METABOLISM OF HALOGENATED COMPOUNDS BY *RHODOCOCCUS* UKMP-5M

JAYASUDHA NAGARAJAN

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY AND DIPLOMA OF IMPERIAL COLLEGE LONDON

DEPARTMENT OF CHEMISTRY
FACULTY OF NATURAL SCIENCES
IMPERIAL COLLEGE LONDON

2014

DECLARATION

I certify that this thesis represents my own work, unless otherwise stated. All external contributions and any information derived from other sources have been acknowledged accordingly.

JAYASUDHA NAGARAJAN

In living memory of my late granny, Madam Thanalaetchumi Appadurai…

Dedicated to Dr. Judit Nagy, an outstanding scientist and the most inspirational woman I have ever met, without whom my PhD journey would have not been possible....may your soul rest in peace…

Dedicated to my dad and mom Mr. & Mrs. Nagarajan Sarujuny, the most precious gift in my life. For their endless support, love and encouragement through all my walks of life…. Thank you pa & ma.

&

To my lovely sister, Jayesree Nagarajan, the little diamond in the family which never fails to shine.. Thank you so much for being there for me whenever I need you.. May you also be motivated to achieve great heights in life…

ACKNOWLEDGEMENT

First and foremost, I'd like to express my deepest gratitude to Professor Tony Cass for accepting me in his research team and for all his patience and valuable inputs to successfully drive this project. I'd like to also thank my co-supervisor, Professor David Leak for his continued support and guidance throughout the project.

Thanks to the Malaysian Ministry of Science, Technology and Innovation for sponsoring my PhD studies in Imperial College London and for giving me an opportunity to be among the best brains in the world! My heartfelt gratitude goes to Dr. Lok Hang Mak for all the guidance, knowledge, ideas and friendship that helped carry me through. To the members, past and present, of Cass Group; Dr. Sanjiv, Dr. Muthu, Dr. Thao, Dr. Chris, Achi, Orada, Sasinee, Ben, Dan, Mahsan, Ascanio and Hannah- thank you for your support encouragement and the 'giggles'. Sarah, Tami and Nawal – thank you so much girls, you are more like sisters to me and thanks a lot for all the 'fun' we had together..I'll cherish those moments forever~

Very special thanks to my family and friend in London, my best friend Hazeeq. You have been a great support to me in many ways especially during all those time when I was down. You made my life in London, very colourful and I do have really great memories to bring back home! BFF forever. To all my other friends in UK, esp. Raja, Joe, Shahrul, Mus and Roobini - thanks guys for the motivation and enlightening words from all of you especially when I need them the most! Keep in touch all…Not forgetting, friends and colleagues in Malaysia esp. Prof. Latif and Dr. Maegala for all the valuable inputs throughout my PhD.

Finally, thank you to my family members, relatives and friends for being there for me. Last but not least, to the Almighty for giving me the strength and perseverance to complete my studies within the stipulated time frame and for keeping me surrounded by these beautiful people!

ABSTRACT

Members of the genus *Rhodococcus* are well known for their high metabolic capabilities to degrade wide range of organic compounds ranging from simple hydrocarbons to more recalcitrant compounds such as polychlorinated biphenyls. Their ability to display novel enzymatic capabilities for the transformation of many hazardous contaminants in the environment makes them a potential candidate for bioremediation. *Rhodococcus* UKMP-5M, an actinomycete isolated in Peninsular Malaysia shows great potential towards degradation of cyanide, hydrocarbons and phenolic compounds. In the present study, the capacity of this strain to degrade halogenated compounds was explored.

Preliminary investigations have proven that *R*. UKMP-5M was not able to utilise any of the halogenated compounds tested as sole carbon and energy source, but the resting cells of *R*. UKMP-5M was able to dechlorinate several compounds which include chloroalkanes, chloroalcohols and chloroacids and the activity was three fold higher when the cells were grown in the presence of 1-Chlorobutane (1-CB). Therefore, 1-CB was chosen as a substrate to unravel the mechanism of dehalogenation in *R*. UKMP-5M. In contrast to the classic hydrolytic route for the assimilation of 1-CB in many organisms, *R*. UKMP-5M was able to metabolise and release chloride from 1-CB, but is unable to use the product from 1-CB metabolism as growth substrate.

On comparing the protein profiles of the induced and non-induced cells of *R*. UKMP-5M, two types of monooxygenases were identified in the induced condition, which were not present in the uninduced sample. The strict oxygen requirement for dechlorination of 1-CB and the identification of monooxygenases in the induced protein extract suggests that 1-CB dehalogenation is likely to be catalysed by a monooxygenase. In addition to these monooxygenases, a protein that was later identified as amidohydrolase (Ah) was also found to be induced when the cells were exposed to 1-CB. Therefore, Ah from *R*. UKMP-5M was cloned and expressed in *E. coli* to test the ability of the purified Ah to release chloride from

1-CB. The heterologous expression of Ah in *E. coli* resulted in the formation of inclusion bodies and the western blot analyses further confirmed that no soluble form of Ah was present. Multiple attempts to obtain a soluble and functionally active Ah were not successful. Therefore, on-column refolding was carried out to obtain a biologically active Ah. A 3D model based on structural homology was predicted as a preliminary step to characterize this protein. However, when assayed with 1-CB, Ah was found not to catalyze dehalogenation. All results of this thesis suggest that metabolism of 1-CB by *R*. UKMP-5M is via γ-butyrolactone which acts as a potent intracellular electrophile that covalently modifies proteins and nucleic acids. The findings from this research are important to determine the metabolic capacity of a Malaysian *Rhodoccoccus* in dehalogenation of halogenated compounds and its potential application in bioremediation.

LIST OF FIGURES

Figure 1.1	Concept of bioremediation	21
Figure 1.2	The most frequently detected organic compounds in the European ground water reported in Environmental Agency Database	24
Figure 1.3	The common routes of horizontal gene transfer	28
Figure 1.4	Degradability of some halogenated compounds by microorganism	31
Figure 1.5	Reaction of hydrolytic dehalogenation	32
Figure 1.6	Mechanism of thiolytic dehalogenation	33
Figure 1.7	Dehalogenation of halomethane catalysed by methyltransferase	34
Figure 1.8	Mechanism of dehalogenation through hydration	34
Figure 1.9	Mechanism of oxidative dehalogenation	35
Figure 1.10	Dehalogenation of lindane by dehydrohalogenase	36
Figure 1.11	Mechanism of dehalogenation through intramolecular substitution reaction	37
Figure 1.12A	Mechanism of reductive dehalogenation of tetrachlorohydroquinone (TCHQ)	38
Figure 1.12B	Reductive dehalogenation of tetrachloroethylene (PCE)	39
Figure 1.13	The increasing pattern of number of publications concerned with degradation of nitriles, aromatic and halogenated compounds by *Rhodococci*	40
Figure 1.14	Different morphological forms of *Rhodococcus*	42
Figure 2.1	Growth of two *Rhodococcus* strains monitored spectrophotometrically at 12 hours interval for 60 hours	56
Figure 2.2	Growth of *R.* UKMP-5M in minimal media containing 1-CB supplemented with other primary substrates	58
Figure 2.3	Growth of *R.* UKMP-5M in minimal media supplemented with 20 mM of various carbon sources	59
Figure 2.4	Calibration curve for chloride assay	60
Figure 2.5	Chloride release from 1-CB by resting cells of *R.* UKMP-5M grown in different carbon sources	61

Figure 2.6	Chloride release from 1-CB by resting cells of *R.* UKMP-5M grown in 20 mM glucose and 20 mM glucose supplemented with 1-CB	64
Figure 2.7	Chloride release from 1-CB by resting cells of *R.* UKMP-5M grown in LB and LB supplemented with 1-CB	65
Figure 2.8	Chloride release from 1-CB by resting cells of *R.* UKMP-5M grown in glucose and glucose supplemented with each component of LB	66
Figure 2.9	Effect of manipulating physical parameters on the degradation of 1-CB by *R.* UKMP-5M	67
Figure 2.10A	Chloride release from 1-CB by resting cells of *R.* UKMP-5M monitored in the presence and absence of oxygen	71
Figure 2.10B	Cells monitored in the absence of oxygen were aerated and re-incubated with 1-CB for their activity recovery	71
Figure 2.11	Chloride release from 1-CB by resting cells of *R.* UKMP-5M monitored in the presence of carbon monoxide	73
Figure 2.12A	Possible metabolic routes for the dehalogenation of 1-CB in *R.* NCIMB 13064	74
Figure 2.12B	Dehalogenation of 1-CB via internal hydroxylation	75
Figure 2.13	Cell density of *R.* UKMP-5M grown on several potential intermediates of 1-CB degradation	76
Figure 2.14	Growth recovery patterns of *R.* UKMP-5M unexposed and exposed to 1-CB	77
Figure 3.1	The schematic workflow of two-dimensional gel electrophoresis	80
Figure 3.2	Calibration curve for protein assay	91
Figure 3.3	The non-reproducibility of 2D GE	92
Figure 3.4	1D GE of protein samples prepared from *R.* UKMP-5M cells grown in non-induced and induced condition	93
Figure 3.5	Protein bands which were excised for identification using nano LC-MS/MS identification.	95
Figure 3.6	A Venn distribution representing the number of unique and common proteins identified through nano LC-MS/MS	96
Figure 3.7	Pie chart showing the biological functions of common proteins identified in both induced and non-induced protein samples	96

Figure 3.8	The matched peptide sequences of Ah, Moh and Alm identified through LC-MS/MS	98
Figure 3.9	Transcript analysis of proteins identified through proteomics	100
Figure 3.10	Gene cluster containing genes identified in proteomics study	104
Figure 4.1	Agarose gel of *Ah* gene from *R*. UKMP-5M	116
Figure 4.2A	The pET-28a(+) vector system containing multiple cloning sites and other components	117
Figure 4.2B	Part of the nucleotide sequence map for pET-28a, b and c(+)	117
Figure 4.3	Analytical digestion of pET-28(a)+ vector	119
Figure 4.4	Agarose gel of Ah insert containing appropriate restriction sites	120
Figure 4.5	A schematic representation of amidohydrolase (*Ah*) cloning into pET-28a(+) vector.	121
Figure 4.6	Agarose gel of colony PCR after transformation of pET-28a_AH into *E. coli* DH5α	122
Figure 4.7A	1D GE of soluble protein from *E. coli* BL21(DE3) containing pET-28a_AHntc	123
Figure 4.7B	1D GE of soluble protein from *E. coli* BL21(DE3) containing pET-28a_AHctc	124
Figure 4.7C	1D GE of soluble protein from *E. coli* BL21(DE3) containing pET-28a_AHnat	124
Figure 4.7D	1D GE of insoluble protein from *E. coli* BL21(DE3) containing recombinant plasmid	125
Figure 4.8	Immunoblotting with anti-6X His-tag antibody	126
Figure 4.9	1D GE of soluble protein and IB extracted from *E. coli* BL21(DE3) containing pET-28a_Ahntc that was grown in medium supplemented with 1% glucose	128
Figure 4.10	1D GE of soluble protein and IB extracted from *E. coli* BL21(DE3)pLysS and Rosetta(DE3) harbouring the recombinant pET-28a_Ahntc	129
Figure 4.11	Vector map of the 4989 bp pGEX-6P-1	130
Figure 4.12	A schematic representation of Ah cloning into pGEX-6P-1 vector	131
Figure 4.13	1D GE of soluble protein and IB extracted from *E. coli* BL21(DE3) containing the recombinant pGEX-6P-1_AH	132
Figure 5.1	Simplified model to demonstrate protein folding	136
Figure 5.2	Circular dichroism (CD) spectra for the each secondary conformation	138

Figure 5.3	SDS-PAGE gel of protein samples obtained at each step during purification of Ah	146
Figure 5.4	FASTA format sequence of hypothetical protein (Ah) with 77% sequence coverage identified by LC-MS/MS	147
Figure 5.5	Amino acid sequence of Ah from *R*. UKMP-5M	149
Figure 5.6	A snapshot view of the predicted distribution of a three-state secondary structure for each amino acid residue in Ah	150
Figure 5.7	A snapshot of the predicted disorder distribution for each amino acid residue in Ah	150
Figure 5.8	Summary of compositions of three-state secondary structure and the solvent accessibility of Ah predicted using RaptorX	151
Figure 5.9	The predicted 3D-model for amidohydrolase (Ah) from *R*. UKMP-5M based on structural homology	153
Figure 5.10	The superposed structure of Ah with five templates which were used to build the 3D model	155
Figure 5.11	The predicted binding site of Ah based on structural homology	157
Figure 5.12	The far-UV spectrum of refolded Ah	158
Figure 5.13	Calibration curve for ammonia measurement	161
Figure 5.14	Ammonia release from five different amides by the refolded Ah	162
Figure 5.15	Ammonia release from 100 mM urea by different concentrations of refolded Ah	162
Figure 5.16	Enzyme activity of refolded Ah in the presence and absence of metal	164
Figure 5.17	Activity of Ah assayed using 100 mM urea and 100 mM 1-CB	165
Figure 6.1	Hydrolytic dehalogenation of 1-Chlorobutane (1-CB) and the metabolism of the resulting products via β-oxidation to produce energy.	170
Figure 6.2	A proposed model for the mechanism of 1-Chlorobutane (1-CB) dehalogenation in *R*. UKMP-5M.	174

LIST OF TABLES

Table 1.1	A summary of the major uses of some chlorinated methanes and ethanes	26
Table 2.1	The amount of cells obtained from 50 ml MSM supplemented with 20 mM of different carbon sources after 24 hours	59
Table 2.2	Dehalogenation activity of 1-CB by resting cells of *R.* UKMP-5M grown in 20 mM of different carbon sources	62
Table 2.3	Biotransformation of other halogenated compounds by resting cell of *R.* UKMP-5M induced with 1-CB	70
Table 3.1	The voltage program used for IEF of 11 cm IPG strip pH 4-7	85
Table 3.2	Thermocycling conditions for PCR amplification using One *Taq* Polymerase	90
Table 3.3	List of proteins present in the uninduced protein sample	97
Table 3.4	List of proteins present in the induced protein sample	97
Table 3.5	List of oligonucleotides used for RT-PCR analyses	99
Table 3.6	Gene annotations for genes present in a gene cluster	104
Table 4.1	List of oligonucleotides used for cloning of *Ah*	113
Table 5.1	List of buffers and their compositions used for the inclusion bodies preparation, refolding and purification of amidoydrolase (Ah)	140
Table 5.2	List of reagents and their compositions for indophenol blue method used to measure ammonia release	143
Table 5.3	Amino acid composition in amidohydrolase (Ah)	145
Table 5.4	List of templates and their information from Protein Data Bank (PDB) that were used to generate a 3D conformation of Ah	149
Table 5.5	Secondary structure composition of refolded amidohydrolase (Ah) predicted from sequence analyses and estimated from the CD spectra	159

ABBREVIATIONS

1-CB	1-Chlorobutane
1D GE	one dimensional gel electrophoresis
2D GE	two-dimensional gel electrophoresis
3D	three-dimensional
Ah	amidohydrolase
Alm	alkane-1-monooxygenase
bp	base pair(s)
BPH	biphenyl
BSA	bovine serum albumin
CD	circular dichroism
cDNA	complementary DNA
CF	chloroform
Cl^-	chloride ion
CM	chloromethane
CO	carbon monooxide
CT	carbon tetrachloride
cyt P450	cytochrome P450
DCA	1,2-dichloroethane
DDT	dichlorodiphenyltrichloroethane
DNA	deoxyribonucleic acid
dNTP	dinucleotide phosphate
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EPA	Environmental Protection Agency
GST	gluthathione-S-transferase
HGT	horizontal gene transfer
His	histidine
HPLC	high perfomance liguid chromatography
IEF	isoelectric focusing
IMAC	immobilised metal affinity chromatography
IPG	immobilized pH gradient
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kDa	kilodalton
LB	luria bertani
LC	liquid chromatography
MCA	monochloroacetic acid
MMO	methane monooxygenase
Moh	monooxygenase hydroxylase
MS	mass spectrometry
MSM	minimal salt media
NADH	nicotinamide adenine dinucleotide
NADPH	nicotinamide adenine dinucleotide phosphate
NCBI	National Centre for Biotechnology
Ni.NTA	nickel-nitrilotriacetic acid
O_2	oxygen
OD_{600nm}	optical density at 600 nm
PAGE	polyacrylamide gel electrophoresis
PCB	polychlorobiphenyls
PCR	polymerase chain reaction
PDB	Protein Data Bank
POP	persistent organic pollutant

RNA	ribonucleic acid
SDS	sodium dodecyl sulphate
TAE	tris-acetate-EDTA
TBST	tris-buffered saline with Tween 20
TCA	1,1,1-trichloroethane
TCA	trichloroacetic acid
TCE	tetrachloroethylene
TCHQ	tetrachlorohydroquinone
T_m	melting temperature
Trs	threonine synthase
U	unit

TABLE OF CONTENTS

DECLARATION **2**

ACKNOWLEDGEMENT **4**

ABSTRACT **5**

LIST OF FIGURES **7**

LIST OF TABLES **11**

ABBREVIATIONS **12**

TABLE OF CONTENTS **14**

CHAPTER 1 GENERAL INTRODUCTION **19**

1.1 Bioremediation 19

1.2 Halogenated compounds 23

1.2.1 Haloalkanes, their properties and routes of exposures 25

1.3 Microbial degradation of halogenated compounds 27

1.3.1 Utilisation of halogenated compounds as carbon and energy source 29

1.3.2 Co-metabolism 30

1.4 Bacterial dehalogenation reactions 32

1.5 The genus *Rhodococcus* 40

1.5.1 Characteristics of *Rhodococcus* 41

1.5.2 Significance of *Rhodococcus* 43

1.5.3 *Rhodococcus* UKMP-5M 45

1.6 Background to the Project 46

1.7 Aim and Objectives 47

CHAPTER 2 DEGRADATION OF 1-CHLOROBUTANE BY *RHODOCOCCUS* UKMP-5M **48**

2.1 Introduction 48

2.2 Materials and methods 52

2.2.1 Bacterial strain and maintenance 52

2.2.2 Glycerol stock preparation 52
2.2.3 Growth media 52
2.2.4 Determination of cell growth 53
2.2.5 Preparation of resting cells of *R*. UKMP-5M 53
2.2.6 Resting cell assay 53
2.2.7 Chloride release assay 54
2.2.8 Test for inducibility of dehalogenation 54
2.2.9 Optimization of reaction conditions 54
2.2.10 Dehalogenation of other halogenated compounds 54
2.2.11 Oxygen dependence of the dehalogenation activity 55
2.2.12 Test for the involvement of cytochrome P450 55
2.3 Results and Discussions 55
2.3.1 Dehalogenation of 1-Chlorobutane by *R.* UKMP-5M 55
2.3.2 Growth of *R.* UKMP on different carbon sources 58
2.3.3 Chloride assay 60
2.3.4 Dehalogenation of 1-CB by resting cells of *R*. UKMP-5M 61
2.3.5 Further inducibility of 1-CB degradation 62
2.3.6 Optimization of reaction conditions 66
2.3.7 Oxygen dependence of the activity 70
2.3.8 Metabolic pathway analyses 74
CHAPTER 3 PROTEOMICS AND TRANCRIPT ANALYSES 79
3.1 Introduction 79
3.2 Materials and methods 84
3.2.1 Bacterial cell preparation 84
3.2.2 Protein sample preparation for 2D GE 84
3.2.3 Protein concentration assay 84
3.2.4 Rehydration of IPG strip 85
3.2.5 First dimension separation - Isoelectric Focusing (IEF) 85
3.2.6 Equilibration of IPG strips prior to second dimension separation 86
3.2.7 Second dimension separation – SDS PAGE 86
3.2.8 Protein gel staining 86
3.2.9 Protein sample preparation for 1D GE 87
3.2.10 One-Dimensional Gel Electrophoresis 87

3.2.11 Mass spectrometry 87
3.2.11.1 Excision of protein bands and proteolytic digestion 87
3.2.11.2 Tandem mass spectromertry (MS/MS) 88
3.2.11.3 Protein Identification 89
3.2.12 Isolation of total *Rhodococcus* RNA 89
3.2.13 Complementary DNA (cDNA) synthesis 89
3.2.14 Polymerase Chain Reaction (PCR) 90
3.3 Results 90
3.3.1 Protein quantification 90
3.3.2 Two-Dimensional Gel Electrophoresis 91
3.3.3 One dimension protein separation 93
3.3.4 Protein identification from one dimension electrophoresis gels 94
3.3.5 Validation of proteomics data through transcript analysis 99
3.4 Discussion 100
CHAPTER 4 MOLECULAR CLONING AND EXPRESSION OF AMIDOHYDROLASE FROM *R.* UKMP-5M INTO *E. COLI* 106
4.1 Introduction 106
4.2 Materials and Methods 108
4.2.1 General procedures for DNA manipulation 108
4.2.1.1 Sequence of amidohydrolase (*Ah*) gene 108
4.2.1.2 *Ah* gene isolation from *R.* UKMP-5M 108
4.2.1.3 Recovery of amplified DNA from agarose gel 109
4.2.1.4 Cloning vectors 109
4.2.1.5 Preparation of insert DNA 109
4.2.1.6 Restriction endonucleases digestion of vector and DNA 110
4.2.1.7 DNA Ligation 110
4.2.1.8 Preparation of chemically competent cells of *E. coli* 110
4.2.1.9 Transformation of recombinant DNA into *E. coli* DH5α 111
4.2.1.10 Colony Polymerase Chain Reaction (PCR) 111
4.2.1.11 DNA Sequencing 112
4.2.2 Small scale recombinant protein expression 114
4.2.2.1 Growth and induction 114
4.2.2.2 Soluble protein extraction 114

4.2.2.3 Analysis of insoluble protein 115
4.2.2.4 Polyacrylamide gel electrophoresis and Coomassie staining 115
4.2.2.5 Western blotting 115
4.3 Results 116
4.3.1 Amplification of *Ah* from *R.* UKMP-5M using Q5 DNA polymerase 116
4.3.2 Analytical digestion of pET-28a+ 117
4.3.3 Preparation of inserts for cloning of *Ah* into pET-28a(+) with histidine tags 119
4.3.4 Colony PCR 122
4.3.5 Protein analysis 123
4.3.6 Western blotting of 6X His-tagged Ah expression 126
4.3.7 Improving the solubility of Ah 127
4.3.7.1 Addition of 1% glucose in growth medium 127
4.3.7.2 Transformation of recombinant plasmid into different host strains 128
4.3.7.3 Cloning of Ah into pGEX-6P-1 plasmid 130
4.4 Discussion 133
CHAPTER 5 PURIFICATION AND REFOLDING OF RECOMBINANT AMIDOHYDROLASE 135
5.1 Introduction 135
5.2 Materials and methods 140
5.2.1 Growth and induction 140
5.2.2 Preparation of inclusion bodies 140
5.2.3 Column preparation, purification and refolding 141
5.2.4 Protein identification using nano LC-MS/MS 1411
5.2.5 Protein dialysis 142
5.2.6 Circular dichroism spectroscopy 142
5.2.7 Activity assay for the purified and refolded Ah 142
5.3 Results 143
5.3.1 Refolding of solubilised inclusion bodies 143
5.3.2 Sequence analyses and structure prediction 147
5.3.3 Circular dichroism (CD) spectroscopic analyses 158
5.3.4 Enzyme Assay 160
5.4 Discussion 165

CHAPTER 6 CONCLUSIONS AND FUTURE WORK 169

6.1 Thesis summary 169

6.1.1 Physiological aspect of 1-CB metabolism in *R*. UKMP-5M (Chapter 2) 169

6.1.2 Protein expression analysis using proteomics approach (Chapter 3) 171

6.1.3 A study on the role of a protein that was identified as amidohydrolase 171

(Chapter 4 and 5) 171

6.2 Future Perspective 172

6.2.1 Transcript analysis using qRT-PCR and *Moh* gene knockout experiment 172

6.2.2 Metabolic pathway identification 173

6.2.3 Prediction of function of amidohydrolase (Ah) 175

APPENDICES 176

REFERENCES 187

CHAPTER 1 GENERAL INTRODUCTION

The acceleration of modern industrial activities in recent years has led to the use and liberation of many xenobiotic compounds into the biosphere. Annually, one thousand new chemical compounds are synthesised and between 60 000 and 95 000 chemicals are used for commercial applications. Based on a report by the Third World Network, approximately half a billion kilograms of toxins are released into the air and water environments globally (Shukla, Singh et al. 2010). Moreover, synthesis of new chemicals for agricultural and industrial purposes and the incorrect disposal of these chemicals have contributed towards contamination of various natural resources including groundwater. Among the major xenobiotic compounds reported to date are pesticides, haloalkyl propellants, polychlorinated biphenyls, hydrocarbons and dioxins (Stuart, Lapworth et al. 2012). As such, long term accumulations of these persistent organic pollutants (POPs) pose a great threat to the ecosystem and human health. Global concern to address this issue has led to a number of approaches such as recycling, landfilling, excavation and incineration (Dua, Singh et al. 2002). Alternatively, a greener strategy known as bioremediation has been proposed as an efficient process to restore contaminated environment (Chauhan and Jain 2010).

1.1 Bioremediation

Bioremediation is an innovative technology which uses biological system(s) such as fungi or bacteria to attenuate or eliminate contaminating substances in the environment (Shukla, Singh et al. 2010). The main goal of bioremediation is to stimulate growth of specific microorganism or microbial consortia that are able to destroy chemical contaminants in the environment. These microorganisms may be indigenous i.e. naturally occurring at the contaminated sites, or they may be isolated elsewhere and brought to the contaminated area to perform detoxification (Kumar, Bisht et al. 2011). During bioremediation, organic environmental contaminants can act as a carbon source to provide necessary energy for the

biomass generation. Bioremediation technology can be classified into two groups; ex-situ and in-situ.

Ex-situ bioremediation technology involves excavation of polluted soil to be treated elsewhere using techniques such as landfarming and composting. For example in landfarming, contaminated soils are collected and placed into lined-beds and tilted periodically to provide sufficient aeration. As a result, enhanced microbial degradation of contaminants can be achieved (Vidali 2001). In a composting system, contaminated soils are excavated and mixed with soil amendments such as fertilisers to stimulate and enhance microbial growth to degrade the pollutants (Tomei and Daugulis 2012).

On the other hand, in-situ bioremediation technologies involve the treatment of the contaminants at the site of contamination through techniques like bioventing and biosparging. This technology is considered cost-effective as compared to ex-situ bioremediation due to the lack of soil excavation and transport costs, and the minimal technological equipment necessary for the addition of nutrients and microorganisms (Tomei and Daugulis 2012). For example, bioventing involves the direct injection of oxygen into the unsaturated zone (area between land and surface of groundwater) whereas biosparging involves the injection of oxygen in the saturated zone (area in an aquifer) to enhance the activity of the indigenous bacteria. In the event that the indigenous bacteria lack metabolic capabilities to degrade a target compound, a microflora which have been isolated and cultivated from other sources (exogenous) can be added to the contaminated sites (groundwater or soil) to stimulate degradation (EPA 2013). This process is known as bioaugmentation.

Microbial degradation of environmental contaminants is possible due to the capacity of the microorganisms to produce relevant intracellular or extracellular catabolic enzymes. These enzymes typically have broad substrate specificity to carry out the conversion of organic contaminants to products which are less toxic.

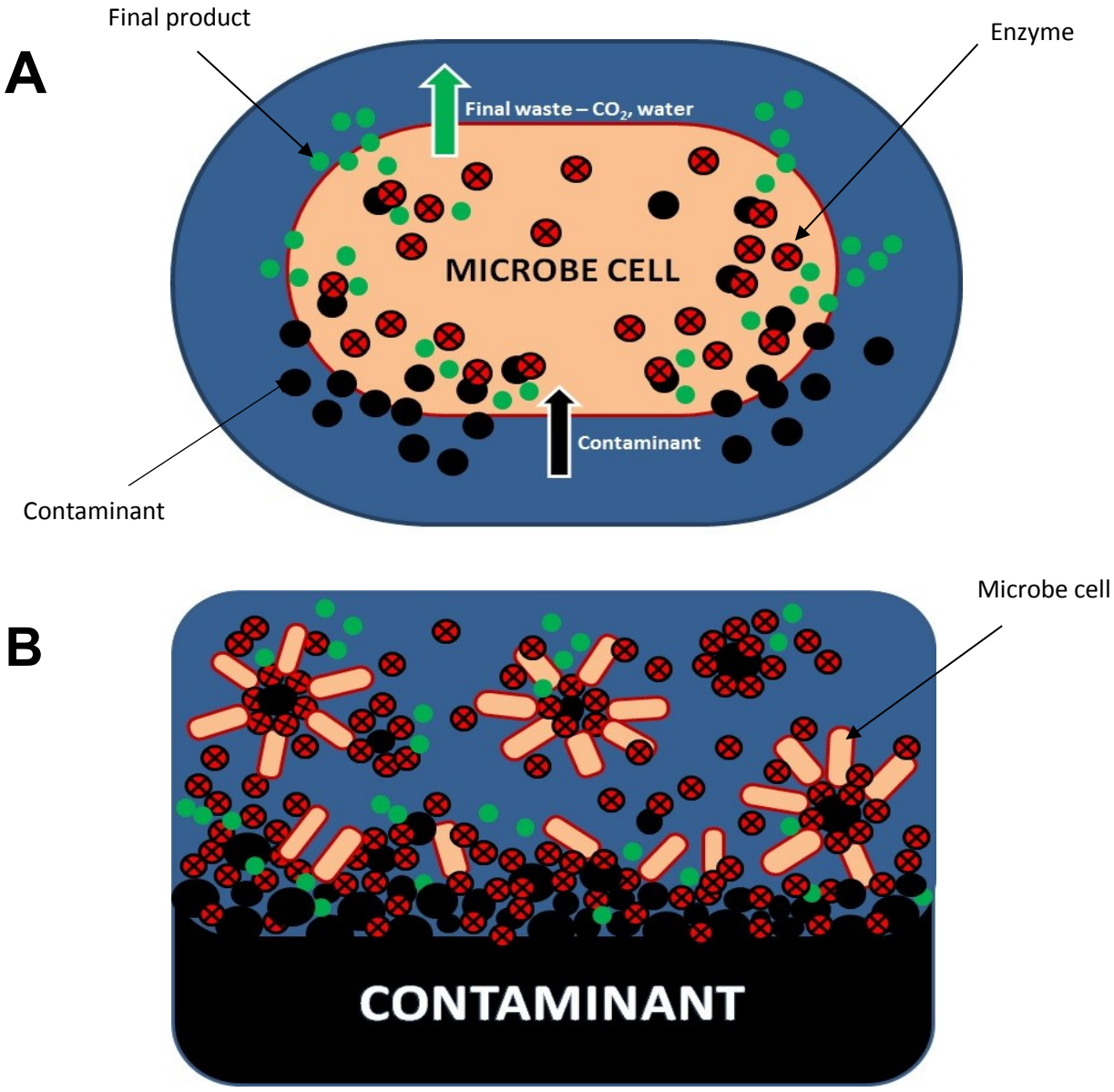


Figure 1.1: Concept of bioremediation. (A) Environmental contaminant enters the microbe cell for intracellular enzyme degradation and the final product is secreted. (B) Extracellular enzymes released by the microbes to digest the contaminants into less toxic form.

The success of bioremediation depends greatly on two factors; the physical conditions of contaminated site or material (e.g. temperature, pH and toxicity level of the contaminant chemical) and the biological factors such as types of microorganisms used and their ability to survive and degrade contaminant chemicals (Mohajeri, Isa et al. 2006).

Microorganisms that have been used to remediate environmental sites can be classified into four main groups: (i) those that are only capable of performing degradation in the presence of oxygen, most of which utilise xenobiotic compounds like pesticides and organic hydrocarbons as their energy source for growth; (ii) those which degrade toxins in the absence of oxygen, majority of which are used for treating halogenated compounds like polychlorobiphenyls and trichloroethylenes (Vidali 2001); (iii) Ligninolytic Fungi – group of fungi (e.g. white rots) which secretes one or more ligninolytic enzymes such as lignin peroxidase and laccase. These enzymes are important for the degradation of recalcitrant compounds such as nitrotoulene and chlorophenols (Novotný, Svobodová et al. 2004); (iv) Methylotrophs – a group of microorganism that is capable of assimilating one-carbon compounds such as methane, methanol and methylamine as their carbon source (Vidali 2001).

For efficient degradation, it is important to ensure that the environmental contaminants and microbe cells are close in proximity. Practically, neither the contaminants nor the microbes are uniformly distributed within the contaminated area. Some bacteria with a chemotactic response sense the contaminant and move towards it to perform degradation. Microbes that are able to grow in a filamentous form can do so towards the contaminant. In cases where this is not possible, the mobilisation of the contaminant can be enhanced by the addition of surfactants to break and evenly distribute the contaminant for microbial attack (Vidali 2001).

1.2 Halogenated compounds

One extremely diverse group of environmental pollutants is halogenated compounds. The occurrence of these compounds in the biosphere can be arbitrarily grouped according to those which are present naturally and those that are foreign to the environment (Field and Sierra-Alvarez 2004). For instance, significant amounts of chloromethane are produced in the ocean by marine algae, wood-rotting fungi and from the natural combustion of plants, wood and soil containing chloride ions. Besides, volcanoes produce significant amount of hydrogen chloride and hydrogen fluoride (up to 3 million tons and 11 million tons per year) and the former may react with other organic compounds to form organochlorines. Examples of such naturally occurring organochlorines include erythrolide C, kumepaloxane, nostocyclophane D, oxypterine, dysidin and polychlorobiphenyls (Gribble 1994).

In addition, anthropogenic impact has also led to the production and discharge of various kinds of halogenated compounds into the biosphere. The stable and inert nature of highly chlorinated compounds appeals in many industrial and agricultural applications. Halogenated compounds like polychlorinated biphenyls (PCBs) are designed to be extremely stable and are often used as plasticizers and flame retardants for household and industrial uses (Posada 2006). It was reported that the amount of PCBs distributed in the environment before their ban in 1970s was approximately 0.6 million metric tonnes. Despite the ban on the production and usage of PCBs, these compounds are still being detected in various environmental samples and living organisms due to their persistence and bioaccumulation properties. Many recent studies have shown that human serum samples from all age groups were found to contain PCBs (Eguchi, Nomiyama et al. 2012, Marek, Thorne et al. 2013).

On the other hand, some commonly used pesticides and insecticides such as atrazine, chloropropionic acid and 2,4-dichlorophenoxyacetate have also posed significant threats to the environment. For example, atrazine which is widely used to control weeds and

grasses is the most frequently detected contaminant in European groundwater (Figure 1.2). Studies have shown that the persistence of atrazine in the environment have caused detrimental impact on the aquatic life (Graymore, Stagnitti et al. 2001) and affects the human reproductive, central nervous and immune system (Liu, Li et al. 2014). In Malaysia, some chlorinated compounds which are used for pest control include dieldrin, heptachlor, lindane and dichlorodiphenyltrichloroethane (DDT). Although the usage of these compounds have been restricted since 1990s, reports have shown that they are still present in the river water at detectable levels (Santhi and Mustafa 2013).

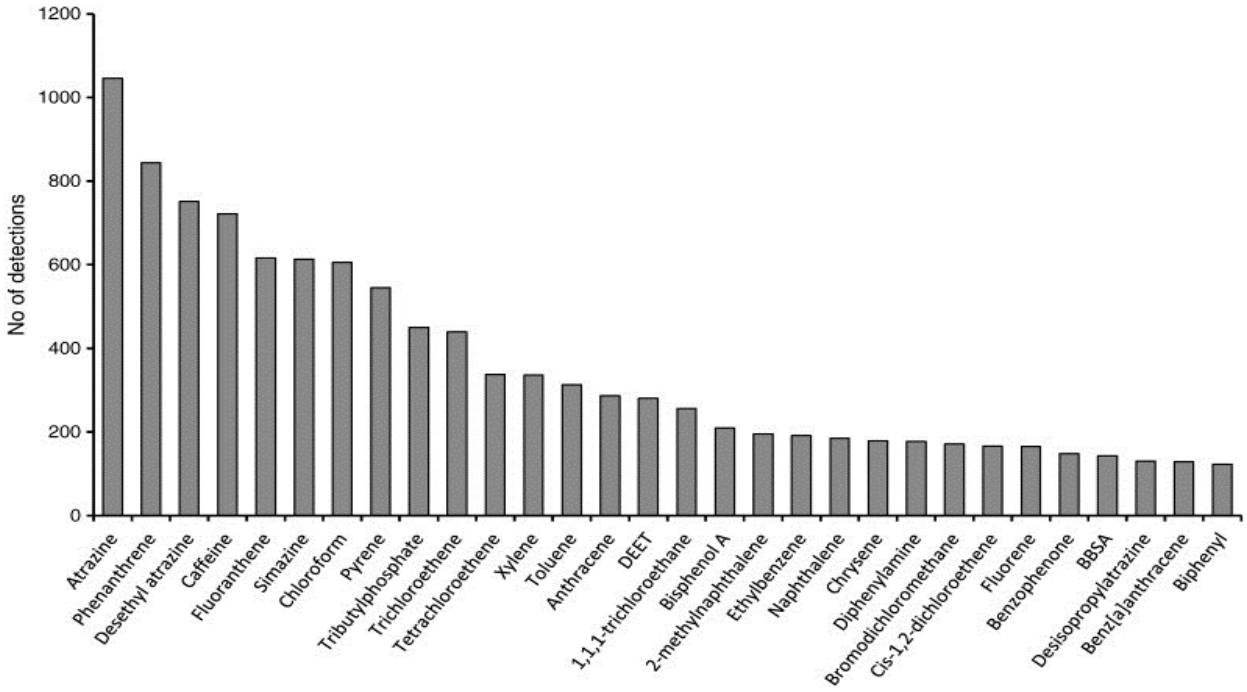


Figure 1.2: The most frequently detected organic compounds in the European groundwater reported in Environmental Agency database (DEET = *N*,*N*-dimethyl-toluamide; BASA = *N*-buthylbenzene sulphonamide; BBSA = *N*-buthylbenzene sulphonamide)(Stuart, Lapworth et al. 2012). *Used with permission from the publisher, Elsevier.*

1.2.1 Haloalkanes, their properties and routes of exposures

Haloalkanes are a group of chemical compounds where one or more hydrogen atoms in the carbon chain have been replaced by halogen atom(s) such as fluorine, chlorine, bromine or iodine. Of the various haloalkanes produced, the simplest are chlorinated methanes with productions of 100 to 550 kilotonnes per year (Field and Sierra-Alvarez 2004). All four chlorinated methanes, chloromethane (CM), dichloromethane (DCM), trichloromethane (chloroform, CF) and tetrachloromethane (carbon tetrachloride, CT) are used as solvents and their industrial production is done through thermal, catalytic or photolytic chlorination (Trotsenko and Torgonskaya 2009). In addition, natural processes such as volcanic eruptions and atmospheric emissions are the main sources of CM occurrence in the biosphere (Gribble 1994). CM is an extremely flammable gas and other chlorinated methanes (DCM, CF and CT) visible-clear oily liquids with sweet odour (Kayser 2001, Faroon, Taylor et al. 2005). Chlorinated ethanes are also produced in large quantities for various industrial purposes. The combined production of chloroethanes in Europe, Asia and United States of America was more than 14.7 million tonnes in year 2006 [reviewed in (Govender 2008)].

The inertness of chloroalkanes makes them persist in the groundwater for long periods of time (Faroon, Taylor et al. 2005) and prolonged exposure to these chemicals may cause adverse effects on both man and aquatic species. The frequency of detection of some haloalkanes in the European groundwater is shown in Figure 1.2 and the major uses of some chlorinated methanes and ethanes are summarised in Table 1.1. In general, there are several routes of exposure to haloalkanes. Humans are exposed to these compounds by: (i) inhalation; (ii) ingestion through contaminated drinking water or (iii) dermal contact through showering and swimming (Kayser 2001, Faroon, Taylor et al. 2005). Common illness reported following human exposed to haloalkanes are often associated with internal organs such as the nervous system, kidneys, liver and lung. Skin absorption of chloromethanes may

cause irritation and gastrointestinal effects like nausea and vomiting (Zarogiannis, Nwaogu et al. 2007).

Table 1.1: A summary of the major uses of some chlorinated methanes and ethanes (Bhatt, Praveena et al. 2007). *Adapted with permission from the publisher, Taylor & Francis Group.*

Haloalkanes	Major Uses
Monochloromethane	Used in the production of silicones, tetramethyllead, methylcellulose and in other methylation reactions
Dichloromethane	Used as degreasing agent, paint remover, pressure mediator in aerosols; extract technology
Trichloromethane and Tetrachloromethane	Used in the production of monochlorodifluoromethane (for the production of tetrafluoroethene, which is used for the manufacture of Hostaflon and Teflon), extractant for pharmaceutical products
Monochloroethane	Used in the production of tetraethyllead, production of ethylcellulose; as ethylating agent for fine chemical production, solvent for extracting processes
1,1- Dichloroethane	Used as feedstock for the production of 1,1,1-trichloroethane
1,2-Dichloroethane	Used in the production of vinyl chloride, production of chlorinated solvents such as 1,1,1-trichloroethane and tri- and tetrachloroethane, synthesis of diethylenediamines
1,1,1-Trichloroethane	Used as dry cleaning, vapor degreasing, solvent for adhesives and metal cutting fluids; textile processing
1,1,2-Trichloroethane	Used as an intermediate for production of 1,1,1-trichloroethane and 1,1-dichloroethane

1.3 Microbial degradation of halogenated compounds

Microorganisms are the primary source of biocatalysts (enzymes) that play a vital role in degradation and utilisation of environmental pollutants. This is possible due to the presence of catabolic genes in microorganisms that encode for different enzymes to carry out dehalogenation reactions. Organic compounds in general can be mineralised by microorganisms to produce carbon dioxide and energy for cellular processes. However, no single organism possesses the ability to degrade every compound present in the biosphere. Some chlorinated compounds such as CF and 1,1,1-trichloroethane are not readily degradable by microorganism. In such cases, prolonged exposure to these xenobiotics provides a selective advantages for the microorganisms that have metabolic pathways to be able to utilise new substrates (Baptista 2008).

The evolution of metabolic pathways in microorganisms can be achieved either; (i) through genetic changes that occur within the microbial genome which includes events like mutations of existing genes or recombination events that lead to changes in gene arrangements; or (ii) by acquisition of novel genes from phylogenetically related or distinct organisms through events like horizontal gene transfer (HGT). HGT contributes to the diversification of microorganisms to degrade xenobiotic contaminants in the environment. Genes that encode for enzymes that are capable of degrading recalcitrant compounds are often found in mobile genetic elements such as plasmids and through methods like conjugation, transformation or transduction, these plasmids can be transferred into other organisms (Fulthorpe and Top 2010).

The most efficient route of HGT is through conjugation, whereby the genetic dissemination occurs when two bacteria are physically in contact and the donor cell transfers the genetic material (i.e. plasmids or DNA segments) to the recipients via sexual pilus. On the other hand, HGT through transformation is the genetic alteration that occurs in the naturally transforming bacterial as a result of the uptake, integration and expression of naked

DNA molecules; and transduction is the genetic material transfer from the donor cell using a bacteriophage as a carrier to infect the recipient cell.

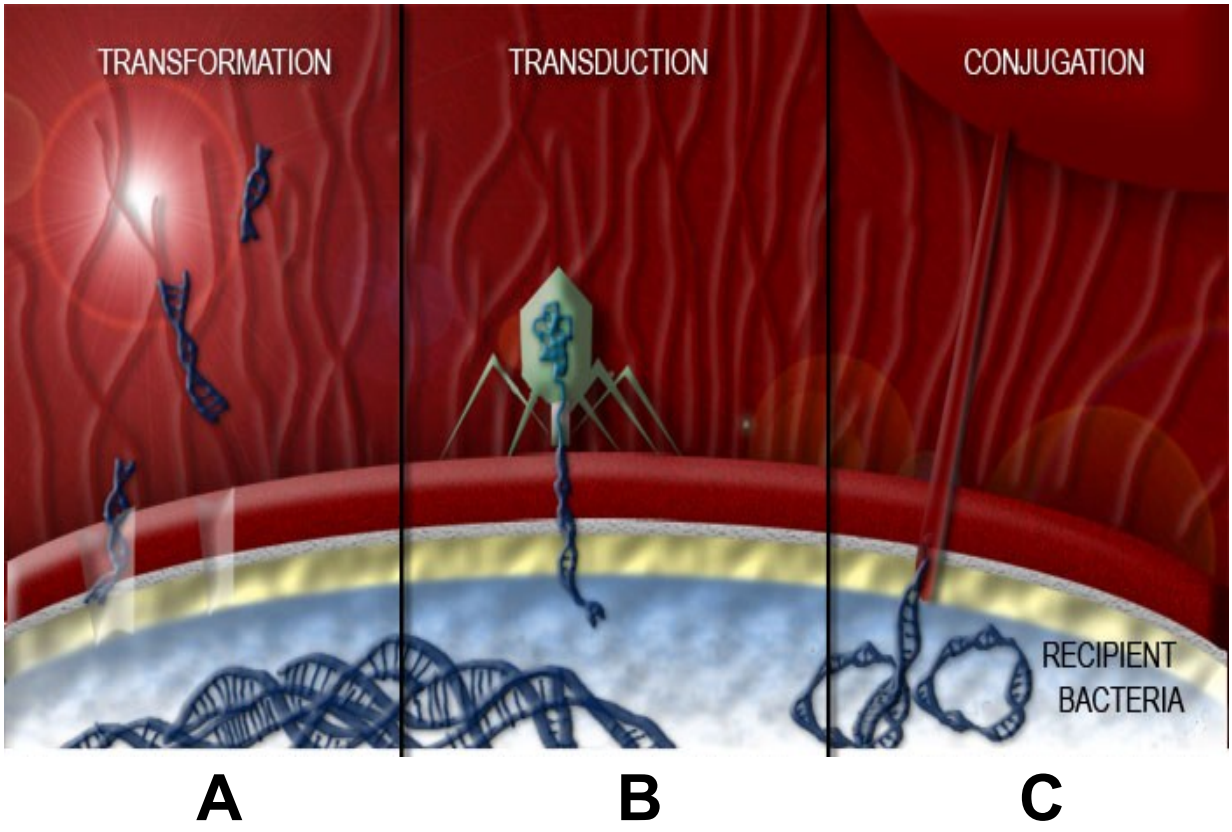


Figure 1.3: The common routes of horizontal gene transfer. Gene transfer through (A) transformation which involves the uptake of extracellular naked DNA molecule by the recipient cell; (B) transduction which requires a bacteriophage to transfer heterologous DNA molecule from the donor cell to the recipient cell; and (C) conjugation that requires a cell-to-cell contact between the donor cells and the recipient cells and the DNA transfer occurs via a sex pilus.

Used from http://amrls.cvm.msu.edu/microbiology (last accessed 25/10/2014).

In general, metabolism of halogenated compounds in microorganism can be classified into two groups: (i) complete mineralization by utilization of the organic compounds as sole carbon and energy source; or (ii) complete or partial degradation of organic compounds through co-metabolism.

1.3.1 Utilisation of halogenated compounds as carbon and energy source.

Some organic compounds can be degraded and utilised as sole carbon and energy source for bacterial growth. Most lower halogenated alkanes i.e. compounds containing one or two halogens can be assimilated for biomass generation. Several studies have demonstrated the ability of facultative methylobacteria to aerobically degrade haloalkanes like chloromethane (CM) and dichloromethane (DCM) for energy production. For example, *Methylobacterium rhodesianum* H13 was reported to completely degrade up to 10 mM DCM in less than 24 hours. It was reported that approximately 11 g dry cell weight of cells per mol DCM was achieved with concomitant release of chloride (Chen, Ouyang et al. 2014). In another study, several *Hyphomicrobium* strains were isolated from enrichment cultures containing 1.3% (v/v) chloromethane as sole carbon source. A CM dependent cell growth consistent with the production of chloride was observed (Nadalig, Farhan Ul Haque et al. 2011). Besides haloalkanes, several other halogenated compounds were also observed to support bacterial growth. For an example, *Xanthobacter autotrophicus* GJ10 was able to grow on monochloroacetic acid (MCA) as sole carbon source. A 1,2-dichloroethane grown inoculum was able to degrade up to 49 mM MCA with the production of chloride and carbon dioxide.(Torz and Beschkov 2005).

1.3.2 Co-metabolism

In this mechanism, the microorganism reacts by only altering the structure of a chemical compound without deriving energy from this catabolic activity. Instead, it grows on alternative substrates that are present (Criddle 1993, Jing 2007). This mechanism takes place only when there is no other route available to metabolise the target chemical (Leisinger 1996). Higher chlorinated compounds like chloroform (CF) and carbon tetrachloride (CT) undergo this type of metabolism (Field and Sierra-Alvarez 2004). For example, CF and 1,1,1-trichloroethane (TCA) were reported to be aerobically degraded by a microbial consortium grown on different co-metabolic substrates such as methane, propane and butane. Among all the carbon sources tested, butane microcosm resulted in the most effective transformation of CF and TCA. The addition of butane and CF or TCA in the culture medium resulted in the complete degradation of CF but only partial degradation of TCA (70%) was achieved. The strong correlation between the rate of transformation with butane utilization and oxygen consumption indicated that aerobic co-metabolism had occurred. (Kim, Semprini et al. 1997).

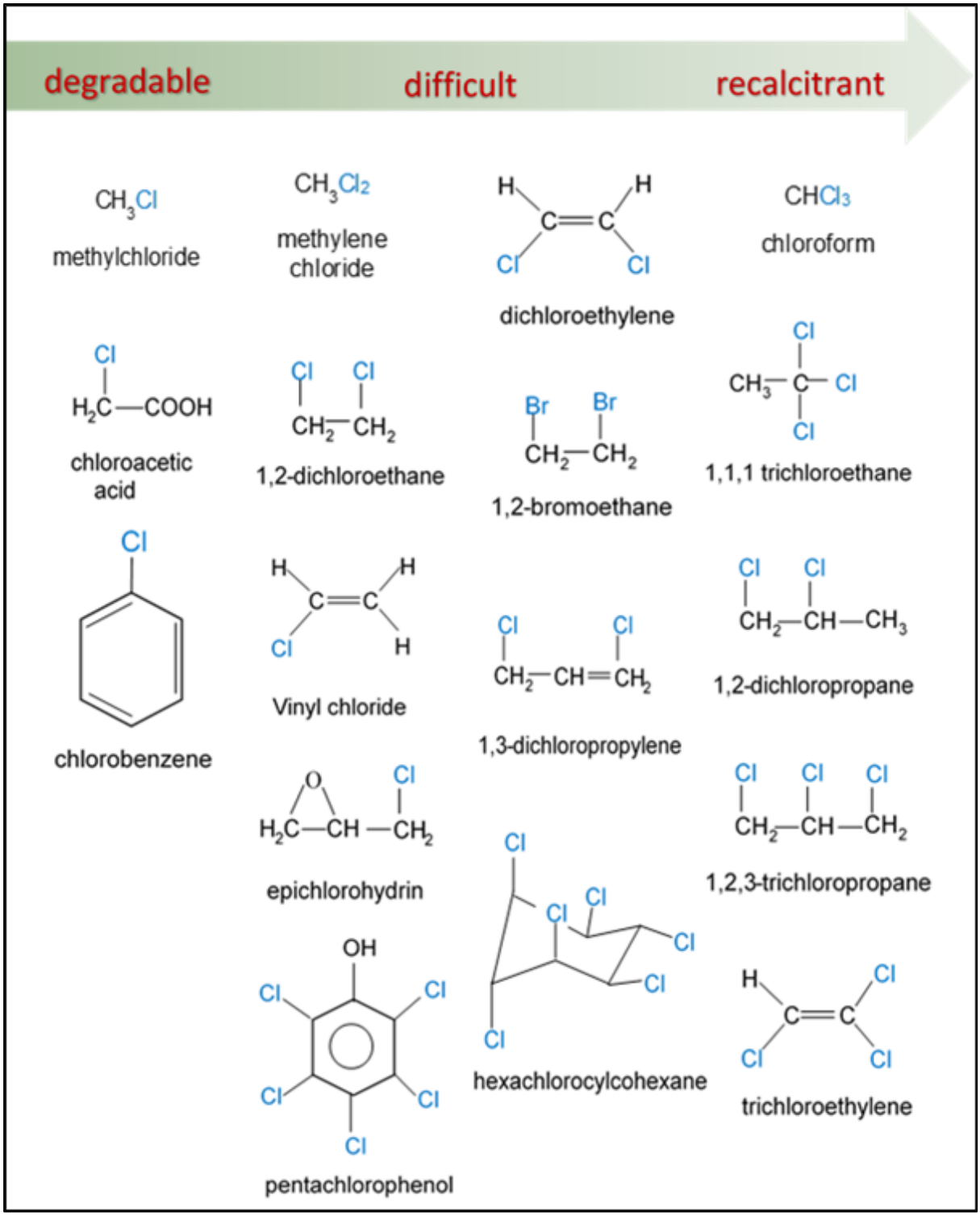


Figure 1.4: Degradability of some halogenated compounds by microorganism. Compounds on the extreme left have been reported to be readily degradable and utilised as carbon sources by microorganisms. The compounds on the extreme right have been shown to be recalcitrant towards degradation. (Dick B. Janssen 2005). *Adapted with permission from the publisher, John Wiley and Sons.*

1.4 Bacterial dehalogenation reactions

Dehalogenation is the process of halogen removal from organic compounds. Microbial degradation of these compounds requires the presence of catabolic enzymes that are able to cleave the carbon-halogen bonds, which can be achieved through various mechanisms. Examples of several mechanisms of dehalogenation are discussed in this section.

(i) Hydrolytic Dehalogenation

One of the well-studied hydrolytic dehalogenases is the haloalkane dehalogenase from *Xanthobacter autotrophicus* GJ10 which is able to degrade and utilise 1,2-dichloroethane as carbon and energy source (Keuning, Janssen et al. 1985). In this reaction, the enzyme catalyses the substitution of the halogen by a hydroxyl group derived from water to yield the corresponding alcohol and no other co-factor is required for the reaction to take place.

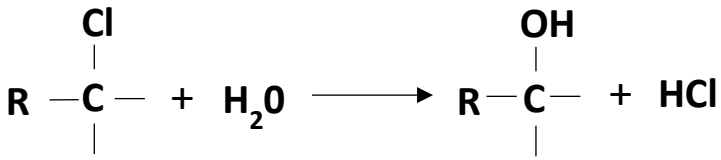


Figure 1.5: Reaction of hydrolytic dehalogenation. In this reaction, the halogen substituent is replaced by a hydroxyl group through nucleophilic displacement reaction.

(ii) Thiolytic dehalogenation

In facultative methylotrophic bacteria such as *Methylophilus sp*. and *Methylobacterium sp*., the dechlorination of dichloromethane (DCM) occurs via thiolytic dehalogenation. In this reaction, the inducible glutathione S-transferase catalyses the formation of the unstable S-chloromethylgluthathione conjugate, which is subsequently hydrolysed to form gluthathione and formaldehyde, with the release of chloride. Formaldehyde, a primary metabolite for methylotrophic bacteria, is then be assimilated for growth (Fetzner 1998).

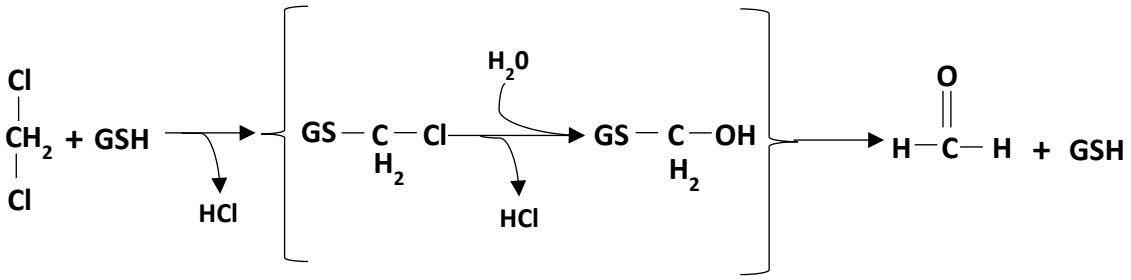


Figure 1.6: Mechanism of thiolytic dehalogenation. In this reaction, gluthathione S-transferase catalyses the formation of formaldehyde via an unstable intermediate known as S-chloromethylgluthathione.

(iii) Dehalogenation by methyl transfer

Dehalogenation by methyl transfer occurs in bacteria which utilise lower chlorinated methanes i.e. chloromethane or dichloromethane as sole carbon source. Examples of aerobic bacteria that are able to use chloromethanes as carbon source include *Methylobacterium chloromethanicum* CM4 and *Hyphomicrobium chloromethanicum* CM2. A methyltransferase isolated from a methylotrophic bacterium CC495 contains a corrinoid-bound cobalt atom. In the active state, cob(II)alamin is oxidised by methylation by halomethane to form the methyl-cob(III)alamin conjugate and in the presence of a methyl group acceptor, this conjugate will be reduced by demethylation.

