IDENTIFICATION OF METAL RESISTANCE GENES IN A STRAIN OF *Enterobacter cloacae*

By

Venkataramana Konda

Submitted in Partial fulfillment of the Requirements

for the degree of

Masters of Science

in the

Chemistry

Program

YOUNGSTOWN STATE UNIVERSITY

AUGUST 2008

IDENTIFICATION OF METAL RESISTANCE GENES IN A STRAIN OF *Enterobacter cloacae*

Venkataramana Konda

I hereby release this thesis to the public. I understand that this thesis will be made available from the OhioLINK ETD Center and the Maag Library Circulation Desk for public access. I also authorize the University or other individuals to make copies of this thesis as needed for scholarly research.

Signature:

Venkataramana Konda, Student Date

Approvals:

Dr. Jonathan J. Caguiat, Thesis Advisor Date

Dr. Timothy R. Wagner, Committee Member Date

Dr. Michael A. Serra, Committee Member Date

Peter J. Kasvinsky, Dean of School of Graduate Studies & Research Date

ABSTRACT

A multi-metal resistant strain of *Enterobacter cloacae* (*E. cloacae*) grows when exposed to toxic salts of mercury, cadmium, zinc, copper and selenite. In general, bacteria respond to toxic metal concentrations using efflux mechanisms, metal transformation (reduction and oxidation), and sequestration. Transposon mutagenesis was used to generate five selenite sensitive, two zinc sensitive and three cadmium sensitive strains of *E. cloacae*. DNA sequencing of the mutagenized genes suggested that a polyphosphate kinase, a sporulation domain protein, a Lon protease and Type-II Secretion protein may be involved in selenite resistance. In addition, a P-type ATPase may be involved in Zn resistance. The sporulation domain protein, tyrosine recombinase and Lon protease may be expressed in response to selenite-induced oxidative stress, the polyphosphate kinase may be involved in selenite reduction and processing and the type-II Sec protein may be involved in selenite efflux. The P-type ATPase may be involved in mercury/cadmium/zinc efflux. Finally, the sequence of two cloned PCR fragments indicated that the *E. cloacae* strain contains genes for copper and mercury resistance. By studying metal-resistance mechanisms, it may be possible to develop strategies to clean metal-contaminated waste sites.

ACKNOWLEDGEMENTS

I am very fortunate to have great parents, Mr. Ramulu and Mrs. Kousalya, whose blessings, love, support and encouragement made me to pursue my higher education and achieve great success in life. I am also thankful to my loving sisters, brother-in-laws and nephews who stood with me in every step I have taken.

I would express my deep gratitude to my thesis advisor, Dr. Jonathan Caguiat for all the support and encouragement he has given to me at all times. I really thank him for all his guidance throughout my academic work and appreciate him for giving me an opportunity to work for him. He created a sound working ambience in the laboratory which encouraged me to do my research work efficiently.

I also thank all my research partners, Danielle DeChant, Misti Mraz, Christi Mraz and Frank Heinselman for helping in my research work by performing mutagenesis technique. I am greatly thankful to Cherise Benton for her help in doing cloning and sequencing of *pco* and *mer* genes which were a part of my project. I extend my thanks to Ed Budde for doing all the sequencing reactions in Department of Biological Sciences.

I heartfully appreciate my friends, Pavan, Ravikiran and Kiran for their mental support and care given during all these years of study. I express my immense thanks to Ramnath for his friendship, mentor support, assistance and calming presence. Ram was always there by leading me in the right track of life. My appreciation also extends to my roommates, Arjun, Srinivas, Sandeep, Varun and Lok for their friendly advices.

I would also thank my friends Sangeetha, Pratima, Samata, Swarnalatha, Anand, and Phani for their help during my academic work.

Finally, I would also like to appreciate the support and guidance of my committee members, Dr. Timothy Wagner and Dr. Michael Serra. I express my gratitude to chemistry department and graduate school of Youngstown State University for funding me through out my course of study. I also thank Department Chair, Dr. Daryl Mincey and Secretary, Kris for helping me in some of the school activities.

TABLE OF CONTENTS

TITLE PAGE ………………………………………………………………………...…..i

SIGNATURE PAGE ………………………………………………………………....ii

ABSTRACT ……………………………………………………………………….…iii

ACKNOWLEDGEMENTS………………………………………………...……...iv-v

TABLE OF CONTENTS ………………………………………………….......vi-vii

LIST OF FIGURES ………………………………………………………..viii-ix

LIST OF TABLES ……………………………………………………………………x

LIST OF SYMBOLS AND ABBREVIATIONS…………………………………xi-xiv

CHAPTER I: INTRODUCTION……………………………………………………..1-4

1.1. Background behind Oakridge, Y-12 plant……...……………………...2

1.2. *Stenotrophomonas maltophilia* strain……………………..……….2-3

1.3. *Enterobacter cloacae* SLD1a-1 strain……………………………….3-4

CHAPTER II: MECHANISMS OF METAL RESISTANCE BACTERIA……...5-26

2.1. Selenium…………………………………………………………6-12

2.1.1. Transport of selenium……………………………..………8-10

2.1.2. Selenocysteine…………...……………………………...10-11

2.1.3. Toxicity of selenium……………...……………………….11-12

2.2. Mercury……………………………………………..………....12-16

2.2.1. Enzymatic reduction of Hg^{2+} to Hg^{0}……………………..…13-16

2.3. Zinc……………………………………………………………......16-20

2.4. Cadmium…………………………………………………….....20-21

2.5. Copper………………………………………………………………21-26

CHAPTER III: HYPOTHESIS…………………………………………………..…27-28

CHAPTER IV: MATERIALS AND METHODS……………………………...29-42

4.1. Bacterial Strains……………………………………………………...30

4.2. Metals……………………………………………………………………30
4.3. Media Preparation……………………………………………………..30-31
4.4. Transformation…………………………………………………...31-32
4.5. Electroporation……………………………………..…………32-33
4.6. Transposon Mutagenesis…………………………………………33-34
4.7. Screening for Metal Sensitive Mutants by Replica Plating…………..34
4.8. Purification of genomic DNA……………………………………34-35
4.9. DNA purification…………………………………………………35-36
4.10. Enzyme digestion for Gene Rescue…………………………………36
4.11. Enzyme Digestion……………………………………………...36-37
4.12. DNA Ligation……………………………………………………...37
4.13. DNA Concentration Determination…………………………….....37
4.14. DNA sequencing…………………………………………….…37-38
4.15. BLAST Analysis…………………………………………………..38
4.16. Polymerase Chain Reaction…………………………………………39
4.17. Primer Design……………………………………………….…39-41
4.18. Agarose Gel Electrophoresis………………………………….…..41-42

CHAPTER V: RESULTS……………………………………………………43-109

CHAPTER VI: DISCUSSIONS………………………………………………110-118
6.1. Lon protease (La protease)…………………………………………...111
6.2. Sporulation Domain protein ……………………………….……111-112
6.3. Polyphosphate kinase ………………………………………....112-113
6.4. P-type ATPases …………………………………………….......113-114
6.5. Tyrosine recombinase …………………………..…………114-115
6.6. Type II secretion protein………………….…………………...…115-116
6.7. Acyl transferase…………………………………….…………116
6.8. Copper and Mercury resistance genes, *pcoA* and *merR*…………..……117

CHAPTER VII: REFERENCES……………………………………………..119-126

LIST OF FIGURES

Figure 1. Blunt end digestion of EZ-Tn*5* mutant genomic DNA……………………….48

Figure 2. Digestion of the transformed DNA……………………………………………51

Figure 3. Digestion of the transformed DNA……………………………………………54

Figure 4. Digestion of the transformed DNA……………………………………………56

Figure 5. Feature map of A3A and L30 mutant……………………………………….…60

Figure 6. Feature map of A3A mutant…………………………………………………...62

Figure 7. Nucleotide sequence of A3A F + R mutant…………………………………...64

Figure 8. Blast result of A3A mutant…………………………………………………….65

Figure 9. Feature map of L31 Mutant……………………………………………………...66

Figure 10. Nucleotide sequence of L31 R + F + R2……………………………….68-69

Figure 11. Blast result of L31 R + F + R2 mutant………………………………………70

Figure 12. Feature map of 8HB mutant…………………………………………………71

Figure 13. Nucleotide sequence of 8HB F + R + R2 mutant………………………...73-74

Figure 14. Blast result of 8HB F + R + F2 mutant………………………………………75

Figure 15. Feature map of F24 R + F mutant………………………….…………...76

Figure 16. Nucleotide sequence of F24 F + R mutant…………………………………...78

Figure 17. Blast result of F24 F + R mutant…………………………...…………...79

Figure 18. Feature map of D21 F +R mutant……………………………………………80

Figure 19. Nucleotide sequence of D21 F+R mutant……………………………...82-84

Figure 20. Blast result of D21 F+R mutant……………………………………………...85

Figure 21. Feature map of 6B mutant……………………………………………………86

Figure 22. Nucleotide sequence of 6B mutant……………………………….…88-89

Figure 23. Blast result of 6B mutant……………………………………………….90

Figure 24. Feature map of Q17 mutant..………………………………...………91

Figure 25. Nucleotide sequence of Q17 R + F mutant………………………………..93

Figure 26. Blast result of Q17 mutant………………………………………………94

Figure 27. Feature map of F34 F…………………………..………………………95

Figure 28. Nucleotide sequence of F34 F mutant…………………...………………...97

Figure 29. Blast result of F34 F mutant………………………………..……………..98

Figure 30. Feature map of *pcoR*…………………………………………..……100

Figure 31. Nucleotide sequence of *pcoR*…………………………………………….101

Figure 32. Blast result of *pcoR*……………………………………………………102

Figure 33. Feature map of *pcoF*……………………………………………...…103

Figure 34. Nucleotide sequence of *pcoF*…………………………………………….104

Figure 35. Blast result of *pcoF*……………………………………………………105

Figure 36. Feature map of *merR* operon……………………………………………..106

Figure 37. Nucleotide sequence of *merR*…………………………………………….108

Figure 38. Blast result of *merR* …………………………………………………...109

LIST OF TABLES

Table 1. Primers used in sequencing reactions……………………………………….40-41

Table 2. Mutants obtained by transformation……………………………………………45

Table 3. Concentrations and sizes of the transformed DNA…………………………….58

LIST OF SYMBOLS AND ABBREVIATIONS

EDTA……………………………………………..Ethylene Diamine Tetra Acetic Acid

HCl…………………………………………………………...Hydrochloric Acid

sec……………………………………………………………………….Seconds

µL………………………………………………………………………...Microliter

µM……………………………………………………………………...Micromolar

$CaCl_2$……………………………………………………………Calcium Chloride

UV……………………………………………………………………...Ultra Violet

NaCl……………………………………………………………Sodium Chloride

$MgCl_2$…………………………………………………...Magnesium Chloride

BSA………………………………………………………Bovine Serum Albumin

DTT……………………………………………………………...Dithioththreitol

mL ………………………………………………………………………milliliters

ssDNA………………………………………Single Stranded Deoxyribonucleic Acid

nm……………………………………………………………………..nanometer

dH_2O…………………………………………………………Deionized water

H_2O……………………………………………………………………...Water

Tris…………………………………………………Tris(hydroxymethyl)aminomethane

v/v…………………………………………………………..volume per volume

dNTP…………………………………………...……deoxynucleotidetriphosphate

TBE…………………………………………………..………Tris-Borate-EDTA

Na_2-EDTA……………………………………………………………….Sodium EDTA

μg………………………………………………………………………….Microgram

μF………………………………………………………………………….Microfarad

kV………………………………………………………………………….kilovolts

Ω…………………………………………………………………………...Ohms

dG……………………………………………………………..Free energy of oligo

% GC………………………………………………percentage of G and C in oligo

mM………………………………………………………………………millimolar

mg………………………………………………………………………...milligram

M………………………………………………………………………...molar

$MgSO_4$ ……………………………………………………………Magnesium sulfate

mL…………………………………………………………………..…milliliters

min……………………………………………………………………….minutes

rRNA……………………………………………...ribosomal ribonucleic acid

SeO_3^{2-}..……………………………………………………………………..Selenite

SeO_4^{2-}…………………………………………………………………..Selenate

Se…………………………………………………………………….....Selenium

Zn…………………………………………………………………………zinc

Cd…………………………………………………………………………cadmium

Cu…………………………………………………………………………copper

pmol……………………………………….……………………......picomoles

LDL……………………………………………………..…Low Density Lipoproteins

ECG…………………………………………………………………..Electrocardiogram

NO_3^- ………………………………………………………………….........Nitrate

DMSe………………………………………………………………...Dimethyl Selenide

DMDSe…………………………………………………………….Dimethyl Diselenide

E. coli………………………………………………………………..*Escherichia coli*

FDH…………………………………………………………Formate dehydrogenase

tRNA………………………………………………………………………transfer RNA

γ ………………………………………………………………………….…gamma

β…………………………………………………………………………………...beta

ATP…………………………………………………………..Adenosine triphosphate

H_2O_2 ……………………………………………………………...Hydrogen peroxide

O_2^- ……………………………………………………….....................Superoxide

GSH ……………………………………………………….................Glutathione

DNA…………………………………………………………...deoxyribonucleic acid

RNA ……………………………………………………….......ribonucleic acid

NADP...nicotinamide adenine dinucleotide phosphate

NADH...nicotinamide adenine dinucleotide dehydrogenase

CH_3Hg…………………………………………………………… methyl mercury

CO_2 …………………………………………………………………Carbondioxide

CH_4 ……………………………………………………………………......Methane

% ……………………………………………………………………………......percentage

V………………………………………………………………………………......volts

kb……………………………………………………………………….………kilobase

bp ………………………………………………………………………………basepairs

ATCC…………………………………………………......American Type Culture Collection

Chapter I: Introduction

1.1. Background behind Oakridge, Y-12 plant

The Y-12 plant at Oak Ridge, Tennessee, a part of the Manhattan project has played a major role in the production of nuclear weapons for the past 60 years. It is situated at the eastern end of the Oakridge Reservation, adjacent to the city of Oakridge, in Anderson County, Tennessee. It is now under the control of the U.S. Department of Energy and was originally constructed in 1943 with the mission of separating fissionable uranium isotopes (U-235) from natural uranium using an electromagnetic process.[1] During World War II it processed uranium to make the first atomic bomb, which was dropped on Hiroshima, Japan in 1945. During the Cold War in the late 1950s, the Y-12 plant focused on processing lithium to make hydrogen bombs. The mission of the Y-12 plant changed from nuclear production to the maintenance and storage of nuclear weapons later on. At present, the Y-12 plant is involved in the receipt, storage and protection of nuclear materials. Y-12 is considered to be an integral part of science based stockpile stewardship, along with Fort Knox, for enriched uranium.[99]

A large amount of heavy metal, mercury (11,000,000 kg) was used as a major component in the lithium separation process involved in making hydrogen bombs. During this process, about 330,000 kg of mercury were assumed to be lost to the environment, contaminating the nearby East Fork Poplar Creek (EFPC).[2, 3] The U.S. Department of Energy has taken up many programs to reduce the mercury concentrations in the water released into the EFPC.[64]

1.2. *Stenotrophomonas maltophilia* strain

Stenotrophomonas maltophilia Oak Ridge Strain O2 (ATCC # 53510) is an aerobic, non fermentative gram-negative bacterium that was isolated from East Fork

Poplar Creek of Y-12 plant. This bacterium grows in toxic levels of metal salts such as mercury, cadmium, zinc, copper and selenite. *S. maltophilia* ORO2, a gamma proteobacterium is capable of reducing selenite (SeO_3^{2-}) to nontoxic elemental selenium.[4] In the process of using 16s rRNA sequencing to identify other metal resistant bacteria from East Fork Poplar Creek, we sequenced a segment of 16s rRNA from our working strain of *S. maltophilia* ORO2 and discovered that it was actually similar to a strain of *Enterobacter*, not *Stenotrophomonas.* Biochemical tests of this strain revealed that it was a strain of *Enterobacter cloacae.* This strain also exhibits resistance to different metal salts of mercury, cadmium, zinc, copper and selenite.

1.3. *Enterobacter cloacae* SLD1a-1 strain

E. cloacae SLD1a-1 is a facultative anaerobic bacterium isolated from Se contaminated water of the San Joaquin Valley, California capable of reducing both selenate (SeO_4^{2-}) and selenite (SeO_3^{2-}) to elemental selenium.[5] *E. cloacae* SLD1a-1 can reduce selenate to elemental selenium through membrane-bound molybdenum dependant selenate reductase under aerobic conditions. Elemental selenium gets deposited at cytoplasmic membrane and then expels out of the cell.[6] This organism is also capable of volatilizing selenium in the presence of selenite (SeO_3^{2-}) to form dimethylselenide (DMSe) apart from reducing selenite and selenate to elemental selenium.[7]

E. cloacae has the potential to be used in bioremediation to remove selenium oxyanion contamination. Selenium oxyanions are transformed to insoluble elemental selenium (Se^0) by bio-reduction process. It may be further converted to volatile forms such as dimethylselenide by methylation. Much more has been reported on the number of microorganisms reducing selenite than on the number of microorganisms that reduce

selenate.[8] Selenite reduction is believed to be carried out by the membrane and periplasmic-bound nitrate and selenate reductase enzymes, even though nitrate reductases are poor reducers of selenate.[6]

Chapter II: Mechanisms of metal resistance bacteria

2.1. Selenium

Selenium is available as a silvery metallic allotrope or red amorphous powder. It is found in sulfide ores bearing atomic number of 34 and is a naturally occurring trace element belonging to group VI A of the periodic table.[9] Although it is essential to living things, it is considered toxic at higher concentrations. It is characterized as a metalloid having the properties of both a metal and nonmetal. Selenium occurs in four oxidation states: selenate [Se (VI)], selenite [Se (IV)], elemental selenium [Se (0)] and selenide [Se (-II)]. Selenate (SeO_4^{2-}) and selenite (SeO_3^{2-}) are toxic water soluble species that bioaccumulate and are found mostly in seleniferous soils and agricultural drainage water.[10] Elemental selenium (Se^0) is essentially nontoxic and insoluble in water.[11] Selenide is also toxic and reactive but can be oxidized to elemental selenium. At certain times, selenide is substituted for sulfur found mostly in sulfide minerals and pyritic coal deposits.[10]

Selenium is an essential nutrient for all living things.[12] Selenium, in combination with vitamin E and sulphur containing amino acids, helps in preventing many nutritional deficiency diseases.[13] It is considered to be an integral part of glutathione peroxidase (GSH-Px) possessing catalytic and structural functions. Several selenoproteins have been identified as having a major role in the treatment of many human diseases. Selenium in the form of selenoprotein such as glutathione peroxidase is used as an antioxidant in the treatment of carcinogenesis and heart diseases. Glutathione peroxidase is used as four types, plasma glutathione peroxidase, phospholipid hydro peroxide and gastrointestinal glutathione peroxidase in treating carcinogenesis and heart diseases-GSHPx-1. Its importance in chronic degenerative diseases in humans was also elucidated.[14]

In certain cases, selenium is used to treat diseases but deficiency of selenium also results in diseases as described below. Selenium deficiency causes ECG (Electrocardiogram) abnormalities, myocardial disease and mulberry heart disease in lambs and pigs.[14] Selenium deficiency is also a major cause of endemic fatal cardiomyopathy (Keshan disease).[15] Selenium deficiency leads to the abnormalities in many functions of the liver, brain, heart, striated muscle, pancreas and genital tract. Selenium is also essential for normal immune function and its deficiency leads to reduced T-cell count and impaired lymphocyte proliferation and responsiveness. Lower concentrations of selenium in the blood lead to coronary heart disease. Deficiency of glutathione peroxidases increases the production rates of hydrogen peroxide and superoxides leading to atherogenesis. Low selenium concentrations in rats have been shown to increase LDL-cholesterol, decrease the production of aortic prostacyclin and increase platelet aggregation. Biosynthesis of prostaglandins, enhancement of thromboxane content of platelets has also been affected by selenium deficiency. In humans, selenium deficiency leads to atherosclerosis.[14]

Microorganisms play an important role in the global cycling of selenium through oxidation, reduction, methylation and demethylation. Of these, *Enterobacter cloacae*, a gram negative organism is involved in the reduction of selenate and selenite to elemental selenium using NO_3^- and SeO_4^{2-} as terminal electron acceptors under anaerobic conditions. Washed-cell suspensions from this bacterium show that it uses membrane bound reductases for the reduction reaction.[8] Elemental selenium can be oxidized into selenate or selenite mainly in soil sediments by biotic process. The reduced selenium can be further reduced into selenide (Se^{2-}) and can be methylated to form dimethylselenide

(DMSe) or dimethyl diselenide (DMDSe) which exist as volatile aqueous species.[8, 16] The biomethylation of selenium from oxyanions or from organic selenium compounds such as selenocysteine, selenocystine and selenomethionine is useful in detoxification and removal of selenium from selenium contaminated sites.[17] *Enterobacter cloacaea* SLD1a-1 is the first reported organism to methylate selenium.[18]

Naturally occurring selenium is also of major concern in the phosphate mining sites of US Western Phosphate Resource Area, Idaho, U.S.A.[19] Selenium is present in higher concentrations in sediments and soils which is a serious threat to the environment. It is also the major source of contamination in many anthropogenic activities such as irrigated agriculture, fossil fuel combustion, petroleum refining and mining operations. Thus, many bioremediation and geochemical process have been implemented to remove Se oxyanions from the seleniferous soils and sediments. Many organisms are involved in the process of removing toxic selenium compounds by the process of reduction, oxidation and methylation. Of these processes, methylation of selenium leading to its volatilization is considered to be the most prominent biotechnology method for the complete removal of selenium from the contaminated sites. The other process involved in removing selenium from drainage water is by gravity and filtration.[10] Oxidation and reduction of selenium have an important impact on the fate and transport of selenium by microorganisms.

2.1.1. Transport of selenium

The specific pathway of selenium import into the cell for protein incorporation is unknown and unclear. Selenium gets incorporated into specific tRNA molecules in *E. coli* and formate dehydrogenases (FDH). Some evidence shows that selenium may be

imported as selenite through the sulfate transport system during the cysteine biosynthesis pathway.[20]

The other pathway by which selenate enters into the cell is through the sulfate permease system by the use of *cysA*, *cysU and cysW* genes that have been observed in *E. coli*. Any alterations in these genes confer selenate resistance. Selenite also uses the same pathway, but because selenite uptake through the sulfate permease transport system has not been inhibited completely, there is a doubt about the existence of an alternate carrier system for selenite.[21] But Muller *et al.* explained that selenite may use sulfate permease transport system at higher concentrations with the help of *cysA* gene.[20] Selenium uptake is a complicated pathway because selenium is required for cell growth but is also toxic. There must be at least two end results under toxic selenite conditions: 1) detoxification and 2) incorporation. In certain cases, selenium replaces sulfur in some proteins and other biomolecules because it is incorporated into enzyme systems responsible for sulfate metabolism.[22] Selenate and sulfate also use the same transport system in a few microorganisms such as *Candida utilis*,[23] *Salmonella typhimurium*,[24] *E. coli*,[25] *Saccharomyces cerevisiae*.[26]

The rumen microorganism, *Selenomonas ruminantium,* transports selenium into the cell in the form of cysteine and methionine sulphur amino acid analogs, selenocysteine and selenomethionine. This organism cannot transport selenate or sulphate.[27] Although the selenium specific pathway is not clearly understood, a novel gene product, *gutS* of *E. coli* is a 43k-Da protein that appears to be associated with the permease and membrane transport proteins involved in selenite metabolism.[28]

The transport of selenium into proteins occurs through the sulphate transport mechanism, but higher affinity was shown to sulphate than selenate and selenite as shown by competitive uptake inhibition. The transport system observed in *Salmonella typhimurium* for selenite indicated that there is a possibility of specific mechanism for the transport system of selenite implying that there might exists a separate transport system for selenite.[28] *Rhodobacter sphaeroides* may transport selenite through a polyol ABC transporter located in its cytoplasmic membrane. The reduced selenite enters through the plasma membrane and cell wall and then accumulates in the cytoplasm.[29, 30]

Selenium oxyanions, selenate and selenite are reduced to selenide utilizing the sulfate reduction pathway and then are incorporated into the amino acid, cysteine as selenocysteine which is then converted to selenomethionine. This process requires cystathione γ-synthase (*metB*), β-cystathionase (*metC*) and methionine synthase (*metE* and *metH*). The major form of selenium incorporation into proteins occurs in form of selenocysteine.[21]

2.1.2. Selenocysteine

Selenocysteine, the 21st amino acid, is essential in the active site of redox proteins such as *E. coli* formate dehydrogenase (FDH). Selenocysteine is also present in eukaryotic glutathione peroxidases and thioredoxin reductases. The incorporation of selenocysteine into proteins is directed by a UGA stop-codon and uses the genes *selA*, *selB*, *selC* and *selD*.[12, 20, 31] Monoselenophosphate synthetase is an enzyme that plays a crucial role in selenium metabolism. It is involved in the synthesis of selenocysteyl-tRNA formed by the reaction of the pyridoxal phosphate-dependant enzyme, selenocysteine synthase (*selA* gene) and seryl-tRNA. Monoselenophosphate is the product of

selenophosphate synthetase (*selD* gene) formed by transferring the γ-phosphate moiety of ATP to selenide. Selenophosphate synthetase is essential for the insertion of selenocysteine into proteins. Mutation in *selD* gene prevents the incorporation of selenium into formate dehydrogenases and tRNA. The free amino acid, selenocysteine esterifies to $tRNA^{cys}$ by not binding to *selC* tRNA, and randomly gets incorporated into proteins replacing cysteine. The gene product, *sel*B is a GTP-dependant translational factor that is an alternative to elongation factor EF-Tu, which transports selenocysteyl-$tRNA^{sec}$ to the ribosome required for the translation. The sequences in the N-terminal domain region of EF-Tu have similar sequences related to the SELB protein, a larger protein than EF-Tu.[12]

2.1.3. Toxicity of Selenium

Oxygen is primarily responsible for selenite-sensitivity in bacteria. Highly toxic substances such as hydrogen peroxide, H_2O_2 and superoxide, O_2^- are produced by the reaction of selenite with glutathione peroxidases causing damage to cell membranes and DNA. The oxidative stress caused by these oxygen species is responsible for the toxicity of selenite. Oxidative stress can also be partially overcome by the synthesis of the proteins such as heat shock proteins, thioredoxin and an iron-containing superoxide dismutase, FeSOD.[30]

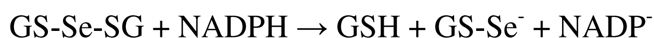
Glutathione, the most abundant thiol found in eukaryotic cells and cyanobacteria has a major role in selenium metabolism. The reduction of selenite to elemental selenium is carried out by the reaction of selenite with sulfhydryl groups of thiol containing molecules such as glutathione with the formation of many different selenium

intermediates, such as selenodiglutathione (GS-Se-SG), unstable selenopersulfide of glutathione (GS-SeH), and hydrogen selenide (HSe^-).[30]

The toxicity of selenite can be explained by a series of reactions carried out by selenite with glutathione. Glutathione (GSH) reacts with selenite to form selenotrisulfides (GS-Se-SG) according to the following reaction:

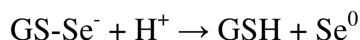
$$4GSH + H_2SeO_3 \rightarrow GS\text{-}Se\text{-}GS + GSSG + 3H_2O$$

The selenotrisulfide also named selenodiglutathione gets reduced to a selenopersulfide of glutathione ($GSe\text{-}Se^-$) using glutathione reductase and NADPH as an electron acceptor.



$$GS\text{-}Se\text{-}SG + NADPH \rightarrow GSH + GS\text{-}Se^- + NADP^-$$

Selenopersulfide of glutathione is unstable and decays to elemental selenium and reduced glutathione.[32]



$$GS\text{-}Se^- + H^+ \rightarrow GSH + Se^0$$

2.2. Mercury

Mercury is a transition element of the periodic table with an atomic number of 80.[9] It is a heavy, silvery d-block metal, represented by the symbol Hg that exists naturally in the environment. Mercury is most widely distributed in the environment in three different forms as 1) elemental or metallic mercury (Hg^0), as 2) inorganic mercury (Hg^{2+}) or as 3) organic mercury. Mercury is present in thermometers as metallic mercury, in dental amalgam fillings as inorganic mercury, and in fish, mostly as methyl mercury.[33] Mercury exists, as a highly toxic vapor and as a less toxic liquid. Because of its volatile nature, mercury is a major source of contamination in air, water and solid wastes. It enters into the environment with evaporation at the start the global mercury cycle. The

major source of elemental air-borne mercury is fossil fuel burning and municipal waste incineration. Mercury is also released into nature from water, sea or land surfaces. Industrial plants such as chlor-alkali plants and scrap metal processing facilities release mercury waste into water ways. Mercury pollution is considered to be a great concern in the Great Lake Regions of United States due to its ability to bioaccumulate in the aquatic food chain.[34] Once it is released into the environment, it can reside for long periods of time in the atmosphere and eventually appears as toxic organic methyl mercury released by bacteria. Accumulation of methyl mercury in fish poses a potential harm to humans. It exerts toxic effects on the central nervous system.[33]

The mercury resistance operon (*mer*) plays a major role in the global cycling of mercury. As these chemicals are toxic to all living organisms, bacteria develop mechanisms of resistance. The genes responsible for resistance are located on plasmids and transposons of gram-negative and gram-positive bacteria.[35, 36] Mercury binding, transport, and reducing proteins are involved in the resistance mechanisms.[36]

2.2.1. Enzymatic reduction of Hg^{2+} to Hg^0:

The mercury resistance genes (*mer*) of transposon Tn*21* from *Shigella flexneri* has been studied in detail. It occupies about 8 kb of the 94 kb plasmid R100, an antibiotic-resistance plasmid that was found in Japan in 1956. The operon consists of five structural genes, *merT*, *merP*, *merC*, *merA*, *merD* and a regulatory transcriptional gene *merR*.[37] Another *mer* operon is found in transposon Tn*501* which was isolated from *Pseudomonas aeruginosa* and consists of a regulatory gene, *merR*, and the structural genes, *merT*, *merP*, *merA* and *merD*.[38] The genes *merT* and *merP* make up a transport system bringing

the extracellular toxic metal (Hg^{2+}) into the cell through cytoplasmic membrane. The MerP protein binds Hg^{2+} in the periplasm and transfers it to a pair of cysteine residues in the inner membrane protein, MerT via a redox rapid exchange mechanism with the other two cysteine residues present on MerT. Mutations of *merT* and *merP* led to a decrease in mercury resistance.[39]

MerC and MerF are alternative inner membrane transport proteins of the operon. MerE is an additional membrane transport protein of unknown function. Highly toxic mercury (Hg^{2+}) is reduced to monoatomic mercury vapor by MerA, the mercuric reductase, a cytosolic flavin disulfide oxidoreductase which uses NADPH as a reductant. This is the last step in the bacterial detoxification of mercury. The reduced and nontoxic metal is then released into the cell cytoplasm as volatile Hg^0.[37]

Detoxification of mercury is carried out by two enzymes namely, oraganomercurial lyase (MerB) and mercuric reductase (MerA). Organomercurial lyase is involved in the cleavage of C-Hg bonds of organomercurial compounds and thereby releasing toxic Hg^{2+}. Mercuric Reductase, MerA is involved in the bacterial detoxification of mercury by catalyzing a 2 electron reduction of Hg^{2+} to Hg^0 by NADPH with the following stoichiometry equation.[35, 40]

$$Hg(SR)_2 + NADPH + H^+ \longrightarrow Hg(0) + NADP^+ + 2RSH$$

The enzyme possesses similar properties to that of pyridine nucleotide disulfide oxidoreductase, lipoamide dehydrogenase and glutathione reductase. The cysteine thiol pairs, $Cys_{558}Cys_{559}$ at the active site of mercuric reductase play an important role in the reduction of mercury.[40] A bacterial strain is said to be broad spectrum resistant if it

possess both organomercurial lyase and mercuric reductase and narrow spectrum resistant if it possesses only mercuric reductase.[41]

