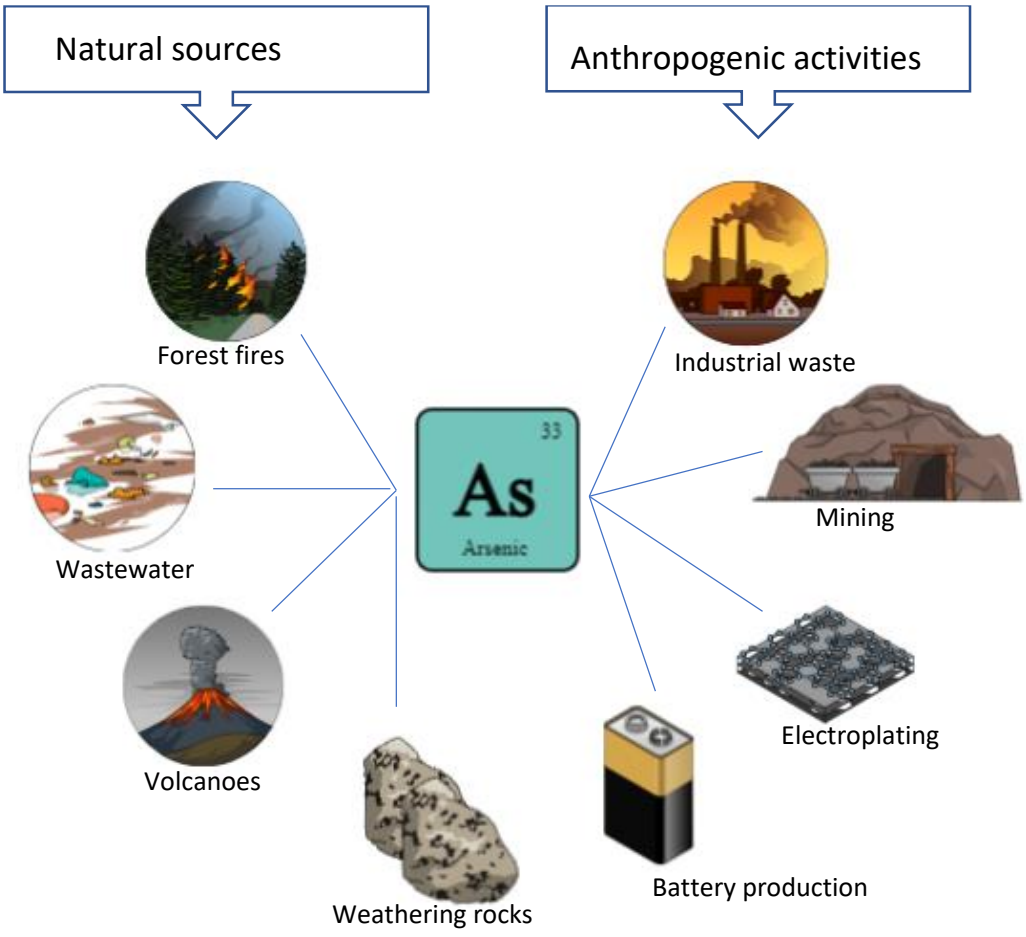


Figure 1.1: Anthropogenic and natural sources of arsenic.[9]

Inorganic arsenic exposure can occur in a variety of industries such as nonferrous smelting companies, glass producing companies, the electronics industry, sectors that use arsenic as a wood preservative, and industries that use and manufacture arsenic-containing insecticides and herbicides. The concentration of arsenic in urine is used to determine the level of exposure of workers. The findings have implications for workplace health and safety policies. As a result, precise arsenic detection techniques are required. When assessing arsenic levels, it is important to keep in mind that different chemical forms of arsenic have different degrees of toxicity. This is best demonstrated by organoarsenic compounds, which are commonly found in various seafoods and can result in high levels of arsenic in urine with little or no risk of arsenic poisoning.[10,11]

1.3 The chemistry of arsenic

Arsenic is a Group 15 element with the symbol "As" on the periodic table, and it is a member of the pnictogens group. It has an atomic weight of 74,921 and is classified as a semi-metal because it exhibits properties of both metals and non-metals. ^{75}As is the only naturally occurring non-radioactive isotope of arsenic. All pnictogen elements have an electronic basic state configuration of ns^2np^3.[12–14]

Figure 1.2 below depicts the orbital diagram of arsenic, which includes the s, p, and d subshells. The s subshell has only one orbital, while the p and d subshells have three and five orbitals, respectively. Each orbital can hold up to two electrons. Arsenic contains five valence electrons, which promotes it to participate in chemical bonding through its unoccupied p-orbital. According to the orbital diagram, the first energy level comprises of two electrons, the second level holds eight electrons, the third level contains eighteen electrons, and the fourth level has five electrons. Although the s-orbital of the 4th energy level is completely full, the p-orbital is only half-filled.[13]

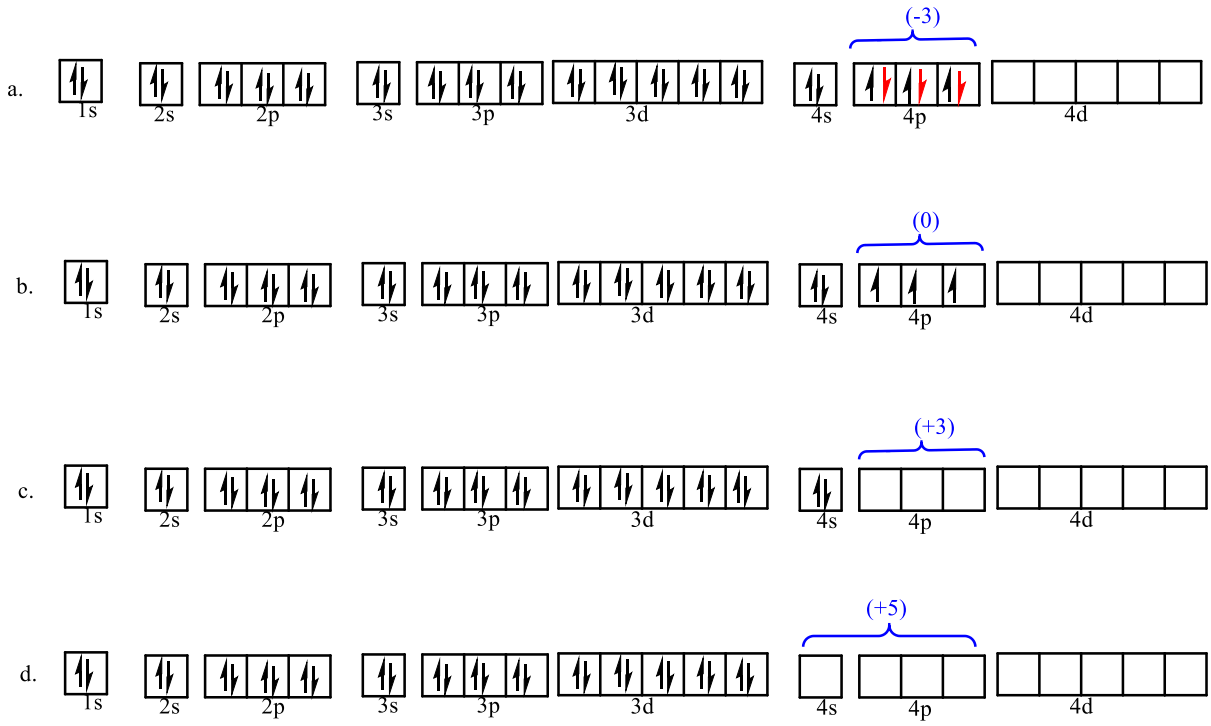


Figure 1.2: Orbital diagram of 75As illustrating the number of electrons inside the orbitals of each energy level where a) represents (-3) Oxidation state, b) represents the elemental arsenic with oxidation state (0), c) represents oxidation state (+3) and lastly d) represents (+5) oxidation state.[14]

Understanding the structure and bonding of a molecule provides insight into its chemistry. A covalent bond is formed when an arsenic atom shares its valence electrons through a bond with another atom. The most commonly observed oxidation states of arsenic are (-3, 0, +3, and +5). The -3-oxidation state is seen when three electrons are added to fill up the 4p subshell resulting in a total of six electrons in the subshell. Arsenic with the valence state of 0 is formed when an arsenic atom shares its three valence electrons in the 4p subshell equally with three other arsenic atoms. This interaction results in a trigonal-pyramidal structure. The +3 valence state is observed when the three valence electrons in the 4p orbitals are attracted to non-metals such as sulfur or oxygen. The +5 oxidation state is obtained when the valence electrons interact with a non-metal using both 4s and 4p electrons.[15]

1.4 Kinetics and metabolism

The metabolic processes of arsenic in the body play a significant role in the determination of arsenic toxicity. The metabolic pathway includes the important steps of arsenic reduction from a pentavalent to a trivalent state and the oxidative methylation to a pentavalent form as shown in Figure 1.4. Trivalent arsenicals are more poisonous than pentavalent arsenicals, even when methylated. Inorganic arsenicals become less toxic as their oxidation state increases, so arsenite (As (+1)) is more toxic than arsenite (As (+3)) and arsenate (As (+5)). Phosphate and arsenate molecules are structurally and electrochemically similar (Figure 1.3). As a result of this similarity, arsenate can be incorporated into metabolic pathways that require the phosphate ion. A notable example of the difficulties associated with this substitution can be found in glycolysis, where the exchange of phosphate by arsenate results in products that are easily hydrolysed to form the next intermediate, which inhibits the production of ATP that is usually formed in this reaction.[16,17]

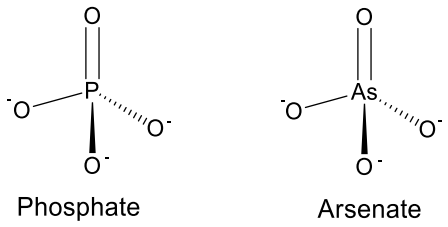


Figure 1.3: Structural similarities between arsenate and phosphate anions.[18]

Arsenate toxicity is caused by partial protein synthesis inhibition and the interference of arsenate with protein phosphorylation, which is especially problematic when phosphate levels are low. In contrast, arsenite is poisonous because it behaves like a metal, establishing strong metal-thiol interactions with cysteine residues of proteins, preventing enzyme activity.[19,20]

Arsenic is metabolized and eliminated from the body primarily through methylation, which occurs in the liver. Figure 1.4 below illustrates the production of monomethylarsonic acid (MMA (V)) from arsenic by the arsenite methyltransferase enzyme, which is then methylated again to produce dimethylarsinic acid DMA (V). The reduction of iAs (V) to iAs (III) is catalysed by an arsenic reductase enzyme.[21,22]

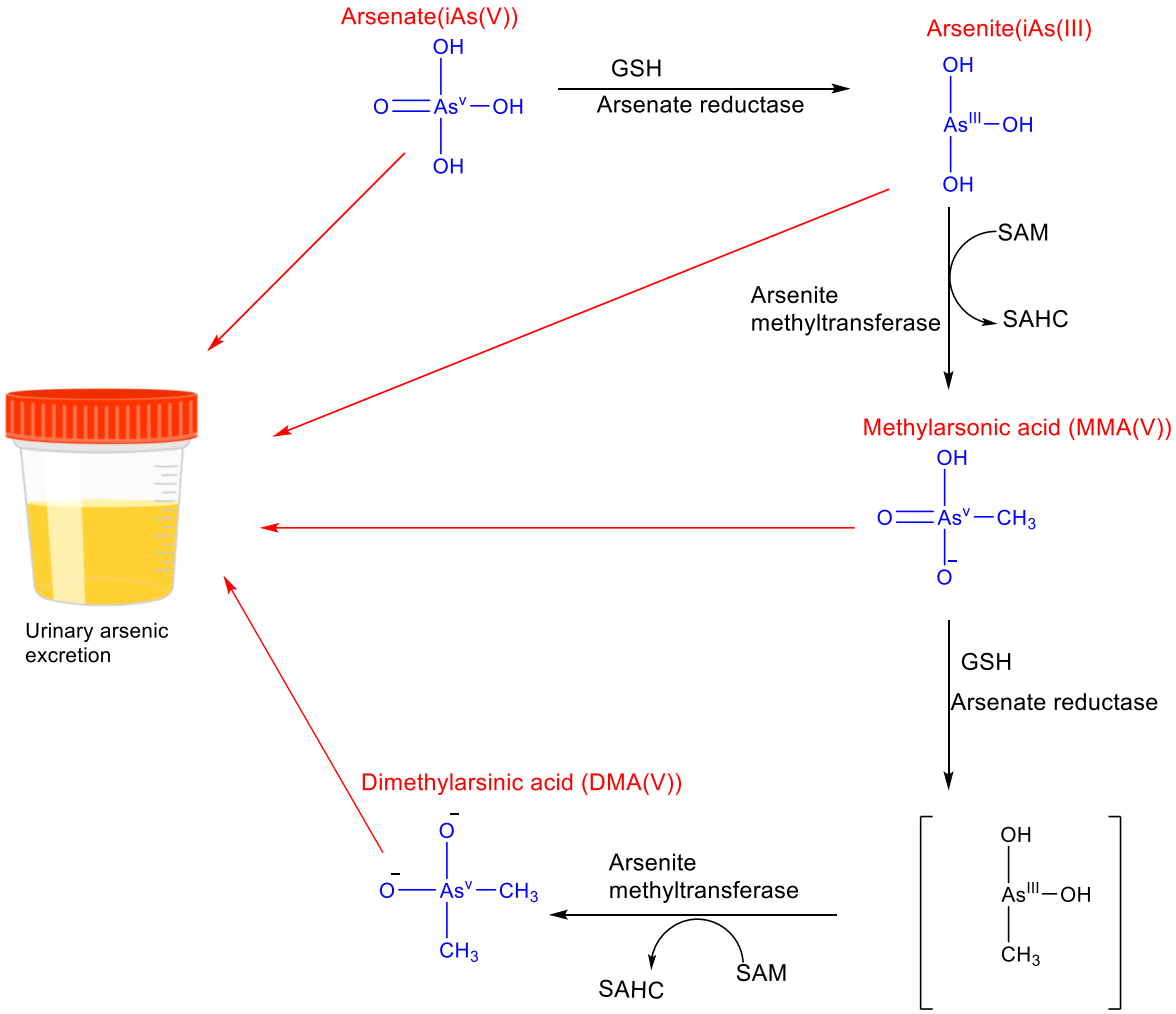


Figure 1.4: Pathway of arsenic biotransformation in humans.[23,24]

Approximately 70-90 percent of inorganic arsenic and 100 percent of organic arsenic is absorbed by the gastrointestinal tract and then transported to other organs. After coming into contact with the skin and other organs, such as the gastrointestinal tract and the epididymis, inorganic arsenic takes a long time to build up in the body. Upon ingestion, the highest arsenic levels are detected in hair, nails, skin, and lungs. The short-term exposure to arsenic can be determined after the metabolism of inorganic arsenic (iAs) in the liver, with the resulting arsenic metabolites excreted in the urine as free inorganic arsenic (As (III) and As (V)) in concentrations of 10-20 percent, MMA in concentrations of 10-15 percent, and DMA in concentrations of 60-80 percent.[23–25]

1.5 Health effects

Arsenic is a chemical toxin that is prevalent throughout the environment and has a significant impact on a broad range of organ systems in the human body. Its chronic and acute health impacts are determined by the kind of arsenic species to which one is exposed, the amount of arsenic absorbed, and the interaction of inherent and extrinsic variables. The mode through which a person is exposed to arsenic also influences whether that individual may have systemic health repercussions such as cancer.[26]

Inorganic arsenic and other arsenic compounds in the air can cause local irritation if they come into contact with the skin or if inhaled. However, systemic effects such as dermatological, cardiovascular, reproductive, and respiratory effects have also been documented. Inhalation and skin contact are the most common routes of exposure in the workplace, and in a significant number of cases, they are associated with carcinogenic health effects.[26]

As a result of the overwhelming evidence that arsenic is a carcinogen in humans, the International Agency for Research on Cancer categorized it as a Class I carcinogen (IARC). The association between chronic arsenic exposure and cancer is strongest in skin, lung, and bladder malignancies among other cancers.[27]

Inflammatory and erosive lesions of the mucosa of the airways, including perforation of the nasal septum have occurred in smelter employees who have been exposed to arsenic at unusual levels for a prolonged period. Long-term exposure to arsenic at smelters and pesticide industries caused many workers to develop lung cancer. Angiosarcoma, a rare form of lung cancer, has been linked/associated with arsenic exposure over a long period.[28]

1.6 Clinical treatment

Chelating drugs administered within hours of arsenic ingestion are such as meso-2,3-dimercaptosuccinic acid (DMSA) are effective in avoiding the full effects of acute arsenic toxicity; therefore, a therapeutic response to arsenic poisoning requires an effective antidote. Previously, 2,3-dimercaptopropanol (dimercaprol) Figure 1.5 was suggested as an antidote for arsenicosis. Dimercaprol is a sulfhydryl reducing agent, its main interaction with arsenic is

through the sulfhydryl bond. The currently proposed therapy is 2,3-dimercapto-1-propanesulfonate (DMPS) or meso-2,3-dimercaptosuccinic acid (DMSA), which are more water-soluble and can be administered orally and has lower toxicity.[29,30]

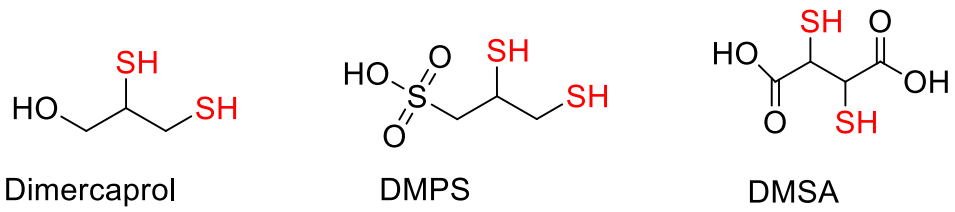


Figure 1.5: Structure of common chelating agents used as for the treatment of arsenic poisoning.

1.7 Biomarkers of exposure

Biological markers (Biomarker) are naturally occurring substances found in biological systems that can be used as qualitative or quantitative indicators of toxin exposure (physiological, or molecular) or clinical disease. Biomarkers are often used in diagnostics to identify the onset, progression, or endpoint of a pathological process or a response to a therapeutic or a pharmacological intervention. To identify and quantify the amount of arsenic exposure, samples of hair, nails, blood, and urine can all be used. Since arsenic has a strong tendency to accumulate in keratin-rich tissue, total arsenic levels in hair, fingernails, and toenails are used in the determination of previous exposure or long-term exposure to arsenic (observed only two weeks after exposure). Lab based extraction methods are necessary with hair, fingernails, and toenails and therefore these samples cannot be used onsite in the field.[31,32]

However, the half-life of arsenic is much shorter in blood, and it is swiftly eliminated from the body. Although arsenic can only be detected in the blood for a few hours (2-3 hours), more than 90% of arsenic exposure is found in the urine 48-72 hours after exposure. In summary, this means that when using a point of care device urine is currently the most reliable matrix for measuring arsenic exposure values as well as arsenic metabolites such as monomethyl arsenic acid (MMA) and dimethyl arsenic acid (DMA).[21]

Because urinary arsenic is rapidly excreted, measurements must be taken as soon as possible after an arsenic exposure event to identify arsenic exposure. After exposure to inorganic arsenic, it takes about 10 hours for arsenite (AsIII) and arsenate (AsV) values to reach their

maximum and a further 20 to 30 hours for the concentrations to return to normal. Urinary MMA and DMA levels typically peak between 40 to 60 hours after exposure and return to normal between 6 to 20 days after exposure.[33]

1.8 Analytical methods and characterization

1.8.1 Detection of inorganic and organic arsenic

A variety of techniques, such as inductively coupled plasma mass spectrometry (ICP-MS), atomic absorption spectrometry (AAS), and colorimetry, are currently used for arsenic detection. It is however difficult to make real-time decisions using these conventional techniques due to the amount of time that lapses between sample collection, transporting of samples to laboratories, samples analysis, and data processing. Thus, there is an increasing need for arsenic point of care diagnostic kits with a quick and efficient detection method. The use of electrochemical methods appears to be a more promising approach in the field since they permit faster analysis than conventional laboratory equipment. Due to its sensitivity and ability to identify heavy metal traces in various states, stripping voltammetry analysis is the most widely used electrochemical method. A variety of materials such as platinum (Pt), gold (Au), and silver (Ag) have been used as electrochemical electrodes to detect arsenic, with Au electrodes being commonly used for the detection of As (III) due to their non-toxicity and high conductivity. Chemically modified electrodes are gaining popularity due to their ability to increase the sensitivity and selectivity of electrochemical analysis methods. In this study, the designed components are compatible with gold electrodes and are discussed in detail in sections 1.9 and 1.10. [34]

1.8.2 Detection of peptides and peptide-arsenic complexes

1.8.2.1 Reversed-phase liquid chromatography

High-performance liquid chromatography (HPLC) is a widely used technique for the separation and characterization of a variety of compounds including proteins and peptides. Separation can result from normal-phase (NP) chromatography or reverse-phase (RP)

chromatography. In this section the fundamentals of reversed-phase HPLC (RP-HPLC) and the basic components of the HPLC system will be introduced. RP-HPLC is used for multiple applications including the purification of large-scale peptides and the separation of peptides from complex sample matrices before detection by UV or mass spectrometry. Its ability to separate peptides of similar structures has made it an attractive technique in the field of proteomics. Additional advantages of the technique include the sensitive detection and ability to hyphenate with other characterization techniques such as mass spectrometry.[35]

A basic HPLC system is composed of the following components shown in Figure 1.6 below: A solvent reservoir, a pump, an injection valve, an analytical column, a detector unit, and a data acquisition system for processing. The solvent, also known as mobile phase or eluent, is delivered by the pump through the system at constant pressure and speed. The injector (injection valve) introduces the sample into the eluent, and the sample analytes are separated in the analytical column. The separated analytes then migrate to the detector, where the retention time (R_t) and concentration of analyte are measured. The detector detects changes in the composition of the eluent and turns the information into an electrical signal, which is then analysed using a data processor or computer.[36]

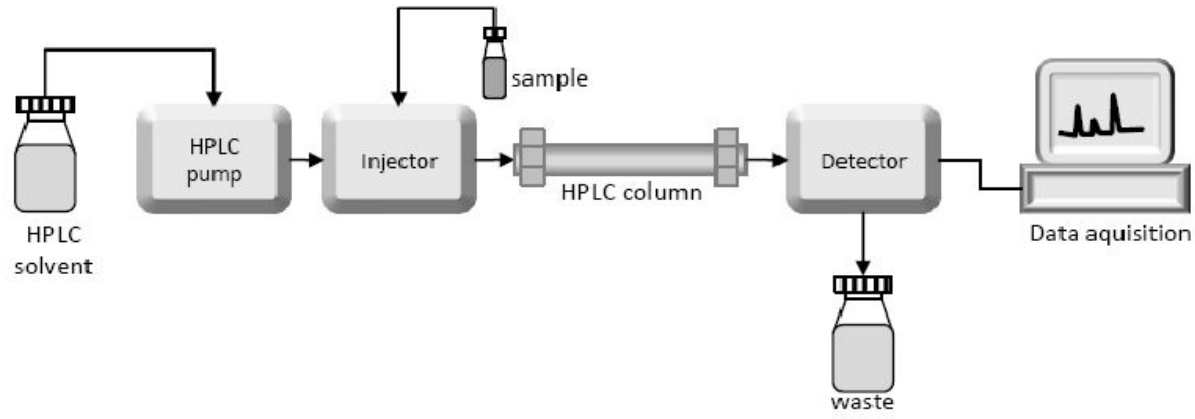


Figure 1.6: Schematic diagram with basic components of HPLC.[37]

The separation of analytes on RP-HPLC is based on the difference in structural and chemical properties of the analytes, including their molecular size, polarity, and electrical charge. During this process, a polar mobile phase and nonpolar stationary phase are used, the distribution and separation of analytes take place between these two phases. Water,

methanol, or acetonitrile are common RP-HPLC eluents that contain ionic additives to help regulate pH, control the ionization of analytes, promote retention by creating ion-pair interactions, minimize the analyte's interactions with the silica surface, or preserve the analyte's solubility or biological activity, among others. The stationary phase is composed of either resin, carbon, or silica particles that have been covalently bound to a non-polar molecule such as alkyl chains or phenyl rings, or a combination of these materials.[38]

Retention of the peptides in the column is caused by the hydrophobic interactions between the hydrophobic component of the peptide known as the "hydrophobic footprint" and the nonpolar stationary phase. Since peptides are large, only the hydrophobic component will temporarily absorb on the stationary phase while the rest of the molecule is still in contact with the mobile phase. When the parameters of the mobile phase are adjusted or changed, the peptide will be desorbed and eluted from the column, which is usually accomplished by increasing the concentration of the organic modifier. Even though peptides still interact with the surface as they go down the column after the first adsorption/desorption, the subsequent interactions are minimal and do not affect the separation. Separation is solely achieved with a single adsorption/desorption as demonstrated in Figure 1.7 below.[39]

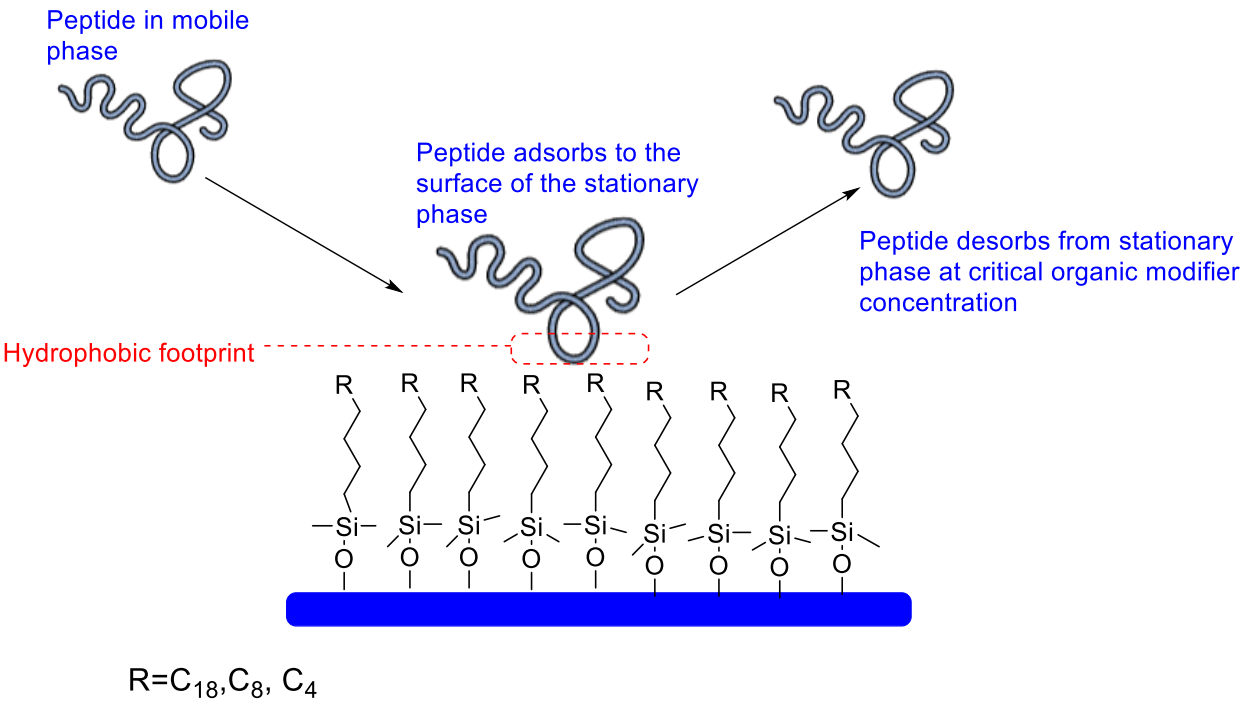


Figure 1.7: Interaction mechanism of peptides with reverse phase HPLC stationary phase.[39]

1.8.2.2 Mass spectrometry of biomolecules

The ability to analyze biomolecules such as peptides and proteins by mass spectrometry has completely transformed the field of proteomics. With chromatographic and spectroscopic methods now available, highly complex peptides can be accurately analyzed. The mass spectrometry technique is useful for studying peptide-ligand interactions, both covalent and noncovalent. This is due to the softness of ionization techniques such as electrospray ionization, which preserves covalent bonds as well as weak noncovalent bonds throughout gas phase transitions. Moreover, mass spectrometry enables the kinetic and mechanistic characterization of target substances, including monitoring substrate consumption and product formation, as well as monitoring peptide-ligand interactions. Chapters 2 and 3 demonstrate the use of mass spectrometry for peptide analysis in greater detail.

1.8.2.3 Mass spectrometer

The overall design of a mass spectrometer is shown in Figure 1.8 below. Mass spectrometers consist of three basic components: the ion source, the mass analyzer, and the detector. While the mass analyzer and detector must be operated under vacuum to be effective in their applications, the ion source may function at atmospheric pressure or in a vacuum depending on the instrument. The electrospray ion source, quadrupole mass analyzer, and time-of-flight mass analyzer will be further reviewed in-depth because they are all integrated for the successful analysis of peptides and their complexes.[40]

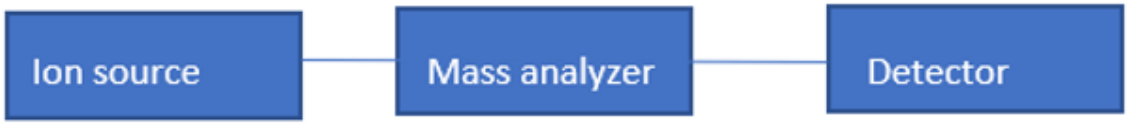


Figure 1.8: The schematic diagram of the fundamental components of a mass spectrometer.[41]

Ion source

The primary function of an Ion source is to generate atomic or molecular ions of analytes. Electrospray Ionization (ESI) is a gentle ionization technique that is frequently used to produce gas-phase ions without fragmenting the molecules. A dilute analyte solution is injected through a needle into a highly charged ESI capillary that is under atmospheric pressure.

Consequently, a fine mist of electrospray (ES) droplets with a high charge (positive or negative) are created from the sample solution as a result of the powerful electric field generated. The nebulizing gas is used to make the spraying process more efficient and helps guide the capillary tip spray to the mass spectrometer. The solvent evaporation or de-solvation process by nitrogen gas reduces the size of the charged droplets. Droplet fission occurs as a result of coulombic repulsion between the charges present in the droplet, resulting in the formation of individual gas-phase analyte ions that are carried on to the mass analyzer through ion optics (Figure 1.9).[42]

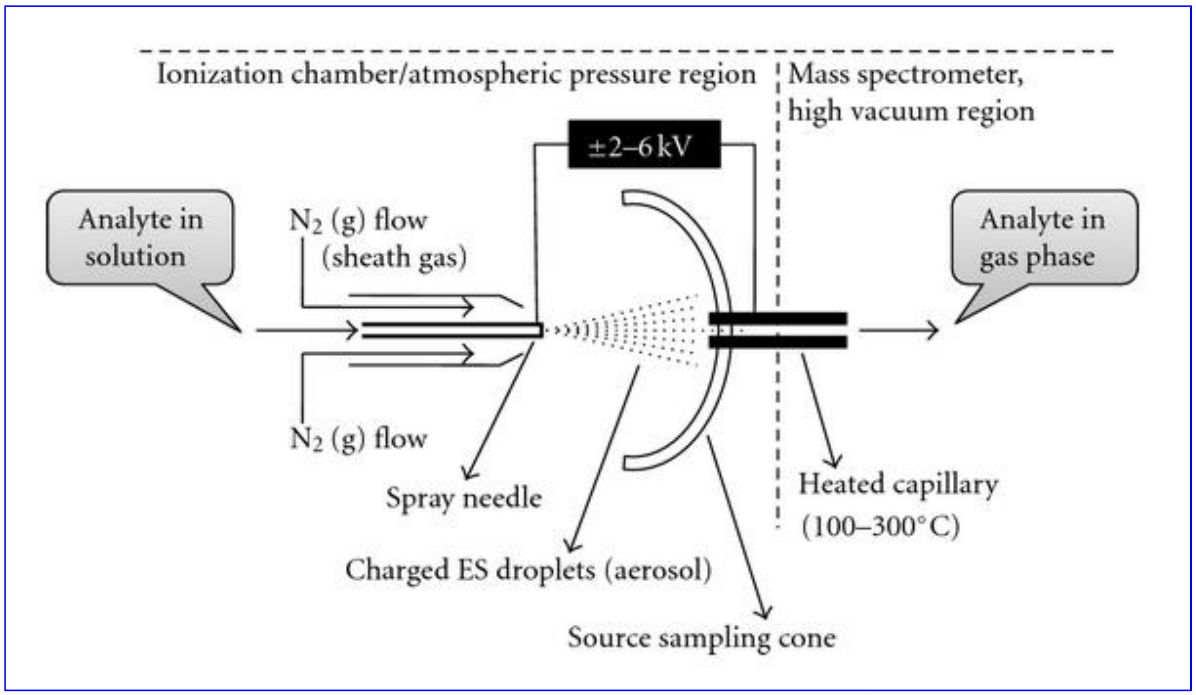


Figure 1.9: Schematic diagram of Electrospray ionization process.[47]

Using the electrospray probe, the analyte ions with low internal energies are guided into the mass spectrometer under the right experimental conditions without any fragmentation as all analytes remain intact. ESI can produce ions that are single- or multiple-charged. There are a variety of factors that influence the charge state of an analyte. These include the size of the analyte molecule, the chemical composition and higher-order structure of the analyte, the composition of the solvent, whether co-solvents were used, and the properties of the analytical equipment. In most cases, when working with compounds with a molecular mass of 2000 Da or less, the ESI generates a single, double, or triple charged ion; however, when working with molecules with a molecular mass greater than 2000 Da, the ESI may generate

an array of numerous charged ions in succession. Peptides and proteins are among the major macromolecules that have been identified as promising candidates for ESI.[44]

Mass analyser

As soon as ions are formed in the ion source, the function of the mass analyser is to sort the different types of ions by separating them based on their mass-to-charge ratio (m/z). Depending on the type of mass analyzer, either an electric or magnetic field is employed to achieve this separation. Neutral molecules are not separated since they have no charge. A few examples of mass analyzers that are currently used include quadrupole ion traps (Q-IT), quadrupoles (Q), and time of flight (TOF). While quadrupoles are scanning devices, time-of-flight analyzers have the benefit of monitoring ions of all masses at the same time, which allows for better sensitivity to dissociation (MSn).[45]

The quadrupole mass analyzer employs oscillating electric fields to distinguish between ions based on their mass-to-charge ratio (m/z ratio). It is made up of four parallel metal rods, and between each pair of rods, a radio frequency (RF) voltage with a DC offset voltage is applied. A mass filter can be used to maintain a stable trajectory by only permitting ions with a specific mass-to-charge ratio (m/z) to pass through it. A variety of ions with varied mass-to-charge ratios hit the rods or walls and are prohibited from passing through the rods or walls as a result of their collision. Time-of-flight mass analysis is based on ions traveling through a flight tube for a predetermined amount of time before reaching a detector. For sources that generate an uninterrupted ion beam, such as the ESI source, the ions must be pulsed and extracted orthogonally for each pulse.[45–47]

By sequentially connecting multiple mass analyzers, it is possible to develop hybrid devices with improved resolution. This method is frequently carried out with the use of quadrupole and time-of-flight mass spectrometers (Q-TOF). This means that in addition to individual mass selection, it also allows for sequential execution of the mass analysis with an interposed collision cell, which can be helpful in terms of providing structural information. In Figure 1.10 the conceptual design of an ESI-Q-TOF device is shown.[47]

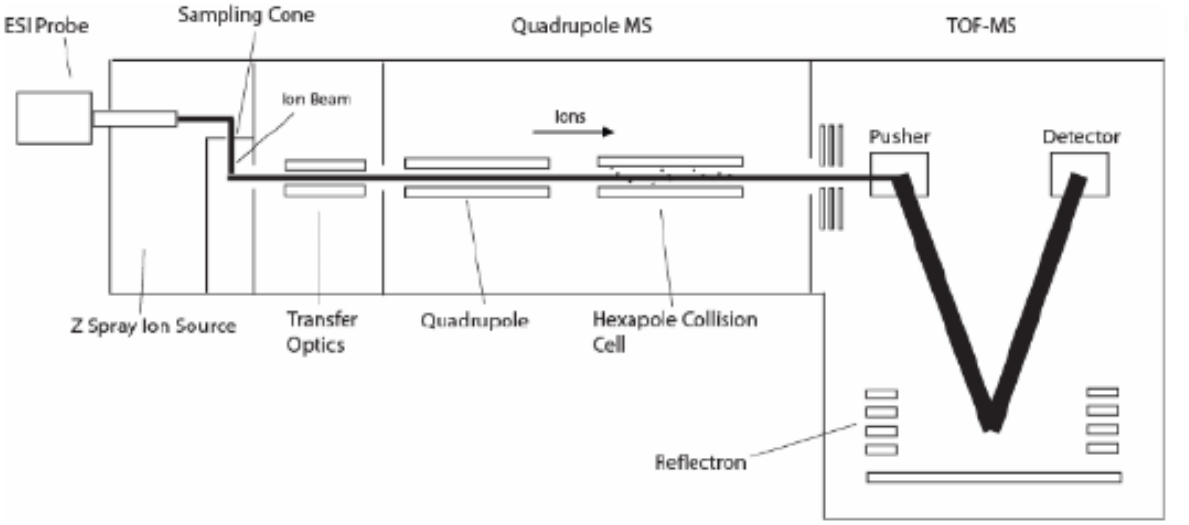


Figure 1.10: Schematic diagram of a Q-TOF mass analyser.[48]

Detector

The primary role of a detector is to record the number of ions produced and separated for each m/z. When they collide with the detector surface, it detects their collisions and amplifies the signal produced by the interactions. For this purpose, the ion beam is focused on a single point detector such as the Faraday Cup, or the scattered ion beam is simultaneously recorded over a plane, as when using a diode array detector. The detector specifications depend on the analyser. Due to the successive separation of the ions by quadrupoles, the detector can only detect one m/z at a time to scan the entire spectrum. Since ions with different m/z ratios do not leave the TOF tube at the same time, single-point detectors can be used with TOF analysers. Instead of a traditional array, TOF devices often use a microchannel plate detector that has numerous single-point detectors.[49]

The following are some examples of binding studies conducted in the field of proteomics using the electrospray ionization mass spectrometry technique: Kitova et al. (2012) investigated the interaction between proteins and their ligands using electrospray ionization mass spectrometry. They were able to bind their proteins to ligands like proteins, carbohydrates, lipids, DNA, or small molecules, and they were also able to determine the stoichiometry and binding affinity using the technique.[50] To determine the presence of arsenic in the peptides, Schmidt et al. (2012) used reversed-phase liquid chromatography coupled to electrospray ionization mass spectrometry to distinguish arsenic-bound peptides from their unbound

original forms. They were able to determine whether a binding site (cysteine residues) for arsenic compounds was available based on their findings, which revealed the presence or absence of structure-stabilizing disulfide bonds. LC-ESI-MS was used to determine partial reaction orders in kinetic studies.[51]

1.9 Biosensor

A biosensor is an interconnected analytical device that is capable of detecting a change in its biological or chemical environment and generating quantitative or semi-quantitative signals proportional to the change. An electrochemical biosensor is a biosensor that uses an electrode as a transducing element. The electrochemical transduction element is connected to a biological recognition element called a bioreceptor. Figure 1.11 below shows the general structure of an electrochemical biosensor where detection occurs in three stages. The analyte first binds to the bioreceptor to produce a biological signal. This is followed by transduction where the formation of a bond or complex between the analyte and the biomolecule is converted into an electrical, chemical, or physical signal. Lastly, the signal is processed, and a computer algorithm is used to convert it into usable information, such as concentration values.[52,53]

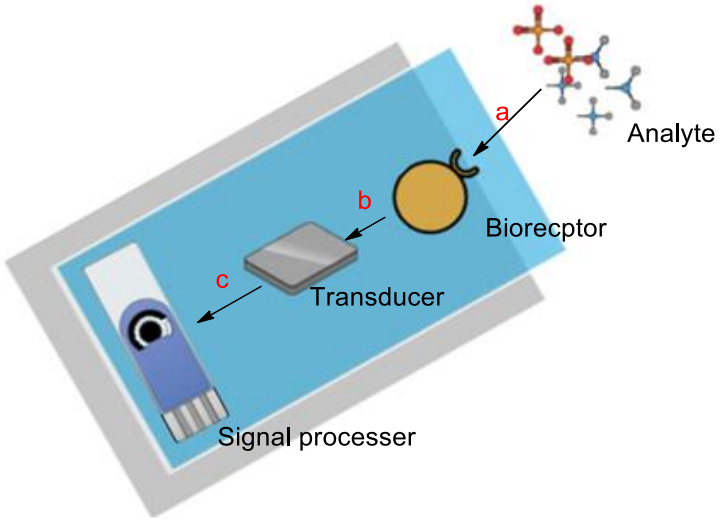


Figure 1.11: Structure of a biosensor: a) Biological signal, b) Physicochemical signal, c) Electrical signal.[54]

An electrochemical detection is a form of transduction where the potential difference of a particular compound or complex in solution is measured in comparison to the reference

electrode. This detection mechanism is based on the production or consumption of electrons or ions during the binding step which causes a change in the electrical properties of the solution. This means that the electrochemical signal observed is influenced by the binding activity of the analyte and not the concentration. This technique also measures the current that is generated during the redox reaction which can either be attributed to the analyte concentration or rate of consumption. The observed electrical signal results from the biorecognition process between the bioreceptor and analyte of interest, and the intensity is proportional to the concentration of the analyte.[55]

1.10 Electrochemical cell

An electrochemical cell is used to conduct experiments during an electrochemical sensor development. The type of electrodes used has a great influence on the performance of the electrochemical cell. The analytical instrument used for the measurements of the cell is called an electrochemical analyser. This device measures the voltage and/or current in an electrochemical cell. The conventional three-electrochemical cell is presented in Figure 1.12 below.[56]

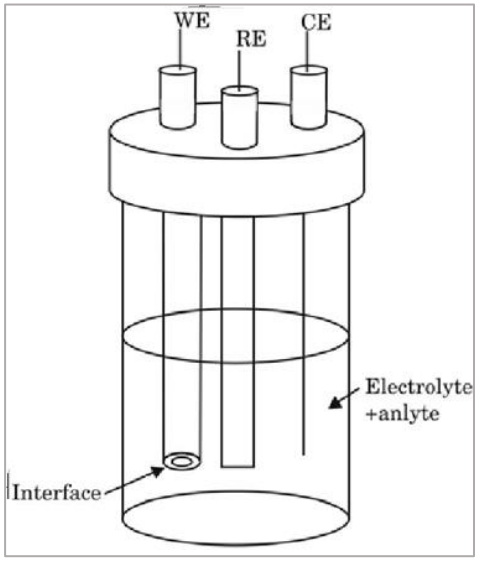


Figure 1.12: Schematic representation of conventional three-electrochemical cell where WE, RE, and CE stands for working electrode, reference electrode, and counter electrode.[56]

The working electrode (WE) is the electrode where the redox reactions of the electrochemical cell take place. The electrode can either be anodic or cathodic depending on whether an oxidation or reduction reaction is taking place. The working electrode needs to produce a

reproducible response as well as high signal-to-noise characteristics. The type and quality of material used for the working electrode will therefore influence the measurements of the voltametric experiments. When selecting a working electrode, one needs to consider the redox behaviour of the analyte of interest, the background current over the potential region, electrical conductivity, surface reproducibility, cost, availability, and toxicity. Currently used materials for working electrodes include glassy carbon electrode, platinum (Pt) electrode, gold (Au) electrode, mercury (Hg) electrode, screen-printed electrode, silver (Ag) electrode, Indium tin oxide-coated glass electrode, carbon paste electrode, etc. The commonly used WE materials for voltammetry include platinum, gold, mercury, and glassy carbon.[55,56]

The reference electrode (RE) is an electrode whose potential is stable enough to serve as a reference standard against which the potential of the other electrodes in the cell can be evaluated. When selecting material for the reference electrode, one should consider using components that are stable over time and which, when exposed to temperature changes, keep activity levels within a defined range as it is important for electrode potentials to remain constant and reproducible. The standard hydrogen electrode (SHE), silver-silver chloride electrode (Ag/AgCl/ KCl), and the Kalomel electrode (Hg/HgCl/KCl) are some of the most commonly used reference electrodes.[55,56]

The last electrode in the electrochemical cell is the counter electrode (CE), also known as the auxiliary electrode. The CE serves as an electron source, enabling electrons to flow from the external circuit through the cell to the internal circuit, thereby closing the current circuit. When selecting the material for the CE, the two most important considerations to keep in mind are that it must be inert and that it must not participate in the electrochemical process in any way. Since the current flows between CE and WE, the CE is expected to have a total surface area that is larger than that of the WE as a smaller surface area may cause it to kinetically become a limiting factor in the electrochemical reactions being investigated. In instances where the WE is acting as an anode, the CE acts as a cathode, and when the WE is acting as a cathode, the CE acts as an anode.[55,56]

A variety of materials such as platinum (Pt), gold (Au), and silver (Ag) have been used as WE in the electrochemical detection of arsenic, with Au electrodes being commonly used for the

detection of As (III) due to their non-toxicity and high conductivity. In this study, the gold electrode will be chemically modified by coupling it with arsenic-binding peptides.[56]

1.11 Peptide-based bioreceptors

Peptides are polymers of amino acids that can be found in nature or synthesized, and they are constructed from the same building blocks as proteins. Since many proteins bind their target ligands with great specificity and selectivity, peptides with specific amino acid sequences can also bind their targeted metabolites with the same selectivity and specificity. Peptides have several advantages over proteins, including reduced chances of denaturing, peptides have better conformational and chemical stability.[57]

Depending on the intended application, the amino acid sequence of peptides is typically derived from a natural protein. The primary barrier preventing full utilization of proteins is the low yields caused by difficulty in protein synthesis and purification. Biological and chemical researchers have greatly benefited from the discovery of peptide synthesis in the 1950s and 60s, and it has since been applied in a variety of applications to overcome the challenges outlined above. Many studies have been carried out to determine whether peptides are suitable probe molecules for biosensors and how they could be used. Due to the availability of chemical and molecular biological methods and the possibility of automated synthesis, peptides make excellent candidates for designing artificial molecular receptors. As an additional benefit, the reagents are relatively inexpensive and very easy to modify to improve the interaction. There have also been reports of increased stability and robustness in the detection process.[58]

Consequently, peptides are emerging as extremely promising synthetic biomimicry agents. The peptide field has advanced rapidly since Merrifield discovered solid-phase peptide synthesis in 1984, and the technique has been applied to a variety of applications due to the high yields and ease of producing highly purified products.[59,60]

One example from the literature of peptides successfully used for binding applications includes the work of Kirk Kitchin and Kathlen Wallace who determined the binding affinity of radioactively labelled arsenite (^{73}As) using synthetic peptides derived from the estrogen

receptor protein. Their peptides were derived from two regions of the protein (the zinc finger portion and the hormone-binding portion). Their synthetic peptides exhibited positive binding and binding constants were determined successfully. Their study reported specific binding as amino acids other than cysteine failed to bind arsenite.[61]

Using the Ars operon, bacteria such as Escherichia coli (E. coli) have developed resistance to arsenic. The resistance is achieved by a repressor protein called ArsR, which dissociates from DNA upon binding to As(III). When bound to As(III) with three cysteine residues, this protein undergoes a critical conformational change. The arsenic resistance characteristics and binding activity of the repressor protein have become the motivation and driving force behind this research.[62] The arsenic binding regions were identified in the protein and all the designed novel peptides were derived from these areas, more details about the peptides are provided in chapter 2.