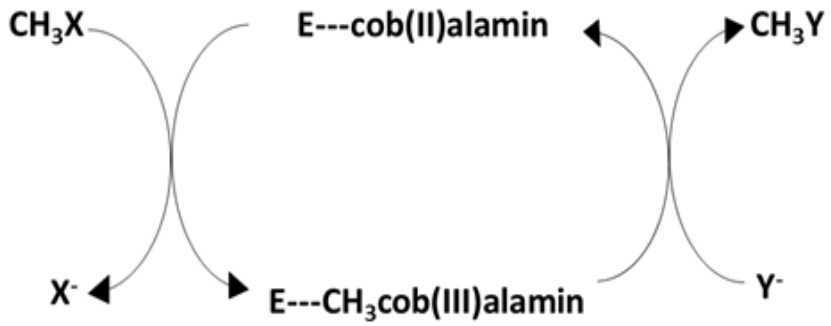


Figure 1.7: Dehalogenation of halomethane catalysed by a methyltransferase. CH_3X can be CH_3Cl, CH_3Br or CH_3I and Y- can be any methyl group acceptor such as Cl^-, Br^- or HS^- (van Pee and Unversucht 2003).

(iv) Dehalogenation by hydration

This type of dehalogenation reaction is more common for haloalkenes. In the reaction shown in Figure 1.8, hydratase catalyses the addition of water molecule to the double bond in a vinylic compound (3-chloroacrylic acid) to form malonate semialdehyde. An example of bacterium that is able to catalyse this reaction is *Pseudomanas chicorii* 170 (Poelarends, Wilkens et al. 1998).

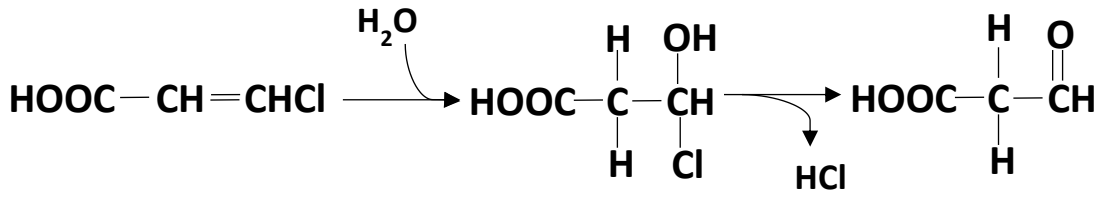
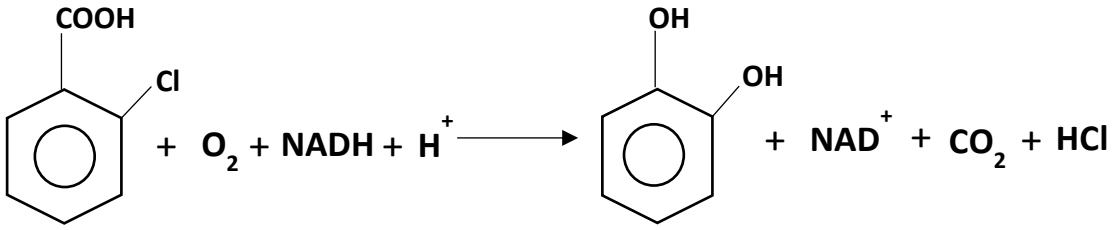


Figure 1.8: Mechanism of dehalogenation through hydration. In this reaction, addition of water molecule on to the unsaturated bond releases chloride from the halogenated compound (van Pee and Unversucht 2003).

(v) Oxidative dehalogenation

Oxidative dehalogenation is an oxygen dependent reaction that can occur in many bacteria including *Rhodococcus (Bondar, Boersma et al. 1999)*. This reaction can be catalysed either by monooxygenases or dioxygenases, which incorporate one or two oxygen atoms into halo-organic compounds, in the presence of cofactors such as NADH or NADPH.

A



B

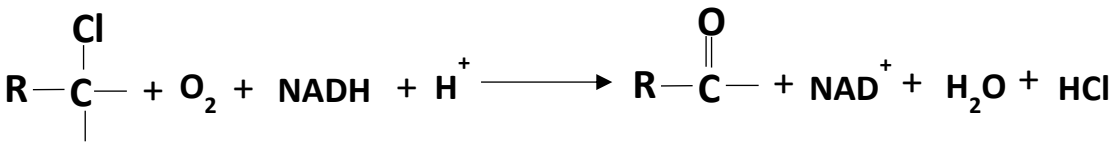


Figure 1.9: Mechanism of oxidative dehalogenation. (A) A dioxygenase catalyses the transformation of 2-chlorobenzoate to catechol through insertion of two atoms from molecular oxygen (Fetzner and Lingens 1994); and (B) A monooxygenase catalyses the oxidation of chlorinated alkane through insertion of one oxygen atom (Janssen, Pries et al. 1994).

(vi) Dehydrohalogenation

In this dehalogenation reaction, dehydrohalogenases eliminates hydrochloric acid from their halogenated substrate, which results in the formation of unsaturated bond. For example, the first two steps in the dehalogenation of γ-hexachlorocyclohexane (lindane) in *Sphingomonas paucimobilis* UT26 are catalysed by dehydrohalogenase (LinA) followed by two hydrolytic steps catalysed by hydrolytic enzymes (LinB) as shown in Figure 1.10.

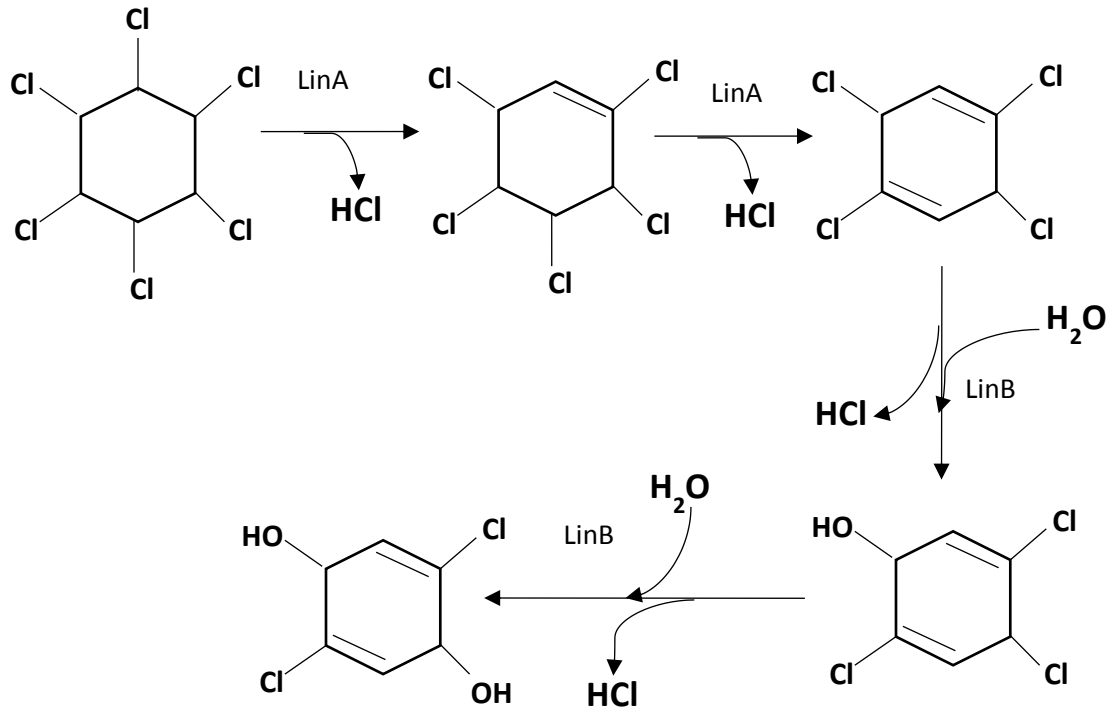


Figure 1.10: Dehalogenation of lindane by dehydrohalogenase. This is an example of dehydrohalogenation that occurs in *Sphingomonas paucimobilis* UT26, that leads to the formation of unsaturated carbon bond (Janssen, Oppentocht et al. 2001).

(vii) Intramolecular substitution

This reaction involves an enzyme known as haloalcohol dehalogenase (also known as halohydrin hydrogen-lyases) which catalyses the intramolecular nucleophilic replacement of a halogen substituent by a vicinal hydroxyl group in haloalcohols to form epoxides. Despite having a sequence homology with the members of short chain dehydrogenases, haloalcohol dehydrogenase does not require a cofactor for dehalogenation reaction. This enzyme is also able to catalyse the reverse reaction, which is the halogenation of epoxides to form haloalcohol in the presence of halide. A well-studied bacterium that is able to carry out this reaction is *Agrobacterium radiobacter* AD1 (de Jong, Tiesinga et al. 2003).

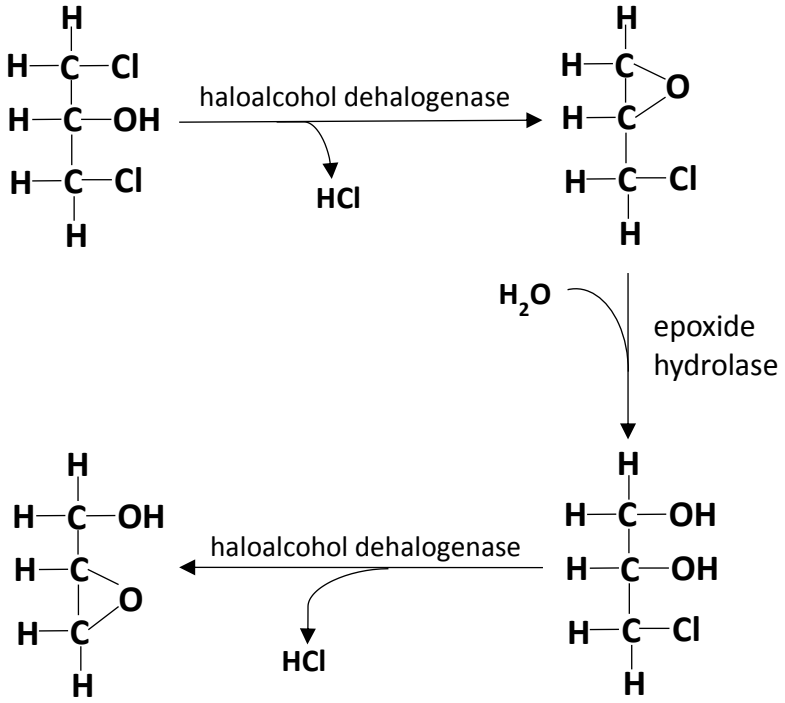


Figure 1.11: Mechanism of dehalogenation through intramolecular substitution reaction. In this reaction, haloalcohol dehalogenase catalyses the removal of halogen by vicinal hydroxyl group in halohydrin to form epoxides (van Pee and Unversucht 2003).

(viii) Reductive dehalogenation

Reductive dehalogenation can occur under both aerobic and anaerobic conditions. This reaction involves the release of halide and its replacement by hydrogen. In highly chlorinated compounds such as hexachlorobenzene and tetrachloroethylene, the reductive dehalogenation in the initial step plays a major role for the metabolism of these compounds. An example of reductive dechlorination in aerobic conditions has been demonstrated in a study of pentachlorophenol degradation by a monooxygenase from *Sphingomonas chlorophenolica* that leads to the formation of tetrachlorohydroquinone (TCHQ). The dehalogenation of TCHQ was catalysed by a TCHQ dehalogenase, a member of the gluthathione S-transferase family, to form trichlorohydroquinone via reductive dehalogenation.

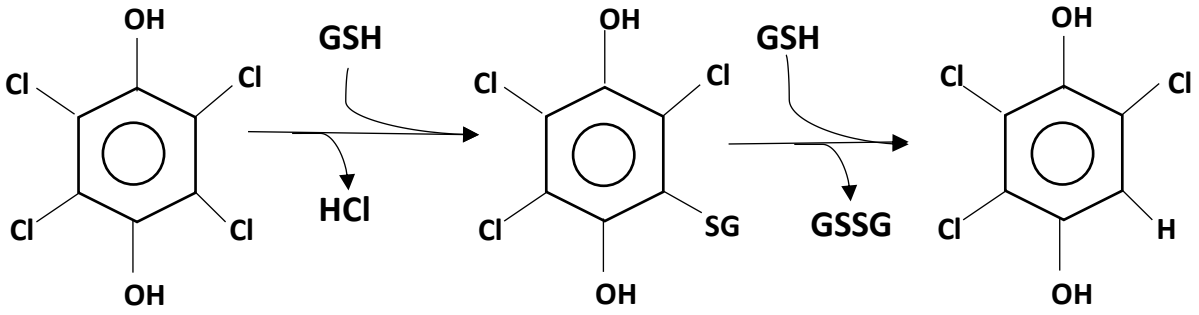


Figure 1.12A: Mechanism of reductive dehalogenation of tetrachlorohydroquinone (TCHQ). This reaction is catalysed by a glutathione S-transferase-type TCHQ dehalogenase (Anandarajah, Kiefer et al. 2000).

On the other hand, some halorespiring bacteria are able to couple the reductive dehalogenation of halogenated organic compounds to energy production and growth under anaerobic conditions. Examples of such bacteria include *Dehalococcoides* sp. and *Desulfitobacterium* sp. In this reaction, halogenated compounds like tetrachloroethylene serve as a terminal electron acceptor during the oxidation of an electron rich compound such as hydrogen.

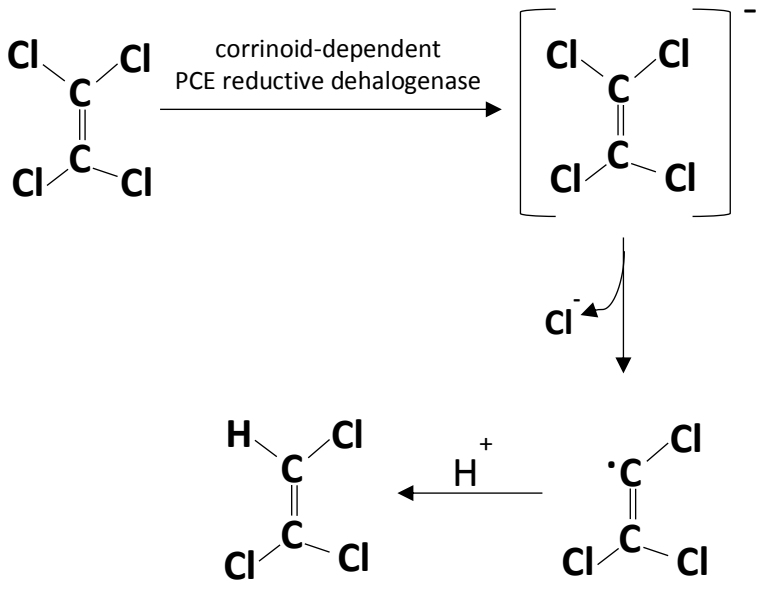


Figure 1.12B: Reductive dehalogenation of tetrachloroethylene (PCE). The reaction is coupled to energy metabolism under anaerobic conditions and catalysed by a corrinoid-dependent reductive dehalogenase (van Pee and Unversucht 2003).

1.5 The genus *Rhodococcus*

Members of the genus *Rhodococcus* are a diverse group of actinomycetes found in various environmental niches such as boreholes, rocks, marine sediments, seawater, groundwater, on the rhizosphere and surfaces of plants and at polluted sites (Bell, Philp et al. 1998). This group of bacteria exhibit a remarkable ability to degrade various recalcitrant compounds ranging from simple hydrocarbons to more complex compounds like polychlorobiphenyl (PCB) (Vellore 2001). Studies on the biodegradation potential of *Rhodococcus* have intensified significantly over the past 25 years as shown by the increasing pattern of publications.

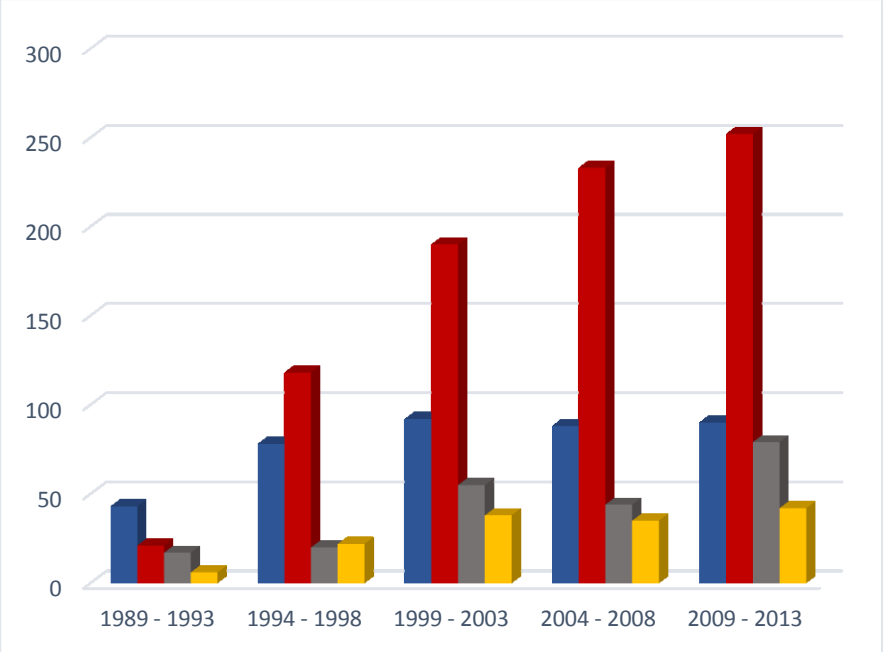


Figure 1.13: The increasing pattern of number of publications concerned with degradation of nitriles, aromatic and halogenated compounds by *Rhodococci* (according to www.scopus.com last accessed 20/10/14). Queries: Title/ Abstract/ Keywords: *Rhodococcus* and (■) nitrile; (■) benzene, toluene, xylene, naphthalene, biphenyl or hydrocarbon; (■) phenol and (■) halogenated compounds.

The name *Rhodococcus* was first proposed by Zoph in 1891, as reviewed in (Bell, Philp et al. 1998). It originates from Greek terms called 'Rhodon' and 'kokkus' which means "rose grain" (Alvarez 2010). Initial classifications of *Rhodococcus* were done based upon their cell morphology and acid-fast staining. There has been some confusion about their taxonomic positions over many years. Goodfellow in 1989 resolved the taxonomical classification of *Rhodococcus* based on their chemotaxanomic characteristics and later, a molecular approach was applied to reclassifying the members in this genus (Larkin, Kulakov et al. 2010). The current genus of *Rhodococcus* is defined on the basis of cell wall composition (particularly the presence of long chain mycolic acids with 34-52 carbon atoms, carbohydrates such as arabinose, galactose and the presence of dihydrogenated menaquinones with 8 isoprene units) and supported by 16S rRNA sequence analysis to confirm the taxonomic positions of the members (Finnerty 1992). Currently, the members of genus *Rhodoccoccus* belong to the mycolic acid containing suborder *Corynebacterineae* in the family *Nocardiaceae*, within the phylum *Actinobacteria* (Alvarez 2010). At the time of writing this thesis, there are 53 established *Rhodococcus* species isolated from various sources (http://www.bacterio.net/*Rhodococcus*.html last accessed 25/10/14). Although few species of *Rhodococcus* such as *Rhodococcus equi* and *Rhodococcus fascians* are reported to be pathogenic, the majority of the members appear to be benign soil inhabitants.

1.5.1 Characteristics of *Rhodococcus*

Rhodococcus is an aerobic, gram positive, non-motile, non-sporulating and catalase positive microorganism with high GC content in its genome (Bell, Philp et al. 1998). These actinomycetes exhibit variety of growth patterns. Depending on their growth rates and the supplemented carbon source, they may form short projecting, elementary branching or extremely branched multinucleated filaments that are capable of undergoing fragmentation into rod shaped or coccoid elements (Ganguly 2005) as shown in Figure 1.14A and 1.14B.

The presence of carotenoids in the bacterial cells gives a distinct red, orange or salmon pink pigmentation to the cells of *Rhodococcus* (Vellore 2001).

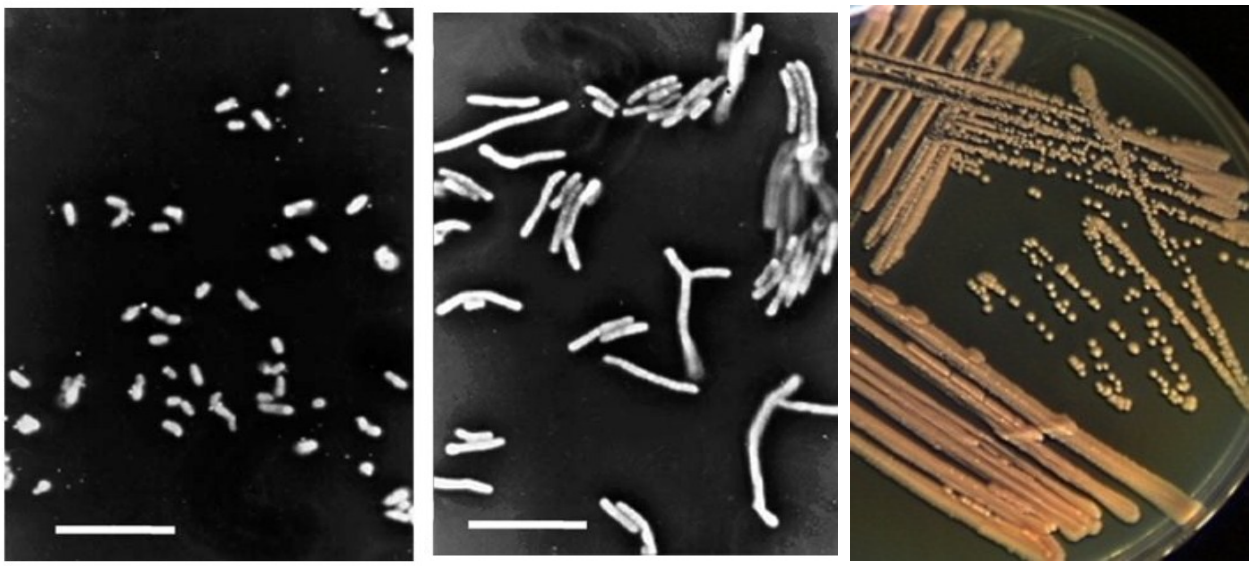


Figure 1.14: Different morphological forms of *Rhodococcus*. (A) Short rods and cocci shapes of *Rhodococcus* NCIMB 12038 when growing at low growth rate, (B) and branched filaments when growing at high growth rate and (C) orange colonies of *Rhodococcus* UKMP-5M on LB agar. *Figures 1.14A and B were used from (Larkin, Kulakov et al. 2010) with permission from the publisher, Springer.*

Cell proliferation of *Rhodococcus* is relatively slow when compared to other biodegradative bacteria like *Pseudomonas* and *Xanthobacter*. Most *Rhodococci* are capable of growing in minimal salt media (MSM) supplemented with a carbon source. However, cells grown in MSM often lack of pigmentation and appear paler in colour than the same cells grown in nutrient rich media such as Luria Bertani (LB) (Larkin, Kulakov et al. 2010). The bacterium forms colonies which are entire, smooth with large diameter ranging from 2 to 4 mm on an LB agar (Vellore 2001). The presence of mycolic acid in the cell wall of

Rhodococci facilitates the uptake of various hydrophobic substrates in to the cells and prevents desiccation of the bacteria for long periods.

These bacteria are also frequently reported to tolerate low temperatures and high pH during the metabolism of toxic chemicals, thus enabling their application in harsh, extremely cold or dry environmental conditions. For example, in a study on phenol biodegradation at low temperature, two strains of *Rhodococcus* were among the four strains which are able to degrade more than 10 mM phenol at temperature as low as 10^0C (Margesin, Fonteyne et al. 2005). Besides that, members of this genus have high adaptability and versatility in degrading wide range of environmental contaminants. The genetic basis of these abilities derives from their capacity to accommodate large plasmids with many catabolic genes; for example, one of the plasmids in *Rhodococcus jostii* RHA 1 has a size of 1.1Mb (McLeod, Warren et al. 2006).

1.5.2 Significance of *Rhodococcus*

The environmental and economic importance of *Rhodococcus* is being increasingly recognised. *Rhodococci* exhibit diverse metabolic capacities to degrade wide class of contaminants in different environments. Their remarkable range of catabolic genes makes them potentially useful for the biotransformation of many chemicals regardless of their recalcitrance or toxicity. Some of these transformations can be exploited for industrial and commercial purposes (Martinkova, Uhnakova et al. 2009).

Crude oil spillage is one of the major causes for chronic pollutions of marine ecosystem by hydrocarbons. For example, in year 2000, two large oil spillage incidents in Cape Town threatened the penguin population in South Africa (Singh and Nagaraj 2006). Studies have demonstrated that *Rhodococcus* has the potential to degrade various oil pollutants. For example, a single strain *R. baikonurensis* EN3 is capable of degrading up to 20 g/L diesel oil (Lee, Kim et al. 2006). The presence of branched alkanes like 2, 6, 10, 14-

tetramethylhexadecane (phytane) and 2, 6, 10, 14-tetramethylpentadecane (pristane) in crude oil further exacerbates its recalcitrance to degradation compared to linear alkanes. A second example of the metabolic versatility of *Rhodococcus* is *Rhodococcus* NTU-1 which is reported to be capable of degrading both hexadecane and pristane via a hydroxylase system at low temperature (Sayavedra-Soto, Chang et al. 2006).

Studies have shown that *Rhodococci* are also capable of degrading various aromatic compounds. Accumulation of aromatic compounds in the environment has become a major threat to the ecosystems. Due to the recalcitrance of aromatic compounds in nature, only a small group of bacteria is capable of mineralising them. *Rhodococcus* sp. RHA1 was isolated from lindane contaminated soil and is capable of degrading aromatic compounds such as PCB (Masai, Yamada et al. 1995), phenylacetate (Navarro-Llorens, Patrauchan et al. 2005), lignin (Ahmad, Roberts et al. 2011) and benzonitrile (Kitagawa, Miyauchi et al. 2001). Research revealed that this bacterium uses a series of oxygenases to substitute the hydroxyl group into the aromatic ring to yield intermediates like catechols which were cleaved through -ortho or -meta cleavage and for entry into different catabolic pathways (Masai, Yamada et al. 1995).

Some studies have also demonstrated the ability of *Rhodococcus* to degrade halogenated compounds. For example, chloropropionic acid is often used as herbicide in the agricultural industry. The long persistence of this compound in the environment may cause carcinogenic effects in animals and humans. *Rhodococcus sp*. HJ1 isolated from the agricultural soil was found to completely degrade and utilise up to 20 mM 3-chloropropionic acid as the sole carbon source. This strain was also tested for its ability to dehalogenate several other halogenated substrates. Among the substrates which were successfully dechlorinated in addition to 3-chloropropionic acid were 3-chlorolactic acid, 2,3-dichloropropionic acid, 3-chlorobutyric acid and 2,2,2-trichlorobutyric acid (Jing 2007). In a recent study, the ability of *Rhodococcus* FP1 to degrade a number of fluorinated compounds was tested. Fluorinated compounds are used extensively in the field of agrochemicals and

pharmaceuticals and their long persistence in the environment is a problem. *Rhodococcus* FP1 was able to utilise up to 0.6 mM 2-fluorophenol as carbon and energy source without any growth inhibition. Some other halogenated compounds degraded by FP1 include 2-fluorobenzene, 2-chlorophenol, 4-chlorophenol and 3-fluorocatechol. The isolation of cell free activity with the addition of FAD as the cofactor suggests that the degradation of 2-flurophenol is mediated by a monooxygenase (Duque, Hasan et al. 2012).

The biodegradative abilities of *Rhodococcus* are also associated with their ability to produce biosurfactants. Surfactants are reagents which disperse the hydrophobic compounds for microbial attack in water and so make them more accessible to degradation (Bell, Philp et al. 1998). Biosurfactants produced by microorganism including *Rhodococci* are less harmful in terms of toxicity and biodegradability compared to the existing synthetic surfactants. Important uses of biosurfactants are in the field of agriculture, pharmaceutical and cosmetics (Pacwa-Plociniczak, Plaza et al. 2011). Moreover, *Rhodococci* also have the ability to produce bioflocculants. Flocculants consist of peptides and lipid, which can cause flocculation of wide variety of suspended solid. Such material is important in the removal of suspended solids in wastewater or effluent treatment (Bell, Philp et al. 1998, Vellore 2001).

1.5.3 *Rhodococcus* UKMP-5M

Rhodococcus UKMP-5M was isolated from petroleum contaminated soil in Negeri Sembilan, Malaysia. The first draft of the genome sequence for this bacterium was completed in collaboration with University of Cambridge in 2008. Analysis of the genome sequence revealed many catabolic genes for the degradation of xenobiotics including phenols, epoxides, cyanide and nitrilated compounds.

These degradative abilities were then shown in several published studies. For example, in a study on degradation of cyanide by several isolates of *Rhodococcus*, *R*. UKMP-5M was reported to exhibit the most efficient removal of cyanide. Among all the

cyanide degrading bacteria reported to date, *R*. UKMP-5M is able to tolerate and degrade up to 312 mg/L cyanide (Maegala Nallapan Maniyam 2011). A second example is the ability of *R*. UKMP-5M to degrade and utilise up to 0.5 g/L phenol (Suhaila, Rosfarizan et al. 2013).

1.6 Background to the Project

With the emerging issues of soil and groundwater contamination by chlorinated compounds such as lindane, chloroform and chloropropionates, it is necessary to identify potential strains that are able to ideally completely mineralise or at least transform them into less toxic forms. As described above, a recently isolated *Rhodococcus* UKMP-5M shows great potential for degradation of toxic chemicals such as cyanide, phenol and various hydrocarbons. However, the capacity of this strain to metabolise halogenated compounds is yet to be explored. Preliminary studies have shown that the *R*. UKMP-5M was not able to assimilate any of the chlorinated compounds tested, but the resting cells of this strain are able to dehalogenate a number of compounds and the dehalogenating activity was further elevated with the addition of 1-Chlorobutane (1-CB) in the growth media. Hence, 1-CB was chosen as a model substrate to study the physiology of R. UKMP-5M in dehalogenation. Dehalogenation of 1-CB in many organisms including *Rhodococcus rhodochrous* NCIMB 13064 has been reported to take place via hydrolysis (Curragh, Flynn et al. 1994). With the aid of a hydrolytic enzyme known as haloalkane dehalogenase, 1-CB can be hydrolysed to its corresponding alcohol which is 1-butanol as shown in the equation below.

$$CH_3CH_2CH_2CH_2Cl + H_2O \longrightarrow CH_3CH_2CH_2CH_2OH + Cl^- + H^+$$

The resulting 1-butanol will be subsequently metabolised as a carbon and energy source via butyric acid. As will be shown in this thesis and in contrast to these reports, *R*. UKMP-5M was found to utilise 1-butanol as a growth substrate but was unable to utilise 1-CB as a

growth substrate. However, the resting cells of *R*. UKMP-5M were able to degrade 1-CB with the liberation of chloride.

1.7 Aim and Objectives

The aim of this study was to unravel the mechanism of 1-CB degradation in *R*. UKMP-5M and to use this understanding to enhance the remediation of chloroalkanes. The objectives of this research were:

(a) To determine the ability of *R*. UKMP-5M to dehalogenate chloroalkanes

(b) To optimise the degradation parameters for 1-CB metabolism in *R*. UKMP-5M

(c) To use a proteomic approach to compare *R*. UKMP-5M grown in the presence and absence of 1-CB

(d) To identify proteins whose expression is induced by 1-CB in *R*. UKMP-5M

(e) To clone and express a protein known as amidohydrolase from *R*. UKMP-5M in *E. coli* that was identified from the analysis of the proteomic data.

(f) To generate a 3D protein structure for the amidohydrolase from *R*. UKMP-5M based on structural homology

(g) To propose a model for the mechanism of 1-CB degradation in *R*. UKMP-5M

CHAPTER 2 DEGRADATION OF 1-CHLOROBUTANE BY *RHODOCOCCUS* UKMP-5M

2.1 Introduction

Halogenated organic compounds are one of the major pollutants introduced to the environment due to natural processes and anthropogenic activities. As mentioned in Chapter 1, the extensive usage of these compounds as herbicides, pesticides, plasticizers, degreasing agents and aerosol propellants has caused public concern over their impacts on the quality of life (Fetzner and Lingens 1994). Many of these halogenated compounds which are not degraded by microorganisms will remain toxic in the environment for an extended period of time. Therefore, identification and application of microorganisms that can destroy these hazardous chemicals is crucial on restoring the ecosystem. Of the diverse halogenated compounds in widespread use, haloalkanes have the simplest configuration and are often chosen as a model substrate to study the microbial metabolism of haloaorganic. In this study, 1-Chlorobutane (1-CB) was used to investigate the ability of *R.* UKMP-5M to degrade halogenated compounds.

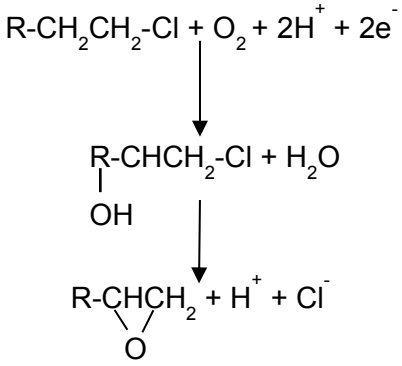
There are two types of dehalogenating enzymes found in nature that reacts on haloalkanes; one is a simple hydrolase that carries out the following reaction:

$$R\text{-}Cl + H_2O \longrightarrow R\text{-}OH + H^+ + Cl^-$$

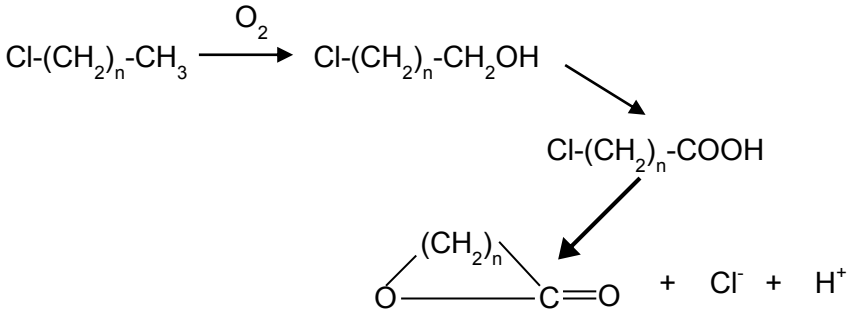
The other is a more complex oxygenase that uses oxygen in an oxidative dehalogenation through the α-hydroxylation:

$$R\text{-}CH_2CH_2\text{-}Cl + \tfrac{1}{2}\, O_2 \longrightarrow R\text{-}CH_2CH(OH)Cl \longrightarrow R\text{-}CH_2CHO + H^+ + Cl^-$$

or through internal hydroxylation followed by intramolecular epoxidation:



or through ω-hydroxylation followed by a stepwise oxidation and intramolecular lactonization:



A number of microorganisms have been reported to efficiently metabolize 1-CB as carbon and energy source and thus, have led to the identification of potential metabolic pathways for this process. It has been reported in many studies that dechlorination of 1-CB was mediated by a single hydrolase-type dehalogenase. For instance, a gammaproteobacteria identified as *Acinetobacter sp*. GJ10 was able to utilise 1-CB along with several other mono and di-substituted haloalkanes as growth substrate. Crude extracts of this bacterium, grown in 1-CB, were able to dehalogenate 1-CB through a hydrolytic reaction to produce 1-butanol. The failure to detect any trace of aldehyde in the reaction suggested the non-involvement of hydroxylase enzyme(s) in the conversion of 1-CB in this bacteria (Janssen, Jager et al. 1987).

Similar observations were also reported with *Corynebacterium* M15-3, an isolate that contains an inducible-type dehalogenase that was able to react with a number of

halogenated compounds. The cell extract's activity towards 1-CB was highly dependent on the presence of 1-CB as an inducer. This was evident through the lack of dehalogenase activity in extracts prepared from cells grown in sodium succinate in the absence of 1-CB. The purified haloalkane dehalogenase from *Corynebacterium* M15-3 has a molecular weight of 33 kDa and a K_m value of 0.18 mM for 1-CB (Yokota, Omori et al. 1987). In contrast to these studies, *Xanthobacter autotrophicus* GJ10, a DCA utilising bacterium, was found to produce dehalogenases constitutively. This strain produced two types of dehalogenases, one reacts specifically with haloalkanes and one with haloacids. However, no evidence of oxygenase involvement in degradation of haloalkanes and haloacids was reported with this strain. The purified haloalkane dehalogenase from this system has a molecular weight of 36 kDa and a K_m value of 1.1 mM for DCA and is able to withstand temperatures up to 50^0C (Keuning, Janssen et al. 1985).

Rhodococcus erythropolis Y2 was isolated by enrichment culture using 1-CB and was able to utilise a number of haloaliphatic compounds ranging from C_3 to C_{18} as its sole carbon source. The utilisation of 1-CB in this strain occurred through a hydrolase-type system which was identified through activity stain-polyacrylamide gel electrophoresis at a molecular mass of 34 kDa (Sallis, Armfield et al. 1990). However, further studies revealed that there were two types of dehalogenation system present in this strain: a hydrolytic type dehalogenase which was induced by C_3 to C_6 1-haloalkanes (with an activity of 44 mU mg^{-1} towards 1-CB) and an oxygenase-type system induced by C_7 to C_{16} 1-haloalkanes and n-alkanes, but that still showed activity towards degradation of 1-CB (3.67 mU mg^{-1}). *R. erythropolis* Y2 was the first isolate reported to contain both a halidohydrolase and an oxygenase-type system for degradation of haloalkanes, though the oxidative metabolic pathway involved in the assimilation of 1-CB was not explored (Armfield, Sallis et al. 1995).

Rhodococcus rhodochrous NCIMB 13064 has been studied extensively for its ability to dehalogenate 1-CB. This strain was isolated from an industrial site in the United Kingdom and is also capable of degrading a wide range of haloalkanes ranging from mono-substituted

haloalkanes to several di-substituted haloalkanes, haloalcohols and haloacids. Catabolism of 1-CB in this strain takes place via two routes. The dominant pathway was through hydrolysis, whereby a hydrolytic enzyme known as haloalkane dehalogenase attacks the C-1 atom of the 1-CB to yield its corresponding alcohol; 1-butanol which was then metabolised through butyric acid. Traces of γ-butyrolactone in the medium further suggest that an oxygenase attack at C-4 atom of 1-CB also occurred. This has led to the formation of 4-chlorobutyric acid which was subsequently either lactonised to form γ-butyrolactone or can be hydrolysed to form 4-hydroxybutyric acid and then subsequently to succinate (Curragh, Flynn et al. 1994).

The isolation and identification of halogenated compound-degrading organisms has played an important role in biological waste treatment especially in the treatment of agricultural wastewater. Huge efforts are still being invested to set up reactors to treat wastewater biologically. For example, a model synthetic wastewater containing 2-fluorobenzoate (2-FB) and dichloromethane (DCM) was created to assess the stability and performance of a continuous up-flow fixed bed reactor. A bacterial consortium consisting of one 2-FB degrading enzyme and one DCM degrader was inoculated in the reactors and the reactors were allowed to operate for 8 months. Efficient removals of the substrates was observed (Emanuelsson, Osuna et al. 2008). Since the utilization of bacterial strains has proven to be practical in biological waste treatment, the need to screen for halogenated compound-degrading microorganisms is of great interest.

In this chapter, the capacity of a recently isolated *R*. UKMP-5M in degrading 1-CB along with several other halogenated compounds was explored. A detailed study on the physiological aspects of this strain in terms of growth patterns in different carbon sources and the ability of the resting cells to degrade 1-CB were demonstrated. In addition, the state of the gene(s) responsible for the biodegradation of 1-CB and a potential repressor of this activity were also identified. The optimal culture conditions for the degradation of 1-CB was also determined by manipulating several physical factors such as temperature and pH to

obtain information which could be employed in designing optimum conditions for actual bioremediation of wastewaters containing halogenated compounds and *R*. UKMP-5M.

2.2 Materials and methods

2.2.1 Bacterial strain and maintenance

Rhodococcus UKMP-5M used in this study was kindly provided by the Unisel Culture Collection (UCC) as part of their extensive effort in gathering various microorganisms from their local resources. The strain was isolated from petroleum contaminated soil in Negeri Sembilan, Malaysia. The strain was grown on Luria broth (LB) (10.0 g/L peptone, 10.0 g/L sodium chloride and 5.0 g/L yeast extract). For agar plates, 15.0 g/L bacterial agar powder was added into the broth prior to autoclaving. Glycerol stocks were prepared for long term storage of the strains.

2.2.2 Glycerol stock preparation

Glycerol stocks for *R*. UKMP-5M were prepared by adding 0.5 ml of sterile 80% glycerol to 0.5 ml of bacterial culture in cryogenic vials. Vials were then inverted few times and stored in -80^{0}C freezer.

2.2.3 Growth media

The ability of *R*. UKMP-5M to degrade 1-CB was first tested by growing the cells in minimal salt media (MSM) supplemented with 5 mM 1-CB. The MSM consists of (/L) 1.6 g $NaH_2PO_4.2H_2O$, 1.4 g Na_2HPO_4, 0.2 g $MgSO_4.7H_2O$, 0.5 g $(NH_4)_2SO_4$, 0.5 g K_2SO_4 and 10.0 ml/L of trace element solution. Trace element solution was made up of (/L) 12.0 g Na_2-EDTA, 0.4 g $ZnSO_4.7H_2O$, 0.4 g $MnSO_4.7H_2O$, 0.1 g $CuSO_4.5H_2O$, 2.0 g $FeSO_4.7H_2O$, 1.0 g $CaSO_4$, 0.1 g $NiSO_4.6H_2O$ and 0.5 ml H_2SO_4, then neutralized to pH 7.0 with 10 M NaOH prior to autoclaving (Hernawan, Soemitro et al. 1999). A 250 ml shake flask containing 50 ml medium was inoculated with 4% of standardised pre-culture ($OD_{600\ nm}$=0.5). The reaction

vessels were sealed with air tight rubber septa prior to injection of 1-CB into the medium which were then incubated at 30^0C with agitation at 180 rpm in an incubator shaker. Growth patterns of *R*. UKMP-5M in the presence and absence of 1-CB were measured spectrophotometrically at 12 hours interval. All experiments were repeated in triplicate.

2.2.4 Determination of cell growth

Samples were withdrawn aseptically by pipetting approximately 1 ml from bacterial broth into 1.5 ml semi-micro disposable cuvettes. Cuvettes were read using a spectrophotometer at 600 nm against blank media at appropriate time intervals.

2.2.5 Preparation of resting cells of *R*. UKMP-5M

Cells were grown in 50 ml MSM containing 20 mM of different carbon sources as described in section 2.2.3. The cells were harvested at their late exponential phase by centrifugation at 4000 rpm. Cells were then washed twice with 40 ml sterile deionised water and resuspended into 7.5 ml of 50 mM Tris-sulphate buffer pH 7.0.

2.2.6 Resting cell assay

Resting cell assays were done in 30 ml glass vials containing 7.5 ml cell suspension. Reaction vessels were sealed with air tight rubber septa and 20 mM 1-CB was injected into the vessels using a glass syringe. Controls containing cell suspension without 1-CB was used to assess endogenous chloride release from the cells and a second control, containing buffer and 20 mM 1-CB without bacterial cells was employed to measure abiotic dehalogenation. In addition to these, another set up was also prepared with the reaction vessels containing 1-CB inoculated with 7.5 ml *E. coli* DH5α grown in 20 mM glucose (prepared as described in 2.2.3 and 2.2.5). All reaction vessels were incubated at 30^0C in an incubator with agitation at 180 rpm. A 1 ml sample was withdrawn every 2 hours for chloride assay.

2.2.7 Chloride release assay

Chloride release was monitored based on a colorimetric method (Bergmann and Sanik 1957). 1 ml sample was added into 100 µl ammonium ferric sulphate in 9 M nitric acid followed by 100 µl mercury (II) thiocyanate saturated ethanol. The solution was then mixed thoroughly and left for 10 minutes for colour development before absorbance measurement at 460 nm. Chloride concentration was determined by comparing the absorbance against a standard curve of known sodium chloride concentrations.

2.2.8 Test for inducibility of dehalogenation

Test for inducibility of 1-CB dehalogenation in *R*. UKMP-5M was done by comparing the chloride release from 1-CB resulting from resting cells which were grown in 20 mM glucose and 20 mM glucose supplemented with 5 mM 1-CB with the experimental procedures described above.

2.2.9 Optimization of reaction conditions

The effects of temperature and pH on the metabolism of 1-CB were tested between 20 and 40^0C and 5 to 9 respectively. Optimal conditions obtained through this experiment were used to determine the effect of different buffer types on the biodegradation of 1-CB.

2.2.10 Dehalogenation of other halogenated compounds

Biotransformation of several other halogenated compounds namely dichloromethane, trichloromethane, tetrachloromethane, 1,2-dichloroethane, 4-chlorobutanol, 3-chloropropionic acid and lindane which were available in the laboratory were tested. 5 mM of these compounds were added to the resting cells instead of 1-CB with the experimental procedures as described in section 2.2.6.

2.2.11 Oxygen dependence of the dehalogenation activity

To determine whether the dehalogenation in *R*. UKMP-5M is an oxygen dependent or independent reaction, a qualitative assay was done. Vials containing cell suspensions were sealed with air tight rubber septa and sparged with oxygen-free nitrogen gas for 30 minutes. Vessels were agitated throughout the sparging to ensure removal of oxygen. 20 mM of 1-CB was then injected into the vessels and reaction mixture was incubated at 30^0C with agitation at 180 rpm. Reactions were repeated in triplicate and chloride release was measured every 2 hours.

2.2.12 Test for the involvement of cytochrome P450

To determine whether the dehalogenation of 1-CB in *R.* UKMP-5M is catalysed by cytochrome P450, vials containing cell suspensions were sealed with air tight rubber septa and sparged with carbon monoxide (CO) for 30 minutes with agitation to ensure saturation. The CO sparging was done in a fume cupboard. 20 mM 1-CB was injected into vials and reaction was incubated in the dark at 30^0C and agitation of 180 rpm. Chloride release was monitored every 2 hours.

2.3 Results and Discussions

2.3.1 Dehalogenation of 1-Chlorobutane by *R.* UKMP-5M

As described earlier, many organisms including *Rhodococcus* species have been reported for their capacity to utilise 1-CB as their sole carbon and energy source. However, the utilisation of 1-CB as growth substrate seems not to occur in *R.* UKMP-5M. This is demonstrated in Figure 2.1 where there was no significant increase in cell population over time in the medium supplemented with 5 mM 1-CB as the carbon source when compared to the growth of a known 1-CB degrader.

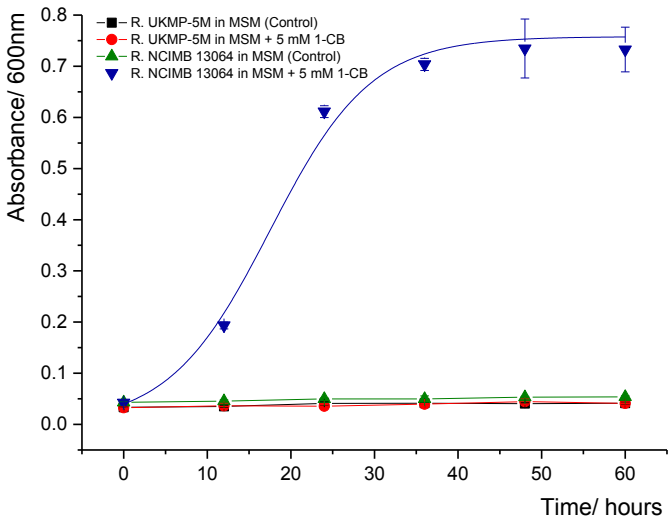


Figure 2.1: Growth of two *Rhodococcus* strains monitored spectrophotometrically at 12 hours interval for 60 hours; *R.* UKMP-5M in minimal media without 1-CB (■) and with 5 mM 1-CB (●). *R.* NCIMB 13064 in minimal media without 1-CB (▲) and with 5 mM 1-CB (▼). Error bars represent standard errors between three determinations.

R. NCIMB 13064 which has been demonstrated to utilise 1-CB as its sole carbon and energy source within 36 hours was used as a positive control in this study. This strain demonstrated a short lag phase (six hours) before growing exponentially to a final absorbance of 0.8 from an initial absorbance of 0.05 which is consistent with a previous study (Curragh, Flynn et al. 1994). The rapid increase in cell density suggests that the cells were degrading and incorporating carbon atoms from 1-CB into biomass. This observation led to the curiosity to explore the reasons for the failure of *R.* UKMP-5M to grow on 1-CB.

Efforts were made to investigate if there were any possible conditions that would lead to cells growing on 1-CB. Multiple studies have shown that addition of yeast extract in growth media aids certain regulation responses by providing growth factors such as vitamins, amino acids and sugars for cell proliferation. For example, addition of yeast extract during cultivation of *Candida boidinii* has significantly shortened the lag phase of its growth by approximately 20 hours (Adachi, Akakabe et al. 2006). However, in this study, addition of 0.1% yeast extract in the growth media still did not help the cells to grow on 1-CB. The inability of *R.* UKMP-5M to utilise 1-CB for biomass generation could be due to several

factors. One main reason could be due to the inability of the organism to possess or sufficiently express the required catabolic system(s) for the utilisation of this substrate. Although the draft genome sequence of this bacterium has shown it to contain genes which encode for possible haloalkane dehalogenases, the substrate specificity of this protein might be different to those identified in previous studies and therefore, was not able to catalyse the degradation of 1-CB. Result of this experiment is shown in Figure 2.2.

Some bacterial systems will depend upon the presence of a primary substrate to be able to degrade and utilise a secondary substrate co-metabolically. For instance, *R. aetherovorans* BCP1 was reported to degrade approximately 633 µM chloroform via aerobic co-metabolism by utilising butane as growth substrate (Frascari, Pinelli et al. 2006). Therefore, ability of 1-CB to support additional growth of *R.* UKMP-5M in the presence of a primary substrate was also tested. Preliminary investigations have shown that ethanol was one of the carbon sources that support the growth of *R.* UKMP-5M without showing a catabolite repression effect onto degradation of nitrilated compounds by *R.* UKMP-5M (unpublished data). However, the addition of 10 mM ethanol as the primary carbon source into the growth medium containing 5 mM 1-CB in the present study did not also contribute to further increase in cell population, as shown in Figure 2.2.

It was then proven that 1-CB could not serve as a growth substrate for *R.* UKMP-5M. The ability of the resting cells grown on alternate carbon sources to dehalogenate 1-CB was investigated. A number of microorganisms have been reported to degrade haloalkanes without growing on them. This includes *R.* NCIMB 13064 which was able to dehalogenate compounds like 1-chlorononane, 1-bromopentane and 1-bromohexane but unable to utilise them as growth substrate although it could both dehalogenate and grow on 1-CB (Curragh, Flynn et al. 1994).

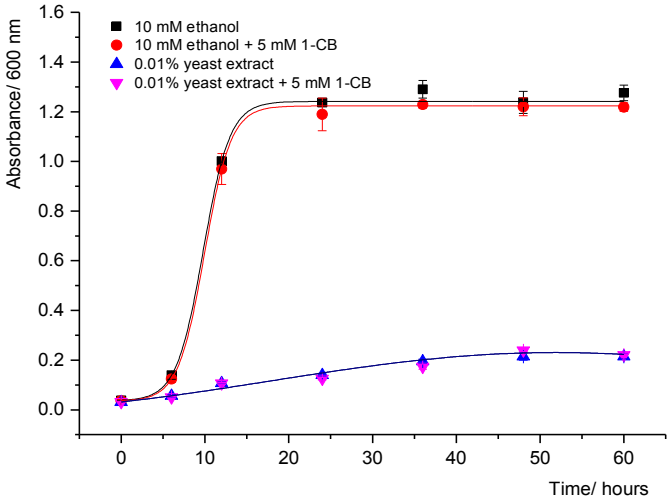


Figure 2.2: Growth of *R.* UKMP-5M in minimal media containing 1-CB supplemented with other primary substrates. Growth with 10 mM ethanol (■); 10 mM ethanol with 5 mM 1-CB (●); 0.01% yeast extract (▲) and 0.01% yeast extract with 5 mM 1-CB (▼). All reactions were incubated at 30^0C with agitation at 180 rpm. Error bars represent standard errors between three determinations.

2.3.2 Growth of *R.* UKMP on different carbon sources

The growth of *R.* UKMP-5M was tested on five different carbon sources. The capacity of this strain to utilise alcohols as growth substrates has not been previously explored. Glucose was generally reported as the best carbon source used to generate biomass of *R.* UKMP-5M which eventually exhibits high activity rate for degradation of substrates like cyanide (Nallapan Maniyam, Sjahrir et al. 2013) and phenol (Suhaila, Rosfarizan et al. 2013). However as Figure 2.3 clearly shows, alcohols such as 1-propanol and 1-butanol can serve as a better growth substrates for *R.* UKMP-5M than glucose.

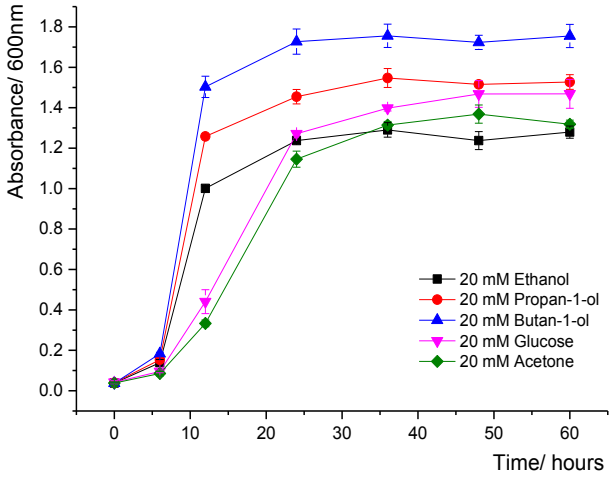


Figure 2.3: Growth of *R.* UKMP-5M in minimal media supplemented with 20 mM of various carbon sources. Ethanol (■); propan-1-ol (●); acetone (♦); butan-1-ol (▲) and glucose (▼). All reactions were incubated at 30^0C with an agitation at 180 rpm. Error bars represent standard errors between three determinations.

Growth patterns for *R.* UKMP-5M on the alcohols were consistent and showed a linear relationship with the number of carbon atoms present in the alcoholic substrates, i.e. as the number of carbon atom increases from two in ethanol to three and four in 1- propanol and 1-butanol, the growth rate increases accordingly, and contributed to the increase in biomass in the order of 1-butanol >1-propanol > ethanol. Cells from the late exponential phase (after 24 hours) were harvested to assess their ability for whole cell dehalogenation of 1-CB.

Table 2.1: The amount of cells obtained from 50 ml MSM supplemented with 20 mM of different carbon sources after 24 hours.

Growth substrate (20 mM)	Wet cell weight after 24 hours (g/50 ml)
1-Butanol	1.47 ± 0.12
1-Propanol	1.20 ± 0.10
Ethanol	1.02 ± 0.10
Glucose	1.18 ± 0.13
Acetone	1.00 ± 0.09

2.3.3 Chloride assay

Biodegradation of haloalkanes will usually produce chloride as one of the products. The colorimetric method for detection and quantification of chloride used in this study was first introduced by Bergmann and Sanik in 1957. This method involves the displacement of thiocyanate ion from mercuric thiocyanate by chloride in the presence of ferric ion as shown in the equation below.

$$2Cl^{-} + Hg(SCN)_2 + 2Fe^{3+} \longrightarrow HgCl_2 + 2\,Fe(SCN)^{2+}$$

The ion displacements will lead to the formation of mercuric chloride, which is an insoluble precipitate. Thiocyanate, being a strong ligand will then bind to ferric ion to form ferric thiocyanate, which is a red complex that is stable for approximately 30 minutes. The intensity of the red complex will be proportional to the chloride concentration.

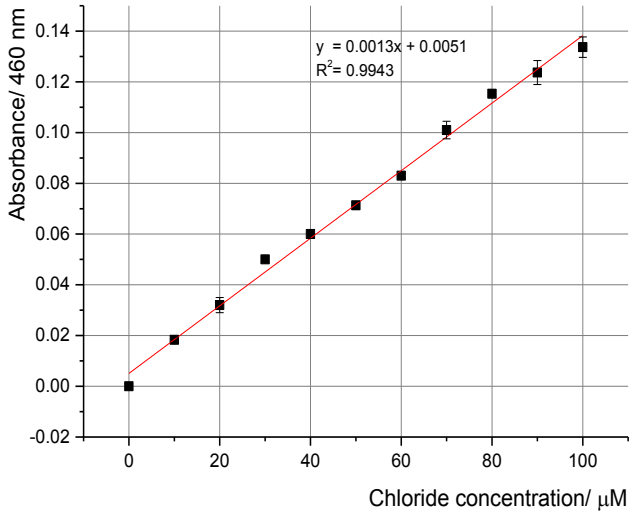


Figure 2.4: Calibration curve for chloride assay. The detection limit for this assay is 10 μM.

2.3.4 Dehalogenation of 1-CB by resting cells of *R.* UKMP-5M

The capacity of *R.* UKMP-5M for whole cell dehalogenation was explored by incubating the resting cells grown in different carbon sources with 20 mM 1-CB. Three types of negative control systems were set up; one to monitor the chloride release due to flux of internal chloride content from cells (endogenous halide release), one to monitor the abiotic dehalogenation and one to monitor the chloride release by a non-degrader of 1-CB, *E. coli* DH5α. As clearly depicted in Figure 2.5, there was insignificant or no chloride accumulation was observed from all the negative controls. However, significant amount of chloride release was only observed in the buffer containing cell suspensions of *R.* UKMP-5M and 1-CB. Results from this assay suggested that *R.* UKMP-5M was capable of metabolising 1-CB but was unable to utilise this substrate for growth.

