IDENTIFICATION OF A POSSIBLE SELENITE SENSOR PROTEIN FROM *Enterobacter* sp. YSU

by

Beatrice Rono

Submitted in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in the

Biological Sciences

Program

YOUNGSTOWN STATE UNIVERSITY

August 2013

IDENTIFICATION OF A POSSIBLE SELENITE SENSOR PROTEIN FROM *Enterobacter* sp. YSU

Beatrice Chepkemoi Rono

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:

Beatrice Chepkemoi Rono, Student Date

Approvals:

Dr. Jonathan Caguiat, Thesis Advisor Date

Dr. David Asch, Committee Member Date

Dr. Xiangjia Min, Committee Member Date

Dr. Sal Sanders, Associate Dean of Graduate Studies Date

ABSTRACT

The Y-12 plant (Oak Ridge, TN) contributed to national defense during World War II and the cold war. The plant processed uranium to make nuclear bombs and later switched to lithium processing to make hydrogen bombs. This process resulted in heavy metal waste deposits in East Fork Poplar Creek and the surrounding environment. *Enterobacter* sp. YSU, which was isolated from this creek, was found to be resistant to metal salts of zinc, cadmium, mercury, selenite, silver, copper and gold. Metal resistant bacteria encode proteins involved in pumping metals out of the cell, in converting them to less toxic forms or in sequestering them. Transposon mutagenesis was used to identify genes involved in selenite resistance. When introduced into the YSU strain, the EZ-Tn5™ <R6Kγ ori/KAN-2>Tnp Transposome™ inserted itself randomly into the genome of this bacterium. One of the transposome transformants, L31, was sensitive to selenite on agar plates. Minimal inhibitory concentration (MIC) experiments in liquid cultures showed that it was sensitive to 40 mM selenite in LB medium, but not in M-9 minimal medium. Gene rescue, DNA sequencing and Basic Local Alignment Search Tool (BLAST) analysis showed that the interrupted gene product is related to a histidine kinase sensor protein. It may regulate a surrounding gene that encodes a protein involved in efflux of toxic metals.

ACKNOWLEDGEMENTS

I am so thankful to God and everyone who contributed to my success at Youngstown State University.

I would like to express my deepest appreciation to my humble thesis advisor Dr. Jonathan Caguiat for giving me the opportunity to work in his lab and also for his help and guidance throughout my thesis research. Thank you for always encouraging me even when I had setbacks in my research. With your help, I have acquired knowledge and good research skills that will enable me advance in the research field. I would also like to thank my committee members Dr. Dave Asch and Dr. Xiangjia Min for taking time to read my thesis and for the good advice and ideas.

I would like to appreciate my parents, Mr. and Mrs. Rono, my sisters Evelyn, Fridah, Vivien, Lucy, Gladys and Phylis, brothers Victor and Vincent for their love and support throughout my studies.

I would also like to thank Prof. Z. Ngalo Otieno-Ayayo for always believing in me and encouraging me to always pursue higher learning. I am also grateful to my friends Mr. Clive Chibaya, Aimable Ngendahimana, David Kwitonda, Ruben Samson, Sheila Sang, Nelly Yegon and Aaron Korir for their encouragement throughout my studies. I am so thankful to my lab colleagues, Valentine Ngendahimana, Robert Giles, Evelyn Akpadock and Sumanun Noina for exchanging research ideas and for providing an excellent environment for research.

Lastly, I would like to thank the entire Proteomics/Genomic Research Group (P/GRG) and the Biology department for their support and for providing funds for my studies.

TABLE OF CONTENTS

COVER PAGE……………………………………………………………………………i

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

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

ACKNOWLEDGEMENTS…………………………………………………………….iv

TABLE OF CONTENTS………………………………………………………….v-vii

LIST OF FIGURES AND TABLES………………………………………………viii

LIST OF SYMBOLS AND ABBREVIATIONS…………………………………....ix

CHAPTER I: INTRODUCTION…………………………………………………1-9

1.1 Background (Y-12 Plant)……………………………………………………........1

1.2 *Stenotrophomonas maltophilia*…………………………………........................1-2

1.3 Microbes and Metal Interaction………………………………………………..2-3

1.4 Microbial Resistance Mechanisms……………………………………………..3-6

1.4.1 Intra-and extra Cellular Exclusion……………………………………………3

1.4.2 Intra-and extra Cellular Sequestration………………………………………..4

1.4.3 Detoxification………………………………………………………........4-5

1.4.4 Active Transport Efflux Pumps………………………………………...5-6

1.4.4.1 Cadmium……………………………………………………....5

1.4.4.2 Copper…………………………………………………….........6

Chapter II: Two-component Signal Transduction System…………………………7-9

2.1 Histidine Kinase Sensor……………………………………………………......7-8

2.2 Transposon Mutagenesis…..9

Chapter III: Hypothesis……..……………………….…………………………......10

Chapter IV: Methods……………………….………………………….……………...11-24

4.1 Growth Medium………………………………………………….....................11

4.2 Bacterial Strains……………………………………………………………...11-12

4.3 Minimal Inhibitory Concentrations (MICs) - Spotting Experiment……………12

4.4 Minimal Inhibitory Concentrations (MICs) - Growth Curves…………………..12

4.5 Genomic DNA Purification…………………………………………………12-13

4.6 Plasmid DNA Purification…………………………………………………...13-14

4.7 Agarose Gel Electrophoresis…………………………………………………....14-15

4.8 DNA Digestion…………………………………………………...................15-16

4.9 Self-Ligation Reactions………………………………………………….......16

4.10 Preparation of Electrocompetent Cells…………………………..……………17

4.11 Preparation of Competent Cells ($CaCl_2$ Method)……………….….……..17-18

4.12 Transformation by Electroporation……………………………………….....18

4.13 Transformation by Heat Shock………………………………………………..18

4.14 DNA Sequencing…………………………………………………………18-20

4.14.1 Sequencing Reaction Cleanup……..……..…………………………19-20

4.14.2 Sequence Analysis……………………………………….……………20

4.15 Polymerase Chain Reaction (PCR)……………………………………...20-21

4.16 Detection of Labeled Probe…………………………………………………21-22

4.17 Southern Blotting……..…………………………………………………..22-24

4.18 Hybridization and Detection……..…..………………………..…………24

Chapter V: Results..……………………..…………………….………………....28-57

5.1 Minimal Inhibitory Concentrations - Spotting Experiment………..…..…28-29

5.2 Minimal Inhibitory Concentrations - Growth Curves………………………..30-33

5.3 Genomic DNA Digestion……………………….……………………….………34-35

5.4 Gene Rescue……….…………….……………………….…………………...36-38

5.5 Sequence Analysis…...……………………………………………………......38-48

5.6 Southern Blot……………………………………………………………………..…48-50

5.7 Multiple Sequence Alignment……………………………………………………….51-55

5.8 Phylogenetic Analysis…………………………………………………………...56-57

Chapter VI: Discussion……………………………………………………………58-63

Chapter VII: References…………………………………………………………..64-72

LIST OF FIGURES AND TABLES

Figure. 1 – MIC Spotting Experiment – LB medium

Figure. 2 – MIC Spotting Experiment – M-9 medium

Figure. 3 – MIC Growth Curve – YSU

Figure. 4 – MIC Growth Curve – L31

Figure. 5 – Genomic DNA Digestions

Figure. 6 – Purified Plasmid DNA

Figure. 7 – BLAST Analysis – Histidine Kinase

Figure. 8 – BLAST Analysis – Transcriptional Regulator, CpxR

Figure. 9 – BLAST Analysis – Periplasmic Repressor, CpxP

Figure. 10 – L31 Feature Map

Figure. 11 – Assembled L31 Sequences

Figure. 12 – Southern Blot Gel

Figure. 13 – Southern Blot Detection

Figure. 14 – Multiple Sequence Alignment (MSA)

Figure. 15 – Phylogenetic Analysis

Table 1. List of Primers used in Sequencing

Table 2. Sizes of Linearized Plasmid DNA

LIST OF SYMBOLS AND ABBREVIATIONS

EFPC……………………..……………………………………………...…East Fork Poplar Creek

Na_2SeO_3……………………………………………………………...…….sodium selenite

NaCl………………………………………………………………..……...sodium chloride

NaOH………………………………………………………………...……sodium hydroxide

$CaCl_2$…………………………………………………………………….calcium chloride

HCl…………………………………………………………...………..hydrogen chloride

$MgCl_2$…………………………………………………………..….magnesium chloride

EDTA……………………………………..................…ethylenediaminetetraacetic acid

TBE………………………………………………….……………..…tris-borate-EDTA

BSA……………………………………………..………………bovine serum albumin

ATP…………………………………………..…………..…..adenosine triphosphate

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

dNTP..deoxyribonucleotide triphosphate

UV……………………………………………………………………..…ultraviolet

M……………………………………………………………………………..Molar

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

µM…………………………………………………………………………...micromolar

ml……………………………………………………………………….……..milliliters

mg…………………………………………………………………….....milligrams

μg……………………………………………………………………………micrograms

μl……………………………………………………………………..……..microliter

μF………………………………………………………………………….microfarad

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

γ……………………………………………………………………………..gamma

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

bp……………………………………………………………………………..basepair

RT-PCR…………………………..……reverse transcription polymerase chain reaction

SBP……………………..………………………………………..……..sulfur-binding protein

CDF…………………...……………………………..……..…cation diffusion facilitator

ICP-MS………………..…………………inductively coupled plasma mass spectrometry

CHAPTER I: INTRODUCTION

1.1 Background (Y-12 Plant)

Y-12 is the code name that was used during the World War II to refer to an electromagnetic isotope separating plant. The Y-12 plant at Oak Ridge, Tennessee played a significant role during World War II as it processed large amounts of uranium used to make nuclear weapons. Large amounts of toxic metals contaminated the surrounding environment due to this manufacturing process. During the Cold War, mercury was a major element in lithium separation to make hydrogen bombs (1). About 75-150 metric tons of mercury were released and most of it was contained in the sediments and floodplain soils of East Fork Poplar Creek, EFPC (8). To control mercury contamination, the US Department of Energy has set up different programs such as pollution control facilities, elimination of untreated discharges, mercury treatment systems and bank stabilization systems (19).

1.2 *Stenotrophomonas maltophilia*

Stenotrophomonas maltophilia Oak Ridge strain 02 (*S. maltophilia* 02) was isolated from East Fork Polar Creek (1). It is an aerobic, non-fermentative, Gram-negative bacterium (15). Some strains of this organism are human opportunistic pathogens that cause respiratory and urinary tract infection (6). Medical therapy is difficult because it is resistance to various antibiotics (15). When compared to the multi-metal resistant bacterium, *Enterobacter* sp. YSU, *S. maltophilia* 02 was more metal resistant and grew well in the presence of toxic levels of metal salts of mercury, gold, copper, selenite, cadmium, lead and chromium (1). Its ability to survive in the presence of

toxic metals is probably due to the gene expression of different resistance mechanisms involved in detoxification, efflux or sequestration.

1.3 Microbes and Metal interactions

Bacteria are the most numerous microorganisms in the natural environment and the functional roles and interactions of some of them are not yet known (7, 11). Their surface area to volume ratio provides them with a large contact area for interaction with its surrounding (11). Bacteria use different strategies to tolerate and survive different extreme environmental conditions such as radiation and toxic chemicals (13).

Bacteria require different elements to carry out their cellular functions. Essential non-metallic elements (H, C, O, P, N) carry the highest percentage in a bacterial cell, as they are the major components of proteins, lipids and nucleic acids. Elemental cations such as Na^{+}, Mg^{++}, K^{+}, Ca^{+} are important for cell function. Essential transitional elements (Cu, Zn, Fe, Co, Ni, Mn) are important for the structural integrity of proteins and nucleic acids and in biochemical pathways such as gene expression (4, 13). Since these different roles are exercised in different areas of the cell, it is necessary for the cell to regulate an appropriate level of these transitional elements. At high concentrations, it is cytotoxic to the cell. Polluted environments can have toxic metals such as Ag, Au, Cd, Cr and Sn. When bacteria encounter such toxic metals that have no beneficial role in its metabolism, it will have a toxic effect on their growth and survival. Therefore, bacteria have developed resistance mechanisms to overcome these challenges.

The different strategies employed by bacteria as a mechanism of survival are all of great interest. Understanding microbial resistance mechanism to heavy metals

deposited in the environment is of great importance for the potential use of microbes in bioremediation (7).

1.4 Microbial Resistance Mechanisms

Microbial metal resistance mechanisms can be summarized systematically in the following manner: (1) Intra- and extra cellular exclusion mechanisms; (2) intra- and extra cellular sequestration; (3) detoxification; (4) excretion via efflux transport (18).

1.4.1 Intra-and extra Cellular Exclusion

Metal exclusion mechanisms in Gram-negative bacteria are more complex than metal exclusion mechanisms in Gram-positive bacteria because Gram-negative bacteria have an outer and inner membrane whereas Gram-positive bacteria only have a single membrane (5). Gram-negative bacteria have outer membrane protein channels called porins that allow metal ions to diffuse through the outer membrane into the periplasm (16). Gram-negative bacteria exclude Cu (II) simply by repressing the expression of Cu (II) protein channels. The outer-membrane can also trap heavy metals by binding to them non-specifically. Bacterial strains such as *Klebsiella aerogenes* and *Pseudomonas putida* bind to metal extracellullary due to a protective coat of exopolysaccharide on the outer membrane (13, 28). The protective coat provides an attachment site for metal cations, hence preventing them from interacting with vital cellular components. In *Pseudomonas sp.* exclusion resistance mechanism is attributed by an operon of four genes *copA, copB, copC* and *copD* that are found in the outer and inner membrane (13).

1.4.2 Intra and –extra Cellular Sequestration

Intracellular sequestration leads to metal accumulation in the cytoplasm, preventing the metal from interacting with essential components of the cell (13). Some bacteria sequester metals such as cadmium, zinc, and copper. Intracellular sequestration of metals involves protein binding. Metallothionein production by *Synechococcus sp* is an example. This bacterial species has the gene, *smtA* that encodes a metallothionine. During high levels of cadmium, zinc and copper, the gene *smtA* is activated and binds to the metals. In the absence these metals, its expression is repressed by the protein, *SmtB. Pseudomonas putida* also displays intracellular sequestration of Cd (II) by producing low molecular weight cysteine, which is said to be related to the *Synechococcus sp* metallothionein (13, 33, 34).

According to Bruins *et al,* extracellular sequestration in bacteria has been hypothesized although a certain strain of *Klebsiella aerogenes* has demonstrated the ability to remove Cd (II) ions from the environment (13, 35).

1.4.3 Detoxification

Detoxification mechanism can be explained by microbial resistance to selenium. Selenium is an element of atomic number 34 and falls in group IV of the periodic table. It is an essential trace element for all living organisms and is therefore found in a number of food products but is considered to be toxic when taken in excess. In mammals, selenium is found in the form of selenocysteine and helps carry out peroxidase activity by the enzyme glutathione peroxidase. In bacteria, it is found as formate dehydrogenase, which helps bacteria carry out anaerobic metabolism (9, 10).

Selenium is transported into cells in the form of its oxyanions, selenite and selenate. The mechanisms for their transport into the cell is not well understood, but once they enter the cell, they are reduced to the less toxic form, elemental selenium, (9) or are incorporated to form the amino acids, selonocysteine and selenomethionine (12). These amino acids are similar to cysteine and methionine, except they contain selenium instead of sulfur.

1.4.4 Active Transport Efflux Pumps

There are different kinds of efflux proteins that pump out metals: (1) P-Type ATPases (CopA, ZntA and CadA), (2) A-Type ATPases (ArsB); (3) Cation diffusion facilitator family (CzcD) proteins and (3) RND (resistance, nodulation, division) efflux proteins. Toxic metals are removed from the bacterial cytoplasm via active transport. Cadmium and copper are examples of metals that are expelled using these mechanisms (3, 14, 17).

1.4.4.1 Cadmium

Cadmium occurs as a minor component in zinc ores and is therefore considered a by-product of zinc production (13, 29). The cell takes it up by divalent ion transport. Once inside the cell, it binds to sulfhydryl groups that are essential functional groups in proteins. In *E. coli*, it interferes with cellular functions and breaks single stranded DNA (13). Cadmium does not undergo enzymatic detoxification and as such it is characterized by the efflux resistance mechanism. A *Staphylococcus aureus* (*S. aureus*) plasmid has been found to contain two cadmium resistance genes (*cadA* and *cadC*). The protein product, CadA, shares amino acid similarity to the ion pump that removes cadmium from

the cell. The protein, CadC, has three metal binding sites and functions by presenting cations to the CadA protein (13).

The Czc system also functions to remove Cd, Zn and Co that enters the cell through a symport fashion. It has an operon that encodes four proteins (CzcA, CzcB, CzcC and CzcD), all which function to form an efflux pump. CzcA pumps heavy metals out of the cytoplasm, CzcB transfers the metal from the inner membrane to the outer membrane, CzcC exports the metals to the external environment and CzcD functions as a sensor protein involved in activating the efflux system (13,14, 35, 36).

1.4.4.2 Copper

Copper (Cu (II)) is an essential element to all living organisms. It is required in trace amounts for bacterial growth although it can be very toxic even at low concentrations (2, 17). The efflux resistance mechanism to copper is through the *cop* operon, which contains four genes (*copA, copB, copZ* and *copY*). The *copA* gene encodes a Cu (II) uptake ATPase, whereas the *copB* gene encodes a P-Type efflux ATPase. Deletion of the *copB* gene causes the microorganism to become sensitive to Cu (II) but deletion of the *copA* gene fails to make it sensitive. Therefore, CopA only plays a role in uptake. (13). The CopZ and CopY are repressor proteins that regulate the *cop* operon (13, 14, 33).

Chapter II: Two-component Signal Transduction System

2.1 Histidine Kinase Sensor Protein

There are different two-component pair systems used by bacteria in order to adapt to the changes in the physical and chemical environment (21, 22, 23). Two-component system functions to control expression of genes that encode proteins involved in pathogenesis and production of toxins (23). The first component is made up of a sensor histidine kinase that is located in the cytoplasmic membrane, and the second component is made up of a response regulator that is located in cytoplasm (21, 22, 23, 37). The sensor kinase functions as a membrane receptor, the receptor consists of a sensory domain that functions in sensing changes in the external environment. It then sends signal to the response regulator that mediates the proper cellular response (21, 23).

There are different kinds of sensor kinases but the best characterized is the *E. coli*, EnvZ protein, which functions in osmoregulation (21, 23). In terms of structure, a typical two-component phosphotransfer system consists of a dimeric transmembrane sensor histidine kinase and cytoplasmic response regulator. The histidine kinase domain has sequence motifs N, F, G1 and G2 boxes (61) that are conserved in members of the sensor kinase. These sequence motifs are located in the ATP-binding domain (21, 23). A histidine residue that becomes phosphorylated defines an H box located on the dimerization domain. The G boxes (G1 and G2) are glycine rich sites and are thought to be the nucleotide binding site with kinase and phosphatase activity (24). According to Alex *et al*, the role of F (phenylalanine) and N (asparagine) region are unknown but a UV-cross-linking experiment carried out by Weihong *et al* showed that mutation of the N-box in the EnvZ sensor kinase affects ATP binding due to misalignment of the γ-

phosphate of ATP with histidine-243 to be phosphorylated. In the same experiment, mutation of the F region also had an effect on ATP binding. Most kinases have all of the 5 conserved motifs but some like CheA of the chemotaxis system lack the H region (23). The EnvZ sensor kinase functions by sensing changes in osmolarity and sends a signal to the response regulator, OmpR which responds to the signal by regulating transcriptions of the porin genes, *ompF* and *ompC* (21).

The two-component system uses energy from ATP hydrolysis, and the flow of information involves protein phosphorylation and dephosphorylation. First, when the sensory domain of the sensor kinase senses external stimuli, it autophosphorylates a specific histidine residue (23, 24). Phosphorylation of the histidine residue is dependent on the γ-phosphoryl group on ATP (22). The phosphoryl group from the phosphorylated histidine (phosphohistidine) is then transferred to a conserved aspartate residue on the response regulator (22, 23, 24). This leads to the activation of an effector domain that triggers the appropriate cellular response (23, 24).

2.2 Transposon Mutagenesis

DNA transposition results in different mutations due to genome insertion and rearrangement (25, 26, 27). The transposition process first involves the binding of the transposase to a specific 19 bp end sequence on the transposon. Second, a synaptic complex is formed due to oligomerization that occurs on the transposable element's end. Third, the synaptic complex is cleaved at the blunt ends resulting to the release of the Transposome. Fourth, the transposon binds to the target DNA. The fifth step involves strand transfer where the 3' – OH ends of the transposon is transferred to the target 5' –

PO_4 groups. The sixth final step involves the release of the transposase from the transposition complex (8, 27).

In this research, the EZ-Tn5™ <R6Kγ *ori*/KAN-2>Tnp Transposome™ was introduced into the genome of *Enterobacter sp.* YSU by transposon mutagenesis. Its introduction into the genome of this bacterium resulted in the interruption of a gene that resulted in a selenite sensitive mutant. The primers KAN-2 FP-1 and R6KAN-2 RP-1 (Table. 1) were used to find the partial sequence of the gene that was interrupted by the transposon. The EZ-Tn5 transposon contains the R6Kγ*ori* origin of replication and a kanamycin resistance gene (KanR), both of which make the transposon useful during the rescue of the interrupted gene. It also contains a 19 bp mosaic end (ME) that is a transposase binding site.