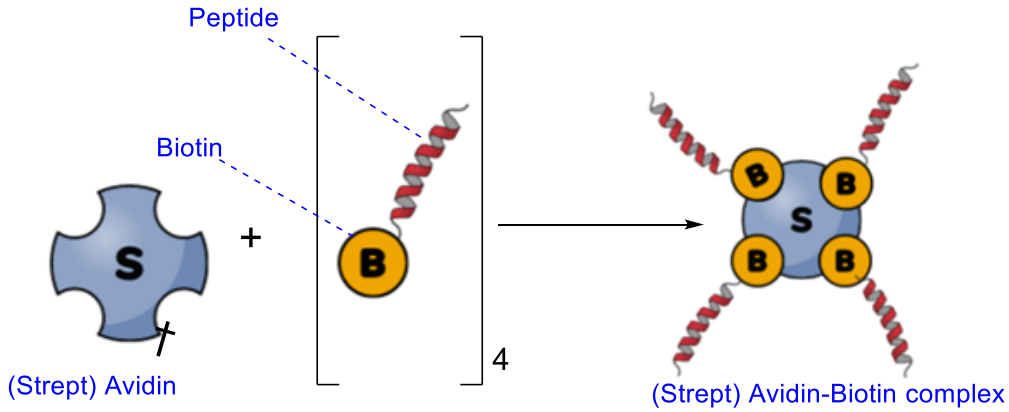
1.12 Modes of peptide immobilization

The immobilization of bioreceptor's on a solid support is an important process during the construction of biosensors that should not be overlooked. Since the current study focuses on the use of peptides as the bioreceptor, the immobilized peptide should have no steric hindrance as a result of its immobilization. The binding site should also remain accessible, allowing for the favourable interaction between the ArsR-derived peptides and the target urinary arsenic metabolites. Additionally, the chemical characteristics of the peptide must be preserved during immobilization, and the peptide binding ability should remain unchanged after immobilization. The immobilization method has an impact on sensor sensitivity and dynamic range. Various strategies have been developed to immobilize a bioreceptor onto a solid surface, including covalent bonding, physical adsorption, and electrostatic/physical integration into a polymer matrix. The biotin (strept) avidin complex is one of the commonly used immobilization strategies, and it is discussed in greater detail below.[63]

1.12.1 Biotin-(Strept) avidin interaction

The (strept) avidin-biotin interaction is a technique that is widely used in immunoassays to immobilize various types of biomolecules in different ways. This procedure entails the biotinylation of the bioreceptor (peptide) and the coating of the solid surface (electrode) with either avidin or streptavidin before performing the experiment. The interaction between avidin and biotin that results in the formation of the (Strept)avidin-biotin complex is depicted in Scheme 1.1 below. It has been shown that the (strept) avidin-biotin complexes are stable at high temperatures, can withstand pH fluctuations, and are not dissociated when exposed to aggressive chemicals. Even after immobilization, the biological function of the biomolecules is maintained. The interaction between avidin and biotin can be used to develop assays that are both robust and highly sensitive.[63,64]

In this study, we anticipate that the designed biotinylated-peptide will bind to the (strept) avidin and form a complex that can be used bind low levels of arsenic in urine.



Scheme 1.1: Avidin-biotin interaction. † Denotes that Avidin is also often conjugated to an immobilized support.[65]

1.13 Background and aims of dissertation

1.13.1 Problem statement

Heavy metal (HMs) pollution in developing countries such as South Africa is a major health and environmental concern. This can result in an adverse range of health effects especially in provinces such as the Northern Cape and Limpopo where HMs such as arsenic have been reported to be as high as 1000 µg/L. Other provinces such as Free State, Eastern Cape, Kwa Zulu-Natal, and Mpumalanga where the country's coal mines are located have also reported large quantities of arsenic exposure.

Instrumental methods are the first point of analysis for arsenic detection; however, there is an increased need for a fast and efficient detection technique that can be used as a diagnostic kit. This is due to the amount of time that lapses between sample collection, analysis, and data processing with the conventional techniques (making it difficult to make real-time decisions). Hence, there is a need for arsenic biosensors that are synthetically designed to be selective, sensitive, economical, convenient, and rapid for the early detection of arsenic poisoning.

1.13.2 Aim

The study aims to qualitatively and quantitatively investigate the binding affinity of arsenic to sulfur-containing ArsR derived peptides using the LC-ESI-MS and to develop components of a ArsR derived peptide-based biosensor which will be used for arsenic monitoring in South African mines.

1.13.3 Objectives

1. To synthesize ArsR derived peptides using the solid phase peptide synthesis strategy.
2. To study the interactions between arsenic and ArsR derived peptides, their mechanism as well as the structural and conformational changes observed in the presence of arsenic using the UHPLC Q-TOF High-resolution mass spectrometry instrument.
3. Develop an LC-ESI-Method for the qualitative and quantitative analysis of unbound vs arsenic-bound ArsR-derived peptides in urine matrix.
4. To determine the binding constants of the peptide-arsenic complexes.
5. To optimise the peptide-arsenic binding by varying pH, temperature, and time.

6. Monitor the selectivity of the peptide in the presence of other metallic species.
7. To immobilize the selective peptide onto a gold electrode using the biotin-(strept) avidin technique

1.13.4 Novelty of study

This dissertation proposes a novel diagnostic device that uses novel ArsR-derived peptides as probe molecules for the detection of urinary arsenic metabolites (iAs(V), (iAs(III), and DMA(V).

References

(1) Nelson, G. Occupational Respiratory Diseases in the South African Mining Industry. *Glob. Health Action* **2013**, *6* (1). https://doi.org/10.3402/gha.v6i0.19520.

(2) Ebenebe, P.; Mahmood Sedibe, M.; Achilonu, M. South African Mine Effluents: Heavy Metal Pollution and Impact on the Ecosystem Citation: Ebenebe PC, Shale K, Sedibe M, et al. South African Mine Effluents: Heavy Metal Pollution and Impact on the Ecosystem South African Mine Effluents: Heavy Metal Pollu. *Artic. Int. J. Chem. Sci.* **2018**, *15* (4), 198.

(3) Guild, R. A Handbook on Occupational Health Practice in the South African Mining Industry . Safety in Mines Research Advisory Committee : Johannesburg, South Africa 2001.

(4) Utembe, W.; Faustman, E. M.; Matatiele, P.; Gulumian, M. Hazards Identified and the Need for Health Risk Assessment in the South African Mining Industry. *Hum. Exp. Toxicol.* **2015**, *34* (12), 1212–1221. https://doi.org/10.1177/0960327115600370.

(5) Dhatrak, S. V; Nandi, S. S. Risk Assessment of Chronic Poisoning among Indian Metallic Miners . *Indian journal of occupational and environmental medicine* . Medknow Publications : India 2009, pp 60–64. https://doi.org/10.4103/0019-5278.55121.

(6) Chung, J. Y.; Yu, S. Do; Hong, Y. S. Environmental Source of Arsenic Exposure. *J. Prev. Med. Public Heal.* **2014**, *47* (5), 253–257. https://doi.org/10.3961/jpmph.14.036.

(7) Das, N. K.; Sengupta, S. R. Arsenicosis: Diagnosis and Treatment . *Indian journal of dermatology, venereology, and leprology* . Medknow Publications and Media Pvt. Ltd : India 2008, p 571. https://doi.org/10.4103/0378-6323.45098.

(8) Epidemiology, E.; Office, T.; Health, P.; Law, L. Information for Health Care Professionals Arsenic Exposure & Toxicity Revised: *Sect. Environ. Epidemiol. Toxicol. Off. Public Heal. Louisiana Dep. Heal. Hosp.* **2008**, *12* (888), 293–7020.

(9) Singh, G.; Singh, A.; Shukla, R.; Karwadiya, J.; Gupta, A.; Naheed, A.; Mishra, V. K. Occurrence, Fate, and Remediation of Arsenic. In *Pollutants and Water Management*; 2021; pp 349–376. https://doi.org/10.1002/9781119693635.ch14.

(10) Hindmarsh, J. T.; McCurdy, R. F.; Savory, J. Clinical and Environmental Aspects of Arsenic Toxicity. *Crit. Rev. Clin. Lab. Sci.* **1986**, *23* (4), 315–347. https://doi.org/10.3109/10408368609167122.

(11) Farmer, J. G.; Johnson, L. R. Assessment of Occupational Exposure to Inorganic Arsenic Based on Urinary Concentrations and Speciation of Arsenic. *Br. J. Ind. Med.* **1990**, *47* (5), 342–348. https://doi.org/10.1136/oem.47.5.342.

(12) Henke, K. Arsenic Chenistry. In *Arsenic: Environmental Chemistry, Health Threats and Waste Treatment*; John Wiley & Sons Ltd: The Atrium, Southern Gate, Chichester, West Sussex, United Kingdom, 2009; pp 9–59.

(13) Neil Burford, Yuen-ying Carpenter, E. C. and C. D. L. S. The Chemistry of Arsenic, Antimony and Bismuth. In *Biological Chemistry of Arsenic, Antimony and Bismuth*; 2011; p 2547.

(14) Flora, S. J. S. Arsenic: Chemistry, Occurrence, and Exposure. In *Handbook of Arsenic Toxicology*; 2015; pp 1–49.

(15) Hue, N. V. Arsenic Chemistry and Remediation in Hawaiian Soils. *Int. J. Phytoremediation* **2013**, *15* (2), 105–116. https://doi.org/10.1080/15226514.2012.683206.

(16) Hughes, M. F. Arsenic Toxicity and Potential Mechanisms of Action. *Toxicol. Lett.* **2002**, *133* (1), 1–16. https://doi.org/10.1016/S0378-4274(02)00084-X.

(17) Booth, J. W.; Guidotti, G. Phosphate Transport in Yeast Vacuoles. *J. Biol. Chem.* **1997**, *272* (33), 20408–20413. https://doi.org/10.1074/jbc.272.33.20408.

(18) Martinez, L. M. Introduction to Arsenic, Visceral Smooth Muscle and Oxidative Stress, 2006. https://doi.org/10.1016/S0079-6123(06)54019-1.

(19) Rosen, B. P. Families of Arsenic Transporters . *Trends in Microbiology* . Elsevier Ltd : England 1999, pp 207–212. https://doi.org/10.1016/S0966-842X(99)01494-8.

(20) Tamaki, S. and Frenkenberger, W. Environmental Biochemistry of Arsenic. *Environ. Contam. Toxicol.* **1992**, 79–110.

(21) Thompson, D. J. A Chemical Hypothesis for Arsenic Methylation in Mammals. *Chem. Biol. Interact.* **1993**, *88* (2–3), 89–114. https://doi.org/10.1016/0009-2797(93)90086-E.

(22) Cullen, W. R.; McBride, B. C.; Reglinski, J. The Reduction of Trimethylarsine Oxide to Trimethylarsine by Thiols: A Mechanistic Model for the Biological Reduction of Arsenicals. *J. Inorg. Biochem.* **1984**, *21* (1), 45–60. https://doi.org/10.1016/0162-0134(84)85038-2.

(23) Khairul, I.; Wang, Q. Q.; Jiang, Y. H.; Wang, C.; Naranmandura, H. Metabolism, Toxicity and Anticancer Activities of Arsenic Compounds. *Oncotarget* **2017**, *8* (14), 23905–23926. https://doi.org/10.18632/oncotarget.14733.

(24) Tchounwou, P. B.; Yedjou, C. G.; Udensi, U. K.; Pacurari, M.; Stevens, J. J.; Patlolla, A. K.; Noubissi, F.; Kumar, S. State of the Science Review of the Health Effects of Inorganic Arsenic: Perspectives for Future Research. *Environ. Toxicol.* **2019**, *34* (2), 188–202. https://doi.org/10.1002/tox.22673.

(25) Hopenhayn-rich, A. C.; Biggs, M. Lou; Smith, A. H.; Kalman, D. A.; Moore, E.; Hopenhayn-rich, C.; Biggs, M. Lou; Smith, A. H.; Kalman, D. A.; Moore, L. E. Methylation Study of a Population Environmentally Exposed to Arsenic in Drinking Water. *Environ. Heal. Perspect.* **2014**, *104* (6), 620–628.

(26) Materials, B. *Toxicological Profile for Arsenic (Draft for Public Comment).*; 1989.

(27) Martinez, V. D.; Vucic, E. A.; Becker-Santos, D. D.; Gil, L.; Lam, W. L. Arsenic Exposure and the Induction of Human Cancers . *Journal of toxicology* . Hindawi Publishing Corporation : Egypt 2011, pp 431213–431287. https://doi.org/10.1155/2011/431287.

(28) Popper H, Thomas LB, Telles NC, Falk H, S. I. Devlopment of Hepatic Angiosarcoma in Man Induced by Vinyl Chloride, Thorotrast, and Arsenic. *Am. J. Pathol.* **1978**, *92* (2), 349–376.

(29) Hafeman, D. M.; Ahsan, H.; Louis, E. D.; Siddique, A. B.; Slavkovich, V.; Cheng, Z.; Van Geen, A.; Graziano, J. H. Association between Arsenic Exposure and a Measure of Subclinical Sensory Neuropathy in Bangladesh. *J. Occup. Environ. Med.* **2005**, *47* (8), 778–784. https://doi.org/10.1097/01.jom.0000169089.54549.db.

(30) Mazumder, D. N. G. Chronic Arsenic Toxicity: Clinical Features, Epidemiology, and Treatment: Experience in West Bengal. *J. Environ. Sci. Heal. - Part A Toxic/Hazardous Subst. Environ. Eng.* **2003**, *38* (1), 141–163. https://doi.org/10.1081/ESE-120016886.

(31) Gupta, R. C. *Biomarkers in Toxicology*; Mica H. Haley, 2014.

(32) Maiga, D. G. A. F. L. J. M. S. B. A. H. Arsenic in African Water. *Water Air Soil Pollut.* **2012**, *226* (302), 1–13.

(33) Wenzel, R. Arsenic Speciation in Urine by Solvent Extraction, Graphite Furnace Atomic Absorption Spectrometry and Capillary Electrophoresis [and] Inductively Coupled Plasma Mass Spectrometry, 2001.

(34) Yang, T.; Zhang, X. X.; Yang, J. Y.; Wang, Y. T.; Chen, M. L. Screening Arsenic(III)-Binding Peptide for Colorimetric Detection of Arsenic(III) Based on the Peptide Induced Aggregation of Gold Nanoparticles. *Talanta* **2018**, *177* (April), 212–216. https://doi.org/10.1016/j.talanta.2017.07.005.

(35) Aguilar, M.-I. *HPLC of Peptides and Proteins*; Humana press, 2003.

(36) Moldoveanu, S.; David, V. *Essentials in Modern HPLC Separations*; 2013. https://doi.org/10.1016/C2010-0-65748-8.

(37) Czaplicki, S. Column Chromatography. In *Chromatography in Bioactivity Analysis of Compounds*; 2014; pp 99–122.

(38) Thermo Electron Corporation. HPLC Analysis of Biomolecules - Technical Guide Successful Separations of Peptides, Proteins and Other Biomolecules. *Guideline*. 2013.

(39) Carr, D. A Guide to the Analysis and Purification of Proteins and Peptides by Reversed-Phase HPLC. *ACE HPLC columns*. 2012, pp 1–7. https://doi.org/10.5402/2012/359572.

(40) Sundqvist, G. Analysis of Noncovalent and Covalent Protein-Ligand Complexes by Electrospray Ionisation Mass Spectrometry, 2008.

(41) Lanças, F. M. The Role of the Separation Sciences in the 21th Century. *J. Braz. Chem. Soc.* **2003**, *14* (2), 183–197. https://doi.org/10.1590/S0103-50532003000200005.

(42) Wilm, M. S.; Mann, M. Electrospray and Taylor-Cone Theory, Dole's Beam of Macromolecules at Last? *Int. J. Mass Spectrom. Ion Process.* **1994**, *136* (2–3), 167–180. https://doi.org/10.1016/0168-1176(94)04024-9.

(43) Chemistry, U. of P.-D. of. Mass Spectrometry Introduction https://www.chem.pitt.edu/facilities/mass-spectrometry/mass-spectrometry-introduction (accessed 2021 -12 -15).

(44) Loo, J. A.; Edmonds, C. G.; Udseth, H. R.; Smith, R. D. Effect of Reducing Disulfide-Containing Proteins on Electrospray Ionization Mass Spectra. *Anal. Chem.* **1990**, *62* (7), 693–698. https://doi.org/10.1021/ac00206a009.

(45) FENN, J. B.; MANN, M.; MENG, C. K.; WONG, S. F.; WHITEHOUSE, C. M. ChemInform Abstract: Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *ChemInform* **1990**, *21* (5). https://doi.org/10.1002/chin.199005359.

(46) Pitt, J. J. J.J.Pitt 2009 - Clin Biochem Rev. Feb; (30) Pages 19–34. **2009**, *30* (February), 19–34.

(47) Banerjee, S.; Mazumdar, S. Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *Int. J. Anal. Chem.* **2012**, *2012*, 1–40. https://doi.org/10.1155/2012/282574.

(48) Roberts, L. D. Defining the Metabolic Effect of Peroxisome Proliferator-Activated Receptor δ Activation A Dissertation Submitted for the Degree of Doctor of Philosophy at the University of Cambridge. **2016**, No. June 2010.

(49) Gross, J. H. Instrumentation. In *Mass Spectrometry*; 2017; Vol. 56, pp 151–292. https://doi.org/10.1021/ac00269a027.

(50) Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. Reliable Determinations of Protein-Ligand Interactions by Direct ESI-MS Measurements. Are We There Yet? *J. Am. Soc. Mass Spectrom.* **2012**, *23* (3), 431–441. https://doi.org/10.1007/s13361-011-0311-9.

(51) Schmidt, A. C.; Mickein, K. Qualitative and Quantitative Characterization of the Arsenic-Binding Behaviour of Sulfur-Containing Peptides and Proteins by the Coupling of Reversed Phase Liquid Chromatography to Electrospray Ionization Mass Spectrometry. *J. Mass Spectrom.* **2012**, *47* (8), 949–961. https://doi.org/10.1002/jms.3025.

(52) Su, W.; Cho, M. S.; Nam, J. Do; Choe, W. S.; Lee, Y. Highly Sensitive Electrochemical Lead Ion Sensor Harnessing Peptide Probe Molecules on Porous Gold Electrodes. *Biosens. Bioelectron.* **2013**, *48*, 263–269. https://doi.org/10.1016/j.bios.2013.04.031.

(53) Serrano, N.; Prieto-Simón, B.; Cetó, X.; Del Valle, M. Array of Peptide-Modified Electrodes for the Simultaneous Determination of Pb(II), Cd(II) and Zn(II). *Talanta* **2014**, *125*, 159–166. https://doi.org/10.1016/j.talanta.2014.02.052.

(54) Shrivastava, A.; Sharma, R. K. Biosensors for the Detection of Mycotoxins. *Toxin Rev.* **2021**, *43* (4), 707–721. https://doi.org/10.1080/15569543.2021.1894175.

(55) Karunakaran Chandran, Kalpana Bhargava, R. B. Introduction to Biosensors. In *Biosensors and Bioelectronics*; 2015; pp 42–45.

(56) Mallappa Kumara Swamy, G. R. R. Electrochemistry of Phytochemicals and Natural Products. In *Phytocompounds: Sources and Bioactivities*; 2019.

(57) Liu, Q.; Wang, J.; Boyd, B. J. Peptide-Based Biosensors. *Talanta* **2015**, *136*, 114–127. https://doi.org/10.1016/j.talanta.2014.12.020.

(58) Fischer, P. M. The Design, Synthesis and Application of Stereochemical and Directional Peptide Isomers: A Critical Review . *Current protein & peptide science* . Bentham Science Publishers Ltd : United Arab Emirates 2003, pp 339–356. https://doi.org/10.2174/1389203033487054.

(59) Mitchell, A. R. Bruce Merrifield and Solid-Phase Peptide Synthesis: A Historical Assessment. *Biopolym. - Pept. Sci. Sect.* **2008**, *90* (3), 175–184. https://doi.org/10.1002/bip.20925.

(60) Pavan, S.; Berti, F. Short Peptides as Biosensor Transducers. *Anal. Bioanal. Chem.* **2012**, *402* (10), 3055–3070. https://doi.org/10.1007/s00216-011-5589-8.

(61) Kitchin, K. T.; Wallace, K. Arsenite Binding to Synthetic Peptides Based on the Zn Finger Region and the Estrogen Binding Region of the Human Estrogen Receptor-α. *Toxicol. Appl. Pharmacol.* **2005**, *206* (1), 66–72. https://doi.org/10.1016/j.taap.2004.12.010.

(62) Touw, D. S.; Nordman, C. E.; Stuckey, J. A.; Pecoraro, V. L. Identifying Important Structural Characteristics of Arsenic Resistance Proteins by Using Designed Three-Stranded Coiled Coils. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (29), 11969–11974. https://doi.org/10.1073/pnas.0701979104.

(63) Jeremy M. Fowler*, Danny K Y Wong, H. Brian Halsall, W. R. H. Recent Developments in Electrochemical Immunoassays and Immunosensors. In *Electrochemical Sensors, Biosensors and their Biomedical Applications*; Academic Press, 2008; pp 115–143.

(64) Scientific, T. F. Avidin-Biotin Interaction https://www.thermofisher.com/za/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/avidin-biotin-interaction.html (accessed 2021 -05 -22).

(65) Technical resource in biotechnology. Avidin Biotin System http://technologyinscience.blogspot.com/2014/01/avidin-biotin-system.html#.YikvUnpBzIV (accessed 2021 -08 -17).

Chapter 2 : Design and synthesis of novel arsenic binding peptides

2.1 Introduction

Researchers have investigated amino acids and peptides as potential recognition elements in electrochemical sensors to better understand how they can be used in the detection of heavy metal ions. Table 2.1 below shows a list of the most common amino acids that appear in the human genetic code, even though there are currently about 500 amino acids identified in nature. In this chapter, the design and synthesis of three ArsR-derived peptides using solid-phase peptide synthesis (SPPS) is presented. We will also discuss Fmoc chemistry, which has been the most widely used technique in recent years. The fundamentals of SPPS as well as the various steps involved (anchoring, deprotection, coupling reaction, and cleavage) along with their mechanisms.[1]

Table 2.1: Common amino acids in nature.

AA name	One letter code	Three letter code
alanine	A	Ala
arginine	R	Arg
asparagine	N	Asn
aspartic acid	D	Asp
cysteine	C	Cys
glutamic acid	E	Glu
glutamine	Q	Gln
glycine	G	Gly
histidine	H	His
isoleucine	I	Ile
leucine	L	Leu
lysine	K	Lys
methionine	M	Met
phenylalanine	F	Phe
proline	P	Pro
serine	S	Ser
threonine	T	Thr
tryptophan	W	Trp
tyrosine	Y	Tyr
valine	V	Val

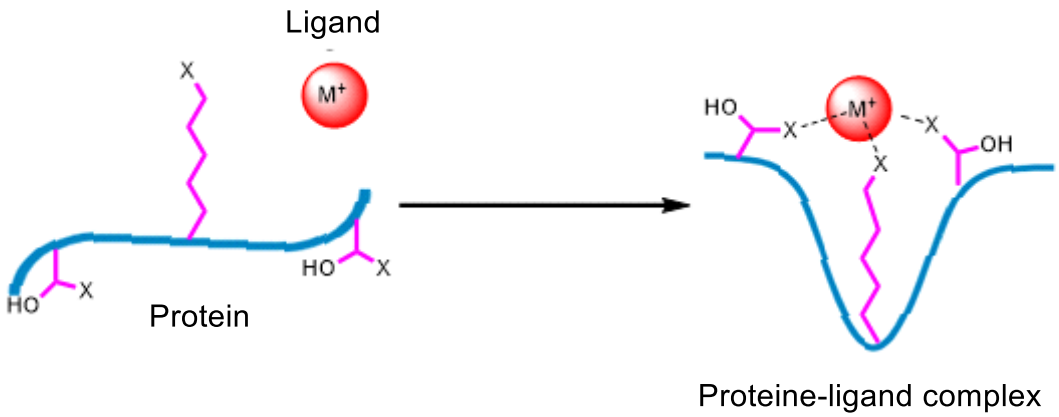
The first step towards the development of a successful peptide-based sensing system involves the careful identification and selection of a sequence of amino acids that have favourable characteristics that can be used in the design of the bioreceptor. Secondly, a simple and reproducible synthetic method should then be developed for the synthesis of the desired bioreceptor. Furthermore, the conformational structure of the bioreceptor/peptide should also be considered, as it can potentially affect the performance and activity of the system. In light of the large number of amino acids available and the large number of peptides that have to be synthesized and tested upon optimization, designing a bioreceptor can be a bit challenging. However, despite these limitations, the synthesis can still be accomplished efficiently with good yields because of the availability of low-cost and well-supported synthetic processes, which allow for the rapid evaluation of a large number of sequences within a short time. Peptides have been identified as attractive alternative bio-receptors due to their competitive nature in sensing applications.[2]

The design and synthesis of genetically encoded peptides with tailored properties can be achieved by using combinatorial libraries. To determine whether the selected peptide can act as the required bioreceptor and, ultimately, which design is truly appropriate for the desired sensing platform, a thorough analysis of the binding kinetics and structural properties of the complex/s should be conducted (Chapter 3). Considering the fact that the binding performance of peptides can vary widely depending on the chemical and physical properties of the peptides, as well as the surrounding environment, it would be prudent to consider the intended use of the peptide sequence (i.e., the biosensor technology) to make sure that the final characteristics of a given peptide are consistent with its intended use.[3,4]

2.2 Relationship between structure and function of a bioreceptor

The three-dimensional structure of proteins and peptides has a profound effect on their ability to bind ligands. During the development of a protein or peptide-based recognition element, the versatility and diversity of the system is largely determined by the amino acid side chains, the flexibility of the polypeptide chain, and the way the amino acid side chains interact with the polypeptide chain. Since proteins are quite flexible molecules, their conformational structure changes when another compound or ligand binds, which ultimately

affects their activity and function. Each protein or peptide binding site has particular abilities to adhere to a specific ligand. This interaction is influenced by the complementary shape and charge distribution of donors or acceptors present on the binding site. Several proteins possess the ability to bind to ligands specifically due to their chemical nature and conformational flexibility. It is because of the chemical nature and conformational flexibility of biomolecules such as proteins and peptides that they are able to bind target ligands with high selectivity and specificity.[4]



Scheme 2.1: General structure of ligand-induced conformational change of a protein.[5]

2.3 Relationship between structure and function of arsenical resistance operon repressor (ArsR)

The high concentrations of arsenic that can be found in the environment have led organisms such as bacteria to develop mechanisms to help them resist arsenic and its effects. Genes associated with arsenic resistance (ars genes) can either be found on the plasmid or the chromosome. Operons found in Escherichia coli (E. Coli) possess five genes (arsR, arsD, arsA, arsB, and arsC) known as the arsRDABC cluster which provides resistance to high levels of arsenic in bacteria. Our study is based on the arsenical resistance operon repressor (ArsR) of plasmid R773, a metalloregulatory protein made up of 117-residues that has shown great selectivity and high affinity towards arsenic. This plasmid has been the source of much of the research on arsenic resistance and control. The ArsR repressor proteins are typically made up of three domains, which include a dimerization domain, DNA-binding domain, and the metal-binding domain.[6–8]

There are two cysteine residues (Cys_{32}, and Cys_{34}) on each subunit of the α-helix, which must be unfolded from one end to the other to bind trivalent inorganic arsenic (iAs(III)) leading to a change in conformation. The repressor protein is then released from the DNA as a result of this conformational change, allowing gene expression to be induced. ArsR binding to iAs(III) is mediated by three cysteine residues (Cys_{32}, Cys_{34}, and Cys_{37}). This means that the helix is positioned in such a way that iAsIII cannot bind until all three thiolates are in position, forming the three-coordinate binding site. When the helix is unwound, it disrupts DNA binding, causing gene expression.[9]

Researchers have identified a region on the arsR protein corresponding to the sequence ELC32VC34DL as a metal-binding domain. This region has self-assembling properties and has interestingly been shown to recognise and bind As(III). Figure 2.1 below illustrates a model of the interaction between iAsIII and the ArsR repressor protein in E.coli. The metal-binding domain of the ArsR repressor protein has attracted our attention and is the primary motivation behind the design and synthesis of peptides with arsenic binding properties for biosensor application.[10]

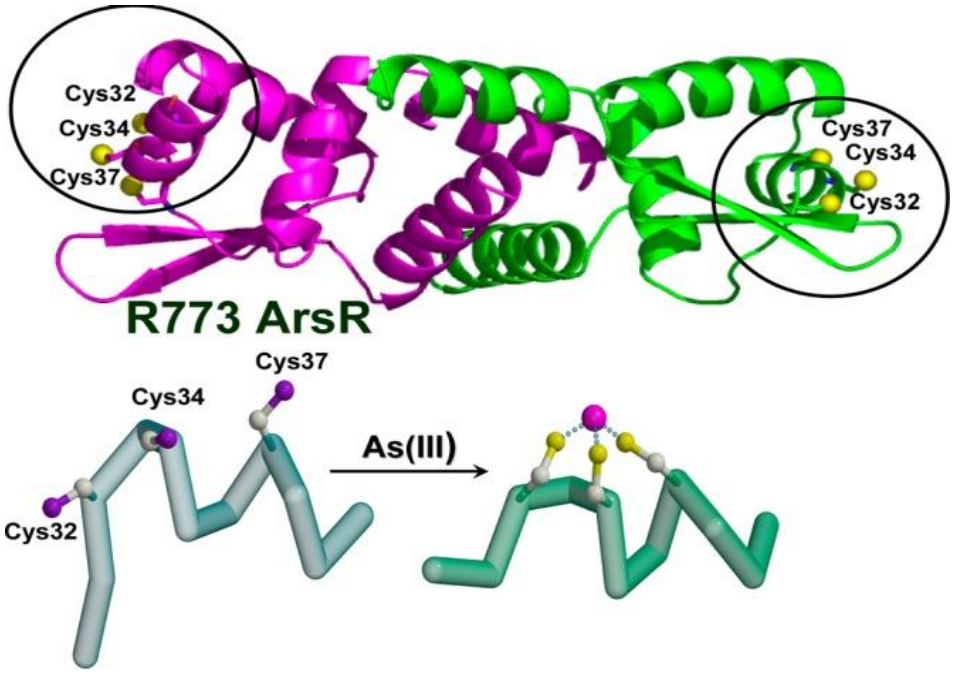


Figure 2.1: A conformational change as a result of iAsIII binding to the E. coli ArsR repressor (P15905).[9]

2.4 Sequence of ArsR-derived peptides

The aim was to design peptides with a conformational structure and activity that closely resembles the ArsR metalloprotein when exposed to iAs (III). When designing the peptides, two different strains of the ArsR repressor protein were considered (Figure 2.2). Although the region with the sequence Cys32-Val-Cys34-Asp-Leu-Cys37 is the main metal-binding domain in both strains, other cysteine-containing regions with seven amino acids in between the second and last cysteine residues were adopted and will be tested for the binding of arsenic metabolites. The research presented in this dissertation is based on the three peptides in Figure 2.3 adopted from the highlighted regions on the protein alignment (Figure 2.2).

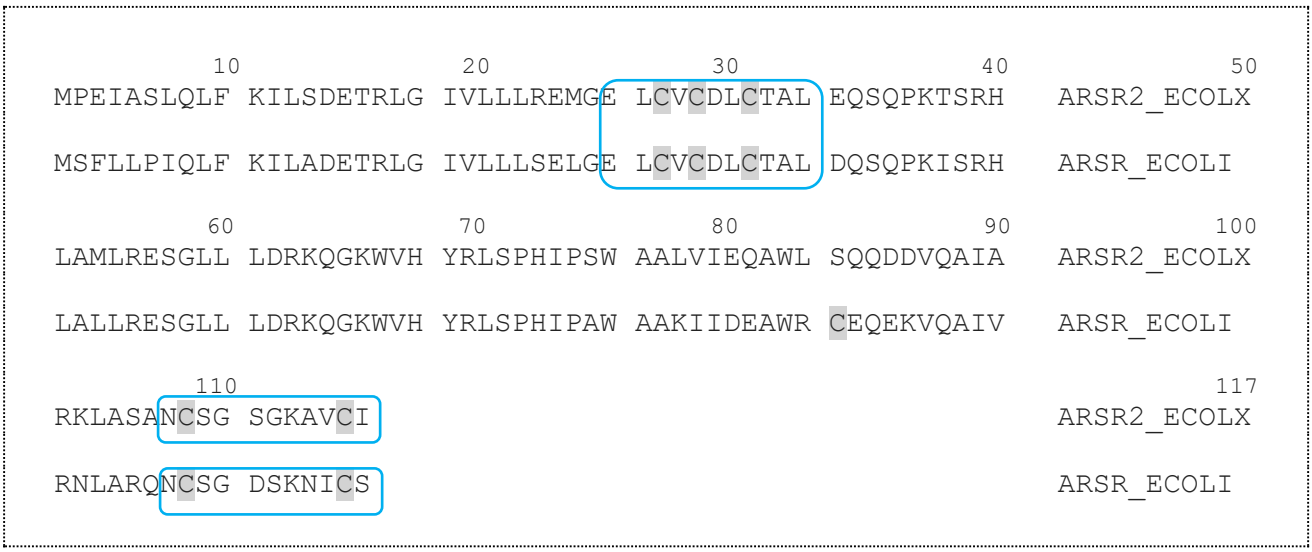


Figure 2.2: Protein alignment of two ArsR proteins ARSR2_ECOLX and ARSR_ECOLI obtained from the Uniprot protein database. ARSR2_ECOLX is the identifier name for ArsR pR773 and ARSR_ECOLI is the identifier name for a strain of pR773 ArsR (strain K12).

PEPTIDE SEQUENCE	PEPTIDE ID
GEL CVC DLC TAL	Pep1-RJM
CVC SGS SKA VCI	Pep2-RJM
CVC SGD SKN ICS	Pep3-RJM

Figure 2.3:Peptide sequence of the ArsR-derived peptides with amino acids derived from the metal-binding domain.

2.5 Basic principles of solid phase peptide synthesis (SPPS)

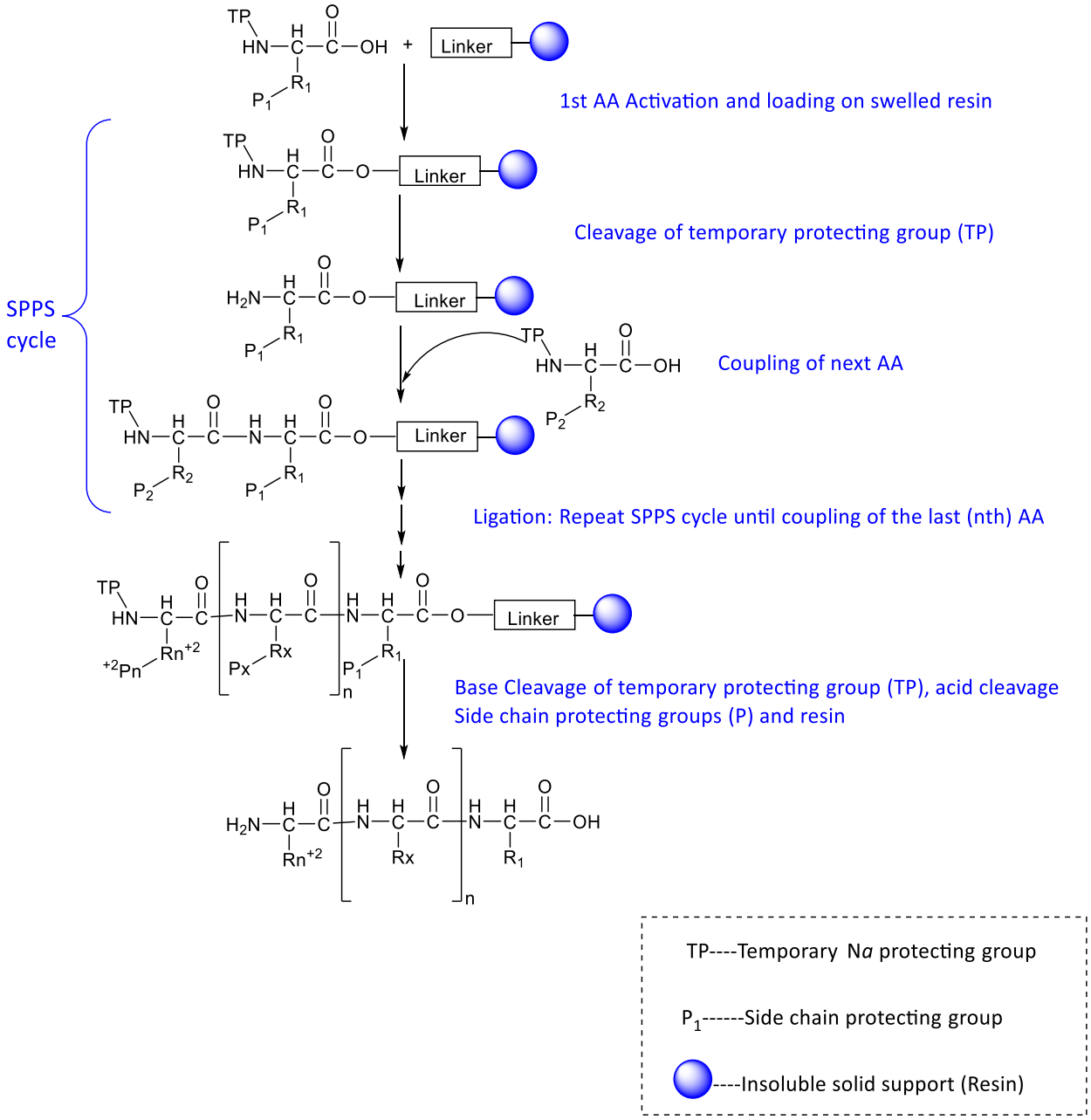
2.5.1 Peptide synthesis

Peptide synthesis is the process of forming an amide bond between two amino acids by coupling the carboxylic acid of one amino acid to the amine group of another. Peptide synthesis can be accomplished by using solution-phase and solid-phase techniques. Peptide synthesis in solution was the initial technique of choice but was associated with several limitations such as tedious and incomplete isolation, purification, and characterization steps required for each coupling step. This approach is not only time-consuming and laborious, but it also results in low yields and low solubility of the intermediate peptides with the increase in the peptide chain length. A new approach was needed to synthesize peptides more efficiently as there was a growing need to synthesize larger and more complex peptides with high yield and high purity.[11]

Finding an alternative approach was of interest for a considerable period, until in 1963, when the ground-breaking work of Nobel prize winner Bruce Merrifield in Solid Phase Peptide Synthesis (SPPS) revolutionized the field and simplified the time-consuming and difficult purification steps associated with the solution-phase synthesis. SPPS technique is based on a peptide chain that is built on an insoluble solid support. One of the most significant advantages associated with SPPS is the ability to easily separate the intermediate peptide from soluble reagents and solvents by washing and filtration, which saves time and makes the labour less tedious than solution synthesis. To help drive the coupling reaction to completion, the reagents can be used in excess or at high concentrations, and the use of insoluble solid support in SPPS can reduce physical product loss during the synthesis. The ability to perform all the synthetic steps in the same reaction vessel without transferring the material is another advantage of using the technique. Merrifield's SPPS paved the way for the widespread use of robotic instrumentation and automation as a result of its invention. Automated peptide synthesis becomes possible after defining a synthesis strategy and programming the amino acid sequence of peptides into a machine. SPPS is now the preferred method for producing peptides, though solution-phase synthesis can still be used for large-scale production.[11]

2.5.2 Basic principle of SPPS

Scheme 2.2 below gives an illustration of the basic principle of SPPS. The synthesis begins with the anchoring of the Carboxyl group of the N-protected amino acid residue to insoluble solid support called a resin. To prevent the functional groups on the amino acid side chains from reacting to form undesired side products, they are protected with acid or base labile groups that remain unaffected by the reaction conditions of the SPPS. The amino- group on the other hand is protected by a temporary protecting group (TP) which is usually a urethane derivative. Under mild conditions, the temporary protecting group (TP) can be easily removed while maintaining peptide integrity. The TP is removed after the first AA has been coupled to the resin to prepare for the coupling of the next N-protected amino acid. An excess of the second amino acid is introduced, and the carboxyl group of the second amino acid is activated to an ester using a coupling reagent for amide bond formation. The desired peptide sequence is synthesized sequentially from the C-terminus to the N-terminus (the C-N strategy) by repeating cycles of N deprotection and amino acid coupling reactions, which are carried out in a single reaction vessel. Deprotection and coupling steps are repeated until the desired sequence is obtained. A variety of conditions are used to finally detach peptides from the resin while simultaneously deprotecting their sidechain protecting groups (P) to produce a C-terminal acid or a C-terminal amide peptide.[12]



Scheme 2.2: Basic principle of SPPS.[11]

2.5.3 Boc/Bzl and Fmoc/tBu strategies

The Boc/Bzl and Fmoc/tBu approaches are the two main SPPS strategies for side-chain protection and temporary N-protection of amino acids. Due to the ease with which it can be cleaved, the Fmoc approach is frequently preferred over the Boc approach. The Boc/Bzl-strategy requires anchoring groups that can tolerate repetitive TFA treatment. For the final cleavage of the peptide-resin bond, a strong acid such as inorganic anhydrous acid such as HF

is typically used. The use of this highly toxic acid places restrictions on the batch size and reactor type (as it requires specialised polytetrafluoroethylene-lined apparatus). The use of harsh acidic conditions can also have a negative impact on the chemical and structural integrity of the peptides, especially those containing fragile sequences.[13]

The Fmoc approach offers mild deprotection conditions and uses an orthogonal protection system. Figure 2.4 below shows the orthogonality of Fmoc/t-Bu protection according to Barany et al. The concept of orthogonality describes how two or more protecting groups can exist in a molecule and how one of them can be removed when the other is still present, under different conditions. In Figure 2.4, a base-labile Fmoc group is used as a temporary protecting group for the amino functional group whereas acid-labile side-chain protecting groups (t-Bu) and acid-labile linkers and resin are used. The ability to remove the TP and P via different mechanisms allows for the use of milder acidic conditions for final deprotection and cleavage of the peptide from the resin, which is one of the method's most significant advantages. All of our peptides will be synthesized using the Fmoc/tBu technique for the reasons and benefits stated above.[14]

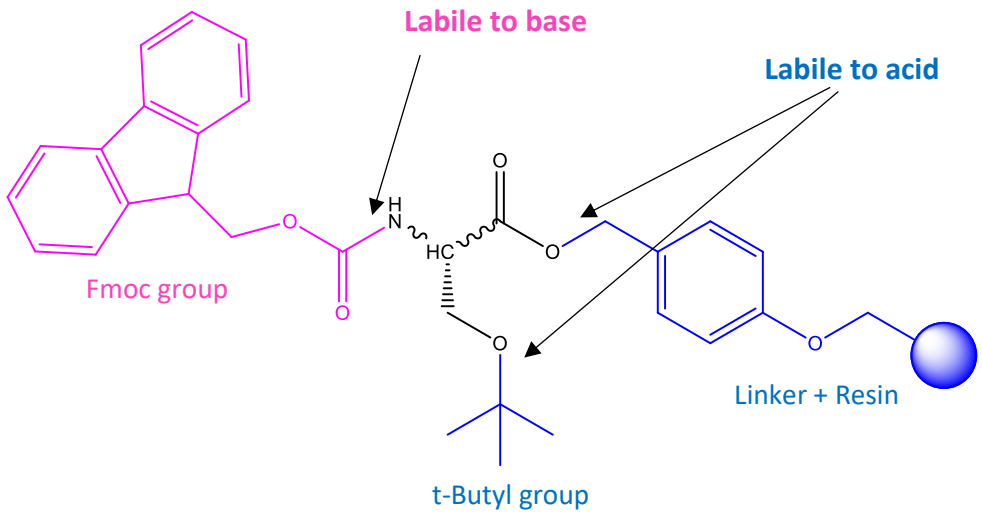


Figure 2.4: The concept of Orthogonality using Fmoc/t-Bu chemistry.[14]

2.5.4 Solid support and linkers in SPPS

The success of SPPS is highly dependent on the type of resin used, the performance and efficiency of the resin amongst other factors. The resin in SPPS is made up of two components, typically referred to as the solid support and the linker. Of these two components, the solid support is regarded as the "soul" of SPPS. The primary criterion for solid support is that it must be chemically stable throughout the synthesis. Additionally, it should facilitate the interaction of the incoming amino acids with the developing peptide that is tethered to the support. Most of these linkers are cleaved in acidic conditions, resulting in a peptide with a changed C-terminus, such as peptide acid, amide, hydrazine, or sulfonamides.