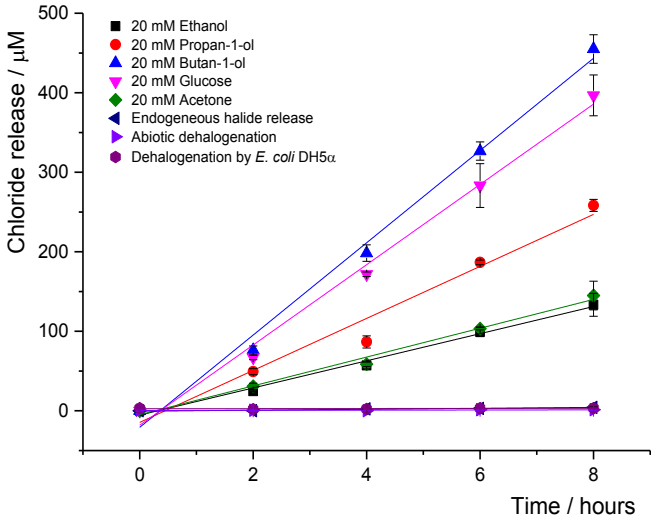


Figure 2.5: Chloride release from 1-CB by resting cells of *R.* UKMP-5M grown in 20 mM of different carbon sources. Ethanol (■); propan-1-ol (●); acetone (♦); butan-1-ol (▲) and glucose (▼). Endogenous chloride release is represented by (◀), abiotic dehalogenation is represented by (▶) and chloride release by *E. coli* DH5α is represented by (●). Error bars represent standard errors between three determinations.

Table 2.2: Dehalogenation activity of 1-CB by resting cells of *R.* UKMP-5M grown in 20 mM of different carbon sources

Carbon source	Specific activity (mU/mg of total protein)
1-Butanol	0.53 ± 0.02
1-Propanol	0.35 ± 0.05
Ethanol	0.26 ± 0.02
Glucose	0.56 ± 0.03
Acetone	0.26 ± 0.02

Highest dehalogenating activity can be observed in the cells grown in 1-butanol with a net chloride release of nearly 450 µM. This is in contrast to the case of *R.* NCIMB 13064 where it was reported that cells of *R.* NCIMB 13064 grown in 1-butanol were incapable of degrading 1-CB, but growth of this strain in 1-CB induces a set of enzyme(s) which allows dehalogenation of a number of halogenated compounds (Curragh, Flynn et al. 1994). The capacity *R.* UKMP-5M's resting cells, prepared in the absence of an inducer, to dehalogenate 1-CB indicates that this activity is expressed constitutively. Although the amount of cell biomass produced from glucose and 1-propanol is similar, glucose grown cells exhibited similar degrading capacity to that of 1-butanol grown cells with a net chloride release of 400 µM when compared to 1-propanol grown cells and no obvious explanation for this is apparent. Given that there were no large differences between the capacities of cells grown on 1-butanol and glucose to dehalogenate 1-CB, glucose was chosen as the substrate to generate biomass for future work in this study.

2.3.5 Further inducibility of 1-CB degradation

The ability of *R.* UKMP-5M to dehalogenate 1-CB in the absence of an inducer as described in section 2.3.4 clearly indicates that the gene(s) responsible for the expression of this activity are constitutive, which is not uncommon in *Rhodococcus*. A constitutive gene is a

gene that is equally expressed in all cells under most if not all environmental conditions (Mandel, Fleming et al. 1995). Multiple studies have shown that *Rhodococcus* has catabolic gene(s) which are constitutively expressed, but further inducible in the presence of a suitable inducer. For example, *R. jostii* RHA1 was demonstrated to constitutively degrade a carcinogenic water contaminant, N-nitrosodimethylamine (NDMA). The degradation of NDMA was further enhanced by approximately 500-fold after growth of the cells on propane. It was reported that propane could elicit the upregulation of gene(s) responsible for oxidation of both propane and substituted benzenes (Sharp, Sales et al. 2007).

The majority of instances for degradation of 1-CB involve inducible enzyme(s) to mediate the reaction. For example, sodium succinate was used as a growth substrate by *Rhodococcus erythropolis* Y2 for biomass generation. The constitutive activity of the hydrolytic dehalogenase towards 1-CB was only 4% of the maximum activity obtained following induction with 1-CB. Among 13 inducers tested, 1-CB was found to yield maximum activity towards its degradation, followed by 1,4-dichlorobutane which exhibits about 93% of activity towards 1-CB (Armfield, Sallis et al. 1995).

In this section, the ability of *R*. UKMP-5M to use the substrate as its inducer was explored. 50 ml growth media containing glucose as a carbon source was supplemented with 5 mM 1-CB. The presence of 1-CB in the media is expected to upregulate the expression of the gene(s) responsible for 1-CB degradation. The resulting cells were harvested after 24 hours and further incubated with 20 mM 1-CB and chloride release was followed. Figure 2.6 clearly shows that the dehalogenation activity in *R.* UKMP-5M is a constitutive activity, but it is further inducible when the cells are grown in the presence of 1-CB. The induced cells of *R.* UKMP-5M were able to accumulate approximately 2 mM chloride from 1-CB within 8 hours of incubation which is a 10% conversion from initial substrate concentration. The specific enzyme activity of the induced cells was 2.7 mU/mg of total protein which is more than four times of the activity observed in non-induced cells.

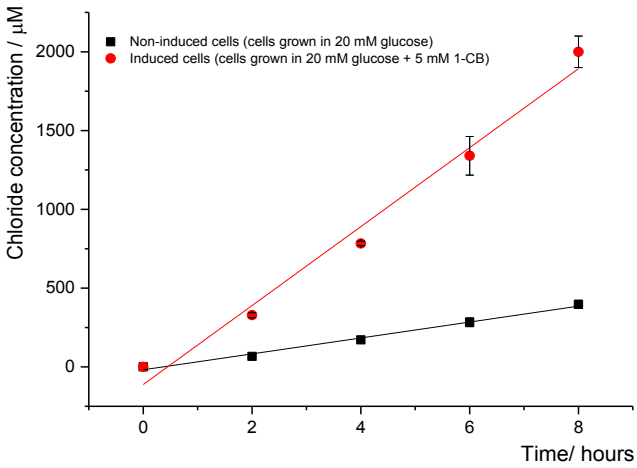


Figure 2.6: Chloride release from 1-CB by resting cells of *R.* UKMP-5M grown in 20 mM of glucose (■) and 20 mM glucose with 5 mM 1-CB (●). Error bars represent standard errors between three determinations.

The capacity of glucose grown *R.* UKMP-5M to dehalogenate 1-CB without catabolite repression was striking. Glucose, in many circumstances acts as a repressor for the synthesis of certain inducible enzymes. The presence of glucose in the growth environment may inhibit the synthesis of cyclic AMP (cAMP) which is required to initiate transcription of any inducible enzyme systems. For example, it was found that biphenyl metabolism in *Rhodococcus jostii* RHA 1 was repressed by the presence of glucose. Results of this investigation were examined at the transcript level for the BPH enzyme gene clusters and it was found that transcription of the targeted genes, which was higher in cells grown in BPH alone, was repressed in the cells grown in the presence of glucose (Araki, Niikura et al. 2011). On the other hand, the role of LB as a catabolite repressor in this study was uncommon. As shown in Figure 2.7, the resting cells of *R.* UKMP-5M which was grown in LB was not able to dehalogenate 1-CB, but the cells grown in LB supplemented with 5 mM 1-CB was able to generate chloride release.

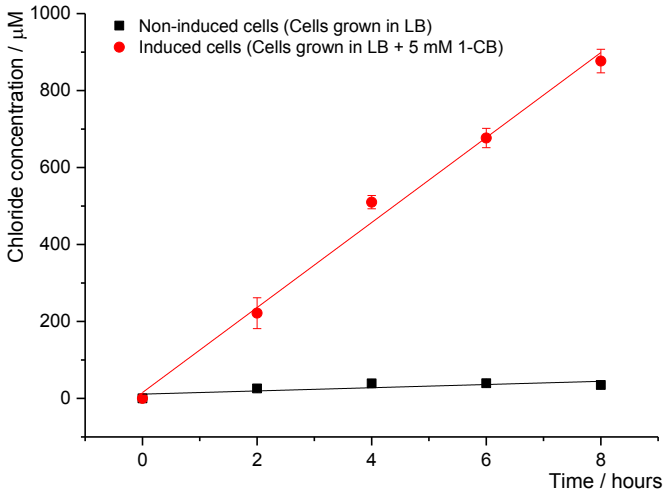


Figure 2.7: Chloride release from 1-CB by resting cells of *R.* UKMP-5M grown in LB (■) and LB supplemented with 5 mM 1-CB (●). All reactions were incubated at 30^0C with an agitation at 180 rpm. Error bars represent standard errors between three determinations.

LB broth, which is made of peptone, sodium chloride and yeast extract, is a medium rich in nutrients and vitamins for the growth of any microorganism. Unlike glucose, no significant studies have been done to demonstrate the role of LB as catabolite repressor in any microorganism. Therefore, each component of LB was analysed individually to determine the source of repression. For example, *R.* UKMP-5M was grown in 20 mM glucose individually supplemented with 10 g/L peptone, 10 g/L sodium chloride and 5 g/L yeast extract. The resulting resting cells were then washed and incubated with 20 mM 1-CB. It was found that the cells grown in the presence of sodium chloride did not show repression effect towards 1-CB. However, cells grown in the presence of either peptone or yeast extract showed activity inhibition of approximately 35% and 50% respectively. Therefore, the combination of these two elements in LB broth has led to the complete inhibition of dehalogenation. This repression effect could be due to the presence of a wide range of vitamins and minerals in both peptone and yeast extract that inhibits the synthesis of other uncommon enzymes but further studies need to be conducted to completely understand the basis of these observation.

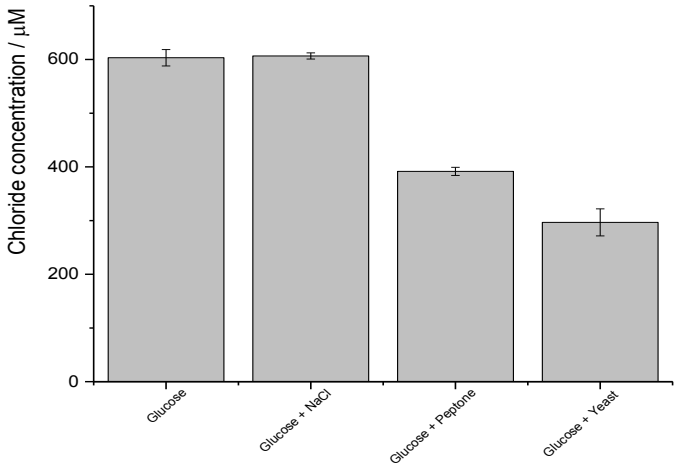


Figure 2.8: Chloride release from 1-CB by resting cells of *R*. UKMP-5M grown in glucose and glucose supplemented with each component of LB. All reactions were incubated at 30^0C with an agitation at 180 rpm. Chloride release was measured after 24 hours. Error bars represent standard errors between three determinations.

2.3.6 Optimization of reaction conditions

It was understood that the utilization of resting cells may have an added advantage in detoxifying environmental pollutants. This is primarily due to the ability of whole cells to accumulate factors that enable them to tolerate a variety of environmental challenges (Maniyam 2013). Practical applications of microorganism in waste treatments sites may introduce conditions, which are unfavourable for the bacteria. Therefore, a range of most desirable conditions needs to be identified in order to design the best approach for degradation of halogenated compounds by this bacterium.

Physical parameters such as temperature and pH play an important role in microbial growth and the secretion of appropriate enzyme(s). In this section, single factor optimisation was done to study the effect of different temperatures, pH, buffer types and inducers on the 1-CB dehalogenation activity in *R.* UKMP-5M. Resting cells of *R*. UKMP-5M grown in the presence of glucose supplemented with 5 mM 1-CB was incubated with 20 mM 1-CB at temperatures ranging from 20 to 40^0C. Chloride release was measured after 8 hours and the specific activity was calculated. The temperature that has given the highest specific activity

was then chosen for the following experiment, to determine the optimum pH for dehalogenation of 1-CB by resting cells of *R*. UKMP-5M.

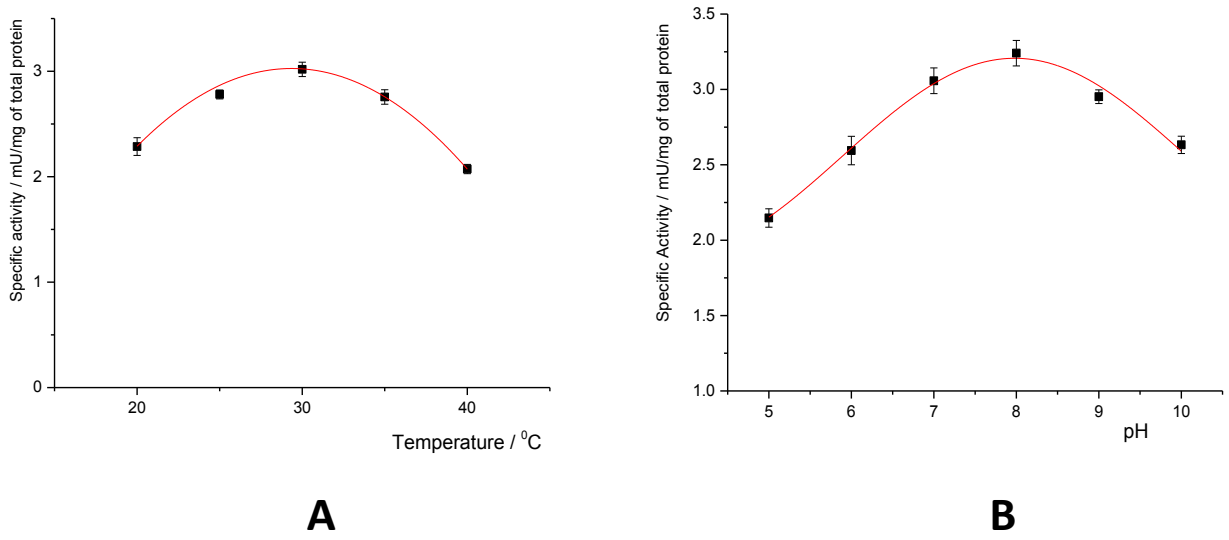


Figure 2.9: Effect of manipulating physical parameters on the degradation of 1-CB by *R*. UKMP-5M. Specific activities when resting cells of *R*. UKMP-5M was incubated with 20 mM of 1-CB at (A) different temperatures, ranging from 20 to 40^{0}C, in a neutral pH buffer and (B) at 30^{0}C in different pH of buffer, ranging from 5 to 10. Error bars represent standard errors between three determinations.

Based on Figure 2.9A and B, it can be concluded that the most appropriate temperature for the maximum removal of 1-CB is 30^{0}C and a pH of 8. *Rhodococcus* in general favours a temperature ranging from 25 to 35^{0}C and an environment close to neutral pH. This is evident through most studies that employed *Rhodococcus* for catabolic activities. For example, *R.* UKMP-5M was shown to exhibit highest activity towards cyanide degradation at temperature of 30^{0}C and a pH of 7. The strain was reported to successfully remove 48% of potassium cyanide in the reaction buffer (Maniyam 2013). From an economic point of view and practical application of waste treatment, this finding is significant since it could benefit the tropical countries such as Malaysia, where the ambient temperature often ranges between 28 to 40°C (Patil and Paknikar 2000). The optimum temperature and pH obtained from this

investigation were then used to identify the most suitable buffer, which was then used to identify the inducer that gives the highest specific activity.

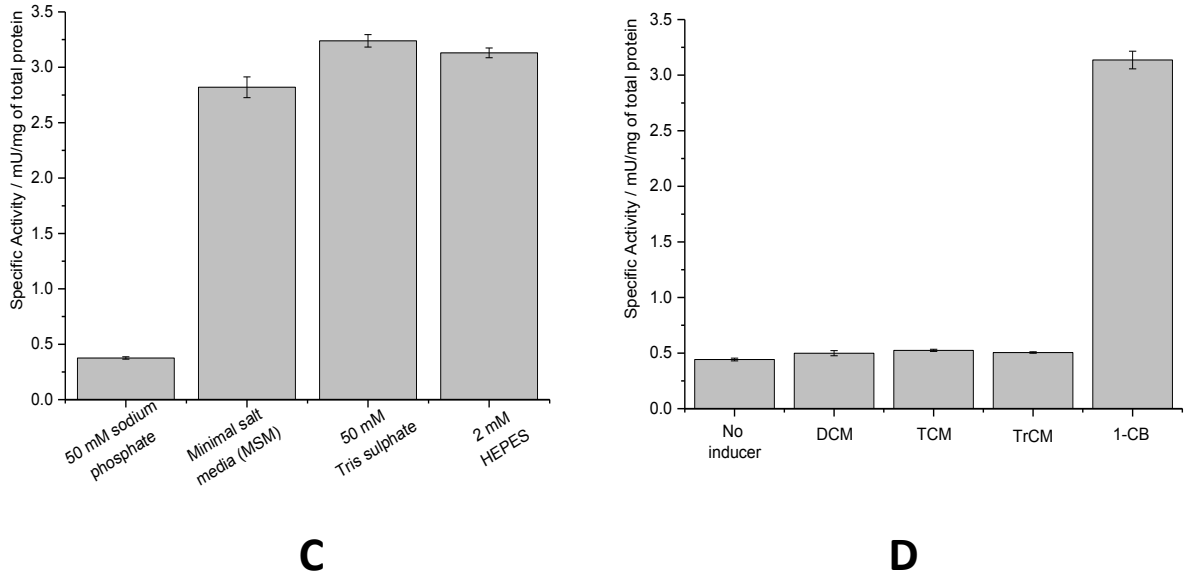


Figure 2.9: Effect of manipulating physical parameters on the degradation of 1-CB by *R*. UKMP-5M. (C) Specific activities when resting cells of *R*. UKMP-5M were incubated with 20 mM of 1-CB at 30^0C in different buffers at pH 8 and (D) Specific activities of resting cells grown in the presence of different inducers on dehalogenation of 1-CB, incubated at 30^0C in 50 mM Tris. sulphate buffer at pH 8.0. Error bars represent standard errors between three determinations.

Based on Figure 2.9C, it was evident that 50 mM sodium phosphate buffer reduces 1-CB dehalogenation by eight-fold when compared to 50 mM Tris. sulphate and 2 mM HEPES buffer. This has never been reported in any previous studies. Analysis on the effect of inducers on the dehalogenation of 1-CB revealed that halomethanes were not able to upregulate the enzyme system(s) in *R.* UKMP-5M to dehalogenate 1-CB. The amount of chloride generation in the absence of an inducer and in the presence of halomethanes as inducer has shown no significant difference. However, in agreement with a published work, 1-CB appeared to be an efficient inducer to induce the activity of dehalogenation towards itself (Armfield, Sallis et al. 1995). Taking the optimum conditions identified into account, the ability of *R.* UKMP-5M to dehalogenate other halogenated compounds was also tested. Cells

which were induced with 1-CB, were incubated with a number of selected halogenated compounds that were available in the laboratory. None of the halogenated compounds listed in Table 2.3 were able to serve as a growth substrate for *R.* UKMP-5M, but the accumulation of chloride in reaction buffer indicates that the compounds are being metabolised. Cells grown on glucose supplemented with halomethanes and haloethanes as inducers showed low levels of dehalogenation activity towards their respective substrates.

Biotransformation of higher chlorinated halomethanes like chloroform (CF) and carbon tetrachloride (CT) in aerobic bacteria mainly occurs through oxidative dehalogenation. For example, the cometabolism of CF in aerobic bacteria is supported by the growth on short chain alkanes such as propane, butane and hexane which induces the activity of monooxygenases for both the metabolism of alkanes and CF. It has been also hypothetically concluded that aerobic degradation of chloroform by monooxygenases leads to the formation of phosgene (Cappelletti, Frascari et al. 2012). On the other hand, aerobic detoxification of CT has been poorly studied. To date, there is no organism isolated that is capable of utilising CT as carbon source for growth. The degradation of CT is always reported as a co-metabolism process. The hydrolytic mineralisation of CT will directly lead to the formation of CO_2. There is no difference in the oxidation state (+4) of the carbon in both CT and CO_2; therefore, no energy release for cellular growth is possible. Most degradation of CT occurs through anaerobic co-metabolism (Penny, Vuilleumier et al. 2010) though in another study, it was found that *Rhodococcus* represented 70% of a consortium capable of dechlorinating CT in the presence of toluene but the mechanism of degradation was not explored (Zhou, Palumbo et al. 1999).

Table 2.3: Biotransformation of other halogenated compounds by resting cell of *R.* UKMP-5M induced with 1-CB

Halogenated compounds	Specific activity (mU/mg total protein)
Dichloromethane (DCM)	0.38 ± 0.05
Trichloromethane (CF)	1.18 ± 0.02
Tetrachloromethane (CT)	0.83 ± 0.02
1, 2-Dichloroethane (DCA)	1.38 ± 0.04
1,1,1-Trichloroethane	0.69 ± 0.03
4-Chlorobutanol	3.06 ± 0.03
Lindane	0
3-Chloropropionic acid	3.47 ± 0.07

2.3.7 Oxygen dependence of the activity

As described in the previous sections, dehalogenation of 1-CB can take place via either hydrolysis by a hydrolase or through a monoxygenase-mediated pathway. Examples of monooxygenases include methane monooxygenase (MMO) and cyctochrome P450 (cyt P450). One criteria that can differentiate between a hydrolase and an oxygenase mediated reactions is the oxygen requirement. Unlike hydrolases, monooxygenase require the presence of oxygen to catalyse a reaction. Methane monooxygenase (MMO) consists of hydroxylase, reductase and a 'B' subunit. The hydroxylase subunit in MMO contains a carboxylate and an oxygen-bridged binuclear iron cluster which will catalyse the activation and insertion of oxygen in a molecule (Smith, Rawat et al. 2011). Therefore, a qualitative investigation was carried out to determine whether the degradation of 1-CB in *R.* UKMP-5M was an oxygen dependent or independent reaction. Resting cells of *R.* UKMP-5M was incubated with 20 mM 1-CB in the presence and absence of oxygen. Chloride release was monitored for 12 hours at two hours interval.

A

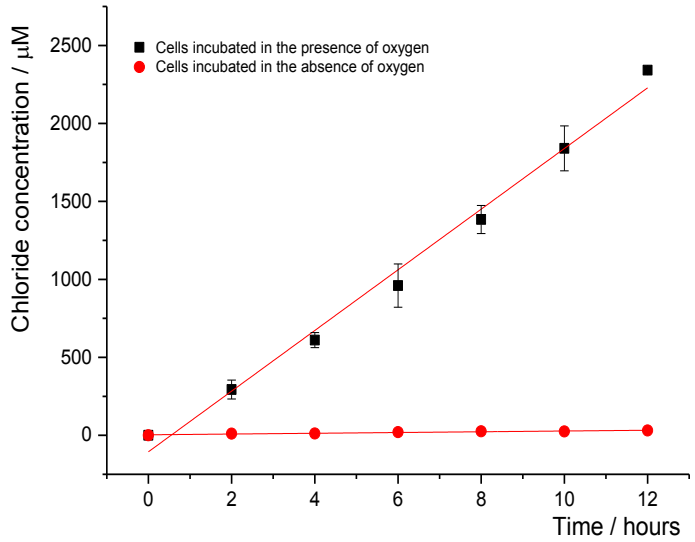


Figure 2.10A: Chloride release from 1-CB by resting cells of *R.* UKMP-5M monitored in the presence of oxygen (■) and in the absence of oxygen (●). Error bars represent standard errors between three determinations.

B

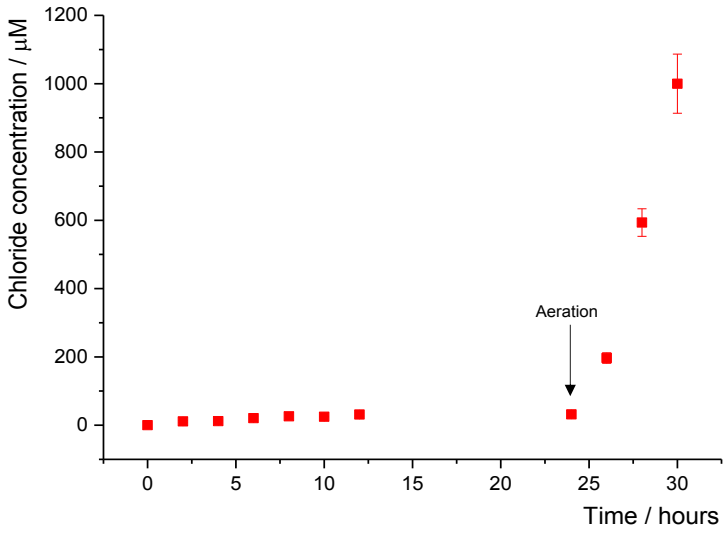


Figure 2.10B: Cells monitored in the absence of oxygen were aerated and re-incubated with 1-CB for their activity recovery. Error bars represent standard errors between three determinations.

Based on the plots in Figure 2.10A, it is evident that dehalogenation of 1-CB in *R.* UKMP-5M is an oxygen dependent activity. As clearly depicted in the graph, there was insignificant chloride release observed in the cells incubated with 1-CB in the absence of oxygen.

However, a linear response can be observed between chloride generation and the time of incubation with the cells incubated in the presence of oxygen. Given that *Rhodococcus* is an aerobic organism, there was a possibility that the removal of oxygen from the environment will eventually kill the cells and therefore, eliminates the dehalogenation activity observed. To address this, vessels containing cells incubated in the absence of oxygen were ventilated to allow entry of oxygen. Upon ventilation, 20 mM of 1-CB was added and chloride release was then monitored.

Results from the recovery test confirmed that the removal of oxygen from the culture environment did not kill the cells and the lack of activity observed in the absence of oxygen supports the involvement of an oxygenase-type enzyme towards dehalogenation of 1-CB. As shown in Figure 2.10B, incubation of the cells with 20 mM 1-CB in the absence of oxygen for 24 hours did not generate chloride in the culture medium and the re-incubation of the cells after aeration has shown 100% dehalogenation activity i.e. amount of chloride released within six hours after aeration was similar to the control which was done in the presence of oxygen from the start. Therefore, it can now be concluded that the dehalogenation activity *R.* UKMP-5M is likely to be an oxygen dependent reaction. A number of microorganisms have been reported for their ability to aerobically degrade haloalkanes. The involvement of monooxygenases in dehalogenation of various haloalkanes has been previously studied. For example, *Pseudomonas sp*. DCA 1 has been found to degrade 1,2-dichloroethane (DCA) as its sole carbon and energy source. The first of this reaction was mediated by a monooxygenase- type enzyme, which led to the formation of the unstable 1,2-dichloroethanol that will then release chloride spontaneously to form chloroacetaldehyde (Hage and Hartmans 1999).

Although many studies have demonstrated the involvement of monooxygenases in aerobic dehalogenation of haloalkanes, not many studies have gone further to identify the types of monooxygenase involve in these reactions. In *Rhodococcus*, the heme-containing cyt P450 are very common. Often in many studies, cytochrome P450 has been identified as

one of the most versatile types of monooxygenase that is able to react with a range of substrates including haloorganic compounds. For example, cytochrome P450 has been reported to initiate the aerobic degradation of cis-dichloroethene in *Polaromonas sp*. JS666 (Nishino, Shin et al. 2013). In the present study, it is likely that the dehalogenation reaction is catalysed by an oxygenase-type system. Therefore, it was worthwhile to test whether cyt P450 was involved in the degradation of 1-CB in *R*. UKMP-5M. The involvement of cyt P450 was tested by incubating the resting cells of *R*. UKMP-5M 1-CB in the presence of carbon monoxide (CO) instead of oxygen and chloride release was then followed.

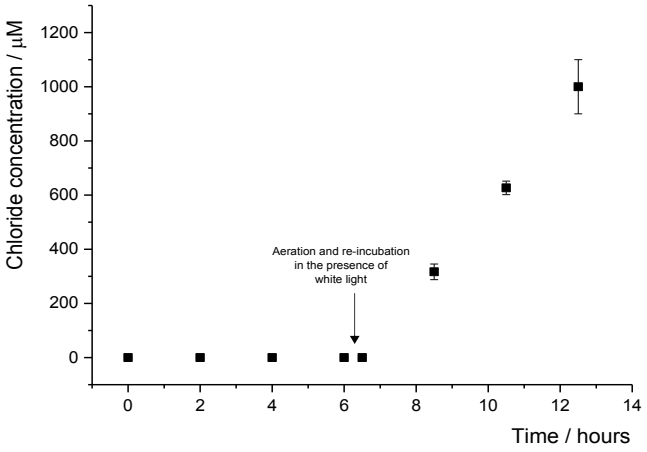


Figure 2.11: Chloride release from 1-CB by resting cells of *R*. UKMP-5M monitored in the presence of carbon monoxide. The vessels were then aerated and re-incubated with 1-CB in the presence of white light for their activity recovery. Error bars represent standard errors between three determinations.

In contrast to the experiment in section 2.3.7, the recovery test should not generate chloride release. This is because, unlike other monooxygenases, cyt P450 contains heme cofactors and the diffusion of CO into the cells will completely occupy the active sites in the reduced cyt P450 and spontaneously interrupts the catalytic cycle of this enzyme. The CO inhibition in cyt P450 is not reversible by a white light source, but is only reversible by blue light irradiation. In the present study, the recovery test, which was done in the presence of white light, was able to stimulate chloride release and 100% activity was recovered after

ventilation. The requirement of oxygen for the degradation of 1-CB suggests the involvement of a monooxygenase type enzyme, but the inability of CO to inactivate the enzyme responsible for degradation suggests that it is not a cyt P450 type.

2.3.8 Metabolic pathway analyses

The possible routes for the degradation of 1-CB by a *Rhodococcus* strain, *R*. NCIMB 13064 have been demonstrated in an earlier study. This involves a dual pathway, i.e. a hydrolytic pathway via the nucleophilic substitution reaction catalysed by a haloalkane dehalogenase and an oxidative pathway catalysed by a monooxygenase that hydroxylates the non-halogenated end of 1-CB (Curragh, Flynn et al. 1994).

A

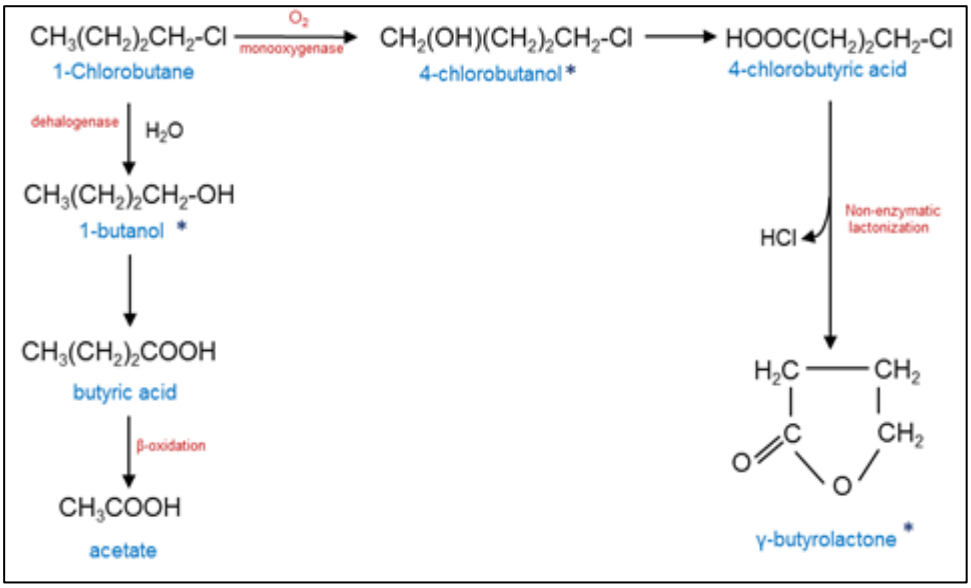


Figure 2.12A: Possible metabolic routes for the dehalogenation of 1-CB in *R*. NCIMB 13064. *Intermediates which were tested for their ability to support the growth of *R*. UKMP-5M. *Adapted with permission from the publisher, Society for General Microbiology.*

Another possible route for oxidative dehalogenation of 1-CB is through internal hydroxylation via the formation of epoxide.

B

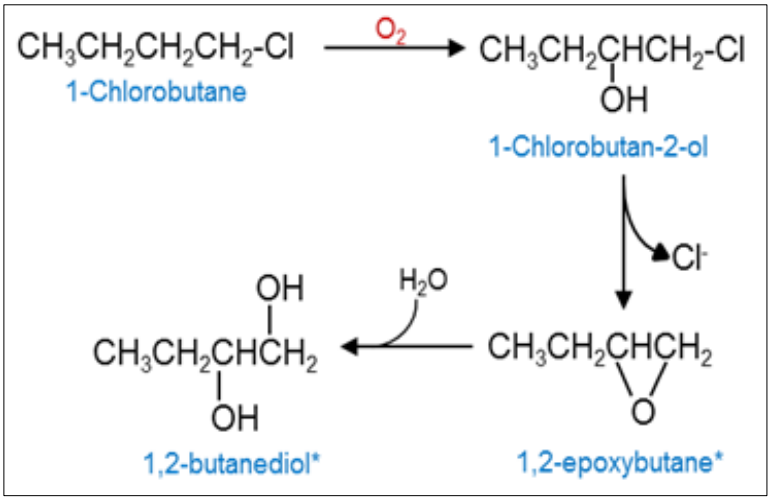


Figure 2.12B: Dehalogenation of 1-CB via internal hydroxylation. This reaction will lead to the formation of 1,2-epoxybutane, which will then be hydrolysed to 1,2-butanediol. *Intermediates which were tested for the ability of *R*. UKMP-5M to grow on.

Given that *R*. UKMP-5M is not able to utilise 1-CB as carbon source for growth, but only dehalogenates it to release chloride, it is speculated that the product (either the final or intermediate) formed after the release of chloride is not able to support the growth. Hence, in order to identify the metabolic pathway involved during the degradation of 1-CB in *R*. UKMP-5M, several possible intermediates according to Figure 2.12A and B were tested for their ability to support the growth of *R*. UKMP-5M. Among the substrates tested were 1-butanol, 4-chlorobutanol, γ-butyrolactone, 1,2-epoxybutane and 1,2-butanediol. Approximately 3 mM chloride was observed when the resting cells were incubated with 1-CB, the potential intermediates were also tested at equal concentrations to determine the ability of *R*. UKMP-5M to utilise them as growth substrates.

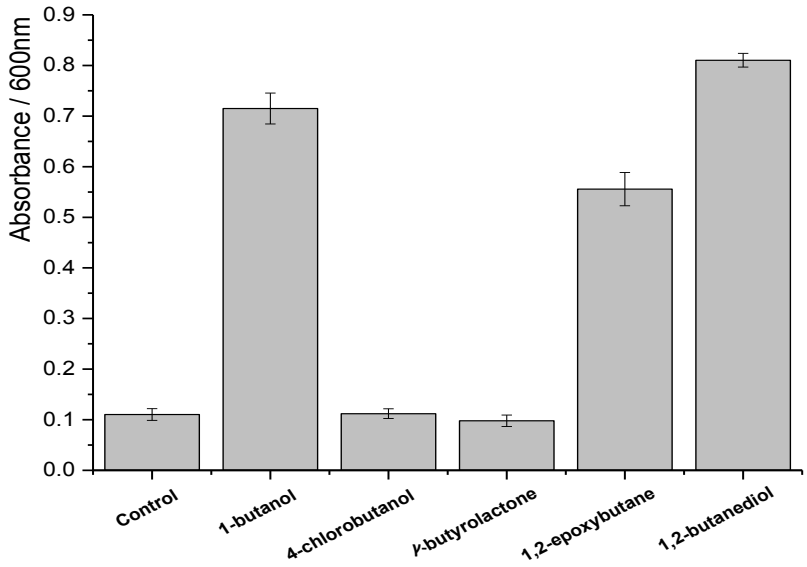


Figure 2.13: Cell density of *R*. UKMP-5M grown on several potential intermediates of 1-CB degradation. Initial absorbance (T= 0) was 0.06. Final absorbance ($OD_{600\ nm}$) was measured after 72 hours of incubation at 30^0C with agitation (180 rpm).

The ability of the cells to grow on 1-butanol, 1,2-epoxybutane and 1,2-butanediol suggest that the dehalogenation of 1-CB is unlikely to take place via the hydrolytic pathway or through the internal hydroxylation. However, the inability of *R*. UKMP-5M to grow on other potential intermediates (4-chlorobutanol, γ-butyrolactone) supports the possibility that the 1-CB degradation could be via the oxidative dehalogenation involving the ω-hydroxylation. When 4-chlorobutanol was assayed for its dehalogenation with *R*. UKMP-5M, chloride release was observed at a similar rate with that of 1-CB dechlorination. This result suggests that the immediate product after dechlorination, which is the γ-butyrolactone in this case, is not supporting the growth of *R*. UKMP-5M, which is in agreement with the results in Figure 2.13. The inability of *R*. UKMP-5M to grow on γ-butyrolactone could be due to the lack of enzymatic system for its assimilation and further accumulation of this compound can cause lethal effect to the cells. In order to test this, a comparative study on growth patterns was done between the 1-CB treated and untreated cells of *R*. UKMP-5M. Cells, which were either exposed or unexposed to 1-CB, were withdrawn at appropriate intervals and inoculated into LB broth to test their growth recovery after 24 hours.

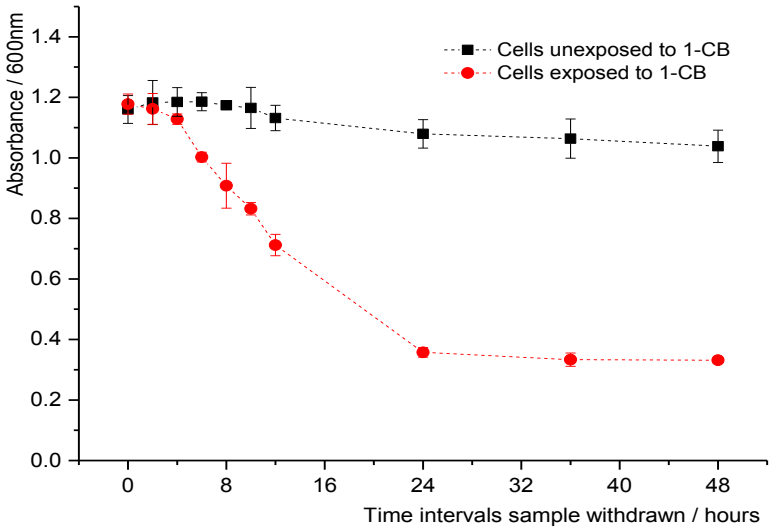


Figure 2.14: Growth recovery patterns of *R*. UKMP-5M unexposed to 1-CB (■) and exposed to 1-CB (●). Spectrophotometric measurement ($OD_{600\ nm}$) was done after 24 hours. Error bars represent standard errors between three determinations.

Plots from Figure 2.14 imply that there is a significant effect of toxicity to the cells following exposure to 1-CB. This lethal effect could either be due to the toxicity of 1-CB or the accumulated product(s) from 1-CB dehalogenation in *R*. UKMP-5M. Unfortunately, further analyses of the metabolic pathway were not successful since the metabolite(s) were not available in their pure forms. The nuclear magnetic resonance (NMR) spectra of the reaction medium showed high background signals, which were due to the residual of glucose (and its metabolites) and the un-degraded 1-CB. Subtraction of the baseline in the NMR spectrum was not able to reveal any product associated with 1-CB and further attempts to eliminate the residual of glucose by thorough washing of the cells before resting cell assay resulted in cell lysis and flux of internal content.

On the other hand, multiple attempts to isolate enzyme activity in the cell free extract were unsuccessful. Different cell lysis methods were used to prepare the crude cell lysate and the amount of protein extracted was approximately 2 mg/ml. Although, the amount of protein obtained was theoretically sufficient to isolate cell lysate's activity, no chloride

release from 1-CB was observed in this study. Assays were also carried out in the presence of cofactors such as NADH and NADPH but no measurable activity was detected.

Hence, in order to identify the enzymes which are responsible for the dehalogenation of 1-CB in *R*. UKMP-5M, a proteomic approach was used to identify the proteins which are induced following the exposure of 1-CB and the results of this investigation are presented in Chapter 3.

CHAPTER 3 PROTEOMICS AND TRANCRIPT ANALYSES

3.1 Introduction

High-throughput genome sequencing is now capable of revealing information about the complete DNA sequence of an organism's genome. Information about the genes and their interactions can be determined and products from the interactions can be predicted. However, this information alone is insufficient to unravel the biological processes in an organism. Approximately 1/3 of the gene sequences present in the genome database have hypothetical or unknown functions (based on sequence homology) (Chauhan and Jain 2010). The traditional relationship that claims one gene encodes for one protein is seen to be a simplification, since a single gene may encode for more than one protein (Singh and Nagaraj 2006). Proteins may undergo different post-translational modifications which alter their diversity and affinity, giving rise to a new protein with different functions (Hardiman 2004). Hence, it is absolutely necessary to study the gene functions directly at the protein level.

The terms 'proteome' and 'proteomics' were first introduced by Marc Wilkins in 1995 (Fulekar and Sharma 2008). The proteome is the set of all proteins expressed by a cell in an organism at one point of time and investigation of the proteome is termed proteomics (Plebani 2005). Every organism possesses one genome but multiple proteomes; because protein expression profiles are dependent on external or internal stimuli of the cell which changes over time (Chauhan and Jain 2010) whereas the genome of an organism is constant regardless of its environment (Mallick and Kuster 2010). Quantitative proteomic methods can be classified as either 'gel-based' or 'gel-free' methods.

The gel-based technique is a powerful and sensitive method to separate and visualise proteins. Two-dimensional gel electrophoresis (2D GE) is a form of electrophoresis that separates a protein mixture in two dimensions. In the first step, proteins are separated

on an immobilized pH gradient (IPG) strip according to their isoelectric points (pI) and in the second step, the focused proteins are separated orthogonally on a sodium dodecyl sulphate-polyacrylamide gel (SDS-PAGE) according to their molecular weights (Abdallah, Dumas-Gaudot et al. 2012). The separated proteins can be visualised after staining with reagents such as coomassie brilliant blue, silver or fluorescent dyes, for example SYPRO ruby or SYPRO orange by digital imaging (Dhingra, Gupta et al. 2005). The protein(s) of interest are then excised and digested with protease trypsin and the peptides are identified using mass spectrometry (MS). MS is a technique used to identify and quantify ions in the gaseous phase based on their mass-to-charge ratio (m/z) and the ion count at the same m/z value (Plebani 2005).

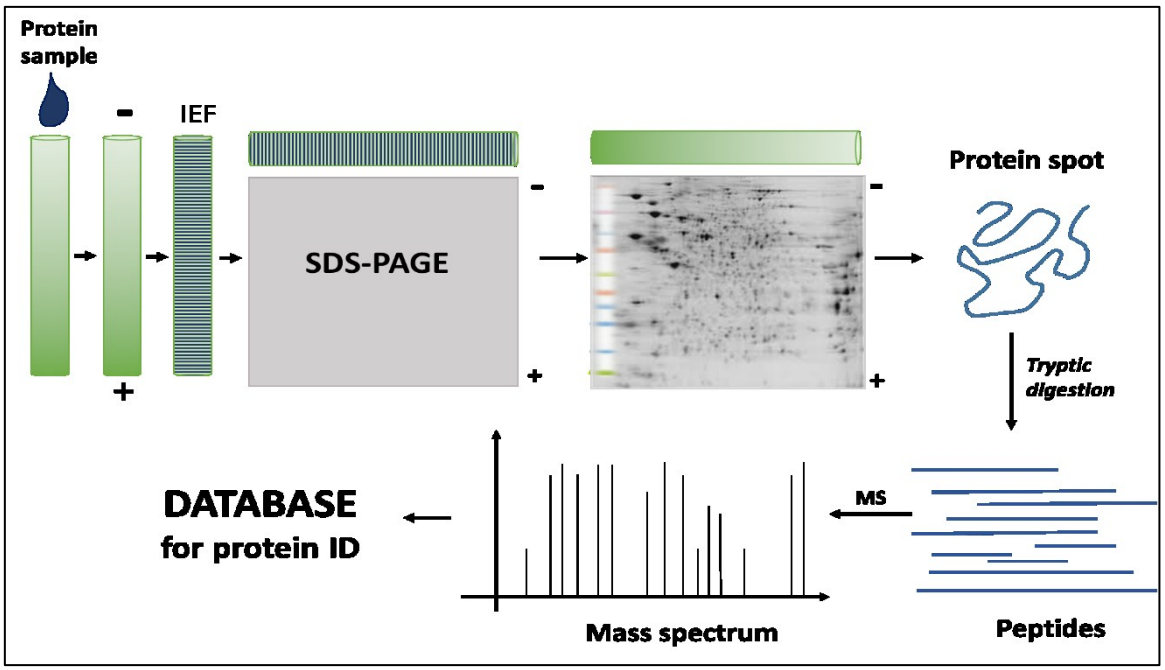


Figure 3.1: The schematic workflow of two-dimensional gel electrophoresis.

Nevertheless, the use of 2D GE in recent years has become less common due to its limitations. The principal disadvantage of this technique is lack of reproducibility between the gels. No two gels are superimposable in terms of their spot patterns, making it difficult to distinguish between the spots arising from experimental variation (due to differences in intra- and inter- laboratory techniques) and biological variations (due to protein abundance in two

different conditions) (Dhingra, Gupta et al. 2005). Furthermore, this technique does not well separate proteins which are highly acidic, alkaline or hydrophobic making it difficult to employ in certain studies. Therefore, an alternative approach is to use a more reproducible and faster technique, one-dimension protein separation based on molecular weight followed by liquid chromatography-mass spectrometry (Abdallah, Dumas-Gaudot et al. 2012).

The gel-free technique is also known as shotgun proteomics and uses the combination of high pressure liquid chromatography (HPLC) and tandem mass spectrometry (MS/MS) for the analysis of complex peptide mixtures (Patterson and Aebersold 2003). In this technique, protein samples are proteolytically digested into peptides. The peptide mixture is then fractionated using reverse phase LC (RPLC) or multidimensional LC (LC/LC), depending on the complexity of the mixture. The separated fractions are then subjected to MS/MS where the peptide ions are identified (McDonald and Yates 2002). Recently, the analysis of a complex protein mixture has used the combination of gel-based and gel-free techniques. For example, after 1D protein separation, the bands of interest or the entire protein lane is excised into slices and each slice is subjected to proteolytic digestion. The peptide mixture is then further separated by LC prior to MS/MS analysis. This approach has greatly reduced the complexity of the protein mixture and increases the chances of identifying low abundance protein (Abdallah, Dumas-Gaudot et al. 2012) which turns out to be beneficial in many studies.

The application of proteomics in the field of bioremediation is gaining recent recognition. Differences in physiological states lead to differences in gene expression in an organism. These changes in expression induce different sets of proteins or enzymes in response to their environment (Fulekar and Sharma 2008). In bioremediation, proteomics plays major roles in studying the bacterial adaptations and responses to various environmental contaminants and in identifying the key protein or set of proteins involve in the degradation and uptake of the toxic compounds (Singh and Nagaraj 2006).

Exposure of microorganisms to different physical conditions such as extreme temperature in the presence of toxic compounds and heavy metals has been shown to trigger different physiological changes. For example, nickel stress on *Pseudomonas putida* UW4 has down-regulated 81 out of 1702 proteins, many of which are vital for cellular activities like protein synthesis and folding, cell division and protein-protein interactions. Moreover, various anti-oxidative stress proteins were upregulated in response to nickel exposure (Cheng, Wei et al. 2009). In another study, *Halobacterium sp.* NRC-1 which is capable of surviving in high concentrations of many halogenated compounds was subjected to high temperature during growth. Through Differential In Gel Electrophoresis (DIGE) technique, five heat shock proteins were identified to be upregulated to allow cells to survive under extreme heat stress (Shukla 2006).

Studies have also shown that proteomics plays a great role in elucidating the degradation pathway for various compounds. For example, difference in the 2D proteome pattern for *Rhodococcus sp*. TFP grown in the presence of phthalate compared to cultures grown in tetralin or naphthalene suggested that different metabolic pathways were involved in the degradation of different monoaromatic compounds and polyaromatic compounds. DIGE analysis of phthalate and glucose grown cells indicated that 56 proteins showed significant difference in expression, 14 of which were responsible for the metabolism of phthalate. Among the 14 proteins, three were found to be responsible for the metabolism of phthalate to protocatechuic acid (PCA) and another four were involved in the production of acetyl-CoA from PCA. Others were chaperonins induced as the stress response (Tomas-Gallardo, Canosa et al. 2006).

In another example, toluene degradation by a consortium of sulphate-reducing anaerobes was studied by proteomics. Combination of 2D GE and a molecular biology technique known as Terminal Restriction Fragment Length Polymorphism (T-RFLP) in this study was helpful in identifying *Desulphobulbaceae* as the main toluene degrader in the consortium (Jehmlich, Kleinsteuber et al. 2010). The use of one-dimensional GE and LC-

MS/MS has also proven to aid in the elucidation of the metabolic pathway of pyrene in *Mycobacterium vanbaalenii* PYR-1. Proteomic and genomic analysis of the pyrene grown cells showed more than two fold upregulation of 142 proteins, 27 of which are potential enzymes for metabolism of pyrene. Analysis from this study concluded that the degradation of pyrene to its central metabolites occur through o-phthalate and β-ketoadipate pathway, giving rise to comprehensive framework of pyrene metabolism in *Mycobacterium vanbaalenii* PYR-1 (Kim, Kweon et al. 2007).

In this chapter, a proteomic approach was adopted to identify enzyme(s) that are responsible for the degradation of 1-CB in *R.* UKMP-5M. As described in Chapter 2, the traditional method of identification through isolation of cell lysate activity was unsuccessful. Given that the dehalogenation of 1-CB in *R.* UKMP-5M was repressed in the cells grown in LB but was induced in the cells grown in LB supplemented with 1-CB, the requirement for a differential expression study (uninduced vs. induced) has now been established. Therefore, the enzyme(s) involved in the degradation of 1-CB are expected to be identified as proteins which are uniquely present in the induced condition. The initial approach to identify these proteins using 2D GE was not convincing due to the reasons mainly described earlier, the lack of reproducibility. Therefore, a one-dimensional protein separation coupled with LC-MS/MS was used to identify proteins which are present in one condition and not in the other. Data was initially analysed using the Mascot database (Perkins, Pappin et al. 1999) and further analysis was performed using Scaffold 4.0 (trial version) (Searle 2010) to display those protein unique to induced or uninduced samples. Proteins thus identified through proteomic studies were then validated by confirming the presence or absence of their transcript in the total RNA.

3.2 Materials and methods

3.2.1 Bacterial cell preparation

A 2 ml preculture (OD_{600nm}=0.5) of *R*. UKMP-5M was inoculated into either 50 ml LB or LB supplemented with 5 mM 1-CB. Reaction flasks were incubated at 30^0C with agitation (180 rpm). The resulting cells were harvested at their late exponential phase (after 24 hours) by centrifugation for 30 minutes at 4000 rpm and washed twice with 40 ml deionised water before protein extraction. All experiments were repeated in triplicate.

3.2.2 Protein sample preparation for 2D GE

Cells were resuspended in 0.5 ml lysis buffer containing 7 M urea, 2 M thiourea, 4% CHAPS, 0.5% IPG buffer and 40 mM DTT (added immediately prior to use). Protein extracts were prepared by disrupting the cells with a sonicator (Jenco Ltd) at 15% amplitude for 3 minutes at 5.0 s pulse. Unbroken cells and debris were removed by centrifugation at 35,000 rpm for 1 hour at 4°C. The crude extract was then stored in -20°C prior to proteomic analysis.

3.2.3 Protein concentration assay

Protein sample concentrations were measured using the Bio-Rad protein assay reagent. The working solution was prepared by diluting one part of the reagent with four parts of deionised water and filtering through Whatman filter. A protein calibration curve was constructed using known concentrations of bovine serum albumin (BSA) from 0 to 0.5 mg/ml and the assay was performed in a BD 96-well plate. 200 µl of working reagent was added to 20 µl of protein samples and the colour was allowed to develop for 10 minutes before absorbance measurement at 595 nm.

3.2.4 Rehydration of IPG strip

11 cm IPG strips (GE Healthcare) were rehydrated on a reswelling tray with 200 µl of rehydration buffer containing 7 M urea, 2 M thiourea, 4% CHAPS, 0.5% IPG buffer and 50 µl of 1% bromophenol blue as the tracking dye. IPG strips were covered with mineral oil to prevent the evaporation of buffer and drying of strips. Rehydration was carried out at room temperature for a minimum of 10-20 hours (normally left overnight).

3.2.5 First dimension separation - Isoelectric Focusing (IEF)

The IEF was performed using Ettan IPGphor3 (GE Healthcare). The rehydrated IPG strips were placed on a manifold ceramic tray covered with mineral oil with the gel side up and the anodic end of the strips orientated towards the anodic side of the machine. 100 µg protein sample was loaded through the cup-loading method on each IPG strip. Moistened paper wicks were placed at both ends of the strips and electrodes were placed on top of the wicks. IEF was carried out using an incremental voltage program depending on the length and pH range of the IPG strips used. The voltage program used in this study is given in Table 3.1

Table 3.1: The voltage program used for IEF of 11 cm IPG strip pH 4-7. The maximum current per strip did not exceed 0.05A.

Voltage mode	Voltage (V)	Duration (hh:min)
Step-and-hold	500	00:30
Gradient	1000	01:30
Step-and-hold	5000	02:50
Gradient	8000	04:00
Step-and-Hold	8000	02:15
Step-and-Hold	300	12:00

3.2.6 Equilibration of IPG strips prior to second dimension separation

The focused IPG strips were equilibrated twice in SDS equilibration buffer prior separation on an SDS gel. The equilibration buffer contains 6 M urea, 75 mM Tris-HCl pH 8.8, 30% (v/v) glycerol, 2% (w/v) SDS and bromophenol blue. In the first equilibration, DTT (0.1 g/10 ml SDS buffer) was added and equilibrated for 15 minutes to reduce the disulphide groups on the proteins. This was then followed by the second equilibration with iodoacetamide (0.25 g/10 ml SDS buffer) for another 15 minutes to alkylate the thiols.

3.2.7 Second dimension separation – SDS PAGE

The equilibrated strips were then inserted into Bis-Tris 4-12% SDS-PAGE pre-cast gradient gels (Bio-Rad) and sealed with 1% (w/v) agarose sealing solution. 1X MES running buffer was added into the electrophoresis tank and 10 µl of a protein marker solution (Precision plus Protein Prestained standards from Bio-Rad) was loaded in the marker well. Two gels were run together at room temperature at 40 V for 30 minutes followed by 120 V for 60 minutes. Electrophoresis was considered done when the tracking dye reached the bottom of the gels.

3.2.8 Protein gel staining

The cassettes were broke open and gels were stained using Coomassie instant blue solution (Expedeon) for an hour at room temperature with mild shaking and then destained overnight in deionised water to reduce the background. The stained gels were then imaged on a gel documentation system (Syngene Diversity Imaging System). The protein expression profiles for *Rhodococcus* strains in both induced and non-induced states were compared and the differentially expressed proteins were located for further analysis (Gorg, 2004).

3.2.9 Protein sample preparation for 1D GE

Cells as described in section 3.2.1 were resuspended in 1 ml of 50 mM Tris. sulphate buffer pH 8.0 and disrupted using a sonicator at 15% amplitude for 3 minutes at 5.0 s pulse. Unbroken cells and debris were removed through centrifugation at 13,000 rpm for 20 minutes at 4^{o}C. The crude extract was then stored in -20^{o}C prior to one dimensional GE.

3.2.10 One-Dimensional Gel Electrophoresis

Protein samples for SDS-PAGE analysis were mixed with 25% NuPAGE LDS Sample Buffer (4X) (Invitrogen) and 1% NuPAGE Reducing Agent (10X) (Invitrogen). Samples were then heated at 70^{0}C for 10 minutes prior to loading into wells of a pre-cast 12% SDS PAGE gel. Electrophoresis was performed at 150 V for 40 minutes. The gels were stained with coomassie instant blue as described in section 3.2.8.

3.2.11 Mass spectrometry

The protein identification by mass spectrometry was performed by Dr. Benjamin Thomas and Dr. Svenja Hester, in the Central Proteomics Facility (CPF), Dunn School of Pathology, University of Oxford.