CHAPTER III: HYPOTHESIS

Transposon mutagenesis was used to create the selenite sensitive mutant, L31. Basic Local Alignment Search Tool (BLAST) analysis showed that the mutant was sensitive to selenite due to the interruption of a putative signal transduction histidine kinase gene. I postulated that the sensor kinase may regulate genes that encode proteins involved the resistance mechanisms described above: (1) intra- and extra cellular exclusion mechanisms, (2) intra- and extra cellular sequestration proteins, (3) detoxification or (4) excretion of metals via efflux transport system

CHAPTER IV: METHODS

4.1 Growth Medium

Lennox LB medium was obtained from Fisher Scientific (Fair Lawn, NJ) and consisted of 10 g/l tryptone, 5 g/l yeast extract and 5 g/l sodium chloride. When required, LB medium was supplemented with 1.6% Agar (Amresco, Inc., Solon, OH) and with 50 μg/ml kanamycin (Amresco, Solon, OH).

M-9 salts were purchased from Becton Dickinson and Company (Sparks, MD). M-9 minimal medium contained 42 mM sodium hydrogen phosphate, 22 mM monopotassium phosphate, 18.7 mM ammonium chloride, 8.5 mM sodium chloride, 1 mM magnesium sulfate, 0.2 % glucose and water. When required, M-9 medium was supplemented with 1.6% Agar (Amresco, Inc., Solon, OH) and with 4 mg/ml cysteine hydrochloride.

SOC medium contained 0.5% (w/v) Yeast Extract, 2% (w/v) Tryptone, 10 mM sodium chloride, 2.5 mM potassium chloride, 10 mM magnesium chloride, 20 mM magnesium sulfate and 20 mM glucose

4.2 Bacterial Strains and Transposome

Enterobacter sp. YSU was isolated from Poplar Creek in Oak Ridge, TN (1). The selenite sensitive mutant, L31 was generated from *Enterobacter* sp. YSU through transposon mutagenesis.

Escherichia coli (*E. coli*) strain ECD100D pir (F– *mcrA* Δ (*mrr-hsd*RMS-*mcr*BC) f80d*lacZ*ΔM15 Δ*lac*X74 recA1 endA1 araD139 Δ (*ara, leu*) 7697 *gal*U *gal*K λ– *rps*L *nup*G pir+ (DHFR)) was purchased from Epicentre (Madison, WI).

4.3 Minimal Inhibitory Concentrations (MICs) - Spotting Experiment

Overnight cultures of L31 mutant and *Enterobacter* sp. YSU were started in M-9 medium and LB medium. Serial dilutions of overnight cultures were prepared in dilutions of 10^{-1} to 10^{-6}. 5 µl of the dilution were spotted on LB and M-9 medium without Na_2SeO_3 and on medium containing 40 mM Na_2SeO_3. The plates were incubated overnight at 30^0C.

4.4 Minimal Inhibitory Concentrations (MICs) - Growth Curves

30 ml of LB broth was added into two sterile 50 ml tubes and 0.6 ml of overnight cultures of *Enterobacter* sp. YSU and L31 mutant was added to the LB medium to make a dilution of 1:50. 5 ml of the diluted cells were added to 11 sterile tubes and different concentrations of selenite (0 mM to 100 mM) were added to the different tubes. The cultures tubes were incubated at 30^0C for 8 hours and turbidity was measured after every hour with a Klett colorimeter to monitor growth. Growth curves were plotted using Microsoft Excel in Klett units against time (minutes) to compare the growth of the wild type and the mutant at different concentrations of selenite.

4.5 Genomic DNA Purification

Genomic DNA was purified using the Wizard Genomic DNA purification kit from Promega (Madison, WI). The purification process was carried out by first centrifuging 1 ml of overnight culture at 15,000 x g for 2 minutes. The supernatant was

then poured off and the pelleted cells were gently resuspended in 600 µl of Nuclei Lysis Solution and incubated at 80^{0}C for 5 minutes in order to lyse the cells. It was cooled to room temperature, mixed with 3 µl of RNase solution (4 mg/ml), inverted 2-5 times and incubated at 37^{0}C for 15-60 minutes. After allowing the lysate to cool to room temperature, it was mixed with 200 µl of protein precipitation solution by vortexing for 20 seconds. It was incubated on ice for 5 minutes, followed by centrifugation for 3 minutes at 15,000 x g. The supernatant obtained after centrifugation was transferred to a sterile 1.5 ml tube containing 600 µl of isopropanol. The mixture was inverted several times until strands of DNA appeared. Centrifuging at 14,000 x g pelleted the DNA and the resulting supernatant was poured off. The DNA was washed by adding 300 µl of 70% ethanol, inverting the tube several times and centrifuging at 14,000 x g for 2 min. The supernatant was poured off and the pellet was air dried for 15 min. Lastly, the DNA was resuspended in 100 µl of DNA Rehydration solution and stored overnight at 4^{0}C.

4.6 Plasmid DNA Purification

Plasmid DNA purification was carried out using Promega's Wizard® Plus SV MiniPrep DNA purification kit. 10 ml of overnight culture was harvested by centrifuging at 8,000 x g for 5 min. The supernatant obtained was poured off and the excess media blotted out using a paper towel. 250 µl of cell resuspension solution containing 50 mM Tris-HCl (pH 7.5), 10 mM EDTA and 100 µg/ml RNase A was used to thoroughly resuspend the pellet by gently pipetting up and down. The resuspended cells were transferred to a sterile 1.5 ml microcentrifuge tube and 250 µl of cell lysis solution containing 0.2 M NaOH and 1% SDS was added to lyse the cells. The tube was inverted gently 4 times. Vortexing was avoided to prevent chromosomal DNA contamination. 10

µl of alkaline protease solution was added, and the tube was inverted gently 4 times, followed by a 5 minutes incubation step at room temperature. Next, 350 µl of Neutralizing solution containing 4.09 M guanidine hydrochloride, 0.759 M potassium acetate and 2.12 M glacial acetic acid was added, and the tube was inverted gently 4 times. This was followed by 10 minutes of centrifugation at 14,000 x g. The spin columns were inserted into 2 ml collection tubes, and the cleared lysate was poured into the spin column. The supernatant was centrifuged at maximum speed for 1 minute, and the flow through was discarded. 750 µl of Column Wash solution containing 162.8 mM potassium acetate, 22.6 mM Tris-HCl pH 7.5, 0.109 EDTA pH 8.0 and 58% ethanol was added to the spin column and centrifuged at 14,000 x g for 1 min. The flow through was discarded and a second wash was done using 250 µl of wash solution, followed by 2 minutes of centrifugation at 14,000 x g. The spin column was gently removed from the collection tube and transferred to a sterile 1.5 ml tube. Finally, 100 µl of nuclease free water was added to the spin column and plasmid DNA was eluted by centrifuging at 14,000 x g for 1 minute. The plasmid DNA was stored at -20 ^{0}C.

4.7 Agarose Gel Electrophoresis

Gel electrophoresis is used to separate proteins and DNA based on size (20). The percentage concentration of agarose to be used depends on the size of the DNA sample. 0.8% agarose has a low percentage of agarose, and hence is used to visualize larger fragment (5-10 kb) compared to 2% agarose that is used to visualize smaller fragments (0.2-1 kb) (20).

A 0.8% agarose gel was prepared by adding 1.04 g of agar (Fisher Scientific, Fair Lawn, NJ) to 130 ml of Tris Borate EDTA (TBE) buffer containing 0.089 Tris M, 0.089

M Borate and 0.002 M EDTA (Amresco, Solon, OH). It was heated in the microwave for 2 minutes with frequent swirling to ensure that the agarose dissolved completely. The mixture was then allowed to cool and dispensed carefully into a casting tray avoiding bubble formation. The combs were inserted and later removed when the gel had solidified hence creating a well.

To load DNA samples onto the gel, the gel was first placed on a gel box and submerged in 1X TBE buffer. DNA sample and EZ-vision dye (Amresco, Solon, OH) were mixed to create a homogenous solution and loaded into the wells. The gel box was then covered and an electric current of 100 V applied to run the gel. The gel was run for approximately 30 minutes and a picture of the gel was taken using an UltraCam Imaging Systems (Ultra-Lum, Inc. Claremont, CA).

4.8 DNA Digestion

Purified genomic DNA was partially digested with *Bfu*C I enzyme by making two mixtures in different tubes. The first mixture consisted of 1 µl of 10 X buffer 4 [50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate and 1 mM, dithiothreitol] (New England BioLabs Inc., MA), 1 µl of 10 X BSA (New England BioLabs Inc., MA), 7 µl of nuclease free water and 1 µl of *Bfu*C I enzyme (New England BioLabs Inc., MA). The second mixture consisted of 1.9 µl of 10 X buffer 4, 1.9 µl of BSA, 15.5 µl of purified genomic DNA and 1 µl of the first mixture hence making a total volume of 20 µl.

Genomic DNA was also digested using 1 µl of restriction enzyme *Eco*R I (New England Biolabs, Inc, MA), 4 µl of 10 X buffer 4, 5 µl of nuclease free water and 10 µl of

genomic DNA. *Bfu*C I digestion mixture was incubated at 37^0C for one hour and heat inactivated at 80^0C for 20 minutes. This was carried out using an Eppendorf Master Cycler. The *Eco*R I digestion mixture was incubated overnight at 37^0C and heat inactivated at 65^0C for 20 minutes.

4.9 Self-ligation Reactions

A 500 µl ligation reaction mixture was prepared by adding 50 µl of 10 X T4 DNA ligase buffer containing 10 mM ATP (New England BioLabs Inc. Beverly, MA), 2 µl of T4 DNA ligase (New England BioLabs Inc. Beverly, MA), 15 µl of digested genomic DNA and 433 µl of nuclease free water. The ligation mixture was incubated overnight at 4^0C.

4.9.1 DNA Ligation Precipitation

50 µl (a tenth of the total volume) of 3 M sodium acetate pH 5.2 and 1 ml 95% ethanol was added to the ligation reaction mixture and incubated at -20^0C for 10 minutes. The ligation mixture was then centrifuge at 14,000 x g for 10 minutes at room temperature. The supernatant was discarded and the pellet washed with 300 µl of 70% ethanol by gently inverting the tube to avoid resuspending the pellet. The tube was then centrifuged at maximum speed for 5 minutes, the supernatant was discarded and excess ethanol was drained by gently blotting on a paper towel. DNA pellet was dried for 10 minutes in a CentriVap (Labconco Corporation, Kansas City, MO). Finally, the pellet was resuspended in 10 µl of nuclease free water and the tube was tapped gently to cover the inside surface with water and centrifuged for 30 seconds to bring the contents to the bottom of the tube. 2 µl of DNA was then used for transformation by electroporation.

4.10 Preparation of Electrocompetent Cells

10 ml overnight culture of *E. coli* strain ECD100D *pir* was added to 160 ml of LB medium. The cells were grown at a 37 ^{0}C in a shaker until they reached an optical density (OD) of between 0.4-0.6 at a wavelength of 600 nm. The cells were then transferred to a sterile 80 ml centrifuge tubes that were pre-chilled on ice. The cells were centrifuged at 5000 x g for 15 minutes at 4^{0}C. The supernatant was poured off. The cells were resuspended in approximately 15 ml of chilled sterile water, mixed with more chilled sterile water up to 80 ml and centrifuged at 5000 x g at 4^{0}C for 15 minutes. This wash step was repeated twice and after the final wash, the cells were resuspended in 200 µl of chilled 10% glycerol and distributed into 1.5 ml microcentrifuge tubes. The cells were then centrifuged at 14,000 x g at a temperature of 4^{0}C for 10 minutes. The supernatant was discarded and the cells resuspended in 250 µl of pre-chilled 10% glycerol and stored at -80^{0}C.

4.11 Preparation of Competent cells ($CaCl_2$ method)

A 10 ml overnight culture of *E. coli* strain ECD100D *pir* was added to 100 ml of LB medium. The cells were grown at 37^{0}C in a shaker until they reached an optical density (OD) of 1.0 at 600 nm. The cells were then transferred to chilled 80 ml centrifuge tubes and cooled on ice for ~15 minutes. The cells were then harvested at 4^{0}C in a chilled centrifuge at 5000 x g for 5 minutes. The supernatant was poured off and the cells were resuspended in approximately 15 ml of sterile 0.15 M NaCl, followed by centrifugation at 5000 x g at 4^{0}C for 5 minutes. The supernatant was poured off and the cells were resuspended in 1 ml of ice-cold transformation buffer containing 15% glycerol, 0.1 M $CaCl_2$, 10 mM Tris-HCl, pH 8.0 and 10 mM $MgCl_2$. 400 µl of the resuspended cells were

distributed to pre-chilled sterile 1.5 ml tubes and incubated overnight on ice in the refrigerator. To make the cells competent, they were frozen at -80^{0}C.

4.12 Transformation by Electroporation

ECD100D *pir* electrocompetent *E. coli* cells were thawed on ice and 40 µl of the thawed cells were transferred to 1.5 ml microcentrifuge tubes. 2 µl of ligated DNA was added to the cells and transferred to an electroporation cuvette obtained from -20^{0}C freezer. The electroporation cuvette was tapped to ensure the cells were at the bottom of the cuvette. Cells were shocked at 25 µF, 200 ohms and 2.5 kV. Immediately, 960 µl of SOC media was added, and the cells were resuspended by pipetting up and down followed by incubation at 37^{0}C in a shaker for 45-60 minutes. The control tube was not shocked. 960 µl of SOC media was added to the control tube with 40 µl of cell and incubated at 37^{0}C. 100 µl of the cells were plated on an LB agar plate containing kanamycin and incubated overnight at 37^{0}C.

4.13 Transformation by Heat Shock

Competent *E. coli* cells were obtained from the -80^{0}C freezer and thawed out on ice. 100 µl of the cells was mixed in a 1.5 ml microcentrifuge tubes with 1 µl of plasmid DNA, followed by 30 minutes of incubation on ice. The microcentrifuge tubes with the cells and DNA were heat shocked in a water bath at 42 ^{0}C for a period of 50 seconds. The cells were put back on ice and 900 µl of LB medium was immediately added and the cells mixed by pipetting up and down. The cells were then incubated at 37^{0}C in a shaker for 45-60 minutes. 100 µl of the cells were plated on LB agar containing kanamycin and incubated overnight at 37^{0}C.

4.14 DNA Sequencing

Sequencing was carried out using the GenomeLab™ Dye Terminator Quick start Kit (Beckman Coulter, Inc. Fullerton, CA). Each reaction contained 50 fmol of DNA. The volume required for each reaction was determined by measuring the concentration of the DNA using a spectrophotometer and the size of the DNA using gel electrophoresis. The calculated volume of DNA was mixed with nuclease free water (NFW) to bring the volume to 10 µl. The DNA and NFW mixture were mixed in the 0.2 ml PCR tubes (BioExpress, Kaysville, UT) and heated at 96 ^{0}C for 1 minute, and then allowed to cool at room temperature. 2 µl of 1.6 µM primer (Table 1) and 8 µl of Dye Terminator Cycle Sequencing (DTCS) Quick Start Master Mix that contained DNA polymerase, pyrophosphate buffer, dNTPs and dye terminator ddNTPs were added to the DNA sample. The sequencing reaction mixture was then incubated in the Eppendorf Master Cycler according to the following program: 96 ^{0}C for 20 seconds, 50 ^{0}C for 20 seconds, 60 ^{0}C for 4 minutes for 30 cycles, and then holding the temperature at 4 ^{0}C.

4.14.1 Sequencing Reaction Cleanup (Ethanol Precipitation)

Sterile 0.6 ml microfuge tubes were prepared and fresh stop solution/glycogen mixture containing 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 µg/mL glycogen were added to the tubes. The sequencing reaction was then transferred to the tubes containing the stop solution/glycogen mixture and pipetted up and down to mix the samples. Ice cold 95% ethanol was added to the mixture and centrifuged at 14,000 x g for 1 minute at 4 ^{0}C. The supernatant was poured off and the pellet was washed twice with 200 µl of ice cold 70% ethanol. This was followed by centrifugation at 14,000 x g for 2 minutes at 4^{0}C after each wash. Excess ethanol was

pipetted off and the tubes were dried for 10 minutes using a CentriVap. The DNA was resuspended using 40 µl of sample loading solution and later analyzed using the Beckman Coulter CEQ 2000XL DNA analysis system (Fullerton, CA).

4.14.2 Sequence Analysis

Basic Local Alignment Search Tool (BLAST) is a program that searches for sequence similarity of a query sequences against a database that contains many other sequences. There are different kinds of BLAST analyses that search either a protein or a nucleotide database (31). To carry out a BLAST analysis, the query can be submitted in FASTA format, gi number or accession number. After submitting a query, a BLAST analysis will begin and matches to your query will be determined. VectorNTI was used to construct a map showing all the identified genes and also to assemble all the sequences by use of ContigExpress.

4.15 Polymerase Chain Reaction (PCR)

Polymerase chain reaction was used to amplify a region of the DNA that had the EZ-Tn5™ transposon insert. Therefore, primers specific for this region were designed and used for the PCR reaction. PCR reaction was used as a probe for Southern Blot but before it was used as a probe, it had to undergo biotinylation. Therefore, PCR reactions were set up and biotin-11-dUTP (Thermo Scientific, Rockford, IL) was added to the other components of the PCR reactions which included 4 µM of forward primer (Kan Probe F), 4 µM of reverse primer (Kan Probe R), 0.025 µg of 2x GoTaq DNA polymerase (Promega Corporation, Madison, WI), 1 µl of DNA and 11.5 µl of nuclease free water. The PCR reactions was carried out in the thermal cycler using the following

program: 95° C for 2 minutes, 35° cycles of 95° C for 1 minute (denatures DNA), 50° C for 1 minute (primer annealing) and 72° C for 1 minute (extension), followed by 72° C for 10 minutes and holding at 10° C.

The PCR reactions were purified as per the protocol outlined in the QIAquick PCR purification kit (Qiagen Sciences, MD). 5 volumes of buffer PB was added to 1 volume of the PCR reaction and mixed by pipetting up and down. The mixture was transferred into a MinElute column and centrifuged at 10, 000 x g for 1 minute to allow the DNA to bind to the membrane of the column. The flow through was discarded and the DNA was washed with 750 µl of PE buffer and centrifuged at 10,000 x g for 1 minute. The flow through was discarded and a second centrifuge was done for 1 minute to remove excess PE buffer. The column was transferred to a sterile 1.5 ml tube, and the DNA eluted by adding 35 µl of EB buffer (10 mM Tris-HCl, pH 8.5).

4.16 Detection of Labeled Probe

Chemiluminescent Nucleic Acid Detection Module (Thermo Scientific, Rockford, IL) was used to detect if the probe was labeled with biotin. It involved spotting 2 µl of biotin labeled ladder, 3 µl of labeled probe and 3 µl of unlabeled PCR product on a piece of Biodyne B Precut Nylon Membrane (0.4 µm, 8 × 12 cm). 16 ml of blocking buffer was added to the membrane and incubated at room temperature for 15 minutes with gentle shaking. The blocking buffer was decanted from the membrane and a conjugate/blocking buffer solution was prepared by adding 50 µl of Streptavidin-Horseradish Peroxidase conjugate to 15 ml of blocking buffer to make a 1:300 dilution. The conjugate/blocking buffer solution was added to the membrane and incubated at room temperature for 15 minutes with gently shaking. The membrane was transferred to a clean container rinsed

with 20 ml of 1 X wash buffer and washed 4 times in 1 X wash solution with gently shaking for 5 minutes. It was again transferred to another clean container and 30 ml of substrate equilibrium buffer was added followed by 5 minutes incubation with gentle shaking.

Chemiluminescent substrate working solution was prepared by mixing equal volumes of Luminol/Enhancer solution and stable peroxidase solution. The membrane was removed from the substrate equilibrium buffer and excess buffer was removed from the membrane by gently blotting the edge of the membrane on a piece of paper towel. The membrane was then transferred to a clean container and the prepared substrate working solution was poured on the membrane making sure it completely covered the entire surface. It was then incubated for 5 minutes without shaking. Lastly, the membrane was blotted on a paper towel, wrapped in a plastic wrap, and exposed to film or on a CCD camera.

4.16 Southern Blotting

Southern blotting is a technique used to detect a specific restriction DNA fragment. It involves the transfer of separated DNA fragments from an agarose gel to a nylon membrane. The DNA is then hybridized to a specific labeled DNA probe. In this research, Southern blotting was used to detect if there was more than one copy insertion of the EZ-Tn5™ Transposon in the L31 genome.

Southern blotting was carried using the TotalBlot™ Southern Kit from Amresco (Solon, OH). Before performing the blot, genomic DNA was extracted from the wild type and the L31 mutant, digested using *Eco*R I restriction endonuclease then resolved on

0.8% Agarose gel. A picture of the gel was taken using the UltraCam Imaging Systems and the section of the gel containing the DNA fragments was cut out and rinsed with distilled water.

