



8-2012

Regulation of Expression of Enterohemolysin Toxin by the Global Repressor HNS and Characterization of a New Class of Protein Produced by Shiga Toxin-Producing Escherichia Coli

Miles T. Rogers
Western Michigan University, milesrogers@gmail.com

Follow this and additional works at: <https://scholarworks.wmich.edu/dissertations>

 Part of the Biochemistry Commons

Recommended Citation

Rogers, Miles T., "Regulation of Expression of Enterohemolysin Toxin by the Global Repressor HNS and Characterization of a New Class of Protein Produced by Shiga Toxin-Producing Escherichia Coli" (2012). *Dissertations*. 73.

<https://scholarworks.wmich.edu/dissertations/73>

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact wmu-scholarworks@wmich.edu.



REGULATION OF EXPRESSION OF ENTEROHEMOLYSIN TOXIN BY THE
GLOBAL REPRESSOR HNS AND CHARACTERIZATION OF
A NEW CLASS OF PROTEIN PRODUCED BY
SHIGA TOXIN-PRODUCING
ESCHERICHIA COLI

by

Miles T Rogers

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Doctor of Philosophy
Department of Biological Sciences
Advisor: Maria Scott, Ph.D.

Western Michigan University
Kalamazoo, Michigan
August 2012

THE GRADUATE COLLEGE
WESTERN MICHIGAN UNIVERSITY
KALAMAZOO, MICHIGAN

Date 7/3/2012

WE HEREBY APPROVE THE DISSERTATION SUBMITTED BY

Miles T. Rogers

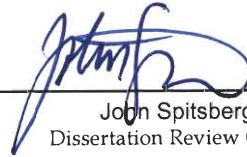
ENTITLED Regulation of the Enterohemolysin Operon by the Histone-like Nucleoid Structuring Protein and Characterization of the Protein YghJ in Shiga Toxin-Producing Escherichia coli

AS PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF Doctor of Philosophy

Biological Sciences

(Department)



John Spitsbergen, Ph.D.
Dissertation Review Committee Chair

Biological Sciences

(Program)



Maria Scott, Ph.D.
Dissertation Review Committee Member



David Huffman, Ph.D.
Dissertation Review Committee Member



Susan Stapleton, Ph.D.
Dissertation Review Committee Member

APPROVED


Dean of The Graduate College

Date August 2012

REGULATION OF EXPRESSION OF ENTEROHEMOLYSIN TOXIN BY THE
GLOBAL REPRESSOR HNS AND CHARACTERIZATION OF
A NEW CLASS OF PROTEIN PRODUCED BY
SHIGA TOXIN-PRODUCING
ESCHERICHIA COLI

Miles T. Rogers, Ph.D.

Western Michigan University, 2012

Shiga toxin-producing *Escherichia coli* (STEC) are worldwide pathogens causing an estimated 200,000 infections per year in the United States. Infections with STEC sometimes progress to kidney failure ending in death. Two studies presented below describe a novel virulence factor and mechanisms required for pathogenesis of STEC seropathotype O91:H21.

First the novel protein, YghJ, was identified and found to be conserved amongst extraintestinal and diarrheal seropathotypes of *E. coli* with similar proteins carried by pathogens such as *Vibrio cholerae*. When the *yghJ* gene was disrupted creating a YghJ null mutant, the defect resulted in a significant growth deficiency in minimal media with L-malate as the sole carbon source. The null mutant growth defect was complemented by replacement of the mutated copy with an intact copy of *yghJ* expressed from a high copy number plasmid. Structure-function analysis of YghJ revealed lipidation restricts the YghJ protein to the inner membrane. Further studies suggest an amino-terminal domain involved in carbon binding and a carboxy-terminal region with mucinase activity. Taken together, the data indicates that YghJ belongs to a new class of protein and plays an

important role in carbon utilization since oxidation of L-malate to oxaloacetate is obligatory for survival of facultative anaerobes like *E. coli*.

The second study focused on regulation of expression of the enterohemolysin toxin gene (*ehxA*) carried by STEC. The role of the HNS protein, a global inhibitor, was investigated. Insertional mutation of the *hns* gene was constructed in STEC O91:H21 and revealed that the amount of *EhxA* increased in the *hns* knockout strain. The HNS protein was found to control expression of the *ehxA* gene through direct binding of the promoter upstream of the enterohemolysin operon (*ehxCABD*). In total, the data indicates that in STEC the 88 bp DNA region upstream of *ehxCABD* contains a cis-acting element to which HNS binds and negatively regulates expression of enterohemolysin.

In conclusion, YghJ participates in a key carbon metabolic pathway and provides a competitive advantage in carbon-limited environments where L-malate is present. Compelling evidence suggests that HNS interferes with RNA polymerase binding blocking transcription of *ehxCABD* in STEC grown *ex vivo*.

© 2010 Miles T. Rogers

ACKNOWLEDGMENTS

I would like to acknowledge the people who have helped me throughout my graduate career at WMU. Most of all, I would like to thank my research advisor Dr. Maria Scott for her unending patience, support, and guidance. I would like to extend special thanks to my dissertation committee, Drs. Susan Stapleton, David Huffman, and John Spitsbergen. I would also like to recognize Dr. Silvia Rossbach for her support.

I am also extremely grateful to my many friends and colleagues, especially Rachel Zimmerman for her contribution to the work described here. I would also like to thank Petra Kohler, Vanessa Revindran, Mary Thwaites, Steven Conrad, Ee Leng Choong, and Vivian Locke who have assisted me in numerous ways. I am also grateful for the love and support of my parents, Donald and Beverly Rogers, and my sister, Lauren. Finally, I would like to thank the Department of Biological Sciences, the WMU Graduate College and the Office of the Vice President for Research at WMU for funding.

Miles T. Rogers

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	ii
LIST OF TABLES.....	vii
LIST OF FIGURES	viii
INTRODUCTION	1
<i>Escherichia coli</i> and its Environments.....	1
Seropathotypes of <i>Escherichia coli</i>	2
Overview of Shiga Toxin-producing <i>E. coli</i>	3
STEC Infections and Associated Shiga Toxin	6
The Type I and Type II Protein Secretion Pathways of STEC	7
The Major Adherence Factor for LEE-Negative STEC is Unknown	8
The Role of HNS in Regulation of Virulence Factors Carried by STEC	9
The Large Virulence Plasmid Encodes the Enterohemolysin Operon of STEC.....	10
What is Known about the Novel Lipoprotein, YghJ?	11
Rationale for These Studies.....	14
MATERIALS AND METHODS.....	16
Bacterial Strains and Growth Conditions.....	16
Measurement of Extracellular Hemolytic Activity	16
Detection of Hemolysis on Blood Agar Plates.....	18
Preparation of Bacterial Lysates for Detection of Enterohemolysin by SDS- polyacrylamide Gel Electrophoresis (SDS-PAGE) and Immunoblot Analysis	18
Cloning of <i>ehxCABD</i> from STEC O91:H21	19

Table of Contents - Continued

Cloning of Partial <i>ehx</i> Operon and Construction of pMC <i>ehxCA'</i> -lacZ Transcriptional Fusion Construct	20
Cloning of <i>hns</i> from Strain B2F1	21
Directional Cloning of <i>hns</i> into the pBAD202/DTOPO Expression Vector	21
Electrophoretic Mobility Shift Assay (EMSA)	22
Preparation of the HNS fusion protein	22
Preparation of the promoter region of <i>ehx</i> and procedures of the EMSA	23
Measurement of β -galactosidase Activity	24
DNA Sequencing.....	25
Comparative Analysis of <i>E. coli</i> Grown on Specific Carbon Sources.....	25
Cellular Fractionation to Determine Subcellular Location of YghJ	25
Globomycin Inhibition of Signal Sequence Cleavage of YghJ.....	27
Cloning the <i>yghJ</i> Gene along with Upstream and Downstream Intervening DNA Sequences	27
Cloning of <i>yghJ</i> from <i>E. coli</i> MG1655 and Generation of the Histidine-Tagged Recombinant YghJ	31
Cloning the open-reading-frame (ORF) of <i>yghJ</i>	31
Directional Cloning of <i>yghJ</i> onto pBAD TOPO Expression Vector to Create a YghJ Fusion Protein Containing a N-terminal Thioredoxin and C-terminal Polyhistidine Tag	32
Induction of the Topo-YghJ-6X His fusion protein from the pBAD202/D-TOPO plasmid	33
Induction and Purification of YghJ-6X His Recombinant Protein Using Clontech HisTalon Gravity Columns	34
Induction and Purification of YghJ-6X His using Pierce B-PER 6XHis Fusion Protein Purification Kit	35
Purification Attempt of YghJ-6X His using Invitrogen Dynabeads His-Tag Isolation Beads	35
Western Blotting of YghJ-6X His.....	36
In Gel Assay to Detect Phosphatase Activity of YghJ.....	36

Table of Contents - Continued

RESULTS	42
Proteomic Analysis and Comparison of Shiga Toxin-Producing <i>Escherichia coli</i> (STEC) O91:H21 and STEC O157:H7 Isolates	42
Comparative Two-Dimensional Gel Analysis of Extracytoplasmic Proteins from LEE-negative versus LEE-positive STEC.....	42
Genetic Analysis of the 160 kDa Protein Identified as YghJ: a Novel Extracytoplasmic Protein Produced by STEC O91:H21	45
Location of the <i>yghJ</i> gene on the chromosome of STEC O91:H21 and <i>E. coli</i> K12 substrain MG1655	45
Structure/Function Analysis of the Novel Protein, YghJ.....	51
Putative Protein phosphatase 2C Domain.	55
In Gel Assay to Test for PP2C Activity of the YghJ protein-.....	61
Analysis of a YghJ null mutant.....	62
Construction of the <i>yghJ</i> null mutant.	62
Capacity of the YghJ knockout mutant to metabolize different carbon sources	63
Complementation of the <i>yghJ</i> null mutant	66
Expression and attempts to purify the YghJ-6xHis fusion protein	68
Expression of YghJ from an inducible vector	68
Purification of YghJ using metal affinity chromatography.....	70
Immunoblot to detect the YghJ-6xHis recombinant protein	76
Localization of YghJ.....	77
Membrane fractionation of YghJ	77
Globomycin Treatment of YghJ.....	78
Comparison of Growth of STEC O91:H21 and STEC O157:H7 in Minimal M9 Media with L-malate as Sole Carbon Source.....	80
Expression of Enterohemolysin	82
EhxA production on SRBC agar	82
Comparison of Enterohemolysin levels in the HNS knockout mutant 34.7	84
Immunoblot analysis of EhxA secreted by the HNS null mutant, 34.7	87
Relative enterohemolysin activity in culture supernatants of 34.7 versus STEC O91:H21	88
HNS binding to the putative ehx promoter region.....	91

Table of Contents - Continued

HNS binding assays to the putative promoter region of the ehx operon carried by STEC O91:H21	91
Transcription of the ehx operon in the presence of HNS	93
Comparison of upstream nucleotide sequence of the ehx operon between STEC O91:H21 and STEC O157:H7.....	95
DISCUSSION AND CONCLUSION	98
Characterization of YghJ a novel protein involved in carbon metabolism and host colonization of Escherichia coli organisms.....	98
Defining the mechanism of HNS repression of the enterohemolysin operon.....	104
REFERENCES	110

LIST OF TABLES

1. Virulence Factors Produced by LEE-Positive STEC as Compared to Virulence Factors Produced by LEE negative STEC.....	38
2. List of strains used in this study.....	39
3. List of plasmids used in this study.....	40
4. List of primers used in this study.....	41

LIST OF FIGURES

1. Schematic representation of the position of yghJ on the chromosome.....	29
2. PCR product and results of screening of putative pCB2YghJ clones.....	30
3. Restriction analysis of pCMGORFyghJ clones.	31
4. Two-Dimensional Gel Analysis of the Proteomes of LEE-negative STEC O91:H21 and LEE positive STEC O157:H7.....	43
5. Two-Dimensional Gel Analysis of the Proteomes of LEE-negative STEC O91:H21 and LEE positive STEC O157:H7.....	44
6. Two-dimensional gel analysis of STEC O91:H21.....	44
7. Phylogenetic relationships of YghJ protein sequences produced by other Gram negative species based on neighbor joining analysis.....	49
8. Phylogenetic relationships of YghJ protein sequences of Escherichia species based on maximum likelihood analysis.....	50
9. Schematic diagram of the novel YghJ protein.....	51
10. Multiple sequence alignment of the DUF4092 domain of YghJ with similar proteins.....	54
11. Multiple sequence alignment of the M60-like domain of YghJ with similar proteins.....	57
12. Multiple alignment of the PP2C superfamily showing the 11 conserved motifs.	60
13. Assay for PP2C activity of the YghJ protein.	62
14. The YghJ null mutant exhibits a normal growth phenotype when grown on M9 minimal media containing 0.1% glucose as the sole carbon source.	64
15. The YghJ null mutant exhibits a normal growth phenotype when grown on M9 minimal media with 0.1% fumarate as the sole carbon source.....	64
16. The YghJ null mutant exhibits a normal growth phenotype when grown on M9 minimal media containing 0.1% succinate as the sole carbon source.	65

List of Figures - Continued

17. The YghJ null mutant exhibits a deficient growth phenotype when grown on 0.1% L-malate as the sole carbon source.	66
18. The YghJ null mutant is able to partially recover growth in 0.1% L-malate when complemented with the yghJ gene expressed from a high copy number plasmid.	67
19. SDS-PAGE analysis of YghJ expressed from the arabinose inducible Pbad202D/Topo vector.	68
20. SDS-PAGE analysis of YghJ expressed from the IPTG inducible pET21a vector.	69
21. SDS-PAGE analysis of an attempted purification of YghJ-6XHis using the Pierce B-PER 6XHis protein purification kit.	71
22. SDS-PAGE analysis of an attempted purification of YghJ-6XHis using the Pierce B-PER 6XHis protein purification kit containing 5 mM β -mercaptoethanol.	71
23. SDS-PAGE gel analysis of a purification performed using the Clonotech HisTALON Gravity Purification Columns.	73
24. SDS-PAGE gel analysis of YghJ purification using the Clonotech HisTALON Gravity Purification Columns.	74
25. SDS-PAGE gel analysis of YghJ purification using the Clonotech HisTALON Gravity Purification Columns.	75
26. SDS-PAGE gel analysis of YghJ purification using Invitrogen Dynabeads His-Tag Isolation Beads.	76
27. SDS-PAGE gel analysis of membrane fractionation of YghJ.	78
28. SDS-PAGE analysis of globomycin treated BL21 (DE3) expressing the YghJ protein from the induced pET21a vector.	79
29. The YghJ negative pathogenic STEC O157:H7 strain 86-24 exhibits a deficient growth phenotype compared to the YghJ positive STEC O91:H21 strain B2F1 when grown on 0.1% L-malate as the sole carbon source.	81
30. Comparison of enterohemolytic phenotypes on SRBC agar.	82

List of Figures - Continued

31. Analysis of representative levels of enterohemolysin produced by different strains.	86
32. Hemolytic tube Assay: Comparison of extracellular hemolytic activity present in supernatants recovered from cultures of wild-type strains and hns mutants.	90
33. Electrophoretic mobility shift assay (EMSA): Binding of HNS to the putative promoter region of ehx from STEC O91:H21.	92
34. Electrophoretic mobility shift assay to determine HNS binding to the putative ehx promoter region of STEC O157:H7.....	92
35. Effect of the hns mutation carried by THK62 on the expression of β -galactosidase from the ehxCA':::lacZ reporter gene fusion.	95
36. Sequence alignment of the putative promoter region of ehx described for STEC O113:H21 strains EH41 which lies upstream of the putative translational start site of ehxC.....	97

INTRODUCTION

***Escherichia coli* and its Environments**

Escherichia coli are Gram-negative rod shaped facultative anaerobic bacteria that adapt and live in a variety of settings. For example, *E. coli* have been isolated from the environment (Nautiyal et al., 2010; Semenov et al., 2010) and non-pathogenic isolates of *E. coli* exist in a mutualistic relationship with the mammalian host. Mutualistic *E. coli* receive nutrients provided through metabolic cast-offs of the host and from by-products of metabolites produced by other commensal bacteria that occupy the same niche within the mammalian intestine (Dave et al., 2012). *E. coli* make-up less than 0.1% of the human gut microbiome, in fact in the healthy human host, levels of *E. coli* are kept in check (Dave et al., 2012). The relationship between the host and *E. coli* is complex and still somewhat mysterious but we do know that *E. coli* produce vitamins and other metabolic products required for human health (Drewke and Leistner, 2001; Moyne et al., 2011; Preidis et al., 2011).

E. coli is not limited to the mammalian gastrointestinal tract but is capable of survival on plants (Ibekwe et al., 2009; Ibekwe et al., 2011). For example, the pathogenic Shiga toxin-producing *E. coli* (STEC) serotype O157:H7 is able to survive and persist after artificial inoculation on lettuce (Moyne et al., 2011). Further analysis reveals that *E. coli* survive on the leaves of lettuce as well as in the root system of plants (rhizosphere) (Ibekwe et al., 2009). Surprisingly, *E. coli* persist in environments irrigated by

contaminated water, suggesting a possible route of contamination (Ibekwe et al., 2004). As we begin to move towards a more global food environment, persistence of pathogenic *E. coli* becomes a larger issue, as seen by the increasing number of infections in the developed and the developing world (Pires et al., 2009; Scallan et al., 2011). Notably the 2011 outbreak in Germany caused by fenugreek shoots contaminated by Shiga toxin-producing *E. coli* serotype O104:H4 is an example of global food markets as a source of contamination (European Food Safety Authority, 2011). The seeds were imported to Germany from Egypt and the shoots served to customers resulted in one of the largest and most lethal STEC outbreaks in Europe to date (European Food Safety Authority, 2011) (Kenny et al., 1997).

Seropathotypes of *Escherichia coli*

Different serogroups of *E. coli* are pathogenic (seropathotypes) and cause infection by different mechanisms. Many seropathotypes of *E. coli* colonize the gastrointestinal tract but some persist outside of the gut causing infection at other sites. Depending on the niche in the host, seropathotypes of *E. coli* are extraintestinal pathogenic *E. coli* (ExPEC) or urogenital pathogens (i.e. uropathogenic) *E. coli* (UPEC) causing urinary tract infections (Patton *et al.* 1991). UPEC form biofilms and adhere to uroepithelial cells of the urinary tract (Kikuchi et al., 2005) as well attach to indwelling medical catheters (Jacobsen et al., 2008). Other ExPEC strains cause bacteremia leading to sepsis or meningitis (Johnson and Russo, 2005).

A great deal is known about the intestinal pathogenic strains of *E. coli* but there are still mechanisms of infection involved in pathogenesis that we do not yet understand. Intestinal *E. coli* seropathotypes cause diarrheal disease through a variety of mechanisms and toxins many times play a major role. Enterotoxigenic *E. coli* (ETEC) strains cause travelers' diarrhea, are noninvasive and produce two toxins, a heat stable toxin and a heat-labile toxin (Isidean et al., 2011). The Enteroaggregative seropathotype of *E. coli* (EAEC) have a unique pattern of aggregation when attached to epithelial cells and produce up to three toxins (Nataro et al., 1995). The Enteroinvasive *E. coli* (EIEC) are similar to *Shigella* because they actively invade and grow inside of host cells causing intestinal epithelial cell death, watery diarrhea as well as a potent inflammatory response described as dysentery (DuPont et al., 1971; Formal and Hornick, 1978). EIEC seropathotypes produce a Shigella-like toxin but the role of the toxin in disease is unclear (DuPont et al., 1971; Formal et al., 1978). Enteropathogenic *E. coli* (EPEC) causes infantile diarrhea and is a severe problem in developing countries where good nutrition is lacking. EPEC carry a large pathogenicity island required for intimate adherence to human epithelial cells (Kenny et al., 1997). The adherence genes encoded by EPEC have also been acquired by certain serotypes of Shiga toxin-producing *E. coli*.

Overview of Shiga Toxin-producing *E. coli*

Shiga toxin-producing *E. coli* (STEC) (Karmali et al., 1983a; Karmali et al., 1983b) encode Shiga toxin and are the subject of this report. The Shiga toxin is a variant of the Shigella toxin produced by *Shigella dysenteriae* (Head et al., 1988; Karmali et al.,

1985;Karmali et al., 1986). Shiga toxin genes are carried by all STEC and the toxin is typically encoded by a lysogenic bacteriophage. Shiga toxin is cytotoxic to epithelial and endothelial cells that carry the globotriaosylceramide (Gb3) receptor (Karmali et al., 1983a;Karmali et al., 1985). The first STEC strains were discovered in 1977 and isolated in samples taken from patients with severe diarrhea (Konowalchuk et al., 1977). The bacteria involved in those infections were found to produce a toxin that destroyed Vero epithelial cells. Later work (O'Brien and LaVeck, 1983;O'Brien et al., 1982) purified the verocytotoxin and it was shown to be identical to the toxin produced by *Shigella dysenteriae*. Kidney damage suffered by patients with hemolytic uremic syndrome (HUS)(Karmali et al., 1983b) is caused by the Shiga toxin produced by STEC.

Zoonotic STEC infections are acquired by ingesting contaminated water or food but STEC are also transferable person-to-person (Melton-Celsa et al., 2012). Serotype STEC O157:H7 is the most commonly isolated STEC serotype that causes food and water borne outbreaks in the United States. There are an estimated 63,153 cases of illness caused by STEC O157:H7 and 112,752 infections caused by non-O157 serotypes in the US annually (Scallan et al., 2011). Based on genetic composition, STEC seropathotypes are divided into two Major groups: (1) STEC O157:H7 and its relatives comprise groups 1 and 2; and (2) STEC O91:H21 and related serotypes include STEC groups 3 and 4 (Melton-Celsa et al., 2012). Due to inadequate detection methods, the exact number of cases of infection and death caused by non-O157:H7 strains are unknown (Henao et al., 2010;Scallan et al., 2011).

Following ingestion of contaminated food or water, STEC groups 1 and 2 utilize specialized adherence factors to move to the site of colonization, the large intestine, and adhere tightly to colonic epithelial cells. Genetic markers that enable intimate adherence of STEC 1 and 2 group members are encoded on a large pathogenicity island known as the locus of enterocyte effacement (**LEE**) (McDaniel et al., 1995). The LEE is 42 kb and codes for genes essential for host colonization. The LEE genes include the translocated intimin receptor (*tir*), the major adhesion, intimin, coded for by the *eae* gene, the type III secretion system and various other effector proteins (Elliott et al., 1998; Elliott et al., 1999). Once contact is made with the host cell, the type III secretion system is assembled and the Tir protein is translocated into the membrane of the host epithelial cell by way of the type III secretion system (Kenny et al., 1997). Other host cell effectors secreted through the type III system and made by STEC instruct host actin to form a pedestal-like structure to which O157:H7 attach. Adherence to the host epithelial cell layer leads to effacement of the microvilli at the site of attachment and formation of a diagnostic scar called the attach and efface lesion (A/E lesion) (Frankel and Phillips, 2008). **Non-O157:H7 STEC groups 3 and 4 isolates do not code for the LEE pathogenicity island and do not form A/E lesions** (Bugarel et al., 2010). However the STEC group 3 and 4 organisms do colonize the host and cause the same disease phenotype as do STEC group 1 and 2 organisms (Scallan et al., 2011; Melton-Celsa et al., 2012).

STEC Infections and Associated Shiga Toxin

STEC infections sometimes present as bloody diarrhea or hemorrhagic colitis (HC) and progress in 5-10% of cases to hemolytic uremic syndrome (HUS); a severe sequelae that 10-15% of the time leads to kidney failure and death (Gould et al., 2009a;Gould et al., 2009a;Gould et al., 2009b). Shiga toxin produced by STEC is the major factor cause of kidney failure caused by toxin-exacerbated thrombocytopenia and endothelial cell apoptotic cell death (Tesh, 2010).

Shiga toxin is released from the bacterial cell by an unknown mechanism and binds to the host cell attaching to globotriaosylceramide (Gb3) receptors present in the plasma membrane of specific epithelial cell types (Abe et al., 2000;Okuda et al., 2006). Following binding, Shiga toxin is phagocytosed brought into the endosome and transported to the endoplasmic reticulum rather than being targeted for degradation (Raa et al., 2009;Sandvig et al., 2010). The A1 subunit of the toxin, the active domain, is proteolytically cleaved and released into the cytoplasm of the eukaryotic cell where it inactivates the large subunit of the host ribosome (Endo and Tsurugi, 1988b). The exact mechanism of inactivation of the ribosome involves removal of amino acid A-4324 adenine base from the 28S rRNA by cleavage of the N-glycosidic bond (Endo et al., 1988;Endo and Tsurugi, 1988a).

Shiga toxin is composed of two major subunits: (1) a 30 kDa A-subunit that contains the active protein domain; and (2) the 35 kDa B-subunit, the binding domain consisting of a pentamer of seven identical 7 kDa monomers ((Fraser et al., 2004). Two major groups of Shiga toxin (Stx) exist which are immunologically distinct; they are

Shiga toxin types 1 and 2 (O'Brien et al., 1992). Variants of Stx2 that are found in STEC include Stx2a, Stx2b, and Stx2c and Stx2dact. STEC O91:H21 and STEC O104:H21 carry the Stx2dact variant (Ito et al., 1990). The Shiga toxin Stx2dact variant is unique in that when it is exposed to intestinal mucous the last two amino acids of the variant is specifically cleaved by the elastase protein found in mucous. Following the elastase cleavage Stx2dact toxicity is increased (Melton-Celsa et al., 2002;Kokai-Kun et al., 2000;Melton-Celsa et al., 2002). The most clinically important Stx2 variants are Stx2c and Stx2dact (Melton-Celsa et al., 2012) and are associated with poor clinical outcomes

The Type I and Type II Protein Secretion Pathways of STEC

The type I secretion apparatus - The type I secretion apparatus was first described as the pathway for hemolysin toxin secretion in *E. coli* (Bauer and Welch, 1996). Hemolysin is a toxin carried by UPEC encoded on the *hly* operon. The *hly* operon contains 4 genes: *hlyC*, *hlyA*, *hlyB*, and *hlyD* (Bauer and Welch, 1996). The *hlyA* gene codes for the hemolysin toxin and the *hlyC* gene-product acylates and activates the HlyA toxin (Bauer and Welch, 1996). Proteins required for HlyA secretion are coded for by *hlyB* and *hlyD*. The HlyB and HlyD together interact with and recruit the type 1 secretion pore, the TolC protein (Bauer and Welch, 1996;Koronakis et al., 2000;Stanley et al., 1998). The type 1 pathway is also carried by STEC organisms and transports the enterohemolysin toxin. The enterohemolysin toxin is discussed below in a different section.

Type II Secretion Pathway - LEE-positive STEC such as O157:H7 carry genes for the type II secretion operon. Few proteins are known to depend on the type II secretion pathway in STEC (Lathem et al., 2002). The type II secretion apparatus consists of up to 13 genes that when expressed assemble into a multiprotein complex (Sandkvist, 2001b). Proteins that depend on the type II secretion apparatus are translocated from the cytoplasm by the Sec secretion system. The nascent protein folds in the periplasm prior to entrance into the type II secretion pathway (Sandkvist, 2001a) but the signal sequence targeting proteins for secretion through the type II pathway is unknown (Johnson et al., 2006). The complex is powered by the ATPase E protein which interacts with protein L at the inner membrane interface. In *Vibrio cholerae*, membrane bound interaction of proteins EpsE and EpsL are dependent on ATPase hydrolysis, therefore EpsE has dual residence in the cytoplasmic and on the inner membrane bound to EpsL (Camberg and Sandkvist, 2005;Patrick et al., 2011;Robien et al., 2003). Protein C is periplasmic and it is believed to be the gatekeeper for type II protein secretion pathway (Nunn, 1999). The outer membrane pore is made up of 12 subunits of EpsD (Brok et al., 1999). It is thought that in *V. cholerae* EpsE, EpsL and EpsM regulate secretion by signaling through ATP hydrolysis (Camberg et al., 2007).

The Major Adherence Factor for LEE-Negative STEC is Unknown

The type III secretion system and the major adherence factors Tir and intimin are encoded by LEE-positive STEC O157:H7 and non-O157:H7 STEC group 2 isolates (Melton-Celsa et al., 2012). **However, LEE-negative STEC O91:H21 and other group**

3 and 4 isolates do not code for the LEE locus. Multiple adherence factors have been identified that are carried by both LEE-positive and LEE-negative isolates. Yet intimin and Tir are the only adherence proteins proven crucial for colonization of the human host. In fact, there is no evidence available to date that explains the mechanism of adherence or colonization for STEC O91:H21 and related isolates.

The Role of HNS in Regulation of Virulence Factors Carried by STEC

Virulence genes carried on the locus of enterocyte effacement (LEE) are tightly regulated by a variety of regulator proteins. The major regulator of the LEE is the LEE encoded regulator or Ler which activates expression of the LEE genes (Berdichevsky et al., 2005); the Ler protein also regulates itself via an autorepressor mechanism (Berdichevsky et al., 2005). The expression of the *ler* gene is regulated by the global regulatory protein called the histone-like nucleoid structuring protein (HNS) (Bustamante et al., 2001). The HNS protein represses LEE genes while the Ler protein counter-acts HNS repression through activation of LEE gene expression (Bustamante et al., 2001). Two other proteins found on the LEE locus have a role in LEE regulation. These proteins are designated as the global regulator of LEE proteins called the GrlA [activator] and GrlR [repressor]. The GrlA and GrlR proteins are encoded by the *grlRA* operon (Deng et al., 2004; Lio and Syu, 2004).

LEE-negative STEC do not carry the LEE pathogenicity island or the Ler and GrlAR proteins. In fact, the structural adherence genes required for host cell attachment by LEE-negative isolates have not been identified. We propose that HNS must have a

role in regulation of key virulence factors because an HNS null mutant is defective in binding human colonic epithelial cells (Scott et al., 2003). Furthermore, the HNS STEC O91:H21 mutant overexpresses its enterohemolysin toxin and is nonmotile (Rogers et al., 2009; Scott et al., 2003). The facts taken together suggest that HNS simultaneously regulates multiple virulence factors carried by STEC O91:H21.

The Large Virulence Plasmid Encodes the Enterohemolysin Operon of STEC

Both LEE-positive and LEE-negative STEC strains possess a large virulence plasmid of approximately 100 kb pairs. The plasmid encodes a variety of virulence genes, including a catalase peroxidase (Brunner et al., 1996), a serine protease (Brunner et al., 1997), and the enterohemolysin operon (*ehxCABD*) (Schmidt et al., 1994). The Enterohemolysin toxin (*EhxA*) is a member of the repeat in toxin (RTX) family of cytolytic toxins which are widely distributed among many different Gram-negative human and animal pathogens (Welch, 1991). These toxins have common features, such as the capacity to disrupt and kill different types of cells through pore formation and lysis (Stanley et al., 1998).

The most intensely studied member of the RTX toxin family is α -hemolysin (HlyA) encoded by *hlyA*. Analysis of the deduced amino acid sequences of HlyA and *EhxA* indicate that the toxins are related with 60% identity (Bauer and Welch, 1996). The α -hemolysin protein is an important virulence factor of uropathogenic *E. coli* (Smith et al., 2008). The enterohemolysin toxin structural gene, *ehxA*, is a component of the *ehxCABD* operon. The operon, when present, is always carried on the large virulence

plasmid of clinical isolates of STEC (Schmidt et al., 1996). The gene products carried on the *ehx* operon function exactly like those of the prototypic *hly* operon (Schmidt et al., 1996). Specifically, the *ehxB* and *ehxD* genes code for the secretion system required for transport of *EhxA* out of the cell (Bauer and Welch, 1996). Functional enterohemolysin requires acylation of *EhxA* by the *EhxC* protein before it is secreted (Bauer and Welch, 1996).

The enterohemolysin operon in LEE-positive STEC species is regulated by Ler (discussed above) in combination with GrlA and GrlR (Iyoda et al., 2011; Saitoh et al., 2008). The enterohemolysin operon is not expressed under normal laboratory conditions; with little or no toxin detected in the cell or in cellular extracts. Expression of the *ehxCABD* operon is repressed by the HNS protein (Scott, et al, 2003) since a mutation in the *hns* gene in LEE negative STEC serotype O91:H21 increased the production of enterohemolysin which was then detectable under normal laboratory conditions (Scott et al. 2003). Therefore, although the mechanism of had not been discerned prior to our investigations, we hypothesized that the HNS protein plays a role in regulation of expression of the *ehx* operon even in the absence of the LEE encoded regulatory proteins.

What is Known about the Novel Lipoprotein, YghJ?

Shiga toxin-expressing isolates of *E. coli* that lack the LEE which contains the type III secretion apparatus and Type III dependent host effector proteins, do not produce and secrete the same proteins as LEE-positive STEC. Analytical comparison of the proteomes of LEE-positive STEC O157:H7 with LEE-negative STEC O91:H21

revealed several novel proteins produced by LEE negative STEC O91:H21 not produced by STEC O157:H7. Following two-dimensional gel electrophoresis and mass spectrometry, one of these proteins was identified as YghJ.

A high-throughput pull-down assay using a histidine-tagged open reading frame library constructed from published genome sequences of *Escherichia coli* strain K-12 used ORFs of 4339 genes cloned onto an expression vector as bait to evaluate protein-protein interactions between histidine-tagged proteins and their interacting partners (Arifuzzaman et al., 2006). This comprehensive shot-gun study indicated that YghJ interacted with several proteins. All the proteins thought to interact with YghJ of localize to the periplasm and many of these proteins bind to ATP. The proteins include: (1) the FAD dependent malate dehydrogenase enzyme MQO; (2) the ATP binding subunit of the D-allose transporter AlsA responsible for the uptake of the monosaccharide D-allose; (3) the transketolase TktA, which possesses a thymine pyrophosphate binding domain which has an important role in the pentose phosphate and glycolytic pathways providing precursors for nucleotide and vitamin biosynthesis; and (4), pitrilysin Ptr which is responsible for cleaving carbohydrate substrates such as glucagon. Other uncharacterized proteins include (1) YliA, a putative glutathione transporter with an ATP binding cassette; (2) YaiN, a formaldehyde induced regulator protein; and (3) protein YmdF of unknown function.

Interestingly, enteropathogenic *E. coli* seropathotypes also carry the *yghJ* gene (see introduction, pg 12). Enteropathogenic *E. coli* (EPEC) seropathotypes carry the *yghJ* gene and in EPEC YghJ may function in biofilm formation of EPEC(Baldi et al., 2012).

In addition, a recent reverse vaccinology study identified in a mouse model of sepsis a homolog of YghJ carried by an ExPEC isolate as a protective antigen (Moriel et al., 2010).