3.2.11.1 Excision of protein bands and proteolytic digestion

The differentially expressed protein bands were identified from the gels. Protein bands of interest were carefully excised using a clean stainless scalpel blade. The pieces of gels were placed into individual eppendorf tubes and destained with 25 mM ammonium bicarbonate in 50% acetonitrile by vortexing gently for 30 minutes at room temperature. The supernatants were discarded and the destaining procedure was repeated 6-7 times. 100% acetonitrile was added to dehydrate the gel pieces. Gel pieces were then dried using a speedvacuum (Thermo Scientific) for 10 minutes and were treated with 100 µl of 10 mM DTT for 30

minutes at 37^0C. Supernatant was discarded and gel pieces were washed twice with 25 mM ammonium bicarbonate solution. Gel pieces were then treated with 100 µl of 55 mM iodoacetamide for 60 minutes in the dark and the supernatant was discarded. Gel pieces were then washed twice with 100 µl of 25 mM ammonium bicarbonate in 50% acetonitrile solution followed by another wash with 100 µl of 100% acetonitrile and dried in a speedvac for 10 minutes. 20 µl of 10 ng/µl sequencing grade modified trypsin (Promega) in 25 mM ammonium bicarbonate was added to the gel pieces and incubated at 37^oC for overnight. The digestion was stopped by the addition of 1 µl formic acid. The supernatant containing peptide fragments were then collected into new tube and used for MS analysis.

3.2.11.2 Tandem mass spectromertry (MS/MS)

Peptide mixtures were analysed on an Ultimate 3000 RSLCnano HPLC system (Dionex) operated in direct injection mode coupled to a QExactive Orbitrap mass spectrometer (Thermo Electron). Samples were resolved on a picotip analytical column (75 µm i.d. x 25 cm) (New Objective) which was packed in-house with ProntoSIL 120-3 C18 Ace-EPS phase, 3 µm bead (Bischoff Chromatography). The system was operated at a flow-rate of 300 nl/min. A 120 min gradient was used to separate the peptides. The mass spectrometer was operated in a “Top 10” data dependent acquisition mode. The precursor scans were performed in the orbitrap at a resolving power of 70,000, from which the ten most intense precursor ions were selected by the quadrupole and fragmented by higher-energy collision dissociation (HCD) at a normalised collision energy of 28%. The quadrupole isolation window was set at *3 m/z.* Charge state +1 ions and undetermined charge state ions were rejected from selection for fragmentation. Dynamic exclusion was enabled for 40 seconds.

3.2.11.3 Protein Identification

The resulting mass spectra data files were converted from .RAW to .MGF using the Proteowizard program (Chambers, Maclean et al. 2012) and were submitted to Mascot (www.matrixscience.com last accessed 25/10/2014). The Mascot software (Perkins, Pappin et al. 1999) matches the experimental data with the proteins in a database. In this study, the MS data were matched with a customised database prepared for *Rhodococcus*. The Mascot results were then imported to Scaffold 4.0 (Searle 2010) to classify the proteins by their molecular functions and to represent the unique proteins present in the induced and uninduced conditions in a Venn diagram.

3.2.12 Isolation of total *Rhodococcus* RNA

Cells as described in section 3.2.1 were resuspended in 0.5 ml Tris-EDTA buffer, pH 8.0 and 1 ml bacteria RNA protect reagent (Qiagen) and incubated at room temperature for 5 minutes. Cells were then recovered by centrifugation at 4000 rpm for 20 minutes. Total RNA was isolated using Qiagen RNeasy RNA mini kit based on manufacturer's protocol. Amount and purity of total RNA isolated was determined using Nanodrop (Thermo Scientific) based on the 260/280 ratios.

3.2.13 Complementary DNA (cDNA) synthesis

The cDNA synthesis was done with reverse transcription system (Promega) using dT primer based on manufacturer's instruction manual. 1 µl of RNA from section 3.2.12 was used in a total reaction volume of 20 µl per sample.

3.2.14 Polymerase Chain Reaction (PCR)

Oligonucleotides were purchased from Invitrogen and all PCR reactions were carried out in sterile 0.2 ml PCR tubes (VWR International) using an Eppendorf Mastercycle Gradient thermal cycler. Reactions using One *Taq* Polymerase (New England Biolabs) were set up based on manufacturer's protocol for a total volume of 25 µl per sample. The thermocycling conditions are given in Table 3.2. Each PCR product was mixed with 5 µl 6X DNA loading dye (New England Biolabs) and run on 1% agarose gel prepared in 1X TAE buffer. A 1 kbp DNA ladder (Bioline) was used and the agarose gel was stained using SYBR safe DNA gel stain (Invitrogen).

Table 3.2: Thermocycling conditions for PCR amplification using One *Taq* Polymerase

Step	Temperature (^{0}C)	Time
Initial Denaturation	94	30 seconds
Denaturation	94	30 seconds
Annealing	61	30 seconds
Extension	68	1 minute
Repeat for 30 cycles		
Final Extension	61	5 minutes
Hold	4	

3.3 Results

3.3.1 Protein quantification

A protein standard was constructed using known concentration of BSA from the range 0 to 0.5 mg/ml. A linear relationship between the absorbance at 595 nm and the concentration of BSA was observed as shown in Figure 3.2. The amount of protein loaded onto gels was determined from this calibration curve.

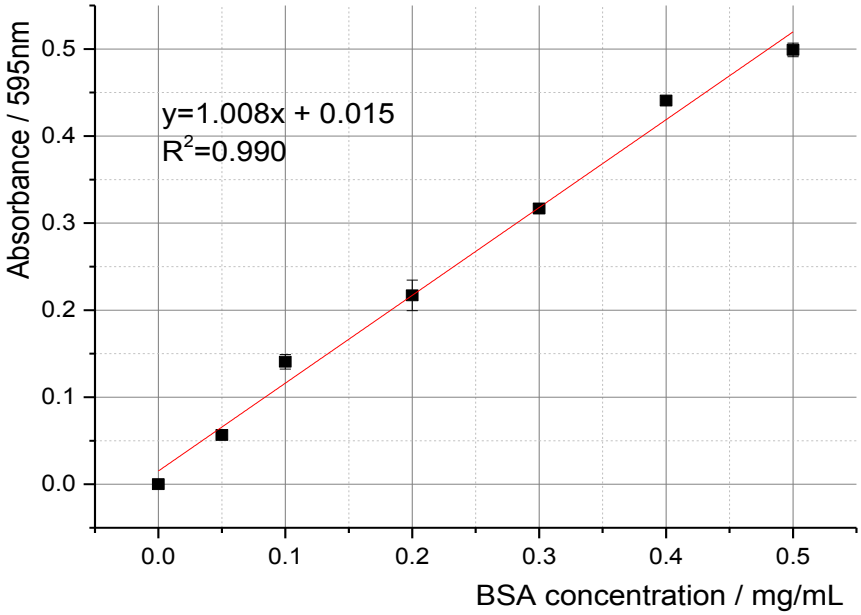


Figure 3.2: Calibration curve for protein assay.

3.3.2 Two-Dimensional Gel Electrophoresis

Initially, multiple attempts were made to achieve a satisfactory 2D separation. However, due to various factors such as interference from contaminating species like mycolic and nucleic acids in the protein sample, the IEF step did not yield a good separation as the majority of the protein samples failed to enter the IPG strip. A number of protein precipitation methods such as a commercial 2D clean-up kit (GE Healthcare) and trichloroacetic acid (TCA) were used to remove the interfering components but resolubilisation of the protein pellets were then difficult and inconsistent, which resulted in irreproducible gels.

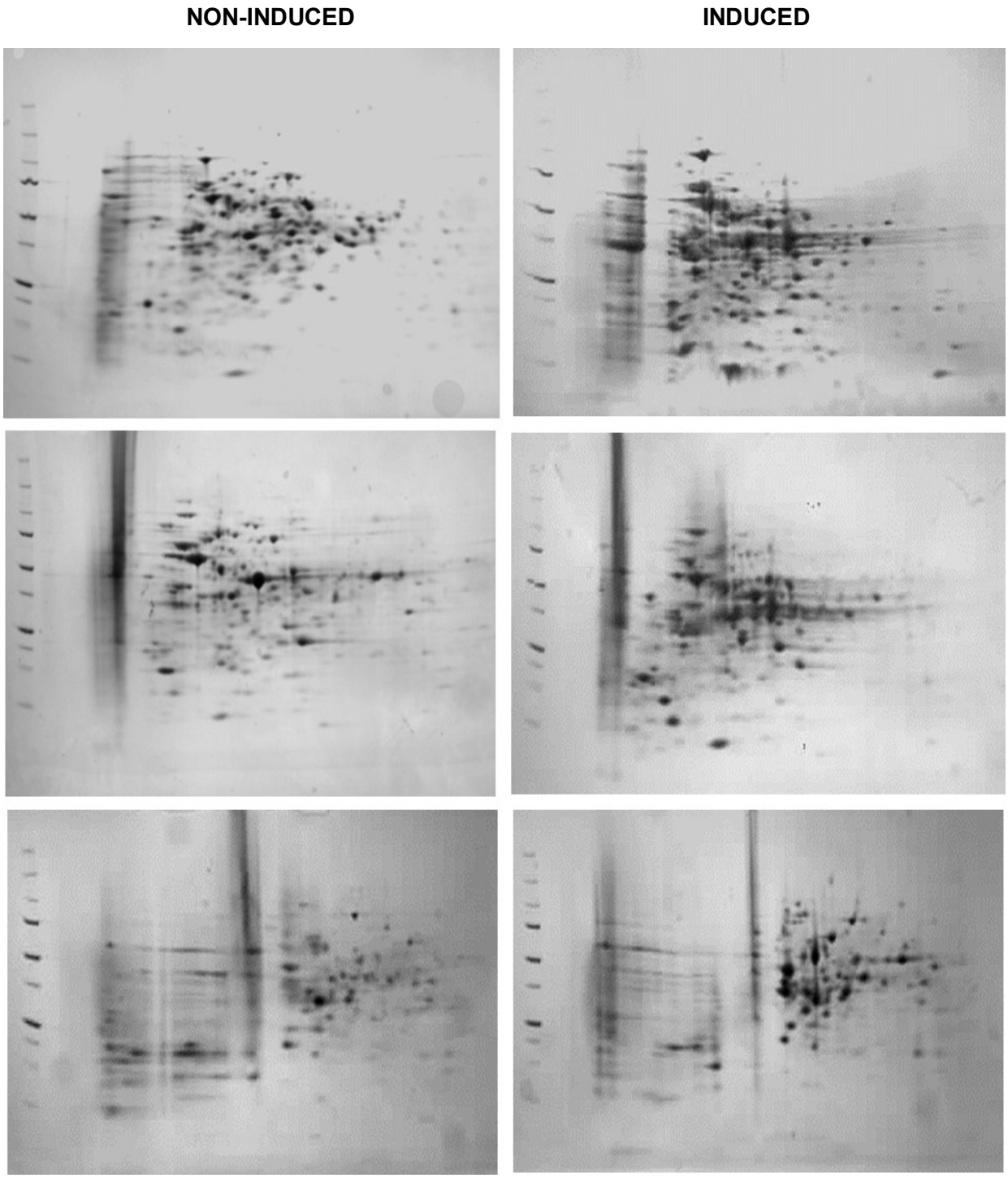


Figure 3.3: The non-reproducibility of 2D GE of protein samples prepared from *R.* UKMP-5M cells grown in non-induced condition (LB) and induced condition (LB supplemented with 5 mM 1-CB). Samples were separated on 11 cm IPG strips pH 4-7 followed by 4-12% SDS-PAGE. Protein detection was with Coomassie instant blue.

The very large variation in the number of spots and the irreproducibility between the gels is illustrated with the examples shown in Figure 3.3 made this technique difficult to use in this study. The problem was exacerbated by the number of overlapping spots due to the size of separation system (11 cm gels) and the inconsistency of the patterns between the gels which will lead to incorrect spot(s) picking, and in turn incorrect protein identification. Therefore, a more reproducible technique, which is 1D protein separation, was chosen for protein identification.

3.3.3 One dimension protein separation

1D SDS-PAGE analysis was performed to identify differences in the profile patterns between the uninduced and induced protein samples of *R.* UKMP-5M. Soluble protein from both conditions was extracted through sonication. Equal amount of proteins (15 µg) were loaded into the wells of the precast gel.

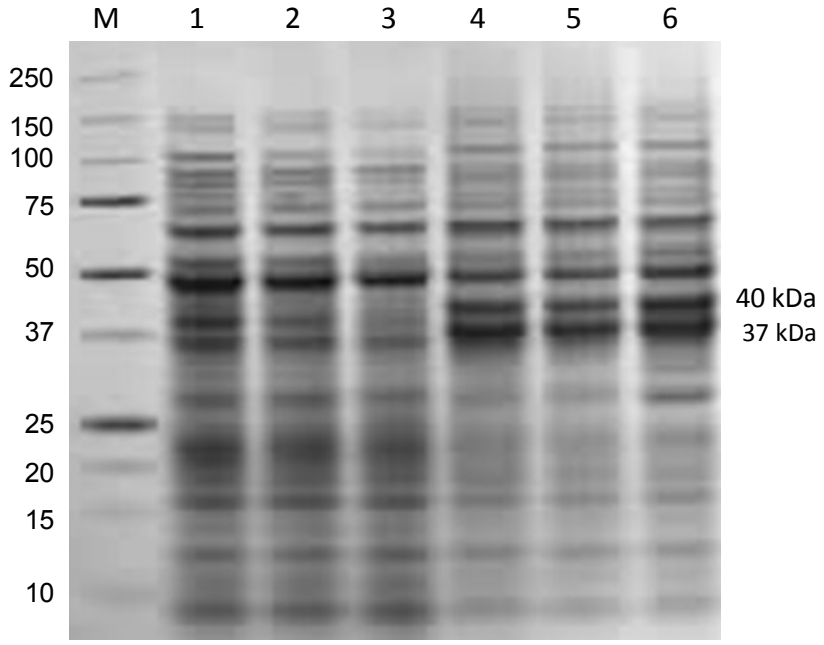


Figure 3.4: 1D GE of protein samples prepared from *R.* UKMP-5M cells grown in non-induced condition (LB) (lane 1-3) and induced condition (LB supplemented with 5mM 1-CB) (lane 4-6). Lane M represents protein standards. Samples were separated on 4-12% SDS-PAGE gradient gel. Protein detection was with Coomassie instant blue. *Pixelation is due to expanding the original .jpeg image.*

On comparing both protein profiles, there were two bands that appeared in the induced conditions which were poorly expressed in the uninduced condition. Distinct variations can be observed in the region of 37 kDa and 40 kDa. The increased staining of the induced bands is consistent with increased protein expression compared to the non-induced condition. Therefore, these bands were considered as the potential bands to be further analysed using nano LC-MS/MS.

3.3.4 Protein identification from one dimension electrophoresis gels

Protein identification in this study was done using Nano Liquid Chromatography-Tandem Mass Spectrometry (nano LC-MS/MS). Two bands from each growth condition extract (Figure 3.5) were excised and were treated with trypsin to generate peptide fragments prior to injection into mass spectrometry. The resulting uninterpreted spectra were searched against a custom database in Mascot for *Rhodococcus* (Mascot data and the database files are available on a CD). The MS/MS ion search for ESI-Quad-Tof data allowed for a maximum of two trypsin misclevages per peptide, a peptide mass tolerance of ±20 ppm, ion fragment mass tolerance of ±0.02 Da, carbamidomethylation of cysteine and variable oxidation of methionine. Analyses were done with a significance threshold of $p<0.05$. Results from the Mascot analysis were further analysed using Scaffold 4.0 program (trial version). The threshold values used for protein identification in Scaffold was highly restrictive i.e. minimum protein identification probability of 99%, minimum peptide number of two and a minimum peptide identification probability of 95%.

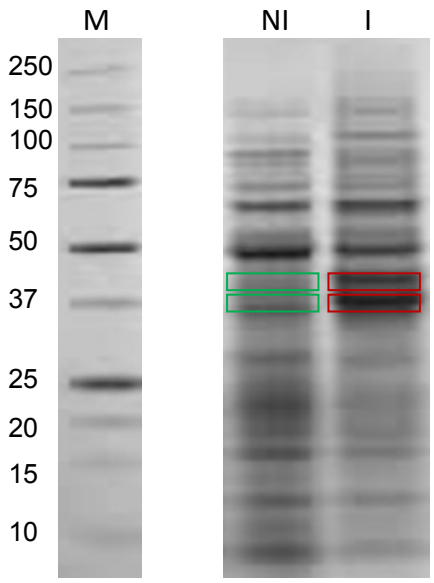


Figure 3.5: Protein bands which were excised for identification using nano LC-MS/MS identification. Bands excision was done in proteomics workstation under clean condition. □- bands from uninduced (NI) condition and □- bands from induced (I) condition.

Unlike 2D gels, identifying a specific protein from 1D protein gels can be challenging since each band contains several proteins depending upon the complexity of the sample. Using the above approach, a total of 107 and 102 proteins were identified in the non-induced and the induced protein samples, respectively. Gene ontology (GO) annotation for *Rhodococcus jostii* RHA1 was obtained from BioCYC database collection (http://biocyc.org/ last accessed on 25/10/2014) and used to assign the biological functions of the proteins (Figure 3.7). A full list of proteins present in both conditions is given in Appendix I and II. Protein hits for keratin and trypsin were eliminated from the list as they were considered to be contaminants introduced during the experiment. List of unique proteins present in the non-induced and induced protein samples of *R.* UKMP-5M is given in the Table 3.3 and Table 3.4.

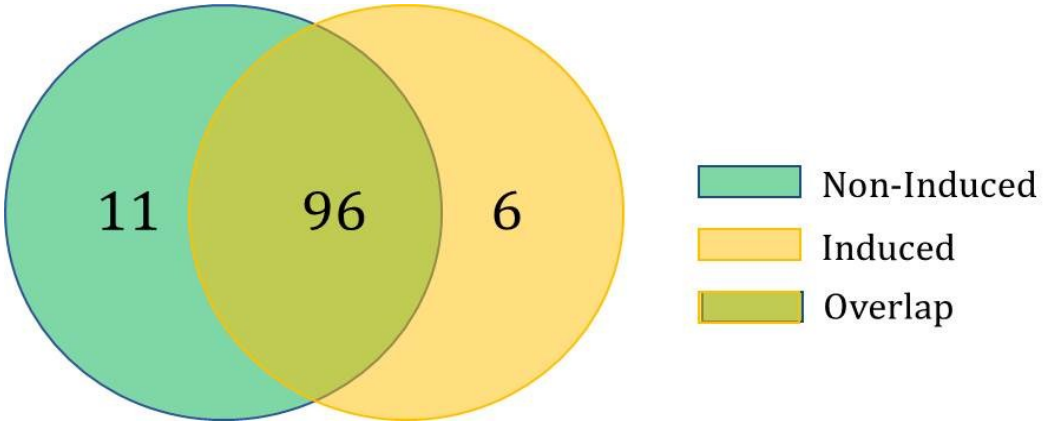


Figure 3.6: A Venn distribution representing the number of unique and common proteins identified through nano LC-MS/MS. 96 proteins are identified as common proteins present in both induced and non-induced samples of *R.* UKMP-5M at 37 and 40 kDa. 11 proteins are uniquely present in the non-induced condition and six proteins are uniquely present in induced conditions.

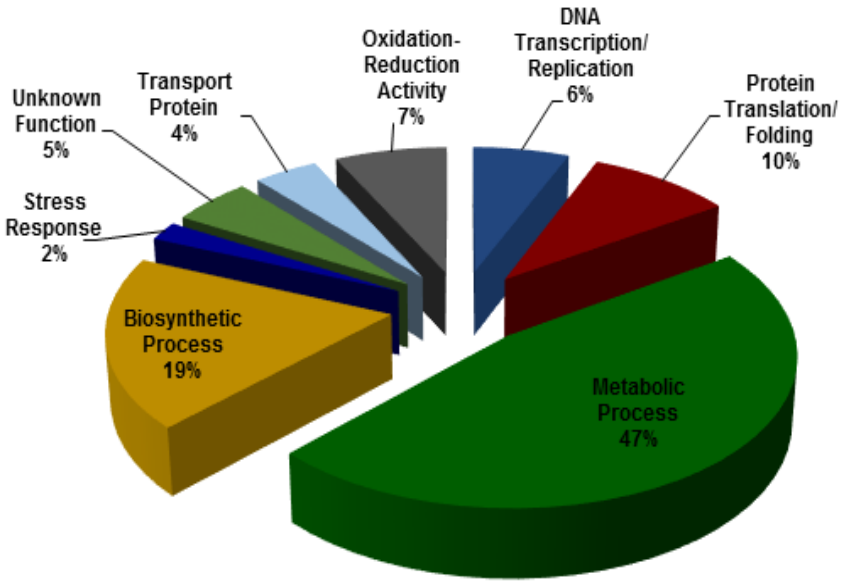


Figure 3.7: Pie chart showing the biological functions of common proteins identified in both induced and non-induced protein samples of *R.* UKMP-5M at 37 and 40 kDa according to GO annotation for *R. jostii* RHA1.

Table 3.3: List of proteins present in the two bands of 1D SDS-PAGE gel in the uninduced protein sample that were not detected in the induced condition.

Acc. No.	Protein Name	MW (kDa)	Match score (%)	No. of Peptides	Sequence Coverage (%)
RHA1_ro07148	aspartate carbamoyltransferase	34	100	2	7.3
RHA1_ro07176	hypothetical protein	35	100	3	14
RHA1_ro01340	acetyl-CoA C-acyltransferase	42	100	2	7.1
RHA1_ro03025	acyl-CoA dehydrogenase	45	100	3	13
RHA1_ro04539	3-HSA hydroxylase, oxygenase	42	100	2	5.8
RHA1_ro07235	probable arginine transport ATPase	36	100	3	9
RHA1_ro05843	probable cystathionine beta-synthase	48	100	2	5.9
RHA1_ro03865	catechol 2,3-dioxygenase	40	100	2	7.5
RHA1_ro01242	phosphate starvation-induced protein	39	100	2	9.5
RHA1_ro00242	possible 3-oxoacyl-[acyl-carrier-protein] reductase	32	100	2	6.7
RHA1_ro01932	probable acyl-CoA dehydrogenase	41	100	2	5.8

Table 3.4: List of proteins present in the two bands of 1D SDS-PAGE gel in the induced protein sample that were not detected in the uninduced condition.

Acc. No.	Protein Name	MW (kDa)	Match score (%)	No. of Peptides	Sequence Coverage (%)
RHA1_ro00445	**hypothetical protein**	**40**	**100**	**8**	**22**
RHA1_ro00443	**monooxygenase hydroxylase**	**42**	**100**	**6**	**20**
RHA1_ro03447	possible myo-inositol-1-phosphate synthase	39	100	5	16
RHA1_ro02534	**alkane 1-monooxygenase**	**47**	**100**	**3**	**7.8**
RHA1_ro06757	possible phage shock protein	29	100	2	11
RHA1_ro10159	possible 5-oxoprolinase	86	100	2	2.8

Since the objective of this approach is to identify enzyme(s) which are responsible for degradation of 1-CB, proteins which were present in the induced protein sample and that were not detected in the uninduced protein sample became the focus of the study. Among the six proteins identified in the induced sample, three proteins seem to be possible candidates for the dehalogenation reaction. As discussed in Chapter 2, 1-CB degradation in *R.* UKMP-5M is an oxygen dependent reaction and from Table 3.3, there were two types of oxygenases present in lysate prepared from cells that were capable of 1-CB dehalogenation;

a monooxygenase hydroxylase (RHA1_ro0043) with a molecular weight of 42 kDa was identified along with an alkane 1-monooxygenase (RHA1_ro02534) with a molecular weight of 47 kDa. In addition, a hypothetical protein (RHA1_ro00445) with a molecular weight of 40 kDa was also identified. The peptide sequence coverage for all the three proteins are shown in Figure 3.8A, B and C.

A

RHA1_ro00445 (100%), 40,166.0 Da
hypothetical protein|RHA1_ro00445|RHA1_ro00445[NC_008268]
8 exclusive unique peptides, 12 exclusive unique spectra, 13 total spectra, 78/348 amino acids (22% coverage)

MYQKDGQKYF IVDGHVHVWD ARESNQKNVH GKQFIDCFYD YHKNLSPEEV VWDYDTYTYY
GSERFERDIF VEGYVDHAIF QATLLSDFYH NGFGRTDEAL ALVAKHPGKL TYNHAYDPRH
EEAGLEQLRK DADRMNLQGV KLYTAEWHGD SRGYKLDDPW SRRYLEECIE LGIKNIHVHK
GPTIRPLDRD AFDVSDIDKV ATDYLDLRFV VEHVGLPRLE DFCWIATQES NVYGGLAVAL
PFIHTRPRYF AQIIGELLYW IGEDKILFGS DYALWTPKWL IEKFVDFQIP EDMLTDYAPI
TALQKQKILG LNAAAALYDI D VPADLQLPEP AGQEGVEVAA GAREVVSS

B

RHA1_ro00443 (100%), 41,696.2 Da
monooxygenase hydroxylase|RHA1_ro00443|RHA1_ro00443[NC_008268]
6 exclusive unique peptides, 6 exclusive unique spectra, 6 total spectra, 74/368 amino acids (20% coverage)

MTATAESKQR SFPKIEFTDS EAGALEFPSS RSRTFTYYTP AKKRSTMYED VTVDVQPDPD
RHLSQGWIYG FGDGPGGYPQ EWTAAKSSNW HAFLDPNEEW DQTIYRNNSK VVHQVELCLS
NAKRARVYDG WNTPWLTFIS RNLGAWMHAE NGLALHVFTS IQRSCPTNMI NTAVAVNAAH
KMRFAQDLAL FNLDLSEATE NFDGTAHKEV WQSAPEWQPT REVVERLTAV PDWCELLFGS
NIVFEQLVGT LFRSELVMQI AAGNGDYITP TIVGTGEHDY DRDLAYTRNL FRLLTRDPEH
GEANKELFGT WLAIWVPRCL DAARALQPIW SQPADKAITF ATSFDAATDK FRSLLEDLGL
DIPKELDQ

C

RHA1_ro02534 (100%), 46,819.0 Da
alkane 1-monooxygenase|alkB|RHA1_ro02534[NC_008268]
3 exclusive unique peptides, 3 exclusive unique spectra, 3 total spectra, 32/410 amino acids (8% coverage)

MTTSNISRMT GEAPPEAWRD RKRYLWLMGL IPPTSLFLAA GLVWAFNQLG WSAAAPVWWW
IGALLLFGLL PLLDRFFGPD GQNPPEEVME QLENDRYYRY CTYVFIPFQM ASLVFACYLW
SADNLSWLGI DGGLGLVSKV GVAFTVAVMG GIGINTAHEM GHKKTELERW LAKITLAQTF
YGHFYIEHNR GHHVRVATPE DPASSRFGES FWTFLPRSVW GSLKSSWELE KTRMQRLGKS
TWSIHNDVLN AWLMSVALFG VLIAIFGPVV IPFLIIQAVY GFSLLETVNY LEHYGLMRQK
TASGRYERCA PSHSWNSDHI VTNIFLYHLQ RHSDHHANPT RRYQTLRSMD GAPNLPSGYA
SMITLAYFPP LWRKVMDHRV LDHYDGDITR VNIQPGKREK VLAKYSGGVR

Figure 3.8: FASTA format sequences. (A) Hypothetical protein with 22% of sequence coverage; (B) monooxygenase hydroxylase with 20% and (C) alkane-1-monooxygenase with 8% peptide sequence coverage. The matched peptide sequences are highlighted in yellow.

3.3.5 Validation of proteomics data through transcript analysis

Possible proteins for 1-CB degradation identified in the induced condition through nano LC-MS/MS experiment were validated by identifying the absence and presence of transcript in both conditions. Total RNA from LB (uninduced) and LB supplemented with 1-CB (induced) grown cells were isolated and analysed by RT-PCR. For each sample, 1 µl of RNA was reverse transcribed to cDNA, which was then used as a template in a PCR reaction. Negative controls were prepared for each condition by using RNA as the template. This is important to ensure the products on the agarose gel are due to the mRNA expression and not the result of genomic DNA contamination. A list of primers used in transcript analysis of hypothetical protein, monooxygenase hydroxylase, alkane 1-monooxygenase and threonine synthase (housekeeping gene) is given in Table 3.5. Result of this analysis is presented in Figure 3.9.

Table 3.5: List of oligonucleotides used for RT-PCR analysis in this section. All oligonucleotides were purchased from Invitrogen. Working primer concentration was 10 µM. F represents forward sequence and R represents reverse primers.

Gene Product	Primer	Amplicon size (bp)	Nucleotide sequence (5' to 3')
Hypothetical protein (Ah)	*Ah* - F	1023	ATGTACGAGAAGGACGGCCAGCAGT
	Ah - R		CGACGACACCGCCTCTCGA
Monooxygenase hydroxylase (Moh)	*Moh* - F	1082	ATGACCGCAACCACCGAGTCCAA
	Moh - R		TCACTGGTCCAACTCCTTCGGGATGT
Alkane 1-monooxygenase (Alm)	*Alm* - F	1071	ATGACCACCAGCGACATCG
	Alm - R		GTGGTACAGGAAGATGTTGGTCA
Threonine synthase (Trs)	*Trs* - F	1021	TGATCGCTGTGACCAGCGTCCTGA
(housekeeping gene)	*Trs* - R		GTGACGGTGCACACCACCTTCA

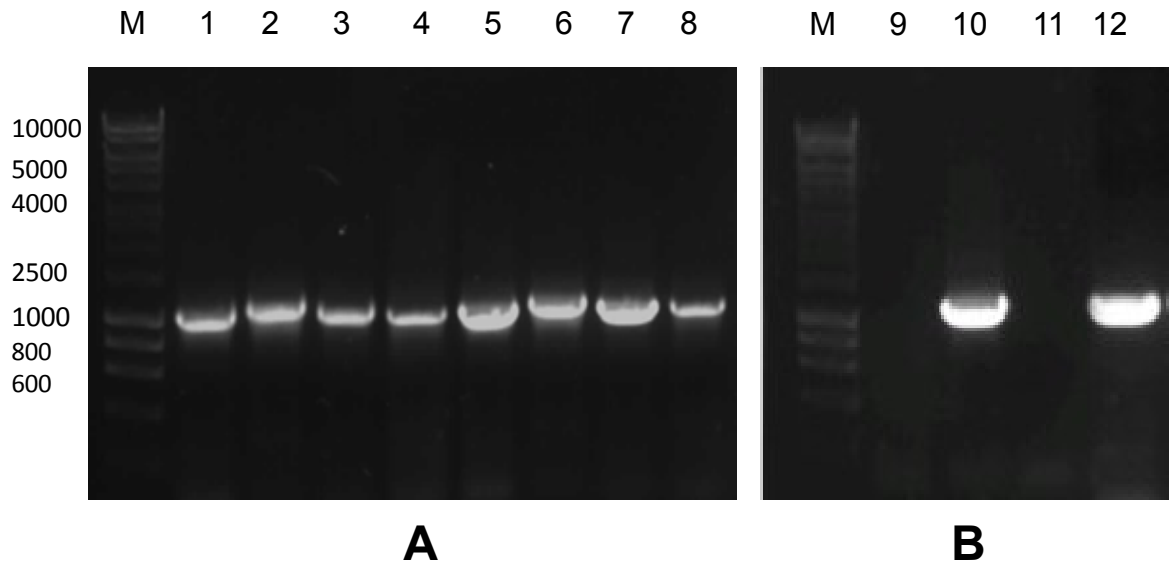


Figure 3.9: Transcript analysis of proteins identified through proteomics. (A) RT-PCR products using cDNA synthesised from RNA extracts of uninduced (lanes 1-4) and induced (lanes 5-8) cells of *R.* UKMP. Ah (lanes 1 and 5), Moh (lanes 2 and 6), Alm (lanes 3 and 7) and Trs (lanes 4 and 8). (B) Lanes 9 and 11 represent the RT-PCR products of Trs obtained using RNA as the template (negative controls) in uninduced sample (lane 9) and induced sample (lane 11) whereas lanes 10 and 12, represent products of Trs obtained using cDNA as the template, in uninduced sample (lane 10) and induced sample (lane 12). All samples were run on 1% agarose gel and stained with SYBR safe DNA gel stain. M represents DNA marker.

3.4 Discussion

The 1D protein analysis was done with the soluble protein extracted from cells of *R.* UKMP-5M grown in LB and LB supplemented with 1-CB, since the analysis of the membrane protein on the SDS-PAGE gel did not show any difference in band intensities between both conditions.

After MS analysis, the uninterpreted data was analysed using the Mascot search engine for protein identification. The Mascot software uses a probability based scoring algorithm to match experimental data with proteins in a database. In this study, a database containing FASTA format of proteins in *R. jostii* RHA1 was used for protein identification

since the draft genome sequence of *R.* UKMP-5M was still incomplete. The fundamental approach of the Mascot search engine is to calculate the probability that the observed match between the experimental data and each sequence database entry is a chance event (Perkins, Pappin et al. 1999). An ion score allocated for each observed peptide is -10Log (P), where P is the probability that the observed match is a random event with a significance threshold of $P<0.05$; and a protein score is the sum of ion scores for a given database entry. A high score represents a low probability that the observed match is a random event. For example, if 1.5×10^5 peptides fell within the mass tolerance window chosen in the search, then the score threshold value can be calculated as below:

$$\text{Threshold value} = (-10)\log(P), \qquad \text{where } P = \left[\frac{1}{20\,(1.5 \times 10^5)}\right] = \mathbf{65}$$

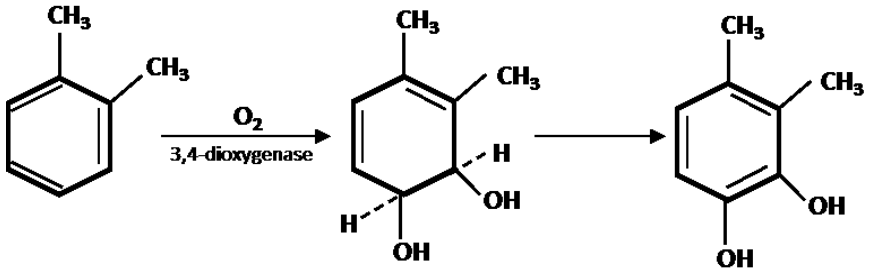
In this study, protein identifications were considered valid when at least two non-identical peptides had a combined ion score of greater than 52. The Mascot results were then further analysed using Scaffold trial version 4.0 that provides a side-by-side view of sample groups and representation of common proteins present in non-induced and induced protein samples through a Venn diagram (Figure 3.6).

The identification of monooxygenase hydroxylase and alkane 1-monooxygenase in the induced protein sample can be considered as an evidence that further supports the observation that the presence of oxygen is necessary to dechlorinate 1-CB. Oxygenases, also known as hydroxylases belong to the class of oxidoreductase enzymes. These enzymes catalyse the activation and insertion of oxygen atom(s) into an organic substrate, using molecular oxygen (O_2) as the oxygen donor (Torres Pazmino, Winkler et al. 2010). Oxygenases can be classified into two types; (i) monooxygenases, which catalyse the insertion of single oxygen atom into the reacting species. For example, the detoxification of DCA in *Pseudomonas sp.* DCA1 was mediated by a monooxygenase-type enzyme which

oxidizes DCA to chloroacetic acid via the formation of an unstable 1,2-dichloroethanol (Hage and Hartmans 1999).

$$CH_2ClCH_2Cl + ½\ O_2 \longrightarrow CH_2ClCH(OH)Cl \longrightarrow CH_2ClCHO + H^+ + Cl^-$$

or (ii) dioxygenases that mediate the insertion of both oxygen atoms from oxygen. For example, the first step in the transformation of xylene in *Rhodococcus sp.* DK17 is catalysed by a dioxygenase to form 3,4-dimethylcatechol (Kim, Chae et al. 2004).



In most studies, oxygenase-catalysed chloroalkane degradation occurs through monooxygenases which requires cofactors such as reduced nicotinamide adenine dinucleotide (NADH) or reduced nicotinamide adenine dinucleotide phosphate [NAD(P)H]. Examples of such monooxygenases are methane monooxygenases and alkane monooxygenase, which were identified in the present study. For instance, dehalogenation of 1-CB in *Mycobacterium vaccae* JOB-5 occurs with the aid of a propane monooxygenase through sub-terminal oxidation, yielding a 2-butanol as product (Vanderberg and Perry 1994). Therefore, the identification of these proteins in the induced sample further supports the hypothesis that 1-CB degradation in *R.* UKMP-5M is catalysed by a type of monooxygenase.

In order to validate the proteomic data, RT-PCR was performed to identify the absence and presence of the transcripts in non-induced and induced cells of *R.* UKMP-5M. The inability of the uninduced cells to dehalogenate 1-CB and the absence of monooxygenases in the uninduced protein sample are consistent. However, results in Figure

3.9 that shows the presence of transcripts in both non-induced and induced cells is inconsistent with the protein expression data. The absence of transcript in the negative controls prepared using RNA as the template confirms that the PCR products obtained are not due to genomic DNA contamination, and most probably due to the mRNA expressions. Multiples studies have demonstrated the lack of correlation between mRNA and protein expression levels. For example, in a study done to investigate the relationship between protein and mRNA abundance in yeast, prediction of the protein expression levels from simple or quantitative mRNA data was found inconsistent, given that some genes had the same mRNA values but the protein levels varied by more than 20 folds and vice versa (Gygi, Rochon et al. 1999). In another example, lack of correspondence was also observed between a set of mRNA and the protein expression profiles in *E. coli* (Lee, Shaw et al. 2003). Based on Figure 3.9, transcripts in the induced condition have stronger signals compared to the non-induced condition which in turn indicates the difference in the gene expression levels. The presence of transcript and the absence of protein in the non-induced condition maybe involves the control of gene expression at the RNA level, a process known as post-transcriptional regulation but further analysis needs to be conducted to understand the mechanism behind this.

In addition, there was an unknown protein identified with high sequence coverage of approximately 22%. Further 'blasting' of this hypothetical protein against National Center for Biotechnology Information (NCBI) database indicates that this protein has sequence homology with amidohydrolase (Ah). Homologous hypothetical protein has also been identified in *Rhodococcus* and several other bacteria. Whilst it is identified as hypothetical protein, the reason for its induction in response to 1-CB exposure in *R.* UKMP-5M was not clear. Therefore, gene cluster analysis based on the genome sequence of *R. jostii* RHA1 was done to identify if any of the gene(s) encoding proteins identified in the induced condition may belong to the same operon.

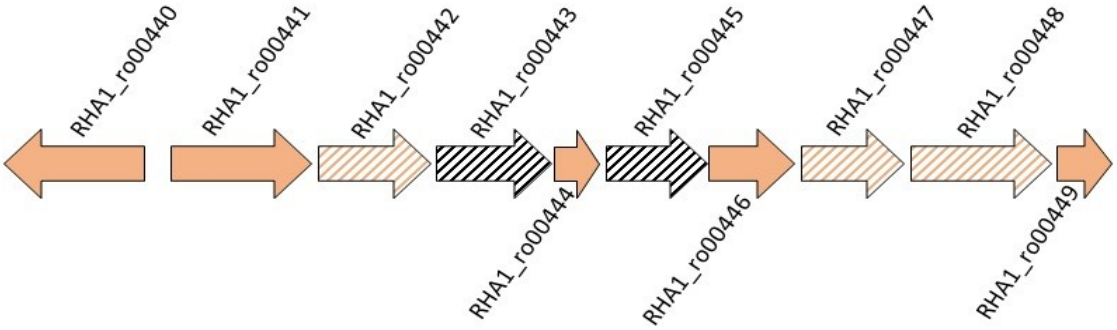


Figure 3.10: Gene cluster containing genes identified in proteomics study. Gene annotations are given in Table 3.6. (□) represents genes which were not identified; (◪) represents genes which were identified in the induced protein sample but with Mascot protein score below 52 and (◪) represent genes which were identified with Mascot protein score greater than 52.

Table 3.6: Gene annotations for genes present in the gene cluster shown in Figure 3.10

Accession No.	Gene Annotation
RHA1_ro00440	Transcriptional Regulator
RHA1_ro00441	Monooxygenase hydroxylase large subunit
RHA1_ro00442	Monooxygenase reductase
RHA1_ro00443	Monooxygenase hydroxylase small subunit
RHA1_ro00444	Monooxygenase coupling protein
RHA1_ro00445	Conserved hypothetical protein
RHA1_ro00446	Conserved hypothetical protein
RHA1_ro00447	Alcohol dehydrogenase
RHA1_ro00448	60kDa Chaperonin GroEL
RHA1_ro00449	Conserved hypothetical protein

Based on the gene cluster analysis in Figure 3.10, it clearly shows that both monoxygenase hydroxylase and the amidohydrolase encoding genes are neighbouring genes and therefore, possibly belong to the same operon. Apart from these two genes, there are also few other proteins from the genes that belong to the same operon identified in the protein analysis but with protein score of lower than the significant threshold. This maybe an indication that the 1-

CB degradation in *R.* UKMP-5M involves the upregulation or induction of a set of protein since most monooxygenase system contains a multisubunit hydroxylase containing diiron centers in its alpha-subunit and the requirement of a coupling protein for its optimal activity on an organic substrate (Coufal, Blazyk et al. 2000).

Although the evidence gathered in Chapter 2 and the results from the proteomic analyses suggest that the 1-CB dechlorination is likely to be an oxygenase-catalysed reaction, the lack of confirmation on the product from 1-CB metabolism and the induction of a protein with unknown function, which was later identified as amidohydrolase (Ah), lead to an indefinite conclusion. Furthermore, members of the Ah superfamily have diverse chemistry and physiologic roles. For example, besides a monooxygenase, a chlorohydrolase (AtzA), which belongs to the amidohydrolase superfamily has been reported to be able to catalyse the dechlorination of atrazine (a chlorinated pesticide) via a hydrolytic pathway to form hydroxyatrazine. Despite having a 98% sequence homology with a melamine deaminase (TriA), AtzA was not able to deaminate melamin and ammelin which were the substrates for TriA, but displays an exclusive ability to catalyse the dehalogenation of halo-substituted triazine ring (Seffernick, de Souza et al. 2001, Govantes, Porrua et al. 2009).

Hence, at this stage, there were two possible experimental designs to identify the enzyme(s) responsible for dehalogenation of 1-CB in *R*. UKMP-5M; one is the cloning and expression of Ah in *E. coli* to test the ability of the purified Ah to dehalogenate 1-CB; and the other is a gene knockout experiment, by constructing mutants with either the Ah, Moh or Alm encoding genes deleted from the genome of *R*. UKMP-5M. To begin with, recombinant protein technology techniques were used to overexpress and purify Ah for further study of the role and properties of this protein in *R.* UKMP-5M. The results of this investigation are presented in the following chapter.

CHAPTER 4 MOLECULAR CLONING AND EXPRESSION OF AMIDOHYDROLASE FROM *R.* UKMP-5M INTO *E. COLI*

4.1 Introduction

Recombinant DNA technology is one of the most common techniques used for functional study of a gene of interest. The concept of DNA cloning was first proposed by Peter Lobban and the first papers describing a successful production and intracellular replication of a recombinant DNA were published in 1972 and 1973 [reviewed in (Gavanji 2013)]. In this technique, the DNA of interest ('insert') is prepared either by isolating it from a DNA source and amplifying it through PCR or is synthetically assembled using individual blocks of oligonucleotides. Next a suitable vector typically a plasmid which is able to replicate in cells independently of cell division is digested using restriction enzymes to form a linear DNA molecule. The recombinant DNA is then created by joining the two pieces (insert and vector) together using a ligase enzyme. The resulting DNA molecule is then transformed into a host cell to allow its propagation to produce identical copies of the same recombinant molecule (Pandey Shivanand 2010) and protein that results from the expression of recombinant DNA is termed a recombinant protein (Gavanji 2013).

The development of this technique has been beneficial in many applications. In the field of bioremediation, molecular cloning techniques have been used widely for the production of enzymes which are capable of degrading environmental pollutants. For example, a naturally occurring microbial culture capable of utilising 1,2,3-trichloropropane (TCP) as growth substrate is yet to be discovered. However, a dehalogenase (DhaA) which has significant activity towards TCP transformation to form 2,3-dichloro-1-propanol (DCP) was cloned and introduced into a DCP-degrading *Pseudomonas putida* MC4. A packed-bed reactor inoculated with this engineered bacterium was able to degrade 95% of TCP (Samin, Pavlova et al. 2014). In another study, a novel arylamidase gene from *Paracoccus sp*. FLN-7 that is capable of hydrolysing amide pesticides was cloned and expressed in *E. coli*. The

purified recombinant arylamidase was characterised and preliminary investigations were done to test its role in amide pesticide hydrolysis (Zhang, Yin et al. 2012).

As mentioned in Chapter 2, a 'hypothetical' protein was found to be induced in response to 1-Chlorobutane (1-CB) exposure in *R.* UKMP-5M. Further analysis of the derived nucleotide sequence of this protein against the National Centre for Biotechnology Information (NCBI) database has shown that this protein shares sequence homology to that of an amidohydrolase from many *Rhodococcus* species. The amidohydrolase (Ah) superfamily comprises a diverse set of enzymes that react with organic compounds containing amide or ester functional groups and are widely distributed in the ecosphere. These enzymes primarily catalyse hydrolytic cleavage of amide bonds through activation of a water molecule via either a mononuclear or a binuclear metallic centre. Examples of Ah-type enzyme that have a divalent metal center include urease, phosphotriesterase and adenosine deaminase (Seibert and Raushel 2005). Although Ahs are enzymes that act upon amides, the role of this protein in degradation of 1-CB in *R.* UKMP-5M is not understood. Furthermore, this protein has often been classified as one of the uncharacterised proteins without a known function in several microorganism including *Rhodococcus* strains.

Hence, in this chapter, *Ah* from *R.* UKMP-5M was overexpressed in *E. coli* using the pET-28a(+) system. The traditional cloning procedures were carried out to yield a recombinant Ah with a His-tag fused at either the N-terminus or C-terminus and without a fusion tag. Small scale protein expression analyses were also done to detect the presence of the heterologous protein in both soluble and insoluble fractions.

4.2 Materials and Methods

4.2.1 General procedures for DNA manipulation

4.2.1.1 Sequence of amidohydrolase (*Ah*) gene

The sequence of *Ah* was obtained from the draft genome sequence of *R.* UKMP-5M which was kindly provided by researchers in King Abdullah University of Science and Technology (KAUST) as part of their effort to construct the complete annotated genome sequence for this bacterium.

4.2.1.2 *Ah* gene isolation from *R.* UKMP-5M

Ah from *R.* UKMP-5M was isolated by PCR using the complementary DNA (cDNA) prepared in section 3.2.13 as the template. This method was chosen as opposed to directly amplifying the *Ah* gene from genomic DNA as the mRNA had already been prepared as part of the transcript analysis. PCR was carried out using Q5 high-fidelity DNA polymerase (New England Biolabs) in a 50 µl reaction containing 1X Q5 reaction buffer, 0.2 mM dNTPs, 0.5 µm each forward and reverse primers (Table 4.1), 0.02 U/µl Q5 high-fidelity DNA polymerase, 1X Q5 high GC enhancer and 1 µg cDNA as the template. Reactions were carried out in Eppendorf Mastercycle Gradient thermal cycler. Initial denaturation was 98^0C for 30 s followed by 30 cycles of denaturation at 98^0C for 10 s, annealing at 72^0C for 30 s and extension at 72^0C for 30 s. A final extension at 72^0C for 2 minutes was included to ensure complete synthesis. Each PCR product was mixed with 10 µl of 6X DNA loading dye (New England Biolabs) and run on 1% agarose gel prepared in 1X TAE buffer. A 1 kbp DNA ladder (Bioline) was used as the marker and the agarose gel was stained using SYBR safe DNA gel stain (Invitrogen).

4.2.1.3 Recovery of amplified DNA from agarose gel

The DNA bands of *Ah* were excised from agarose gel with a sharp knife and purified using gel purification kit (Qiagen) based on the manufacturer's instructions. The resulting DNA was eluted with 50 µl of elution buffer provided with the kit.

4.2.1.4 Cloning vectors

Two types of vectors were used to clone *Ah* in this study. The first vector used was pET-28a(+) vector (Novagen) which carries an N-terminal 6x histidine residues (His-Tag) and an optional C-terminal His-Tag. The second vector used was pGEX-6P-1 (GE Healthcare) which produces a Glutathione S-transferase (GST) fusion tag at the N-terminal of the target protein.

4.2.1.5 Preparation of insert DNA

In total, three approaches to *Ah* cloning were done with the pET-28a system. The polyhistidine fusion tags were manipulated to yield a target protein with a fusion tag at either the N-terminus, C-terminus or in the absence of the His-Tag (native). Cloning with the pGEX-6P-1 system yields a fusion *Ah* with a GST moiety at the N-terminus. PCR reactions containing the respective primers were designed to introduce the relevant restriction sites. Reactions were carried out as described in section 4.2.1.2 with annealing at 5^0C below the primers' T_m for 30 s. A full list of primers used in this chapter is given in Table 4.1. PCR products were run on 1% agarose gel and purified from the gel as described in section 4.2.1.3.

4.2.1.6 Restriction endonucleases digestion of vector and DNA

Restriction enzymes were purchased from New England Biolabs and enzyme digestions were performed according to the manufacturer's protocol. All reactions were incubated at 37^0C for at least two hours. The digested vectors were dephosphorylated with calf intestinal alkaline phosphatase (CIP) (New England Biolabs) based on the manufacturer's protocol to prevent self-ligation. The 30 minutes CIP treatment at 37^0C was followed by inactivation of the enzyme at 75^0C for 10 minutes. Products were then purified with the same method described in 4.2.1.3.

4.2.1.7 DNA Ligation

The linearized vectors were ligated with their respective inserts using T4 DNA ligase (New England Biolabs) in a total volume of 15 µl according to the manufacturer's protocol with slight modifications. Ligations were carried out with vector: insert ratios of 1:1 and 1:3. Reactions were incubated at 65^0C for 5 minutes and cooled on ice for 2 minutes prior to adding 1.5 µl of T4 DNA ligase and 1 µl of T4 DNA ligase buffer. Reactions were then mixed gently and incubated at room temperature for at least two hours. The recombinant plasmids were stored in -20^0C until required.

4.2.1.8 Preparation of chemically competent cells of *E. coli*

Four strains of *E. coli* were used in this study. *E. coli* DH5α which is used for plasmid propagation and long term maintenance of plasmid; BL21(DE3) which is a non-pLysS containing host used for higher expression levels of recombinant protein; BL21(DE3)plysS which produces T7 lysozyme to reduce the basal expression of a heterologous protein and Rosetta(DE3) which is used to enhance the expression of proteins that contain codons rarely used in *E. coli*. Chemically competent cells were prepared by picking a colony of each strain from LB plates and growing them overnight in 5 ml of LB broth. About 100 µl of cultures were

then transferred into 100 ml LB broth in 1 L shake flasks and cultures were incubated for four hours to reach an absorbance of 0.4 to 0.5 at 600 nm. Cells were then pelleted by centrifugation at 4000 rpm for 10 minutes at 4^0C. Supernatants were discarded and cells were resuspended in 30 ml ice-cold buffer containing 80 mM magnesium chloride and 20 mM calcium chloride. Cells were then pelleted through centrifugation and resuspended again in 100 mM calcium chloride solution (2 ml per original 50 ml culture). 2 ml of 50% glycerol solution was added to the cell suspension and aliquoted into 50 µl volumes before storing at -80^0C for future transformations.

4.2.1.9 Transformation of recombinant DNA into *E. coli* DH5α

All recombinant plasmids were transformed into *E. coli* DH5α through heat shock transformation. Briefly, 15 µl of recombinant vectors were added into 50 µl of *E. coli* DH5α competent cells and the mixture was incubated on ice for 30 minutes. The cells were then subjected to heat shock at 42^0C for 90 seconds and placed immediately on ice. SOC medium (0.9 ml) was added into each vial and the vials were shaken (220 rpm) at 37^0C for one hour. Aliquots were then spread onto selective agar plates containing kanamycin at a concentration of 50 µg/ml and incubated overnight at 37^0C.

4.2.1.10 Colony Polymerase Chain Reaction (PCR)

A few colonies from agar plates were selected to determine the presence or absence of inserts in plasmid constructs. Half of a single colony was picked and suspended into 9 µl of nuclease-free water in PCR tubes. A 10 µl of *OneTaq* Quick-Load 2X Master Mix with GC Buffer (New England biolabs) was added into each reaction mix along with 0.2 mM primers (Table 4.1) and amplifications were carried out. The initial denaturation was 94^0C for 8 minutes followed by 30 cycles of denaturation at 94^0C for 30 s, annealing at 68^0C for 60 s and extension at 68^0C for 60 s. A final extension at 68^0C for 5 minutes was included to

ensure the complete synthesis. PCR products were electrophoresed on 1% agarose gel and imaged using a gel documentation system.

4.2.1.11 DNA Sequencing

Colonies were picked and cultivated in 5 ml LB broth containing appropriate antibiotics overnight. Cells were then pelleted by centrifugation at 4000 rpm for 15 minutes and treated with QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer's protocol. About 50 to 100 ng of the resulting recombinant plasmid DNA was sent to Eurofins (Germany) for DNA sequencing.

Table 4.1: List of oligonucleotides used for *Ah* isolation from *R.* UKMP-5M, *Ah* cloning into pET-28a(+) and pGEX-6P-1 vectors, DNA sequencing and for colony PC*R.* Restriction enzyme (RE) sites are underlined. All nucleotides were purchased from Invitrogen.

Primers	Sequence 5' to 3'	RE sites	Notes and purpose
*Ah*_fw	ATGTACGAGAAGGACGGCCAGCAGT	-	Forward primer to amplify *Ah* gene from *R.* UKMP-5M
*Ah*_rv	CGACGACACCGCCTCTCGA	-	Reverse primer to amplify *Ah* gene from *R.* UKMP-5M
pET28a_AHntcfw	ATATATATGAATTCATGTACGAGAAGGACGGCCAGCAGTAC	*Eco*RI	Clone full length *Ah* into pET-28a(+) with His-tag only at N-terminus
pET28a_AHntcrv	ATATATATAAGCTTTCACGACGACACCGCCTCTCGAG	*Hind*III	Clone full length *Ah* into pET-28a(+) with His-tag only at N-terminus
pET28a_AHctcfw	ATATATATCCATGGATGTACGAGAAGGACGGCCAGCAGTACTTCAT CGTGGACGGCCACGTGCACAT	*Nco*I	Clone full length *Ah* into pET-28a(+) with His-tag only at C-terminus
pET28a_AHctcrv	ATATATACCATGGTCAATGATGATGATGATGATGCGACGACACCGC CTCTCGAGCTCCCGCC	*Nco*I	Clone full length *Ah* into pET-28a(+) with His- tag only at C-terminus
pET28a_AHnatfw	ATATATACCATGGATGTACGAGAAGGACGGCCAGCAGT	*Nco*I	Clone full length *Ah* into pET-28a(+) without fusion tags
pET28a_AHnatrv	ATATATACCATGGTCACGACGACACCGCCTCTCGA	*Nco*I	Clone full length *Ah* into pET-28a(+) without fusion tags
pET28a_AHcolfw	TATGCGACTCCTGCATTAGG	-	Primers for colony PCR of pET-28a_AH
pET28a_AHcolrv	AACCCGTTGTGGTAGAACTC	-	Primers for colony PCR of pET-28a_AH
pET28aforseq	ATGCGTCCGGCGTAGA	-	Sequence pET-28a_AH
pET28arevseq	CAAGGGGTTATGCTAGTTATTGC	-	Sequence pET-28a_AH
pGEX6P1_AHfw	ATATATGGATCCATGTACGAGAAGGACGGCCAGCAGTAC	*Bam*HI	Clone full length *Ah* into pGEX-6P-1 with GST tag at only at N-terminus
pGEX6P1_AHrv	ATATATCTCGAGCGACGACACCGCCTCTCGA	*Xho*I	Clone full length *Ah* into pGEX-6P-1 with GST tag at only at N-terminus
pGEX6P1forseq	ATAGCATGGCCTTTGCAGG	-	Sequence pGEX-6P-1_AH and primers for colony PCR
pGEX6P1revseq	GAGCTGCATGTGTCAGAGG	-	Sequence pGEX-6P-1_AH and primers for colony PCR

fw = forward sequence and rv = reverse sequence

4.2.2 Small scale recombinant protein expression

4.2.2.1 Growth and induction

Recombinant plasmid was transformed into either *E. coli* BL21(DE3), *E. coli* BL21(DE3)pLysS or Rosetta(DE3) as described in section 4.2.1.9 for heterologous protein expression. Overnight cultures of *E. coli* BL21(DE3) carrying the recombinant DNA plasmid were inoculated into 5 ml LB with kanamycin at 50 µg/ml. Cultures of *E. coli* BL21(DE3)pLysS or Rosetta(DE3) were inoculated into 5ml LB with 34 µg/ml of chloramphenicol. Cultures were incubated at either 37^0C or 18^0C in a shaking incubator (220 rpm) until the absorbance at 600 nm was between 0.5 and 0.6. Expression of the cloned genes was chemically induced by using Isopropyl- β-D-1-thiogalactopyranoside (IPTG) at various concentrations ranging from 0 to 1 mM final concentration. The culture was incubated for a further three to four hours and the bacterial cells were pelleted by centrifugation at 4000 rpm for 15 min.

4.2.2.2 Soluble protein extraction

Soluble protein from *E. coli* was extracted using Bugbuster mastermix (Novagen) based on manufacturer's protocol with slight modifications. Briefly, the cell pellet was resuspended in 250 µl bugbuster solution. The suspensions were then transferred into 1.5 ml eppendorf tubes and incubated at room temperature for 30 minutes on a rotary shaker. Suspensions were then centrifuged at 16,000 rpm for 15 minutes at 4^0C to remove debris which was kept for analysis of insoluble proteins. Supernatant was then transferred to a fresh tube and stored at -20^0C for PAGE analysis.

