The role of naturally occurring alleles of *rpoS* in *Escherichia coli*

By

Daniel Gyewu, BSc

A Thesis

Submitted to the School of Graduate Studies in

Partial Fulfillment of the Requirements

for the Degree

Master of Science

McMaster University

Abstract

Sigma S (RpoS), encoded by *rpoS*, is a subunit of RNA polymerase holoenzyme that controls the expression of many genes in stationary phase of various gram negative bacteria. *Escherichia coli* expresses these genes to withstand environmental stress and nutrient starvation. Several naturally occurring mutant alleles of the gene have been reported and indicate key differences from laboratory strains. We sought to explore the role of natural alleles of the *rpoS* gene (from non- K12 strains) and thus the sigma subunit relative to the K12 allele. To study the effect of the *rpoS* polymorphism on gene expression of RpoS regulon members, *rpoS* alleles from ECOR- 21, ECOR-28, ECOR-37 and ECOR-40 as well as MG1655 were cloned into the same background, MG1655 Δ*rpoS*:cat *osmY-lacZ*. Sequence analysis showed *rpoS* alleles from all the natural strains tested were different from MG1655 and each other. The strain with *rpoS* allele from ECOR-28 had increased expression of *osmY* and *katE* similar to MG1655. In contrast, *rpoS* allele from ECOR- 37 showed low expression of *osmY* but not as low as ECOR-21 and ECOR-40 which had expression similar to the *rpoS* mutant. Not surprisingly, recombinant strains with *rpoS* alleles from ECOR-21, ECOR-37 and ECOR-40 showed no expression of *katE* (HPII). These suggest that RpoS in ECOR-28 has high activity similar to wildtype K12 strain while RpoS in ECOR-21, ECOR-37 and ECOR-40 has very low or no activity. We conclude that natural *E. coli* strains have polymorphism in their *rpoS* ORF which cause variation in the regulatory activities of RpoS on its regulon.

MASTER OF SCIENCE (2008)   McMaster University
(Biology)   Hamilton, Ontario

TITLE: The role of naturally occurring alleles of *rpoS* in *Escherichia coli*

AUTHOR: Daniel Gyewu, B.Sc. (Brock University)

SUPERVISOR: Dr H.E. Schellhorn

NUMBER OF PAGES: xiii, 88

Acknowledgements

I thank Dr Schellhorn for the opportunity to work in his lab and for his supervision and guidance. I have acquired a lot of skills in the lab I consider valuable and this would have been impossible without him. I also thank Dr Sorger who was on my committee for his helpful comments and suggestions.

My life in the lab would have been very difficult without all the wonderful members of the Schellhorn lab. Each member provided invaluable assistance at one point or another, and I am truly grateful for their help. Most people are interested in this page of a thesis to see what the candidate meant to them. It was difficult to write this because everyone in the lab has been very wonderful to me, so I put your names in a hat and reasons for my gratitude in another hat and this is the order of names and matching reasons that I randomly picked.

I appreciate the fellowship of S. Chiang and for providing very useful and detailed insight on my project. I enjoyed coffee time and talking about life in general. Thanks to T. Dong for answering my many questions and for the fun times we shared comparing Ghanaian and Chinese cultures. My thanks to Thuraya for her encouragement and humour which got me through challenging times. Charlie kept me laughing as well as thinking about my work, and I really appreciate that. My gratitude also goes to Daniel my brother, or Li, as I call him, for his friendship and advice.

I can't forget Mirella, Eva, Nicole, Beckie, Rosemary and Ali. Finally, I thank my family for their unconditional love and support.

TABLE OF CONTENTS

List of Figures

List of Tables

1.0 Introduction

Prokaryotic RNA polymerase (RNAP) is a multi-subunit enzyme which comprises a core enzyme ($\alpha_2 \beta\beta'\omega$) with catalytic activity, as well as a sigma subunit required for promoter recognition and DNA melting, thus playing a role in transcription of a particular set of genes (Ishihama, 2000). Each sigma factor controls a set of genes. The number of sigma factors a bacterium possesses depends on its ability to adapt to its niche (Typas and Hengge, 2006). Sigma 70 (RpoD), encoded by *rpoD*, regulates most housekeeping genes expressed during exponential phase of growth (as reviewed in Hengge-Aronis, 1996). As the major sigma factor in *Escherichia coli*, the 70kDa protein is essential for its survival and therefore RpoD mutants are not viable (Keener and Nomura, 1993).

In *E. coli*, there are six alternative sigma factors which respond to different conditions including, stationary phase/starvation/stress ($\sigma^{38/S}$), heat shock ($\sigma^{32/H}$), nitrogen depletion ($\sigma^{54/N}$), flagellin requirement ($\sigma^{28/F}$), extracytoplasmic stress ($\sigma^{24/E}$) and requirement of citrate-dependent iron transport ($\sigma^{19/FecI}$) (Gruber and Gross, 2003). Each sigma factor enables RNA polymerase to recognize specific DNA sequences and hence target specific promoters. The relative level of each sigma factor is dependent on a number of factors including cell growth conditions and affinity of each sigma for the core enzyme (Colland *et al.*, 2002; Jishage *et al.*, 1996). For a group of genes under the control of the same species of sigma factor, the order and magnitude of transcription is determined by initiation efficiency and promoter strength (as reviewed in Ishihama, 2000).

In terms of amino acid sequence and modular organization, all the alternative sigma factors except σ^{54} share similarities with σ^{70} (Lonetto *et al.*, 1992). Even more similar in sequence, structure and molecular function to RpoD is RpoS (encoded by *rpoS*), although they have very distinct and complementary physiological functions (as reviewed in Hengge-Aronis, 1996). These two sigma factors control different sets of genes but have been found to recognize identical key promoter elements (Gaal *et al.*, 2001).

1.1 RpoS, The Master regulator of stress

Bacteria have developed signaling mechanisms that efficiently sense their surrounding environment and adjust cell physiology accordingly (Hirsch and Elliott, 2005). One of the key regulators in this process is RpoS (σ^s), a general stress sigma factor in *E. coli* that is induced in response to starvation and external stress usually encountered upon entry into stationary phase (as reviewed in Hengge-Aronis, 2002). RpoS regulates over 400 genes (Patten *et al.*, 2004) including genes involved in oxidative stress (Lange *et al.*, 1995), osmotic shock (McCann *et al.*, 1993), near-UV radiation (Sammartano *et al.*, 1986) and heat shock (Lange *et al.*, 1995; McCann *et al.*, 1993). Because of its numerous functions, several groups independently identified what is now known as the *rpoS* gene and named it differently (Hengge-Aronis, 2000). It was named *nuv,* because of its involvement in near-UV resistance (Tuveson and Jonas, 1979), *csi-2*, as a starvation inducible gene (Lange and Hengge-Aronis, 1991), then *katF, xthA* and *appR* for its regulation of catalase HPII (Loewen and Triggs, 1984), exonuclease III (Sak *et al.*, 1989) and acidic phosphatase (Touati *et al.*, 1986) respectively. All these genes were found to

be alleles of the same gene and thus named *rpoS* because of its importance in stress and stationary phase conditions (Lange and Hengge-Aronis, 1991).

The *rpoS* gene is 993 bp which codes for 330 amino acids and a 38kDa protein. The protein has 4 conserved regions as in proteins in the σ^{70} family, designated regions 1-4 and each region can be further divided into subregions : 1.1, 1.2, 2.1-2.4, 3.1, 3.2, 4.1 and 4.2 (Lonetto *et al.*, 1992). Genetic and biochemical analysis have been used to decipher the role of some of these subregions. In the σ^{70} family, region 1 is the least conserved while regions 2 and 4 are the most conserved (Ohnuma *et al.*, 2000). In the absence of the core RNA polymerase , region 1.1 prevents binding of the sigma factor to free DNA (Lonetto *et al.*, 1992). Region 4.2 is involved in the recognition of the -35 and -10 promoter while regions 2.3 and 2.4 facilitate binding of the sigma to the core enzyme (RNAP) and promoter melting respectively (Paget and Helmann, 2003). Regions 2.3, 2.4 and 4.2 of RpoD are highly homologous (over 70%) to RpoS, however RpoS promoters usually contain only a single DNA recognition element (Lee and Gralla, 2002).

With the high sequence similarities between RpoD and RpoS, it is expected that they will also recognize similar promoter sequences. Several RpoS-dependent promoter sequences show conservation at the -10 region while most of these promoters have weak or no -35 regions relative to RpoD, which possesses a consensus -10 and -35 regions (Becker and Hengge-Aronis, 2001; Espinosa-Urgel *et al.*, 1996). RpoS dependent promoters have the consensus sequence TATACT at their -10 region which is quite similar to the TATAAT of RpoD (Espinosa-Urgel *et al.*, 1996; Loewen and Hengge-

Aronis, 1994). In addition to their -10 region, RpoS-dependent promoters also have an intrinsic curvature, a physical characteristic that is not shared among RpoD-dependent promoters (Espinosa-Urgel and Tormo, 1993). These features differentiate between promoters recognized by the two sigma factors. That is, certain promoters could be recognized by both RpoS and RpoD , while others are recognized solely by either RpoS or RpoD depending on the presence of both -10 and -35 regions or a -10 region together with an intrinsic curvature (Loewen and Hengge-Aronis, 1994; Tanaka *et al.*, 1993).

1.2 Components of the RpoS regulon

There are about 2000 core RNA polymerase enzyme molecules in a typical *E. coli* cell (Ishihama, 1981), which are not sufficient to transcribe the approximately 4000 genes in its genome (Blattner *et al.*, 1997). Therefore, the particular genes transcribed, as well as the timing and frequency of transcription by RNA polymerase is very important (Ishihama, 1988). Sigma factors are switched depending on gene transcription requirement, providing an efficient mechanism to modify promoter recognition properties of the core enzyme (as reviewed in Ishihama, 2000). One of the first genes reported to be regulated by RpoS was *katE* (Schellhorn and Hassan, 1988). Currently, over 400 genes have been reported to be under the control of RpoS (Patten *et al.*, 2004). The RpoS regulon contains genes with diverse functions. However, the functions of some of the regulon members are still unknown. Functional categories of RpoS-dependent genes include stress management (11%), metabolic enzymes (19%), transport and/or membrane proteins (14%), regulatory proteins (8%), protein processing (5%) and others (43%) (Weber *et al.*, 2005). RpoS, in conjunction with RNA polymerase, regulates transcription

of about 10% of *E. coli* genes (Schellhorn *et al.*, 1998; Patten *et al.*, 2004). The current

report of the number of RpoS-dependent genes is consistent with an estimate by

Schellhorn *et al* (1998b), based on mutational analysis using random *lacZ* insertion

fusions. RpoS dependent genes are dispersed over the *E. coli* genome. However, a few

clusters can be found with numerous RpoS-dependent genes condensed in a small area. A

typical example is a 91kb region around 79.3 min on the genome that houses 29 RpoS-

dependent genes (Weber *et al.*, 2005).

There are more genes that are regulated negatively by RpoS than positively, in

stationary phase of a wild type *E. coli* strain (Patten *et al.*, 2004). The negatively

regulated genes can be classified into 3 groups including flagellar structural and

accessory signaling protein genes, energy and metabolism genes, and a cluster of genes

around Rac prophage region (Patten *et al.*, 2004).

1.3 Polymorphism in *rpoS*

The evolution of *rpoS* has become a popular subject for researchers because of its

predominance in natural *E. coli* and Salmonella populations (Ferenci, 2003). Distinct

alleles of the *rpoS* gene have been found in different isolates of *E. coli* (Atlung *et al.*,

2002), suggesting that evolutionary selection may alter *rpoS* gene expression and/or

activity. These alleles carry point mutations or deletions and/or insertions of extra DNA

sequences in various strains of the bacteria (Atlung *et al.*, 2002). Natural *E. coli* isolates

possess *rpoS* polymorphisms that are distinct from those found in K12 strains (Jishage

and Ishihama, 1997; Waterman and Small, 1996). The mutation rate in *E. coli* is 5×10^{-10}

per base per generation (Drake *et al.*, 1998). Considering the size of *rpoS* (almost 10^3 bp),

if for example there are 10^8 cells in exponential phase, there would be approximately 50

cells that would have a mutation in the ORF of *rpoS*. This number will increase as the

cells continue to grow. Early reports of polymorphism in *rpoS* were published when the

gene was known as *katF*. Several base differences were found in the *katF* gene in six *E.*

coli strains resulting in variation in phenotypes including catalase HPII synthesis,

carotenoid synthesis and sensitivity to acid and heat (Ivanova *et al*., 1992). Amino acid

sequences of RpoS isolated from clinical and environmental isolates such as O157:H7

and O55:H7 show some variations, while laboratory strains DH1 and K12 had even more

variable sequences relative to the majority of other *E. coli* sequences (Ferreira *et al*.,

1999). Even the region in the genome between *rpoS* and *mutS* is highly polymorphic

(Herbelin *et al*., 2000). MutS mutants represent more than 1% of independent isolates, a

finding that could explain the quick development of antibiotic resistance since these

mutant alleles increase mutation rates and recombination among various species (Ferenci,

2003; LeClerc *et al*., 1996).

One possible explanation for *rpoS* polymorphism is competition between RpoD and

RpoS for core RNA polymerase (Ferenci, 2003; Notley-McRobb *et al*., 2002). Mutations

in *rpoS* which affect the proper function of the sigma factor, results in more core RNA

polymerase available for RpoD. This leads to expression of more RpoD-dependent genes

and thus better nutrient utilization of otherwise poor substrates (King *et al*., 2004) while

making cells prone to external stresses like heat shock and low pH (Dodd and Aldsworth,

2002). Mutations in *rpoS* led to enhanced expression of genes involved in high-affinity

glucose utilization system necessary for the hunger response (Notley-McRobb *et al*.,

2002). When acid stress was applied to a glucose-limited culture, the phenotype and frequency of *rpoS* mutations decreased (Notley-McRobb *et al.*, 2002). Although it is difficult to explain why such an important gene for bacterial survival and fitness as *rpoS* is highly polymorphic, it is not surprising since bacteria freely accumulate mutations that will favour and improve their survival in adverse conditions (Chen *et al.*, 2004; Ferenci, 2003). For example, some *rpoS* mutants have the ability to outgrow wildtype cells after a long incubation time, a growth advantage that allows the mutants to take over the population in stationary phase (Zambrano and Kolter, 1996). RpoS mutants exhibit better growth on succinate because of enhanced metabolism of TCA cycle intermediates (Chen *et al.*, 2004). Similarly, strains with *rpoS* mutations form more biofilm (Corona-Izquierdo and Membrillo-Hernandez, 2002) and outcompete wildtype strains in colonization of mouse intestine (Krogfelt *et al.*, 2000). Many natural isolates of *E. coli* have lost or partially lost RpoS function (as reviewed in Ferenci, 2003). Chen *et al* (2004) proposed that selection for loss of RpoS function is evolutionarily advantageous because it permits cells to "turn off" expression of the regulon when it is not needed and turn it on again when necessary (Figure 1).

A frequently reported *rpoS* mutation in *E. coli* is found at codon 33. The codons CAG, CTG, TAT, TCG and TAG have all been found in various K12 strains, while non-K12 strains possess GAG at this position (Atlung *et al.*, 2002; Ferreira *et al.*, 1999). Thus, several K12 strains have the amber codon TAG and most of the rest have CAG, which codes for glutamine rather than glutamate, GAG, in the non K12 strains. Atlung *et. al.* (2002) suggested a likely evolutionary connection between these mutants (Figure 2).

GAG is the ancestral codon since it is found in all non-K12 strains (Atlung et al., 2002;

Subbarayan and Sarkar, 2004a). This codon (GAG) mutated to TAG in the original K12

strains followed by mutations to restore RpoS function by pseudo-reversion or mediation

by an amber suppressor (Atlung *et al.*, 2002). Most of the pseudo-revertants had CAG

while others include TAT, TCG and CTG. A suppressor mutation is a second mutation at

a distinct site that modifies the phenotypic effect of a previous mutation. In certain

mutants, a point mutation can change a codon which encodes an amino acid into a chain

termination code such as TAG described above. A mutated tRNA which has a single base

change in its anticodon will read the termination codon as an amino acid, and in the case

above, restore RpoS function (Goodman *et al.*, 1968).

Increased metabolic ability
Loss of RpoS function

Decreased metabolic ability
Gain of RpoS function
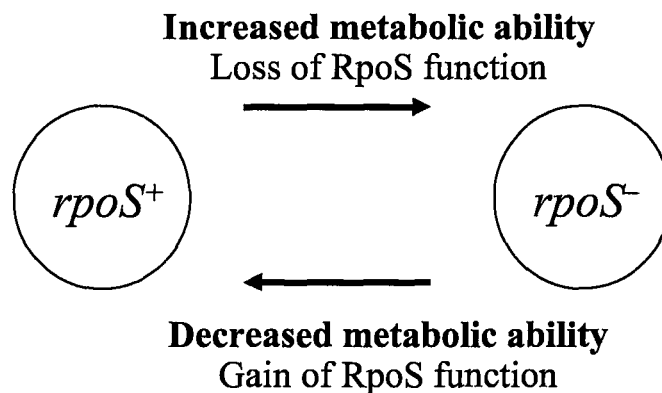
Figure 1: A molecular switch of *E. coli* which selects for loss and gain of RpoS function
by mutations (Chen *et al.*, 2004)

Strains with the amber codon, TAG, as codon 33 have moderate catalase activity while

those with CAG and TAT have a relatively higher catalase activity (Subbarayan and

Sarkar, 2004a; Visick and Clarke, 1997). This amber codon results in a non-functional 32

amino acid long peptide, however, an internal secondary translational initiation region

(STIR) has been hypothesized to initiate translation of a truncated but functional RpoS (Subbarayan and Sarkar, 2004b).

TAT (Y)

GAG (E) ⟶ TAG (am) ⟶ CAG (Q) ⟶ CTG (L)

TCG (S)
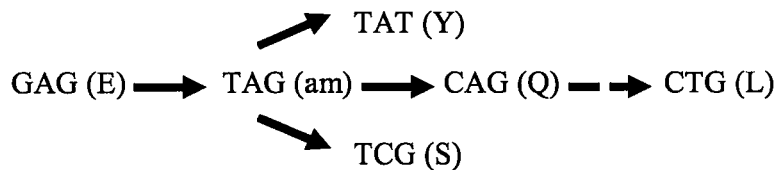
Figure 2: Possible evolutionary relationship of codon 33 in *E. coli* strains (Atlung *et al.*, 2002).

Polymorphism in *rpoS* is not restricted to laboratory strains but also found in natural populations of *E. coli* including pathogenic strains. Characterization of *rpoS* alleles in various natural *E. coli* isolates show several nucleotide changes, some leading to changes in amino acid while others remain silent mutations. Sequence analysis of *rpoS* from four *E. coli* isolates from the Hamilton Habour, ON, showed a C to G (relative to K12) mutation in all the strains, in addition to 11 silent mutations in one strain and a TGGAG deletion at codon 86 in another strain, leading to a premature stop codon after 10 codons (Chiang and Schellhorn, unpublished). A study of *rpoS* in ten clinical *E. coli* isolates showed a C to G mutation in all the strains (relative to K12) and at least one mutation in six of them leading to changes in amino acid (Figure 3). In addition, the strain N994390 had a premature stop codon at nucleotide position 733 leading to 242 amino acids instead of 330 (Figure 3). The shiga-like toxin producing *E. coli* (SLTEC) strain PS12 serotype O157:H7 has an 11 base pair duplication at nucleotide 131 resulting in an open reading frame (ORF) shift and eventually a stop codon at position 223, while O78:K80 has a deletion of glutamic acid (aa 74-76) within a conserved motif made up of three glutamic

acid residues (Ferreira *et al.*, 1999; Waterman and Small, 1996). The pathogenic strain O157:H7 (W2-2) has a duplication region at -97 to -87, and CCG (Pro) at position 214-216 while ACG (Thr) is found in other O157:H strains such as EDL933 and H1071 as well as K12 strain DH1 (Robey *et al.*, 2001). Sequence analysis of *rpoS* from *E. coli* O55:H7, O157: NM and O78:K80 showed variation between these strains and O157:H7 isolates at position 543 (Ferreira *et al.*, 1999). The polymorphism of *rpoS* sequence and function still remains an enigma in laboratory as well as natural populations of *E. coli* and gram negative bacteria in general (Notley-McRobb *et al.*, 2002). The relationship between these different alleles and how they are/were selected for are yet to be elucidated. Alleles of *rpoS* are widely distributed and there must be several others present that are still not known, since the phenotypic differences between the various alleles are not well known.

1.4 Regulation of RpoS

When a bacterium encounters a stressful condition, RpoS is induced, and coordinates a regulon in a signaling network with several downstream targets, initiating a response accordingly. The hierarchical nature of this network makes way for numerous control points and secondary regulators (as reviewed in Hengge-Aronis, 2002). Stress regulation in *E. coli* is primarily dependent on the cellular levels of RpoS, which is regulated by the sigma factor's synthesis and accumulation (as reviewed in Hengge-Aronis, 2002). The synthesis and accumulation of RpoS is dependent on factors and mechanisms that affect the transcription, translation and proteolysis of the sigma subunit, as well as the assembly of the holoenzyme (Loewen *et al.*, 1998).

Protein levels of RpoS are very low during exponential phase, however relatively high amounts of its mRNA are available that do not respond to stresses in the presence of high RpoS concentration. This suggests that most of its regulation is posttranscriptional although transcriptional control does exist (Lange and Hengge-Aronis, 1994a). In stationary phase, the protein level of RpoS increases to 30-50% of RpoD, which maintains a concentration of 50-80 fmol/μg in both exponential and stationary phase (Jishage and Ishihama, 1995). Transcription of *rpoS* is driven by several promoters. Two weak promoters, *nlpDp1 and nlpDp2* located upstream of *nlpD* (encodes a lipoprotein) direct the expression of a growth phase independent basal level of *rpoS* (Lange and Hengge-Aronis, 1994b). In stationary phase however, the major promoter *rpoSp,* positioned within *nlpD* generates a monocistronic mRNA of *rpoS* (Lange *et al.*, 1995). Positive regulators of *rpoS* transcription include guanosine 3'5'-bispyrophosphate (ppGpp) and polyphosphate whiles cAMP-CRP and EIIA (Glc) negatively regulate this process (as reviewed in Hengge-Aronis, 2002; Loewen *et al.*, 1998). Other reported regulators include BarA, homoserine lactone, homocysteine thiolactone, *oxyS* RNA, UDP-glucose,cellular NADH to NAD^+ ratio as well as acetate and other weak acids (as reviwed in Hengge-Aronis, 2002).