The YghJ protein is likely homologous to the AcfD protein found in *Vibrio cholerae* since the two proteins share 64% similarity. A possible role for YghJ in bacterial-host interaction was described in *Vibrio cholerae*. Specifically, transposon mutagenesis of *V. cholerae* resulted in null mutants of AcfD that had an insertion within the *acfD* gene, causing a disruption of function of AcfD, were unable to successfully colonize the intestine of an infant mouse, the preferred animal model of cholera infection (Peterson and Mekalanos, 1988).

Regulatory sequences upstream of *yghJ* control transcription of the type II secretion pathway in an Enterotoxigenic *E. coli* strain H10407 (Yang et al., 2007). In this ETEC strain, the *yghJ* gene is encoded on the chromosome upstream of the type II Secretion operon (Yang et al., 2007). A transcriptional fusion revealed that *yghJ* is tightly regulated by HNS in a temperature dependent manner (Yang et al., 2007). In addition, reverse transcriptase PCR indicated that *yghJ* was cotranscribed with the other genes in the ETEC type II secretion operon resulting in a 14 gene transcriptional unit (Yang et al., 2007). This group also suggests that YghJ is a substrate for the type II secretion. Evidence presented herein does not support these findings.

The mission of the Human Microbiome Project is to describe the collections of bacteria, fungi, and archaeae that inhabit the human body. Subsequently, a novel protein called M60-like BT4244 produced by the human intestinal commensalist *Bacteroides*

thetaitaomicron (Collison et al., 2012; Nakjang et al., 2012) was identified and the BT4244 protein was shown to degrade mucin *in vitro* (Nakjang et al., 2012). Further analysis revealed an enzymatic domain near the carboxy-terminus on the protein (Nakjang et al., 2012). The BT4244 protein is similar to a metalloprotease made by Baculovirus called enhancin (Wang and Granados, 1997b). Baculovirus enhancin and the BT4244 protein both possess a mucinase domain and a zinc-binding motif (Wang and Granados, 1997b; Wang and Granados, 1997a; Nakjang et al., 2012). YghJ also contains the mucinase domain as well as the zinc-binding motif. Therefore we speculate that YghJ may have similar mucinase activity.

Mucin is routinely scavenged by bacteria as a carbon source (Collison et al., 2012). Proteins that contain the M60-like domain are commonly detected in microbes that inhabit a mucosa-associated environment which suggests a role for these proteins in host colonization (Nakjang et al., 2012).

Rationale for These Studies

There are a number of major differences between virulence factors produced by LEE-positive and LEE-negative STEC (Table 1). However, both types of STEC persist in the host, colonize and cause identical infections. We have knowledge of many virulence factors involved in LEE positive STEC infection but we do not know what is behind LEE negative STEC pathogenesis. It is essential we begin to dissect the molecular mechanisms important in host pathogenesis of LEE-negative STEC. As our food production system becomes more industrialized and global, there are continual threats to

our food supply from pathogens like STEC. The natural reservoir for STEC organisms was thought restricted to bovine with infection occurring through consumption of contaminated meat. More recently we find that STEC persist on vegetables that we grow and consume, suggesting that we have not identified all sources of infection and novel vectors of infection will continually arise. Discrimination between commensal *E. coli* and non-O157:H7 isolates of STEC is difficult and inefficient which puts the population at great risk. As a result, research focused on understanding mechanisms that underpin pathogenicity of LEE-negative STEC must be intensified because more atypical STEC serotypes are recognized as potent pathogens. Case in point is the recent LEE-negative O104:H4 outbreak in Europe.

The work presented here is an attempt to identify differences in gene regulation of the *ehxCABD* operon and to identify novel virulence factors that define mechanisms of infection for LEE-negative STEC. By filling in the gap in knowledge we can leverage this information in order to develop new treatments and detection methods. Here we examine regulation of the enterohemolysin operon carried by LEE-negative STEC. We found that the global silencer, HNS, directly regulates the *ehxCABD* operon of LEE-negative STEC by binding to its promoter region. Surprisingly we find that HNS also binds to the same region of the *ehx* operon carried by LEE-positive STEC O157:H7. We characterized the product of the *yghJ* gene carried by LEE-negative STEC and found that the YghJ protein provides a competitive growth advantage for LEE-negative STEC O91:H21 when grown on carbon sources like L-malate. This growth advantage may be important in carbon limited environments like the large intestine. This growth advantage

may be important for survival on plants. We are intrigued by the recent discovery of the mucinase domain in YghJ. We hypothesize that the capacity to degrade mucin into useful carbohydrate metabolites such as glucosamine will likely play a strong role in host colonization.

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions

Table 2 describes pertinent facts about strains of bacteria and plasmids used in this study. Bacterial strains were routinely grown to mid-logarithmic or stationary phase in Luria broth (LB) at 37°C as specified. Antibiotics were added as needed and applied at final concentrations for ampicillin (Ap) at 100 µg/ml, chloramphenicol (Cm) at 30 µg/ml, and kanamycin (Km) at 50 µg/ml. The strain transformed with pT-*hns*/His was grown in minimal media with 4% casamino acids then induced with 0.02% arabinose for expression of *hns*. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was used for plasmids with IPTG inducible promoters at concentrations required to induce maximal gene expression.

Measurement of Extracellular Hemolytic Activity

The activity of enterohemolysin produced and secreted by each strain was quantified by a tube assay modified from other procedures (Rennie et al., 1974; Welch and Falkow, 1984; Williams, 1979). Bacterial strains were grown overnight in LB broth at 37°C with aeration. The bacterial culture was diluted 1:50 into LB broth and grown

with aeration until an optical density at 600 nm of 1.0 was reached. Concurrently, sheep red blood cells (SRBC) were washed three times in sterile Dulbecco's PBS supplemented with 10 mM CaCl₂ and 0.1% BSA. A 2% SRBC solution of the washed cells was prepared in Dulbecco's PBS supplemented with 10 mM CaCl₂ and 0.1% BSA. Once the STEC strains to be tested reached the desired optical density, the supernatants were separated from the cells by centrifugation at 5,000 x g for 15 minutes. The supernatants were harvested and filtered to remove any remaining bacterial cells or debris. Next, serial 2-fold dilutions of the supernatant filtrates were made in Dulbecco's PBS supplemented with 10 mM CaCl₂ and 0.1% BSA. A 500 µl sample of the diluted supernatant filtrates was added to an equal volume of washed 2% SRBC solution in an eppendorf tube. The tubes were gently rocked to mix the suspension for 4 hours at 37°C, with inversion of the tubes every 30 minutes. The tubes were then spun in a microcentrifuge for 30 seconds at 13,000 x g. The resultant supernatant was removed from each tube and the optical density at 540 nm measured. The SRBC buffer was the control for background levels of spontaneous hemolysis. Supernatants recovered from *E. coli* strains DH5α, DH10B, or plasmid-cured strain S11 served as negative controls. Maximum lysis was measured in samples that included 500 µl of water added to 500 µl of the 2% SRBC solution. Relative per cent lysis was calculated as follows: $100 \times (A_{540} \text{ of sample} - A_{540} \text{ of background}) / (A_{540} \text{ of total} - A_{540} \text{ of background})$.

Detection of Hemolysis on Blood Agar Plates

Hemolysis activity by wild-type strains of *E. coli* that produce Shiga toxin can not be detected on commercially prepared SRBC plates (Beutin et al., 1989). To enhance the hemolytic activity of enterohemolysin on blood agar plates, SRBC were prepared as follows. The SRBC used in the assay plates were washed three times in 10 mM PBS at pH 7.5 (Beutin et al., 1989). Tryptose blood agar was supplemented with 5% pre-washed SRBC and 10 mM CaCl₂. Bacteria were applied to the blood agar plates incubated overnight at 37°C in a 5% CO₂ atmosphere. The hemolytic phenotypes of the bacterial strains were assessed following overnight incubation.

Preparation of Bacterial Lysates for Detection of Enterohemolysin by SDS-polyacrylamide Gel Electrophoresis (SDS-PAGE) and Immunoblot Analysis

Bacterial strains were grown overnight with aeration at 37°C in LB broth supplemented with antibiotics as needed. The bacterial cultures were then diluted 1:50 into LB broth and grown with aeration to log phase. Bacterial cultures were equalized by optical density readings at 600 nm. Next, half the culture of each strain was removed and centrifuged at 5,000 x g at 4°C for 10 minutes. The supernatant was removed and placed on ice. The bacterial pellets were resuspended in 10 ml of 0.1M PBS. The resuspended pellets were sonicated to lyse the cells. The proteins were precipitated from each of the sonically-disrupted samples by incubation with 10% final concentration of trichloroacetic acid with rocking at 4°C overnight. Subsequently, samples were centrifuged at 10,000 x g for 10 minutes to recover the precipitated proteins. The supernatants were discarded and the pellets washed with cold acetone.

The pellets were dried and resuspended in SDS-PAGE loading buffer and frozen at -20°C.

Samples were heated to 100°C before loading onto 4-12% gradient SDS-PAGE gels. Gels were run at 200 Volts for 50 minutes then stained with colloidal blue (Invitrogen, Carlsbad, CA) or immunoblotted with *Bordetella pertussis* monoclonal antibody, CyaA(9D4) specific for CyaA toxin CyaA(9D4) [Santa Cruz Biotechnology, Inc., Santa Cruz, CA]. CyaA antibody binds to the nonapeptide repeat region of HlyA and cross-reacts with *EhxA* (Bauer and Welch, 1996). Secondary antibody was goat anti-mouse IgG (H + L) conjugated to horseradish peroxidase (Pierce, Rockford, IL). Western blots were developed with a chemiluminescent substrate (Pierce, Rockford, IL).

Cloning of *ehxCABD* from STEC O91:H21

Clinical isolates of STEC harbor a large plasmid (Levine et al., 1987; Schmidt et al., 1996), the size of which varies among isolates (Brunner et al., 2006). Representative plasmids carried by EHEC O157:H⁻ strain 3072/96 and by STEC O113:H21 strain EH41 have been sequenced (Brunner et al., 2006) (Leyton et al., 2003). The *ehx* operon was cloned from STEC O91:H21 (B2F1) with sequence information available for STEC O113:H21 (EH41) (Leyton et al., 2003). The primer pairs used to generate the ~7kb PCR fragment that contained the entire operon of *ehx* are listed in Table 4. PCR amplification was done with PFU polymerase (Stratagene, La Jolla, CA) and 100 ng of template genomic DNA isolated from B2F1. The

concentration of each primer was 0.5 μ M per 50 μ l reaction. Amplification of *ehx* was conducted under the following conditions: 1 minute at 95°C and 25 cycles of 30 seconds at 95°C, 50 seconds at 47°C, and 8 minutes at 72°C with a final extension for 10 minutes at 72°C. To create plasmid pCR*ehx*, the resultant blunt-end PCR fragment that contained *ehxCABD* was cloned into the *SrfI* restriction site of pPCR-Script Amp SK (+) (Stratagene, La Jolla, CA). In addition to the open-reading frame of *ehx*, this construct contained 126 bp of DNA sequence upstream of the start codon of *ehx*.

Cloning of Partial *ehx* Operon and Construction of pMC*ehxCA'*-*lacZ* Transcriptional Fusion Construct.

A clone of part of the *ehx* operon from strain STEC O91:H21 was generated to contain 126 bp of DNA upstream of the start codon of *ehxC*, the entire coding region of *ehxC*, and the first 695 bp of the coding region of *ehxA*. Primers were designed to generate the partial *ehx* operon flanked by *BamHI* and *EcoRI* restriction enzyme sites (Table 4). PCR with PFU polymerase (Stratagene) generates blunt-end PCR products, so the partial *ehx* PCR product was blunt-end cloned into the *EcoRV* site of pCR-Script Amp SK (+) to create pCR*ehxCA'*. The PCR reaction was done as follows: 1 minute at 95°C and 25 cycles of 30 seconds at 95°C, 50 seconds at 65°C, and 1.5 minutes at 72°C and a final extension for 10 minutes at 72°C.

To create a transcriptional fusion of the putative regulatory sequences upstream of the *ehxCA'* with the promoterless β -galactosidase gene (*lacZ*), the *ehx* fragment was removed from the pCR-Script Amp SK (+) by sequential digestion with *EcoRI* and *BamHI* followed by ligation of the *ehx* fragment into the compatible *EcoRI* and *BamHI*

sites of the pMC1403 vector that contains the promoterless *lacZ* (Casadaban et al., 1980) to generate pMC*ehxCA'*-*lacZ*. In the pMC1403 plasmid, the first eight codons of the amino-terminal end of the *lacZ* gene that code for β -galactosidase has been removed and unique *Bam*HI, *Eco*RI and *Sma*I restriction sites inserted so that DNA fragments with regulatory cis-elements and intact amino-terminal regions of any gene of interest, when ligated into that site on the plasmid, will yield a hybrid protein that retains the enzymatic β -galactosidase function. The pMC1403 vector has been used extensively to identify cis-acting DNA elements required for transcriptional control (Casadaban et al., 1980).

Cloning of *hns* from Strain B2F1

PCR was used to isolate the wild-type *hns* gene from B2F1. The primers used to generate the *hns* gene are listed in Table 4. The primers were designed so that the PCR product of the *hns* gene was flanked by *Eco*RI and *Bam*HI restrictions sites. The *hns* gene was blunt-end cloned into the *Srf*I site of the pCRScript Amp SK (+) cloning vector. Then the *hns* gene was transferred from pCRScript Amp SK (+) by double restriction enzyme digestion with *Eco*RI and *Bam*HI, gel purified and cloned into the *Eco*RI and *Bam*HI cut low copy number cloning vector pMMB66EH to create pMS1001.

Directional Cloning of *hns* into the pBAD202/DTOPO Expression Vector

Primer pairs were designed such that the forward primer contained a four base overhang complementary (CACC) to the expression vector to permit directional insertion

of *hns* into the pBAD expression vector, pBAD202/D-TOPO (Invitrogen, Carlsbad, CA). Plasmid pMS1001 that contained the *hns* coding sequence was used as DNA template for the PCR reaction. The primer pairs are listed in Table 4. The forward primer did not contain the start codon so that a translational fusion with the thioredoxin protein could occur. The stop codon in the reverse primer was replaced with the alanine codon so that the V5 epitope and the 6Histidine-tag (6His-tag) was fused in-frame to *hns*; stop codons are present at the end of the 6His-tag sequence within the vector. PCR amplification was done in a manner similar to that described above except that the extension time at 72°C during amplification cycles was 0.5 minute. Amplification yielded a 426 bp fragment that included the open-reading frame of the gene which codes for HNS. The gene-product was designed to contain the HNS fusion protein with the thioredoxin protein along with the V5 epitope at its amino-terminus and the 6His-tag at its carboxy-end carried on the expression vector designated as *pt-hns/his*. This construct yields an HNS-fusion protein product which is 31.4 kDa (HNS is 15.4 kDa; thioredoxin is 13 kDa and the V5 plus 6His is 3 kDa).

Electrophoretic Mobility Shift Assay (EMSA).

Preparation of the HNS fusion protein- E. coli- LMG194 carries the plasmid *pt-hns/His* that encodes the recombinant clone of HNS described above. Cultures of LMG194 (*pt-hns/His*) were grown overnight with aeration in LB broth with 50 µg/ml of kanamycin. A 100-fold dilution of the overnight bacterial culture was made into LB broth and grown at 37°C with shaking until an optical density at 600 nm of 0.2

was reached. The culture was induced with 0.02% arabinose and allowed to grow for 4 additional hours, and bacteria were pelleted by centrifugation at 13,000 x G for 5 minutes. Cells were resuspended in 500 μ l of lysis buffer that contained 50mM potassium phosphate, 400 mM NaCl, 100mM KCl, 10% glycerol, 0.5% Triton X-100, 10 mM imidazole. Lysis of the bacteria was done by freezing the resuspended cells in an ethanol bath, then thawing the frozen organisms at 42°C. After lysis, the pellet was separated from the supernatant by centrifugation at 13,000 x g for one minute. To determine the amount of protein contained in the sample, two-fold serial dilutions were made and compared to an albumin standard in a bicinchoninic acid (BCA) protein assay (Pierce, Rockford, IL). The HNS fusion protein was extracted from the mixture with Nickel-Chelated Agarose reagent by virtue of the 6His-tag (Pierce, Rockford, IL).

Preparation of the promoter region of ehx and procedures of the

EMSA- Plasmid pCREhx was the source of the *ehx* putative promoter region. Specifically, plasmid pCREhx was digested with *EcoRI* and *PmeI*, a process which yield an 88 bp DNA fragment that contained the putative promoter region. The 88 bp fragment was gel purified and exacted with the Qiagen Gel Extraction Kit (Qiagen, Valenica, CA). Next, 40 μ M of DNA that contained 88 bp of DNA upstream of the *ehxC* start codon was added to samples that contained decreasing concentrations of his-tagged HNS at 0.2 μ M, 0.1 μ M, 0.05 μ M, 0.025 μ M, 0.0122 μ M, and 0.006 μ M in binding buffer to a final volume of 10ul. Binding buffer consisted of 750 mM KCl, 0.5 mM dithiothreitol, 0.5 mM EDTA, 50 mM Tris at pH 7.4. A negative control

sample was included that contained only buffer added to the *ehx* DNA fragment. The reaction mixture was incubated at 25°C or 37°C for 20 minutes and separated by 6% polyacrylamide non-denaturing gel electrophoresis (Invitrogen, Carlsbad, CA). The DNA within the gel was stained with Sybr-Green (Invitrogen, Carlsbad, CA) for 20 minutes and visualized with the Kodak Gel Logic 1500 Imaging System. Experiments were repeated at least three times at both 25°C and 37°C. As a negative control, a 100 bp fragment of DNA from pCRScript Amp SK (+) was reacted with HNS under identical conditions. In addition to the 88 bp fragment, the region of DNA 126 bp upstream of the start codon of *ehx* cloned onto pMC*ehxCA'-lacZ* was also tested by EMSA for HNS binding.

Measurement of β -galactosidase Activity

A microassay was used to quantitate the amount of β -galactosidase enzyme produced in strains that carried the transcriptional fusion construct. The High Sensitivity β -galactosidase Assay kit (Invitrogen) uses the galactosidase analog chlorophenol red- β -D-galactopyranoside (CPRG) which is 10 times more sensitive than the standard O-nitrophenyl- β -D-galactopyranoside (ONPG) substrate under similar assay conditions. The CRPG substrate is extremely stable as it is resistant to proteolytic enzymes typically present in bacterial lysates. The maximum absorbance of the chromophore is 570 nm. Bacteria were grown in LB medium with or without antibiotics at 37°C with aeration and harvested at mid-logarithmic growth. Lysates were prepared and assayed at pH 7.5 and 37°C for 2.5 hours. Optical density readings were done at

570 nm. All assays were performed in triplicate and repeated a minimum of three times as per the manufacturer's instructions. The specific activity of β -galactosidase is expressed as nmol of chlorophenol red formed \times minute⁻¹ \times mg of total protein⁻¹ which is equal to the Units (U) of β -galactosidase activity per mg of total protein.

DNA Sequencing

The PCR fragment cloned into pCREhx was sequenced by Northwoods DNA, Inc., Solway, MN. The PCR fragment was sequenced from both strands of DNA, and each strand of DNA was sequenced twice with 100% agreement between each trial. The PCR fragment cloned into pCB2YghJ was sequenced in the same manner.

Comparative Analysis of *E. coli* Grown on Specific Carbon Sources

Bacteria were grown overnight in M9 minimal media (20mM NH₄Cl, 25mM KH₂PO₄, 40mM Na₂HPO₄ at pH 7.4) supplemented with 0.2% of glucose at 37° C with shaking except where otherwise indicated. The cultures were diluted 1:100 into M9 minimal media supplemented with 0.1% of the indicated carbohydrate and grown for 12 or 24 hours at 37°C. All assays were performed in triplicate and repeated multiple times. Over a 24 hour period, the relative growth of the cultures was determined by measuring the optical density at 600 nm and plotting a standard growth curve of optical density over time.

Cellular Fractionation to Determine Subcellular Location of YghJ

Spheroplast preparation and sonication- Overnight bacterial cultures grown at 37°C were harvested by centrifugation at 2,500 x g and spheroplasts prepared

using a modified procedure based on a method by Osborn et al.(1992). The cell pellet was resuspended in 200mM Tris buffer (pH 8.0) containing 0.2 mM dithithreitol (DTT). An equal volume of 1 M sucrose containing 10 mg/ml lysozyme was added to the cell solution and the mixture incubated on ice for five minutes. Then an equal volume of ice cold distilled water was added to the solution. After incubation on ice for 10 minutes, the cells were lysed by sonication and unbroken cells were removed by centrifugation at 2,400 x g.

Triton X-100 subcellular fractionation - Inner membrane proteins were extracted with triton X-100. Triton X-100 is a nonionic detergent that solubilizes inner membrane proteins leaving outer membrane proteins intact (Filip *et al.* 1973). The spheroplasts were prepared as described above. The pellet was recovered and washed twice in 50 mM Tris containing 0.2 mM DTT. After incubation on ice for 30 minutes, the sample was split into two separate aliquots that included (1) the total membrane extract and (2) soluble proteins from the cytoplasm and periplasm recovered by centrifugation at 160,000 x g for 4 hours. After centrifugation, the triton X-100 inner membrane fraction was extracted from the membrane pellet using 2% triton X-100 in 50 mM Tris (pH 8.0) containing MgCl and 0.2 mM DTT. The membrane fraction containing only the outer membrane proteins was separated by centrifugation at 160,000 x g for one hour This resulted in the triton X-100 soluble inner membrane proteins and the urea soluble outer membrane proteins. The pellet, which contains outer membrane proteins, was solubilized prior to PAGE analysis in 50 mM Tris-HCl (pH 7.4) containing 5 M urea. All subcellular fractions were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis

(SDS-PAGE). The cytoplasmic and non-attached periplasmic proteins, triton x-100 soluble inner membrane proteins, and the urea extracted outer membrane fractions are shown as: (cyto), the combined cytoplasmic and periplasmic soluble proteins; IM; and OM respectively.

Globomycin Inhibition of Signal Sequence Cleavage of YghJ

The antibiotic globomycin specifically inhibits the signal peptidase II enzyme (LspA), which cleaves the nascent prolipoprotein at its diacylglycerol modified cysteine residue (Tokunaga *et al* 1982). The histidine-tagged recombinant YghJ protein expressed from the inducible pET21a expression vector was used in this analysis. Strains were grown to an optical density at 600 nm to 0.600 prior to addition of globomycin. Globomycin was added to the strains above at the following concentrations: 200 mM, 100 mM, and 50 mM. Samples were taken every 2 hours for 6 hours. After treatment with globomycin, cells were harvested by centrifugation and proteins solubilized in 50 mM Tris (pH 7.4) containing 0.1% SDS. The protein concentration of each sample was determined, samples equalized and then analyzed by SDS-PAGE.

Cloning the *yghJ* Gene along with Upstream and Downstream Intervening DNA Sequences: Construction of Plasmid pCRB2yghJ.2 that Contains the Full Transcriptional Unit of *yghJ* from STEC O91:H21 Strain B2F1

Plasmid pCRB2yghJ contains full-length *yghJ* cloned by PCR from STEC O91:H21. In STEC O91:H21, the *yghJ* gene is flanked by intervening DNA sequences; specifically 583 bp of DNA downstream of *glcA* and 276 bp DNA upstream of *pppA*. In the commensal *E. coli* K-12 substrain MG1655, the location of *yghJ* is the same but the

amount of intervening DNA is different. See Figure 1 below which compares the location of *yghJ* in the two strains. Therefore we used primers that code for sequences at the end of *glcA* and fall within *pppA*. Gene releaser (Bioventures, Murfreesboro, TN) was used to extract genomic DNA from both STEC O91:H21 and MG1655 used in the PCR reaction.

Genome sequence data was acquired from accession #U00069 which includes the whole genome sequence of *Escherichia coli* K12- substrain MG1655 (Figure 1). The *yghJ* gene is located at 3,112,572-3,117,134 on the MG1655 chromosome. Comparison of the amino acid sequence of YghJ carried by MG1655 and STEC O91:H21 indicates that the proteins are 99% identical. Further studies to characterize the function of the YghJ protein was done in a *E. coli* K-12 background to avoid inherent problems in handling the STEC O91:H21 pathogen.

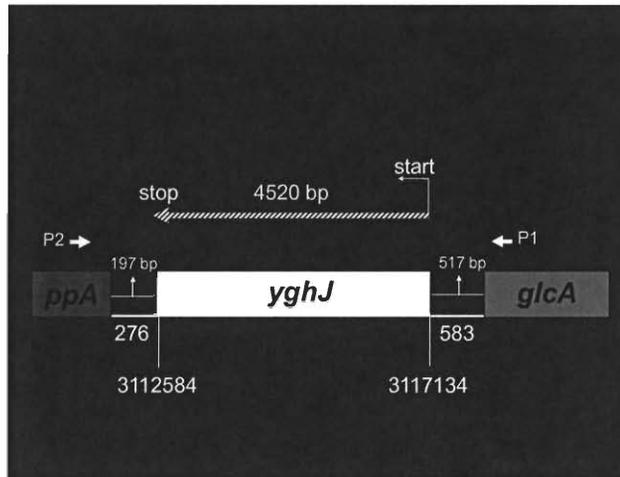


Figure 1. Schematic representation of the position of *yghJ* on the chromosome. The intervening Sequence for STEC O91:H21 is shown in white and MG1655 is shown in gold.

PCR reactions were done with Phusion a high-fidelity DNA polymerase (Finnzymes-New England Biolabs, Ipswich, MA). PCR with Phusion results in a blunt-ended DNA product. PCR with primers CTG-10F and CTG-20R generated the full-length *yghJ* transcriptional unit flanked by *XbaI* and *KpnI* enzymes. The PCR fragment generate from STEC O91:H21 contained 5439 bp of DNA. This fragment of DNA contains the *yghJ* as well as non-coding intervening sequences downstream of *glcA* and upstream of *pppA* genes. The PCR reaction was done by the two-step method as follows: 30s at 98°C and 25 cycles of 30 s at 98°C, 2 minutes at 72 °C with a final extension for 10 minutes at 72 °C. The PCR fragment was blunt-end cloned into the EcoRV site of pPCRScript Amp SK (+) to create pCRB2yghJ.

Figure 2 is a picture of the PCR product and results of screening of putative clones that carry the plasmid of interest. The DNA fragments recovered after PCR from

B2F1 and MG1655 were of the expected size. The *yghJ* PCR product from DH5 α is much smaller but correlates with the reported size of the *yghJ* gene-fragment carried by this strain of *E. coli*.

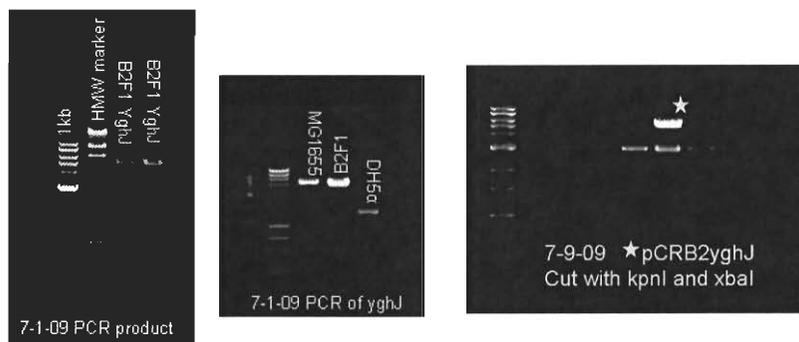


Figure 2. PCR product and results of screening of putative pCB2YghJ clones. Panel A is PCR product from STEC O91:H21 strain B2F1, Panel B is a comparison PCR product using the same primer set for genomic DNA of MG1655, Panel C is a restriction analysis of pCB2YghJ clones cut with *KpnI* and *XbaI*.

The ligation mixture was transformed into DH5 α and selected on LB agar containing ampicillin. Clones surviving selection were tested for the presence of the plasmid using the Qiagen Plasmid Extraction Kit (Qiagen, Valenica, CA). Isolated plasmids were cut with *XbaI* and *KpnI* to determine if an insert was present. Plasmid pCB2YghJ, shown in panel C above contains the insert. This plasmid was sequenced by overlapping primer design. This construct contains the complete *yghJ* transcriptional unit as well as DNA sequences that flank the *yghJ* gene.

Cloning of *yghJ* from *E. coli* MG1655 and Generation of the Histidine-Tagged Recombinant YghJ

Cloning the open-reading-frame (ORF) of yghJ- PCR reaction was done using genomic DNA prepared from *E. coli* MG1655. Primer pairs YghJ-5FA and YghJ-6RA were used to create a modified blunt-end DNA product containing the protein-coding DNA sequence of *yghJ* (ORF-*yghJ*) flanked by *Bam*HI and *Eco*RI restriction sites. The PCR reaction was done by the two-step method as follows: 30s at 98°C and 25 cycles of 30 s at 98°C, 2 minutes at 72 °C with a final extension for 10 minutes at 72 °C.

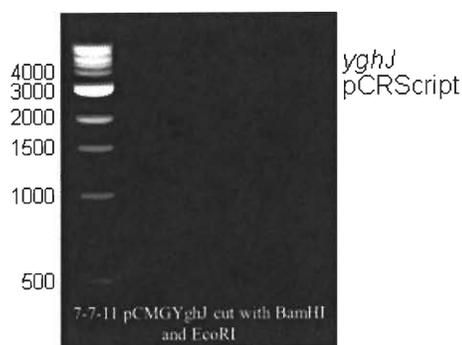


Figure 3. Restriction analysis of pCMGORFyghJ clones.

The modified *yghJ* sequence obtained contained the translational start codon but without a stop codon. The modified *yghJ* ORF sequence was cloned by blunt-end ligation into the *Srf*I restriction site of pPCRScript Amp SK (+) to create pCMGYghJ. The pPCRScript Amp SK (+) plasmid provides the stop codon for the cloned *yghJ* fragment.

Clones were screened by restriction analysis with *Bam*HI and *Eco*RI enzymes. Clones 3, 7 and 9 all contained *yghJ* DNA of the correct size, shown in Figure 3. Clone

pCRMGyghJ.7 was used for cloning into pET21a to prepare his-tagged YghJ recombinant protein.

Preparation of recombinant his-tagged YghJ protein-Plasmid pCRMGyghJ.7 was cut sequentially. First the plasmid was cut with *EcoRI* which liberated a 4560 bp DNA fragment that contained *yghJ*. This fragment was gel extracted then cut with *BamHI*. The resulting *yghJ* DNA fragment flanked by *BamHI* and *EcoRI* was cloned inframe into *BamHI* and *EcoRI* restricted sites of the T7 expression vector pET21a; creating plasmid pETORFMGyghJ. Plasmids were screened by PCR analysis using forward primer YghJ-5FA and reverse primer YghJ-6RA. Clones 1, 2, 3, and 5 contained the *yghJ* ORF. These clones were selected and frozen. The ORF of *yghJ* was moved from pPCRScript Amp SK (+) into the pET21a vector.

Directional Cloning of *yghJ* onto pBAD TOPO Expression Vector to Create a YghJ Fusion Protein Containing a N-terminal Thioredoxin and C-terminal Polyhistidine Tag

This construct was generated in an attempt to produce a thioredoxin-YghJ fusion protein that would remain cytoplasmic. The hope was that this construct would be stable enough because of the thioredoxin tag and could be purified by metal chelating resins. The construct was designed so that the thioredoxin protein could be cleaved after purification releasing an intact YghJ protein without the thioredoxin fusion partner. This was done because the YghJ-his-tagged recombinant proteins cloned from both STEC O91:H21 and MG1655 and expressed from the pET21a vector did not stick to any of the metal chelating resins we used for purification. **NOTE:** Sequencing across the joint

between the *yghJ* gene and the histidine codons indicated that the *yghJ* gene is inframe. However, Western blot analysis of the recombinant Ygh-his-tag fusion protein indicated that the histidine tag, if present, could not be detected by this method.

To create a thioredoxin-YghJ fusion protein, the ORF of *yghJ* was acquired by PCR from *E. coli* MG1655. The blunt-end DNA fragment was the expected size at 4560 bp. The DNA was cloned inframe with the thioredoxin gene by inclusion of four nucleotides, CACC, in the forward primer sequence. Directional cloning was obtained by CACC base pairing of the PCR product with the GTCC sequence in the pBAD202/D-TOPO vector. Blunt-end PCR reaction was done with PFU ultra II Fusion HS DNA polymerase (Agilent Technologies, Santa Clara, CA). The reaction procedure included: Denaturing step at 95°C for 5 minutes, followed by 25 cycles of 95°C for 1 minute, then 30 seconds at 55°C, extension for 2.25 minutes at 72°C; ending with 10 minutes at 72°C.

Five clones were selected, analyzed and frozen prior to protein expression studies. The *Hind*III restriction enzyme occurs once in the expression vector and does not occur within the *yghJ* gene, so restriction with *Hind*III should linearize the plasmid. *Hind*III restriction resulted in one band of the expected size of 9010 bp.

Induction of the Topo-YghJ-6X His fusion protein from the pBAD202/D-TOPO plasmid

Cultures of *E. coli* LMG194 (PbadMGYghJ) were grown overnight with aeration in LB broth with 50 µg/ml of kanamycin. A 100-fold dilution of the overnight bacterial culture was made into LB broth and grown at 37°C with shaking until an OD_{600 nm} of 0.2 was reached. The culture was induced with varying concentrations of arabinose and

allowed to grow for 8 additional hours. Bacteria were pelleted by centrifugation at 6,000 x G for 5 minutes. Samples were lysed by resuspension in SDS loading buffer. Proteins were separated by SDS-PAGE in a 4-12% Bis-Tris gel. Proteins were visualized using Coomassie Brilliant Blue R-250.

Induction and Purification of YghJ-6X His Recombinant Protein Using Clontech HisTalon Gravity Columns

Bacterial cultures were grown overnight in LB media containing 100 µg/ml ampicillin then diluted 1:100 into LB media with antibiotic. Diluted cultures were grown to an OD_{600 nm} of 0.600 and cultures were induced for expression of *yghJ* by addition of 100 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). Following addition of IPTG, cultures were grown for 6 hours to produce YghJ. Cells were harvested by centrifugation for 15 minutes at 2,400 x g. Bacteria were lysed using the Talon xTractor Buffer (Clontech, CA) containing 1 mg/ml lysozyme and 1 µg/ml DNase I and EDTA free protease inhibitor cocktail (Sigma-Aldrich) according to manufacturer's specifications. The HisTalon gravity metal affinity column (Clontech, CA) was used for protein purification. Briefly, the sample was loaded onto an equilibrated column and washed with 3x column volumes of wash buffer (50 mM sodium phosphate, 300 mM NaCl) before eluting with buffer containing 400 mM imidazole at pH 7.4. A second attempt was made at purification using the same conditions with buffers containing 0.1% Triton X-100. Protein levels were quantified and equalized before SDS-PAGE analysis.

Induction and Purification of YghJ-6X His using Pierce B-PER 6XHis Fusion Protein Purification Kit

Bacterial cultures were grown overnight in LB media containing as described above.