4.2.2.3 Analysis of insoluble protein

The protein pellets from section 4.2.2.2 were resuspended in 50 µl lysis buffer containing 65 mM Tris, 2% (w/v) SDS, 1 mM β-mercaptoethanol, 4.5% (v/v) glycerol and bromophenol blue. Suspensions were mixed thoroughly and heated at 80^0C for 10 minutes on a 'Dri block'. Cell debris was removed through centrifugation and the supernatant was stored at -20^0C for PAGE analysis.

4.2.2.4 Polyacrylamide gel electrophoresis and Coomassie staining

Soluble proteins prepared in section 4.2.2.2 were mixed with 25% (v/v) NuPAGE LDS Sample Buffer (4X) (Invitrogen) and 1% NuPAGE Reducing Agent (10X) (Invitrogen). Samples were heated at 70^0C for 10 minutes prior to loading into the 4-12% SDS PAGE pre-cast gradient gels. Protein samples from section 4.2.2.3 were loaded directly into the gels. Electrophoresis was performed at 150 V for 40 minutes. The gels were stained with Coomassie instant blue as described in section 3.2.8.

4.2.2.5 Western blotting

For western blots, samples were first electrophoresed on 4-12% SDS-PAGE gradient gel without staining. Following this, proteins were electroblotted using a mini Trans-blot cell (Bio-Rad) onto a nitrocellulose membrane (GE Healthcare). A wet transfer was carried out at 4^0C with a constant voltage of 30 V for overnight in a transfer buffer containing 40 mM glycine, 50 mM Tris, 1 mM SDS and 20% (v/v) methanol. The membrane was blocked with 5% skimmed milk in TBST buffer (50 mM Tris, 150 mM NaCl, 0.1% Tween 20, pH 7.5) at room temperature for one hour. The blocking solution was discarded and the nitrocellulose membrane was incubated with anti-6X his tag antibody (Abcam) in skimmed milk solution (1:1000 dilution) on a rocking platform at room temperature for another one hour. After incubation with the primary antibody, nitrocellulose membrane was washed three times with TBST for 10 min and incubated with horseradish peroxidase conjugated antimouse (Bio-

Rad) in skimmed milk (1:2000 dilution) for an additional one hour before washing with TBST three times for 10 minutes. Finally, blots were developed by enhanced chemiluminescene (ECL) (Amersham Biosciences) according to the manufacturer's instructions. Developing reagents were mixed at a ratio of 1:1 and the solution was allowed to cover the membrane. After 5 minutes of incubation, developing reagent was discarded and the nitrocellulose membrane was wrapped in a cling film and visualised using an imaging system (UVP).

4.3 Results

4.3.1 Amplification of *Ah* from *R.* UKMP-5M using Q5 DNA polymerase

Q5 DNA polymerases are thermostable DNA polymerases, which have 3' → 5' exonuclease proofreading activity, resulting in an accurate DNA amplification. The error rate for these polymerases is more than 100-fold lower than that of *Taq* DNA polymerases and twelve-fold lower than the *Pfu* DNA polymerases and thus are ideal for cloning and can be used for long or difficult amplicons. Therefore, *Ah* gene from *R.* UKMP-5M was isolated using Q5 High-Fidelity DNA polymerase using *Ah*_fw and *Ah*_rv primers (Table 4.1).

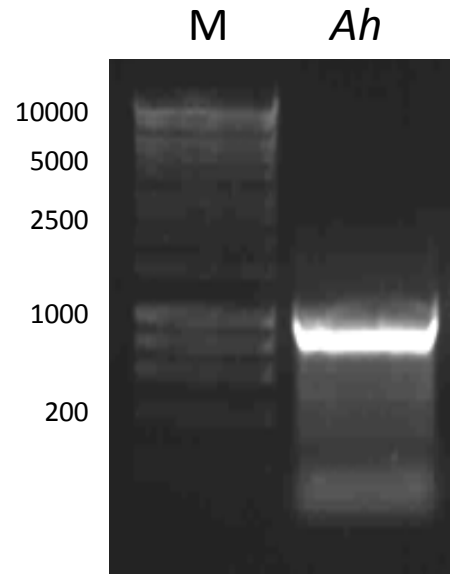


Figure 4.1: Agarose gel of *Ah* gene (1041 bp) isolated from *R.* UKMP-5M. PCR was done using Q5 High-Fidelity DNA polymerase. Primers used to amplify the gene were *Ah*_fw and *Ah*_rv. M represents DNA marker and Ah represents amidohydrolase.

4.3.2 Analytical digestion of pET-28a+

One of the most common and powerful systems developed for cloning and expression of recombinant proteins in *E. coli* is the pET system. This system was originally developed by Studier and colleagues in 1986. The transcription of the target genes cloned in pET vectors are controlled by a strong T7 promoter which is highly specific for bacteriophage T7 RNA polymerases which is chromosomally encoded in *E. coli* BL21(DE3) (Studier and Moffatt 1986).

A number of studies have shown successful overexpression of genes from *Rhodococcus* in *E. coli* using the pET systems. For example, flavin reductase from *R. erythropolis* D-1 which was one of the four genes responsible for the degradation of dibenzothiophene was successfully cloned and overexpressed in *E. coli* using pET-21a (Matsubara, Ohshiro et al. 2001). In another study, pET-21a was also used to clone a gene for an Ah-type enzyme known as barbiturase from *Rhodococcus erythropolis* JCM 3132 into *E. coli*. The heterologous protein was produced in a soluble form and the specific activity towards substrate was 10-fold higher than that of the *R. erythropolis* JCM 3132 (Soong, Ogawa et al. 2002).

In this study, cloning of *Ah* from *R.* UKMP-5M in *E. coli* was done in the pET-28a(+) system (Figure 4.2). This system is driven by the T7 promoter and the recombinant protein carries a 6X histidine fusion tag (His-tag) at the N-terminus or optionally at the C-terminus. The presence of his-tag in the expressed protein aids the purification process through immobilised metal affinity chromatography (IMAC) based on its interaction with divalent metal ions like nickel, cobalt and zinc (Terpe 2003). Analytical digestion was performed in order to confirm that the plasmid is intact. Two single and one double digestion were performed to ensure the resulting DNA fragments are corresponding to their theoretical lengths. Result of the analytical digestion of pET-28a(+) is shown in Figure 4.3.

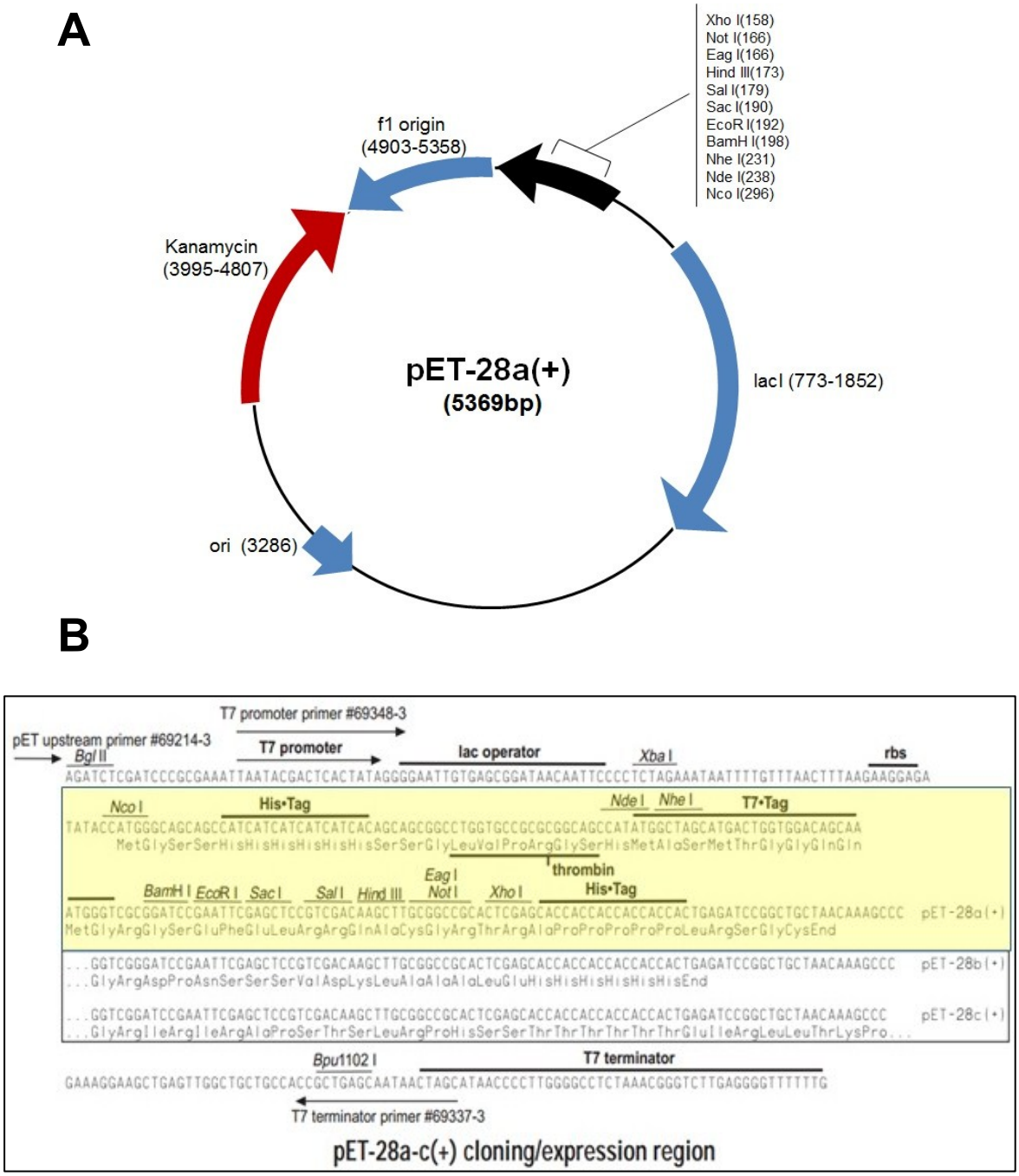


Figure 4.2: (A) The 5369 bp pET-28a(+) vector system containing multiple cloning sites and other components; (B) Part of the nucleotide sequence map for pET-28a, b and c(+). The maps for pET-28b(+) and pET-28c(+) are the same as pET-28a(+) with the following exceptions: pET-28b(+) is a 5368 bp plasmid; subtract 1 bp from each site beyond *Bam*HI at 198. The pET-28c (+) is a 5367 bp plasmid; subtract 2 bp from each site beyond *Bam*HI at 198. The pET-28a(+) region containing restriction sites and His-tags used in this study is given in the yellow box (Novagen 2002).

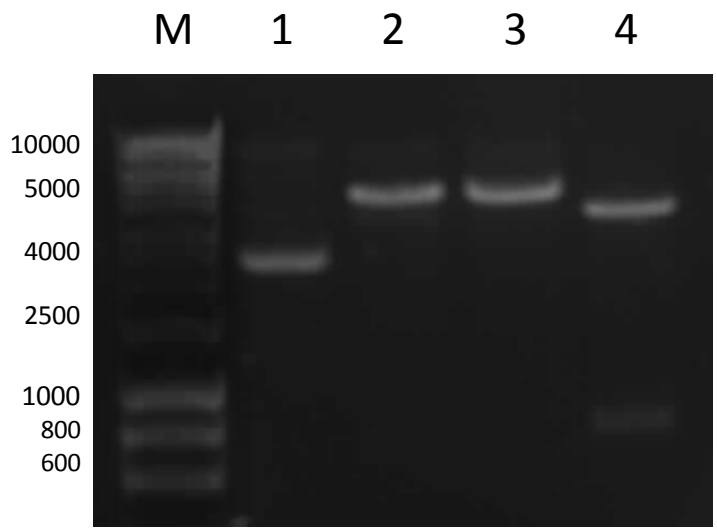


Figure 4.3: Analytical digestion of pET-28a(+). Agarose gel electrophoresis of undigested circular pET-28a(+) (lane 1). The pET-28a(+) underwent single cleavage with either *Hind*III (lane 2) or *Eco*RI (lane 3) which yields a linearized DNA vector with 5369 bp. The pET-28a(+) was then subjected to double digestion with *Eco*RI and *Mlu*I, resulting in two DNA fragments at 4438 bp and 931 bp (lane 4). M represents DNA ladder.

4.3.3 Preparation of inserts for cloning of *Ah* into pET-28a(+) with histidine tags

In comparison to other fusion tags like gluthathione-S-transferase (GST) and maltose binding proteins (MBP), His-tags are relatively small in size, which rarely affects the folding, and solubility of a target protein. Furthermore, characterisation of recombinant proteins via protein crystallography and NMR structures has shown that the presence or absence of a His-tag does not interfere with the conformation of a protein, resulting in similar three-dimensional structures (Graslund, Nordlund et al. 2008).

Many studies have shown that the N-terminus His-tag is preferred over the C-terminus for a number of reasons. An N-terminal fusion tag ensures the gene transcription and translation in bacteria always encounter the 5' region, and thus the N-terminus sequence is often compatible with RNA synthesis and protein expression. Besides, given that the N-terminal His-tag can be 'chopped off' using specific proteases, the structure and

functional aspect of the target protein are not significantly altered (Graslund, Nordlund et al. 2008). However, the N-terminal fusion proteins do not always yield soluble protein. At times, a protein may require fusion to other affinity tags to be expressed in their soluble form (Hammarstrom, Hellgren et al. 2002) and in some cases, the C-terminal his-tagged protein was found to have improved soluble expression compared to the N-terminal fused proteins (Woestenenk, Hammarstrom et al. 2004).

Since the optimal placement of a fusion-tag is highly specific to the protein of interest, a pragmatic approach was opted for by carrying out the cloning with three different *Ah* inserts containing restriction sites that (i) introduce a N-terminal his-tag, (ii) introduce a C-terminal his-tag and (iii) have no his-tag introduced due to the introduction of a stop codon before the C-terminal his-tag sequence. The corresponding inserts with their respective restriction sites were extracted and purified from the agarose gel and ligation was then performed with the digested vector to construct the corresponding recombinant plasmids as shown in Figure 4.5.

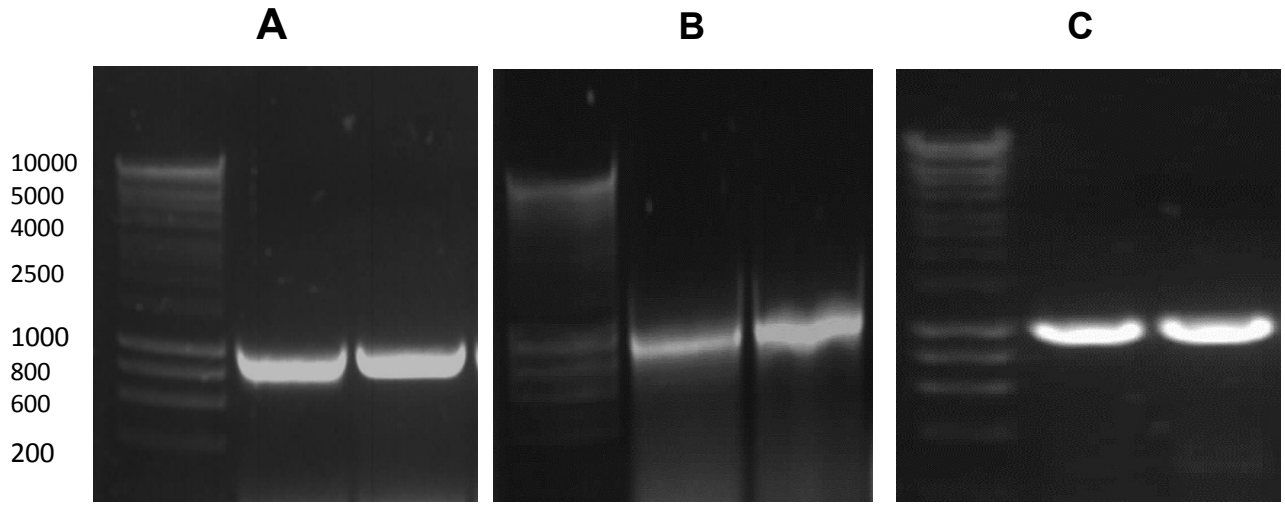


Figure 4.4: Agarose gel of *Ah* gene containing (A) restriction sites for *Eco*RI (5') and *Hind*III (3') to yield a heterologous *Ah* expression with His-fusion tag at the N-terminus; (B) restriction sites for *Nco*I (5' and 3') for *Ah* expression with His-fusion tag at the C-terminus and; (C) restriction sites for *Nco*I (5' and 3') to yield *Ah* expression in the absence of fusion tag. Amplifications were done using Q5 High-Fidelity DNA polymerase. Primers used to prepare the inserts are given in Table 4.1.

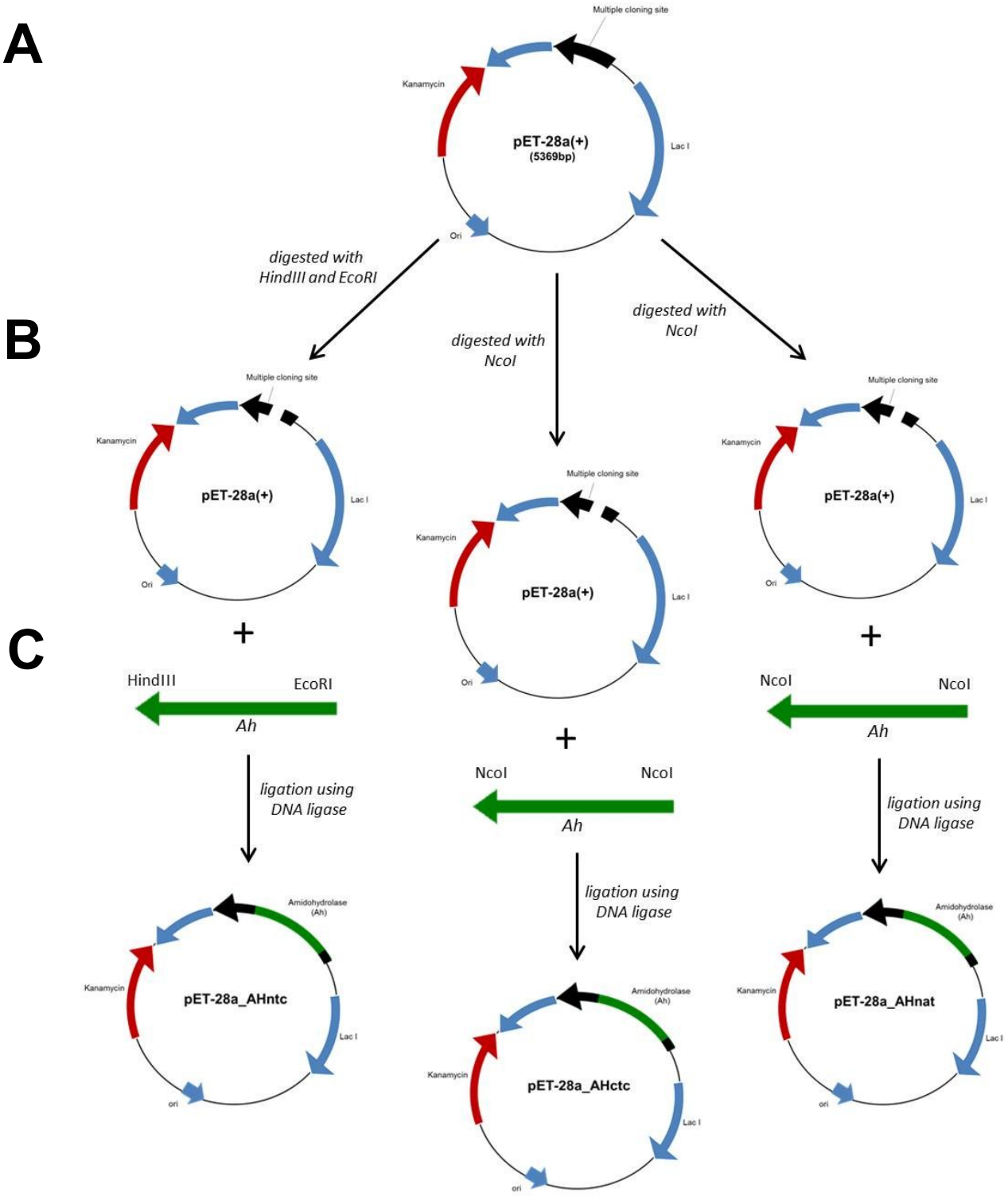


Figure 4.5: A schematic representation of amidohydrolase (*Ah*) cloning into pET-28a(+) vector. (A) Intact pET-28a(+) underwent digestion with appropriate restriction enzyme(s) (B). (C) The digested vector and *Ah* insert containing appropriate restriction sites were subjected to ligation to form the recombinant plasmid, pET-28a_AHntc (His-tag at the N-terminus of Ah), pET-28a_AHctc (His-tag at the C-terminus of Ah) and pET-28a_AHnat (Ah without His-tag).

4.3.4 Colony PCR

The recombinant plasmids shown in Figure 4.5 were introduced into *E. coli* DH5α through heat shock transformation. Colony PCR was carried out using a vector specific primer at the 5' and an insert specific primer at the 3' end to determine the presence or absence of insert DNA in the plasmid constructs and also to confirm the correct orientation of the inserts.

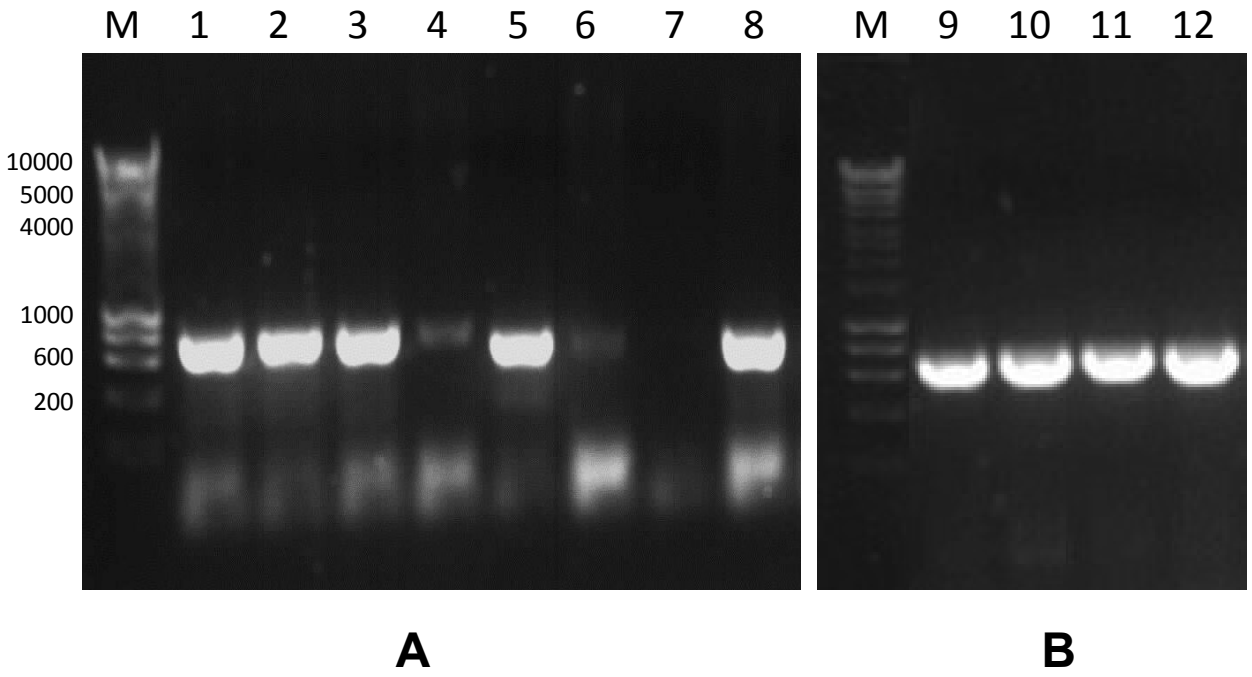


Figure 4.6: Agarose gel of colony PCR after transformation of pET-28a_AH into *E. coli* DH5α. (A) Positive clones for Ah expression with His-fusion tag at the N-terminus (lane 1-3) and at the C-terminus (lane 5 and 8). (B) Positive clones for Ah expression in the absence of fusion tag (lane 9-12) Amplifications were done using *OneTaq* Quick-Load 2X Master Mix with GC Buffer (Qiagen). Primers used were pET-28a_AHcolfw and pET-28a_AHcolrv and their sequences are given in Table 4.1.

Positives clones for each type of construct were identified and DNA minipreps were done to isolate and purify the recombinant plasmid DNA for each construct. The resulting DNA was sent to Eurofins (Germany) for DNA sequencing. The primers used for DNA sequencing were pET28aforseq and pET28arevseq and their sequences are given in Table 4.1. The DNA sequencing results are provided in the Appendix IV, V and VI.

4.3.5 Protein analysis

Plasmid DNA from clones which were confirmed to contain the correct insert by DNA sequencing were transformed into *E. coli* BL21(DE3). A small-scale test extraction was done to determine which of the positive clones produces soluble protein and to establish optimal conditions for efficient expression of the target protein. One problem with heterologous over-expression is the formation of inclusion bodies and a well-known technique for the expression of soluble recombinant proteins is by reducing the growth temperature for the bacteria. Therefore, cells harbouring the recombinant plasmids were grown at either 18^0C or 37^0C and induced with different concentrations of IPTG. The recombinant Ah is expected to have a molecular weight approximately 45 kDa.

A

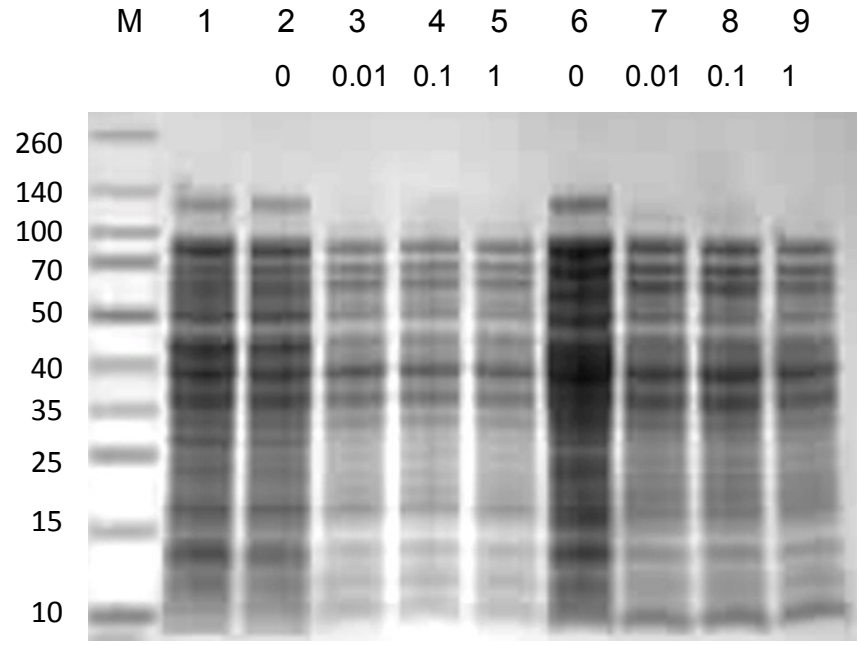
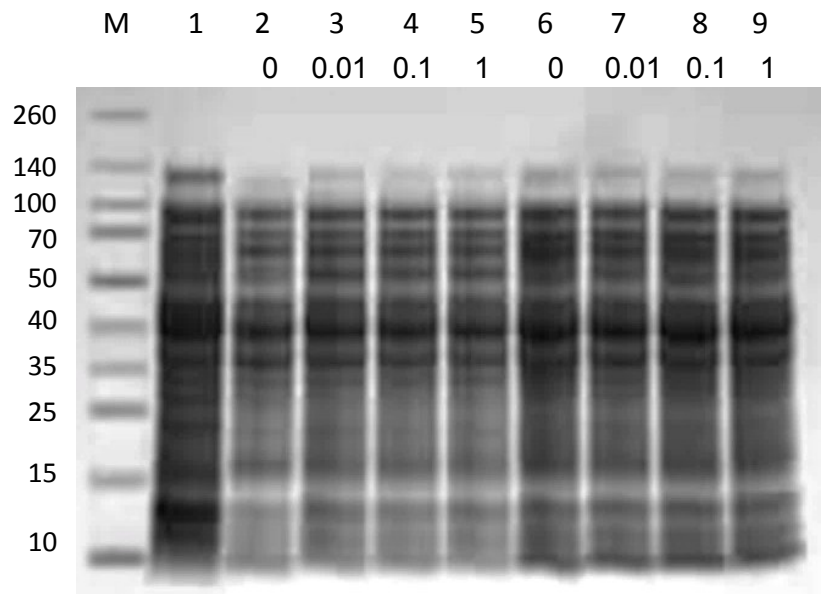


Figure 4.7A: 1D GE of soluble protein extracted using bugbuster master mix from *E. coli* BL21(DE3) harbouring the native pET-28a(+) (lane 1) and pET-28a_AHntc (lanes 2-9).

Lanes 2-5 contains protein from cells grown at 37^0C whilst lanes 6-9 contain protein from cells grown at 18^0C. The numbers below the lane number are the concentration of IPTG (in mM) used to induce the expression. M represents protein standards. Samples were separated on 4-12% SDS-PAGE gradient gel. Protein detection was with Coomassie instant blue.

B



C

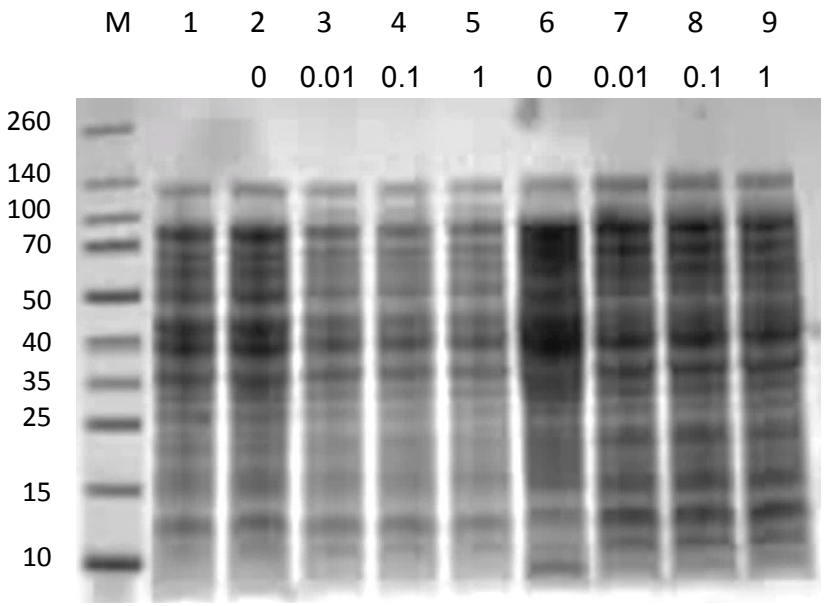


Figure 4.7B: 1D GE of soluble protein extracted using bugbuster master mix from *E. coli* BL21(DE3) harbouring the native pET-28a (+) (lane 1) and pET-28a_AHctc (lanes 2-9).

Figure 4.7C: 1D GE of soluble protein extracted using bugbuster master mix from *E. coli* BL21(DE3) harbouring the native pET-28a (+) (lane 1) and pET-28a_AHnat (lanes 2-9).

Lanes 2-5 contains protein from cells grown at 37^0C whilst lanes 6-9 contain protein from cells grown at 18^0C. The numbers below the lane number are the concentration of IPTG (in mM) used to induce the expression. M represents protein standards. Samples were separated on 4-12% SDS-PAGE gradient gel. Protein detection was with Coomassie instant blue.

D

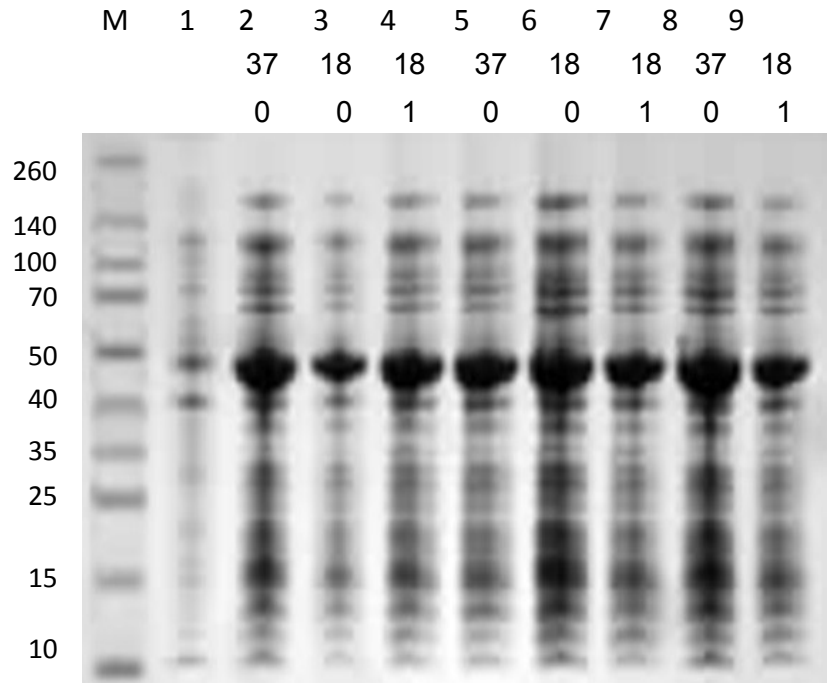


Figure 4.7D: 1D GE of insoluble protein extracted using SDS lysis buffer from pellet of *E. coli* BL21(DE3) harbouring the native pET-28a(+) (lane 1); pET-28a_AHntc (lanes 2-4); pET-28a_AHctc (lanes 5-7) and pET-28a_AHnat (lanes 8 and 9).

The numbers below the lane number are the temperature used to grow the cells containing recombinant plasmids followed by the concentration of IPTG (in mM) used to induce the expression. M represents protein standards. Samples were separated on 4-12% SDS-PAGE gradient gel. Protein detection was with Coomassie instant blue.

The expression of the heterologous Ah in *E. coli* was successful with all three plasmid constructs. This indicates that the presence of a His-fusion tag on the target protein and its placement at either the N-terminus or C-terminus does not influence the expression of the protein which agrees with a previous study (Lundstrom 2006). However, the protein was not present in the soluble fraction (Figure 4.7A, B and C), but only expressed as insoluble protein or inclusion bodies (IB) (Figure 4.7D). This was the case even when no his-tag was present. Although many studies have shown that lowering the temperature during the expression of recombinant *E. coli* promotes the production of soluble protein (Vasina and Baneyx 1997, San-Miguel, Perez-Bermudez et al. 2013), in this case this approach did not overcome the problem.

4.3.6 Western blotting of 6X His-tagged Ah expression

Since a commercial antibody for Ah is not available to further validate the results obtained in the small-scale protein expression analysis, western blotting was carried out to confirm the presence of the His-tag using an anti-6X His-tag antibody. Soluble and insoluble protein of *E. coli* BL21(DE3) containing the recombinant plasmids and also the native vector, pET-28a(+) was separated on an SDS-PAGE gel and transferred to a nitrocellulose membrane, followed by immunoblotting.

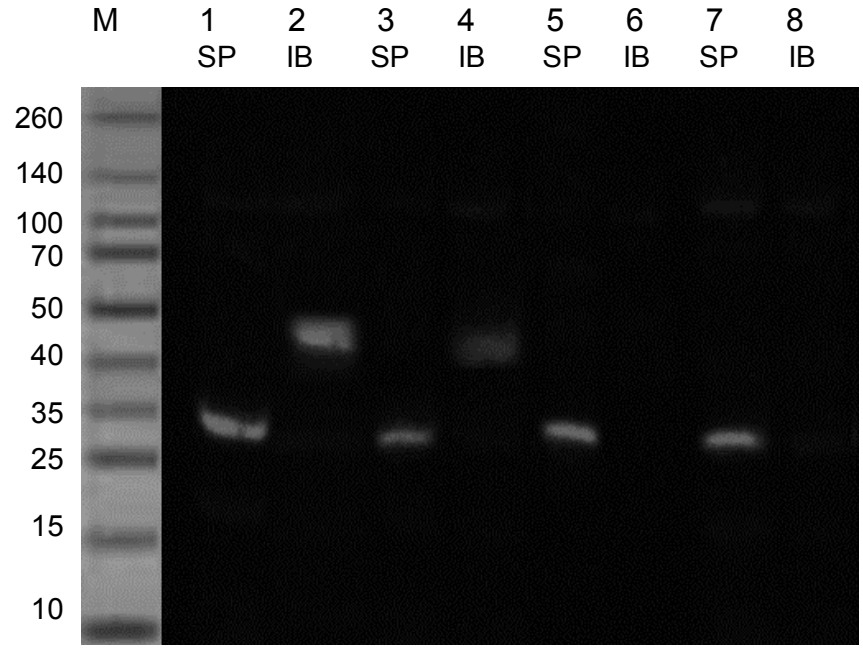


Figure 4.8: Immunoblotting with anti-6X His-tag antibody. Protein samples were extracted from *E. coli* BL21(DE3) harbouring pET-28a_AHntc (lanes 1 and 2); pET-28a_AHctc (lanes 3 and 4); pET-28a_AHnat (lanes 5 and 6) and the native pET-28a (+) (lanes 7 and 8). All protein samples were extracted from cells grown at 37^0C with the addition of 1 mM IPTG to induce the expression. Protein samples were separated on 4-12% SDS-PAGE gradient gel and transferred onto a nitrocellulose membrane. M = protein marker, SP = soluble protein and IB = inclusion bodies.

Figure 4.8 confirms that there is no expression of soluble Ah by the recombinant constructs of pET-28a_AHntc and pET-28a_AHctc and the protein aggregates to form inclusion bodies with the expected molecular weight as shown by the signal at 45 kDa. Expression of Ah by

cells transformed with pET-28a_AHnat will not be detected through the western blot since the protein does not possess a His-fusion tag to be identified by the polyhistidine antibody. One disadvantage of using anti polyhistidine tag antibodies is their lack of specificity. As shown in Figure 4.8, there is a signal at 35 kDa in all the soluble protein samples, including protein extracted from cells transformed with the native vector (vector without Ah insert). This could be due to interaction of the antibody with an endogenous protein rich in histidine residues when eluting the protein with high concentrations of imidazole during IMAC purification (Vashist 2013).

4.3.7 Improving the solubility of Ah

Since the cloning of Ah with all three different inserts resulted in the formation of insoluble protein (IB), several attempts were made to obtain a soluble form of Ah. Some of the approaches are discussed in this section.

4.3.7.1 Addition of 1% glucose in growth medium

Many studies have shown that altering the growth medium with the addition of glucose prevents 'leaky' expression of recombinant protein by repressing the induction of T7 promoters. As a result, the transcription and synthesis of recombinant protein occurs at a slower rate, and may allow the protein to fold in its correct configuration (Morris 2001). Hence, an investigation was done by growing the cells containing the recombinant plasmid in LB supplemented with 1% glucose. The cells were allowed to grow until the cell density was between 0.4 and 0.6, and was induced with IPTG before further incubation at two different temperatures. Unfortunately, the expressed protein still formed IB. The result of this investigation is shown in Figure 4.9.

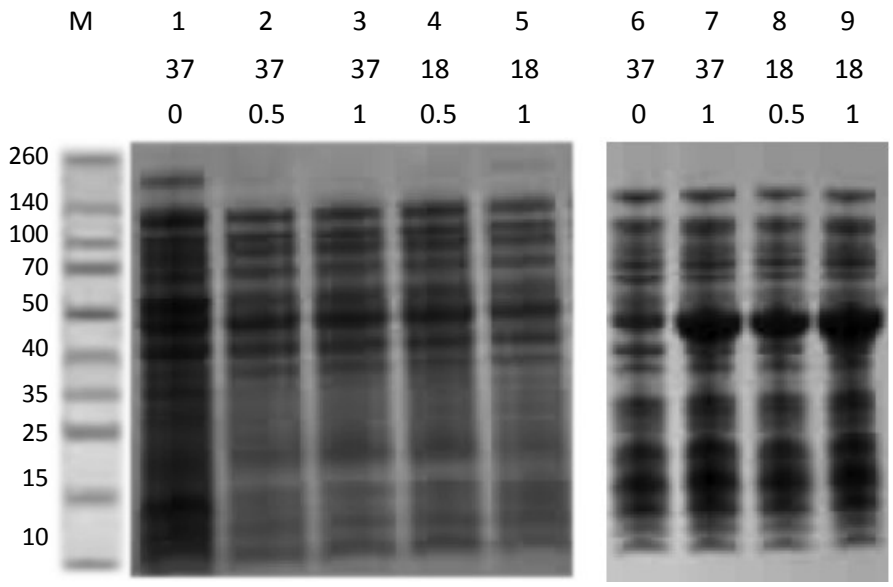


Figure 4.9: 1D GE of soluble protein (lanes 1-5) and IB (lanes 6-9) extracted from *E. coli* BL21(DE3) harbouring the recombinant pET-28a_AHntc that was grown in medium supplemented with 1% glucose.

The numbers below the lane number are the temperature used to grow the cells containing recombinant plasmids followed by the concentration of IPTG (in mM) used to induce the expression. M represents protein standards. Samples were separated on 4-12% SDS-PAGE gradient gel. Protein detection was with Coomassie instant blue.

4.3.7.2 Transformation of recombinant plasmid into different host strains

The recombinant pET-28a_Ahntc was transformed into two different host cells; *E. coli* BL21(DE3)pLysS and Rosetta(DE3). Both the strains are derivatives of BL21(DE3), with the former designed to contain a pLysS plasmid that produces T7 lysozyme to reduce basal level expression of a foreign protein and the latter, to express protein (mostly eukaryotic) that contains codons rarely used in *E. coli*. Both the strains confer resistance towards chloramphenicol (Novagen 2004). Protein analyses of these transformed cells are shown in Figure 4.10. The recombinant Ah produced in these cells was also forming IB.

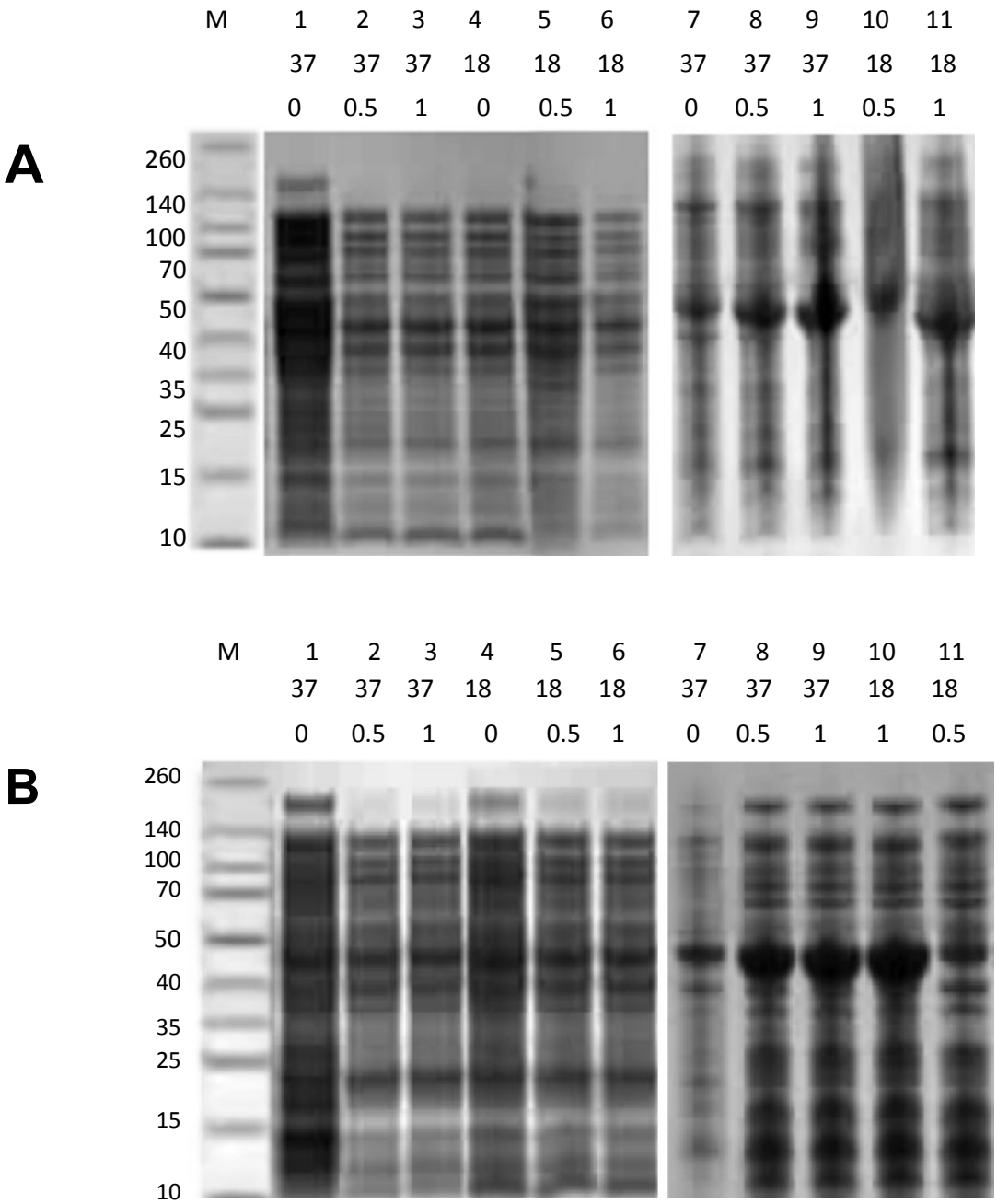


Figure 4.10: 1D GE of soluble protein (lanes 1-6) and IB (lanes 7-11) extracted from (A) *E. coli* BL21(DE3)pLysS and (B) Rosetta(DE3) harbouring the recombinant pET-28a_Ahntc.

The numbers below the lane number are the temperature used to grow the cells containing recombinant plasmids followed by the concentration of IPTG (in mM) used to induce the expression. M represents protein standards. Samples were separated on 4-12% SDS-PAGE gradient gel. Protein detection was with Coomassie instant blue.

4.3.7.3 Cloning of Ah into pGEX-6P-1 plasmid

A pGEX-6P-1 plasmid contains a gene for 26 kDa Gluthathione S-transferase (GST) at the N-terminus of the recombinant protein that facilitates the detection of GST-tagged protein. Protein expression with this system is controlled by the *tac* promoter, which is inducible with IPTG. This vector is also designed to contain a *lac* I gene, whose product is a repressor protein that binds to the operator region of the *tac* promoter to prevent expression until the induction by IPTG. Hence, it provides a tight control over the expression of the recombinant gene. All pGEX vectors confer resistance to ampicillin (Healthcare 2002)

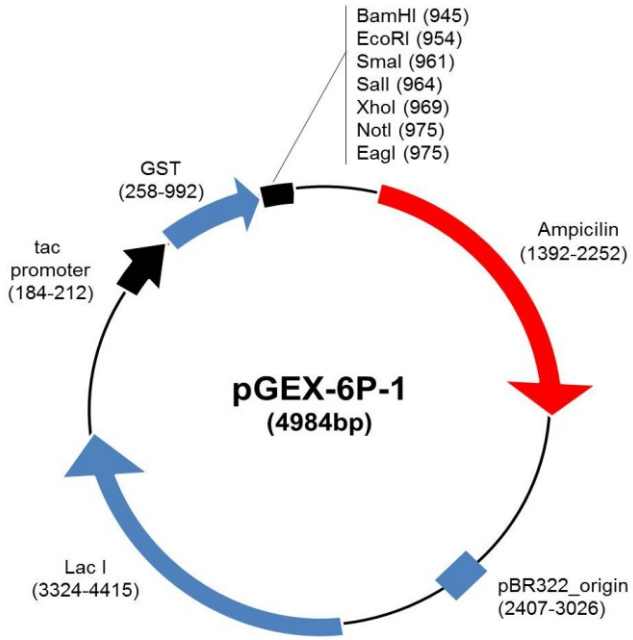


Figure 4.11: Vector map of the 4984 bp pGEX-6P-1.

Adapted from (Healthcare 2002).

Inserts of Ah containing appropriate restriction sites was prepared and cloned into pGEX-6P-1. The in-frame cloning was confirmed through DNA sequencing and the result is provided in Appendix VII. The recombinant plasmid was then transformed into *E. coli* BL21(DE3) for protein expression analysis and the results are provided in Figure 4.12. The recombinant Ah with the GST-fusion tag is expected to have a molecular weight of approximately 68 kDa.

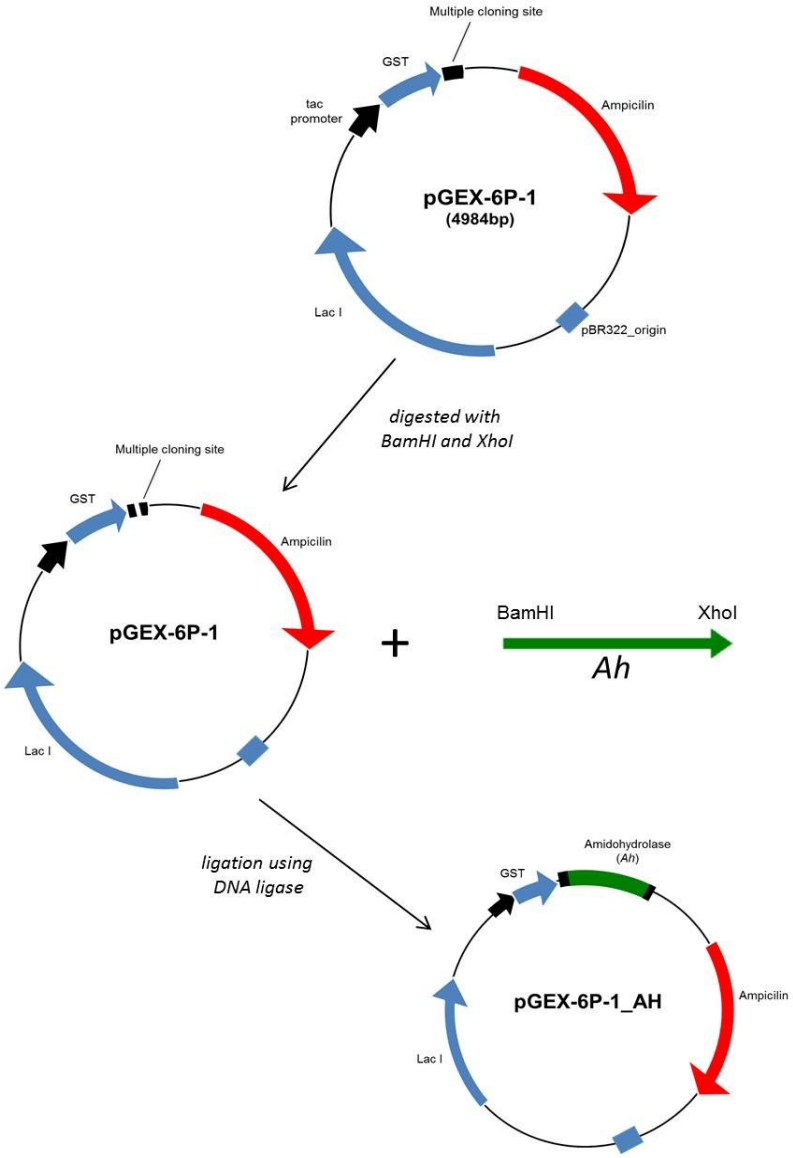


Figure 4.12: A schematic representation of Ah cloning into pGEX-6P-1 vector. Intact pGEX-6P-1 underwent digestion with *Bam*HI and *Xho*I. The digested vector and Ah insert containing restrictions sites were subjected to ligation to form recombinant plasmid, pGEX-6P-1_AH.

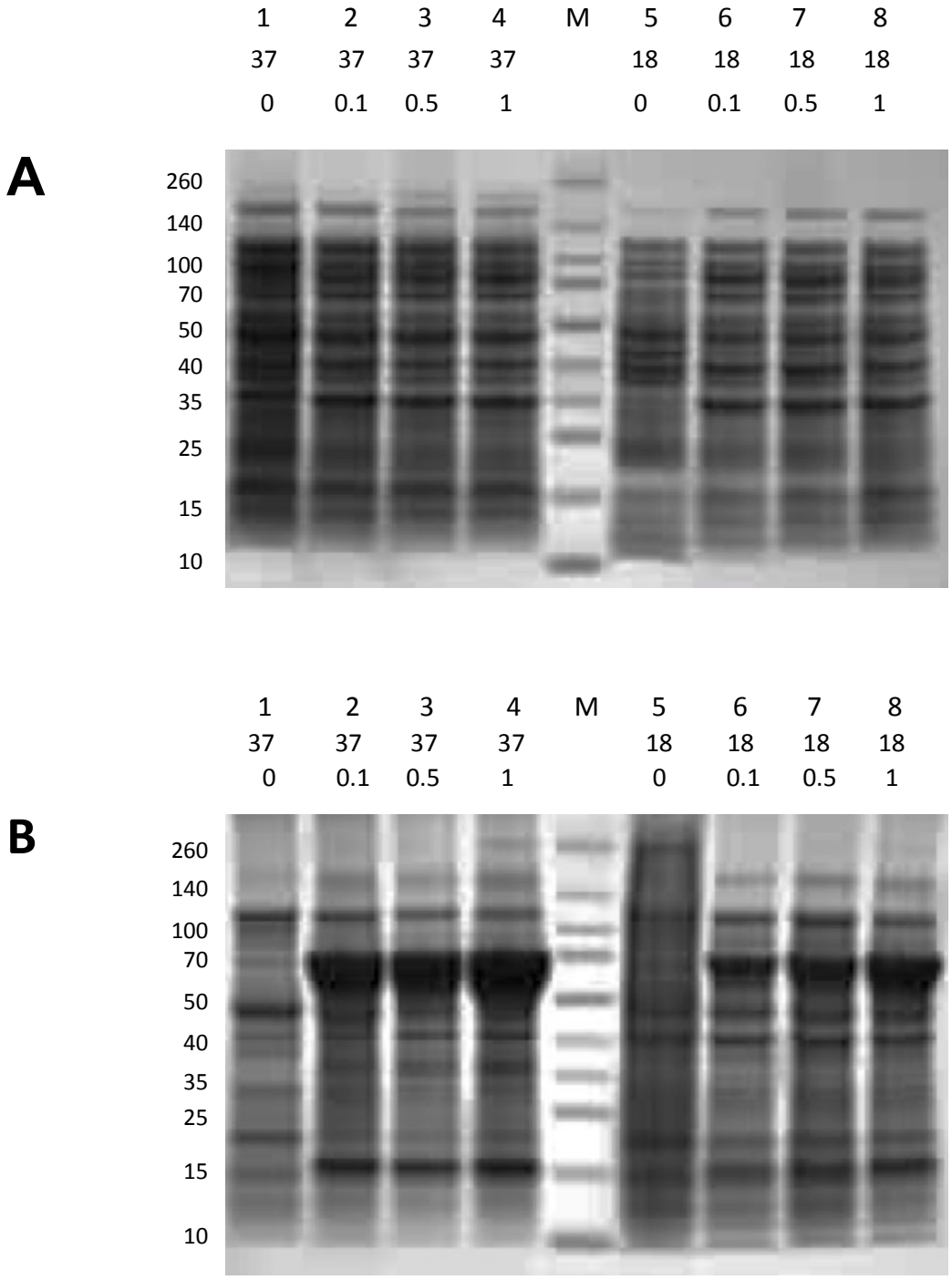


Figure 4.13: 1D GE of (A) soluble protein (lanes 1-6) and (B) IB (lanes 7-11) extracted from *E. coli* BL21(DE3) harbouring the recombinant pGEX-6P-1_AH.

The numbers below the lane number are the temperature used to grow the cells containing recombinant plasmids followed by the concentration of IPTG (in mM) used to induce the expression. M represents protein standards. Samples were separated on 4-12% SDS-PAGE gradient gel. Protein detection was with Coomassie instant blue.

4.4 Discussion

A 'hypothetical' protein was found to be induced in response to 1-CB exposure in *R.* UKMP-5M. Further sequence analysis of this protein against the NCBI database revealed that it has a sequence homology with amidohydrolases (Ah) in *Rhodococcus*. The *Ah* gene was amplified from *R.* UKMP-5M mRNA and successfully cloned into *E. coli*. Heterologous expression of Ah in a soluble form in *E. coli* was not achieved; instead the protein forms inclusion bodies (IB). IBs are insoluble protein aggregates that often occur during high-level expression of a heterologous protein. Recombinant protein synthesis in *E. coli* especially when driven by a strong promoter such as T7, happens at a faster rate which leads to misfolding and accumulation of insoluble proteins in the cells (Singh and Panda 2005) .

In this study, the cloning of the *Ah* gene was done in a pET vector system. Due to the high transcription efficiency of this system, a high basal level of activity known as 'leaky' expression often encountered (Gruber, Pieribone et al. 2008) which accounts for the high expression of Ah even without induction with IPTG (0 mM) as shown in Figure 4.7D. Multiple attempts were made to reduce the level of protein expression and promote the solubility of the heterologous Ah in *E. coli*. One was by altering the growth medium with the addition of 1% glucose. Many studies have shown that addition of glucose to the growth medium significantly tightens the regulation of bacterial expression by repressing the induction of the lac promoters (Matthew J. Giacalone and Carl W. Gunderson). However, this technique did not help protein solubilisation in this study.