Translation of *rpoS* mRNA is mainly dependent on its secondary structure (Loewen *et al.*, 1998). The ribosome-binding site and initiation codon, the major translational control spots, are situated in locations which make it difficult for the ribosome to access (Lange and Hengge-Aronis, 1994a). Therefore regulation of ribosomal access to these two sites either through secondary structure or protein binding is the main mechanism in *rpoS*

regulation via translation (Loewen *et al.*, 1998). Hfq, an RNA-binding protein enhances translation initiation by disrupting the secondary structure of *rpoS* mRNA (Brown and Elliott, 1996; Muffler *et al.*, 1997b). Contrarily, H-NS, a nucleoid histone-like protein, functions to reduce RpoS levels by inhibiting Hfq (Barth *et al.*, 1995). Other regulatory factors of *rpoS* mRNA structure and hence translation include HU, DsrA, LeuO, RprA, DnaK, DksA, CspC, CspE, EIIA (Glc) and UDP-glucose (Hengge-Aronis, 2002).

RpoS is usually degraded by ClpXP in exponential phase as a way of regulating its accumulation. However, this is abolished when cells enter stationary phase (Zgurskaya *et al.*, 1997). This regulation is facilitated by SprE, also known as RssB (Muffler *et al.*, 1996), because ClpPX has a relatively constant concentration and activity (Schweder *et al.*, 1996). DnaK, on the other hand, protects RpoS from ClpPX, thereby increasing its accumulation (Muffler *et al.*, 1997a). RpoS becomes more stabilized in conditions such as osmotic shock, high cell density and depletion of nutrients like carbon, phosphate and ammonia (as reviewed in Hengge-Aronis, 2002).

1.5 *In vivo* cloning of recombinant DNA

DNA engineering and the discovery of plasmids, non-chromosomal DNA in bacteria, have made significant contributions to biology as well as launch a whole new subject area called molecular cloning (Li and Elledge, 2005; Smith and Wilcox, 1970). Introducing foreign DNA into *E. coli* and the late twentieth century pioneering methods of DNA ligase mediated ligation of DNA into a vector are still widely used today (Zhang *et al.*, 2000). The advent of sequencing and thus availability of genome information of organisms has created the need to analyze several genes simultaneously. To meet this

demand, new techniques such as microarray analysis and recombinant DNA cloning

methods have been developed (Li and Elledge, 2005).

Several forms of recombinant DNA cloning techniques have emerged including ET

cloning (Zhang *et al.*, 1998), the Gateway cloning system (Suzuki *et al.*, 2005) and

MAGIC (Li and Elledge, 2005). ET cloning makes use of indigenous bacterial

homologous recombination through *rec*E and *rec*T proteins, hence the name "ET-

cloning" (Zhang *et al.*, 1998). This method eliminates the limitation in using restriction

sites for cloning, and thus allowing easy manipulation of DNA (Zhang *et al.*, 1998).

Gateway cloning system is based on *in vitro* bacteriophage λ recombination system

(Suzuki *et al.*, 2005). The method uses PCR to generate a fragment with homology

regions for recombination. The fragment is then introduced into a vector by *in vitro*

recombination catalyzed by BP clonase. The vector is then ligated through intramolecular

recombination and the recombinant plasmid is transformed into *E. coli* (Suzuki *et al.*,

2005). Mating-assisted genetically integrated cloning, MAGIC, involves the use of

bacterial conjugation, *in vivo* site specific restriction endonucleases and homologous

recombination to facilitate the transfer of DNA on a donor vector from one bacterium to a

recipient plasmid in another strain (Li and Elledge, 2005). The donor vector is incapable

of replicating in the recipient strain because of a conditional origin of replication, ori

R6Kγ, which only functions in *pir*[+] strains. After conjugation of the donor vector into the

recipient strain, the DNA of interest which is flanked by homology regions is relocated

into the recipient plasmid by homologous recombination facilitated by *I-SceI*

endonuclease and a lambda recombinase system, both under the control of an inducible

promoter. Multiple approaches are used to reduce background non-recombinants

including placing *lacO* on the transfer vector and the use of a suicide marker, *pheS,*

which is lethal in the presence of chlorophenylalanine. The *I-SceI* endonuclease system

ensures only re-ligated fragments of the recipient plasmid will be replicated and thus

survive (Li and Elledge, 2005).

The natural role of homologous recombination is to maintain genome integrity and

plasticity, then to facilitate cellular replication and DNA repair (Volodin *et al.*, 2005).

The major recombination pathway of *E. coli*, RecBCD, becomes inactivated when the

bacteria is infected by phage λ. The phage then recombines through the Red

recombination pathway encoded by λ (Hill *et al.*, 1997). In the laboratory, the

bacteriophage λ Red recombination system provides a way to genetically modify genes or

segments of chromosomes with double stranded DNA (Murphy, 1998). The λ Red system

comprises of the Gam, Exo (α) and Beta (β) proteins. All three proteins are required for

recombination involving double stranded DNA whiles only the Beta protein is necessary

for single stranded DNA recombination (Ellis *et al.*, 2001; Yu *et al.*, 2000). Red

mediated molecular events take place on both sides of a DNA fragment to produce a

recombinant. The mechanism is initiated by digestion of the 5' ended strand by λ

endonuclease, exposing a 3' ended single strand. RecA and λ β protein then facilitates

invasion of the unbroken double strand by the 3' ended strand. This interaction results in

the formation of a three stranded junction which is converted into a holliday junction by

RuvAB and/or RecQ helicase. The intermediate is then resolved into a recombinant

molecule by RuvC (Poteete, 2001).

Homologous recombination offers an alternative approach to manipulation of DNA which eliminates the use of restriction enzymes and DNA preparation, saving time, effort and money (Zhang *et al.*, 1998; Li and Elledge, 2005). It allows the generation of recombinant genomes by the combination of already existing ones. As a genetic tool, the use of homologous recombination includes rescuing a plasmid between two linear fragments (Oliner *et al.*, 1993), incorporation of a fragment on a temperature sensitive plasmid into *E. coli* chromosome to replace or delete a gene (Hamilton *et al.*, 1989), and recombination of a linear DNA fragment into a circular target (Jasin and Schimmel, 1984). Homologous recombination between PCR fragments and yeast genomes or yeast artificial chromosome (YAC) are very common because of its high efficiency and simplicity (Baudin *et al.*, 1993; Zhang *et al.*, 1998). Cloning methods based on *in vivo* homologous recombination appear to be more efficient than conventional *in vitro* methods (Li and Elledge, 2005). In addition, the range of products it could generate is not achievable by *in vitro* means, such as in long cloned chromosome fragments (Volodin *et al.*, 2005).

1.6 Enzyme kinetics

Enzymes are protein catalysts that speed up the rate of a chemical reaction without being consumed in the process. This is achieved by binding to a substrate and lowering the activation energy required to convert it to a product. Factors which affect enzyme activity include concentration of substrate molecules, enzyme concentration, temperature, pH and presence of inhibitors. Enzyme assays are commonly used in the laboratory to determine enzyme activity.When using enzyme assays it is important to use optimal conditions to

eliminate artifacts. This condition should provide a linear response in the assay. That is,

the activity of the enzyme (rate) should increase with increasing enzyme concentration.

Optimal condition also implies a substrate concentration that results in a maximum

activity for an uninhibited enzyme. The activity of an uninhibited enzyme will increase

with increasing substrate concentration until a maximum velocity (Vmax) is reached

when the enzyme is fully saturated with substrate molecules. The desired substrate

concentration is this plateau when the enzyme is functioning at a constant velocity,

independent of substrate concentration (London *et al.*, 1975).

Catalase is a commonly assayed water soluble enzyme found in many bacteria and

eukaryotes that accelerates the lysis of hydrogen peroxide into water and oxygen (Li and

Schellhorn, 2007; Switala and Loewen, 2002). The rate of H_2O_2 decomposition is linear

with catalase levels of 0.05 to 1.0 units in a reaction volume of 250µl (Li and Schellhorn,

2007). The reaction obeys Michaelis-Menten kinetics and in a pH range of 4.7 to 10.5,

catalytic activity is independent of pH (Jones and Suggett, 1968).

1.7 Project objectives

Polymorphisms in *rpoS* are common in *E. coli*. In spite of the widespread distribution of

rpoS alleles, only a few studies have examined the effect of these polymorphisms on gene

expression. Most of these studies use and provide information obtained from laboratory

or K12 strains. Several questions remain unanswered that are yet to be elucidated in this

area of molecular genetics and genomics. These include the mechanism of selection of

rpoS alleles, the phenotypic difference between these alleles, the reason *rpoS* is highly

polymorphic if it is important for cell survival, the amount of functional similarity

between natural and laboratory *E. coli* strains, and the effect of *rpoS* polymorphism on gene expression in the large RpoS regulon.

This project examined how closely related *rpoS* in natural *E. coli* populations function relative to the laboratory strains. We obtained and used 10 strains from the ECOR collection of natural *E. coli* isolates. ECOR is a group of 72 reference *E. coli* strains isolated from various hosts and geographic locations, which are representative of the diversity in the genotype of the species and therefore established as the standard reference strains of *E. coli* from natural populations (Ochman and Selander, 1984). Exposure of the ECOR strains to hydrogen peroxide showed diverse levels of catalase activity which meant there was a possibility that *katE* (which encodes HPII) could be regulated differently by RpoS in those strains (Table 1). Sequence analysis of the open reading frame of these strains showed some mutations in nucleotide sequences leading to differences in codons or amino acid as well as silent mutations which maintained the amino acid sequences (Figure 3)(Dong and Schellhorn, unpublished data). We hypothesize that variation in the regulation of RpoS-dependent genes in natural isolates of *E. coli* is a result of *rpoS* polymorphism.

The objective of this project is to investigate the effect of *rpoS* polymorphism on gene expression within the large RpoS regulon. To achieve this, we placed selected natural *rpoS* alleles into a common background, a laboratory wildtype K12 strain with a *lacZ* fusion to *osmY*, a RpoS-dependent gene (as a reporter of gene expression) and tested the expression of the alleles using β-galactosidase and catalase assays.

Table 1: Catalase activity of ECOR strains

Strain	Genotype	Catalase activity
ECOR-1	ON:HN	+++
ECOR-2	ON:H32	+++
ECOR-3	O1:NM	+++
ECOR-4	ON:HN	+++
ECOR-7	O85:HN	++
ECOR-10	O6:H10	+++
ECOR-21	O121:HN	+
ECOR-28	O104:NM	++++
ECOR-37	ON:HN	+
ECOR-40	O7: NM	+
MG1655	F- λ- rph	++++
HS2718	F- λ- rph $\Delta rpoS$	+

Catalase activity was determined by rate of evolution of O_2 after dropping 5µl of H_2O_2 onto individual colonies.

| | | DNA melting | |
| Function | Inhibition | Core binding | -10 Recognition | Core binding | -35 Recognition |

| Conserved Region | 1.1 | 1.2 | 2.1 | 2.2 | 2.3 | 2.4 | 3.1 | 3.2 | 4.1 | 4.2 |

1 .... 261 262 .... 514 .... 750 751 .... 990 nt

		97[a]	163[b]		264[b]	354[b]	387[b]	429[b]	431[c]	462[b]	543[b]	574[b]	729[d]	733[b]	761[b]	821[b]	944[b]	953[b]
K12	MG1655.	ATG...CAG...T..........C......C...C...A..ACA.T.......C..T...-..C.........C...G...C..G.....
	ECOR1	ATG...GAG...T..........C......C...C...A..ACA.T.......C..T...-..C.........C...G...T..G.....
	ECOR2	ATG...GAG...T..........C......C...C...A..ACA.T.......C..T...-..C.........C...G...T..G.....
	ECOR3	ATG...GAG...T..........C......C...C...A..ACA.T.......C..T...-..C.........C...G...T..G.....
	ECOR4	ATG...GAG...T..........C......C...C...A..ACA.T.......C..T...-..C.........C...G...C..G.....
	ECOR7	ATG...GAG...T..........C......C...C...G..ACA.T.......C..T...-..C.........T...G...C..A.....
Natural	ECOR10	ATG...GAG...T..........C......C...C...A..ACA.T.......C..T...-..C.........C...G...T..G.....
isolates	ECOR21	ATG...GAG...T..........C......C...C...A..ACA.T.......C..T...A.C.........C...G...T..G.....
	ECOR28	ATG...GAG...T..........C......C...C...A..ACA.T.......C..T...-..C.........C...G...C..G.....
	ECOR37	ATG...GAG...T..........C......C...T...A..ATA.T.......C..T...-..C.........C...A...C..G.....
	ECOR40	ATG...GAG...C..........T......T...C...A..ACA.C.......C..C...-..T.........C...A...C..G.....
	CL3	ATG...GAG...T..........C......C...C...A..ACA.T.......C..T...-..C.........C...A...C..G.....
	EDL933	ATG...GAG...T..........C......C...C...A..ACA.T.......A..A...-..C.........C...A...C..G.....
	R82F2	ATG...GAG...A..........C......C...C...A..ACA.T.......C..T...-..C.........C...G...C..G.....
	N004859	ATG...GAG...T..........C......C...C...A..ACA.T.......C..T...-..C.........C...G...C..G.....
Pathogenic	EC3377	ATG...GAG...T..........C......C...C...A..ACA.T.......C..T...-..C.........C...G...C..G.....
isolates	EC2044	ATG...GAG...T..........C......C...C...A..ACA.T.......C..T...-..C.........C...G...C..G.....
	N004067	ATG...GAG...T..........C......C...C...A..ACA.T.......C..T...-..C.........C...G...C..A.....
	EC6484	ATG...GAG...T..........C......C...C...A..ACA.T.......C..T...-..C.........C...G...C..A.....
	CL106	ATG...GAG...T..........C......C...C...A..ACA.T.......C..T...-..C.........C...A...C..G.....
	N994390	ATG...GAG...T..........C......C...C...A..ACA.C.......C..C....TAG

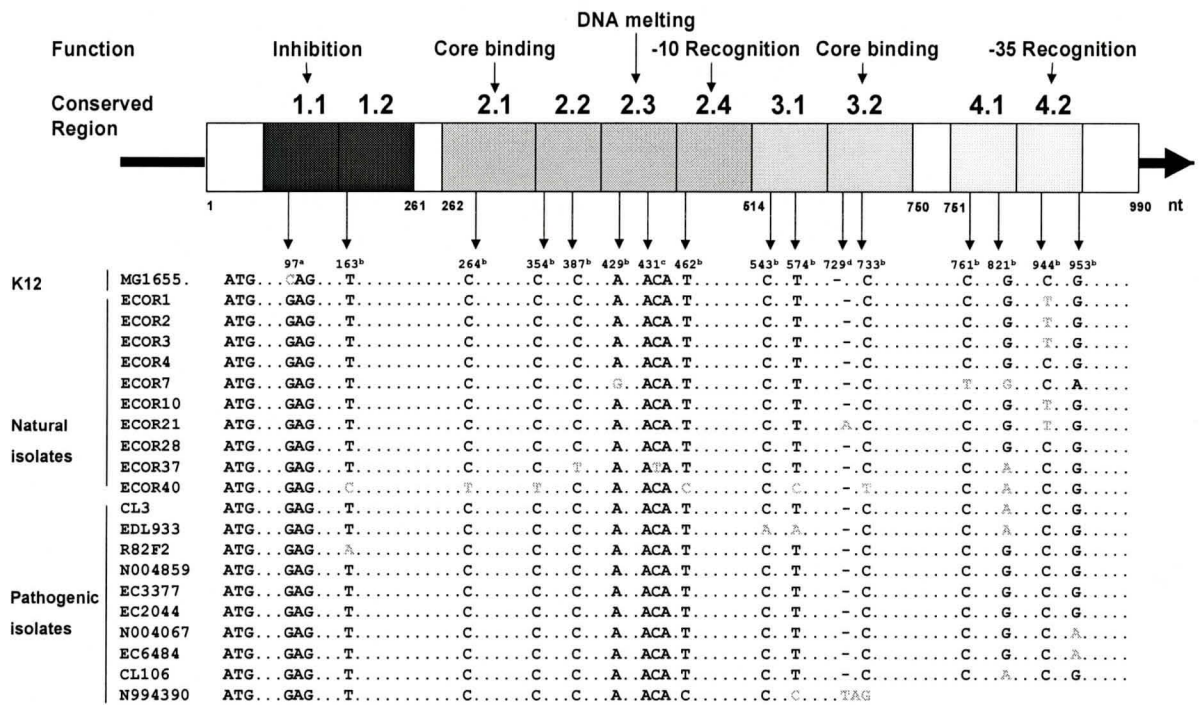**Figure 3:** Mutations in *rpoS* of natural and pathogenic strains of *E. coli* in comparison to a laboratory strain, MG1655. a, b – mutations leading to change in amino acids. c- silent mutations. d- insertions causing truncation of amino acid sequence (252 aa in ECOR 21 and 242 aa in N994390). a- Mutation in this codon has been reported by (Atlung *et al.*, 2002). DNA were amplified and sequenced by Tao Dong.

2.0 Materials and Methods

2.1 Media and growth conditions

Strains were routinely grown in Luria-Bertani (LB) broth containing the appropriate

antibiotics at 37°C with shaking at 200 rpm. Strains with pKD46 were cultured at 30°C.

The antibiotics used were ampicillin (150 μg/ml), kanamycin (50 μg/ml),

chloramphenicol (20 μg/ml) and spectinomycin (50 μg/ml).

2.2 Bacterial strains and plasmids

Bacterial strains and plasmids used and constructed in this study are listed in tables 2 and

3 respectively.

Table 2: Bacterial strains used in this study

Strain	Genotype or Description	Source/Reference
ECOR-1	ON:H	(Ochman and Selander, 1984)
ECOR-2	ON:H32	(Ochman and Selander, 1984)
ECOR-3	O1:NM	(Ochman and Selander, 1984)
ECOR-4	ON:HN	(Ochman and Selander, 1984)
ECOR-7	O85:H	(Ochman and Selander, 1984)
ECOR-10	O85:H	(Ochman and Selander, 1984)
ECOR-21	O121:HN	(Ochman and Selander, 1984)
ECOR-28	O104:NM	(Ochman and Selander, 1984)
ECOR-37	ON:HN	(Ochman and Selander, 1984)
ECOR-40	O7: NM	(Ochman and Selander, 1984)
HS2719	as HS2718 but *rpoS* [ECOR37]	This study
HS2720	as HS2718 but *rpoS* [ECOR21]	This study
HS2721	as HS2718 but *rpoS* [ECOR40]	This study
HS2722	as HS2718 but *rpoS* [MG1655]	This study
HS2723	as HS2718 but *rpoS* [ECOR28]	This study
MG1655	prototrophic *E.coli* K-12,F- λ- rph	Yale University
HS2213	as MG1655 but Δ*rpoS::cat*	Mark Kirchhof
HS2703	as HS2213 but *osmY-lacZ*	This study
HS2718	as HS2703 but Δcat	This study
BW19851	F-, RP4-2(Km::Tn7,Tc::Mu-1),ΔuidA3::pir+, recA1-, endA1-, hsdR17-	(Goulian *et al*., 2006)
DH10βF'DOT	recA1 endA1 rpsL umuC::pir116-frt F'(lac+pro+ΔoriT::Tc)	(Li and Elledge, 2005)

Table 3: Plasmids used in this study

Plasmid	Genotype or Description	Source/Reference
pMAGIC1	oriR6Kγ,oriT and PGK promoter-T7-His6-MY3. KanR	(Li and Elledge, 2005)
pBSPheSH3H4	*pheS* flanked by H3 and H4 sites. ampR	(Li and Elledge, 2005)
pMarMar4	R6K ori, lacZ-gfp3.1-lacZ, SacB, mob RP4. cat	(Goulian *et.al.*, 2006)
pKD46	λ Red recombinase . bla	(Datsenko and Wanner, 2000)
pCP20	γcI857(ts) ts-rep, FLP recombinase. bla	(Datsenko and Wanner, 2000)
pDOG1	ampR and pheS region from pBSPheS cloned into NotI-NotI cut pMAGIC1	This study
pDOG3	same as pDOG1 with deletion of extra H4 region	This study
pMar21	pMarMar4 with rpoS from ECOR-21 replacing SacI-Bsu361 region	This study
pMar28	pMarMar4 with rpoS from ECOR-28 replacing SacI-Bsu361 region	This study
pMar37	pMarMar4 with rpoS from ECOR37 replacing SacI-Bsu361 region	This study
pMar40	pMarMar4 with rpoS from ECOR-40 replacing SacI-Bsu361 region	This study
pMarMG	pMarMar4 with rpoS from MG1655 replacing SacI-Bsu361 region	This study

2.4 Construction of general conjugation vector

The plasmid pBSpheSH3H4 (600ng) was linearized by digestion with *Bsp120I* (40 units)

and 1X Buffer B (Fermentas Canada Inc.) following manufacturer's instructions. The

linear fragment (100ng) was then digested partially by addition of 10units of *BspHI* (

PagI), 1X Buffer O (Fermentas Canada Inc.) and 0.2mg/ml ethidium bromide, then

incubated at room temperature for 30 min. The mixture was run on a 1% agarose gel for

45min at 100V to separate the fragments. The fragment size 3.1kb which had *ampR* and

pheS was isolated using the QIAEX II gel extraction kit (QIAGEN Inc.). pMAGIC1

(240ng) was digested with *NotI* (2 units) and *NcoI* (4 units) (Fermentas Canada Inc)

following manufacturer's instructions. The digested pMAGIC1 fragment and partially

digested pBSpheS H3H4 were then ligated in a 7:9 ratio using T4 DNA ligase and

instructions from Fermentas Inc. A *pir*[+] host strain, BUN21, was transformed with the

ligation mixture (20ng) using TSS as previously described (Chung *et al.*, 1989).

