DOCTOR OF PHILOSOPHY

The cloning, expression and characterisation of PbRR from Pseudomonas putida KT2440

Ansir, Nagwa Hassan

Award date:
2014

Awarding institution:
Bangor University

The Cloning, Expression and Characterisation of PbrR from *Pseudomonas putida* KT2440

A thesis submitted to Bangor University in candidature for the degree of Doctor of Philosophy in Chemistry

School of Chemistry, Bangor University, Gwynedd, LL 57UW, UK

By

Nagwa Hassan Ansir

2014

ABSTRACT:

Highly selective lead-binding proteins (PbrR family) have been characterised in relatively few biological systems. This thesis reports the first research on cloning, expression and characterisation of native and mutant PbrR from *Pseudomonas putida* KT2440 (PbrR(Pp)). Three cysteine residues are believed to play a vital role in ligand binding and discrimination (Cys77, Cys112, and Cys119). To investigate the role of these residues, cysteine was mutated to single or multiple mutants of either alanine, (C77A, C77A/C112A, and C77A/C112/119A) or aspartic acid (C77D, C77D/C112D, and C77D/C112D/119 D).

A recombinant DNA vector was generated by cloning the PbrR(Pp) gene into the pET28 (+) plasmid. Site directed mutagenesis was utilised to generate the desired mutants using the recombinant DNA plasmid as a template. The recombinant and all mutants were successfully expressed in the *E. coli* strain BL21 (DE3), as evidenced by DNA sequencing and SDS PAGE analysis, and purified using HiTrap affinity columns.

UV/Vis spectroscopy, atomic absorption spectroscopy and protein concentration assays showed that lead was partially incorporated into the as-isolated PbrR protein from *Pseudomonas putida* KT2440 (PbrR(Pp)) with a protein:lead ratio of (1:0.9) Reconstitution of the holoprotein, either via dialysis or titration, resulted in formation of a Pb-PbrR(Pp) complex, indicating that each PbrR(Pp) dimer binds one Pb^{2+} ion. A characteristic charge transfer band in the UV/Vis spectrum of Pb-PbrR(Pp) was observed at 332 nm (ε_{max} = 3,720 M^{-1} cm^{-1}). For the mutant proteins, only the Cys77D (aspartic acid) mutant had a nearly 1:1 lead:protein with percentage lead saturation in Pb-PbrR-C77D was 88% after reconstitution with lead, and an associated blue shifted charge transfer band appeared at 318 nm (ε_{max}=3889 M^{-1} cm^{-1}). The other mutants did not exhibit any clearly defined charge transfer band and the percentage lead saturation were range (15%–40%).

Currently, ^{207}Pb NMR spectroscopy has not been widely used by many research groups to study lead enzyme complexes. In this thesis, ^{207}Pb NMR spectroscopy has

been used to characterise a model system, a Pb-glutathione complex, and Pb-PbrR(Pp). Comparison of the observed chemical shifts (singlet centred at 2765 ppm) confirmed that the Pb-PbrR(Pp) has a Pb-S_3 coordination geometry.

Table of Contents

List of Figures

Chapter 1:

List of Tables

ABBREVIATIONS

3c **Tri-coordinate**

4c **Tetra-coordinate**

bp **Base pair**

CT **Charge transfer**

Da **Dalton**

dNTP **Deoxyribonucleoside triphosphate**

IPTG **Isopropyl-β-D-thiogalactopyranoside**

kDa **Kilo Dalton**

Km **Kanamycin**

LB **Luria Bertani broth**

NCBI **National Centre for Biotechnology Information**

NMR **Nuclear Magnetic Resonance**

OD **Optical density**

PAGE **Polyacrylamide gel electrophoresis**

PCR **Polymerase chain reaction**

pET **Plasmid for expression by T7 RNA polymerase**

PbrR(Pp) **PbrR protein from *Pseudomonas putida* KT2440**

PbrR(Rm) **PbrR(Rm) PbrR protein from *Ralstonia metallidurans* CH34**

SDS **Sodium dodecyl sulphate**

TEMED ***N-N-N`-N`*-tetramethylethylenediamine**

UV **Ultraviolet**

Vis **Visible**

ACKNOWLEDGMENTS

I am particularly grateful to my supervisor Dr Loretta Murphy for her great help and support. From the very beginning of my project, she has been supportive and helpful throughout the course of writing the material. She does indeed deserve special thanks for reading the early drafts of the project and for making some insightful comments.

I would like also to thank all the staff for their encouraging comments and productive questions and feedback, and especially Dr. Christopher Gwenin (School of Chemistry) and Dr. Thomas Caspari (School of Biological Sciences) for encouraging me to pursue this project.

Many people have helped me during this research in many different ways. I am very grateful to my colleagues for their invaluable comments and suggestions. Numerous colleagues have provided me with helpful criticism and advice at various stages and I have also benefited from some useful discussions with them. Special thanks go to the School of Chemistry staff for their support, especially Gwynfor Davies, Mike Lewis, Sam Page, Nicholas Welsby, Denis Williams and Glynne Evans.

I do not think this work could have been accomplished without the support of my family in Libya especially my mother (God have mercy on her) and my father. Nothing in my life could have been accomplished without his guidance and support and to whom I owe everything. I would like to dedicate this thesis to my husband (Allafi Gawad) for all the support, help and for listening to my day-to-day stories and coping with my frustrations. To Ghaida and Rama, my lovely girls, who brought lots of laughs and good times in my life. Thanks to all my family in Libya for all the support

CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW

1.1. Heavy metals

The term *heavy metal* has no exact definition. In general, it pertains to metals having a high relative atomic mass between 63.546 and 200.590 and specific gravity greater than 4.0. This term also can refer to any metallic chemical elements that are toxic to the environment and human health, including lead (Pb), mercury (Hg) and cadmium (Cd)[1]. The ambiguity lies in the fact that some elements, such as gold (Au), for instance, have high atomic weights (196.97 g/mol), yet are non-toxic, while others, such as Cd, have relatively low atomic weights (112.41 g/mol) but are toxic. Still others are beneficial to human health at low doses; copper (Cu), for instance, is essential to well-being as it is an essential component of cytochrome c oxidase, a respiratory enzyme complex. However, at high doses, particularly in pure form, Cu is toxic.[2] The situation is further complicated in that the toxicity of a metal is in part a function of it oxidation state/speciation; hexavalent chromium (Cr), for example, is lethal, but trivalent Cr is not; indeed, it is a human nutrient. For these reasons, IUPAC (International Union of Pure and Applied Chemistry) describes the assignation of *heavy metals* as "meaningless" in a technical report.[3]

Nonetheless, heavy metals can be defined as elements with a specific gravity five or more times higher than water (1 at 4°C).[4] Moreover, despite problems over definition, some metals are toxic: Pb, for instance, as discussed below, has no human nutritional importance and is deadly in high doses. In short, some metals in the environment can be hazardous to human health, and for simplicity one can term these *heavy metals*. Previous studies suggested that the following 11 elements had potential health risks: arsenic (As), Cd, cobalt (Co), Cr, Cu, Hg, manganese (Mn), nickel (Ni), Pb, tin (Sn), and thallium (Tl). However, living organisms require trace amounts of elements such as Mn, Zn and Cu,[5,6] but the situation is delicate, for high concentrations of otherwise beneficial metal ions are toxic. The toxicity of metal ions within cells is therefore, in part, a function of their concentration.[7] Table 1.1 lists the maximum recommended concentration levels of heavy metals in the environment.[8]

Table: 1. 1. Maximum contamination levels for heavy metal concentrations in air, soil and water.

Heavy Metals	*Max Conc. in Air (mg/m^3)*	*Max. Conc. In sludge (soil) ppm*	*Max. Conc. In Drinking Water (mg/l)*	*Max. Conc. In H_2O Supporting Aquatic Life (mg/l)*
Cd^{2+}	0.1-0.2	85	0.005	0.008
Pb^{2+}	--	420	0.01	0.0058
Zn^{2+}	1,5	7500	5.00	0.0766
Hg^{+}	--	1	0.002	0.05
Ag^{+}	0.01	--	0.0	0.1
As^{2+}	--	--	0.01	--

Source: Duruibe *et al.*[8]

Some of the metals are not only directly dangerous to human health;[9] they are also subject to corrosion (e.g. Cu),[10] and some cause harm in other ways (e.g. as pollutants).[11] Some, including Cu, Pb, Cd, are known to bio-accumulate in marine life,[12] which means that low environmental levels may nonetheless prove toxic.[13-15] A further problem is that some such metals persist in the environment.

The toxicity of heavy metals is a function of their binding to functional groups. In sites within the cell, some (e.g. Cd) tend to bind with the thiol (SH) groups in glutathione and produce a bis-glutathione complex.[16] The binding of heavy metals such as Hg (II), Pb(IV), Tl (III), and Pd (II) to functional groups could, for example, promote the formation of free radical reactions; these can displace the methyl group in vitamin B12.[17] This action may block transfer of fluids within the cell, and in turn deleteriously affect the functioning of the cell. In this regard, free radicals may be implicated as a cause of cancer.[18,19] High concentrations of heavy metals also have effects on such organs as the lungs, the liver, the kidneys and the brain.[20] High concentrations of heavy metals have also been implicated in such degenerative disorders as Alzheimer's disease,[21] Parkinson's disease,[22] muscular dystrophy, [23]and multiple sclerosis.[24] The following sub-sections itemise the more problematic heavy metals.

1.1.1 Arsenic

Arsenic (As) is abundant in the atmosphere, soils, rocks, natural waters and organisms.[25,26] It is the most common cause of acute heavy metal poisoning in adults[27] and is released into the environment *via* smelting of Cu or Zn ores. Two zinc smelters in Palmerton, PA have highly contaminated the surrounding soil and vegetation with Zn, Cd, Cu and Pb.[28] Arsenic is also present in some pesticides, paints, rat poisons, fungicides and wood preservers. Arsenic has contaminated water supplies and has poisoned fish. It is a particularly nasty poison, affecting the skin, the central nervous system, the digestive system, the blood and the kidneys.[29]

1.1.2. Mercury

Mercury (Hg) poisoning can be fatal as it deteriorates mental health and has hazardous effects on the nervous, digestive and immune systems and the lungs and kidneys.[29,30] Research indicates that the molecular interactions occurring between mercury ions and different reaction points in the kidney are responsible for the specific compound-related toxicity and hazards. The nucleophilic reaction sites present in the vicinity and on the kidney cells take part in these interactions.[31,32] A further look into the details of the toxic mechanism reveals that the sulphur-containing groups that constitute the major cellular components of the kidney (like albumin, metallothionein, glutathione, and cysteine) possess strong attractions for the Hg-containing compounds. Hence, molecules containing sulphydryl and other sulphur-containing groups (e.g. thiol groups) react with the Hg ions, cause damage to the proximal convoluted tubules and induce considerable damage to the kidneys. [33,34]

Some of the most notable cases alarmed people during the 1950s to 1960s, with the appearance especially of the potentially fatal Minamata disease, which was caused by methyl-mercury poisoning due to bio-concentration of Hg in fish/shellfish in the human food chain. For this reason, this metal is legally prohibited in a number of various uses. Although its use is tightly regulated it is still used in some applications e.g. thermometers, thermostats and dental amalgams. Mercury poisoning may also occur through inhalation, from the gaseous discharges from mining and paper industries. [35]

1.1.3. Cadmium

Cadmium (Cd) is a common by product of the mining and smelting of Pb and Zn ores, occurring as CdTe and CdS.[36] It is found in batteries, PVC, paints, fertilisers, cigarette tobacco and insecticides; hence, it is present in agricultural soils. Other environmental sources include vehicle exhaust, dental alloys and motor oil. It has been found in reservoirs containing shellfish. Cadmium poisoning damages the liver, placenta, kidneys, bones, lungs and brain.[29,30] Cadmium can cause a large variety of toxic reactions, majorly targeting the bones, lungs and kidneys.

Within the living system, Cd^{2+} ions compete with Ca^{2+} ions and lower the normal calcium levels in the body. This, in turn, prevents the effective absorption of calcium ions into the bones. The ineffective and incomplete absorption of calcium causes bones to lose their normal toughness and to become fragile and highly prone to bone damage and fractures. [37] Similar to mercury, Cd also interacts with different sites in the kidney and its accumulation there leads to different kidney diseases. [38] The commonly seen diseases that occur due to Cd accumulation in the kidney include proteinuria, glomerular filtration rate lowering, increased occurrence of kidney stone formation and even renal cancers.[39] The oral and inhalation routes are also responsible for the accumulation of Cd and its related toxicity to the lungs in the form of severe bronchiolitis and emphysema.

1.1.4. Lead

After As, lead (Pb) is the most hazardous of the heavy metals (supporting data are shown in Table 1.1). Lead contamination has negative effects both on the environment and on human health.[40-42] This is unfortunate because Pb is a soft metal, and therefore has numerous industrial and practical applications in drains and other pipes, for instance, and in soldering. At present, approximately 2.5 million tons are released into in the environment every year. [43] Lead appears to play no essential role in biological processes;[44] however, it may interfere with the functioning of diverse biologically essential metals such as Ca, Fe, and Zn.[45]In this regard, Pb can bind to the same binding sites and ligands as biologically essential metals [S donor ligands within aminolevulinate dehydratase (ALAD), for example].[46] This binding changes the functioning of the molecules and this may lead to failure to complete essential

reactions, including, for example, enzyme production.[47]

Lead accounts for most cases of heavy metal poisoning in children. Environmental sources include buildings built prior to 1940; such buildings may contain Pb for piping and use Pb-based paints. The main industrial sources are Pb used in batteries. Lead also continues to be used in some countries for paints, plumbing, cable coverings and ammunition (for bullets). Poisoning affects the bones, the brain, blood, kidneys and thyroid gland.[48]

Lead is a common component of medical and municipal waste and has been shown to leach into soil and water.[49]The Pb is thereafter transferred to the people via drinking water and food. Such Pb pollution is a matter of major concern, not only because of its danger to human health (especially to children and foetuses) but also because of the persistence of Pb in soil.[50-52] Although use of Pb in gasoline (as an anti-knock agent) and in paint has been discontinued, Pb remains a significant public health threat.[53,54]

However, there is a caveat. Although Pb pollution appears to be a cause of environmentally-induced disease,[55-57]and it remains important to monitor and regulate Pb levels in the environment. The severity of the pollution, at least in the West, appears to be declining. In London (UK), for instance, Pb concentrations in the air fell by 87 % in the period 1976–1993, threats from other potentially toxic metals also decreased. Levels of Cr, Cu, Zn, Cd, and Ni in London air fell by between 50 % and 75% (according to metal); and US trends appear similar. [43] In the period of 1990-2010, the emission of Pb was decreased (89 %) in the EU, although this percentage slightly increased (by 9.1 %) in 2009-2010. [58] In any event, metal pollution, including that of Pb, appears to be waning. This, of course, does not preclude the need for continued vigilance.

1.2 Electronic configuration of lead

Lead is a soft, heavy, toxic and malleable metal. Compared with other metals, it has poor electrical conductivity. It has four stable, naturally occurring isotopes: ^{204}Pb, ^{206}Pb, ^{207}Pb, and ^{208}Pb. The oxidation states and electronic configurations of Pb are Pb(0) $[Xe]4f^{14}5d^{10}6s^{2}6p^{2}$, Pb (II) $[Xe]4f^{14}5d^{10}6s^{2}$, and Pb(IV) $[Xe]4f^{14}5d^{10}$. Of these

three oxidation states, the interaction of divalent Pb (Pb^{2+}) with biological samples has been especially well studied. [59] This oxidation state binds preferentially to thiol and phosphate groups in nucleic acids, proteins and cell membranes. [60] The oxidised Pb significantly destabilises the double-helix structure of DNA.[61] Pb is a post-transition metal: the lone pair of electrons on the 6s orbital of Pb^{2+} resist removal and participate in covalent bond formation and hydrogen bonding. This appears to be due to a relativistic effect that causes the 6s orbital to contract, which in turn increases the energy required to remove or interact with the 6s pair of electrons. Meanwhile, the d and f orbitals are destabilised; they expand radially as a result of screening from nuclear attraction by the s and p electrons. This phenomenon results in a stable, relatively inert outer lone pair of electrons. This pair effects a nonspherical charge distribution around the Pb^{2+}cation; the disposition of ligands around the cation results in an identifiable void. [62]

The coordination chemistry of Pb(II) is unique. This metal ion has a stereochemically active lone pair that is resistant to engagement with bonding to ligands. In consequence, in low coordinate Pb(II) complexes (three-coordinate or four-coordinate), the lone pair is exposed and the Pb (II) centre adopts a hemidirected geometry in which all ligands are clustered on one side of the metal. Shimoni-Livny and colleagues have used crystallography to examine the role of the lone pair of electrons of Pb (II) as regards coordination geometry and *ab initio* (i.e., from the beginning) molecular orbital optimisations.[62] The work addressed contributory factors to the disposition of ligands around the Pb with geometries that are (a) holodirected (the bonds to ligand atoms are distributed throughout the surface of an encompassing globe) and (b) hemidirected (the bounds to ligand atoms are directed throughout only part of an encompassing globe). Figure 1.1 shows examples of lead coordination geometries.

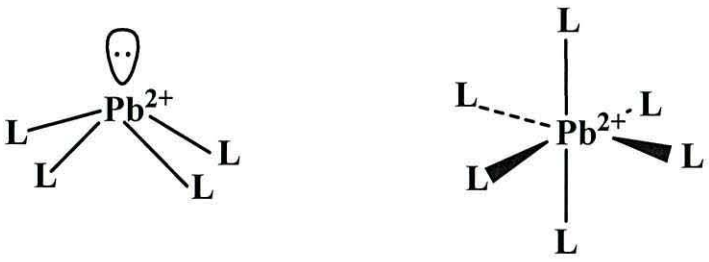


Figure: 1.1. Coordination geometry of Pb hemidirected and holodirected coordination. [62]

The Pb^{2+} low coordination numbers (2–5) are hemidirected and holodirected for high coordination numbers (9–10). Shimoni-Livny *et al.* found the preferred coordination numbers to be 4 for Pb^{4+} and 4 and 6 for Pb^{2+}.[62] All Pb^{2+} structures in the Cambridge Structure Database (CSD) of compounds appeared hemidirected for low coordination numbers (2–5) and holodirected for high coordination numbers (9–10); however, examples of each type of stereochemistry appeared present for intermediate coordination numbers. The energy cost for converting a hemidirected to a constrained holodirected structure was of the order 8–12 Kcal mol^{-1} in the absence of a strong inter-ligand interaction. [62] Meanwhile, Pb (II) can bind to the same site in the peptides by adopting a three-coordinate trigonal planar, which is fundamentally different from the natural coordination of zinc in peptides (Figure 1.2). Peptide conformation undergoes a dramatic change after Pb binding, leading to improper protein folding. These effects account for the disruption of peptide structures by Pb, and may provide the molecular mechanism for the toxicity of Pb. [63]

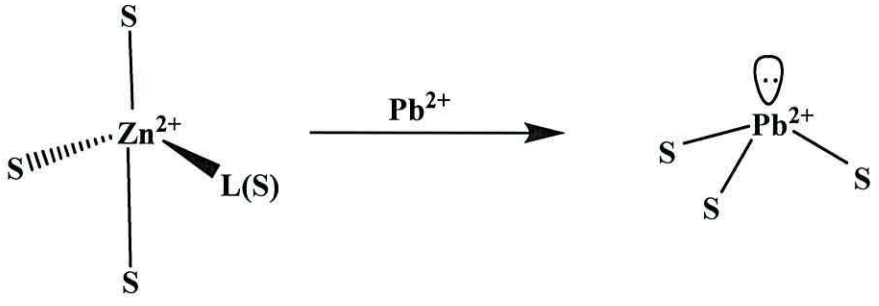


Figure: 1.2. Coordination preference of Zinc (II) vs. Pb (II) in cysteine-rich domains. [63]

1.2.1. Coordination sphere of lead

The Pb^{2+} ion has versatile behaviour. It can be complexed by a combination of S, N, O or P-containing ligands with a coordination number in the range 2–9. A trigonal pyramidal geometry appears the most preferred geometry for Pb (II) in a thiol-rich environment; in this, a lone pair occupies the apical position (hemidirected). [63]

Proteins are known to bind metal ions almost exclusively through coordination via three types of atoms: sulphurs (Cys and Met), nitrogens (His), and oxygens (Asp, Glu and Tyr). The metal ion binding sites in proteins adopt a variety of *chelate structures* (atoms within a molecule surrounding the same central molecular structure). Coordination numbers (n) for transition metals in biological systems are usually in the range of 3–6; coordination geometries range through trigonal (n = 3), tetrahedral, square planar, trigonal pyramidyl (n= 4), trigonal bipyramidal (n = 5) and octahedral (n = 6). Specific metal ions appear to prefer specific coordination numbers and chelate geometries in addition to specific ligand types. Research suggests these preferences govern metal selectivity in biological systems. [64]

1.2.2. Pb-binding protein

PbrR(Rm) is a MerR type protein found in the gram-negative *Ralstonia metallidurans* (*Cupriavidus metallidurans* strain CH34) bacterium, a bacterium adapted to survive in heavy metal environments. PbrR(Rm) has a very high selectivity for binding lead (II). The PbrR(Rm) protein is composed of two protein subunits, giving a molecular weight of 44 kDa. [65-,67]

The metal coordination geometry appears hemidirected from analysis of the x-ray absorption edge spectrum of the Pb(II)-loaded PbrR(Rm) and resembles that of a PbS_3 model. [65] Thus, the PbrR protein utilises a unique, hemidirected geometry to selectively recognise Pb(II). As a homologue of MerR protein, the dimer of PbrR, also binds one equivalent of Pb(II). [66,67] Figures 1.3 and 1.4 show the binding of Pb (II) to PbrR.

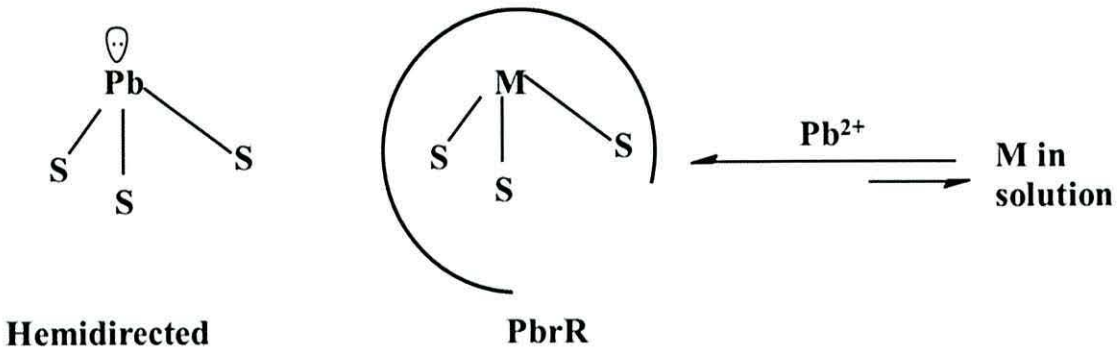


Figure: 1.3. Lead (II) prefers a hemidirected geometry in 3- or 4-coordinate complexes. [67]

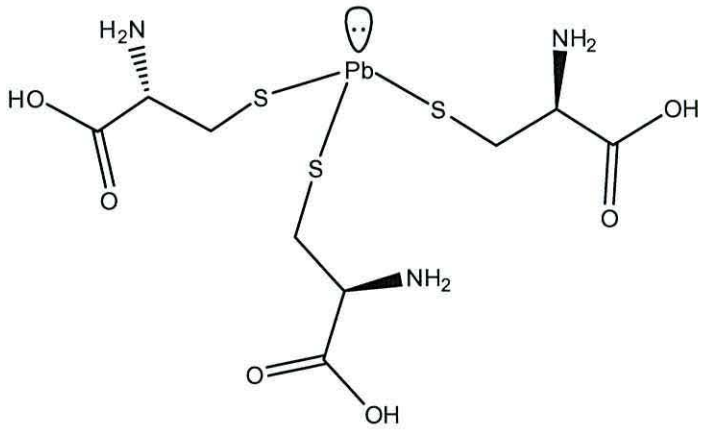


Figure: 1.4. Proposed structure of Pb (II) bound to the thiol groups of cysteines in the PbrR protein.

1.2.3. Electronic absorption spectroscopy of PbrR

The UV–visible (UV/Vis) spectra of Pb^{2+}-S solutions ($Pb(GSH)_3$) have suggested strong transitions between 317 and 335 nm; at 317 nm in the UV-vis spectrum and mean Pb-S and Pb-(N/O) bond, these have been assigned to the combination of S $3p \rightarrow Pb^{2+}$ 6p ligand-to-metal charge-transfer (LMCT) and intra-atomic Pb^{2+} $6s^2 \rightarrow Pb^{2+}$ 6p transitions. [42,59,63,68] Similar electronic transitions have been observed for three-coordinate PbS_3 complexes to cysteine-rich metalloregulatory proteins, [68,69,70]and to structural Zn^{2+}-binding domains.

1.3. Lead uptake, transport and binding in biological organisms

Heavy metals damage cells. The underlying mechanisms vary, and include interruption of essential metal ions homeostasis, perturbation of membrane potentials

and inhibition of metalloenzyme's active sites. [71] In humans, cells protect the body against heavy metal toxicity by synthesising thiol-rich small molecules; these include glutathione (GSH) and metallothionein, each of which can sequester toxic metal ions before they can cause damage. [72.73] GSH is a tripeptide; it contains glutamic acid, cysteine, and glycine (γ-Glu-Cys-Gly) residues. It is found in millimolar concentrations (0.1 to 10mM) in erythrocytes, [74] making it the most abundant nonproteinaceous thiol in cellular systems.

Glutathione (GSH) has various cell roles. It is an antioxidant; for example, it protects cells from reactive O species that may damage DNA and RNA. Heavy metals are also thiophilic and as such exhibit an affinity for the thiol of cysteine; thus, GSH can act as a chelator of heavy metals in cellular compartments. [75] The interaction between GSH and Pb (II) is relevant to understanding the toxicity of Pb.

Heavy metals target Zn-finger protein gene regulators. Research has identified several Zn-dependent proteins; in these, the Zn appears to function either catalytically or structurally. [76] The most common and best studied structural Zn-dependent proteins are the Zn fingers, which are of three types[77]: (a) Cys_2His_2, or the cellular/transcription factor type; (b) $Cys_2HisCys$, or the retroviral type, represented by the retroviral nucleocapsid proteins; and (c) Cys_4, or the steroid/thyroid hormone receptor type (see Figure 1.5). [78]

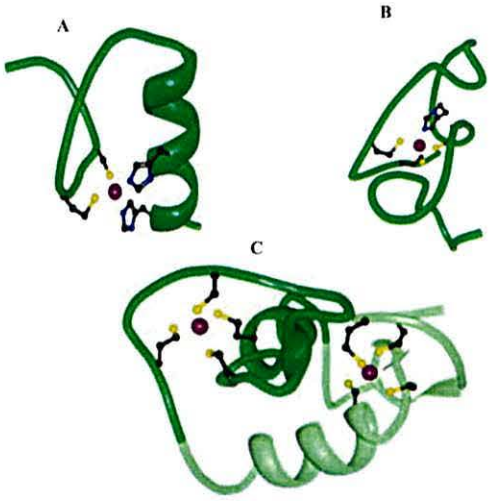


Figure: 1.5. Peptide ribbon diagrams and Zn^{2+}-coordinating side chains of selected zinc finger protein: (a) Cys2His2 zinc-finger for (PDB 1SP2), (b) Cys2His Cys zinc-finger for (PDB 1PXE), (c) Cys4 zinc-finger for (PDB 2GDA). The thiol group of Cys is highlighted in yellow.[78]

Previous study suggests that As(III), Cd(II), Hg(II) and Pb (II) inhibit DNA repair, and that this arises from the displacement of the Zn ion in the Zn finger structures of DNA repair proteins.[79,80]

Berg and colleagues designed a consensus zinc- finger peptide containing 26 amino acid residues[81,82] based on sequence alignment of 131 sequences of transcription factor proteins. Following this, Godwin and colleagues conducted Pb binding studies on a series of Zn-binding domains of consensus peptides (CP-CCHC, CP-CCHH and CPCCCC) and the HIV nucleocapsid protein (HIV-CCHC); the authors suggested that the domains are potential targets for Pb(II).[59,83] The model peptides, with Cys_3 and Cys_4 motifs, suggest that Pb(II) has a higher binding affinity than Zn(II) does; however, the results also suggest that Pb(II) cannot compete with Zn(II) binding to the His2Cys2 site.[83] This may be attributable to the improper folding of the peptide; also, histidine may not tightly coordinate the Pb(II) ion.

1.3.1 ALAD: Aminolevulinic acid dehydratase

Lead is required to maintain a soluble form between the blood, major organs, gastrointestinal tract and lungs for excretion by urinary or faecal routes. The levels of the lead in blood are believed to represent an essential indicator of Pb toxicity in every individual.[84] The movement of Pb in blood and its movement into hard and soft tissues indicates its attachment to a soluble carrier like a metal-binding protein.[85] The bioavailability and toxic potential of lead is mediated by soluble Pb-binding proteins (PbBPs) present in blood [84,86] and target tissues.[87,88] Many studies have been conducted on the mechanism of protein-mediated toxicity for metals like calcium. PbBPs exhibit an attenuating effect of lead on aminolevulinic acid dehydratase (ALAD) and govern intranuclear transport and chromatin binding [85,89] of lead in specific cell nuclei from the kidney.[90,91] ALAD is an enzyme that regulates the haem biosynthetic pathway.[92] In haem formation, ALAD condenses two molecules of aminolevulinic acid (ALA) to produce porphobilinogen (PBG), a haemoglobin precursor (Figure 1.6).[93]

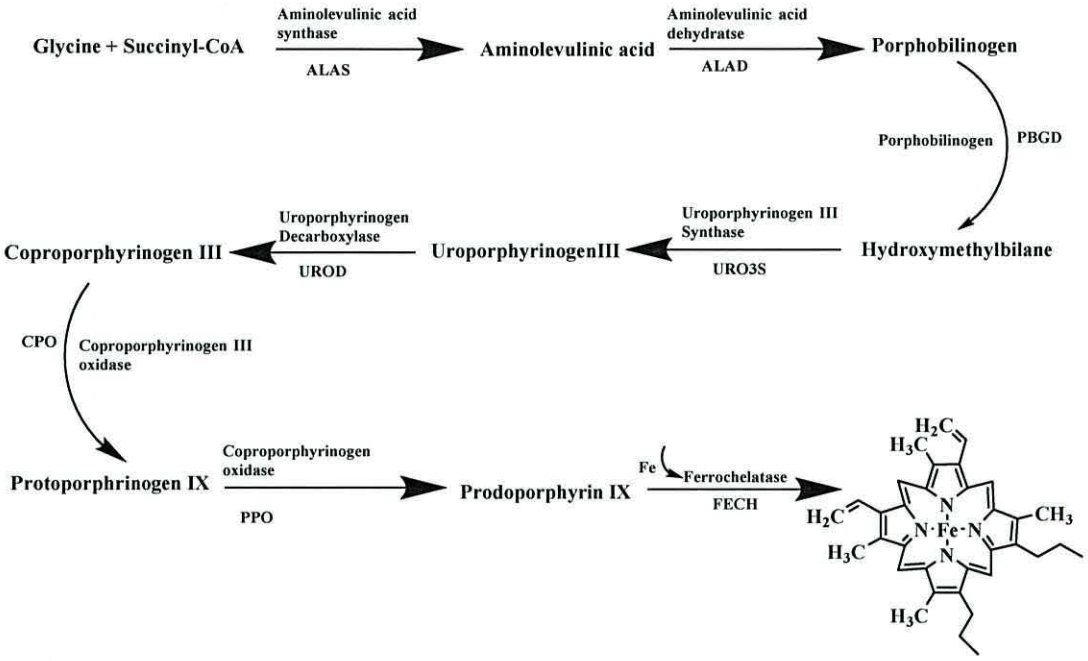


Figure: 1.6. The haem biosynthetic pathway.

Source: Ajioka, *et al.*[93]

Inhibition of ALAD by lead occurs when lead replaces zinc stoichiometrically.[94,95] Haemotoxicity can also result due to lead inhibition of haem formation. This occurs when lead inhibits three enzymes in the biosynthesis pathway: ALAD, ferrochelatase and coporphyrinogen oxidase. The impact of Pb on ALAD is more obvious than on the other two enzymes,[96] so that the level of ALAD inhibition is utilised in measuring lead toxicity. Zinc ions are replaced by lead at the ALAD metal-binding site[97] and enzyme inhibition is due to a consequent change in the quaternary structure of enzyme. When ALAD is altered, a level of 10 μg/dl of aminolevulinic acid can be detected in the plasma and urine. In the nervous system, aminolevulinic acid stimulates γ-aminobutyric acid receptors due to its similarity to γ-aminobutyric acid; this is the leading cause of lead-induced neurotoxicity.[97,98] The formation of the haem metabolite known as zinc protoporphyrin (ZnPP) is also altered by Pb contamination, as Pb will again compete at the same binding site as Zn.

In mammals, Pb^{2+} poisoning of ALAD blocks the synthesis of haemoglobin, causing anaemia. In addition, the high affinity of Pb^{2+} for cysteine thiolates is presumably due to the high enthalpy of Pb–S bond formation and the preferred PbS_3 coordination geometry in thiolate-rich sites of proteins. Chelation of iron with protoporphyrin is

the final stage of reaction in the haemoglobin biosynthetic pathway. Zinc is used as a substitute in ferrochelatase when iron is not available or is insufficient, and this leads to elevated levels of ZnPP in red blood cells (RBCs). This zinc-for-iron replacement happens mainly in bone marrow; therefore, the ZnPP/haem ratio shows the position of iron. The competitive inhibition of heme oxygenase[99] by ZnPP also regulates the catabolism of haem. Initially, ZPP was thought to be produced non-enzymatically via chelation of zinc by the protoporphyrin that accumulates when ferrochelatase is altered by lead or by iron deficiency Within mitochondria,[100] an enzymatic formation of ZPP occurs through the haem synthetic enzyme: this inserts either zinc or ferrous iron into protoporphyrin.[101] In this biosynthetic pathway, haem synthesis plays a key role in regulating haemoglobin levels in RBCs. The inhibition of ALAD within erythrocyte cells can therefore be used as a bioassay to find acute and chronic lead exposure in humans,[102] in other mammals[103] and in birds[104] by measuring Pb-induced anaemia. [105]

1.4. Metallo-regulatory proteins

Metallothionein (MT) is a metal-binding protein that contains abundant cysteine residues. Although its purpose still unknown, this enzyme has been well studied.[106,107] The very first idea of metal-regulated gene expression was presented about thirty years ago, when Piscator showed that administering Cd (ll) could induce the expression of MT [108] Experiments carried out a decade later further validated a role for cadmium and zinc, with the support of actinomycin D, in the transcriptional regulation of MT.[109,110] Further investigations in intact animals and mammalian cells have led to the belief that some hormones and cytokines also contribute to the regulation of MT genes, even though all the previously studied systems showed regulation of MT genes at the transcriptional level by metals.[111]

All eukaryotes have similar elements of the MT regulatory system.[112] Multiple metal response elements (MREs) are a particular example. These MREs of the MT promoter and a metallo-regulatory protein (transcription factor) work in tandem, resulting in commencement of MT gene transcription. Metal occupancy by this factor is likely essential. Studies recognising these proteins have categorised them in yeast, whereas the mammalian proteins did not indicate any metal occupancy. They

are called ACEI (for activation of CUP 1 expression) or AMTl (for activation of metallothionein transcription) in yeasts and MTF-l (for metal-dependent transcription factor), MBF-l or MRE-BP (for metal response element-binding factor), MEP-l (for metal element-binding protein), or ZAP (for zinc-activated protein) in mammals .[113]

Relevant processes here include the production of metal ion-binding proteins, efflux (removal of toxic chemicals) and detoxification (neutralising otherwise toxic chemicals); these processes can allow bacteria to flourish in otherwise highly hostile environments. In this regard, metallo-regulatory proteins (for example the ArsR or MerR families) appear to regulate these processes. Recent molecular genetics studies have identified a number of putative and known regulators of metal toxins.[64]

Table 1.2 shows the MerR, ArsR, DtxR, Fur and NikR families of metallo-regulators and example metals that they regulate.

Table: 1.2. Main metallo-regulatory families of proteins in prokaryotes

Family	Members	Example metals regulated
MerR	MerR	Hg
	CueR	Cu
	ZntR	Zn
	PbrR	Pb
ArsR/SmtB	ArsR	As, Sb
	CadC	Cd, Pb
DtxR	SmtB,CzrR	Zn
	NmtR	Ni
	BxmR	Cu, Zn
Fur	IdeR	Fe
	MntR	Mn
	Zur	Zn
NikR	NikR	Ni

Source: Adapted from Pennella *et al.*[64]

The current literature sheds light on metal-specific regulators that sense distinct metal ions in four of the five protein families listed in Table 1.2. The NikR family,

for instance, has regulators bearing only nickel. Similarly, the MerR (ZntR), ArsR (SmtB and CzrA) and Fur (Zur) families contain zinc-specific metal sensor proteins. Detailed information is available regarding the ability of different types of protein from the same family to distinguish a particular metal ion from other metal ions. The negative allosteric regulation of the promoter/operator DNA by metal binding is stressed, with the subsequent transcription of the resistance/uptake operons. [64]

1.4.1. ArsR/SmtB

The ArsR/SmtB (or ArsR for short) family has been extensively studied and comprises what is plausibly the largest and most diverse family of metallo-regulatory proteins. [114] It is named after the *Escherichia coli* As(III)/Sb(III) sensor ArsR and the *Synechococcus* PCC7942 Zn(II) sensor SmtB. Members of the family appear in practically all bacterial taxonomies; the number of ArsR-encoding genes appears to be greater than 500. The species *Mycobacterium tuberculosis* and *Streptomyces coelicolor*, for instance, encode 10 and 13 ArsR proteins, respectively. [115]

Members of this family can sense diverse metal ions, from essential metals Zn(II) and Ni(II), for instance, to toxic pollutants As(III), Cd(II), and Pb(II) (for examples, refer to Table 1.2). The ArsR family regulates genes that affect the effluxing, scavenging, or detoxification of excess metal ions within the intracellular fluid (or cytosol). [116]All ArsR/SmtB proteins are dimeric and consist of a similar fold with a winged helix-turn-helix motif (α3-turn-αR) which plays a role in DNA binding. Also indicated on the structure of a representative ArsR/SmtB repressor, *S. aureus* pI258 CadC, [117] are the apo-structure of the α3N Cd(II)/ Pb(II) sensor of *S. aureus* CadC [117] as well as a solution structure of the Cd(II)-bound α4C Cd(II)/Pb (II) sensor *M. tuberculosis* CmtR. [118] The CadC and *Synechococcus* SmtB can be illustrated as being flat or open. Their winged helical domain forms an important component of the dimer. The N-terminal α1 and C-terminal α5 helices are the main interface of the dimer (Figure 1.7).

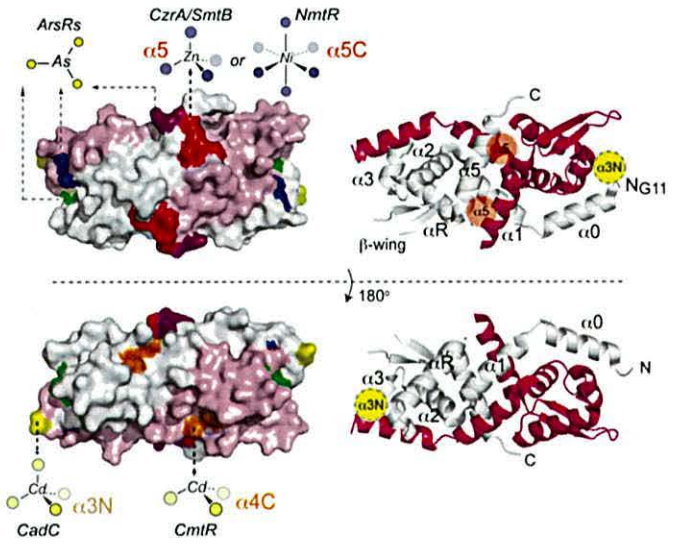


Figure: 1.7. The known binding areas of the ArsR/SmtB family of repressors for the structure of the *S. aureus* pI258 CadC homodimer. [117]

Left: two views of space filling models carrying a ribbon on right; a single promoter is shaded in pink and the other is grey, with the α-helices labelled consecutively from the N-terminus, α0-α5, of the ribbon diagrams. The figure also gives the schematic locations of the α3N (yellow) and α5 (shaded red) sensing sites on every side of the dimer. [117] On the left, the estimated regions and schematic renderings of representative coordination complexes of distinct sensing sites are given, which correspond to *S. aureus* pI258 CadC (yellow, α3N), *M. tuberculosis* CmtR (orange, α4C), *E. coli* plasmid R773 ArsR (green, α3), *S. aureus* CzrA/*Synechococcus* SmtB and *M. tuberculosis* NmtR (red, α5 and α5C, respectively), *C. glutamicum* ArsR1 (blue) and *A. ferrooxidans* ArsR (purple).