The gel was transferred to a clean container and 2 gel volumes of depurination solution containing 0.25 M HCl was added to the gel and incubated at room temperature for 30 minutes with gentle shaking. The depurination solution functions to remove purines (adenine and guanine bases) and also breaks the DNA into small pieces, facilitating transfer during blotting. After depurination, the gel was rinsed with distilled water and the DNA was denatured by gently shaking the gel slab for 20 minutes in 2 gel volumes of denaturation solution containing 1.5 M NaCl and 0.5 M NaOH. Denaturation solution was poured off and the pH was lowered to below 9 by adding 2 gel volumes of neutralization solution containing 1.5 M NaCl and 1 M Tris, pH 7.0. After 20 minutes of incubation with gently shaking, the neutralization solution was poured off and fresh neutralization solution was added for another 20 minutes.

An upward capillary transfer method was used for the blot whereby a solid support was placed in a reservoir filled with 20X SSC (sodium saline citrate) buffer containing 3 M NaCl and 300 mM sodium citrate. Whatman 3MM blotting paper was saturated with the buffer and placed on the solid support with the ends submerged in the 20X SSC buffer. The gel was then placed gently on the filter paper making sure that no air bubbles formed. A piece of the Biodyne B Nylon Membrane was first cut almost the same size as the gel, wetted with distilled water and equilibrated in 20X SSC for 5 minutes then laid gently on top of the gel making sure no air bubbles formed. 3 sheets of Whatman 3MM blotting paper which were almost the same size as the membrane were

laid on top of the membrane followed by many layer of paper towel about 5 cm tall. Weight was added on top of the paper towels to ensure good contact throughout the stack. The stack was then left overnight.

The next day, the membrane was recovered from the stack and using a pen, the position of the wells was marked and a small notch was made on the left side of the membrane as a mark of orientation. The membrane was rinsed in 2X SSC buffer then placed of a sheet of Whatman 3MM blotting paper and allowed to dry by baking at 80 ^{0}C for 30 minutes.

4.18 Hybridization and Detection

The blot was transferred to a 50 ml tube and prehybridized using 0.1 ml per cm^2 of hybridization buffer at 55° C for 30 minutes with gentle rotation to ensure that the buffer completely covered the blot. The biotinylated probe was denatured by heating at 100° C for 10 minutes and placed on ice for 5 minutes. Approximately 30 ng of probe per milliliter of hybridization buffer was added to the prehybridized blot and incubated overnight at 55° C. The next day, the hybridization buffer was poured off and the blot was washed three times using 10 ml of 1X hybridization stringency wash buffer at 55° C for 15 minutes. The probe was then detected using the probe detection method described earlier.

Table 1. List of primers used in Sequencing

Primers	Nucleotide Sequence
Sen F2	5' – CTG TGG CTG CCG CTG TAT – 3'
Sen R	5' – ATA CAG CGG CAG CCA CAG C – 3'
Sen F3	5' – GGT AGA GAT GGT CGG TCC TT – 3'
Sen R3	5' – AGA AAG GAC CGA CCA TCT CTA – 3'
Sen F4	5' – GGT GCG GCT GAC TAT TTA C – 3'
Sen R4	5' – GTA AAT AGT CAT CCG CAC C – 3'
Sen F5	5' – GCC TTC AGT GCG TTT AGC CAG – 3'
Sen R5	5' – CTG GCT AAA CGC ACT GAA GGC – 3'
Sen F6	5' – CCT GTT TTC CTT GCC ATA GAC AC- 3'
Sen R6	5'–GTG TCT ATG GCA AGG AAA ACA GG-3'
Sen F7	5' – GAT TCA CTG GTG GAT ATT GC – 3'
SOD F1	5' – CTG AAC AGG CGG ACC TCT T – 3'
SOD R1	5' – AAG AGG TCC GCC TGT TCA – 3'
Fe Efflux F	5'– GTT GGC GTG GTC GTA AAC GGT AGT – 3'
Fe Efflux R	5' – CAA CCT GTC TGG TGG TGC GTC TAC T – 3'
S Binding F	5' – CAG GTA ACC GGC GAT AAC GTG G – 3'
S Binding R	5' – TTG CCA GAA CGC TGG TCG ATG – 3'
Sensor F1	5' – AAC GCT GTG GCT GCC GCT GTA – 3'
Sensor R1	5' – ATG ACG AGC AGT ATC ACC GCC – 3'
CDF F1	5' – CCT GGT ATG GCT GGC ATC GG – 3'
CDF F2	5' – TCG CTG AAC AGG TGG AGC AGG – 3'
PFK F1	5' – GTT GAC GG CTG GCG AAG TAC AT – 3'
PFK R1	5' – ATG TAC TTC GCC AGC TCG TCA AC – 3'
SBP F1	5' – GGTVTAG CTT TCT GGC ACG CTG – 3'
SBP R1	5' – CAG CGT GCC AGA AAG CTA ACC – 3'
SBP F2	5' – GTC TCA CGG TGG TTC TGG CAA – 3'

Primers	Nucleotide Sequence
SBP R2	5’ – TTG CCA GAA CCA CCG TGA GAC – 3’
SBP F3	5’ – GCA ACC AAC ACC TTC GTC GAA C – 3’
SBP R3	5’ – GTT CGA CGA AGG TGT TGG TTG C – 3’
KKR R	5’ – GCC AGA ACC ACC GTG AGA CT – 3’
KKR F	5’ – GCG TGG CCG TAT CGA CAA AA – 3’
Se Sen R	5’ – GGC TTT CAC CGT TTC GCT CAC A – 3’
Se Sen F	5’ – TGA AAG CCG ATG ACA GCC CG – 3’
ZZR F	5’ – CGC AAC CAA CAC CTT CGT CG – 3’
ZZR R	5’ – CGA CGA AGG TGT TGG TTG CG – 3’
JJR F	5’ – TAC CAC ATC AGA CCG CCG AG – 3’
JJR R	5’ – CGT AAC GAT GCC GCT TCA GG – 3’
Sensor F	5’ – CGG GCT GTC ATC GGC TTT CA – 3’
Sensor R	5’ – GAA GCG GAG CAG ATG GGC AA – 3’
Repressor F	5’ – AGC AGC AAG CGG TTT TGA ATA CCA – 3’
Repressor R	5’ – GGG CGT TAG CAG GTG GAA CAT CT – 3’
7BF R	5’ – GAA ACC AAA CGC GAA ACC CG – 3’
PFK F2	5’ – TTT GGT ATC TAT GAC GGT TAC CTG – 3’
PFK R2	5’ – GTG CAG AAC CGA GGA AAG TA – 3’
PFK F3	5’ – ACC GTT CTG GGT CAC ATT C – 3’
PFK R3	5’ – TTC GCC AGC TCG TCA AC – 3’
PFK R4	5’ – GGC GAA CCA CCA CGT TAT – 3’
SBP F4	5’ – ACC AAC GAG CTG GGT AAA – 3’
SBP R4	5’ – TTA GTC CTT GCT GAT TC – 3’
SBP R5	5’ – GGC GAA TAA ACG ATT CAF GA – 3’
CDF F3	5’ – CTT CTT GAC CGT GCG CTT C – 3’
CDF F4	5’ – GGT GGA GCA GGC GAT TT – 3’
R6KAN-2	5’ – CTA CCC TGT GGA ACA CCT ACA TCT – 3’
KAN – 2 F	5’ – ACC TAC AAC AAA GCT CTC ATC AAC C – 3’

Primers	Nucleotide Sequence
Kan Probe F	5' – GGT ATA AAT GGG CTC GCG ATA A – 3'
Kan Probe R	5' – CCG ACT CGT CCA ACA TCA ATA C – 3'

CHAPTER V: RESULTS

The L31 mutant used in this research was generated from a previous study through transposon mutagenesis (38). Minimal inhibitory concentration experiments previously done on this mutant in our lab showed that it was sensitive to 60 mM selenite but resistant to 240 µM cadmium, 750 µM zinc, 10 µM mercury and 3 mM copper (38). To confirm the mutant's phenotype in response to selenite, further MICs experiments were conducted on solid and liquid medium.

5.1 Minimal Inhibitory Concentration – Spotting Experiment

Minimal inhibitory concentration (MIC) experiments determine the lowest concentration that prevents bacterial growth (30). The MIC for the L31 mutant on solid medium was 40 mM. To confirm these results, additional spotting experiment were conducted at dilutions of 10^{-1} to 10^{-6}. Results obtained on LB and M-9 medium without selenite showed that the L31 mutant grew as well as the wild type (Fig. 1). When the medium was supplemented with 40 mM selenite, all dilutions (10^{-1} to 10^{-6}) of the L31 mutant failed to grow compared to all dilutions of the wild type which did grow (Fig. 1 top row).

The same results were observed in M-9 medium containing glucose, although there was minimal growth at dilutions of 10^{-1} and 10^{-2}, the mutant still showed sensitivity to selenite in all the other dilutions compared to the wild type which showed resistance in all dilutions (Fig. 2).

LB Medium

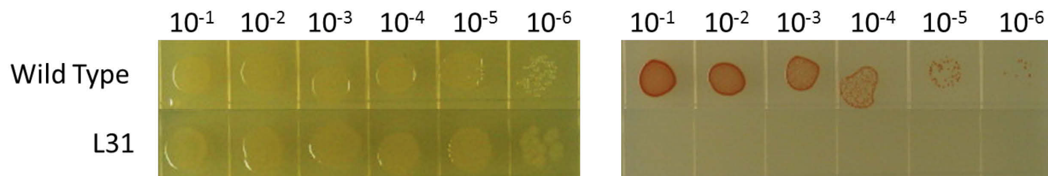

Figure. 1 Comparison of the wild type (top row) to the L31 mutant (bottom row) in LB medium without selenite (Left) and LB medium supplemented with 40 mM selenite (Right) at dilutions of 10^{-1} to 10^{-6}.

M-9 Medium

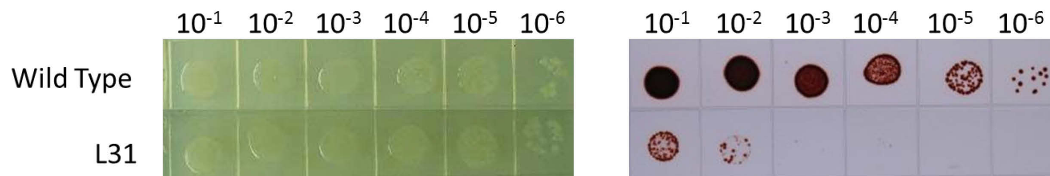

Figure. 2 Comparison of the wild type (top row) to the L31 mutant (bottom row) in M-9 medium without selenite (Left) and M-9 medium supplemented with selenite (Right) at dilutions of 10^{-1} to 10^{-6}.

5.2 Minimal Inhibitory Concentration – Growth Curves

To further study the phenotype of the mutant, MIC growth curve experiments were done in liquid medium containing concentrations of 0 mM to 100 mM selenite. The growth rate of the mutant was compared to the wild type by measuring turbidity every hour for eight hours using a Klett colorimeter. The experiment was repeated three times and the average turbidity values were plotted as shown in figures 3 and 4. Error was calculated using the student t test with a 95% confidence level.

The final average turbidities for the wild type (Fig. 3) grown in the presence of 0, 10, 20, 30, 40, 50, 60, 70, 80, 90 and 100 mM selenite were 197 ± 27, 272 ± 99, 245 ± 58, 229 ± 35, 213 ± 39, 193 ± 59, 180 ± 42, 174 ± 38, 167 ± 19, 163 ± 19 and 153 ± 25 Klett units, respectively, and the final average turbidities for the mutant (Fig. 4) were 174 ± 25, 231 ± 30, 191 ± 41, 143 ± 67, 115 ± 74, 89 ± 73, 65 ± 75, 60 ± 76, 53 ± 59, 47 ± 40 and 45 ± 37 Klett units, respectively. Although not statistically different, the final wild type turbidities of 272, 245, 229 and 213 Klett units for the 10, 20, 30 and 40 mM selenite cultures, respectively, were higher than the turbidity of 197 Klett units for the no selenite control. This was caused by the precipitation of red elemental selenium, rather than by an increase in cell number. For the L31 mutant, only the final turbidities of 231 and 191 Klett units for the 10 and 20 mM selenite cultures, respectively, were higher than the L31 turbidity of 174 Klett units for the no selenite control. Generally, all the turbidities for the L31 mutant at all selenite concentrations were lower than the corresponding turbidities of the wild type. At 70 mM selenite, the turbidity of 60 ± 76 Klett units for the L31 mutant is significantly lower than the turbidity of 167 ± 19 for the

wild type strain. Thus, the L31 mutant has a lower tolerance for 70 mM than the wild type strain.

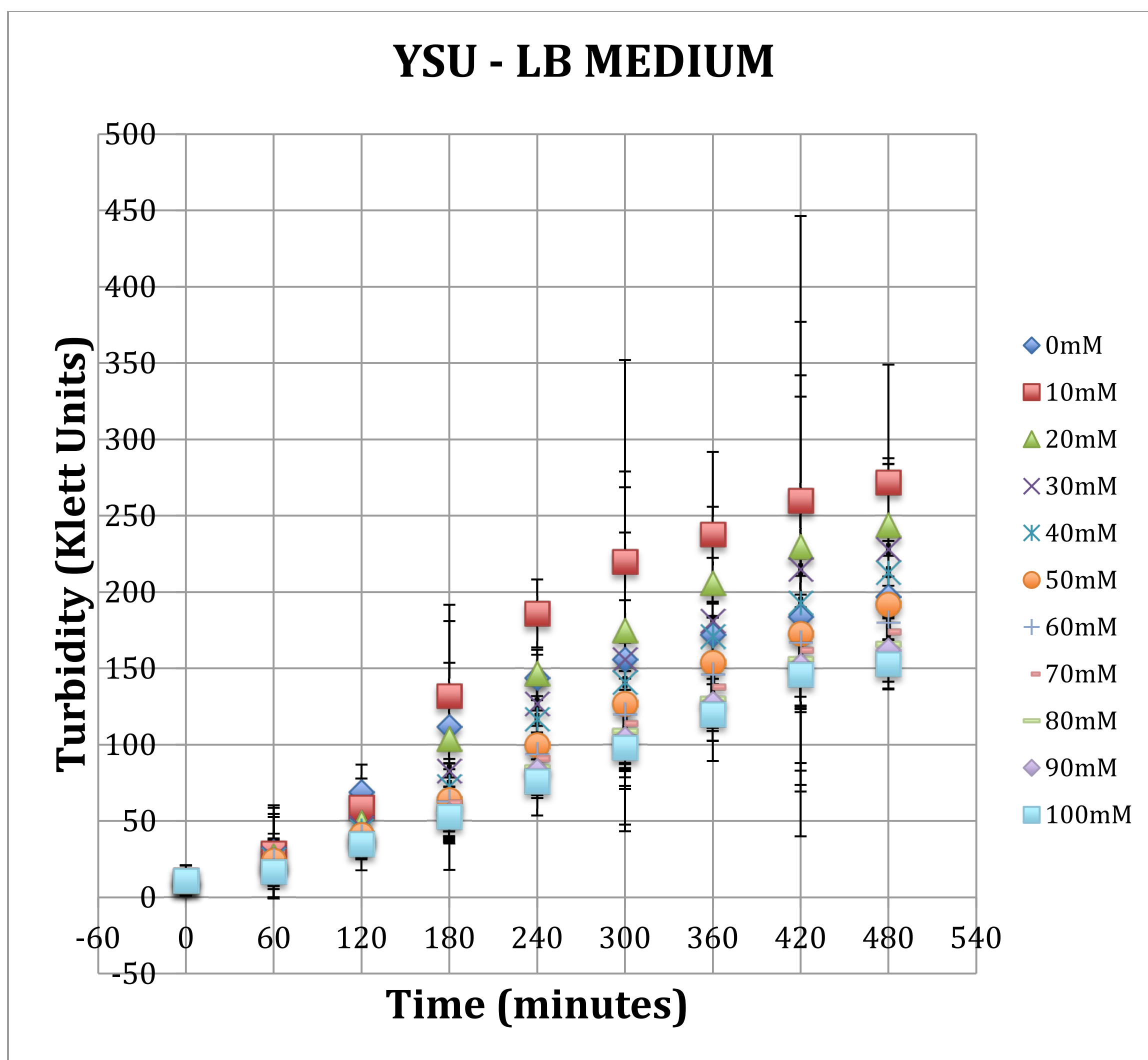

Figure 3. MIC Growth Curve for *Enterobacter* sp. YSU, Wild Type Strain. Dark blue diamonds – no selenite control; red squares – 10 mM selenite; green triangles – 20 mM selenite; purple x – 30 mM selenite; blue asterisks – 40 mM selenite; orange circles – 50 mM selenite; grey crosses – 60 mM selenite; pink hyphens – 70 mM selenite; green dashes – 80 mM selenite; purple diamonds – 90 mM selenite and light blue squares – 100 mM selenite. The growth of the wild type in LB medium showed a significant increase in cell number at high selenite concentrations of 40 mM to 100 mM with time. The values at different time points represent average turbidity values of 3 experiments with error calculated using the student t test at 95% confidence level.

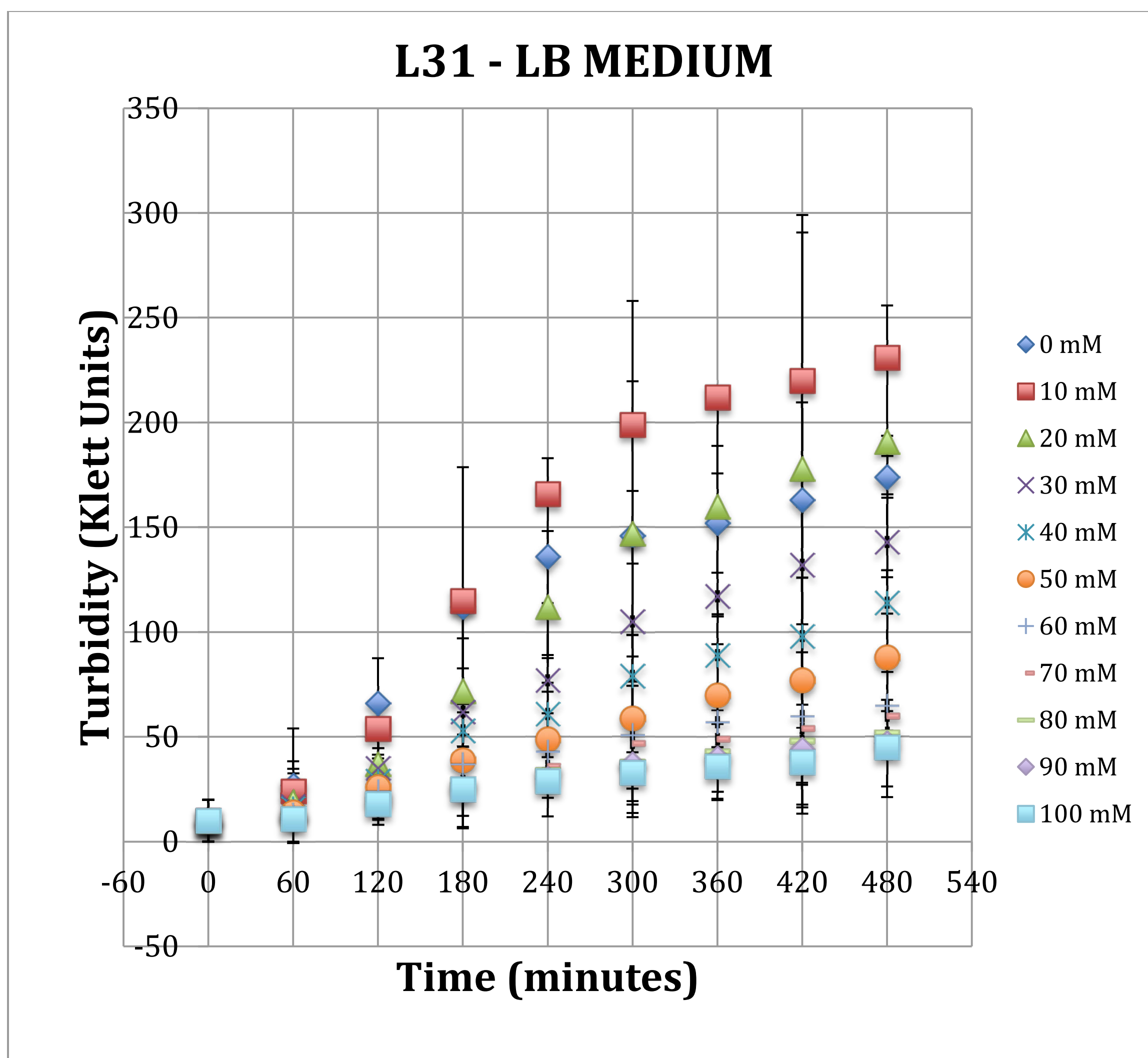

Figure 4. L31 – MIC Growth Curve for *Enterobacter* sp. YSU, L31 Mutant. Dark blue diamonds – no selenite control; red squares – 10 mM selenite; green triangles – 20 mM selenite; purple x – 30 mM selenite; blue asterisks – 40 mM selenite; orange circles – 50 mM selenite; grey crosses – 60 mM selenite; pink hyphens – 70 mM selenite; green dashes – 80 mM selenite; purple diamonds – 90 mM selenite and light blue squares – 100 mM selenite. Compared to the wild type strain, the mutant demonstrated significant growth inhibition at 70, 80, 90 and 100 mM selenite.