MerR is a transcriptional regulatory protein that regulates *mer* operon expression. MerR functions as a dimer and binds to the operator region of the *mer* operon. In the absence of mercury, MerR represses the expression of the *merTPCAD*. In the presence of mercury, MerR binds to mercury, undergoes a conformational change and activates *merTPCAD*. Transcription of *merR* is repressed irrespective of the presence or absence of mercury.[42, 43]

The role of MerD, a small cysteine-rich and low abundance protein is not well understood but may be involved in regulation. It might functions as a coregulator due to its N-terminal amino acid residue similarity to MerR.[43] It binds to *merO*, the *mer* operator region, and restores the MerR repression state when Hg (II) has been removed from the cells environment. Mutations in the *merD* gene increases mercury resistance by two-fold.[44]

There are other mechanisms by which bacteria can modify mercury. Hydroperoxidase catalase, KatG, of *E. coli* can oxidize Hg^0 to Hg^{2+} which then combines with sulfhydryl groups and imino-nitrogen ligands in proteins and other biological molecules.[45] Methylation of mercury is observed in *Desulfovibrio desulfuricans* by the enzyme methyltransferase by transferring a methyl group from methyl-tetrahydrofolate to methylcobalamine.[35]

Microbial CH_3Hg degradation is also carried out using oxidative demethylation (OD). Methanogens and sulfate reducers use this protein to oxidize CH_3Hg to Hg (0), CO_2 and CH_4.[46]

2.3. Zinc

Zinc is available as a bluish-white metal that is brittle at ambient temperatures but malleable at 100 - 150 ^{o}C. It is one of the transition elements of the periodic table and has an atomic number of 30.[9] It is an ubiquitous essential trace element and a cofactor of many enzymes involved in metabolism. It plays a major role in catalysis, in the maintenance of protein structure, and in the regulation of gene expression.[47, 48]

Zinc is an essential trace element for most bacteria and is present in the active site of many bacterial enzymes.[49] However, since an excess of zinc is toxic to living organisms, some bacteria have developed resistance mechanisms of sequestration and of efflux for this metal. In the presence of high concentrations, they must establish a balance or homeostasis of how much is incorporated into cell mass and how much is excluded by the resistance mechanisms.[50]

The zinc transport systems are categorized into primary and secondary transport systems depending upon the energy source used for exporting the metal ions across the membrane. The primary transport systems use an ATP chemical energy source for transporting metal ions across the cytoplasmic membrane whereas the secondary transport systems utilize the energy of an electrochemical gradient.[47]

Resistance, nodulation and division (RND) efflux transporters, P-type ATPases, cation diffusion facilitators, and ATP binding cassette (ABC) transporters also play major roles in the uptake and efflux during Zn^{2+} homeostasis.[49]

The RND system from *Ralstonia metallidurans* consists of proteins, CzcA, CzcB and CzcC which are involved in transporting Zn^{2+}, Co^{2+} and Cd^{2+} from the cytoplasm and periplasm into the growth medium.[49] The efflux system, CzcCBA protects the cytoplasmic membrane and the periplasm from these toxic metals. The heavy metal cations are exported from the periplasm by the action of uptake and a CBA transenvelope efflux system.[51] CzcA is located in the cytoplasm and functions as a cation proton antiporter. CzcB, acts as an acriflavin export pump. CzcC is involved in the formation of CzcABC protein complex by connecting CzcB to the outer membrane. CzcS and CzcR functions as histidine sensors and kinase regulators of the zinc transport expression.[49]

The protein, CzcD, is involved in developing resistance to higher concentrations of Zn^{2+} by the metal cation efflux pump.[49] CzcD is a membrane-bound protein from the gram negative bacterium, *Ralstonia metallidurans* CH34, and is located in the cytoplasmic membrane. It belongs to the cation diffusion facilitator protein family. CzcD is also involved in the expression of $CzcCB_2A$ efflux pump.[52]

ZRC1p and COT1p from *Saccharomyces cerevisiae* are yeast proteins that belong to cation diffusion facilitators (CDF) family. CDF proteins act as a cation efflux pump by catalyzing the accumulation of heavy metals and detoxify cobalt ions by binding to the proteins. These CDF proteins are also assumed to act as a heavy metal buffer using exporting and importing systems to maintain optimal metal concentrations inside the cell.[52]

P-type ATPases are another family of cation transporting membrane proteins involved in metal homeostasis. P-type ATPases are soft metal transporters, possess an ATP binding domain, and are phosphorylated at an aspartate amino acid residue.[49]

The *zntA* gene encodes a P-type ATPase responsible for conferring zinc resistance. This gene is involved in zinc metabolism by catalyzing ATP dependant zinc efflux from *E. coli*. Mutation of this gene in *Proteus mirabilis* causes a defect in swarming which is responsible for the urinary tract pathogenecity along with the combination of zinc homeostasis. *CadA* and *CadC* are other genes belonging to the P-type ATPase family and are located on staphylococcal plasmid pI258 which confers Zn^{2+} and Cd^{2+} resistance. The CadA/C transport system is also present in gram-negative bacterium, *Stenotrophomonas maltophilia*.[49]

ABC transporters are trans-membrane proteins involved in transporting metals into or out of the cell by forming a pore and utilizing ATP as the source of energy. The ZnuABC system contains a set of proteins that belong to the ABC transporter family and are involved in transporting the metals out of the cytoplasm.[47] ZnuA is located outside the cytoplasmic membrane. The periplasmic binding protein of Mn (II) transport system of *Synechocystis* species has an identical ZnuA amino acid residue sequence which makes it efficient for metal ion binding. The other proteins of ABC transporters include the hydrophobic protein, ZnuB, which is a membrane component and ZnuC, which possesses a motif that is similar to the ATPase subunit of ATP transporters. This transport system has a high affinity for Zn (II) uptake.[49]

Zur is a protein which regulates Zn^{2+} uptake systems and has an amino acid residue sequence that is similar to that of the iron uptake regulator, Fur. Zur is distributed

evenly in gram-positive and cyanobacteria and functions as a repressor. The cytoplasmic protein, Zur is more active in the presence of reduced thiols than in the presence of oxidized disulfides. Due to the poor binding of oxidized Zur to Zn^{2+}, it binds to only one of the nine cysteines present in Zur. The function of Zur as a regulatory protein and zinc metabolism varies with the concentration of zinc present in the environment. In *B. subtilis*, the sequence similarity of Zur and Fur, along with the third protein, YqfV may allow it to act as Zur. In some organisms such as *Salmonella* strains, *Klebsiella pneumoniae, Yersinia pestis, Vibrio cholerae, Bordetellapertussis, Caulobacter crescentus, Pseudomonas aeruginosa*, and *Neisseria* strains, the Zur system is similar to Znu and possesses the same regulatory function of Zn^{2+} uptake.[49]

Another transport system having a major role in zinc homeostasis and cell signaling is YiiP, a membrane transport system involved in the export of Zn^{2+}/H^{+} across the inner membrane of *E. coli*.[53]

ziaA and *ziaR* are the genes of *synechocystis* PCC 6803 and confer tolerance to high concentration of zinc. *ziaA* initiates efflux of Zn^{2+} from the cytosol to the periplasm while *ziaR* regulates the expression of *ziaA*. ZiaR is a highly Zn (II) specific sensor that is similar to SmtB found in *synechococcus* species.[54]

In *synechococcus* strain PCC7942, the *smt* operon encodes the genes, *smtA* and *smtB*. They play a major role in maintaining zinc homeostasis when excess zinc is present. The protein, SmtA, is a class II metallothionein that sequesters zinc. Another protein, SmtB, initiates the transcription of *smtA* gene by acting as a trans-acting repressor.[55]

The MerR and SmtB/ArsR family of proteins are metal sensing proteins in prokaryotes. MerR functions as a repressor in the absence of metal and as an activator upon binding to a metal. The SmtB/ArsR family is involved in regulating sequestration or efflux of metal ions in gram-negative bacteria.[56]

ZupT is an additional transport system belonging to the ZIP protein family and is involved in transporting zinc into the cytosol of *E. coli*.[47, 57] PZP1 is another periplasmic Zn (II) metallo-chaperone from *Haemophilus influenza* similar to ZnuA and is involved in zinc uptake. ZraP and YdaE are other *E. coli* metallo-chaperones helpful in binding to zinc under high zinc concentrations. ZitB is another CDF family protein that pumps out zinc.[47]

Zinc fingers are protein domains that participate in eukaryotic metabolism by interacting with DNA, RNA, proteins and lipids. The *E. coli* protein, GatA, possesses a zinc-finger-like structure but acts as a metallothionein by sequestering excess Zn^{2+}.[58]

2.4. Cadmium

Cadmium is a non-essential metal and is recognized as a type I carcinogenic element.[59] It has an atomic number of 48.[9] Cadmium resistance in *Staphylococcus aureus* is regulated by cadmium efflux system encoding *cadA* and *cadB* genes on S. *aureus* plasmid pI258.[47] CadA is a Cd^{2+}/ATPase transporter belonging to the class of P-type ATPases.[60] CadA is an integral membrane protein acting as an electro-neutral antiporter involved in catalyzing the exchange of one Cd (II) for two protons in the cytosol.[47] The gene, *cadA* is responsible for conferring resistance to Cd^{2+} and Zn^{2+}.[60] The *cadB* is another gene located on plasmid pI258, and protects the cell by binding to Cd^{2+}.[61] CadC

is a soluble protein encoding metallo-regulatory repressor protein of the ArsR family involved in the negative regulation of *cad* operon.[47, 61] The *cadR* is a gene in *Pseudomonas aeruginosa* similar to z*ntA* in *E. coli* and encodes the transcriptional regulatory protein, CadR. This protein induces expression in response to Cd^{2+} at its cognate promoter, P_{cadA} and at P_{zntA}, in *E. coli*. CadR/P_{cadA} belongs to the MerR regulatory protein family but the mechanism of action is unknown.[62] Another new gene, *cadD*, is similar to that of the *cadB* gene which confers resistance to cadmium by sequestering toxic Cd^{2+} ions.[60]

In gram negative bacteria, a CBA transport system encoded by three different proteins in a single operon protects the periplasm from the damage being caused by metals. This system acts as a defensive layer protecting the cytoplasm by translocating these metals across the outer membrane. This transport system includes an RND proteins acting as a central pump along with two other components, a membrane fusion protein (MFP) and an outer membrane factor (OMF), which together are involved in export of metal ions, xenobiotics and drugs. A Czc transport system of *Ralstonia metallidurans*, encoded by the genes *czcA*, *czcB* and *czcC*, pumps out cadmium, zinc and cobalt, and is considered to be the best characterized metal CBA transport system.[63]

2.5. Copper

Copper is a reddish, malleable and ductile metal with an atomic number of 29.[9] Copper ions exist in two stages as oxidized Cu (II) and reduced Cu (I).[63] Copper serves as a cofactor in various redox enzymes such as lysyl oxidase, cytochrome *c* oxidase, and superoxide dismutase or dopamine β-hydroxylase.[65] It is also found in multicopper

oxidases, amine oxidase or lysine oxidase which is considered to be active in various processes such as respiration, iron transport, oxidative stress protection, blood clotting and pigmentation. As copper is a redox-active transition metal, it is highly toxic even at low concentrations. Thus, copper homeostasis is needed for copper metabolism because it causes oxidative stress through Fenton-like reactions by generating superoxide or other reactive oxygen species.[63] For this purpose, copper resistance genes that are involved in copper homeostasis have been identified and studied.[65]

Copper homeostasis in the Gram-negative bacterium, *E. coli* and the gram-positive bacterium, *Enterococcus hirae*, has been studied and is mediated by four genes: *copA, copB, copY* and *copZ* which together make up a *cop* operon.[66] The *cop* operon of *Pseudomonas syringae* encodes the genes, *copA*, *copB*, *copC* and *copD* on plasmid pPT23D which is responsible for copper resistance.[67] In *E. coli*, chromosomal and plasmid-borne resistance genes involved in copper homeostasis have been identified. The proteins that are involved in copper homeostasis and transport include CueO (multicopper oxidase), CopA (Cu [I]-translocating P-type ATPase), CusCFBC, PcoABCD, PcoE (plasmid borne system) each having their own identity and function. The high toxic nature of copper in the digestive tract of warm-blooded animals (where *E. coli* live) influenced enteric bacteria to develop resistance mechanisms for copper. Two regulatory proteins, CusR and CusS, regulate the *cusCFBA* genes and CueR regulates the *copA* and *cueO* genes.[63]

The protein CopA is considered to be the central component of cytoplasmic copper homeostasis in *E. coli.* CopA is a Cu (I) translocating P-type ATPase controlled by CueR. The main function of the protein CopA in *E. coli* is to extrude excess copper

present in the cytoplasm.[63] The function of CopA in *Enterococcus hirae* is opposite and imports copper when it is deficient.[66] CopA is also present as an outer membrane protein encoding *cop* operon in *Pseudomonas syringae* performing sequestration and compartmentalization of copper in the periplasm and outer membrane.[67] The function of the CopB protein in *E. hirae* is to remove excess copper present in the cytoplasm.[66] The specific function of CopB protein in *E. coli* and *Pseudomonas syringae* are not yet defined.[63, 67] The genes, *copA* and *copB*, in *E. hirae* transport copper using ATPases while *copY* acts as a copper responsive repressor and *copZ* functions in the transport of intracellular copper.[66]

CopC is another periplasmic outer membrane protein found in *P. syringae*. It performs sequestration of copper in periplasm along with CopA which transports copper along with CopD, an inner membrane protein.[67] CopY is an *E. hirae* protein that binds to the promoter region of the *cop* operon. It functions as a repressor involved in the expression of the *cop* operon. CopZ is a metallo chaperone that passes copper to the CopY repressor.[66] The two separate regulatory genes, *copR* and *copS*, encoded by *P. syringae* form a signal transduction system in regulating the expression of *copABCD*. The operon system of *E. coli*, *pcoABCDRS* differs from *P. syringae copABCDRS* due to the different expression systems the bacteria uses to adapt and survive.[67]

Cus is another transport system responsible for copper homeostasis in *E. coli* and consists of two operons on the chromosome and are expressed in opposite directions. Two regulatory component system CusR/S senses excess copper in the periplasm and regulates the *cus* expression. CusR functions as a transcription regulatory factor for *cusCFBA* and CusS is a membrane bound histidine kinase involved in identifying copper

ions located in the periplasm. The Cus transport system exports copper ions from the periplasm across the outer membrane. A careful review of the two transport systems, CzcCBA and CusCFBA postulated that the transport of copper occurs from cytoplasm or periplasm. CusA is considered to be the central component of the Cus system and exports copper out of the cell. CusB and CusC are two other important proteins mediating copper resistance and play a major role in the function of Cus system. CusF is a unique periplasmic metallochaperone that binds to copper in periplasm and transfers it to the CusCBA efflux pump.[63]

The Cue transport system consists of the *copA* gene which encodes a copper efflux P-type ATPase and CueO, a multi copper oxidase.[63] CueO is a periplasmic protein possessing laccase activity. It participates in the biosynthesis of antibiotics, sporulation, tolerance to copper, morphogensis and oxidation of manganese. CueO is responsible for protecting the periplasm from copper induced damage. CueO detoxifies Cu(I) by sequestering it, transporting it to the periplasm and oxidizing it to the less toxic Cu (II).[63]

The *pco* operon encodes copper resistance, contains the seven genes, *pcoABCDRSE*, and was found on plasmid pRJ1004 from the gut flora of pigs. The gene products are responsible for copper efflux. PcoA belongs to the family of multi-copper oxidases and is considered to be the central protein of the *pco* system. In *E. coli*, PcoA sometimes replaces CueO due to the identical oxidase activity of two proteins. PcoB is another outer membrane protein. A combination of PcoA and PcoB is assumed to confer more copper resistance than the individual proteins alone. PcoC and PcoD confer maximal resistance and function together for copper uptake in *P. syringae*. They are determined to be more efficient as fusion genes termed as *ycn*J than as single genes. The

N- terminal region of PcoC and C-terminal region of PcoD form the protein YcnJ. The function of PcoC is to transport copper present in the periplasm to PcoD in the cytoplasm which delivers it to PcoA. Although *pcoE* is associated to the *pco* determinant, it is not considered to be a part of *pcoABCD* operon. It is situated downstream of the regulatory systems CusRS in plasmid pRJ1004.[63]

The mechanisms by which copper homeostasis takes place in *E. coli* involves six genes, *cutA, cutB, cutC, cutD, cutE,* and *cutF* that are involved in uptake, intracellular storage, delivery, efflux and copper metabolism All these genes also assist in reducing methionine in CueO, insuring proper protein folding in the periplasm and in alteration of pores.[63] The locus of *cutA* includes two operons, one operon with a single open reading frame encoding a cytoplasmic protein and the other operon possessing two genes, *cutA2* and *cutA3*, which encode inner membrane proteins. CutC, a cytosolic protein and CutF (NlpE), an outer membrane lipoprotein are also implicated in conferring copper resistance. The protein, CutE (*lnt*), encoding apolipoprotein *N*-acyltransferase, is an inner membrane protein involved in catalyzing the final reaction of Braun's lipoprotein, a major lipoprotein. Mutations in this gene in *Salmonella typhimurium* alters the minor lipoproteins which are responsible for copper tolerance and protection of the cell.[68]

Copper functions as a cofactor in various oxidases and hydrolases in electron transport system due to its oxidation-reduction properties. The toxicity of copper is explained by the production of reactive oxygen species in *E. coli*. The toxicity of copper is increased by the enzyme cupric reductase and NADH which reduce of Cu (II) to Cu (I). This Cu (I) is then oxidized by hydrogen peroxide causing damage to the respiratory system in *E. coli*. In *E. coli,* ubiquinone is the electron acceptor of NADH which is

considered to be the major site of the copper mediated damage by hydrogen peroxide. $NADH_2$ also reduces Cu (II) in the presence of FAD or quinone. Enzymes such as succinate, D-lactate dehydrogenases, and other thiols containing proteins are also involved in electron transfer by NADH.[69]

The other mechanisms for copper homeostasis include influx and efflux pathways, modification in the cytoplasm and sequestration by metallothionines.[65]

The *synechococcus* PCC 7942 thylakoid proteins, PacS and CtaA, are P-type ATPases which are involved in the transport of copper from the cytosol to the external medium or to an inner compartment. Atx1 is a *synechocystis* metallochaperone that incorporates copper ions into proteins.[70]

Chapter III: Hypothesis

The present research focuses on a multi-metal resistant strain of *Enterobacter cloacae* and its metal resistance genes. Some of these genes were identified by randomly mutagenizing this strain with a transposon and identifying some of the mutants that are interrupted in this strain. As discussed earlier, we expected to find some of the genes involved in reducing oxidative stress, in acting as an efflux pump or in transforming metal ions. Understanding the mechanisms of resistance to heavy metals may provide valuable information on using microorganisms to clean up some of the metal contaminated sites.

Chapter IV: Materials and Methods

4.1. Bacterial Strains:

Enterobacter cloacae (*E. cloacae*) is a mercury, cadmium, zinc and selenite resistant strain from an unknown origin. It was believed to be *Stenotrophomonas maltophilia* ORO2 (ATCC # 53510), but 16s rRNA sequencing and biochemical tests identified it as *E. cloacae*. Since its origin in unknown, this strain will be referred to as *E. cloacae* UNK. EC 100D p*ir* [F^- *mcr*A Δ (*mrr-hsd*RMS-*mcr*BC) ø80d/acZ ΔM15 Δ/acX74 *rec*A1 *ara*D139 Δ (*ara, leu*) 7697 *gal*U *gal*K 7^- *rps*L *nup*G *pir*-116 (DHFR)} and EC 100D pir-116 [F^- *mcr*A Δ (*mrr-hsd*RMS-*mcr*BC) ø80d/acZΔM15 Δ/acX74 *rec*A1 *end*A1 *ara*D139 Δ (*ara, leu*) 7697 *gal*U *gal*K 7^- *rps*L *nup*G *pir*-116 (DHFR)} were used for gene rescue and were purchased from Epicentre Biotechnologies (Madison, WI). Both *E. coli* strains contain genes for a trans-acting II protein (*pir* gene product) that allow plasmids with R6Kγ*ori* replication origins to replicate.[71]

4.2. Metals:

Sodium selenite was purchased from MP Bio Medicals LLC, (Solon, Ohio). Mercuric chloride, copper sulfate, zinc chloride, potassium dichromate, lead nitrate and cadmium chloride were purchased from Fisher Scientific (Fair Lawn, New Jersey).

4.3. Media preparation:

Bacterial cells were grown at 37°C in Luria Bertani (LB) medium (Fisher Scientific, Fair Lawn, New Jersey) which consisted of 10 grams of Bacto Tryptone, 5 grams of yeast extract and 5 grams of NaCl per liter of deionized water. When required, media were supplemented with 1.6% Agar (Amresco, Solon, Ohio) and 50 µg/mL kanamycin sulfate (Amresco, Solon, Ohio).

M-9 minimal medium[72] contained 0.24 M anhydrous disodium phosphate, 0.11 M monopotassium phosphate, 0.04 M sodium chloride, 0.09 M ammonium chloride, 0.45 M $MgSO_4$, 9 % Glucose, 0.225 % of Thiamine, and water. When required, M-9 minimal medium was supplemented with 4 mg/mL cysteine hydrochloride (Fisher Scientific, Fair Lawn, New Jersey).

SOC medium[73] contained 2% tryptone (Amresco, Solon, Ohio), 0.5% yeast extract (Fisher Scientific, Fair Lawn, New Jersey), 10 mM sodium chloride (Fisher Scientific, Fair Lawn, New Jersey), 2.5 mM potassium chloride (Amresco, Solon, Ohio), 10 mM magnesium chloride (Fisher Scientific, Fair Lawn, New Jersey), 10 mM magnesium sulfate (Fisher Scientific, Fair Lawn, New Jersey) and 20 mM glucose (Amresco, Solon, Ohio) per liter solution of deionized water.

A modified Tris-R3A medium[74] was prepared to minimize metal precipitation. It contained 0.1% yeast extract (Fisher Scientific, Fair Lawn, New Jersey), 0.1% Difco Protease Peptone no. 3 (Difco Laboratories, Sparks, MD), 0.1% casamino acids (Amresco, Solon, Ohio), 0.1% Glucose (Amresco, Solon, Ohio), 0.1% soluble starch (Difco Laboratories, Sparks, MD), 0.05% sodium pyruvate (Fisher Scientific, Fair Lawn, New Jersey), 0.01% $MgSO_4.7H_2O$ (Fisher Scientific, Fair Lawn, New Jersey) and 10 mM of Tris, pH of 7.5 (Amresco, Solon, Ohio) per liter.

4.4. Transformation

Transformation of competent *E. coli* cells was performed using a $CaCl_2$ technique.[72] Single colonies were inoculated into 3 mL of LB medium and grown overnight at 37^oC in a TC8 roller drum (New Brunswick Scientific, Edison, New Jersey).

The overnight cultures were then diluted 1:50 into fresh LB medium and grown at 37°C with shaking in a C24 Incubator Shaker (New Brunswick Scientific, Edison, New Jersey) until the cells reached to an optical density of 1.0 at 600 nm as determined by an Eppendorf BioPhotometer spectrophotometer (Eppendorf, Westbury, NY). The cells were then chilled to 4°C and harvested by centrifuging at a speed of 6000 x g in an Eppendorf 5810 R micro centrifuge (Westbury, NY) at 4°C. The cells were resuspended in 40 mL of 0.15 M NaCl and centrifuged again at a speed of 6000 x g at 4°C. The cells were resuspended in 1 mL of transformation buffer containing 0.1 M $CaCl_2$ (Fisher Scientific, Fair Lawn, New Jersey), 15% glycerol, 0.01 M tris-HCl, pH 8.0 and 0.01 M $MgCl_2$. After incubating them on ice overnight, the cells were frozen and stored at -80°C.

The competent *E. coli* cells (100 µL) were thawed on ice and mixed with 1 µL of approximately 1 µg DNA. After incubating them on ice for 30 min, the cells were heat shocked at 42°C for 50 sec and placed back on ice. LB media was added to a final volume of 1 mL, and the cells were incubated with shaking for 45 – 120 minutes. Volumes ranging from 10 to 1000 µL of cells were spread on LB-agar plates containing the appropriate antibiotic and incubated at 37°C overnight. The number of colonies that grew the next day were counted and recorded.

4.5. Electroporation

Electroporation uses an electric shock to transform *E. coli* cells with DNA[72] Single colonies were inoculated into 3 mL of LB medium and grown overnight at 37°C in a TC8 roller drum (New Brunswick Scientific, Edison, NJ). The overnight cultures were then diluted 1:50 into 250 mL of fresh LB medium and grown at 37°C with shaking in a

C24 Incubator Shaker (New Brunswick Scientific, Edison, New Jersey) until the cells reached to an optical density between 0.4 and 0.6 at 600 nm (approximately 2hours) as determined by an Eppendorf BioPhotometer spectrophotometer (Westbury, NY). The cells were then chilled to 4°C and harvested by centrifuging at a speed of 8000 x g in an Eppendorf 5810 R centrifuge (Westbury, New York) at 4°C. Cells were resuspended in equal volumes (250 mL) of ice cold water and centrifuged. The cold water washes were repeated twice followed by a wash in a 1:5 volume (50 mL) of ice cold 10% glycerol. The cells were resuspended in 0.5 mL of 10% ice cold glycerol and frozen at -80 ^{o}C. The electrocompetent *E. coli* cells (40 µL) were thawed on ice, mixed with 0.4-1 µL of approximately 1 µg DNA and electroporated in 2 mm cuvettes using a Bio-Rad Gene Pulser II (Bio-Rad, Hercules, CA) set at 25 µF, 2.5 kV and 200 Ω. The cells were immediately resuspended in 960 µL of SOC medium. After incubating them at 37°C in an Isotemp Incubator (Fisher Scientific, Fair Lawn, NJ) for 45-120 minutes, 10 - 1000 µL of cells were spread on LB plates containing the appropriate antibiotic and incubated at 37°C overnight. The number of colonies that appeared on the plates the next day were counted and recorded.

4.6. Transposon Mutagenesis

Transposon mutagenesis was performed using the EZ-Tn*5* <R6Kγ*ori*/KAN-2> Tnp Transposome Kit.[75]

A volume of 0.4 µL of EZ-Tn*5* Transposome was mixed with 40 µL of electrocompetent cells and transformed by electroporation. The electroporated cells were mixed immediately with 960 µL of SOC medium, incubated for 45 - 120 min in an

environmental shaker at 37°C, spread on LB plates containing 50 µg/mL kanamycin and incubated overnight at 37°C.

4.7. Screening for Metal Sensitive Mutants by Replica Plating[76]

The colonies that grew on the LB-kanamycin plates above were spotted as a grid of 50 on fresh kanamycin plates and allowed to grow at 37°C overnight. The cells were transferred to a Scienceware Velveteen Square (Bel-Art, Pequannock, NJ) that was mounted on a Scienceware Replica-Plating Tool (Bel-Art, Pequannock, NJ) solid cylinder. The cells were then transferred to R3A-agar, M-9 agar or LB-agar plates containing different metals. Growth or lack of growth in the presence of the different metals was recorded in Table 1.

4.8. Purification of genomic DNA

The genomic DNA was purified using a Wizard Genomic DNA Purification Kit purchased from Promega (Madison, WI) and all the ingredients were supplied with this kit. The kit contains Cell Lysis Solution, Nuclei Lysis Solution, Protein Precipitation Solution, DNA Rehydration Solution and RNase Solution. The other materials not supplied with the kit were sterile 1.5 mL microcentrifuge tubes, a water bath set at 37°C, isopropanol, and 70% ethanol.

An overnight culture of 1 mL was centrifuged at 13,000-16,000 x g for 2 min to pellet the cells. Next, the pellet was resuspended in 600 µL of Nuclei Lysis Solution by gently pipetting and incubated at 80°C for 5 min to lyse the cells. The sample was cooled to room temperature and the tubes were inverted 2-5 times after adding 3 µL of RNase solution and incubated at 37°C for 15-60 min. After the sample was cooled to room

temperature, 200 µL of Protein Precipitation Solution was added and the cell lysate was vortexed vigorously. After incubating the preparation on ice for 5 minutes, the samples were centrifuged at 13,000-16,000 x g for 3 minutes. The supernatant was transferred to a clean 1.5 mL microcentrifuge tube containing 600 µL of room temperature isopropanol to precipitate the DNA. The DNA was pelleted by centrifugation at 13,000- 16,000 x g and the supernatant was discarded. The pellet was washed with room temperature 70 % ethanol and centrifuged at 13,000- 16,000 x g for 2 minutes. Again the supernatant was discarded and the pellet was air dried on clean absorbent paper and resuspended in 100 µL of DNA Rehydrating Solution. The DNA was stored at 4°C.

4.9. DNA Purification

Eppendorf Perfect prep Plasmid Mini kit was used to purify DNA plasmids. The kit was supplied with Solutions 1, 2 (sodium hydroxide) and 3, spin Columns, DNA binding matrix (guanidinium chloride), purification solution concentrate, elution buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) and collection tubes. The pellet obtained by centrifuging 3 ml of overnight culture in an Eppendorf 5415D at 12,000-16,000 x g for 20 seconds was resuspended in 200 µL of solution 1 by vortexing. Next, 200 µL of solution 2 was added and the tube was inverted several times to lyse the cells. The lysate was immediately mixed with 200 µL of solution 3 to neutralize the bacterial lysate. The preparation was centrifuged at 12,000-16,000 x g for 2 minutes to pellet contaminating lipids, polysaccharides, proteins and chromosomal DNA. The supernatant was mixed with 450 µL of DNA binding matrix and transferred to a spin column in a collection tube. The solution in spin column/collection tube assembly was mixed by pipetting or vigorously inverting the assembly. The spin column was centrifuged at 12,000- 16,000 x

g for 30 seconds and the filtrate was discarded. Next, the retained resin in the spin column was resuspended in 400 µL of diluted purification solution and mixed. After centrifuging the spin column at 12,000- 16,000 x g for 30 seconds, the filtrate was discarded. The spin column was centrifuged again at 12,000- 16,000 x g for 60 sec to remove any remaining diluted purification solution. The spin column was then transferred to a new collection tube, and 60 µL of elution buffer, heated at 70°C, was added directly to the DNA binding matrix in the spin column. The spin column was then centrifuged at 12,000-16,000 x g for 60 sec to elute the plasmid DNA which was stored at -20°C.

The concentration was then determined using the eppendorf Bio Photometer (Eppendorf, Westbury, NY) by taking the absorbance of the samples at 260 nm followed by DNA analysis.[72]

4.10. Enzyme Digestion for Gene Rescue

Genomic DNA from the metal sensitive strains were digested in New England Biolabs (Beverly, MA) buffer 3 and 1 X BSA (0.1 mg/mL) and contained the blunt–end cutting enzymes (0.5 µL each) *Sca*I (20,000 U/mL), *Pvu*II (10,000 U/ml), *Bsr*BI (10,000 U/mL) and *Eco*RV (20,000 U/mL). After digesting the DNA at 37°C, it was incubated at 80°C to inactivate the enzymes and ligated using T4 DNA ligase as described below.

4.11. Enzyme Digestion

Purified plasmid DNA was digested using the restriction digestion endonuclease enzyme, *Pst*I (0.5 µL), 1 µL of 10 x NEBuffer 3 (50 mM Tris-HCl, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM Dithiothreitol pH 7.9), 1 µL of 10 x BSA (1 mg/mL) obtained from

New England Biolabs, Beverly, MA, and 7.5 µL of purified DNA. The digested DNA (5 µL) was then run on agarose gel electrophoresis to estimate the size of DNA.