Diluted cultures were grown to an OD₆₀₀ nm of 0.600 and cultures were induced for expression of *yghJ* by addition of 100 mM IPTG. Following addition of IPTG, cultures were grown for 6 hours for YghJ production. Cells were harvested by centrifugation for 15 minutes at 2,400 x g. Cultures were lysed using the Pierce Bacterial Protein Extract Reagent containing EDTA free protease inhibitor cocktail (Sigma –Aldrich, CA) overnight at 4° C. The lysate was incubated with nickel-chelate agarose for one hour. The flowthrough was removed by centrifugation and the nickel-chelate agarose was washed three times with the wash buffer supplied with the kit. The protein was eluted as per manufacturer's instructions using the included elution buffer. Protein levels were quantified and equalized before SDS-PAGE analysis. In some cases, β-mercaptoethanol was included as an antioxidant.

Purification Attempt of YghJ-6X His using Invitrogen Dynabeads His-Tag Isolation Beads

Bacterial cultures were grown overnight in LB media as described above.

Expression of *yghJ* induced by addition of 100 mM IPTG as described above and cultures were grown for 6 hours for YghJ production. Cells were harvested by centrifugation for 15 minutes at 2,400 x g. Bacteria were lysed using the Talon xTractor Buffer (Clontech, CA) containing 1 mg/ml lysozyme and 1 ug/ml DNase I and EDTA free protease inhibitor cocktail (Sigma- Aldrich) according to manufacturer's specifications. Membrane debris was removed by centrifugation at 20,000 rpm for 30 minutes. The cleared lysate

was incubated with 2 mg of dynabeads with rocking at 4° C for one hour. The dynabeads were immobilized using a magnet and the lysate was removed. Beads were then washed three times in a 50 mM sodium phosphate buffer. Proteins were eluted using a 50 mM sodium phosphate buffer containing 200 mM imidazole. The purification was analyzed by SDS-PAGE gel electrophoresis.

Western Blotting of YghJ-6X His

A Western blot was done to detect the recombinant protein. Cultures of BL21(DE3) containing both plasmids pETORFB2YghJ-His or pETORFMGYghJ were induced. Lysates prepared and proteins separated on an 8% polyacrylamide gel. Proteins were electrophoretically transferred to a nitrocellulose membrane post SDS-PAGE and blocked overnight in 5% milk. The membrane was incubated with the primary antibodies separately, either India His-HRP (Pierce, Rockford, IL) at 400 ug/ml in PBS and Covance anti-His antibody (Covance, Princeton, NJ.) at 100 ug/ml in PBS. The Covance anti-His was then treated with a goat anti-mouse HRP conjugated antibody (Pierce, Rockford, IL). Western blots were developed with a chemiluminescent substrate (Pierce, Rockford, IL) and the results compared with the Kodak Gel-Logic 1500 Imaging System.

In Gel Assay to Detect Phosphatase Activity of YghJ

The fluorogenic substrate 4-methylumbelliferyl phosphate (MUP) can be used to visualize phosphatase activity from proteins present in a polyacrylamide gel. Dephosphorylation of the substrate causes a fluorescent band to form where phosphatase enzyme localizes following separation in the PAGE gel. Total lysate or extracted

membrane protein from *E. coli* K12 containing *yghJ* expressed from a high copy number plasmid was separated into a denaturing 4-12% Bis-Tris gel. Post-run, the gel was washed twice with 20% isopropanol in 50 mM Tris-HCl (pH 7.5) for one hour per wash. A further step was done in denaturation buffer containing 50 mM Tris-HCl (pH 7.5), 5 mM β mercaptoethanol, and 8 M urea for one hour at room temperature. The proteins were renatured in a buffer containing 50 mM Tris-HCl, 0.02% Tween, 20 mM β mercaptoethanol and 1 mM $MgCl_2$ for 16-20 hours. After renaturation, the gel was incubated with the substrate containing 50 mM Tris-HCl (pH 8.0) 0.1 mM EGTA, 0.01% Tween, 2 mM DTT, 20 mM $MnCl_2$, and 0.5 mM MUP. The gel was for 15 minutes before observing fluorescent bands.

Table 1: Virulence Factors Produced by LEE-Positive STEC as compared to Virulence Factors Produced by LEE negative STEC

Protein	LEE positive	LEE negative	
	O157:H7	O91:H21	O104:H4
HNS	+	+	+
GrlA	+	-	-
GrlR	+	-	-
Ler	+	-	-
Enterohemolysin	+	+	+
Shiga toxin	+	+	+
intimin	+	-	-
Tir	+	-	-
Type III secretion	+	-	+
Type II secretion	+	-	+
Plasmid encoded toxin (Pet)	-	-	+

Table 2. List of strains used in this study

<i>E. coli</i> Strain	Relevant Features	Reference
DH10B	Cloning strain, non-hemolytic	(Grant et al., 1990)
DH5 α	K-12 cloning strain, non-hemolytic	(Silhavy and Beckwith, 1983)
WAF100 (pSF4000)	Strain HB101 with α - <i>hly</i> CABD	(Welch and Falkow, 1984)
DH10B (pB2F1)	Plasmid pB2F1 [encodes <i>ehx</i>]	This study
DH10B (pCR <i>ehx</i>)	<i>ehx</i> operon from B2F1 cloned into pCRScript	This study
TOP10	<i>araD139</i> Δ (<i>ara-leu</i>)7697	Invitrogen
LMG194	deleted for <i>araBADC</i>	Invitrogen
STEC O91:H21 str. B2F1	<i>eae</i> negative, <i>ehx</i> ⁺ , <i>hns</i> ⁺	(Karmali et al., 1986)
STEC O91:H21 str. S11	plasmid-cured derivative of B2F1, <i>ehx</i> ⁻ , <i>hns</i> ⁺	(Scott et al., 2003)
STEC O91:H21 str. 34.7	<i>hns</i> ::mini-Tn5phoACmr, <i>ehx</i> ⁺ , hyper-hemolytic	(Scott et al., 2003)
THK62	<i>hns</i> null mutant	(Donato et al., 1997)
BL21(DE3)	T7 promoter	(Studier and Moffatt, 1986)
K-12 MG1655	Wild type YghJ	(Blattner, 1997)
B strain BW25113	Wild type YghJ	(Datsenko and Wanner, 2006)
STEC O157:H7	LEE positive, YghJ negative	(Karmali 1985)

Table 3. List of plasmids used in this study

Plasmid	Relevant Features	Reference
pBAD202/D-TOPO	Expression cloning vector	Invitrogen
pCRScript Amp SK(+)	Cloning vector, <i>bla</i> ^r	Stratagene
pSF4000	pACYC184 vector with α - <i>hly</i> CABD	(Welch and Falkow, 1984)
pMMB66EH	Low copy IPTG inducible vector	(Furste et al., 1986)
pMS1001	<i>hns</i> from STEC O91:H21 str. B2F1 carried on pMMB66EH	(Scott, 2003)
pCRehx	<i>ehx</i> carried on pCRScript Amp SK(+)	This study
pT- <i>hns</i> /His	<i>hns</i> fused with his-tag and thioredoxin on pBAD202/D-TOPO	This study
pMC1403	promoterless <i>lacZ</i> vector for transcriptional fusions	(Casadaban et al., 1980)
pCRehxCA'	upstream regulatory sequences, <i>ehxC</i> and part of <i>ehxA</i> on pCRScript Amp SK(+)	This study
pMC <i>ehxCA'</i> - <i>lacZ</i>	promoter region, complete <i>ehxC</i> gene partial <i>ehxA</i> fused with <i>lacZ</i> gene on pMC1403	This study
pET21a	IPTG inducible vector	Novagen
pCB2 <i>yghJ</i>	<i>yghJ</i> from B2F1 including upstream and downstream sequence on pCRScript Amp SK(+)	This study
pCORFB2 <i>yghJ</i>	open reading frame of <i>yghJ</i> from B2F1 on pCRScript Amp SK(+)	This study
pCORFMG <i>yghJ</i>	open reading frame of <i>yghJ</i> from MG1655 on pCRScript Amp SK(+)	This study
pETB2 <i>yghJ</i>	<i>yghJ</i> from B2F1 including upstream and downstream sequence on pET21a	This study
pETORFB2 <i>yghJ</i>	open reading frame of <i>yghJ</i> from B2F1 on pET21a	This study
pETORFMG <i>yghJ</i>	open reading frame of <i>yghJ</i> from MG1655 on pET21a	This study
pBADMG <i>yghJ</i>	open reading frame of <i>yghJ</i> from MG1655 on pBAD202/D-Topo	This study

Table 4. List of primers used in this study

Gene	Primer Sequence
<i>ehx</i> operon	EhxF:5'- TACGTATAGATGCTTCTTGCTTAAAAGGATATAAT-3'
	EhxR:5'- CCTAGGTATACTAACGTTACGTAAACT-3'
partial <i>ehxCA</i>	EhxF:5'-GAATTCTTGGGAGGGATTCTCCTTGCGGTAGC-3'
	EhxR:5'-GGATCCCCGGAGACAACATCCAGCCCAT-3'
<i>hns</i>	HNSF:5'-AGAATTCATGAGCGAAGCACTTAAAATTCT-3'
	HNSR:5'-AGGATCCTTATTGCTTGATCAGGAAATC-3'
<i>hns</i> -his	HNS1F: 5'-ACCAGCGAAGCACTTAAAATTCTGAACAACATCCGTA-3'
	HNS2R: 5'-GAAATTGCTTGATCAGGAAATCGTCGAGGGATTTACC-3'
pCB2YghJ	CTG-10F 5' TCTAGACATGGTGGGCGGAGAATCTGAAC-3'
	CTG-20R 5'GGTACCACATTTACCGTGAAACTCC-3'
ORFMGYghJ & ORFB2YghJ	YGHJ-5FA 5'-CGGGATCCTCACTTGCGTTATTAATGAATAAGAAATTT-3'
	YGHJ-6RA 5'TAGAATTCAAACTCGGCAGACATCTTATGCTC-3'
YghJ-Topo	YghJ-BAD1F 5'-CACCATGAATAAGAAATTTAAATATAAGAAATCGCTTT-3'

RESULTS

Proteomic Analysis and Comparison of Shiga Toxin-Producing *Escherichia coli* (STEC) O91:H21 and STEC O157:H7 Isolates

Comparative Two-Dimensional Gel Analysis of Extracytoplasmic Proteins from LEE-negative versus LEE-positive STEC-

Extracellular proteins were recovered from two STEC isolates: (1) STEC O91:H21 a LEE-negative strain and (2) STEC O157:H7 the LEE-positive isolate. The protein profiles were examined by SDS-PAGE gel. Then separated by molecular weight in a one-dimensional gel and visualized by silver staining (Figure 4). After separation, a high molecular weight protein of approximately 160 kDa was visible in the STEC O91:H21 isolates but was not present in the sample from STEC O157:H7. In addition HNS did not appear to inhibit expression of the 160 kDa protein because it is present in the STEC O91:H21 HNS knockout mutant designated as 34.7. Furthermore, the novel gene-product is not carried on the large virulence plasmid but is coded for on the chromosome since the plasmid-cured variant of STEC O91:H21 called S11 still produced the protein. However, there appears to be lower levels of the protein in S11 which suggests but does not prove that gene-products encoded on the large virulence plasmid may play a role in its expression. See Figure 4 which is representative of the data discussed above.

To ensure that proteins detected were distinct, two-dimensional gel electrophoresis was used. A representative gel is shown in Figure 5. The protein profiles of STEC O91:H21 as compared to O157:H7 showed several differences. Arrow one

indicate the 160 kDa protein of interest. Other differences in the protein profile are indicated by numbered arrows.

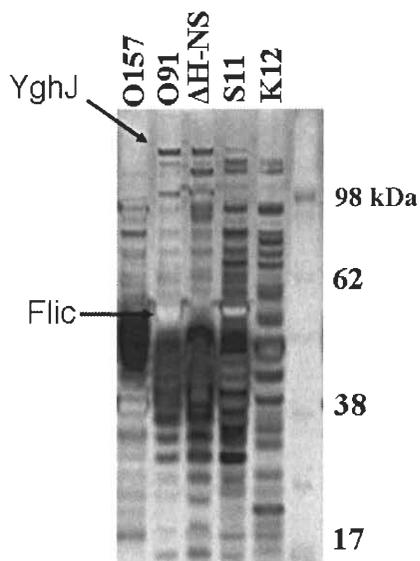


Figure 4. Two-Dimensional Gel Analysis of the Proteomes of LEE-negative STEC O91:H21 and LEE positive STEC O157:H7. YghJ is indicated by an arrow. The flagellin protein FliC, expressed at high levels in wildtype O91:H21 and not expressed in the HNS mutant indicated by an arrow.

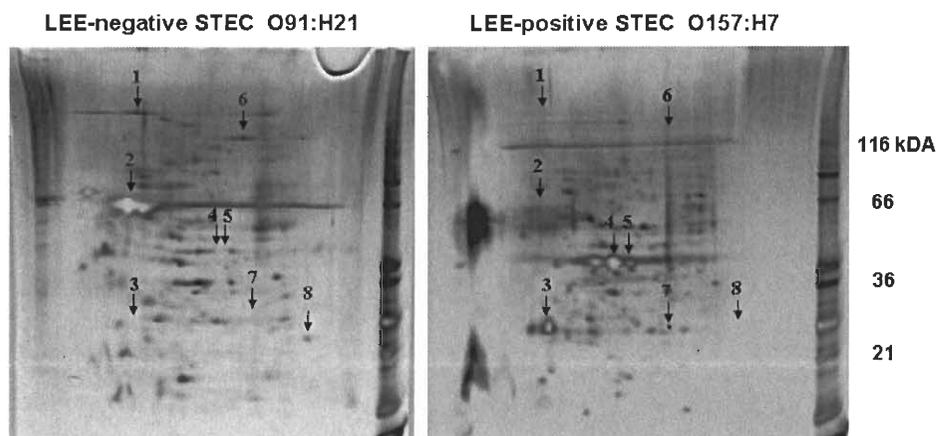


Figure 5. Two-Dimensional Gel Analysis of the Proteomes of LEE-negative STEC O91:H21 and LEE positive STEC O157:H7. Red numbers indicate proteins reduced or are not present in STEC O157:H7 and Blue numbers indicate proteins reduced or not present in STEC O91:H21.

In Figure 6 STEC O91:H21 two-dimensional gels are shown. In panel A a portion of a coomassie stained gel is shown; an arrow marks the location of the protein of interest which was cut out and sent for mass spectrometry to identify the protein (Chad Trumble Master's Thesis 2009). Panel B is the silver stained 2-D gel used to confirm the presence and location of the 160 kDa protein.

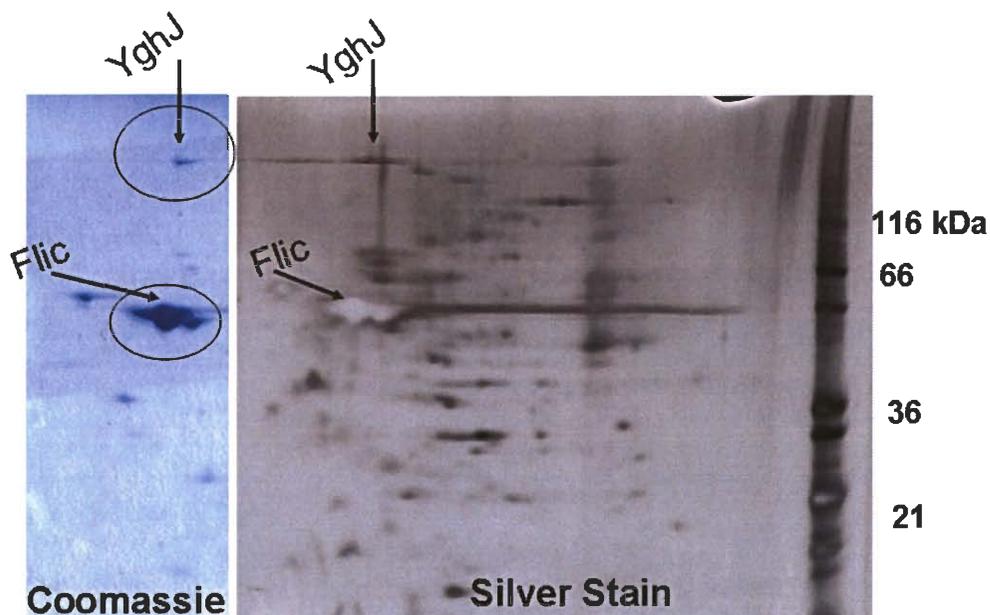


Figure 6. Two-dimensional gel analysis of STEC O91:H21. YghJ and the flagellar protein FliC are indicated with arrows. Bands of interest were excised from the coomassie stained gel and sent for LC/MS/MS^E to identify the proteins. The silver stain was used for protein visualization.

Genetic Analysis of the 160 kDa Protein Identified as YghJ: a Novel Extracytoplasmic Protein Produced by STEC O91:H21

Location of the yghJ gene on the chromosome of STEC O91:H21 and E. coli K12 substrain MG1655-

LC/MS/MSE analysis revealed the 160 kDa protein to be an uncharacterized protein called YghJ (Chad Trumble Master's Thesis, 2009). Colony blot analysis of the wild-type O91:H21 and its plasmid cured variant S11 confirmed the presence of yghJ in both strains, indicating that *yghJ* is present on the chromosome (Chad Trumble Master's Thesis, 2009). The genome sequence is unpublished for STEC O91:H21, so the commensal *E. coli* K12 substrain MG1655 (Blattner et al., 1997) was used for further genetic analysis. Accession# U00096 describes the genome sequence of MG1655. The *yghJ* gene is located on the chromosome 197 bp upstream of the pre-pillin peptidase A (*pppA*) gene and 517 bp downstream of the glycolate permease A (*glcA*) gene. See Figure 1 on page 7, a schematic that shows the location of the *yghJ* gene on the chromosomes of MG1655 and STEC O91:H21.

We amplified the *yghJ* gene from STEC O91:H21 and MG1655 as described above using primers which were complementary to the 3'-end of the *glcA* gene and the 5'-end at the beginning of the *pppA* gene, indicating that *yghJ* is located between the *glcA* and *pppA* in both MG1655 and STEC O91:H21. However, the amount of intervening DNA nucleotides varies between the two *E. coli* strains. Specifically STEC O91:H21 contains 79 more DNA base pairs than does MG1655 in the intergenic space between *pppA* and *yghJ*, and 66 more base pairs between the *glcA* gene and the *yghJ* gene.

However, the amino acid sequence of the protein itself shares a 99% identity between the two isolates. Therefore, to prevent any risks associated with the use of pathogens, MG1655 was used throughout this study.

Phylogenetic analysis of YghJ-

Similar YghJ proteins found in other Gram-negative bacterial species

The Basic Local Alignment Search Tool (BLAST) was used to search for proteins similar to YghJ in other bacterial species with sequence data from the *E. coli* K12 strain MG1655 (Accession YP_026189). Results of this comparison are shown in Figure 7. YghJ is related to the AcfD protein of *Vibrio cholerae*, which is known to play a role in colonization of the large intestine in the infant mouse model (Peterson and Mekalanos, 1988). The AcfD proteins from *V. cholerae* appears to cluster when using the nearest neighbor joining method, indicating a close relationship between these proteins; the proteins share a 67% identity with the YghJ protein from *E. coli* MG1655. AcfD proteins from other *Vibrio* species which share a slightly lower identity to YghJ are also found in other *Vibrio* species including *V. vulnificus*, a marine bacteria, *V. parahaemolyticus* and *V. mimicus*, pathogenic isolates which cause gastroenteritis but do not produce cholera toxin (Chowdhury et al., 1987). *V. harveyi* and *V. fischeri*, two bioluminescent bacteria often associated with marine animals also carry the *acfD*. Other bacterial isolates encoding *yghJ*-like genes include *Grimontia hollisae*, a pathogen formerly classified as *Vibrio hollisae* that causes gastroenteritis, *Photobacterium damsela*, known to cause infection and illness and death in fish, and two *Shewanella* isolates, *S. halifaxensis* and *S. pealeana*, both marine bacteria carry homologues of YghJ. Interestingly, all of the genera

that contain the *yghJ* gene have two distinct lifestyles, surviving in such disparate environments as a salty marine environment and the intestines of warm-blooded mammals.

YghJ is distributed throughout both commensal and pathogenic *E. coli* strains Figure 8 shows the Phylogenetic relationships of YghJ protein sequences from *Escherichia* organisms based on BLAST analysis. Numbers shown at nodes represent values based on 100 bootstrap replicates. Bootstrap replicates indicate the probability of the same species being placed into the same node in a set of 100 predicted phylogenetic trees. YghJ is present in several *Escherichia* species, including *E. albertii*, a strain associated with diarrheal disease (Huys et al., 2003) as well as *E. fergusonii*, a species which is often associated with wound infection and bacteremia. YghJ is also present in the pathogenic *Shigella* sp. D9. Among the *E. coli* species, there appear to be three distinct subtree groupings. However, no immediate relationship between species in the three groups can be identified.

YghJ appears in pathogenic strains including uropathogenic *E. coli* (UPEC), enterotoxigenic *E. coli* (ETEC), enteropathogenic *E. coli* (EPEC), enteroinvasive *E. coli* (EIEC), enteroaggregative *E. coli* (EAEC) and extraintestinal *E. coli* (ExPEC) isolates. YghJ proteins from seropathogenic strains do not appear to bear any significant differences in the proteins produced relative to YghJ proteins produced in nonpathogenic or commensal *E. coli*. Thus, there are no identifiable phylogenetic pathogenic groupings detectable. All bootstrap values are above 98, indicating a high relationship between sequenced YghJ proteins in *Escherichia* species and high similarity of proteins ranging

from 87% identity in *E. fergusonii* to between 98% to 100% identity among most *E. coli* isolates.

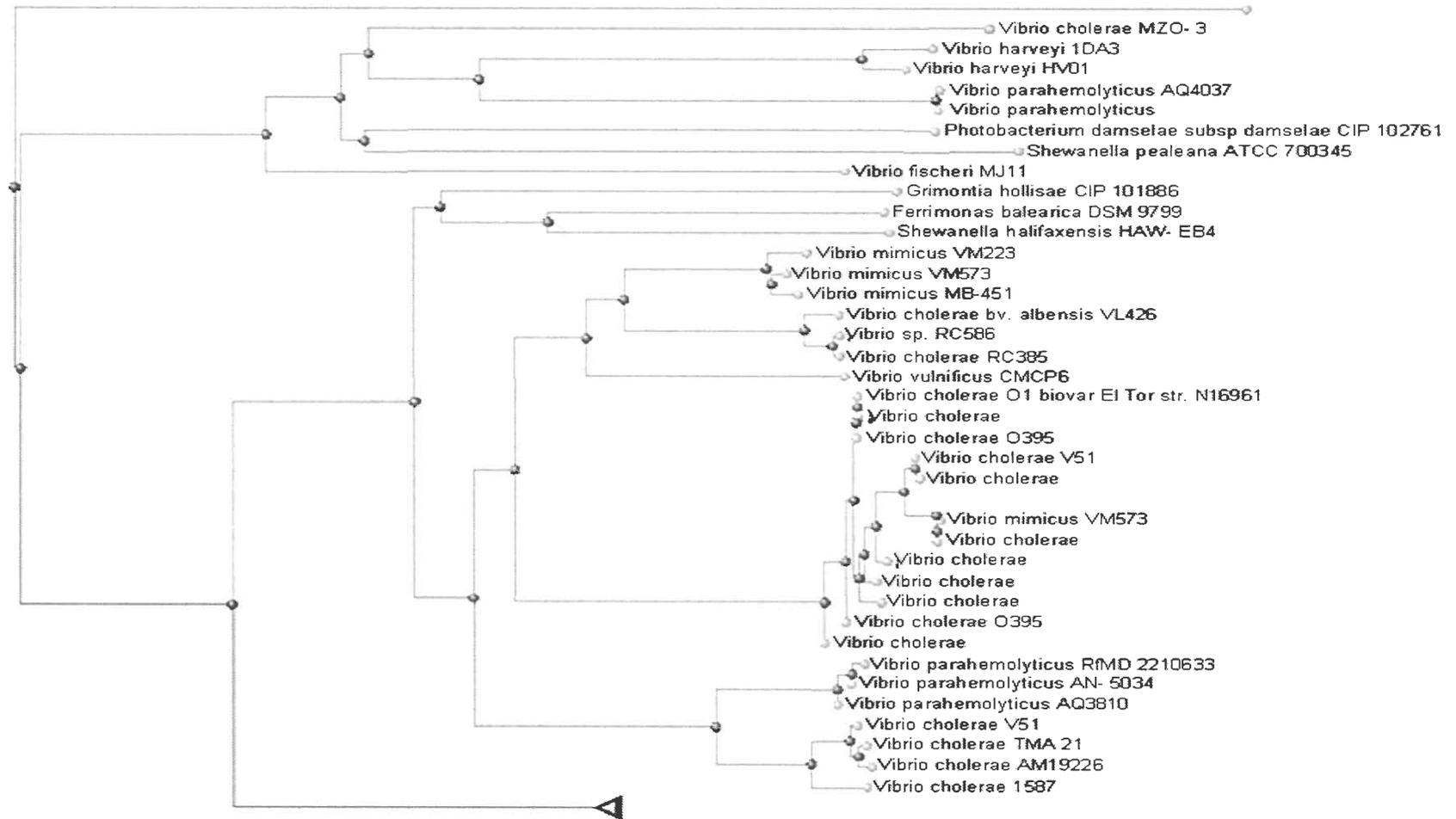


Figure 7. Phylogenetic relationships of YghJ protein sequences produced by other Gram negative species based on neighbor joining analysis.

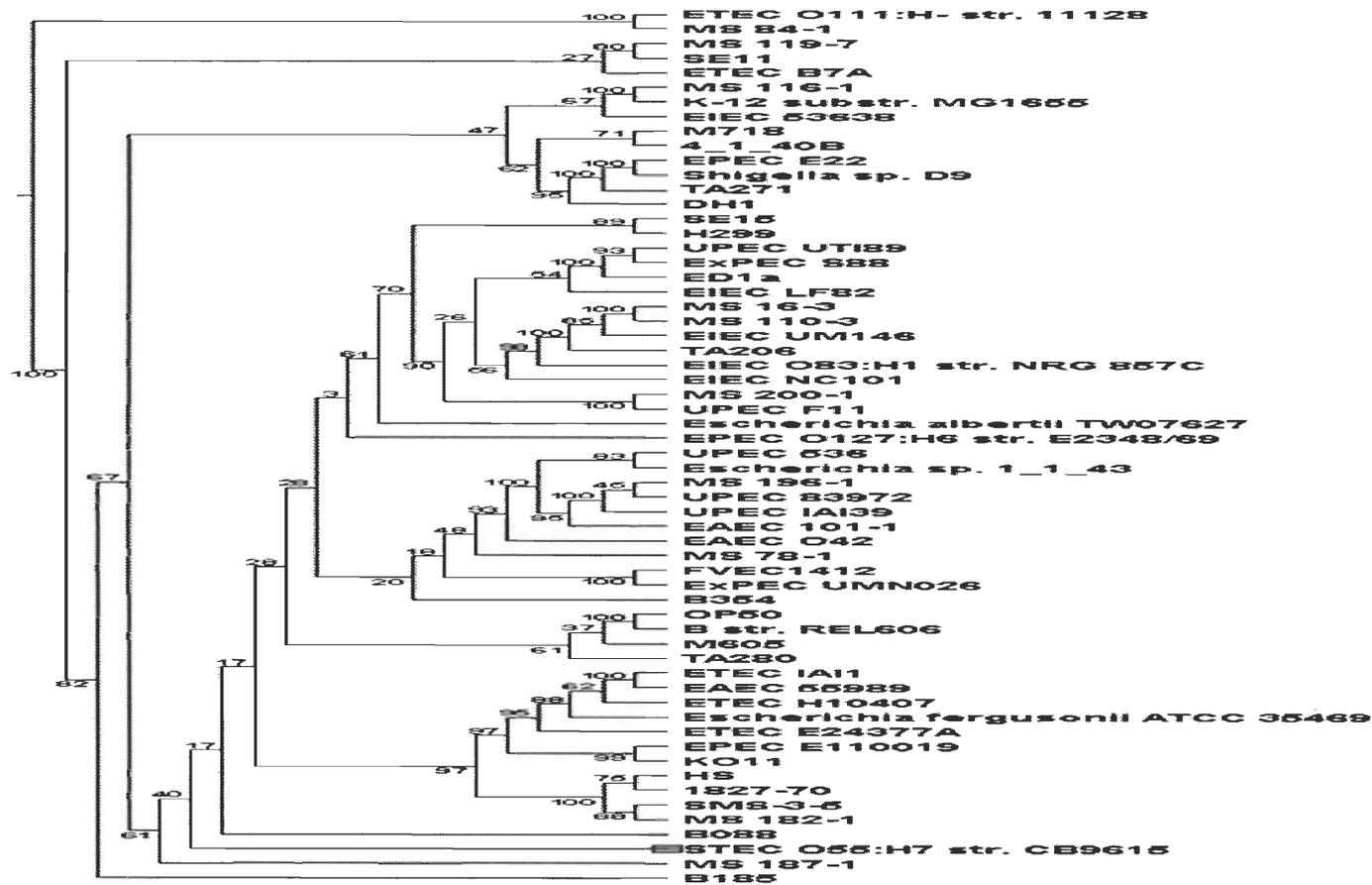


Figure 8. Phylogenetic relationships of YghJ protein sequences of *Escherichia* species based on maximum likelihood analysis. Numbers shown at nodes represent values based on 100 bootstrap replicates. Pathogenic strains are indicated. PEC: Uropathogenic *E. coli*. EPEC: Enteropathogenic *E. coli*. ETEC: Enterotoxigenic *E. coli*. EIEC: Enteroinvasive *E. coli*. ExPEC: Extraintestinal pathogenic *E. coli*. STEC: Shiga toxin-producing *E. coli*

Structure/Function Analysis of the Novel Protein, YghJ.

Structural and Functional Domains of YghJ-Protein alignment analysis using The Basic Local Alignment Search Tool (BLAST) and the Simple Modular Architecture Research Tool (SMART) revealed that YghJ contains a lipoprotein signal sequence which spans residues 1-26. The amino acid sequence is [MNKKFKYKKSLLAAILSATLLLAGCDG] Sutcliffe and Russell (1995). This sequence contains the lipoprotein processing consensus motif, or lipobox, LAG/C¹D²G³. The YghJ protein also contains a DUF4092 (Pfam13322) domain of unknown function from amino acid 416 to 585 (Figure 9). This domain also appears in other proteobacteria such as *V. cholerae* and *V. parahemolyticus* AcfD proteins. Proteins with this domain may be involved with metabolic pathways including nucleotide metabolism.

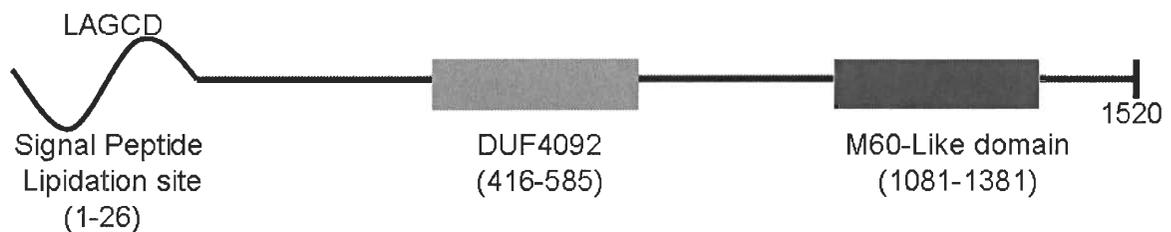


Figure 9. Schematic diagram of the novel YghJ protein. Predicted functional domains are indicated. The curved line indicates the putative lipobox. The amino acid sequence LAGCD is an LOL pathway avoidance signal contained in the lipobox. DUF4092 is a domain of unknown function. The M60-Like domain is a putative metallopeptidase domain with mucinase activity. Figure is not drawn to scale.

Blast analysis also indicates a M60-like domain (Pfam13402) from amino acid 1081 to 1381 (alignment shown in Figure 11). This domain is related to the Enhancin peptidase family, which was originally identified from the invertebrate insect pathogen

Baculovirus. Baculovirus enhancin proteins possess mucin degrading activity both *in vivo* and *in vitro*. The mucin degrading activity of this protein enhances *Baculovirus* infection of Lepidopteran insects (Wang and Granados, 1997b). This domain contains two HEXXH zinc binding motifs and has metallopeptidase activity (Wang and Granados, 1997a). The HEXXH motif forms an alpha helix in which the two histidine residues coordinate a zinc ion in the active site of the enzyme (Fukasawa et al., 1999). Recently the M60-like containing protein BT4244 was purified from the human gut mutualist *Bacteroides thetaiotaomicron* and shown to have mucin degrading activity (Nakjang et al., 2012).

Figure 10. Multiple sequence alignment of the DUF4092 domain of YghJ with similar proteins. Alignment was performed using ClustlW 18.1 (Thompson *et al.* 1994) using *Escherichia coli* YghJ as a prototype. Identical residues are shown with a red background, while similar residues are indicated by a red letter. Conserved motifs are surrounded by a blue box. Species names and accession numbers are as follows: *Escherichia coli* YP_026189.1, *Escherichia coli* Q707L9, *Vibrio fischeri* Q5E705, *Vibrio vulnificus* Q8DE36, *Vibrio mimicus* ZP_05718328.1, *Vibrio cholerae* ZP_01949281.1, *Vibrio parahemolyticus* ZP_01988669.1, *Shewanella halifaxensis* YP_001675698.1.

When treated with EDTA, the capacity of the protein to degrade mucin was reduced. This indicates that the metal binding domain is necessary for protein function. Using the CLUSTALW 18.1 multi-alignment software YghJ was aligned with several M60-like metallopeptidases, including BT4244 (Figure 8). The YghJ protein has the conserved [HEVGH] domain at amino acid 1,303 through 1,307.

Putative Protein phosphatase 2C Domain- Amino acid sequence analysis of the YghJ protein using the Simple Modular Architecture Research Tool (SMART) protein database tools (Schultz et al., 2000) revealed the presence of a putative serine threonine phosphatase domain which is related to the protein phosphatase type 2C (PP2C) superfamily (Figure 12). Protein phosphatase 2C type enzymes remove inorganic phosphate groups from phosphorylated proteins, typically from serine or threonine residues. The PP2C family, also known as the PPM family, is comprised of Mg^{2+} or Mn^{2+} dependent protein phosphatase class of enzymes (Casadaban et al., 1980). The PP2C enzymes are exemplified by the human PP2C protein that possesses 11 distinct functional motifs that allow coordination of a divalent metal cation such as Mg^{2+}/Mn^{2+} crucial for the removal of the phosphate group from proteins phosphorylated on serine or threonine residue (Das et al., 1996). This enzyme family is distributed across all kingdoms of life, with members found in eukaryotes, prokaryotes, and archaea. Despite the divergence between members of this family, several conserved motifs are recognized in PP2C proteins.