There has been significant progress in the development and application of numerous supports for peptide synthesis. A number of these advances were derived from the well-known crosslinked polystyrene (PS) and divinylbenzene copolymers that were developed by Merrifield, while others were derived from crosslinked polyamides (PAs) and composite PS-polyethylene glycol (PEG-PS)-based materials, or a combination of polystyrene and these two materials.[11]

PS-based resins have excellent solvation characteristics that allow reactants to diffuse into the polymer matrix and also offers great accessibility of the linker sites that are buried within the beads which have made them widely accepted and widely used. In comparison to the standard cross-linked PS resins, the PA- and PEG-based resins are much more hydrophilic and have a lower loading capacity. As a result, they can be used as an alternative to PS resin in the synthesis of long peptide sequences and complicated sequences. Commercially available linkers can be linked to either one of the available matrices (PS, PA, PEG-PS). Common resins for the synthesis of C-terminal peptide acids using the Fmoc/tBu approach are shown/presented in Figure 2.5. The bead (ball) symbol used in Figure 2.5 is generic and does not refer to a particular matrix.[11]

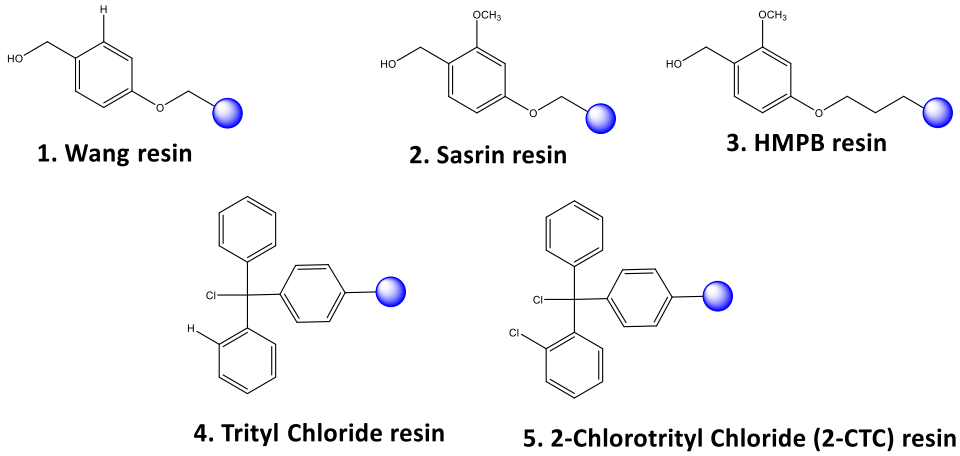


Figure 2.5: Common resins used for Fmoc-based SPPS. Resins (1-5) are used for the synthesis of peptide acids.[11]

.

In this study, all our peptides were synthesized using the 2-chlorotrityl resin (2-CTC) as it is one of the most utilized resins for the SPPS of C-terminal peptide acids and it is usually based on crosslinked polystyrene (Ps) or modified polystyrenes (e. g. tentagel, polystyrene grafted with polyethyleneglycol). The three main advantages of the 2-CTC are that it can be cleaved at mild acidic conditions, it reduces racemization during peptide synthesis of sequences that contains residues in the C terminus such as Cys, His, and Pro, and it also minimizes the formation of diketopiperazine (DKP) during fmoc deprotection.[14]

2.5.5 Amino acid activation and coupling reagents

The activation of the carboxyl residue of the incoming AA is the first and most crucial step in peptide ligation. An excess of activated AA is added to a reaction vessel containing swelled resin, usually by 2-10 times the amount of resin in the vessel. A high concentration of the activated AA is required for a successful coupling reaction. Although methods involving carboxyl group activation are common in organic synthesis, their tendency to cause side reactions makes them unsuitable for the synthesis of peptides. As a result, it has been necessary to use less abrasive activation techniques, such as the formation of active esters, or the activation of carboxyl groups in the peptide in situ. During the coupling step, the amount of time required depends on the activation method used, the nature of the activated species, the sequence of the peptide that is already bound to the resin, and the concentration of the coupling reagent

Coupling procedures that involve activation of the carboxylic acid in situ are the most common. In general, the most widely used coupling reagents are classified into four categories: carbodiimides, uronium/aminium salts, phosphonium salts, and immonium salts, each of which has a structure shown in

Figure 2.*6*. PyBOP is favored for phosphonium-based activation while HBTU is favoured for aminium/uronium-based activation (Figure 2.6). Their mode of action involves transforming the N-protected amino acid into their activated OBt esters. This reaction also involves the use of a tertiary amine usually DIPEA to form the carboxylate of the N-protected amino acid which will further react with the coupling reagent. More recently, it has been reported that the OAt esters generated by HATU and PyAOP are more efficient and that epimerization is reduced as a result of these reactions. Since HBTU and HATU are efficient coupling reagents, they will be used in this study for peptide synthesis.[15]

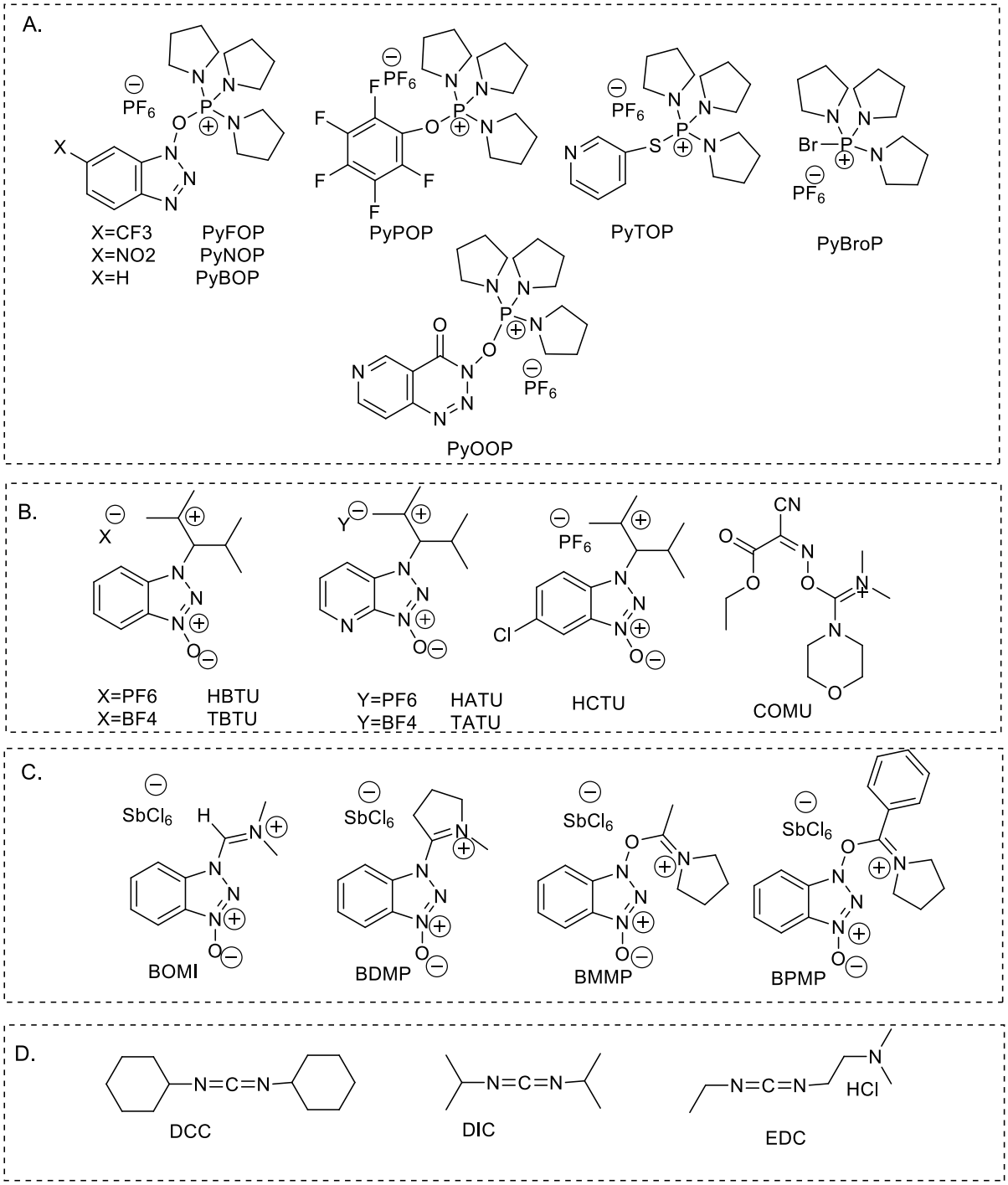


Figure 2.6: Structures of commercially available coupling reagents for SPPS. Where A. Phosphonium salts, B. Uronium/ Aminium salts, C. Immonium salts, and D. Carbidiimides.

2.5.6 Peptide cleavage

After the successful synthesis of a protected peptide, the next step involves the difficult task of simultaneously detaching the peptide from the resin support and the removal of all of the side-chain protecting groups of the amino acid residues to obtain the desired peptide. These

steps are typically carried out in the Fmoc SPPS strategy using TFA treatment of the peptidyl resin as the starting point. This process generates highly reactive cationic species from the protecting groups which can potentially react with and modify residues that contain nucleophilic functional groups such as Trp, Met, Tyr, and Cys if they are not captured. Nucleophilic reagents (also known as scavengers) are added to the TFA to quench these ions and prevent side reactions.[16]

2.6 Methods

2.6.1 Synthesis of Pep1-RJM, Pep2-RJM, and Pep3-RJM

The ArsR-derived peptides (Pep1-RJM, Pep2-RJM, and Pep3-RJM) were synthesised manually, using the basic setup illustrated in Figure 2.7 below. The glass reaction vessel was designed in-house by the glassblower technician and fitted with a porous frit. Nitrogen gas (N_2) was connected to the reaction vessel through plastic tubing, to allow for mixing of reagents during coupling reactions. A vacuum pump was connected to the filtration flask through plastic tubing for the collection of unreacted reagents and solvent waste.

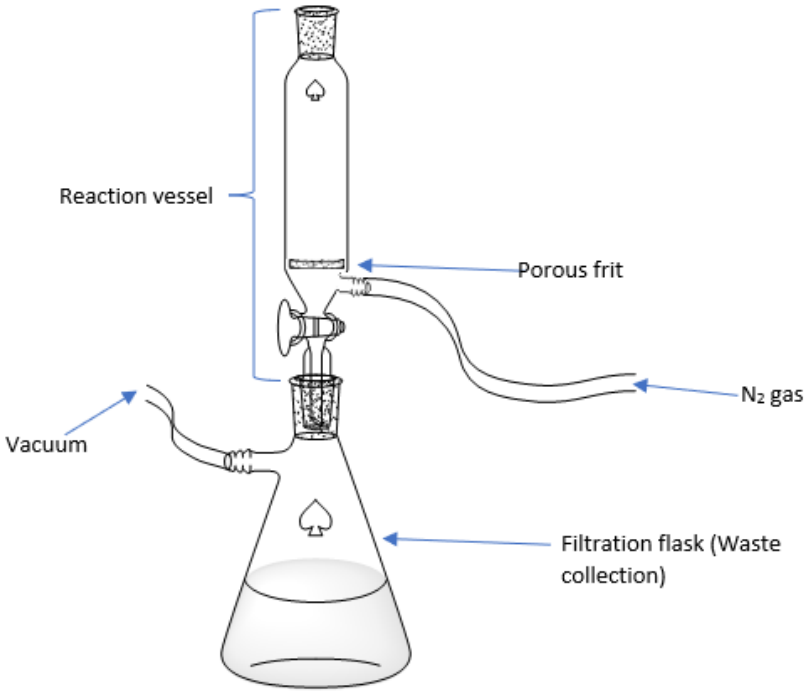
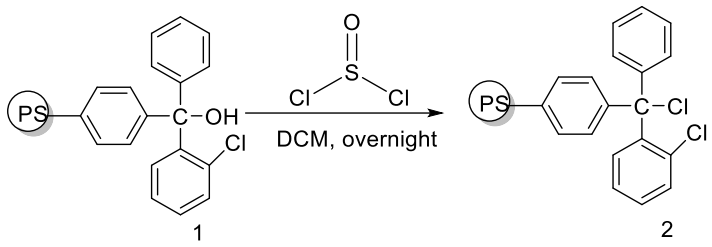


Figure 2.7:Basic equipment for SPPS.

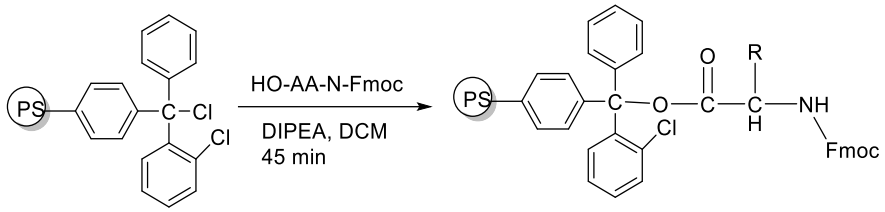
The Fluorenylmethyloxycarbonyl (Fmoc-) solid-phase peptide synthesis (SPPS) strategy was utilized using a 2-CTC resin to afford a carboxylic C-terminal ending.

2.6.2 Swelling and activation of 2-CTC resin



All our peptides were synthesized using the 2-CTC resin. The 2-CTC resin takes up moisture and is hydrolysed to 2-Chlorotrityl alcohol resin (1) during storage. To regenerate it into its activated chloride form (2), the resin was suspended in dry DCM with thionyl chloride ($SOCl_2$). The resin's terminal hydroxyl group was restored to its chloride form (2) making it reusable for further synthesis according to the reaction procedure above. The resin was swelled overnight in DCM to expose the active sites required for peptide synthesis since synthesis occurs inside the beads of the resin. During the first attempt and method development, 0.3g of the 2-CTC resin was used for the synthesis of all three peptides. Repeats of the reactions were performed on 1g of resin.[16]

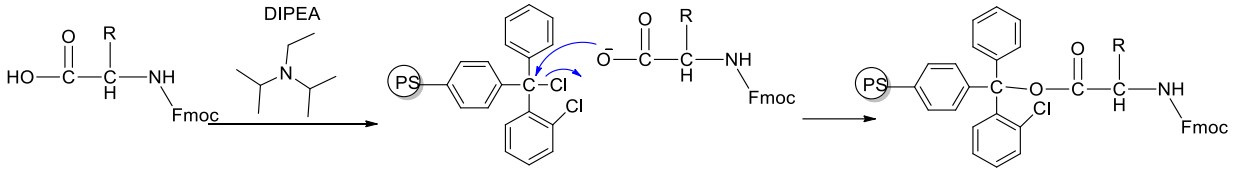
2.6.3 Loading the first amino acid on a 2-CTC resin



The basic procedure for SPPS outlined in Scheme 2.2 was used for the synthesis of the ArsR-derived peptides. The first AA being glycine in the pep1-RJM sequence and Cysteine in pep2-RJM and pep3-RJM was added to a reaction vessel with dry dichloromethane and DIPEA. DIPEA was used to extract a proton from the hydroxyl group to make it a good nucleophile. Loading of the first AA occurs on the chlorine of the 2-CTC resin as a result of the nucleophilic attack of the resin electrophilic site by the nucleophilic oxygen of the Fmoc AA. The reaction

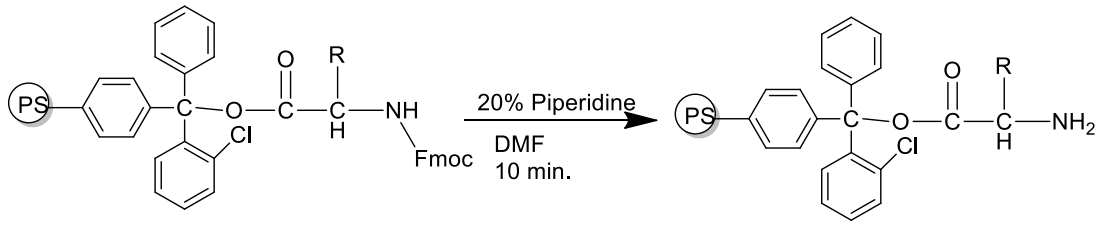
mixture was mixed gently by bubbling nitrogen gas for 45 minutes. To ensure that all active sites are occupied on the resin, double coupling was done for this step. The anchoring reaction was performed in anhydrous dichloromethane and all reagents containing water were dried before use.[17]

Reaction



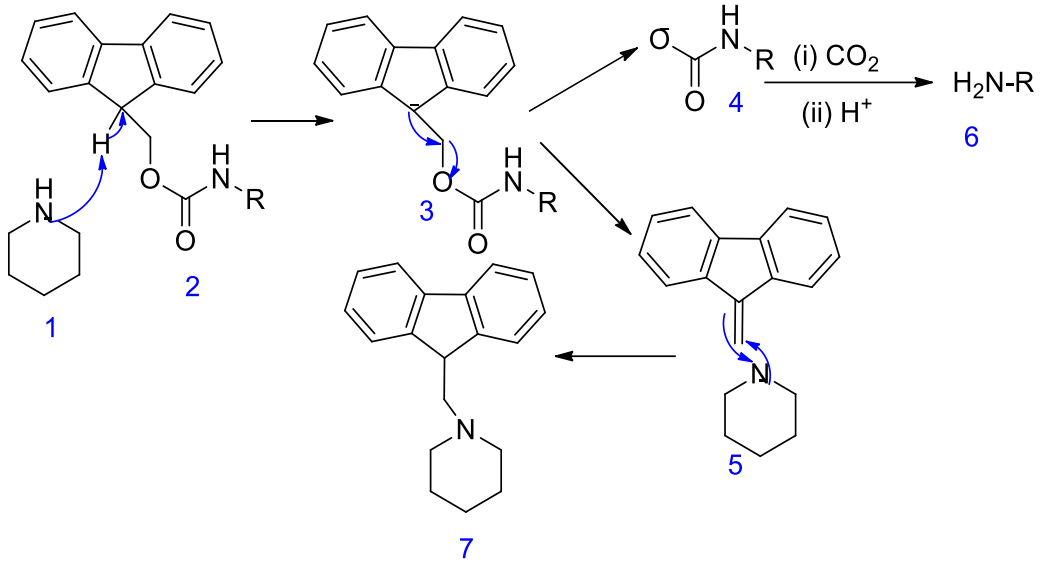
Scheme 2.3: Mechanism for anchoring 1st AA.

2.6.4 Fmoc deprotection



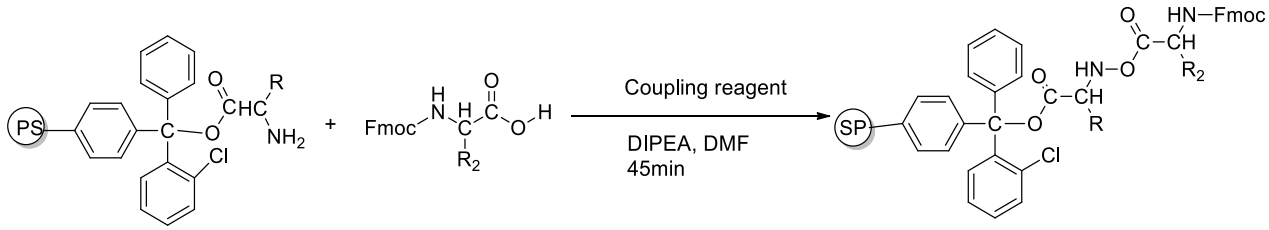
For the step-by-step ligation of the peptides, the temporary protecting group (Fmoc) was cleaved off with a base (20% piperidine in DMF) for 10 min to expose the primary amine which will allow for the coupling of the successive Fmoc AA on the peptide sequence of the three peptides. Piperidine was selected as a good candidate for cleaving Fmoc because it forms a stable adduct (7) with the dibenzofulvene (5) by-product (Scheme 2.4), minimizing any side reactions with the deprotected substrate (6).[18]

Reaction



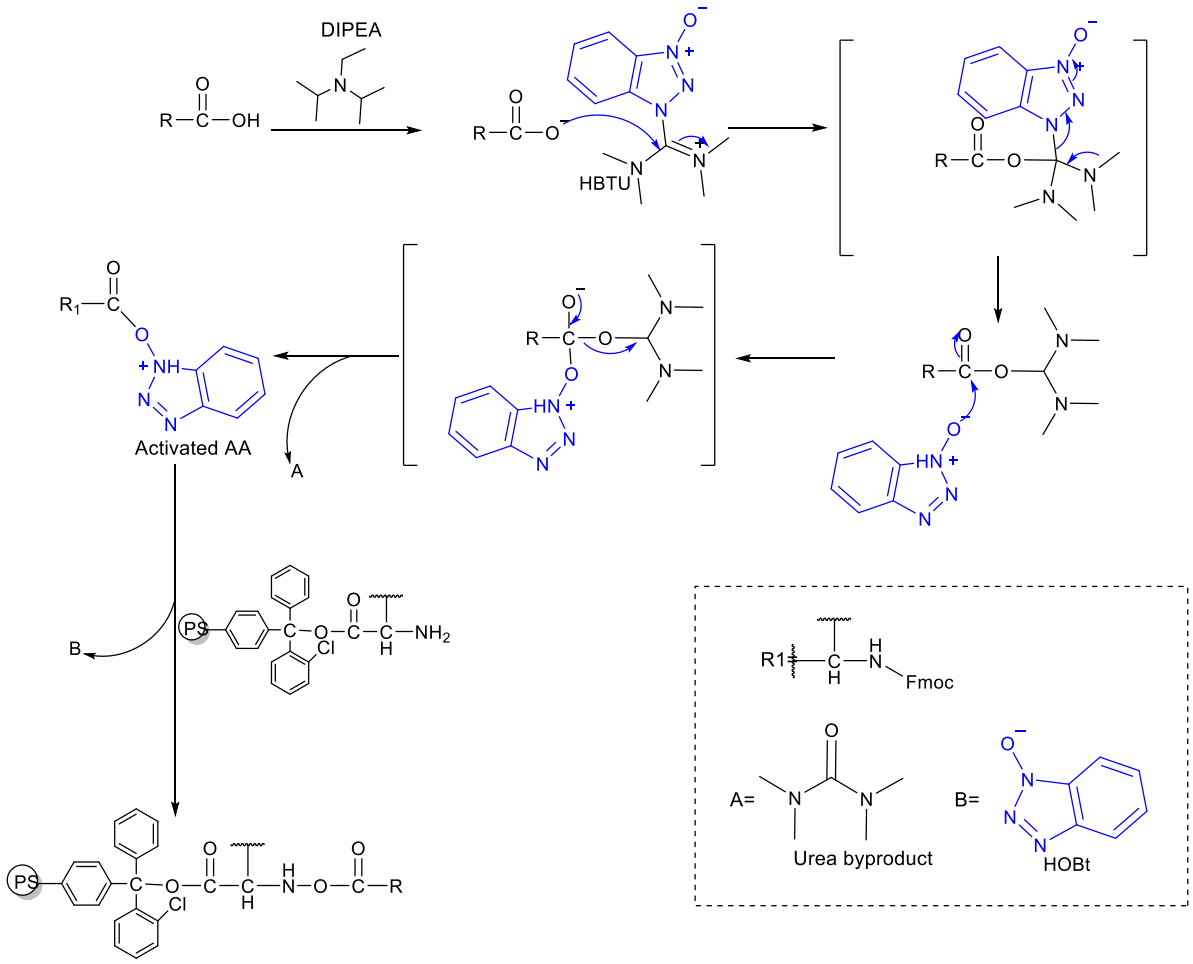
Scheme 2.4: Fmoc deprotection mechanism using 20% piperidine where R represents the amino acid sequence.

2.6.5 Peptide ligation

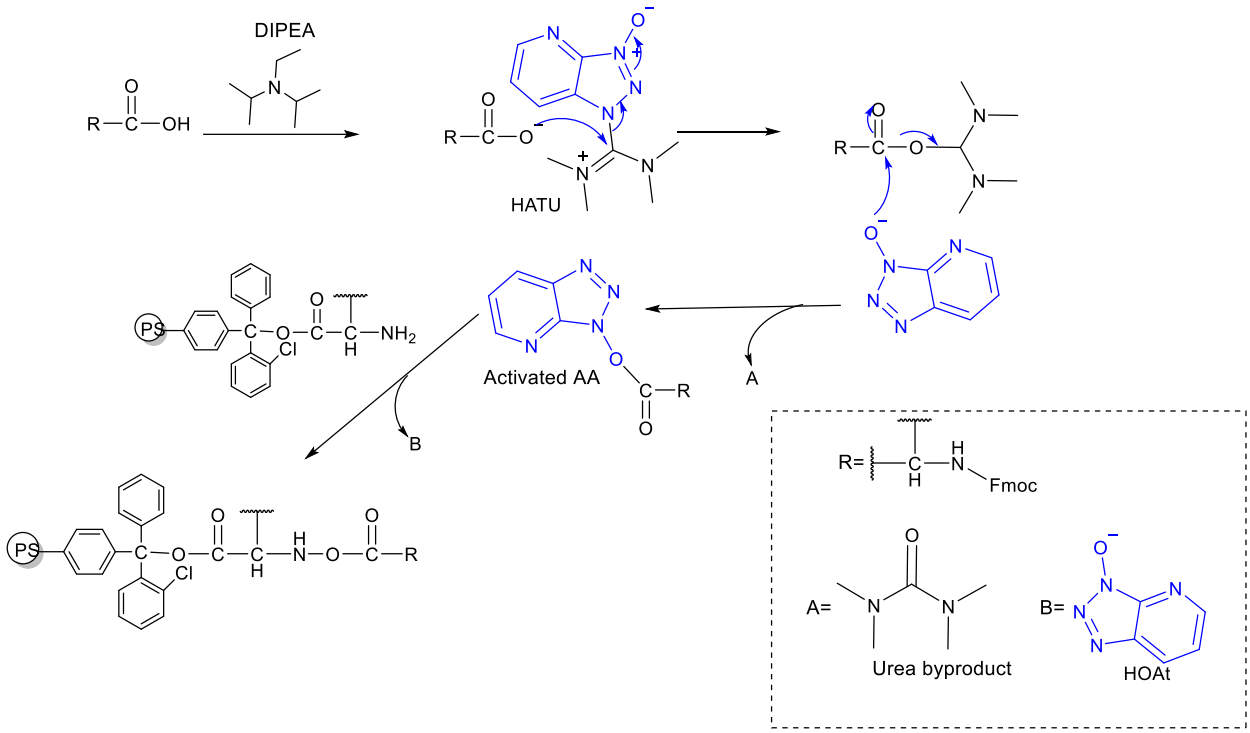


Once the Fmoc group was cleaved, the activation and ligation step occurred on the free primary amine according to the reaction scheme above. The activation step involved the use of coupling reagents to convert the incoming amino acid into an active ester by forming a stabilized HOBt leaving group (Hydroxybenzotriazole) when using HBTU, (see Scheme 2.5 below) or by forming a stabilized HOAt leaving group when using HATU (Scheme 2.6). During the first attempt with 0.3g resin, HBTU was used as a coupling reagent in the presence of DIPEA as a base in DMF. HBTU was initially selected because it is relatively cheap and commercially available. Upon scaling up to 1g resin, HBTU was only used to couple the first 9 AA, then HATU was used to couple the last three amino acids for all three peptides. Once the second amino acid was coupled, the SPPS cycle (Fmoc deprotection, next AA coupling) is

repeated until the last amino acid has been coupled. The last Fmoc protecting group was deprotected before the peptide was cleaved from the resin.



Scheme 2.5: Mechanism for the activation of the Fmoc AA by HBTU, producing an OBt leaving group.



Scheme 2.6: Mechanism for the activation of the Fmoc AA by HATU, producing an OAt leaving group.

2.6.6 Peptide cleavage

Upon complete synthesis of the peptide, it was cleaved from the resin for two hours, using a cleaving cocktail consisting of 95% TFA. The scheme below illustrates the mechanism for cleaving a peptide from 2-CTC resin under highly acidic conditions.



The side-chain protecting groups are also removed under these conditions to produce a fully deprotected peptide. The side-chain deprotection reaction can form cationic species that are highly reactive. To keep these species from reacting and causing undesired changes to the synthesized peptides, several nucleophilic reagents, known as "scavengers," were added to the cleavage mixture. Triisopropylsilane (TIS), water (H_2O), and 1, 2-ethanedithiol (EDT) molecules are the most commonly employed scavengers. The cleavage cocktail composition was as follows (94.5% TFA, 2.5% H_2O, 2.5% EDT and 1% TIS). The cleaved resin was filtered, and the peptide was precipitated using cold diethyl ether. The peptides were then dissolved in appropriate LC-MS compatible solvents based on their polarity.[16]

2.6.7 Peptide purification and characterization

Purification was done on a C18 reverse-phase column using semi-preparative High Performance Liquid Chromatography (prep-HPLC). The run time was selected based on the developed method whereas the solvent system and gradient elution method was selected based on the polarity of the peptide. For reaction monitoring during peptide synthesis and characterization of the final peptide, Liquid Chromatography tandem Mass Spectrometry (LCMS) techniques were used. LC component of the instrument was used to separate the analytes (desired peptide from impurities) in the sample by using a reverse phase column with a solvent system chosen based on the polarity of the peptide. The MS component ionizes the analytes that were separated in the LC and offers sensitive and selective detection of the ionised analytes.

2.7 Results and discussion

2.7.1 Peptide synthesis

The peptides were cleaved from the resin and analysed using liquid chromatography-mass spectrometry (LC-MS), all three peptides were successfully synthesized using 0.3 grams of the resin and HBTU as a coupling reagent. HBTU is one of the most commonly used coupling reagents and was initially selected because it is cheap and works well for most simple sequences. However, a low yield of the final peptide was observed for Pep1-RJM when the reaction was scaled up to 1g of 2-CTC resin and coupling occurred only until the 9^{th} amino acid (GELCVCDLC-Fmoc). Figure 2.8 below shows the mass spectra with the observed m/z of 1277.5006 of incomplete Pep1-RJM corresponding to $[M+H]^+$. The desired mass was also not obtained for Pep2-RJM (Figure 2.9) and Pep3-RJM (Figure 2.10) when using 1g of resin as coupling also only occurred until the 9^{th} amino acid (Pep2-RJM CVCSGSSKA-Fmoc, $[M+H]^+$=1063.4166). and (Pep3-RJM CVCSGDSKN-Fmoc, [M]=1133.4592). The coupling of the subsequent amino acids was unsuccessful.

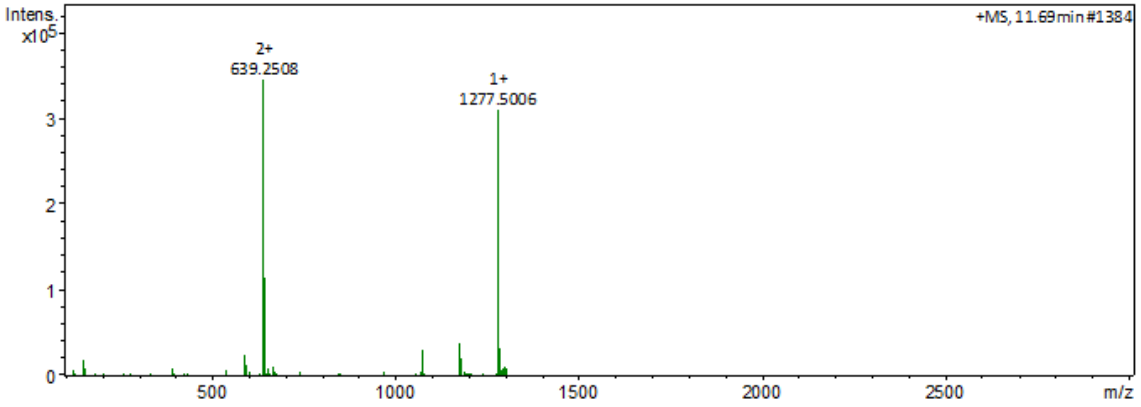


Figure 2.8: Mass spectra of Fmoc protected Pep1-RJM coupling up to 9th AA. $[M+H]^{+}$=1277.5006 and $[M+2H]^{2+}$=639.2508 corresponds to GELCVCDLC-Fmoc sequence.

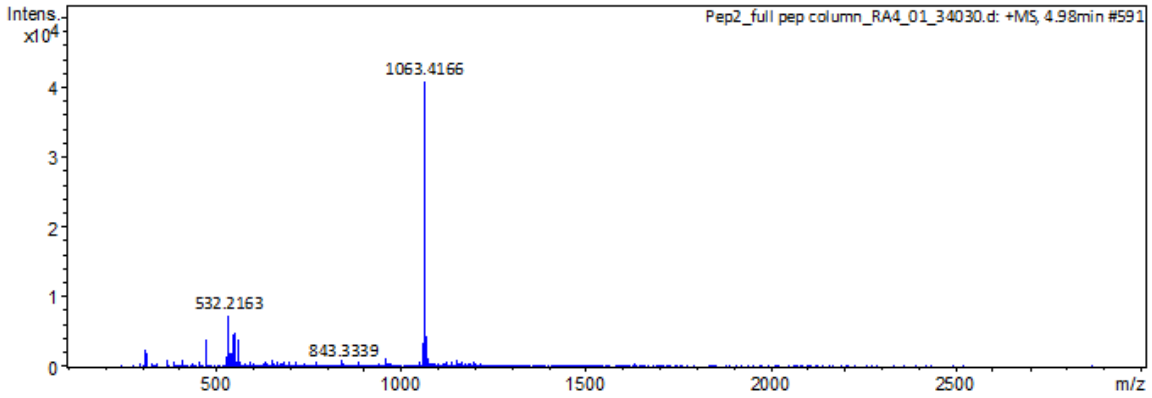


Figure 2.9: Mass spectra of Fmoc protected Pep2-RJM coupling up to 9th AA. $[M+H]^{+}$= 1063.4166 and $[M+2H]^{2+}$=532.2163 corresponds to CVCSGSSKA-Fmoc sequence.

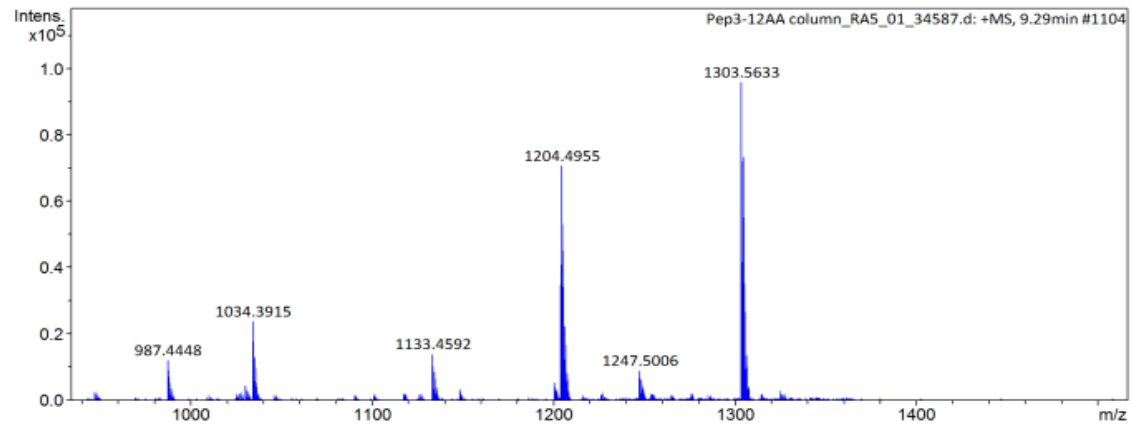


Figure 2.10: Mass spectra of Fmoc protected Pep3-RJM coupling up to 9th AA. [M]=1133.4592 corresponds to the CVCSGDSKN-Fmoc sequence.

The reaction coupling times were varied between 40 minutes and 60 minutes, the reaction was also performed using single and double coupling of the amino acids. However, despite

varying synthesis conditions, the LC-MS showed the molecular ion peak corresponding to 9 amino acids. Another attempt was made to synthesize the peptides where the same method was used but HBTU was replaced with HATU. Table 2.2 below shows a summary of the results that were obtained. The desired mass was observed for all three peptides and the reaction was cleaner with less impurities. Although HBTU worked perfectly for the synthesis of our peptides at a small scale, we were unable to grow our peptide beyond the 9th amino acid at the 1g resin scale. The success of HATU can be attributed to the fact that although HBTU and HATU both convert carboxylic acids into an active ester in the presence of a base, the coupling mechanism of HBTU generates an OBt ester whereas HATU generates an OAt ester (Scheme 2.6 and 2.6). The OAt ester is more reactive than their OBt counterparts owing to the lower pKa of HOAt compared to HOBt. Furthermore, HOAt has the added benefit of the pyridine nitrogen which provides anchiomeric assistance to the coupling reaction, making HATU a more efficient coupling reagent than HBTU.

HATU is expensive despite its exceptional properties. The aim was to develop a low-cost bioreceptor using a simple and inexpensive synthetic method. In our final attempt, we used HBTU to couple the first nine amino acids of each peptide and HATU to couple the last three amino acids. This method was selected as the final synthetic method.

Table 2.2: Summary of synthesized ArsR-derived peptides.

Peptide ID	Peptide sequence	Formula	Theoretical/ expected [M]	Detected $[M+H]^+$	Detected $[M+2H]^{2+}$	R_t/min
Pep1-RJM	GEL CVC DLC TAL	$C_{50}H_{86}N_{12}O_{18}S_3$	1238.53	1239.4910	620.2628	6.40
Pep2-RJM	CVC SGS SKA VCI	$C_{45}H_{81}N_{13}O_{16}S_3$	1155.51	**	578.7493	3.11
Pep3-RJM	CVC SGD SKN ICS	$C_{45}H_{78}N_{14}O_{19}S_3$	1214.47	1215.2454	608.3062	6.73

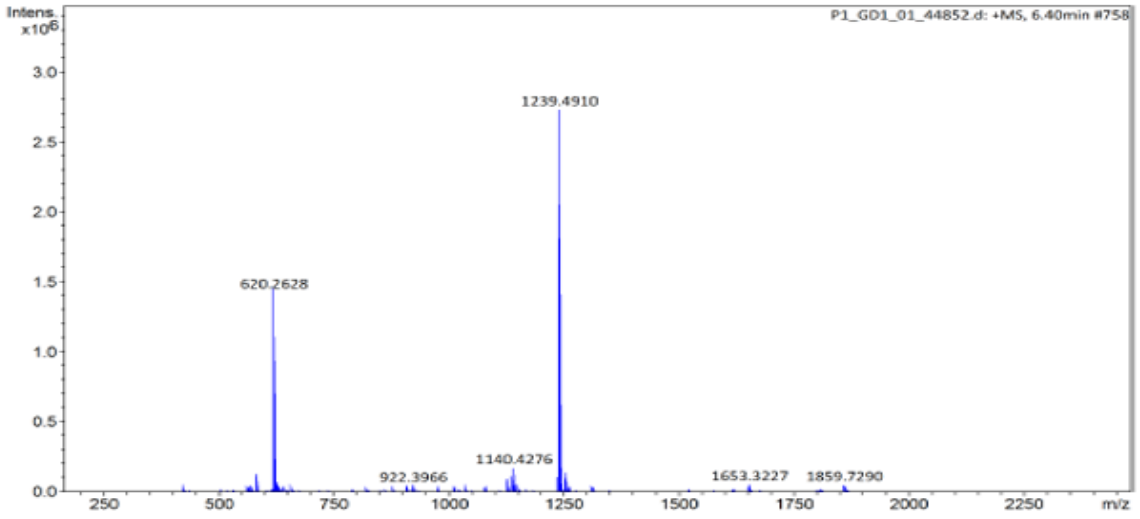


Figure 2.11: Mass spectra of Pep1-RJM with the complete sequence GELCVCDLCTAL-NH₂.

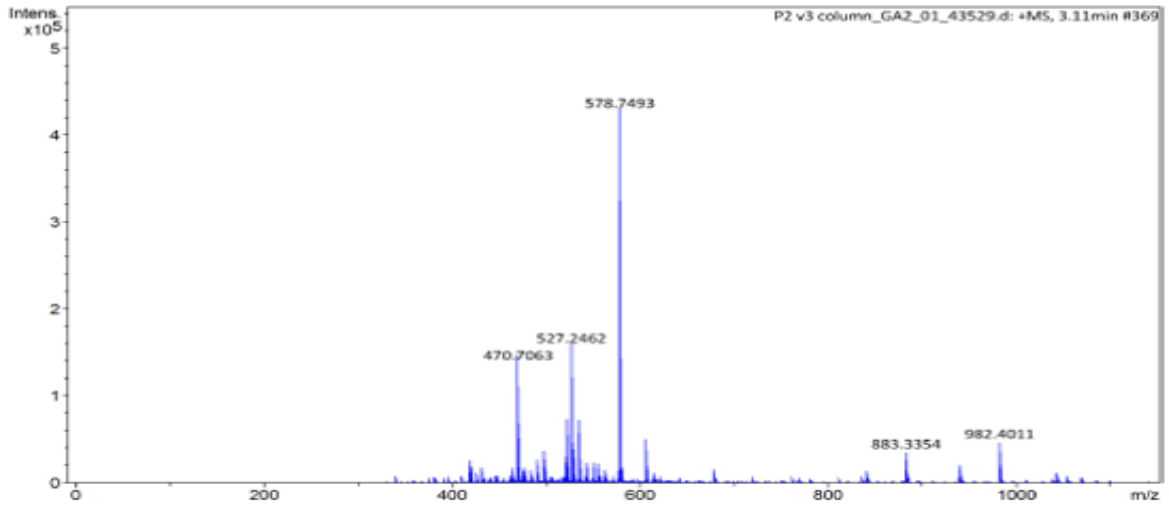
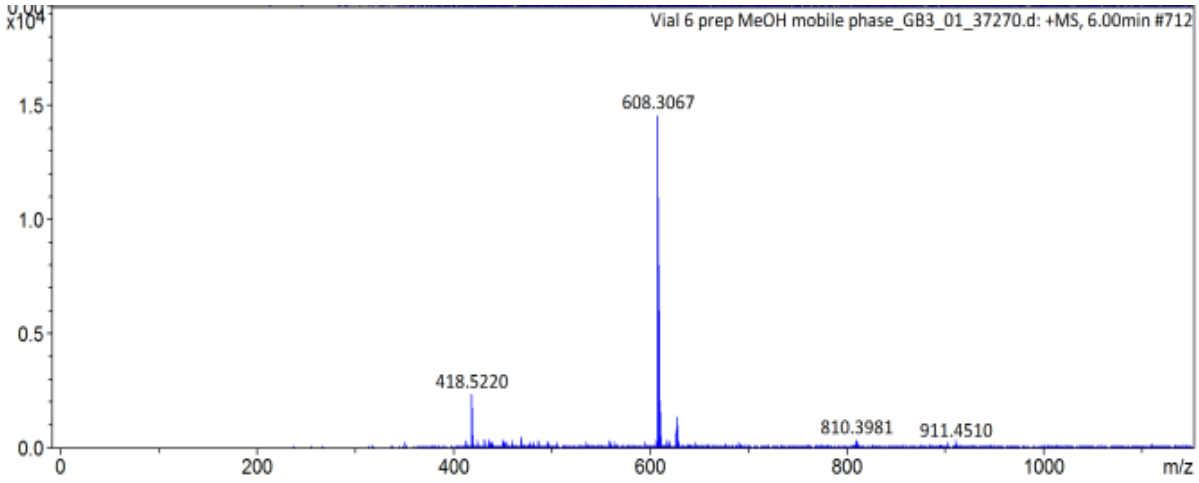


Figure 2.12: Mass spectra of Pep2-RJM with the complete sequence CVCSGSSKAVCI-NH₂.