4.12. DNA Ligation

The digested DNA was ligated using 1 µL of 10 x T4 DNA Ligase Reaction buffer (50 mM Tris-HCl, 10 mM DTT, pH 7.5, 10 mM $MgCl_2$, 1 mM ATP), 1.5 µL of T4 DNA ligase (400,000 U/mL) obtained from New England Biolabs (Beverly, MA), 6.5 µL of digested DNA and 1 µL of deionized water. The mixture was incubated overnight at 4^oC. The ligated DNA was then transformed into *E. coli*.

4.13. DNA Concentration Determination

DNA concentration was determined using the Lambert-Beer equation $A=\varepsilon c l$, where A is the absorbance of sample at a particular wavelength, ε is the extinction coefficient usually 50 µg/mL for dsDNA, c is the concentration in µg/mL, *l* is the path length of spectrophotometer cuvette (1 cm). The absorbance of DNA samples were taken using an Eppendorf BioPhotometer set at wavelength of 260 nm. The concentration of DNA obtained can be detected in a range of 1- 50 µg/ml.[72]

4.14. DNA sequencing

DNA sequencing was performed using the GenomeLab™ Dye Terminator Cycle Sequencing with Quick Start Kit purchased from Beckman Coulter, Inc. (Fullerton, CA). The volume of DNA samples used for DNA sequencing was calculated depending upon the size and concentration of DNA fragment used following the kit instructions. The sequencing reaction was set by mixing thoroughly sufficient quantities of distilled H_2O,

DNA template, 0.16 μM KAN-2 FP-1 Forward Primer (5' - ACCTACAACAAAGCTCTCATCAACC - 3') or R6KAN-2 RP-1 Reverse Primer (5' - CTACCCTGTGGAACACCTACATCT - 3') obtained from Epicentre Biotechnologies, and DTCS Quick Start Master Mix. The reactions were incubated in an Eppendorf Master Cycler according to the following program: 96°C for 20 sec, 50°C for 20 sec and 60°C for 4 min for 30 cycles followed by holding at 4°C.

The next day, DNA was precipitated and washed according to the Beckman Coulter protocol. Freshly prepared 5 μL of stop solution/glycogen mixture (2 μL of 3M sodium acetate, pH 5.2; 2 μL of 100 mM Na_2-EDTA, pH 8.0 and 1 μL of 20 mg/mL of glycogen) was mixed thoroughly with the sequencing reaction. The DNA was then precipitated by adding 60 μL of cold 95% (v/v) ethanol/distilled H_2O and centrifuged at 14,000 x g for 2 minutes in an Eppendorf 5415D centrifuge. The pelleted DNA was washed twice with 200 μL of 70% (v/v) ethanol/distilled H_2O followed by centrifugation at 14,000 x g for 2 minutes. Finally the pellet was air dried, resuspended in 40 μL of Sample Loading Solution and analyzed using a Beckman Coulter CEQ 2000 XL DNA Analysis System (Fullerton, California) in the Department of Biological Sciences at Youngstown State University.

4.15. BLAST Analysis

BLAST is a Basic Local Alignment Search Tool[75, 77] used to identify gene families by comparing the nucleotide or protein sequences with the reference sequences or library of sequences at the National Library of Medicine. The query sequence is entered either in Accession number, gi or FASTA format using the blastp or blastn program.

4.16. Polymerase Chain Reaction

The polymerase chain reaction (PCR) was performed using Fisher Scientific Taq Polymerase (Fair Lawn, NJ), 10 mM dNTPs, 10 x PCR Buffer and *mer* primers 5'-GGGAGATCTAAAGCACGCTAAGGC[G or A]TA-3' and 5'-GGGGAATTCTTGAC[T or A]GTGATCGGGCA-3' or *pco* primers 5’ CGTCTCGACGAACTTTCCTG 3’ and 5’ GGACTTCACGAAACATTCCC 3’[78] obtained from Integrated DNA Technologies, Inc. (Coralville, IA). The PCR reaction was set up using the above ingredients and incubated overnight in a PCR Thermal Cycler set with the following program: 98°C for 1 min, 55°C for 20 sec, and 72°C for 30 sec for 30 cycles followed by holding at 4°C.

4.17. Primer Design

The primers used for DNA sequencing were determined using Vector NTI based on region of DNA sequence, length of the product, melting temperature of primers, % GC, and maximum and minimum length of primers. The default variables such as 50 millimolar of salt concentration, 25 pmol of Probe concentration, and dG temperature of 25°C were used to calculate the primer melting temperatures (T_m).

Table 1. Primers used in sequencing reactions

Primers	Nucleotide Sequence
D21 F	5'- TTT ATA TCG CAC CTG AAT CC - 3'
D21 R	5'- CAG AAG ATG GCG AAA GTG GG - 3'
D21 F2	5'- TTT ATA TCG CAC CTG AAT CC - 3'
D21 R2	5'- CAG AAG ATG GCG AAA GTG GG - 3'
F24 F	5'- TTT TTA CCG ACG GCG CAA - 3'
F24 R	5'- CGT GAC GAT GCG AAA GAC G - 3'
F24 F2	5'- TTC CTC GTC ACC ACG CTG CT - 3'
F24 R2	5'- GGT ACA CCG TCC TGC ATC AC - 3'
F34 F	5'- CCA TTT GGA CTG GCC TGC T - 3'
F34 R	5'- GAT GCC CGC AGC CTT TGG - 3'
L31 F	5'- CTT GTG AGC GAA ACG GTG - 3'
L31 R	5'- CAT GAC CTT GAT ACG CGA - 3'
6B F	5'- TTA GAC GAA CTG CTC AGC TG - 3'
6B R	5'- ACT CCA CGC TGC CAA TTG - 3'
6B F2	5'- ATT AAA AAG CAC CTG CCG AA - 3'
6B R2	5'- AGA TCG CCC TGA ACT TCA AC - 3'
6B F3	5'- AAT TTC ACC GCC TAC CAC AC - 3'
6B R3	5'- CAT CCA TTA ACT CTT CTT CG - 3'

8HB F	5'- GCA GTA GAC GGT GCT GCA TG - 3'
8HB R	5'- ACT GGC CCG GCG ACA CAT- 3'
8HB F2	5'- ACG GCT ACG CAC GAC GAA TC - 3'
8HB R2	5'-CTG GAT TGA GCG CGT ATG AC - 3'
Q17 F2	5'- AAC CTT GTA GAA CTC ATC CA - 3'
Q17 R2	5'- AAG GCG TGG TCG GTG AAC AA - 3'

Below were the primers used for sequencing reactions of PCR fragments of copper and mercury resistance genes, *pcoA* and *merR*.

M13 Forward primer: 5'-ACTGGCCGTCGTTTTACAA-3'

M13 Reverse primer: 5'-GGAAACAGCTATGACCATG-3'

4.18. Agarose Gel Electrophoresis

Agarose gel electrophoresis[72] was used to verify the purity of DNA. Agarose gel electrophoresis uses an electric field and an agarose matrix for separating DNA by size. Once the electric field is applied, the DNA begins to migrate towards the anode due to negatively charged phosphate groups. Friction during movement through the agarose matrix caused size separation with small molecules traveling through the gel faster than the larger ones. The bands are visualized using 1% ethidium bromide as a staining reagent which fluoresces after it intercalated between the DNA bases.

DNA samples were separated in 1% agarose gels using an electrophoresis apparatus purchased from Embi Tec (San Diego, CA), and a 1 kb DNA ladder purchased from New England Biolabs (Beverly, MA) was used as a reference to determine the size

of DNA fragments. 1 gram of Agarose (Fisher Scientific, Fair lawn, NJ) was dissolved in 100 mL of 1 x Tris Borate EDTA Buffer [Tris base 0.089 M, Borate 0.089 M, EDTA 0.002 M at pH of 8.3] (Fisher Scientific, Fair Lawn, NJ) and heated in a microwave until all the agarose was melted. 5 µL of 1% ethidium bromide (Fisher Scientific, Fair Lawn, NJ) was added to the solution for the visualization of DNA under UV light. The casting gels were prepared by pouring the agarose solution in casting trays avoiding the air bubbles in it and a comb was placed in it for the formation of wells. The combs were taken out of the trays after the agarose solidified.

The casting gels were submerged in 1 x TBE buffer before loading the samples into the wells. Then, 5 µL of sample was mixed with the 1 µL of loading dye consisting of 15 % ficoll, light blue dye #1, indigo dye #2, and magenta dye #3, and added to the wells of the gels. An electric current of 100 V was applied to run the gel for 30 min-1 hour. After the blue dye migrated 2/3 the length of the gel, it was removed from the electrophoresis apparatus and visualized under UV light. The pictures of the gels were taken using UltraCam Imaging Systems purchased from Ultra-Lum, Inc, (Claremont, CA) and saved on a computer for further analysis.

CHAPTER V: RESULTS

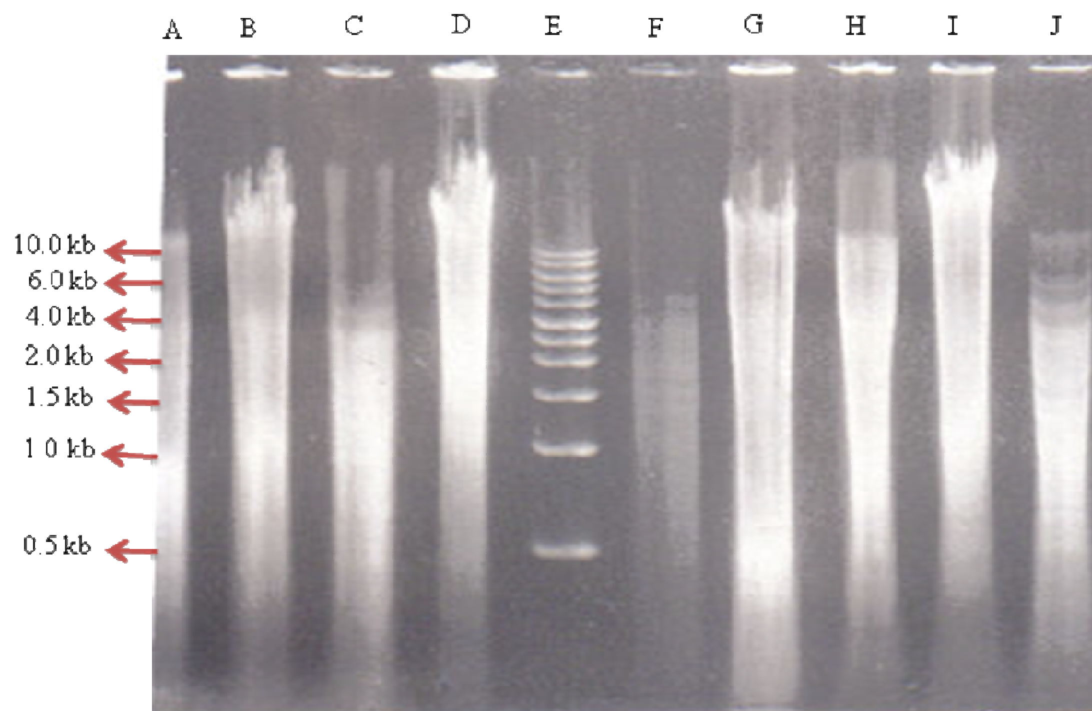
E. cloacae UNK was transformed by the EZ-Tn*5* transposome to generate metal sensitive mutants. One thousand kanamycin resistant colonies were replica plated on R3A agar medium that contained Cd (240 µM), Se (60 mM), Zn (750 µM), Hg (10 µM) or Cu (3 mM). These were the minimal inhibitory concentrations (MICs) on R3A medium for the *E. coli* strain HB101 which was used as the negative control. Of the 1000 colonies tested, two were sensitive to zinc, three were sensitive to cadmium and five were sensitive to selenite. The mutants, their ability to grow in the presence of each metal and the genes that were interrupted by the transposon are listed in the Table 2.

Table 2. Mutants obtained by transformation

Mutant	Cd	Se	Zn	Hg	Cu	Interrupted Gene	Accession number
Wild Type	+++	+++	+++	+++	+++	________________	
A3A	+++	_	+++	+++	+++	Lon protease	ZP_02751147.1
F24	_	_	_	+++	+++	P-type ATPase	YP_001178579.1
F34	_	++	+++	+	+++	Acyl transferase	YP_001746189.1
L31	+++	_	+++	+++	+++	Type II sec protein	ZP_02138252.1
6B	+	_	+++	+++	+++	Sporulation domain protein	ABP62456.1
D21	+++	_	+++	+++	+++	Lon protease	YP_001175641.1
8HB	_	_	_	+++	+++	Tyrosine recombinase	CP000653.1
Q17B	+++	_	+++	+++	+++	Polyphosphate kinase	YP_001177705.1
10A	_	_	_	+++	+++	Hypothetical protein	

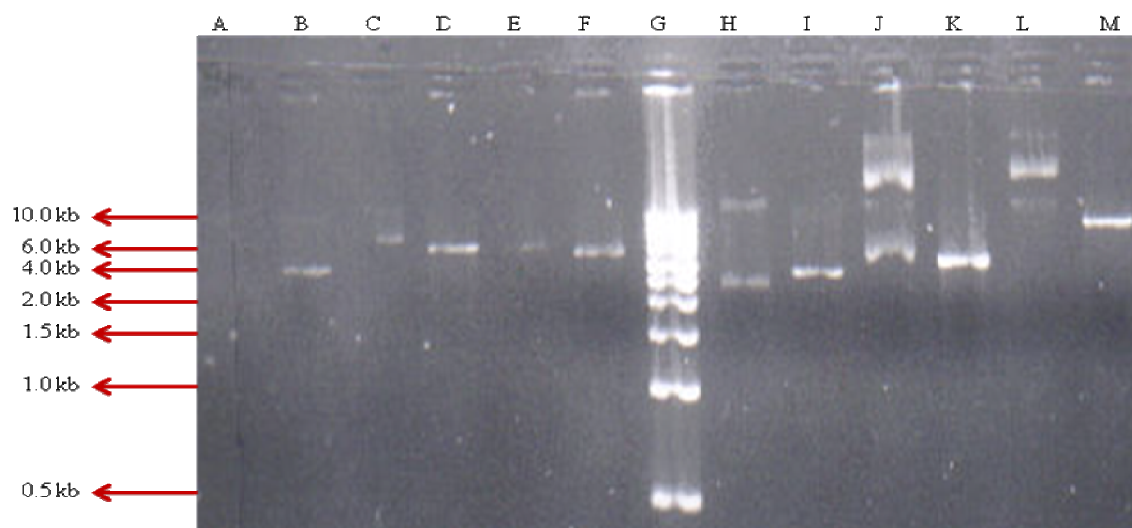
The mutant colonies obtained by transformations were retested for different metal sensitivities using replica plating. Then, the interrupted genes from these mutants were identified by gene rescue. Genomic DNA from each metal sensitive mutant was purified and digested with the blunt end generating enzymes, *Sca* I, *Pvu* II, *Bsr*B I and *Eco*R V. The digested DNA was analyzed by agarose gel electrophoresis as shown in Figure 1. The smears in lanes A-D and F-J showed that the DNA was completely digested.

Figure 1. Blunt end digestion of EZ-Tn*5* mutant genomic DNA. Lane 1(A): A3A; Lane 2(B): C-18, Lane 3(C): F-34; Lane 4(D): G-5; Lane 5(E): DNA ladder (1 kb); Lane 6(F): G-27; Lane 7(G): L-30; Lane 8(H): L-31; Lane 9(I): Q-17; Lane 10(J): Q-34.


A
B
C
D
E
F
G
H
I
J
10.0 kb
6.0 kb
4.0 kb
2.0 kb
1.5 kb
1.0 kb
0.5 kb

These enzymes did not cut within the inserted transposon and created blunt ended DNA fragments which allowed for easy ligation. In addition, the transposome contains a kanamycin selection marker and the R6K*γori* which allowed for replication in the ECD100D *pir E. coli* strain. Thus, the ligation produced new plasmids that contained the transposome and flanking *E. cloacae* regions that were interrupted by the transposome. Transformation of the ligation mixture into *E. coli* strain ECD100D *pir* generated 5 transformants for mutant A3A, numerous transformants for L31, 4 transformants for mutant L30, 13 transformants for Q17, 5 transformants for F24, transformants for F34, 85 transformants for 6B, 3 transformants for D21, 2 transformants for 8HB, 7 transformants for 10A. The new plasmids were then purified using a Fast Plasmid MiniPrep and digested using enzyme, *Pst* I which cuts at only one position of the transposome and was helpful in obtaining linearized DNA fragments. The digested samples were then analyzed by agarose gel electrophoresis (Figures 2-4).

Figure 2. Digestion of the transformed DNA. Lane 1(A): Undigested A3A; Lane 2(B): Digested A3A; Lane 3(C): Undigested C18C; Lane 4(D): Digested C18C; Lane 5(E): Undigested L30 A; Lane 6(F): Digested L30; Lane 7(G): DNA Ladder (1 kb); Lane 8(H): Undigested L30 B; Lane 9(I): Digested L30 B; Lane 10(J): Undigested L31 A; Lane 11(K): Digested L31 A; Lane 12(L): Undigested Q17 B; Lane 13(M): Digested Q17B


A
B
C
D
E
F
G
H
I
J
K
L
M
10.0 kb
6.0 kb
4.0 kb
2.0 kb
1.5 kb
1.0 kb
0.5 kb

Undigested and digested samples were loaded to ensure that the DNA was digested completely. For example in figure 2, Lane L is the undigested plasmid from mutant Q17 and lane M is the digested sample. The plasmid was clearly digested in lane M because is demonstrates a different migration pattern that the one in lane L. Lane A appears empty because residual ethanol from the plasmid preparation caused the sample to migrate out of the well after it was loaded. The sizes of the DNA fragments were estimated using the DNA ladder in lane G. In this lane, the 4th band from the bottom is 2 kb in size, and the size of each band increases by 1 kb with each band above the 2 kb band. Thus, the 5th band from the bottom is 3 kb in size. The ladder in the 3 gels show that rescued plasmids range in sizes from 3 kb to 6 kb as listed in Table 2. Since the transposome is 2 kb in length, the length of *E. cloacae* DNA associated with the transposome ranged from 1-4 kb in length. By determining the sequence of the *E. cloacae* DNA that flanked the transposome, it was possible to identify the mutated genes using the KAN-2 FP-1 Forward Primer or the R6KAN-2 RP-1 Reverse Primer (see methods section) that were homologous to the transposome.

Figure 3. Digestion of the transformed DNA. Lane 1(A): Undigested A3A; Lane 2(B): Digested A3A; Lane 3(C): Undigested A3C; Lane 4(D): Digested A3C; Lane 5(E): Undigested C18C; Lane 6(F): Empty; Lane 7(G): Digested C18C; Lane 8(H): DNA Ladder (1 kb); Lane 9(I): Undigested F34; Lane 10(J): Digested F34; Lane 11(K): Undigested L 30A; Lane 12(L): Digested L30A.

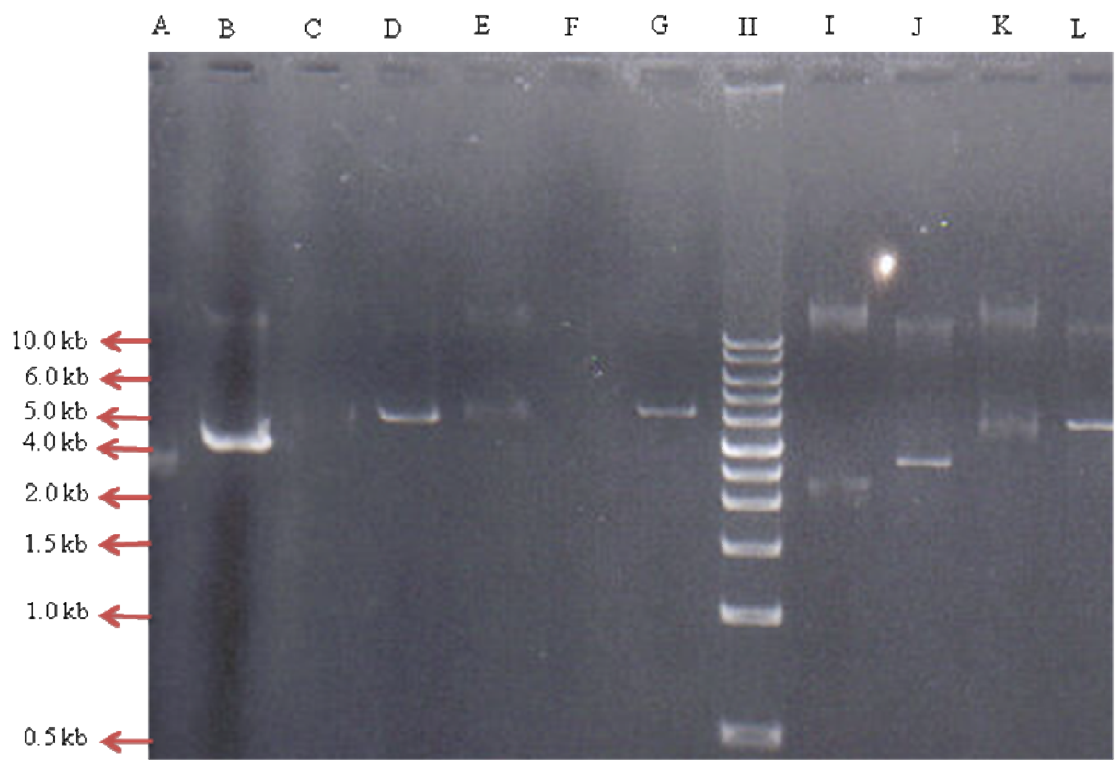
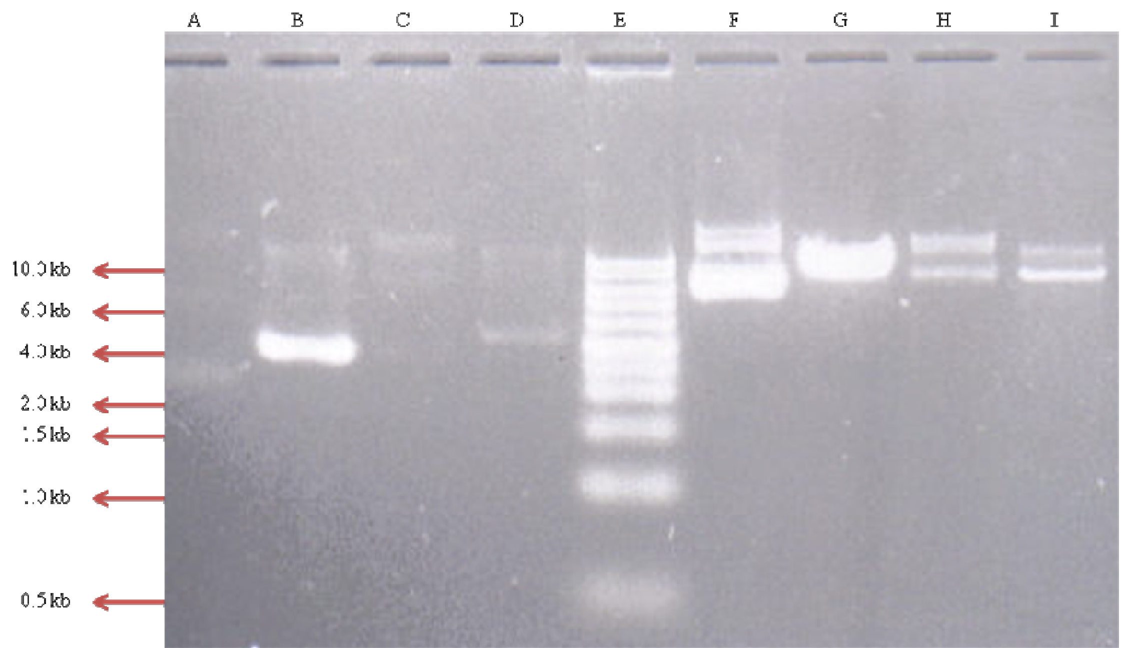

A B C D E F G H I J K L
10.0 kb
6.0 kb
5.0 kb
4.0 kb
2.0 kb
1.5 kb
1.0 kb
0.5 kb

Figure 4. Digestion of the transformed DNA. Lane 1(A): Undigested L30 B; Lane 2(B): Digested L30 B; Lane 3(C): Undigested L31 A; Lane 4(D): Digested L31 A; Lane 5(E): DNA Ladder (1 kb); Lane 6(F): Undigested Q 17 B; Lane 7(G): Digested Q 17 B; Lane 8(H): Undigested Q 17C; Lane 9(I): Digested Q 17C.


A
B
C
D
E
F
G
H
I
10.0 kb
6.0 kb
4.0 kb
2.0 kb
1.5 kb
1.0 kb
0.5 kb

All sequencing reactions with a particular primer resolve between 400 to 800 bp. To resolve a larger section, additional primers homologous to a region downstream of the original primer must be designed. Thus, a 1,000 bp region required at least two primers and a 4,000 bp region required at least 8 primers to be resolved. The additional primers used are listed in Table 1 and the concentrations and sizes of DNA samples used in sequencing are listed in Table 3.

Table 3. Concentrations and sizes of the transformed DNA

DNA fragment	Concentrations (ng/µL)	Sizes (kb)
A3A	106	4
F24	97	4
F34	117	4
C18C	93	5
L30	40	3
L31	89	5
Q17B	59	8
8HB	155	5
6B	124	8
10A	64	2

When using the KAN-2 FP-1 Forward Primer and the R6KAN-2 RP-1 Reverse Primer for sequencing, part of the sequence obtained contained a segment from the transposome. These segments were removed from the data and then assembled with data from other sequencing reactions into continuous sequences using the Contig computer program from Invitrogen's (Carlsbad, CA) software package, Vector NTI. The continuous sequences were analyzed using the Basic Local Alignment Search Tool (BLAST) to predict a possible function for each gene that was interrupted. Figures 5-29 show maps of the obtained sequences and some of the BLAST results.

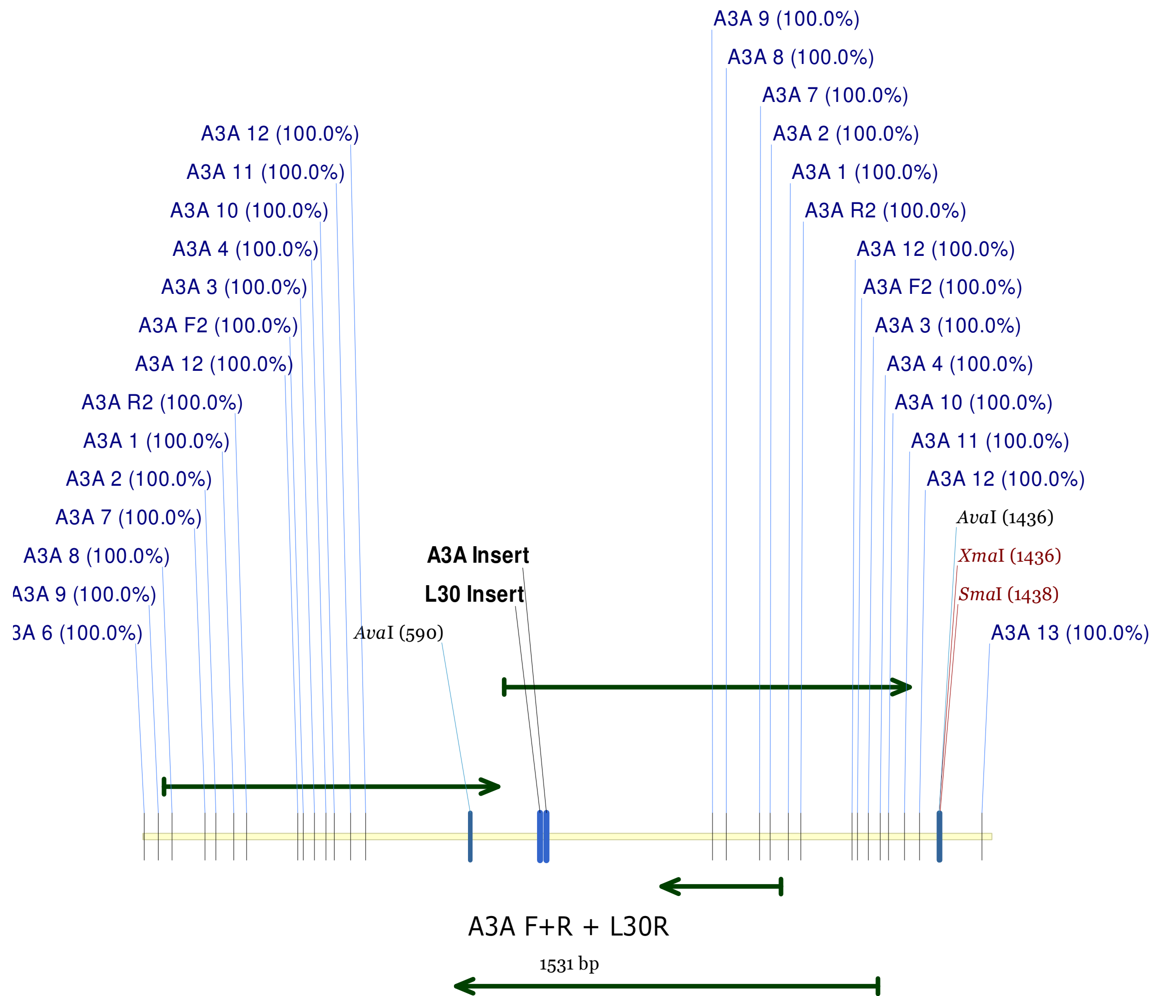


Figure 5. Feature map of A3A and L30 mutant

Green Arrows – Open Reading frames, Yellow line – DNA sequence,

Blue lines - Primers

Figure 5 is a sequence map drawn by Vector NTI of *E. cloacae* DNA from two selenite-sensitive mutants, A3A and L30. It appears that the transposome inserted itself into the same *E. cloacae* gene at two different positions. The yellow line represents the DNA sequence and the green arrows represent the sequences of DNA fragment that may encode a protein as predicted by Vector NTI. These sequences are called open reading frames (ORFs). The blue lines represent the primers that were designed or used for sequencing additional segments of the interrupted gene. A Vector NTI scan of the yellow sequence with different primers showed that the primers are homologous to the yellow sequence in more than one place. For instance, identical sequences for the A3A-1 primer can be found on both ends of the yellow sequence. Thus, the transposome has inserted itself into a repetitive DNA sequence. This repetition made it difficult to obtain additional sequence data on the A3A and L30 mutants. The feature map also shows restriction endonuclease recognition sites. The sites, *Xma*I, *Sma*I or *Ava*I, are places in the DNA that these particular restriction endonuclease cut the DNA. If the enzyme is red, then it only cuts the analyzed DNA sequence once. If the enzyme is black, it cuts the analyzed DNA sequence at more than one site. Thus, *Xma*I and *Sma*I cut the DNA sequence at only one place, and *Ava*I cuts the DNA sequence in at least two places.