1 10 20 30 40 50 60 70
 Vibcho_/1087-1392 GNRQRF¹GG¹⁰AKVAQOEF²⁰SVAS³⁰GGEDS⁴⁰FV⁵⁰IT⁶⁰IALADDL⁷⁰TGREKHE⁸⁰LOLKA⁹⁰FFPRM¹⁰⁰SKSE¹¹⁰VIGG¹²⁰ENH¹³⁰FTSK¹⁴⁰T
 Bac¹⁵⁰thu_/78-367 SPHEE¹⁶⁰GG¹⁷⁰IXAK¹⁸⁰PFDEE¹⁹⁰IVVE²⁰⁰IKG²¹⁰SGS²²⁰..IK²³⁰AEI²⁴⁰OTRAY²⁵⁰DKREG²⁶⁰PK²⁷⁰EPDL²⁸⁰NFGK²⁹⁰.....HI
 Pae³⁰⁰sym_/30-314 CELO³¹⁰RF³²⁰GFY³³⁰IS³⁴⁰GOE³⁵⁰IK³⁶⁰IVH³⁷⁰EG³⁸⁰ST³⁹⁰GS⁴⁰⁰YH⁴¹⁰AV⁴²⁰IG⁴³⁰Y⁴⁴⁰PEL⁴⁵⁰JK⁴⁶⁰..PE⁴⁷⁰NY⁴⁸⁰LL⁴⁹⁰T⁵⁰⁰T⁵¹⁰OL⁵²⁰.....HK
 Bac⁵³⁰the_/357-656 ...E⁵⁴⁰FT⁵⁵⁰GIS⁵⁶⁰V⁵⁷⁰KAG⁵⁸⁰DDI⁵⁹⁰IV⁶⁰⁰VL⁶¹⁰GD⁶²⁰TV⁶³⁰GS⁶⁴⁰MI⁶⁵⁰SH⁶⁶⁰CI⁶⁷⁰X⁶⁸⁰RE⁶⁹⁰T⁷⁰⁰GE⁷¹⁰V⁷²⁰K⁷³⁰QT⁷⁴⁰AS⁷⁵⁰SD⁷⁶⁰V⁷⁷⁰Y⁷⁸⁰ML⁷⁹⁰N⁸⁰⁰FG⁸¹⁰.....VHK
 Tri⁸²⁰vaq_/141-445 ..YH⁸³⁰ET⁸⁴⁰LV⁸⁵⁰AK⁸⁶⁰EG⁸⁷⁰ELI⁸⁸⁰TE⁸⁹⁰IE⁹⁰⁰FA⁹¹⁰V⁹²⁰GN⁹³⁰LV⁹⁴⁰LV⁹⁵⁰LN⁹⁶⁰RQ⁹⁷⁰AF⁹⁸⁰SH⁹⁹⁰GD¹⁰⁰⁰LT¹⁰¹⁰OR¹⁰²⁰Y¹⁰³⁰EN¹⁰⁴⁰LOC¹⁰⁵⁰SET¹⁰⁶⁰LN¹⁰⁷⁰.....AK¹⁰⁸⁰KVT
 Akk¹⁰⁹⁰muc_/277-579 GGX¹¹⁰⁰HS¹¹¹⁰ET¹¹²⁰LV¹¹³⁰AF¹¹⁴⁰FO¹¹⁵⁰AE¹¹⁶⁰RS¹¹⁷⁰CS¹¹⁸⁰LS¹¹⁹⁰GA¹²⁰⁰PN¹²¹⁰GG¹²²⁰IS¹²³⁰VS¹²⁴⁰RI¹²⁵⁰CG¹²⁶⁰HT¹²⁷⁰DS¹²⁸⁰LH¹²⁹⁰KL¹³⁰⁰DE¹³¹⁰M¹³²⁰RR¹³³⁰V¹³⁴⁰FE¹³⁵⁰IT¹³⁶⁰MO¹³⁷⁰UF¹³⁸⁰AG¹³⁹⁰RR¹⁴⁰⁰.....VK
 Esccol_/1081-1381 GHK¹⁴¹⁰QS¹⁴²⁰ET¹⁴³⁰LV¹⁴⁴⁰AF¹⁴⁵⁰FO¹⁴⁶⁰AE¹⁴⁷⁰RS¹⁴⁸⁰CS¹⁴⁹⁰LS¹⁵⁰⁰GA¹⁵¹⁰PN¹⁵²⁰GG¹⁵³⁰IS¹⁵⁴⁰VS¹⁵⁵⁰RI¹⁵⁶⁰CG¹⁵⁷⁰HT¹⁵⁸⁰DS¹⁵⁹⁰LH¹⁶⁰⁰KL¹⁶¹⁰DE¹⁶²⁰M¹⁶³⁰RR¹⁶⁴⁰V¹⁶⁵⁰FE¹⁶⁶⁰IT¹⁶⁷⁰MO¹⁶⁸⁰UF¹⁶⁹⁰AG¹⁷⁰⁰RR¹⁷¹⁰.....TVK

80 90 100 110 120 130
 Vibcho_/1087-1392 FTV⁸¹PYG⁹¹GL¹⁰¹YV¹¹¹AG¹²¹QO¹³¹..SE¹⁴¹LV¹⁵¹SL¹⁶¹IT¹⁷¹FG¹⁸¹IT¹⁹¹ID²⁰¹DF²¹¹FL²²¹X²³¹KE²⁴¹Q²⁵¹WE²⁶¹H²⁷¹Q²⁸¹LD²⁹¹GS³⁰¹FA³¹¹FI³²¹GE³³¹V³⁴¹V³⁵¹SH³⁶¹GF³⁷¹V³⁸¹AF³⁹¹K⁴⁰¹A⁴¹¹H
 Bac⁴²¹thu_/78-367 ISS⁴³¹PS⁴⁴¹GG⁴⁵¹ILL⁴⁶¹Y⁴⁷¹F⁴⁸¹Y⁴⁹¹H⁵⁰¹MM⁵¹¹..TG⁵²¹ET⁵³¹T⁵⁴¹AT⁵⁵¹IT⁵⁶¹SG⁵⁷¹GT⁵⁸¹HE⁵⁹¹FF⁶⁰¹LE⁶¹¹FL⁶²¹GL⁶³¹HT⁶⁴¹KK⁶⁵¹Q⁶⁶¹WD⁶⁷¹AM⁶⁸¹LE⁶⁹¹K⁷⁰¹Y⁷¹¹KN⁷²¹PF⁷³¹AI⁷⁴¹EL⁷⁵¹KG⁷⁶¹DR⁷⁷¹SL
 Pae⁷⁸¹sym_/30-314 PTK⁷⁹¹NE⁸⁰¹GG⁸¹¹LL⁸²¹LF⁸³¹T⁸⁴¹HH⁸⁵¹MM⁸⁶¹..MG⁸⁷¹Y⁸⁸¹KK⁸⁹¹IL⁹⁰¹X⁹¹¹SE⁹²¹LO⁹³¹K⁹⁴¹IF⁹⁵¹SP⁹⁶¹KL⁹⁷¹HE⁹⁸¹TH⁹⁹¹TD¹⁰⁰¹WE¹⁰¹¹HM¹⁰²¹MA¹⁰³¹SY¹⁰⁴¹SD¹⁰⁵¹AF¹⁰⁶¹V¹⁰⁷¹VO¹⁰⁸¹LS¹⁰⁹¹SE¹¹⁰¹RA¹¹¹¹I
 Bac¹¹²¹the_/357-656 LTK¹¹³¹KG¹¹⁴¹GG¹¹⁵¹QL¹¹⁶¹LV¹¹⁷¹MY¹¹⁸¹HT¹¹⁹¹EL¹²⁰¹HT¹²¹¹AK¹²²¹PI¹²³¹KE¹²⁴¹R¹²⁵¹IF¹²⁶¹LG¹²⁷¹SG¹²⁸¹T¹²⁹¹NG¹³⁰¹FF¹³¹¹ED¹³²¹LE¹³³¹SK¹³⁴¹TD¹³⁵¹DE¹³⁶¹K¹³⁷¹Y¹³⁸¹AE¹³⁹¹LL¹⁴⁰¹KN¹⁴¹¹ST¹⁴²¹HK¹⁴³¹Y¹⁴⁴¹FC¹⁴⁵¹I¹⁴⁶¹R¹⁴⁷¹GE¹⁴⁸¹NI
 Tri¹⁴⁹¹vaq_/141-445 FGYP¹⁵⁰¹PL¹⁵¹¹GG¹⁵²¹Y¹⁵³¹ND¹⁵⁴¹IK¹⁵⁵¹CK¹⁵⁶¹ID¹⁵⁷¹TF¹⁵⁸¹..LH¹⁵⁹¹GV¹⁶⁰¹VE¹⁶¹¹II¹⁶²¹IT¹⁶³¹GS¹⁶⁴¹IR¹⁶⁵¹IF¹⁶⁶¹Y¹⁶⁷¹RY¹⁶⁸¹GA¹⁶⁹¹ES¹⁷⁰¹DD¹⁷¹¹WE¹⁷²¹KE¹⁷³¹ET¹⁷⁴¹R¹⁷⁵¹HY¹⁷⁶¹VA¹⁷⁷¹PL¹⁷⁸¹TE¹⁷⁹¹Y¹⁸⁰¹DT¹⁸¹¹GH¹⁸²¹V¹⁸³¹EC
 Akk¹⁸⁴¹muc_/277-579 NVN¹⁸⁵¹PN¹⁸⁶¹GG¹⁸⁷¹LL¹⁸⁸¹Y¹⁸⁹¹V¹⁹⁰¹Y¹⁹¹¹Q¹⁹²¹OR¹⁹³¹FR¹⁹⁴¹..RG¹⁹⁵¹IL¹⁹⁶¹ER¹⁹⁷¹Y¹⁹⁸¹Q¹⁹⁹¹IS²⁰⁰¹GA²⁰¹¹VE²⁰²¹ST²⁰³¹PL²⁰⁴¹V²⁰⁵¹NG²⁰⁶¹TT²⁰⁷¹FE²⁰⁸¹OX²⁰⁹¹AE²¹⁰¹OLE²¹¹¹N²¹²¹.....TR²¹³¹AP²¹⁴¹Q²¹⁵¹OE
 Esccol_/1081-1381 PKV²¹⁶¹PYG²²⁶¹GL²³⁶¹Y²⁴⁶¹V²⁵⁶¹AG²⁶⁶¹QO²⁷⁶¹..TH²⁸⁶¹ET²⁹⁶¹LV³⁰⁶¹SL³¹⁶¹IT³²⁶¹FG³³⁶¹V³⁴⁶¹V³⁵⁶¹SH³⁶⁶¹GF³⁷⁶¹V³⁸⁶¹AF³⁹⁶¹K⁴⁰⁶¹A⁴¹⁶¹H⁴²⁶¹LD⁴³⁶¹NS⁴⁴⁶¹FA⁴⁵⁶¹FI⁴⁶⁶¹GE⁴⁷⁶¹V⁴⁸⁶¹V⁴⁹⁶¹SH⁵⁰⁶¹GF⁵¹⁶¹V⁵²⁶¹AF⁵³⁶¹K⁵⁴⁶¹A⁵⁵⁶¹H⁵⁶⁶¹LD⁵⁷⁶¹NS⁵⁸⁶¹FA⁵⁹⁶¹FI⁶⁰⁶¹GE⁶¹⁶¹V⁶²⁶¹V⁶³⁶¹SH⁶⁴⁶¹GF⁶⁵⁶¹V⁶⁶⁶¹AF⁶⁷⁶¹K⁶⁸⁶¹A⁶⁹⁶¹H⁷⁰⁶¹LD⁷¹⁶¹NS⁷²⁶¹FA⁷³⁶¹FI⁷⁴⁶¹GE⁷⁵⁶¹V⁷⁶⁶¹V⁷⁷⁶¹SH⁷⁸⁶¹GF⁷⁹⁶¹V⁸⁰⁶¹AF⁸¹⁶¹K⁸²⁶¹A⁸³⁶¹H⁸⁴⁶¹LD⁸⁵⁶¹NS⁸⁶⁶¹FA⁸⁷⁶¹FI⁸⁸⁶¹GE⁸⁹⁶¹V⁹⁰⁶¹V⁹¹⁶¹SH⁹²⁶¹GF⁹³⁶¹V⁹⁴⁶¹AF⁹⁵⁶¹K⁹⁶⁶¹A⁹⁷⁶¹H⁹⁸⁶¹LD⁹⁹⁶¹NS¹⁰⁰⁶¹FA¹⁰¹⁶¹FI¹⁰²⁶¹GE¹⁰³⁶¹V¹⁰⁴⁶¹V¹⁰⁵⁶¹SH¹⁰⁶⁶¹GF¹⁰⁷⁶¹V¹⁰⁸⁶¹AF¹⁰⁹⁶¹K¹¹⁰⁶¹A¹¹¹⁶¹H¹¹²⁶¹LD¹¹³⁶¹NS¹¹⁴⁶¹FA¹¹⁵⁶¹FI¹¹⁶⁶¹GE¹¹⁷⁶¹V¹¹⁸⁶¹V¹¹⁹⁶¹SH¹²⁰⁶¹GF¹²¹⁶¹V¹²²⁶¹AF¹²³⁶¹K¹²⁴⁶¹A¹²⁵⁶¹H¹²⁶⁶¹LD¹²⁷⁶¹NS¹²⁸⁶¹FA¹²⁹⁶¹FI¹³⁰⁶¹GE¹³¹⁶¹V¹³²⁶¹V¹³³⁶¹SH¹³⁴⁶¹GF¹³⁵⁶¹V¹³⁶⁶¹AF¹³⁷⁶¹K¹³⁸⁶¹A¹³⁹⁶¹H¹⁴⁰⁶¹LD¹⁴¹⁶¹NS¹⁴²⁶¹FA¹⁴³⁶¹FI¹⁴⁴⁶¹GE¹⁴⁵⁶¹V¹⁴⁶⁶¹V¹⁴⁷⁶¹SH¹⁴⁸⁶¹GF¹⁴⁹⁶¹V¹⁵⁰⁶¹AF¹⁵¹⁶¹K¹⁵²⁶¹A¹⁵³⁶¹H¹⁵⁴⁶¹LD¹⁵⁵⁶¹NS¹⁵⁶⁶¹FA¹⁵⁷⁶¹FI¹⁵⁸⁶¹GE¹⁵⁹⁶¹V¹⁶⁰⁶¹V¹⁶¹⁶¹SH¹⁶²⁶¹GF¹⁶³⁶¹V¹⁶⁴⁶¹AF¹⁶⁵⁶¹K¹⁶⁶⁶¹A¹⁶⁷⁶¹H¹⁶⁸⁶¹LD¹⁶⁹⁶¹NS¹⁷⁰⁶¹FA¹⁷¹⁶¹FI¹⁷²⁶¹GE¹⁷³⁶¹V¹⁷⁴⁶¹V¹⁷⁵⁶¹SH¹⁷⁶⁶¹GF¹⁷⁷⁶¹V¹⁷⁸⁶¹AF¹⁷⁹⁶¹K¹⁸⁰⁶¹A¹⁸¹⁶¹H¹⁸²⁶¹LD¹⁸³⁶¹NS¹⁸⁴⁶¹FA¹⁸⁵⁶¹FI¹⁸⁶⁶¹GE¹⁸⁷⁶¹V¹⁸⁸⁶¹V¹⁸⁹⁶¹SH¹⁹⁰⁶¹GF¹⁹¹⁶¹V¹⁹²⁶¹AF¹⁹³⁶¹K¹⁹⁴⁶¹A¹⁹⁵⁶¹H¹⁹⁶⁶¹LD¹⁹⁷⁶¹NS¹⁹⁸⁶¹FA¹⁹⁹⁶¹FI²⁰⁰⁶¹GE²⁰¹⁶¹V²⁰²⁶¹V²⁰³⁶¹SH²⁰⁴⁶¹GF²⁰⁵⁶¹V²⁰⁶⁶¹AF²⁰⁷⁶¹K²⁰⁸⁶¹A²⁰⁹⁶¹H²¹⁰⁶¹LD²¹¹⁶¹NS²¹²⁶¹FA²¹³⁶¹FI²¹⁴⁶¹GE²¹⁵⁶¹V²¹⁶⁶¹V²¹⁷⁶¹SH²¹⁸⁶¹GF²¹⁹⁶¹V²²⁰⁶¹AF²²¹⁶¹K²²²⁶¹A²²³⁶¹H²²⁴⁶¹LD²²⁵⁶¹NS²²⁶⁶¹FA²²⁷⁶¹FI²²⁸⁶¹GE²²⁹⁶¹V²³⁰⁶¹V²³¹⁶¹SH²³²⁶¹GF²³³⁶¹V²³⁴⁶¹AF²³⁵⁶¹K²³⁶⁶¹A²³⁷⁶¹H²³⁸⁶¹LD²³⁹⁶¹NS²⁴⁰⁶¹FA²⁴¹⁶¹FI²⁴²⁶¹GE²⁴³⁶¹V²⁴⁴⁶¹V²⁴⁵⁶¹SH²⁴⁶⁶¹GF²⁴⁷⁶¹V²⁴⁸⁶¹AF²⁴⁹⁶¹K²⁵⁰⁶¹A²⁵¹⁶¹H²⁵²⁶¹LD²⁵³⁶¹NS²⁵⁴⁶¹FA²⁵⁵⁶¹FI²⁵⁶⁶¹GE²⁵⁷⁶¹V²⁵⁸⁶¹V²⁵⁹⁶¹SH²⁶⁰⁶¹GF²⁶¹⁶¹V²⁶²⁶¹AF²⁶³⁶¹K²⁶⁴⁶¹A²⁶⁵⁶¹H²⁶⁶⁶¹LD²⁶⁷⁶¹NS²⁶⁸⁶¹FA²⁶⁹⁶¹FI²⁷⁰⁶¹GE²⁷¹⁶¹V²⁷²⁶¹V²⁷³⁶¹SH²⁷⁴⁶¹GF²⁷⁵⁶¹V²⁷⁶⁶¹AF²⁷⁷⁶¹K²⁷⁸⁶¹A²⁷⁹⁶¹H²⁸⁰⁶¹LD²⁸¹⁶¹NS²⁸²⁶¹FA²⁸³⁶¹FI²⁸⁴⁶¹GE²⁸⁵⁶¹V²⁸⁶⁶¹V²⁸⁷⁶¹SH²⁸⁸⁶¹GF²⁸⁹⁶¹V²⁹⁰⁶¹AF²⁹¹⁶¹K²⁹²⁶¹A²⁹³⁶¹H²⁹⁴⁶¹LD²⁹⁵⁶¹NS²⁹⁶⁶¹FA²⁹⁷⁶¹FI²⁹⁸⁶¹GE²⁹⁹⁶¹V³⁰⁰⁶¹V³⁰¹⁶¹SH³⁰²⁶¹GF³⁰³⁶¹V³⁰⁴⁶¹AF³⁰⁵⁶¹K³⁰⁶⁶¹A³⁰⁷⁶¹H³⁰⁸⁶¹LD³⁰⁹⁶¹NS³¹⁰⁶¹FA³¹¹⁶¹FI³¹²⁶¹GE³¹³⁶¹V³¹⁴⁶¹V³¹⁵⁶¹SH³¹⁶⁶¹GF³¹⁷⁶¹V³¹⁸⁶¹AF³¹⁹⁶¹K³²⁰⁶¹A³²¹⁶¹H³²²⁶¹LD³²³⁶¹NS³²⁴⁶¹FA³²⁵⁶¹FI³²⁶⁶¹GE³²⁷⁶¹V³²⁸⁶¹V³²⁹⁶¹SH³³⁰⁶¹GF³³¹⁶¹V³³²⁶¹AF³³³⁶¹K³³⁴⁶¹A³³⁵⁶¹H³³⁶⁶¹LD³³⁷⁶¹NS³³⁸⁶¹FA³³⁹⁶¹FI³⁴⁰⁶¹GE³⁴¹⁶¹V³⁴²⁶¹V³⁴³⁶¹SH³⁴⁴⁶¹GF³⁴⁵⁶¹V³⁴⁶⁶¹AF³⁴⁷⁶¹K³⁴⁸⁶¹A³⁴⁹⁶¹H³⁵⁰⁶¹LD³⁵¹⁶¹NS³⁵²⁶¹FA³⁵³⁶¹FI³⁵⁴⁶¹GE³⁵⁵⁶¹V³⁵⁶⁶¹V³⁵⁷⁶¹SH³⁵⁸⁶¹GF³⁵⁹⁶¹V³⁶⁰⁶¹AF³⁶¹⁶¹K³⁶²⁶¹A³⁶³⁶¹H³⁶⁴⁶¹LD³⁶⁵⁶¹NS³⁶⁶⁶¹FA³⁶⁷⁶¹FI³⁶⁸⁶¹GE³⁶⁹⁶¹V³⁷⁰⁶¹V³⁷¹⁶¹SH³⁷²⁶¹GF³⁷³⁶¹V³⁷⁴⁶¹AF³⁷⁵⁶¹K³⁷⁶⁶¹A³⁷⁷⁶¹H³⁷⁸⁶¹LD³⁷⁹⁶¹NS³⁸⁰⁶¹FA³⁸¹⁶¹FI³⁸²⁶¹GE³⁸³⁶¹V³⁸⁴⁶¹V³⁸⁵⁶¹SH³⁸⁶⁶¹GF³⁸⁷⁶¹V³⁸⁸⁶¹AF³⁸⁹⁶¹K³⁹⁰⁶¹A³⁹¹⁶¹H³⁹²⁶¹LD³⁹³⁶¹NS³⁹⁴⁶¹FA³⁹⁵⁶¹FI³⁹⁶⁶¹GE³⁹⁷⁶¹V³⁹⁸⁶¹V³⁹⁹⁶¹SH⁴⁰⁰⁶¹GF⁴⁰¹⁶¹V⁴⁰²⁶¹AF⁴⁰³⁶¹K⁴⁰⁴⁶¹A⁴⁰⁵⁶¹H⁴⁰⁶⁶¹LD⁴⁰⁷⁶¹NS⁴⁰⁸⁶¹FA⁴⁰⁹⁶¹FI⁴¹⁰⁶¹GE⁴¹¹⁶¹V⁴¹²⁶¹V⁴¹³⁶¹SH⁴¹⁴⁶¹GF⁴¹⁵⁶¹V⁴¹⁶⁶¹AF⁴¹⁷⁶¹K⁴¹⁸⁶¹A⁴¹⁹⁶¹H⁴²⁰⁶¹LD⁴²¹⁶¹NS⁴²²⁶¹FA⁴²³⁶¹FI⁴²⁴⁶¹GE⁴²⁵⁶¹V⁴²⁶⁶¹V⁴²⁷⁶¹SH⁴²⁸⁶¹GF⁴²⁹⁶¹V⁴³⁰⁶¹AF⁴³¹⁶¹K⁴³²⁶¹A⁴³³⁶¹H⁴³⁴⁶¹LD⁴³⁵⁶¹NS⁴³⁶⁶¹FA⁴³⁷⁶¹FI⁴³⁸⁶¹GE⁴³⁹⁶¹V⁴⁴⁰⁶¹V⁴⁴¹⁶¹SH⁴⁴²⁶¹GF⁴⁴³⁶¹V⁴⁴⁴⁶¹AF⁴⁴⁵⁶¹K⁴⁴⁶⁶¹A⁴⁴⁷⁶¹H⁴⁴⁸⁶¹LD⁴⁴⁹⁶¹NS⁴⁵⁰⁶¹FA⁴⁵¹⁶¹FI⁴⁵²⁶¹GE⁴⁵³⁶¹V⁴⁵⁴⁶¹V⁴⁵⁵⁶¹SH⁴⁵⁶⁶¹GF⁴⁵⁷⁶¹V⁴⁵⁸⁶¹AF⁴⁵⁹⁶¹K⁴⁶⁰⁶¹A⁴⁶¹⁶¹H⁴⁶²⁶¹LD⁴⁶³⁶¹NS⁴⁶⁴⁶¹FA⁴⁶⁵⁶¹FI⁴⁶⁶⁶¹GE⁴⁶⁷⁶¹V⁴⁶⁸⁶¹V⁴⁶⁹⁶¹SH⁴⁷⁰⁶¹GF⁴⁷¹⁶¹V⁴⁷²⁶¹AF⁴⁷³⁶¹K⁴⁷⁴⁶¹A⁴⁷⁵⁶¹H⁴⁷⁶⁶¹LD⁴⁷⁷⁶¹NS⁴⁷⁸⁶¹FA⁴⁷⁹⁶¹FI⁴⁸⁰⁶¹GE⁴⁸¹⁶¹V⁴⁸²⁶¹V⁴⁸³⁶¹SH⁴⁸⁴⁶¹GF⁴⁸⁵⁶¹V⁴⁸⁶⁶¹AF⁴⁸⁷⁶¹K⁴⁸⁸⁶¹A⁴⁸⁹⁶¹H⁴⁹⁰⁶¹LD⁴⁹¹⁶¹NS⁴⁹²⁶¹FA⁴⁹³⁶¹FI⁴⁹⁴⁶¹GE⁴⁹⁵⁶¹V⁴⁹⁶⁶¹V⁴⁹⁷⁶¹SH⁴⁹⁸⁶¹GF⁴⁹⁹⁶¹V⁵⁰⁰⁶¹AF⁵⁰¹⁶¹K⁵⁰²⁶¹A⁵⁰³⁶¹H⁵⁰⁴⁶¹LD⁵⁰⁵⁶¹NS⁵⁰⁶⁶¹FA⁵⁰⁷⁶¹FI⁵⁰⁸⁶¹GE⁵⁰⁹⁶¹V⁵¹⁰⁶¹V⁵¹¹⁶¹SH⁵¹²⁶¹GF⁵¹³⁶¹V⁵¹⁴⁶¹AF⁵¹⁵⁶¹K⁵¹⁶⁶¹A⁵¹⁷⁶¹H⁵¹⁸⁶¹LD⁵¹⁹⁶¹NS⁵²⁰⁶¹FA⁵²¹⁶¹FI⁵²²⁶¹GE⁵²³⁶¹V⁵²⁴⁶¹V⁵²⁵⁶¹SH⁵²⁶⁶¹GF⁵²⁷⁶¹V⁵²⁸⁶¹AF⁵²⁹⁶¹K⁵³⁰⁶¹A⁵³¹⁶¹H⁵³²⁶¹LD⁵³³⁶¹NS⁵³⁴⁶¹FA⁵³⁵⁶¹FI⁵³⁶⁶¹GE⁵³⁷⁶¹V⁵³⁸⁶¹V⁵³⁹⁶¹SH⁵⁴⁰⁶¹GF⁵⁴¹⁶¹V⁵⁴²⁶¹AF⁵⁴³⁶¹K⁵⁴⁴⁶¹A⁵⁴⁵⁶¹H⁵⁴⁶⁶¹LD⁵⁴⁷⁶¹NS⁵⁴⁸⁶¹FA⁵⁴⁹⁶¹FI⁵⁵⁰⁶¹GE⁵⁵¹⁶¹V⁵⁵²⁶¹V⁵⁵³⁶¹SH⁵⁵⁴⁶¹GF⁵⁵⁵⁶¹V⁵⁵⁶⁶¹AF⁵⁵⁷⁶¹K⁵⁵⁸⁶¹A⁵⁵⁹⁶¹H⁵⁶⁰⁶¹LD⁵⁶¹⁶¹NS⁵⁶²⁶¹FA⁵⁶³⁶¹FI⁵⁶⁴⁶¹GE⁵⁶⁵⁶¹V⁵⁶⁶⁶¹V⁵⁶⁷⁶¹SH⁵⁶⁸⁶¹GF⁵⁶⁹⁶¹V⁵⁷⁰⁶¹AF⁵⁷¹⁶¹K⁵⁷²⁶¹A⁵⁷³⁶¹H⁵⁷⁴⁶¹LD⁵⁷⁵⁶¹NS⁵⁷⁶⁶¹FA⁵⁷⁷⁶¹FI⁵⁷⁸⁶¹GE⁵⁷⁹⁶¹V⁵⁸⁰⁶¹V⁵⁸¹⁶¹SH⁵⁸²⁶¹GF⁵⁸³⁶¹V⁵⁸⁴⁶¹AF⁵⁸⁵⁶¹K⁵⁸⁶⁶¹A⁵⁸⁷⁶¹H⁵⁸⁸⁶¹LD⁵⁸⁹⁶¹NS⁵⁹⁰⁶¹FA⁵⁹¹⁶¹FI⁵⁹²⁶¹GE⁵⁹³⁶¹V⁵⁹⁴⁶¹V⁵⁹⁵⁶¹SH⁵⁹⁶⁶¹GF⁵⁹⁷⁶¹V⁵⁹⁸⁶¹AF⁵⁹⁹⁶¹K⁶⁰⁰⁶¹A⁶⁰¹⁶¹H⁶⁰²⁶¹LD⁶⁰³⁶¹NS⁶⁰⁴⁶¹FA⁶⁰⁵⁶¹FI⁶⁰⁶⁶¹GE⁶⁰⁷⁶¹V⁶⁰⁸⁶¹V⁶⁰⁹⁶¹SH⁶¹⁰⁶¹GF⁶¹¹⁶¹V⁶¹²⁶¹AF⁶¹³⁶¹K⁶¹⁴⁶¹A⁶¹⁵⁶¹H⁶¹⁶⁶¹LD⁶¹⁷⁶¹NS⁶¹⁸⁶¹FA⁶¹⁹⁶¹FI⁶²⁰⁶¹GE⁶²¹⁶¹V⁶²²⁶¹V⁶²³⁶¹SH⁶²⁴⁶¹GF⁶²⁵⁶¹V⁶²⁶⁶¹AF⁶²⁷⁶¹K⁶²⁸⁶¹A⁶²⁹⁶¹H⁶³⁰⁶¹LD⁶³¹⁶¹NS⁶³²⁶¹FA⁶³³⁶¹FI⁶³⁴⁶¹GE⁶³⁵⁶¹V⁶³⁶⁶¹V⁶³⁷⁶¹SH⁶³⁸⁶¹GF⁶³⁹⁶¹V⁶⁴⁰⁶¹AF⁶⁴¹⁶¹K⁶⁴²⁶¹A⁶⁴³⁶¹H⁶⁴⁴⁶¹LD⁶⁴⁵⁶¹NS⁶⁴⁶⁶¹FA⁶⁴⁷⁶¹FI⁶⁴⁸⁶¹GE⁶⁴⁹⁶¹V⁶⁵⁰⁶¹V⁶⁵¹⁶¹SH⁶⁵²⁶¹GF⁶⁵³⁶¹V⁶⁵⁴⁶¹AF⁶⁵⁵⁶¹K⁶⁵⁶⁶¹A⁶⁵⁷⁶¹H⁶⁵⁸⁶¹LD⁶⁵⁹⁶¹NS⁶⁶⁰⁶¹FA⁶⁶¹⁶¹FI⁶⁶²⁶¹GE⁶⁶³⁶¹V⁶⁶⁴⁶¹V⁶⁶⁵⁶¹SH⁶⁶⁶⁶¹GF⁶⁶⁷⁶¹V⁶⁶⁸⁶¹AF⁶⁶⁹⁶¹K⁶⁷⁰⁶¹A⁶⁷¹⁶¹H⁶⁷²⁶¹LD⁶⁷³⁶¹NS⁶⁷⁴⁶¹FA⁶⁷⁵⁶¹FI⁶⁷⁶⁶¹GE⁶⁷⁷⁶¹V⁶⁷⁸⁶¹V⁶⁷⁹⁶¹SH⁶⁸⁰⁶¹GF⁶⁸¹⁶¹V⁶⁸²⁶¹AF⁶⁸³⁶¹K⁶⁸⁴⁶¹A⁶⁸⁵⁶¹H⁶⁸⁶⁶¹LD⁶⁸⁷⁶¹NS⁶⁸⁸⁶¹FA⁶⁸⁹⁶¹FI⁶⁹⁰⁶¹GE⁶⁹¹⁶¹V⁶⁹²⁶¹V⁶⁹³⁶¹SH⁶⁹⁴⁶¹GF⁶⁹⁵⁶¹V⁶⁹⁶⁶¹AF⁶⁹⁷⁶¹K⁶⁹⁸⁶¹A⁶⁹⁹⁶¹H⁷⁰⁰⁶¹LD⁷⁰¹⁶¹NS⁷⁰²⁶¹FA⁷⁰³⁶¹FI⁷⁰⁴⁶¹GE⁷⁰⁵⁶¹V⁷⁰⁶⁶¹V⁷⁰⁷⁶¹SH⁷⁰⁸⁶¹GF⁷⁰⁹⁶¹V⁷¹⁰⁶¹AF⁷¹¹⁶¹K⁷¹²⁶¹A⁷¹³⁶¹H⁷¹⁴⁶¹LD⁷¹⁵⁶¹NS⁷¹⁶⁶¹FA⁷¹⁷⁶¹FI⁷¹⁸⁶¹GE⁷¹⁹⁶¹V<

Figure 11. Multiple sequence alignment of the M60-like domain of YghJ with similar proteins. Alignment was performed using ClustlW 18.1 (Thompson *et al.* 1994) *Bacteroides fragilis* lipoprotein as a prototype. Identical residues are shown with a red background, while similar residues are indicated by a red letter. Conserved motifs are surrounded by a blue box. The zinc binding motif HEXXH is surrounded by the green box. The full species names and protein accession numbers are as follows : *Vibrio cholerae* ZP_06941093.1, *Bacteroides thetaiotaomicron* NP_813155.1, *Trichomonas vaginalis* XP_001330197.1, *Akkermansia muciniphila* YP_001878116.1, *Escherichia coli* YP_026189.1

See Figure 12 of the ClustalW 18.1 multialignment software we generated that shows these conserved motifs. The motifs include important amino acid residues (Das et al., 1996) and a conserved secondary structure which yield a characteristic $\beta\beta\beta\alpha\beta$ repeat motif (Bork et al., 1996). A comparison of this region, along with approximately 100 amino acids of flanking sequence on either side of the PP2C region revealed that several of the most important motifs are present in YghJ. There are eleven important motifs used to predict the PP2C domain yet only 7 of them possess any conserved amino acid residues; these are motifs 1, 2, 4, 5, 8, and 11. The amino acid residues found in the remaining motifs are not well conserved and the similarities between the PP2C domains can only be observed in the secondary structure of these domains. In addition to the recognized motifs, as mentioned above, (Das et al., 1996) identified an arginine residue at approximately position 33 of the PP2C domain of the prototype human PP2C α . This residue is important in catalysis and is conserved in YghJ. An aspartic acid at approximately position 40 of the human PP2C α (Motif 1), an aspartic acid at approximately residue 65 (Motif 2), and aspartic acid (Motif 8) are believed to coordinate the metal. Motifs 4 through 7 are indicated on the figure but the role of these residues is unknown. At Motif 11, a conserved aspartic acid residue followed by an asparagine residue located at about position 440 is changed in YghJ and this small change may indicate a different metal binding capability (Huang et al., 1997).