The winged helix domain is also packed against the N-terminal α0 helix, although the winged helix forms a folded subdomain in other metal sensor families. In ArsR-family sensors, metal-binding residues are virtually always taken from opposite protomers within the homodimer, which results in pairs of symmetry-related metal sites. An example of this is the metalloregulatory α5 sites that make use of the ligands from adjacent N- and C-terminal regions of the α5 helix (Figure 1.8). Similarly, cysthiolates are used by the α3N and α4C Cd/Pb binding sites. Cysthiolates are also derived from the specific α-helices present in the core of the molecule (α3 or α4) as well as the N-terminal and C-terminal tails of the opposite protomer. The quaternary structural transitions in the dimer play a vital role in achieving allosteric negative regulation of DNA operator binding by inducing metal ions. Placement of these sites over the dimer interface is most favourable for achieving these structural transitions. [119]

One of the characteristics of well-characterised individual ArsR/SmtB family members is the presence of regulatory metal-binding sites of the characteristic metal-liganding donor set available in various regions, both inside the primary structure and other projected secondary, tertiary and quaternary structural folds. For instance, the α3/α3N and α4C metal-binding sites use cysteine residues to work with metal ions, which result in attachment of highly polarisable, thiophilic or soft metals like Pb(II), Cd(II) and As(III) (Figure 1.7). In these instances, where Zn(II) is bound to α3N sites to conduct regulation (i.e. in the cyanobacterial Zn(II) sensors *Anabaena* AztR, [114] *O. brevis* Bxm [120] and *Synechocystis* ZiaR), [121] these remnants replace one of the Cys residues to create a S3N donor set. This set of donors can be differentiated from S3 [Pb (II)] and S4 [Cd(II)] donor sites of the associated Cd(II)/Pb (II) sensor, *S. aureus* CadC.[122-124] In the end, what arises to differentiate a trigonal As (III) α3 sensing site[125] from the Cd (II)/Pb(II)-sensing α3N site is metal co-ordination by a key Cys residue from the area of the N-terminal opposite subunit, Cys7 in *S. aureus* CadC. The Cys7 is a known key allosteric residue for Cd(II), Pb(II) and Bi(III), as substitution lowers the capability of CadC to sense Cd(II) in *vitro* [122,126] and in vivo.[127] *Bona fide* ArsRs linked with *E. coli* plasmid R773 lacks this area of the N-terminal and is regulated via As(III) in a DNA binding assay; it is insensitive to the availability or absence of Cys7 binding through inducing metal ions. One of the current exceptions is *Acidithiobacillus ferrooxidans* ArsR (*Af*ArsR): in this organism, a homology model was suggested where co-ordination of three cysteines occurred from the C-terminus of one promoter As (III), but the degree to which this occurs in the actual structure is not yet known.

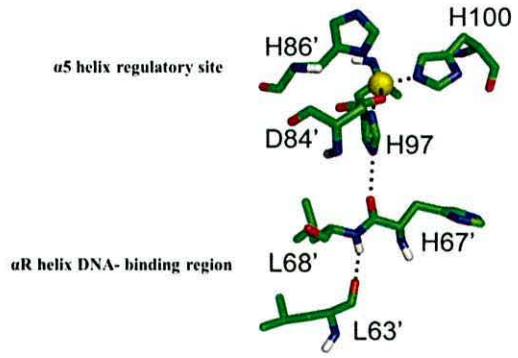


Figure: 1.8. The proposed model for a hydrogen bonding network in Zn (II)-bound CzrA, which is associated with regulatory sites of the DNA binding helices.[128,129]

1.4.2. Fur

The Fur family of proteins sense Fe. Detecting and controlling Fe concentration is important for bacterial metabolism. As Fe concentrations build up to near hazardous levels, excess Fe binds to Fur; this activates the protein to bind with DNA, which then turns off expression of Fe-importing proteins. Nearly all gram-negative bacteria encode Fur in their genomes. Not only does Fur regulate Fe levels within the cell; it is also implicated in the control of oxidative stress and acid tolerance.[130] Some Fur homologues and orthologues can sense other metals. These include an Mn(II) sensor (*Mur*), a Zn(II) sensor (*Zur*), and a Ni(II) sensor (*Nur*). This raises the possibility that, should metals other than Fe activate Fur, Fe importation could be turned off, thereby starving the cell of Fe.[131]

The same protein fold as FurB, along with an N-terminal winged helix, was observed in the crystallographic structures of *P. aeruginosa* Fur,[132] B. *subtilis* PerR33,[133] *S. coelicolor* Nur[134] and *M. tuberculosis* Zur[135] (formerly annotated as FurB) along with a DNA-binding domain linked to a C-terminal dimerisation domain by a flexible linker. The quantity and action of metal sites differ in individual Fur family members, but a consensus might be emerging on a single metallo-regulatory site or area probably shared by the Fur protein, which can adopt a number a co-ordination geometries dictated by metal type.[128,136] The majority of Fur-family repressors possess a structural Zn (II) site that adopts a tetrahedral S4 coordination complex, produced due to four cysteine residues gained from the dimerisation domain; *P. aeruginosa* lacks a structural site in Fur (Figure 1.9).[128,135] The NMR studies conducted on Fur from *E. coli* suggested that Zn(II) bound at this tetrathiolate site firmly stabilised the functional dimer. [137] Examination of the crystallographic structure of *M. tuberculosis* Zur revealed an S4 site in addition to two bound Zn(II) ions that correspond to the location of two sites found in the Zn(II)-complex structures of the Fe(II)-sensor *P. aeruginosa* Fur. Only two metal sites were found after conducting biochemical and spectroscopic experiments on *E. coli* Zur: only protein denaturation can remove the structural S4 and regulatory Zn(II) sites, which adopt a tetrahedral mixed S-N/O coordination complex.[138] This region is analogous to the regulatory site in PerR that brings remnants from the C-terminal and N-terminal domain. These are separate in the apo-repressor (Figure 1.9), in near

approximation. This model orients the relative positions of two NA-binding domains (Figure 1.9), forming a conformation with high DNA binding affinity.[135] Current studies have shown that apo-Fur binding to DNA is based on a single nucleotide change in the DNA sequence. [139] Further biochemical and structural studies are needed to understand the phenomenon underlying the DNA-binding mode in Fur-family regulation.

The H_2O_2 sensor, *B. subtilis* PerR, also carries a structural site for S4 Zn(II) and binds Fe(II) or Mn(II) to a regulatory site that can bridge the domains of N- and C-termini. One of the current reports suggests the existence of crystal structures of an oxidised PerR (PerR-Zn-ox) and PerR-Zn-Mn complex, which favour the allosteric regulation mechanism put forward for *M. tuberculosis* Zur (Figure 1.9). The production of square pyramidal Mn (II) coordination, which is penta-coordinate, carries a regulatory site. In this, the dimer has a greater affinity for binding to DNA, which leads to production of an open coordination site for H_2O_2. The clusters sense H_2O_2 by a PerR-mediated reaction, as compared to other oxidative stress sensors that make use of cysteine residues or [Fe-S] clusters. This is a distinct Fe (III)-catalysed oxidation reaction in which one of the two histidine residues carries a metallic regulatory site, like His 91 which belongs to the dimerisation domain and is transformed into 2-oxo-His, giving rise to an oxidised protein that can bind to DNA (Figure 1.9) The other is His37, which also works in the same manner. [140] Both Zn-His37ox and PerR-Zn-His37A can bind to Mn(II) and show a molecular affinity that is 20-fold lower than seen for wild-type PerR. His37, which is derived from the DNA-binding domain, is a key allosteric residue,[128,141] so that substitution or modification causes failure to properly orient the N- and C-terminal domains for DNA binding. [133] In contrast, His91 oxidation lowers the affinity of Mn(II)/Fe(II) binding to undetectable levels, causing dissociation of the metal and ultimate dissociation form the DNA operator.[133] The complete structure of PerR-Zn-Mn shows a rationality of structure where His37 and His91 are subject to Fe(III)-catalysed oxidation, whereas the Mn(II) ligand His93 is refractory to Fe(III)-catalysed oxidation. An axial position is occupied by His93, which carries an open co-ordination site that is positioned directly and will bind to H_2O_2 but has no access to the locally liberated hydroxyl radical.[136,142]

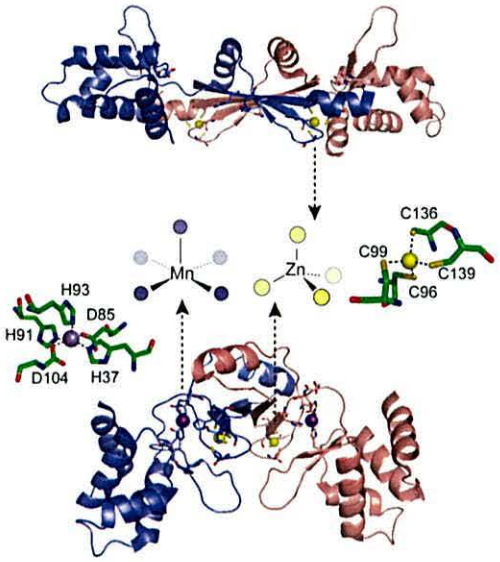


Figure: 1.9. A ribbon diagram of two crystallographic structures of *B. subtilis* PerR. The subunits are shaded in blue and red.

1.4.3. DtxR

The DtxR family of proteins controls the levels of Fe and Mn; as such, it comprises two sub-families: Fe(II) sensors and Mn(II) sensors.[143] The former is found in *Corynebacterium diphtheria* (DtxR) and regulates expression of diphtheria toxin. The latter is found in *Bacillus subtilis* (MntR) and regulates Mn transportation. In effect, because it controls concentrations of Fe within the cell, DtxR performs the same role in Actinobacteria, which are gram-positive, as that played by Fur in gram-negative bacteria.[114]MntR is found in a wide range of gram-positive and gram-negative bacteria. It is also found in certain members of the Archaea kingdom of microorganisms.

Metal-sensing regulators, including those for iron, are important in determining the virulence of certain pathogenic bacteria.[144]

Genes that take part in encoding proteins and mediate the uptake and storage of iron are regulated by *C. diphtheria* DtxR and its homologue IdeR from *M. tuberculosis.*[145] These genes are expressed under certain iron-limiting conditions, whereas an increase in cytosolic iron causes repression mediated by DtxR/IdeR. *In vivo*, this transcriptional response is particular for Fe (II), but *in vitro* experiments show that Ni(II) or Co(II), but not Mn(II), can perform role of activator for binding DNA.

Therefore, Ni(II) or Co(II) are used as co-repressors and are utilised *in vitro* experiments, particularly those involving structural work. [146] None of the structures are reported for DtxR bound to its cognate inducer Fe(II), and the Ni(II)-dependent conformational alterations listed are important but not sufficient to support the robust transcriptional regulatory mechanism of cells. Regulators of DtxR/IdeR carry an N-terminal winged helix DNA-binding domain, which is followed by a helical dimerisation domain and a C-terminal SH3-like domain not present in the Mn (II) sensor MntR.[147] This SH3-like domain is believed to boost the DNA binding affinity through stabilisation of intra- and/or inter-subunit protein-protein interactions.[148,149]

The two different variations of metal-binding site observed in DtxR are designated the primary (regulatory) and ancillary sites. The ancillary sites comprise ligands (His79, Glu83, His98 and two solvent molecules) obtained from the dimerisation helices, while the primary metal site also includes a residue from the N-terminal α-helix in the DNA-binding domain.[150] When a metal binds to a site other than its binding site or ancillary site, various structural changes take place both in the binding metal and in the binding site involved; for example, changes occur in the N-terminal residue(s) of the protein that take part in the primary binding site and a water-mediated interaction with Leu4 as a result of an α1 helix-coil transition in the bound metal.[147,151]This N-terminal unstructured residue in the apo-protein also halts the binding of DNA, especially in unfavourable steric clashes; these changes in the N-terminal helix can be crucial for Ni(II)-dependent allosteric activation of DNA-binding.[147] A slight domain closure in the N-terminal DBNA is noticed in metal-binding domains in the dimer, giving rise to a suitable conformation for DNA recognition.[147] The level of global quaternary structural conformations the Ni(II)-activated and the apo-protein state of DtxR or IdeR dimer vary in solution and are not yet known. The extent to which domain closure (akin to that which occurs mainly in PerR[133] and perhaps most other Fur family regulators)[135] takes part in allosteric activation is not determined yet.

The *B. subtilis* MntR is the founding member of the DtxR-family regulators. MntR is capable of regulating the process of transcription and shows a greater affinity for manganese uptake systems encoded by *mntABCD* and *mntH*.[143] The expression is diminished by MntR when the levels of cytosolic Mn(II) are high, in a manner that is

specific for Mn(II) and Cd(II) over other divalent metal ions like Ca(II), Fe(II), Mg(II), Co(II), Ni(II) and Zn(II). [143,152] Atomic resolution crystallographic studies of wild-type and mutant MntRs attached to many metal ions show an overall architecture quite similar to the Fe(II)-bound MntR structure, which shows a binuclear Mn(II) cluster developed by two ions of Mn(II) ions situated at a distance of 3.3 Å. These were termed MnA and MnB, and they both possess an octahedral or distorted octahedral coordination geometry.[153] The crystals formed at room temperature exhibit Mn(II) binding sites separated by 4.4 Å, known as MnA and Mnc.[152] The co-ordination geometry of every individual Mn(II) ions is the same, and shows a great variation in two structures of the bridging and bidentate ligands. The newly produced MnA-MnC conformer is more consistent with the Mn (II)-Mn (II) distance given by solution EPR studies and this is believed to be related biologically.[154] Some structures, like Cd(II)- , Ca(II)-, and Zn(II)-bound MntR, are also resolved.[152] In the structure, only a single Zn(II) is attached at a particular area, as in MnA, but it adopts a non-native tetrahedral coordination geometry. This indicates the non-effectiveness of Zn(II) as an allosteric activator of operator DNA binding.[147] The MeA-MeC (Me=metal) structures are formed by both Ca(II) and Cd(II); these are structurally analogous to those produced by Mn(II). This explains why Cd(II) acts as an *in vivo* effector. A slight contradiction occurs with Ca(II), where a relationship is expected between co-ordination geometry and the metal sensitivity. A current report[155] suggests that the morphology of apo-MntR permits a contrast with metal-bound state that takes part actively in binding DNA.

These studies help in supporting the proposed mechanism for allosteric regulation of operator binding of MntR by Mn (II).[155] This shows that the DNA binding at N-terminal domains present in the apo-protein dimer have the ability to achieve a variety of orientations relative to the dimerisation domain.. The EPR studies conducted on a spin-labelled MntR support this assessment. The domain closure is conducted by the α4 helix that links the two domains and implicates many chief residues that co-ordinate with both Mn (II) ions. The ligands of Mn (II) from the N-terminal α-helix provide a major contribution to the conformational change; this resembles the model put forward for DtxR and IdeR.[155] The MntR dynamics probed due to exchange of hydrogen-deuterium mass spectrometry showed that the amide

groups in the α4 helix are protected dimers, so that any exchange with solvent upon Mn (II) binding leads to global rigidification of the whole protein. This is thought to lower the entropic price of DNA binding.[156]The findings indicate shared analogous characteristic dynamics of AntR, an MntR homologue from *Bacillus anthracis.* Metal-binding reduces the area between the two DNA-binding helices in the dimer and in the backbone.[157]

1.4.4. NikR

NikR was first identified in *E. coli*.[158] It regulates expression of those proteins implicated in Ni uptake; it also regulates the transcription of those enzymes that require Ni.[114] Diverse microorganisms require Ni as a metallo-enzyme cofactor (i.e., to activate enzymes).[159] However, Ni can damage cells;[160] thus, cells control levels of Ni within them.

NikR not only binds Ni(II), it also binds with other divalent metals [e.g. Cu (II), Zn (II), and Cd (II)].[161] Binding of some ions, such as Ni, appears to induce less protein denaturation and thereby less damage,[162] which suggests the possibility of two binding sites for two different types of regulation. NikR has been found in microbial species other than *E. coli*, and is arguably the most studied of all metallo-regulators.[114] The research has identified three allosteric states: metal free-, Ni (II)-, and DNA-Ni (II)-bound conformers are available.[114]

The transcription of the *nik* operon (*nik*ABCDE) is regulated by *E. coli* NikR which encodes a greater affinity nickel-specific uptake ABC-transporter. The Ni(II)-bound NikR attaches to the *nik* operator-promoter DNA with greater affinity and decreases transcription under conditions of replete Ni(II). The apo-NikR is also weakly and non-specifically bound to the operator, making Ni(II) an obligate compressor.[163] The crystalline structure of apo-NikR demonstrates that the C-terminus has a tetrameric core flanked via canonical N-terminal RHH DNA-binding domains (Figure 1.10).[164] The tetrameric regulatory domain of the C-terminus is morphologically homologous to that of aspartokinase chorismate mutase (ACT), and the TyrA-domain. This serves the purpose of a small effector molecule and sensing of amino acids allosterically, which regulate the enzyme's role. Better knowledge about the allosteric regulation mechanism of NikR by Ni(II) ions will highlight any common regulatory

characteristics of the ACT-domain enzymes and proteins.[165]

Biochemical studies show that NikR carries two sets of Ni(II)-binding sites: one high affinity binding site situated in the tetrameric C-terminal regulatory domain and a low affinity binding site which remains the subject of ongoing investigations. Allosteric activation of nickel operator-promoter binding occurs when Ni(II) binds to four symmetry-related high affinity regions in the tetrameric domain of ATC. The low affinity regions that are occupied show orientation of two DNA binding domains to the tetramer in a close cis-type conformation, which further enables the affinity of DNA binding chiefly on entropic grounds (Figure 1.10).[166]

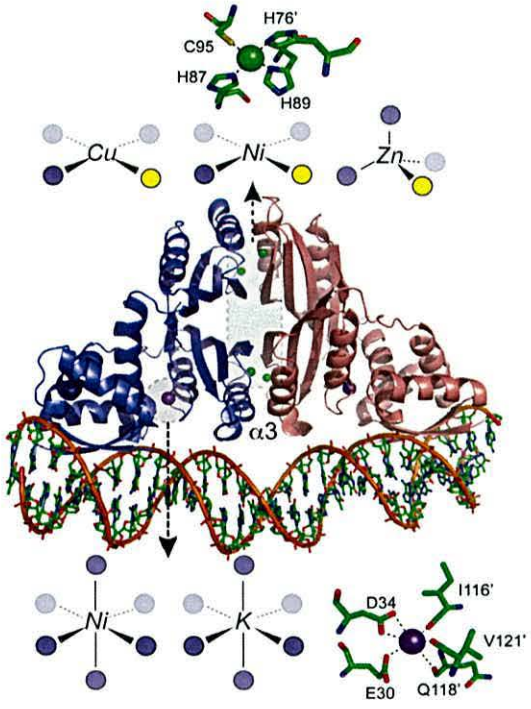


Figure: 1.10. A crystallographic structure of the *E. coli* NikR-DNA complex.

The *E. coli* NikR-DNA complex along with Ni (II) (shaded green) binds with greater affinity at C-terminal sites and K+ ions (shaded purple) bind at the low affinity sites near the DNA. Schematic representations of the coordination complexes formed by other metal ions bound in each site.

In the beginning, the structure of 100% Ni(II) bound NikR had not been crystallographically determined, but the structure of the isolated C-terminal domain attached to Ni(II) gave detailed information on co-ordination of the structure of the high affinity site as well as the conformational change in the domain upon Ni(II)

binding.[164] The Ni(II) adopts an N3S square planar co-ordination geometry, which is supported by low-spin d^8 Ni(II), produced due to three histidine residues(His87, His89 and His76) and one cysteine residue (Cys95) across the tetramer interface (Figure 1.10). NikR responds to hydrogen bonds in its surroundings. Hence, the different structural changes taking place in the DNA ultimately lead to the appeasement and stabilisation of nickel molecules. The fully formed α3 helix is produced and this stabilises Ni(II) binding.[164]The structure of the Ni(II)-NikR-DNA complex also provides additional insights into Ni(II) regulation through plausible mechanisms of allosteric activation and the nature and quantity of low affinity Ni(II) regions which can promote the affinity of NikR for the operator.[167] The complex structure, along with the structure of Ni(II)-bound full length NikR and apo-protein structure, clearly indicate that dramatic conformational changes occurring in tetramer are related to these allosteric states.[168]This allosteric model of Ni(II) binding to a greater affinity square plane of the Ni(II)-bound full length NikR co-complex produces a localised changes in the structure, which incorporates loop and helix (α3) formation.

These changes are important but they do not entirely allow NikR to bind to the operator. A dramatic orientation of the RHH domain is formed by the low affinity NI (II) binding site on the tetramer-DNA complex; this is entirely based upon the structure[169] as this domain adopts a closed *cis*-type conformation (Figure 1.10). Inside the Nikk-DNA complex, these sites with less affinity are denoted by octahedral co-ordination geometry and are completely filled with potassium (K+) ions rather than the expected Ni (II) ions. This site for co-ordination is unusual and inside the ligand it has mainly side-chain backbone carbonyl oxygens, which arise from the C-terminal metal-binding domain as well as the N-terminal DNA-binding domain. This outcome is consistent with the key role played by this region in causing conformational changes in an optimised high DNA-binding affinity state (Figure 1.10).[168]

Many studies have examined the regulatory metal binding in NikR due to the two sets of regulatory metal-binding sites carrying a high affinity square planar Ni(II) regions in the C-terminal domain. These studies on metal binding have shown that the high affinity site is capable of co-ordinating many other divalent metal ions with

an affinity ranking that follows the Irving-Williams series i.e., Mn(II)<Co(II)<Ni(II)<Cu(II)≥Zn(II). The two divalent ions, Cu(II) and Zn(II), which are present in large amounts in the body, have strong binding affinities for each other and for NikR. This binding also serves as one of the interesting mechanisms involved in their action in the body. The expected co-ordination geometry, which covers metal availability in right oxidation state, works in collaboration as a key determinant for the biological metal selectivity. At present, x-ray absorption spectroscopy has shown that various metals work in different co-ordination geometries at the high affinity site, with square planar co-ordination geometry produced for Ni (II) and Cu (II), octahedral for Co (II), tetrahedral for Zn (II) and trigonal for reduced Cu (I) (Figure 1.10).[170] The coordination geometry of the low affinity sites were examined by formulating and characterising bimetallic NikR samples bound to DNA. These data showed that when Cu (II) occupies the high affinity sites, there is an adoption of low affinity sites by an average octahedral (N/O) 6 coordination geometry with Ni (II); this gives the first structural insight into the structure of the low affinity sites when attached to Ni (II) (Figure 1.10).

The biochemical studies show that only Ni (II) and Cu (II) adopt similar square planar co-ordination geometry and bring about the conformational changes needed for *in vitro* allosteric activation. This gives rise to a compelling correlation between the production of right co-ordination geometry and a metal-specific allosteric response. Under the reducing situations in the cytosol and with free Cu available in the Cu (I) oxidation state, the content of Cu (II) becomes small. The presence of Cu (I) is likely to be less because of the action of copper chaperones and metallo-regulatory proteins that bind Cu (I) with very high affinity, which highlights the fact that Cu(I) adopts a non-native original coordination geometry in NikR, further ensuring that NikR will be selective for Ni(II) in the cytosol.[170]

1.4.5. MerR

The term *MerR* refers to a diverse family of regulators that includes proteins that respond to various signals; as such, it includes metallo-regulatory proteins.[171]MerR was among the first metallo-regulatory proteins to be described.[172]

Table: 1.3. Number of MerR family genes in a selection of bacterial species.

Species	NO. MerR genes	Species	NO. MerR genes
Acidithiobacillus ferrooxidans	4	*Haemophilus influenza* KW20	4
Agrobacterium tumefaciens	6	*Lactococcus lactis*	4
Bacillus stearothermophillus	6	*Magnetospirillum magnetotacticum*	9
*Bacillus halodurans*C-125	6	*Mycobacterium avium*	4
Bacillus subtilis	10	*Mycobacterium smegmatis*	6
Bordetella bronchiseptica	5	*Pseudomonas aeruginosa*	6
Bordetella parapertussis	5	*Pseudomonas putida* KT2440	10
Bordetella pertussis	4	*Ralstonia. metallidurans*	9
Brucella melitensis biovar. suis	4	*Rhodobater sphaeroides*	6
Burkholderia cepacia	12	*Salmonella enterica subsp. enter*	3
Clostridium acetobutylicum	3	*Salmonella paratyphi*	4
Desulfitobacteriumhafniense	7	*S. melilon*	23
Escherichia coli K12	5	*Thermomonas porafusca*	8
E.coli 0157:H7 ELD9333	4	*Vibrio cholera* EI T or N16961	5

Source: Brown *et al.*[171]

The family is characterised as a group of transcriptional factors that control metal ion, radical, and small organic molecule concentrations inside bacterial cells. The MerR proteins are present in most bacterial genomes (see Table 1.3).

MerR proteins typically regulate defensive systems against toxic or high levels of metal ions. Members of the family are known to sense and control levels of Cd^{2+}, Zn^{2+}, Co^{2+}, Cu^{+} , Ag^{+} , Au^{+} , Hg^{2+}, and Pb^{2+}ions inside bacteria.[173] The MeR family includes several groups: Hg^{2+} (MerR), Cu^{+} (CueR), Cd^{2+} (CadR), Pb^{2+}(PbrR), and Zn^{2+} (ZntR).[66]

MerR is associated with mercury resistance *operons* (i.e. functioning units of genomic DNA controlled by a single regulatory signal); these are present on transposable elements of DNA (e.g., Tn21 and Tn501).[174] Members of the MerR family detect metal ion excess and promote the transcription of metal resistance genes. Typically, this regulation utilises promoter elements, unusually long promoter spacing, and MerR-like regulators that bind to a site between the –35 and –10 promoter elements; this induces a remodelling of the promoter region of the DNA (Figure 1.11).[175]

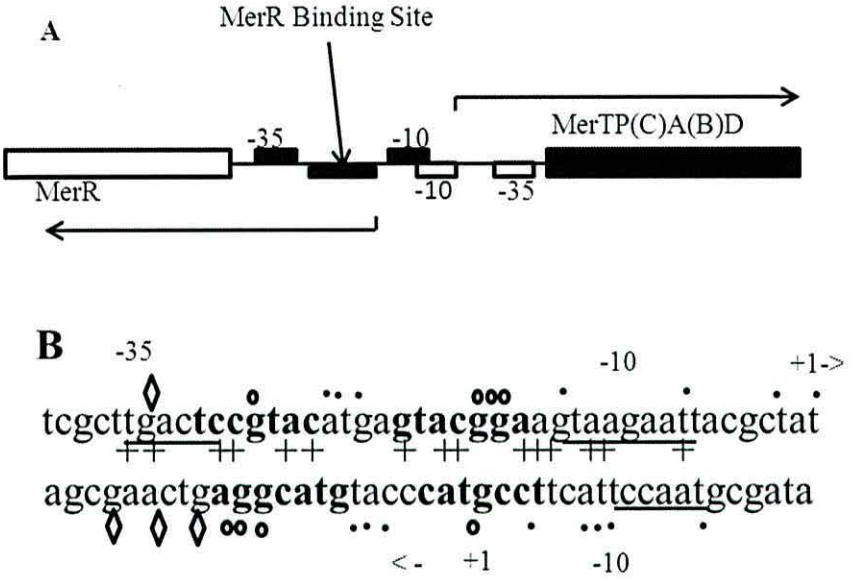


Figure: 1.11. Tn21/Tn501 Mer operator-promoter region (Boldface type indicates palindrome). Source: Summers.[175]

Members of this family include the MerR (Hg resistance regulator), CueR (Cu efflux regulator), ZntR (Zn transport regulator), PbrR (Pb resistance regulator), CoaR or CorR (Co responsive regulator), and CadR (Cd resistance regulator) (See Table 1.4).

Table: 1.4. Members of the MerR family

Protein	ID	Organism	Metal inducers Ligands
MerR	P22853	*B. cereus*	Hg^{2+}
CueR	1Q05(pdb)	*E. coli*	Cu^{+},Ag^{+}, Au^{+}
ZntR	1Q0A	*E. coli*	Zn^{2+}
PbrR	YP145623	*Cupriavidus metallidurans* CH34	Pb^{2+}
CoaR	NP442632	*Synechocystis* PCC 6803	Co^{2+}
CadR	AAK48830	*Pseudomonas putida* 06909	Cd^{2+}, Zn^{2+}, Hg^{2+}

Source: Mergeay *et al.*[176]

The proliferation of MerR family members is illustrated by *Cupriavidus metallidurans* CH34 (a gram-negative bacterium); this organism is metal-tolerant and appears to possess at least seven metal-sensing paralogues associated with efflux and resistance systems[176] Similarly, the large genome of the gram-positive *Streptomyces coelicolor* bacterium may contain 23 MerR homologues.[171]The best described MerR proteins come from Hg resistance operons encoded on the Tn21 and Tn501 transposons. The Tn21 MerR operon comprises five structural genes (MerTPCAD), from which MerR is divergently transcribed. MerTPCAD are implicated in Hg detoxification by a mechanism that, curiously, involves Hg import into the cell. MerP binds the Hg in the periplasm and transfers it to MerT (which is a membrane-bound transporter); this delivers Hg^{2+} to MerA. The MerA then reduces the Hg^{2+} to the volatile form Hg^{0}. The Hg^{0} then diffuses through the cell membrane and so disperses into the environment.[174]See Figure 1.12

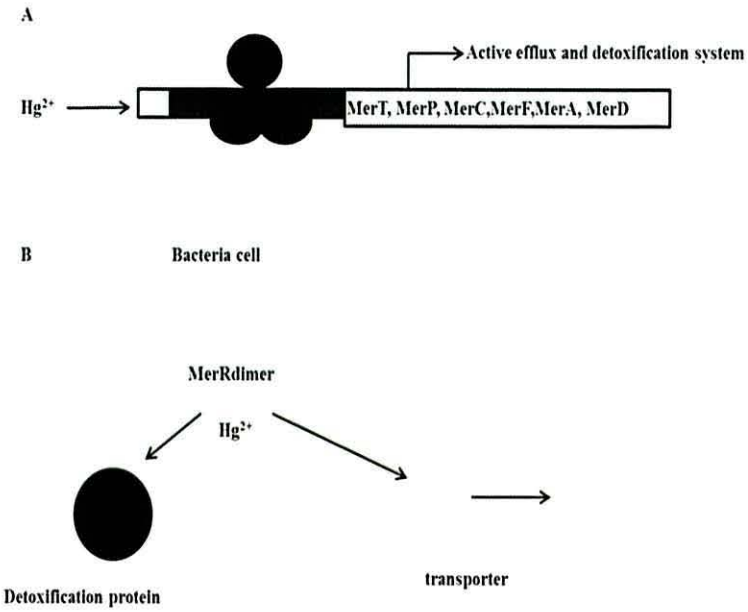


Figure: 1.12. Function of the MerR family of proteins.

(A)The archetype of the MerR family of transcriptional activators is the regulation of gram-negative mercury resistance. Addition of Hg^{2+} to the MerR dimer will enable RNA polymerase (RNAP) to form an open complex with the promoter sites and thus initiate the transcription of the downstream mercury(II) detoxification and efflux genes. (B) Upon binding to the target molecule (Hg^{2+}), MerR will induce a conformational change that sends a signal to activate metal detoxification systems in bacterial cells, including detoxification proteins and membrane transporters.

The MerR family of proteins thus controls metal ion, radical, and small organic molecule concentrations inside bacterial cells.[171,177-180]As such, the proteins act as defences against toxins and high concentrations of metal ions. The binding of metal ions to the protein causes a conformational change, and this in turn activates metal detoxification or efflux genes.[66] The prototype is the MerR protein, which activates a Hg defence system upon binding to Hg^{2+}ions.[181]MerR type proteins also regulate the efflux of toxic organic molecules. For instance, the transcription activator BmrR in the *bmr* gene of the gram-positive bacterium *Bacillus subtilis* recognises diverse organic drug molecules and, in doing so, activates multidrug transporters.[180,182] Similarly, SoxR, a MerR-type protein present in, *E. coli*, activates defence mechanisms against oxygen radicals.[171]

MerR proteins can recognise Hg^{2+} at a concentration as low as 10^{-8} M, and may do so even when millimolar concentrations of competing small molecular thiol ligands are

present.[179,183] Similarly, CueR responds to Cu^{1+} concentration of 10^{-21} M. MerR proteins typically exhibit greater than 100-fold binding selectivity towards target metal ions over other metal ions. Thus, the MerR family is one of the best templates for the development of selective metal ion bioprobes.

Changela and colleagues tested the metal sensitivity and selectivity of the MerR family by solving the x-ray crystal structures of *E. coli* CueR and *E. coli* ZntR; these discriminate between metal ions with M^{1+} and M^{2+} charge.[184] The authors determined the structures of Cu^{+}-, Ag^{+}-, and Au^{+}-bound forms of CueR to 2.2, 2.1, and 2.5 Å resolution, respectively; and they determined the structure of an N-terminal truncated fragment of ZntR bound to Zn^{2+} to 1.9 Å resolution.[184] The overall structure of the CueR dimer was the same in all three metal-bound states. Each monomer could be divided into three functional domains: (a) a dimerisation domain; (b) a DNA-binding domain that flanks the first; and (c) a metal-binding domain.

The structures of the dimerisation and DNA-binding domains appear typical of the MerR family. They also have similar topology to that of two other MerR-family transcriptional activators, BmrR[179]and MtaN;[183,185] these are sensitive to the presence of organic substrates. (See Figure 1.13).

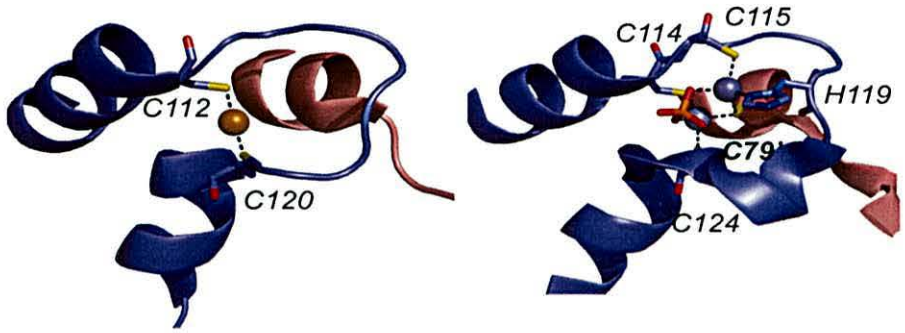


Figure: 1.13 The crystal structures of a MerR family transcriptional regulator protein.

The metal-binding loops of the metallo-regulator MerR family are given by a ribbon. In all situations, only a single symmetry-related metal site is present shaded blue and a single promoter shaded red; the indicated donor ligand is shown in stick form. The structures shown are the Cu(I)-bound form of *E. coli* CueR and the Zn(II) sulphate anion (shown in red/orange)-bridged binuclear structure of *E. coli* ZntR.[114]

Sources: Ma *et. al.*[114]

Kidd and colleagues performed a phylogenetic analysis of the MerR family. The PbrR-related regulators comprised a distinct clade from that of metal-responsive MerR regulators[186] (see Figure 1.14).

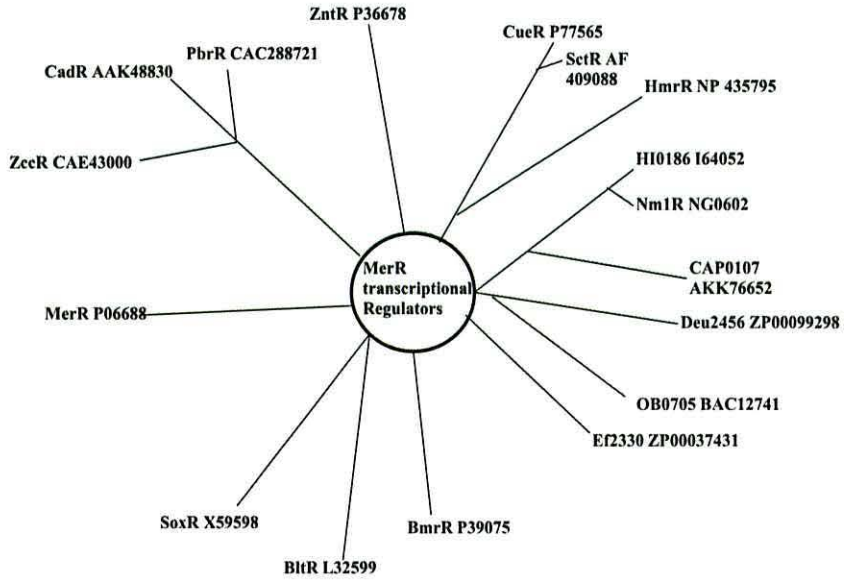


Figure: 1.14.The unrooted phylogenetic tree of MerR transcriptional regulators.

Source: Kidd et. al.[186]

1.5. PbrR protein

The PbrR protein regulates the Pb(II) efflux pumps implicated in Pb detoxification inside *R. metallidurans* (Figure1.15).[187] Research suggests that PbrR binds Pb(II) nearly 1000-fold more selectively over other metal ions.[188]This protein, along with its two close homologues (also in *R. metallidurans*), is the only known naturally occurring Pb(II)-specific regulatory protein.

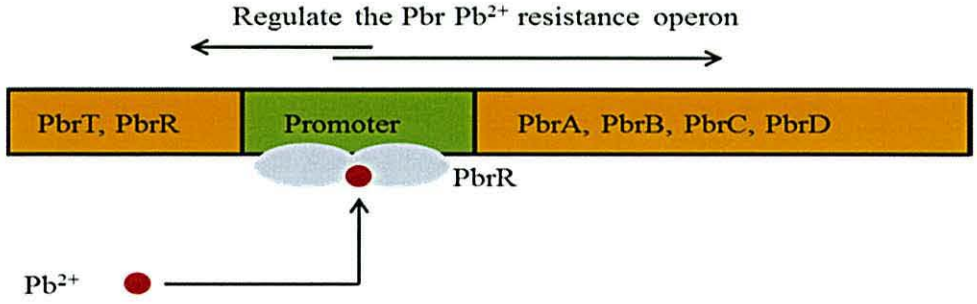


Figure: 1.15. The Pb resistance operon in *Ralstonia metallidurans* strain CH34 (PbrR).

The resistance is regulated by PbrR, a protein that mediates Pb^{2+}-inducible transcription from its divergent promoter.[188]

Source: Chen. *et.al.*[188]

1.6. Advanced spectroscopic methods for studying lead proteins.

1.6.1. Atomic absorption spectroscopy

Ideally, any measurement technique should be accurate, in expensive and easy to use. Fortunately, at least as with regards to Pb and other metal pollutants, atomic absorption spectroscopy (AAS) meets all three criteria.[189] Although several variants of AAS now exist, the principle behind AAS is fairly simple. Atoms within an analyte (substance to be analysed) can be promoted to excited states, albeit only for a few nanoseconds, by their absorption of energy. The wavelength of the light absorbed is specific to each element and the amount of light absorbed is proportional to the concentration of the analyte.[190,191]

1.6.2. NMR spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is a technique that utilises the magnetic properties of atomic nuclei. NMR is used to determine the properties of atoms and molecules. It is commonly used to determine the physical and chemical properties of organic molecules; however, NMR spectroscopy may be used to investigate any atomic nuclei possessing spin. Its use has had a profound effect on the sciences because of the wide range of samples to which it can be applied.[192]

When in a magnetic field, nuclei (e.g. ^{13}C) absorb photons at a characteristic

frequency of the isotope. Different nuclei resonate at different frequencies as a function of their nuclear magnetic moments. Different isotopes of the same element have different nuclear magnetic moments.[193]

1.6.2.1. Lead NMR-Spectroscopy

^{207}Pb NMR is a powerful tool in the study of metallo-enzymes and model complexes; this is because it characterises the interaction of the ligands and the metal in aqueous environments and realistic physiological conditions. Nevertheless, Pb NMR is not widely used, due largely to the difficulties in the measurement of Pb NMR spectra, [194] which show a wide range in the chemical shifts (see Table 1.5).

Table: 1.5. ^{207}Pb NMR chemical shifts reported for lead (II) coordination sites containing sulphur ligands

Coordination environment	Chemical shift
PbS_3	2818–2868
PbS_2N	2873
$PbSN_2$	2357
PbS_2O_2	1506–1555
PbS_3O_3	1422–1463

Source: Mah and Jalilehvand.[195]

1.6.2.2. ^{207}Pb NMR spectroscopy on biological molecules.