Transformants were selected on LB agar containing kanamycin and ampicillin. The

recombinant plasmid was sequenced by Mobix Lab, McMaster University (Hamilton,
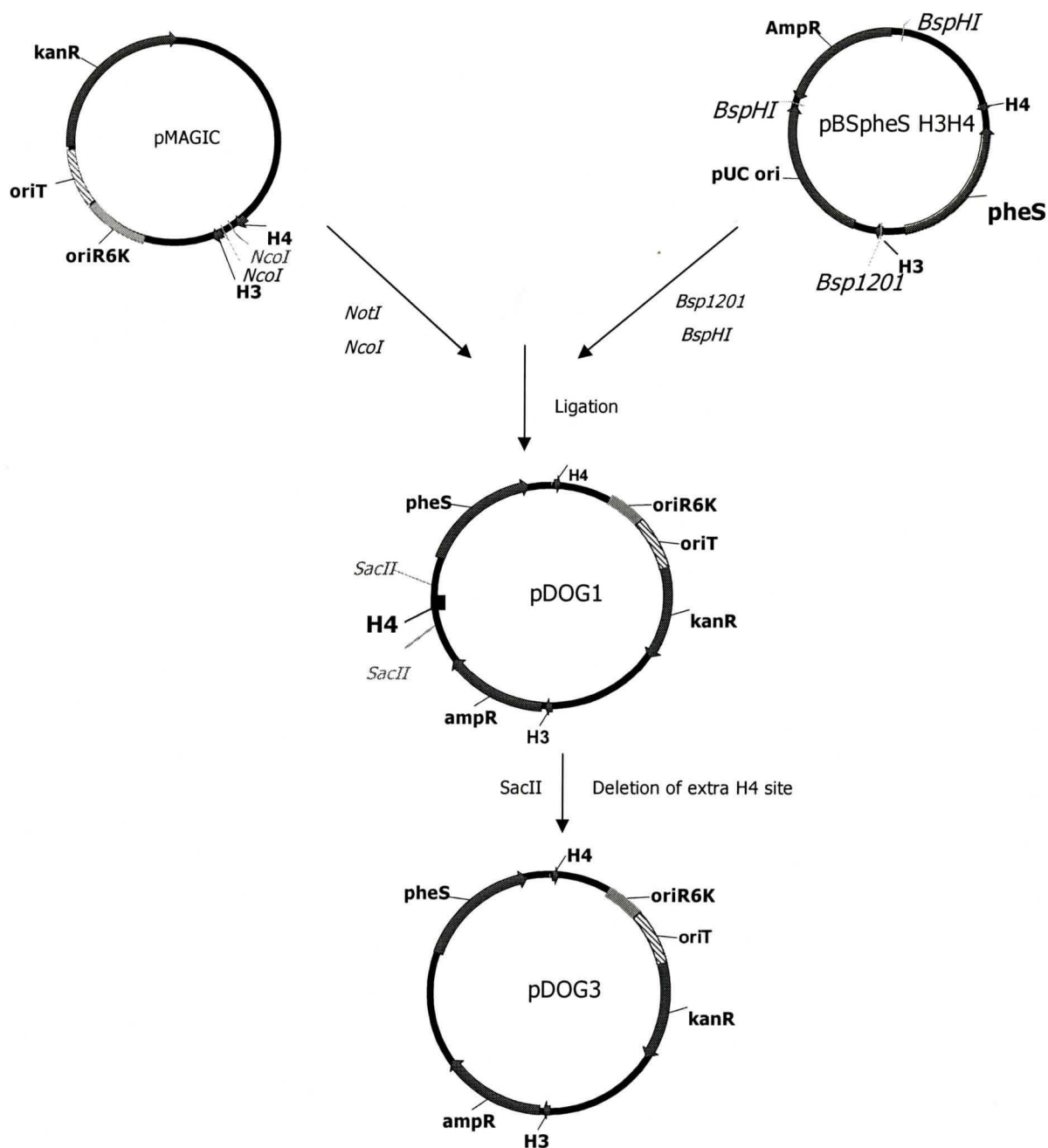
ON) for confirmation.

Figure 4: Construction of the general conjugation vector, pDOG3 from pMAGIC and pBSPheSH3H4. pBSPheS was digested with *BspHI* and *Bsp1201* and ligated to corresponding compatible ends of *NcoI* and *NotI* cut pMAGIC respectively. The extra H4 site in the resulting plasmid, pDOG1 was deleted by *SacII* digestion forming pDOG3.

2.5 Transduction of recipient strain with P1 lysate (Miller, 1992)

Transduction was used to transfer *osmY-lacZ* from a P1 lysate (prepared by Mirella Younes) to the recipient strain, MG1655 Δ*rpoS*::cat Δ*lacZYA*. A 50ml flask containing 10ml of 1XLB with 5mM $CaCl_2$ was inoculated with a single colony of the recipient strain. After incubation and shaking at 37°C overnight, cells were washed and resuspended in 1X LB containing 5mM $CaCl_2$. A mixture containing 100µl of the cell suspension and 100µl P1 lysate was incubated for 15 minutes along with each component separately to serve as controls. The cell suspension-P1 lysate mixture was then washed twice with 10ml of 1X LB containing 100mM sodium citrate resuspended in 10ml of the LB-sodium citrate solution and incubated for 45 min at 37°C. The cells were centrifuged and resuspended in 200µl of the remaining media in the tube. A volume of 100µl of the cell resuspension was transferred onto LB/agar plates containing kanamycin (50 µg/ml). Plates were incubated at 37°C overnight.

2.6 Deletion of antibiotic resistance marker (Datsenko and Wanner, 2000)

MG1655 Δ*rpoS::cat osmY-lacZ* cells in early exponential phase (OD_{600} 0.3) were transformed with pCP20 using TSS (Chung *et al.*, 1989). Ampicillin resistant transformants were selected at 30°C. The next day, a few colonies were purified on LB/agar plates with 50 µg/ml kanamycin (without ampicillin) at 42°C. Strains were tested for loss of antibiotic marker by growth on ampicillin and chloramphenicol plates. This was confirmed by PCR using *rpoS* flanking region primers.

2.7 Cloning *rpoS* into a vector, pMarMar4

SacI and *Bsu36I* restriction sites were introduced onto *rpoS* through PCR and the product

was digested with the two restriction enzymes. The vector, pMarMar4 (Figure 5) was

also digested with *SacI* and *Bsu36I*, run on a 0.8% agarose gel for 1hr at 100V and

stained with 0.001% crystal violet for 30min. The linearized vector was isolated and

purified using the Nucleospin nucleic acid purification kit, following manufacturer's

instructions. The purified linear fragment was then treated with SAP (from Fermentas

Inc, following manufacturer's instructions) to dephosphorylate it. The SAP treated

linearized vector was mixed with the *SacI /Bsu36I* digested *rpoS* PCR product and T4

DNA ligase in a 1:3 ratio to ligate them at 16°C overnight. BW18951 (a *pir*[+] host strain)

cells in early exponential phase (OD_{600} 0.4) were washed 3X with ice-cold 10% glycerol

to make them electrocompetent, then transformed with 20ng of heat inactivated (65°C for

15 min) ligation reaction product using electroporation. Transformed cells were

incubated at 37°C in 1ml 1X LB supplemented with 20mM glucose for 1hr to allow gene

expression. Cells were centrifuged, resuspended and plated on LB/agar with 20 µg/ml

chloramphenicol. After growth overnight, transformants were purified and characterized

by isolation of plasmids and digestion with *SacI* . Insertion of *rpoS* was confirmed by

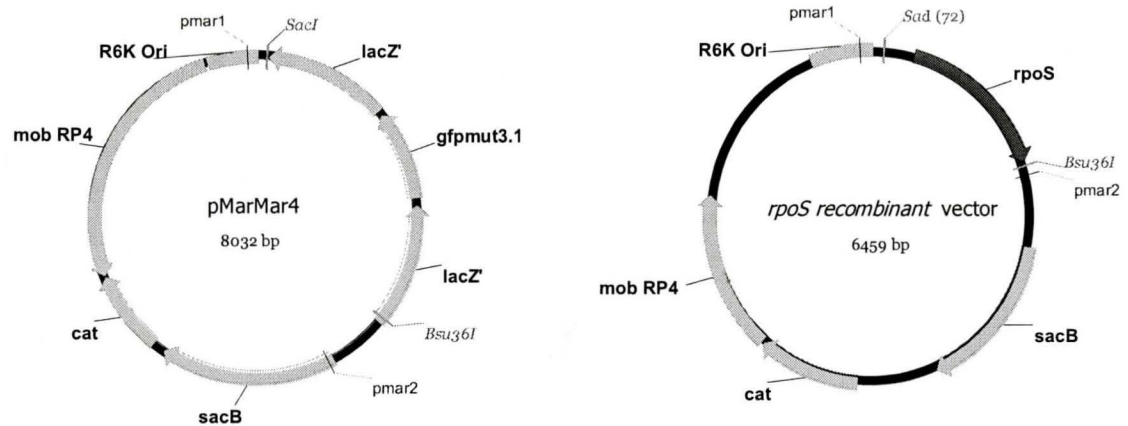PCR amplification using primers pmar1 and pmar2 (Figure 5).

Figure 5: Map of pMarMar4 (left) showing restriction sites for *SacI* and *BsuI* used in cloning *rpoS* into that region generating *rpoS* recombinant vector (right). The *rpoS* recombinant vectors were characterized by restriction digestion with *SacII* as well as PCR using primers pmar1 and pmar2.

2.8 Integration of *rpoS* recombinant vectors into chromosome of recipient strain

Early exponential phase (OD_{600} 0.4) cells of the recipient strain, MG1655 $\Delta rpoS$ *osmY-lacZ* pKD46 were washed with 10% ice cold glycerol and transformed with 100ng of the *rpoS* recombinant vector by electroporation. Cells were incubated at 37°C in 1ml 1X LB supplemented with 20mM glucose for 1hr to allow gene expression. Cells were centrifuged, resuspended and plated on LB/agar with 20μg/ml chloramphenicol. Recombinants were selected on LB/agar containing chloramphenicol (20μg/ml) plates. Recombinants were characterized by PCR using a primer upstream of the *rpoS* ORF (and the cloned region), 697rp and a reverse primer downstream *rpoS* on the vector, pmar2 (Figure 6).
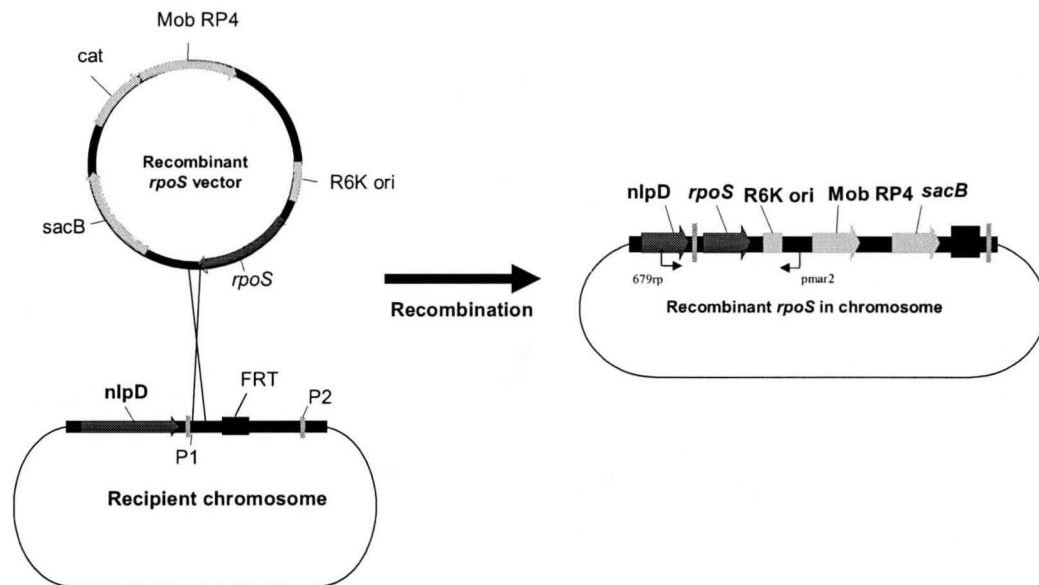
Figure 6: Integration of recombinant *rpoS* vector into chromosome of recipient strain. Homologous recombination occurs through homology regions upstream of *rpoS* ORF facilitate by Red gam proteins expressed from pKDD46.

2.9 Quantitative β- galactosidase assay of *rpoS* recombinant strains

Beta-galactosidase activity of duplicate samples from three replicates of each test and control strain was assayed as previously described (Miller, 1992). Overnight cultures of each strain were diluted 10,000X and subcultured in 1X LB to avoid carryover of stationary phase cells. Samples (1ml in duplicate) were collected at OD_{600} 0.3 (exponential phase) and OD_{600} 1.5 (stationary phase) following growth at $37^{o}C$ and shaking at 200rpm. In borosilicate tubes, 200μl of each culture was mixed by vortexing with 800μl Z-buffer, 25μl 0.1% SDS and 50μl chloroform. Addition of 200μl freshly prepared 4mg/ml ONPG (ortho-nitrophenyl-β-D-galactopyranoside) started the reaction after they were placed in a $28^{o}C$ water bath. Upon the appearance of a yellow colour (OD_{420} 0.2-0.9), the reaction was stopped by addition of 500μl of 1M Na_2CO_3 solution.

Absorbance of samples were taken at 420nm using a microtiter spectrophotometer

(Multiskan Spectrum, Thermo Scientific, Ottawa ON). Units of activity of β-

galactosidase were calculated and expressed in Miller Units as follows:

[1000 x OD_{420}] / [time (min) x volume (ml) x OD_{600}].

2.10 Preparation of cell extracts for catalase activity gels

Overnight cultures from a single colony isolate of desired strains were diluted 1000X and

grown to OD_{600} 1.5 (stationary phase) following growth at 37°C and shaking at 200rpm.

Protein synthesis was arrested with 150μg/ml chloramphenicol and cells were washed 3X

with 50mM phosphate buffer. The cells were resuspended in 0.5ml of the phosphate

buffer and sonicated until completely lysed. The cellular lysate supernatant was purified

by pelleting debris via centrifugation at 13,000 rpm for 15min. Protein concentrations

were calculated using a Bradford Assay (Bradford, 1976). Protein concentrations of the

unknown samples were determined using the equation of a standard curve generated from

bovine serum albumin (BSA, 2-10 μg/μl).

2.11 Catalase activity gel staining

HPI and HPII catalases were separated by running 5μg of each protein extract on a 10%

non-denaturing polyacrylamide gel for 120mins at 100V. Staining for catalase activity

was performed as previously described (Schellhorn and Stones, 1992). Gels were soaked

in 0.05 mg/ml horseradish peroxidase for 15min, followed by 5.0mM H_2O_2 for 10min,

then 0.5mg/ml diaminobenzidine till appearance of light brown background with clear

HPII and HPII bands. As a loading control, another gel was run in parallel and stained with Coomassie Brilliant Blue G-250.

2.12 Catalase Assay of cell extracts

Catalase activity of the cellular protein extracts were assayed as previously described (Li and Schellhorn, 2007) . A 5µl aliquot of each protein extract was mixed with 250µl of the hydrogen peroxide substrate solution (0.5ml of 30% H_2O_2 in 250 ml 50 mM phosphate buffers) in a 96 well UV- transparent microtiter plate. The plate was immediately scanned every at 240nm every 10s for 5min at 22^oC. Catalase activity was calculated using a standard curve generated from 0.05 to 1 unit of commercial catalase, following normalization to total protein extract in the cellular extract. Catalase activity was expressed as units per mg of protein.

3.0 Results

3.1 Construction of the general conjugation suicide vector

A general conjugation vector was constructed by combining genes and fragments from two plasmids, pMAGIC1 and pBSpheS H3H4 (Li and Elledge, 2005) . The plasmid pBSpheS H3H4 was partially digested with BspHI and Bsp1201 to obtain several fragments, one of which contained *ampR* and *pheS*. The other plasmid, pMAGIC1, was cut with *NcoI* and *NotI*, since these are compatible with *BspHI* and *Bsp1201* respectively. The fragment that contained *ampR* and *pheS* was ligated into the *NcoI* and *NotI* digested pMAGIC1 fragment to obtain the general conjugation vector, pDOG1. This vector (pDOG1) had two H4 homology regions between *ampR* and *pheS* with one flanked by two *SacII* sites (Figure 4). The vector was further digested with *SacII* to delete the extra H4 region. Linearization of the recombinant vector (pDOG3) confirmed its expected size of 6kb (Figure 7). *SalI* digested pMAGIC and pBSPheS H3H4 showed bands at 3.1 and 4.3kb respectively (Figure 7). After cloning, the general conjugation suicide vector was sequenced for confirmation.

3.2 Construction of the recipient strain, MG1655 Δ*rpoS osmY-lacZ*

To investigate the effect of natural *rpoS* polymorphism on gene expression within the RpoS regulon, we constructed a recipient strain to serve as an isogenic background for donor *rpoS* alleles. We placed selected natural *rpoS* alleles into the chromosome of this common recipient strain, a wildtype laboratory K-12 strain. This strain has a deletion of *rpoS* that allows recombination of donor *rpoS* alleles. A *lacZ* fusion to *osmY*, an RpoS-

dependent gene that encodes for a periplasmic protein, was used as a reporter of gene expression. We used a previously constructed strain (by Mark Kirchhof), HS2213, which has the *rpoS* ORF replaced with the CAT gene (Δ*rpoS*:: CAT) as well as a complete deletion of *lacZYA*. Strain HS2213 was transduced with P1 lysate containing *osmY-lacZ*. The *lacZ* fusion contains a kanamycin resistance gene marker, and this allowed the selection of transductants on LB media containing 50μg/ml kanamycin. Transductants streaked onto LB plates containing X-gal showed a light blue colony phenotype relative to the original pale yellow phenotype suggesting the recombination of the *osmY-lacZ* fusion into the chromosome (Figure 8). The CAT gene cassette that replaced *rpoS* was deleted to avoid interference with downstream cloning and selection on chloramphenicol. The *rpoS* ORF precise deletion was performed using the gene disruption technique by Wanner and Datsenko (2000). To eliminate the antibiotic resistance marker, CAT, a helper plasmid that expresses FLP recombinase was used (Datsenko and Wanner, 2000). FLP recombinase acts on FRT sites, which flanked the CAT gene in this case, looping it out. The strain MG1655 Δ*rpoS*::CAT *osmY-lacZ* was transformed with the temperature sensitive and ampicillin resistant FLP recombinase plasmid, pCP20. Transformants were purified nonselectively at 42°C to cure strain of the plasmid. PCR using *rpoS* flanking primers confirmed the deletion of the CAT cassette. Amplification of the *rpoS* region of the original HS2213 strain (MG1655 Δ*rpoS*::CAT Δ*lacZYA* and the transduced HS2213 strain with *osmY-lacZ* generated a 1.2kb product. However, in MG1655 Δ*rpoS osmY-lacZ*, the product was 300 bp, indicating the deletion of the CAT cassette, leaving an FRT scar (Figure 9).

3.3 Cloning *rpoS* into pMarMar4

To facilitate their transfer into the chromosome of the recipient strain, *rpoS* alleles from

ECOR-21, ECOR-28, ECOR-37, ECOR-40 and MG1655 were cloned into pMarMar4

(Figure 5), an 8kb vector with an R6K conditional origin of replication, cat which confers

chloramphenicol resistance, and *sacB* which confers sensitivity to sucrose (Goulian and

van der, 2006). The vector pMarMar4 was cut with restriction enzymes *SacI* and *Bsu36I*,

generating a 5.2 kb fragment and looping out the *lacZ-gfp-lacZ* region (Figure 5). PCR

amplified *rpoS* alleles (1.2kb) were similarly digested with both restriction enzymes and

ligated to the cut vector. BW18951, a *pir*$^+$ strain, was transformed with the ligation

products to allow replication of the recombinant plasmids. Plasmids isolated from

transformants were digested with *SacI* to linearize them, and this generated 6.4 kb

recombinant fragments (Figure 10), confirming the insertion of the *rpoS* region in the

vector. Insertion of *rpoS* was also confirmed by PCR amplification using primers pmar1

and pmar2 (see figure 5 for primer positions).

3.4 Integration of *rpoS* vector into chromosome of recipient strain

The recombinant *rpoS* vector has several features to facilitate integration into the

chromosome of the recipient strain. It has an R6K origin of replication, which allows the

plasmid to only replicate in *pir*$^+$ strains. Therefore in a non-pir recipient strain (like

MG1655), the plasmid will not be able to replicate, resulting in integration of the plasmid

into the chromosome through a homology region. The presence of a CAT gene which

conferred resistance to chloramphenicol allowed selection of primary integrants. To

facilitate recombination of the recombinant *rpoS* vector, the recipient strain MG1655

Δ*rpoS osmY-lacZ,* was transformed with pDK46, a plasmid that expresses λ Red

recombinase (Datsenko and Wanner, 2000).The non-*pir* recipient strain, MG1655 Δ*rpoS*

osmY-lacZ, containing pKD46 was transformed with the *rpoS* recombinant vector. The

rpoS vector, unable to replicate in the recipient strain, integrated into the chromosome

through the 174 bp homology region at the 5' region upstream of *rpoS* open reading

frame (Figure 6). Recombinants were selected on LB medium containing

chloramphenicol. PCR using a primer upstream of the *rpoS* ORF (and the cloned region),

697rp, and a reverse primer downstream *rpoS* on the vector, pmar2, confirmed the

integration of the whole *rpoS* vector into the chromosome of the recipient strain (Figure

11). A 1.8kb band was observed in all the *rpoS* recombinant constructs but not in

MG1655 or the no template control.

Sequence analysis of PCR fragments from integrated *rpoS* region confirmed an

MG1655 background indicated by an "AC" insertion at position -134bp upstream of *rpoS*

ORF which is not present in the ECOR strains. In addition, there was a "G" at nucleotide

position 97 (*rpoS* ORF) characteristic of the ECOR strains instead of a "C" in MG1655 as

well as the other mutations in the respective *rpoS* alleles. Strain *rpoS* [MG1655] had 4

additional mutations different from the original wildtype MG1655 strain. These include

silent mutations A→G, C→A and A →G at positions 131, 558 and 597 respectively, as

well as an A→G at position 580 which lead to a switch from serine to glycine at codon

193 (Figure 12). Strain *rpoS* [ECOR-40] also had a mutation M93V (Figure 12). These

additional mutations may have occurred either during cloning of the alleles or during

replication of DNA after integration of these alleles into the chromosome of the recipient

strain.

3.5 Phenotype of recombinant strains with natural *rpoS* alleles

The resulting recombinant *rpoS* strains had a common MG1655 background with a

transcriptional *osmY-lacZ* fusion. This allowed us to compare the effect of different *rpoS*

alleles on the RpoS regulon without other confounding genomic variations. Since *osmY* is

highly dependent on RpoS, the expression of the *osmY-lacZ* fusion relates to the status of

RpoS in the strain (Hengge-Aronis *et al.*, 1993). A qualitative β-galactosidase activity

assay was performed by streaking out recombinant strains with *rpoS* alleles from ECOR

21, 28, 37, 40 and MG1655, as well as the recipient strain (Δ *rpoS*) on LB agar

containing X-gal. The recombinant strains *rpoS* [ECOR-28] and *rpoS* [MG1655]

appeared more blue on X-gal than all the other strains (Figure 13). Strain *rpoS* [ECOR-

40] was slightly more blue than *rpoS* [ECOR-21] and *rpoS* [ECOR-37] which were

comparable to the recipient strain.

3.6 Quantitative beta galactosidase assay of *rpoS* recombinant strains

Streaking out strains on X-gal is a quick, qualitative way to test the expression level of

any *lacZ* fusion (Bremer *et al.*, 1984). For a more conclusive comparison of the

expression level of the *osmY-lacZ* fusion driven by RpoS from the recombinant strains, a

quantitative β- galactosidase assay was performed. Sampler of liquid cultures of the

recombinant strains and the recipient strain (as control) at exponential phase ($O.D_{600}=$

0.3) and stationary phase ($O.D_{600}=1.5$) were permeabilized with chloroform and SDS,

releasing the enzyme into the solution. Addition of the substrate, ONPG started the reaction and units of β-galactosidase activity were calculated in Miller Units (MU) by measuring the OD_{420} of the samples in a microtiter plate using a Multiskan spectrophotometer. Similar to the observation on the X-gal plates, in stationary phase, recombinant strains *rpoS* [MG1655] and *rpoS* [ECOR-28] had high β-galactosidase activity of 166.4 ± 14.8 and 167.8 ± 19.7 M.U., respectively (Figure 14). The other strains had a comparably low activity. The strain *rpoS* [ECOR-21] had 17.2 M.U. while *rpoS* [ECOR-40] had 17.2 M.U., similar to the *rpoS* mutant recipient strain, which had 18.5 M.U. The recombinant strain *rpoS* [ECOR-37] on the other hand, had about double the level of the low expression strains at 36.4 ±13.1 M.U. In contrast, there was no significant difference in the β-galactosidase activity of the strains in exponential phase. The highest activity was from *rpoS* [ECOR-37] which had 12.8 M.U. ± 5.4, while the least activity was from *rpoS* [ECOR-40] with 6.3 ± 3.7.