F*igure 2.13: Mass spectra of Pep3-RJM with complete sequence CVCSGDSKNICS-NH₂.*

2.7.2 UV absorbance of ArsR-derived peptides

The synthesized ArsR-derived peptides were subjected to LC-UV analysis to determine their UV absorption range. The detection of peptides purified by reversed-phase prep-HPLC is typically performed at a UV wavelength of 214-215nm. In this wavelength range, the peptide bond absorbs well, allowing for sensitive detection of a wide range of peptides. Only Pep1-RJM had good UV absorption out of the three peptides (Figure 2.14). The peptides Pep2-RJM and Pep3-RJM did not exhibit any UV absorption when exposed to ultraviolet light.

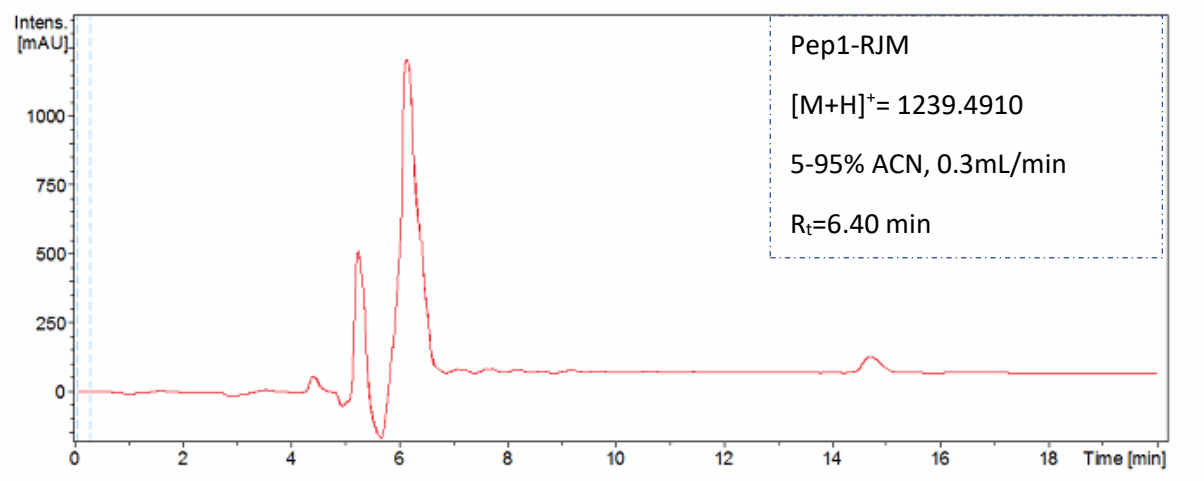


Figure 2.14: Chromatogram showing the UV absorbance of Pep1-RJM at R_t=6.40 min.

The two peptides were then analysed at wavelengths of 190, 200, 205, and 210 nm. In the range of 190-205nm, small peaks began to appear, and Pep2-RJM and Pp3-RJM showed poor but better absorption at 195 nm and 190 nm, respectively, when compared to the other wavelengths (Appendix I SD-8 and SD-9).

2.7.3 Purification and characterization of Pep1-RJM

Pep1-RJM was purified using a binary solvent system consisting of water buffered with 0.1% formic acid and acetonitrile buffered with 0.1% formic acid. A gradient elution method was developed. The acetonitrile ratio was allowed to increase from 5% to 95% in 20 minutes and fractions of the sample that were separated through reverse-phase liquid chromatography

using C_{18} semi-preparatory HPLC were collected in various vials. The peptide was separated from its impurities thus the peptide was successfully purified with 92% purity. After the successful purification, the collected peak was analysed using the LC-MS and it was eluted at a retention time R_t=6.40 min. The expected mass to charge ratio (m/z) was 1238.53 with a chemical formula of $C_{50}H_{86}N_{12}O_{18}S_3$.

2.7.4 Purification and characterization of Pep2-RJM

In our first attempt to purify Pep2-RJM the gradient elution method of 5-95% acetonitrile was used over 20 min at a flow of 20mL/min. Most of the impurities that were observed in the crude product were still present in the fractions. Slight modifications were then made to the prep HPLC method where a flow of 15mL/min and a run time of 30 min was used. Column fractions were collected and analysed using the LC-MS. Peptide fractions containing Pep2-RJM were concentrated using rotary evaporation. The persistent impurity peak with m/z=470.6995 appearing at R_t=5.86 min and it corresponds to the half mass of the sequence (CVCSGSSKAV-NH_2) which was a result of incomplete coupling beyond the 10^{th} amino acid. The major peak appearing at R_t=6.31 with m/z=578.7464 corresponds to the half mass of the desired Pep2-RJM (Figure 2.15).

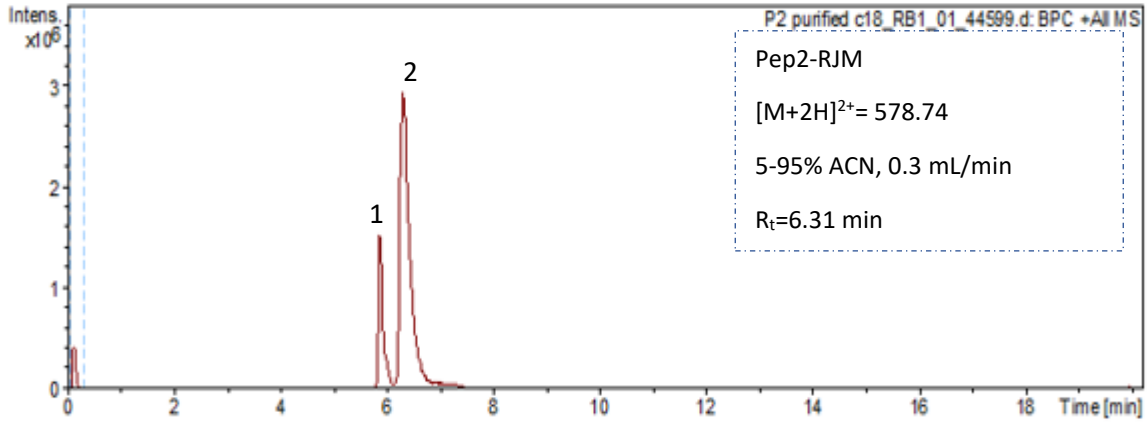


Figure 2.15: HPLC chromatogram of purified Pep2-RJM. Peak 1 at 5.86 is an impurity from the synthesis whereas peak 2 at 6.31 minutes is Pep2-RJM.

After various prep-HPLC method modifications to try and improve the peptide purity. The acetonitrile ratio was ramped from 5% to 40 % in 5 minutes and 40% to 80% in 25 minutes at a flow rate of 15 mL/min in the final purification method. The collected fractions were analysed using the LC-MS and did not contain the problematic impurity peak, however, the yield was very small at the end of all the purification cycles. The mass of pep2-RJM appeared in three different peaks with R_t= 5.23, 5.68, and 6.34 minutes. The difference in retention times could be a result of conformational isomers.

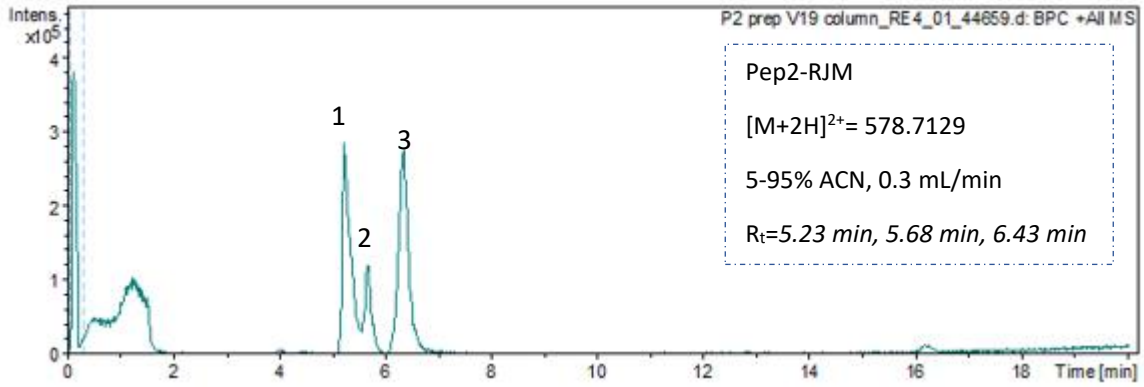


Figure 2.16: HPLC chromatogram of purified Pep2-RJM. Peak 1-3 appearing at 5.23 min, 5.68 min and 6.43 min correspond to [M+2H]=578.7129.

Since the purification of Pep2-RJM was troublesome and involved many purification cycles, the purification was not reproducible and all our other attempts to purify the peptide were unsuccessful and the sample contained impurities that closely resembled the peptide either due to deletion sequences or incomplete coupling. As a result, Pep2-RJM will not be used in the qualitative studies in Chapter 3.

2.7.5 Purification and characterization of Pep3-RJM

Several modifications were applied to the purification method that was used to purify Pep2-RJM in an attempt to purify Pep3-RJM. However, none of the collected fractions which were analysed through the LC-MS (

Figure 2.*17*) yielded good peptide separation because impurities were still coeluting with with the desired peptide. Although the mobile phase system was changed from acetonitrile to

methanol as well as the prep-HPLC column, there was no significant difference in the obtained results. It was discovered that reducing the mobile phase flow rate to 10 mL/min and using the gradient elution method of 3-65% acetonitrile in 30 minutes resulted in slightly better separation because some impurities were eliminated, but the purity remained low. Hence, Pep3-RJM will not be used in a qualitative study in Chapter 3

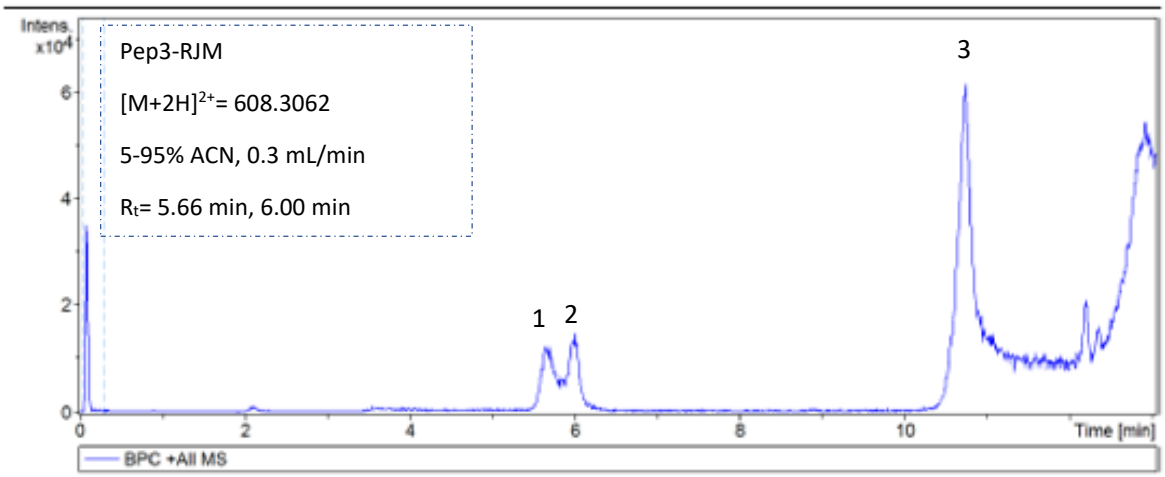


Figure 2.17: HPLC chromatogram of purified Pep3-RJM. Peak 1-2 at 5.66 min and 6.00 min correspond to Pep3-RJM, whereas peak 3 was an impurity from synthesis.

2.8 Concluding remarks

The novel ArsR-derived peptides (Pep1-RJM, Pep2-RJM, and Pep3-RJM) were successfully designed, synthesized and characterized. The developed synthetic method for manual SPPS was successfully optimized by replacing HBTU with HATU when coupling the last three amino acids for large-scale synthesis. The purification of peptides is highly dependent on their UV absorbance. Pep1-RJM was successfully purified.

References

(1) Burtis, C. A.; Bruns, D. E. Tietz Fundamentals of Clinical Chemistry And Molecular Diagnostics. *J. Chem. Inf. Model.* **2015**, 1075.

(2) Lewandowski, C. M. Peptidesynthesis Review. *Eff. Br. mindfulness Interv. acute pain Exp. An Exam. Individ. Differ.* **2015**, *1*.

(3) Ullman, C. G.; Frigotto, L.; Cooley, R. N. In Vitro Methods for Peptide Display and Their Applications. *Brief. Funct. Genomics* **2011**, *10* (3), 125–134. https://doi.org/10.1093/bfgp/elr010.

(4) Xiao, X.; Kuang, Z.; Slocik, J. M.; Tadepalli, S.; Brothers, M.; Kim, S.; Mirau, P. A.; Butkus, C.; Farmer, B. L.; Singamaneni, S.; Hall, C. K.; Naik, R. R. Advancing Peptide-Based Biorecognition Elements for Biosensors Using in-Silico Evolution. *ACS Sensors* **2018**, *3* (5), 1024–1031. https://doi.org/10.1021/acssensors.8b00159.

(5) Negami, T.; Shimizu, K.; Terada, T. Coarse-Grained Molecular Dynamics Simulation of Protein Conformational Change Coupled to Ligand Binding. *Chem. Phys. Lett.* **2020**, *742* (January), 137144. https://doi.org/10.1016/j.cplett.2020.137144.

(6) Xu, C.; Zhou, T.; Kuroda, M.; Rosen, B. P. Metalloid Resistance Mechanisms in Prokaryotes. *J. Biochem.* **1998**, *123* (1), 16–23. https://doi.org/10.1093/oxfordjournals.jbchem.a021904.

(7) Neil Burford, Yuen-ying Carpenter, E. C. and C. D. L. S. The Chemistry of Arsenic, Antimony and Bismuth. In *Biological Chemistry of Arsenic, Antimony and Bismuth*; 2011; p 2547.

(8) Branco, R.; Chung, A. P.; Morais, P. V. Sequencing and Expression of Two Arsenic Resistance Operons with Different Functions in the Highly Arsenic-Resistant Strain Ochrobactrum Tritici SCII24T. *BMC Microbiol.* **2008**, *8*, 1–12. https://doi.org/10.1186/1471-2180-8-95.

(9) Shen, S.; Li, X.-F.; Cullen, W. R.; Weinfeld, M.; Le, X. C. Arsenic Binding to Proteins . *Chemical reviews* . American Chemical Society : United States 2013, pp 7769–7792. https://doi.org/10.1021/cr300015c.

(10) Prabaharan, C.; Kandavelu, P.; Packianathan, C.; Rosen, B. P.; Thiyagarajan, S. Structures of Two ArsR As(III)-Responsive Transcriptional Repressors: Implications for the Mechanism of Derepression. *J. Struct. Biol.* **2019**, *207* (2), 209–217. https://doi.org/10.1016/j.jsb.2019.05.009.

(11) Stawikowski M, F. G. Introduction to Peptide Synthesis. *Curr Protoc Protein Sci.* **2012**. https://doi.org/10.1002/0471140864.ps1801s69.

(12) Amblard, M.; Fehrentz, J.-A.; Martinez, J.; Subra, G. Methods and Protocols of Modern Solid Phase Peptide Synthesis . *Molecular biotechnology* . Humana Press : Totowa 2006, pp 239–254. https://doi.org/10.1385/MB:33:3:239.

(13) BARLOS, K.; CHATZI, O.; GATOS, D.; STAVROPOULOS, G. 2-Chlorotrityl Chloride Resin : Studies on Anchoring of Fmoc-Amino Acids and Peptide Cleavage . *International journal of peptide and protein research* . Munksgaard : Copenhagen 1991, pp 513–520. https://doi.org/10.1111/j.1399-3011.1991.tb00769.x.

(14) Góngora-Benítez, M.; Tulla-Puche, J.; Albericio, F. Handles for Fmoc Solid-Phase Synthesis of Protected Peptides. *ACS Comb. Sci.* **2013**, *15* (5), 217–228. https://doi.org/10.1021/co300153c.

(15) Han, S.-Y.; Kim, Y.-A. Recent Development of Peptide Coupling Reagents in Organic Synthesis . *Tetrahedron* . Elsevier Ltd 2004, pp 2447–2467. https://doi.org/10.1016/j.tet.2004.01.020 .

(16) García-Martín, F.; Bayó-Puxan, N.; Cruz, L. J.; Bohling, J. C.; Albericio, F. Chlorotrityl Chloride (CTC) Resin as a Reusable Carboxyl Protecting Group. *QSAR Comb. Sci.* **2007**, *26* (10), 1027–1035. https://doi.org/10.1002/qsar.200720015.

(17) Spare, L. K.; Menti, M.; Harman, D. G.; Aldrich-Wright, J. R.; Gordon, C. P. A Continuous Flow Protocol to Generate, Regenerate, Load, and Recycle Chlorotrityl Functionalised Resins. *React. Chem. Eng.* **2019**, *4* (7), 1309–1317. https://doi.org/10.1039/c8re00318a.

(18) Pedersen, S. L.; Tofteng, A. P.; Malik, L.; Jensen, K. J. Microwave Heating in Solid-Phase Peptide Synthesis. *Chem. Soc. Rev.* **2012**, *41* (5), 1826–1844. https://doi.org/10.1039/c1cs15214a.

Chapter 3 : Qualitative determination of arsenic metabolites bound to novel Ars-R derived peptides in urine matrix by HPLC-ESI-MS

3.1 Introduction

As explained in Chapter 2, peptides are made up of a variety of amino acid building blocks that are linked by a covalent peptide bond. Each of these amino acids has its own chemical characteristics. The binding of biomolecules such as proteins and peptides to small molecules or metal ligands is a common occurrence in cellular processes. The discovery of novel metal ion binding sites in peptide sequences has received attention for two reasons: (a) it provides insight into the molecular basis of metal ion specificity in peptides, and (b) it provides information about the chemistry of the binding analytes for biomedical and technical applications. Researchers have used several established techniques to study the presence of these biomolecule-metal complexes, including gel filtration chromatography, ultracentrifugation, calorimetry, scattering (X-ray diffraction, circular dichroism), fluorescence quenching, ultraviolet (UV) spectroscopy, and nuclear magnetic resonance (NMR). Among the techniques available, each one has its own strengths and weaknesses. Some of the limitations encountered include the need for large amounts of samples, lengthy experiments, lack of specificity of the method for a given complex, and poor mass resolution.[1]

As part of the sensor development process, electrospray ionisation mass spectrometry (ESI-MS) in conjunction with high-performance liquid chromatography afford new opportunities to analyze intact peptides and their complexes. Trivalent arsenicals such as arsenite (AsIII) and its metabolites dimethylarsinous acid (DMAIII) and phenyl arsine oxide (PAO) are not easily detected using ESI-MS because of their poor ionizability, however, they have a high affinity for thiols. The peptide-ligand chemistry in this study was developed based on existing knowledge of the strong arsenic-thiol interaction. The easily ionisable ArsR-derived peptides were used as binding agents for trivalent arsenicals before they were analyzed on the LC-ESI-MS (Figure 3.1).

The analysis of the peptides with individual arsenic species will provide information regarding the distribution, behavior, and favorable binding conditions of each arsenic species. The aim of this chapter is to provide a detailed overview of the methods investigated during sample preparation and arsenic-binding studies. Several reaction conditions such as the solution pH, the reaction temperature, the initial arsenic(III) concentration, reaction time, and initial peptide concentration were also evaluated to determine their effect on the binding efficiency of the peptides.

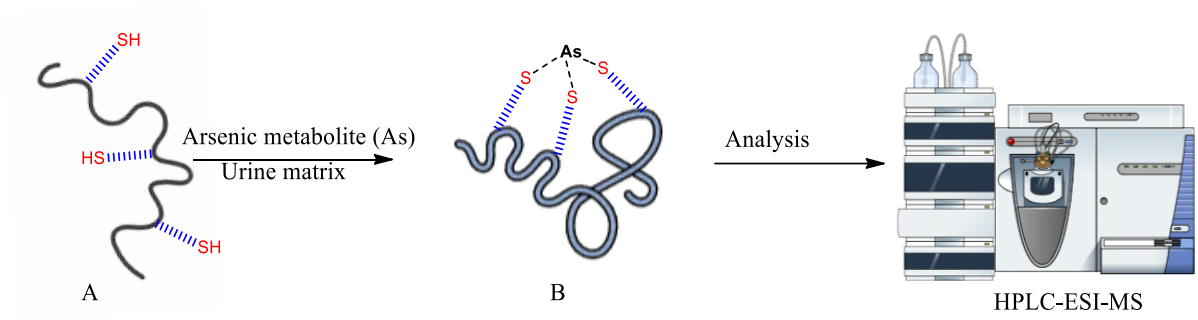


Figure 3.1: The covalent binding of cysteine-containing peptides with urinary arsenic metabolites where A represents the novel ArsR-derived peptides and B represents the covalently bonded Peptide-arsenic complex.

3.2 Arsenic speciation

Figure 3.3.2 below shows the structures of four urinary arsenic metabolites that are detected in urine upon exposure to inorganic arsenic. Among the arsenic metabolites found in urine is inorganic arsenate (iAs(v)), arsenite (iAs(iii)), methylarsonic acid (MMA(v)), and dimethyl arsenic acid (DMA(v)). As discussed in chapter 1, arsenic in urine can either exist as a +3 or +5 oxidation state, and much of the chemistry of these urinary arsenic metabolites depends on the redox conversion between these states. Pentavalent arsenic (As(V)) metabolites undergo a two-electron reduction under acidic conditions, to produce the trivalent arsenic (As(III) which can get oxidized under basic conditions returning to its pentavalent state. The trivalent metabolites (iAs(III), DMA(III), and MMA(III) are stable under reducing conditions, whereas the pentavalent forms (iAs(V), DMA(V), and MMA(V)are stable under oxidizing conditions.

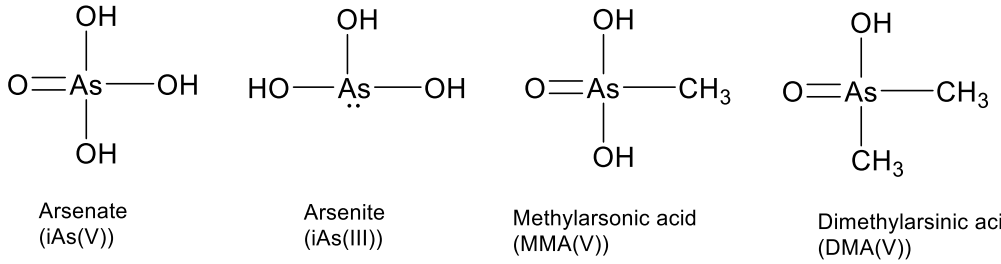


Figure 3.3.2: Structure of arsenic metabolites found in urine.[2]

As(III) exists as a neutral species in the form of arsenous acid (H_3AsO_3) at pH <8. At pH values >8, As(III) occurs as monovalent ($H_2AsO_3^-$), divalent ($H_1AsO_3^{2-}$) and trivalent (AsO_3^{3-}) species. As(III) species are stable under reducing conditions. As(V) mainly occurs as an anionic species over a wide range of pH values. At a neutral pH (pH=7), As(v) can either exist as a monovalent or divalent anion ($H_2AsO_4^-$ and $HAsO_4^{2-}$). MMA(v) exists as a neutral species $(CH_3)As(O)(OH)_2$ at a pH ≤4 and it occurs as a monovalent anion between pH 4-9 and exists as a divalent anion at pH ≥9. DMA exists as a neutral species $(CH_3)_2As(O)(OH)$ at a pH <6 and monovalent anion $(CH_3)_2As(O)(O^-)$ at pH >6. The distribution of organic and inorganic species as a function of pH is presented in Figure 3.3.3 below. As(V) species are stable in oxidizing conditions.[3]

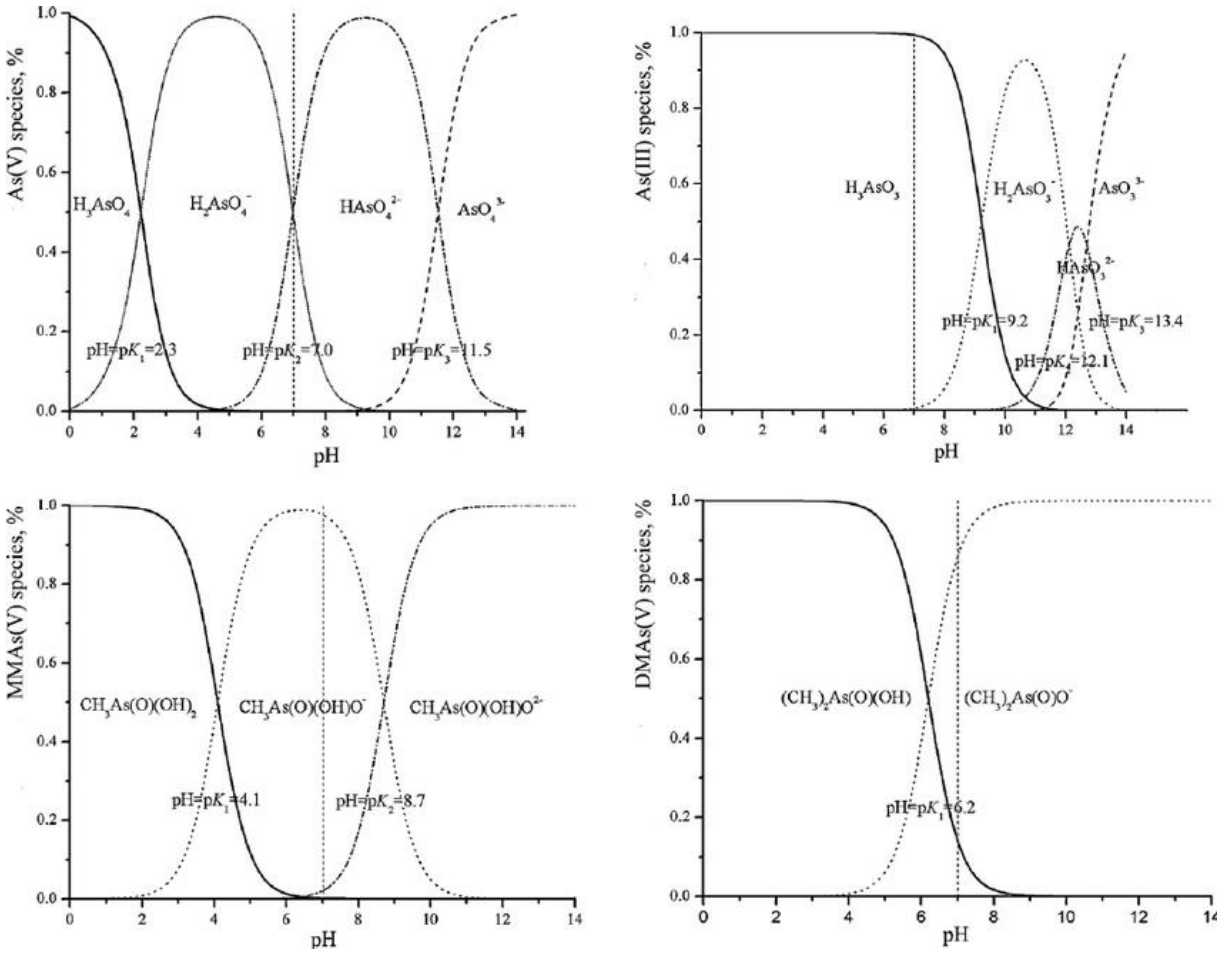
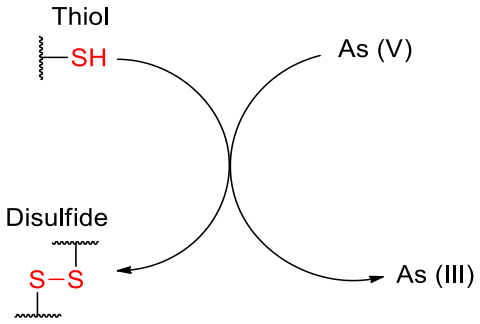


Figure 3.3: The distribution of organic and inorganic arsenic as a function of pH.[3]

Extensive research has been conducted on the use of various materials for arsenic monitoring and remediation through binding or absorption.[4] These materials include proteins, peptides, bacteria, ion-pairing agents, and nanomaterials. Many studies are being carried out to determine the most effective material, not only in terms of binding affinity and efficiency but also in terms of its applicability across a wide pH range, taking into consideration that arsenic can take on molecular or ionic forms depending on the pH.[5] Most of these materials are used for environmental applications and very little is known about their use in biological matrices such as urine. This study aims to address the gap by developing a method to bind and detect arsenic metabolites in urine matrix using three thiol-containing ArsR-derived peptides. While different materials use different mechanisms to bind arsenic metabolites, trivalent arsenic metabolites are more effectively bound and detected than pentavalent arsenicals when using thiol-containing peptides. As a result, the prereduction of As(v) to As(iii) is crucial in improving binding.

3.3 Reduction of As(V) to As(III)

Several methods have been reported for the reduction of pentavalent arsenic to its trivalent state. The potassium iodide solution in acidic medium is one of the most widely used reagents, however, there are several limitations associated with it. According to recent research, the treatment of samples containing oxidizing agents with potassium iodide results in the production of high amounts of I_2. For this reason, a high concentration of potassium iodide is required. The undesirable I_2 could lead to the loss of As(III)-iodine adduct in some applications. Ascorbic acid is typically added to prevent the inevitable I^-oxidation by atmospheric oxygen. As(V) reduction to As(III) is very slow, even in the presence of very high acid concentrations, making this method time-consuming. Thiol-containing agents are more popular nowadays. These agents reduce pentavalent arsenicals to produce trivalent arsenicals and disulfide bonds as shown in Scheme 3.1 below. More than one sulfhydryl group is required as this reduction is a two-electron process. L-cystine (L-cys), Glutathione (GSH), and other similar compounds with the –SH group have been tested and used as the reducing agents.[6]



Scheme 3.1: Reduction of As(V) to As(III) using thiol-containing agents. [6]

3.4 Peptide disulfide bond reduction

Cysteine-containing peptides tend to form disulfide bonds when exposed to oxidizing conditions. This reaction occurs when an S^- ion from a single sulfhydryl group acts as a nucleophile, attacking one of the side chains of the second cysteine, which in turn leads to oxidation of the reduced sulfhydryl groups of the cysteine (S-H), thus resulting in the formation of a disulfide bond (S-S). Molecular disulfide bonds can form by intra- or

intermolecular interactions, and the number of disulfide bridges in a molecule can vary depending on the number of cysteine residues present. The primary function of a disulfide bond in peptides and proteins is to provide both stability and rigidity to the tertiary structure.[7] Figure 3.4 below shows the conformational folding of a peptide, positioning the cysteine residues to be in special proximity to form a disulfide bond (B). The presence of these disulfide bridges can block the active site of the peptide thus hindering its activity and function. As a result, efficient reduction of disulfide bridges is a requirement for many applications of Cysteine-containing molecules in the fields of chemistry and biochemistry. Among the thiol-based reducing reagents, reduction using DTT is reported to be faster. In addition, DTT is widely accepted as a reducing agent in the field of chemistry and biochemistry, mainly due to its ease of synthesis.

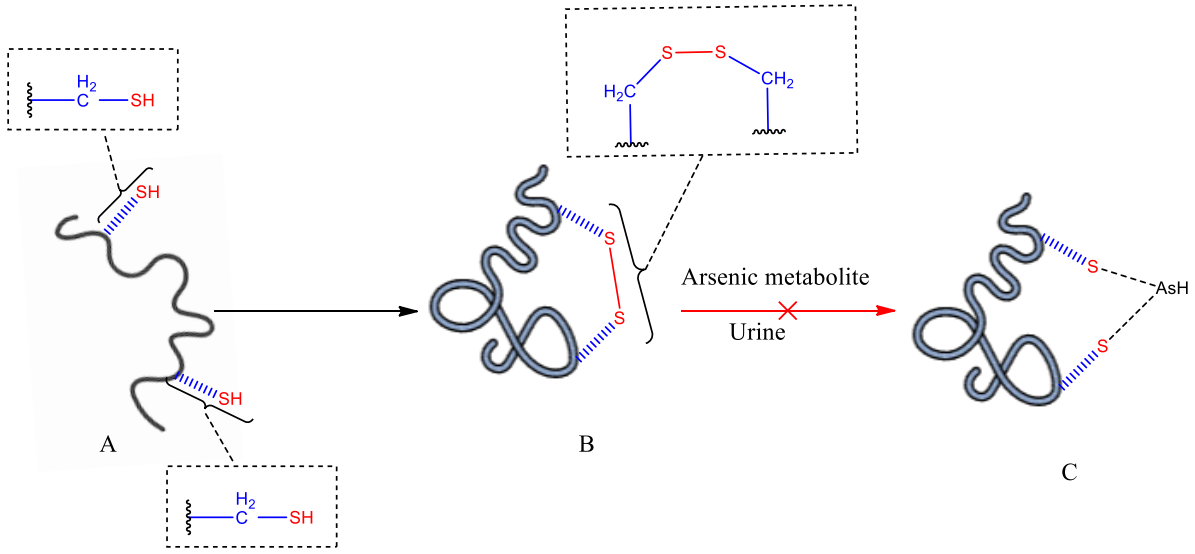


Figure 3.4: conformational folding as a result of disulfide bond formation. Diagram A represents a peptide with reduced thiol groups on the cysteine side chains. In structure B, the peptide folds to bring the two cysteine residues closer for the formation of a disulfide bond.

During the last few decades, thiol-based and phosphine-based reducing agents have been used in molecular biology to reduce disulfide bonds in proteins and peptides. Figure 3.5 below shows a list of the commonly used reducing agents. Thiol-based reducing agents are usually preferred over phosphine-based reducing agents because they are less expensive and allow for easy and specific reduction of disulfide bonds in peptides.[8]

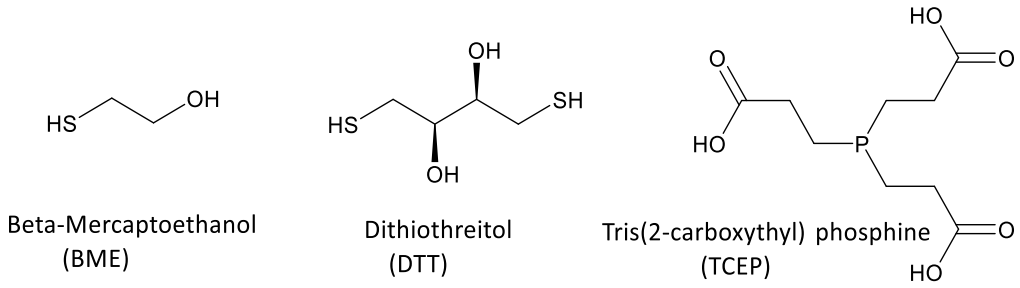


Figure 3.5: Structure of commonly used reducing reagents. BME and DTT are thiol-based reducing reagents and TCEP is a phosphine-cased reducing reagent.[8]

3.5 Materials and methods

3.5.1 Reagents

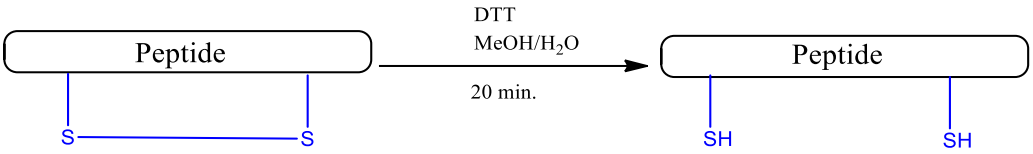
Peptides derived from the ArsR protein (Pep1-RJM, Pep2-RJM, Pep3, RJM) were synthesized in chapter 2 using SPPS. The Sigmatrix urine diluent, Stannous Chloride ($SnCl_2$), sodium (meta)arsenite ($NaAsO_2$) standard, Sodium arsenate dibasic heptahydrate ($Na_2HAsO_4.7H_2O$) standard, and cacodylic acid ($(CH_3)_2AsO_2H$) standard were supplied by Sigma Aldrich. Potassium iodide (KI), (Hydrochloric acid (HCl), and Sodium hydroxide were purchased from Merck.

3.5.2 Sample preparation

Method development for binding studies was conducted in deionised water and optimized using sigmatrix (synthetic urine). Pep1-RJM was dissolved in 100% MeOH to make a stock solution of 807.8 µM in a 15 mL volumetric flask. Pep2-RJM and Pep3-RJM were dissolved in ACN and MeOH (1:1 (v/v)) to make up a stock solution of 865.8µM and 823.7 µM respectively in 15 mL volumetric flasks. The peptide stock solutions were aliquoted into 2 mL portions to avoid contamination and constant freezing and thawing between reaction days. The peptide aliquots were stored at -20 °C. The study only included iAs(III), iAs(V), and DMA(V) urinary arsenic metabolites because monomethyl arsenic acid (MMA) was not commercially available and phenyl arsine oxide was used as an external standard.

3.5.3 Reduction of peptides

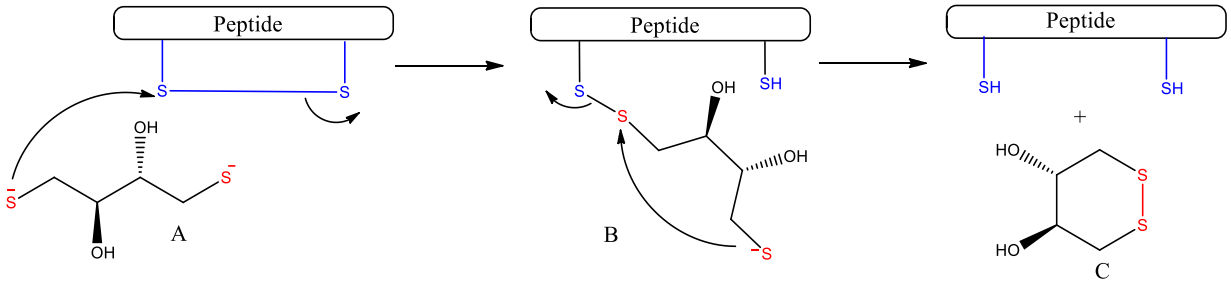
All three peptides were reduced using fivefold molar excess of DTT in deionised water and methanol (8:2 (v/v)) and the reaction was incubated for 20 min before the binding studies.



Scheme 3.2: Dithiothreitol induced disulfide bond reduction in Ars-derived peptides.

During the reduction reaction, a disulfide exchange mechanism takes place, and this process is pH-dependent as it only occurs under basic conditions. This is because the thiol group of the reducing agent must first be deprotonated to create the negatively charged thiolate (A) (Scheme 3.3), which is a better nucleophile. The thiolate then attacks one sulfhydryl group from the peptide forming an intermediate (B) with one reduced S-H group. The reduced S-H group forms a new disulfide bond (S-S) with the reducing agent. A nucleophilic attack by the second thiolate anion results in the formation of a stable six-membered ring with an internal disulfide bond and reduced peptide (Scheme 3.2).

Reaction



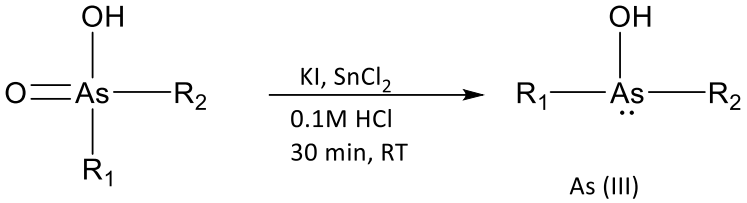
Scheme 3.3: Mechanistic representation of a disulfide exchange reaction using DTT.[9]

3.5.4 Reduction As(V) to As(III)

Three different methods were tested for the prereduction of As(V) to As(III) before proceeding with the binding studies. The first method involved reducing arsenic with potassium iodide under acidic conditions. Methods 2 and 3 were carried out using thiol-based

reducing agents, in which ArsR-derived peptides were tested as possible thiol-containing reducing agents in method 2, and smaller thiol-containing compounds like EDT, DTT, and L-cysteine were tested as potential reducing agents in method 3.

3.5.4.1 Method 1: Reduction by KI [10]



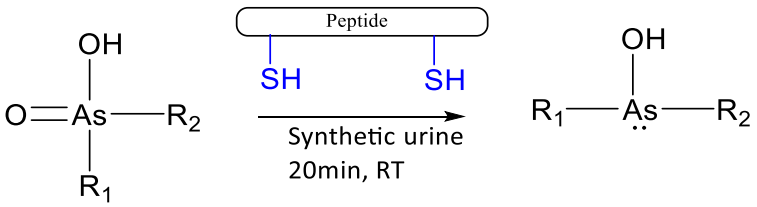
As (V)

As (III)

R_1= OH, CH_3, Ph
R_2=OH, CH_3

As(V) metabolites were treated with potassium iodide (KI) and stannous chloride under acidic conditions. 100 µM of pentavalent arsenic metabolites (iAs(V), DMA(V)) were added into a clean amber bottle from freshly prepared stock solutions. 5 mL of 0.1 M HCl and 2mL of KI were also added into the bottle and the reaction was stirred at room temperature. 0.5 mL of $SnCl_2$ was added with thorough mixing of the solution and left in the dark for 15, 20, 30, 40, and 60min for the reduction of arsenic pentavalent to the trivalent state.

3.5.4.2 Method 2: Reduction by thiol-containing peptides



As (V)

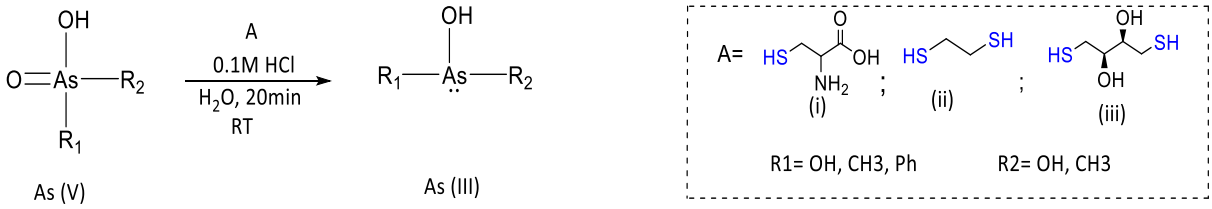
As (III)

R1= OH, CH3, Ph
R2=OH, CH3

In the presence of high concentrations of reduced thiol, there is a possibility that some of the pentavalent arsenicals could be reduced to trivalent arsenicals. Since all our designed peptides have three cysteine residues, this experiment involved the use of ArsR-derived peptides (Pep1-RJM, Pep2-RJM, and Pep3-RJM) as thiol-based arsenic reducing agents in synthetic urine. A series of experiments were conducted in round-bottomed flasks on a reactor where the following As(v): Peptide ratios (1:1, 1:2, 1:5) were used.

3.5.4.3 Method 3: Reduction by Cysteine, EDT, and DTT [11]

Three thiol-based reagents (L-cysteine (i), EDT (ii), and DTT (iii)) were screened for the ability to reduce As(V) to As(III).



Scheme 3.4: Prereduction of As(V) to As(III) where reagent A=L-cystine, EDT, or DTT.

(i) L-cysteine

This method was adopted from literature, Feng, Y. lai et al (1994), and a few modifications were made. 125mg of L-cysteine (i) was dissolved in 15 mL of deionised water and added into a 25 mL volumetric flask. 2.5 mL of 0.1M HCl was added to the solution followed by 2.5 mL of H_2O. An arsenic (V) solution (14.49 mM) was added to the volumetric flask and the solution was diluted to the mark. The solution was left to stand for 20 min before the standards were added to the peptides in the binding reaction setup described below.

Gregus, Z. et al (2009), proposed the equation below for the reduction of arsenic(V) to arsenic(III) using L-cysteine.[12]

$$R\text{-}SH + As(V) \longrightarrow RS\text{-}As^{V} + H_2O \quad (1)$$

$$RS\text{-}As^{V} + RSH \longrightarrow R\text{-}S\text{-}S\text{-}R + As^{III} \quad (2)$$

The method was repeated with slight modifications, L-cysteine was replaced with ethane-1,2-dithiol (ii) in the second experiment and dithiothreitol (iii) in the third experiment.

(ii) EDT

EDT (2 mL) was added to 13 mL of Millipore water in a 25 mL volumetric flask. 2.5 mL of 0.1M HCl was added followed by 2.5mL of water. DMA(V) solution (14.49 mM) was added to the

volumetric flask and the solution was diluted to the mark. The solution was shaken vigorously and left on a shaker for 20 minutes before the peptide-binding studies.

(iii) DTT

DTT (125 mg) was dissolved in 15 mL of Millipore water and added to a 25 mL volumetric flask. 2.5 mL of 0.1M HCl was added followed by 2.5mL of water. DMA(V) solution (14.49 mM) was added to the volumetric flask and the solution was diluted to the mark. The solution was left to stand for 20 minutes before the peptide-binding studies.

3.5.5 Binding reaction setup

All binding reactions were conducted using a six-vessel reactor in an inert atmosphere as shown in Figure 3.6 below. The reactor is equipped with reaction vessels and a stirring hotplate which allows for temperature control and stirring to ensure adequate mixing of reactants during a reaction. The setup not only offers a closed system but also allows for multiple reactions to run simultaneously under the same conditions without sample or solvent loss.



Figure 3.6: 6-vessel reactor where A. rubber seal, B. PTFE cap, C. round bottom flask, and D. Stirring hot plate.

The arsenic(III) binding was studied using a batch of experiments which were carried out by spiking the peptide samples with a known amount of As(III) in synthetic urine matrix using the reactor shown above. The reactor consists of rubber seals which are used to avoid a change in concentration due to evaporation. Each round bottom flask was equipped with magnetic stirrers and the solution was mixed using a stirring plate on the reactor for a predetermined time. Once the reaction was complete, 1.5 ml of the solution was transferred into HPLC vials for analysis. Due to the complexity and variability of urine, it is highly important to investigate the binding ability of the peptides under various conditions. These experiments were conducted considering five variables including the pH, reaction temperature, initial arsenic(III) concentration, reaction time, and initial peptide concentration. Based on the results of these experiments, the optimal experimental conditions were determined and used in the quantitative analysis in Chapter 4.