Vogel and colleagues used ^{207}Pb NMR spectroscopy (using isotopically enriched ^{207}Pb) to study Pb^{2+} binding to the Ca^{2+} site of calcium-binding proteins, including calmodulin (CaM).[196] This appears to be the only example of the use of ^{207}PbNMR as a probe in metallo-proteins. No ^{207}Pb spectra have been reported for sulphur-rich metallo-proteins. However, some small synthetic molecules, with or without mixed O, S, and N donor ligands (e.g. S_2O_2, S_2N_2, N_2O_4, N_3O_3, N_4, N_6), have been characterised.[42,197-201]

The ^{207}Pb NMR signal for the thiol-rich binding sites is shifted further downfield than is that of the O- and N-rich Ca-binding sites in proteins. [202] Thus, distinguishing

between possible PbS3 and PbS3O coordination was easy.[197] The coordination number and geometry of the Pb^{2+} ion can also be analysed.[42] Dean and colleagues have synthesised [Ph_4As] [$Pb(SPh)_3$] and have characterised complexes in non-aqueous media using ^{207}Pb NMR spectroscopy.[203-205] Neupane and Pecoraro presented ^{207}Pb NMR spectra for a physiologically relevant coordination environment of thiolate-rich metallo-peptides in the preferred homoleptic trigonal pyramidal geometry for Pb ions; these authors used three-strand coiled-coil peptides. This appeared to be the first report of ^{207}Pb NMR spectroscopy used in a Cys3 motif and it promised to be a direct probe for those thiol-rich metallo-enzymes (e.g. ALAD) that are implicated in human Pb poisoning.[202]

^{207}Pb NMR revealed a Pb resonance at 2818 ppm; this was further downfield than Pb complexes with N and O ligation. There is evidence of three-coordinate Pb-binding in proteins with cysteine-rich metal-binding sites, and this appears to be a model for Pb sites in cells.[206]

This preference plausibly arises from the high affinity of Pb for the thiolate functional group in cysteine and may reflect the high enthalpy of formation for Pb-S bonds.[195] Recent research suggests that the preferred coordination mode of Pb in thiol-rich sites in proteins is three-coordinate. Pb plausibly assumes a trigonal-pyramidal geometry in these sites, with the fourth "open" coordination site of the tetrahedron occupied by the stereochemically active $6s^2$ lone pair.[63]

1.7. Scope and aims of this work

PbrR691 from *Ralstionia metallidurans* is a lead regulatory protein which has a high selectivity for Pb^{2+} ion. The aim of this project is to better understand the structural features that lead to high selectivity for lead by:

(a) using a bioinformatics approach to select an appropriate homologue for PbrR691 protein,
(b) cloning and overexpressing the target protein and mutants,
(c) mutating specific active site cysteine residues (Cys77, Cys112 and Cys119) to single, double and triple mutants based on cysteine to alanine or cysteine to aspartic acid mutations.

(d) Charactersing the lead binding properties of the recombinant native and mutant proteins site using a mixture of biochemical and spectroscopic methods, including ^{207}Pb NMR, UV/Vis and atomic absorption spectroscopy.

1.9 References

1. I. Sherameti and A. Varma, *Soil heavy metals*, Heidelberg: Springer Verlag. 2010.

2. G. Brewer, Copper toxicity in the general population, *Clinical Neurophysiology*, 2010, **121**, 459-460.

3. J.Duffus, IUPAC Heavy metals a meaningless term? *Pure and Applied Chemistry*, 2002, **74,** 793–807.

4. S. Gaur, M. Joshi, H. Jos and S. Saxena, Analytical study of water safety parameters in ground water samples of Uttarakhand in India, *Indian Journal of Public Health Research and Development,* 2013, **4**, 185-189.

5. D. Nies, Microbial heavy-metal resistance, *Applied Microbiology and Biotechnology,* 1999, **6**, 730.

6. J. Hall, Cellular mechanisms for heavy metal detoxification and tolerance, *Journal of Experimental Botany,* 2002, **53**, 1-11.

7. S. Lippard and J. Berg, *Principles of Bioinorganic Chemistry*, Mill Valley, CA: University Science Books, 1994, 103–174.

8. O. Duruibe, C. Ogwuegbu and N. Egwurugwu, Heavy metal pollution and human biotoxic effects, *International Journal of Physical Sciences,* 2007, **2**, 112-118.

9. X. Wang, T. Sato, B. Xing and S. Tao, Health risks of heavy metals to the general public in Tianjin, China via consumption of vegetables and fish, *the Science of the Total Environment,* 2005, **350**, 28-37.

10. G. Vastag, E. Szocs, A. Shaban and E. Kalman, New inhibitors for copper corrosion, *Pure and Applied Chemistry,* 2001,**73**, 1861-1869.

11. M. Ibrahimpasic, A. Strkalj, A.Vukovic, I. Boras and A. Skrobica, Determination of low arsenic concentration in light gasoline by GFAAS technique, *Izvorni znanstveni rad / Original scientific paper,*2013, **52**, 35-40.

12. A. Jakimska, P. Konieczka, J.Namiesnik and K. Skora, Bioaccumulation of metals in tissues of marine animals, part I: The role and impact of heavy metals on organisms, *Polish Journal of Environmental Studies,* 2011, 1117-1125.

13. B. Korenekova, M. Skalicka and P. Nad, Concentration of some heavy metals in

cattle reared in the vicinity of a metallurgic industry, *Veterinarski Arhiv,* 2002, **72,** 259-268.

14. M. Ikram, A. Ismail, V. Yap and N. Azwady, Levels of heavy metals (Zn, Cu, Cd, and Pb) in mudskippers (Periophthalmodonschlosseri) and sediments collected from intertidal areas at Morib and Remis, Peninsular Malaysia, *Toxicological and Environmental Chemistry,* 2010, **92,** 1471-1486.

15. G. Bryan, M. Waldichuk, R. Pentreath and A. Darracott, Bioaccumulation of marine pollutants and discussion, *Philosophical Transactions Royal Society London. B*, 1979, **286**, 483–505.

16. H. Brodie, Y. Suzuki, H. McAdams and L. Andersen, Whole-genome transcriptional analysis of heavy metal stresses in Caulobactercrescentus, *Journal of Bacteriology,* 2005, 8437-8449.

17. M. Wood and K. Wang, Microbial resistance to heavy metals, *Environmental Science and Technology,* 1983, **17**, 582-590.

18. D. Dreher and A. Junod, Role of oxygen free radicals in cancer development, *European Journal of Cancer,* 1996, **32**, 30-38.

19. S. Flora, M. Mittal and A. Mehta, Heavy metal induced oxidative stress and it's possible reversal by chelation therapy, *Indian Journal of Medical Research*, 2008, **128**, 501–523.

20. M. McCally, *Life support: the environment and human health*, Cambridge, Mass: MIT Press, 2002.

21. M. Basha, W. Wei, S. Bakheet, N. Benitez, H. Siddiqi, Y. Ge, K. Lahiri and N. Zawia, The fetal-basis of amyloidogenesis: exposure to lead and latent over-expression of amyloid precursor protein and β-amyloid in the aging brain, *Journal of Neurophysiology science*, 2005, **25**, 823–829.

22. B. Montgomery, Heavy metals and the etiology of Parkinson's disease and other movement disorders, *Toxicology,* 1995, **97**, 3-9.

23. T. Rando, R. Crowley, E. Carlson, C. Epstein and P. Mohapatra, Overexpression of copper/zinc superoxide dismutase: a novel cause of murine muscular dystrophy, *Annals of Neurology,* 1998, **44**, 381-386.

24. R. Bharti and A. Singh, Heavy Metal Concentrations in Pharmaceutical Effluents of Industrial Area of Dehradun (Uttarakhand), *International Journal of Environmental Sciences and Research,* 2013,**2**, 140-145.

25. S. Murcott, *Arsenic contamination in the world: an international sourcebook*, London: IWA Pub, 2012.

26. L. Smedley, Sources and distribution of arsenic in groundwater and aquifers, *International Association of Hydrogeologists Publication*, 2008, 4-32.

27. World Health Organization (WHO) Arsenic in drinking-water. Background document for preparation of WHO Guidelines for drinking-water quality, 2003, 308-309.

28. M. Buchauer, Contamination of soil and vegetation near a zinc smelter by zinc, cadmium, copper, and lead, *Environmental Science and Technology*, 1973, **7**, 131-135.

29. Agency for Toxic Substances and Disease Registry list of the 10 most hazardous chemicals, 2011, Accessed 16 July 2013 from http://www.atsdr. cdc.gov/ SPL/ index.html

30. F. Oehme, *Toxicity of Heavy Metals in the Environment*, New York: M. Dekker, 1978, **1**, 1-970.

31. T. Clarkson, The pharmacology of mercury compounds. *Annual Review Pharmacology and Toxicology*, 1972, **12,** 375–406.

32. World Health Organization, *Environmental Health Criteria 118: Inorganic Mercury*. World Health Organization, Geneva. 1991.

33. Agency for Toxic Substance and Disease Registry, *Toxicological Profile for Mercury* US Department of Health and Human Services, Public Health Service, Agency for Toxic Substance and Disease Registry, Bethesda, MD, 1999.

34. L. Lash and R. Zalups, Activities of enzymes involved in renal cellular glutathione metabolism after uninephrectomy in the rat, *Archives of Biochemistry and Biophysics*, 1994, **309,** 129–138.

35. K. Eto, M. Marumoto and M. Takeya, The pathology of methylmercury poisoning (Minamata disease). The 50th Anniversary of Japanese Society of

Neuropathology, *Neuropathology,* 2010, **30**, 471-479.

36. V. Fthenakis and C. Kim, CdTe photovoltaics: Life cycle environmental profile and comparisons, *Thin Solid Films,* 2007, **515**, 5961-5963.

37. Dartmouth, Center for Environmental Health Sciences; Hanover, NH: Toxic Metals Research Program, 2006.

38. NTP. Report on Carcinogens U.S. Department of Health and Human Services, Public Health Service, National Toxicology Program. Cadmium and cadmium compounds; 2004, 42–44.

39. J. Huff, R. Lunn, M. Waalkes, L. Tomatis and P. Infante, Cadmium-induced Cancers in Animals and in Humans, *International Journal of Environmental and Occupational Health*, 2007, **13**, 202–212.

40. L. Jarup, Hazards of heavy metal contamination, *British Medical Bulletin,* 2003, **68**, 167-182.

41. S. Srianujata, Lead the toxic metal to stay with humans, *the Journal of Toxicological Sciences,* 1998, **23**, 237-240.

42. E. Claudio, H. Godwin and J. Magar, Fundamental coordination chemistry, environmental chemistry and biochemistry of Pb(II), *Inorganic Chemistry*, 2003, **51**, 1–144.

43. E. Murphy and H. Gene, Sources of lead in homes served by private wells in Hunterdon County, NJ. Trenton, N.J. New Jersey Dept. of Environmental Protection, Division of Science Research and Technology, 2006.

44. R. Pawar, V. Raut, M. Bhise, G. Murhekar, M. Wadekar, V. Banewar and D. Gulwade, Metal ion complexes of Sr (II), Cd (II), Pb (II), Zn (II) with substituted pyrazoles, isoxazoles and their antimicrobial activities, *Oriental Journal of Chemistry,* 2009, **25**, 1117-1120.

45. C. Andreini, I. Bertini, G. Cavallaro, G. Holliday and J. Thornton, Metal ions in biological catalysis: from enzyme databases to general principles, *Journal of Biological Inorganic Chemistry*, 2008. **13**, 1205-1218.

46. M. Johansson and A. Larsson, Effects of inorganic lead on delta-aminolevulinic acid dehydratase activity and hematological variables in the rainbow trout, *Salmo*

gairdnerii. Archives of Environmental Contamination and Toxicology, 1979, **8**, 419-431.

47. R. Prasanthi, C. Devi, D. Basha, N.Reddy, and G. Reddy, Calcium and zinc supplementation protects lead (Pb)-induced perturbations in antioxidant enzymes and lipid peroxidation in developing mouse brain, *International Journal of Developmental Neuroscience.* 2010, **28**, 161-167.

48. B. Joseph and C. Stroup, *Lead poisoning: exposure, abatement, regulation*, Boca Raton: Lewis Publishers,1995.

49. M. Lone, H. Zhen, P. Stoffella and X. Yang, Phytoremediation of heavy metal polluted soils and water: progresses and perspectives, *Journal of Zhejiang University, Science*, 2008, **9**, 210-220.

50. S. Buckley, R. Sawyer, C. Koshland and D. Lucas, Measurements of Pb vapor and particulate in flames and post-flame gases, *Combustion and Flame*, 2002, **128**, 435–446.

51. A. Manceau, M. Boisset, G. Sarret, J. Hazemann, M. Mench, P. Cambier and R. Prost, Direct determination of Pb speciation in contaminated soil by EXAFS spectroscopy, *Environmental Science and Technology*, 1996, **30**, 1540–1552.

52. P. Shiwani and A. Rai, Screening of brick-kiln area soil for determination of heavy metal Pb using LIBS, *Environmental Monitoring and Assessment*, 2009, **148**, 437-447

53. P. Landrigan, Toxicity of Pb at low dose, *British Journal of Industrial Medicine,* 1989, **46**, 593–596.

54. P. Landrigan, Lead in the modern workplace (Editorial). *American Journal. Public Health*, 1990, **80**, 907–908.

55. P. Landrigan and A. Todd, Lead poisoning, *the Western Journal of Medicine*, 1994, **161**, 153–159.

56. A. Todd, J. Wetmur, J. Moline, J. Godbold, S. Levin and P. Landrigan, Unraveling the chronic toxicity of Pb: an essential priority for environmental health, *Environmental Health Perspectives*, 1996, **104**, 141–146.

57. C. Zhang, Y. Tang, L. Luo and W. Xu, Outlier identification and visualization for

Pb concentrations in urban soils and its implications for identification of potential contaminated land, *Environmental Pollution*. 2009, **157**, 3083-3090.

58. EEA Technical report 8/2012, European Union emission inventory report 1990–2010 under the UNECE Convention on Long-range Transboundary Air Pollution (LRTAP), Copenhagen, 2012.

59. A. Ghering, M. Jenkins, B. Schenck, S. Deo, R. Mayer, M. Pikaart, J. Omichinski and H. Godwin, Spectroscopic and functional determination of the interaction of Pb^{2+} with GATA proteins, *Journal of the American Chemical Society*, 2005, **127**, 3751–3759.

60. B. L. Vallee and D. Ulmer, Biochemical effects of mercury, cadmium, and lead, *Annual Review of Biochemistry*, 1972, **41**, 91-128.

61. J. Zhuang L. Fu, M. Xu, Q. Zhou, G. Chen and D. Tang, DNAzyme-based magneto-controlled electronic switch for picomolar detection of lead (II) coupling with DNA-based hybridization chain reaction, *Biosensors and Bioelectronics,* 2013, **45**, 52-57.

62. L. Shimoni-Livny, J. Glusker and C. Bock, Lone pair functionality in divalent lead compounds, *Inorganic Chemistry*, 1998, **37**, 1853–1867.

63. J. Magyar, T. Weng, C. Stern, D. Dye, B. Rous, J. Payne, B. Bridgewater, A. Mijovilovich, G. Parkin, J. Zaleski, J. Penner-Hahn and H. Godwin, Reexamination of Pb (II) coordination preferences in sulfur-rich sites: Implications for a critical mechanism of Pb poisoning, *Journal of the American Chemical Society*, 2005, **127**, 9495–9505.

64. M. Pennella and D. Giedroc, Structural determinants of metal selectivity in prokaryotic metal-responsive transcriptional regulators, *BioMetals*, 2005, **18**, 413–428.

65. G. Zampella, K. Neupane, L. Gioia and V. Pecoraro, The importance of stereochemically active lone pairs for influencing Pb (II) and As (III) protein binding, *Chemistry a European Journal,* 2012, **18**, 2040-2050.

66. P. Chen and C. He, Selective recognition of metal ions by metalloregulatory proteins, *Current Opinion in Chemical Biology*, 2008, **12**, 214–221.

67. P. Chen, E. Wasinger, J. Zhao, D. Lelie, L. Chen and C. He, Spectroscopic insights into lead(II) coordination by the selective lead(II)-binding protein PbrR691, *Journal of American Chemical Society*, 2007, **129**, 12350–12351.

68. Y. Wang, L. Hemmingsen and D. Giedroc, Structural and functional characterization of mycobacterium tuberculosis CmtR, a PbII/CdII-sensing SmtB/ArsR metalloregulatory repressor, *Biochemistry*, 2005, **44**, 8976–8988.

69. T. Liu, J. Golden and D. Giedroc, A zinc (II)/Pb (II)/cadmium (II)-inducible operon from the *Cyanobacterium anabaena* is regulated by AztR, an alpha3N ArsR/SmtB metalloregulator, *Biochemistry*, 2005, **44**, 8673–8683.

70. L. Busenlehner, N. Cosper, R. Scott, B. Rosen, M. Wong and D. Giedroc, Spectroscopic properties of the metallo-regulatory Cd(II) and Pb(II) sites of S. aureus pI258 CadC, *Biochemistry*, 2001, **40**, 4426–4436.

71. L. Patrick, Lead toxicity: A review of the literature, Part II: The role of free radical damage and the use of antioxidants in the pathology and treatment of Pb toxicity, *Alternative Medicine Review,* 2006, **11**, 114–127.

72. E. Perpetuo, C. Souza and C. Nascimento, Progress in Molecular and Environmental Bioengineering-From Analysis and Modeling to Technology Applications, INTECH Open Access Publisher, 2011, chapter 28.

73. M. Hossain, P. Piyatida, A. Jaime, T. Silva, and M. Fujita1, Molecular mechanism of heavy metal toxicity and tolerance in plants: central role of glutathione in detoxification of reactive oxygen species and methylglyoxal and in heavy metal chelation, *Journal of Botany*, 2012, **2012**, 1-37.

74. C Sen, Nutritional biochemistry of cellular glutathione, *the Journal of Nutritional Biochemistry*, 1997, **8**, 660–672.

75. L. Kromidas, L. Trombetta and I. Jamall, The protective effects of glutathione against methylmercury cytotoxicity, *Toxicology Letters*, 1990, **51**, 67–80.

76. J. Coleman, Zinc proteins: enzymes, storage proteins, transcription factors, and replication proteins, *Annual Review of Biochemistry*, 1992, **61**, 897–946.

77. R. Kaptein, Zinc fingers: *Current Opinion in Structural Biology*, 1991, **1**, 63–70.

78. A. Rich, E. Bombarda, A. Schenk, P. Lee, E. Cox, A. Spuches, L. Hudson, B. Kieffer and D. Wilcox, Thermodynamics of Zn^{2+} Binding to Cys2His2 and Cys2HisCys Zinc Fingers and a Cys4 Transcription Factor Site, *Journal of the American Chemical Society*, 2012, **134**, 10405−10418.

79. H. Meilin, D. Krepkiy, W. Hu and D. Petering, Zn-, Cd-, and Pb- transcription factor IIIA: Properties, DNA binding, and comparison with TFIIIA-finger 3 metal complexes. *Journal of Inorganic Biochemistry*, 2004, **98**, 775–785.

80. A. Hartwig, Zinc finger proteins as potential targets for toxic metal ions: Differential effects on structure and function. *Antioxidants and Redox Signaling*, 2001, **3**, 625–634.

81. B. Krizek, B. Amann, V. Kilfoil, D. Merkle, and J. Berg, A consensus zinc finger peptide: Design, high-affinity metal binding, a pH-dependent structure, and a his to cys sequence variant, *Journal of the American Chemical Society*, 1991, **113**, 4518–4523.

82. B. Krizek, D. Merkle and J. Berg, Ligand variation and metal ion binding specificity in zinc finger peptides, *Inorganic Chemistry*, 1993, **32**: 937–940.

83. J. Payne, M. Horst and H. Godwin, Lead fingers: Pb^{2+}-binding to structural zinc-binding domains determined directly by monitoring lead-thiolate charge-transfer bands, *Journal of the American Chemical Society*, 1999, **121**, 6850–6855.

84. Y. Lolin, P. O'Gorman, An intra-erythrocyte low molecular weight lead-binding protein in acute and chronic lead exposure and its possible protective role in lead toxicity, *Annals of Clinical Biochem*istry, 1988, **25**, 688-697.

85. P. Mistry, W. Lucier and A. Fowler, High affinity lead-binding proteins in rat kidney cytosol: mediate cell-free nuclear translocation of lead, *Journal of Pharmacology Experimental Therapeutics,* 1985, **232**, 462-469.

86. R. Ragavan and H. Gonick, Isolation of low-molecular weight lead-binding protein from human erythrocytes, *Experimental Biology and Medicine*, 1977, **155**, 164-167.

87. G. DuVal and B. Fowler, Preliminary purification and characterization studies of a low molecular weight, high affinity cytosolic lead-binding protein in rat brain,

Biochemical and Biophysical Research, 1989,159, 177-184.

88. R.Smith, W.Khang, B. Quintanilla and B. Fowler High affinity renal lead-binding proteins in environmentally-exposed humans, *Chemico-Biological Interactions,* 1998, **115**, 39-52.

89. P. Mistry C. Mastri and B. Fowler, Influence of metal-ions on renal cytosolic lead-binding proteins and nuclear uptake of lead in the kidney, *Biochemistry and Pharmacology,* 1986, **35**, 711-713.

90. P. Goering and B. Fowler, Regulation of lead inhibition of delta aminolevulinic acid dehydratase by a low molecular weight, high affinity renal lead-binding protein, *Journal of Pharmacology and Experimental Therapeutics*, 1984, 231, 66-71.

91. P. Goering, P. Mistry and B. Fowler, A low molecular weight lead-binding protein in brain attenuates lead inhibition of delta-aminolevulinic acid dehydratase: comparison with a renal lead-binding protein, *Journal of Pharmacology and Experimental Therapeutics,* 1986, **237**, 220-225.

92. H. Fujita, N. Ishida and R. Akagi, Delta-Aminolevulinate dehydratase deficiency, *Nihon Rinsho, Japanese Journal of Clinical Medicine*, 1995, **53**, 1408-17.

93. R. Ajioka, J. Phillips and J. Kusher, Biosynthesis of heme in mammals, *Biochimica et Biophysica Acta*, 2006, 723-736.

94. J. Chisom, D. Thomas and T. Hamill, Erythrocyte porphobilinogen synthase activity as an indicator of lead exposure in children, *Clinical Chemistry, 1985,* **31**, 601-605.

95. P. Erskine, N Senior, S.Awan, R. Lambert, G. Lewis, I. Tickle, M. Sarwar, P. Spencer, P. Thomas, M. Warren and Shoolingin, P. X-ray structure of 5-aminolaevulinate dehydratase, a hybrid aldolase, *Nature Structural Biology*, 1997, **4**, 1025-31.

96. A. Kappas S. Sassa, R. Galbraith and Y. Nordmann, *the Metabolic and Molecular Bases of Inherited Diseases*, 7th Ed. New York: McGraw-Hill, 1995, 2103–2160.

97. M. Warren, J. Cooper, S. Wood and P. Shoolingin, Lead poisoning, haem

synthesis and 5-aminolaevulinic acid dehydratase, *Trends in Biochemical Sciences,* 1998, **23**, 217–221.

98. M. Brennan and R. Cantrill, δ-Aminolaevulinic acid is a potent agonist for GABA autoreceptors, *Nature,* 1979, **280**, 514–515.

99. R. Labbe, H. Vreman and D. Stevenson, Zinc protoporphyrin: A metabolite with a mission, *Clinical Chemistry*, 1999, **45**, 2060-2072.

100. R. Labbe and L. Rettmer, Zinc Protoporphyrin: a product of iron-deficient erythropoiesis, *Seminars in Hematology*, 1989, **26**, 40–46.

101. M. Jones and O. Jones, The structural organization of haem synthesis in rat liver mitochondria, *Biochemistry Journal*, 1969,**113**, 507–514.

102. S. Hernberg, J. Nikkanen, G. Mellin and H. Lilius, Deltaamino-levulinic acid dehydratase as a measure of lead exposure, *Environmental. Health,* 1970, **21**, 140-145.

103. D. Mouw, K. Kalitis, M. Anver, J. Schwartz, A. Conston, R. Hartun, B. Cohen and D. Ringler, Lead: Possible toxicity in urban vs rural rats, *Environmental Health,* 1975. **30**, 276-280.

104. M. Dieter, and M.Finley, Aminolevulinic acid dehydratase enzyme activity in blood, brain, and liver of lead-dosed ducks, *Environmental Research*, 1979, **19**, 127-135.

105. H. Waldron, The anemia of lead poisoning: A review, *British Journal of Industrial Medicine,* 1966, 23, 83-100.

106. R. Cousins, Absorption, transport and hepatic metabolism of copper and zinc: special reference to metallothionein and ceruloplasmin, *Physiological Reviews*, 1985, **65**, 238-309.

107. D. Hamer, Metallothionein, *Annual Review of Biochemistry*, 1986, **55**, 913-51.

108. M. Piscator, On Cadmium in normal human kidneys with a report on the isolation of metallothioneine from cadmium-exposed rabbit livers, *Nordisk Hygienisk Tidskrift*, 1964, **45**, 76-82.

109. M. Richards and R. Cousins, Mammalian zinc homeostasis: Requirement for

RNA and metallothionein synthesis, *Biochemical and Biophysical Research Communications*, 1975, **64**, 1215-1223.

110. K. Squibb and R. Cousins, Control of cadmium binding protein synthesis in rat liver, *Environmental Physiology and Biochemistry*, 1974, **4**, 24-30.

111. M. Dunn, T. Blalock and R. Cousins, Metallothionein, *Experimental Biology and Medicine*, 1987, **185**, 107-119.

112. D. Thiele, Metal-regulated transcription in eukaryotes, *Nucleic Acids Research,* 1992, **20**, 1183-1191.

113. K. Gorski, M. Carneiro and U. Schibler, Tissue-specific in vitro transcription from the mouse albumin promoter, *Cell*, 1986, **47**,767-76.

114. Z. Ma, F. Jacobsen and D. Giedroc, Metal transporters and metal sensors: How coordination chemistry controls bacterial metal homeostasis, *Chemical Reviews,* 2009, **109**, 4644–4681.

115. M. Pennella, E. Shokes, N. Cosper, R. Scott and D. Giedroc, Structural Elements of Metal Selectivity in Metal Sensor Proteins, *Proceedings of the National Academy of Sciences U.S.A.,* 2003, **100**, 3713-3718.

116. J. Wu and B. Rosen, The ArsR protein is a trans-acting regulatory protein, *Molecular Microbiology*, 1991, **5**, 1331-1336.

117. J. Ye, A. Kandegedara, P. Martin and B. Rosen, Crystal structure of the Staphylococcus aureus pI258 CadC Cd(II)/Pb(II)/Zn(II)-responsive repressor, *Journal of Bacteriology,* 2005,**187**, 4214-4221.

118. L. Banci, I. Bertini, F. Cantini, S. Ciofi-Baffoni, J. Cavet, C. Dennison, A. Graham, D. Harvie and N. Robinson, NMR structural analysis of cadmium sensing by winged helix repressor CmtR, *Journal of Biological Chemistry,* 2007, **282**, 30181-30188.

119. J. Qin, F. Hl, J. Ye, K. Bencze, T. Stemmler, D. Rawlings and B. Rosen, Convergent evolution of a new arsenic binding site in the ArsR/SmtB family of metalloregulators, *Journal of Biological Chemistry,* 2007, **282**, 34346-34355.

120. T. Liu, X. Chen, Z. Ma, J. Shokes, L. Hemmingsen, R. Scott and D. Giedroc, A Cu (I)-sensing ArsR family metal sensor protein with a relaxed metal selectivity

profile, *Biochemistry*, 2008, **47**, 10564-10575.

121. C. Thelwell, N. Robinson and J. Turner-Cavet, an SmtB-like repressor from Synechocystis PCC 6803 regulates a zinc exporter, *Proceedings of the National Academy of Sciences U.S.A.*, 1998, **95**, 10728-10733.

122. L. Busenlehner, T. Weng, J. Penner-Hahn and D. Giedroc, Elucidation of Primary (alpha(3)N) and Vestigial (alpha(5)) Heavy Metal-binding Sites in Staphylococcus aureus pI258 CadC: Evolutionary Implications for Metal Ion Selectivity of ArsR/SmtB Metal Sensor Proteins, *Journal of Molecular Biology,* 2002, **319**, 685-701.

123. L. Busenlehner and D. Giedroc, Kinetics of metal binding by the toxic metal-sensing transcriptional repressor Staphylococcus aureus pI258 CadC, *Journal of Inorganic Biochemistry,* 2006, **100**, 1024-1034.

124. L. Busenlehner, S. Codreanu, P. Holm, P. Bhakat, H. Hebert, R. Morgenstern and R. Armstrong, Stress sensor triggers conformational response of the integral membrane protein microsomal glutathione transferase, *Biochemistry,* 2004, **43**, 11145-11152.

125. W. Shi, J. Dong, R. Scott, M. Ksenzenko and B. Rosen, The role of arsenic-thiol interactions in metalloregulation of the ars operon, *Journal of Biological Chemistry*, 1996, **271**, 9291-9297.

126. L. Busenlehner, J. Apuy and D. Giedroc, Characterization of a metallo-regulatory bismuth (III) site in *Staphylococcus aureus* pI258 CadC repressor, *Journal of Biological Inorganic Chemistry,* 2002, **7**, 551-559.

127. Y. Sun, M. Wong and B. Rosen, Role of cysteinyl residues in sensing Pb(II), Cd(II), and Zn(II) by the plasmid pI258 CadC repressor, *Journal of Biological Chemistry*, 2001, **276**, 14955-60.

128. D. Giedroc, A. Arunkumar and D. Trans, Metal sensor proteins: nature's metalloregulated allosteric switches, *Biochemistry and Biophysics*, 2007, **29**, 3107.

129. C. Eicken, M. Pennella, X. Chen, K. Koshlap, M. VanZile, J. Sacchettini and D. Giedroc, A Metal-Ligand-mediated Intersubunit Allosteric Switch in Related SmtB/ArsR Zinc Sensor Proteins, *Journal of Molecular Biology*, 2003, **333**, 683-695.

130. S. Patzer and K. Hantke, Dual repression by Fe^{2+}-Fur and Mn^{2+}-MntR of the mntH gene, encoding an NRAMP-like Mn^{2+} transporter in Escherichia coli, *Journal of Bacteriology*, 2001, **183**, 4806-4813.

131. S. Mills, and M. Marletta, Metal binding characteristics and role of iron oxidation in the ferric uptake regulator from Escherichia coli, *Biochemistry,* 2005, **44**, 13553-13559.

132. I. Loftin, S. Franke, N. J. Blackburn and M. McEvoy, Unusual Cu(I)/Ag(I) coordination of *Escherichia coli* CusF as revealed by atomic resolution crystallography and X-ray absorption spectroscopy, *Protein Science*, 2007, **16**, 2287-2293.

133. D. Traore, El A. Ghazouani, L. Jacquamet, F. Borel, L. Ferrer, D. Lascoux, L. Ravanat, M. Jaquinod, G. Blondin, C.Caux-Thang, V. Duarte and J. Latour, Structural and functional characterization of 2-oxo-histidine in oxidized PerR protein, *Nature Chemical Biology*, 2009, **5**, 53-59.

134. Y. An, B. Ahn, A. Han, H. Kim, K.Chung, J.Shin, Y. Cho, J. Roe and S. Cha, Structural basis for the specialization of Nur, a nickel-specific Fur homolog, in metal sensing and DNA recognition, *Nucleic Acids Research,* 2009, **37**, 3442-3451.

135. D. Lucarelli, S. Russo, E. Garman,. A. Milano, W. Meyer-Klaucke and E. Pohl, Crystal structure and function of the zinc uptake regulator FurB from Mycobacterium tuberculosis, *Journal of Biological Chemistry,* 2007, **282**, 9914-9922.

136. L. Jacquamet, D. Traore, J. Ferrer, O. Proux, D. Testemale, J. Hazemann, E. Nazarenko, A. El Ghazouani, C. Caux-Thang, V. Duarte and J. Latour, Structural characterization of the active form of PerR : insights into the metal-induced activation of PerR and Fur proteins for DNA binding, *Molecular Microbiology*, 2009, **73**, 20-31.

137. L. Pecqueur, B. D'Autréaux, J. Dupuy, Y. Nicolet, L. Jacquamet, B. Brutscher, I. Michaud-Soret and B. Bersch, Structural changes of *Escherichia coli* ferric uptake regulator during metal-dependent dimerization and activation explored by NMR and X-ray crystallography, *Journal of Biological Chemistry*, 2006, **281**, 21286-21295.

138. C. Outten, D. Tobin, J. Penner-Hahn and T. O'Halloran, Characterization of the metal receptor sites in Escherichia coli Zur, an ultrasensitive zinc (II)

metalloregulatory protein, *Biochemistry,* 2001, **40**, 10417-10423.

139. B. Carpenter, H. Gancz, R. Gonzalez-Nieves, A. West, J. Whitmire, S. Michel and D. Merrell, A single nucleotide change affects fur-dependent regulation of sod B in H. pylori, *PloS One,* 2009, **4**, 5369-5381.

140. L. Jin-Won and J. Helmann, The PerR transcription factor senses H_2O_2 by metal-catalysed histidine oxidation, *Nature*, 2006, **440**, 363–367.

141. P. Alphonse, I. Arunkumar and D. Giedroc, Individual Metal Ligands Play Distinct Functional Roles in the Zinc Sensor *Staphylococcus aureus* CzrA, *Journal of Molecular Biology*, 2006, **356**, 1124-1136.

142. D. Giedroc, Hydrogen peroxide sensing in Bacillus subtilis: it is all about the (metallo) regulator, *Molecular Microbiology*, 2009, **73**, 1-4.

143. M. Schmitt, Analysis of a DtxR-like metalloregulatory protein, MntR, from *Corynebacterium diphtheriae* that controls expression of an ABC metal transporter by an Mn^{2+}-dependent mechanism, *Journal of Bacteriology*, 2002, **184**, 6882-6892.

144. N. Schiering, X. Tao and H. Zeng, Structures of the apo-and the metal ion-activated forms of the diphtheria tox repressor from *Corynebacterium diphtheria*, *Proceedings of the National Academy of Sciences, U.S.A.*, 1995, **92**, 9843–9850

145. S. Andrews, A. Robinson and F. Rodriguez-Quinones, Bacterial iron homeostasis, *FEMS Microbiology Reviews*, 2003, **27**, 215-237.

146. M. Spiering, Bioinorganic Chemistry Special Feature: Metal stoichiometry and functional studies of the diphtheria toxin repressor, *Proceedings of the National Academy of Sciences U.S.A.*, 2003, **100**, 3808-3813.

147. A. White, X. Ding, J. VanderSpek, J. Murphy and D. Ringe, Structure of the metal-ion-activated diphtheria toxin repressor/tox operator complex, *Nature – Londo*n, 1998, **6692**, 502-506.

148. G. Wylie, V. Rangachari, E. Bienkiewicz, V. Marin, N. Bhattacharya, J. Love, J. Murphy and T. Logan, Prolylpeptide binding by the prokaryotic SH3-like domain of the diphtheria toxin repressor: a regulatory switch, *Biochemistry*, 2005, **44**, 40-51.

149. C. Liu, K. Mao, M. Zhang, Z. Sun, W. Hong, C. Li, B. Peng and Z. Chang, The SH3-like domain switches its interaction partners to modulate the repression activity

of mycobacterial iron-dependent transcription regulator in response to metal ion fluctuations, *Journal of Biological Chemistry*, 2008, **283**, 2439-2453.

150. J. DAquino, Mechanism of metal ion activation of the diphtheria toxin repressor DtxR, *Proceedings of the National Academy of Sciences U.S.A.,* 2005, **102**, 18408-18413.

151. J. DAquino, J. Lattimer, A. Denninger, K. DAquino and Ringe, D. Role of the N-terminal helix in the metal ion-induced activation of the diphtheria toxin repressor DtxR, *Biochemistry*, 2007, **46**, 11761-11770.

152. J. Kliegman, S. Griner, J. Helmann, R. Brennan and A. Glasfeld, Structural basis for the metal-selective activation of the manganese transport regulator of Bacillus subtilis, *Biochemistry*, 2006, **45**, 3493-3505.

153. G. Guedon, J. Helmann and R. Brennan, Structure of the manganese-bound manganese transport regulator of *Bacillus subtilis*, *Nature Structural Biology*, 2003, **10**, 652-657.

154. K. Sen, A. Sienkiewicz, J. Love, J. VanderSpek, P. Fajer and T. Logan, Mn(II) binding by the anthracis repressor from Bacillus anthracis, *Biochemistry*, 2006, **45**, 4295-303.

155. M. Dewitt, J. Kliegman, J. Helmann, R. Brennan, D. Farrens and A. Glasfeld, The Conformations of the manganese transport regulator of *bacillus subtilis* in its metal-free state, *Journal of Molecular Biology,* 2007, **365**, 1257-1265.

156. M. Golynskiy, S, Li, V. Woods and S. Cohen, Conformational studies of the manganese transport regulator (MntR) from Bacillus subtilis using deuterium exchange mass spectrometry, *Journal of Biological Inorganic Chemistry*, 2007, **12**, 699–709.

157. K. Sen, T. Logan and P. Fajer, Protein dynamics and monomer-monomer interactions in AntR activation by electron paramagnetic resonance and double electron-electron resonance, *Biochemistry*, 2007, **46**, 11639-11649.

158. K. De Pina, V. Desjardin, M. Mandrand-Berthelot, G. Giordano and L. Wu, Isolation and characterization of the nikR gene encoding a nickel-responsive regulator in *Escherichia coli*, *Journal of Bacteriology*, 1999, **181**, 670-674.

159. S. Mulrooney and R. Hausinger, Nickel uptake and utilization by micro-organisms, *FEMS Microbiology Review*, 2003, **27**, 239–261.

160. A. Sigel and H. Sigel, *Metal ions in Biological Systems: in Nickel and its Role in Biology*, New York: M. Dekker, 1988.

161. S. Wang, A. Dias, S. Bloom and D. Zamble, Selectivity of metal binding and metal-induced stability of Escherichia coli NikR, *Biochemistry*, 2004, **43**, 10018–10028.

162. C. Bahlawane, C. Dian, C. Muller, A. Round, C. Fauquant, K. Schauer, de H. Reuse, L. Terradot and I. Michaud-Soret, Structural and mechanistic insights into Helicobacter pylori NikR activation, *Nucleic Acids Research,* 2010, **38**, 3106-3118.

163. S. Bloom and D. Zamble, Metal-selective DNA-binding response of *Escherichia coli* NikR, *Biochemistry*, 2004, **43**, 10029-10038.

164. R. Schreiter, M. Sintchak, Y. Guo, P. Chivers, R. Sauer and C. Drennan, Crystal structure of the nickel-responsive transcription factor NikR, *Nature Structural Biology*, 2003, **10**, 794-799.

165. M. Bradley, P. Chivers and N. Baker, Molecular dynamics simulation of the Escherichia coli NikR protein: equilibrium conformational fluctuations reveal interdomain allosteric communication pathways, *Journal of Molecular Biology*, 2008, **378**, 1155-1173.

166. L. Busenlehner and D. Giedroc, The SmtB/ArsR family of metalloregulatory transcriptional repressors: Structural insights into prokaryotic metal resistance, *FEMS Microbiology Reviews*, 2003, **27**, 2-3.

167. P. Chivers, and R. Sauer, NikR repressor: high-affinity nickel binding to the C-terminal domain regulates binding to operator DNA, *Chemistry and Biology*, 2002, **9**, 1141-8.

168. E. Schreiter, S. Wang, D. Zamble and C. Drennan, NikR-operator complex structure and the mechanism of repressor activation by metal ions, *Proceedings of the National Academy of Sciences U.S.A.*, 2006, **103**, 13676-13681

169. T. Chivers and T. Tahirov, Structure of *Pyrococcus horikoshii* NikR: Nickel Sensing and Implications for the Regulation of DNA Recognition, *Journal of*

Molecular Biology, 2005, **348**, 597-607.

170. S. Leitch, M. Bradley, J. Rowe, P. Chivers, and M. Maroney, Nickel-specific response in the transcriptional regulator, Escherichia coli NikR, *Journal of the American Chemical Society*, 2007, **129**, 5085-5095.

171. N. Brown, J. Stoyanov, S. Kidd, and J. Hobman, The MerR family of transcriptional regulators, *FEMS Microbiology Reviews,* 2003, **27**,145-163.

172. T. O'Halloran and C. Walsh, Metalloregulatory DNA-binding protein encoded by the MerR gene: Isolation and characterization, *Science*, 1987, **235**, 211–214.

173. D. Nies and S. Silver, Molecular microbiology of heavy metals, *Microbiology Monographs*, 2007, **6**, 357-370.

174. B. Tamar, S. Miller and A. Summers, Bacterial mercury resistance from atoms to ecosystems, *FEMS Microbiology Reviews*, 2003, **27**, 355–384.

175. A. Summers, Untwist and shout: A heavy metal-responsive transcriptional regulator, *Journal of Bacteriology*, 1992, **174**, 3097–3101.

176. M. Mergeay, S. Monnchy, T. Vallaeys, V. Auquier, A. Benotmane, P. Bertin, S. Taghavi, J. Dunn, D. Lelie, R. Wattiez, *Ralstonia metallidurans*, a bacterium specifically adapted to toxic metals: Towards a catalogue of metal-responsive genes, *FEMS Microbiology Reviews,* 2003, **27**, 385–410.