10 20 30 40 50 60 70 80 90 100 110 120 130 140 150

MOTIF 1 MOTIF 2

PppA MV---PNSVSYR-SA-S---YSH-VGMV-RKVN-CA-SLDAP-E-AG-LNV-VADG-G-GHA-AGDFV-SBL-IVDYLR-IPA-ASBL
 Predator cc-----cccccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----hhhh-----cccc-----cccc-----
 StpI M---QRTEPWR-SA-A---RTD-RGV-RARNE-CA-FLDCP-D-RG-LWA-VADG-G-GNQ-AGOLA-SRM-IYTSIAE-LPA-RAGT
 Predator c-----cccccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----cchh-----hhhh-----cccc-----cccc-----
 PpH M---RIEVA-G---G---STR-VGNK-RNNE-DN-FLMLP-E-EF-LFC-VADG-G-GHS-SGEIA-SRI-AVDELGEF-YLCTSKDCCYV
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----hhhh-----hhhh-----cccc-----cccc-----
 s111771 MT---EVNLSV/SCSST-G---KTD-PGLV-RQYVQ-DS-FVLDP-EG-RF-YI-VADG-G-GHA-QGEA-SRI-AVERVRD-Y-L---DYTW
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----cchh-----hhhh-----cccc-----cccc-----
 PP2Ca MGAFLDKPKGKGN-AQ-C---CGNGLRYGLSSGQGWVEE-CAITA-VIGLPSGLE-S-YS-FFA-VYDG-A-SCQ-VAYVC-CRH-LLDRITRN-QDFKQ-SAGA
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----cchh-----hhhh-----cccc-----cccc-----
 PpP TERLARQKLARSBAKGINRFEDLNAVLD-IGLV-RANQ-D-LIAAIRVWTP-SH-VGNPTTAMA-EDGQ-G-GNQ-DGKQC-ATI-ALSTLYS-L-INK-RSDP
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----cchh-----hhhh-----cccc-----cccc-----
 YghJ AK---EIDTAICA-KTD-GCN---EARK-FSLTT-RVNV-D-CCGQGVINKLWQVQVTRYKSVKGF-V-EDS-SN-FYGTGNA-RQJAV-VNISNAAFFILHARDNGYLAFG-EGRADWQKELAY
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----cchh-----hhhh-----cccc-----cccc-----
 SpoII QSI---TRKVIYIPGTVERLQ-E---CCQ-YARKI-KDVAQVQ-SSMVFHALS-E-SFATFYQASDQ-SSVDLFLSKITERSGCTCYGGRROWQVDFKTDLH-QQVM-LETEEKYASNR
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----hhhh-----hhhh-----cccc-----cccc-----

161 171 181 191 201 211 221 231 241 251 261 271 281 291 301

MOTIF 3 MOTIF 4 MOTIF 5

PppA PA-YV---GALRTOLAQVHE-RVRC-EAGL-RGV-SVMSISVLLAAR-GMQ-ASCLMAG-DSRLYR-LR-QGVLEAT-S-R-D-HS-YVQELLDN-AL-IS-EEKA
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----hhhh-----hhhh-----cccc-----cccc-----
 StpI DE-RL---NAMRQCLHWLNR-RLGE-ELTSLAERQ-RIDSSVVALVVE-GSR-AACVWVG-DSRYL-WR-QORLYQL-S-R-D-HS-LLQRLVDE-RA-ND-IEQA
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----hhhh-----hhhh-----cccc-----cccc-----
 PpH PF---KMDKTRNDENRLAIGVLANA-RIFE-KACSES-KY-KGMPITIVTVHFS-QSA-VYVGVG-DSRVYV-FR-QGALKQV-T-E-D-ES-LLNDYLNK-KK-LS-PEEI
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----cchh-----hhhh-----cccc-----cccc-----
 s111771 QS-EI-T---SEQLLRDALMDANE-GILEDCQINLE-R-RMPTTAVLIAFR-EDG-AMRAVVG-DSRLYR-LR-NQQLERV-T-E-D-HT-NVARALKM-GC-ID-PAQA
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----hhhh-----hhhh-----cccc-----cccc-----
 PP2Ca PSVENV---KNGIRTGFLIEDRVMSSKGGAD-R-SGSAVGVGLIS-POE-TYFVCG-DSRGLL-CR-NRUVVFF-T-Q-D-HK-F-ENPLEK-ER-IQ-NAGG
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----cchh-----hhhh-----cccc-----cccc-----
 PpP FE---SRLKATLEANS-VVYD-YAKE-H-GSLSAVIIEGSA-PVIVVVG-DSRIYS-FSL-DCGLTAL-S-S-D-DS-L-EALOCR-GPGLLVIGMGE
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----hhhh-----hhhh-----cccc-----cccc-----
 YghJ ITEAPS-LVE---PENVRD-TATFNL-PFISLQVGGG-KL-MVIGNPFRNSILRCFNGVSWNGVKNKGGCTLNDFDDMGNFM-ENVLYL-S-D-EKWKPKAASNTVGTMLDTVYFKRQ-CQ-VTGNBAJF
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----cccc-----cccc-----cccc-----cccc-----
 SpoII LIOGTFQCYCSKMGVVELLIEDELAKHHAH-LTLK-KVQDSR-RLVARQLVSEVMADFREIKRRECFLOEPIIRA-LQVFLIEIQVVEYSLRQGNIDEMTTFVSGHGESEKILAPMLSDILEQ-IL-VKAI
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----hhhh-----hhhh-----cccc-----cccc-----

312 322 332 342 352 362 372 382 392 402 412 422 432 442 452

MOTIF 6 MOTIF 7 MOTIF 8 MOTIF 9 MOTIF 10 MOTIF 11

PppA RHHNRA---N---VVRVAVGV---HEQLE---LSEAAH-VL-PQ-D-SFLI---CSQGL-RTAD-D-SKLRDV---LSE-SD-PYAV---VRSIVELGLT-RGAP-D<16>NI
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----hhhh-----hhhh-----cccc-----cccc-----
 StpI RAYPSA-R---ALTRVGA---SEQLV---LDVLEL-VK-PQ-D-VTLL---CSQGL-YEELS-A-DALGRA---LSL-AS-PQVA---VERLFDGALS-GAAR-D<16>ch
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----hhhh-----hhhh-----cccc-----cccc-----
 PpH ENFPHK---N---VIVR-LGM---KENVQ---VDVSRVE-PQ-ED-D-VFLI---CSQGL-SQVST-D-AQDSEI---LQR-TPELEKAC-SQ-LIDMNA-AQGN-D<16>hh
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----cchh-----hhhh-----cccc-----cccc-----
 s111771 KVHFWR---H---VLPQLS---RQDLNF---EVVALL-AQ-PQ-D-TPMI---CSQGL-TEEV-D-MLIEKI---LIG-QGCGDQA-VQ-LIEEAG-AGGS-D<16>cc
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----cccc-----cccc-----cccc-----cccc-----
 PP2Ca SVMQKVY---NGSLAVSR-LADYKCVGKQPTTEQLVSE---PEVNDIER-SF-ED-D-QITIL---AC-DGI-MDVWG-N-EEICDFVRSR-LEV-TDDLERVC-NE-VVDTCLY-RGSR-D<16>cc
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----cchh-----hhhh-----cccc-----cccc-----
 PpP SIKPHINILDKNNKNIILT-DGTA---FISHSAFELLSSSDFSTSAQRIAQVQVCGANVNASVGIY-NQDYE-NSLNSHKDYGVELWDP-RGN-LHM-WGQYPAQNYFSQIVDQDQK-EPSPI-D<16>NLIL
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----cccc-----cccc-----cccc-----cccc-----
 YghJ DFHDFPA-GISVEHLS-YGDLD---PQKMLP---LILNGFYVYQ-VGN-D-PYATP---LPL-TE-KPLTQQGVIDLAYLNKGSVLIEMVMSLKEES---ASGFVLLDA-AGLSML-D<16>NK
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----cccc-----cccc-----cccc-----cccc-----
 SpoII QNSPHN---HG-YSRVLPSS---TKSYVST---GAARAAKQGLVSG-D-SYSEELQARKYAAAE-DMGARARHP-ENETIKL---LEKILESDIEKIAIKTINSLSLRTDIEYST-D<16>IIML
 Predator cc-----cccccc-----ccc-----ccc-----ccc-----ccc-----cc-----cc-----ccc-----ccc-----hhhh-----hhhh-----cccc-----cccc-----

Figure 12. Multiple alignment of the PP2C superfamily showing the 11 conserved motifs. The alignment was constructed using CLUSTALW 18.1 (Thompson *et al.* 1994) with PP2C α , *Homo sapiens* protein phosphatase 2C isoform α , Accession No. P35813 as the prototype. The first column contains the protein code: PppA, *Pseudomonas aeruginosa* protein phosphatase A, Accession No. NP_248765; PphI, *Myxococcus xanthus* protein phosphatase I, Accession No. ABF93082; sll1771, *Synechocystis sp.* PCC6803 protein phosphatase A, Accession No. BAA17671; PrpZ, *Salmonella typhi* strain CT18 protein phosphatase 2C, Accession No. CAD06946; YghJ, *Escherichia coli* predicted lipoprotein, Accession No. Q46837; SpoIIE, *Bacillus subtilis* stage II sporulation protein E, Accession No. P37475. Residues conserved between proteins are shown in red, residues with high homology are shown in blue, hydrophobic residues are shown in green. The arrow indicates an arginine thought to be important in catalysis. The alignment is given as 11 motifs numbered along the top and indicated by a red box. Motifs: #1 (D), #2 (DG), #4 (G-T), #5 (GD), #6 (G), #7 (P), #8 (DG), #11 (DN). The secondary alignment was constructed independently for each protein using the PREDATOR alignment tool (Frishman and Argos, 1996). The secondary sequence prediction is immediately beneath each protein: H/h=Helix, e=Sheet, c=Coil.

In Gel Assay to Test for PP2C Activity of the YghJ protein-

The 4-methylumbelliferyl phosphate (MUP) substrate fluoresces after the substrate is dephosphorylated. So MUP was used in an in-gel format for *in situ* detection of PP2C activity. The in-gel assay was used to determine if the YghJ protein possessed PP2C activity. YghJ was expressed from the high copy number plasmid called PCB2YghJ and the gene was under control of its native promoter. The positive control was the known PP2C protein Peptidylprolyl Isomer (PIAI) [Accession# AEV61503.1] from *Pseudomonas aeruginosa*. Protein lysates were prepared and separated by native polyacrylamide gel electrophoresis. Then proteins were renatured in the gel before being incubated with MUP for 15 minutes. Results are seen in Figure 13. The fluorescent bands seen in each of the *E. coli* samples are alkaline phosphatase. No other fluorescent bands are seen in the strain with the pCB2YghJ plasmid which indicates that YghJ may not possess PP2C activity. The fluorescent band seen in the *P. aeruginosa* sample is the 20.1 kDa PpiA protein.

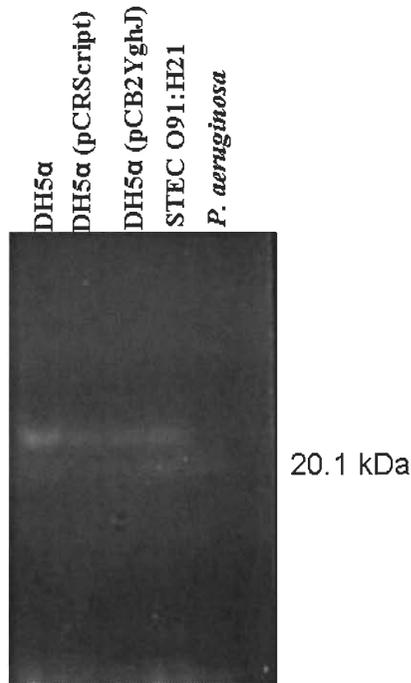


Figure 13. Assay for PP2C activity of the YghJ protein. Samples were incubated with the fluorogenic substrate MUP for 15 minutes before being analyzed. The positive result in Lane 5, indicated by the arrow, contains the PPIA protein from *P. aeruginosa*, a known PP2C.

Analysis of a YghJ null mutant

Construction of the yghJ null mutant-A knockout of the *yghJ* gene in the *Escherichia coli* strain BW25113 was obtained from the Keio collection (Baba et al., 2006). *E. coli* BW25113 is related to *E. coli* strain K12 substrain MG1655; both strains carry homologs of *yghJ*. The *E. coli* K12 DH α is also related to BW25113 and MG1655 but carries a variant of *yghJ* which contains a sizable deletion (Lukjancenko et al., 2010). The Keio collection mutants were created using the FLP recombinase system. PCR products containing a kanamycin resistance cassette were flanked by FLP recognition targets around 50-bp homologies to the desired chromosomal sequences. These constructs were inserted into the genome using the Lambda red recombinase system

(Datsenko and Wanner, 2000). The cassette was then excised using the FLP recombinase. This procedure created an in-frame deletion of the gene from the second codon removing intervening DNA except the terminal seven codons from the 3'-end of the gene. The start codon and the translational signal for any downstream genes are left intact.

Capacity of the YghJ knockout mutant to metabolize different carbon sources-To determine a phenotype for the *yghJ* deletion mutant, growth studies were done in minimal media with a specified carbon source. Bacteria were grown in batch culture with the sole carbon source indicated. A closed system was used, meaning that the only sugar available to the bacteria was added at the initial inoculation, and was depleted over the course of the study. Figure 14 shows the growth rate of the wild type BW25113 compared to the YghJ null mutant in M9 minimal media supplemented with 0.1% glucose over 24 hours. The YghJ null mutant exhibits no significant difference in growth rate with glucose as the sole carbon source. A lag of approximately 5 hours was observed before an exponential growth phase began which lasted for 8 hours. Stationary phase lasted until approximately 20 hours, before the bacteria began to die; likely due to the depletion of the glucose. Other common carbon sources were also tested. Figure 15 shows a growth curve comparing the wild type strain and the YghJ null mutant over 12 hours in M9 minimal media containing 0.1% fumarate, a 4-carbon sugar. No significant difference in growth rate was observed.

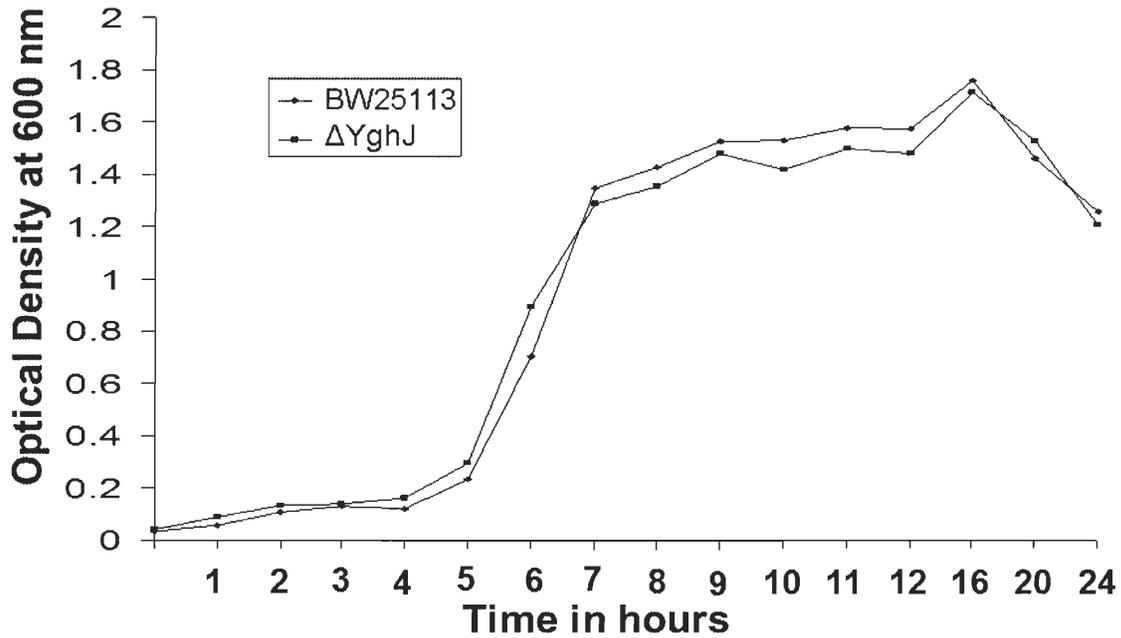


Figure 14. The YghJ null mutant exhibits a normal growth phenotype when grown on M9 minimal media containing 0.1% glucose as the sole carbon source. Strains were grown at 37° C with shaking.

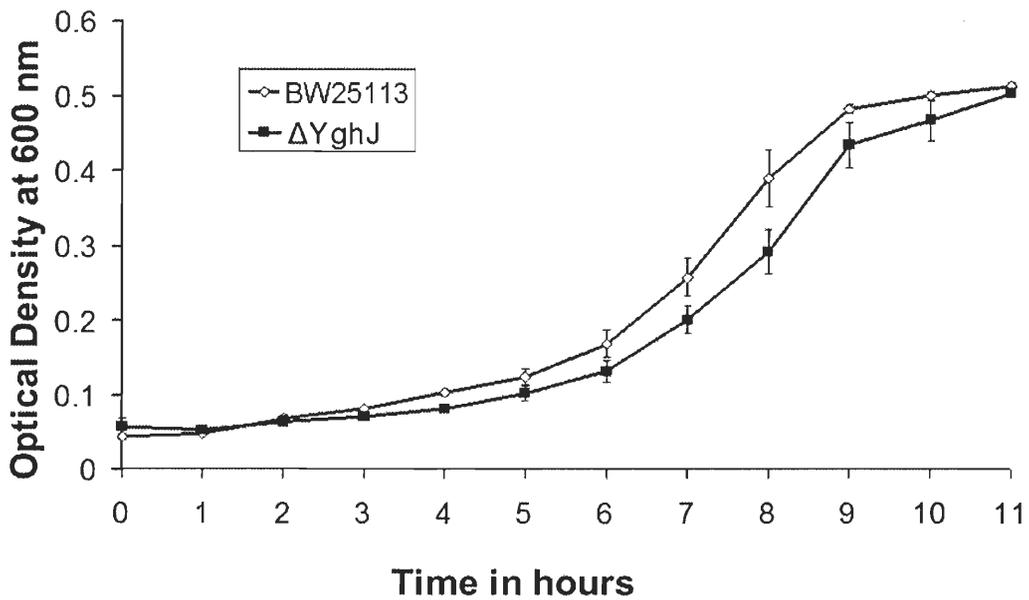


Figure 15. The YghJ null mutant exhibits a normal growth phenotype when grown on M9 minimal media with 0.1% fumarate as the sole carbon source. Strains were grown at 37° C with shaking. Error bars indicate standard error of the mean

The mutant was tested for its ability to grow on 0.1% succinate (Figure 16). The YghJ null mutant grew as well as the wild type strain did on fumarate and succinate but the overall cell mass was significantly reduced when bacteria were grown on these poor carbon sources. For example, lag phase was approximately 5 hours long and log phase lasted for only about 4 hours before stationary phase was achieved.

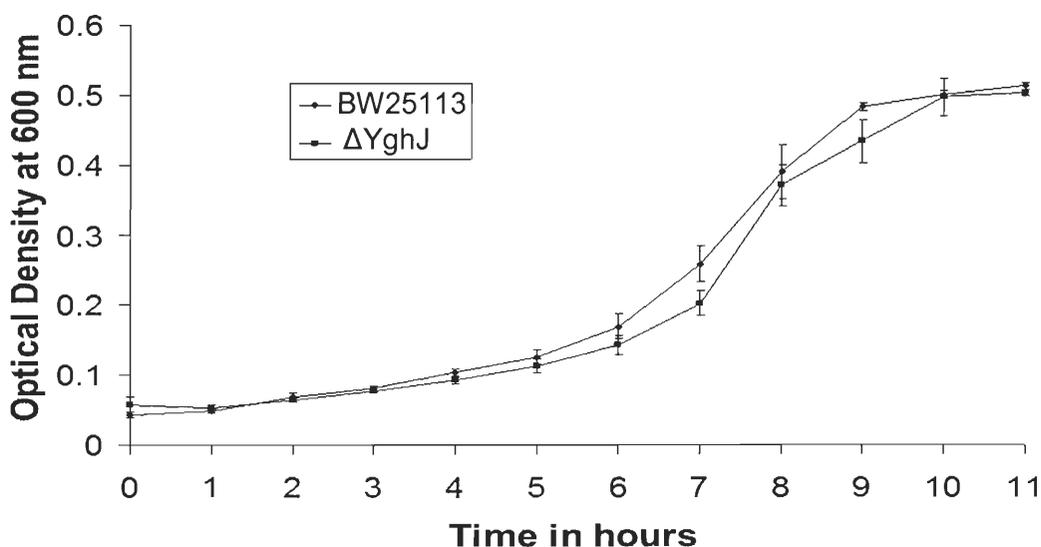


Figure 16. The YghJ null mutant exhibits a normal growth phenotype when grown on M9 minimal media containing 0.1% succinate as the sole carbon source. Strains were grown at 37° C with shaking. Error bars indicate standard error of the mean.

In contrast, the *yghJ* null mutant grown in M9 minimal media with 0.1% malate as the sole carbon source exhibited a severe growth deficiency. The mutant strain grew to an OD₆₀₀ nm of ~0.2 after 4 hours and the cell mass remained stationary at that level for 24 hours. The wild type strain grew at the normal rate, entering exponential growth phase

between 4-5 hours and stationary phase after 9 hours. The cell mass of the population began to decline after 20 hours (Figure 17).

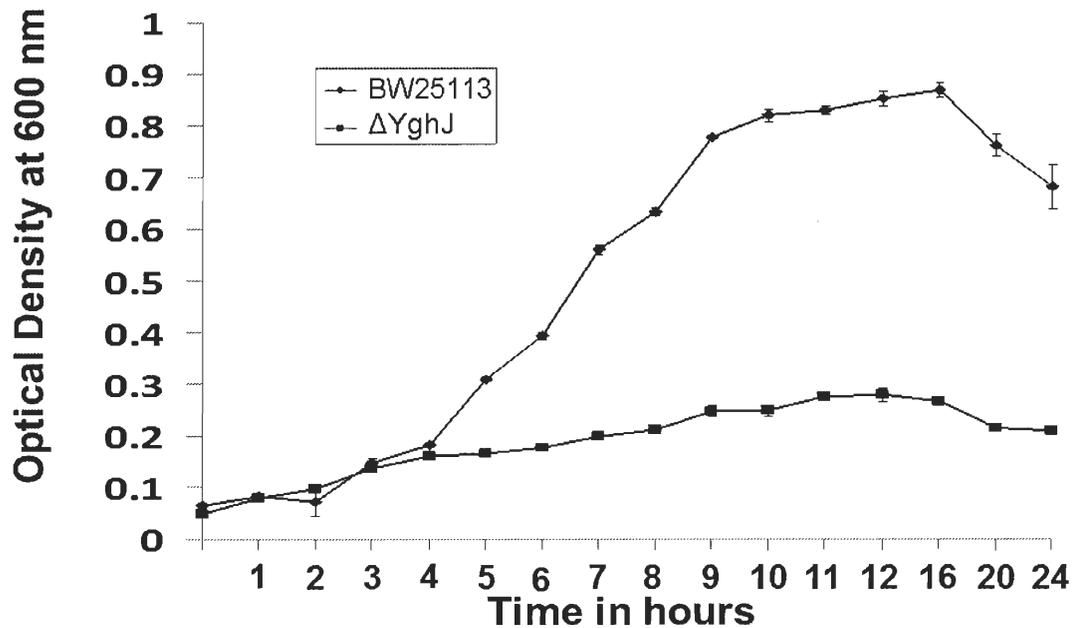


Figure 17. The YghJ null mutant exhibits a deficient growth phenotype when grown on 0.1% L-malate as the sole carbon source. Strains were grown at 37° C with shaking. Error bars indicate the standard error of the mean.

Complementation of the yghJ null mutant-The YghJ null mutant of BW25113 was complemented by replacement of the YghJ protein by expression of the protein *in trans* from the pCRB2YghJ plasmid. This clone contained the full length *yghJ* gene as well as the intergenic space between *yghJ* and the upstream *glcA* gene and the downstream *pppA* gene. No induction was used because the pCRB2YghJ construct contains the native promoter region, as well as any other regulatory regions present

outside the open reading frame upstream of the gene itself. Figure 18 is representative the complementation results obtained. The lag phase of the wild type strain was approximately 5 hours and exponential phase lasted for 8 hours, reaching an OD_{600 nm} of ~0.7. The YghJ null mutant did not obtain any significant growth; growth was blocked at an OD_{600 nm} of ~0.2. The growth defect in 0.1% malate was complemented by expression of YghJ from plasmid pCB2yghJ. The complemented mutant grew significantly better than the YghJ null mutant, after a lag of 5 hours reaching an OD_{600 nm} of ~ 0.6 after 12 hours; exponential growth phase lasted for about 11 hours.

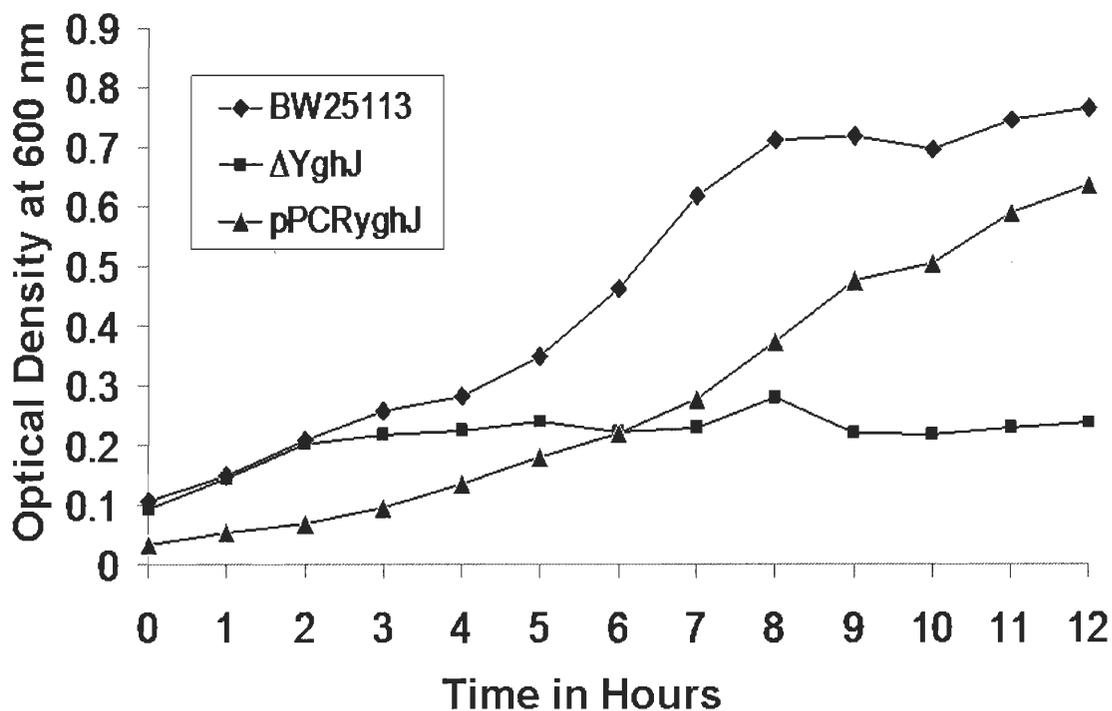


Figure 18. The YghJ null mutant is able to partially recover growth in 0.1% L-malate when complemented with the yghJ gene expressed from a high copy number plasmid. Strains were grown at 37° C with shaking. The strain containing pPCRYghJ was grown in ampicillin at 100 µg/ml.

Expression and attempts to purify the YghJ-6xHis fusion protein

Expression of YghJ from an inducible vector-

pBAD202D/TOPO expression vector

The open reading frame of the *yghJ* gene was cloned into the arabinose inducible expression vector pBAD202D/Topo vector (Invitrogen). The plasmid contains the arabinose inducible *araC* promoter. Figure 19 shows SDS-PAGE analysis of the induction study. Cultures were induced for the indicated amount of time with the different concentrations of arabinose. No YghJ was seen in any of the samples. However, a low molecular weight band of approximately 14 kDa was visible on the gel. This band matches the size of the thioredoxin tag present on the carboxy-terminus of the YghJ protein when expressed from the pBAD202D/Topo vector. The protein may have been degraded due to improper folding leading to proteolytic cleavage of the recombinant protein.

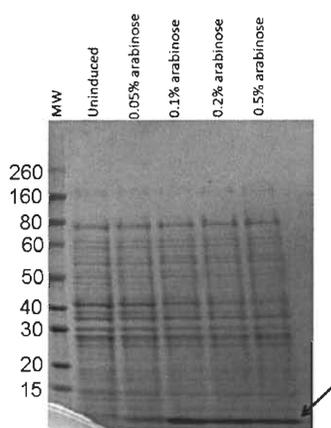


Figure 19. SDS-PAGE analysis of YghJ expressed from the arabinose inducible Pbad202D/Topo vector. Samples were separated in in a 4-12% Novex Bis-Tris gel. All samples were induced for 6 hours with the indicated concentration of arabinose and grown at 37 °C with shaking in the presence of 30 µg/ml kanamycin The small molecular weight band is indicated by the arrow.

Expression of YghJ from the pET21a vector in strain BL21(DE3)

The open reading frame of the *yghJ* gene was cloned into the inducible expression vector pET21a. This plasmid contains the IPTG inducible T7 promoter. Cultures were grown to an OD₆₀₀ of 0.4 then induced with IPTG. Figure 20 shows SDS-PAGE analysis of an induction study using various IPTG concentrations. Cultures were induced with 50 mM, and 100 mM IPTG and grown for 8 hours; samples (1 ml) were taken at 2, 4, 6, and 8 hours. The YghJ protein expressed successfully with both concentrations of IPTG as protein was detected within the first 2 hours and every hour thereafter.

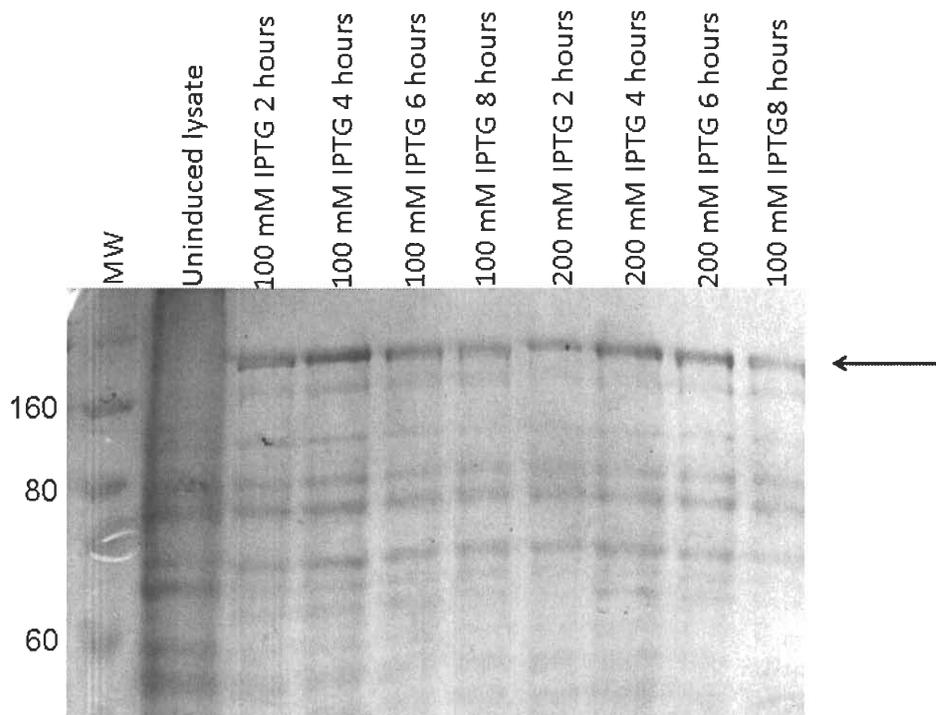


Figure 20. SDS-PAGE analysis of YghJ expressed from the IPTG inducible pET21a vector. Samples were separated in an 8% Bis-Tris polyacrylamide gel. Samples were induced with the indicated IPTG concentration for the indicated time at 37 °C with shaking. The arrow indicates YghJ.

Purification of YghJ using metal affinity chromatography-

Pierce B-PER 6XHis Fusion Protein Purification Kit

The Pierce B-PER 6XHis Fusion Protein Purification Kit uses the Bacterial Protein Extract Reagent to lyse cells and extract soluble proteins. After removing cellular debris by centrifugation, the lysate was incubated with nickel chelate agarose to bind the His-tag YghJ recombinant protein. Agarose beads were washed by centrifugation to remove superfluous unbound proteins before the YghJ-6xHis protein was eluted using imidazole. Figure 21 shows the result of the purification attempt. The protein was eluted in 4 fractions that were pooled before being analyzed by SDS-PAGE. Lane 1 contains the molecular weight marker. Lane 2 contains the induced lysate before reaction with the nickel chelate agarose. The YghJ-6xHis protein is present at high levels, indicating a successful induction. Lane 3 contains the flow-through obtained after the YghJ-6xHis recombinant protein was reacted with the nickel chelate agarose. A large amount of YghJ is present in the flow-through, which means that YghJ-6xHis did not stick well to the nickel chelate agarose. Lanes 4-6 are the column washes. A small amount of YghJ-6xHis is visible in the first wash, but no YghJ-6xHis is visible in the proceeding washes. Lanes 7-10 are the elution fractions. Proteins were eluted into a 1 ml volume of elution buffer and 20 μ l were loaded onto the gel. No YghJ-6xHis is visible in any of the elution fractions because the protein did not stick to the nickel chelate agarose.

To potentially improve the stability of the YghJ-6xHis protein, the antioxidant 2-mercaptoethanol was added to the lysis buffer, wash buffer, and elution buffer before purification. The lysate was purified as above and the results can be seen in Figure 22.

Lane 2 contains lysate that with the induced YghJ-6xHis protein. Lane 3 contains the initial flow-through. Again, a large amount of YghJ-6xHis is seen in the flow-through but none is seen in any of the washes or elution fractions.