3.5.6 HPLC-ESI-MS

The UHPLC-ESI-MS was used to monitor the reaction during method development and arsenic binding studies. The UHPLC component (Thermo Scientific Ultimate 3000) was equipped with a binary solvent system where solvent A consisted of H_2O and 0.1% FA (v/v) and solvent B consisted of ACN and 0.1% FA (v/v). The samples were placed on an autosampler at 25 °C and 20 µL of the samples were injected through a reversed-phase C18 analytical column that separated the peptides from the other components in the synthetic urine matrix using a gradient elution method of 5 to 95% ACN in 12 minutes. The separated analytes were detected by a Q-Time-Of-Flight high-resolution mass spectrometer (Bruker Daltonics, Bremen, Germany). The samples were analysed in the positive ionization mode and the MS was scanned at m/z 100-3000 range during separation and detection. Nitrogen gas (N_2) was used as the dry and nebulizer gas. The nebuliser gas was operated at a pressure of 1.8 bars whereas the drying gas was operated at a flow rate of 9 L/min. To guarantee the reliability of the results, the ESI source was cleaned and calibrated every week.

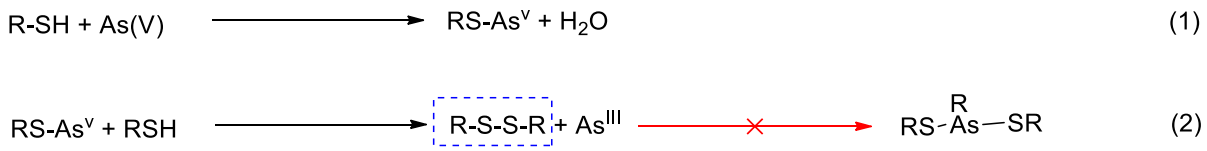
3.6 Results and discussion

3.6.1 Arsenic(V) reduction to As(III)

The first attempt at reducing As(V) to As(III) involved using KI and $SnCl_2$ under acidic conditions. The reaction was photosensitive therefore the reaction vessels were covered with foil and incubated at room temperature (25 °C) for a maximum of 60 min. To determine if the reaction was successful, the samples were spiked with 10 µM, 20 µM, and 30 µM of reduced ArsR peptides and incubated for another 30 min. The peptides only bind to As(III) therefore the successful reduction of arsenic can be confirmed by the formation of the peptide-arsenic complex. The samples were diluted in deionised water and analysed using LC-ESI-MS.

According to thiol-arsenic chemistry, reduced thiols should react with As(III) metabolites. However, the results obtained showed only the free peptide and no peptide-arsenic complexes. So, this means that the reduction was unsuccessful. This could be attributed to the following reasons: The iodide in the reaction can easily get oxidised by atmospheric oxygen which would then require the use of ascorbic acid to prevent the oxidation. The reduction of As(V) to As(III) using KI requires acid concentrations greater than 0.3M but the peptide could produce unwanted side reactions at high acid concentrations; hence, the acid concentration of the reaction cannot be increased. Reduction may also take up to four to five hours using KI, which is a lengthy process.[6]

The second attempt involved using the ArsR synthetic peptides as reducing agents prior to the actual binding studies. The pentavalent arsenic metabolites (iAs(V) and DMA(V) were spiked with 10 µM, 20 µM, and 30 µM of reduced ArsR peptides and analysed on the LC-ESI-MS. The mass obtained corresponded to that of a cyclised peptide and there was no complex detected. According to equation 3.1-(2) below, the reduction of As(V) to As(III) results in the formation of a disulfide bond which cannot undergo further binding since the binding site is not available for binding. It was hard to unambiguously confirm whether the reduction was successful or not using this method. This method requires excess amount of peptide as some of the peptides will used to reduce the arsenic while the remainder will be used to bind As(III).



Equation 3.1: Thiol dependant reduction of As(V) to As(III).[12]

Three thiol-containing compounds were investigated for their ability to reduce As(V) to As(III). The reaction was carried out in three different experiments where As(V) metabolites were incubated with L-cys, EDT, and DTT in excess. These samples were further spiked with 10 µM, 20 µM, and 30 µM of the reduced peptide to facilitate binding. The samples were diluted in deionised water and analysed using LC-ESI-MS. The sample treated with EDT consisted of two separate layers as EDT is not soluble in water and as a result, the reduction was not successful. In both the samples treated with L-cys and the one treated with DTT, the peptide-arsenic complex was detected. Further experiments were conducted to determine the most effective reducing agent. As the relative intensity of the peptide-arsenic complex detected for samples treated with L-cysteine was much higher than that of those treated with DTT, L-cysteine was used as a reducing agent in all our binding experiments.

3.6.2 Characterization

The binding of arsenic compounds to ArsR-derived peptides was determined by LC-ESI-MS. In principle, the binding between arsenic metabolites and the peptides should lead to a shift in mass. The binding of iAs(III) to the peptides can be characterised by an increase in mass by 76 Da, 74 Da, or 72 Da for each bound residue of iAs(III) depending on whether it is bound through one, two, or three covalent bonds. The binding of DMA(III) to the peptides is characterised by an increase in mass by 104 Da and PAO(III) binding is characterised by an increase in mass by 152 Da or 150 Da depending on whether one or two covalent bonds were formed. Table 3.1 below shows the results obtained during the binding studies where the complexes are either detected in the (+1) charge state as $[M+H]^{+}$ or in the (+2) charge state as $[M+2H]^{2+}$. A similar trend was depicted by all three peptides upon interacting with the arsenic metabolites, see supporting documents (SD-10 to SD-18, pages 125-133).

Table 3.1: Detailed summary of the complexes formed during arsenic binding to ArsR-derived peptides.

Complex name	Formula	Theoretical/ expected Mass	Detected $[M+H]^+$	Detected $[M+2H]^{2+}$	R_t/ min
Pep1-iAs	$C_{50}H_{83}AsN_{12}O_{18}S_3$	1310.43	1311.4415	656.2261	5.33
Pep1-DMA	$C_{52}H_{91}AsN_{12}O_{18}S_3$	1342.50	1343.4643	**	6.12
Pep1-PAO	$C_{56}H_{89}AsN_{12}O_{18}S_3$	1388.48	1389.4892	695.2492	6.27
Pep2-iAs	$C_{45}H_{78}AsN_{13}O_{16}S_3$	1227.41	1228.3069	614.6632	4.55
Pep2-DMA	$C_{47}H_{86}AsN_{13}O_{16}S_3$	1259.47	1260.5449	630.7795	8.77
Pep2-PAO	$C_{51}H_{84}AsN_{13}O_{16}S_3$	1305.45	1306.4577	653.7375	5.08
Pep3-iAs	$C_{45}H_{75}AsN_{14}O_{19}S_3$	1286.37	1287.3756	644.1955	4.36
Pep3-DMA	$C_{47}H_{83}AsN_{14}O_{19}S_3$	1318.43	1319.4793	660.2485	5.72
Pep3-PAO	$C_{51}H_{81}AsN_{14}O_{19}S_3$	1364.42	1365.4278	683.2197	5.06

The three peptides produced a complex with a mass shift of +72 Da when reacted with iAs(III). Although traces of peptides with a mass shift of +76 were observed in some cases, the major product for all peptides showed coordination to iAs(III) metabolite through all three sulfur groups. The three-coordinate binding of an iAs(III) to a Pep1-RJM has been illustrated in Figure 3.7 and the corresponding mass spectra can be seen in Figure 3.8.

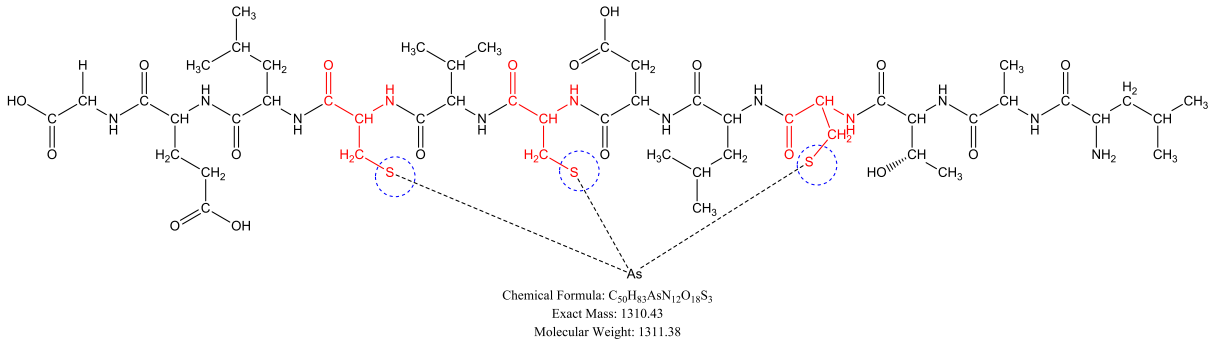


Figure 3.7: A typical example of iAs binding using Pep1-RJM. iAs is coordinated to the three thiol groups of the cysteine residues on the peptide.

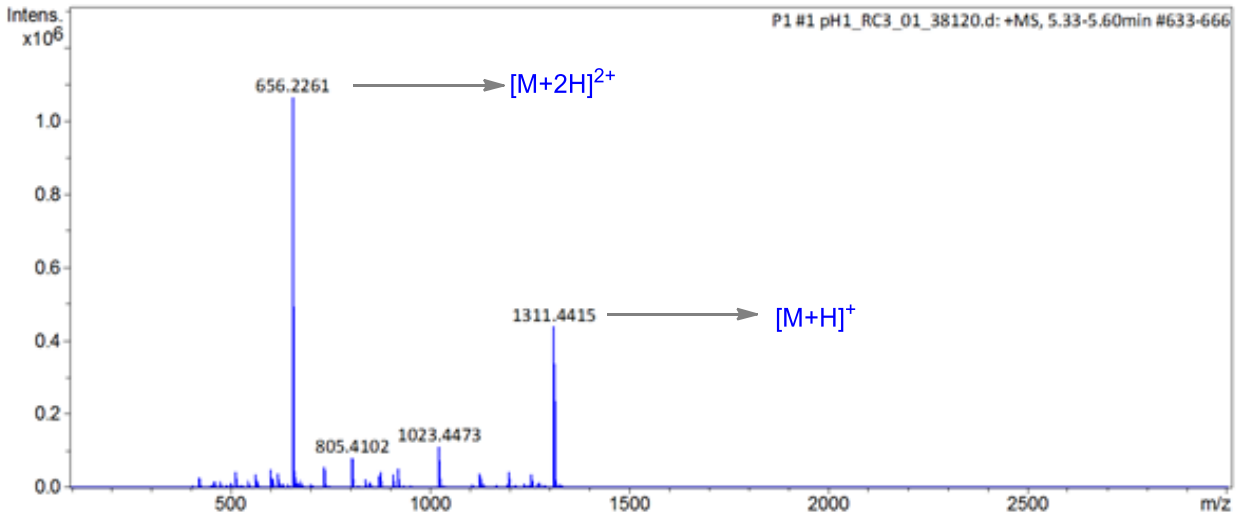


Figure 3.8: Mass spectrum of Pep1-iAs complex.

DMA(III) which already has two of its coordination sites occupied by methyl groups can only bind one thiol group and its binding was confirmed by the mass shift of +104 Da. PAO(III) which has one of its coordination sites occupied by a phenyl ring showed a mass shift of +150 Da which is the expected mass shift when Pep(x)-PAO complex is formed (Figure 3.9).

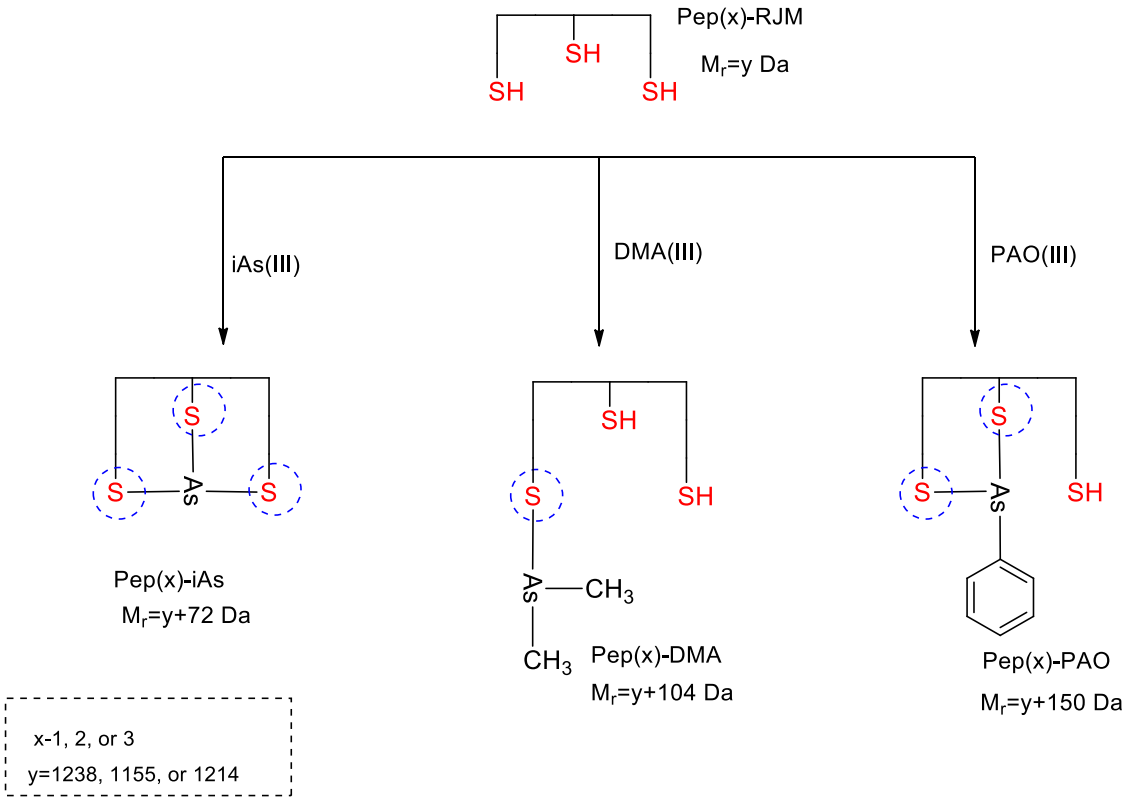


Figure 3.9: Interaction between iAs(III), DMA(III), and PAO(III) with ArsR-derived peptides containing three sulfur atoms from three different cysteine residues.

3.6.3 Effect of pH

The relative intensity of the arsenic-peptide complexes at different pH (2, 4, 6, 8. 10, 12, 14) were studied using 20 µM of peptide concentration at room temperature 25 °C, reaction time of 30 min and an initial arsenic concentration of 100 µM. Equations 3a, 3b, and 3c below describe the reactions of ArsR-peptides with iAs(III), DMA, and PAO, respectively.

$$P(SH)_3 + AsH_3O_3 \rightleftharpoons P_1(S)_2\text{-As-S} + 3H_2O \quad (3a)$$

$$P(SH)_3 + C_2H_7AsO \rightleftharpoons (SH)_2P_1\text{-S-As-}(CH_3)_2 + H_2O \quad (3b)$$

$$P(SH)_3 + C_6H_7AsO_2 \rightleftharpoons (SH)_1P_1\text{-}(S)_2\text{-As-}(C_6H_7) + 2H_2O \quad (3c)$$

There was no change in solution pH after the addition of the peptides into the reaction vessel. The samples were prepared and analysed using the HPLC-ESI-MS in triplicates. The average was used to plot a graph and the results are presented in Figure 3.10 below which shows the relative intensity of the arsenic-peptide complex formed at the specified pH range. All three peptides showed a similar trend, and as the pH of the solution increased from 2 to 6, arsenic binding gradually increased. Between pH 6 to 8, there was a noticeable increase in the formation of the arsenic-peptide complex with the relative intensity of the compounds on average increasing by 54%, 87%, and 71% respectively for iAs(III), DMA(III), and PAO(III). There is no significant increase in binding beyond pH 8 and the plot flattens at pH greater than 8. Although binding was observed across the studied pH range of 2-14, it is evident that the maximum relative intensity for arsenic binding occurred at pH 8. This means that basic conditions are required to achieve optimum results for the binding reactions. The binding of arsenic decreases with a decrease in pH and this may be due to the following: At lower pH, the concentration of H^+ in the urine matrix is high and the thionyl group on the peptide needs to be deprotonated before it can undergo nucleophilic attack of the arsenic metabolite. At higher pH, the urine matrix has a higher concentration of OH^- ions in solution which can catalyse the binding by abstracting a proton from the sulfhydryl group to produce an S^- anion which can act as a nucleophile and attack the arsenic metabolite.

$$R\text{-}SH + OH^- \longrightarrow R\text{-}S^- + H_2O \qquad (4)$$

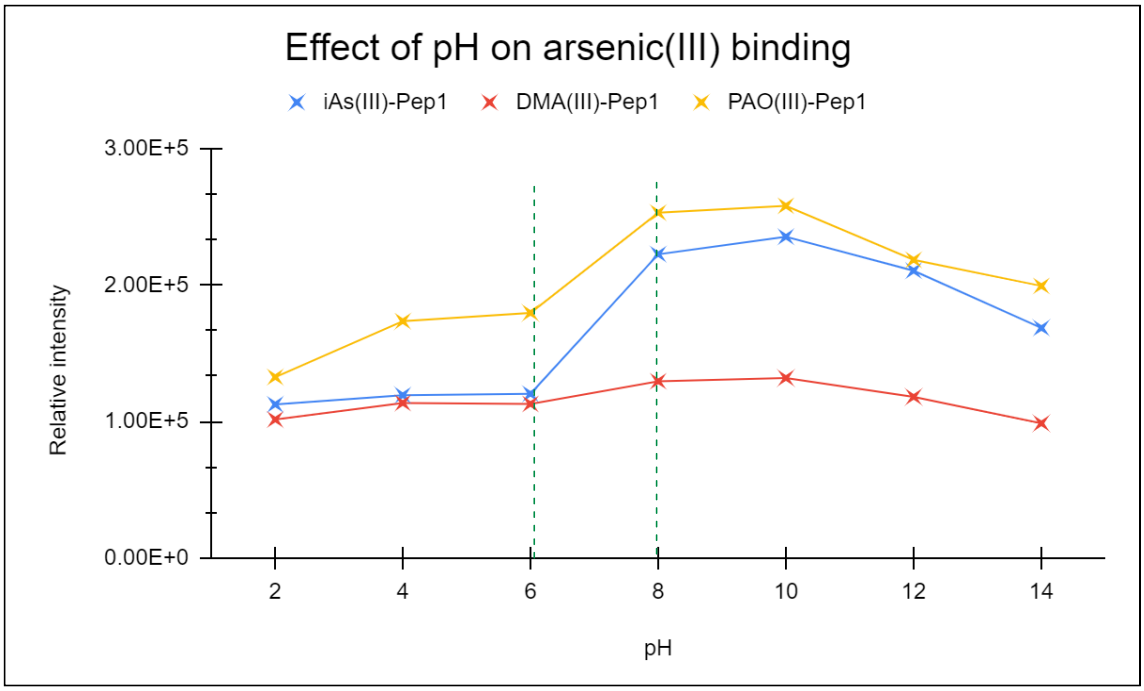


Figure 3.10: Effect of pH on the binding of arsenic(III) by Pep1-RJM.

3.6.4 Effect of temperature

The effect of temperature upon arsenic binding was studied by monitoring the relative intensity of the arsenic-peptide complex formed at various temperatures (25, 30, 35, 40, 45, 50, 55) °C. This was studied using 20 µM of peptide concentration at pH8, a reaction time of 30, min and an initial arsenic concentration of 100 µM. The results are represented in Figure 3.11 below and it was observed that there was a gradual increase in arsenic binding between 25-40 ˚C. Temperatures higher than 40 ˚C were observed to have a negative effect on the binding and this was mostly because the three peptides tend to form disulfide bridges upon heating thus the thiols are unavailable for binding. The further decrease in binding at temperatures >50 ˚C was due to the degradation of the peptide at these temperatures. The binding of arsenic to the three peptides was initially found to increase with an increase in temperature but shortly dropped upon temperatures higher than 40 ˚C, indicating that arsenic binding on the ArsR-derived peptides is favoured at lower temperatures. The highest binding was observed at 40°C for all 3 peptides, and this will be used as the optimum temperature.

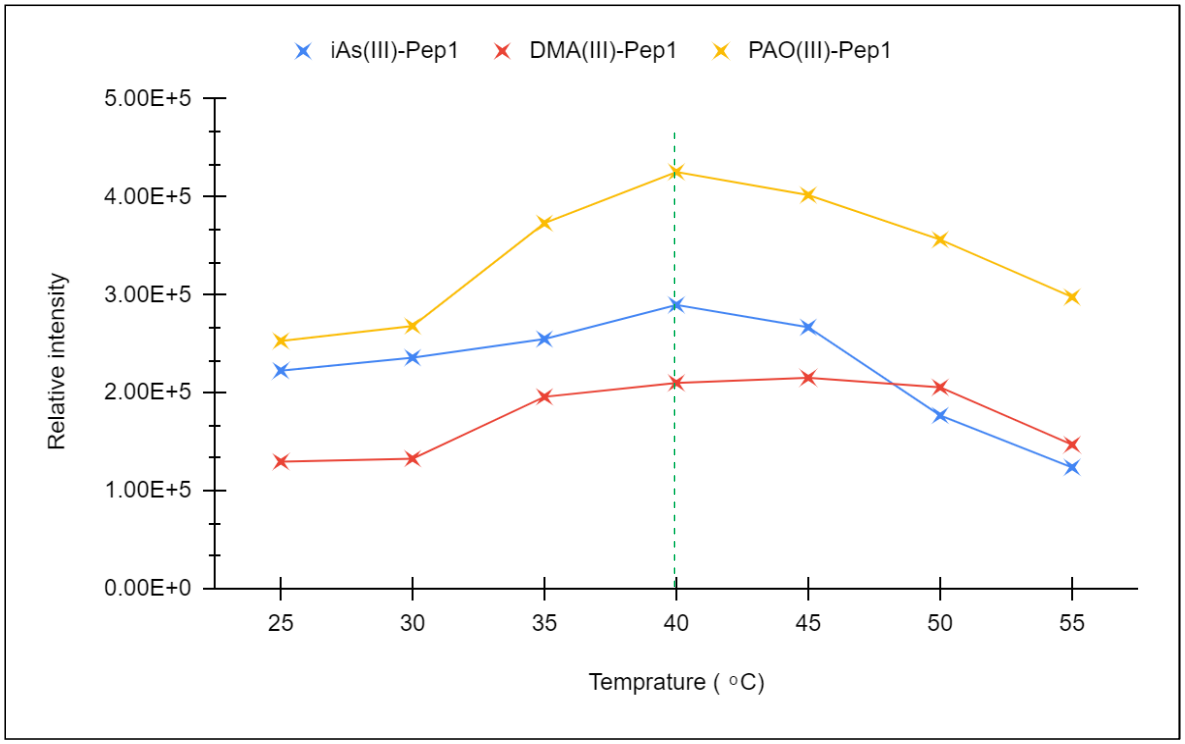


Figure 3.11: Effect of temperature on the binding of arsenic(III) by Pep1-RJM.

3.6.5 Effect of reaction time

Reaction time is an important parameter in binding studies as it determines the equilibrium time of the binding reaction. The binding was studied as a function of reaction time in the range of 0 to 30 min using 100 µM As(III) solution at 40 °C, at pH 8 and 20 µM of Pep1-RJM. The results presented in Figure 3.12 below show similar binding behaviour between the three arsenic species where As(III) binding increased with an increase in reaction time up to 15 min. There is a sharp increase of As(III) binding to Pep1-RJM between t=5 to t=15 and this is caused by a large amount of available binding sites for As(III) binding. As the reaction time elapsed, the available binding sites are saturated with As(III) species resulting in the flattening of the curve at t>15. Therefore, based on these results, 15 min was taken as the equilibrium time for further experiments.

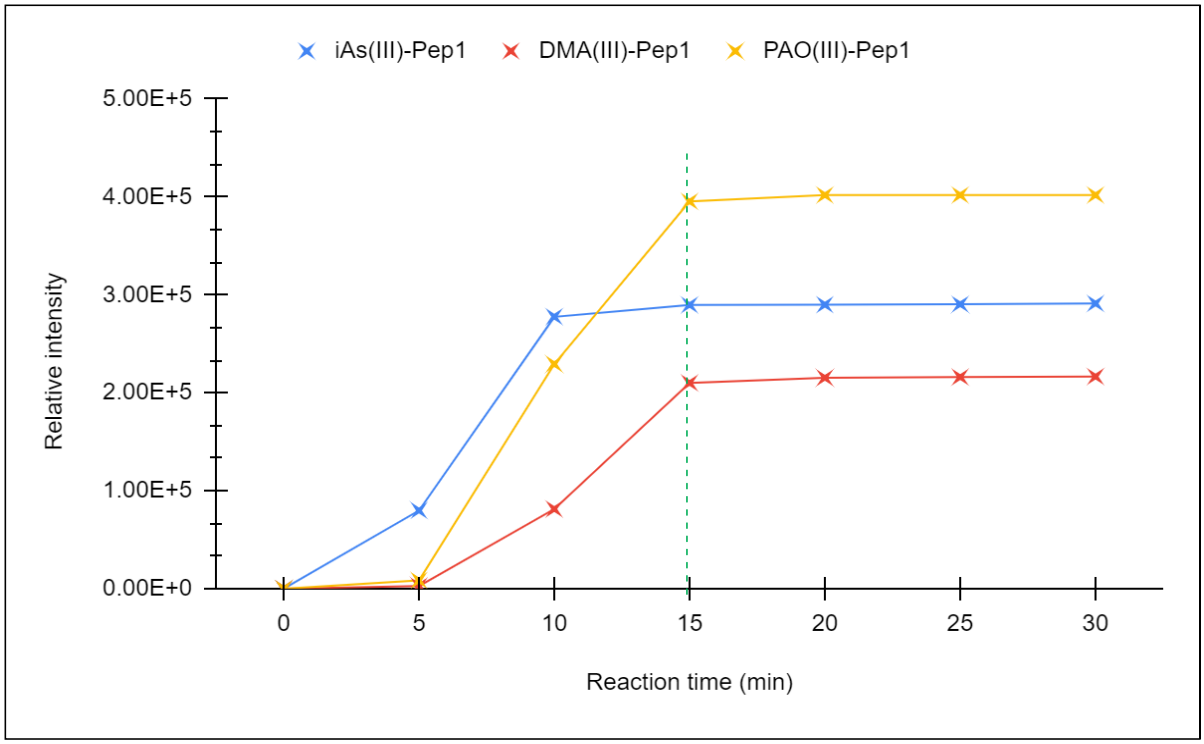


Figure 3.12: Effect of reaction time on the binding of arsenic(III) by Pep1-RJM.

3.6.6 Effect of initial As(III) concentration

The effect of the initial concentration of arsenic binding on ArsR-derived peptides was investigated by monitoring the relative intensity of arsenic-peptide complex formed at various initial arsenic concentrations (0, 10, 20, 40, 60, 80, 100) µM. This was studied using 20µM of peptide concentration at pH8, a reaction time of 15 min, and at 40 ˚C. The results are presented in Figure 3.13 below. When the initial concentration was varied from 10 to 60 µM, the binding capacity of Pep1-RJM increased from 11.5 to 28.5 % for iAs(III), 8.5 to 15.9 % for DMA(III), and 19.2 to 26.5% for PAO(III). The observed increase in binding capacity with the increase in As(III) concentration is probably due to higher interaction between As(III) species and Pep1-RJM. Hence the optimum concentration that can bind to 20µM of the peptide is 60µM. The next experiment will look at the effect of peptide concentration on As(III) binding.

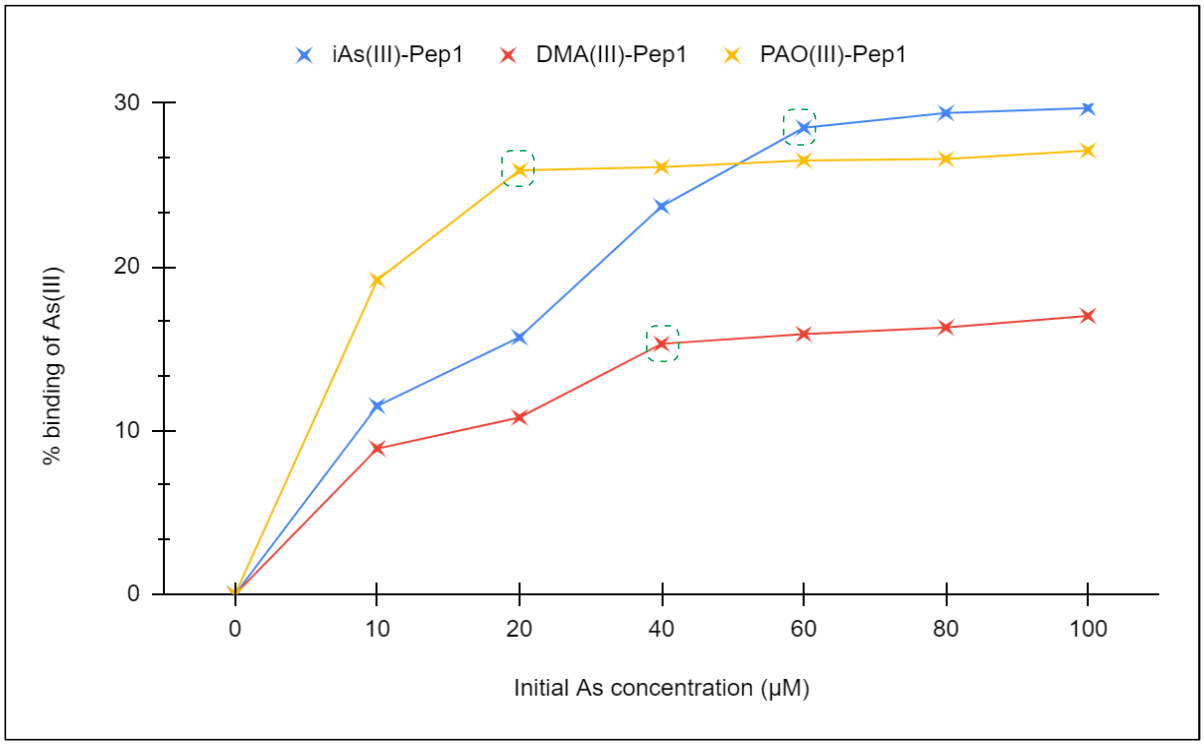


Figure 3.13: Effect of initial As(III) concentration on the binding of arsenic(III) by Pep1-RJM.

3.6.7 Effect of peptide dosage

The effect of peptide concentration is an important parameter to evaluate because it can help us determines the binding capacity of the peptide for a given As(III) concentration. This experiment was conducted using eight different peptide concentrations (5,10,15,20,25, 30, 35, and 40 µM). During the reaction, the pH of the solution was kept at 8.0. A reaction time of 15 minutes was used, with an arsenic concentration of 60 µM at a temperature of 40°C. The results for this study are presented in Figure 3.14 below. When the peptide concentration was increased from 0 to 20 µM, the binding of arsenic increased with an increase in peptide concentration attaining a maximum at 20 µM for DMA(III) and PAO(III) and a maximum of 25 µM for iAs(III) of peptide concentration. Higher concentration of peptide greater number of binding sites for arsenic metabolites. It was noted that at concentrations >25 µM there was no significant increase in the binding of As(III) instead the plot remains constant for iAs(III) and a slight decrease is observed for DMA(III) and PAO(III). This may be due to peptide aggregation and overlapping active binding sites at higher concentrations. As a result, 25 µM was selected as the optimum concentration and was used for further study.

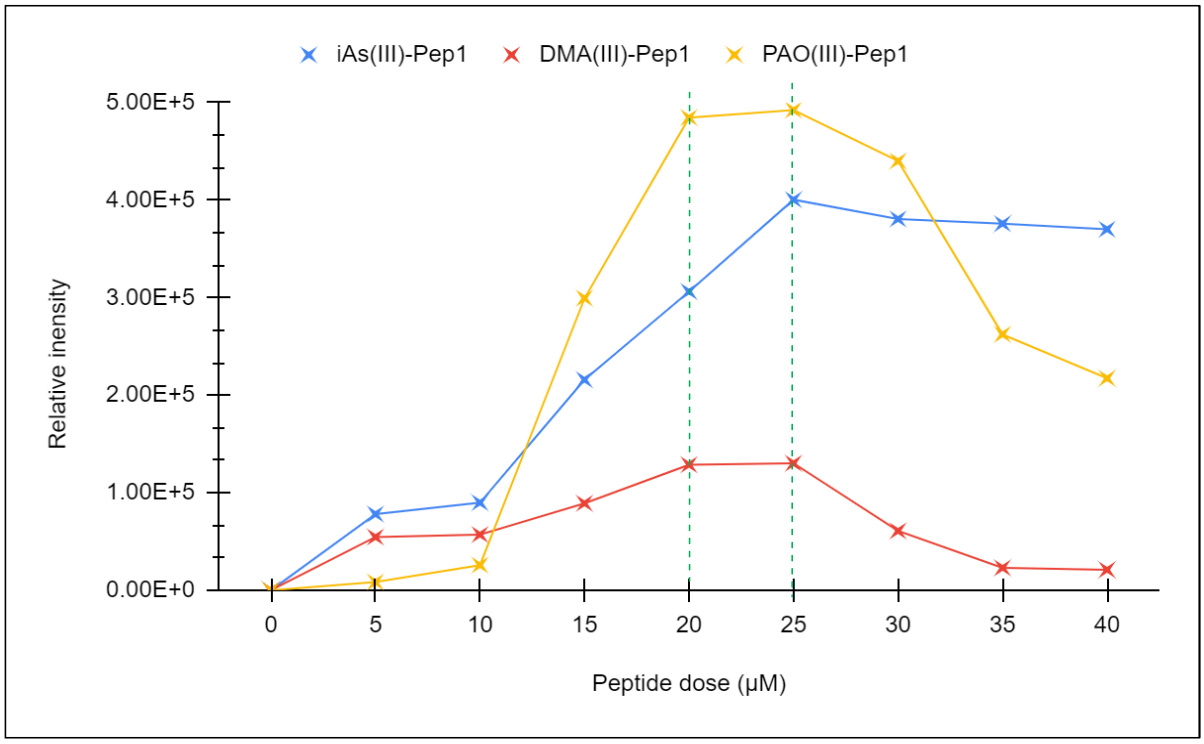


Figure 3.14: Effect of peptide concentration on the binding of arsenic(III) by Pep1-RJM.

3.6.8 Competitive binding studies

A study was conducted to examine the effect of competing heavy metal ions upon the binding of As(III). Nickel (Ni(II), lead (Pb(II), chromium (Cr(II), and copper (Cu(II)) species were used as the competing ions. The initial concentration of As(III) ions was maintained at 60 µM while the initial concentrations of heavy metal ions were varied between 10, 20, 40, 80, and 100 µM. This experiment was conducted at a pH of 8.0, a Pep1-RJM concentration of 20 µM, and a temperature of 40˚C for 15 min. The presence of metal ions in the solution resulted in minor interferences where a fraction of Pep1-RJM was converted into its oxidized form. However, none of the detected masses corresponded to the binding of the peptides with Ni(II), Pb(II), Cr(II), and Cu(II). Arsenic binding was still strongly observed in the presence of competing metal-ion species.

3.6 Concluding remarks

The study presented here showed that ArsR-derived peptides (Pep1-RJM, Pep2-RJM, and Pep3-RJM) can be used as promising bioreceptors for the binding of As(III) from urine. The binding experiments showed that As(III) could be bound with great selectivity in a relatively short time of 15 minutes. The LC-ESI-MS has shown to be a rapid and sensitive tool for the determination of peptide-arsenic binding. The technique can be used to determine the number of arsenic residues attached to each peptide. As(III) binding to Pep1-RJM was studied by conducting a batch of experiments and was found to be strongly dependent on the pH of the solution. The initial concentration of As(III), the peptide concentration and temperature also had an effect on the binding reaction. According to the above experiments, pH 8.0, reaction temperature of 40 °C, reaction time of 15 minutes, initial As(III) concentration of 60 µM, and peptide concentration of 20 µM were found to be the optimal conditions for the arsenic binding reaction. The next chapter will present the quantitative analysis and binding constant of Pep1-RJM with As(III) species under these optimum binding conditions.

References

(1) Kožíšek, M.; Svatoš, A.; Buděšínský, M.; Muck, A.; Bauer, M. C.; Kotrba, P.; Ruml, T.; Havlas, Z.; Linse, S.; Rulíšek, L. Molecular Design of Specific Metal-Binding Peptide Sequences from Protein Fragments: Theory and Experiment. *Chem. - A Eur. J.* **2008**, *14* (26), 7836–7846. https://doi.org/10.1002/chem.200800178.

(2) Khairul, I.; Wang, Q. Q.; Jiang, Y. H.; Wang, C.; Naranmandura, H. Metabolism, Toxicity and Anticancer Activities of Arsenic Compounds. *Oncotarget* **2017**, *8* (14), 23905–23926. https://doi.org/10.18632/oncotarget.14733.

(3) Issa, N. Ben; Rajaković-Ognjanović, V. N.; Marinković, A. D.; Rajaković, L. V. Separation and Determination of Arsenic Species in Water by Selective Exchange and Hybrid Resins. *Anal. Chim. Acta* **2011**, *706* (1), 191–198. https://doi.org/10.1016/j.aca.2011.08.015.

(4) Byambaa, E.; Seon, J.; Kim, T. H.; Kim, S. D.; Ji, W. H.; Hwang, Y. Arsenic (V) Removal by an Adsorbent Material Derived from Acid Mine Drainage Sludge. *Appl. Sci.* **2021**, *11* (1), 1–13. https://doi.org/10.3390/app11010047.

(5) Nicomel, N. R.; Leus, K.; Folens, K.; Van Der Voort, P.; Du Laing, G. Technologies for Arsenic Removal from Water: Current Status and Future Perspectives. *Int. J. Environ. Res. Public Health* **2015**, *13* (1), 1–24. https://doi.org/10.3390/ijerph13010062.

(6) Chen, H.; Brindle, I. D.; Le1, X. chun. Prereduction of Arsenic(V) to Arsenic(III), Enhancement of the Signal, and Reduction of Interferences by L-Cysteine in the Determination of Arsenic by Hydride Generation. *Anal. Chem.* **1992**, *64* (6), 667–672. https://doi.org/10.1021/ac00030a018.

(7) Kastin, A. J. Handbook of Biologically Active Peptides. In *Handbook of Biologically Active Peptides*; 2013; pp 1721–1729.

(8) Mthembu, S. N.; Sharma, A.; Albericio, F.; de la Torre, B. G. Breaking a Couple: Disulfide Reducing Agents. *ChemBioChem* **2020**, *21* (14), 1947–1954. https://doi.org/10.1002/cbic.202000092.

(9) Ágoston, V.; Cemazar, M.; Kaján, L.; Pongor, S. Graph-Representation of Oxidative Folding Pathways. *BMC Bioinformatics* **2005**, *6* (February 2005). https://doi.org/10.1186/1471-2105-6-19.

(10) Quináia, S. P.; Rollemberg, M. D. C. E. Selective Reduction of Arsenic Species by Hydride Generation - Atomic Absorption Spectrometry. Part 2 - Sample Storage and Arsenic Determination in Natural Waters. *J. Braz. Chem. Soc.* **2001**, *12* (1), 37–41. https://doi.org/10.1590/S0103-50532001000100004.

(11) Feng, Y. lai; Cao, J. ping. Simultaneous Determination of Arsenic(V) and Arsenic(III) in Water by Inductively Coupled Plasma Atomic Emission Spectrometry Using Reduction of Arsenic(V) by l-Cysteine and a Small Co-Centric Hydride Generator without a Gas-Liquid Separator. *Anal. Chim. Acta* **1994**, *293* (1–2), 211–218.

https://doi.org/10.1016/0003-2670(94)00069-7.

(12) Gregus, Z.; Roos, G.; Geerlings, P.; Németi, B. Mechanism of Thiol-Supported Arsenate Reduction Mediated by Phosphorolytic-Arsenolytic Enzymes: II. Enzymatic Formation of Arsenylated Products Susceptible for Reduction to Arsenite by Thiols. *Toxicol. Sci.* **2009**, *110* (2), 282–292. https://doi.org/10.1093/toxsci/kfp113.

Chapter 4 : Quantitative characterization of arsenic(III)-binding to Pep1-RJM using Liquid-chromatography hyphenated to electrospray ionization quadrupole time-of-flight mass spectrometry.

4.1 Introduction

Peptide regions with specific sequences have been used by researchers to mimic the interaction between proteins and their target analytes. The quantitative analysis of peptide-metal complexes has particularly shown to be useful in environmental, pharmaceutical, toxicological, and bioanalytical sciences. The electrospray ionisation mass spectrometry (ESI-MS) has been shown to provide a great deal of information thus emerging as an important technique for quantitatively understanding the nature of peptide-metal complexes due to its speed, sensitivity, and selectivity. Additionally, ESI-MS experiments have been used to determine the binding stoichiometry of peptide-metal complexes, including the degree of association by determining the binding constants. During the ESI process, only ions preserved in the gas phase can be detected, and the covalent bonds remain intact during desolvation and ionization. However, nonspecific cluster ions have also been reported during the ESI process, which does not exist in the original sample solution. Although the use of ESI-MS has been previously reported in the majority of binding studies, systematic parameter variations have not been reported in the determination of the binding constants. Outmost care must be taken with the quantitative binding studies to ensure that reliable and reproducible data is obtained. Research articles have presented ESI-MS-based experiments of various peptide-metal complexes, but studies focusing on the quantification of arsenic binding to peptides are rare or limited.[1–5]

A consensus report was developed by a subcommittee of the Ligand Binding Assay Bioanalytical Focus Group (LBABFG) in response to the lack of guidance for method development and validation of bioanalytical methods for biomolecules. This document provides specific recommendations on how ligand-binding studies can be used to support biomarker concentration, toxicokinetic, or pharmacokinetic assessments of biomolecules in

biological matrices such as urine. There are three major stages in the quantitative assessment of peptide-ligand interactions. These stages include method development, pre-study validation, and in-study validation. The experimental concept is designed and tested during the method development phase which is confirmed during the pre-study validation phase and is subsequently implemented during the in-study validation.[6,7]

The binding behaviour of the urinary arsenic metabolites to Pep1-RJM has been studied qualitatively using ESI-MS in chapter 3. The purpose of this chapter is to develop and optimise an ESI-MS method for the quantitative analysis of ArsR peptide-arsenic complexes using the LC-ESI-MS. Additionally, this chapter aims to determine the binding stoichiometry of the reactions between thiol-containing Pep1-RJM, to obtain reliable apparent binding constants, and to describe the operating conditions, limitations, and suitability of this method for detecting arsenic in urine.

4.2 Binding constants from LC-ESI-MS

According to the IUPAC, equilibrium constants are quantities that characterize chemical reactions at equilibrium. The equilibrium constant (K) is also known as the formation constant or binding constant, and it compares the ratio of products to reactants. In essence, this constant measures how tightly a ligand is bound to a receptor. A higher constant implies a closer binding of the ligand to the biomolecule or a higher affinity between the biomolecule and the ligand. Schmidt et. al., (2012) demonstrated a method to calculate the equilibrium constants for the reaction of peptides with ligands, which was used to determine the binding constants for the reaction of peptides with different ligands.[5] Equation 4.1-4.2 gives an estimate of the equilibrium constants of equilibrium systems.

$$P + L \rightleftharpoons P\text{-}L + H_2O \qquad (4.1)$$

$$K = \frac{[P-L]_{eq}\ x\ [H_2O]}{[P]_{eq}\ x\ [L]_{eq}} \qquad (4.2)$$

The approximate calculation for the binding constant is based on the changing peak areas of mass signals derived from the concentration series.

The symbols for equations 4.1-4.2 above and equations 4.3-4.4 below are defined as follows:

$[P\text{-}L]_{eq}$ is the equilibrium concentration of ligand-bound biomolecule,

$[P]_{eq}$ is the equilibrium concentration of the unbound biomolecule,

$[L]_{eq}$ is the equilibrium concentration of ligands of interest.

$[P]_0$ is the initial concentration of the biomolecule,

PA is the peak area,

b represents the sensitivity coefficient for the ESI-MS detection.

The mass signal for the detected reactants or products were either in (+1) or (+2) charge state.

This study was conducted using the titration method where a fixed concentration of the peptide was titrated against different concentrations of the ligand (iAs(III), DMA(III), PAO). The concentration of the formed peptide-arsenic complexes as a result of the decreasing reactant concentration was calculated using the ion signals from the ESI-MS. This study does not report any thermodynamic constants; however, it provides apparent values that are sufficient for comparison of the binding affinity of Pep1 with different urinary arsenic species under similar experimental conditions. Calculations of binding constants will be based on peak areas (PA) obtained from LC-ESI-MS.