177. T. O'Halloran, Transition metals in control of gene expression, *Science*, 1993, **261**, 715–725.

178. E. Hidalgo, H. Ding and B. Demple, Redox signals transduction via iron-sulfur clusters in the SoxR transcription activator, *Trends Biochemical Sciences*, 1997, **22**, 207–210.

179. J. Stoyanov, J. Hobman and N. Brown, CueR (YbbI) of Escherichia coli is a MerR family regulator controlling expression of the copper exporter CopA, *Molecular Microbiology*, 2001, **39**, 502–511.

180. E. Heldwein and R. Brennan, Crystal structure of the transcription activator BmrR bound to DNA and a drug, *Nature*, 2001, **409**, 378–382.

181. L. Shewchuk, J. Helmann, W. Ross, S. Park, A. Summers and C. Walsh,

Transcriptional switching by the MerR protein: Activation and repression mutants implicate distinct DNA and mercury (II) binding domains, *Biochemistry*, 1989, **28**, 2340–2344.

182. M. Ahmed, C. Borsch, S. Taylor, N. Vázquez-Laslop and A. Neyfakh, A protein that activates expression of a multidrug efflux transporter upon binding the transporter substrates, *Journal of Biological Chemistry*, 1994, **269**, 28506–28513.

183. Q. Zeng, C. Stalhandske, M. Anderson, R. Scott and A. Summers, The core metal-recognition domain of MerR, *Biochemistry*, 1998, **37**, 15885–15895.

184. A. Changela, Molecular basis of metal-ion selectivity and zeptomolar sensitivity by CueR, *Science*, 2003, **301**, 1383–1387.

185. M. Godsey, N. Baranova, A. Neyfakh and R. Brennan, Crystal structure of MtaN, a global multidrug transporter gene activator, *Journal of Biological Chemistry*, 2001, **276**, 47178–47184.

186. S. Kidd, A. Potter, M. Apicella, M. Jennings and A. McEwan, NmlR of Neisseria gonorrhoea: A novel redox responsive transcription factor from the MerR family, *Molecular Microbiology*, 2005, **57**, 1676–1689.

187. B. Borremans, J. Hobman, A. Provoost, N. Brown and D. van Der Lelie, Cloning and functional analysis of the Pbr lead resistance determinant of *Ralstonia metallidurans* CH34, *Journal of Bacteriology*, 2001, **183**, 5651–5658.

188. P. Chen, B. Greenberg, S. Taghavi, C. Roman, D. Lelie and C. He, An exceptionally selective Pb(II)-regulatory protein from *Ralstonia metallidurans*: Development of a fluorescent Pb(II) probe, *Angewantde Chemie* 2005, **44**, 2715–2719.

189. J. Uddin, *Macro to nano spectroscopy*, Rijeka: InTech. http://www. Intech-open.com/ books/macro-to-nano-spectroscopy, 2012.

190. S. Deo and H. Godwin, Selective, ratiometric fluorescent sensor for Pb^{2+}, *Journal of the American Chemistry Soc*iety, 2000, **122**, 174–175.

191. A. Czarnik, Desperately seeking sensors, *Chemistry and Biology*, 1995, **2**, 423–428.

192. D. Becker, High resolution NMR theory and chemical applications, (3rd Ed),

New York, 2000.

193. N. Stone, The value of electronic hyperfine anomalies to nuclear and solid state physics, *Journal De Physique Colloques*, 1973, **34**, 69-75.

194. K. Neupane and V. Pecoraro, ^{207}Pb- NMR spectroscopy reveals that Pb (II) coordinates with glutathione (GSH) and tris cysteine zinc finger proteins in a PbS3 coordination environment, *Journal of Inorganic Biochemistry*, 2011, **105**, 1030-1034.

195. V. Mah and F. Jalilehvand, Lead (II) complex formation with glutathione, *Inorganic Chemistry*, 2012, **51**, 6285–6298.

196. J. Aramini, T. Hiraoki, M. Yazawa, T. Yuan, M. Zhang and J. Vogel, Lead-207 NMR: a novel probe for the study of calcium-binding proteins, *Journal of Biological Inorganic Chemistry*, 1996, **1**, 39–48.

197. B. Wrackmeyer and K. Horchler, Pb-NMR parameters, *Annual Reports on NMR Spectroscopy*, New York: *Academic Press*, 1990, **22,** 249–306.

198. S. Rupprecht, S. Franklin and K. Raymond, Synthesis of monothiohydroxamic ligands and their lead complexes, Structures of N-methyl-3-pyridothiohydroxamic acid, *bis* (N-methyl-3-pyridothiohydroxamato) lead(II) and *bis*(N-cyclohexyl-phenylacetothiohydroxamato) lead (II), *Inorganica Chimica Acta*, 1995, **235**, 185–194.

199. S. Rupprecht, K. Langemann, T. Lugger, J. McCormick and K. Raymond, Coordination chemistry of bis-thiohydroxamic acids: Synthesis and characterization of their lead(II) complexes and stability constant determination, *Inorganica Chimica Acta*, 1996, **243**, 79–90.

200. R. Pedrido, M. Bermejo, M. Romero, M. Vazquez, A.Gonzalez-Noya, M. Maneiro, M. Rodríguez and M. Fernandez, Syntheses and X-ray characterization of metal complexes with the pentadentate thiosemicarbazone ligand bis(4-N-methylthiosemicarbazone)-2,6-diacetylpyridine. The first pentacoordinate lead (II) complex with a pentagonal geometry, *Dalton Trans*, 2005, **7**, 572–579.

201. D. Reger, Y. Ding, A. Rheingold and R. Ostrander, Reactions of $\{(C_5H_5)$ Co [P (O) $(OC_2H_5)_2]^3\}$ with MCl_3 and M (CH_3) Cl_2 (M= Ga, In). Crystal and molecular

structure of [{(C_5H_5) Co [P (O) (OC_2H_5)$_2$]^3} 2Ga] Ga (CH_3) Cl_3, *Polyhedron*, 1994, **13**, 3053–3058.

202. K. Neupane and V. Pecoraro, Probing a Homoleptic PbS_3 Coordination Environment in a Designed Peptide Using Pb NMR Spectroscopy: Implications for Understanding the Molecular Basis of Lead Toxicity, *Angewandte Chemie*, 2010, **49**, 8177–8180.

203. I. Arsenault and A. Dean, Preparative and multinuclear nuclear magnetic resonance spectroscopic study of $As(SPh)_x(SePh)_{3-x}$(x=0–3), $Sb(SPh)_x(SePh)_{3-x}$ (x=0–3), $Bi(SPh)_3$,$Bi(SePh)_3$,$[Sn(SPh)_x(SePh)_{3-x}]^-$(x=0–3), $Pb(SePh)^{3-}$, and $Pb(SPh)^{3-}$ and some related thiolatoplumbates(II), *Canadian Journal of Chemistry*, 1983, **61**, 1516–1523.

204. P. Dean, J. Vittal and N. Payne, Discrete trigonal-pyramidal lead (II) complexes: syntheses and x-ray structure analyses of $[(C_6H_5)_4As][Pb(EC_6H_5)_3]$ (E = S, Se), *Inorganic Chemistry,*1984, **23**, 4232–4236.

205. G. Christou, K. Folting and J. Huffman, Mononuclear three-coordinate metal thiolates: Preparation and crystal structures of $[NBu^n_4][Hg(SPh)_3]$ and $[NPr^n_4]$ $[Pb(SPh)_3]$, *Polyhedron*, 1984, **3**, 1247–1253.

206. R. Andersen, R. diTargiani, R. Hancock, C. Stern, D. Goldberg and H. Godwin, Characterization of the first N_2S(alkylthiolate)Pb compound: a model for three-coordinate Pb in biological systems, *Inorganic Chemistry*, 2006, **45**, 6574–6576.

CHAPTER 2: MATERIALS AND METHODS

2.1. Introduction to gene cloning

In order to study proteins from organisms, the proteins must first be produced in sufficient quantity. The proteins and the genes that encode them must be identified. After identification and isolation from the genomic DNA, a given gene can be inserted into a self-replicating genetic element; this is termed the *vector producing recombinant DNA*. The vector can be derived from a bacterial plasmid or virus. This process lies at the heart of recombinant DNA technology (or gene cloning). The recombinant DNA is then transplanted into a suitable host cell, and the protein can be expressed thereafter.[1]

The remainder of this second chapter provides a description and explanation of gene cloning theory and an account of the system that was used for the production of recombinant PbrR from *Pseudomonas putida* KT2440.

2.1.1. DNA structure

DNA is a polymer comprising two polynucleotide chains that are held together by weak thermodynamic forces in a double helix structure bonded with hydrogen bonds (H-bonds). The DNA molecule consists of four different nucleotides containing one of four nitrogenous bases: thymine (T), adenine (A), guanine (G) or cytosine (C), along with a molecule of a pentose sugar, 2-deoxyribose, and a phosphate group. Within the double helix, A always links with T and C always links with G; these are known as complementary base pairs (A-T and C-G) (in RNA, it is C-U). The A-T base pairs are held together by two H-bonds, while the C-G base pairs are bound by three H-bonds. Between the 3-hydroxyl group of one pentose and the 5-hydroxyl of the next are phosphodiester bonds that linking the two, as shown in Figure 2.1.[2]

Unlike many polymers, DNA chains are paired in an anti-parallel arrangement in which the 5-end of one strand is paired with the 3-end of the other. The hydrophilic deoxyribophosphate of each thread is positioned on the outside of the molecule, while the hydrophobic nitrogen bases are found inside, perpendicular to the axis of the helix. This double helical formation avoids the relaxation of torsional stress by

rotation across a single covalent bond. The formation of the of the bases, one on top of another, gives the DNA significant flexional rigidity (Figures2.1 and 2.2).[3,4]

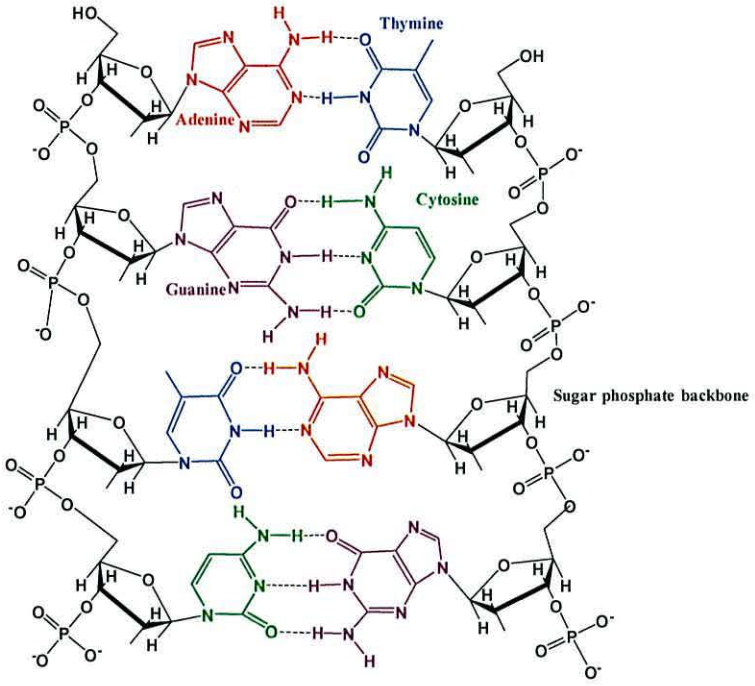


Figure: 2. 1. The chemical structure of DNA

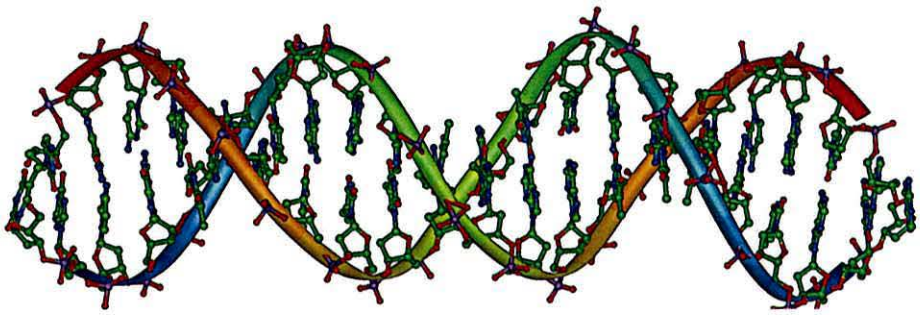


Figure: 2.2. DNA double helix.

[oxygen atom (red), Nitrogen atom (blue), carbon atom (green), phosphate group (purple)].

Source: pdb1D65

2.1.2. The gene

In a prokaryote, the entire sequence of DNA consists of a circular chromosome, containing all the genetic information needed by that organism; this is the prokaryote's *genome*. Genomic DNA varies considerably in size, from 5224 base pairs in the Simian virus (SV40)[5] to approximately 2.91 billion base pairs in

humans.[6] The locations within the genomic DNA that are responsible for encoding the production of a particular protein or are accountable for a particular organism's characteristics are termed *genes*.[7] Protein synthesis is encoded by genes in two main ways, namely:

Transcription of the gene into complementary RNA molecules (mRNA, tRNA, rRNA), each of which has an explicit purpose.
Translation. This starts when the mRNA is read from the 5-end to the 3-end in blocks of three nucleotide bases (called *codons*).The tRNA decodes the codons, which then interact with mRNA, and are translated to protein in ribosomes.
Each amino acid is specified by one or more specific codons. Because there are four different nucleotides, with each codon being made up of three nucleotides, sixty four different codons are possible. Each of these also has an anticodon to recognise the codons in the mRNA. This means there are multiple codons available to specify most of the twenty different amino acid monomers from which proteins are made.[8]

2.1.3. Cloning vectors

A portion of DNA that has a foreign piece of DNA inserted and yet remains viable within the cell is termed a *cloning vector*. These vectors can be plasmids (as in this work), bacteriophages, viruses, or even small artificial chromosomes.[9]All these vectors have a position within their sequence known as the origin of replication, the ability to amplify replication according to an antibiotic resistance gene, and at least two areas (restriction sites) that allow fragments of exogenous DNA to be inserted after being cut by restriction endonucleases.[9,10] The vector and the exogenous DNA have to be cut by identical restriction enzymes. This permits the insertion of the exogenous DNA fragment into the appropriate restriction site. If not enough exogenous genomic DNA fragments are available for cloning into a vector, the quantity of gene specific copies can be increased using the polymerase chain reaction technique.[10]

2.1.4. Polymerase chain reaction

The polymerase chain reaction (PCR) was introduced *in vitro* (1985).[11]This technique is powerful; however, the method is simple. The PCR automation was

developed in 1988; in addition, the use of a thermo–stable Taq polymerase allowed the procedure to be simplified.[12] PCR is conducted from the known sequence of the target DNA molecule to be copied. PCR amplifies small DNA targets 100-1000 base pairs (bp) long; targets >5000 bp are difficult to amplify. A pair of single stranded oligonucleotide primers must be synthesised that possess DNA sequences in the flanking areas of the targeted sequences. These primers are complementary to every end of the targeted arrangement but must persist in opposite strands. Primers carry 20-30 nucleotides and attach to the complementary flanking area at the 3' end. Primers are designed to flank the sequence of interest.[1]

2.1.5 Electrophoresis methods

2.1.5.1 Agarose gel electrophoresis for characterestion of nucleic acids

Seaweed is a source of agarose, a natural linear polymer that is capable of producing a gel matrix through hydrogen-bonding when treated with buffer and left to cool. Agarose gel preparation is an easy and quick process.[13] Many applications require only a single-component agarose, without any need for any polymerisation catalysts. Agarose gels are commonly used to separate moderate and large-sized nucleic acids and have good range of separation but a relatively low resolving power.[14]The concentration of agarose gel can be expressed as a percentage of agarose to volume of buffer (w/v); the concentration range of agarose gels is generally from 0.2% to 3%.[15]However, the choice of percentage of agarose gel utilised is entirely based on the fragments that need to be resolved. DNA fragments migrate at a faster rate in low concentration agarose gels, so separation of small fragments of DNA requires a high concentration of agarose, whereas larger fragments require a low concentration of agarose for separation. The DNA in the agarose gel can be localised via direct staining with a low concentration of intercalating fluorescent ethidium bromide dye and viewing under ultraviolet light. The dye is usually included in the gel and running buffer tank, or the gel can be stained once the DNA is separated.[16]

2.1.5.2 SDS-PAGE for characterestion of protein

The choice of acrylamide concentration is critical for effective separation of molecules.[17] controlling the pore size of an acrylamide gel requires appropriate

selection of concentrations of acrylamide and the cross-linking agent methylene bisacrylamide. A linear decrease occurs in the pores size with increasing concentration (T) of the monomer (acrylamide plus cross-linker) in the gel. The linkage between of the monomer and the cross-linker (C) is more complex. Smaller molecules are separated in higher percentage gels (higher T) with smaller pores. Polyacrylamide gel electrophoresis aids in effective separation of proteins when the pH within the matrix is maintained. The electrophoresis buffers are integral for determining protein molecular weight. The electrophoretic mobility of protein is altered due to composition and ionic strength (salt content) of the electrophoresis buffer.[18] In the absence of salt, the electrical conductance is little and the protein exhibits poor migration within the gel. When the ionic strength of the buffer is increased, then electrical conductance is efficient and an optimum quantity of heat is released.[16]

2.2. Molecular biology methods.

Table: 2.1. Equipment used in this chapter.

Equipment	Manufacturer/model
Centrifuge (expression protein)	Hettich Rotanta 460R
UV Visible spectroscopy	Perkin Elmer Lambda 35
Shaker	1-Bamstead/lab-line E class 2-Gallenkamp Shaker incubator (orbital incubator shaker 2010 England
Atomic absorption spectroscopy	Atomic absorption spectrometer – Varian-220FS
Centrifuge (purification plasmid)	Eppendorf centrifuge 5415D

2.2.1. The plasmid vector pET28a (+)

The *P. putida* KT2440 PbrR gene was cloned using pET28a(+) (Novagen) as a plasmid vector (Figure 2.3).

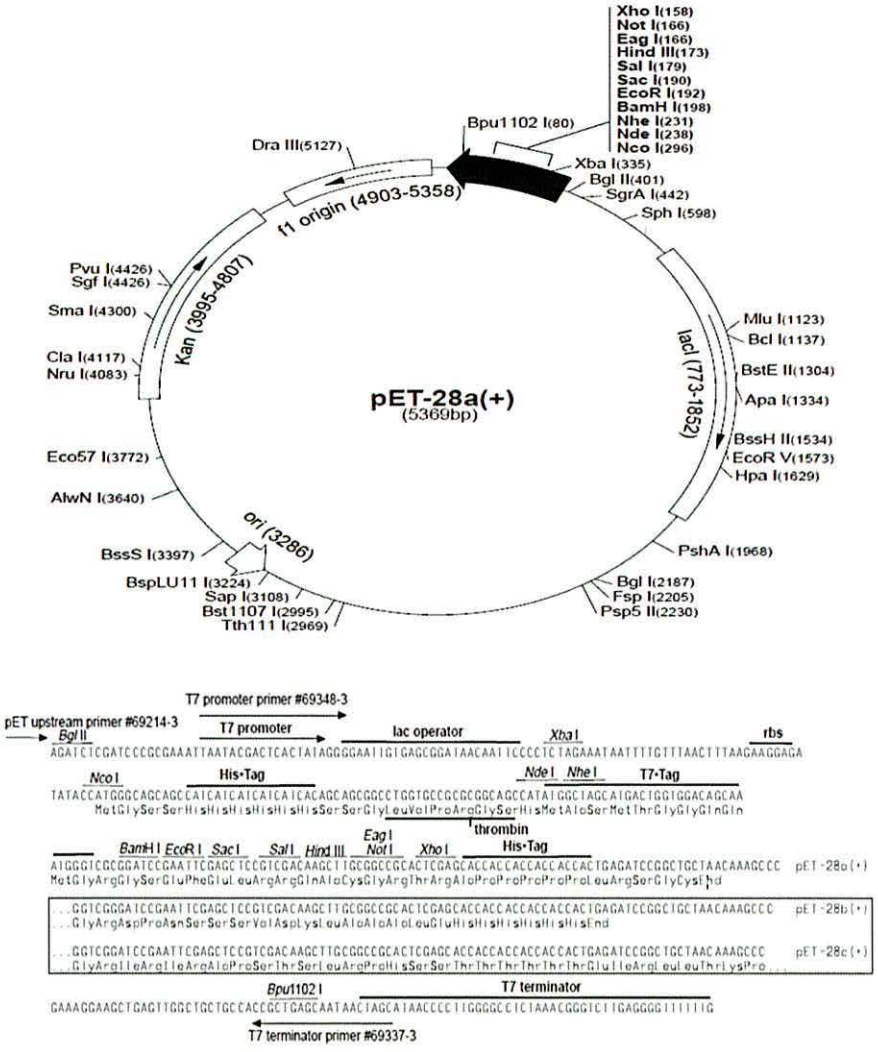


Figure: 2.3. The pET28a (+) expression vector (Novagen).

Code: (a) the vector map; (b) cloning and expression regions. [19]

2.2.2. Isolation of genomic DNA from *Pseudomonas putida* KT2440

Sterile/autoclaved tools and materials were used to inoculate 5 mL nutrient broth E medium with 2.5 µL 50 µM kanamycin and a single colony of *P. putida* KT2440 taken from a nutrient agar streak-plate culture strain. The inoculated broth E was placed overnight in a 37 °C incubator shaker. The genomic DNA was isolated by

placing the overnight solution in tubes; each tube was then centrifuged at 13,000 x *g* and 4 °C for 2 minutes in an Eppendorf centrifuge 5415 D. The supernatant was removed and the cell pellet was resuspended and placed in 600 µL of Nuclei Lysis solution (75 mM NaCl, 24 mM EDTA; pH 8.0).

This mixture was then incubated for 5 minutes at 80°C using a block heater and then allowed to cool to room temperature. After addition of 3 µL of 10 mg/ml RNase solution, the mixture incubated at 37 °C for 60 minutes. The solution was then allowed to cool to room temperature and 200 µL of a protein precipitation solution was added to the samples and vortexed; the solution was then incubated on ice for 5 minutes. It was then centrifuged at 13,000 x *g* for 10 minutes at 4° C. This resulted in a lysed cell solution. This cell lysate was transferred to a 1.5 mL capacity Eppendorf tube and centrifuged in a Sanyo Hawk 15/05 refrigerated bench-top microfuge at 12,000 x *g* and 4 °C for 2 minutes. The resulting top aqueous layer containing the genomic DNA was transferred to a clean 1.5 ml Eppendorf tube.

2.2.3. Construction of the expression plasmid

The Polymerase Chain Reaction (PCR) was used to amplify the PbrR from *P. putida* KT2440. The pET28a (+) vector (Novagen) was used for the PCR polymerisation. The cloning strategy comprised cloning PbrR into the plasmid vector pET28a (+), using the forward and reverse primers shown in Table 2.2.

Table: 2.2. Oligonucleotide primers used for the amplification and cloning of the PbrR gene into the expression vector.

Primer	Nucleotide sequence (5`→ 3`)	Restriction Enzyme
F-PbrR (forward)	CTAGTGGAATTCATGAAGATCGGAGAACTGGC	*EcoRI*
R-PbrR (reverse)	CATCACAAGCTTGATGCCCGTGACTCCGCCCCAC	*HindIII*

The stop codon TAA was removed from the PET28a (+); this was in order to include the His-tag in the protein. The restriction sites had been engineered as appropriate. The amplified fragments were cloned into the pET28a (+) using the *EcoRI* and *HindIII* restriction enzymes.

2.2.4. Amplification of the PbrR gene using PCR.

DNA polymerase enzyme (Novagen) was chosen for PCR amplification of the PbrR gene from *P. putida* KT2440 genomic DNA. The PCR reaction mixture contained 5 μL of a 10 mM mixture of each of dATP, dCTP, dGTP, and dTTP (dNTPs); 2.5 μL 10 μM forward and the reverse primers, 10 ng (2 μL) DNA template, 10 μL 5 × Phusion GC buffer (1 X buffer provides 1.5 mM $MgCl_2$), 27 μL sterile water and 1 μL Phusion polymerase buffer.

The PCR machine (Techgene) was programmed to treat the reaction mixture as follows:

Initialisation step: 95 °C for 2 minutes; this was to activate the polymerase enzyme.
Denaturation step: 95 °C for 20 seconds; this was to separate the DNA strands, principally through the heat-disruption of H-bonds between complementary bases.
Annealing step: Lowest primer T_m °C for 10 seconds; this was to alter the primers to single stranded DNA.
Extension step: 70 °C for 90 seconds; this was to add the complimentary dNTPs to the template; 70 °C is the optimum temperature of the polymerase enzyme when treated for 90 seconds; at this point, new DNA strands complement the DNA template strand, which is synthesised by adding complementary dNTPs to the template in 5' to 3' directions.[1]
Elongation step: 70 °C for 10 minutes; this was to enable any uncompleted double strand to be completed; this ensures that any remaining single stranded DNA is fully extended.
The thermal cycler machine was then programmed to repeat the (Denaturation, Annealing, and Extension) steps 25 times.

2.2.5. Agarose gel electrophoresis

The PCR product was assessed by agarose gel electrophoresis. A 1% (w/v) agarose solution was prepared by melting 1 g of agarose in 100 mL of 1 × Tris-borate ethylenediamine tetra-acetic acid (TBE) buffer in a microwave oven for 3–5 minutes. The melted agarose was then stirred and 3 μL of ethidium bromide was added to the gel (as part of the staining procedure).

The solution mixture was then placed into a gel rack; a comb was inserted at one side of the gel, and the gel was cooled to room temperature. The gel and the rack were then placed in a tank and sufficient 1 × TBE buffer was added to cover the gel. A 3 μL volume of loading buffer (0.25 % w/v bromophenol blue) was then added to 10 μL of the PCR product, and the mixture loaded into the well. Bromophenol blue allows monitoring of progress of the electrophoresis; the high density of sucrose inhibits loss of the sample from the sample well.

A 200 bp DNA ladder (Promega) was routinely run in another lane in the gel to calibrate the size of the base pairs in the PCR products (Figure 2.4). A voltage of 120 V was applied for a period of 45 minutes. Agarose gels were visualised under UV light in a dark room. Once band(s) of the correct molecular weight were detected on the gel, the corresponding solutions from the PCR were purified as shown in chapter 4.

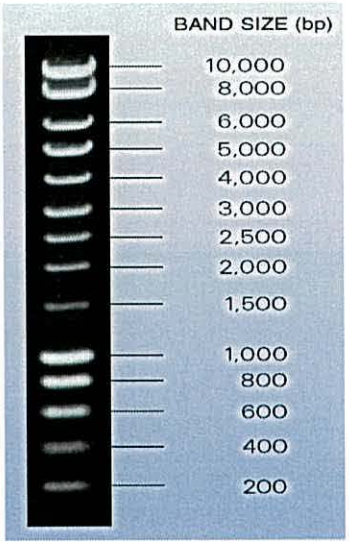


Figure: 2.4. The DNA ladder (Promega).

Fragment sizes ranged from 200 to 10000 base pairs; these indicated the corresponding sizes of the DNA fragments from the PCR product

2.2.6. PCR product purification

The final PCR product was gel purified to remove any contaminating PCR buffers, primers, and smaller fragment by-products. The PCR product was applied to a 1 × TBE buffer agarose gel for this purpose.

The resolved PCR DNA product bands on the gel were checked using UV trans-illumination with the Bio-Rad Gel Doc 2000 Imager System. The appropriate bands were removed from the gel with a clean scalpel blade; each gel slice was then moved to an empty 1.5 mL plastic Eppendorf microfuge. The QIA-quick Gel Extraction Kit Protocol[20] was then applied for purification purposes. The procedure was as follows:

Three volumes (3 ×) of the gel (agarose) solubilisation QG solution (1.0 M NaCl, 50 mM MOPS (3-morpholinopropane-1-sulfonic acid), pH 7.0, 15 % isopropanol, v/v) were added to an Eppendorf tube (100 mg: 300 µl).

The tube was placed in a hot-block set at 50 °C for 10 minutes incubation; during this period the tube was removed periodically (approx. every 2 minutes) and briefly vortexed.

While the agarose was being solubilised, GenElute binding column G was placed into 2 mL collection tubes (the number of samples and therefore the number of tubes varied); 800 µl of the column preparation solution was then added to each binding column and centrifuged for 1 minute at 10,000 rpm. The flow-through liquid was then discarded; 500 µl of the gel solubilisation QG solution was then added to each binding column, and the mixture centrifuged for 1 minute.

The solubilised gel solution mixture was then loaded into the binding column and centrifuged for 1 minute; the flow-through liquid was then discarded.

Next, 750 µL of wash solution was added to the binding column and centrifuged for 1 minute; the flow-through liquid was then discarded. This was the wash step procedure.

The wash step procedure was repeated and the empty column assembly was re-spun at 10,000 rpm for 1 minute; this removed any residual traces of wash solution from the column.

Finally, the column was placed in a fresh/clean 1.5 mL capacity plastic Eppendorf microfuge tube; 50 µL of kit elution buffer (10 mM Tris-Cl, pH 8.5) was then added to the centre column; the tube was centrifuged at 10,000 rpm for 1 minute and the flow-through liquid (i.e., the PCR DNA product) was saved.

2.2.7. Gel extraction and purification of DNA fragments

A QIA quick gel extraction kit (Qiagen, UK) [20] was used to extract the desired gene from the gel. After initial identification of the fragment by viewing the agarose gel under UV light the corresponding band was isolated. The DNA was extracted from the gel. The DNA was stained using purified salts; this procedure followed the manufacturer's instructions.[21]

2.2.8. Nucleic acid digestion

Complementary ends were generated by doubly digesting the purified PCR product and the plasmid vector separately with restriction enzymes. *EcoRI* and *HindIII* restriction enzymes were used to generate pET28a (+) PbrR. A recombinant DNA construction reaction was performed in the final volume (10 µL) of a nucleic acid mixture. In order to double digest the plasmid, the mixture was prepared in two tubes. The first contained 2 µL of pET28a (+), 1 µL of each restriction enzymes (*EcoRI* and *HindIII*), 1 µL of appropriate supplied buffer E (6 mM Tris-HCl at pH 7.5, 6 mM $MgCl_2$, 100 mM KCl, 1 mM DTT) and 5 µL nuclease-free water. The second tube contained 7 µL of purified PCR product, 1 µL of each restriction enzyme (*EcoRI* and *HindIII*), and 1 µL of appropriate supplied buffer (buffer E). Each tube was mixed and incubated at 37 °C for 1 hour; 50 µL binding buffer (PB) [a high concentration of guanidine hydrochloride and isopropanol (5 M Gu-HCl, 30 % isopropanol)] was then added to each tube. The contents of each first and second tube were then combined (using a pipette) and the resultant mixture placed in spin column (120 µL) and spun for 1 minute; 750 µL washing buffer (PE) (10 mM Tris-HCl at pH 7.5, 80 % ethanol) was added and spun for 1 minute; 30 µL elution buffer (EB) (10 mM Tris-Cl, pH 8.5) buffer was then added and the mixture spun for a further 1 minute.

2.2.9. DNA ligation reaction

The ligation of the gene ends to the plasmid was required to produce the recombinant DNA. A purified product, as described in the previous sub-section, was frozen in liquid nitrogen and vacuum dried for 45 minutes; the product was then re-suspended in 8 µL of nuclease-free water; 1 µL T4 ligase enzyme (Promega) and 1 µL T4 ligase

buffer (Promega) were added to the resuspended DNA and incubated overnight at 16 °C. The following day, 1 µL sacI restriction enzyme and 1 µL J buffer (10 mM Tris-HCl at pH 7.5, 7 mM $MgCl_2$, 50 mM KCl, 1 mM DTT) were added to the re-suspended DNA and incubated at 37 °C for 1 hour. [22]

2.2.10. Checking the cloned gene size and sequencing

In order to check whether the PbrR gene was inserted into the pET28a (+) plasmid and the recombinant plasmid DNA has been successfully constructed, 7 µL of the sample from the above sub-section (2.2.9) was double digested with 1 µL of each appropriate restriction enzyme and 1 µL of suitable supplied buffer (buffer E) at 37 °C for 1 hour. Then, 2 µL of loading buffer were added to the digested mixture and run on 1 % (w/v) agarose gel electrophoresis alongside the DNA molecular weight marker. This allowed determination of the size (length) of the cloned gene and the original plasmid. Once the correct length was observed on the gel, 15 µL of the same sample and two tubes containing 15 µL from 2 ng/ml forward and reverse primers were sent to Eurofins MWG Operon, UK for sequencing.

2.2.11. Site-directed mutagenesis

Site-directed mutagenesis is a powerful tool for the examination of structure-function relationships, as a mutation can be created at a defined site in the plasmid DNA. Site-directed mutagenesis is performed by first designing two synthetic oligonucleotide primers (see Table 2.3); these effect a change in at least one codon (from that observed in the native DNA). The DNA polymerase extends the primers on the template plasmid during temperature cycling. This results in the production of a mutated plasmid DNA.

Table: 2.3. Oligonucleotide primers for site-directed mutagenesis.

Mutated nucleotides shown in green, primers were designed based on the wild-type DNA plasmid.

Primer	Sequence 5'-3'
Cys 77 A	For:CCCGACGACGCG**GCC**GGCAGCGTCAATGCG Rev:CGCATTGACGCTGCC**GGC**CGCGTCGTCGGG
Cys 112 A	For:CTGCGGCGGCGC**GCC**AACGCG CAG GGGAGT Rev:ACTCCCCTGCGCGTT**GGC**GCGCCGCCGCAG
Cys112A/Cys119A	For:CTGCGGCGGCGC**GCC**AACGCGCAGGGGAGT GAA**GCC**GCGATCTTGCAGCAA Rev:TTGCTGCAAGATCGC**GGC**TTCACTCCCCTGC GC**GTT**GGCGCGCCGCCG CAG
Cys 77D	For:CCC GACGACGCG**GAC**GGCAGCGTCAATGCG Rev:CGCATTGACGCTGCC**GTC**CGCGTCGTCGGG
Cys 112D	For:CTGCGGCGGCGC**GAC**AACGCGCAGGGGAGT Rev:ACTCCCCTGCGCGTT**GTC**GCGCCGCCGCAG
Cys112D/Cys119D	For:CTGCGGCGGCGGCGC**GAC**AACGCGCAGGGGA GTGAA**GAC**GCGATCTTGCAG Rev:CTGCAAGATCGC**GTC**TTCACTCCCCTGCGCGTT **GTC**GCGCCGCCGCCGCAG

Three steps were used for the mutagenesis. The first involved construction of the initial fragment PCR1. This was made from F1, F2 (F1 = 20 μL 5 × PG buffer, 10 μL 2 mM dNTPs, 2 μL DNA template 1 μLphusion polymerase, 5 μL PbrR forward primer, 5 μL mutant reverse primers, 57 μL deionised water (dH_2O) (separated into two tubes 50 μL each tube), F2= 20 μL 5 × PG buffer, 10 μL 2 mM dNTPs, 2 μL 10 ng DNA template, 1 μL phusion polymerase, 5 μL mutant forward primers, 5 μL PbrR reverse primers and 57 μL dH_2O separated into two tubes, with 50 μL in each tube. The purification PCR1 product was purified following the procedure described above.

The second involved construction of the second fragment PCR2. This was made from 20 μL 5 × GC buffer, 10 μL 2 mM dNTPs, 10 μL F1, 10 μL F2, 1 μL phusion polymerase, and 49 μL dH_2O.

The third involved augmenting PCR2. The augmentation comprised 30 µL 5 × GC buffer, 20 µL 2 mM dNTPs, 10 µL PbrR forward primer, 10 µL PbrR reverse primer, 2 µL phusion polymerase, 28 µL dH_2O. These were added immediately after the cycling of PCR2 had finished and placed in the PCR machine (Techgene).

2.3. Microbiology methods

2.3.1. Bacterial strains, growth media, and antibiotics

Bacterial strains used in this study were *E. coli* BL21 (DE3) (Promega), for the over-expression of the protein PbrR, and *E. coli* DH5α (Invitrogen), for cloning and general plasmid amplifications.

Luria-Bertani broth (LB) (10 g NaCl, 5 g yeast extract, and 10 g tryptone in 1 L of dH_2O) was autoclaved at 121°C for 15 minutes and used for over expression. Agar testing was used to assess routine growth of bacteria. Bacterial colonies were cultivated on a solid medium. This comprised 40 g of test-sensitive agar per litre of dH_2O autoclaved at 121°C for 15 minutes. Sterile-filtered antibiotics were added to a final concentration of 50 µg/mL of kanamycin (Km) after cooling the medium to 22 °C. The mixture was then poured into plates (100 mm diameter). Kanamycin was dissolved in dH_2O. Super optimal broth with catabolite repression (SOC) medium (0.5 % yeast extract, 2 % tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose) was used for the recovery of the *E. coli* competent cells.

2.3.2. Heat shock transformation of *E. coli* competent cells

Transfection of *E. coli* with plasmid DNA involved two steps. The first involved binding the DNA to the competent cell surface. For this, 5 µL DNA recombinants or mutants were added to a tube containing 200 µL competent *E. coli* DH5α, mixed gently and placed on ice for 30 minutes.

The second involved allowing the DNA plasmids to enter the cell cytosol. The DNA plasmid was subject to a heat-pulse in the range 0–42 °C by placing the tube in a 42 °C plate for exactly 52 seconds. To recover the *E. coli*, the tube was placed on ice for 3 minutes, and 400 µL of SOC medium, containing no antibiotics, was added and

incubated at 37 °C for 1 hour. This step promoted the synthesis of enough plasmid-encoded antibiotic resistance prior to plating cells on appropriate antibiotic-containing LB agar plates (agar sensitive test).

2.3.3. Selecting transformed cells

Aliquots of 50, 75, 100 or 200 µL were spread onto Petri dishes of LB agar plates containing kanamycin (50 µg/mL); the dishes were then incubated (37 °C) overnight. This medium allows bacteria carrying the recombinant plasmid to grow. This is because the kanamycin resistance gene is an integral part of pET28a (+) and untransformed cells die on the same plate.

The single colonies that grew were transferred to 5 mL of liquid LB medium containing (50 µg/mL) kanamycin. To further ensure the growth of resistant cells, culture incubation at 37 °C was completed overnight; this resulted in a high-copy plasmid.

2.3.4. Extraction and purification of high-copy plasmids

To extract the high copy plasmid, a QIA-prep spin miniprep kit (Qiagen, UK) was used; the manufacturer's protocol was followed. Cells from the overnight culture (described above) were harvested, suspended, and lysed under alkaline conditions in 250 µL lysis buffer P2 [200 mM NaOH, 1% SDS (w/v)], in the presence of RNAse 250 µL of buffer P1 (50 mM Tris-HCl, pH 8.0; 10 mM EDTA; 100 µg/ml RNase A).

The lysate was then neutralised and adjusted to a high salt binding concentration by adding 350 µL neutralisation buffer N3 (3.0 M potassium acetate, pH 5.5); the mixture was thoroughly and gently mixed (the mixture contained high salt concentration precipitates, denatured chromosomal DNA, proteins, cellular debris, and SDS) while renaturing the plasmid DNA. The solution was then centrifuged for 10 minutes at 13,000 rpm; a compact white pellet formed as a result. The supernatant was then transferred to the supplied spin column and centrifuged at the same speed for 1 minute to remove RNA, cellular proteins, and metabolites; plasmid DNA remained on the silica membrane.

The supplied columns have a silica membrane for selective absorption of plasmid DNA in high-salt buffer and elution in low-salt buffer. The silica membrane was washed with 750 μL buffer PE, centrifuged for 1 minute, and the flow-through discarded; the spin column was then placed in a clean 1.5 mL Eppendorf tube. Plasmid DNA was then eluted with 30 μL nuclease-free water and collected by centrifuging for 1 minute.[23] The resultant solution contained pET28a (+) with the cloned PbrR gene. To check the cloning for plasmid 7 μL of the recombinant DNA plasmid or the mutant plasmid, 1 μL restriction enzyme, and 1 μL buffer E, the mixture was incubated at 37 °C for 1 hour and then run in an agarose gel and the bands were visualised under the UV illuminator as in chapter 4.

2.4. PbrR overproduction

Overnight pre-cultures were grown aerobically at 37 °C and shaken at 1700 rpm. The pre-cultures were then used to inoculate 1 L of LB medium containing kanamycin (50 μg/mL). The cells were grown until the $O.D_{600}$ was 0.6. Next, IPTG (1 mM) was added. The cells were grown overnight at room temperature (20–25 °C). The cells were then harvested by centrifugation and stored at –20 °C.

2.5. Extraction and purification of recombinant and mutant PbrR

2.5.1 Preparation of the crude extract

Cells were harvested from the culture for 10 minutes by centrifugation at 8000 rpm and 4 °C. The cells were then broken open by sonication.