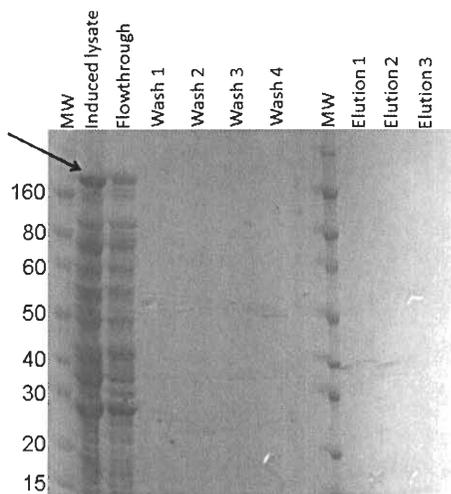


Figure 21. SDS-PAGE analysis of an attempted purification of YghJ-6XHis using the Pierce B-PER 6XHis protein purification kit. Samples were separated in a 4-12% Novex Bis-Tris polyacrylamide gel. The arrow indicates YghJ. All proteins appear to be present in the flowthrough.

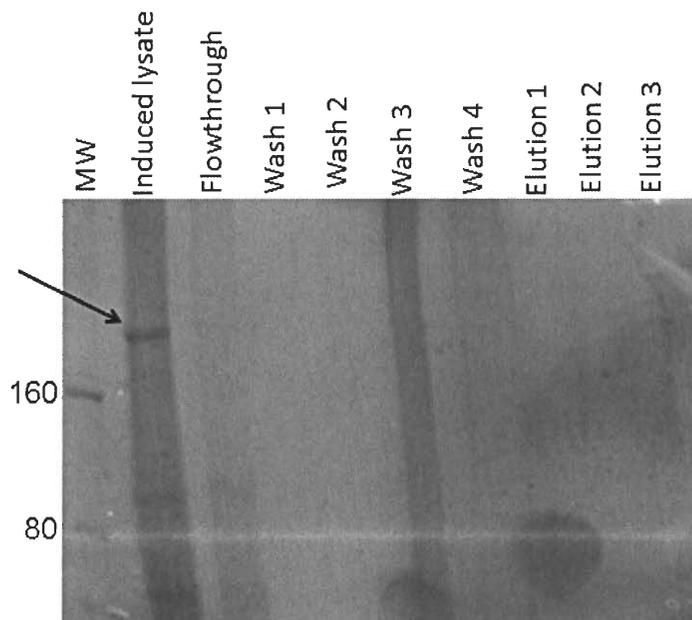


Figure 22. SDS-PAGE analysis of an attempted purification of YghJ-6XHis using the Pierce B-PER 6XHis protein purification kit containing 5 mM β -mercaptoethanol.

Samples were separated using an 8% Bis-Tris polyacrylamide gel. The arrow indicates YghJ.

Clonetech HisTALON Gravity Purification Columns

Clonetech HisTALON Gravity Purification Columns replaced the other matrices to see if YghJ-6xHis would stick better to this column. These columns used a cobalt chelate agarose as opposed to the nickel chelate agarose in the Pierce kit which is reported to be more effective. The protein was reacted with the column then washed using gravity rather than centrifugation. We first attempted to purify YghJ-6xHis using the B-PER buffer from the Pierce kit above. A 100 ml culture was lysed overnight in B-PER buffer before being applied to the column. The column was washed four times in sodium phosphate wash buffer before being eluted with 100 mM imidazole. Figure 23 shows an SDS-PAGE gel analysis of the attempted purification. Lane 1 contains the induced YghJ fraction. Lane 2 contains the flow-through from the column. Lanes 3-5 contain the washes. Lanes 6-8 contain elution fractions. Again, most of the YghJ protein is seen in the initial flow-through fraction, meaning that the YghJ-6xHis fusion protein did not bind to the column. However, a smaller protein is visible in the elution fractions, meaning either some non-specific binding occurred or that the protein of interest was breaking down on the column.

To ensure that YghJ-6xHis was fully denatured and no other proteins were binding and blocking access to the 6XHis epitope, the lysate and buffers were treated with 0.1% Triton-X 100 detergent. Purification was performed as above, and the results can be seen in Figure 24. Lane 1 contains uninduced sample, Lane 2 contains induced YghJ-6xHis

and Lane 3 contains the flow-through from the column. An almost equal amount of YghJ is seen in the induced fraction as the flow-through, indicating that YghJ-6xHis did not stick to the column. The elution fractions in lanes 8-10 again contain the smaller unknown protein.

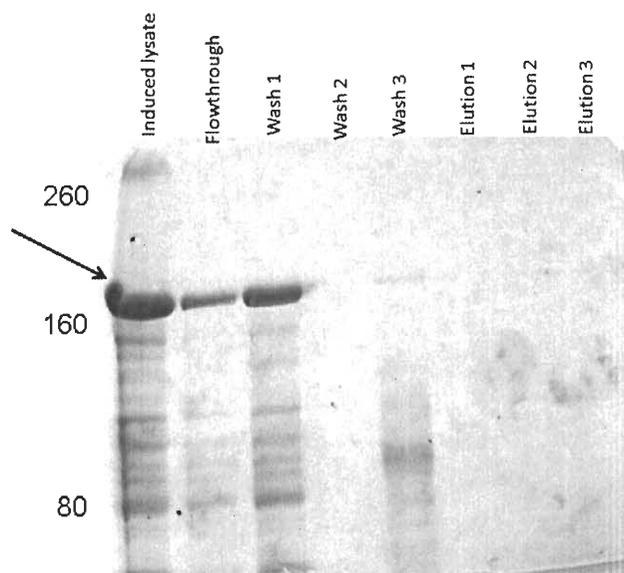


Figure 23. SDS-PAGE gel analysis of a purification performed using the Clontech HisTALON Gravity Purification Columns. The samples were lysed using the Pierce Bacterial Protein Extraction Reagent. Samples were separated in an 8% Bis-Tris polyacrylamide gel. The arrow indicates YghJ.

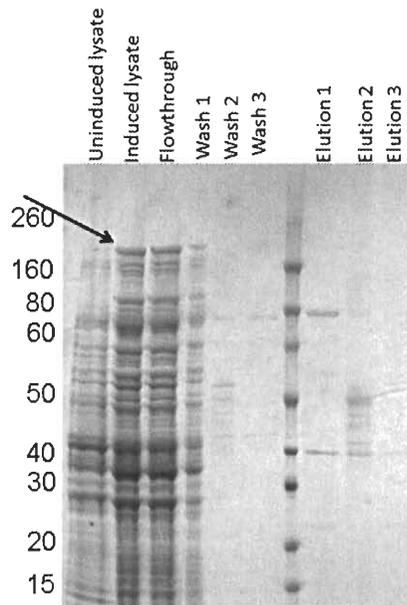


Figure 24. SDS-PAGE gel analysis of YghJ purification using the Clonetechn HisTALON Gravity Purification Columns. Samples were separated in a 4-12% Novex Bis-Tris polyacrylamide gel. Samples were lysed in the the Pierce B-PER buffer containing 0.1% Triton-X 100. The arrow indicates YghJ.

The Talon xTractor buffer can be used to extract membrane bound proteins. The manufacturer recommends the use of 1mg/ml lysozyme combined with 1 ug/ml DNase I for the extraction of membrane bound proteins. Figure 25 shows the results of a purification using these recommended reagents. Again, a similar amount of YghJ-6xHis is seen in the lysate and flow-through, while the unknown smaller protein is present in the elution.

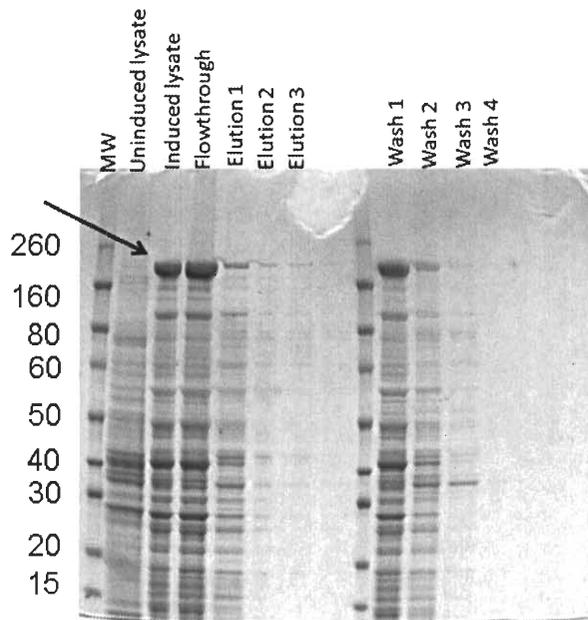


Figure 25. SDS-PAGE gel analysis of YghJ purification using the Clontech HisTALON Gravity Purification Columns. Samples were separated in a 4-12% Novex Bis-Tris polyacrylamide gel. Samples were lysed in the Clontech xTractor buffer containing 1 $\mu\text{g/ml}$ DNase I and 1 mg/ml lysozyme. The arrow indicates YghJ.

Invitrogen Dynabeads His-Tag Isolation Beads

The Invitrogen Dynabeads are a cobalt magnetic chelate that binds histidine-tagged proteins. Beads are incubated in the lysate and removed using a magnetic system. A 100 ml lysate was prepared using the Talon xTractor buffer containing lysozyme and DNase I. The lysate was incubated with the Dynabeads for 2 hours before the beads were removed from the lysate using a magnet and washed 3 times using a sodium phosphate buffer. Bound proteins were eluted from the beads using a sodium phosphate buffer containing 100 mM imidazole. Figure 26 shows the results of this purification. Lane 5 is the induced lysate containing the YghJ-6xHis protein before any purification. Lane 6 contains the lysate which was removed from the beads after the 2 hour incubation period. Lanes 9-11 contain the three wash fractions, while lanes 1-4 contain the elution. Little to

no YghJ-6xHis is seen in the elutions, meaning that YghJ-6xHis did not successfully bind to the Dynabeads. The uninduced fraction is in the final lane.

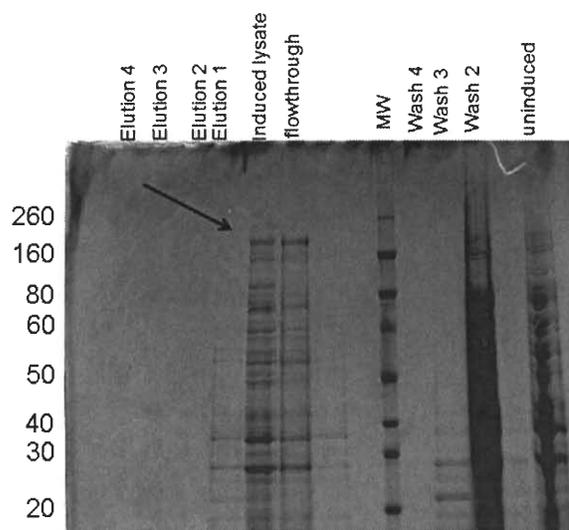


Figure 26. SDS-PAGE gel analysis of YghJ purification using Invitrogen Dynabeads His-Tag Isolation Beads. Samples were separated in a Novex 4-12% Bis-Tris polyacrylamide gel. Samples were lysed in the the Clonetech xTractor buffer containing 1 μ g/ml DNase I and 1 mg/ml lysozyme. The arrow indicates YghJ.

Immunoblot to detect the YghJ-6xHis recombinant protein- To determine if the 6XHis tag was available for binding to the nickel or cobalt columns we attempted to use for purification, an immunoblot was done using two antibodies specific to the 6XHis tag. An induced lysate was separated by SDS-PAGE then transferred onto nitrocellulose and treated with antibody specific to the 6XHis epitope. No protein was detectable with the commercially available India anti-His HRP conjugated antibody (Pierce) or the Covance anti-His antibody (Covance). A control immunoblot was performed using the same procedure with the YhcM-6XHis tagged fusion protein. The YhcM-6XHis protein is another protein we discovered and cloned into the pET21a expression vector and purified using the Dynabeads. Both antibodies successfully

detected the 47 kDa fusion protein (Data not shown); which shows that the reagents are functional and the techniques used were appropriate.

Localization of YghJ

Membrane fractionation of YghJ- To test the hypothesis that YghJ is localized to the inner membrane by virtue of its predicted inner membrane localization signal sequence, the inner and outer membranes of the bacteria were separated by ultracentrifugation and the inner membrane proteins solubilized with Triton-X 100. The results are shown in Figure 27. The strain BL21(DE3)carrying pETMGORFYghJ was used. The strain was induced for 6 hours before separation was done. The uninduced sample was untreated with IPTG. After separation by ultracentrifugation, the majority of the induced YghJ is in the inner membrane fraction, indicating that YghJ preferentially localizes to the inner membrane (lane 2). Small amounts of YghJ were detected in the cell fraction liberated after sucrose treatment and sonication; this fraction contains cytoplasm and periplasmic proteins. The small amount of YghJ seen is probably nascent YghJ protein that had not anchored to the inner membrane (lane 6). No YghJ was visible in the outer membrane fraction (lane 4). As expected, YghJ was not seen in any of the fractions from the uninduced samples.

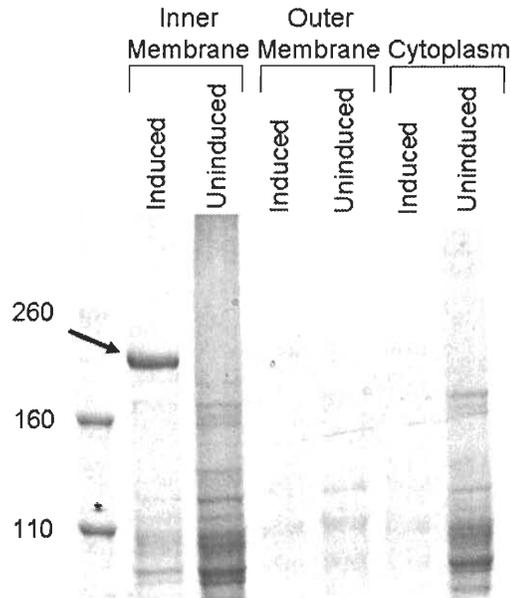


Figure 27. SDS-PAGE gel analysis of membrane fractionation of YghJ. Samples were separated in an 8% Bis-Tris polyacrylamide gel Lane 1: Molecular weight marker Lane 2: inner membrane fraction induced with IPTG. Lane 3: Uninduced inner membrane fraction. Lane 4: Uninduced outer membrane fraction Lane 5: Induced soluble fraction Lane 6: Uninduced soluble fraction The arrow indicates the position of YghJ.

Globomycin Treatment of YghJ- The putative signal sequence spanning the first 23 amino acids of YghJ includes a predicted lipobox sequence. This sequence is cleave by the signal peptidase II LspA in the periplasm before being lipidated and the proteins anchored in the membrane. To test the hypothesis that YghJ is lipidated, the protein was treated with the antibiotic globomycin. The antibiotic globomycin specifically inhibits the signal peptidase II enzyme LspA, which cleaves the nascent prolipoprotein at its diacylglycerol modified cysteine residue (Wu et al., 1983) leaving the first 23 amino acids attached to the protein and stops the addition of the final acyl group onto the protein. This results in a longer protein which will not anchor in the

membrane. After treatment with nonlethal concentrations of globomycin, two species of all lipoproteins in the cell are present including YghJ.

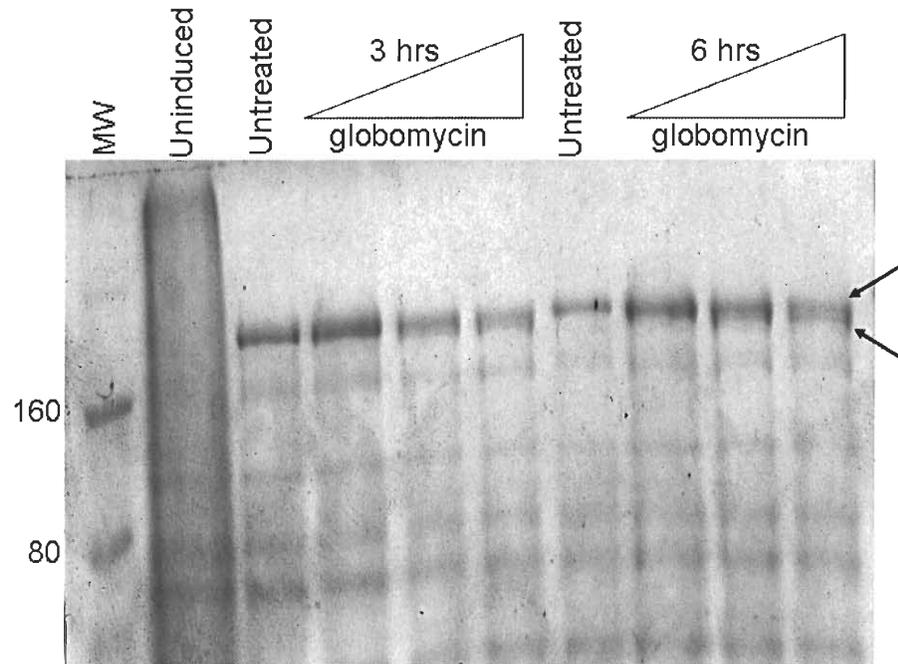


Figure 28. SDS-PAGE analysis of globomycin treated BL21 (DE3) expressing the YghJ protein from the induced pET21a vector. Samples were separated in an 8% Bis-Tris polyacrylamide gel. Samples were treated with the increasing concentrations of globomycin: 50 mM, 100 mM, and 200 mM for 3 or 6 hours. The arrows indicate the two variants of the YghJ protein.

E. coli strain BL21(DE3)pETMGORFYghJ.7 containing the *yghJ* gene on the IPTG inducible pET21a vector was used to perform the experiment. The strain was induced for one hour before treatment with the indicated concentration of globomycin, then grown for three hours and analysed by SDS-PAGE (Figure 28). The uninduced sample contained no YghJ (lane 2) and the induced sample which was NOT treated with

the antibiotic contained YghJ of the proper molecular weight (lane 3). As the concentration of globomycin increased, a second YghJ variant became visible on the gel, most evident in the 200 mM globomycin treated sample (lane 6). The increasing amount of the smaller variant indicates that YghJ is nonlipidated in presence of globomycin. The generation of the smaller variant during globomycin treatment indicates that under normal conditions the YghJ protein is lipidated.

Comparison of growth of STEC O91:H21 and STEC O157:H7 in Minimal M9 Media with L-malate as Sole Carbon Source

The YghJ protein was initially identified through a proteomics screen to identify differential protein expression in the LEE-positive STEC O157:H7 strain 86-24 and the LEE- negative STEC O91:H21 strain B2F1. The LEE-positive strain lacked the *yghJ* gene but the LEE-negative strain B2F1 carried the *yghJ* gene and expressed the protein under normal laboratory conditions. Once we knew the YghJ null mutant was deficient for growth on L-malate, we tested the wild type LEE-positive STEC O157:H7 for its capacity to utilize L-malate as the sole carbon source. See results in Figure 29.

YghJ positive STEC O91:H21 grew much more efficiently on L-malate compared to *yghJ*-negative STEC O157:H7. STEC O91:H21 reached exponential growth phase after 4 hours and entered stationary phase after only 8 hours. STEC O157:H7 had an extended lag phase of 13 hours delay as compared to STEC O91:H21. STEC O157:H7 did not reach stationary phase until 22 hours. STEC O157:H7 recovered after the lag phase and reached a similar cell mass at stationary phase once growth resumed.

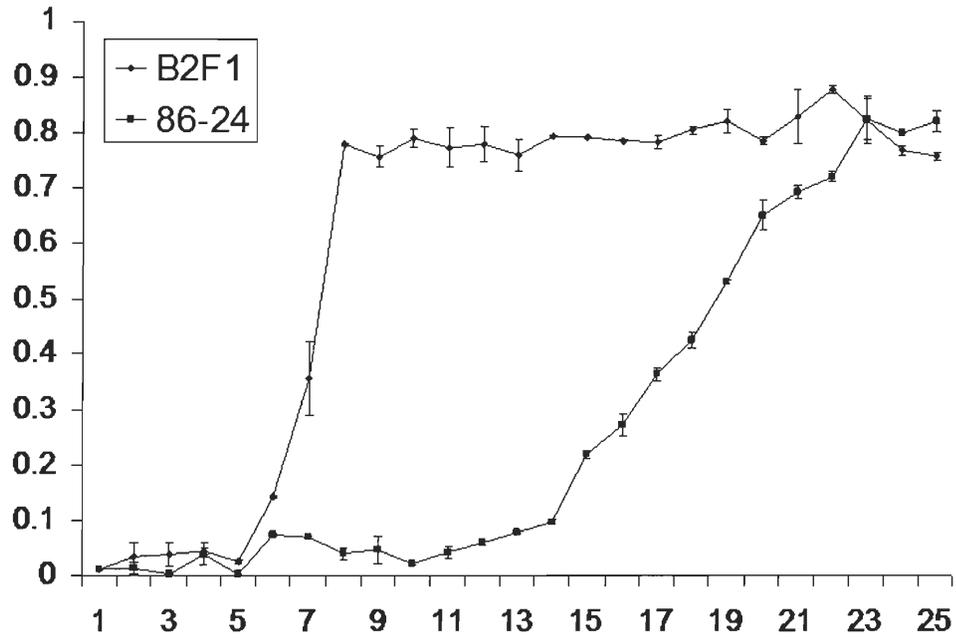


Figure 29. The YghJ negative pathogenic STEC O157:H7 strain 86-24 exhibits a deficient growth phenotype compared to the YghJ positive STEC O91:H21 strain B2F1 when grown on 0.1% L-malate as the sole carbon source. Error bars indicate the standard error of the mean.

Expression of Enterohemolysin

EhxA production on SRBC agar- Detection of enterohemolysin activity by wild-type strains of LEE-positive and LEE-negative STEC organisms requires that the sheep red blood cells added to the agar are pre-washed and the agar plates are incubated overnight until small turbid zones of hemolysis around colonies are evident. Clear hemolytic halos were seen around the colonies of 34.7, and these zones of hemolysis were much larger than those associated with B2F1. The amount of *EhxA* activity detected in Figure 30 is greater in 34.7 than that detected for the parental strain.

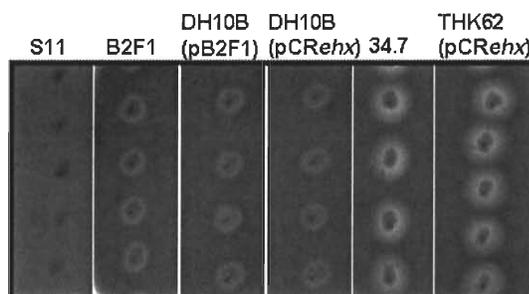


Figure 30. Comparison of enterohemolytic phenotypes on SRBC agar. Bacteria tested are individually labeled above each panel. Cultures were applied to tryptose agar that contained 5% washed sheep red blood cells and 10 mM CaCl then incubated at 37°C in a 5% CO₂ atmosphere for 18 hours.

To determine if transcriptional regulation of the *ehx* operon required chromosomal genes unique to STEC O91:H21, the large plasmid (pB2F1) was transferred into the non-pathogenic and non-hemolytic *E. coli* K-12 laboratory strain DH10B. As seen in Figure 30, the hemolytic phenotype of DH10B (pB2F1) is comparable to that noted for wild-type STEC O91:H21 as indicated by the similarity in

size of the turbid hemolytic halos around STEC O91:H21 and DH10B (pB2F1). These results suggested that the *ehx* operon present on the large plasmid is regulated in a like manner, whether in the STEC O91:H21 or DH10B background.

To determine if an increase in copy-number of *ehxCABD* would affect the amount of EhxA produced, we cloned the *ehxCABD* operon, including 88 bp of DNA upstream of the *ehxC* start codon that we surmised contains the promoter region, onto the high-copy number vector pCRScript Amp SK (+) to create pCR*ehx*. This plasmid construct was then moved into DH10B. The level of hemolysis was comparable to that seen for DH10B (pB2F1) and STEC O91:H21 (Figure 30). These results suggest that the expression of the *ehx* operon is under strict control since the presence of multiple copies of the operon, as in the case of DH10B (pCR*ehx*), within the bacterial cell did not result in a detectable increase in hemolytic activity. These findings further implied that there were no other genes present on the plasmid of pB2F1 required for regulation of *ehx* and that a chromosomal factor found in both strain STEC O91:H21 and DH10B was involved in *ehx* regulation.

We then investigated whether HNS might be the major factor involved in *ehx* regulation. To accomplish this aim, we transferred the multi-copy number plasmid pCR*ehx* into strain THK62. This is a non-pathogenic, non-hemolytic laboratory strain of *E. coli* that contains a tetracycline antibiotic insertion that disrupts the functionality of *hns* (Donato et al., 1997). The *hns* mutant THK62 (pCR*ehx*) reverted from non-hemolytic to a hemolytic phenotype, as evidenced by large and clear hemolytic halos on SRBC agar (Figure 30). The hemolytic activity detected is comparable to that noted for 34.7. Taken

together, the data indicate that lack of the HNS protein resulted in de-regulation of *ehx* expression and that plasmid-copy-number does not play a significant role in how much active hemolysin is detected on SRBC agar.

Comparison of Enterohemolysin levels in the HNS knockout mutant 34.7-

To determine whether more EhxA was made or whether more was secreted by strain 34.7, the relative amounts of EhxA protein present in the supernatants of cultures, within the cell and on the cell membrane were evaluated. Log-phase cultures of 34.7 and STEC O91:H21 were pelleted and supernatants recovered and set aside for analysis (supernatant was designated as [S]). Pelleted bacterial cells were re-suspended in fresh media and then disrupted by sonication. The resulting lysate that contained membrane, cytoplasmic and periplasmic proteins was concentrated by centrifugation and designated as pellet [P]. Proteins in the pellet lysates and the supernatant samples were precipitated with trichloroacetic acid. The negative control consisted of the plasmid-cured derivative, S11, of the parental strain STEC O91:H21 which is *ehx*-negative by virtue of the loss of the plasmid (Scott et al., 2003). The positive control strain was *E. coli* WAF100 (pSF4000) (Welch and Falkow, 1984). Plasmid pSF4000 carries the α -hemolysin operon (*hlyACBD*) originally cloned from the pathogenic urinary tract isolate of *E. coli* strain J96 (Welch et al., 1983). Previous reports indicate that HlyA is 110 kDa and EhxA is approximately 107 kDa (Felmlee T et al., 1985) (Schmidt et al., 1995).

Figure 31A shows the results of SDS-PAGE gel analysis when enterohemolysin was detected by colloidal blue stain. In Figure 31A the horizontal arrow indicates migration of the HlyA protein in 4-12% gradient gels whereas the star points to the

position of the *EhxA* protein. Lane 1 Figure 31A contains the HlyA protein present in the supernatant of WAF100 (pSF4000) while lane 2 represents membrane associated HlyA contained in the pellet fraction. Lane 2 Figure 31A contains the supernatant fraction recovered from 34.7 cultures. Lane 6 Figure 31A contains the supernatant fraction recovered from STEC O91:H21.

Comparison of *EhxA* recovered from the supernatant fractions of 34.7 and STEC O91:H21 revealed that 34.7 secreted more enterohemolysin than did STEC O91:H21; lanes 3 and 6 of Figure 31A respectively. It is typical that wildtype STEC secrete little enterohemolysin (Chart et al., 1998). No cell-associated *EhxA* was detected in pellet samples of 34.7 run on colloidal blue-stained SDS-PAGE gel (31A lane 4). A star marks the migration of *EhxA* detected in the pellet fraction of STEC O91:H21 (Figure 31A lane 7).

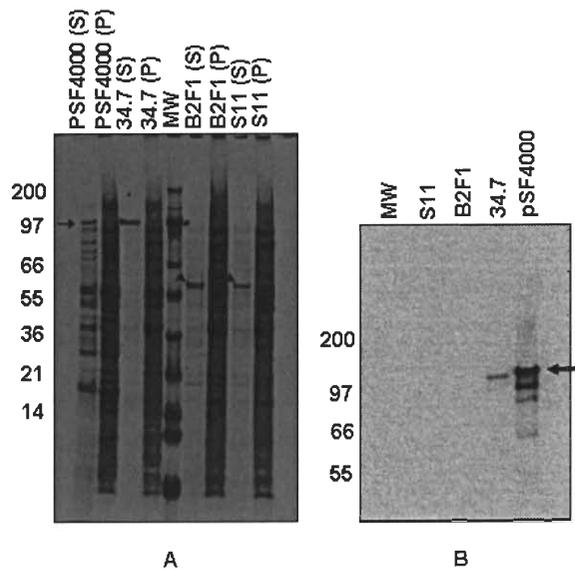


Figure 31. Analysis of representative levels of enterohemolysin produced by different strains. Samples marked with an (S) or (P) stands for supernatant and pellet; respectively. (A) SDS-PAGE gel comparison of enterohemolysin produced by wild-type strains versus the *hns* mutant. Various strains are listed across the top of the gel. The horizontal arrow points to the α -Hly protein which migrates at 110 kDa and the star points to the enterohemolysin present in both 34.7 and B2F1. The positions of the molecular weight markers in lane 7 are indicated. The arrowhead marks migration of FliC protein present in B2F1 and S11 but absence in 34.7. (B) Western blot analysis of extracellular enterohemolysin. All lanes contain supernatants recovered from samples analyzed by SDS-PAGE gel shown in panel A. Immunoblot of the supernatants were run at the same time as the SDS-PAGE gels. Molecular weight protein markers are indicated on the gel.

SDS-PAGE gel analysis suggests that the amount of *EhxA* secreted by 34.7 exceeded that of STEC O91:H21. Note that the level of *EhxA* produced and secreted by STEC O91:H21 is quite low and difficult to detect by colloidal blue stain. As expected no *Ehx* was detected in any of the plasmid-cured S11 samples (Figure 31A lanes 8-9). Bacterial cultures were equalized prior to SDS-PAGE gel and Western blot analysis. The protein profile of the *hns* mutant, 34.7, is somewhat different from the parental strain and is due to a change in protein expression as a result of the *hns* defect. This difference is

exemplified by the lack of FliC protein in 34.7. In contrast, the FliC protein is detected in the supernatant of STEC O91:H21 and S11. The migration of the FliC protein is denoted by an arrowhead in STEC O91:H21 and S11 samples; lanes 6 and 8 respectively. The absence of flagella in *E. coli* isolates that have knockouts in the *hns* gene has been noted by other investigators (Bertin et al., 1994).

Immunoblot analysis of EhxA secreted by the HNS null mutant, 34.7-
EhxA as well as HlyA belong to the repeat in toxin (RTX) family of toxins and hemolysins HlyA and ApxII are most closely related to *EhxA* (Bauer and Welch, 1996). Members of this family including the HlyA of *E. coli* and the adenylate cyclase-hemolysin (CyaA) of *Bordetella pertussis* show genetic relatedness and antigenic similarity. Some anti-CyaA monoclonal antibodies cross-react with *EhxA* within an invariant region of the protein called the nonapeptide repeat region (Bauer and Welch, 1996). Commercially available antibody CyaA(9D4):sc-13581 specific for the adenylate cyclase-hemolysin (CyaA) was shown to cross-react with *EhxA* (Bauer and Welch, 1996). Therefore, we used the CyaA(9d4):sc-13581 antibody to detect the presence of *EhxA* in culture supernatants. Previous reports indicate that the *EhxA* protein runs on SDS-PAGE gels as an ~107 kDa protein band (Brunder et al., 2006). The same supernatant samples analyzed by SDS-PAGE and shown in Figure 30A were used in Western blot analysis. In Figure 31B lane 5 the positive control HlyA protein known to migrate to 110 kDa molecular weight and correlates with the *EhxA* protein known to be of similar size (Figure 31B lane 4) (Felmlee T et al., 1985) (Schmidt et al., 1995). The

Western blot confirms the SDS-PAGE results that showed a band similar in size to HlyA which corresponds to *EhxA*. Taken together the SDS-PAGE gel and the Western blot analysis indicate that 34.7 produced and secreted more enterohemolysin than did the parental wild-type strain, STEC O91:H21. The small amount of *EhxA* secreted by STEC O91:H21 and detected with colloidal blue stain was not picked-up by the cross-reacting primary antibody (lane 3 of Figure 31B). As expected, no *EhxA* protein was noted in samples of plasmid-cured S11 (Figure 31B lane 2).

Relative enterohemolysin activity in culture supernatants of 34.7 versus STEC O91:H21- To quantify the relative amount of *EhxA* activity present in supernatants recovered from the *hns* mutant 34.7 as compared to STEC O91:H21, a standard hemolysis tube assay was done with supernatants recovered from log-phase cultures (Welch et al., 1983). The α -hemolysin protein expressed by WAF100 (pSF4000) served as the positive control strain for cytolytic activity and DH10B that contained the vector without the insert was the toxin-negative control (Figure 32A). With this standard hemolysis tube assay, no hemolytic activity was measurable in bacterial supernatants recovered from the wild-type strain STEC O91:H21. Likewise, culture supernatants from DH10B (pB2F1) which contained the large plasmid that encodes the *ehx* operon transferred from STEC O91:H21 was non-hemolytic in the tube assay. These results suggest that HNS repressed the level of enterohemolysin produced by STEC O91:H21 and DH10B (pB2F1). Similarly, other investigators have reported that STEC O157:H7 organisms secreted only trace amounts of enterohemolysin into the surrounding media during growth (Chart et al., 1998).

In contrast, in the HNS deleted background of strain 34.7, the level of enterohemolysin activity detectable in the supernatants increased considerably such that an undiluted sample of culture supernatant yielded approximately 70% lysis in a standard tube assay and activity persisted (at ~15% lysis) even when supernatants were diluted 1:8 as shown in Figure 29. Similarly, as depicted in Figure 32B, when plasmid pCRE ehx was moved into the *hns* knockout, THK62, the culture supernatant contained high levels of enterohemolysin activity (undiluted sample was capable of 90% lysis) comparable to that of 34.7. In sum, the results suggest that it is the presence of the HNS protein that represses transcription of the *ehx* operon. Furthermore, the multi-copy plasmid pCRE ehx carried by DH10B (wild-type for HNS) did not produce and transport enough enterohemolysin out of the cell to be detectable in the tube assay (Figure 32B). The tube assay data supports the SDS-PAGE gel and Western blot analyses of supernatants recovered from STEC O91:H21 and 34.7 (Figure 31A and 31B). The data indicates that the amount of *EhxA* protein secreted by STEC O91:H21 cultures is low when compared to 34.7 and suggests that the production of enterohemolysin is regulated.