4.3 Materials and methods

4.3.1 Sample preparation

The Pep1-RJM peptide (80.7 µM, 0.263 pmol) was first reduced using excess DTT (1.5 pmol,5 equiv, 5 mL). It was then incubated with 10, 20, 30, 40, 50, and 60 µM of each arsenic metabolite (Table 4.1) using the 6-vessel reactor presented in chapter 3. These reactions were conducted at 40 °C and at pH 8 for 15 minutes. The LC-ESI-MS measurements were carried out immediately after the 15 min incubation period. All sample measurements were obtained three times.

4.3.2 LC-ESI-MS

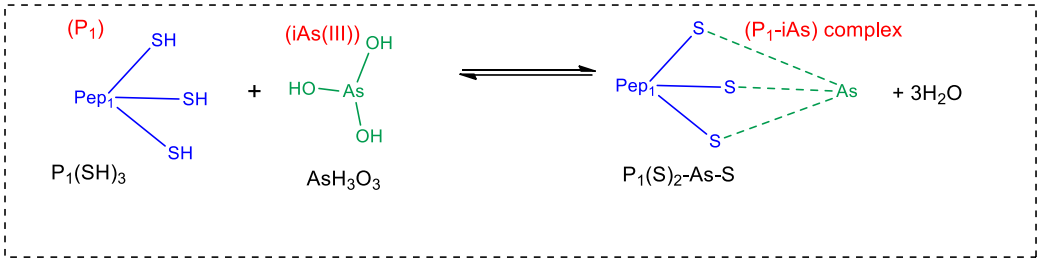
The UHPLC-ESI-MS was used to monitor the reaction during method development and arsenic during binding studies. The UHPLC component (Thermo Scientific Ultimate 3000) was equipped with a binary solvent system where solvent A consisted of H_2O and 0.1% FA (v/v) and solvent B consisted of ACN and 0.1% FA (v/v). The samples were placed on an autosampler at 25 °C and 20 µL of the samples were injected onto a reversed-phase C18 analytical column that separated the peptides from the other components in the synthetic urine matrix using a gradient elution method of 5 to 95% ACN in 12 minutes. The separated analytes were detected by a Quadrupole-Time-Of-Flight high-resolution mass spectrometer (Bruker Daltonics, Bremen, Germany). The samples were analysed in the positive ionization mode and the MS was scanned at m/z 100-3000 range during separation and detection. Nitrogen gas (N_2) was used as the dry and nebulizer gas. The nebuliser gas was operated at a pressure of 1.8 bars whereas the drying gas was operated at a flow rate of 9 L/min. To guarantee the reliability of the results, the ESI source was cleaned and calibrated every week.

Table 4.1: Concentrations of pep1-RJM with individual As(III) species used in the titration reactions.

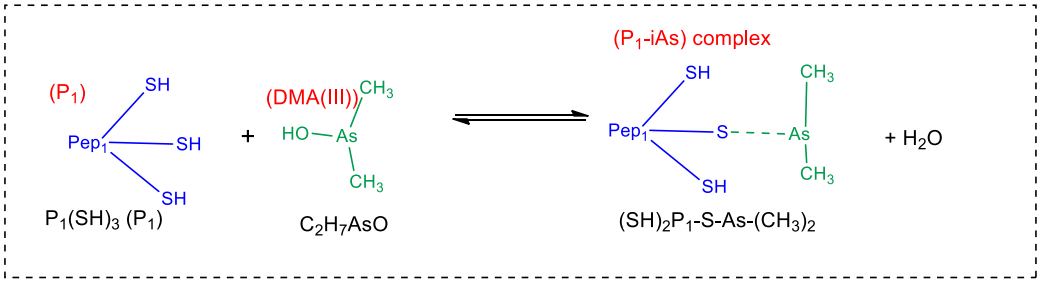
Pep1-RJM (µM)	iAs(III) (µM)	DMA(III)	PAO(III)
20	10	10	10
20	20	20	20
20	30	30	30
20	40	40	40
20	50	50	50
20	60	60	60

4.3.3 Binding calculations[5]

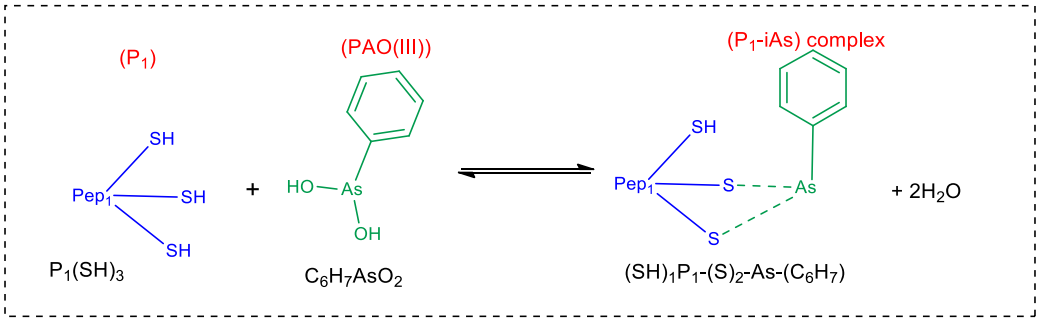
The condensation reactions of Pep1-RJM with arsenic metabolites (iAs, DMA, PAO) are represented in equations 4.3(a-c) below.



(a)



(b)



(c)

The calculation for binding constant (K) involves dividing the reaction products [P_1-As] by the reactants $P_1(SH)_3$ and As(III) at equilibrium. For simplicity, $P_1(SH)_3$ will be represented as [P_1] throughout the chapter and As(III) represents iAs (III), DMA(III) or PAO(III). Water is generated as a reaction product in the above reactions but the concentration of water emerging from these reactions is significantly small and can be neglected.[3]

$$K = \frac{[P_1 - As]_{eq}}{[P_1]_{eq} . [As(III)]_{eq}} \quad (4.4)$$

Eq 4.5a was used to determine the concentration of unreacted peptide at equilibrium using 25 µM as the initial peptide concentration $[P_1]_0$ and the peak areas (PA) of the unbound peptide and complex formed were obtained from the extracted ion chromatogram (EIC) from the LC-ESI-MS analysis. Eq 4.5b was used to calculate the concentration of complex formed during the condensation reaction.

$$[P_1]_{eq} = \frac{[P_1]_0 . PA(EIC)_{P1\ unbound}}{(PA(XIC)_{P1,unbound} + PA(XIC)_{P1-As})} \quad (4.5a)$$

$$[(P_1 - As)_{complex}]_{eq} = \frac{[P_1]_0 . PA(EIC)_{(P_1-As),complex}}{(PA(XIC)_{P1,unbound} + PA(XIC)_{(P1-As),complex}} \quad (4.5b)$$

The concentration of unreacted arsenic metabolites at equilibrium was obtained by using eq 4.5c below. The concentration of the complex formed was subtracted from known initial concentration of arsenic (10, 20, 30, 40, 50, or 60).

$$[As]_{eq} = [As]_0 - [P_1 - As]_{eq} \quad (4.5c)$$

The second condition presumes a similar ionization efficiency during the electrospray process resulting in a similar sensitivity coefficient b for the unbound peptide and its reaction products with the arsenic compounds (eq 4.6)

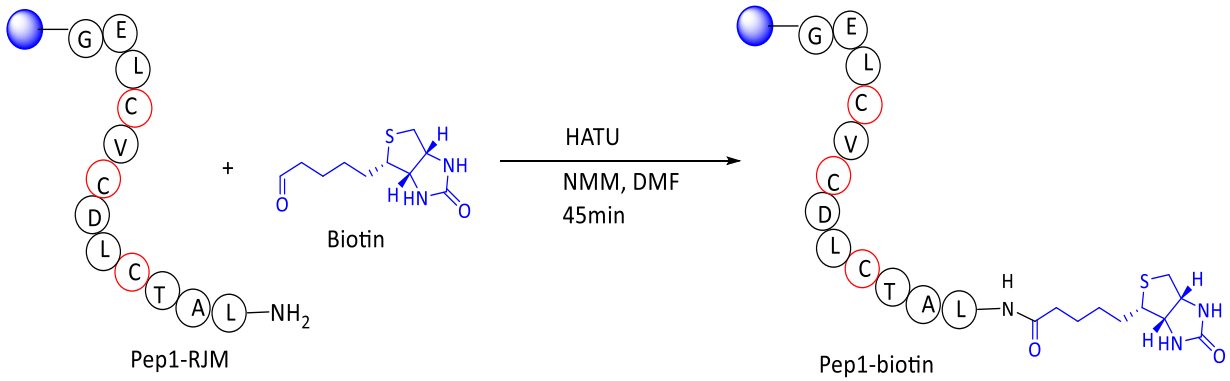
$$b_{\mathrm{P1,unbound}} = b_{(\mathrm{P1-As})} = b \quad (4.6)$$

$$b_{\mathrm{P1,unbound}} = \frac{[P_1]_{eq}}{PA(XIC)_{P1,unbound}} \quad (4.6a)$$

$$b_{(\mathrm{P1-As})} = \frac{[P_1 - \mathrm{As}]_{eq}}{PA(XIC)_{P1-As}} \quad (4.6b)$$

4.3.3 Pep1-RJM biotinylation

The resin-bound Pep1-RJM was Fmoc deprotected using 20% piperidine in DMF as described in chapter 3. Biotin (1.20 mmol, 0,2 M) was dissolved in DMSO and coupled to Pep1-RJM using HATU (1.14 mmol, 0.19M) as a coupling reagent and N-methylmorpholine (NMM) (10 mmol, 1M) in DMF for 45 minutes (Scheme 4.1). The coupling reaction was repeated. The biotinylated Pep1 (Pep1-Bio) was cleaved from the resin for 2 hrs using the TFA cleaving cocktail described in chapter 2. The crude Pep1-bio was analysed using the HPLC-ESI-MS equipped with a C18 column and purified using a semi-preparative HPLC with UV/VIS detector.



Scheme 4.1: Pep1-RJM biotinylation.

4.4 Results and discussion

The formation of the covalent complex resulting from the condensation reaction between Pep1 and As(III) at different molar ratios was analysed using the LC-ESI-MS. It was observed that at different molar ratios of the As(III) the coordination number of the resulting complex was fixed at 1:1 (Figure 4.1, structure A) , which implies the coordination number of the final product is independent of the initial molar concentration of the As(III). Based on this, we infer that the resulting covalent complexes are not formed by a cluster of arsenic species (Figure 4.1, structure B).

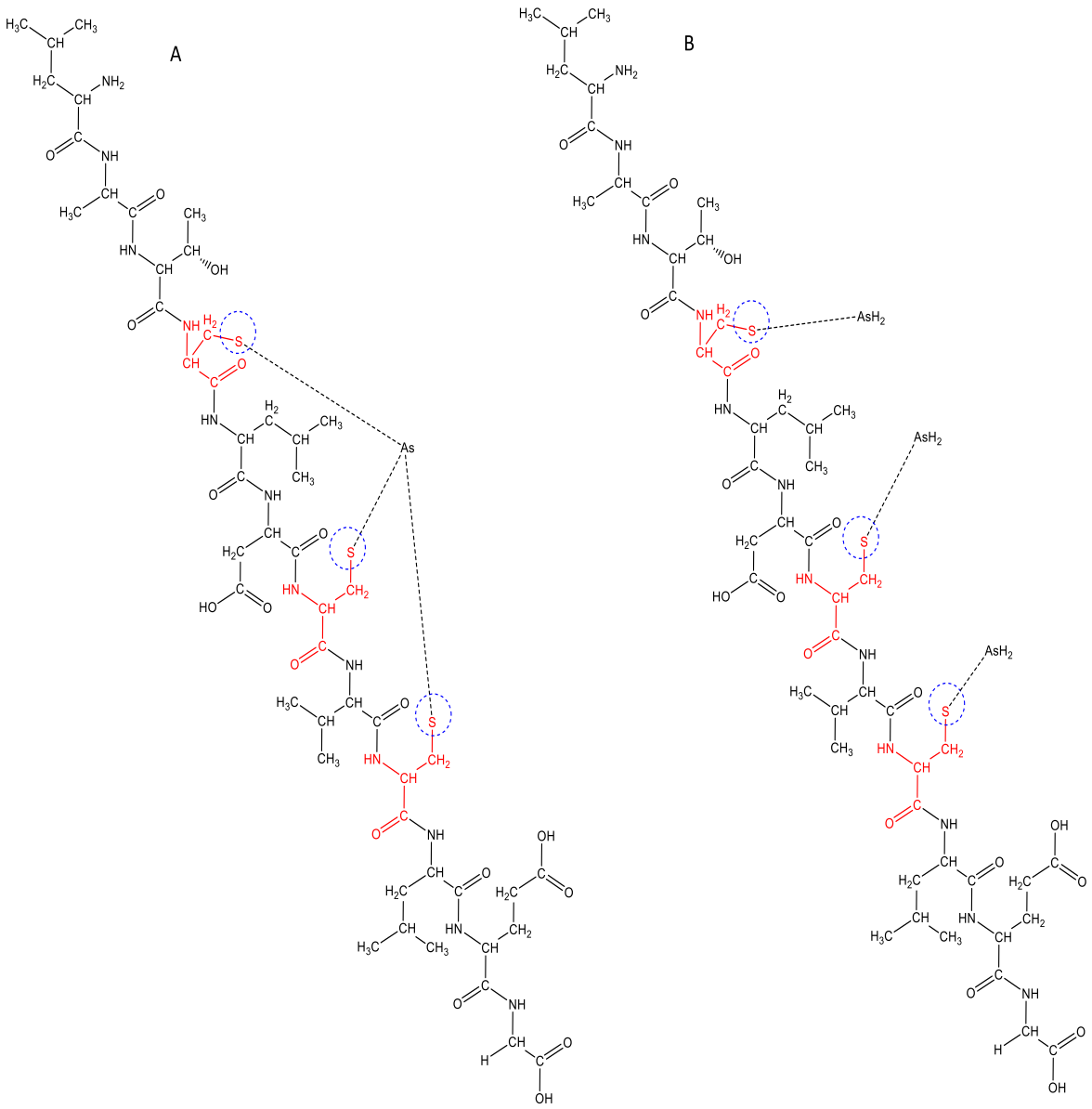


Figure 4.1: structural example of Pep1-RJM coordination to iAs(III). Structure A shows a typical 1:1 coordination whereas structure B shows a cluster coordination.

4.4.1 Binding constant calculations

The following study was conducted using the titration method. A series of samples with a constant concentration of Pep1-RJM (25 µM) and an increasing concentration of As(III) from 10-60 µM were mixed according to Table 4.1 above. The peak areas from the extracted ion chromatogram of the unbound pep1 and formed complexes obtained from the titration measurements were then converted into concentrations using equations 4.5a-c.

It can be seen that the concentration of unbound peptide decreases as the concentration of As(III) increases. Table 4.3 to Table 4.5 below shows a summary of the estimated concentrations and the apparent binding constants from the series of conducted reactions. The final binding constant (K) results for the P_1-iAs, P_1-DMA, P_1-PAO complexes are listed in the far-right column of each table. The three tables show that the binding strength for complexes of Pep1 with iAs(III), DMA(III), and PAO(III) decrease in the following order P_1-PAO> P_1-iAs> P_1-DMA. The calculated binding constants for iAs and PAO are characterised by an increasing trend with an increase in iAs or PAO. This observation is possibly due to the strongly increasing peak areas of the peptide-arsenic complexes as the initial As(III) concentration is increased and this is consistent with the findings in chapter 3 , where the binding capacity of Pep1-RJM increased with increasing As(III) concentration from 11.5 to 28.5 % for iAs(III), and 19.2 to 26.5% for PAO(III).

Table 4.2: Observed masses of Pep1-As complexes.

Complex name	**Formula**	**Theoretical/ expected Mass**	**Detected $[M+H]^+$**	**Detected $[M+2H]^{2+}$**	**R_t/ min**
Pep1-iAs	$C_{50}H_{83}AsN_{12}O_{18}S_3$	1310.43	1311.4415	656.2261	5.33
Pep1-DMA	$C_{52}H_{91}AsN_{12}O_{18}S_3$	1342.50	1343.4643	**	6.12
Pep1-PAO	$C_{56}H_{89}AsN_{12}O_{18}S_3$	1388.48	1389.4892	695.2492	6.27

The calculated binding constants for DMA showed no significant trend with an increase in DMA concentration. According to the masses observed in Chapter3 and structures presented in Figure 4.2 below DMA(III) is presumed to bind to one sulphur group on the cysteine residue of Pep1RJM whereas iAs(III) and PAO(III) bind to three and two sulphur groups respectively (Figure 4.2). This could potentially be one of the reasons for the poor binding behaviour of DMA(III) that is observed throughout this study. Future computational and nuclear magnetic resonance (NMR) spectroscopy studies will afford more information about the favourable conformation of the binding site as well as give insight about the bond lengths of each As-S formed in the complexes.

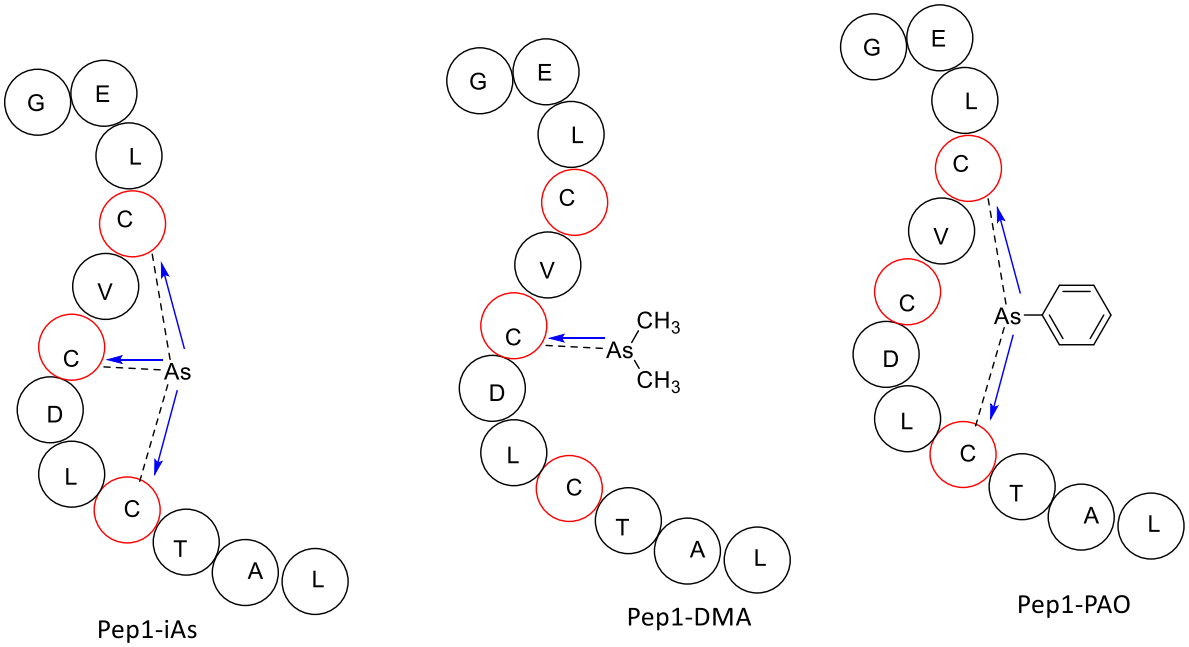


Figure 4.2: Structure of Pep1-As complexes with covalent bonds contributing to the binding affinity of the complex.

Complexes with binding constants less than 10^4 M^{-1} are considered to be weak interactions.[8] The titration experiments conducted in this study produced binding constants that are significantly greater with the highest binding constants obtained being 3.47 x 10^8, 2.00 x 10^8, and 4.11 x 10^8 M^{-1} for iAs(III), DMA(III), and PAO(III) respectively. Although the peptides show different degrees of binding with each As(III) species, all the binding constants are relatively good and comparable to binding constants presented by Schmidt, A. C. et al (2009, 2012) during their arsenic binding studies [3,5] Models with an RSD absolute error of ≤10% are considered acceptable, all the calculated concentrations reported in this study were reported with an <10% RSD error, thus fit within the criteria.[6]

Table 4.3: Binding constant reaction of Pep1-RJM with iAs(III) obtained by means of peak areas using LC-ESI-MS.

$[P_1]_0$ (µM)	$[iAs(III)]_0$ (µM)	LC-ESI-MS PA(EIC) (Unbound)	%RSD	LC-ESI-MS PA(EIC) (Complex)	% RSD	$[P]_{eq}$ (µM)	$[P\text{-}L]_{eq}$ (µM)	$[L]_{eq}$ (µM)	K (M^{-1})
25	0	423128.95	1.84	-	-	-	-	-	-
25	10	378325.33	2.78	46274.20	2.73	22.28	2.72	7.28	1.68×10^8
25	20	333325.35	3.18	94134.83	3.64	19.49	5.51	14.49	1.95×10^8
25	30	271194.17	1.35	157073.26	2.17	15.83	9.17	20.83	2.78×10^8
25	40	234653.50	2.19	196154.65	4.71	13.62	11.38	28.62	2.92×10^8
25	50	204403.14	3.91	228321.98	2.02	11.81	13.19	36.81	3.03×10^8
25	60	166917.98	2.96	259625.64	3.37	9.78	15.22	44.78	3.47×10^8

Table 4.4: Binding constant reaction of Pep1-RJM with DMA(III) obtained by means of peak areas using LC-ESI-MS.

$[P_1]_0$ (µM)	$[DMA(III)]_0$ (µM)	LC-ESI-MS PA(EIC) (Unbound)	%RSD	LC-ESI-MS PA(EIC) (Complex)	%RSD	[P]eq (µM)	[P-L]eq (µM)	[L]eq (µM)	K (M^{-1})
25	0	431208.63	2.62	-	-	-	-	-	-
25	10	379953.42	3.29	46113.61046	4.82	22.29	2.71	7.29	1.66×10^8
25	20	331013.78	3.40	95488.82015	2.89	19.40	5.60	14.40	2.00×10^8
25	30	319151.13	3.87	115034.4684	2.08	18.38	6.62	23.38	1.54×10^8
25	40	279489.88	3.56	141718.7476	1.77	16.59	8.41	31.59	1.61×10^8
25	50	259761.06	2.93	179724.9614	3.95	14.78	10.22	39.78	1.74×10^8
25	60	236908.50	2.43	194073.3423	2.01	13.74	11.26	48.41	1.68×10^8

Table 4.5: Binding constant reaction of Pep1-RJM with PAO(III) obtained by means of peak areas using LC-ESI-MS.

$[P_1]_0$ (µM)	$[PAO(III)]_0$ (µM)	LC-ESI-MS PA(EIC) (Unbound)	%RSD	LC-ESI-MS PA(EIC) (Complex)	%RSD	[P]eq (µM)	[P-L]eq (µM)	[L]eq (µM)	K (M^{-1})
25	0	433594.23	4.96	-	-	-	-	-	-
25	10	374953.44	3.53	58641.02	4.32	21.62	3.38	6.62	2.36×10^8
25	20	311013.78	2.14	107580.55	3.91	18.57	6.43	13.57	2.55×10^8
25	30	299151.1	3.72	199442.69	2.79	15.00	10.00	20.00	3.33×10^8
25	40	229489.88	4.29	228154.35	2.06	12.54	12.46	27.54	3.61×10^8
25	50	183813.06	3.02	239781.17	3.95	10.85	14.15	35.85	3.64×10^8
25	60	160608.50	4.49	289585.73	3.16	8.92	16.08	43.92	4.11×10^8

4.4.2 Pep1-RJM biotinylation

The biotinylation of peptides is important for various applications including biosensor application. The biotin-labelled Pep1-RJM in this study will be immobilised for future studies on to a streptavidin-coated gold electrode for As(III) detection. Figure 4.3 shows the structure of the biotinylated Pep1-RJM and *Figure* 4.4 below shows the mass spectra with the observed m/z of 1465.5503 and 733.2820 of the biotin-bound Pep1-RJM corresponding to $[M+H]^{+}$ and $[M+2H]^{2+}$.

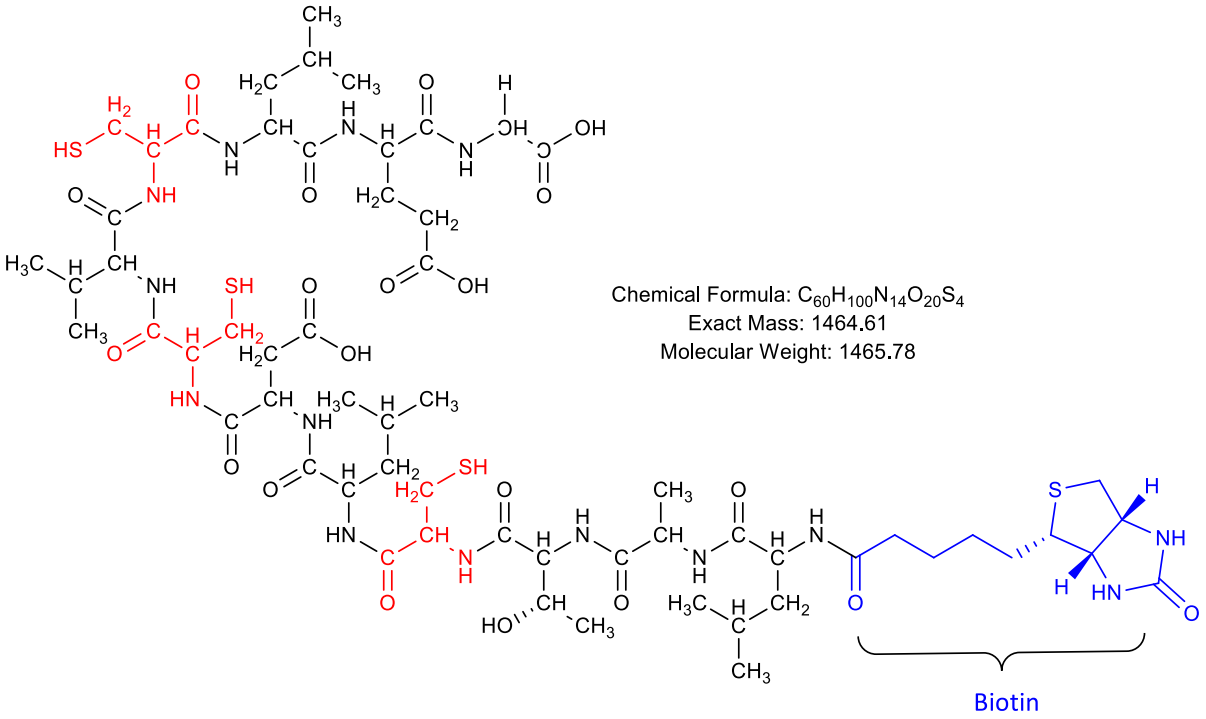


Figure 4.3: Structure of biotinylated Pep1-RJM.

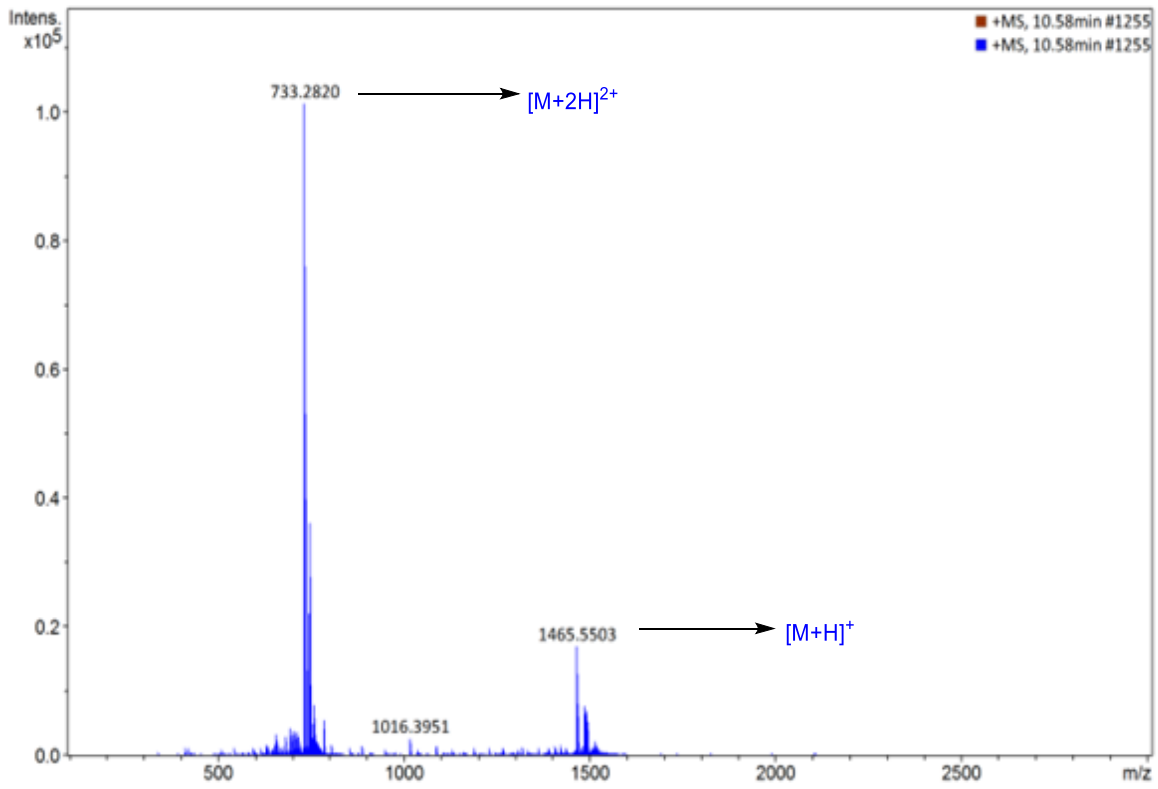


Figure 4.4: Mass spectrum of biotinylated Pep1-RJM.

4.5 Concluding remarks

The quantitative binding analysis of iAs(III), DMA(III), and PAO(III) to thiol-containing Pep1-RJM was successfully conducted using the LC-ESI-MS. The reaction between iAs, DMA, PAO and the ArsR-derived peptide resulted in the formation of As-S covalent bonds. Using the peak areas obtained from EIC, the concentrations of the resultant complexes were determined. Binding constants for condensation reactions can be successfully calculated using the concentrations that were determined from the EIC peak areas. A comparative analysis of the reactions of the different arsenic metabolites (iAs, DMA, and PAO) with thiol reactant (Pep1-RJM) revealed that the designed peptide has greater affinity for PAO(III) and in terms of the studied urinary arsenic metabolites, the peptide showed higher affinity for iAs(III) compared to DMA(III).

References

(1) Matzapetakis, M.; Farrer, B. T.; Weng, T. C.; Hemmingsen, L.; Penner-Hahn, J. E.; Pecoraro, V. L. Comparison of the Binding of Cadmium(II), Mercury(II) and Arsenic(III) to the de Novo Designed Peptides TRI L12C and TRI L16C. *J. Am. Chem. Soc.* **2002**, *124* (27), 8042–8054. https://doi.org/10.1021/ja017520u.

(2) Cline, D. J.; Thorpe, C.; Schneider, J. P. Effects of As(III) Binding on α-Helical Structure. *J. Am. Chem. Soc.* **2003**, *125* (10), 2923–2929. https://doi.org/10.1021/ja0282644.

(3) Schmidt, A. C.; Fahlbusch, B.; Otto, M. Size Exclusion Chromatography Coupled to Electrospray Ionizationmass Spectrometry for Analysis and Quantitative Characterization of Arsenic Interactions with Peptides and Proteins. *J. Mass Spectrom.* **2009**, *44* (6), 898–910. https://doi.org/10.1002/jms.1563.

(4) Carlton, D. D.; Schug, K. A. A Review on the Interrogation of Peptide-Metal Interactions Using Electrospray Ionization-Mass Spectrometry. *Anal. Chim. Acta* **2011**, *686* (1–2), 19–39. https://doi.org/10.1016/j.aca.2010.11.050.

(5) Schmidt, A. C.; Mickein, K. Qualitative and Quantitative Characterization of the Arsenic-Binding Behaviour of Sulfur-Containing Peptides and Proteins by the Coupling of Reversed Phase Liquid Chromatography to Electrospray Ionization Mass Spectrometry. *J. Mass Spectrom.* **2012**, *47* (8), 949–961. https://doi.org/10.1002/jms.3025.

(6) Desilva, B.; Smith, W.; Weiner, R.; Kelley, M.; Smolec, J. M.; Lee, B.; Khan, M.; Tacey, R.; Hill, H.; Celniker, A. Recommendations for the Bioanalytical Method Validation of Ligand-Binding Assays to Support Pharmacokinetic Assessments of Macromolecules. *Pharm. Res.* **2003**, *20* (11), 1885–1900. https://doi.org/10.1023/B:PHAM.0000003390.51761.3d.

(7) Booth, B. P.; Simon, W. C. Analytical Method Validation. *New Drug Dev. Regul. Paradig. Clin. Pharmacol. Biopharm.* **2016**, No. May, 138–159. https://doi.org/10.1201/9780203026427-15.

(8) Deng, L.; Sun, N.; Kitova, E. N.; Klassen, J. S. Direct Quantification of Protein-Metal Ion Affinities by Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2010**, *82* (6),

2170–2174. https://doi.org/10.1021/ac902633d.

Chapter 5 : Conclusion and recommendations

The constant exposure to heavy metals such as inorganic arsenic in the occupational environment remains a great challenge and poses a huge threat to the health of the workers. The development of a novel diagnostic device with high selectivity for arsenic metabolites is necessary for the early detection and pre-diagnosis of arsenicosis for on-site usage. The main aim of this study was to qualitatively and quantitatively investigate the binding affinity of arsenic to sulfur-containing ArsR derived peptides using the LC-ESI-MS. The following chapter presents a summary of the results obtained from this study, the conclusion and recommendations for future work.

5.1 Summary

A solid-phase peptide synthesis method was successfully developed for the synthesis of three ArsR-derived peptides (Pep1-RJM, Pep2-RJM, and Pep3-RJM) using the Fmoc approach. The successful synthesis depended mainly on the coupling reagents. HBTU was proven to be a good coupling reagent for the small-scale synthesis of the peptides but HATU was much more efficient at large-scale and capable of coupling all 12 amino acids. The peptide with good UV absorbance (Pep1-RJM) was purified with ease yielding a peptide purity of 92%. The purification of Pep2-RJM and Pep3-RJM remained a challenge due to unwanted side products and deletion sequences. These peptides also showed poor UV absorbance which resulted in poor purification in preparative HPLC.

The successfully synthesized thiol-containing peptides were used to test for the binding characteristics of the peptides to urinary arsenic metabolites. All three peptides showed positive binding towards arsenic metabolites in their trivalent state. Pre-reduction of As(V) was therefore a crucial step in arsenic-peptide binding studies. The results from the development of a prereduction method of As(V) to As(III) showed that the order of the reducing reagents efficiency was as follows: L-cysteine> DTT> EDT> Peptide≈KI.

To explore the optimum conditions for the binding reactions, the As(III) binding to the Pep1-RJM was conducted using a batch of experiments, however, the binding conditions for Pep2-RJM and Pep3-RJM could not be investigated because the peptides were obtained in low yield

and purity. The obtained binding results for Pep1-RJM showed that the efficiency of the binding reaction was highly dependent on various factors such as reaction time, temperature, arsenic concentrations, initial concentration of the peptide, and pH of the solution. Although the binding mechanism is not yet fully understood, the pH studies have shown that binding is favourable under basic conditions and the highest binding was obtained at pH=8. The binding capacity of Pep1-RJM increased with increasing concentrations of arsenic and this was attributed to the moderate affinity of the pep1-RJM for arsenic. Temperatures higher than 40°C were not favourable for the peptides used in this study. The maximum arsenic binding was 28.5%, 15.9%, and 26.5% for iAs, DMA, and PAO respectively for a peptide at a concentration of 25 µM. The screened peptide also showed great selectivity towards the arsenic metabolites in the presence of other metals such as Nickel (Ni(II), lead (Pb(II), chromium (Cr(II), and copper (Cu(II)).

The LC-ESI-MS method developed in this study allowed for the investigation of the interaction between Pep1-RJM and the arsenic metabolites [iAs(III), DMA(III), and PAO(III)] and also allowed for the determination of apparent binding constants of the condensation reactions. The comparison of equilibrium constants calculated based on EIC peak areas obtained from LC-ESI-MS binding experiments showed the order of relative binding affinities of Peptide-arsenic complexes as Pep1-PAO>Pep1-iAs>Pep1-DMA. The consistency of this trend which was observed throughout this study demonstrates that the LC-ESI-MS is a reliable method and can be used to study peptide-arsenic interactions.

The developed SPPS method was successfully modified for the peptide biotinylation procedure, and the biotinylated peptide was successfully purified and characterized using LC-ESI-MS.

5.2 Conclusion

The novel ArsR-derived peptides were successfully designed, synthesized, and characterized. Liquid chromatography coupled with electrospray ionization mass spectrometry technique was successfully used to develop and optimize a method for studying the interaction of arsenic metabolites with thiol groups of ArsR-derived peptides in the urine matrix.

Pep1-RJM gave better results during synthesis and purification among the three designed and synthesized peptides. Despite challenges encountered for the purification of Pep2-RJM and Pep3-RJM, all three peptides showed an appreciable capacity to bind trivalent arsenic species in the urine matrix. The As(III) binding was found to occur through a condensation reaction resulting in the formation of covalent bonds.

The equilibrium constants obtained from these binding experiments are greatly influenced by (i) the successful optimization of the sample conditions during binding as well as (ii) the interpretation of the obtained mass spectrometry spectra under various conditions.

Pep1-RJM is capable of selectively and efficiently binding arsenic metabolites in urine matrix and the LC-ESI-MS is an efficient technique for the qualitative and quantitative analysis of peptides and peptide-arsenic complexes.

5.3 Future scope of research

Due to time constraints and delays caused by lab construction, the biotinylated Pep1 could not be immobilized onto the gold electrode for practical applications of the designed peptide. Future research will focus on the immobilization of the biotinylated Pep1-RJM on a gold biotin coated electrode and its ability to bind and detect arsenic metabolites in urine will be evaluated. The stability of the peptide-gold nanoparticle conjugates will also be investigated in order to determine the shelf life of the peptide-coated gold electrodes.

The current method has been proven to promote selective binding of arsenic metabolites and the binding constants of peptide-arsenic complexes showed the following trend: PAO(III)>iAs(III)>DMA(III). This maybe the result of structural conformations that the peptide assumes in the urine matrix at pH8 and at a temperature of 40 °C. Once the conformational structure is established through computational and NMR studies, the knowledge can be used to design peptidomimetics with shorter, affordable sequences and appropriate conformations. The information can further be used in the design of chelating drugs and masks additives.

Different peptide-arsenic complexes will be studied using computational studies. This includes determining the minimized structures of each formed complex and the bond length

between each thiol group and the respective arsenic metabolites in the Peptide-arsenic complexes. The computational analysis of the peptide with the various arsenicals will provide a better understanding of the peptide folding and the weak, but cumulative non-covalent bonds formed with the arsenicals. This study will enable us to have a better understanding of the relationship between the structure and function of these peptides. The information obtained from this study can further assist to determine the binding mechanism of the condensation reaction that occurs between the peptide and arsenic species.

The successful determination of the binding mechanism and conformational structure will lay a foundation for designing analyte-specific peptides and consequently, increasing the sensitivity of the designed biosensor.

Structural modifications of Pep2-RJM and Pep3-RJM by adding amino acids such as tryptophan, tyrosine, or phenylalanine to improve UV absorbance and purification should be implemented. The binding characteristics of these peptides to urinary arsenic metabolites can then be explored under the same conditions as Pep1-RJM. Structural modifications of the peptides can also be implemented where cysteine residues are replaced with another sulfur-containing amino acid such as methionine.

The comparison of binding characteristics and binding affinities of the three ArsR-derived peptides will be useful in selecting a better performing bio-receptor for the sensor.

5.4 Research contribution

The work presented in this dissertation (literature review, method development, and research findings) contributes to the overall body of scientific research and can potentially open up new research avenues in the future. Our first contribution is an analysis of the literature regarding the various methods and techniques that are used in the detection of arsenic metabolites including their limitations. Among the available materials used for arsenic detection, peptide-based bio-receptors have not been fully explored for their ability to bind and detect arsenic metabolites in urine. There are currently no studies that have utilized the combination of materials and methods used in this study (ArsR-derived peptides, SPPS, urinary arsenic metabolites, and the LC-ESI-MS). A proven set of techniques presented in this dissertation can be used by more researchers.

Chapter 6 : Experimental

This chapter provides a list of reagents and a description of the major equipment and instrumentation that was used to conduct this research.

6.1 Reagents

All reagents and solvents were obtained from commercial suppliers. Fmoc protected amino acids, 2-chlorotrityl resin , N,N,N',N'-Tetramethyl-O-(1H-benzotriazol-1-yl)uranium hexafluorophosphate (HBTU), Hexafluorophosphate Azabenzotriazole Tetramethyl Uronium (HATU), Diisopropylethylamine (DIPEA), piperidine, thionyl chloride ($SOCl_2$), Triisopropylsilane (TIS) were purchased from DLD Scientific (Durban, South Africa). Arsenic acid sodium (iAs(V)), Cacodylic acid DMA(V)), DL-Dithiothreitol, Diisopropylethylamine (DIPEA), 1,2-ethanedithiol (EDT), formic acid, hydrochloric acid (HCl), sodium hydroxide (NaOH), phenylarsine oxide (PAO(III)), piperidine, sigmatrix urine diluent, sodium M-Arsenite(iAs(III)), triisopropylsilane (TIS) and trifluoroacetic acid (TFA) were purchased from Sigma-Aldrich (South Africa). Organic solvents such as acetonitrile (ACN), dimethyl sulfoxide (DMSO), methanol (MeOH) and diethyl ether were supplied by Pyramid Scientific (Johannesburg, South Africa). and Radchem (Johannesburg, South Africa).

6.2 Instrumentation

Purification of synthetic peptides was carried out on an Agilent 1260 Infinity semi-preparative HPLC instrument with a UV/VIS detector and an automated fraction collector. The semi-preparative HPLC system is equipped with a binary solvent system where solvent A consisted of H_2O and 0.1% Formic acid (v/v), and solvent B consisted of Acetonitrile and 0.1% Formic acid (v/v). 300µL of the crude samples was injected onto a Kinetix® 5 µm B C18, 100 Å (250 ×230 nm) column. The system flow rate and UV detector were set at 15-20 mL/min and wavelengths of 190, 195, 215, 254, and 300 nm.

Analytical characterization was performed on an Ultra-High-Performance Liquid Chromatography (Thermo Scientific Ultimate 3000, RS diode array detectors) with a Diode

Array (190, 195, 215, 254, and 300 nm) coupled to a Bruker Compact Q-TOF high-resolution mass spectrometer (Figure 6.1). The samples were analysed using a binary solvent system where solvent A consisted of H_2O and 0.1% Formic acid (v/v), and solvent B consisted of Acetonitrile and 0.1% Formic acid (v/v). 20 µL of the sample was injected onto a C18 column (5 µm, 100 Å, 4.60 mm × 150 mm). The system flow rate was set at 0.3 mL/min in the positive mode. The mass spectrum analysis was processed using the Bruker Daltonics data analysis software.

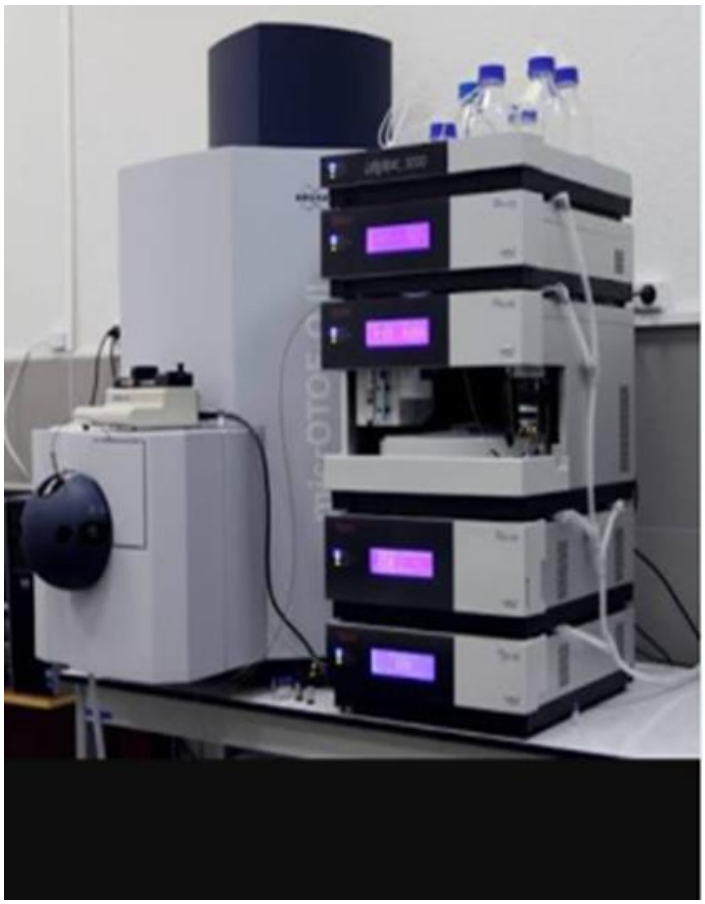


Figure 6.1: Ultra-High-Performance Liquid Chromatography (Thermo Scientific Ultimate 3000, RS diode array detectors) with a Diode Array (190, 195, 215, 254, and 300 nm) coupled to a Bruker Compact Q-TOF high-resolution mass spectrometer.

The ultra-pure water used in this study was obtained from the Millipore Direct-Q®3 UV water purification (ZRQSVPO30) system.