2.5.1.1. Sonication

This method is used to disrupt biological membranes of bacteria by exposing them to high-frequency sound waves. [24] After suspending cells in 10ml of lysis buffer [0.1 M Phosphate buffer pH 7.4 (1 M K_2HPO_4, 1 M KH_2PO_4, 70 μL BME, diluted to 1 L), 2 M imidazole, 1 mM 2-mercaptoethanol (BME)], the suspension solution was sonicated four times for 30 seconds; this broke open the cell walls, whilst avoiding overheating the solution. The solution was then centrifuged at 20,000 rpm for 1 hour at 4 °C. The supernatant was then separated and collected for further analysis.

2.5.2. Protein purification

The protein in the supernatant from the section (2.5.1.1) above was purified by the His-tag column (HiTrap affinity column, GE Healthcare) with 10 ml binding buffer (0.1 M phosphate buffer, 2 M imidazole), then 1 ml nickel sulphate (0.1 M) was passed through the column to bind with His, and then 5 ml distilled water was passed to wash the column; all supernatants of the PbrR and mutant proteins were passed through the column, and 500mM imidazole salts were added to collect the pure protein; a PD-10 desalting column was used to remove all salts in the protein solution.

2.5.3 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

The purity of the protein samples was assessed using SDS-PAGE (as described by Laemmli).[25] This, together with a molecular weight marker, permitted estimation of the protein molecular weight.

The acrylamide resolving gel (15 %) was prepared using the procedure described .[26]

Preparation involved mixing double deionised water (ddH_2O) (3.4 ml), degassed acrylamide (30 %, 4.0 ml) (BioRad), gel buffer; Tris (Tris HCl (conc.), dH_2O). pH 8.8 (2.5 ml), and (10 %, 100 μL) sodium dodecyl sulphate (SDS). Preparation of the stacking gel was the same, save that the pH of the gel buffer was 6.8. Immediately before the gel was poured, ammonium persulphate (APS) (10 %, 100 μL) was added to the acrylamide resolving gel and N, N, N', N'-tetramethylethylenediamine (TEMED) (15 μL) was added to the stacking gel. Stirring TEMED facilitates gel polymerisation.

The resolving gel was poured into the plate until it was approximately 1 cm from the top. Then water-saturated ethanol was poured on top of the gel to prevent the gel from contact with the air. After the resolving gel had set, distilled water was used to wash away the ethanol. Next, the stacking gel was poured onto the resolving gel, and a comb pushed into the stacking gel.

Each sample (100 µL) was mixed with (900 µL) loading buffer; this buffer consisted of electrophoresis buffer, glycerol (3.0 ml), 0.5 % bromophenol blue (0.2 ml), mercaptoethanol (0.8 ml) and dH_2O (5.0 ml). The resultant mixture (samples and loading buffer; i.e. new samples) were then heated to 95 °C for 5 minutes in order to denature the protein. The new samples were then loaded onto the gel, and run at 150 V for 45 minutes in an electrolyte solution of 1 × running buffer. The gel was then stained with Coomassie solution for 30 minutes. Coomassie solution consisted of Coomassie Blue (1.25 g), methanol (500 ml) and acetic acid (100 ml) diluted to 1 L with distilled water. The gel was immersed in de-staining solution. De-staining solution consisted of acetic acid (50 ml) and industrial methylated spirits (IMS) (100 ml) diluted to 500 mL with distilled water.

2.7. Biochemical methods

2.7.1. Preparation of glutathione/Pb and Pb-PbrR protein for UV/Vis and NMR study.

2.7.1.1 UV/Vis spectroscopy Pb-GSH.

Tris-HCl buffer (100 mM, pH 8.5) was prepared and degassed with nitrogen gas for 1 hour before the preparation of the stock solution. A 60 µM GSH solution was prepared and 1/3 equivalent of Pb $(NO_3)_2$ of the total GSH concentration [20 µM Pb $(NO_3)_2$] was added to the GSH solution at pH 8.5. The spectra were recorded over the range of 200 nm to the 450 nm, using Tris-HCl buffer as a blank.

2.7.1.2. UV/Vis spectroscopy of the Pb-PbrR protein.

A 5 ml volume of 25 µM PbrR apo-protein (either recombinant mutant protein) was added to a dialysis bag (MWCO 10 kDa) and dialysed against 1 L of the lead loading buffer (1 mM $Pb(NO_3)_2$, 20 mM Bis-Tris pH 7.0, 500 mM $NaNO_3$, 1 mM 2-mercaptoethanol , and 5 % v/v glycerol) at 4 °C for 3 hours. The outside of the dialysis bag was thoroughly washed before the sample had been dialysed against 1 L of Lead Washing Buffer (20 mM Bis-Tris pH 7.0, 500 mM $NaNO_3$, 1 mM ME, and 5 % v/v glycerol) at 4 °C for 1 hour. This washing step was repeated three times. The protein sample was then removed from the dialysis bag and concentrated down to a

volume of approximately 200 µL using an Amicon cell (viva spin molecular weight cut off 10 kDa) at 4 °C. (All buffers were purged with nitrogen gas overnight prior to being used). UV/Vis spectrophotometry was used to assess the purity, determine the concentration, and characterise the recombinant and mutant PbrR. All spectrophotometric experiments were performed at room temperature (20–25 °C) using a spectrophotometer (Perkin Elmer Lambda 35).The blank UV/Vis spectrum was first taken against a lead washing buffer. Next, the absorption spectra of the Pb-PbrR complex were recorded, as were the absorption spectra of the recombinant apo-PbrR and the mutation protein.

2.7.1.3. Pb^{2+} titrations with recombinant native PbrR(Pp)

The spectra were recorded in the range of 250 nm to 450 nm. Metal binding titrations of the PbrR protein was carried out by adding small aliquots of (5µL 40µM) Pb(NO3)2 into a PbrR Protein solution (1.0 mL; 2 μ M) The absorbance due to the protien was subtracted from each spectrum so that the observed spectrum is due to the formation.

2.7.3. Preparation of Pb-GSH and Pb-PbrRp sample for NMR Study.

The ^{207}Pb NMR spectra were collected at 300 K and resonance frequencies of 62.95 and 300.14 MHz using a Bruker AMX 300 spectrometer equipped with a 10 mm broad-band probe. The ^{207}Pb NMR data were acquired using a 90° pulse, a sweep width of 62.8 kHz, and 64K data points with a 1 s recycle delay between scans. Approximately 10000–40000 scans were accumulated. Spectra were processed using exponential line broadening between 25 and 200 Hz .

2.7.3.1 NMR spectroscopy of Pb-GSH.

50 mL of Pb-GSH complex solution containing 270 mM GSH and 90 mM $Pb(NO_3)_2$ in 100 mM Tris-HCl buffer at pH 8.5 was prepared. The white precipitate initially formed on mixing $Pb(NO_3)_2$ and GSH in the above buffer solution was redissolved by adjusting the pH of the solution from pH= 2.24 to pH 8.5, using 5 M NaOH. Then the clear solution was made up to the final volume (50 mL) using 100 mM Tris-HCl

buffer at pH 8.5. 100 µl of 10 % D_2O/ 90 % H_2O (100 µl D_2O in 900 µl H_2O) was added to 900 µl of the above Pb-GSH complex solution.

2.7.3.2. NMR spectroscopy of Pb-PbrR.

The Pb (II)-PbrR complex was prepared as described above (2.7.1.2); then 100 µl of 10 % D_2O/ 90 % H_2O was added to 900 µl of the complex. Pb-PbrRp was run in an NMR machine for roughly 16 h.

2.8. References

1. T. Brown, *Gene cloning and DNA analysis: an Introduction*, Oxford: Blackwell Pub. 2006.

2. A. Lehninger, D. Nelson and M.Cox, *Principles of Biochemistry*, (5thEd.), New York: Worth, 2008.

3. P. Yakovchuk, E. Protozanova and M. Frank Base-stacking and base-pairing contributions into thermal stability of the DNA doublehelix, *Nucleic Acids Research*, 2006, **34**, 564–574.

4. K. Edwards, D. Brown, N. Spink, J. Skelly and S. Neidle, Molecularstructure of the B-DNA dodecamer $(CGCAAATTTGCG)_2$: An examination of propeller twist and minor-groove water structure at 2·2Å resolution, *Journal of Molecular Biology*, 1992, **226**, 1161–1173.

5. W. Fiers, R. Contreras, G. Haegemann, R. Rogiers, A. Van de voord, H. Van heuverswyn, J. Van herreweghe, G. Vokkaert and M. Ysebaert, Complete nucleotide sequence of SV40 DNA, *Nature*, 1978, **273**,113–120.

6. J. Venter M. Adams and E.Myers, The sequence of the human genome, *Science,* 2001, **291**, 1304–1305.

7. C. Mathews, K. Holde, and K. Ahern, *Biochemistry*, (3rd ed.), San Francisco, CA: Benjamin/Cummings, 2000.

8. L. Wang, and P. Schultz, Expanding the Genetic Code, *Angewandte Chemie International Edition*. 2004, **44**, 34-66.

9. H. Horton, L. Moran, K. Scrimgeour, M. Perry and J. Rawn, *Principles of Biochemistry*, (3rdEd.), Upper Saddle River, NJ: Pearson Prentice Hall, 2002.

10. S. Hadi and R. Yuan, Complementation in vitro by mutant restriction enzymes from *Escherichia coli* K, *Journal of Biological Chemistry*, 1974. **249**, 4580–4586.

11. M. Joshi and J. Deshpande, Polymerase chain reaction: methods, principles and application, *International Journal of Biomedical Research*. 2011, **1**, 81-88.

12. R. Saiki, D. Gelfand, S. Stoffel, S. Scharf, R. Higuchi, G. Horn, K. Mullis and K. Erlich, Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase, *Science*, 1988, **239**, 487-491.

13. H. Chawla, *Basic techniques, Introduction to Plant Biotechnology*, (2nd Ed), Science Publishers, Inc. Enfield, NH, U.S.A., 2004.

14. N. Stellwagen, *DNA gel electrophoresis. Nucleic Acid Electrophoresis Laboratory Manual,* Springer Verlag, Berlin-Heidelberg-New York, 1998.

15. D. Smith, *Agarose gel electrophoresis. Methods in Molecular Biology: Transgenesis Techniques*, (D Murphy and DA Carter, Ed.). Humana Press Inc., Totowa, NJ, 1993.

16. P. Barril and S. Nates, *Introduction to Agarose and Polyacrylamide Gel Electrophoresis Matrices with Respect to Their Detection Sensitivities*, INTECH Open Access Publisher, 2012.

17. B. Kurien and R. Scofield, *Protein Electrophoresis: Methods and Protocols*, New York: Humana press, 2012.

18. S. Magdeldin, *Gel Electrophoresis - Principles and Basics*, INTECH. http://www.intechopen.com/books/gel-electrophoresis-principles-and-basics. 2012.

19. X. Jiang, C. Wang, P. Zhang and Z. He, Cloning and expression of Mycobacterium bovis secreted protein MPB83 in *Escherichia coli*, *Journal of Biochemistry and Molecular Biology,* 2006, **39**, 22-25.

20. QIA-quick® *spin handbook*, February (2008).

21. N. Qiagen. *HotStarrTaq® PCR handbook.* Venlo, Netherlands, February 2008.

22. Promega Corporation. *Protocols and Application guide,* (3rd ed.), 1996.

23. N. Qiagen, *QIA-Prep® (miniprep) mini prep handbook*, February 2008.

24. L. Benov and J. Al-Ibraheem, Disrupting *Escherichia coli*: A comparison of methods, *Journal of Biochemistry and Molecular Biology.* 2002, **35**, 428-431.

25. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature,* 1970, **227**, 680–685.

26. H. Schagger, Tricine–SDS-PAGE, *Nature Protocols*, 2006, **1**, 16–22.

CHAPTER 3: BIOINFORMATICS

3.1. Bioinformatics

Bioinformatics can be defined as a management information system for molecular biology. It has numerous applications.[1] It can be more broadly explained as informing the theory of Biology in terms of molecular interactions and implementing "informatics techniques" (originated from subjects like statistics, computer science and applied maths) so that information linked with molecules is comprehended and arranged on a larger level.

Lately, bioinformatics has gained much interest through different subjects like non-traditional biological sciences, mathematics, and information technology.[2] This is because of the vast collection of private and public biological data available and the need for these data to be modified into beneficial information and knowledge. The key tasks in bioinformatics are to understand the patterns, structures and correlations in biological data. Later on, there can be sensible use of the information and knowledge from these subjects for applications that include biological control, genome analysis and drug discovery.[3]

Thus, bioinformatics can be regarded as the mix of certain scientific fields, comprising biochemistry, biology, computer science and mathematics. It requires the use of statistical tools and computer technology to handle and evaluate a vast collection of biological data. Specifically the data are mainly about protein interactions, protein sequences, gene expression profiles, protein structures, RNA and DNA sequences and structures. Numerous databases are currently used to store and retrieve biological data.[4] The data are stored in formats that can be interrogated and compared using a further comprehensive array of computational packages.[5]

Bioinformatics is dependent on a number of subdisciplines and activities: (a) gathering

biological data into structural databases such as the Protein Data Bank PDB, [6] and Macromolecular Structure Database (MSD),[7] DDBJ, GenBank and NCBI (http://www.ncbi.nih.gov/), (b) collecting of biological data like the Human Genome Project, gene discovery and gene expression,[8] (c) interrogating biological data structures; for example, by using structural bioinformatics,[9] (d) probing functional and structural biological databases like SWISS-PROT, EMBI, and EMBL,[10] (e) using biological modelling like molecular modelling,[11] probabilistic modelling,[12] comparative modelling,[13]and Hidden Markov model HMM,[14] (f) biological data management; for instance, bioinformatics data warehousing [15] and Sequence Retrieval Systems (SRS),[16] (g) data analysis and investigation like bioinformatics data mining, and biological comprehension; for instance, pattern matching, machine learning and visualisation of biological sequences,[17] (h) biological processes, and (i) sequence evaluation: sequence assembly and alignment.[18]

The overall purpose of these approaches is to analyse and ascertain relationships in biological data and to use statistical techniques to examine, understand, and mine datasets. The basic description of bioinformatics in presented in figure 3.1 [19-21]

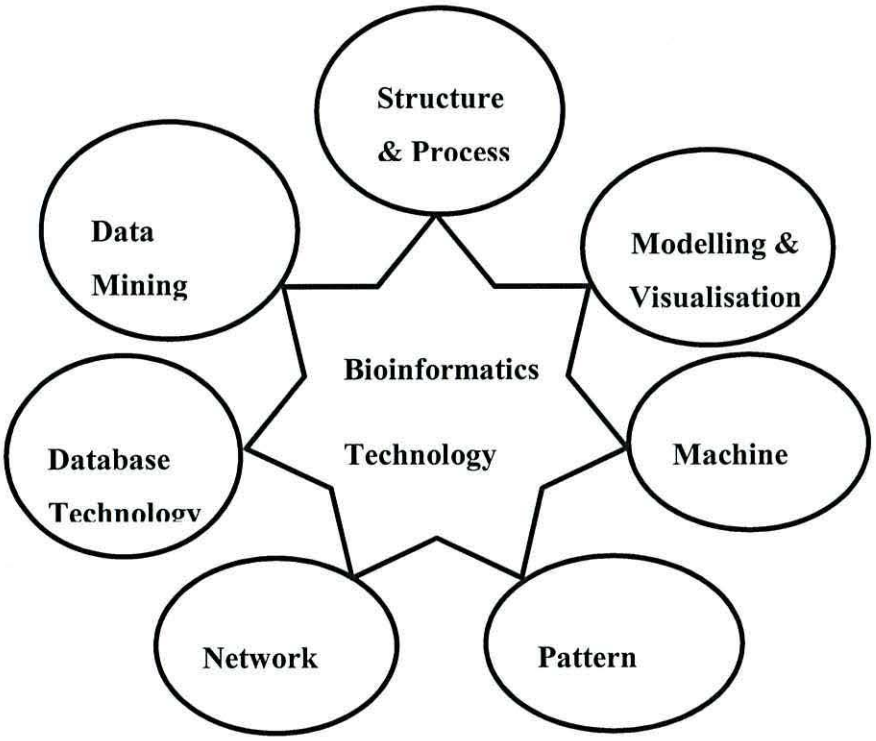


Figure 3.1. Technologies within bioinformatics.

Source Chen, Yi-Ping Phoebe. 2005 [22]

The study of major biological issues increasingly involves the field of bioinformatics. This is due to the exponential burst of sequence and structural information as time has passed.[23]

Huge volumes of data are created with the rise of high-throughput technologies like DNA microarrays and whole genome sequencing. A need exists for proficient management of these biological data. Nevertheless, one sole database is not able to deliver solutions to the complicated questions posed by biologists in certain situations. Handling and combining the prevailing biological databases is a problem for data management.[24] Thus, some other main difficulties in bioinformatics are incorporation or

assembling of information from many databases to be able to determine solutions and discover new knowledge.[25,26]

3.2. An Overview of Bioinformatics Technologies

In the course of this thesis, a number of bioinformatics tools have been used to identify new organisms containing genes that code for highly specific lead-binding proteins similar to those already discovered in *Ralstonia metallidurans* CH34. Specifically, gene and protein sequence comparisons were made in the initial stages of target selection and for verifying successful cloning and expression of novel protein and mutants. In addition, the structural and functional properties of related zinc- and lead-binding systems were further studied by looking at 3D structures available in the PDB databank. This review will now look at some of the key bioinformatics tools utilised in more detail.[22]

3.2.1. Sequence analysis

Sequence analysis determines similar structural and functional factors, and discrepancies between various biological sequences. This analysis involves comparison of thoroughly studied and annotated (known) gene or protein sequences with new (unknown) sequences. Sequence alignment, which is a suggested technique for sequence comparison, is a method involving base-by-base/ amino acid residue by residue comparison of two (pair-wise) or more (multiple) sequences.[27] It is performed by looking for a range of individual characters or character patterns that are arranged in the same order in the sequences. The method of string matching is commonly used to identify an identical character or character patterns.[28]

3.2.1.1. Protein Sequences

The study of proteins and their interaction is structural analysis. Proteins are a mixture of

amino acids arranged in a specific order. They are large molecules which are needed for the function, structure and management of the body's organs, tissues, and cells.

Distinctive families of proteins are recognised, which can be grouped either on a functional or structural basis. [29] Hierarchical structures of proteins exist, comprising primary, secondary and tertiary structures. This means that proteins can be seen as 3D structures at the molecular level. At present, research on protein structural analysis includes comparison and prediction of protein structures. [30] Scientists have long recognised that two similar amino acid sequences may fold into a similar 3D structure and have an identical functional role in different regulatory or biochemical pathways. These are stated to be homologous sequences if two similar sequences are from separate organisms.[31] Ascertaining the nature of a protein requires determination of the existence of homologous sequences, which in turns helps a lot in the formulation of new medicines and even in the presentation of phylogenetic analysis. Interestingly, very different protein sequences can occur that have the same functionality; e.g. the nitrite reductase family which consists of family members that are either copper based proteins [32] or haeme cd^1 proteins. [33]

3.3. Software and Database tools for Gene and Protein Analysis

3.3.1. The EMBL nucleotide sequence data library

The European Bioinformatics Institute (EMBL–EBI) has given public access to many typical sequence analysis applications since 1998. [34,35]These comprise multiple sequence alignment tools (http://www.ebi.ac.uk/Tools/sequence.html), such as T-Coffee, [36] Kalign, [37] MUSCLE,[38] MAFFT,[39] and ClustalW,[40] and sequence similarity search services (http://www.ebi.ac.uk/Tools/similarity.html) such as InterProScan, [41] BLAST [42,43]and FASTA. [44] The PERL-CGI job dispatcher framework regulates these services and handles job submission and result representation. More than 16 million jobs were

managed through this infrastructure in 2009. [45]

3.3.2. GenBank

GenBank is a wide-ranging database that includes publicly accessible nucleotide sequences for more than 300,000 organisms. These are mainly attained via batch submissions from bigger sequencing projects such as whole genome shotgun (WGS) and environmental sampling projects and from individual laboratories submissions. They are named at the genus level or lower. The majority of submissions are made through standalone Sequin programs [46] or Web-based BankIt. [47] GenBank can be accessed via the NCBI Entrez retrieval system, which combines data from the key DNA and protein sequence databases and also includes domain information, taxonomy, protein structure, genome mapping and even biomedical journal literature through PubMed. Sequence similarity searches of GenBank and other sequence databases are done through BLAST. Daily updates of the GenBank database and bi-monthly releases are obtainable by FTP.[48] Global coverage is ensured through regular data exchange with the DNA Data Bank of Japan and the European Molecular Biology Laboratory Nucleotide Sequence Database in Europe.

3.3.3. BLAST

In this thesis, the Basic Local Alignment Search Tool (BLAST) program was used to obtain the PbrR amino acid sequence for the gene in *Ralstonia metallidurans* CH34, to compare basic biological sequences, and to identify other genes with similarity to it. In bioinformatics, BLAST is an algorithm that compares basic biological sequence information like nucleotide or amino acid sequences. [49] The initial clues regarding the function of a newly sequenced gene are provided through detection of sequence homology to an identified protein or family of proteins. However, to differentiate chance similarities from biologically important relationships, all tools utilise some means of

quantifying, via a Figure of Merit, the level of similarity among sequences. The majority of sequence comparison programmes utilise some variation of the same complex programming algorithm.[50,51] Complex programming algorithms are not feasible for searching large databases without employing a supercomputer [52] or other special purpose hardware [53] due to their computational needs. Creation of quick heuristic algorithms, [54] which work on estimating the above techniques, now enable searching of large databases on standard desktop computers via web-based interfaces. For instance, locally similar areas between two sequences, as per identities but not gaps, are initially found through the FASTA program. [55] These areas are then rescored through a measure of similarity among the residues, like the PAM matrix[56] that enables conservative replacements and even identities to increase the similarity points. Heuristic techniques, like FASTA, have been well utilised and have determined several biologically important relations, [6] even though they provide quite indirect estimations of minimal evolution measures.

3.3.4. Clustal

The comparison or placement of protein sequences is one of the keystones of modern bioinformatics. Biologists are able to examine the sequence patterns that were preserved through evolution and the ancestral relations among organisms with the help of multiple sequence placements. Alignment of sequences can either be across their complete length (global alignment) or only in specific areas (local alignment). For global multiple sequence alignment, the most extensively used programs are the Clustal range of programs. The first Clustal program was formatted particularly to work proficiently on personal computers, the computing capabilities of which were quite weak when compared to current standards.[57] This included a memory-efficient dynamic programming algorithm [58]and a progressive alignment strategy. [59,60] A range of pair-wise alignments allowed gradual build-up of a multiple alignment according to the branching order in a guide tree. The Unweighted Pair Group Method with Arithmetic

Mean (UPGMA) approach was used to create the guide tree [61] and a rapid word-based alignment algorithm was used for the preliminary pre-comparison. A new release named ClustalV, [62] was developed in 1992. It integrated profile alignments (alignments of present alignments) and provided the facility to create trees from the multiple alignments by using the Neighbour-Joining (NJ) technique. ClustalW, [63] released in 1994, was the third generation of the series. It included many enhancements to the alignment algorithm, such as the automatic option of an appropriate residue comparison matrix at every stage in the multiple alignment, position-specific gap penalties and sequence weighting. A more sensitive dynamic programming algorithm was also used in place of the approximate word search used for the pre-comparison phase. NJ took the place of the dendogram construction by UPGMA. .

3.3.5. SWISS-PROT

The existing knowledge in the Life Sciences is linked with amino acid sequences through the SWISS-PROT protein knowledge base (http:// www.expasy.org/sprot/ and http://www.ebi.ac.uk/swissprot/). SWISS-PROT [64] is a protein sequence and knowledge database that is appreciated for its high quality annotation, minimal redundancy, direct links to specialised databases and the utilisation of standardised terminology. SWISS-PROT provides annotated entries for all species. However, to ensure the existence of high quality data for representative members of all protein families, more focus is placed on the explanation of entries from the Human Proteomics Initiative (HPI project) [10] and other model organisms. This is similar to what is done for microbes by the High-quality Automated and Manual Annotation of microbial Proteomes (HAMAP) project; a portion of the annotation can be passed to other family members. As an addition, those protein sequences not represented in SWISS-PROT are included in TrEMBL.[65] Sending findings to SWISS-PROT at swiss-prot@expasy.org allows researchers to add their knowledge to the scientific community. The format of EMBL Nucleotide Sequence

Database for standardisation purposes is followed by SwissProt. [66] The data in SWISS-PROT originate from translations of DNA sequences in the EMBL database revised from the Protein Identification Resource collection [67] (PIR, Washington, D.C.), taken from the literature and directly sent by researchers. Every 3 months, SWISS-PROT is distributed on CDROM and magnetic tape. The EBI network servers allow new entries to be recovered between releases. [68]

3.3.6. Protein Data Bank

The Protein Data Bank (PDB) is a library for biological macromolecular crystal structures. It was formed in 1971 at Brookhaven National Laboratories (BNL).[69] Improvement in the technology for all features of the crystallographic process, modifications in the views of community regarding data sharing, and new structures found through the nuclear magnetic resonance (NMR) method led to a considerable increase in the number of deposited structures in the 1980s. Most of the journals required a PDB accession code prior to publication of crystallographic structures by the beginning of 1990s. The instructions published by the International Union of Crystallography (IUCr), requiring data deposition for all structures, were implemented by at least one funding agency (National Institute of General Medical Sciences).[70]

The use of the PDB was initially restricted to a limited group of experts working on structural research. At present, depositors to the PDB have different proficiencies regarding the methods of X-ray crystal structure determination, theoretical modelling, cryoelectron microscopy and NMR. Those utilising this data bank are a diverse set of researchers in chemistry, biology and computer science, including students and educators at different levels. The scientific community recognises a need for new methods of collection, organisation and distribution of data. This is because of the rapid increase in data that will occur due to the structural genomics initiative and the rise in acknowledgement of the value of the data as it will help in understanding the biological

function.[71]

3.3.7 Collaborative Computational Project, Number 4

The Collaborative Computational Project, Number 4 software suite is a collection of programs and associated data and software libraries which can be used for macromolecular structure determination by X-ray crystallography. [72]

3.9. References

1. N. Luscombe, D. Greenbaum and M. Gerstein, What is Bioinformatics? A Proposed Definition and Overview of the Field. *Methods of Information in Medicine,* 2001, **40**, 346-358.

2. U. Shrivastava, Incorporating Bioinformatics into Biological Science in Nepal: Prospects and Challenges, *Academic Voices: A Multidisciplinary Journal*, 2013, **2**, 78-85.

3. R. Clarke, H. Ressom, A. Wang, V. Xuan, M. Liu, E. Gehan and Y. Wang, The properties of high-dimensional data spaces: implications for exploring gene and protein expression data, *Nature Reviews*, 2008, **8**, 37-49.

4. B. Pierre and S. Brunak, *Bioinformatics: the Machine Learning Approach*, 2nd Ed, Cambridge, Mass: MIT Press, 2001.

5. R. Grant, *Computational Genomics: Theory and Application*, Wymondham: Horizon Bioscience, 2004.

6. H. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, H. Weissig, I. Shindyalov and P. Bourne, The Protein Data Bank, *Nucleic Acids Research*, 2000, **28**, 235-242.

7. D. Moss, S. Jelaska, and S. Pongor, *Essays in Bioinformatics*, Amsterdam: IOS Press. 2005.

8. M. Adams, J. Kelley, J. Gocayne, M. Dubnick, M. Polymeropoulos, H. Xiao, C. Merril, A. Wu, B. Olde, R. Moreno, A. Kerlavag, W. Mccombie and J. Venter, Complementary DNA sequencing: expressed sequence tags and human genome project, *Science,* 1991,**252**, 1651-1656.

9. M. Cohen, Structural bioinformatics-based design of selective, irreversible kinase inhibitors, *Science,* 2005, **308**, 1318-1321.

10. R. Apweiler and A. Bairoch, The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000, *Nucleic Acids Research*, 2000, **28**, 45-48.

11. U. Pieper, N. Eswar, H. Braberg, M. Madhusudhan, F. Davis, A. Stuart, N. Mirkovic, N. Mirkovic, A. Rossi, M. Marti-Renom, A. Fiser, B. Webb, D. Greenblatt, C. Huang , T. Ferrin and A. Sali, MODBASE, a database of annotated comparative protein structure models and associated resources, *Nucleic Acids Research,* 2004, **32**, 217.

12. S. Eddy, A new generation of homology search tools based on probabilistic inference, *Genome Informatics,* 2009, **23**, 205-211.

13. C. Combet, M. Jambon, G. Deleeage and C. Geourjon, Geno3D: automatic comparative molecular modelling of protein, *Bioinformatics*, 2002.**18**, 213-214.

14. P. Baldi, Y. Chauvin, T. Hunkapiller and M. McClure, Hidden Markov models of biological primary sequence information. *Proceedings of the National Academy of Sciences U. S. A.,* 1994, **91**, 1059-1063.

15. C. Schonbach, P. Kowalski-Saunders and V. Brusic, Data warehousing in molecular biology, *Briefings in Bioinformatics,* 2000, **1**, 190-198.

16. G. Perriere and M. Gouy, WWW-query: An on-line retrieval system for biological sequence banks, *Biochimie,* 1996, **78**, 364-369.

17. I. Witten and E. Frank, *Data mining: Practical Machine Learning Tools and Techniques*, 2nd Ed, Amsterdam: Morgan Kaufman, 2005.

18. D. Dean, M. Patton and R. Stephens, Direct sequence evaluation of the major outer

membrane protein gene variant regions of *Chlamydia trachomatis* subtypes D', I', and L2''', *Infection and Immunity,* 1991, **59**, 1579-1582.

19. A. Baxevanis and B. Ouellette, *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*, 2nd Ed, New York: Wiley-Interscience, 2001.

20. D. Kuonen, Challenges in bioinformatics for statistical data miners, *Bulletin of the Swiss Statistical Society*, 2003, **46**, 10-17.

21. D. Westhead, J. Parish and R. Twyman, *Bioinformatics*, Instant Notes in Bioinformatics, BIOS Scientific Publishing, 2002.

22. T. Attwood, A. Eriksson and E. Bongcam-Rudloff, Concepts, historical milestones and the central place of bioinformatics in modern biology: *A* European Perspective, www. Intechopen. Com, 2011.

23. L. Ohno-Machado, S. Vinterbo and G. Weber, Classification of gene expression data using fuzzy logic, *Journal of Intelligent and Fuzzy Systems*, 2002, **12**, 19-24.

24. S. Ng and L. Wong, Accomplishments and challenges in bioinformatics, *IT Professional*, 2004, **6**, 44- 50.

25. L. Wong, Technologies for integrating biological data, *Briefings in Bioinformatics*, 2002, **3**, 389-404.

26. W. Kleisli, A functional query system, *Journal of Functional Programming*, 2000, **10**, 19-56.

27. J. Nilsson, B. Persson and G. Heijne, Comparative analysis of amino acid distributions in integral membrane proteins from 107 genomes. *Proteins: Structure, Function, and Bioinformatics,* 2005, **60**, 606-616.

28. G. Benson, Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research,* 1999, **27**, 573-580.

29. A. Andreeva, D. Howorth, J. Chandonia, S. Brenner, T. Hubbard, C. Chothia and A. Murzin, Data growth and its impact on the SCOP database: new developments, *Nucleic Acids Research*, 2008, **36**, 419-425.

30. M. Wiederstein, and M. Sippl, ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, *Nucleic Acids Research,* 2007, **35**, 407-410.

31. M. Simmons and H. Ochoterena, Gaps as characters in sequence-Based phylogenetic analyses, *Systematic Biology*, 2000, **49**, 369-381.

32. F. Fenderson, S.Kumar, E. Adman, M. Liu, W. Payne and J. LeGall, Amino acid sequence of nitrite reductase: a copper protein from *Achromobacter cycloclastes*, *Biochemistry,* 1991, **30**, 7180-7185.

33. B. Oakley, C. Francis, K. Roberts, C. Fuchsman, S. Srinivasan and J. Staley, Analysis of nitrite reductase (nirK and nirS) genes and cultivation reveal depauperate community of denitrifying bacteria in the Black Sea suboxic zone, *Environmental Microbiology,* 2007, **9**, 118-130.

34. H. McWilliam, F. Valentin, M. Goujon, W. Li, M. Narayanasamy, J. Martin, T. Miyar and R. Lopez, Web services at the European Bioinformatics Institute 2009, *Nucleic Acids Res*, 2009, **37**, 6-10.

35. C. Brooksbank, G. Cameron and J. Thornton, The European Bioinformatics Institute's data resources, *Nucleic Acids Res*, 2010 , **38**, 17–25.

36. C. Notredame, D. Higgins and J. Heringa, T-Coffee: A novel method for multiple

sequence alignments, *Journal Molecular Biology*, 2000, **302**, 205–217.

37. T. Lassmann and E. Sonnhammer, Kalign an accurate and fast multiple sequence alignment algorithm, *BMC Bioinformatics*, 2005, **6**, 298.

38. R. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Research*, 2004, **32**, 1792–1797.

39. K. Katoh, G. Asimeno and H. Toh, Multiple alignment of DNA sequences with MAFFT. *Methods Moleculr Biology*, 2009, **537**, 39–64.

40. M. Larkin, G. Blackshields, N. Brown, R. Chenna, P. McGettigan, H. McWilliam, F. Valentin, I. Wallace, A. Wilm, R. Lopez, J. Thompson, T. Gibson and D. Higgins, ClustalW2 and ClustalX version 2.0, *Bioinformatics*, 2007, **23**, 2947–2948.

41. E. Quevillon, V. Silventoinen, S. Pillai, N. Harte, N. Mulder, R. Apweiler, and R. Lopez, InterProScan: protein domains identifier. *Nucleic Acids Research*, 2005, **33**, 116–120.

42. R. Lopez, V. Silventoinen, S. Robinson, A. Kibria and W. Gish, WU-Blast2 server at the European Bioinformatics Institute, *Nucleic Acids Research*, 2003, **31**, 3795–3798.

43. S. Altschul, T. Madden, A. Schaffer, J. Zhang, Z. Zhang, W. Miller and D. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 1997, **25**, 3389–3402.

44. W. Pearson, and D. Lipman, Improved tools for biological sequence comparison, *Proceedings of the NationalAcademic Science,* USA, 1988, **85**, 2444–2448.

45. M. Goujon, H. McWilliam, W. Li, F. Valentin, S. Squizzato, J. Paern and R. Lopez, *A* new bioinformatics analysis tools framework at EMBL–EBI, *Nucleic Acids Research*,

2010, **38**, 695–699.

46. S. Celniker, D. Wheeler, B. Kronmiller, J. Carlson, A. Halpern, S. Patel, M. Adams, M. Champe, S. Dugan, E. Frise, A. Hodgson, R. George, R. Hoskins, T. Laverty, D. Muzny, C. Nelson, J. Pacleb, S. Park, B. Pfeiffer, S. Richards, E. Sodergren, R. Svirskas, P. Tabor, K. Wan, M. Stapleton, G. Sutton, C. Venter, G. Weinstock, S. Scherer, E. Myers, R. Gibbs and J. Rubin, Finishing a whole-genome shotgun: Release 3 of the Drosophila melanogaster euchromatic genome sequence. *Genome Biology,* 2002, **3**, 1-14.

47. D. Benson, I. Karsch-Mizrachi, D. Lipman, J. Ostell and D. Wheeler GenBank, *Nucleic Acids Research,* 2008, **36**, 25-30.

48. D. Benson, I. Karsch-Mizrachi, D. Lipman, J. Ostell and E. Sayers, GenBank, *Nucleic Acids Research*, 2010, **38**, 46–51.

49. F. Altschul, W. Gish, W. Miller, E. Myers and D. Lipman, Basic local alignment search tool, *Journal of Molecular Biology*, 1990, **215**, 403-410.

50. S. Needleman, C. Wunsch, Algorithm for sequence similarity searches, *Journal Molecular Biol*ogy, 1970, 48:443-453.

51. P. Sellers, On the theory and computation of evolutionary distances, *Society for Industrial and Applied Mathematics,* 1974, 26, 787–793.

52. O. Gotoh and Y. Tagashira, Sequence search on a supercomputer, *Nucleic Acids Research,* 1986, **14**, 57-64.

53. A. Coulson, J. Collins and A. Lyall, protein and nuclic acid sequence database searching: a sutable case for parallel processing, *Computer Journal*, 1987, **30**, 420-424.

54. S. Waterman, Multiple sequence alignment by consensus, *Nucleic Acids Research*, 1986, **14**, 9095-9102.

55. A. Pevzner, Statistical distance between texts and filtration methods in sequence comparison, *Bioinformatics*, 1992, **8**, 121-127.

56. S. Altschul, T. Madden, A. Schaeffer and J. Zhang, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Research*, 1997, **25**, 3389.

57. G. Higgins and P. Sharp, CLUSTAL: a package for performing multiple sequence alignment on a microcomputer, *Gene*, 1988, **73**, 237-244.

58. E. Myers and W. Miller, Optimal alignments in linear space, *Bioinformatics*, 1988, **4,** 11-17.

59. D. Feng and R. Doolittle, Progressive sequence alignment as a prerequisite to correct phylogenetic trees, *Journal of Molecular Evolution*, 1987, **25**, 351-360.

60. T. Smith, The art of matchmaking: sequence alignment methods and their structural implications, *Structure*, 1999, **7**, 7-12.

61. G. Gall, C. Arnauld, E. Boilletot, J. Morisse and D. Rasschaert, Molecular epidemiology of rabbit haemorrhagic disease virus outbreaks in France during 1988 to 1995, *Journal of General Virology,* 1998, **79**,11-16.

62. A. Griffin and H. Griffin, CLUSTAL V: multiple alignment of DNA and protein sequences, *Methods Molecular Biology,* 1994, **25**, 307–318.

63. J. Thompson, D. Higgins and T. Gibson, CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting position-specific gap penalties and weight matrix choice, *Nucleic Acids Research*, 1994, **22**, 4673–4680.

64. C. O'Donovan, High-quality protein knowledge resource: SWISS-PROT and TrEMBL, *Briefings in Bioinformatics,* 2002, **3**, 275-284.

65. B. Apweiler, R. Schneider, M. Blatter, A. Estreicher, E. Gasteiger, M. Martin, K. Michoud, C. O'Donovan, I. Phan, S. Pilbout and M. Schneider, The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003, *Nucleic Acids Research,* **31**, 365-370.

66. G. Stoesser, W. Baker, A. van den Broek, E. Camon, M. Garcia-Pastor, C. Kanz and T. Kulikova, The EMBL Nucleotide Sequence Database, *Nucleic Acids Research,* 2002, **30**, 21-26.

67. D. Brett, H. Pospisil, J. Valcarcel, J. Reich and P. Bork, Alternative splicing and genome complexity, *Nature Genetics,* 2002, **30**, 29-30.

68. D. Emmert, P. Stoehr, G. Stoesser and G. Cameron, The European Bioinformatics Institute (EBI) databases, *Nucleic Acids Research,* 1994, **22**, 3445-3449.

69. F. Bernstein, T. Koetzle, G. Williams, E. Meyer, M. Brice, J. Rodgers, O. Kennard, T. Shimanouchi and M. Tasumi, The Protein Data Bank, A computer-based archival file for macromolecular structures, *Journal of Molecular Biology,* 1977, **112**, 535-542.

70. H. Berman, T. Battistuz, T. Bhat, W. Bluhm, P. Bourne, K. Burkhardt, Z. Feng, G. Gilliland, L. Iype, S. Jain, P. Fagan, J. Marvin, D. Padilla, V. Ravichandran, B. Schneider, N. Thanki , H. Weissig, J. Westbrook and C. Zardecki, The Protein Data Bank, *Acta Crystallographica Section D Biological Crystallography,* 2002, **58**, 899-907.

71. H. Berman, J.Westbrook, Z. Feng, G.Gilliland, T. Bhat, H.Weissig, I. Shindyalov and P. Bourne, *International Tables for Crystallography*, 2006, **F** 675-684.

72. M. Winn, C. Ballard, K. Cowtan, E. Dodson, P. Emsley, P. Evans, R. Keegan, E. Krissinel, A. Leslie, A. McCoy, S. McNicholas, G. Murshudov, N. Pannu, E. Potterton, H. Powell, R. Read, A. Vagin and K. Wilson, Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D Biological crystallography,* 2011, **67**, 235–242.

CHAPTER 4: TARGET SELECTION, CLONING, EXPRESSION, PURIFICATION AND GENERATION OF PbrR(Pp) RECOMBINANT AND MUTANT PROTEIN

4.1 Introduction

In this thesis, bioinformatics tools have been used to identify new organisms containing genes that code for highly specific lead-binding proteins similar to those already discovered in *Ralstonia metallidurans* CH34. After the target selection step, the generation of an expression clone that is capable of producing the protein of interest in a soluble form and in high yield is the rate-limiting step in protein production. Protein expression in heterologous systems often fails when the protein is expressed as insoluble aggregates, or cannot be purified by standard methods.[1] Thus, returning to the cloning stages to make a new construct with a different purification tag or moving the target gene from one vector to another or expressing the protein in a different host may be required to enable the system to express the protein. This is why protein expression pipelines can be time consuming.[2]

To date, many gene cloning and expression systems have been developed.[3] Usually the gene of interest is amplified and optionally modified by the polymerase chain reaction (PCR) and cloned into a plasmid, which in most cases is selected and replicated in *E. coli*. The cloned plasmid can then either be used for protein expression in an appropriate host itself or as a shuttle vector.[4]

4.2. Overview of Research Detailed in this Chapter:

Bioinformatics methods using BLAST [5] as a search tool identified a suitable target PbrR protein sequence from *Pseudomonas putida* KT2440. The PbrR gene was then isolated from *Pseudomonas putida* KT2440 by PCR. The forward and reverse DNA primers to amplify DNA by PCR were designed using restriction enzymes (*EcoRI* and *HindIII*).