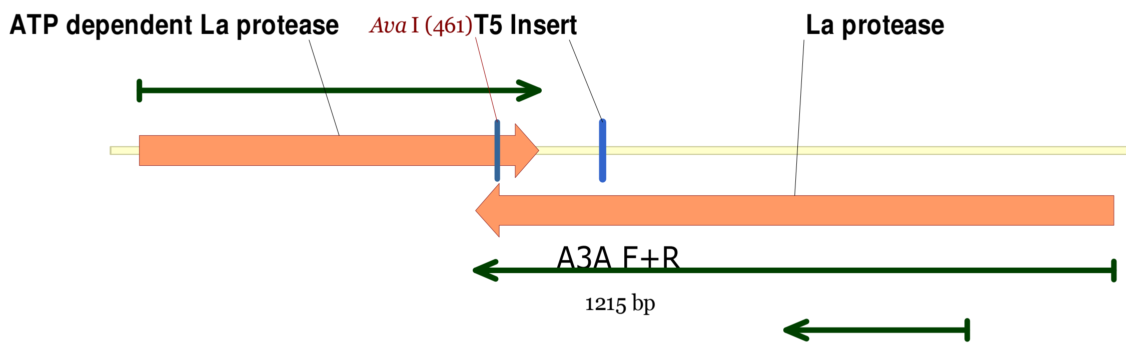


Figure 6. Feature map of A3A mutant

Green arrows – Open Reading frames, Yellow line – DNA sequence,

Solid orange arrow – Interrupted gene

Figure 6 is a feature map of the A3A DNA showing the open reading frames (solid orange arrows) that were translated into amino acid residue sequences and analyzed by BLAST. In addition, the nucleotide sequence of the A3A DNA with the open reading frames represented by the blue line is shown in Figure 7. BLAST analysis showed that 157 amino acid residues of ORFs appeared to be identical to a La protease in *E. coli*. Due to the repetitive nature of the DNA, additional reactions to obtain the sequence of the A3A and L30 mutant plasmids were not attempted.

```
                                    A3A F+R

  +1
   1  ATCCGCGGCC ATCGTCGCAC CTATATCGGT TCCATGCCGG GCAAGCTGAT CCAGAAGATG GCGAAAGTGG GGGTTAAAAA CCCGCTGTTC CTGCTCGATG
      TAGGCGCCGG TAGCAGCGTG GATATAGCCA AGGTACGGCC CGTTCGACTA GGTCTTCTAC CGCTTTCACC CCCAATTTTT GGGCGACAAG GACGAGCTAC
  +1
 101  AGATCGACAA AATGTCTTCG GACATGCGTG GCGATCCGGC GTCGGCGCTG CTGGAAGTGC TTGATCCAGA ACAGAACGTG GCGTTCAGCG ATCACTACCT
      TCTAGCTGTT TTACAGAAGC CTGTACGCAC CGCTAGGCCG CAGCCGCGAC GACCTTCACG AACTAGGTCT TGTCTTGCAC CGCAAGTCGC TAGTGATGGA
  +1
 201  GGAAGTGGAC TACGACCTGA GCGATGTGAT GTTCGTGGCG ACCTCCAACT CCATGAACAT TCCGGCCCCG CTGCTGGACC GTATGGAAGT GATTCGTCTC
      CCTTCACCTG ATGCTGGACT CGCTACACTA CAAGCACCGC TGGAGGTTGA GGTACTTGTA AGGCCGGGGC GACGACCTGG CATACCTTCA CTAAGCAGAG
  +1
 301  TCCGGTTATA CCGAAGATGA GAAGCTGAAC ATTGCTAAGC AGCACCTGCT GCCGAAACAG ATTGAGCGTA ACGCGCTGAA AGCCAACGAG CTGACCGTCG
      AGGCCAATAT GGCTTCTACT CTTCGACTTG TAACGATTCG TCGTGGACGA CGGCTTTGTC TAACTCGCAT TGCGCGACTT TCGGTTGCTC GACTGGCAGC
                                                                         AvaI
  +1
 401  AGGACAGCGC GATTATCGGC ATCATTCGCT ATTACACCCG TGAAGCGGGC GTTCGTAGCC TCGAGCGTGA GATCTCTAAG CTGTGCCGTA AGGCGGTGAA
      TCCTGTCGCG CTAATAGCCG TAGTAAGCGA TAATGTGGGC ACTTCGCCCG CAAGCATCGG AGCTCGCACT CTAGAGATTC GACACGGCAT TCCGCCACTT
  -2
  +1
 501  ACAGACTATC TGAACGAGCA AATGAAAGCC ATTCAGAAAG AGCTCGGCGA GATGGACGAT GCGCCGGATG AAAACGAAGC GCTGAAGGCG CTGAAGCGTA
      TGTCTGATAG ACTTGCTCGT TTACTTTCGG TAAGTCTTTC TCGAGCCGCT CTACCTGCTA CGCGGCCTAC TTTTGCTTCG CGACTTCCGC GACTTCGCAT
  -2
 601  AGATCGACGC GGCGAAGATG CCGAAAGAGG CGAAAGAGAA AGCCGAAGCT GAGTTGCAGA AGCTGAAAAT GATGTCTCCG ATGTCGGCAG AGGCGACCGT
      TCTAGCTGCG CCGCTTCTAC GGCTTTCTCC GCTTTCTCTT TCGGCTTCGA CTCAACGTCT TCGACTTTTA CTACAGAGGC TACAGCCGTC TCCGCTGGCA
  -2
 701  CGTGCGCGGC TATATTGAGT GGATGGTGCA GGTTCCGTGG AATGCCCGCA GCAAGGTCAA AAAAGACCTG CGTCAGGCGC AGGAAATCCT CGATACCGAC
      GCACGCGCCG ATATAACTCA CCTACCACGT CCAAGGCACC TTACGGGCGT CGTTCCAGTT TTTTCTGGAC GCAGTCCGCG TCCTTTAGGA GCTATGGCTG
  -2
 801  CACTACGGCC TGGAACGCGT GAAAGACCGC ATTCTTGAGT ACCTCGCGGT ACAGAGCCGT GTGAACAAAA TCAAAGGCCC AATCCTGTGC CTGGTAGGGC
      GTGATGCCGG ACCTTGCGCA CTTTCTGGCG TAAGAACTCA TGGAGCGCCA TGTCTCGGCA CACTTGTTTT AGTTTCCGGG TTAGGACACG GACCATCCCG
  -3
  -2
 901  CTCCAGGCGT GGGTAAAACC TCTCTGGGCC AGTCCATCGC CAAAGCGACC GGACGTAAGT ATATCCGTAT GGCGCTGGGC GGCGTACGCG ATGAAGCGGA
      GAGGTCCGCA CCCATTTTGG AGAGACCCGG TCAGGTAGCG GTTTCGCTGG CCTGCATTCA TATAGGCATA CCGCGACCCG CCGCATGCGC TACTTCGCCT
  -3
  -2
.001  AATCCGCGGC CATCGTCGCA CCTATATCGG TTCCATGCCG GGCAAGCTGA TCCAGAAGAT GGCGAAAGTG GGGGTTAAAA ACCCGCTGTT CCTGCTCGAT
      TTAGGCGCCG GTAGCAGCGT GGATATAGCC AAGGTACGGC CCGTTCGACT AGGTCTTCTA CCGCTTTCAC CCCCAATTTT TGGGCGACAA GGACGAGCTA
  -3
  -2
ay  June 24  2008 11:33:13 AM Page 1

                                    A3A F+R

1101  GAGATCGACA AAATGTCTTC GGACATGCGT GGCGATCCGG CGTCGGCGCT GCTGGAAGTG CTTGATCCAG AACAGAACGT GGCGTTCAGC GATCACTACC
      CTCTAGCTGT TTTACAGAAG CCTGTACGCA CCGCTAGGCC GCAGCCGCGA CGACCTTCAC GAACTAGGTC TTGTCTTGCA CCGCAAGTCG CTAGTGATGG
  -2
1201  TGGAAGTGGA CTACG
      ACCTTCACCT GATGC
```

Figure 7. Nucleotide sequence of A3A F + R mutant

```
>ref|ZP_02751147.1|  DNA-binding ATP-dependent protease La [Escherichia coli O157:H7
str. EC4206]
Length=736

 Score =  313 bits (801),  Expect = 3e-84, Method: Compositional matrix adjust.
 Identities = 152/157 (96%), Positives = 154/157 (98%), Gaps = 0/157 (0%)

Query  1    MPGKLIQKMAKVGVKNPLFLLDEIDKMSSDMRGDPASALLEVLDPEQNVAFSDHYLEVDY  60
            MPGKLIQKMAKVGVKNPLFLLDEIDKMSSDMRGDPASALLEVLDPEQNVAFSDHYLEVDY
Sbjct  354  MPGKLIQKMAKVGVKNPLFLLDEIDKMSSDMRGDPASALLEVLDPEQNVAFSDHYLEVDY  413

Query  61   DLSDVMFVATSNSMNIPAPLLDRMEVIRLSGYTEDEKLNIAKQHLLPKQIERNALKANEL  120
            DLSDVMFVATSNSMNIPAPLLDRMEVIRLSGYTEDEKLNIAK+HLLPKQIERNALK  EL
Sbjct  414  DLSDVMFVATSNSMNIPAPLLDRMEVIRLSGYTEDEKLNIAKRHLLPKQIERNALKKGEL  473

Query  121  TVEDSAIIGIIRYYTREAGVRSLEREISKLCRKAVKQ  157
            TV+DSAIIGIIRYYTREAGVR LEREISKLCRKAVKQ
Sbjct  474  TVDDSAIIGIIRYYTREAGVRGLEREISKLCRKAVKQ  510
```

Figure 8. Blast result of A3A mutant

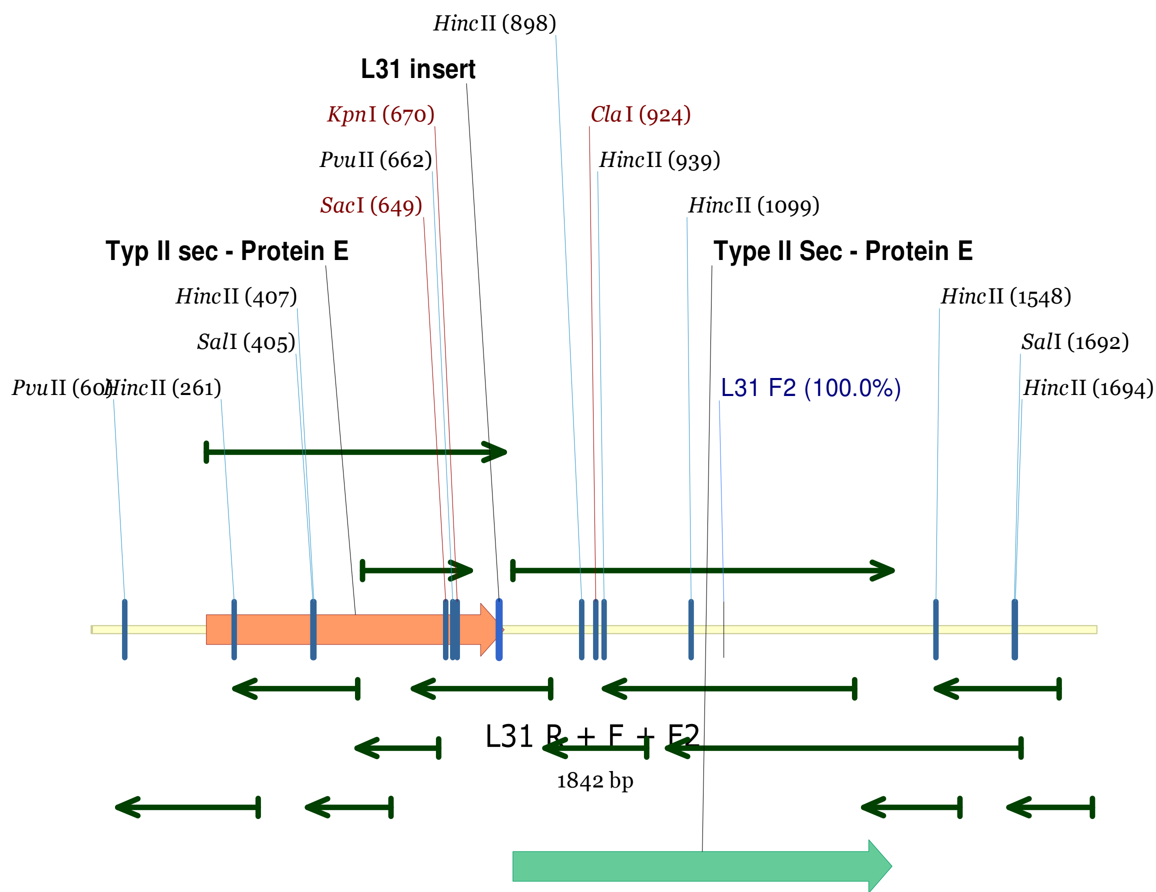


Figure 9. Feature map of L31 mutant

Green arrows – Open Reading frames, Yellow line – DNA sequence,

Solid orange and green arrows – Interrupted genes, Blue lines - Primers

Figure 9 is the feature map of DNA from the selenite sensitive L31 mutant showing two open reading frames that were translated (orange and green solid lines, respectively) into amino acid residue sequences. BLAST analysis of these polypeptide sequences suggested that they were related to an ATPase found in type II secretion complex (Fig 11). The nucleotide sequence of the L31 mutant is shown in Figure 10 along with the open reading frame sequences denoted by the blue arrows. The dark red line indicates the L31 F2 primer that was used to obtain additional sequence of the mutated DNA.

L31 R + F + F2

Figure 10. Nucleotide sequence of L31 R + F + R2

```
>ref|ZP_01064893.1|  Type II secretory pathway, ATPase EpsE [Vibrio sp. MED222]
 gb|EAQ53764.1|  Type II secretory pathway, ATPase EpsE [Vibrio sp. MED222]
Length=500

 Score =  241 bits (614),  Expect = 4e-62, Method: Compositional matrix adjust.
 Identities = 122/212 (57%), Positives = 161/212 (75%), Gaps = 9/212 (4%)

Query  8    KNALVSETVKANHLWHEVLDNA---AFEAEQMGKSFTVNFPPGPWPIDLMALTEEIPDNE  64
            KN    + + A+    ++ D     + EA Q+ +        G    D  +L EE+P +E
Sbjct  56   KNGFTPQAISADEFDKKLTDAYQRDSSEARQLMEDI------GADNDDFFSLAEELPQDE  109

Query  65   DLLDNDENSPVIRLINAILGEAVKDGASDIHIETFERTLSIRFRVDGVLRPVLQPARKLA  124
            DLL++++++P+I+LINA+LGEA+K+GASDIHIETFE++LSIRFR+DGVLR VL P+RKLA
Sbjct  110  DLLESEDDAPIIKLINAMLGEAIKEGASDIHIETFEKSLSIRFRIDGVLRDVLAPSRKLA  169

Query  125  PLLVSRIKVMSKLDIAEKRLPQDGRISLRIGRKAIDVRVSTIPSQYGERVVMRLLDKSNL  184
            PLLVSR+KVM+KLDIAEKR+PQDGRISLRIG +A+DVRVST+PS +GERVVMRLLDK+
Sbjct  170  PLLVSRVKVMAKLDIAEKRVPQDGRISLRIGGRAVDVRVSTMPSSHGERVVMRLLDKNAT  229

Query  185  KPDINKLGLIDEELEKLKGLIDRPHGIILVTG  216
            + D++ LG+  E  E  + LI RPHGIILVTG
Sbjct  230  RLDLHSLGMTAENHENFRKLIQRPHGIILVTG  261
```

Figure 11. Blast result of L31 R + F + R2 mutant

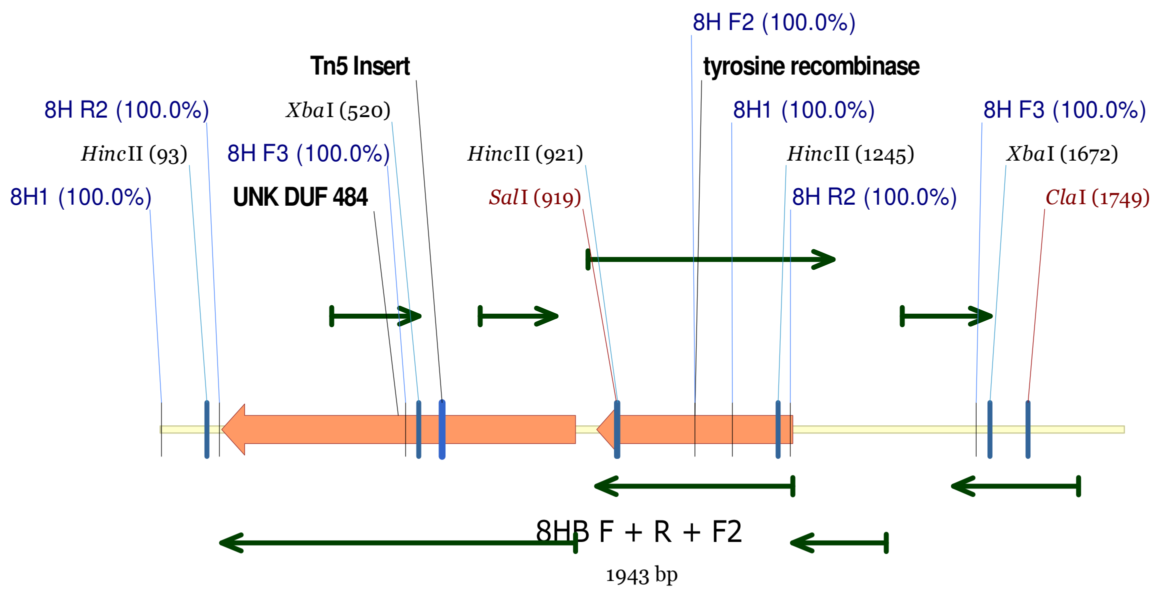


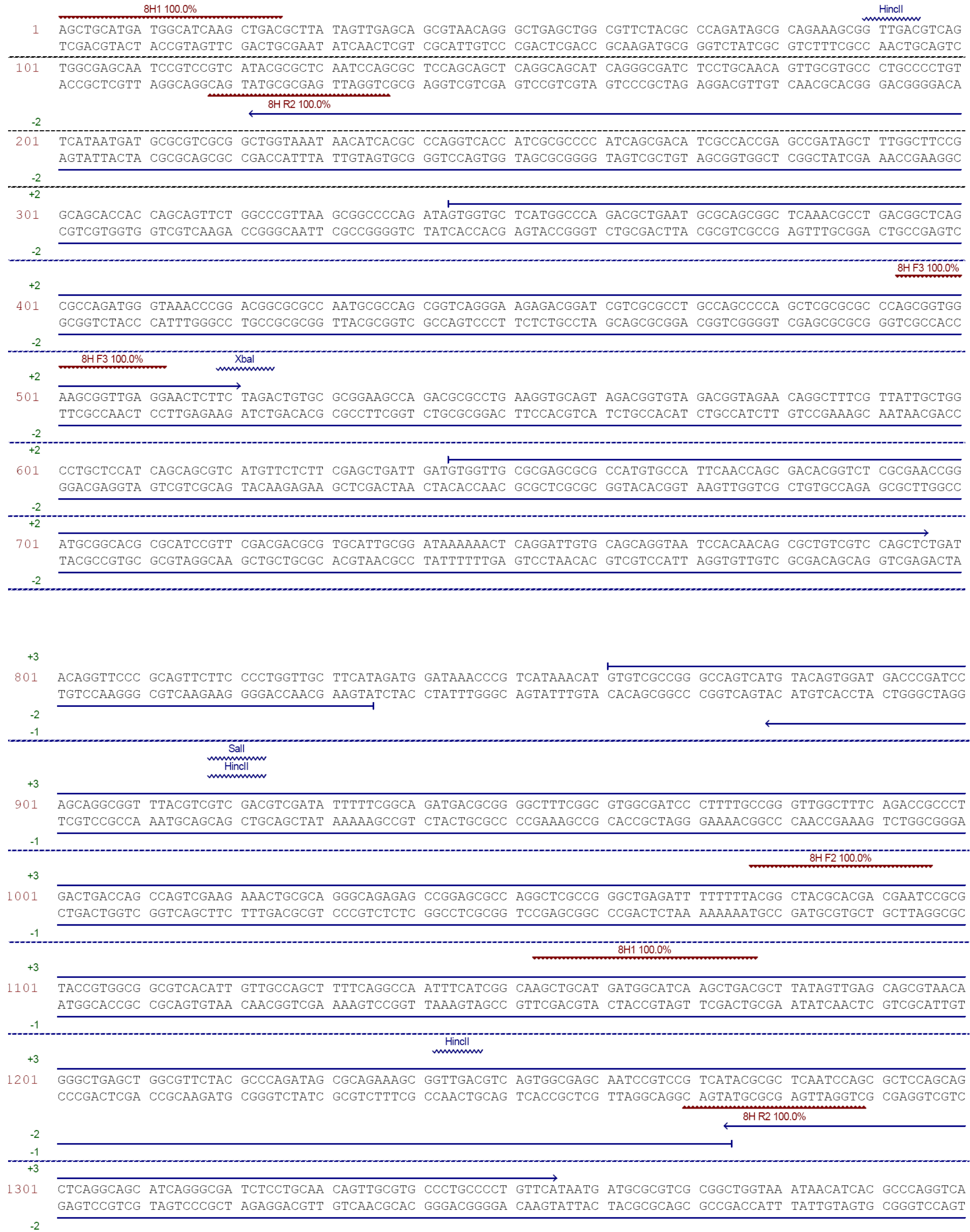
Figure 12. Feature map of 8HB Mutant

Green arrows – Open Reading frames, Yellow line – DNA sequence,

Solid orange arrow – Interrupted gene, Blue lines - Primers

Figure 12 represents the feature map of DNA from the selenite sensitive 8HB mutant showing all the restriction endonucleases, *Sal*I and *Cla*I in red, and *Hinc*II and *Xba*I in black. It shows all the possible ORFs in green arrows and the nucleotides of ORFs that were analyzed by BLAST. BLAST analysis showed that one was related to a DUF 484 protein of unknown function. The sequence of the second ORF was also of a protein of unknown function, however, BLAST analysis of the nucleotide sequence revealed that it may encode a XerC subunit belonging to the family of tyrosine recombinase (Figure 14). The nucleotide sequence of 8HB mutant is listed in Figure 13 along with the open reading frame sequences denoted by the blue arrows. The dark red lines indicate 8H 1, 8H F2 and 8H F3 primers that were used to obtain additional sequences of the mutated DNA. Additional reactions could not be performed on the 8HB mutant DNA due to the repetitive nature of the DNA sequence in that region.

8HB F + R + F2



```
1401   CCATCGCGCC CCATCAGCGA CATCGCCACC GAGCCGATAG CTTTGGCTTC CGGCAGCACC ACCAGCAGTT CTGGCCCGTT AAGCGGCCCC AGATAGTGGT
       GGTAGCGCGG GGTAGTCGCT GTAGCGGTGG CTCGGCTATC GAAACCGAAG GCCGTCGTGG TGGTCGTCAA GACCGGGCAA TTCGCCGGGG TCTATCACCA
  -2
  +2
1501   GCTCATGGCC CAGACGCTGA ATGCGCAGCG GCTCAAACGC CTGACGGCTC AGCGCCAGAT GGGTAAACCC GGACGGCGCG CCAATGCGCC AGCGGTCAGG
       CGAGTACCGG GTCTGCGACT TACGCGTCGC CGAGTTTGCG GACTGCCGAG TCGCGGTCTA CCCATTTGGG CCTGCCGCGC GGTTACGCGG TCGCCAGTCC
  -1
                                                                    8H F3 100.0%                Xbal
  +2
1601   GAAGAGACGG ATCGTCGCGC CTGCCAGCCC CAGCTCGCGC GCCCAGCGGT GGAAGCGGTT GAGGAACTCT TCTAGACTGT GCGCGGAAGC CAGACGCGCC
       CTTCTCTGCC TAGCAGCGCG GACGGTCGGG GTCGAGCGCG CGGGTCGCCA CCTTCGCCAA CTCCTTGAGA AGATCTGACA CGCGCCTTCG GTCTGCGCGG
  -1
                                                     ClaI
1701   TGAAGGTGCA GTAGACNGTC TGTCTCTTAT ACACATCTCA ACCATCATCG ATGANTNGCT TCGTNAATAC AGATGTAGGT GTNNCACAGG GTAGCCAGCA
       ACTTCCACGT CATCTGNCAG ACAGAGAATA TGTGTAGAGT TGGTAGTAGC TACTNANCGA AGCANTTATG TCTACATCCA CANNGTGTCC CATCGGTCGT
  -1
1801   GCATCCTGCG ATGCAGATNC GGATGCCATT TCATNACCTC TTTCTCGCAC CNGACATNNA TCNGANGATC AGCAGTTCAN CNTGTNGATA GTACGTACTA
       CGTAGGACGC TACGTCTANG CCTACGGTAA AGTANTGGAG AAAGAGCGTG GNCTGTANNT AGNCTNCTAG TCGTCAAGTN GNACANCTAT CATGCATGAT
  -1
1901   AGCTCTCATG TNNCACGNAC TANGCNCTCA TGTTNANCGT ACT
       TCGAGAGTAC ANNGTGCNTG ATNCGNGAGT ACAANTNGCA TGA
```

Figure 13. Nucleotide sequence of 8HB F + R + R2 mutant

```
>gb|CP000653.1|  Enterobacter sp. 638, complete genome
Length=4518712

                                                          Sort alignments for this subject sequence by:
                                                            E value  Score  Percent identity
                                                            Query start position  Subject start position

 Features in this part of subject sequence:
   tyrosine recombinase XerC subunit
   protein of unknown function DUF484

 Score =  913 bits (494),  Expect = 0.0
 Identities = 775/907 (85%), Gaps = 33/907 (3%)
 Strand=Plus/Plus

Query  447      GCTGCATGATGGCAT-CAAGCTGACGCTTATAGTTGAGCAGCGTAA-CAGGGCTGAGCTG  504
                |||||||||| |||| ||| ||| |||| |||||| |||||||||| | || ||||||||
Sbjct  4315982  GCTGCATGATCGCATCCAA-CTGGCGCTGATAGTTCAGCAGCGTAATC-GGACTGAGCTG  4316039

Query  505      GCGTTCTACGCCCAGATAGCGCAGAAAGCG-GTTGACGTCAGTG-GC---GAGCAATCCG  559
                |||||| | ||| ||||| |||||||| || |  ||||  || | ||   |||  |  ||
Sbjct  4316040  GCGTTCGATGCCGAGATAACGCAGAAAACGCG-CGACG-AAG-GAGCTAAGAG-GA-GCG  4316094

Query  560      TCCGTCATACGCGCTCAATCCAGCGCTCCAGCAGCTCAGGCAGCATCAGGGCGATCTCCT  619
                |||||||||||||||||||||||||||||||||||||||||||||| |||||||| ||||
Sbjct  4316095  TCCGTCATACGCGCTCAATCCAGCGCTCCAGCAGCTCAGGCAGCATTAGGGCGATTTCCT  4316154

Query  620      GCAACAGTTGCGTGCCCTGCCCCTGTTCATAATGATGCGCGTCGCGGCTGGTAAATAACA  679
                ||| ||  |||||||||||||||||||| || || ||||||||||||||||| |||||||
Sbjct  4316155  GCAGCAAATGCGTGCCCTGCCCCTGTTCGTAGTGGTGCGCGTCGCGGCTGGTGAATAACA  4316214

Query  680      TCACGCCCAGGTCACCATCGCGCCCCATCAGCGACATCGCCACCGAGCCGATAGCTTTGG  739
                | |||||||| || || ||||||||||||||||||||||||||| |||||  || |||||||
Sbjct  4316215  TTACGCCCAGATCCCCTTCGCGCCCCATCAGCGACATCGCCACTGAGCCAATCGCTTTGG  4316274

Query  740      CTTCCGGCAGCACCACCAGCAGTTCTGGCCCGTTAAGCGGCCCCAGATAGTGGTGCTCAT  799
                |||||||||| ||||||||||||||| || ||||| || |||| |||||||||| |||||||
Sbjct  4316275  CTTCCGGCAGTACCACCAGCAGTTCAGGTCCGTTCAGTGGCCCAAGATAGTGATGCTCAT  4316334

Query  800      GGCCCA-GACGCTGAATGCGCAGCGGCTCAAACGCCTGACGGCTCA-GCGCCAGATGGGT  857
                | |||| ||||||| |||||||||| |||||||  |||||  ||| | |||||||||||
Sbjct  4316335  G-CCCAAGACGCTGGATGCGCAGCGGTTCAAACGTTTGACGATTCAAG-GCCAGATGGGT  4316392

Query  858      AAACCCGGACGGCGCGCCAATGCGCCAGCGGTCAGGGAAGAGACGGATCGTCGCGCCTGC  917
                ||| || |||||||||||| ||||||||||| || ||||| ||||| || ||||| || ||
Sbjct  4316393  AAAACCAGACGGCGCGCCGATGCGCCAGCGATCGGGGAACAGACGAATGGTCGCTCCCGC  4316452

Query  918      CAGCCCCAGCTCGCGCGCCCAGCGGTGGAAGCGGTTGAGGAACTCTTCTAGACTGTGCGC  977
                ||| || |  |||||||||||||||||||| ||| |||| |||||||| || ||||||||
Sbjct  4316453  CAGGCCTATTTCGCGCGCCCAGCGGTGGAAACGGCTGAGAAACTCTTCCAGGCTGTGCGC  4316512

Query  978      GGAAGCCAGACGCGCCTG-AAGGTGCAGTAGACGGTGTAGACGGTAGAACAGGCTTTCGT  1036
                 || || || |||||||| ||  | ||| ||||| |  | |    | | |||||||||||
Sbjct  4316513  CGACGCGAGCCGCGCCTGCAAATT-CAGCAGACGAT--A-A----A-A-CAGGCTTTCGT  4316562

Query  1037     TATTGCTGGCCTGCTCCATCAGCAGCGTCATGTTCTCTTCGAGCTGATTGATGTGGTTGC  1096
                ||||||||||||| |||||||||||||||||||||| ||||| ||||||| |||||| || |
Sbjct  4316563  TATTGCTGGCCTGTTCCATCAGCAGCGTCATGTTTTCTTCAAGCTGATCGATGTGTTTAC  4316622

Query  1097     GCGAGCGCGCCATGTGCCATTCAACCAGCGACACGGTCTCGCGAACCGGATGCGGCACGC  1156
                ||||||||||||||||||||||||||||||||||||| |||||| || |||||||||||||
Sbjct  4316623  GCGAGCGCGCCATGTGCCATTCAACCAGCGACACGGTATCGCGCACGGGATGCGGCACGC  4316682

Query  1157     GCATCCGTTCGACGACGCGTGCATTGCGGATAAAAAACTCAGGATTGTGCAGCAGGTAAT  1216
                ||||  |||| || ||||||||||||||||||||||||||||||||| |||||| ||||
Sbjct  4316683  GCATTTGTTCAACCACGCGTGCATTGCGGATAAAAAACTCAGGATTGCGCAGCAGATAAT  4316742

Query  1217     CCACAACAGCGCTGTCGTCCAGCTCTGATACAGG-TTCCC-GCAGTTCTTCCCCTGGTTG  1274
                | |||||||| | ||| |||| || | | ||   || || |||||||||||||||||||
Sbjct  4316743  CAACAACAGCCCGGTCATCCAGTTCCG-T-CATTATTTCCTGCAGTTCTTCCCCTGGTTG  4316800

Query  1275     CTTCATAGATGGATAAACCCGTCATAAACATGTGTCGCCGGGCCAGTCATGTACAGTGGA  1334
                 ||||||||||||||||||||||||||||||||| |||||||||||||||  |||| |||
Sbjct  4316801  TTTCATAGATGGATAAACCCGTCATAAACATGTGCCGCCGGGCCAGTCATAAACAGCGGA  4316860

Query  1335     TGACCCG  1341
                |||||||
Sbjct  4316861  TGACCCG  4316867
```