Another approach is to use a different host cell for protein expression. A number of host cells are available which have been reported to be efficient in soluble protein expression. In this study, the two types of host strains used in addition to *E. coli* BL21(DE3) are *E. coli* BL21(DE3)pLysS and Rosetta(DE3). The former is often used for high protein expression for any gene under the control of the T7 promoter. It contains a plasmid that carries a gene for T7 lysozyme, which is a natural inhibitor of T7 RNA polymerase. Low

levels of lysozyme expressed by this system are sufficient to stabilise the plasmid, but still results in high level of protein expression upon induction (Studier and Moffatt 1986). Rosetta(DE3) is designed to enhance eukaryotic proteins that contains codons like AGG, AGA, GGA, AUA, CUA, and CCC which are not commonly used by *E. coli*. Thus, Rosetta cells provide universal translation of proteins which are limited by the codon usage in *E. coli*. Given that genes from *R.* UKMP-5M are rich in GC content, an attempt to test the protein expression in this strain was thought worthwhile. However, the recombinant protein whilst still produced again formed inclusion bodies.

A different attempt was then made to obtain a soluble protein expression of Ah by cloning the *Ah* gene into a different plasmid vector. A pGEX-6P-1 was chosen for expression of Ah. A GST-fused protein can be purified using immobilised gluthathione affinity chromatography (Terpe 2003). It was reported that most GST-fused proteins are expressed in the soluble form rather than forming IB (Harper and Speicher 2011), however other studies have shown that the GST-fused protein can still form insoluble aggregates (Tao, Liu et al. 2010, Dae-Wook Park 2011). Removal of the GST-tag can be done with site-specific protease, thus does not interferes with the activity of the protein after purification. However, protein expression from this recombinant DNA was found to still form inclusion bodies, even after varying the growth temperatures and IPTG concentrations.

Since the commonly used methods to express soluble protein were not successful with Ah, a more laborious technique known as IB solubilisation and refolding was performed to try and obtain Ah in its active form. This technique is time consuming and often yields a poor recovery (Singh and Panda 2005), but many studies to-date have shown that it is possible to successfully refold IB proteins to their biologically active states. The IB from Ah was solubilised in denaturants such as guanidine hydrochloride and on-column refolding was carried out using decreasing gradient concentrations of urea. The results of this approach are presented in Chapter 5.

CHAPTER 5 PURIFICATION AND REFOLDING OF RECOMBINANT AMIDOHYDROLASE

5.1 Introduction

Recombinant DNA technology techniques have been playing a vital role in large-scale production of proteins essential for both research and industrial applications. However, the expression of a recombinant protein in a prokaryotic host like *E. coli* often results in the formation of insoluble protein known as inclusion bodies (Basu, Li et al. 2011). Inclusion bodies (IB) are protein aggregates that accumulate in either the cytoplasmic or periplasmic compartments of *E. coli* during expression of a heterologous protein. This form of protein is insoluble and assumed to contain unfolded or highly misfolded polypeptides that did not achieve their native confirmation and thus, making IB biologically inactive (García-Fruitós, Vázquez et al. 2012).

As discussed in Chapter 4, a few common approaches were attempted to express soluble form of recombinant amidohydrolase (Ah) in *E. coli*. The protein translation rate was manipulated by lowering the growth temperature to allow time for protein folding to take place and the growth medium was altered with the addition of 1% glucose to repress the 'leaky' expression of the target protein. In addition, fusion tags such as polyhistidine (His-tag) and gluthathione S-transferase (GST-tag) were also used to try and help the solubilisation of the protein. Different host cells like BL21(DE3)pLysS and Rosetta(DE3) were also tested for soluble expression of Ah. Since none of these efforts were successful in the production of soluble Ah in *E. coli*, refolding of IB was carried out to obtain Ah in its bioactive form.

Briefly, the IB fraction is isolated from the cells and resuspended in buffer containing guanidium hydrochloride or urea to allow complete denaturation and solubilisation of the protein aggregates. Protein refolding can be done either by conventional methods like

dilution and dialysis followed by purification procedures; or by on-column refolding using ion-exchange or immobilised metal affinity (IMAC). Conventionally, refolding is done by reducing the concentration of the denaturant in the IB sample either through dialysis or dilution to a point where protein refolding can occur. Protein is exposed to this condition for an extended period of time to allow refolding to complete and any insoluble aggregate material to flocculate. Following this, the refolded material is then purified by removing impurities and concentrated through chromatographic techniques like ion-exchange chromatography before functional analyses (Burgess 2009). In both cases the, the main challenge is to achieve the correct native conformation despite the misfolding and aggregation that may possibly occur as shown in Figure 5.1.

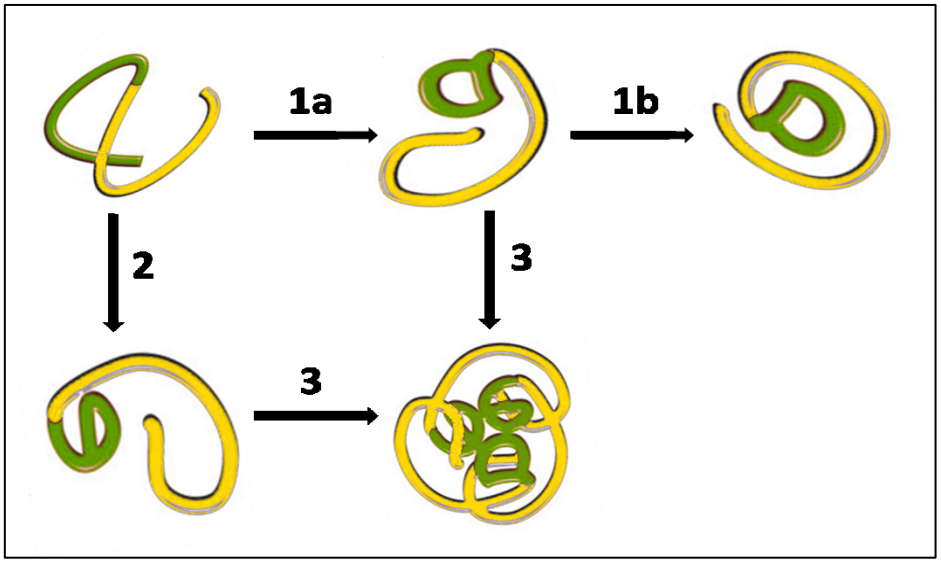


Figure 5.1: A Simplified model to demonstrate protein folding. The correct folding pathway of a protein through formation of correct intermediate (1), misfolding of protein (2) and aggregation between the correct intermediate and misfolded protein to form inclusion bodies (3).

Although conventional techniques in general are laborious and often result in poor recovery (Yamaguchi and Miyazaki 2014), a number of studies have successfully shown refolding of IB into biologically active protein. For example, N-carbamoyl-D-amino acid amidohydrolase (DCaseH) from *Agrobacterium radiobacter* was overexpressed in *E. coli*.

The resulting protein formed both native and denatured forms in the soluble fraction and partly as IB. The denatured His-tagged DCaseH was purified by single step purification from crude lysate and the IB were solubilised using 8 M urea. A dilution coupled with additives refolding was done in both cases. It was found that refolding done in the presence of 50% glycerol was able to restore the activity of DCaseH and the native-like secondary structure of the denatured DCaseH (Chen, Lin et al. 2005).

In another example, a recombinant L-asparaginase was expressed in the form of IB in *E. coli*. In order to obtain this protein in its functional form, urea was used to denature the protein and dilution was used to facilitate protein refolding. Purification of the refolded asparaginase was done through ion exchange and gel-filtration chromatography. Through this technique, approximately 50% of the protein was found to be active and the specific activity was approximately 190 U/mg which is comparable to that of the commercially available asparaginase (Upadhyay, Singh et al. 2014).

In the case of IB containing polyhistidine (His) fusion tags, on-column refolding using immobilised metal affinity chromatography (IMAC) has been used successfully for protein refolding. In this technique, urea denatured proteins are bound to metal charged resins packed in a column and exposed to a detergent wash to prevent misfolding. The denaturants are then removed by washing the column with either descending gradient concentration of denaturant (Zhu, Li et al. 2005) or with low molecular additives like β-cyclodextrins (Oganesyan, Kim et al. 2005) to remove detergent and promote folding as well as reagents to allow formation of disulphide bonds. The refolded protein is then eluted using imidazole. An example of this method is the refolding of His-tagged superoxidase dismutase (SOD) which was produced as inactive inclusion bodies in *E. coli.* The urea denatured protein was allowed to bind to nickel-NTA sepharose and an on-column refolding was carried out using buffer containing linear decreasing urea gradient in the presence of a mixture of reduced and oxidised forms of gluthathione (GSH) to produce biologically active SOD (He, Yuan et al. 2002).

In order to determine if the refolded protein has a defined secondary structure, circular dichroism (CD) spectroscopy is often carried out. CD spectroscopy is a useful technique to study the conformation of a biological molecule based on the interaction of circularly polarised light with chiral molecules. Most biological samples are chiral, i.e. all amino acids are chiral with the exception of glycine. When polarised light is passed through a chiral molecule, the difference in the absorbance between the left-handed (L) and right-handed (R) circularly polarised light will result in elliptical polarisation which can be measured as a function of wavelength (Greenfield 2006). The CD of the secondary structure is different from that of the side chains. This technique is quite informative in the identification of secondary structural elements in a protein.

Different types of secondary structure have different spectra in the far-UV. An illustration of CD spectrum for each common secondary structure element is shown in Figure 5.2. The information obtained from the CD spectra can then be analysed using a number of algorithms to estimate the secondary structure compositions in a protein.

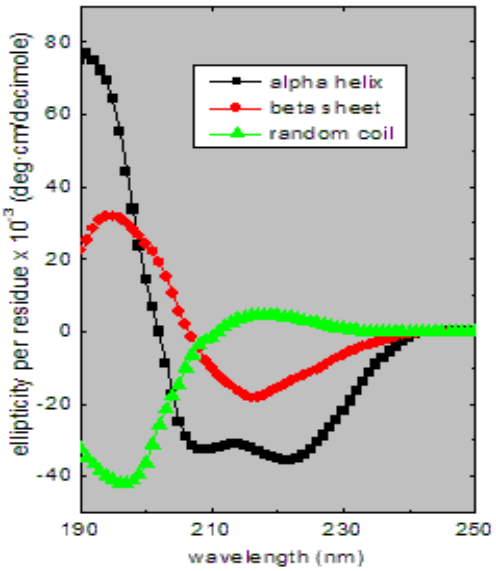


Figure 5.2: Circular dichroism (CD) spectra for the each secondary conformation. ■– alpha helix, ■– beta-sheet and ■- random coil (Used from: http://www.ap-lab.com/circular_dichroism.htm last accessed 25/10/14)

Protein modelling is a widely used technique for structural and functional prediction of a protein sequence. The traditional method of predicting a three dimensional (3D) structure of a protein is by developing a comparative model, which involves four main steps; (i) identification of homologous template(s) that have a sequence similarity of more than 30%, (ii) alignment of target sequence with that of identified template(s), (iii) 3D model building and (iv) evaluation of the model quality (Arnold, Bordoli et al. 2006). The quality of a predicted 3D structure depends upon the sequence similarity between the target protein and the template(s) used to generate the model.

In the case of a target sequence with a similarity of less than 30% with the templates, protein modelling using sequence homology will result in a poor model. Instead a protein threading approach can be applied using multi-remote templates. RaptorX is one package developed to address the issue of 3D homology modelling using low sequence similarity templates. Protein threading using RaptorX provides information on the structural alignment between the target protein and the distantly related templates (remote templates) based on fold recognition of a number of solved protein structures deposited in the Protein Data Bank (PDB) (Källberg, Wang et al. 2012). A number of publications have stated that functional prediction of a protein using structural similarity is more accurate than sequence homology as the structures are more evolutionarily conserved than sequences (Maiti, Van Domselaar et al. 2004).

In this chapter, His-tagged Ah from *R*. UKMP-5M that was expressed as inclusion bodies as shown in Chapter 4, was refolded to obtain a functional protein. The IB were solubilised in 6 M guanidine hydrochloride and on-column protein refolding was carried out by IMAC using decreasing concentrations of urea. The refolded Ah was collected in fractions and analysed with one-dimensional SDS-PAGE gel. Liquid chromatography –tandem mass spectrometry (LC-MS/MS) technique was used to confirm the amino acid sequence of the protein. Circular dichroism (CD) spectral measurement was used to analyse the secondary structure of the refolded protein and its activity with a number of amides were also assayed.

Structural analysis and the three-dimensional structure of Ah was predicted using RaptorX software.

5.2 Materials and methods

5.2.1 Growth and induction

Overnight cultures of *E. coli* BL21(DE3) carrying the pET-28a_AHntc were inoculated into 1 L LB with kanamycin at 50 µg/mL. Cultures were incubated at 37^0C in a shaking incubator (220 rpm) until reaching an absorbance (600 nm) between 0.5 and 0.6. Expression of Ah was induced by using Isopropyl- β-D-1-thiogalactopyranoside (IPTG) at 1 mM final concentration. The culture was incubated for a further three to four hours and the bacterial cells were pelleted by centrifugation at 4000 rpm for 15 min.

5.2.2 Preparation of inclusion bodies

Table 5.1: List of buffers and their compositions used for the inclusion bodies preparation, refolding and purification of amidoydrolase (Ah).

Buffer	Composition
Lysis buffer I	20 mM Tris. Hcl (pH8.0)
Lysis buffer II	20 mM Tris. Hcl, 2 M urea, 0.5 M NaCl, 2% Triton X-100 (pH 8.0)
Binding buffer	20 mM Tris. Hcl, 0.5 M NaCl, 5 mM imidazole, 6 M guanidine hydrochloride, 1 mM 2-mercaptoethanol (pH 8.0)
Washing buffer	20 mM Tris. Hcl, 0.5 M NaCl, 20 mM imidazole, 6 M urea, 1 mM 2-mercaptoethanol (pH 8.0)
Refolding buffer	20 mM Tris. Hcl, 0.5 M NaCl, 20 mM imidazole, 1 mM 2-mercaptoethanol (pH 8.0)
Elution buffer	20 mM Tris. Hcl, 0.5 M NaCl, 250 mM imidazole, 1 mM 2-mercaptoethanol (pH 8.0)

The bacterial pellet was resuspended in 40 ml of lysis buffer I and disrupted by sonication on ice (4 x 10 sec) and centrifuged at 4000 rpm for 10 minutes at 4^0C. Supernatant was discarded and the pellet containing insoluble protein was resuspended in 30 ml of cold lysis buffer II and sonicated as above. Pellet containing IB was separated through centrifugation and subjected to a second round of urea wash. The pellet was then resuspended in 50 ml binding buffer and stirred for 60 minutes at room temperature. Centrifugation was used at 11,000 rpm for 20 minutes to clarify the solution and the supernatant was collected for purification.

5.2.3 Column preparation, purification and refolding

A clean column was loaded with 10 ml Ni-NTA His.Bind resin (Novagen) and washed with 50 ml distilled water. The column was then equilibrated with 50 ml binding buffer and the supernatant collected in section 5.2.2 was loaded into the column. The column was first washed with 100 ml binding buffer followed by a second wash with 100 ml washing buffer.

Refolding of the bound protein was performed with 100 ml of refolding buffer containing urea at a descending gradient concentration (wash buffer was diluted with refolding buffer), starting with 6 M and finishing without urea. The column was continued washing with 50 ml of buffer without urea after the gradient wash. The refolded recombinant protein was eluted using 10 ml of elution buffer collected in 1 ml fractions and each fraction was analysed on an SDS-PAGE gel as described in section 3.2.10 (Chapter 3). All fractions were pooled and concentrated using an Amicon ultra-15 centrifugal filter unit (Merck Milipore) and the concentration of the protein was measured as described in section 3.2.3 (Chapter 3).

5.2.4 Protein identification using nano LC-MS/MS

A protein band of a purified fraction from SDS-PAGE gel in section 5.2.3 was excised in a clean condition and in-gel digestion was performed as described in section 3.2.11 (Chapter

3). Protein identification was done using nano LC-MS/MS as described in section 3.1.12 (Chapter 3).

5.2.5 Protein dialysis

Protein dialysis was performed to remove the imidazole from protein sample using D-Tube dialyser kit (Merck Milipore), based on manufacturer's protocol. The dialysis buffer consists of 20 mM Tris.HCl, 0.5 M NaCl and 1 mM 2-mercaptoethanol pH 8.0.

5.2.6 Circular dichroism spectroscopy

Circular dichroism (CD) spectra were obtained on a JASCO J715 CD system using a 1-mm path length cell from 260 to 200 nm. Protein concentration used was 0.2 mg/mL and each spectrum obtained was an average of three individual scans. Baselines obtained with dialysis buffer in which the protein was dissolved were subtracted from protein sample spectra.

5.2.7 Activity assay for the purified and refolded Ah

The refolded Ah was tested for its activity towards a number of amides. Assays were done on a 96-well plate (BD Falcon). 50 µg of Ah was added into 100 µl of reaction buffer containing 20 mM sodium phosphate (pH 8.0), 0.5 M NaCl, 1 mM 2-mercaptoethanol and 100 mM substrate. Reaction was incubated at 37^0C for an hour. A colorimetric indophenol blue method was used to detect and measure ammonia release from the amide substrates (Fawcett and Scott 1960). The detailed preparations for each reagent are given in Table 5.2 and all procedures for this method was carried out in a safety fume hood. To each well (containing 100 µl of sample), 4 µl of phenolic alcohol reagent and 4 µl sodium nitroprusside solution were added and left to stand for 5 minutes. Following this, 10 µl of oxidising solution

was added to the mixture and incubated in the dark at room temperature for 1 hour before reading the absorbance at 630 nm. Concentration of ammonia was determined by comparing the absorbance against a standard curve of known concentrations of ammonium chloride.

Table 5.2: List of reagents and their compositions for indophenol blue method used to measure ammonia release. After addition of reagents, samples were incubated in dark for an hour before absorbance measurement at 630nm. All preparations and procedures were carried out in a safety fume hood.

	Stock Reagents	Preparation / composition
1.	Phenol-alcohol reagent	1 g of phenol was dissolved in 95% ethyl alcohol to a final volume of 10 ml.
2.	Sodium nitroprusside	0.1 g of sodium nitroprusside was dissolved in 20 ml of deionised water
3.	Alkaline complexing reagent	10 g of trisodium citrate and 0.5 g of sodium hydroxide was dissolved in 50 ml of deionised water
4.	Sodium hypochlorite solution	5 ml of deionised water was added to 5 ml of 10-15% sodium hypochlorite
5.	Oxidising solution	250 µl of sodium hypochlorite solution (4) was added to 1 ml of alkaline complexing reagent (3). This solution was prepared fresh prior to use.

5.3 Results

5.3.1 Refolding of solubilised inclusion bodies

Isolation of recombinant protein produced as inclusion bodies often results in high degree of purification since majority of the contaminating species such as soluble and membrane proteins have been eliminated (Vallejo and Rinas 2004). In this study, mechanical disruption using sonicator was performed to remove protein contaminants that may co-precipitate with other insoluble cell material during inclusion bodies extraction. Cell pellet containing the

inclusion bodies of Ah was solubilised in 6 M guanidine hydrochloride (GdnHCl) to induce unfolding by eliminating the non-covalent interactions between protein molecules. Although some studies have shown that urea is a better destabiliser than GdnHCl (Tsumoto, Ejima et al. 2003) and (Lim, Rösgen et al. 2009), urea solubilisation of inclusion bodies is pH dependent (i.e. high alkaline pH values are required), and therefore optimal pH conditions for solubilisation varies among proteins (Estapé and Rinas 1996).

In addition, the presence of 2-mercaptoethanol in the solubilisation buffer is necessary to reduce the intra- and intermolecular disulfide bonds and to maintain the protein molecule in the unfolded state (Yamaguchi and Miyazaki 2014). The denatured inclusion bodies which carry a histidine fusion tag were purified by metal affinity chromatography using nickel-charged nitrilotriacetic acid (NTA) resins packed in a column. The string of histidine residues at the N-terminus of recombinant Ah contains imidazole functional groups that will bind to immobilised metal ions such as copper, nickel, zinc and cobalt. This interaction forms a complex that can be easily purified from other endogenous proteins.

Solubilisation of inclusion bodies will result in soluble protein that has no structure and thus, is functionally inactive. This protein has to be subjected to refolding conditions that allow the native-like conformation to be formed. The recovery of a biologically active protein is highly dependent on the gradual removal of denaturants, as the unfolded protein will tend to aggregate again in the absence of denaturant. The application of descending linear gradient of urea solution in a previous study was found to efficiently assist protein refolding (Gu, Su et al. 2001). In the present study, buffer containing decreasing concentration of urea (from 6 M to one without urea) was used to refold Ah. The refolded protein was then washed with buffer containing a low concentration of imidazole (5 mM) to remove other contaminating proteins that weakly interact with the nickel resin. The elution of refolded Ah was done with excess of imidazole (250 mM) and collected in 1 ml fractions. An SDS-PAGE was done to analyse the tail fractions. The molecular weight of Ah is 40 kDa and the purified

His-tagged Ah is expected to have a molecular weight of 45 kDa as a result of the combined expression of Ah, His-tag and a short region of plasmid before Ah expression.

Table 5.3: Amino acid composition in amidohydrolase (Ah). Total number of amino acid residue in the protein is 347, the molecular weight of the protein is 40 kDa and predicted isoelectric point (pI) is 4.98

Amino acid	Abbreviation	No. of residues / Percentage
Alanine	Ala / A	24 / 6.92
Cysteine	Cys / C	3 / 0.86
Aspartic acid	Asp / D	29 / 8.36
Glutamic acid	Glu / E	32 / 9.22
Phenylalanine	Phe / F	16 / 4.61
Glycine	Gly / C	25 / 7.20
Histidine	His / H	16 / 4.61
Isoleucine	Ile / I	20 / 5.76
Lysine	Lys / K	17 / 4.90
Leucine	Leu / L	31 / 8.93
Methionine	Met / M	4 / 1.15
Asparagine	Asn / N	9 / 2.59
Proline	Pro / P	15 / 4.32
Glutamine	Gln / Q	13 / 3.75
Arginine	Arg / R	19 / 5.48
Serine	Ser / S	10 / 2.88
Threonine	Thr / T	14 / 40.35
Valine	Val / V	21 / 60.69
Tryptophan	Trp / W	8 / 2.30
Tyrosine	Tyr / Y	21 / 60.69

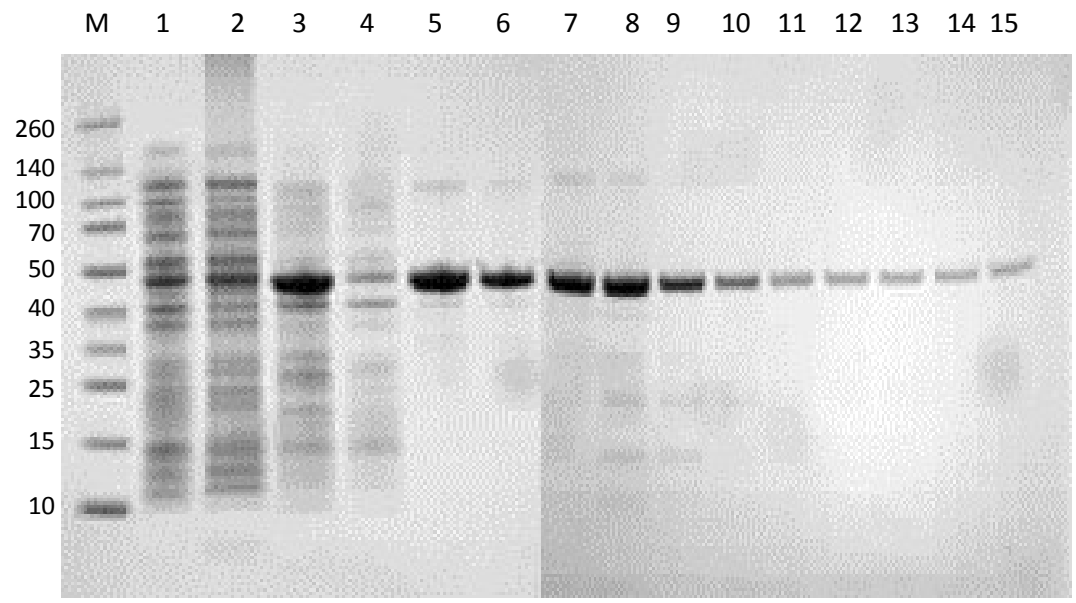


Figure 5.3: SDS-PAGE gel of protein samples obtained at each step during purification of Ah. *E. coli* BL21(DE3) containing recombinant plasmid pET-28a_AHntc was lysed for soluble protein (lane 1) and membrane protein extraction (lane 2). The cell debris was resuspended in 6 M GdnHCl to solubilise the inclusion bodies (lane 3) and removed through centrifugation. The supernatant was loaded onto column packed with Ni-NTA and the flow through was collected (lane 4). The refolding of Ah was done using decreasing linear concentrations of urea and the refolded protein was collected in fractions of 11 x 1 ml (lanes 5 to 15). Lane M represents protein standards. Samples were separated on 4-12% SDS-PAGE gradient gel. Protein staining was with Coomassie instant blue.

As shown in Figure 5.3, approximately 85% of the IB protein (lane 3) was recovered after the refolding process (lane 5). Bands at 45 kDa from lanes 3 to 15 are likely to represent the purified Ah. Although the protein samples extracted from soluble and membrane fractions do have a slightly strong band at the same molecular weight (lanes 1 and 2), western blotting with anti-6X His tag antibody has shown that these bands do not represent the His-tagged protein, and therefore may contain other endogenous proteins. To further confirm that the purified fractions represent the recombinant Ah, the band at 45 kDa in lane 8 was excised and in-gel digestion with trypsin was performed. The peptide fragments were then analysed using nano LC-MS/MS and the amino acid sequence coverage was approximately 77%.

Hypothetical protein (100%), 40,166.0 Da
Custom sequence
43 exclusive unique peptides, 97 exclusive spectra, 268 total spectra, 267/347 amino acids (77% coverage)

MYEKDGQQYF IVDGHVHIWD GRESNQKNVH GRQFIDCFYD YHKNLSPEEE VWDYDTYTYY
GGERLMKDLF VDGHVDHAIF QATLLSEFYH NGFGQTEEAF ELTRKHPDKL TYNHAYDPRH
GEAGLEQLRR DAERMNLQGV KLYTAEWHGD SRGYKLDDPW SRRYLEECIE LGIKNIHVHK
GPTIRPLDRD AFDVADVDKV ATDYLELRFV VEHVGLPRLE DFCWIATQES NVYGGLAVAI
PFIHTRPRYF AQIIGELLYW LGEDKILFSS DYALWTPKWL IEKFVDFQIP EDMTEYAPIT
IEQKKKILGL NAAALYDIDV PEHLRLPEPA GQEGVEVAAG AREAVSS

Figure 5.4: FASTA format sequence for amidohydrolase (Ah) with 77% of sequence coverage identified by LC-MS/MS. The matched peptide sequences are highlighted in yellow.

5.3.2 Sequence analyses and structure prediction

The nucleotide sequence of a 'hypothetical protein' obtained from the draft genome sequence of *R*. UKMP-5M was analysed against the National Centre for Biotechnology Information (NCBI) database using blastX, to identify proteins with significant sequence similarity. The result of this analysis revealed that this protein has 100% amino acid sequence homology to a metallo-dependent hydrolase, called amidohydrolase (Ah) in a number of *Rhodococcus* species. Members of this superfamily show conservation in their (TIM)-barrel fold and their active sites. In most cases, Ah-type proteins have been reported to contain a metal binding site, which is essential for catalytic activity (Marchler-Bauer and Bryant 2004). An example of a well-studied metal-dependent hydrolase is the urease-alpha from *Bacillus pasteuri*, whose structural details have been determined using X-ray crystallography (Benini, Rypniewski et al. 1999). However, Ah from *R*. UKMP-5M only shares 19% sequence similarity with this protein.

There are a number of web based and software packages available for secondary and tertiary structure prediction of a protein. One of the most common methods to obtain a reliable prediction is by generating a comparative model. The concept of comparative

modelling is based on the use of evolutionarily related protein structure(s) which have been solved previously. These structures having a significant sequence homology (more or equal to 30%) with the target protein serve as templates for the building of a three-dimensional (3D) model of the protein of interest (Ginalski 2006). An example of an automated protein structure prediction server that does a 3D-homology modelling is Swiss-Model (Arnold, Bordoli et al. 2006). However, this approach is limited to the identification of templates that have more than 30% sequence similarity to the protein of interest. In the case of target proteins which have less sequence homology to any previously solved protein structures in Protein Data Bank (PDB), a low-homology protein threading approach can be applied to generate a 3D model. This technique uses the sequence information (if available) and structural information based on fold recognition against a number of solved protein sequences in a database to generate a 3D model (Peng and Xu 2010).

In this study, the template(s) identification for Ah performed with the Protein Data Bank (PDB) was revealed to contain a highest sequence similarity of only 18% to an uncharacterised hydrolase from *Bordetella bronchiseptica*. As a consequence, the generated 3D model based on sequence homology was not of a good quality. Hence, a protein threading approach using fold recognition was performed. RaptorX, a protein structure and function prediction package (Peng and Xu 2011) was used to predict the structure information and generate a 3D conformation of Ah. A multitemplate approach based on information from inter-template similarity was used to obtain an improved alignment as compared to the pairwise alignments. The templates used to predict the tertiary structure of Ah is given in Table 5.4. To evaluate the quality of the model predicted, RaptorX assigns confidence scores such as P-value, global distance test (GDT) and un-normalised GDT (uGDT) for relative global quality, absolute global quality and modelling error at each amino acid residue respectively (Vallejo and Rinas 2004).

Table 5.4: List of structural templates and their information from Protein Data Bank (PDB) that were used to generate a 3D conformation of Ah.

Template ID	Protein Name	Source
3irsA	Uncharacterised protein BB4693	*Bordetella bronchispetica*
3ij64	Uncharacterised metal-dependent hydrolase	*Lactobacillus acidophilus*
4infA	Metal-dependent hydrolase	*Novosphingobium aromaticivorans* DSM 12444
4dlfA	Amidohydrolase 2	*Burkholderia multivorans* ATCC 17616
3s4tA	Amidohydrolase 2	*Polaromonas sp*. JS666

MYEKDGQQYFIVDGHVHIWDGRESNQKNVHGRQFIDCFYDYHKNLSPEEEVWDYDTYTYYGGERLMKDLFVD
GHVDHAIFQATLLSEFYHNGFGQTEEAFELTRKHPDKLTYNHAYDPRHGEAGLEQLRRDAERMNLQGVKLYT
AEWHGDSRGYKLDDPWSRRYLEECIELGIKNIHVHKGPTIRPLDRDAFDVADVDKVATDYLELRFVVEHVGLP
RLEDFCWIATQESNVYGGLAVAIPFIHTRPRYFAQIIGELLYWLGEDKILFSSDYALWTPKWLIEKFVDFQIPEDM
TEYAPITIEQKKKILGLNAAALYDIDVPEHLRLPEPAGQEGVEVAAGAREAVSS

Figure 5.5: Amino acid sequence of amidohydrolase (Ah) from *R*. UKMP-5M. Residues in black are modelled and residues in red are not modelled due to lack of suitable templates in the PDB.

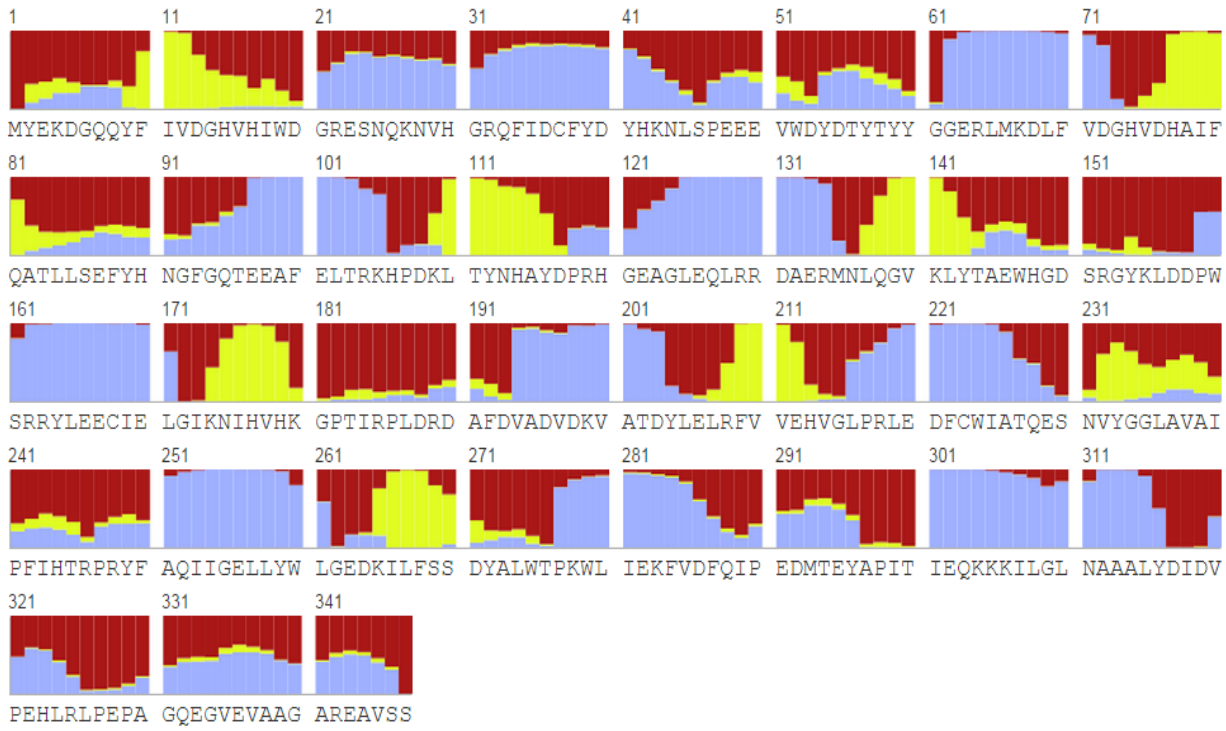


Figure 5.6: A snapshot view of the predicted distribution of a three-state secondary structure for each amino acid residue in amidohydrolase (Ah) ■-represents α- helix, ■-represents β-sheet and ■-represents loop.

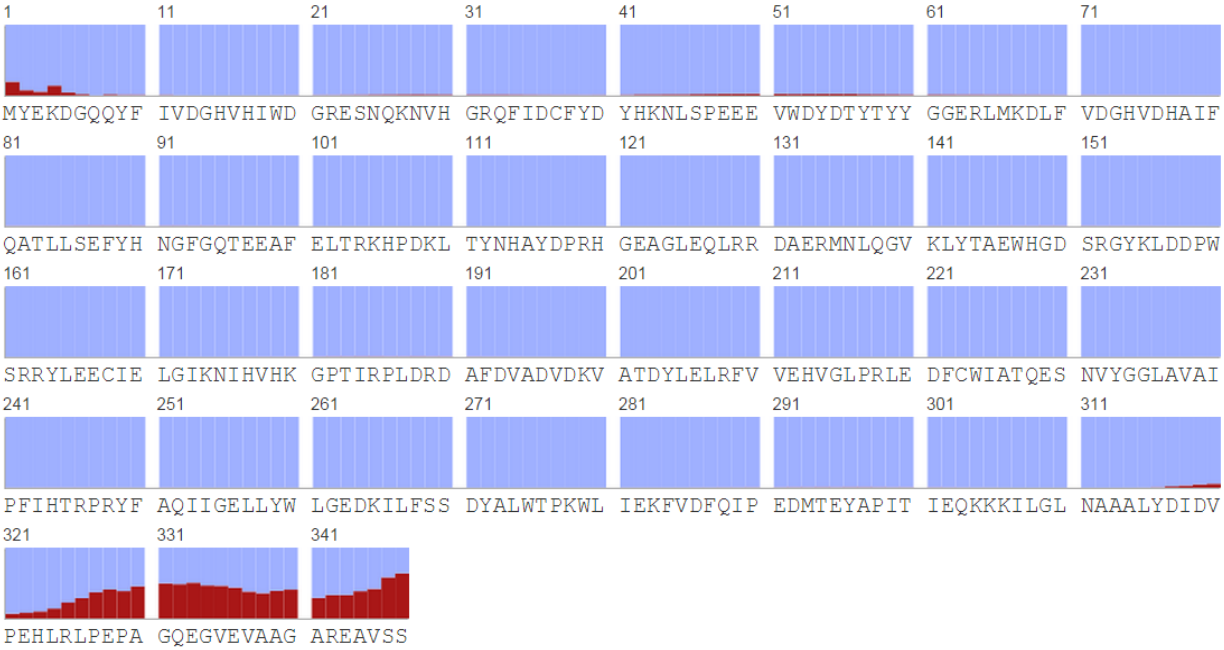
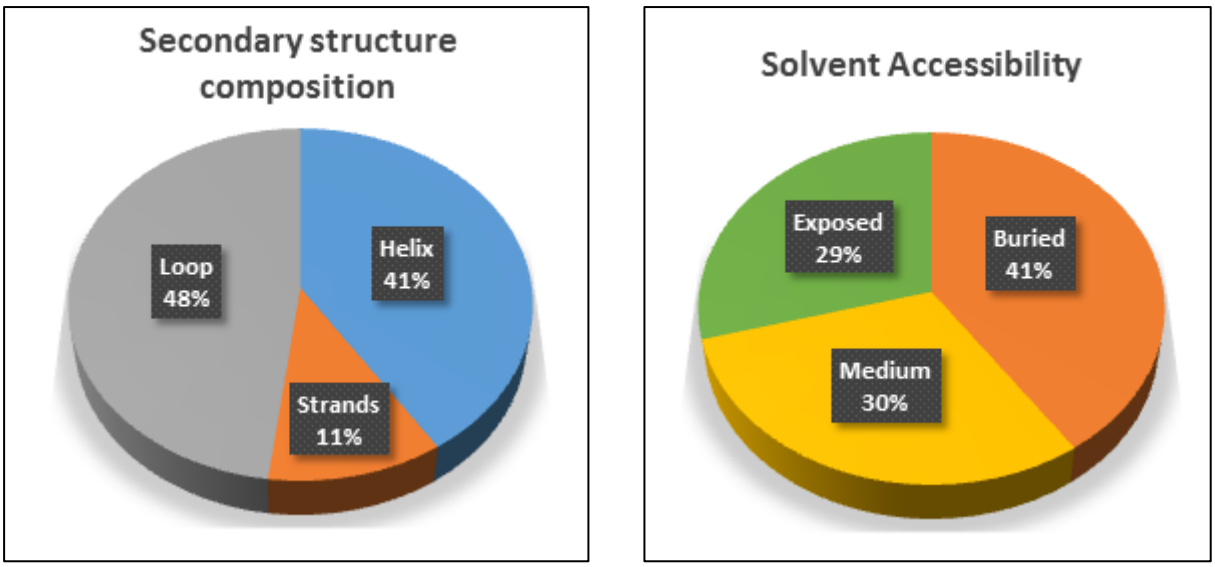


Figure 5.7: A snapshot of the predicted disorder distribution for each amino acid residue in amidohydrolase (Ah) during 3D-modelling using RaptorX. ■- non-disorder and ■-represents disorder. Result indicates that the majority of the residues are ordered in the 3D structure predicted.



A B

Figure 5.8: Summary of (A) the compositions of three-state secondary structure and (B) the solvent accessibility of Ah predicted from structural homology using RaptorX.

Among the 347 amino acid residues in Ah, 324 residues were modelled and 23 were not since there were no homologous templates available for this region (Figure 5.5). The structural alignment of Ah with five different templates predicts that Ah predominantly consists of loops (48%) that connect together the β-strands, strands to α-helices or helices to each other. The majority of the secondary structures in the protein are α-helices, which correspond to more than 40% of the total composition of the protein followed by β-strands with 11% (Figure 5.6 and Figure 5.8A). An essential prerequisite to protein analysis is the prediction of disordered regions. Disordered regions are part of a protein that does not have a 'fixed' secondary structure in their native state. Therefore, characterising this region can be difficult (Plaxco and Gross 2001). The protein disorder prediction for Ah shown in Figure 5.7 suggests that partially unstructured regions can be observed at the C-terminus of Ah (last 23 residues) and therefore, tertiary structure for this region could not be predicted

Solvent accessibility of a protein is a measure of exposed surface of an entire protein to the solvent surrounding. Folded proteins have side chains which are either polar or non-polar that would determine the degree of solvent accessibility. Formation of the tertiary

structure of a protein is driven by the hydrophobic interactions between the side chains and the solvent in its environment (Chen and Zhou 2005). The solvent accessible surface area can be classified into three groups, buried surfaces that are not solvent-accessible, exposed surfaces that are solvent-accessible and intermediate, which has a mixture of both, buried and exposed surfaces. In the present study, Ah is predicted to contain a balanced distribution of solvent accessibility per residue with buried surface area being slightly dominant over the exposed and intermediate surfaces (Figure 5.8B). Using the information from the secondary structures, a three-dimensional (3D) model for Ah from *R*. UKMP-5M was predicted.

A quality evaluation based on scoring functions was carried out as a measurement of compatibility between the target sequence and the template structures. This would discriminate between a good and a bad model and thus, affects the ability to predict protein conformations (Xiang 2006). When a model falls within the satisfactory level of quality, hypotheses about some local features such as possible mechanisms, active sites positions, catalytic residues, ligand-binding motifs or the class of substrate bound can be derived from the template, which in turn may be useful in predicting the functional information of the target sequence (Soding, Biegert et al. 2005). On the other hand, when a model fails to meet the required quality, the protein conformation prediction becomes incorrect, as there would be parts of undetermined protein fold in the generated 3D model.

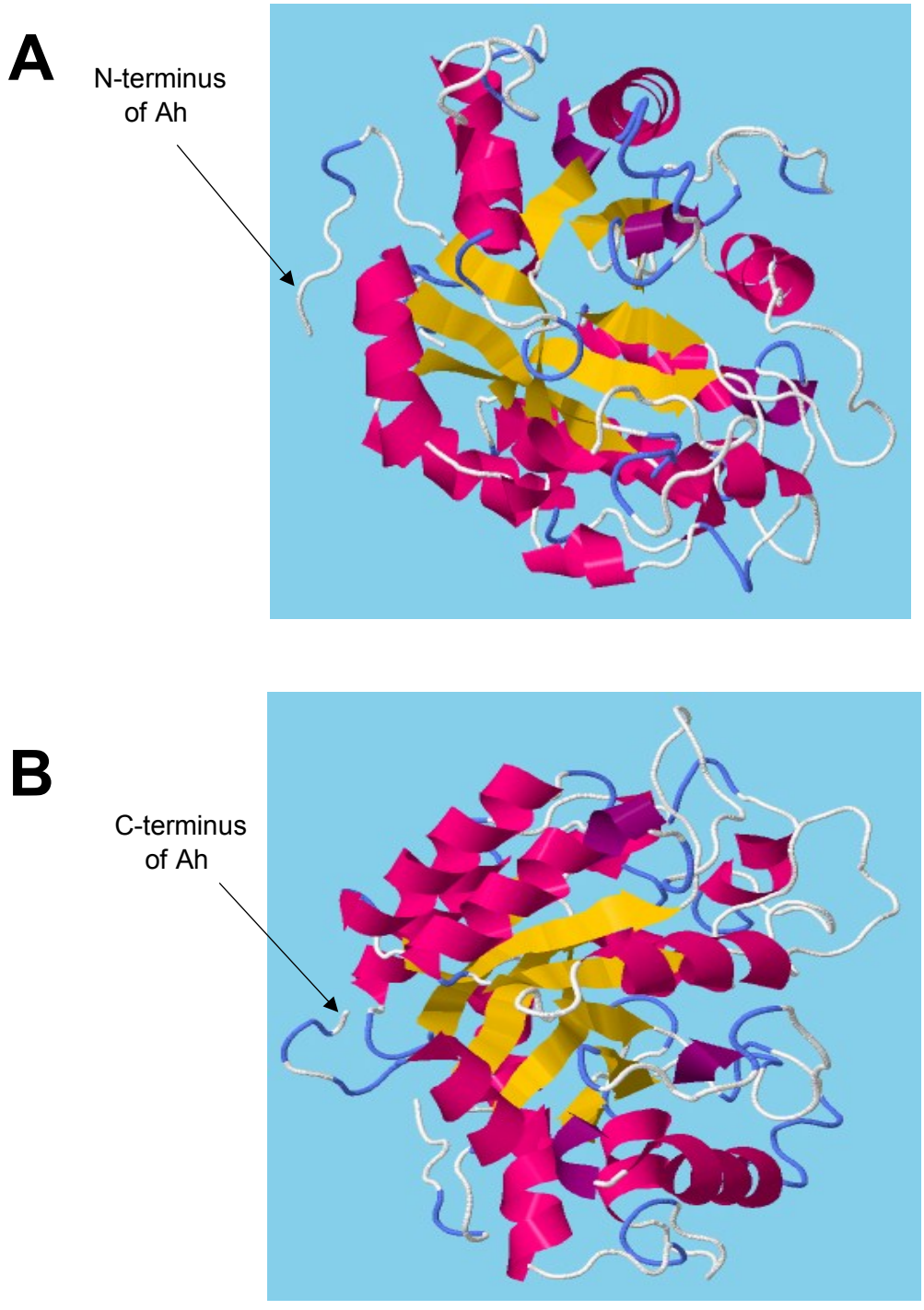


Figure 5.9: The predicted 3D-model for amidohydrolase (Ah) from *R*. UKMP-5M based on structural homology. Model was generated using multitemplate protein threading approach in RaptorX. (A) The front view and (B) the back view of the model. The templates used are given in Table 5.4. The α-helices in the structure are shown in pink, β-sheets are shown in yellow and turns are shown in magenta.

RaptorX assigns several measures such as alignment score, P-value and uGDT(GDT) to assess the quality of a 3D model. As far as a template-based modelling is concerned, alignment between target protein and template is an essential step to obtain a reliable model. An alignment score in RaptorX indicates the number of residues that are able to align. The generated model has a score of 244/347, which means almost 70% of the amino acids were aligned. Besides that, a P-value is assigned to represent the relative quality of a model. It is a measure of the probability of a predicted model being worse than the best of a set of randomly generated models for a target protein. Hence, the smaller the P-value, the higher is the quality of a model. For protein containing mainly alpha-helices (which is the case in Ah), P-values below 10^{-3} applies for a good quality model. GDT stands for Global Distance Test, which is the score, assigned to assess a structural alignment of a target protein with its template(s) and uGDT is the unnormalized GDT, which represents the absolute model quality. For a protein with >100 residues, UGDT >50 is likely to be accurate (Peng and Xu 2011). In this study, the generated model has a P-value of 1.6×10^{-8} and uGDT(GDT) value of 187(58) which falls within the range of a good quality model.

The model structures shown in Figure 5.9 contains eight α-helices and eight β-strands which are linked in an alternate pattern by loops and turns along the protein backbone, which resembles a typical α/β TIM-barrel structure. Example of Ahs which have been reported to have similar structural fold include urease, phosphotriesterase and adenosine deaminase (Marchler-Bauer and Bryant 2004). Similar to homology modelling which provides information on the evolution, catalytic function and interaction through sequence alignment, threading-based modelling provides these information through structural alignment and this can be done by superimposing the target protein with the templates (Maiti, Van Domselaar et al. 2004). A structural alignment was done between the predicted 3D model of Ah and the templates using RaptorX and the outcome was visualised using RasMol molecular viewer software.

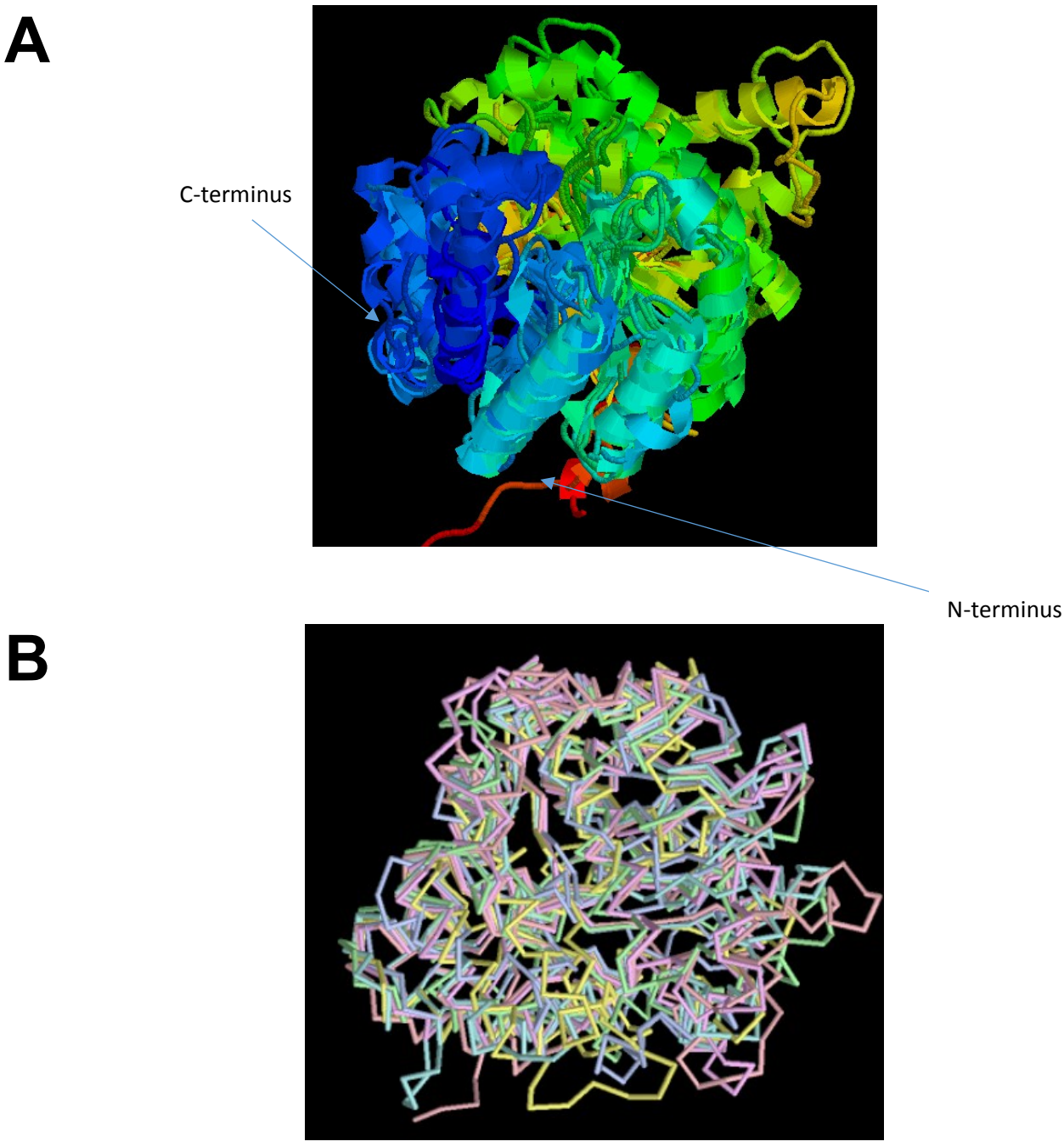


Figure 5.10: The superposed structure of Ah with five templates used to build the 3D model. (A) Ribbon representation of the alignment with colour shown according to residue numbering. (The N-terminus is represented by red and the C-terminus by blue). (B) Backbone of the six proteins (one target and 5 templates) superposed to each other using RaptorX. The structural alignment shows that the generated 3D model has extensive similarity in the folding with its templates (Tm score: 0.735 = 90% chance of to have similar fold) and therefore, may have similar evolution history.

Predicting the function of an uncharacterised protein can be challenging, especially when there is no activity information available on other homologous protein. In such cases, analysing the structurally conserved binding sites can be useful. Binding sites are evolutionarily conserved regions in protein molecules and are more directly linked to their functions. Hence, identifying similar binding site through similar residue patterns in other solved proteins may provide clues to predict its biological role (Konc, Hodoscek et al. 2013). In metalloproteins, they often contain a metal binding site responsible for biological functions. In general, metal-dependent Ah are reported to possess active site which is coordinated by up to four histidines and one aspartate residue (Marchler-Bauer and Bryant 2004). The majority of the members of Ah superfamily are hydrolytic enzymes that contains a mono- or binuclear metal center and is occupied by one or two metal ions (Seibert and Raushel 2005). However, some members of Ah with trinuclear metal center and without cofactor at all have been also discovered (Liu and Huo 2001).

The type of metal ion present in a protein is found to be enzyme-specific. For example, urease from *Helicobacter pylori* UreF contains metal site for two-nickel ions. This active site is connected by a lysine residue and is coordinated by histidine and aspartic residues along with a hydroxide ion that acts as a nucleophile during catalytic activity (Zambelli, Berardi et al. 2014). On the other hand, *Helicobacter mustelae*, in response to gastric neutralization, was found to produce another urease which is an iron-dependent in addition to a nickel containing urease (Carter, Tronrud et al. 2011). Therefore, identification of conserved binding site is the key step to predict the function of a protein.

Figure 5.11 shows the predicted binding site for Ah from *R*. UKMP-5M. The enzyme is likely to be a zinc-dependent enzyme as the template structures contain a single Zn^{2+} ion surrounded by two histidine residues and one aspartic acid, resembling the feature of a typical metal-dependent hydrolase.

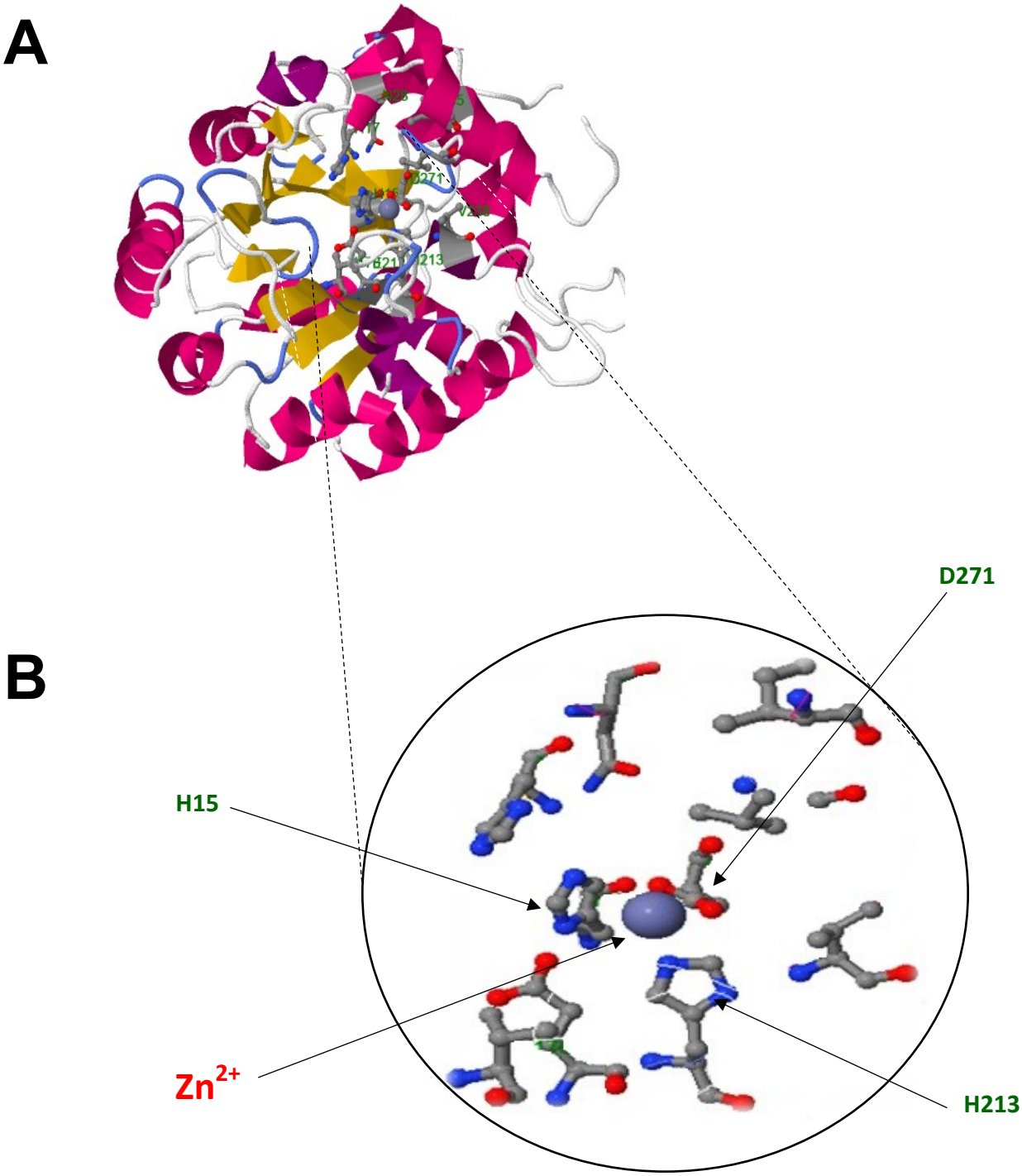


Figure 5.11: The predicted binding site of amidohydrolase (Ah) based on structural homology. Model was generated using multitemplate protein threading approach in RaptorX. (A) Full view of the 3D model and (B) Expanded view of the model which shows the presence of a single Zn^{2+}, with 2 histidine residues (H15 and H213) and one aspartic acid residue (D271). The α-helices in the structure are shown in pink, β-sheets are shown in yellow.