After this, insertion of the gene into the *pET28a (+)* plasmid, containing a His-tag N-terminal coding sequence, after which it was transformed into the host bacteria *BL21* (DE3). The protein was then expressed by transferring the single colony to *E. coli* (*Rosetta*) and the single colony in *Rosetta* is expressed to get overproduction of protein after induction overnight at room temperature. The molecular weight of the protein subunits was determined by SDS-PAGE.

4.2.1. Target selection (BLAST, Clustal W).

The BLAST programme was used to identify other organisms containing the PbrR gene sequence that would code for a PbrR protein amino acid sequence. The PbrR amino acid sequence from *Cupriavidus metallidurans* CH34 was used to compare with other species. *Pseudomonas putida* KT2440 had a predicted 57% amino acid identity with *Ralstonia (Cupriavidus) metallidurans*. (table 4.1). This was selected for further study as it was considered more interesting to look at a related homologue that had the putative metal binding site coded, but was not fully identical to the already isolated and partially characterised systems. In addition our group had ready access to *P. putida* KT2440 from the group of Dr Gwenin.

To visualise any key differences in the protein sequences between different PbrR family members ClustalW [6] was used to compare and align different amino acids sequences (predicted or already isolated) in different organisms(figure 4.1). In addition Figure 4.2 shows the amino acid alignment of related MerR family members CueR in *E. coli* K12 (1Q05), ZntR 1Q08 in *E. Coli,* PbrR in *Pseudomonas putida* KT2440, Cad R in *Pseudomonas putida* 06909, MerR (Hg) MerR P22853.

Table 4.1. Predicted amino identity (%) and similarity (%) of PbrR from *Cupriavidus metallidurans* CH34 with other homologues.

Organism	Accession number	Amino acid length	Identities (%)	Similarities (%)
Cupriavidus metallidurans CH34	YP_584450.1	132	100	100
Pseudomonas aeruginosa LES431	YP_008942828.1	132	99	100
Acidovorax ebreus TPSY	YP_002553599.1	132	94	100
Hydrogenophaqa Sp.PBC	WP_009516069.1	132	80	100
Alcaligenes faecalis	WP_003804130.1	146	57	96
Polynucleobacter susp	YP_001156455.1	149	58	99
Pseudomonas chlorraphis	WP_009051350.1	149	85	96
Pseudomonas sp.CF149	WP_019828097.1	149	57	96
Pseudomonas sp. TJI-51	WP_009686167.1	150	61	96
Klebsiella pneumoniae subsp	NP_943470.1	144	56	96
Pseudomonas protegens Pf-5	YP_262951.1	149	57	96
Pseudomonas sp.S9	WP_010490645.1	150	61	96
Psedomonas chlororaphis	WP_023967953.1	149	57	96
Pseudomonas putida KT2440	NP_747241.1	147	57	96
Pseudomonas putida H8234	YP_008097902.1	147	57	96
Acinetobacter	WP_004641918.1	146	56	96

A

```
PbrR1 MEIRIGDLAKRSGCEVVTIRYYEKEGLLPKPARSGGNFRLYGEAHIERLQFIRHCRSLDM 60
PbrR2 MEIRIGDLAKRSGCEVVTIRYYEKEGLLPKPARSGGNFRLYGEAHIERLQFIRHCRSLDM 60
PbrR3 -MMRIGELGKKADCLVQTVRFYESEGLLPEPARSEGNFRLYDEVHLQRLLFIRRCRAKDM 59
PbrR4 --MKIGELAKATDCAVETIRYYERENLLPEPARSEGNYRLYTQAHVERLTFIRNCRTLDM 58
        ::**:*.* :.* * *:*:** *.***:**** **:*** :.*::** ***.**: **

PbrR1 TLSEIRALLGLRDNPMQDCGEVNTLLEAHIQQVEMRVSALLQLKRHLVDLREKCSGSRSV 120
PbrR2 TLSEIRALLGLRDNPMQDCGEVITLLEAHIQQVEMRVSALLQLKRHLVDLREKCSGSRSV 120
PbrR3 TLDEIRQLLNLRDRPELGCGEVNALVDAHIAQVRTKMKELRALERELMDLRRSCDSARTS 119
PbrR4 TLDEIRSLLRLRDSPDDACGSVNALIDEHIEHVQARIDGLVALQEQLVELRRRCNTQGS- 117
      **.*** ** *** *   **.* :*:: ** :*. ::. *  *:..*::**. *.   :

PbrR1 EACGILQGLGNCNCHGESATNSQTSG----- 146
PbrR2 EACGILQGLG--NCHGESATNSQTSG----- 144
PbrR3 RECGILNSLA------EPA------------ 132
PbrR4 -ECAILQQLETNGAVSVPDTEHSHVGRSHGH 147
        *.**: *         .
```

The (*) (asterisk) indicates positions which have a single, fully conserved residue, (:) (Colon) indicates conservation between groups of strongly similar properties, and (.) (Period) indicates conservation between groups of weakly similar properties.

B.

	PbrR1	PbrR2	PbrR3	PbrR4
PbrR1	100.00	99.31	56.82	52.11
PbrR2	99.31	100.00	56.06	52.14
PbrR3	56.82	56.06	100.00	57.36
PbrR4	52.11	52.14	57.36	100.00

Figure 4.1. Clustal W Multiple Alignment amino acid sequences of PbrR from different bacterial species: *Alcaligenes faecalis* (Q6WB35)(PbrR1), *Klebsiella pneumoniae* subsp(Q6U5Q1_KLEPN)(PbrR2)) (*Ralstonia Cupriavidus metallidurans* CH34(PbrR3), *Pseudomonas* putida KT2440(PbrR4). (A) Predicted amino acid alignment (B) Pairwise matrix alignment scores

Figure 4.1 shows that there is a significant number of residues that are highly conserved withing the PbrR family. The minimim sequence identity for the four sequences aligned was 52.1% between PbrR(Pp) (PbrR4) and the *Alcaligenes faecalis* PbrR (PbrR1). However in all cases there are three fully conserved cysteine residues (Cys 77, Cys 112 and Cys 119) (based on the PbrR(Pp) amino acid numbering sequence). Interestingly Cys 112 and Cys 119 area also fully conserved within the MerR family (Figure 4.2).

A.

```
CadR --MKIGELAKATDCAVETIRYYEREQLLPEPARSDGNYRLYTQAHVERLTFIRNCRTLDM 58
PbrR --MKIGELAKATDCAVETIRYYERENLLPEPARSEGNYRLYTQAHVERLTFIRNCRTLDM 58
MerR MKFRIGELADKCGVNKETIRYYERLGLIPEPERTEKGYRMYSQQTVDRLHFIKRMQELGF 60
CueR --MNISDVAKITGLTSKAIRFYEEKGLVTPPMRSENGYRTYTQQHLNELTLLRQARQVGF 58
ZntR -MYRIGELAKMAEVTPDTIRYYEKQQMMEHEVRTEGGFRLYTESDLQRLKFIRHARQLGF 59
        .*.::*.      .:**:**.  ::    *:: .:* *::  ::.* :::. : :.:

CadR TLDEIRSLLRLR-DSPDDSCGSVNALIDEHIEHVQARIDGLVALQEQLVELRRRC---NA 114
PbrR TLDEIRSLLRLR-DSPDDACGSVNALIDEHIEHVQARIDGLVALQEQLVELRRRC---NT 114
MerR TLNEIDKLLGVV-DRDEAKCRDMYDFTILKIEDIQRKIEDLKRIERMLMDLKERCPE-NK 118
CueR NLEESGELVNLFNDPQRHSADVKR-RTLEKVAEIERHIEELQSMRDQLLALANACPG--D 115
ZntR SLESIRELLSIRIDPEHHTCQESKGIVQERLQEVEARIAELQSMQRSLQRLNDACCGTAH 119
     .*:.  .*: :  *     .          :: .:: :*  *  :.  *  *   *

CadR QGAECAILQQLETNGAVSVPETEHSHVGRSHGH 147
PbrR QGSECAILQQLETNGAVSVPDTEHSHVGRSHGH 147
MerR DIYECPIIETLMKK------------------- 132
CueR DSADCPIIENLS-GCCHHRAG------------ 135
ZntR SSVYCSILEALEQGASGVKSGC----------- 141
     .   *.*:: *
```

The (*) (asterisk) indicates positions which have a single, fully conserved residue, (:) (Colon) indicates conservation between groups of strongly similar properties, and (.) (Period) indicates conservation between groups of weakly similar properties.

B.

	CadR	PbrR	MerR	CueR	ZntR
CadR	100.00	95.92	39.06	30.30	36.03
PbrR	95.92	100.00	39.84	29.55	36.03
MerR	39.06	39.84	100.00	33.86	32.06
CueR	30.30	29.55	33.86	100.00	37.04
ZntR	36.03	36.03	32.06	37.04	100.00

Figure 4.2. Diagram of the alignment amino acid sequences of MerR family group. (CueR in *E. coli* K12 (1Q05), ZntR 1Q08 in *E. Coli,* PbrR in *Pseudomonas putida* KT2440, Cad R in *Pseudomonas putida* 06909, MerR (Hg) MerR.P22853.

However in CueR the equivalent residue to Cys77 is an alanine. PbrR(Pp*)* shows very high sequence identity to CadR (95.9%) also indicating that it may bind (Cd, Zn and Hg (see section 1.4.5). Cys 77, 112 and 119 form the metal binding site in *Ralstonia* CH34[7] and thus are key residues to mutate to determine the effect on the metal binding properties of the PbrR protein family.

4.3 Construction of plasmid vectors

The PbrR gene was isolated from *Pseudomonas putida* KT2440 by PCR. In this construct, the PbrR gene was cloned between two restriction sites *EcoRI* and *HindIII* using two primers, F-PbrR691 and R-PbrR691, and vector pET28a(+) (using the experimental methods described in Chapter 2, section 2.2.3; Figure 2.2). To include the His-tag for purification purposes, the reverse primer was designed to exclude the stop codon (TAA) of the PbrR gene while the forward primer F-PbrR691 was designed with an *EcoRI* restriction site and R-PbrR691 with a *HindIII* restriction site (see Table 2.1; chapter 2). To visualise the PCR product, both replicates were run on a 1% (w/v) agarose gel, as described in section 2.2.6. Fragments of ~444 bp were detected on the gel under a UV illuminator (Figure 4.3).

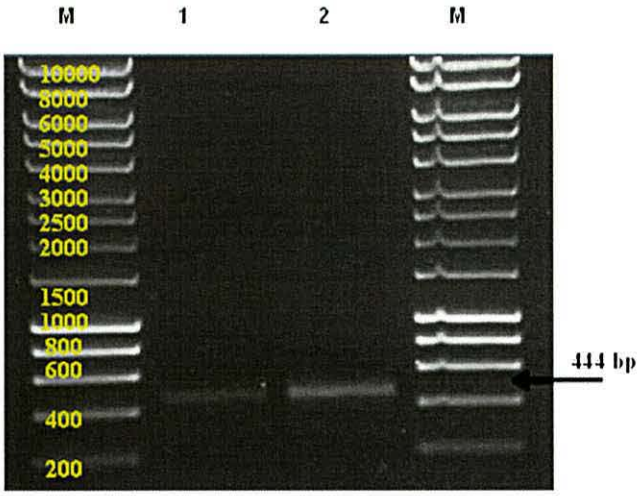


Figure 4.3. Agarose gel electrophoresis (1% (w/v)) of a PCR amplified fragment of the PbrR gene. The DNA molecular marker is shown in lane M (lanes 1 & 2).

The band of interest on the gel was purified using a QIAquick PCR purification kit (Qiagen, UK), as described in detail in section 2.2.6. In order to clone the PbrR gene fragments obtained from the PCR into the expression vector *pET28a(+),* complementary ends of the gene and the plasmid needed to be produced. To do so, the purified PCR product and the vector were digested separately using restriction enzymes *EcoRI* and *HindIII*, as described in section 2.2.7. The complementary ends on the gene fragments and on the plasmid vector were produced and the fragments and vector were mixed together and further purified (see section 2.2.8). The complementary ends were joined using a T4 ligase enzyme with an appropriate buffer and incubated overnight at 16°C (see section 2.2.9).

The ligation product was digested with a *SacI* restriction enzyme at 37°C for an hour then transformed into DH5α competent cells by heat shock (see section 2.3.2). Aliquots of different volumes were plated onto the LB agar Petri dishes containing kanamycin and incubated overnight at 37°C. Since the kanamycin resistance gene was integrated as

part of the *pET28a (+)*, only transformed bacteria carrying the recombinant plasmid grew as it had kanamycin resistance (see section 2.3.3). For plasmid preparation, the day after, single colonies were transferred from LB-agar plates into 5 different tubes containing LB medium and the kanamycin at the same concentration and incubated overnight at 37°C. Cells were then harvested and plasmids were extracted (see section 2.3.4). The plasmids were double digested with *EcoRI*, and *HindIII*. The digested product was then run on a 1% (w/v) agarose gel electrophoresis to confirm the insertion of the correct gene into the plasmid (see section 2.2.6 in chapter 2) and (Figure 4.4).

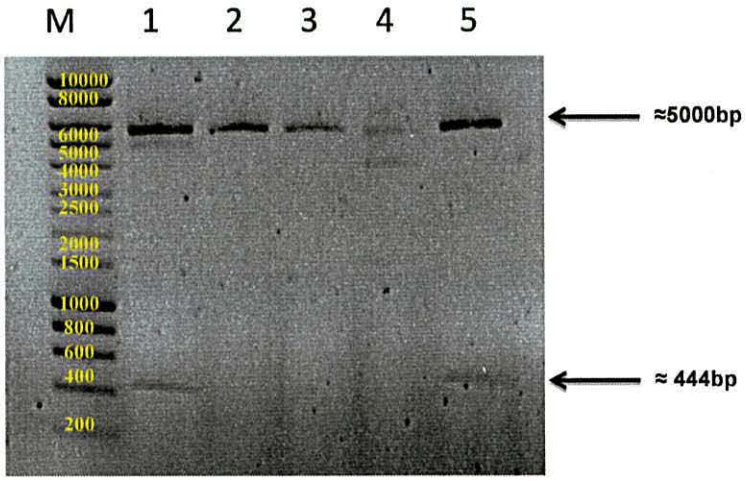


Figure 4.4. Agarose gel electrophoresis (1% (w/v)) of the double digested recombinant plasmid PbrR-*PET28a(+)* (lanes 1-5 correspond to clones 1-5 respectively).

Plasmids of clones 1 and 5 indicated two bands, one at ~444 bp corresponding to the PbrR gene and another at ~5000 bp corresponding to the *pET28a (+)* plasmid. These results demonstrated the successful cloning of the vector in clones 1 and 5. These two clones were sent for sequencing to Eurofins MWG operon, UK (see section 7.2.1, 7.2.2). The sequencing result confirmed gene cloning into *pET28a (+)* (see figure 4.5). The gene may not have been cloned into the vector in the other samples, or the bands at ~444 bp were too faint to be detected under the UV illuminator.

Figure 4.5. DNA sequence of produced after ligation the PbrR gene with *PET28a (+)* plasmid.

4.4 PbrR(Pp) over-production

Competent cells of the high level expression *E. coli* strain BL21 (DE3) were heat shock transformed (see section 2.3.2) with the *pET28a (+)*-PbrR plasmid which also carried a His-tag. Since the plasmid *pET28a (+)* is kanamycin (Km) resistant, all growth media contained kanamycin.

4.5 PbrR(Pp) purification

Recombinant PbrR(Pp) and all other mutants were purified using the method described in section 2.5. After harvesting cells from the growth medium, *via* centrifugation the crude extract was lysed in binding buffer and sonicated (section 2.5.1.1). The cell debris was removed by centrifugation and the supernatant was purified on a His-tag column, and imidazole salt was used to elute the protein. A PD-10 desalting column was then used to remove all the salts from the protein.

An overproduction expression protein experiment began by inoculating two flasks containing LB media and kanamycin with an overnight culture. The first flask was used

as a control (no IPTG added) and the second flask was grown until an $O.D_{600}$ reached 0.6 at 37°C. It was then induced for 4 hours with IPTG. The recombinant PbrR overproduction was visualised after centrifuging the culture, re-suspending, sonicating, centrifuging again and purifying the protein on a His-tag column and eluting with increasing concentrations of imidazole (50, 100, 150, 200, 300, 500) mM. The resulting fractions were then run on an SDS-PAGE gel (Figure 4.6). Although it was clear that some PbrR did start to elute with 300 mM imidazole, the intensity of the band observed by eluting with 500 mM imidazole was the strongest compared to the other imidazole concentrations. Therefore the amount of protein eluted would be greatest using 500 mM imidazole concentration in subsequent experiments.

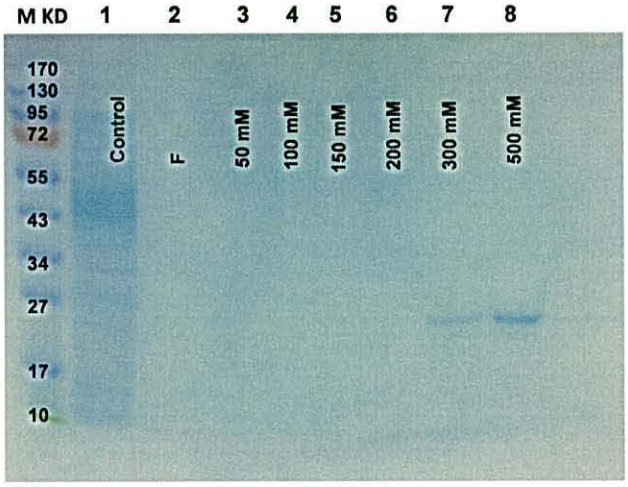


Figure 4.6. SDS-PAGE (12%) of the His-tag column purified PbrR gene product. from *PET28a (+)*. Lane (M): marker, lane 1: uninduced, lane 2: flow through fraction, lanes 3-8: flow through samples with increasing imidazole concentrations (50-500mM).

Then next sets of experiments were carried out by preparing three flasks. The first was un-induced and the control, the second was an untransformed *Rosetta*, and the third was grown until an OD_{600} reached 0.6 and then induced for 4 hours. The culture was

centrifuged, re-suspended, sonicated, centrifuged again, and the protein was purified with a His-tag column and eluted with 500mM imidazole. The re-suspended protein was then run on a 12% (w/v) SDS-PAGE gel (Figure 4.7). The uninduced and an transformated extracts exhibited multiple bands over arrange of molecular weights, but no strong bands at ~26 kDa. The post His-tage colomn protein samples showed one major band at the expected molecular weight for the PbrR subunit (~26 kDa).

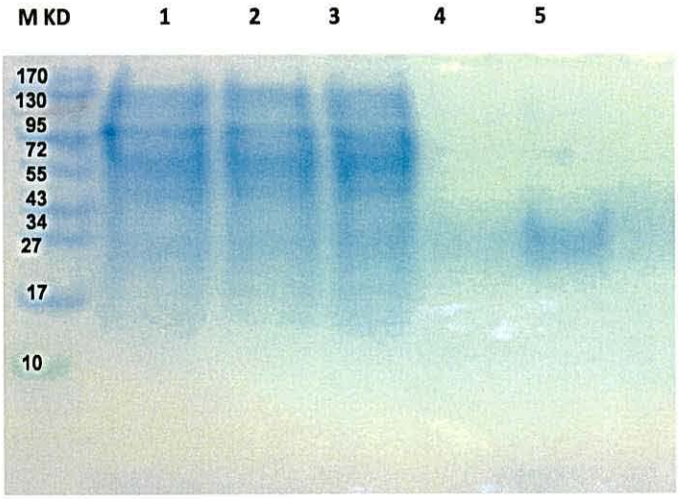


Figure 4.7. SDS-PAGE (12%) of PbrR(Pp) Protein purification test. Lane M: marker, lane 1: uninduced, lane 2: untransformed Rosetta, lane 3: IPTG induced cell free extract (unpurified), lane 4: flow through fraction, lane 5: flow through sample eluted of the His-tag column with 500 mM imidazole.

A large scale experiment was conducted as follows: four flasks containing LB with kanamycin were inoculated with an overnight culture. The first two flasks were grown at 37°C until the OD_{600} reached 0.6. One flask was then induced for four hours and the other was induced overnight at room temperature. The second two flasks were grown at 37°C until the OD_{600} reached 0.4. One was then induced for four hours and the other was induced overnight at room temperature. The cells were re-suspended, sonicated, centrifuged again and the protein was purified by a His-tag column and eluted with 500

mM imidazole. The re-suspended protein was then run on a 12% (w/v) SDS-PAGE gel (Figure 4.8).

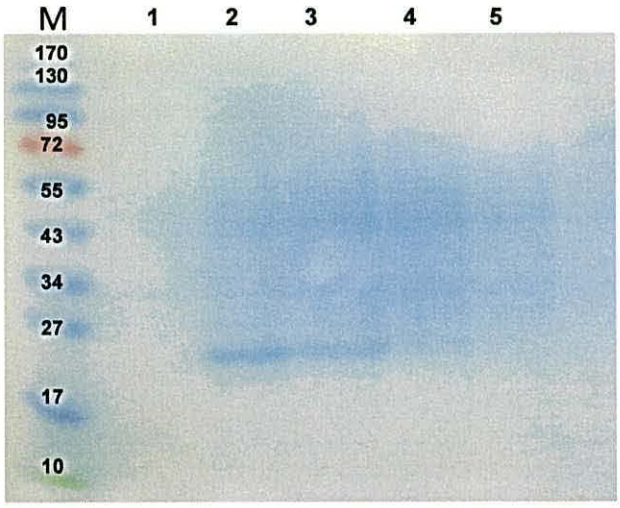


Figure 4.8. The SDS-PAGE (12%) of PbrR(Pp) protein purified using His-tag column eluted with 500 mM imidazole salt; Lane 1: flow through lane 2: bacteria were grown at 37°C until the $O.D_{600}$ reached 0.6 and induced overnight at room temperature; lane 3: grown at 37°C until the $O.D_{600}$ reached 0.6 and induced for 4 hours at 37°C; lanes 4 and 5: grown until the $O.D_{600}$ reached 0.4 at 37°C of bacteria in lane 4 were induced for four hours with 1 mM IPTG; while bacteria in lane 5 were induced overnight at room temperature.

Although the molecular weight of PbrR is 52 kDa as a dimer, the loading buffer of SDS-PAGE denatures the protein and prevents electrostatic interactions between the two monomers. A band was seen at ~26 kDa in lane 2 (Figure 4.8) and was interpreted to be the recombinant PbrR.

Table: 4.2. Optimisation of conditions for the production of recombinant PbrR.

Samples	Induced time	Conc. (µg/ml)	Conc. (µM)	Amount (µg)
37 °C (IPTG 1 mM) O.D$_{600}$ ≈ 0.6	4h	39	750	273
	Overnight≈16h	62	1192	434
37 °C (IPTG 1 mM) O.D$_{600}$ ≈ 0.4	4h	29.8	537	208
	Overnight≈16h	_	_	_

Conc.: concentration

The highest yield was in the LB medium when the cells were grown at 37°C, induced and had continuous agitation overnight at 20°C.

Table 4.2 shows that recombinant PbrR was not expressed in LB medium and at different conditions in bacteria grown to O.D$_{600\ nm}$ ≈ 0.4 with agitation overnight at 20°C, while the amount of PbrR expression increased with agitation for 4h at 37°C. The highest yield of PbrR recombinant protein was obtained at O.D$_{600}$ ≈ 0.6 with agitation overnight at room temperature (≈20 °C).

4.6. Generation and overproduction of mutants

4.6.1 Introduction

PbrR protein contains three amino acid side chains that form the active site consisting of the cysteine thiol groups: C77, C112 and C119. These sulphur donors ligate with lead, appearing as one band at 332 nm in the UV/Visible spectrum associated with a Pb-S charge transfer band. To better understand the role of each of the cysteines in forming the highly specific and efficient lead binding site the mutation of those residues was undertaken. Cysteine was mutated to alanine, aspartic acid, in mono mutants (C77A,

C77D), double mutants (which used the mono mutant as template) (C112A, C112D), and a triple mutant (which used the mono mutant as a template) (C112/119A, C112/119D). The effect of the mutations was monitored using UV/Vis spectroscopy to examine any changes to the charge transfer band, atomic absorption and UV/Vis spectroscopy measurements to determine the metal to protein stoichiometry and ^{207}Pb NMR spectroscopy to examine the ligand environment surrounding the lead ion.

4.6.2 Results and Discussion

The mutant proteins of cysteine (C77, C112, and C119) to alanine and aspartic acid were generated successfully and their correct sequences were confirmed by sequencing (see section 2.2.10). Agarose gel electrophoresis of the PCR amplified genes of the C77A-PbrR and C77D-PbrR single mutants showed clear strong bands at 444 bp after using PbrR(Pp) gene as template (figure 4.9 A). This corresponds to the expected weight of the DNA templates for those mutants.

To ligate the single mutant genes into pET28a(+) the PCR products for each mutant were purified, digested with restriction enzymes (*EcoRI* and *HindIII*), and then ligated using T4 ligase enzyme (see section 2.2.9). The overnight product was digested with restriction enzyme and then run on a 1% agarose gel (figure 4.9 B).

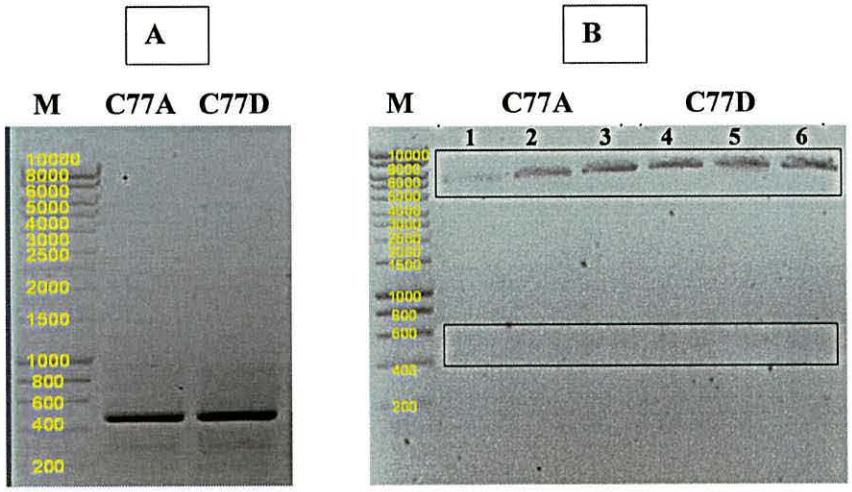


Figure 4.9. Agarose gel electrophoresis (1% (w/v)) of (A) PCR amplification products and (B) double digested ligation products (in *PET28a(+)*) of C77A-PbrR (lanes 1-3) and C77D-PbrR (lanes 4-6) genes (M= DNA molecular marker).

These results showed successful digestion as indicated by two bands, one at 444 bp corresponding to the expected molecular weight of the C77A-PbrR and C77D-PbrR genes, and one at 5000 bp corresponding to the digested vector (*pET28a (+)*). These clones were sent for sequencing to Eurofins MWG operon, UK. The sequences were confirmed as being correct for the above clones (see appendices section 7.2.3.1.1, 7.2.3.1.2 for C77A-PbrR and 7.2.3.2.1, 7.2.3.2.2 for C77D-PbrR).

The amplification of double mutation gene C77A/C112A-PbrR and C77D/C112D-PbrR was successful as judged by analysis of DNA sequencing results (see appendices 7.2.3.1.3, 7.2.3.1.4 for C77A/C112A-PbrR (+) and 7.2.3.2.3, 7.2.3.2.4 for C77D/C112D-PbrR). The PCR amplified products were visualised using agarose gel electrophoresis (figure 4.10 A). Strong clear bands were visible at 444 bp, which corresponds to the expected weight of the DNA templates for those mutants.

Plasmids of clones C77A/C112A-PbrR-pET28a(+) and C77D/C112D-PbrR-pET28a(+) were double digested and run on an agarose gel (see figure 4.9 B). For each mutant two

bands are visible, one at ~444 bp corresponding to the C77A/C112A-PbrR gene and C77D/C112D-PbrR another bands at ~5000 bp corresponding to the pET28a (+) plasmid. These results demonstrated the successful cloning of the vector in clones (see figure 4.10 B). These clones were sent for sequencing to Eurofins MWG operon, UK. The sequences were confirmed as being correct for the above clones (see appendices section 7.2.3.1.3, 7.2.3.1.4 for C77A/C112A-PbrR and 7.2.3.2.3, 7.2.3.2.4 for C77D/C112D-PbrR).

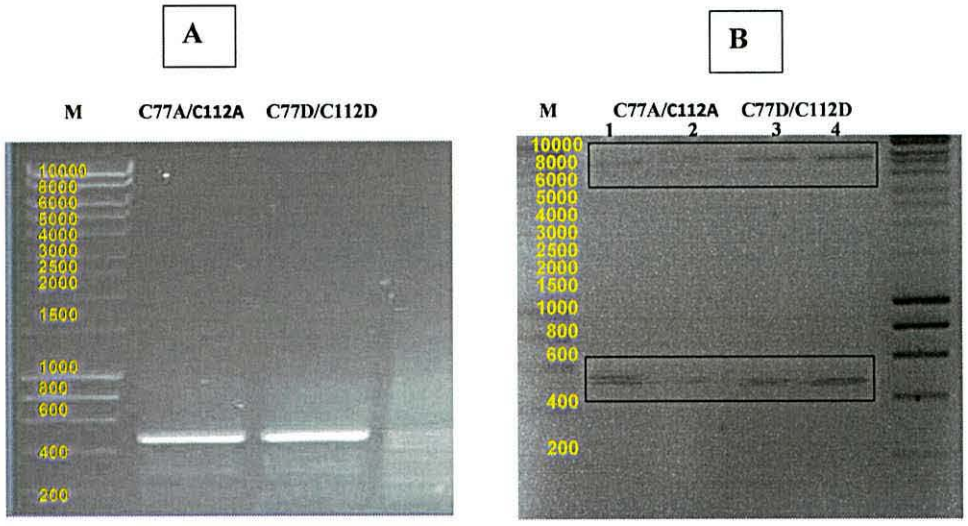


Figure 4.10. Agarose gel electrophoresis (1% (w/v)) of (A) PCR amplification products and (B) double digested ligation products (in *PET28a (+)*) of C77A/C112A-PbrR (lanes 1-2) and C77D/C112D-PbrR (lanes 3-4) genes (M= DNA molecular marker).

As for the previous single and double gene mutants, the triple mutants C77A/C112A/C119A-PbrR and C77D/C112D/C119D-PbrR were amplified by using PCR and visualised in a 1% agarose gel (see figure 4.11A). The gel showed a band at 444 bp for each mutant. The PCR product was ligated into pET28a(+) plasmid and the double digestion products produced two bands each when run on 1% agarose gel (figure

4.11 (B)). The observed bands at 444 bp (quite faint) corresponded to the DNA template of each mutant and the other bands at 5000 corresponding to the pET28a(+)..

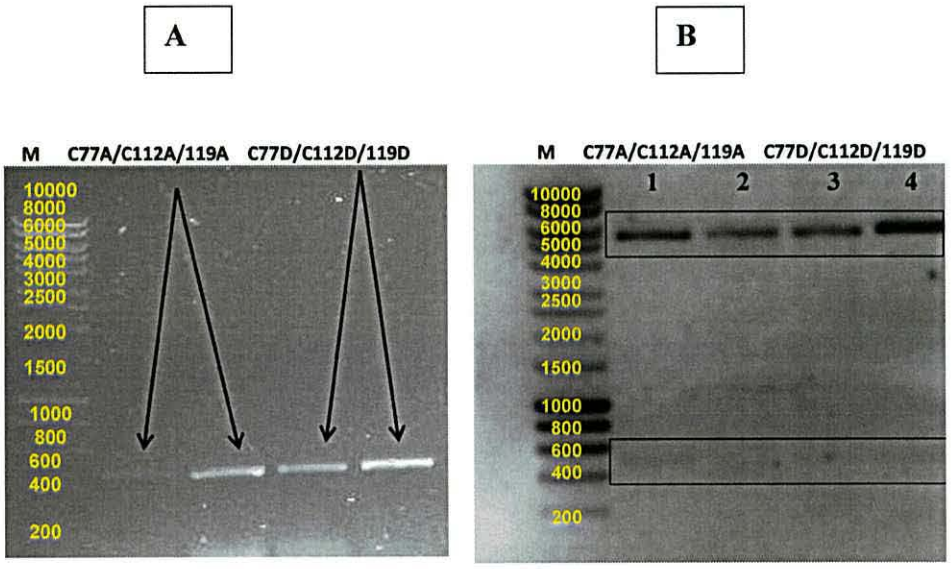


Figure 4.11. Agarose gel electrophoresis (1% (w/v)) of (A) PCR amplification products and (B) double digested ligation products (in *PET28a (+)*) of C77A/C112A/C119A-PbrR (lanes 1-2) and C77D/C112D/C119D-PbrR (lanes 3-4) genes (M= DNA molecular marker).

These clones were sent for sequencing to Eurofins MWG operon, UK. The sequences were confirmed as being correct for the above clones (see appendix section 7.2.3.1.5, 7.2.3.1.6 for C77A/C112A/C119A-PbrR-PET28a (+) and 7.2.3.2.5, 7.2.3.2.6 for C77D/C112D/C119D-PbrR-PET28a (+)).

Mutant PbrR protein samples were purified using the same method as was used for recombinant protein (see section 2.5). Harvested cells from the growth medium centrifuged and sonicated. The cell debris was removed by centrifugation and the supernatant was purified on a His-tag column, and imidazole salt was used to elute the protein. A PD- 10 desalting column was then used to remove all the salts from the mutant proteins. The molecular weight of the purified mutant proteins subunits were

determined by using SDS-PAGE method (reducing conditions gives the subunit molecular weight rather than the intact protein molecular weight).

The molecular weight of the strongest bands on the SDS-PAGE gels for all the mutant proteins, C77A, C77A/C112A, C77A/C112A/C119A) and (C77D, C77D/C112D, C77/C112D/C119D), were ≈26 kDa (see figure 4.12). This corresponded to the expected molecular weight of the subunits of PbrR. The additional bands visible post his-tag purification may have resulted from aggregation of protein at high concentrations.

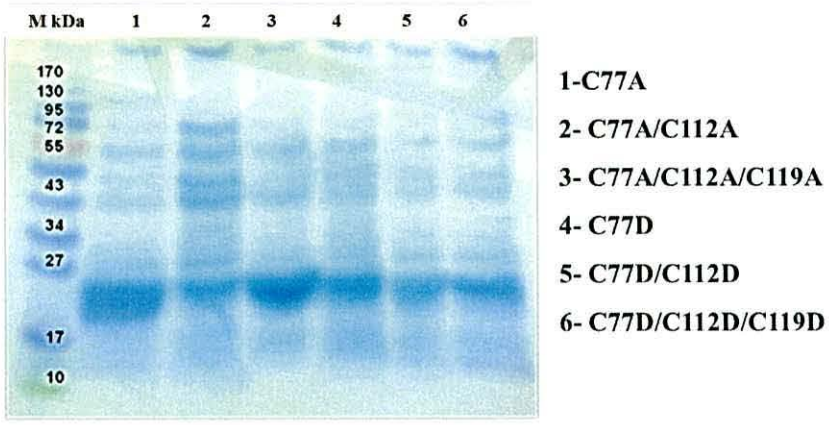


Figure 4.12. 12% SDS-PAGE of PbrR(Pp) single, double and triple mutant proteins overproduced in the *E. coli* BL21 (DE3) strain. M: molecular weight markers.

4.7 Conclusions

Pseudomonas putida KT2440 had 57% similarity with *Ralstonia (Cupriavidus) metallidurans* when the BLAST program was used to compare predicted PbrR amino acid sequences. Forward and reverse primers were designed and amplified the PbrR gene using the PCR technique.

Successful production of recombinant PbrR protein and mutants was achieved using the *pET28a (+)* plasmid as a vector. In this plasmid, the PbrR gene was cloned between the two restriction sites *EcoRI* and *HindIII* of the pET28a (+) vector to produce the PbrR-PET28a (+) plasmid containing an additional His-tag N-terminus sequence. The production of large amounts of the native and mutant PbrR protein was highly dependent on the induction time. Using LB medium and increasing the induction time to overnight at room temperature yielded the highest production of PbrR protein. The protein was purified using a His-tag column and imidazole salts were removed with a PD-10 desalting column. The molecular weight of the purified recombinant and mutated proteins subunits were checked by SDS-PAGE to confirm the approximate size at 26 kDa. Other groups studies on related systems show PbrR protein family members to be dimers, so the expected molecular weight of the PbrRp protein is approximately 52 kDa. This could be confirmed on the whole intact protein using gel filtration chromatography or native PAGE in future studies.

4.6 References

1. J. Lodge, P. Lund and S. Minchin, *Gene Cloning Principles and Applications*, Taylor and Francis group, 2007.

2. A. Doyle, *High Throughput Protein Expression and Purification: Methods and Protocols*, Totowa, N.J. Humana Press. (6^{th} Ed), 2009.

3. L, Hartley, Cloning technologies for protein expression and purification, *Current Opinion in Biotechnology*, 2006, **17**, 359-366.

4. F. Freuler, T. Stettler, M. Meyerhofer, L. Leder and L. Mayr, Development of a novel Gateway-based vector system for efficient, multiparallel protein expression in *Escherichia coli*, *Protein Expression and Purification*, 2008, **59**, 232-241.

5. F. Altschul, W. Gish, W. Miller, E. Myers and D. Lipman, Basic local alignment search tool, *Journal of Molecular Biology*, 1990, **215**, 403-410.

6. D. Feng and R. Doolittle, Progressive sequence alignment as a prerequisite to correct phylogenetic trees, *Journal of Molecular Evolution*, 1987, **25**, 351-360.

7. P. Chen, E. Wasinger, J. Zhao, D. Lelie, L. Chen and C. He, Spectroscopic insights into lead (II) coordination by the selective lead (II)-binding protein PbrR691, *Journal of the American Chemical Society*, 2007, **129**, 12350–12351

CHAPTER 5: ELECTRONIC SPECTROSCOPY AND NMR STUDIES OF RECOMBINANT PbrR (Pp) AND C77, C112 AND C119 MUTANTS.

5.1. Introduction

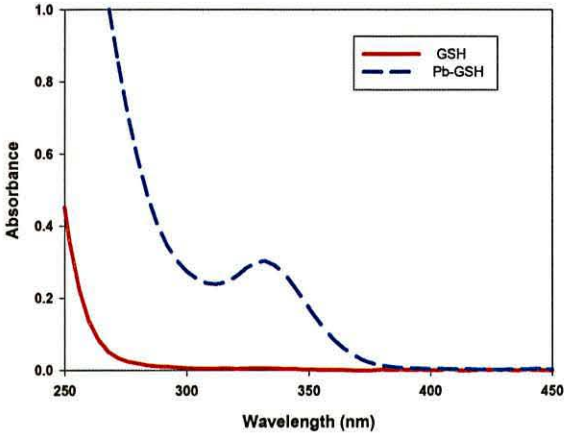
This chapter focuses on the characterisation of the PbrR native and mutant proteins using electronic spectroscopy and ^{207}Pb NMR. In addition, GSH-lead complexes are studied as a model system for characterisation. Previous electronic absorption spectroscopy literature[1] on the properties of $[Pb(GS)_3]^-$ at different pH values suggested that the absence of an absorption band at 334 nm is due to the incomplete deprotonation of cysteine thiolates and the consequential inability to form Pb-S(cys) bonds. Therefore, in order to select conditions suitable for preparing a GSH-Pb model complex for characterisation *via* ^{207}Pb NMR and UV/Vis spectroscopy, a more alkaline pH of pH 8.5 was selected (see section 2.7.3.1).

The ^{207}Pb NMR spectroscopy of Pb-GSH was used as a model to compare with the ^{207}Pb NMR of Pb-PbrR(Pp) protein. The coordination environment around the Pb^{2+} ion in solution of the Pb-PbrR(Pp) was evaluated using Pb^{2+} NMR spectroscopy. The Pb^{2+} nucleus provide excellent sensitivity to changes in the metal coordination environment although that nuclei is only 22% abundant.