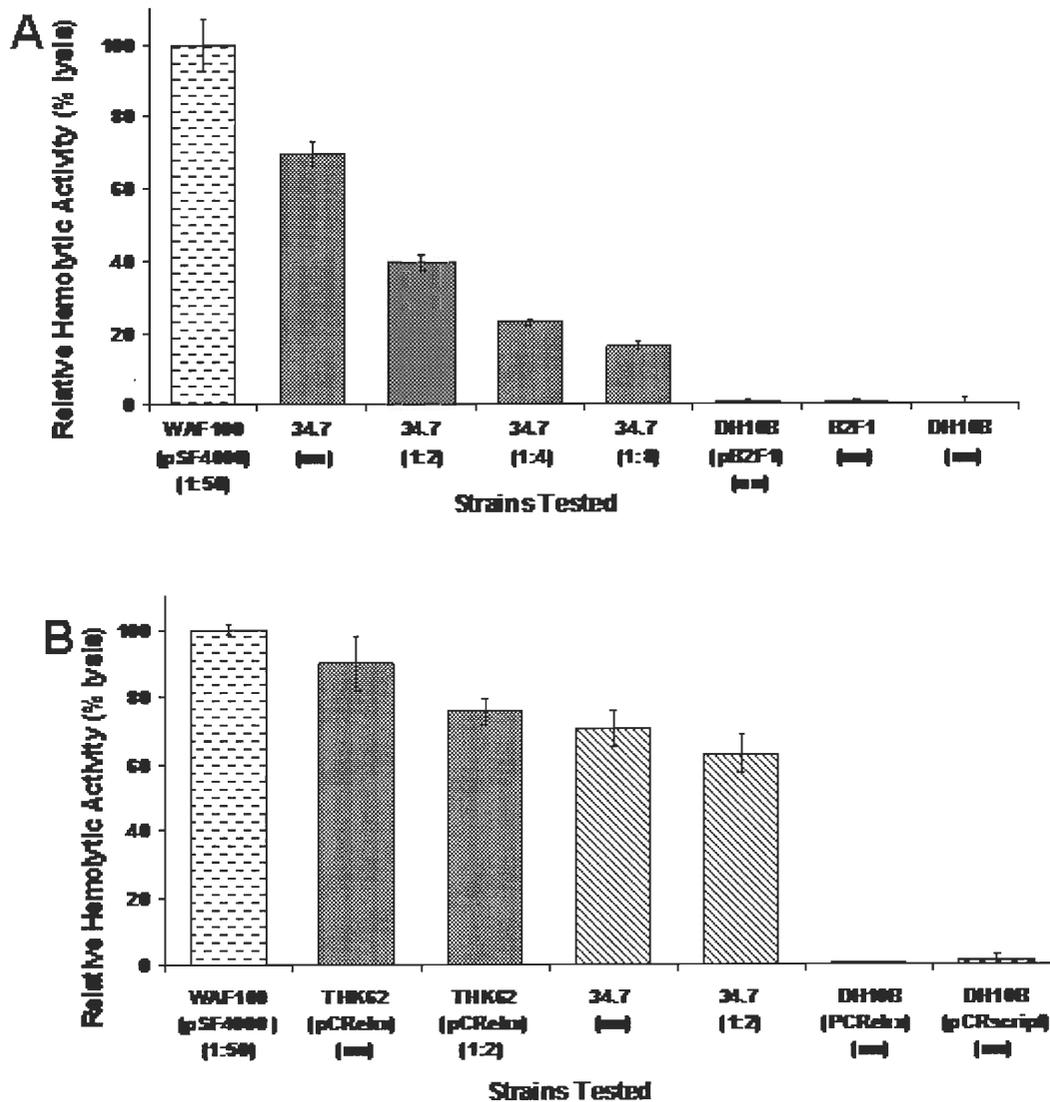


Figure 32. Hemolytic tube Assay: Comparison of extracellular hemolytic activity present in supernatants recovered from cultures of wild-type strains and *hns* mutants. Panel A: The samples of supernatants recovered from each bacterial isolate were used as undiluted (un), serial 2-fold dilutions (1:2); (1:4); or (1:8), or diluted 1:50 as shown. Panel B: Bacterial cultures were equalized and the supernatant from the logarithmic-phase cultures recovered by centrifugation and filtered, then 2% washed SRBC suspension was added. The mixture was incubated for 4 hours at 37°C. Post incubation the optical density at 540 nm of supernatants was determined. Error bars indicate the standard error from the mean.

HNS binding to the putative ehx promoter region

HNS binding assays to the putative promoter region of the ehx operon carried by STEC O91:H21- Based on the evidence described above, we hypothesized that the HNS protein interacts directly with the putative promoter region of the *ehx* operon to repress gene transcription. To test this hypothesis, we performed electrophoretic DNA mobility shift assays (EMSA) at both 25°C and 37°C to determine if temperature had a significant effect on HNS binding. The EMSA results shown are for reactions done at 37°C with a DNA fragment which includes 126 nucleotides just upstream of the *ehxC* translational start site (Figure 33A). In addition, we ran EMSA with a slightly smaller fragment that contained only 88 nucleotides of DNA upstream of the *ehxC* start codon. The result of EMSA with the 126 bp fragment of DNA is shown below. The HNS-his-tagged fusion protein purified by nickel-affinity chromatography (Figure 33B lane 2) migrated to the expected size of 31 kDa. Figure 30B lane 2 is a sample of the purity of the HNS-his-tagged fusion protein used in the electrophoresis mobility shift assay (EMSA).

The target DNA region, 88 bp or 126 bp upstream of the start codon of *ehxC*, used in the EMSA did not include any of the *ehx* DNA coding sequence. A constant amount of DNA (20µM) that contained the putative promoter region of *ehx* and varying concentrations of HNS were reacted together for 30 minutes at 25°C or 37°C. The results shown in Figure 31A are the EMSA done at 37°C with the 126 bp fragment. A HNS:Promoter-*ehx* complex with significantly retarded mobility was formed upon addition of HNS at both 25°C (data not shown) and 37°C.

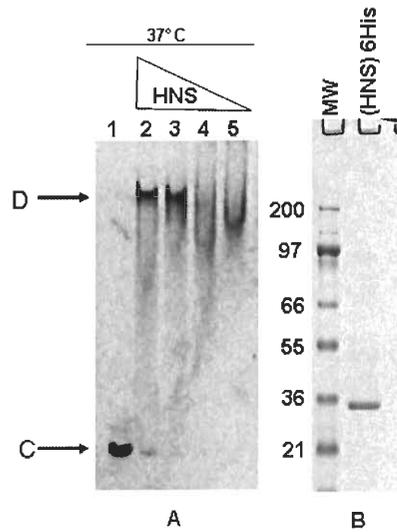


Figure 33. Electrophoretic mobility shift assay (EMSA): Binding of HNS to the putative promoter region of *ehx* from STEC O91:H21. Panel A lane 1 contains the 126 bp fragment of DNA without HNS added. Migration of 20 μM of the 126 bp fragment is marked by an arrow labeled as F. Lanes 2 - 5 contains 0.40 μM, 0.20 μM, 0.10 μM, and 0.05 μM of HNS reacted with the 126 bp DNA fragment held constant at 20 μM. Arrow C indicates the untreated *ehx* fragment that is unbound to HNS. Arrow D indicates *ehx* fragment that is bound to HNS. Panel B is the colloidal blue stained SDS-PAGE gel that contains the HNS histidine-tagged 31 kDa fusion protein recovered, shown in lane 2, following nickel-chelating affinity chromatography. Molecular weight markers are indicated for panel B.

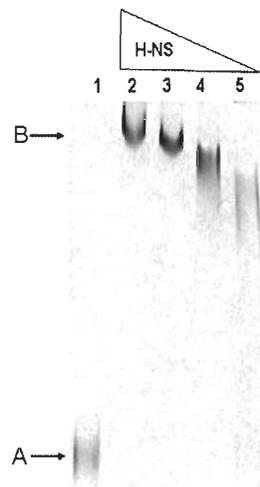


Figure 34. Electrophoretic mobility shift assay to determine HNS binding to the putative *ehx* promoter region of STEC O157:H7. Lane 1 contains the putative 112 bp promoter

region with no HNS added. Lanes 2-5 contain the same 112 bp region incubated with decreasing concentrations of HNS; 0.40 μM , 0.20 μM , 0.10 μM , and 0.05 μM . DNA concentration was kept constant at 20 μM . Arrow A indicates the untreated *ehx* fragment that is unbound to HNS. Arrow B indicates *ehx* fragment that is bound to HNS.

Maximal DNA/HNS complex was formed when HNS was added at 0.2 μM and 0.10 μM (Figure 33A lanes 2-3). However, when the concentration of HNS was reduced to 0.050 μM or 0.025 μM the DNA/HNS complex began to fall apart or did not form (Figure 33A lanes 4-5). Both fragments of DNA that corresponded to the upstream region of *ehxC* yield virtually identical results. The result suggests that the regulatory cis-acting element which is sensitive to HNS repression lies within 88 bp of DNA upstream of the *ehxC* start codon and that HNS binds this fragment of DNA at both 25°C and 37°C. Direct binding of HNS to target DNA indicates a role for HNS regulation of transcription of the *ehx* operon.

Regulation of the *ehx* operon in LEE-positive STEC strains is controlled by the LEE encoded regulatory protein GrlA. However, when an identical study was performed using the region upstream of the *ehx* operon from the LEE-positive STEC O157:H7. The HNS protein appeared to bind the same region of DNA in the same manner in STEC O157:H7 as in STEC O91:H21. The results are shown in Figure 34.

Transcription of the *ehx* operon in the presence of HNS

To demonstrate that the HNS protein represses transcription of the *ehx* operon, a gene fusion between the putative promoter region of the *ehx* operon and the promoter-less β -galactosidase gene was constructed. We observed β -galactosidase activity from

vector pMC*ehxCA'-lacZ* in the wild-type background (DH10B) and in the HNS null background (THK62). Vector pMC*ehxCA'-lacZ* contains 126 bp of DNA upstream of the *ehxC* start codon along with the open-reading-frame of *ehxC* and includes 695 bp of the *ehxA* coding region. The two *E. coli* strains used, DH10B and THK62 are both defective in β -galactosidase. When plasmid pMC*ehxCA'-lacZ* was transferred into DH10B, the level of detectable β -galactosidase was negligible (Figure 35). In contrast, in the HNS null mutant strain that contained the transcriptional fusion construct, THK62 (pMC*ehxCA'-lacZ*), ~10-fold more β -galactosidase was produced as compared to THK62 without the *lacZ* transcriptional fusion (Figure 35). In sum, the data supports the idea that HNS regulates transcription of the *ehx* operon by direct binding to DNA sequences within 126 bp region of DNA upstream of the translational start site of the *ehx* operon, that HNS binding results in suppression of transcription of the *ehx* operon. This 126 bp region of contains the promoter of *ehx*.

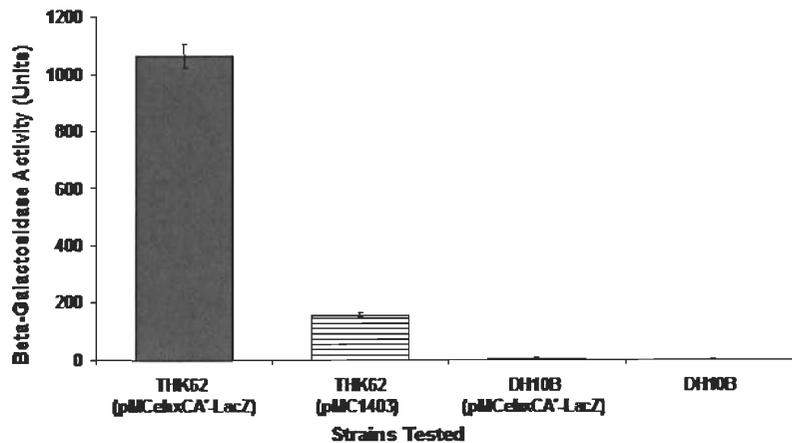


Figure 35. Effect of the *hns* mutation carried by THK62 on the expression of β -galactosidase from the *ehxCA'*::lacZ reporter gene fusion. All strains were grown to mid-log phase at 37°C and assayed for β -galactosidase activity. The data is an average of three independent experiments, with each experiment done in triplicate. The β -galactosidase activity measured as enzymatic activity/mg of protein/time. Detection was done at a maximum absorbance at optical density 570 nm. Error bars indicate the standard error of the mean.

Comparison of upstream nucleotide sequence of the *ehx* operon between STEC O91:H21 and STEC O157:H7

A fragment that contained 100 nucleotides of DNA upstream of *ehxC* and included 516 nucleotides of the open reading frame of the *ehxC* gene carried by STEC O91:H21 strain STEC O91:H21 was sequenced. The large plasmid carried by STEC O113:H21 strain EH41 has been sequenced (Leyton et al., 2003) and the large plasmids carried by four different strains of STEC have been sequenced (Brunner et al., 2006) (Burland et al., 1998) (Makino et al., 1998). To determine if the non-coding upstream regulatory sequences of the *ehx* operon cloned from strain STEC O91:H21 was similar to the same region of DNA in STEC O157 we used the STEC O91:H21 sequence as the

query and subjected it to the Basic Local Alignment Search Tool (Blast). We found that the region of DNA 100 bp upstream of the start codon of *ehxC* present in four different strains of STEC O157 was 95% identical to the same region of DNA cloned from STEC O91:H21. The alignment was virtually identical in each of the four strains of STEC O157 with no gaps (Figure 36). The EC4115 representative O157 strain is shown but the results of all three O157 isolates were similar. The boxed nucleotides are those which are different from the sequence found in STEC O91:H21. The strains of STEC O157:H7 identified were: EC4115, 3072/96, EDL933, and Sakai. As expected, alignment between the related STEC O113:H21 and O91:H21 isolates was higher at 97% with 95% of the DNA fragments tested. An almost identical 100 bp DNA sequence exists upstream of the start codon of *ehxC* in both LEE-negative and LEE-positive STEC. We showed that this region serves as a cis-regulatory element important for HNS binding and gene transcription in STEC, therefore we suggest that this region of DNA found in STEC organisms function in a like manner.

Serotype of STEC	Strain	Alignment of DNA Sequence Upstream of Start Codon of <i>ehxC</i>
O91:H21	B2F1	-95 TGTTTTAGATGCTTCTTGCTTAAAAGAATATAAATTCCTGTTCTTT
O113:H21	EH41	-95 TGTTTTAGATGCTTCTTGCTTAAAAGGATATAAATTCCTGTTCTTT
O157:H7	EC4115	-95 TGTTTTAGATGCTTCTTGCTTAAAAGGATATAAATTCCTGTTCTTT
O91:H21	B2F1	ATATAGAGTTCTTTACATTTACTGAATTTTCTATTGTTTAAACATTATG ^{start codon}
O113:H21	EH41	ATATAGAGTTCTTTACATTTACTGAATTTTCTATTGTTTAAACATTATG
O157:H7	EC4115	ATATAGGTTCTTTACATTTACTGAGTTTTCTATTATTAAACATTATG

Figure 36. Sequence alignment of the putative promoter region of *ehx* described for STEC O113:H21 strains EH41 which lies upstream of the putative translational start site of *ehxC*. Significant alignments with DNA sequences derived from STEC O91:H21 strain B2F1 using the Basic Local Alignment Search Tool (BLAST) database. Boxed nucleotides are those different from strain B2F1. Matching DNA sequence carried on the large plasmid of five different strains of Shiga toxin-expressing *E. coli*: O113:H21 strain EH41 [accession # NC_007365.1]; O157:H- strain 3072/96 [accession # NC_009602.1]; O157:H7 strain EC4115 [accession # NC_011350.1]; O157:H7 EDL933 [accession # NC_007414.1]; and O157:H7 strain Sakai [accession # NC_002128.1]

DISCUSSION AND CONCLUSION

Characterization of YghJ a novel protein involved in carbon metabolism and host colonization of *Escherichia coli* organisms

E. coli is capable of existing in surprisingly diverse environments, thus breaking the paradigm that *E. coli* exists entirely within a mammalian host. We no longer believe that *E. coli* found in other environments is a consequence of fecal contamination. *E. coli* associate with roots and leaves of plants, (Ibekwe et al., 2009) sands of beaches (Hartz et al., 2008) (Hartz, A. 2008) and even in marine environments (Ibekwe et al., 2004), 2004). What specific attributes of *E. coli* enhance this versatility? How can *E. coli* adapt to the rigors of life in the host and to the harshness of the environment? Pathogenic strains of *E. coli* appear deceptively facile yet we have yet to break the code for how LEE-negative STEC undertake and complete the cycle of infection within the host. We do not stay one step ahead but are continually surprised by the rise of new vectors of infection. Prior to my arrival, our group identified through a comparative proteomics study thirty novel extracytoplasmic uncharacterized proteins. [Note: Extracytoplasmic proteins are defined as proteins transported across the inner membrane with residency in the periplasm, outer membrane or secreted from the cell.] Then using publically available functional genomics programs we determined that ten of the proteins we identified did not fall into recognizable functional/structural groups. Further analysis of available data allowed us to divided these proteins into groups and recognize that ten of the thirty novel proteins had the potential of being important in virulence.

The YghJ protein was one of the proteins we discovered which appeared promising. We reached this conclusion partially because in 1988 a study by Peterson and Mekalanos reported that a transposon insertion into the *acfD* gene, carried by *Vibrio cholerae* with ~49% identity to *yghJ*, resulted in a *V. cholerae* mutant defective in colonization of the infant mouse; the accepted *V. cholerae* animal model (Peterson and Mekalanos, 1988). All the information we had at the time was that the YghJ protein resided outside of the cytoplasm and was likely membrane attached and was coded for by LEE-negative but not LEE-positive STEC.

We determined by phylogenetic analysis that similar proteins are present not only in *E. coli* and *V. cholerae* but *yghJ*-like genes are carried by a broad range of Gram-negative bacteria. Almost all of these bacteria adapt to multiple environments. For example *Vibrio* sp. exists as marine symbionts, in planktonic form, as biofilm and as infective agents within the human host. *Shewanella* persist in both marine environments as well as in soil. *E. coli* is not typically thought of to be ‘showy’ pathogens like *Vibrio* or *Shewanella* but diarrheal pathogenic *E. coli* like STEC typically carry large genetic islands which allow them to colonize.

In the last six months a reverse vaccinology report involving a YghJ-like protein shows that this protein is important in the survival of an ExPEC isolate within an animal model (Moriel et al., 2010). Specifically, this group identified YghJ as a protective antigen after animals experimentally infected with an ExPEC (pathogenic *E. coli* strain causing sepsis) isolate carrying the *yghJ* gene [*yghJ* homolog = ECOK1_3385] survived infection with little symptoms of disease (Moriel et al., 2010), *et al.*2010). This strongly

supports our findings that YghJ provides LEE-negative STEC O91:H21 a significant growth advantage in poor carbon environments over STEC O157:H7 which does not carry *yghJ*. Our data together with the *Vibrio* and the ExPEC studies imply that YghJ is an essential host colonization factor.

The lower large intestine in mammals is carbon limited because the majority of carbohydrates are absorbed in the small intestine and depleted in route to the large intestine. Due to the great array of microbes present in the large intestine the competition is fierce for available carbon sources. The capacity to take advantage of poor and scarce carbon sources would provide a significant advantage to bacteria within the host. We show that the YghJ protein confers STEC O91:H21 the capacity to utilize L-malate efficiently outgrowing STEC O157:H7 when grown in the same media. This advantage could translate inside the host where YghJ will provide a competitive growth advantage to organisms that possess it. Additionally, *ex vivo* L-malate is an important molecule in the phyllosphere and rhizosphere of plants, where it functions in plant metabolism and homeostasis (Nunes-Nesi et al., 2005). Might YghJ enable *E. coli* to survive and thrive on and around plants due to a heightened ability to metabolize L-malate?

A broad shotgun protein-protein interaction study (Arifuzzaman et al., 2006) indicates that YghJ may interact with the membrane-bound FAD-dependent malate dehydrogenase MQO, found in the *E. coli* periplasm (Kretzschmar et al., 2002). We show that YghJ localizes to the inner membrane after export to the periplasm, so it exists in the same cellular compartment as MQO, but we have not shown any interaction between YghJ and MQO proteins. However, our functional analysis indicates that YghJ possesses

a putative DUF4092 domain. The DUF4092 domain is present in many inner membrane bound lipoproteins encoded by γ -proteobacteria. For example, one protein, a putative aldehyde dehydrogenase, (Accession # C5DGE1.1) that belongs to the same superfamily produced by *Lachancea thermotolerans*, contains the DUF042 domain and is involved in carbohydrate metabolism. Not enough is known about the DUF4092 domain to hypothesize exactly what attribute this domain provides to YghJ. The protein-protein interaction study was crude but we were surprised by the numerous proteins that are suspected of interacting with YghJ that contain ATP binding domains. We also noted that the same region within the *yghJ* gene that overlaps with the proposed DUF4092 domain has a rudimentary PP2C-like domain. We proposed that YghJ had PP2C activity but were unable to demonstrate this activity; this may be because we did not have the correct substrate or the domain may look somewhat similar to PP2C but have an entirely different function.

The YghJ protein also possesses a putative M60-like domain. We were able to identify by computer modeling that YghJ contained the conserved zinc binding motif HEVGH, essential for activity. This zinc binding metallopeptidase domain was first identified as a viral enhancin in the insect virus *Baculovirus*. The enhancin protein is commonly found in gut-associated microbes and has been proven to have mucinase activity in both the insect pathogen *Baculovirus* (Wang and Granados, 1997a) and more recently found in the human gut associated microbe *Bacteroides thetaiotamicron* (Nakjang et al., 2012). We are attempting to produce a cytoplasmic version of YghJ without its signal sequence and test this construct of YghJ for mucinase activity.

The large intestine is an environment rich in mucin which helps to form a protective layer over the epithelial cells and limit the movement of bacteria throughout the host. (Belley et al., 1999). Mucinase activity conferred by the M60-like domain may enhance colonization of STEC by allowing it to penetrate the mucin layer and come into direct contact with the host colonic microvilli, or by allowing the bacterial cells to move more freely in the intestinal environment. Additionally, one of the major mucin subunits is glucosamine which is also an important component of the bacterial cell wall (Hill et al., 1992). During glycolysis, fructose-6-phosphate is generated and can be redirected to create UDP-N-acetyl-D-glucosamine. Alternatively, D-glucosamine from the environment can be transported into the cytoplasm of the cell by the mannose PTS permease (Kornberg, 1975). This process results in D-glucosamine-6-phosphate and a molecule of pyruvate. This pathway ultimately results in a molecule of UDP-N-acetyl-D-glucosamine, which can be used for peptidoglycan cell wall synthesis or lipid biosynthesis. Degradation of mucin may provide the bacteria with an important carbon source and allow it to conserve energy by absorbing glucosamine from the environment rather than synthesizing the molecule itself from D-fructose-6-phosphate.

We tested the role of HNS in expression of *yghJ* but found that HNS had no influence over expression of *yghJ*. We re-cloned the promoter region of *yghJ* onto a vector to form a transcriptional fusion but have not been able to demonstrate regulation to date. We were surprised by our data since another group reported that *yghJ* plays a role in transcription of the type II apparatus in ETEC. The *yghJ* gene in this species is located upstream of the operon for type II secretion. In contrast, to our knowledge STEC

O91:H21 does not possess the type II secretion operon. In the *E. coli* K12 background *yghJ* is located on its own transcriptional unit separate from the *pppA* and *glcA* flanking genes. Another report on EPEC strains shows that YghJ is secreted by the type II system and is probably not involved in transcription of the type II secretion operon (Baldi et al., 2012). Instead they suggest that YghJ is essential in formation of EPEC biofilm. The *yghJ* null mutant they created was unable to successfully form biofilms under laboratory conditions (Baldi et al., 2012). YghJ produced by STEC O91:H21 is not a secreted protein.

We were unable to purify the YghJ protein and Western blot data indicated that the 6XHis tag was not detectable by two separate antibodies. This may indicate that YghJ possesses an undetected cleavage sequence on the carboxy-terminus of the protein or that the protein possesses intramolecular bonds or is otherwise folded in a manner that renders the His tag inaccessible. We are creating a His-tagged variant of the YghJ protein lacking the signal sequence to determine if the protein can be purified from the cytoplasm maybe then we can use metal chelate to purify the protein.

The *Bacteroides thetaiotamicron* enhancin-like protein/the M60-like protein is also an inner membrane protein yet all the experiments shown in the Nakjang manuscript the protein did not contain the signal sequence thereby re-directing it to the cytoplasm (Nakjang et al., 2012). The Baldi group did not attempt purification but showed a protein preparation from lysate and the Yang group never demonstrated expression of the protein at all, although an RT-PCR did indicate transcription of the gene to mRNA. We suspect

but can not prove that these groups could not isolate the YghJ by His-tag or Gst-tag immunoaffinity purification methods.

There is strong evidence from our group and from other reports that the YghJ protein holds great promise as a unique protein with a significant role in host colonization. Looking forward we predict that L-malate utilization only one of the many functions of YghJ.

In addition to studies to identify and characterize novel virulence factors, one of my goals was to define the mechanism behind HNS repression of the enterohemolysin operon. The enterohemolysin operon is a well-defined group of genes that have been extensively studied by various groups yet its specific function in virulence of LEE-negative STEC is still unknown. However, previous data does suggest a role in the disease process (Smith et al., 2008).

Defining the mechanism of HNS repression of the enterohemolysin operon

Both LEE-negative and LEE-positive STEC strains seem to tightly regulate the expression of enterohemolysin. For example, reports indicate difficulty in detecting liquid-phase hemolytic activity in culture supernatants recovered from isolates of Shiga toxin-expressing *E. coli* (Beutin et al., 1989) (Chart et al., 1998) (Schmidt et al., 1995) Moreover, contact hemolytic activity requires at least 16 hours incubation on blood agar plates before hemolysis is detectable (Chart et al., 1998) (Schmidt et al., 1995). Data presented here supports the idea that in STEC, transcription of the *ehx* operon is tightly controlled by the global regulator HNS. Why would such tight restrictions be placed on

expression of *ehxA*? We know that HNS is important for adequate adherence and motility since null mutations in *hns* disrupt both of these function (Scott et al., 2003). A recent animal study by (Smith et al., 2008) shows that α -hemolysin is the direct cause of sloughing of epithelial cells from the urethral lining of mice within 24 hours post bacterial inoculation. Thus α -hemolysin causes significant damage to the urethral epithelial mucosa and contributes to disease manifestations of infection (Smith et al., 2008). We suggest that a significant amount of *EhxA* present during the initial stages of colonization when adherence is most important would be counter-productive. Therefore, we propose that the reason *EhxA* is down-regulated at 37°C may be to allow bacterial-host cell attachment to occur without the risk of host cell damage by *EhxA* during colonization. We now know that expression of enterohemolysin is repressed at 37°C but we also know that at some point during infection *EhxA* must be made because convalescent sera of patients that recover from STEC infections contain antibody to *EhxA* (Schmidt et al., 1995). As of yet, the precise role of *EhxA* in STEC disease has not been determined.

The HNS protein acts to silence or repress gene expression of several diverse and unlinked genes in *E. coli* and other Gram-negative bacteria (Dorman, 2004;Dorman, 2007). It is thought that HNS binds to DNA and creates a nucleoprotein structure which may be a blockade to RNA polymerase and repress transcription of specific genes (Dorman, 2004;Dorman, 2006) . The precise mechanism as to how HNS silences gene transcription remains a debate (Dorman, 2004). This small 137 amino acid protein contains a nucleic-acid binding domain at its carboxy-terminus attached by a flexible

linker to an oligomerization domain at its amino-terminal end. Investigation of binding of HNS to DNA reveals that it does not typically bind to a specific consensus sequence but binds to A + T-rich regions of curved DNA(Dorman, 2006). However, a study by Lang *et al* 2007 identified a discrete DNA sequence, 5'-TCGATATATT-3', in a subset of genes to which HNS bound (Lang et al., 2007). We have not found this specific DNA sequence in the promoter region of *ehxC*.

Some A+ T-rich regions of DNA may be sensitive to environmental factors such as osmolarity and temperature. Changes in these factors can affect curvature of DNA at the promoter of HNS regulated genes causing HNS to bind or to be displaced. Moreover, HNS itself can also cause change in curvature of DNA at select promoters. The precise environmental cues that stimulate binding of HNS to the cis-acting element upstream of *ehxC* are not known. Also, factors that cause HNS to be displaced from the promoter region have not been explored. Our studies show that HNS can bind to the promoter region at both 25°C and 37°C which suggests that temperature may not play a large role in regulation of *ehx* expression. Furthermore, during analysis of the *in vivo* effect of an *hns* null mutation on the expression of *ehx*, we found that HNS was essential for repression of transcription of *ehx* in *E. coli* cells grown to mid-log phase at 37°C. In contrast, a study with an STEC O157 isolate by Li *et al.* showed that HNS inhibition of *ehx* was maximal when cultures were grown at 30 °C and had reached stationery phase (Li et al., 2008). Our data are based on findings for LEE-negative STEC which probably explains the discrepancy in the conclusions between our two groups.

It is well documented that temperature and osmolarity are environmental factors that regulate expression of plasmid-encoded genes of the *hlyCABD* virulence operon cloned from the transmissible plasmid pHly152 carried by the extra-intestinal uropathogenic *E. coli* strain PM152 (Juarez et al., 2000), (Madrid et al., 2002). HNS is a crucial factor in regulation of the *hly* operon as it binds to two essential cis-acting elements located upstream of the coding region of *hlyCABD* (Godessart et al., 1988) (Vogel et al., 1988). In the case of *hly*, protein-protein interactions between HNS and the small regulator protein Hha control expression of the *hly* operon in response to environmental changes in osmolarity and temperature (Nieto et al., 1991; Nieto et al., 1997; Nieto et al., 2000). We found that there is no DNA sequence identity between the osmo-thermoregulatory cis-acting elements found upstream of *hlyCABD* and the promoter region of *ehxCABD* encoded by STEC O91:H21 and STEC O113:H21.

It is widely accepted that the *ehx* operon is transcribed from a single promoter located upstream of the *ehxC* gene (Schmidt et al., 1996) (Brunner et al., 2006) (Schmidt et al., 1995) (Bauer and Welch, 1996). Sequence analysis of the *ehx* operon located on the large 165 kbp plasmid of STEC O113:H21 strain EH41 revealed that 302 nucleotides lie between the stop codon of the previous open-reading-frame and the ATG start site of the *ehx* operon (Leyton et al., 2003). By comparison, the STEC O157:H7 Sakai strain contains 564 bp of DNA upstream of the start of the open-reading-frame of *ehxC* (Makino et al., 1998). When we compared the sequence of DNA 100 bp upstream of the *ehxC* gene encoded by strain B2F1 to four different isolates of STEC O157, we found

that the DNA sequences were almost identical. We determined that HNS binds to the same region of the putative *ehx* promoter in STEC O157:H7 as in STEC O91:H21. This could indicate that the LEE was acquired by LEE positive STEC after the large virulence plasmid, and that before acquisition of the LEE, the *ehx* operon was regulated by the same mechanism in LEE positive and LEE negative strains.

Although regulation of the *ehx* operon carried by STEC isolates involves activation by GrlA, we propose that HNS plays a significant role in *ehx* regulation. For example, GrlR negatively regulates GrlA through direct interactions (Creasey et al., 2003) (Iyoda and Watanabe, 2005) (Jobichen et al., 2007). As STEC O157:H7 organisms reach stationary growth phase, concentration of the negative regulator GrlR drops with increased levels of free GrlA as a result (Iyoda and Watanabe, 2005). The GrlA protein is then free to activate expression of *ehx* (Saitoh et al., 2008). We suggest that activation of *ehx* by GrlA probably occurs through GrlA inhibition of HNS binding to the promoter region of *ehx* thereby mitigating the repression of enterohemolysin by HNS. Disruption of HNS repression by GrlA and the mechanism by which it is achieved remains to be proven. However, several models exist in nature which involves displacement or blockade of HNS (Stoebel et al., 2008). For example, in *Vibrio cholerae* the ToxT regulatory protein silences HNS by displacement of HNS from the promoter of genes that the ToxT regulator activates (Hulbert and Taylor, 2002) (Yu and DiRita, 2002).

In this study, we showed that transcriptional control of the *ehx* operon by HNS is due to binding of HNS within an 88 bp region of DNA located upstream of the

translational start codon of *ehxC*. Previously, we found that HNS played a role in *ehxA* gene expression (Scott et al., 2003) but it was not known until now that HNS binds to cis-acting elements within the promoter region to directly repress expression of *ehxCABD* in STEC O91:H21. We also show that the almost identical nucleotide sequence within the promoter region of both LEE-negative and LEE-positive STEC is bound by HNS in the same manner. In the future the specific nucleotides involved in HNS inhibition of the *ehx* operon will be identified. Evidence indicates that HNS regulation occurs in LEE-negative STEC serotypes through direct DNA-binding to promoter DNA sequences.

REFERENCES

1. Abe, A., L.J.Arend, L.Lee, C.Lingwood, R.O.Brady, and J.A.Shayman. 2000. Glycosphingolipid depletion in fabry disease lymphoblasts with potent inhibitors of glucosylceramide synthase. *Kidney Int.* 57:446-454.
2. Arifuzzaman, M., M.Maeda, A.Itoh, K.Nishikata, C.Takita, R.Saito, T.Ara, K.Nakahigashi, H.C.Huang, A.Hirai, K.Tsuzuki, S.Nakamura, M.taf-UI-Amin, T.Oshima, T.Baba, N.Yamamoto, T.Kawamura, T.Ioka-Nakamichi, M.Kitagawa, M.Tomita, S.Kanaya, C.Wada, and H.Mori. 2006. Large-scale identification of protein-protein interaction of Escherichia coli K-12. *Genome Res* 16:686-691.
3. Baba, T., T.Ara, M.Hasegawa, Y.Takai, Y.Okumura, M.Baba, K.A.Datsenko, M.Tomita, B.L.Wanner, and H.Mori. 2006. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* 2:2006.
4. Baldi, D.L., E.E.Higginson, D.M.Hocking, J.Praszkiel, R.Cavaliere, C.E.James, V.nett-Wood, K.I.Azzopardi, L.Turnbull, T.Lithgow, R.M.Robins-Browne, C.B.Whitchurch, and M.Tauschek. 2012. The type II secretion system and its ubiquitous lipoprotein substrate, SslE, are required for biofilm formation and virulence of enteropathogenic Escherichia coli. *Infect. Immun.* 80:2042-2052.
5. Bauer, M.E. and R.A.Welch. 1996. Characterization of an RTX toxin from enterohemorrhagic Escherichia coli O157:H7. *Infect. Immun.* 64:167-175.
6. Belley, A., K.Keller, M.Gottke, and K.Chadee. 1999. Intestinal mucins in colonization and host defense against pathogens. *Am. J. Trop. Med. Hyg.* 60:10-15.
7. Berdichevsky, T., D.Friedberg, C.Nadler, A.Rokney, A.Oppenheim, and I.Rosenshine. 2005. Ler is a negative autoregulator of the LEE1 operon in enteropathogenic Escherichia coli. *J. Bacteriol.* 187:349-357.
8. Bertin, P., E.Terao, E.H.Lee, P.Lejeune, C.Colson, A.Danchin, and E.Collatz. 1994. The HNS protein is involved in the biogenesis of flagella in Escherichia coli. *J. Bacteriol.* 176:5537-5540.
9. Beutin, L., M.A.Montenegro, I.Orskov, F.Orskov, J.Prada, S.Zimmermann, and R.Stephan. 1989. Close association of verotoxin (Shiga-like toxin) production with enterohemolysin production in strains of Escherichia coli. *J. Clin. Microbiol.* 27:2559-2564.