Figure 14. Blast result of 8HB F + R + F2 mutant

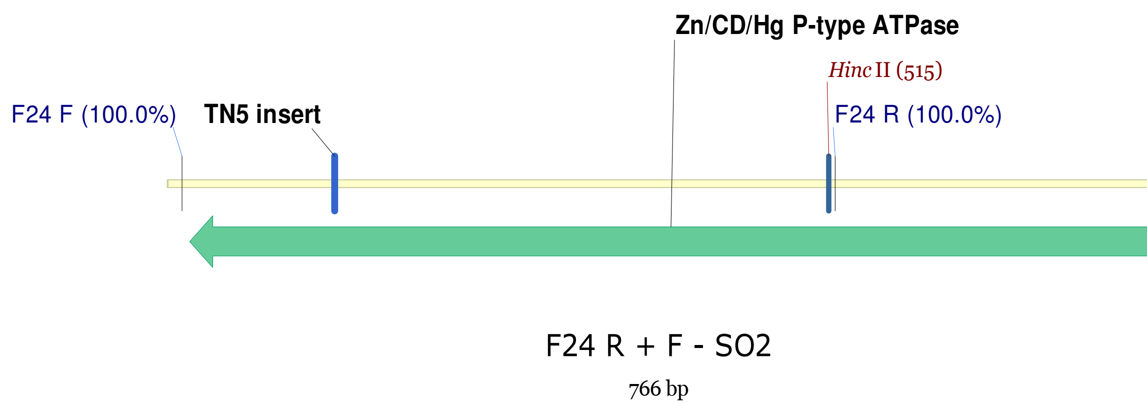


Figure 15. Feature map of F24 R + F mutant

Yellow line – DNA sequence, Solid green arrow – Interrupted gene,

Blue lines - Primers

Figure 15 represents the feature map of the zinc/cadmium/selenite sensitive F24 mutant along with one ORF (solid green arrow) which was translated into an amino acid residues sequence and analyzed by BLAST. BLAST analysis revealed that the putative protein was similar to a zinc, cadmium and mercury P-type ATPase transport system belonging to *Enterobacter* sp 638 (Figure 17). The nucleotide sequence of F24 mutant is listed in Figure 16 along with primer, F24 F and F24 R, represented a red line and used to resolve larger segments of the sequence.

```
                              F24 R + F - SO2

             F24 F 100.0%
  1  CCACCGGGCA TTTTTACCGA CGGCGCAACA GCCGCAACGC GTTCGCCGTA ACCAGCACCG TCGCCCCCGT ATCCGCCAGC ACCGCCAGCC ACAGGCCGGT
     GGTGGCCCGT AAAAATGGCT GCCGCGTTGT CGGCGTTGCG CAAGCGGCAT TGGTCGTGGC AGCGGGGGCA TAGGCGGTCG TGGCGGTCGG TGTCCGGCCA
----------------------------------------------------------------------------------------------------------------
101  CATGCCGAGC AGCGTGGTGA CGAGGAATAT GAGGAATATC CCCTTCAGCC CCAGCGCAAT CCCAATGTTC TGGCGGATGT TGGCGCGCGT CGCCCGCGCC
     GTACGGCTCG TCGCACCACT GCTCCTTATA CTCCTTATAG GGGAAGTCGG GGTCGCGTTA GGGTTACAAG ACCGCCTACA ACCGCGCGCA GCGGGCGCGG
----------------------------------------------------------------------------------------------------------------
201  AGGCTAATCA TCTGCGCCAG CCCGGTCAGG CGGTTGTGGG TCAGCGCCGC ATCCGCCGTC TCCAGCGCCA CATCGGTGCC GCTGCCCATC GCAATGCCGA
     TCCGATTAGT AGACGCGGTC GGGCCAGTCC GCCAACACCC AGTCGCGGCG TAGGCGGCAG AGGTCGCGGT GTAGCCACGG CGACGGGTAG CGTTACGGCT
----------------------------------------------------------------------------------------------------------------
301  TGGTGGAAGC CTTCATTGCC GGGGCATCGT TTATCCCGTC CCCGACCATC GCCAGCGGGG CATGACCGTT CAGCTCGGTT ACCGCGCTGA CTTTATCCGC
     ACCACCTTCG GAAGTAACGG CCCCGTAGCA AATAGGGCAG GGGCTGGTAG CGGTCGCCCC GTACTGGCAA GTCGAGCCAA TGGCGCGACT GAAATAGGCG
----------------------------------------------------------------------------------------------------------------
401  AGGCAACAAT CCGGCCTTAA ACTCCAGCCC CAGCTCGCCG GCAATCGCCG CCGCCGCACG CGGGTTATCA CCGGTCAGGA TCACGCCCTG TACCCCCAGC
     TCCGTTGTTA GGCCGGAATT TGAGGTCGGG GTCGAGCGGC CGTTAGCGGC GGCGGCGTGC GCCCAATAGT GGCCAGTCCT AGTGCGGGAC ATGGGGGTCG
----------------------------------------------------------------------------------------------------------------
               HincII
501  CTGTGCAGCG CGTCGACGGC GTCTTTCGCA TCGTCACGCA GGGTGTCGCG CAGCGCCAGC ATGCCCATCG GTACACCGTC CTGCATCACG GCAACGACCG
     GACACGTCGC GCAGCTGCCG CAGAAAGCGT AGCAGTGCGT CCCACAGCGC GTCGCGGTCG TACGGGTAGC CATGTGGCAG GACGTAGTGC CGTTGCTGGC
                             F24 R 100.0%
----------------------------------------------------------------------------------------------------------------
601  TCTGCCCGGC CTGCTCTAAT GCCTCAACCT GTGGATNCGG GAACGCGCCT GCCGCGGCGA TCAGCACTTT TTTGCCTTCG ACCACAGCCT CAATCCCTGA
     AGACGGGCCG GACGAGATTA CGGAGTTGGA CACCTANGCC CTTGCGCGGA CGGCGCCGCT AGTCGTGAAA AAACGGAAGC TGGTGTCGGA GTTAGGGACT
----------------------------------------------------------------------------------------------------------------
701  CCCGACCAGC GCTCGCTGGG CGGTTGCCGG AGGGATGGTC AGCCCGCGCC CCTNCGCTTC ACGCAC
     GGGCTGGTCG CGAGCGACCC GCCAACGGCC TCCCTACCAG TCGGGCGCGG GGANGCGAAG TGCGTG
----------------------------------------------------------------------------------------------------------------
```

Figure 16. Nucleotide sequence of F24 F + R mutant

```
>ref|YP_001178579.1| G zinc/cadmium/mercury/lead-transporting ATPase [Enterobacter sp.
638]
 gb|ABP62528.1| G heavy metal translocating P-type ATPase [Enterobacter sp. 638]
Length=728

 GENE ID: 5110600 zntA | zinc/cadmium/mercury/lead-transporting ATPase
[Enterobacter sp. 638]

 Score =  396 bits (1018),  Expect = 6e-109, Method: Compositional matrix adjust.
 Identities = 212/254 (83%), Positives = 225/254 (88%), Gaps = 7/254 (2%)

Query  1    VREAXGRGLTIPPATAQRALVGSGIEAVVEGKKVLIAAAGAFPXP----QVEALEQAGQT  56
            VREA  RGLT+P A+ QRALVGSGIEA V+GKKVLI AA  FP      Q+  LEQAGQT
Sbjct  478  VREAQSRGLTLPAASEQRALVGSGIEADVDGKKVLITAADKFPSQALGQQISELEQAGQT  537

Query  57   VVAVMQDGVPMGMLALRDTLRDDAKDAVDALHRLGVQGVILTGDNPRAAAAIAGELGLEF  116
            V+ V  DGV  G+LALRDTLRDDAK+AV ALH+LG+QGVILTGDNPRAAAAIA ELGLEF
Sbjct  538  VIIVAVDGVAKGVLALRDTLRDDAKEAVAALHQLGIQGVILTGDNPRAAAAIAHELGLEF  597

Query  117  KAGLLPADKVSAVTELNGHAPLAMVGDGINDAPAMKASTIGIAMGSGTDVALETADAALT  176
            KAGLLPADKV AVT+LNG APLAMVGDGINDAPAMKASTIGIAMGSGTDVALETADAALT
Sbjct  598  KAGLLPADKVQAVTKLNGQAPLAMVGDGINDAPAMKASTIGIAMGSGTDVALETADAALT  657

Query  177  HNRLTGLAQMISLARATRANIRQNIGIALGLKGIFLIFLVTTLLGMTGLWLAVLADTGAT  236
            HNRLTGLAQMI LARATRANIRQNIGIALGLKG   IFLVTTLLG+TGLWLAVLADTGAT
Sbjct  658  HNRLTGLAQMIGLARATRANIRQNIGIALGLKG---IFLVTTLLGITGLWLAVLADTGAT  714

Query  237  VLVTANALRLLRRR  250
            VLVTANALRLLR++
Sbjct  715  VLVTANALRLLRKK  728
```

Figure 17. Blast result of F24 F + R mutant

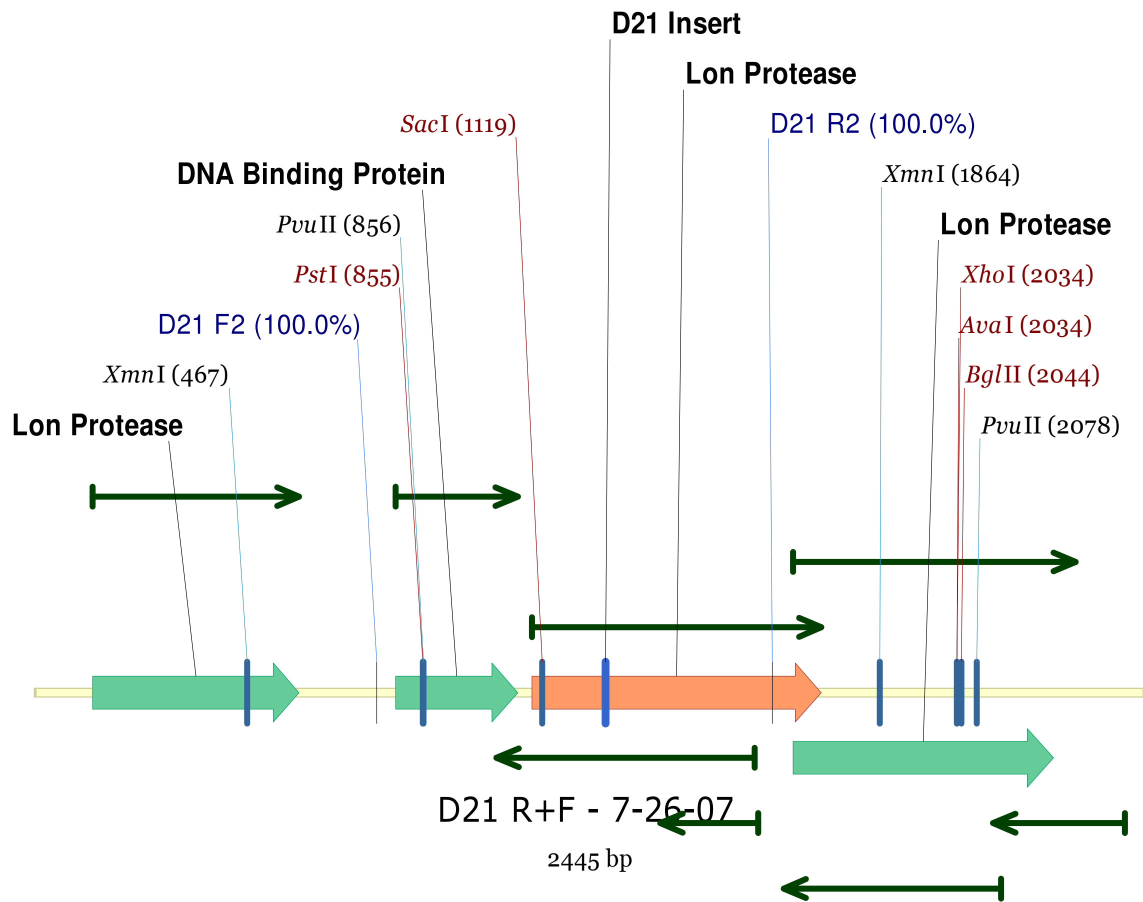


Figure 18. Feature map of D21 F +R mutant

Green arrows – Open Reading frames, Yellow line – DNA sequence,

Solid orange and green arrows – Interrupted gene, Blue lines - Primers

Figure 18 represents a feature map of the selenite sensitive D21 mutant showing the restriction sites, *Pst*I, *Bgl*II, *Xmn*I and *Pvu*II and the ORFs. The nucleotide sequence of ORFs (solid orange and green arrows) analyzed by BLAST was identical to be similar to a Lon protease gene of an *E. coli* is shown in Figure 20. Another ORF (smallest solid green arrow) of D21 mutant translated into amino acid residue, analyzed by BLAST was suggested to be a DNA binding protein of an *E. coli*. The nucleotide sequence of D21 mutant is listed in Figure 19 along with the open reading frames denoted by a blue arrow. The dark red line indicates the D21 F2 primer that was used to obtain additional sequence information of the mutated DNA.

D21 R+F - 7-26-07

```
1    AAACCCGCGN TNGNTCAGNT GACCCGNTCT NGCATGGACG GAANGNGNCG GGCGATCTGC TGACCATCNN ACCCGCCTGT GTNCCGGGCA AAGGCNAGCT
     TTTGGGCGCN ANCNAGTCNA CTGGGCNAGA NCGTACCTGC CTTNCNCNGC CCGCTAGACG ACTGGTAGNN TGGGCGGACA CANGGCCCGT TTCCGNTCGA

+3
101  GACCTACACC GGTTCACTGG NTGAAGTGAT GCAGGAGTCC ATTCAGGCGG CGCTGACCGT GGTGCGCGCG CGCGCGGAAA AACTGGGTAT CAACCCGGAT
     CTGGATGTGG CCAAGTGACC NACTTCACTA CGTCCTCAGG TAAGTCCGCC GCGACTGGCA CCACGCGCGC GCGCGCCTTT TTGACCCATA GTTGGGCCTA

+3
201  TTCTACGAAA AACGCGACAT TCACGTTCAC GTTCCGGAAG GGGCGACGCC GAAAGATGGC CCAAGCGCGG GTATTGCCAT GTGTACGGCG CTGGTTTCCT
     AAGATGCTTT TTGCGCTGTA AGTGCAAGTG CAAGGCCTTC CCCGCTGCGG CTTTCTACCG GGTTCGCGCC CATAACGGTA CACATGCCGC GACCAAAGGA

+3
301  GCCTGACAGG GAACCCGGTT CGTGCTGATG TGGCCATGAC CGGTGAAATT ACGCTGCGTG GTCAGGTACT GCCTATCGGC GGGTTAAAAG AAAAACTGCT
     CGGACTGTCC CTTGGGCCAA GCACGACTAC ACCGGTACTG GCCACTTTAA TGCGACGCAC CAGTCCATGA CGGATAGCCG CCCAATTTTC TTTTTGACGA

                                                                        XmnI
+3
401  GGCGGCACAC CGCGGTGGGA TCAAAACCGT ATTGATCCCT TACGAAAATA AACGCGATCT GGAAGAGATT CCGGACAACG TTATTGCCGA TTTGCAGATC
     CCGCCGTGTG GCGCCACCCT AGTTTTGGCA TAACTAGGGA ATGCTTTTAT TTGCGCTAGA CCTTCTCTAA GGCCTGTTGC AATAACGGCT AAACGTCTAG

+3
501  CATCCGGTGA AACGCATTGA GGAAGTTCTG ACTCTCGCAC TGCAAAATGA ACCGTCCGGA ATGCAGGTTG TAACCGCAAA ATAGTGACCT CGCGCAAAGA
     GTAGGCCACT TTGCGTAACT CCTTCAAGAC TGAGAGCGTG ACGTTTTACT TGGCAGGCCT TACGTCCAAC ATTGGCGTTT TATCACTGGA GCGCGTTTCT

601  GCGCCAATAA AAACAAGGCT GGTAAGTCAT TTCGTACTTG CCAGCCTTTT TTTGTATAGC TAATTTAGAT TGCTGATTGG GTTAGCCATC AACAACGGGT
     CGCGGTTATT TTTGTTCCGA CCATTCAGTA AAGCATGAAC GGTCGGAAAA AAACATATCG ATTAAATCTA ACGACTAACC CAATCGGTAG TTGTTGCCCA

+3
701  GTTGTAAGGT CATAGCAGGA CTGATATAAC TGCTGCGCGG TCGCGCTGTG AAGGATTCAG GTGCGATATA AATTATAAAG AGAGGAAGAG AACAGTGAAT
     CAACATTCCA GTATCGTCCT GACTATATTG ACGACGCGCC AGCGCGACAC TTCCTAAGTC CACGCTATAT TTAATATTTC TCTCCTTCTC TTGTCACTTA
                                                            D21 F2 100.0%

                                                          PvuII
                                                        PstI
+3
801  AAATCTCAAC TGATTGACAA AATTGCTGCG GGTGCTGATA TTTCTAAAGC TGCAGCTGGA CGTGCTTTAG ATGCTTTAAT TGCTTCTGTT ACTGAATCTC
     TTTAGAGTTG ACTAACTGTT TTAACGACGC CCACGACTAT AAAGATTTCG ACGTCGACCT GCACGAAATC TACGAAATTA ACGAAGACAA TGACTTAGAG

+3
901  TGCAAGCTGG GGACGACGTT GCACTGGTAG GCTTCGGTAC TTTTGCGGTT AAAGAGCGTG CAGCCCGTAC TGGCCGCAAC CCTCAAACCG GTAAAGAGAT
     ACGTTCGACC CCTGCTGCAA CGTGACCATC CGAAGCCATG AAAACGCCAA TTTCTCGCAC GTCGGGCATG ACCGGCGTTG GGAGTTTGGC CATTTCTCTA

+3
+1
1001 CACCATTGCT GCCGCTAAAG TGCCGGGTTT CCGTGCAGGT AAAGCGCTGA AAGACGCAGT AAACTGATTG CTTTCCCGAC TATCTGAACG AGCAAATGAA
     GTGGTAACGA CGGCGATTTC ACGGCCCAAA GGCACGTCCA TTTCGCGACT TTCTGCGTCA TTTGACTAAC GAAAGGGCTG ATAGACTTGC TCGTTTACTT
-1

                SacI
+1
1101 AGCCATTCAG AAAGAGCTCG GCGAGATGGA CGATGCGCCG GATGAAAACG AAGCGCTGAA GCGTAAGATC GACGCGGCGA AGATGCCGAA AGAGGCGAAA
     TCGGTAAGTC TTTCTCGAGC CGCTCTACCT GCTACGCGGC CTACTTTTGC TTCGCGACTT CGCATTCTAG CTGCGCCGCT TCTACGGCTT TCTCCGCTTT
-1

+1
1201 GAGAAAGCCG AAGCTGAGTT GCAGAAGCTG AAAATGATGT CTCCGATGTC GGCAGAGGCG GCAGAGGCGA CCGTCGTGCG CGGCTATATT GAGTGGATGG
     CTCTTTCGGC TTCGACTCAA CGTCTTCGAC TTTTACTACA GAGGCTACAG CCGTCTCCGC CGTCTCCGCT GGCAGCACGC GCCGATATAA CTCACCTACC
-1

+1
1301 TGCAGGTTCC GTGGAATGCC CGCAGCAAGG TCAAAAAAGA CCTGCGTCAG GCGCAGGAAA TCCTCGATAC CGACCACTAC GGCCTGGAAC GCGTGAAAGA
     ACGTCCAAGG CACCTTACGG GCGTCGTTCC AGTTTTTTCT GGACGCAGTC CGCGTCCTTT AGGAGCTATG GCTGGTGATG CCGGACCTTG CGCACTTTCT
-2
-1

+1
1401 CCGCATTCTT GAGTACCTCG CGGTACAGAG CCGTGTGAAC AAAATCAAAG GCCCAATCCT GTGCCTGGTA GGGCCTCCAG GCGTGGGTAA AACCTCTCTG
     GGCGTAAGAA CTCATGGAGC GCCATGTCTC GGCACACTTG TTTTAGTTTC CGGGTTAGGA CACGGACCAT CCCGGAGGTC CGCACCCATT TTGGAGAGAC
-2
-1

+1
1501 GGCCAGTCCA TCGCCAAAGC GACCGGACGT AAGTATATCC GTATGGCGCT GGGCGGCGTA CGCGATGAAG CGGAAATCCG CGGCCATCGT CGCACCTATA
     CCGGTCAGGT AGCGGTTTCG CTGGCCTGCA TTCATATAGG CATACCGCGA CCCGCCGCAT GCGCTACTTC GCCTTTAGGC GCCGGTAGCA GCGTGGATAT
-2
-1
```

D21 R+F - 7-26-07

```
1     AAACCCGCGN TNGNTCAGNT GACCCGNTCT NGCATGGACG GAANGNGNCG GGCGATCTGC TGACCATCNN ACCCGCCTGT GTNCCGGGCA AAGGCNAGCT
      TTTGGGCGCN ANCNAGTCNA CTGGGCNAGA NCGTACCTGC CTTNCNCNGC CCGCTAGACG ACTGGTAGNN TGGGCGGACA CANGGCCCGT TTCCGNTCGA
+3
101   GACCTACACC GGTTCACTGG NTGAAGTGAT GCAGGAGTCC ATTCAGGCGG CGCTGACCGT GGTGCGCGCG CGCGCGGAAA AACTGGGTAT CAACCCGGAT
      CTGGATGTGG CCAAGTGACC NACTTCACTA CGTCCTCAGG TAAGTCCGCC GCGACTGGCA CCACGCGCGC GCGCGCCTTT TTGACCCATA GTTGGGCCTA
+3
201   TTCTACGAAA AACGCGACAT TCACGTTCAC GTTCCGGAAG GGGCGACGCC GAAAGATGGC CCAAGCGCGG GTATTGCCAT GTGTACGGCG CTGGTTTCCT
      AAGATGCTTT TTGCGCTGTA AGTGCAAGTG CAAGGCCTTC CCCGCTGCGG CTTTCTACCG GGTTCGCGCC CATAACGGTA CACATGCCGC GACCAAAGGA
+3
301   GCCTGACAGG GAACCCGGTT CGTGCTGATG TGGCCATGAC CGGTGAAATT ACGCTGCGTG GTCAGGTACT GCCTATCGGC GGGTTAAAAG AAAAACTGCT
      CGGACTGTCC CTTGGGCCAA GCACGACTAC ACCGGTACTG GCCACTTTAA TGCGACGCAC CAGTCCATGA CGGATAGCCG CCCAATTTTC TTTTTGACGA
                                                                   XmnI
+3
401   GGCGGCACAC CGCGGTGGGA TCAAAACCGT ATTGATCCCT TACGAAAATA AACGCGATCT GGAAGAGATT CCGGACAACG TTATTGCCGA TTTGCAGATC
      CCGCCGTGTG GCGCCACCCT AGTTTTGGCA TAACTAGGGA ATGCTTTTAT TTGCGCTAGA CCTTCTCTAA GGCCTGTTGC AATAACGGCT AAACGTCTAG
+3
501   CATCCGGTGA AACGCATTGA GGAAGTTCTG ACTCTCGCAC TGCAAAATGA ACCGTCCGGA ATGCAGGTTG TAACCGCAAA ATAGTGACCT CGCGCAAAGA
      GTAGGCCACT TTGCGTAACT CCTTCAAGAC TGAGAGCGTG ACGTTTTACT TGGCAGGCCT TACGTCCAAC ATTGGCGTTT TATCACTGGA GCGCGTTTCT
601   GCGCCAATAA AAACAAGGCT GGTAAGTCAT TTCGTACTTG CCAGCCTTTT TTTGTATAGC TAATTTAGAT TGCTGATTGG GTTAGCCATC AACAACGGGT
      CGCGGTTATT TTTGTTCCGA CCATTCAGTA AAGCATGAAC GGTCGGAAAA AAACATATCG ATTAAATCTA ACGACTAACC CAATCGGTAG TTGTTGCCCA
+3
701   GTTGTAAGGT CATAGCAGGA CTGATATAAC TGCTGCGCGG TCGCGCTGTG AAGGATTCAG GTGCGATATA AATTATAAAG AGAGGAAGAG AACAGTGAAT
      CAACATTCCA GTATCGTCCT GACTATATTG ACGACGCGCC AGCGCGACAC TTCCTAAGTC CACGCTATAT TTAATATTTC TCTCCTTCTC TTGTCACTTA
                                                             D21 F2 100.0%
                                                        PvuII
                                                    PstI
+3
801   AAATCTCAAC TGATTGACAA AATTGCTGCG GGTGCTGATA TTTCTAAAGC TGCAGCTGGA CGTGCTTTAG ATGCTTTAAT TGCTTCTGTT ACTGAATCTC
      TTTAGAGTTG ACTAACTGTT TTAACGACGC CCACGACTAT AAAGATTTCG ACGTCGACCT GCACGAAATC TACGAAATTA ACGAAGACAA TGACTTAGAG
+3
901   TGCAAGCTGG GGACGACGTT GCACTGGTAG GCTTCGGTAC TTTTGCGGTT AAAGAGCGTG CAGCCCGTAC TGGCCGCAAC CCTCAAACCG GTAAAGAGAT
      ACGTTCGACC CCTGCTGCAA CGTGACCATC CGAAGCCATG AAAACGCCAA TTTCTCGCAC GTCGGGCATG ACCGGCGTTG GGAGTTTGGC CATTTCTCTA

+3
+1
1001  CACCATTGCT GCCGCTAAAG TGCCGGGTTT CCGTGCAGGT AAAGCGCTGA AAGACGCAGT AAACTGATTG CTTTCCCGAC TATCTGAACG AGCAAATGAA
      GTGGTAACGA CGGCGATTTC ACGGCCCAAA GGCACGTCCA TTTCGCGACT TTCTGCGTCA TTTGACTAAC GAAAGGGCTG ATAGACTTGC TCGTTTACTT
-1
                SacI
+1
1101  AGCCATTCAG AAAGAGCTCG GCGAGATGGA CGATGCGCCG GATGAAAACG AAGCGCTGAA GCGTAAGATC GACGCGGCGA AGATGCCGAA AGAGGCGAAA
      TCGGTAAGTC TTTCTCGAGC CGCTCTACCT GCTACGCGGC CTACTTTTGC TTCGCGACTT CGCATTCTAG CTGCGCCGCT TCTACGGCTT TCTCCGCTTT
-1
+1
1201  GAGAAAGCCG AAGCTGAGTT GCAGAAGCTG AAAATGATGT CTCCGATGTC GGCAGAGGCG GCAGAGGCGA CCGTCGTGCG CGGCTATATT GAGTGGATGG
      CTCTTTCGGC TTCGACTCAA CGTCTTCGAC TTTTACTACA GAGGCTACAG CCGTCTCCGC CGTCTCCGCT GGCAGCACGC GCCGATATAA CTCACCTACC
-1
+1
1301  TGCAGGTTCC GTGGAATGCC CGCAGCAAGG TCAAAAAAGA CCTGCGTCAG GCGCAGGAAA TCCTCGATAC CGACCACTAC GGCCTGGAAC GCGTGAAAGA
      ACGTCCAAGG CACCTTACGG GCGTCGTTCC AGTTTTTTCT GGACGCAGTC CGCGTCCTTT AGGAGCTATG GCTGGTGATG CCGGACCTTG CGCACTTTCT
-2
-1
+1
1401  CCGCATTCTT GAGTACCTCG CGGTACAGAG CCGTGTGAAC AAAATCAAAG GCCCAATCCT GTGCCTGGTA GGGCCTCCAG GCGTGGGTAA AACCTCTCTG
      GGCGTAAGAA CTCATGGAGC GCCATGTCTC GGCACACTTG TTTTAGTTTC CGGGTTAGGA CACGGACCAT CCCGGAGGTC CGCACCCATT TTGGAGAGAC
-2
-1
+1
1501  GGCCAGTCCA TCGCCAAAGC GACCGGACGT AAGTATATCC GTATGGCGCT GGGCGGCGTA CGCGATGAAG CGGAAATCCG CGGCCATCGT CGCACCTATA
      CCGGTCAGGT AGCGGTTTCG CTGGCCTGCA TTCATATAGG CATACCGCGA CCCGCCGCAT GCGCTACTTC GCCTTTAGGC GCCGGTAGCA GCGTGGATAT
-2
-1
```

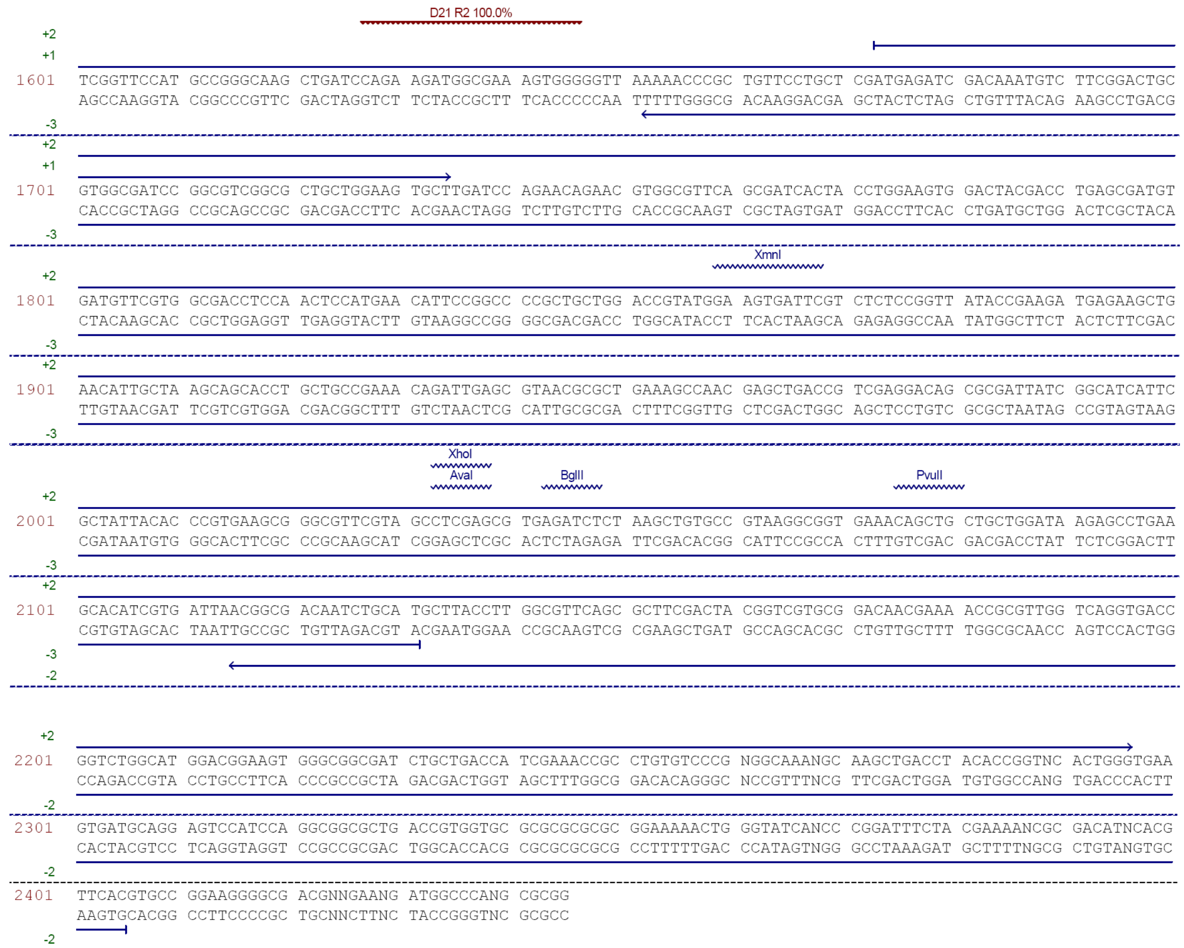


Figure 19. Nucleotide sequence of D21 F+R mutant

```
>ref|YP_001175641.1| G DNA-binding ATP-dependent protease La [Enterobacter sp. 638]
 gb|ABP59590.1| G ATP-dependent protease La [Enterobacter sp. 638]
Length=784

 GENE ID: 5111095 Ent638_0906 | DNA-binding ATP-dependent protease La
[Enterobacter sp. 638]

 Score =  369 bits (948),  Expect = 6e-101, Method: Compositional matrix adjust.
 Identities = 192/197 (97%), Positives = 194/197 (98%), Gaps = 3/197 (1%)

Query  1    MKAIQKELGEMDDAPDENEALKRKIDAAKMPKEAKEKAEAELQKLKMMSPMSAEAAEATV  60
            MKAIQKELGEMDDAPDENEALKR+IDAAKMPKEAKEKAEAELQKLKMMSPMSAEA   TV
Sbjct  234  MKAIQKELGEMDDAPDENEALKRRIDAAKMPKEAKEKAEAELQKLKMMSPMSAEA---TV  290

Query  61   VRGYIEWMVQVPWNARSKVKKDLRQAQEILDTDHYGLERVKDRILEYLAVQSRVNKIKGP  120
            VRGYIEWMVQVPWNARSKVKKDLRQAQEILDTDHYGLERVKDRILEYLAVQSRVNK+KGP
Sbjct  291  VRGYIEWMVQVPWNARSKVKKDLRQAQEILDTDHYGLERVKDRILEYLAVQSRVNKLKGP  350

Query  121  ILCLVGPPGVGKTSLGQSIAKATGRKYIRMALGGVRDEAEIRGHRRTYIGSMPGKLIQKM  180
            ILCLVGPPGVGKTSLGQSIAKATGRKYIRMALGGVRDEAEIRGHRRTYIGSMPGKLIQKM
Sbjct  351  ILCLVGPPGVGKTSLGQSIAKATGRKYIRMALGGVRDEAEIRGHRRTYIGSMPGKLIQKM  410

Query  181  AKVGVKNPLFLLDEIDK  197
            AKVGVKNPLFLLDEIDK
Sbjct  411  AKVGVKNPLFLLDEIDK  427
```