5.3 Genomic DNA Digestions

Genomic DNA was isolated from the L31 mutant and partially digested with the restriction enzymes, *Bfu*C I and *Hin*d III (Figure 5). In addition, a complete digestion was carried out using *Eco*R I. *Bfu*C I recognizes the site GATC and cuts both inside and outside the transposome. It was therefore used in partial a digestion to shear the genomic DNA which was ligated and transformed into ECD100D *pir* electrocompetent *E. coli* cells. The resulting transformants contained a plasmid with the transposon and a host DNA region flanking the transposon insertion site. After obtaining part of the sequence that contained the transposon interrupted gene, *Hin*d III and *Eco*R I were used to identify a larger segment of the interrupted region. .

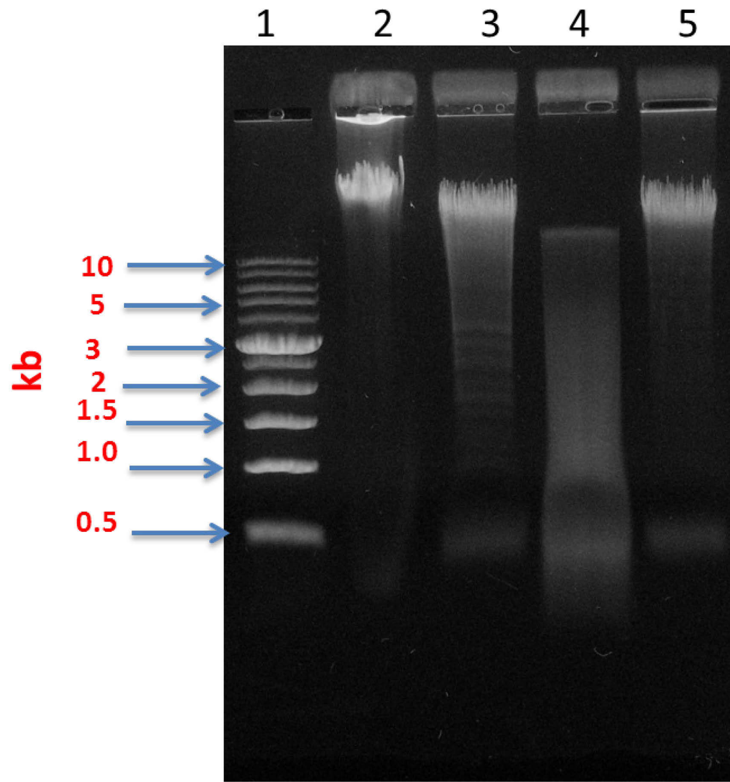

Figure 5. Genomic DNA Digestions. Lane 1 – 1kb ladder, Lane 2 – undigested L31, Lane 3 – *Hin*d III digestion, Lane 4 – *Bfu*C I *digestion*, Lane 5 – *Eco*R I digestion

5.4 Gene Rescue

Transformation into electrocompetent *E.coli* cells resulted in 251 transformants for *Eco*R I , 11 transformants for *Bfu*C I and 22 transformants for *Hin*d III. Plasmid DNA was purified from the transformants and linerized using *Xho* I restriction enzyme as illustrated in Figure 6. Digestion of plasmid DNA was carrried out in order to determine the actual size of linerized plasmid, helping to determine the volume of plasmid DNA required for sequencing reactions. Using the ladder in lane 1, the sizes of each plasmid were estimated as listed in Table 2. As mentioned earlier, *Eco*R I and *Hin*d III were used to obtain larger fragments of the interrupted region.The undigested *Hin*d III samples generated plasmids which were larger than 10 kb (lane 6 and 8) whereas the digested samples generated plasmids with 2 bands (lane 7 and 9). Therefore, the actual sizes of the plasmids were determined by adding the sizes of the two bands. Undigested *Eco*R I-L31 sample generated a 6 kb plasmid (lane 2) and the digested sample generated an 8 kb plasmid (lane 3) whereas the undigested *Bfu*C I-L31 sample generated a 4 kb plasmid and the digested sample generated a 5 kb plasmid.

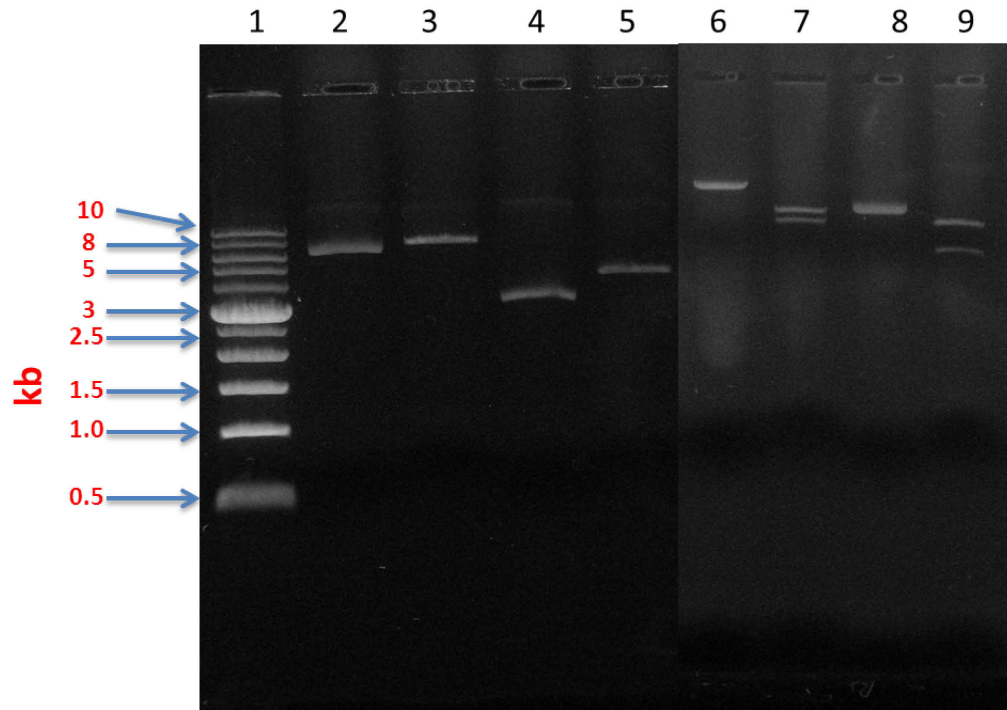

Figure. 6. Purified Plasmid DNA. The enzyme *Xho* I was used to linearize the purified plasmid DNA. Lane 1 – 1kb ladder, Lane 2 – undigested *Eco*R I-L31 plasmid, Lane 3 – digested *Eco*R I-L31 plasmid, Lane 4 – undigested *Bfu*C I-L31 plasmid, Lane 5 digested *Bfu*C I-L31 plasmid, Lane 6 – undigested *Hin*d III-L31 plasmid, Lane 7 – digested *Hin*d III-L31 plasmid, Lane 8 – undigested *Hin*d III-L31 plasmid, Lane 9 digested *Hin*d III-L31 plasmid.

Table. 2 Sizes of Linerized Plasmid DNA

		Undigested Plasmid	***Xho* I Digested plasmid**
Fig. 6 lane 3	*Eco*R I – L31 plasmid	**6kb**	**8kb**
Fig. 6 lane 5	*Bfu* CI – L31 plasmid	**4kb**	**5kb**
Fig. 6 lane 8	*Hin*d III – L31 plasmid	**>10kb**	**(10 + 11) 21**
Fig. 6 lane 10	*Hin*d III – L31 plasmid	**>10kb**	**(10 + 6) 16**

5.5 Sequence Analysis

In order to identify the sequence of the gene that was interrupted by the EZ-Tn5 transposon, the primers KAN-2 FP-1 and R6KAN-2 RP-1 (Table 1) that were specific for the transposon were used. BLAST analysis showed that the interrupted gene in the selenite sensitive mutant, L31, was similar to a signal transduction histidine kinase in *Enterobacter cloacae* subsp. *cloacae* (Figure 7).

Enterobacter cloacae subsp. cloacae NCTC 9394 draft genome
Sequence ID: **emb|FP929040.1|** Length: 4908759 Number of Matches: 1
Range 1: 66514 to 67028

Score	Expect	Identities	Gaps	Strand	Frame
941 bits(509)	0.0()	513/515(99%)	0/515(0%)	Plus/Plus	

Features:
Signal transduction histidine kinase

```
Query  1      ATGCTGTCCAGGCGATGCGCTTCCGTCTCAATACGCTCCAGCTCTTTGCTCTCACCGCTG  60
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Sbjct  66514  ATGCTGTCCAGGCGATGCGCTTCCGTCTCAATACGCTCCAGCTCTTTACTCTCACCGCTG  66573

Query  61     CGGCGGCGCAGCAGGGCGGTACCCAGCTGTAAGCGCGTGAGCGGCGTCCGCAGCTCGTGG  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  66574  CGGCGGCGCAGCAGGGCGGTACCCAGCTGTAAGCGCGTGAGCGGCGTCCGCAGCTCGTGG  66633

Query  121    GAGATATCCGACAGCAGGCGTTGCTGCGCGGTCATCATGCGGTCAAGAGCAGAGACCATC  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  66634  GAGATATCCGACAGCAGGCGTTGCTGCGCGGTCATCATGCGGTCAAGAGCAGAGACCATC  66693

Query  181    TGGTTAAAACTGGTTCCGGCGGCCAGGAACTCCTGCGGCCCGGATTCCAGCTCCGGATGC  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  66694  TGGTTAAAACTGGTTCCGGCGGCCAGGAACTCCTGCGGCCCGGATTCCAGCTCCGGATGC  66753

Query  241    TGCCTCAGGTTGCCCTGCGCCACTTCGTCCGCCGCATTTTTCAGCTTACGCGCCGGTTTC  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  66754  TGCCTCAGGTTGCCCTGCGCCACTTCGTCCGCCGCATTTTTCAGCTTACGCGCCGGTTTC  66813

Query  301    GCCAGGCTCCACGCCAGCCATAACAGCAGCGGGGAACTGACCAGCATCGTGACAATCAAC  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  66814  GCCAGGCTCCACGCCAGCCATAACAGCAGCGGGGAACTGACCAGCATCGTGACAATCAAC  66873

Query  361    AGCAAAAGCGGACGGTCAAACAGCAGGTTGATAAAGTCGGACTGTGAACTGCTCGCAGGG  420
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Sbjct  66874  AGCAAAAGCGGACGGTCAAACAGCAGGTTGATAAAGTCGGACTGGGAACTGCTCGCAGGG  66933

Query  421    CGAATCAGATAGAGCTGGTAATTGTCTTCCCCATCCCGCACCGAGAAAGGACCGACCATC  480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  66934  CGAATCAGATAGAGCTGGTAATTGTCTTCCCCATCCCGCACCGAGAAAGGACCGACCATC  66993

Query  481    TCTACCCGACCATATTTCTTTTTCTGTGGGTGATC  515
              |||||||||||||||||||||||||||||||||||
Sbjct  66994  TCTACCCGACCATATTTCTTTTTCTGTGGGTGATC  67028
```

Figure 7. BLAST analysis-Histidine Kinase. The gene that was interrupted by the transposon was similar to a signal transduction histidine kinase in *Enterobacter cloacae subsp. cloacae* with 99% sequence similarity. The “Query” was the submitted sequence from the L31 mutant, and the “Sbjct” was the sequence that matched to the submitted sequence.

The DNA sequence of the sensor protein was used to design primers (Table. 1) to further resolve the DNA sequence of the genes that flanked the sensor kinase. Additional sequencing primers were designed as addition regions around the sensor kinase was resolved. (Table. 1). BLAST analysis of the obtained sequences indicated that some of the genes that flanked the sensor kinase were similar to the DNA-binding transcriptional regulator, *cpxR* (Figure 8) and periplasmic repressor, *cpxP* (Figure 9). Other segments matched to genes for a ferrous iron efflux protein F and 6-phosphofructokinase.

Enterobacter cloacae subsp. cloacae ENHKU01, complete genome
Sequence ID: **gb|CP003737.1|** Length: 4726582 Number of Matches: 1
Range 1: 4647654 to 4648352

Score	Expect	Identities	Gaps	Strand	Frame
1075 bits(582)	0.0()	660/699(94%)	0/699(0%)	Plus/Minus	

Features:
DNA-binding transcriptional regulator CpxRtwo-component sensor protein

```
Query  1        TCATGAAGCGGAAACCATCAGATAACCGCGACCACGCAGGGTTTTAAACCAGGGATGACC  60
                |||||||||||||||||||||||| |||||||||||||||||||||||||| || |||||
Sbjct  4648352  TCATGAAGCGGAAACCATCAGATAGCCGCGACCACGCAGGGTTTTAAACCATGGGTGACC  4648293

Query  61       GTCTTTACGCTCCGGCAGCTTACGGCGCAGGTTAGAGATGTGCATGTCGATGGCGCGATC  120
                |||||||||||| |||||||| ||||||||||||||||||||||||||||||||||| ||
Sbjct  4648292  GTCTTTACGCTCTGGCAGCTTGCGGCGCAGGTTAGAGATGTGCATGTCGATGGCGCGGTC  4648233

Query  121      AAACGGGGTGAGGCGTTTGCCGAGCACTTCCTGGCTTAAGTGTTCACGCGACACCACCTG  180
                |||||||||||| |||||||| ||||||||||| || | |||||||||||| ||||||||
Sbjct  4648232  AAACGGGGTGAGACGTTTGCCCAGCACTTCCTGACTCAGGTGTTCACGCGATACCACCTG  4648173

Query  181      GCCGAGATGCTGCGCCAGCAGATACAGCAGGGTGAACTCGGTGCCGGTTAGCTCCAGCGT  240
                |||||| ||||||||||||||||| |||||||||||||| |||||||| | |||||| ||
Sbjct  4648172  GCCGAGGTGCTGCGCCAGCAGATAAAGCAGGGTGAACTCTGTGCCGGTCAACTCCAGGGT  4648113

Query  241      TTCACCGTCGAAGCTCGCTTCCTGGCGGCCCGGGTTCAGGCTCAGGGAGTCCACTTCCAG  300
                 | ||||||||||||||||||||| || || |||||||||||||||||||||||||||||
Sbjct  4648112  CTGACCGTCGAAGCTCGCTTCCTGACGACCTGGGTTCAGGCTCAGGGAGTCCACTTCCAG  4648053

Query  301      CGTAGGTGAGCTGTTGTCGGTATTCTGCTGCTGCTCGCTCCAGTGAGAGCGGCGCAGGAT  360
                ||| ||||||||||||||||||||||||||||||||||||||||| || |||||||||||
Sbjct  4648052  CGTGGGTGAGCTGTTGTCGGTATTCTGCTGCTGCTCGCTCCAGTGGGAACGGCGCAGGAT  4647993

Query  361      CGCGCGAATACGGGCCACCAGTTCACGGTCGTTGAACGGCTTAGGTAAATAGTCATCCGC  420
                 |||||||||||||||||||||||||| ||||||||||| || |||||||||||||||||
Sbjct  4647992  AGCGCGAATACGGGCCACCAGTTCACGATCGTTGAACGGTTTCGGTAAATAGTCATCCGC  4647933

Query  421      ACCCAGCTCAAGGCCGAGTACGCGGTCAAGCTCGCTGCCGCGTGCAGTCAGCATAATGAC  480
                |||||||||||||||||| ||||||||||| ||||||||||| || |||||||| || ||
Sbjct  4647932  ACCCAGCTCAAGGCCGAGAACGCGGTCAAGTTCGCTGCCGCGGGCGGTCAGCATGATAAC  4647873

Query  481      GGGGGTCTGGTGTGTCTGGCGAAGCTCTTTCAACGTATCAATACCGTTTTTCTTCGGCAT  540
                |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Sbjct  4647872  GGGGGTCTGGTGTGTCTGGCGAAGCTCTTTCAGCGTATCAATACCGTTTTTCTTCGGCAT  4647813

Query  541      CATGACGTCGAGCAAAAGTAAATCGATGCTGTCGTCAAGGAGACTCAGCGCCTGCTCGCC  600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4647812  CATGACGTCGAGCAAAAGTAAATCGATGCTGTCGTCAAGGAGACTCAGCGCCTGCTCGCC  4647753

Query  601      ATCATGGGCAACCAGGACGTTGAAACCTTCCATGTCGAGCAACTCTTTTAAAAGAGATGT  660
                ||||||||||||||||||||||||||||||||||||||||||||| |||||||| |||||
Sbjct  4647752  ATCATGGGCAACCAGGACGTTGAAACCTTCCATGTCGAGCAACTCCTTTAAAAGGGATGT  4647693

Query  661      GAGCTCTCGGTCATCATCAACTAACAGGATTTTATTCAT  699
                |||||||||||||||||||||||||||||||||||||||
Sbjct  4647692  GAGCTCTCGGTCATCATCAACTAACAGGATTTTATTCAT  4647654
```

Figure 8. BLAST analysis-Transcriptional Regulator, CpxR. The region flanking the sensor kinase was sequenced and one of the identified genes was similar to the DNA-binding transcriptional regulator, *cpxR*, found in *Enterobacter cloacae* subsp. *cloacae* with 94% sequence similarity.

Enterobacter cloacae subsp. dissolvens SDM, complete genome
Sequence ID: gb|**CP003678.1**| Length: 4968248 Number of Matches: 1
Range 1: 4865405 to 4865912

Score	Expect	Identities	Gaps	Strand	Frame
761 bits(412)	0.0()	478/510(94%)	4/510(0%)	Plus/Minus	

Features:
periplasmic repressor CpxP

```
Query  1        ATGCGCAAAGTTACCGCTGCCGTCATGGCCTCAACGCTGGCCTTCAGTGCGTTTAGCCAG  60
                ||||||||||||||||||||||||||||||||||||||||| |||||||| |||||||||
Sbjct  4865912  ATGCGCAAAGTTACCGCTGCCGTCATGGCCTCAACGCTGGCGTTCAGTGCTTTTAGCCAG  4865853

Query  61       GCTGCTGTAGCTATCATCAGCGATAACGGTTCCTCAGCAGAGGGCGCAACGCAGCACAGC  120
                || ||||||||||||||| ||||||||| |||||||  |||||||  |||||||||||||
Sbjct  4865852  GCCGCTGTAGCTATCATCGGCGATAACGAGTCCTCACAAGAGGGCATAACGCAGCACAGC  4865793

Query  121      AGCCAAAGCCATATGTTTGACGGCATAAGTTTAACCGAACATCAGCGTCAACAGATGCGA  180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4865792  AGCCAAAGCCATATGTTTGACGGCATAAGTTTAACCGAACATCAGCGTCAACAGATGCGA  4865733

Query  181      GATCTGATGCAGAGGGCGAGACACGACCAGCCCCCTGTTAATGTTAGCGAAATGGAGACA  240
                |||||||||||||||||  ||||||| ||||||||||||||||||||||||||||||| |
Sbjct  4865732  GATCTGATGCAGAGGGCACGACACGATCAGCCCCCTGTTAATGTTAGCGAAATGGAGATA  4865673

Query  241      ATGCATCGCCTTGTCACCGCAGAAAATTTTGACGAAAGCGCTGTACGCGCTCAGGCCGAA  300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4865672  ATGCATCGCCTTGTCACCGCAGAAAATTTTGACGAAAGCGCTGTACGCGCTCAGGCCGAA  4865613

Query  301      AAAATGGCGCAGGAACAGGTTGCCCGTCAGGTAGAGATGGCGAAGGTTCGCAACCAGATG  360
                |||||||| ||||| ||||||||||| |||||||| ||||||||||| ||||||||||||
Sbjct  4865612  AAAATGGCACAGGAGCAGGTTGCCCGCCAGGTAGAAATGGCGAAGGTCCGCAACCAGATG  4865553

Query  361      TTCCACCTGCTAACGCCCGAGCAGCAAGCGGTTTTGAATACCA-AA-CATCAGCAACGTA  418
                ||||||||| || |||||||||||||||||||||||||  ||| || ||||||||||| |
Sbjct  4865552  TTCCACCTGTTATCGCCCGAGCAGCAAGCGGTTTTGAA--CCAGAAACATCAGCAACGAA  4865495

Query  419      TGAATCAGTTGCGTGAGGTTGCACGGATGCAGCGAAGCTCAGATATGACGCTTTTCAGTA  478
                || | ||||||||||||||||||||||||||||||||||||||   |||| |||||||||
Sbjct  4865494  TGGACCAGTTGCGTGAGGTTGCACGGATGCAGCGAAGCTCAGAAGCGACGTTTTTCAGTA  4865435

Query  479      GCAATAGCAGTACCCGTAGTAACCAGTAAA  508
                ||||||||||||||||||||||||||||||
Sbjct  4865434  GCAATAGCAGTACCCGTAGTAACCAGTAAA  4865405
```

Figure 9. BLAST analysis-Periplasmic Repressor, CpxP. The other identified gene was the periplasmic repressor, *cpxP* found in *Enterobacter cloacae* subsp. *dissolvens* with 94% sequence similarity.