10. Blattner, F.R., G.Plunkett, III, C.A.Bloch, N.T.Perna, V.Burland, M.Riley, J.Collado-Vides, J.D.Glasner, C.K.Rode, G.F.Mayhew, J.Gregor, N.W.Davis, H.A.Kirkpatrick, M.A.Goeden, D.J.Rose, B.Mau, and Y.Shao. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* 277:1453-1462.
11. Bork, P., N.P.Brown, H.Hegy, and J.Schultz. 1996. The protein phosphatase 2C (PP2C) superfamily: detection of bacterial homologues. *Protein Sci.* 5:1421-1425.
12. Brok, R., G.P.Van, M.Winterhalter, U.Ziese, A.J.Koster, C.H.de, M.Koster, J.Tommassen, and W.Bitter. 1999. The C-terminal domain of the *Pseudomonas* secretin XcpQ forms oligomeric rings with pore activity. *J. Mol. Biol.* 294:1169-1179.
13. Brunder, W., H.Karch, and H.Schmidt. 2006. Complete sequence of the large virulence plasmid pSFO157 of the sorbitol-fermenting enterohemorrhagic *Escherichia coli* O157:H- strain 3072/96. *Int. J. Med. Microbiol.* 296:467-474.
14. Brunder, W., H.Schmidt, and H.Karch. 1996. KatP, a novel catalase-peroxidase encoded by the large plasmid of enterohaemorrhagic *Escherichia coli* O157:H7. *Microbiology* 142 (Pt 11):3305-3315.
15. Brunder, W., H.Schmidt, and H.Karch. 1997. EspP, a novel extracellular serine protease of enterohaemorrhagic *Escherichia coli* O157:H7 cleaves human coagulation factor V. *Mol. Microbiol.* 24:767-778.
16. Bugarel, M., L.Beutin, A.Martin, A.Gill, and P.Fach. 2010. Micro-array for the identification of Shiga toxin-producing *Escherichia coli* (STEC) seropathotypes associated with Hemorrhagic Colitis and Hemolytic Uremic Syndrome in humans. *Int. J. Food Microbiol.* 142:318-329.
17. Burland, V., Y.Shao, N.T.Perna, G.Plunkett, H.J.Sofia, and F.R.Blattner. 1998. The complete DNA sequence and analysis of the large virulence plasmid of *Escherichia coli* O157:H7. *Nucleic Acids Res.* 26:4196-4204.
18. Bustamante, V.H., F.J.Santana, E.Calva, and J.L.Puente. 2001. Transcriptional regulation of type III secretion genes in enteropathogenic *Escherichia coli*: Ler antagonizes HNS-dependent repression. *Mol. Microbiol.* 39:664-678.
19. Camberg, J.L., T.L.Johnson, M.Patrick, J.Abendroth, W.G.Hol, and M.Sandkvist. 2007. Synergistic stimulation of EpsE ATP hydrolysis by EpsL and acidic phospholipids. *EMBO J.* 26:19-27.
20. Camberg, J.L. and M.Sandkvist. 2005. Molecular analysis of the *Vibrio cholerae* type II secretion ATPase EpsE. *J. Bacteriol.* 187:249-256.

21. Casadaban, M.J., J.Chou, and S.N.Cohen. 1980. In vitro gene fusions that join an enzymatically active beta-galactosidase segment to amino-terminal fragments of exogenous proteins: Escherichia coli plasmid vectors for the detection and cloning of translational initiation signals. *J. Bacteriol.* 143:971-980.
22. Chart, H., C.Jenkins, H.R.Smith, D.Hedges, and B.Rowe. 1998. Haemolysin production by strains of Verocytotoxin-producing Escherichia coli. *Microbiology* 144 (Pt 1):103-107.
23. Chowdhury, R., A.Ray, P.Ray, and J.Das. 1987. Replication and packaging of cholera phage phi 149 DNA. *J. Virol.* 61:3999-4006.
24. Collison, M., R.P.Hirt, A.Wipat, S.Nakjang, P.Sanseau, and J.R.Brown. 2012. Data mining the human gut microbiota for therapeutic targets. *Brief. Bioinform.* Epub: DOI:10.1093/bib/bbs002
25. Creasey, E.A., R.M.Delahay, S.J.Daniell, and G.Frankel. 2003. Yeast two-hybrid system survey of interactions between LEE-encoded proteins of enteropathogenic Escherichia coli. *Microbiology* 149:2093-2106.
26. Das, A.K., N.R.Helps, P.T.Cohen, and D.Barford. 1996. Crystal structure of the protein serine/threonine phosphatase 2C at 2.0 Å resolution. *EMBO J.* 15:6798-6809.
27. Datsenko, K.A. and B.L.Wanner. 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A* 97:6640-6645.
28. Dave, M., P.D.Higgins, S.Middha, and K.P.Rioux. 2012. The human gut microbiome: current knowledge, challenges, and future directions. *Transl. Res.* Epub: DOI: 10.1016/j.trsl.2012.05.003.
29. Deng, W., J.L.Puente, S.Gruenheid, Y.Li, B.A.Vallance, A.Vazquez, J.Barba, J.A.Ibarra, P.O'Donnell, P.Metalnikov, K.Ashman, S.Lee, D.Goode, T.Pawson, and B.B.Finlay. 2004. Dissecting virulence: systematic and functional analyses of a pathogenicity island. *Proc. Natl. Acad. Sci. U. S. A* 101:3597-3602.
30. Donato, G.M., M.J.Lelivelt, and T.H.Kawula. 1997. Promoter-specific repression of fimB expression by the Escherichia coli nucleoid-associated protein HNS. *J. Bacteriol.* 179:6618-6625.
31. Dorman, C.J. 2006. DNA supercoiling and bacterial gene expression. *Sci. Prog.* 89:151-166.

32. Dorman, C.J. 2004. HNS: a universal regulator for a dynamic genome. *Nat. Rev. Microbiol.* 2:391-400.
33. Dorman, C.J. 2007. HNS, the genome sentinel. *Nat. Rev. Microbiol.* 5:157-161.
34. Drewke, C. and E.Leistner. 2001. Biosynthesis of vitamin B6 and structurally related derivatives. *Vitam. Horm.* 61:121-155.
35. DuPont, H.L., S.B.Formal, R.B.Hornick, M.J.Snyder, J.P.Libonati, D.G.Sheahan, E.H.LaBrec, and J.P.Kalas. 1971. Pathogenesis of Escherichia coli diarrhea. *N. Engl. J. Med.* 285:1-9.
36. Elliott, S.J., L.A.Wainwright, T.K.McDaniel, K.G.Jarvis, Y.K.Deng, L.C.Lai, B.P.McNamara, M.S.Donnenberg, and J.B.Kaper. 1998. The complete sequence of the locus of enterocyte effacement (LEE) from enteropathogenic Escherichia coli E2348/69. *Mol. Microbiol.* 28:1-4.
37. Elliott, S.J., J.Yu, and J.B.Kaper. 1999. The cloned locus of enterocyte effacement from enterohemorrhagic Escherichia coli O157:H7 is unable to confer the attaching and effacing phenotype upon E. coli K-12. *Infect. Immun.* 67:4260-4263.
38. Endo, Y. and K.Tsurugi. 1988a. The RNA N-glycosidase activity of ricin A-chain. *Nucleic Acids Symp. Ser.* 139-142.
39. Endo, Y. and K.Tsurugi. 1988b. The RNA N-glycosidase activity of ricin A-chain. *Nucleic Acids Symp. Ser.* 139-142.
40. Endo, Y., K.Tsurugi, T.Yutsudo, Y.Takeda, T.Ogasawara, and K.Igarashi. 1988. Site of action of a Vero toxin (VT2) from Escherichia coli O157:H7 and of Shiga toxin on eukaryotic ribosomes. RNA N-glycosidase activity of the toxins. *Eur. J. Biochem.* 171:45-50.
41. European Food Safety Authority. 2011. Tracing Seeds, in Particular Fenugreek (*Trigonell foenum-graecum*) Seeds, in Relation to the Shiga Toxin-Producing *E. coli* (STEC) O104:H4 2011 Outbreaks in Germany and France. *EFSA* 1-23.
42. Felmler T, Pellett S, and Welch RA. 1985. Nucleotide sequence of an *Escherichia coli* chromosomal hemolysin. *J. Bacteriol.* 163:94-105.
43. Formal, S.B. and R.B.Hornick. 1978. Invasive Escherichia coli. *J. Infect. Dis.* 137:641-644.

44. Formal, S.B., A.O'Brien, P.Gemski, and B.P.Doctor. 1978. Invasive Escherichia coli. *J. Am. Vet. Med. Assoc.* 173:596-598.
45. Frankel, G. and A.D.Phillips. 2008. Attaching effacing Escherichia coli and paradigms of Tir-triggered actin polymerization: getting off the pedestal. *Cell Microbiol.* 10:549-556.
46. Fraser, M.E., M.Fujinaga, M.M.Cherny, A.R.Melton-Celsa, E.M.Twiddy, A.D.O'Brien, and M.N.James. 2004. Structure of shiga toxin type 2 (Stx2) from Escherichia coli O157:H7. *J. Biol. Chem* 279:27511-27517.
47. Fukasawa, K., K.M.Fukasawa, H.Iwamoto, J.Hirose, and M.Harada. 1999. The HELLGH motif of rat liver dipeptidyl peptidase III is involved in zinc coordination and the catalytic activity of the enzyme. *Biochemistry* 38:8299-8303.
48. Furste, J.P., W.Pansegrau, R.Frank, H.Blocker, P.Scholz, M.Bagdasarian, and E.Lanka. 1986. Molecular cloning of the plasmid RP4 primase region in a multi-host-range tacP expression vector. *Gene* 48:119-131.
49. Godessart, N., F.J.Munoz, M.Regue, and A.Juarez. 1988. Chromosomal mutations that increase the production of a plasmid-encoded haemolysin in Escherichia coli. *J. Gen. Microbiol.* 134:2779-2787.
50. Gould, L.H., C.Bopp, N.Strockbine, R.Atkinson, V.Baselski, B.Body, R.Carey, C.Crandall, S.Hurd, R.Kaplan, M.Neill, S.Shea, P.Somsel, M.Tobin-D'Angelo, P.M.Griffin, and P.Gerner-Smidt. 2009a. Recommendations for diagnosis of shiga toxin--producing Escherichia coli infections by clinical laboratories. *MMWR Recomm. Rep.* 58:1-14.
51. Gould, L.H., L.Demma, T.F.Jones, S.Hurd, D.J.Vugia, K.Smith, B.Shiferaw, S.Segler, A.Palmer, S.Zansky, and P.M.Griffin. 2009b. Hemolytic uremic syndrome and death in persons with Escherichia coli O157:H7 infection, foodborne diseases active surveillance network sites, 2000-2006. *Clin. Infect. Dis.* 49:1480-1485.
52. Grant, S.G., J.Jessee, F.R.Bloom, and D.Hanahan. 1990. Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants. *Proc. Natl. Acad. Sci. U. S. A* 87:4645-4649.
53. Hartz, A., M.Cuvelier, K.Nowosielski, T.D.Bonilla, M.Green, N.Esiobu, D.S.McCorquodale, and A.Rogerson. 2008. Survival potential of Escherichia coli and Enterococci in subtropical beach sand: implications for water quality managers. *J. Environ. Qual.* 37:898-905.

54. Head, S.C., M.A.Karmali, M.E.Roscoe, M.Petric, N.A.Strockbine, and I.K.Wachsmuth. 1988. Serological differences between verocytotoxin 2 and shiga-like toxin II. *Lancet* 2:751.
55. Henao, O.L., E.Scallan, B.Mahon, and R.M.Hoekstra. 2010. Methods for monitoring trends in the incidence of foodborne diseases: Foodborne Diseases Active Surveillance Network 1996-2008. *Foodborne. Pathog. Dis.* 7:1421-1426.
56. Hill, J., D.A.Hutton, G.G.Green, J.P.Birchall, and J.P.Pearson. 1992. Culture of human middle ear mucosal explants; mucin production. *Clin. Otolaryngol. Allied Sci.* 17:491-496.
57. Huang, H.B., A.Horiuchi, J.Goldberg, P.Greengard, and A.C.Nairn. 1997. Site-directed mutagenesis of amino acid residues of protein phosphatase 1 involved in catalysis and inhibitor binding. *Proc. Natl. Acad. Sci. U. S. A* 94:3530-3535.
58. Hulbert, R.R. and R.K.Taylor. 2002. Mechanism of ToxT-dependent transcriptional activation at the *Vibrio cholerae* tcpA promoter. *J. Bacteriol.* 184:5533-5544.
59. Huys, G., M.Cnockaert, J.M.Janda, and J.Swings. 2003. *Escherichia albertii* sp. nov., a diarrhoeagenic species isolated from stool specimens of Bangladeshi children. *Int. J. Syst. Evol. Microbiol.* 53:807-810.
60. Ibekwe, A.M., C.M.Grieve, S.K.Papiernik, and C.H.Yang. Persistence of *Escherichia coli* O157:H7 on the rhizosphere and phyllosphere of lettuce. *Lett. Appl. Microbiol.* 49, 784-790. 2009.
Ref Type: Journal (Full)
61. Ibekwe, A.M., S.K.Papiernik, C.M.Grieve, and C.H.Yang. 2011. Quantification of Persistence of *Escherichia coli* O157:H7 in Contrasting Soils. *Int. J. Microbiol.* 2011.
62. Ibekwe, A.M., P.M.Watt, P.J.Shouse, and C.M.Grieve. 2004. Fate of *Escherichia coli* O157:H7 in irrigation water on soils and plants as validated by culture method and real-time PCR. *Can. J. Microbiol.* 50:1007-1014.
63. Isidean, S.D., M.S.Riddle, S.J.Savarino, and C.K.Porter. 2011. A systematic review of ETEC epidemiology focusing on colonization factor and toxin expression. *Vaccine* 29:6167-6178.
64. Ito, H., A.Terai, H.Kurazono, Y.Takeda, and M.Nishibuchi. 1990. Cloning and nucleotide sequencing of Vero toxin 2 variant genes from *Escherichia coli*

- O91:H21 isolated from a patient with the hemolytic uremic syndrome. *Microb. Pathog.* 8:47-60.
65. Iyoda, S., N.Honda, T.Saitoh, K.Shimuta, J.Terajima, H.Watanabe, and M.Ohnishi. 2011. Coordinate control of the locus of enterocyte effacement and enterohemolysin genes by multiple common virulence regulators in enterohemorrhagic *Escherichia coli*. *Infect. Immun.*
 66. Iyoda, S. and H.Watanabe. 2005. ClpXP protease controls expression of the type III protein secretion system through regulation of RpoS and GrlR levels in enterohemorrhagic *Escherichia coli*. *J. Bacteriol.* 187:4086-4094.
 67. Jacobsen, S.M., D.J.Stickler, H.L.Mobley, and M.E.Shirtliff. 2008. Complicated catheter-associated urinary tract infections due to *Escherichia coli* and *Proteus mirabilis*. *Clin. Microbiol. Rev* 21:26-59.
 68. Jobichen, C., M.Li, G.Yerushalmi, Y.W.Tan, Y.K.Mok, I.Rosenshine, K.Y.Leung, and J.Sivaraman. 2007. Structure of GrlR and the implication of its EDED motif in mediating the regulation of type III secretion system in EHEC. *PLoS. Pathog.* 3:e69.
 69. Johnson, J.R. and T.A.Russo. 2005. Molecular epidemiology of extraintestinal pathogenic (uropathogenic) *Escherichia coli*. *Int. J. Med. Microbiol.* 295:383-404.
 70. Johnson, T.L., J.Aabendroth, W.G.Hol, and M.Sandkvist. 2006. Type II secretion: from structure to function. *FEMS Microbiol. Lett* 255:175-186.
 71. Juarez, A., J.M.Nieto, A.Prenafeta, E.Miquelay, C.Balsalobre, M.Carrascal, and C.Madrid. 2000. Interaction of the nucleoid-associated proteins Hha and HNS to modulate expression of the hemolysin operon in *Escherichia coli*. *Adv. Exp. Med. Biol.* 485:127-131.
 72. Karmali, M.A., M.Petric, C.Lim, P.C.Fleming, G.S.Arbus, and H.Lior. 1985. The association between idiopathic hemolytic uremic syndrome and infection by verotoxin-producing *Escherichia coli*. *J. Infect. Dis.* 151:775-782.
 73. Karmali, M.A., M.Petric, C.Lim, P.C.Fleming, and B.T.Steele. 1983a. *Escherichia coli* cytotoxin, haemolytic-uraemic syndrome, and haemorrhagic colitis. *Lancet* 2:1299-1300.
 74. Karmali, M.A., M.Petric, S.Louie, and R.Cheung. 1986. Antigenic heterogeneity of *Escherichia coli* verotoxins. *Lancet* 1:164-165.

75. Karmali, M.A., B.T.Steele, M.Petric, and C.Lim. 1983b. Sporadic cases of haemolytic-uraemic syndrome associated with faecal cytotoxin and cytotoxin-producing *Escherichia coli* in stools. *Lancet* 1:619-620.
76. Kenny, B., R.DeVinney, M.Stein, D.J.Reinscheid, E.A.Frey, and B.B.Finlay. 1997. Enteropathogenic *E. coli* (EPEC) transfers its receptor for intimate adherence into mammalian cells. *Cell* 91:511-520.
77. Kikuchi, T., Y.Mizunoe, A.Takade, S.Naito, and S.Yoshida. 2005. Curli fibers are required for development of biofilm architecture in *Escherichia coli* K-12 and enhance bacterial adherence to human uroepithelial cells. *Microbiol. Immunol.* 49:875-884.
78. Kokai-Kun, J.F., A.R.Melton-Celsa, and A.D.O'Brien. 2000. Elastase in intestinal mucus enhances the cytotoxicity of Shiga toxin type 2d. *J. Biol. Chem.* 275:3713-3721.
79. Konowalchuk, J., J.I.Speirs, and S.Stavric. 1977. Vero response to a cytotoxin of *Escherichia coli*. *Infect. Immun.* 18:775-779.
80. Koronakis, V., A.Sharff, E.Koronakis, B.Luisi, and C.Hughes. 2000. Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature* 405:914-919.
81. Kretzschmar, U., A.Ruckert, J.H.Jeoung, and H.Gorisch. 2002. Malate:quinone oxidoreductase is essential for growth on ethanol or acetate in *Pseudomonas aeruginosa*. *Microbiology* 148:3839-3847.
82. Lang, B., N.Blot, E.Bouffartigues, M.Buckle, M.Geertz, C.O.Gualerzi, R.Mavathur, G.Muskhelishvili, C.L.Pon, S.Rimsky, S.Stella, M.M.Babu, and A.Travers. 2007. High-affinity DNA binding sites for HNS provide a molecular basis for selective silencing within proteobacterial genomes. *Nucleic Acids Res.* 35:6330-6337.
83. Lathem, W.W., T.E.Grys, S.E.Witowski, A.G.Torres, J.B.Kaper, P.I.Tarr, and R.A.Welch. 2002. StcE, a metalloprotease secreted by *Escherichia coli* O157:H7, specifically cleaves C1 esterase inhibitor. *Mol. Microbiol.* 45:277-288.
84. Levine, M.M., J.G.Xu, J.B.Kaper, H.Lior, V.Prado, B.Tall, J.Nataro, H.Karch, and K.Wachsmuth. 1987. A DNA probe to identify enterohemorrhagic *Escherichia coli* of O157:H7 and other serotypes that cause hemorrhagic colitis and hemolytic uremic syndrome. *J. Infect. Dis.* 156:175-182.

85. Leyton, D.L., J.Sloan, R.E.Hill, S.Doughty, and E.L.Hartland. 2003. Transfer region of pO113 from enterohemorrhagic *Escherichia coli*: similarity with R64 and identification of a novel plasmid-encoded autotransporter, EpeA. *Infect. Immun.* 71:6307-6319.
86. Li, H., A.Granat, V.Stewart, and J.R.Gillespie. 2008. RpoS, HNS, and DsrA influence EHEC hemolysin operon (*ehxCABD*) transcription in *Escherichia coli* O157:H7 strain EDL933. *FEMS Microbiol. Lett.* 285:257-262.
87. Lio, J.C. and W.J.Syu. 2004. Identification of a negative regulator for the pathogenicity island of enterohemorrhagic *Escherichia coli* O157:H7. *J. Biomed. Sci.* 11:855-863.
88. Lukjancenko, O., T.M.Wassenaar, and D.W.Ussery. 2010. Comparison of 61 sequenced *Escherichia coli* genomes. *Microb. Ecol.* 60:708-720.
89. Madrid, C., J.M.Nieto, S.Paytubi, M.Falconi, C.O.Gualerzi, and A.Juarez. 2002. Temperature- and HNS-dependent regulation of a plasmid-encoded virulence operon expressing *Escherichia coli* hemolysin. *J. Bacteriol.* 184:5058-5066.
90. Makino, K., K.Ishii, T.Yasunaga, M.Hattori, K.Yokoyama, C.H.Yutsudo, Y.Kubota, Y.Yamaichi, T.Iida, K.Yamamoto, T.Honda, C.G.Han, E.Ohtsubo, M.Kasamatsu, T.Hayashi, S.Kuhara, and H.Shinagawa. 1998. Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhagic *Escherichia coli* O157:H7 derived from Sakai outbreak. *DNA Res.* 5:1-9.
91. McDaniel, T.K., K.G.Jarvis, M.S.Donnenberg, and J.B.Kaper. 1995. A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *Proc. Natl. Acad. Sci. U. S. A* 92:1664-1668.
92. Melton-Celsa, A., K.Mohawk, L.Teel, and A.O'Brien. 2012. Pathogenesis of Shiga-toxin producing *Escherichia coli*. *Curr. Top. Microbiol. Immunol.* 357:67-103.
93. Melton-Celsa, A.R., J.F.Kokai-Kun, and A.D.O'Brien. 2002. Activation of Shiga toxin type 2d (Stx2d) by elastase involves cleavage of the C-terminal two amino acids of the A2 peptide in the context of the appropriate B pentamer. *Mol. Microbiol.* 43:207-215.
94. Moriel, D.G., I.Bertoldi, A.Spagnuolo, S.Marchi, R.Rosini, B.Nesta, I.Pastorello, V.A.Corea, G.Torricelli, E.Cartocci, S.Savino, M.Scarselli, U.Dobrindt, J.Hacker, H.Tettelin, L.J.Tallon, S.Sullivan, L.H.Wieler, C.Ewers, D.Pickard, G.Dougan, M.R.Fontana, R.Rappuoli, M.Pizza, and L.Serino. 2010. Identification of protective and broadly conserved vaccine antigens from the genome of

- extraintestinal pathogenic *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A* 107:9072-9077.
95. Moyne, A.L., M.R.Sudarshana, T.Blessington, S.T.Koike, M.D.Cahn, and L.J.Harris. 2011. Fate of *Escherichia coli* O157:H7 in field-inoculated lettuce. *Food Microbiol.* 28:1417-1425.
 96. Nakjang, S., D.A.Ndeh, A.Wipat, D.N.Bolam, and R.P.Hirt. 2012. A novel extracellular metallopeptidase domain shared by animal host-associated mutualistic and pathogenic microbes. *PLoS. One.* 7:e30287.
 97. Nataro, J.P., Y.Deng, S.Cookson, A.Cravioto, S.J.Savarino, L.D.Guers, M.M.Levine, and C.O.Tacket. 1995. Heterogeneity of enteroaggregative *Escherichia coli* virulence demonstrated in volunteers. *J. Infect. Dis.* 171:465-468.
 98. Nautiyal, C.S., A.Rehman, and P.S.Chauhan. 2010. Environmental *Escherichia coli* occur as natural plant growth-promoting soil bacterium. *Arch. Microbiol.* 192:185-193.
 99. Nieto, J.M., M.Carmona, S.Bolland, Y.Jubete, C.F.de la, and A.Juarez. 1991. The hha gene modulates haemolysin expression in *Escherichia coli*. *Mol. Microbiol.* 5:1285-1293.
 100. Nieto, J.M., C.Madrid, A.Prenafeta, E.Miquelay, C.Balsalobre, M.Carrascal, and A.Juarez. 2000. Expression of the hemolysin operon in *Escherichia coli* is modulated by a nucleoid-protein complex that includes the proteins Hha and HNS. *Mol. Gen. Genet.* 263:349-358.
 101. Nieto, J.M., M.Mourino, C.Balsalobre, C.Madrid, A.Prenafeta, F.J.Munoz, and A.Juarez. 1997. Construction of a double hha hns mutant of *Escherichia coli*: effect on DNA supercoiling and alpha-haemolysin production. *FEMS Microbiol. Lett.* 155:39-44.
 102. Nunes-Nesi, A., F.Carrari, A.Lytovchenko, A.M.Smith, M.E.Loureiro, R.G.Ratcliffe, L.J.Sweetlove, and A.R.Fernie. 2005. Enhanced photosynthetic performance and growth as a consequence of decreasing mitochondrial malate dehydrogenase activity in transgenic tomato plants. *Plant Physiol* 137:611-622.
 103. Nunn, D. 1999. Bacterial type II protein export and pilus biogenesis: more than just homologies? *Trends Cell Biol.* 9:402-408.
 104. O'Brien, A.D. and G.D.LaVeck. 1983. Purification and characterization of a *Shigella dysenteriae* 1-like toxin produced by *Escherichia coli*. *Infect. Immun.* 40:675-683.

105. O'Brien, A.D., G.D.LaVeck, M.R.Thompson, and S.B.Formal. 1982. Production of Shigella dysenteriae type 1-like cytotoxin by Escherichia coli. *J. Infect. Dis.* 146:763-769.
106. O'Brien, A.D., V.L.Tesh, A.Donohue-Rolfe, M.P.Jackson, S.Olsnes, K.Sandvig, A.A.Lindberg, and G.T.Keusch. 1992. Shiga toxin: biochemistry, genetics, mode of action, and role in pathogenesis. *Curr. Top. Microbiol. Immunol.* 180:65-94.
107. Okuda, T., N.Tokuda, S.Numata, M.Ito, M.Ohta, K.Kawamura, J.Wiels, T.Urano, O.Tajima, K.Furukawa, and K.Furukawa. 2006. Targeted disruption of Gb3/CD77 synthase gene resulted in the complete deletion of globo-series glycosphingolipids and loss of sensitivity to verotoxins. *J. Biol. Chem* 281:10230-10235.
108. Patrick, M., K.V.Korotkov, W.G.Hol, and M.Sandkvist. 2011. Oligomerization of EpsE coordinates residues from multiple subunits to facilitate ATPase activity. *J. Biol. Chem.*
109. Peterson, K.M. and J.J.Mekalanos. 1988. Characterization of the Vibrio cholerae ToxR regulon: identification of novel genes involved in intestinal colonization. *Infect. Immun.* 56:2822-2829.
110. Pires, S.M., E.G.Evers, P.W.van, T.Ayers, E.Scallan, F.J.Angulo, A.Havelaar, and T.Hald. 2009. Attributing the human disease burden of foodborne infections to specific sources. *Foodborne. Pathog. Dis.* 6:417-424.
111. Preidis, G.A., C.Hill, R.L.Guerrant, B.S.Ramakrishna, G.W.Tannock, and J.Versalovic. 2011. Probiotics, enteric and diarrheal diseases, and global health. *Gastroenterology* 140:8-14.
112. Raa, H., S.Grimmer, D.Schwudke, J.Bergan, S.Walchli, T.Skotland, A.Shevchenko, and K.Sandvig. 2009. Glycosphingolipid requirements for endosome-to-Golgi transport of Shiga toxin. *Traffic.* 10:868-882.
113. Rennie, R.P., J.H.Freer, and J.P.Arbutnott. 1974. The kinetics of erythrocyte lysis by Escherichia coli haemolysin. *J. Med. Microbiol.* 7:189-195.
114. Robien, M.A., B.E.Krumm, M.Sandkvist, and W.G.Hol. 2003. Crystal structure of the extracellular protein secretion NTPase EpsE of Vibrio cholerae. *J. Mol. Biol.* 333:657-674.
115. Rogers, M.T., R.Zimmerman, and M.E.Scott. 2009. Histone-like nucleoid-structuring protein represses transcription of the *ehx* operon carried by locus of

- enterocyte effacement-negative Shiga toxin-expressing *Escherichia coli*. *Microb. Pathog.* 47:202-211.
116. Saitoh, T., S.Iyoda, S.Yamamoto, Y.Lu, K.Shimuta, M.Ohnishi, J.Terajima, and H.Watanabe. 2008. Transcription of the *ehx* enterohemolysin gene is positively regulated by GrlA, a global regulator encoded within the locus of enterocyte effacement in enterohemorrhagic *Escherichia coli*. *J. Bacteriol.* 190:4822-4830.
 117. Sandkvist, M. 2001a. Biology of type II secretion. *Mol. Microbiol.* 40:271-283.
 118. Sandkvist, M. 2001b. Type II secretion and pathogenesis. *Infect. Immun.* 69:3523-3535.
 119. Sandvig, K., J.Bergan, A.B.Dyve, T.Skotland, and M.L.Torgersen. 2010. Endocytosis and retrograde transport of Shiga toxin. *Toxicon* 56:1181-1185.
 120. Scallan, E., R.M.Hoekstra, F.J.Angulo, R.V.Tauxe, M.A.Widdowson, S.L.Roy, J.L.Jones, and P.M.Griffin. 2011. Foodborne illness acquired in the United States—major pathogens. *Emerg. Infect. Dis.* 17:7-15.
 121. Schmidt, H., L.Beutin, and H.Karch. 1995. Molecular analysis of the plasmid-encoded hemolysin of *Escherichia coli* O157:H7 strain EDL 933. *Infect. Immun.* 63:1055-1061.
 122. Schmidt, H., H.Karch, and L.Beutin. 1994. The large-sized plasmids of enterohemorrhagic *Escherichia coli* O157 strains encode hemolysins which are presumably members of the *E. coli* alpha-hemolysin family. *FEMS Microbiol. Lett.* 117:189-196.
 123. Schmidt, H., C.Kernbach, and H.Karch. 1996. Analysis of the EHEC hly operon and its location in the physical map of the large plasmid of enterohaemorrhagic *Escherichia coli* O157:H7. *Microbiology* 142 (Pt 4):907-914.
 124. Schultz, J., R.R.Copley, T.Doerks, C.P.Ponting, and P.Bork. 2000. SMART: a web-based tool for the study of genetically mobile domains. *Nucleic Acids Res* 28:231-234.
 125. Scott, M.E., A.R.Melton-Celsa, and A.D.O'Brien. 2003. Mutations in *hns* reduce the adherence of Shiga toxin-producing *E. coli* O91:H21 strain B2F1 to human colonic epithelial cells and increase the production of hemolysin. *Microb. Pathog.* 34:155-159.

126. Semenov, A.M., A.A.Kuprianov, and A.H.van Bruggen. 2010. Transfer of enteric pathogens to successive habitats as part of microbial cycles. *Microb. Ecol.* 60:239-249.
127. Silhavy, T.J. and J.Beckwith. 1983. Isolation and characterization of mutants of Escherichia coli K12 affected in protein localization. *Methods Enzymol.* 97:11-40.
128. Smith, Y.C., S.B.Rasmussen, K.K.Grande, R.M.Conran, and A.D.O'Brien. 2008. Hemolysin of uropathogenic Escherichia coli evokes extensive shedding of the uroepithelium and hemorrhage in bladder tissue within the first 24 hours after intraurethral inoculation of mice. *Infect. Immun.* 76:2978-2990.
129. Stanley, P., V.Koronakis, and C.Hughes. 1998. Acylation of Escherichia coli hemolysin: a unique protein lipidation mechanism underlying toxin function. *Microbiol. Mol. Biol. Rev* 62:309-333.
130. Stoebel, D.M., A.Free, and C.J.Dorman. 2008. Anti-silencing: overcoming HNS-mediated repression of transcription in Gram-negative enteric bacteria. *Microbiology* 154:2533-2545.
131. Studier, F.W. and B.A.Moffatt. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* 189:113-130.
132. Tesh, V.L. 2010. Induction of apoptosis by Shiga toxins. *Future. Microbiol.* 5:431-453.
133. Vogel, M., J.Hess, I.Then, A.Juarez, and W.Goebel. 1988. Characterization of a sequence (hlyR) which enhances synthesis and secretion of hemolysin in Escherichia coli. *Mol. Gen. Genet.* 212:76-84.
134. Wang, P. and R.R.Granados. 1997a. An intestinal mucin is the target substrate for a baculovirus enhancin. *Proc. Natl. Acad. Sci. U. S. A* 94:6977-6982.
135. Wang, P. and R.R.Granados. 1997b. Molecular cloning and sequencing of a novel invertebrate intestinal mucin cDNA. *J. Biol. Chem* 272:16663-16669.
136. Welch, R.A. 1991. Pore-forming cytolysins of gram-negative bacteria. *Mol. Microbiol.* 5:521-528.
137. Welch, R.A. and S.Falkow. 1984. Characterization of Escherichia coli hemolysins conferring quantitative differences in virulence. *Infect. Immun.* 43:156-160.

138. Welch, R.A., R.Hull, and S.Falkow. 1983. Molecular cloning and physical characterization of a chromosomal hemolysin from *Escherichia coli*. *Infect. Immun.* 42:178-186.
139. Williams, P.H. 1979. Determination of the molecular weight of *Escherichia coli* alpha-haemolysin. *FEMS Microbiol. Lett.* 5:21-24.
140. Wu, H.C., M.Tokunaga, H.Tokunaga, S.Hayashi, and C.Z.Giam. 1983. Posttranslational modification and processing of membrane lipoproteins in bacteria. *J. Cell Biochem.* 22:161-171.
141. Yang, J., D.L.Baldi, M.Tauschek, R.A.Strugnell, and R.M.Robins-Browne. 2007. Transcriptional regulation of the yghJ-pppA-yghG-gspCDEFGHIJKLM cluster, encoding the type II secretion pathway in enterotoxigenic *Escherichia coli*. *J. Bacteriol.* 189:142-150.
142. Yu, R.R. and V.J.DiRita. 2002. Regulation of gene expression in *Vibrio cholerae* by ToxT involves both antirepression and RNA polymerase stimulation. *Mol. Microbiol.* 43:119-134.

WESTERN MICHIGAN UNIVERSITY

Recombinant DNA Biosafety Committee



Recombinant DNA Biosafety Committee

Project Approval Certification

For rDNA Biosafety Committee Use Only

Project Title: Characterization of the Novel Phosphatase-Like Protein YghJ

Principal Investigator: Maria Scott

IBC Project Number: 12-MSb

Date Received by the rDNA Biosafety Committee: October 31, 2011

Reviewed by the rDNA Biosafety Committee

Approved

Approval not required

Karin Esau

Vice Chair of rDNA Biosafety Committee Signature

12/01/2011
Date