3.7 Catalase activity of ECOR strains and recombinant strains with natural *rpoS* alleles

E. coli has two catalases, hydroperoxidase I (HPI) and hydroperoxidase II (HPII) encoded by *katG* and *katE* respectively (Loewen, 1984; Loewen *et al.*, 1985). *katG* is controlled by OxyR (Christman *et al.*, 1985), while *katE* is under the control of RpoS (Mulvey *et al.*, 1990; Schellhorn and Hassan, 1988) . Thus, catalase activity corresponds to the status of RpoS in *E. coli*. Hence, catalase expression of the ECOR strains was characterized by qualitative catalase activity. Cells from these strains in stationary phase ($O.D._{600}=1.5$) were sonicated to release cellular protein extracts. The protein extracts were stained for

catalase activity sequentially with horseradish peroxidase, hydrogen peroxide and diaminobenzidine. ECOR-1, ECOR-2, ECOR-3, ECOR-4, ECOR- 7, ECOR-28, ECOR-40 and MG1655 had both HPII and HPI activity (Figure 15). However, ECOR-28, ECOR-40 and MG1655 had very low levels of HPI. ECOR 10, 21, 37 and MG1655 Δ*rpoS* on the other hand, had HPI but little HPII activity. There were also unknown dark bands on the catalase gel which migrated faster than both HPI and HPII bands. The unknown dark band from ECOR-7, ECOR-37 and ECOR-28 migrated even faster than all the others. A protein gel run as a loading control showed that the samples had similar protein concentrations (Figure 15).

The recombinant *rpoS* strains were also tested for catalase activity using protein extract from the strains. The strains *rpoS* [MG1655] and *rpoS* [ECOR-28] had both HPII and HPI activity while *rpoS* [ECOR- 21], *rpoS* [ECOR-37] and *rpoS* [ECOR-40] had HPI but no HPII activity similar to the *rpoS* null mutant (Figure 16). This suggests no RpoS activity in *rpoS* [ECOR- 21], *rpoS* [ECOR-37] and *rpoS* [ECOR-40], while *rpoS* [ECOR-28] has RpoS activity similar to *rpoS* [MG1655]. We compared the catalase activity of the recombinant strains with their corresponding natural strains by staining protein extracts from all the strains (Figure 17). Activity of the recombinant strains which had HPII bands, *rpoS* [MG1655] and *rpoS* [ECOR-28] were about 10-fold less in intensity than their wildtype strains. The protein gel run as a loading control showed that the samples had similar protein concentration (Figure 17). There may be an unknown factor affecting the expression of *katE* in the *rpoS* recombinant strains.

3.8 Catalase assay of recombinant and wildtype strains with natural *rpoS* alleles

Another indicator of the status of RpoS via catalase activity is the rate of H_2O_2 decomposition (O_2 gas development). The RpoS status of a strain can be determined by the degree of bubbling after exposure to H_2O_2. Strains with no or defect *rpoS* have a lower rate of H_2O_2 decomposition and thus reduced bubbling (Schellhorn and Hassan, 1988; Visick and Clarke, 1997). An aliquot of 30% hydrogen peroxide was dropped on overnight cultures of natural and recombinant *rpoS* strains to ascertain their RpoS status in a qualitative catalase assay (Table 4). For the wildtype strains, ECOR-28 had high catalase activity (bubbled vigorously) similar to MG1655, while ECOR-21, ECOR-37 and ECOR-40 had relatively less activity almost to the level of the *rpoS* null mutant (Table 4). When the *rpoS* alleles were put into the same background, *rpoS* [ECOR-28] like *rpoS* [MG1655] maintained the high catalase activity while *rpoS* [ECOR- 21], *rpoS* [ECOR-37] and *rpoS* [ECOR-40] had catalase levels comparable to the *rpoS* mutant (Table 4).

We also characterized the total catalase activity of the *rpoS* recombinant strains as well as their corresponding wildtype using a quantitative catalase assay (Figure 18). The total catalase activity ranged between 0.094 Units/µg for ECOR-21 to 0.036 Units/µg for MG1655. There was no correlation or pattern in total catalase activity between the strains since the activity was variable (Figure 18).

4.0 Figures and tables for results

Table 4: Catalase assay of recombinant and wildtype strains with different *rpoS* alleles

Strain	Genotype	Catalase activity
ECOR-21	O121:HN	++
ECOR-28	O104:NM	++++
ECOR-37	ON:HN	++
ECOR-40	O7: NM	++
MG1655	F- λ- rph	++++
HS2719	*rpoS* [ECOR-21]	++
HS2720	*rpoS* [ECOR-28]	++++
HS2721	*rpoS* [ECOR-37]	++
HS2722	*rpoS* [ECOR-40]	++
HS2723	*rpoS* [MG1655]	++++
HS2718	MG1655Δ*rpoS*	+

Catalase activity was determined by rate of evolution of O_2 after dropping 5µl of H_2O_2 onto individual colonies.
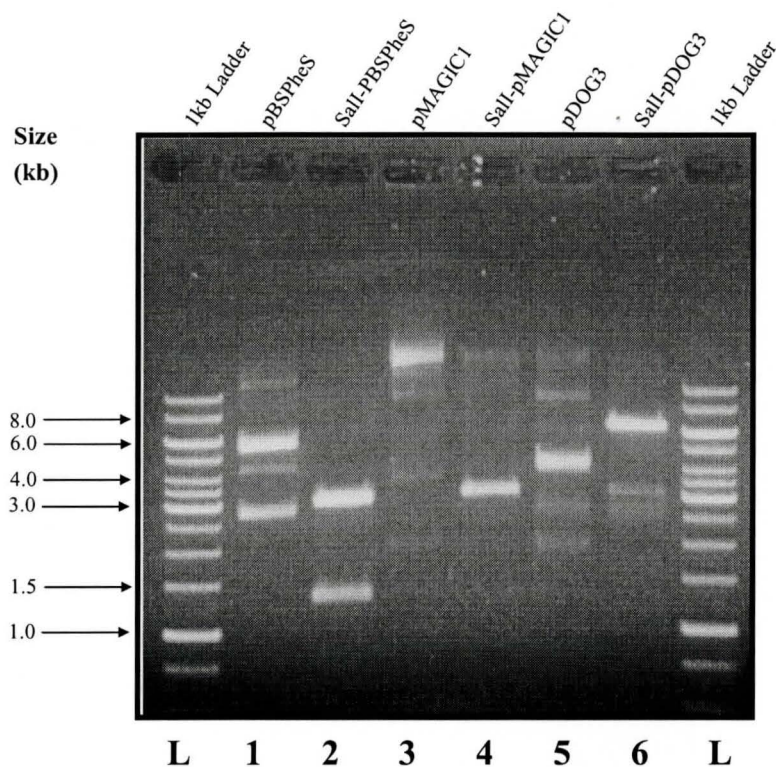
Figure 7: Characterization of pDOG3 by restriction digestion with *SalI*. pBSPheS contains two *SalI* sites while pDOG3 and pMAGIC1 each has one. Each plasmid is followed by its *SalI* digested sample. Samples (150ng) were separated on a 0.8% agarose gel and visualized by staining with ethidium bromide.
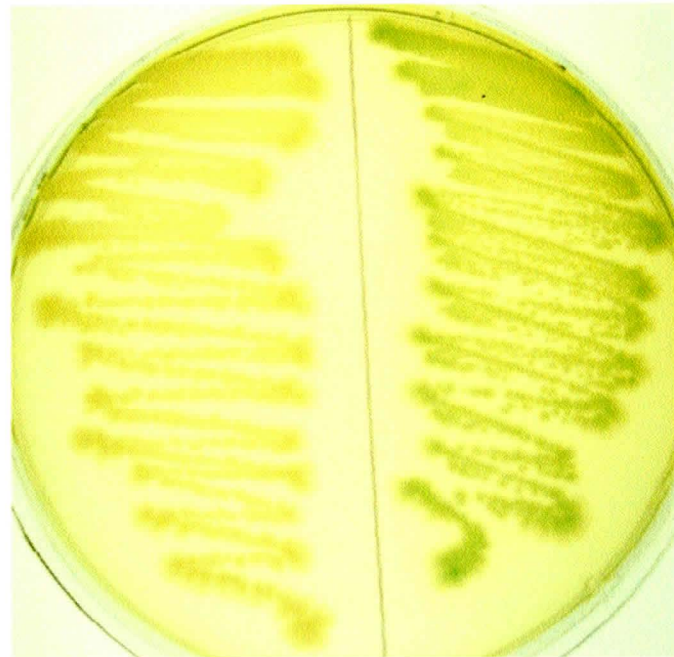
MG1655 Δ*rpoS* Δ*lacZYA*   **MG1655 Δ*rpoS* *osmY-lacZ***
(HS2213)   **(HS2703)**

Figure 8: Confirmation of insertion of *osmY-lacZ* into MG1655 Δ*rpoS::cat* Δ*lacZYA* (HS2213). The recipient strain was transduced with P1 lysate containing *osmY-lacZ*. A single colony of each strain was streaked on an LB agar plate containing 50μg X-gal.
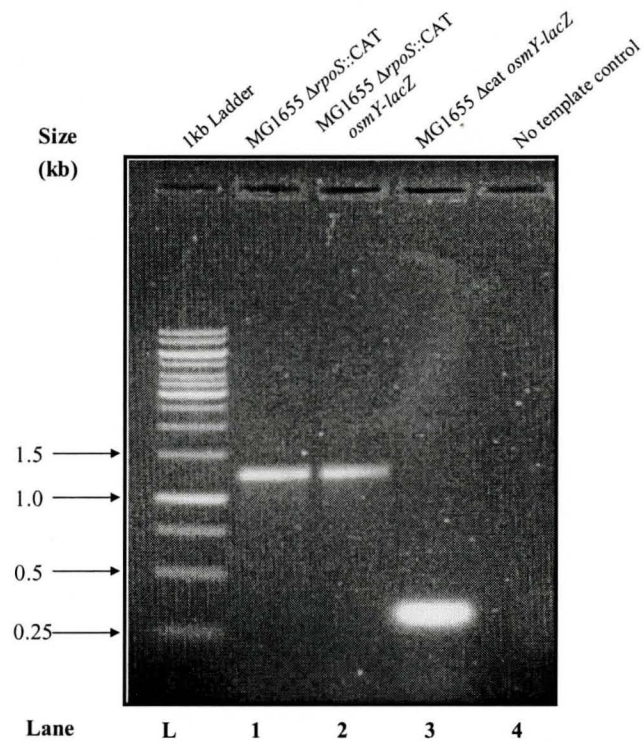
Figure 9: Confirmation of deletion of the resistance marker, CAT, from the genome of the recipient strain by PCR amplification. Primers fw2rp (126 bp upstream *rpoS* ORF) and rv2rp (61 downstream *rpoS* ORF) were used in PCR amplification. PCR products (100ng) were separated on a 0.8% agarose gel and visualized by staining with Ethidium bromide.
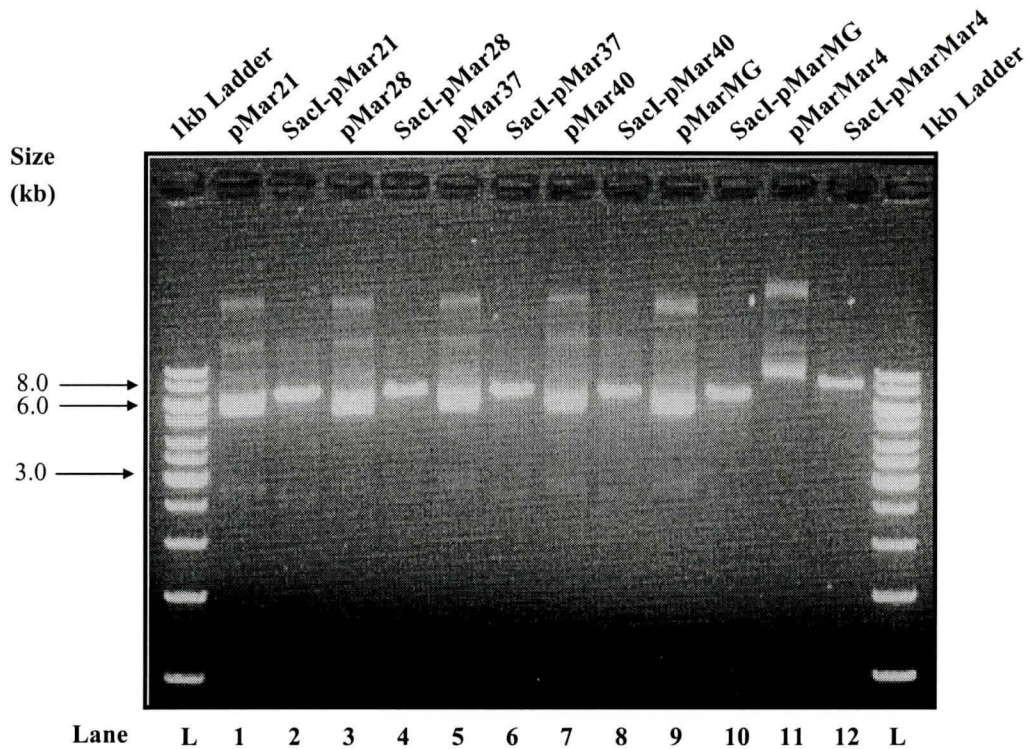
Figure 10: Characterization of *rpoS* recombinant vectors by restriction digestion with *SacI*. All the vectors have one *SacI* site. Samples (200ng) were separated on a 0.8% agarose gel and visualized by staining with ethidium bromide. Each vector is followed by its *SacI* digested sample. Vector names are designated "pmar" and the name of the ECOR strain from which the *rpoS* allele was obtained. For example "pmar21" contains *rpoS* from ECOR-21.
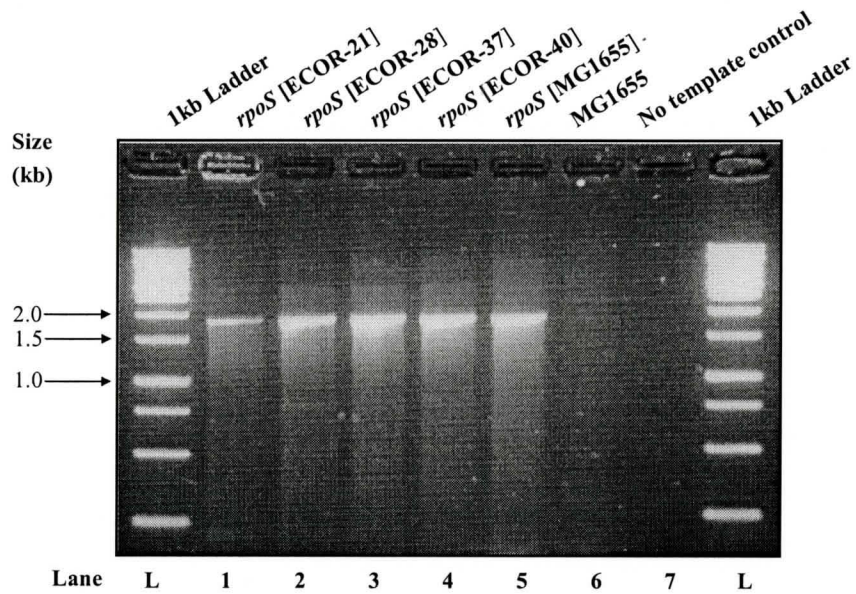
Figure 11: Confirmation of integration of *rpoS* vectors into chromosome of recipient strain by PCR amplification. Primers 679rp, 679bp upstream *rpoS* ORF, and pmar2, downstream R6K ori on the vector, were used (see Figure 5). PCR products (200ng) were separated on a 0.8% agarose gel and visualized by staining with ethidium bromide.

Function		**Inhibition**		**Core binding**		**DNA melting** **-10 Recognition**		**Core binding**		**-35 Recognition**	
Conserved Region		1.1	1.2	2.1	2.2	2.3	2.4	3.1	3.2	4.1	4.2

1 · · · · · · · · · · · · · · · · · · · 261 262 · · · · · · · · · · · · · · · · 514 · · · · · · · · · 750 751 · · · · · · · · · · 990 nt

	MG1655	ATG..CAG.G....T.........C..A...C..C...A..ACA.T...C..A..T..G..G...-..C..C..G...C..G...
rpoS allele	ECOR21	ATG..GAG.A....T.........C..A...C..C...A..ACA.T...C..C..T..A..A...A..C..C..G...T..G...
	ECOR28	ATG..GAG.A....T.........C..A...C..C...A..ACA.T...C..C..T..A..A...-..C..C..G...C..G...
	ECOR37	ATG..GAG.A....T.........C..A...C..T...A..ATA.T...C..C..T..A..A...-..C..C..G...T..G...
	ECOR40	ATG..GAG.A....C.........T..G...T..C...A..ACA.C...C..C..C..A..A...-..T..C..A...C..G...

rpoS donor	Type of mutation	Mutation and position	Effect on amino acid sequence
	Transversion	G97C[1]	Glutamic acid →Glutamine
	Transition	A131G[2]	None
MG1655	Transversion	C558A[2]	None
	Transversion	A580G[2]	None
	Transition	A597G[2]	Serine→Glycine
ECOR-21	Insertion	-729A	Premature stop codon at 756 nt
	Transition	C944T	None
	Transition	C387T	None
ECOR-37	Transition	C431T	Threonine→Isoleucine
	Transition	C944T	None
	Transition	T163C	None
	Transition	C264T	None
	Transition	A279G[2]	Methionine→Valine
ECOR-40	Transition	C354T	None
	Transition	T462C	None
	Transition	T574C	None
	Transition	C733T	None
	Transition	G821A	None