Figure 20. Blast result of D21 F+R mutant

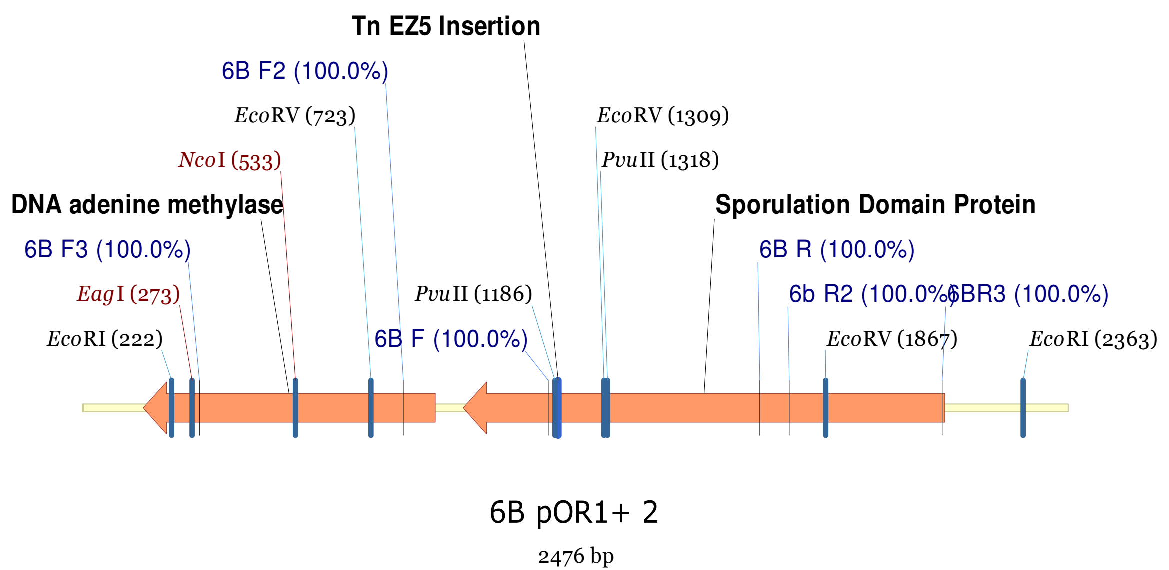


Figure 21. Feature map of 6B mutant

Yellow line – DNA sequence, Solid orange arrow – Interrupted gene, Blue lines - Primers

Figure 21 represents the feature map of the selenite sensitive 6B mutant showing the restriction sites, *Pvu*II, *Eag*I, *EcoR*V and *EcoR*I and open reading frames (solid orange arrows) which were translated into amino acid residues and analyzed by BLAST. BLAST analysis showed that the putative 408- amino acid residue polypeptide may be related to a sporulation domain protein of *Enterobacter* sp. 638 (Figure 23). In addition another ORF (small solid orange arrow) analyzed by BLAST was revealed to be DNA adenine methylase. The nucleotide sequence of 6B mutant is shown in Figure 22 along with the sequence of primers, 6B F, 6B F2 and 6B F3, used to obtain additional sequence of mutant DNA.

6B pOR1+ 2

```
   1  GACGNTCCAC AGGCNGNTCA TNCANTNACT NTTCNTCGNA ATCAANCTCC TCGTCNTCTC GCTCATTTCG CGCGGTACGG CTACGCGGAG GGCGACGATC
      CTGCNAGGTG TCCGNCNAGT ANGTNANTGA NAAGNAGCNT TAGTTNGAGG AGCAGNAGAG CGAGTAAAGC GCGCCATGCC GATGCGCCTC CCGCTGCTAG
 101  GTCTGCATCC AGTTCGACAT CGTCCAAAAT GATCTGCNGC TCGTTATCAC GTTCTGAAGA TTGACGAGAA CGACCAGTAC GACGATCGCT GGGATCGGGT
      CAGACGTAGG TCAAGCTGTA GCAGGTTTTA CTAGACGNCG AGCAATAGTG CAAGACTTCT AACTGCTCTT GCTGGTCATG CTGCTAGCGA CCCTAGCCCA
                                EcoRI                                                                   EagI
 201  TTCAGCTCGT CTTCTGGTTT GAATTCTNTT GCTGACCAGC TTTTCCGCCA TCTCAGCCAG ACGAGCCTGC TCGGCCGGGC TGAAGCTATT GGTGTGGTAG
      AAGTCGAGCA GAAGACCAAA CTTAAGANAA CGACTGGTCG AAAAGGCGGT AGAGTCGGTC TGCTCGGACG AGCCGGCCCG ACTTCGATAA CCACACCATC
                                                                                                         6B F3 100.0%
 301  GCGGTGAAAT TTGCCGTGGC AGACAGCGGC GCGTAAGGCG GGTCGCAGTA GACCACCGAG TTTACGCCCG CCAGATCCAT GCACTCTTCA TAGGAGAGAC
      CGCCACTTTA AACGGCACCG TCTGTCGCCG CGCATTCCGC CCAGCGTCAT CTGGTGGCTC AAATGCGGGC GGTCTAGGTA CGTGAGAAGT ATCCTCTCTG
        6B F3 100.0%
 401  AGTAGAACTC GGCATTCTGC GCTTTTTCAG CAAAGTGATA CAGTTCGTCC TGTGGGAAAT AGGGACGCTT ATAGCGGCCA AACGGCACGT TAAACTCGCC
      TCATCTTGAG CCGTAAGACG CGAAAAAGTC GTTTCACTAT GTCAAGCAGG ACACCCTTTA TCCCTGCGAA TATCGCCGGT TTGCCGTGCA ATTTGAGCGG
                                           NcoI
 501  ACGCAGATTA TACCGGCACA GACCGTTGTA GCCATGGCGG TTGAGATAAA GAAACAACAG CGCGCGACGG AACGGGTCCT GGCTTTGATT AAACTCAGCG
      TGCGTCTAAT ATGGCCGTGT CTGGCAACAT CGGTACCGCC AACTCTATTT CTTTGTTGTC GCGCGCTGCC TTGCCCAGGA CCGAAACTAA TTTGAGTCGC
 601  CGGAACTGAT AGTAGACGTC CGGGTTGTTG TTCTCTGGCG TAAACAGCTT CCGCGCCTCT TCCACATACT CATCGGTACG CAGTTTGACG ATGTTATAGA
      GCCTTGACTA TCATCTGCAG GCCCAACAAC AAGAGACCGC ATTTGTCGAA GGCGCGGAGA AGGTGTATGA GTAGCCATGC GTCAAACTGC TACAATATCT
                             EcoRV
 701  GGCTGATAAG GTCGCTGTTG ATATCCGCCA GGATATAACG AGAAAAATCG GTATTCAGGA ACACCGATCC AGCACCCACG AAGGGCTCGA TAAGACACTC
      CCGACTATTC CAGCGACAAC TATAGGCGGT CCTATATTGC TCTTTTTAGC CATAAGTCCT TGTGGCTAGG TCGTGGGTGC TTCCCGAGCT ATTCTGTGAG
 801  GCCTTTCGGC AGGTGCTTTT TAATATCGTC GAGCAGGGGG TATTTCCCCC CTGCCCATTT CAGAAAAGCG CGATTTTTTT TCATGCTGAC TAACTGATTA
      CGGAAAGCCG TCCACGAAAA ATTATAGCAG CTCGTCCCCC ATAAAGGGGG GACGGGTAAA GTCTTTTCGC GCTAAAAAAA AGTACGACTG ATTGACTAAT
                   6B F2 100.0%
 901  CACCTTCTCC GGCTGTGGAG AAAGCTCCGA CAGCATCCTG CGCTTTAAAC ATTACTTCAG ATCGGCCTGC ACCTGATGAA TCGGCTTCGC CCACGGGTTT
      GTGGAAGAGG CCGACACCTC TTTCGAGGCT GTCGTAGGAC GCGAAATTTG TAATGAAGTC TAGCCGGACG TGGACTACTT AGCCGAAGCG GGTGCCCAAA
1001  TTCGCCTGCA CGTCGGCTGG CAGCGTGGAA ACGGCACGTT TTGCTTCATC TTTGGAAGCA TAAATACCAC TCACCAGCAC ATACCACGGT TGACCGTTAC
      AAGCGGACGT GCAGCCGACC GTCGCACCTT TGCCGTGCAA AACGAAGTAG AAACCTTCGT ATTTATGGTG AGTGGTCGTG TATGGTGCCA ACTGGCAATG
                                                                                     6B F 100.0%
                                                                                              PvuII
1101  GGGTCGTCTG ATAAACCACG TAGTTTTTCA GATTCGATTT CTTCGCCCAG GCGTTGAGGT TGTCATAGTT AGACGAACTG CTCAGCTGAA GGGATAGTAA
      CCCAGCAGAC TATTTGGTGC ATCAAAAAGT CTAAGCTAAA GAAGCGGGTC CGCAACTCCA ACAGTATCAA TCTGCTTGAC GAGTCGACTT CCCTATCATT

1201  ACGTCCGGGG TTGTTGTTCT CTGGCGAAAA CAGCTNCCGC GCCTCTTCCA AATACTCATC GGTACGCAGT TTGACGATGT TATAGAGGCT GATAAGGTCG
      TGCAGGCCCC AACAACAAGA GACCGCTTTT GTCGANGGCG CGGAGAAGGT TTATGAGTAG CCATGCGTCA AACTGCTACA ATATCTCCGA CTATTCCAGC
            EcoRV      PvuII
1301  CTGTTGATAT CCGCCAGCTG AAGGGTGTAG TTGCTGGAAG AGGCCGACTG CAAAGAACCC ACGTTGCCCG TCGTTTTACC TGATGCAGCG CCCGTTTTGG
      GACAACTATA GGCGGTCGAC TTCCCACATC AACGACCTTC TCCGGCTGAC GTTTCTTGGG TGCAACGGGC AGCAAAATGG ACTACGTCGC GGGCAAAACC
1401  CTGCGGGAGC CGCAGGTGTT GTTGCAGAAG GTGTTGTTGC AGCAGGCGCG GTTGTGGTCG CCGTTGCTTT CGGTGCAGTG GTTGCCGCAG GCGTTGGCGC
      GACGCCCTCG GCGTCCACAA CAACGTCTTC CACAACAACG TCGTCCGCGC CAACACCAGC GGCAACGAAA GCCACGTCAC CAACGGCGTC CGCAACCGCG
1501  TTTAACCGGC TGAGTCGCCG CAGGTTTTGT CTGCGCTGGC GCTTTTGCCA CGGTTTGCGG TTTCGTCTCA CGCTTCGGCT CAATCACAGC CTGCTTACGC
      AAATTGGCCG ACTCAGCGGC GTCCAAAACA GACGCGACCG CGAAAACGGT GCCAAACGCC AAAGCAGAGT GCGAAGCCGA GTTAGTGTCG GACGAATGCG
1601  TCCTGGCGTG GCGCAGTCTG CGTCTGGCGC GGTTTGGTTT CCGTTGTCGC CGTTTGCTGC TGGGCATTGC TGCCATGAAC TGGCGCAACG GTGGCTGGTT
      AGGACCGCAC CGCGTCAGAC GCAGACCGCG CCAAACCAAA GGCAACAGCG GCAAACGACG ACCCGTAACG ACGGTACTTG ACCGCGTTGC CACCGACCAA
                                                                                              6b R2 100.0%
1701  CAATTGGCAG CGTGGAGTTC ACCACGGCAG CGTTAACCTG CTCCTGATTC TGCGGCTGCA TCAGCGCGTT ATTCAGATCG CCCTGAACTT CAACACGCTG
      GTTAACCGTC GCACCTCAAG TGGTGCCGTC GCAATTGGAC GAGGACTAAG ACGCCGACGT AGTCGCGCAA TAAGTCTAGC GGGACTTGAA GTTGTGCGAC
           6B R 100.0%
                                                                    EcoRV
1801  CTGGCCTTCA GGCGTCGCAG GCGCCTGGCC CTGAGTCGGG GTCGCAGAGA CAGGCGGCAG AGAGATATCC TGGCCAGCGC TGGTGTTACC TGCGGTCTGC
      GACCGGAAGT CCGCAGCGTC CGCGGACCGG GACTCAGCCC CAGCGTCTCT GTCCGCCGTC TCTCTATAGG ACCGGTCGCG ACCACAATGG ACGCCAGACG
1901  TCTGCGGAAG TGGTGCCCGG CGCAGGCTGA GCGCCATTCG CCTGGTTGCT CGCATCGTTA CCAGACAGGT CGATGCTCTT CTCGGCAGAC GCCGTCTGCT
      AGACGCCTTC ACCACGGGCC GCGTCCGACT CGCGGTAAGC GGACCAACGA GCGTAGCAAT GGTCTGTCCA GCTACGAGAA GAGCCGTCTG CGGCAGACGA
2001  CGGTAGAATT GGTAGAAGGT GCTTTTAGCG CGGACCCAAT GCCGACAATC AGCAAGACCA GCACCAGAAC GCCAAGGCCC ATCATGATGT ACTGACGGGA
      GCCATCTTAA CCATCTTCCA CGAAAATCGC GCCTGGGTTA CGGCTGTTAG TCGTTCTGGT CGTGGTCTTG CGGTTCCGGG TAGTACTACA TGACTGCCCT
                                                                      6BR3 100.0%
2101  NGCCNGTTTT GCCGCTACCG CTTTTTNACG TTTGCGCGGA CGACGTTCAA CAGGCTGTTC ATCCATTAAC TCTTCTTCGG ATTCATACTC CTCGTCTTCT
      NCGGNCAAAA CGGCGATGGC GAAAAANTGC AAACGCGCCT GCTGCAAGTT GTCCGACAAG TAGGTAATTG AGAAGAAGCC TAAGTATGAG GAGCAGAAGA
2201  CGCTCATTTC GCGCGTNACG GCTACGCGAA GGGCGACGAT CGTCTGCATC CAGTTCGACA TCGTCAAAAT NGATCTGCNG CTCGTNATCA CGTTCTGAAG
      GCGAGTAAAG CGCGCANTGC CGATGCGCTT CCCGCTGCTA GCAGACGTAG GTCAAGCTGT AGCAGTTTTA NCTAGACGNC GAGCANTAGT GCAAGACTTC
                                                                   EcoRI
2301  ATTGACGAGA NCGACCAGTA CGACGATCGC TGNGATCGGG TTTCAGCTCG TCTTCTGGTN TGAATTCTTT GGCTGACCAG CNTNNTCGCC ATCTGGGCCT
      TAACTGCTCT NGCTGGTCAT GCTGCTAGCG ACNCTAGCCC AAAGTCGAGC AGAAGACCAN ACTTAAGAAA CCGACTGGTC GNANNAGCGG TAGACCCGGA
2401  GACGAGCCTG CTCNGCCGGG CTGAAGCTAT GGGTGTGGTA GGCGGTGAAN TTTNNCGTGG CAGACAGCGG CGCGTA
      CTGCTCGGAC GAGNCGGCCC GACTTCGATA CCCACACCAT CCGCCACTTN AAANNGCACC GTCTGTCGCC GCGCAT
```

```
6B pOR1+ 2

   1  GACGNTCCAC AGGCNGNTCA TNCANTNACT NTTCNTCGNA ATCAANCTCC TCGTCNTCTC GCTCATTTCG CGCGGTACGG CTACGCGGAG GGCGACGATC
      CTGCNAGGTG TCCGNCNAGT ANGTNANTGA NAAGNAGCNT TAGTTNGAGG AGCAGNAGAG CGAGTAAAGC GCGCCATGCC GATGCGCCTC CCGCTGCTAG
 101  GTCTGCATCC AGTTCGACAT CGTCCAAAAT GATCTGCNGC TCGTTATCAC GTTCTGAAGA TTGACGAGAA CGACCAGTAC GACGATCGCT GGGATCGGGT
      CAGACGTAGG TCAAGCTGTA GCAGGTTTTA CTAGACGNCG AGCAATAGTG CAAGACTTCT AACTGCTCTT GCTGGTCATG CTGCTAGCGA CCCTAGCCCA
                             EcoRI                                                      EagI
 201  TTCAGCTCGT CTTCTGGTTT GAATTCTNTT GCTGACCAGC TTTTCCGCCA TCTCAGCCAG ACGAGCCTGC TCGGCCGGGC TGAAGCTATT GGTGTGGTAG
      AAGTCGAGCA GAAGACCAAA CTTAAGANAA CGACTGGTCG AAAAGGCGGT AGAGTCGGTC TGCTCGGACG AGCCGGCCCG ACTTCGATAA CCACACCATC
                                                                                                        6B F3 100.0%
 301  GCGGTGAAAT TTGCCGTGGC AGACAGCGGC GCGTAAGGCG GGTCGCAGTA GACCACCGAG TTTACGCCCG CCAGATCCAT GCACTCTTCA TAGGAGAGAC
      CGCCACTTTA AACGGCACCG TCTGTCGCCG CGCATTCCGC CCAGCGTCAT CTGGTGGCTC AAATGCGGGC GGTCTAGGTA CGTGAGAAGT ATCCTCTCTG
        6B F3 100.0%
 401  AGTAGAACTC GGCATTCTGC GCTTTTTCAG CAAAGTGATA CAGTTCGTCC TGTGGGAAAT AGGGACGCTT ATAGCGGCCA AACGGCACGT TAAACTCGCC
      TCATCTTGAG CCGTAAGACG CGAAAAAGTC GTTTCACTAT GTCAAGCAGG ACACCCTTTA TCCCTGCGAA TATCGCCGGT TTGCCGTGCA ATTTGAGCGG
                                        NcoI
 501  ACGCAGATTA TACCGGCACA GACCGTTGTA GCCATGGCGG TTGAGATAAA GAAACAACAG CGCGCGACGG AACGGGTCCT GGCTTTGATT AAACTCAGCG
      TGCGTCTAAT ATGGCCGTGT CTGGCAACAT CGGTACCGCC AACTCTATTT CTTTGTTGTC GCGCGCTGCC TTGCCCAGGA CCGAAACTAA TTTGAGTCGC
 601  CGGAACTGAT AGTAGACGTC CGGGTTGTTG TTCTCTGGCG TAAACAGCTT CCGCGCCTCT TCCACATACT CATCGGTACG CAGTTTGACG ATGTTATAGA
      GCCTTGACTA TCATCTGCAG GCCCAACAAC AAGAGACCGC ATTTGTCGAA GGCGCGGAGA AGGTGTATGA GTAGCCATGC GTCAAACTGC TACAATATCT
                            EcoRV
 701  GGCTGATAAG GTCGCTGTTG ATATCCGCCA GGATATAACG AGAAAAATCG GTATTCAGGA ACACCGATCC AGCACCCACG AAGGGCTCGA TAAGACACTC
      CCGACTATTC CAGCGACAAC TATAGGCGGT CCTATATTGC TCTTTTTAGC CATAAGTCCT TGTGGCTAGG TCGTGGGTGC TTCCCGAGCT ATTCTGTGAG
 801  GCCTTTCGGC AGGTGCTTTT TAATATCGTC GAGCAGGGGG TATTTCCCCC CTGCCCATTT CAGAAAAGCG CGATTTTTTT TCATGCTGAC TAACTGATTA
      CGGAAAGCCG TCCACGAAAA ATTATAGCAG CTCGTCCCCC ATAAAGGGGG GACGGGTAAA GTCTTTTCGC GCTAAAAAAA AGTACGACTG ATTGACTAAT
                 6B F2 100.0%
 901  CACCTTCTCC GGCTGTGGAG AAAGCTCCGA CAGCATCCTG CGCTTTAAAC ATTACTTCAG ATCGGCCTGC ACCTGATGAA TCGGCTTCGC CCACGGGTTT
      GTGGAAGAGG CCGACACCTC TTTCGAGGCT GTCGTAGGAC GCGAAATTTG TAATGAAGTC TAGCCGGACG TGGACTACTT AGCCGAAGCG GGTGCCCAAA
1001  TTCGCCTGCA CGTCGGCTGG CAGCGTGGAA ACGGCACGTT TTGCTTCATC TTTGGAAGCA TAAATACCAC TCACCAGCAC ATACCACGGT TGACCGTTAC
      AAGCGGACGT GCAGCCGACC GTCGCACCTT TGCCGTGCAA AACGAAGTAG AAACCTTCGT ATTTATGGTG AGTGGTCGTG TATGGTGCCA ACTGGCAATG
                                                                                     6B F 100.0%
                                                                                              PvuII
1101  GGGTCGTCTG ATAAACCACG TAGTTTTTCA GATTCGATTT CTTCGCCCAG GCGTTGAGGT TGTCATAGTT AGACGAACTG CTCAGCTGAA GGGATAGTAA
      CCCAGCAGAC TATTTGGTGC ATCAAAAAGT CTAAGCTAAA GAAGCGGGTC CGCAACTCCA ACAGTATCAA TCTGCTTGAC GAGTCGACTT CCCTATCATT

1201  ACGTCCGGGG TTGTTGTTCT CTGGCGAAAA CAGCTNCCGC GCCTCTTCCA AATACTCATC GGTACGCAGT TTGACGATGT TATAGAGGCT GATAAGGTCG
      TGCAGGCCCC AACAACAAGA GACCGCTTTT GTCGANGGCG CGGAGAAGGT TTATGAGTAG CCATGCGTCA AACTGCTACA ATATCTCCGA CTATTCCAGC
          EcoRV     PvuII
1301  CTGTTGATAT CCGCCAGCTG AAGGGTGTAG TTGCTGGAAG AGGCCGACTG CAAAGAACCC ACGTTGCCCG TCGTTTTACC TGATGCAGCG CCCGTTTTGG
      GACAACTATA GGCGGTCGAC TTCCCACATC AACGACCTTC TCCGGCTGAC GTTTCTTGGG TGCAACGGGC AGCAAAATGG ACTACGTCGC GGGCAAAACC
1401  CTGCGGGAGC CGCAGGTGTT GTTGCAGAAG GTGTTGTTGC AGCAGGCGCG GTTGTGGTCG CCGTTGCTTT CGGTGCAGTG GTTGCCGCAG GCGTTGGCGC
      GACGCCCTCG GCGTCCACAA CAACGTCTTC CACAACAACG TCGTCCGCGC CAACACCAGC GGCAACGAAA GCCACGTCAC CAACGGCGTC CGCAACCGCG
1501  TTTAACCGGC TGAGTCGCCG CAGGTTTTGT CTGCGCTGGC GCTTTTGCCA CGGTTTGCGG TTTCGTCTCA CGCTTCGGCT CAATCACAGC CTGCTTACGC
      AAATTGGCCG ACTCAGCGGC GTCCAAAACA GACGCGACCG CGAAAACGGT GCCAAACGCC AAAGCAGAGT GCGAAGCCGA GTTAGTGTCG GACGAATGCG
1601  TCCTGGCGTG GCGCAGTCTG CGTCTGGCGC GGTTTGGTTT CCGTTGTCGC CGTTTGCTGC TGGGCATTGC TGCCATGAAC TGGCGCAACG GTGGCTGGTT
      AGGACCGCAC CGCGTCAGAC GCAGACCGCG CCAAACCAAA GGCAACAGCG GCAAACGACG ACCCGTAACG ACGGTACTTG ACCGCGTTGC CACCGACCAA
                                                                                           6b R2 100.0%
1701  CAATTGGCAG CGTGGAGTTC ACCACGGCAG CGTTAACCTG CTCCTGATTC TGCGGCTGCA TCAGCGCGTT ATTCAGATCG CCCTGAACTT CAACACGCTG
      GTTAACCGTC GCACCTCAAG TGGTGCCGTC GCAATTGGAC GAGGACTAAG ACGCCGACGT AGTCGCGCAA TAAGTCTAGC GGGACTTGAA GTTGTGCGAC
           6B R 100.0%
                                                                     EcoRV
1801  CTGGCCTTCA GGCGTCGCAG GCGCCTGGCC CTGAGTCGGG GTCGCAGAGA CAGGCGGCAG AGAGATATCC TGGCCAGCGC TGGTGTTACC TGCGGTCTGC
      GACCGGAAGT CCGCAGCGTC CGCGGACCGG GACTCAGCCC CAGCGTCTCT GTCCGCCGTC TCTCTATAGG ACCGGTCGCG ACCACAATGG ACGCCAGACG
1901  TCTGCGGAAG TGGTGCCCGG CGCAGGCTGA GCGCCATTCG CCTGGTTGCT CGCATCGTTA CCAGACAGGT CGATGCTCTT CTCGGCAGAC GCCGTCTGCT
      AGACGCCTTC ACCACGGGCC GCGTCCGACT CGCGGTAAGC GGACCAACGA GCGTAGCAAT GGTCTGTCCA GCTACGAGAA GAGCCGTCTG CGGCAGACGA
2001  CGGTAGAATT GGTAGAAGGT GCTTTTAGCG CGGACCCAAT GCCGACAATC AGCAAGACCA GCACCAGAAC GCCAAGGCCC ATCATGATGT ACTGACGGGA
      GCCATCTTAA CCATCTTCCA CGAAAATCGC GCCTGGGTTA CGGCTGTTAG TCGTTCTGGT CGTGGTCTTG CGGTTCCGGG TAGTACTACA TGACTGCCCT
                                                                           6BR3 100.0%
2101  NGCCNGTTTT GCCGCTACCG CTTTTTNACG TTTGCGCGGA CGACGTTCAA CAGGCTGTTC ATCCATTAAC TCTTCTTCGG ATTCATACTC CTCGTCTTCT
      NCGGNCAAAA CGGCGATGGC GAAAAANTGC AAACGCGCCT GCTGCAAGTT GTCCGACAAG TAGGTAATTG AGAAGAAGCC TAAGTATGAG GAGCAGAAGA
2201  CGCTCATTTC GCGCGTNACG GCTACGCGAA GGGCGACGAT CGTCTGCATC CAGTTCGACA TCGTCAAAAT NGATCTGCNG CTCGTNATCA CGTTCTGAAG
      GCGAGTAAAG CGCGCANTGC CGATGCGCTT CCCGCTGCTA GCAGACGTAG GTCAAGCTGT AGCAGTTTTA NCTAGACGNC GAGCANTAGT GCAAGACTTC
                                                                    EcoRI
2301  ATTGACGAGA NCGACCAGTA CGACGATCGC TGNGATCGGG TTTCAGCTCG TCTTCTGGTN TGAATTCTTT GGCTGACCAG CNTNNTCGCC ATCTGGGCCT
      TAACTGCTCT NGCTGGTCAT GCTGCTAGCG ACNCTAGCCC AAAGTCGAGC AGAAGACCAN ACTTAAGAAA CCGACTGGTC GNANNAGCGG TAGACCCGGA
2401  GACGAGCCTG CTCNGCCGGG CTGAAGCTAT GGGTGTGGTA GGCGGTGAAN TTTNNCGTGG CAGACAGCGG CGCGTA
      CTGCTCGGAC GAGNCGGCCC GACTTCGATA CCCACACCAT CCGCCACTTN AAANNGCACC GTCTGTCGCC GCGCAT
```

Figure 22. Nucleotide sequence of 6B mutant

```
>ref|YP_001178507.1| G hypothetical protein Ent638_3801 [Enterobacter sp. 638]
 gb|ABP62456.1| G Sporulation domain protein [Enterobacter sp. 638]
Length=426

 GENE ID: 5110845 Ent638_3801 | hypothetical protein [Enterobacter sp. 638]

 Score =  446 bits (1148),  Expect = 1e-123, Method: Compositional matrix adjust.
 Identities = 276/408 (67%), Positives = 303/408 (74%), Gaps = 55/408 (13%)

Query  1    MDEQPVERRPRKRXKAVAAKXASRQYIMMGLGVLVLVLLIVGIGSALKAPSTNSTEQTAS  60
            +DE+P+ERRPRKR K VAAK ASRQY+MMGLGV VL+LLI+GIGSALKAPST S EQTAS
Sbjct  70   LDEEPLERRPRKRKK-VAAKPASRQYVMMGLGVFVLLLLIIGIGSALKAPSTPSNEQTAS  128

Query  61   AEKSIDLSGN--DASNQANGAQPAPGTTSAEQTAGN-TSAGQDISLPPVSATPTQGQAPA  117
             EKSI+LSGN  DA++QANGAQPAPGTTSAEQTAGN T+A QD+SLPPVSATPTQGQA A
Sbjct  129  TEKSINLSGNNNDAADQANGAQPAPGTTSAEQTAGNPTNAPQDVSLPPVSATPTQGQASA  188

Query  118  TPEGQQRVEVQGDLNNALMQPQNQEQVNAAVVNSTLPIEPATVAPVHGSNAQQQTATTET  177
             PEGQQRVEVQGDLNNAL Q  N +QVN  V NSTLP EPATVAP+ G NAQ QTA TET
Sbjct  189  APEGQQRVEVQGDLNNALTQ--NPDQVNNVVANSTLPTEPATVAPIRGGNAQPQTAATET  246

Query  178  KPRQTQTA-PRQERKQAVIEPKRETKPQTVAKAPAQTKPAATQPVKAPTPAATTAPKATA  236
            KPRQTQ A  R ER++AVIEPKRETKPQ VAKA  + KP A QP   PT  AT+AP  T
Sbjct  247  KPRQTQAATSRPERQKAVIEPKRETKPQAVAKA-TEAKPVA-QP--KPTETATSAPVKTP  302

Query  237  TTTAPAATTPSATTPAAPAAKTGAASGKTTGNVGSLQSASSSNYTLQLADINSDLISLYN  296
              TA    T +  TPAA A    G ASGK+ GNVG++++A S+NYTL
Sbjct  303  AATAAPKVTAATPTPAATATAPGTASGKSAGNVGAMKAAPSNNYTL--------------  348

Query  297  IVKLRTDEYLEEARXLFSPENNNPGRLLSLQLSSSSNYDNLNAWAKKSNLKNYVVYQTTR  356
                                          QLSSSSNYDNLN WAKKSNLKNYVVYQ+TR
Sbjct  349  ------------------------------QLSSSSNYDNLNNWAKKSNLKNYVVYQSTR  378

Query  357  NGQPWYVLVSGIYASKDEAKRAVSTLPADVQAKNPWAKPIHQVQADLK  404
            NGQPWY LVSG+YASKDEAKRAV+TLPADVQAKNPWAKPIHQVQADLK
Sbjct  379  NGQPWYTLVSGVYASKDEAKRAVATLPADVQAKNPWAKPIHQVQADLK  426
```

Figure 23. Blast result of 6B mutant

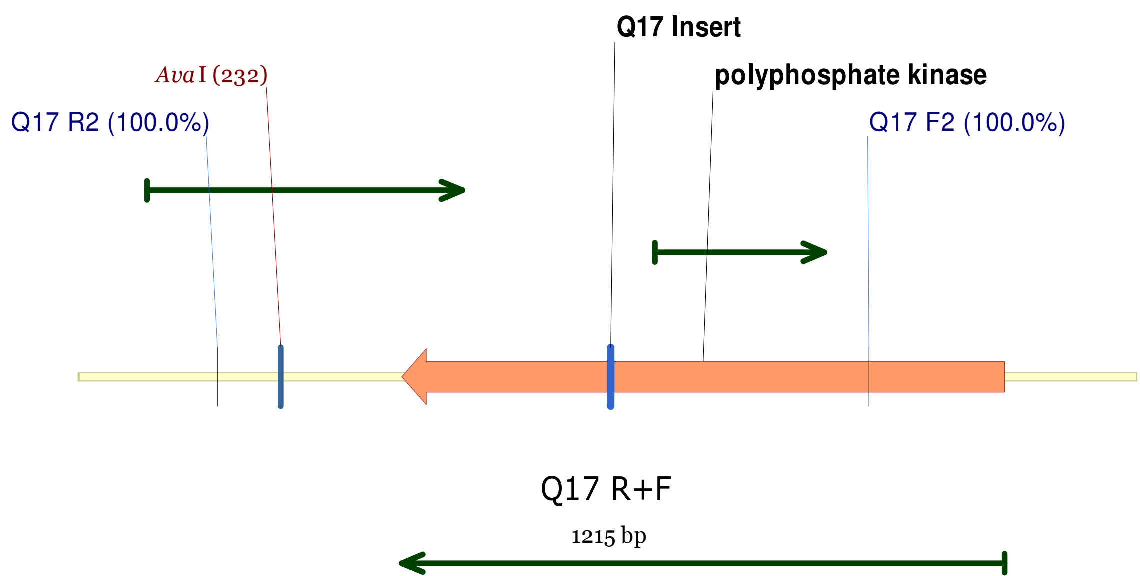


Figure 24. Feature map of Q17 mutant

Green arrows – Open Reading frames, Yellow line – DNA sequence,

Solid orange arrow – Interrupted gene, Blue lines - Primers

Figure 24 represents the feature map of selenite sensitive Q17 mutant showing the open reading frame (solid orange arrow) which was sequenced. The ORF was translated into a polypeptide and analyzed by BLAST. BLAST analysis revealed that the putative 212-amino acid residue polypeptide was nearly identical to polyphosphate kinase of *Enterobacter* sp. 638 (Figure 26). The nucleotide sequence of Q17 mutant is listed in Figure 25 along with the open reading frames denoted by blue arrow. The dark red line indicates the Q17 F2 primer that was used to obtain additional sequence of the mutated DNA.