5.2. Results and Discussion

5.2.1 The UV/Vis and ^{207}Pb NMR spectral properties of GSH and Pb-GSH

The UV/Vis spectrum of Pb-GSH at pH 8.5 clearly shows that the coordination of all cysteine thiolates to Pb (II) is complete (Figure 5.1). The LMCT band centred at 334 nm becomes more prominent and the observed extinction coefficient (ε_{334} = 3560 $M^{-1}cm^{-1}$) is very close to that obtained for PbS_3 in homoleptic three stranded coiled coil peptides. [2]



GSH (20 μM)
Pb-GSH in (100 mM) Tris-HCl (pH 8.5)

Figure: 5.1. UV/Vis spectra of GSH and Pb-GSH solutions. at room temperature

Figure 5.2 shows the ^{207}Pb NMR spectra of the Pb-GSH complex (Pb: GSH 1:3 ratio) at pH 8.5. The observed ^{207}Pb signal is at 2764.69 ppm. At high pH (pH 8.5), all cysteine thiolates are deprotonated and able to bind Pb^{2+}, so a sharp signal appears with a relatively narrow band width.[2] This suggests that the major species present in the solution is $[Pb(GS)_3]^-$ where Pb^{2+} is coordinated in a trigonal pyramidal geometry (hemidirected) with a lone pair occupying the apical position.

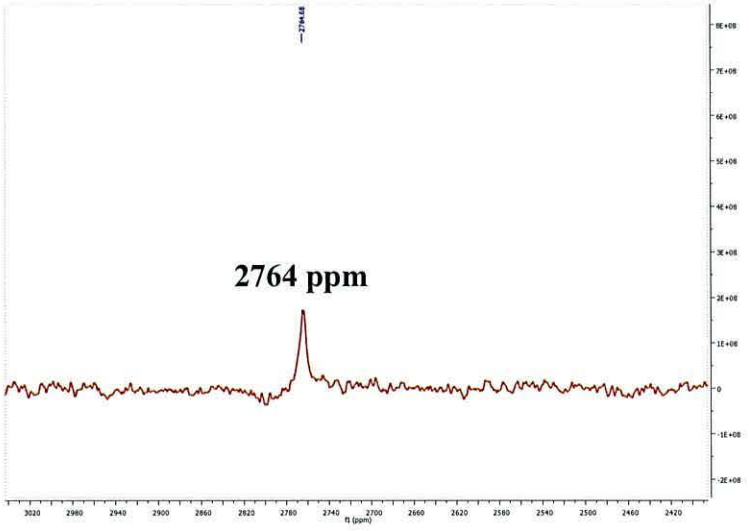


Figure: 5.2. NMR spectrum of an aqueous solution Pb-GSH complex containing 10% D_2O.

The ^{207}Pb NMR yields narrow signals over an extremely wide chemical shift range. A number of literature studies have examined lead complexes with a variety of

coordination spheres ranging from Pb-S_3 to Pb-N_4 and have also used either UV/Vis or Pb NMR to characterise those complexes (Table 5.1).

Table 5.1. ^{207}Pb NMR Data for selected lead compounds.

Compound	Coordination	Chemical shift (ppm)	UV/Vis (nm)	References
GSH	PbS_3	2793	335	1
	PbS_3	5750 (2790*)	334	2
	PbS_3	2764	334	This thesis
$[(PhS)_3Pb](Ph_4As)$	Pb-$(ArS)_3$	5828(2868*)	-	3
$[PATH\text{-}Pb]^{+(a)}$	Pb-N_2(RS)	5318(2358*)	-	4
Bis(thiohydroxamato)lead	Pb-S_2O_2	4100-4500	-	5, 6
$[H_2B(PZ)_2]Pb$	Pb-N_4	2821	-	7
$[Pb(EDTA)]^{2-}$	PbO_4N_2	2441	-	8
$Pb(EDTAN_2)^{(b)}$	PbO_3N_3	2189	-	
$[Pb(EDTA\text{-}N_4)]^{2+(c)}$	PbO_3N_3	1764	-	

*Relative to Pb $(NO_3)_2$ defined as reference at -2960 ppm (a) 2-methyl-1-[methyl(2-pyridin-2-ylethyl)amino]propane-2-thiolatolead (b) Pb(II) Ethylenediamine-N,N'-bis(acetic acid)-N,N'-bis(acetamide), (c) Pb(II) Ethylenediamine tetra acetayl lead.

In liquid state NMR, the chemical shift of $Pb(NO_3)_2$ is often used as an NMR reference compound. Some inconsistencies appear in the literature with respect to how the chemical shifts are reported relative to the $Pb(NO_3)_2$ standard solution.[2-5] In some studies, the chemical shifts have been cited by using the $Pb(NO_3)_2$ reference as defined as 0 ppm[1] and others have defined it as -2960 ppm (this thesis -2955 ppm). Table 5.1 shows the chemicals shifts as reported in the literature, but where appropriate the adjusted chemical shifts for $Pb(NO_3)_2$, defined as a reference at -2960 ppm, have been calculated and are asterisked and shown within the table in brackets. Unlike ^{1}H-NMR spectroscopy, ^{207}Pb resonances become more de-shielded and shift downfield as the electronegativity of the donor atom decreases (downfield shift/desheildng: O < N < S).[9] Furthermore, the presence of a stereochemically active

lone pair shifts the lead NMR resonance downfield.[10] A rough correlation exists between the coordination number of lead and the ^{207}Pb chemical shift.[11] Unfortunately, not many Pb-X coupling constants have been reported so far. Coupling constants provide a measure of bond strength and steric configuration.[9]

^{207}Pb NMR of $[PATH\text{-}Pb]^+$ exhibits a Pb resonance at 2358 ppm* (Table 5.1). This is further downfield than previously observed either for Pb complexes, in which Pb is bound by N atoms only [7] or for bis- (thiohydroxamato)lead complexes, which contain both S and O in the coordination sphere [4,6] The Pb-alkanethiolate bond is similar to that of the $[(PhS)_3Pb]$ anion, which contains three arylthiolates bound to Pb.[3] These data suggest that the ^{207}Pb NMR resonance for Pb in Zn sites in proteins (which typically contain more than one cysteine residue and, hence, more than one alkanethiolate ligand) will likely be even further downfield ca. 2700-3000 ppm versus $Pb(NO_3)_2$.

The trend observed for the ^{207}Pb chemical shifts can potentially be used to distinguish between different binding modes.[8,12] However, the trend that is observed for the chemical shifts also was shown to reflect the change going from negatively charged to neutral and finally positively charged species. As the number of nitrogen atoms in the ligand (not necessarily directly coordinating the Pb ion) is increased in $[Pb(EDTA)]^{2-}$ to $Pb(EDTAN_2)$ to $[Pb(EDTA\text{-}N4)]^{2+}$, the charge on the metal complex becomes more positive and the ^{207}Pb resonance shifts upfield. In the Pb 1st coordination sphere of these complexes, the ratio of nitrogen:oxygen atoms was 2:4 for $[Pb(EDTA)]^{2-}$ and 3:3 for both Pb(EDTA-N2) and $[Pb(EDTA\text{-}N4)]^{2+}$.

5.3 PbrR(Pp) protein

The method for reconstituting the apo-PbrR(Pp) protein with lead has been described in chapter 2 (materials and experimental). The apo-PbrR(Pp) was dialysed against a lead-containing buffer overnight, with the dialysis solution being exchanged for fresh solution after 12 h.Weakly bound lead was removed by using a lead-free buffer as a final dialysis step and this buffer was exchanged three times. This method was used

several times to prepare duplicate batches of lead-loaded protein for characterisation using UV/Vis spectroscopy as shown in section 2.7.2 and Table 5.2 and ^{207}Pb NMR in section 2.7.3. The protein concentrations were determined using the BioRad assay and the metal concentrations using atomic absorption spectroscopy. For the apo-PbrR(Pp) protein the ratio of Pb^{2+}: PbrR(Pp) was 1:1 thus, each PbrR(Pp) dimer binds one Pb^{2+} ion. This is as expected on the basis of studies of related PbrR family members (see section 1.2.2)

Table 5.2. The stoichiometry of PbrR(Pp) protein to lead metal

Protein	No Moles* Protein (moles)	No Moles* Pb^{2+} (moles)	λmax (nm)	Ratio Pb:PbrR(Pp)
Apo-PbrR(Pp)	2×10^{-4}	3.5×10^{-7}	-	0.002
Pb-treated PbrR*	4.4×10^{-5}	4.7×10^{-5}	332	1.067

*After the lead loading procedure shown in section 5.2

* The concentration of protein was measured by the BioRad assay and the concentration of Pb^{2+} was measured by atomic absorption spectroscopy.

5.3.1. The UV/Vis and ^{207}Pb NMR spectral properties of apo-PbrR(Pp) and Pb-PbrR(Pp) recombinant proteins

Spectroscopic studies were performed on the lead (II)-loaded PbrR(Pp) and the apo-protein. Figure 5.3 shows the UV/Vis spectrum of apo-PbrR(Pp); no absorbance band appears in the region 300-600 nm , but following the procedure described in section 2.7.2, a Pb loaded sample (Pb-PbrR(Pp)) was prepared which exhibited a strong absorption band at 332 nm. This was assigned to the combination of both S 3p

to Pb^{2+} 6p ligand-to- metal charge-transfer (LMCT) and intra-atomic Pb^{2+} $6s^2$ to Pb^{2+} 6p transitions [13] (see Figure 5.3).

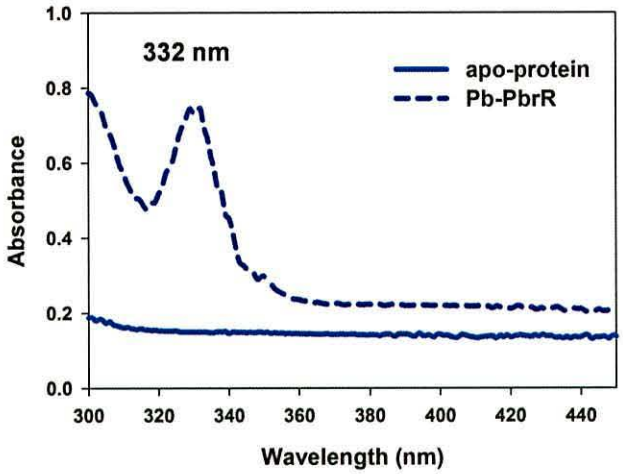


Figure: 5.3. UV/Vis spectra of the independently prepared samples of the 0.2 mM PbrR(Pp) and Pb^{2+}-PbrR(Pp) complex.

The extinction coefficient of the major absorption band at 332 nm was calculated to be 3720 M^{-1} cm^{-1} (Table 5.3).

Table 5.3. Calculation of the extinction coefficient from the Pb-PbrR(Pp) complex protein

Protein	**Absorption maximum (332 nm)**	**Protein concentration (μM)**	**Extinction coefficient (M^{-1} cm^{-1}) ***
Pb-PbrR(Pp)	0.744	200	3720

*Calculated using the Beer-Lambert Law

This was expected from related studies in the literature on PbrR691[14] and from studies on glutathione-lead complexes.[1] On the basis of these literature studies, the absorption at 332 nm in Figure 5.3 was assigned as a thiolate-lead (II) charge transfer band. This assignment was confirmed and the coordination environment of the Pb^{2+}

centre in PbrR(Pp) was probed by performing NMR studies of the Pb-PbrR(Pp) complex.[3]

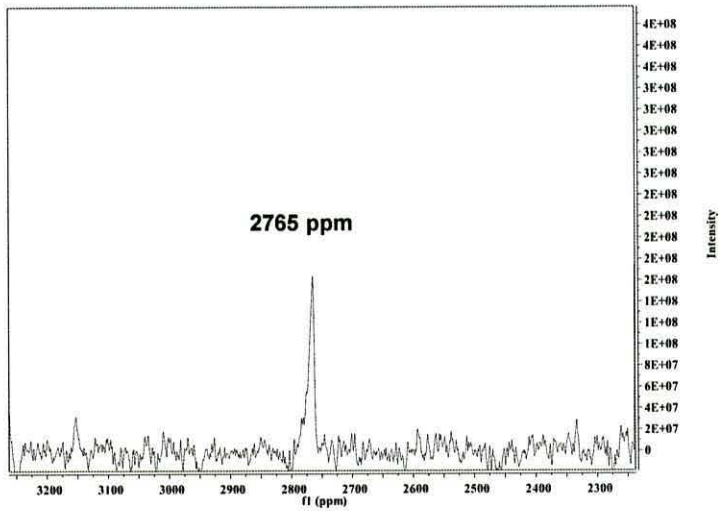


Figure: 5.4. NMR spectrum of an aqueous solution (5mM) Pb-PbrR(Pp) protein complex containing 10% D_2O.

However, information is still lacking, in particular for lead (II) thiolate complexes, to make the ^{207}Pb chemical shift scale reliably useful in a similar way to Pb-GSH because the proteins containing cysteine are major biological targets for heavy metals. The ^{207}Pb chemical shifts are reported for Pb^{2+} sulphur coordinated complexes. A Pb NMR signal at 2765 ppm was observed for the [Pb-PbrR(Pp)] complex at 298 K.

When coordinated PbS_3 complexes predominate, a ^{207}Pb NMR chemical shift of 2765 ppm, and a LMCT band at 332 nm occurs (see table 5.1). The Pb-PbrR(Pp) complex probably has trigonal-pyramidal PbS3 geometry because the coordination sphere is hemidirected with a stereochemically active inert electron pair numbers. A preference for PbS_3 coordination environments was previously observed for Pb^{2+} bound to large proteins.

The similarity of the wavelength maximum and extinction coefficient of the Pb-PbrR(Pp) to the those of the Pb-GSH complexes prepared herein and reported in the literature suggests that the first coordination spheres ligands and geometries are alike.

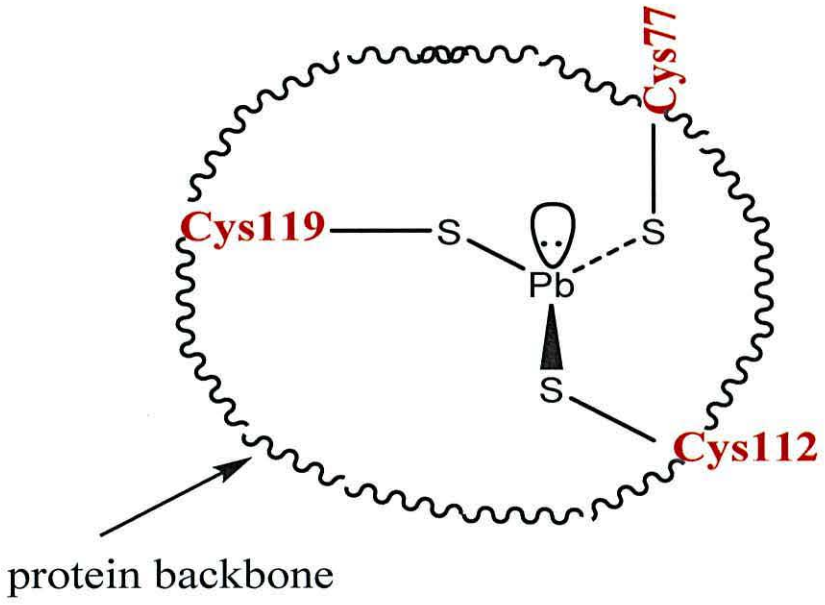


Figure: 5.5. Proposed structure of Pb(II)-bound thiol group in cysteine (Pb-PbrR(Pp)).

5.4. The mutation of cysteine 77, 112 and 119 to (alanine or aspartic acid)

The effect of mutating the amino acid residues responsible for Pb^{2+} binding was determined by performing site-directed mutagenesis to alter each of the three cysteine residues in the *Pseudomonas putida* KT2440 PbrR(Pp) protein to alanine, or aspartic acid both individually and as double and triple mutants (Table 5.4). The overall aim was to study the metal binding and spectroscopic properties of the perturbed active site, as mutation was expected to disrupt the binding of Pb^{2+} to the PbrR protein.

Table 5.4. The type of PbrR mutation and amino acid residuse number.

Mutation	Position
Mono mutants	C77A, C77D,
Double mutants	C77A/C112A, C77D/C112D,
Triple mutants	C77A/C112A/C119A, C77D/C112D/C119D

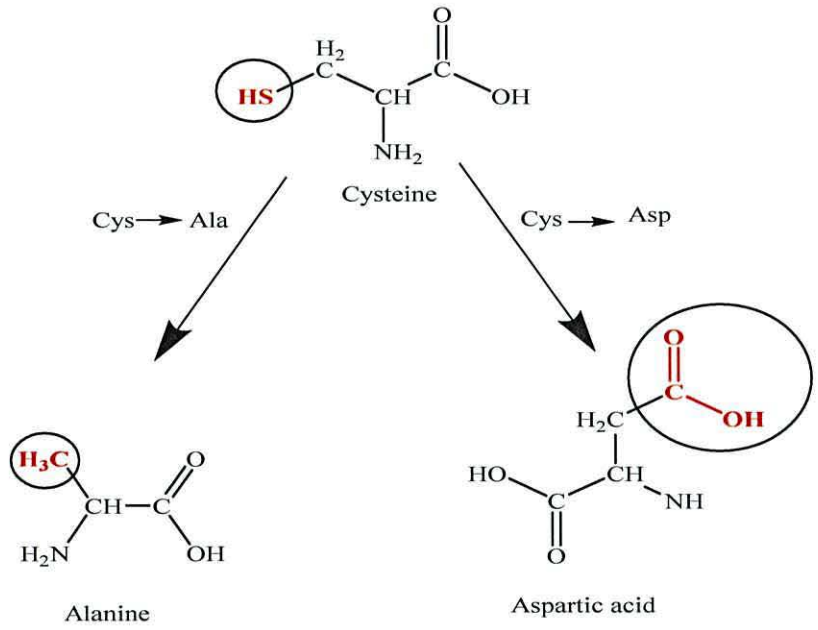


Figure: 5.6. The structures of cysteine, alanine, and aspartic acid.

5.4.1 The spectral properties of apo-PbrR(Pp) and Pb-PbrR(Pp) mutants

Tables 5.5 and 5.6 show the concentrations of apo-protein and 'holo' protein after the lead reconstitution procedure had been attempted, as measured by the Bio-Rad protein concentration assay. The metal concentrations were determined by atomic absorption spectroscopy.[14]

Table 5.5. The λ_{max} and the ratio of Pb: apo-PbrR(Pp) in the as isolated mutant proteins

Protein	No Moles* Protein (moles)	No Moles* Pb^{2+}(moles)	λ_{max} (nm)	Ratio Pb: Apo-PbrR
Apo-PbrR-C77A	2.7×10^{-4}	5×10^{-8}	336	2×10^{-4}
Apo-PbrR-C77A/ C112A	3×10^{-4}	5×10^{-8}	317, 342	2×10^{-4}
Apo-PbrR-C77A/C112A/ C119A	2.54×10^{-4}	3.3×10^{-9}	322	1×10^{-5}
Apo-PbrR-C77D	2.4×10^{-4}	1.39×10^{-7}	328	5×10^{-4}
Apo-PbrR-C77D/ C112D	2.67×10^{-4}	1.1×10^{-7}	330	4×10^{-4}
Apo-PbrR-C77D/C112D/ C119D	2.3×10^{-4}	5×10^{-8}	330	2×10^{-4}

* The concentration of protein was measured by the BioRad assay and the concentration of Pb^{2+} was measured by atomic absorption spectroscopy.

The concentration of mutant apo-proteins and lead treated 'holo' complexes were determined by using the Bio Rad assay and atomic absorption spectroscopy.[14] For the mutant proteins the percentage saturation of lead to protein was less than 50% for all mutants except PbrR-C77D was ~ 88% as shown in Table (5.6).

Table 5.6. The λ_{max} and the ratio of Pb: apo-PbrR(Pp) in the lead treated mutant proteins

Protein	No Moles* Protein (moles)	λ_{max} (nm)	No Moles* Pb^{2+}(moles)	Ratio Pb:Treated-protein
Pb-PbrR-C77A	2.57×10^{-5}	323	1.04×10^{-5}	0.40
Pb-PbrR-C77A/C112A	2.57×10^{-5}	323	2×10^{-6}	0.25
Pb-PbrR-C77A/C112A/C119A	3.2×10^{-5}	323	0.48×10^{-5}	0.15
Pb-PbrR-C77D	7.5×10^{-6}	318	6.6×10^{-6}	0.88
Pb-PbrR-C77D/C112D	8.7×10^{-6}	338	2.8×10^{-6}	0.32
Pb-PbrR-C77D/C112D/C119D	8.8×10^{-6}	331	2.3×10^{-6}	0.26

*Lead treated mutant protein PbrR(Pp). *The concentration of protein was measured by the BioRad assay and the concentration of Pb^{2+} was measured by atomic absorption spectroscopy.

5.4.2. Spectral properties of PbrR-C77A, PbrR-C77A/C112A, PbrR-C77A/C112A/ C119A mutated proteins and complexes

For C77A, C77A/C112A and C77A/C112A/C119A mutants the UV/Vis spectra of both the apo and lead treated forms showed very broad, asymmetric absorption bands with λ_{max} in the range 317-342 nm (Figure 5.7). A long tail appeared on the low energy end of each of the spectra. These measurements were repeated on several different batches of these mutants (and on two separate spectrometers) and they produced the same spectra each time (data not shown). Visible inspection of these samples revealed no turbidity. In addition, the SDS-PAGE gels of these samples showed a single band in all cases (chapter 4, Figure 4.13). Some differences were noted in the overall shape and intensities of the UV/Vis spectra of the corresponding pairs of apo and lead-treated mutants. However, none of these mutants showed the presence of the well-defined Pb-S (Cys) charge transfer band at 332 nm that was observed with the lead-loaded native protein (Figure 5.3). This was interpreted to show that mutation of cysteine to alanine caused significant disruption of the lead coordination sphere at the binding site. Interestingly, analysis of data in Table 5.6 shows evidence for partial lead loading in the lead-treated mutants of C77A, C77A/C112A and C77A/C112A/C119 when percentage saturation was less than 50% (Pb:Protein = 0.4, 0.25, 0.15, respectively).

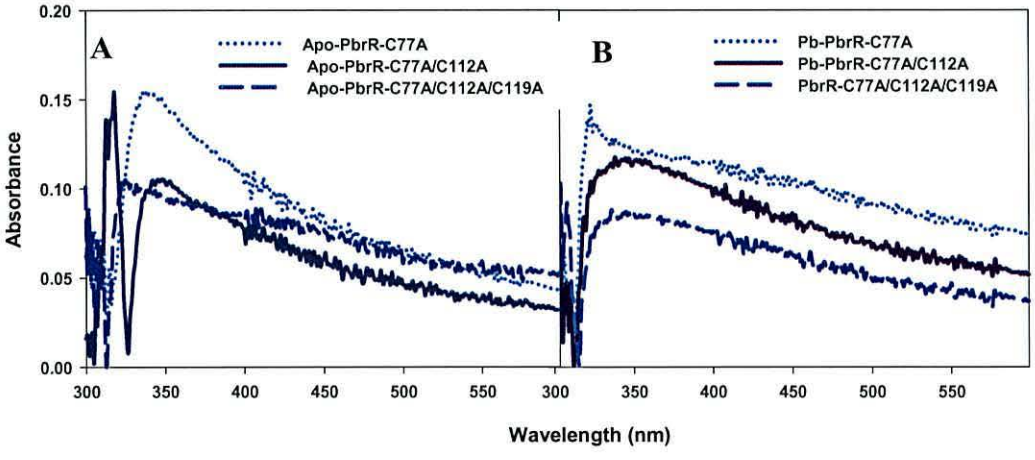


Figure: 5.7. Electronic spectra of the mutated proteins (0.2 mM), (A) As isolated, (B) Lead treated.

The UV/Vis spectral data were inconclusive as to whether the lead was binding at the native site or at another site. The side chain of alanine is a hydrophobic CH_3 group and this will not provide a ligand for a metal centre. Thus, even a single mutation of cysteine to alanine was expected to perturb the metal-binding site. Indeed, as the series of mutants progresses from C77A, C77A/C112A and C77A/C112A/C119, the ratio of Pb:Protein decreases, indicating a decreased ability of this site to bind lead. The ^{207}Pb NMR studies were attempted to better understand the nature of the lead-binding site in the mutants. In each case, the ^{207}Pb NMR chemical shift range from 2500 – 3000 ppm was recorded (native Pb-PbrR has a signal at 2765 ppm). However, no signal in that region was observed – a typical spectrum for the alanine mutants is shown in Figure 5.8.

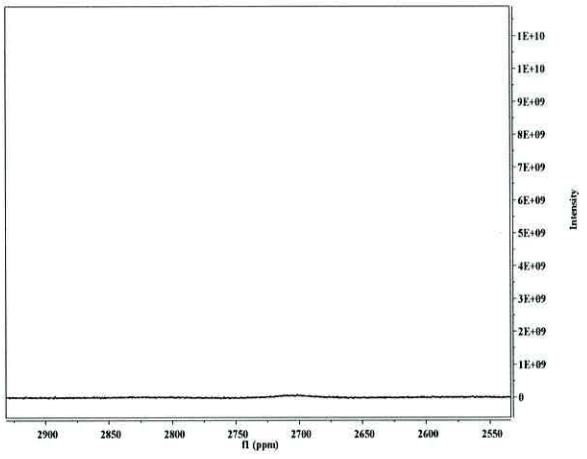


Figure: 5.8. NMR spectra of aqueous solution 5mM Pb-PbrR(Pp)-C77A protein complex containing 10% D_2O

On the basis of NMR chemical shifts (Table 5.1) the possibility of formation of nitrogen- or oxygen-bound species can be ruled out, as a distinct up field chemical shift has previously been observed for mixed-donor ligand types [(PbN_2S, δ = 5318 ppm; PbS_2O_2, δ = 4100–4500 ppm)].[5] Therefore, the observed ^{207}Pb signal could be due to these species. These results also showed that Pb-PbrR(Pp)-C77A also had no signals in that range either. Overall, the observed stoichiometric data, taken with the UV/Vis and ^{207}Pb NMR data, do not provide a clear picture of the lead binding modes in the alanine mutants. The metal:protein ratios indicate partial lead loading

that decreases with increasing mutation of the number of active site cysteine residues to alanine. The lead is not bound at a site which gives rise to the classic Pb-S(cys) charge transfer band or singlet NMR signal, at *ca.* 2765 ppm, associated with the PbS_3 (hemidirected) geometry observed for PbS_3-(GSH) and for native Pb-PbrR(Pp). No ^{207}Pb NMR studies have been conducted on Pb-PbrR(Rm); however, the UV/Vis spectrum for that enzyme is very similar to what has been observed for the Pb-PbrR(Pp) herein. However, the apo-PbrR protein showed higher binding affinity and the selective binding of Pb(II) with thiol group at 2765 ppm corresponds to the PbS_3 range of chemical shift (2577-2853); these results were in agreement with those reported in previous work.[1]

5.4.3. Spectral properties of C77D, C112D and C112/119D protein and complexes

The amino acid residues responsible for Pb^{2+} binding were determined by site-directed mutagenesis to alter each of the three cysteine residues in the *Pseudomonas putida* KT2440 PbrR(Pp) protein to aspartic acid (Figure 5.9). Cysteine residues (C77, C112, C119) were mutated to aspartic acid in several mutants as single, double and triple mutants. For C77D/C112D and C77D/C112D/C119D mutants, the UV/Vis spectra of both the apo and lead-treated forms of C77D/C112D and C77D/C112D/C119D showed very broad, asymmetric absorption bands with λ_{max} in the range 328-338 nm.

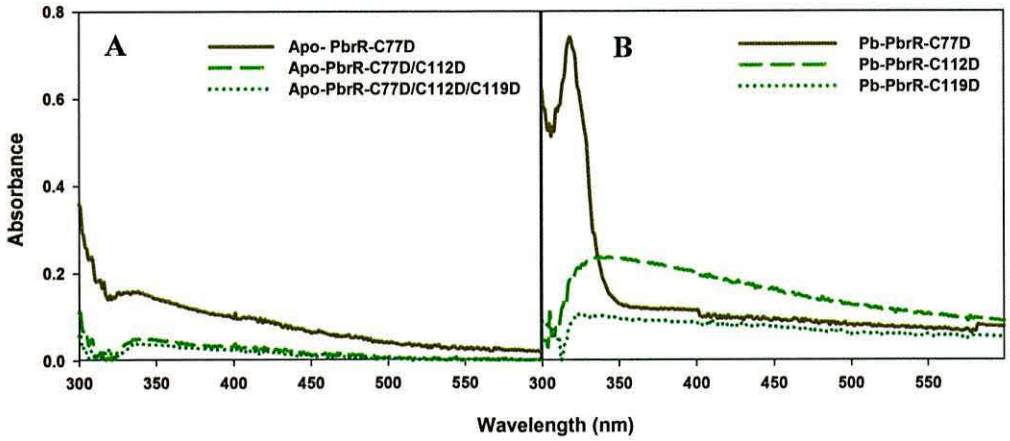


Figure: 5.9. Electronic spectra of the mutant proteins (0.2 mM at pH 7) in the absence (A) and presence (B) of lead.

Given the anomalous shape of the UV/Vis spectra for these mutant (Figure 5.9), the measurements were repeated again to check that the instrumentation was functioning correctly (two different spectrometers were used) and that the samples themselves had not precipitated (no turbidity) and were homogenous (a single band on SDS-PAGE) (Chapter 4, Figure 4.14). Minor differences were seen in the overall shape and intensities of the UV/Vis spectra of the corresponding pairs of apo and lead-treated mutants. A well-defined Pb-S(Cys) charge transfer band at 332 nm was not observed for the double and triple aspartate mutants. The C77D mutant shows a significant difference in the UV/Vis spectra of the apo and lead-loaded forms (Figure 5.9), with the Pb^{2+}-treated PbrR-C77D showing an intense, sharp band at 318nm.

Interestingly, the analysis of data in table 5.5 reveals high lead loading in the lead-treated C77D and partial lead loading in the lead-treated mutants of C77D/C112D and C77D/C112D/C119D. The percentage lead saturation in Pb-PbrR-C77D was 88% but for C77D/C112D and C77D/C112D/C119D were 32% and 26%, respectively. The UV/Vis spectral data are inconclusive regarding whether the lead is binding at the native site or at another site. The carboxyl group on the side-chains of aspartic acid is de-protonated and possesses a negative charge. Thus, even a single mutation of cysteine to aspartic acid was expected to perturb the metal-binding site.

Indeed, as the series of mutant's progresses from C77D, C77D/C112D and C77D/C112D/C119D, the ratio of Pb:Protein decreases, indicating decreased ability of this site to bind lead.

The active site of protein C77 mutated to aspartic acid (D) in the mono mutated C77D and using the purified PCR native protein as a template showed a clear band charge transfer at 318 nm. Aspartic acid has a carboxylic group side chain and may provide an oxygen ligand for lead. This could form a modified binding site with cysteine side chains in the 112 and 119 positions giving rise to the observed extinction coefficient (ε_{318}=3889 M^{-1} cm^{-1}). The ^{207}Pb NMR studies were attempted to better understand the nature of the lead-binding site in the mutants. In each case, a ^{207}Pb NMR chemical shift ranging from 2500 – 3000 ppm was recorded (native PbPbrR has a signal at 2765 ppm). However, no signal in that region was observed for lead-treated C77D/C112D and C77D/C112D/C119D (data not shown). The NMR spectrum of Pb-PbrR-C77D (Figure 5.10) does not show a single strong narrow signal in the region 2764-2868 ppm as expected for a Pb-S3 coordination geometry (Table 5.1).

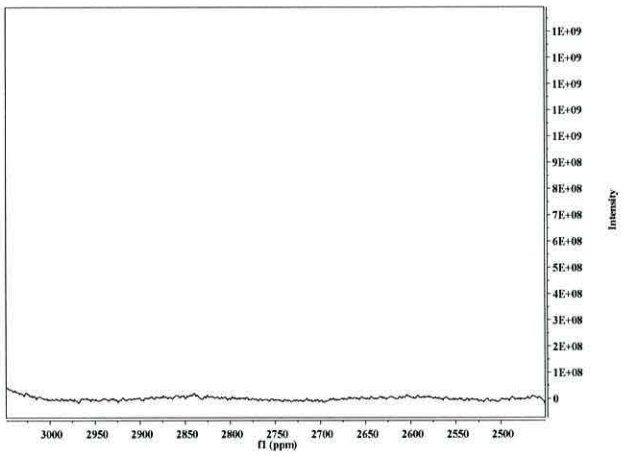


Figure: 5.10. NMR spectra of aqueous solution 5mM Pb-PbrR(Pp)- C77D protein complex containing 10% D_2O.

Table 5.1 indicates that the likely region where a signal from a PbS_2O species would be observed is the region from 1140 to 1555 ppm.[5,6] Due to time constraints in

finishing the laboratory work, re-running the ^{207}Pb NMR over that scan region was not possible (N.B. each ^{207}Pb NMR run took ca. 16 h and was only permitted on weekends due to other user requirements in the School).

The UV/Vis spectroscopy studies previously described showed that the C77D mutant protein possessed an absorption band at 318 nm with an extinction coefficient of $\varepsilon 318=3889\ M^{-1}\ cm^{-1}$. Given that other literature studies have shown that a Pb- S_2O coordination sphere in a GSH has also exhibited a blue shifted band at 317 nm[1], it seems plausible that C77D has a similar coordination sphere. This is not unreasonable given that remaining cysteine residues are likely to form, with the oxygen from aspartate, a suitable metal-binding site.

5.5 Lead nitrate titrations

5.5.1 Pb^{2+} titrations with recombinant native PbrR(Pp)

The UV/Vis spectrum offers an opportunity to characterise lead (II)-binding properties of PbrR(Pp). Binding of lead (II) to PbrR(Pp) was monitored by recording absorption spectra as a function of Pb^{2+}concentrations, as shown in Figure 5.11. Lead-binding titrations of the PbrR(Pp) were carried out by adding small aliquots of $Pb(NO_3)_2$ into protein solution (2μM). In general, this method is restricted by the sensitivity of the spectrometer, and therefore the dynamic range for the concentration measurement of a chromophore is substantially smaller as compared to potentiometric methods. Additionally, titrations such as these that are carried out at constant pH provide only limited information about the solution chemistry.[15]

The electronic absorption spectra of apo-PbrR(Pp) at pH 7 shows an increase in intensity of the LMCT band centred at 332 nm with increasing concentration of $Pb(NO_3)_2$ (Figure 5.11).

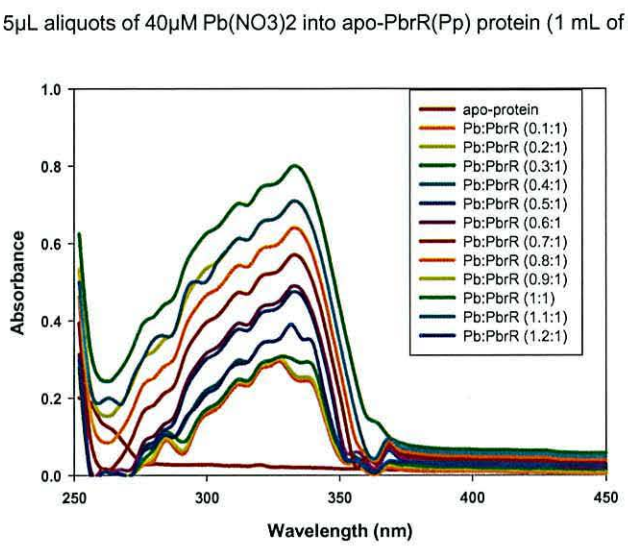


Figure: 5.11. UV/Vis titration of $Pb(NO_3)_2$ into apo-PbrR(Pp) protein.

The UV/Vis data was plotted to show the increase in absorbance at 332 nm vs the ratio of the number of moles of Pb:Protein (Figure 5.12). The data show a steady increase until the metal:protein molar ratio is 1:1 and then the absorbance plateaus with additional aliquots of lead (II) nitrate. This is interpreted as showing that PbrR (Pp) forms a 1:1 complex with Pb^{2+}, which indicates that each PbrR(Pp) protein dimer binds one lead (II) ion.

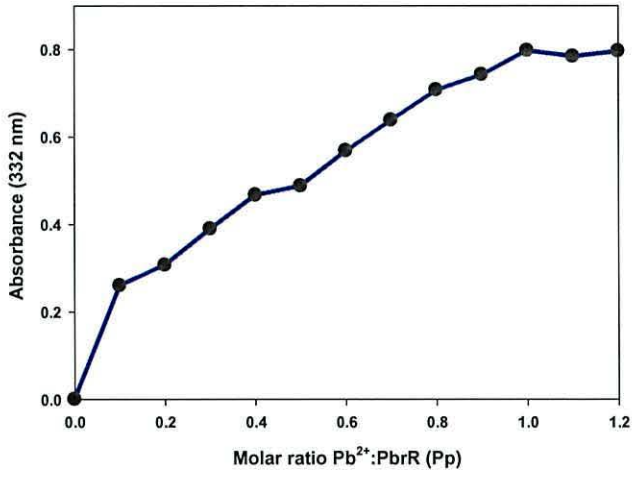


Figure: 5.12. Absorbance change at 332 nm vs $Pb^{2+:}$ PbrR(Pp) protein molar ratio.

This result agrees with the PbrR691 protein after addition of different concentrations of lead and the result of 1:1 shows that PbrR691 forms a 2:1 complex with Pb^{2+} ions; thus, each PbrR691 dimer binds one Pb^{2+} ion.[15]

5.6 Conclusions

In this study, we utilised UV/Vis spectroscopy, NMR spectroscopy, and metal to protein stoichiometry measurements to probe Pb^{2+} binding to the PbrR(Pp) protein and active site cysteine mutants. The results of the ^{207}Pb NMR spectroscopic study and the UV/Vis spectroscopic study on lead treated Pb-PbrR(Pp) unambiguously showed data that were consistent with a PbS_3 coordination environment. This was also consistent with metal to protein stoichiometry and titration data which showed approximately 1 mole of Pb binds to 1 mole of PbrR(Pp). However, most of the cysteine mutants only had partial loading of lead but showed no clear evidence that the lead was bound at the modified active site. The Cys77Asp mutant had partial lead loading and a UV/Vis spectrum consistent with a mixed S/O ligand set, which was expected on the basis of the mutation. Also the ^{207}Pb NMR spectrum of Cys77Asp mutant did not show a signal associated with a PbS_3 species. However due to difficulties with identifying the appropriate chemical shift range to measure data over, because of the extended chemical shift range for the ^{207}Pb nucleus, it was not possible to locate the resonance associated with a OS_2 putative binding site. The double and triple mutated systems do not appear to have active sites viable for binding lead.

5.7. References

1. V. Mah and F. Jalilehvand, Lead (II) complex formation with glutathione, *Inorganic Chemistry*, 2012, **51**, 6285–6298.

2. K. Neupane and V. Pecoraro, Pb-207 NMR spectroscopy reveals that Pb(II) coordinates with glutathione (GSH) and tris cysteine zinc finger proteins in a PbS3 coordination environment, *Journal of Inorganic Biochemistry*, 2011,**105**, 1030-1034.

3. P. Dean, J. Vittal and N. Payne, Discrete trigonal-pyramidal lead (II) complexes: syntheses and x-ray structure analyses of $[(C_6H_5)_4As][Pb(EC_6H_5)_3]$ (E = S, Se), *Inorganic Chemistry,* 1984, **23**, 4232-4236.

4. R. Andersen, R. Ditargiani, R. Hancock, C. Stern, D. Goldberg and H. Godwin, Characterization of the First N_2S(alkylthiolate) lead compound: A model for three-coordinate lead in biological systems, *Inorganic Chem*istry, 2006, **45**, 6574-6576.

5. S. Rupprecht, K. Langemann, T. Lugger, J. McCormick and K. Raymond, Coordination chemistry of *bis*-thiohydroxamic acids: Synthesis and characterization of their lead complexes, determination of their stability constants, *Inorganica Chimica Acta*, 1996, **243**, 79-90.

6. S. Rupprecht, S. Franklin and K. Raymond, Synthesis of monothiohydroxamic ligands and their lead complexes, Structures of N-methyl-3-pyridothiohydroxamic acid, *bis* (*N*-methyl-3-pyridothiohydroxamato)lead(II) and *bis* (*N*-cyclohexyl-phenylacetothio hydroxamato) lead(II), *Inorganica Chimica Acta*, 1995, **235**, 185-194.

7. D. Reger and Y. Ding, Lead(II) complexes containing two different polydentate ligands. Crystal and molecular structure of [HB(3,5-Me_2pz)3]Pb(3,5-Me_2pzH)Cl (pz = Pyrazolyl Ring), a cationic- anionic, double-coordination complex, *Inorganic Chemistry*, 1994,**33**, 4226-4230.

8. E. Claudio, M. ter Horst, C. Forde, C. Stern, M. Zart and H. Godwin, ^{207}Pb-^{1}H two-dimensional NMR spectroscopy: a useful new tool for probing lead (II) coordination chemistry, *Inorganic Chemistry*, 2000, **39**, 1391-1397.