The obtained sequences of these genes were assembled using ContigExpress from the Vector NTI Advance® 11.5.0 software package and a segment of the related *Enterobacter cloacae* subspecies cloacae NCTC9393 sequence (Accession Number FP929040) was used as a guide (Figure 10). The solid green arrows in figure 10 represent DNA sequence of the sensor kinase and the genes flanking the sensor kinase. It also designated the directions of transcription. As illustrated, the sensor kinase is transcribed in the same direction as *CpxR* whereas the other genes were transcribed in the opposite direction. Further downstream is a sulfur binding protein whose complete sequence was not identified, hence not indicated in Figure 8. The EZ-Tn5 Transposon inserted near the 5' end of the histidine kinase sensor (Figure 8). Identification of the genes flanking the sensor kinase was important in determining if the sensor kinase played a role in regulating nearby genes.

Vector NTI was also used to show the location of the different primers (Table 1) used in sequencing reactions together with the recognition sites of the different restriction endonuleases. (Figure 11). The words in red represent the primers used in sequencing reactions whereas the blue words represent the restriction endonuclease sites. The DNA sequence of the histidine kinase sensor is represented by the base pairs 1,503 to 2876 whereas the transposon insert is between base pairs 2,017 and 2,025.

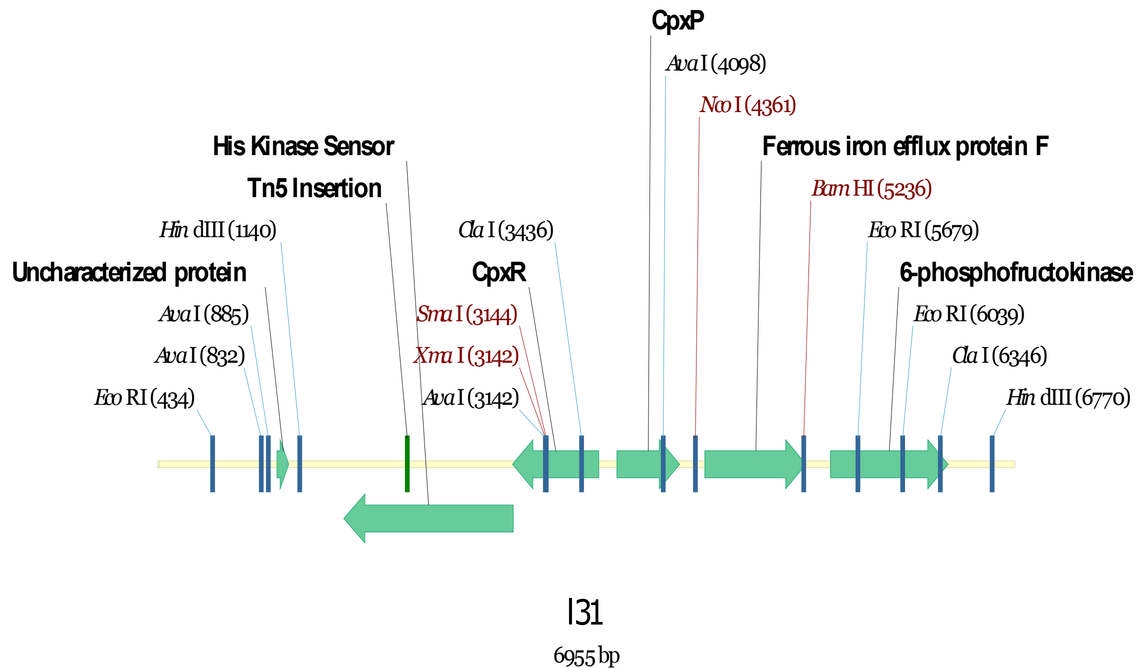

Figure 10. L31 Feature Map. Vector NTI Advance® 11.5.0 software was used to construct a map showing the EZ-Tn5 transposon insertion site and some of the genes that flank the histidine kinase sensor. The green arrows represent the sequence of the identified genes, and the blue lines represent different restriction endonuclease recognition sites. Digestions of this region with *Eco*R I should result in a 5,605 bp fragment.

131

```
                                                  EcoRI/AvaI/HindIII/primer annotations shown inline
   1  ATTTTCAATA TCATTTAATT AACTATAATG AACCAACTGC TTACGCGGCA TTAACAGCTG TGCCGCCCGA CAATAATGGA GAGGATTATG AGTTATACAC
      TAAAAGTTAT AGTAAATTAA TTGATATTAC TTGGTTGACG AATGCGCCGT AATTGTCGAC ACGGCGGGCT GTTATTACCT CTCCTAATAC TCAATATGTG
 101  TGCCATCCCT GCCGTATGCC TACGACGCAC TGGAACCGCA TTTCGACAAG CAGACGATGG AAATCCATCA CACTAAACAC CACCAGACCT ATGTGAACAA
      ACGGTAGGGA CGGCATACGG ATGCTGCGTG ACCTTGGCGT AAAGCTGTTC GTCTGCTACC TTTAGGTAGT GTGATTTGTG GTGGTCTGGA TACACTTGTT
 201  CGCGAATGCT GCGCTGGAAA GCCTGCCAGA GTTCGCTAAT CTGCCTGTTG AAGAGCTGAT CACCAAACTG GACCAGCTGC CGGCAGACAA GAAAACCGTG
      GCGCTTACGA CGCGACCTTT CGGACGGTCT CAAGCGATTA GACGGACAAC TTCTCGACTA GTGGTTTGAC CTGGTCGACG GCCGTCTGTT CTTTTGGCAC
 301  CTGCGCAACA ACGCGGGCGG TCACGCTAAC CACAGCCTGT TCTGGAAAGG CCTGAAAACC GGCACCACCC TTCAGGGCGA CCTGAAAGCG GCTATCGAGC
      GACGCGTTGT TGCGCCCGCC AGTGCGATTG GTGTCGGACA AGACCTTTCC GGACTTTTGG CCGTGGTGGG AAGTCCCGCT GGACTTTCGC CGATAGCTCG
                                         EcoRI
 401  GTGACTTCGG TTCCGTTGAC AACTTCAAAG CGGAATTCGA AAAAGCCGCT GCAACCCGTT TCGGCTCTGG CTGGGCGTGG CTGGTCCTGA AAGGTGACAA
      CACTGAAGCC AAGGCAACTG TTGAAGTTTC GCCTTAAGCT TTTTCGGCGA CGTTGGGCAA AGCCGAGACC GACCCGCACC GACCAGGACT TTCCACTGTT
 501  ACTGGCGGTC GTTTCTACCG CTAACCAGGA CTCCCCGCTG ATGGGTGAAG CTATCTCTGG CGCATACGAC TACCCAAGTA TGCGCCCGCA AGTGGGGGAA
      TGACCGCCAG CAAAGATGGC GATTGGTCCT GAGGGGCGAC TACCCACTTC GATAGAGACC GCGTATGCTG ATGGGTTCAT ACGCGGGCGT TCACCCCCTT
 601  CACGCGTACT AAACAAAGTA CCAAAACCGT CGCAATAACG ACACCAAAGC AAACCAAAAC CTAGGGAACG GAGACAAAGC ACCAGCCCGC AGCGCCGCTA
      GTGCGCATGA TTTGTTTCAT GGTTTTGGCA GCGTTATTGC TGTGGTTTCG TTTGGTTTTG GATCCCTTGC CTCTGTTTCG TGGTCGGGCG TCGCGGCGAT
 701  AAAAAAAAAA TAGCATTGAA GCACGCAAGA AAAGAGAAGG ATAACACGCG CTTTTTGGAA CTGAAGGAGT GGAGATGCAT TACCCGGTGA ATGTGTTTAC
      TTTTTTTTTT ATCGTAACTT CGTGCGTTCT TTTCTCTTCC TATTGTGCGC GAAAAACCTT GACTTCCTCA CCTCTACGTA ATGGGCCACT TACACAAATG
                                        AvaI                                                   AvaI
 801  AGGCAAGGTA AGGGAGTACG AAGGCAGCCG CCCGAGCGCC ATCGACCAAA GTGCAGGTCA ACGGTGAGCT GACGTTAACC GATCTCGGGC TTGCGGGTGA
      TCCGTTCCAT TCCCTCATGC TTCCGTCGGC GGGCTCGCGG TAGCTGGTTT CACGTCCAGT TGCCACTCGA CTGCAATTGG CTAGAGCCCG AACGCCCACT
 901  CCAGCAGGCT GAAAAGAAAA TCCACGGCGG GCCCGATCGC GCGCTGTGCC ACTATCCGCG CGAGCACTAT CAGCACTGGA AAGCCGAATT TCCTGAACAG
      GGTCGTCCGA CTTTTCTTTT AGGTGCCGCC CGGGCTAGCG CGCGACACGG TGATAGGCGC GCTCGTGATA GTCGTGACCT TTCGGCTTAA AGGACTTGTC
1001  GCGGACCTCT TCGTCGCCCC CGCGTTTGGC GAAAATCTCT CAACGGAGGG GCTGACGGAG AAGAACGTCT TTATCGGCGA CATTTACCGC TGGGGCGATG
      CGCCTGGAGA AGCAGCGGGG GCGCAAACCG CTTTTAGAGA GTTGCCTCCC CGACTGCCTC TTCTTGCAGA AATAGCCGCT GTAAATGGCG ACCCCGCTAC
                                                  HindIII
1101  CTTTGATTCA GGTCACCCAG CCGCGCTCAC CGTGCTTTAA GCTTAATTAC CATTTCGGCA TTCAGGATAT GTCGGCCCAG CTGCAAAACG CGGGTAAAAC
      GAAACTAAGT CCAGTGGGTC GGCGCGAGTG GCACGAAATT CGAATTAATG GTAAAGCCGT AAGTCCTATA CAGCCGGGTC GACGTTTTGC GCCCATTTTG
1201  CGGCTGGCTG TATCGCGTTG TGCAGGCGGG ACAGGTTTCG GCGGATGCGC CGCTTGAGCT GGCTTCGCGC TTGAGTGAGG TGTCGGTATA CGAGGCCTGC
      GCCGACCGAC ATAGCGCAAC ACGTCCGCCC TGTCCAAAGC CGCCTACGCG GCGAACTCGA CCGAAGCGCG AACTCACTCC ACAGCCATAT GCTCCGGACG
1301  GCAATTGCCT GGCATATGCC GTTTGATGAC GAGCAGTATC ACCGCCTGCT GTCAGCGGCG GGATTATCCA CCAGCTGGAC CAGAACGATG CAGAAGCGGC
      CGTTAACGGA CCGTATACGG CAAACTACTG CTCGTCATAG TGGCGGACGA CAGTCGCCGC CCTAATAGGT GGTCGACCTG GTCTTGCTAC GTCTTCGCCG
1401  GGTTAAGCGG CAAGATCGAG AGCAGTTCGC GGAGATTGTG GGGGAAAGAG ATTCCGTAGA AAGTAGGCCG GGTAAGGCGA AGCCGCCACC CGGCACATAC
      CCAATTCGCC GTTCTAGCTC TCGTCAAGCG CCTCTAACAC CCCCTTTCTC TAAGGCATCT TTCATCCGGC CCATTCCGCT TCGGCGGTGG GCCGTGTATG
                    Sen R 100.0%
1501  GGTTACGAAC GCTTATACAG CGGCAGCCAC AGCGTTAACC TCAGCCCGCC CAGCGGGCTG TCATCGGCTT TCACCCAGCC ACGGTGCTGT TGCATGGCGG
      CCAATGCTTG CGAATATGTC GCCGTCGGTG TCGCAATTGG AGTCGGGCGG GTCGCCCGAC AGTAGCCGAA AGTGGGTCGG TGCCACGACA ACGTACCGCC
                    Sen F2 100.0%
1601  TTTCAACAAT CGCCAGACCT AGCCCCGTAC CACCAGATTC ACGGTCGCGC GCCTCGTCGG TGCGGTAGAA CGGACGGAAA ATCTGCTCGC GATCTTCCGG
      AAAGTTGTTA GCGGTCTGGA TCGGGGCATG GTGGTCTAAG TGCCAGCGCG CGGAGCAGCC ACGCCATCTT GCCTGCCTTT TAGACGAGCG CTAGAAGGCC
1701  GCTGACGCCA GGACCATCGT CATCGACAAT GACGGTAATC CCGTCTTTAT CCACCGAAAA CGCCACTTCA ATCTTCGTGT GCGAATAGCG CAGGGCGTTA
      CGACTGCGGT CCTGGTAGCA GTAGCTGTTA CTGCCATTAG GGCAGAAATA GGTGGCTTTT GCGGTGAAGT TAGAAGCACA CGCTTATCGC GTCCCGCAAT
1801  CGCACGATAT TCTCCAGCGC GCTTTCCAGC GTGTTGGGGT TACCGTAGAG CGGCCACGGG CCCGGCGGGA AGTTGACGGT GAACGATTTG CCCATCTGCT
      GCGTGCTATA AGAGGTCGCG CGAAAGGTCG CACAACCCCA ATGGCATCTC GCCGGTGCCC GGGCCGCCCT TCAACTGCCA CTTGCTAAAC GGGTAGACGA
1901  CCGCTTCGAA CGCCGCGTTG TCCAGCACCT CATGCCAGAG GTGATTGGCT TTCACCGTTT CGCTCACAAG CGCGTTTTTC TGCTGATTGC GAGACATCAC
      GGCGAAGCTT GCGGCGCAAC AGGTCGTGGA GTACGGTCTC CACTAACCGA AAGTGGCAAA GCGAGTGTTC GCGCAAAAAG ACGACTAACG CTCTGTAGTG
2001  CAGCAGATCG TTGATCATGC TGTCCAGGCG ATGCGCTTCC GTCTCAATAC GCTCCAGCTC TTTGCTCTCA CCGCTGCGGC GGCGCAGCAG GGCGGTACCC
      GTCGTCTAGC AACTAGTACG ACAGGTCCGC TACGCGAAGG CAGAGTTATG CGAGGTCGAG AAACGAGAGT GGCGACGCCG CCGCGTCGTC CCGCCATGGG
2101  AGCTGTAAGC GCGTGAGCGG CGTCCGCAGC TCGTGGGAGA TATCCGACAG CAGGCGTTGC TGCGCGGTCA TCATGCGGTC AAGAGCAGAG ACCATCTGGT
      TCGACATTCG CGCACTCGCC GCAGGCGTCG AGCACCCTCT ATAGGCTGTC GTCCGCAACG ACGCGCCAGT AGTACGCCAG TTCTCGTCTC TGGTAGACCA
2201  TAAAACTGGT TCCGGCGGCC AGGAACTCCT GCGGCCCGGA TTCCAGCTCC GGATGCTGCC TCAGGTTGCC CTGCGCCACT TCGTCCGCCG CATTTTTCAG
      ATTTTGACCA AGGCCGCCGG TCCTTGAGGA CGCCGGGCCT AAGGTCGAGG CCTACGACGG AGTCCAACGG GACGCGGTGA AGCAGGCGGC GTAAAAAGTC
2301  CTTACGCGCC GGTTTCGCCA GGCTCCACGC CAGCCATAAC AGCAGCGGGG AACTGACCAG CATCGTGACA ATCAACAGCA AAAGCGGACG GTCAAACAGC
      GAATGCGCGG CCAAAGCGGT CCGAGGTGCG GTCGGTATTG TCGTCGCCCC TTGACTGGTC GTAGCACTGT TAGTTGTCGT TTTCGCCTGC CAGTTTGTCG
                                                                                               CDF F3 78.9%
2401  AGGTTGATAA AGTCGGACTG TGAACTGCTC GCAGGGCGAA TCAGATAGAG CTGGTAATTG TCTTCCCCAT CCCGCACCGA GAAAGGACCG ACCATCTCTA
      TCCAACTATT TCAGCCTGAC ACTTGACGAG CGTCCCGCTT AGTCTATCTC GACCATTAAC AGAAGGGGTA GGGCGTGGCT CTTTCCTGGC TGGTAGAGAT
                                                                                                Sen F3 100.0%
2501  CCCGACCATA TTTCTTTTTC TGTGGGTGAT CGGCATTATC CGCCTGGCCA ATAAAGTTGC GGATAATCTG CATTTCATTG CGATCTGCCC CAATCACGCG
      GGGCTGGTAT AAAGAAAAAG ACACCCACTA GCCGTAATAG GCGGACCGGT TATTTCAACG CCTATTAGAC GTAAAGTAAC GCTAGACGGG GTTAGTGCGC
      Sen F3 100.0
```

```
2601  GCCTTCACTG GTGACCAGTA AAAGACGTTG TCCGGGGGGC GCCCACTTGT CGATAGCGCG AAACAACCTG CGCCACCACA TTAAATCGTT CGGCGGATCG
      CGGAAGTGAC CACTGGTCAT TTTCTGCAAC AGGCCCCCCG CGGGTGAACA GCTATCGCGC TTTGTTGGAC GCGGTGGTGT AATTTAGCAA GCCGCCTAGC

2701  TTTGCCAGCT CCGCTTCCAC GTGCTGCTCG ATCATCACGC CCTGACGTTG CTCGCTGTCG AGTAGCTCCG TCATCTGGCG TGAGTCGAGT TTTGGCAACA
      AAACGGTCGA GGCGAAGGTG CACGACGAGC TAGTAGTGCG GGACTGCAAC GAGCGACAGC TCATCGAGGC AGTAGACCGC ACTCAGCTCA AAACCGTTGT

2801  TCAAAACGAG CATTAAAACC AGTGCCAGCG TCAGCCAGAA GATGGCGAAG ATGCGGGCGG TTAAGCTGCC TATCATGAAG CGGAAACCAT CAGATAACCG
      AGTTTTGCTC GTAATTTTGG TCACGGTCGC AGTCGGTCTT CTACCGCTTC TACGCCCGCC AATTCGACGG ATAGTACTTC GCCTTTGGTA GTCTATTGGC

2901  CGACCACGCA GGGTTTTAAA CCAGGGATGA CCGTCTTTAC GCTCCGGCAG CTTACGGCGC AGGTTAGAGA TGTGCATGTC GATGGCGCGA TCAAACGGGG
      GCTGGTGCGT CCCAAAATTT GGTCCCTACT GGCAGAAATG CGAGGCCGTC GAATGCCGCG TCCAATCTCT ACACGTACAG CTACCGCGCT AGTTTGCCCC

3001  TGAGGCGTTT GCCGAGCACT TCCTGGCTTA AGTGTTCACG CGACACCACC TGGCCGAGAT GCTGCGCCAG CAGATACAGC AGGGTGAACT CGGTGCCGGT
      ACTCCGCAAA CGGCTCGTGA AGGACCGAAT TCACAAGTGC GCTGTGGTGG ACCGGCTCTA CGACGCGGTC GTCTATGTCG TCCCACTTGA GCCACGGCCA

                                                 SmaI
                                                 XmaI
                                                 AvaI
3101  TAGCTCCAGC GTTTCACCGT CGAAGCTCGC TTCCTGGCGG CCCGGGTTCA GGCTCAGGGA GTCCACTTCC AGCGTAGGTG AGCTGTTGTC GGTATTCTGC
      ATCGAGGTCG CAAAGTGGCA GCTTCGAGCG AAGGACCGCC GGGCCCAAGT CCGAGTCCCT CAGGTGAAGG TCGCATCCAC TCGACAACAG CCATAAGACG

3201  TGCTGCTCGC TCCAGTGAGA GCGGCGCAGG ATCGCGCGAA TACGGGCCAC CAGTTCACGG TCGTTGAACG GCTTAGGTAA ATAGTCATCC GCACCCAGCT
      ACGACGAGCG AGGTCACTCT CGCCGCGTCC TAGCGCGCTT ATGCCCGGTG GTCAAGTGCC AGCAACTTGC CGAATCCATT TATCAGTAGG CGTGGGTCGA
                                                                                         Sen F4 94.7%

3301  CAAGGCCGAG TACGCGGTCA AGCTCGCTGC CGCGTGCAGT CAGCATAATG ACGGGGGTCT GGTGTGTCTG GCGAAGCTCT TTCAACGTAT CAATACCGTT
      GTTCCGGCTC ATGCGCCAGT TCGAGCGACG GCGCACGTCA GTCGTATTAC TGCCCCCAGA CCACACAGAC CGCTTCGAGA AAGTTGCATA GTTATGGCAA

                                      ClaI
3401  TTTCTTCGGC ATCATGACGT CGAGCAAAAG TAAATCGATG CTGTCGTCAA GGAGACTCAG CGCCTGCTCG CCATCATGGG CAACCAGGAC GTTGAAACCT
      AAAGAAGCCG TAGTACTGCA GCTCGTTTTC ATTTAGCTAC GACAGCAGTT CCTCTGAGTC GCGGACGAGC GGTAGTACCC GTTGGTCCTG CAACTTTGGA

3501  TCCATGTCGA GCAACTCTTT TAAAAGAGAT GTGAGCTCTC GGTCATCATC AACTAACAGG ATTTTATTCA TTGTTTAAAT ACCTCCGAGG CAGAAATTAC
      AGGTACAGCT CGTTGAGAAA ATTTTCTCTA CACTCGAGAG CCAGTAGTAG TTGATTGTCC TAAAATAAGT AACAAATTTA TGGAGGCTCC GTCTTTAATG

3601  GACATCAAGG CTATCTAATC CATGACTTTA CGTTGTTTTA CACCCCCTGA CGCATGTTTG CAGCCTGAAT CGTAGACTGT CTCTCGTTGA ATCGCGACAC
      CTGTAGTTCC GATAGATTAG GTACTGAAAT GCAACAAAAT GTGGGGGACT GCGTACAAAC GTCGGACTTA GCATCTGACA GAGAGCAACT TAGCGCTGTG

3701  GAAAGATTTT GGGAGCAAGT GATGCGCAAA GTTACCGCTG CCGTCATGGC CTCAACGCTG GCCTTCAGTG CGTTTAGCCA GGCTGCTGTA GCTATCATCA
      CTTTCTAAAA CCCTCGTTCA CTACGCGTTT CAATGGCGAC GGCAGTACCG GAGTTGCGAC CGGAAGTCAC GCAAATCGGT CCGACGACAT CGATAGTAGT
                                                                          Sen R5 100.0%

3801  GCGATAACGG TTCCTCAGCA GAGGGCGCAA CGCAGCACAG CAGCCAAAGC CATATGTTTG ACGGCATAAG TTTAACCGAA CATCAGCGTC AACAGATGCG
      CGCTATTGCC AAGGAGTCGT CTCCCGCGTT GCGTCGTGTC GTCGGTTTCG GTATACAAAC TGCCGTATTC AAATTGGCTT GTAGTCGCAG TTGTCTACGC

3901  AGATCTGATG CAGAGGGCGA GACACGACCA GCCCCCTGTT AATGTTAGCG AAATGGAGAC AATGCATCGC CTTGTCACCG CAGAAAATTT TGACGAAAGC
      TCTAGACTAC GTCTCCCGCT CTGTGCTGGT CGGGGGACAA TTACAATCGC TTTACCTCTG TTACGTAGCG GAACAGTGGC GTCTTTTAAA ACTGCTTTCG

                                                                                                          AvaI
4001  GCTGTACGCG CTCAGGCCGA AAAAATGGCG CAGGAACAGG TTGCCCGTCA GGTAGAGATG GCGAAGGTTC GCAACCAGAT GTTCCACCTG CTAACGCCCG
      CGACATGCGC GAGTCCGGCT TTTTTACCGC GTCCTTGTCC AACGGGCAGT CCATCTCTAC CGCTTCCAAG CGTTGGTCTA CAAGGTGGAC GATTGCGGGC

      AvaI
4101  AGCAGCAAGC GGTTTTGAAT ACCAAACATC AGCAACGTAT GAATCAGTTG CGTGAGGTTG CACGGATGCA GCGAAGCTCA GATATGACGC TTTTCAGTAG
      TCGTCGTTCG CCAAAACTTA TGGTTTGTAG TCGTTGCATA CTTAGTCAAC GCACTCCAAC GTGCCTACGT CGCTTCGAGT CTATACTGCG AAAAGTCATC

                                           Sen F6 100.0%
4201  CAATAGCAGT ACCCGTAGTA ACCAGTAAAC CCTGTTTTCC TTGCCATAGA CACCATCCCT GTCTTCCCCC ACATGATGTG GGGGTTTTTT TTGCCCATTG
      GTTATCGTCA TGGGCATCAT TGGTCATTTG GGACAAAAGG AACGGTATCT GTGGTAGGGA CAGAAGGGGG TGTACTACAC CCCCAAAAAA AACGGGTAAC
                                           Sen R6  100.0%

                                                                   NcoI
4301  TTCAGGTTTG TTAACGCTTT ATTCATCCTT TCACTCCACC ACCATCCGTT ATACTAGCGC CATGGCATGA CAGGAGTGTT TATGAATCAA TCCTATGGAC
      AAGTCCAAAC AATTGCGAAA TAAGTAGGAA AGTGAGGTGG TGGTAGGCAA TATGATCGCG GTACCGTACT GTCCTCACAA ATACTTAGTT AGGATACCTG

                                                                                         Sen F3 70.0%
4401  GGCTGGTAAG CCGGGCCGCA ATAGCCGCGA CGGTGATGGC GTCGTGTTTG CTGATCATTA AAATTTTCGC GTGGTGGTAC ACCGGGTCGG TCAGTATTCT
      CCGACCATTC GGCCCGGCGT TATCGGCGCT GCCACTACCG CAGCACAAAC GACTAGTAAT TTTAAAAGCG CACCACCATG TGGCCCAGCC AGTCATAAGA

4501  GGCGGCACTG GTGGATTCAC TGGTGGATAT TGCCGCCTCG CTGACCAACC TGCTGGTGGT GCGCTACTCG CTGCAACCGG CGGATGAAGA GCATACGTTT
      CCGCCGTGAC CACCTAAGTG ACCACCTATA ACGGCGGAGC GACTGGTTGG ACGACCACCA CGCGATGAGC GACGTTGGCC GCCTACTTCT CGTATGCAAA

4601  GGCCACGGCA AAGCGGAATC GCTGGCCGCG CTGGCACAAA GCATGTTTAT TTCGGGCTCT GCGCTTTTCC TGTTTTTGAC CGGTATTCAG CATCTTATTT
      CCGGTGCCGT TTCGCCTTAG CGACCGGCGC GACCGTGTTT CGTACAAATA AAGCCCGAGA CGCGAAAAGG ACAAAAACTG GCCATAAGTC GTAGAATAAA

                                   Fe Efflux F 75.0%
4701  CGCCTTCTCC GATGAACGAT CCGGGCGTTG GCGTGGTCGT AACGGTAGTT GCACTTATAA GCACACTTGT TCTTGTAACG TTCCAGCGCT GGGTTGTACG
      GCGGAAGAGG CTACTTGCTA GGCCCGCAAC CGCACCAGCA TTGCCATCAA CGTGAATATT CGTGTGAACA AGAACATTGC AAGGTCGCGA CCCAACATGC

4801  CAAAACACAA AGCCAGGCTG TACGGGCCGA TATGCTTCAT TATCAGTCTG ATGTTATGAT GAATGGGGCT ATTCTTATTG CGCTTGGTCT GGCCTGGTAT
      GTTTTGTGTT TCGGTCCGAC ATGCCCGGCT ATACGAAGTA ATAGTCAGAC TACAATACTA CTTACCCCGA TAAGAATAAC GCGAACCAGA CCGGACCATA
```