Q17 R+F

```
+3
1     GTGTAGGGTA CGGGCGAGAC GGGAATACGG GGAATAGCAA TGCGNNAAGC ATTGCCGTCG GCATCGACCA TTTCATAATG GCCTTGCATG GTACCCAGCG
      CACATCCCAT GCCCGCTCTG CCCTTATGCC CCTTATCGTT ACGCNNTTCG TAACGGCAGC CGTAGCTGGT AAAGTATTAC CGGAACGTAC CATGGGTCGC
+3
101   GCGTTTCAAT GACCGCGCCG CTGGTGTACT GATACTCTTC GCCAGGGGCG ATATGGGGTT GTTCACCGAC CACGCCTTCG CCCTGGACTT CAATTTCACG
      CGCAAAGTTA CTGGCGCGGC GACCACATGA CTATGAGAAG CGGTCCCCGC TATACCCCAA CAAGTGGCTG GTGCGGAAGC GGGACCTGAA GTTAAAGTGC
                                                                  Q17 R2 100.0%
                                  AvaI
+3
201   GCCATTGCCG TTGGTGATAA GCCAGTAGCG CCCGAGCAGT TGCACAGGCA TCCGCCCCAG ATTGCGAATG GTTACGGTAT AAGCAAAGAC AAAACGTTCT
      CGGTAACGGC AACCACTATT CGGTCATCGC GGGCTCGTCA ACGTGTCCGT AGGCGGGGTC TAACGCTTAC CAATGCCATA TTCGTTTCTG TTTTGCAAGA
+3
301   TCATCCGGTG TGGATTGGGA CTCAATGTAG ACGCTTTGAA CCTGAACACA GACGCGGGGC GAATCAATCA TGGCTTAGCT CTCCTTCGGC GGTGCATTGT
      AGTAGGCCAC ACCTAACCCT GAGTTACATC TGCGAAACTT GGACTTGTGT CTGCGCCCCG CTTAGTTAGT ACCGAATCGA GAGGAAGCCG CCACGTAACA
-3
+3
401   CGCTGATGTA ATTAGCCAGC TTGCAGTATC GTCCAGGCAG TAGCGCAGGA TGTTATCCAG CAGGATCATC GGCTTACGCC GGCGCGGGGT TTCCGGCGGC
      GCGACTACAT TAATCGGTCG AACGTCATAG CAGGTCCGTC ATCGCGTCCT ACAATAGGTC GTCCTAGTAG CCGAATGCGG CCGCGCCCCA AAGGCCGCCG
-3
501   AGGTTTACAA AGCGTGGGAC TTTATCAGAC GGTATTTCCA GCAACGCGTA GTTAATGGTC TCACCACGAA TGATTTCTAC CGCCAGATAG GTGTAATCAT
      TCCAAATGTT TCGCACCCTG AAATAGTCTG CCATAAAGGT CGTTGCGCAT CAATTACCAG AGTGGTGCTT ACTAAAGATG GCGGTCTATC CACATTAGTA
-3
+2
601   CTTTCAGGAA CTTCAGGAAC TGCACCAGAT CGGTTTCACG GTTGATCAGA ATCGGGGTAA TGTGCTGGCG CAGATAATGT TTGAAATAGT GGCGCAGCCA
      GAAAGTCCTT GAAGTCCTTG ACGTGGTCTA GCCAAAGTGC CAACTAGTCT TAGCCCCATT ACACGACCGC GTCTATTACA AACTTTATCA CCGCGTCGGT
-3
+2
701   GTTTTGCTGA TTCGCGGAGA GCTGGCGTTC GTTGATCAGG AAGATTTGGT TGCGTGCCAT TTCCAGCAGC AGCTCGTTAT AGAGGCCATC AAATTCCTGA
      CAAAACGACT AAGCGCCTCT CGACCGCAAG CAACTAGTCC TTCTAAACCA ACGCACGGTA AAGGTCGTCG TCGAGCAATA TCTCCGGTAG TTTAAGGACT
-3
+2
801   TCGGCTTTCA TGACGCGAGA CTGGATTTTT CCAAGCAGAT GCCGCGAGTG GGAGTTTAAG CCCTGTTCTT CGCTGATGAT GATTCGGCGT TTCAGTTCGG
      AGCCGAAAGT ACTGCGCTCT GACCTAAAAA GGTTCGTCTA CGGCGCTCAC CCTCAAATTC GGGACAAGAA GCGACTACTA CTAAGCCGCA AAGTCAAGCC
-3

              Q17 F2 100.0%
901   CGAAGCGAAC CTTGTAGAAC TCATCCAGAT TGTTGGAATA GATGCCTAAA AAACGCATAC GTTCAATCAG CGGGTTGCTT TTATCGGCCG CTTCCTGAAG
      GCTTCGCTTG GAACATCTTG AGTAGGTCTA ACAACCTTAT CTACGGATTT TTTGCGTATG CAAGTTAGTC GCCCAACGAA AATAGCCGGC GAAGGACTTC
-3
1001  TACACGTTCG TTGAATGCTA ACCAGCTTAG CTCTTTCTCG ATGTATAACT TTTCCTGACC CATTACAGCT CACACTCCAG TTCAATCACA GGACGTGGTA
      ATGTGCAAGC AACTTACGAT TGGTCGAATC GAGAAAGAGC TACATATTGA AAAGGACTGG GTAATGTCGA GTGTGAGGTC AAGTTAGTGT CCTGCACCAT
-3
1101  AATCCGTCTC GTCCTTATTA TGGCGAGCAT TTCCACGATA TGTCCAACAG TGCCAGAAAA GTATGACAGT NATTTTTTTG TTGGGGAATN TAGAGGGTNG
      TTAGGCAGAG CAGGAATAAT ACCGCTCGTA AAGGTGCTAT ACAGGTTGTC ACGGTCTTTT CATACTGTCA NTAAAAAAAC AACCCCTTAN ATCTCCCANC
1201  GGGGAGGGGG GGTCG
      CCCCTCCCCC CCAGC
```

Figure 25. Nucleotide sequence of Q17 R + F mutant

```
>ref|YP_001177705.1|  polyphosphate kinase [Enterobacter sp. 638]
 gb|ABP61654.1|  Polyphosphate kinase [Enterobacter sp. 638]
Length=686

 GENE ID: 5112526 Ent638_2990 | polyphosphate kinase [Enterobacter sp. 638]

 Score =  416 bits (1069),  Expect = 7e-115, Method: Compositional matrix adjust.
 Identities = 204/212 (96%), Positives = 207/212 (97%), Gaps = 3/212 (1%)

Query  1    MGQEKLYIEKELSWLAFNERVLQEAADKSNPLIERMRFLGIYSNNLDEFYKVRFAELKRR  60
            MGQEKLYIEKELSWLAFNERVLQEAADKSNPLIERMRFLGIYSNNLDEFYKVRFAELKRR
Sbjct  1    MGQEKLYIEKELSWLAFNERVLQEAADKSNPLIERMRFLGIYSNNLDEFYKVRFAELKRR  60

Query  61   IIISEEQGLNSHSRHLLGKIQSRVMKADQEFDGLYNELLLEMARNQIFLINERQLSANQQ  120
            IIISEEQGLNSHSRHLLGKIQ+RV+KADQEFDGLYNELLLEMARNQIFLINERQLSANQQ
Sbjct  61   IIISEEQGLNSHSRHLLGKIQARVLKADQEFDGLYNELLLEMARNQIFLINERQLSANQQ  120

Query  121  NWLRHYFKHYLRQHITPILINRETDLVQFLKFLKDDYTYLAVEIIRGETINYALLEIPSD  180
             WLRHYFKHYLRQHITPILINRETDLVQ   FLKDDYTYLAVEIIRGE+I YALLEIPSD
Sbjct  121  TWLRHYFKHYLRQHITPILINRETDLVQ---FLKDDYTYLAVEIIRGESIRYALLEIPSD  177

Query  181  KVPRFVNLPPETPRRRKPMILLDNILRYCLDD  212
            KVPRFVNLPPETPRRRKPMILLDNILRYCLDD
Sbjct  178  KVPRFVNLPPETPRRRKPMILLDNILRYCLDD  209
```

Figure 26. Blast result of Q17 mutant

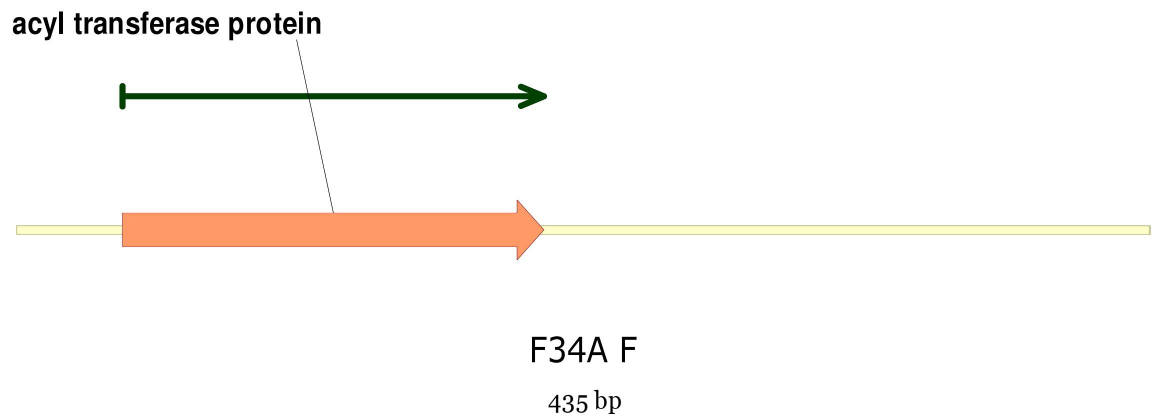


Figure 27. Feature map of F34 F

Green arrow – Open Reading frame, Yellow line – DNA sequence,

Solid orange arrow – Interrupted gene

Figure 27 represents the feature map of selenite sensitive F34 mutant showing an open reading frame (green arrow) translated into amino acid residue sequence (orange solid arrow). BLAST of this putative peptide revealed that it may be related to a segment of an E. *coli* acyl transferase (Figure 29). The nucleotide sequence of the F34 mutant is listed in Figure 28 along with the open reading frame denoted by a dark blue arrow.

```
                                   F34A F

  +2
   1  GGTCCCCACG ATCGTCGTAT NANCCAGACA GCTGGCCTGG GTGCCTTTCC ATTTGGACTG GCCTGCTGTG GCGCTGGATA TGCCGTTTAT GAAACGCTAT
      CCAGGGGTGC TAGCAGCATA NTNGGTCTGT CGACCGGACC CACGGAAAGG TAAACCTGAC CGGACGACAC CGCGACCTAT ACGGCAAATA CTTTGCGATA
  +2
 101  TCACGCAGCT ATTTAATTCG TCATCCGGAA CGCCGCGGTA AGGACGTGGA AACCACGCGC CGTTCTTGCG AGAAGTTTCG CGCGCATCCC NACCACCATT
      AGTGCGTCGA TAAATTAAGC AGTAGGCCTT GCGGCGCCAT TCCTGCACCT TTGGTGCGCG GCAAGAACGC TCTTCAAAGC GCGCGTAGGG NTGGTGGTAA
  +2
 201  GTTAACNTTC GGTCGAAGGA TCACGCTTTA CCGAAGAGAA GCGCCAGCCA GACTCGTTCT CCCTTATCAG AANCCTGTTG CCGCCCCAAA GGCTGCGGGC
      CAATTGNAAG CCAGCTTCCT AGTGCGAAAT GGCTTCTCTT CGCGGTCGGT CTGAGCAAGA GGGAATAGTC TTNGGACAAC GGCGGGGTTT CCGACGCCCG
 301  ATCGCAATGG CGCTCAANCG GTGCTGGGTG AGCAGTTNCG ATANAATTGT TGAACGTAAC GCTCTGCTAT CCGGAAANAA TGACAGGANC GCCCGGTTTC
      TAGCGTTACC GCGAGTTNGC CACGACCCAC TCGTCAANGC TATNTTAACA ACTTGCATTG CGAGACGATA GGCCTTTNTT ACTGTCCTNG CGGGCCAAAG
 401  TACGAATNAT GCCTCAGCCG GGCAANGCCT GAACC
      ATGCTTANTA CGGAGTCGGC CCGTTNCGGA CTTGG
```

Figure 28. Nucleotide sequence of F34 F mutant

```
>ref|YP_001746189.1| G acyltransferase domain protein [Escherichia coli SMS-3-5]
 gb|ACB15761.1| G acyltransferase domain protein [Escherichia coli SMS-3-5]
Length=318

 Score = 86.7 bits (213),  Expect = 5e-16, Method: Composition-based stats.
 Identities = 41/50 (82%), Positives = 43/50 (86%), Gaps = 1/50 (2%)

Query  1    VPFHLDWPAVALDMPFMKRYSRSYLIRHPERRGKDVETTRRSCEKFRAHP  50
            VPF L     ALDMPFMKRYSR+YL+RHPERRGKDVETTRRSCEKFR HP
Sbjct  142  VPF-LGLACWALDMPFMKRYSRAYLLRHPERRGKDVETTRRSCEKFRLHP  190
```

Figure 29. Blast result of F34 F mutant

In addition to using transposon mutagenesis to identify metal resistance genes, primers for known metal resistance genes, *pco* and *mer*[79] were used in PCR reactions to detect the presence of the mercury and copper resistance genes in *E. cloacae* UNK. A 1,500 bp fragment of *pcoA* and a 1,100 bp fragment of *mer* were cloned into the plasmid pSC-A and sequenced on each end. BLAST analysis of the short sequences showed that *E. cloacae* most likely contained these two genes. Figures 30 and 33 contain a map of the *pco* sequences on each end. As suggested by the BLAST analysis, the open reading frames (solid orange arrows) contained DNA sequences for the copper resistance protein of the CopA family (CP000946.1). The nucleotide sequences of each end of the cloned fragments are shown in Figures 31 and 34 and are referred to as *pcoF* and *pcoR* (forward and reverse primers).

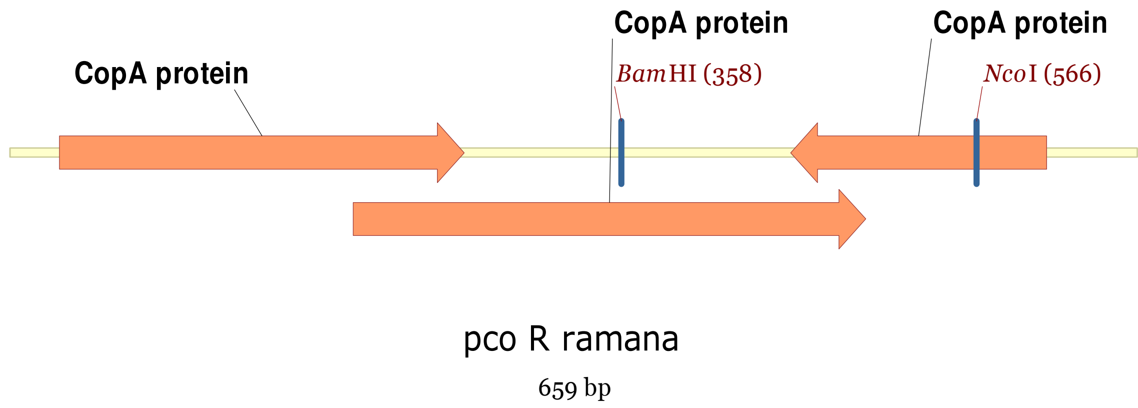


Figure 30. Feature map of *pcoR*

Yellow line – DNA sequence, Solid orange arrow – Protein

```
                                  pco R ramana

  1  GGACTTCACG AAACATTCCC ATTTCCATGT GATAGAGCAA ATGGCAGTGA TACGCCCAGC GGCCAAGCGC ATCTGCTGTC ACTCTGTAAC TGCGTTTTGT
     CCTGAAGTGC TTTGTAAGGG TAAAGGTACA CTATCTCGTT TACCGTCACT ATGCGGGTCG CCGGTTCGCG TAGACGACAG TGAGACATTG ACGCAAAACA
---------------------------------------------------------------------------------------------------------------
101  ACCAGGGGGA ACATCTATTG TGTGTTTACG AACCATGAAA TTACCGTTTT CATCTTCCAG ATCGCTCCAC ATACCATGCA GGTGAATGGG GTGAGTCATC
     TGGTCCCCCT TGTAGATAAC ACACAAATGC TTGGTACTTT AATGGCAAAA GTAGAAGGTC TAGCGAGGTG TATGGTACGT CCACTTACCC CACTCAGTAG
---------------------------------------------------------------------------------------------------------------
201  ATGGTATCGT TGATCAGCGT GATCCTGAGC CGCTCACCGT ATTTCAGCAG CACCGGTGCG GCATCTGAAA ACTTGATTCC GTTAAATGAC CAGGCAAACT
     TACCATAGCA ACTAGTCGCA CTAGGACTCG GCGAGTGGCA TAAAGTCGTC GTGGCCACGC CGTAGACTTT TGAACTAAGG CAATTTACTG GTCCGTTTGA
---------------------------------------------------------------------------------------------------------------
                                                                BamHI
301  TTTCCATGTG GCCGGTTAAA TGCAGTTCTA TGGTACGGCC AGGTTCACGT CCGTCAGGAT CCTCAAAGCG GCTTTTCAAA TCCGCGTACG TGAGAACCTT
     AAAGGTACAC CGGCCAATTT ACGTCAAGAT ACCATGCCGG TCCAAGTGCA GGCAGTCCTA GGAGTTTCGC CGAAAAGTTT AGGCGCATGC ACTCTTGGAA
---------------------------------------------------------------------------------------------------------------
401  TCTTCCGTTA TTTCGAAGAC CAATACCCGG ATCATTTAAT TTCGGAGAGA CGCTCATCGC CTGCATATCA ACCAGTGGGT TATCCGTTTC TGACGCAGGA
     AGAAGGCAAT AAAGCTTCTG GTTATGGGCC TAGTAAATTA AAGCCTCTCT GCGAGTAGCG GACGTATAGT TGGTCACCCA ATAGGCAAAG ACTGCGTCCT
---------------------------------------------------------------------------------------------------------------
                                                                             NcoI
501  TGACTTTGCA TACCCGGCAT TCCGGCCATC CGGGAATGAT CCATACCGGC CATGCTGCTG TGATCCATGG GCGCGGAGGA TGTCCCGCTA TCCGGANNGT
     ACTGAAACGT ATGGGCCGTA AGGCCGGTAG GCCCTTACTA GGTATGGCCG GTACGACGAC ACTAGGTACC CGCGCCTCCT ACAGGGCGAT AGGCCTNNCA
---------------------------------------------------------------------------------------------------------------
601  CAGCACCGTC CATAGACATC ATCTCTCCGC TGTTATCCAT GCCTCCCATC TGGCTGTGG
     GTCGTGGCAG GTATCTGTAG TAGAGAGGCG ACAATAGGTA CGGAGGGTAG ACCGACACC
---------------------------------------------------------------------------------------------------------------
```

Figure 31. Nucleotide sequence of *pcoR*

```
>gb|CP000783.1| D Enterobacter sakazakii ATCC BAA-894, complete genome
Length=4368373

 Features in this part of subject sequence:
   hypothetical protein

 Score =  555 bits (300),  Expect = 5e-155
 Identities = 300/300 (100%), Gaps = 0/300 (0%)
 Strand=Plus/Minus

Query  1        ATGGTATCGTTGATCAGCGTGATCCTGAGCCGCTCACCGTATTTCAGCAGCACCGGTGCG  60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4216578  ATGGTATCGTTGATCAGCGTGATCCTGAGCCGCTCACCGTATTTCAGCAGCACCGGTGCG  4216519

Query  61       GCATCTGAAAACTTGATTCCGTTAAATGACCAGGCAAACTTTTCCATGTGGCCGGTTAAA  120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4216518  GCATCTGAAAACTTGATTCCGTTAAATGACCAGGCAAACTTTTCCATGTGGCCGGTTAAA  4216459

Query  121      TGCAGTTCTATGGTACGGCCAGGTTCACGTCCGTCAGGATCCTCAAAGCGGCTTTTCAAA  180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4216458  TGCAGTTCTATGGTACGGCCAGGTTCACGTCCGTCAGGATCCTCAAAGCGGCTTTTCAAA  4216399

Query  181      TCCGCGTACGTGAGAACCTTTCTTCCGTTATTTCGAAGACCAATACCCGGATCATTTAAT  240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4216398  TCCGCGTACGTGAGAACCTTTCTTCCGTTATTTCGAAGACCAATACCCGGATCATTTAAT  4216339

Query  241      TTCGGAGAGACGCTCATCGCCTGCATATCAACCAGTGGGTTATCCGTTTCTGACGCAGGA  300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4216338  TTCGGAGAGACGCTCATCGCCTGCATATCAACCAGTGGGTTATCCGTTTCTGACGCAGGA  4216279
```

Figure 32. Blast result of *pcoR*

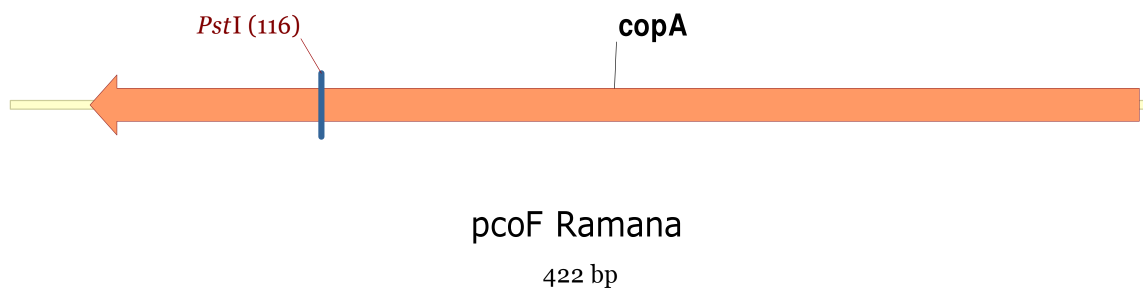


Figure 33. Feature map of *pcoF*

Yellow line – DNA sequence, Solid orange arrow – Protein

```
                                   pcoF Ramana

  1  CGTCTCGACG AACTTTCCTG AAGGGGTTAA CCCTCTCTGG CGTAGCCGGA AGTCTTGGCG TATGGAGTTT CAATGCGCGT TCCAGTCTGA GCCTGCCAGT
     GCAGAGCTGC TTGAAAGGAC TTCCCCAATT GGGAGAGACC GCATCGGCCT TCAGAACCGC ATACCTCAAA GTTACGCGCA AGGTCAGACT CGGACGGTCA
--------------------------------------------------------------------------------------------------------------------
                 PstI
101  TGCCGCATCC CTGCAGGGTA CTCAGTTTGA CCTGACCATT GGTGAAACGG CCGTCAATAT CACGGGCAGT GAGCGTCAGG CCAAAACAAT CAATGGAGGC
     ACGGCGTAGG GACGTCCCAT GAGTCAAACT GGACTGGTAA CCACTTTGCC GGCAGTTATA GTGCCCGTCA CTCGCAGTCC GGTTTTGTTA GTTACCTCCG
--------------------------------------------------------------------------------------------------------------------
201  CTGCCGGGGC CCGTTCTTCG CTGGAAAGAA GGTGACACCA TTACCCTGAA GGTCAAAAAC CGTCTTAATG AACAGACGTC CATTCACTGG CACGGCATTA
     GACGGCCCCG GGCAAGAAGC GACCTTTCTT CCACTGTGGT AATGGGACTT CCAGTTTTTG GCAGAATTAC TTGTCTGCAG GTAAGTGACC GTGCCGTAAT
--------------------------------------------------------------------------------------------------------------------
301  TTCTTCCGGC CAATATGGAT GGTGTTCCGG GGCTGAGTTT TATGGGCATA GAGCCTGATG ATACCTACGT TTACACCTTT AAGGTTAAGC AGAACGGGAC
     AAGAAGGCCG GTTATACCTA CCACAAGGCC CCGACTCAAA ATACCCGTAT CTCGGACTAC TATGGATGCA AATGTGGAAA TTCCAATTCG TCTTGCCCTG
--------------------------------------------------------------------------------------------------------------------
401  TTACTGGTAC CACAGCCATT CC
     AATGACCATG GTGTCGGTAA GG
--------------------------------------------------------------------------------------------------------------------
```

Figure 34. Nucleotide sequence of *pcoF*

```
>gb|CP000783.1| D Enterobacter sakazakii ATCC BAA-894, complete genome
Length=4368373

 Features in this part of subject sequence:
   hypothetical protein

 Score =  721 bits (390),  Expect = 0.0
 Identities = 390/390 (100%), Gaps = 0/390 (0%)
 Strand=Plus/Plus

Query  1        ACCCTCTCTGGCGTAGCCGGAAGTCTTGGCGTATGGAGTTTCAATGCGCGTTCCAGTCTG  60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4215017  ACCCTCTCTGGCGTAGCCGGAAGTCTTGGCGTATGGAGTTTCAATGCGCGTTCCAGTCTG  4215076

Query  61       AGCCTGCCAGTTGCCGCATCCCTGCAGGGTACTCAGTTTGACCTGACCATTGGTGAAACG  120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4215077  AGCCTGCCAGTTGCCGCATCCCTGCAGGGTACTCAGTTTGACCTGACCATTGGTGAAACG  4215136

Query  121      GCCGTCAATATCACGGGCAGTGAGCGTCAGGCCAAAACAATCAATGGAGGCCTGCCGGGG  180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4215137  GCCGTCAATATCACGGGCAGTGAGCGTCAGGCCAAAACAATCAATGGAGGCCTGCCGGGG  4215196

Query  181      CCCGTTCTTCGCTGGAAAGAAGGTGACACCATTACCCTGAAGGTCAAAAACCGTCTTAAT  240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4215197  CCCGTTCTTCGCTGGAAAGAAGGTGACACCATTACCCTGAAGGTCAAAAACCGTCTTAAT  4215256

Query  241      GAACAGACGTCCATTCACTGGCACGGCATTATTCTTCCGGCCAATATGGATGGTGTTCCG  300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4215257  GAACAGACGTCCATTCACTGGCACGGCATTATTCTTCCGGCCAATATGGATGGTGTTCCG  4215316

Query  301      GGGCTGAGTTTTATGGGCATAGAGCCTGATGATACCTACGTTTACACCTTTAAGGTTAAG  360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4215317  GGGCTGAGTTTTATGGGCATAGAGCCTGATGATACCTACGTTTACACCTTTAAGGTTAAG  4215376

Query  361      CAGAACGGGACTTACTGGTACCACAGCCAT  390
                ||||||||||||||||||||||||||||||
Sbjct  4215377  CAGAACGGGACTTACTGGTACCACAGCCAT  4215406
```