5.3.3 Circular dichroism (CD) spectroscopic analyses

In this study, CD spectroscopic measurements were performed to determine if (i) there is any difference in the structural composition of Ah obtained before and after the refolding, and (ii) the refolded protein has secondary structures as predicted by the threading modelling. The CD measurement in this study was done between 260 to 200 nm as the instrument was not able to collect data below 200 nm. Initially, the CD spectra were collected before performing a buffer exchange to remove the imidazole in the protein sample. As a result, the strong imidazole absorbance was found to generate high levels of background in the region below 220 nm. Therefore, dialysis against buffer lacking imidazole was done and the baseline was improved. As shown in Figure 5.12, there is a significant difference in the structure composition between the proteins assessed before and after refolding.

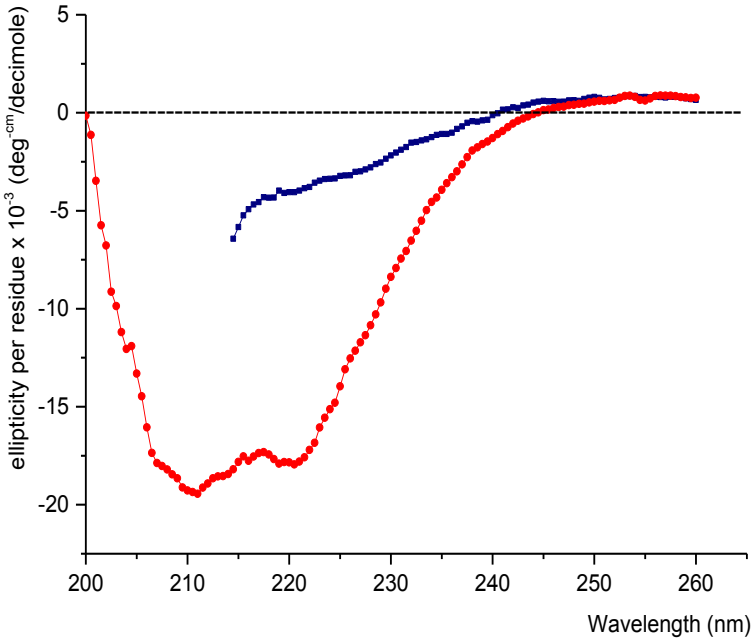


Figure 5.12: The far-UV spectrum of amidohydrolase (Ah) measured from 260-200 nm using 1 mm path-length cuvette. ■ – represents protein before refolding and ■ – after refolding. Baselines were corrected with a blank containing the same buffer.

Protein sample before refolding were solubilised using GdnHCl. This sample was purified and concentrated without performing the refolding experiment and hence, may typically contain unfolded structures. Protein sample after refolding is the protein that has undergone renaturation with decreasing concentrations urea to form its native-like structure. The predicted secondary structure for Ah (based on results in section 5.6 and 5.8A) indicates that this protein is primarily composed of loops followed by α-helices and a small percentage of β-strands. Analysis of the CD data using the online CD analysis, Dichroweb (http://dichroweb.cryst.bbk.ac.uk/ last accessed 25/10/14) revealed that the refolded Ah mainly contains disordered region (loops) which is consistent with the structural prediction done. However, a significant portion of strands was found to be misfolded and only a small percentage of helices were recovered as given in Table 5.5.

Table 5.5: Secondary structure composition of refolded amidohydrolase (Ah) predicted from sequence analyses and estimated from the CD spectra shown in Figure 5.12 using the online analysis site, dichroweb. CD data for denatured Ah could not be analysed.

Conformation	CD estimated Percentage (%)	Predicted Percentage (%)
Helix	12.5	41
Strand	32	11
Turns	12	48
Disordered	43	

Nevertheless, studies with similar observation have shown that even small percentage of secondary structure recovery is sufficient to activate a protein. For example, the refolding of N-carbamyol-D-amino acid amidohydrolase (DCaseH) produced as IB in *E. coli* was only able to recover approximately 7-8% of its helix structure. When compared to the native conformation of this protein, it was found that a significant portion of β-strands in this protein was misfolded. Despite the incomplete restoration of its native structure, a measurable level

of activity was still observed (Chen, Lin et al. 2005). Results from the CD analysis were useful to determine if the refolding experiment has given significant changes to the structural composition of Ah. Given that the protein has a defined conformation after refolding, enzyme assays can now be carried out to investigate its biological activity.

5.3.4 Enzyme Assay

The biological activity of Ah was investigated by testing its ability to release ammonia from amide containing compounds. An indophenol blue method was employed to quantify the ammonia released. This method is based on the reaction of ammonia with phenolate in alkaline solution to form indophenol (a blue complex) in the presence of an oxidising agent such as sodium hypochlorite. The addition of a metal containing catalyst like sodium nitroprusside helps to increase the rate of indophenol formation. This blue complex can be measured using a spectrophotometer at 630 nm. The Ah protein sample in this section was prepared in 20 mM Tris. Hcl, 0.5 M NaCl, 1 mM β-mercaptoethanol. However, the colorimetric indophenol blue method was not compatible with this buffer and a calibration curve prepared in this buffer was not successful. Therefore, buffer exchange through dialysis was performed with 20 mM phosphate buffer. A standard curve was constructed using known concentrations of ammonium chloride from a range of 0 to 50 µM. A linear relationship between the absorbance and concentration of ammonia was observed.

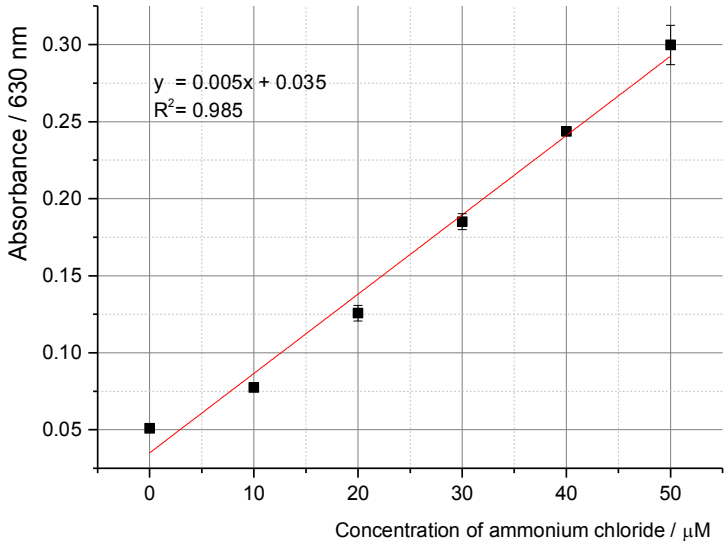


Figure 5.13: Calibration curve for ammonia measurement

Since homologues of Ah in a number of organisms remain uncharacterised, functional prediction and the correct substrate identification for this protein were difficult. Based on several studies for amidohydrolase-type protein in the literature, a number of amides were used as substrates to test the ability of Ah to release ammonia. Examples of substrates used were acetamide, propionamide, butyramide, valeramide and urea. Activity assay was also performed with whole cells containing inclusion bodies since a number of studies have reported that part of inclusion bodies can still be active (Wu, Xing et al. 2011, Huang, Zhang et al. 2013). Among the five substrates tested, detectable level of activity was observed only with urea as shown in Figure 5.14. The specific enzyme activity observed was 0.333 U/mg.

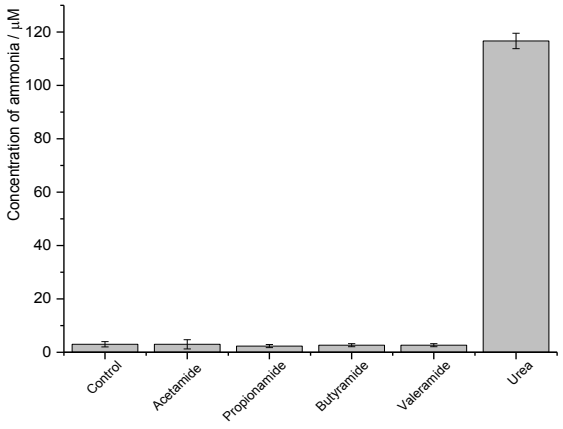


Figure 5.14: Ammonia release from five different amides by amidohydrolase (Ah). Concentration of each substrate used was 100 mM and enzyme reaction was carried out at 37^0C for 60 minutes.

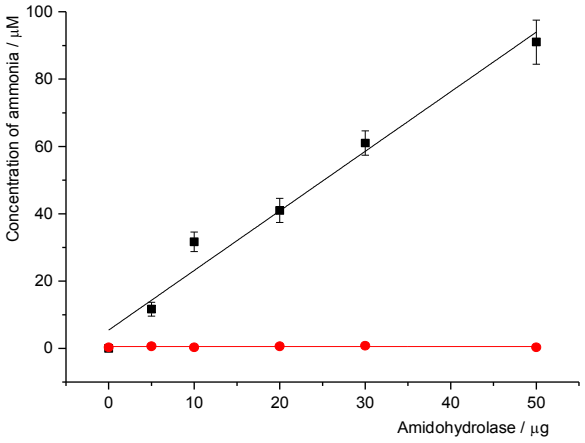


Figure 5.15: Ammonia release from 100 mM urea by different concentrations amidohydrolase (Ah). ■ – represents refolded Ah and ● – represents IB of Ah. Enzyme reaction was carried out at 37^0C for 60 minutes. Concentration of ammonia released was measured with indophenol blue method.

The ability of Ah to release ammonia from urea is a good indication that the protein is (either completely or partially) functional. The level of activity shown by Ah towards 100 mM urea is very low as compared to a typical jack bean-urease which has a specific activity of 1000 U/mg protein towards the same substrate (Abdallah, Dumas-Gaudot et al. 2012). The low levels of activity of this refolded protein can be due to several reasons. One of them could be

the incorrect substrate chosen for enzyme assay. Although the sequence analysis of Ah in this study shows homology to metallo-dependent hydrolases from *Rhodococcus*, this protein family comprises of a diverse set of enzymes such as urease alpha, adenosine deaminase, allantoinases, hydantoinases, chlorohydrolases and each enzyme has different substrate specificity. Hence, prediction of the most suitable substrate was challenging.

Besides that, poor recovery of a native-like structure during the refolding of Ah can also lead to a poor activity of an enzyme. As shown in Table 5.5, the refolded Ah contains a significant fragment of α-helix and β-strands, which are not properly folded when compared to the predicted composition of the secondary structures. This incomplete restoration might have happened due to the absence of correct metal ions in the refolding buffer. Some studies have shown that the presence of metal ions is essential to maintain the conformation and biological function of a protein. For example, the crystal structure of histone deacetylase (HDAH) was shown to contain a Zn^{2+} and the interaction between Zn^{2+} and the residues in the active site was found to facilitate the correct refolding of HDAH (Kern, Riester et al. 2007).

In this study, Ah was classified as a metallo-dependent hydrolase and 3D-structural model predicted that a Zn^{2+} ion is present around the binding site of Ah and hence, might be required for the correct orientation during protein refolding. However, some studies have shown that addition of metal in the buffer during enzyme assay can enhance the activity of a refolded protein. As such, five different metals were added in the assay buffer to a final concentration of 1 mM to study the effect of metal on the enzyme activity of refolded protein.

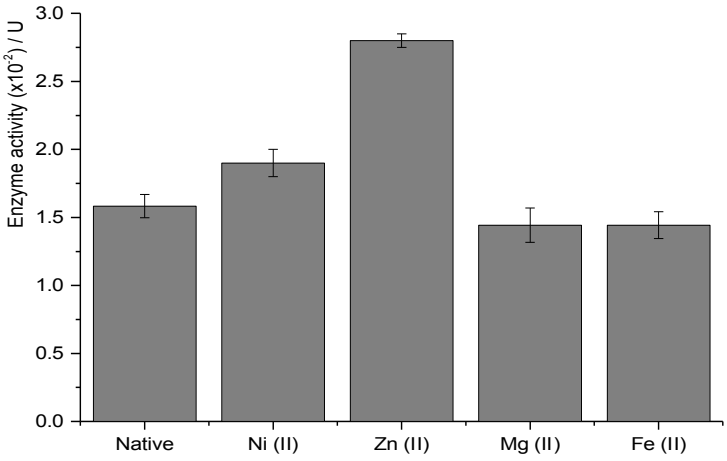


Figure 5.16: Enzyme activity of refolded Ah in the presence and absence of metal. Enzyme assay was carried out in the native condition (without the addition of metal) and with the addition of different metals at a final concentration of 1 mM. All reactions were incubated at 37^0C and concentration of ammonia released was measured with indophenol blue method.

As predicted, enzyme activity of Ah was enhanced in the presence of Zn^{2+}. However, slight increase in activity was also observed with the addition of Ni^{2+} in the reaction buffer and not with other metal ions. This implies that the activity of this protein is cofactor-specific and Zn^{2+} plays a crucial role in the catalytic function of the protein. Increase in enzyme activity with the addition of Ni^{2+}suggests that this metal can also act as a cofactor for this protein, but a weaker one as compared to Zn^{2+}. The absence of metal during refolding of Ah may have led to the significant misfolding, which is in agreement with a study done on refolding of histone deacetylase HDAH (Kern, Riester et al. 2007). The addition of Zn^{2+} in the refolding buffer during refolding process of Ah in future may help to prevent aggregation and improve the yield of a native-like conformation that is thus, biologically functional.

The purified and refolded protein was assayed for its activity towards 1-CB. Initially, the presence of sodium chloride in the protein buffer was found incompatible with the chloride release assay. Therefore, protein dialysis was performed with buffer in the absence of sodium chloride and the protein was reassayed for its ability to dechlorinate 1-CB. Ah, at a concentration of 0.5 mg/ml was added into air-tight glass vessels and sealed with rubber septa. 1-CB was added at a final concentration of 100 mM and chloride release was

followed. Similar set up was prepared with 100 mM urea as a positive control. Result of this investigation is shown in Figure 5.17. There was no chloride release observed from incubation with 1-CB but significant levels of ammonia were detected when Ah was incubated with urea.

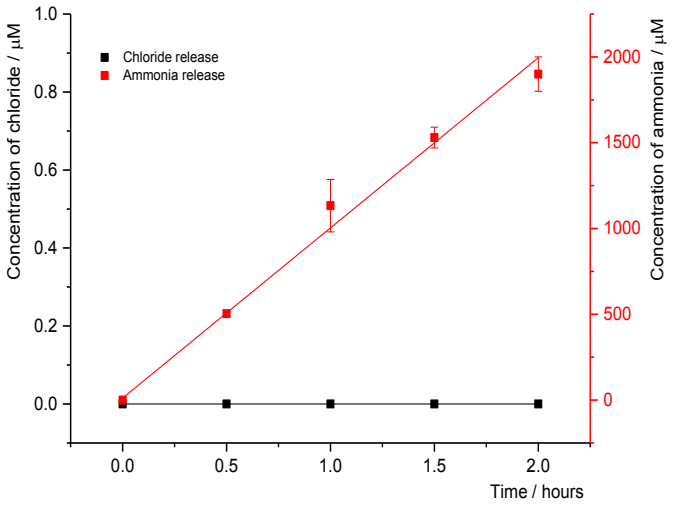


Figure 5.17: Activity of Ah assayed using 100 mM urea and 100 mM 1-CB. Enzyme reaction was carried out at 37^0C in a total volume of 1 ml. Concentration of ammonia released was measured with indophenol blue method and chloride release was measured using Bergman and Sanik method.

5.4 Discussion

The amidohydrolase (Ah) superfamily was first identified by Sander and Holm (Holm and Sander 1997). This superfamily comprises a set of functionally diverse enzymes that exhibit a variety of catalytic activities such as hydrolysis of amide, deamination of nucleic acids, decarboxylation and isomerization (Fan, Hitchcock et al. 2013). Their conservation of the (TIM) barrel fold and their metal binding site are the key features for the enzyme belonging to this group. The majority of the members of Ah were shown to have metal centres for one or two divalent ions. For example, adenosine and cytosine deaminase contains mononuclear metal centre for a single Zn^{2+} ion but urease contains a binuclear metal center which binds two Ni^{2+} ions (Liaw, Chen et al. 2003). In a typical hydrolysis

reaction, a metal is required for deprotonation of water molecule to generate hydroxide ions (nucleophiles) and the subsequent nucleophilic attack will result in the cleavage of scissile bond of the substrate (Seibert and Raushel 2005).

In the present study, a hypothetical protein was induced in response to the exposure to 1-Chlorobutane (1-CB) in *R*. UKMP-5M. Sequence analysis against NCBI has shown that this protein has homology to amidohydrolase (Ah) from several *Rhodococcus* strains but the involvement of Ah-type enzyme during degradation of haloalkanes has never been reported before and the upregulation of this protein during degradation of 1-CB by *R*. UKMP-5M was not understood. Moreover, in a number of *Rhodococcus* and several other strains, this protein remains uncharacterised. Therefore, a three-dimensional (3D) conformation of Ah from *R*. UKMP-5M was predicted based on fold recognition to attempt to predict information on structural and functional properties of this protein.

The predicted model in Figure 5.9 shows that, Ah contains $(\alpha/\beta)_8$ which resembles a typical (TIM) barrel fold that most other members of the Ah superfamily possess. As shown in Figure 5.11 this enzyme is predicted to contain a mononuclear metal center that most probably binds a single Zn^{2+} ion coordinated by two histidines (H15 and H213) and one aspartate residue (D271). The stimulation of catalytic activity by Zn^{2+} as the required cofactor by Ah provides some insight into the prediction of function in this enzyme. The zinc binding site in enzymes can be categorised into three groups: structural, catalytic and co-catalytic. (i) The structural binding site is coordinated by four protein ligands and no bound water is involved. The most preferred ligand is Cys>His>Glu>Asp. (ii) The catalytic binding site also involves four protein ligands but in the order His>Cys>Glu>Asp. In this site, zinc forms a complex with water and any three nitrogen, oxygen or sulphur donor from the ligand. In addition to the amino acid residue, a water molecule is also accepted as ligand. (iii) The co-catalytic zinc sites contain two or more metal ions, where at least one of them is zinc. These metals are bridged by a ligand like water molecule or carboxylate ligand. The common protein ligands observed in this catalytic site is Asp and His (Hernick and Fierke 2005).

Zinc is one of the most frequently used cofactors in metallohydrolases. One example of a member of Ah superfamily that has a metal center assembly similar to Ah in this study is the barbiturase (Bbr) from *Rhodococcus erythropolis* JCM 3132. Bbr was shown to have 48% sequence identity with a cyanuric acid amidohydrolase but has no significant sequence homology to other members of Ah superfamily. Bbr catalyses the hydrolysis of barbituric acid to ureidomalonic acid and the metal content analysis revealed that this protein contains approximately 4.4 mol zinc per mol of enzyme. The zinc-binding motif is located at the carboxyl terminal of this protein and contains two histidine (H322 and H324) and one aspartate residues (D320). Results from the site-directed mutagenesis of the histidine residues revealed that they are essential for both zinc binding and catalytic activity of this protein (Soong, Ogawa et al. 2002). Since Ah in the present study has a binding site similar to that of Bbr in *R. erythropolis* JCM 3132, the activity assay for Ah can also be done with barbituric acid as the substrate. However, due to time limitation, this assay was not set up and the assays were done using other amides that were readily available in the laboratory.

Given that Ah from *R*. UKMP-5M is predicted to contain a Zn^{2+} cofactor binding site, it is important to ensure the presence of this metal in all functional assays. The poor recovery of the native-like conformation may have resulted in the low level of activity. Consistent with the predicted metal and binding site for Ah from the 3D model, the enzyme activity was found to be enhanced in the presence of Zn^{2+}. Activity was also slightly higher in the presence of Ni^{2+} but not with other metal ions. However, the increased levels of the enzyme activity were not too much greater than the activity obtained in the absence of metal.

On the other hand, although induced by 1-CB, the role of Ah (if any) in response to 1-CB degradation is still poorly understood and no information is available in the literature to demonstrate this possibility. As shown in Figure 5.17, no chloride release was observed when the purified and refolded Ah was assayed with 1-CB. Hence, Ah is unlikely to be the enzyme that catalyses the degradation of 1-CB in *R*. UKMP-5M. The strict oxygen requirement for the chloride release from 1-CB is a strong evidence that 1-CB degradation is

most likely mediated by a monooxygenase-type protein, and this was further confirmed when monooxygenases were identified in the proteomic analyes. The induction of Ah in the induced protein samples is speculated to be a result of coexpression. Given that monooxygenase hydroxylase (Moh) and Ah identified in Chapter 3 are neighbouring genes that may belong to the same operon, there are high possibilities that these genes are transcribed and expressed together. Similar observations have been reported in many studies. For example, a transcriptomic analysis was done on *R. jostii* RHA 1 that was grown on propane and is able to degrade N-nitrosodimethylamine (NDMA). Results revealed that the metabolism of these compounds have not only upregulated the proteins responsible for degradation, but has induced the upregulation of a series of gene that belong to the same operon (Sharp, Sales et al. 2007).

Therefore, in order to provide concrete evidence that the dehalogenation of 1-CB in *R*. UKMP-5M is catalysed by monooxygenase, a gene knockout experiment should be carried out. Due to time constraints, this experiment was not conducted, but is discussed in the future perspective section in the following chapter.

CHAPTER 6 CONCLUSIONS AND FUTURE WORK

6.1 Thesis summary

The increased anthropogenic activities in recent years have led to the use and liberation of a number of xenobiotic compounds into the biosphere. Halogenated compounds are one of the more abundant environmental pollutants. The diverse use of these compounds mainly as biocides, degreasing solvents, plasticizers and intermediates for chemical syntheses has resulted in their widespread dissemination in soil and groundwater (Janssen, Oppentocht et al. 2001). In many studies microbial degradation of environmental contaminants, known as bioremediation, has been demonstrated as a potential approach to restore the environment (Paul, Pandey et al. 2005, Shankar, Kansrajh et al. 2014).

6.1.1 Physiological aspect of 1-CB metabolism in *R*. UKMP-5M (Chapter 2)

In the present study, the ability of an actinomycete isolated in Malaysia and designated as *Rhodococcus* UKMP-5M to degrade a number of halogenated compounds was tested. This bacterium was not able to utilise any of the chlorinated compounds tested in this study as a sole carbon source, but the resting cells of *R*. UKMP-5M were able to degrade a number of chlorinated compounds such as dichloromethane, trichloromethane, tetrachloromethane, 1,2-dichloroethane, 1,1,2-trichloroethane, 1-Chlorobutane (1-CB), 2-chloropropionic acid and 4-chlorobutanol (Table 2.3). The dechlorination rate of these compounds was found to be three times higher when the cells were grown in the presence of 1-CB. Therefore, 1-CB was chosen as the substrate to unravel its metabolism in *R*. UKMP-5M.

Based on several reports in the literature, dehalogenation of 1-CB in many organisms including *Rhodococcus* has been demonstrated to be a hydrolytic reaction catalysed by a class of enzyme known as haloalkane dehalogenase (Yokota, Omori et al. 1987, Sallis, Armfield et al. 1990, Curragh, Flynn et al. 1994, Erable, Goubet et al. 2005). 1-CB being one

of the lower chlorinated alkanes can be readily assimilated for growth by many microorganisms.

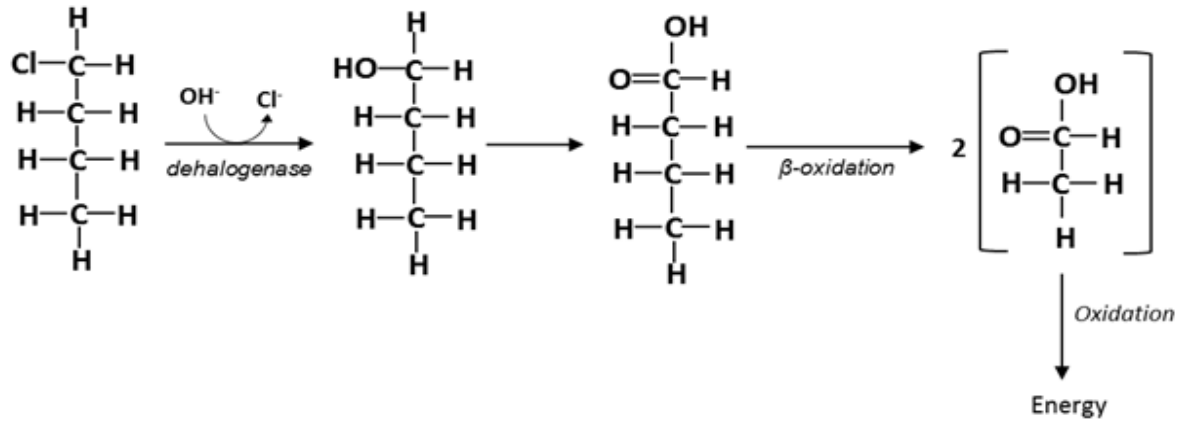


Figure 6.1: Hydrolytic dehalogenation of 1-Chlorobutane (1-CB) and the metabolism of the resulting products via β-oxidation to produce energy.

In contrast to these reports, the ability of *R*. UKMP-5M to dehalogenate 1-CB without growing on it was a notable observation that merited further research. The behaviour of *R*. UKMP-5M in being able to utilise 1-butanol as growth substrate but not able to grow on 1-CB, despite dehalogenating it to release chloride suggested that it has a metabolic pathway different from Figure 6.1. Dehalogenation of 1-CB in *R*. UKMP-5M was found to be constitutive as chloride release was observed even with glucose grown cells. This activity was further induced approximately four fold when using cells which were preexposed to 1-CB during growth (Figure 2.6). On the other hand, LB was identified as a repressor for dehalogenation of 1-CB in *R*. UKMP-5M. Cells which were grown in LB were not able to dehalogenate 1-CB but the cells grown in LB supplemented with 1-CB did release chloride. Moreover, the strict oxygen requirement for the dehalogenation of 1-CB in *R*. UKMP-5M (Figure 2.10A and B) further provides evidence that the metabolism of 1-CB is unlikely to take place via the hydrolytic mechanism, but may involve enzyme(s) that require molecular oxygen, such as oxygenases.

6.1.2 Protein expression analysis using proteomics approach (Chapter 3)

The traditional approach for identification of a specific enzyme is by isolating its activity from a cell free extract, followed by purification using chromatographic methods. Since this approach was unsuccessful with *R*. UKMP-5M, a proteomics method was used to attempt to identify proteins which are responsible for the degradation of 1-CB. The method employed was to compare protein expression patterns between cells exposed to 1-CB and those not. Consistent with the results obtained from the oxygen dependence test, analysis of protein samples extracted from the exposed and unexposed cells revealed that two monooxygenases namely monooxygenase hydroxylase and alkane 1-monooxygenase were present in the former samples but not the latter (Figure 3.5, Tables 3.3 and 3.4). The identification of another monooxygenase subunit (monooxygenase reductase) and several other proteins which belong to the same operon as monooxygenase hydroxylase (Figure 3.10 and Table 3.6) are consistent with dehalogenation of 1-CB being an oxygenase-mediated reaction.

6.1.3 A study on the role of a protein that was identified as amidohydrolase (Chapter 4 and 5)

In addition to the monooxygenases, an uncharacterised protein which was later identified as amidohydrolase (Ah) was also induced with the exposure to 1-CB and the role of this protein is obscure. There is only one example of a member of the Ah superfamily that dehalogenates haloorganic compounds (Govantes, Porrua et al. 2009). With the growing list of members of Ah superfamily to date, many of them remained uncharacterised with unknown function. Hence, in order to study the possible role of this protein in the degradation of 1-CB and in *R*. UKMP-5M as a whole, Ah was cloned and expressed in *E. coli*.

The heterologous expression of Ah in *E. coli* gave rise to inclusion bodies. Multiple attempts to obtain a soluble Ah were unsuccessful (Chapter 4) and therefore, on-column protein refolding was done to obtain a biologically active Ah (Chapter 5). The predicted 3D structure for this protein suggested that it contains a mononuclear metallic center which most probably binds a single Zn^{2+} ion and is coordinated by two histidines and one aspartate molecule (Figure 5.11). Consistent with this prediction, highest Ah activity was observed when Zn^{2+} was added into the assay buffer when compared to four other metals tested (Figure 5.16) and the absence of Zn^{2+} during the refolding of Ah may have resulted in only partial recovery of catalytic activity. Nevertheless, the predicted 3D model and the ability of the refolded Ah to deaminate urea were the preliminary steps to characterise a protein with an unknown function in *R*. UKMP-5M. However, the inability of the refolded protein to release chloride from 1-CB (Figure 5.17) suggests that this protein is probably not responsible for the degradation of 1-CB in *R*. UKMP-5M and the reason for its induction is as a result of coexpression of neighbouring genes.

6.2 Future Perspective

With regard to 1-CB metabolism in *R*. UKMP-5M, there are several questions remain to be addressed. Some potential future work to unravel the complete mechanism is briefly discussed in this section.

6.2.1 Transcript analysis using qRT-PCR and *Moh* gene knockout experiment

The identification of monooxygenase hydroxylase (Moh) and other proteins encoded by genes from the same gene cluster in the induced protein samples of *R*. UKMP-5M and the strict oxygen requirement for the dehalogenation of 1-CB provides a significant piece of evidence for the metabolic pathway elucidation. For example, the activity of a methane monooxygenase (MMO) on an alkane involves either its hydroxylation or epoxidation through activation of molecular oxygen. The catalytic activity of MMO typically requires the

interaction of multisubunit hydroxylase containing diiron centers and a coupling protein (Coufal, Blazyk et al. 2000).

The absence of Moh in the uninduced protein sample of *R*. UKMP-5M (protein extracted from cells which are not able to dehalogenate 1-CB) but the presence of its transcript identified through RT-PCR experiment is not understood. Based on Figure 3.9 (Chapter 3), there is a difference in the level of mRNA expression between the induced samples and the uninduced samples. To better quantify the differential expression of genes encoding for Moh and other proteins identified in the proteomics study, transcript analysis using quantitative real time PCR (qRT-PCR) should be conducted. Furthermore, to confirm that the Moh is a catalyst for dehalogenation of 1-CB in *R*. UKMP-5M, knockout mutant strains with the deletion of *Moh* should be constructed. It is expected that the deletion of *Moh* in the mutant strain will abolish the ability of *R*. UKMP-5M to dehalogenate 1-CB when compared to wild-type *R*. UKMP-5M. These investigations will provide both transcriptomic and phenotypic evidences for the involvement of Moh in dehalogenation of 1-CB in *R*. UKMP-5M.

6.2.2 Metabolic pathway identification

Clearly investigations are now required to understand and elucidate the metabolic pathway involved in the degradation of 1-CB in *R*. UKMP-5M. As discussed in Chapter 2, multiple attempts to identify products from 1-CB dehalogenation apart from chloride were unsuccessful. Metabolite analyses using nuclear magnetic resonance (NMR) showed high levels of background signals which belong to the residues of glucose (due to inefficient washing of cells before resting cell assay) and the residues of 1-CB. When these backgrounds were eliminated, no signal was observed in the NMR spectrum. The inability of the cells to grow on 1-CB but metabolises it to release chloride with no detectable organic compound (associated with 1-CB) secreted, leads to a speculation that the organic product

may be incorporated within the cell to form adducts with biological material such as DNA, RNA or protein.

The formation of DNA adduct involves the covalent binding of a chemical to a piece of DNA which results in the damage or malfunction of the gene. The potential intermediates of 1-CB degradation (as shown in section 2.3.8, Chapter 2) which can form adducts are either the epoxide compound i.e. 1, 2-epoxybutane (González-Pérez, Gómez-Bombarelli et al. 2012) or the lactones i.e. γ-butyrolactone (Lawrence, McGown et al. 2001). Since, *R*. UKMP-5M is able to metabolise 1,2-epoxybutane as a growth substrate (Figure 2.13, Chapter 2), the focus on the potential product can now be narrowed down to γ-butyrolactone. In most cases, the formation of adducts exerts mutagenic and toxicity effect which in turn may kill the cell. In agreement with this, the metabolism of 1-CB was found to cause a toxicity effect in *R*. UKMP-5M as demonstrated by the 'death curve' in Figure 2.14 (Chapter 2).

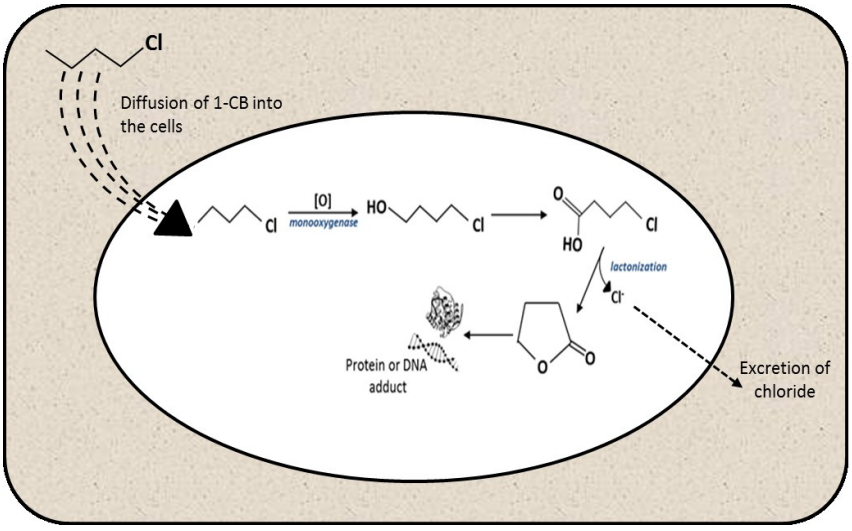


Figure 6.2: A proposed model for the mechanism of 1-Chlorobutane (1-CB) dehalogenation in *R*. UKMP-5M. The first step in the mechanism is catalysed by a monooxygenase that inserts an oxygen atom at the non-halogenated end of 1-CB to form 4-chlorobutanol which will then be subsequently oxidised to 4-chlorobutyrate. Due to non-enzymatic lactonization, the 4-chlorobutyrate will spontaneously release chloride to form γ-butyrolactone. It is speculated that in *R*. UKMP-5M, the γ-butyrolactone covalently binds to the protein or nucleic acid in the cells to form adducts.

In order to test this hypothesis, experiments to detect DNA or protein adducts need to be conducted. A number of techniques have been established for efficient detection of adducts. One simple and precise approach to achieve this is by using radiolabelled compounds coupled with mass spectrometry analysis. In this technique, ^{14}C labelled 1-CB is fed to the cells of *R*. UKMP-5M and after a suitable period of time, the DNA of the cells are extracted and analysed using accelerator mass spectrometry to determine the difference in radiocarbon content between the treated and the untreated samples. This is calculated from the ratios of ^{14}C / ^{12}C or ^{13}C isotope (Brown 2012).

6.2.3 Prediction of function of amidohydrolase (Ah)

The predicted 3D model for Ah in this study has provided some insight into the characterisation of a protein with an unknown function in *R*. UKMP-5M. Many of the members of Ah superfamily which have high sequence similarity with the Ah in this study remained uncharacterised. Given that the binding site predicted for this protein contains a mononuclear metal center for Zn^{2+} ion which is coordinated by two histidines and one aspartate amino acid residues, the inclusion body refolding process should now be performed in the presence of Zn^{2+} to obtain a higher recovery of its native-like conformation. In addition, the effect of metal chelators can also be investigated to determine the metal dependence of the activity in refolded Ah since some Ahs have been reported to be metal-independent (Kwon, Lin et al. 1997, Liu and Huo 2001). Activity assay can also be performed with other metals such as nickel, iron, manganese and cobalt to determine the cofactor specificity of this protein.

Such investigations are expected to yield valuable insights in the field of microbial dehalogenation and bioremediation as a whole.

APPENDICES

APPENDIX I

LIST OF PROTEINS IDENTIFIED IN THE UNINDUCED EXTRACT OF *R.* UKMP-5M

Accession Number	Protein Name	MW (Da)	Match score (%)
RHA1_ro06162	DNA-directed RNA polymerase alpha subunit	38	100
RHA1_ro02146	60kDa chaperonin GroEL	57	100
RHA1_ro02497	alcohol dehydrogenase	39	100
RHA1_ro01487	threonine synthase	37	100
	Cluster of heat shock protein Hsp70	66	100
RHA1_ro05497	heat shock protein Hsp70	66	100
RHA1_ro06936	heat shock protein Hsp70	66	100
RHA1_ro05100	acyl-CoA dehydrogenase	41	100
RHA1_ro06488	ketol-acid reductoisomerase	36	100
RHA1_ro05536	fructose-bisphosphate aldolase	37	100
RHA1_ro07158	methionine adenosyltransferase	43	100
RHA1_ro01472	H(+)-transporting two-sector ATPase beta subunit	52	100
RHA1_ro01921	elongation factor EF1A	44	100
RHA1_ro00979	30S ribosomal protein S1	54	100
RHA1_ro06190	chaperone protein	56	100
RHA1_ro04292	aspartate-semialdehyde dehydrogenase	36	100
RHA1_ro01080	cell division initiation protein	30	100
RHA1_ro06463	electron transfer flavoprotein alpha-subunit	32	100
RHA1_ro07177	glyceraldehyde 3-phosphate dehydrogenase	36	100
RHA1_ro06578	30S ribosomal protein S2	32	100
RHA1_ro01455	acetyl-CoA C-acetyltransferase	41	100
RHA1_ro04998	citrate (Si)-synthase	48	100
RHA1_ro06285	thiosulfate sulfurtransferase	33	100
RHA1_ro01875	acyl-CoA dehydrogenase	41	100
RHA1_ro06490	3-isopropylmalate dehydrogenase	35	100
RHA1_ro07142	3-dehydroquinate synthase	41	100
RHA1_ro06097	fatty acid CoA dehydrogenase	41	100
RHA1_ro01922	elongation factor G	77	100
RHA1_ro05962	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase	33	100
RHA1_ro07207	aconitate hydratase	101	100
RHA1_ro05865	fructose-bisphosphatase, class II	36	100
RHA1_ro01085	cell division protein, FtsZ	41	100
RHA1_ro05841	probable beta-ketoadipyl CoA thiolase	42	100
RHA1_ro01473	H(+)-transporting two-sector ATPase gamma subunit	35	100

RHA1_ro01483	peptide chain release factor RF1	39	100
RHA1_ro05777	phosphopyruvate hydratase	45	100
RHA1_ro07178	phosphoglycerate kinase	42	100
RHA1_ro06238	isocitrate dehydrogenase (NADP+)	45	100
	Cluster of malate dehydrogenase (oxaloacetate decarboxylating)		
RHA1_ro06000	malate dehydrogenase	41	100
RHA1_ro08928	malate dehydrogenase	42	97
RHA1_ro01247	heat-inducible transcription repressor HrcA	37	100
RHA1_ro06136	50S ribosomal protein L2	30	100
RHA1_ro06200	possible IMP dehydrogenase/GMP reductase	39	100
RHA1_ro06442	ribonucleotide reductase	37	100
RHA1_ro06765	recombinase A	37	100
RHA1_ro03390	possible butyryl-CoA dehydrogenase	41	100
RHA1_ro05986	citrate (pro-3S)-lyase beta subunit	35	100
RHA1_ro06590	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase	40	100
RHA1_ro06139	30S ribosomal protein S3	30	100
RHA1_ro06217	probable ribonucleoside-diphosphate reductase	37	100
RHA1_ro04993	citrate (Si) synthase	44	100
RHA1_ro05573	succinate--CoA ligase beta subunit	41	100
RHA1_ro02104	hypothetical protein	28	100
RHA1_ro01132	probable ubiquinol-cytochrome c reductase iron-sulfur subunit	42	100
RHA1_ro01929	probable acyl-CoA dehydrogenase	42	100
RHA1_ro06501	DNA-binding protein	23	100
RHA1_ro01474	H(+)-transporting two-sector ATPase alpha subunit	58	100
RHA1_ro06073	probable FMN-dependent (S)-2-hydroxy-acid oxidase	42	100
RHA1_ro05879	probable GTP binding protein	39	100
RHA1_ro02122	isocitrate lyase	47	99
RHA1_ro04841	phosphoribosylformylglycinamidine cyclo-ligase	38	100
RHA1_ro02262	probable oxidoreductase	37	100
RHA1_ro01373	FKBP-type bacterial trigger factor	50	100
RHA1_ro05069	O-acetylhomoserine aminocarboxypropyltransferase	46	100
RHA1_ro00971	hypothetical protein	34	100
RHA1_ro01541	probable acyl-CoA dehydrogenase	41	100
RHA1_ro06307	mannose-1-phosphate guanylyltransferase	37	100
RHA1_ro06787	long-chain-acyl-CoA dehydrogenase	42	100
RHA1_ro06796	adenosylmethionine-8-amino-7-oxononanoate transaminase	44	100
RHA1_ro07151	carbamoyl-phosphate synthase small subunit	41	100
RHA1_ro04496	ABC transporter, ATP-binding component	34	100
RHA1_ro04599	probable acetyl-CoA C-acetyltransferase	40	100
RHA1_ro04423	ATP-binding subunit of ATP-dependent Clp protease	94	100
RHA1_ro07148	aspartate carbamoyltransferase	34	100
RHA1_ro02587	mycothiol-dependent formaldehyde dehydrogenase	38	100

RHA1_ro00447	alcohol dehydrogenase	36	100
RHA1_ro03654	probable chromosome partitioning protein	38	100
RHA1_ro02204	probable 5,10-methylenetetrahydromethanopterin reductase	37	100
RHA1_ro02052	hydroxymethylbilane synthase	33	100
RHA1_ro04822	C4-dicarboxylate transporter, DAACS family	48	100
RHA1_ro05045	Na(+):solute symporter, SSF family	57	100
RHA1_ro07176	hypothetical protein	35	100
RHA1_ro06853	probable thiolase	42	100
RHA1_ro01340	acetyl-CoA C-acyltransferase	42	100
RHA1_ro00860	hypothetical protein	34	100
RHA1_ro04777	phosphoribosylamine--glycine ligase	42	100
RHA1_ro01876	acetyl-CoA C-acetyltransferase	41	98
RHA1_ro01538	probable enoyl-CoA hydratase	28	99
RHA1_ro01183	hypothetical protein	40	100
	Cluster of acyl-CoA dehydrogenase		
RHA1_ro03025	acyl-CoA dehydrogenase	45	100
RHA1_ro10128	possible acyl-CoA dehydrogenase, C-terminal	20	100
RHA1_ro02050	porphobilinogen synthase	37	100
RHA1_ro05870	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	36	100
RHA1_ro11166	probable glucose-6-phosphate 1-dehydrogenase	37	100
RHA1_ro02396	glycerol-3-phosphate dehydrogenase (NAD(P)+)	35	100
RHA1_ro03041	4-hydroxyphenylpyruvate dioxygenase	44	100
RHA1_ro04317	probable As(2+)-transporting ATPase	41	100
RHA1_ro00952	ornithine carbamoyltransferase	34	100
RHA1_ro05964	succinyl-diaminopimelate desuccinylase	38	100
RHA1_ro04539	3-HSA hydroxylase, oxygenase	42	100
RHA1_ro07235	probable arginine transport ATPase	36	100
RHA1_ro05843	probable cystathionine beta-synthase	48	100
RHA1_ro03865	catechol 2,3-dioxygenase	40	100
RHA1_ro03951	probable thiolase	41	100
RHA1_ro06894	holliday junction DNA helicase	39	100
RHA1_ro03431	single-strand binding protein	18	100
RHA1_ro01242	phosphate starvation-induced protein	39	100
RHA1_ro00242	possible 3-oxoacyl-[acyl-carrier-protein] reductase	32	100
RHA1_ro01932	probable acyl-CoA dehydrogenase	41	100

APPENDIX II

LIST OF PROTEINS IDENTIFIED IN THE INDUCED EXTRACT OF *R.* UKMP-5M

	Accession Number	Protein Name	MW (Da)	Match score (%)
1	RHA1_ro06162	DNA-directed RNA polymerase alpha subunit	38	100
2	RHA1_ro02146	60kDa chaperonin GroEL	57	100
3	RHA1_ro02497	alcohol dehydrogenase	39	100
4	RHA1_ro01487	threonine synthase	37	100
5		Cluster of heat shock protein Hsp70		
5.1	RHA1_ro05497	heat shock protein Hsp70	66	100
5.2	RHA1_ro06936	heat shock protein Hsp70	66	100
6	RHA1_ro05100	acyl-CoA dehydrogenase	41	100
7	RHA1_ro06488	ketol-acid reductoisomerase	36	100
8	RHA1_ro05536	fructose-bisphosphate aldolase	37	100
9	RHA1_ro07158	methionine adenosyltransferase	43	100
10	RHA1_ro01472	H(+)-transporting two-sector ATPase beta subunit	52	100
11	RHA1_ro01921	elongation factor EF1A	44	100
12	RHA1_ro00979	30S ribosomal protein S1	54	100
13	RHA1_ro06190	chaperone protein	56	100
14	RHA1_ro04292	aspartate-semialdehyde dehydrogenase	36	100
15	RHA1_ro01080	cell division initiation protein	30	100
16	RHA1_ro06463	electron transfer flavoprotein alpha-subunit	32	100
17	RHA1_ro07177	glyceraldehyde 3-phosphate dehydrogenase	36	100
18	RHA1_ro06578	30S ribosomal protein S2	32	100
19	RHA1_ro01455	acetyl-CoA C-acetyltransferase	41	100
20	RHA1_ro04998	citrate (Si)-synthase	48	100
21	RHA1_ro00445	hypothetical protein	40	100
22	RHA1_ro06285	thiosulfate sulfurtransferase	33	100
23	RHA1_ro01875	acyl-CoA dehydrogenase	41	100
24	RHA1_ro06490	3-isopropylmalate dehydrogenase	35	100
25	RHA1_ro07142	3-dehydroquinate synthase	41	100
26	RHA1_ro06097	fatty acid CoA dehydrogenase	41	100
27	RHA1_ro01922	elongation factor G	77	100
28	RHA1_ro05962	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase	33	100
29	RHA1_ro07207	aconitate hydratase	101	100
30	RHA1_ro05865	fructose-bisphosphatase, class II	36	100
31	RHA1_ro01085	cell division protein, FtsZ	41	100
32	RHA1_ro05841	probable beta-ketoadipyl CoA thiolase	42	100
33	RHA1_ro01473	H(+)-transporting two-sector ATPase gamma subunit	35	100
34	RHA1_ro01483	peptide chain release factor RF1	39	100
35	RHA1_ro05777	phosphopyruvate hydratase	45	100

36	RHA1_ro07178	phosphoglycerate kinase	42	100
37	RHA1_ro06238	isocitrate dehydrogenase (NADP+)	45	100
38		Cluster of malate dehydrogenase (oxaloacetate decarboxylating)		
38.1	RHA1_ro06000	malate dehydrogenase	41	100
38.2	RHA1_ro08928	malate dehydrogenase	42	100
39	RHA1_ro01247	heat-inducible transcription repressor HrcA	37	100
40	RHA1_ro06136	50S ribosomal protein L2	30	100
41	RHA1_ro06200	possible IMP dehydrogenase/GMP reductase	39	100
42	RHA1_ro06442	ribonucleotide reductase	37	100
43	RHA1_ro06765	recombinase A	37	100
44	RHA1_ro03390	possible butyryl-CoA dehydrogenase	41	100
45	RHA1_ro05986	citrate (pro-3S)-lyase beta subunit	35	100
46	RHA1_ro06590	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase	40	100
47	RHA1_ro06139	30S ribosomal protein S3	30	100
48	RHA1_ro06217	probable ribonucleoside-diphosphate reductase	37	100
49	RHA1_ro04993	citrate (Si) synthase	44	100
50	RHA1_ro05573	succinate--CoA ligase beta subunit	41	100
51	RHA1_ro02104	hypothetical protein	28	100
52	RHA1_ro01132	probable ubiquinol-cytochrome c reductase iron-sulfur subunit	42	100
53	RHA1_ro01929	probable acyl-CoA dehydrogenase	42	99
54	RHA1_ro06501	DNA-binding protein	23	100
55	RHA1_ro01474	H(+)-transporting two-sector ATPase alpha subunit	58	100
56	RHA1_ro06073	probable FMN-dependent (S)-2-hydroxy-acid oxidase	42	100
57	RHA1_ro05879	probable GTP binding protein	39	100
58	RHA1_ro02122	isocitrate lyase	47	100
59	RHA1_ro04841	phosphoribosylformylglycinamidine cyclo-ligase	38	99
60	RHA1_ro02262	probable oxidoreductase	37	100
61	RHA1_ro01373	FKBP-type bacterial trigger factor	50	100
62	RHA1_ro05069	O-acetylhomoserine aminocarboxypropyltransferase	46	100
63	RHA1_ro00971	hypothetical protein	34	100
64	RHA1_ro01541	probable acyl-CoA dehydrogenase	41	100
65	RHA1_ro06307	mannose-1-phosphate guanylyltransferase	37	100
66	RHA1_ro06787	long-chain-acyl-CoA dehydrogenase	42	100
67	RHA1_ro06796	adenosylmethionine-8-amino-7-oxononanoate transaminase	44	100
68	RHA1_ro07151	carbamoyl-phosphate synthase small subunit	41	100
69	RHA1_ro04496	ABC transporter, ATP-binding component	34	100
70	RHA1_ro04599	probable acetyl-CoA C-acetyltransferase	40	100
71	RHA1_ro04423	ATP-binding subunit of ATP-dependent Clp protease	94	100
72	RHA1_ro02587	mycothiol-dependent formaldehyde dehydrogenase	38	100
73	RHA1_ro00443	monooxygenase hydroxylase	42	100
74	RHA1_ro00447	alcohol dehydrogenase	36	100
75	RHA1_ro03447	possible myo-inositol-1-phosphate synthase	39	100

76	RHA1_ro03654	probable chromosome partitioning protein	38	99
77	RHA1_ro02204	probable 5,10-methylenetetrahydromethanopterin reductase	37	100
78	RHA1_ro02052	hydroxymethylbilane synthase	33	100
79	RHA1_ro04822	C4-dicarboxylate transporter, DAACS family	48	100
80	RHA1_ro05045	Na(+):solute symporter, SSF family	57	100
81	RHA1_ro06853	probable thiolase	42	100
82	RHA1_ro00860	hypothetical protein	34	100
83	RHA1_ro04777	phosphoribosylamine--glycine ligase	42	99
84	RHA1_ro01876	acetyl-CoA C-acetyltransferase	41	100
85	RHA1_ro01538	probable enoyl-CoA hydratase	28	100
86	RHA1_ro01183	hypothetical protein	40	100
87	RHA1_ro02050	porphobilinogen synthase	37	95
88	RHA1_ro02534	alkane 1-monooxygenase	47	100
89	RHA1_ro05870	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	36	99
90	RHA1_ro11166	probable glucose-6-phosphate 1-dehydrogenase	37	100
91	RHA1_ro02396	glycerol-3-phosphate dehydrogenase (NAD(P)+)	35	99
92	RHA1_ro03041	4-hydroxyphenylpyruvate dioxygenase	44	100
93	RHA1_ro04317	probable As(2+)-transporting ATPase	41	100
94	RHA1_ro00952	ornithine carbamoyltransferase	34	100
95	RHA1_ro05964	succinyl-diaminopimelate desuccinylase	38	100
96	RHA1_ro06757	possible phage shock protein	29	100
97	RHA1_ro03951	probable thiolase	41	99
98	RHA1_ro06894	holliday junction DNA helicase	39	99
99	RHA1_ro03431	single-strand binding protein	18	99
100	RHA1_ro10159	possible 5-oxoprolinase	86	100

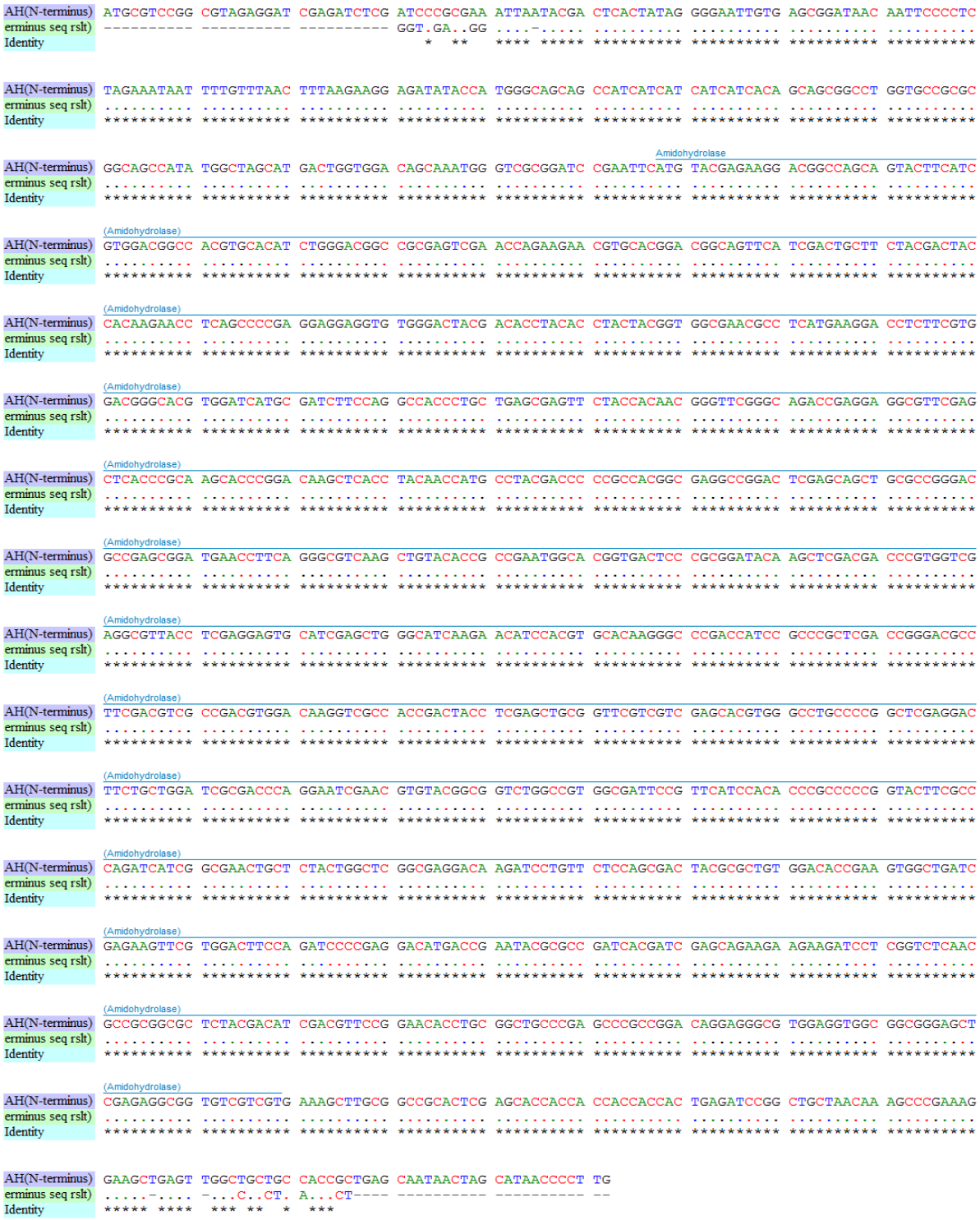
APPENDIX III

DNA AND AMINO ACID SEQUENCE OF AMIDOHYDROLASE

```
     M  Y  E  K  D  G  Q  Q  Y  F  I  V  D  G  H  V  H  I  W  D  G  R  E  S  N  Q  K  N  V  H  G  R  Q  F
0001 ATGTACGAGA AGGACGGCCA GCAGTACTTC ATCGTGGACG GCCACGTGCA CATCTGGGAC GGCCGCGAGT CGAACCAGAA GAACGTGCAC GGACGGCAGT

      I  D  C  F  Y  D  Y  H  K  N  L  S  P  E  E  E  V  W  D  Y  D  T  Y  T  Y  Y  G  G  E  R  L  M  K
0101 TCATCGACTG CTTCTACGAC TACCACAAGA ACCTCAGCCC CGAGGAGGAG GTGTGGGACT ACGACACCTA CACCTACTAC GGTGGCGAAC GCCTCATGAA

      D  L  F  V  D  G  H  V  D  H  A  I  F  Q  A  T  L  L  S  E  F  Y  H  N  G  F  G  Q  T  E  E  A  F
0201 GGACCTCTTC GTGGACGGGC ACGTGGATCA TGCGATCTTC CAGGCCACCC TGCTGAGCGA GTTCTACCAC AACGGGTTCG GGCAGACCGA GGAGGCGTTC

     E  L  T  R  K  H  P  D  K  L  T  Y  N  H  A  Y  D  P  R  H  G  E  A  G  L  E  Q  L  R  R  D  A  E  R
0301 GAGCTCACCC GCAAGCACCC GGACAAGCTC ACCTACAACC ATGCCTACGA CCCCCGCCAC GGCGAGGCCG GACTCGAGCA GCTGCGCCGG GACGCCGAGC

      M  N  L  Q  G  V  K  L  Y  T  A  E  W  H  G  D  S  R  G  Y  K  L  D  D  P  W  S  R  R  Y  L  E  E
0401 GGATGAACCT TCAGGGCGTC AAGCTGTACA CCGCCGAATG GCACGGTGAC TCCCGCGGAT ACAAGCTCGA CGACCCGTGG TCGAGGCGTT ACCTCGAGGA

      C  I  E  L  G  I  K  N  I  H  V  H  K  G  P  T  I  R  P  L  D  R  D  A  F  D  V  A  D  V  D  K  V
0501 GTGCATCGAG CTGGGCATCA AGAACATCCA CGTGCACAAG GGCCCGACCA TCCGCCCGCT CGACCGGGAC GCCTTCGACG TCGCCGACGT GGACAAGGTC

     A  T  D  Y  L  E  L  R  F  V  V  E  H  V  G  L  P  R  L  E  D  F  C  W  I  A  T  Q  E  S  N  V  Y  G
0601 GCCACCGACT ACCTCGAGCT GCGGTTCGTC GTCGAGCACG TGGGCCTGCC CCGGCTCGAG GACTTCTGCT GGATCGCGAC CCAGGAATCG AACGTGTACG

      G  L  A  V  A  I  P  F  I  H  T  R  P  R  Y  F  A  Q  I  I  G  E  L  L  Y  W  L  G  E  D  K  I  L
0701 GCGGTCTGGC CGTGGCGATT CCGTTCATCC ACACCCGCCC CCGGTACTTC GCCCAGATCA TCGGCGAACT GCTCTACTGG CTCGGCGAGG ACAAGATCCT

      F  S  S  D  Y  A  L  W  T  P  K  W  L  I  E  K  F  V  D  F  Q  I  P  E  D  M  T  E  Y  A  P  I  T
0801 GTTCTCCAGC GACTACGCGC TGTGGACACC GAAGTGGCTG ATCGAGAAGT TCGTGGACTT CCAGATCCCC GAGGACATGA CCGAATACGC GCCGATCACG

     I  E  Q  K  K  K  I  L  G  L  N  A  A  A  L  Y  D  I  D  V  P  E  H  L  R  L  P  E  P  A  G  Q  E  G
0901 ATCGAGCAGA AGAAGAAGAT CCTCGGTCTC AACGCCGCGG CGCTCTACGA CATCGACGTT CCGGAACACC TGCGGCTGCC CGAGCCCGCC GGACAGGAGG

      V  E  V  A  A  G  A  R  E  A  V  S  S
1001 GCGTGGAGGT GGCGGCGGGA GCTCGAGAGG CGGTGTCGTC G
```