1. This mutation has also been reported by Atlung *et. al.* (2002)
2. Additional mutation not found in original strains

Figure 12: Mutations in *rpoS* alleles cloned into MG1655 recipient strain background. Top figure shows location of mutations in the functional regions of *rpoS*, and bottom chart shows the type of mutations and their effect on amino acid sequence.
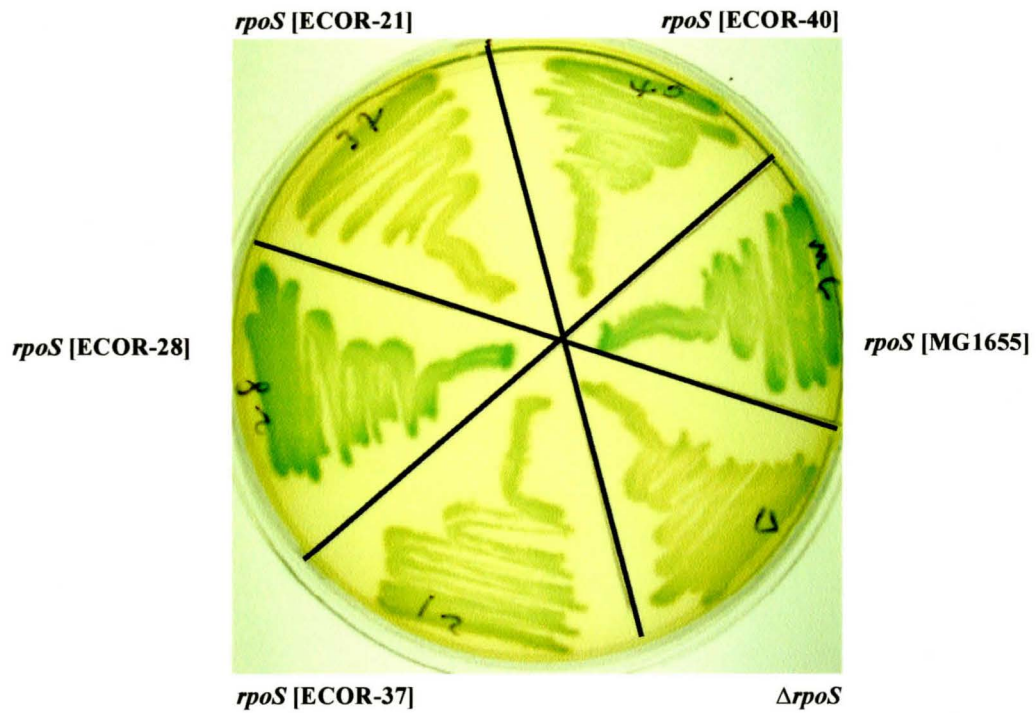
Figure 13: Phenotype of recombinant strains with natural *rpoS* alleles showing the expression of an *osmY-lacZ* fusion on X-gal. A single colony of each strain was streaked onto LB agar plates containing 50 µg/ml X-gal.
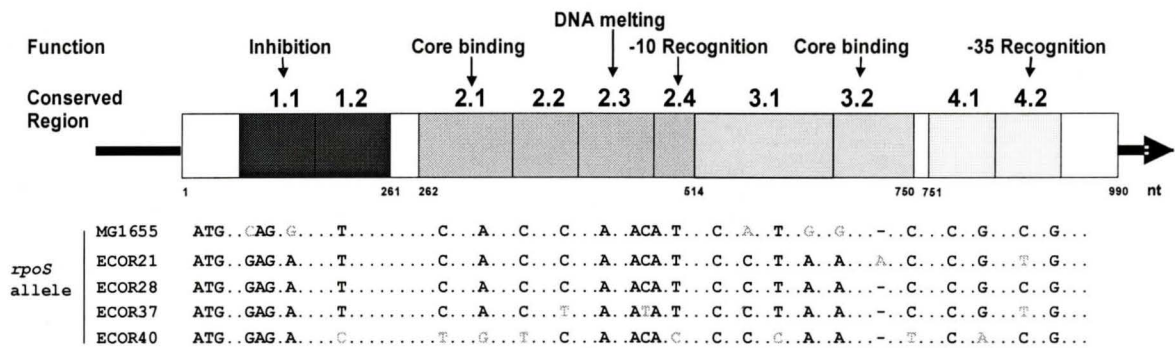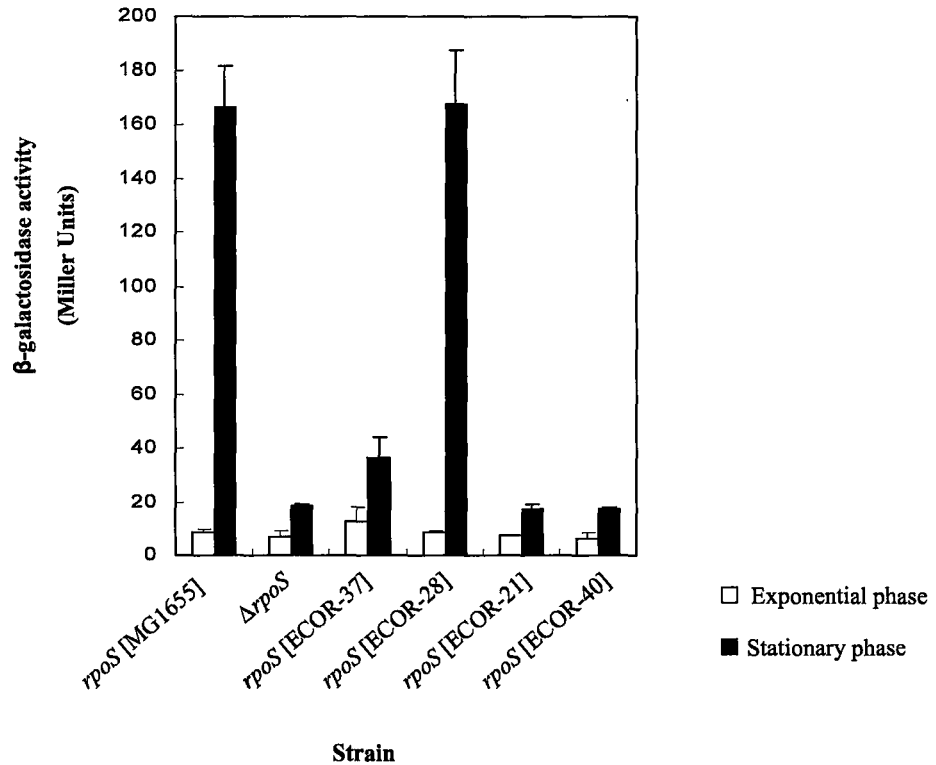
Figure 14: RpoS-dependent regulation of the *osmY-lacZ* fusions measured by quantitative β-galactosidase activity assay (Miller units) in recombinant strains with natural *rpoS* alleles. Liquid cultures of each strain at exponential phase (O.D$_{600}$= 0.3) and stationary phase (O.D$_{600}$=1.5) were permeabilized and assayed for β-galactosidase activity. Data shown are averages of duplicate samples from three independent replicates of each strain. Error bars indicate standard error.
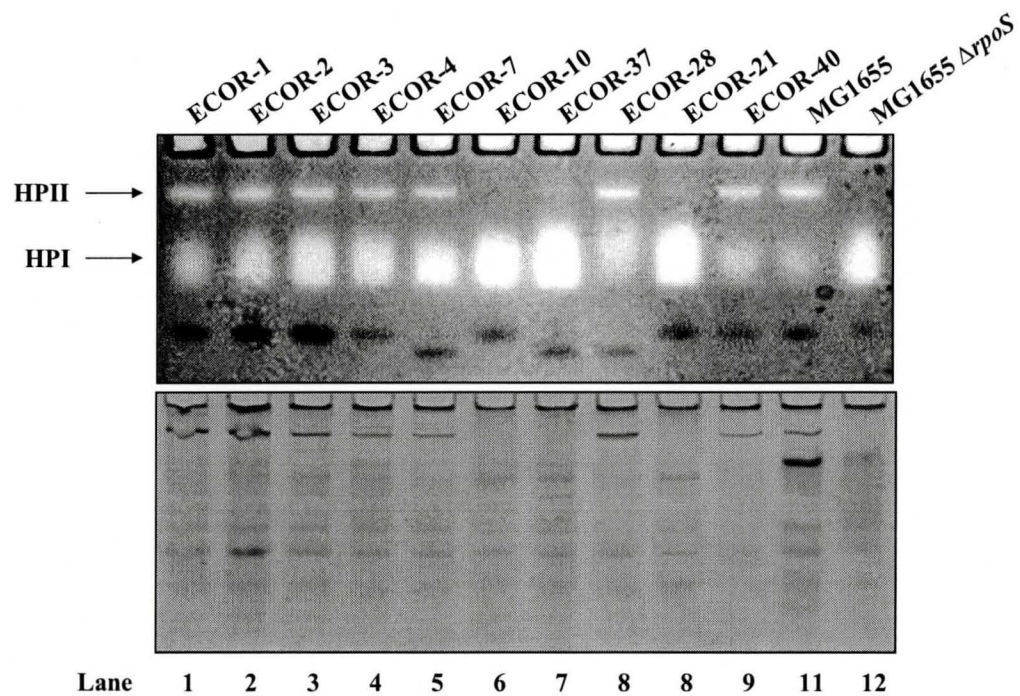
Figure 15: Catalase activity of ECOR strains from stationary phase (OD_{600} 1.5) cultures. Protein extracts (5μg) were run on a 10% non-denaturing polyacrylamide gel for 120min at 100V and stained with 0.05 mg/ml horseradish peroxidase, 5.0mM H_2O_2 and 0.5mg/ml diaminobenzidine. Protein gel (10μg of protein extract) below was run in parallel and stained with Coomasie Brilliant Blue G-250 as a loading control.

Figure 16: Catalase activity of recombinant natural *rpoS* strains from 10μg of protein extract. Protein extract from stationary phase cultures (OD$_{600}$1.5) were run on a 10% non-denaturing polyacrylamide gel for 120min at 100V and stained with 0.05 mg/ml horseradish peroxidase, 5.0mM H$_2$O$_2$ and 0.5mg/ml diaminobenzidine. For a loading control, a similar gel was run in parallel and stained with Coomassie Brilliant Blue G-250 (protein gel below).
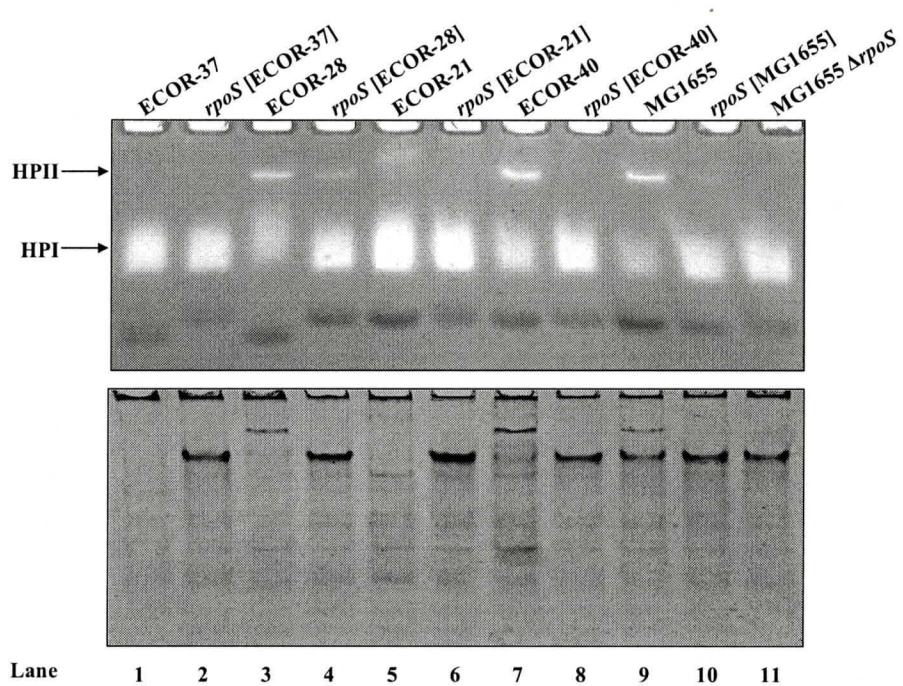
Figure 17: Catalase activity of recombinant natural *rpoS* constructs in comparison to their corresponding natural strains. Protein extracts (5μg) from stationary phase (OD$_{600}$1.5) cultures were run on a 10% non-denaturing polyacrylamide gel for 120min at 100V and stained with 0.05 mg/ml horseradish peroxidase, 5.0mM H$_2$O$_2$ and 0.5mg/ml diaminobenzidine. The figure below is a protein gel (10μg of protein extract) which was run in parallel and stained with Coomassie Brilliant Blue G-250 as a loading control.

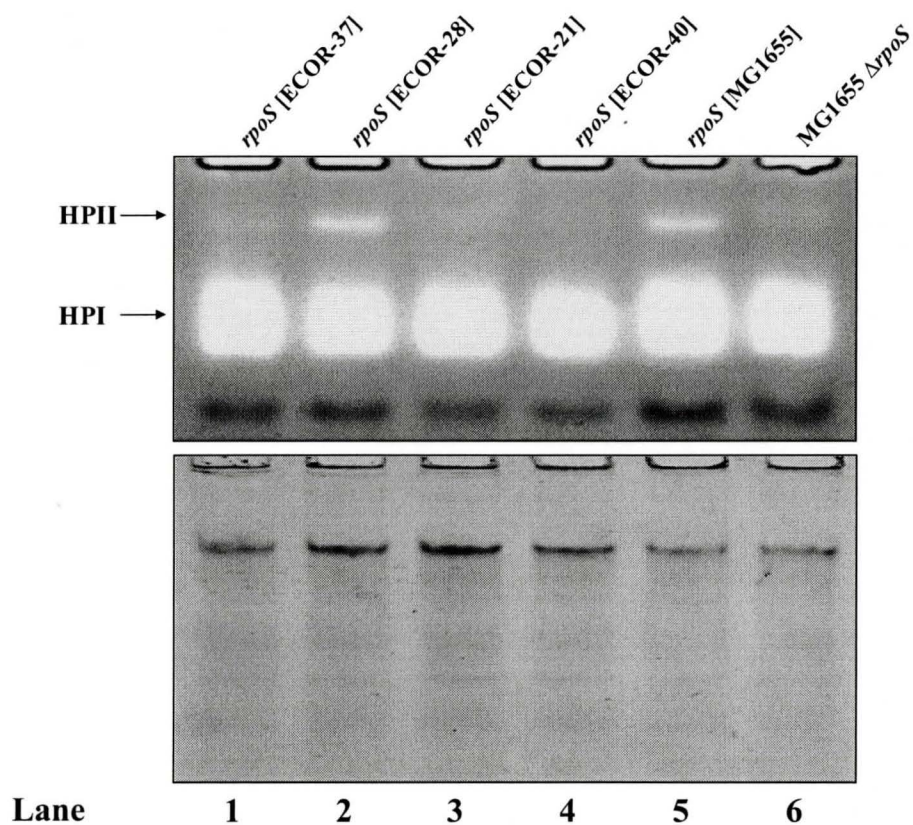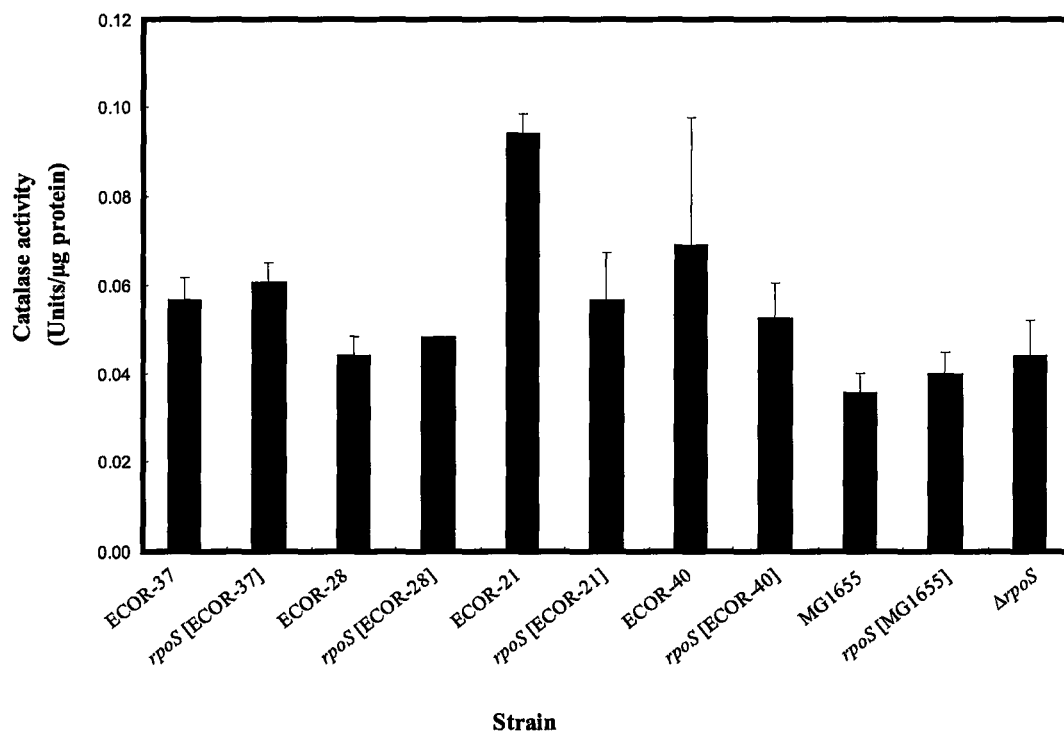**Figure 18**: Quantitative catalase assay of recombinant natural *rpoS* constructs in comparison to their corresponding natural strains. Protein extracts (5µg) from stationary phase (OD_{600}1.5) cultures were mixed with 250µl of hydrogen peroxide substrate solution in a 96 well UV- transparent microtiter plate. The plate was scanned every at 240nm every 10s for 5min at 22^0C. Each bar is an average of 4 replicates from each strain. Error bars are standard error.

5.0 Discussion

5.1 The *in vivo* cloning system

We have constructed a general conjugation suicide vector (Figure 4) which will facilitate cloning any gene of interest *in vivo*. The *in vivo* cloning method relies on indigenous conjugal mating between bacteria with the objective of easily and effectively transferring genes into a recipient strain of choice. The gene of interest is amplified by PCR using primers that will introduce H3 and H4 homology regions at both ends of the fragment. The host strain of the conjugation vector (donor strain) is transformed with the PCR product of the gene of interest, and with the help of a plasmid, pML300 which possesses genes for lambda Redα, Redβ and Gam under a *rhaB* promoter (Li and Elledge, 2005; Zhang *et al.*, 2002), the gene of interest will recombine into the general conjugation vector (Figure 19). The recombination event will replace (and thus eliminating) *ampR* and *pheS* on the conjugation vector with the gene of interest. The *pheS* gene encodes for the tRNA synthetase for phenylalanine. Its allele, *pheS* Gly294 is toxic to cells in the presence of chlorophenylalanine, Cl-Phe, since the allele charges Phe-tRNAs with this compound (Li and Elledge, 2005). This (*pheS* Gly294) will thus be used as a counter selection marker to minimize background and increase the number of transformants that have actually undergone homologous recombination to acquire the gene of interest. Transformants with inserts will also be sensitive to ampicillin, which is another testable characteristic of the recombinants. The general conjugation vector with the recombined gene of interest will then be transferred to a non *-pir* recipient strain by conjugation and the gene of interest will recombine into the chromosome of the recipient strain through

homology regions at its 5' and 3' ends. The vector contains an origin of transfer (*oriT*) obtained from a neighboring plasmid with an F factor system which possesses a trans-acting transfer operon (F') that will facilitate transfer of the conjugation vector but not the F'plasmid itself (Li and Elledge, 2005). The conjugation vector also has a conditional origin of replication from R6K, oriγ, and so will only initiate replication in the presence of π, a trans-acting factor coded by *pir1* or its allele *pir1-116* (Metcalf *et al.*, 1994). Therefore the general conjugation vector will only replicate in the donor strain which expresses *pir1-166* but not the non-*pir* recipient strain. Selection of recombinants at this stage will depend on the gene being cloned. For example for *rpoS*, recombinants will have active RpoS which will drive the expression of a *lacZ* fusion to an RpoS-dependent gene producing β-galactosidase and hence the ability to utilize lactose. Thus, recombinants can be selected on lactose minimal media containing X-gal. This *in vivo* cloning system will be a useful tool for testing the expression of *rpoS* alleles, introduce a library of genes into *rpoS* fusions to identify potential transacting regulators and to place genes into overexpressing plasmids or under controllable promoter. Similarly, it could be used for expression studies of any gene of interest.
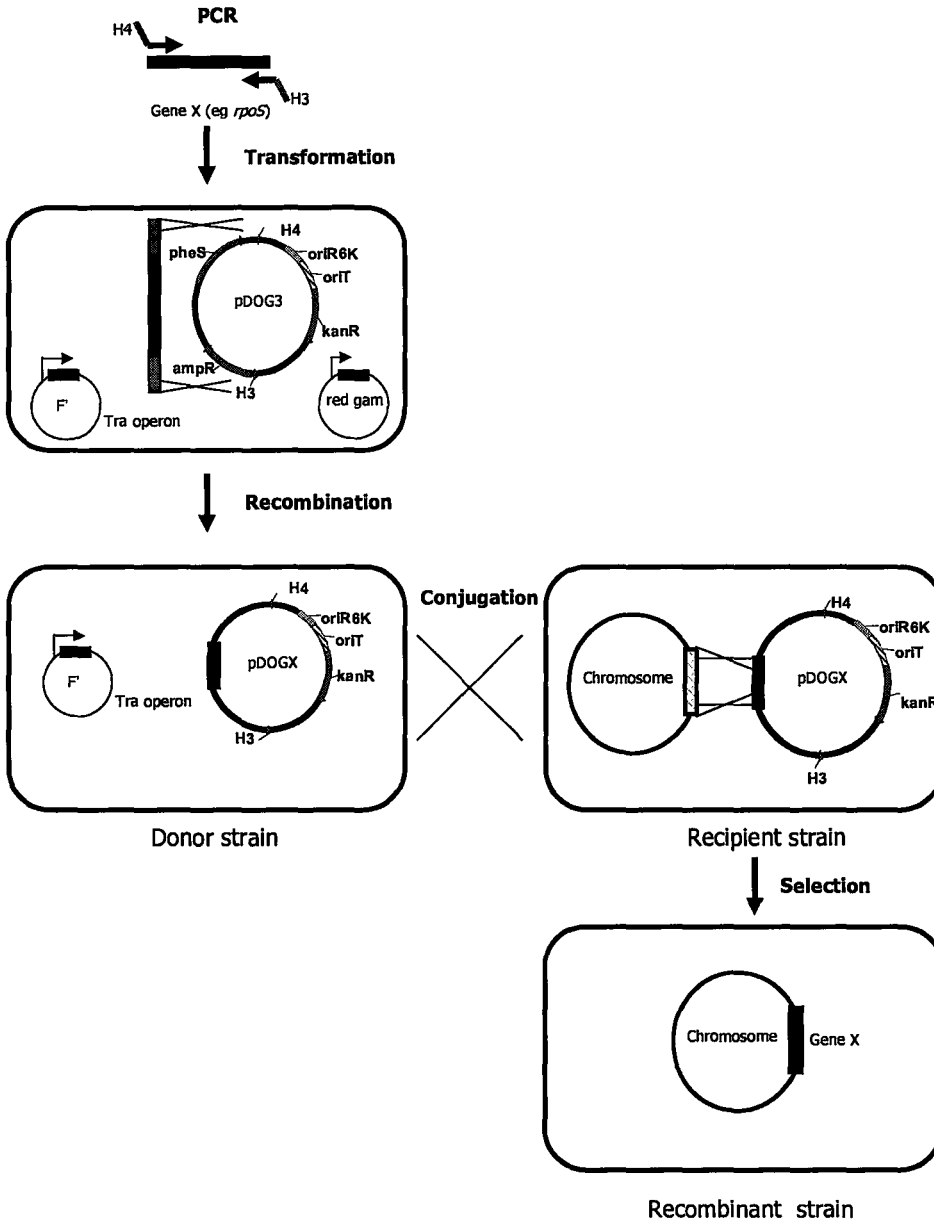
Figure 19: Scheme of the *in vivo* cloning system. A gene X (eg *rpoS*) is amplified by PCR using H4 and H3 flanking primers. The donor host strain is transformed with the PCR product which then recombines into the general conjugation vector through the H3 and H4 sites followed by counter selection against *pheS* marker on Cl-Phe. The general conjugation vector carrying gene X is conjugated into a recipient strain where gene X recombines into the chromosome and recombinants are selected using an appropriate marker.

5.2 Construction of *rpoS* recombinant strains

To study the effect of the *rpoS* polymorphism on downstream processes, *rpoS* alleles

from ECOR 21, 28, 37 and 40 were cloned into the same background, MG1655

Δ*rpoS*:cat *osmY-lacZ*. The *rpoS* allele from MG1655, a K12 wildtype was also cloned

for comparison. This prevented all confounding variables and allowed the attribution of

any phenotypic differences to polymorphism of the *rpoS* ORF. Alleles of *rpoS* from this

group of ECOR strains were selected because of the diversity in the degree and type of

mutations. The allele from ECOR-21 has an insertion at nucleotide 728 causing a

truncation of the amino acid at 252 instead of the complete 330aa (Figure 3). ECOR-37

has two silent mutations and one that leads to a change in amino acid while ECOR-40 has

7 silent mutations. The *rpoS* allele from ECOR-28 on the other hand had no additional

mutations beside the C33G that all the natural strains possess.

   The features and use of the cloning vector, pMarMar4 provided a robust system for

transferring *rpoS* alleles into another genetic background. Its R6K origin of replication

which allows the vector to replicate exclusively in *pir*⁺ host strains, forced the

recombinant vector to integrate into the chromosome of the non-pir recipient strain,

MG1655. The presence of a CAT gene allowed selection of recombinant recipient strains

with primary integration (by homologous recombination facilitated by Red gam proteins

expressed from a plasmid, pKD46) of the whole vector on media containing

chloramphenicol. Since *rpoS* was the only gene of interest for recombination, attempts

were made to eliminate the rest of the plasmid. This was also to prevent any potential

effects of the other parts of the vector on further studies on the strains. An additional

feature of the cloning vector is the *sacB* gene which confers sensitivity to sucrose in many bacteria. A secondary recombination event was initiated by selecting for recombinant strains which were resistant to sucrose due to loss of *sacB*. However, all the recombinants tested were revertants that had lost the whole plasmid including the *rpoS* region (data not shown). A secondary recombination of primary integrants should have two possible outcomes. One is reversion to the original strain genotype and the second is loss of the rest of the plasmid leaving *rpoS in the* chromosome. The relative percentage of colonies of each outcome depends on how favorable each recombinant event is, which in turn is dependent on the distance between the two homology regions and the length of homology. One limitation of the later recombination outcome is that, the homology region for the 3' ends of *rpoS* (one 3' end on plasmid and the other homology region on the chromosome) to recombine out the rest of the plasmid was relatively short, 23 bp, and thus less favorable relative to the 5' ends recombining out the whole plasmid (more than 100 bp homology region). The reason why such short homology region was used is that the sequence following the 3' end of *rpoS* ORF from the ECOR strains is highly polymorphic and thus it was difficult to design consensus primers for all the strains. The primer that was used was just 23 bp immediately following the ORF. So, in the secondary recombination, only 23 bp were available as a homology region. A primer that was 23 bp downstream the ORF (thus 46 bp region of homology) could not amplify all the strains (data not shown).