Figure 35. Blast result of *pcoF*

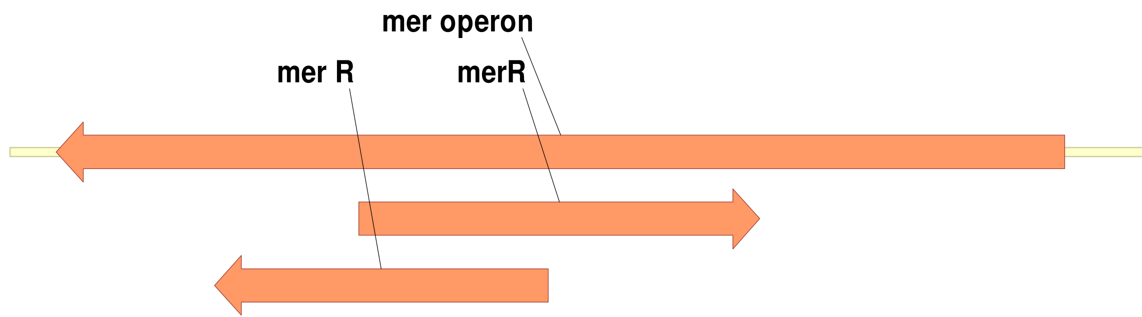


Figure 36. Feature map of *merR* operon

Yellow line – DNA sequence, Solid orange arrows - Genes

Figure 36 represents the feature map of the mercury resistance gene, *merR* showing the open reading frames. BLAST analysis of the short nucleotide sequence contained *mer* gene identical to mercury resistance transposable element from a strain of *Enterobacter cloacae* (Figure 38). The nucleotide sequence of *merR* is listed in Figure 37.

```
                                  mer R ramana

  1  GGGAGATCTA AAGCACGCTA AGGCGTAGTC ACCCCGTGAC TCCCCCGCGC CGATGCAGCG AGCTTCGTTC CGTCTTGCAG TGACGCAATC AGCGGGCAGG
     CCCTCTAGAT TTCGTGCGAT TCCGCATCAG TGGGGCACTG AGGGGGCGCG GCTACGTCGC TCGAAGCAAG GCAGAACGTC ACTGCGTTAG TCGCCCGTCC
-------------------------------------------------------------------------------------------------------------------
101  AAACGTTCCC TTTCCGCGCA TGGCAGGCGC ACACCAGTTC AGACAGCACG GCCTCCATGC GTGCCAAGTC GGCCATCTTC TCGCGCACAT CCTTGAGCTT
     TTTGCAAGGG AAAGGCGCGT ACCGTCCGCG TGTGGTCAAG TCTGTCGTGC CGGAGGTACG CACGGTTCAG CCGGTAGAAG AGCGCGTGTA GGAACTCGAA
-------------------------------------------------------------------------------------------------------------------
201  GTGCTCGGCC AGGCCGCTGG CTTCCTCGCA ATGGGTGCCA TCCTCCAGCC GCAGTAGCTC GGCGATTTCG TCCAGGCTAA AGCCCAGCCG CTGGGCCGAT
     CACGAGCCGG TCCGGCGACC GAAGGAGCGT TACCCACGGT AGGAGGTCGG CGTCATCGAG CCGCTAAAGC AGGTCCGATT TCGGGTCGGC GACCCGGCTA
-------------------------------------------------------------------------------------------------------------------
301  TTCACGAACC GCACTCGTGT TACATCCGCC TCGCCATAGC GGCGAATGCT GCCATAGGGC TTGTCTGGCT CCGGCAGCAG GCCCTTGCGC TGGTAGAACC
     AAGTGCTTGG CGTGAGCACA ATGTAGGCGG AGCGGTATCG CCGCTTACGA CGGTATCCCG AACAGACCGA GGCCGTCGTC CGGGAACGCG ACCATCTTGG
-------------------------------------------------------------------------------------------------------------------
401  GGATGGTCTC CACATTGACC CCGGCCGCCT TGGCAAAAAC GCCAATGGTC AGATTCTCAA AATTAATTTG CATATCGCTT GACTCCGTAC ATAACTACGG
     CCTACCAGAG GTGTAACTGG GGCCGGCGGA ACCGTTTTTG CGGTTACCAG TCTAAGAGTT TTAATTAAAC GTATAGCGAA CTGAGGCATG TATTGATGCC
-------------------------------------------------------------------------------------------------------------------
501  AAGTAAGCT
     TTCATTCGA
-------------------------------------------------------------------------------------------------------------------
```

Figure 37. Nucleotide sequence of *merR*

```
>emb|Y09025.1|ECMERTREL  Enterobacter cloacae DNA, mosaic mercury resistance transposable
element (mer-operon)
Length=8012

 Score =  837 bits (453),  Expect = 0.0
 Identities = 453/453 (100%), Gaps = 0/453 (0%)
 Strand=Plus/Plus

Query  1    AGGCGTAGTCACCCCGTGACTCCCCCGCGCCGATGCAGCGAGCTTCGTTCCGTCTTGCAG  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  37   AGGCGTAGTCACCCCGTGACTCCCCCGCGCCGATGCAGCGAGCTTCGTTCCGTCTTGCAG  96

Query  61   TGACGCAATCAGCGGGCAGGAAACGTTCCCTTTCCGCGCATGGCAGGCGCACACCAGTTC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  97   TGACGCAATCAGCGGGCAGGAAACGTTCCCTTTCCGCGCATGGCAGGCGCACACCAGTTC  156

Query  121  AGACAGCACGGCCTCCATGCGTGCCAAGTCGGCCATCTTCTCGCGCACATCCTTGAGCTT  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  157  AGACAGCACGGCCTCCATGCGTGCCAAGTCGGCCATCTTCTCGCGCACATCCTTGAGCTT  216

Query  181  GTGCTCGGCCAGGCCGCTGGCTTCCTCGCAATGGGTGCCATCCTCCAGCCGCAGTAGCTC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  217  GTGCTCGGCCAGGCCGCTGGCTTCCTCGCAATGGGTGCCATCCTCCAGCCGCAGTAGCTC  276

Query  241  GGCGATTTCGTCCAGGCTAAAGCCCAGCCGCTGGGCCGATTTCACGAACCGCACTCGTGT  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  277  GGCGATTTCGTCCAGGCTAAAGCCCAGCCGCTGGGCCGATTTCACGAACCGCACTCGTGT  336

Query  301  TACATCCGCCTCGCCATAGCGGCGAATGCTGCCATAGGGCTTGTCTGGCTCCGGCAGCAG  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  337  TACATCCGCCTCGCCATAGCGGCGAATGCTGCCATAGGGCTTGTCTGGCTCCGGCAGCAG  396

Query  361  GCCCTTGCGCTGGTAGAACCGGATGGTCTCCACATTGACCCCGGCCGCCTTGGCAAAAAC  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  397  GCCCTTGCGCTGGTAGAACCGGATGGTCTCCACATTGACCCCGGCCGCCTTGGCAAAAAC  456

Query  421  GCCAATGGTCAGATTCTCAAAATTAATTTGCAT  453
            |||||||||||||||||||||||||||||||||
```

Figure 38. Blast result of *merR*

CHAPTER VI: DISCUSSION

Transposon mutagenesis was performed on the multimetal-resistant strain, *Enterobacter cloacae* UNK, to study resistance mechanisms towards heavy metals such as mercury, cadmium, zinc, copper and selenite. Previous research on metal resistances suggested that *E. cloacae* UNK may express proteins involved in sequestration, transformation, efflux, and oxidative stress reduction in response to toxic concentrations of these metals. Some of identified proteins are discussed in the next few sections.

6.1. Lon protease (La protease)

It is one of the ATP dependent protease belonging to AAA^+ family of proteins.[80] Lon possesses two domains: an ATPase domain and proteolytic domain.[81] *E. coli* Lon protease is an oligomeric endoprotease with three functional domains: a variable N-terminal domain, an ATPase domain and a C-terminal proteolytic domain.[80]

The mutants, A3A, D21 and L30 of *E. cloacae* have interruptions in La protease conferring sensitivity to selenite. The *E. coli lon* gene encoding 159-amino acid protein may be involved in degrading short-lived regulatory proteins and thereby help maintain homeostasis during protein metabolism. They also participate in cell growth balance, in reducing external stress.[81] and in the maintenance of protein quality.[82] There are no references available for La protease responding to selenite. Our assumption is that when *E. coli* is exposed to toxic concentrations of selenite, La protease may be involved in degrading proteins that were damaged by the oxidative stress caused by selenite.

6.2. Sporulation Domain protein

The sporulation domain protein was one of the unexpected proteins (6B mutant) obtained by random mutagenesis of *Enterobacter cloacae* with the transposome because

this strain does not form endospores. It may have been involved in reducing oxidative stress when the bacterial strain, *E. cloacae* was exposed to the toxic concentrations of selenite. When microorganisms use oxygen during respiration, oxygen is reduced to form reactive oxygen species such as superoxide (O_2^-), hydrogen peroxide (H_2O_2) and hydroxide (OH^-). These reactive oxygen species cause oxidative stress which can be overcome by the formation of spores (Sporulation).[83] CueO encoding multi-copper oxidase confers resistance to copper by influencing the formation of spores.[63] The cellular structure of bacterial endospores are designed to protect the bacterial cell against the extreme conditions such as heat, radiation, UV light and oxidizing agents by destroying vegetative cells.[84]

6.3. Polyphosphate kinase

This protein is a 219-amino acid polypeptide having a similar sequence to that of polyphosphate kinase belonging to an *Enterobacter* species (accession number YP_001177705.1). The Q17 mutant appears to contain an interruption in a polyphosphate kinase gene that may confer resistance to selenite. The role of polyphosphate kinase varies in different organisms.[85] It is of vital importance in microorganisms in increased resistance to heavy metals.[86] Polyphosphate kinase has a major role in many living organisms, animals and plants by synthesizing inorganic polyphosphate (Poly P).[87]

Poly P is a multifunctional metabolite regulating the cell balance in bacteria. It is involved in transport, metabolism of orthophosphate (Pi) and feedback inhibition of poly P metabolism. Poly P also acts as a phosphate and energy reserve, participates in membrane channel formation and cell envelop development, controls gene expression,

and mediates stress response and cell survival during the stationary phase of bacterial growth. Bacterial poly P is located in the cytoplasm, cell surface, the periplasm and plasma membrane.[87] Microorganisms make use of the detoxification mechanism to sequester the heavy metals.[86] Bacterial cells respond to the heavy metals by stimulating the activity of exopolyphosphatase as soon as heavy metal cations enter the cell. Polyphosphate then sequesters these metals by the formation of phosphate-metal complex from Poly P which releases Pi. These metal complexes are then transported out of the cells.[87] From this, it may be inferred that *Enterobacter* species may also use this detoxification mechanism to reduce toxic selenite to non-toxic selenium.

The other mechanism by which detoxification of metals occurs is through hydrolysis of polyphosphate. Polyphosphate kinase catalyzes the formation of polyphosphate by transferring the terminal phosphate of ATP to a long chain polyphosphate (Poly P) inside the cell of microorganisms.[86]. In *Enterobacter* species, polyphosphate kinase may be involved in phosphorylating selenite before it enters and then may reduce it. Other levels of evidence also prove that heavy metals degrade intracellular polyphosphate during the growth of bacterial strains such as *Klebsiella aerogenes, Stichococcus bacillaris* and *Anacystis nidulans*.

6.4. P-type ATPases

The cadmium, zinc and selenite sensitive mutant, F24, appeared to contain an interrupted P-type ATPase, which may transport these metals out of the cell. (Accession Number: YP_001178579.1). The BLAST search resulted in 83% amino acid residue sequence similarity to a Cd^{2+}, Zn^{2+} and Hg^{2+} transporting P-type ATPase ZntA belonging

to *E. cloacae*. This protein is known to catalyze the efflux of Zn (II) or Cd (II) using ATP as an energy source.[88] It is interesting that this mutant is sensitive to zinc and cadmium but not to mercury. It is sensitive to selenite instead. Perhaps the sequence differences play a role in metal specificity.

6.5. Tyrosine recombinase

The cadmium, selenite and zinc sensitive mutant, F34, appeared to contain an insert in a subunit of the tyrosine recombinase gene, *xerC*. XerC and XerD are two members of the Xer site specific recombinases belonging to the integrase/tyrosine recombinase family. XerC and XerD encoding *xerCD* genes are found in both gram-positive bacteria such as *Bacillus subtilis*, *Lactobacillus leichmannii*, and *Staphylococcus aureus* and gram-negative bacteria, such as *E. coli*, *Enterobacteriaceae* species, *Pseudomonas aeruginosa*, *Haemophilus influenza* and *Vibrio cholera*.[89]

The site specific recombinases, XerC and XerD, act on specific sites such as *cer*, *ckr*, *nmr*, *parB* and *psi* found on plasmids or the *dif* site found on *E. coli* chromosome to convert multimeric replicons to monomeric state. *E. coli xerC* and *xerD* genes are functionally equivalent to *E. cloacae xerCD* genes.[90] The main functions of these site specific recombinases include cleavage of double stranded DNA and rejoining at *dif* site necessary for normal chromosome segregation during cell division. The recombination site, *dif* is bounded by *xerC* on the left half site and x*erD* on the right half site.[91] Both these recombinases are involved in performing intermolecular and intramolecular recombination.[91] As the main function of tyrosine recombinase is recombination, our assumption is that it may be involved in DNA repair due to the oxidative stress caused by

selenite. The nucleotide sequence of 8HB mutant is 85% homologous to *xerC* subunit belonging to tyrosine recombinase family of *Enterobacter* species.

6.6. Type II secretion protein

The L31 mutant appeared to contain an insert of an ATPase found in a Type II secretion system (Accession Number: ZP_01064893.1). The Type II secretion pathway is a unique transport system used by gram-negative bacteria to transport proteins across the periplasm or outer membrane into the extracellular environment[92, 93] It is the main terminal branch of general secretory pathway (GSP).[94]

Type II secretion occurs in two different steps: In the first phase, the unfolded proteins are transported across the cytoplasmic membrane into the periplasm by targeting them to either Sec or Tat machinery.[92] In the Sec machinery, the peptides are hydrolyzed by an ATP-hydrolyzing proteins, SecA and SecYEG translocon located in the Sec machinery. Tat components are also involved in folding the proteins in the cytoplasm. Tat system is used as an alternative for Sec- independent step for feeding secreton. Sec and Tat machinery routes converge at the translocation of protein across the outer membrane.[94] In the second phase, the unfolded proteins oligomerize, undergo post-translational modifications and convert to fully folded proteins in the periplasm. The components present in Type II Secretion pathway (T2S) are involved in translocation of the fully folded proteins across outer membrane.[92, 94] In the individual components, T2S proteins gather together to form cell envelope based upon protein-protein interactions between this components.[92]

Eps system so called as Type II secretion system is an assembly employed by *Vibrio cholerae* to fight against the diseases caused by pathogens, is involved in secreting cholera toxins from the periplasm into the lumen of gastro-intestinal tract of the host. It also plays a vital role in designing therapeutic agents for such diseases.[93] From the above functions performed by Type II secretion pathway, the cells may use this system to pump out selenite or a selenite-protein complex. Without the energy source provided by the ATPase, the cells may not be able to pump out the selenite or selenite protein complexes and are sensitive to it.

6.7. Acyltransferase

Acyltransferases participate mainly in lipid metabolism. The protein, RssC encoding acyltransferase family of *Serratia marcescens* is involved in the regulation of swarming behavior.[95] Homeserine transsuccinylase, HTS encoding *metA* gene and homoserine transacetylase, HTA encoding *metB* gene are responsible for the biosynthesis of methionine by making use of acyltransferase.[96] ADP1 of *A. calcoaceticus* is the first bacterial long chain acyltransferase involved in the catalysis of triacyl glycerols (TAG) and wax esters (WE) metabolism. It has a role in the acylation of diacylglycerides and fatty acids.[97] Ict1p encoding acyl-CoA lysophosphatidic acyltransferase contains a hydrolase/acyltranferase domain belonging to *Saccharomyces cerevisiae*. The major role of Ict1p in the biosynthesis of phosphatidic acid is useful in tolerating excess organic solvent stress by increasing the synthesis of phosphatidic acid.[98] Perhaps *Enterobacter cloacae* UNK, uses a similar mechanism.

6.8. Copper and Mercury resistance genes, *pcoA* and *merR*

Finally, DNA segments for *pcoA* and *merR*, copper and mercury resistance genes, respectively, were amplified by the polymerase chain reaction (PCR) and cloned into the pSCA plasmid to search for other metal resistance genes in *E. cloacae* UNK. The length of nucleotide sequence was 100% homologous to copper resistance gene belonging to the *copA* family from *E. coli*. The *merR* gene has 100% sequence similarity to mercury resistance *mer* operon of an *E. coli* strain may be encoded with *merR*, *merT* and partial *merP* genes. These genes were not detected by the transposon mutagenesis technique. Maybe the transposon just did not insert itself into any of these genes. On the other hand, if there are multiple mechanisms for resistances to these metals, interference of one gene would not result in metal-sensitivity. Likewise, if the genes for resistance are located on a multicopy plasmid, an interruption in one copy would not inactivate all the copies in the cell. Thus, this PCR approach is an important alternative to using transposon mutagenesis.

Conclusion: This work focused mainly on identifying the genes in the multi-resistant bacterium, *Enterobacter cloacae* and their functional roles such as efflux pumps, sequestration and metal transformations towards heavy metals such as mercury, cadmium, zinc, copper and selenite. The genes identified through this work are the Lon protease (also known as La protease), a sporulation domain protein which may be involved in response to oxidative stress, a P-type ATPase which may act as a mercury/cadmium/zinc transporter, an acyl transferase, a Type II Sec protein which may be involved in selenite efflux, a tyrosine recombinase, and a polyphosphate kinase which

may reduce selenite. Understanding resistance mechanisms of the heavy metals can be used to clean up some of the contaminated sites.

CHAPTER VII: REFERENCES

(1) Love, L. O. *Science* **1973**, *182*, 343-352.

(2) Widner, T. E.; Ripple, S. R.; Buddenbaum, J. E. *Health Phys.* **1996**, *71*, 457-469.

(3) Campbell, R. *Environmental Toxicology and Chemistry* **1998**, *17*, 1191.

(4) Nathaniel W. Revis, Suzanne B. Benson, Tanya R. Osborne 4826602, 1989.

(5) Yee, N.; Ma, J.; Dalia, A.; Boonfueng, T.; Kobayashi, D. Y. *Appl. Environ. Microbiol.* **2007**, *73*, 1914-1920.

(6) Watts, C. A.; Ridley, H.; Condie, K. L.; Leaver, J. T.; Richardson, D. J.; Butler, C. S. *FEMS Microbiol. Lett.* **2003**, *228*, 273-279.

(7) Dungan, R. S.; Frankenberger Jr, William T. *Soil Biol. Biochem.* **2000**, *32*, 1353-1358.

(8) Losi, M. E.; Frankenberger, W. T. *Appl. Environ. Microbiol.* **1997**, *63*, 3079-3084.

(9) Emsley, a. c. J. In *The elements;* Clarendon Press ; Oxford University Press: Oxford Oxfordshire New York, 1988; .

(10) Frankenberger, W. T.,Jr; Arshad, M. *Biofactors* **2001**, *14*, 241-254.

(11) Turner , Raymond J; Weiner , Joel H; Taylor , Diane E *Biometals* **1998**, *11*, 223-227.

(12) Stadtman, T. C. *Annu. Rev. Biochem.* **1996**, *65*, 83-100.

(13) Schwarz, K.; Foltz, C. M. *Nutrition* **1999**, *15*, 255.

(14) Alissa, E. M.; Bahijri, S. M.; Ferns, G. A. *Med. Sci. Monit.* **2003**, *9*, RA9-18.

(15) Yang, G. Q.; Chen, J. S.; Wen, Z. M.; Ge, K. Y.; Zhu, L. Z.; Chen, X. C.; Chen, X. S. *Adv. Nutr. Res.* **1984**, *6*, 203-231.

(16) Siddique, T.; Zhang, Y.; Okeke, B. C.; Frankenberger, W. T.,Jr *Bioresour. Technol.* **2006**, *97*, 1041-1049.

(17) Ranjard, L.; Nazaret, S.; Cournoyer, B. *Appl. Environ. Microbiol.* **2003**, *69*, 3784-3790.

(18) Dungan, R. S.; Frankenberger, J.,William T. *Biogeochemistry* **2001**, *55*, 73-86.

(19) Knotek-Smith, H. M.; Crawford, D. L.; Moller, G.; Henson, R. A. *J. Ind. Microbiol. Biotechnol.* **2006**, *33*, 897-913.

(20) Muller, S.; Heider, J.; Bock, A. *Arch. Microbiol.* **1997**, *168*, 421-427.

(21) Turner, R. J.; Weiner, J. H.; Taylor, D. E. *Biometals* **1998**, *11*, 223-227.

(22) Lindblow-Kull, C.; Kull, F. J.; Shrift, A. *J. Bacteriol.* **1985**, *163*, 1267-1269.

(23) Benitez, J. A.; Alonso, A.; Delgado, J.; Kotyk, A. *Folia Microbiol. (Praha)* **1983**, *28*, 6-11.

(24) Brown, T. A.; Shrift, A. *Can. J. Microbiol.* **1980**, *26*, 671-675.

(25) Brown, T. A.; Shrift, A. *Can. J. Microbiol.* **1982**, *28*, 307-310.

(26) McCready, R. G.; Din, G. A. *FEBS Lett.* **1974**, *38*, 361-363.

(27) Hudman, J. F.; Glenn, A. R. *Arch. Microbiol.* **1984**, *140*, 252-256.

(28) Guzzo, J.; Dubow, M. S. *Appl. Environ. Microbiol.* **2000**, *66*, 4972-4978.

(29) Sarret, G.; Avoscan, L.; Carriere, M.; Collins, R.; Geoffroy, N.; Carrot, F.; Coves, J.; Gouget, B. *Appl. Environ. Microbiol.* **2005**, *71*, 2331-2337.

(30) Bebien, M.; Chauvin, J. P.; Adriano, J. M.; Grosse, S.; Vermeglio, A. *Appl. Environ. Microbiol.* **2001**, *67*, 4440-4447.

(31) Leinfelder, W.; Forchhammer, K.; Veprek, B.; Zehelein, E.; Bock, A. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 543-547.

(32) Kessi, J.; Hanselmann, K. W. *J. Biol. Chem.* **2004**, *279*, 50662-50669.

(33) Virtanen, J. K.; Rissanen, T. H.; Voutilainen, S.; Tuomainen, T. *J. Nutr. Biochem.* **2007**, *18*, 75-85.

(34) Sastry *Environ. Prog.* **2002**, *21*, 231.

(35) Barkay, T.; Miller, S. M.; Summers, A. O. *FEMS Microbiol. Rev.* **2003**, *27*, 355-384.

(36) Silver, S.; Phung le, T. *J. Ind. Microbiol. Biotechnol.* **2005**, *32*, 587-605.

(37) Osborn, A. M.; Bruce, K. D.; Strike, P.; Ritchie, D. A. *FEMS Microbiol. Rev.* **1997**, *19*, 239-262.

(38) Moore, M. J.; Walsh, C. T. *Biochemistry (N. Y.)* **1989**, *28*, 1183-1194.

(39) Hamlett, N. V.; Landale, E. C.; Davis, B. H.; Summers, A. O. *J. Bacteriol.* **1992**, *174*, 6377-6385.

(40) Miller, S. M.; Moore, M. J.; Massey, V.; Williams, C. H.,Jr; Distefano, M. D.; Ballou, D. P.; Walsh, C. T. *Biochemistry* **1989**, *28*, 1194-1205.

(41) Wang, Y.; Mahler, I.; Levinson, H. S.; Halvorson, H. O. *J. Bacteriol.* **1987**, *169*, 4848-4851.

(42) Champier, L.; Duarte, V.; Michaud-Soret, I.; Coves, J. *Mol. Microbiol.* **2004**, *52*, 1475-1485.

(43) Summers, A. O. *J. Bacteriol.* **1992**, *174*, 3097-3101.

(44) Mukhopadhyay, D.; Yu, H. R.; Nucifora, G.; Misra, T. K. *J. Biol. Chem.* **1991**, *266*, 18538-18542.

(45) Smith, T.; Pitts, K.; McGarvey, J. A.; Summers, A. O. *Applied & Environmental Microbiology* **1998**, *64*, 1328.

(46) Marvin-DiPasquale, M. C.; Oreml, R. S. *Environ. Sci. Technol.* **1998**, *32*, 2556-2563.

(47) Blencowe, D. K.; Morby, A. P. *FEMS Microbiol. Rev.* **2003***, 27*, 291-311.

(48) Vallee, B. L.; Auld, D. S. *Biochemistry* **1990***, 29*, 5647-5659.

(49) Hantke, K. *Biometals* **2001***, 14*, 239-249.

(50) Chao *J. Biol. Chem.* **2004***, 279*, 17173.

(51) Nies, D. H. *FEMS Microbiol. Rev.* **2003***, 27*, 313-339.

(52) Anton, A.; Grosse, C.; Reissmann, J.; Pribyl, T.; Nies, D. H. *J. Bacteriol.* **1999***, 181*, 6876-6881.

(53) Lu, M.; Fu, D. *Science* **2007***, 317*, 1746-1748.

(54) Thelwell, C.; Robinson, N. J.; Turner-Cavet, J. S. *Proc. Natl. Acad. Sci. U. S. A.* **1998***, 95*, 10728-10733.

(55) VanZile, M. L.; Cosper, N. J.; Scott, R. A.; Giedroc, D. P. *Biochemistry* **2000***, 39*, 11818-11829.

(56) Eicken, C.; Pennella, M. A.; Chen, X.; Koshlap, K. M.; VanZile, M. L.; Sacchettini, J. C.; Giedroc, D. P. *J. Mol. Biol.* **2003***, 333*, 683-695.

(57) Grass, G.; Wong, M. D.; Rosen, B. P.; Smith, R. L.; Rensing, C. *J. Bacteriol.* **2002***, 184*, 864-866.

(58) Blindauer, C. A.; Harrison, M. D.; Robinson, A. K.; Parkinson, J. A.; Bowness, P. W.; Sadler, P. J.; Robinson, N. J. *Mol. Microbiol.* **2002***, 45*, 1421-1432.

(59) Wang, B.; Schneider, S. N.; Dragin, N.; Girijashanker, K.; Dalton, T. P.; He, L.; Miller, M. L.; Stringer, K. F.; Soleimani, M.; Richardson, D. D.; Nebert, D. W. *Am. J. Physiol. Cell. Physiol.* **2007***, 292*, C1523-35.

(60) Oger, C.; Mahillon, J.; Petit, F. *FEMS Microbiol. Ecol.* **2003***, 43*, 173-183.

(61) Nucifora, G.; Chu, L.; Misra, T. K.; Silver, S. *Proc. Natl. Acad. Sci. U. S. A.* **1989***, 86*, 3544-3548.

(62) Brocklehurst, K. R.; Megit, S. J.; Morby, A. P. *Biochem. Biophys. Res. Commun.* **2003***, 308*, 234-239.

(63) Rensing, C.; Grass, G. *FEMS Microbiol. Rev.* **2003***, 27*, 197-213.

(64) Mercury Treatability Study Final Report, Oak Ridge Y-12 Plant, Oak Ridge, Tennessee.

(65) Odermatt, A.; Suter, H.; Krapf, R.; Solioz, M. *J. Biol. Chem.* **1993***, 268*, 12775-12779.

(66) Magnani, D.; Solioz, M. *Biometals* **2005***, 18*, 407-412.

(67) Cooksey, D. A. *FEMS Microbiol. Rev.* **1994***, 14*, 381-386.

(68) Gupta, S. D.; Wu, H. C.; Rick, P. D. *J. Bacteriol.* **1997***, 179*, 4977-4984.

(69) Rapisarda, V. A.; Montelongo, L. R.; Farias, R. N.; Massa, E. M. *Arch. Biochem. Biophys.* **1999***, 370*, 143-150.

(70) Cavet, J. S.; Borrelly, G. P. M.; Robinson, N. J. *FEMS Microbiol. Rev.* **2003***, 27*, 165-181.

(71) Metcalf, W. W.; Jiang, W.; Wanner, B. L. *Gene* **1994***, 138*, 1-7.

(72) Ausubel, F. M., Ed.; In *Short protocols in molecular biology : a compendium of methods from Current protocols in molecular biology;* Wiley: New York, 1997; , pp 1 v. (various.

(73) Hanahan, D. *J. Mol. Biol.* **1983***, 166*, 557-580.

(74) Reasoner, D. J.; Geldreich, E. E. *Appl. Environ. Microbiol.* **1985***, 49*, 1-7.

(75) Goryshin, I. Y.; Jendrisak, J.; Hoffman, L. M.; Meis, R.; Reznikoff, W. S. *Nat. Biotechnol.* **2000***, 18*, 97-100.

(76) Miller, J. H. In *A short course in bacterial genetics : a laboratory manual and handbook for Escherichia coli and related bacteria;* Cold Spring Harbor Laboratory Press: Plainview, N.Y., 1992; .

(77) Altschul, S. F.; Gish, W.; Miller, W.; Meyers, E. W.; Lipman, D. J. *J. Mol. Biol.* **1990***, 215*, 403-410.

(78) Hobman, J. L. In *molecular techniques for the study of toxic metal resistance mechanisms in bacteria;* 2007; pp 1166.

(79) Hurst, C. J.; Crawford, R. L., Eds.; In *Manual of environmental microbiology;* ASM Press: Washington, D.C., 2007; , pp 1293.

(80) Botos, I.; Melnikov, E. E.; Cherry, S.; Tropea, J. E.; Khalatova, A. G.; Rasulova, F.; Dauter, Z.; Maurizi, M. R.; Rotanova, T. V.; Wlodawer, A.; Gustchina, A. *J. Biol. Chem.* **2004***, 279*, 8140-8148.

(81) Park, S. C.; Jia, B.; Yang, J. K.; Van, D. L.; Shao, Y. G.; Han, S. W.; Jeon, Y. J.; Chung, C. H.; Cheong, G. W. *Mol. Cells* **2006***, 21*, 129-134.

(82) Rotanova, T. V.; Botos, I.; Melnikov, E. E.; Rasulova, F.; Gustchina, A.; Maurizi, M. R.; Wlodawer, A. *Protein Sci.* **2006***, 15*, 1815-1828.

(83) Inaoka, T.; Matsumura, Y.; Tsuchido, T. *J. Bacteriol.* **1998***, 180*, 3697-3703.

(84) Martins, L. O.; Soares, C. M.; Pereira, M. M.; Teixeira, M.; Costa, T.; Jones, G. H.; Henriques, A. O. *J. Biol. Chem.* **2002***, 277*, 18849-18859.

(85) Geissdorfer, W.; Ratajczak, A.; Hillen, W. *Appl. Environ. Microbiol.* **1998***, 64*, 896-901.

(86) Keasling, J. D. *Ann. N. Y. Acad. Sci.* **1997***, 829*, 242-249.

(87) Kulaev, I. S.; Vagabov, V. M. *Adv. Microb. Physiol.* **1983**, *24*, 83-171.

(88) Gatti, D.; Mitra, B.; Rosen, B. P. *J. Biol. Chem.* **2000**, *275*, 34009-34012.

(89) Villion, M.; Szatmari, G. *FEMS Microbiol. Lett.* **2003**, *226*, 65-71.

(90) Sirois, S.; Szatmari, G. *J. Bacteriol.* **1995**, *177*, 4183-4186.

(91) Blakely, G. W.; Davidson, A. O.; Sherratt, D. J. *J. Mol. Biol.* **1997**, *265*, 30-39.

(92) Johnson, T. L.; Abendroth, J.; Hol, W. G.; Sandkvist, M. *FEMS Microbiol. Lett.* **2006**, *255*, 175-186.

(93) Abendroth, J.; Bagdasarian, M.; Sandkvist, M.; Hol, W. G. *J. Mol. Biol.* **2004**, *344*, 619-633.

(94) Voulhoux, R.; Ball, G.; Ize, B.; Vasil, M. L.; Lazdunski, A.; Wu, L. F.; Filloux, A. *EMBO J.* **2001**, *20*, 6735-6741.

(95) Soo, P. C.; Horng, Y. T.; Fu, Y. H.; Lu, C. C.; Lai, H. C. *Biochem. Biophys. Res. Commun.* **2008**, *371*, 462-467.

(96) Zubieta, C.; Arkus, K. A.; Cahoon, R. E.; Jez, J. M. *J. Biol. Chem.* **2008**, *283*, 7561-7567.

(97) Stoveken, T.; Steinbuchel, A. *Angew. Chem. Int. Ed Engl.* **2008**, *47*, 3688-3694.

(98) Ghosh, A. K.; Ramakrishnan, G.; Rajasekharan, R. *J. Biol. Chem.* **2008**, *283*, 9768-9775.

(99) www.securityglobal.org.