CDF F3 100.0

4901 GGCTGGCATC GGGCCGATGC TTTGTTTGCG TTAGGGATAG GGATCTATAT TTTATACAGC GCCTTGCGGA TGGGGTATGA AGCGGTACAG TCGCTTCTTG
CCGACCGTAG CCCGGCTACG AAACAAACGC AATCCCTATC CCTAGATATA AAATATGTCG CGGAACGCCT ACCCCATACT TCGCCATGTC AGCGAAGAAC

CDF F3 100.0%

5001 ACCGTGCGCT TCCGGACGCA GAACGTGATG AAATTTATGC CATCGTGACC AACTGGCCTG GCGTCAGTGG TGCTCACGAT CTTCGTACGC GGCAGTCAGG
TGGCACGCGA AGGCCTGCGT CTTGCACTAC TTTAAATACG GTAGCACTGG TTGACCGGAC CGCAGTCACC ACGAGTGCTA GAAGCATGCG CCGTCAGTCC

5101 GGCCGACCCG CTTTATTCAG ATTCATTTGG AAATGGAAGA CAACCTGCCA CTGGGTTCAG GCGCATATGG TCGCTGAACA GGTGGAGCAG GCGATTTTGC
CCGGCTGGGC GAAATAAGTC TAAGTAAACC TTTACCTTCT GTTGGACGGT GACCCAAGTC CGCGTATACC AGCGACTTGT CCACCTCGTC CGCTAAAACG

BamHI

5201 AGCGTTTCCC TGGGTCAGAC GTCATTATTC ATCAGGATCC GTGCTCTGTC GTACCCAGGG CGTTTTGAGC CTTCGTAATT CGTGGTAAAA AAGTGAGCAA
TCGCAAAGGG ACCCAGTCTG CAGTAATAAG TAGTCCTAGG CACGAGACAG CATGGGTCCC GCAAAACTCG GAAGCATTAA GCACCATTTT TTCACTCGTT

5301 GAGCAGCATT TTGTGGATAA ATTACCACCA TATGGCCTGA CCTGAATCAA TTCAGCTGGA AGGGATTGAT ATACTATTTG CAGTATTCGA CGGTTGAACA
CTCGTCGTAA AACACCTATT TAATGGTGGT ATACCGGACT GGACTTAGTT AAGTCGACCT TCCCTAACTA TATGATAAAC GTCATAAGCT GCCAACTTGT

5401 GTTTCTTCCG GCAGCAGATT TCAATTTTGC ATTCCTAAGT TCAGAGGTAG TCATGATTAA GAAAATCGGT GTGTTGACAA GCGGCGGTGA TGCGCCGGGC
CAAAGAAGGC CGTCGTCTAA AGTTAAAACG TAAGGATTCA AGTCTCCATC AGTACTAATT CTTTTAGCCA CACAACTGTT CGCCGCCACT ACGCGGCCCG

5501 ATGAACGCGG CAATCCGTGG GGTTGTCCGT GCAGCGCTGA CGGAAGGTCT GGAAGTTTTT GGTATCTATG ACGGTTACCT GGGTCTGTAT GAAGACCGCA
TACTTGCGCC GTTAGGCACC CCAACAGGCA CGTCGCGACT GCCTTCCAGA CCTTCAAAAA CCATAGATAC TGCCAATGGA CCCAGACATA CTTCTGGCGT

EcoRI

5601 TGGTTCAGCT CGACCGCTAC AGCGTGTCTG ACATGATCAA CCGTGGCGGT ACTTTCCTCG GTTCTGCACG CTTCCCGGAA TTCCGTGACG AACACATCCG
ACCAAGTCGA GCTGGCGATG TCGCACAGAC TGTACTAGTT GGCACCGCCA TGAAAGGAGC CAAGACGTGC GAAGGGCCTT AAGGCACTGC TTGTGTAGGC

5701 TGAAGTGGCT ATCGAAAACA TGAAAAAACG GGGTCTGGAT GCGCTGGTGG TTATCGGCGG CGACGGCTCC TACATGGGTG CAAAACGTCT GACTGAAATG
ACTTCACCGA TAGCTTTTGT ACTTTTTTGC CCCAGACCTA CGCGACCACC AATAGCCGCC GCTGCCGAGG ATGTACCCAC GTTTTGCAGA CTGACTTTAC

5801 GGATTCCCGT GCATCGGCCT GCCTGGCACC ATCGACAATG ACATCAAAGG CACCGACTAC ACCATCGGTT ACTTCACCGC GCTGGGTACC GTTGTGGAAG
CCTAAGGGCA CGTAGCCGGA CGGACCGTGG TAGCTGTTAC TGTAGTTTCC GTGGCTGATG TGGTAGCCAA TGAAGTGGCG CGACCCATGG CAACACCTTC

5901 CGATTGACCG CCTGCGTGAC ACCTCTTCCT CTCACCAGCG TATTTCTATT GTTGAAGTGA TGGGCCGTTA CTGCGGTGAC CTGACTCTGG CGGCGGCAAT
GCTAACTGGC GGACGCACTG TGGAGAAGGA GAGTGGTCGC ATAAAGATAA CAACTTCACT ACCCGGCAAT GACGCCACTG GACTGAGACC GCCGCCGTTA

EcoRI

6001 TGCCGGTGGC TGTGAGTTCG TGGTGGTGCC GGAAGTGGAA TTCAGCCGTG AAGATCTGGT CGCTGAAATC AAAGCCGGTA TCGCGAAAGG TAAGAAACAC
ACGGCCACCG ACACTCAAGC ACCACCACGG CCTTCACCTT AAGTCGGCAC TTCTAGACCA GCGACTTTAG TTTCGGCCAT AGCGCTTTCC ATTCTTTGTG

6101 GCCATCGTTG CTATCACCGA GCACATCTGT GACGTTGACG AGCTGGCGAA GTACATCGAA ACCGAAACCA AACGCGAAAC CCGCGCGACC GTTCTGGGTC
CGGTAGCAAC GATAGTGGCT CGTGTAGACA CTGCAACTGC TCGACCGCTT CATGTAGCTT TGGCTTTGGT TTGCGCTTTG GGCGCGCTGG CAAGACCCAG

6201 ACATTCAGCG TGGTGGTTCC CCAGGCCCGT ACGACCGTAT CCTGGCGTCC CGCATGGGCG CGTACGCGAT CGAGCTACTG CTTCAGGGCC ACGGCGGCCG
TGTAAGTCGC ACCACCAAGG GGTCCGGGCA TGCTGGCATA GGACCGCAGG GCGTACCCGC GCATGCGCTA GCTCGATGAC GAAGTCCCGG TGCCGCCGGC

ClaI

6301 CTGCGTCGGT ATTCAGAACG AGAAACTGGT TCACCATGAC ATCATCGATG CCATTGAAAA CATGAAGCGT CCGTTCAAAG GTGACTGGCT GGACTGCGCG
GACGCAGCCA TAAGTCTTGC TCTTTGACCA AGTGGTACTG TAGTAGCTAC GGTAACTTTT GTACTTCGCA GGCAAGTTTC CACTGACCGA CCTGACGCGC

6401 AAAAAACTGT ACTGATTGGC TCTGTTGCCC GGTGGCGGCA AGGTAAATGC AAAACGGTAA CCCTTGGGTT ACCGTTTTTT ATGGTTGCTC CCTCTCCCGT
TTTTTTGACA TGACTAACCG AGACAACGGG CCACCGCCGT TCCATTTACG TTTTGCCATT GGGAACCCAA TGGCAAAAAA TACCAACGAG GGAGAGGGCA

6501 GGGAGAGGGT TGGGGTGAGG GCATCAGGCC GCTCATTCAC TTTTTTTTAT TCCTCCACGT TATAGCCAAT CTTTTTTTAT TCTTTAATCA TCGGTTAGCT
CCCTCTCCCA ACCCCACTCC CGTAGTCCGG CGAGTAAGTG AAAAAAAATA AGGAGGTGCA ATATCGGTTA GAAAAAAATA AGAAATTAGT AGCCAATCGA

6601 TTCTGGCACG CTGCATTCAT CAAAACACTA CACAAGAGAG CTGGGCGATG AATAAATGGG GCGTGGGGTT AACATTATTG CTGGCATCGA CCAGCGTTCT
AAGACCGTGC GACGTAAGTA GTTTTGTGAT GTGTTCTCTC GACCCGCTAC TTATTTACCC CGCACCCCAA TTGTAATAAC GACCGTAGCT GGTCGCAAGA

HindIII

6701 GGCAAAGGAC ATTCAGTTAC TGAACGTGTC GTACGATCCG ACGCGTGAAC TATACGACCA GTACAACAAA GCTTTCGCGG CGCACTGGAA ACAGGAAACC
CCGTTTCCTG TAAGTCAATG ACTTGCACAG CATGCTAGGC TGCGCACTTG ATATGCTGGT CATGTTGTTT CGAAAGCGCC GCGTGACCTT TGTCCTTTGG

6801 GGCGATAACG TGGTGGTTCG CCAGTCTCAC GGTGGTTCTG GCAAGCAGGC GACCTCCGTC ATCAACGGTA TTGAAGCCGA CGTGGTGACC CTGGCGCTGG
CCGCTATTGC ACCACCAAGC GGTCAGAGTG CCACCAAGAC CGTTCGTCCG CTGGAGGCAG TAGTTGCCAT AACTTCGGCT GCACCACTGG GACCGCGACC

6901 CCTACGATGT TGACGCTATC GCCGAGCGTG GCCGTATCGA CAAAAACTGG ATCAA
GGATGCTACA ACTGCGATAG CGGCTCGCAC CGGCATAGCT GTTTTTGACC TAGTT

Figure 11. **Assembed L31 Sequences.** The words in red represent the location of the different primers used in sequencing and the words in blue represent the sites for the different restriction endonuclease. The histidine kinase sensor is located between base pairs 1,503 and 2,876 wheras the EZ-Tn5 Transposon inserted between the base pairs 2,017 and 2,025.

5.6 Southern Blot

A Southern Blot was performed to show that the transposon inserted itself into the histidine kinase gene only and not into any other gene of the L31 mutant. *Eco*R I digested wild type and L31 genomic DNA were separated on a 0.8% agaraose gel (Figure 12) and blotted onto a positively charged Biodyne B Nylon membrane. Then, it was hybridized with a biotin labeled segment of the kanamycin resistance gene from the transposome. A 5.6 kb, *Eco*R I fragment contains histidine kinase gene (Fig 10). Since the transposon is 2 kb and does not contain an *Eco*R I site, the transposon-interrupted histidine kinase gene of L31 should be found in a 7.6 kb *Eco*R I fragment. In lane 2 of figure 13, the probe hybridized to the the positive control at an approximate size of 550 bp. In addition, there was no signal in the negative control in lane 3 which contained digested wild type genomic DNA. This strain lacked an EZ-Tn5^{TM} Transposon insert. As expected, the probe hybridized to a band of approximately 8,000 bp and no other bands in lane 4. Thus, the transposon only inserted itself at one site in the L31 mutant.

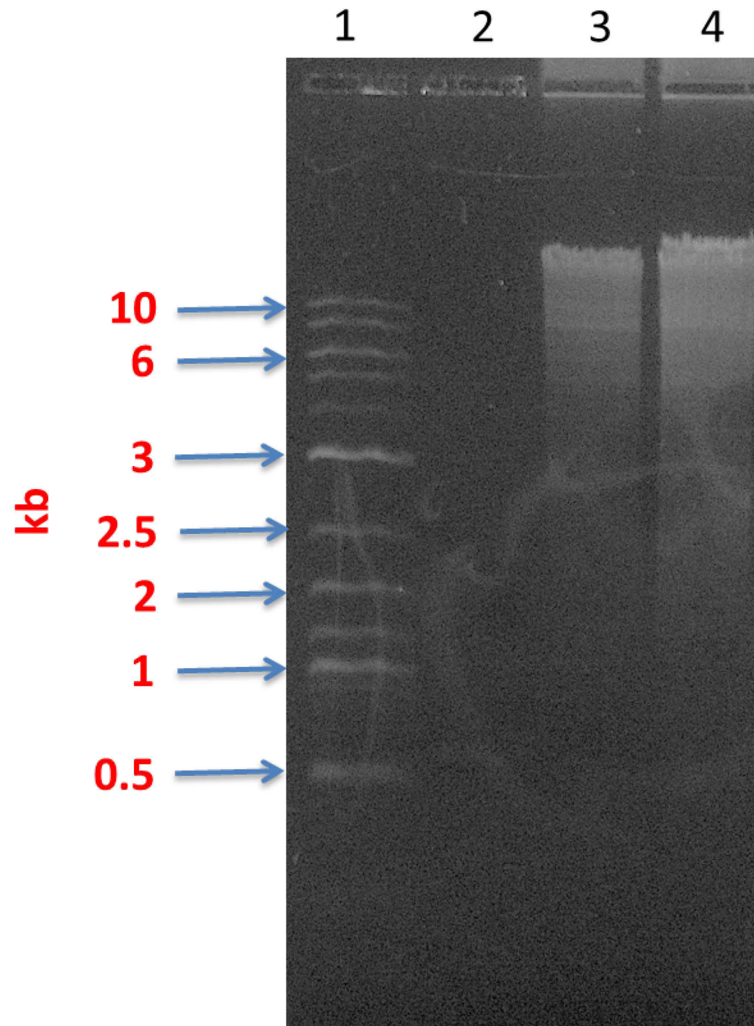

Figure 12. Southern Blot Gel. Lane 1 – Biotin labeled ladder; Lane 2 – PCR product; Lane 3 – *Eco*R I genomic digestion of wild type, *Enterobacter* sp. YSU; Lane 4 – *Eco*R I genomic digestion of the mutant, L31.