In addition to the natural mutations in *rpoS* in the ECOR strains, sequence analysis also showed additional mutations in some of the constructed *rpoS* recombinant strains

(Figure 12). The strain *rpoS* [ECOR-40] had a mutation M93V while *rpoS* [MG1655]

had 4 mutations different from the original wildtype MG1655 strain. These include silent

mutations A→G, C→A and A →G at positions 131, 558 and 597 respectively, as well as

a change in amino acid mutation, S193G. These additional mutations could be a result of

several events. One is DNA damage which might have occurred from exposure to

chemicals and the many enzymatic procedures and during cloning. Other sources include

errors during PCR amplification of the alleles or DNA replication after cloning.

5.3 Polymorphism of *rpoS* in natural strains of *E. coli*

We amplified and sequenced 4 natural *rpoS* alleles from the ECOR collection of natural

E. coli isolates. These sequences confirmed previous data from our lab, figure 3 (Dong

and Schellhorn, unpublished data). Sequence analysis showed a G at position 97 in all the

natural strains instead of a C as in laboratory wildtype strains like MG1655, confirming

previous reports of GAG as codon 33 in non-K12 strains (Atlung *et al.*, 2002). This

codon is quite variable in many *E. coli* strains. Codon 33 variants include CAG, CTG,

TAT, TCG and TAG in various K12 strains, while non-K12 strains possess GAG at this

position (Atlung *et al.*, 2002; Ferreira *et al.*, 1999). There could be a possible

evolutionary connection between strains with those codons and an explanation for the

highly polymorphic nature of that segment of the *rpoS* gene. It has been suggested that

GAG is the ancestral codon which mutated to the amber codon TAG in the pioneering

K12 strains, followed by various mutations to restore RpoS function for cell survival in

stationary phase (Atlung *et al.*, 2002). Between the isolation of *E. coli* strain from a

diphtheria patient in 1922 to the original K12 strain by E.L. Tatum, the *rpoS* gene

acquired mutations (Belin, 2003).

The *rpoS* allele in ECOR-21 haboured an "A" insertion at nucleotide 728 causing a

truncation of the amino acid sequence to 252 aa (Figures 12). The K12 strains with

amber mutations at codon 33 also synthesize a truncated protein, with 278 aa

(Subbarayan and Sarkar, 2004a). That is, RpoS from ECOR-21 has a deletion of 78

amino acids on its C-terminus while strains with amber codon 33 lack the first 53 amino

acids in the N-terminal. Natural *rpoS* alleles with amber mutations leading to partial

amino acid sequences have been reported. The *rpoS* allele of the prototype O157:H7

EHEC EDL933[a] has recently been found to possess a point mutation causing a premature

stop codon (Coldewey *et al.*, 2007), a finding that conflicts previously published database

sequence (Perna *et al.*, 2001). Likewise, the shiga-like toxin producing *E. coli* (SLTEC)

strain PS2 serotype O157:H7 has a single base pair deletion at nucleotide 48 in its ORF

causing an amber codon 29 (Waterman and Small, 1996).

The other mutations in *rpoS* in the ECOR strains are all silent mutations except a C →T

at position 432 which changes codon 144 from Threonine to Isoleuncine in ECOR-37

(Figure 12). It should be noted that all the mutations are transitions (i.e. C↔T or A↔G).

The mutations are not concentrated in one functional region, but rather spread out across

the entire gene. Contrary to previous reports, codon 29 and 196 in the ECOR strains were

CCG (Pro) and GAA (Val) respectively as in K12, instead of of GCC (Val) and (Glu)

respectively, in other natural strains (Ferreira *et al.*, 1999). However, position 387 in

ECOR-37 has a C →T mutation as in most O157:H7, O55:H7 and O157: NM strains ,

while the other ECOR strains tested in this study, K12, EDL933, O111:H1, and O26:H11

possess a C at that position (Bhagwat et $al.$, 2006; Coldewey et $al.$, 2007; Ferreira et $al.$,

1999; Robey et $al.$, 2001). Similarly, position 819 in both ECOR-37 and ECOR-40 has a

G →A mutation as in O157:H7, O55:H7 and O124 strains but a G in all the ECOR

strains from this study, K12, O26:11, and O111:H11 strains (Bhagwat et $al.$, 2006;

Ferreira et $al.$, 1999; Robey et $al.$, 2001).

Laboratory $E.$ $coli$ and $Salmonella$ populations acquire mutators at frequencies of 10^{-5}

to 10^{-6} (Boe et $al.$, 2000; LeClerc et $al.$, 1998; Mao et $al.$, 1997). Natural $E.$ $coli$ and

$Salmonella$ populations in the environment have even higher frequencies of 1-5%,

increasing mutation rate (LeClerc et $al.$, 1998; Matic et $al.$, 1997). If the mutation is

beneficial, it can contribute immensely to evolutionary change (Hegreness et $al.$, 2006).

Polymorphisms in $rpoS$ are common since it is located at a highly mutable site of the $E.$

$coli$ genome (Atlung et $al.$, 2002; Ferenci, 2003). The high polymorphic nature of $rpoS$

can also be explained by a trade off between nutrition and stress resistance. In this

hypothesis, $rpoS$ mutations decrease the expression of RpoS, easing sigma competition

and thus increasing the expression of RpoD-dependent genes for enhanced nutrient

uptake (Ferenci, 2003; Jishage and Ishihama, 1999). $rpoS$ mutants have growth advantage

over wildtype cells after a long incubation time, allowing the mutants to take over the

population in stationary while RpoS mutants exhibit better growth on succinate because

of enhanced metabolism of TCA cycle intermediates (Chen et $al.$, 2004; Zambrano and

Kolter, 1996). Loss of RpoS function is evolutionarily advantageous because it allows

cells to "turn off" expression of the regulon when it is not needed and switches it on

again when necessary (Chen *et al.*, 2004). What still remain unclear are exactly how these different alleles are selected for and the relationship between them.

5.4 Effect of *rpoS* polymorphism on gene expression in the RpoS regulon

In spite of the ubiquity of *rpoS* mutations, only a few studies have examined the effect of *rpoS* polymorphisms on gene expression. Most of these studies use and provide information obtained from laboratory or K12 strains (Notley-McRobb *et al.*, 2002; Subbarayan and Sarkar, 2004a; Visick and Clarke, 1997). In such studies, whether in K12 or natural strains, *rpoS* alleles are maintained in their natural genetic background, introducing several confounding variables for their experiments. A few studies have done complementation studies where *rpoS* alleles are placed on a vector but this does not give physiological amounts of RpoS (Atlung *et al.*, 2002; Rajkumari and Gowrishankar, 2002). The unique feature of this study is that, we placed selected natural *rpoS* alleles into the chromosome of a common background, a laboratory wildtype K12 strain with an RpoS-dependent reporter, *osmY-lacZ* which allows (i) testing the regulatory activities of RpoS by measuring expression levels of reporter genes and (ii) elimination of any confounding variable, permitting the direct attribution of any effects exclusively on differences in *rpoS* ORF. We were therefore able to robustly investigate the effects of *rpoS* polymorphism on gene expression within the RpoS regulon.

We have shown that there is variation in the regulatory activities of RpoS due to mutations in natural *rpoS* alleles, using two independent gene expression analyses, β-galactosidase and catalase activity assays. The recombinant strain with *rpoS* from ECOR-28 had increased expression qualitatively and quantitatively of β-galactosidase from

osmY-lacZ similar to the wildtype K12 strain, MG1655 (Figures 13 and 14). The ECOR-37 *rpoS* allele on the other hand, showed low β-galactosidase activity but not to the level of the *rpoS* null mutant. In contrast, the level of β-galactosidase activities from ECOR-21 and 40 *rpoS* alleles were as low as the *rpoS* null mutant. In an independent test, recombinant strains carrying *rpoS* alleles from ECOR-21, ECOR-37 and ECOR-40 showed little expression of *katE*, HPII (Figure 16). However, the strain with *rpoS* from ECOR-28 had an HPII signal similar to the wildtype K12 allele. This was confirmed by a catalase bubble test (Table 4).Taken in toto, these results suggest that RpoS in ECOR-28 has high activity which is similar to wildtype K12 strain, while ECOR-37 has intermediate RpoS activity and ECOR-21 and 40 have little RpoS activity. Since the only differences between the strains are several nucleotide changes in *rpoS* ORF, we can conclude that these variations are effects of *rpoS* mutations in the strains. Allelic variation in *rpoS* resulting in differences in RpoS phenotype is not uncommon. Many natural isolates of *E. coli* have lost or partially lost RpoS function (Ferenci, 2003). Catalase activity of two pathogenic K1 strains with nonsense mutations were 15 fold less than a K12 strain, while that of another K1 strain with a C-to-G change at codon 33 had a two fold increase than the wildtype K12 strain (Wang and Kim, 2000). Similarly, the RpoS phenotype of six non-K12 strains with only silent mutation differences in *rpoS*, varied significantly more than that of K12 strains with respect to catalase, acid phosphatase and glycogen accumulation levels (Atlung *et al.*, 2002). In a study examining the abilities of 58 Shiga-like toxin producing *E. coli* (SLTEC) to survive at pH 2.5, 13 of them were subnormal in acid resistance (Waterman and Small, 1996).

Sequence analysis of some of those strains showed mutations in *rpoS* and complementation with *rpoS* on a plasmid restored acid resistance function (Waterman and Small, 1996). In contrast, Coldewey *et. al.* (2007) reported that RpoS activity is not always dependent on the *rpoS* genotype, after finding a fully functional *rpoS* gene was defective in acid resistance. Such a conclusion is questionable since each *rpoS* allele in that study was in a different background introducing many confounding variables.

While our results suggest variation in RpoS activity, we do not know the exact mechanism of this observation. Some mutations in *rpoS* can lead to changes in encoded amino acids which lead to changes in RpoS primary structure and possible changes in overall structure affecting its stability and functionality. This will explain variation in RpoS phenotype of *rpoS* alleles with mutations resulting in amino changes. Silent mutations on the other hand do not change amino acid sequences although codons are different and thus primary structure of the protein is maintained. Nevertheless, it has been shown in several species that there is a relationship between gene expression levels and codon usage, and codon usage in turn correlates with the amount of corresponding tRNA in a cell (Ikemura, 1985). To put this in context, O157:H7 contains 100 tRNAs while the laboratory K12 strain, MG1655 has 88 (Rocha, 2004). *E. coli* cells make optimal use of their tRNA resources to meet the codon demands of frequently expressed genes which encode almost 90% of proteins in the cell (Solomovici *et al.*, 1997). Thus, a favoured or frequent codon will be translated faster than a rare one because its cognate tRNA is more abundant (Sorensen and Pedersen, 1991). The rate of mRNA translation by ribosomes can be controlled by codon usage, since a collection of frequent codons will

result in relatively quicker translation and vice versa (Deana *et al.*, 1996; Deana *et al.*, 1998) . The rate of translation can in turn affect protein folding. Silent mutations of EgFABP1 (*Echinococcus granulosus* fatty acid binding protein1) significantly decreased the solubility of the protein when expressed in *E. coli* and caused a strong activation of a reporter gene designed to detect misfolded proteins (Cortazzo *et al.*, 2002). Silent nucleotide changes can also affect mRNA levels and half-life. Using Northern blot analysis, greater than a 10-fold reduction in the steady-state amounts and more than a 4-fold reduction in the half-life of silently mutated mRNA of *ompA* was observed (Deana *et al.*, 1998). If the functionality of RpoS is influenced, its specificity for certain promoters will be altered causing reduction or enhancement in its regulatory activities. The level at which mutations affects RpoS activity (in this study) is not known, but this could have been resolved by real time PCR to quantify the amount of mRNA transcripts available in the strains for translation, and a Western blot analysis to show the amount of protein translated. Variation in mRNA level would indicate a defect in transcription of alleles with lower transcripts, while differences in protein concentrations would imply that a similar amount of mRNA transcripts are available but the rate of translation of some of the alleles are low. Alternatively, there could be similar amounts of mRNA transcripts and protein but the protein could be misfolded in some of the strains.

The location of a mutation in the *rpoS* allele is important in determining its effect on the function of the sigma factor. Functional regions 2 and 4 are the most conserved because of their importance in -10 and -35 recognition and therefore mutations in those areas could have serious effects on the function of RpoS (Ohnuma *et al.*, 2000). The *rpoS*

allele from ECOR-21, ECOR-37 and ECOR-40 all have mutations in their region 2 and 4 which could explain why RpoS acivity in their recombinant strains were low. Contrarily, *rpoS* alleles from ECOR-28 as well as MG1655 had no mutations in their region 2 and 4, and thus had relatively high RpoS activity. This supports the explanation of a mutation in a functionally important or conserved area being deleterious.

One part of our results that couldn't be explained is the lower expression of HPII in the two constructed recombinant strains that had HPII (*rpoS* [ECOR-28] and *rpoS* [MG1655]) compared to their corresponding wildtype strains (Figure 17). It is possible that the recipient strain has mutations in *katE* or a gene required for HPII activity or expression, as previously reported by Robbe-Saule *et. al.* (2003). They found mutations in *katE* when they couldn't link differences in catalase activity to *rpoS* status. This will not be surprising in our case since the recipient strain was transduced with a P1 lysate which increases the chance of illegitimate insertions or recombinations at various points in the recipient's genome. Alternatively, it could be an effect from the presence of the cloning vector just after the 3' end (23 nucleotides) of *rpoS* in the constructs, interfering with the terminator and in turn affecting mRNA stability. Although, they encoded fully functional proteins, two strains with insertions (2100bp and 2900bp) downstream *rpoS* exhibited partially mutant RpoS phenotype (Atlung *et al.*, 2002). It is also not clear why there is an HPII band in ECOR-40 but not when it is cloned into the MG1655 background (Figure 17). It is possible that *katE* in ECOR-40 is under the control of a different regulator or sigma factor besides RpoS. That is, the *rpoS* allele in ECOR-40 is defective, but HPII is produced due to regulation of *katE* by a different regulator.

In conclusion, we have shown that natural *E. coli* strains have polymorphisms in their *rpoS* ORF. The polymorphism in natural *rpoS* alleles causes variation in the regulation of RpoS-dependent genes, specifically *osmY* and *katE*. This study has begun a discussion on the direct effect of *rpoS* polymorphism on gene expression within the large RpoS regulon. This is important information on the regulatory activities of RpoS and its relevance to cell survival during stress. Given that RpoS is a virulence factor in *Salmonella* sp. (Fang *et al.*, 1991), insights from this study can offer a general application for medical research in gram-negative pathogenic bacteria. Future studies can investigate the effect of *rpoS* polymorphism on gene expression by assaying other genes in the regulon using *lacZ* fusions and microarray. It would also be interesting to use a pathogenic recipient strain to explore how the *rpoS* alleles used in this project and others would influence its virulence and pathogenesis.

6.0 Appendices

6.1 Point form methods

6.11 PCR Amplification of DNA

1. Add the following in a 0.2ml thin-walled PCR tube

Reagent	Volume (µl)	Final Concentration
ddH$_2$O	38	-
10X buffer	5	1X
dNTPs	1	400µM
Primer I	2	1µM
Primer II	2	1µM
Taq Polymerase	1	1U
Template	1	-

(For whole colony PCR, boil a colony in 10µl of ddH$_2$O at 95^{0}C for 5 mins and add 3µl of suspension to PCR mix).

2. As a control, have a tube with all reagents except template.

3. Place tubes into a Thermocyler according to the following program:

Step	Temperature (^{o}C)	Time	Cycles
Initial denaturation	95	2:00	1
Cycle denaturation	95	0:30	30
Annealing	5^{0}C below T$_m$ of primers	0:30	30
Extension	72	1 min/ kb	30
Final extension	72	7:00	1

4. Store at -20°C until ready to use
5. Check 5µl of each sample on 0.8% agaraose gel.

6.12 Restriction digestion and Ligation for cloning

1. Introduce restriction sites into fragment of interest using PCR.
2. Purify PCR product using a purification kit
3. Check PCR product on a gel to estimate concentration
4. Digest ≈200ng of purified PCR product with 10u of appropriate restriction enzyme, 1X restriction buffer and ddH$_2$O up to 20ul final volume. Check manufacturer's protocol for optimal buffer and conditions for double digest
5. Incubate mixture at 37°C for 1 hr, for most enzymes. For complete digestion, incubate overnight. Check for manufacturer's protocol for optimal temperature and time.

6. Run digested fragment on gel along with undigested fragment as control
7. Purify digested fragment and run on gel again to check for purity and estimate concentration.
8. Digest ≈200ng of cloning vector with 10u of appropriate restriction enzyme(s), 1X restriction buffer and ddH$_2$O up to 20µl final volume. Check manufacturer's protocol for optimal buffer for double digest
9. Incubate mixture at 37°C for 1 hr, for most enzymes. For complete digestion, incubate overnight. Check for manufacturer's protocol for optimal buffer and conditions for double digest

10. Run linearized vector on 0.8% agarose gel along with undigested vector as control
11. Cut out linearized vector and isolate fragment from agarose using a DNA purification kit.
12. Run purified DNA on gel to check for purity and estimate concentration
13. If necessary, treat linearized vector with Shrimp Alkaline phosphatase to dephosphorylate it and incubate for 1hr at 37°C.
14. Inactivate SAP by incubation at 65^{0}C for 10mins.
15. Run DNA on gel to check for purity and estimate concentration
16. For ligation of DNA insert to cloning vector, use 3:1 insert to vector molar ratio
17. Add 1U of T4 DNA ligase, 1X DNA ligase buffer and ddH$_2$O to a final volume of 20µl.
18. Incubate mixture at 16^{0}C overnight
19. Check sample on 0.8% agarose gel along with unligated fragments
20. Inactivate T4 DNA ligase by incubation at 65^{0}C for 15mins
21. Transform host strain with 2µl of ligation product and screen for tranformants.

6.13 Transduction of recipient strain with P1 lysate (Miller, 1992)
1. Inoculate a 50ml flask containing 10ml of 1XLB supplemented with 5mM CaCl$_2$, with a single colony of the recipient strain.
2. Incubate the culture at 37°C overnight with shaking at 200rpm.
3. Wash cells with 1X LB containing 5mM CaCl$_2$ by transferring into a 15ml red cap polypropylene tube and centrifuging for 15min at 3000 rpm. Then discard supernatant and resuspend pellet in 300µl of 1X LB containing 5mM CaCl$_2$.

4. Mix 100µl of cell suspension with 100µl of P1 donor lysate in 15ml red cap polypropylene tube. Tap tube gently to mix and incubate at for 15mins at 37°C along with bacterial strain and P1 lysate separately as controls.

5. Wash cell suspension-P1 lysate mixture twice with 10ml of 1X LB containing 100mM sodium citrate by centrifuging for 15min at 3000 rpm.
6. Resuspend mixture in 10ml of 1X LB containing 100mM sodium citrate. Add 10ml of 1X LB containing 100mM sodium citrate to the two control tubes.

7. Incubate tubes at 37°C for 45mins on rotating wheel at 200 rpm.

8. Centrifuge tubes at 3000rpm for 15min, discard supernatant and resuspend pellet in 200μl of remaining media in tube.

9. Transfer 100μl of the cells onto LB agar plates containing appropriate antibiotics and incubate at 37°C overnight.

6.14 Preparation of electrocompentent cells for electroporation

1. Inoculate 10ml of 1X LB in Erlenmeyer flask containing the appropriate antibiotic with a single colony of the host strain to be transformed.

2. Incubate culture at 37°C overnight with shaking at 200 rpm.

3. Inoculate 50ml of 1X LB in a 250ml Erlenmeyer flask containing the appropriate antibiotic with 500μl of the overnight culture.

4. Incubate cells at 37°C (or 30°C if temperature sensitive) until OD_{600} 0.4.

5. Transfer culture to a 50ml centrifuge tube and leave on ice for 30min.

6. Centrifuge tube at 4200rpm for 15min at 4^0C. Discard supernatant.

6. Wash pellet 3X with 50ml ice-cold 10% glycerol by centrifuging at 4200rpm and 4^0C for 15 min each time. Use 5ml ice-cold 10% glycerol in the last wash.

7. Resuspend cells in 500μl ice-cold 10% glycerol.

8. Make 125μl aliquots in 1.5ml centrifuge tubes, flask freeze with liquid nitrogen and store at -80°C.

6.15 Electroporation of competent cells

1. Transfer 50μl of competent cells into a pre-chilled electroporation cuvette and place on ice.

2. Add 10-1000ng (in < 5μl) of DNA to the electroporation cuvette. Tap the cuvette to mix properly.

3. Set the Gene pulsar (electroporation equipment) to 2.25kV and 25μF.

4. Place cuvette in chamber and press the two red buttons to deliver electric pulse. Release buttons at the sound on a tone and appearance of "pls" on screen.

5. Immediately add 1ml of 1X LB supplemented with 20mM glucose (without antibiotic) to the cuvette.

6. Transfer contents of cuvette into a 1.5ml centrifuge tube and incubate at 37°C (or 30°C for temperature sensitive plasmid) for 1hr.

7. Centrifuge cells at 13 000rpm for 2min and resuspend in 200μl of 1X LB.

8. Spread plate 100μl of cell suspension on LB agar containing the appropriate antibiotics.

9. Incubate plates at 37°C (or 30°Cfor temperature sensitive strains) overnight.

6.16 β-galactosidase Assay (Miller, 1992)

1. Inoculate 10ml of 1X LB in Erlenmeyer flask containing the appropriate antibiotic with a single colony of the strain to be assayed.

2. Dilute culture 10 000X in 1X LB to avoid carry over of stationary phase cells and incubate at 37°C with shaking at 200 rpm.

3. Take out two aliquots of 1ml at both OD_{600} 0.3 (exponential phase) and OD_{600} 1.5

4. Add 4.4µl of 34mg/ml chloramphenicol to the samples taken out and mix to stop protein synthesis

5. Take out 200µl from each 1ml sample and add 800µl Z-buffer, 25µl 0.1% SDS and 50µl chloroform. Vortex to mix and incubate at room temperature for 5 min. (Z-buffer is 60mM $NaH_2PO_4 7H_2O$, 40mM $NaH_2PO_4H_2O$, 10mM KCl, 1mM $MgSO_4$ $7H_2O$, 50mM β-mercaptoethanol)

6. Start reaction by adding 200µl freshly prepared 4mg/ml ONPG (ortho-nitrophenyl-β-D-galactopyranoside), vortex samples, place in 28°C water bath and start the timer.

7. Add a control containing all reagents except bacterial culture.

8. Upon appearance of a yellow colour (OD_{600}0.2-0.9), stop reaction by adding 500µl of 1M Na_2CO_3.

9. Stop the timer, record the time

10. Centrifuge test tubes at 3000 rpm for 15 min.

11. Transfer 200µl of samples into a microtiter plate and measure absorbance at OD_{420}.

12. Calculate units of β-galactosidase as follows:

$$= [1000 \times OD_{420}] / [time (min) \times volume (ml) \times OD_{600}]$$

6.17 Preparation of cells for protein extracts

1. Inoculate 10ml of 1X LB in Erlenmeyer flask containing the appropriate antibiotic with a single colony of the strain to be assayed.