9. D. Reger, M. Huff, A. Rheingold and B. Haggertyt, Control of Structure in Lead(II) Complexes Using Poly(pyrazolyl)borate Ligands, Stereochemically Inactive Lone Pair in Octahedral $[HB(3,5\text{-}Me_2pz)_3]_2Pb$ (pz=Pyrazolyl), *Journal of the American Chemical Society*, 1992, **114**, 579-584.

10. G. Fallon, L. Spiccia, B. West and Q. Zhang, The structure of hydrated lead(ll) dimethylacetate,$[Pb_6(O_2CCH(CH_3)_2)_{12}]4H_2O$ a hexanuclear carboxylate, *Polyhedron*, 1997, ***16***, 19-23

11. B. Wrackmeyer, K. Horchler, *Annul Report NMR Spectroscopy, ^{207}Pb-NMR parameters,* Academic presses inc, 1989, **22,** 249-306.

12. S. Magyar, C. Weng, M. Stern, F. Dye, W. Rous, C. Payne, M. Bridgewater, A. Mijovilovich, G. Parkin, M. Zaleski, J. E. Penner-Hahn and A. Godwin, Reexamination of lead(II) coordination preferences in sulfur-rich sites: Implications for a critical mechanism of lead poisoning, *Journal of the American Chemical Society*, 2005, **127,** 9495-9505.

13. P. Chen, B. Greenberg, S. Taghavi, C. Romano, D. van der Lelie and C. He, An Exceptionally Selective Lead (II)-Regulatory Protein from *Ralstonia Metallidurans*: Development of a Fluorescent Lead (II) Probe, *Angewandte Chemie*, **2005,** 117, 2715-2719.

14. P. Chen, E. Wasinger, J. Zhao, D. Lelie, L. Chen and C. He, Spectroscopic insights into lead (II) coordination by the selective lead (II)-binding protein PbrR691, *Journal of the American Chemical Society*, 2007, **129**, 12350–12351.

15. C. Fahrni and T. O'Halloran, Aqueous coordination chemistry of quinoline-based fluorescence probes for the biological chemistry of zinc, *Journal of the American Chemical Society*, 1999. **121**, 11448-11458.

CHAPTER 6: CONCLUSIONS AND FUTURE WORK

This thesis was divided into two main parts. The objective of first part was to produce PbrR protein from *Pseudomonas putida* KT2440. The sequence comparison program BLAST was used to identify the DNA sequence similarities of PbrR gene in different micro-organisms (98 species) compared to *Ralstonia metallidurans* CH34. *Pseudomonas putida* KT2440 showed 57% of similarity to the PbrR gene from *Ralstonia metallidurans* CH34 and was selected as an appropriate target. The DNA sequences were isolated from *Pseudomonas putida* KT2440 followed by amplification, ligation (between restriction sites *EcoRI and HindIII*) and cloning into vector pET28a (+). Plasmid pET28a (+) was used to produce pure and adequate quantities of his-tagged PbrR(Pp). The successful overexpression was performed by transforming the PbrR-PET28a (+) plasmid into the Rosetta expression system.

The overproduction of large amounts of the native and mutant PbrR(Pp) was highly dependent on the induction time. Induction for a period of 4 hours by IPTG did not result in the production of the PbrR. Hence, auto-induction was employed; its duration seems to have a significant effect on protein production. 12 hours of auto-induction did not result in the production of the mature PbrR, while longer periods (16-20 h) produced sufficient quantities of the apo-protein (62μg/ml). This may be due to the time required for the processing and maturation of the protein.

Overproduced PbrR(Pp) protein was purified by a His-tag column and removing imidazole salts with a PD-10 desalting protein for pure protein. The molecular weight of recombinant and mutated proteins subunits were checked by SDS-PAGE to confirm the approximate size (26 kDa monomer protein) and purity.

The PbrR gene was used as a template for the generation of two different single mutants: C77A and C77D. Multiple mutants were prepared using a PbrR-C77A template to give PbrR-C77A /C112A and PbrR-C77A /C112A/C119A. A PbrR-C77D template was used to prepare PbrR-C77D/C112D and PbrR-C77D

/C112D/C119D mutants. All the mutants have been successfully overproduced and show strong bands at the expected molecular weight for the protein subunit of 26 kDa on SDS-PAGE gels.

The bioinformatics and molecular biology section of the project was then followed by the second part experiments to study the Pb^{2+} binding properties of PbrR(Pp) protein and active site cysteine mutants using UV/Vis spectroscopy, protein concentration assays, atomic absorption spectroscopy and ^{207}Pb NMR spectroscopy.

Metal to protein stoichiometry measurements indicated that as the as isolated protein (wild type) contained typical Pb:protein ratios (1:0.99) and that either dialysis or titration vs aqueous Pb^{2+} was required to get near to 1:1 stoichiometries for PbrR (Pp) and PbrR-C77D only. However, the results of studies on PbrR-C77A, PbrR-C77A/C112A, PbrR-C77A /C112/119A, PbrR-C77D/C112D, and PbrR-C77/C112/119D revealed a partial lead binding.

A Pb-glutathione model complex was prepared which had been previously established to consist of a very similar coordination sphere (PbS_3) to that expected in Pb-PbrR(Pp). The Pb-GSH complex clearly showed a strong charge transfer band at 334 nm in the UV/Vis spectrum and a single signal at 2764 ppm ^{207}Pb NMR spectrum. These two spectroscopic features where used to infer the nature of the coordination sphere in Pb-PbrR (Pp) and the mutant proteins.

The ^{207}Pb NMR and UV/Vis spectra of Pb-PbrR (Pp) clearly showed the presence of a single NMR signal at 2765 ppm and a strong charge transfer band at 332 nm, respectively. The conclusion was that these results were consistent with coordination geometry of PbS_3 (Cys). This was also consistent with the observed coordination geometry known for the related system Pb-PbrR(Rm). However, similar studies conducted on the lead-treated PbrR–C77D mutant showed that there was only partial lead binding (protein to lead ratio of 1:0.88). The NMR results indicated that there was not an observable signal in the range expected for PbS_3 (Cys) coordination sphere. The UV/Vis spectroscopic studies on PbrR–C77D did, however, show a

strong band at 318 nm. Similar blue shifted bands have been observed with a PbS_2O/N coordination geometry (see section 1.2.3)

All the other mutants proteins PbrR-C77A, PbrR-C77A/C112A, PbrR-C77A/C112A/C119A, PbrR-C77D/C112D, and PbrR-C77D/C112D/C119D exhibited no well-defined UV/Vis absorption band in the region 315 – 350 nm and no signal in the (2000-3000) ppm by ^{207}Pb NMR spectroscopy.

Future work

There are some experiments that with more time would be useful to revisit to improve the data obtained and to try and a clearer picture of the nature of the binding site in the PbrR-C77D mutant. For example it would be important to explore a large range of chemical shifts in the ^{207}Pb NMR experiments to locate the signal associated with the putative OS_2 species. The affinity between the metal and a protein could be determined by additional experiments to determine the binding constant of Pb^{2+} to native and mutant (C77D) PbrR protein (e.g equilibrium analytical ultracentrifugation (AU)). In addition it would be important to check if any other metals e.g. Cd^{2+} would bind to native or mutant PbrR(Pp).

It would be useful to improve on the yield of recombinant native protein and mutants obtained in this study, conduct the purification an aerobically and try a bigger of condition e.g. induction time, temperature, different method of cell breakage (e.g. French press). This would permit further investigation using more sophisticated techniques such as protein crystallography and x-ray absorption spectroscopy. X-ray crystallography has the potential to provide the high resolution 3D structures of native and C77D. Pb L-edge x-ray absorption spectroscopy would provide accurate bond lengths within 5 Å of the Pb and permit studies to be done under a wider range of conditions than possible for protein crystallography. Both of these techniques would provide information to complement the data obtained by the ^{207}Pb NMR measurements.

The NMR measurements could be further expanded to include 2D heteronuclear single-quantum correlation spectroscopy (HMQC); this method is more detailed than one dimensional of NMR spectra. It is also has the potential to provide a detailed structure of a molecule.

7. APPENDIX

7.1.Amino acid sequence of PbrR *Pseudomonas putida* KT2440 (147aa)

```
       10         20         30         40         50         60

MKIGELAKAT DCAVETIRYY ERENLLPEPA RSEGNYRLYT QAHVERLTFI RNCRTLDMTL
         70         80         90        100        110        120
DEIRSLLRLR DSPDDACGSV NALIDEHIEH VQARIDGLVA LQEQLVELRR RCNTQGSECA

       130        140
ILQQLETNGA VSVPDTEHSH VGRSHGH
```

7.1.1. Nucleotides sequence of PbrR *Pseudomonas putida* KT2440 (444 bp) (5866225-5866668)

```
5866201 aagccacttg ctggagattt cgtgatgaag atcggagaac tggccaaagc caccgactgc

5866261 gcggtggaaa ccatccgcta ctacgagcgt gaaaacctgc tgccagagcc ggcgcgcagc

5866321 gagggcaact accggttgta cacccaggcc catgtggagc ggctgacctt catccgcaac

5866381 tgccgcacgc tggacatgac cctggacgaa attcgcagcc tgctacgcct gcgcgacagc

5866441 cccgacgacg cgtgcggcag cgtcaatgcg ctgatcgacg agcatatcga gcatgttcag

5866501 gcgcggatcg atggcttggt ggcattgcag gagcagctgg tggagctgcg gcggcgctgc

5866561 aacgcgcagg ggagtgaatg cgcgatcttg cagcaactgg agacaaacgg ggcggtatcg

5866621 gtaccggata ccgaacattc ccatgtgggg cggagtcacg ggcattgagc caaaagcttc

5866681 gcgggtgaac ccgcgaagag gtcagttcta acaatagaga tctatcagac cgccatcggc
```

7.2. SEQUENCES CHECK

7.2.1. PbrR(Pp)-DNA gene (Forward primer) (Length=986)

```
Score = 815 bits (441), Expect = 0.0, Identities = 441/441 (100%),
Gaps = 0/441 (0%), Strand=Plus/Plus

Query  1    ATGAAGATCGGAGAACTGGCCAAAGCCACCGACTGCGCGGTGGAAACCATCCGCTACTAC  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 164   ATGAAGATCGGAGAACTGGCCAAAGCCACCGACTGCGCGGTGGAAACCATCCGCTACTAC  223
```

```
Query 61   GAGCGTGAAAACCTGCTGCCAGAGCCGGCGCGCAGCGAGGGCAACTACCGGTTGTACACC  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 224  GAGCGTGAAAACCTGCTGCCAGAGCCGGCGCGCAGCGAGGGCAACTACCGGTTGTACACC  283

Query 121  CAGGCCCATGTGGAGCGGCTGACCTTCATCCGCAACTGCCGCACGCTGGACATGACCCTG  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 284  CAGGCCCATGTGGAGCGGCTGACCTTCATCCGCAACTGCCGCACGCTGGACATGACCCTG  343

Query 181  GACGAAATTCGCAGCCTGCTACGCCTGCGCGACAGCCCCGACGACGCGTGCGGCAGCGTC  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 344  GACGAAATTCGCAGCCTGCTACGCCTGCGCGACAGCCCCGACGACGCGTGCGGCAGCGTC  403

Query 241  AATGCGCTGATCGACGAGCATATCGAGCATGTTCAGGCGCGGATCGATGGCTTGGTGGCA  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 404  AATGCGCTGATCGACGAGCATATCGAGCATGTTCAGGCGCGGATCGATGGCTTGGTGGCA  463

Query 301  TTGCAGGAGCAGCTGGTGGAGCTGCGGCGGCGCTGCAACGCGCAGGGGAGTGAATGCGCG  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 464  TTGCAGGAGCAGCTGGTGGAGCTGCGGCGGCGCTGCAACGCGCAGGGGAGTGAATGCGCG  523

Query 361  ATCTTGCAGCAACTGGAGACAAACGGGGCGGTATCGGTACCGGATACCGAACATTCCCAT  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 524  ATCTTGCAGCAACTGGAGACAAACGGGGCGGTATCGGTACCGGATACCGAACATTCCCAT  583

Query  421  GTGGGGCGGAGTCACGGGCAT  441
             ||||||||||||||||||||
Sbjct  584  GTGGGGCGGAGTCACGGGCAT  604
```

7.2.2.Apo-PbrR(Pp)- DNA gene (Reverse primer) (Length=738)

```
Score = 815 bits (441), Expect = 0.0, Identities = 441/441 (100%),
Gaps = 0/441 (0%), Strand=Plus/Minus

Query  1    ATGAAGATCGGAGAACTGGCCAAAGCCACCGACTGCGCGGTGGAAACCATCCGCTACTAC  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 451   ATGAAGATCGGAGAACTGGCCAAAGCCACCGACTGCGCGGTGGAAACCATCCGCTACTAC  392

Query 61    GAGCGTGAAAACCTGCTGCCAGAGCCGGCGCGCAGCGAGGGCAACTACCGGTTGTACACC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 391   GAGCGTGAAAACCTGCTGCCAGAGCCGGCGCGCAGCGAGGGCAACTACCGGTTGTACACC  332

Query 121   CAGGCCCATGTGGAGCGGCTGACCTTCATCCGCAACTGCCGCACGCTGGACATGACCCTG  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 331   CAGGCCCATGTGGAGCGGCTGACCTTCATCCGCAACTGCCGCACGCTGGACATGACCCTG  272

Query 181   GACGAAATTCGCAGCCTGCTACGCCTGCGCGACAGCCCCGACGACGCGTGCGGCAGCGTC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 271   GACGAAATTCGCAGCCTGCTACGCCTGCGCGACAGCCCCGACGACGCGTGCGGCAGCGTC  212

Query 241   AATGCGCTGATCGACGAGCATATCGAGCATGTTCAGGCGCGGATCGATGGCTTGGTGGCA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 211   AATGCGCTGATCGACGAGCATATCGAGCATGTTCAGGCGCGGATCGATGGCTTGGTGGCA  152

Query 301   TTGCAGGAGCAGCTGGTGGAGCTGCGGCGGCGCTGCAACGCGCAGGGGAGTGAATGCGCG  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 151   TTGCAGGAGCAGCTGGTGGAGCTGCGGCGGCGCTGCAACGCGCAGGGGAGTGAATGCGCG  92

Query 361   ATCTTGCAGCAACTGGAGACAAACGGGGCGGTATCGGTACCGGATACCGAACATTCCCAT  420
```

```
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 91    ATCTTGCAGCAACTGGAGACAAACGGGGCGGTATCGGTACCGGATACCGAACATTCCCAT  32

Query  421  GTGGGGCGGAGTCACGGGCAT  441
            |||||||||||||||||||||
Sbjct  31   GTGGGGCGGAGTCACGGGCAT  11
```

7.2.3. Apo-PbrR (Pp)-mutants DNA gene

Mono mutant C77A, C77D double mutant C112A, C112D and used gene mono mutant as template , and triple mutant Cys112/119 A and C112/119D at the same time and used gene mono mutant as template.

7.2.3.1. Cysteine (TGC) mutated to alanine (GCC)

```
       10         20         30         40         50         60

MKIGELAKAT DCAVETIRYY ERENLLPEPA RSEGNYRLYT QAHVERLTFI RNCRTLDMTL
         70         80         90        100        110        120
DEIRSLLRLR DSPDDAAGSV NALIDEHIEH VQARIDGLVA LQEQLVELRR RANTQGSEAA

       130        140
ILQQLETNGA VSVPDTEHSH VGRSHGH
```

7.2.3.1.1. PbrR(Pp)-C77A DNA gene (Forward primer)(Length: 444)

Score = 815 bits (441), Expect = 0.0, Identities = 441/441 (100%),Gaps =0/441 (0%), Strand=Plus/Plus

```
Query 176  ATGAAGATCGGAGAACTGGCCAAAGCCACCGACTGCGCGGTGGAAACCATCCGCTACTAC  235
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 1    ATGAAGATCGGAGAACTGGCCAAAGCCACCGACTGCGCGGTGGAAACCATCCGCTACTAC  60

Query 236  GAGCGTGAAAACCTGCTGCCAGAGCCGGCGCGCAGCGAGGGCAACTACCGGTTGTACACC  295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 61   GAGCGTGAAAACCTGCTGCCAGAGCCGGCGCGCAGCGAGGGCAACTACCGGTTGTACACC  120

Query 296  CAGGCCCATGTGGAGCGGCTGACCTTCATCCGCAACTGCCGCACGCTGGACATGACCCTG  355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 121  CAGGCCCATGTGGAGCGGCTGACCTTCATCCGCAACTGCCGCACGCTGGACATGACCCTG  180

Query 356  GACGAAATTCGCAGCCTGCTACGCCTGCGCGACAGCCCCGACGACGCGGCCGGCAGCGTC  415
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 181  GACGAAATTCGCAGCCTGCTACGCCTGCGCGACAGCCCCGACGACGCGGCCGGCAGCGTC  240

Query 416  AATGCGCTGATCGACGAGCATATCGAGCATGTTCAGGCGCGGATCGATGGCTTGGTGGCA  475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 241  AATGCGCTGATCGACGAGCATATCGAGCATGTTCAGGCGCGGATCGATGGCTTGGTGGCA  300

Query 476  TTGCAGGAGCAGCTGGTGGAGCTGCGGCGGCGCTGCAACGCGCAGGGGAGTGAATGCGCG  535
```

```
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 301   TTGCAGGAGCAGCTGGTGGAGCTGCGGCGGCGCTGCAACGCGCAGGGGAGTGAATGCGCG  360

Query 536   ATCTTGCAGCAACTGGAGACAAACGGGGCGGTATCGGTACCGGATACCGAACATTCCCAT  595
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 361   ATCTTGCAGCAACTGGAGACAAACGGGGCGGTATCGGTACCGGATACCGAACATTCCCAT  420

Query 596   GTGGGGCGGAGTCACGGGCAT  616
              ||||||||||||||||||||
Sbjct 421   GTGGGGCGGAGTCACGGGCAT  441
```

7.2.3.1.2.PbrR(Pp)-C77A DNA gene (Reverse primer) (Length=444)

Score = 815 bits (441), Expect = 0.0,Identities = 441/441 (100%), Gaps= 0/441 (0%), Strand=Plus/Minus

```
Query 26    ATGCCCGTGACTCCGCCCCACATGGGAATGTTCGGTATCCGGTACCGATACCGCCCCGTT  85
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 441   ATGCCCGTGACTCCGCCCCACATGGGAATGTTCGGTATCCGGTACCGATACCGCCCCGTT  382

Query 86    TGTCTCCAGTTGCTGCAAGATCGCGCATTCACTCCCCTGCGCGTTGCAGCGCCGCCGCAG  145
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 381   TGTCTCCAGTTGCTGCAAGATCGCGCATTCACTCCCCTGCGCGTTGCAGCGCCGCCGCAG  322

Query 146   CTCCACCAGCTGCTCCTGCAATGCCACCAAGCCATCGATCCGCGCCTGAACATGCTCGAT  205
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 321   CTCCACCAGCTGCTCCTGCAATGCCACCAAGCCATCGATCCGCGCCTGAACATGCTCGAT  26

Query 206   ATGCTCGTCGATCAGCGCATTGACGCTGCCGGCCGCGTCGTCGGGGCTGTCGCGCAGGCG  265
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 261   ATGCTCGTCGATCAGCGCATTGACGCTGCCGGCCGCGTCGTCGGGGCTGTCGCGCAGGCG  202

Query 266   TAGCAGGCTGCGAATTTCGTCCAGGGTCATGTCCAGCGTGCGGCAGTTGCGGATGAAGGT  325
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  201  TAGCAGGCTGCGAATTTCGTCCAGGGTCATGTCCAGCGTGCGGCAGTTGCGGATGAAGGT  14

Query 326   CAGCCGCTCCACATGGGCCTGGGTGTACAACCGGTAGTTGCCCTCGCTGCGCGCCGGCTC  385
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 141   CAGCCGCTCCACATGGGCCTGGGTGTACAACCGGTAGTTGCCCTCGCTGCGCGCCGGCTC  82

Query 386   TGGCAGCAGGTTTTCACGCTCGTAGTAGCGGATGGTTTCCACCGCGCAGTCGGTGGCTTT  445
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 81    TGGCAGCAGGTTTTCACGCTCGTAGTAGCGGATGGTTTCCACCGCGCAGTCGGTGGCTTT  22

Query 446   GGCCAGTTCTCCGATCTTCAT  466
            ||||||||||||||||||||
Sbjct  21    GGCCAGTTCTCCGATCTTCAT  1
```

7.2.3.1.3. PbrR(Pp)-C77A/C112A DNA gene(forward primer)(Length: 444)

Score = 808 bits (437), Expect = 0.0, Identities = 440/441 (99%),Gaps =1/441 (0%), Strand=Plus/Plus

```
Query 106   ATGAAGATCGGAGAACTGGCCAAAGCCACCGACTGCGCGGTGGAAACCATCCGCTACTAC  165
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 1     ATGAAGATCGGAGAACTGGCCAAAGCCACCGACTGCGCGGTGGAAACCATCCGCTACTAC  60
```

```
Query 166  GAGCGTGAAAACCTGCTGCCAGAGCCGGCGCGCAGCGAGGGCAACTACCGGTTGTACACC  225
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 61   GAGCGTGAAAACCTGCTGCCAGAGCCGGCGCGCAGCGAGGGCAACTACCGGTTGTACACC  120

Query 226  CAGGCCCATGTGGAGCGGCTGACCTTCATCCGCAACTGCCGCACGCTGGACATGACCCTG  285
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 121  CAGGCCCATGTGGAGCGGCTGACCTTCATCCGCAACTGCCGCACGCTGGACATGACCCTG  180

Query 286  GACGAAATTCGCAGCCTGCTACGCCTGCGCGACAGCCCCGACGACGCGG-CGGCAGCGTC  344
           |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Sbjct 181  GACGAAATTCGCAGCCTGCTACGCCTGCGCGACAGCCCCGACGACGCGGCCGGCAGCGTC  240

Query 345  AATGCGCTGATCGACGAGCATATCGAGCATGTTCAGGCGCGGATCGATGGCTTGGTGGCA  404
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 241  AATGCGCTGATCGACGAGCATATCGAGCATGTTCAGGCGCGGATCGATGGCTTGGTGGCA  300

Query 405  TTGCAGGAGCAGCTGGTGGAGCTGCGGCGGCGCGCCAACGCGCAGGGGAGTGAATGCGCG  464
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 301  TTGCAGGAGCAGCTGGTGGAGCTGCGGCGGCGCGCCAACGCGCAGGGGAGTGAATGCGCG  360

Query 465  ATCTTGCAGCAACTGGAGACAAACGGGGCGGTATCGGTACCGGATACCGAACATTCCCAT  524
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 361  ATCTTGCAGCAACTGGAGACAAACGGGGCGGTATCGGTACCGGATACCGAACATTCCCAT  420

Query 525  GTGGGGCGGAGTCACGGGCAT  545
           ||||||||||||||||||||
Sbjct 421  GTGGGGCGGAGTCACGGGCAT  441
```

7.2.3.1.4.PbrR(Pp)-C77A/C112A DNA gene(Reverse primer)(Length: 444))

Score = 808 bits (437),Expect = 0.0,Identities =440/441 (99%), Gaps= 1/441(0%), Strand=Plus/Minus

```
Query 16   ATGCCCGTGACTCCGCCCCACATGGGAATGTTCGGTATCCGGTACCGATACCGCCCCGTT  75
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 441  ATGCCCGTGACTCCGCCCCACATGGGAATGTTCGGTATCCGGTACCGATACCGCCCCGTT  382

Query 76   TGTCTCCAGTTGCTGCAAGATCGCGCATTCACTCCCCTGCGCGTTGGCGCGCCGCCGCAG  135
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 381  TGTCTCCAGTTGCTGCAAGATCGCGCATTCACTCCCCTGCGCGTTGGCGCGCCGCCGCAG  322

Query 136  CTCCACCAGCTGCTCCTGCAATGCCACCAAGCCATCGATCCGCGCCTGAACATGCTCGAT  195
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 321  CTCCACCAGCTGCTCCTGCAATGCCACCAAGCCATCGATCCGCGCCTGAACATGCTCGAT  262


Query 196  ATGCTCGTCGATCAGCGCATTGACGCTGCCG-CCGCGTCGTCGGGGCTGTCGCGCAGGCG  254
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Sbjct 261  ATGCTCGTCGATCAGCGCATTGACGCTGCCGGCCGCGTCGTCGGGGCTGTCGCGCAGGCG  202

Query 255  TAGCAGGCTGCGAATTTCGTCCAGGGTCATGTCCAGCGTGCGGCAGTTGCGGATGAAGGT  314
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 201  TAGCAGGCTGCGAATTTCGTCCAGGGTCATGTCCAGCGTGCGGCAGTTGCGGATGAAGGT  142

Query 315  CAGCCGCTCCACATGGGCCTGGGTGTACAACCGGTAGTTGCCCTCGCTGCGCGCCGGCTC  374
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 141  CAGCCGCTCCACATGGGCCTGGGTGTACAACCGGTAGTTGCCCTCGCTGCGCGCCGGCTC  82
```

```
Query 375  TGGCAGCAGGTTTTCACGCTCGTAGTAGCGGATGGTTTCCACCGCGCAGTCGGTGGCTTT  434
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 81   TGGCAGCAGGTTTTCACGCTCGTAGTAGCGGATGGTTTCCACCGCGCAGTCGGTGGCTTT  22

Query 435  GGCCAGTTCTCCGATCTTCAT  455
           ||||||||||||||||||||
Sbjct 21   GGCCAGTTCTCCGATCTTCAT  1
```

7.2.3.1.5.PbrR(Pp)-C77A/C112A/C119A DNA gene(forward primer) (Length: 444)

Score = 399 bits (216), Expect = 1e-115, Identities = 216/216(100%), Gaps =0/216 (0%), Strand=Plus/Plus

```
Query 262  GCGGCCGGCAGCGTCAATGCGCTGATCGACGAGCATATCGAGCATGTTCAGGCGCGGATC  321
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 226  GCGGCCGGCAGCGTCAATGCGCTGATCGACGAGCATATCGAGCATGTTCAGGCGCGGATC  285

Query 322  GATGGCTTGGTGGCATTGCAGGAGCAGCTGGTGGAGCTGCGGCGGCGCGCCAACGCGCAG  381
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 286  GATGGCTTGGTGGCATTGCAGGAGCAGCTGGTGGAGCTGCGGCGGCGCGCCAACGCGCAG  345

Query 382  GGGAGTGAAGCCGCGATCTTGCAGCAACTGGAGACAAACGGGGCGGTATCGGTACCGGAT  441
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 346  GGGAGTGAAGCCGCGATCTTGCAGCAACTGGAGACAAACGGGGCGGTATCGGTACCGGAT  405

Query 442  ACCGAACATTCCCATGTGGGGCGGAGTCACGGGCAT  477
           ||||||||||||||||||||||||||||||||||||
Sbjct 406  ACCGAACATTCCCATGTGGGGCGGAGTCACGGGCAT  441
```

7.2.3.1.6.PbrR(Pp)-C77A/C112A/C119A DNA gene (Reverse primer) (Length: 444)

Score = 399 bits (216), Expect = 1e-115, Identities = 216/216(100%), Gaps =0/216(0%), Strand=Plus/ Minus

```
Query 17   ATGCCCGTGACTCCGCCCCACATGGGAATGTTCGGTATCCGGTACCGATACCGCCCCGTT  76
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 441  ATGCCCGTGACTCCGCCCCACATGGGAATGTTCGGTATCCGGTACCGATACCGCCCCGTT  382

Query 77   TGTCTCCAGTTGCTGCAAGATCGCGGCTTCACTCCCCTGCGCGTTGGCGCGCCGCCGCAG  136
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 381  TGTCTCCAGTTGCTGCAAGATCGCGGCTTCACTCCCCTGCGCGTTGGCGCGCCGCCGCAG  322

Query 137  CTCCACCAGCTGCTCCTGCAATGCCACCAAGCCATCGATCCGCGCCTGAACATGCTCGAT  196
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 321  CTCCACCAGCTGCTCCTGCAATGCCACCAAGCCATCGATCCGCGCCTGAACATGCTCGAT  262

Query 197  ATGCTCGTCGATCAGCGCATTGACGCTGCCGGCCGC  232
           ||||||||||||||||||||||||||||||||||||
Sbjct 261  ATGCTCGTCGATCAGCGCATTGACGCTGCCGGCCGC  226
```

7.2.3.2. Cysteine (TGC) mutated to aspartic acid (GAC)

```
       10         20         30         40         50         60

MKIGELAKAT DCAVETIRYY ERENLLPEPA RSEGNYRLYT QAHVERLTFI RNCRTLDMTL
         70         80         90        100        110        120
DEIRSLLRLR DSPDDADGSV NALIDEHIEH VQARIDGLVA LQEQLVELRR RDNTQGSEDA

       130        140
ILQQLETNGA VSVPDTEHSH VGRSHGH
```

7.2.3.2.1.PbrR(Pp)-C77D DNA gene (forward primer) (Length: 444)

Score = 815 bits (441), Expect = 0.0, Identities = 441/441 (100%), Gaps =1/441 (0%), Strand=Plus/Plus

```
Query 176  ATGAAGATCGGAGAACTGGCCAAAGCCACCGACTGCGCGGTGGAAACCATCCGCTACTAC  235
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 1    ATGAAGATCGGAGAACTGGCCAAAGCCACCGACTGCGCGGTGGAAACCATCCGCTACTAC  60

Query 236  GAGCGTGAAAACCTGCTGCCAGAGCCGGCGCGCAGCGAGGGCAACTACCGGTTGTACACC  295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 61   GAGCGTGAAAACCTGCTGCCAGAGCCGGCGCGCAGCGAGGGCAACTACCGGTTGTACACC  120

Query 296  CAGGCCCATGTGGAGCGGCTGACCTTCATCCGCAACTGCCGCACGCTGGACATGACCCTG  355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 121  CAGGCCCATGTGGAGCGGCTGACCTTCATCCGCAACTGCCGCACGCTGGACATGACCCTG  180

Query 356  GACGAAATTCGCAGCCTGCTACGCCTGCGCGACAGCCCCGACGACGCGGACGGCAGCGTC  415
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 181  GACGAAATTCGCAGCCTGCTACGCCTGCGCGACAGCCCCGACGACGCGGACGGCAGCGTC  240

Query 416  AATGCGCTGATCGACGAGCATATCGAGCATGTTCAGGCGCGGATCGATGGCTTGGTGGCA  475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 241  AATGCGCTGATCGACGAGCATATCGAGCATGTTCAGGCGCGGATCGATGGCTTGGTGGCA  300

Query 476  TTGCAGGAGCAGCTGGTGGAGCTGCGGCGGCGCTGCAACGCGCAGGGGAGTGAATGCGCG  535
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 301  TTGCAGGAGCAGCTGGTGGAGCTGCGGCGGCGCTGCAACGCGCAGGGGAGTGAATGCGCG  360

Query 536  ATCTTGCAGCAACTGGAGACAAACGGGGCGGTATCGGTACCGGATACCGAACATTCCCAT  595
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 361  ATCTTGCAGCAACTGGAGACAAACGGGGCGGTATCGGTACCGGATACCGAACATTCCCAT  420

Query 596  GTGGGGCGGAGTCACGGGCAT  616
           ||||||||||||||||||||
Sbjct 421  GTGGGGCGGAGTCACGGGCAT  441
```

7.2.3.2.2. PbrR(Pp)-C77D DNA gene (Reverse primer)(Length: 444)

Score = 809 bits (438), Expect = 0.0, Identities = 440/441 (99%), Gaps =0/441 (0%), Strand=Plus/Minus

```
Query 23   ATGCCCGTGACGCCGCCCCACATGGGAATGTTCGGTATCCGGTACCGATACCGCCCCGTT  82
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 441  ATGCCCGTGACTCCGCCCCACATGGGAATGTTCGGTATCCGGTACCGATACCGCCCCGTT  382

Query 83   TGTCTCCAGTTGCTGCAAGATCGCGCATTCACTCCCCTGCGCGTTGCAGCGCCGCCGCAG  142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 381  TGTCTCCAGTTGCTGCAAGATCGCGCATTCACTCCCCTGCGCGTTGCAGCGCCGCCGCAG  322

Query 143  CTCCACCAGCTGCTCCTGCAATGCCACCAAGCCATCGATCCGCGCCTGAACATGCTCGAT  202
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 321  CTCCACCAGCTGCTCCTGCAATGCCACCAAGCCATCGATCCGCGCCTGAACATGCTCGAT  262

Query 203  ATGCTCGTCGATCAGCGCATTGACGCTGCCGTCCGCGTCGTCGGGGCTGTCGCGCAGGCG  262
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 261  ATGCTCGTCGATCAGCGCATTGACGCTGCCGTCCGCGTCGTCGGGGCTGTCGCGCAGGCG  202

Query 263  TAGCAGGCTGCGAATTTCGTCCAGGGTCATGTCCAGCGTGCGGCAGTTGCGGATGAAGGT  322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 201  TAGCAGGCTGCGAATTTCGTCCAGGGTCATGTCCAGCGTGCGGCAGTTGCGGATGAAGGT  142

Query 323  CAGCCGCTCCACATGGGCCTGGGTGTACAACCGGTAGTTGCCCTCGCTGCGCGCCGGCTC  382
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 141  CAGCCGCTCCACATGGGCCTGGGTGTACAACCGGTAGTTGCCCTCGCTGCGCGCCGGCTC  82

Query 383  TGGCAGCAGGTTTTCACGCTCGTAGTAGCGGATGGTTTCCACCGCGCAGTCGGTGGCTTT  442
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 81   TGGCAGCAGGTTTTCACGCTCGTAGTAGCGGATGGTTTCCACCGCGCAGTCGGTGGCTTT  22

Query 443  GGCCAGTTCTCCGATCTTCAT  463
           ||||||||||||||||||||
Sbjct 21   GGCCAGTTCTCCGATCTTCAT  1
```

7.2.3.2.3.PbrR(Pp)-C77D/C112D DNA gene(forward primer) (Length: 444)

Score = 815bits (441), Expect = 0.0, Identities = 441/441 (100%), Gaps =0/441 (0%), Strand=Plus/Plus

```
Query 178  ATGAAGATCGGAGAACTGGCCAAAGCCACCGACTGCGCGGTGGAAACCATCCGCTACTAC  237
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 1    ATGAAGATCGGAGAACTGGCCAAAGCCACCGACTGCGCGGTGGAAACCATCCGCTACTAC  60

Query 238  GAGCGTGAAAACCTGCTGCCAGAGCCGGCGCGCAGCGAGGGCAACTACCGGTTGTACACC  297
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 61   GAGCGTGAAAACCTGCTGCCAGAGCCGGCGCGCAGCGAGGGCAACTACCGGTTGTACACC  120

Query 298  CAGGCCCATGTGGAGCGGCTGACCTTCATCCGCAACTGCCGCACGCTGGACATGACCCTG  357
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 121  CAGGCCCATGTGGAGCGGCTGACCTTCATCCGCAACTGCCGCACGCTGGACATGACCCTG  180

Query 358  GACGAAATTCGCAGCCTGCTACGCCTGCGCGACAGCCCCGACGACGCGGACGGCAGCGTC  417
```

```
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 181   GACGAAATTCGCAGCCTGCTACGCCTGCGCGACAGCCCCGACGACGCGGACGGCAGCGTC   240

Query 418   AATGCGCTGATCGACGAGCATATCGAGCATGTTCAGGCGCGGATCGATGGCTTGGTGGCA   477
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 241   AATGCGCTGATCGACGAGCATATCGAGCATGTTCAGGCGCGGATCGATGGCTTGGTGGCA   300

Query 478   TTGCAGGAGCAGCTGGTGGAGCTGCGGCGGCGCGACAACGCGCAGGGGAGTGAATGCGCG   537
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 301   TTGCAGGAGCAGCTGGTGGAGCTGCGGCGGCGCGACAACGCGCAGGGGAGTGAATGCGCG   360

Query 538   ATCTTGCAGCAACTGGAGACAAACGGGGCGGTATCGGTACCGGATACCGAACATTCCCAT   597
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 361   ATCTTGCAGCAACTGGAGACAAACGGGGCGGTATCGGTACCGGATACCGAACATTCCCAT   420

Query 598   GTGGGGCGGAGTCACGGGCAT   618
            ||||||||||||||||||||
Sbjct 421   GTGGGGCGGAGTCACGGGCAT   441
```

7.2.3.2.4.PbrR(Pp)-C77D/C112D DNA gene(Reverse primer)(Length: 444)

Score = 815bits (441), Expect = 0.0, Identities = 441/441 (100%), Gaps =0/441 (0%), Strand=Plus/Minus

```
Query 25    ATGCCCGTGACTCCGCCCCACATGGGAATGTTCGGTATCCGGTACCGATACCGCCCCGTT   84
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 441   ATGCCCGTGACTCCGCCCCACATGGGAATGTTCGGTATCCGGTACCGATACCGCCCCGTT   382

Query 85    TGTCTCCAGTTGCTGCAAGATCGCGCATTCACTCCCCTGCGCGTTGTCGCGCCGCCGCAG   144
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 381   TGTCTCCAGTTGCTGCAAGATCGCGCATTCACTCCCCTGCGCGTTGTCGCGCCGCCGCAG   322

Query 145   CTCCACCAGCTGCTCCTGCAATGCCACCAAGCCATCGATCCGCGCCTGAACATGCTCGAT   204
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 321   CTCCACCAGCTGCTCCTGCAATGCCACCAAGCCATCGATCCGCGCCTGAACATGCTCGAT   262

Query 205   ATGCTCGTCGATCAGCGCATTGACGCTGCCGTCCGCGTCGTCGGGGCTGTCGCGCAGGCG   264
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  261   ATGCTCGTCGATCAGCGCATTGACGCTGCCGTCCGCGTCGTCGGGGCTGTCGCGCAGGCG   20

Query 265   TAGCAGGCTGCGAATTTCGTCCAGGGTCATGTCCAGCGTGCGGCAGTTGCGGATGAAGGT   324
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  201   TAGCAGGCTGCGAATTTCGTCCAGGGTCATGTCCAGCGTGCGGCAGTTGCGGATGAAGGT   14

Query 325   CAGCCGCTCCACATGGGCCTGGGTGTACAACCGGTAGTTGCCCTCGCTGCGCGCCGGCTC   384
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 141   CAGCCGCTCCACATGGGCCTGGGTGTACAACCGGTAGTTGCCCTCGCTGCGCGCCGGCTC   82

Query 385   TGGCAGCAGGTTTTCACGCTCGTAGTAGCGGATGGTTTCCACCGCGCAGTCGGTGGCTTT   444
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 81    TGGCAGCAGGTTTTCACGCTCGTAGTAGCGGATGGTTTCCACCGCGCAGTCGGTGGCTTT   22

Query  445  GGCCAGTTCTCCGATCTTCAT   465
            ||||||||||||||||||||
Sbjct  21   GGCCAGTTCTCCGATCTTCAT   1
```

7.2.3.2.5.PbrR(Pp)-C77D/C112D/C119D DNA gene (Forward primer) (Length: 444)

Score = 381bits (206), Expect = 6e-110, Identities = 215/219(98%), Gaps =1/219 (0%), Strand=Plus/Plus

```
Query 364  GACGCGGCCGGCAGCGTCAATGCGCTGATCGACGAGCATATCGAGCATGTTCAGGCGCGG  423
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 223  GACGCGGACGGCAGCGTCAATGCGCTGATCGACGAGCATATCGAGCATGTTCAGGCGCGG  282

Query 424  ATCGATGGCTTGGTGGCATTGCAGGAGCAGCTGGTGGAGCTGCGGCGGCGCG-CCACGCG  482
           |||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Sbjct 283  ATCGATGGCTTGGTGGCATTGCAGGAGCAGCTGGTGGAGCTGCGGCGGCGCGACAACGCG  342

Query 483  CAGGGGAGTGAAGCCGCGATCTTGCAGCAACTGGAGACAAACGGGGCGGTATCGGTACCG  542
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 343  CAGGGGAGTGAAGACGCGATCTTGCAGCAACTGGAGACAAACGGGGCGGTATCGGTACCG  402

Query 543  GATACCGAACATTCCCATGTGGGGCGGAGTCACGGGCAT  581
           |||||||||||||||||||||||||||||||||||||||
Sbjct 403  GATACCGAACATTCCCATGTGGGGCGGAGTCACGGGCAT  441
```

7.2.3.2.6.PbrR(Pp)-C77D/C112D/119D DNA gene (Reverse primer)(Length: 444)

Score = 154 bits (83), Expect = 2e-42, Identities = 85/86(99%), Gaps =0/86(0%), Strand=Plus/Minus

```
Query 1    TCCTGCAATGCCACCAAGCCATCGATCCGCGCCTGAACATGCTCGATATGCTCGTCGATC  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 308  TCCTGCAATGCCACCAAGCCATCGATCCGCGCCTGAACATGCTCGATATGCTCGTCGATC  249

Query 61   AGCGCATTGACGCTGCCGGCCGCGTC  86
           ||||||||||||||||| |||||||
Sbjct 248  AGCGCATTGACGCTGCCGTCCGCGTC  223
```