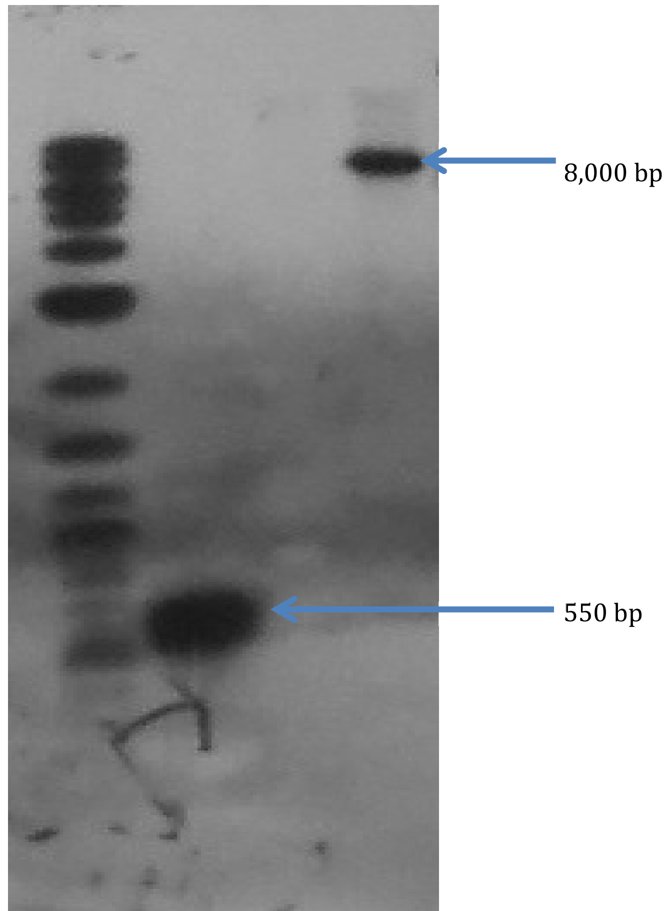

Figure 13. Southern Blot. Lane 1 – Biotin labeled ladder; Lane 2 – labeled PCR product, ~550 bp; Lane 3 – *Enterobacter* sp. YSU, wild type did not give any signal since it lacks the EZ-Tn5^{TM} Transposon insert; Lane 4 – L31 mutant indicating a single copy insertion of the EZ-Tn5^{TM} Transposon that is approximately 8,000 bp

5.7 Multiple Sequence Alignement

Multiple sequence alignment (MSA) was performed to compare the protein sequence of L31 histidine kinase and its homologs. Seven homologs were selected from a blastp search of a translated L31 histidine kinase DNA sequence. Alignment was performed using clustalX and the ouput analysed using the GeneDoc format as shown in figure 14. The location of the conservered domains were determined using PROSITE protein domain database. The output of PROSITE scan results indicated two main domains. The first domain, HAMP (histidine kinases, adenylyl cyclases, methyl-accepting chemo- taxis proteins, and phosphatases), is located between amino acid residue 184 and 237 and the other domain, histidine kinase, is located between amino acid residue 245 and 455 in the MSA (Figure 14). This domain organization is characteristic of class I sensor kinases (59). In addition to these two domains, class I sensor kinases have a periplasmic sensory domain and two transmebrane domains (TM1 and TM2) located on either side of the sensory domain. The sensory domain in the MSA (Figure 14) is located between amino acid residues 30 and 163 whereas the TM1 between 8 and 29 and TM2 between 164 and 183. The histidine kinase domain has sequence motifs that are conserved in members of the class I sensor kinase (39, 45, 62). These consevered sequence motifs are represented by H, N, G1, F and G2 boxes (44). The location of these boxes in the MSA in figure 14 was identified from studies carried out by Kim *et al*, 1989. The H box is represented by amino acid residues between 248 and 254, the N box between 356 and 361, the G1 box between 375 and 382, the F box between 359 and 403 and the G2 box between 416 and 421 amino acid residues. The EZ-Tn5 Transposon inserted between the amino acid residue 285 and 287, circled with a red box.

```
                       *        20          *        40
L31_His_Ki  : MIGSLTARIFAIFWLTLALVLMLVLMLPKLDSRQMTELLDS :  41
gi|4792709  : MIGSLTARIFAIFWLTLALVLMLVLMLPKLDSRQMTELLDS :  41
gi|3659727  : MIGSLTARIFAIFWLTLALVLMLVLMLPKLDSRQMTELLDS :  41
gi|4017657  : MIGSLTARIFAIFWLTLALVLMLVLMLPKLDSRQMTELLDS :  41
gi|4859039  : MIGSLTARIFAIFWLTLALGLMLVLMLPKLDSRQMTELLDS :  41
gi|4857290  : MIGSLTARIFAIFWLTLALVLMLVLMLPKLDSRQMTELLDS :  41
gi|4465025  : MIGSLTARIFAIFWLTLALVLMLVLMLPKLDSRQMTELLDS :  41
gi|3752582  : MIGSLTARIFAIFWLTLALVLMLVLMLPKLDSRQMTELLES :  41
              MIGSLTARIFAIFWLTLALvLMLVLMLPKLDSRQMTELLdS
```

Transmembrane domain (TM1)

```
                      *        60          *        80
L31_His_Ki  : EQRQGVMIEQHVEAELANDPPNDLMWWRRLFRAIDKWAPPG :  82
gi|4792709  : EQRQGVMIEQHVEAELANDPPNDLMWWRRLFRAIDKWAPPG :  82
gi|3659727  : EQRQGVMIEQHVEAELANDPPNDLMWWRRLFRAIDKWAPPG :  82
gi|4017657  : EQRQGVMIEQHVEAELANDPPNDLMWWRRLFRAIDKWAPPG :  82
gi|4859039  : EQRQGLMIEQHVEAELANDPPNDLMWWRRLFRAIDKWAPPG :  82
gi|4857290  : EQRQGLMIEQHVEAELANDPPNDLMWWRRLFRAIDKWAPPG :  82
gi|4465025  : EQRQGLMIEQHVEAELANDPPNDLMWWRRLFRAIDKWAPPG :  82
gi|3752582  : EQRQGIMIEQHVEAELANDPPNDLMWWRRLFRAIDKWAPPG :  82
              EQRQG6MIEQHVEAELANDPPNDLMWWRRLFRAIDKWAPPG

                     *        100         *        120
L31_His_Ki  : QRLLLVTSEGRVIGADRNEMQIIRNFIGQADNADHPQKKKY : 123
gi|4792709  : QRLLLVTSEGRVIGADRNEMQIIRNFIGQADNADHPQKKKY : 123
gi|3659727  : QRLLLVTSEGRVIGADRNEMQIIRNFIGQADNADHPQKKKY : 123
gi|4017657  : QRLLLVTSEGRVIGADRNEMQIIRNFIGQADNADHPQKKKY : 123
gi|4859039  : QRLLLVTTEGRVIGAERSEMQIIRNFIGQADNADHPQKKKY : 123
gi|4857290  : QRLLLVTTEGRVIGAERSEMQIIRNFIGQADNADHPQKKKY : 123
gi|4465025  : QRLLLVTSEGRVIGAERSEMQIIRNFIGQADNADHPQKKKY : 123
gi|3752582  : QRLLLVTSEGRVIGAERNEMQIIRNFIGQADNADHPQKKRY : 123
              QRLLLVT3EGRVIGA R EMQIIRNFIGQADNADHPQKK4Y

                    *        140         *        160
L31_His_Ki  : GRVEMVGPFSVRDGEDNYQLYLIRPASSSQSDFINLLFDRP : 164
gi|4792709  : GRVEMVGPFSVRDGEDNYQLYLIRPASSSQSDFINLLFDRP : 164
gi|3659727  : GRVEMVGPFSVRDGEDNYQLYLIRPASNSQSDFINLLFDRP : 164
gi|4017657  : GRVEMVGPFSVRDGEDNYQLYLIRPASSSQSDFINLLFDRP : 164
gi|4859039  : GRVELVGPFSVRDGEDNYQLYLIRPASSSQSDFINLLFDRP : 164
gi|4857290  : GRVELVGPFSVRDGEDNYQLYLIRPASSSQSDFINLLFDRP : 164
gi|4465025  : GRVEMVGPFSVRDGEDNYQLYLIRPASSSQSDFINLLFDRP : 164
gi|3752582  : GRLEMVGPFSVRDGEDNYQLYLIRPANTSQSDFINLLFDRP : 164
              GR6E6VGPFSVRDGEDNYQLYLIRPAs SQSDFINLLFDRP
```

```
                    *        180         *        200
L31_His_Ki : LLLLIVTMLVSSPLLLWLAWSLAKPARKLKNAADEVAQGNL : 205
gi|4792709 : LLLLIVTMLVSSPLLLWLAWSLAKPARKLKNAADEVAQGNL : 205
gi|3659727 : LLLLIVTMLVSSPLLLWLAWSLAKPARKLKNAADEVAQGNL : 205
gi|4017657 : LLLLIVTMLVSSPLLLWLAWSLAKPARKLKNAADEVAQGNL : 205
gi|4859039 : LLLLIVTMLVSTPLLLWLAWSLAKPARKLKNAADEVAQGNL : 205
gi|4857290 : LLLLIVTMLVSTPLLLWLAWSLAKPARKLKNAADEVAQGNL : 205
gi|4465025 : LLLLIVTMLVSSPLLLWLAWSLAKPARKLKNAADEVAQGNL : 205
gi|3752582 : LLLLIVTMLVSAPLLLWLAWSLAKPARKLKNAADEVAQGNL : 205
             LLLLIVTMLVS PLLLWLAWSLAKPARKLKNAADEVAQGNL
```

Transmembrane domain (TM 2) HAMP domain

```
                  *        220         *        240
L31_His_Ki : RQHPELESGPQEFLAAGTSFNQMVSALDRMMTAQQRLLSDI : 246
gi|4792709 : RQHPELESGPQEFLAAGTSFNQMVSALDRMMTAQQRLLSDI : 246
gi|3659727 : RQHPELEAGPQEFLAAGTSFNQMVSALDRMMTAQQRLLSDI : 246
gi|4017657 : RQHPELEAGPQEFLAAGTSFNQMVSALDRMMTAQQRLLSDI : 246
gi|4859039 : RQHPELEAGPQEFLAAGASFNQMVTALERMMTSQQRLLSDI : 246
gi|4857290 : RQHPELEAGPQEFLAAGASFNQMVTALERMMTSQQRLLSDI : 246
gi|4465025 : RQHPELEAGPQEFLAAGASFNQMVTALERMMTSQQRLLSDI : 246
gi|3752582 : RQHPELEAGPQEFLAAGTSFNQMVTALERMMSSQHRLLSDI : 246
             RQHPELEaGPQEFLAAG SFNQMV3AL RMM3 QqRLLSDI
```

HAMP domain

```
                   *       260          *        280
L31_His_Ki : SHELRTPLTRLQLGTALLRRRSGESKELERIETEAHRLDSM : 287
gi|4792709 : SHELRTPLTRLQLGTALLRRRSGESKELERIETEAHRLDSM : 287
gi|3659727 : SHELRTPLTRLQLGTALLRRRSGESKELERIETEAHRLDSM : 287
gi|4017657 : SHELRTPLTRLQLGTALLRRRSGESKELERIETEAHRLDSM : 287
gi|4859039 : SHELRTPLTRLQLGTALLRRRSGESKELERIETEAQRLDSM : 287
gi|4857290 : SHELRTPLTRLQLGTALLRRRSGEGKELERIETEAQRLDSM : 287
gi|4465025 : SHELRTPLTRLQLGTALLRRRGGESKELERIETEAQRLDSM : 287
gi|3752582 : SHELRTPLTRLQLGTALLRRRSGESKELERIEMEAHRLDSM : 287
             SHELRTPLTRLQLGTALLRRRsGEsKELERIEtEA RLDSM
```

H box

```
                  *        300         *        320
L31_His_Ki : INDLLVMSRNQQKNALVSETVKANHLWHEVLDNAAFEAEQM : 328
gi|4792709 : INDLLVMSRNQQKNALVSETVKANHLWHEVLDNAAFEAEQM : 328
gi|3659727 : INDLLVMSRNQQKNALVSETVKANHLWHEVLDNAAFEAEQM : 328
gi|4017657 : INDLLVMSRNQQKNALVSETVKANHLWHEVLDNAAFEAEQM : 328
gi|4859039 : INDLLVMSRNQQKNALVSETIKANQLWSEVLDNAAFEAEQM : 328
gi|4857290 : INDLLVMSRNQQKNALVSETIKANQLWSEVLDNAAFEAEQM : 328
gi|4465025 : INDLLVMSRNQQKNALVSETMKANQLWGEVLDNAAFEAEQM : 328
gi|3752582 : INDLLVMSRNQAKTALVSETVKANQLWGEVLDNAAFEAEQM : 328
             INDLLVMSRNQqKnALVSET6KAN LW EVLDNAAFEAEQM
```

```
               *        340          *        360
L31_His_Ki  : GKSFTVNFPPGPWPLYGNPNTLESALENIVRNALRYSHTKI : 369
gi|4792709  : GKSFTVNFPPGPWPLYGNPNTLESALENIVRNALRYSHTKI : 369
gi|3659727  : GKSFTVNFPPGPWPLYGNPNTLESALENIVRNALRYSHTKI : 369
gi|4017657  : GKSFTVNFPPGPWPLYGNPNTLESALENIVRNALRYSHTKI : 369
gi|4859039  : GKSLTVNFPPGPWPLYGNPNALESALENIVRNALRYSHTKI : 369
gi|4857290  : GKSLTVNFPPGPWPLYGNPNALESALENIVRNALRYSHTKI : 369
gi|4465025  : GKSLTVNYPPGPWPLYGNPNALESALENIVRNALRYSHTKI : 369
gi|3752582  : GKSFTVEYPPGPWPLYGNPNALESALENIVRNALRYSHTKI : 369
              GKS TVn5PPGPWPLYGNPN LESALENIVRNALRYSHTKI
                                          N box

              *        380          *        400          *
L31_His_Ki  : EVAFSVDKDGITVIVDDDGPGVSPEDREQIFRPFYRTDEAR : 410
gi|4792709  : EVAFSVDKDGITVIVDDDGPGVSPEDREQIFRPFYRTDEAR : 410
gi|3659727  : EVAFSVDKDGITIIVDDDGPGVSPEDREQIFRPFYRTDEAR : 410
gi|4017657  : EVAFSVDKDGITVIVDDDGPGVSPEDREQIFRPFYRTDEAR : 410
gi|4859039  : EVGFAVDKDGITITVDDDGPGVSPEDREQIFRPFYRTDEAR : 410
gi|4857290  : EVGFAVDKDGITITVDDDGPGVSPEDREQIFRPFYRTDEAR : 410
gi|4465025  : EVCFSVDKDGITITVDDDGPGVSPEDREQIFRPFYRTDEAR : 410
gi|3752582  : SVSFSVDKDGITVNVDDDGPGVSPEDREQIFRPFYRTDEAR : 410
              eV FsVDKDGIT6 VDDDGPGVSPEDREQIFRPFYRTDEAR
                  G1 box                  F box

                 420          *        440          *
L31_His_Ki  : DRESGGTGLGLAIVETAMQQHRGWVKADDSPLGGLRLTLWL : 451
gi|4792709  : DRESGGTGLGLAIVETAMQQHRGWVKADDSPLGGLRLTLWL : 451
gi|3659727  : DRESGGTGLGLAIVETAMQQHRGWVKADDSPLGGLRLTLWL : 451
gi|4017657  : DRESGGTGLGLAIVETAMQQHRGWVKADDSPLGGLRLTLWL : 451
gi|4859039  : DRESGGTGLGLAIVETAIQQHRGWVKAEDSPLGGLRLVIWL : 451
gi|4857290  : DRESGGTGLGLAIVETAIQQHRGWVKAEDSPLGGLRLVIWL : 451
gi|4465025  : DRESGGTGLGLAIVESAMQQHRGWVKADDSPLGGLRLTLWL : 451
gi|3752582  : DRESGGTGLGLAIVEAAIQQHRGWVKADDSPLGGLRLTLWL : 451
              DRESGGTGLGLAIVE A6QQHRGWVKAdDSPLGGLRLt6WL
                 G2 box

L31_His_Ki  : PLYKRS : 457
gi|4792709  : PLYKRS : 457
gi|3659727  : PLYKRS : 457
gi|4017657  : PLYKRS : 457
gi|4859039  : PLYKRS : 457
gi|4857290  : PLYKRS : 457
gi|4465025  : PLYKRT : 457
gi|3752582  : PLYKRT : 457
              PLYKR3
```

Figure 14. Multiple Sequence Alignment. The GeneDoc output showing the protein sequence alignment of L31 hitidine kinase and its homologs – gi 4792709 *Enterobacter cloacae* subsp. *cloacae* NCTC 9394; gi 3659727 *Enterobacter cloacae* EcWSU1; gi 4017657 *Enterobacter cloacae* subsp. *cloacae* ENHKU01; gi 4859039 *Escherichia coli*, gi 4857290 *Escherichia coli*; gi 4465025 *Salmonella enterica*; gi 3752582 *Klebsiella oxytoca* KCTC 1686. The blue lines represent the transmembrane domains (TM1 and TM2), HAMP domain and the conserved motifs represented by the H, N, G1, F and G2 boxes. The protein sequence between TM1 and TM2 represent the sensory domain and the red circle represent the transposon insertion site.

5.8 Phylogenetic Analysis

A phylogenetic tree was used to compare the evolutionary relationship between L31 histidine kinase with other families of histidine kinases in *E. coli* strain K-12 (Figure 15). The protein sequence of 14 different histidine kinases in *E. coli* were downloaded and aligned with Mega5. The alignment output was then used to draw a phylogenetic tree. The tree branched into 5 different groups implying there were 5 different types of histidine kinases. Type 1 and type 2 histidine kinases appeared to be closely related with type 1 consisting of many members that further fall into subtypes, type 1A, 1B and 1C. The L31 histidine kinase falls into type 1A with a close relationship with CpxA compared to other type 1A members, EnvZ and BaeS. Type 3 and 4 also appeared to be closely related. CheA histidine kinase was classified as type 5 since it does not branch to form a group with any of the other type of histidine kinases, although it appears to be closely related to type 1 and 2 members compared to type 3 and 4 members.

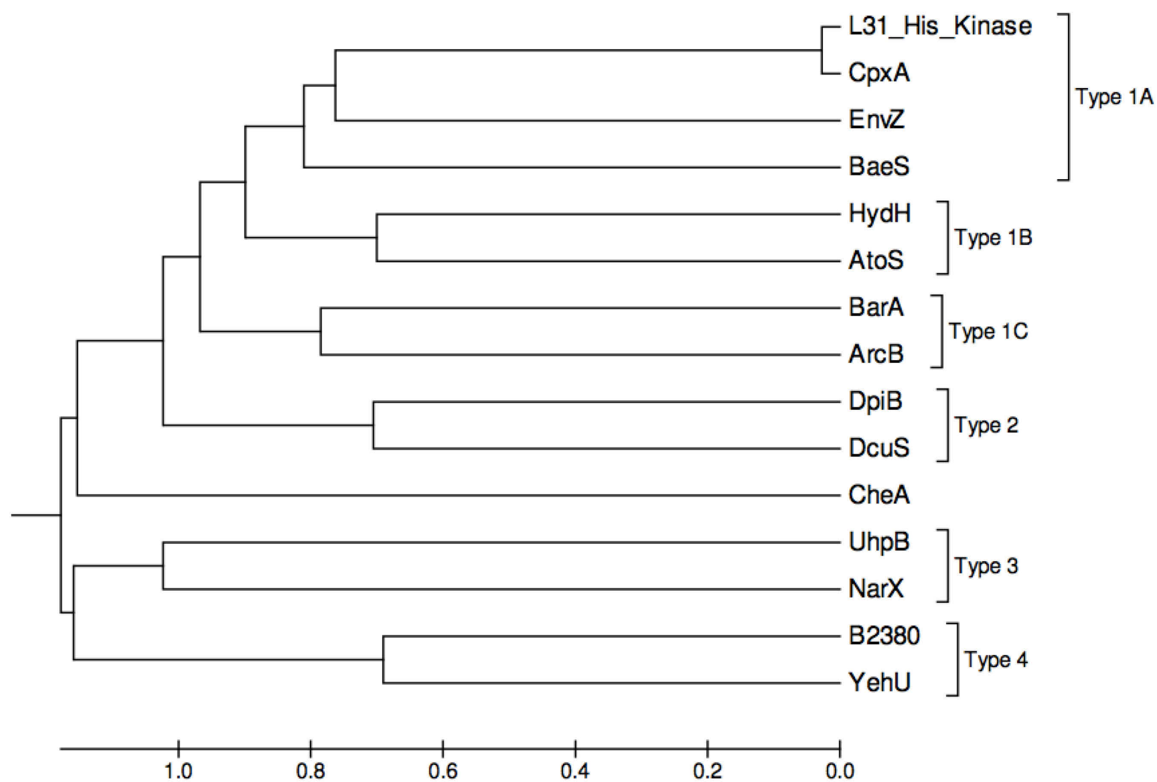

Figure 15. Phylogenetic Analysis. A Phylogenetic tree showing the evolutionary relationship of L31 histidine kinase with other families of histidine kinases in *E. coli* K-12. P0AE82 CpxA; P0AEJ4 EnvZ; P30847 BaeS; Q61U37 HydH; Q060067 AtoS; P0AEC5 BarA; P0AEC3 ArcB, P77510 DpiB; P0AEC8 DcuS; P07363 CheA; P09835 UhpB; P0AFA2 NarX; P0AA93 B2380; POAD14 YehU.

CHAPTER VI: DISCUSSION

The MIC results indicated that the *Enterobacter* sp. YSU mutant, L31 was more sensitive to selenite compared to the wild type. Therefore, the region containing the transposon was sequenced. We expected that the transposon interrupted gene would express a protein involved in efflux, sequestration or detoxification, but instead found a gene that encoded for a protein that appears to regulate one of these mechanisms. From the MSA alignment, it was evident that *cpxA* is the gene product of the L31 histidine kinase illustrated in figure 10. Genes regulated by CpxAR system encode proteins involved in envelope protein folding and in reducing the concentrations of cellular toxic molecules (48). So far, at least 50 genes in 34 operons have been found to be CpxAR-regulated (48). Some CpxAR-regulated genes like *cpxP, syp, ebr, ybaJ, yccA, ycfS, ydeH, yecl, yqjA* and JW1832 are induced in response to copper (48, 52, 53, 54).