2. Subculture overnight culture 1:1000 and grow to 15-20 OD_{600} units.

3. Add 150µg/ml chloramphenicol to sample and leave on ice for 30 min.

4. Centrifuge sample at 3000rpm for 15min at 4°C. Discard supernant and resuspend pellet.

4. Wash cells 3X with 50mM phosphate (pH 7.0) by centrifuging at 3000rpm for 15min at 4°C. [50mM Phosphate buffer is 5.35g K_2HPO_4 and KH_2PO_4 in 1L ddH_2O]

5. Resuspend cells in 100µl of 50mM phosphate and transfer to a blue cap polystyrene tube.

6. Sonicate samples using the sonicator (See sonicator instruction manual)

7. Centrifuge sonicated samples in a 1.5ml tube at 13000rpm for 15min at 4°C.

8. Transfer supernatant to a new tube and store at 4°C

6.18 Bradford Assay for protein concentration determination

1. Aliquot 800µl ddH_2O into 18 borosilicate test tubes

2. Label each tube and add in triplicate, 0, 2, 4, 6, 8, and 10µl 1mg/ml BSA protein to the test tubes

3. Pipette 200µl of BioRad stock solution to each test tube, vortex and incubate at room temperature for 5min

4. Measure absorbance of each sample at OD_{600} after blanking spectrophotometer with water.

5. Plot standard curve of absorbance vs protein concentration

6. Repeat above procedure with 5µl of unknown protein extract in triplicate

7. Determine concentration of unknown using equation from the standard curve.

6.19 Polyacrylamide gel electrophoresis of protein extracts

1. Calculate and dilute all protein samples with appropriate volume of 50mM phosphate buffer and 1/3 volume 3X loading dye to obtain a final protein concentration of 1µg/ul

2. Prepare a 10% polyacrylamide gel by mixing 7ml ddH$_2$O, 3.75ml 1.5M Tris-HCl, 4.75ml 30% acrylamide/0.8% bisacrylamide, 150µl 10% Aps and 15µl TEMED.

3. Pour gel into Biorad MiniProtean III system plates. Allow gel to solidify.

4. Place plates with gel into gel box and fill with 1X running buffer. [Running buffer is 6.32g Tris, 3.94g glycine in 1L ddH$_2$O, pH 8.9]

6. Load polyacrylamide gel wells with 10µl (10µg) of protein samples. Run two gels in parallel and use one as a loading control.

7. Run gel at 100V for 2hr

8. Stain gel as required

6.20 Staining for catalase activity (Schellhorn and Stones, 1992)

1. Soak polyacylamide gel in 0.05mg/ml horseradish peroxidase solution for 15min. [0.05mg/ml horseradish peroxidase solution is 2.5mg horseradish peroxidase in 50ml 50mM phosphate buffer]

2. Pour out peroxidase solution and wash gel 2X with ddH$_2$O.

3. Soak gel in 5.0mM H$_2$O$_2$ for 10min. [5.0mM H$_2$O$_2$ is 45µl 30% H$_2$O$_2$ in 50ml 50mM phosphate buffer]

4. Pour out H$_2$O$_2$ solution and wash gel 2X with ddH$_2$O.

5. Soak gel in 0.5mg/ml diaminobenzidine solution until gel is brown and catalase is clear. [0.5mg/ml diaminobenzidine solution is 25mg 0.5mg/ml diaminobenzidine powder in 50ml 50mM phosphate buffer]

6. Take a picture of the gel.

6.21 Protein staining for polyacrylamide gels

1. Place gel in Coomassie blue R-250 staining solution for 1 hr. [Coomassie blue staining solution is 0.25g Coomassie blue R-250,10ml glacial acetic acid and 60ml water]

2. Pour out Coomassie blue solution and rinse gel with ddH$_2$O.

3. Place gel in destaining solution for 1hr. [Destaining solution is 30% ethanol and 10% glacial acetic acid]

4. Pour out decanting solution

5. Add fresh destaining solution and leave for 30mins

5. Take picture of the gel

6.22 Microassay for catalase activity (Li and Schellhorn, 2007)

1. Transfer concentrations of Bovine liver catalase from 0.01 to 2.00 units into a 96 well-flat bottom UV transparent microtiter plate.

2. Add 250µl 5mM hydrogen peroxide to each well using a repeating pipette.

3. Place the plate in a microtiter spectrophotometer using a program that takes absorbance reading at λ=240nm at 10 sec intervals for 5min.

4. Repeat procedure using triplicate of 5µl of protein extracts.

5. Determine catalase activity of unknown samples using rate of H_2O_2 decomposition of the standard, bovine liver catalase.

6.2 Raw results data

Determination of protein concentrations for catalase activity

Strain			OD595	[Protein] (ug/ul)	Final volume (ul)	Total protein (ug)	Sample volume (ul)	3x Native dye (ul)	Phophaste buffer (ul)
ECOR21	0.7867	0.8317	0.8092	3.541738	100	100	28.23473	33.33	38.435272
ECOR28	0.8378	0.8572	0.8475	4.194208	100	100	23.8424	33.33	42.827595
ECOR37	0.6956	0.7171	0.70635	1.789608	100	100	55.87815	33.33	10.791847
ECOR40	0.7862	0.7951	0.79065	3.225724	100	100	31.00079	33.33	35.669208
MG1655	0.852	0.8922	0.8721	4.613288	100	100	21.67651	33.33	44.993486
O21	0.8905	0.9236	0.90705	5.208688	100	100	19.19869	33.33	47.471308
O28	0.9306	0.893	0.9118	5.289608	100	100	18.90499	33.33	47.765008
O37	0.7744	0.7971	0.78575	3.142249	100	100	31.82434	33.33	34.845657
O40	0.836	0.8072	0.8216	3.752981	100	100	26.64548	33.33	40.024517
OMG	0.9379	0.9552	0.94655	5.881601	100	100	17.00217	33.33	49.667828
ΔrpoS	0.8461	0.9079	0.877	4.696763	100	100	21.29126	33.33	45.378741

Catalase sssay Data

Strain	Slope (units)				Units of Catalase				Replicate (units/ug)				Average	Standard Deviation	Standard Error
	1	2	3	4	1	2	3	4	1	2	3	4			
ECOR37	0.0003	0.0004	0.0004	0.0003	0.243	0.326	0.326	0.243	0.0485	0.0652	0.0652	0.0485	0.0568	0.0096	0.0048
rpoS[ECOR-37]	0.0003	0.0004	0.0004	0.0004	0.243	0.326	0.326	0.326	0.0485	0.0652	0.0652	0.0652	0.0610	0.0083	0.0042
ECOR28	0.0002	0.0003	0.0003	0.0003	0.159	0.243	0.243	0.243	0.0318	0.0485	0.0485	0.0485	0.0443	0.0083	0.0042
rpoS[ECOR-28]	0.0003	0.0003	0.0003	0.0003	0.243	0.243	0.243	0.243	0.0485	0.0485	0.0485	0.0485	0.0485	0.0000	0.0000
ECOR21	0.0005	0.0006	0.0006	0.0006	0.409	0.493	0.493	0.493	0.0818	0.0985	0.0985	0.0985	0.0943	0.0083	0.0042
rpoS[ECOR-21]	0.0003	0.0005	0.0004	0.0002	0.243	0.409	0.326	0.159	0.0485	0.0818	0.0652	0.0318	0.0568	0.0215	0.0108
ECOR40	0.0003	0.0004	0.0001	0.0009	0.243	0.326	0.076	0.743	0.0485	0.0652	0.0152	0.1485	0.0693	0.0567	0.0284
rpoS[ECOR-40]	0.0002	0.0004	0.0004	0.0003	0.159	0.326	0.326	0.243	0.0318	0.0652	0.0652	0.0485	0.0527	0.0160	0.0080
MG1655	0.0002	0.0002	0.0003	0.0002	0.159	0.159	0.243	0.159	0.0318	0.0318	0.0485	0.0318	0.0360	0.0083	0.0042
rpoS[MG1655]	0.0002	0.0003	0.0003	0.0002	0.159	0.243	0.243	0.159	0.0318	0.0485	0.0485	0.0318	0.0402	0.0096	0.0048
ΔrpoS	0.0002	0.0003	0.0004	0.0002	0.159	0.243	0.326	0.159	0.0318	0.0485	0.0652	0.0318	0.0443	0.0160	0.0080

β-galactosidase assay data

β-galactosidase activity assay of *osmY-lacZ* fusions in *rpoS* recombinant strains from stationary phase cultures

Strain	Volume	Time(min)	Rep 1			Rep 2			Rep 3			Activity			Ave Activity	Standard Dev	Standard Error
			OD600	OD420a	OD420b	OD600	OD420a	OD420b	OD600	OD420a	OD420b	Rep1	Rep 2	Rep 3			
rpoS [MG1655]	0.2	13	1.596	0.656	0.969	1.536	0.641	0.591	1.492	0.495	0.662	195.802	154.247	149.129	166.392	25.598	14.779
Δ*rpoS*	0.2	58	1.480	0.291	0.343	1.480	0.280	0.303	1.438	0.336	0.330	18.449	16.990	19.969	18.469	1.490	0.860
rpoS [ECOR-21]	0.2	47	1.631	0.682	0.494	1.642	0.603	0.890	1.542	0.307	0.344	38.362	48.351	22.463	36.392	13.056	7.538
rpoS [ECOR -28]	0.2	13	1.414	0.544	0.910	1.579	0.594	0.843	1.600	0.469	0.619	197.743	174.988	130.690	167.807	34.098	19.687
rpoS [ECOR- 37]	0.2	58	1.528	0.307	0.224	1.592	0.325	0.272	1.480	0.342	0.362	14.988	16.158	20.524	17.224	2.917	1.684
rpoS [ECOR -40]	0.2	58	1.423	0.336	0.280	1.536	0.342	0.284	1.560	0.248	0.313	18.674	17.568	15.490	17.244	1.616	0.933

β-galactosidase activity assay of *osmY-lacZ* fusions in *rpoS* recombinant strains from exponential phase cultures

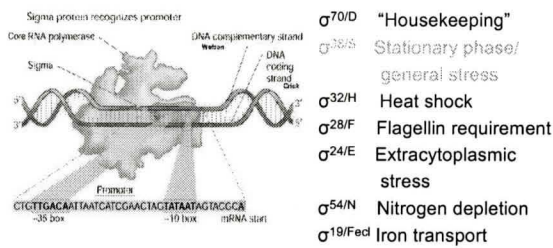| Strain	Volume	Time(min)	Rep 1			Rep 2			Rep 3			Activity			Ave Activity	Standard Dev	Standard Error
			OD600	OD420a	OD420b	OD600	OD420a	OD420b	OD600	OD420a	OD420b	Rep1	Rep 2	Rep 3			
rpoS [MG1655]	0.2	588	0.404	0.367	0.488	0.384	0.240	0.293	0.567	0.625	0.794	8.998	5.901	10.640	8.513	2.406	1.389
Δ*rpoS*	0.2	588	0.257	0.335	0.363	0.279	0.154	0.182	0.279	0.234	0.088	11.553	5.124	4.909	7.195	3.775	2.180
rpoS [ECOR-21]	0.2	588	0.245	0.196	0.193	0.359	0.921	1.062	0.452	0.426	0.430	6.742	23.475	8.046	12.754	9.307	5.373
rpoS [ECOR -28]	0.2	588	0.404	0.435	0.469	0.490	0.525	0.500	0.325	0.336	0.256	9.516	8.900	7.739	8.718	0.902	0.521
rpoS [ECOR- 37]	0.2	588	0.442	0.415	0.349	0.459	0.454	0.379	0.442	0.448	0.340	7.348	7.720	7.589	7.552	0.189	0.109
rpoS [ECOR -40]	0.2	588	0.450	0.569	0.552	0.378	0.225	0.182	0.353	0.219	0.098	10.594	4.580	3.818	6.331	3.712	2.143

6.3 PowerPoint slides for Oral Defense

The role of naturally occurring
alleles of *rpoS* in *Escherichia coli*

Daniel Gyewu
Thesis Defense
May 15,08

Outline

- Introduction
- Polymorphism in *rpoS*
- Project objective and rationale
- Methods and results
- Discussion
- Conclusions

Sigma factors in *E. coli*

σ^{70}/D	"Housekeeping"
$\sigma^{38/S}$	Stationary phase/ general stress
σ^{32}/H	Heat shock
σ^{28}/F	Flagellin requirement
σ^{24}/E	Extracytoplasmic stress
σ^{54}/N	Nitrogen depletion
$\sigma^{19}/Fecl$	Iron transport

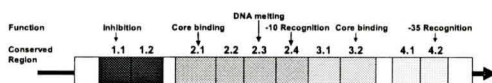http://fig.cox.miami.edu/~cmallery/150/gene/mol_gen.htm

RpoS: Master regulator of stress

- Oxidative stress
- Osmotic shock
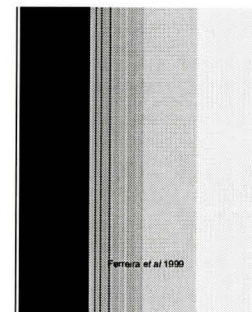- Near UV-radiation
- Heat shock
- Starvation
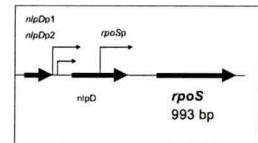
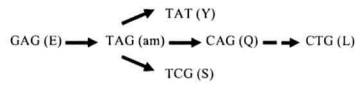Hengge-Aronis, 2002

Conserved regions of RpoS

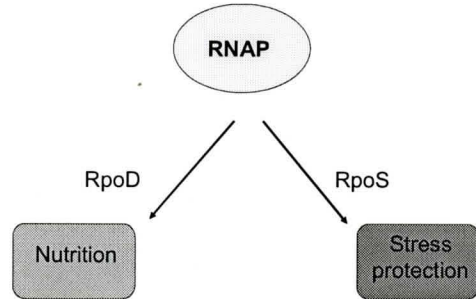Lonetto *et. al.*, 1992

Polymorphism in *rpoS*

Ferreira *et al* 1999

Possible evolutionary relationship of codon 33

TAT (Y)

GAG (E) → TAG (am) → CAG (Q) ⇢ CTG (L)

TCG (S)

Atlung et. al., 2002

What drives acquisition of mutations in *rpoS*?

RNAP

RpoD   RpoS

Nutrition   Stress protection

Chen et al 2004, Ferenci 2003

Variation in regulation of RpoS-dependent genes

- 13 out of 58 pathogenic strains were defective in acid resistance (Waterman and Small 1996)
- Six non-K12 strains had significant variation in acid phosphatase, catalase and glycogen accumulation levels (Atlung et. al., 2002)

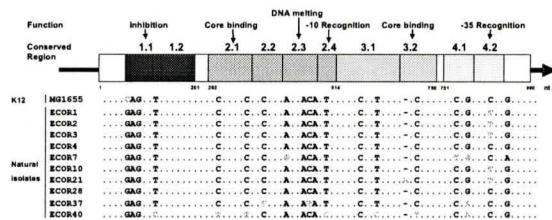| Strain	Genotype	Catalase activity
ECOR-1	ON:HN	+++
ECOR-2	ON:H32	+++
ECOR-3	O1:NM	+++
ECOR-4	ON:HN	+++
ECOR-7	O85:HN	++
ECOR-10	O6:H10	+++
ECOR-21	O121:HN	+
ECOR-28	O104:NM	++++
ECOR-37	ON:HN	+
ECOR-40	O7: NM	+
MG1655	F- λ- rph	++++
HS2718	F- λ- rph Δ*rpoS*	+

Main Objective

To investigate the effect of *rpoS* polymorphism on gene expression within the RpoS regulon

Polymorphism in *rpoS* of ECOR strains

Hypothesis

Variation in the regulation of RpoS-dependent genes in natural isolates of *E. coli* is a result of *rpoS* polymorphism

Lonetto et. al., 1992   Sequences are courtesy of T. Dong

Methodology

1. Construct recipient strain:
   MG1655 Δ*rpoS osmY-lacZ*
2. Clone selected natural *rpoS* alleles into a vector
3. Integrate *rpoS on vector* into chromosome of recipient strain
4. Comparative expression studies of the selected alleles using β-galactosidase assay and catalase activity

Construction of recipient strain

- Transduced MG1655 Δ*rpoS* Δ*lacZYA* with P1 lysate containing *osmY-lacZ*

MG1655 Δ*rpoS* Δ*lacZYA*   MG1655 Δ*rpoS osmY-lacZ*

Cloning *rpoS* alleles into a vector

Integration of *rpoS* into chromosome of recipient strain

Characterization of *rpoS* recombinants

Phenotype of recombinant natural *rpoS* strains on X-gal

β-galactosidase activity of osmY-lacZ in recombinant rpoS strains
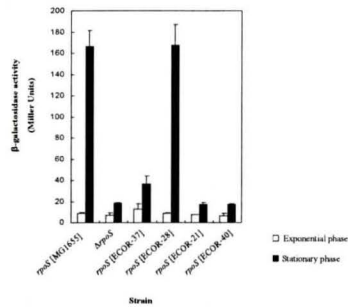
Catalase activity of recombinant natural rpoS strains
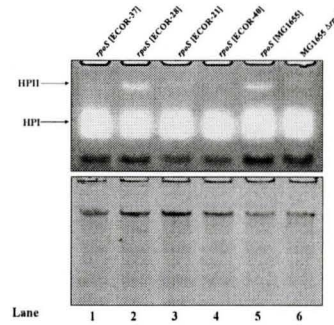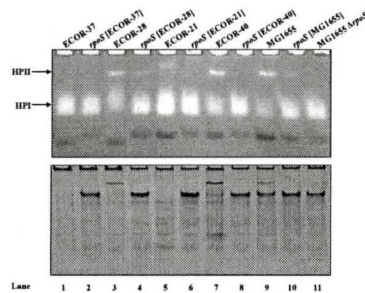

Comparison of catalase activity between recombinant and natural strains

Effects of missense mutations

- Change in amino acid
- Modification of RpoS primary structure
- Potential change in overall structure
- Effect on stability and function of RpoS

Effects of silent mutations

- The rate of mRNA translation by ribosomes can be controlled by codon usage
- The rate of translation can in turn affect protein folding
- Effect on RpoS stability and function

How does rpoS polymorphism affect RpoS activity?

- Alteration in RpoS function
  - affect specificity for certain promoters
  - reduction or enhancement in its regulatory activities
- Alteration of RpoS stability
  - Effect on regulation at post-translational level

Location of mutations

Conclusions

- Natural *E. coli* strains have polymorphism in their *rpoS* ORF different from K12 and each other
- Polymorphism in natural *rpoS* alleles causes variation in the regulation of RpoS-dependent genes, in our case *osmY* and *katE*

Significance of project

- Insight on the direct effect of *rpoS* polymorphism on gene expression within the regulon
- Provide additional information on the regulatory activities of RpoS and its relevance to cell survival during stress

Future directions

- Examine the effect of *rpoS* polymorphism on gene expression by assaying other genes in the regulon using *lacZ* fusions and microarray
- Determine the strength of *rpoS* promoters in the natural strains
- Investigate the role of *rpoS* polymorphism in pathogenic *E. coli* strains by using a pathogenic recipient strain in a similar study

Acknowledgments

- Dr H.E.Schellhorn
- Dr G. Sorger
- Dr J.P. Xu
- Members of the Schellhorn Lab
- Everyone present

Reference List

Atlung,T., Nielsen,H.V., and Hansen,F.G. (2002). Characterisation of the allelic variation in the *rpoS* gene in thirteen K12 and six other non-pathogenic *Escherichia coli* strains. Mol. Genet. Genomics *266*, 873-881.

Barth,M., Marschall,C., Muffler,A., Fischer,D., and Hengge-Aronis,R. (1995). Role for the histone-like protein H-NS in growth phase-dependent and osmotic regulation of sigma S and many sigma S-dependent genes in *Escherichia coli*. J. Bacteriol. *177*, 3455-3464.

Baudin,A., Ozier-Kalogeropoulos,O., Denouel,A., Lacroute,F., and Cullin,C. (1993). A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. Nucleic Acids Res. *21*, 3329-3330.

Becker,G. and Hengge-Aronis,R. (2001). What makes an *Escherichia coli* promoter sigma(S) dependent? Role of the -13/-14 nucleotide promoter positions and region 2.5 of sigma(S). Mol. Microbiol. *39*, 1153-1165.

Belin,D. (2003). Why are suppressors of amber mutations so frequent among *Escherichia coli* K12 strains?. A plausible explanation for a long-lasting puzzle. Genetics *165*, 455-456.

Bhagwat,A.A., Tan,J., Sharma,M., Kothary,M., Low,S., Tall,B.D., and Bhagwat,M. (2006). Functional heterogeneity of RpoS in stress tolerance of enterohemorrhagic *Escherichia coli* strains. Appl. Environ. Microbiol. *72*, 4978-4986.

Blattner,F.R., Plunkett,G., III, Bloch,C.A., Perna,N.T., Burland,V., Riley,M., Collado-Vides,J., Glasner,J.D., Rode,C.K., Mayhew,G.F., Gregor,J., Davis,N.W., Kirkpatrick,H.A., Goeden,M.A., Rose,D.J., Mau,B., and Shao,Y. (1997). The complete genome sequence of *Escherichia coli* K-12. Science *277*, 1453-1474.

Boe,L., Danielsen,M., Knudsen,S., Petersen,J.B., Maymann,J., and Jensen,P.R. (2000). The frequency of mutators in populations of *Escherichia coli*. Mutat. Res. *448*, 47-55.

Bremer,E., Silhavy,T.J., Weisemann,J.M., and Weinstock,G.M. (1984). Lambda p*lac*Mu: a transposable derivative of bacteriophage lambda for creating *lacZ* protein fusions in a single step. J. Bacteriol. *158*, 1084-1093.

Brown,L. and Elliott,T. (1996). Efficient translation of the RpoS sigma factor in *Salmonella typhimurium* requires host factor I, an RNA-binding protein encoded by the hfq gene. J. Bacteriol. *178*, 3763-3770.

Chen,G., Patten,C.L., and Schellhorn,H.E. (2004). Positive selection for loss of RpoS function in *Escherichia coli*. Mutat. Res. *554*, 193-203.

Chiang, S. and Schellhorn, H.E. Unpublished data.

Christman,M.F., Morgan,R.W., Jacobson,F.S., and Ames,B.N. (1985). Positive control of a regulon for defenses against oxidative stress and some heat-shock proteins in *Salmonella typhimurium*. Cell *41*, 753-762.

Chung,C.T., Niemela,S.L., and Miller,R.H. (1989). One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. Proc. Natl. Acad. Sci. U. S. A *86*, 2172-2175.

Coldewey,S.M., Hartmann,M., Schmidt,D.S., Engelking,U., Ukena,S.N., and Gunzer,F. (2007). Impact of the rpoS genotype for acid resistance patterns of pathogenic and probiotic *Escherichia coli*. BMC. Microbiol. *7*, 21.