Binding studies were conducted using the Carousel 6 plus reaction station which can simultaneously heat and stir samples.



Figure 6.6.2: Carousel 6 plus reaction station

All pH studies presented in this dissertation were conducted using a pH meter from Peak instruments Inc. (T-710L).



Figure 6.6.3: Peak instruments pH meter used in As(III) binding studies

6.3 Activation of 2-chlorotrityl resin

2-chorotrityl resin (1 g) was swelled and activated in a mixture of dry DCM and $SOCl_2$ (10:1 (v/v), 11ml) in a 15 mL centrifuge tube and left to shake overnight. The solution was transferred to a 70 mL glass reaction vessel with a fritted filter. DCM was removed by suction and the resin was washed with DCM (5 mL) three times with mixing by bubbling inert nitrogen gas (N_2) for 30 seconds per wash.

6.4 Coupling of the first amino acid

The Fmoc-protected amino acid (2.63 mmol, 0.2 M) was dissolved in dry DCM (9.6 mL) and DIPEA (10 mmol, 1M (3.4 mL)) was added to the centrifuge tube. The mixture was added to the reaction vessel containing the 2-CTC resin and mixed by gently bubbling with N_2 for 45 minutes. The solvent was filtered at the end of the coupling cycle and the resin was washed with (3 x 5 mL) DCM. The coupling step was repeated to ensure adequate loading of the first amino acid on the resin.

6.5 General procedure for determining loading capacity

Resin beads coupled to the first amino acid were dried under vacuum for 60 minutes. Approximately 5 to 10 mg of resin beads were weighed and transferred into two 2 ml Eppendorf tubes. A solution of 20% Piperidine in DMF (1 mL) was added to each tube containing the resin beads. The sample tubes were then shaken in a shaker at 90 revolutions per minute (rpm) for 20 minutes. The samples were then centrifuged for 2 minutes and 100 µL of the supernatant was transferred into DMF (10 ml) in a 15 ml centrifuge tube. Using an ultraviolet/visible spectrometer (UV/VIS), 2mL of blank DMF was analysed to zero the instrument before analysing the samples for absorbance at wavelength 301 nm. The analysis was conducted in duplicates and the average was used to determine the loading capacity using the formula provided below.

$$Loading\ capacity\left(\frac{mmol}{g}\right) of\ resin = \frac{(101 \times\ (A))}{(7.8\ \times\ (W))}$$

Where; A= Average absorbance and W=mg of resin.

Table 6.1: Absorbance values used to determine the loading capacity of the 2-CTC resin.

Sample No.	Mass of resin (mg) (W)	Absorbance	Average absorbance (A)
1	9.6	A_1= 0.3520	0.3650
		A_2= 0.3780	
2	9.8	A_1= 0.2295	0.2405
		A_2= 0.2514	

6.6 Coupling of the 2nd-9th amino acid

The first amino acid was then deprotected with a (2×10 ml) 20% piperidine solution in DMF (v/v) the reaction was mixed by bubbling N_2 gas for 10 minutes. After deprotection, the resin was washed with (3×5 mL) DMF. The moles of the subsequent amino acids were determined from the loading capacity. Single coupling was conducted for the subsequent amino acids at 45 min per coupling cycle.

Fmoc protected amino acid (2.63 mmol, 0.2 M) and HBTU coupling reagent (2.50 mmol, 0.19 M) were dissolved in a solution of DIPEA (10 mmol, 1M) in DMF (13mL). The solution was transferred to the reaction vessel with the resin and the reaction was mixed by gently bubbling N_2 gas for 45 minutes. At the end of 45 minutes, the solvent was removed through suction and the resin was washed with 3 x 5 mL DMF. The cycle of coupling and deprotection was repeated until the 9^{th} AA.

6.7 Coupling of the 10th -12th amino acid

The Fmoc group of the 9th amino acid was then deprotected with a (2×10 ml) 20% piperidine solution in DMF (v/v) for 10 minutes, the reaction was mixed by gently bubbling N2 gas. After deprotection, the resin was washed with (3×5 mL) DMF. Fmoc protected amino acid (2.63 mmol, 0.2 M) and HATU coupling reagent (2.50 mmol, 0.19 M) were dissolved in a solution of

DIPEA (10 mmol, 1M) in DMF (13 mL). The cycle of coupling and deprotection was repeated until the 12th amino acid. After the last coupling, the Fmoc was deprotected for 10 minutes three times (3 X 10 minutes)using the 20% piperidine solution in DMF (10ml). The resin-bound peptide was washed using (3 x 5 mL) DMF and (3 x 5 mL) DCM and then dried under vacuum for 60 minutes. The peptides were cleaved using the general cleavage procedure outlined in section 6.8 below. The peptides were either dissolved in methanol, acetonitrile or a combination (1:1, (v/v)) and analysed using the LC-MS. Table 6.2 below presents a summary of the LC-MS results of the synthesized peptides.

Table 6.2: LCMS results for ArsR derived peptides.

Peptide ID	Peptide sequence	Formula	Theoretical [M]	Detected $[M+H]^+$	Detected $[M+2H]^{2+}$	R_t/min
Pep1-RJM	GEL CVC DLC TAL	$C_{50}H_{86}N_{12}O_{18}S_3$	1238.53	1239.4910	620.2628	6.40
Pep2-RJM	CVC SGS SKA VCI	$C_{45}H_{81}N_{13}O_{16}S_3$	1155.51	**	578.7493	3.11
Pep3-RJM	CVC SGD SKN ICS	$C_{45}H_{78}N_{14}O_{19}S_3$	1214.47	1215.2454	608.3062	6.73

6.8 General procedure for resin cleavage from peptide

The resin-bound peptide was thoroughly washed with (3 x 5 mL) DMF and (3 x 5 mL) DCM after Fmoc deprotection. The resin was then dried under vacuum for 90 minutes. The dried resin was subjected to a 10 mL cleavage cocktail containing TFA, H_2O, EDT, and TIS (9.4: 2.5: 2.5: 1 (v/v)) for 2hrs in a 50 mL centrifuge tube. The filtrate was divided into two equal portions and transferred into 50 ml centrifuge tubes. Cold ether was then added (5-10 times more than the sample volume) to the filtered solution to yield a white precipitate of the crude peptide which was centrifuged for 10 minutes at 5000 rpm. The solvent was then decanted leaving the crude peptide at the bottom of the centrifuge tube. The process of washing with cold ether was repeated twice.

6.9 Biotinylation of Pep1-RJM

Pep1-RJM was synthesised using procedures 6.3 - 6.7 above. The resin-bound Pep1-RJM was Fmoc deprotected using 20% piperidine in DMF (2 x 10 mL). Biotin (2.63 mmol, 0.2 M) was

dissolved in DMSO and coupled to Pep1-RJM using HATU (2.50 mmol, 0.19 M)) and N-methylmorpholine (NMM) (10 mmol, 1M) in DMF for 45 minutes. The coupling reaction was repeated. The biotinylated Pep1 (Pep1-Bio) was cleaved from the resin for 2 hrs using the TFA cleaving cocktail described in procedure 6.8 above. The crude Pep1-bio was analysed using the LC-ESI-MS and purified using a semi-preparative HPLC.

Observed m/z of 1465.5503 and 733.2820 of the biotin-bound Pep1-RJM corresponding to $[M+H]^+$ and $[M+2H]^{2+}$ expected, Rt.10.58 min

6.10 Disulfide bond reduction

In a 5 mL volumetric flask, a stock solution of DTT (6.48 mM, 0.03 mmol) was freshly prepared. From the stock solution, a concentration of 0.3 nM (1.5 pmol,5 equiv, 5 mL) was prepared and added to sample vial 1 containing Pep1-RJM (403.70 nmol, 80.74 µM), sample vial 2 containing Pep2-RJM (432 nmol, 86.54 µM), and sample vial 3 containing Pep3-RJM (411 nmol, 82.34 µM) all three vials were incubated for 20 minutes.

6.11 As(V) to As(III) pre-reduction

6.11.1 Reduction by KI

A stock solution of DMA(III) (7.25 mM, 1.45 mols) was freshly prepared in a 5 mL volumetric flask. DMA(V) (100 µM , (97 µL)) was added into a clean amber bottle followed by 5 mL of 0.1 M HCl and 2mL of KI (6.02 µM). $SnCl_2$ (8.40 µM, mols, 0.5mL) was added with thorough mixing of the reaction. The reaction was then left in the dark for 15, 20, 30, 40, and 60 minutes for the reduction of arsenic pentavalent to the trivalent state.

6.11.2 Reduction by peptide

The ability of Pep1-RJM, Pep2-RJM, and Pep3-RJM to reduce DMA(V) was tested by adding DMA(V) (0.15 mol, 724 µM) into different round-bottomed flasks containing Pep1-RJM, Pep2-RJM, and Pep3-RJM (100 µM, 0.5 µmol). The reactions were also tested with two-fold and five-fold molar excess of the peptides.

6.11.3 Reduction by DTT, EDT, L-cysteine

L-Cysteine

L-cysteine (125 mg) was dissolved in 15 mL of Millipore water and added to a 25 mL volumetric flask. 2.5 mL of 0.1M HCl was added followed by 2.5mL of water. DMA(V) solution (14.49 mM, 0.36 mmol) was added to the volumetric flask and the solution was diluted to the mark. The solution was left to stand for 20 minutes before the peptide-binding studies.

DTT

DTT (125 mg) was dissolved in 15 mL of Millipore water and added to a 25 mL volumetric flask. 2.5 mL of 0.1M HCl was added followed by 2.5mL of water. DMA(V) solution (14.49 mM, 0.36 mmol) was added to the volumetric flask and the solution was diluted to the mark. The solution was left to stand for 20 minutes before the peptide-binding studies.

EDT

EDT (2 mL) was added to 13 mL of Millipore water in a 25 mL volumetric flask. 2.5 mL of 0.1M HCl was added followed by 2.5mL of water. DMA(V) solution (14.49 mM, 0.36 mmol) was added to the volumetric flask and the solution was diluted to the mark. The solution was shaken vigorously and left on a shaker for 20 minutes before the peptide-binding studies.

6.12 Procedure for the binding of Pep1-RJM to iAs(III), DMA(III), and PAO(III)

Fresh stock solutions of iAs(III) (7.70 mM, 0.078mmols), DMA(7.25 mM, 0.072mmols), and PAO(III) (5.95 mM, 0.06mmols) were prepared in water in a 10 mL volumetric flasks. Reduced Pep1-RJM (20 µM, 100 nmols) was added into seven different round-bottomed flasks each flask contained synthetic urine which had been spiked with either 0, 10, 20, 40, 60, 80, 100 µM initial arsenic concentration. The flasks were incubated for 20 minutes. The binding ability of Pep1-RJM was investigated under various pH conditions, reaction temperature, initial arsenic(III) concentration, contact time, and initial peptide concentration.

6.12.1 Effect of solution pH

The effect of solution pH was monitored by changing the initial pH of the solution from 2 to 14. The pH was adjusted using 0.1M HCl or 0.1 M NaOH and was measured using the Peak instruments pH meter. The initial Pep1-RJM concentration was fixed at 20 µM and the reaction was conducted at room temperature (25°C) and pH 8.

6.12.2 Effect of reaction temperature

The effect of the reaction temperature on the As(III) binding was investigated by incubating the solutions at different binding temperatures at 25, 30, 35, 40, 45, 50, 55 °C using the Carousel 6 plus reaction station temperature control system while keeping other parameters such as the Pep1-RJM concentration at 20 µM, the reaction pH at 8 and the initial As(III) concentration at 100 µM.

6.12.3 Effect of reaction time

The effect of the reaction time on the As(III) binding was examined by incubating the Pep1 and As(III) mixture at various reaction times from 0 to 30 minutes and the reaction was monitored at 5-minute intervals. In this experiment the concentration of Pep1-RJM was kept constant at 20 µM, the initial concentration of arsenic was 100 µM, and the pH of the reaction was kept at 8. The reaction was conducted at 40°C.

6.12.4 Effect of initial arsenic concentration

To study the effect of initial arsenic concentration, 20 µM of Pep1-RJM was spiked with different concentrations (0, 10, 20, 40, 60, 80, 100 µM) of each arsenic metabolite. The reaction was incubated for 15 minutes at 40°C and the pH was maintained at 8.

6.12.5 Effect of Pep1-RJM concentration

The effect of Pep1-RJM concentration on arsenic binding was investigated by varying the concentration of Pep1-RJM (0, 5, 10, 15, 20, 25, 30, 35, 40) µM. The reaction was incubated for 15 minutes at 40°C and the pH was maintained at 8.

6.12.6 Binding affinity by titration method

The methodology for determining the binding constant was similar to the one used for the above batch experiments except that the optimum binding conditions presented in Table 6.3 were used for the titration experiments.

Table 6.3: Optimum As(III) binding conditions.

Biomolecule	pH	Temp (˚C)	Reaction time (min)	iAs(III) conc (µM)	Pep1-RJM conc (µM)
Pep1-RJM	8	40	15	60	25

During the titration experiments, a series of samples were prepared with a constant concentration of the peptide whilst varying the concentrations of the arsenic species from 10-60 µM using the 6-vessel reactor and the formation of the peptide-As(III) complex was monitored on the LC-ESI-MS as the concentration of the reactants decreased.

APPENDICES

Appendix I- Supplementary information for Chapter 2

Appendix II- Supplementary information for Chapter 3

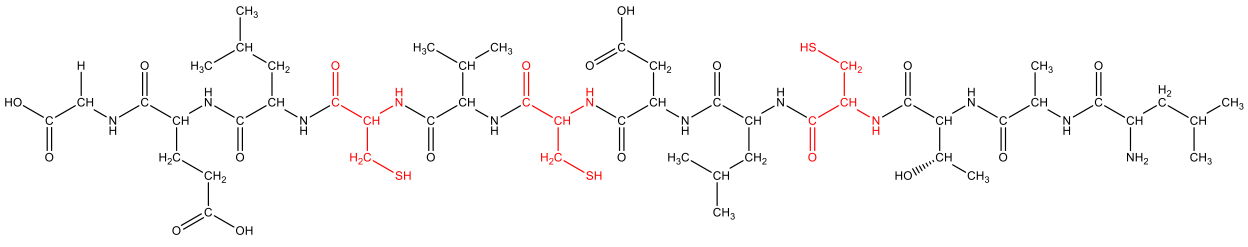
Appendix III- Supplementary information for Chapter 4

Appendix I- Supplementary information for Chapter 2

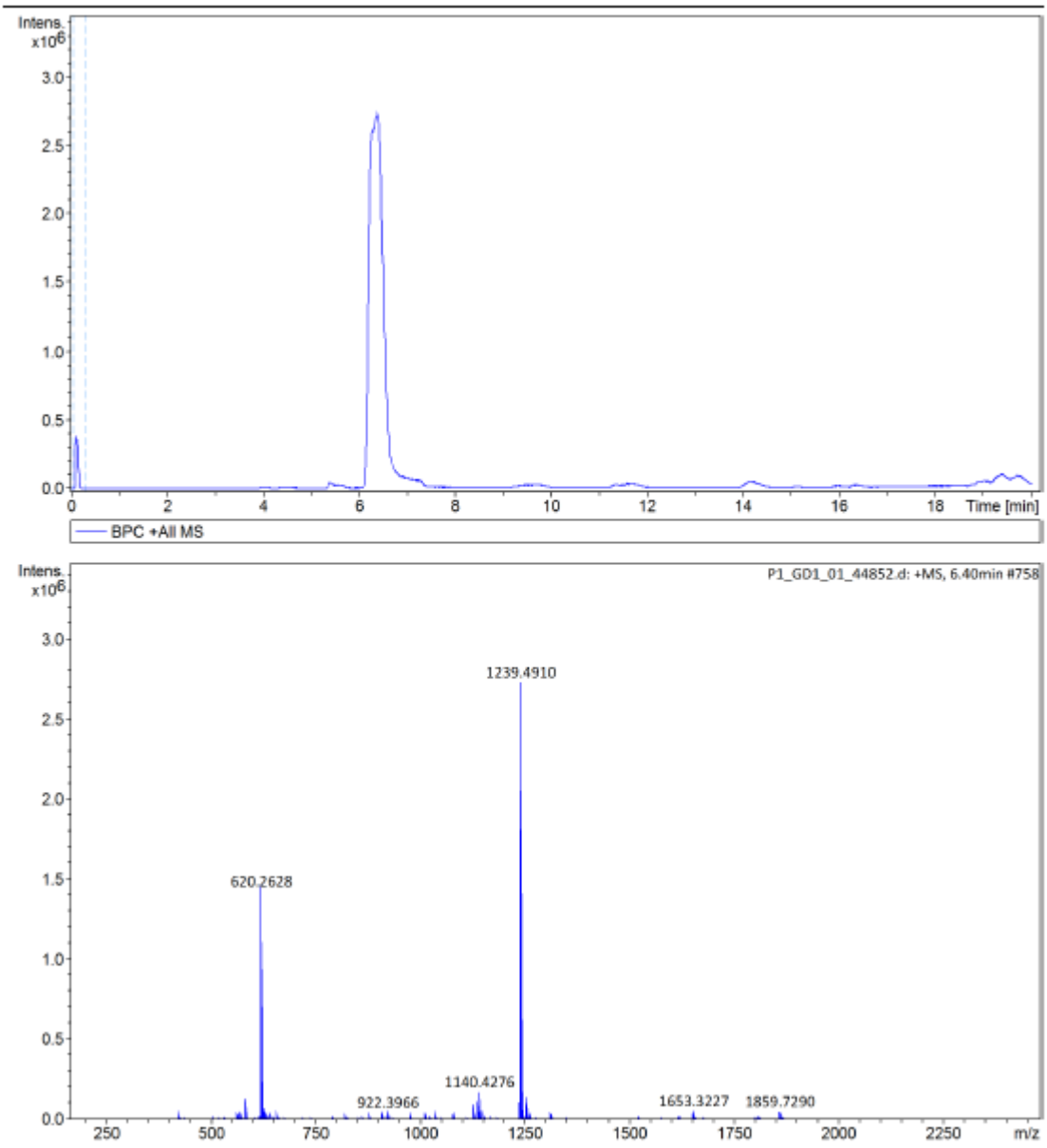
SD-1: Table 0.1:Absorbance values used for determining loading capacity of 2-CTC.

Sample No.	Mass of resin (mg) (W)	Absorbance	Average absorbance (A)
1	9.6	A_1= 0.3520	0.3650
		A_2= 0.3780	
2	9.8	A_1= 0.2295	0.2405
		A_2= 0.2514	

Pep1-RJM sequence: GEL CVC DLC TAL

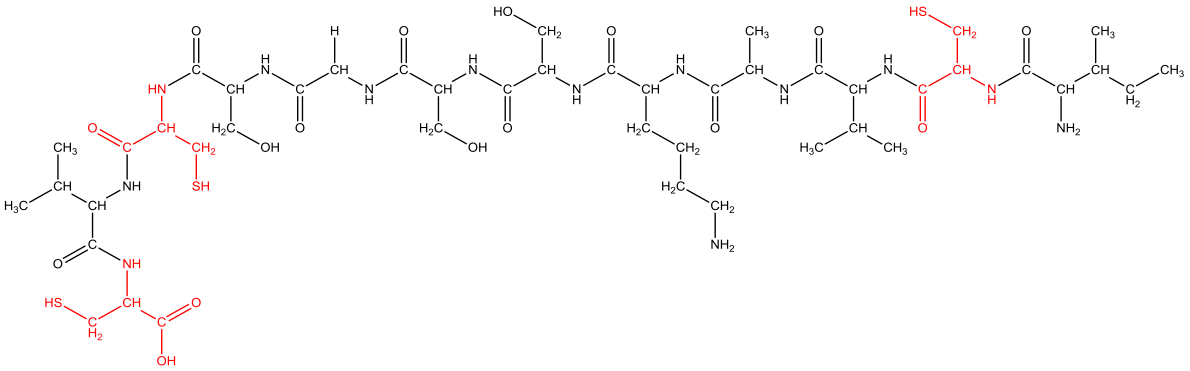


Chemical Formula: $C_{50}H_{86}N_{12}O_{18}S_3$
Exact Mass: 1238,53
Molecular Weight: 1239,48

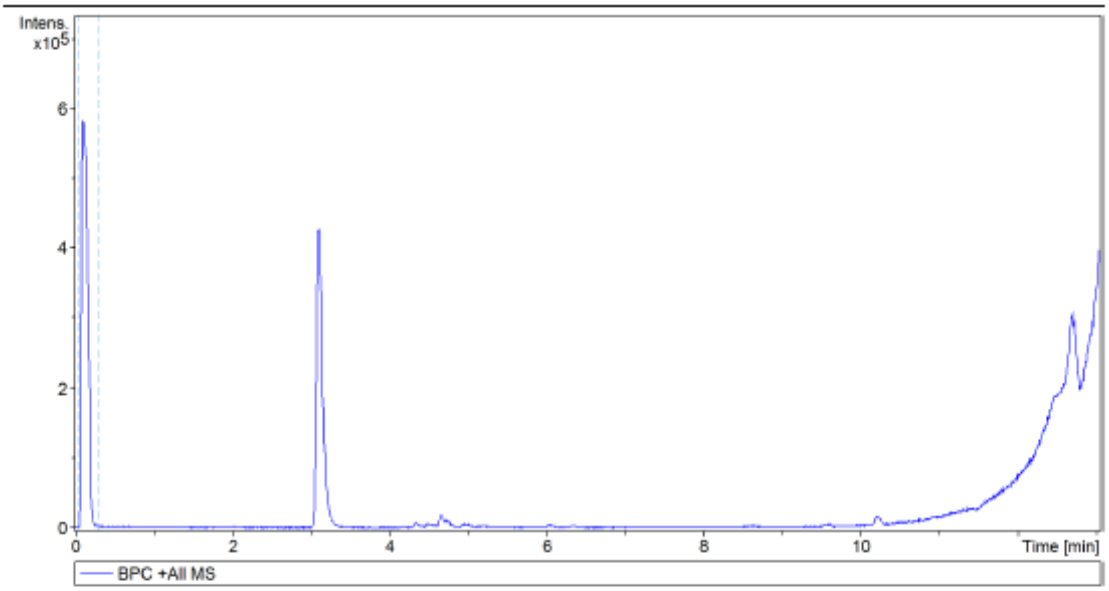


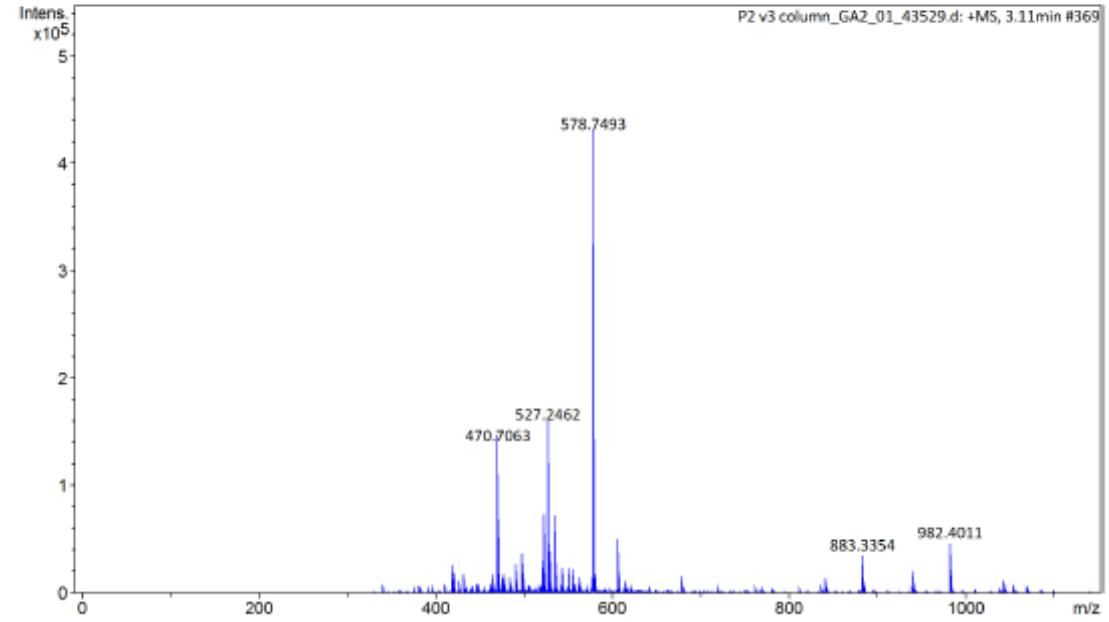
SD-2: Figure 0.1: Mass spectra of Pep1-RJM with sequence GELCVCDLCTAL-NH2.

Pep2-RJM sequence: CVC SGS SKA VCI

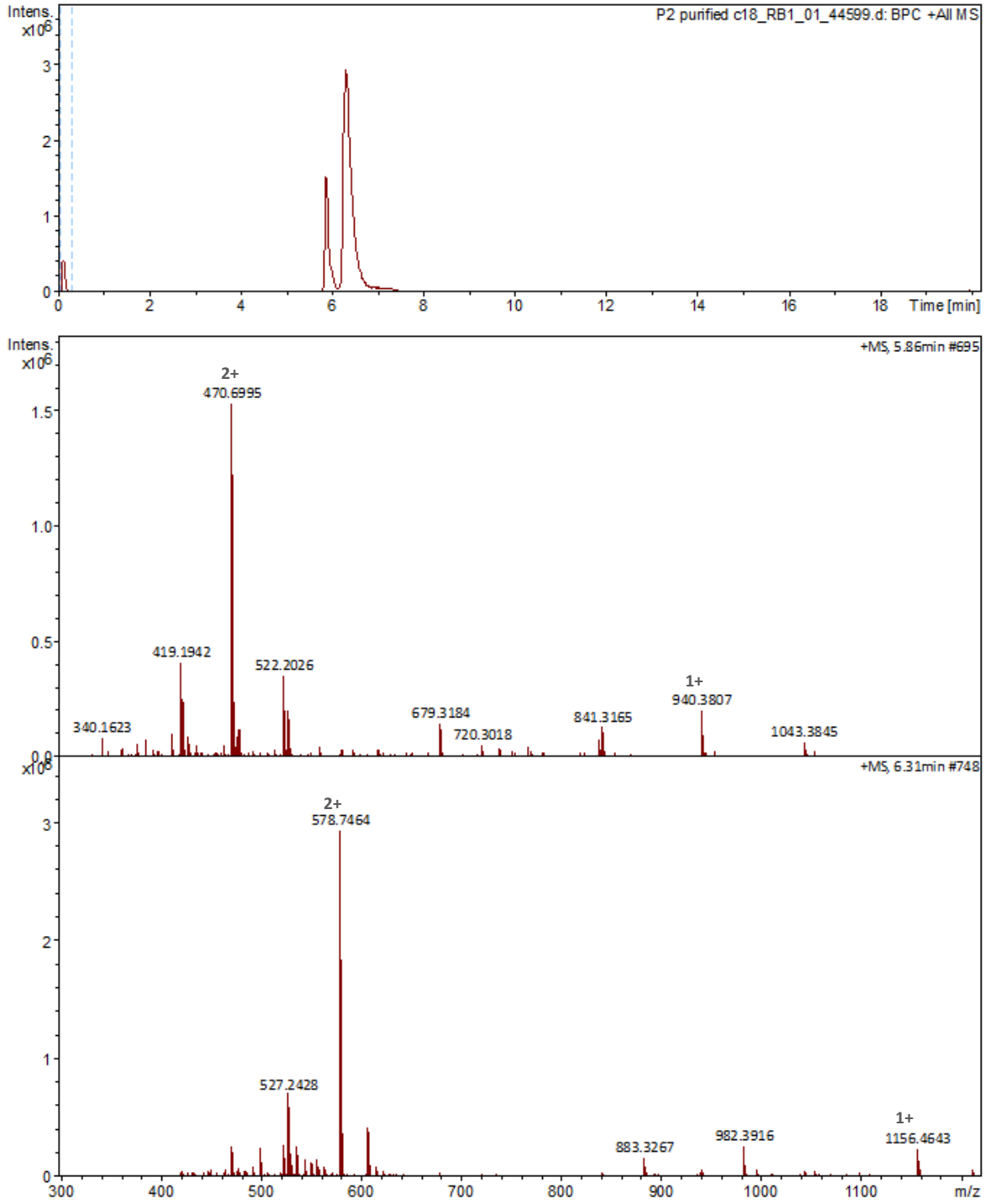


Chemical Formula: $C_{45}H_{81}N_{13}O_{16}S_3$
Exact Mass: 1155,51
Molecular Weight: 1156,40

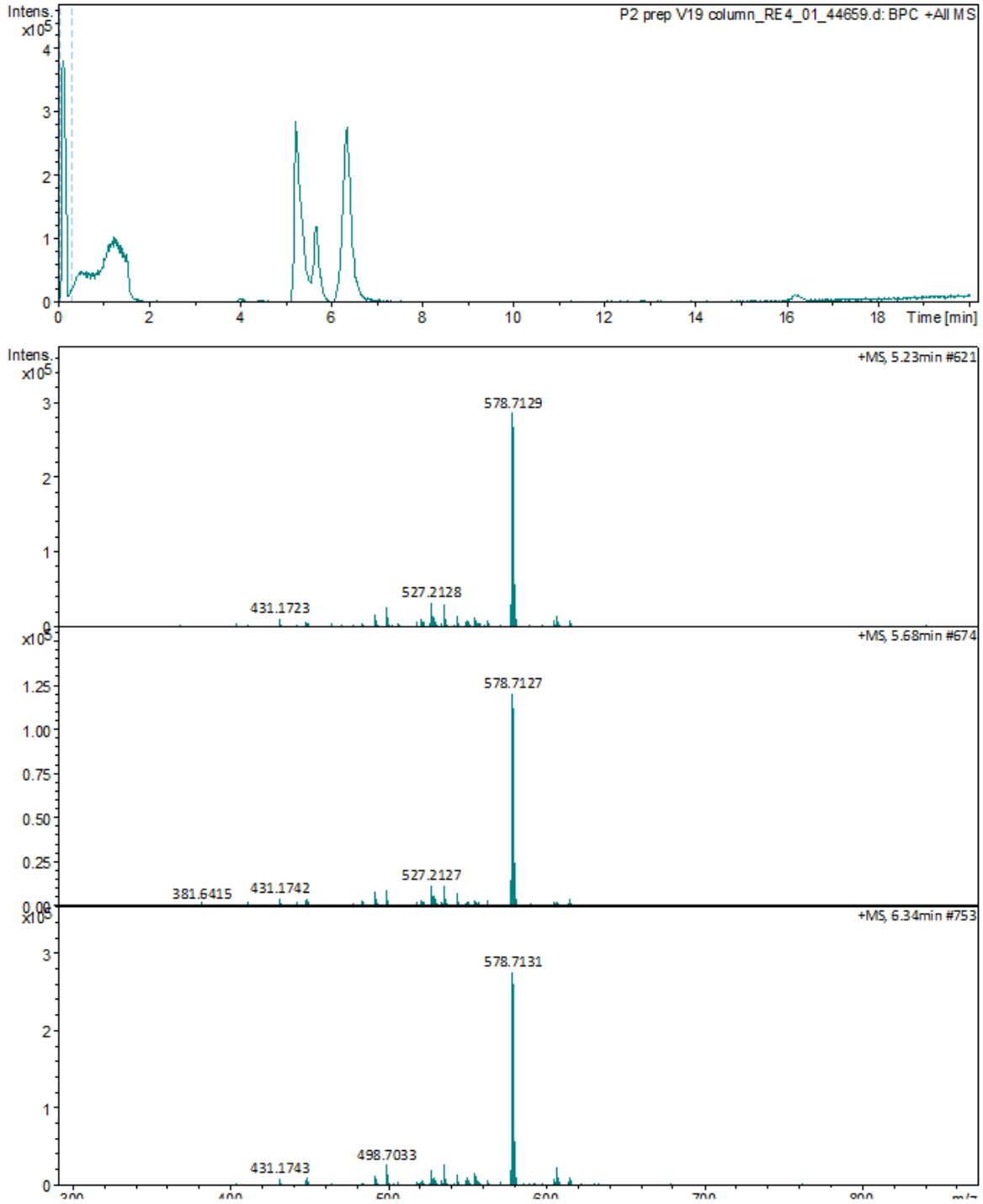




SD-3:Figure 0.2: Mass spectra of Pep2-RJM with sequence CVCSGSSKAVCI-NH2.

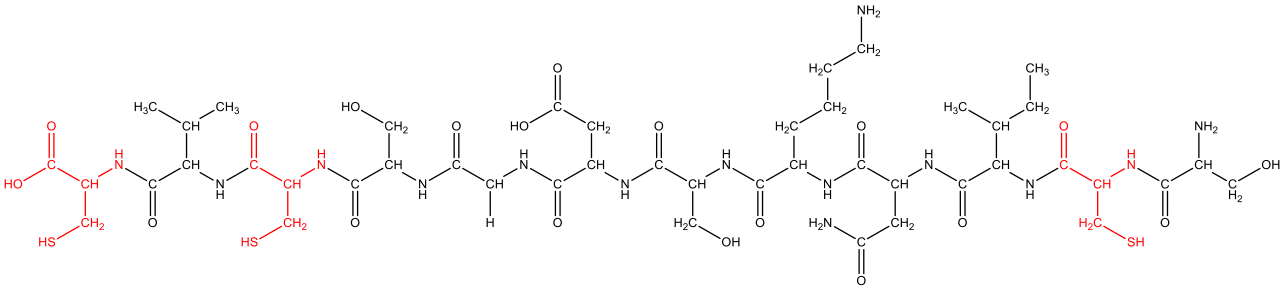


SD-4: Figure 0.3: HPLC-MS chromatogram of Pep2-RJM with impurity.

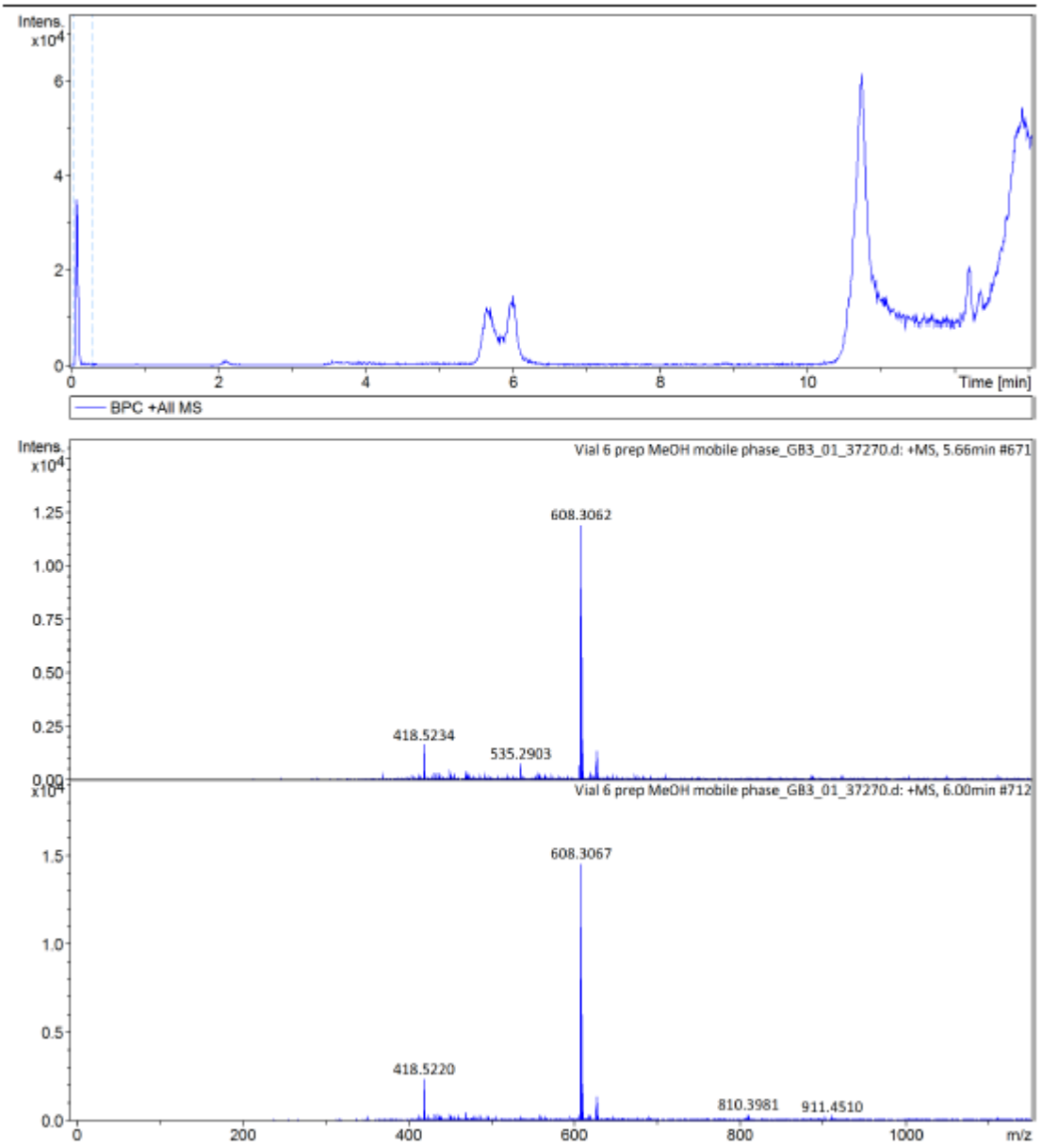


SD-5: Figure 0.4: HPLC-MS chromatogram of purified Pep2-RJM.

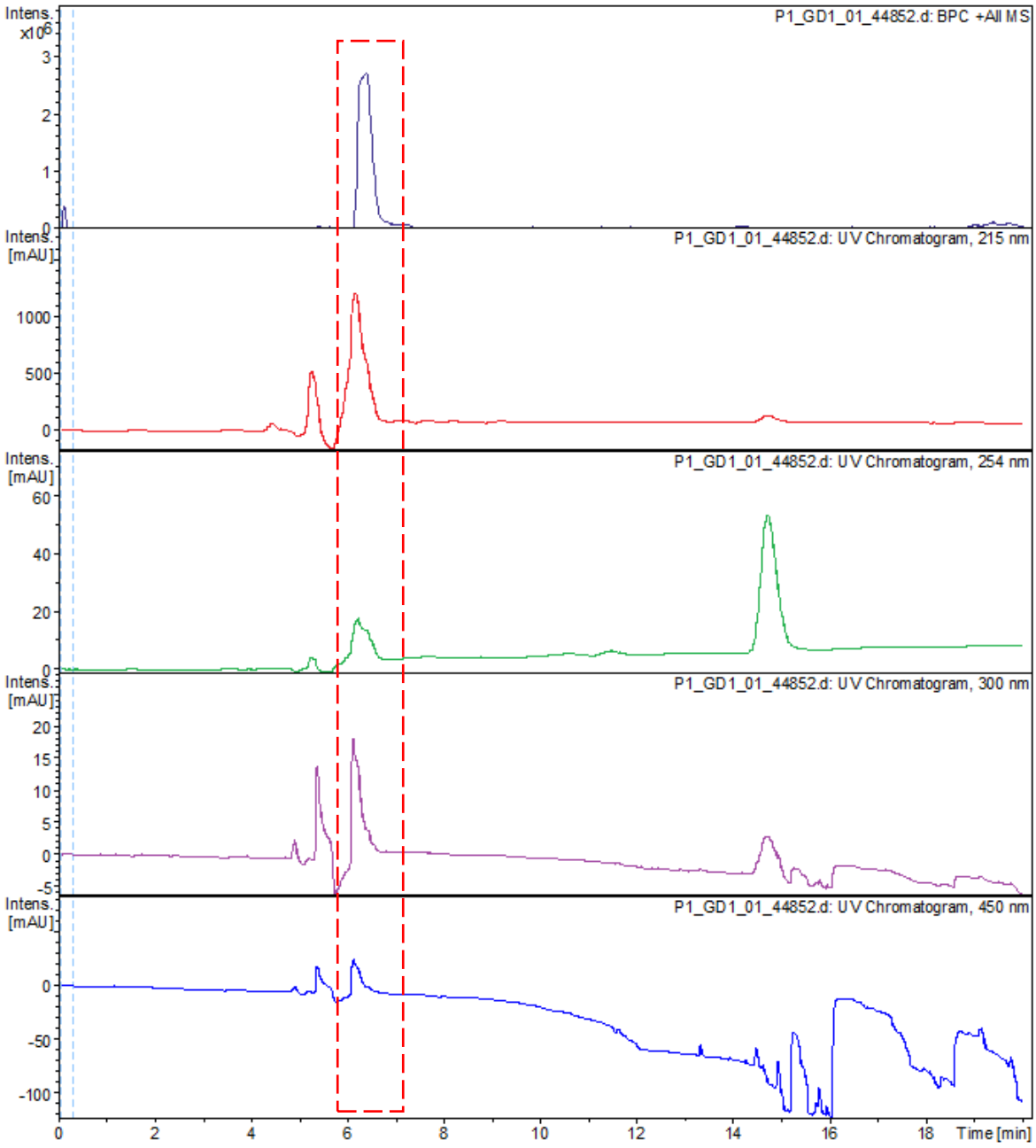
Pep3-RJM sequence: CVC SGD SKN ICS



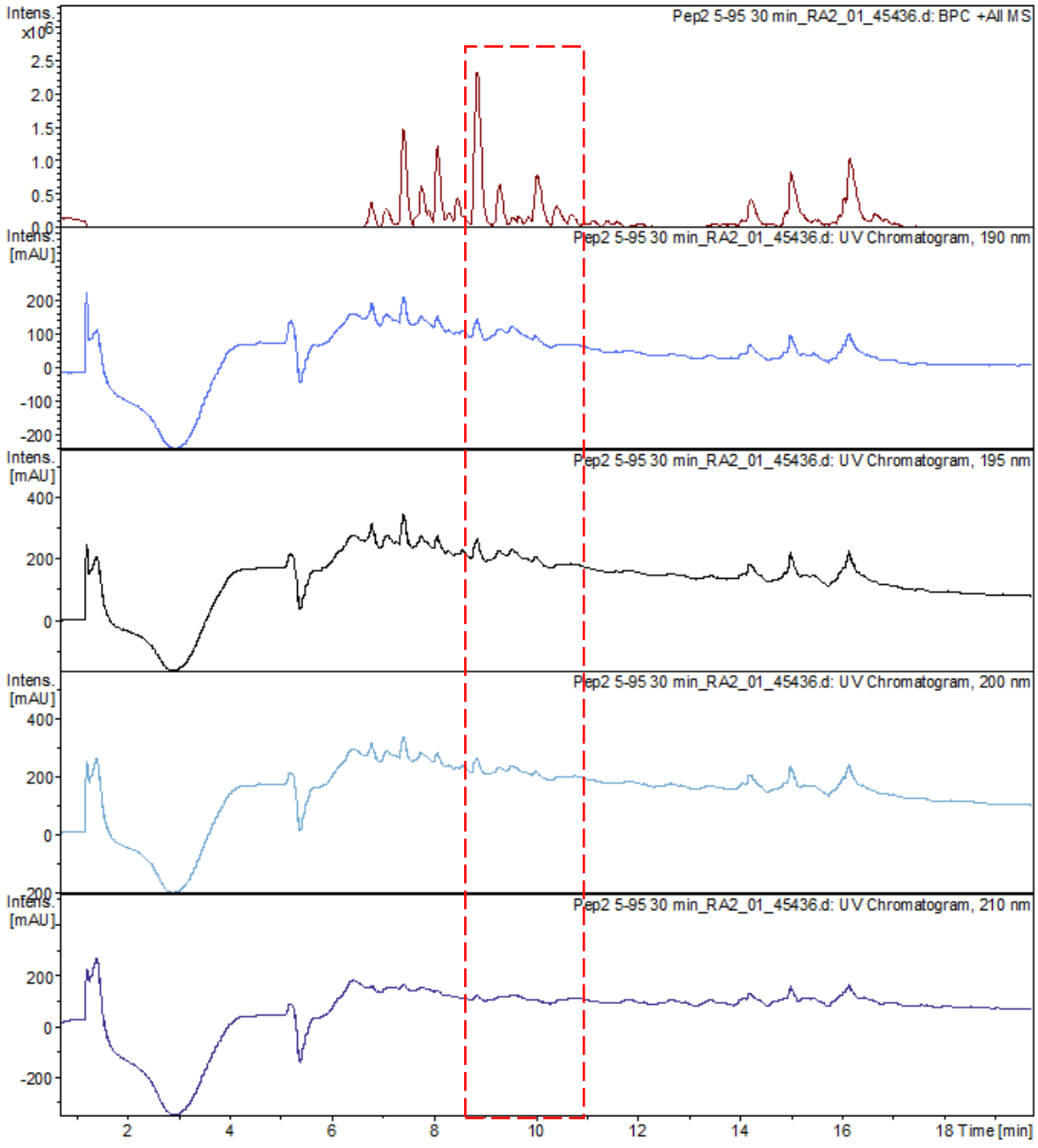
Chemical Formula: $C_{45}H_{78}N_{14}O_{19}S_3$
Exact Mass: 1214,47
Molecular Weight: 1215,38



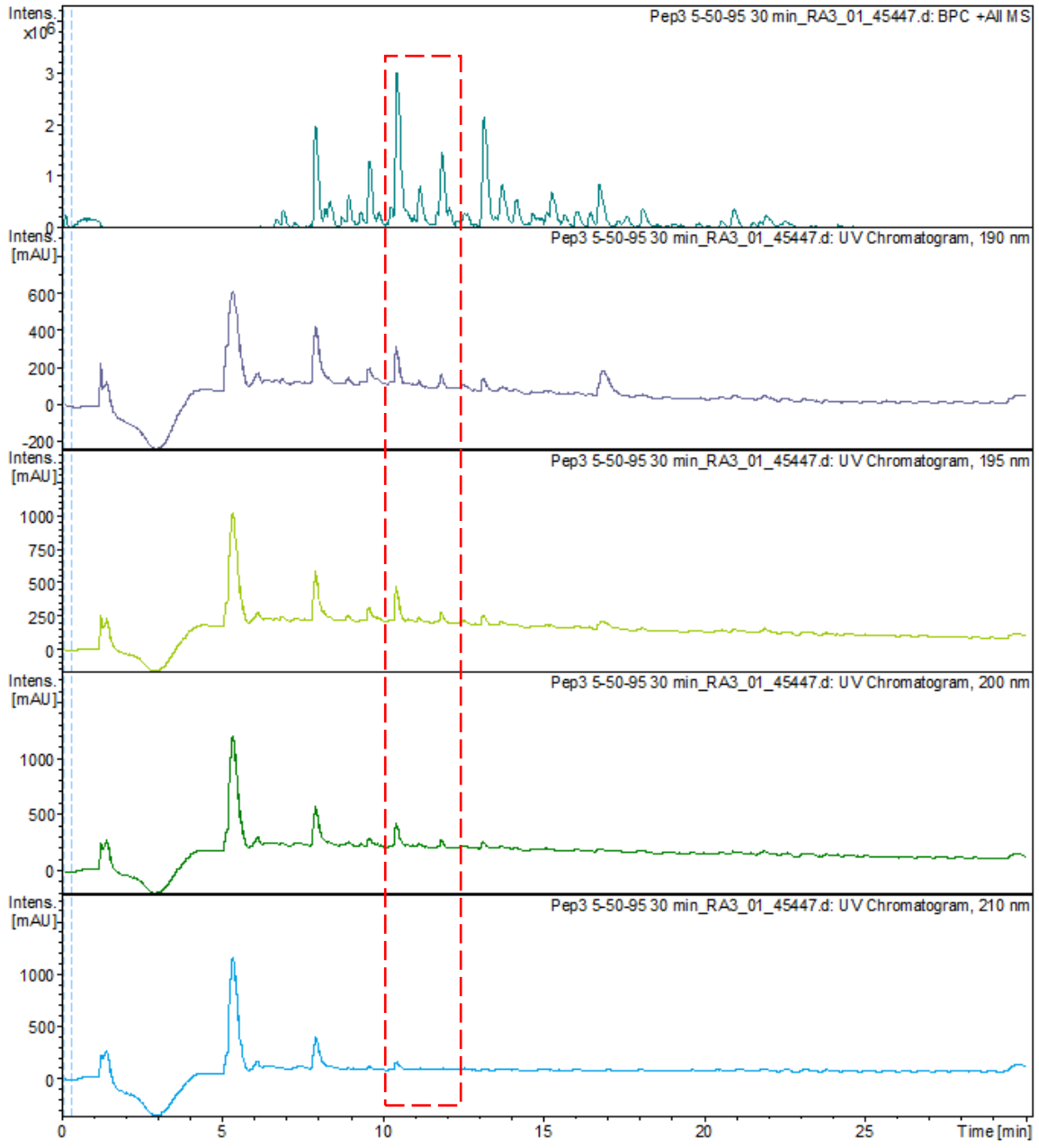
SD-6: Figure 0.5: Mass spectra of Pep3-RJM with sequence CVCSGDSKNICS-NH_2.