Cpx Envelope Stress System

The genes *cpxA* and *cpxR* are part of the Cpx-two component envelope stress system (39, 40, 41, 42, 43). There are several stimuli that cause protein misfolding, hence inducing the Cpx system (39, 40, 41). The metals zinc and copper are known inducers of the Cpx system (39). CpxA functions by sensing external stimuli and phosphorylates itself at a specific conservered histidine residue (55). CpxR functions as the response regulator that gets phosphorylated on an aspartate residue (39, 40, 60). Phosphorylation of the response regulator results in transcription of the Cpx regulon (40) that contain genes involved in protein folding and degradation such as the periplasmic endoprotease, DegP (42, 55, 56). DegP is important for resistance to reactive oxgen species that cause protein damage (57). Selenite toxicity is attributed to oxygen species (58), therefore

interuption of CpxA in the L31 mutant could have effected the activity of DegP since expression of DegP is dependent of the Cpx system, although further tests needs to be conducted to supported this statement. In addition to the two-component system is a periplasmic stress response protein, CpxP (40). CpxP functions by controling the activities of CpxA and also fuctions as a periplasmic adaptor protein that helps in the degradation of misfolded proteins (40, 43, 56). When CpxP binds to CpxA, it results to dephosphorylation of CpxA, preventing the activation of CpxR (39). When a stress molecule is present CpxP binds to the molecule, keeping CpxP from binding to CpxA. This allows CpxA to activate CpxR and induce a response to the stress molecule (selenite). The response can also be initiated in the absence of CpxP but it is not clear how.

Characterization of the Cpx regulon by Price *et al,* 2009 indicated that most of the Cpx- regulated genes are those induced due to misfolded proteins and those regulated due to copper stress. Although the function of most of these identified genes are unknown (48). A recent microarray study on Cpx-regulated genes identified more genes in the Cpx regulon, some of the identified genes affect antibiotic resistance (50) and some are induced in response to zinc (49). Yamamoto and Ishihama, 2006 compared the sensitivity of *E .coli* wild type strain and its cpxAR null mutant to metal salts of lithium, sodium, magnesium, potassium, calcium, chromate, iron cobalt, nickel, copper, zinc, rubidium, strontium, caesium, barium, lead and silver. Results indicated that the mutant was highly sensitive to copper followed by zinc but sensitivity to other metals was the same as the wild type (49).

Induction of Cpx-regulated genes due to copper toxicity

Twelve identified genes in the CpxAR regulon are induced in response to copper (49, 52). The 12 genes, *cpxP, syp, ebr, ybaJ, yccA, ycfS, ydeH, yecl, yqjA* and JW1832 were identified using a microarray (52). Studies conducted on *cpxA** (gain-of-function mutant) mutant with autokinase activity but defective in phosphatase activity showed increased sensitivity to copper (51, 56). These results were later supported by research in *E. coli* carried out by Yamamoto and Ishihama, 2005 on the involvement of CpxAR in response to elevated levels of external copper (52).

The CpxAR regulated gene, CpxP share high protein sequence homology to the metal binding protein, ZraP, and the metal sensor, CnrX. ZraP is induced in a TCS in response to zinc or lead whereas CnrX function to bind cobalt, nickel and copper (40). This structural similarity could imply a possible common function of these three genes in periplasmic binding (40).

Multiple Sequence Alignment Analysis

The EZ-Tn5 Transposon insertion site was between the leucine and isoleucine residues located at position 284 and 288 respectively in the MSA alignment in figure 14. The insertion location at an X-region located between the conserved H box and N box. Studies conducted by Hsing *et al*, 1998 on the phosphatase and kinase activity of *envZ* gene in *E. coli* indicated that mutation of the X region located 40 amino acid residues downstream of the phosphorylated histidine resulted in a weak kinase activity and a decrease phosphatase activity. Mutation in this region also resulted in a significant conformational change in the the EnvZ sensor kinase (21). Since EnvZ belongs to the

same class as the L31 histidine kinase, mutation at this region therefore affected the kinase activity of the L31 mutant. Tyrosine residue located in this X-region is a common mutation site resulting in substitution of 3 amino acid residues (21). In *E. coli* EnvZ , the tyrosine residue is located at position 287 whereas in the MSA in figure 14, the tyrosine residue in this region is located close to the transposon insertion site at position 279. Another amino acid considered important in this region is the Arg^{289} in *E.coli* EnvZ (21). This amino acid residue is located at position 298 in the MSA in figure 14. Mutation at this region results in a weak phosphatase activity (21). Therefore, the X region was considered important in the interaction between the sensor and the response regulator (21).

The MIC growth curves indicated that the L31 mutant was sensitive to selenite in solid LB and M-9 minimal medium but in liquid cultures, it was only sensitive in LB medium but not in M-9 minimal medium. This suggests that there are multiple mechanisms for resistance and they grow differently on solid and liquid medium. It is not possible to determine the difference at this time. The mutation in the gene for the sensor protein probably just caused the L31 mutant to be inhibited by selenite rather than be killed by it.

Phylogenetic analysis indicated that the L31 histidine kinase was closely related to CpxA, EnvZ and BaeS. BaeS and CpxA are similar in terms of function as they are both envelope stress response proteins (63). EnvZ, on the other hand, is a sensor kinase involved in osmoregulation (21). An alignment of the different groups indicated that type 1 and type 2 members have kinase domain with N, G1, F and G2 conserved sequence motifs. The N1 of the N box in type 3 is replaced with a glycine residue whereas for type

4 its replaced with a proline residue. The other difference is that type 3 and 4 member lack the F box and the G2 box is trancated (45).

Phylogenetic comparison of L31 histidne kinase and other hisidine kinases may outline difference that could further help understand the role sensor kinases play in allowing bacteria adapt to various environmental conditions.

Future work

Gene for a putative cation diffusion facilitator (CDF) and a putative sulfur binding proteins (SBP) were found to flank the sensor kinase. These proteins may function in metal resistance mechanisms. Therefore, RT-PCR or qPCR can be conducted to determine if they are induced in the presence of selenite. If these gene are induced in the presence of selenite, inverted membrane experiments can be conducted to see if the CDF can pump selenite into vesicles. Selenite containing vesicles can be digested and the selenite content can be measured by inductively coupled plasma mass spectrometry (ICP-MS).

The sulfur binding protein and CpxP can also be purified to see if they are selenite binding proteins. These proteins can be mixed with different concentrations of selenite to allow for binding. Then, protein/selenite complexes can be digested and selenite content can be measured by ICP-MS.

Conclusion

Bacteria have developed different resistance mechanisms to survive in heavy metal contaminated environments. Transposon mutagenesis was used to identify genes involved in selenite resistance. The EZ-Tn5 Transposon was randomly inserted into the genome of *Enterobacter sp.* YSU and a selenite sensitive mutant, L31, was generated. Sequencing of the region interupted by the transposon indicated a signal transduction histidine kinase. There are different kinds of histidine kinases but from a multiple sequence alignment, it was evident that the histidine kinase, *cpxA* was the gene product of the L31 histidine kinase. So far, the metals copper and zinc are known inducers of the Cpx system, although these two metals were not tested. Sequencing of the regions flanking the sensor kinase identified two gene that are part of the Cpx system, *cpxR* and *cpxP*. Sulfur binding protein and cation diffusion facilitator were also sequenced although further experiments such RT-PCR or qPCR will be conducted to determine if they are induced in the presence of selenite.

CHAPTER VII: REFERENCES

1. Holmes A. Vinayak A, Benton C, Esbenshade A, Heiselman C, Frankland D, Kulkarni S, Kurtanich A, Caguiat J: **Comparison of two multimetal resistant bacterial strains: *Enterobacter* sp. YSU and *Stenotrophomonas maltophilia* ORO2:** Curr Microbiol. 2009 Nov;59(5):526-31. Epub 2009 Aug.

2. Andreazza R, Okeke BC, Pieniz S, Camargo FA: **Characterization of copper-resistant rhizosphere bacteria from *Avena sativa* and *Plantago lanceolata* for copper bioreduction and biosorption:** Biol Trace Elem Res. 2012 Apr;146(1):107-15. Epub 2011 Oct 15.

3. Argüello JM, González-Guerrero M, Raimunda D: **Bacterial transition metal P(1B)-ATPases: transport mechanism and roles in virulence:** Biochemistry. 2011 Nov 22;50(46):9940-9. Epub 2011 Oct 31.

4. Ballatori N: **Transport of toxic metals by molecular mimicry:** Environ Health Perspect. 2002 Oct;110 Suppl 5:689-94.

5. Brul S, Coote P: **Preservative agents in foods. Mode of action and microbial resistance mechanisms:** Int J Food Microbiol. 1999 Sep 15;50(1-2):1-17.

6. Cantagrel S, Mereghetti L, Marchand S, Van der Mee N, Besnier J M, Laugier J, Quentin R: **Spread of *Stenotrophomonas maltophilia* colonization in a pediatric intensive care unit detected by monitoring tracheal bacterial carriage and molecular typing:** ISSN: 1198-743X, DOI: 10.1046/j.1469-0691.2003.00785.x

7. Haferburg G, Kothe E: **Microbes and metals: interactions in the environment.** J Basic Microbiol. 2007 Dec;47(6):453-67.

8. Han FX, Su Y, Monts DL, Waggoner CA, Plodinec MJ: **Binding, distribution, and plant uptake of mercury in a soil from Oak Ridge, Tennessee, USA:** Sci Total Environ. 2006 Sep 15;368(2-3):753-68. Epub 2006 Mar 29.

9. Jasenec A, Barasa N, Kulkarni S, Shaik N, Moparthi S, Konda V, Caguiat J: **Proteomic profiling of L-cysteine induced selenite resistance in *Enterobacter* sp. YSU:** Proteome Sci. 2009 Aug 28;7:30.

10. Kramer GF, Ames BN: **Mechanisms of mutagenicity and toxicity of sodium selenite (Na_2SeO_3) in *Salmonella typhimurium*:** Mutat Res. 1988 Sep;201(1):169-80.

11. Lacerda CM, Reardon KF: **Environmental proteomics: applications of proteome profiling in environmental microbiology and biotechnology:** Brief Funct Genomic Proteomic. 2009 Jan;8(1):75-87. Epub 2009 Mar 11.

12. Lacourciere GM, Levine RL, Stadtman TC: **Direct detection of potential selenium delivery proteins by using an *Escherichia coli* strain unable to incorporate selenium from selenite into proteins:** Proc Natl Acad Sci U S A. 2002 Jul 9;99(14):9150-3. Epub 2002 Jun 25.

13. Mark R. Bruins, Sanjay Kapil, Frederick W. Oehme: **Microbial Resistance to Metals in the Environment:** Received 27 May 1999. Available online 25 March 2002.

14. Nies DH: **Efflux-mediated heavy metal resistance in prokaryotes:** FEMS Microbiol Rev. 2003 Jun;27(2-3):313-39.

15. Rocco F, De Gregorio E, Colonna B, Di Nocera PP: ***Stenotrophomonas maltophilia*** **genomes: a start-up comparison:** Int J Med Microbiol. 2009 Dec;299(8):535-46. Epub 2009 Jul 1

16. Rouch DA, Lee BT, Morby AP: **Understanding cellular responses to toxic agents: a model for mechanism-choice in bacterial metal resistance:** J Ind Microbiol. 1995 Feb;14(2):132-41.

17. Su CC, Long F, Yu EW: **The Cus efflux system removes toxic ions via a methionine shuttle:** Protein Sci. 2011 Jan;20(1):6-18.

18. Waldron KJ, Robinson NJ: **How do bacterial cells ensure that metalloproteins get the correct metal?** Nat Rev Microbiol. 2009 Jan;7(1):25-35.

19. Widner TE, Ripple SR, Buddenbaum JE.Health Phys: **Identification and screening evaluation of key historical materials and emission sources at the Oak Ridge Reservation:** 1996 Oct;71(4):457-69.

20. Anderson J, Wright D, Meksem K: **Agarose gel electrophoresis and polyacrylamide gel electrophoresis for visualization of simple sequence repeats:** Methods Mol Biol. 2013;1006:167-77. doi: 10.1007/978-1-62703-389-3_12.

21. Hsing W, Russo FD, Bernd KK, Silhavy TJ: **Mutations that alter the kinase and phosphatase activities of the two-component sensor EnvZ.** J: Bacteriol. 1998 Sep;180(17):4538-46.

22. Hoch JA, Silhavy TJ, Bourret RB: **Two-component signal transduction**: 1995 - molbio1.princeton.edu

23. Klumpp S, Krieglstein J: **Phosphorylation and dephosphorylation of histidine residues in proteins:** Eur J Biochem. 2002 Feb;269(4):1067-71.

24. Alex LA, Simon Mi: **Protein histidine kinases and signal transduction in prokaryotes and eukaryotes:** Trends Genet. 1994 Apr;10(4):133-8.

25. Reznikoff WS, Goryshin IY, Jendrisak JJ: **Tn5 as a molecular genetics tool: In vitro transposition and the coupling of in vitro technologies with in vivo transposition**: Methods Mol Biol. 2004;260:83-96.

26. Goryshin IY, Jendrisak J, Hoffman LM, Meis R, Reznikoff WS: **Insertional transposon mutagenesis by electroporation of released Tn5 transposition complexes:** Nat Biotechnol. 2000 Jan;18(1):97-100.

27. Goryshin IY, Reznikoff WY: **Tn5 in vitro transposition:** Journal of Biological Chemistry, 1998 - ASBMB Vol. 273, No. 13, Issue of March 27, pp. 7367–7374, 1998

28. Scott, J. A.Palmer, S. J: **Sites of cadmiun uptake in bacteria used for biosorption Applied Microbiology and Biotechnology**: Volume 33, issue 2 (May 1990), p. 221 – 225

29. Nies D. H: **Resistance to cadmium, cobalt, zinc, and nickel in microbes**: Plasmid, Volume 27, issue 1 (January, 1992), p. 17-28

30. Mann CM, Markham JL: **A new method for determining the minimum inhibitory concentration of essential oils:** J Appl Microbiol. 1998 Apr;84(4):538-44.

31. McGinnis S, Madden TL: **BLAST: at the core of a powerful and diverse set of sequence analysis tools:** Nucleic Acids Res. 2004 Jul 1;32(Web Server issue):W20-5.

32. Silver S, Walderhaug M: **Gene regulation of plasmid- and chromosome-determined inorganic ion transport in bacteria:** Microbiol Rev 1992 Mar;56(1):195-228

33. Silver S ; Phung LT: **Bacterial heavy metal resistance: new surprises:** Annual Review Of Microbiology Annu Rev Microbiol] 1996; Vol. 50, pp. 753-89.

34. Joho, M, Inouhe, M, Tohoyama, H, Murayama, T: **Nickel resistance mechanisms in yeasts and other fungi. Journal of Industrial Microbiology**: Volume 14, issue 2 (February 1995), p. 164 – 168

35. Rensing C, Pribyl T, Nies DH: **New functions for the three subunits of the CzcCBA cation-proton antiporter**: J Bacteriol. 1997 Nov;179(22):6871-9.

36. Van der Lelie D, Schwuchow T, Schwidetzky U, Wuertz S, Baeyens W, Mergeay M, Nies DH: **Two-component regulatory system involved in transcriptional control of heavy-metal homoeostasis in** ***Alcaligenes eutrophus***: Mol Microbiol. 1997 Feb;23(3):493-503.

37. Parkinson JS, Kofoid EC: **Communication modules in bacterial signaling proteins:** Annu Rev Genet. 1992;26:71-112.

38. Venkataramana Konda. **Identification of Metal Resistance Genes in a Strain of Enterobacter cloacae** (2008)

39. Hunke S, Keller R, Müller VS. **Signal integration by the Cpx-envelope stress system.** FEMS Microbiol Lett. 2012 Jan;326(1):12-22. doi: 10.1111/j.1574-6968.2011.02436.x. Epub 2011 Nov 8**.**

40. Thede GL, Arthur DC, Edwards RA, Buelow DR, Wong JL, Raivio TL, Glover JN. **Structure of the periplasmic stress response protein CpxP.** J Bacteriol. 2011 May;193(9):2149-57. doi: 10.1128/JB.01296-10. Epub 2011 Feb 11.

41. Ruiz N, Silhavy TJ. **Sensing external stress: watchdogs of the *Escherichia coli* cell envelope.** Curr Opin Microbiol. 2005 Apr;8(2):122-6.

42. Buelow DR, Raivio TL. **Cpx signal transduction is influenced by a conserved N-terminal domain in the novel inhibitor CpxP and the periplasmic protease DegP.** J Bacteriol. 2005 Oct;187(19):6622-30.

43. Isaac DD, Pinkner JS, Hultgren SJ, Silhavy TJ. **The extracytoplasmic adaptor protein CpxP is degraded with substrate by DegP.** Proc Natl Acad Sci U S A. 2005 Dec 6;102(49):17775-9. Epub 2005 Nov 22

44. Foussard M, Cabantous S, Pédelacq J, Guillet V, Tranier S, Mourey L, Birck C, Samama J. **The molecular puzzle of two-component signaling cascades.** Microbes Infect. 2001 Apr;3(5):417-24.

45. Kim D, Forst S. **Genomic analysis of the histidine kinase family in bacteria and archaea.** Microbiol Rev. 1989 Dec;53(4):450-90.

46. Cheung J, Hendrickson WA**. Sensor domains of two-component regulatory systems.** Curr Opin Microbiol. 2010 Apr;13(2):116-23. doi: 10.1016/j.mib.2010.01.016. Epub 2010 Mar 10.

47. MacRitchie DM, Buelow DR, Price NL, Raivio TL. **Two-component signaling and gram negative envelope stress response systems.** Adv Exp Med Biol. 2008;631:80-110. doi: 10.1007/978-0-387-78885-2_6.

48. Price NL, Raivio TL. **Characterization of the Cpx regulon in *Escherichia coli* strain MC4100.** J Bacteriol. 2009 Mar;191(6):1798-815. doi: 10.1128/JB.00798-08. Epub 2008 Dec 19.

49. Yamamoto K, Ishihama A**. Characterization of copper-inducible promoters regulated by CpxA/CpxR in *Escherichia coli.*** Biosci Biotechnol Biochem. 2006 Jul;70(7):1688-95.

50. Raivio TL, Leblanc SK, Price NL. **The *Escherichia coli* Cpx Envelope Stress Response Regulates Genes of Diverse Function That Impact Antibiotic Resistance and Membrane Integrity**. J Bacteriol. 2013 Jun;195(12):2755-67. doi: 10.1128/JB.00105-13. Epub 2013 Apr 5.

51. De Wulf P, Lin EC. **Cpx two-component signal transduction in *Escherichia coli*: excessive CpxR-P levels underlie CpxA* phenotypes.** J Bacteriol. 2000 Mar;182(5):1423-6.

52. Yamamoto K, Ishihama A. **Transcriptional response of *Escherichia coli* to external copper.** Mol Microbiol. 2005 Apr;56(1):215-27.

53. Cao J, Woodhall MR, Alvarez J, Cartron ML, Andrews SC. **EfeUOB (YcdNOB) is a tripartite, acid-induced and CpxAR-regulated, low-pH Fe2+ transporter that is cryptic in *Escherichia coli* K-12 but functional in *E. coli* O157:H7.** Mol Microbiol. 2007 Aug;65(4):857-75. Epub 2007 Jul 11.

54. De Wulf P, Kwon O, Lin EC. **The CpxRA signal transduction system of *Escherichia coli*: growth-related autoactivation and control of unanticipated target operons**. J Bacteriol. 1999 Nov;181(21):6772-8.

55. Raivio TL, Silhavy TJ. **Transduction of envelope stress in *Escherichia coli* by the Cpx two-component system.** J Bacteriol. 1997 Dec;179(24):7724-33.

56. Raivio TL, Popkin DL, Silhavy TJ. **The Cpx envelope stress response is controlled by amplification and feedback inhibition.** J Bacteriol. 1999 Sep;181(17):5263-72.

57. Pogliano J, Lynch AS, Belin D, Lin EC, Beckwith J. **Regulation of *Escherichia coli* cell envelope proteins involved in protein folding and degradation by the Cpx two-component system.** Genes Dev. 1997 May 1;11(9):1169-82.

58. Bebien M, Chauvin JP, Adriano JM, Grosse S, Verméglio A. **Effect of selenite on growth and protein synthesis in the phototrophic bacterium Rhodobacter sphaeroides.** Appl Environ Microbiol. 2001 Oct;67(10):4440-7.

59. Dutta R, Qin L, Inouye M. **Histidine kinases: diversity of domain organization.** Mol Microbiol. 1999 Nov;34(4):633-40.

60. Fleischer R, Heermann R, Jung K, Hunke S. **Purification, reconstitution, and characterization of the CpxRAP envelope stress system of *Escherichia coli*.** J Biol Chem. 2007 Mar 23;282(12):8583-93. Epub 2007 Jan 26.

61. West AH, Stock AM. **Histidine kinases and response regulator proteins in two-component signaling systems.** Trends Biochem Sci. 2001 Jun;26(6):369-76.

62. Buckler DR, Zhou Y, Stock AM. **Evidence of intradomain and interdomain flexibility in an OmpR/PhoB homolog from *Thermotoga maritima*.** Structure. 2002 Feb;10(2):153-64.

63. Leblanc SK, Oates CW, Raivio TL. **Characterization of the induction and cellular role of the BaeSR two-component envelope stress response of *Escherichia coli*.** J Bacteriol. 2011 Jul;193(13):3367-75. doi: 10.1128/JB.01534-10. Epub 2011 Apr 22.