Colland,F., Fujita,N., Ishihama,A., and Kolb,A. (2002). The interaction between sigmaS, the stationary phase sigma factor, and the core enzyme of *Escherichia coli* RNA polymerase. Genes Cells *7*, 233-247.

Corona-Izquierdo,F.P. and Membrillo-Hernandez,J. (2002). A mutation in *rpoS* enhances biofilm formation in *Escherichia coli* during exponential phase of growth. FEMS Microbiol. Lett. *211*, 105-110.

Cortazzo,P., Cervenansky,C., Marin,M., Reiss,C., Ehrlich,R., and Deana,A. (2002). Silent mutations affect in vivo protein folding in *Escherichia coli*. Biochem. Biophys. Res. Commun. *293*, 537-541.

Datsenko,K.A. and Wanner,B.L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc. Natl. Acad. Sci. (USA) *97*, 6640-6645.

Deana,A., Ehrlich,R., and Reiss,C. (1996). Synonymous codon selection controls in vivo turnover and amount of mRNA in *Escherichia coli bla* and *ompA* genes. J. Bacteriol. *178*, 2718-2720.

Deana,A., Ehrlich,R., and Reiss,C. (1998). Silent mutations in the *Escherichia coli ompA* leader peptide region strongly affect transcription and translation in vivo. Nucleic Acids Res. *26*, 4778-4782.

Dodd,C.E. and Aldsworth,T.G. (2002). The importance of RpoS in the survival of bacteria through food processing. Int. J. Food Microbiol. *74*, 189-194.

Dong, T and Schellhorn, H.E. Unpublished data.

Drake,J.W., Charlesworth,B., Charlesworth,D., and Crow,J.F. (1998). Rates of spontaneous mutation. Genetics *148*, 1667-1686.

Ellis,H.M., Yu,D., DiTizio,T., and Court DL (2001). High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. Proc. Natl. Acad. Sci. U. S. A *98*, 6742-6746.

Espinosa-Urgel,M., Chamizo,C., and Tormo,A. (1996). A consensus structure for sigma S-dependent promoters. Mol. Microbiol. *21*, 657-659.

Espinosa-Urgel,M. and Tormo,A. (1993). Sigma s-dependent promoters in Escherichia coli are located in DNA regions with intrinsic curvature. Nucleic Acids Res. *21*, 3667-3670.

Fang,F.C., Krause,M., Roudier,C., Fierer,J., and Guiney,D.G. (1991). Growth regulation of a *Salmonella* plasmid gene essential for virulence. J. Bacteriol. *173*, 6783-6789.

Ferenci,T. (2003). What is driving the acquisition of *mutS* and *rpoS* polymorphisms in *Escherichia coli*? Trends Microbiol. *11*, 457-461.

Ferreira,A., Rendano,L., Wiedmann,M., and Boor,K.J. (1999). Characterization of *rpoS* alleles in *Escherichia coli* O157:H7 and in other *E. coli* serotypes. J. Appl. Microbiol. *86*, 295-301.

Gaal,T., Ross,W., Estrem,S.T., Nguyen,L.H., Burgess,R.R., and Gourse,R.L. (2001). Promoter recognition and discrimination by EsigmaS RNA polymerase. Mol. Microbiol. *42*, 939-954.

Goodman,H.M., Abelson,J., Landy,A., Brenner,S., and Smith,J.D. (1968). Amber suppression: a nucleotide change in the anticodon of a tyrosine transfer RNA. Nature *217*, 1019-1024.

Goulian,M. and van der,W.M. (2006). A simple system for converting *lacZ* to gfp reporter fusions in diverse bacteria. Gene *372*, 219-226.

Gruber,T.M. and Gross,C.A. (2003). Multiple sigma subunits and the partitioning of bacterial transcription space. Annu. Rev. Microbiol. *57*, 441-466.

Hamilton,C.M., Aldea,M., Washburn,B.K., Babitzke,P., and Kushner,S.R. (1989). New method for generating deletions and gene replacements in *Escherichia coli*. J. Bacteriol. *171*, 4617-4622.

Hegreness,M., Shoresh,N., Hartl,D., and Kishony,R. (2006). An equivalence principle for the incorporation of favorable mutations in asexual populations. Science *311*, 1615-1617.

Hengge-Aronis,R. (1996). Back to log phase: sigma S as a global regulator in the osmotic control of gene expression in *Escherichia coli*. Mol. Microbiol. *21*, 887-893.

Hengge-Aronis,R. (2000). The general stress response in *Escherichia coli*. In Bacterial stress response, G.Storz and R.Hengge-Aronis, eds. (Washinton, D.C.: ASM press), pp. 161-178.

Hengge-Aronis,R. (2002). Signal transduction and regulatory mechanisms involved in control of the sigma(S) (RpoS) subunit of RNA polymerase. Microbiol. Mol. Biol. Rev. *66*, 373-95, table.

Hengge-Aronis,R., Lange,R., Henneberg,N., and Fischer,D. (1993). Osmotic regulation of rpoS-dependent genes in *Escherichia coli*. J. Bacteriol. *175*, 259-265.

Herbelin,C.J., Chirillo,S.C., Melnick,K.A., and Whittam,T.S. (2000). Gene conservation and loss in the *mutS-rpoS* genomic region of pathogenic *Escherichia coli*. J. Bacteriol. *182*, 5381-5390.

Hill,S.A., Stahl,M.M., and Stahl,F.W. (1997). Single-strand DNA intermediates in phage lambda's Red recombination pathway. Proc. Natl. Acad. Sci. U. S. A *94*, 2951-2956.

Hirsch,M. and Elliott,T. (2005). Stationary-phase regulation of RpoS translation in *Escherichia coli*. J. Bacteriol. *187*, 7204-7213.

Ikemura,T. (1985). Codon usage and tRNA content in unicellular and multicellular organisms. Mol. Biol. Evol. *2*, 13-34.

Ishihama,A. (1981). Subunit of assembly of *Escherichia coli* RNA polymerase. Adv. Biophys. *14*, 1-35.

Ishihama,A. (1988). Promoter selectivity of prokaryotic RNA polymerases. Trends Genet. *4*, 282-286.

Ishihama,A. (2000). Functional modulation of *Escherichia coli* RNA polymerase. Annu. Rev. Microbiol. *54*, 499-518.

Ivanova,A., Renshaw,M., Guntaka,R.V., and Eisenstark,A. (1992). DNA base sequence variability in katF (putative sigma factor) gene of *Escherichia coli*. Nucleic Acids Res. *20*, 5479-5480.

Jasin,M. and Schimmel,P. (1984). Deletion of an essential gene in *Escherichia coli* by site-specific recombination with linear DNA fragments. J. Bacteriol. *159*, 783-786.

Jishage,M. and Ishihama,A. (1995). Regulation of RNA polymerase sigma subunit synthesis in *Escherichia coli*: intracellular levels of sigma 70 and sigma 38. J. Bacteriol. *177*, 6832-6835.

Jishage,M. and Ishihama,A. (1997). Variation in RNA polymerase sigma subunit composition within different stocks of *Escherichia coli* W3110. J. Bacteriol. *179*, 959-963.

Jishage,M. and Ishihama,A. (1999). Transcriptional organization and *in vivo* role of the *Escherichia coli rsd* gene, encoding the regulator of RNA polymerase sigma D. J. Bacteriol. *181*, 3768-3776.

Jishage,M., Iwata,A., Ueda,S., and Ishihama,A. (1996). Regulation of RNA polymerase sigma subunit synthesis in *Escherichia coli*: intracellular levels of four species of sigma subunit under various growth conditions. J. Bacteriol. *178*, 5447-5451.

Jones,P. and Suggett,A. (1968). The catalse-hydrogen peroxide system. Kinetics of catalatic action at high substrate concentrations. Biochem. J. *110*, 617-620.

Keener,J. and Nomura,M. (1993). Dominant lethal phenotype of a mutation in the -35 recognition region of *Escherichia coli* sigma 70. Proc. Natl. Acad. Sci. U. S. A *90*, 1751-1755.

King,T., Ishihama,A., Kori,A., and Ferenci,T. (2004). A regulatory trade-off as a source of strain variation in the species *Escherichia coli*. J. Bacteriol. *186*, 5614-5620.

Krogfelt,K.A., Hjulgaard,M., Sorensen,K., Cohen,P.S., and Givskov,M. (2000). *rpoS* gene function is a disadvantage for *Escherichia coli* BJ4 during competitive colonization of the mouse large intestine. Infect. Immun. *68*, 2518-2524.

Lange,R., Fischer,D., and Hengge-Aronis,R. (1995). Identification of transcriptional start sites and the role of ppGpp in the expression of *rpoS*, the structural gene for the sigma S subunit of RNA polymerase in *Escherichia coli*. J. Bacteriol. *177*, 4676-4680.

Lange,R. and Hengge-Aronis,R. (1991). Growth phase-regulated expression of *bolA* and morphology of stationary-phase *Escherichia coli* cells are controlled by the novel sigma factor sigma S. J. Bacteriol. *173*, 4474-4481.

Lange,R. and Hengge-Aronis,R. (1994a). The cellular concentration of the sigma S subunit of RNA polymerase in *Escherichia coli* is controlled at the levels of transcription, translation, and protein stability. Genes Dev. *8*, 1600-1612.

Lange,R. and Hengge-Aronis,R. (1994b). The *nlpD* gene is located in an operon with *rpoS* on the *Escherichia coli* chromosome and encodes a novel lipoprotein with a potential function in cell wall formation. Mol. Microbiol. *13*, 733-743.

LeClerc,J.E., Li,B., Payne,W.L., and Cebula,T.A. (1996). High mutation frequencies among *Escherichia coli* and *Salmonella* pathogens. Science *274*, 1208-1211.

LeClerc,J.E., Payne,W.L., Kupchella,E., and Cebula,T.A. (1998). Detection of mutator subpopulations in *Salmonella typhimurium* LT2 by reversion of his alleles. Mutat. Res. *400*, 89-97.

Lee,S.J. and Gralla,J.D. (2002). Promoter use by sigma 38 (rpoS) RNA polymerase. Amino acid clusters for DNA binding and isomerization. J. Biol. Chem. *277*, 47420-47427.

Li,M.Z. and Elledge,S.J. (2005). MAGIC, an *in vivo* genetic method for the rapid construction of recombinant DNA molecules. Nat. Genet. *37*, 311-319.

Li,Y. and Schellhorn,H.E. (2007). Rapid kinetic microassay for catalase activity. J. Biomol. Tech. *18*, 185-187.

Loewen,P.C. (1984). Isolation of catalase-deficient *Escherichia coli* mutants and genetic mapping of *katE*, a locus that affects catalase activity. J. Bacteriol. *157*, 622-626.

Loewen,P.C. and Hengge-Aronis,R. (1994). The role of the sigma factor sigma S (KatF) in bacterial global regulation. Annu. Rev. Microbiol. *48*, 53-80.

Loewen,P.C., Hu,B., Strutinsky,J., and Sparling,R. (1998). Regulation in the *rpoS* regulon of *Escherichia coli*. Can. J. Microbiol. *44*, 707-717.

Loewen,P.C., Switala,J., and Triggs-Raine,B.L. (1985). Catalases HPI and HPII in *Escherichia coli* are induced independently. Arch. Biochem. Biophys. *243*, 144-149.

Loewen,P.C. and Triggs,B.L. (1984). Genetic mapping of katF, a locus that with *katE* affects the synthesis of a second catalase species in *Escherichia coli*. J. Bacteriol. *160*, 668-675.

London,J.W., Shaw,L.M., Fetterolf,D., and Garfinkel,D. (1975). A systematic approach to enzyme assay optimization illustrated by aminotransferase assays. Clin. Chem. *21*, 1939-1952.

Lonetto,M., Gribskov,M., and Gross,C.A. (1992). The sigma 70 family: sequence conservation and evolutionary relationships. J. Bacteriol. *174*, 3843-3849.

Mao,E.F., Lane,L., Lee,J., and Miller,J.H. (1997). Proliferation of mutators in A cell population. J. Bacteriol. *179*, 417-422.

Matic,I., Radman,M., Taddei,F., Picard,B., Doit,C., Bingen,E., Denamur,E., and Elion,J. (1997). Highly variable mutation rates in commensal and pathogenic *Escherichia coli*. Science *277*, 1833-1834.

McCann,M.P., Fraley,C.D., and Matin,A. (1993). The putative sigma factor KatF is regulated posttranscriptionally during carbon starvation. J. Bacteriol. *175*, 2143-2149.

Metcalf,W.W., Jiang,W., and Wanner,B.L. (1994). Use of the rep technique for allele replacement to construct new *Escherichia coli* hosts for maintenance of R6K gamma origin plasmids at different copy numbers. Gene *138*, 1-7.

Miller,J.H. (1992). A short course in bacterial genetics: A laboratory manual and handbookfor *Escherichia coli* and related bacteria., Cold Spring Harbor Press, ed. (Cold Spring Harbor, N.Y..

Muffler,A., Barth,M., Marschall,C., and Hengge-Aronis,R. (1997a). Heat shock regulation of sigmaS turnover: a role for DnaK and relationship between stress responses mediated by sigmaS and sigma32 in *Escherichia coli*. J. Bacteriol. *179*, 445-452.

Muffler,A., Fischer,D., Altuvia,S., Storz,G., and Hengge-Aronis,R. (1996). The response regulator RssB controls stability of the sigma(S) subunit of RNA polymerase in *Escherichia coli*. EMBO J. *15*, 1333-1339.

Muffler,A., Traulsen,D.D., Fischer,D., Lange,R., and Hengge-Aronis,R. (1997b). The RNA-binding protein HF-I plays a global regulatory role which is largely, but not exclusively, due to its role in expression of the sigmaS subunit of RNA polymerase in *Escherichia coli*. J. Bacteriol. *179*, 297-300.

Mulvey,M.R., Switala,J., Borys,A., and Loewen,P.C. (1990). Regulation of transcription of *katE* and *katF* in *Escherichia coli*. J. Bacteriol. *172*, 6713-6720.

Murphy,K.C. (1998). Use of bacteriophage lambda recombination functions to promote gene replacement in *Escherichia coli*. J. Bacteriol. *180*, 2063-2071.

Notley-McRobb,L., King,T., and Ferenci,T. (2002). *rpoS* mutations and loss of general stress resistance in *Escherichia coli* populations as a consequence of conflict between competing stress responses. J. Bacteriol. *184*, 806-811.

Ochman,H. and Selander,R.K. (1984). Standard reference strains of Escherichia coli from natural populations. J. Bacteriol. *157*, 690-693.

Ohnuma,M., Fujita,N., Ishihama,A., Tanaka,K., and Takahashi,H. (2000). A carboxy-terminal 16-amino-acid region of sigma(38) of *Escherichia coli* is important for transcription under high-salt conditions and sigma activities *in vivo*. J. Bacteriol. *182*, 4628-4631.

Oliner,J.D., Kinzler,K.W., and Vogelstein,B. (1993). *In vivo* cloning of PCR products in *E. coli*. Nucleic Acids Res. *21*, 5192-5197.

Paget,M.S. and Helmann,J.D. (2003). The sigma70 family of sigma factors. Genome Biol. *4*, 203.

Patten,C.L., Kirchhof,M.G., Schertzberg,M.R., Morton,R.A., and Schellhorn,H.E. (2004). Microarray analysis of RpoS-mediated gene expression in *Escherichia coli* K-12. Mol. Genet. Genomics *272*, 580-591.

Perna,N.T., Plunkett,G., III, Burland,V., Mau,B., Glasner,J.D., Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A., Posfai,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L., Grotbeck,E.J., Davis,N.W., Lim,A., Dimalanta,E.T., Potamousis,K.D., Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C., Welch,R.A., and Blattner,F.R. (2001). Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. Nature *409*, 529-533.

Poteete,A.R. (2001). What makes the bacteriophage lambda Red system useful for genetic engineering: molecular mechanism and biological function. FEMS Microbiol. Lett. *201*, 9-14.

Rajkumari,K. and Gowrishankar,J. (2002). An N-terminally truncated RpoS (sigma(S)) protein in *Escherichia coli* is active in vivo and exhibits normal environmental regulation even in the absence of *rpoS* transcriptional and translational control signals. J. Bacteriol. *184*, 3167-3175.

Robey,M., Benito,A., Hutson,R.H., Pascual,C., Park,S.F., and Mackey,B.M. (2001). Variation in resistance to high hydrostatic pressure and *rpoS* heterogeneity in natural isolates of *Escherichia coli* O157:H7. Appl. Environ. Microbiol. *67*, 4901-4907.

Rocha,E.P. (2004). Codon usage bias from tRNA's point of view: redundancy, specialization, and efficient decoding for translation optimization. Genome Res. *14*, 2279-2286.

Sak,B.D., Eisenstark,A., and Touati,D. (1989). Exonuclease III and the catalase hydroperoxidase II in Escherichia coli are both regulated by the *katF* gene product. Proc. Natl. Acad. Sci. U. S. A *86*, 3271-3275.

Sammartano,L.J., Tuveson,R.W., and Davenport,R. (1986). Control of sensitivity to inactivation by H2O2 and broad-spectrum near-UV radiation by the *Escherichia coli katF* locus. J. Bacteriol. *168*, 13-21.

Schellhorn,H.E., Audia,J.P., Wei,L.I.C., and Chang,L. (1998). Identification of conserved, RpoS-dependent stationary-phase genes of *Escherichia coli*. Journal of Bacteriology *180*, 6283-6291.

Schellhorn,H.E. and Hassan,H.M. (1988). Transcriptional regulation of *katE* in *Escherichia coli* K-12. J. Bacteriol. *170*, 4286-4292.

Schellhorn,H.E. and Stones,V.L. (1992). Regulation of *katF* and *katE* in *Escherichia coli* K-12 by weak acids. J. Bacteriol. *174*, 4769-4776.

Schweder,T., Lee,K.H., Lomovskaya,O., and Matin,A. (1996). Regulation of *Escherichia coli* starvation sigma factor (sigma s) by ClpXP protease. J. Bacteriol. *178*, 470-476.

Smith,H.O. and Wilcox,K.W. (1970). A restriction enzyme from *Hemophilus influenzae*. I. Purification and general properties. J. Mol. Biol. *51*, 379-391.

Solomovici,J., Lesnik,T., and Reiss,C. (1997). Does *Escherichia coli* optimize the economics of the translation process? J. Theor. Biol. *185*, 511-521.

Sorensen,M.A. and Pedersen,S. (1991). Absolute in vivo translation rates of individual codons in *Escherichia coli*. The two glutamic acid codons GAA and GAG are translated with a threefold difference in rate. J. Mol. Biol. *222*, 265-280.

Subbarayan,P.R. and Sarkar,M. (2004a). A comparative study of variation in codon 33 of the *rpoS* gene in *Escherichia coli* K12 stocks: implications for the synthesis of sigma(s). Mol. Genet. Genomics *270*, 533-538.

Subbarayan,P.R. and Sarkar,M. (2004b). *Escherichia coli rpoS* gene has an internal secondary translation initiation region. Biochem. Biophys. Res. Commun. *313*, 294-299.

Suzuki,Y., Kagawa,N., Fujino,T., Sumiya,T., Andoh,T., Ishikawa,K., Kimura,R., Kemmochi,K., Ohta,T., and Tanaka,S. (2005). A novel high-throughput (HTP) cloning strategy for site-directed designed chimeragenesis and mutation using the Gateway cloning system. Nucleic Acids Res. *33*, e109.

Switala,J. and Loewen,P.C. (2002). Diversity of properties among catalases. Arch. Biochem. Biophys. *401*, 145-154.

Tanaka,K., Takayanagi,Y., Fujita,N., Ishihama,A., and Takahashi,H. (1993). Heterogeneity of the principal sigma factor in *Escherichia coli*: the *rpoS* gene product, sigma 38, is a second principal sigma factor of RNA polymerase in stationary-phase *Escherichia coli*. Proc. Natl. Acad. Sci. U. S A *90*, 8303.

Touati,E., Dassa,E., and Boquet,P.L. (1986). Pleiotropic mutations in *appR* reduce pH 2.5 acid phosphatase expression and restore succinate utilisation in CRP-deficient strains of *Escherichia coli*. Mol. Gen. Genet. *202*, 257-264.

Tuveson,R.W. and Jonas,R.B. (1979). Genetic control of near-UV (300-400 NM) sensitivity independent of the *recA* gene in strains of *Escherichia coli* K12. Photochem. Photobiol. *30*, 667-676.

Typas,A. and Hengge,R. (2006). Role of the spacer between the -35 and -10 regions in sigmas promoter selectivity in *Escherichia coli*. Mol. Microbiol. *59*, 1037-1051.

Visick,J.E. and Clarke,S. (1997). RpoS- and OxyR-independent induction of HPI catalase at stationary phase in *Escherichia coli* and identification of *rpoS* mutations in common laboratory strains. J. Bacteriol. *179*, 4158-4163.

Volodin,A.A., Voloshin,O.N., and Camerini-Otero,R.D. (2005). Homologous recombination and RecA protein: towards a new generation of tools for genome manipulations. Trends Biotechnol. *23*, 97-102.

Wang,Y. and Kim,K.S. (2000). Effect of *rpoS* mutations on stress-resistance and invasion of brain microvascular endothelial cells in *Escherichia coli* K1. FEMS Microbiol. Lett. *182*, 241-247.

Waterman,S.R. and Small,P.L. (1996). Characterization of the acid resistance phenotype and *rpoS* alleles of shiga-like toxin-producing *Escherichia coli*. Infect. Immun. *64*, 2808-2811.

Weber,H., Polen,T., Heuveling,J., Wendisch,V.F., and Hengge,R. (2005). Genome-wide analysis of the general stress response network in *Escherichia coli*: sigmaS-dependent genes, promoters, and sigma factor selectivity. J. Bacteriol. *187*, 1591-1603.

Yu,D., Ellis,H.M., Lee,E.C., Jenkins,N.A., Copeland,N.G., and Court DL (2000). An efficient recombination system for chromosome engineering in *Escherichia coli*. Proc. Natl. Acad. Sci. U. S. A *97*, 5978-5983.

Zambrano,M.M. and Kolter,R. (1996). GASPing for life in stationary phase. Cell *86*, 181-184.

Zgurskaya,H.I., Keyhan,M., and Matin,A. (1997). The sigma S level in starving *Escherichia coli* cells increases solely as a result of its increased stability, despite decreased synthesis. Mol. Microbiol. *24*, 643-651.

Zhang,P., Li,M.Z., and Elledge,S.J. (2002). Towards genetic genome projects: genomic library screening and gene-targeting vector construction in a single step. Nat. Genet. *30*, 31-39.

Zhang,Y., Buchholz,F., Muyrers,J.P., and Stewart,A.F. (1998). A new logic for DNA engineering using recombination in *Escherichia coli*. Nat. Genet. *20*, 123-128.

Zhang,Y., Muyrers,J.P., Testa,G., and Stewart,A.F. (2000). DNA cloning by homologous recombination in *Escherichia coli*. Nat. Biotechnol. *18*, 1314-1317.