SD-7: Figure 0.6: Chromatograms of Pep1_RJM after purification. A) Base peak chromatogram, b) UV spectra 215 nm, c) UV spectra 254 nm, d) UV spectra 300 nm and e) UV spectra 450 nm.



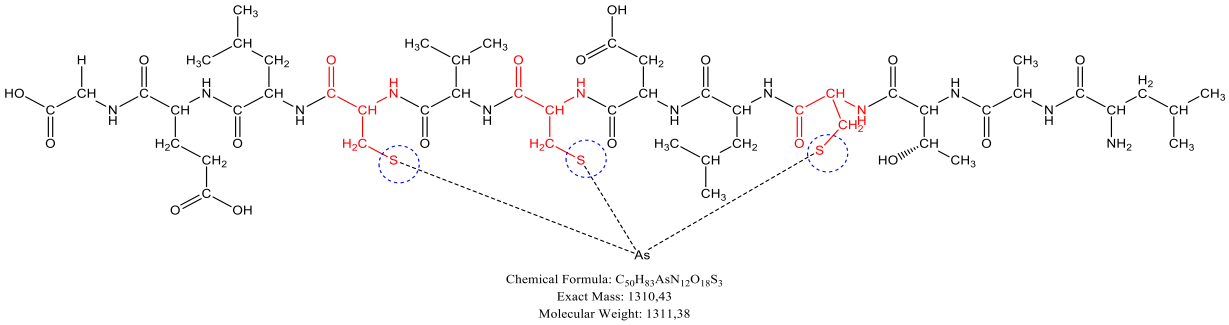
SD-8: Figure 0.7: Chromatograms of crude Pep2_RJM before purification. A) Base peak chromatogram, b) UV spectra 190 nm, c) UV spectra 195 nm, d) UV spectra 200 nm and e) UV spectra 200 nm

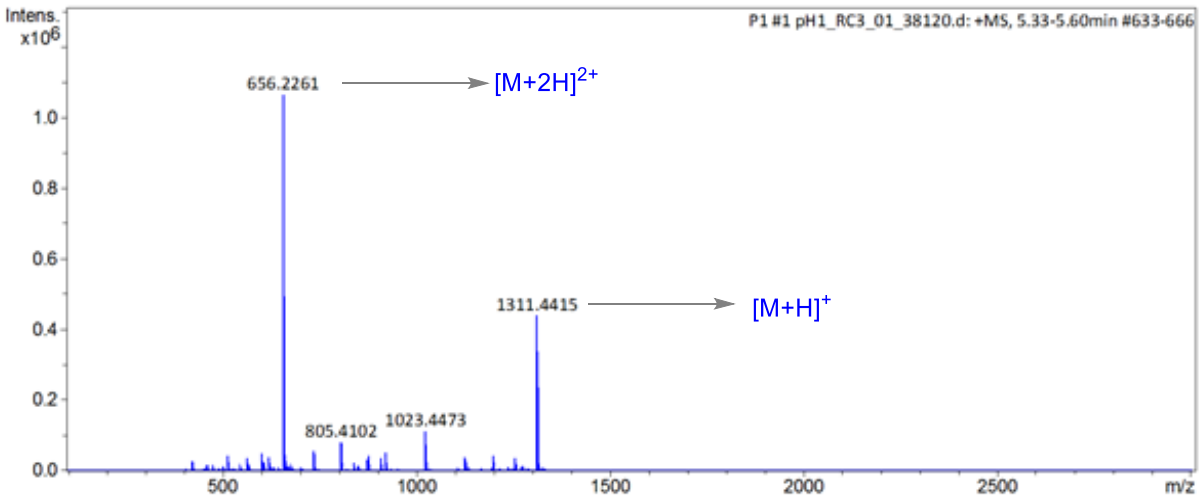


SD-9: Figure 0.8: Chromatograms of crude Pep3_RJM before purification. a) Base peak chromatogram, b) UV spectra 190 nm, c) UV spectra 195 nm, d) UV spectra 200 nm and e) UV spectra 200 nm

Appendix II- Supplementary information for Chapter 3

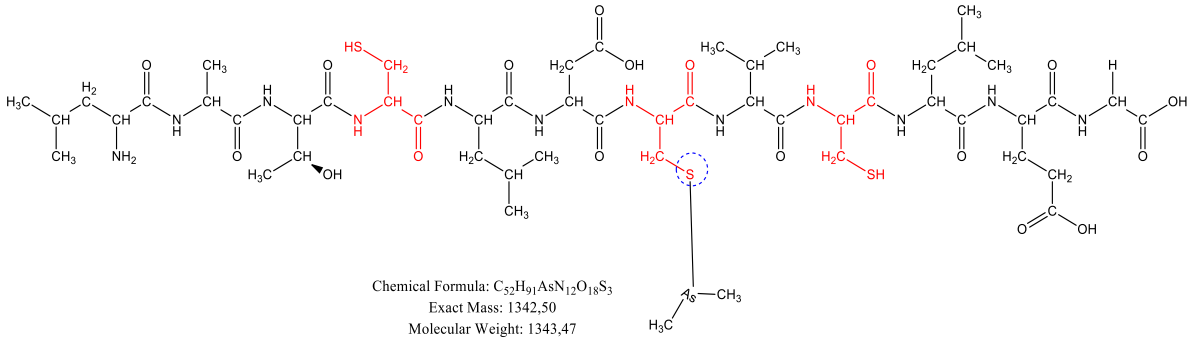
Pep1-RJM binding to iAs(III)

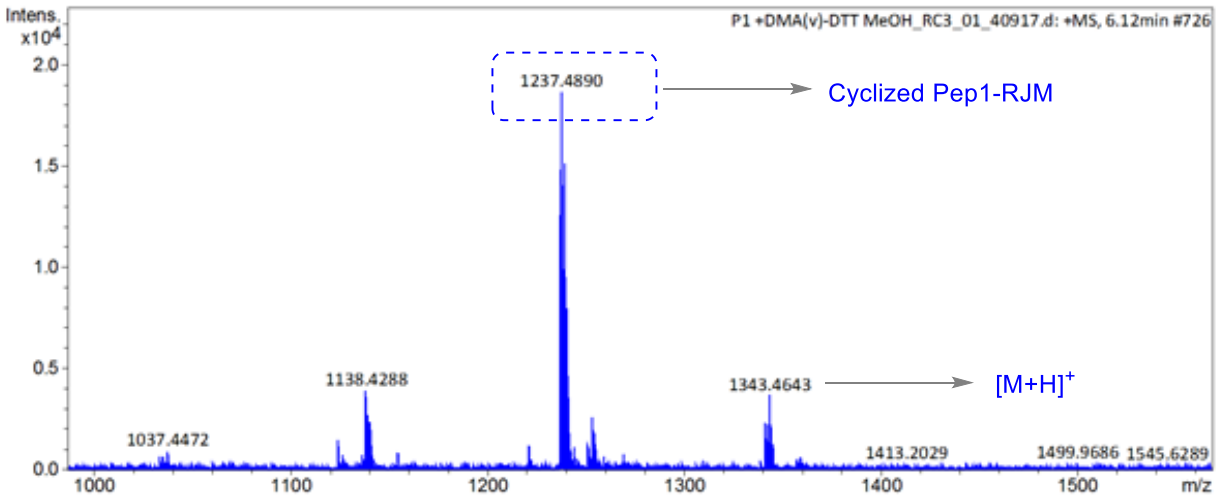




SD-10: Figure 0.9: Mass spectrum of iAs-Pep1 complex

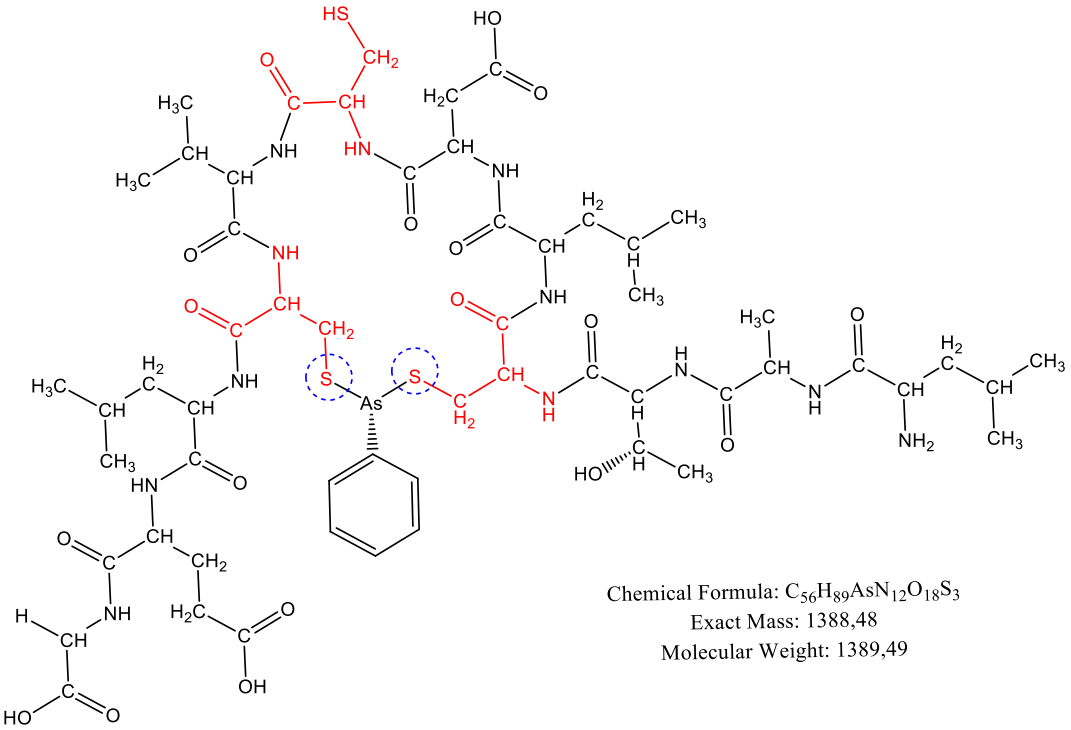
Pep1-RJM binding to DMA(III)



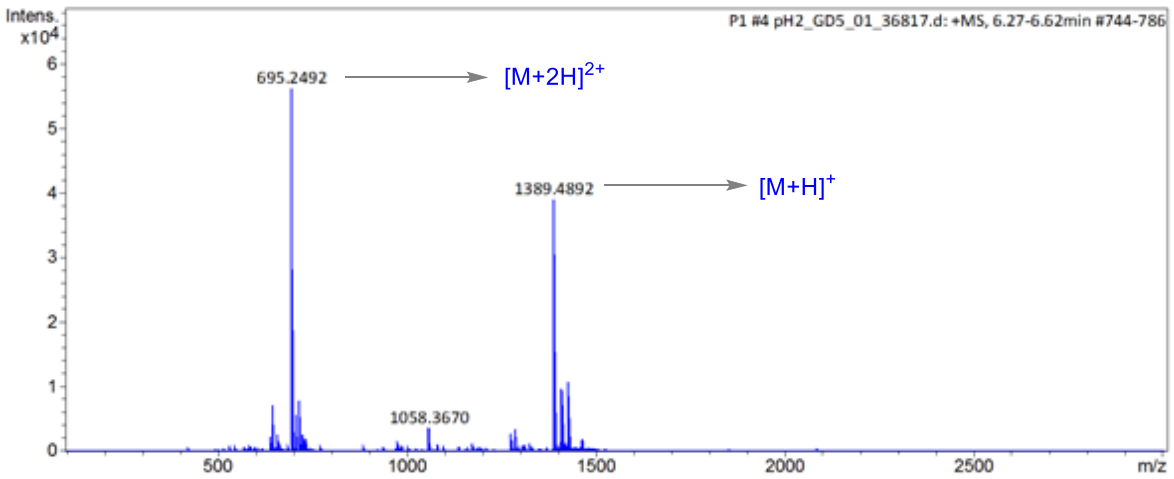


SD-11:Figure 0.10: Mass spectrum of DMA-Pep1 complex

Pep1-RJM binding to PAO(III)

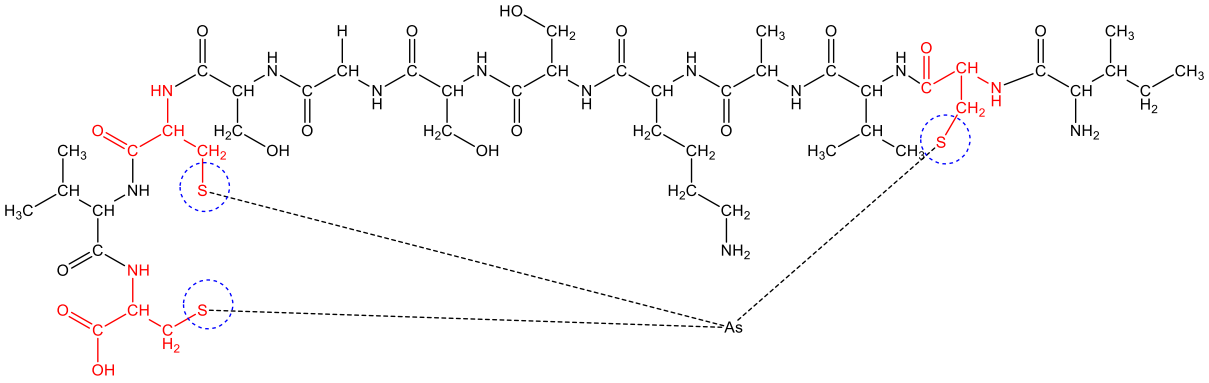


Chemical Formula: $C_{56}H_{89}AsN_{12}O_{18}S_3$
Exact Mass: 1388,48
Molecular Weight: 1389,49

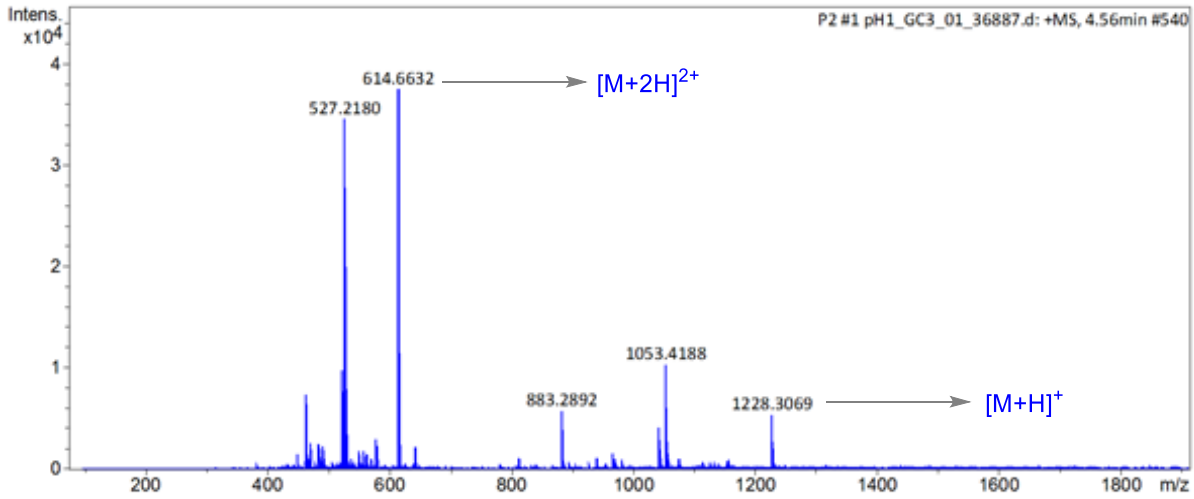


SD-12: Figure 0.11: Mass spectrum of PAO-Pep1 complex.

Pep2-RJM binding to iAs(III)

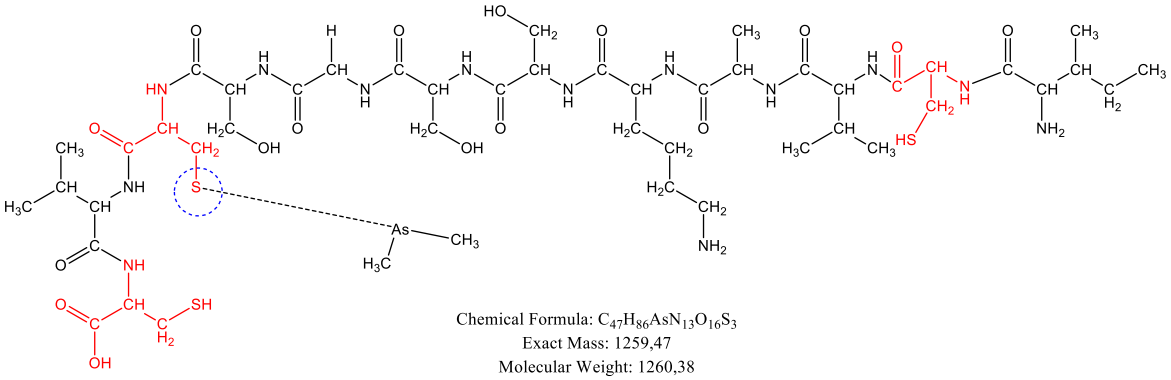


Chemical Formula: $C_{45}H_{78}AsN_{13}O_{16}S_3$
Exact Mass: 1227,41
Molecular Weight: 1228,29

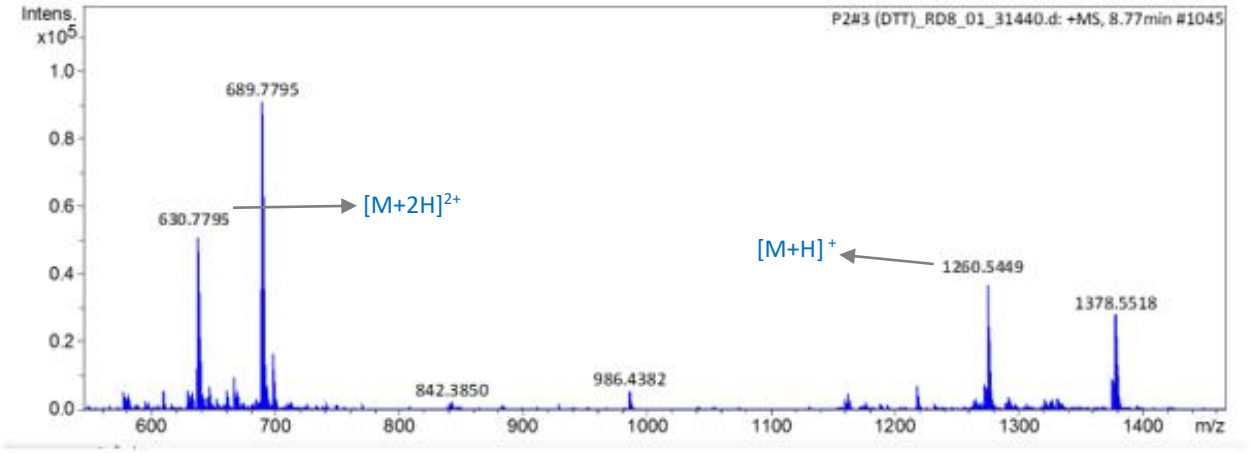


SD-13: Figure 0.12: Mass spectrum of iAs-Pep2 complex.

Pep2-RJM binding to DMA(III)

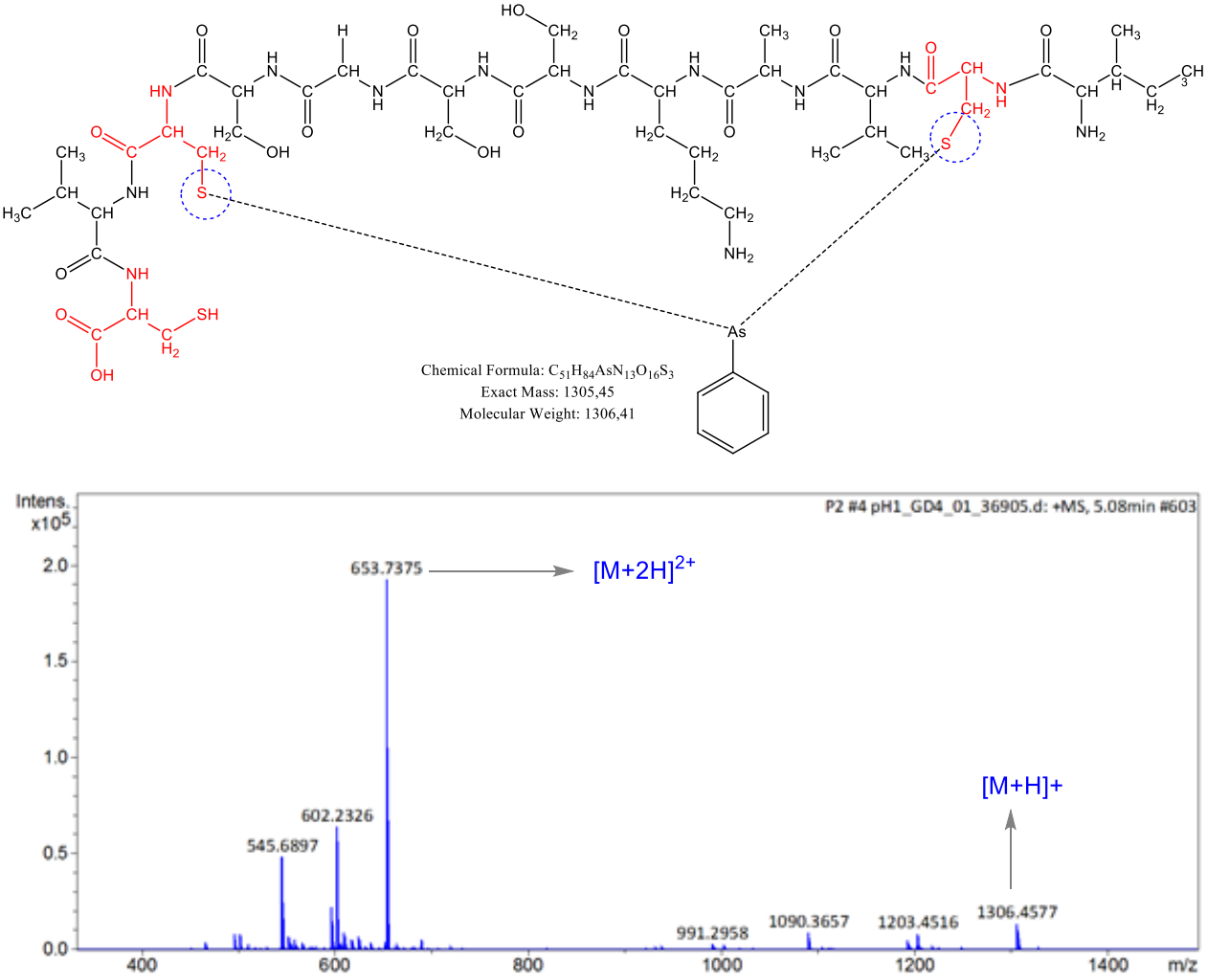


Chemical Formula: $C_{47}H_{86}AsN_{13}O_{16}S_3$
Exact Mass: 1259,47
Molecular Weight: 1260,38



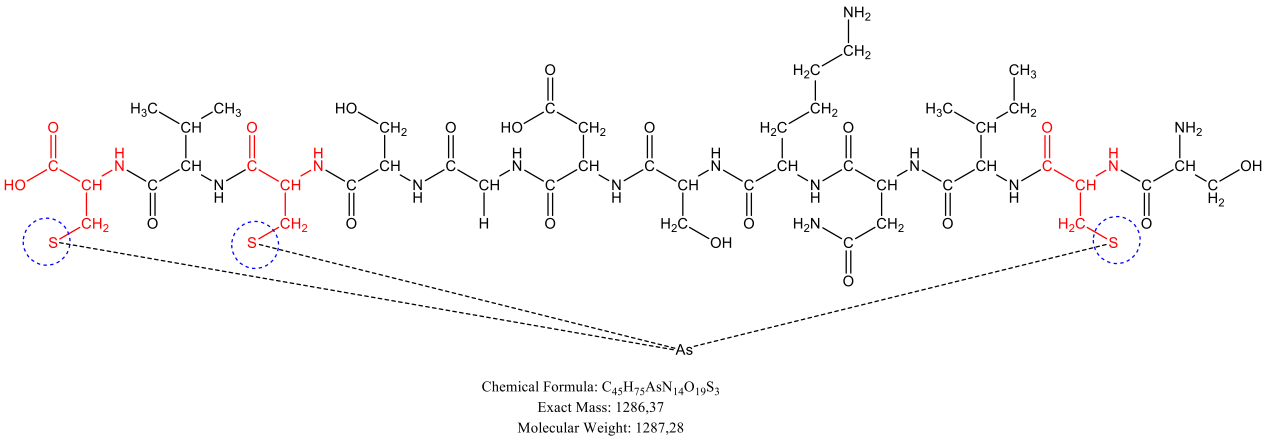
DS-14: Figure 0.13: Mass spectrum of DMA-Pep2 complex.

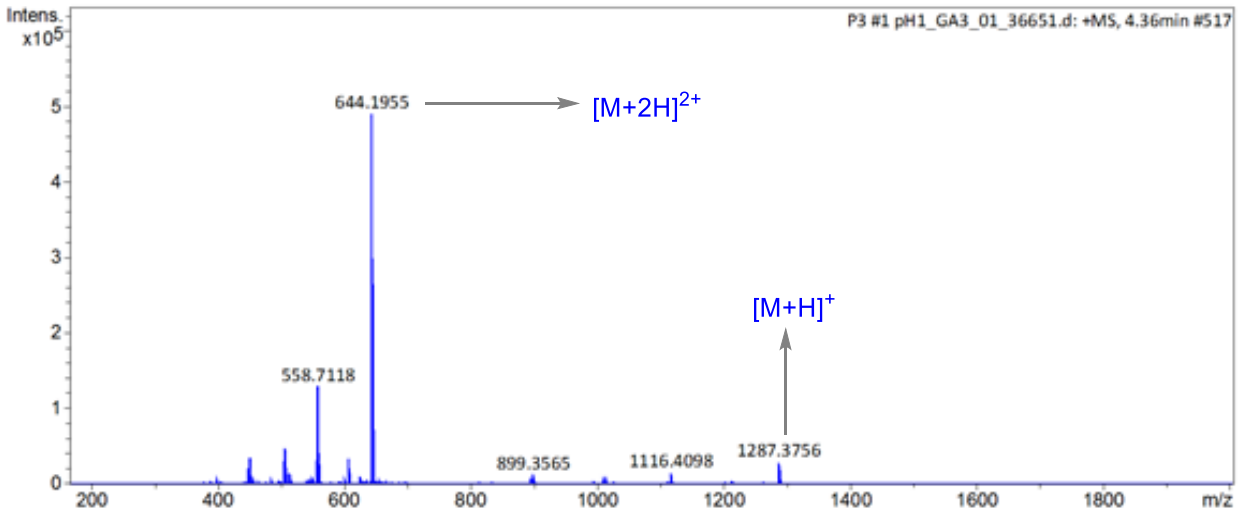
Pep2-RJM binding to PAO(III)



SD-15: Figure 0.14: Mass spectrum of PAO-Pep2 complex.

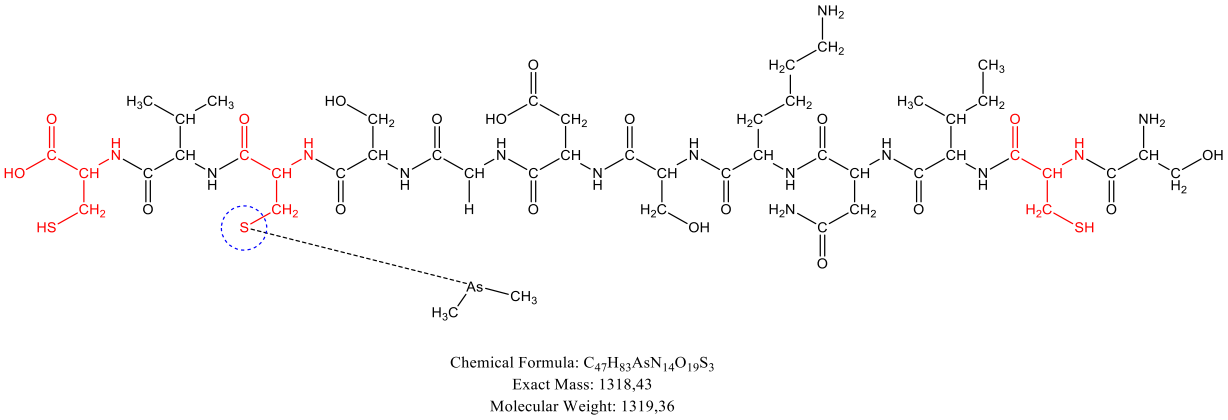
Pep3-RJM binding to iAs(III)



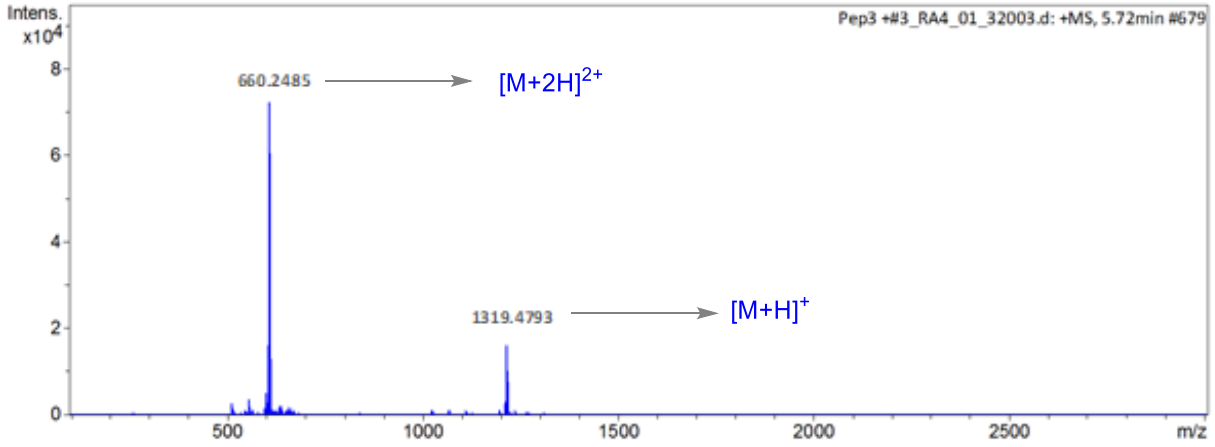


SD-16: Figure 0.15: Mass spectrum of iAs-Pep3 complex.

Pep3-RJM binding to DMA(III)

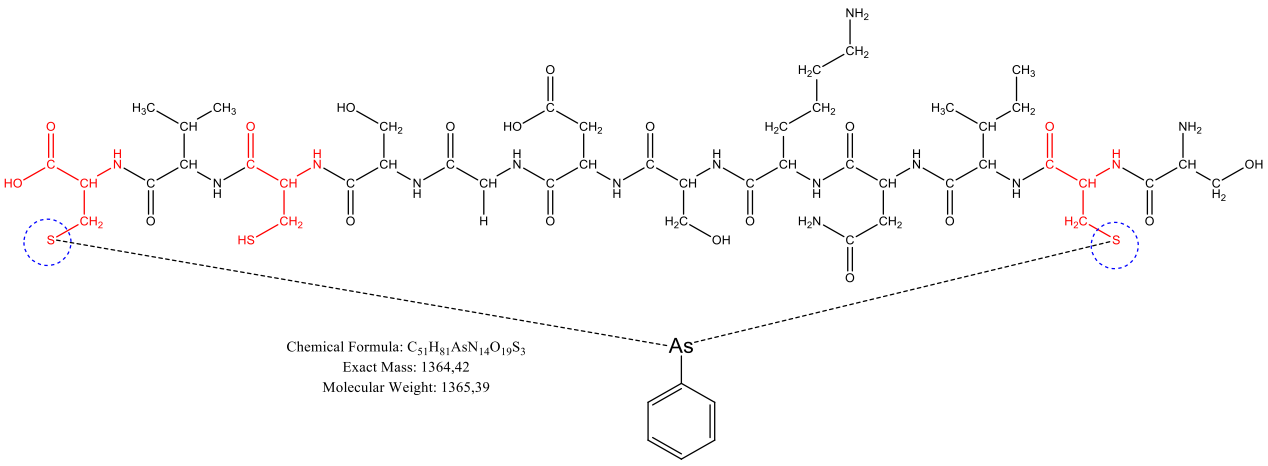


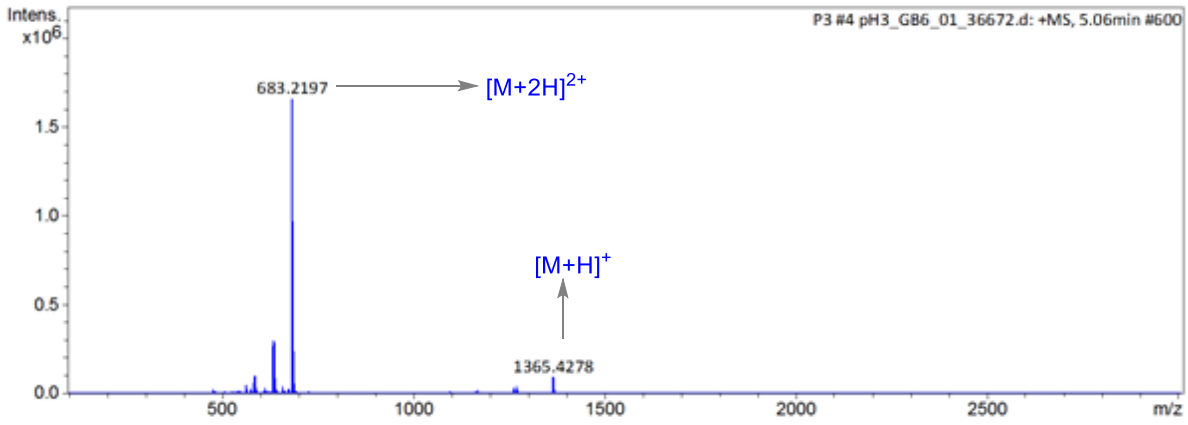
Chemical Formula: $C_{47}H_{83}AsN_{14}O_{19}S_3$
Exact Mass: 1318,43
Molecular Weight: 1319,36



SD-17: Figure 0.16: Mass spectrum for DMA-Pep2 complex.

Pep3-RJM binding to PAO(III)





SD-18: Figure 0.17: Mass spectrum of PAO-Pep3 complex.

SD-19: Table 0.2: Effect of pH on arsenic(III) binding.

pH	Relative intensity					
	iAs(III)-Pep1	%RSD	DMA(III)-Pep1	%RSD	PAO(III)-Pep3	%RSD
2	112510.18	3.61	101510.77	4.35	132512.60	4.64
4	119321.71	2.98	113587.19	3.68	173519.46	3.25
6	120325.33	4.65	112832.93	5.42	179532.51	3.62
8	222457.87	3.56	129524.54	2.27	252871.77	4.18
10	235247.39	3.89	131914.91	4.18	257943.59	3.49
12	210423.53	4.02	118010.62	2.53	218340.76	6.77
14	168551.32	3.41	98680.70	3.66	199279.94	4.94

SD-20: Table 0.3: Effect of reaction temperature on binding.

Temp (˚C)	Relative intensity					
	iAs(III)-Pep1	%RSD	DMA(III)-Pep1	%RSD	PAO(III)-Pep3	%RSD
25	222457.8	2.30	129524.5	3.81	252871.7	2.84
30	235798.6	5.29	132762.3	4.39	268036.4	3.40
35	254891.1	2.61	195889.2	2.54	373215.9	3.93
40	289624.0	3.48	210068.6	3.25	425361.2	2.70
45	266864.6	4.59	215320.3	2.85	401995.3	5.04
50	176848.9	5.57	205630.9	4.72	356375.5	2.31
55	123794.3	2.63	147048.0	4.14	297752.8	5.97

SD-21: Table 0.4: Effect of reaction time on binding.

Reaction time (min)	Relative intensity					
	iAs(III)-Pep1	%RSD	DMA(III)-Pep1	%RSD	PAO(III)-Pep3	%RSD
0	0	**	0	**	0	**
5	79800.62	3.29	2620.39	5.94	8360.42	4.35
10	277560.91	3.69	81310.46	2.43	229280.45	5.57
15	289624.08	5.61	210068.60	3.27	395361.21	2.40
20	290000.40	3.90	215320.33	3.53	401995.35	3.44
25	290325.15	4.72	216120.35	4.66	401995.59	2.96
30	291095.58	2.93	216597.59	3.77	401995.15	3.55

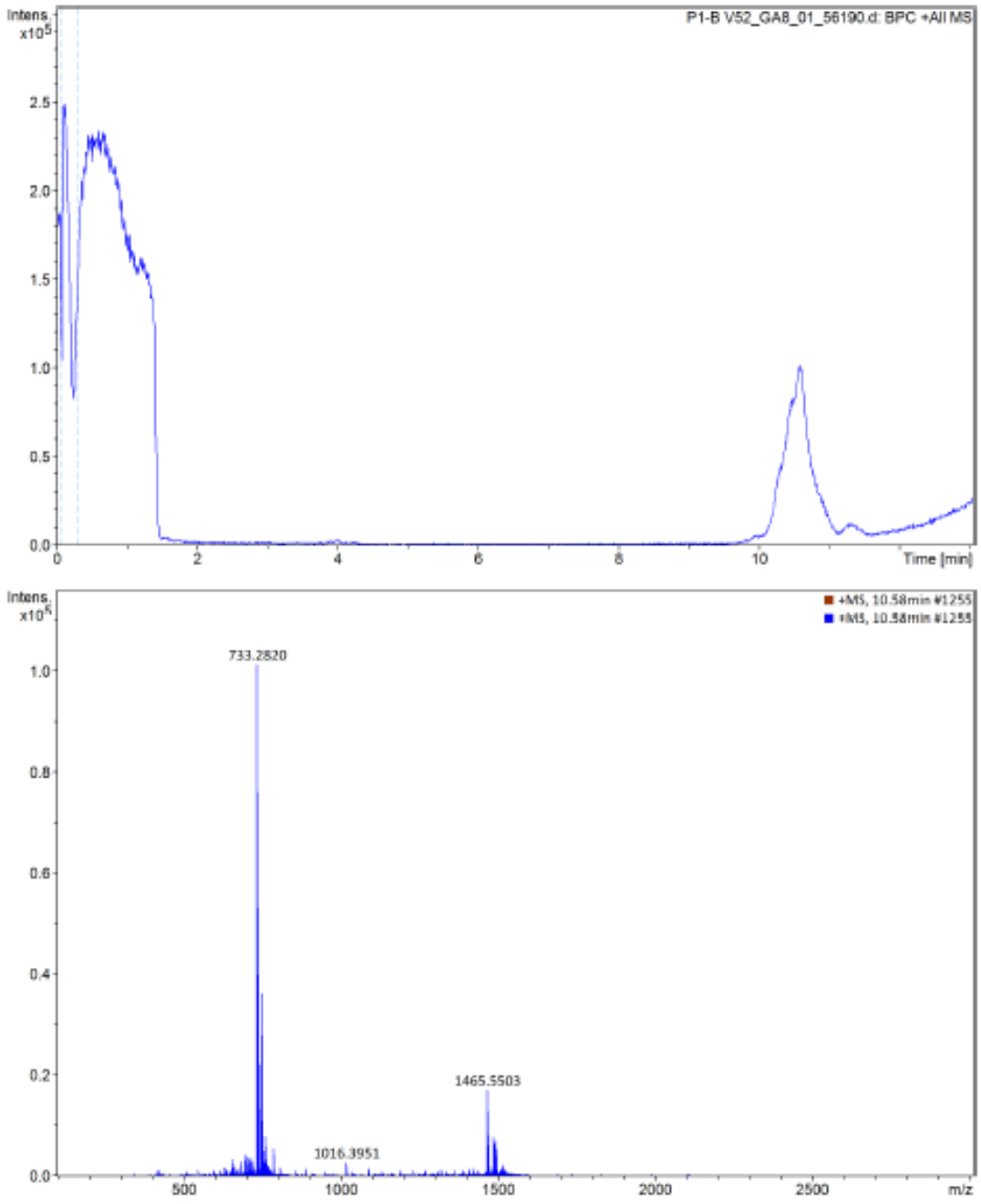
SD-22: Table 0.5: Effect of initial As(III) on binding.

initial As(III) Concentration (µM)	% binding of As(III)					
	iAs(III)-Pep1	%RSD	DMA(III)-Pep1	%RSD	PAO(III)-Pep3	%RSD
2	0	**	0	**	0	**
4	11.50	3.42	8.90	3.08	19.23	2.82
6	15.71	5.21	10.81	4.72	25.91	5.57
8	23.79	2.39	15.34	3.89	26.17	2.76
10	28.58	4.57	15.98	4.35	26.52	5.01
12	29.81	3.92	16.32	5.29	26.68	3.99
14	29.62	2.31	17.01	3.52	27.14	4.12

SD-23: Table 0.6: Effect of peptide dose on As(III) binding.

Pep1-RJM Concentration (μM)	Relative intensity					
	iAs(III)-Pep1	%RSD	DMA(III)-Pep1	%RSD	PAO(III)-Pep3	%RSD
0	0	**	0	**	0	**
5	78000.64	2.53	54600.40	3.18	8360.45	5.98
10	89700.91	3.07	56910.24	3.73	25614.63	2.43
15	215805.35	4.15	89041.38	4.53	299401.33	3.84
20	306323.84	3.98	128578.81	2.75	484665.82	4.39
25	400452.72	2.90	130168.62	4.67	492294.51	3.03
30	380625.67	4.25	60853.34	2.39	440094.51	3.15
35	375829.81	5.02	22987.12	3.66	262294.51	2.87
40	370028.13	3.60	20987.12	2.52	217295.51	4.41

Appendix III- Supplementary information for Chapter 4



SD-24: Figure 0.18: LC-MS of biotinylated Pep1-RJM.