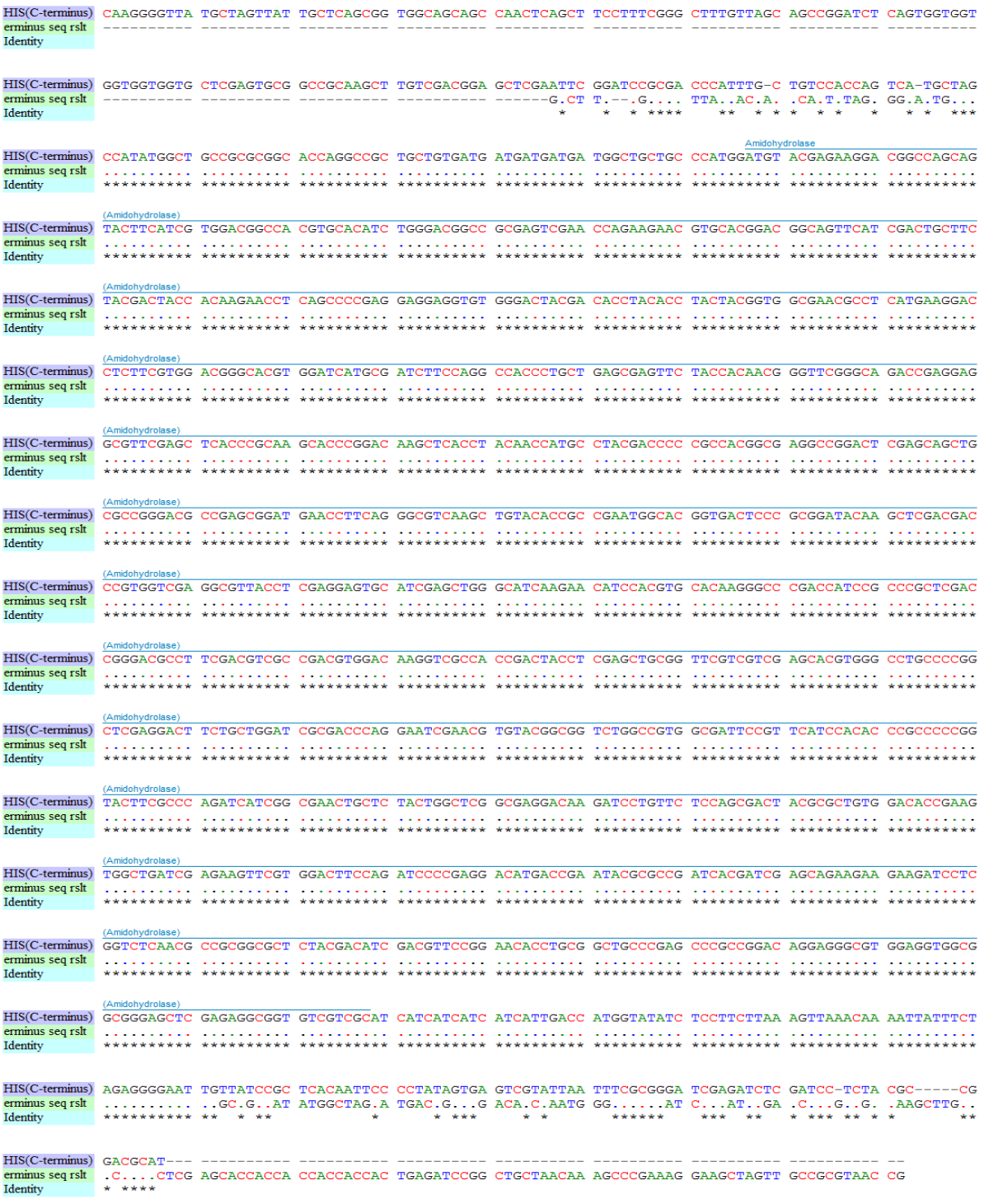
APPENDIX IV

RESULT OF DNA SEQUENCING OF AMIDOHYDROLASE WITH HISTIDINE (HIS) TAG AT THE N-TERMINUS (pET-28a_AHntc)



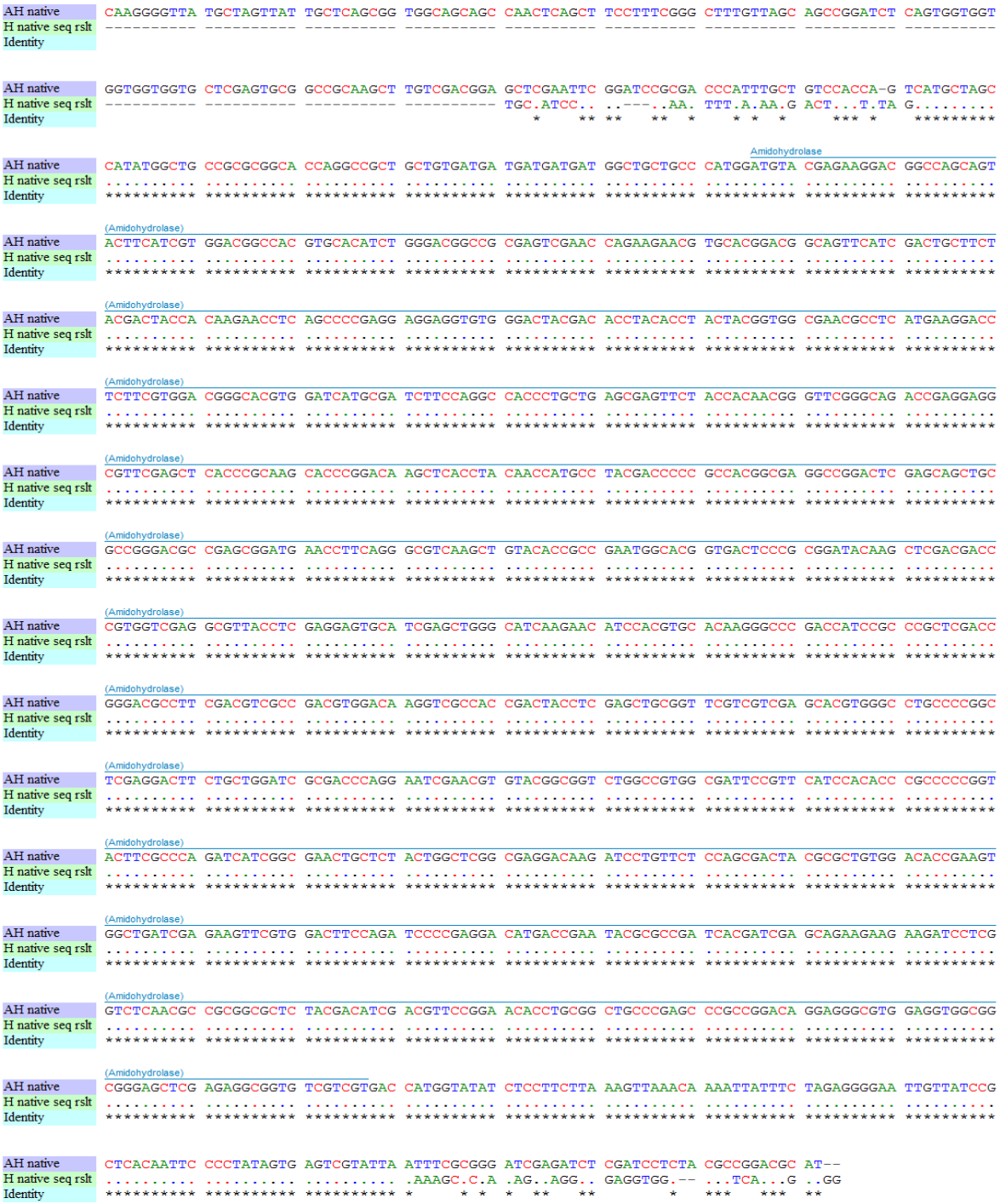
APPENDIX V

RESULT OF DNA SEQUENCING OF AMIDOHYDROLASE WITH HISTIDINE (HIS) TAG AT THE C-TERMINUS (pET-28a_AHctc)



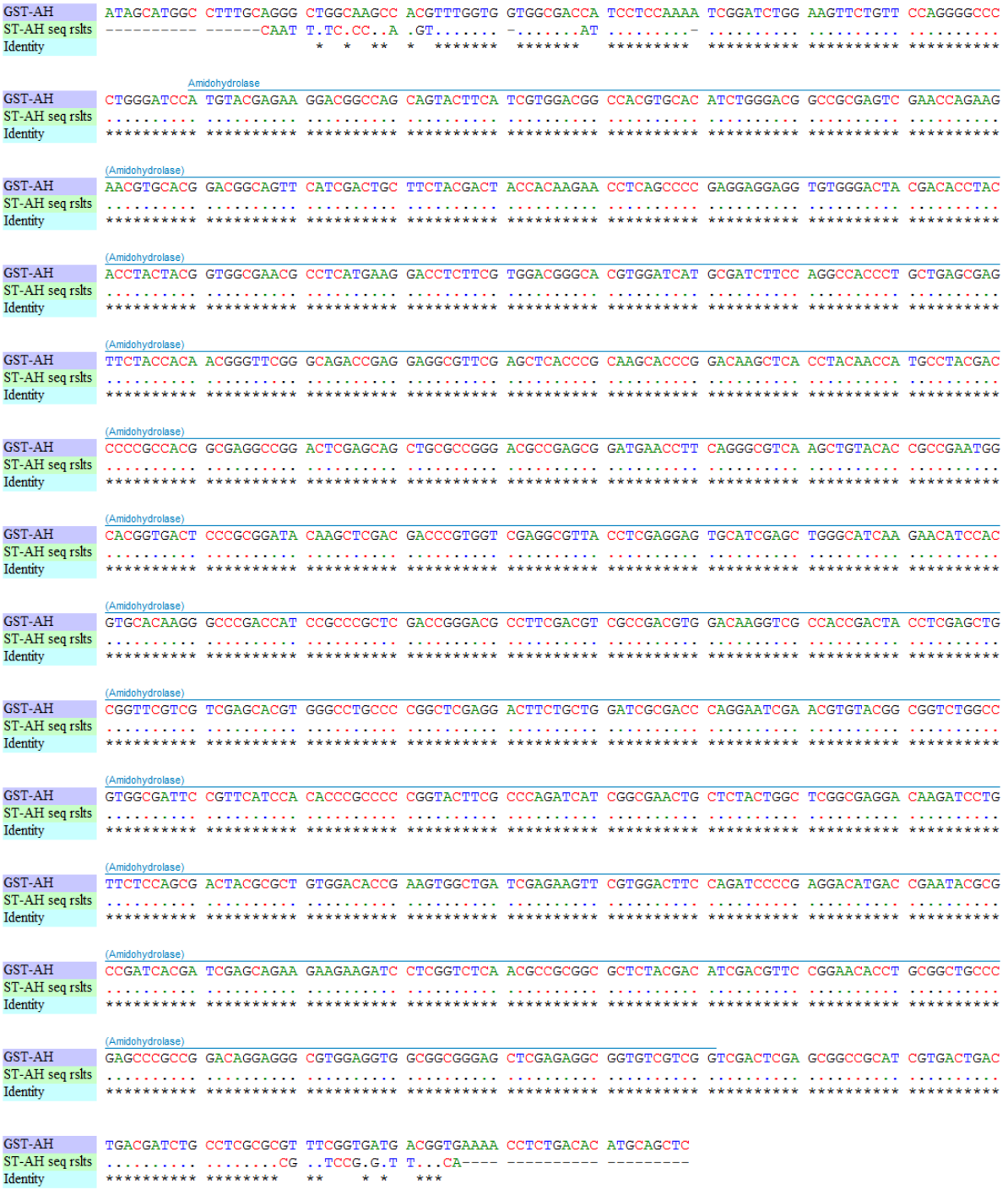
APPENDIX VI

RESULT OF DNA SEQUENCING OF NATIVE AMIDOHYDROLASE (WITHOUT FUSION TAG) (pET-28a_AHnat)



APPENDIX VII

RESULT OF DNA SEQUENCING OF AMIDOHYDROLASE WITH GLUTHATHIONE S-TRANSFERASE (GST) TAG AT THE N-TERMINUS (pGEX-6P-1_AH)



REFERENCES

1. Abdallah, C., E. Dumas-Gaudot, J. Renaut and K. Sergeant (2012). "Gel-Based and Gel-Free Quantitative Proteomics Approaches at a Glance." International Journal of Plant Genomics **2012**: 17.

2. Adachi, O., Y. Akakabe, H. Toyama and K. Matsushita (2006). "Growth promoting substance in yeast extract for methylotrophic growth of Candida boidinii." Biosci Biotechnol Biochem **70**(8): 2007-2009.

3. Ahmad, M., J. N. Roberts, E. M. Hardiman, R. Singh, L. D. Eltis and T. D. H. Bugg (2011). "Identification of DypB from Rhodococcus jostii RHA1 as a Lignin Peroxidase." **50**(23): 5096-5107.

4. Alvarez, H. M. (2010). Biology of Rhodococcus: Microbiology Monographs 16. DE, Springer Verlag.

5. Anandarajah, K., P. M. Kiefer, Jr., B. S. Donohoe and S. D. Copley (2000). "Recruitment of a double bond isomerase to serve as a reductive dehalogenase during biodegradation of pentachlorophenol." Biochemistry **39**(18): 5303-5311.

6. Araki, N., Y. Niikura, K. Miyauchi, D. Kasai, E. Masai and M. Fukuda (2011). "Glucose-mediated transcriptional repression of PCB/biphenyl catabolic genes in Rhodococcus jostii RHA1." J Mol Microbiol Biotechnol **20**(1): 53-62.

7. Armfield, S., P. Sallis, P. Baker, A. Bull and D. Hardman (1995). "Dehalogenation of haloalkanes by Rhodococcus erythropolis Y2." Biodegradation **6**(3): 237-246.

8. Arnold, K., L. Bordoli, J. Kopp and T. Schwede (2006). "The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling." Bioinformatics **22**(2): 195-201.

9. Baptista, I. I. R. (2008). Biomolecular approach to the study of microbial dynamics during biodegradation of halogenated compounds. PhD thesis, Imperial College London.

10. Basu, A., X. Li and S. Leong (2011). "Refolding of proteins from inclusion bodies: rational design and recipes." Applied Microbiology and Biotechnology **92**(2): 241-251.

11. Bell, K. S., J. C. Philp, D. W. Aw and N. Christofi (1998). "The genus Rhodococcus." J Appl Microbiol **85**(2): 195-210.

12. Benini, S., W. R. Rypniewski, K. S. Wilson, S. Miletti, S. Ciurli and S. Mangani (1999). "A new proposal for urease mechanism based on the crystal structures of the native and inhibited enzyme from Bacillus pasteurii: why urea hydrolysis costs two nickels." Structure **7**(2): 205-216.

13. Bergmann, J. G. and J. Sanik (1957). "Determination of Trace Amounts of Chlorine in Naphtha." **29**(2): 241-243.

14. Bhatt, Praveena, K. Suresh, M, Mudliar, Sandeep, Chakrabarti and G. Tapan (2007). Biodegradation of chlorinated compounds: A review. Philadelphia, PA, ETATS-UNIS, Taylor & Francis. **37:** 34.

15. Bondar, V. S., M. G. Boersma, W. J. van Berkel, Z. I. Finkelstein, E. L. Golovlev, B. P. Baskunov, J. Vervoort, L. A. Golovleva and I. M. Rietjens (1999). "Preferential oxidative dehalogenation upon conversion of 2-halophenols by Rhodococcus opacus 1G." FEMS Microbiol Lett **181**(1): 73-82.

16. Brown, K. (2012). "Methods for the detection of DNA adducts." Methods Mol Biol **817**: 207-230.

17. Burgess, R. R. (2009). Chapter 17 Refolding Solubilized Inclusion Body Proteins. Methods in Enzymology. R. B. Richard and P. D. Murray, Academic Press. **Volume 463:** 259-282.

18. Cappelletti, M., D. Frascari, D. Zannoni and S. Fedi (2012). "Microbial degradation of chloroform." Appl Microbiol Biotechnol **96**(6): 1395-1409.

19. Carter, E. L., D. E. Tronrud, S. R. Taber, P. A. Karplus and R. P. Hausinger (2011). "Iron-containing urease in a pathogenic bacterium." Proceedings of the National Academy of Sciences **108**(32): 13095-13099.

20. Chambers, M. C., B. Maclean, R. Burke, D. Amodei, D. L. Ruderman, S. Neumann, L. Gatto, B. Fischer, B. Pratt, J. Egertson, K. Hoff, D. Kessner, N. Tasman, N. Shulman, B. Frewen, T. A. Baker, M.-Y. Brusniak, C. Paulse, D. Creasy, L. Flashner, K. Kani, C. Moulding, S. L. Seymour, L. M. Nuwaysir, B. Lefebvre, F. Kuhlmann, J. Roark, P. Rainer, S. Detlev, T. Hemenway, A. Huhmer, J. Langridge, B. Connolly, T. Chadick, K. Holly, J. Eckels, E. W. Deutsch, R. L. Moritz, J. E. Katz, D. B. Agus, M. MacCoss, D. L. Tabb and P. Mallick (2012). "A cross-platform toolkit for mass spectrometry and proteomics." Nat Biotech **30**(10): 918-920.

21. Chauhan, A. and R. K. Jain (2010). "Biodegradation: gaining insight through proteomics." Biodegradation **21**(6): 861-879.

22. Chen, D. Z., D. J. Ouyang, H. X. Liu, J. Chen, Q. F. Zhuang and J. M. Chen (2014). "Effective utilization of dichloromethane by a newly isolated strain Methylobacterium rhodesianum H13." Environ Sci Pollut Res Int **21**(2): 1010-1019.

23. Chen, H.-M., K.-Y. Lin and C.-H. Lu (2005). "Refolding and activation of recombinant N-carbamoyl-d-amino acid amidohydrolase from Escherichia coli inclusion bodies." Process Biochemistry **40**(6): 2135-2141.

24. Chen, H. and H. X. Zhou (2005). "Prediction of solvent accessibility and sites of deleterious mutations from protein sequence." Nucleic Acids Res **33**(10): 3193-3199.

25. Cheng, Z., Y. Y. Wei, W. W. Sung, B. R. Glick and B. J. McConkey (2009). Proteomic analysis of the response of the plant growth-promoting bacterium Pseudomonas putida UW4 to nickel stress. Proteome Sci. England. **7:** 18.

26. Coufal, D. E., J. L. Blazyk, D. A. Whittington, W. W. Wu, A. C. Rosenzweig and S. J. Lippard (2000). "Sequencing and analysis of the Mmethylococcus capsulatus (Bath) solublemethane monooxygenase genes." Eur J Biochem **267**(8): 2174-2185.

27. Criddle, C. S. (1993). "The kinetics of cometabolism." Biotechnol Bioeng **41**(11): 1048-1056.

28. Curragh, H., O. Flynn, M. J. Larkin, T. M. Stafford, J. T. Hamilton and D. B. Harper (1994). "Haloalkane degradation and assimilation by Rhodococcus rhodochrous NCIMB 13064." Microbiology **140 (Pt 6)**: 1433-1442.

29. Dae-Wook Park, S.-S. K., Min-Kyung Nam (2011). Improved recovery of active GST-fusion proteins from insoluble aggregates: solubilization and purification conditions using PKM2 and HtrA2 as model proteins. Korea, Sogang University**:** 279-284.

30. de Jong, R. M., J. J. W. Tiesinga, H. J. Rozeboom, K. H. Kalk, L. Tang, D. B. Janssen and B. W. Dijkstra (2003). "Structure and mechanism of a bacterial haloalcohol dehalogenase: a new variation of the short-chain dehydrogenase/reductase fold without an NAD(P)H binding site." The EMBO Journal **22**(19): 4933-4944.

31. Dhingra, V., M. Gupta, T. Andacht and Z. F. Fu (2005). "New frontiers in proteomics research: A perspective." International Journal of Pharmaceutics **299**(1-2): 1-18.

32. Dick B. Janssen, I. J. T. D., Gerrit J. Poelarends and Peter Terpstra (2005). "Bacterial degradation of xenobiotic compounds: evolution and distribution of novel enzyme activities." Environmental Microbiology **7**(12): 1868–1882

33. Dua, M., A. Singh, N. Sethunathan and A. K. Johri (2002). "Biotechnology and bioremediation: successes and limitations." Appl Microbiol Biotechnol **59**(2-3): 143-152.

34. Duque, A. F., S. A. Hasan, V. S. Bessa, M. F. Carvalho, G. Samin, D. B. Janssen and P. M. Castro (2012). "Isolation and characterization of a Rhodococcus strain able to degrade 2-fluorophenol." Appl Microbiol Biotechnol **95**(2): 511-520.

35. Eguchi, A., K. Nomiyama, G. Devanathan, A. Subramanian, K. A. Bulbule, P. Parthasarathy, S. Takahashi and S. Tanabe (2012). "Different profiles of anthropogenic and naturally produced organohalogen compounds in serum from residents living near a coastal area and e-waste recycling workers in India." Environment International **47**(0): 8-16.

36. Emanuelsson, M. A., M. B. Osuna, J. Sipma and P. M. Castro (2008). "Treatment of halogenated organic compounds and monitoring of microbial dynamics in up-flow fixed bed reactors under sequentially alternating pollutant scenarios." Biotechnol Bioeng **99**(4): 800-810.

37. EPA, U. S. (2013). "Introduction to in situ bioremediation of groundwater." Office of Solid Waste and Emergency Response: 86.

38. Erable, B., I. Goubet, S. Lamare, A. Seltana, M. D. Legoy and T. Maugard (2005). "Nonconventional hydrolytic dehalogenation of 1-chlorobutane by dehydrated bacteria in a continuous solid-gas biofilter." Biotechnol Bioeng **91**(3): 304-313.

39. Estapé, D. and U. Rinas (1996). "Optimized procedures for purification and solubilization of basic fibroblast growth factor inclusion bodies." Biotechnology Techniques **10**(7): 481-484.

40. Fan, H., D. S. Hitchcock, R. D. Seidel, 2nd, B. Hillerich, H. Lin, S. C. Almo, A. Sali, B. K. Shoichet and F. M. Raushel (2013). "Assignment of pterin deaminase activity to an enzyme of unknown function guided by homology modeling and docking." J Am Chem Soc **135**(2): 795-803.

41. Faroon, O., J. Taylor, N. Roney, E. M. Fransen, S. Bogaczyk and G. Diamond (2005). Toxicological Profile for Carbon Tetrachloride, U.S. Department of Health and Human Services**:** 1-309.

42. Fawcett, J. K. and J. E. Scott (1960). "A rapid and precise method for the determination of urea." J Clin Pathol **13**: 156-159.

43. Fetzner, S. (1998). "Bacterial dehalogenation." Appl Microbiol Biotechnol **50**(6): 633-657.

44. Fetzner, S. and F. Lingens (1994). "Bacterial Dehalogenases: Biochemistry, Genetics, and Biotechnological Applications." Microbiological Reviews **58**(4): 641-685.

45. Field, J. A. and R. Sierra-Alvarez (2004). "Biodegradability of chlorinated solvents and related chlorinated aliphatic compounds." Reviews in Environmental Science and Biotechnology **3**(3): 185-254.

46. Finnerty, W. R. (1992). "The biology and genetics of the genus Rhodococcus." Annu Rev Microbiol **46**: 193-218.

47. Frascari, D., D. Pinelli, M. Nocentini, S. Fedi, Y. Pii and D. Zannoni (2006). "Chloroform degradation by butane-grown cells of Rhodococcus aetherovorans BCP1." Appl Microbiol Biotechnol **73**(2): 421-428.

48. Fulekar, M. H. and J. Sharma (2008). "Bioinformatics Applied in Bioremediation." Innovative Romanian Food Biotechnology **2**(2): 28-36.

49. Fulthorpe, R. R. and E. M. Top (2010). Evolution of New Catabolic Functions Through Gene Assembly by Mobile Genetic Elements. Handbook of Hydrocarbon and Lipid Microbiology. K. Timmis, Springer Berlin Heidelberg**:** 1219-1233.

50. Ganguly, S. (2005). Enhanced Stabilisation of Nitrile Hydratase Enzyme from Rhodococcus sp. DAP 96253 and R. Rhodochrous DAP 96622. PhD, Georgia State University.

51. García-Fruitós, E., E. Vázquez, C. Díez-Gil, J. L. Corchero, J. Seras-Franzoso, I. Ratera, J. Veciana and A. Villaverde (2012). "Bacterial inclusion bodies: making gold from waste." Trends in Biotechnology **30**(2): 65-70.

52. Gavanji, S. (2013). Aplication of Recombinant DNA Technology- A review. Applied Science Reports, Young Researchers Club, Khorasgan Branch, Islamic Azad University, Khorasgan, Isfahan, Iran. **2:** 29-31.

53. Ginalski, K. (2006). "Comparative modeling for protein structure prediction." Current Opinion in Structural Biology **16**(2): 172-177.

54. González-Pérez, M., R. Gómez-Bombarelli, J. Arenas-Valgañón, M. T. Pérez-Prior, M. P. García-Santos, E. Calle and J. Casado (2012). "Connecting the Chemical and Biological Reactivity of Epoxides." Chemical Research in Toxicology **25**(12): 2755-2762.

55. Govantes, F., O. Porrua, V. Garcia-Gonzalez and E. Santero (2009). "Atrazine biodegradation in the lab and in the field: enzymatic activities and gene regulation." Microb Biotechnol **2**(2): 178-185.

56. Govender, A. (2008). Characterisation of Chlorinated-Hydrocarbon-Degrading Genes of Bacteria. PhD, University of KwaZulu-Natal.

57. Graslund, S., P. Nordlund, J. Weigelt, B. M. Hallberg, J. Bray, O. Gileadi, S. Knapp, U. Oppermann, C. Arrowsmith, R. Hui, J. Ming, S. dhe-Paganon, H. W. Park, A. Savchenko, A. Yee, A. Edwards, R. Vincentelli, C. Cambillau, R. Kim, S. H. Kim, Z. Rao, Y. Shi, T. C. Terwilliger, C. Y. Kim, L. W. Hung, G. S. Waldo, Y. Peleg, S. Albeck, T. Unger, O. Dym, J. Prilusky, J. L. Sussman, R. C. Stevens, S. A. Lesley, I. A. Wilson, A. Joachimiak, F. Collart, I. Dementieva, M. I. Donnelly, W. H. Eschenfeldt, Y. Kim, L. Stols, R. Wu, M. Zhou, S. K. Burley, J. S. Emtage, J. M. Sauder, D. Thompson, K. Bain, J. Luz, T. Gheyi, F. Zhang, S. Atwell, S. C.

Almo, J. B. Bonanno, A. Fiser, S. Swaminathan, F. W. Studier, M. R. Chance, A. Sali, T. B. Acton, R. Xiao, L. Zhao, L. C. Ma, J. F. Hunt, L. Tong, K. Cunningham, M. Inouye, S. Anderson, H. Janjua, R. Shastry, C. K. Ho, D. Wang, H. Wang, M. Jiang, G. T. Montelione, D. I. Stuart, R. J. Owens, S. Daenke, A. Schutz, U. Heinemann, S. Yokoyama, K. Bussow and K. C. Gunsalus (2008). "Protein production and purification." Nat Methods **5**(2): 135-146.

58. Graymore, M., F. Stagnitti and G. Allinson (2001). "Impacts of atrazine in aquatic ecosystems." Environment International **26**(7–8): 483-495.

59. Greenfield, N. J. (2006). "Using circular dichroism spectra to estimate protein secondary structure." Nat Protoc **1**(6): 2876-2890.

60. Gribble, G. W. (1994). "The natural production of chlorinated compounds." Environmental Science & Technology **28**(7): 310A-319A.

61. Gruber, D. F., V. A. Pieribone, B. Porton and H. T. Kao (2008). "Strict regulation of gene expression from a high-copy plasmid utilizing a dual vector system." Protein Expr Purif **60**(1): 53-57.

62. Gu, Z., Z. Su and J.-C. Janson (2001). "Urea gradient size-exclusion chromatography enhanced the yield of lysozyme refolding." Journal of Chromatography A **918**(2): 311-318.

63. Gygi, S. P., Y. Rochon, B. R. Franza and R. Aebersold (1999). "Correlation between protein and mRNA abundance in yeast." Mol Cell Biol **19**(3): 1720-1730.

64. Hage, J. C. and S. Hartmans (1999). "Monooxygenase-mediated 1,2-dichloroethane degradation by Pseudomonas sp. strain DCA1." Appl Environ Microbiol **65**(6): 2466-2470.

65. Hammarstrom, M., N. Hellgren, S. van Den Berg, H. Berglund and T. Hard (2002). "Rapid screening for improved solubility of small human proteins produced as fusion proteins in Escherichia coli." Protein Sci **11**(2): 313-321.

66. Hardiman, G. (2004). "Introduction to Proteomics: tools for the new biotechnology." Expert Review Proteomics **1**(1): 9-10.

67. Harper, S. and D. W. Speicher (2011). "Purification of proteins fused to glutathione S-transferase." Methods Mol Biol **681**: 259-280.

68. He, H. J., Q. S. Yuan, G. Z. Yang and X. F. Wu (2002). "High-level expression of human extracellular superoxide dismutase in Escherichia coli and insect cells." Protein Expr Purif **24**(1): 13-17.

69. Healthcare, GE. (2002). GST Gene Fusion System, GE Healthcare**:** 130.

70. Hernawan, T., S. Soemitro, H. Dewayani, Tetih and I. W.T. (1999). "Isolation and Purification of Dehalogenase from Pseudomonas cepacia UK7WS1 strain." Jurnal Matematika dan Sains **4**(2): 70-82.

71. Hernick, M. and C. A. Fierke (2005). "Zinc hydrolases: the mechanisms of zinc-dependent deacetylases." Archives of Biochemistry and Biophysics **433**(1): 71-84.

72. Holm, L. and C. Sander (1997). "An evolutionary treasure: unification of a broad set of amidohydrolases related to urease." Proteins **28**(1): 72-82.

73. Huang, Z., C. Zhang, S. Chen, F. Ye and X.-H. Xing (2013). "Active inclusion bodies of acid phosphatase PhoC: aggregation induced by GFP fusion and activities modulated by linker flexibility." Microbial Cell Factories **12**(1): 25.

74. Janssen, D. B., D. Jager and B. Witholt (1987). "Degradation of n-haloalkanes and alpha, omega-dihaloalkanes by wild-type and mutants of Acinetobacter sp. strain GJ70." Appl Environ Microbiol **53**(3): 561-566.

75. Janssen, D. B., J. E. Oppentocht and G. J. Poelarends (2001). "Microbial dehalogenation." Curr Opin Biotechnol **12**(3): 254-258.

76. Janssen, D. B., F. Pries and J. R. van der Ploeg (1994). "Genetics and biochemistry of dehalogenating enzymes." Annu Rev Microbiol **48**: 163-191.

77. Jehmlich, N., S. Kleinsteuber, C. Vogt, D. Benndorf, H. Harms, F. Schmidt, M. Von Bergen and J. Seifert (2010). "Phylogenetic and proteomic analysis of an anaerobic toluene-degrading community." J Appl Microbiol **109**(6): 1937-1945.

78. Jing, N. H. (2007). Isolation of Local Bacterial Capable of Degrading Halogenated Compounds and Analysis of Putative Haloacid Permease Gene. Master, Universiti Teknologi Malaysia.

79. Källberg, M., H. Wang, S. Wang, J. Peng, Z. Wang, H. Lu and J. Xu (2012). "Template-based protein structure modeling using the RaptorX web server." Nat. Protocols **7**(8): 1511-1522.

80. Kayser, M. (2001). Genes and Associated with Dichloromethane Metabolism in Methylobacterium Dichloromethanicum DM4. PhD, Swiss Federal Institute of Technology Zurich.

81. Kern, S., D. Riester, C. Hildmann, A. Schwienhorst and F.-J. Meyer-Almes (2007). "Inhibitor-mediated stabilization of the conformational structure of a histone deacetylase-like amidohydrolase." FEBS Journal **274**(14): 3578-3588.

82. Keuning, S., D. B. Janssen and B. Witholt (1985). "Purification and characterization of hydrolytic haloalkane dehalogenase from Xanthobacter autotrophicus GJ10." J Bacteriol **163**(2): 635-639.

83. Kim, D., J. C. Chae, G. J. Zylstra, Y. S. Kim, S. K. Kim, M. H. Nam, Y. M. Kim and E. Kim (2004). "Identification of a novel dioxygenase involved in metabolism of o-xylene, toluene, and ethylbenzene by Rhodococcus sp. strain DK17." Appl Environ Microbiol **70**(12): 7086-7092.

84. Kim, S. J., O. Kweon, R. C. Jones, J. P. Freeman, R. D. Edmondson and C. E. Cerniglia (2007). Complete and integrated pyrene degradation pathway in Mycobacterium vanbaalenii PYR-1 based on systems biology. J Bacteriol. United States. **189:** 464-472.

85. Kim, Y., L. Semprini and D. J. Arp (1997). "Aerobic Cometabolism of Chloroform and 1,1,1-Trichloroethane by Butane-Grown Microorganisms." Bioremediation Journal **1**(2): 135-148.

86. Kitagawa, W., K. Miyauchi, E. Masai and M. Fukuda (2001). "Cloning and Characterization of Benzoate Catabolic Genes in the Gram-Positive Polychlorinated Biphenyl Degrader Rhodococcus sp. Strain RHA1." J. Bacteriol. **183**(22 %U http://jb.asm.org/cgi/content/abstract/183/22/6598 %8 November 15, 2001): 6598-6606.

87. Konc, J., M. Hodoscek, M. Ogrizek, J. Trykowska Konc and D. Janezic (2013). "Structure-based function prediction of uncharacterized protein using binding sites comparison." PLoS Comput Biol **9**(11): e1003341.

88. Kumar, A., B. S. Bisht, V. D. Joshi and T. Dhewa (2011). "Review of Bioremediation of Polluted Environment: A Management Tool." International Journal of Environmental Sciences **1**(6): 1079-1093.

89. Kwon, D. S., C. H. Lin, S. Chen, J. K. Coward, C. T. Walsh and J. M. Bollinger, Jr. (1997). "Dissection of glutathionylspermidine synthetase/amidase from Escherichia coli into autonomously folding and functional synthetase and amidase domains." J Biol Chem **272**(4): 2429-2436.

90. Larkin, M. J., L. A. Kulakov, C. C. R. Allen, M. J. Larkin, L. A. Kulakov and C. C. R. Allen (2010). Rhodococcus. Handbook of Hydrocarbon and Lipid Microbiology. K. N. Timmis and K. N. Timmis, Springerlink**:** 1840-1852.

91. Lawrence, N. J., A. T. McGown, J. Nduka, J. A. Hadfield and R. G. Pritchard (2001). "Cytotoxic Michael-type amine adducts of alpha-methylene lactones alantolactone and isoalantolactone." Bioorg Med Chem Lett **11**(3): 429-431.

92. Lee, M., M. K. Kim, I. Singleton, M. Goodfellow and S. T. Lee (2006). Enhanced biodegradation of diesel oil by a newly identified Rhodococcus baikonurensis EN3 in the presence of mycolic acid. J Appl Microbiol. England. **100:** 325-333.

93. Lee, P. S., L. B. Shaw, L. H. Choe, A. Mehra, V. Hatzimanikatis and K. H. Lee (2003). "Insights into the relation between mrna and protein expression patterns: ii. Experimental observations in Escherichia coli1." Biotechnology and Bioengineering **84**(7): 834-841.

94. Leisinger, T. (1996). Biodegradation of chlorinated aliphatic compounds. Curr Opin Biotechnol. England. **7:** 295-300.

95. Liaw, S. H., S. J. Chen, T. P. Ko, C. S. Hsu, C. J. Chen, A. H. Wang and Y. C. Tsai (2003). "Crystal structure of D-aminoacylase from Alcaligenes faecalis DA1. A novel subset of amidohydrolases and insights into the enzyme mechanism." J Biol Chem **278**(7): 4957-4962.

96. Lim, W. K., J. Rösgen and S. W. Englander (2009). "Urea, but not guanidinium, destabilizes proteins by forming hydrogen bonds to the peptide group." Proceedings of the National Academy of Sciences **106**(8): 2595-2600.

97. Liu, A. and L. Huo (2001). Amidohydrolase Superfamily. eLS, John Wiley & Sons, Ltd.

98. Liu, X., W.-J. Li, L. Li, Y. Yang, L.-G. Mao and Z. Peng (2014). "A label-free electrochemical immunosensor based on gold nanoparticles for direct detection of atrazine." Sensors and Actuators B: Chemical **191**(0): 408-414.

99. Lundstrom, K. H. (2006). Structural Genomics on Membrane Proteins, Taylor & Francis.

100. Maegala Nallapan Maniyam, F. S., and Abdul Latif Ibrahim (2011). "Bioremediation of Cyanide by Optimized Resting Cells of Rhodococcus Strains Isolated from Peninsular Malaysia " International Journal of Bioscience, Biochemistry and Bioinformatics **1**: 98-101.

101. Maiti, R., G. H. Van Domselaar, H. Zhang and D. S. Wishart (2004). "SuperPose: a simple server for sophisticated structural superposition." Nucleic Acids Research **32**(suppl 2): W590-W594.

102. Mallick, P. and B. Kuster (2010). Proteomics: a pragmatic perspective. Nat Biotechnol. United States. **28:** 695-709.

103. Mandel, T., A. J. Fleming, R. Krahenbuhl and C. Kuhlemeier (1995). "Definition of constitutive gene expression in plants: the translation initiation factor 4A gene as a model." Plant Mol Biol **29**(5): 995-1004.

104. Maniyam, M. N. (2013). BIODETOXIFICATION OF CYANIDE BY A MALAYSIAN ISOLATE RHODOCOCCUS UKMP-5M. PhD thesis, Universiti Selangor.

105. Marchler-Bauer, A. and S. H. Bryant (2004). "CD-Search: protein domain annotations on the fly." Nucleic Acids Res **32**(Web Server issue): W327-331.

106. Marek, R. F., P. S. Thorne, K. Wang, J. Dewall and K. C. Hornbuckle (2013). "PCBs and OH-PCBs in serum from children and mothers in urban and rural U.S. communities." Environ Sci Technol **47**(7): 3353-3361.

107. Margesin, R., P. A. Fonteyne and B. Redl (2005). "Low-temperature biodegradation of high amounts of phenol by Rhodococcus spp. and basidiomycetous yeasts." Res Microbiol **156**(1): 68-75.

108. Martinkova, L., B. Uhnakova, M. Patek, J. Nesvera and V. Kren (2009). Biodegradation potential of the genus Rhodococcus. Environ Int. United States. **35:** 162-177.

109. Masai, E., A. Yamada, J. M. Healy, T. Hatta, K. Kimbara, M. Fukuda and K. Yano (1995). "Characterization of biphenyl catabolic genes of gram-positive polychlorinated biphenyl degrader Rhodococcus sp. strain RHA1." Appl Environ Microbiol **61**(6): 2079-2085.

110. Matsubara, T., T. Ohshiro, Y. Nishina and Y. Izumi (2001). "Purification, characterization, and overexpression of flavin reductase involved in dibenzothiophene desulfurization by Rhodococcus erythropolis D-1." Appl Environ Microbiol **67**(3): 1179-1184.

111. Matthew J. Giacalone, A. M. G., Brian T. Lovitt, Neil L. Berkley, and a. M. W. S. Carl W. Gunderson "Toxic protein expression in Escherichia coli using a rhamnose-based tightly regulated and tunable promoter system." BioTechniques **40**: 355-364.

112. McDonald, W. H. and J. R. Yates, 3rd (2002). "Shotgun proteomics and biomarker discovery." Dis Markers **18**(2): 99-105.

113. McLeod, M. P., R. L. Warren, W. W. Hsiao, N. Araki, M. Myhre, C. Fernandes, D. Miyazawa, W. Wong, A. L. Lillquist, D. Wang, M. Dosanjh, H. Hara, A. Petrescu, R. D. Morin, G. Yang, J. M. Stott, J. E. Schein, H. Shin, D. Smailus, A. S. Siddiqui, M. A. Marra, S. J. Jones, R. Holt, F. S. Brinkman, K. Miyauchi, M. Fukuda, J. E. Davies, W. W. Mohn and L. D. Eltis (2006). The complete genome of Rhodococcus sp. RHA1 provides insights into a catabolic powerhouse. Proc Natl Acad Sci U S A. United States. **103:** 15582-15587.

114. Mohajeri, L., M. H. Isa and H. A. Aziz (2006). "Factors Affecting Bioremediation of Hydrocarbons in Terrestrial Environment." Civil Engineering Colloquium: 1-9.

115. Morris, R. N. a. B. (2001) "Use of glucose to control basal expression in the pET System." **13**, 8-10.

116. Nadalig, T., M. Farhan Ul Haque, S. Roselli, H. Schaller, F. Bringel and S. Vuilleumier (2011). "Detection and isolation of chloromethane-degrading bacteria from the Arabidopsis thaliana phyllosphere, and characterization of chloromethane utilization genes." FEMS Microbiol Ecol **77**(2): 438-448.

117. Nallapan Maniyam, M., F. Sjahrir, A. Ibrahim and A. G. Cass (2013). "Biodegradation of cyanide by Rhodococcus UKMP-5M." Biologia **68**(2): 177-185.

118. Navarro-Llorens, J. M., M. A. Patrauchan, G. R. Stewart, J. E. Davies, L. D. Eltis and W. W. Mohn (2005). Phenylacetate catabolism in Rhodococcus sp. strain RHA1: a central pathway for degradation of aromatic compounds. J Bacteriol. United States. **187:** 4497-4504.

119. Nishino, S. F., K. A. Shin, J. M. Gossett and J. C. Spain (2013). "Cytochrome P450 initiates degradation of cis-dichloroethene by Polaromonas sp. strain JS666." Appl Environ Microbiol **79**(7): 2263-2272.

120. Novagen (2002). pET-28a-c(+) Vectors. M. Milipore.

121. Novagen (2004) "Competent cells." 1-23.

122. Novotný, C., K. Svobodová, P. Erbanová, T. Cajthaml, A. Kasinath, E. Lang and V. Sasek (2004). "Ligninolytic fungi in bioremediation: extracellular enzyme production and degradation rate." Soil Biology and Biochemistry **36**(10): 1545-1551.

123. Oganesyan, N., S. H. Kim and R. Kim (2005). "On-column protein refolding for crystallization." J Struct Funct Genomics **6**(2-3): 177-182.

124. Pacwa-Plociniczak, M., G. A. Plaza, Z. Piotrowska-Seget and S. S. Cameotra (2011). "Environmental applications of biosurfactants: recent advances." Int J Mol Sci **12**(1): 633-654.

125. Pandey Shivanand, S. N. (2010). "RECOMBINANT DNA TECHNOLOGY AND GENETIC ENGINEERING:

126. A SAFE AND EFFECTIVE MEANING FOR PRODUCTION VALUABLE BIOLOGICALS." International Journal of Pharmaceutical Sciences Review and Research **1**(1): 14-19.

127. Patil, Y. B. and K. M. Paknikar (2000). "Biodetoxification of silver-cyanide from electroplating industry wastewater." Lett Appl Microbiol **30**(1): 33-37.

128. Patterson, S. D. and R. H. Aebersold (2003). Proteomics: the first decade and beyond. Nat Genet. United States. **33 Suppl:** 311-323.

129. Paul, D., G. Pandey, J. Pandey and R. K. Jain (2005). "Accessing microbial diversity for bioremediation and environmental restoration." Trends in Biotechnology **23**(3): 135-142.

130. Peng, J. and J. Xu (2010). "Low-homology protein threading." Bioinformatics **26**(12): i294-i300.

131. Peng, J. and J. Xu (2011). "Raptorx: Exploiting structure information for protein alignment by statistical inference." Proteins: Structure, Function, and Bioinformatics **79**(S10): 161-171.

132. Penny, C., S. Vuilleumier and F. Bringel (2010). "Microbial degradation of tetrachloromethane: mechanisms and perspectives for bioremediation." FEMS Microbiology Ecology **74**(2): 257-275.

133. Perkins, D. N., D. J. Pappin, D. M. Creasy and J. S. Cottrell (1999). "Probability-based protein identification by searching sequence databases using mass spectrometry data." Electrophoresis **20**(18): 3551-3567.

134. Plaxco, K. W. and M. Gross (2001). "Unfolded, yes, but random? Never!" Nat Struct Mol Biol **8**(8): 659-660.

135. Plebani, M. (2005). Proteomics: the next revolution in laboratory medicine? Clin Chim Acta. Netherlands. **357:** 113-122.

136. Poelarends, G. J., M. Wilkens, M. J. Larkin, J. D. van Elsas and D. B. Janssen (1998). "Degradation of 1,3-dichloropropene by pseudomonas cichorii 170." Appl Environ Microbiol **64**(8): 2931-2936.

137. Posada, A. C. (2006). "Biphenyl Degradation Genes in Rhodococcus Sp. 124." Massachusetts Institute of Technology Undergraduate Research Journal **14**(Fall 2006): 18-21.

138. Sallis, P. J., S. J. Armfield, A. T. Bull and D. J. Hardman (1990). "Isolation and characterization of a haloalkane halidohydrolase from Rhodococcus erythropolis Y2." J Gen Microbiol **136**(1): 115-120.

139. Samin, G., M. Pavlova, M. I. Arif, C. P. Postema, J. Damborsky and D. B. Janssen (2014). "A Pseudomonas putida Strain Genetically Engineered for 1,2,3-Trichloropropane Bioremediation." Appl Environ Microbiol **80**(17): 5467-5476.

140. San-Miguel, T., P. Perez-Bermudez and I. Gavidia (2013). "Production of soluble eukaryotic recombinant proteins in is favoured in early log-phase cultures induced at low temperature." Springerplus **2**(1): 89.

141. Santhi, V. and A. Mustafa (2013). "Assessment of organochlorine pesticides and plasticisers in the Selangor River basin and possible pollution sources." Environmental Monitoring and Assessment **185**(2): 1541-1554.

142. Sayavedra-Soto, L. A., W.-N. Chang, T.-K. Lin, C.-L. Ho and H.-S. Liu (2006). "Alkane Utilization by Rhodococcus Strain NTU-1 Alone and in Its Natural Association with Bacillus fusiformis L-1 and Ochrobactrum sp." **22**(5): 1368-1373.

143. Searle, B. C. (2010). "Scaffold: a bioinformatic tool for validating MS/MS-based proteomic studies." Proteomics **10**(6): 1265-1269.

144. Seffernick, J. L., M. L. de Souza, M. J. Sadowsky and L. P. Wackett (2001). "Melamine deaminase and atrazine chlorohydrolase: 98 percent identical but functionally different." J Bacteriol **183**(8): 2405-2410.

145. Seibert, C. M. and F. M. Raushel (2005). "Structural and Catalytic Diversity within the Amidohydrolase Superfamily†." Biochemistry **44**(17): 6383-6391.

146. Shankar, S., C. Kansrajh, M. G. Dinesh, R. S. Satyan, S. Kiruthika and A. Tharanipriya (2014). "Application of indigenous microbial consortia in bioremediation of oil-contaminated soils." International Journal of Environmental Science and Technology **11**(2): 367-376.

147. Sharp, J. O., C. M. Sales, J. C. LeBlanc, J. Liu, T. K. Wood, L. D. Eltis, W. W. Mohn and L. Alvarez-Cohen (2007). "An inducible propane monooxygenase is responsible for N-nitrosodimethylamine degradation by Rhodococcus sp. strain RHA1." Appl Environ Microbiol **73**(21): 6930-6938.

148. Shukla, H. D. (2006). Proteomic analysis of acidic chaperones, and stress proteins in extreme halophile Halobacterium NRC-1: a comparative proteomic approach to study heat shock response. Proteome Sci. England. **4:** 6.

149. Shukla, K. P., N. K. Singh and S. Sharma (2010). "Bioremediation: Developments. Current Practices and Perspectives." Genetic Engineering and Biotechnology Journal(GEBJ-3): 1-20.

150. Singh, O. V. and N. S. Nagaraj (2006). "Transcriptomics, proteomics and interactomics: unique approaches to track the insights of bioremediation." Briefings in Functional Genomics & Proteomics **4**(4): 355-362.

151. Singh, S. M. and A. K. Panda (2005). "Solubilization and refolding of bacterial inclusion body proteins." J Biosci Bioeng **99**(4): 303-310.

152. Smith, S. M., S. Rawat, J. Telser, B. M. Hoffman, T. L. Stemmler and A. C. Rosenzweig (2011). "Crystal structure and characterization of particulate methane monooxygenase from Methylocystis species strain M." Biochemistry **50**(47): 10231-10240.

153. Soding, J., A. Biegert and A. N. Lupas (2005). "The HHpred interactive server for protein homology detection and structure prediction." Nucleic Acids Res **33**(Web Server issue): W244-248.

154. Soong, C. L., J. Ogawa, E. Sakuradani and S. Shimizu (2002). "Barbiturase, a novel zinc-containing amidohydrolase involved in oxidative pyrimidine metabolism." J Biol Chem **277**(9): 7051-7058.

155. Stuart, M., D. Lapworth, E. Crane and A. Hart (2012). "Review of risk from potential emerging contaminants in UK groundwater." Science of The Total Environment **416**(0): 1-21.

156. Studier, F. W. and B. A. Moffatt (1986). "Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes." J Mol Biol **189**(1): 113-130.

157. Suhaila, Y. N., M. Rosfarizan, S. A. Ahmad, I. Abdul Latif and A. B. Ariff (2013). "Nutrients and culture conditions requirements for the degradation of phenol by Rhodococcus UKMP-5M." J Environ Biol **34**(3): 635-643.

158. Tao, H., W. Liu, B. N. Simmons, H. K. Harris, T. C. Cox and M. A. Massiah (2010). "Purifying natively folded proteins from inclusion bodies using sarkosyl, Triton X-100, and CHAPS." Biotechniques **48**(1): 61-64.

159. Terpe, K. (2003). "Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems." Appl Microbiol Biotechnol **60**(5): 523-533.

160. Tomas-Gallardo, L., I. Canosa, E. Santero, E. Camafeita, E. Calvo, J. A. Lopez and B. Floriano (2006). "Proteomic and transcriptional characterization of aromatic degradation pathways in Rhodoccocus sp. strain TFB." Proteomics **6 Suppl 1**: S119-132.

161. Tomei, M. C. and A. J. Daugulis (2012). "Ex Situ Bioremediation of Contaminated Soils: An Overview of Conventional and Innovative Technologies." Critical Reviews in Environmental Science and Technology **43**(20): 2107-2139.

162. Torres Pazmino, D. E., M. Winkler, A. Glieder and M. W. Fraaije (2010). "Monooxygenases as biocatalysts: Classification, mechanistic aspects and biotechnological applications." J Biotechnol **146**(1-2): 9-24.

163. Torz, M. and V. Beschkov (2005). "Biodegradation of monochloroacetic acid used as a sole carbon and energy source by Xanthobacter autotrophicus GJ10 strain in batch and continuous culture." Biodegradation **16**(5): 423-433.

164. Trotsenko, Y. and M. Torgonskaya (2009). "The aerobic degradation of dichloromethane: Structural-functional aspects (a review)." Applied Biochemistry and Microbiology **45**(3): 233-247.

165. Tsumoto, K., D. Ejima, I. Kumagai and T. Arakawa (2003). "Practical considerations in refolding proteins from inclusion bodies." Protein Expression and Purification **28**(1): 1-8.

166. Upadhyay, A. K., A. Singh, K. J. Mukherjee and A. K. Panda (2014). "Refolding and purification of recombinant L-asparaginase from inclusion bodies of E. coli in to active tetrameric protein." Frontiers in Microbiology **5**.

167. Vallejo, L. F. and U. Rinas (2004). "Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins." Microb Cell Fact **3**(1): 11.

168. van Pee, K. H. and S. Unversucht (2003). "Biological dehalogenation and halogenation reactions." Chemosphere **52**(2): 299-312.

169. Vanderberg, L. A. and J. J. Perry (1994). "Dehalogenation by Mycobacterium vaccae JOB-5: role of the propane monooxygenase." Can J Microbiol **40**(3): 169-172.

170. Vashist, S. (2013). "Anti-Polyhistidine Antibodies." Materials and Methods **3**: 180.

171. Vasina, J. A. and F. Baneyx (1997). "Expression of aggregation-prone recombinant proteins at low temperatures: a comparative study of the Escherichia coli cspA and tac promoter systems." Protein Expr Purif **9**(2): 211-218.

172. Vellore, J. M. (2001). Iron acquisition in Rhodococcus erythropolis: the mutant(s) that do not produce a Siderophore. MSc, East Tennessee State University.

173. Vidali, M. (2001). "Bioremediation: An Overview." Pure and Applied Chemistry **73**(7): 1163-1172.

174. Woestenenk, E. A., M. Hammarstrom, S. van den Berg, T. Hard and H. Berglund (2004). "His tag effect on solubility of human proteins produced in Escherichia coli: a comparison between four expression vectors." J Struct Funct Genomics **5**(3): 217-229.

175. Wu, W., L. Xing, B. Zhou and Z. Lin (2011). "Active protein aggregates induced by terminally attached self-assembling peptide ELK16 in Escherichia coli." Microb Cell Fact **10**(1): 9.

176. Xiang, Z. (2006). "Advances in Homology Protein Structure Modeling." Curr Protein Pept Sci **7**(3): 217-227.

177. Yamaguchi, H. and M. Miyazaki (2014). "Refolding techniques for recovering biologically active recombinant proteins from inclusion bodies." Biomolecules **4**(1): 235-251.

178. Yokota, T., T. Omori and T. Kodama (1987). "Purification and properties of haloalkane dehalogenase from Corynebacterium sp. strain m15-3." J Bacteriol **169**(9): 4049-4054.

179. Zambelli, B., A. Berardi, V. Martin-Diaconescu, L. Mazzei, F. Musiani, M. Maroney and S. Ciurli (2014). "Nickel binding properties of Helicobacter pylori UreF, an accessory protein in the nickel-based activation of urease." JBIC Journal of Biological Inorganic Chemistry **19**(3): 319-334.

180. Zarogiannis, P., T. A. Nwaogu, N. Tuffnel and B. Lucas (2007). Impact Assessment of Potential Restrictions on the Marketing and Use of Dichloromethane in Paint Strippers, RPA**:** 1-266.

181. Zhang, J., J. G. Yin, B. J. Hang, S. Cai, J. He, S. G. Zhou and S. P. Li (2012). "Cloning of a novel arylamidase gene from Paracoccus sp. strain FLN-7 that hydrolyzes amide pesticides." Appl Environ Microbiol **78**(14): 4848-4855.

182. Zhou, J., A. Palumbo and J. Strong (1999). "Phylogenetic characterization of a mixed microbial community capable of degrading carbon tetrachloride." Applied Biochemistry and Biotechnology **80**(3): 243-253.

183. Zhu, X.-Q., S.-X. Li, H.-J. He and Q.-S. Yuan (2005). "On-column Refolding of an Insoluble His6-tagged Recombinant EC-SOD Overexpressed in Escherichia coli." Acta Biochimica et Biophysica Sinica **37**(4): 265-279.